*AMARETTO for network biology and medicine:

linking diseases, drivers, targets and drugs

via graph-based fusion of multi-omics, clinical, imaging and perturbation data

Nathalie Pochet, Ph.D.

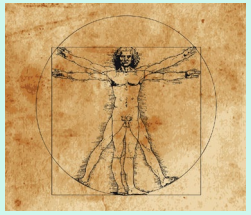
Assistant Professor, Harvard Medical School
Associate Scientist, Brigham and Women's Hospital
Associate Member, Broad Institute of MIT and Harvard

http://portals.broadinstitute.org/pochetlab/
http://portals.broadinstitute.org/pochetlab/amaretto.html
npochet@broadinstitute.org, npochet@bwh.harvard.edu

Funded by NIH NCI ITCR

Big Data: multi-omics, clinical, imaging, perturbations,… across biological systems

Patient Studies



Decipher disease heterogeneity

Multi-omics: (epi)genetics & functional genomics

Driver discovery via regulatory network inference

Bulk & single-cell multi-omics

Radiographic & histopathology imaging

Model Systems

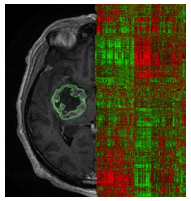
Perturbation studies

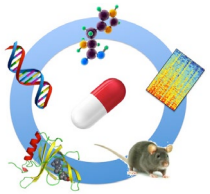
Genetic perturbations for driver discovery

Chemical perturbations for drug discovery

Multi-omics & imaging

Genetic & chemical & etiology perturbations





Disease discovery

Driver discovery

Drug discovery

Data-driven hypothesis generators based on multimodal and multiscale big data fusion?

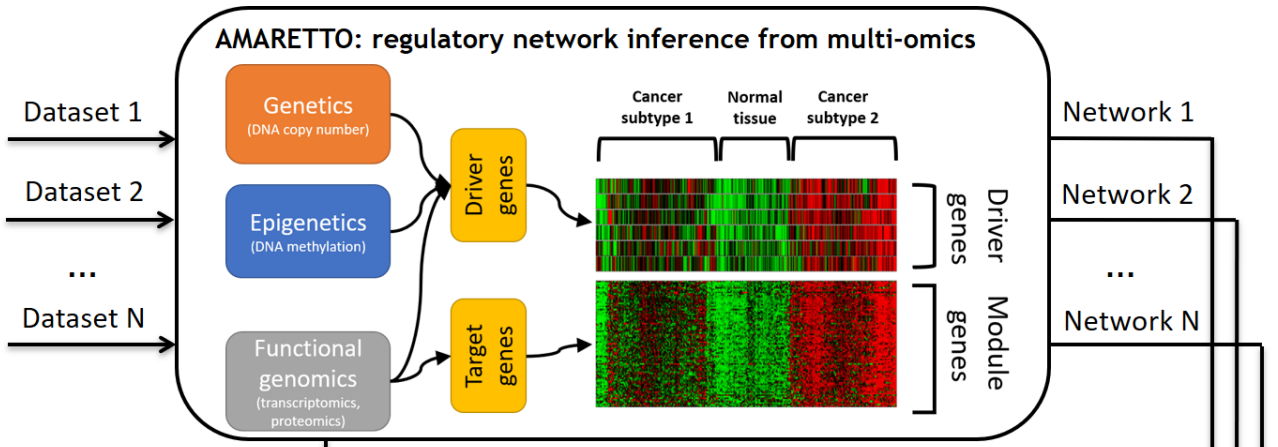
*AMARETTO

links **diseases**, **drivers**, **targets** and **drugs**

via network graph-based fusion
of **multi-omics**, **clinical**, **imaging** and **perturbation** data
across model systems and patient studies of complex disease

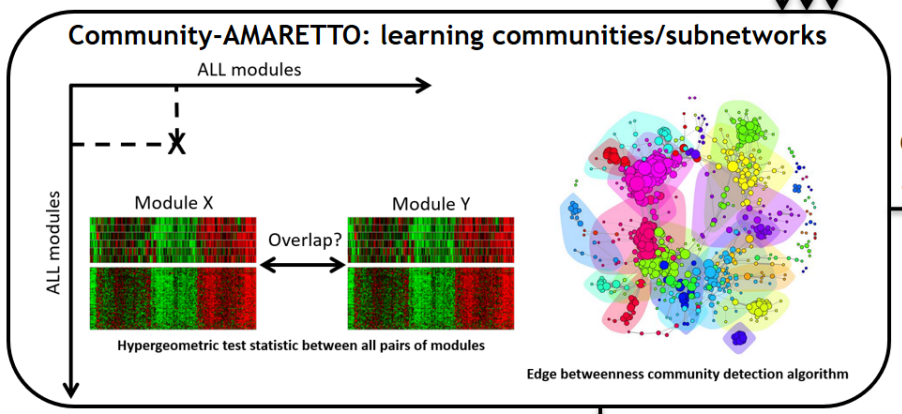
software toolbox for **network biology and medicine**
towards developing a **data-driven platform** for **diagnostic, prognostic and therapeutic decision-making in complex disease**

*AMARETTO framework



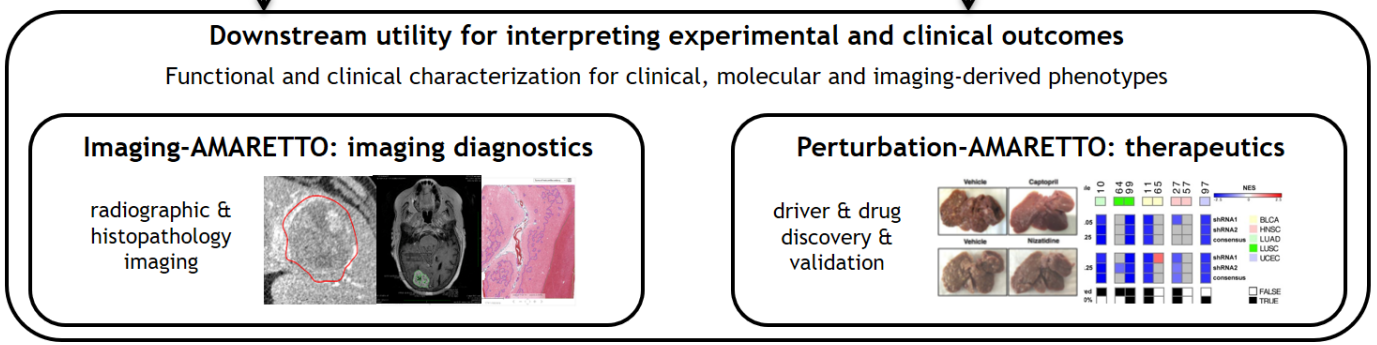
Core: AMARETTO

Multimodal data fusion within biological systems



Core: Community-AMARETTO

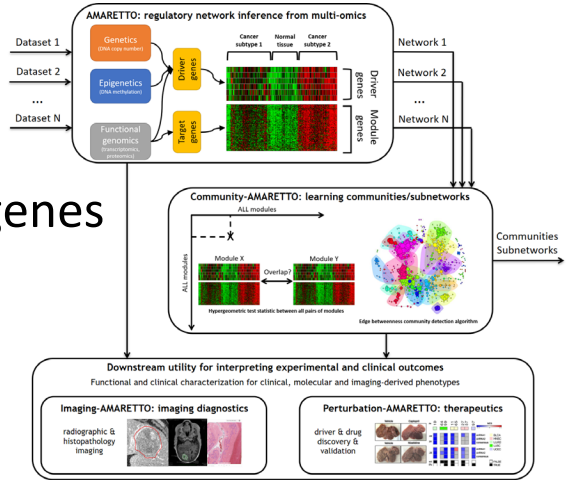
Multiscale data fusion across biological systems



Downstream utility:

Interpreting experimental & clinical outcomes

*AMARETTO discovers drugs reversing drivers and targets in complex disease



AMARETTO

- ❑ learns networks of regulatory circuits (modules) - circuits of drivers and target genes
- ❑ infers networks from functional genomics or multi-omics data
- ❑ associates circuits to clinical, molecular and imaging-derived phenotypes
- ❑ learns networks within each biological system (e.g., model systems or patients)

Community-AMARETTO

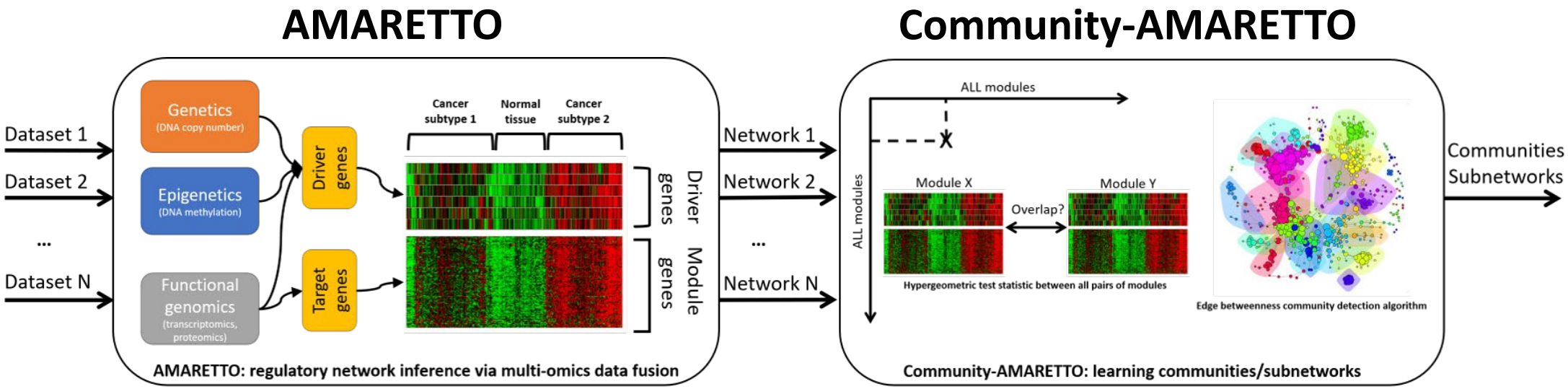
- ❑ learns subnetworks (communities) of regulatory circuits shared or distinct across multiple networks
- ❑ learns networks across multiple biological systems (e.g., model systems and patients, cohorts and individuals, diseases and etiologies, *in vitro* and *in vivo* systems)

Perturbation-AMARETTO

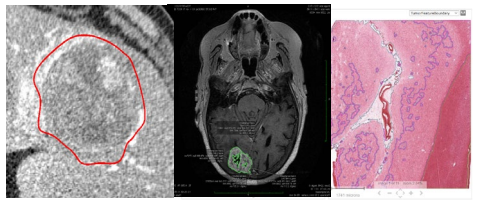
- ❑ maps driver (genetic) perturbations in model systems onto patient-derived networks
- ❑ maps drug (chemical) perturbations in model systems onto patient-derived networks
- ❑ identifies perturbations reversing disease-associated behavior, not affecting normal behavior
- ❑ prioritizes lead drivers, targets and drugs for follow-up with experimental validation

➢ ***AMARETTO links diseases – drivers – targets – drugs**

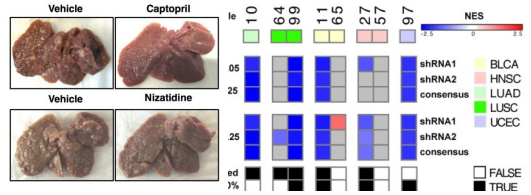
The *AMARETTO framework



non-invasive & histopathology imaging



driver & drug discovery & validation

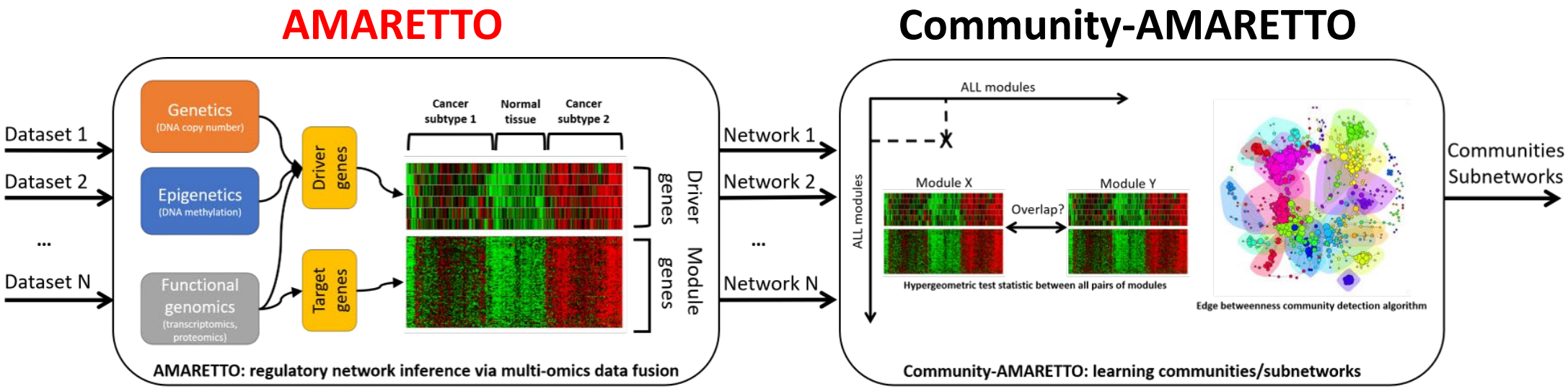


The *AMARETTO framework:

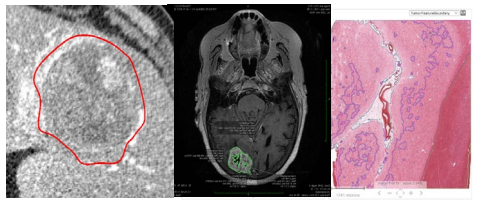
1. the **AMARETTO algorithm** for inferring regulatory networks via multi-omics and imaging data fusion

2. the **Community-AMARETTO algorithm** for learning subnetworks shared/distinct across systems and diseases

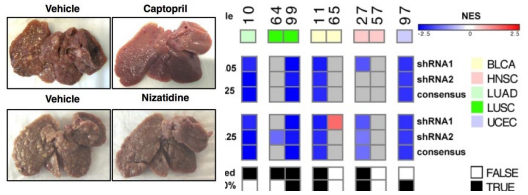
The *AMARETTO framework



non-invasive & histopathology imaging



driver & drug discovery & validation

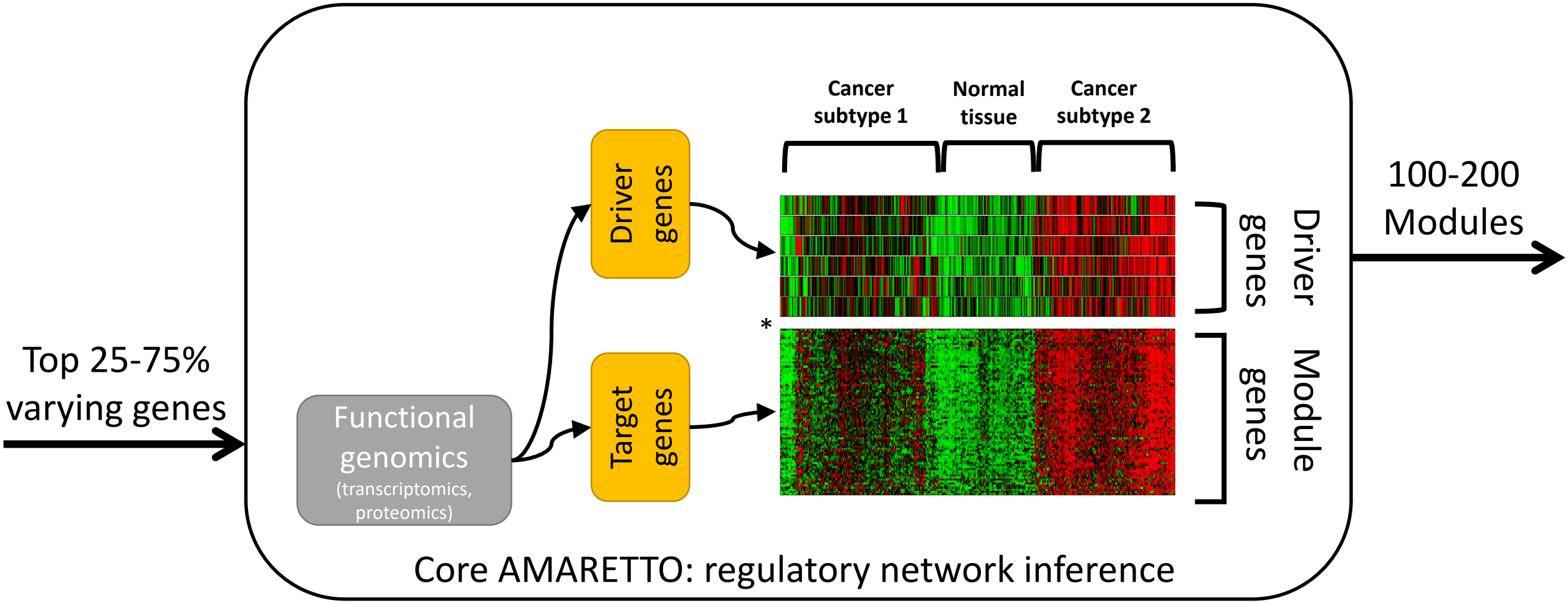


The *AMARETTO framework:

1. the **AMARETTO algorithm** for inferring regulatory networks via multi-omics and imaging data fusion

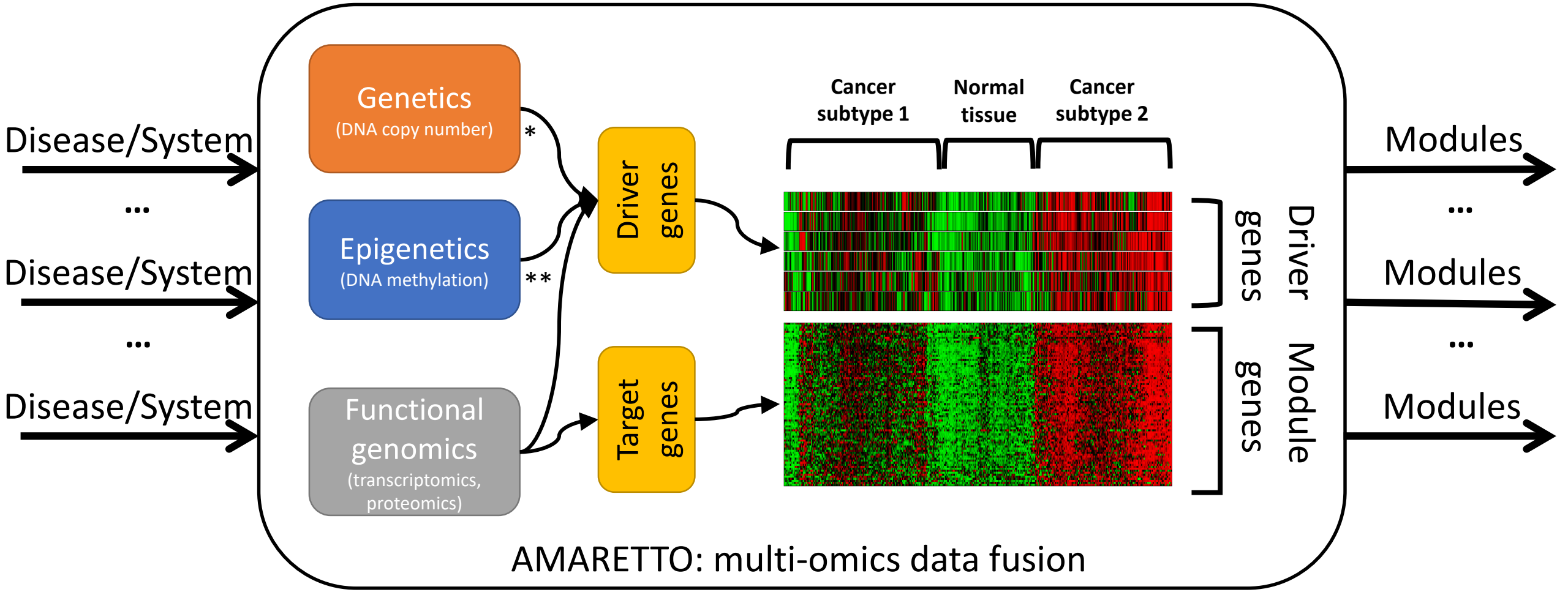
2. the **Community-AMARETTO algorithm** for learning subnetworks shared/distinct across systems and diseases

*AMARETTO for regulatory network inference within systems and diseases



(*) Regularized regression: Lee *et al.,* PLoS Genetics 2009; Zou and Hastie, J R Stat Soc 2005; Tibshirani, J R Stat Soc 1996

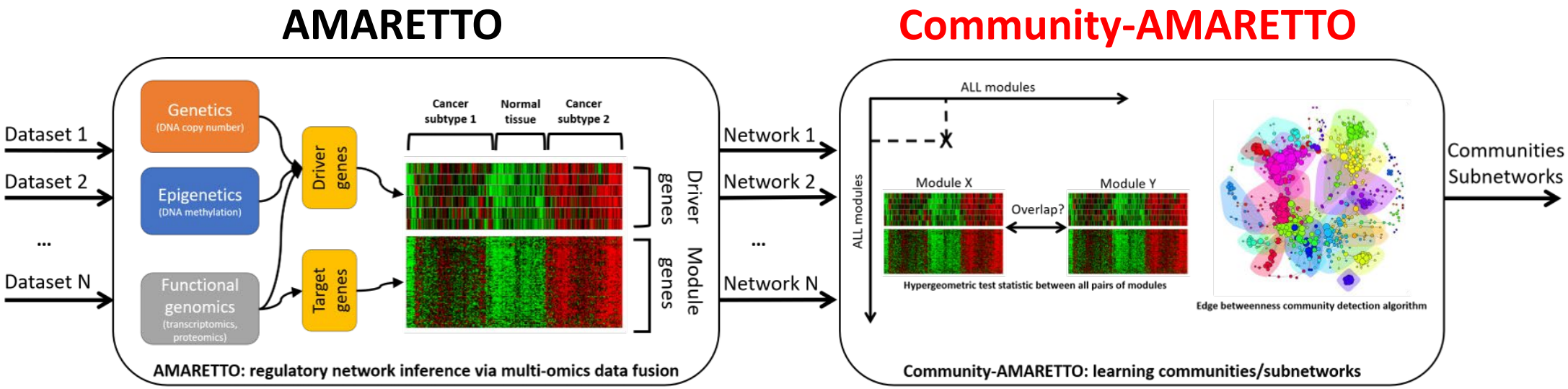
*AMARETTO for multi-omics data fusion in multiple systems and diseases



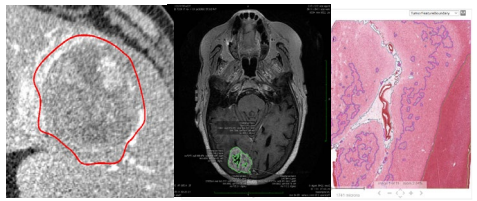
(*) GISTIC: Mermel *et al.*, Genome Biology 2011; Beroukhim *et al.*, Nature 2010

(**) MethylMix: Gevaert, Bioinformatics 2015; Gevaert *et al.*, Genome Biology 2015; Cedoz *et al.*, Bioinformatics 2018

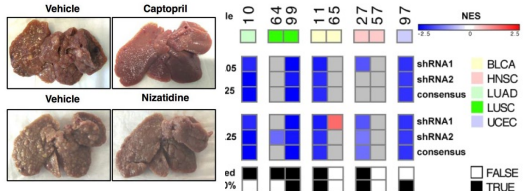
The *AMARETTO framework



non-invasive & histopathology imaging



driver & drug discovery & validation

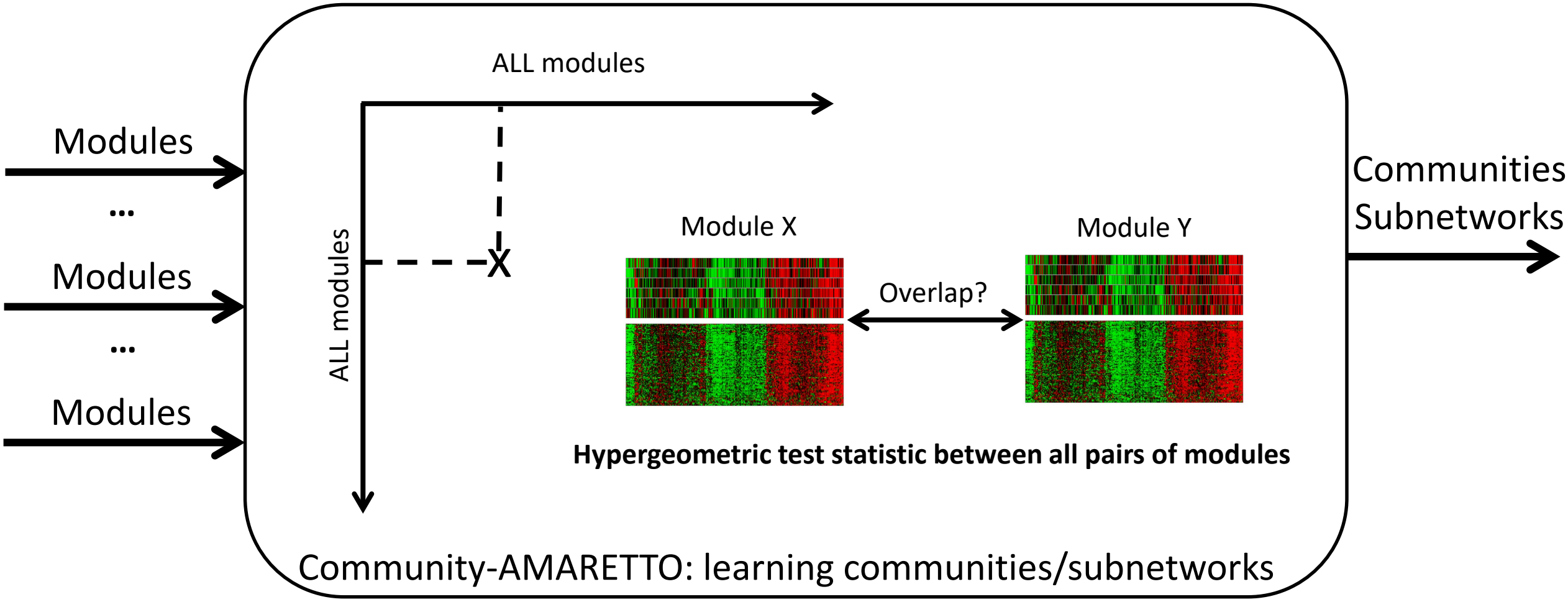


The *AMARETTO framework:

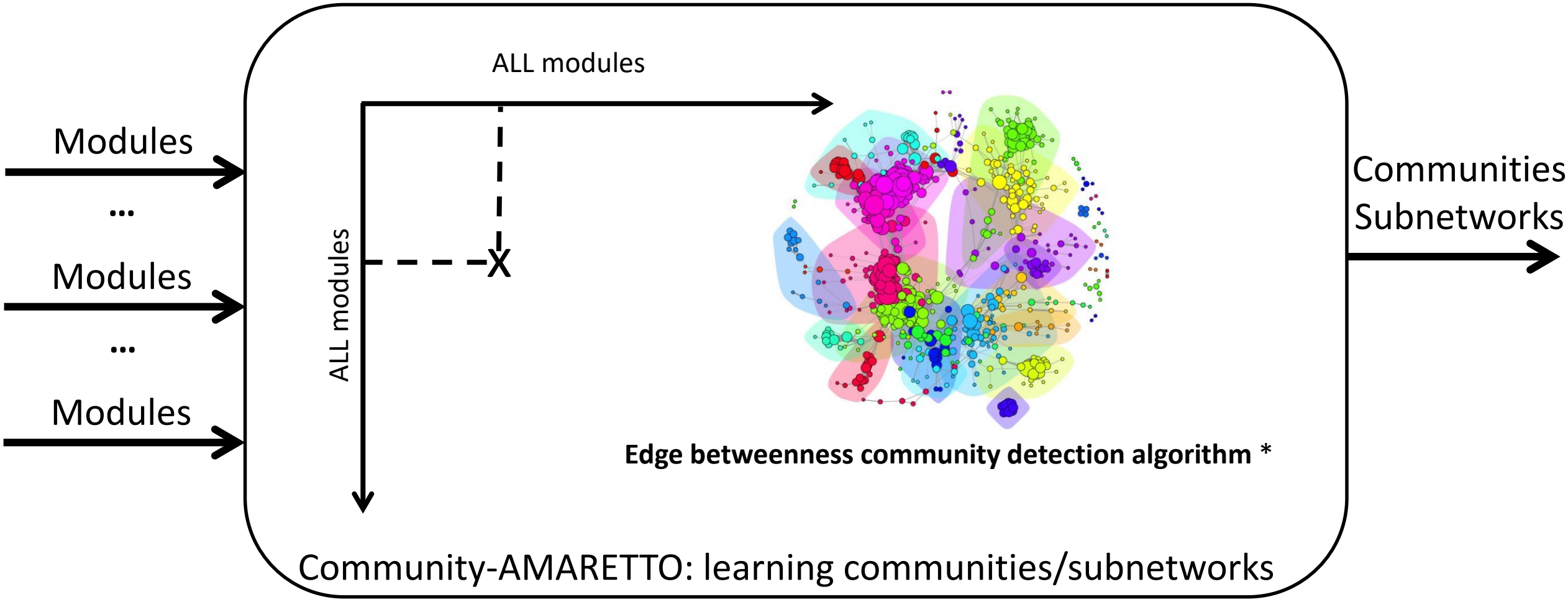
1. the **AMARETTO algorithm** for inferring regulatory networks via multi-omics and imaging data fusion

2. the **Community-AMARETTO algorithm** for learning subnetworks shared/distinct across systems and diseases

*AMARETTO for learning subnetworks across systems and diseases



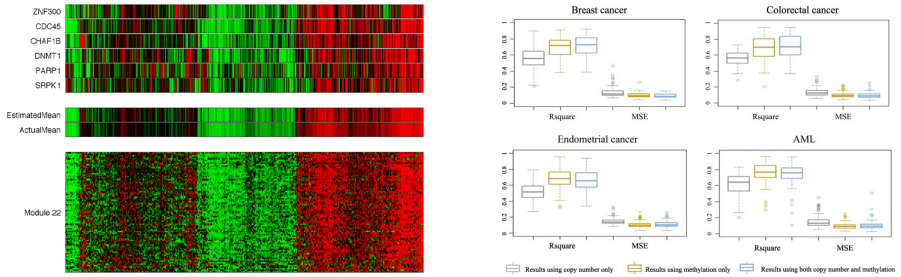
*AMARETTO for learning subnetworks across systems and diseases



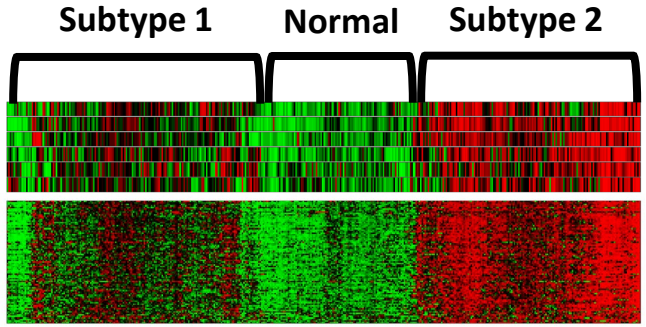
(*) Girvan and Newman, Physical Review E. 2004

Functionalities for optimization and downstream analytics

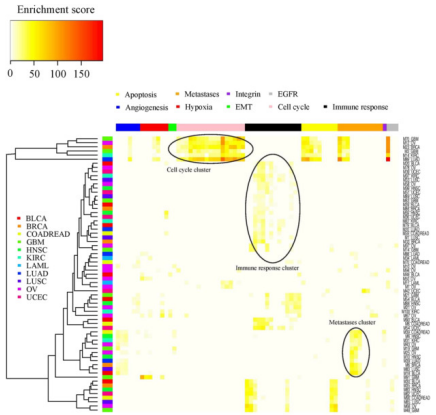
Optimal generalization performance



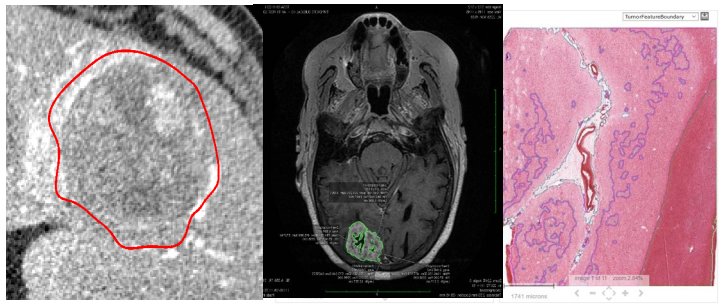
Stratification for disease phenotypes



Annotation of functional categories

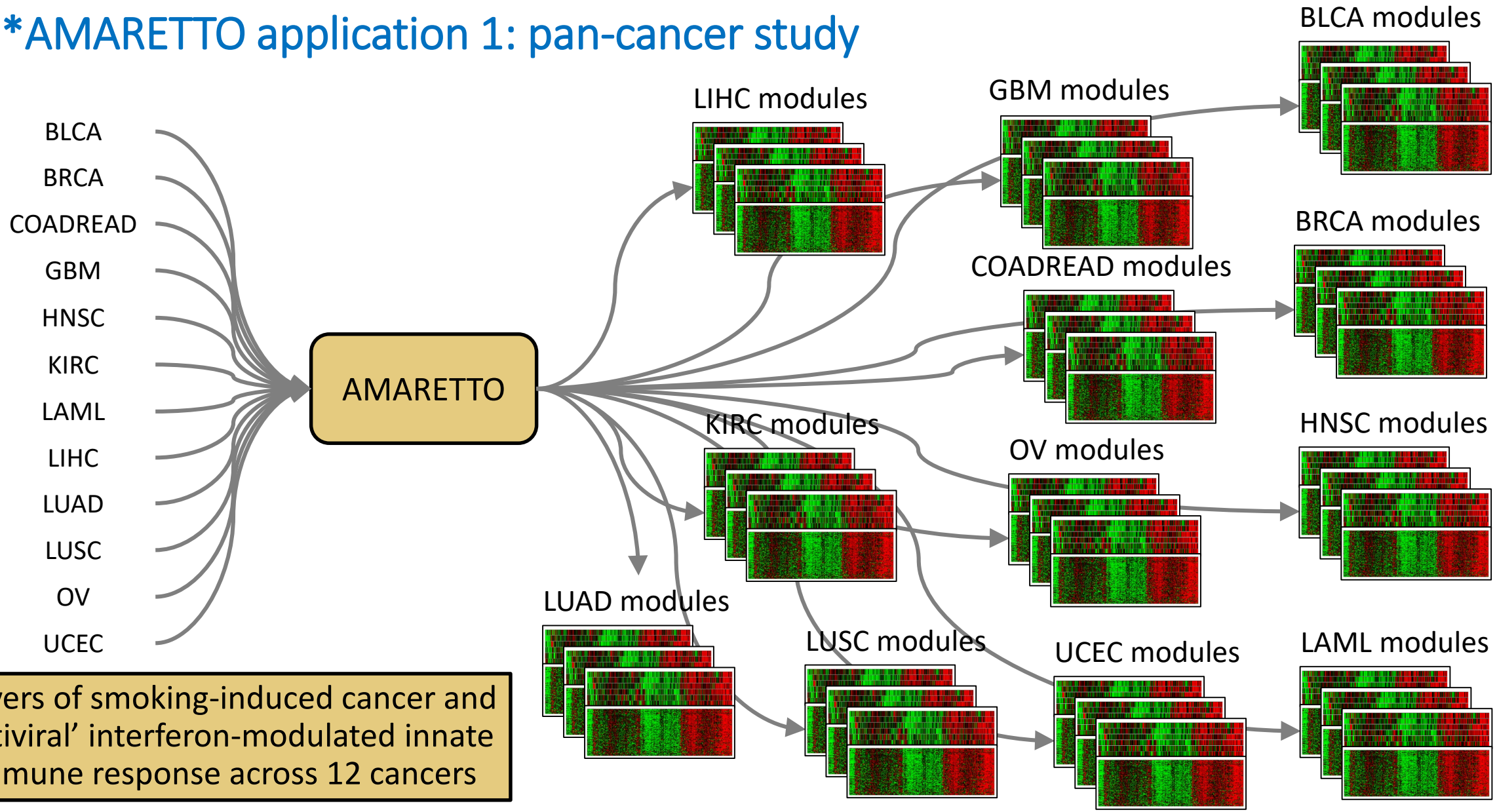


Association with imaging features



radiographic & histopathology imaging

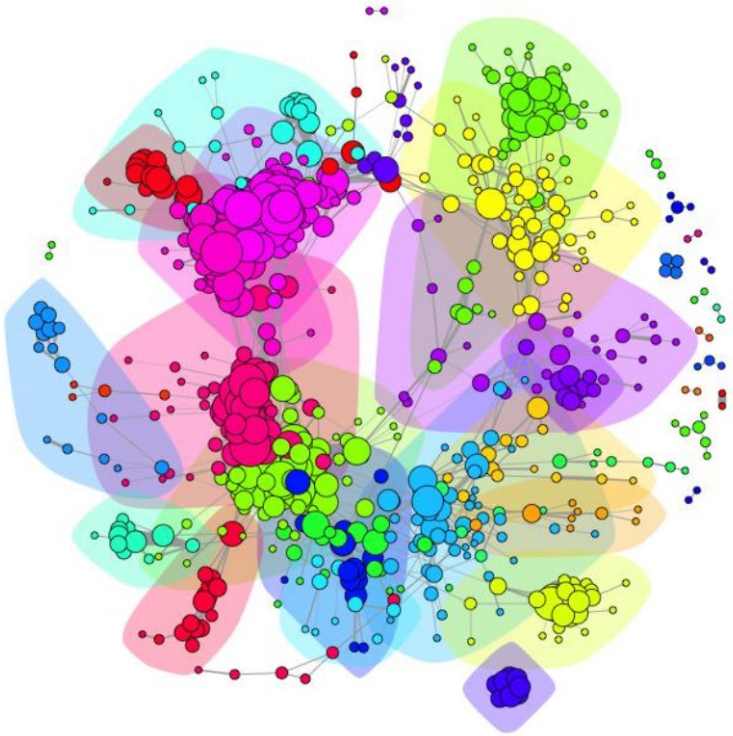
*AMARETTO application 1: pan-cancer study



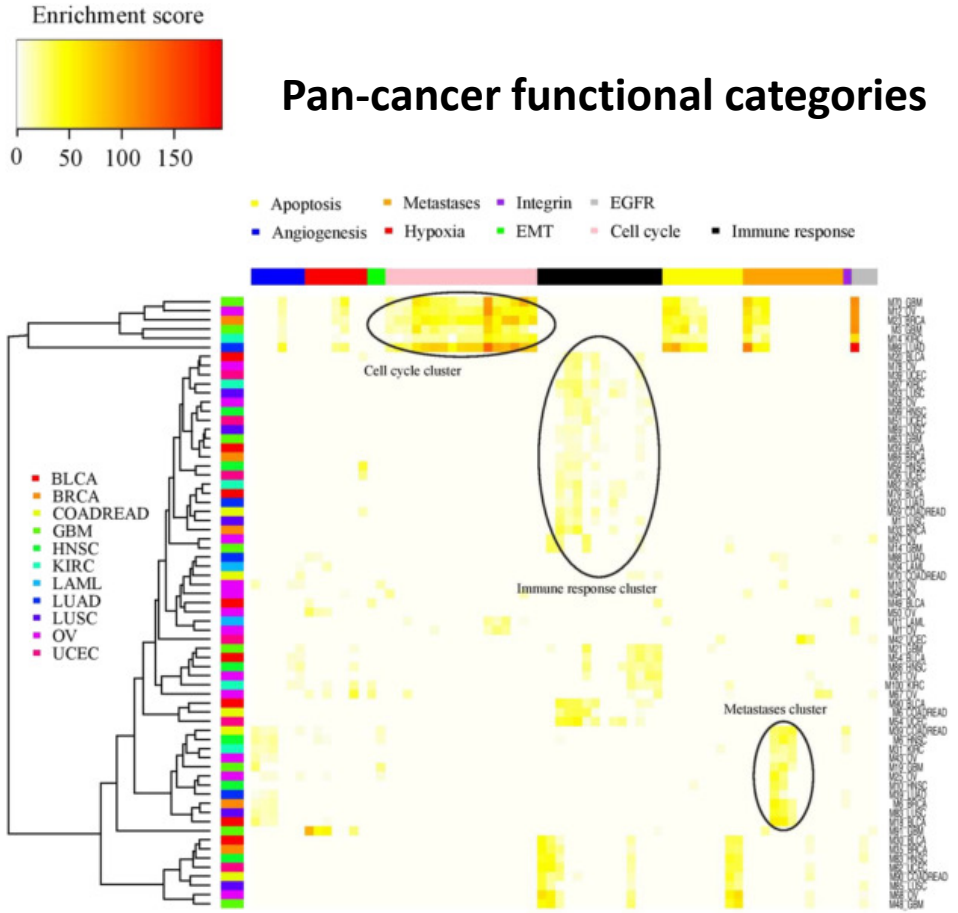
Drivers of smoking-induced cancer and 'antiviral' interferon-modulated innate immune response across 12 cancers

*AMARETTO application 1: pan-cancer study

Pan-cancer communities or subnetworks



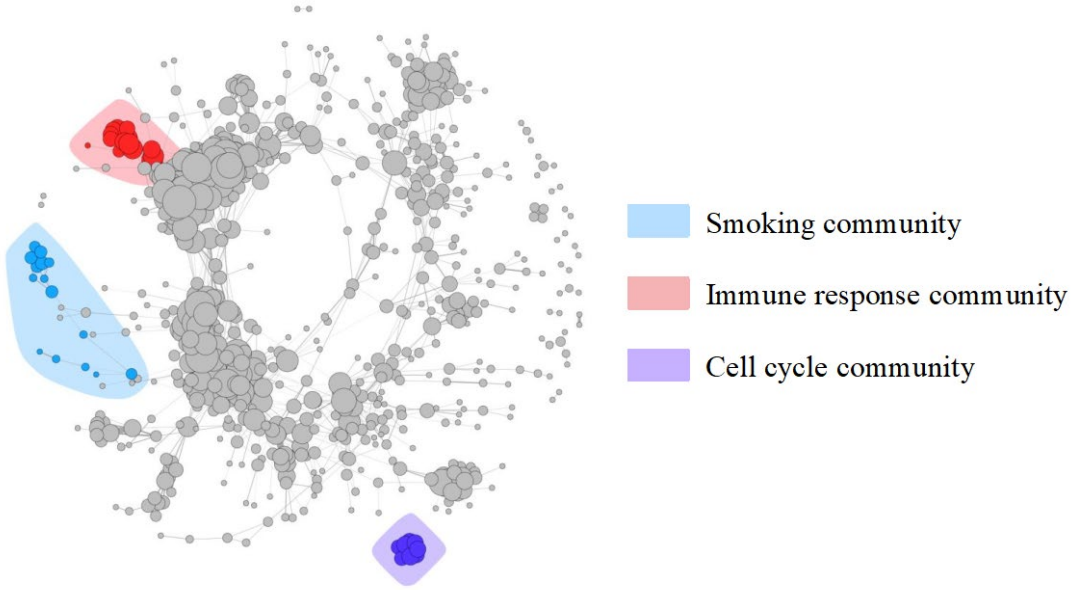
Pan-cancer functional categories



⇒ **AMARETTO captures hallmarks of cancer**

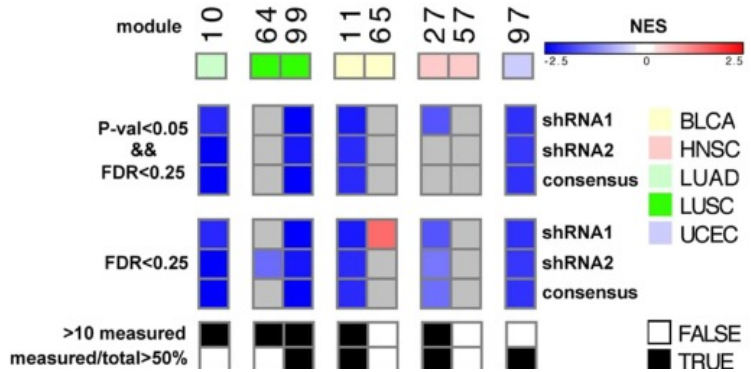
*AMARETTO application 1: pan-cancer study

Driver discovery



- OAS2 pan-cancer driver of 'antiviral' interferon-modulated innate immune response
- GPX2 pan-cancer driver of smoking-induced cancer

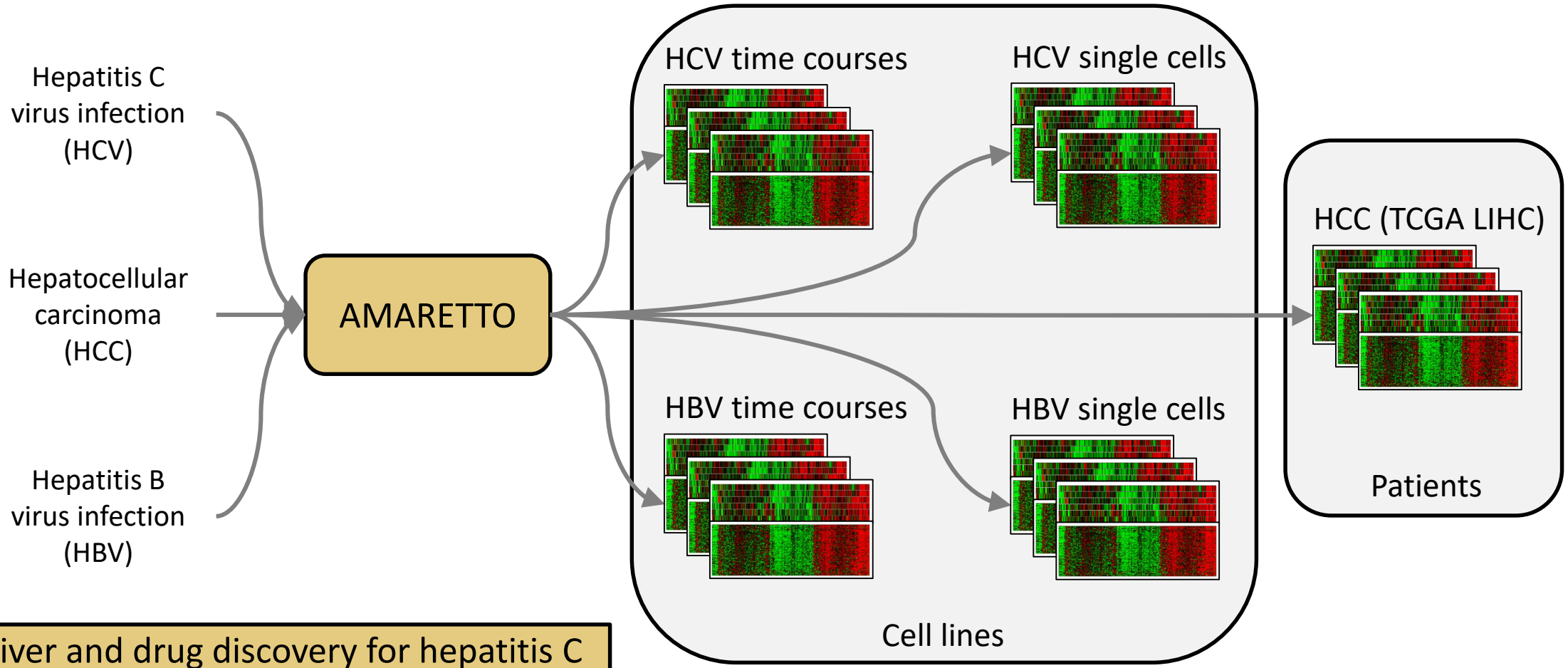
Driver validation



Genetic perturbation of GPX2 in the A549 (LUAD) cell line
⇒ Knocking down GPX2 represses target genes in GPX2-regulated modules

⇒ AMARETTO facilitates identification of known and novel cancer drivers and their targets

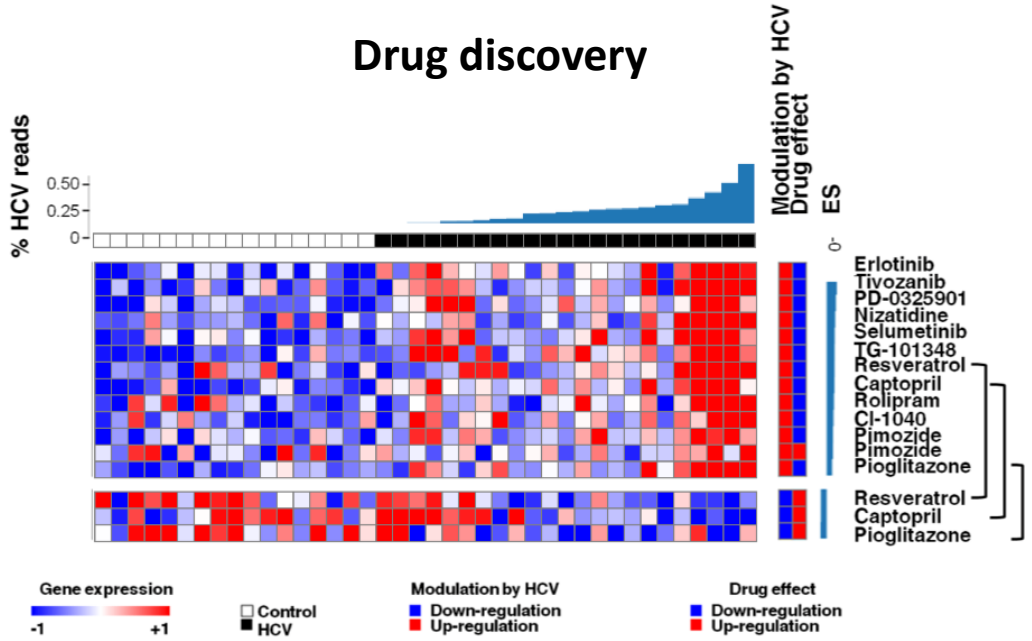
*AMARETTO application 2: virus-induced cancer



Driver and drug discovery for hepatitis C (HCV) and hepatitis B (HBV) virus-induced hepatocellular carcinoma (HCC)

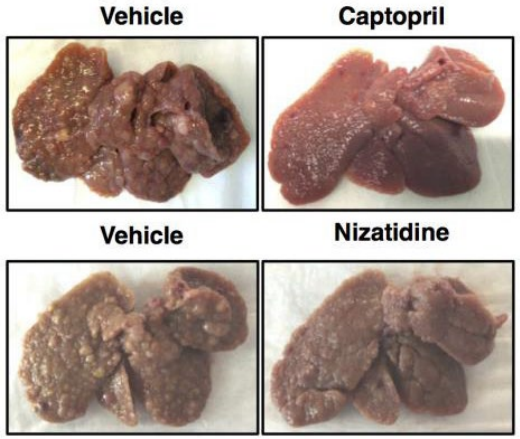
*AMARETTO application 2: virus-induced cancer

Drug discovery



Chemical perturbations in cell lines
Predict which drugs can reverse disease-associated modules
Alternative treatments with less severe adverse effects

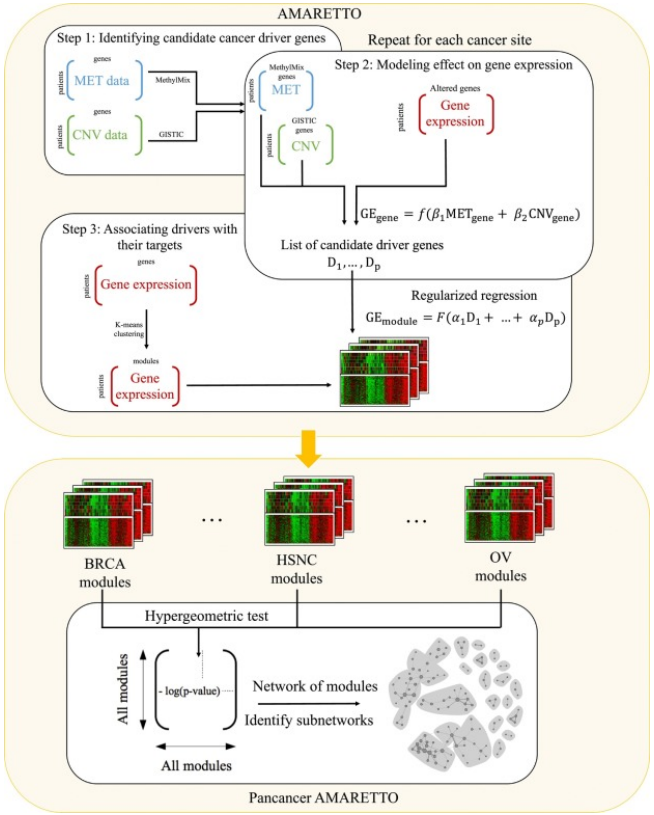
Drug validation

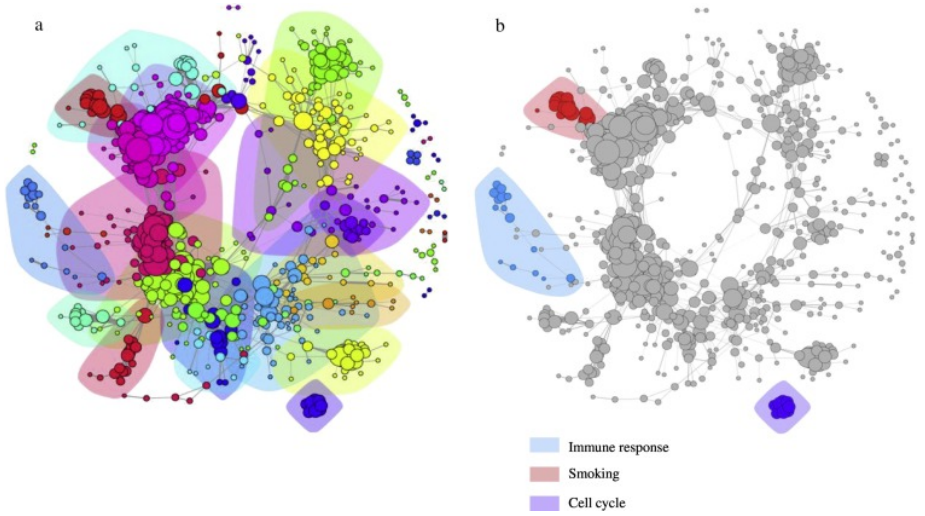


Experimental validation of drugs in rat models
⇒ Two novel compounds attenuate HCC development
⇒ Safe and low-cost approach for chemoprevention of HCC?

⇒ AMARETTO facilitates identification of known and novel drug compounds and how they modulate cancer drivers and their targets

*AMARETTO





Champion *et al.*, EBioMedicine 2018

***AMARETTO:**

1. Captures hallmarks of cancer

2. Facilitates identification of known and novel cancer drivers and their targets

3. Facilitates identification of known and novel drug compounds and how they modulate cancer drivers and their targets

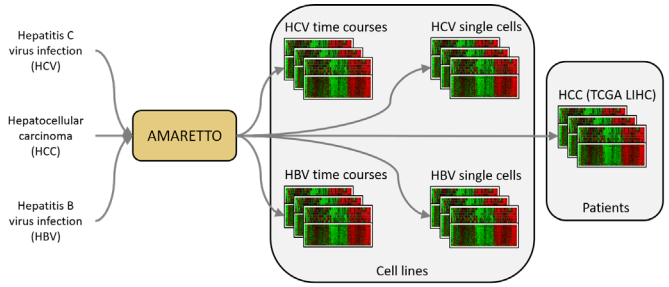
Use Cases: integrating multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer

1. A study of hepatitis C and B virus-induced hepatocellular carcinoma (LIHC) with driver and drug discovery for chemoprevention across pan-etiologies of hepatocellular carcinoma, experimentally validated in rat models

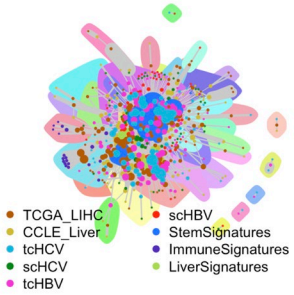
Use Case 1: Studying virus-induced hepatocellular carcinoma

Driver prediction for hepatitis C and B virus-induced hepatocellular carcinoma across subnetworks derived from >6 systems validated in cell lines, and prediction of chemopreventive treatments modulating disease-associated subnetworks using chemical perturbations in cell lines, experimentally validated in rat models

Driver and drug discovery for hepatitis C (HCV) and hepatitis B (HBV) virus-induced hepatocellular carcinoma (HCC)

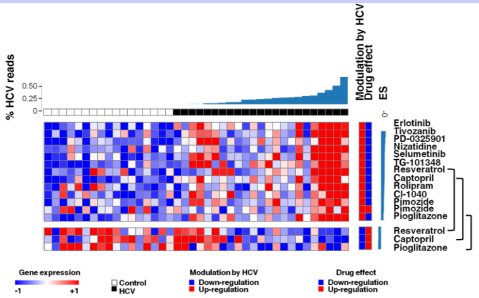


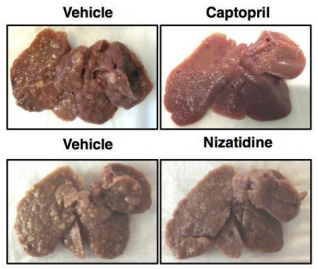
Pan-etiology of cancer communities or subnetworks



Drug discovery:
Chemical perturbations in cell lines
Predict which drug compounds can reverse disease-associated circuits
Alternative treatments with less severe adverse effects?

Drug validation:
Experimental validation of drugs in rat models
⇒ Two novel compounds attenuate HCC development
⇒ Safe and low-cost approach for chemoprevention of HCC?

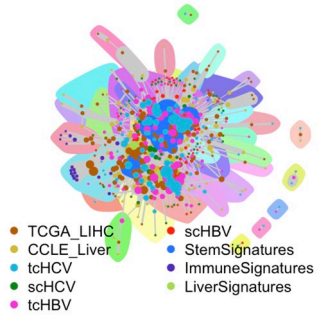


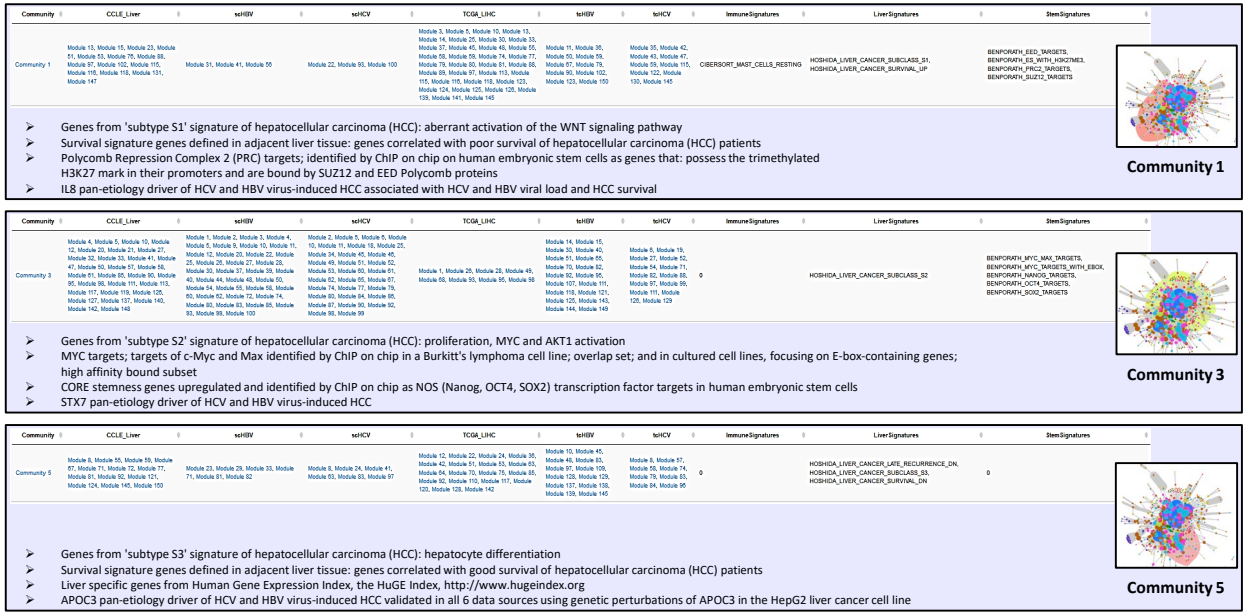


Nathalie Pochet and Thomas Baumert, submitted

Driver and drug discovery for chemoprevention of hepatitis C (HCV) and hepatitis B (HBV) virus-induced hepatocellular carcinoma (HCC)
⇒ **AMARETTO facilitates identification of known and novel drug compounds and how they modulate cancer drivers and their targets**

Community-AMARETTO





Community-AMARETTO report: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/index.html
NDEx network visualization: http://www.ndexbio.org/#/network/f50b3ecb-7b47-11e9-848d-0ac135e8bacf

Use Case 1: Studying virus-induced hepatocellular carcinoma

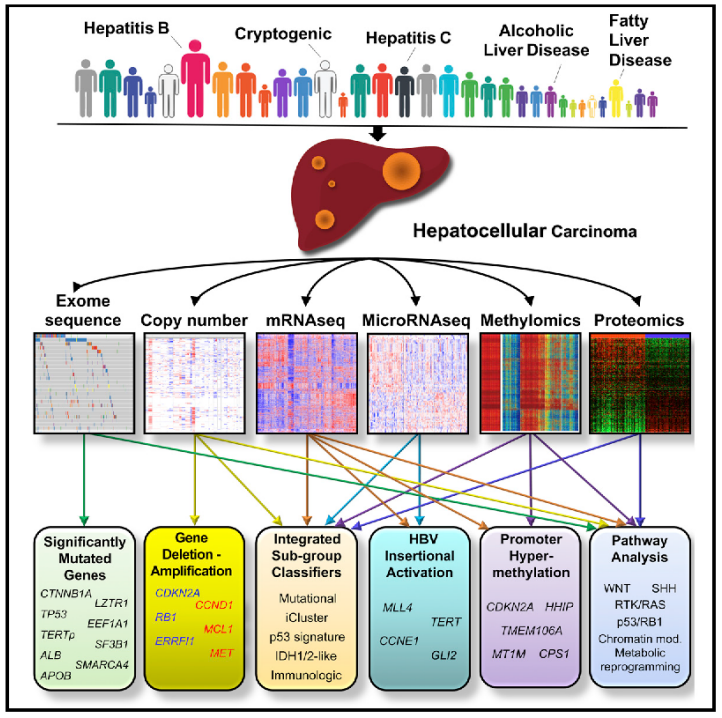
Driver prediction for hepatitis C and B virus-induced hepatocellular carcinoma across subnetworks derived from >6 systems validated in cell lines, and prediction of chemopreventive treatments modulating disease-associated subnetworks using chemical perturbations in cell lines, experimentally validated in rat models

Resource

Cell

Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma

Graphical Abstract



Authors

The Cancer Genome Atlas Research Network

Correspondence

wheeler@bcm.edu (David A. Wheeler), roberts.lewis@mayo.edu (Lewis R. Roberts)

In Brief

Multiplex molecular profiling of human hepatocellular carcinoma patients provides insight into subtype characteristics and points toward key pathways to target therapeutically.

Highlights

- Analysis of hepatocellular carcinomas integrates data of

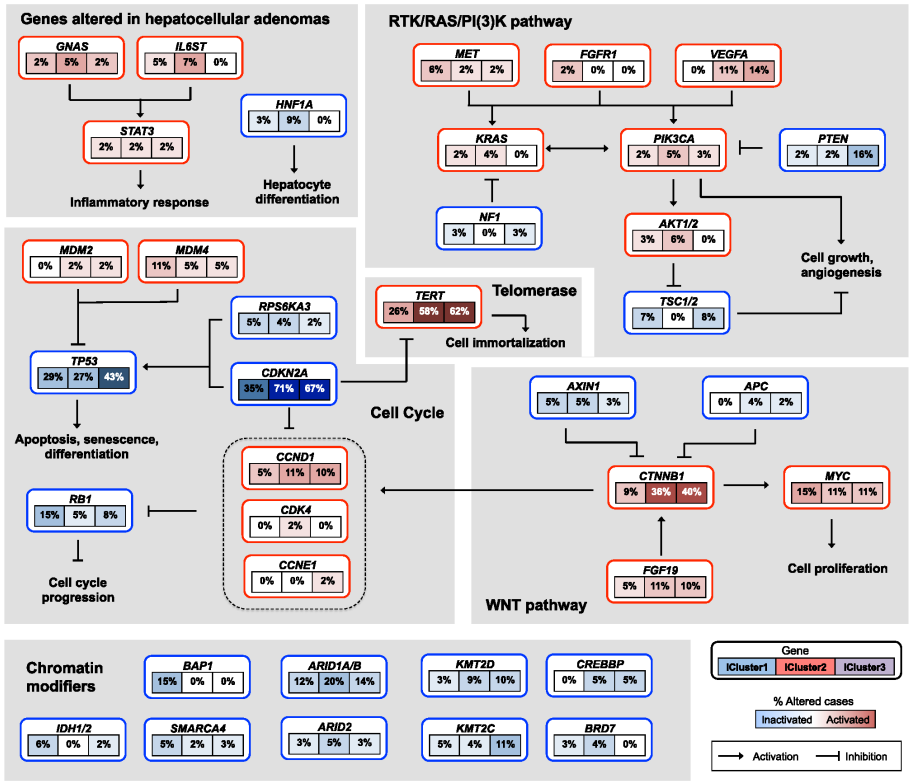


Figure 6. Integrated Molecular Comparison of Somatic Alterations in Signaling Pathways across iCluster Groups

Each gene box includes three percentages representing the frequency of activation or inactivation in iClusters 1, 2, and 3 based on the core 196 sample HCC dataset. All somatic changes are tallied together in calculating the percentages of altered cases within each of the iCluster sample groups. Somatic alterations include mutations and copy-number changes (homozygous deletion and high-level amplifications), as well as epigenetic silencing of *CDKN2A*. Missense mutations are only counted if they have known oncogenic function, have been reported in COSMIC, or occur at known mutational hotspots. Genes are grouped by signaling pathways, with edges showing pairwise molecular interactions.

See also Figure S6.

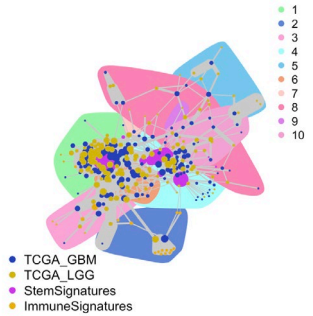
Use Cases: integrating multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer

1. A study of hepatitis C and B virus-induced hepatocellular carcinoma (LIHC) with driver and drug discovery for chemoprevention across pan-etiologies of hepatocellular carcinoma, experimentally validated in rat models

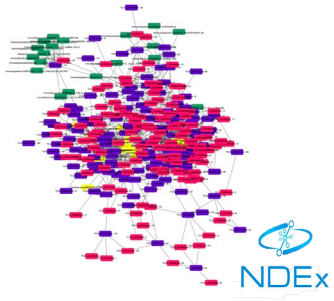
2. A study of glioblastoma multiforme (GBM) and low-grade glioma (LGG) with driver discovery for diagnostic and prognostic molecular subclasses associated with imaging-derived features for non-invasive imaging diagnostics

Use Case 2: Studying multi-omics and imaging of gliomas

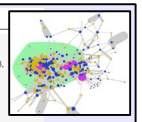
Driver prediction for multi-omics subnetworks associated with imaging-derived features representing prognostic molecular subclasses of gliomas and glioblastoma multiforme



Community-AMARETTO
TCGA GBM vs LGG



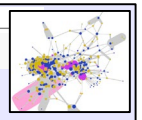
Community	TCGA_GBM	TCGA_LGG	ImmuneSignatures	StemSignatures
Community 1	Module 3, Module 6, Module 18, Module 22, Module 23, Module 24, Module 25, Module 27, Module 33, Module 37, Module 38, Module 39, Module 41, Module 43, Module 48, Module 54, Module 60, Module 62, Module 66, Module 67, Module 68, Module 71, Module 72, Module 75, Module 78, Module 80, Module 81, Module 82, Module 85, Module 86, Module 87, Module 88, Module 90, Module 91, Module 97, Module 98, Module 99, Module 102, Module 107, Module 110, Module 114, Module 116, Module 117, Module 120, Module 122, Module 129, Module 132, Module 137, Module 138, Module 142, Module 147, Module 150	Module 2, Module 7, Module 8, Module 16, Module 18, Module 19, Module 23, Module 26, Module 27, Module 30, Module 34, Module 35, Module 37, Module 40, Module 43, Module 44, Module 47, Module 50, Module 51, Module 53, Module 55, Module 59, Module 60, Module 61, Module 63, Module 64, Module 66, Module 69, Module 72, Module 79, Module 80, Module 81, Module 83, Module 84, Module 85, Module 86, Module 87, Module 88, Module 89, Module 93, Module 99, Module 102, Module 103, Module 104, Module 110, Module 112, Module 117, Module 120, Module 123, Module 127, Module 133, Module 134, Module 135, Module 137, Module 138, Module 141, Module 142, Module 147, Module 150	CIBERSORT_MAST_CELLS_RESTING	BENPORATH_EED_TARGETS, BENPORATH_ES_WITH_H3K27ME3, BENPORATH_NOS_TARGETS, BENPORATH_OCT4_TARGETS, BENPORATH_PRC2_TARGETS, BENPORATH_SUZ12_TARGETS



Community 1

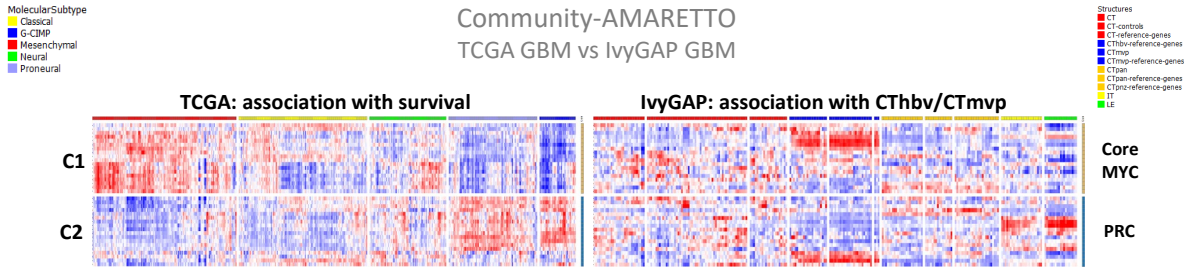
- VERHAAK GLIOBLASTOMA PRONEURAL: Genes correlated with proneural type of glioblastoma multiforme tumors
- NOUSHMEHR GBM SILENCED BY METHYLATION: Top 50 most differentially hypermethylated and down-regulated genes in proneural G-CIMP (a CpG island methylator phenotype) GBM (glioblastoma multiforme) tumors
- Molecular markers: PDGFRA, IDH1 and MGMT
- Methylation-driven drivers shared between GBM (Module 90) and LGG (Module 77): FBXO17, XKR8, RAB34/36
- Survival analysis: hypermethylated drivers, repressed expression of drivers and targets, and association with better survival rates in the G-CIMP and proneural molecular subclasses

Community	TCGA_GBM	TCGA_LGG	ImmuneSignatures	StemSignatures
Community 3	Module 8, Module 55, Module 70, Module 77, Module 93, Module 125, Module 131, Module 133, Module 134, Module 141	Module 22, Module 33, Module 38, Module 73, Module 90, Module 118, Module 121	CIBERSORT_EOSINOPHILS, CIBERSORT_MACROPHAGES_M0, CIBERSORT_MACROPHAGES_M2, CIBERSORT_MONOCYTES, CIBERSORT_NEUTROPHILS	0



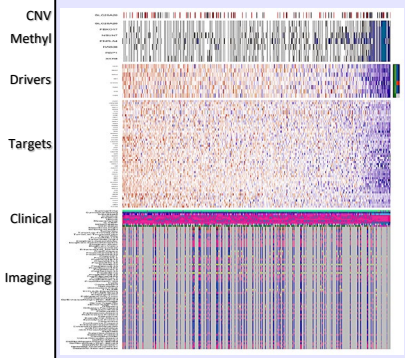
Community 3

- VERHAAK GLIOBLASTOMA MESENCHYMAL: Genes correlated with mesenchymal type of glioblastoma multiforme tumors
- Enriched for immune cell type-specific signatures
- Associated with imaging feature Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)
- Survival analysis: induced expression of drivers and targets and association with poorer survival rates in the mesenchymal molecular subclass



Community-AMARETTO
TCGA GBM vs IvyGAP GBM

Community-AMARETTO report: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/index.html
NDEx network visualization: http://www.ndexbio.org/#/network/c1f0fccf-80b6-11e9-848d-0ac135e8bacf



GBM Module 90

Functional Enrichment

NOUSHMEHR GBM SILENCED BY METHYLATION
Top 50 most differentially hypermethylated and down-regulated genes in proneural G-CIMP (a CpG island methylator phenotype) GBM (glyoblastoma multiforme) tumors

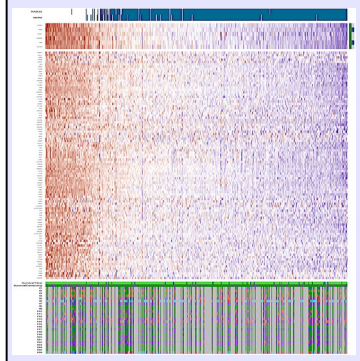
Phenotype Association

Survival Analysis
lower expression, better survival

Molecular Subclass
G-CIMP

Molecular Markers
IDH1 & MGMT

Community 1



LGG Module 150

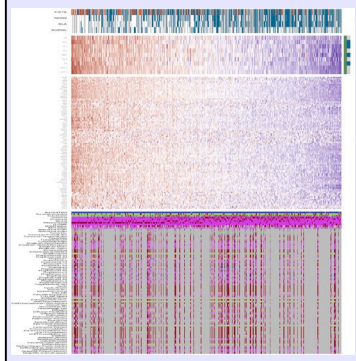
Functional Enrichment

Phenotype Association

Survival Analysis
lower expression, better survival

Imaging
- Proportion nCET (f6)
lower expression, higher proportion of non-enhancing tumor (not edema)
- Proportion Enhancing (f5)
higher expression, higher proportion of enhancing tumor

Community 1



GBM Module 77

Functional Enrichment

VERHAAK GLIOBLASTOMA MESENCHYMAL
Genes correlated with mesenchymal type of glioblastoma multiforme tumors

Phenotype Association

Survival Analysis
higher expression, poorer survival

Molecular Subclass
Mesenchymal

Molecular Markers
IDH1

Imaging
- Proportion nCET (f6)
lower expression, higher proportion of non-enhancing tumor (not edema)

Community 3

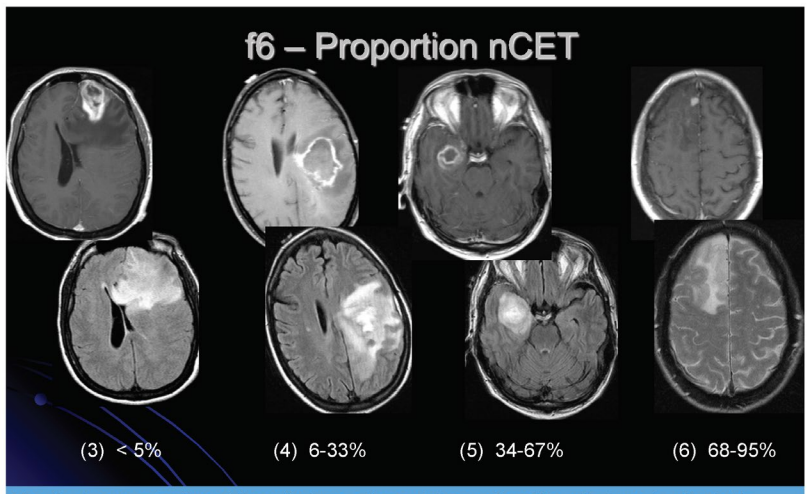
Use Case 2: Studying multi-omics and imaging of gliomas

Driver prediction for multi-omics subnetworks associated with imaging-derived features representing prognostic molecular subclasses of gliomas and glioblastoma multiforme

Vasari MRI Visual Feature Guide

Sample images derived from the Rembrandt MRI examinations contributed by Thomas Jefferson University Hospital

Rev 1.1



Visually, when scanning through the entire tumor volume, what proportion of the entire tumor is estimated to represent non-enhancing tumor (not edema)? Non-enhancing tumor is defined as regions of T2W hyperintensity (less than the intensity of cerebrospinal fluid, with corresponding T1W hypointensity) that are associated with mass effect and architectural distortion, including blurring of the gray-white interface.(Assuming that the the entire abnormality may be comprised of: (1) an enhancing component, (2) a non-enhancing component, (3) a necrotic component and (4) a edema component.)

The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 JUNE 25, 2015 VOL. 372 NO. 26

Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas

The Cancer Genome Atlas Research Network*

ABSTRACT

Cell
Vol 455|23 October 2008|doi:10.1038/nature07385
Resource

The Somatic Genomic Landscape of Glioblastoma

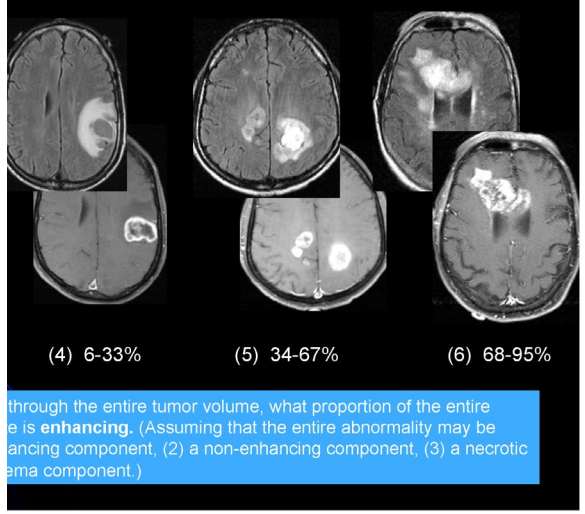
Cameron W. Brennan,[1,2,40,*] Roel G.W. Verhaak,[3,11,40] Aaron McKenna,[4,40] Benito Campos,[5,6] Houtan Noushmehr,[7,8] Sofie R. Salama,[9] Siyuan Zheng,[3] Debyani Chakravarty,[1] J. Zachary Sanborn,[9] Samuel H. Berman,[1] Rameen Beroukhim,[4,5] Brady Bernard,[10] Chang-Jiun Wu,[11] Giannicola Genovese,[11] Ilya Shmulevich,[10] Jill Barnholtz-Sloan,[12] Lihua Zou,[4] Rahulsimham Vegesna,[3] Sachet A. Shukla,[5] Giovanni Ciriello,[13] W.K. Yung,[14] Wei Zhang,[15] Carrie Sougnez,[4] Tom Mikkelsen,[16] Kenneth Aldape,[15] Darell D. Bigner,[17] Erwin G. Van Meir,[18] Michael Prados,[19] Andrew Sloan,[20] Keith L. Black,[21] Jennifer Eschbacher,[22] Gaetano Finocchiaro,[23] William Friedman,[24] David W. Andrews,[25] Abhijit Guha,[26] Mary Iacocca,[27] Brian P. O'Neill,[28] Greg Foltz,[29] Jerome Myers,[30] Daniel J. Weisenberger,[7] Robert Penny,[31] Raju Kucherlapati,[32] Charles M. Perou,[33] D. Neil Hayes,[33] Richard Gibbs,[34] Marco Marra,[35] Gordon B. Mills,[36] Eric Lander,[4] Paul Spellman,[37] Richard Wilson,[38] Chris Sander,[13] John Weinstein,[3] Matthew Meyerson,[4,5] Stacey Gabriel,[4] Peter W. Laird,[7] David Haussler,[9,39] Gad Getz,[4] Lynda Chin,[4,11,*] and TCGA Research Network
[1]Human Oncology and Pathogenesis Program, Brain Tumor Center, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA
[2]Department of Neurosurgery, Memorial Sloan-Kettering Cancer Center, Department of Neurological Surgery, Weill Cornell Medical Center, New York, NY 10065, USA
[3]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[4]Cancer Program, The Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[5]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[6]Division of Experimental Neurosurgery, Department of Neurosurgery, Heidelberg University Hospital, 69120 Heidelberg, Germany
[7]University of Southern California Epigenome Center, University of Southern California, Keck School of Medicine, Los Angeles, CA 90033, USA
[8]Department of Genetics, Center for Integrative System Biology, Faculty of Medicine at Ribeirão Preto, University of São Paulo, 14049-900 Ribeirão Preto, São Paulo, Brazil

Comprehensive genomi defines human glioblas core pathways

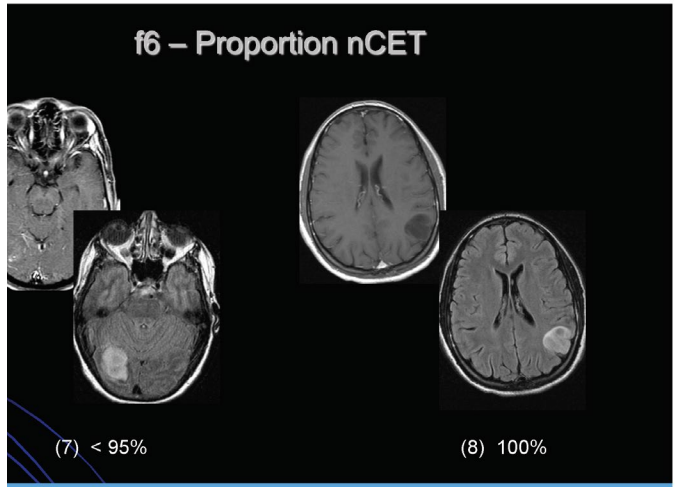
s Research Network*

y harbour multiple chromosomal aberra ignant transformation. The Cancer Genc onal analysis of these molecular charact Here we report the interim integrative a 206 glioblastomas—the most common t of the 206 glioblastomas. This analysis

f5 – Proportion Enhancing



through the entire tumor volume, what proportion of the entire e is **enhancing.** (Assuming that the entire abnormality may be ancing component, (2) a non-enhancing component, (3) a necrotic ema component.)



nning through the entire tumor volume, what proportion of the entire tumor is estimated to ancing tumor (not edema)? Non-enhancing tumor is defined as regions of T2W hyperintensity (less than the intensity of cerebrospinal fluid, with corresponding T1W hypointensity) that are associated with mass effect and architectural distortion, including blurring of the gray-white interface.(Assuming that the the entire abnormality may be comprised of: (1) an enhancing component, (2) a non-enhancing component, (3) a necrotic component and (4) a edema component.)

y of California Santa Cruz,
/030, USA
OH 44106, USA
30, USA
A
ol of Medicine, Emory University,
and, OH 44106, USA
la
30, USA
239, USA

http://portals.broadinstitute.org/pochetlab/demo/Imaging_Features/Vasari%20MR%20Feature%20Guide_v1.1.pdf

Use Cases: integrating multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer

1. A study of hepatitis C and B virus-induced hepatocellular carcinoma (LIHC) with driver and drug discovery for chemoprevention across pan-etiologies of hepatocellular carcinoma, experimentally validated in rat models

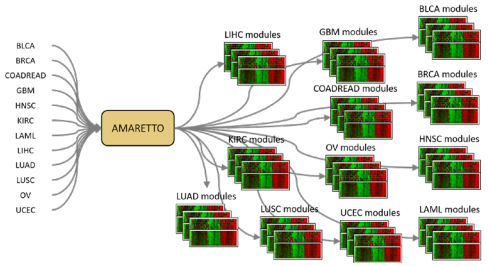
2. A study of glioblastoma multiforme (GBM) and low-grade glioma (LGG) with driver discovery for diagnostic and prognostic molecular subclasses associated with imaging-derived features for non-invasive imaging diagnostics

3. A pan-cancer study across twelve cancer sites with driver discovery of pan-cancer drivers of smoking-induced and 'antiviral' interferon-modulated innate immune response cancer

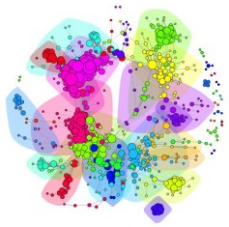
Use Case 3a: Pan-cancer driver discovery

Driver prediction for pan-cancer multi-omics subnetworks across 12 cancer (sub)types validated using genetic perturbations in cell lines

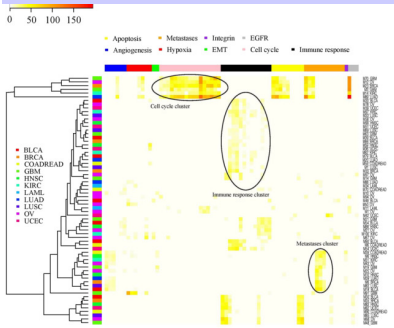
Drivers of smoking-induced cancer and ‘antiviral’ interferon-modulated innate immune response across 12 cancer (sub)types



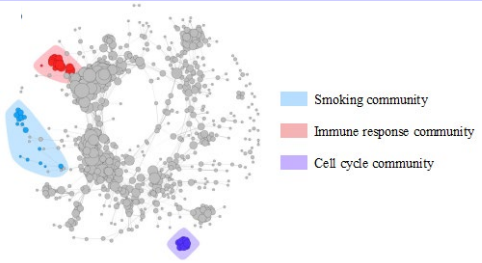
Pan-cancer communities or subnetworks



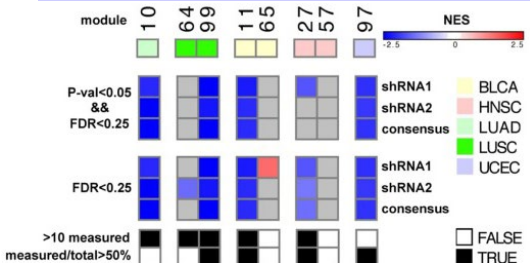
Pan-cancer functional categories
⇒ AMARETTO captures hallmarks of cancer



Driver discovery:
- OAS2 pan-cancer driver of ‘antiviral’ interferon-modulated innate immune response
- GPX2 pan-cancer driver of smoking-induced cancer

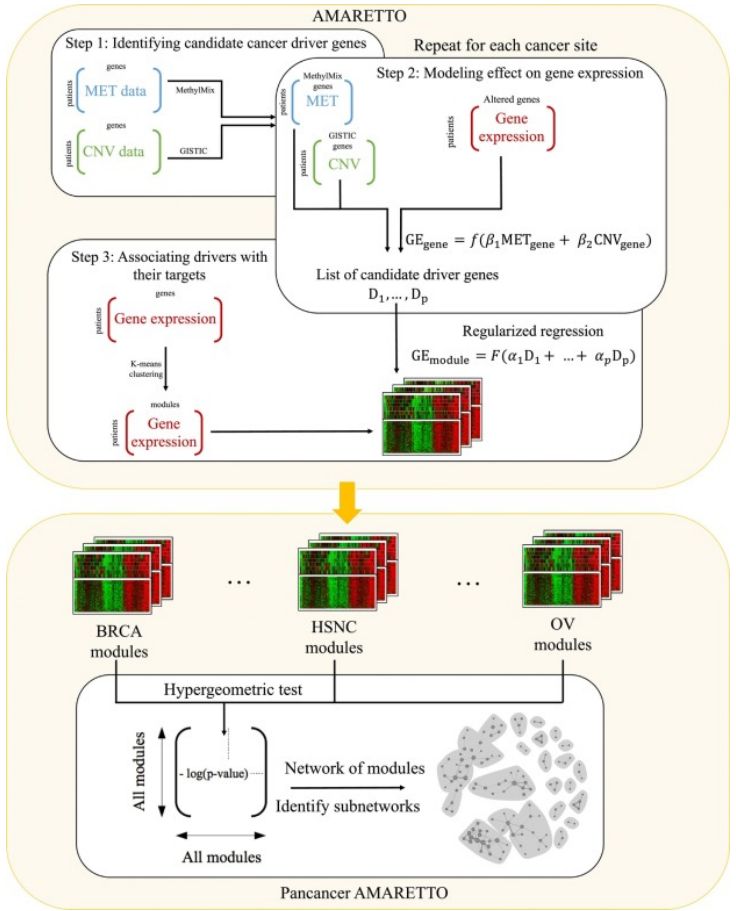


Driver validation:
Genetic perturbations of GPX2 in the A549 (LUAD) cell line
⇒ Knocking down GPX2 represses target genes in GPX2-regulated circuits



Nathalie Pochet and Olivier Gevaert, *EBioMedicine* 2018

Drivers of smoking-induced cancer and ‘antiviral’ interferon-modulated innate immune response across 12 cancer (sub)types (GBM, LIHC)
⇒ **AMARETTO facilitates identification of known and novel cancer drivers and their targets**



Workflow of *AMARETTO:

First, AMARETTO infers regulatory networks within each biological system via multi-omics data fusion. Specifically, AMARETTO identifies potential cancer drivers by identifying genes whose genetic and epigenetic cancer aberrations have a direct functional impact on their own transcriptomic or proteomic expression. AMARETTO then connects these drivers with modules of co-expressed target genes that they putatively control, defined as regulatory circuits, using a penalized regulatory program. Second, Community-AMARETTO learns communities or subnetworks by connecting the regulatory circuits inferred from different systems to identify drivers across diseases or biological systems.

Use Cases: integrating multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer

1. A study of hepatitis C and B virus-induced hepatocellular carcinoma (LIHC) with driver and drug discovery for chemoprevention across pan-etiologies of hepatocellular carcinoma, experimentally validated in rat models

2. A study of glioblastoma multiforme (GBM) and low-grade glioma (LGG) with driver discovery for diagnostic and prognostic molecular subclasses associated with imaging-derived features for non-invasive imaging diagnostics

3. A pan-cancer study across twelve cancer sites with driver discovery of pan-cancer drivers of smoking-induced and 'antiviral' interferon-modulated innate immune response cancer

3. A pan-cancer study of squamous cell carcinoma (SCC) across five SCC cancer sites, in particular, lung (LUSC), head and neck (HNSC), esophageal (ESCA), cervical (CESC) and bladder (BLCA)

Use Case 3b: Pan-squamous cell carcinoma driver discovery

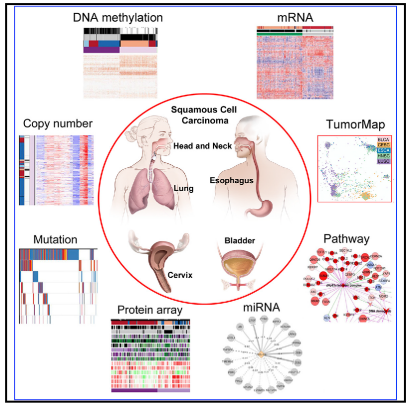
Driver prediction for pan-squamous cell carcinoma multi-omics subnetworks across 5 cancer sites, i.e., in lung (LUSC), head and neck (HNSC), esophageal (ESCA), cervical (CESC) and bladder (BLCA), validated using genetic perturbations in cell lines

Resource

Cell Reports

Genomic, Pathway Network, and Immunologic Features Distinguishing Squamous Carcinomas

Graphical Abstract



Highlights

- SCCs show chromosome or methylation alterations affecting multiple related genes
- These regulate squamous stemness, differentiation, growth, survival, and inflammation
- Copy-quiet SCCs have hypermethylated (*FANCF*, *TET1*) or mutated (*CASP8*, *MAPK-RAS*) genes
- Potential targets include ΔNp63, WEE1, IAPs, PI3K-mTOR/MAPK, and immune responses

Authors

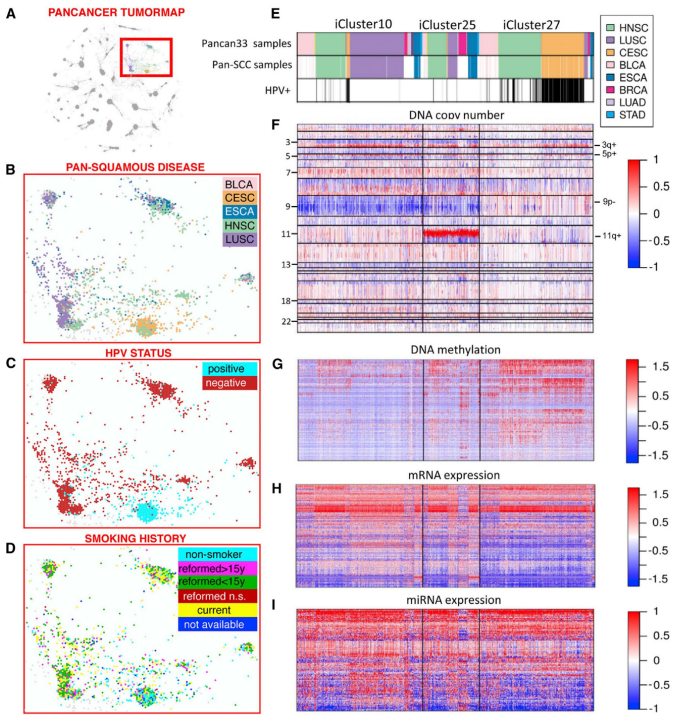
Joshua D. Campbell, Christina Yau, Reanne Bowlby, ..., Curtis R. Pickering, Zhong Chen, Carter Van Waes

Correspondence

chenz@nidcd.nih.gov (Z.C.), vanwaesc@nidcd.nih.gov (C.V.W.)

In Brief

Campbell et al. reveal that squamous cell cancers from different tissue sites may be distinguished from other cancers and subclassified molecularly by recurrent alterations in chromosomes, DNA methylation, messenger and microRNA expression, or by mutations. These affect squamous cell pathways and programs that provide candidates for therapy.



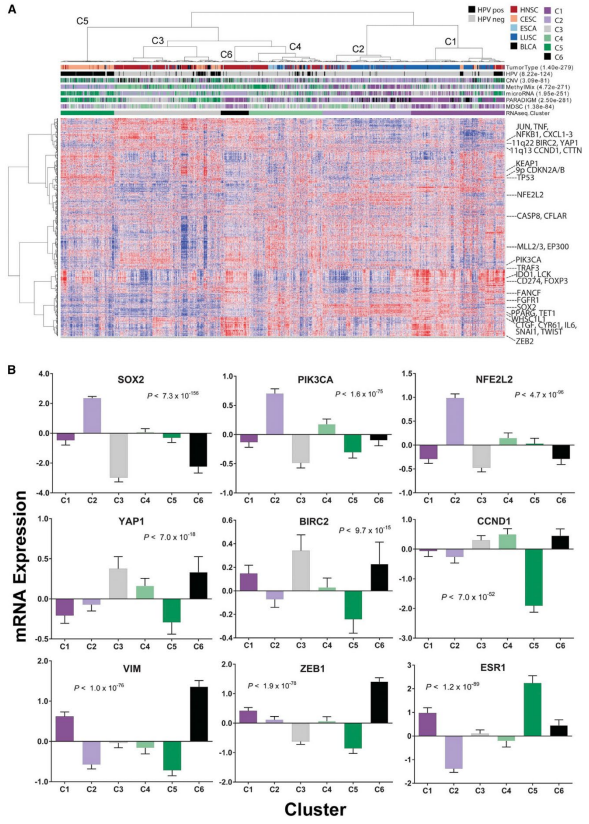


Figure 1. TumorMap and iCluster of Squamous Cancers from PanCancer-33 Analysis

(A) TumorMap analysis visualizing close mapping of LUSC, HNSC, ESCA, CESC, and BLCA among 28 PanCancer-33 islands.
(B) Higher resolution view of TM islands and distribution of SCC from 5 sites.
(C) HPV status showing the majority of HPV(+) CESC and HNSC map around a distinct island.
(D) Smoking history of SCC. Each spot in the map represents a sample. The colors of the sample spots represent attributes as described for each panel.
(E–I) Summary of iCluster analysis (E), DNA copy-number (F), methylation (G), mRNA (H), and miRNA (I) expression. PanCancer-33 SCC and other tumors and Pan-SCC from 5 sites identified by histopathologic diagnosis cluster within iC10, iC25, and iC27. Annotation bars show cancer type and HPV status, and keys show an increase (red) or decrease (blue) in features as indicated: DNA copy number, copy-number log ratio (tumor versus normal); DNA methylation, normalized beta values; miRNA expression, normalized log expression counts; miRNA expression, normalized log expression counts.

(legend on next page)

Campbell et al., 2018, Cell Reports 23, 194–212
April 3, 2018
https://doi.org/10.1016/j.celrep.2018.03.063

AMARETTO reports for case studies

Case Study 1 (virus-induced LIHC):

- TCGA LIHC: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html
- CCLE liver: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/CCLE_Liver/AMARETTOhtmls/index.html
- Time-course HCV: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/tcHCV/AMARETTOhtmls/index.html
- Single-cell HCV: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/scHCV/AMARETTOhtmls/index.html
- Time-course HBV: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/tcHBV/AMARETTOhtmls/index.html
- Single-cell HBV: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/scHBV/AMARETTOhtmls/index.html

Case Study 2 (gliomas GBM and LGG):

- TCGA GBM: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_GBM/AMARETTOhtmls/index.html
- TCGA LGG: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_LGG/AMARETTOhtmls/index.html

Case Study 3 (pan-squamous BLCA, CESC, ESCA, HNSC, LUSC):

- TCGA BLCA: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_BLCA/AMARETTOhtmls/index.html
- TCGA CESC: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_CESC/AMARETTOhtmls/index.html
- TCGA ESCA: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_ESCA/AMARETTOhtmls/index.html
- TCGA HNSC: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_HNSC/AMARETTOhtmls/index.html
- TCGA LUSC: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_LUSC/AMARETTOhtmls/index.html

Community-AMARETTO reports for case studies

Case Study 1 (virus-induced LIHC):

- TCGA LIHC & CCLE liver & Time-course HCV & Single-cell HCV & Time-course HBV & Single-cell HBV: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/index.html

Case Study 2 (gliomas GBM and LGG):

- TCGA GBM & LGG: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/index.html

Case Study 3 (pan-squamous BLCA, CESC, ESCA, HNSC, LUSC):

- TCGA BLCA & CESC & ESCA & HNSC & LUSC: http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index.html

Perturbation-AMARETTO reports for case studies

Case Study 1 (virus-induced LIHC):

- Driver discovery: https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers
- Drug discovery: https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drugs_Diseases

Case Study 2 (gliomas GBM and LGG):

- Driver discovery: https://pochetlab.shinyapps.io/pAMARETTO_Brain_2DS_Drivers

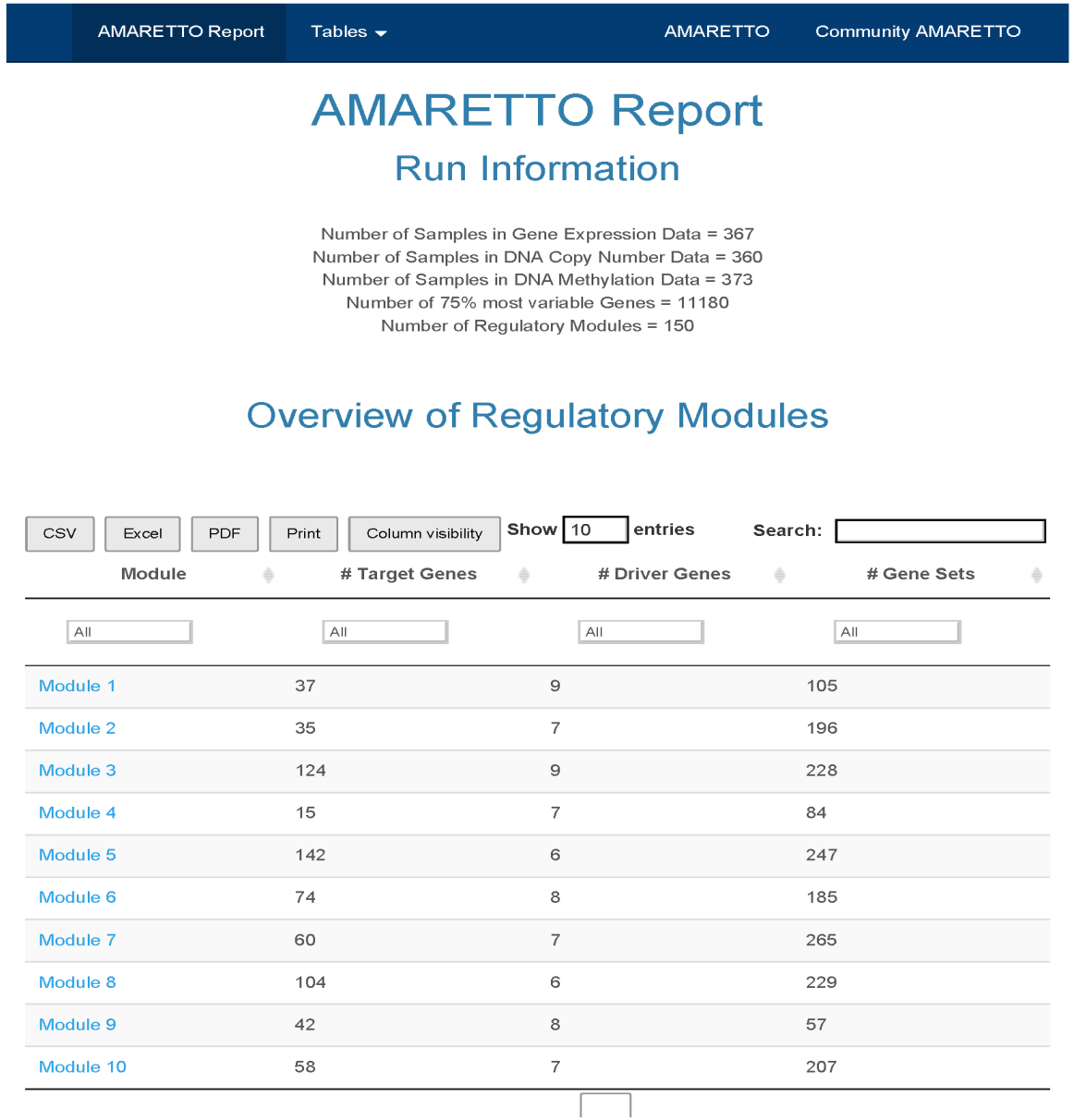
Case Study 3 (pan-squamous BLCA, CESC, ESCA, HNSC, LUSC):

- Driver discovery: https://pochetlab.shinyapps.io/pAMARETTO_AMARETTO_PanCancer_5DS_Drivers

Case Study 1

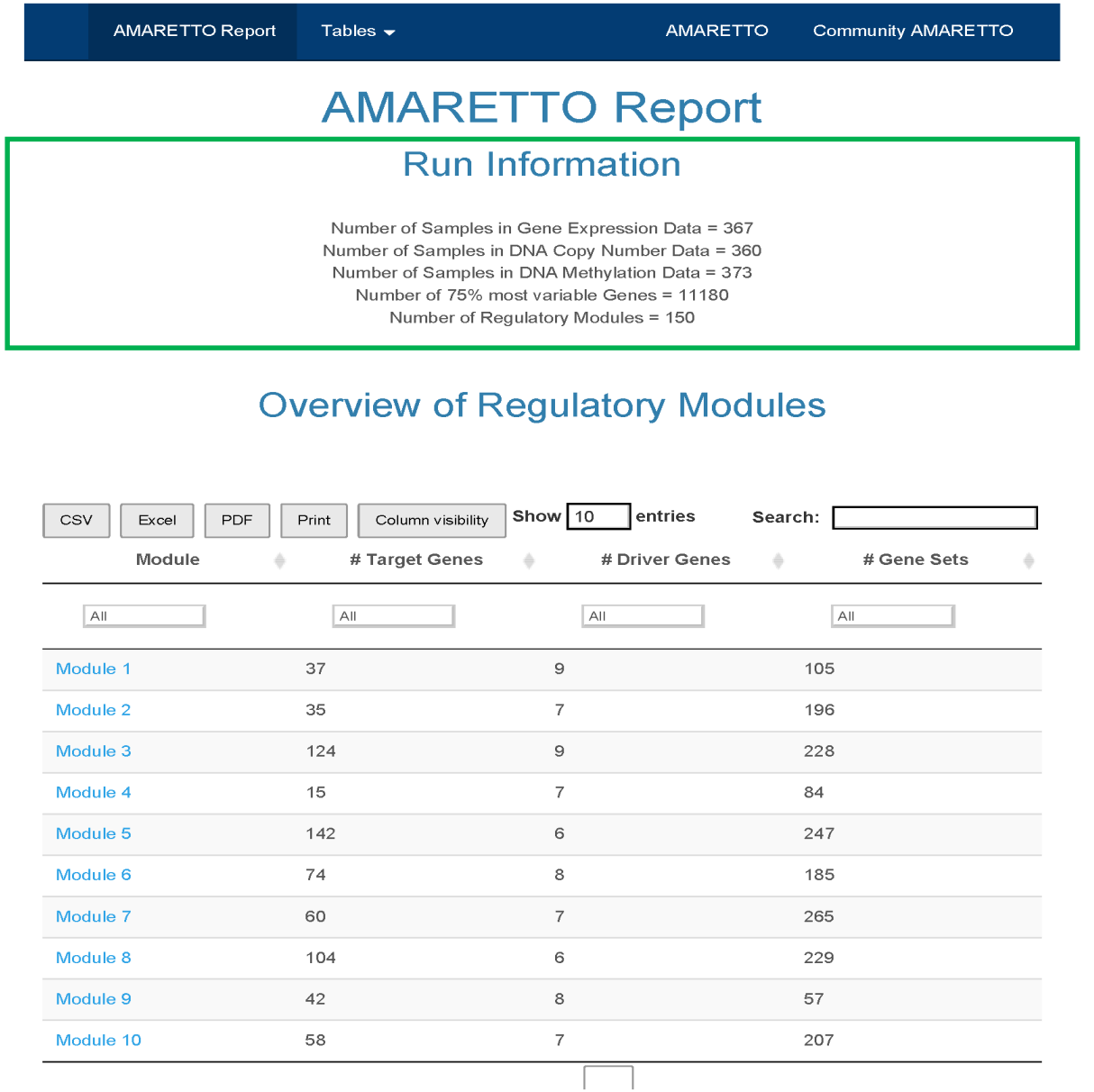
Hepatitis C and B virus-induced Hepatocellular Carcinoma (LIHC)

AMARETTO report LIHC



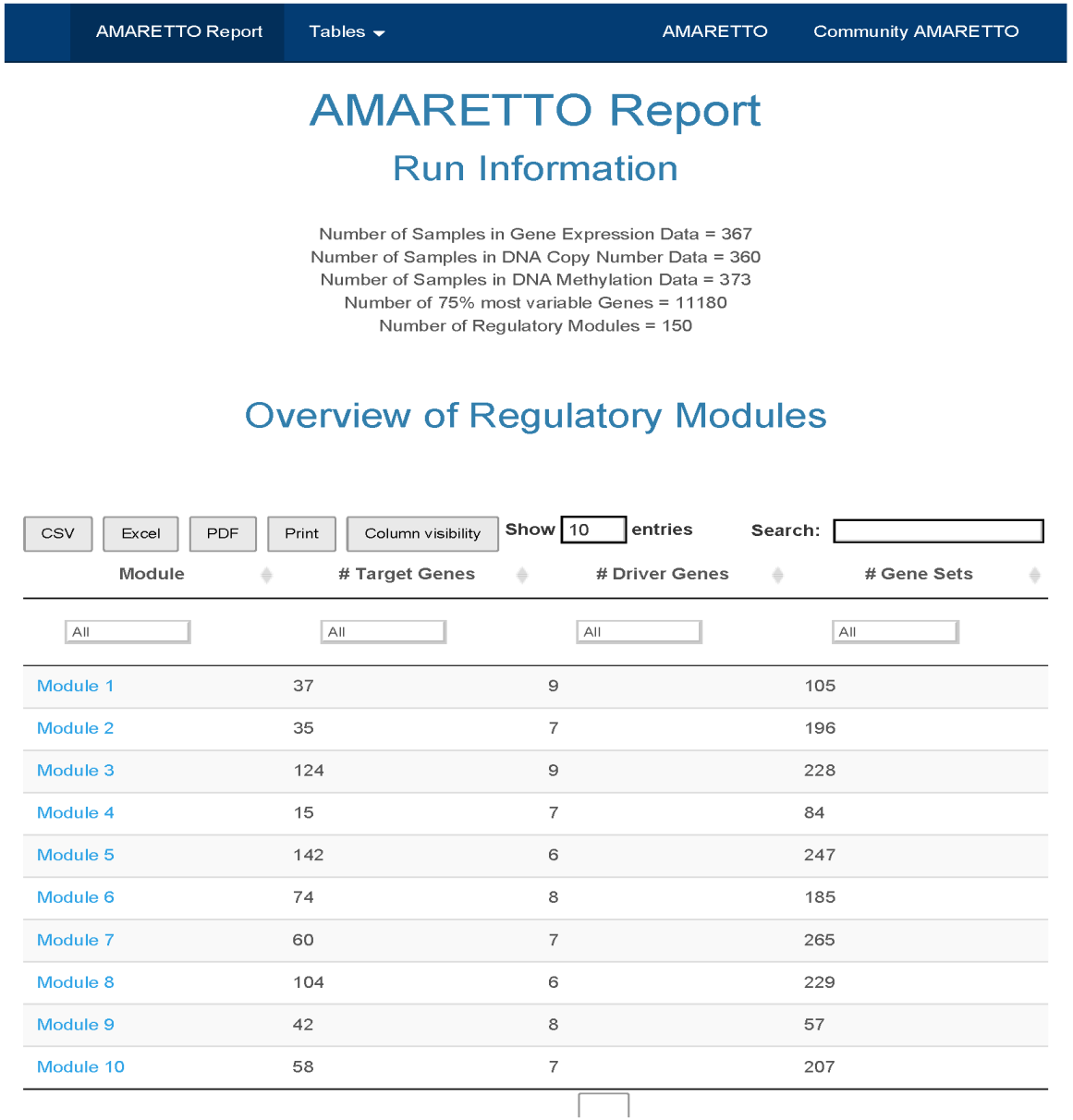
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

AMARETTO report LIHC



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

AMARETTO report LIHC



Module	# Target Genes	# Driver Genes	# Gene Sets
All	All	All	All
Module 1	37	9	105
Module 2	35	7	196
Module 3	124	9	228
Module 4	15	7	84
Module 5	142	6	247
Module 6	74	8	185
Module 7	60	7	265
Module 8	104	6	229
Module 9	42	8	57
Module 10	58	7	207

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

AMARETTO report LIHC

Overview of Regulatory Modules

AMARETTO Report | Tables | AMARETTO | Community AMARETTO

AMARETTO Report
Overview of Regulatory Modules

CSV | Excel | PDF | Print | Column visibility | Show 200 entries | Search:

Module	# Target Genes	# Driver Genes	# Gene Sets
	All	All	All	All
Module 1	37	9	105
Module 2	35	7	196
Module 3	124	9	228
Module 4	15	7	84
Module 5	142	6	247
Module 6	74	8	185
Module 7	60	7	265
Module 8	104	6	229
Module 9	42	8	57
Module 10	58	7	207
Module 11	61	7	170
Module 12	83	6	200
Module 13	82	8	227
Module 14	117	10	66
Module 15	75	7	404
Module 16	25	9	106
...			
Module 147	73	6	198
Module 148	49	6	62
Module 149	103	7	300
Module 150	107	5	255

Showing 1 to 150 of 150 entries | Previous | 1 | Next

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

AMARETTO report LIHC

All tables: functionalities for querying, saving and viewing results

AMARETTO Report | Tables | AMARETTO | Community AMARETTO

AMARETTO Report
Overview of Regulatory Modules

CSV | Excel | PDF | Print | Column visibility | Show 200 entries | Search:

Module	# Target Genes	# Driver Genes	# Gene Sets
All	All	All	All
Module 1	37	9	105
Module 2	35	7	196
Module 3	124	9	228
Module 4	15	7	84
Module 5	142	6	247
Module 6	74	8	185
Module 7	60	7	265
Module 8	104	6	229
Module 9	42	8	57
Module 10	58	7	207
Module 11	61	7	170
Module 12	83	6	200
Module 13	82	8	227
Module 14	117	10	66
Module 15	75	7	404
Module 16	25	9	106
...			
Module 147	73	6	198
Module 148	49	6	62
Module 149	103	7	300
Module 150	107	5	255

Showing 1 to 150 of 150 entries | Previous | 1 | Next

Overview of Regulatory Modules

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

AMARETTO report LIHC

Assignments of Genes to Regulatory Modules

AMARETTO Report | Tables | AMARETTO | Community AMARETTO

AMARETTO Report

Assignments of Genes to Regulatory Modules

CSV | Excel | PDF | Print | Column visibility | Show 100 entries | Search:

Gene	Module	Gene Type
A1BG	Module 53	Target
A1CF	Module 64	Target
A2LD1	Module 64	Target
A2M	Module 81	Target
A4GALT	Module 123	Target
AACS	Module 104	Target
AADAC	Module 22	Target
AADAT	Module 70	Target
AAK1	Module 89	Target
AARS	Module 145	Target
AARSD1	Module 94	Target
AASS	Module 85	Target
AASS	Module 85	Driver
AATK	Module 59	Target
ABAT	Module 70	Target
ABCA1	Module 101	Target
...		
ZXDB	Module 55	Target
ZYG11A	Module 47	Target
ZYG11B	Module 70	Target
ZYX	Module 93	Target
ZZEF1	Module 134	Target
ZZEF1	Module 134	Driver

Showing 12,101 to 12,183 of 12,183 entries

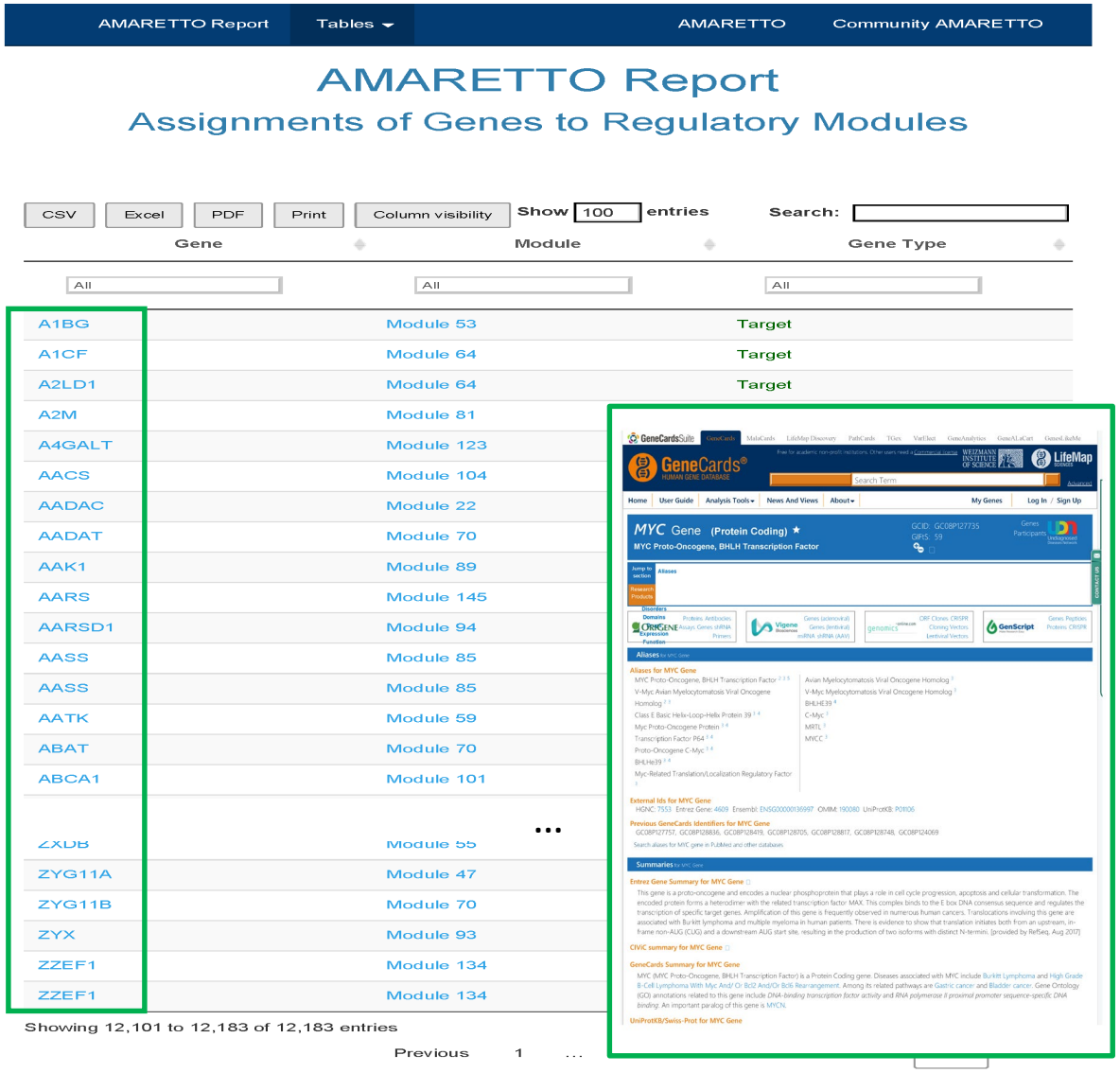
Previous 1 ... 118 119 120 121 122 Next

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

AMARETTO report LIHC

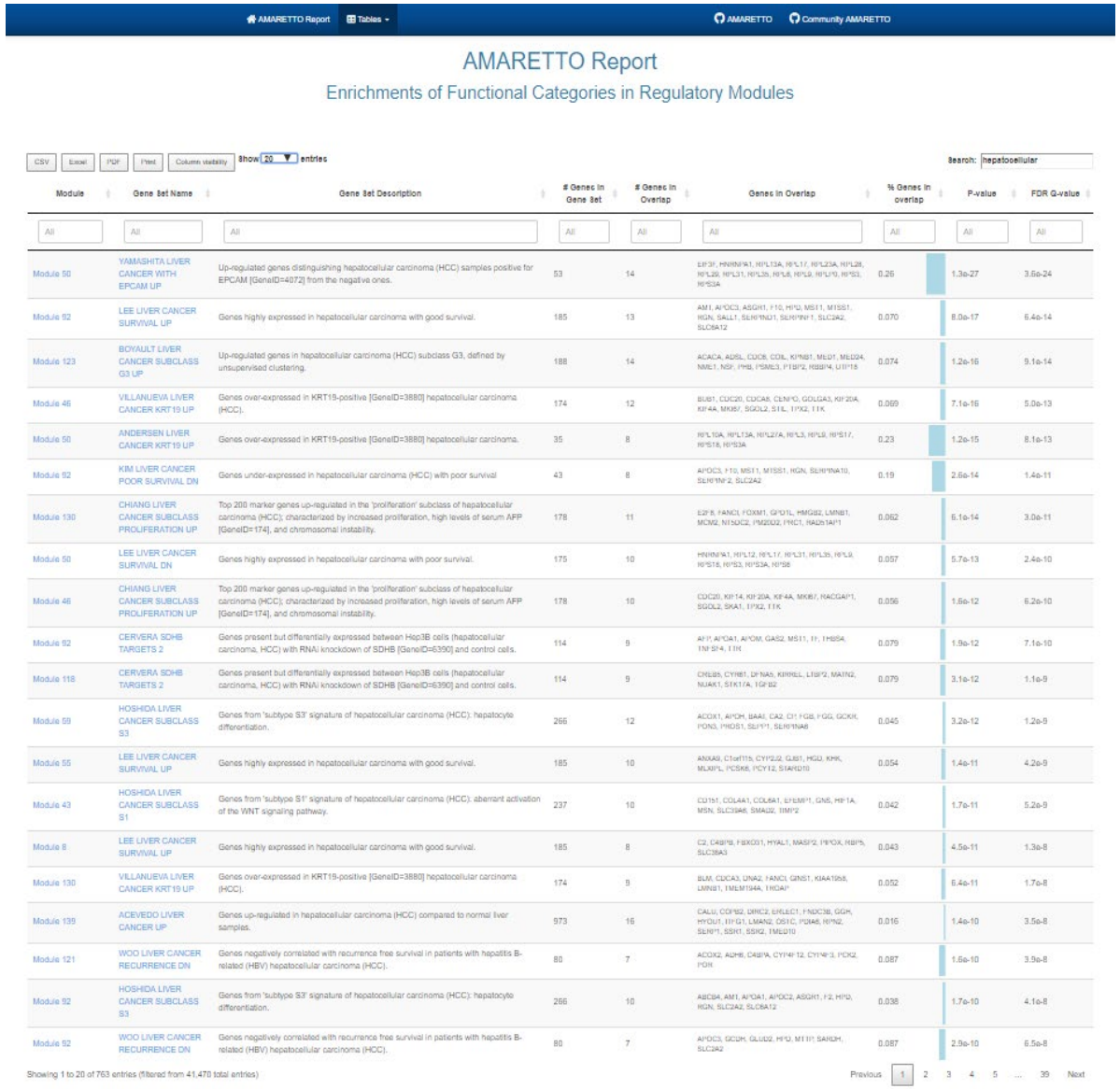
Assignments of Genes to Regulatory Modules

Links to html pages with gene descriptions from GeneCards



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

AMARETTO report LIHC

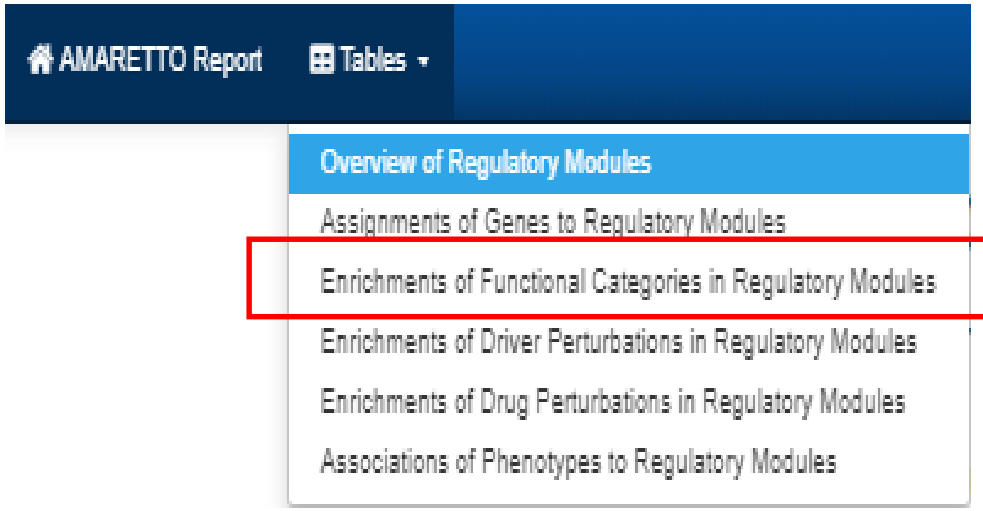


Enrichments of Functional Categories in Regulatory Modules

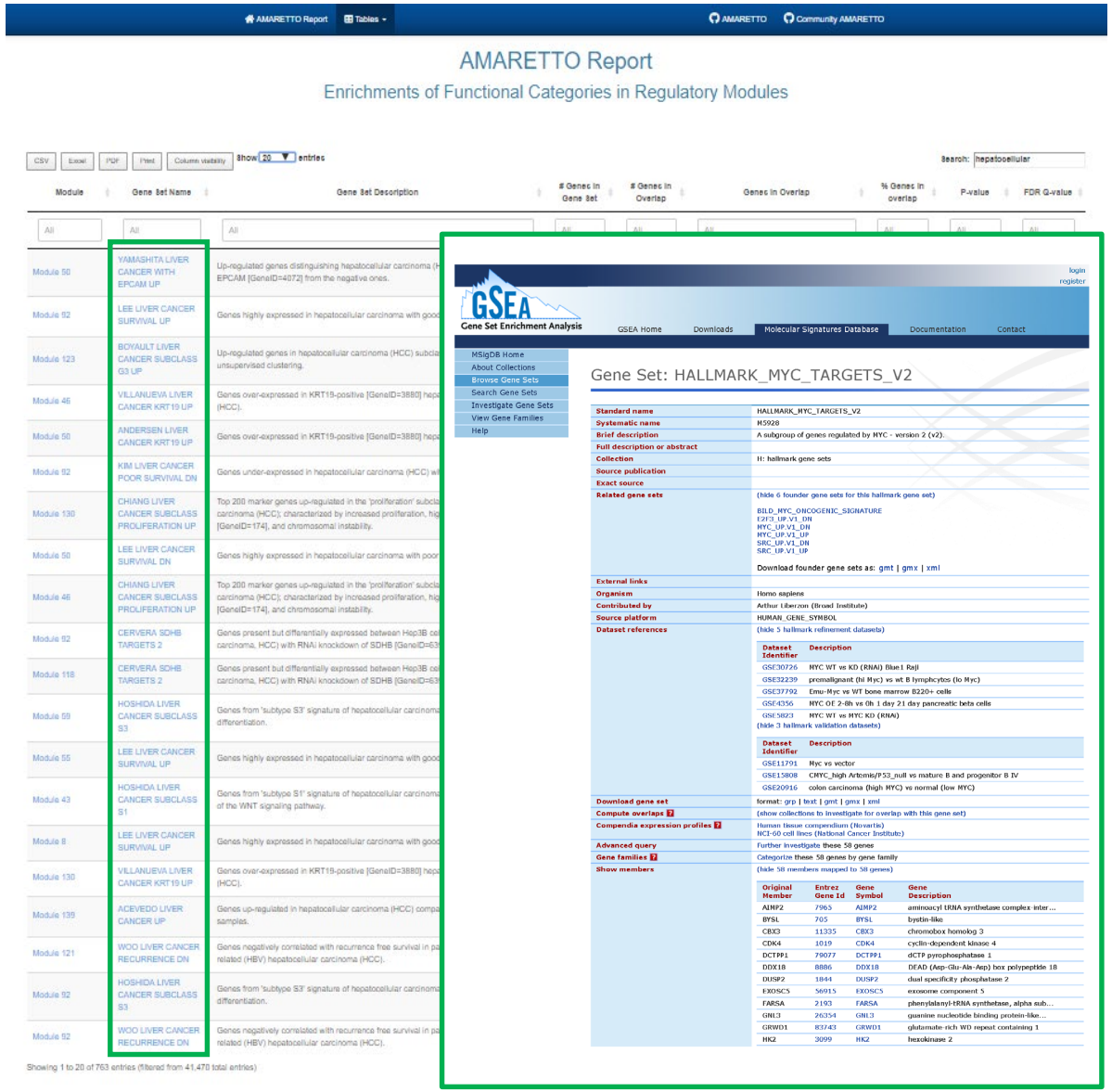
- Functional characterization

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment.html

AMARETTO report LIHC



Links to html descriptions of gene signatures from MSigDB (H+C2)

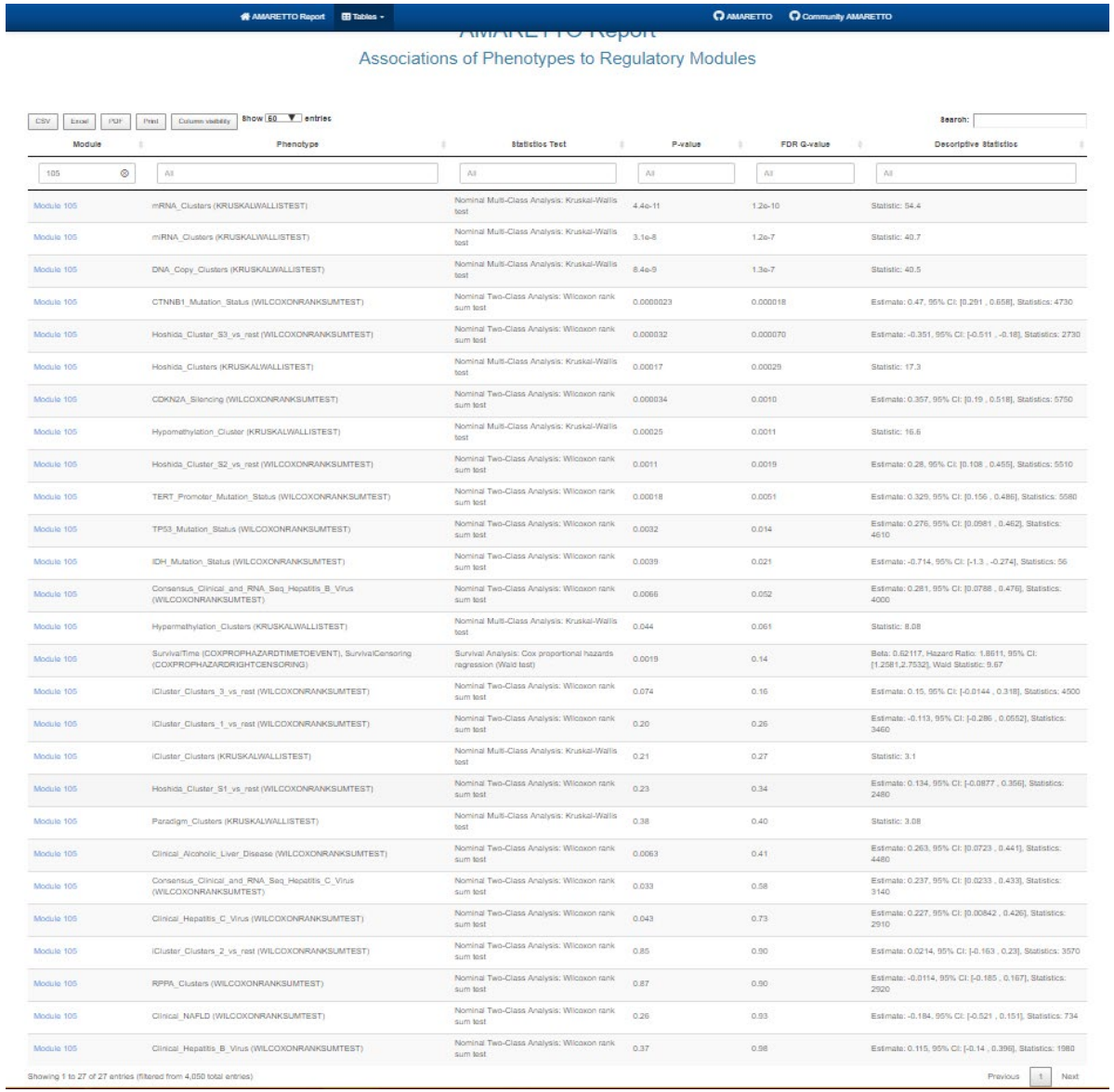


Enrichments of Functional Categories in Regulatory Modules

- Functional characterization

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment.html

AMARETTO report LIHC



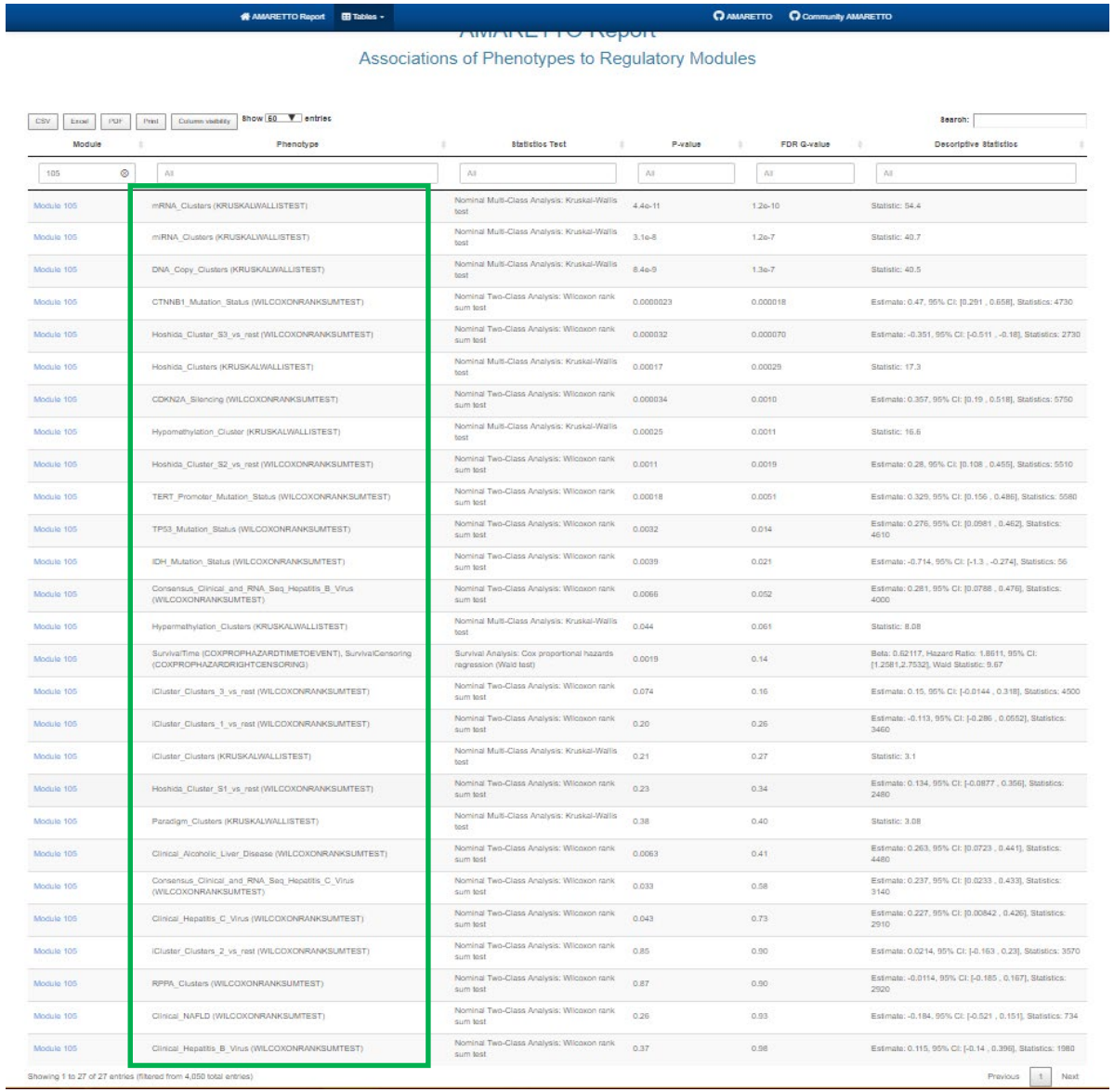
Associations of Phenotypes to Regulatory Modules

- Clinical characterization for clinical, molecular and imaging-derived phenotypes

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_PhenoAssociation.html

AMARETTO report LIHC

Clinical, molecular & imaging-derived phenotypes from TCGA/TCIA

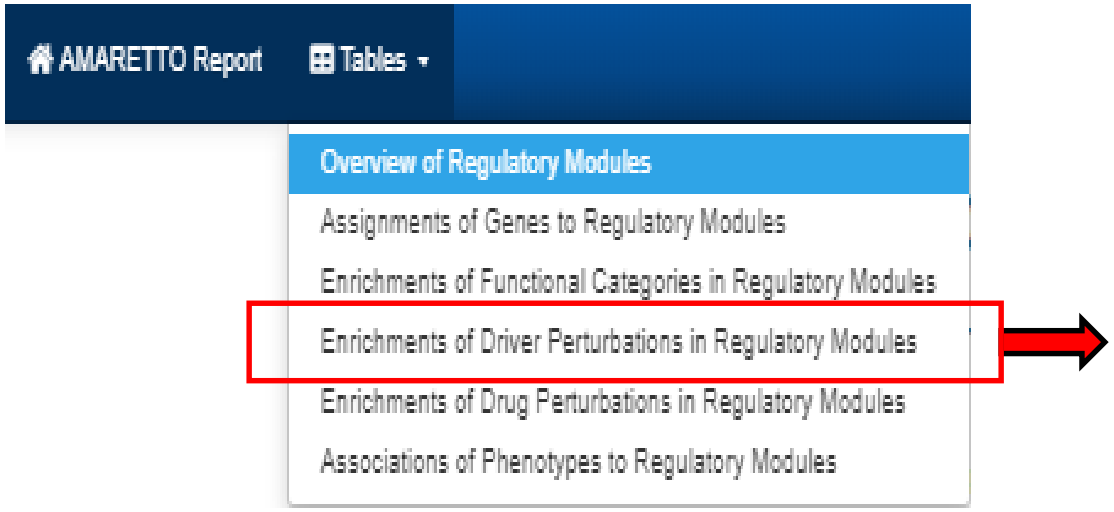


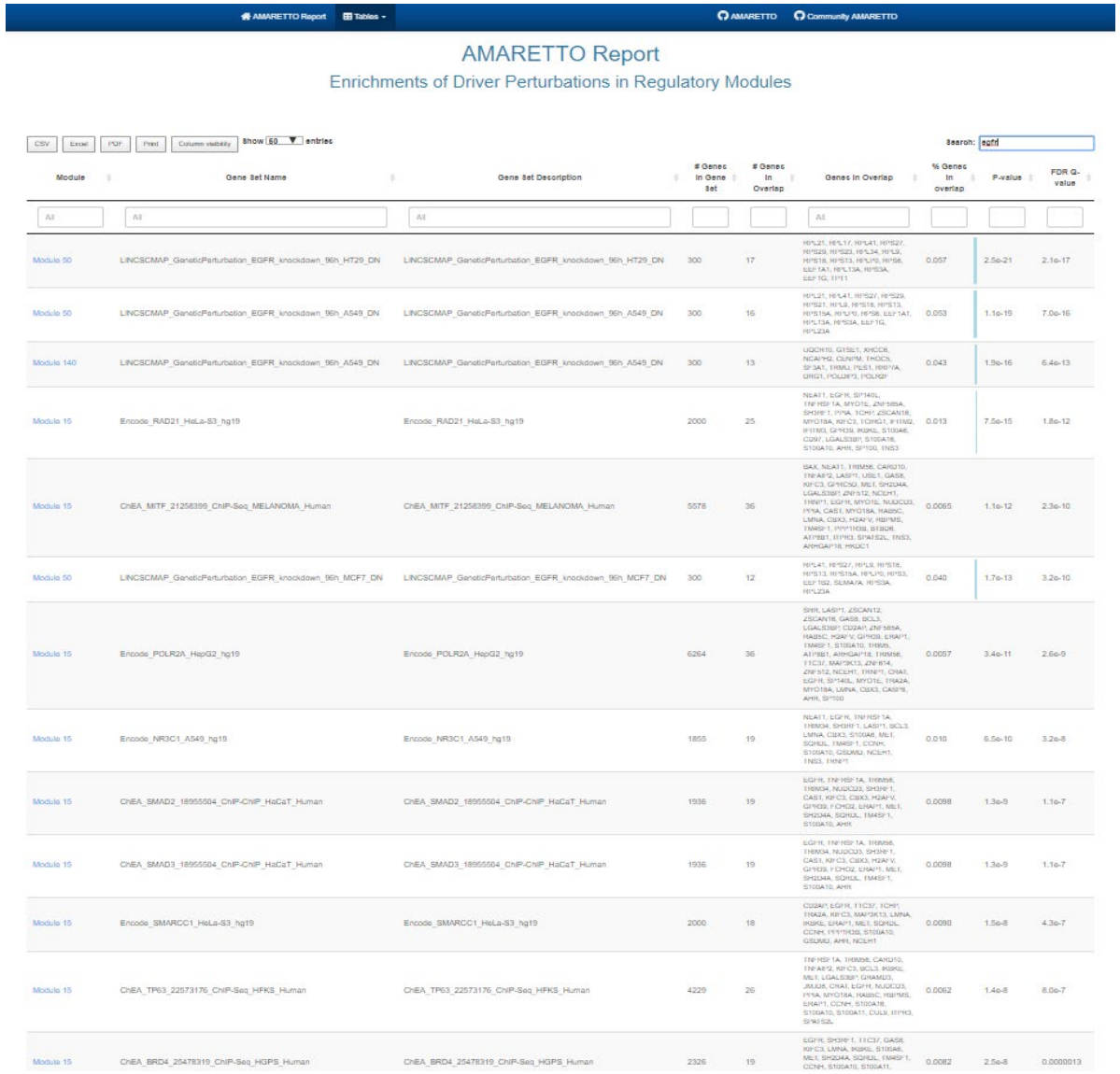
Associations of Phenotypes to Regulatory Modules

- Clinical characterization for clinical, molecular and imaging-derived phenotypes

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_PhenoAssociation.html

AMARETTO report LIHC



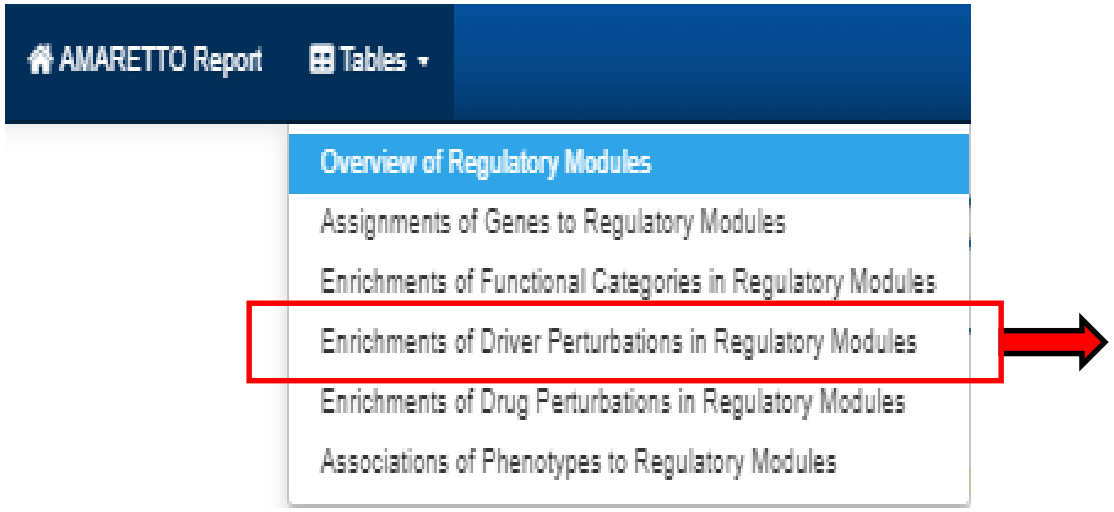


Enrichments of Driver Perturbations in Regulatory Modules

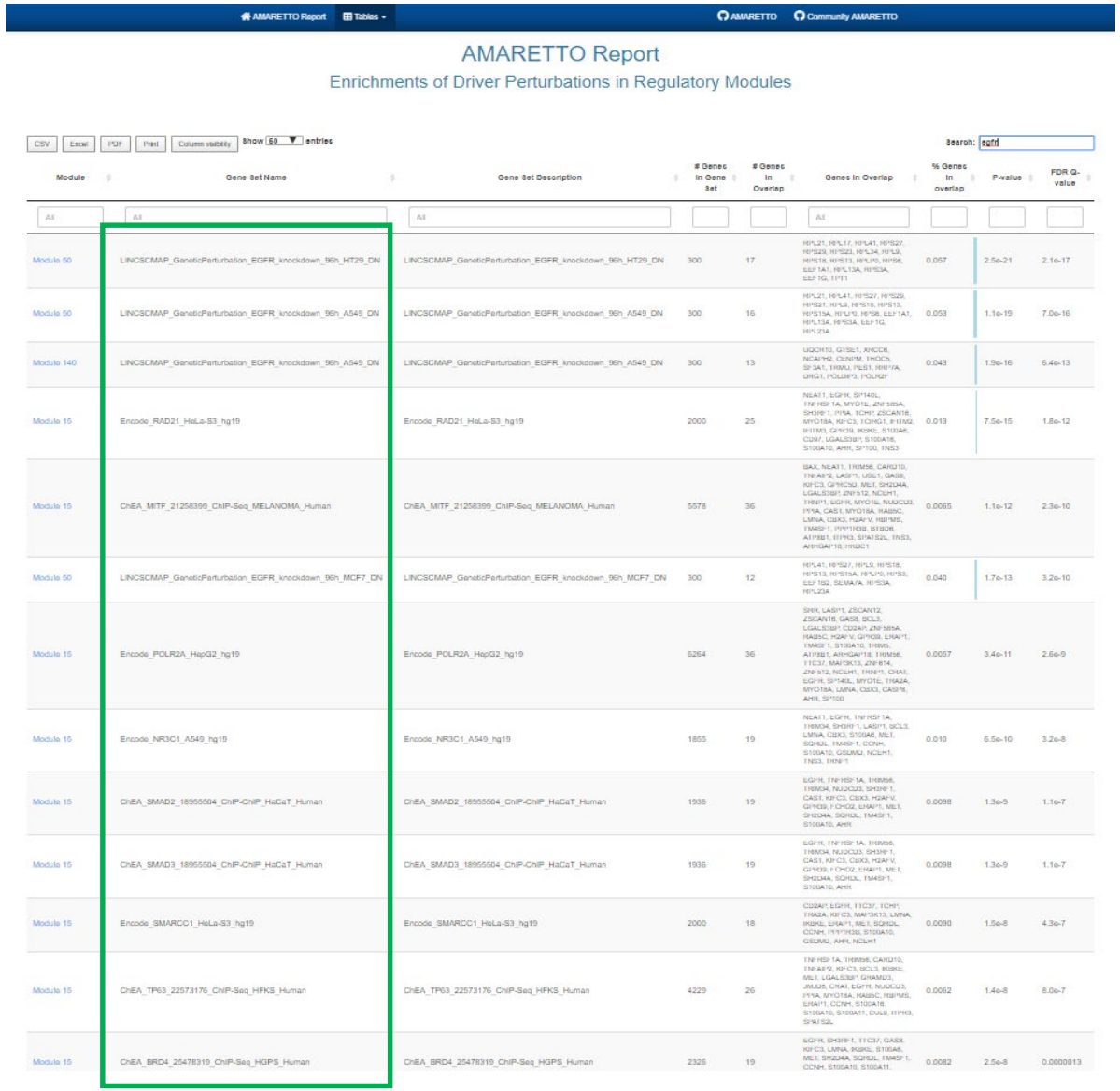
- Perturbation-AMARETTO v1 for driver discovery using genetic perturbations in model systems

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment_gp.html

AMARETTO report LIHC



Genetic perturbations from Encode, ChEA, LINCS/CMAP

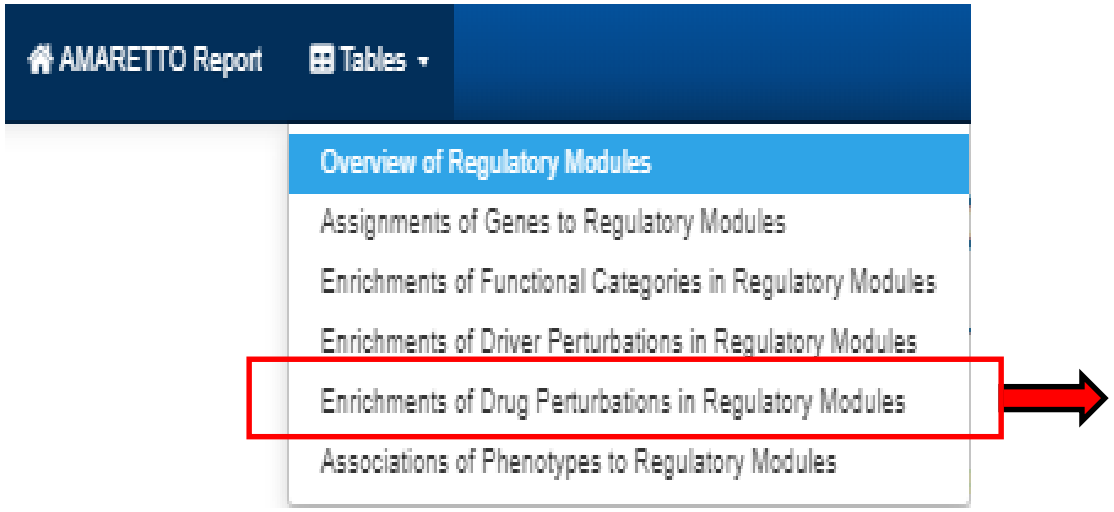


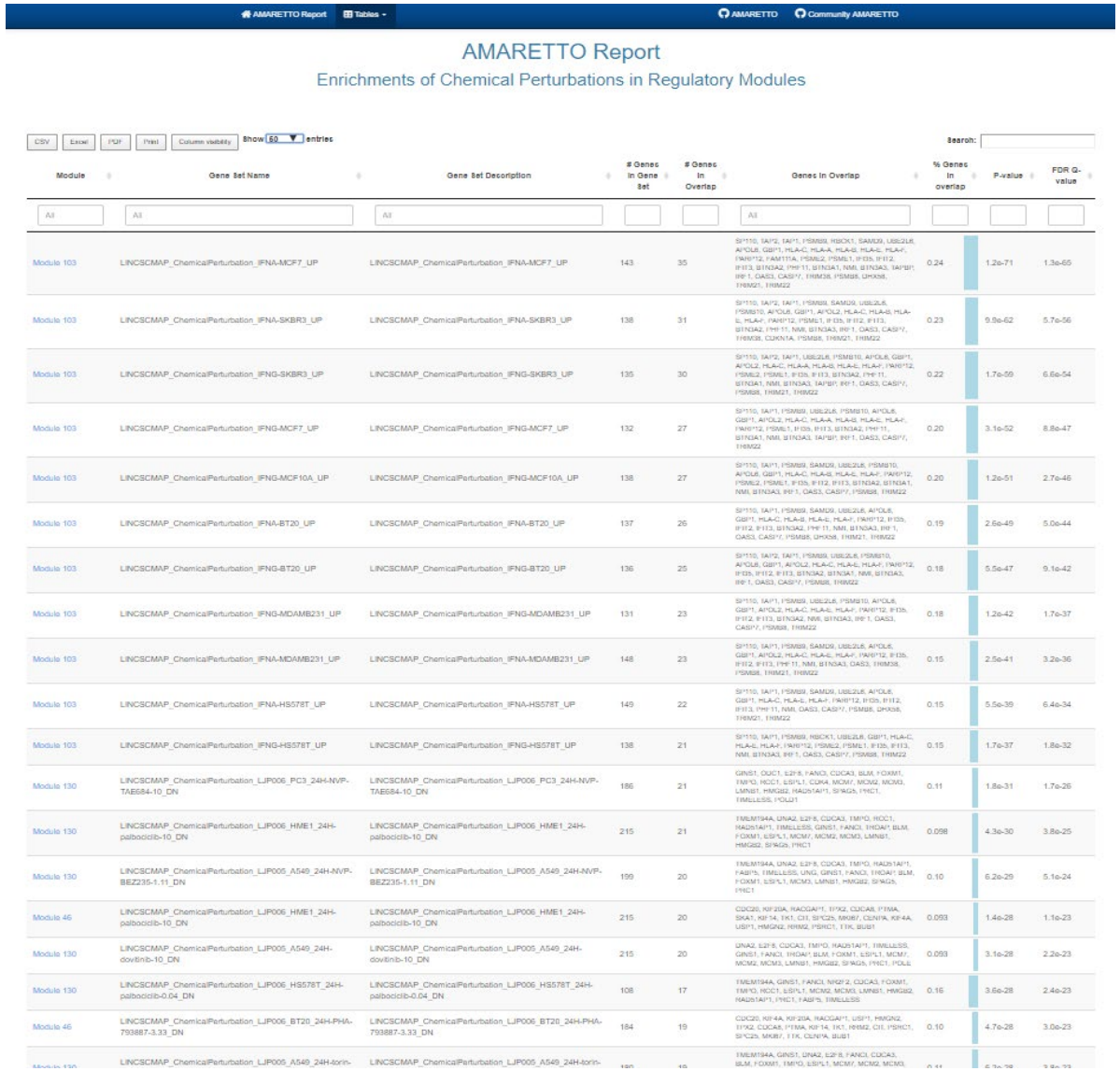
Enrichments of Driver Perturbations in Regulatory Modules

- Perturbation-AMARETTO v1 for driver discovery using genetic perturbations in model systems

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment_gp.html

AMARETTO report LIHC



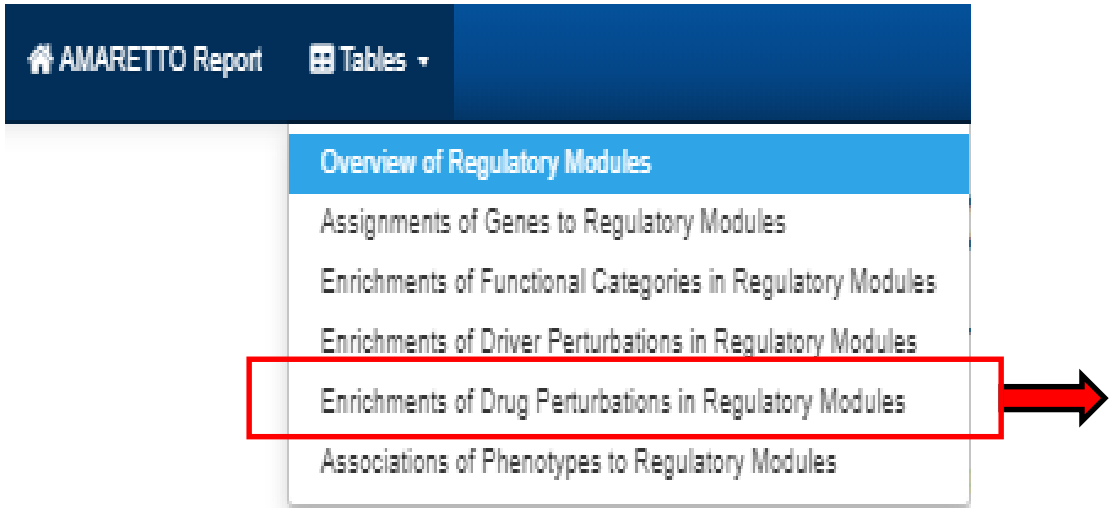


Enrichments of Drug Perturbations in Regulatory Modules

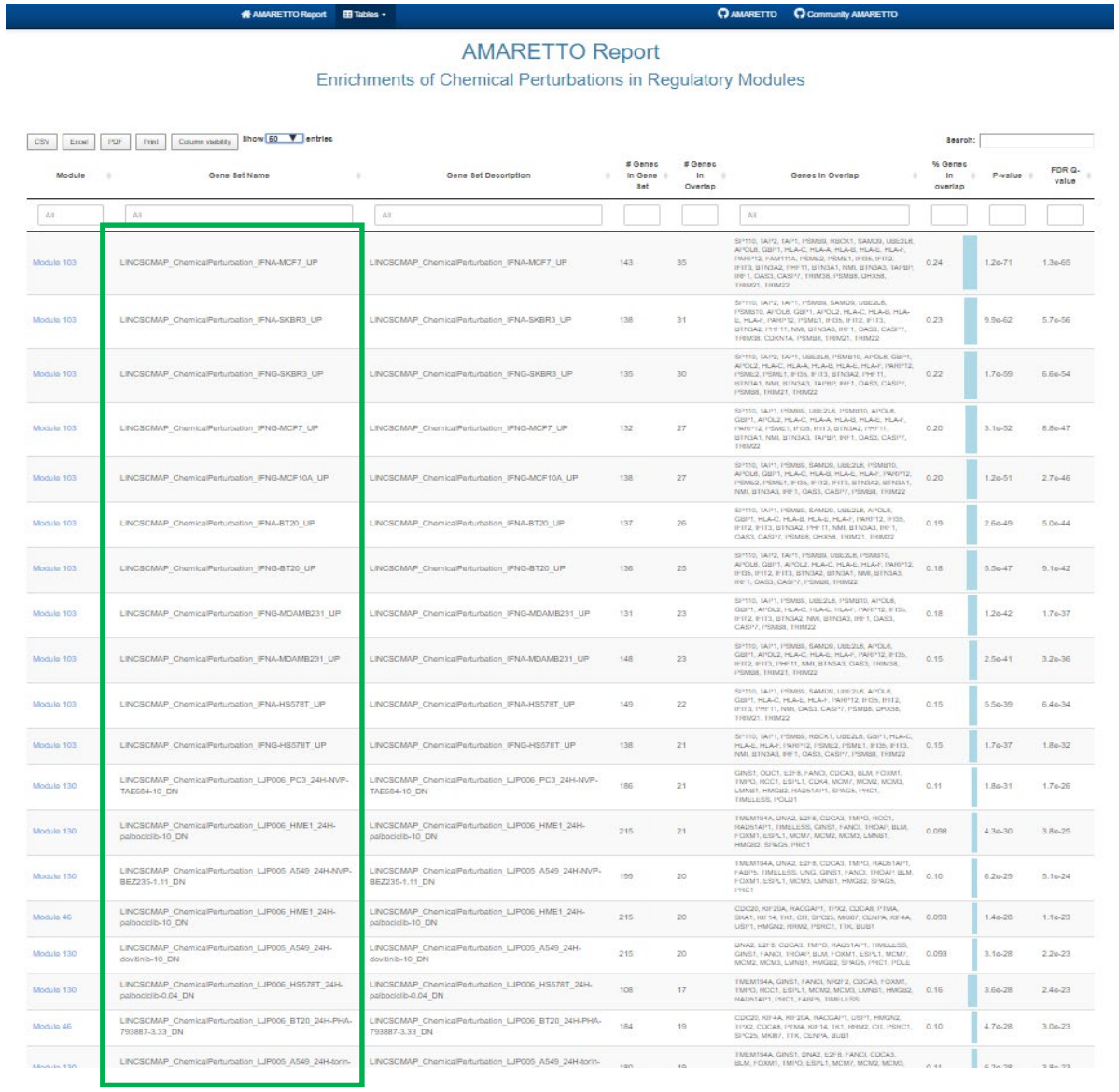
- Perturbation-AMARETTO v1 for drug discovery using chemical perturbations in model systems

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment_cp.html

AMARETTO report LIHC



Chemical perturbations from LINCS/CMAP



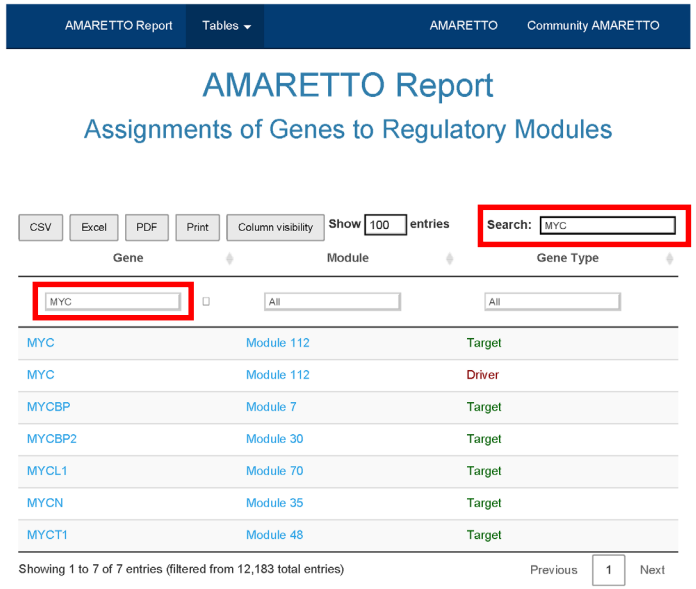
Enrichments of Drug Perturbations in Regulatory Modules

- Perturbation-AMARETTO v1 for drug discovery using chemical perturbations in model systems

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_GenesetsEnrichment_cp.html

AMARETTO report LIHC: Module(s) regulated by MYC?

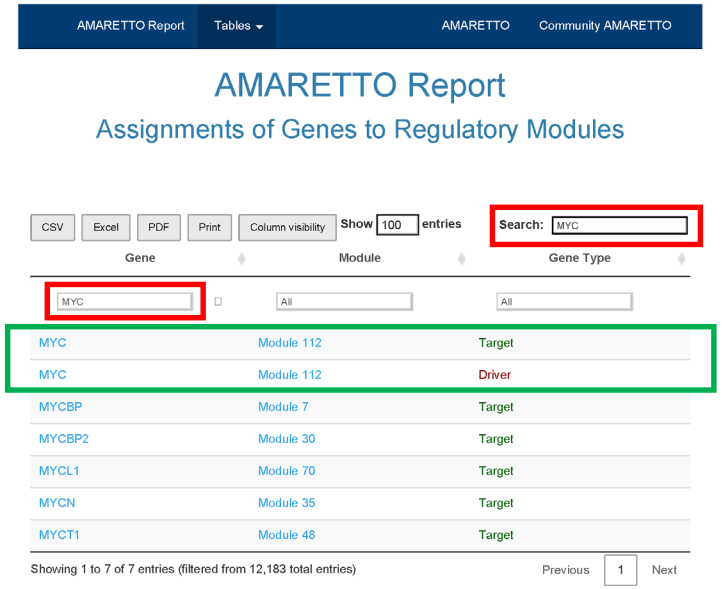
AMARETTO report LIHC: Module(s) regulated by MYC?



Search for module(s) regulated by MYC

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

AMARETTO report LIHC: Module(s) regulated by MYC



Search for module(s) regulated by MYC

⇒ Module 112 is regulated by MYC

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

AMARETTO report LIHC: Module(s) regulated by MYC

AMARETTO Report | Tables | AMARETTO | Community AMARETTO

AMARETTO Report

Assignments of Genes to Regulatory Modules

CSV | Excel | PDF | Print | Column visibility | Show 100 entries | Search: MYC

Gene	Module	Gene Type
MYC	All	All
MYC	Module 112	Target
MYC	Module 112	Driver
MYCBP	Module 7	Target
MYCBP2	Module 30	Target
MYCL1	Module 70	Target
MYCN	Module 35	Target
MYCT1	Module 48	Target

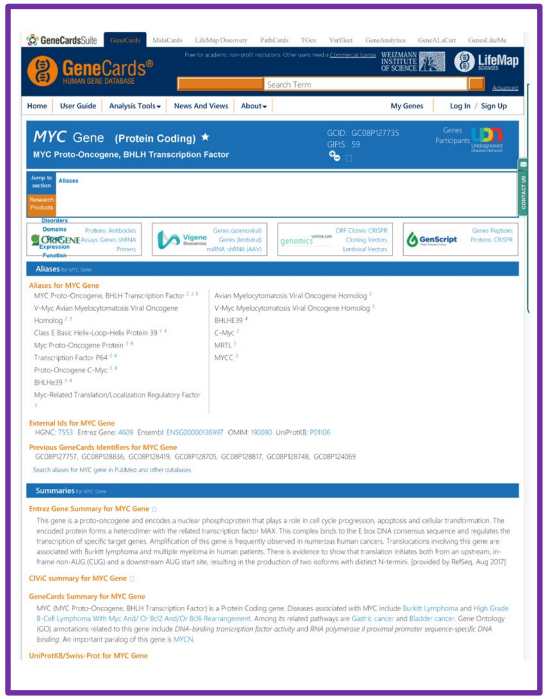
Showing 1 to 7 of 7 entries (filtered from 12,183 total entries) Previous 1 Next

Search for module(s) regulated by MYC

⇒ Module 112 is regulated by MYC

Link to GeneCards description of MYC

GeneCards



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

AMARETTO report LIHC: Module(s) regulated by MYC

AMARETTO Report | Tables | AMARETTO | Community AMARETTO

AMARETTO Report

Assignments of Genes to Regulatory Modules

CSV | Excel | PDF | Print | Column visibility | Show 100 entries | Search: MYC

Gene	Module	Gene Type
MYC	All	All
MYC	Module 112	Target
MYC	Module 112	Driver
MYCBP	Module 7	Target
MYCBP2	Module 30	Target
MYCL1	Module 70	Target
MYCN	Module 35	Target
MYCT1	Module 48	Target

Showing 1 to 7 of 7 entries (filtered from 12,183 total entries) Previous 1 Next

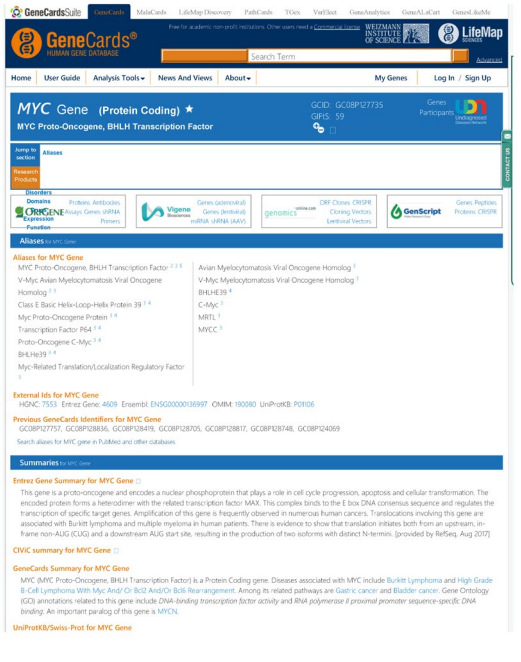
Search for module(s) regulated by MYC

⇒ Module 112 is regulated by MYC

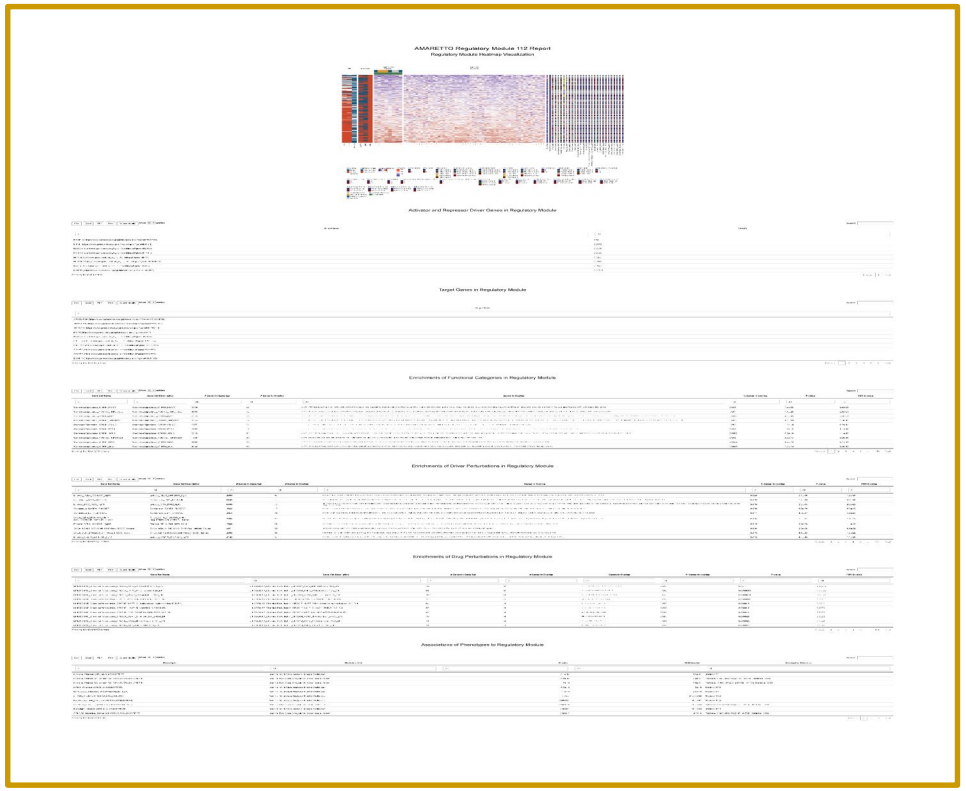
Link to GeneCards description of MYC

Link to Module 112 report page

GeneCards



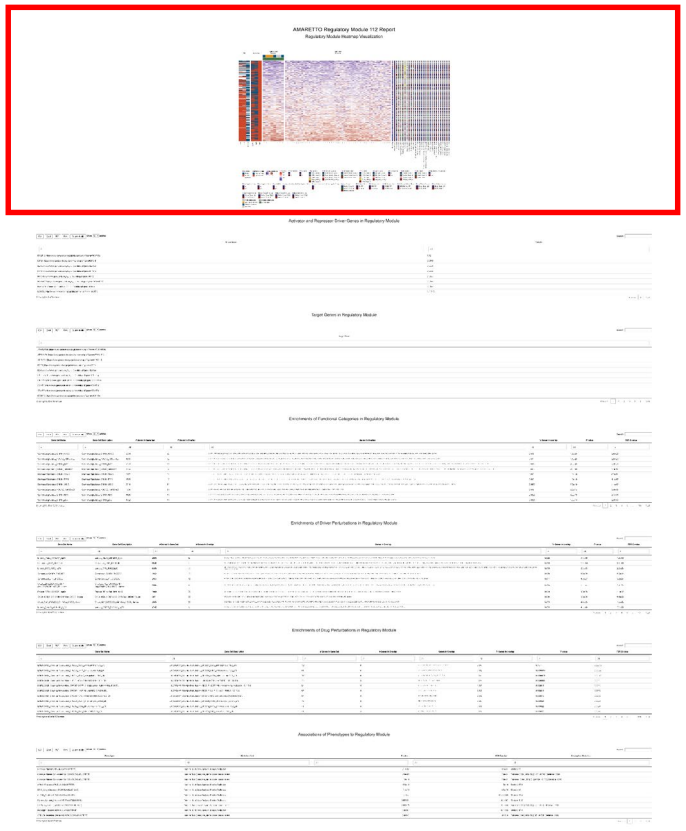
Detailed report of MYC-driven Module 112



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/modules/module112.html

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index_AllGenes.html

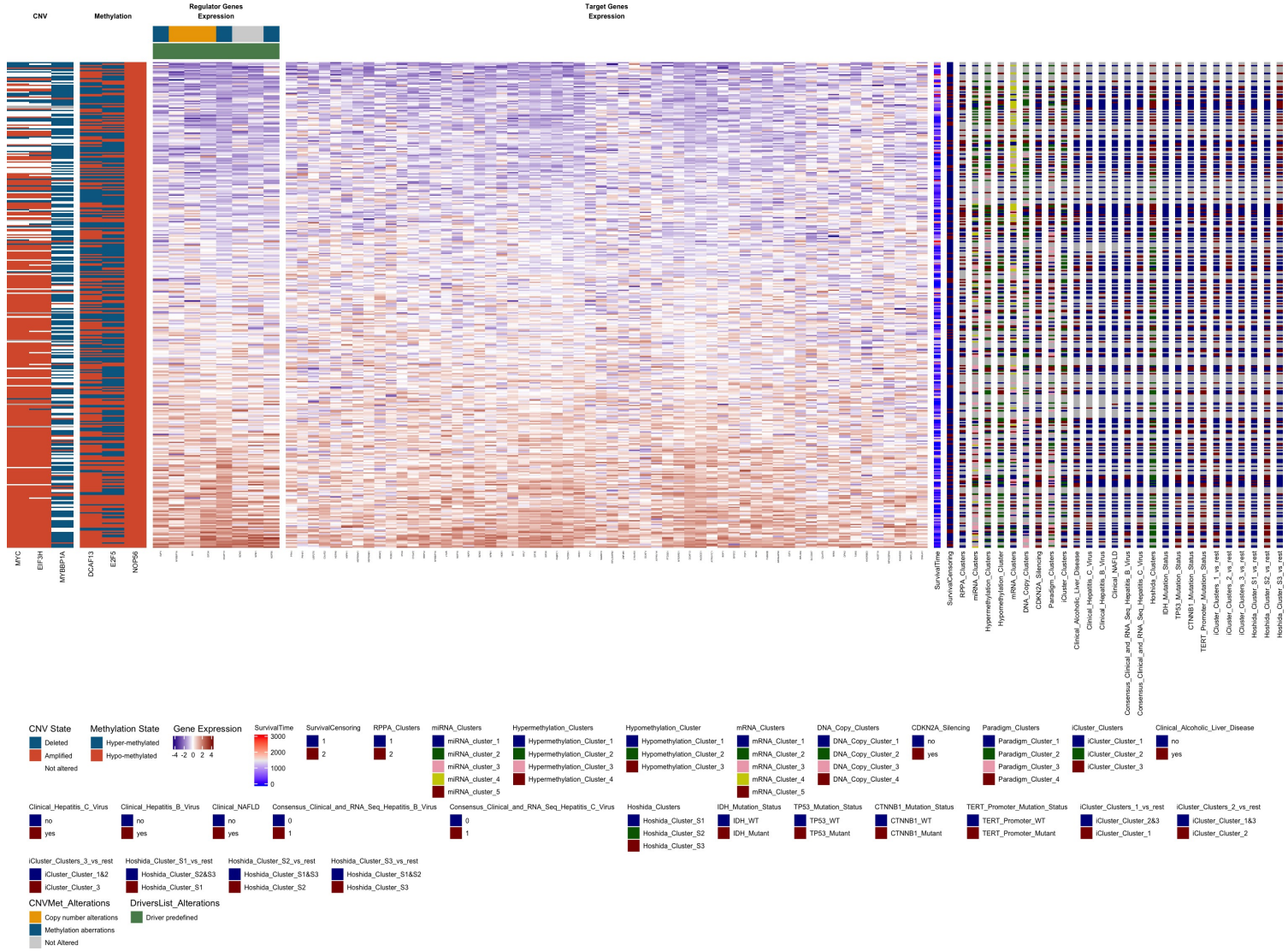
AMARETTO report LIHC



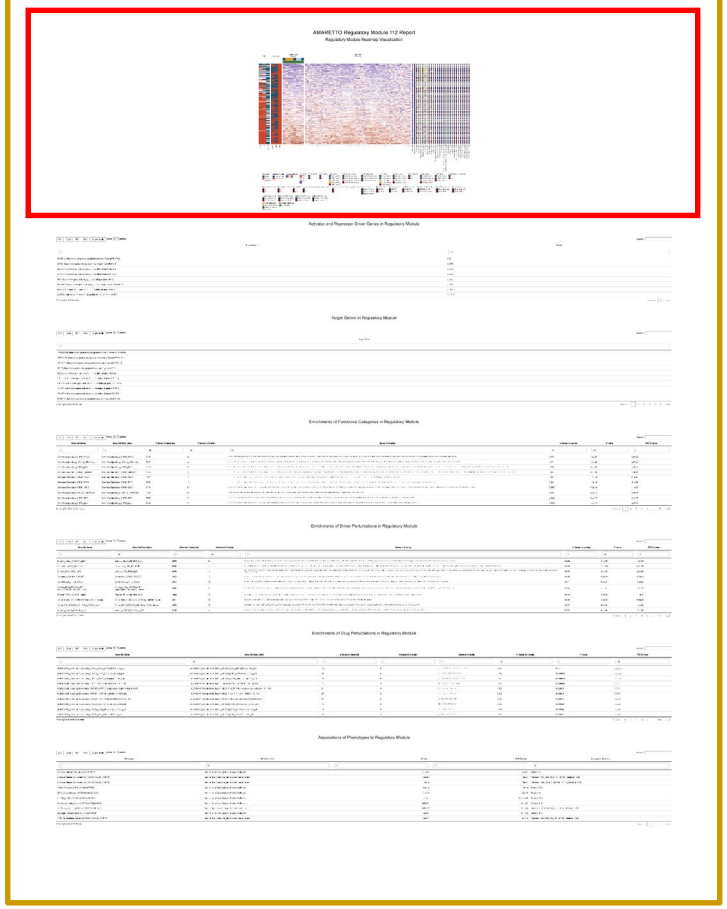
Detailed report of MYC-driven Module 112:
heatmap visualization

AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization



AMARETTO report LIHC



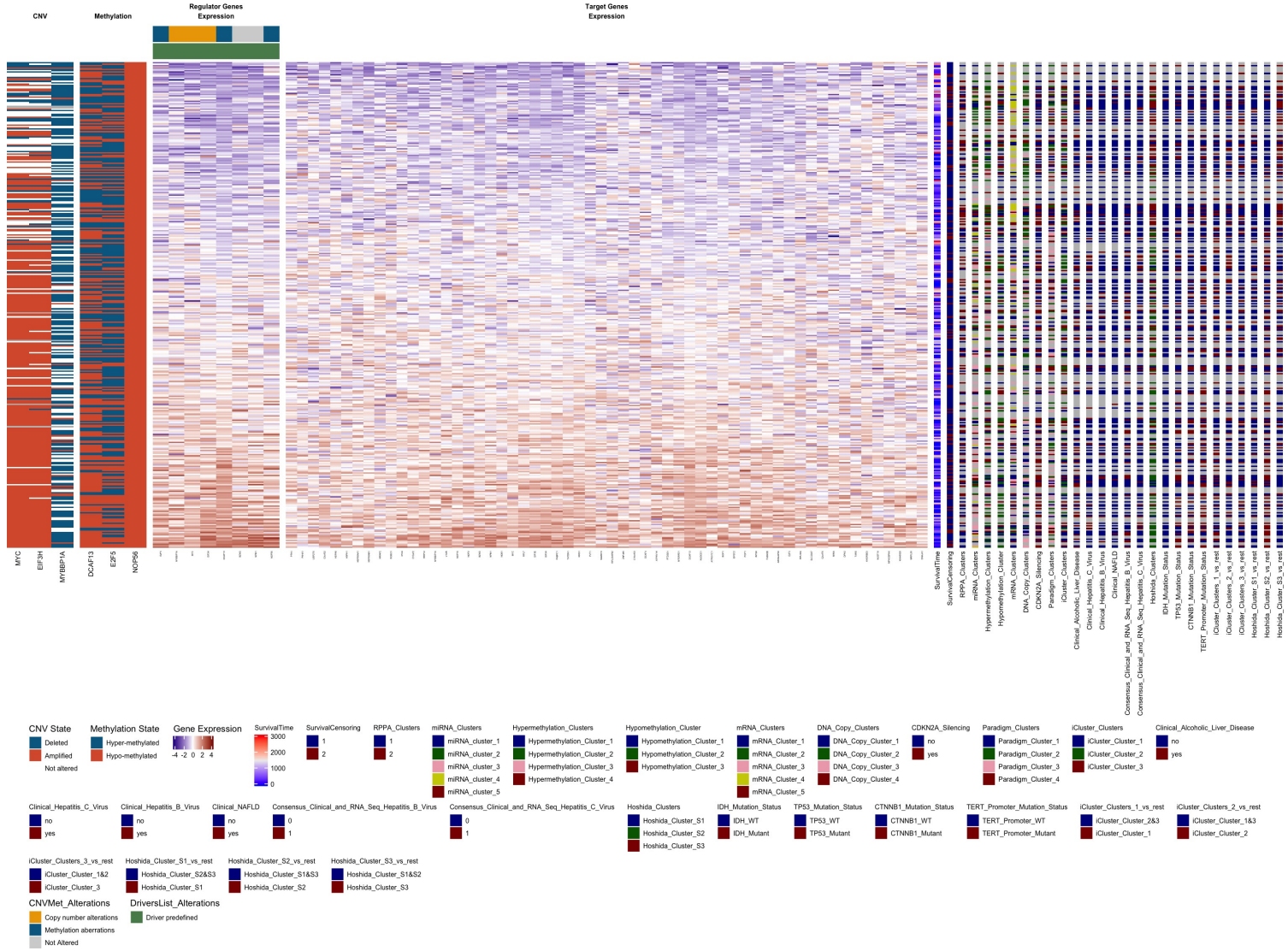
Detailed report of MYC-driven Module 112:
heatmap visualization

AMARETTO Regulatory Module 112 Report

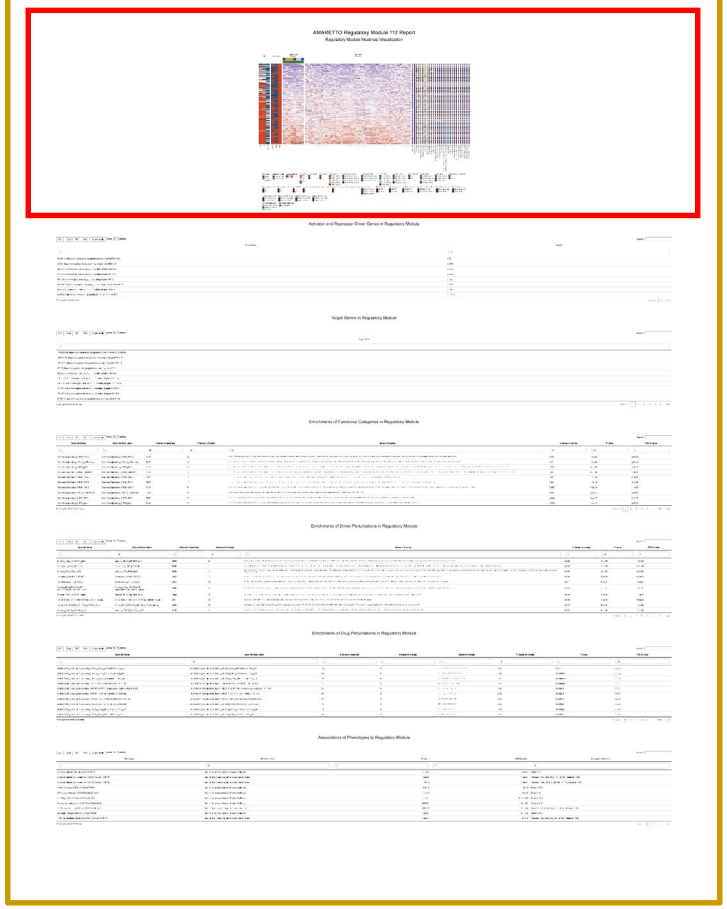
Regulatory Module Heatmap Visualization

Rows =

Patient Tumor Samples



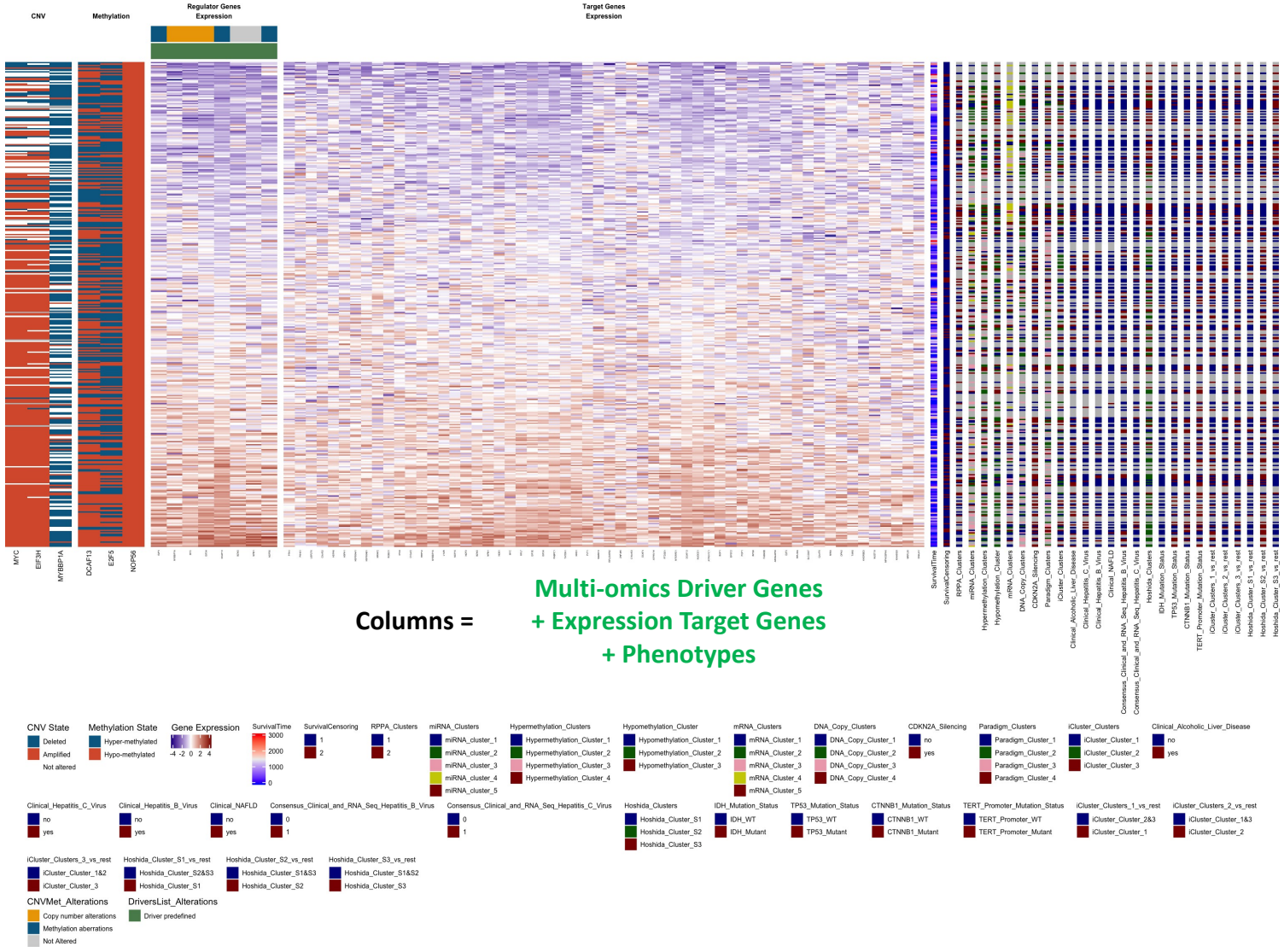
AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
heatmap visualization

AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization



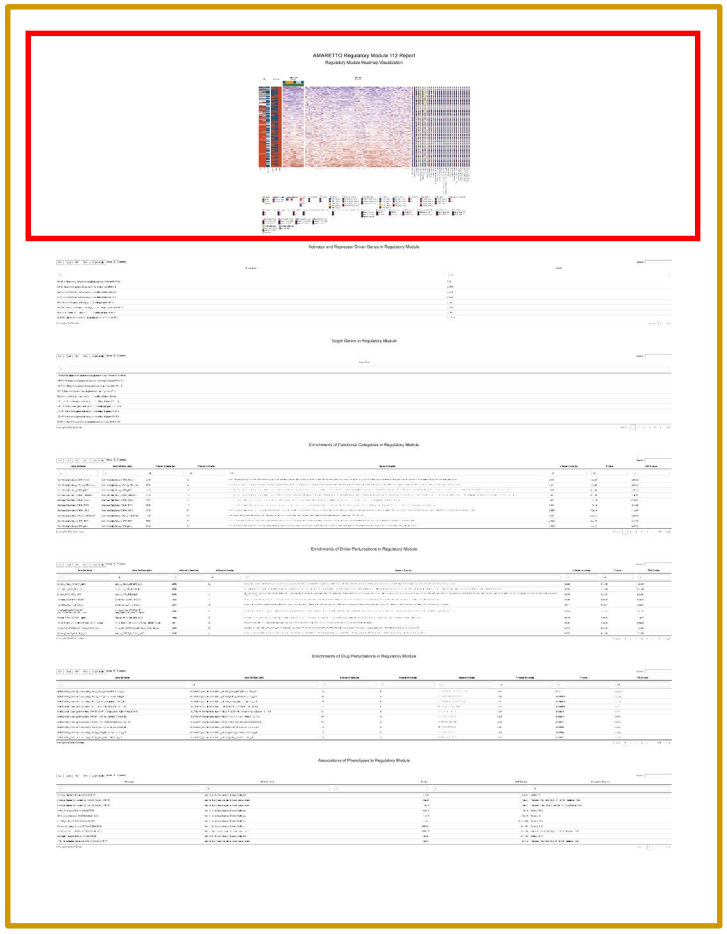
Rows =

Patient Tumor Samples

Columns =

Multi-omics Driver Genes + Expression Target Genes + Phenotypes

AMARETTO report LIHC

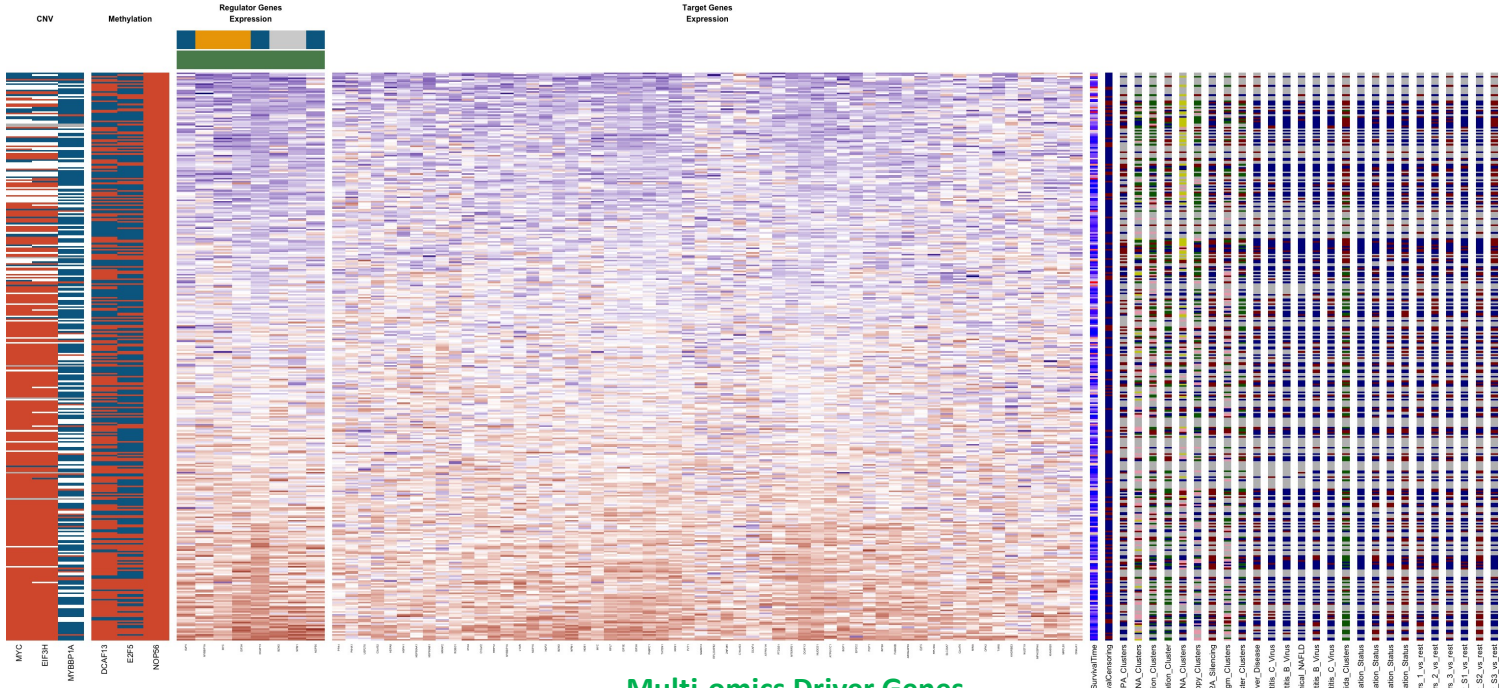


Detailed report of MYC-driven Module 112:
heatmap visualization

AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization

Target Genes Expression

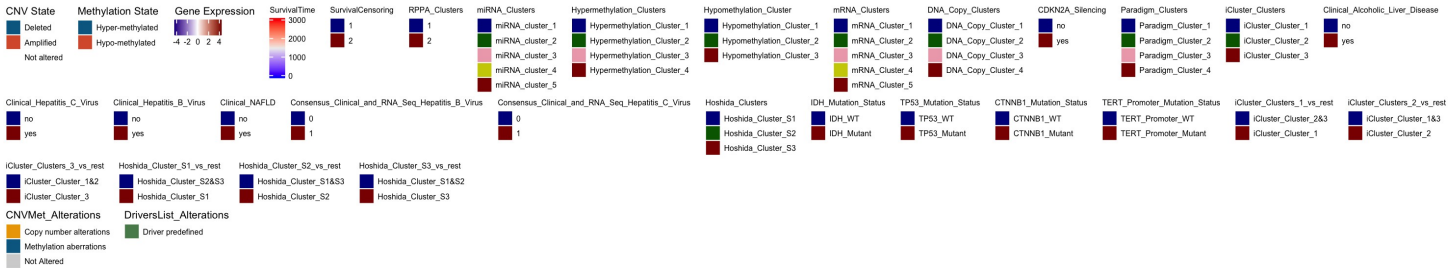


Rows =

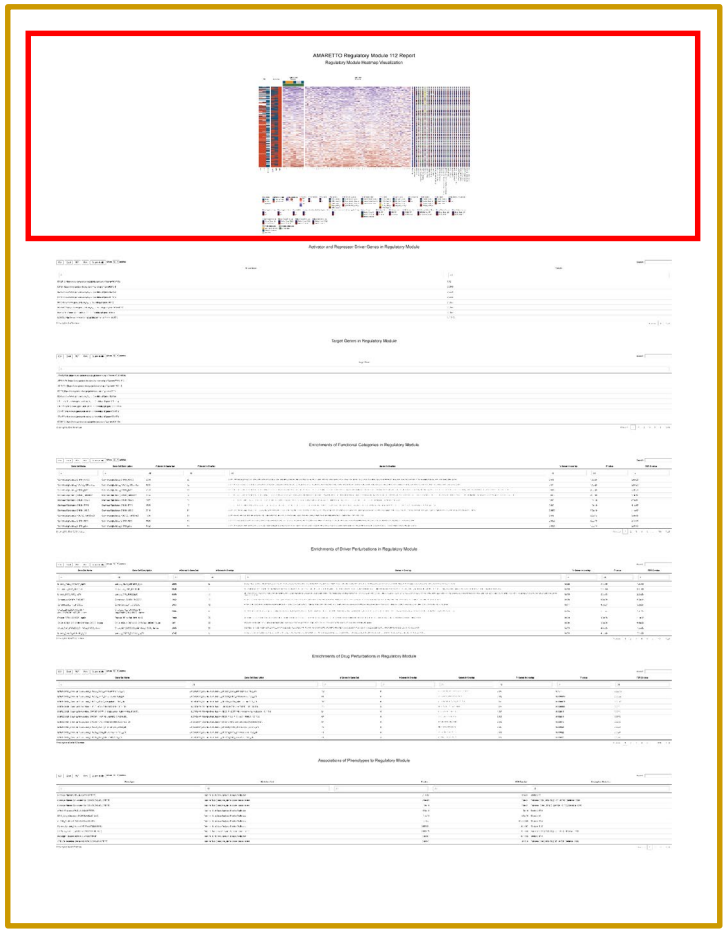
Patient Tumor Samples

Columns =

Multi-omics Driver Genes + Expression Target Genes + Phenotypes



AMARETTO report LIHC



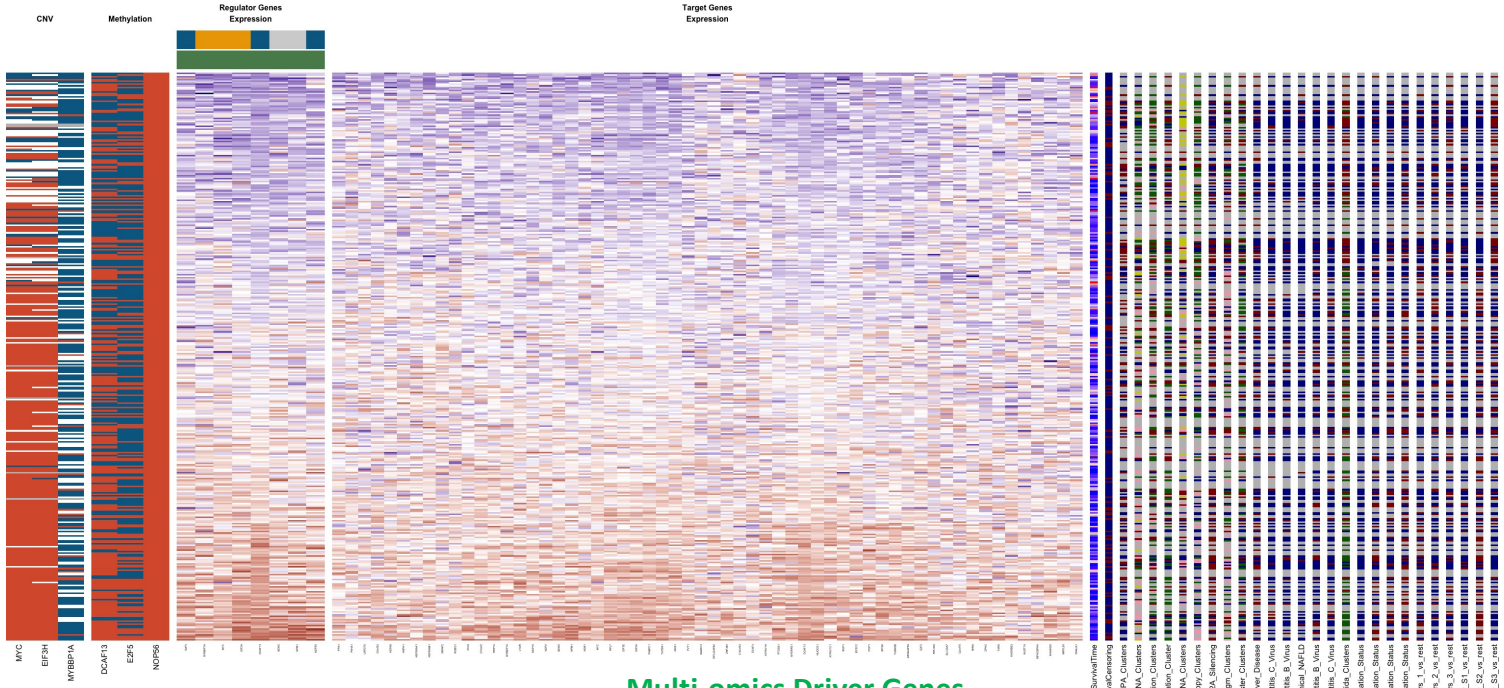
Detailed report of MYC-driven Module 112:
heatmap visualization

AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization

Driver Genes
CNV MET Expression

Target Genes
Expression

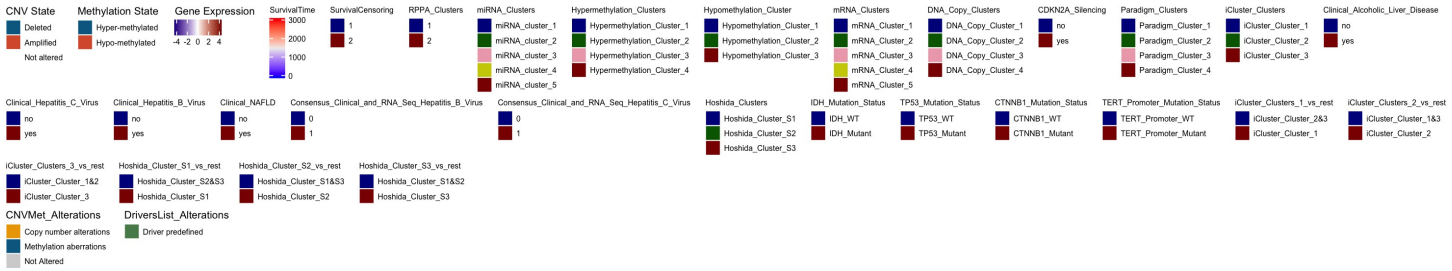


Rows =

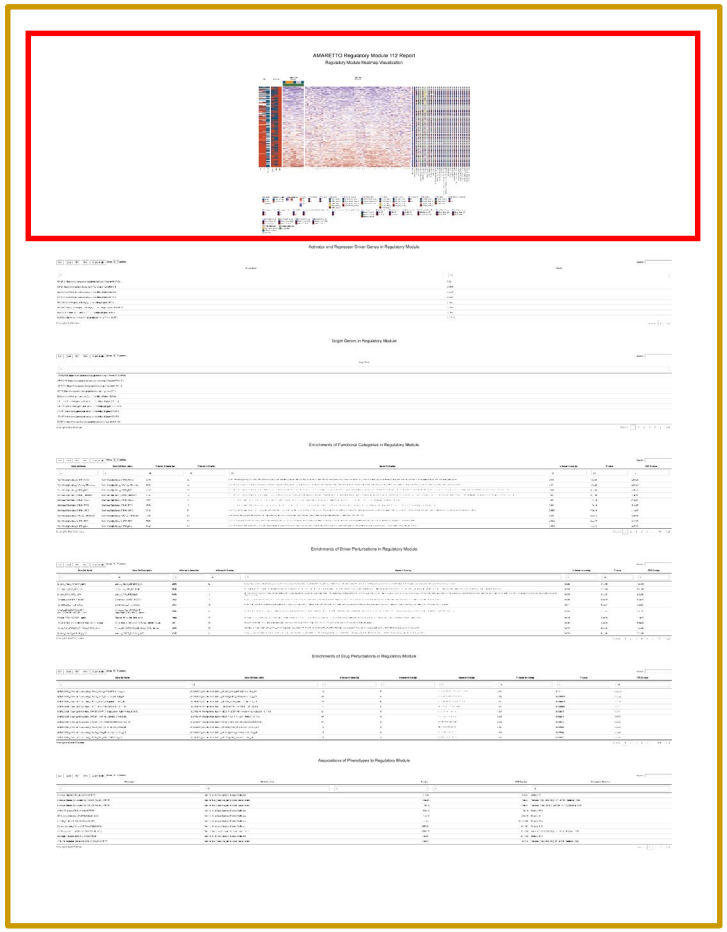
Patient Tumor Samples

Columns =

Multi-omics Driver Genes
+ Expression Target Genes
+ Phenotypes



AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
heatmap visualization

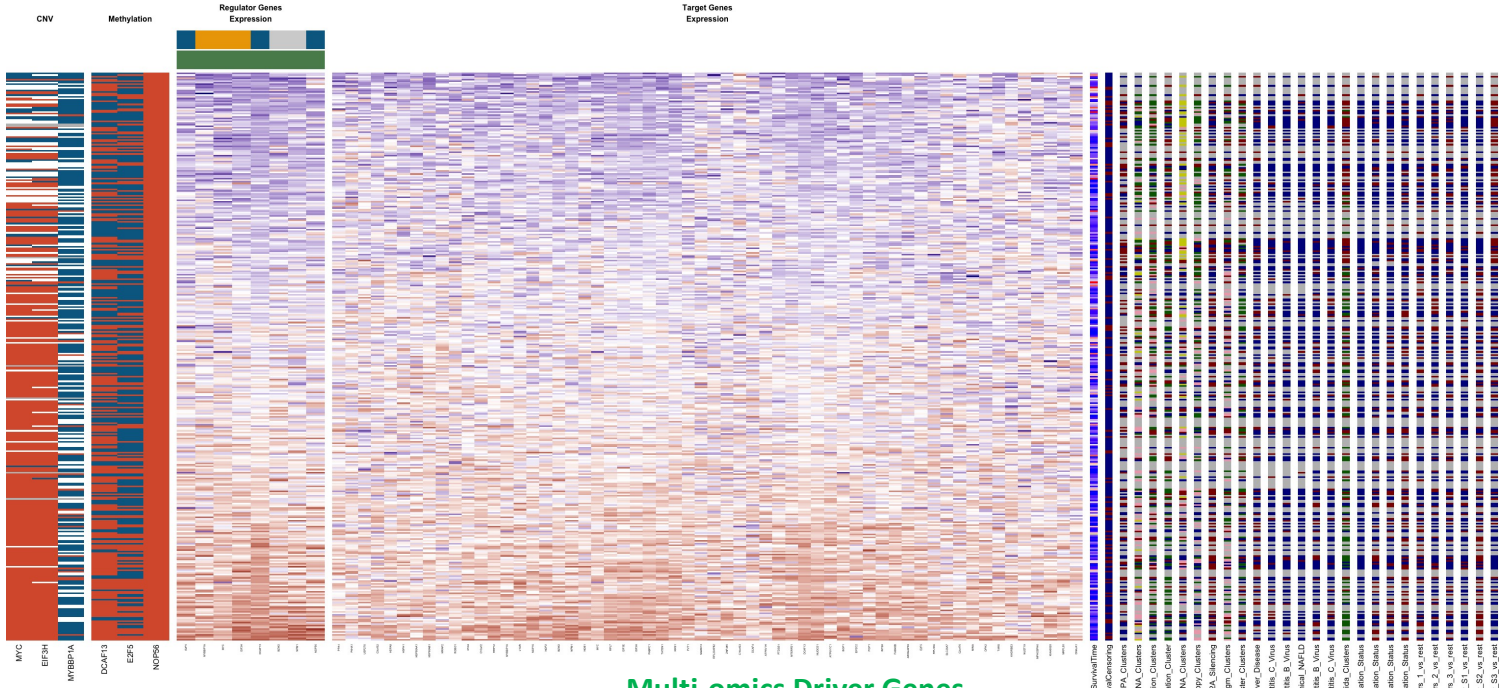
AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization

Driver Genes
CNV MET Expression

Target Genes
Expression

Phenotypes
Clinical, molecular, imaging

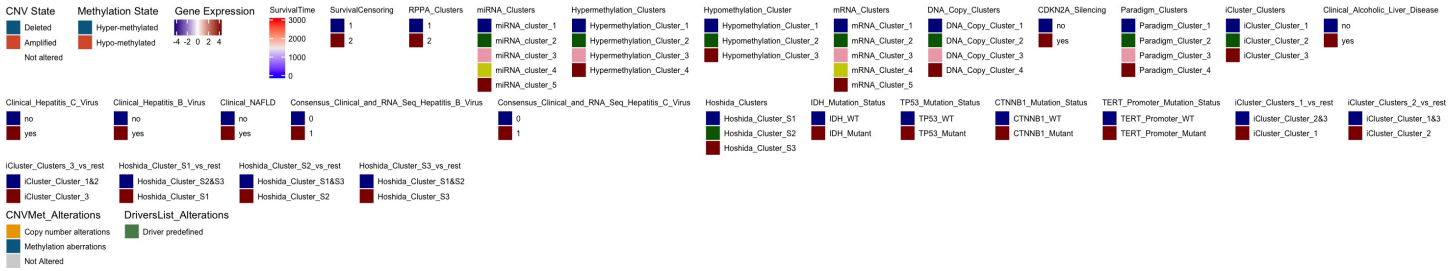


Rows =

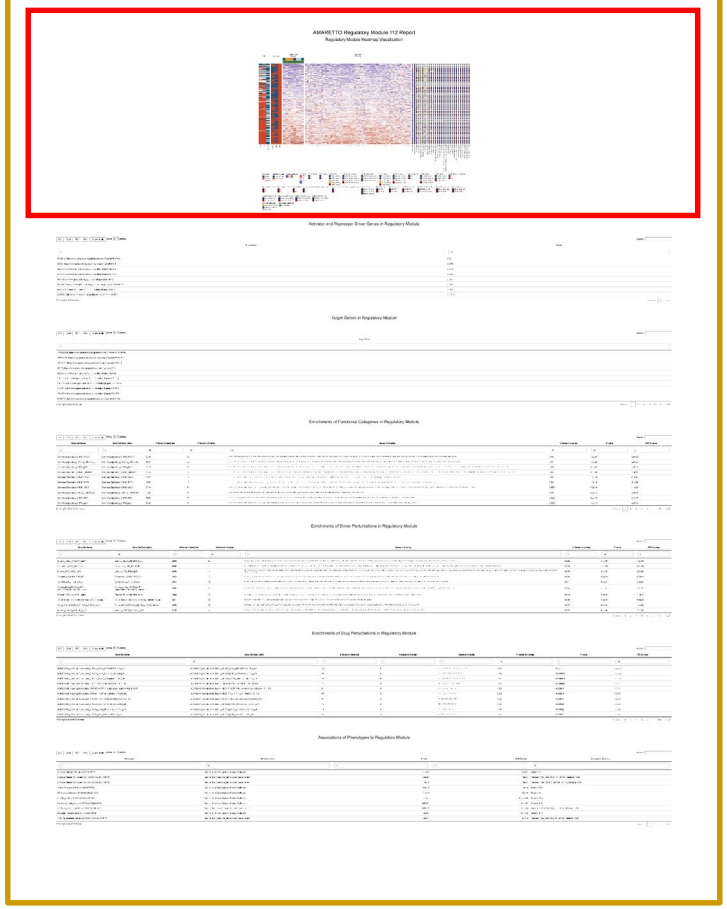
Patient Tumor Samples

Columns =

Multi-omics Driver Genes
+ Expression Target Genes
+ Phenotypes



AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
heatmap visualization

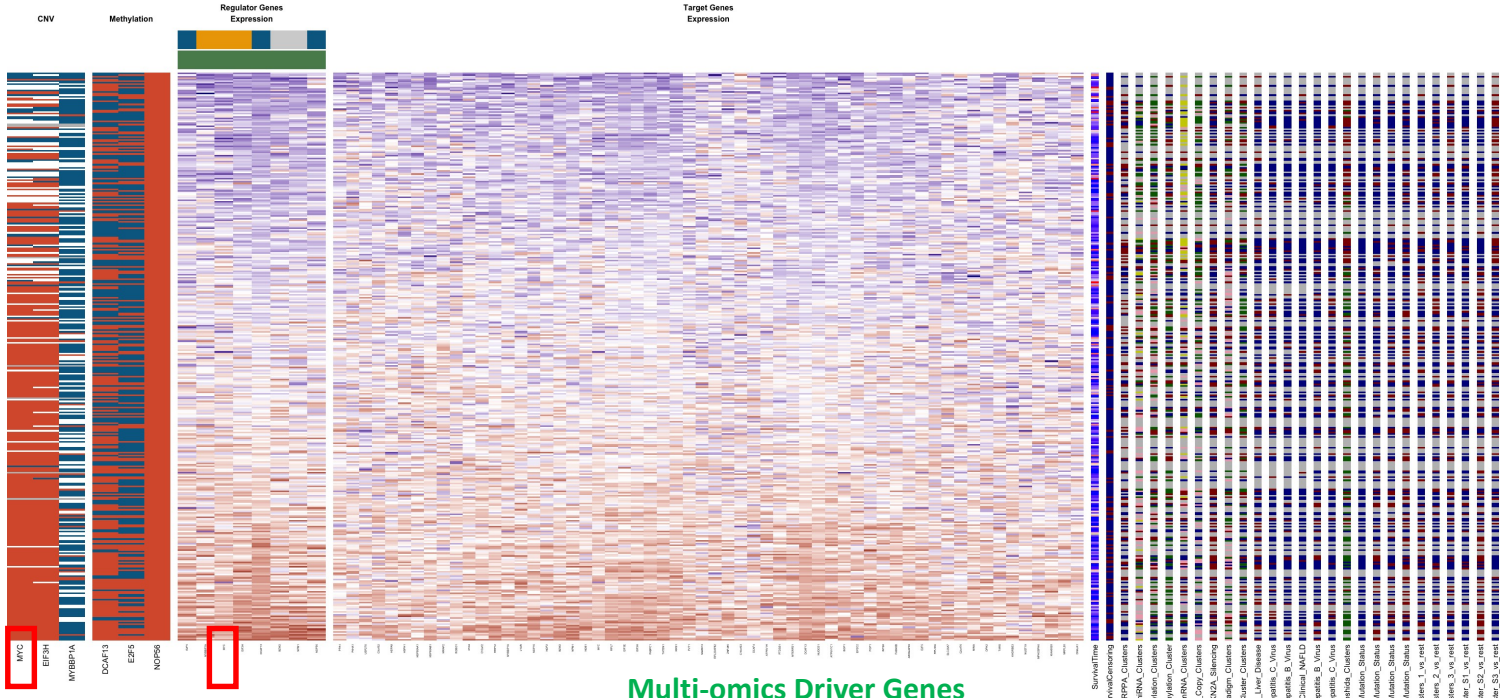
AMARETTO Regulatory Module 112 Report

Regulatory Module Heatmap Visualization

Driver Genes
CNV MET Expression

Target Genes
Expression

Phenotypes
Clinical, molecular, imaging



Rows =

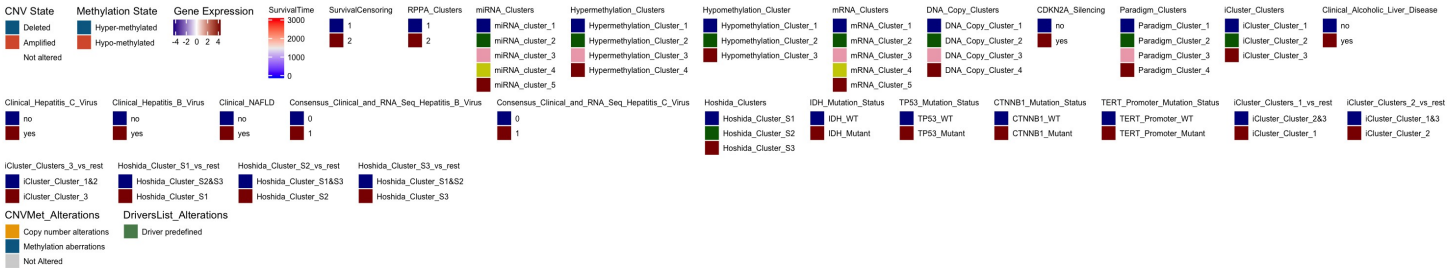
Patient Tumor Samples

MYC CNV

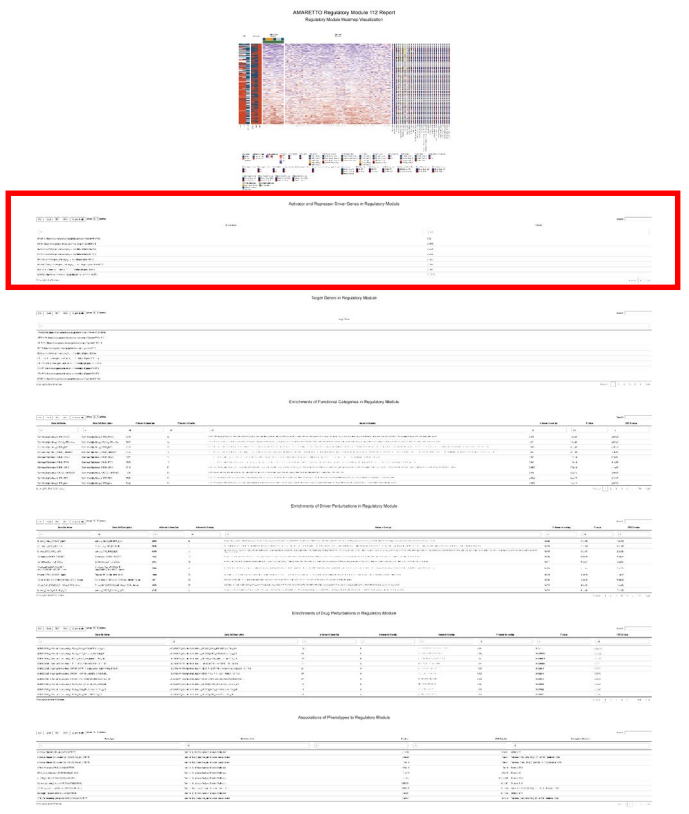
MYC EXPR

Columns =

Multi-omics Driver Genes
+ Expression Target Genes
+ Phenotypes



AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
activator and repressor driver genes

Activator and Repressor Driver Genes in Regulatory Module

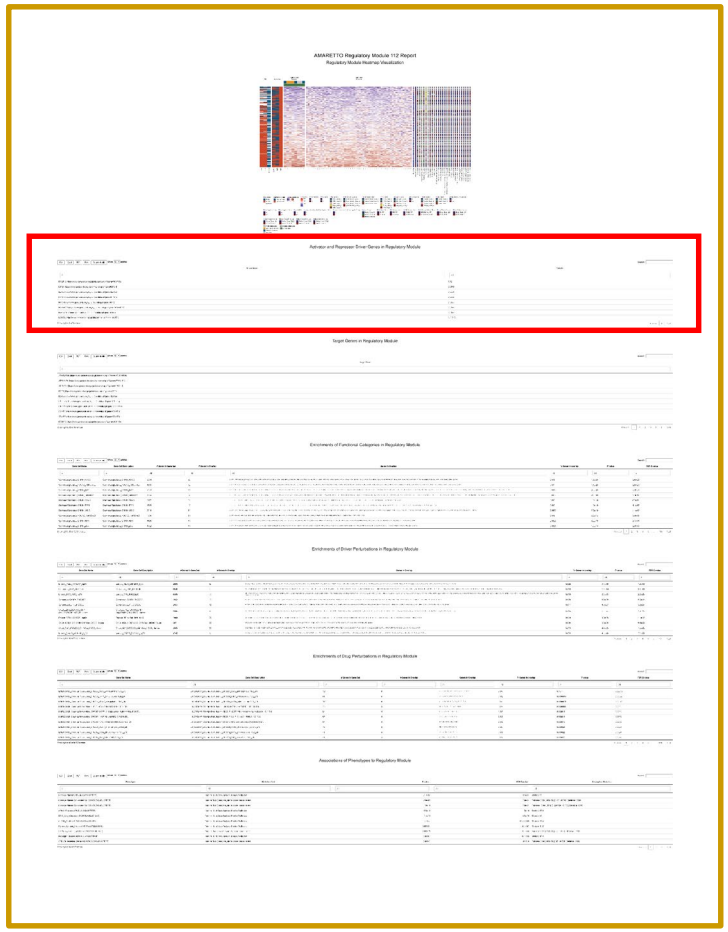
CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

Driver Gene	Weight
DCAF13	0.29
EIF3H	0.0348
BZW2	0.0206
NPM1	0.0205
MYC	0.0179
MYBBP1A	0.0148
E2F5	0.0129
NOP56	0.00215

Showing 1 to 8 of 8 entries

Previous 1 Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
activator and repressor driver genes

Activator and Repressor Driver Genes in Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

Driver Gene	Weight
All	All
DCAF13	0.29
EIF3H	0.0348
BZW2	0.0206
NPM1	0.0205
MYC	0.0179
MYBBP1A	0.0148
E2F5	0.0129
NOP56	0.00215

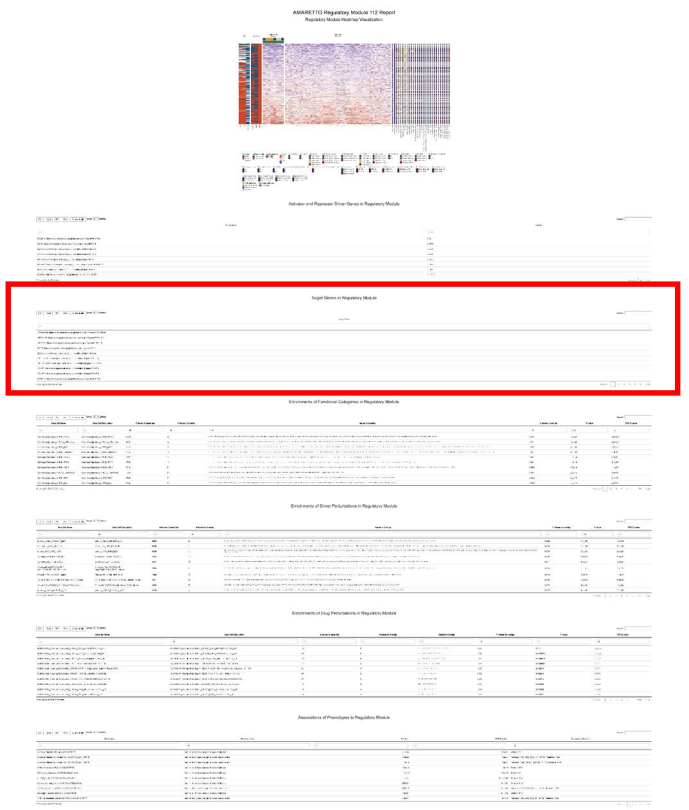
MYC

Showing 1 to 8 of 8 entries

Previous 1 Next

Activator Driver Genes (weight > 0)
Repressor Driver Genes (weight < 0)

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
target genes

Target Genes in Regulatory Module

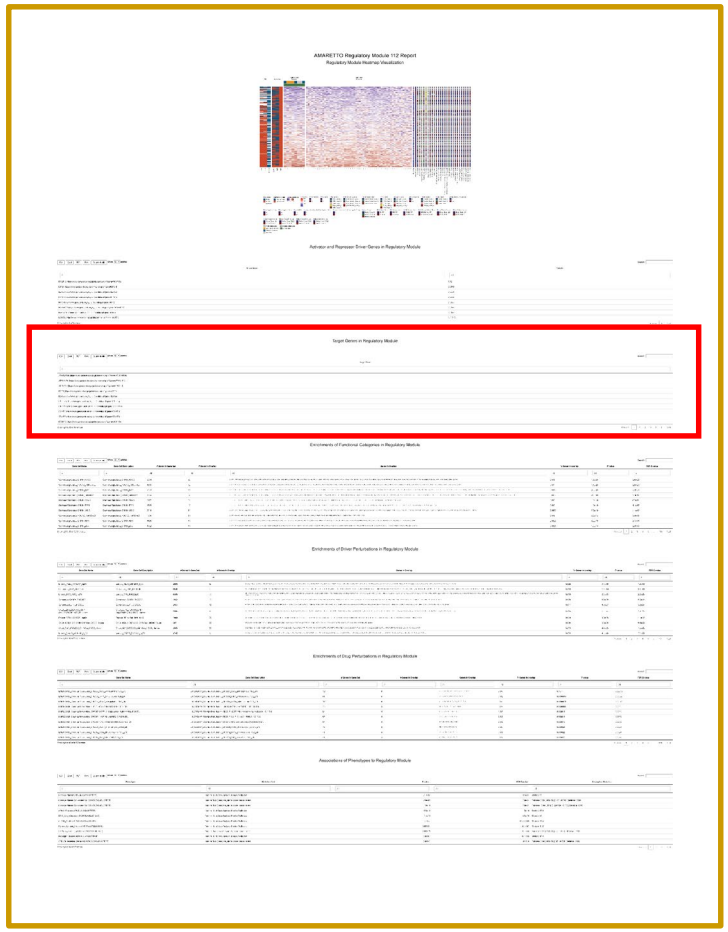
CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

Target Gene
All
ARHGAP39
ATP6V1C1
ATP6V1H
BOP1
BZW2
C10orf2
C14orf33
C2orf76
C3orf32
DCAF13

Showing 1 to 10 of 58 entries

Previous 1 2 3 4 5 6 Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
target genes

Target Genes in Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

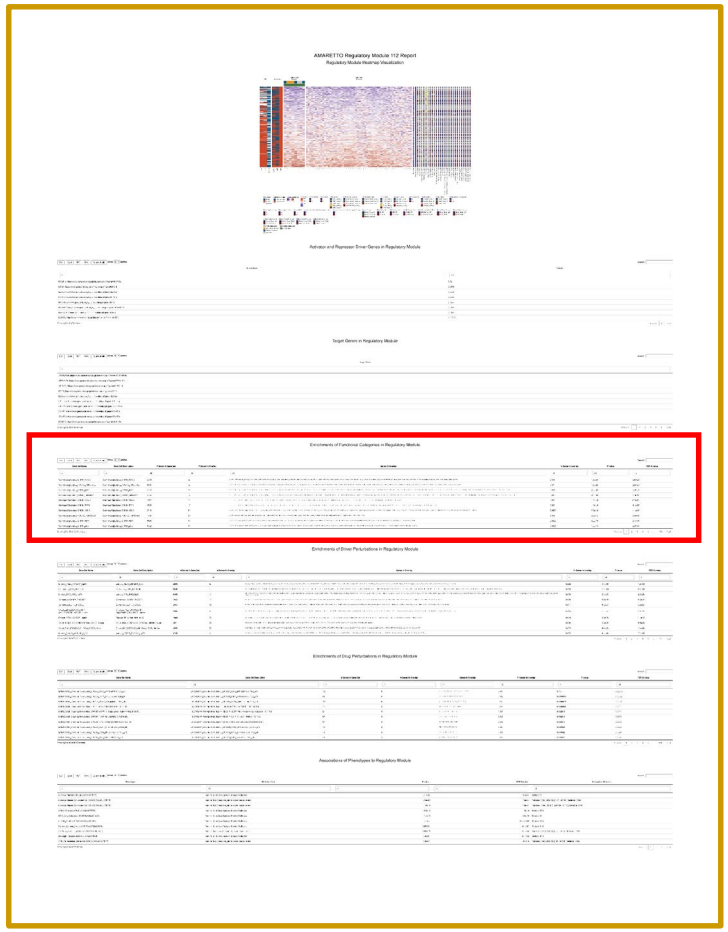
Target Gene
All
ARHGAP39
ATP6V1C1
ATP6V1H
BOP1
BZW2
C10orf2
C14orf33
C2orf76
C3orf32
DCAF13

Showing 1 to 10 of 58 entries | Previous | 1 | 2 | 3 | 4 | 5 | 6 | Next

58 Target Genes in Module 112:

ARHGAP39, ATP6V1C1, ATP6V1H, BOP1, BZW2, C10orf2, C14orf33, C2orf76, C3orf32, DCAF13, DCAF4, DNAJA1, DPH2, E2F5, EIF2C2, EIF3E, EIF3H, FAM49B, HSP90AA1, HSP90AB1, HSPA8, HSPH1, INTS8, IPO4, KHDRBS3, KIAA0020, LYAR, MINA, MPHOSPH6, MRAP2, MRPL50, MTERFD1, MYBBP1A, MYC, NOB1, NOP16, NOP2, NPM1, NUDCD1, NUDT19, P4HA1, PABPC1, POP1, PPA1, PTDSS1, PVT1, ROBO1, RPL23AP82, RPL36A, RPL7, RRP12, RRS1, SAMD13, SLC26A7, TARS, TATDN1, USP27X, ZNF485

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
functional characterization

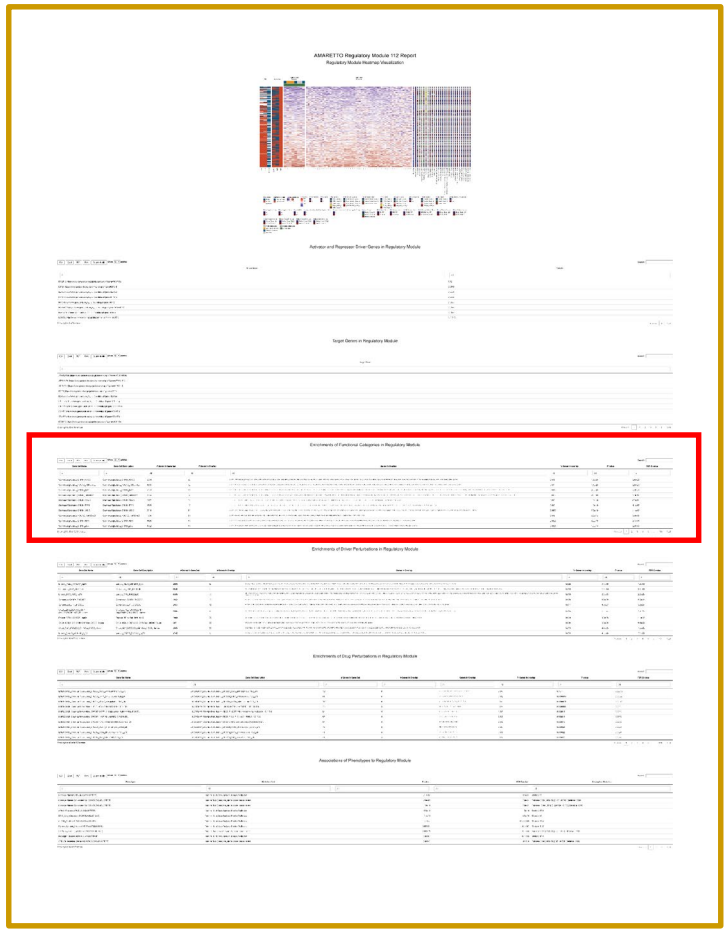
Enrichments of Functional Categories in Regulatory Module

CSV | Excel | PDF | Print | Column visibility Show 10 entries Search: MYC

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
MYC	All						
StemnessSignatures_ORKIN_MYC	StemnessSignatures_ORKIN_MYC	355	9	COX8A, DDB1, GTF3C4, INCENP, IPO7, MRPL49, PRPF19, XPO5, ZNHIT2	0.025	1.4e-7	0.0000012
WEI MYCN TARGETS WITH E BOX	Genes whose promoters contain E-box motifs and whose expression changed in MYCN-3 cells (neuroblastoma) upon induction of MYCN [GeneID=4613].	795	11	AHCTF1, C11orf83, CDCA5, GTF3C4, MTCH2, NSUN2, PRMT3, QSER1, RNF219, SAAL1, TIMM10	0.014	0.0000022	0.00014
StemnessSignatures_WEINBERG_MYC_MAX_TARGETS	StemnessSignatures_WEINBERG_MYC_MAX_TARGETS	775	8	DDB1, SNX15, GTF3C4, BAZ1B, UBXN1, HNRNPL, ARFIP2, CSTF3	0.010	0.00042	0.0015
DANG BOUND BY MYC	Genes whose promoters are bound by MYC [GeneID=4609], according to MYC Target Gene Database.	1103	11	ARFGAP2, ARFIP2, BAZ1B, CLP1, CSTF3, DDB1, GTF3C4, MEN1, NAT10, TIMM10, UBXN1	0.010	0.000046	0.0015
SCHLOSSER SERUM RESPONSE AUGMENTED BY MYC	Cluster 2: genes up-regulated in B493-6 cells (B lymphocytes) by serum alone or in combination with MYC [GeneID=4609] but not by MYC alone.	108	4	KAT5, OTUB1, PRPF19, TAF6L	0.037	0.00011	0.0030
BENPORATH MYC MAX TARGETS	Set 'Myc targets2': targets of c-Myc [GeneID=4609] and Max [GeneID=4149] identified by ChIP on chip in a Burkitt's lymphoma cell line; overlap set.	775	8	ARFIP2, BAZ1B, CSTF3, DDB1, GTF3C4, HNRNPL, SNX15, UBXN1	0.010	0.00042	0.0075
BILD MYC ONCOGENIC SIGNATURE	Genes selected in supervised analyses to discriminate cells expressing c-Myc [GeneID=4609] from control cells expressing GFP.	206	4	C11orf48, SNHG1, WDR74, XPO5	0.019	0.0013	0.016
HALLMARK MYC TARGETS V2	A subgroup of genes regulated by MYC - version 2 (v2).	58	2	PRMT3, WDR74	0.035	0.0079	0.050

Showing 1 to 8 of 8 entries (filtered from 1,137 total entries) Previous 1 Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
functional characterization

Enrichments of Functional Categories in Regulatory Module

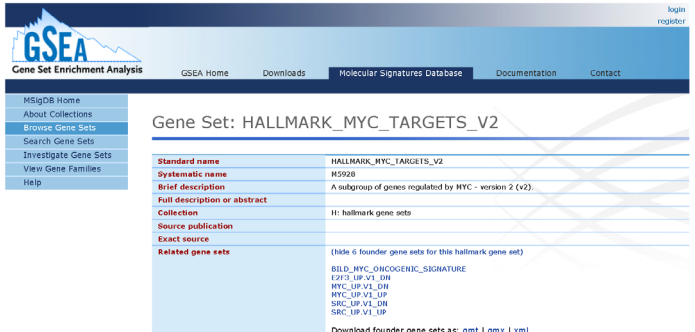
CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search: MYC

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
MYC	All						
StemnessSignatures_ORKIN_MYC	StemnessSignatures_ORKIN_MYC	355	9	COX8A, DDB1, GTF3C4, INCENP, IPO7, MRPL49, PRPF19, XPO5, ZNHIT2	0.025	1.4e-7	0.0000012
WEI MYCN TARGETS WITH E BOX	Genes whose promoters contain E-box motifs and whose expression changed in MYCN-3 cells (neuroblastoma) upon induction of MYCN [GeneID=4613].	795	11	AHCTF1, C11orf83, CDCA5, GTF3C4, MTCH2, NSUN2, PRMT3, QSER1, RNF219, SAAL1, TIMM10	0.014	0.0000022	0.00014
StemnessSignatures_WEINBERG_MYC_MAX_TARGETS	StemnessSignatures_WEINBERG_MYC_MAX_TARGETS	775	8	DDB1, SNX15, GTF3C4, BAZ1B, UBXN1, HNRNPL, ARFIP2, CSTF3	0.010	0.00042	0.0015
DANG BOUND BY MYC	Genes whose promoters are bound by MYC [GeneID=4609], according to MYC Target Gene Database.	1103	11	ARFGAP2, ARFIP2, BAZ1B, CLP1, CSTF3, DDB1, GTF3C4, MEN1, NAT10, TIMM10, UBXN1	0.010	0.000046	0.0015
SCHLOSSER SERUM RESPONSE AUGMENTED BY MYC	Cluster 2: genes up-regulated in B493-6 cells (B lymphocytes) by serum alone or in combination with MYC [GeneID=4609] but not by MYC alone.	108	4	KAT5, OTUB1, PRPF19, TAF6L	0.037	0.00011	0.0030
BENPORATH MYC MAX TARGETS	Set 'Myc targets2': targets of c-Myc [GeneID=4609] and Max [GeneID=4149] identified by ChIP on chip in a Burkitt's lymphoma cell line; overlap set.	775	8	ARFIP2, BAZ1B, CSTF3, DDB1, GTF3C4, HNRNPL, SNX15, UBXN1	0.010	0.00042	0.0075
BILD MYC ONCOGENIC SIGNATURE	Genes selected in supervised analyses to discriminate cells expressing c-Myc [GeneID=4609] from control cells expressing GFP.	206	4	C11orf48, SNHG1, WDR74, XPO5	0.019	0.0013	0.016
HALLMARK MYC TARGETS V2	A subgroup of genes regulated by MYC - version 2 (v2).	58	2	PRMT3, WDR74	0.035	0.0079	0.050

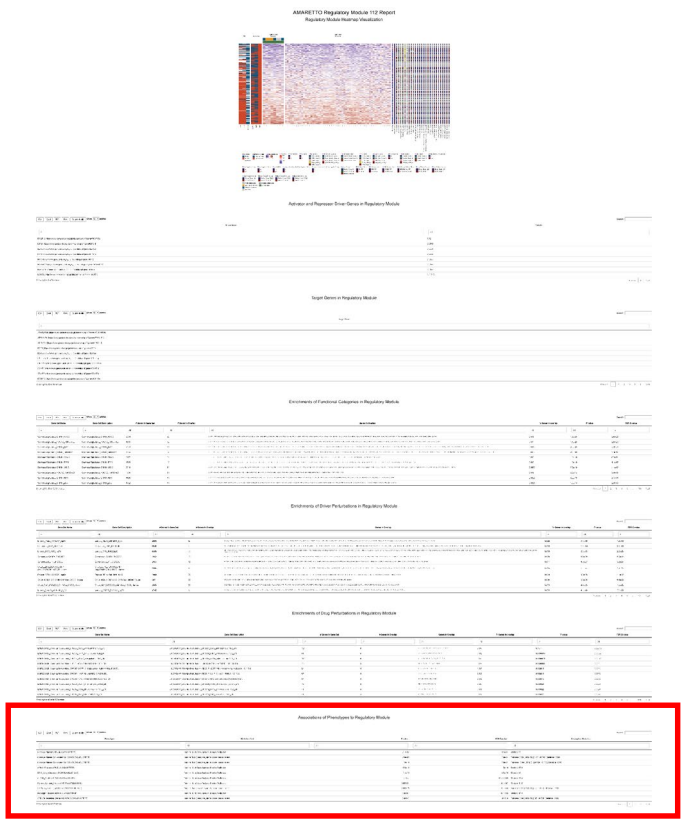
Showing 1 to 8 of 8 entries (filtered from 1,137 total entries) Previous 1 Next

Filter for significant results & Search for MYC gene signatures

Link to MSigDB description of Hallmark MYC targets:



AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
clinical characterization

Associations of Phenotypes to Regulatory Module

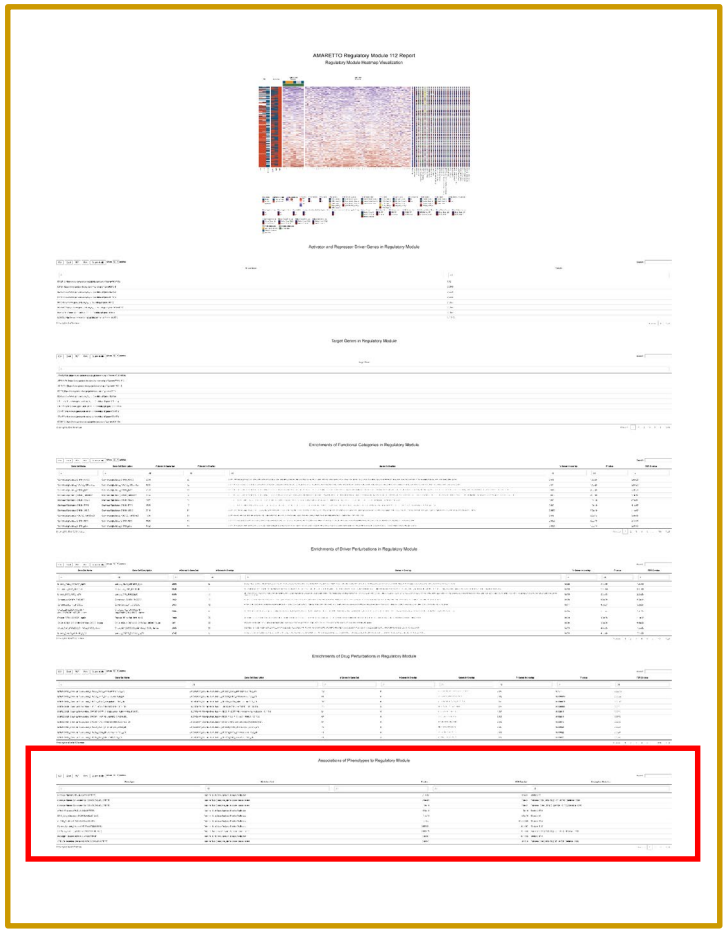
CSV | Excel | PDF | Print | Column visibility | Show 20 entries | Search:

Phenotype	Statistics Test	P-value	FDR Q-value	Descriptive Statistics
All	All	0.000000	All	All
Hoshida_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	3.1e-12	1.1e-11	Statistic: 53
Hoshida_Cluster_S2_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	6.9e-12	3.0e-11	Estimate: 0.581, 95% CI: [0.423 , 0.726], Statistics: 6820
Hoshida_Cluster_S3_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	1.5e-11	6.5e-11	Estimate: -0.564, 95% CI: [-0.709 , -0.407], Statistics: 1790
mRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	6.8e-11	1.8e-10	Statistic: 53.5
DNA_Copy_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	1.1e-11	5.6e-10	Statistic: 54
miRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	9.7e-7	0.0000028	Statistic: 33.4
Hypomethylation_Cluster (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.000069	0.00041	Statistic: 19.2
CDKN2A_Silencing (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.000021	0.00078	Estimate: 0.374, 95% CI: [0.207 , 0.54], Statistics: 5800
Paradigm_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.00060	0.00085	Statistic: 17.4
CTNNB1_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00040	0.0016	Estimate: 0.362, 95% CI: [0.17 , 0.554], Statistics: 4360
TERT_Promoter_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00011	0.0041	Estimate: 0.33, 95% CI: [0.169 , 0.502], Statistics: 5620
TP53_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0028	0.013	Estimate: 0.3, 95% CI: [0.0977 , 0.496], Statistics: 4620
Consensus_Clinical_and_RNA_Seq_Hepatitis_B_Virus (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0014	0.019	Estimate: 0.323, 95% CI: [0.129 , 0.517], Statistics: 4150
iCluster_Clusters_1_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.027	0.046	Estimate: 0.213, 95% CI: [0.0224 , 0.386], Statistics: 4680
RPPA_Clusters (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.031	0.054	Estimate: -0.208, 95% CI: [-0.402 , -0.0173], Statistics: 2370
SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013	0.14	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34

Showing 1 to 16 of 16 entries (filtered from 27 total entries)

Previous 1 Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
clinical characterization

Clinical and molecular phenotypes from TCGA

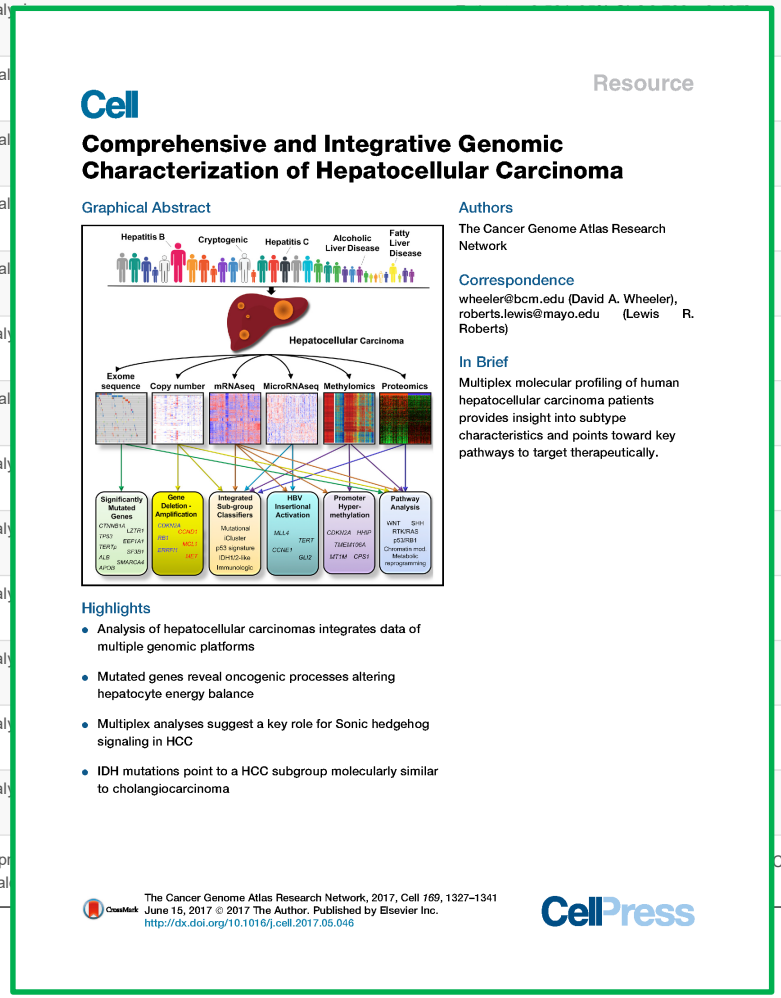
Associations of Phenotypes to Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 20 entries | Search:

Phenotype	Statistics Test	P-value	FDR Q-value	Descriptive Statistics
Hoshida_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	3.1e-12	1.1e-11	Statistic:53
Hoshida_Cluster_S2_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	6.9e-12	3.0e-11	Estimate: 0.581, 95% CI: [0.423 , 0.726], Statistics: 6820
Hoshida_Cluster_S3_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
mRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Anal Kruskal-Wallis test			
DNA_Copy_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Anal Kruskal-Wallis test			
miRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Anal Kruskal-Wallis test			
Hypomethylation_Cluster (KRUSKALWALLISTEST)	Nominal Multi-Class Anal Kruskal-Wallis test			
CDKN2A_Silencing (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
Paradigm_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Anal Kruskal-Wallis test			
CTNNB1_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
TERT_Promoter_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
TP53_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
Consensus_Clinical_and_RNA_Seq_Hepatitis_B_Virus (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
iCluster_Clusters_1_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
RPPA_Clusters (WILCOXONRANKSUMTEST)	Nominal Two-Class Analy Wilcoxon rank sum test			
SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox pr hazards regression (Wal			CI:

Showing 1 to 16 of 16 entries (filtered from 27 total entries)

Next



Resource

Cell

Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma

Graphical Abstract

Authors
The Cancer Genome Atlas Research Network

Correspondence
wheeler@bcm.edu (David A. Wheeler), roberts.lewis@mayo.edu (Lewis R. Roberts)

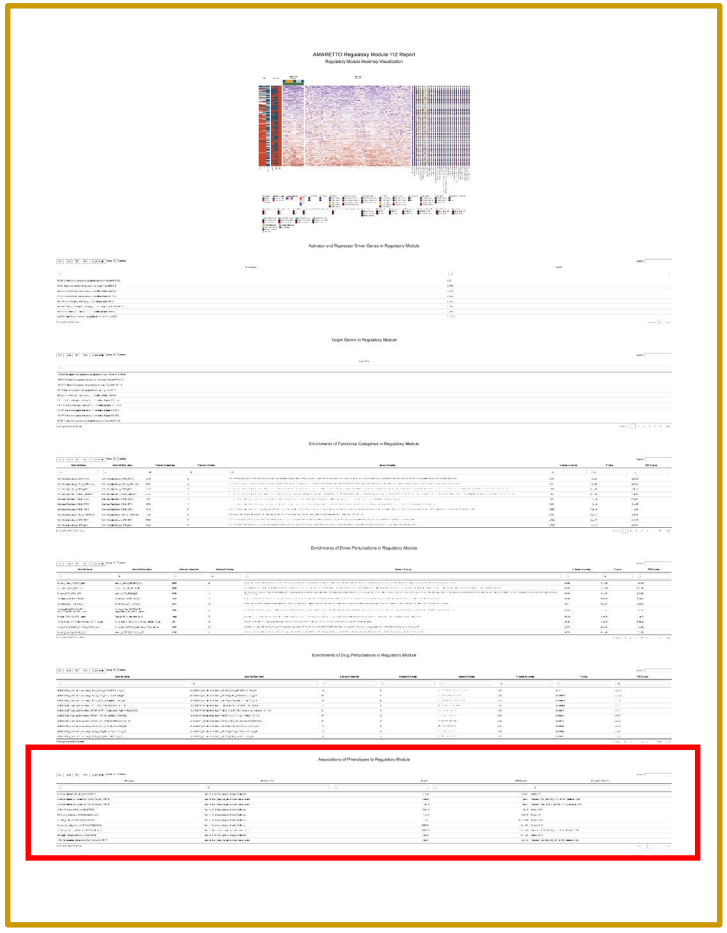
In Brief
Multiplex molecular profiling of human hepatocellular carcinoma patients provides insight into subtype characteristics and points toward key pathways to target therapeutically.

Highlights
- Analysis of hepatocellular carcinomas integrates data of multiple genomic platforms
- Mutated genes reveal oncogenic processes altering hepatocyte energy balance
- Multiplex analyses suggest a key role for Sonic hedgehog signaling in HCC
- IDH mutations point to a HCC subgroup molecularly similar to cholangiocarcinoma

The Cancer Genome Atlas Research Network, 2017, Cell 169, 1327–1341
June 15, 2017 © 2017 The Author. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2017.05.046

CellPress

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
clinical characterization

Associations of Phenotypes to Regulatory Module

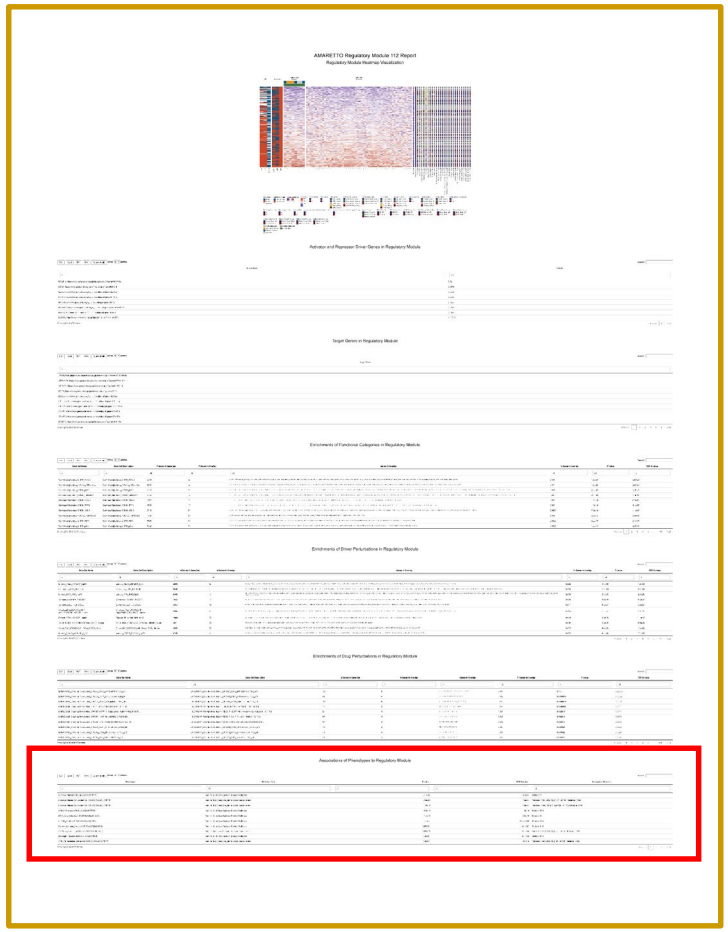
CSV | Excel | PDF | Print | Column visibility | Show 20 entries | Search:

Phenotype	Statistics Test	P-value	FDR Q-value	Descriptive Statistics
All	All	0.000000	All	All
Hoshida_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	3.1e-12	1.1e-11	Statistic: 53
Hoshida_Cluster_S2_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	6.9e-12	3.0e-11	Estimate: 0.581, 95% CI: [0.423 , 0.726], Statistics: 6820
Hoshida_Cluster_S3_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	1.5e-11	6.5e-11	Estimate: -0.564, 95% CI: [-0.709 , -0.407], Statistics: 1790
mRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	6.8e-11	1.8e-10	Statistic: 53.5
DNA_Copy_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	1.1e-11	5.6e-10	Statistic: 54
miRNA_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	9.7e-7	0.0000028	Statistic: 33.4
Hypomethylation_Cluster (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.000069	0.00041	Statistic: 19.2
CDKN2A_Silencing (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.000021	0.00078	Estimate: 0.374, 95% CI: [0.207 , 0.54], Statistics: 5800
Paradigm_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.00060	0.00085	Statistic: 17.4
CTNNB1_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00040	0.0016	Estimate: 0.362, 95% CI: [0.17 , 0.554], Statistics: 4360
TERT_Promoter_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00011	0.0041	Estimate: 0.33, 95% CI: [0.169 , 0.502], Statistics: 5620
TP53_Mutation_Status (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0028	0.013	Estimate: 0.3, 95% CI: [0.0977 , 0.496], Statistics: 4620
Consensus_Clinical_and_RNA_Seq_Hepatitis_B_Virus (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0014	0.019	Estimate: 0.323, 95% CI: [0.129 , 0.517], Statistics: 4150
iCluster_Clusters_1_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.027	0.046	Estimate: 0.213, 95% CI: [0.0224 , 0.386], Statistics: 4680
RPPA_Clusters (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.031	0.054	Estimate: -0.208, 95% CI: [-0.402 , -0.0173], Statistics: 2370
SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013	0.14	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34

Showing 1 to 16 of 16 entries (filtered from 27 total entries)

Previous 1 Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
clinical characterization

Associations of Phenotypes to Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 20 entries | Search:

Phenotype	Statistics Test	P-value	FDR Q-value	Descriptive Statistics
Hoshida_Clusters (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	3.1e-12	1.1e-11	Statistic:53
Hoshida_Cluster_S2_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	6.9e-12	3.0e-11	Estimate: 0.581, 95% CI: [0.423 , 0.726], Statistics: 6820
Hoshida_Cluster_S3_vs_rest (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	1.5e-11	6.5e-11	Estimate: -0.564, 95% CI: [-0.709 , -0.407], Statistics: 1790
mRNA_Clusters (KRUSKALWALLISTEST)				
DNA_Copy_Clusters (KRUSKALWALLISTEST				
miRNA_Clusters (KRUSKALWALLISTEST)				
Hypomethylation_Cluster (KRUSKALWALLIST				
CDKN2A_Silencing (WILCOXONRANKSUMTE				
Paradigm_Clusters (KRUSKALWALLISTEST)				
CTNNB1_Mutation_Status (WILCOXONRANK				
TERT_Promoter_Mutation_Status (WILCOXO				
TP53_Mutation_Status (WILCOXONRANKSU				
Consensus_Clinical_and_RNA_Seq_Hepatitis_ (WILCOXONRANKSUMTEST)				
iCluster_Clusters_1_vs_rest (WILCOXONRAN				
RPPA_Clusters (WILCOXONRANKSUMTEST				
SurvivalTime (COXPROPHAZARDTIMETOEV SurvivalCensoring (COXPROPHAZARDRIGH				

Showing 1 to 16 of 16 entries (filtered from 27 to

GSEA Gene Set Enrichment Analysis — login / register

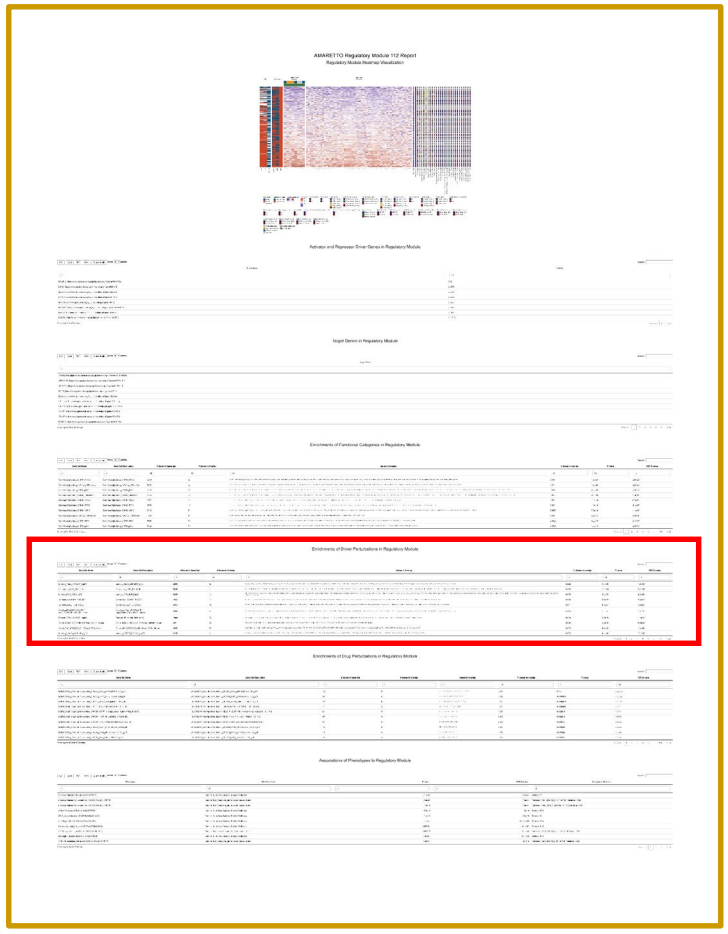
GSEA Home | Downloads | Molecular Signatures Database | Documentation | Contact

- MSigDB Home
- About Collections
- Browse Gene Sets
- Search Gene Sets
- Investigate Gene Sets
- View Gene Families
- Help

Gene Set: HOSHIDA_LIVER_CANCER_SUBCLASS_S2

Standard name	HOSHIDA_LIVER_CANCER_SUBCLASS_S2
Systematic name	M7995
Brief description	Genes from 'subtype S2' signature of hepatocellular carcinoma (HCC): proliferation, MYC and AKT1 [GeneID=4609;207] activation.
Full description or abstract	Hepatocellular carcinoma (HCC) is a highly heterogeneous disease, and prior attempts to develop genomic-based classification for HCC have yielded highly divergent results, indicating difficulty in identifying unified molecular anatomy. We performed a meta-analysis of gene expression profiles in data sets from eight independent patient cohorts across the world. In addition, aiming to establish the real world applicability of a classification system, we profiled 118 formalin-fixed, paraffin-embedded tissues from an additional patient cohort. A total of 603 patients were analyzed, representing the major etiologies of HCC (hepatitis B and C) collected from Western and Eastern countries. We observed three robust HCC subclasses (termed S1, S2, and S3), each correlated with clinical parameters such as tumor size, extent of cellular differentiation, and serum alpha-fetoprotein levels. An analysis of the components of the signatures indicated that S1 reflected aberrant activation of the WNT signaling pathway, S2 was characterized by proliferation as well as MYC and AKT activation, and S3 was associated with hepatocyte differentiation. Functional studies indicated that the WNT pathway activation signature characteristic of S1 tumors was not simply the result of beta-catenin mutation but rather was the result of transforming growth factor-beta activation, thus representing a new mechanism of WNT pathway activation in HCC. These experiments establish the first consensus classification framework for HCC based on gene expression profiles and highlight the power of integrating multiple data sets to define a robust molecular taxonomy of the disease. [Cancer Res 2009;69(18):7385-92].
Collection	C2: curated gene sets CGP: chemical and genetic perturbations
Source publication	Pubmed 19723656 Authors: Hoshida Y,Nijman SM,Kobayashi M,Chan JA,Brunet JP,Chiang DY,Villanueva A,Newell P,Ikeda K,Hashimoto M,Watanabe G,Gabriel S,Friedman SL,Kumada H,Llovet JM,Golub TR
Exact source	Table 3S: Subtype=S2
Related gene sets	(show 2 additional gene sets from the source publication) (show 300 gene sets from the same authors)
External links	
Organism	Homo sapiens
Contributed by	Jessica Robertson (Broad Institute)
Source platform	EntrezGeneIds
Dataset references	
Download gene set	format: grp | text | gmt | gmx | xml
Compute overlaps	(show collections to investigate for overlap with this gene set)
Compendia expression profiles	Human tissue compendium (Novartis)

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
driver validation & discovery

➢ Perturbation-AMARETTO v1

Enrichments of Driver Perturbations in Regulatory Module

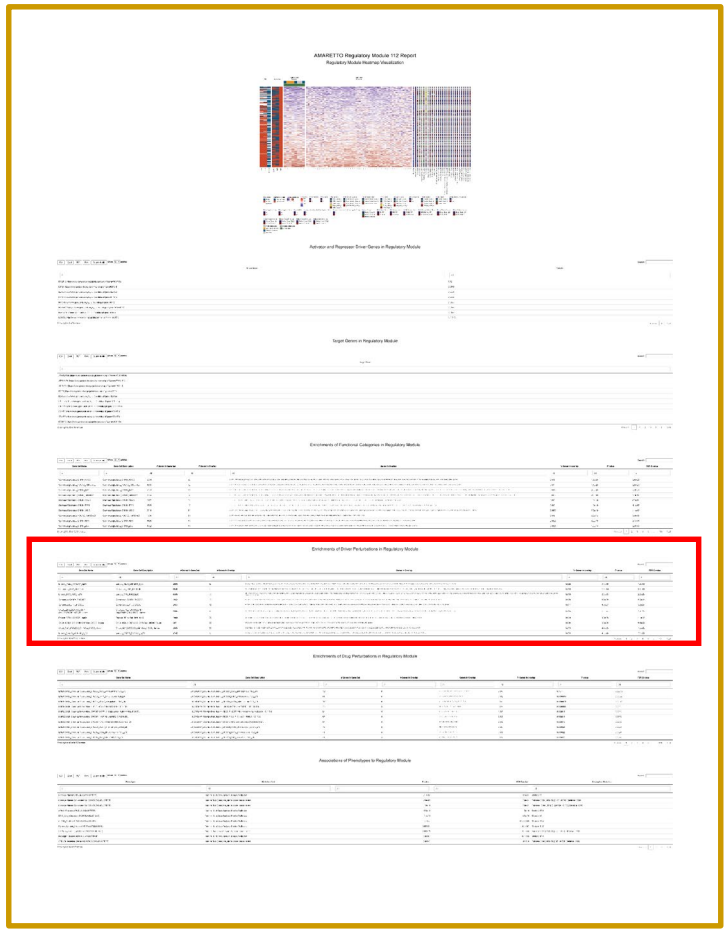
CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search: MYC

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
All	All			All			
Encode_MYC_K562_hg19	Encode_MYC_K562_hg19	6800	70	SF1, MTCH2, CWF19L2, MRPS17, RRP8, CPSF7, VPS37C, BANF1, EIF4G2, ZFP91, MAP4K2, PRMT3, FAU, DNAJC24, HNRNPA1L2, CDCA5, EIF1AD, EIF3M, ATF7IP, SSRP1, TAF6L, KAT5, FTSJ2, GTF3C4, TMEM41B, CLP1, MAD2L1BP, METTL12, CLPTM1L, FEN1, MARK2, TRIM41, COX8A, NUP188, POLA2, WDR74, POM121C, IPO7, PDHX, CSTF3, AHCTF1, NSUN2, UTP3, MGA, INTS5, ZDHHC5, SAAL1, SNHG1, PRPF19, BAZ1B, RNF219, INCENP, DDB1, NAT10, HNRNPL, ZNHIT2, KBTBD4, XPO5, CAPRIN1, KDM3B, PSMC3, TUT1, MRPL49, HNRNPUL2, PDS5B, PDS5A, NDUFS3, TIMM10, CKAP5, ZNF195	0.010	2.2e-33	2.2e-29
Encode_MYC_HeLa-S3_hg19	Encode_MYC_HeLa-S3_hg19	3080	43	MTCH2, MRPS17, RRP8, IPO7, ARFIP2, CSTF3, UBXN1, NSUN2, UTP3, MGA, INTS5, ZDHHC5, EIF4G2, ZFP91, SNHG1, PRPF19, PRMT3, FAU, DNAJC24, BAZ1B, HNRNPA1L2, INCENP, DDB1, EIF3M, NAT10, QSER1, ATF7IP, SSRP1, ZNHIT2, KAT5, XPO5, KDM3B, PSMC3, TUT1, MRPL49, GTF3C4, HNRNPUL2, PDS5A, TMEM41B, CLP1, ARFGAP2, FEN1, ZNF195	0.014	2.3e-23	4.6e-20
Encode_MYC_MCF-7_hg19	Encode_MYC_MCF-7_hg19	5003	50	SF1, SDHAF2, TRIM41, COX8A, MRPS17, NUP188, POLA2, RRP8, POM121C, IPO7, ARFIP2, CPSF7, VPS37C, UTP3, MGA, INTS5, ZDHHC5, SAAL1, PRPF19, FAU, DNAJC24, BAZ1B, HNRNPA1L2, CDCA5, DDB1, ZNF431, EIF1AD, TBC1D14, ATF7IP, SSRP1, ZNHIT2, KAT5, KBTBD4, KDM3B, TMEM33, TUT1, MRPL49, FTSJ2, GANAB, HNRNPUL2, PDS5B, PDS5A, NDUFS3, SF3B2, TIMM10, ZNF107, METTL12, ARFGAP2, CKAP5, ZNF195	0.010	3.4e-21	4.2e-18
Encode_MYC_GM12878_hg19	Encode_MYC_GM12878_hg19	2000	31	BAZ1B, HNRNPA1L2, RNF219, DDB1, EIF1AD, COX8A, MRPS17, NUP188, OTUB1, RRP8, ARFIP2, KAT5, KBTBD4, KDM3B, TUT1, UTP3, MRPL49, FTSJ2, MGA, INTS5, HNRNPUL2, PDS5A, NDUFS3, CLP1, SF3B2, TIMM10, SAAL1, FEN1, FAU, DNAJC24, ZNF195	0.015	8.7e-18	5.0e-15
ChEA_MYC_18358816_ChIP-ChIP_MESCs_Mouse	ChEA_MYC_18358816_ChIP-ChIP_MESCs_Mouse	3413	38	COX8A, NUP188, POLA2, WDR74, MTA2, IPO7, CSTF3, AHCTF1, DDAH1, NSUN2, MGA, INTS5, BANF1, TGFBRAP1, PCNXL3, MEN1, PRPF19, MAP4K2, FAU, FIZ1, CDCA5, INCENP, DDB1, NAT10, ATF7IP, OTUB1, ZNHIT2, XPO5, PSMC3, TMEM33, TUT1, MRPL49, GTF3C4, GANAB, TMEM41B, MAD2L1BP, TIMM10, FEN1	0.011	4.0e-17	2.0e-14
ChEA_MYC_19030024_ChIP-ChIP_MESCs_Mouse	ChEA_MYC_19030024_ChIP-ChIP_MESCs_Mouse	3868	40	MTCH2, COX8A, CWF19L2, NUP188, POLA2, WDR74, IPO7, ARFIP2, CSTF3, AHCTF1, DDAH1, NSUN2, MGA, ZDHHC5, BANF1, EIF4G2, ZFP91, PRPF19, PRMT3, FAU, FIZ1, BAZ1B, CDCA5, INCENP, DDB1, NAT10, ATF7IP, SSRP1, OTUB1, TAF6L, XPO5, CAPRIN1, TMEM33, GTF3C4, NDUFS3, TMEM41B, MAD2L1BP, TIMM10, CLPTM1L, FEN1	0.010	5.5e-17	2.7e-14
Consensus_MYC_ENCODE	Consensus_MYC_ENCODE	1515	24	BAZ1B, NSUN2, TUT1, UTP3, DDB1, EIF1AD, GTF3C4, EIF3M, MGA, NAT10, MRPS17, HNRNPUL2, BANF1, WDR74, HNRNPL, CLP1, RRP8, IPO7, ZFP91, SNHG1, CSTF3, PRMT3, XPO5, CAPRIN1	0.016	5.3e-14	5.9e-12
ChEA_MYC_18555785_ChIP-Seq_MESCs_Mouse	ChEA_MYC_18555785_ChIP-Seq_MESCs_Mouse	1200	20	SF1, CDCA5, INCENP, DDB1, MTCH2, COX8A, NAT10, BANF1, NDUFS3, TGFBRAP1, WDR74, SF3B2, TAF6L, EIF4G2, CSTF3, PRPF19, PRMT3, XPO5, FAU, FIZ1	0.017	3.6e-12	6.3e-10
ChEA_MYCN_18555785_ChIP-Seq_MESCs_Mouse	ChEA_MYCN_18555785_ChIP-Seq_MESCs_Mouse	2261	25	DDB1, MTCH2, CWF19L2, NUP188, POLA2, SSRP1, OTUB1, IPO7, KBTBD4, XPO5, DDAH1, NSUN2, MRPL49, FTSJ2, GTF3C4, GANAB, ZDHHC5, TMEM41B, TIMM10, DGKZ, EIF4G2, CLPTM1L, FEN1, FIZ1, MARK2	0.011	3.7e-11	5.0e-9
Encode_MYC_MCF 10A_hg19	Encode_MYC_MCF 10A_hg19	3382	29	SDHAF2, MTCH2, MRPS17, POM121C, ARFIP2, CPSF7, VPS37C, INTS5, SAAL1, EIF4G2, PRPF19, FIZ1, HNRNPA1L2, CDCA5, RNF219, EIF3M, TMEM223, SNRNP35, KAT5, KBTBD4, TUT1, FTSJ2, GTF3C4, GANAB, PDS5B, NDUFS3, TMEM41B, ZNF107, METTL12	0.0086	3.1e-10	1.7e-8

Showing 1 to 10 of 24 entries (filtered from 7,061 total entries)

Previous | 1 | 2 | 3 | Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
driver validation & discovery

➢ Perturbation-AMARETTO v1

Enrichments of Driver Perturbations in Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search: MYC

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
All	All			All			
Encode_MYC_K562_hg19	Encode_MYC_K562_hg19	6800	70	SF1, MTCH2, CWF19L2, MRPS17, RRP8, CPSF7, VPS37C, BANF1, EIF4G2, ZFP91, MAP4K2, PRMT3, FAU, DNAJC24, HNRNPA1L2, CDCA5, EIF1AD, EIF3M, ATF7IP, SSRP1, TAF6L, KAT5, FTSJ2, GTF3C4, TMEM41B, CLP1, MAD2L1BP, METTL12, CLPTM1L, FEN1, MARK2, TRIM41, COX8A, NUP188, POLA2, WDR74, POM121C, IPO7, PDHX, CSTF3, AHCTF1, NSUN2, UTP3, MGA, INTS5, ZDHHC5, SAAL1, SNHG1, PRPF19, BAZ1B, RNF219, INCENP, DDB1, NAT10, HNRNPL, ZNHIT2, KBTBD4, XPO5, CAPRIN1, KDM3B, PSMC3, TUT1, MRPL49, HNRNPUL2, PDS5B, PDS5A, NDUFS3, TIMM10, CKAP5, ZNF195	0.010	2.2e-33	2.2e-29
Encode_MYC_HeLa-S3_hg19	Encode_MYC_HeLa-S3_hg19	3080	43	MTCH2, MRPS17, RRP8, IPO7, ARFIP2, CSTF3, UBXN1, NSUN2, UTP3, MGA, INTS5, ZDHHC5, EIF4G2, ZFP91, SNHG1, PRPF19, PRMT3, FAU, DNAJC24, BAZ1B, HNRNPA1L2, INCENP, DDB1, EIF3M, NAT10, QSER1, ATF7IP, SSRP1, ZNHIT2, KAT5, XPO5, KDM3B, PSMC3, TUT1, MRPL49, GTF3C4, HNRNPUL2, PDS5A, TMEM41B, CLP1, ARFGAP2, FEN1, ZNF195	0.014	2.3e-23	4.6e-20
Encode_MYC_MCF-7_hg19	Encode_MYC_MCF-7_hg19	5003	50	SF1, SDHAF2, TRIM41, COX8A, MRPS17, NUP188, POLA2, RRP8, POM121C, IPO7, ARFIP2, CPSF7, VPS37C, UTP3, MGA, INTS5, ZDHHC5, SAAL1, PRPF19, FAU, DNAJC24, BAZ1B, HNRNPA1L2, CDCA5, DDB1, ZNF431, EIF1AD, TBC1D14, ATF7IP, SSRP1, ZNHIT2, KAT5, KBTBD4, KDM3B, TMEM33, TUT1, MRPL49, FTSJ2, GANAB, HNRNPUL2, PDS5B, PDS5A, NDUFS3, SF3B2, TIMM10, ZNF107, METTL12, ARFGAP2, CKAP5, ZNF195	0.010	3.4e-21	4.2e-18
Encode_MYC_GM12878_hg19	Encode_MYC_GM12878_hg19	2000	31	BAZ1B, HNRNPA1L2, RNF219, DDB1, EIF1AD, COX8A, MRPS17, NUP188, OTUB1, RRP8, ARFIP2, KAT5, KBTBD4, KDM3B, TUT1, UTP3, MRPL49, FTSJ2, MGA, INTS5, HNRNPUL2, PDS5A, NDUFS3, CLP1, SF3B2, TIMM10, SAAL1, FEN1, FAU, DNAJC24, ZNF195	0.015	8.7e-18	5.0e-15
ChEA_MYC_18358816_ChIP-ChIP_MESCs_Mouse	ChEA_MYC_18358816_ChIP-ChIP_MESCs_Mouse	3413	38	COX8A, NUP188, POLA2, WDR74, MTA2, IPO7, CSTF3, AHCTF1, DDAH1, NSUN2, MGA, INTS5, BANF1, TGFBRAP1, PCNXL3, MEN1, PRPF19, MAP4K2, FAU, FIZ1, CDCA5, INCENP, DDB1, NAT10, ATF7IP, OTUB1, ZNHIT2, XPO5, PSMC3, TMEM33, TUT1, MRPL49, GTF3C4, GANAB, TMEM41B, MAD2L1BP, TIMM10, FEN1	0.011	4.0e-17	2.0e-14
ChEA_MYC_19030024_ChIP-ChIP_MESCs_Mouse	ChEA_MYC_19030024_ChIP-ChIP_MESCs_Mouse	3868	40	MTCH2, COX8A, CWF19L2, NUP188, POLA2, WDR74, IPO7, ARFIP2, CSTF3, AHCTF1, DDAH1, NSUN2, MGA, ZDHHC5, BANF1, EIF4G2, ZFP91, PRPF19, PRMT3, FAU, FIZ1, BAZ1B, CDCA5, INCENP, DDB1, NAT10, ATF7IP, SSRP1, OTUB1, TAF6L, XPO5, CAPRIN1, TMEM33, GTF3C4, NDUFS3, TMEM41B, MAD2L1BP, TIMM10, CLPTM1L, FEN1	0.010	5.5e-17	2.7e-14
Consensus_MYC_ENCODE	Consensus_MYC_ENCODE	1515	24	BAZ1B, NSUN2, TUT1, UTP3, DDB1, EIF1AD, GTF3C4, EIF3M, MGA, NAT10, MRPS17, HNRNPUL2, BANF1, WDR74, HNRNPL, CLP1, RRP8, IPO7, ZFP91, SNHG1, CSTF3, PRMT3, XPO5, CAPRIN1	0.016	5.3e-14	5.9e-12
ChEA_MYC_18555785_ChIP-Seq_MESCs_Mouse	ChEA_MYC_18555785_ChIP-Seq_MESCs_Mouse	1200	20	SF1, CDCA5, INCENP, DDB1, MTCH2, COX8A, NAT10, BANF1, NDUFS3, TGFBRAP1, WDR74, SF3B2, TAF6L, EIF4G2, CSTF3, PRPF19, PRMT3, XPO5, FAU, FIZ1	0.017	3.6e-12	6.3e-10
ChEA_MYCN_18555785_ChIP-Seq_MESCs_Mouse	ChEA_MYCN_18555785_ChIP-Seq_MESCs_Mouse	2261	25	DDB1, MTCH2, CWF19L2, NUP188, POLA2, SSRP1, OTUB1, IPO7, KBTBD4, XPO5, DDAH1, NSUN2, MRPL49, FTSJ2, GTF3C4, GANAB, ZDHHC5, TMEM41B, TIMM10, DGKZ, EIF4G2, CLPTM1L, FEN1, FIZ1, MARK2	0.011	3.7e-11	5.0e-9
Encode_MYC_MCF 10A_hg19	Encode_MYC_MCF 10A_hg19	3382	29	SDHAF2, MTCH2, MRPS17, POM121C, ARFIP2, CPSF7, VPS37C, INTS5, SAAL1, EIF4G2, PRPF19, FIZ1, HNRNPA1L2, CDCA5, RNF219, EIF3M, TMEM223, SNRNP35, KAT5, KBTBD4, TUT1, FTSJ2, GTF3C4, GANAB, PDS5B, NDUFS3, TMEM41B, ZNF107, METTL12	0.0086	3.1e-10	1.7e-8

Showing 1 to 10 of 24 entries (filtered from 7,061 total entries)

Previous | 1 | 2 | 3 | Next

Experiments validating MYC and MYC-regulated genes in Module 112:
Encode and ChEA ChIP-Seq experiments

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: driver validation & discovery using genetic perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Driver discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility (https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/#) | Show 500 entries | Search:

PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	LIHC
All	All	["MYC","BZW2","E2F5","EI	All	All	All	escore-pval-padj
CGS001_A375_96H:BZW2:1	A375	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_A375_96H:MYC:1	A375	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A375_96H:NPM1:1	A375	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
OEB005_A375_96H:BRDN0000408975:-666	A375	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_A375_96H:EIF3H:1	A375	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_A549_96H:EIF3H:1	A549	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_A549_96H:MYC:1	A549	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A549_96H:E2F5:1	A549	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj
CGS001_HA1E_96H:MYC:1.5	HA1E	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HA1E_96H:EIF3H:1.5	HA1E	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_HA1E_96H:BZW2:1.5	HA1E	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HCC515_96H:MYC:2	HCC515	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:MYC:1.5	HEPG2	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:BZW2:1.5	HEPG2	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
OEB005_HEPG2_96H:BRDN0000408975:-666	HEPG2	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_HT29_96H:BZW2:1	HT29	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HT29_96H:MYC:1	HT29	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_HT29_96H:E2F5:1	HT29	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj-zscore
CGS001_MCF7_144H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_144H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_96H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_96H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_144H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_MCF7_96H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_MCF7_96H:EIF3H:2	MCF7	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_NPC_96H:BZW2:1.5	NPC	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_NPC_96H:MYC:1.5	NPC	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_PC3_96H:MYC:2	PC3	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
OEB003_PC3_96H:BRDN0000405602:-666	PC3	BZW2	28969	trt_oe	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_PC3_96H:EIF3H:2	PC3	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_VCAP_120H:NPM1:5	VCAP	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_VCAP_120H:MYC:5	VCAP	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_VCAP_120H:EIF3H:5	VCAP	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore

Showing 1 to 33 of 33 entries (filtered from 55,753 total entries) Previous 1 Next

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: driver validation & discovery using genetic perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Driver discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility (https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/#) | Show 500 entries | Search:

PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	LIHC
All	All	["MYC","BZW2","E2F5","EI	All	All	All	escore-pval-padj
CGS001_A375_96H:BZW2:1	A375	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_A375_96H:MYC:1	A375	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A375_96H:NPM1:1	A375	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
OEB005_A375_96H:BRDN0000408975:-666	A375	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_A375_96H:EIF3H:1	A375	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_A549_96H:EIF3H:1	A549	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_A549_96H:MYC:1	A549	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A549_96H:E2F5:1	A549	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj
CGS001_HA1E_96H:MYC:1.5	HA1E	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HA1E_96H:EIF3H:1.5	HA1E	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_HA1E_96H:BZW2:1.5	HA1E	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HCC515_96H:MYC:2	HCC515	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:MYC:1.5	HEPG2	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:BZW2:1.5	HEPG2	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
OEB005_HEPG2_96H:BRDN0000408975:-666	HEPG2	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_HT29_96H:BZW2:1	HT29	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HT29_96H:MYC:1	HT29	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_HT29_96H:E2F5:1	HT29	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj-zscore
CGS001_MCF7_144H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_144H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_96H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_96H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_144H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_MCF7_96H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_MCF7_96H:EIF3H:2	MCF7	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_NPC_96H:BZW2:1.5	NPC	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_NPC_96H:MYC:1.5	NPC	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_PC3_96H:MYC:2	PC3	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
OEB003_PC3_96H:BRDN0000405602:-666	PC3	BZW2	28969	trt_oe	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_PC3_96H:EIF3H:2	PC3	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_VCAP_120H:NPM1:5	VCAP	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_VCAP_120H:MYC:5	VCAP	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_VCAP_120H:EIF3H:5	VCAP	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore

Showing 1 to 33 of 33 entries (filtered from 55,753 total entries) Previous 1 Next

Experiment | Cell line | Gene perturbed | KO or OE | Validation status in LIHC

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: driver validation & discovery using genetic perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Driver discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility (https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/#) | Show 500 entries | Search:

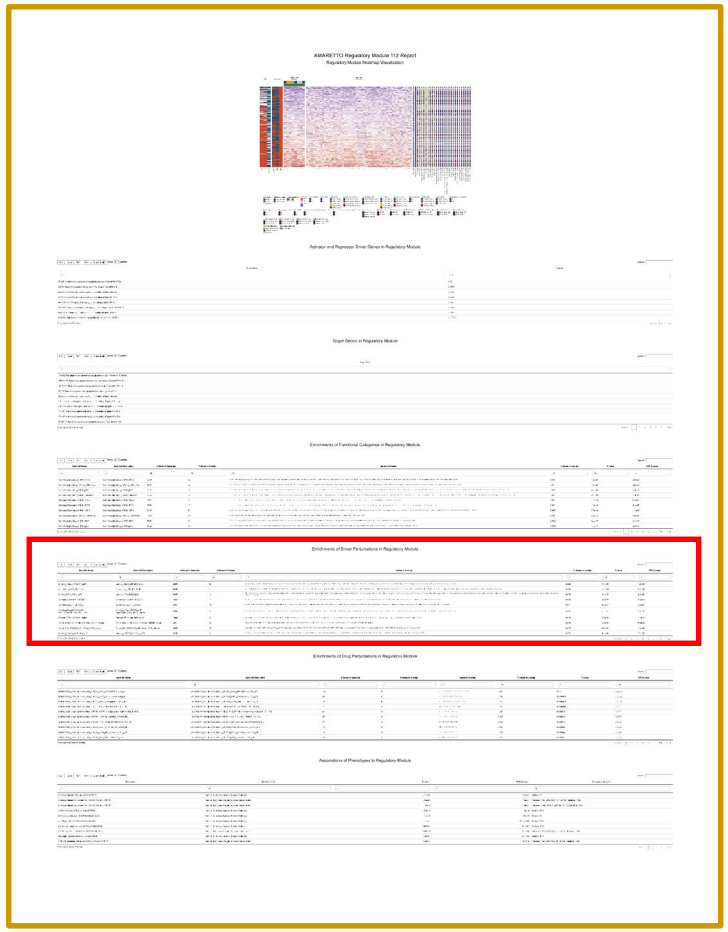
PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	LIHC
All	All	["MYC","BZW2","E2F5","EI	All	All	All	escore-pval-padj
CGS001_A375_96H:BZW2:1	A375	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_A375_96H:MYC:1	A375	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A375_96H:NPM1:1	A375	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
OEB005_A375_96H:BRDN0000408975:-666	A375	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_A375_96H:EIF3H:1	A375	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_A549_96H:EIF3H:1	A549	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_A549_96H:MYC:1	A549	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_A549_96H:E2F5:1	A549	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj
CGS001_HA1E_96H:MYC:1.5	HA1E	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HA1E_96H:EIF3H:1.5	HA1E	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_HA1E_96H:BZW2:1.5	HA1E	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HCC515_96H:MYC:2	HCC515	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:MYC:1.5	HEPG2	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_HEPG2_96H:BZW2:1.5	HEPG2	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
OEB005_HEPG2_96H:BRDN0000408975:-666	HEPG2	NPM1	4869	trt_oe	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_HT29_96H:BZW2:1	HT29	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_HT29_96H:MYC:1	HT29	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_HT29_96H:E2F5:1	HT29	E2F5	1875	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0129) , escore-pval-padj-zscore
CGS001_MCF7_144H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_144H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_96H:BZW2:2	MCF7	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_MCF7_96H:MYC:2	MCF7	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_MCF7_144H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_MCF7_96H:NPM1:2	MCF7	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj
CGS001_MCF7_96H:EIF3H:2	MCF7	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore
CGS001_NPC_96H:BZW2:1.5	NPC	BZW2	28969	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_NPC_96H:MYC:1.5	NPC	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
CGS001_PC3_96H:MYC:2	PC3	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj-zscore
OEB003_PC3_96H:BRDN0000405602:-666	PC3	BZW2	28969	trt_oe	landmark	Module 112 : A_D (w = 0.0206) , escore-pval-padj-zscore
CGS001_PC3_96H:EIF3H:2	PC3	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj
CGS001_VCAP_120H:NPM1:5	VCAP	NPM1	4869	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0205) , escore-pval-padj-zscore
CGS001_VCAP_120H:MYC:5	VCAP	MYC	4609	trt_sh.cgs	landmark	Module 112 : A_D (w = 0.0179) , escore-pval-padj
CGS001_VCAP_120H:EIF3H:5	VCAP	EIF3H	8667	trt_sh.cgs	best inferred	Module 112 : A_D (w = 0.0348) , escore-pval-padj-zscore

Showing 1 to 33 of 33 entries (filtered from 55,753 total entries) | Previous | 1 | Next

Search for drivers of Module 112 validated using genetic perturbations from LINCS/CMAP: MYC, BZW2, E2F5, EIF3H, NPM1

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
drug discovery

➢ Perturbation-AMARETTO v1

Enrichments of Drug Perturbations in Regulatory Module

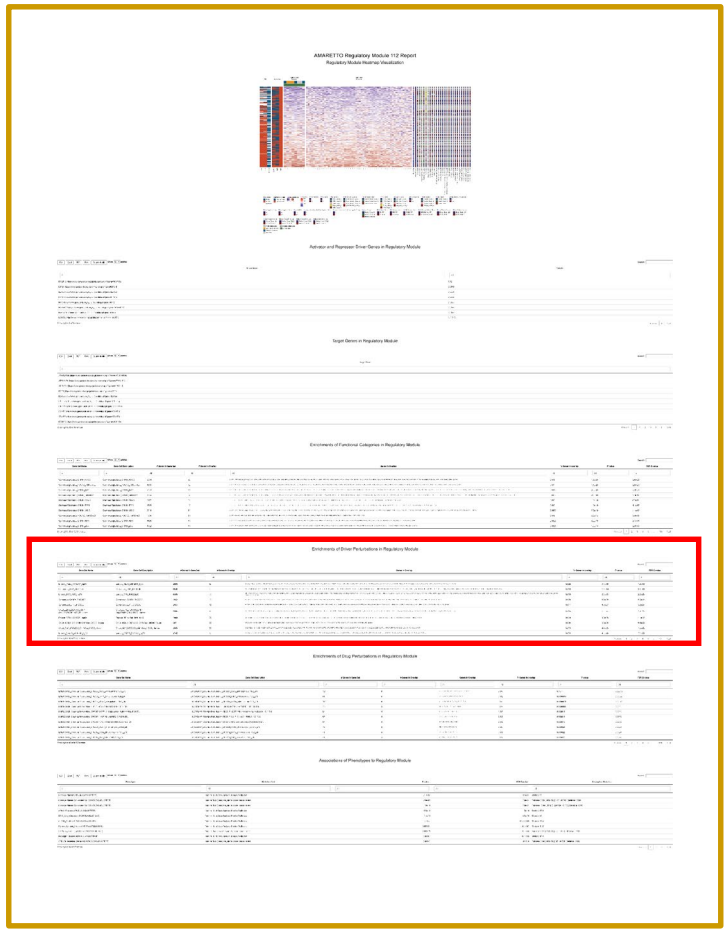
CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
All	All						
LINCSCMAP_ChemicalPerturbation_CPC007_A375_24H-t5247673-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC007_A375_24H-t5247673-10.0_DN	127	6	CPSF7, HNRNPL, BAZ1B, CAPRIN1, ZNHIT2, NDUFS3	0.047	5.1e-7	0.00013
LINCSCMAP_ChemicalPerturbation_CPC006_THP1_6H-teniposide-1.25_DN	LINCSCMAP_ChemicalPerturbation_CPC006_THP1_6H-teniposide-1.25_DN	81	5	ATF7IP, BAZ1B, CAPRIN1, PDS5B, ZNHIT2	0.062	0.0000013	0.00027
LINCSCMAP_ChemicalPerturbation_CPC006_VCAP_24H-piplartine-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC006_VCAP_24H-piplartine-10.0_DN	162	6	CPSF7, HNRNPL, FTSJ2, RRP8, ZNHIT2, KDM3B	0.037	0.0000021	0.00041
LINCSCMAP_ChemicalPerturbation_LJP005_HT29_24H-XMD16-144-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP005_HT29_24H-XMD16-144-3.33_DN	114	5	PDHX, WDR74, PRPF19, NAT10, FEN1	0.044	0.0000068	0.0010
LINCSCMAP_ChemicalPerturbation_CPC017_MCF7_24H-	LINCSCMAP_ChemicalPerturbation_CPC017_MCF7_24H-	64	4	DNAJC24, ATF7IP,	0.063	0.000015	0.0018
LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-vu0418946-2-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-vu0418946-2-10.0_DN	67	4	ATF7IP, CAPRIN1, OTUD1, ZNHIT2	0.060	0.000018	0.0020
LINCSCMAP_ChemicalPerturbation_CPC013_NPC_24H-	LINCSCMAP_ChemicalPerturbation_CPC013_NPC_24H-	67	4	HNRNPL, ATF7IP,	0.060	0.000018	0.0020
LINCSCMAP_ChemicalPerturbation_CPC004_VCAP_6H-mitomycin_c-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC004_VCAP_6H-mitomycin_c-10.0_DN	70	4	PDHX, ARCIP2, ATF7IP, ZNHIT2	0.057	0.000021	0.0023
LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-	LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-	71	4	ATF7IP, CAPRIN1,	0.056	0.000022	0.0024
LINCSCMAP_ChemicalPerturbation_CPC014_ASC_24H-a-1065-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC014_ASC_24H-a-1065-10.0_DN	71	4	HNRNPL, EIF2M, RRP8, ZNHIT2	0.056	0.000022	0.0024

Showing 1 to 10 of 2,167 entries (filtered from 9,505 total entries)

Previous | 1 | 2 | 3 | 4 | 5 | … | 217 | Next

AMARETTO report LIHC



Detailed report of MYC-driven Module 112:
drug discovery

➢ Perturbation-AMARETTO v1

Enrichments of Drug Perturbations in Regulatory Module

CSV | Excel | PDF | Print | Column visibility | Show 10 entries | Search:

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
All	All						
LINCSCMAP_ChemicalPerturbation_CPC007_A375_24H-t5247673-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC007_A375_24H-t5247673-10.0_DN	127	6	CPSF7, HNRNPL, BAZ1B, CAPRIN1, ZNHIT2, NDUFS3	0.047	5.1e-7	0.00013
LINCSCMAP_ChemicalPerturbation_CPC006_THP1_6H-teniposide-1.25_DN	LINCSCMAP_ChemicalPerturbation_CPC006_THP1_6H-teniposide-1.25_DN	81	5	ATF7IP, BAZ1B, CAPRIN1, PDS5B, ZNHIT2	0.062	0.0000013	0.00027
LINCSCMAP_ChemicalPerturbation_CPC006_VCAP_24H-piplartine-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC006_VCAP_24H-piplartine-10.0_DN	162	6	CPSF7, HNRNPL, FTSJ2, RRP8, ZNHIT2, KDM3B	0.037	0.0000021	0.00041
LINCSCMAP_ChemicalPerturbation_LJP005_HT29_24H-XMD16-144-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP005_HT29_24H-XMD16-144-3.33_DN	114	5	PDHX, WDR74, PRPF19, NAT10, FEN1	0.044	0.0000068	0.0010
LINCSCMAP_ChemicalPerturbation_CPC017_MCF7_24H-	LINCSCMAP_ChemicalPerturbation_CPC017_MCF7_24H-	64	4	DNAJC24, ATF7IP,	0.063	0.000015	0.0018
LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-vu0418946-2-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-vu0418946-2-10.0_DN	67	4	ATF7IP, CAPRIN1, OTUD1, ZNHIT2	0.060	0.000018	0.0020
LINCSCMAP_ChemicalPerturbation_CPC013_NPC_24H-	LINCSCMAP_ChemicalPerturbation_CPC013_NPC_24H-	67	4	HNRNPL, ATF7IP,	0.060	0.000018	0.0020
LINCSCMAP_ChemicalPerturbation_CPC004_VCAP_6H-mitomycin_c-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC004_VCAP_6H-mitomycin_c-10.0_DN	70	4	PDHX, ARID1B, ATF7IP, ZNHIT2	0.057	0.000021	0.0023
LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-	LINCSCMAP_ChemicalPerturbation_CPC012_HA1E_6H-	71	4	ATF7IP, CAPRIN1,	0.056	0.000022	0.0024
LINCSCMAP_ChemicalPerturbation_CPC014_ASC_24H-a-1065-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC014_ASC_24H-a-1065-10.0_DN	71	4	HNRNPL, EIF2M, RRP8, ZNHIT2	0.056	0.000022	0.0024

Showing 1 to 10 of 2,167 entries (filtered from 9,505 total entries) Previous 1 2 3 4 5 … 217 Next

Search for current LIHC treatments Sorafenib and Regorafenib:

Search: orafenib

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
All	All						
LINCSCMAP_ChemicalPerturbation_LJP006_HEPG2_24H-sorafenib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HEPG2_24H-sorafenib-10_DN	91	2	DDAH1, FEN1	0.022	0.019	0.11
LINCSCMAP_ChemicalPerturbation_LJP009_MCF7_24H-regorafenib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP009_MCF7_24H-regorafenib-10_DN	98	2	BANF1, FEN1	0.020	0.021	0.11

Showing 1 to 2 of 2 entries (filtered from 9,505 total entries) Previous 1 Next

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: drug discovery using chemical perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Drug discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility | Show [redacted] entries | Search: [redacted]

Dataset	Module	sig_id	Cell_Line	pert_iname	pert_type	pval_perturbation	qval_perturbation	ES	NES	nMoreExtreme	size	Phenotypes	Statistical_Test	pval_phenotype	qval_phenotype	Descriptive_Statistics	Chemical_Phenotype_Direction
LIHC	Module-112	CPC013_HEPG2_6H:BRD-K49810818-001-01-0:10	HEPG2	sorafenib	trt_cp	0.0114	0.0445	-0.4431	-1.8227	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0017	0.0341	-0.5687	-1.9697	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0038	0.0341	-0.4938	-2.0441	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_24H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0203	0.0702	-0.4375	-1.6111	6	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_6H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0054	0.0341	-0.5196	-1.9476	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPD002_MCF7_6H:BRD-K70401845-001-03-3:10	MCF7	erlotinib	trt_cp	0.0039	0.0341	-0.5105	-2.007	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	LJP002_MCF7_24H:BRD-K70401845-001-04-1:10	MCF7	erlotinib	trt_cp	0.0206	0.0702	-0.3627	-1.5352	3	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_PHH_24H:BRD-K46652470-001-02-6:10	PHH	nizatidine	trt_cp	0.0039	0.0341	-0.5339	-2.1747	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed

Showing 1 to 8 of 8 entries (filtered from 329,000 total entries)

Previous 1 Next

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drugs_Diseases/ (under development)

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: drug discovery using chemical perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Drug discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility | Show [redacted] entries | Search: [redacted]

Dataset	Module	sig_id	Cell_Line	pert_iname	pert_type	pval_perturbation	qval_perturbation	ES	NES	nMoreExtreme	size	Phenotypes	Statistical_Test	pval_phenotype	qval_phenotype	Descriptive_Statistics	Chemical_Phenotype_Direction
LIHC	Module-112	CPC013_HEPG2_6H:BRD-K49810818-001-01-0:10	HEPG2	sorafenib	trt_cp	0.0114	0.0445	-0.4431	-1.8227	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0017	0.0341	-0.5687	-1.9697	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0038	0.0341	-0.4938	-2.0441	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_24H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0203	0.0702	-0.4375	-1.6111	6	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_6H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0054	0.0341	-0.5196	-1.9476	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPD002_MCF7_6H:BRD-K70401845-001-03-3:10	MCF7	erlotinib	trt_cp	0.0039	0.0341	-0.5105	-2.007	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	LJP002_MCF7_24H:BRD-K70401845-001-04-1:10	MCF7	erlotinib	trt_cp	0.0206	0.0702	-0.3627	-1.5352	3	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_PHH_24H:BRD-K46652470-001-02-6:10	PHH	nizatidine	trt_cp	0.0039	0.0341	-0.5339	-2.1747	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed

Showing 1 to 8 of 8 entries (filtered from 329,000 total entries) Previous 1 Next

Data set, Module Experiment, Cell line, Compound, Statistics Phenotype: Survival Statistics Reversed?

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drugs_Diseases/ (under development)

Perturbation-AMARETTO report LIHC

Perturbation-AMARETTO v2: drug discovery using chemical perturbations from LINCS/CMAP

Case Study 1: virus-induced hepatocellular carcinoma

Drug discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility | Show [redacted] entries | Search: [redacted]

Dataset	Module	sig_id	Cell_Line	pert_iname	pert_type	pval_perturbation	qval_perturbation	ES	NES	nMoreExtreme	size	Phenotypes	Statistical_Test	pval_phenotype	qval_phenotype	Descriptive_Statistics	Chemical_Phenotype_Direction
LIHC	Module-112	CPC013_HEPG2_6H:BRD-K49810818-001-01-0:10	HEPG2	sorafenib	trt_cp	0.0114	0.0445	-0.4431	-1.8227	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0017	0.0341	-0.5687	-1.9697	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0038	0.0341	-0.4938	-2.0441	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_24H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0203	0.0702	-0.4375	-1.6111	6	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC013_MCF7_6H:BRD-K49810818-001-01-0:10	MCF7	sorafenib	trt_cp	0.0054	0.0341	-0.5196	-1.9476	1	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPD002_MCF7_6H:BRD-K70401845-001-03-3:10	MCF7	erlotinib	trt_cp	0.0039	0.0341	-0.5105	-2.007	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	LJP002_MCF7_24H:BRD-K70401845-001-04-1:10	MCF7	erlotinib	trt_cp	0.0206	0.0702	-0.3627	-1.5352	3	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_PHH_24H:BRD-K46652470-001-02-6:10	PHH	nizatidine	trt_cp	0.0039	0.0341	-0.5339	-2.1747	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed

Showing 1 to 8 of 8 entries (filtered from 329,000 total entries) Previous 1 Next

Search drug treatments reversing survival-associated Module 112 using chemical perturbations from LINCS/CMAP, Query: Sorafenib, Erlotinib, Nizatidine

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drugs_Diseases/ (under development)

AMARETTO report LIHC

MYC-driven Module 112

Summary of MYC-regulated Module 112:

MYC CNV amplification, associated with induced MYC expression, and MYC is activator of its target genes

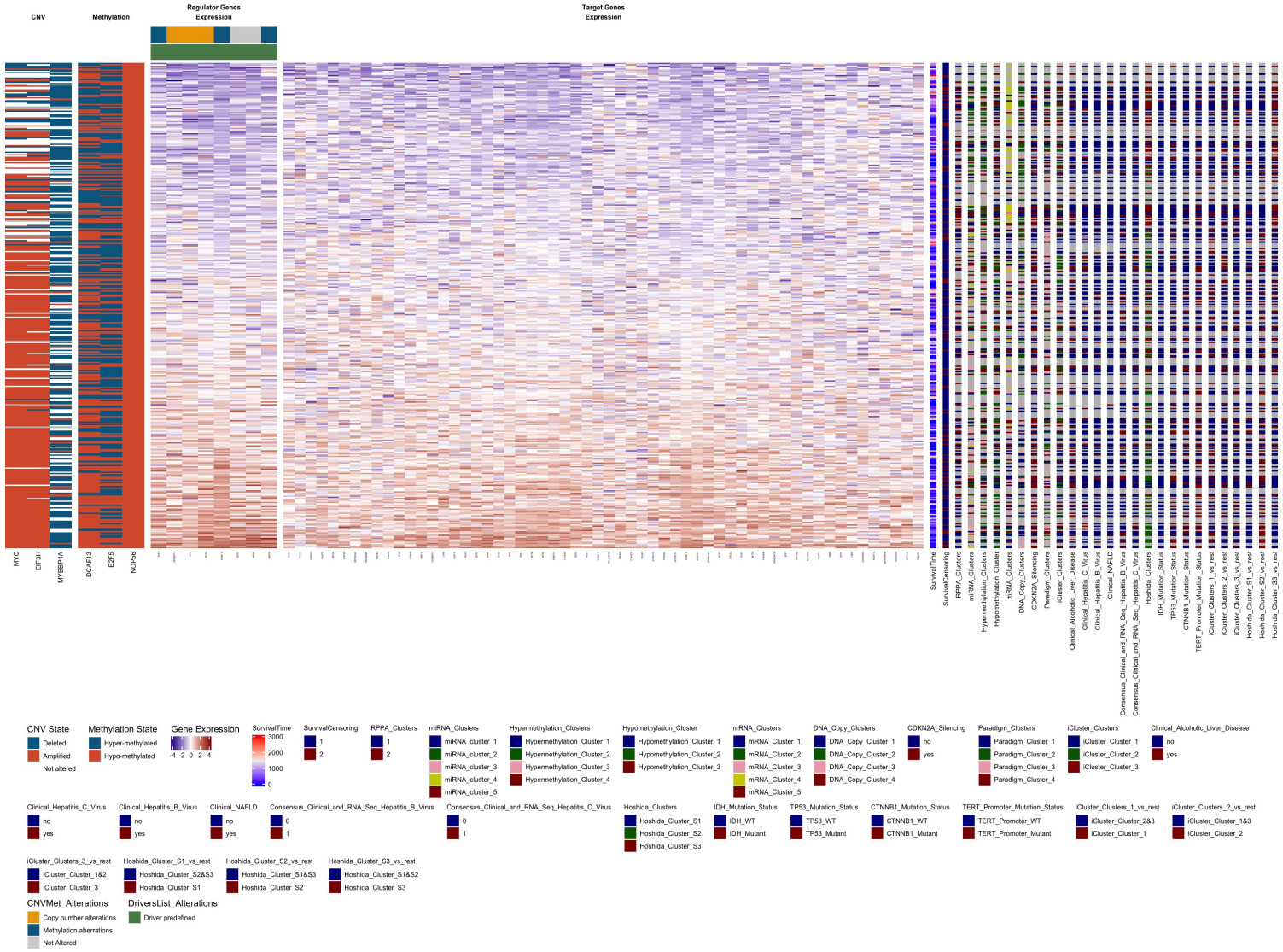
Associated with survival: higher expression, poorer survival

Enriched for gene signature HOSHIDA_LIVER_CANCER_SUBCLASS_S2 (Genes from 'subtype S2' signature of hepatocellular carcinoma (HCC): proliferation, MYC and AKT1 activation.)

Drivers validated:

- MYC: ENCODE and ChEA ChIP-Seq, bound to its target genes
- MYC, BZW2, E2F5, EIF3H, NPM1: LINCS/CMAP genetic perturbations, modulating drivers modulates its target genes

Drugs: Sorafenib, Regorafenib, Erlotinib, Nizatidine,… reverse survival-associated behavior of driver and target genes of Module 112

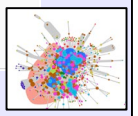


http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/TCGA_LIHC/AMARETTOhtmls/index.html

Community-AMARETTO report virus-induced LIHC

Community	CCLE_Liver	scHBV	scHCV	TCGA_LIHC	tcHBV	tcHCV	ImmuneSignatures	LiverSignatures	StemSignatures
Community 1	Module 13, Module 15, Module 23, Module 51, Module 53, Module 75, Module 88, Module 97, Module 102, Module 115, Module 116, Module 118, Module 131, Module 147	Module 31, Module 41, Module 56	Module 22, Module 93, Module 100	Module 3, Module 5, Module 10, Module 13, Module 14, Module 25, Module 30, Module 33, Module 37, Module 45, Module 48, Module 55, Module 58, Module 59, Module 74, Module 77, Module 79, Module 80, Module 81, Module 88, Module 89, Module 97, Module 113, Module 115, Module 116, Module 118, Module 123, Module 124, Module 125, Module 126, Module 139, Module 141, Module 145	Module 11, Module 36, Module 50, Module 59, Module 79, Module 80, Module 90, Module 102, Module 123, Module 150	Module 35, Module 42, Module 43, Module 47, Module 59, Module 115, Module 122, Module 130, Module 145	CIBERSORT_MAST_CELLS_RESTING	HOSHIDA_LIVER_CANCER_SUBCLASS_S1, HOSHIDA_LIVER_CANCER_SURVIVAL_UP	BENPORATH_EED_TARGETS, BENPORATH_ES_WITH_H3K27ME3, BENPORATH_PRC2_TARGETS, BENPORATH_SUZ12_TARGETS

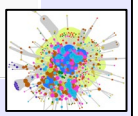
- Genes from 'subtype S1' signature of hepatocellular carcinoma (HCC): aberrant activation of the WNT signaling pathway
- Survival signature genes defined in adjacent liver tissue: genes correlated with poor survival of hepatocellular carcinoma (HCC) patients
- Polycomb Repression Complex 2 (PRC) targets; identified by ChIP on chip on human embryonic stem cells as genes that: possess the trimethylated H3K27 mark in their promoters and are bound by SUZ12 and EED Polycomb proteins
- IL8 pan-etiology driver of HCV and HBV virus-induced HCC associated with HCV and HBV viral load and HCC survival



Community 1

Community	CCLE_Liver	scHBV	scHCV	TCGA_LIHC	tcHBV	tcHCV	ImmuneSignatures	LiverSignatures	StemSignatures
Community 3	Module 4, Module 5, Module 10, Module 12, Module 20, Module 21, Module 27, Module 32, Module 33, Module 41, Module 47, Module 50, Module 57, Module 58, Module 91, Module 95, Module 95, Module 50, Module 95, Module 98, Module 111, Module 113, Module 117, Module 119, Module 120, Module 127, Module 137, Module 140, Module 142, Module 148	Module 1, Module 2, Module 3, Module 4, Module 5, Module 9, Module 10, Module 11, Module 12, Module 20, Module 22, Module 25, Module 26, Module 27, Module 28, Module 30, Module 37, Module 39, Module 40, Module 44, Module 48, Module 50, Module 54, Module 55, Module 58, Module 60, Module 62, Module 72, Module 74, Module 80, Module 83, Module 85, Module 93, Module 99, Module 100	Module 2, Module 5, Module 8, Module 10, Module 11, Module 18, Module 25, Module 34, Module 45, Module 46, Module 49, Module 51, Module 52, Module 53, Module 60, Module 61, Module 62, Module 65, Module 67, Module 74, Module 77, Module 79, Module 80, Module 84, Module 85, Module 87, Module 90, Module 92, Module 98, Module 99	Module 1, Module 25, Module 28, Module 49, Module 50, Module 93, Module 95, Module 98	Module 14, Module 15, Module 30, Module 40, Module 51, Module 55, Module 70, Module 82, Module 92, Module 95, Module 107, Module 111, Module 118, Module 121, Module 125, Module 143, Module 144, Module 149	Module 5, Module 19, Module 27, Module 52, Module 54, Module 71, Module 82, Module 85, Module 97, Module 99, Module 111, Module 125, Module 129	0	HOSHIDA_LIVER_CANCER_SUBCLASS_S2	BENPORATH_MYC_MAX_TARGETS, BENPORATH_MYC_TARGETS_WITH_EBOX, BENPORATH_NANOG_TARGETS, BENPORATH_OCT4_TARGETS, BENPORATH_SOX2_TARGETS

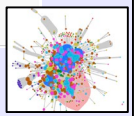
- Genes from 'subtype S2' signature of hepatocellular carcinoma (HCC): proliferation, MYC and AKT1 activation
- MYC targets; targets of c-Myc and Max identified by ChIP on chip in a Burkitt's lymphoma cell line; overlap set; and in cultured cell lines, focusing on E-box-containing genes; high affinity bound subset
- CORE stemness genes upregulated and identified by ChIP on chip as NOS (Nanog, OCT4, SOX2) transcription factor targets in human embryonic stem cells
- STX7 pan-etiology driver of HCV and HBV virus-induced HCC



Community 3

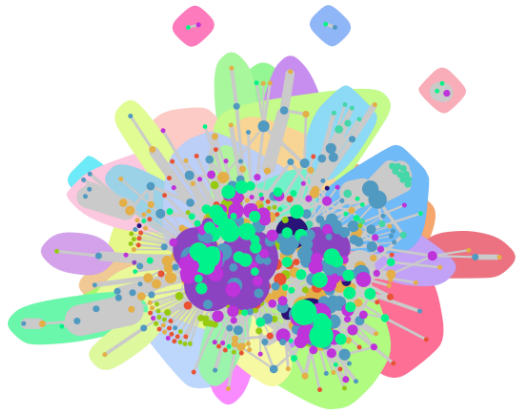
Community	CCLE_Liver	scHBV	scHCV	TCGA_LIHC	tcHBV	tcHCV	ImmuneSignatures	LiverSignatures	StemSignatures
Community 5	Module 8, Module 55, Module 59, Module 67, Module 71, Module 72, Module 77, Module 91, Module 92, Module 121, Module 124, Module 145, Module 150	Module 23, Module 29, Module 33, Module 71, Module 81, Module 82	Module 8, Module 24, Module 41, Module 63, Module 83, Module 97	Module 12, Module 22, Module 24, Module 36, Module 42, Module 51, Module 53, Module 63, Module 64, Module 70, Module 75, Module 85, Module 92, Module 110, Module 117, Module 123, Module 128, Module 142	Module 10, Module 45, Module 48, Module 53, Module 97, Module 109, Module 128, Module 129, Module 137, Module 138, Module 139, Module 145	Module 8, Module 57, Module 58, Module 74, Module 79, Module 83, Module 84, Module 96	0	HOSHIDA_LIVER_CANCER_LATE_RECURRENCE_DN, HOSHIDA_LIVER_CANCER_SUBCLASS_S3, HOSHIDA_LIVER_CANCER_SURVIVAL_DN	0

- Genes from 'subtype S3' signature of hepatocellular carcinoma (HCC): hepatocyte differentiation
- Survival signature genes defined in adjacent liver tissue: genes correlated with good survival of hepatocellular carcinoma (HCC) patients
- Liver specific genes from Human Gene Expression Index, the HuGE Index, http://www.hugeindex.org
- APOC3 pan-etiology driver of HCV and HBV virus-induced HCC validated in all 6 data sources using genetic perturbations of APOC3 in the HepG2 liver cancer cell line



Community 5

Community Network Visualization



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41

TCGA_LIHC
CCLE_Liver
tcHCV
scHCV
tcHBV
scHBV
LiverSignatures
ImmuneSignatures
StemSignatures

NDEx

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Liver_6DS/index.html

Perturbation-AMARETTO report virus-induced LIHC:

Driver validation & discovery: across modules in tcHCV, scHCV, tcHBV, scHBV, CCLE and LIHC

Case Study 1: virus-induced hepatocellular carcinoma

Driver discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility | Show 500 entries | Search:

PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	DataSetFrequency	CCLE	LIHC	scHBV	scHCV	tcHBV	tcHCV
CGS001_HEPG2_96H:AGT:1.5	HEPG2	AGT	183	trt_sh.cgs	best inferred	6	Module 59 : T_CD (w = 0) , escore-pval-padj-zscore	Module 63 : T_CD (w = 0) , escore-pval-padj-zscore	Module 71 : A_D (w = 0.3526) , escore-pval-padj-zscore	Module 22 : R_D (w = -0.0062) , pval-padj-zscore; Module 63 : T_CD (w = 0) , escore-pval-padj-zscore	Module 83 : A_D (w = 0.118) , escore-pval-padj-zscore	Module 96 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HEPG2_96H:APOC3:1.5	HEPG2	APOC3	345	trt_sh.cgs	inferred	6	Module 92 : T_CD (w = 0) , escore-pval-padj-zscore	Module 53 : A_D (w = 0.0681) , escore-pval-padj-zscore	Module 23 : A_D (w = 0.0448) , escore-pval-padj-zscore	Module 8 : A_D (w = 3e-04) , escore-pval-padj-zscore	Module 97 : A_D (w = 0.1519) , escore-pval-padj-zscore	Module 79 : A_D (w = 0.1892) , escore-pval-padj-zscore
CGS001_PC3_96H:CBR1:2	PC3	CBR1	873	trt_sh.cgs	landmark	6	Module 73 : T_CD (w = 0) , escore-pval-padj-zscore	Module 75 : T_CD (w = 0) , escore-pval-padj-zscore	Module 20 : T_CD (w = 0) , escore-pval-padj-zscore	Module 83 : A_D (w = 0.0286) , escore-pval-padj-zscore	Module 99 : R_D (w = -0.2644) , zscore; Module 114 : T_CD (w = 0) , escore-pval-padj-zscore	Module 54 : T_CD (w = 0) , escore-pval-padj-zscore
OEC001_PC3_72H:CCSBBROAD304_03269:-666	PC3	CD320	51293	trt_oe	landmark	6	Module 147 : T_CD (w = 0) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore	Module 80 : A_D (w = 0.3247) , escore-pval-padj-zscore; Module 91 : A_D (w = 0.0021) , escore-pval-padj-zscore	Module 48 : T_CD (w = 0) , escore-pval-padj-zscore	Module 60 : T_CD (w = 0) , escore-pval-padj-zscore	Module 142 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_MCF7_96H:IL8:2	MCF7	CXCL8	3576	trt_sh.cgs	best inferred	6	Module 83 : T_CD (w = 0) , escore-pval-padj-zscore	Module 115 : A_D (w = 0.0079) , escore-pval-padj-zscore	Module 16 : A_D (w = 0.0266) , escore-pval-padj-zscore; Module 41 : T_CD (w = 0) , escore-pval-padj-zscore	Module 93 : A_D (w = 0.0606) , escore-pval-padj-zscore	Module 36 : A_D (w = 0.2325) , escore-pval-padj-zscore	Module 130 : A_D (w = 4e-04) , escore-pval-padj-zscore
CGS001_PC3_96H:DRAP1:2	PC3	DRAP1	10589	trt_sh.cgs	landmark	6	Module 78 : A_D (w = 0.4877) , escore-pval-padj-zscore; Module 121 : R_D (w = -0.0039) , pval-padj-zscore	Module 94 : T_CD (w = 0) , escore-pval-padj-zscore	Module 15 : A_D (w = 0.001) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.0907) , escore-pval-padj-zscore	Module 79 : T_CD (w = 0) , escore-pval-padj-zscore	Module 67 : T_CD (w = 0) , escore-pval-padj-zscore	Module 36 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_A549_96H:FAH:1	A549	FANCA	2175	trt_sh.cgs	best inferred	6	Module 80 : T (w = 0) , escore-pval-padj-zscore; Module 84 : T (w = 0) , escore-padj-zscore	Module 75 : T (w = 0) , escore-pval-padj-zscore; Module 90 : T (w = 0) , escore-pval-padj-zscore	Module 82 : T (w = 0) , escore-pval-padj-zscore	Module 35 : T (w = 0) , escore-pval-padj-zscore; Module 51 : T (w = 0) , zscore	Module 19 : T (w = 0) , escore-pval-padj-zscore; Module 60 : T (w = 0) , escore-pval-padj-zscore	Module 52 : T (w = 0) , escore-pval-padj-zscore; Module 131 : T (w = 0) , escore-pval-padj-zscore
OEB005_HT29_96H:BRDN0000410732:-666	HT29	GABPB1	2553	trt_oe	landmark	6	Module 94 : T_CD (w = 0) , escore-pval-padj-zscore	Module 69 : T_CD (w = 0) , escore-pval-padj-zscore	Module 89 : A_D (w = 0.0038) , escore-pval-padj-zscore	Module 55 : A_D (w = 0.1145) , escore-pval-padj-zscore	Module 65 : T_CD (w = 0) , escore-pval-padj-zscore	Module 30 : T_CD (w = 0) , escore-pval-padj-zscore; Module 94 : T_CD (w = 0) , escore-padj-zscore
OEC001_A375_96H:CCSBBROAD304_03340:-666	A375	LSR	51599	trt_oe	landmark	6	Module 71 : T (w = 0) , escore-pval-padj-zscore	Module 42 : T (w = 0) , escore-pval-padj-zscore	Module 12 : T (w = 0) , escore-pval-padj-zscore	Module 65 : T (w = 0) , escore-pval-padj-zscore	Module 109 : T (w = 0) , escore-pval-padj-zscore	Module 66 : T (w = 0) , escore-pval-padj-zscore
OEB006_A549_96H:BRDN0000462006:-666	A549	LSR	51599	trt_oe	landmark	6	Module 71 : T (w = 0) , escore-pval-padj-zscore	Module 42 : T (w = 0) , escore-pval-padj-zscore	Module 12 : T (w = 0) , escore-pval-padj-zscore	Module 65 : T (w = 0) , escore-pval-padj-zscore	Module 109 : T (w = 0) , escore-pval-padj-zscore	Module 66 : T (w = 0) , escore-pval-padj-zscore
CGS001_HCC515_96H:MRPL12:2	HCC515	MRPL12	6182	trt_sh.cgs	landmark	6	Module 123 : T (w = 0) , escore-pval-padj-zscore	Module 120 : T (w = 0) , escore-pval-padj-zscore	Module 28 : T (w = 0) , escore-pval-padj-zscore	Module 30 : T (w = 0) , escore-pval-padj-zscore	Module 93 : T (w = 0) , escore-pval-padj-zscore	Module 44 : T (w = 0) , escore-pval-padj-zscore
CGS001_PC3_144H:MRPL12:2	PC3	MRPL12	6182	trt_sh.cgs	landmark	6	Module 123 : T (w = 0) , escore-pval-padj-zscore	Module 120 : T (w = 0) , escore-pval-padj-zscore	Module 28 : T (w = 0) , escore-pval-padj-zscore	Module 30 : T (w = 0) , escore-pval-padj-zscore	Module 93 : T (w = 0) , escore-pval-padj-zscore	Module 44 : T (w = 0) , escore-pval-padj-zscore
CGS001_MCF7_144H:MTHFD2:2	MCF7	MTHFD2	10797	trt_sh.cgs	landmark	6	Module 79 : T_CD (w = 0) , escore-pval-padj-zscore	Module 3 : T_CD (w = 0) , escore-pval-padj-zscore	Module 71 : T_CD (w = 0) , escore-pval-padj-zscore	Module 87 : T_CD (w = 0) , escore-pval-padj-zscore	Module 74 : T_CD (w = 0) , escore-pval-padj-zscore	Module 148 : A_D (w = 0.0083) , escore-pval-padj-zscore
OEC001_A375_96H:CCSBBROAD304_01093:-666	A375	NFKBIA	4792	trt_oe	landmark	6	Module 60 : T_CD (w = 0) , escore-pval-padj-zscore	Module 85 : T_CD (w = 0) , escore-pval-padj-zscore	Module 39 : T_CD (w = 0) , escore-pval-padj-zscore	Module 80 : T_CD (w = 0) , escore-pval-padj-zscore	Module 112 : T_CD (w = 0) , escore-pval-padj-zscore	Module 114 : T_CD (w = 0) , escore-pval-padj-zscore
OEB001_HA1E_96H:BRDN0000398867:-666	HA1E	NFKBIA	4792	trt_oe	landmark	6	Module 60 : T_CD (w = 0) , escore-pval-padj-zscore	Module 85 : T_CD (w = 0) , escore-pval-padj-zscore	Module 39 : T_CD (w = 0) , escore-pval-padj-zscore	Module 80 : T_CD (w = 0) , escore-pval-padj-zscore	Module 112 : T_CD (w = 0) , escore-pval-padj-zscore	Module 114 : T_CD (w = 0) , escore-pval-padj-zscore
OEB001_HCC515_96H:BRDN0000398867:-666	HCC515	NFKBIA	4792	trt_oe	landmark	6	Module 60 : T_CD (w = 0) , escore-pval-padj-zscore	Module 85 : T_CD (w = 0) , escore-pval-padj-zscore	Module 39 : T_CD (w = 0) , escore-pval-padj-zscore	Module 80 : T_CD (w = 0) , escore-pval-padj-zscore	Module 112 : T_CD (w = 0) , escore-pval-padj-zscore	Module 114 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HA1E_96H:PCNA:1.5	HA1E	PCNA	5111	trt_sh.cgs	landmark	6	Module 95 : T (w = 0) , escore-pval-padj-zscore	Module 149 : T (w = 0) , escore-pval-padj-zscore	Module 44 : T (w = 0) , escore-pval-padj-zscore	Module 35 : T (w = 0) , escore-pval-padj-zscore	Module 60 : T (w = 0) , escore-pval-padj-zscore	Module 85 : T (w = 0) , escore-pval-padj-zscore
OEB005_MCF7_96H:BRDN0000409395:-666	MCF7	PCNA	5111	trt_oe	landmark	6	Module 95 : T (w = 0) , escore-pval-padj-zscore	Module 149 : T (w = 0) , escore-pval-padj-zscore	Module 44 : T (w = 0) , escore-pval-padj-zscore	Module 35 : T (w = 0) , escore-pval-padj-zscore	Module 60 : T (w = 0) , escore-pval-padj-zscore	Module 85 : T (w = 0) , escore-pval-padj-zscore

CCLE TCGA scHBV scHCV tcHBV tcHCV

Experiment, Cell line, Gene perturbed, KO or OE

Validation status in 6 liver disease data sets

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drivers/

Perturbation-AMARETTO report virus-induced LIHC:

Drug discovery: Nizatidine reverses disease-associated modules in scHCV & scHBV (viral load), and LIHC (survival)

Case Study 1: virus-induced hepatocellular carcinoma

Drug discovery across 6 data sets

CSV | Excel | PDF | Print | Column visibility | Show 500 entries | Search:

Dataset	Module	sig_id	Cell_Line	pert_iname	pert_type	pval_perturbation	qval_perturbation	ES	NES	nMoreExtreme	size	Phenotypes	Statistical_Test	pval_phenotype	qval_phenotype	Descriptive_Statistics	Chemical_Phenotype_Direction
LIHC	Module-105	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0038	0.0535	-0.5305	-2.1799	0	36	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0018766	0.140745	Beta: 0.62117, Hazard Ratio: 1.8611, 95% CI: [1.2581,2.7532], Wald Statistic: 9.67	reversed
LIHC	Module-112	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0017	0.0341	-0.5687	-1.9697	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-112	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0038	0.0341	-0.4938	-2.0441	0	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0013013	0.140745	Beta: 0.64749, Hazard Ratio: 1.9107, 95% CI: [1.2877,2.8352], Wald Statistic: 10.34	reversed
LIHC	Module-129	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0025	0.0155	0.3926	1.5686	1	78	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.049614	0.6179884615384615	Beta: -0.30619, Hazard Ratio: 0.73625, 95% CI: [0.54234,0.99948], Wald Statistic: 3.85	reversed
LIHC	Module-145	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0114	0.0909	-0.3827	-1.5842	2	38	SurvivalTime (COXPROPHAZARDTIMETOEVENT), SurvivalCensoring (COXPROPHAZARDRIGHTCENSORING)	Survival Analysis: Cox proportional hazards regression (Wald test)	0.0044566	0.22283	Beta: 0.60976, Hazard Ratio: 1.84, 95% CI: [1.2087,2.8011], Wald Statistic: 8.09	reversed
scHCV	Module-8	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0386	0.1512	0.3664	1.4524	13	53	ViralLoad (SPEARMANCORRTEST)	Continuous or Ordinal Analysis: Spearman rank correlation analysis	0.000086042	0.000215105	Correlation: -0.64, Statistic: 17500	reversed
scHCV	Module-48	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0034	0.1901	-0.4854	-1.9181	0	30	ViralLoad (SPEARMANCORRTEST)	Continuous or Ordinal Analysis: Spearman rank correlation analysis	3.5371e-7	0.0000176855	Correlation: 0.706, Statistic: 3130	reversed
scHCV	Module-57	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0082	0.1919	-0.3861	-1.6928	1	44	ViralLoad (SPEARMANCORRTEST)	Continuous or Ordinal Analysis: Spearman rank correlation analysis	0.011602	0.04834166666666667	Correlation: 0.395, Statistic: 6450	reversed
scHBV	Module-25	CPC018_HEPG2_6H:BRD-K46652470-001-02-6:10	HEPG2	nizatidine	trt_cp	0.0042	0.1899	-0.3878	-1.7613	0	51	ViralLoad (SPEARMANCORRTEST)	Continuous or Ordinal Analysis: Spearman rank correlation analysis	0.040756	0.5576	Correlation: 0.265, Statistic: 26500	reversed
scHBV	Module-47	CPC015_HEPG2_6H:BRD-K73589491-001-05-4:10	HEPG2	nizatidine	trt_cp	0.0179	0.2449	0.4806	1.6566	6	27	ViralLoad (SPEARMANCORRTEST)	Continuous or Ordinal Analysis: Spearman rank correlation analysis	0.034134	0.5576	Correlation: -0.274, Statistic: 45900	reversed

Showing 1 to 10 of 10 entries (filtered from 329,000 total entries)

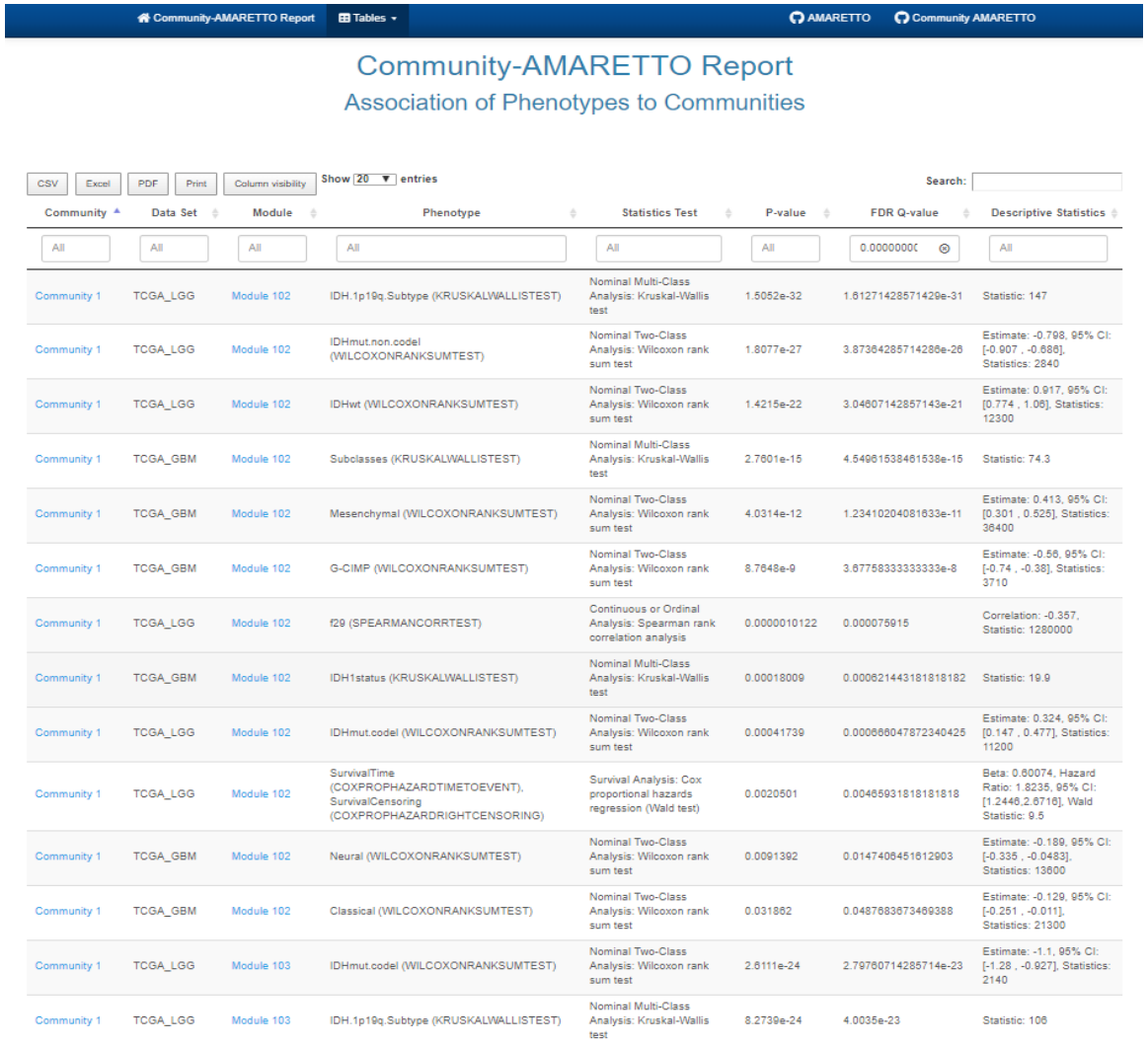
Previous 1 Next

https://pochetlab.shinyapps.io/pAMARETTO_Liver_6DS_Drugs_Diseases/ (under development)

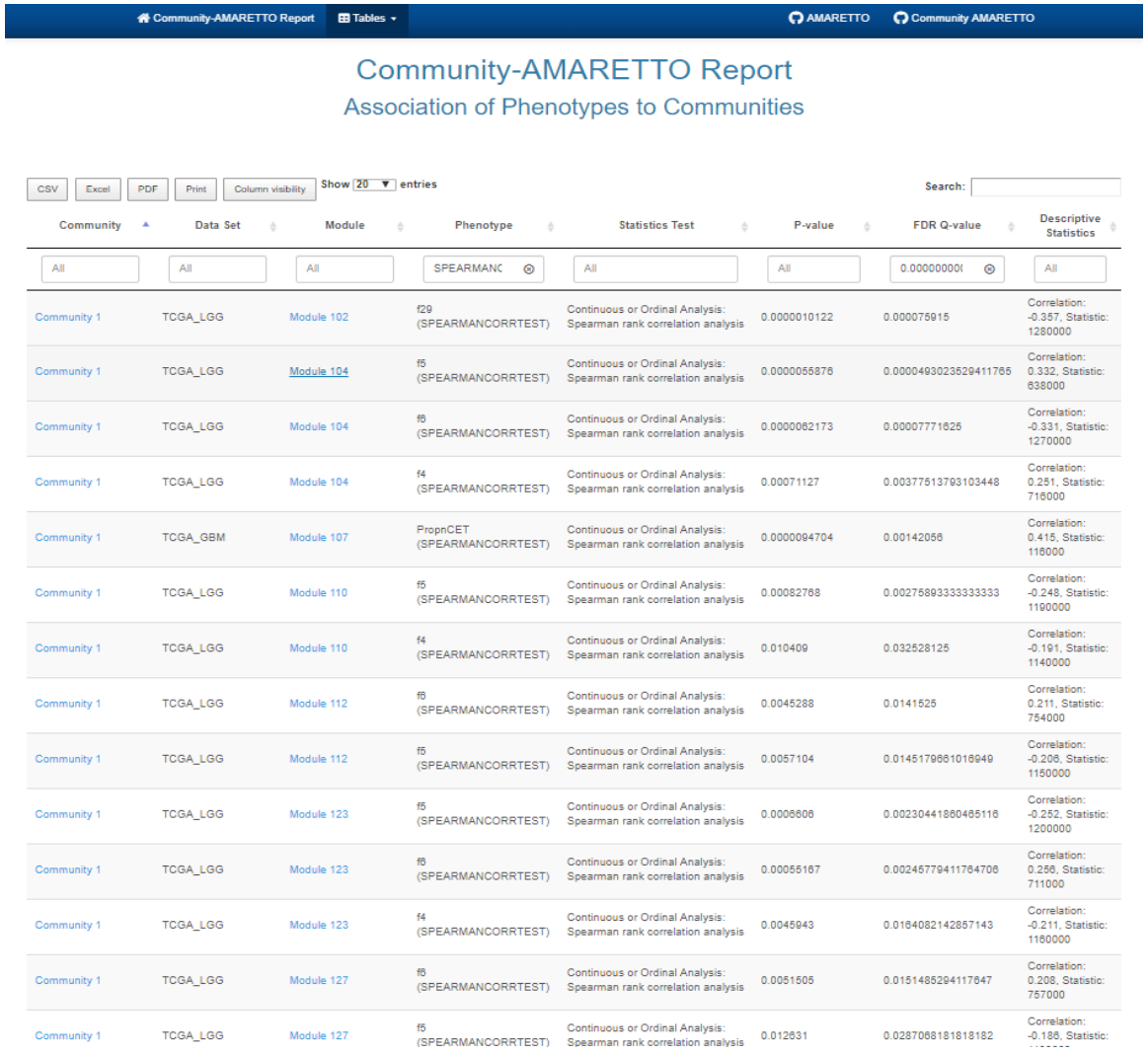
Case Study 2

Glioblastoma Multiforme (GBM)
& Low-grade Glioma (LGG)

Community-AMARETTO report GBM/LGG



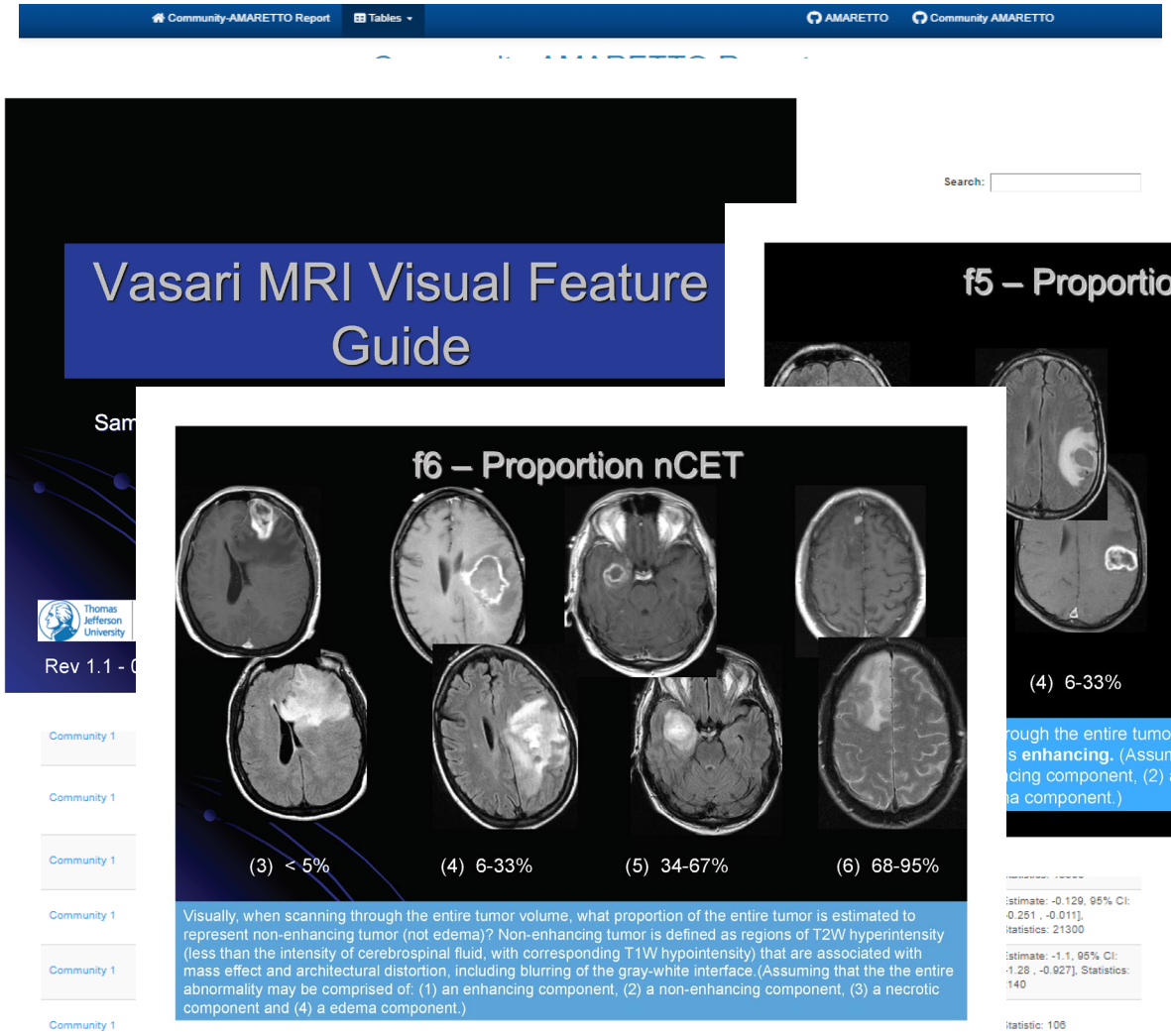
Clinical, molecular and imaging-derived phenotypes from TCGA



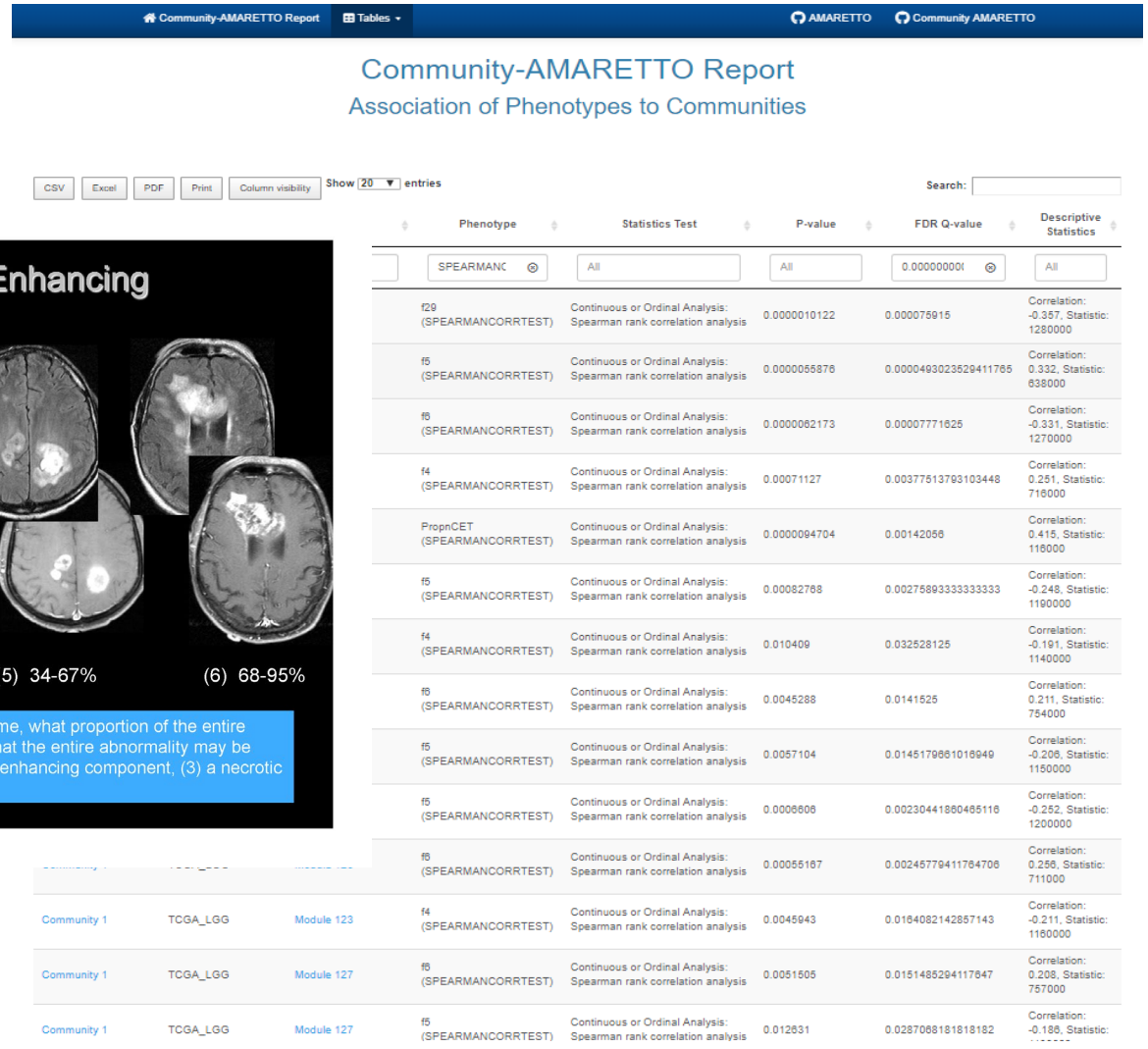
Imaging-derived phenotypes from TCGA/TCIA VASARI/Rembrandt

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/index_PhenoAssociation.html

Community-AMARETTO report GBM/LGG



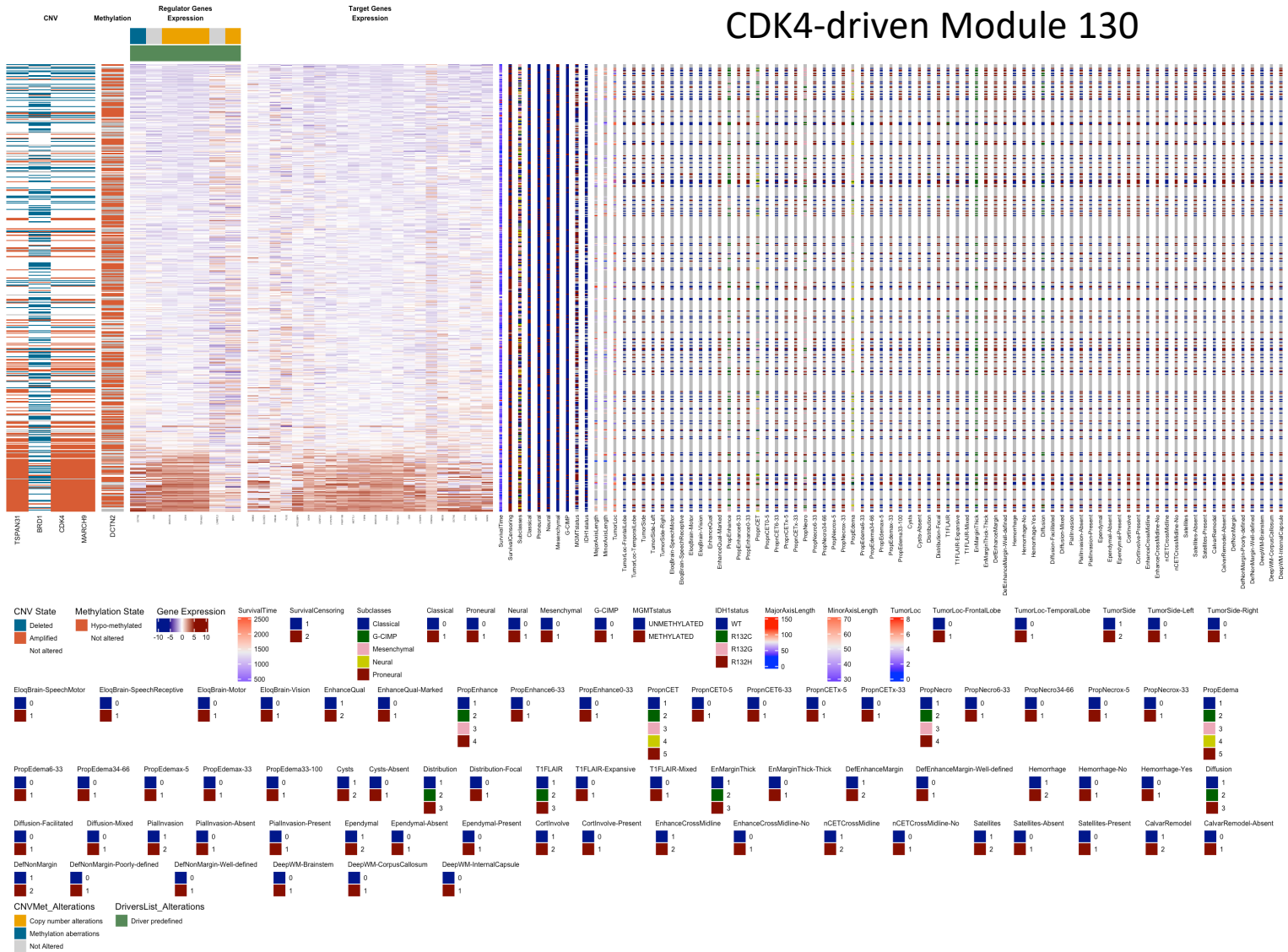
Clinical, molecular and imaging-derived phenotypes from TCGA



Imaging-derived phenotypes from TCGA/TCIA VASARI/Rembrandt

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/index_PhenoAssociation.html

AMARETTO report GBM



Module 130 regulated by CDK4

CDK4 amplifications/deletions, associated with induced/repressed CDK4 expression levels

CDK4 is activator of its target genes
Target genes CDKN2A and MDM2

Represents proneural molecular subclass of GBM (higher expression)

Enriched for functional categories:

- TCGA GLIOBLASTOMA COPY NUMBER UP (Genes up-regulated and displaying increased copy number in glioblastoma samples)
- KEGG GLIOMA (Glioma)
- PID RB 1 PATHWAY (Regulation of retinoblastoma protein)
- KEGG P53 SIGNALING PATHWAY (p53 signaling pathway)

Drivers validated (across GBM and related LGG modules): CDK4, CDKN2A, MDM2: LINCS/CMAP

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_GBM/AMARETTOhtmls/modules/module130.html

Perturbation-AMARETTO report GBM/LGG

Perturbation-AMARETTO v2: drug discovery using chemical perturbations from LINCS/CMAP

Case Study 2: glioblastoma multiforme and low-grade glioma

Driver discovery across 2 data sets

CSV | Excel | PDF | Print | Column visibility | Show 500 entries | Search:

PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	DataSetFrequency	GBM	LGG
All	All	["CDK4","C	All	All	All	["2","1"]	Module 130 :	All
CGS001_A375_96H:CDK4:1	A375	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_A549_96H:CDK4:1	A549	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HA1E_96H:CDK4:1:5	HA1E	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HCC515_96H:CDK4:2	HCC515	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HT29_96H:CDK4:1	HT29	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_MCF7_96H:CDK4:2	MCF7	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_PC3_96H:CDK4:2	PC3	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_VCAP_120H:CDK4:5	VCAP	CDK4	1019	trt_sh.cgs	landmark	2	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HT29_96H:CDKN2A:1	HT29	CDKN2A	1029	trt_sh.cgs	landmark	2	Module 130 : T_CD (w = 0) , escore-pval-padj-zscore	Module 53 : T_CD (w = 0) , escore-pval-padj-zscore
OEB005_HCC515_96H:BRDN0000410000:-666	HCC515	CDK4	1019	trt_oe	landmark	1	Module 130 : A_D (w = 0.198) , zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
CGS001_HEPG2_96H:CDK4:1:5	HEPG2	CDK4	1019	trt_sh.cgs	landmark	1	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , escore-padj-zscore
OEB005_HEPG2_96H:BRDN0000410000:-666	HEPG2	CDK4	1019	trt_oe	landmark	1	Module 130 : A_D (w = 0.198) , padj-zscore	Module 76 : T_CD (w = 0) , escore-pval-padj-zscore
OEB005_MCF7_96H:BRDN0000410000:-666	MCF7	CDK4	1019	trt_oe	landmark	1	Module 130 : A_D (w = 0.198) , escore-pval-padj-zscore	Module 76 : T_CD (w = 0) , pval-padj-zscore
CGS001_HA1E_96H:CDKN2A:1:5	HA1E	CDKN2A	1029	trt_sh.cgs	landmark	1	Module 130 : T_CD (w = 0) , escore-pval-padj-zscore	Module 53 : T_CD (w = 0) , escore-zscore
CGS001_HCC515_96H:CDKN2A:2	HCC515	CDKN2A	1029	trt_sh.cgs	landmark	1	Module 130 : T_CD (w = 0) , escore-pval-padj-zscore	Module 53 : T_CD (w = 0) , zscore
CGS001_MCF7_144H:CDKN2A:2	MCF7	CDKN2A	1029	trt_sh.cgs	landmark	1	Module 130 : T_CD (w = 0) , escore-pval-padj	Module 53 : T_CD (w = 0) , none
CGS001_HT29_96H:MDM2:1	HT29	MDM2	4193	trt_sh.cgs	best inferred	1	Module 130 : T_CD (w = 0) , escore-pval-padj-zscore	Not_in_AMARETTO

Showing 1 to 17 of 17 entries (filtered from 55,753 total entries) Previous 1 Next

https://pochetlab.shinyapps.io/pAMARETTO_Brain_2DS_Drivers/ (under development)

AMARETTO report GBM

Summary of methylation-driven GBM Module 90:

Drivers: methylation-driven RBP1, PNPLA4, NSUN7, SLC25A20, FBXO17, XKR8, RAB36

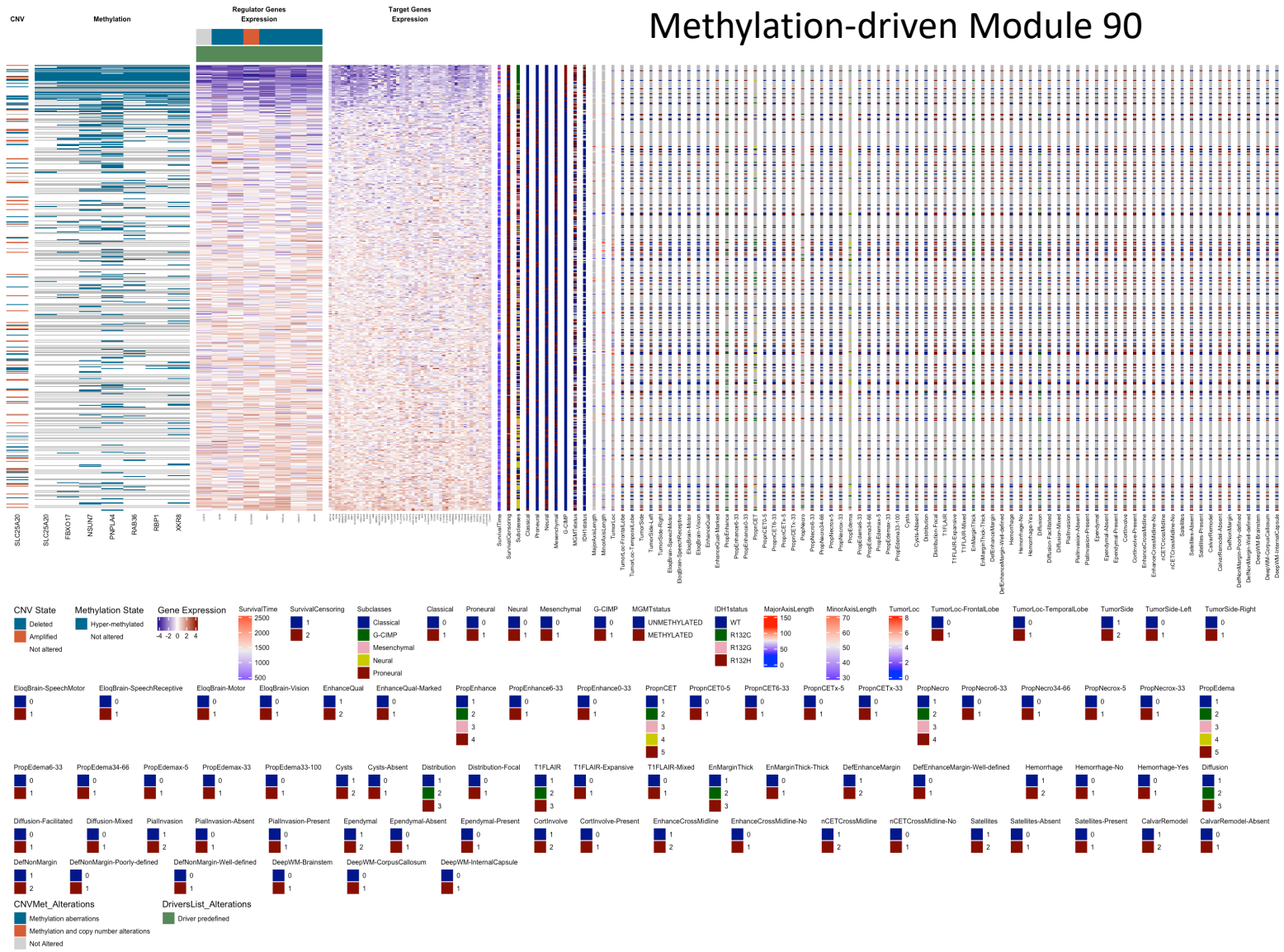
Hypermethylation of drivers, associated with their repressed gene expression levels

Drivers are activators of their targets

Associated with:
- Survival (lower expression, better survival)
- Molecular subclass G-CIMP (lower expression)
- Molecular markers IDH1 and MGMT

Enriched for NOUSHMEHR GBM SILENCED BY METHYLATION (Top 50 most differentially hypermethylated and down-regulated genes in proneural G-CIMP (a CpG island methylator phenotype) GBM (glioblastoma multiforme) tumors)

Methylation-driven Module 90



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_GBM/AMARETTOhtmls/modules/module90.html

AMARETTO report LGG

Summary of methylation-driven LGG Module 150:

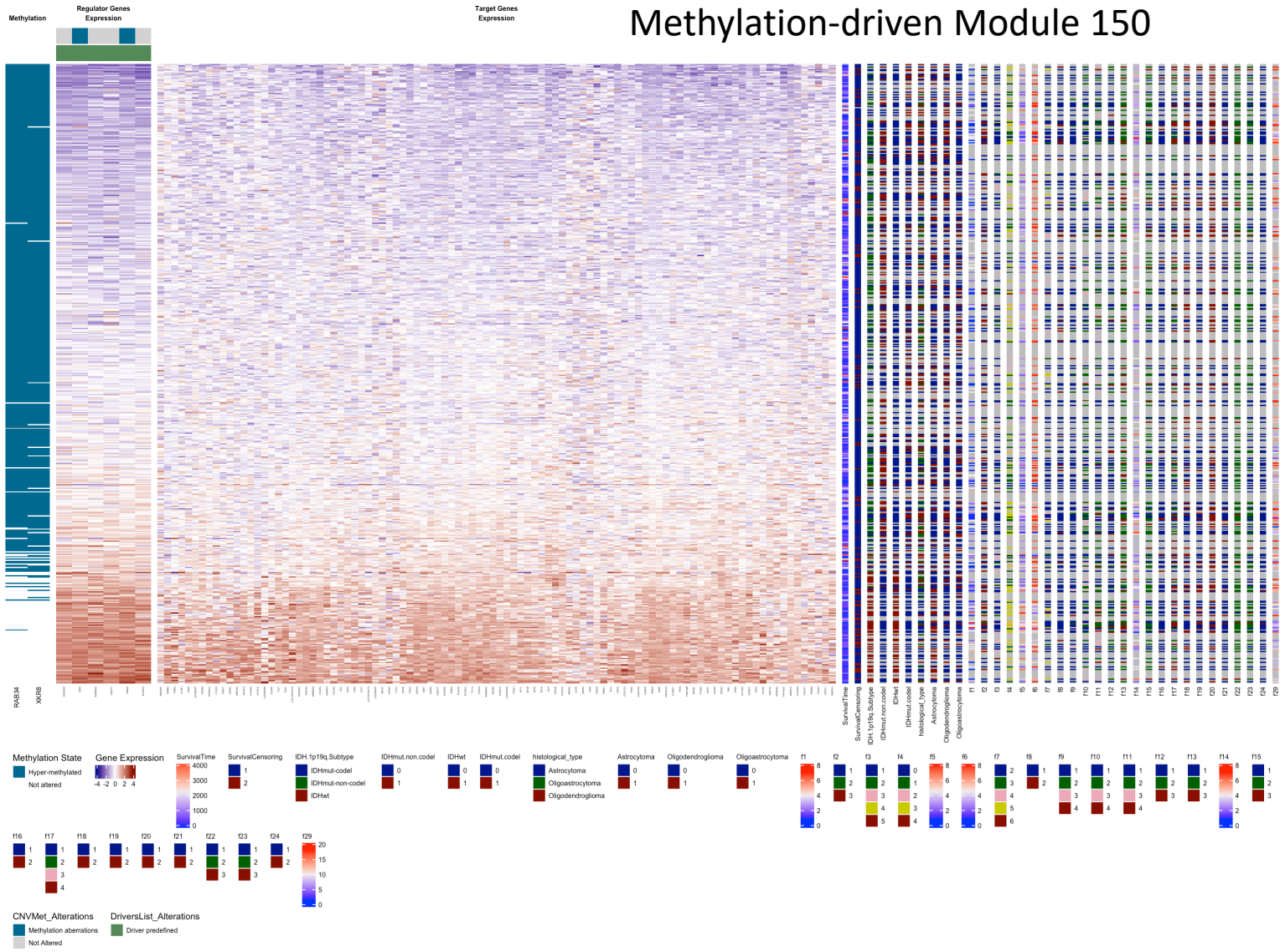
Drivers: methylation-driven, shared drivers with GBM module 90 (Community 1)

Hypermethylation of drivers, associated with their repressed gene expression levels

Drivers are activators of their targets

Associated with:
- Survival (lower expression, better survival)
- Imaging:
  - Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)
  - Proportion Enhancing (f5): higher expression, higher proportion of enhancing tumor

Methylation-driven Module 150



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_LGG/AMARETTOhtmls/modules/module150.html

AMARETTO report LGG

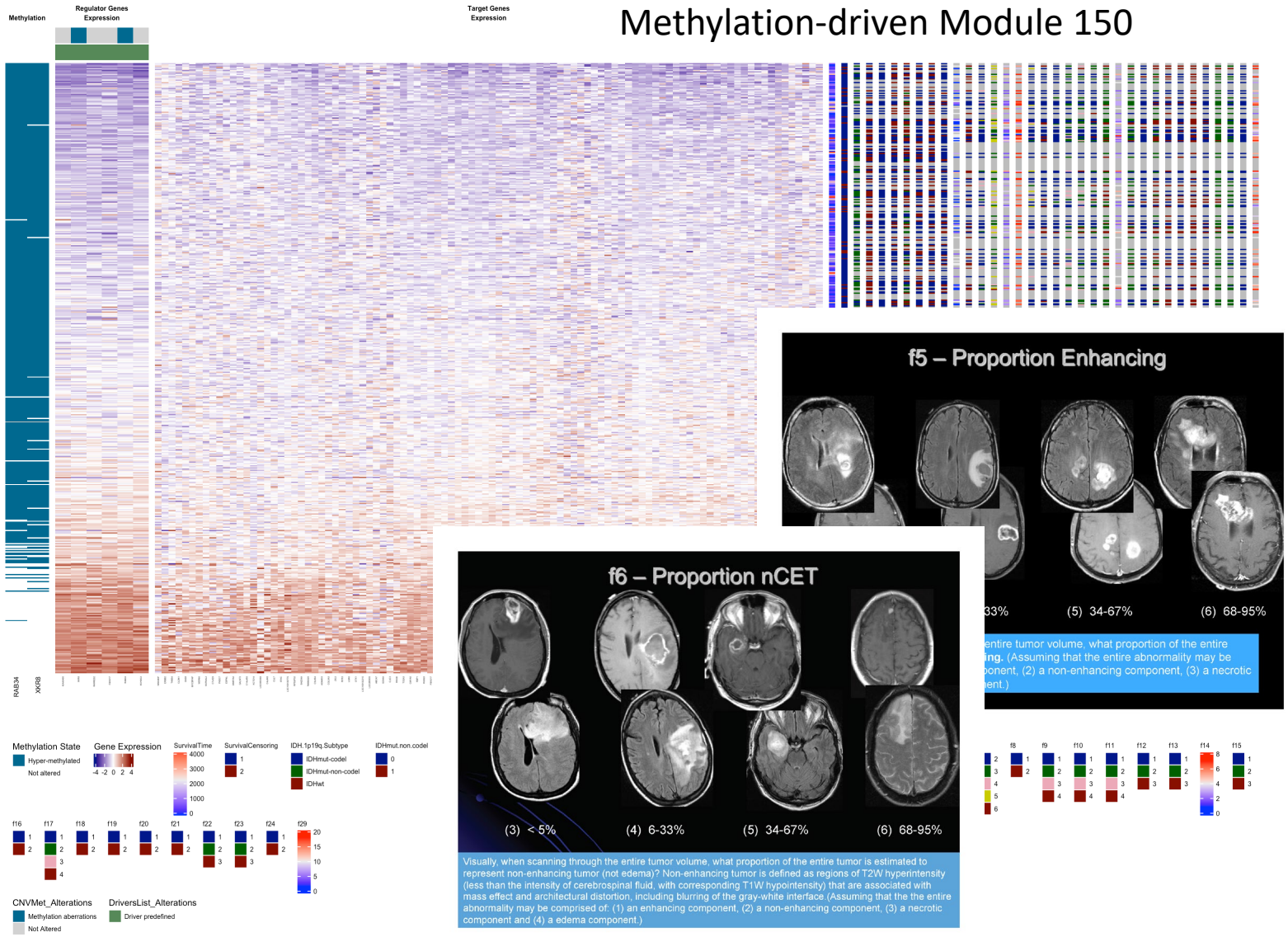
Summary of methylation-driven LGG Module 150:

Drivers: methylation-driven, shared drivers with GBM module 90 (Community 1)

Hypermethylation of drivers, associated with their repressed gene expression levels

Drivers are activators of their targets

Associated with:
- Survival (lower expression, better survival)
- Imaging:
  - Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)
  - Proportion Enhancing (f5): higher expression, higher proportion of enhancing tumor



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_LGG/AMARETTOhtmls/modules/module150.html

AMARETTO report GBM

Summary of methylation-driven GBM Module 77:

Drivers: methylation-driven

Hypermethylation of drivers, associated with their repressed gene expression levels

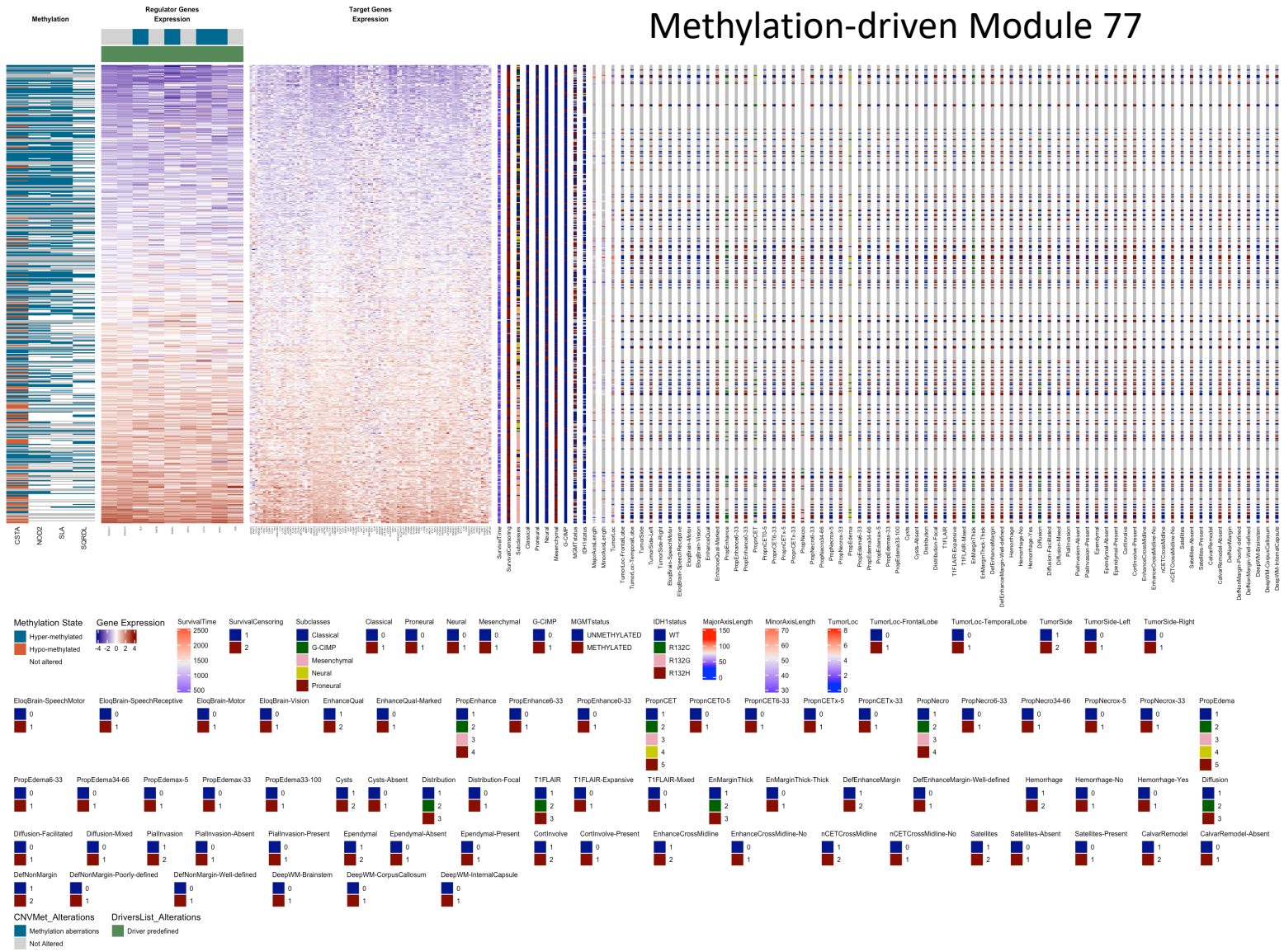
Drivers are activators of their targets

Associated with:

- Survival (higher expression, poorer survival)
- Molecular subclass Mesenchymal (higher expression)
- Molecular marker IDH1
- Imaging:
    - Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)

Enriched for VERHAAK GLIOBLASTOMA MESENCHYMAL (Genes correlated with mesenchymal type of glioblastoma multiforme tumors)

Methylation-driven Module 77



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_GBM/AMARETTOhtmls/modules/module77.html

AMARETTO report GBM

Summary of methylation-driven GBM Module 77:

Drivers: methylation-driven

Hypermethylation of drivers, associated with their repressed gene expression levels

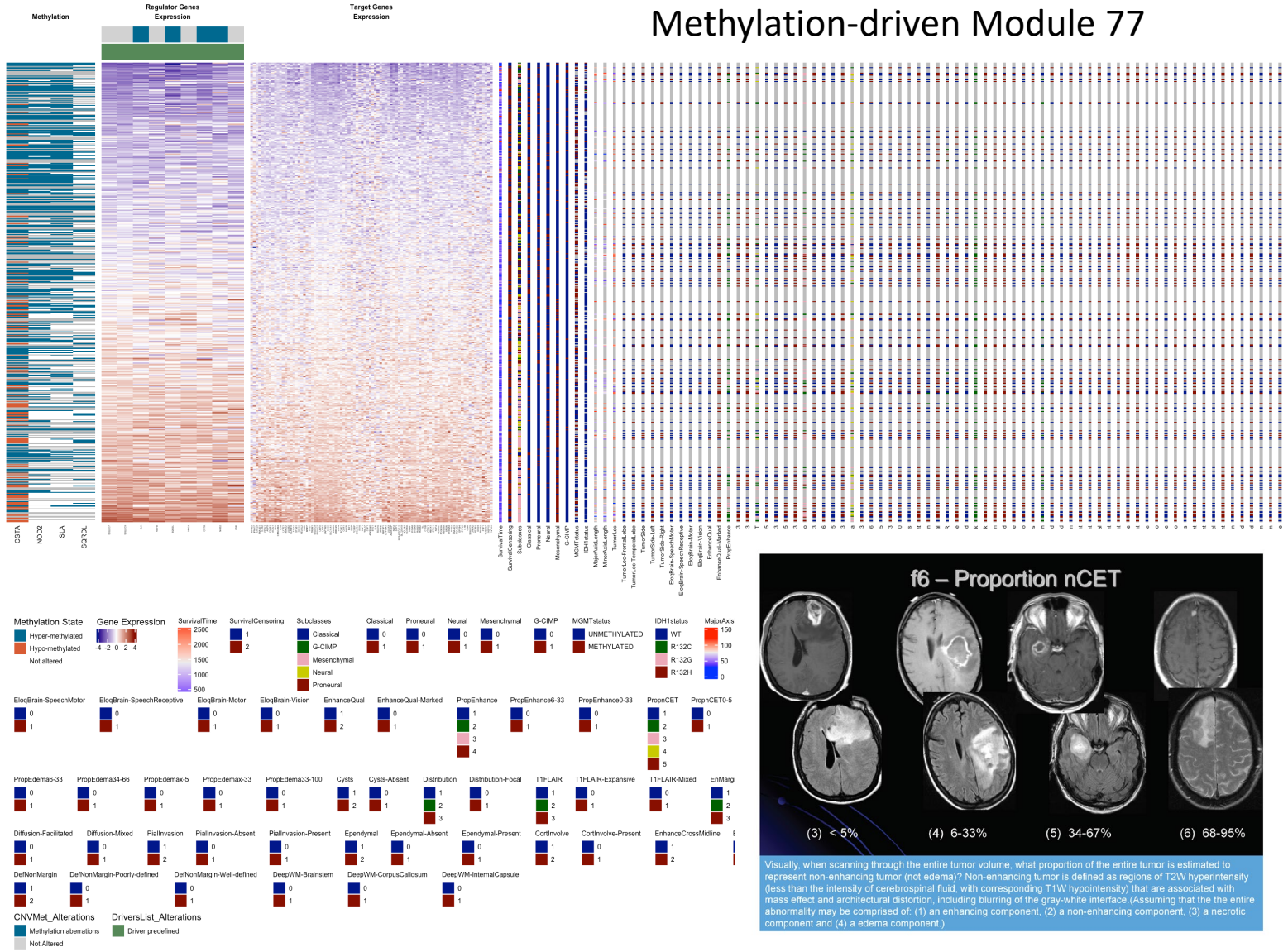
Drivers are activators of their targets

Associated with:

- Survival (higher expression, poorer survival)
- Molecular subclass Mesenchymal (higher expression)
- Molecular marker IDH1
- Imaging:
  - Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)

Enriched for VERHAAK GLIOBLASTOMA MESENCHYMAL (Genes correlated with mesenchymal type of glioblastoma multiforme tumors)

Methylation-driven Module 77

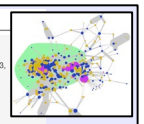


Visually, when scanning through the entire tumor volume, what proportion of the entire tumor is estimated to represent non-enhancing tumor (not edema)? Non-enhancing tumor is defined as regions of T2W hyperintensity (less than the intensity of cerebrospinal fluid, with corresponding T1W hypointensity) that are associated with mass effect and architectural distortion, including blurring of the gray-white interface.(Assuming that the the entire abnormality may be comprised of: (1) an enhancing component, (2) a non-enhancing component, (3) a necrotic component and (4) a edema component.)

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/TCGA_GBM/AMARETTOhtmls/modules/module77.html

Community-AMARETTO report GBM/LGG

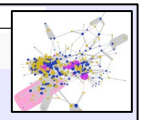
Community	TCGA_GBM	TCGA_LGG	ImmuneSignatures	StemSignatures
Community 1	Module 3, Module 6, Module 18, Module 22, Module 23, Module 24, Module 25, Module 27, Module 33, Module 37, Module 38, Module 39, Module 41, Module 43, Module 48, Module 54, Module 60, Module 62, Module 66, Module 67, Module 68, Module 71, Module 72, Module 75, Module 78, Module 80, Module 81, Module 82, Module 85, Module 86, Module 87, Module 88, Module 90, Module 91, Module 97, Module 98, Module 99, Module 102, Module 107, Module 110, Module 114, Module 116, Module 117, Module 120, Module 122, Module 129, Module 132, Module 137, Module 138, Module 142, Module 147, Module 150	Module 2, Module 7, Module 8, Module 16, Module 18, Module 19, Module 23, Module 26, Module 27, Module 30, Module 34, Module 35, Module 37, Module 40, Module 43, Module 44, Module 47, Module 50, Module 51, Module 53, Module 55, Module 59, Module 60, Module 61, Module 63, Module 64, Module 66, Module 68, Module 72, Module 79, Module 80, Module 81, Module 83, Module 84, Module 85, Module 86, Module 87, Module 88, Module 89, Module 93, Module 99, Module 102, Module 103, Module 104, Module 110, Module 112, Module 117, Module 120, Module 123, Module 127, Module 133, Module 134, Module 135, Module 137, Module 138, Module 141, Module 142, Module 147, Module 150	CIBERSORT_MAST_CELLS_RESTING	BENPORATH_EED_TARGETS, BENPORATH_ES_WITH_H3K27ME3, BENPORATH_NOS_TARGETS, BENPORATH_OCT4_TARGETS, BENPORATH_PRC2_TARGETS, BENPORATH_SUZ12_TARGETS



Community 1

- VERHAAK GLIOBLASTOMA PRONEURAL: Genes correlated with proneural type of glioblastoma multiforme tumors
- NOUSHMEHR GBM SILENCED BY METHYLATION: Top 50 most differentially hypermethylated and down-regulated genes in proneural G-CIMP (a CpG island methylator phenotype) GBM (glioblastoma multiforme) tumors
- Molecular markers: PDGFRA, IDH1 and MGMT
- Methylation-driven drivers shared between GBM (Module 90) and LGG (Module 77): FBXO17, XKR8, RAB34/36
- Survival analysis: hypermethylated drivers, repressed expression of drivers and targets, and association with better survival rates in the G-CIMP and proneural molecular subclasses

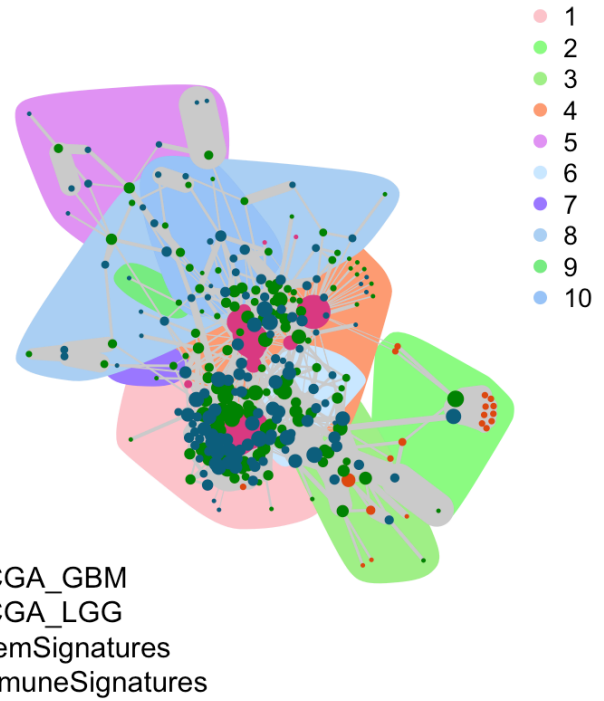
Community	TCGA_GBM	TCGA_LGG	ImmuneSignatures	StemSignatures
Community 3	Module 8, Module 55, Module 70, Module 77, Module 93, Module 125, Module 131, Module 133, Module 134, Module 141	Module 22, Module 33, Module 38, Module 73, Module 90, Module 118, Module 121	CIBERSORT_EOSINOPHILS, CIBERSORT_MACROPHAGES_M0, CIBERSORT_MACROPHAGES_M2, CIBERSORT_MONOCYTES, CIBERSORT_NEUTROPHILS	0



Community 3

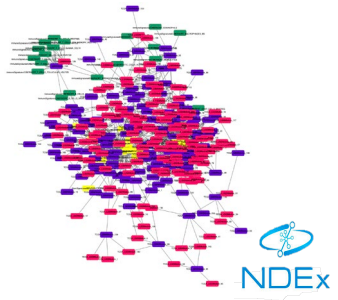
- VERHAAK GLIOBLASTOMA MESENCHYMAL: Genes correlated with mesenchymal type of glioblastoma multiforme tumors
- Enriched for immune cell type-specific signatures
- Associated with imaging feature Proportion nCET (f6): lower expression, higher proportion of non-enhancing tumor (not edema)
- Survival analysis: induced expression of drivers and targets and association with poorer survival rates in the mesenchymal molecular subclass

Community Network Visualization



Perturbation-AMARETTO (under development):

- driver validation & discovery: https://pochetlab.shinyapps.io/pAMARETTO_Brain_2DS_Drivers/
- drug discovery: https://pochetlab.shinyapps.io/pAMARETTO_Brain_2DS_Drivers



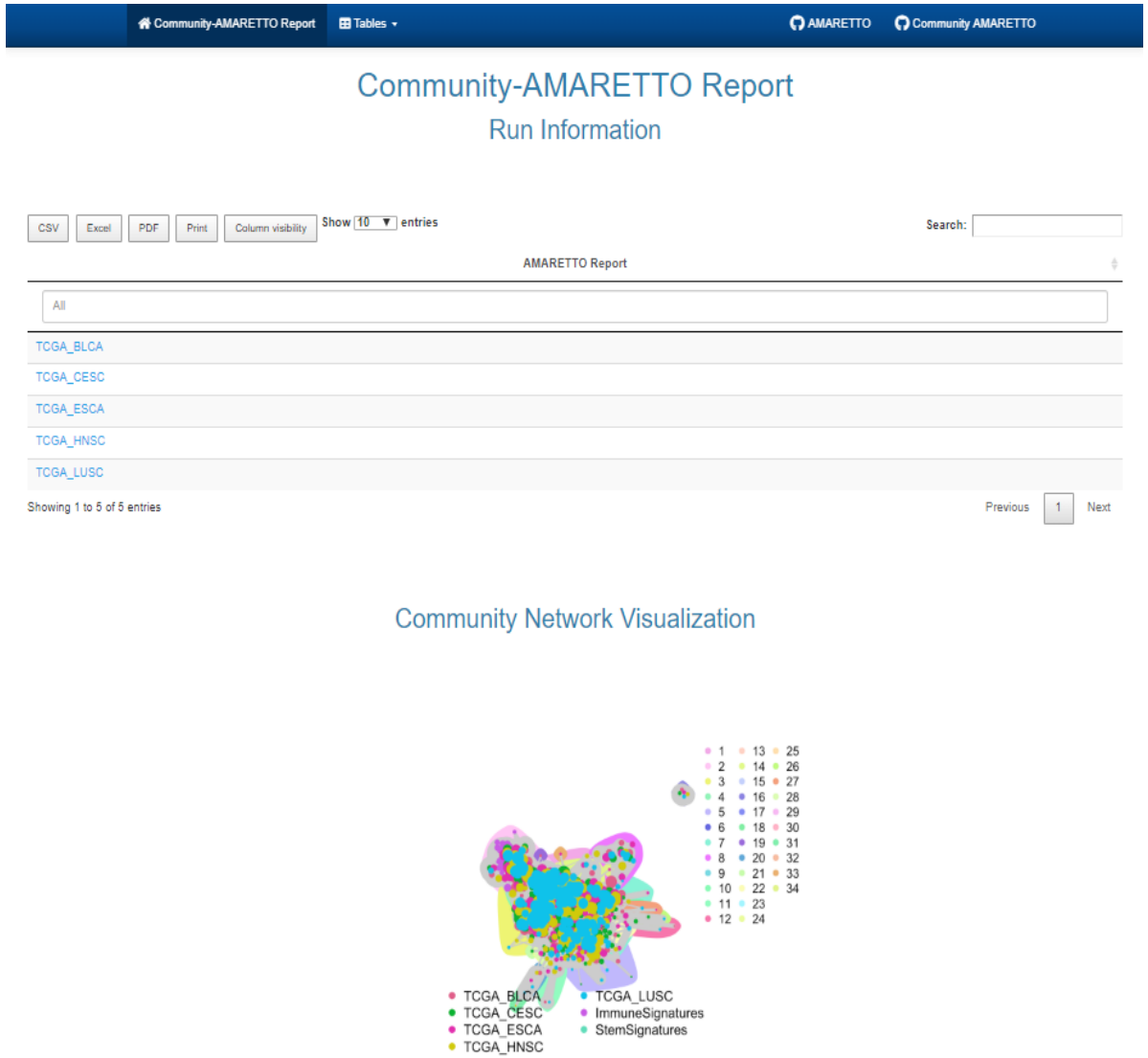
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_Brain_2DS/index.html

Case Study 3

Pan-squamous cell carcinoma (SCC)

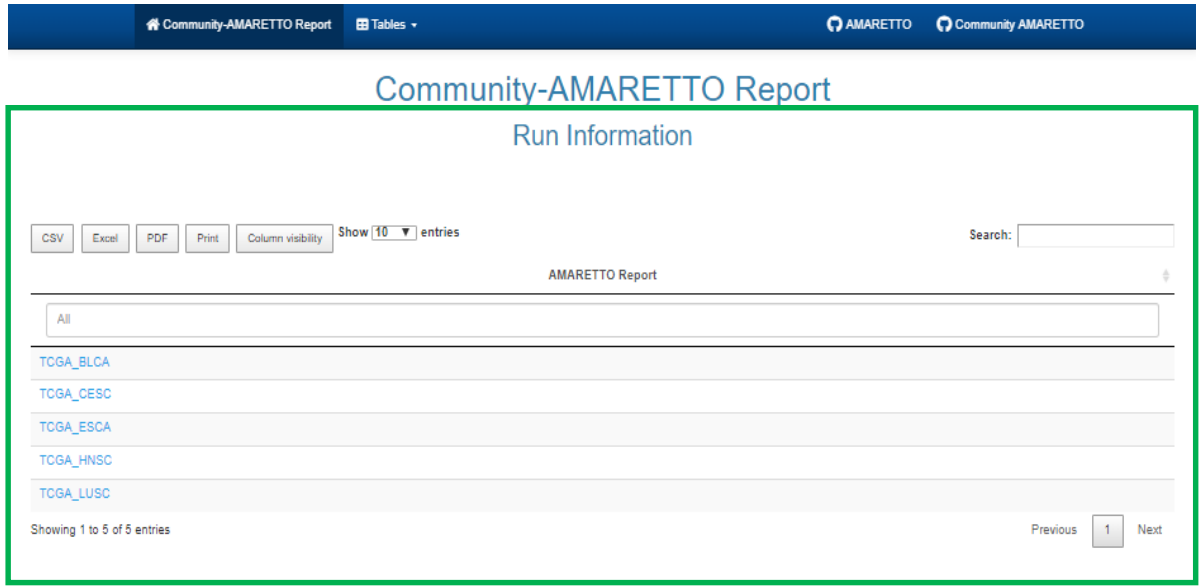
across 5 SCC cancer sites: lung (LUSC), head and neck (HNSC), esophageal (ESCA), cervical (CESC) and bladder (BLCA)

Community-AMARETTO report SCC



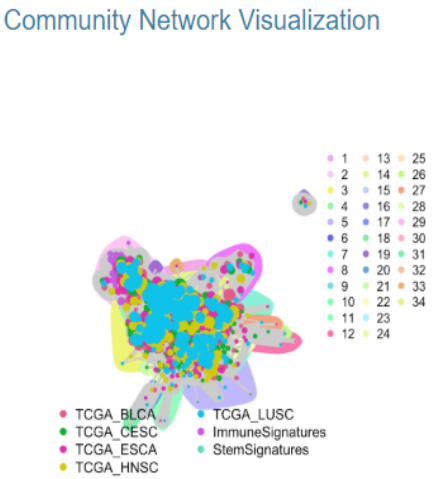
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index.html

Community-AMARETTO report SCC



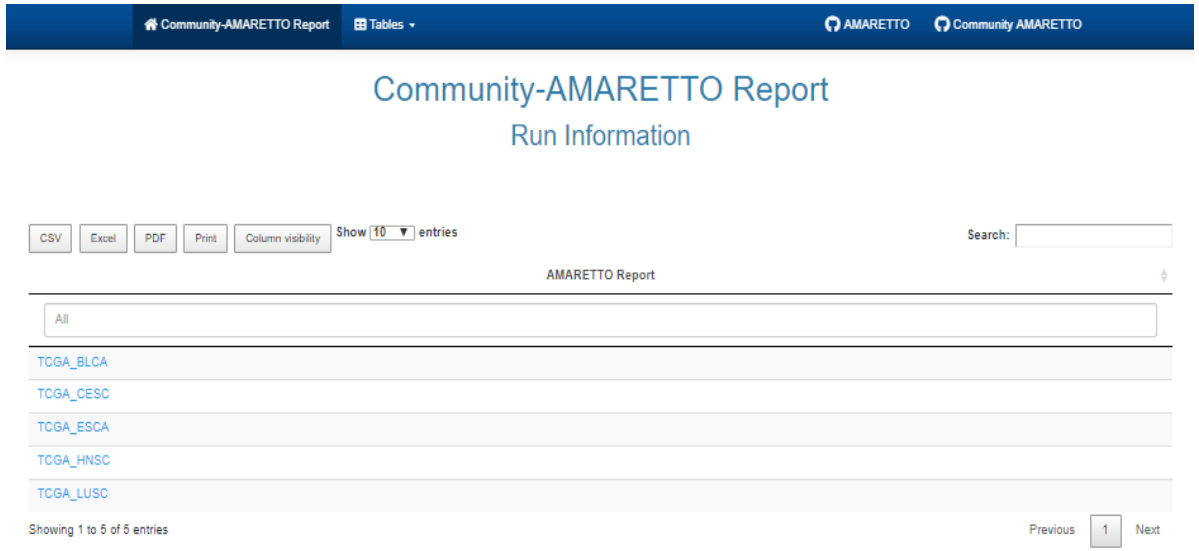
Run Information: links to AMARETTO reports combined in Community-AMARETTO report

Community Network Visualization



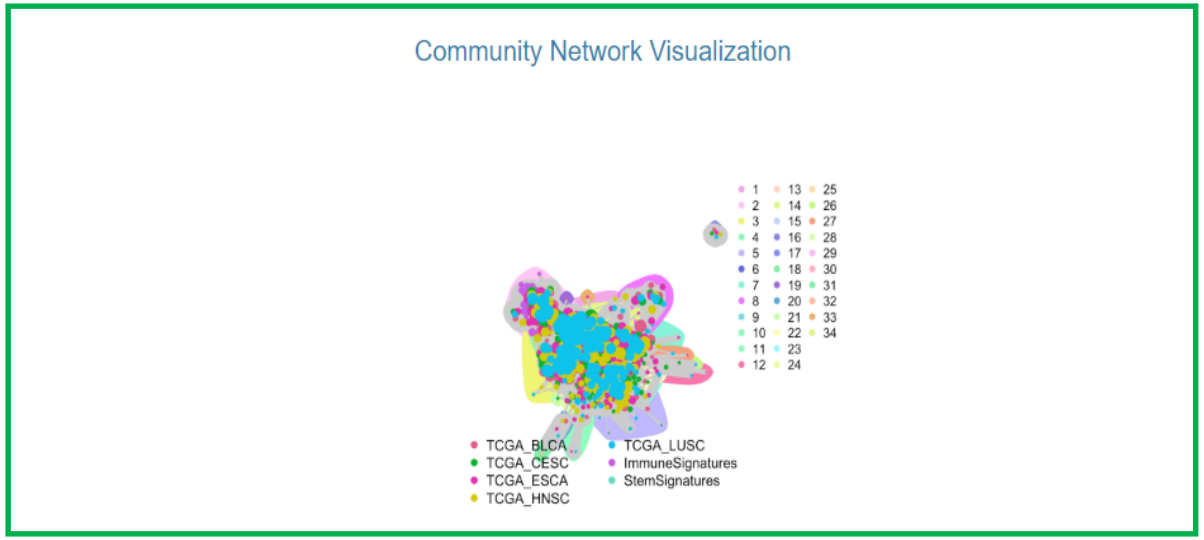
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index.html

Community-AMARETTO report SCC



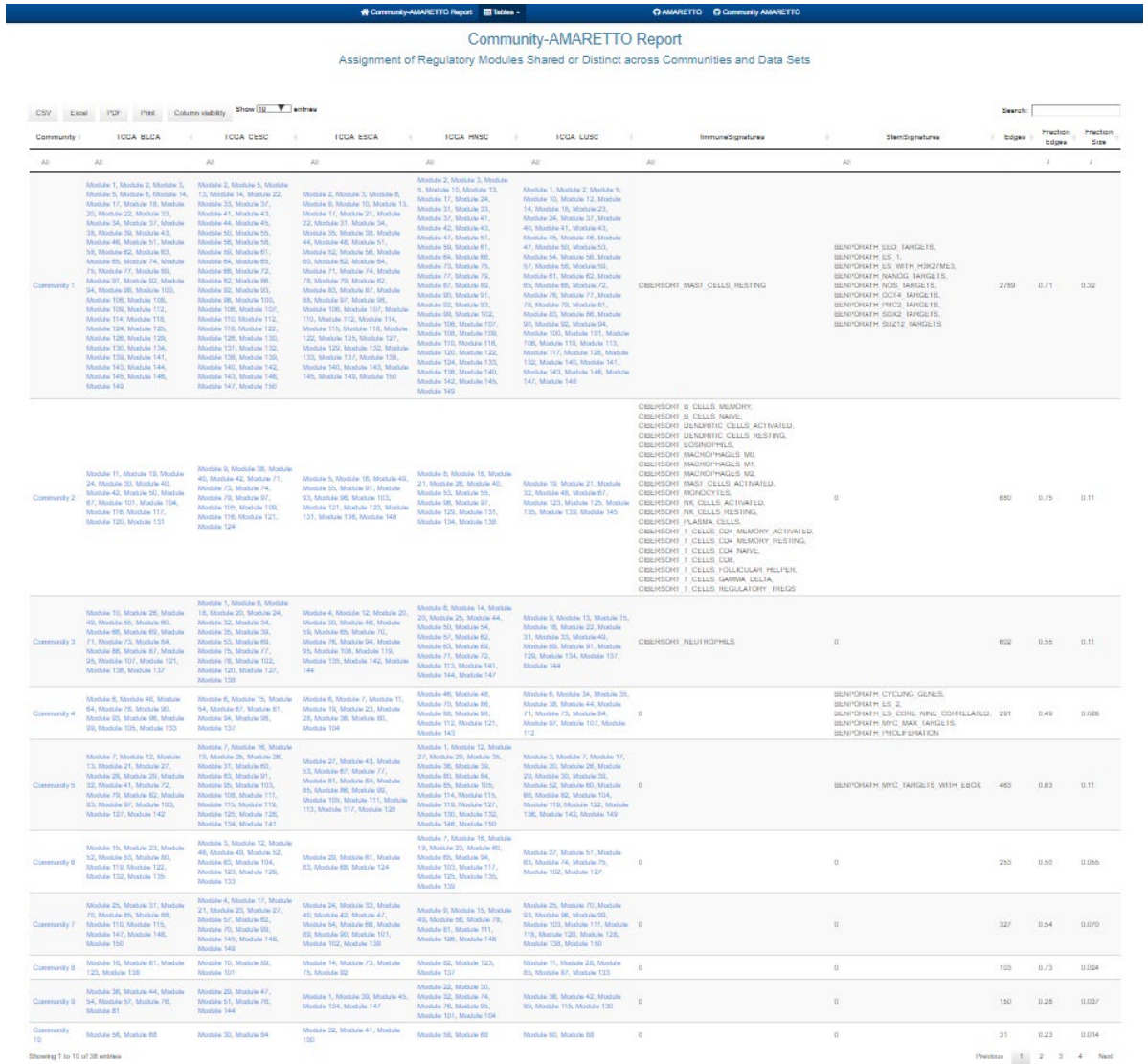
Run Information: links to AMARETTO reports combined in Community-AMARETTO report

Community Network Visualization



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index.html

Community-AMARETTO report SCC

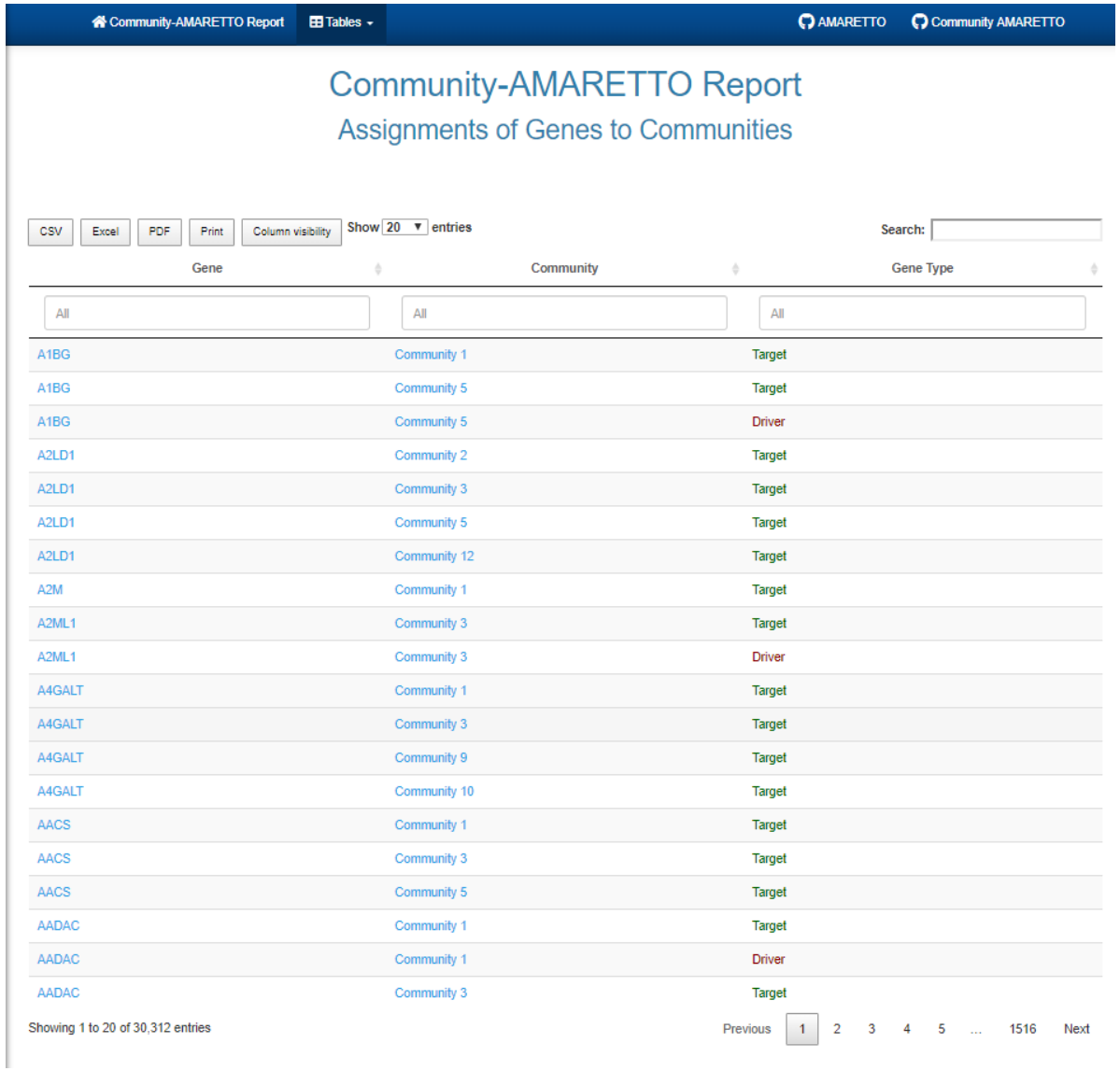


Assignments of Regulatory Modules Shared or Distinct across Communities and Data Sets

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_AllCommunities.html

Community-AMARETTO report SCC

Assignments of Genes to Communities



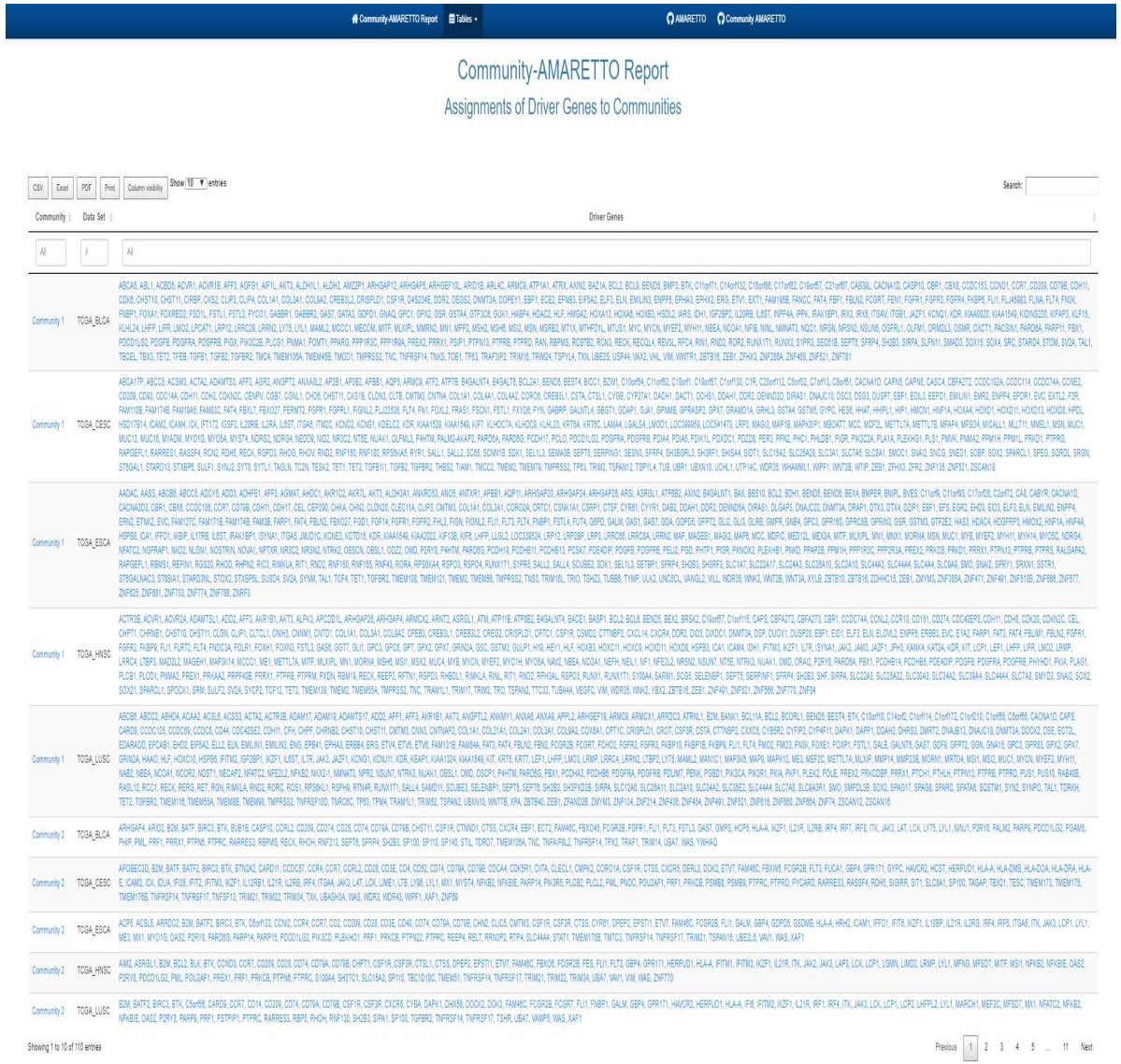
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_AllGenes.html

Community-AMARETTO report SCC

Assignments of Driver Genes Shared or Distinct across Communities and Data Sets

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_Drivers.html

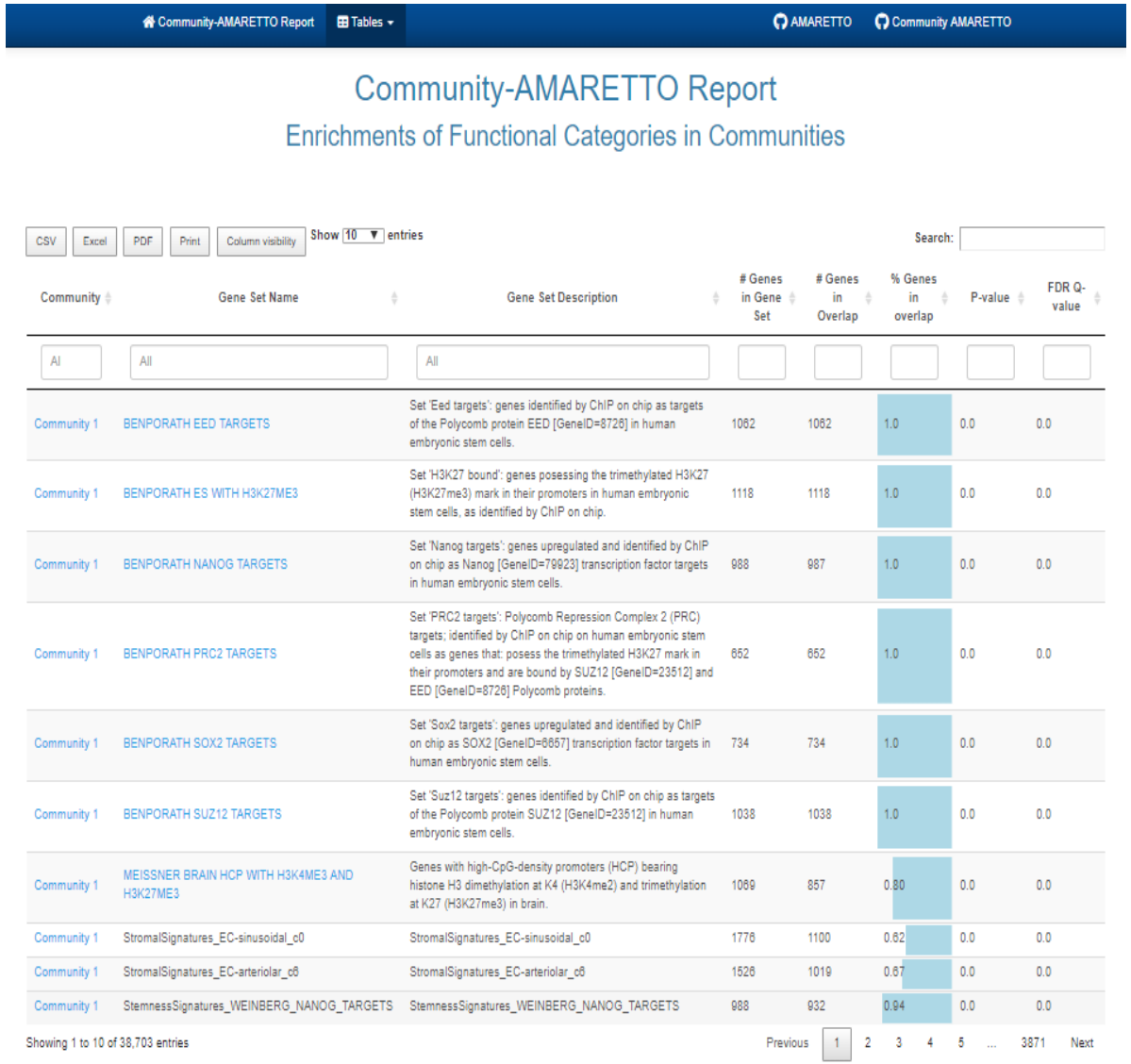
Community-AMARETTO report SCC



Assignments of Driver Genes to Communities

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_Drivers2.html

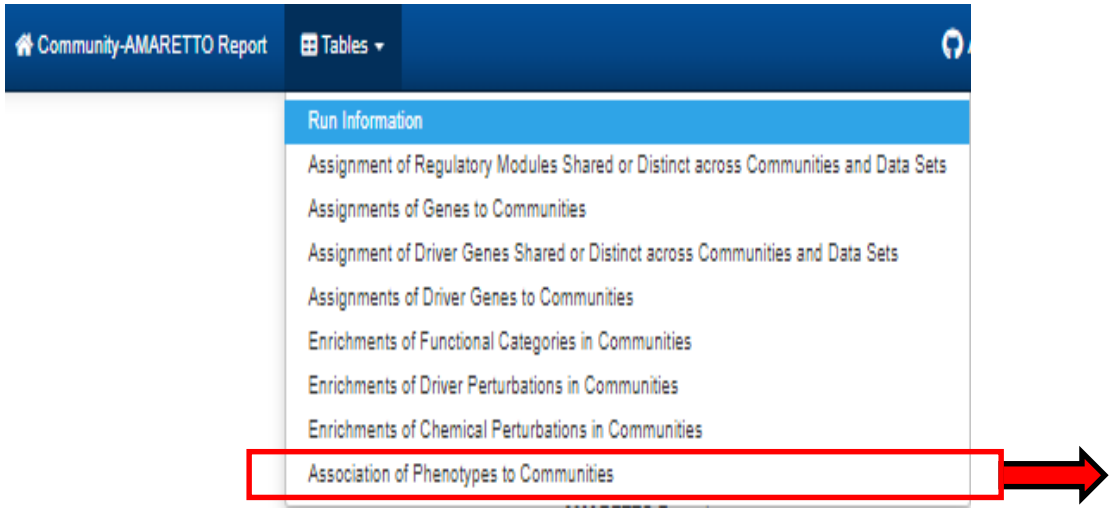
Community-AMARETTO report SCC



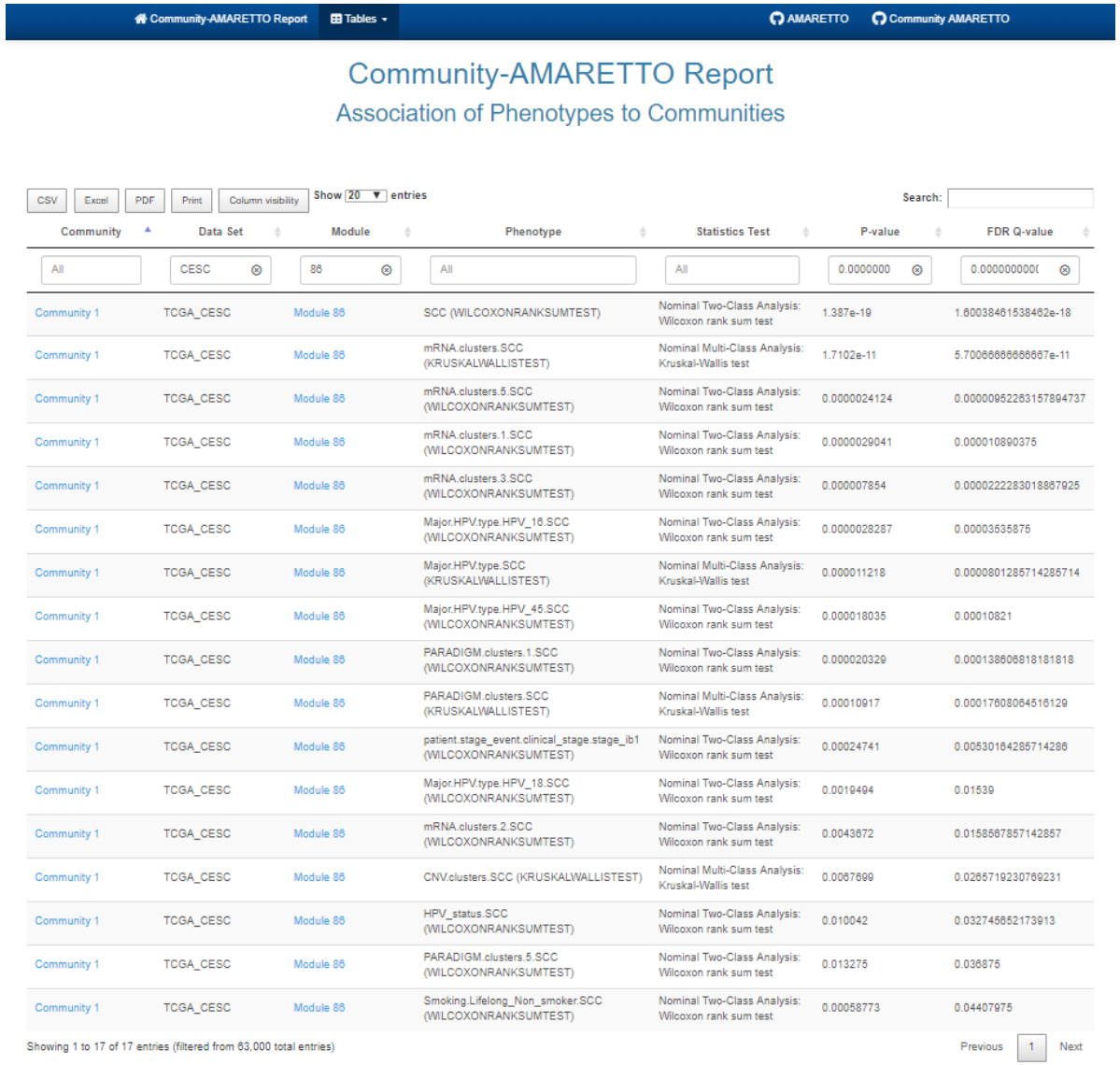
Enrichments of Functional Categories in Communities

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_GenesetsEnrichment.html

Community-AMARETTO report SCC



Association of Phenotypes to Communities



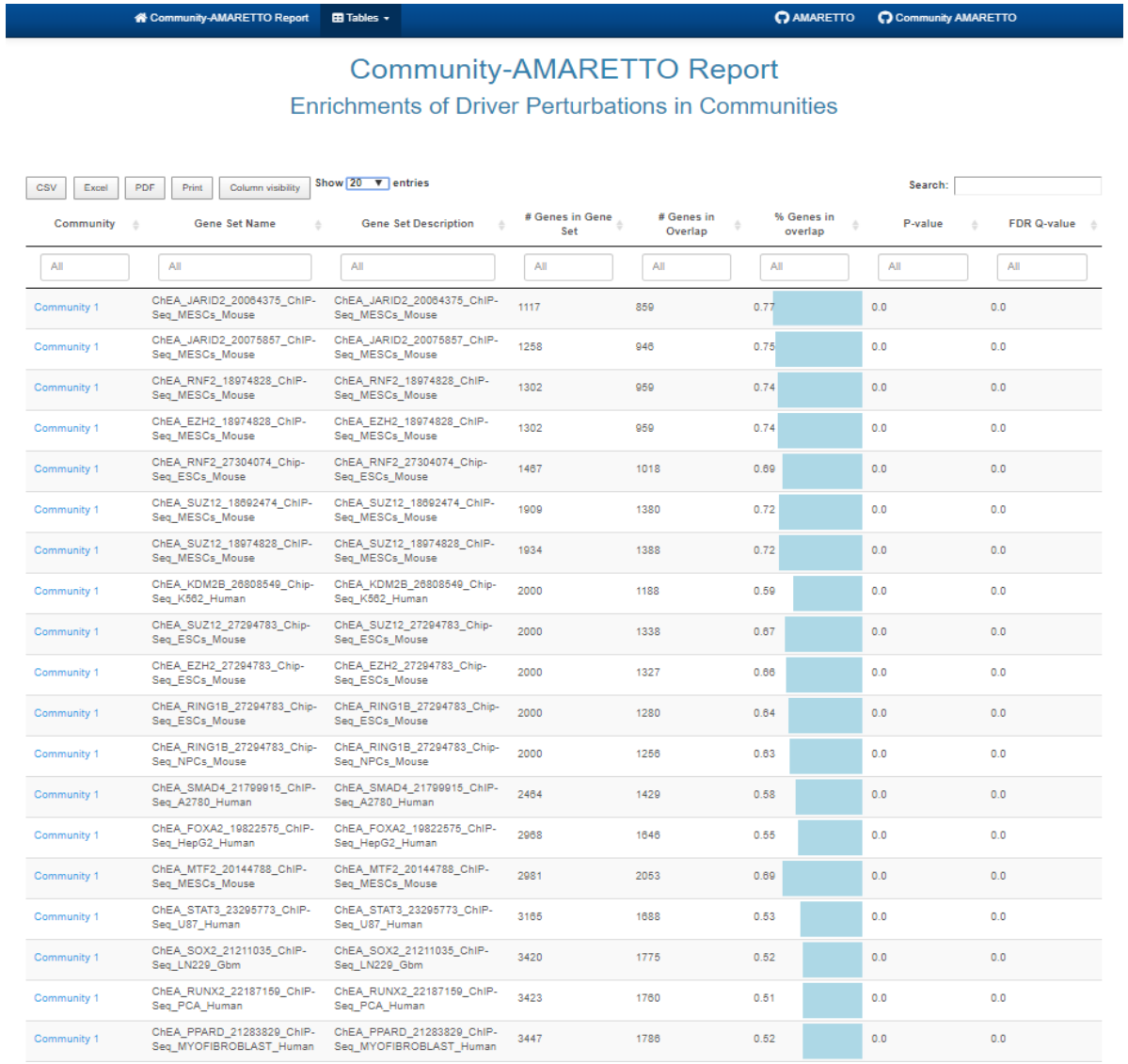
Community	Data Set	Module	Phenotype	Statistics Test	P-value	FDR Q-value
All	CESC	86	All	All	0.0000000	0.000000000(
Community 1	TCGA_CESC	Module 86	SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	1.387e-19	1.6003846153846 2e-18
Community 1	TCGA_CESC	Module 86	mRNA.clusters.SCC (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	1.7102e-11	5.70066666666667e-11
Community 1	TCGA_CESC	Module 86	mRNA.clusters.5.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0000024124	0.000009522631578947 37
Community 1	TCGA_CESC	Module 86	mRNA.clusters.1.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0000029041	0.000010890375
Community 1	TCGA_CESC	Module 86	mRNA.clusters.3.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.000007854	0.000022228301886792 5
Community 1	TCGA_CESC	Module 86	Major.HPV.type.HPV_16.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0000028287	0.00003535875
Community 1	TCGA_CESC	Module 86	Major.HPV.type.SCC (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.000011218	0.000080128571428571 4
Community 1	TCGA_CESC	Module 86	Major.HPV.type.HPV_45.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.000018035	0.00010821
Community 1	TCGA_CESC	Module 86	PARADIGM.clusters.1.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.000020329	0.000138068181818 18
Community 1	TCGA_CESC	Module 86	PARADIGM.clusters.SCC (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.00010917	0.000176080645161 29
Community 1	TCGA_CESC	Module 86	patient.stage_event.clinical_stage.stage_ib1 (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00024741	0.005301642857142 86
Community 1	TCGA_CESC	Module 86	Major.HPV.type.HPV_18.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0019494	0.01539
Community 1	TCGA_CESC	Module 86	mRNA.clusters.2.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.0043672	0.015856785714285 7
Community 1	TCGA_CESC	Module 86	CNV.clusters.SCC (KRUSKALWALLISTEST)	Nominal Multi-Class Analysis: Kruskal-Wallis test	0.0067699	0.026571923076923 1
Community 1	TCGA_CESC	Module 86	HPV_status.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.010042	0.032745652173913
Community 1	TCGA_CESC	Module 86	PARADIGM.clusters.5.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.013275	0.036875
Community 1	TCGA_CESC	Module 86	Smoking.Lifelong_Non_smoker.SCC (WILCOXONRANKSUMTEST)	Nominal Two-Class Analysis: Wilcoxon rank sum test	0.00058773	0.04407975

Showing 1 to 17 of 17 entries (filtered from 63,000 total entries)

Previous 1 Next

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_PhenoAssociation.html

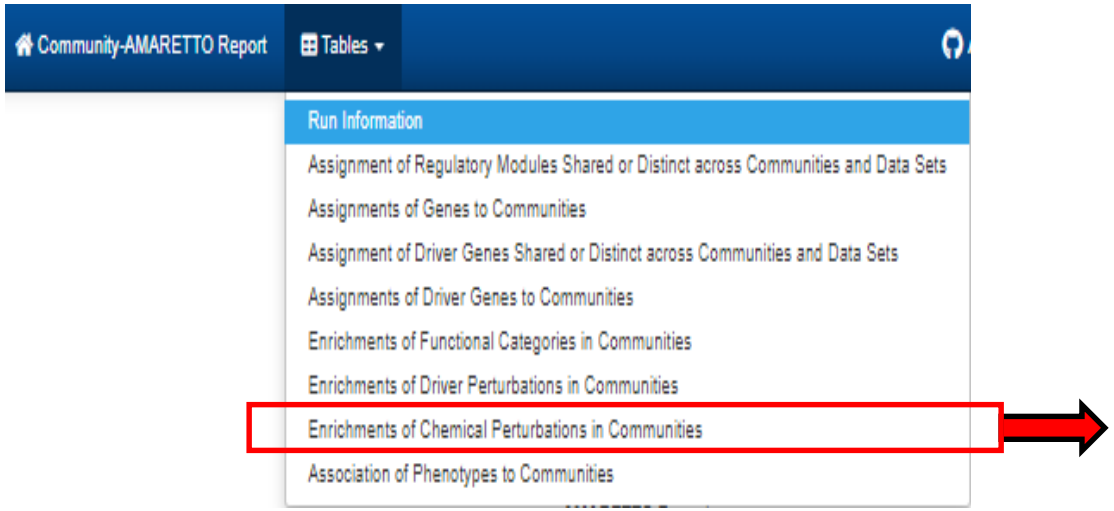
Community-AMARETTO report SCC



Enrichments of Driver Perturbations in Communities

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_GenesetsEnrichment_gp.html

Community-AMARETTO report SCC



Enrichments of Drug Perturbations in Communities

Community-AMARETTO Report

Enrichments of Chemical Perturbations in Communities

CSV | Excel | PDF | Print | Column visibility | Show 20 entries | Search:

Community	Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HME1_24H-palbociclib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HME1_24H-palbociclib-10_DN	215	185	0.86	4.8e-191	1.6e-186
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HME1_24H-palbociclib-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HME1_24H-palbociclib-3.33_DN	183	169	0.92	5.5e-187	9.0e-183
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_PC3_24H-NVP-TAE684-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_PC3_24H-NVP-TAE684-10_DN	186	164	0.88	3.2e-173	3.5e-169
Community 4	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-dovitinib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-dovitinib-10_DN	215	175	0.81	7.2e-172	5.8e-168
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_LNCAP_24H-GDC-0941-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP006_LNCAP_24H-GDC-0941-3.33_DN	185	161	0.87	6.0e-168	3.9e-164
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_LNCAP_24H-mitoxantrone-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP006_LNCAP_24H-mitoxantrone-0.37_DN	195	165	0.85	1.0e-167	5.5e-164
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_A375_24H-mitoxantrone-0.12_DN	LINCSCMAP_ChemicalPerturbation_LJP006_A375_24H-mitoxantrone-0.12_DN	222	175	0.79	3.6e-167	1.7e-163
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HCC515_24H-PHA-793887-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HCC515_24H-PHA-793887-3.33_DN	186	160	0.86	4.2e-165	1.7e-161
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_BT20_24H-PHA-793887-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP006_BT20_24H-PHA-793887-3.33_DN	184	157	0.85	8.5e-161	3.1e-157
Community 4	LINCSCMAP_ChemicalPerturbation_LJP005_MCF10A_24H-mitoxantrone-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP005_MCF10A_24H-mitoxantrone-0.37_DN	192	158	0.82	1.2e-156	4.0e-153
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_A375_24H-palbociclib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_A375_24H-palbociclib-10_DN	154	141	0.92	1.5e-154	4.3e-151
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HT29_24H-palbociclib-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HT29_24H-palbociclib-0.37_DN	195	158	0.81	1.6e-154	4.3e-151
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_MCF10A_24H-mitoxantrone-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP006_MCF10A_24H-mitoxantrone-3.33_DN	181	152	0.84	1.9e-153	4.9e-150
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HS578T_24H-palbociclib-1.11_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HS578T_24H-palbociclib-1.11_DN	156	141	0.90	1.5e-152	3.5e-149
Community 4	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-torin-2-0.12_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-torin-2-0.12_DN	180	151	0.84	2.8e-152	6.0e-149
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_BT20_24H-palbociclib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_BT20_24H-palbociclib-10_DN	168	146	0.87	4.4e-152	8.9e-149
Community 4	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-NVP-BEZ235-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A549_24H-NVP-BEZ235-0.37_DN	207	161	0.78	5.1e-152	9.8e-149
Community 4	LINCSCMAP_ChemicalPerturbation_LJP006_HS578T_24H-palbociclib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP006_HS578T_24H-palbociclib-10_DN	152	137	0.90	8.2e-148	1.5e-144

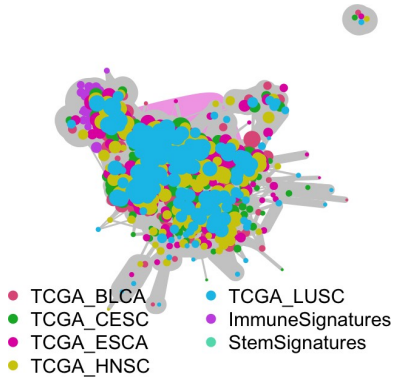
http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/index_GenesetsEnrichment_cp.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

Community-AMARETTO Community 1

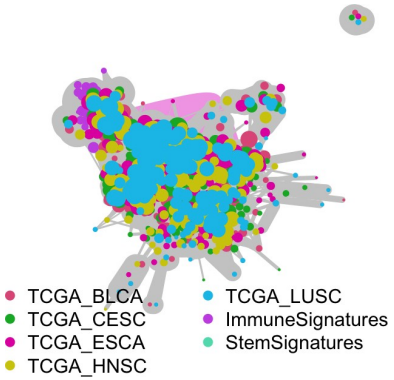


Data Set	Module	Gene	Gene Type
TCGA_BLCA	Module 8	SOX2	Target
TCGA_CESC	Module 14	SOX2	Target
TCGA_CESC	Module 14	SOX2	Driver
TCGA_ESCA	Module 83	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Driver
TCGA_HNSC	Module 106	SOX2	Driver
TCGA_LUSC	Module 18	SOX2	Target
TCGA_LUSC	Module 18	SOX2	Driver
TCGA_LUSC	Module 37	SOX2	Driver
TCGA_LUSC	Module 41	SOX2	Driver
TCGA_LUSC	Module 61	SOX2	Driver
TCGA_LUSC	Module 94	SOX2	Driver

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

Community-AMARETTO Community 1



Enrichments of Functional Categories in Community

Show 5 entries Search: SOX2

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
TARGETS	Polycomb protein EED [GeneID=8726] in human embryonic stem cells.					
BENPORATH NANOG TARGETS	Set 'Nanog targets': genes upregulated and identified by ChIP on chip as Nanog [GeneID=79923] transcription factor targets in human embryonic stem cells.	988	987	1.0	0.0	0.0
BENPORATH SOX2 TARGETS	Set 'Sox2 targets': genes upregulated and identified by ChIP on chip as SOX2 [GeneID=6657] transcription factor targets in human embryonic stem cells.	734	734	1.0	0.0	0.0
BENPORATH SUZ12 TARGETS	Set 'Suz12 targets': genes identified by ChIP on chip as targets of the Polycomb protein SUZ12 [GeneID=23512] in human embryonic stem cells.	1038	1038	1.0	0.0	0.0
MEISSNER BRAIN HCP WITH H3K4ME3 AND H3K27ME3	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation at K4 (H3K4me2) and trimethylation at K27 (H3K27me3) in brain.	1069	857	0.80	0.0	0.0

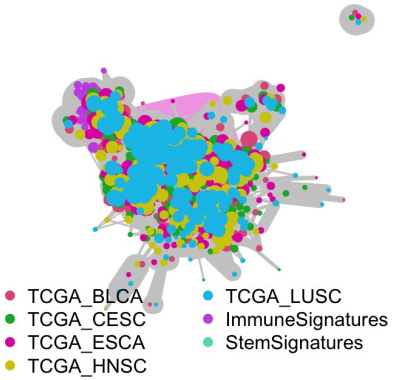
Showing 1 to 5 of 115 entries (filtered from 4,907 total entries) Previous 1 2 3 4 5 … 23 Next

Data Set	Module	Gene	Gene Type
TCGA_BLCA	Module 8	SOX2	Target
TCGA_CESC	Module 14	SOX2	Target
TCGA_CESC	Module 14	SOX2	Driver
TCGA_ESCA	Module 83	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Driver
TCGA_HNSC	Module 106	SOX2	Driver
TCGA_LUSC	Module 18	SOX2	Target
TCGA_LUSC	Module 18	SOX2	Driver
TCGA_LUSC	Module 37	SOX2	Driver
TCGA_LUSC	Module 41	SOX2	Driver
TCGA_LUSC	Module 61	SOX2	Driver
TCGA_LUSC	Module 94	SOX2	Driver

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

Community-AMARETTO Community 1



Enrichments of Functional Categories in Community

Show 5 entries Search: SOX2

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value
TARGETS	Polycomb protein EED [GeneID=8726] in human embryonic stem cells.				
BENPORATH NANOG TARGETS	Set 'Nanog targets': genes upregulated and identified by ChIP on chip as Nanog [GeneID=79923] transcription factor targets in human embryonic stem cells.	988	987	1.0	0.0
BENPORATH SOX2 TARGETS	Set 'Sox2 targets': genes upregulated and identified by ChIP on chip as SOX2 [GeneID=6657] transcription factor targets in human embryonic stem cells.	734	734	1.0	0.0
BENPORATH SUZ12 TARGETS	Set 'Suz12 targets': genes identified by ChIP on chip as targets of the Polycomb protein SUZ12 [GeneID=23512] in human embryonic stem cells.	1038	1038	1.0	0.0
MEISSNER BRAIN HCP WITH H3K4ME3 AND H3K27ME3	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation at K4 (H3K4me2) and trimethylation at K27 (H3K27me3) in brain.	1069	857	0.80	0.0

Showing 1 to 5 of 115 entries (filtered from 4,907 total entries) Previous 1 2 3 4 5 ...

Enrichments of Driver Perturbations in Community

CSV Excel PDF Print Column visibility Show 10 entries Search:

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
ChEA_SOX2_21211035_ChIP-	ChEA_SOX2_21211035_ChIP-	3420	1775	0.52	0.0	0.0
ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	2564	1343	0.52	3.0e-292	5.3e-291
Consensus_SOX2_CHEA	Consensus_SOX2_CHEA	775	556	0.72	5.8e-208	3.0e-206
ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	1278	755	0.59	4.7e-202	3.6e-201
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-	3319	1408	0.42	2.5e-190	1.5e-189
ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human	ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human	2000	940	0.47	6.8e-159	2.6e-158
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-	1991	937	0.47	7.7e-159	2.9e-158
ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse	ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse	2000	922	0.46	6.2e-149	1.9e-148
ChEA_SOX2_18555785_Chip-	ChEA_SOX2_18555785_Chip-	2000	900	0.45	3.5e-137	1.0e-136
ChEA_SOX2_21211035_ChIP-Seq_LN229_Human	ChEA_SOX2_21211035_ChIP-Seq_LN229_Human	2000	875	0.44	2.3e-124	5.2e-124

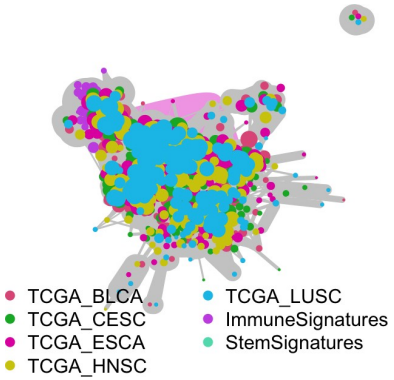
Showing 1 to 10 of 14 entries (filtered from 9,045 total entries) Previous 1 2 Next

Data Set	Module	Gene	Gene Type
TCGA_BLCA	Module 8	SOX2	Target
TCGA_CESC	Module 14	SOX2	Target
TCGA_CESC	Module 14	SOX2	Driver
TCGA_ESCA	Module 83	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Driver
TCGA_HNSC	Module 106	SOX2	Driver
TCGA_LUSC	Module 18	SOX2	Target
TCGA_LUSC	Module 18	SOX2	Driver
TCGA_LUSC	Module 37	SOX2	Driver
TCGA_LUSC	Module 41	SOX2	Driver
TCGA_LUSC	Module 61	SOX2	Driver
TCGA_LUSC	Module 94	SOX2	Driver

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

Community-AMARETTO Community 1



Enrichments of Functional Categories in Community

Show 5 entries Search: SOX2

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value
TARGETS	Polycomb protein EED [GeneID=8726] in human embryonic stem cells.				
BENPORATH NANOG TARGETS	Set 'Nanog targets': genes upregulated and identified by ChIP on chip as Nanog [GeneID=79923] transcription factor targets in human embryonic stem cells.	988	987	1.0	0.0
BENPORATH SOX2 TARGETS	Set 'Sox2 targets': genes upregulated and identified by ChIP on chip as SOX2 [GeneID=6657] transcription factor targets in human embryonic stem cells.	734	734	1.0	0.0
BENPORATH SUZ12 TARGETS	Set 'Suz12 targets': genes identified by ChIP on chip as targets of the Polycomb protein SUZ12 [GeneID=23512] in human embryonic stem cells.	1038	1038	1.0	0.0
MEISSNER BRAIN HCP WITH H3K4ME3 AND H3K27ME3	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation at K4 (H3K4me2) and trimethylation at K27 (H3K27me3) in brain.	1069	857	0.80	0.0

Showing 1 to 5 of 115 entries (filtered from 4,907 total entries) Previous 1 2 3 4 5 …

Enrichments of Driver Perturbations in Community

CSV Excel PDF Print Column visibility Show 10 entries Search:

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
ChEA_SOX2_21211035_ChIP-	ChEA_SOX2_21211035_ChIP-	3420	1775	0.52	0.0	0.0
ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	2564	1343	0.52	3.0e-292	5.3e-291
Consensus_SOX2_CHEA	Consensus_SOX2_CHEA	775	556	0.72	5.8e-208	3.0e-206
ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	1278	755	0.59	4.7e-202	3.6e-201
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-	3319	1408	0.42	2.5e-190	1.5e-189
ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human	ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human	2000	940	0.47	6.8e-159	2.6e-158
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-	1991	937	0.47	7.7e-159	2.9e-158
ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse	ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse	2000	922	0.46	6.2e-149	1.9e-148
ChEA_SOX2_18555785_Chip-	ChEA_SOX2_18555785_Chip-	2000	900	0.45	3.5e-137	1.0e-136
ChEA_SOX2_21211035_ChIP-Seq_LN229_Human	ChEA_SOX2_21211035_ChIP-Seq_LN229_Human	2000	875	0.44	2.3e-124	5.2e-124

Showing 1 to 10 of 14 entries (filtered from 9,045 total entries) Previous 1 2 Next

Data Set	Module	Gene	Gene Type
TCGA_BLCA	Module 8	SOX2	Target
TCGA_CESC	Module 14	SOX2	Target
TCGA_CESC	Module 14	SOX2	Driver
TCGA_ESCA	Module 83	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Driver
TCGA_HNSC	Module 106	SOX2	Driver
TCGA_LUSC	Module 18	SOX2	Target
TCGA_LUSC	Module 18	SOX2	Driver
TCGA_LUSC	Module 37	SOX2	Driver
TCGA_LUSC	Module 41	SOX2	Driver
TCGA_LUSC	Module 61	SOX2	Driver
TCGA_LUSC	Module 94	SOX2	Driver

Enrichments of Chemical Perturbations in Community

CSV Excel PDF Print Column visibility Show 10 entries Search: SOX2

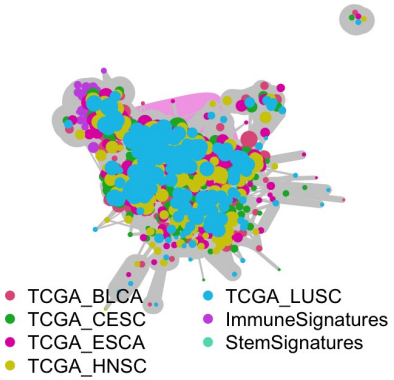
Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-buparlisib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-buparlisib-10_DN	147	111	0.76	7.6e-46	4.8e-43
LINCSCMAP_ChemicalPerturbation_IGF1-MCF7_UP	LINCSCMAP_ChemicalPerturbation_IGF1-MCF7_UP	172	122	0.71	2.4e-45	1.3e-42
LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-pelitinib-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-pelitinib-0.37_DN	73	57	0.78	1.5e-25	4.0e-24
LINCSCMAP_ChemicalPerturbation_LJP005_SKBR3_24H-CGP-60474-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP005_SKBR3_24H-CGP-60474-3.33_DN	114	61	0.54	1.4e-14	7.5e-14
LINCSCMAP_ChemicalPerturbation_CPC009_A549_24H-BRD-A85712510-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC009_A549_24H-BRD-A85712510-10.0_DN	115	52	0.45	3.5e-9	8.6e-9

Showing 1 to 5 of 5 entries (filtered from 33,147 total entries) Previous 1 Next

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report SCC: Module(s) regulated by SOX2?

Community-AMARETTO Community 1



Enrichments of Functional Categories in Community

Show 5 entries Search: SOX2

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value
TARGETS	Polycomb protein EED [GeneID=8726] in human embryonic stem cells.				
BENPORATH NANOG TARGETS	Set 'Nanog targets': genes upregulated and identified by ChIP on chip as Nanog [GeneID=79923] transcription factor targets in human embryonic stem cells.	988	987	1.0	0.0
BENPORATH SOX2 TARGETS	Set 'Sox2 targets': genes upregulated and identified by ChIP on chip as SOX2 [GeneID=6657] transcription factor targets in human embryonic stem cells.	734	734	1.0	0.0
BENPORATH SUZ12 TARGETS	Set 'Suz12 targets': genes identified by ChIP on chip as targets of the Polycomb protein SUZ12 [GeneID=23512] in human embryonic stem cells.	1038	1038	1.0	0.0
MEISSNER BRAIN HCP WITH H3K4ME3 AND H3K27ME3	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation at K4 (H3K4me2) and trimethylation at K27 (H3K27me3) in brain.	1069	857	0.80	0.0

Showing 1 to 5 of 115 entries (filtered from 4,907 total entries) Previous 1 2 3 4 5 …

Enrichments of Driver Perturbations in Community

CSV Excel PDF Print Column visibility Show 10 entries Search:

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
SOX2						
ChEA_SOX2_21211035_ChIP-	ChEA_SOX2_21211035_ChIP-	3420	1775	0.52	0.0	0.0
ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	ChEA_SOX2_20726797_ChIP-Seq_SW620_Human	2564	1343	0.52	3.0e-292	5.3e-291
Consensus_SOX2_CHEA	Consensus_SOX2_CHEA	775	556	0.72	5.8e-208	3.0e-206
ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	ChEA_SOX2_16153702_ChIP-ChIP_HESCs_Human	1278	755	0.59	4.7e-202	3.6e-201
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-	3319	1408	0.42	2.5e-190	1.5e-189
ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human	ChEA_SOX2_19829295_ChIP-Seq_ESCs_Human					
ChEA_SOX2_18692474_ChIP-	ChEA_SOX2_18692474_ChIP-					
ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse	ChEA_SOX2_27498859_Chip-Seq_STOMACH_Mouse					
ChEA_SOX2_18555785_Chip-	ChEA_SOX2_18555785_Chip-					
ChEA_SOX2_21211035_ChIP-Seq_LN229_Human	ChEA_SOX2_21211035_ChIP-Seq_LN229_Human					

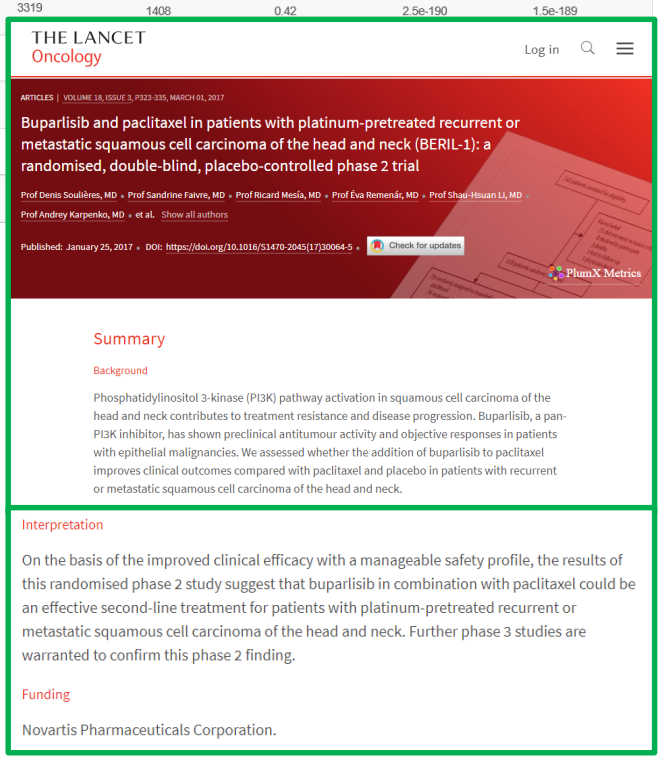
Showing 1 to 10 of 14 entries (filtered from 9,045 total entries)

Enrichments of Chemical Perturbations in Community

CSV Excel PDF Print Column visibility Show 10 entries Search: SOX2

Gene Set Name	Gene Set Description	# Genes in Gene Set	# Genes in Overlap	% Genes in overlap	P-value	FDR Q-value
LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-buparlisib-10_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-buparlisib-10_DN	147	111	0.76	7.6e-46	4.8e-43
LINCSCMAP_ChemicalPerturbation_IGF1-MCF7_UP	LINCSCMAP_ChemicalPerturbation_IGF1-MCF7_UP	172	122	0.71	2.4e-45	1.3e-42
LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-pelitinib-0.37_DN	LINCSCMAP_ChemicalPerturbation_LJP005_A375_24H-pelitinib-0.37_DN	73	57	0.78	1.5e-25	4.0e-24
LINCSCMAP_ChemicalPerturbation_LJP005_SKBR3_24H-CGP-60474-3.33_DN	LINCSCMAP_ChemicalPerturbation_LJP005_SKBR3_24H-CGP-60474-3.33_DN	114	61	0.54	1.4e-14	7.5e-14
LINCSCMAP_ChemicalPerturbation_CPC009_A549_24H-BRD-A85712510-10.0_DN	LINCSCMAP_ChemicalPerturbation_CPC009_A549_24H-BRD-A85712510-10.0_DN	115	52	0.45	3.5e-9	8.6e-9

Showing 1 to 5 of 5 entries (filtered from 33,147 total entries) Previous 1 Next


THE LANCET Oncology Log in

ARTICLES | VOLUME 18, ISSUE 3, P323-335, MARCH 01, 2017

Buparlisib and paclitaxel in patients with platinum-pretreated recurrent or metastatic squamous cell carcinoma of the head and neck (BERIL-1): a randomised, double-blind, placebo-controlled phase 2 trial

Prof Denis Soulières, MD • Prof Sandrine Faivre, MD • Prof Ricard Mesía, MD • Prof Éva Remenár, MD • Prof Shau-Hsuan Li, MD • Prof Andrey Karpenko, MD • et al. Show all authors

Published: January 25, 2017 • DOI: https://doi.org/10.1016/S1470-2045(17)30064-5 • Check for updates

PlumX Metrics

Summary

Background

Phosphatidylinositol 3-kinase (PI3K) pathway activation in squamous cell carcinoma of the head and neck contributes to treatment resistance and disease progression. Buparlisib, a pan-PI3K inhibitor, has shown preclinical antitumour activity and objective responses in patients with epithelial malignancies. We assessed whether the addition of buparlisib to paclitaxel improves clinical outcomes compared with paclitaxel and placebo in patients with recurrent or metastatic squamous cell carcinoma of the head and neck.

Interpretation

On the basis of the improved clinical efficacy with a manageable safety profile, the results of this randomised phase 2 study suggest that buparlisib in combination with paclitaxel could be an effective second-line treatment for patients with platinum-pretreated recurrent or metastatic squamous cell carcinoma of the head and neck. Further phase 3 studies are warranted to confirm this phase 2 finding.

Funding

Novartis Pharmaceuticals Corporation.

Data Set	Module	Gene	Gene Type
TCGA_BLCA	Module 8	SOX2	Target
TCGA_CESC	Module 14	SOX2	Target
TCGA_CESC	Module 14	SOX2	Driver
TCGA_ESCA	Module 83	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Target
TCGA_HNSC	Module 51	SOX2	Driver
TCGA_HNSC	Module 106	SOX2	Driver
TCGA_LUSC	Module 18	SOX2	Target
TCGA_LUSC	Module 18	SOX2	Driver
TCGA_LUSC	Module 37	SOX2	Driver
TCGA_LUSC	Module 41	SOX2	Driver
TCGA_LUSC	Module 61	SOX2	Driver
TCGA_LUSC	Module 94	SOX2	Driver

http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/communities/Community_1.html

Community-AMARETTO report pan-SCC - AMARETTO report LUSC

Summary of SOX2-regulated LUSC Module 37:

Drivers: SOX2, TP63, PIK3CA

SOX2 CNV amplification, associated with induced SOX2 expression

SOX2, TP63, PIK3CA are activators of their targets

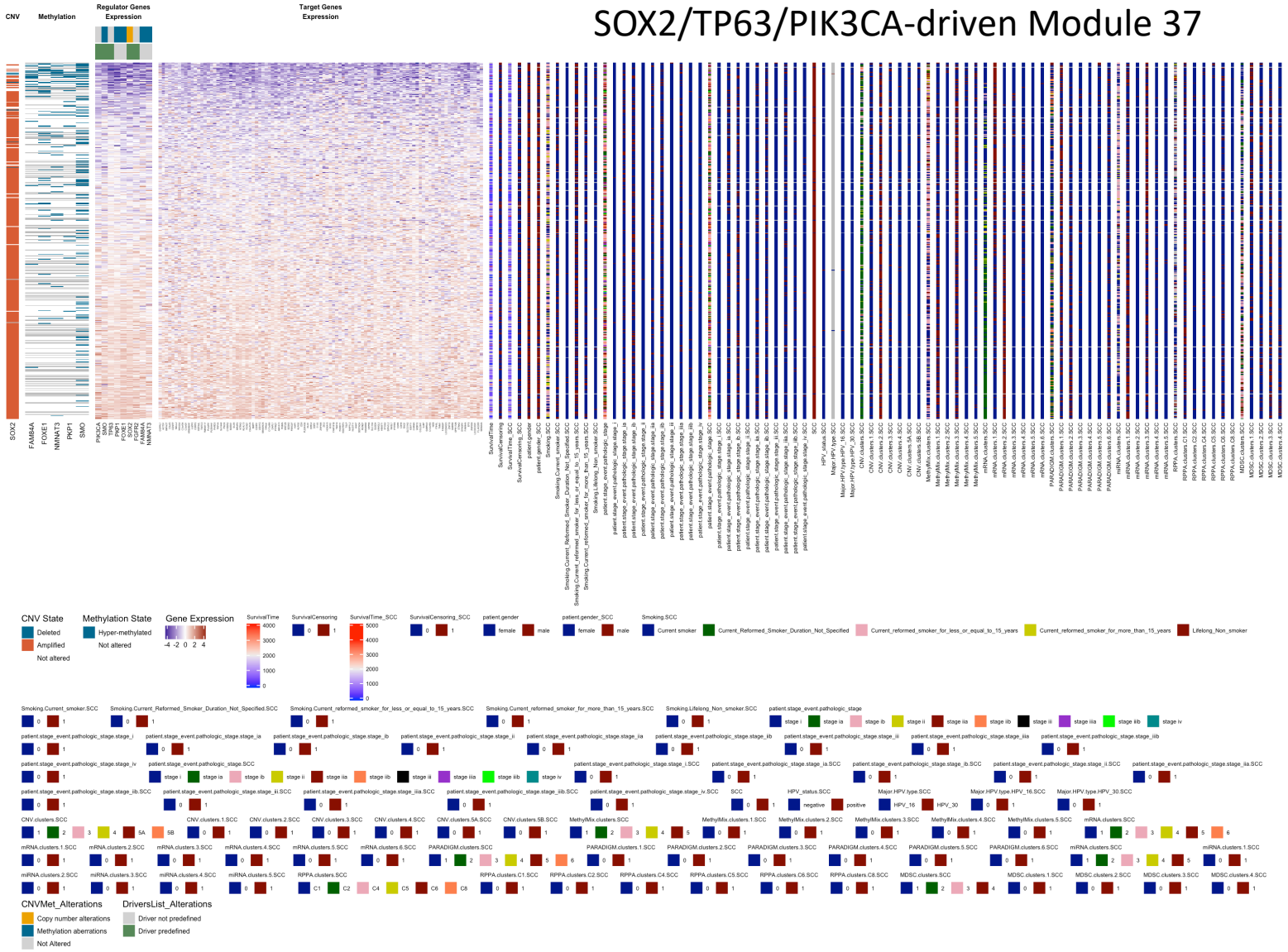
Associated with survival (lower expression, poorer survival) and TCGA multi-omics clusters (CNV)

Enriched for PI3K pathway, stemness and squamous-specific gene signatures

Drivers validated:

- SOX2 and TP63: ENCODE and ChEA ChIP-Seq, bound to its target genes
- SOX2 and PIK3CA: LINCS/CMAP genetic perturbations, modulating drivers modulates its target genes

SOX2/TP63/PIK3CA-driven Module 37



http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_LUSC/AMARETTOhtmls/modules/module37.html

Perturbation-AMARETTO report SCC/LUSC

Perturbation-AMARETTO v2: driver validation & discovery using genetic perturbations from LINCS/CMAP

Case Study 3: squamous cell carcinoma across 5 cancer sites

Driver discovery across 5 data sets

CSV | Excel | PDF | Print | Column visibility | Show 500 entries | Search:

PerturbationID	Cell_Line	GeneSymbol	EntrezID	PerturbationType	Type	DataSetFrequency	BLCA	CESC	ESCA	HNSC	LUSC
All	,	["SC		All		All	All			All	Module 37 :
OEC001_A375_48H:CCSBBROAD304_01579:-666	A375	SOX2	6657	trt_oe	landmark	5	Module 8 : T_CD (w = 0) , escore-pval-padj-zscore	Module 14 : A_D (w = 0.0737) , escore-pval-padj-zscore	Module 83 : T_CD (w = 0) , escore-pval-padj-zscore	Module 51 : A_D (w = 0.1625) , escore-pval-padj-zscore; Module 76 : R_D (w = -0.0149) , escore-pval-padj-zscore; Module 106 : A_D (w = 0.0034) , escore-pval-padj-zscore	Module 18 : A_D (w = 0.2285) , escore-pval-padj-zscore; Module 37 : A_D (w = 0.0975) , escore-pval-padj-zscore; Module 41 : A_D (w = 0.0128) , escore-pval-padj-zscore; Module 61 : A_D (w = 0.0357) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.0547) , escore-zscore; Module 94 : A_D (w = 0.0282) , escore-pval-padj-zscore
OEB002_HT29_96H:BRDN0000401187:-666	HT29	SOX2	6657	trt_oe	landmark	5	Module 8 : T_CD (w = 0) , escore-pval-padj-zscore	Module 14 : A_D (w = 0.0737) , escore-pval-padj-zscore	Module 83 : T_CD (w = 0) , escore-pval-padj-zscore	Module 51 : A_D (w = 0.1625) , escore-pval-padj-zscore; Module 76 : R_D (w = -0.0149) , zscore; Module 106 : A_D (w = 0.0034) , escore-pval-padj-zscore	Module 18 : A_D (w = 0.2285) , escore-pval-padj-zscore; Module 37 : A_D (w = 0.0975) , escore-pval-padj-zscore; Module 41 : A_D (w = 0.0128) , escore-pval-padj-zscore; Module 61 : A_D (w = 0.0357) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.0547) , pval-padj-zscore; Module 94 : A_D (w = 0.0282) , escore-pval-padj-zscore
OEB002_MCF7_96H:BRDN0000401187:-666	MCF7	SOX2	6657	trt_oe	landmark	5	Module 8 : T_CD (w = 0) , escore-pval-padj-zscore	Module 14 : A_D (w = 0.0737) , escore-pval-padj-zscore	Module 83 : T_CD (w = 0) , escore-pval-padj-zscore	Module 51 : A_D (w = 0.1625) , escore-pval-padj-zscore; Module 76 : R_D (w = -0.0149) , escore-zscore; Module 106 : A_D (w = 0.0034) , escore-pval-padj-zscore	Module 18 : A_D (w = 0.2285) , escore-pval-padj-zscore; Module 37 : A_D (w = 0.0975) , escore-pval-padj-zscore; Module 41 : A_D (w = 0.0128) , escore-pval-padj-zscore; Module 61 : A_D (w = 0.0357) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.0547) , escore-zscore; Module 94 : A_D (w = 0.0282) , escore-pval-padj-zscore
OEB002_PC3_96H:BRDN0000401187:-666	PC3	SOX2	6657	trt_oe	landmark	5	Module 8 : T_CD (w = 0) , escore-pval-padj-zscore	Module 14 : A_D (w = 0.0737) , escore-pval-padj-zscore	Module 83 : T_CD (w = 0) , escore-pval-padj-zscore	Module 51 : A_D (w = 0.1625) , escore-pval-padj-zscore; Module 76 : R_D (w = -0.0149) , escore-padj-zscore; Module 106 : A_D (w = 0.0034) , escore-pval-padj-zscore	Module 18 : A_D (w = 0.2285) , escore-pval-padj-zscore; Module 37 : A_D (w = 0.0975) , escore-pval-padj-zscore; Module 41 : A_D (w = 0.0128) , escore-pval-padj-zscore; Module 61 : A_D (w = 0.0357) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.0547) , escore-pval-padj-zscore; Module 94 : A_D (w = 0.0282) , escore-pval-padj-zscore
CGS001_A375_96H:PIK3CA:1	A375	PIK3CA	5290	trt_sh.cgs	landmark	4	Not_in_AMARETTO	Module 94 : A_D (w = 0.1491) , escore-pval-padj-zscore	Module 19 : A_D (w = 0.0365) , escore-pval-padj-zscore	Module 16 : A_D (w = 0.0398) , escore-zscore; Module 98 : A_D (w = 0.0123) , escore-pval-padj-zscore	Module 37 : A_D (w = 0.0034) , escore-pval-padj-zscore; Module 51 : A_D (w = 0.002) , escore-zscore; Module 73 : A_D (w = 0.1852) , escore-pval-padj-zscore
CGS001_HA1E_96H:PIK3CA:1.5	HA1E	PIK3CA	5290	trt_sh.cgs	landmark	4	Not_in_AMARETTO	Module 94 : A_D (w = 0.1491) , escore-pval-padj-zscore	Module 19 : A_D (w = 0.0365) , escore-pval-padj-zscore	Module 16 : A_D (w = 0.0398) , escore-zscore; Module 98 : A_D (w = 0.0123) , escore-pval-padj-zscore	Module 37 : A_D (w = 0.0034) , escore-pval-padj-zscore; Module 51 : A_D (w = 0.002) , escore-pval-padj-zscore; Module 73 : A_D (w = 0.1852) , escore-zscore
CGS001_HT29_96H:PIK3CA:1	HT29	PIK3CA	5290	trt_sh.cgs	landmark	4	Not_in_AMARETTO	Module 94 : A_D (w = 0.1491) , escore-pval-padj-zscore	Module 19 : A_D (w = 0.0365) , escore-pval-padj-zscore	Module 16 : A_D (w = 0.0398) , zscore; Module 98 : A_D (w = 0.0123) , escore-pval-padj-zscore	Module 37 : A_D (w = 0.0034) , escore-zscore; Module 51 : A_D (w = 0.002) , escore-zscore; Module 73 : A_D (w = 0.1852) , escore-pval-padj-zscore
CGS001_MCF7_96H:PIK3CA:2	MCF7	PIK3CA	5290	trt_sh.cgs	landmark	4	Not_in_AMARETTO	Module 94 : A_D (w = 0.1491) , escore-pval-padj-zscore	Module 19 : A_D (w = 0.0365) , escore-pval-padj-zscore	Module 16 : A_D (w = 0.0398) , zscore; Module 98 : A_D (w = 0.0123) , escore-pval-padj-zscore	Module 37 : A_D (w = 0.0034) , pval-padj-zscore; Module 51 : A_D (w = 0.002) , zscore; Module 73 : A_D (w = 0.1852) , escore-pval-padj-zscore

https://pochetlab.shinyapps.io/pAMARETTO_AMARETTO_PanCancer_5DS_Drivers/ (under development)

Community-AMARETTO report pan-SCC - AMARETTO report HNSC

Summary of SOX2/GPX2-regulated HNSC Module 51:

SOX2 CNV amplification, associated with induced SOX2 expression

GPX2 hypo/hyper-methylation, associated with induced/repressed GPX2 expression

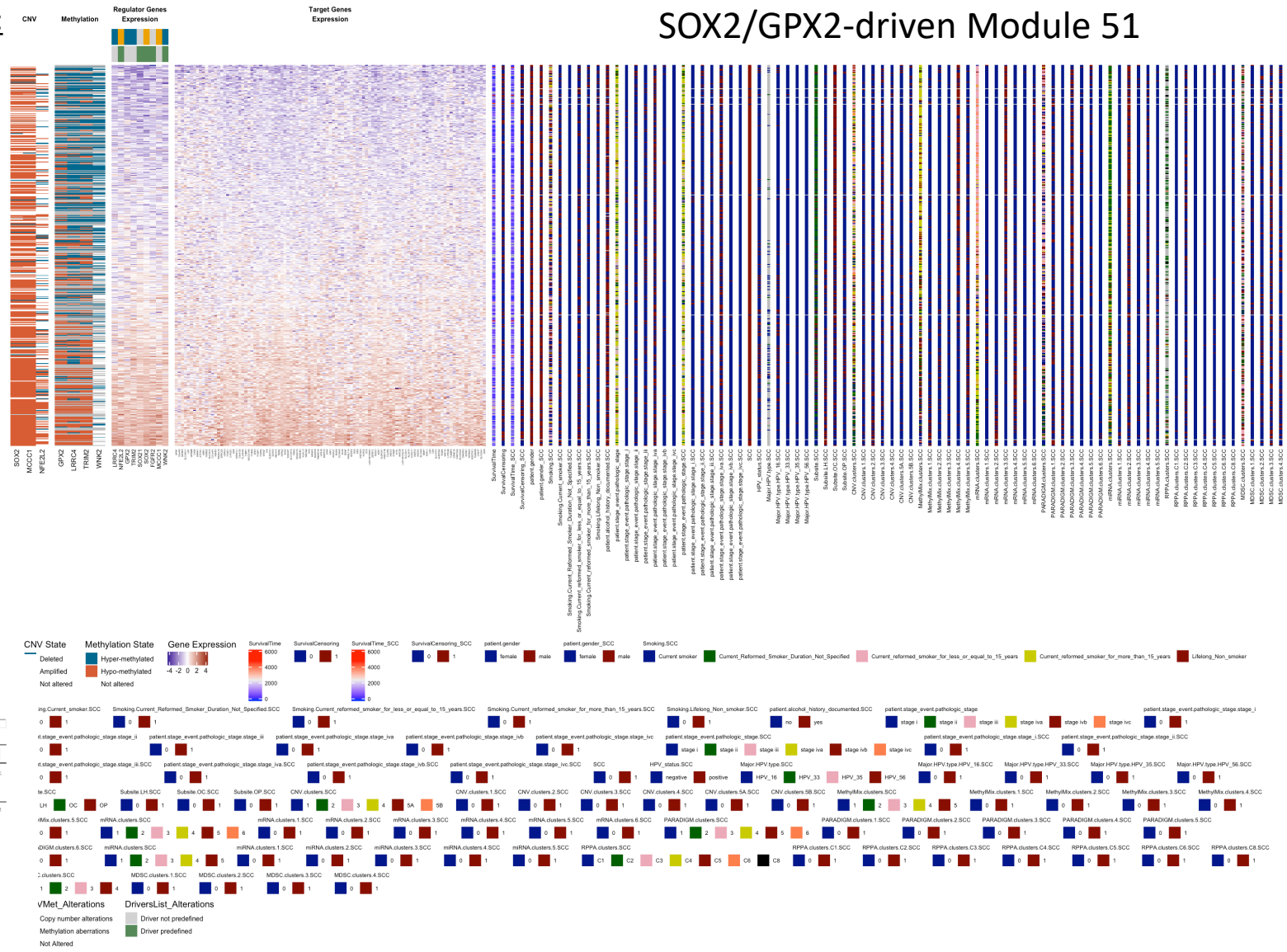
SOX2 and GPX2 are activators of their targets

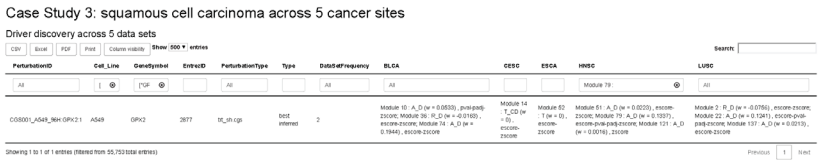
Associated with smoking and HPV, and TCGA multi-omics clusters (CNV and methylation)

Enriched for HNSC and SCC-specific gene signatures

SOX2 and GPX2 validated for HNSC/LUSC modules

SOX2/GPX2-driven Module 51





http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_HNSC/AMARETTOhtmls/modules/module51.html

Community-AMARETTO report pan-SCC - AMARETTO report HNSC

Summary of GPX2-regulated HNSC Module 79:

GPX2 hypo/hyper-methylation, associated with induced/repressed GPX2 expression

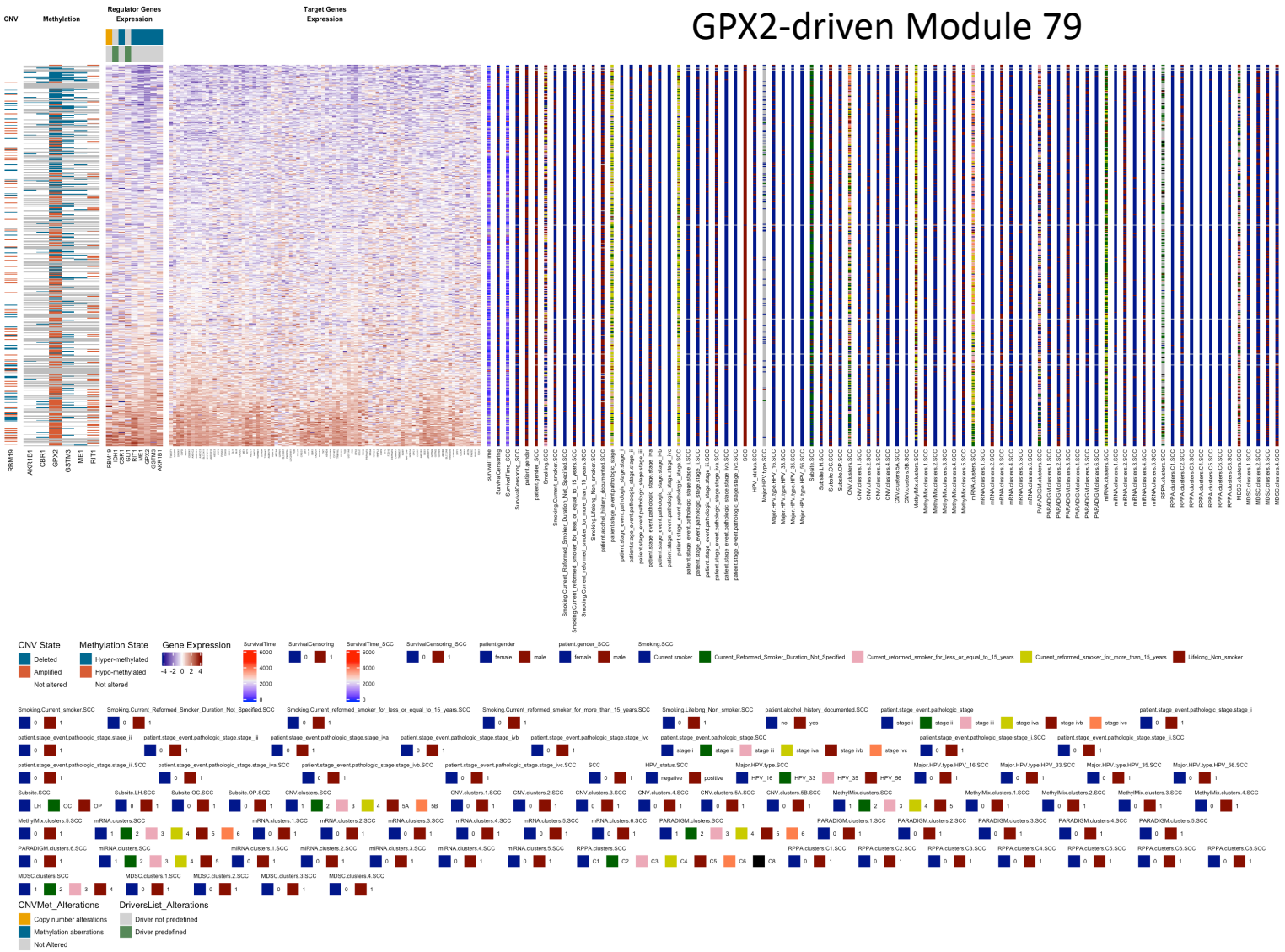
GPX2 is an activator of its target genes

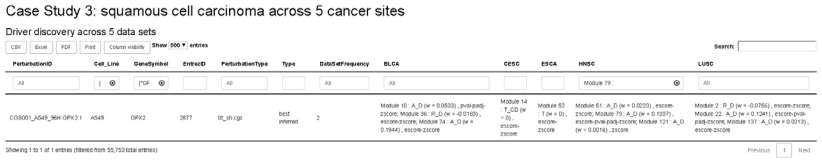
Associated with smoking (lower expression ~non-smoking) and HPV (lower expression ~HPV), and TCGA multi-omics clusters (methylation)

Enriched for HNSC and SCC-specific gene signatures

SOX2 and GPX2 validated using KD experiments in A549 cell line for HNSC/LUSC modules

GPX2-driven Module 79





http://portals.broadinstitute.org/pochetlab/demo/cAMARETTO_PanCancer_5DS/TCGA_HNSC/AMARETTOhtmls/modules/module79.html

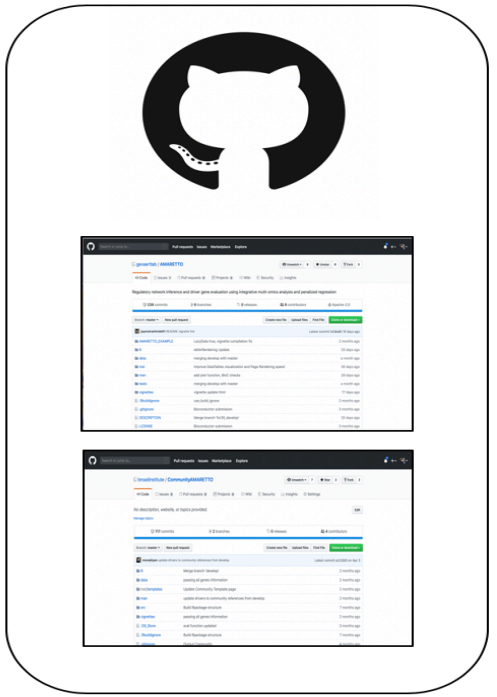
*AMARETTO source code, tools & notebooks

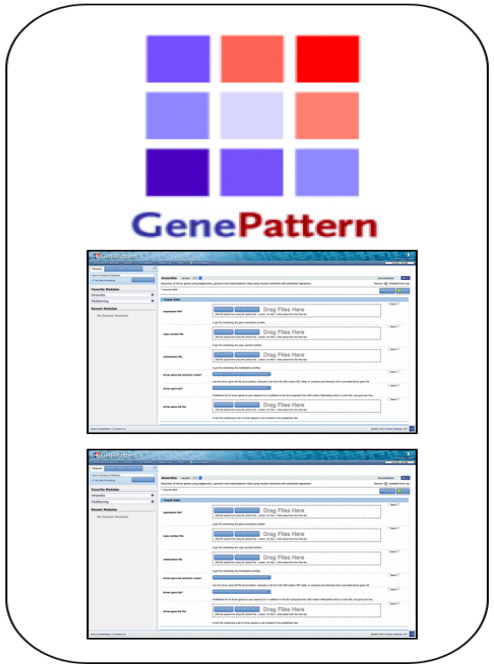
*AMARETTO is available via:

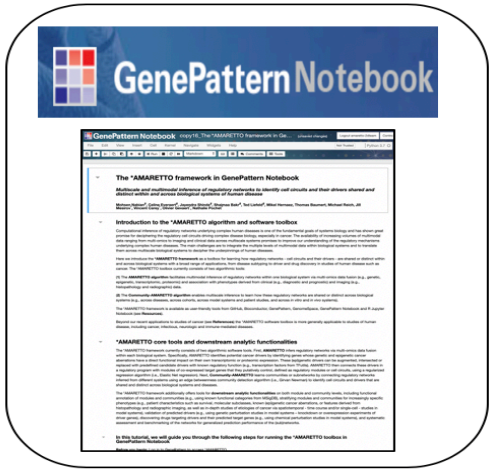
- GitHub
- Bioconductor
- Jupyter Notebook
- GenePattern
- GenomeSpace
- GenePattern Notebook

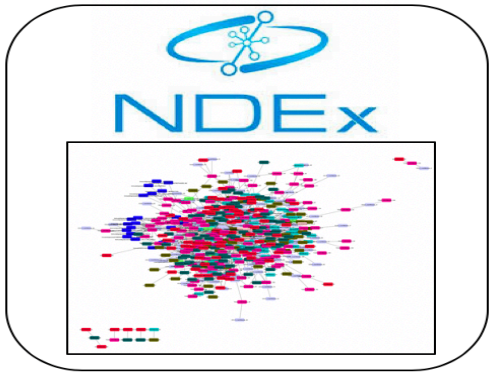
Tools and resources:

http://portals.broadinstitute.org/pochetlab/amaretto.html

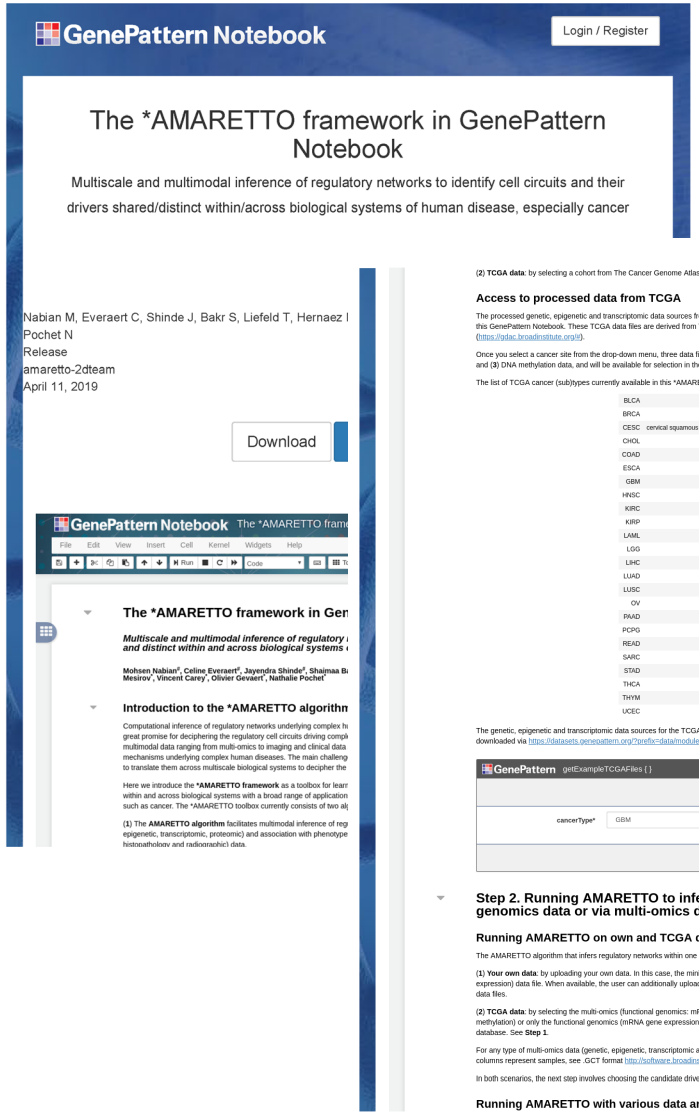


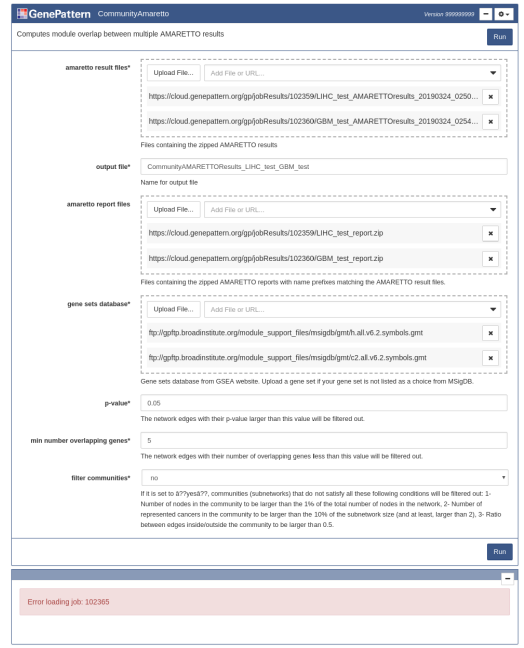


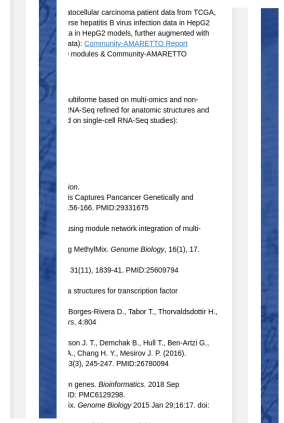


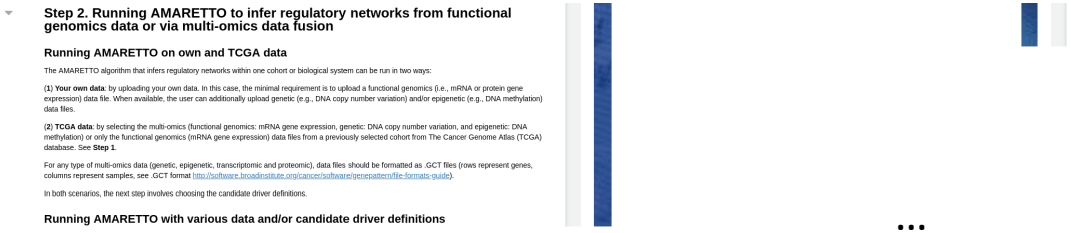


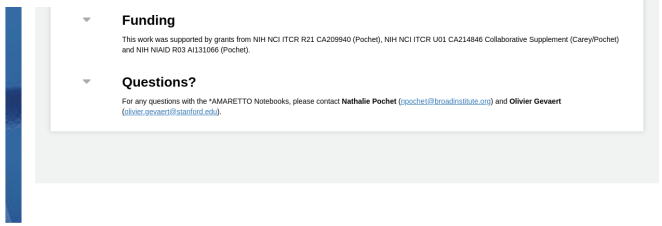
*AMARETTO GenePattern Notebook











https://notebook.genepattern.org/services/sharing/notebooks/334/preview/

*AMARETTO R Jupyter Notebook

The *AMARETTO framework in R via GitHub and Biocond

Multiscale and multimodal inference of regulatory networks to identify cell circuits and their drivers biological systems of human disease

Mohsen Nabian[#], Jayendra Shinde[#], Celine Everaert[#], Shaimaa Bakr[#], Ted Liefeld, Thorin Tabor, Charles Blatti, Th Mikel Hernaez[*], Vincent Carey[*], Olivier Gevaert[*], Nathalie Pochet[*]

Introduction to the *AMARETTO algorithm and software toolbox

Computational inference of r
promise for deciphering the
ranging from multi-omics to
complex human diseases. Tl
multiscale biological system

Here we introduce the ***AMA**
across biological systems w
*AMARETTO toolbox current

(1) The **AMARETTO algorith**
epigenetic, transcriptomic, p
radiographic) data.

(2) The **Community-AMARE**
(e.g., across diseases, acros

The *AMARETTO framework
Notebook (see **Resources**).

Beyond our recent applicatic
including cancer, infectious,

*AMARETTO core

The *AMARETTO framework
each biological system. Spe
direct functional impact on t
candidate drivers with know
modules of co-expressed tar
Elastic Net regression). Next
an edge betweenness comm
systems and diseases.

The *AMARETTO framework
of modules and communitie
(e.g., patient characteristics

For running the Notebook on **preloaded TCGA data** please continue in this **Step 2**.
https://datasets.genepattern.org/?prefix=data/module_support_files/Amaretto/ in
automatically read from this link and how they are converted to data matrices in R

For running the Notebook on **example data** you can immediately proceed to **Step 3**
and also available for download from https://www.broadinstitute.org/~npochet/No
example data are available here: (1) for 150 modules and 75% variation filtering (se
https://www.broadinstitute.org/~npochet/NotebookExample/ExampleResults/resu
https://www.broadinstitute.org/~npochet/NotebookExample/ExampleResults/resu
https://www.broadinstitute.org/~npochet/NotebookExample/ExampleResults/resu
https://www.broadinstitute.org/~npochet/NotebookExample/ExampleResults/resu

Access to processed data from TCGA

The processed genetic, epigenetic and transcriptomic data sources from TCGA are
Notebook. These TCGA data files are derived from The Cancer Genome Atlas (TCG
(https://gdac.broadinstitute.org/).

From https://datasets.genepattern.org/?prefix=data/module_support_files/Amaret
downloaded: (1) mRNA gene expression data (MA), (2) DNA copy number variatior

The list of TCGA cancer (sub)types currently available in this *AMARETTO in R Not

BLCA	bladder urothelial carcinoma
BRCA	breast invasive carcinoma
CESC	cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	cholangiocarcinoma
COAD	colon adenocarcinoma
ESCA	esophageal carcinoma
GBM	glioblastoma multiforme
HNSC	head and neck squamous cell carcinoma
KIRC	kidney renal clear cell carcinoma
KIRP	kidney renal papillary cell carcinoma
LAML	acute myeloid leukemia
LGG	brain lower grade glioma
LIHC	liver hepatocellular carcinoma
LUAD	lung adenocarcinoma
LUSC	lung squamous cell carcinoma
OV	ovarian serous cystadenocarcinoma
PAAD	pancreatic adenocarcinoma
PCPG	pheochromocytoma and paraganglioma
READ	rectum adenocarcinoma
SARC	sarcoma
STAD	stomach adenocarcinoma
THCA	thyroid carcinoma

...

Step 3. Running AMARETTO on first example study: infer networks via multi-omics data fusion for TCGA LIHC pati

The AMARETTO algorithm that infers regulatory networks within one cohort or biological system can be run in multi genetic, epigenetic and functional genomics data are available (see example in this **Step 3** for multi-omics data fron are available (see example in next **Step 4** for transcriptomic data from CCLE).

When **either genetic** (e.g., DNA copy number variation) **or epigenetic** (e.g., DNA methylation) data **or both** are availal transcriptomic or proteomic) data, there are various options for defining candidate drivers for analysis by the AMAR

In case **only functional genomics** (i.e., mRNA or protein gene expression) data are available, a predefined list of can AMARETTO algorithm.

The AMARETTO algorithm can take vario

(1) Select **computed lists of candidate dr** data files are uploaded);

(2) Select or upload **predefined lists of ca**
https://bioconductor.org/packages/relea
http://software.broadinstitute.org/gsea/r
http://software.broadinstitute.org/gsea/r
data("Driver_Genes"));

(3) Take the **union or intersection** betwee

For computed lists of candidate drivers f for TCGA data, however, for processing o recurrent DNA copy number aberrations (recurrent DNA methylation aberrations (h association for DNA copy number aberra

Step 3.a. Preparing data ar

Loading multi-omics data from TC

Loading Gene Expression (MA) data fron

Step 4. Running AMARETTO on second examp networks from RNA-Seq data from CCLE liver

The AMARETTO algorithm that infers regulatory networks within one cohort or biological syste genetic, epigenetic and functional genomics data are available (see example in previous **Step** data are available (see example in this **Step 4** for transcriptomic data from CCLE). See **Step 3**

Step 4.a. Preparing data and parameter settings for running A

Loading RNA-Seq data from CCLE liver cell lines

Loading Gene Expression (MA) data from CCLE liver cell lines (Required)

```
MA_matrix_CCLE <- readRDS(url("https://www.broadinstitute.org/~npochet/NotebookEx
ProcessedData_CCLE = list(MA_matrix=MA_matrix_CCLE, CNV_matrix=NULL, MET_matrix=NU
```

Defining List(s) of Candidate Driver Genes (Required)

In this example, we precompiled a list of candidate driver genes that takes the union of TCGA list of candidate drivers as in **Step 3**)

```
candidate_drivers_CCLE <- readRDS(url("https://www.broadinstitute.org/~npochet/No
```

Setting parameters for running AMARETTO (Required)

Core parameters that can be set by the user for running AMARETTO. See **Step 3** for more deta

```
NrModules = 150
VarPercentage = 75
```

Setting parameters for generating HTML results reports (Optional)

Additional parameters that can be set by the user for running AMARETTO. See **Step 3** for more detailed information.

```
genesets_database_reference <- "H_C2_genesets.gmt"
download.file(url="https://www.broadinstitute.org/~npochet/NotebookExample/ExampleData/H_C2_genesets

output_directory_CCLE = "./AMARETTO_report_CCLE/"
dir.create(output_directory_CCLE)
```

Step 5. Running Community-AMARETTO to combine botl identifying regulatory subnetworks or communities share TCGA and CCLE cohorts

The Community-AMARETTO algorithm takes as input results from two or more previous AMARETTO analyses to ide (i.e., cell circuits and their drivers) that are shared and distinct across multiple datasets, cohorts, biological systems

Step 5.a. Preparing data and parameter settings for running Community-AM

Loading two or more results from AMARETTO, in this example the previous TCGA and CCLE resul

Selecting AMARETTO analyses for Community-AMARETTO analysis. The user can submit the .rds files that represe more previous AMARETTO analyses (see above, run in **Steps 3 and 4**).

```
AMARETTOresults_TCGA <- readRDS(file="TCGA_AMARETTOresults.rds")

AMARETTOresults_CCLE <- readRDS(file="CCLE_AMARETTOresults.rds")

HTMLsAMARETTOlist <- c("TCGA"=output_directory_TCGA,"CCLE"=output_directory_CCLE)
```

Loading additional networks as a set of signatures in .GMT format (Optional)

One or more additional networks can be submitted as signatures files in .GMT format and combined by running the Community-AMARETTO as separate networks. In this example, we submit previously published signatures and/or n CiberSort, stemness signatures from *Ben-Porath et al.*, and diagnostic and prognistic liver cancer signatures from H be analyzed together with the liver cancer networks derived from TCGA in **Step 3** and CCLE in **Step 4**.

If additional networks are submitted, please run following cell code to include them in the analysis.

```
ImmuneSignatures <- "ImmuneSignatures.gmt"
download.file(url="https://www.broadinstitute.org/~npochet/NotebookExample/ExampleData/ImmuneSignatu

StemSignatures <- "StemSignatures.gmt"
download.file(url="https://www.broadinstitute.org/~npochet/NotebookExample/ExampleData/StemSignature

LiverSignatures <- "LiverSignatures.gmt"
download.file(url="https://www.broadinstitute.org/~npochet/NotebookExample/ExampleData/LiverSignatur

list_additional_networks = list(ImmuneSignatures = "ImmuneSignatures.gmt", StemSignatures = "StemSig
```

Otherwise set to NULL.

```
list_additional_networks = NULL
```

Setting parameters for generating HTML results reports (Optional)

11. Reich M., Liefeld T., Ocana M., Jang D., Bistline J., Robinson J., Carr P., Hill B., McLaughlin J., Pochet N., Borge Mesirov J.P. (2013). GenomeSpace: an environment for frictionless bioinformatics. *F1000Posters*, 4:804. (htt

Funding

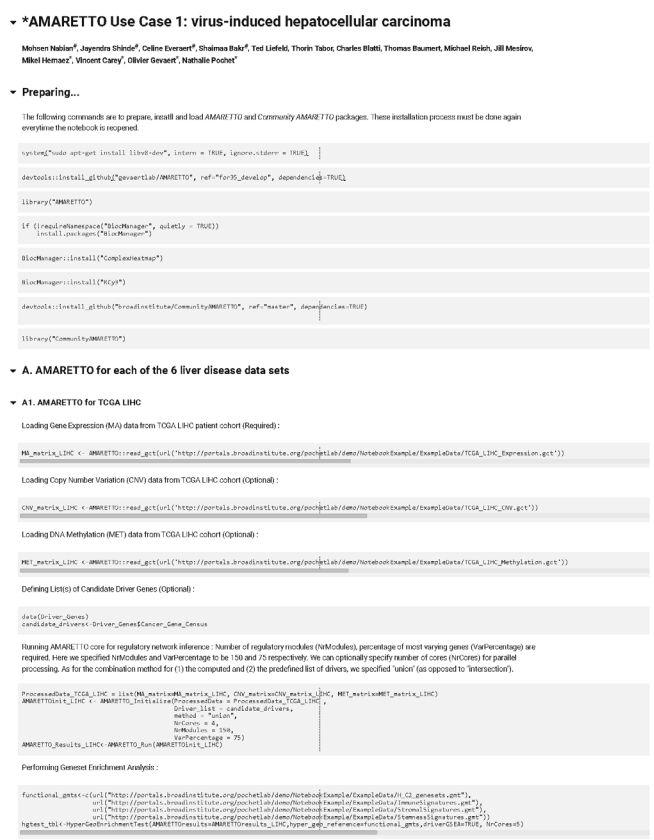
This work was supported by grants from NIH NCI ITCR R21 CA209940 (Pochet), NIH NCI ITCR U01 CA214846 Collal NIAID R03 AI131066 (Pochet).

Questions?

For any questions with the *AMARETTO Notebooks, please contact **Nathalie Pochet** (npochet@broadinstitute.org) a

https://colab.research.google.com/drive/1JfnRoNgTVX_7VEGAAmjGjwP_yX2tdDxs

*AMARETTO R Jupyter Notebook Use Case 1



*AMARETTO Use Case 1: virus-induced hepatocellular carcinoma

Mohsen Nabian#, Jayendra Shinde#, Celine Everaert#, Shaimaa Bakr#, Ted Liefeld, Thorin Tabor, Charles Blatti, Thomas Baumert, Michael Reich, Jill Mesirov, Mikel Hernaez*, Vincent Carey*, Olivier Gevaert*, Nathalie Pochet*

Preparing...

The following commands are to prepare, install and load *AMARETTO* and *Community AMARETTO* packages. These installation process must be done again everytime the notebook is reopened.

```
system("sudo apt-get install libv8-dev", intern = TRUE, ignore.stderr = TRUE)

devtools::install_github("gevaertlab/AMARETTO", ref="for35_develop", dependencies=TRUE)

library("AMARETTO")

if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

BiocManager::install("ComplexHeatmap")

BiocManager::install("RCy3")

devtools::install_github("broadinstitute/CommunityAMARETTO", ref="master", dependencies=TRUE)

library("CommunityAMARETTO")
```

A. AMARETTO for each of the 6 liver disease data sets

A1. AMARETTO for TCGA LIHC

Loading Gene Expression (MA) data from TCGA LIHC patient cohort (Required) :

```
MA_matrix_LIHC <- AMARETTO::read_gct(url('http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LIHC_Expression.gct'))
```

Loading Copy Number Variation (CNV) data from TCGA LIHC cohort (Optional) :

```
CNV_matrix_LIHC <-AMARETTO::read_gct(url('http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LIHC_CNV.gct'))
```

Loading DNA Methylation (MET) data from TCGA LIHC cohort (Optional) :

```
MET_matrix_LIHC <-AMARETTO::read_gct(url('http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LIHC_Methylation.gct'))
```

Defining List(s) of Candidate Driver Genes (Optional) :

```
data(Driver_Genes)
candidate_drivers<-Driver_Genes$Cancer_Gene_Census
```

Running AMARETTO core for regulatory network inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are required. Here we specified NrModules and VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel processing. As for the combination method for (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
ProcessedData_TCGA_LIHC = list(MA_matrix=MA_matrix_LIHC, CNV_matrix=CNV_matrix_LIHC, MET_matrix=MET_matrix_LIHC)
AMARETTOinit_LIHC <- AMARETTO_Initialize(ProcessedData = ProcessedData_TCGA_LIHC,
                                         Driver_list = candidate_drivers,
                                         method = "union",
                                         NrCores = 4,
                                         NrModules = 150,
                                         VarPercentage = 75)
AMARETTO_Results_LIHC<-AMARETTO_Run(AMARETTOinit_LIHC)
```

Performing Geneset Enrichment Analysis :

```
functional_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/H_C2_genesets.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StromalSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StemnessSignatures.gmt"))
hgtest_tbl<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_LIHC,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

urbations :

```
te.org/pochetlab/demo/NotebookExample/ExampleData/GeneticPerturbationSignatures.gmt"),
te.org/pochetlab/demo/NotebookExample/ExampleData/ChIPABindingSignatures.gmt"),
te.org/pochetlab/demo/NotebookExample/ExampleData/ConsensusBindingSignatures.gmt"),
te.org/pochetlab/demo/NotebookExample/ExampleData/EncodeBindingSignatures.gmt"))
sults=AMARETTOresults_LIHC,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

rbations :

```
bationSignatures.gmt"))
TOresults=AMARETTOresults_LIHC,hyper_geo_reference=chemical_gmts,driverGSEA=TRUE, NrCores=5)
```

orming performing the phenotype association tests :

```
institute.org/pochetlab/demo/NotebookExample/ExampleData/LIHC_All_Phenotypes_data.csv"))
lm.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/LIHC_All_Statistics_data.csv"))
(AMARETTO_statistical_test(AMARETTOinit_LIHC,AMARETTO_Results_LIHC,samples_LIHC,phenotype_tests_LIHC))
```

```
_hgtest_tbl,
TO_results/LIHC_Report',
e_LIHC,

ble = LIHC_phenotype_tests_all)
```

atient cohort (Required) :

```
dinstitute.org/pochetlab/demo/NotebookExample/ExampleData/CCLE_data_mat.rds"))
```

```
us
```

ice : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are
ge to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel
computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
LE, CNV_matrix=NULL, MET_matrix=NULL)
edData=ProcessedData_CCLE,
list = candidate_drivers,
= "union",
= 4,
es = 150,
entage = 75,
seeds = c(42,42))
CCLE)
```

```
itute.org/pochetlab/demo/NotebookExample/ExampleData/H_C2_genesets.gmt"),
itute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt"),
itute.org/pochetlab/demo/NotebookExample/ExampleData/StromalSignatures.gmt"),
itute.org/pochetlab/demo/NotebookExample/ExampleData/StemnessSignatures.gmt"))
ts=AMARETTOresults_CCLE,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

Performing Geneset Enrichment Analysis for driver perturbations :

```
genetic_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/GeneticPerturbationSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ChIPABindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ConsensusBindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/EncodeBindingSignatures.gmt"))
hgtest_tbl_gpc-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_CCLE,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

Performing Geneset Enrichment Analysis for drug perturbations :

```
chemical_gmts<-c(url("./signature/ChemicalPerturbationSignatures.gmt"))
GBM_hgtest_tbl_cpt-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_CCLE,hyper_geo_reference=chemi
```

```
ppt<-AMARETTO_HTMLreport(AMARETTOinit_CCLE,
                         AMARETTOresults_CCLE,
                         ProcessedData_CCLE,
                         hyper_geo_reference=CCLE_hgtest_tbl,
                         genetic_pert_hyper_geo_reference = CCLE_hgtest_tbl_gp,
                         chem_pert_hyper_geo_reference = CCLE_hgtest_tbl_cp,
                         output_address='./AMARETTO_results/CCLE_Report',
                         show_row_names = TRUE,
                         driverGSEA = TRUE)
```

A3. AMARETTO for tcHCV

Loading Gene Expression (MA) data from TCGA LIHC patient cohort (Required) :

```
eadRDS(file = url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/tcHCVnormdata.rds"))
```

Candidate Driver Genes (Optional) :

```
s)
s<-Driver_Genes$Cancer_Gene_Census
```

TO core for regulatory network inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are
specified NrModules and VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel
the combination method for (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
HCV = list(MA_matrix=tcHCVnormdata, CNV_matrix=NULL, MET_matrix=NULL)
CV <- AMARETTO_Initialize(ProcessedData=ProcessedData_tcHCV,
                          Driver_list = candidate_drivers,
                          method = "union",
                          NrCores = 4,
                          NrModules = 150,
                          VarPercentage = 75,
                          random_seeds = c(0,0))
tcHCV<-AMARETTO_Run(AMARETTOinit_tcHCV)
```

et Enrichment Analysis :

```
-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/H_C2_genesets.gmt"),
  url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt"),
  url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StromalSignatures.gmt"),
  url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StemnessSignatures.gmt"))
rGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_tcHCV,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

et Enrichment Analysis for driver perturbations :

```
url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/GeneticPerturbationSignatures.gmt"),
url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ChIPABindingSignatures.gmt"),
url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ConsensusBindingSignatures.gmt"),
url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/EncodeBindingSignatures.gmt"))
yperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_tcHCV,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

et Enrichment Analysis for drug perturbations :

```
(url("./signature/ChemicalPerturbationSignatures.gmt"))
yperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_tcHCV,hyper_geo_reference=chemical_gmts,driverGSEA=TRUE, NrCores=5)
```

es and statistical tests data, and performing performing the phenotype association tests :

```
V<-read.csv(read_delim(url('http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/Samples_tcHCV_phenotypes.txt'),'\t')
tcHCV<-read.csv(read_delim(url('http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/Samples_tcHCV_statistics.txt'),delim="\t",col_
tests_all<-AMARETTO_unite_results(AMARETTO_statistical_test(AMARETTOinit_tcHCV,AMARETTOresults_tcHCV,samples_tcHCV,phenotype_tests_tcHCV))
```

TO HTML report :

```
Report(AMARETTOinit_tcHCV,
       AMARETTOresults_tcHCV,
       ProcessedData_tcHCV,
       hyper_geo_reference=tcHCV_hgtest_tbl,
       genetic_pert_hyper_geo_reference=tcHCV_hgtest_tbl_gp,
       chem_pert_hyper_geo_reference=tcHCV_hgtest_tbl_cp,
       output_address='./AMARETTO_results/tcHCV_Report',
       SAMPLE_annotation=samples_tcHCV_HCV,
       file='sample',
       show_row_names=TRUE,
       phenotype_association_table =tcHCV_phenotype_tests_all,
       driverGSEA = TRUE)
```

l for tcHBV

ression (MA) data from TCGA LIHC patient cohort (Required) :

```
geo_reference = cAMARETTO_hgtest_tbl,
per_reference_gp = cAMARETTO_hgtest_tbl_gp,
per_reference_cp = cAMARETTO_hgtest_tbl_cp,
_address= "./community_Amaretto_report/",
eneTablesList=TCGA_LIHC=LIHC_phenotype_tests_all,CCLE_Liver=NULL,tcHCV=tcHCV_phenotype_tests_all,tcHBV=tcHBV_p
```

al inference of regulatory networks using *AMARETTO. *In preparation for submission.*

, Oosterwijk T, Gentles A.J., Pochet N., Gevaert O. (2018). Module Analysis Captures Pancancer Genetically and Epigenetically r Genes for Smoking and Antiviral Response. *EBioMedicine*, 27:156-166. PMID:29331675 PMCIDPMC5828545.

Sikic B.I., Plevritis S.K. (2014). Identification of ovarian cancer driver genes by using module network integration of multi-omics data. 30013. PMID:24511378 PMCIDPMC3915833.

Lenkat K., Goldsmith A.J., Gevaert O. (2014). CaMoDi: a new method for cancer module discovery. *BMC Genomics*, 15 Suppl 10:S8. MC4304219.

013). Identifying master regulators of cancer and their downstream targets by integrating genomic and epigenomic features. *Pacific ting*, 2013:123-34. PMID:23424118 PMCIDPMC3911770.

renman K., Gevaert O. (2018). MethylMix 2.0: an R package for identifying DNA methylation genes. *Bioinformatics*, 34(17):3044-3046. MC6129298.

Plevritis S.K. (2015). Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biology*, 16(1):17. MC4305593.

lMix: an R package for identifying DNA methylation-driven genes. *Bioinformatics*, 31(11):1839-41. PMID:25609794

man S., Glass K., Pochet N., Everaert C., Raby B., Carey V. (2019). TFutils: Data structures for transcription factor bioinformatics. https://doi.org/10.12688/f1000research.17976.1).

F., Thorvaldsdottir H., Liefeld T., Ocana M., Borges-Rivera D., Pochet N., Robinson J.T., Demchak B., Hull T., Ben-Artzi G., Blankenberg uhn R.M., Nekrutenko A., Segal E., Ideker T., Reich M., Regev A., Chang H.Y., Mesirov J.P. (2016). Integrative genomic analysis by matics tools in GenomeSpace. *Nature Methods*, 13(9):745-747. PMID:26780094 PMCID:PMC4976623.

11. Reich M., Liefeld T., Ocana M., Jang D., Bistline J., Robinson J., Carr P., Hill B., McLaughlin J., Pochet N., Borges-Rivera D., Tabor T., Thorvaldsdottir H., Regev A., Mesirov J.P. (2013). GenomeSpace: an environment for frictionless bioinformatics. *F1000Posters*, 4:804. (https://f1000research.com/posters/1093972).

Funding

This work was supported by grants from NIH NCI ITCR R21 CA209940 (Pochet), NIH NCI ITCR U01 CA214846 Collaborative Supplement (Carey/Pochet) and NIH NIAID R03 AI131066 (Pochet).

Questions?

nsert (cgevaert@stanford.edu)

...

https://colab.research.google.com/drive/17GieTfYriTVlbKchl-OEb5nI_NA2vvjQ#scrollTo=uMfrol3e-0NU

*AMARETTO R Jupyter Notebook Use Case 2

*AMARETTO Use Case 2: glioblastoma multiforme and low-grade glioma

Mohsen Nabian#, Jayendra Shinde#, Celine Everaert#, Shaimaa Bakr#, Ted Liefeld, Thorin Tabor, Charles Blatti, Thomas Baumert, Michael Reich, Jill Mesirov, Mikkel Hemaez*, Vincent Carey*, Olivier Gevaert*, Nathalie Pochet*

Preparing...

The following commands are to prepare, install and load *AMARETTO* and *Community AMARETTO* packages. These installation process must be done again everytime the notebook is reopened.

```
system("sudo apt-get install libv8-dev", intern = TRUE, ignore.stderr = TRUE)
```

```
devtools::install_github("gevaertlab/AMARETTO", ref="for35_develop", dependencies=TRUE)
```

```
library("AMARETTO")
```

```
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
```

```
BiocManager::install("ComplexHeatmap")
```

```
BiocManager::install("rDyn")
```

```
devtools::install_github("broadinstitute/CommunityAMARETTO", ref="master", dependencies=TRUE)
```

```
library("CommunityAMARETTO")
```

A. AMARETTO for each of the 2 brain disease data sets

A1. AMARETTO for TCGA GBM

Loading Multi-Omics data including Gene Expression (MA), Copy Number Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

```
MA_matrix_GBM <-AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_GBM_Expression.gct"))
CNV_matrix_GBM <-AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_GBM_CNV.gct"))
MET_matrix_GBM <-AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_GBM_Methylation.gct"))
```

Defining List(s) of Candidate Driver Genes (Optional) :

```
data(Driver_Genes)
candidate_drivers<-Driver_Genes$Cancer_Gene_Census
```

Running AMARETTO core for regulatory network inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are required. Here we specified NrModules and VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel processing. As for the combination method for (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
ProcessedData_TCGA_GBM = list(MA_matrix=MA_matrix_GBM, CNV_matrix=CNV_matrix_GBM, MET_matrix=MET_matrix_GBM)
AMARETTOinit_GBM <- AMARETTO_Initialize(ProcessedData = ProcessedData_TCGA_GBM,
                                        Driver_list = all_drives_GBM,
                                        method = "union",
                                        NrModules = 150,
                                        VarPercentage = 75,
                                        NrCores = 4)

AMARETTO_Results_GBM<-AMARETTO_Run(AMARETTOinit_GBM)
```

Loading phenotypes and statistical tests data, and performing performing the phenotype association tests :

```
sample_annotations_GBM<-read.csv.delim(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/Samples_GBM_phenotypes.txt"),delim="\t")
phenotype_tests_GBM<-read.csv.delim(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/Samples_GBM_statistics.txt"),delim="\t",col_name
GBM_phenotype_tests_all<-AMARETTO_unite_results(AMARETTO_ystatistical_test(AMARETTOinit_GBM,AMARETTOresults_GBM,sample_annotations_GBM_all,phenotype_tests_GBM_all))
```

Performing Geneset Enrichment Analysis :

```
functional_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/H_C2_genesets.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StromalSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StemnessSignatures.gmt"))
hgtest_tbl<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_GB,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

Creating AMARETTO HTML report :

```
AMARETTO_HTMLreport(AMARETTOinit=AMARETTOinit_LGG,
                    AMARETTOresults=AMARETTOresults_LGG,
                    ProcessedData_TCGA_LGG,
                    hyper_geo_reference=LGG_hgtest_tbl,
                    output_address='./AMARETTO_results/LGG_Report',
                    SAMPLE_annotations=sample_annotations_LGG_all,
                    ID='Sample',
                    show_row_names=FALSE,
                    phenotype_association_table = LGG_phenotype_tests_all)
```

B. Community-AMARETTO for integrating 2 brain disease data sets (cfr. cAMARETTO old Notebook)

Multiscale and multimodal inference of regulatory networks to identify cell circuits and their drivers shared and distinct within and across biological systems of human disease

Loading additional networks as a set of signatures in .GMT format (Optional)

```
ab/demo/NotebookExample/ExampleData/GeneticPerturbationSignatures.gmt"),
ab/demo/NotebookExample/ExampleData/ChIPBindingSignatures.gmt"),
ab/demo/NotebookExample/ExampleData/ConsensusBindingSignatures.gmt"),
ab/demo/NotebookExample/ExampleData/EncodeBindingSignatures.gmt"))
results_GBM,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

```
gs.gmt"))
RTTOresults_GBM,hyper_geo_reference=chemical_gmts,driverGSEA=TRUE, NrCores=5)
```

```
est_tbl_gp,
_tbl_cp,
ort',
ll,

e_tests_all)
```

er Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

```
stitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LGG_Expression.gct"))
nstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LGG_CNV.gct"))
nstitute.org/pochetlab/demo/NotebookExample/ExampleData/TCGA_LGG_Methylation.gct"))
```

regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are
d 75 respectively. We can optionally specify number of cores (NrCores) for parallel
) the predefined list of drivers, we specified 'union' (as opposed to "intersection").

```
rix=CNV_matrix_LGG, MET_matrix=MET_matrix_LGG)
eaData_TCGA_LGG,
ves_LGG,


2))
```

ng the phenotype association tests :

```
institute.org/pochetlab/demo/NotebookExample/ExampleData/sample_annotations_LGG_all.csv"))
titute.org/pochetlab/demo/NotebookExample/ExampleData/phenotype_tests_LGG_all.csv"))
stical_test(AMARETTOinit_LGG,AMARETTOresults_LGG,sample_annotations_LGG_all,phenotype_tests_LGG_all))
```

```
functional_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/H_C2_genesets.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StromalSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StemnessSignatures.gmt"))
hgtest_tbl<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_LGG,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

Performing Geneset Enrichment Analysis for driver perturbations :

```
genetic_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/GeneticPerturbationSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ChIPBindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ConsensusBindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/EncodeBindingSignatures.gmt"))
hgtest_tbl_gp<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_LGG,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

Performing Geneset Enrichment Analysis for drug perturbations :

```
chemical_gmts<-c(url("./signature/ChemicalPerturbationSignatures.gmt"))
GBM_hgtest_tbl_cp<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_LGG,hyper_geo_reference=chemical_gmts,driver
```

```
portals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/ImmuneSignatures.gmt")
rtals.broadinstitute.org/pochetlab/demo/NotebookExample/ExampleData/StemSignatures.gmt")
s = StemSignatures, ImmuneSignatures = ImmuneSignatures)
```

orithm which computes the extent of overlaps between all regulatory modules learned across above data cohorts.

```
ults(cAMARETTOresults = list(TCGA_GBM=AMARETTOresults_GBM,TCGA_LGG=AMARETTOresults_LGG),
                gmt_filelist = gmt_filelist,
                NrCores = 4,
                output_dir = "./",
                drivers=TRUE)
```

nodule networks across data cohorts: Additional parameters that can be set by the user in this step include:

fes or communities based on p-value significance. Edges in the network with p-values larger than the cutoff value will be filtered
05.

fes or communities based on the minimum number overlapping genes. Edges in the network with the number of overlapping
l be filtered out. The default cutoff value is 5 overlapping genes.

```
duleNetwork(cAMARETTOresults,
                        pvalue = 0.10,
                        inter = 5)
entifyCom(cAMARETTOnetworkM,filtercomm = FALSE)
```

alysis on communities :

```
/AMARETTO::CreateHyperGeoTest1l(cAMARETTOresults,cAMARETTOnetworkM,cAMARETTOnetworkC,hyper_geo_reference=functional_gmts,driverGSEA=TRUE,N
```

alysis for driver perturbations on communities :

```
ity/AMARETTO::CreateHyperGeoTest1l(cAMARETTOresults,cAMARETTOnetworkM,cAMARETTOnetworkC,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE,N
```

alysis for chemical perturbations on communities :

```
ity/AMARETTO::CreateHyperGeoTest1l(cAMARETTOresults,cAMARETTOnetworkM,cAMARETTOnetworkC,hyper_geo_reference=chemical_gmts,driverGSEA=TRUE,
```

TML report :

```
esults,
etworkM,
etworkC,
ITCGAlist = c("TCGA_GBM"="./AMARETTO_results/GBM_Report","TCGA_LGG"="./AMARETTO_results/LGG_Report"),
TTOReport = TRUE,
_reference = c(AMARETTO_hgtest_tbl,
_reference_gp = c(AMARETTO_hgtest_tbl_gp,
_reference_cp = c(AMARETTO_hgtest_tbl_cp,
browse "./community_Amaretto_report/",
tables1=list1=list(TCGA_GBM=GBM_phenotype_tests_all,TCGA_LGG=LGG_phenotype_tests_all))
```

ools of the current *AMARETTO toolbox and future developments are available from GitHub, Bioconductor, GenePattern, ook and R Jupyter Notebook.

: https://www.bioconductor.org/packages/devel/bioc/html/AMARETTO.html
Bioconductor: in preparation for submission

...

***AMARETTO in GitHub**

- **AMARETTO in GitHub**: https://github.com/gevaertlab/AMARETTO
- **Community-AMARETTO in GitHub**: https://github.com/broadinstitute/CommunityAMARETTO

***AMARETTO in GenePattern**

- **AMARETTO in GenePattern**: https://cloud.genepattern.org/gp/pages/index.jsf?lsid=urn:lsid:broad.mit.edu:cancer.software.genepattern.module.analysis:00378
- **Community-AMARETTO in GenePattern**: https://cloud.genepattern.org/gp/pages/index.jsf?lsid=urn:lsid:broad.mit.edu:cancer.software.genepattern.module.analysis:00381

***AMARETTO in GenomeSpace**

The AMARETTO and Community AMARETTO modules in GenoFiltam are also available within GenomeSpace: http://www.genomespace.org/

***AMARETTO Notebooks**

The ***AMARETTO in GenePattern and R Notebooks** provide users with a complete analysis pipeline that enables running AMARETTO on one or multiple data cohorts and connecting them using Community-AMARETTO. Each AMARETTO and Community-AMARETTO analysis generates a detailed report of genome-wide networks inferred from one cohort and/or shared/distinct across multiple cohorts. These reports include queryable tables and visualizations (heatmaps and network graphs) of shared/distinct cell circuits and their drivers, as well as their functional and phenotypic characterizations.

- ***AMARETTO in GenePattern Notebook**: GenePattern Notebook from https://notebook.genepattern.org/
- ***AMARETTO in R via GitHub and Bioconductor**: Jupyter Notebook from https://colab.research.google.com/

***AMARETTO example reports**

Studying hepatitis C & B virus-induced hepatocellular carcinoma using AMARETTO & Community-AMARETTO:

- An example report that learns regulatory networks from multi-omics data for hepatocellular carcinoma based on integrating genetic, epigenetic and functional genomics data from TCGA: AMARETTO Report
- An example report that integrates regulatory networks derived from >6 liver data sources (multi-omics hepatocellular carcinoma patient data from TCGA, ~25 liver cell line models from CCLE, time course hepatitis C virus infection data in Huh7 models, time course hepatitis B virus infection data in HepG2 models, single cell hepatitis C virus infection data in Huh7 models, single cell hepatitis B virus infection data in HepG2 models, further augmented with previously published prognostic network models that were derived from hepatocellular carcinoma patient data): Community-AMARETTO Report
- An example of ongoing work on developing gene-level ontology network representations from AMARETTO modules & Community-AMARETTO communities: Shiny App

Multi-omics & imaging data fusion for glioblastoma multiforme using AMARETTO:

- An example report that integrates imaging data into the multi-omics regulatory networks for glioblastoma multiforme based on multi-omics and non-invasive imaging data from TCGA/TCIA (that we will later connect with networks learned from integrating RNA-Seq refined for anatomic structures and stem cells with histopathology imaging data from IvyGAP and that we will subsequently further refine based on single-cell RNA-Seq studies): AMARETTO Report

References

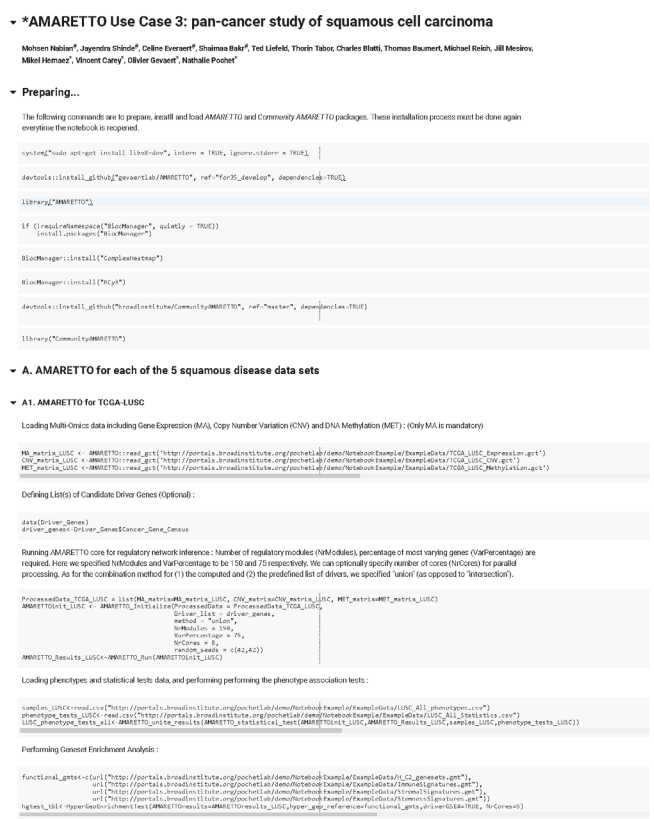
1. Multiscale and multimodal inference of regulatory networks using *AMARETTO. *In preparation for submission.*
2. Champion M., Brennan K., Croonenborghs T., Gentles A.J., Pochet N., Gevaert O. (2018). Module Analysis Captures Pancancer Genetically and Epigenetically Deregulated Cancer Driver Genes for Smoking and Antiviral Response. *EBioMedicine*, 27:156-166. PMID:29331675 PMCID:PMC5828545.
3. Gevaert O., Villalobos V., Sikic B.I., Plevritis S.K. (2014). Identification of ovarian cancer driver genes by using module network integration of multi-omics data. *Interface Focus*, 3(4):20130013. PMID:24511378 PMCID:PMC3915833.
4. Manuchakian A., Ochoa I., Venkat K., Goldsmith A.J., Gevaert O. (2014). CaMoDi: a new method for cancer module discovery. *BMC Genomics*, 15 Suppl 10:S8. PMID:25566933 PMCID:PMC4304219.
5. Gevaert O., Plevritis S. (2013). Identifying master regulators of cancer and their downstream targets by integrating genomic and epigenomic features. *Pacific Symposium on Biocomputing*, 2013:123-34. PMID:23424118 PMCID:PMC3911770.
6. Cedoz P.L., Prunello M., Brennan K., Gevaert O. (2018). MethylMix 2.0: an R package for identifying DNA methylation genes. *Bioinformatics*, 34(17):3044-3046. PMID:29668835 PMCID:PMC6129298.
7. Gevaert O., Tibshirani R., Plevritis S.K. (2015). Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biology*, 16(1):17. PMID:25631659 PMCID:PMC4365533.
8. Gevaert O. (2015). MethylMix: an R package for identifying DNA methylation-driven genes. *Bioinformatics*, 31(11):1839-41. PMID:25609794 PMCID:PMC4443673.
9. Stubbs B.J., Gopalakrishnan S., Glass K., Pochet N., Everaert C., Raby B., Carey V. (2019). TFutils: Data structures for transcription factor bioinformatics. *F1000Research*, 8:152. (https://doi.org/10.12688/f1000research.17976.1).
10. Qu K., Garamszegi S., Wu F., Thorvaldsdottir H., Liefeld T., Ocana M., Borges-Rivera D., Pochet N., Robinson J.T., Demchak B., Hull T., Ben-Artzi G., Blankenberg D., Barber G.P., Lee B.T., Kuhn R.M., Nekrutenko A., Segal E., Ideker T., Reich M., Regev A., Chang H.Y., Mesirov J.P. (2016). Integrative genomic analysis by interoperation of bioinformatics tools in GenomeSpace. *Nature Methods*, 13(3):245-247. PMID:26780094 PMCID:PMC4767623.
11. Reich M., Liefeld T., Ocana M., Jang D., Bistline J., Robinson J., Carr P., Hill B., McLaughlin J., Pochet N., Borges-Rivera D., Tabor T., Thorvaldsdottir H., Regev A., Mesirov J.P. (2013). GenomeSpace: an environment for frictionless bioinformatics. *F1000Posters*, 4:804. (https://f1000research.com/posters/1092972).

Funding

This work was supported by grants from NIH NCI ITCR R21 CA209940 (Pochet), NIH NCI ITCR U01 CA214846 Collaborative Supplement (Carey/Pochet) and NIH

https://colab.research.google.com/drive/11Q0GH6YHCoTZrDfHrUavixAi_f_qGBau#scrollTo=LujO14znmO0J

*AMARETTO R Jupyter Notebook Use Case 3



*AMARETTO Use Case 3: pan-cancer study of squamous cell carcinoma

Mohsen Nabian#, Jayendra Shinde#, Celine Everaert#, Shaimaa Bakr#, Ted Liefeld, Thorin Tabor, Charles Blatti, Thomas Baumert, Michael Reich, Jill Mesirov, Mikkel Hemaez*, Vincent Carey*, Olivier Gevaert*, Nathalie Pochet*

Preparing...

The following commands are to prepare, install and load AMARETTO and Community AMARETTO packages. These installation process must be done again everytime the notebook is reopened.

```
system("sudo apt-get install libv8-dev", intern = TRUE, ignore.stderr = TRUE)

devtools::install_github("gevaertlab/AMARETTO", ref="for35_develop", dependencies=TRUE)

library("AMARETTO")

if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

BiocManager::install("ComplexHeatmap")

BiocManager::install("RCy3")

devtools::install_github("broadinstitute/CommunityAMARETTO", ref="master", dependencies=TRUE)

library("CommunityAMARETTO")
```

A. AMARETTO for each of the 5 squamous disease data sets

A1. AMARETTO for TCGA-LUSC

Loading Multi-Omics data including Gene Expression (MA), Copy Number Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

```
MA_matrix_LUSC <- AMARETTO::read_gct("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_LUSC_Expression.gct")
CNV_matrix_LUSC <-AMARETTO::read_gct("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_LUSC_CNV.gct")
MET_matrix_LUSC <-AMARETTO::read_gct("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_LUSC_Methylation.gct")
```

Defining List(s) of Candidate Driver Genes (Optional) :

```
data(Driver_Genes)
driver_genes<-Driver_Genes$Cancer_Gene_Census
```

Running AMARETTO core for regulatory network inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are required. Here we specified NrModules and VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel processing. As for the combination method for (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

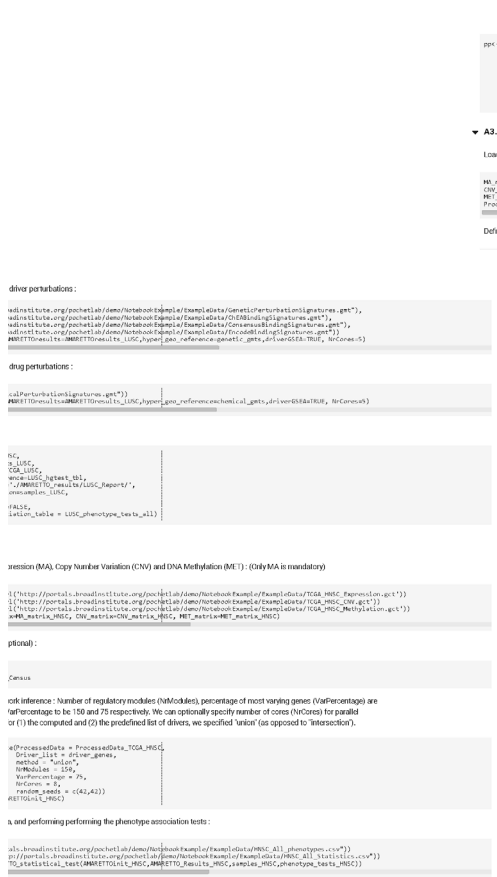
```
ProcessedData_TCGA_LUSC = list(MA_matrix=MA_matrix_LUSC, CNV_matrix=CNV_matrix_LUSC, MET_matrix=MET_matrix_LUSC)
AMARETTOInit_LUSC <- AMARETTO_Initialize(ProcessedData = ProcessedData_TCGA_LUSC,
                                         Driver_list = driver_genes,
                                         method = "union",
                                         NrModules = 150,
                                         VarPercentage = 75,
                                         NrCores = 8,
                                         random_seeds = c(42,42))
AMARETTO_Results_LUSC<-AMARETTO_Run(AMARETTOInit_LUSC)
```

Loading phenotypes and statistical tests data, and performing performing the phenotype association tests :

```
samples_LUSC<-read.csv("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/LUSC_All_phenotypes.csv")
phenotype_tests_LUSC<-read.csv("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/LUSC_All_Statistics.csv")
LUSC_phenotype_tests_all<-AMARETTO_unite_results(AMARETTO_statistical_test(AMARETTOInit_LUSC,AMARETTO_Results_LUSC,samples_LUSC,phenotype_tests_LUSC))
```

Performing Geneset Enrichment Analysis :

```
functional_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/H_C2_genesets.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ImmuneSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StromalSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StemnessSignatures.gmt"))
hgtest_tbl<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_LUSC,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```



driver perturbations :

```
ndinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/GeneticPerturbationSignatures.gmt"),
ndinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ChIPBindingSignatures.gmt"),
ndinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ConsensusBindingSignatures.gmt"),
ndinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/EncodeBindingSignatures.gmt"))
AMARETTOresults=AMARETTOresults_LUSC,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

drug perturbations :

```
calPerturbationSignatures.gmt"))
AMARETTOresults=AMARETTOresults_LUSC,hyper_geo_reference=chemical_gmts,driver=GSEA=TRUE, NrCores=5)
```

```
SC,
s_LUSC,
CGA_LUSC,
ence=LUSC_hgtest_tbl,
"./AMARETTO_results/LUSC_Report/',
ons=samples_LUSC,

rFALSE,
iation_table = LUSC_phenotype_tests_all)
```

xpression (MA), Copy Number Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

```
1("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_HNSC_Expression.gct"))
1("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_HNSC_CNV.gct"))
1("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_HNSC_Methylation.gct"))
a=MA_matrix_HNSC, CNV_matrix=CNV_matrix_HNSC, MET_matrix=MET_matrix_HNSC)
```

ptional) :

```
Census
```

ork inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are
VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel
or (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
e(ProcessedData = ProcessedData_TCGA_HNSC,
  Driver_list = driver_genes,
  method = "union",
  NrModules = 150,
  VarPercentage = 75,
  NrCores = 8,
  random_seeds = c(42,42))
#ETTOInit_HNSC)
```

a, and performing performing the phenotype association tests :

```
als.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/HNSC_All_phenotypes.csv"))
ps://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/HNSC_All_Statistics.csv"))
TO_statistical_test(AMARETTOInit_HNSC,AMARETTO_Results_HNSC,samples_HNSC,phenotype_tests_HNSC))
```

```
functional_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/H_C2_genesets.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ImmuneSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StromalSignatures.gmt"),
                   url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StemnessSignatures.gmt"))
hgtest_tbl<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_HNSC,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

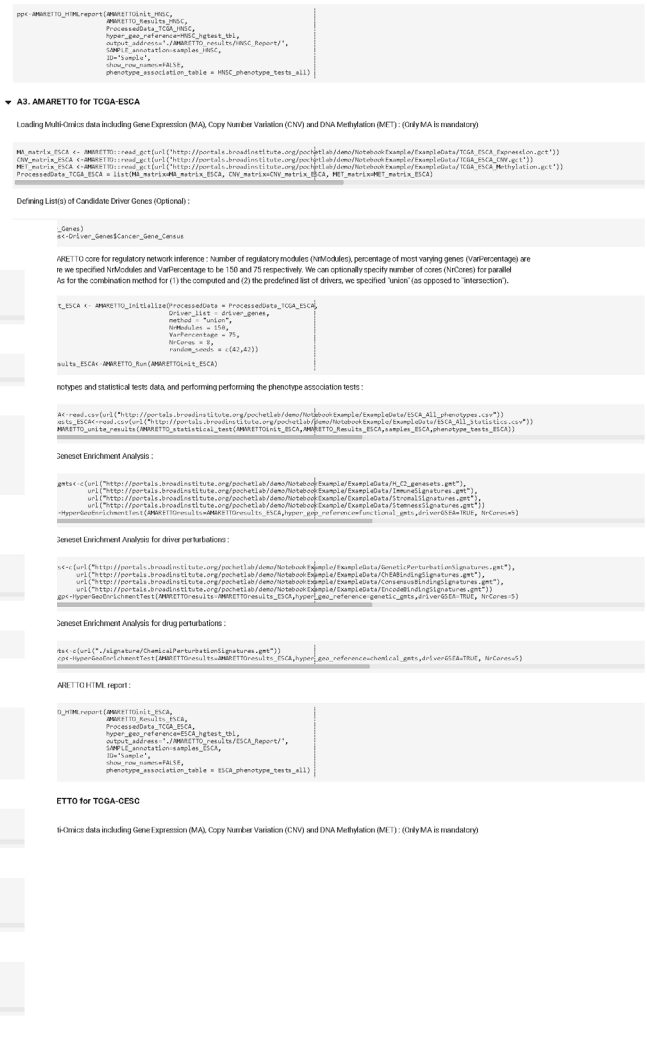
Performing Geneset Enrichment Analysis for driver perturbations :

```
genetic_gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/GeneticPerturbationSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ChIPBindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ConsensusBindingSignatures.gmt"),
                url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/EncodeBindingSignatures.gmt"))
hgtest_tbl_gpc<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_HNSC,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

Performing Geneset Enrichment Analysis for drug perturbations :

```
chemical_gmts<-c(url("./signature/ChemicalPerturbationSignatures.gmt"))
hgtest_tbl_cpc<-HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_HNSC,hyper_geo_reference
```

Creating AMARETTO HTML report :



```
ppc<-AMARETTO_HTMLreport(AMARETTOInit_HNSC,
                         AMARETTO_Results_HNSC,
                         ProcessedData_TCGA_HNSC,
                         hyper_geo_reference=HNSC_hgtest_tbl,
                         output_address='./AMARETTO_results/HNSC_Report/',
                         SAMPLE_annotation=samples_HNSC,
                         ID='Sample',
                         show_row_names=FALSE,
                         phenotype_association_table = HNSC_phenotype_tests_all)
```

A3. AMARETTO for TCGA-ESCA

Loading Multi-Omics data including Gene Expression (MA), Copy Number Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

```
MA_matrix_ESCA <- AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_ESCA_Expression.gct"))
CNV_matrix_ESCA <-AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_ESCA_CNV.gct"))
MET_matrix_ESCA <-AMARETTO::read_gct(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/TCGA_ESCA_Methylation.gct"))
ProcessedData_TCGA_ESCA = list(MA_matrix=MA_matrix_ESCA, CNV_matrix=CNV_matrix_ESCA, MET_matrix=MET_matrix_ESCA)
```

Defining List(s) of Candidate Driver Genes (Optional) :

```
_Genes)
s<-Driver_Genes$Cancer_Gene_Census
```

ARETTO core for regulatory network inference : Number of regulatory modules (NrModules), percentage of most varying genes (VarPercentage) are
re we specified NrModules and VarPercentage to be 150 and 75 respectively. We can optionally specify number of cores (NrCores) for parallel
As for the combination method for (1) the computed and (2) the predefined list of drivers, we specified 'union' (as opposed to 'intersection').

```
t_ESCA <- AMARETTO_Initialize(ProcessedData = ProcessedData_TCGA_ESCA,
                              Driver_list = driver_genes,
                              method = "union",
                              NrModules = 150,
                              VarPercentage = 75,
                              NrCores = 8,
                              random_seeds = c(42,42))
sults_ESCA<-AMARETTO_Run(AMARETTOInit_ESCA)
```

notypes and statistical tests data, and performing performing the phenotype association tests :

```
A<-read.csv(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ESCA_All_phenotypes.csv"))
sts_ESCA<-read.csv(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ESCA_All_Statistics.csv"))
AMARETTO_unite_results(AMARETTO_statistical_test(AMARETTOInit_ESCA,AMARETTO_Results_ESCA,samples_ESCA,phenotype_tests_ESCA))
```

Geneset Enrichment Analysis :

```
gmts<-c(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/H_C2_genesets.gmt"),
        url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ImmuneSignatures.gmt"),
        url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StromalSignatures.gmt"),
        url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/StemnessSignatures.gmt"))
=HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_ESCA,hyper_geo_reference=functional_gmts,driverGSEA=TRUE, NrCores=5)
```

Geneset Enrichment Analysis for driver perturbations :

```
s<-c(url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/GeneticPerturbationSignatures.gmt"),
     url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ChIPBindingSignatures.gmt"),
     url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/ConsensusBindingSignatures.gmt"),
     url("http://portals.broadinstitute.org/pochetlab/demo/Notebook-Example/ExampleData/EncodeBindingSignatures.gmt"))
gpc=HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_ESCA,hyper_geo_reference=genetic_gmts,driverGSEA=TRUE, NrCores=5)
```

Geneset Enrichment Analysis for drug perturbations :

```
tsc<-c(url("./signature/ChemicalPerturbationSignatures.gmt"))
cpc=HyperGeoEnrichmentTest(AMARETTOresults=AMARETTOresults_ESCA,hyper_geo_reference=chemical_gmts,driverGSEA=TRUE, NrCores=5)
```

ARETTO HTML report :

```
O_HTMLreport(AMARETTOInit_ESCA,
             AMARETTO_Results_ESCA,
             ProcessedData_TCGA_ESCA,
             hyper_geo_reference=ESCA_hgtest_tbl,
             output_address='./AMARETTO_results/ESCA_Report/',
             SAMPLE_annotation=samples_ESCA,
             ID='Sample',
             show_row_names=FALSE,
             phenotype_association_table = ESCA_phenotype_tests_all)
```

ETTO for TCGA-CESC

ti-Omics data including Gene Expression (MA), Copy Number Variation (CNV) and DNA Methylation (MET) : (Only MA is mandatory)

...

sborghs T., Gentles A.J., Pochet N., Gevaert O. (2018). Module Analysis Captures Pancancer Genetically and Epigenetically for Smoking and Antiviral Response. *EBioMedicine*, 27:156-166. PMID:29331675 PMCID:PMC5828545.

, Plevritis S.K. (2014). Identification of ovarian cancer driver genes by using module network integration of multi-omics data. PMID:24511378 PMCID:PMC3915833.

., Goldsmith A.J., Gevaert O. (2014). CoMoDE: a new method for cancer module discovery. *BMC Genomics*, 15 Suppl 10:S8. 4219.

ntifying master regulators of cancer and their downstream targets by integrating genomic and epigenomic features. *Pacific* 13:123-34. PMID:23424118 PMCID:PMC3911770.

K., Gevaert G. (2016). MethylMix 2.0: an R package for identifying DNA methylation genes. *Bioinformatics*, 34(17):3044-3046. 9298.

S.K. (2015). Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biology*, 16(1):17. 5533.

R package for identifying DNA methylation-driven genes. *Bioinformatics*, 31(11):1839-41. PMID:25609794

Glass K., Pochet N., Everaert C., Raby B., Carey V. (2019). TFutils: Data structures for transcription factor bioinformatics. 2.org/10.12688/f1000research.17976.1).

valdsdottir H., Liefeld T., Ocana M., Borges-Rivera D., Pochet N., Robinson J.T., Demchak B., Hull T., Ben-Artzi G., Blankenberg t., Nekrutenko A., Segal E., Ideker T., Reich M., Regev A., Chang H.Y., Mesirov J.P. (2016). Integrative genomic analysis by ools in GenomeSpace. *Nature Methods*, 13(9):245-247. PMID:26780094 PMCID:PMC4767623.

ng D., Bistline J., Robinson J., Carr P., Hill B., McLaughlin J., Pochet N., Borges-Rivera D., Tabor T., Thorvaldsdottir H., Regev A., ce: an environment for frictionless bioinformatics. *F1000Posters*, 4:804. (https://f1000research.com/posters/1093972).

n NIH NCI ITCR R21 CA209940 (Pochet), NIH NCI ITCR U01 CA214846 Collaborative Supplement (Carey/Pochet) and NIH

l Notebooks, please contact **Nathalie Pochet** (npochet@broadinstitute.org) and **Olivier Gevaert** (ogevaert@stanford.edu).

https://colab.research.google.com/drive/17RwBxwWWnXJMRI_VZl-X-hztJTwvlFjl#scrollTo=WY-C5lDxdDYw

Team: Lab & Collaborators

Pochet Lab (BWH/HMS/Broad)

Mohsen Nabian

Celine Everaert

Rileen Sinha

Tom Croonenborghs

Carey Lab (BWH/HMS/Broad)

Vincent Carey

Regev Lab (MIT/Broad)

Aviv Regev

Brian Haas

Gevaert Lab (Stanford/Broad)

Olivier Gevaert

Jayendra Ravindra Shinde

Shaimaa Hesham Bakr

Andrew Gentles

Kevin Brennan

Magali Champion

Mesirov Lab (UCSD/Broad)

Jill Mesirov

Michael Reich

Ted Liefeld

Thorin Tabor

Hernaez Lab (Illinois)

Mikel Hernaez

Baumert Lab (Strasbourg)

Thomas Baumert

Joachim Lupberger

Eloi Verrier

Quintana Lab (BWH/HMS/Broad)

Francisco Quintana

Krichevsky Lab (BWH/HMS)

Anna Krichevsky

Funding Sources

NIH NCI ITCR R21 CA209940 (Pochet)

NIH NIAID R03 AI131066 (Pochet)

NIH NCI ITCR U01 CA214846 collaborative set-aside (Carey/Pochet)