Translational Bioinformatics: Connecting Genes with Drugs

Aik Choon Tan, Ph.D.
Associate Professor of Bioinformatics
Division of Medical Oncology
Department of Medicine
aikchoon.tan@ucdenver.edu

11/13/2018

http://tanlab.ucdenver.edu/labHomePage/teaching/CANB7640

Outline

- Motivation
- Gene-Drug Databases
  - Knowledge-bases: PharmGKB, DrugBank
  - Drugs/Compounds resources: : NCI-60, CCLE, GDSC, CTRP
- Rational Combination of Drugs
  - Chemical Screening
  - Functional Genetic Screening
    - Loss-of-Function (LOF) Screens (Synthetic lethality and Essential Gene)
    - Gain-of-Function (GOF) Screens
- Connectivity Map
  - Concept
  - Applications
  - LINCS (Library of Integrated Network Cell Signaling)
  - K-Map

R&D Model for the Discovery and Development of A New Oncology Drug

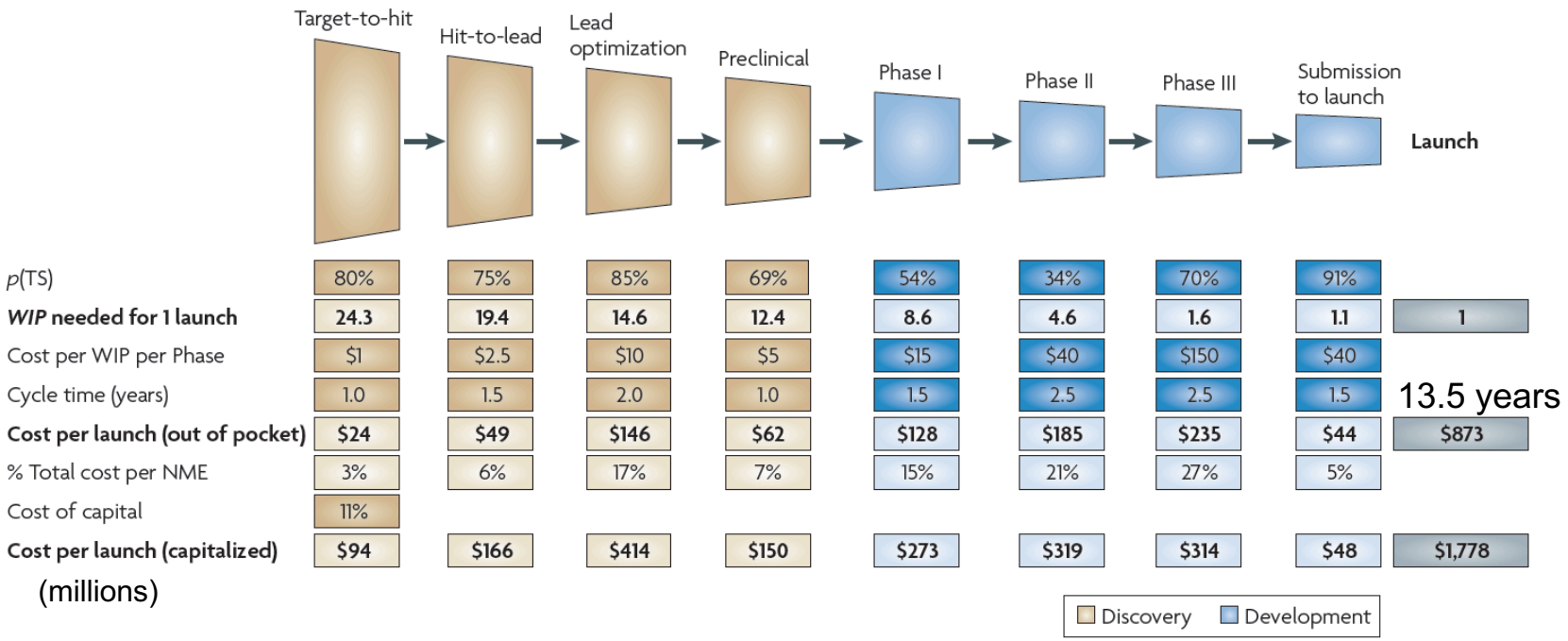
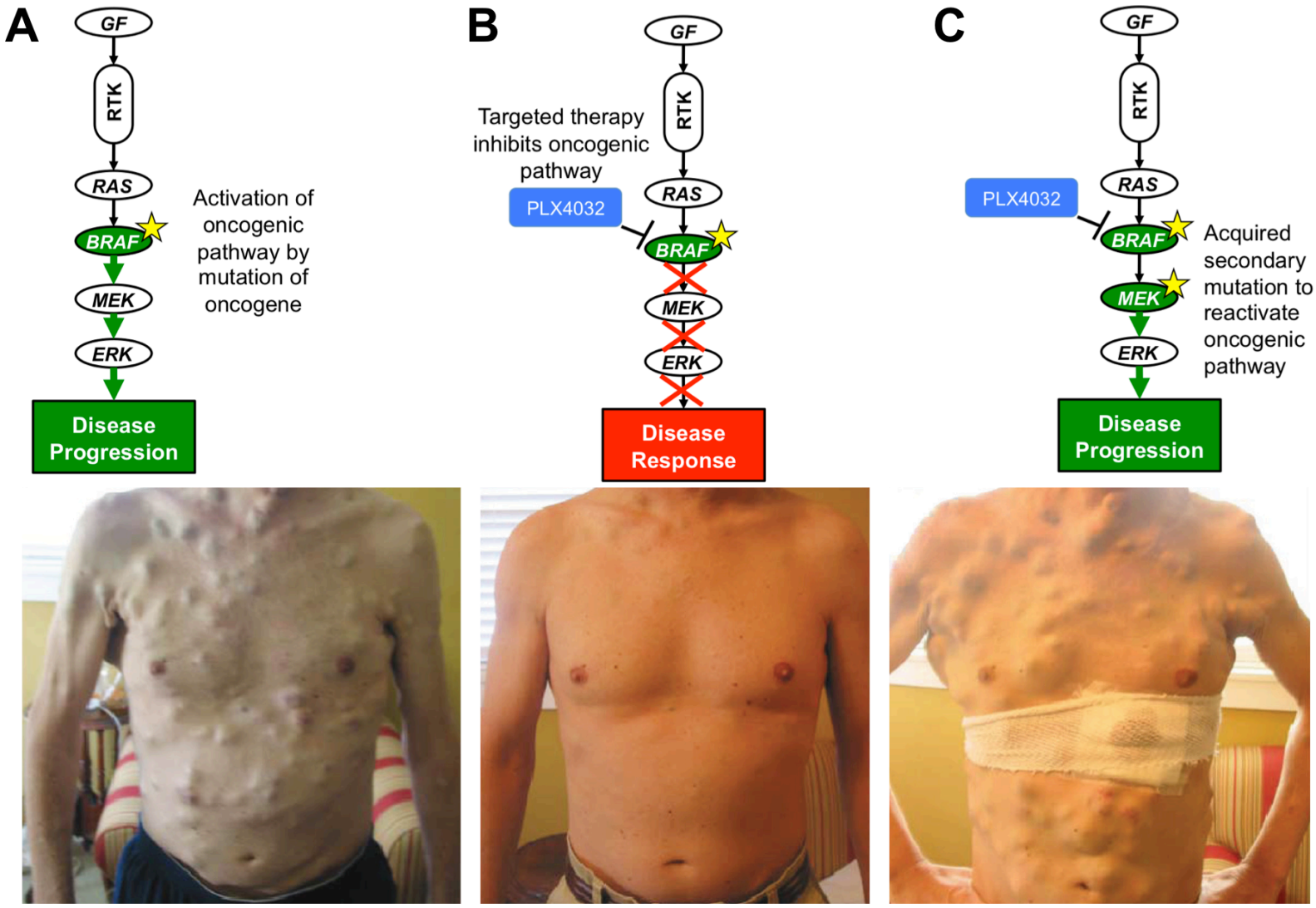


Figure 2 | **R&D model yielding costs to successfully discover and develop a single new molecular entity.** The model defines the distinct phases of drug discovery and development from the initial stage of target-to-hit to the final stage, launch. The model is based on a set of industry-appropriate R&D assumptions (industry benchmarks and data from Eli Lilly and Company) defining the performance of the R&D process at each stage of development (see Supplementary information S2 (box) for details). R&D parameters include: the probability of successful transition from one stage to the next (*p*(TS)), the phase cost for each project, the cycle time required to progress through each stage of development and the cost of capital, reflecting the returns required by shareholders to use their money during the lengthy R&D process. With these inputs (darker shaded boxes), the model calculates the number of assets (work in process, *WIP*) needed in each stage of development to achieve one new molecular entity (NME) launch. Based on the assumptions for success rate, cycle time and cost, the model further calculates the 'out of pocket' cost per phase as well as the total cost to achieve one NME launch per year (US$873 million). Lighter shaded boxes show calculated values based on assumed inputs. Capitalizing the cost, to account for the cost of capital during this period of over 13 years, yields a 'capitalized' cost of $1,778 million per NME launch. It is important to note that this model does not include investments for exploratory discovery research, post-launch expenses or overheads (that is, salaries for employees not engaged in R&D activities but necessary to support the organization).

(Paul et al, 2010. Nature Rev Drug Discovery 9:203-214)

Targeting mutated genes with Targeted Therapy



Pre-treatment
(38 yo melanoma patient with BRAF mut)

Post-treatment
15 weeks on PLX4032 (BRAFi)

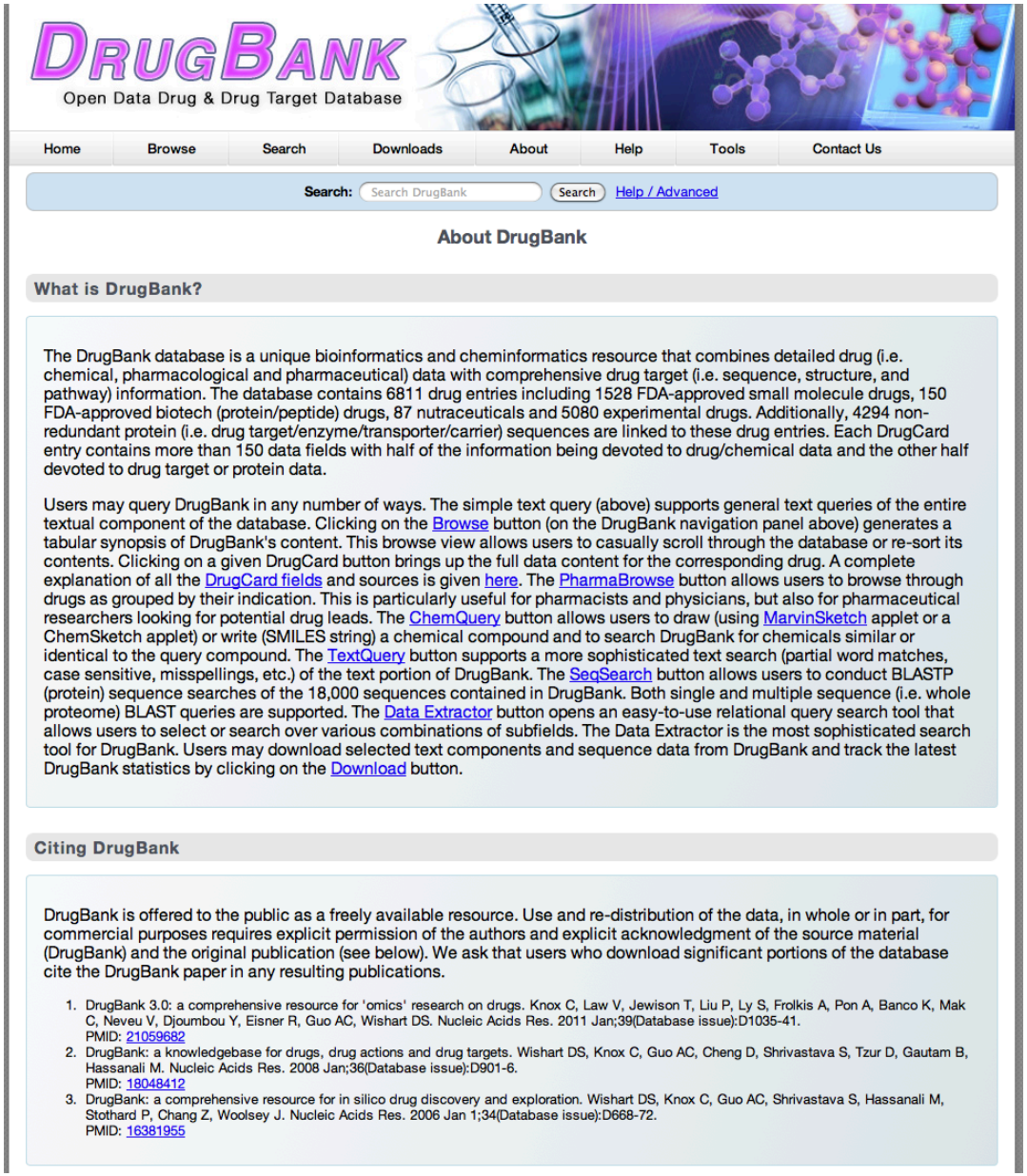
Relapse (acquired resistance)
23 weeks on PLX4032 (BRAFi)

(Patient Image From: *Wagle et al JCO 2011*)

Knowledge-Bases
Drug-Gene Resources

- DrugBank
- PharmGKB
- OncoKB
- Other (Not covered):
  - KEGG-DRUG
  - STITCH
  - PubChem
  - ChEMBL

DrugBank http://www.drugbank.ca/



DRUGBANK
Open Data Drug & Drug Target Database

Home | Browse | Search | Downloads | About | Help | Tools | Contact Us

Search: Search DrugBank Search Help / Advanced

About DrugBank

What is DrugBank?

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. The database contains 6811 drug entries including 1528 FDA-approved small molecule drugs, 150 FDA-approved biotech (protein/peptide) drugs, 87 nutraceuticals and 5080 experimental drugs. Additionally, 4294 non-redundant protein (i.e. drug target/enzyme/transporter/carrier) sequences are linked to these drug entries. Each DrugCard entry contains more than 150 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

Users may query DrugBank in any number of ways. The simple text query (above) supports general text queries of the entire textual component of the database. Clicking on the Browse button (on the DrugBank navigation panel above) generates a tabular synopsis of DrugBank's content. This browse view allows users to casually scroll through the database or re-sort its contents. Clicking on a given DrugCard button brings up the full data content for the corresponding drug. A complete explanation of all the DrugCard fields and sources is given here. The PharmaBrowse button allows users to browse through drugs as grouped by their indication. This is particularly useful for pharmacists and physicians, but also for pharmaceutical researchers looking for potential drug leads. The ChemQuery button allows users to draw (using MarvinSketch applet or a ChemSketch applet) or write (SMILES string) a chemical compound and to search DrugBank for chemicals similar or identical to the query compound. The TextQuery button supports a more sophisticated text search (partial word matches, case sensitive, misspellings, etc.) of the text portion of DrugBank. The SeqSearch button allows users to conduct BLASTP (protein) sequence searches of the 18,000 sequences contained in DrugBank. Both single and multiple sequence (i.e. whole proteome) BLAST queries are supported. The Data Extractor button opens an easy-to-use relational query search tool that allows users to select or search over various combinations of subfields. The Data Extractor is the most sophisticated search tool for DrugBank. Users may download selected text components and sequence data from DrugBank and track the latest DrugBank statistics by clicking on the Download button.

Citing DrugBank

DrugBank is offered to the public as a freely available resource. Use and re-distribution of the data, in whole or in part, for commercial purposes requires explicit permission of the authors and explicit acknowledgment of the source material (DrugBank) and the original publication (see below). We ask that users who download significant portions of the database cite the DrugBank paper in any resulting publications.

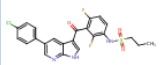
1. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, Djoumbou Y, Eisner R, Guo AC, Wishart DS. Nucleic Acids Res. 2011 Jan;39(Database issue):D1035-41.
   PMID: 21059682
2. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B, Hassanali M. Nucleic Acids Res. 2008 Jan;36(Database issue):D901-6.
   PMID: 18048412
3. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D668-72.
   PMID: 16381955

DrugBank http://www.drugbank.ca/

Identification | Taxonomy | Pharmacology | Pharmacoeconomics | Properties | References | Interactions

targets (1) | enzymes (3) | transporters (2) | carriers (2) | Show Drugs with Similar Structures

Identification

Name	**Vemurafenib**
Accession Number	**DB08881**
Type	small molecule
Groups	approved
Description	Vemurafenib is a BRAF enzyme inhibitor developed by Plexxikon and Genentech for the treatme melanoma. [Wikipedia] The cobas® 4800 BRAF B600 mutation test provided by Roche Molecul diagnostic test to confirm eligibility for treatment. FDA approved on August 17, 2011 under the c Roche.
Structure	 Download: MOL | SDF | SMILES | InChI Display: 2D Structure | 3D Structure
Synonyms	BRAF(V600E) Kinase Inhibitor RO5185426 PLX4032 RG7204 RO5185426
Salts	Not Available
Brand names	Name: Zelboraf — Company: Hoffman La Roche
Brand mixtures	Not Available
Categories	• Antineoplastic Agents
CAS number	918504-65-1
Weight	Average: 489.922 Monoisotopic: 489.072546264
Chemical Formula	$C_{23}H_{18}ClF_2N_3O_3S$
InChI Key	InChIKey=GPXBXXGIAQBQNI-UHFFFAOYNA-N
InChI	InChI=1/C23H18ClF2N3O3S/c1-2-9-33(31,32)29-19-8-7-18(25)20(21(19)26)22(30)17-12-28-23-23)13-3-5-15(24)6-4-13/h3-8,10-12,29H,2,9H2,1H3,(H,27,28)

Targets

1. B-Raf proto-oncogene serine/threonine-protein kinase

Pharmacological action: **yes**
Actions: **inhibitor**

Involved in the transduction of mitogenic signals from the cell membrane to the nucleus. May play a role in the postsynaptic responses of hippocampal neuron

Organism class: **human**
UniProt ID: P15056
Gene: BRAF
Protein Sequence: FASTA
Gene Sequence: FASTA
SNPs: SNPJam Report

References:

1. Jordan EJ, Kelly CM: Vemurafenib for the treatment of melanoma. Expert Opin Pharmacother. 2012 Dec;13(17):2533-43. doi: 10.1517/14656566.2012.737780. Epub 2012 Oct 24. Pubmed

Enzymes

1. Cytochrome P450 1A2

Actions: **inhibitor**

Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics. Most active in catalyzing 2-hydroxylation. Caffeine is metabolized primarily by cytochrome CYP1A2 in the liver through an initial N3-demethylation. Also acts in the metabolism of aflatoxin B1 and acetaminophen

UniProt ID: P05177
Gene: CYP1A2
Protein Sequence: FASTA
Gene Sequence: FASTA
SNPs: SNPJam Report

References:

1. Jordan EJ, Kelly CM: Vemurafenib for the treatment of melanoma. Expert Opin Pharmacother. 2012 Dec;13(17):2533-43. doi: 10.1517/14656566.2012.737780. Epub 2012 Oct 24. Pubmed

2. Cytochrome P450 2D6

Actions: **inhibitor**

Responsible for the metabolism of many drugs and environmental chemicals that it oxidizes. It is involved in the metabolism of drugs such as antiarrhythmics, adrenoceptor antagonists, and tricyclic antidepressants

UniProt ID: P10635
Gene: CYP2D6
Protein Sequence: FASTA
Gene Sequence: FASTA
SNPs: SNPJam Report

References:

1. Jordan EJ, Kelly CM: Vemurafenib for the treatment of melanoma. Expert Opin Pharmacother. 2012 Dec;13(17):2533-43. doi: 10.1517/14656566.2012.737780. Epub 2012 Oct 24. Pubmed

PharmGKB

Pharmacogenomics Knowledge for Personalized Medicine

M Whirl-Carrillo[1], EM McDonagh[1], JM Hebert[1], L Gong[1], K Sangkuhl[1], CF Thorn[1], RB Altman[1,2] and TE Klein[1]

Clinical Pharmacology & Therapeutics (2012) 92(4): 414-417.

http://www.pharmgkb.org

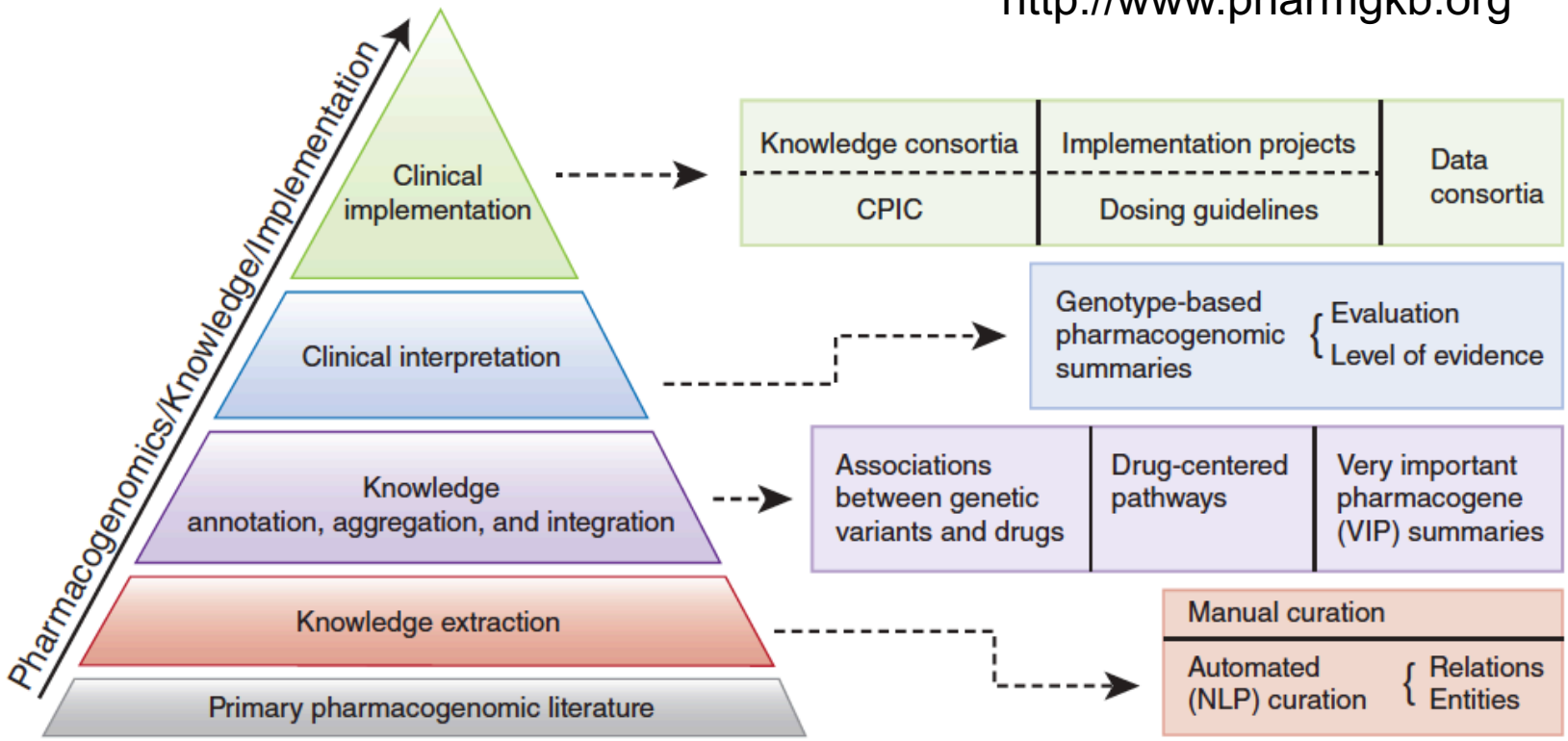


Figure 1 The PharmGKB Knowledge Pyramid. CPIC, Clinical Pharmacogenetics Implementation Consortium; NLP, natural-language processing. From PharmGKB with the permission of PharmGKB and Stanford University. Copyright PharmGKB.

PharmGKB

DRUG/SMALL MOLECULE:
vemurafenib

Clinical PGx | Overview | Properties | Pathways | Is Related To | Publications | Downloads/LinkOuts

Dosing Guidelines (0) | **Drug Labels (1)** | **Clinical Annotations (0)** | **Genetic Tests (0)**

PharmGKB gathers information regarding PGx on FDA drug labels from the FDA's "Table of Pharmacogenomic Biomarkers in Drug Labels", and from FDA-approved FDA and EMA-approved (European Medicines Agency) EMA labels brought to our attention. Excerpts from the label and downloadable highlighted label PDFs are manually curated by PharmGKB.

Please note that some drugs may have been removed from or added to the FDA's "Table of Pharmacogenomic Biomarkers in Drug Labels" without our knowledge. We periodically check the table for additions to this table and update PharmGKB accordingly.

There is currently no such list for European drug labels - we are working with the EMA to establish a list of European Public Assessment Reports (EPAR)s that contain PGx information. We are constructing this list by initially searching for drugs for which we have PGx-containing FDA drug labels - of these 44 EMA EPARs were identified and are being curated for pgx information.

We welcome any information regarding drug labels containing PGx information approved by the FDA, EMA or other Medicine Agencies around the world - please contact feedback.

view legend

FDA Label for vemurafenib and BRAF *last updated 10/25/2013*

This label is on the FDA Biomarker List

Genetic testing required

Summary

Vemurafenib is a kinase inhibitor used to treat patients with unresectable or metastatic melanoma ONLY in cases where the BRAF V600E mutation is found by an FDA-approved test. *In vitro* evidence points to BRAF wild-type cells proliferation with exposure to BRAF inhibitors. Therefore, testing is required.

Annotation

Excerpt from the vemurafenib (*Zelboraf*) drug label:

> ZELBORAF is indicated for the treatment of patients with unresectable or metastatic melanoma with BRAFV600E mutation as detected by an FDA-approved test.
>
> Limitation of Use: ZELBORAF is not recommended for use in patients with wild-type BRAF melanoma.
>
> Confirm the presence of BRAF V600E mutation in tumor specimens prior to initiation with ZELOBRAF.

For the complete drug label text with sections containing pharmacogenetic information highlighted, see the vemurafenib (*Zelboraf*) drug label PDF.

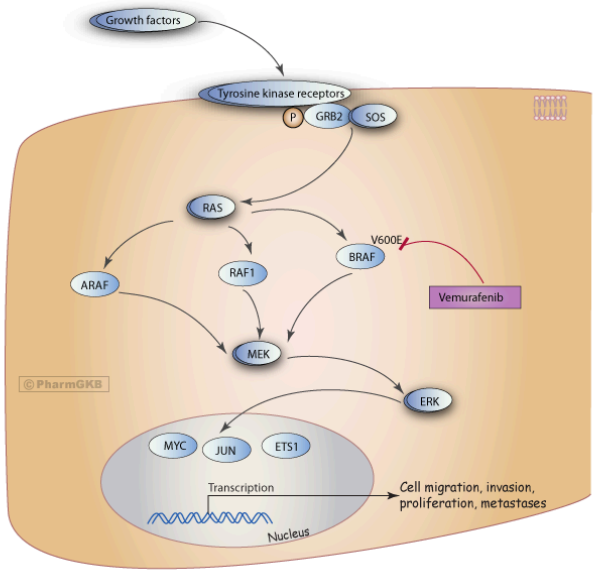
*Disclaimer: The contents of this page have not been endorsed by the FDA and are the sole responsibility of PharmGKB.

Full label available at DailyMed

Genes and/or phenotypes found in this label

- BRAF
  - Indications & usage section, Dosage & administration section, Warnings and precautions section, efficacy, toxicity
  - source: FDA Label
- CYP1A2
  - Drug interactions section, metabolism/PK
  - source: FDA Label
- CYP3A4
  - Drug interactions section, metabolism/PK
  - source: FDA Label

Overview | Components | Downloads/LinkOuts



- Legend
- Pathway diagram
- How to cite this pathway?
- All PharmGKB pathways

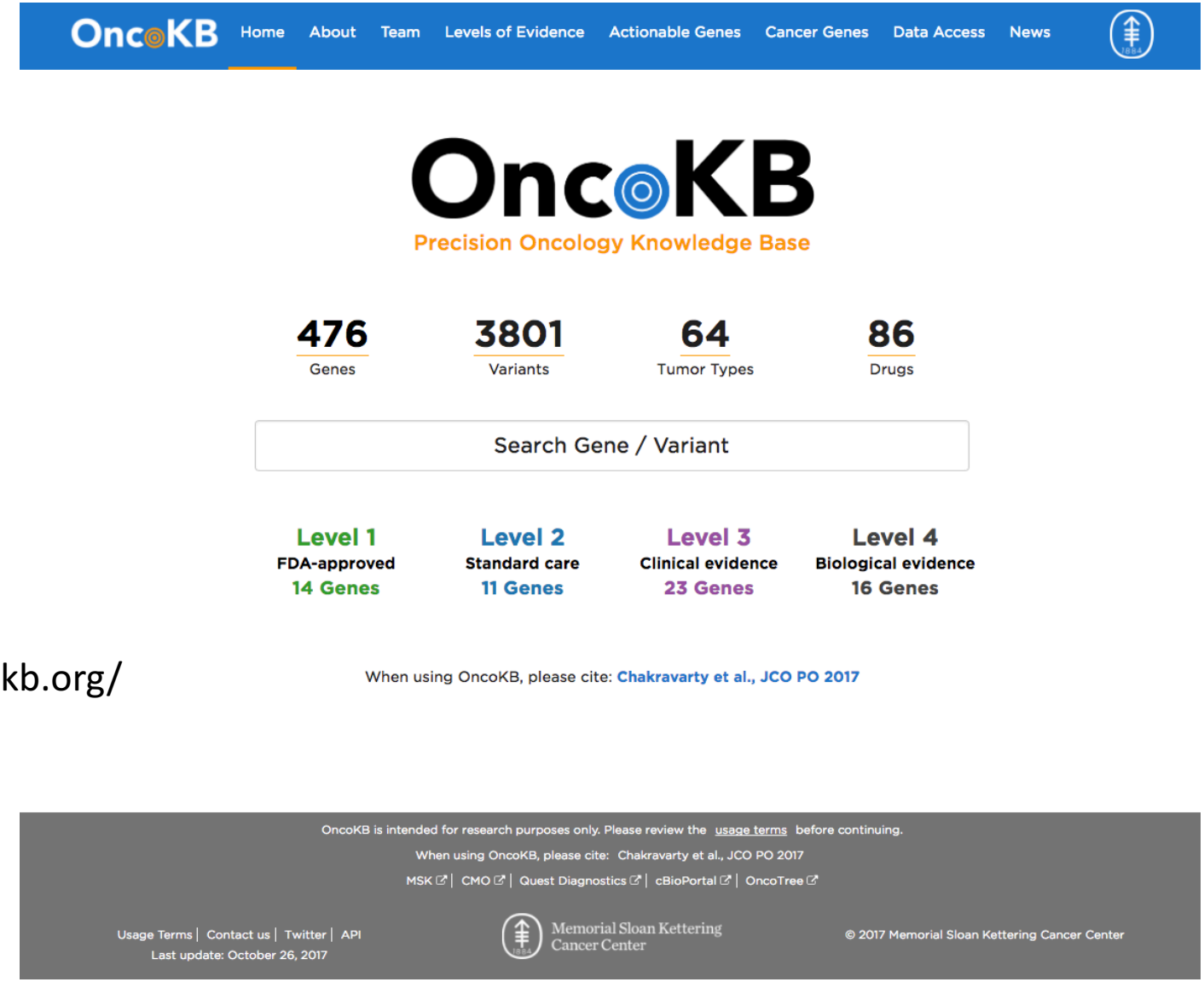
Click icons in pathway for more info: gene | drug | pathway

Description

One of the established signaling pathways in cancer development is the mitogen-activated protein kinase (MAPK) pathway. This pathway links extracellular stimuli, such as growth factors and hormone, to gene expression in the nucleus. Growth factor signals progress through the proteins RAS, BRAF, MEK and ERK, leading to cell proliferation. Mutations in the BRAF gene cause elevated kinase activity of the BRAF protein, which is seen in 8% of all solid tumors and in 50% of melanoma cases. The BRAF V600E mutation (rs113488022), which involves the substitution of valine (V) by glutamate (E) within the activation segment of the kinase domain , represents the vast majority of all BRAF mutations in cancer. Vemurafenib, approved by the FDA in August of 2011, is an oral BRAF inhibitor that specifically targets the V600E isoform [Articles:20818844, 22028422]. Clinical studies showed that vemurafenib improved rates of overall and progression-free survival in patients with previously untreated melanoma with the BRAF V600E mutation [Article:21639808].

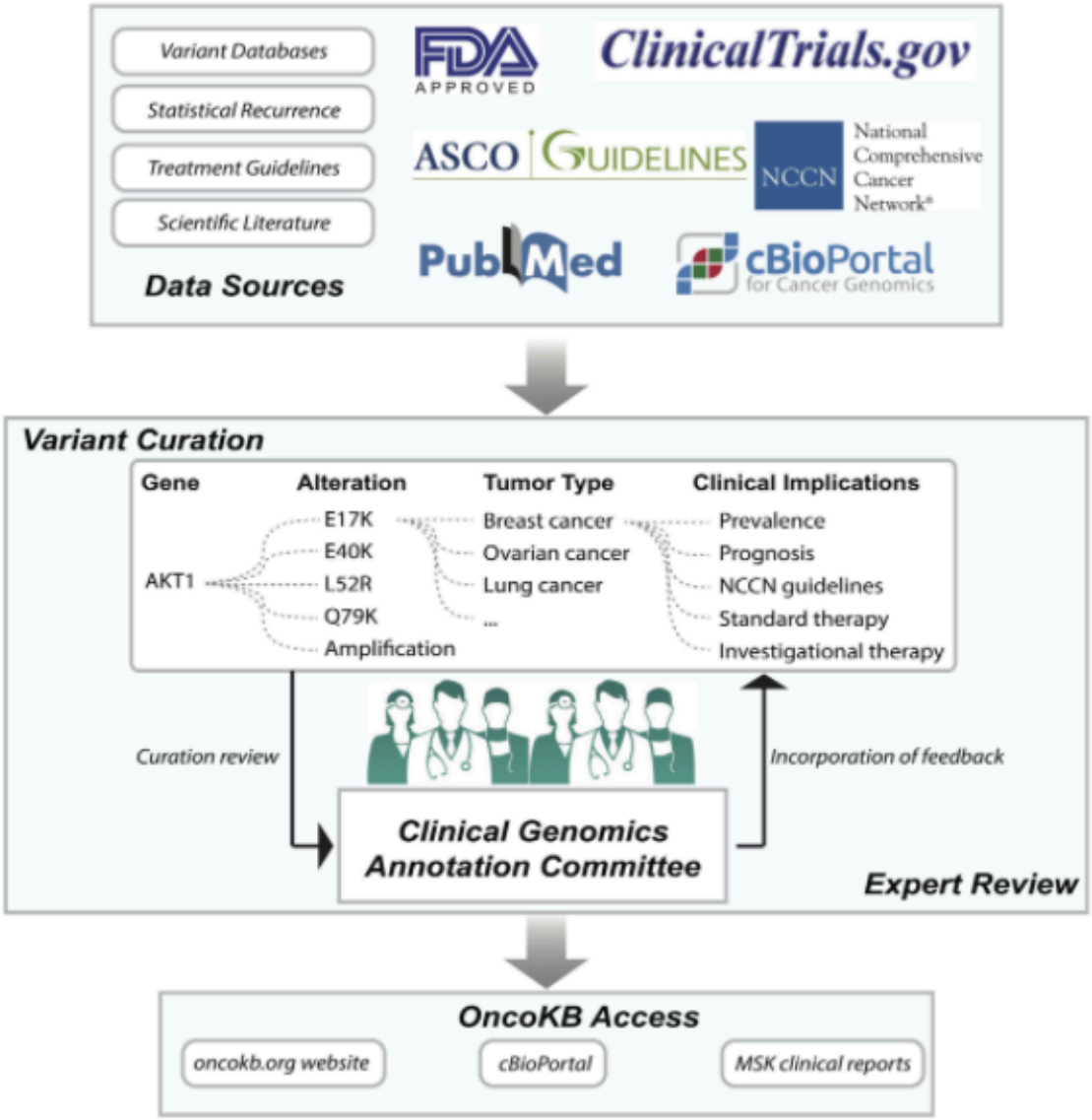
The RAF family is a set of serine/threonine kinases, which include BRAF, ARAF and CRAF (also called RAF1), that signals through phosphorylation of other downstream kinases. Upstream from RAF is RAS, the guanine nucleotide-binding protein activated by the binding of GTP. In the normal signaling pathway, an extracellular signal binds to a receptor tyrosine kinase on the cell surface. The types of receptors that signal through tyrosine kinases include those for fibroblast growth factors (FGFR), plate-derived growth factors (PDGFR), vascular epidermal growth factors (VEGFR), epidermal growth factors (EGFR), and nerve growth factors (NGFR), among others. When bound to the ligand, the receptor tyrosine kinase undergoes autophosphorylation of its tyrosine residues. This results in the phosphorylation of GRB2, an adapter protein that binds to the guanine nucleotide exchange factor, SOS. The GRB2/SOS complex stimulates inactive RAS by replacing bound GDP with GTP [Article:11294822]. Activated RAS causes BRAF to form a dimer with the other RAF isoforms ARAF and CRAF [Article:22083257]. Dimerization activates and phosphorylates MEK, which is a dual-specificity tyrosine and serine/threonine kinase. The main downstream target of MEK is ERK (also called MAPK). Upon phosphorylation by MEK, ERK can directly phosphorylate and activate a variety of transcription factors, such as Ets-1, c-Jun and c-Myc [Article:16854453].

OncoKB

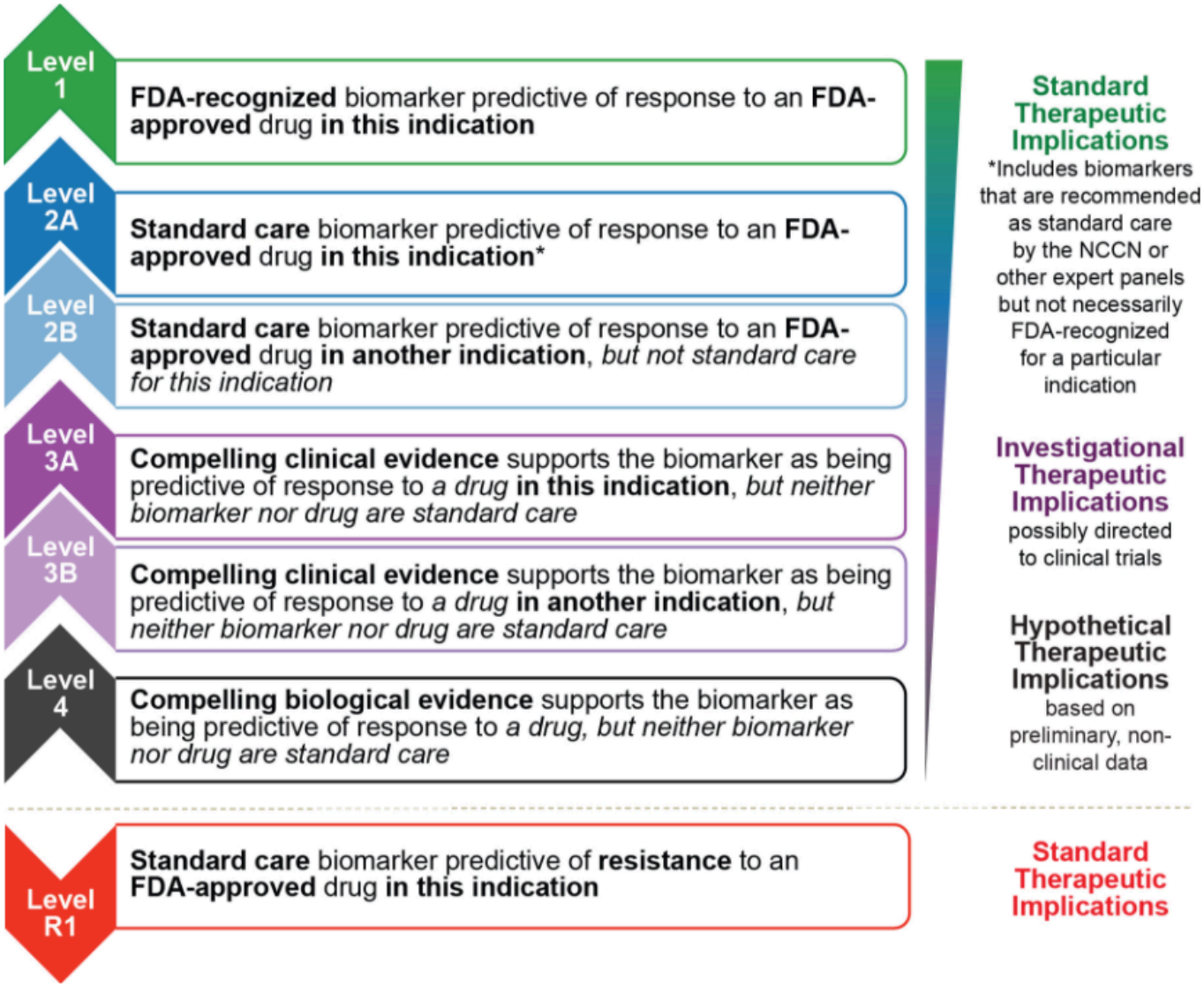


http://oncokb.org/

OncoKB


Variant Databases
Statistical Recurrence
Treatment Guidelines
Scientific Literature
Data Sources
FDA APPROVED
ClinicalTrials.gov
ASCO GUIDELINES
NCCN National Comprehensive Cancer Network
PubMed
cBioPortal for Cancer Genomics
Variant Curation
Gene
Alteration
Tumor Type
Clinical Implications
AKT1
E17K
E40K
L52R
Q79K
Amplification
Breast cancer
Ovarian cancer
Lung cancer
...
Prevalence
Prognosis
NCCN guidelines
Standard therapy
Investigational therapy
Curation review
Incorporation of feedback
Clinical Genomics Annotation Committee
Expert Review
OncoKB Access
oncokb.org website
cBioPortal
MSK clinical reports

OncoKB


Level 1
FDA-recognized biomarker predictive of response to an FDA-approved drug in this indication
Level 2A
Standard care biomarker predictive of response to an FDA-approved drug in this indication*
Level 2B
Standard care biomarker predictive of response to an FDA-approved drug in another indication, but not standard care for this indication
Level 3A
Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication, but neither biomarker nor drug are standard care
Level 3B
Compelling clinical evidence supports the biomarker as being predictive of response to a drug in another indication, but neither biomarker nor drug are standard care
Level 4
Compelling biological evidence supports the biomarker as being predictive of response to a drug, but neither biomarker nor drug are standard care
Level R1
Standard care biomarker predictive of resistance to an FDA-approved drug in this indication
Standard Therapeutic Implications
*Includes biomarkers that are recommended as standard care by the NCCN or other expert panels but not necessarily FDA-recognized for a particular indication
Investigational Therapeutic Implications
possibly directed to clinical trials
Hypothetical Therapeutic Implications
based on preliminary, non-clinical data
Standard Therapeutic Implications

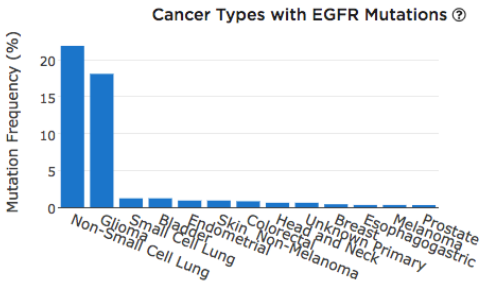
OncoKB

EGFR 127 annotated variants

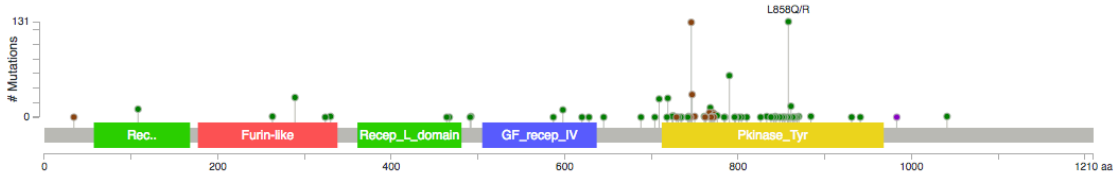
Oncogene
Highest level of evidence: Level 1
Also known as PIG61, ERBB1, mENA, ERBB, HER1, NISBD2
Isoform: ENST00000275493 RefSeq: NM_005228.3

EGFR, a receptor tyrosine kinase, is altered by amplification, mutation and/or overexpression in various cancers, most frequently in lung and brain cancers.

See EGFR background



Annotated Mutation Distribution in MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., Nature Medicine, 2017)



Clinically Relevant Variants (19) | All Annotated Variants (127)

If you notice any mistakes or missing variants / citations, please send an email to feedback@oncokb.org.

Search:

Variant	Cancer Type	Drug(s)	Level	Citations
Exon 19 deletion/insertion	Non-Small Cell Lung Cancer	Afatinib Erlotinib Gefitinib	1	12 references
Exon 19 deletion	Non-Small Cell Lung Cancer	Afatinib Erlotinib Gefitinib	1	12 references
Exon 19 insertion	Non-Small Cell Lung Cancer	Afatinib Erlotinib Gefitinib	1	1 reference
E709K	Non-Small Cell Lung Cancer	Afatinib Erlotinib Gefitinib	1	6 references
G719C	Non-Small Cell Lung Cancer	Afatinib Erlotinib Gefitinib	1	11 references

Perturbagen Drug Resources

- NCI-60
- Genomics of Drug Sensitivity in Cancer (GDSC / COSMIC)
- Cancer Cell Lines Encyclopedia (CCLE)
- Cancer Therapeutics Response Portal

Cancer Therapeutics Response Portal (CTRP)

Resource

Cell

An Interactive Resource to Identify Cancer Genetic and Lineage Dependencies Targeted by Small Molecules

Amrita Basu,[1,4] Nicole E. Bodycombe,[1,4] Jaime H. Cheah,[1,4] Edmund V. Price,[1] Ke Liu,[1] Giannina I. Schaefer,[1] Richard Y. Ebright,[1] Michelle L. Stewart,[1] Daisuke Ito,[1,5] Stephanie Wang,[1] Abigail L. Bracha,[1] Ted Liefeld,[1] Mathias Wawer,[1] Joshua C. Gilbert,[1] Andrew J. Wilson,[2] Nicolas Stransky,[1,6] Gregory V. Kryukov,[1] Vlado Dancik,[1] Jordi Barretina,[1,7] Levi A. Garraway,[1] C. Suk-Yee Hon,[1] Benito Munoz,[1] Joshua A. Bittker,[1] Brent R. Stockwell,[3] Dineo Khabele,[2] Andrew M. Stern,[1] Paul A. Clemons,[1,*] Alykhan F. Shamji,[1,*] and Stuart L. Schreiber[1,*]

[1]The Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

[2]Vanderbilt University School of Medicine, Nashville, TN 37232, USA

[3]Columbia University, New York, NY 10027, USA

[4]These authors contributed equally to this work

[5]Present address: H3 Biomedicine, Cambridge, MA 02139, USA

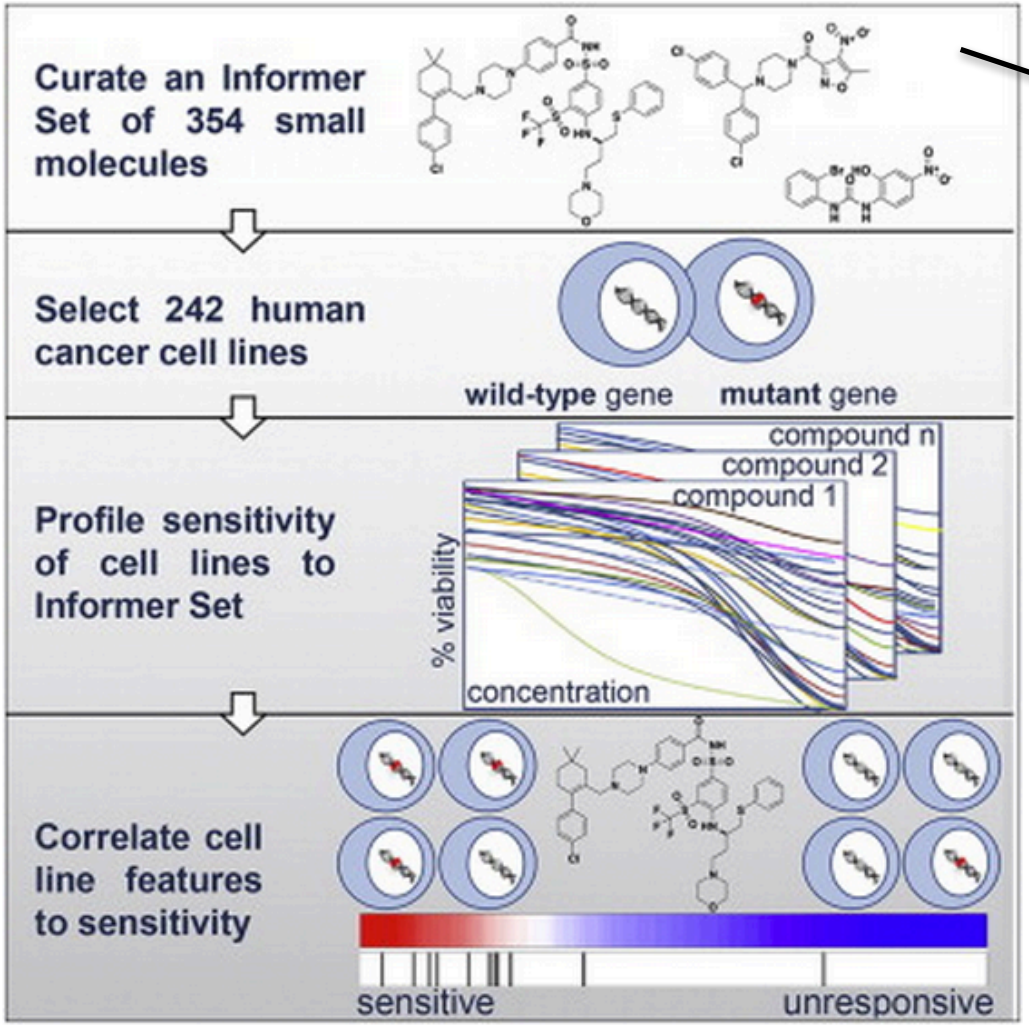
[6]Present address: Blueprint Medicines, Cambridge, MA 02142, USA

[7]Present address: Novartis Institutes for Biomedical Research (NIBR), Cambridge, MA 02139, USA

*Correspondence: pclemons@broadinstitute.org (P.A.C.), ashamji@broadinstitute.org (A.F.S.), stuart_schreiber@harvard.edu (S.L.S.)

http://dx.doi.org/10.1016/j.cell.2013.08.003

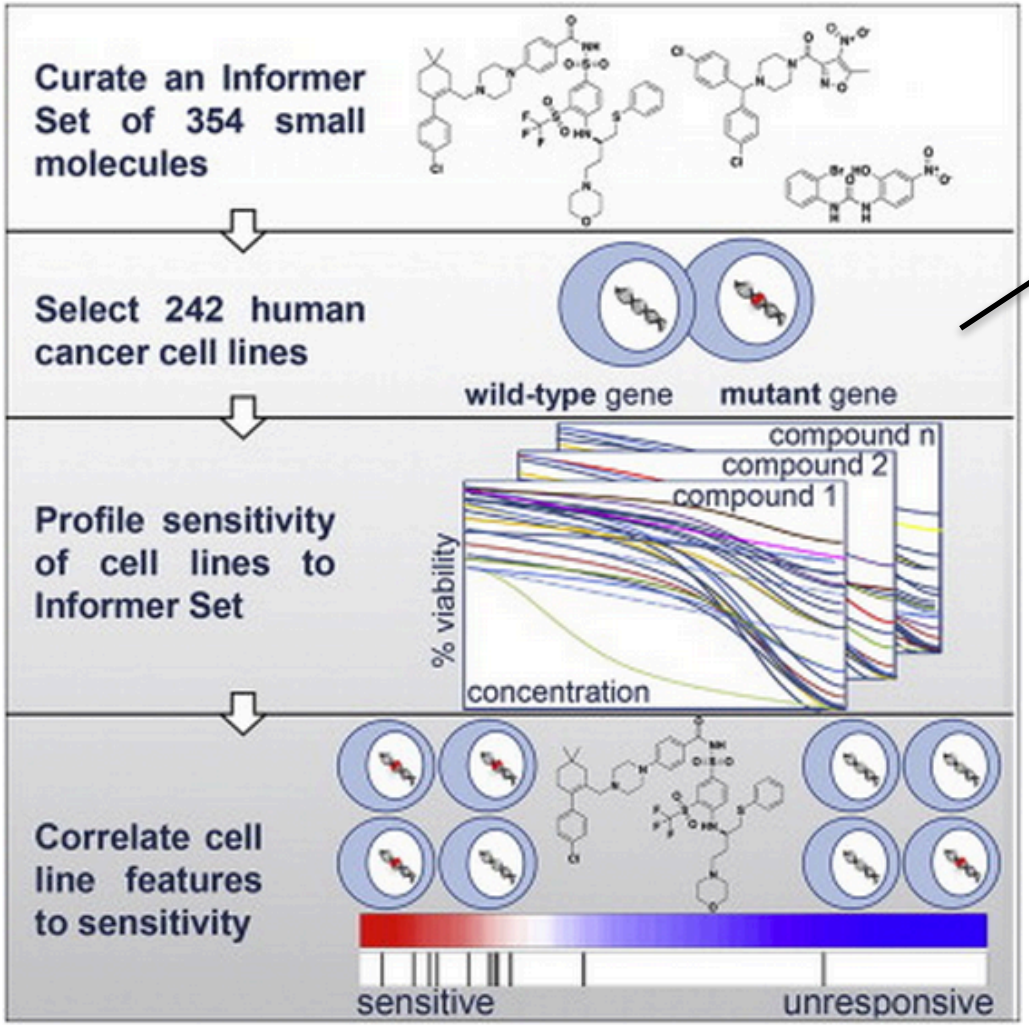
Cancer Therapeutics Response (CTRP)



Two main considerations:
1) High selectivity for their targets
2) Targeting of many distinct nodes in cell circuitry

35 FDA-approved drugs
54 clinical candidates
265 probes

Cancer Therapeutics Response (CTRP)

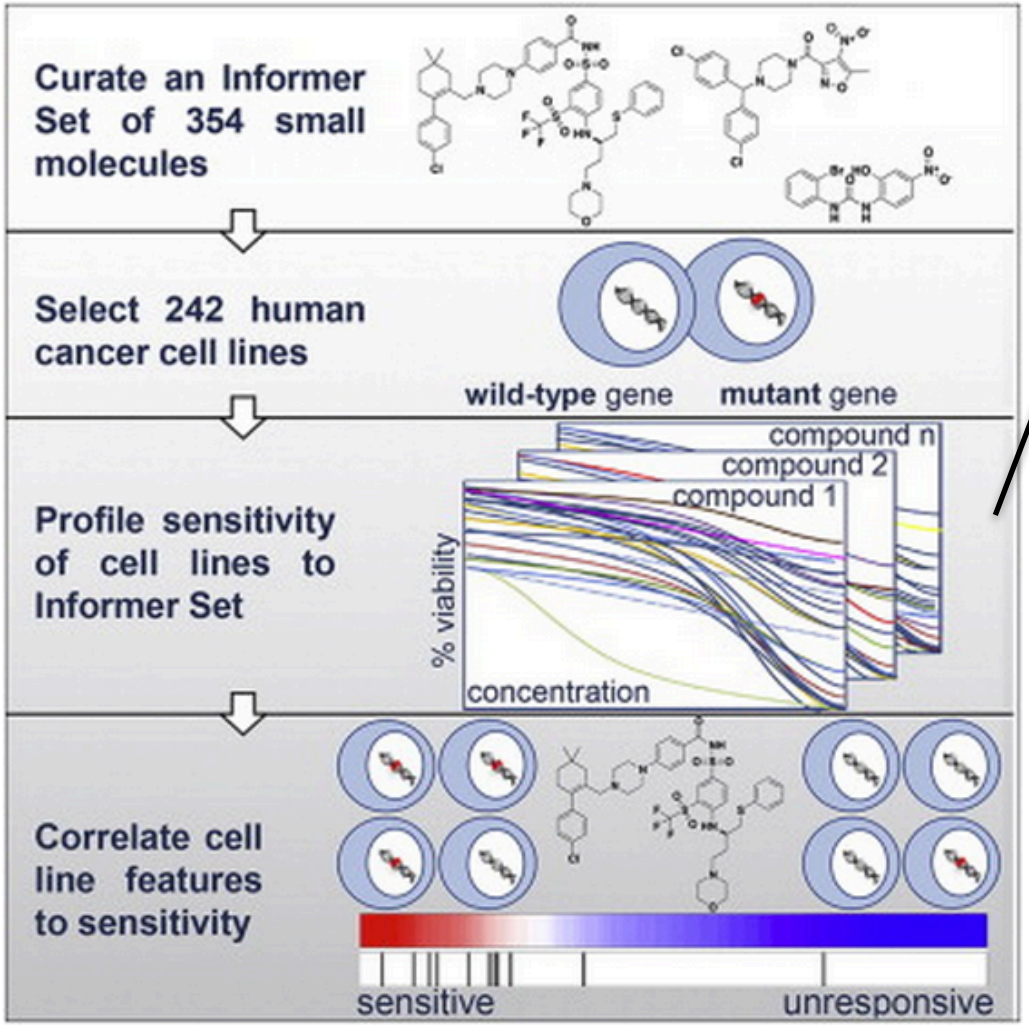


Chosen to align with lineages studied by TCGA and in published genome-wide RNAi screens

Subset of CCLE

Genetically characterized :
Gene expression
Amplifications/deletions
Somatic mutations in 1645 cancer genes
Lineage/histological subtypes

Cancer Therapeutics Response (CTRP)



Each CCL was grown in its preferred media

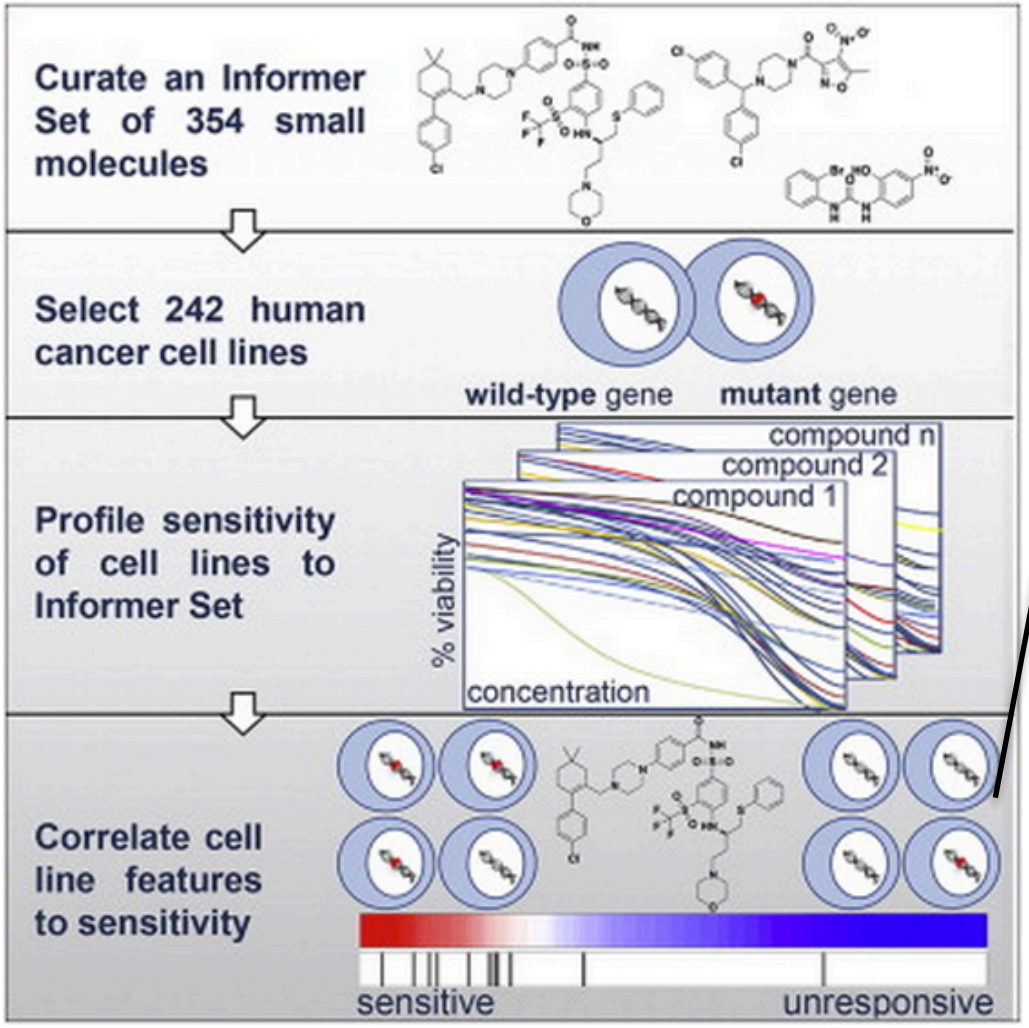
Treated with compound at 8 concentrations for 72 hr

Sensitivity assayed using CellTiter-Glo to measure cellular ATP levels as a surrogate for cell number and growth

Area under percent viability curves (AUC) was computed as a measure of sensitivity

AUC reflects both relative potency and total level of inhibition observed for a compound across CCLs

Cancer Therapeutics Response (CTRP)

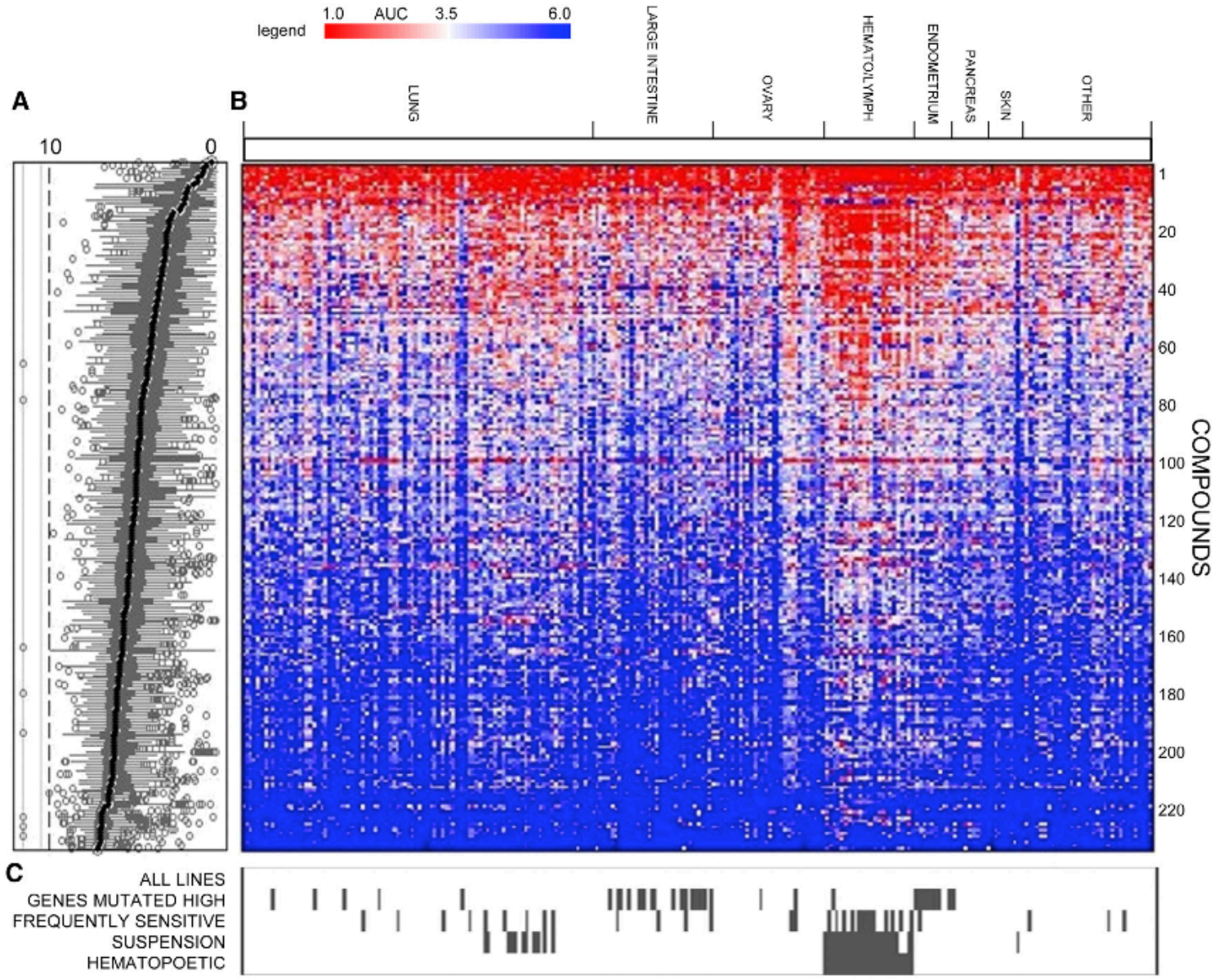


Use AUC as cut-off:
Sensitive (AUC < 3.5) CCLs
Unresponsive (AUC > 5.5) CCLs

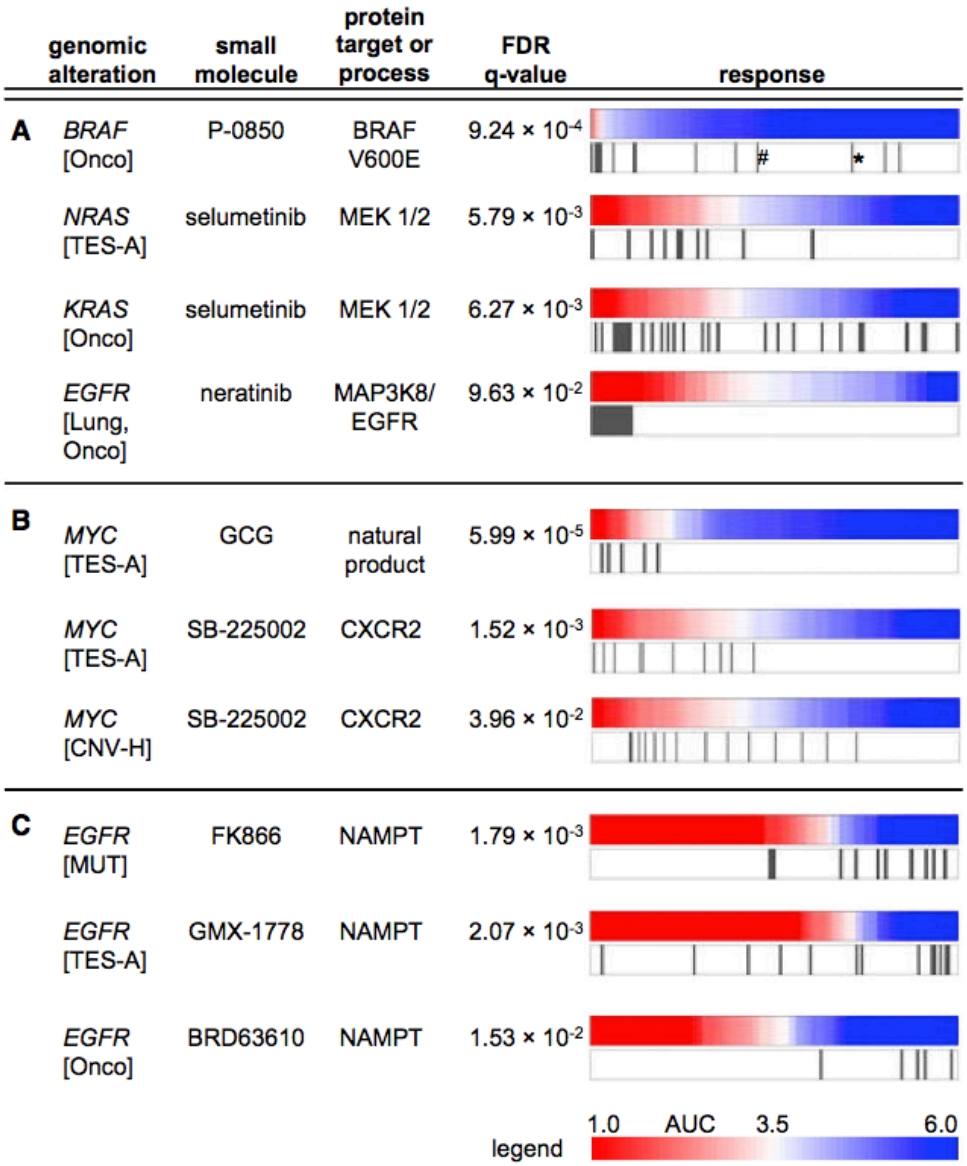
For each compound, performed statistics-based enrichment analysis that combined rank-based and parametric tests to identify genetic alterations and cellular features that are significantly enriched in sensitive or unresponsive CCLs.

>involves individual genetic alterations relative to ranked sensitivities measured for a single compound across many cell lines (p-values)
>elastic net regression analysis across multiple genetic features

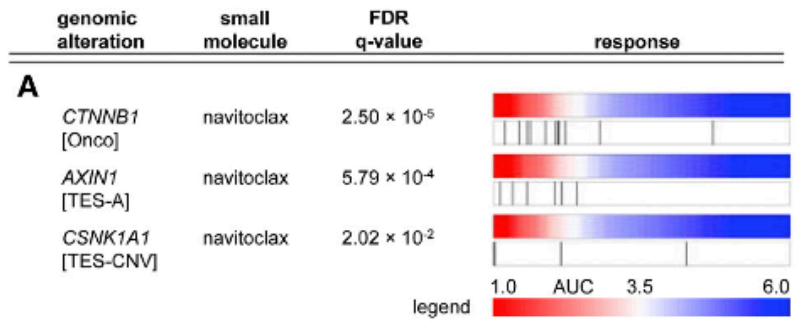
Response of CCLs to Informer Set

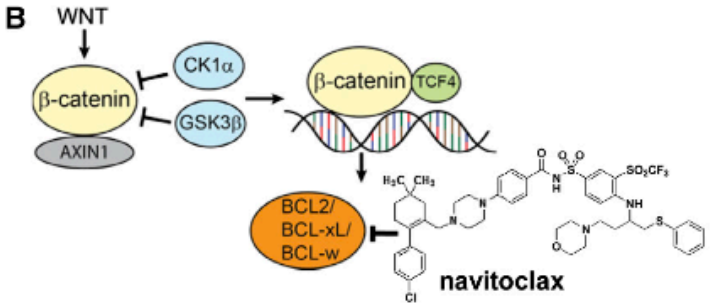


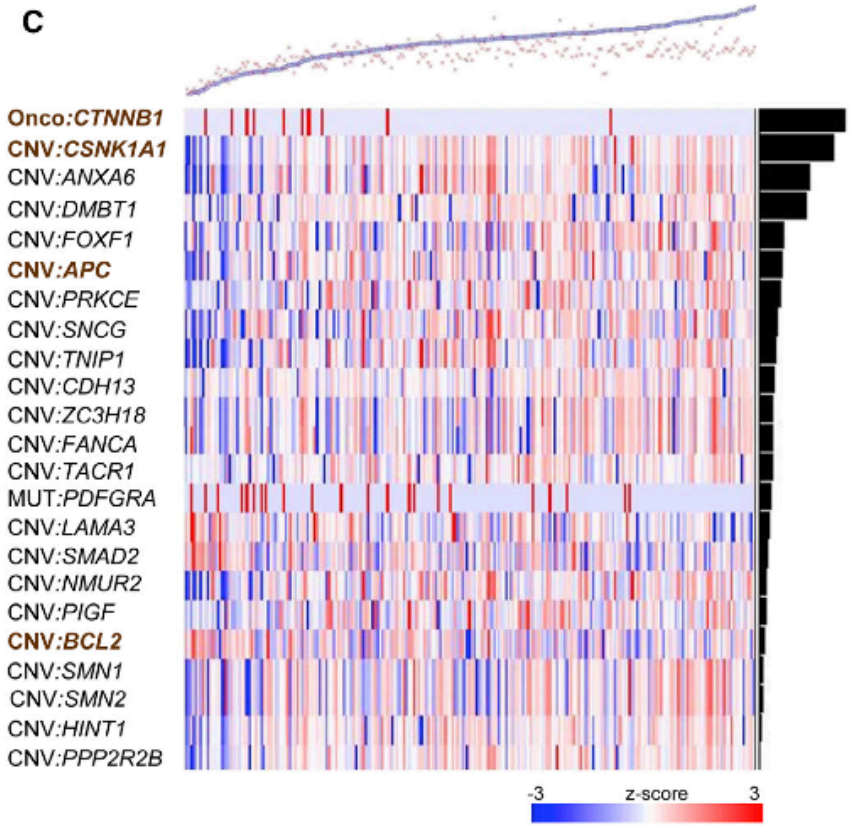
Genetic Dependencies Targeted by Small Molecules

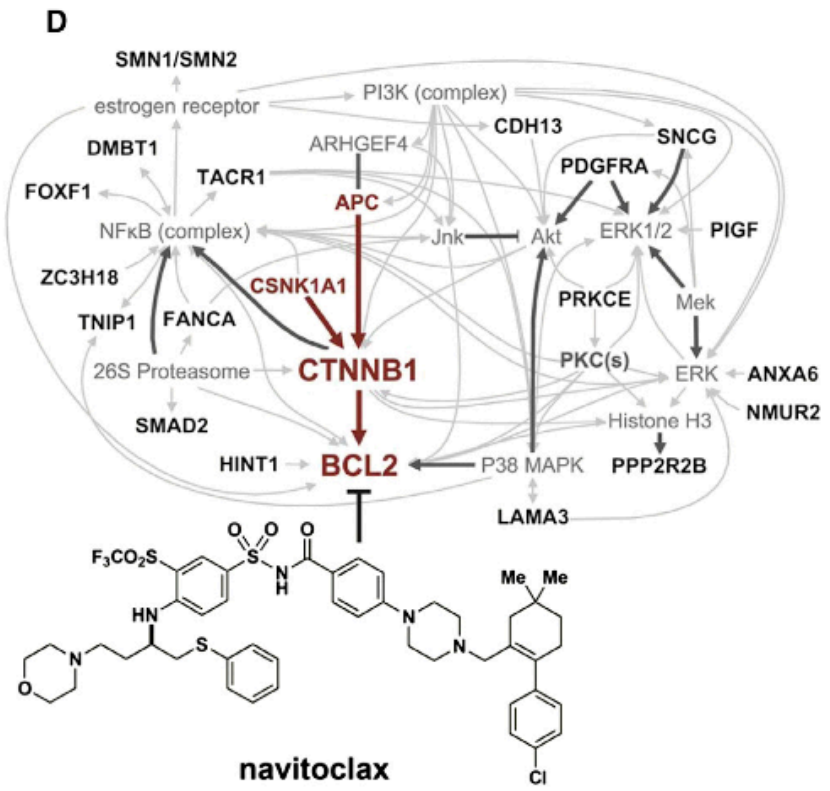


Mutations in b-Catenin Associate with Sensitivity to Navitoclax









Cancer Therapeutics Response Portal

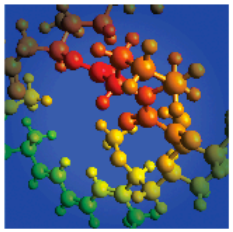
HOME | COMPOUNDS | FEATURES | TARGETS | CONTACT

Profiling cancer cell-line sensitivities with small molecules

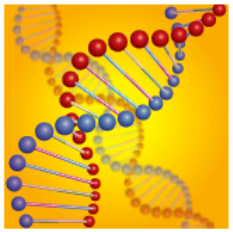
an NCI CTD^2 Network project

Several protein kinase-targeting drugs are yielding high clinical response rates when matched to cancer patients with specific genomic alterations in their cancers. Several other cancer drugs yield similarly high response rates within a particular cancer lineage. These clinical successes have prompted our efforts to identify more systematically additional genetic and lineage context-dependent small-molecule sensitivities.

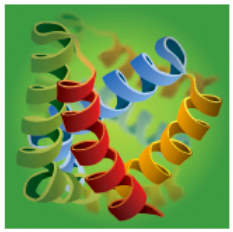
We have generated a novel 'Informer Set' of small-molecule probes and drugs that each selectively target a distinct node in cell circuitry and that collectively modulate a broad array of cell processes. By profiling the impact of this small-molecule collection on a panel of cancer cell lines for which extensive genetic characterizations are publicly available, we have generated a dataset that can be used to identify comprehensively relationships between genetic and lineage features of human cancer cell lines and small-molecule sensitivities.



Small Molecules



Enriched Features



Targets

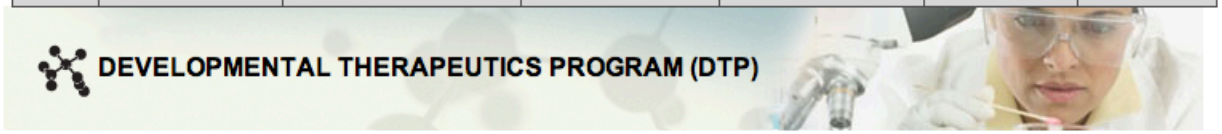
http://www.broadinstitute.org/ctrp/

The Cancer Therapeutics Response Portal provides open access to the results obtained through quantitatively measuring the sensitivity of 242 genetically characterized cancer-cell lines to a 354-member 'Informer Set' of small-molecule probes and drugs. Statistically significant correlations identified between the genetic and lineage features of the cell lines and small-molecule sensitivities are accessible through the portal. With this dataset, users can mine for genetic correlations in a lineage-specific context and control for potential confounding factors. We anticipate continuing to expand the dataset in the portal, providing a living resource for the cancer-research community. We hope that the Portal can be used to develop novel therapeutic hypotheses and to accelerate future discovery of drugs matched to patients based on their cancer genotype and lineage.

Querying Navitoclax in CTRP

Rational Combinations

- Chemical Screening Strategy for Drug-Drug Combinations
  - NCI-60
- Functional Genetic Screening Strategy
  - Loss-of-Function (LOF) Screens
    - Synthetic Lethality Screens
    - Essential Gene Screens
  - Gain-of-Function (GOF) Screens



- Introduction
- Major Ongoing Initiatives
- Current Funding Opportunities
- Tools and Resources
- Scientific Advances

DCTD Programs

Cancer Diagnosis Program

Cancer Imaging Program

Cancer Therapy Evaluation Program

Developmental Therapeutics Program

Radiation Research Program

Translational Research Program

Biometric Research Branch

Office of Cancer Complementary and Alternative Medicine

MAJOR ONGOING INITIATIVES

Last Updated: 04/25/2012

NCI-60 Combination Screening Matrix of Approved Anticancer Drugs

Anticancer drugs are rarely curative as single agents, thus most treatment regimens utilize a combination of agents. Selection of rational combinations based upon presumed mechanisms of action is an active area of research. Nonetheless, treatment regimens may have built up over time by addition of new agents to existing standards of care. To establish a framework for extending the understanding of combination therapy, we have initiated an additional approach: the systematic testing of all pairwise combinations of anticancer drugs approved the Food and Drug Administration (FDA) in the NCI-60 panel of human tumor cell lines.

About 100 small molecule drugs are approved for cancer treatment worldwide. Combinations are tested at 3 or more concentrations of each agent, and single agents are tested on the same plates. About 100 pairs of drugs have been tested in the NCI-60 in a pilot phase. Some combinations of drugs show better than single agent activity in nearly all cell lines tested. Other combinations show more restricted benefit, with some cell lines showing enhanced growth inhibition or cell kill, while in other cell lines the combination is antagonistic or identical to the more active single agent.

The NCI-60 panel of cell lines has been extensively molecularly characterized, with publicly available data including gene mutation, DNA copy number, DNA methylation, and expression of mRNA, protein and microRNA. The patterns of cell line sensitivity for a particular drug combination can be used to probe these molecular characterization data, generating hypotheses about potential predictive markers and/or mechanisms of action. The most promising combinations identified from NCI-60 testing are being tested *in vivo* to determine the therapeutic index. The combination screening data, along with tools to analyze the data, will be made available to the public through the NCI website.

^ Back to Top

< Previous | Next >

NCI ALMANAC

Therapeutics, Targets, and Chemical Biology

Cancer Research

The National Cancer Institute ALMANAC: A Comprehensive Screening Resource for the Detection of Anticancer Drug Pairs with Enhanced Therapeutic Activity

Susan L. Holbeck[1], Richard Camalier[1], James A. Crowell[1], Jeevan Prasaad Govindharajulu[2], Melinda Hollingshead[1], Lawrence W. Anderson[1], Eric Polley[1], Larry Rubinstein[1], Apurva Srivastava[2], Deborah Wilsker[2], Jerry M. Collins[1], and James H. Doroshow[1,3]

Abstract

To date, over 100 small-molecule oncology drugs have been approved by the FDA. Because of the inherent heterogeneity of tumors, these small molecules are often administered in combination to prevent emergence of resistant cell subpopulations. Therefore, new combination strategies to overcome drug resistance in patients with advanced cancer are needed. In this study, we performed a systematic evaluation of the therapeutic activity of over 5,000 pairs of FDA-approved cancer drugs against a panel of 60 well-characterized human tumor cell lines (NCI-60) to uncover combinations with greater than additive growth-inhibitory activity. Screening results were compiled into a database, termed the NCI-ALMANAC (A Large Matrix of Anti-Neoplastic Agent Combinations), publicly available at https://dtp.cancer.gov/ncialmanac. Subsequent *in vivo* experiments in mouse xenograft models of human cancer confirmed combinations with greater than single-agent efficacy. Concomitant detection of mechanistic biomarkers for these combinations *in vivo* supported the initiation of two phase I clinical trials at the NCI to evaluate clofarabine with bortezomib and nilotinib with paclitaxel in patients with advanced cancer. Consequently, the hypothesis-generating NCI-ALMANAC web-based resource has demonstrated value in identifying promising combinations of approved drugs with potent anticancer activity for further mechanistic study and translation to clinical trials. *Cancer Res; 77(13); 3564–76. ©2017 AACR.*

Examples of Chemical Screening of Drug-Drug Combinations

Preclinical Development

Molecular Cancer Therapeutics

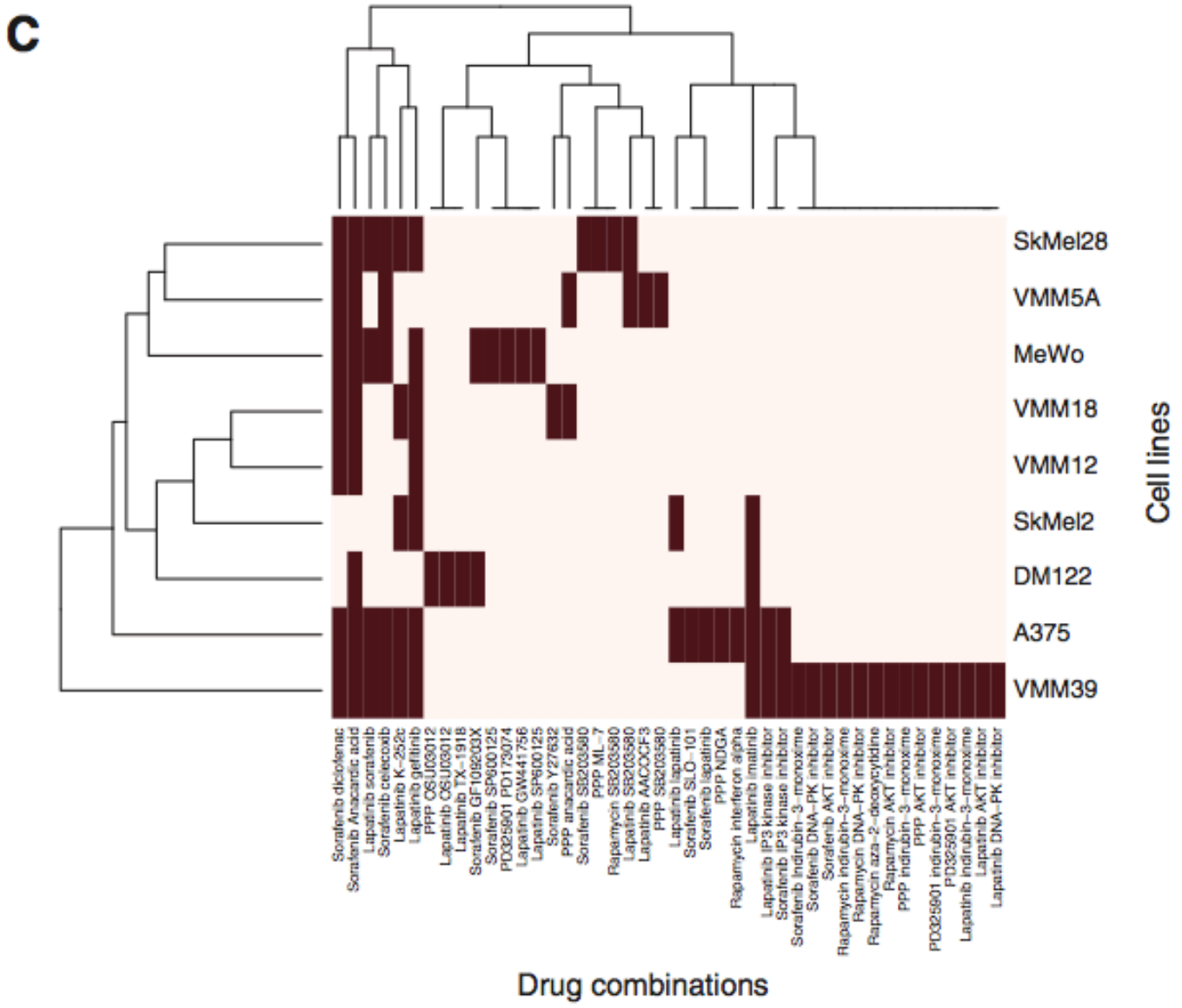
Synthetic Lethal Screening with Small-Molecule Inhibitors Provides a Pathway to Rational Combination Therapies for Melanoma

Devin G. Roller[1], Mark Axelrod[1], Brian J. Capaldo[2], Karin Jensen[3], Aaron Mackey[2,4], Michael J. Weber[1,4], and Daniel Gioeli[1,4]

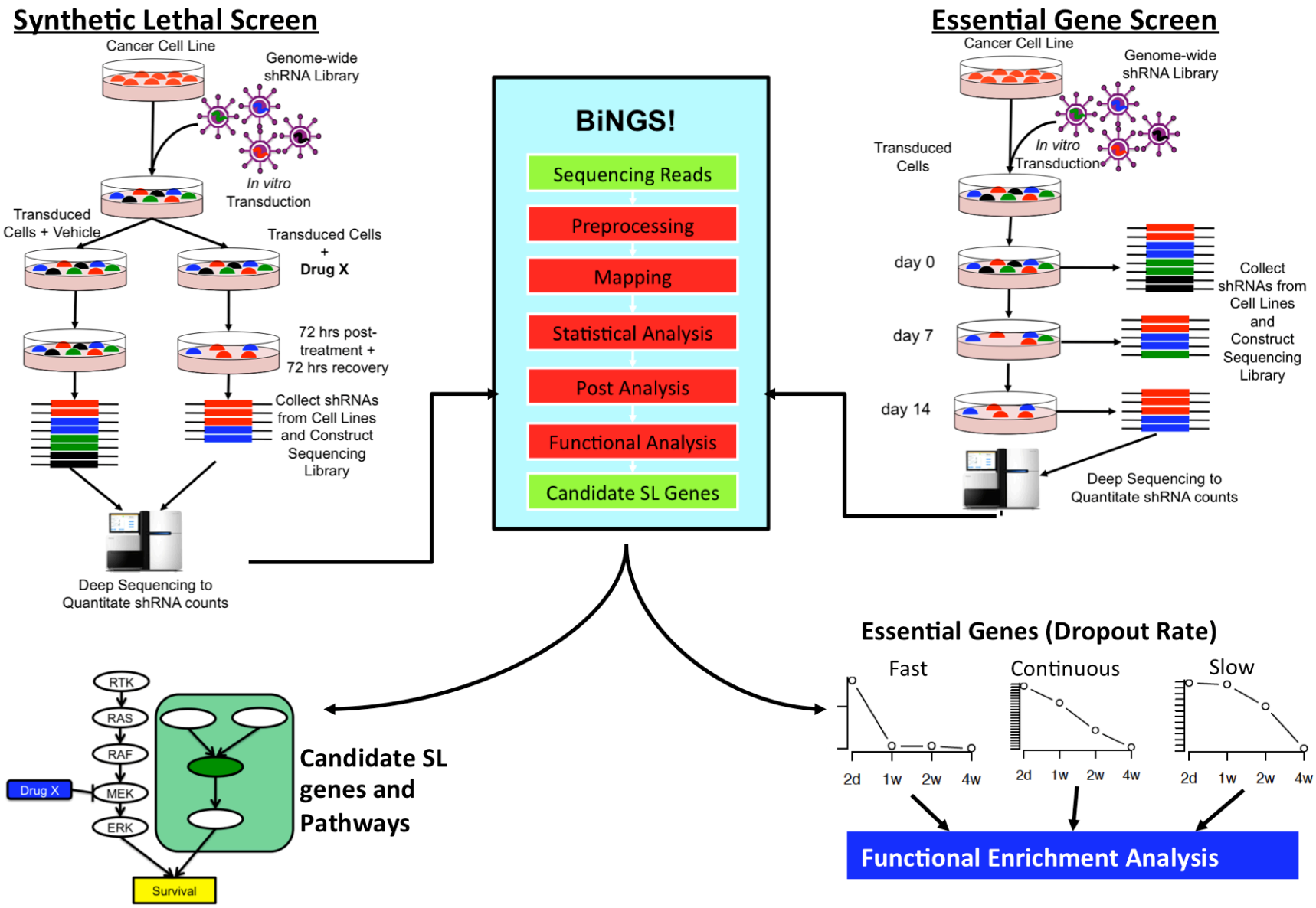
Abstract

Recent data show that extracellular signals are transmitted through a network of proteins rather than hierarchical signaling pathways, suggesting that the inhibition of a single component of a canonical pathway is insufficient for the treatment of cancer. The biologic outcome of signaling through a network is inherently more robust and resistant to inhibition of a single network component. In this study, we conducted a functional chemical genetic screen to identify novel interactions between signaling inhibitors that would not be predicted on the basis of our current understanding of signaling networks. We screened over 300 drug combinations in nine melanoma cell lines and have identified pairs of compounds that show synergistic cytotoxicity. The synergistic cytotoxicities identified did not correlate with the known *RAS* and *BRAF* mutational status of the melanoma cell lines. Among the most robust results was synergy between sorafenib, a multikinase inhibitor with activity against RAF, and diclofenac, a nonsteroidal anti-inflammatory drug (NSAID). Drug substitution experiments using the NSAIDs celecoxib and ibuprofen or the MAP–ERK kinase inhibitor PD325901 and the RAF inhibitor RAF265 suggest that inhibition of COX and mitogen-activated protein kinase signaling are targets for the synergistic cytotoxicity of sorafenib and diclofenac. Cotreatment with sorafenib and diclofenac interrupts a positive feedback signaling loop involving extracellular signal–regulated kinase, cellular phospholipase A2, and COX. Genome-wide expression profiling shows synergy-specific downregulation of survival-related genes. This study has uncovered novel functional drug combinations and suggests that the underlying signaling networks that control responses to targeted agents can vary substantially, depending on unexplored components of the cell genotype. *Mol Cancer Ther; 11(11); 2505–15. ©2012 AACR.*

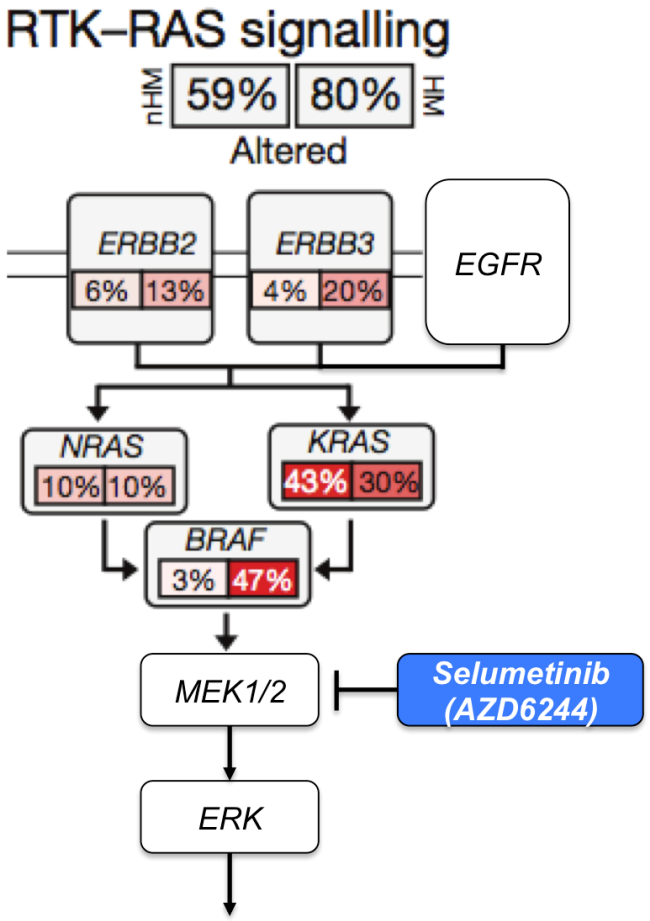
Examples of Chemical Screening of Drug-Drug Combinations



Functional Genetic Screens



Rational Combinations of MEKi and WNTi



Proliferation, cell survival, translation

Cancer Therapy: Preclinical

Clinical Cancer Research

Rational Combination of a MEK Inhibitor, Selumetinib, and the Wnt/Calcium Pathway Modulator, Cyclosporin A, in Preclinical Models of Colorectal Cancer

Anna Spreafico[1], John J. Tentler[1], Todd M. Pitts[1], Aik Choon Tan[1], Mark A. Gregory[2], John J. Arcaroli[1], Peter J. Klauck[1], Martine C. McManus[3], Ryan J. Hansen[6], Jihye Kim[1], Lindsey N. Micel[1,5], Heather M. Selby[1], Timothy P. Newton[1], Kelly McPhillips[1], Daniel L. Gustafson[6], James V. DeGregori[2], Wells A. Messersmith[1], Robert A. Winn[4], and S. Gail Eckhardt[1]

Abstract

Purpose: The mitogen-activated protein kinase (MAPK) pathway is a crucial regulator of cell proliferation, survival, and resistance to apoptosis. MEK inhibitors are being explored as a treatment option for patients with KRAS-mutant colorectal cancer who are not candidates for EGFR-directed therapies. Initial clinical results of MEK inhibitors have yielded limited single-agent activity in colorectal cancer, indicating that rational combination strategies are needed.

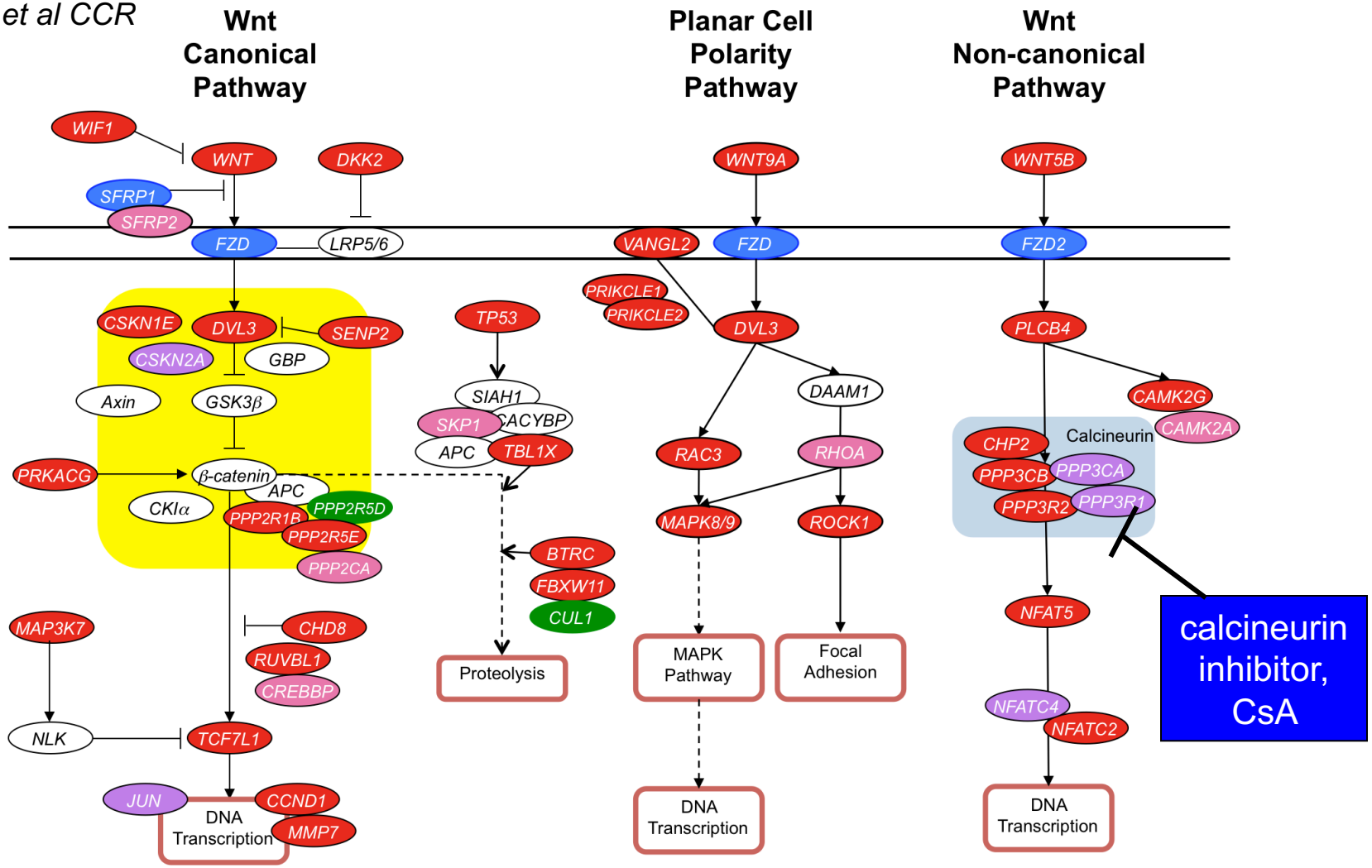
Experimental Design: In this study, we conducted unbiased gene set enrichment analysis and synthetic lethality screens with selumetinib, which identified the noncanonical Wnt/Ca++ signaling pathway as a potential mediator of resistance to the MEK1/2 inhibitor selumetinib. To test this, we used shRNA constructs against relevant WNT receptors and ligands resulting in increased responsiveness to selumetinib in colorectal cancer cell lines. Further, we evaluated the rational combination of selumetinib and WNT pathway modulators and showed synergistic antiproliferative effects in *in vitro* and *in vivo* models of colorectal cancer.

Results: Importantly, this combination not only showed tumor growth inhibition but also tumor regression in the more clinically relevant patient-derived tumor explant (PDTX) models of colorectal cancer. In mechanistic studies, we observed a trend toward increased markers of apoptosis in response to the combination of MEK and WntCa^{++} inhibitors, which may explain the observed synergistic antitumor effects.

Conclusions: These results strengthen the hypothesis that targeting both the MEK and Wnt pathways may be a clinically effective rational combination strategy for patients with metastatic colorectal cancer. *Clin Cancer Res; 1–14. ©2013 AACR.*

SLS for Selumetinib (MEKi) in CRC

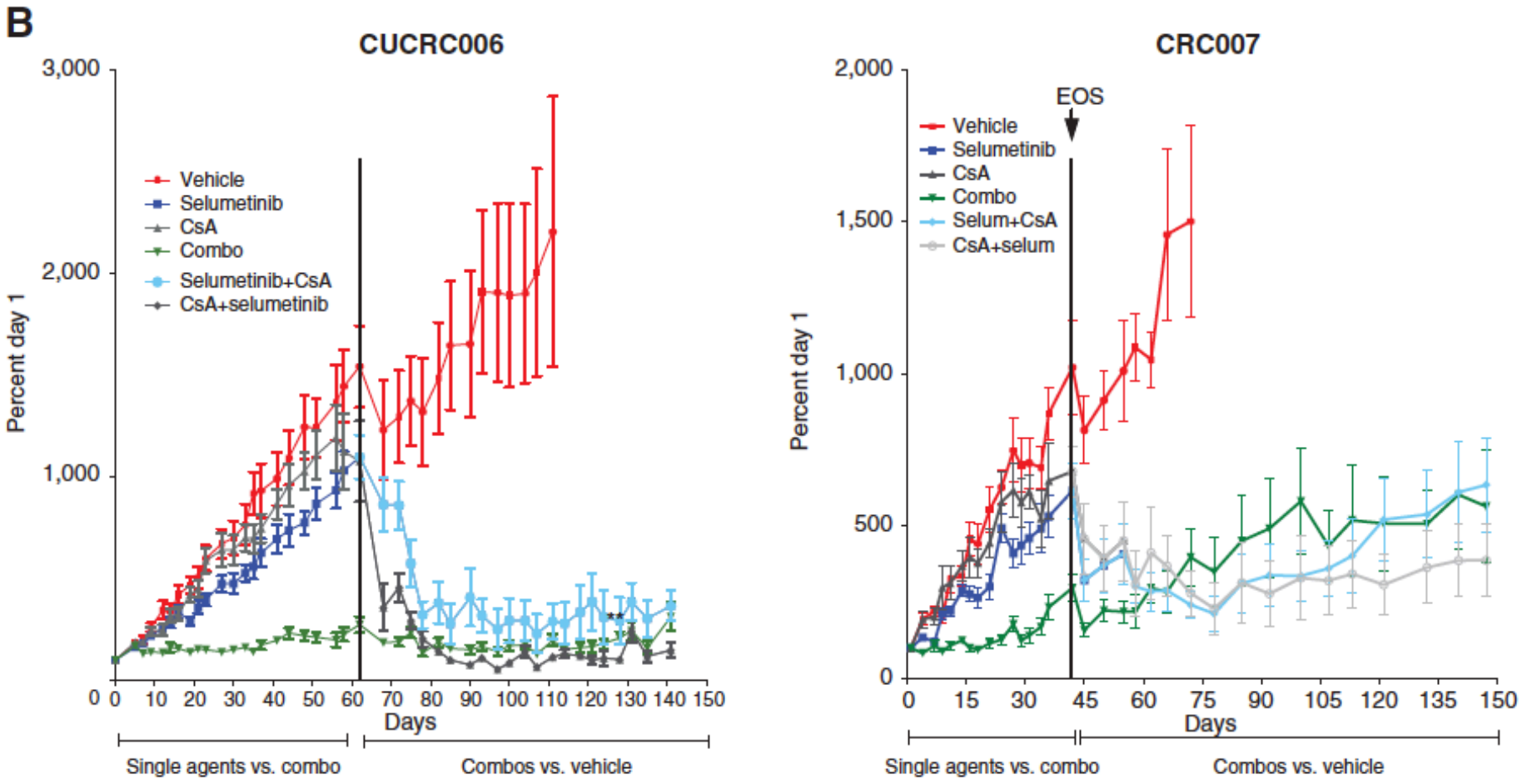
(Spreafico et al CCR 2013)



Legend:

- Core genes by GSEA
- SL hits in SW480
- SL hits in RKO
- Core genes of GSEA and SL hits in SW480
- Core genes of GSEA and SL hits in RKO
- No hit by GSEA or SL

In Vivo Validation study



Translation into Clinical Trial

Selumetinib and Cyclosporine in Treating Patients With Advanced Solid Tumors or Advanced or Metastatic Colorectal Cancer

This study is currently recruiting participants. (see Contacts and Locations)

Verified August 2014 by National Cancer Institute (NCI)

Sponsor:
National Cancer Institute (NCI)

Information provided by (Responsible Party):
National Cancer Institute (NCI)

ClinicalTrials.gov Identifier:
NCT02188264

First received: July 10, 2014
Last updated: November 3, 2014
Last verified: August 2014
History of Changes

Purpose

This phase I trial studies the side effects and best dose of selumetinib when given together with cyclosporine in treating patients with solid tumors or colorectal cancer that have spread to other places in the body and cannot be cured or controlled with treatment. Selumetinib may stop the growth of tumor cells by blocking some of the enzymes needed for cell growth. Biological therapies, such as cyclosporine, use substances made from living organisms that may stimulate or suppress the immune system in different ways and stop tumor cells from growing. Giving selumetinib and cyclosporine may be a better treatment for solid tumors or colorectal cancer.

Condition	Intervention	Phase
Recurrent Colon Cancer Recurrent Rectal Cancer Stage IIIA Colon Cancer Stage IIIA Rectal Cancer Stage IIIB Colon Cancer Stage IIIB Rectal Cancer Stage IIIC Colon Cancer Stage IIIC Rectal Cancer Stage IVA Colon Cancer Stage IVA Rectal Cancer Stage IVB Colon Cancer Stage IVB Rectal Cancer Unspecified Adult Solid Tumor, Protocol Specific	Drug: selumetinib Drug: cyclosporine Other: laboratory biomarker analysis Other: pharmacological study	Phase 1

Study Type: Interventional
Study Design: Endpoint Classification: Safety Study
Intervention Model: Single Group Assignment
Masking: Open Label
Primary Purpose: Treatment

Official Title: A Phase IB Study of the Combination of AZD6244 Hydrogen Sulfate (Selumetinib) and Cyclosporin A (CsA) in Patients With Advanced Solid Tumors With an Expansion Cohort in Metastatic Colorectal Cancer

Clinical Trial Results

Translational Science

Cancer Research

Phase Ib Results of the Rational Combination of Selumetinib and Cyclosporin A in Advanced Solid Tumors with an Expansion Cohort in Metastatic Colorectal Cancer

Anuradha Krishnamurthy[1], Arvind Dasari[2], Anne M. Noonan[3], Janice M. Mehnert[4], Albert C. Lockhart[5], Stephen Leong[1], Anna Capasso[1], Mark N. Stein[4], Hanna K. Sanoff[6], James J. Lee[7], Aaron Hansen[8], Usha Malhotra[4], Sarah Rippke[1], Daniel L. Gustafson[9], Todd M. Pitts[1], Kim Ellison[1], S. Lindsey Davis[1], Wells A. Messersmith[1], S. Gail Eckhardt[1,10], and Christopher H. Lieu[1]

Abstract

MEK inhibition is of interest in cancer drug development, but clinical activity in metastatic colorectal cancer (mCRC) has been limited. Preclinical studies demonstrated Wnt pathway overexpression in *KRAS*-mutant cell lines resistant to the MEK inhibitor, selumetinib. The combination of selumetinib and cyclosporin A, a noncanonical Wnt pathway modulator, demonstrated antitumor activity in mCRC patient-derived xenografts. To translate these results, we conducted a NCI Cancer Therapy Evaluation Program–approved multicenter phase I/IB trial (NCT02188264) of the combination of selumetinib and cyclosporin A. Patients with advanced solid malignancies were treated with the combination of oral selumetinib and cyclosporin A in the dose escalation phase, followed by an expansion cohort of irinotecan and oxaliplatin-refractory mCRC. The expansion cohort utilized a single-agent selumetinib "run-in" to evaluate FZD2 biomarker upregulation and *KRAS*-WT and *KRAS*-MT stratification to identify any potential predictors of efficacy. Twenty and 19 patients were enrolled in dose escalation and expansion phases, respectively. The most common adverse events and grade 3/4 toxicities were rash, hypertension, and edema. Three dose-limiting toxicities (grade 3 hypertension, rash, and increased creatinine) were reported. The MTD was selumetinib 75 mg twice daily and cyclosporin A 2 mg/kg twice daily on a 28-day cycle. *KRAS* stratification did not identify any differences in response between *KRAS*-WT and *KRAS*-MT cancers. Two partial responses, 18 stable disease, and 10 progressive disease responses were observed. Combination selumetinib and cyclosporin A is well tolerated, with evidence of activity in mCRC. Future strategies for concept development include identifying better predictors of efficacy and improved Wnt pathway modulation.

Significance: These findings translate preclinical studies combining selumetinib and cyclosporin into a phase I first-in-human clinical trial of such a combination in patients with advanced solid malignancies. *Cancer Res; 78(18); 5398–407.* ©2018 AACR.

Table 1. Patient baseline characteristics

Characteristics	Escalation cohort ($n = 20$) No. of patients (%)	Expansion cohort ($n = 19$) No. of patients (%)
Age (years)		
18–64	13 (68.4%)	12 (57.1%)
65+	6 (31.6%)	9 (42.8%)
Sex		
Male	9 (47.4%)	10 (47.6%)
Female	10 (52.6%)	11 (52.4%)
Tumor primary site		
Colorectal	13	
Renal	1	
Prostate	1	
Pancreas	1	
ECOG Performance Status		
0	7 (36.8%)	10 (47.6%)
1	12 (63.2%)	11 (52.4%)

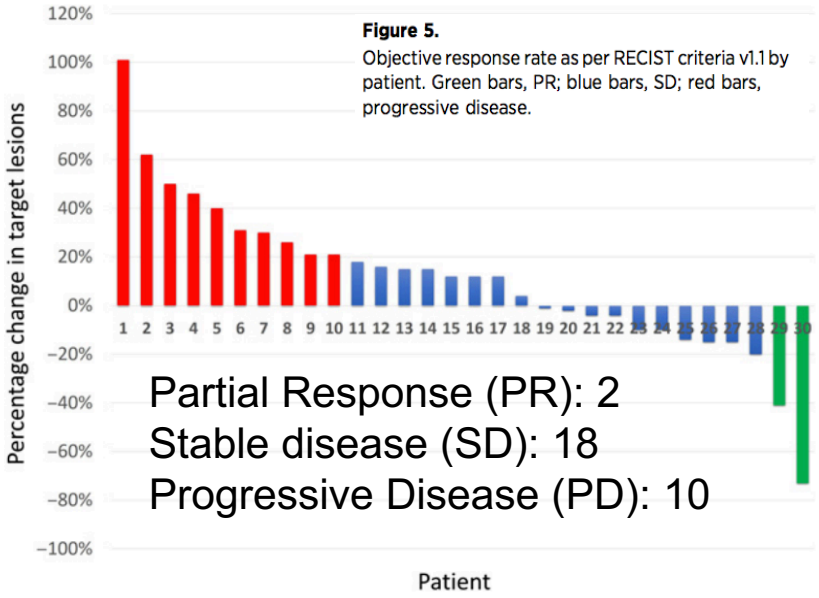


Figure 5.
Objective response rate as per RECIST criteria v1.1 by patient. Green bars, PR; blue bars, SD; red bars, progressive disease.

Partial Response (PR): 2
Stable disease (SD): 18
Progressive Disease (PD): 10

Combined Clinical Benefit Rate (CR+PR+SD): 67%

Essential Gene Screen

RESEARCH ARTICLE

Essential Gene Profiles in Breast, Pancreatic, and Ovarian Cancer Cells

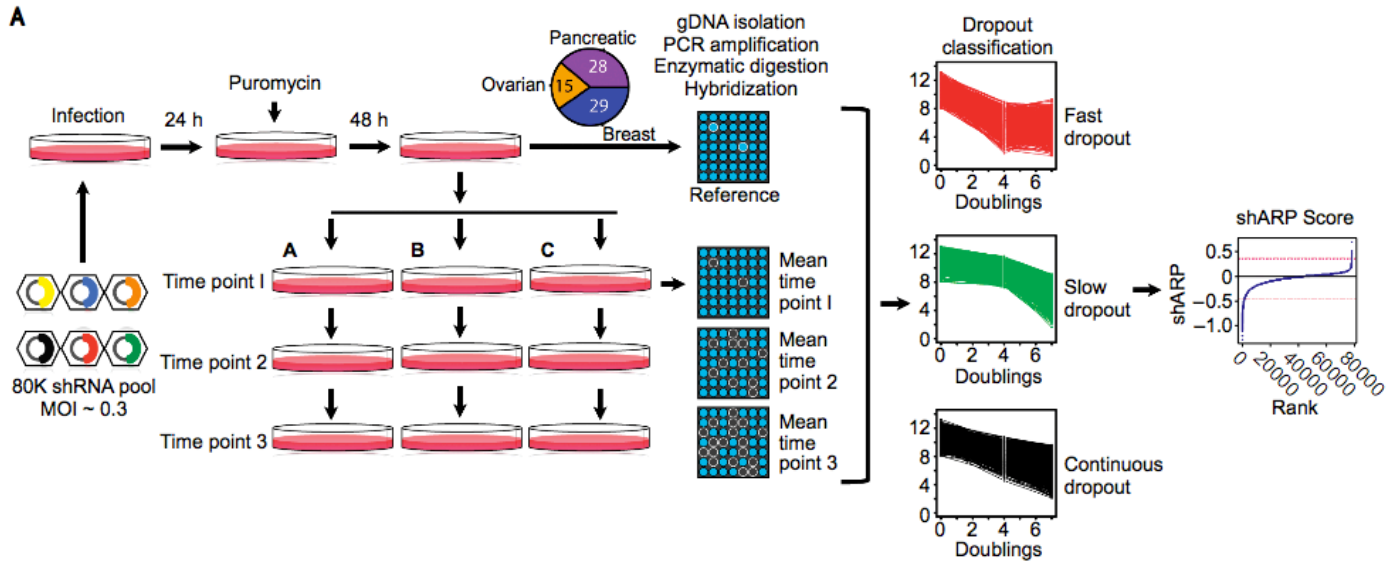
Richard Marcotte[4], Kevin R. Brown[1], Fernando Suarez[4], Azin Sayad[1], Konstantina Karamboulas[1], Paul M. Krzyzanowski[4], Fabrice Sircoulomb[4], Mauricio Medrano[3,4], Yaroslav Fedyshyn[1], Judice L.Y. Koh[1], Dewald van Dyk[1], Bohdana Fedyshyn[1], Marianna Luhova[1], Glauber C. Brito[1], Franco J. Vizeacoumar[1], Frederick S. Vizeacoumar[5], Alessandro Datti[5,7], Dahlia Kasimer[1], Alla Buzina[1], Patricia Mero[1], Christine Misquitta[1], Josee Normand[4], Maliha Haider[4], Troy Ketela[1], Jeffrey L. Wrana[2,5], Robert Rottapel[3,4,6], Benjamin G. Neel[3,4], and Jason Moffat[1,2]

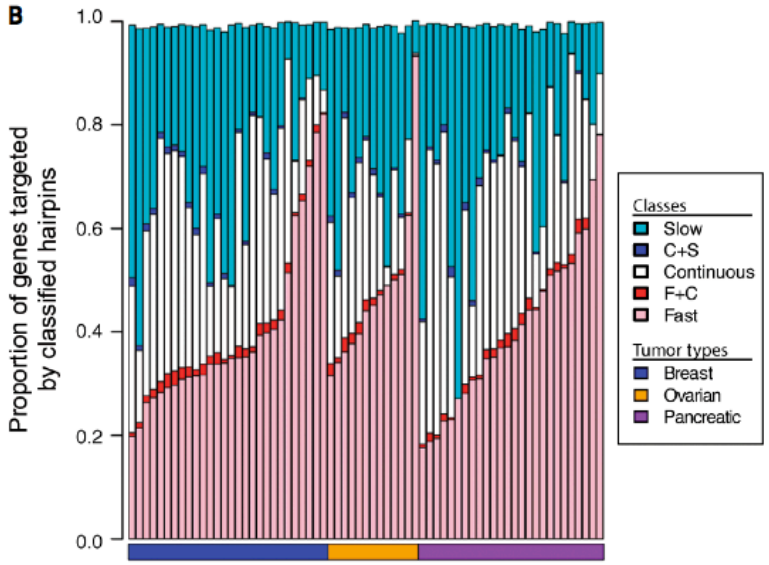
ABSTRACT

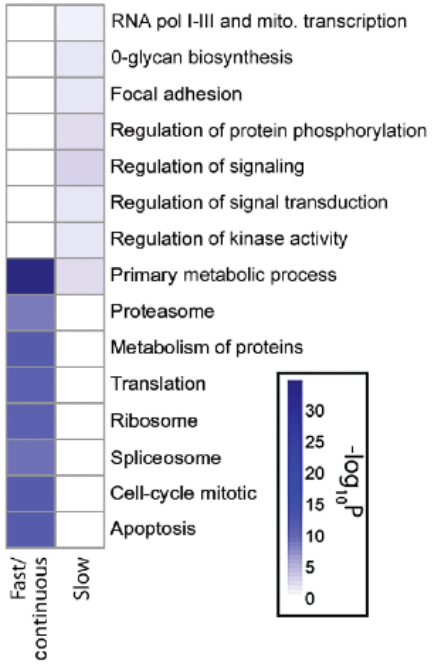
Genomic analyses are yielding a host of new information on the multiple genetic abnormalities associated with specific types of cancer. A comprehensive description of cancer-associated genetic abnormalities can improve our ability to classify tumors into clinically relevant subgroups and, on occasion, identify mutant genes that drive the cancer phenotype ("drivers"). More often, though, the functional significance of cancer-associated mutations is difficult to discern. Genome-wide pooled short hairpin RNA (shRNA) screens enable global identification of the genes essential for cancer cell survival and proliferation, providing a "functional genomic" map of human cancer to complement genomic studies. Using a lentiviral shRNA library targeting ~16,000 genes and a newly developed, dynamic scoring approach, we identified essential gene profiles in 72 breast, pancreatic, and ovarian cancer cell lines. Integrating our results with current and future genomic data should facilitate the systematic identification of drivers, unanticipated synthetic lethal relationships, and functional vulnerabilities of these tumor types.

SIGNIFICANCE: This study presents a resource of genome-scale, pooled shRNA screens for 72 breast, pancreatic, and ovarian cancer cell lines that will serve as a functional complement to genomics data, facilitate construction of essential gene profiles, help uncover synthetic lethal relationships, and identify uncharacterized genetic vulnerabilities in these tumor types. *Cancer Discovery; 2(2);* 172–89. ©2011 AACR.

Essential Gene Screen





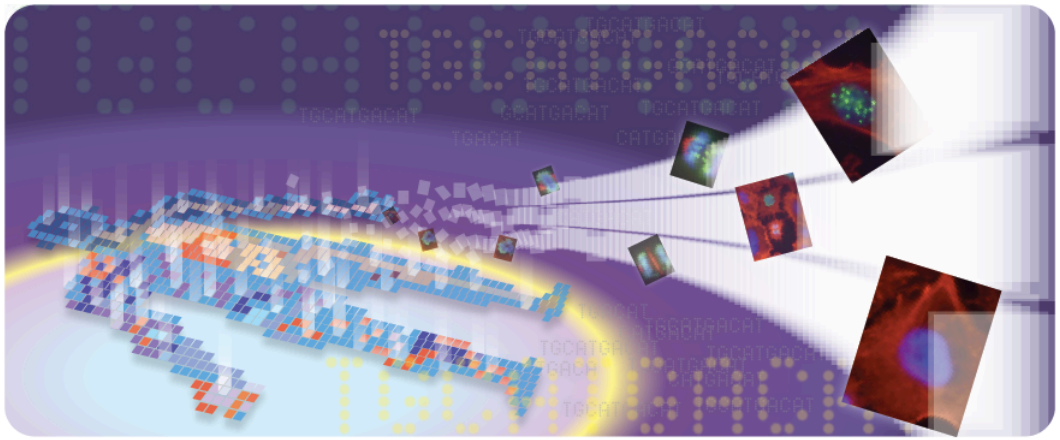


Project Achilles

Project Achilles | BROAD INSTITUTE

Login

Data | Resources | About | Contact



Project Achilles is a systematic effort aimed at identifying and cataloging gene essentiality across hundreds of genomically characterized cancer cell lines. The project uses genome-scale RNAi and CRISPR-Cas9 genetic perturbation reagents to silence or knockout individual genes and identify those genes that affect cell survival. By linking these genetic dependencies to the genetic or molecular features of the tumors, this project is providing the foundation for a "Cancer Dependency Map" (https://depmap.org).

Search Genes | Search

Citing Achilles

Aviad Tsherniak, Francisca Vazquez, Phil G. Montgomery, Barbara A. Weir, *et al.*, Defining a cancer Dependency Map. *Cell.* July 27, 2017. DOI: j.cell.2017.06.010

Aguirre, A.J., Meyers, R.M., Weir, B.A., Vazquez, F., Zhang, C.Z., *et al.* (2016). Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer Discov* **6**, 914-929. June 3, 2016.

Cowley, G.S., Weir, B.A., Vazquez, F., Tamayo, P., *et al.* (2014) Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Nature Scientific Data* **1**, Article number: 140035. September 30, 2014.

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» RNAi
» Cancer genomics

Received: 20 May 2014
Accepted: 22 August 2014
Published: 30 September 2014
Updated: 11 November 2014

Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies

Glenn S. Cowley[1,2,*], Barbara A. Weir[1,2,*], Francisca Vazquez[1,2,*], Pablo Tamayo[1], Justine A. Scott[1], Scott Rusin[1], Alexandra East-Seletsky[1], Levi D. Ali[1], William F.J. Gerath[1], Sarah E. Pantel[1], Patrick H. Lizotte[1], Guozhi Jiang[1], Jessica Hsiao[1], Aviad Tsherniak[1], Elizabeth Dwinell[1], Simon Aoyama[1], Michael Okamoto[1], William Harrington[1], Ellen Gelfand[1], Thomas M. Green[1], Mark J. Tomko[1], Shuba Gopal[1], Terence C. Wong[1], Hubo Li[1], Sara Howell[1], Nicolas Stransky[1], Ted Liefeld[1], Dongkeun Jang[1], Jonathan Bistline[1], Barbara Hill Meyers[1], Scott A. Armstrong[1], Ken C. Anderson[1], Kimberly Stegmaier[1,3], Michael Reich[1], David Pellman[1], Jesse S. Boehm[1], Jill P. Mesirov[1], Todd R. Golub[1], David E. Root[1] & William C. Hahn[1,2,4,5]

Using a genome-scale, lentivirally delivered shRNA library, we performed massively parallel pooled shRNA screens in 216 cancer cell lines to identify genes that are required for cell proliferation and/or viability. Cell line dependencies on 11,000 genes were interrogated by 5 shRNAs per gene. The proliferation effect of each shRNA in each cell line was assessed by transducing a population of 11M cells with one shRNA-virus per cell and determining the relative enrichment or depletion of each of the 54,000 shRNAs after 16 population doublings using Next Generation Sequencing. All the cell lines were screened using standardized conditions to best assess differential genetic dependencies across cell lines. When combined with genomic characterization of these cell lines, this dataset facilitates the linkage of genetic dependencies with specific cellular contexts (e.g., gene mutations or cell lineage). To enable such comparisons, we developed and provided a bioinformatics tool to identify linear and nonlinear correlations between these features.

Design Type(s)	genotyping design • cell type comparison design • RNAi screening • loss-of-function screening by pooled shRNA
Measurement Type(s)	SNP interrogation genotyping • cell viability assay
Technology Type(s)	microfluidics platform • next generation sequencing
Factor Type(s)	tumor subtype • growth medium • doubling time • study personnel
Sample Characteristic(s)	*Homo sapiens* • A2780 cell • BJHTERT • C2BBe1 cell • COLO-783 • EFO-21 cell • GP2D cell • IGROV-1 cell • JHESOAD1 • KM12 • LN215 • LN319 • LN382 • NCI-H1792 cell • OAW42 cell • ... • ZR-75-30 cell

[1]Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Department of Medical, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. [3]Department of Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. [4]Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. [5]Departments of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [6]Blueprint Medicines, Inc. 215 1st Street, Cambridge, Massachusetts 02142, USA. [7]Memorial Sloan Kettering, 1275 York Ave, New York, New York 10065, USA. *These authors contributed equally to this work.

Correspondence and requests for materials should be addressed to D.E.R. (email: droot@broadinstitute.org) or to W.C.H. (email: William_Hahn@dfci.harvard.edu).

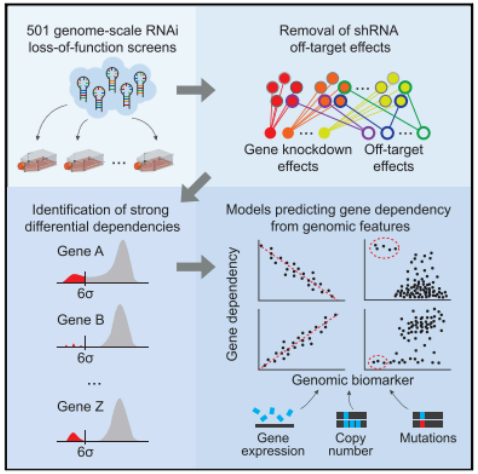
Cancer Dependency Map

Resource

Cell

Defining a Cancer Dependency Map

Graphical Abstract



Authors

Aviad Tsherniak, Francisca Vazquez, Phil G. Montgomery, ..., Todd R. Golub, Jesse S. Boehm, William C. Hahn

Correspondence

william_hahn@dfci.harvard.edu

In Brief

A large-scale analysis of 501 cancer cell lines reveals new vulnerabilities that will help prioritize therapeutic targets

Highlights

- The DEMETER computational model segregates on- from off-target effects of RNAi
- 769 strong differential dependencies were identified in 501 cancer cell lines
- Predictive models for 426 dependencies were found using 66,646 molecular features
- This cancer dependency map facilitates the prioritization of therapeutic targets

Tsherniak et al., 2017, Cell *170*, 564–576
July 27, 2017 © 2017 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2017.06.010

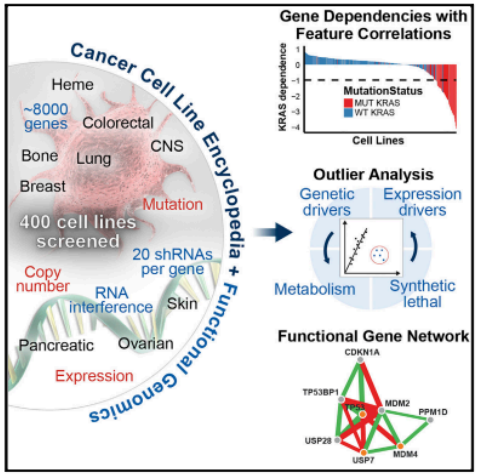
CellPress

Resource

Cell

Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening

Graphical Abstract



Authors

E. Robert McDonald III, Antoine de Weck, Michael R. Schlabach, ..., Francesco Hofmann, Tobias Schmelzle, William R. Sellers

Correspondence

rob.mcdonald@novartis.com (E.R.M.), tobias.schmelzle@novartis.com (T.S.)

In Brief

A large-scale RNAi screen in 398 cancer cell lines reveals vulnerabilities of specific cancer subtypes.

Highlights

- Project DRIVE: deep RNAi interrogation of viability effects in cancer
- ~8,000 genes were targeted by a median of 20 shRNAs per gene in ~400 CCLE models
- Data robustly define cancer dependency genes falling into distinct outlier classes
- Enables genetic networks that reveal protein complexes and biological pathways

McDonald et al., 2017, Cell *170*, 577–592
July 27, 2017 © 2017 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2017.07.005

CellPress

How to Connect Drugs Automatically?

AML

+cytarabine

TP53^{WT}

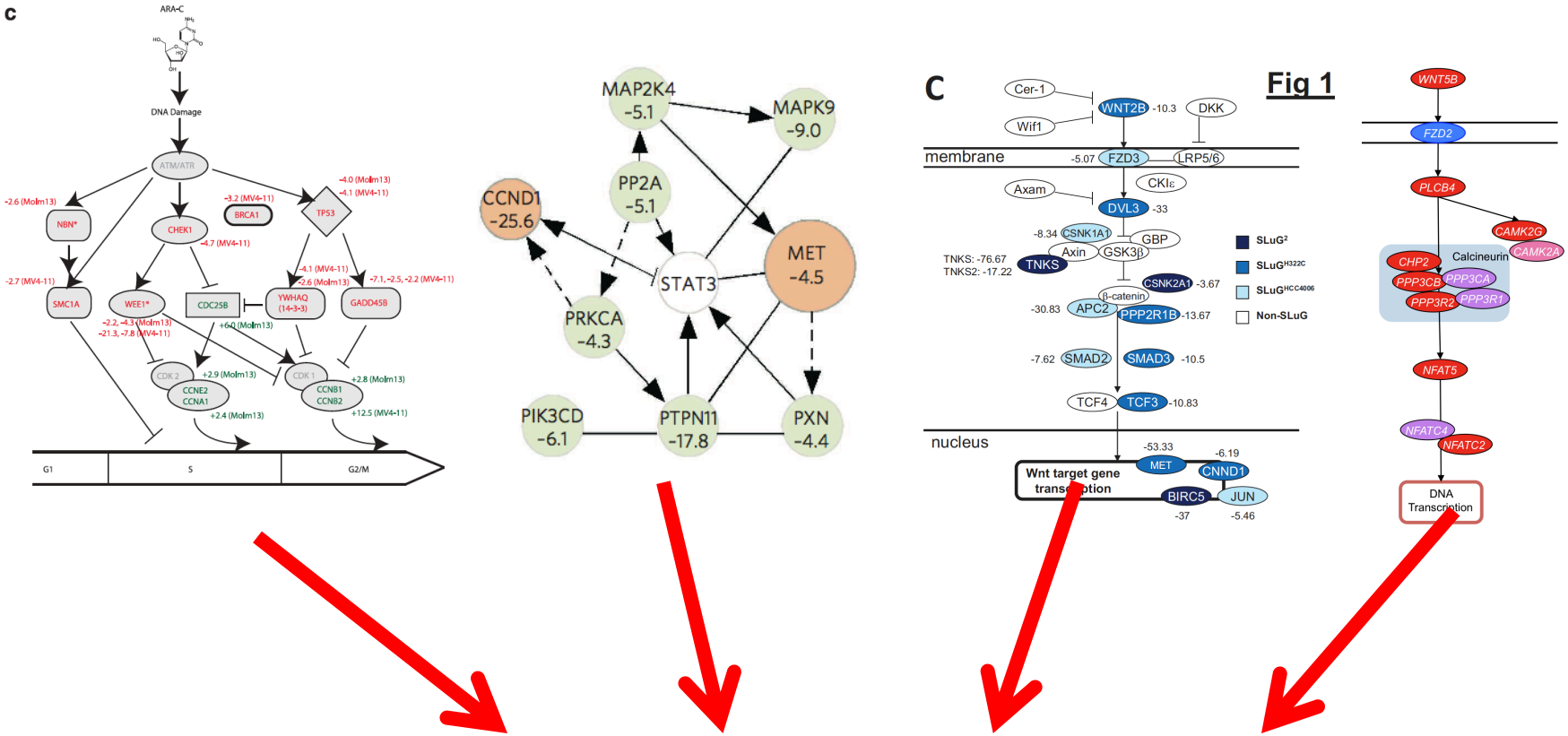
+nutlin-3

NSCLC

+gefitinib

CRC

+selumetinib



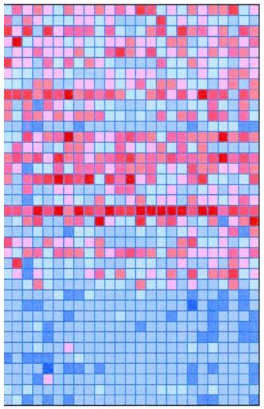
How to find Drugs Automatically?

Google-like Approach

Google Search | I'm Feeling Lucky

Advanced Search
Preferences
guage Tools

Input: Gene expression profile



Output: Therapeutic options

Rank	Drug
1	TAXOTERE (docetaxel)
2	GEMZAR gemcitabine HCl (for injection) + Tarceva erlotinib tablets
3	Xeloda capecitabine
4	VELCADE (bortezomib) FOR INJECTION

The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease

Justin Lamb,[1*] Emily D. Crawford,[1†] David Peck,[1] Joshua W. Modell,[1] Irene C. Blat,[1] Matthew J. Wrobel,[1] Jim Lerner,[1] Jean-Philippe Brunet,[1] Aravind Subramanian,[1] Kenneth N. Ross,[1] Michael Reich,[1] Haley Hieronymus,[1,2] Guo Wei,[1,2] Scott A. Armstrong,[2,3] Stephen J. Haggarty,[1,4] Paul A. Clemons,[1] Ru Wei,[1] Steven A. Carr,[1] Eric S. Lander,[1,5,6] Todd R. Golub[1,2,3,5,7*]

To pursue a systematic approach to the discovery of functional connections among diseases, genetic perturbation, and drug action, we have created the first installment of a reference collection of gene-expression profiles from cultured human cells treated with bioactive small molecules, together with pattern-matching software to mine these data. We demonstrate that this "Connectivity Map" resource can be used to find connections among small molecules sharing a mechanism of action, chemicals and physiological processes, and diseases and drugs. These results indicate the feasibility of the approach and suggest the value of a large-scale community Connectivity Map project.

Connectivity Map

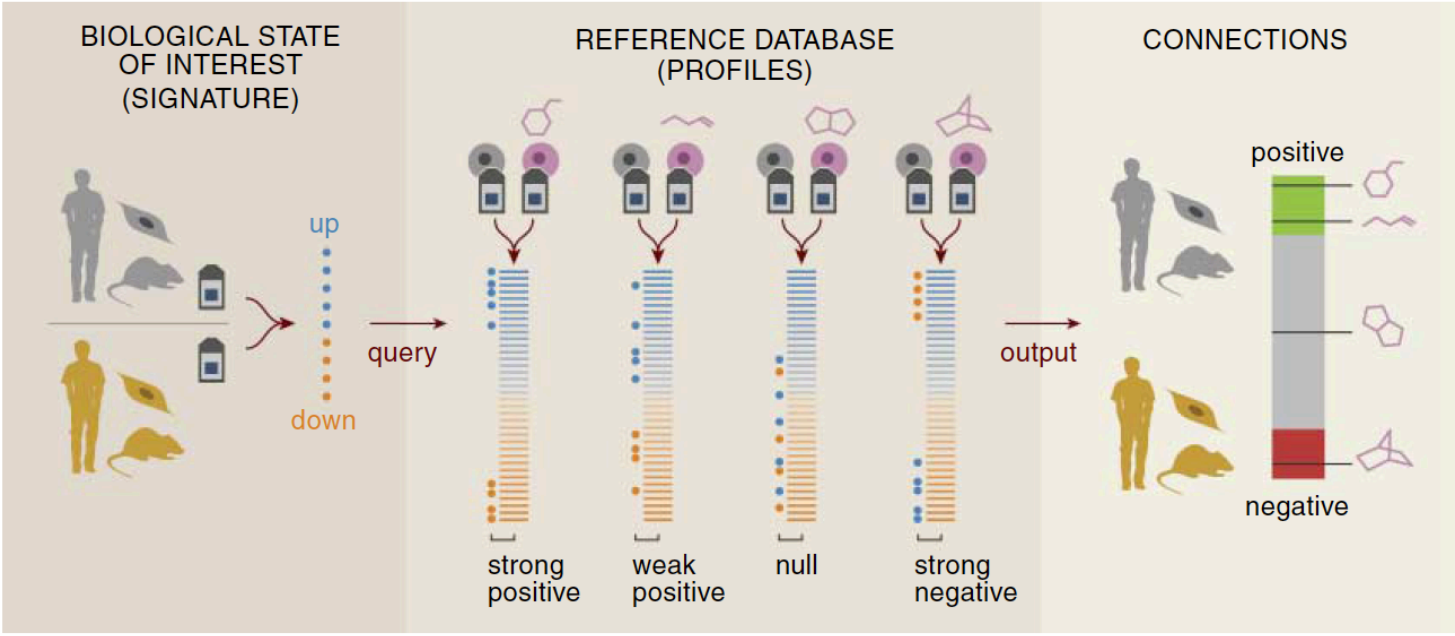


Fig. 1. The Connectivity Map Concept. Gene-expression profiles derived from the treatment of cultured human cells with a large number of perturbagens populate a reference database. Gene-expression signatures represent any induced or organic cell state of interest (**left**). Pattern-matching algorithms score each reference profile for the direction and strength of enrichment with the query signature (**center**). Perturbagens are ranked by this "connectivity score"; those at the top ("positive") and bottom ("negative") are functionally connected with the query state (**right**) through the transitory feature of common gene-expression changes.

Connectivity Map

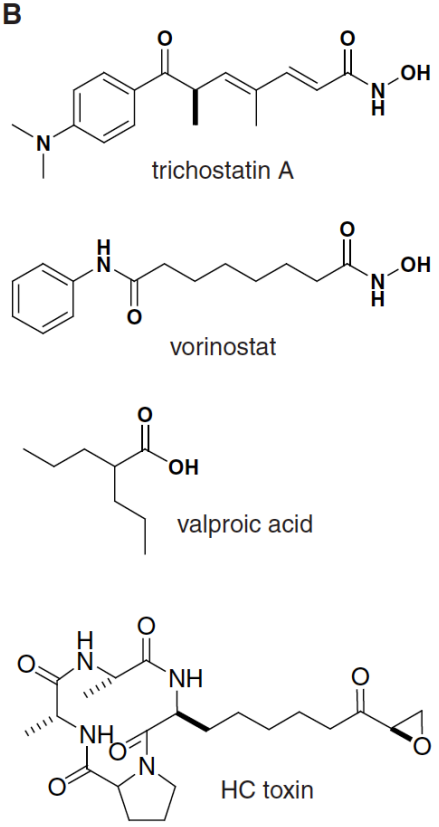
HDAC inhibitors. We first determined whether a query signature derived from a class of small molecules could recover those same compounds in the Connectivity Map. A recent report (14) described gene-expression responses of T24 (bladder), MDA 435 (breast carcinoma), and MDA 468 (breast carcinoma) cells treated with three histone deacetylase (HDAC) inhibitors: vorinostat (also known as suberoylanilide hydroxamic acid or SAHA), MS-27-275, and trichostatin A. The authors of this study defined a 13-gene signature (8 up-regulated and 5 downregulated genes; Signature S1) that was used to query our database.

Fig. 2. HDAC Inhibitors. (**A**) HDAC inhibitors are highly ranked with an external HDAC inhibitor signature. The "barview" is constructed from 453 horizontal lines, each representing an individual treatment instance, ordered by their corresponding connectivity scores with the Glaser *et al.* (*14*) signature (+1, top; –1, bottom). All valproic acid (n = 18), trichostatin A (n = 12), vorinostat (n = 2), and HC toxin (n = 1) instances in the data set are colored in black. Colors applied to the remaining instances reflect the sign of their scores (green, positive; gray, null; red, negative). The rank, name [instance id], concentration, cell line, and connectivity score for each of the selected HDAC inhibitor instances is shown. Unabridged results from this query are provided as Result S1. (**B**) Chemical structures.

A

rank	perturbagen	dose	cell	score
1	vorinostat [1000]	10 μM	MCF7	1
2	trichostatin A [873]	1 μM	MCF7	0.969
3	trichostatin A [992]	100 nM	MCF7	0.931
4	trichostatin A [1050]	100 nM	MCF7	0.929
5	vorinostat [1058]	10 μM	MCF7	0.917
6	trichostatin A [981]	1 μM	MCF7	0.915
7	HC toxin [909]	100 nM	MCF7	0.914
8	trichostatin A [1112]	100 nM	MCF7	0.908
9	trichostatin A [1072]	1 μM	MCF7	0.906
10	trichostatin A [1014]	1 μM	MCF7	0.893
11	trichostatin A [332]	100 nM	MCF7	0.882
12	trichostatin A [331]	100 nM	MCF7	0.846
13	trichostatin A [448]	100 nM	PC3	0.788
14	valproic acid [345]	10 mM	MCF7	0.743
15	valproic acid [23]	1 mM	MCF7	0.735
16	valproic acid [1047]	1 mM	MCF7	0.733
17	trichostatin A [413]	100 nM	ssMCF7	0.725
18	valproic acid [410]	10 mM	HL60	0.725
19	valproic acid [458]	1 mM	PC3	0.680
33	valproic acid [409]	1 mM	HL60	0.634
39	valproic acid [1020]	500 μM	MCF7	0.619
52	valproic acid [346]	2 mM	MCF7	0.582
61	valproic acid [1078]	500 μM	MCF7	0.563
71	valproic acid [629]	1 mM	SKMEL5	0.539
72	valproic acid [347]	500 μM	MCF7	0.539
73	valproic acid [989]	1 mM	MCF7	0.538
76	valproic acid [433]	1 mM	PC3	0.528
89	trichostatin A [364]	100 nM	HL60	0.507
92	valproic acid [497]	1 mM	ssMCF7	0.501
297	valproic acid [348]	50 μM	MCF7	0
388	valproic acid [994]	200 μM	MCF7	0
403	valproic acid [1002]	50 μM	MCF7	0
419	valproic acid [1060]	50 μM	MCF7	-0.537

B



Connectivity Map Webpage

CONNECTIVITY MAP 02

instances | query | results | admin | downloads | help

permuted results

HDACi [99730] (06-NOV-13)

total instances: 6100 , *signature:* HDACi (Glaser), *export:* Excel

search:

by name | by name and cell line | by ATC code

rank	cmap name	mean	n	enrichment	p	specificity	% non-null
1	vorinostat	0.865	12	0.973	0.00000	0.0201	100
2	trichostatin A	0.786	182	0.895	0.00000	0.0095	97
3	geldanamycin	0.484	15	0.705	0.00000	0.0163	100
4	fluphenazine	0.388	18	0.629	0.00000	0.0155	88
5	trifluoperazine	0.392	16	0.625	0.00000	0.0625	87
6	thioridazine	0.440	20	0.599	0.00000	0.1278	85
7	tanespimycin	0.431	62	0.574	0.00000	0.0259	87
8	sirolimus	0.337	44	0.491	0.00000	0.0542	77
9	LY-294002	0.324	61	0.486	0.00000	0.0738	68
10	valproic acid	0.304	57	0.359	0.00000	0.0263	61
11	CP-690334-01	0.507	8	0.735	0.00002	0.0121	87
12	rifabutin	0.735	3	0.971	0.00004	0.0052	100
13	5707885	0.549	4	0.913	0.00004	0.0000	100
14	pioglitazone	-0.337	11	-0.646	0.00004	0.0061	72
15	6-bromoindirubin-3'-oxime	-0.532	7	-0.770	0.00008	0.0047	85
16	withaferin A	0.542	4	0.896	0.00010	0.0632	100
17	wortmannin	0.382	18	0.501	0.00010	0.1355	77
18	ivermectin	0.461	5	0.858	0.00012	0.0215	100
19	prochlorperazine	0.362	16	0.524	0.00014	0.1262	68
20	suloctidil	0.553	4	0.888	0.00016	0.0182	100

result detail

vorinostat

rank	batch	cmap name	dose	cell	score	up	down	instance_
4	725	vorinostat	10 µM	MCF7	.983	.887	-.888	521
10	506	vorinostat	10 µM	MCF7	.970	.860	-.892	100
11	757	vorinostat	10 µM	MCF7	.970	.855	-.896	558
42	765	vorinostat	10 µM	MCF7	.911	.869	-.776	698
49	767	vorinostat	10 µM	MCF7	.905	.845	-.790	693
56	626	vorinostat	10 µM	MCF7	.900	.784	-.841	164
63	513	vorinostat	10 µM	MCF7	.891	.863	-.744	105
84	727	vorinostat	10 µM	PC3	.864	.746	-.814	444
128	603	vorinostat	10 µM	PC3	.795	.680	-.756	122
139	602	vorinostat	10 µM	HL60	.777	.496	-.908	116
153	750	vorinostat	10 µM	HL60	.726	.490	-.820	617

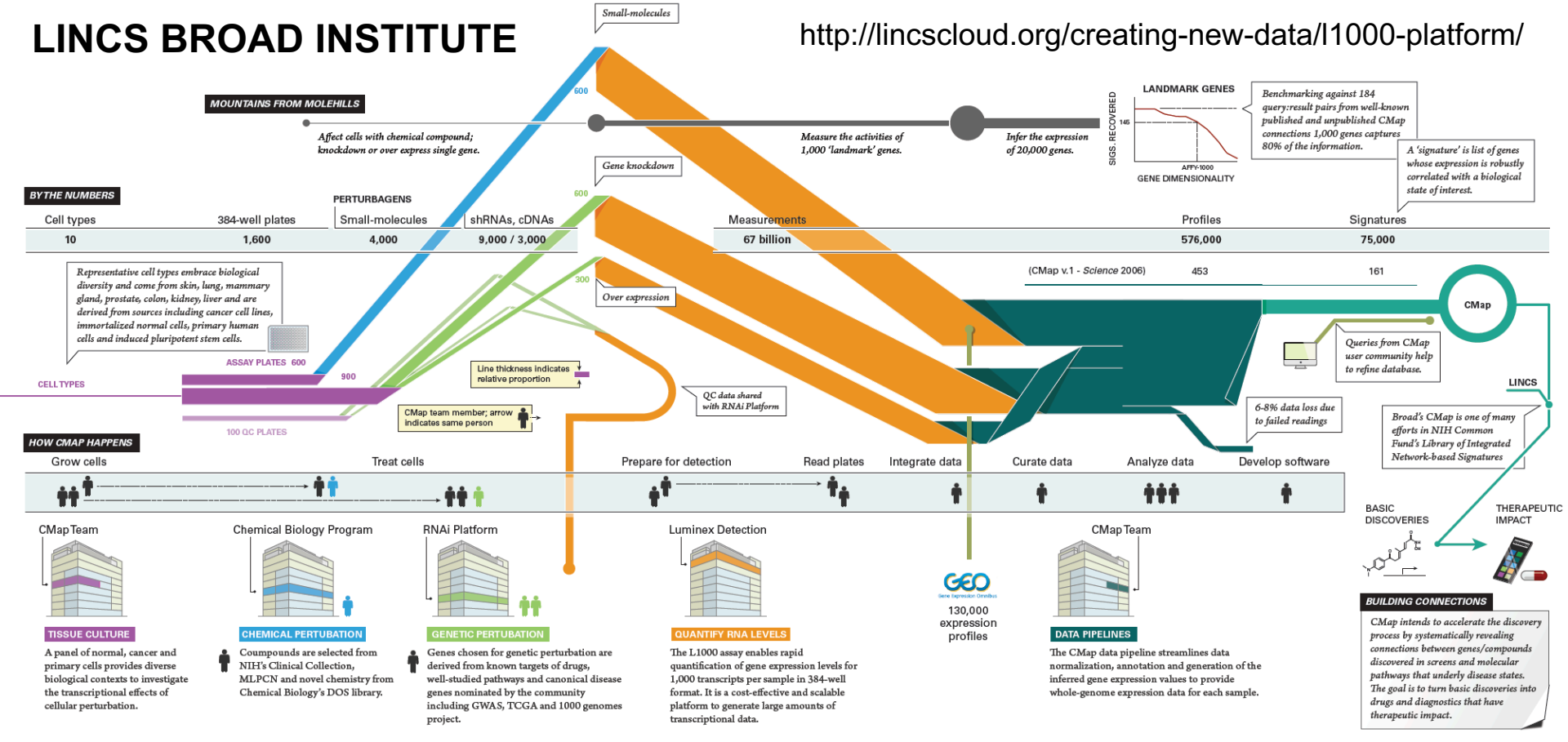
http://www.broadinstitute.org/cmap/

Library of Integrated Network-Based Cellular Signatures (NIH LINCS Program)

LINCS aims to create a network-based understanding of biology by cataloging changes in gene expression and other cellular processes that occur when cells are exposed to a variety of perturbing agents, and by using computational tools to integrate this diverse information into a comprehensive view of normal and disease states that can be applied for the development of new biomarkers and therapeutics. By generating and making public data that indicates how cells respond to various genetic and environmental stressors, the LINCS project will help us gain a more detailed understanding of cell pathways and aid efforts to develop therapies that might restore perturbed pathways and networks to their normal states.

LINCS BROAD INSTITUTE

http://lincscloud.org/creating-new-data/l1000-platform/



LINCS BROAD INSTITUTE (http://lincscloud.org/)

A LINCS RESOURCE FOR FUNCTIONAL GENOMICS

THE Connectivity Map

SIGN IN | CREATE AN ACCOUNT

SEARCH THE SITE

beta for LINCS

THE DATA SET
- Current data
- The Landmark Genes
- Perturbagens assayed
- Cell types profiled
- How data were prepared

EXPLORING THE DATA

WEB APPS
- Browse datasets
- Browse connections
- Query with signatures
- Visual workbench
- Integrative analysis
- Examples of data/analysis

FOR DEVELOPERS
- Data API
- Code API
- Cloud integration
- Inference challenge

CREATING NEW DATA
- L1000 platform
- Protocol
- Collaborations

ABOUT
- Projects
- Publications
- Acknowledgements

Video

Welcome

The Connectivity Map (or CMap) is a catalog of gene-expression data collected from human cells treated with chemical compounds and genetic reagents. Computational methods to reduce the number of necessary genomic measurements along with streamlined methodologies enable the current effort to significantly increase the size of the CMap database and along with it, our potential to connect human diseases with the genes that underlie them and the drugs that treat them.

CMap intends to accelerate the discovery process by systematically revealing connections between genes/compounds discovered in screens and molecular pathways that underlie disease states. The goal is to turn basic discoveries into drugs and diagnostics that have therapeutic impact.

NEWS

SEPTEMBER 25, 2013
The first public conference on LINCS will be held at the Broad Institute in November 2013. Register here.

MARCH 11, 2013
The Connectivity Map website is now hosted at Amazon Web Services. Users will need to reregister for the site. Details.

NOVEMBER 15, 2012
The beta version of the new Connectivity Map website, lincscloud.org, is launched.

OCTOBER 4, 2012
Broad LINCS project compiles over 950K+ profiles, 15 cell lines, and over 20k+ perturbagens.

NIH LINCS PROGRAM

BROAD INSTITUTE

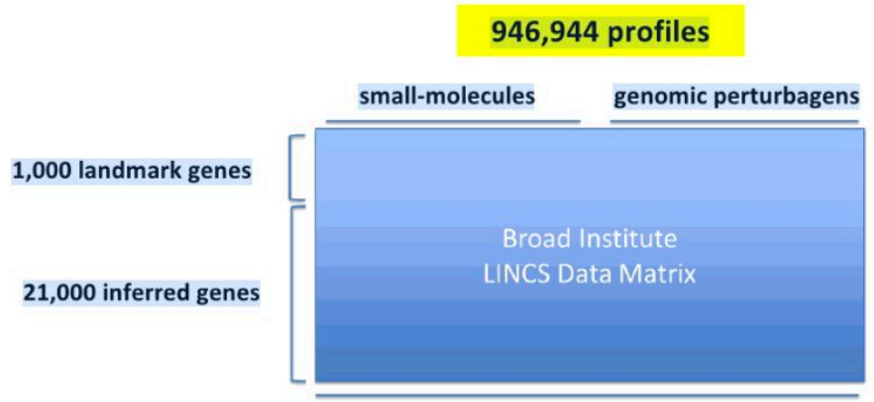
The current data

The types of perturbagens in the dataset are:

- **Small-molecule compounds.**
  The compounds were selected from multiple sources, including known drugs, pathway-specific tool compounds, and compounds of interest identified in NIH-sponsored small-molecule screening efforts in addition to nominations from the research community.
- **Human genes perturbed using lentivirally-delivered shRNAs.**
- **Genes profiled for the effect of over-expression.**
  Genes for the genetic perturbations were chosen to include known targets of FDA-approved drugs, drug-target pathway members, candidate disease genes, and genes nominated by the research community.

Accessing the data: See Data API

List of perturbagens profiled: See Perturbagens assayed



5,178 compounds

- 1,300 off-patent FDA-approved drugs
- 700 bioactive tool compounds
- 2,000+ screening hits (MLPCN and others)

3,712 genes (shRNA + cDNA)

- targets/pathways of FDA-approved drugs (n=900)
- candidate disease genes (n=600)
- community nominations (n=500+)

15 cell types

- Banked primary cell types
- Cancer cell lines
- Primary hTERT immortalized
- Patient derived iPS cells
- 5 community nominated

LINCS BROAD INSTITUTE (https://clue.io/)

A Next Generation Connectivity Map: L1000 platform and the first 1,000,000 profiles

Aravind Subramanian[1,9], Rajiv Narayan[1,9], Steven M. Corsello[1,2,3,9], David D. Peck[1], Ted E. Natoli[1], Xiaodong Lu[1], Joshua Gould[1], John F. Davis[1], Andrew A. Tubelli[1], Jacob K. Asiedu[1], David L. Lahr[1], Jodi E. Hirschman[1], Zihan Liu[1], Melanie Donahue[1], Bina Julian[1], Mariya Khan[1], David Wadden[1], Ian Smith[1], Daniel Lam[1], Arthur Liberzon[1], Courtney Toder[1], Mukta Bagul[1], Marek Orzechowski[1], Oana M. Enache[1], Federica Piccioni[1], Alice H. Berger[1,2,3,10], Alykhan Shamji[1], Angela N. Brooks[1,2,3,10], Anita Vrcic[1], Corey Flynn[1], Jacqueline Rosains[1,10], David Takeda[1,2,3], Desiree Davison[1], Justin Lamb[1,10], Kristin Ardlie[1], Larson Hogstrom[1], Nathanael S. Gray[1,3,4], Paul A. Clemons[1], Serena Silver[1], Xiaoyun Wu[1], Wen-Ning Zhao[1,3,5], Willis Read-Button[1,10], Xiaohua Wu[1], Stephen J. Haggarty[1,3,5], Lucienne V. Ronco[1,10], Jesse S. Boehm[1], Stuart L. Schreiber[1,6,7], John G. Doench[1], Joshua A. Bittker[1], David E. Root[1], Bang Wong[1], Todd R. Golub[1,3,7,8,*]

[1]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[3]Harvard Medical School, Boston, MA 02115, USA
[4]Department of Cancer Biology, Dana-Farber Cancer Institute, Bo
[5]Department of Neurology, Massachusetts General Hospital, Bos
[6]Department of Chemistry and Chemical Biology, Harvard Univer
[7]Howard Hughes Medical Institute, Chevy Chase, MD 20815, US
[8]Department of Pediatric Oncology, Dana-Farber Cancer Institute
[9]These authors contributed equally to this work
[10]Present address: University of Washington, Seattle, WA, USA (
CA, USA (A.N.B.); Seven Bridges Genomics, Cambridge, MA, US
USA (W.R., J.L.); Fulcrum Therapeutics, Cambridge, MA, USA (L

* Correspondence: golub@broadinstitute.org

https://www.biorxiv.org/content/biorxiv/early/2017/05/10/136168.full.pdf

SUMMARY

We previously piloted the concept of a Connectivity Map (CMap), whereby genes, drugs and disease states are connected by virtue of common gene-expression signatures. Here, we report more than a 1,000-fold scale-up of the CMap as part of the NIH LINCS Consortium, made possible by a new, low-cost, high throughput reduced representation expression profiling method that we term L1000. We show that L1000 is highly reproducible, comparable to RNA sequencing, and suitable for computational inference of the expression levels of 81% of non-measured transcripts. We further show that the expanded CMap can be used to discover mechanism of action of small molecules, functionally annotate genetic variants of disease genes, and inform clinical trials. The 1.3 million L1000 profiles described here, as well as tools for their analysis, are available at https://clue.io.

KEYWORDS

Functional genomics; gene expression profiling; chemical biology

HIGHLIGHTS

- A new gene expression profiling method, L1000, dramatically lowers cost
- The Connectivity Map database now includes 1.3 million publicly accessible L1000 perturbational profiles
- This expanded Connectivity Map facilitates discovery of small molecule mechanism of action and functional annotation of genetic variants
- The work establishes feasibility and utility of a truly comprehensive Connectivity Map

LINCS BROAD INSTITUTE (https://clue.io/)

CLUE

Tools Projects Partnering | Log in

ConnectivityMap

BROAD INSTITUTE

Unravel biology with the world's largest perturbation-driven gene expression dataset.

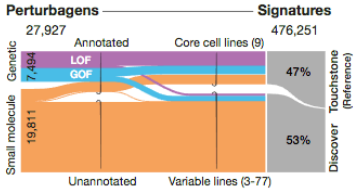
> TYPE COMPOUND, GENE, MoA, OR PERTURBAGEN CLASS TO SEE OVERVIEW
> TYPE A SLASH CHARACTER "/" TO SEE LIST OF COMMANDS

DATA VERSION: 1.1.1.1 / SOFTWARE VERSION: 1.1.1.19

Data and Tools

The CMap dataset of cellular signatures catalogs transcriptional responses of human cells to chemical and genetic perturbation. Here you can find the 1.3M L1000 profiles and the tools for their analysis.

A total of 27,927 perturbagens have been profiled to produce 476,251 expression signatures. About half of those signatures make up the Touchstone (reference) dataset generated from testing well-annotated genetic and small-molecular perturbagens in a core panel of cell lines. The remainder make up the Discover dataset, generated from profiling uncharacterized small molecules in a variable number of cell lines.



Start exploring the data by using the text-box on this page to look up perturbagens of interest in Touchstone. To see the suite of tools, including apps to query your gene expression signatures and analyze resulting connections, click on Tools in the menu bar.

BROAD INSTITUTE

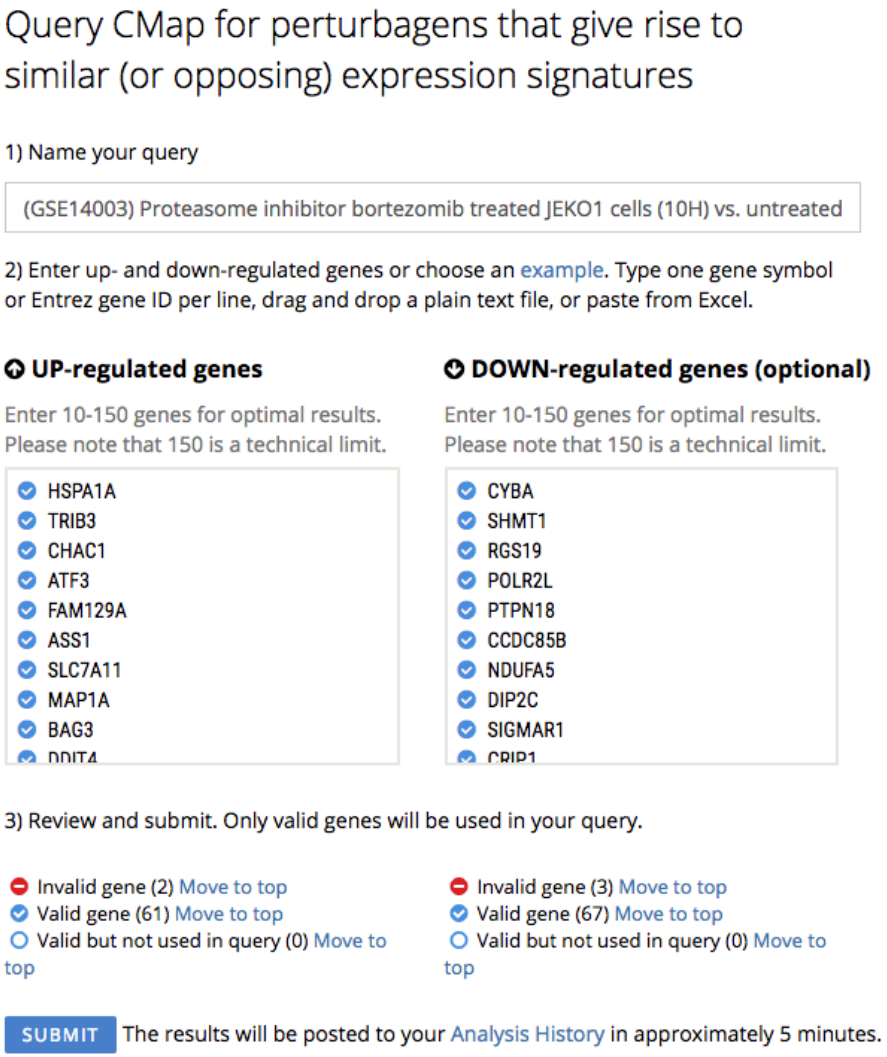
vis Skunkworks AT BROAD INSTITUTE

About
CMap
CLUE
Team
Careers
Acknowledgements

Support
Contact Us
Connectopedia: Clue Knowledge Base
Office hours

Branch: master Build Date: Tue Nov 07 2017 Revision: 2f2d6ecf326e9ef3762f6b5631630ec1bd79a34c Short: 2f2d6ec

LINCS BROAD INSTITUTE (https://clue.io)



LINCS BROAD INSTITUTE (https://clue.io/)

CLUE.IO has replaced lincscloud.org

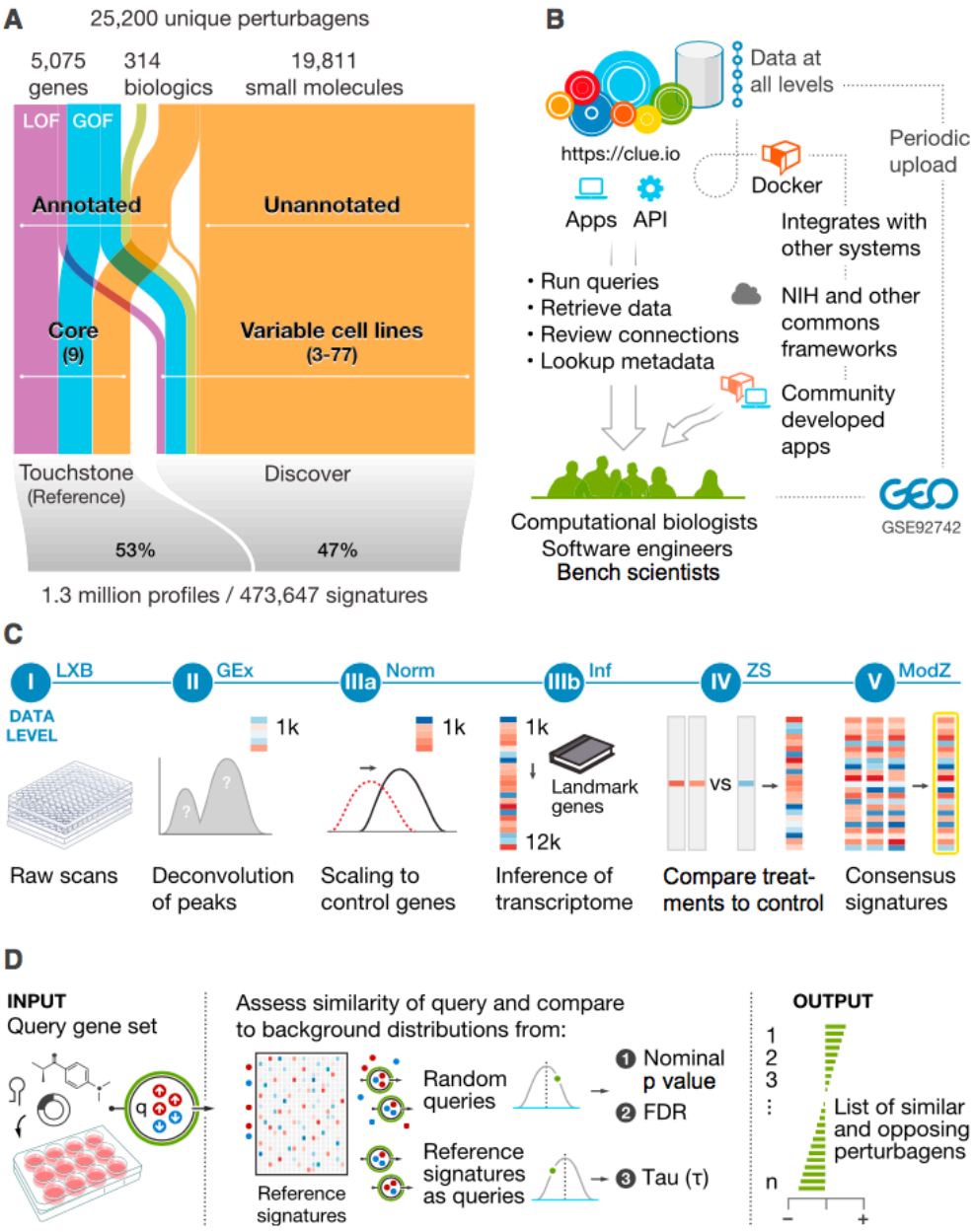
Question: Where can I download LINCS L1000 datasets?
For LINCS Phase I data see GEO GSE92742
For LINCS Phase II data see GEO GSE70138
For CMAP-HBS-LINCS contest data, see GEO GSE92743

Question: I'd rather not download all the data - do you have analysis tools?
Analysis tools from the Connectivity Map project are now available on the Broad Institute's CLUE software platform. Sign up at clue.io

Question: Is there an API? A new API is now available. See clue.io/api

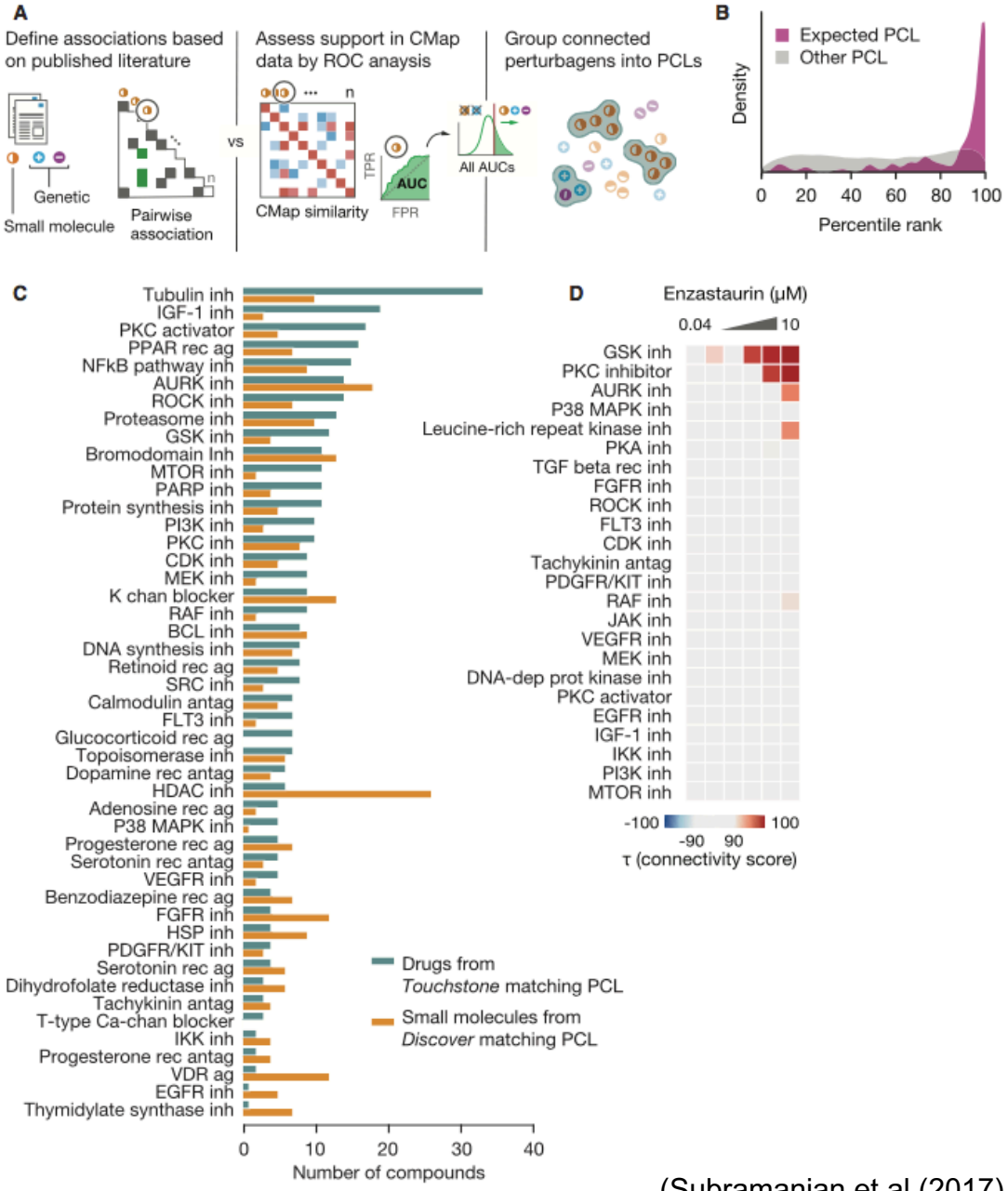
Please contact clue@broadinstitute.org with any questions.

Fig 2: L1000 Dataset Coverage, Signature Generation & Data Access



(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

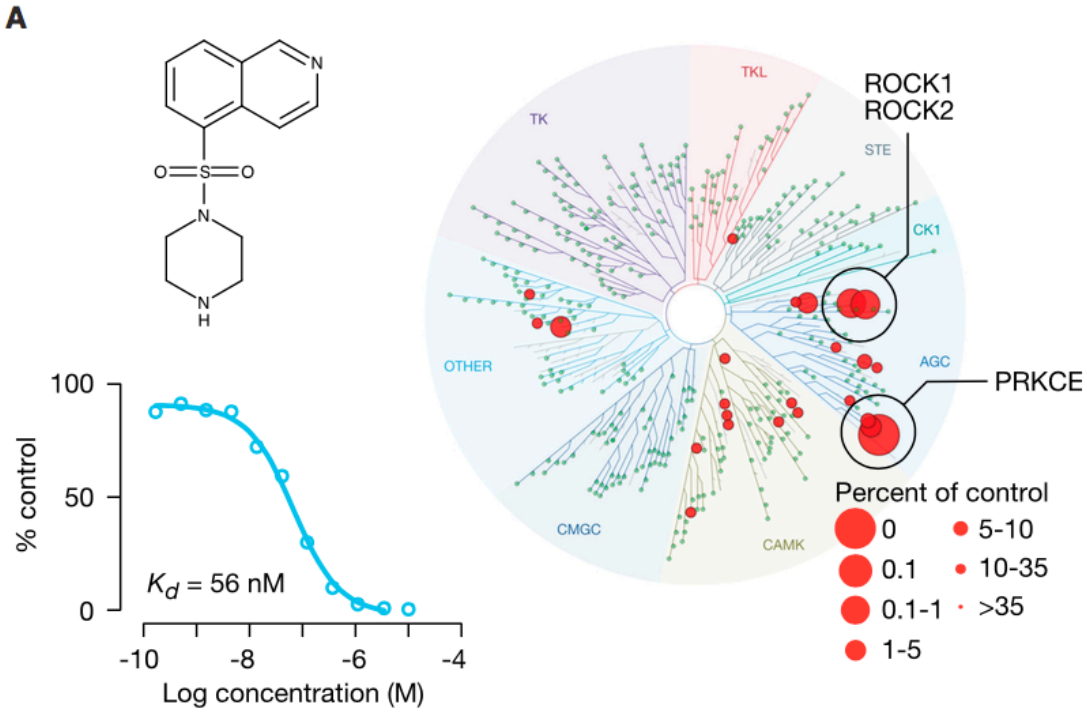
Fig 4. Reference Perturbagen Classes for CMap Discovery



(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

Fig 6A: Discovery of MOA

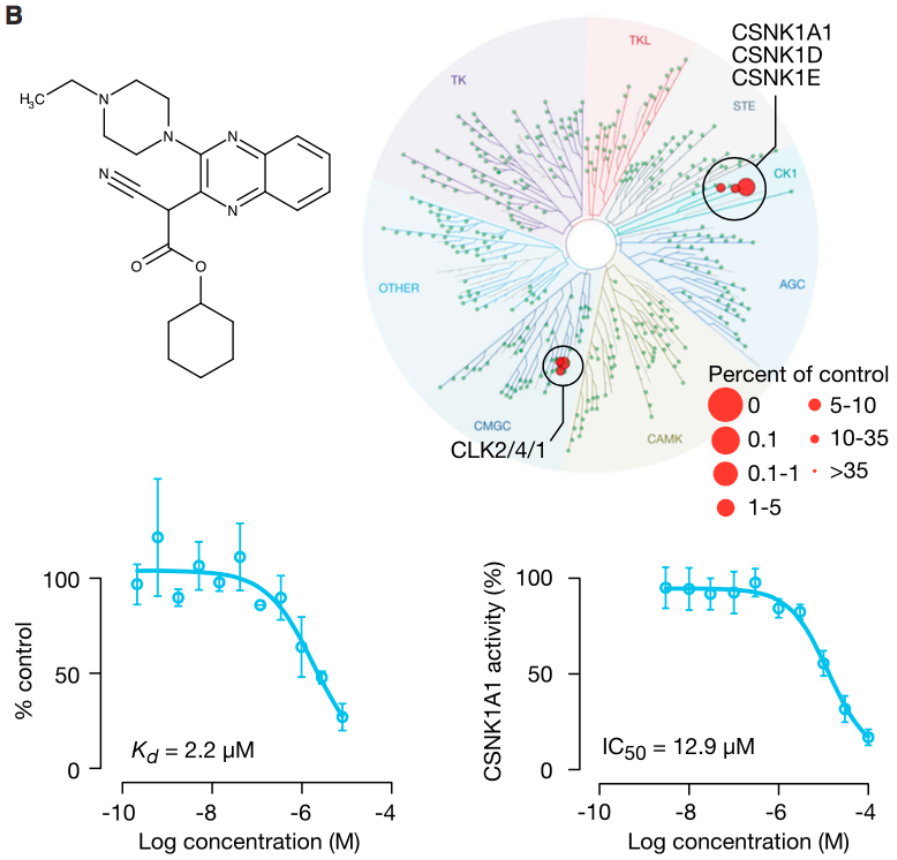
Query: Find compounds that induce the similarity gene expression profiles as query signature.
Unannotated compound BRD-2751 showed strong connectivity to the Rho-associated protein kinase (ROCK) PCL, suggesting that it might in fact be a ROCK inhibitor.



(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

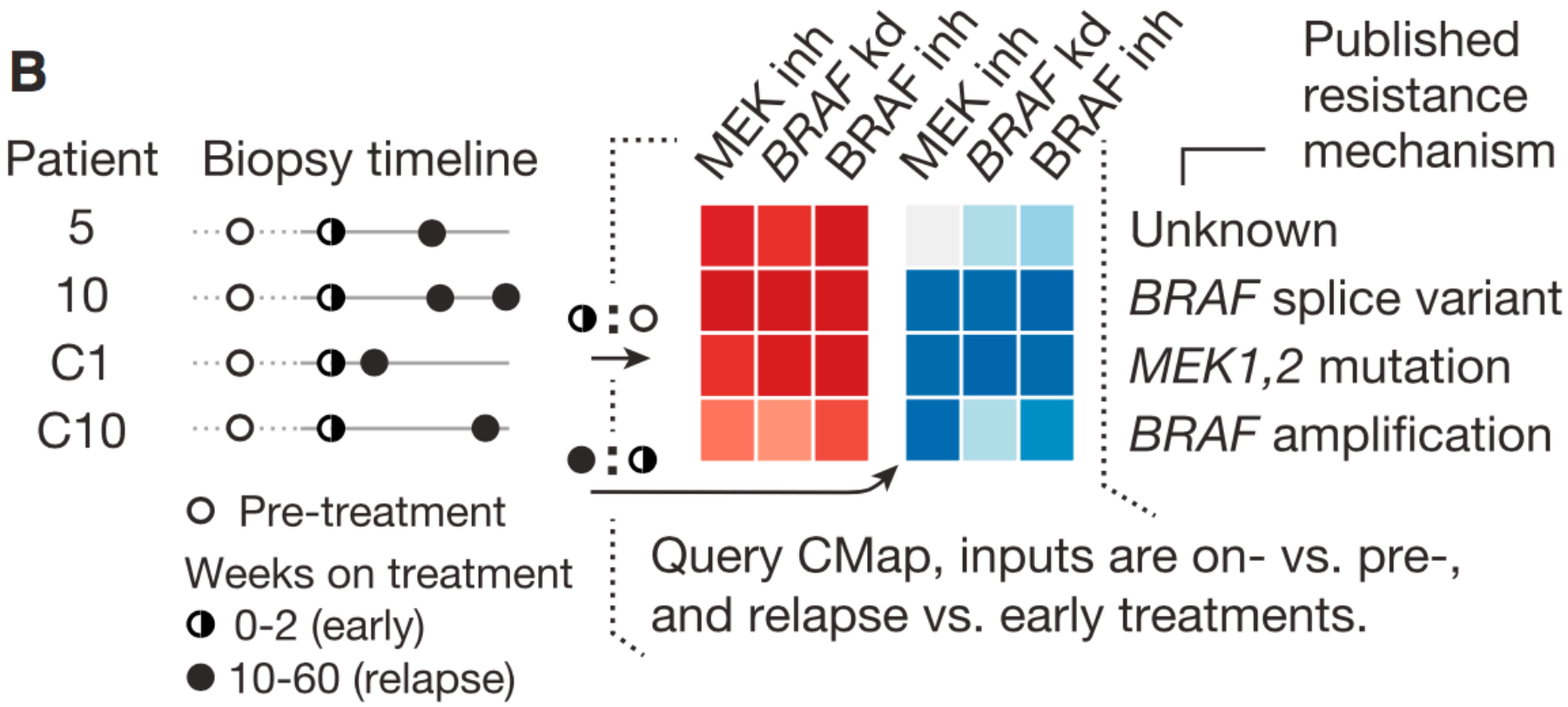
Fig 6B: Discovery of Selective Compound

Query: Find compounds that induce the similarity gene expression profiles as Loss of function (shRNAs CSKN1A1).
Results: One unannotated compound BRD-1868 showed strong connectivity to CSNK1A1 knock-down in two cell types.



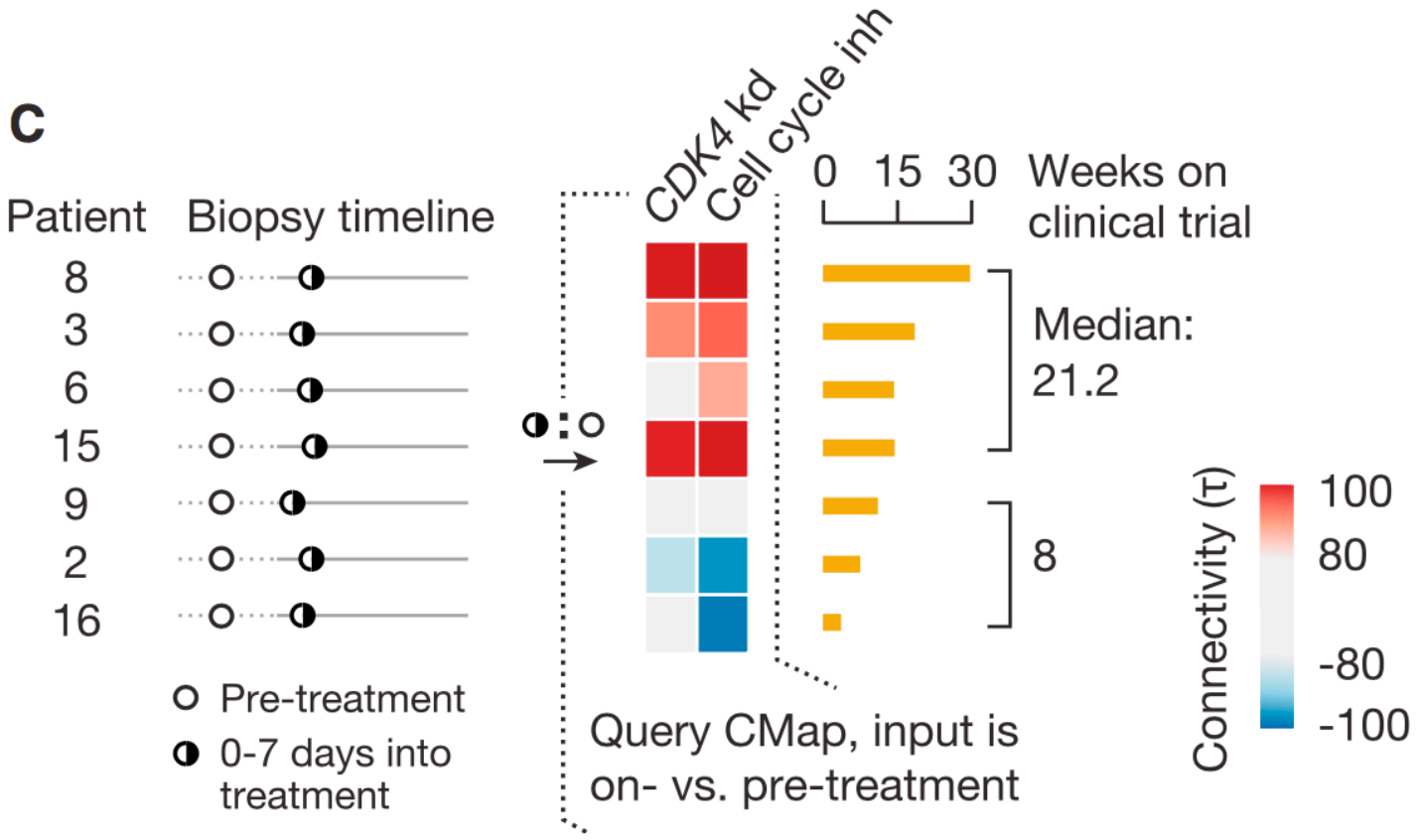
(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

Fig 7B: Connecting Patients data to explain resistance mechanisms



(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

Fig 7C: Connecting Patients data to predict therapeutic efficacy



(Subramanian et al (2017), ***Cell*** 171, 1437 – 1452)

Take Home Message

- Knowledge-based of pharmacogenomics, drug-gene interactions / relationships provide hypothesis-generation for developmental therapeutics research.
- Perturbagen-based resources capture rich phenotypic-gene interactions that could be further analyzed to reveal gene-drug interactions.
- Acquired resistance and escape pathways can be identified by chemical and functional genomics screens.
- From high-throughput screens, hits could be inhibited by rational combination of drugs.
- Gene expression changes could be used as the ‘universal language’ to connect between biological systems, genes, and drugs.
- Connectivity map (Cmap) concept provides an innovative approach to connect between biological systems, genes and drugs.