Pilot Program: Developing Reagents, Protocols, and Tools to Enable Translation of Discoveries from Large Scale Cancer Genomics Programs (TCGA, TARGET, Others)

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

Henry Rodriguez, Ph.D., MBA
Clinical Proteomic Technologies for Cancer

Anna D. Barker, Ph.D.
Deputy Director
National Cancer Institute

NATIONAL INSTITUTES OF HEALTH

NATIONAL CANCER INSTITUTE

NCI Board of Scientific Advisors
June 22, 2009

Background for the Pilot

- Large scale genome characterization/sequencing programs are well under way (TCGA, TARGET, other whole genome sequencing programs) – creating large complex data sets for mining by the communities

- The programs are already identifying new genomic alterations – but perhaps of more importance – are providing unprecedented opportunities to analyze the multi-dimensional data for new potential cancer "signatures/targets" While some high-quality reagents exist for highly studied targets – the next era of cancer discovery and development will depend in large measure on the success of functional studies

- Looking ahead, the numbers of potential "targets" will be large – but reagents and tools will be limited, costly and difficult to access for many

- There is a critical need for highly-characterized reagents, protocols and other tools to support cancer researchers as they move to functionalize biologically and clinically important genomic alterations

Overall Goals of the Pilot

- To develop a community-based process that will prioritize "signature/target" candidates **(targets)** from large-scale genomic programs

- To enable and accelerate functional studies through the development of broadly available highly-characterized reagents, protocols and tools to the selected targets

- To regularly evaluate the pilot – determine a longer term model for the future

TCGA = depth + data integration...
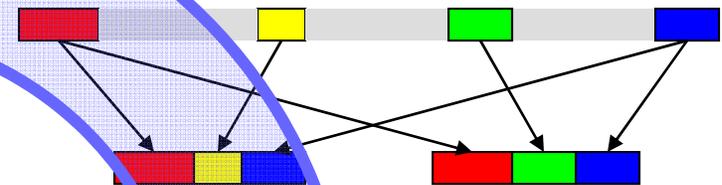
THE CANCER GENOME ATLAS
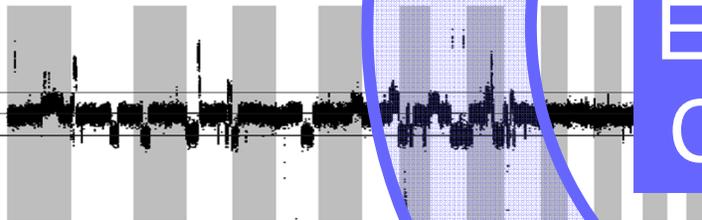
From Ron DePinho
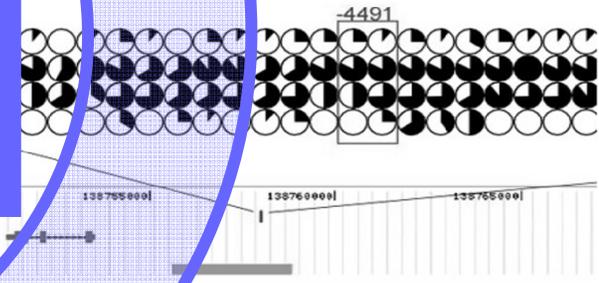
Aneuploidy; Re-arrangement; Translocation
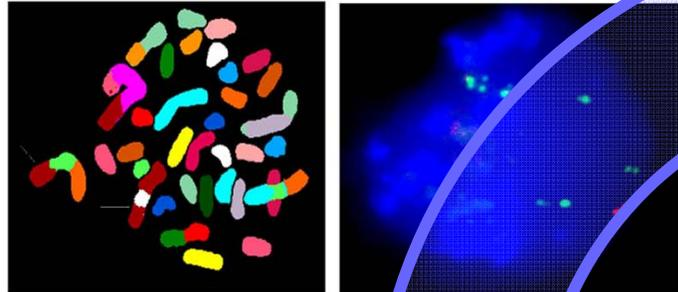
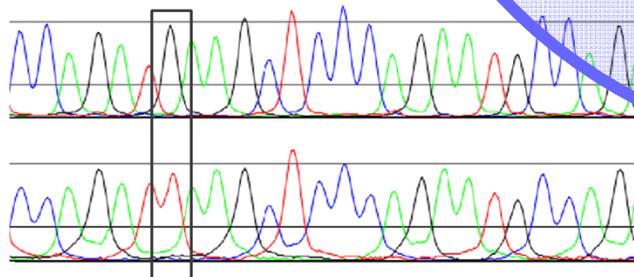**Gene Splicing Alterations**

Copy number aberrations
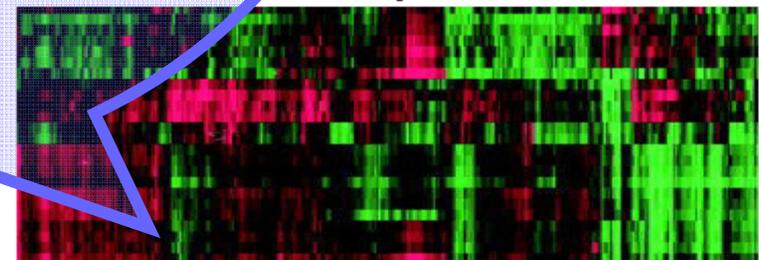
Methylation or histone modification

Each Sample Clinical Data

Somatic mutations

Altered expression
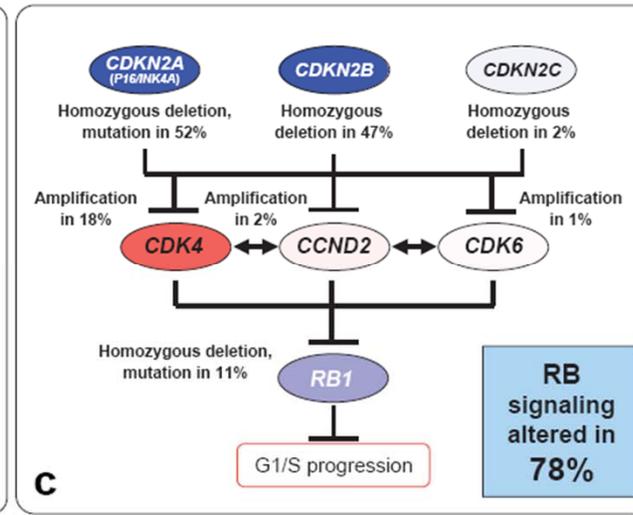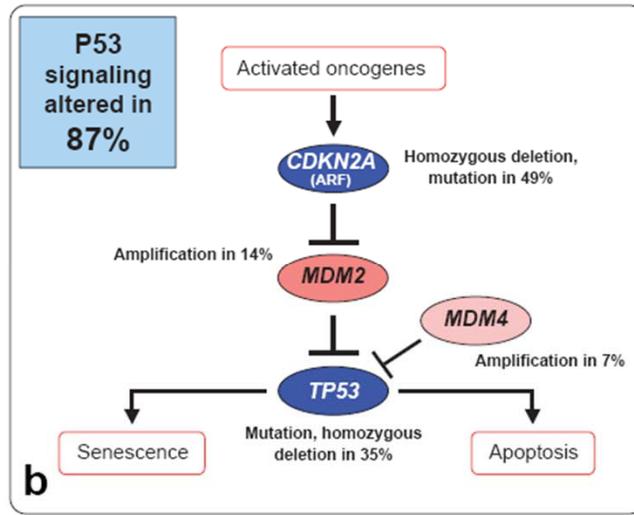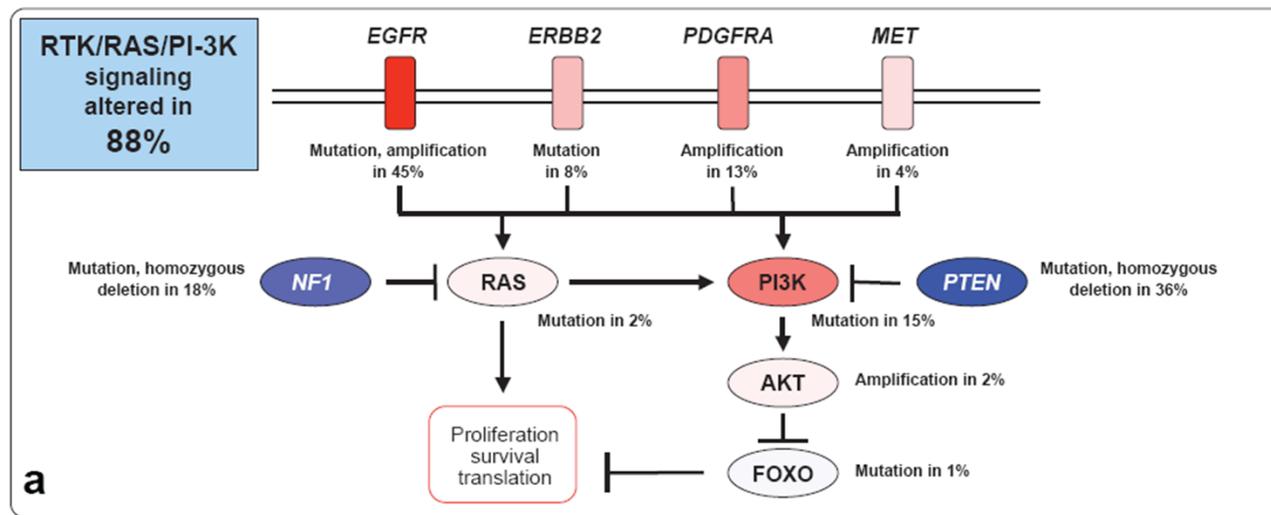
Adopted from Cameron Brennan

GBM Pathways (Characterization/Sequencing Continuing)

TCGA: *Nature* 2008

Frequency of Mutations in Some Genes – Likely Specific for GBM Subtypes

Table 3a. Distribution of frequently-mutated genes across GBM subtypes.

Gene	Proneural n=37	Neural n=19	Classical n=22	Mesenchymal n=38	Total # Mut	χ^2	p-value
TP53	20	4	0	12	36	14.1	0.003 ←
PTEN	6	4	5	12	27	2.5	0.47
NF1	2	3	1	14	20	13.3	0.004 ←
EGFR	6	5	7	2	20	6.8	0.078
PIK3R1	7	2	1	0	10	8.5	0.035
RB1	1	1	0	5	7	5.7	0.127
ERBB2	2	3	1	1	7	3.9	0.272
EGFRvIII	1	0	5	1	7	10.8	0.013
PIK3CA	3	1	1	1	6	1.1	0.789
PDGFRA	4	0	0	0	4	8.5	0.037

Verhaak et al., submitted

# Relevant New Discoveries in Pediatric Cancers -TARGET

JAK mutations in "BCR-ABL1-like" ALL (TARGET)

- High risk childhood ALL - 67% cases have lesions in B-cell development pathway genes; IKZF1 (IKAROS) alterations in ~ 30% of cases

  - JAK2 (n=16): 10 R683G; 3 non-R683G pseudokinase domain; 3 kinase domain
  - JAK1 (n=3): 3 pseudokinase domain
  - JAK3 (n=1): uncertain functional consequences

Genomic Approaches/Reagents Exist for Historical Targets – New Insights Require New Reagents

- *PIK3R1* had anecdotal reports in the literature
- Cluster of mutations/indels in contact amino acids known to be important for interaction with *PIK3CA*
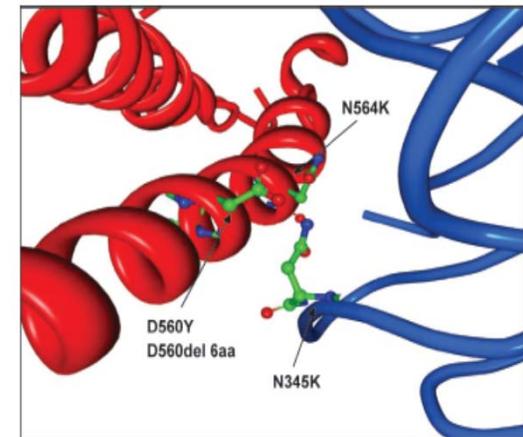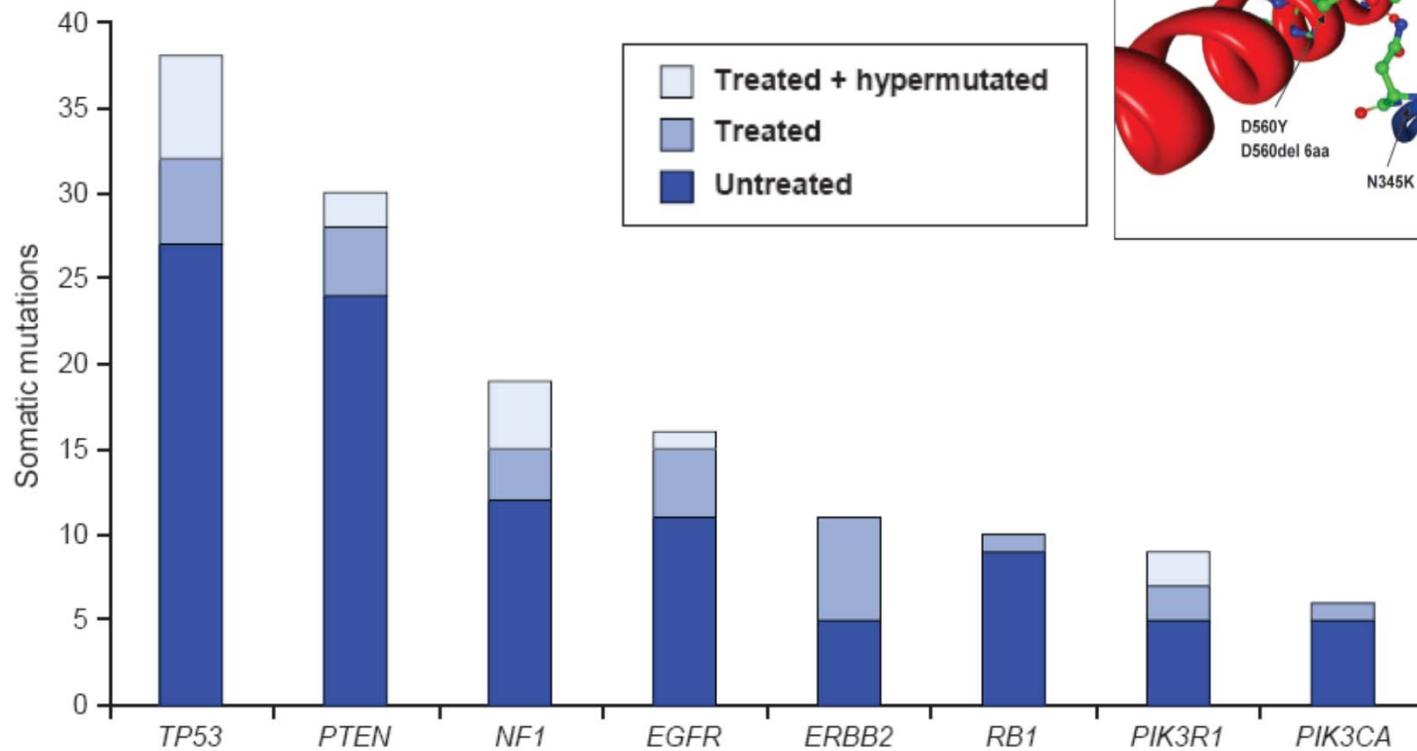
Need to Expand Successful Mutation-Specific Antibody Reagent Models

- **85% to 90% of NSCLC-associated EGFR mutations are:**

  - **In-frame deletions in exon 19 (E746_A750del) or point mutation in exon 21 (L858R)**

- **Monoclonal antibodies against synthetic peptides matching above aberrations recently developed and shown effective in IHC assays**

- **Such unique reagents are critical as patients with similar mutations shown responsive to EGFR inhibitors including gefitinib and erlotinib.**

Control EGFR mAb   L858R specific mAb   dEGFR specific mAb
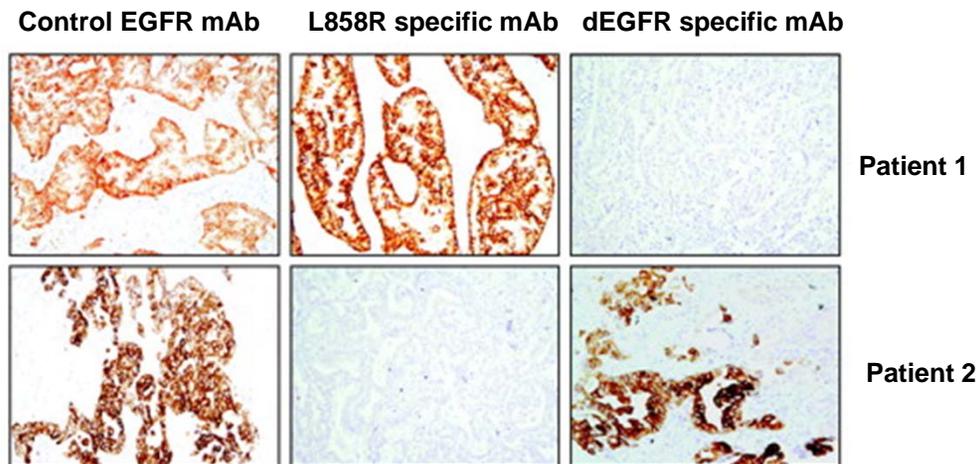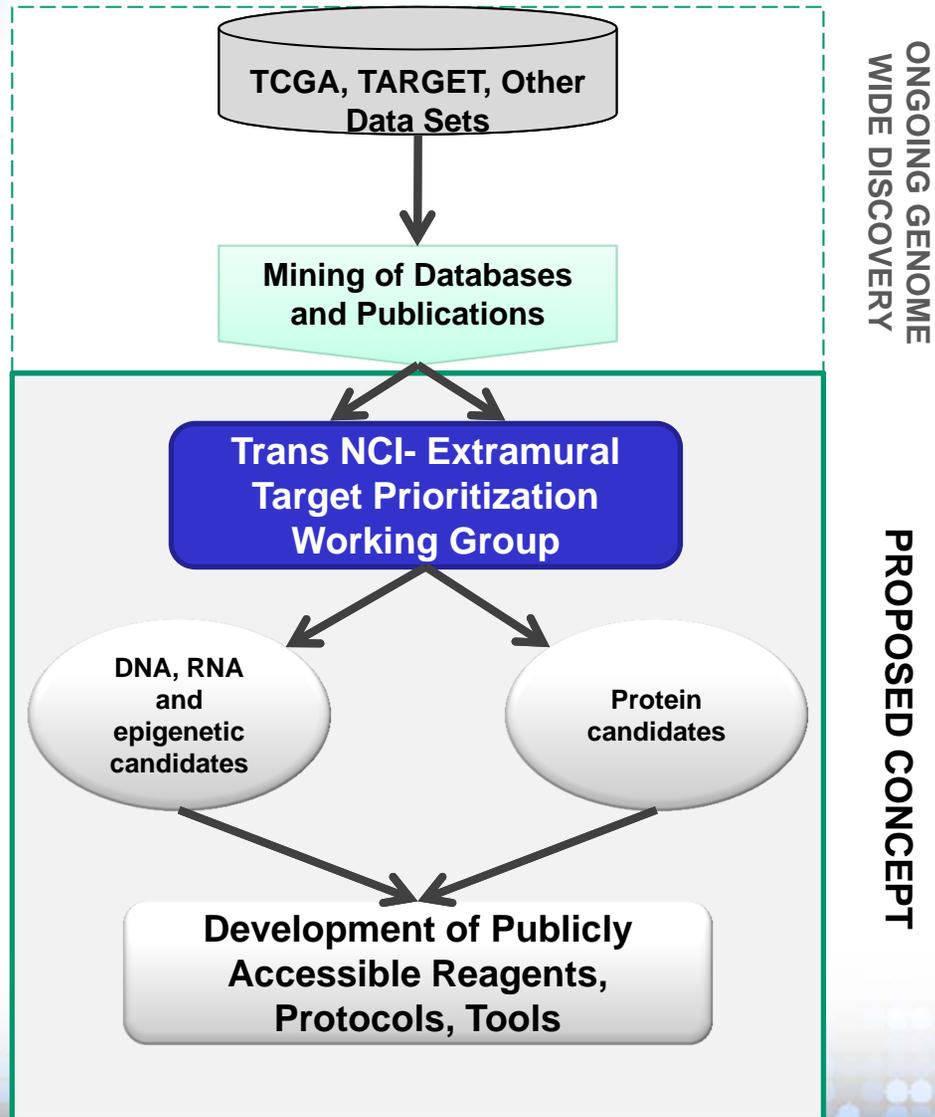

Patient 1

Patient 2

Mutation-specific EGFR IHC assays using anti-peptide monoclonal antibodies

- NSCLC patient samples with unknown genotype were stained with above mAbs.
- DNA sequence analysis confirmed the presence of the L858R mutation in Patient 1 and Patient 2

Yu et al, Clin. Cancer Res. 2009; 15 (9) 3023- 3028

Overall Concept Design

Vision for Target Selection Process

- RFI to seek candidates from the broader scientific community
- ongoing, iterative process -

- NCI-Extramural Target Prioritization Working Group (Representatives from the extramural community together with NCI division/program leadership) receive candidates along with supporting evidence – ensure synergy avoid duplication (Group to define process)

- Prioritized candidates vetted for functional evidence – also vetted in terms of feasibility for reagent production (academic and private laboratories)

- Final selection based on supporting technical evidence, community need and feasibility

- Ongoing evaluation to evaluate process – and future model (if feasible, high value and financially viable)

Reagent/Protocol/Tools Development for Genomic/Proteomic Targets – Target Dependent

- Genomics targets/signatures - develop protocols to characterize specific/further define genomic targets (e.g., multiplex sequencing, digital mRNA profiling, etc.)

- Proteomic targets/signatures - Develop and comprehensively characterize (western blotting, IHC, epitope maps) monoclonal antibodies (mAbs); develop mAbs against peptide and/or selected protein fragments to functionalize targets

- Make resources available on all relevant websites (TCGA, TARGET, CPTAC, Other Portals)

Value of Doing the Pilot Now

- With Nex-gen technologies, integrated multi-dimensional genomic data sets on large numbers of types (and subtypes) will drive large numbers of discoveries – defining and functionalizing potential targets will become rate limiting

- The pilot program will leverage knowledge from individuals and groups to enable the development of reagents, protocols and tools for broader community – may be prerequisite for effective translation

- Targets can be identified *throughout* pilot - "just-in-time" approach could speed discovery and development

- The pilot will support collaborative early scientific pipelines – to drive new target validation and development of more effective interventions

- We can evaluate the process before we are buried under data

RFP Mechanism Chosen for the Pilot

- To ensure meeting milestones and deliverables

- Facilitate making of reagents and accompanying data freely available to the scientific community

- Best approach to include interested and qualified experts form both academic and private sectors

- Can be re-directed if needed

How Might this Develop – Future Possibilities

- **Possibility 1** – for a number of reasons (information sharing is limited; targets are easy – everyone makes everything they need; targets are really difficult to qualify and only a few people can proceed to leverage the data; or it's too expensive, etc. – **we decide not to continue the pilot program**

- **Possibility 2** - we decide that it is needed and of high value and we want to optimize it and scale it up – several scenarios are possible:
    - We scale it up at government expense (likely not a popular choice)
    - We scale it up using a cost recovery model
    - We develop the concept as a public-private partnership (could also be cost recovery)

With these possibilities in mind – the pilot must pay close attention to feasibility, time required, cost and ease of access and use by all of the relevant communities

Proposed Timeline & Budget

Initiative	Est. # Awards	FY10	FY11	$
RFP1: *Genomics*	2 - 4	$1.25 M	$1.25 M	**2.5 million**
RFP2: *Proteomics*	2 - 3	$1.25 M	$1.25 M	**2.5 million**
Total:		$2.5 M	$2.5 M	5 million