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

NCI Board of Scientific Advisors
March 1, 2011 Meeting

Daniela S. Gerhard, Ph.D.
Director, Office of Cancer Genomics

Large Projects Examples of NIH Investment in Genomic Research

Office of Cancer Genomics

- Therapeutically Applicable Research to Generate Effective Treatment (TARGET)
- The Cancer Genome Atlas (TCGA)
- Cancer Genome Anatomy Project/Cancer Genome Characterization Initiative (CGAP/CGCI)
- Genome-wide association studies (GWAS) of common and complex diseases and follow-up (~60/450 grants are cancer-related)

Data generated is made publicly available

- ~20% of NIH ARRA funded genomic projects

Molecular Characterization of Cancer is Essential but not Sufficient

- **Each tumor has hundreds to thousands genomic alterations**
  - Chromosomal changes: amplifications, deletions, translocations
  - Epigenetic changes
  - Mutations

- **Little is known about the cellular function of most genes, much less how sequence variants and mutations affect them**
  - Distinguishing initiating vs. driver vs. passenger mutations
    - Drivers are defined as genes involved in tumor maintenance
    - Evidence is accumulating that multiple subclones exist within a tumor and their frequency varies between patients
  - Genomic alterations result in cancer within specific context
    - Cell of origin
    - Other molecular alterations in genes that may have synergistic or antagonistic impact

CML and Gleevec: 40 Years From Discovery to Delivery

Office of Cancer Genomics

Empirical → Target-Driven

1960s

1960: Shortened chromosome 22 "Philadelphia" discovered

1970s

1973: t(9;22)(q34;q11) defined

1980s

1984: Breakpoint region cloned; fused *BCR* and *ABL* genes

1985: BCR-ABL tyrosine kinase activity described

1990s

1993: Preclinical testing on STI571 begins

1998: Phase I clinical trials begin

2000s

2001: Gleevec® approved by the FDA for treating patients who failed IFN therapy

2002: Approved for all stages of CML, GIST

Clinical information

- CML treated with Busulfan (inhibits DNA replication)
- Cytogenetics becomes a standard diagnostic
- Hydroxyurea treatment gains popularity
- Stem cell transplants pioneered
- PCR and FISH gain popularity in diagnosing CML
- Stem cell transplants become standard
- Treatment with Interferon introduced
- Interferon treatment correlated with positive prognosis
- Development of ST1571 (Gleevec) to target BCL-ABL kinase begins


cancer
genomics
cancer
therapeutics

ARRA Opportunity

Question:

Can a network be formed that would effectively address a current major scientific challenge: efficient transition from patient-based large multi-dimensional genomic data → target validation → small molecule modulators → (therapy, not part of the initiative)

How to advantage the flood of genomic data and accelerate the transition to treatments of patients based on the genomic profile of their cancer?

ARRA Request for Application

- **To utilize the molecular data to accelerate translation into the clinic:**
  - Mine genomic data sets with new approaches to identify targets in context of pathways
  - Innovate models to qualify and validate new targets that optimize both the biological context of the potential target with relevance to the clinic
  - New approaches to chemical genomics and compound synthesis
- **The Network members would define collaborations to take advantage of the strength of each component Center**
- **Share knowledge, experience and results with the research community**

A year and 6 months later the answer to the question:

Yes, a network was formed that is an innovative, efficient and highly-collaborative

ARRA Cancer Target Discovery and Development (CTD^2) Network Centers

Office of Cancer Genomics

Selected by review and complementarities of functions

- **Broad Institute, Cambridge, Massachusetts**
  PI: Stuart Schreiber, Ph.D.

- **Cold Spring Harbor Laboratory, Long Island, New York**
  PI: Scott Powers, Ph.D., co-PI: Scott Lowe, Ph.D.

- **Columbia University, New York, New York**
  PI: Andrea Califano, Ph.D.

- **Dana-Farber Cancer Institute, Boston, Massachusetts**
  PIs: William Hahn, M.D., Ph.D., L. Chin, M.D. and R. DePinho, M.D.

- **University of Texas Southwestern Medical Center, Dallas, Texas**
  PI: Michael Roth, Ph.D., co-PIs: M. White, Ph.D., J. Minna, M.D.

http://ocg.cancer.gov/programs/ctdd.asp

CTD^2: A Bridge from Genomics to Therapeutics

ARRA CTD^2 Network

- **Brought together highly motivated and outstanding investigators who do cutting edge science**
  - Each application included up to 3 mature projects and while the Centers made impressive progress on those, they will not be discussed here
  - Wrote a manuscript in which outlined the vision of the science (Appendix A)
- **Functional network formed rapidly**
  - Meet monthly via teleconference and once a year in person
  - Component centers share results “in real time” (pre-competitive)
- **Established an ethos of data and resource sharing with scientific community upon validation**
  - IT WG developed file formats for data sharing compatible with Cancer Data Standards Registry and Repository (caDSR) within caBIG
- **Enabled experiments, using new data generated by the molecular characterization projects to identify candidate targets, small molecule modulators and mechanisms: example TCGA’s caOv**

One Example of an Ongoing ARRA CTD^2 Collaborative Project

Office of Cancer Genomics

- **New targets in ovarian cancer discovered within a very short period of time**
- **A subset is already validated and more are in the pipeline**

- **The rapid progress was enabled by the Network Centers:**
  - Using experimental approaches that are complementary and
  - Sharing of results in real time

Ovarian Cancer (caOv) Background: TCGA and Other

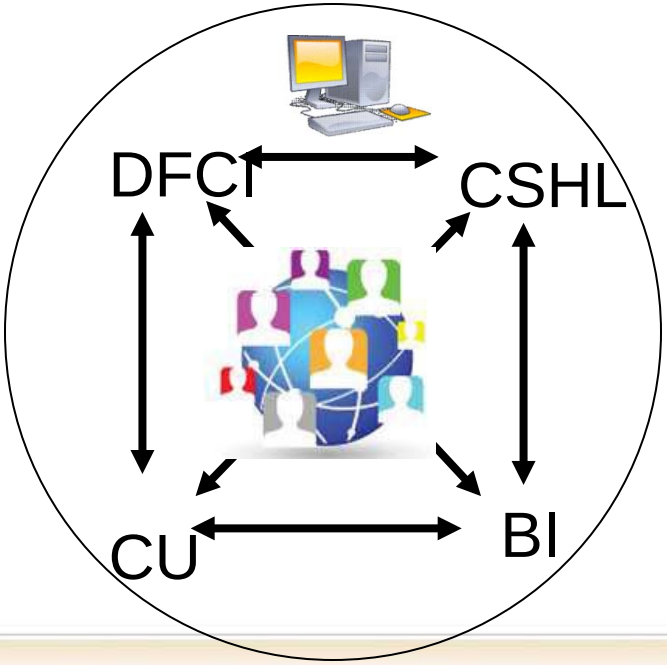
- **Tothill et al., (2008) generated expression profiles of ~200 caOv cases**
  - Identified 6 expression subgroups, 4 of which are specific for high grade ovarian cancer
- **The results of TCGA's molecular characterization of caOv* data:**
  - Confirmed the 4 expression subgroups of high-grade ovarian cancer
  - The tumors are highly rearranged and aneuploid
    - **1200 expressed genes map within the amplified regions**
  - Confirmed the importance of p53 and BRCA1/2 genes
    - At least 95% of tumors have 1 or 2 p53 hits (mutation or deletion)
    - About 30% of tumors have a BRCA1/2 hit (somatic or germline mutation, deletion or epigenetic silencing)
    - The p53/BRCA1 loss of function probably explain the high level of chromosomal rearrangement
  - Did not find many genes mutated along their length at high frequency (>5%)
  - Identified involvement of FOXM1 (~80%), RB (~60%), RAS/PI3K (~40%) and NOTCH (~20%) pathways
  - TCGA manuscript will be published shortly in Nature

*** ~500 caOv cases Agilent expression chips, Affymetrix SNP 6.0, Illumina 27K methylation chip; exome sequencing of ~310 cases. The tissue passed stringent criteria of pathology, tumor cellularity, and nucleic acid integrity**

Example of ARRA CTD^2 Network Collaborations: caOv

Only one Center included a specific aim on caOv in their ARRA application, the rest started working on the problem after discussions within Network which defined the opportunity to make an impact.

On the next 4 slides, the contributions of each Center are summarized



ARRA CTD^2 Network's Discovery and Confirmation of Genes Important in caOv

OFFICE of CANCER GENOMICS

Experiments at DFCI

- ID4, PAX8, ERBB2 & 3, KRAS & ~60 other genes were identified as essential for proliferation and survival of a subset of 25 ovarian cell lines in a shRNA screen

  ID4 induces the expression of the oncogenic NUP98-HOXA9 gene set

  - **Down regulation of ID4 resulted in decrease of HOXA9 expression signature in cell lines, and decreased tumor growth in mice**
  - PAX8 is the most differentially expressed gene between ovarian and non-ovarian cell lines
  - PAX8 shRNAs reduce the viability of ovarian cancer cells which in PAX8 is either amplified or over-expressed
- Analysis of TCGA data by GISTIC
  - ID4 is amplified in ~32% of ovarian cancer cases
  - PAX8 is amplified in ~16% of cases
  - ID4 regulated genes are found over-expressed in ovarian cancer tissues
- **Synthesized ID4 in vitro and showed it binds E proteins**

ARRA CTD^2 Network's Discoveries and Development of Genes Important in caOv

OFFICE of CANCER GENOMICS

Experiments at CSHL

- Analysis of TCGA CNA data by FOCAL (different algorithm from GISTIC) identified a few hundred candidate cancer genes, some overlap with DFCI's list, including PAX8, ERBB2, CDK2
- **The BR5 mouse ovarian surface epithelial cells (p53-/-, Brca1-/-), are used to confirm the transformation capacity of the candidate genes upon transfection and injection into nude mice**
  - **To date PAX8, ERBB2, FOXM1 and 33 other genes are confirmed**
- The Reactome analysis suite, with "hand-curated" annotation is used to identify pathways necessary for ovarian transformation from the exome mutation data set
- "Speedy" mouse model(s) are planned once the Centers complete their target identification phase—
  - The selection of the genotype(s) will made after a discussion with the Network members

ARRA CTD^2 Network's Discoveries and Development of Genes Important in caOv

Analysis at CU

- Expression data was analyzed by ARACNe to identify the protein interaction network and their "master" regulators
  - Identified a signature of >200 genes differentially expressed between patients with best and worse prognosis
    - The poor prognosis signature was enriched for mesenchymal genes
    - Using MARINa, algorithm for identifying transcription factors that are "master regulators" of the worse prognosis group discovered STATs, FOXM1 and C/EBPs
  - **Identified candidate signature for patients that recur within 6 months of platinum or platinum/taxane treatment**
  - **Identified about 30 genes which are candidates for follow-up as ovarian oncogenes**
  - About 50% of the genes from the various analyses overlap with the sets identified by CSHL & DFCI

ARRA CTD^2 Network's Discoveries and Development of Genes Important in caOv

OFFICE of CANCER GENOMICS

➢ HTS at BI

- **Increased the number of compounds with unique characteristics, such as stereochemical structure and the ability to modify side groups by a simple chemical reaction**
- Improved screening of small molecule microarrays (SMM) that has been shown to allow to identify compounds which bind proteins, such as transcription factors (TFs)
  - Developed Luminex assay to detect expression changes of ~1000 TFs regulated genes
- **SMM screen identified STAT3 binding proteins**
  - **Assays under way at CU to confirm specificity of compound function as well as a counter screen**
  - Functional ID4 prepared last month by DFCI, it is in the queue to be screened by SMM
  - In vitro synthesized C/EBPβ protein is being tested for function and when confirmed will be screened for small molecule modulators
- Improvements in high-content, high-throughput automated cell-based assay small molecule screens

Summary of the ARRA CTD^2 Network caOv Results

The power of the network: made rapid progress by sharing data, working together and taking advantage of complementary, non-overlapping expertise to carry out the experiments. Each Center contributed to the results :

- Identified candidate signature to stratify patients into best and worst prognostic groups
- Identified candidate targets for therapeutic development
  - Confirmed a subset of candidates by in vitro and ex vivo experiments
- Identified candidate small molecules for a subset of confirmed targets
- Plan to generate mouse models for in vivo screening of other candidate genes within a specific genetic context
- These results will be utilized by the scientific community
  - Thereby saving a lot of people a lot of work
- Experiments are ongoing

Critical lesson: collaborative efforts to integrate several methods can yield exponential gains relative to the incremental gains achieved through improving any single method (united they are more than a sum of parts)

ARRA CTD^2 Network's Approaches: caOv Example

OFFICE of CANCER GENOMICS

Streamlines the development an efficient process in which the preclinical discovery phase is directed by patients' cancer genetic profile

- **Each Center had a unique contribution to the results by performing experiments in:**
  - Integrated genomics systems biology
  - High-throughput shRNA screening in cancer cell lines identified genes important in subsets of ovarian cancers
  - Expression of genes within a given context (ex vivo) identified genes which are essential in tumor progression
  - Development of mouse model(s) – initial stage(s)
  - Identification of small molecule modulators for the dependencies modulated by transcription factors (not classical drug targets)
  - Provides leads or the discovery of associated biomarkers

What are Among the Successes of the ARRA CTD^2 Network?

- **Developed a process in which to translate newly generated genomic data immediately into series of experiments resulting in new and validated targets, small molecules that modulate them as well as identification context signatures**
  - Collaborations were established “on the fly” as the status of the genomics datasets warranted—the caOv is one of many
    - The concept document includes other examples of collaborations
    - New collaborations were developed a month ago at the SC meeting
- **Improvements in methodologies were rapidly implemented in each Center**
  - Enhanced data quality/interpretability
  - Immediate positive impact on CTD^2 projects
  - Obtained economy of scale

Rationale for a New CTD^2 Network Initiative

- **Build on the success of the ARRA pilot and utilize the lessons learned to address issues vital to the integral mission of NCI by nimbly responding to science opportunities as the genomic data for cancers is generated by the large scale projects**
  - Follow-up targets for validation which are “non-traditional” including those that function through protein-protein interactions, transcriptome factors and others
  - Combine systematic genetic, chemical and bioinformatic approaches
  - Improve the process to define combination of targets for therapy
  - Improve the process to identify synthetic lethals, i.e. effect of a molecule within context of another mutation; will result in improved specificities of treatment and reduced side effects
  - Adaptation of methods improvements as they are developed within any one of the Centers
  - The Center infrastructure results in speedy generation of results, economy of scale and cost efficiency not easily possible in most small laboratories

Rationale for a New CTD^2 Network Initiative - Continued

- **The pre-competitive collaborations, the development and utilization of novel, cost-effective methods, public data and resource sharing upon validation**
- **Encourage hand-offs and interactions, examples**
  - Potential to feed the pipeline at DCTD—NeXT
  - Preclinical testing through NCI's CAPER
- **Fills a niche not represented by other programs (portfolio analysis)**

Goals for the New Network

- **Accelerate the translation of patient genomic data into clinical application**
  - Innovate the integration of computational mining large scale genomic data analyses
    - Make tools available through web
  - Identify and confirm new therapeutic target candidates
  - Identify and confirm novel modulators within specific cancer context (cellular or mutational) in vitro (cell lines) or in vivo (cancer models)
    - Small, stereochemically “interesting“ molecules
      - Use of novel organist chemistry – molecules more “natural products-like”
      - Mature molecules: optimize activity, structure activity relationship, systematic variation of stereochemistry
    - siRNAs
  - Multi-expertise team
  - Share models and reagents with the scientific community
  - Share data and methods with the scientific community through the web

- **As genomic data become available from TARGET, TCGA etc.,: be nimble, flexible and open to new opportunities**

RFA/U24: Cons and Pros

- In a period of strained budget, it "takes" funds from R01 grants
- The research would get done without it – through R01 or P01 research
- The rich get richer – many of the labs working in this space are already well funded

- The Institute participation ensures that the genomic data sets that will be generated in the next few years for many cancers will be candidates for systematic target discovery and development
- The U24 provides a process to share data and reagents within and with the rest of the scientific community
- Pre-competitive collaborations accelerates the generation of results
  - The total is greater than sum of its parts
- Allows the cost-efficiency of scale

Mechanism and Cost

- **Mechanism**
  - U24 Cooperative Agreement Grants
    - Critical for pre-competitive collaborations
    - Essential for communication
    - Important for governance; e.g. allows for inclusion of an external scientific group
- **Open competition**
  - **No presumption of current Centers**
  - **Will be reviewed by a Special Emphasis Panel convened by DEA**
  - **Establish the best network possible from proposed grants**
- **Budget: $10M for year 1: 2 options**
  - Fund up to 8 Centers
    - Concentrates the research in a few highly functional Centers and promotes cost-efficiency

    OR
  - Fund up to 16 Centers
    - Provides an opportunity for smaller labs to participate and build their expertise

Program Evaluation Criteria: Examples

- **The number and quality of publications**
- **Number of validated probes and/or targets**
- **Impact of the program on the biomedical research community, such as:**
  - How many times were published manuscripts cited
  - Are the results, methods, tools etc. developed by the Centers used in academia and industry
  - Frequency of data portal visits and data downloads
  - How do the results influence the number of proposals received at the NIH as following up of CTD^2 findings
- **Were the results of the projects transitioned into preclinical testing**
- **Other appropriate specific evaluation parameters will be determined once the projects are defined**

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

NCI Board of Scientific Advisors
March 1, 2011

Daniela S. Gerhard, Ph.D.
Director, Office of Cancer Genomics

CTD^2 Centers Other Results: Examples

- Identified small molecule modulators of IDH1 R132H
- Identified small molecule modulators of STAT3 (important TF of mesenchymal subtypes of cancer) which are being test in cell-based assays
- Identified proteasome-degradation regulator of STAT3 and C/EBPs
- Identified mechanism of glucocorticoid resistance treatment in T-ALL
- Identified candidate oncogenes that cause transformation in GBM (collaboration ongoing w/ others within the Network)
- Mouse model of AML and understanding of myb oncogene addition
- Identified small molecules which selectively affect growth of cell lines that have activated KRAS mutation and mutated STK11
- Identified a small molecule that increases ROS in cancer, but not normal cells
- Identified candidate “master regulators” of TKs for a set of NSCLC cell lines and candidate addiction points

CTD^2 Network Research Mission

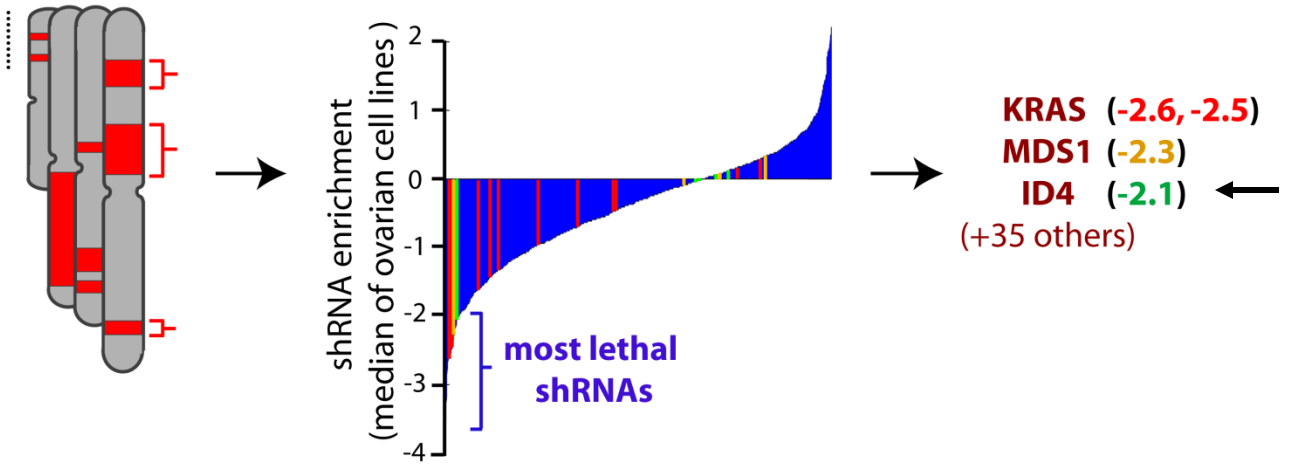
- Shift current research paradigms in translation pathway of patient-derived multidimensional genetic data to the clinic and utilize novel concepts, approaches and methodologies
- Accelerate the translation of these data to the patients benefit
- Innovate in all areas and adopt new technologies as is scientifically warranted
- Develop research that will exert a sustained influence on the field
- Develop a pre-competitive culture to ensure sharing of data, methods (analytical, experimental) and reagents within the network and the scientific community at large

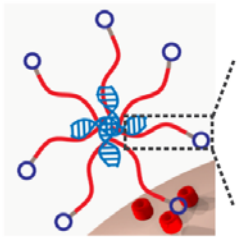
Additional Example: CTD² Discovery of Novel Drug Candidates for Ovarian Cancer

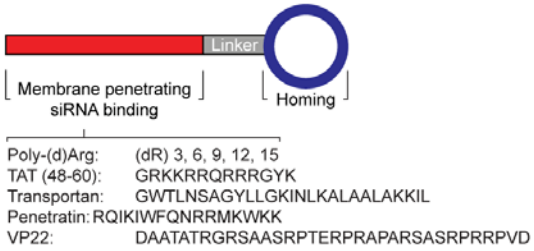
270 amplified genes in ovarian tumors

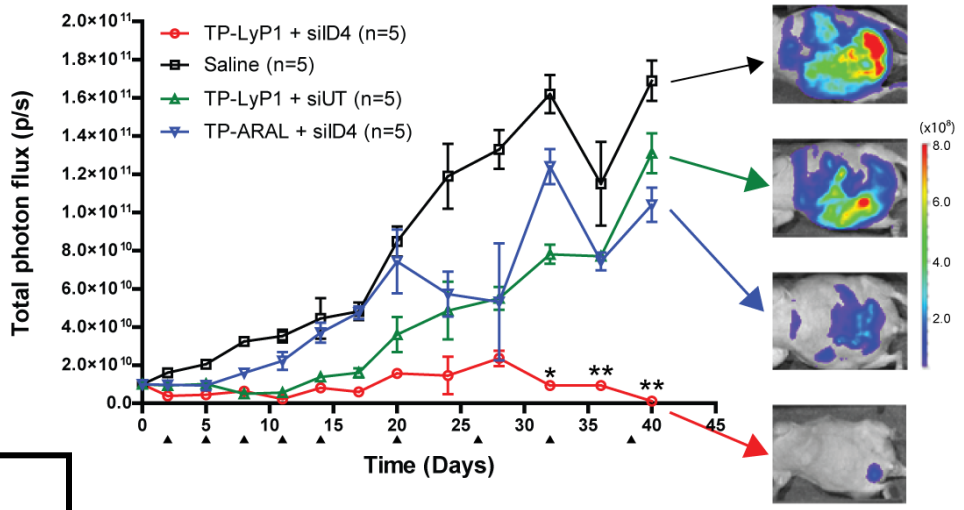
1350 shRNAs targeting amplified genes in ovarian cell lines

Essential and amplified genes





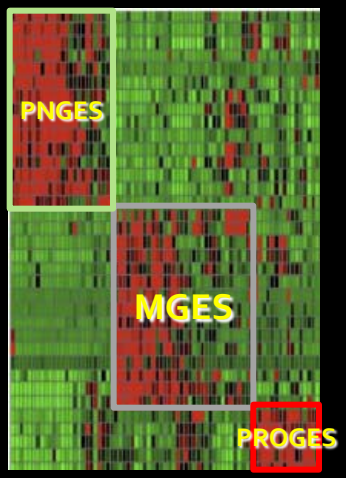




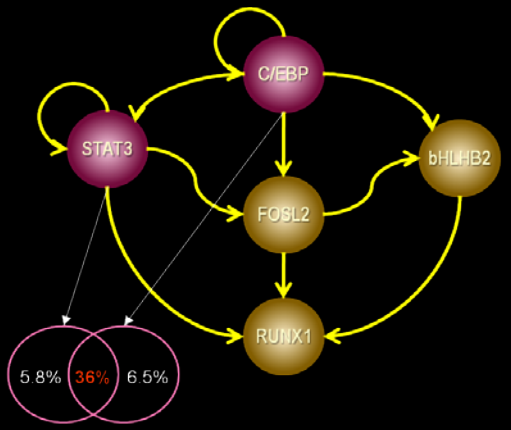
Status:
DFCI developed a construct to synthesize the protein and small molecular array screen is under way.

W. Hahn et al.,

Example: CTD^2 Discovery of Small Molecules which Bind to Transcription Factors



STARTING POINT



C/EBP and STAT3 are synergistic master regulators of the MEGS (Mesenchymal Signature) of GBM

Carro MS et al. (2010) Nature 2010 Jan 21;463(7279):318-25

STAT3 and C/EBPβ or δ are novel targets

- ARACNe analysis of GBM expression data → identified a signature of a mesenchymal subtype of tumors with poor prognosis
- Ectopic expression of C/EBPβ and STAT3C in mouse neural stem cells caused the cells to express mesenchymal genes
- Transfecting shRNAs targeting STAT3, C/EBPβ into a human GBM xenograph line generated from a mesenchymal subtype tumor resulted in significant decrease of invading cells

A. Califano et al.,

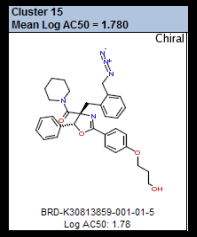
.

Example: CTD^2 Discovery of small Molecules which Bind to Transcription Factors

Collaboration with BI:

Status: STAT3

- SMM assay and screening identified ~30 compounds
  - The protein used binds DNA
  - Structure of one compound:



- The compounds are being tested in a cell-based biological assay (luciferase expression from a STAT3-promoter construct)
- Those compounds that inhibit biological function will be analyzed further:
  - Expression profile of cells exposed to the compounds
  - Determination of binding constants

Status:

- C/EBPβ protein to be obtained
- Small molecule array (SMM) screen will be done

A. Califano et al.,