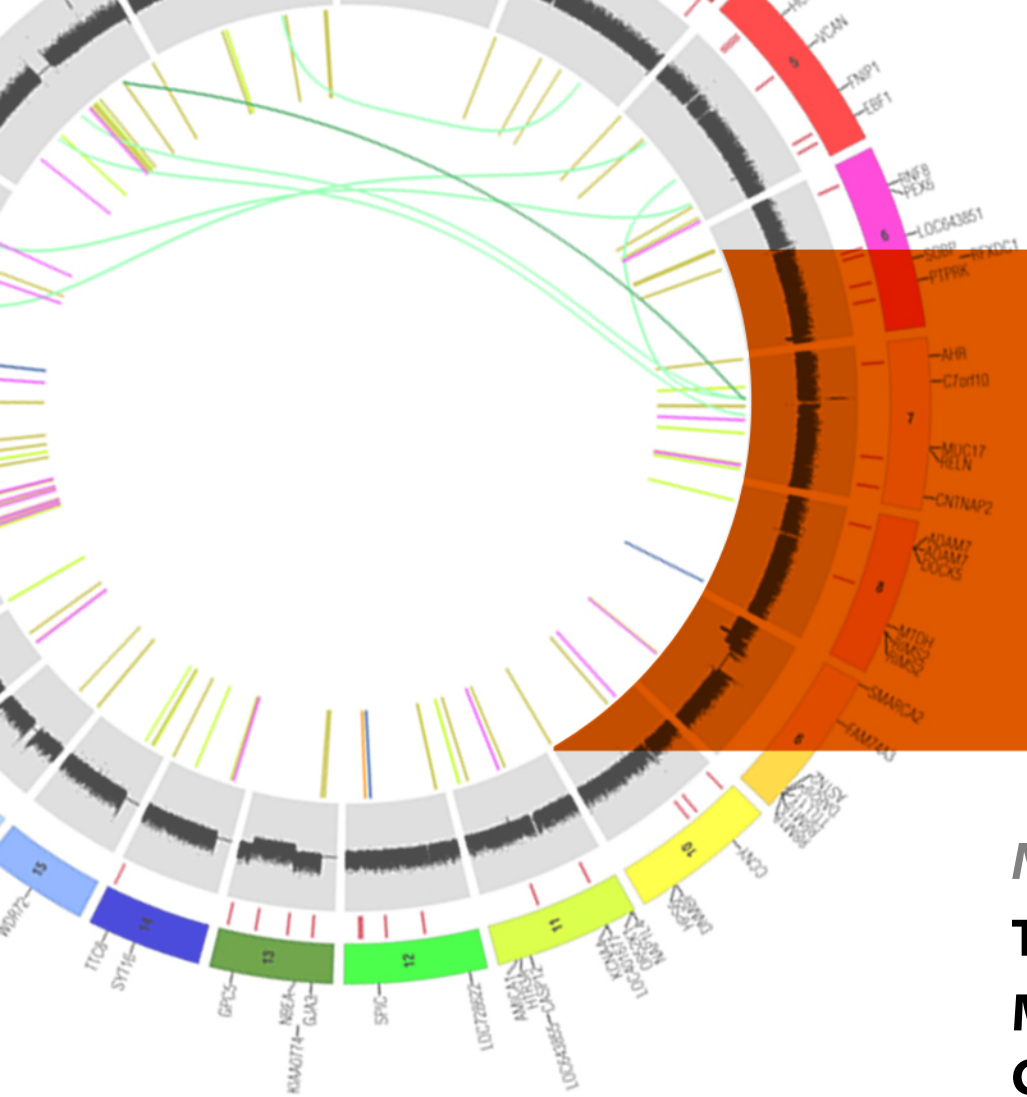


The Cancer Genome Atlas

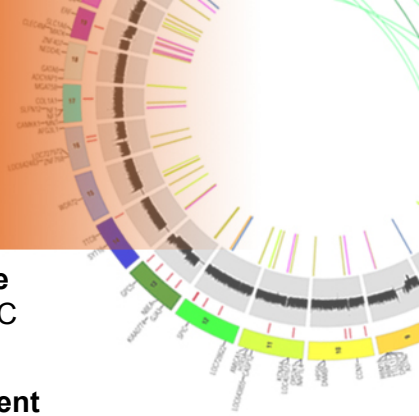
Lung adenocarcinoma genomics

November 28, 2012

TCGA 2nd Annual Symposium

Matthew Meyerson, Ramaswamy Govindan, Steve Baylin, co-chairs

Key participants in TCGA lung cancer analysis group



DNA methylation analysis
Leslie Cope, Johns Hopkins
Ludmila Danilova, Johns Hopkins
Steve Baylin, Johns Hopkins

Gene expression and transcriptome
Neil Hayes, North Carolina
Matt Wilkerson, North Carolina
Gordon Robertson, UBC
Lauren Byers, MD Anderson
Gordon Mills, MD Anderson

DNA sequence analysis
Andrey Sivachenko, Broad
Gad Getz, Broad
Mike Lawrence, Broad
Carrie Sougnez, Broad
Stacey Gabriel, Broad
Eric Lander, Broad
Bryan Hernandez, Broad
Marcin Imielinski, Broad
Elena Helman, Broad
Alice Berger, Broad
Mara Rosenberg, Broad
Juliann Chmielecki Dana-Farber/Broad
Angela Hadjipanayis , Harvard
Raju Kucherlapati, Harvard

Copy number analysis
Gad Getz, Broad
Gordon Saksena, Broad
Andy Cherniack, Broad

Clinical contributors
Bill Travis, MSKCC
Dennis Wigle, Mayo Clinic

Cross-platform Analysis
Chad Creighton, Baylor
Eric Collisson, UCSF
Sam Ng, UCSC
Jacob Kaufman, Vanderbilt
Rileen Sinha, MSKCC
Ronglai Shen, MSKCC
Niki Schultz, MSKCC
Ron Bose, WUSL

Biospecimen Core
Joe Paulauskis, IGC
Bob Penny, IGC

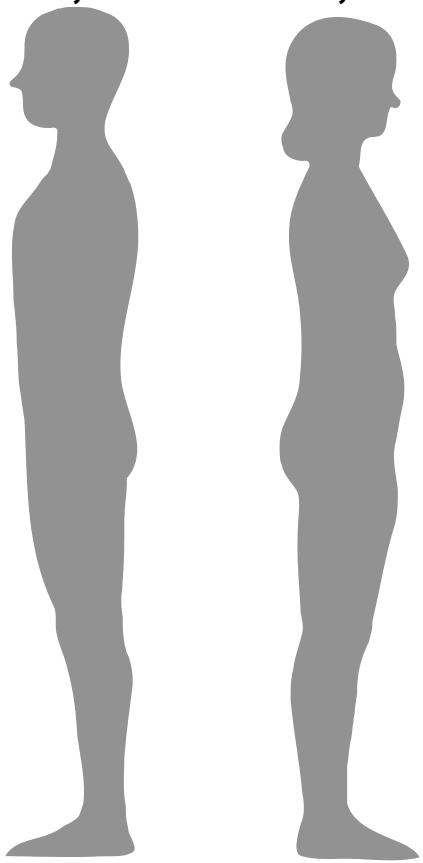
Project management
Kenna Shaw, NCI
Laura Dillon, NCI
Margi Sheth, NCI
Ram Iyer, NCI
Brad Ozenberger, NCI

Tissue collaborators
Malcolm Brock, Johns Hopkins
Ming Tsao, Toronto
Dennis Wigle, Mayo
Val Rusch, Memorial Sloan Kettering
Peter Goldstraw, Royal Brompton
Kwun Fong, Prince Charles
Andrew Godwin, Fox Chase
Maria Raso, MD Anderson
Rajiv Dhir, Pitt
Carl Morrison, Roswell Park

Working group tri-chairs
Ramaswamy Govindan, Washington U
Steve Baylin, Johns Hopkins
Matthew Meyerson, Dana-Farber/Broad

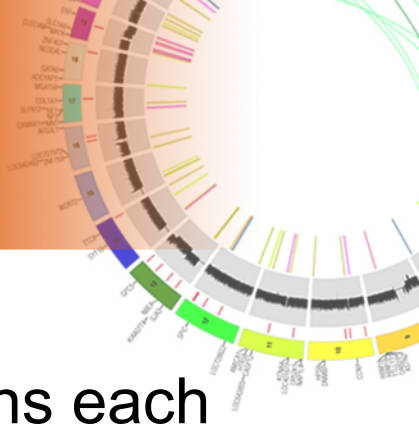
Lung cancers account for over 25% of cancer deaths in the U.S. each year

Men 292,540		Women 269,800	
Lung & bronchus	30%	26%	Lung & bronchus
Prostate	9%	15%	Breast
Colon & rectum	9%	9%	Colon & rectum
Pancreas	6%	6%	Pancreas
Leukemia	4%	5%	Ovary
Liver & intrahepatic bile duct	4%	4%	Non-Hodgkin lymphoma
Esophagus	4%	3%	Leukemia
Urinary bladder	3%	3%	Uterine corpus
Non-Hodgkin lymphoma	3%	2%	Liver
Kidney & renal pelvis	3%	2%	Brain/nervous system
All other sites	25%	25%	All other sites



Source: American Cancer Society, 2009.

Lung adenocarcinoma is the most common form of lung cancer



- Lung cancer kills more than 150,000 Americans each year and more than one million people world-wide
- Major lung cancer histologies are lung adenocarcinoma, squamous cell lung carcinoma, and small cell lung carcinoma
- Lung adenocarcinoma accounts for ~40% of lung cancer diagnoses and ~65,000 deaths each year in the United States.
- While lung cancer is generally associated with smoking, lung adenocarcinoma uniquely often occurs in non-smokers

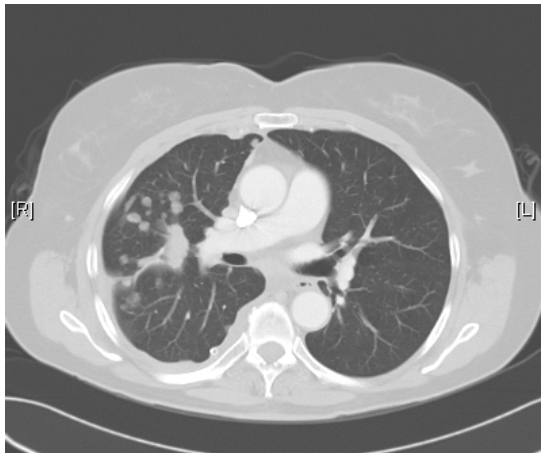
Lung adenocarcinoma: paradigm for molecular subtyping

In recent years, treatments for lung adenocarcinoma have shifted from histology-based strategies to molecular-based strategies.

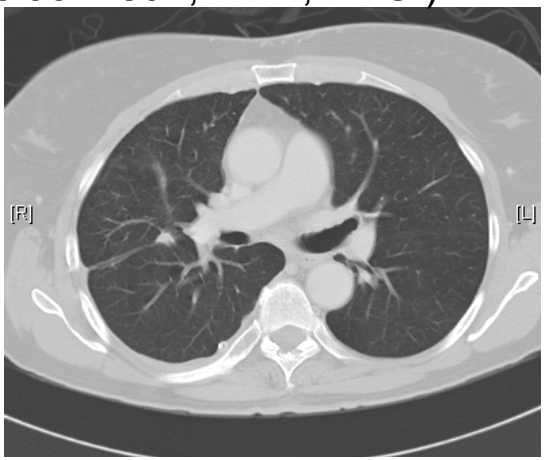
We have made major advances in treatment for lung adenocarcinoma with targeted inhibitors of EGFR (gefitinib, erlotinib) and ALK (crizotinib) thanks to genomic discoveries

Example: a patient with lung adenocarcinoma, with a somatic *EGFR* deletion mutant in exon 19 (thanks to Bruce Johnson, M.D., DFCI)

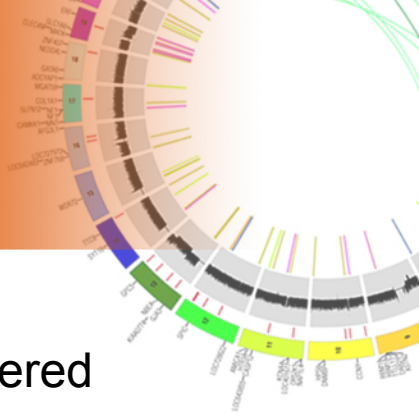
Before treatment



After 2 months erlotinib treatment



Lung adenocarcinoma: previous comprehensive genomic studies



Weir et al., Nature, 2007: copy number analysis of 371 cases, discovered *NKX2-1* and *TERT* amplifications

Ding, Getz et al., Nature, 2008: mutation analysis of 188 cases, discovered mutations of *NF1*, *ATM*, *APC*

Shedden et al., Nat Med, 2008: expression classification of 448 cases

Govindan et al., Cell, 2012: whole genome sequencing of 17 cases, identified smoking/non-smoking signatures

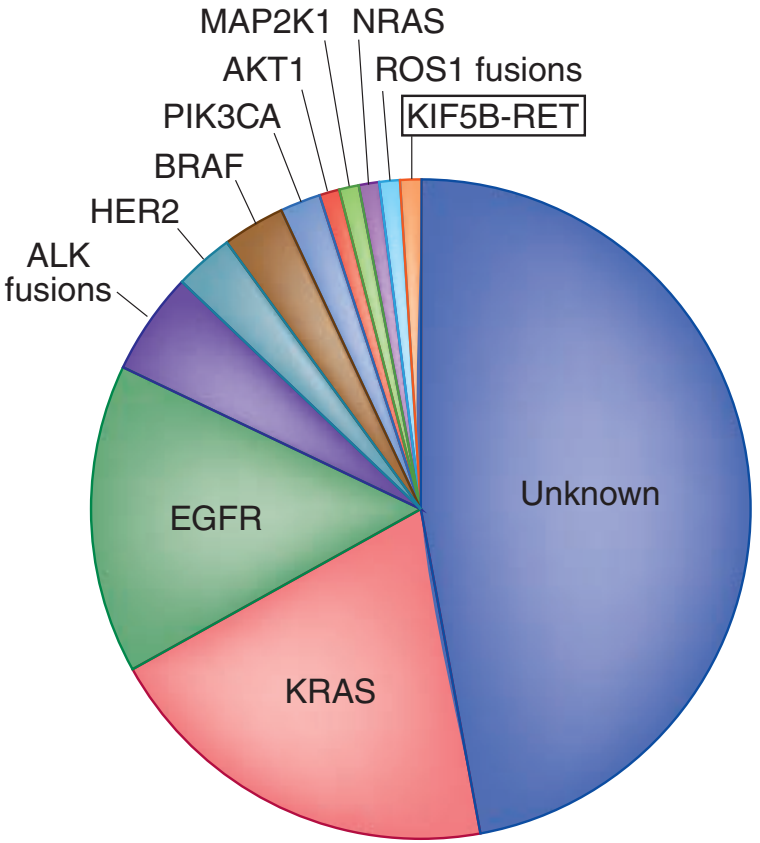
Imielinski, Berger et al., Cell, 2012: whole exome sequencing of 183 cases, identified mutations of *RBM10*, *U2AF1*

Seo et al., Genome Research, 2012: transcriptome sequencing identified recurrent *MET* splicing alterations

Lung adenocarcinoma therapeutic targets: 2012

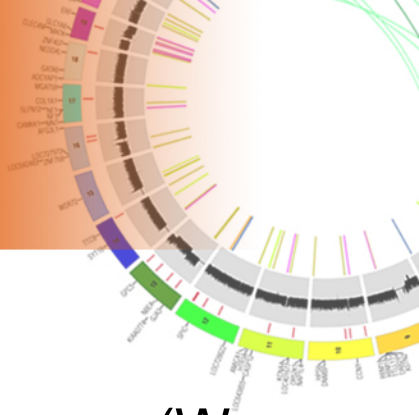
Lung adenocarcinoma drivers

Despite the identification of molecular subsets, more than half of all lung adenocarcinomas lack an identifiable driver mutation.



Adopted from Pao and Hutchinson, 2012

TCGA lung adenocarcinoma project status



- 303 samples collected
- Adenocarcinoma pathology was confirmed for all cases (W. Travis, MSKCC)
- 230 samples included within the data freeze (10/2/12)
  - Majority of samples excluded were due to pathology review—these cases will be included in a subsequent pan-NSCLC report
- High-quality data across multiple platforms for all samples in freeze
  - Next-gen DNA sequencing, RNA-seq, methylation arrays, proteomic analysis, fusion discovery
- 38 sample pairs with whole genome sequence data (planned)
- First face-to-face meeting tomorrow
- Goal: manuscript submission in February to April, 2013

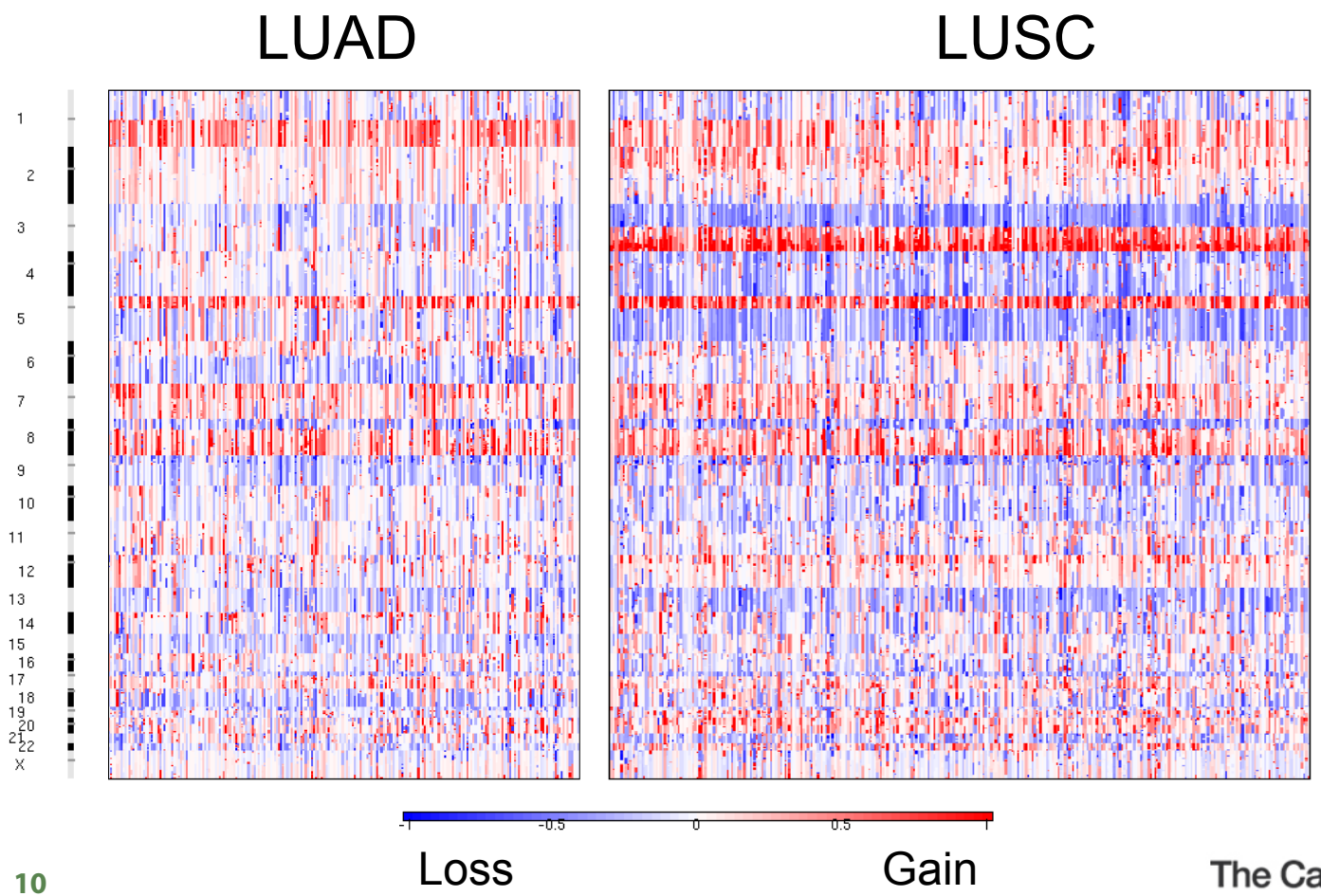
Copy number analysis of lung adenocarcinoma

- Andrew Cherniack, Broad Institute
- Gad Getz, Broad Institute

- 230 tumor/normal DNA pairs, analyzed on Affymetrix SNP 6.0 arrays

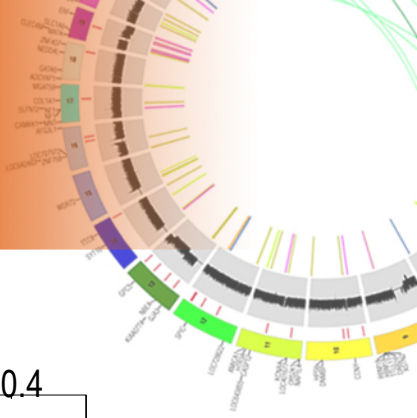
Chromosome arm level copy number in lung adenocarcinoma

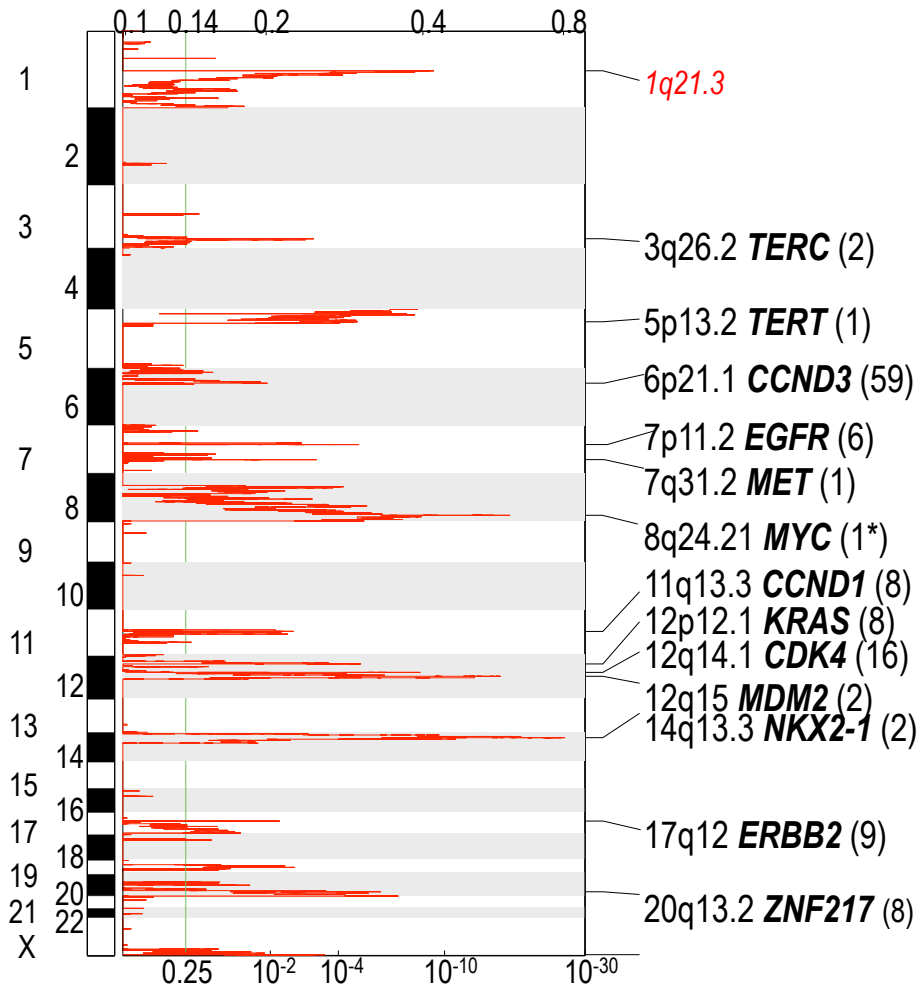
Overall Comparison of Copy Number Changes in TCGA Lung Adenocarcinoma and Squamous Cell Carcinoma

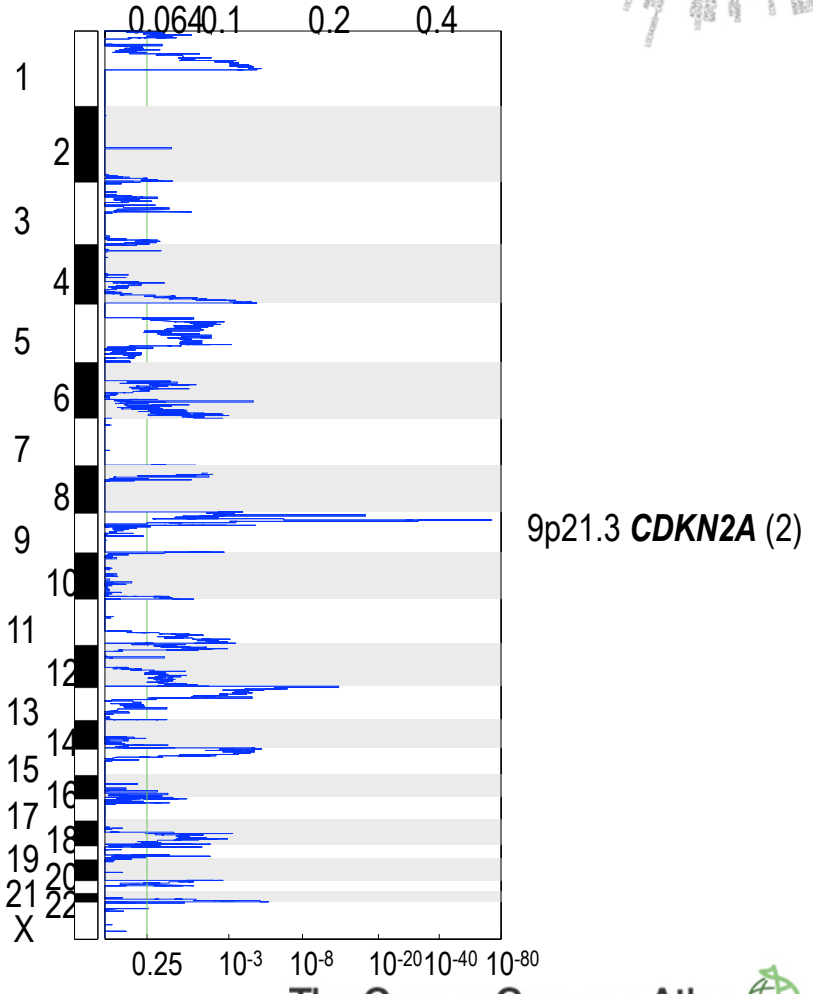


Some differences between LUSC and LUAD.

Focal copy number alterations in lung adenocarcinoma (GISTIC 2.0)







Andy Cherniack

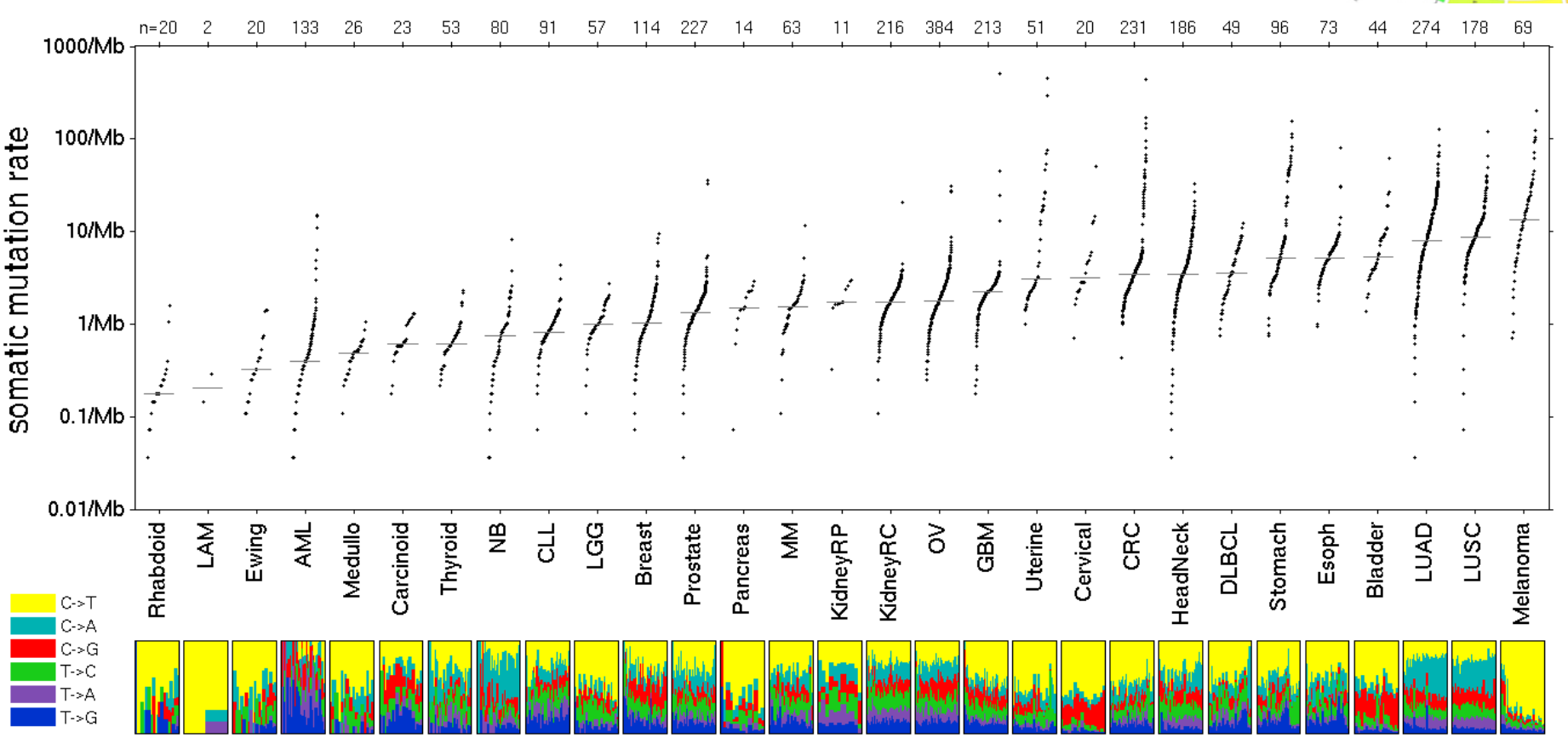
The Cancer Genome Atlas

Exome and RNA sequence analysis of lung adenocarcinoma

- Juliann Chmielecki, Dana-Farber Cancer Institute/Broad Institute
- Mara Rosenberg, Broad Institute
- Matt Wilkerson, University of North Carolina
- Marcin Imielinski, Broad Institute
- Bryan Hernandez, Broad Institute
- Michael Lawrence, Broad Institute
- Neil Hayes, University of North Carolina
- Gad Getz, Broad Institute

- 230 tumor/normal DNA pairs and 230 tumor RNAs, on Illumina paired-end sequencing

Lung adenocarcinoma has a very high rate of somatic mutations



Mike Lawrence, Gad Getz

The high mutation rate poses a major problem in identifying significantly mutated genes

- Known recurrently mutated genes (e.g. *ERBB2*, *CTNNB1*) do not show up as significant regardless of method used
- Expression filtering enriches for real genes
- However, we need to consider a variety of alternative approaches including…
  - Inclusion of functional significance analysis
  - Two-stage statistical analysis
- In the end, a much larger sample size may be required for elucidation of the full population of lung adenocarcinoma causative mutations

Top 21 mutated genes in lung adenocarcinoma (expression-filtered)

	gene	# of mutations	# of patients	# of sites	p value	Median expression
	KEAP1	40	40	38	3.33E-16	10.47
	TP53	113	105	92	6.66E-16	10.50
	STK11	42	40	37	5.55E-15	9.58
	KRAS	69	68	5	7.11E-15	10.47
	RBM10	19	19	18	1.14E-13	10.11
	EGFR	45	33	28	6.01E-12	10.04
*	*ITGAL*	18	17	18	5.52E-09	9.33
	RB1	10	10	10	3.00E-05	9.96
	BRAF	23	22	12	3.68E-05	7.35
*	*HAX1*	6	6	3	6.14E-05	10.88
	ARID1A	17	16	17	9.96E-05	11.51
*	*IL32*	5	5	2	1.31E-04	11.24
	SMARCA4	14	13	14	3.58E-04	11.58
	NF1	30	26	30	2.25E-03	10.83
	U2AF1	8	8	1	3.01E-03	10.36
*	*MGA*	22	19	22	3.14E-03	9.67
*	*BCL9L*	9	9	9	3.62E-03	11.43
	CDKN2A	9	9	9	4.80E-03	7.11
*	*PPPDE1*	2	2	2	5.01E-03	10.77
*	*NKD2*	5	5	5	6.63E-03	7.03
*	*MKI67IP*	5	5	5	6.67E-03	9.66

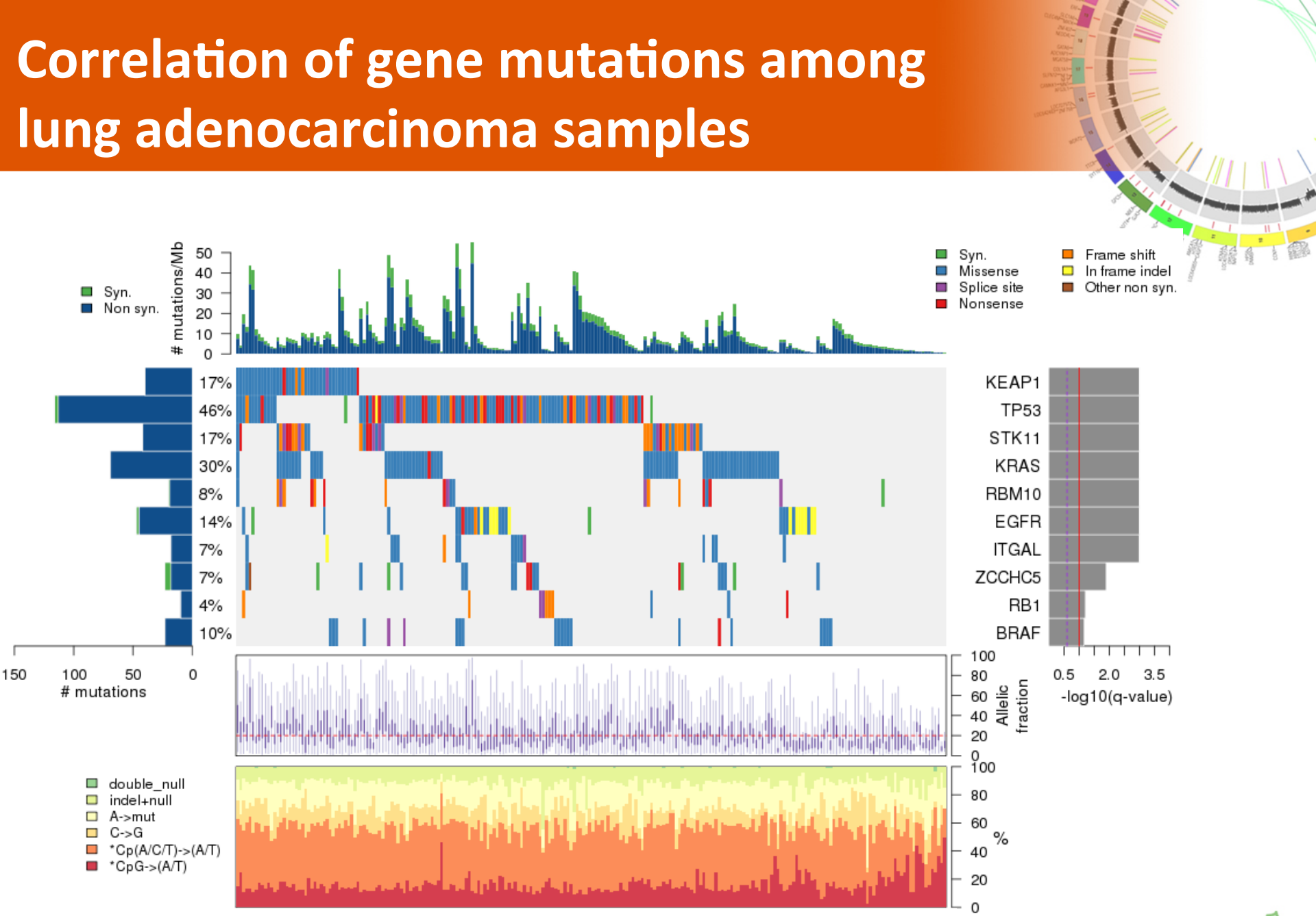
*candidate novel mutated genes

Juliann Chmielecki, Mara Rosenberg

Intriguing mutated gene candidates in lung adenocarcinoma

- *BCL9L*—homolog, *BCL9*, is translocated in B-cell lymphoma and is reported to encode a protein interacting with beta-catenin
- *MGA*—reported suppressor of *MYC*, recently reported to be subject to inactivating mutations in B-cell leukemia/lymphoma
- *MKI67IP*—encodes protein that interacts with Ki-67, encoded by *MKI67*, which is mutated in endometrial cancer

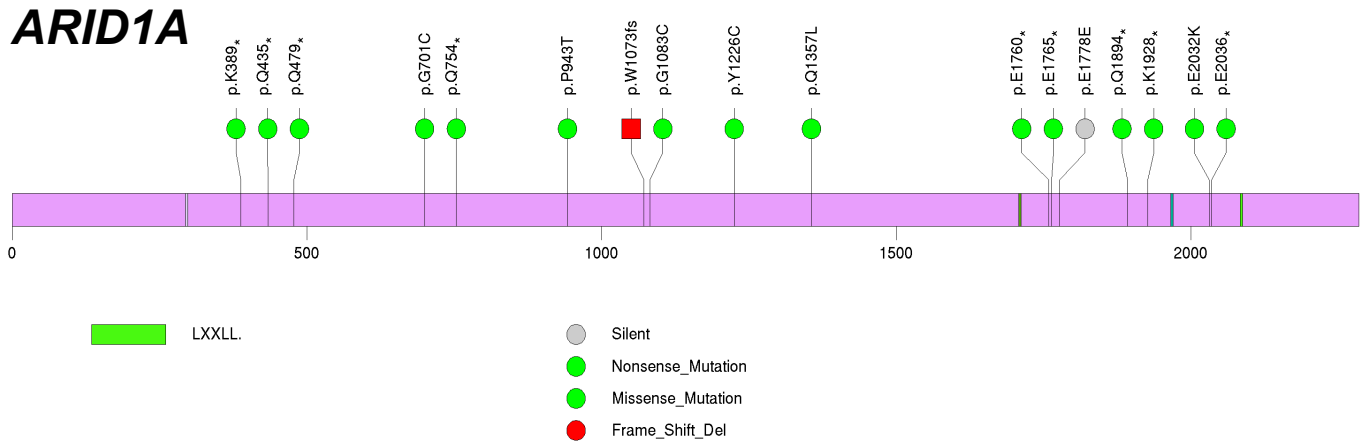
Correlation of gene mutations among lung adenocarcinoma samples

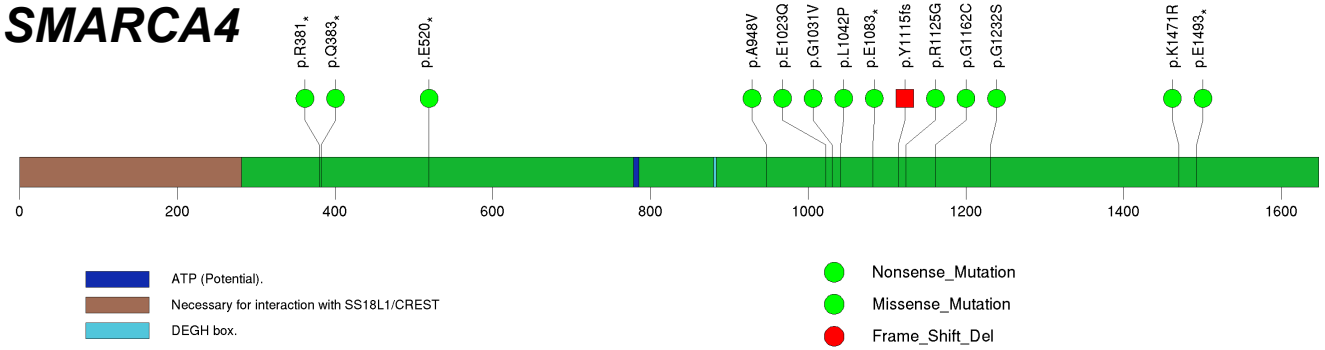


Juliann Chmielecki, Mara Rosenberg

The Cancer Genome Atlas

Recurrent mutations in SWI/SNF chromatin remodeling genes



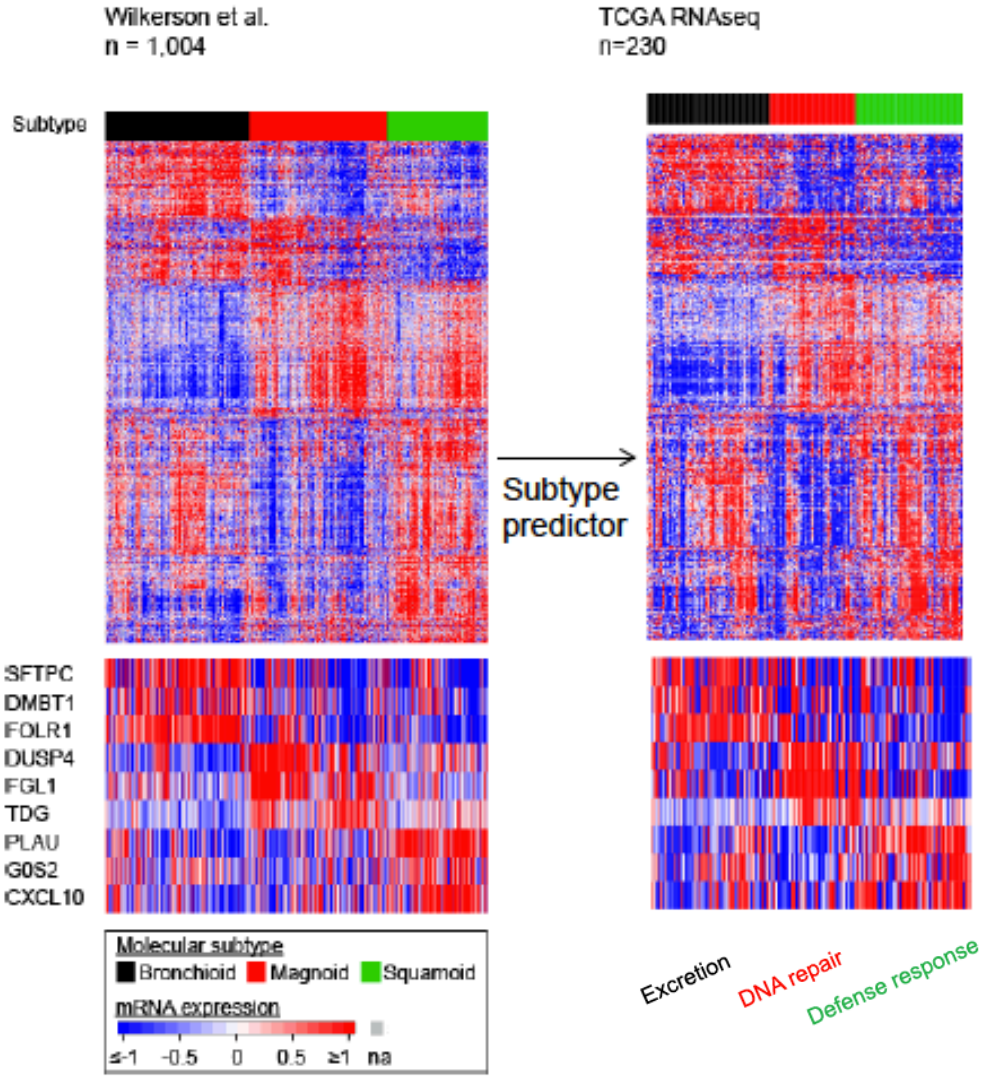


Expression-based classification of lung adenocarcinoma

- Matt Wilkerson, University of North Carolina
- Neil Hayes, University of North Carolina

- 230 tumor RNAs, on Illumina paired-end sequencing

Expression clustering of lung adenocarcinoma shows reproducible classes



Matt Wilkerson,
Neil Hayes

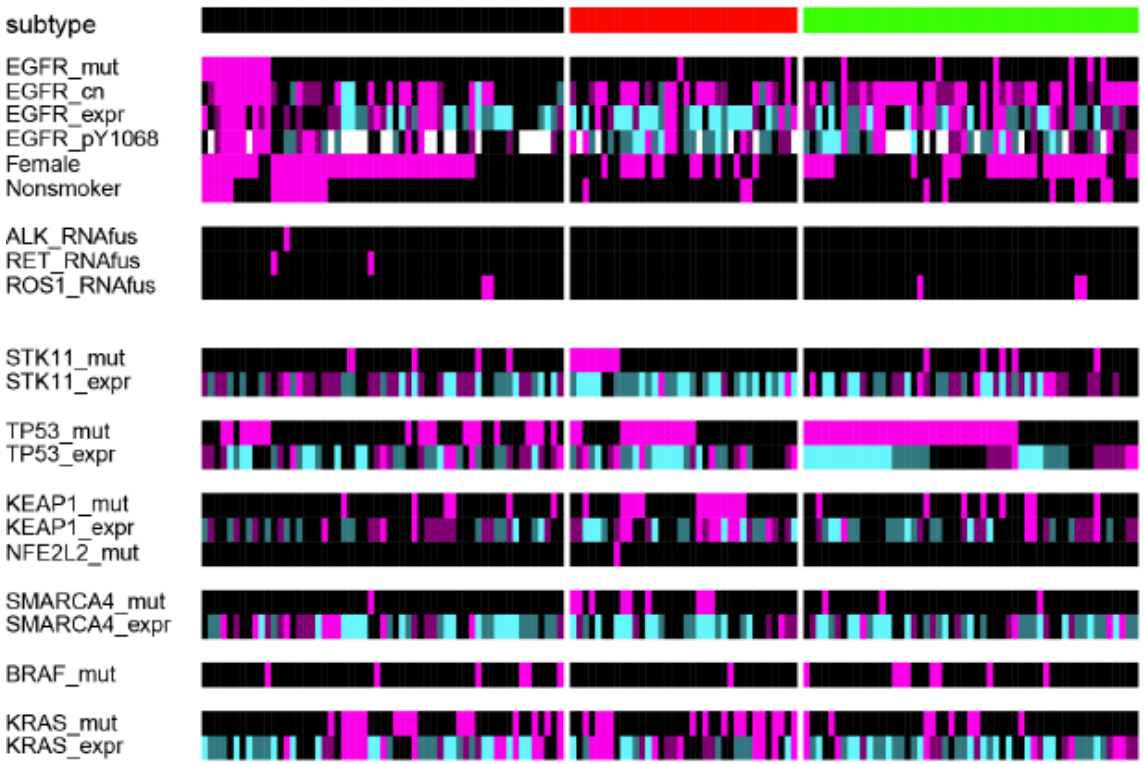
Bronchioid

Magnoid

Squamoid

e Cancer Genome Atlas

Expression subtype integrative analysis



Bronchioid

Magnoid

Squamoid

n=146
Note other ALK fusions
Are in bronchioid,
Not in 146 group.

Matt Wilkerson,
Neil Hayes

Low pass whole genome analysis of lung adenocarcinoma

- Angela Hadjipanayis, Harvard Medical School
- Raju Kucherlapati, Harvard Medical School
- Matt Wilkerson, UNC
- Neil Hayes, UNC

133 tumor/normal DNA pairs for low-pass WGS.

230 tumors for RNA-seq analysis.

Reads were analyzed for structural rearrangements; expression of rearrangements was validated in RNA-seq data.

Fusions identified from RNA-seq involve known fusion partners

– ***ALK***

• TCGA-67-6215	*EML4~ALK*	Bronchioid
• TCGA-67-6216	*EML4~ALK*	Bronchioid
• TCGA-78-7163	*EML4~ALK*	Bronchioid

– ***ROS1***

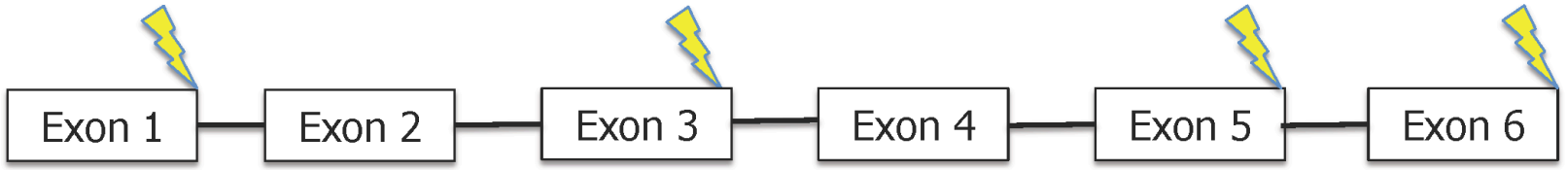
• TCGA-44-2665	*ROS1~CLTC*	Squamoid
• TCGA-05-4426	*SLC34A2~ROS1*	Squamoid
• TCGA-55-6986	*EZR~ROS1*	Bronchioid
• TCGA-64-1680	*CD74~ROS1*	Bronchioid

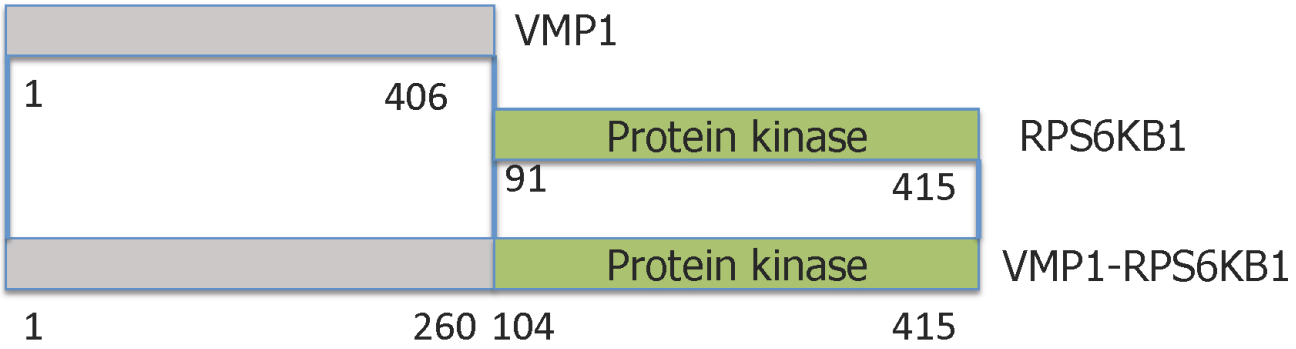
– ***RET***

• TCGA-55-6543	*TRIM33~RET*	Bronchioid
• TCGA-75-6203	*~RET*	Bronchioid

Recurrent VMP1-RPS6KB1 fusion t(17;17)(q23.1;q23)

RPS6KB1:ribosomal protein S6 kinase, 70kDa, polypeptide 1
VMP1: Vacuole Membrane Protein 1



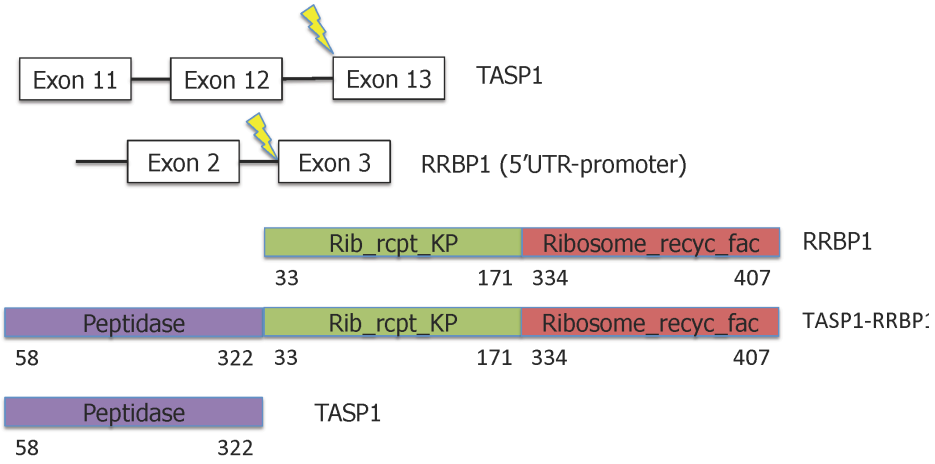


Detected by DNA Sequencing-BreakDancer
7 Tumor Samples/114 RNASeq Samples= ~6.3%

Peptidase fusions in lung adenocarcinoma

TASP1-RRBP1

Split Read Detected
2 Tumor Samples/114 RNASeq files

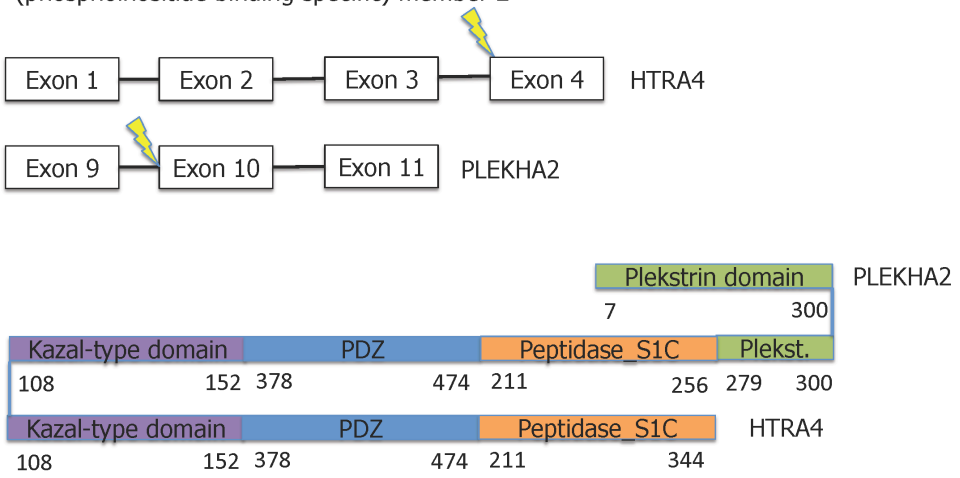


Overexpression of Peptidase?

HTRA4-PLEKHA2

Split Read Detected
2 Tumor Samples/114 RNASeq files

HTRA4: HtrA serine peptidase 4
PLEKHA2:pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 2



Overexpression of Peptidase?

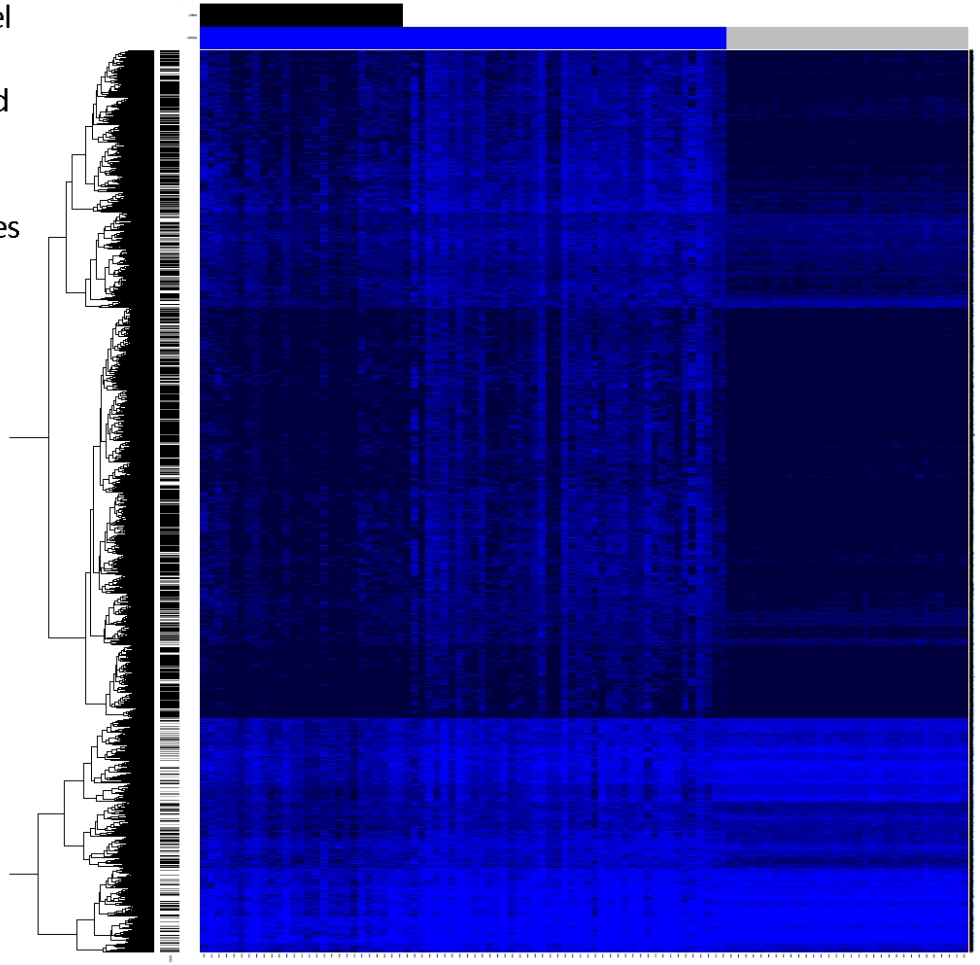
DNA methylation array analysis of lung adenocarcinoma

- Leslie Cope, Johns Hopkins
- Ludmila Danilova, Johns Hopkins
- Steve Baylin, Johns Hopkins

- 181 tumor samples/18 normal DNA pairs, analyzed on Illumina 450K whole genome methylation arrays

CDKN2A inactivated by multiple genomic mechanisms in lung adeno

Compare methylation level in samples where p16 homozygously deleted and methylated
There are around nine thousand significant probes



brighter blue, higher methylation

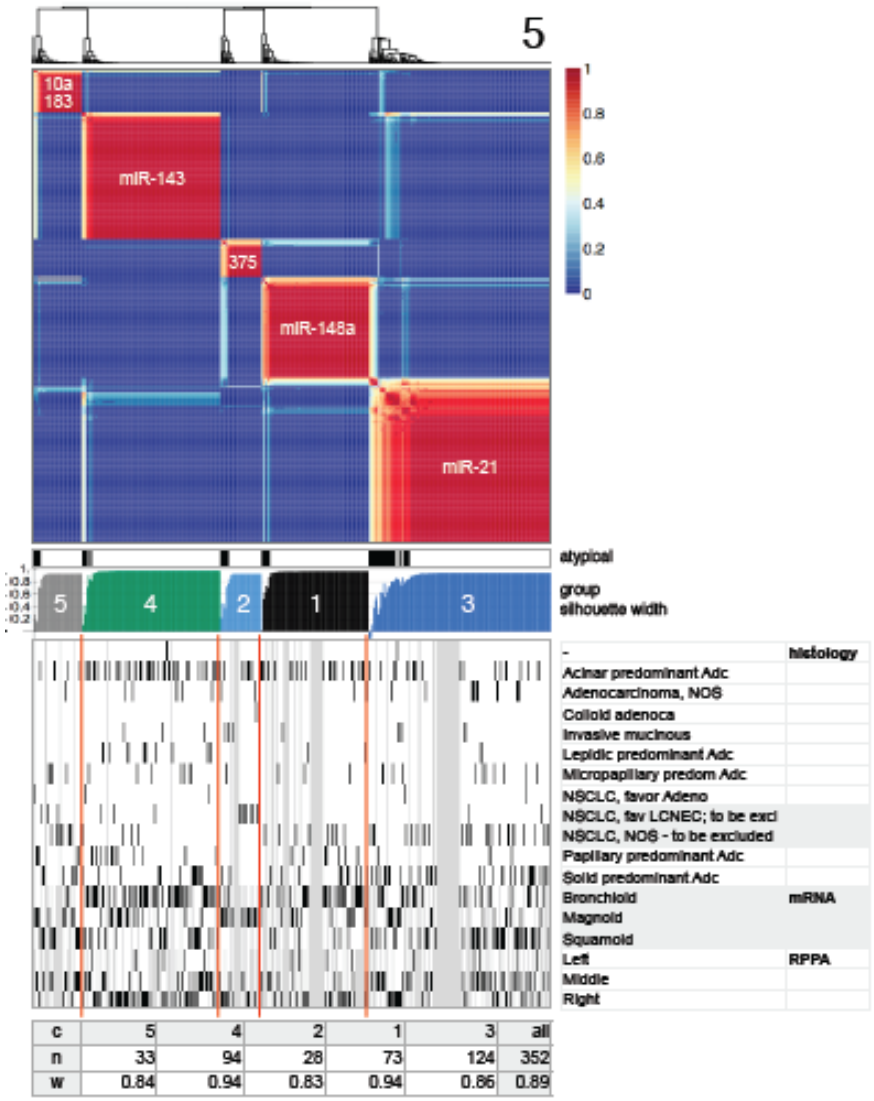
Lung adenocarcinomas frequently lose p16 expression via deletion or methylation

miRNA clustering in lung adenocarcinoma

- Gordon Robertson, BC Cancer Agency
- Andy Chu, BC Cancer Agency

Unsupervised clustering of miRNA sequencing from 352 tumor samples suggested 5 groups.

miRNA clustering in lung adenocarcinoma



c	5	4	2	1	3	all
n	33	94	28	73	124	352
w	0.84	0.94	0.83	0.94	0.86	0.89

miR-10a/183, 143, 375, 148a and 21 discriminate these groups,
and are abundant enough that they are likely biologically active.

miR21 defines one large subset of LUAD

Oncogene Negative Analysis

- Alice Berger, Broad Institute
- Eric Collisson, UCSF
- William Lee, MSKCC
- Marc Ladanyi, MSKCC

- Examined mutational events in tumors lacking RTK activation and other “defining” events (e.g. *H/N/KRAS, EGFR, ERBB2, BRAF* mutation; *ALK*, *RET, ROS* fusion negative)

MutSigCV analysis of "oncogene-positive" and "oncogene-negative" sample sets

Onc pos sample list (n = 139) q < 0.1

rank	gene	q	rank Dneg	npat (pos)	npat (neg)
1	STK11	3.73E-11	1	22	18
2	KRAS	3.73E-11	>5000	67	1
3	TP53	3.73E-11	2	52	53
4	RBM10	3.73E-11	527	15	4
5	EGFR	3.73E-11	>5000	28	5
6	KEAP1	7.07E-05	3	18	22
7	BRAF	2.72E-02	1099	17	5
8	RB1	6.14E-02	161	6	4
9	TMEM169	8.58E-02	4202	4	1

Enriched in oncogene positive group

Onc neg sample list (n = 91) q < 0.1

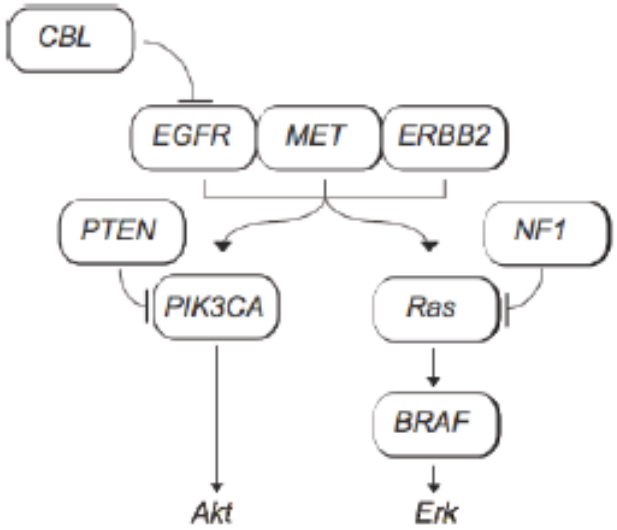
rank	gene	q	rank Dpos	npat (neg)	npat (pos)
1	STK11	5.88E-11	1	18	22
2	TP53	5.88E-11	3	53	52
3	KEAP1	2.42E-10	6	22	18
4	NF1	6.95E-04	>5000	21	5
5	ROPN1L	6.89E-02	2129	4	1

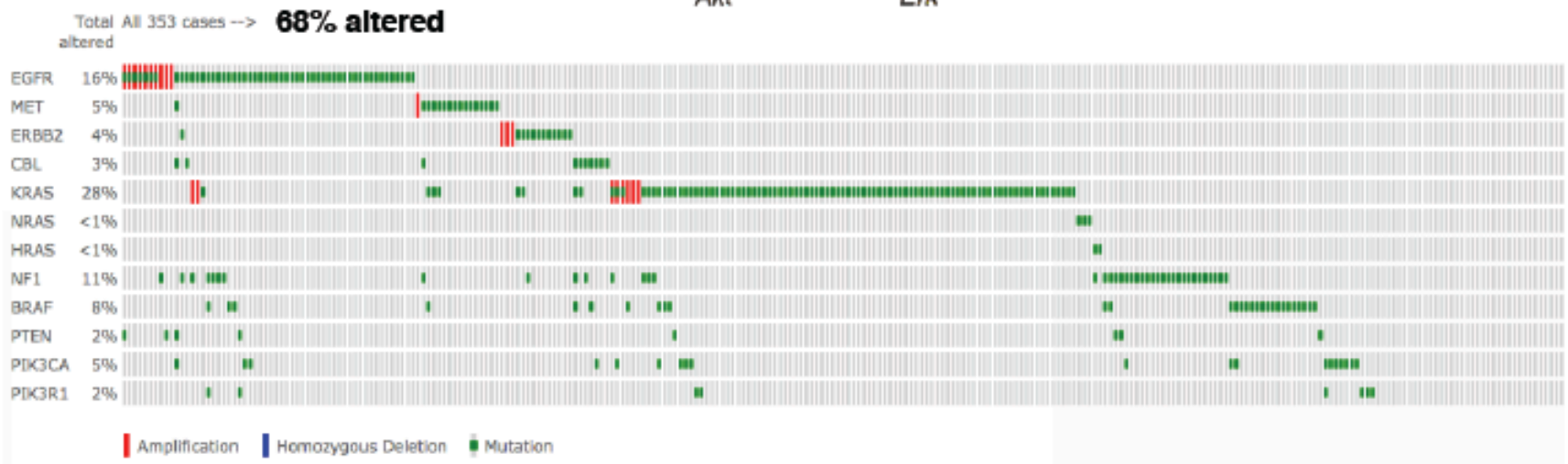
Enriched in oncogene negative group

Integrative cross-platform analysis of lung adenocarcinoma

- Chad Creighton, Baylor
- Eric Collisson, UC San Francisco
- Ron Bose, Washington University
- Niki Schultz, Memorial Sloan-Kettering Cancer Center
- Ted Goldstein, UCSC
- Sam Ng, UCSC

Major deregulation of RTK/RAS/RAF and PI3K/AKT in lung adenocarcinoma





The Cancer Genome Atlas

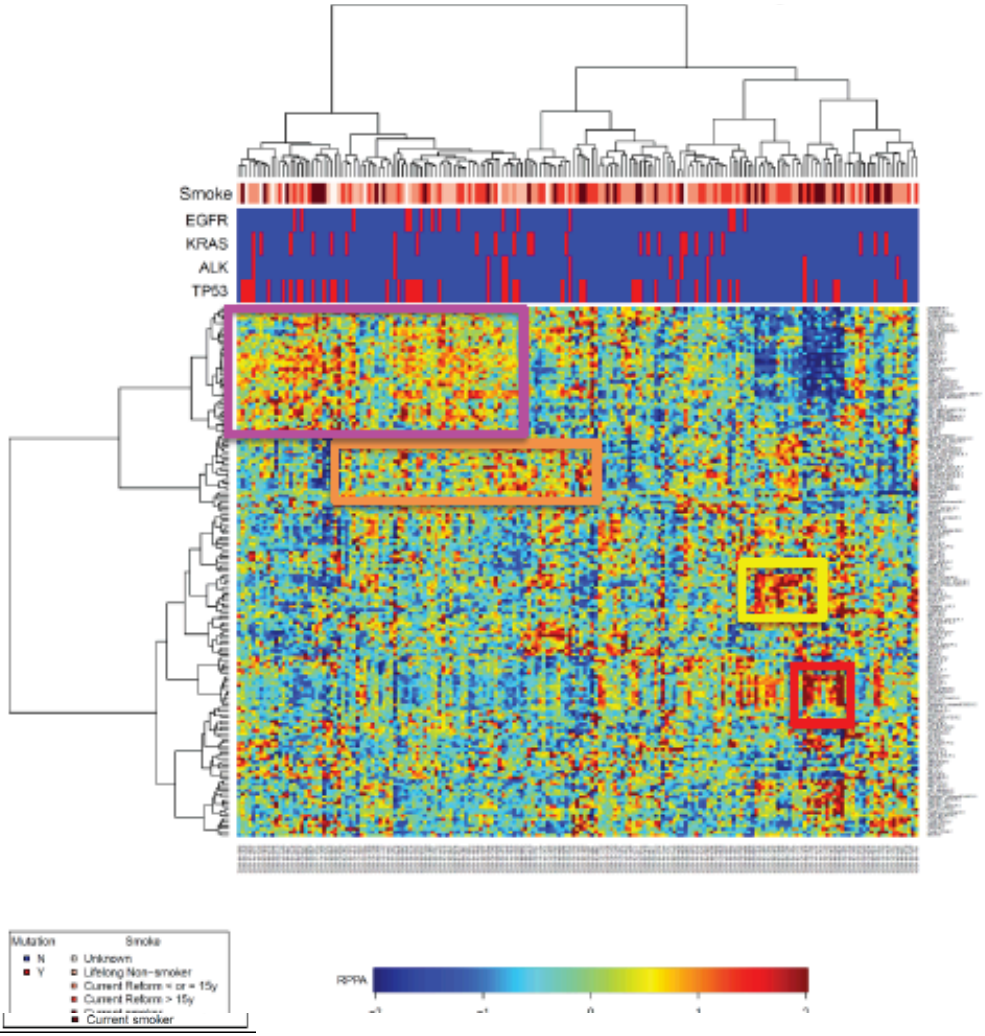
RPPA in lung adenocarcinoma

- Lauren Byers, MD Anderson Cancer Center
- Lixia Diao, MD Anderson Cancer Center
- Gordon Mills, MD Anderson Cancer Center

167 total and phosphorylated proteins quantified by RPPA (reverse phase protein array) in 183 patient tumors.

Tumors cluster into distinct groups that are independent of smoking status.

Lung adeno clusters include RTK activation, MEK activation, and DNA repair groups



RTK activated

YB.1_pS102.R.V
p90RSK_pT359_S363.R.C
p38_pT180_Y182.R.V
Bad_pS112.R.V
PRAS40_pT246.R.V
mTOR_pS2448.R.C
JNK_pT183_pT185.R.V
c.Jun_pS73.R.C
Akt_pT308.R.V
Akt_pS473.R.V
Rb_pS807_S811.R.V
X4E.BP1_pT37.R.V
S6_pS240_S244.R.V
S6_pS235_S236.R.V
Src_pY527.R.V
Src_pY416.R.C
HER2_pY1248.R.V
EGFR_pY1068.R.V

MEK activated

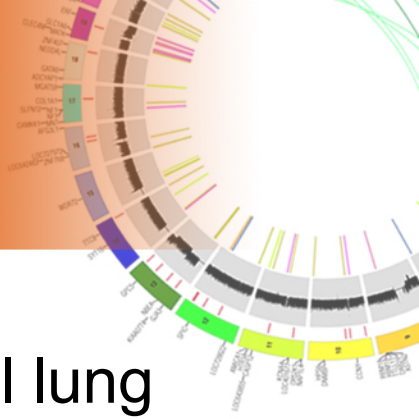
C.Raf_pS338.R.C
MEK1_pS217_S221.R.V
MAPK_pT202_Y204.R.V

DNA repair activated

Stathmin.R.V
Chk1.R.C
Rad51.M.C
TAZ.R.C
Chk2_pT68.R.C
N.Cadherin.R.V
Mre11.R.C
HER3_pY1298.R.C
Bcl.X.R.C
Caspase.9_cleavedD330.R.C

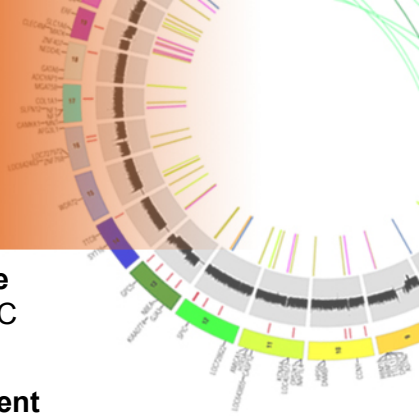
p70S6K.R.V
X53BP1.R.C
P.Cadherin.R.C
eEF2.R.V
ACC1.R.C
ACC_pS79.R.V
beta.Catenin.R.V
alpha.Catenin.M.V
HER2.M.V
PTEN.R.V
JNK2.R.C
eEF2K.R.V
mTOR.R.V
Ku80.R.C
Rad50.M.C
ATM.R.C
Tuberin.R.C
ERK2.R.C
Akt.R.V
STAT5.alpha.R.V
XIAP.R.C
GAB2.R.V
PI3K.p85.R.V
AMPK_pT172.R.V
PDK1_pS241.R.V
GSK3.alpha.beta.M.V
GSK3_pS9.R.V
GSK3.alpha.beta_pS21_S9.R.V
NF.kB.p65_pS536.R.C
Axl.M.C
NF2.R.C
KEAP1.R.C
PKC.alpha_pS657.R.V
PKC.alpha.M.V
PKC.delta_pS664.R.V
PI3K.p110.alpha.R.C
LKB1.M.C
C.Raf.R.V
Syk.M.V
Lck.R.V

Lung adenocarcinoma: conclusions from TCGA analyses thus far



- Both lung adenocarcinoma and squamous cell lung carcinoma have similar copy number profiles.
- Very high mutation rate—challenge to identify novel mutated genes including *MGA.*
- Three distinct expression subtypes identified from RNA-sequencing data.
- Multiple fusions are expressed in lung adenocarcinoma.
- Multiple mechanisms for *CDKN2A* inactivation.
- Distinct miRNA and proteomic clusters.
- Mutational differences between “oncogene positive” and “oncogene negative” subtypes including enrichment of *NF1* mutation in oncogene-negative group.

Key participants in TCGA lung cancer analysis group



DNA methylation analysis
Leslie Cope, Johns Hopkins
Ludmila Danilova, Johns Hopkins
Steve Baylin, Johns Hopkins

Gene expression and transcriptome
Neil Hayes, North Carolina
Matt Wilkerson, North Carolina
Gordon Robertson, UBC
Lauren Byers, MD Anderson
Gordon Mills, MD Anderson

DNA sequence analysis
Andrey Sivachenko, Broad
Gad Getz, Broad
Mike Lawrence, Broad
Carrie Sougnez, Broad
Stacey Gabriel, Broad
Eric Lander, Broad
Bryan Hernandez, Broad
Marcin Imielinski, Broad
Elena Helman, Broad
Alice Berger, Broad
Mara Rosenberg, Broad
Juliann Chmielecki, Dana-Farber/Broad
Angela Hadjipanayis , Harvard
Raju Kucherlapati, Harvard

Copy number analysis
Gad Getz, Broad
Gordon Saksena, Broad
Andy Cherniack, Broad

Clinical contributors
Bill Travis, MSKCC
Dennis Wigle, Mayo Clinic

Cross-platform Analysis
Chad Creighton, Baylor
Eric Collisson, UCSF
Sam Ng, UCSC
Jacob Kaufman, Vanderbilt
Rileen Sinha, MSKCC
Ronglai Shen, MSKCC
Niki Schultz, MSKCC
Ron Bose, WUSL

Biospecimen Core
Joe Paulauskis, IGC
Bob Penny, IGC

Project management
Kenna Shaw, NCI
Laura Dillon, NCI
Margi Sheth, NCI
Ram Iyer, NCI
Brad Ozenberger, NCI

Tissue collaborators
Malcolm Brock, Johns Hopkins
Ming Tsao, Toronto
Dennis Wigle, Mayo
Val Rusch, Memorial Sloan Kettering
Peter Goldstraw, Royal Brompton
Kwun Fong, Prince Charles
Andrew Godwin, Fox Chase
Maria Raso, MD Anderson
Rajiv Dhir, Pitt
Carl Morrison, Roswell Park

Working group tri-chairs
Ramaswamy Govindan, Washington U
Steve Baylin, Johns Hopkins
Matthew Meyerson, Dana-Farber/Broad