The Cancer Target Discovery and Development (CTD^2) Network RFA Concept

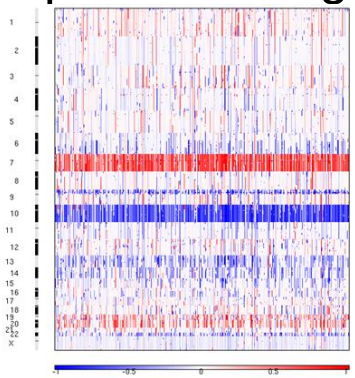
NCI Board of Scientific Advisers
March 29, 2016

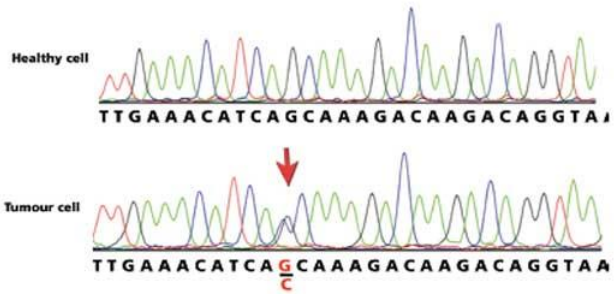
Daniela S. Gerhard, Ph.D.
OCG, CCG

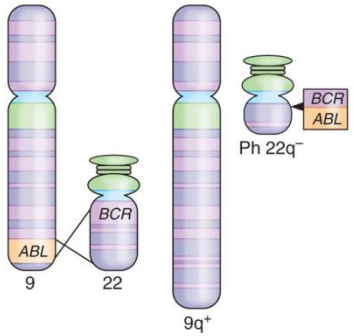
NIH NATIONAL CANCER INSTITUTE

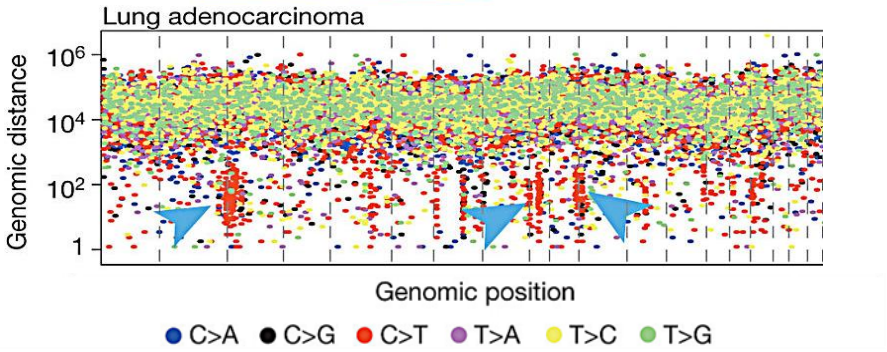
Molecular Characterization is Essential but not Sufficient for Precision Oncology

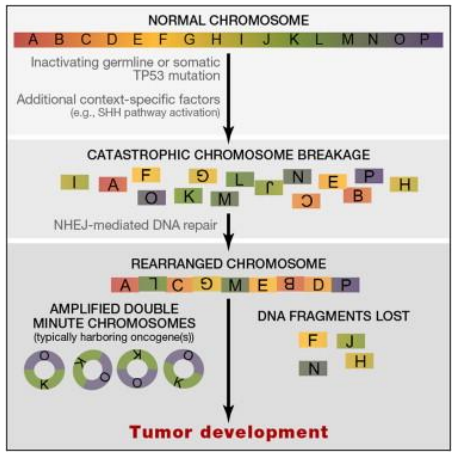
- Each tumor has hundreds to thousands genomic alterations
  - Amplifications, deletions, translocations, chromothripsis, kataegis, epigenetic changes and mutations all impacting expression of genes and behavior of cells











Molecular Characterization is Essential but not Sufficient for Precision Oncology

- Little is known about the cellular function of most genes, much less how the cancer-associated alternations affect it
  - Distinguishing initiating vs. driver vs. contributing vs. passenger alterations in the context of intra- and inter-tumor heterogeneity
    - Drivers are genes involved in tumor maintenance;



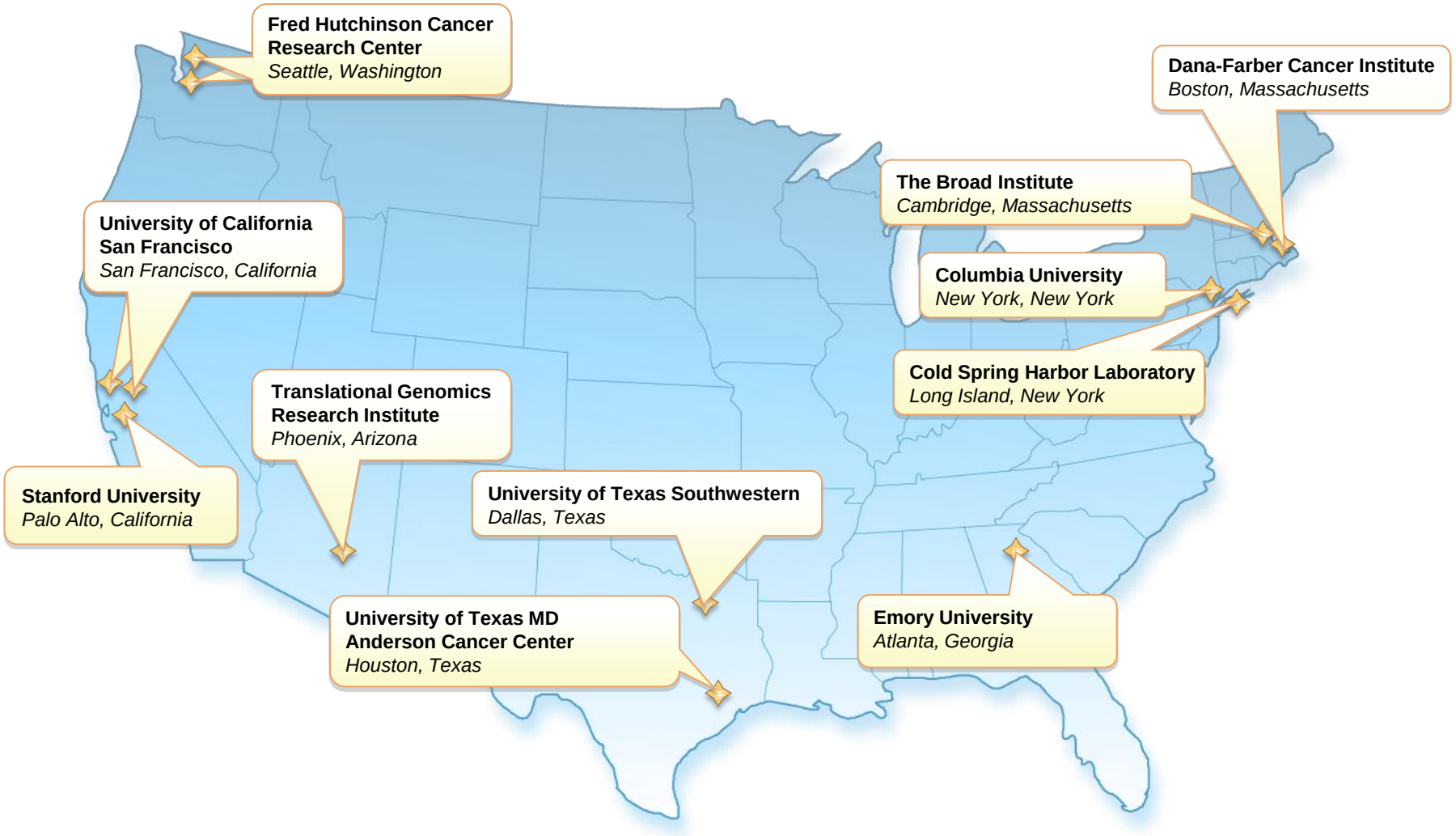
  - Cancer alterations have context-specific impact
    - Context includes cell of origin, other molecular alterations in genes that may have synergistic or antagonistic impact
      - For example, NOTCH can be an oncogene or TSG
  - Adaptation of cancer to environmental stimuli
    - Therapy
    - Signals from adjacent tissues

Response to the RFA: CTD² Multi-disciplinary Teams Were Established

To use genome-scale experimental approaches to address these challenges and identify cancer drivers, therapeutic targets, predictive biomarkers, pertubagens (e.g. small molecules, RNAi, etc.)

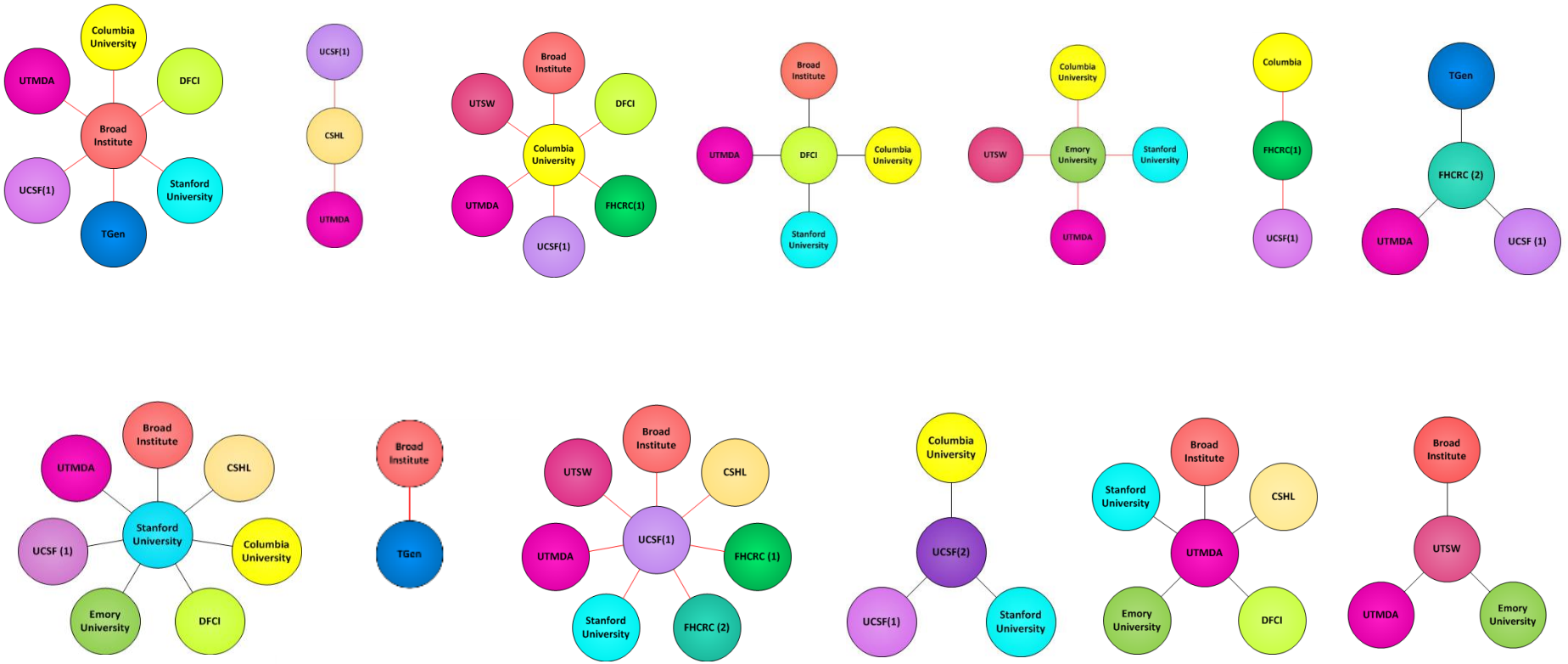
- Computation across comprehensive data sets
  - Single organ cancer type
  - Across multiple cancers
- High-throughput, high-content screening
  - Small molecules
  - RNAi
  - Protein-protein interactions
  - Combination of small molecules
- Integrate results and iterate to improve predictions
- Collaborate intra-Network
- Share all outcomes through easily accessible web pages and reagents through distributors

CTD² Network: 13 Centers



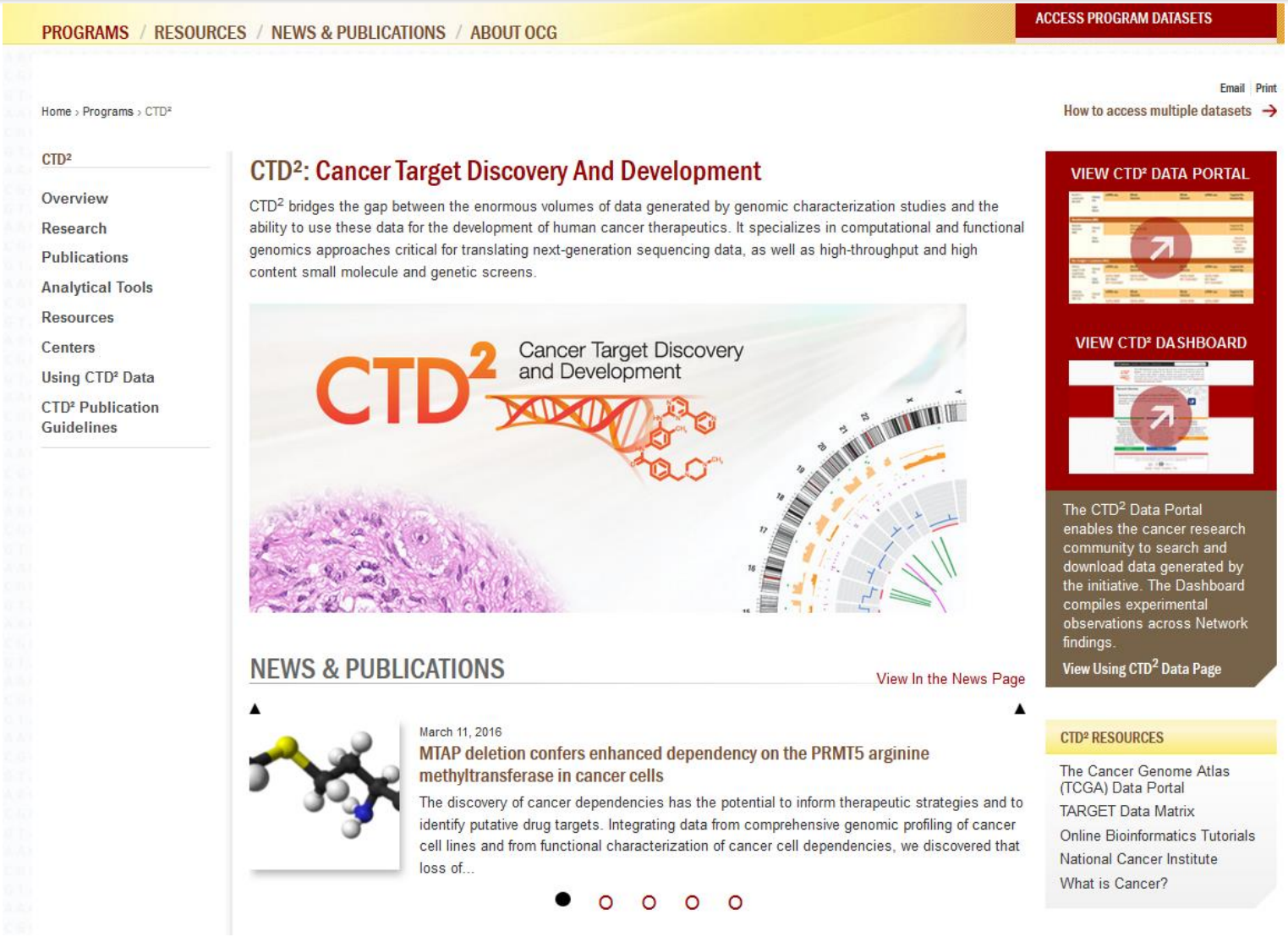
CTD^2 Network Centers

CTD² Centers Collaborate: RFA Requirement



2015

CTD² Web Home Page

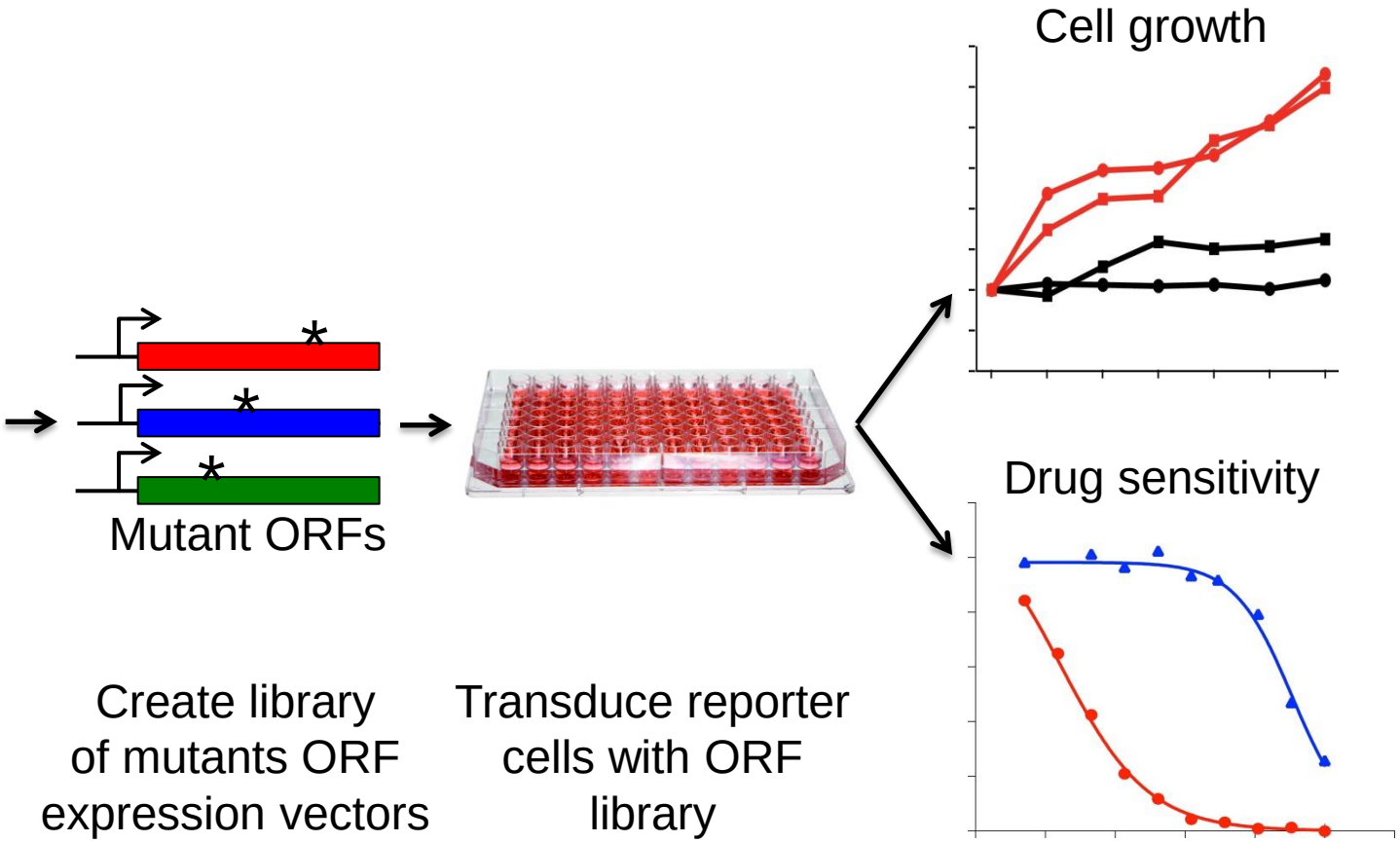


https://ocg.cancer.gov/programs/ctd2

CTD^2 Initiative

- A highly functional network that effectively addresses a major scientific challenge in cancer research
  - Efficient transition from patient-based large multi-dimensional genomic data → target validation → small molecule modulators → therapy
- Continuously improves and innovates approaches
- Forms rapid, pre-competitive collaborations
- 132 publications and counting (20 cited >40 times)
- Shifts current research paradigms in translation of patient-derived multidimensional genetic data to the clinic and utilize novel concepts, approaches and methodologies
  - Neomorphic functions are the norm, MDACC
  - New generation of CRISPRs tools, UCSF-1

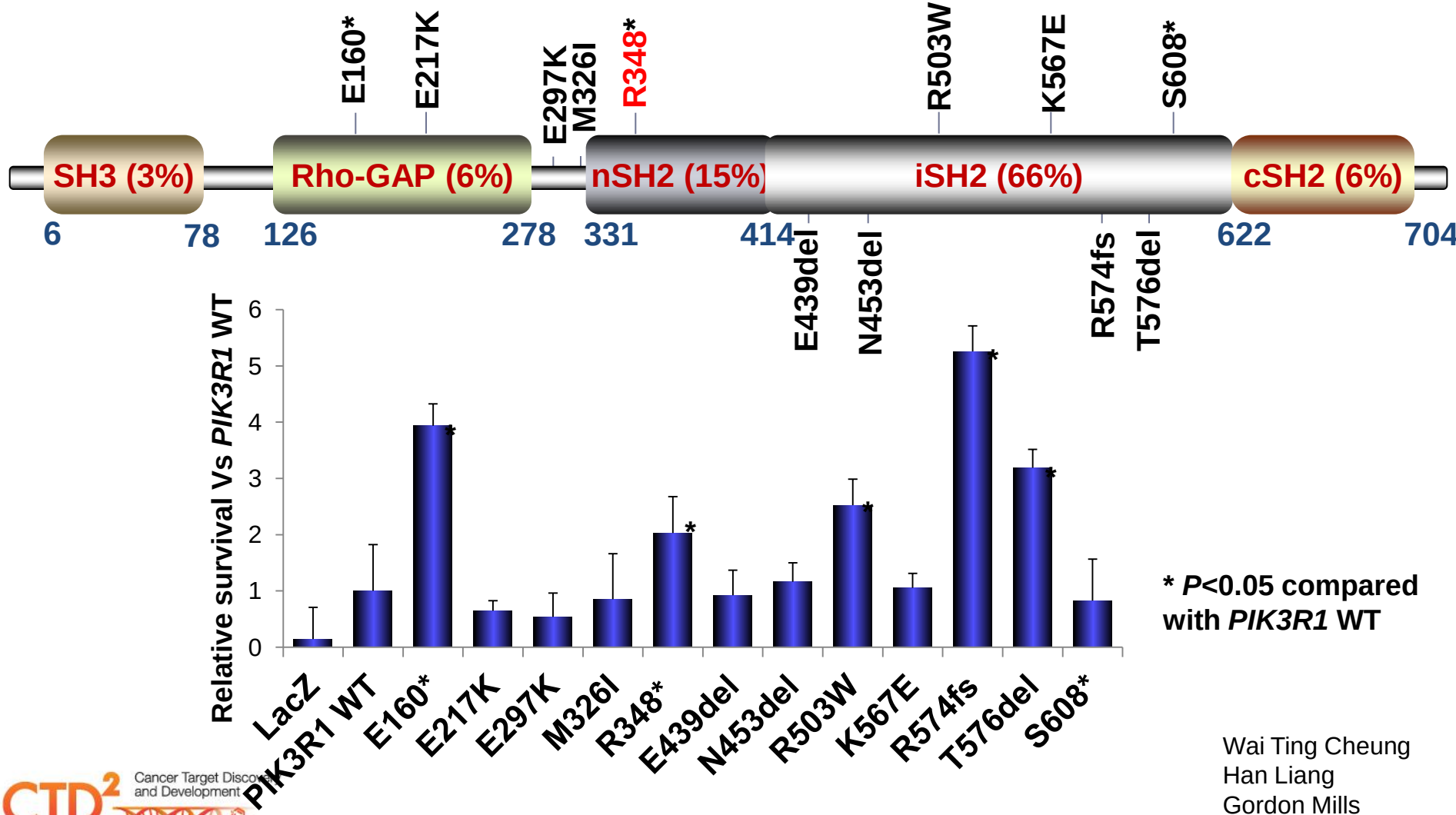
Functional Validation of Cancer Variants



Adapted by L. Staudt from G. Mills

PIK3R1 Functional Mutations:

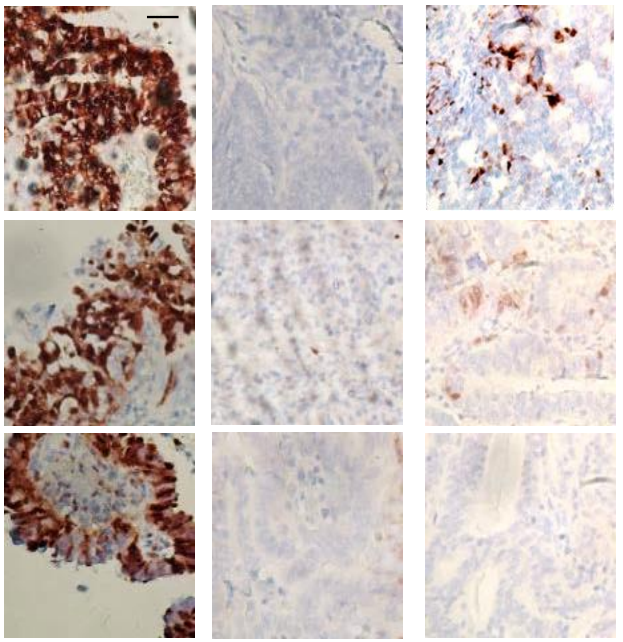
p85 binding partner of p110 catalytic domain of PIK3CA



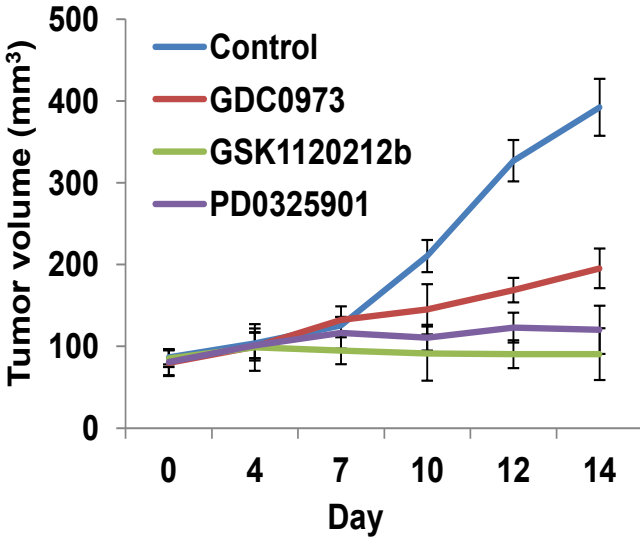
Wai Ting Cheung
Han Liang
Gordon Mills

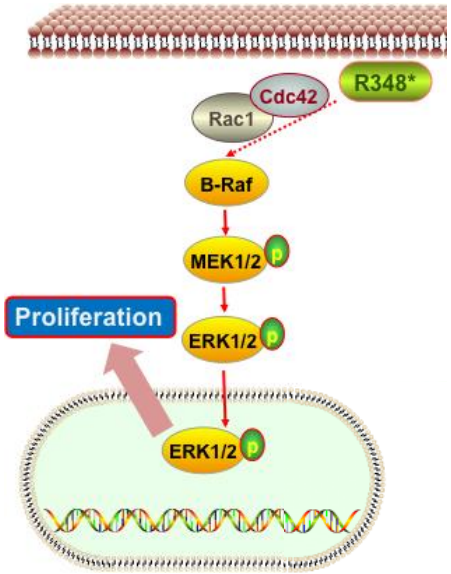
PIK3R1 R348* is a Neomorphic Mutation Causes Addiction to MEK-ERK Signaling

p-ERK



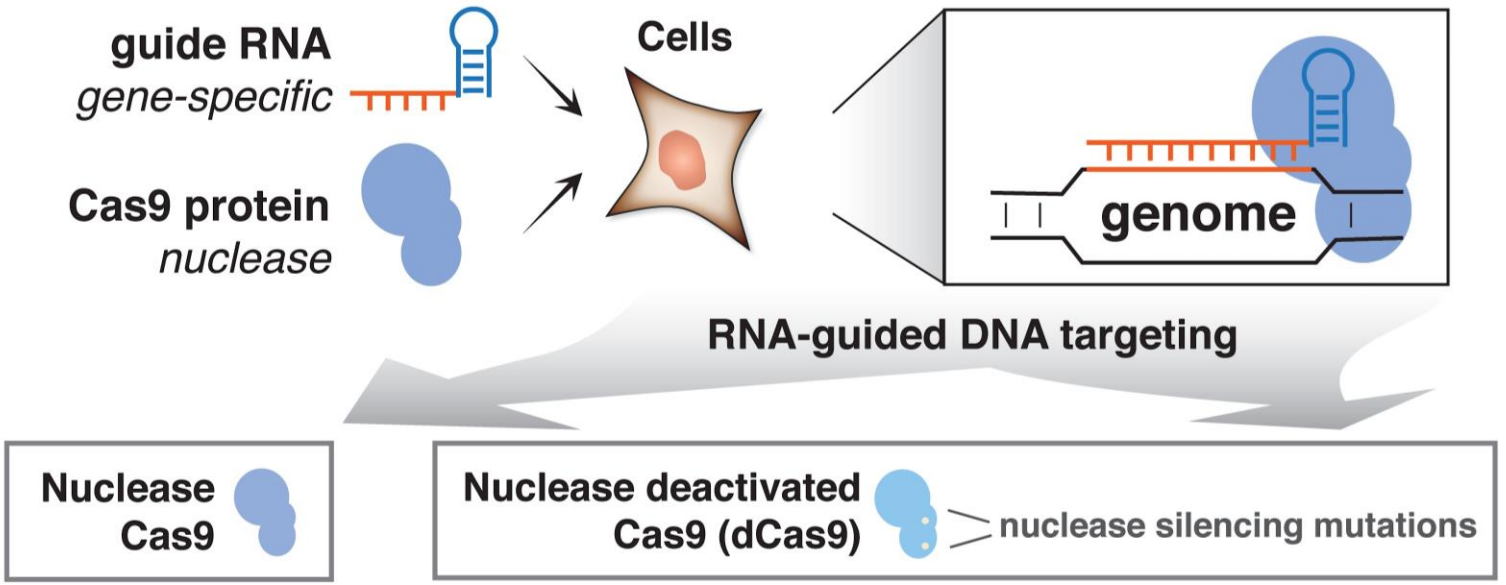
Response to MEK inhibitors



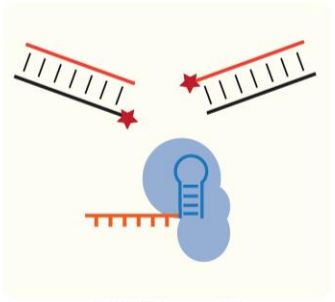


Wai Ting Cheung, Jane LI, Russell Broaddus, Gordon Mills

CRISPRs: Gene Regulation by Excision



Genome editing

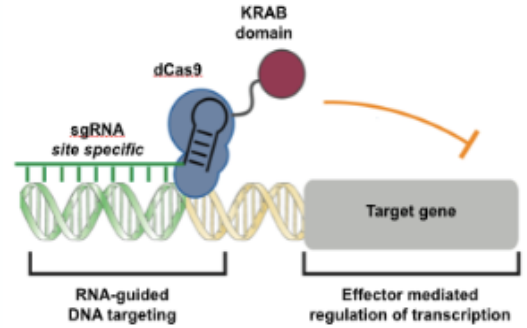


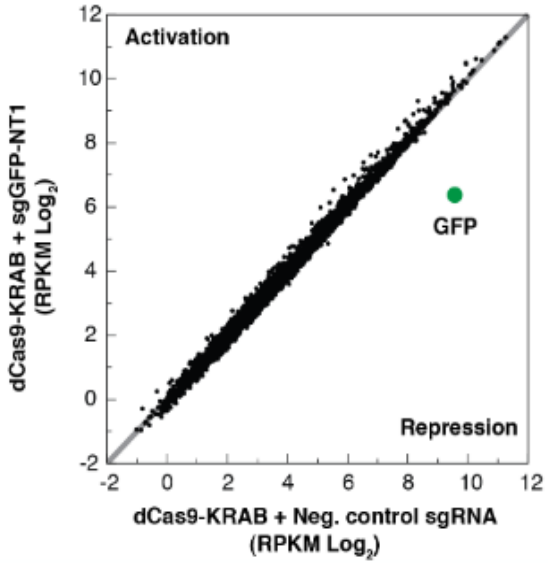
DNA cutter

March 10, 2016, Nature cover

Retooling CRISPRs: Turn Genes On or Off

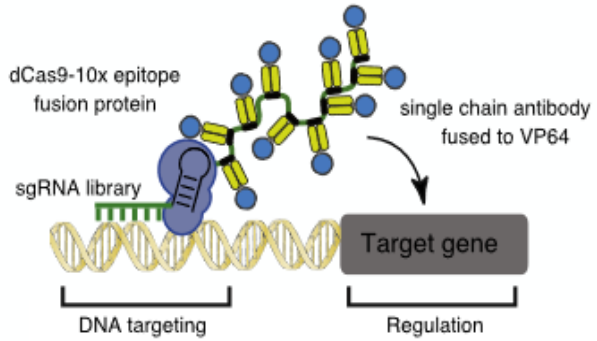
CRISPRi

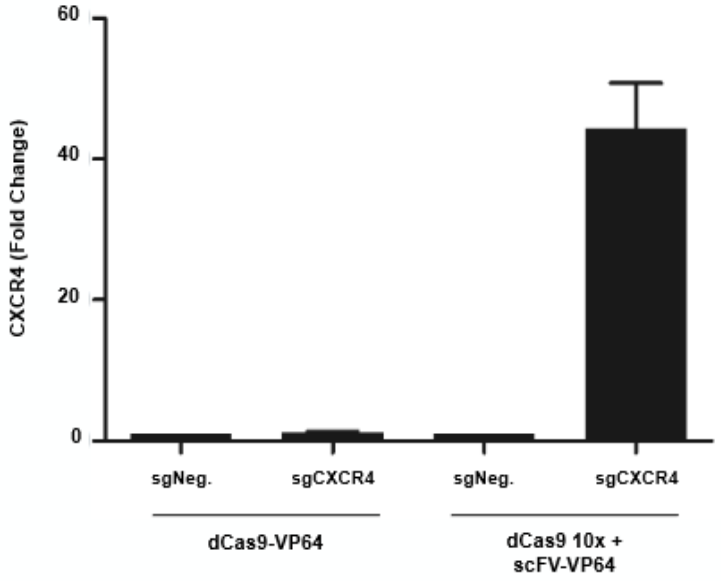




Gilbert et al., *Cell* 2013

CRISPRa



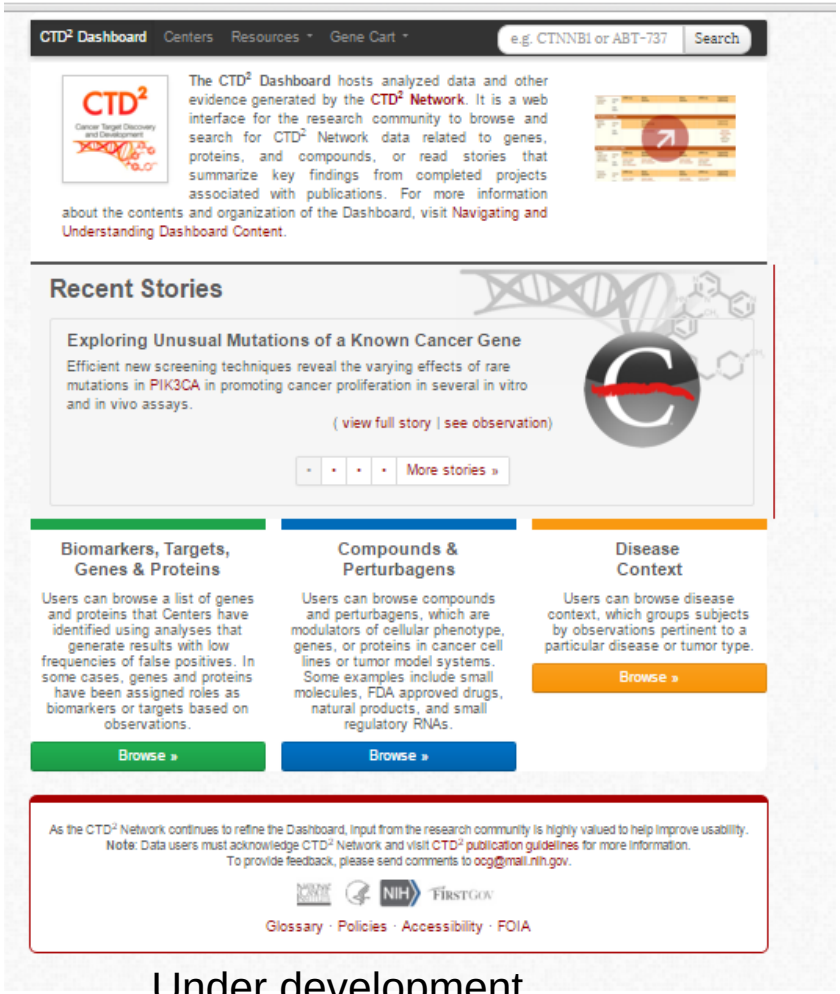


Tanenbaum et al., *Cell* 2014

CTD² Network: Shares Data

https://ctd2.nci.nih.gov/dataPortal/

http://ctd2-dashboard.nci.nih.gov/



Under development

CTD² Intra-Network Data Portal

- Makes collaborations easier
- Allows rapid graduation of data to Public Data Portal

CTD² INTRA-NETWORK DATA PORTAL

Last Updated: March 10, 2016

For target annotated CTD² Informer Set and other resources please see HERE

Project Title	Experimental Approaches	Data	Principal Investigator	Contact Name
Cold Spring Harbor Laboratory (CSHL) : Computational and Functional Approaches to Validate Cancer Genome Targets				
ORF and shRNA Screening for Driver Genes and Tumor Cell Dependencies	• ORF and shRNA screening • mouse model technology	Raw/Analyzed Data (DCC)	Scott Powers, Ph.D.	Jinyu Li
Columbia University : Systems Biology of Tumor Progression and Drug Resistance				
Master Regulator Genes of the FL Transformation to DLBCL	• transcriptional-level signal transduction network reverse engineering (ARACNe) • master regulator analysis (MARINa)	*coming soon*	Andrea Califano, Ph.D.	Kenneth Smith
Analysis of Regulatory Networks Determining the Mesenchymal Subtype of Glioblastoma (GBM)	• master regulator analysis (MARINa) • functional copy number variation analysis • modulator (MINDy) analysis	*coming soon*		
Master Regulators of Ovarian Cancer	• master regulator analysis (MARINa)	*coming soon*		
Master Regulators of Neuroblastoma Subtypes	• master regulator analysis (MARINa)	*coming soon*		
Emory University : High-Throughput Protein-Protein Interaction Interrogation in Cancer				
High-Throughput Protein-Protein Interaction Dataset for Lung Cancer-Associated Genes	• high-throughput orthogonal protein-protein interaction screening	Raw/Analyzed Data (DCC) Dashboard Submission(s)	Haian Fu, Ph.D.	Andrei Ivanov
mRNA-seq Analysis of FFPE Prostate Cancer Tissues Long, Xu et al. (Cancer Res)	• expression profiling by mRNA-seq	Raw/Analyzed Data (GEO,SRA)		

Goals of Proposed Renewal

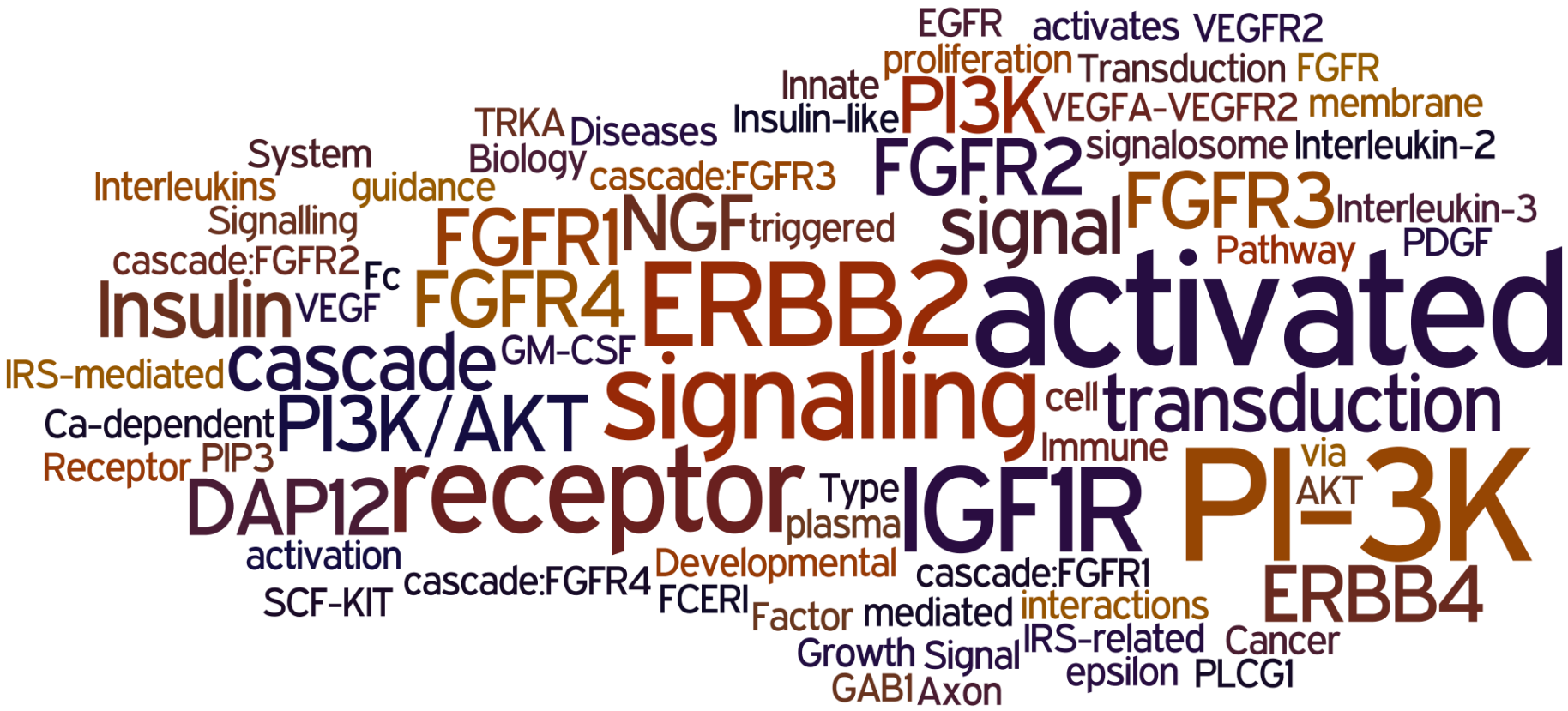
- New genomic data will be integrated into the bioinformatic component, which currently includes TCGA, TARGET, CGCI, such as NCI supported clinical trials and studies (Alchemist, Match (adult and pediatric), Exceptional Responders, etc.)
  - Nimble, flexible and open to new opportunities

- Accelerate the translation of patient genomic data into clinical application
  - Innovate the integration of computational mining large scale genomic data analyses and N of 1 application
  - Identify and confirm novel therapeutic target candidates
  - Identify and confirm novel modulators within specific cancer context (cellular or mutational) in vitro (cell lines) or in vivo (novel cancer models, e.g. organoids or conditionally reprogramed cells, others)
    - Pharmacogenomic screens to understand mechanism of action
    - CRISPRs, RNAi
  - Continue to share models, methods, data and resources with the scientific community through web site(s) and distributors

Examples of Resources to be Developed & Scientific Questions to be Addressed

- Collaborative activities to maximize useful shared data
  - Use of the 320 compound "CTD^2 informer set" in screening campaigns, analyze and make data available through the web sites and publications
  - Develop methods to interpret results from different types of experimental read-outs
  - Modify informer set to include compounds which target cellular metabolites, enzymes

Top 50 Pathways Targeted by the CTD^2 Informer Set



Jeff Kiefer, TGEN

Examples of Resources to be Developed & Scientific Questions to be Addressed

- Collaborative activities to maximize useful shared data - continued
  - Results from all CRISPR screens will be assembled in the CTD^2 Data Portal
  - Use of “control” screening models/cell lines to improve interpretation
  - Compare the various CAS9 enzyme constructs and the impact on interpretation of the results
- Use new cancer models for high-throughput functional studies to define biologically relevant targets, modulators or biomarkers
  - Development of methodologies for the community
  - Determine if the growth conditions impact on interpretation of results of CRISPR, small molecule, RNAi and cDNA screens
  - Share results which can serve as bases for preclinical testing or the next phase
- Define pathway and gene redundancies
  - Identify approaches to overcome them
- High-throughput combination screening of small molecules and CRISPRs
  - Identify genes which overcome resistance to precision treatments
- Other

Questions from the BSA Sub-committee-1

- What evaluation criteria will be used to ascertain the quality of the data? And
- Are established metrics in place to evaluate high throughput data?
  - Development of the "Tier" concepts and publishing the document for others to use if they are so inclined
    - PI subgroup (rotating appointment) reviews Dashboard submissions
  - D-HIP group reviews submissions to ensure uniformity of metadata
    - All metadata is registered at NCI's caDSR (dictionary with definitions)
  - Codification of QC metrics for the various high-throughput screens; each one will have their own technical issues to consider—in progress. Will be shared through the Data Portal for each project—be transparent. Examples:
    - Use of multiple RNAi(s) and cell lines in a screening campaign
      - Development of software to identify seed sequences which would lead to off-target effects
    - Replicate reproducibility and either remove failures or repeat a screen
    - Remove data with "low/high signal" depending on the assay
    - Small molecule HTS use dilution series; use of a compound with known function

Questions from the BSA Sub-committee-2

- How are current networks organized? And
- What is your vision of the types of networks that will be established?
  - Centers deposit “raw” prepublication data into intra-Network Portal
    - Use of “uniform controls” will allow cross-Center analyses
  - Centers collaborate to accelerate their research
  - Continue monthly teleconferences in which pre-publication data are discussed

- Going forward, how do you anticipate using TCGA data?
  - Expect that all NCI (and other) large scale genomic data will be used by the next set of Centers
  - NCI will launch Genomic Data Commons June 1, 2016 making high-content genomic data easier to access
    - Includes clinical, sample, molecular data

Mechanism and Cost

- Mechanism:
  - U01 Cooperative Agreement Grants
    - Critical for pre-competitive collaborations
    - Essential for communication
    - Important for governance
- Open competition
  - No presumption of current Centers
  - Establish the best network possible from proposed grants
- Budget: $12M for year 1
  - Fund up to 12 centers

Program Evaluation Criteria: Examples

The program will be evaluated by a number of parameters:

- The number of publication as well as journal's H index
  - How many times the manuscripts were cited
- Are the results, methods, tools, etc. developed by scientists used in academia and industry?
  - Frequency of data portal visits and data downloads
  - How do the results impact on the number of proposals received at the NIH following up of CTD^2 findings
- Number of validated probes and/or targets
  - Were the results of the projects transitioned into preclinical testing?
  - Inclusion of CTD^2 results as a basis of an early phase clinical trial
- Other appropriate specific evaluation parameters will be determined once the composition of network is known

NATIONAL CANCER INSTITUTE

www.cancer.gov

www.cancer.gov/espanol

CTD² Web Home Page

PROGRAMS / RESOURCES / NEWS & PUBLICATIONS / ABOUT OCG

ACCESS PROGRAM DATASETS

Email | Print

Home › Programs › CTD²

How to access multiple datasets →

CTD²

- Overview
- Research
- Publications
- Analytical Tools
- Resources
- Centers
- Using CTD² Data
- CTD² Publication Guidelines

CTD²: Cancer Target Discovery And Development

CTD² bridges the gap between the enormous volumes of data generated by genomic characterization studies and the ability to use these data for the development of human cancer therapeutics. It specializes in computational and functional genomics approaches critical for translating next-generation sequencing data, as well as high-throughput and high content small molecule and genetic screens.

VIEW CTD² DATA PORTAL

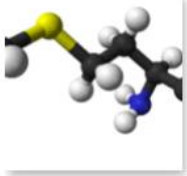
VIEW CTD² DASHBOARD

The CTD² Data Portal enables the cancer research community to search and download data generated by the initiative. The Dashboard compiles experimental observations across Network findings.

View Using CTD² Data Page

NEWS & PUBLICATIONS

View In the News Page



March 11, 2016

MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells

The discovery of cancer dependencies has the potential to inform therapeutic strategies and to identify putative drug targets. Integrating data from comprehensive genomic profiling of cancer cell lines and from functional characterization of cancer cell dependencies, we discovered that loss of...

CTD² RESOURCES

- The Cancer Genome Atlas (TCGA) Data Portal
- TARGET Data Matrix
- Online Bioinformatics Tutorials
- National Cancer Institute
- What is Cancer?

CTD² Web Home Page

PROGRAMS / RESOURCES / NEWS & PUBLICATIONS / ABOUT OCG

ACCESS PROGRAM DATASETS

Email | Print

Home › Programs › CTD²

How to access multiple datasets →

CTD²

- Overview
- Research
- Publications
- Analytical Tools
- Resources
- Centers
- Using CTD² Data
- CTD² Publication Guidelines

CTD²: Cancer Target Discovery And Development

CTD² bridges the gap between the enormous volumes of data generated by genomic characterization studies and the ability to use these data for the development of human cancer therapeutics. It specializes in computational and functional genomics approaches critical for translating next-generation sequencing data, as well as high-throughput and high content small molecule and genetic screens.

VIEW CTD² DATA PORTAL

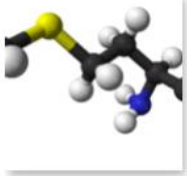
VIEW CTD² DASHBOARD

The CTD² Data Portal enables the cancer research community to search and download data generated by the initiative. The Dashboard compiles experimental observations across Network findings.

View Using CTD² Data Page

NEWS & PUBLICATIONS

View In the News Page



March 11, 2016

MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells

The discovery of cancer dependencies has the potential to inform therapeutic strategies and to identify putative drug targets. Integrating data from comprehensive genomic profiling of cancer cell lines and from functional characterization of cancer cell dependencies, we discovered that loss of...

CTD² RESOURCES

- The Cancer Genome Atlas (TCGA) Data Portal
- TARGET Data Matrix
- Online Bioinformatics Tutorials
- National Cancer Institute
- What is Cancer?

CTD² Cancer Target Discovery and Development

CTD² Web Home Page

PROGRAMS / RESOURCES / NEWS & PUBLICATIONS / ABOUT OCG

ACCESS PROGRAM DATASETS

Email | Print

Home › Programs › CTD²

How to access multiple datasets →

CTD²

- Overview
- Research
- Publications
- Analytical Tools
- Resources
- Centers
- Using CTD² Data
- CTD² Publication Guidelines

CTD²: Cancer Target Discovery And Development

CTD² bridges the gap between the enormous volumes of data generated by genomic characterization studies and the ability to use these data for the development of human cancer therapeutics. It specializes in computational and functional genomics approaches critical for translating next-generation sequencing data, as well as high-throughput and high content small molecule and genetic screens.

VIEW CTD² DATA PORTAL

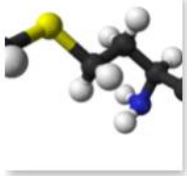
VIEW CTD² DASHBOARD

The CTD² Data Portal enables the cancer research community to search and download data generated by the initiative. The Dashboard compiles experimental observations across Network findings.

View Using CTD² Data Page

NEWS & PUBLICATIONS

View In the News Page



March 11, 2016

MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells

The discovery of cancer dependencies has the potential to inform therapeutic strategies and to identify putative drug targets. Integrating data from comprehensive genomic profiling of cancer cell lines and from functional characterization of cancer cell dependencies, we discovered that loss of...

CTD² RESOURCES

- The Cancer Genome Atlas (TCGA) Data Portal
- TARGET Data Matrix
- Online Bioinformatics Tutorials
- National Cancer Institute
- What is Cancer?

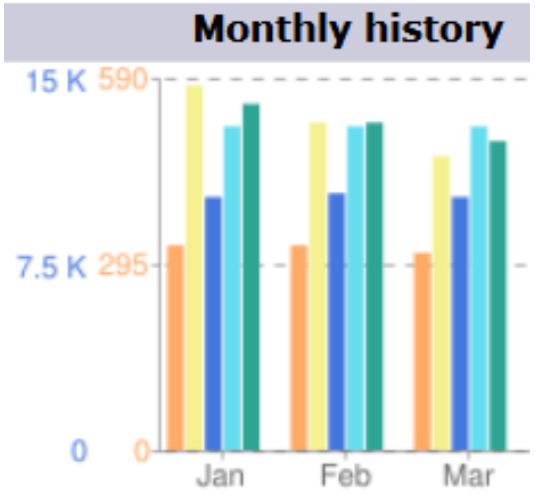
CTD² Cancer Target Discovery and Development

CTD^2 Network Funded Tools

Name	Function	Usages	Downloads
ATARiS: Analytic Technique for Assessment of RNAi by Similarity	*Reduces off-target effects from data in phenotypic screens using multiple RNAi reagents*	*Avg 18 users/month*	
CTRP: Cancer Therapeutics Response Portal	*Connects cellular features to small molecule sensitivities for >800 cell lines*	*Avg 600 users/month*	
DecoRNAi: Deconvolution Analysis of RNAi Screening data	*Quantitates and annotates off-target effects of primary RNAi screening datasets*	*No data available*	
DeMAND: Detecting Mechanism of Action based Network Dysregulation	*Uses gene expression profiles of treated and untreated cell lines to determine mechanism of action of small molecules*		*326*
DIGGIT: Driver-gene Inference by Genetical-genomics and Information Theory	*Integrates patient-matched genomic mutation and gene expression data with corresponding gene regulatory networks to identify candidate driver mutations*		*1190*
EDDY: Evaluation of Differential DependencY	*Estimates differential dependencies for a set of genes between two conditions*	*No data available*	
FuSiOn: Functional Signature Ontology	*Searchable ontology map built from gene expression data from human kinome screens*	*No data available*	
MethylMix	*Identifies differentially and transcriptionally predictive methylated genes within a disease*	*Avg 99 users/month*	*2125*
MINDy2/ CINDy: Modulator Inference by Network Dynamics/ Conditional Inference of Network Dynamics	*Assess the effect of candidate proteins on a transcription factor of interest*		*129*
PiHelper	*Integrates drug target and antibody target interactions from publicly available resources to facilitate research in systems pharmacology, perturbation biology, and proteomics*	*No data available*	
PARIS: Probability Analysis by Ranked Information Score	*Uses a mutual information-based metric to rank data, such as shRNA/gene dependencies in cell lines, genomic features, and chemical sensitivities*	*Avg 87 users/month*	
Project Achilles Portal	*Uses genome-wide pooled shRNA screens to identify and catalog genetic vulnerabilities associated with genetic or epigenetic changes across hundreds of cancer cell lines*	*2000 users/month*	
The Cancer Genome Atlas Clinical Explorer	*A platform to query TCGA data to identify clinical-genomic associations*	*Avg 500 users/month*	
VIPER/ MARINa: Virtual Inference of Protein-activity by Enriched Regulon analysis/Master Regulator Inference algorithm	*MARINa uses the transcriptional targets of each transcription factor as a multiplexed reporter assay to infer the transcription factors controlling the transition between related cellular states.* *VIPER extends MARINa to single samples and any regulatory protein*		*3359*

CTD^2 Network Data Portal: Metrics

Do we have sufficient volume of traffic, and more importantly qualified traffic that is actually using our data sets?



Orange bar and numbers on left Y axis- unique visitors
Yellow bar-number of visits
Blue bar and number on left Y axis- pages accessed
Cyan blue bar- number of hits
Green bar and numbers on right Y axis- bandwidth

Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2016	334	589	10,909	13,999	152.31 MB
Feb 2016	332	532	11,041	14,007	142.91 MB
Mar 2016	321	475	10,953	13,989	135.90 MB

NCI is working on to allow tracking actual downloads—in progress