Functionalizing the Cancer Genome

Lynda Chin
Harvard Medicine School
Belfer Institute for Applied Cancer Science
Dana-Farber Cancer Institute
Broad Institute

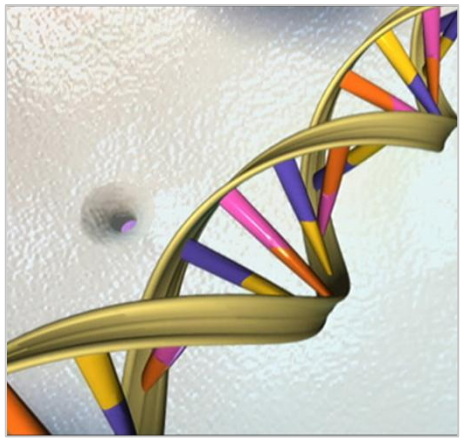
Disclosure

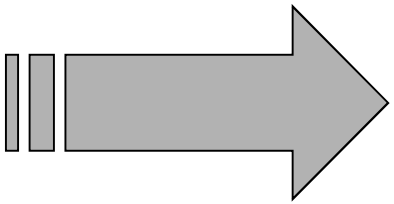
AVEO Pharmaceuticals: co-founder and advisor
Metamark Genetics: founder, director and advisor
Eden; Epizyme; Agios: Consultants;
GSK: Sponsored Research;
Merck; sanofi-aventis: Corporate alliance partnerships

Major Goals in Cancer Medicine

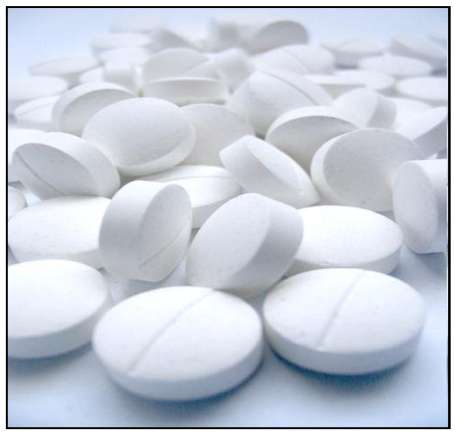
- Prevention
- Detection
- Intervention

Genome Science

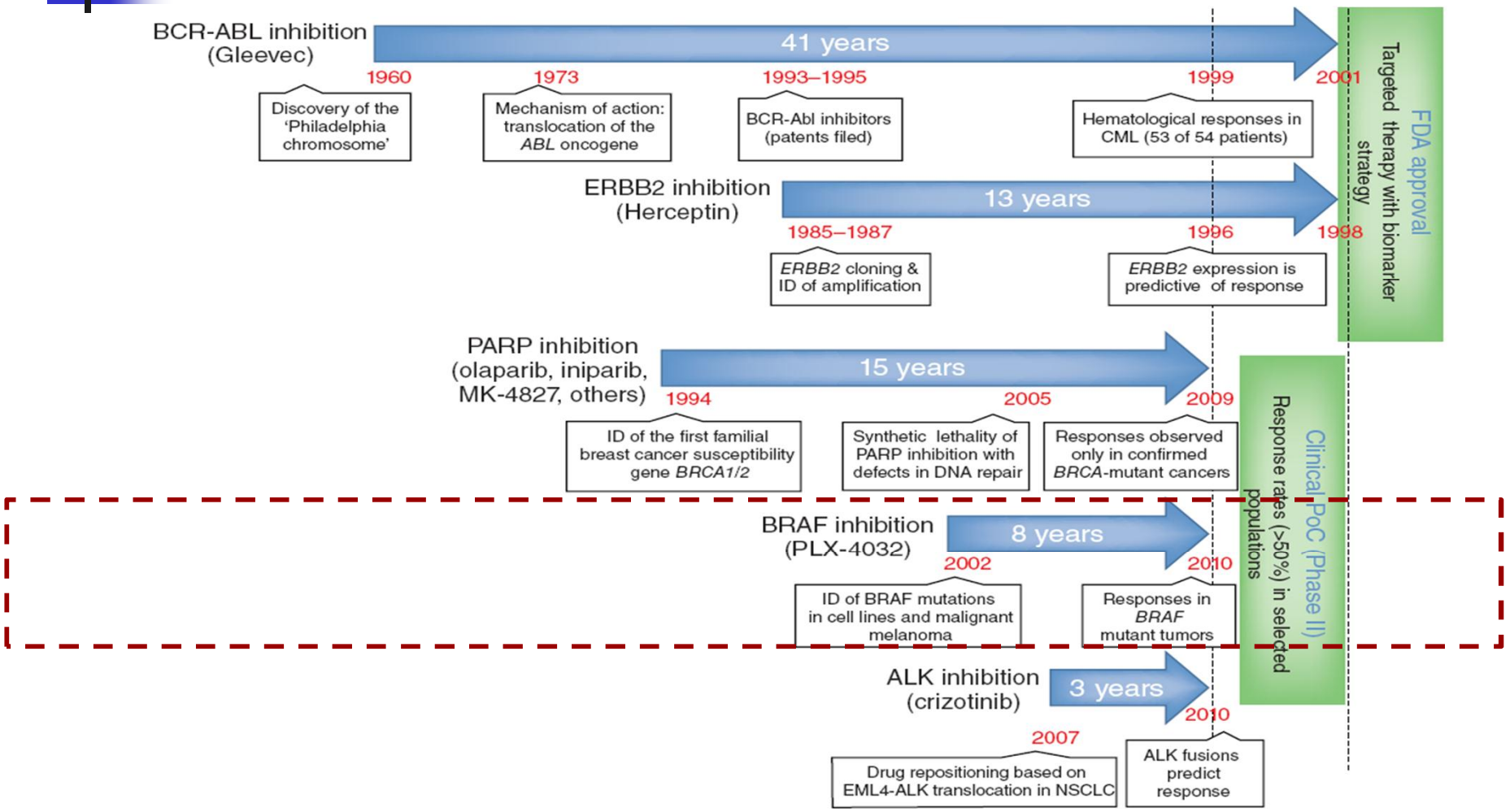




Personalized Medicine



Impacting on Cancer Medicine



Chin et al "Cancer Genomics: from discovery science to personalized medicine" *Nature Medicine* in press

Potential of Cancer Genomics

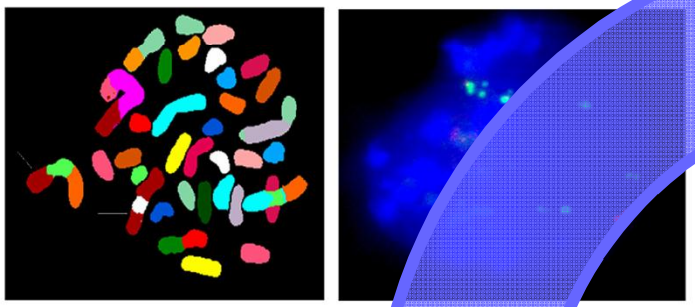
- Enable prevention
  - Understanding the underlying etiology → strategy
- Facilitate early detection é
  - Identify risk alleles / genomic events for screening
  - Early events may be detectable in serum or by imaging
- Guide evidence-based interventioné
  - Stratify high vs low risk patients to treat or not
  - Identify new therapeutic targets for drug discovery
  - Inform selection of the right patient for the right drug
  - Define combination / co-extinction strategies

→ Personalized Cancer Medicine

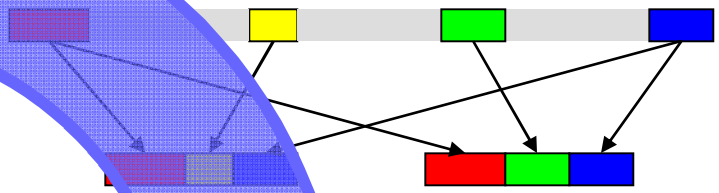
TCGA Pilot (2006-2009)

THE CANCER GENOME ATLAS

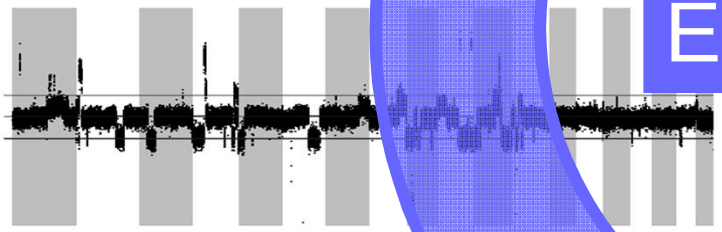
Aneuploidy; Re-arrangement; Translocation



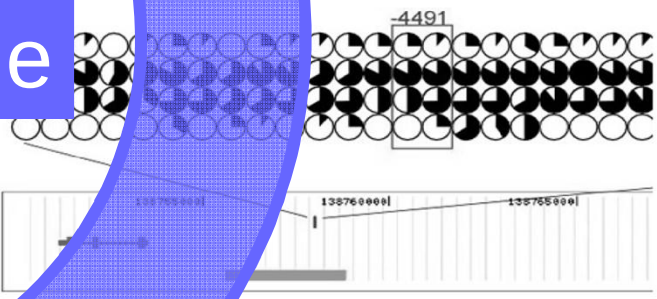
Gene Splicing Alterations



Copy number aberrations

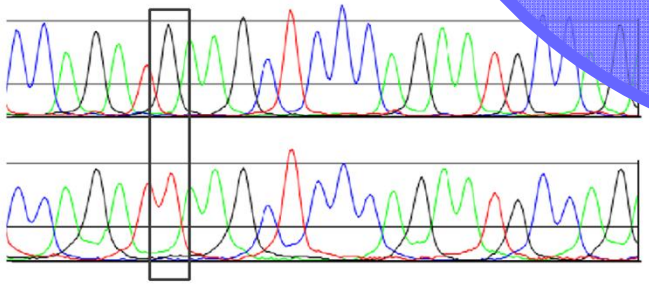


Methylation or histone modification

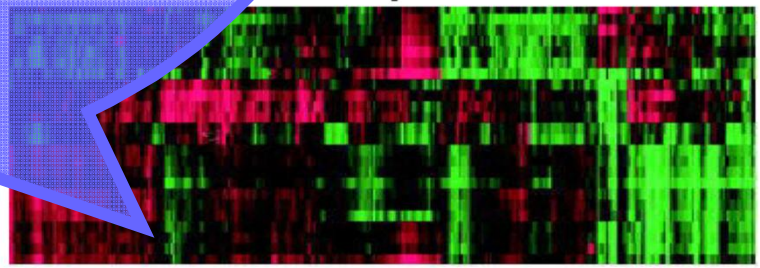


Each Sample

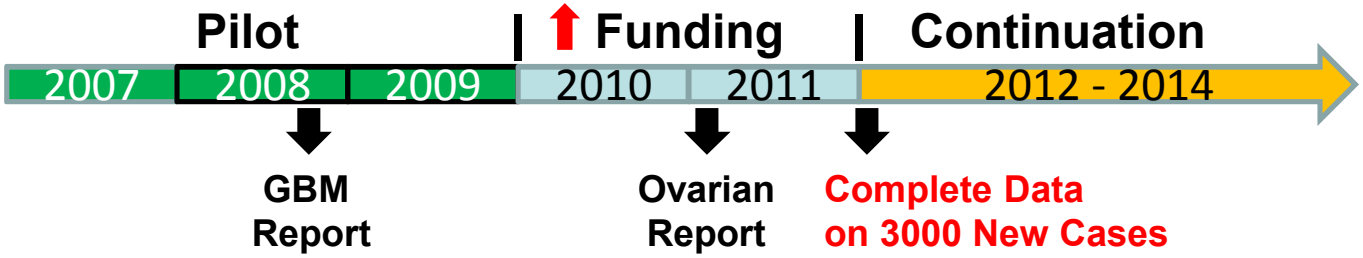
Somatic mutations

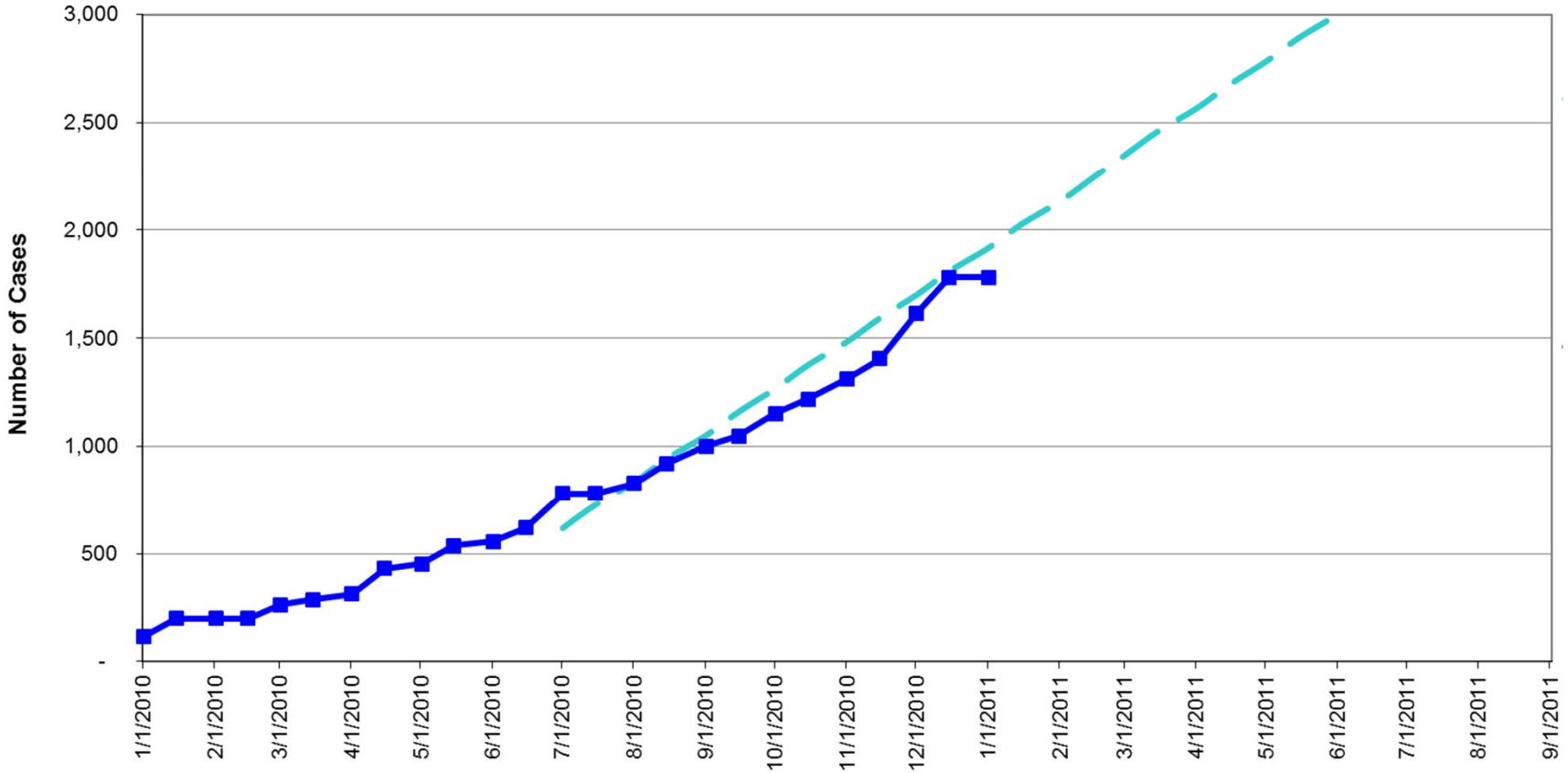


Altered expression



TCGA Phase II





TCGA Phase II Projects

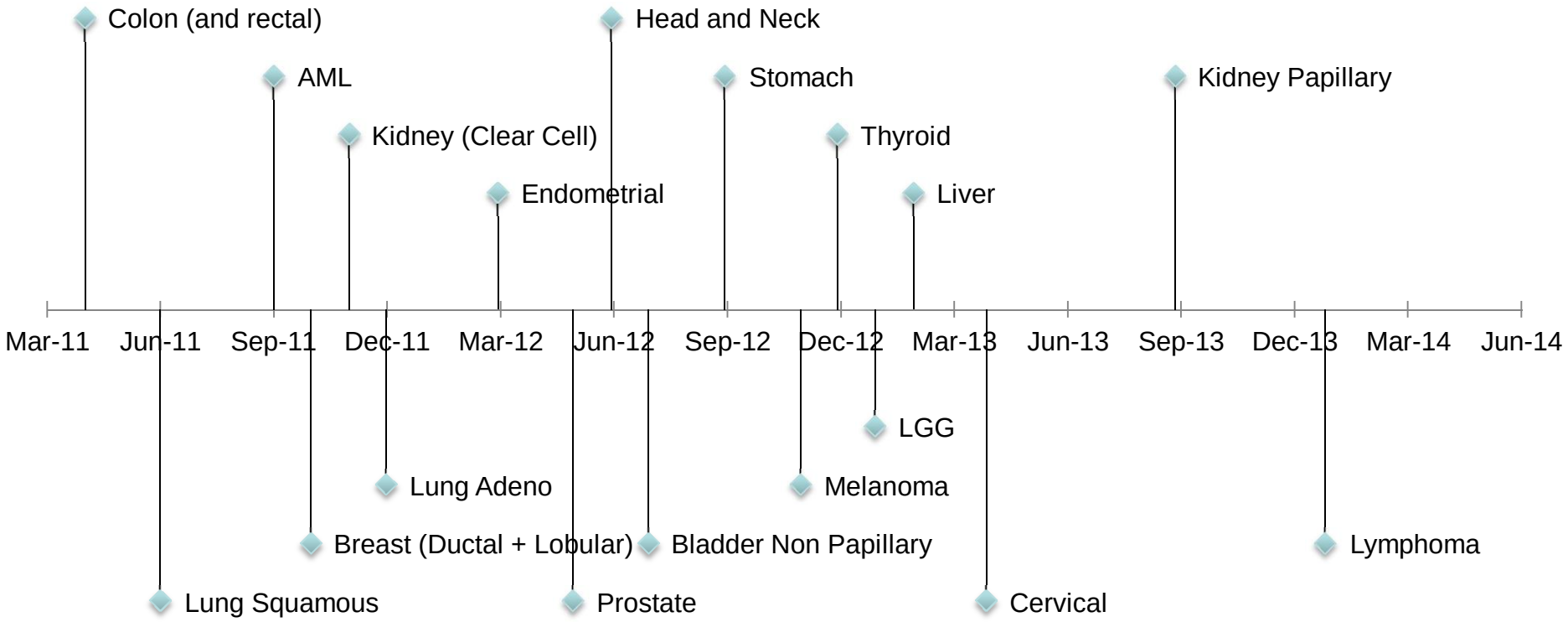
THE CANCER GENOME ATLAS

Brain	GBM and low-grade gliomas
Breast	Ductal & lobular breast adenocarcinomas
Stomach	Intestinal-type gastric adenocarcinoma
Liver	Hepatocellular carcinoma
Intestine	Colon and rectal adenocarcinomas
Gynecologic	Serous ovarian adenocarcinoma; endometrial and cervical squamous carcinomas
Prostate	Prostate adenocarcinoma
Bladder	Non-papillary bladder cancer
Head and Neck	Squamous cell and thyroid papillary carcinomas
Hematopoietic	Acute myeloid leukemia
Skin	Metastatic cutaneous melanoma
Lung	Non-small cell lung cancer, adenocarcinoma and squamous subtypes
Kidney	Renal clear cell and renal papillary carcinomas
Pancreas	Pancreatic adenocarcinoma

Active Tumor Projects

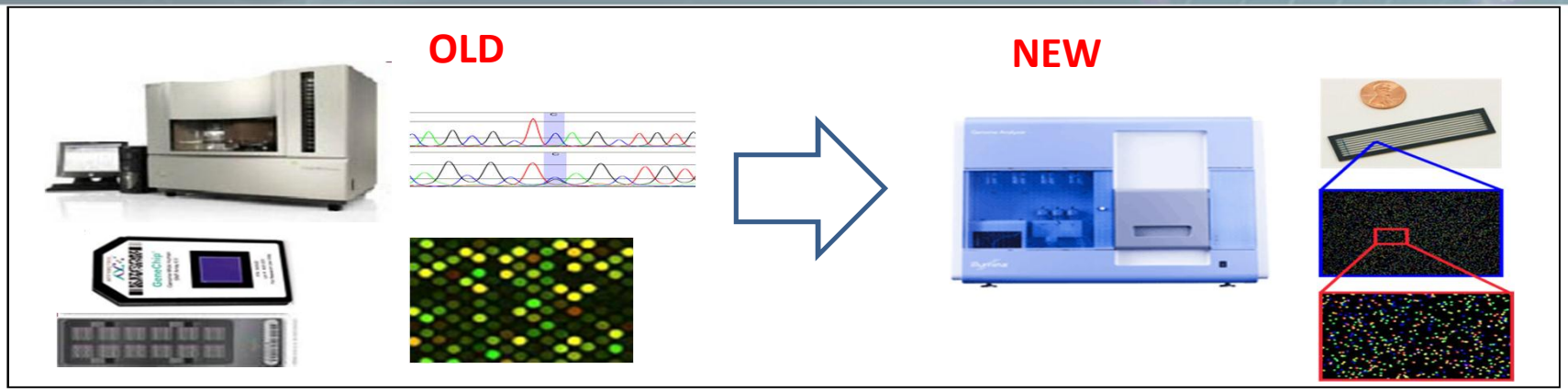
THE CANCER GENOME ATLAS

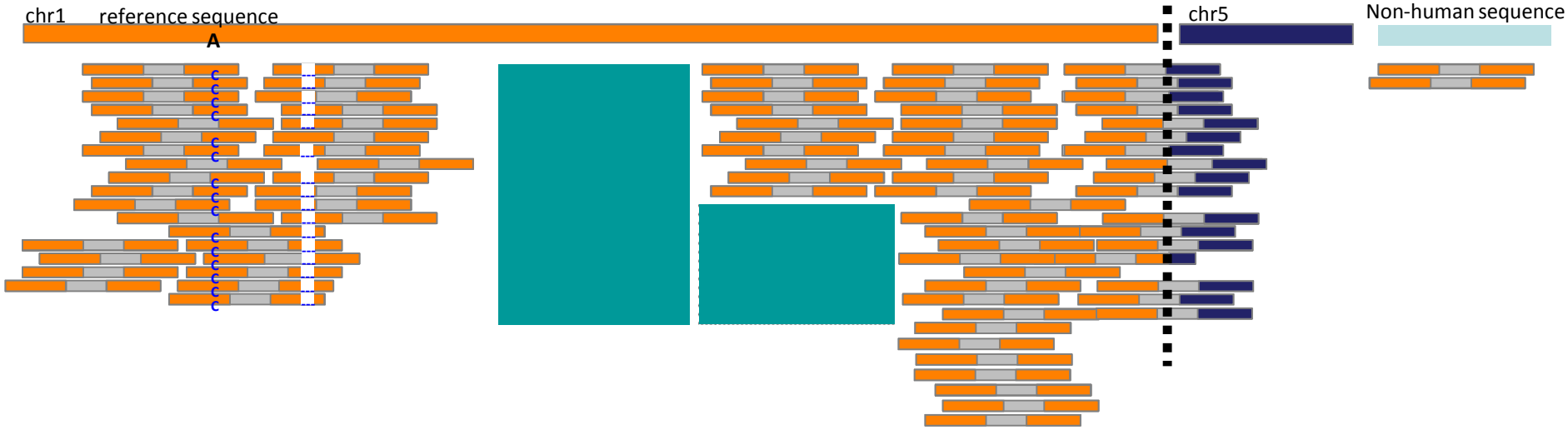
Timeline to Completion of Comprehensive Analysis for Each Tumor Project



Massively Parallel Sequencing

THE CANCER GENOME ATLAS





Gad Getz

Scale of Growth is unprecedented

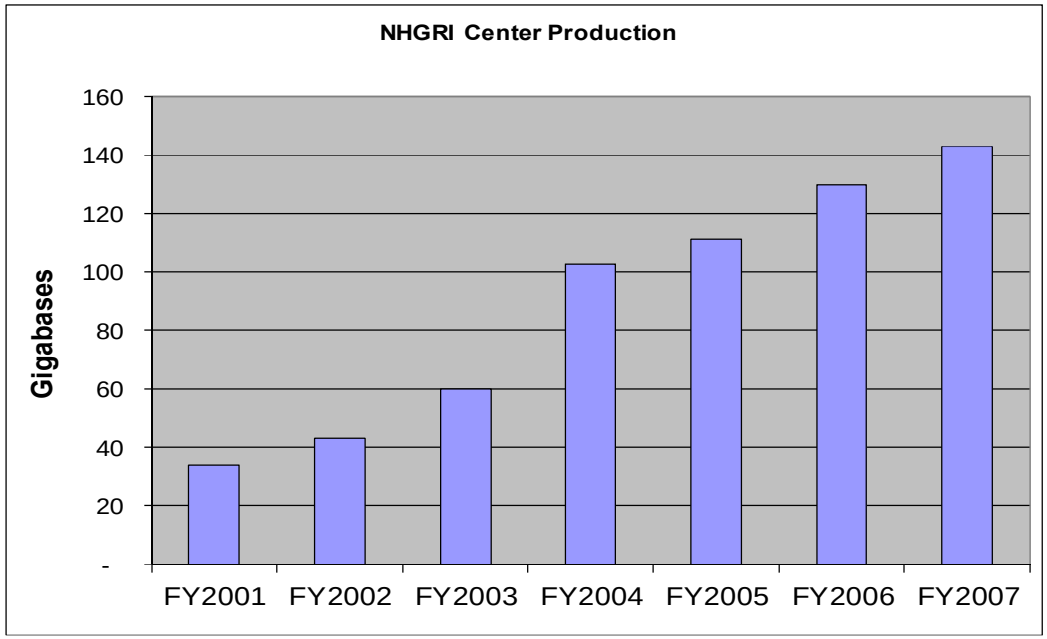
THE CANCER GENOME ATLAS

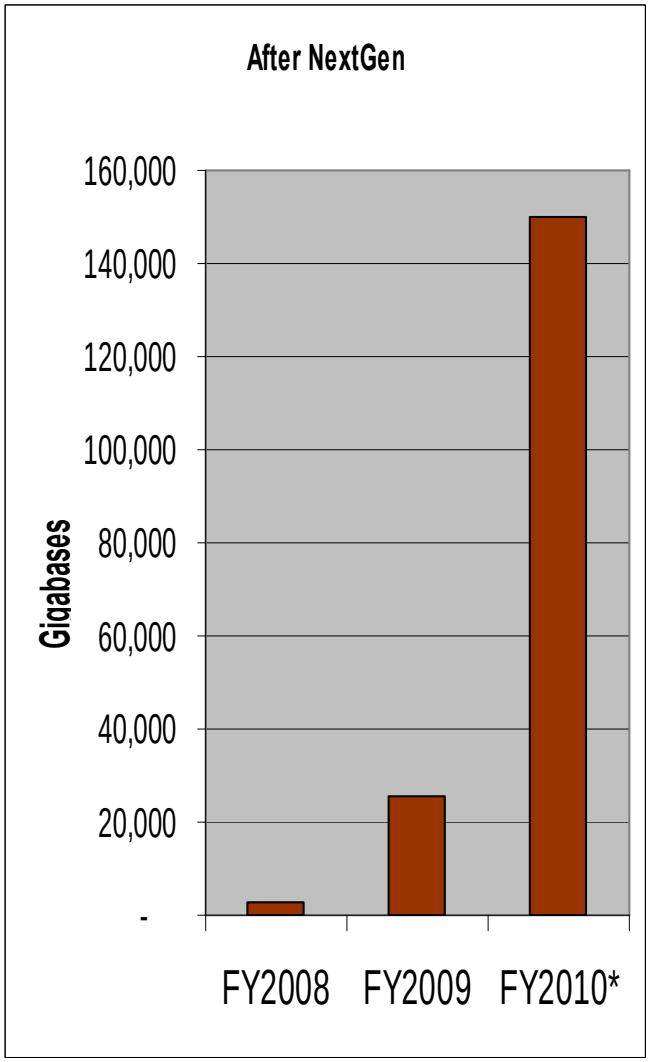
- 240 tumor cases/mo = 480 exome files (38 – 52 Mbases)
- 48 whole genomes (3,000 Mbases)

Exome (16 Gb) → **7,700 Gbytes**

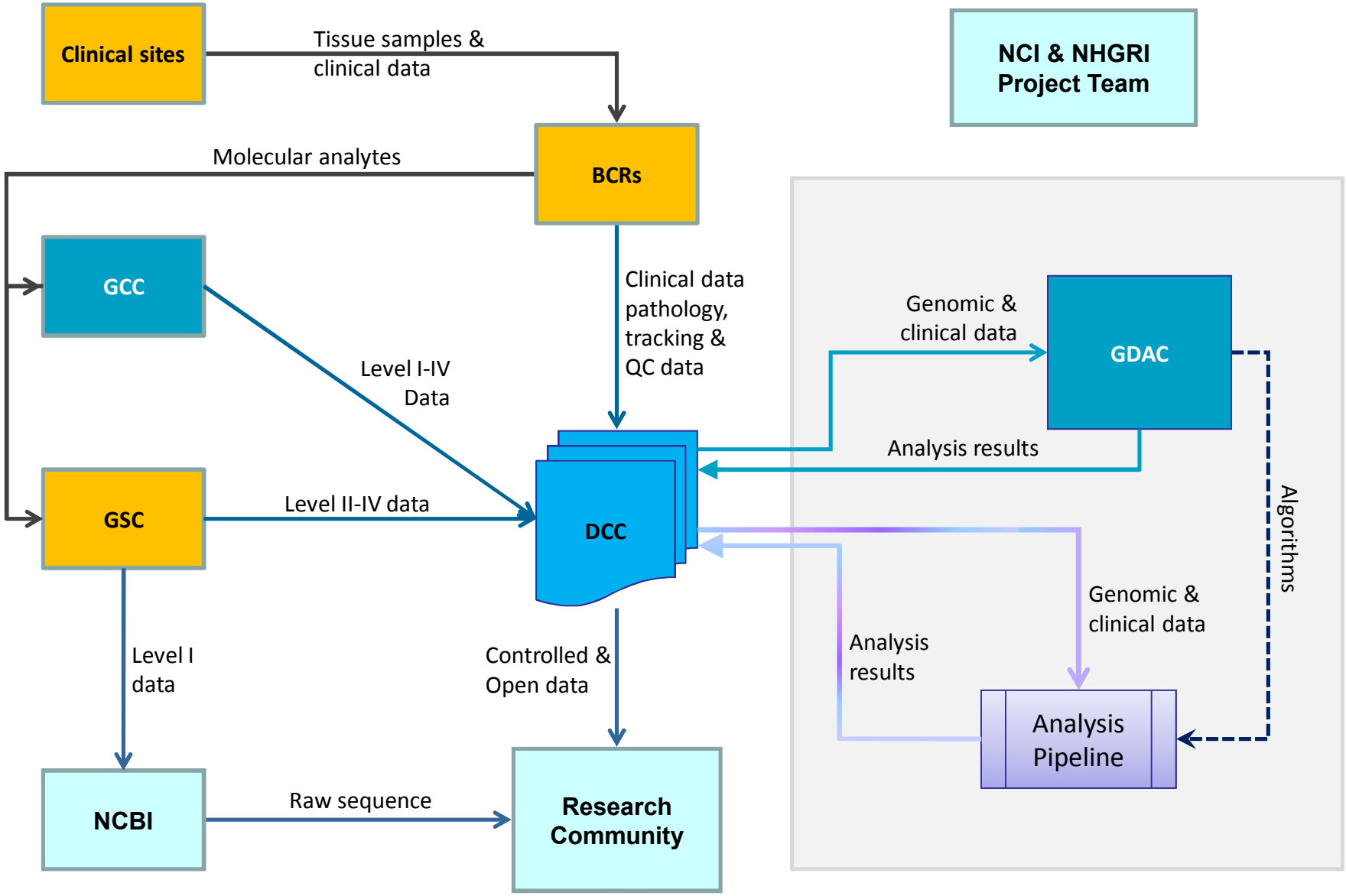
Genome (200 Gb) → **9,600 Gbytes**

Per month **17.3 Terabytes TOTAL**





TCGA Phase II Research Network

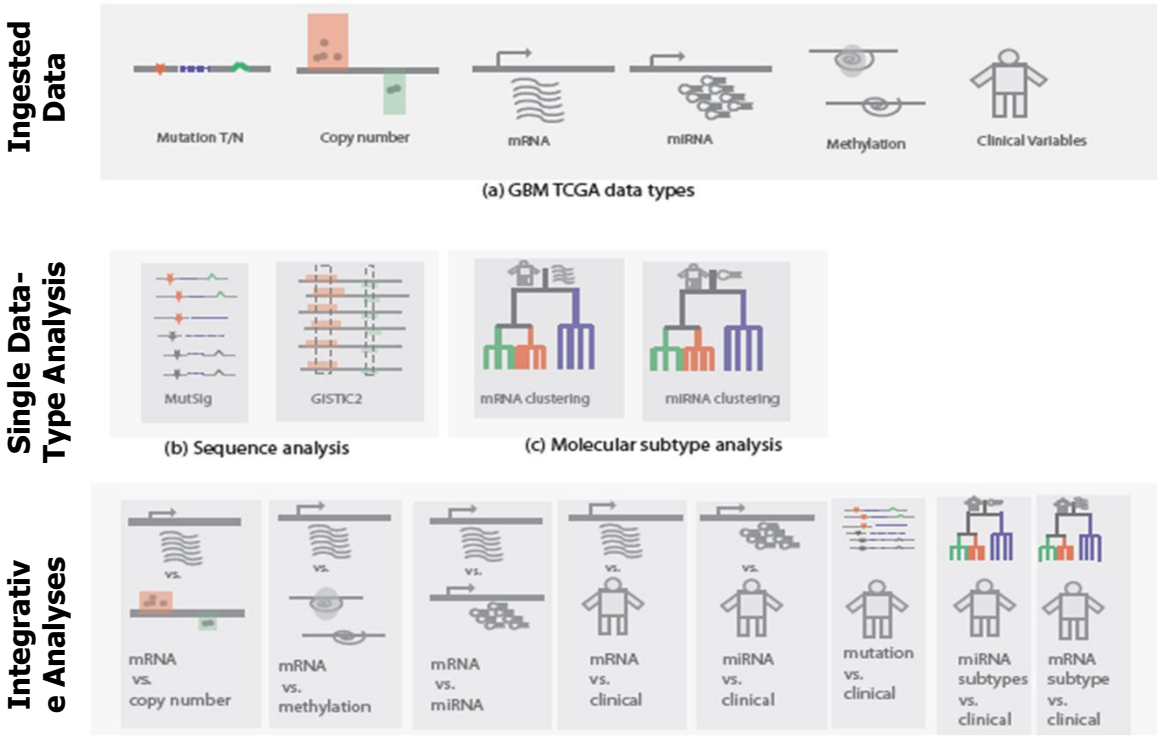


Summary of TCGA Tumor Data
Ingested into Broad GDAC Pipeline
01/14/2011 Run

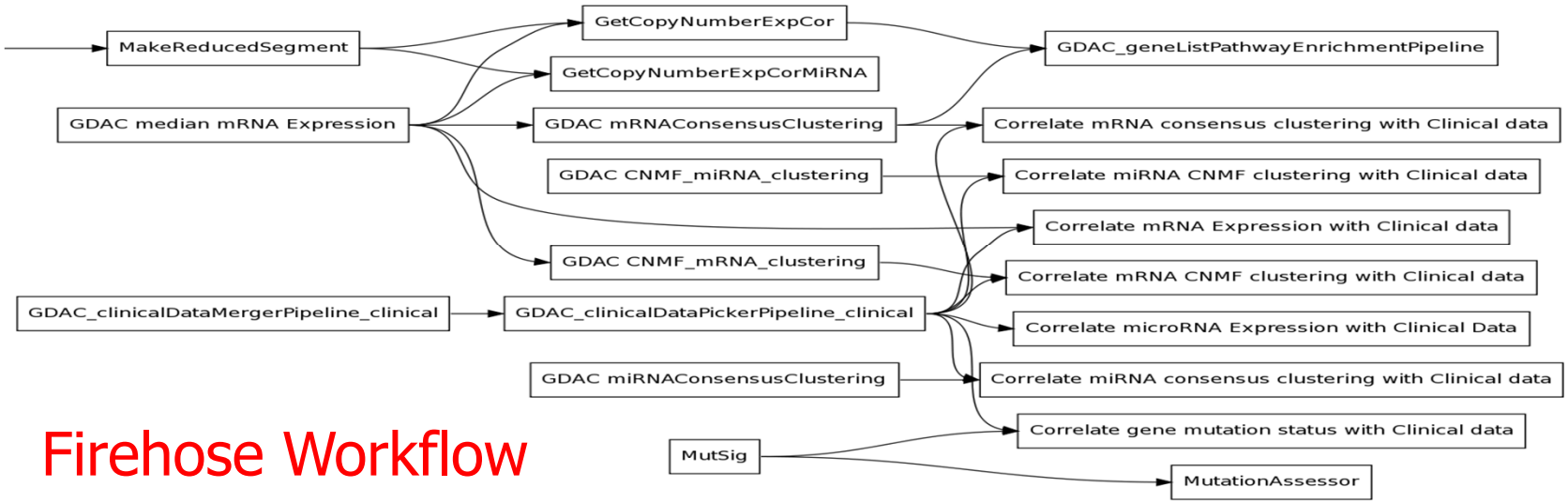
TumorType	Biospecimen	Any_Level_1	Clinical	CNA	Methylation	mRNA	miR	MAF
BRCA	346	186	244	265	186	280	0	0
COAD	203	151	130	137	167	155	0	64
GBM	508	448	490	466	288	444	415	169
HNSC	39	0	0	0	0	0	0	0
KIRC	355	39	19	254	219	41	0	0
KIRP	48	39	0	16	36	41	0	0
LAML	202	0	0	0	188	0	0	0
LGG	30	0	0	0	0	0	0	0
LUAD	128	21	11	56	128	33	0	0
LUSC	160	116	42	117	133	116	0	0
OV	584	570	532	519	425	519	566	384
READ	79	52	72	51	69	69	0	12
STAD	82	35	0	81	82	0	0	0
UCEC	145	24	0	114	70	0	0	0
Totals	2909	1681	1540	2076	1991	1698	981	629

Status of TCGA Analysis Pipeline (Jan 14, 2010 Run)

Mike Noble; Doug Voet



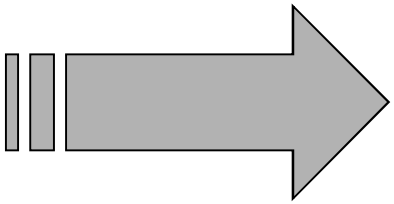
- Pre-defined analyses
- Automated and Fast, producing standard human-readable summary reports
- Reproducible
  - Associated input, algorithms and parameters are tracked
  - Uniform intermediate data files for higher level analyses

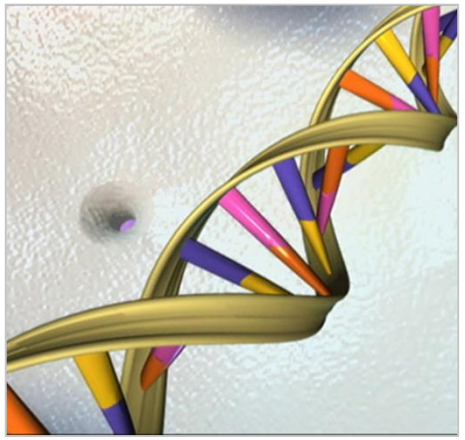


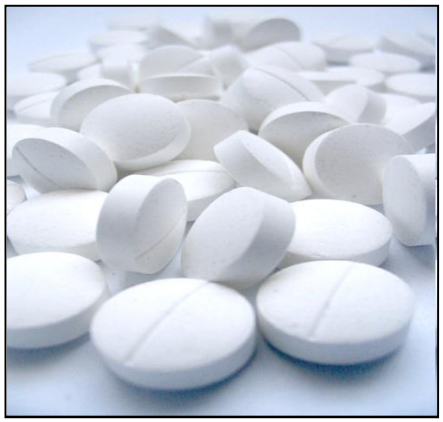
Firehose Workflow

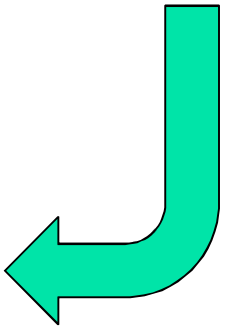
Complete catalogues will be generated

Complete Compendia







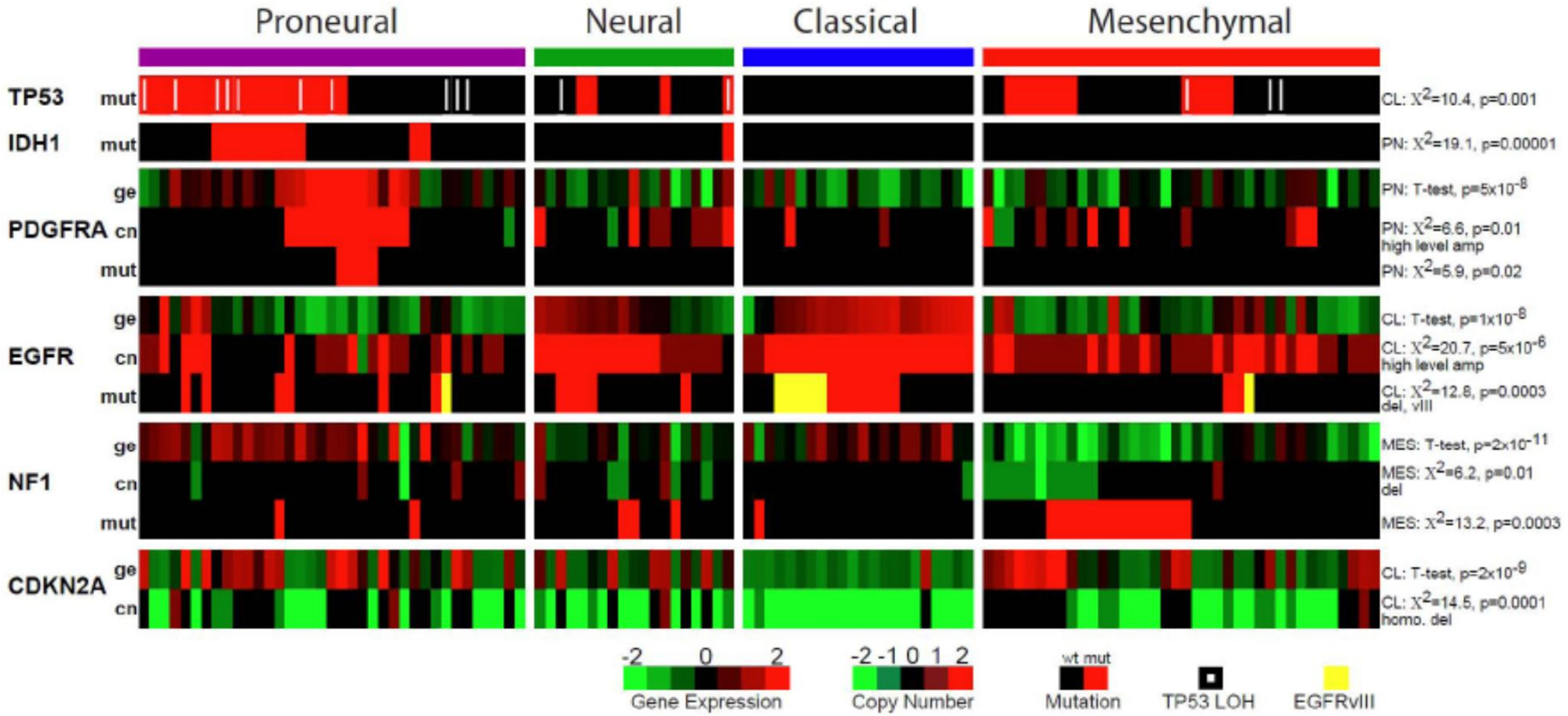


What does it take?

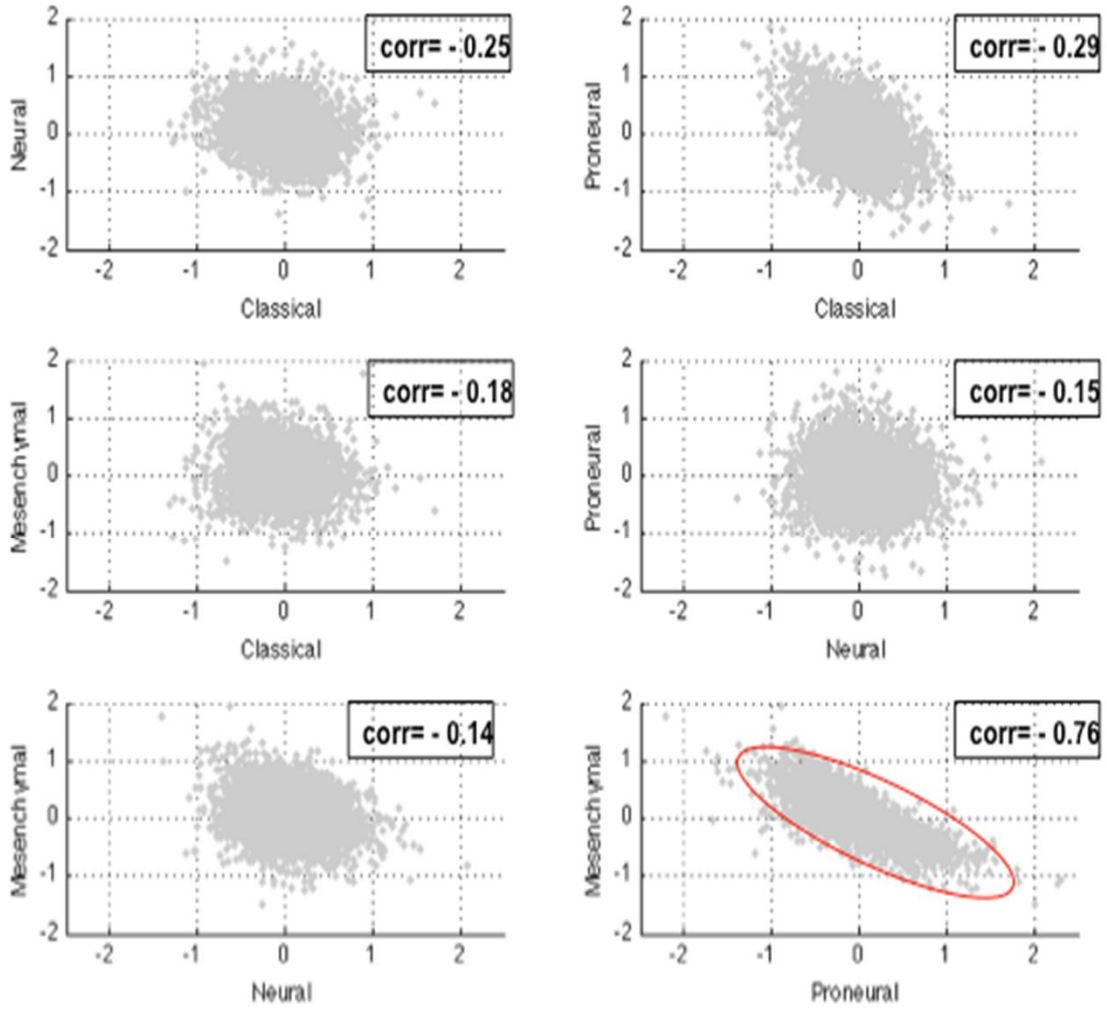
Cancer Cell
Article

Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*

Roel G.W. Verhaak,[1,2,17] Katherine A. Hoadley,[3,4,17] Elizabeth Purdom,[7] Victoria Wang,[8] Yuan Qi,[4,5] Matthew D. Wilkerson,[4,5] C. Ryan Miller,[4,6] Li Ding,[9] Todd Golub,[1,10] Jill P. Mesirov,[1] Gabriele Alexe,[1] Michael Lawrence,[1,2] Michael O'Kelly,[1,2] Pablo Tamayo,[1] Barbara A. Weir,[1,2] Stacey Gabriel,[1] Wendy Winckler,[1,2] Supriya Gupta,[1] Lakshmi Jakkula,[11] Heidi S. Feiler,[11] J. Graeme Hodgson,[12] C. David James,[12] Jann N. Sarkaria,[13] Cameron Brennan,[14] Ari Kahn,[15] Paul T. Spellman,[11] Richard K. Wilson,[9] Terence P. Speed,[7,16] Joe W. Gray,[11] Matthew Meyerson,[1,2] Gad Getz,[1] Charles M. Perou,[3,4,8] D. Neil Hayes,[4,5,*] and The Cancer Genome Atlas Research Network



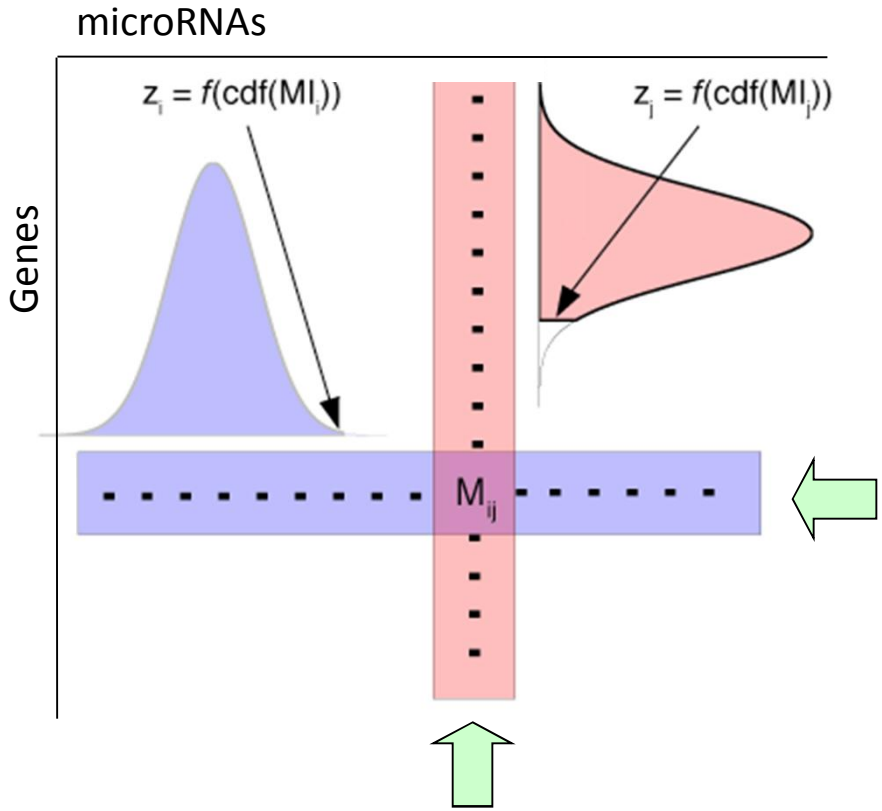
What is driving the molecular difference among subtypes?

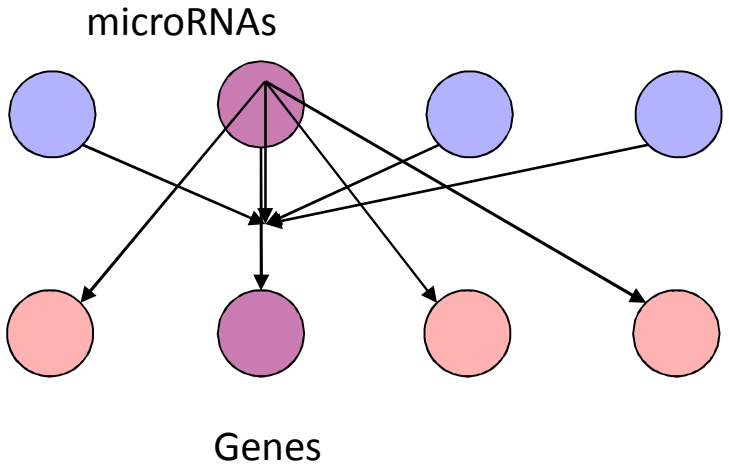


The most significant difference is observed between PN and MS subtypes

CLR- Context Likelihood of Relatedness

- Faith et al, PLoS Biology, 2007
- Extension of relevance networks
- Based on Mutual Information scores

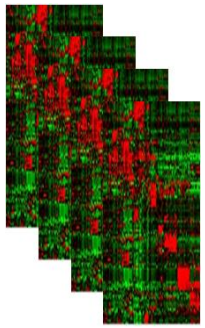




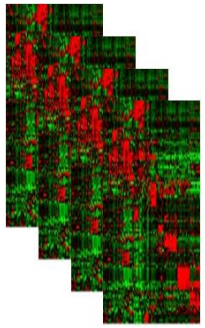
Ayla Ergun; Jim Collins

Define miRNA-mRNA regulatory network

microRNA expression



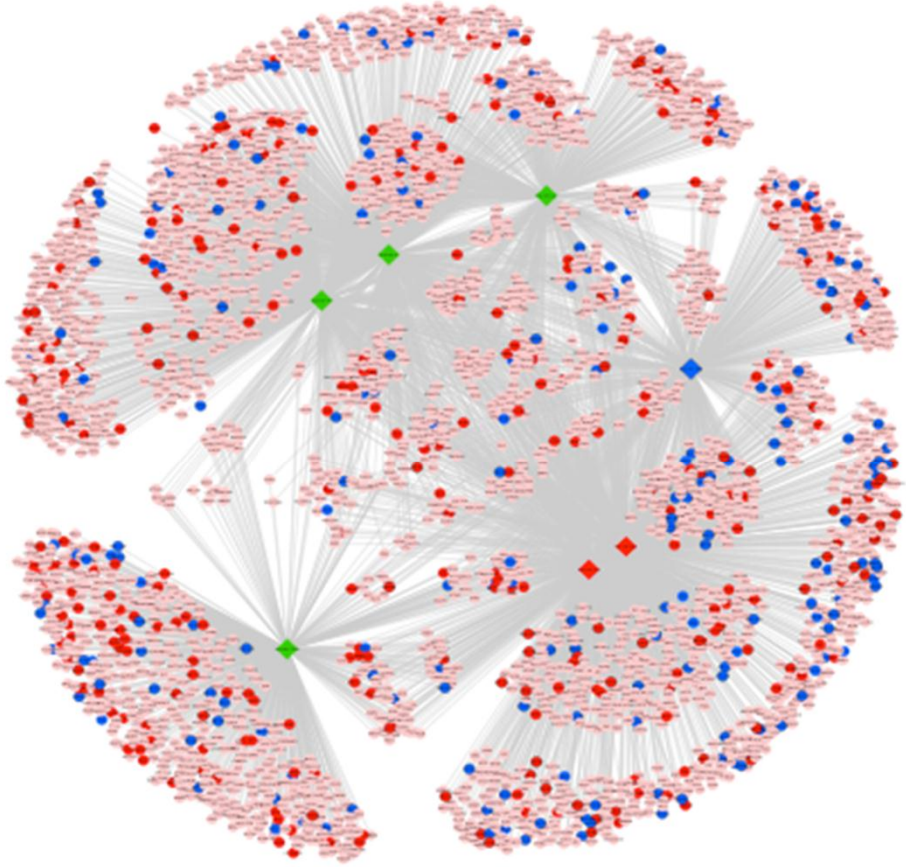
mRNA expression



194 matched samples

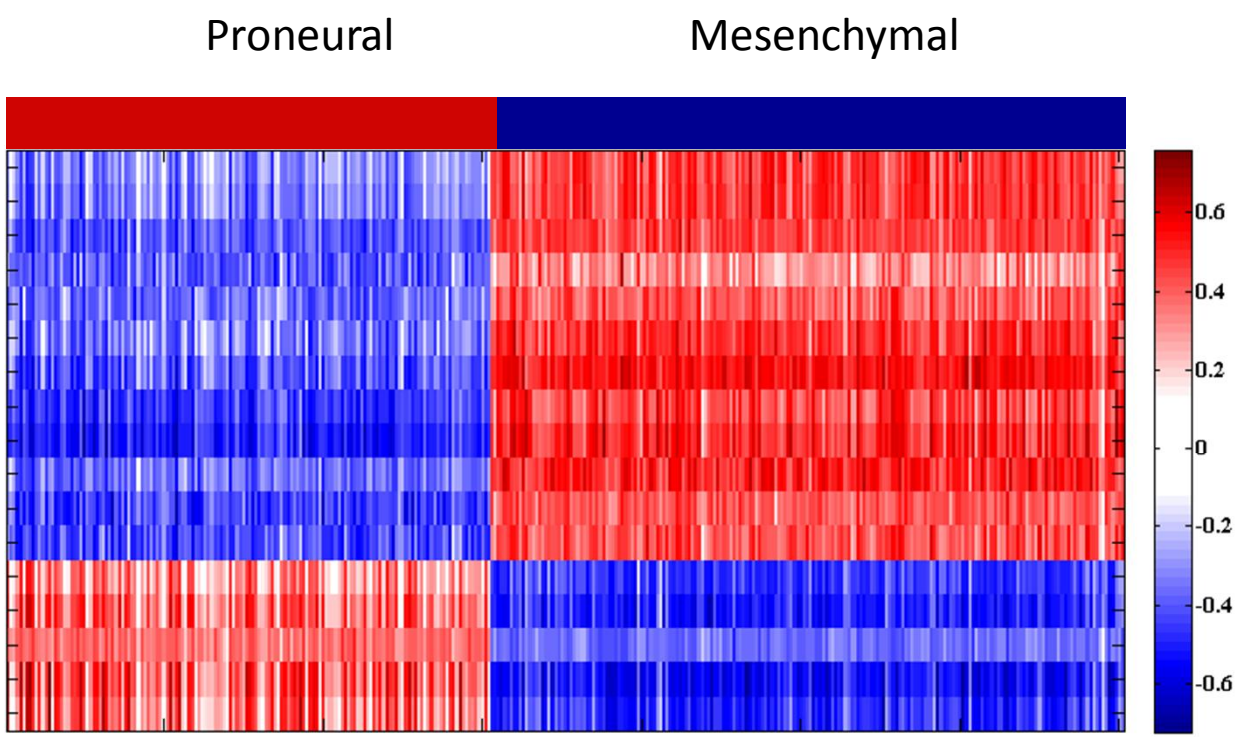
Å TCGA Glioblastoma data

ï 534 miRNA, 19692 mRNAs



➔ 29610 edges: 252 miRNA and 7373 mRNAs

A subset of the miRNAs show strong correlations with subtype signature genes



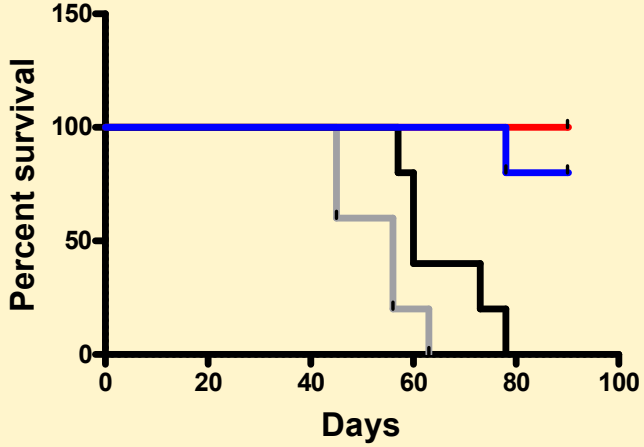
17 miRNA with strong correlations with the proneural and mesenchyaml signature genes

miR34a as a candidate determinant of PN molecular subtype

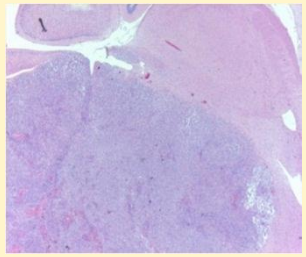
- Integration with copy number reveals miR34a resides in region of loss
- miR34a is low in PN subtype GBM
- PN signature is enriched for miR34a edges defined by CLR

miR34a is tumor suppressive in human GBM models *in vivo*

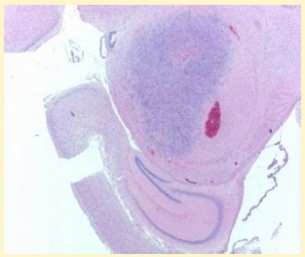
LOF studies in Human GBM Cells



LN319 scr n=5
LN319 mir-34a n=5
A1207 scr n=5
A1207 mir-34a n=5

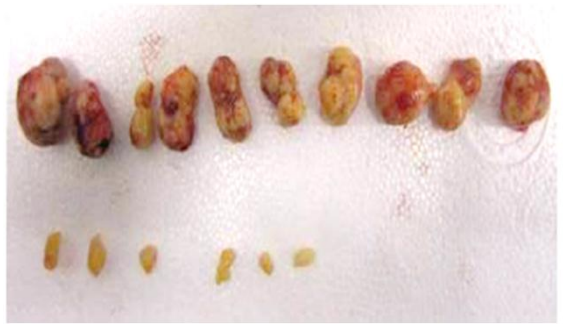


A1207 pHAGE ct



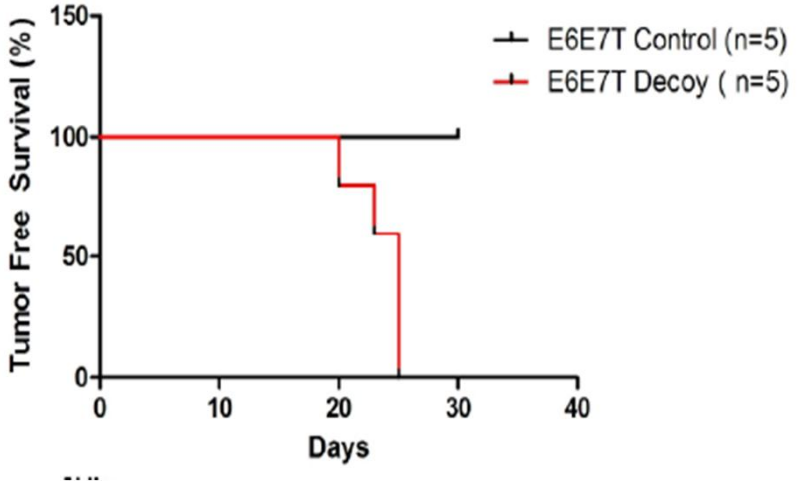
A1207 pHAGE mir-34a

GOF studies in E6/E7 Astrocytes

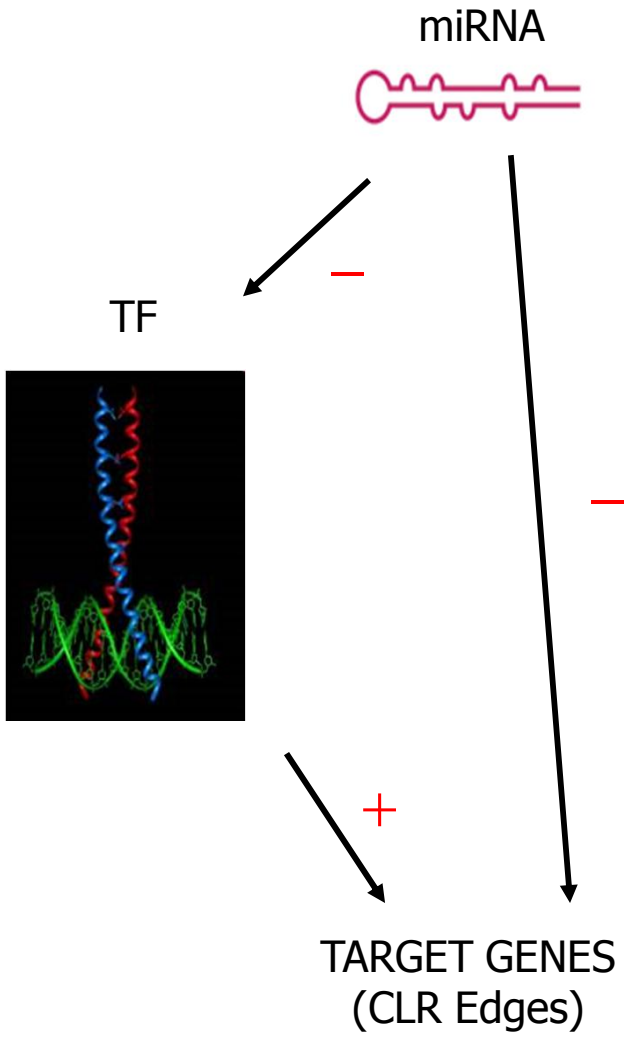


Decoy

Ctrl



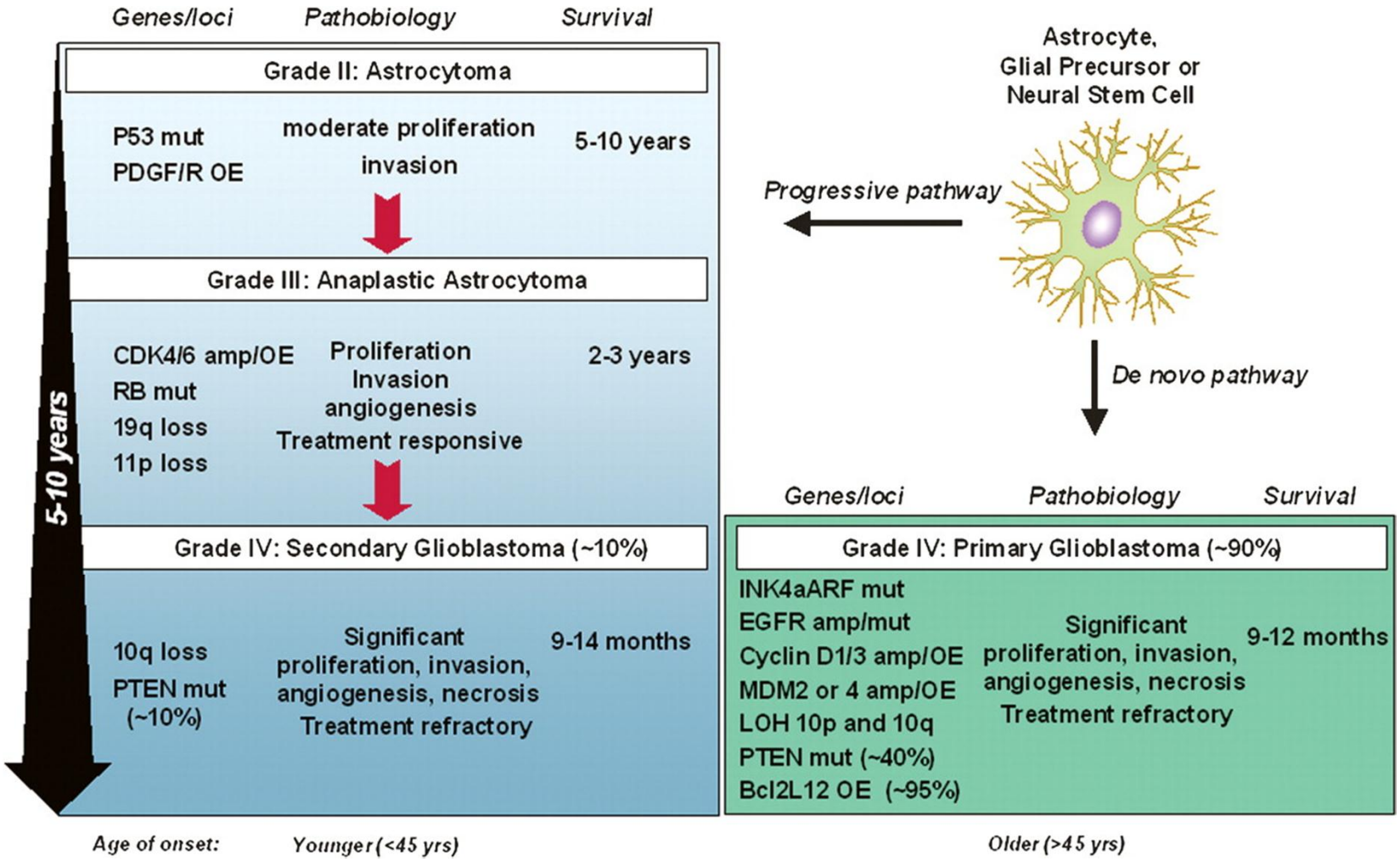
How does miR34a regulate the PN/MS transcriptomic signatures?



- 3ôUTR Luciferase reporter ï direct regulation of PDGFRa and DLL1 by miR34a
- Modulation of miR34a regulates PDGFRa and DLL1 (as well as Notch downstream targets) in human astrocytes and GBM cells

➔ Relevance?

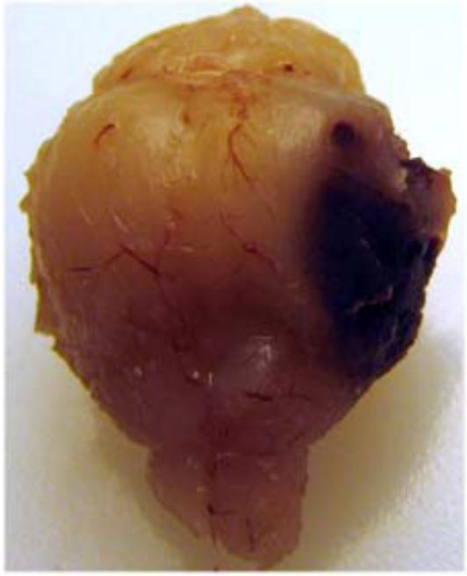
Chromosomal and genetic aberrations involved in the genesis of glioblastoma.

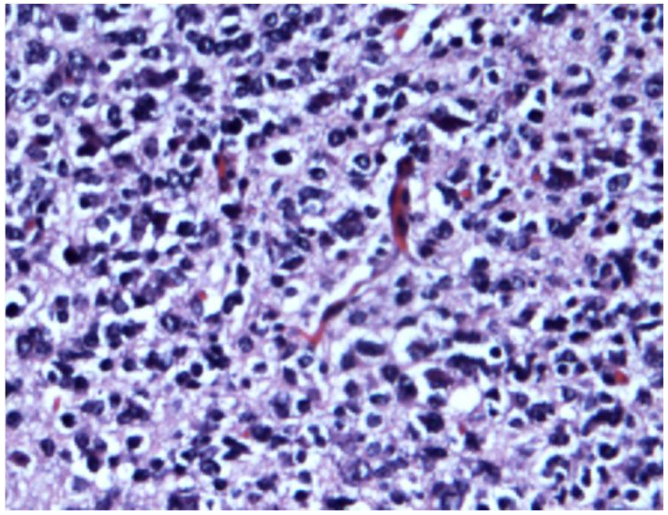


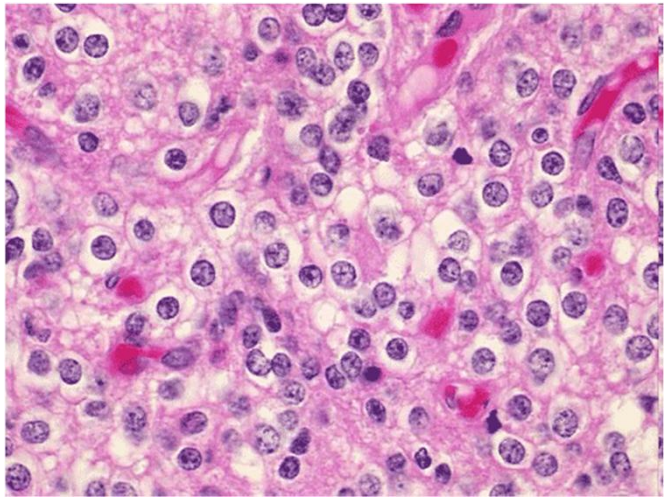
Furnari F B et al. Genes Dev. 2007;21:2683-2710

TCGA GBM cohort shows enrichment of NOTCH in Classical and PN subtypes

p53 and *Pten loss in neural progenitor cells results in malignant gliomas*

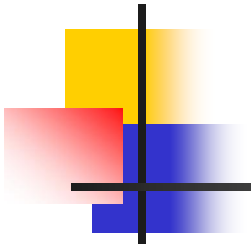






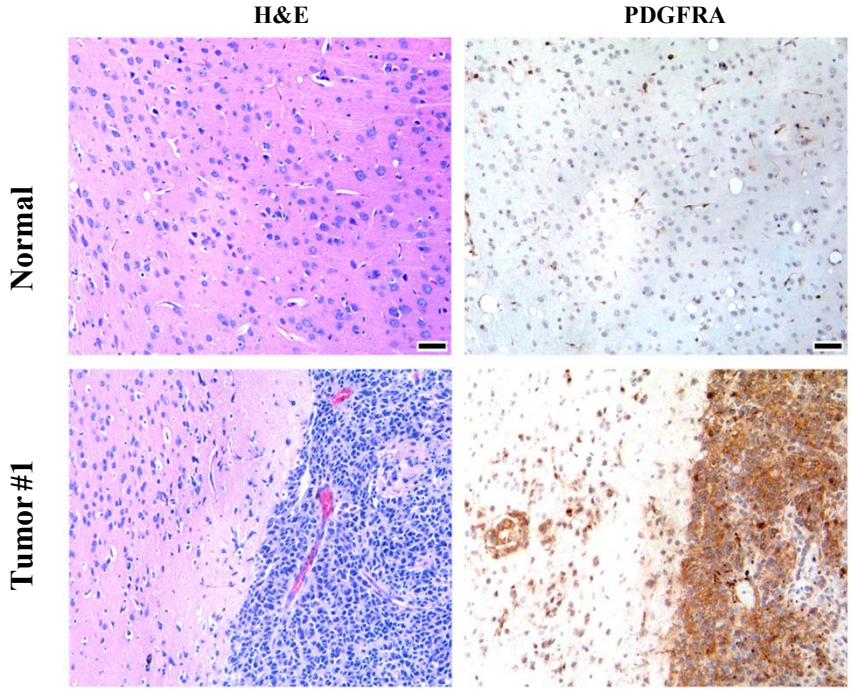
- 42/57 (73%) p53L/L Pten L/+ mice:
- Acute neurological symptoms
- 28 Grade III; 14 Grade IV
- astrocytic morphology (95%)
- diffuse & proliferative
- necrosis
- glioma markers

Zheng (DePinho), Nature 2008



P53/Pten GBM model is PN-like

Pdgfra overexpression is a hallmark of mouse PN GBMs



(Zheng et al., Nature 2008)

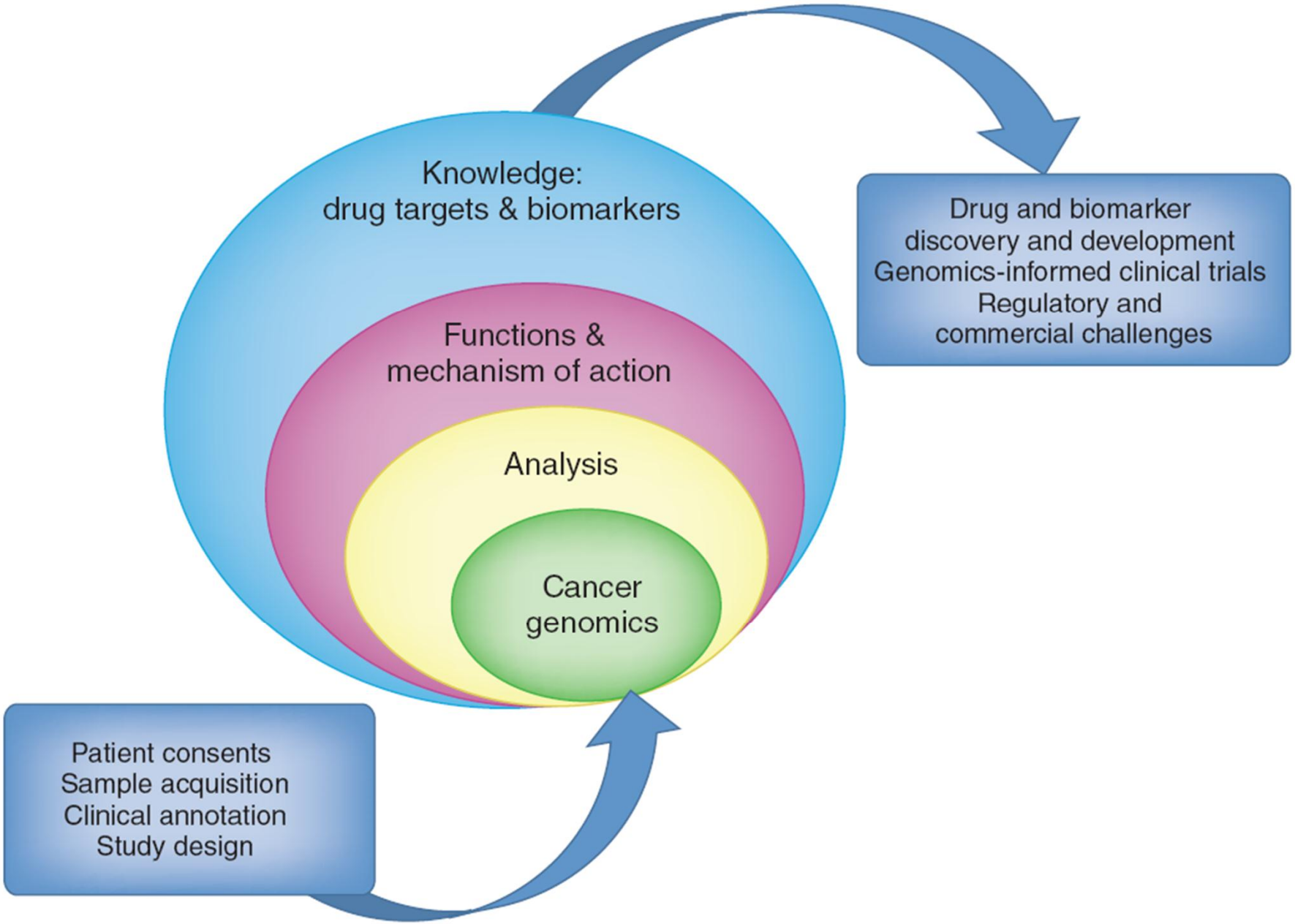
- miR34a modulates Pdgfra expression
- Pdgfra is functionally epistatic to miR34a

miR34a-low GEMM tumors show Notch activation in vivo

- miR34a modulates Notch activity and its downstream target gene expression

miR34a is a determinant of PN molecular phenotype in GBM

- Integrative genomic data set enables
  - network modeling to generate testable hypothesis
  - development of framework for understanding complex cancer genomics data
- miR34a defines a subset of GBM with concurrent PDGFRa and Notch activation

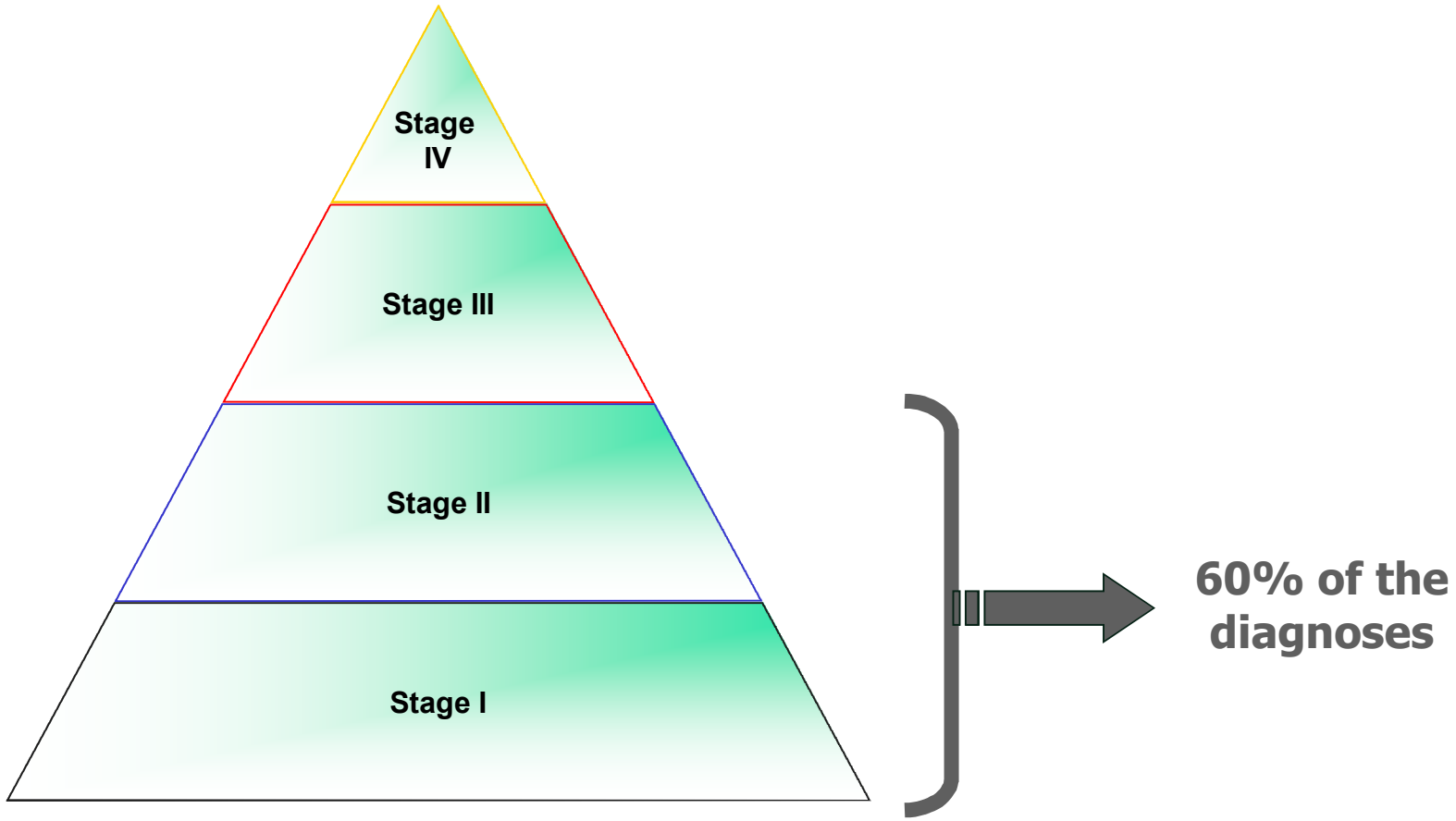


Chin et al "Cancer Genomics: from discovery science to personalized medicine" *Nature Medicine* in press

Promise of Cancer Genomics

- Enable prevention
  - Understanding the underlying etiology → strategy
- Facilitate early detection é
  - Identify risk alleles / genomic events for screening
  - Early events may be detectable in serum or by imaging
- Guide evidence-based interventioné
  - Stratify high vs low risk patients to treat or not
  - Identify new therapeutic targets for drug discovery
  - Inform selection of the right patient for the right drug
  - Define combination / co-extinction strategies

Early-staged patients make up the majority of US cancer diagnoses



doi:10.1038/nature09677

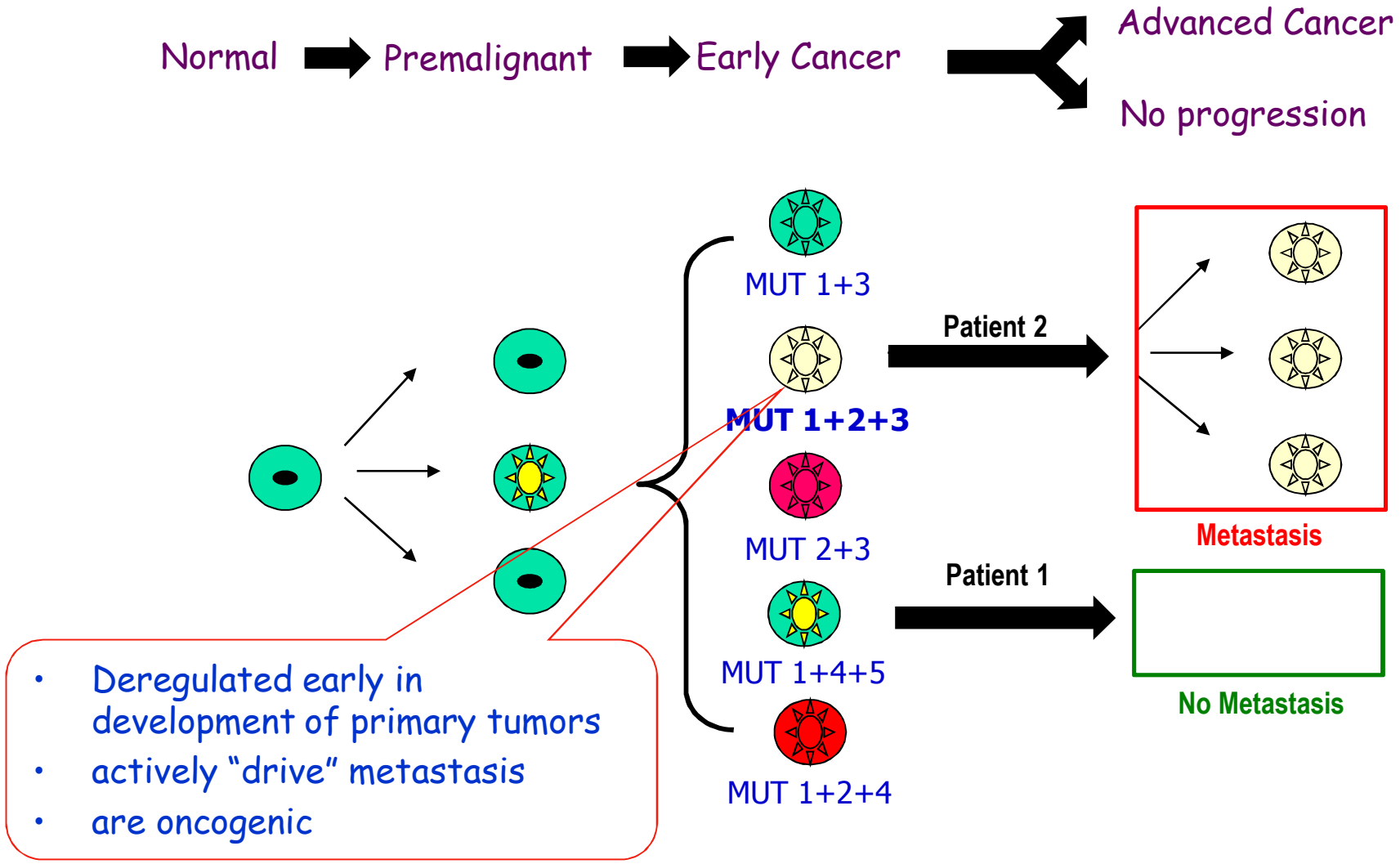
SMAD4-dependent barrier constrains prostate cancer growth and metastatic progression

Zhihu Ding[1,2,3,4], Chang-Jiun Wu[1,2,3,4]*, Gerald C. Chu[1,2,5]*, Yonghong Xiao[1,2], Dennis Ho[1,2,3,4], Jingfang Zhang[6], Samuel R. Perry[1,2], Emma S. Labrot[1,2], Xiaoqiu Wu[2,7], Rosina Lis[2,7], Yujin Hoshida[8,9], David Hiller[10], Baoli Hu[1,2], Shan Jiang[1,2], Hongwu Zheng[1,2,3,4], Alexander H. Stegh[1,2,3,4], Kenneth L. Scott[1,2,3,4], Sabina Signoretti[11], Nabeel Bardeesy[12], Y. Alan Wang[1,2], David E. Hill[3,13], Todd R. Golub[8,9], Meir J. Stampfer[15,16,17], Wing H. Wong[10], Massimo Loda[2,5,7], Lorelei Mucci[15,17], Lynda Chin[1,2,3,4,14] & Ronald A. DePinho[1,2,3,4]

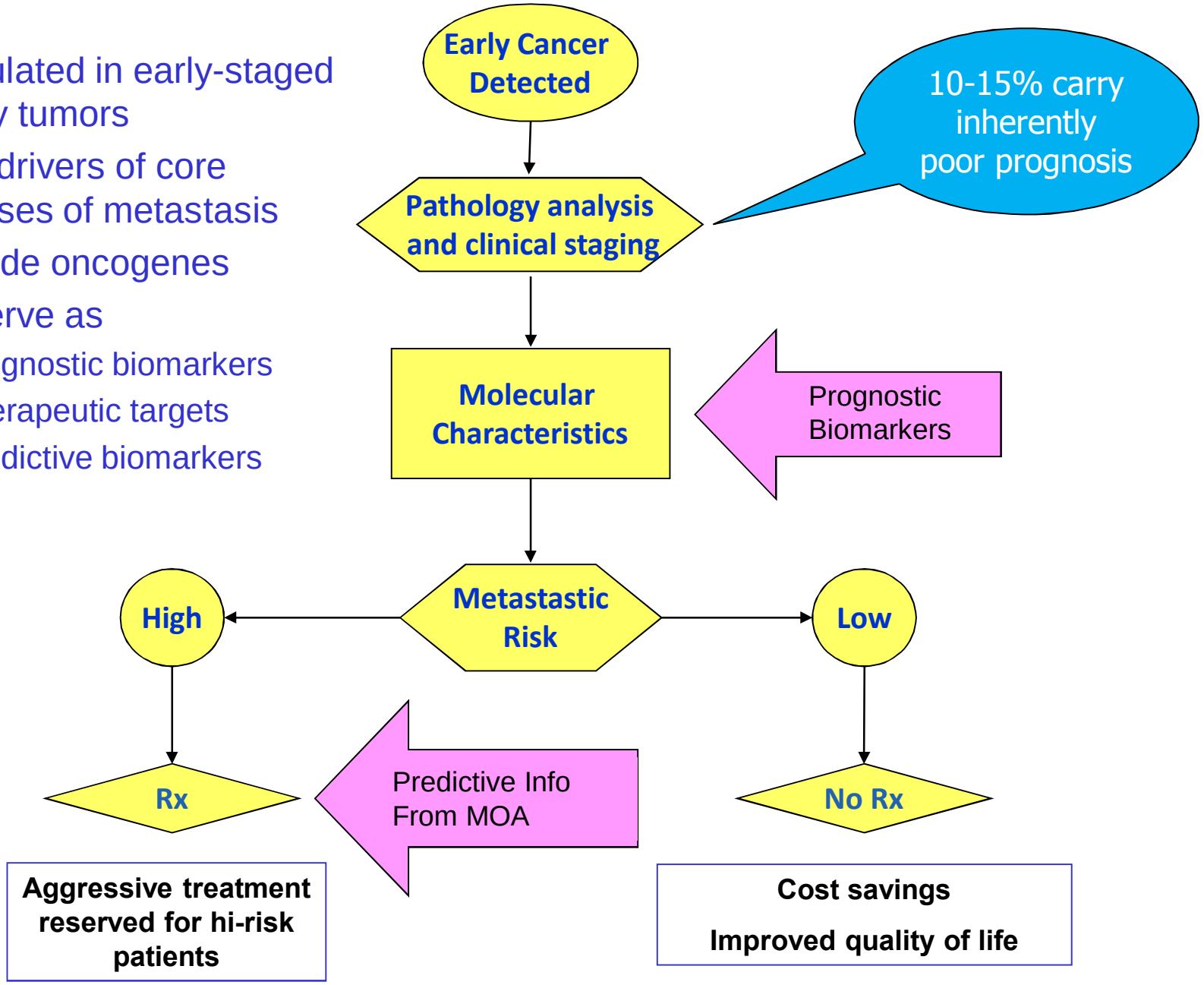
In Physicianôs Health Cohort (n=405)

- 4-marker outperforms Gleason in predicting lethal disease
  - Gleason-only C Index = 0.774;
  - 4-marker only C Index = 0.829
- Carries molecular information not captured by clinical parameters
  - 4-marker + Gleason C Index =0.882, p = 0.015 for improvement

Metastatic potential of a primary tumor can be determined early on in evolution



- Deregulated in early-staged primary tumors
- Active drivers of core processes of metastasis
- Bona fide oncogenes
- May serve as
  - Prognostic biomarkers
  - Therapeutic targets
  - Predictive biomarkers



Adapted from: Rethinking Screening for Breast and Prostate Cancer. JAMA, 2009

Cancer Genomics → Genomic Medicine

- Enable prevention
  - Understanding the underlying etiology → strategy
- Facilitate early detection é
  - Identify risk alleles / genomic events for screening
  - Early events may be detectable in serum or by imaging
- Guide evidence-based interventioné
  - Stratify high vs low risk patients to treat or not
  - Identify new therapeutic targets for drug discovery
  - Inform selection of the right patient for the right drug
  - Define combination / co-extinction strategies

Acknowledgement

Cancer Genomics
Alexei Protopopov
Elena Ivanova
Maria Alimova
Ilana Perna Otey
Georgia Ren

Bioinformatics
Yonghong Xiao
Juihua Zhang
Spring Liu
Sachet Shukla
Hailei Zhang
Terrence Wu

Molecular Pathology
Gerry Chu
Perry Samuels

At the Bench:

Denise Spring
Nate Goldstein
Steven Quayle
Larry Kwong
Papia Ghosh
Gianni Genovese
Chengyin Min
Sharmistha Sarka
Kunal Rai
Nina Seitzer
Erik Uhlmann
Terrence Wu
Ian Watson
Yonathan Lissanu Deribe
Benito Campos
Robert Dewan
Huiyu Liu

Ruprecht Wiedemeyer
Kenneth Scott
Omar Kabbarah
Cristina Nogueira
Tim Heffernan

Metastasis oncogenes

Jason Hanna
David Rimm (Yale)

GOLPH3:

Kwok-kin Wong (DFCI)
Mei-Chih Liang

David Rimm (Yale)
Elsa Anagnastou

CLR Network Modeling:

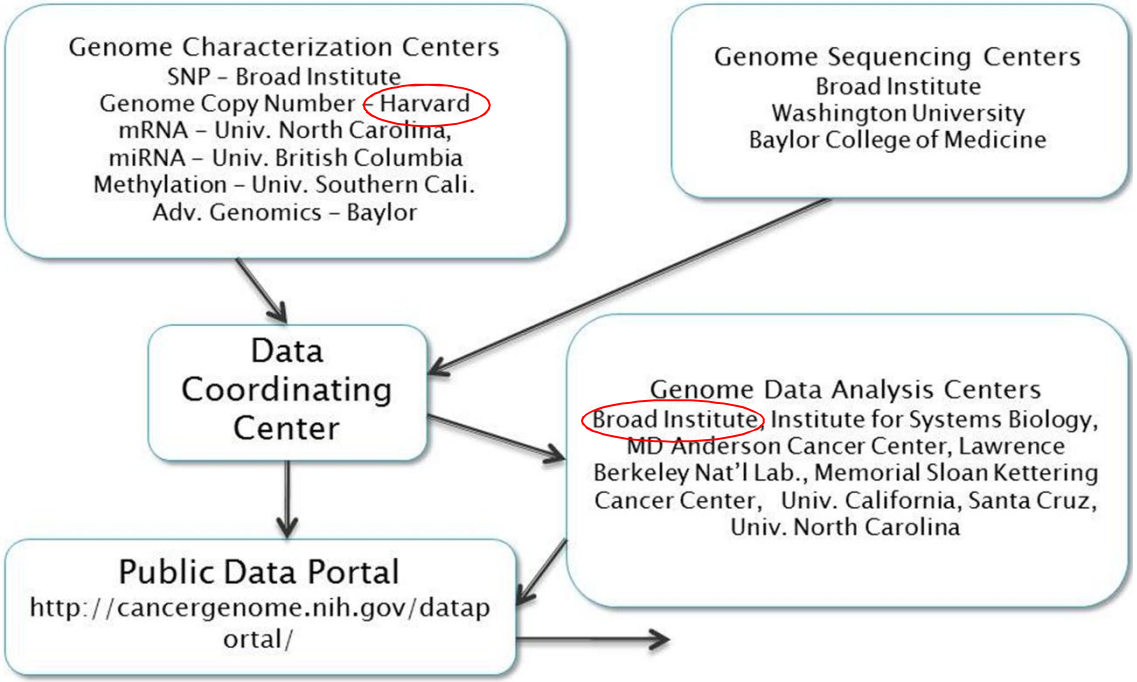
Ayla Ergun
Jim Collins (BU)

RNAi / Human ORFeome

Jesse Bohem (Broad)
Bill Hahn

Marc Vidal (DFCI)
Dave Hills

Acknowledgement



Broad GDAC

Gaddy Getz
Mike Noble
Doug Voet
Gordon Saksena
Mike Lawrence
Lihua Zou
Rui Jing

Juihua Zhang
Spring Liu
Sachet Shukla
Hailei Zhang
Terrence Wu

Nils Gehlenborg
Richard Park
Peter Park

Harvard GCC

Raju Kucherlapati
Jon Seidman
Peter Park

Alexei Protopopov
Ilana Perna
Georgia Ren
Juinhua Zhang
Sachet Shukla

Juihua Zhang
Hailei Zhang

N Sathiamoorthy
Oleg Iartchouk