



<https://www.cbioportal.org/>

Ino de Bruijn

Bioinformatics Engineer

Knowledge Systems Group (Head: Nikolaus Schultz)

Memorial Sloan-Kettering Cancer Center



Memorial Sloan-Kettering
Cancer Center



Dana-Farber
Cancer Institute



The Children's Hospital of Philadelphia

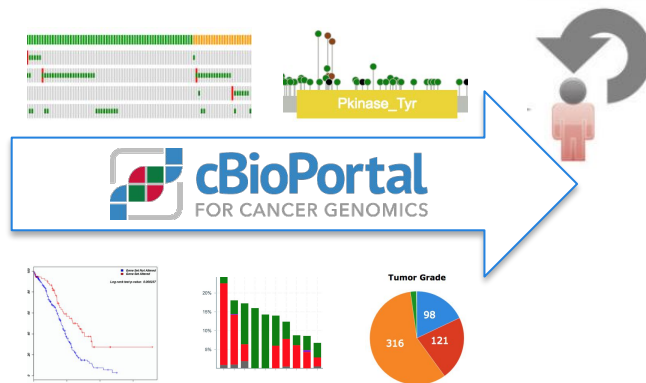
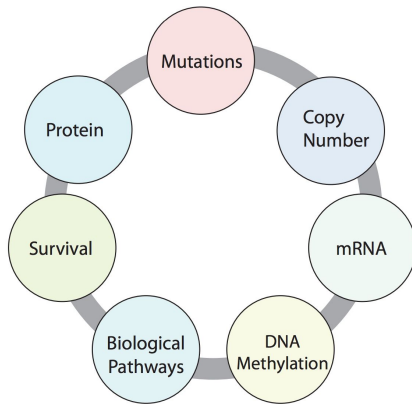


Toronto General
Toronto Western
Princess Margaret
Toronto Rehab



Goal of cBioPortal

Make complex cancer genomic data **accessible** and **interpretable** for cancer biologists and clinicians



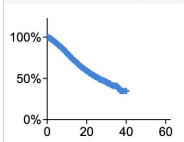
Biological
discovery

Clinical
interpretation

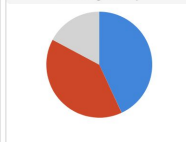
Cohort exploration

Cancer Type	#	Freq
Non-Small Cell Lung Cancer	1,668	15.2%
Breast Cancer	1,324	12.1%
Colorectal Cancer	1,007	9.2%
Prostate Cancer	717	6.6%
Glioma	553	5.1%
Pancreatic Cancer	502	4.6%
Soft Tissue Sarcoma	443	4.0%
Bladder Cancer	423	3.9%
Melanoma	365	3.3%
Renal Cell Carcinoma	361	3.3%
Hepatobiliary Cancer	355	3.2%

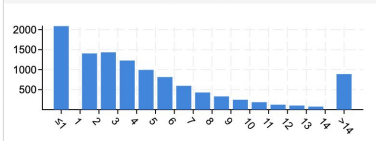
KM Plot: Overall Survival (months)



Smoking History

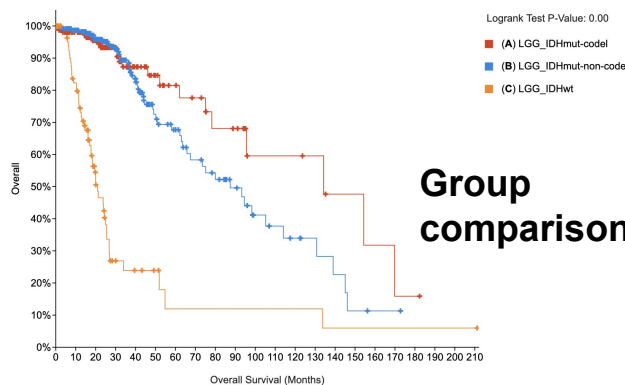


Mutation Count

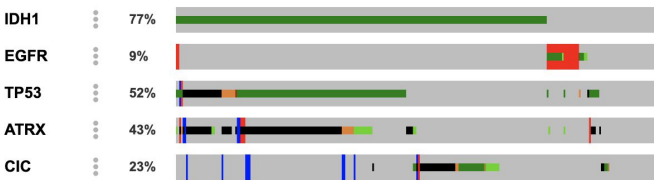


Mutated Genes (10945 profiled samples)

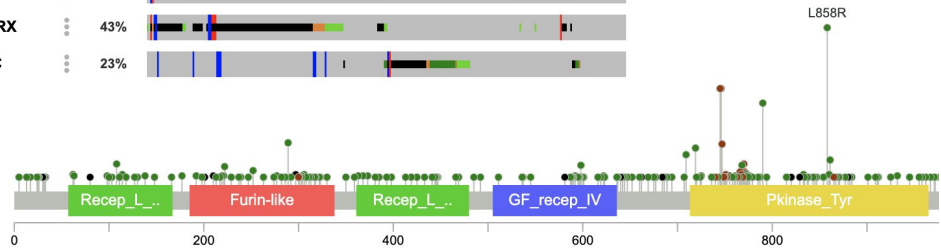
Gene	# Mut	#	Freq
TP53	4,954	4,538	41.5%
KRAS	1,670	1,643	15.0%
TERT	1,549	1,460	13.3%
PIK3CA	1,517	1,355	12.4%
APC	1,674	1,121	10.2%
ARID1A	1,072	875	8.0%
KMT2D	1,210	851	7.8%
PTEN	764	665	6.1%
KMT2C	818	642	5.9%
EGFR	799	636	5.8%
FAT1	773	595	5.4%



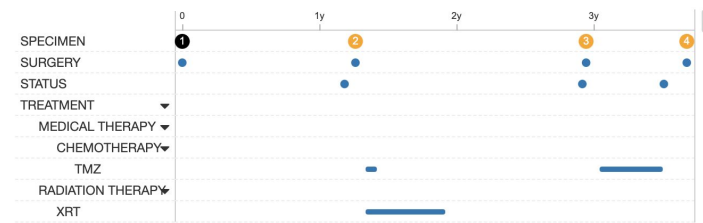
Group comparison



Gene-centric queries

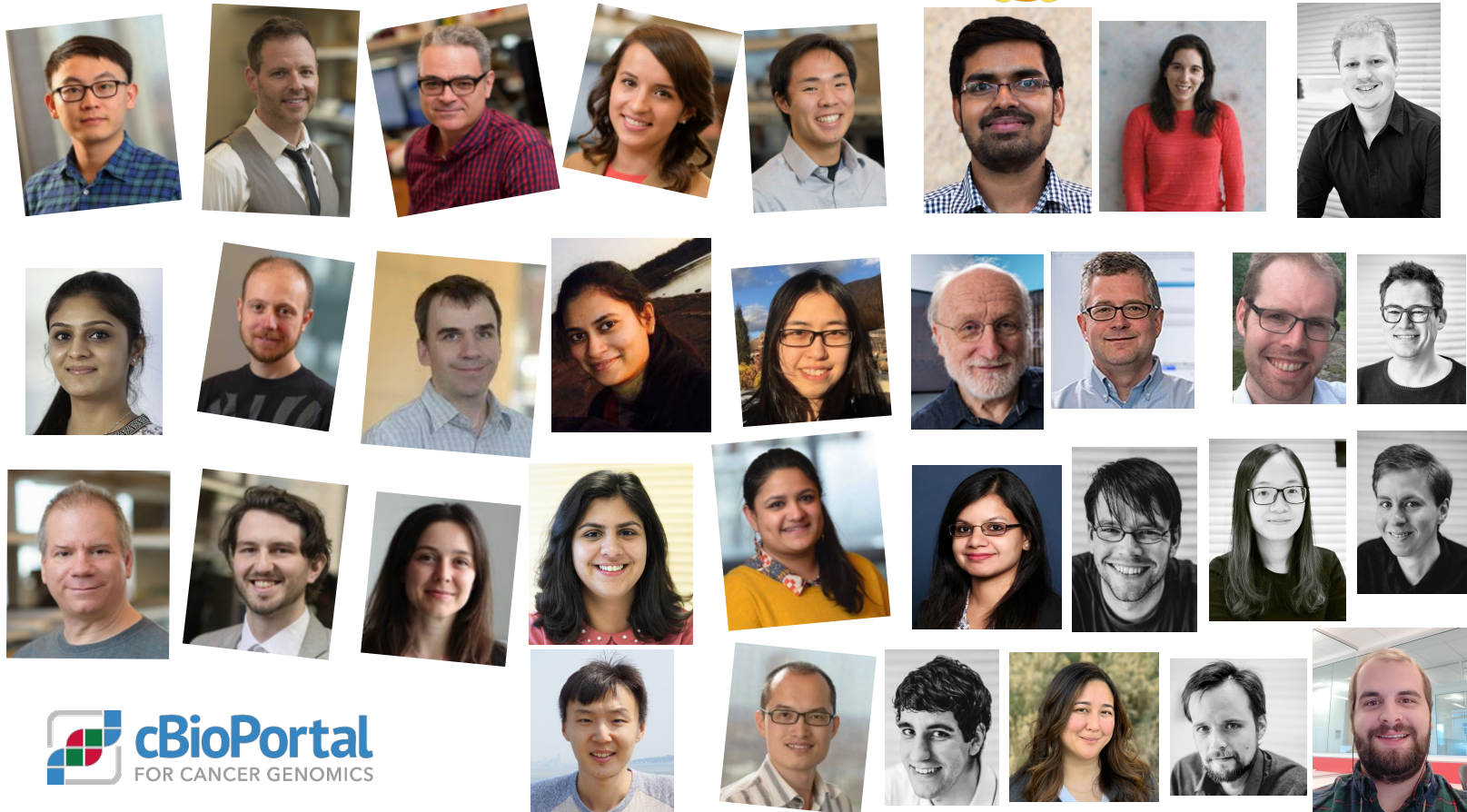


Patient visualization

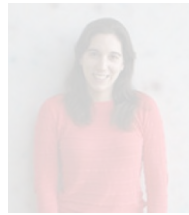


Samples	Gene	Protein Change	Annotation
1 2 3 4	IDH1	R132C	⊕ ⊖ ⚡ 🔥
1 2 3 4	TP53	R248Q	⊕ ⊖ ⚡ 🔥
3	ERC1	L283Fs*20	⊕
3	HSP90AB1	K72E	⊕
3	SPRNT	F404Lfs*3	⊕
1	OR10V1	R273W	⊕
1 2 3 4	TEAD3	P51L	⊕

cBioPortal has an awesome team 🤗



Say hi to us (Poster #39) 🙌



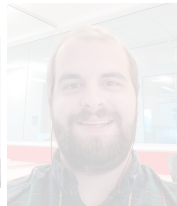
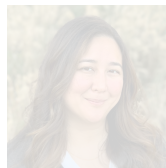
Aaron Lisman



Onur Sumer



Ino de Bruijn



Open source and open data



Code is open <https://github.com/cBioPortal/>

Data is open <https://github.com/cBioPortal/datahub/>

Free to download and install

Modifications and contributions are welcome

Maintained by multiple institutions



Memorial Sloan Kettering
Cancer Center



DANA-FARBER
CANCER INSTITUTE



UHN
Toronto General
Toronto Western
Princess Margaret
Toronto Rehab



The Children's Hospital of Philadelphia



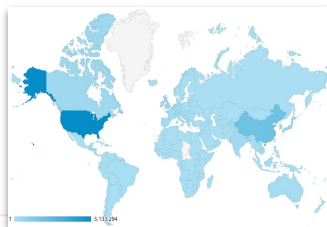
thehyve

cBioPortal is used a lot

Monthly Users ~34k

Total Citations ~18k

Curated Studies 325



> 7,000,000 visits to cBioPortal.org since 2011



>70 institutional instances globally

Weekly user visits

30,000
20,000
10,000

2006 2008 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021



Start of TCGA



First SU2C cBioPortal

DISCOVERY

Cerami et al. Cancer Discov.



Gap et al. Sci Signal. cBioPortal @MSK



cBioPortal @PMCC



cBioPortal @DFCI



Project GENE



cBioPortal @CHOP



ITCR funding



Start of HTAN



Count Me In



Webinars

Integrations with

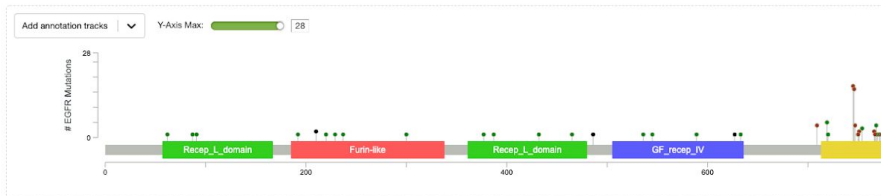
6



Projects

1. Clinical Interpretation of Variants In Cancer (CIViC)

EGFR



137 Mutations (page 1 of 6)

Sample ID	Cancer Type	Protein Change	Annotat
LUAD-NYU1142-Tumor	Lung Adenocarcinoma	T790M	↓
TCGA-49-4494-01	Lung Adenocarcinoma	T790M	↓
TCGA-L9-A50W-01	Lung Adenocarcinoma	T790M	↓
LUAD-F00134-Tumor	Lung Adenocarcinoma	L858R	↓
LUAD-NYU1142-Tumor	Lung Adenocarcinoma	L858R	↓
LUAD-NYU704-Tumor	Lung Adenocarcinoma	L858R	↓
LUAD-RT-501771-T...	Lung Adenocarcinoma	L858R	↓
LUAD-RT-501856-T...	Lung Adenocarcinoma	L858R	↓
LUAD_E01419-Tumor	Lung Adenocarcinoma	L858R	↓
TCGA-38-4827-01	Lung Adenocarcinoma	L858R	↓
TCGA-49-4494-01	Lung Adenocarcinoma	L858R	↓

CIViC Variants

EGFR - EGFR is widely recognized for its importance in cancer. Amplification and mutations have been shown to be driving events in many cancer types. Its role in non-small cell lung cancer, glioblastoma and basal-like breast cancers has spurred many research and drug development efforts. Tyrosine kinase inhibitors have shown efficacy in EGFR amplified tumors, most notably gefitinib and erlotinib. Mutations in EGFR have been shown to confer resistance to these drugs, particularly the variant T790M, which has been functionally characterized as a resistance marker for both of these drugs. The later generation TKI's have seen some success in treating these resistant cases, and targeted sequencing of the EGFR locus has become a common practice in treatment of non-small cell lung cancer. Overproduction of ligands is another possible mechanism of activation of EGFR. ERBB ligands include EGF, TGF- α , AREG, EPG, BTC, HB-EGF, EPR and NRG1-4 (for detailed information please refer to the respective ligand section).

T790M Entries: predictive: 38, prognostic: 2.

EGFR T790M was one of the very first mutations recognized to confer resistance to targeted therapies in non-small cell lung cancer. While successful in amplified EGFR, the efficacy of the first and second generation TKI's (erlotinib, gefitinib, neratinib) in treating patients harboring this mutation before treatment is notably lower. This lack of

CIViC

Misense

EGFR

RefSeq: NM_005228
 Ensembl: ENST00000275493
 CDS: CCDS5514
 UniProt: EGFR_HUMAN

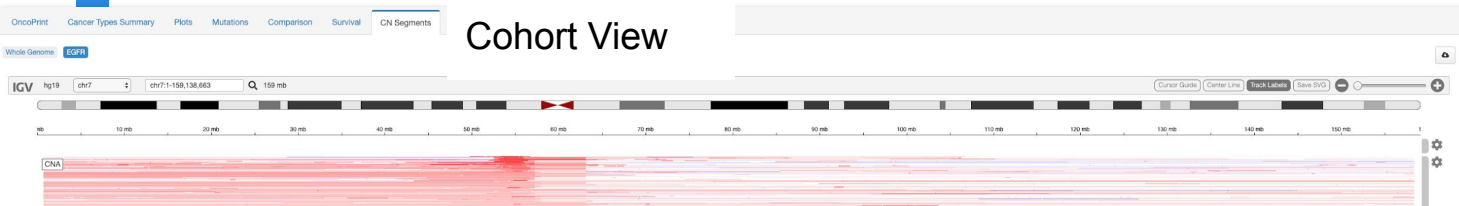
Somatic Mutation Frequency **10.2%**

82 Missense 5 Truncating
 50 Inframe 0 Other

View 3D Structure

Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
Amp	355	0.18	49
Gain	355	0.52	169
Gain	355	0.56	34
Gain	2247	0.20	65
Amp	2247	0.31	49
Gain	2247	0.30	48
Gain	2247	0.26	33
Amp	2247	0.68	57
ShallowDel	2247	0.22	71
Gain	2247	0.12	30
Gain	2247	0.20	68

2. Integrative Genomics Viewer (IGV)



3. Next Generation Clustered Heat Map (NG-CHM)

Summary Clinical Data Heatmaps CN Segments Selected: 92 patients | 92 samples

THE UNIVERSITY OF TEXAS
MD Anderson
Cancer Center

TCGA Next-Generation Clustered Heat Map (NG-CHM) Compendium

Click on any thumbnail image shown below to open in the NG-CHM viewer.

Single Study Maps
Choose one or more criteria:

Cancer Type
ACC - Adrenocortical carcinoma

Platform
Select platform

Heatmap Type
Select heat_map_type

See other Heat Map Collections
Bookmark Link for Current View

NG-CHM Viewer Help
Quick User Guide (Video)
Other User Guides / Videos

	Gene/Probe vs Sample	Gene/Probe vs Gene/Probe	Sample vs Sample
mRNA Expression	 tcga_maseqv2_acc_v2.0_gene_sample	 tcga_maseqv2_acc_v2.0_gene_gene	 tcga_maseqv2_acc_v2.0_sample_sample
Reverse Phase Protein Array	 tcga_rppa_acc_v2.0_protein_sample	 tcga_rppa_acc_v2.0_protein_protein	 tcga_rppa_acc_v2.0_sample_sample
miRNA Expression	 tcga_mirna_acc_v2.0_mirna_sample	 tcga_mirna_acc_v2.0_mirna_mirna	 tcga_mirna_acc_v2.0_sample_sample

4. Cancer Digital Slide Archive (CDSA)

Patient: TCGA-BK-A0CC, 69 years old, Endometrial Cancer (Uterine Serous Carcinoma/Uterine Papillary Serous Carcinoma), LIVING (10 months), DiseaseFree (10 months)
Samples: TCGA-BK-A0CC-01, Primary, Stage III

Uterine Corpus Endometrial Carcinoma (TCGA, Nature 2013)

Summary Clinical Data Pathology Report Tissue Image

CANCER Digital Slide Archive

EMORY WINSHIP CANCER INSTITUTE

Pathology Report


Layer Default Layer

Slides

TCGA-BK-A0CC

1/2 (6 slides)

- TCGA-BK-A0CC-01A-01-BS1.2
- TCGA-BK-A0CC-01A-02-BS2.8
- TCGA-BK-A0CC-01B-03-BS3.7
- TCGA-BK-A0CC-01B-04-BS4.5
- TCGA-BK-A0CC-01Z-00-DX1.4



5. R Package cBioPortalData - Bioconductor

cBioPortalData

cBioPortal data and MultiAssayExperiment

Overview

This project aims to import all cBioPortal datasets as [MultiAssayExperiment](#) objects in Bioconductor. It offers some advantages over the CDGS-R package:

1. The MultiAssayExperiment class explicitly links all assays to the patient clinical/pathological data
2. The MultiAssayExperiment class provides a [flexible API](#) including harmonized subsetting and reshaping to convenient wide and long formats.
3. It provides complete datasets, not just for subsets of genes
4. It provides automatic local caching, thanks to BioFileCache.

MultiAssayExperiment Cheatsheet

Summary of the MultiAssayExperiment API (Ramos et al. Can. Res. 2017; DOI: 10.1158/0008-5472.CAN-17-0344)

Category and Function	Description	Returned class
Constructors		
MultiAssayExperiment	Create a MultiAssayExperiment object	MultiAssayExperiment
ExperimentList	Create an ExperimentList from a List or list	ExperimentList
Accessors		
colData	Get or set data that describe patients / biological units	DataFrame
experiments	Get or set the list of experimental data objects as original classes	ExperimentList
assays	Get the list of experimental data numeric matrices	SimpleList
assay	Get the first experimental data numeric matrix	matrix, matrix-like
sampleMap	Get or set the map relating observations to subjects	DataFrame
metadata	Get or set additional data descriptions	list
rownames	Get row names for all experiments	CharacterList
colnames	Get column names for all experiments	CharacterList
Subsetting		
mae[i, j, k]	Get rows, columns, and/or experiments	MultiAssayExperiment
mae[, i,]	i: GRanges, character, integer, logical, List, list	MultiAssayExperiment
mae[, , j,]	j: character, integer, logical, List, list	MultiAssayExperiment
mae[, , , k]	k: character, integer, logical	MultiAssayExperiment
mae[[n]]	Get or set object of arbitrary class from experiments	(varies)
mae[[n]]	n: character, integer, logical	(varies)
mae\$column	Get or set colData column	vector (varies)
getWithColData(mae, k)	Extract a single assay with associated colData; k: character, integer	(varies)
Management		
complete.cases	Identify subjects with complete data in all experiments	vector (logical)
deduplicated.replicated*	Identify subjects with replicate observations per experiment	list of LogicalLists
anyReplicated	Displays whether there are any replicate observations in each assay	vector (logical)
mergeReplicates	Merge replicate observations within each experiment, using function	MultiAssayExperiment
intersectRows	Return features that are present for all experiments	MultiAssayExperiment
intersectColumns	Return subjects with data available for all experiments	MultiAssayExperiment
prepMultiAssay	Troubleshoot common problems when constructing main class	list

Links

Browse source code at
<https://github.com/waldronlab/cBioPortalData/>

Report a bug at
<https://github.com/waldronlab/cBioPortalData/issues>

License

AGPL_3

Developers

Levi Waldron

Author

Marcel Ramos

Author, maintainer

Dev status

build passing

build failing

6. NDEx Pathway Analysis



Modify Query **MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)** Queried genes are altered in 5494 (53%) of queried patients
Samples with mutation and CNA data (10336 patients / 10945 samples) - CDKN2A, MDM2 & 2 other genes 5746 (52%) of queried samples

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Comparison/Survival CN Segments **Pathways** Download

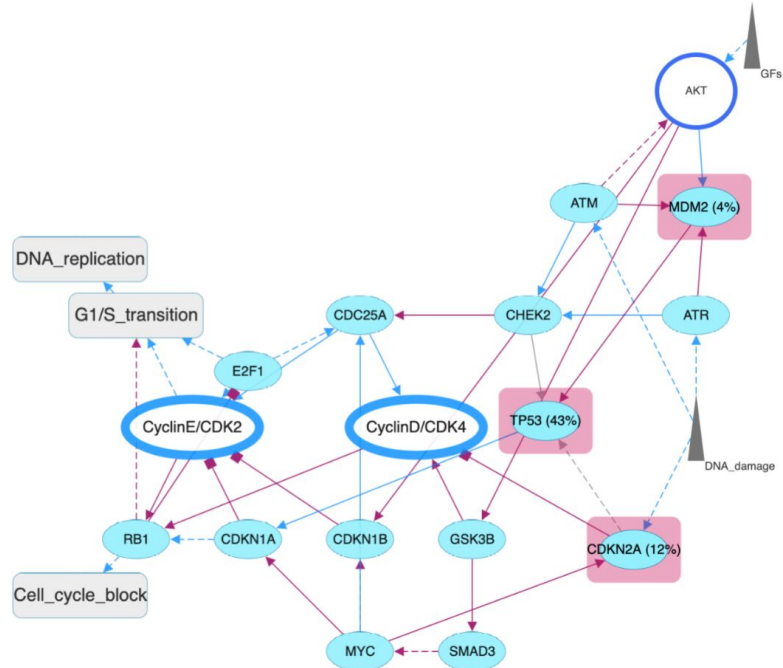
Choose Pathway Source: [PathwayMapper](#) **NDEx**

Sort by Similarity p-Value Overlap

- WP2261 - Glioblastoma signaling pathways - Homo sapiens**
4 / 82 unique genes Similarity: 0.24
CDKN2A MDM2 MDM4 TP53
- Cell cycle G1/S phase transition**
3 / 21 unique genes Similarity: 0.24
CDKN2A MDM2 TP53
- WP2516 - ATM signaling pathway - Homo sapiens**
3 / 40 unique genes Similarity: 0.23
MDM2 MDM4 TP53
- Malignant Melanoma**
3 / 31 unique genes Similarity: 0.21
CDKN2A MDM2 TP53
- p53 pathway**
4 / 57 unique genes Similarity: 0.20
CDKN2A MDM2 MDM4 TP53
- WP1742 - TP53 network - Homo sapiens**
3 / 20 unique genes Similarity: 0.19
CDKN2A MDM2 TP53
- Luminal Breast Cancer**
3 / 39 unique genes Similarity: 0.18
CDKN2A MDM2 TP53
- Cell cycle G2/M phase transition**
3 / 34 unique genes Similarity: 0.16
CDKN2A MDM2 TP53

Signor: Cell cycle: G1/S phase transition

Layout Default



6. NDEx Pathway Analysis



Modify Query **MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)**
Samples with mutation and CNA data (10336 patients / 10945 samples) - CDKN2A, MDM2 & 2 other genes

Queried genes are altered in 5494 (53%) of queried patients
5746 (52%) of queried samples

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Comparison/Survival CN Segments **Pathways** Download

Choose Pathway Source: **PathwayMapper** **NDEx**

Sort by Similarity p-Value Overlap

WP2261 - Glioblastoma signaling pathways - Homo sapiens
4 / 82 unique genes Similarity: 0.24
CDKN2A MDM2 MDM4 TP53

Cell cycle G1/S phase transition
3 / 21 unique genes Similarity: 0.24
CDKN2A MDM2 TP53

WP2516 - ATM signaling pathway - Homo sapiens
3 / 40 unique genes Similarity: 0.23
MDM2 MDM4 TP53

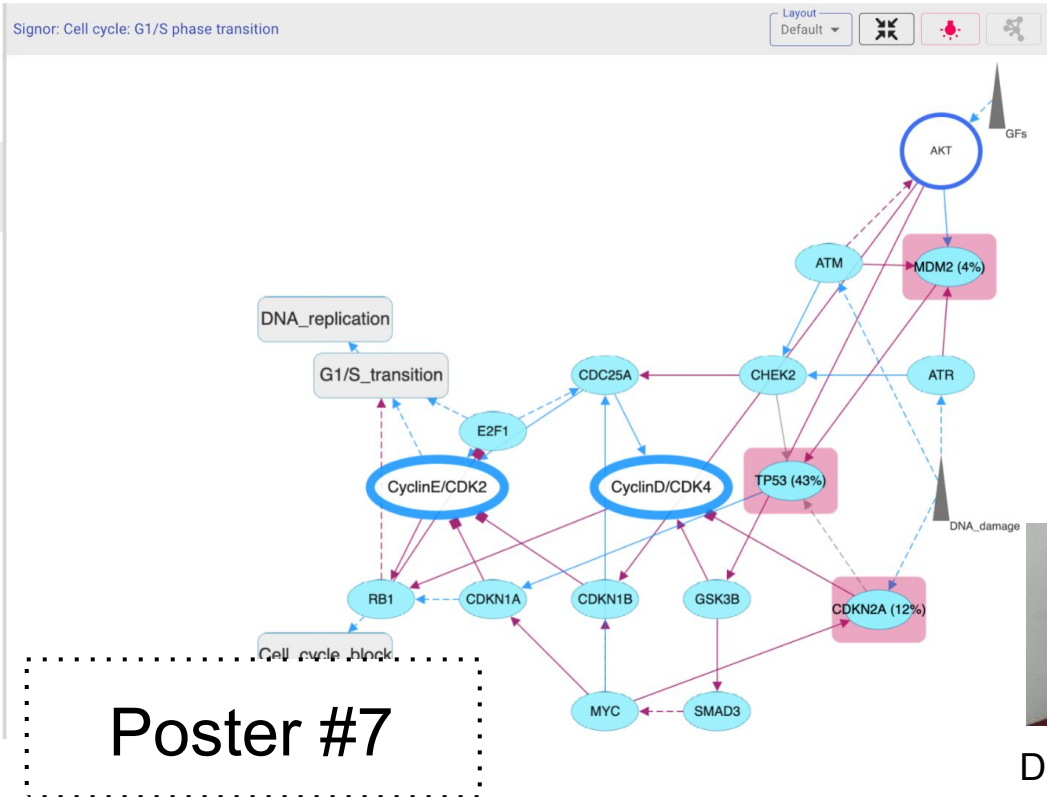
Malignant Melanoma
3 / 31 unique genes Similarity: 0.21
CDKN2A MDM2 TP53

p53 pathway
4 / 57 unique genes Similarity: 0.20
CDKN2A MDM2 MDM4 TP53

WP1742 - TP53 network - Homo sapiens
3 / 20 unique genes Similarity: 0.19
CDKN2A MDM2 TP53

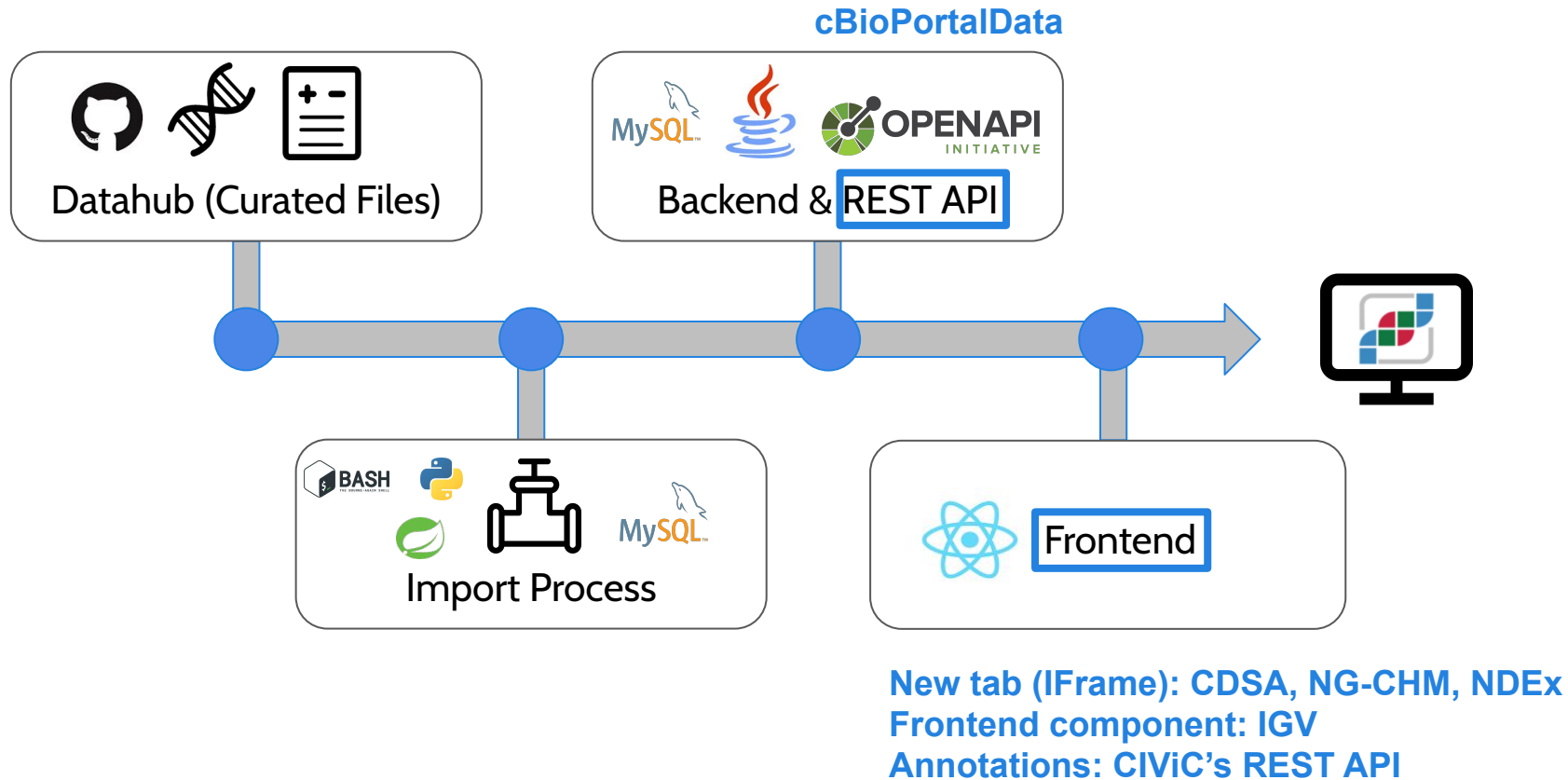
Luminal Breast Cancer
3 / 39 unique genes Similarity: 0.18
CDKN2A MDM2 TP53

Cell cycle G2/M phase transition
3 / 34 unique genes Similarity: 0.16
CDKN2A MDM2 TP53

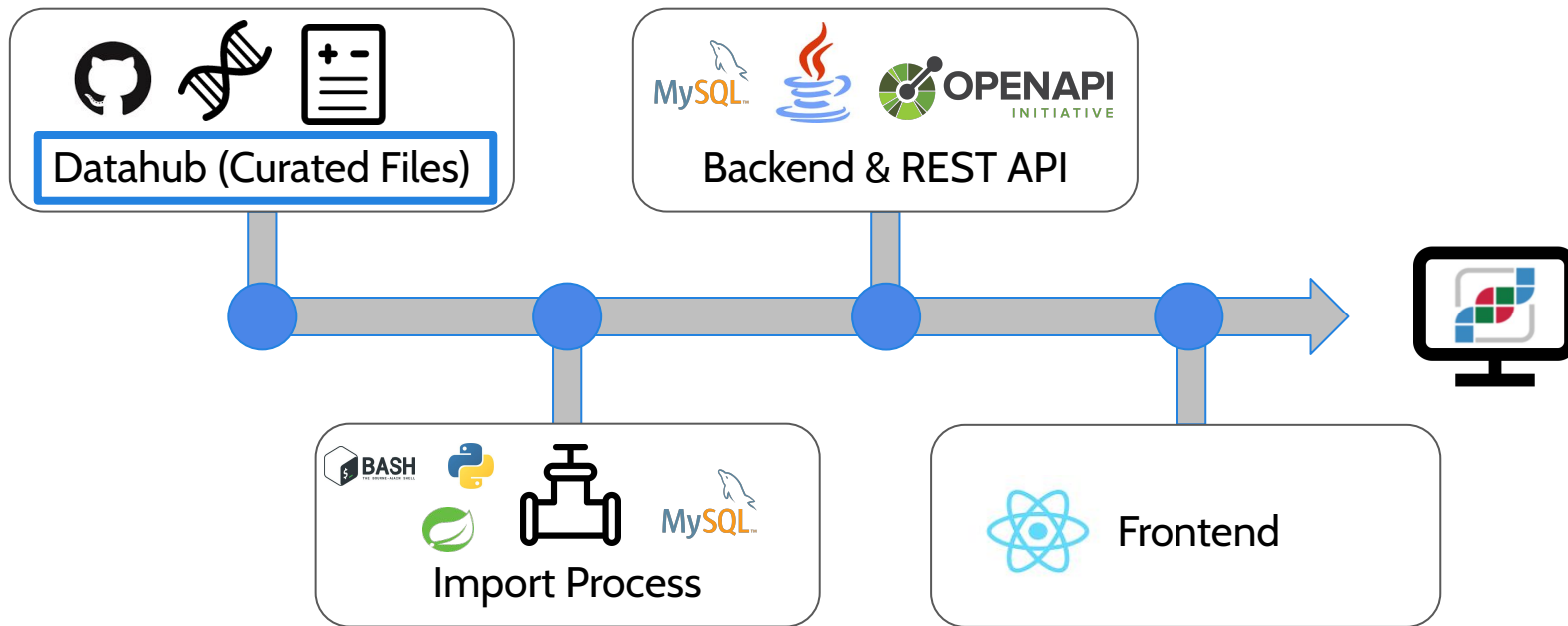


Dexter Pratt

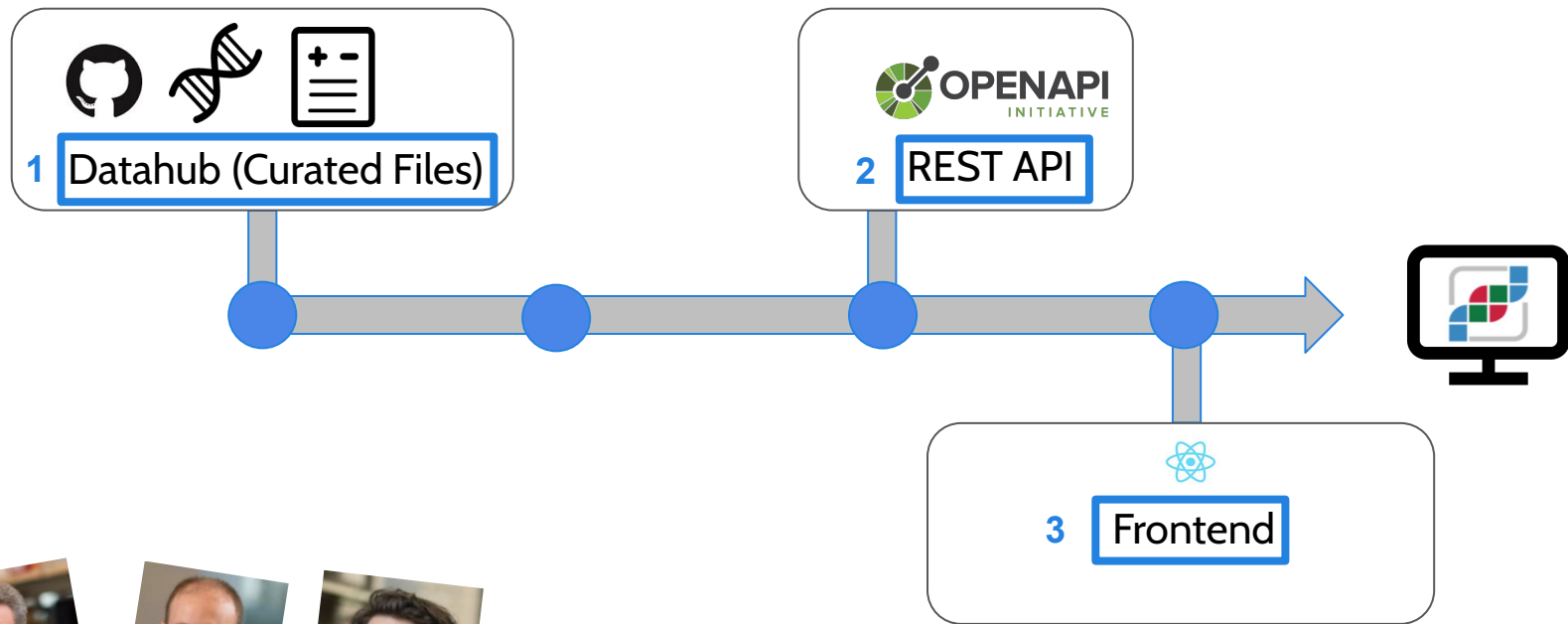
Existing **ITCR** Collaborations



Potential **ITCR** Collaborations



Potential **ITCR** Collaborations



Poster #39 🙌

