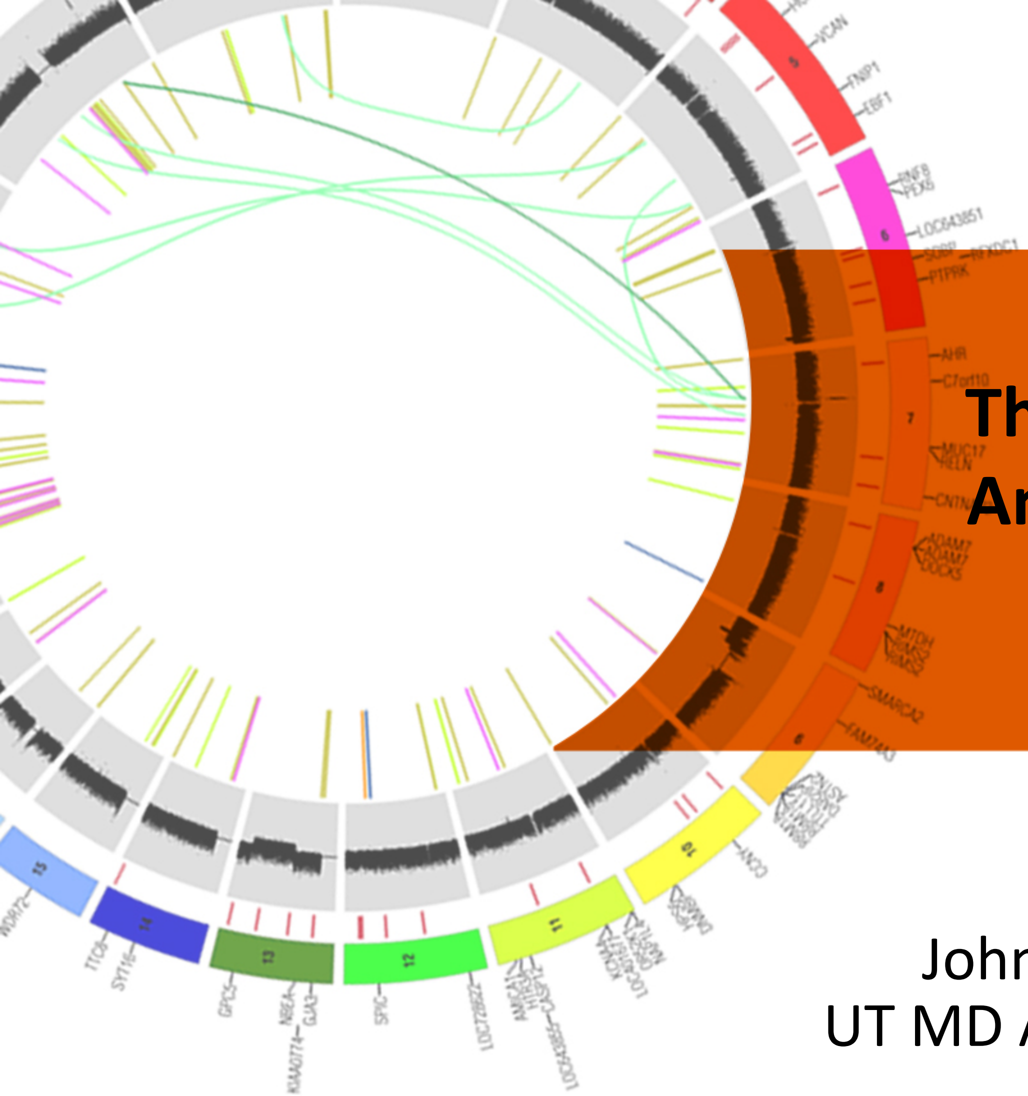
The Cancer Genome Atlas



The Bladder Cancer (BLCA) Analysis Working Group: a Progress Report

John N. Weinstein, M.D., Ph.D.
UT MD Anderson Cancer Center GDAC

THE UNIVERSITY OF TEXAS
MD Anderson
~~Cancer~~ Center

American Cancer Society Statistics on BLCA*

- 90% over age 55, > half are over age 73
- Four times more likely in men than women
- Lifetime risk for men 1 in 27; Women 1 in 85
- 4th most common cancer in men

U.S. spends 2.2 billion dollars a year in health care for bladder cancer patients compared with 1.4 billion for prostate cancer**

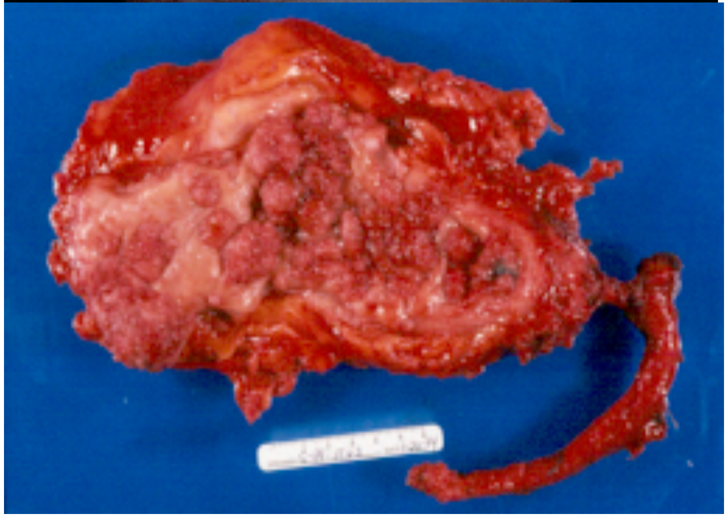
*American Cancer Society. Cancer Facts and Figures 2008. Atlanta, Ga: American Cancer Society; 2008

**Agency for Health Care Policy & Research, 1995 and 1999, U.S. Public Health Service (HHS)

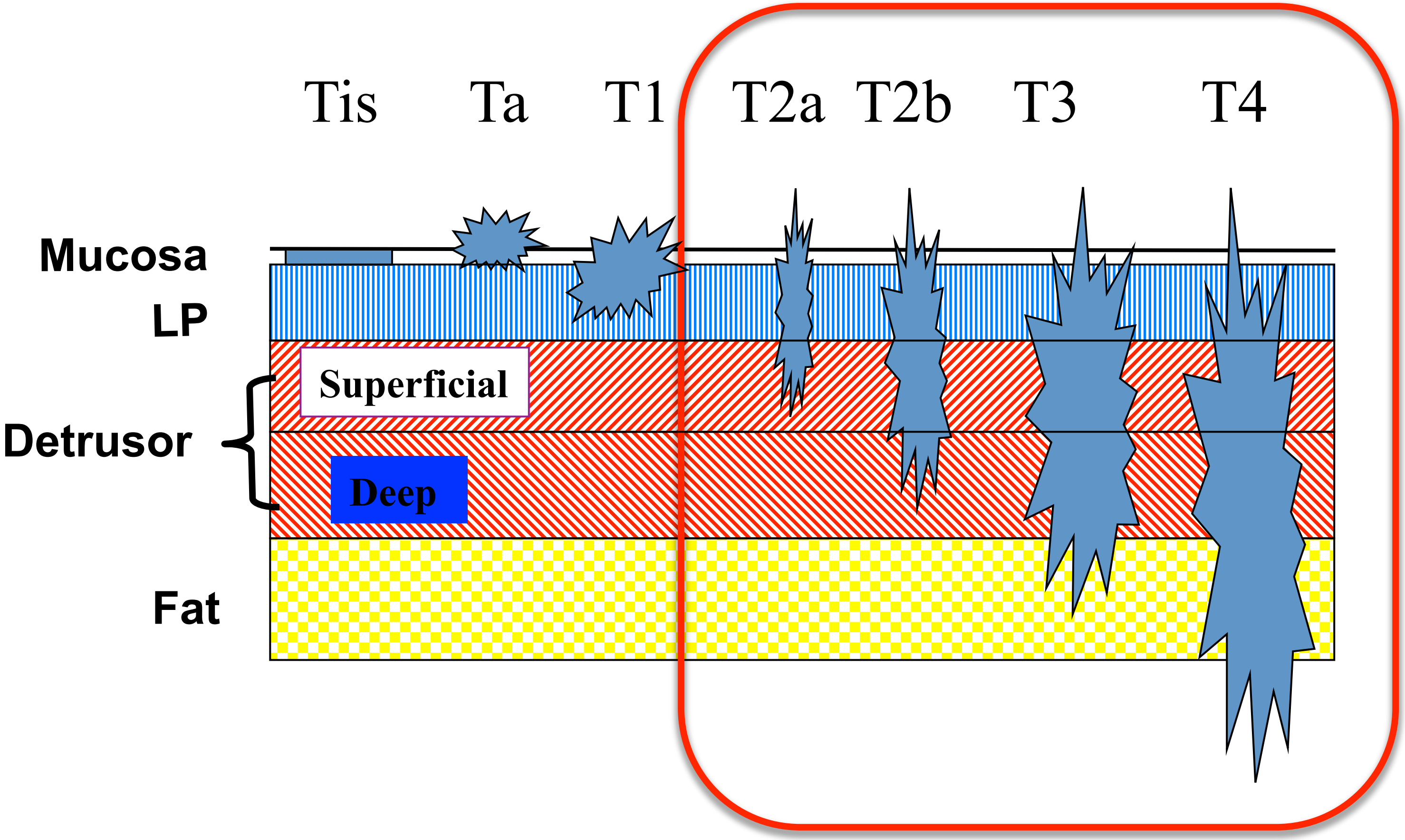
Low- and High-Grade BLCA

Low-grade: superficial, less likely to invade or metastasize, frequently reappears after resection but amenable to therapy, low mortality

High-grade: propensity to invade and metastasize, high mortality when invasive, but good response to treatment if detected early



TNM Staging



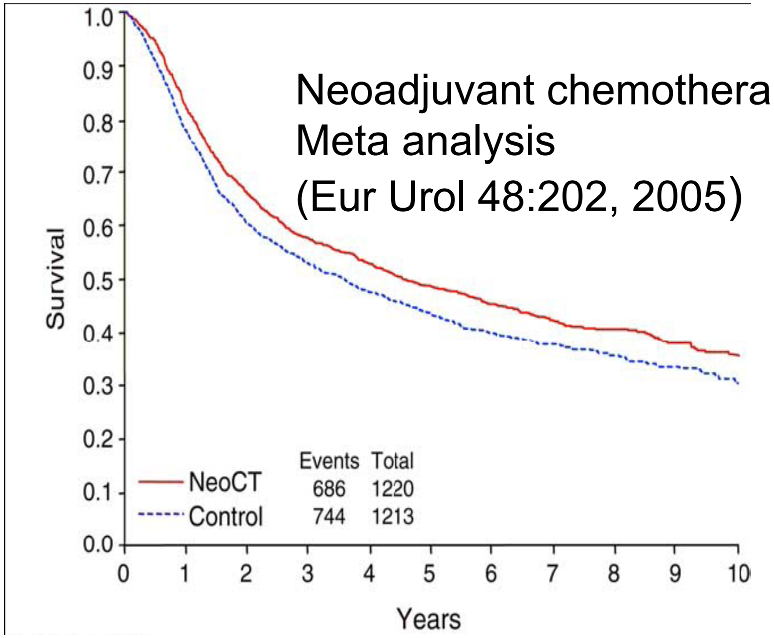
Muscle-invasive

Muscle Invasive Bladder Cancer

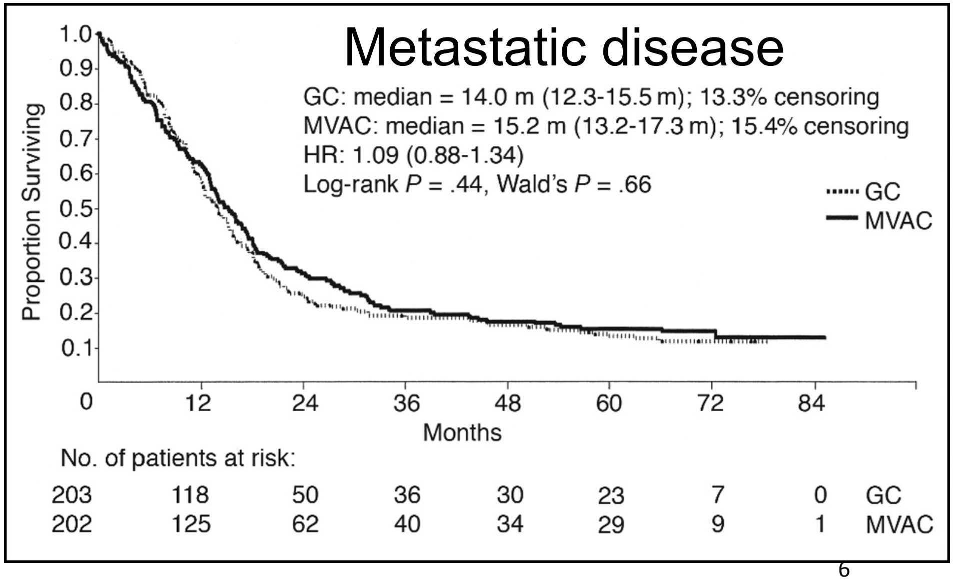
- 15-20% of patients with Ta, T1 or Tis cancer progress to muscle invasion
- 80% of patients with muscle invasive cancer present de novo
- Distant metastases most common cause of treatment failure
  - Present at the time of cystectomy
  - Occurs in 40-50% within 2 years without additional therapy

Bladder Cancer Treatment

- Cisplatin based multi-agent chemotherapy standard of care for neoadjuvant prior to cystectomy and for measurable metastatic disease
- No new FDA approved drugs for muscle invasive BLCA cancer in over two decades!



Neoadjuvant chemotherapy
Meta analysis
(Eur Urol 48:202, 2005)


Metastatic disease

JCO 23:4602, 2005

BLCA Working Group Chairs and Coordinators

Co-Chairs:	Seth Lerner John Weinstein
Data Coordinator:	Chad Creighton
Analysis Coordinator:	Rehan Akbani
Admin. Coordinator:	Margi Sheth
Manuscript Coordinator:	Maggie Morgan
Chromatin Remodeling Chairs:	David Kwiatkowsky, Jonathan Rosenberg, Peter Laird

BLCA Working Group Members

- Seth Lerner (co-chair)
- John Weinstein (co-chair)
- Rehan Akbani
- Hikmat Al-Ahmadie
- Pavana Anur
- Jessie Au
- Keith Baggerly
- Dean Bajorin
- Steve Baylin
- Michael Blute
- Bernard Bochner
- Reanne Bowlby
- Andrew Cherniack
- John Cheville
- Andy Chu
- Chad Creighton
- Bogdan Czerniak
- Kyle Ellrott
- Igor Frank
- Gad Getz
- Angela Hadjipanayis
- Donna Hansel
- Neil Hayes
- Toshinori Hinoue
- Katie Hoadley
- Peter Jones
- Jaegil Kim
- William Kim
- Theresa Koppie
- Raju Kucherlapati
- David Kwiatkowski
- Peter Laird
- Semin Lee
- Bradley Leibovich
- Monica Liebert
- Yuexin Lu
- David McConkey
- Matthew Meyerson
- Gordon Mills
- Matt Milowsky
- Andy Mungall
- Sam Ng
- Peter O'Donnell
- Alex Pastuszak
- Evan Paull
- Alexei Protopopov
- Victor Reuter
- Brian Robinson
- Jonathan Rosenberg
- Gordon Robertson
- Niki Schultz
- Hui Shen
- Ilya Shmulevich
- Payal Sipahimalani
- Eila Skinner
- Carolyn Smith
- Carrie Sougnez
- Paul Spellman
- Dominik Stoll
- Josh Stuart
- Xiaoping Su
- Youting Sun
- Roel Verhaak
- Nils Weinhold
- Da Yang
- Li Zhang
- Lihua Zou

BCR Representatives

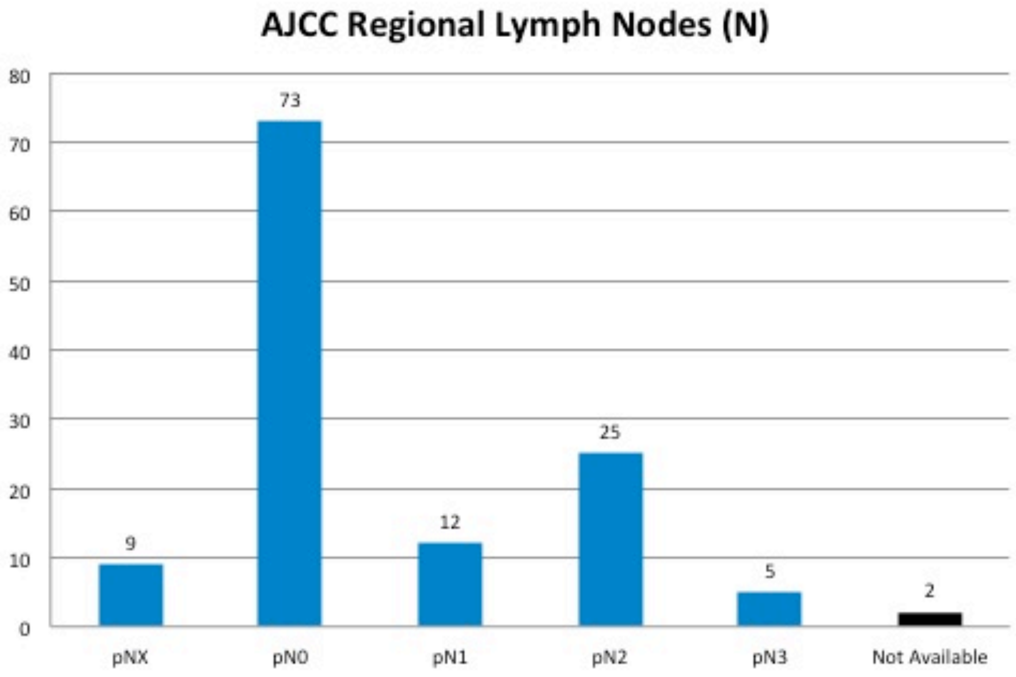
- Jay Bowen (BCR)
- Julie Gastier-Foster (BCR)
- Tara Lichtenberg (BCR)
- Bob Penny (BCR)

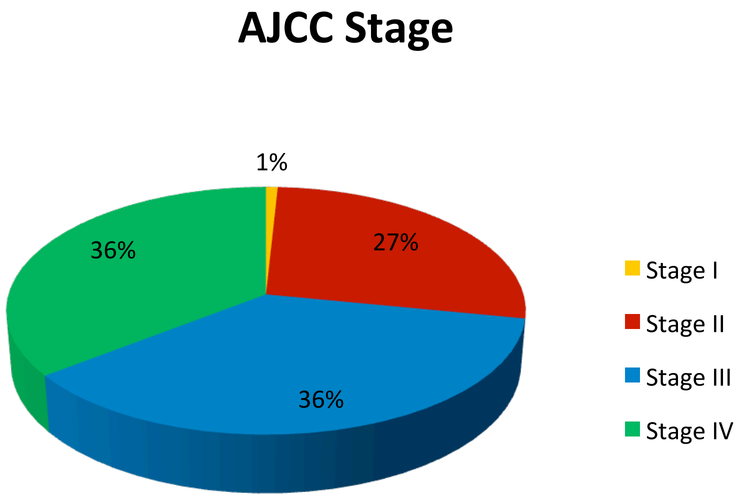
Clinical Data on TCGA Samples

- Muscle invasive urothelial cancer
  - Mixed histology allowed up to 9%
- 126 samples in data freeze for marker paper
  - 153 qualified; 138 in pipeline
- Gender: male 72%; female 28%
- Caucasian 85%
- Median age 69 (34-88)
- Follow up (n = 126)
  - Median 209 days (0-131.2 months)
- Event rate
  - Progression – 10 (data available for 31 patients)
  - Deaths 35 – (data available for all 126 patients)

Staging of TCGA BLCA Cases (N = 124)

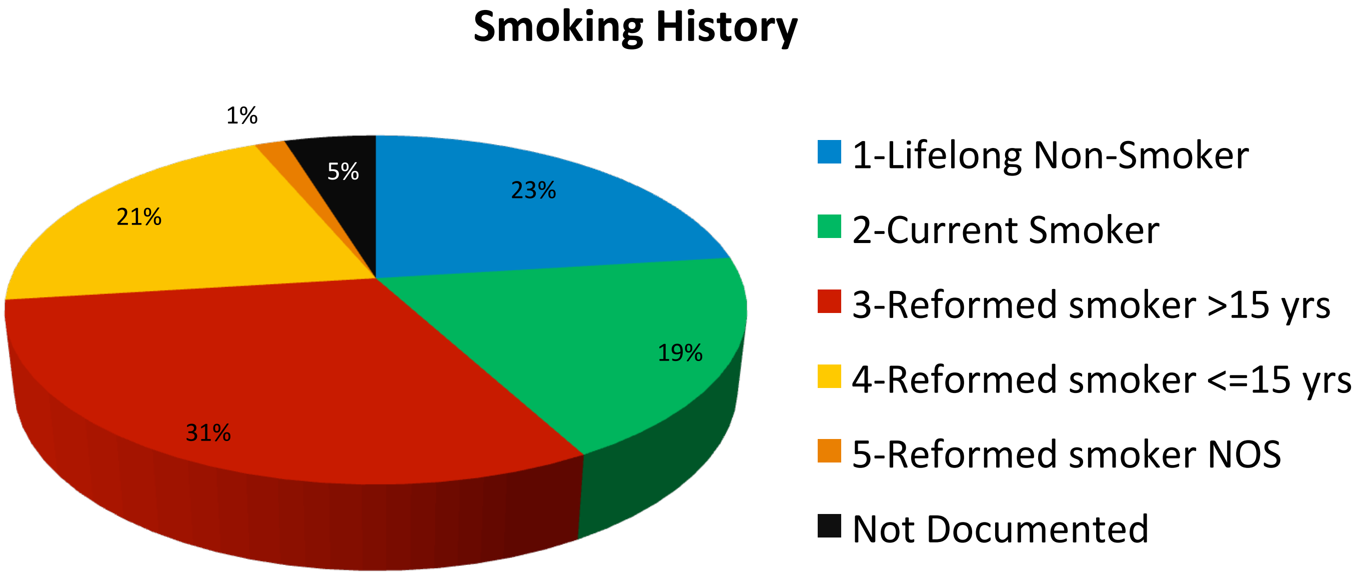
	N0	N+	Nx
pT1	0	0	1
pT2	26	8	8
pT3	41	24	-
pT4	6	10	-
Total	73	42	9





Cigarette Smoking History

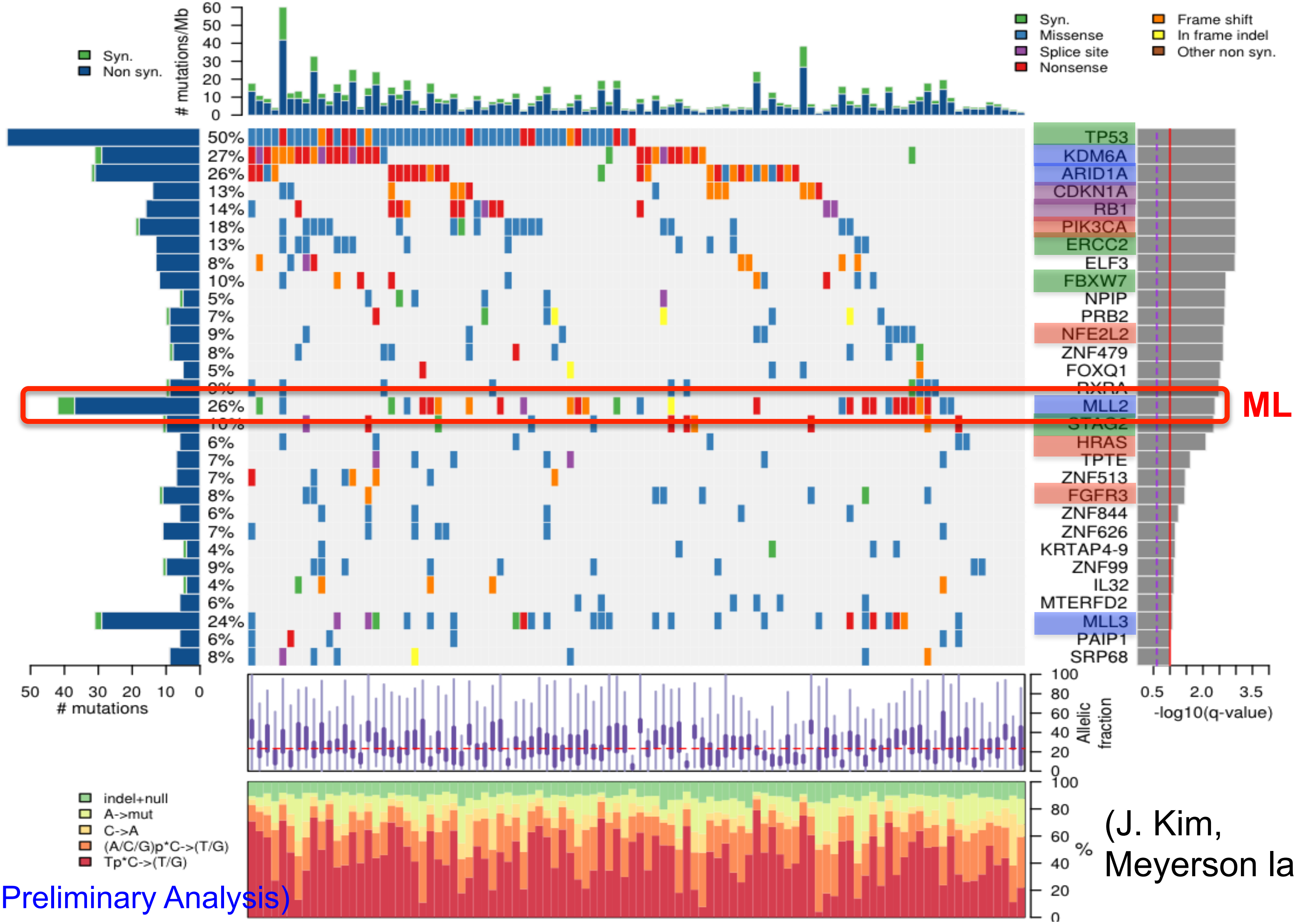
N = 126



Status of the BLCA Project

- Accrual a limiting factor but recently accelerated.
- Productive Face-to-Face held in Houston in October 2012
- Data freeze at 126 tumor samples (plus normals) in December 2012
- So data and analyses presented here will be an "interim look"
- Fast progress on Marker Paper

Significantly mutated genes



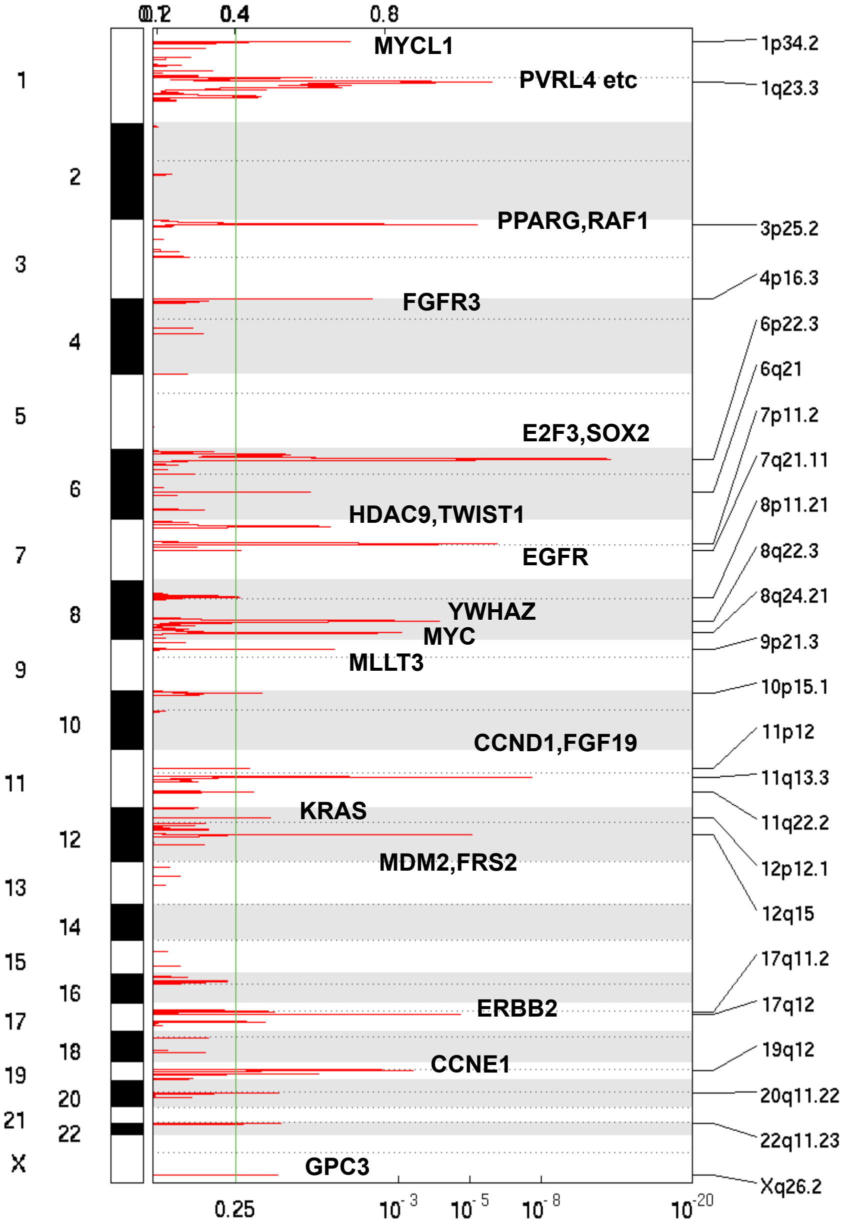
MLL2

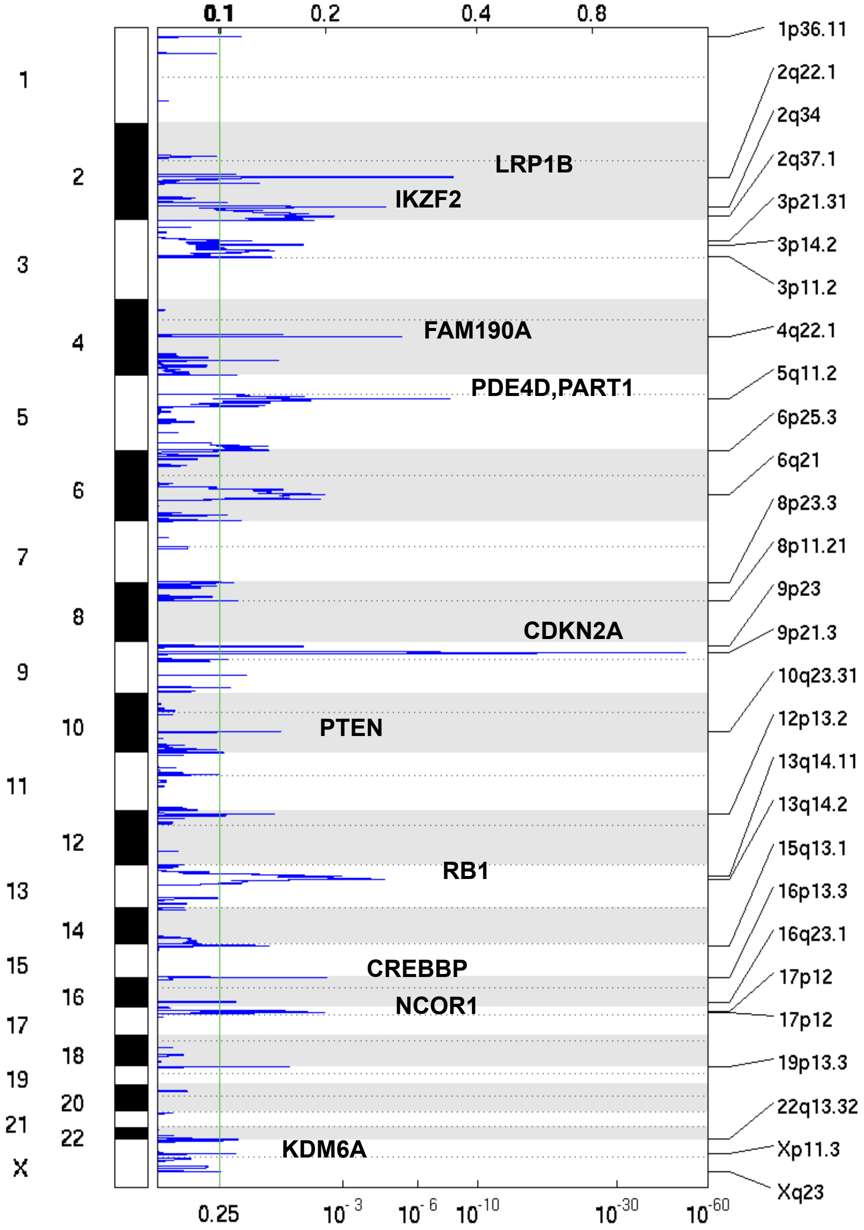
(J. Kim, Meyerson lab)

(Preliminary Analysis)

GISTIC2.0: focal events

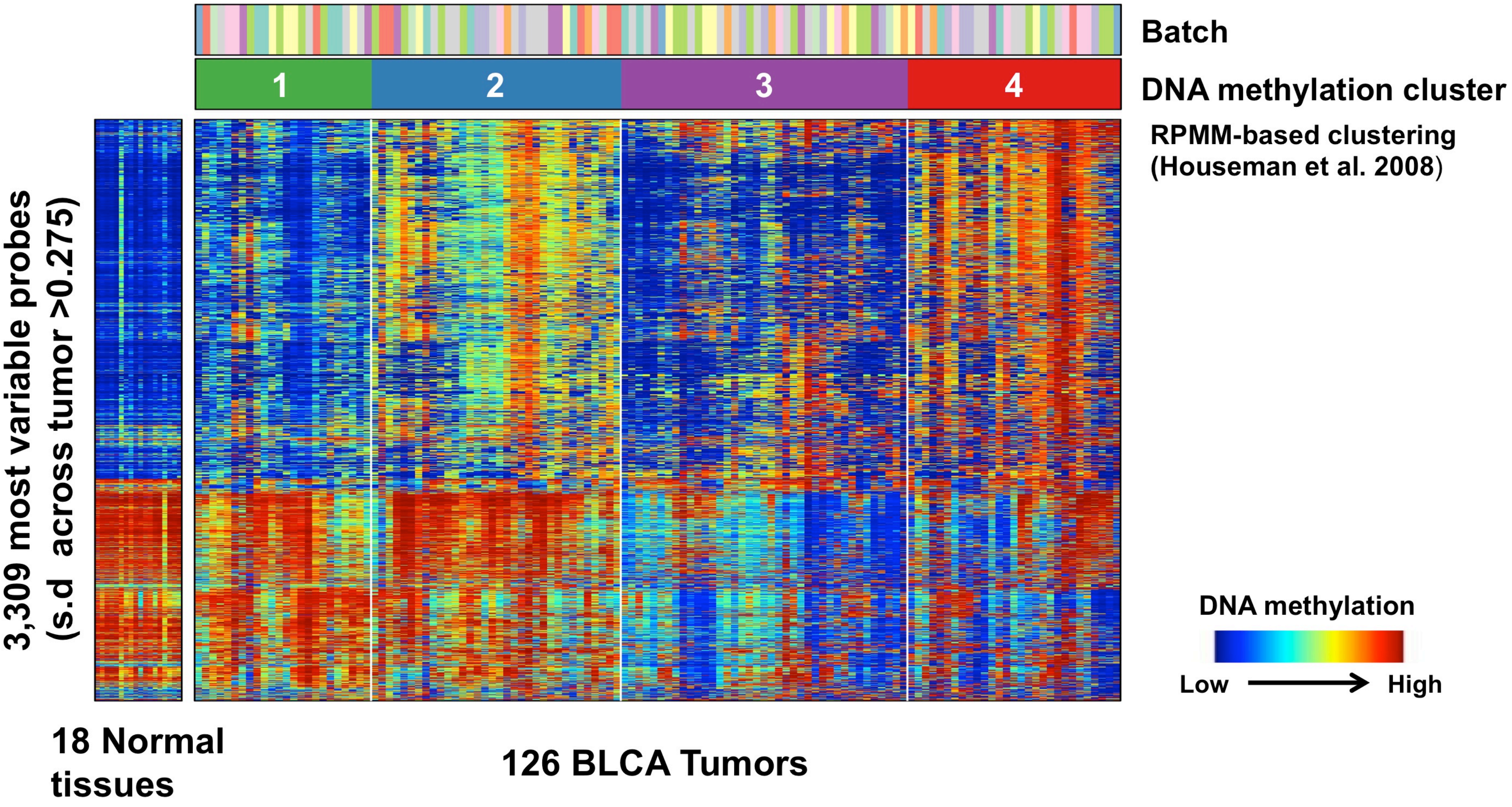
(J. Kim, Meyerson lab)





(Preliminary Analysis)

Unsupervised Clustering of Methylation Data

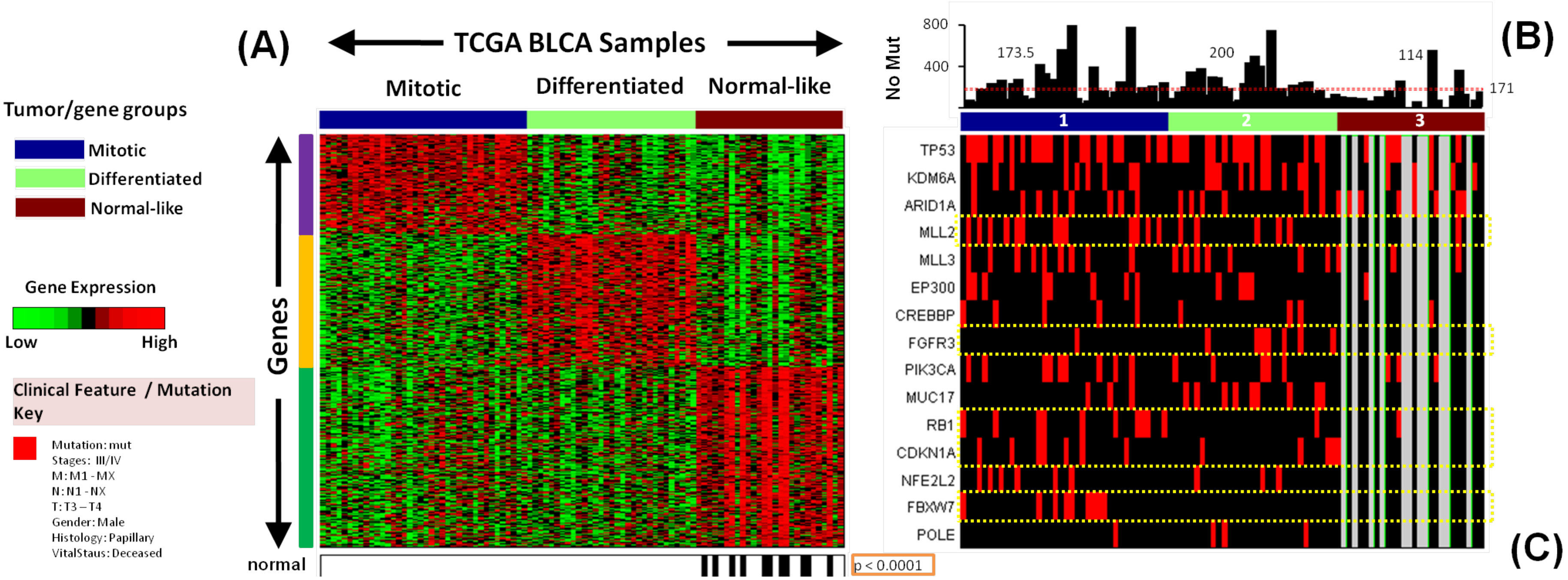


RPMM-based clustering (Houseman et al. 2008)

(Preliminary Analysis)

(T Hinoue, Laird lab)

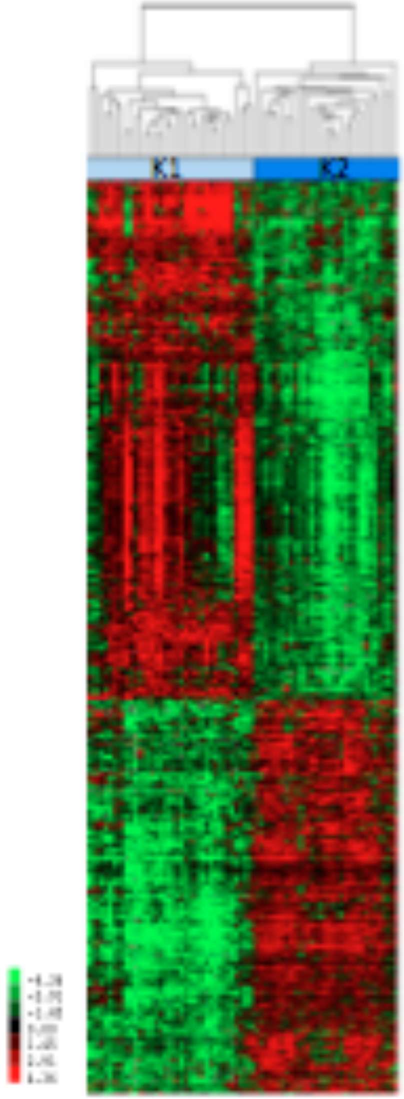
mRNA-based subtypes



(Preliminary Analysis)

(Y Liu, W Zhang)

SuperCluster Results in Relation to GO Terms

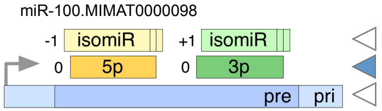


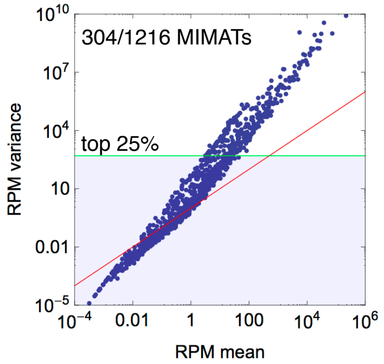
Gene Ontology Term	FDR
Immunoglobulin subtype	2.49E-21
chemotaxis	6.72E-21
regulation of leukocyte activation	6.00E-13
regulation of lymphocyte activation	6.00E-13
regulation of T cell activation	1.50E-10
Small chemokine, interleukin-8-like	3.13E-12
chemokine activity	3.26E-11
hydroxylysine	2.03E-08
hydroxyproline	2.89E-07
regulation of lymphocyte mediated immunity	3.32E-06
regulation of leukocyte mediated immunity	2.09E-05
regulation of immune effector process	2.84E-05
positive regulation of lymphocyte proliferation	0.002465
positive regulation of mononuclear cell proliferatic	0.003025
positive regulation of leukocyte proliferation	0.003025
positive regulation of T cell proliferation	0.003913
Tumor Necrosis Factor	0.002465
tumor necrosis factor receptor superfamily bindin	0.197701
adaptive immune response	0.018576
lymphocyte mediated immunity	0.034558
leukocyte mediated immunity	0.063276

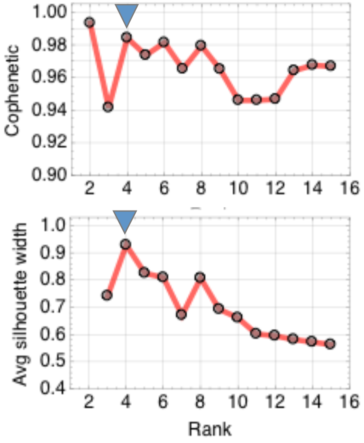
(Preliminary Analysis)

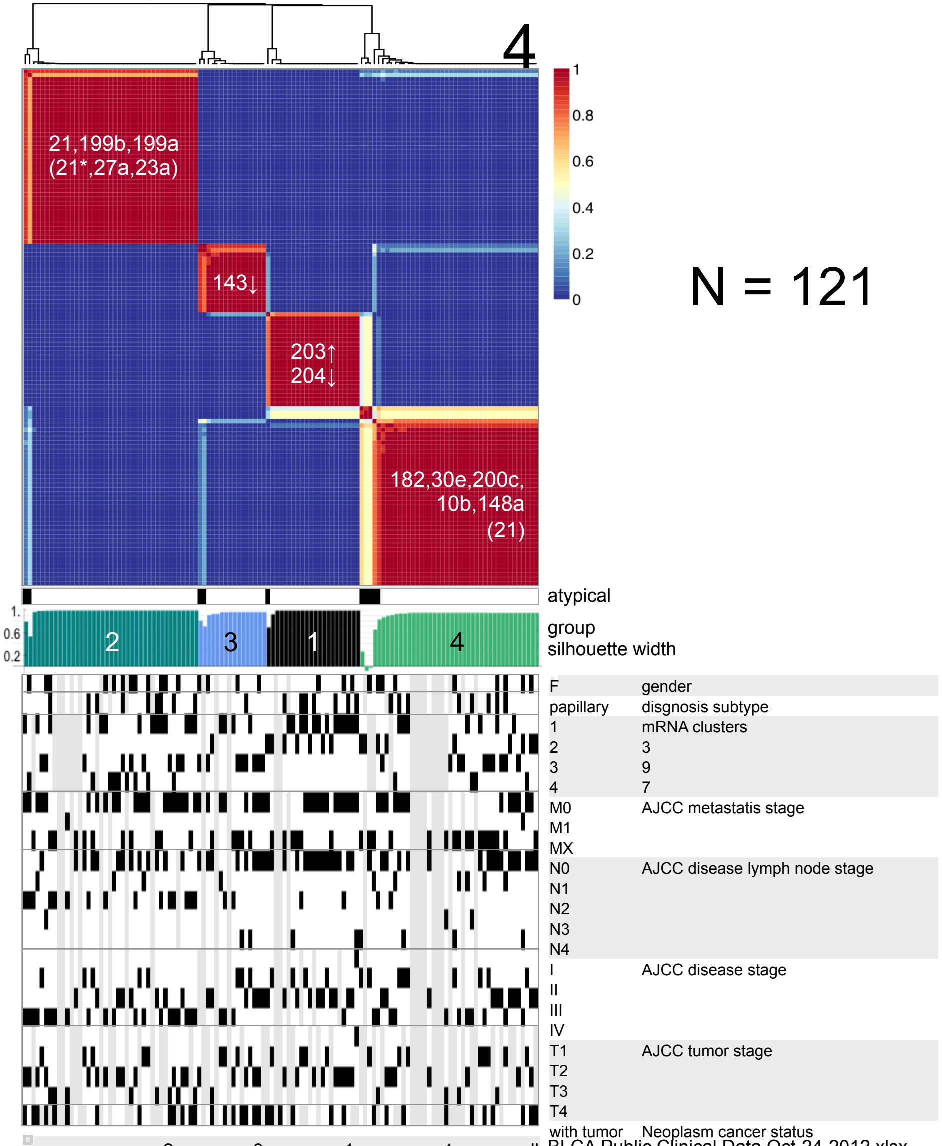
(K Hoadley, B Kim, et al.)

Unsupervised NMF Consensus Clustering of miRNA data









BLCA Public Clinical Data Oct-24-2012.xlsx

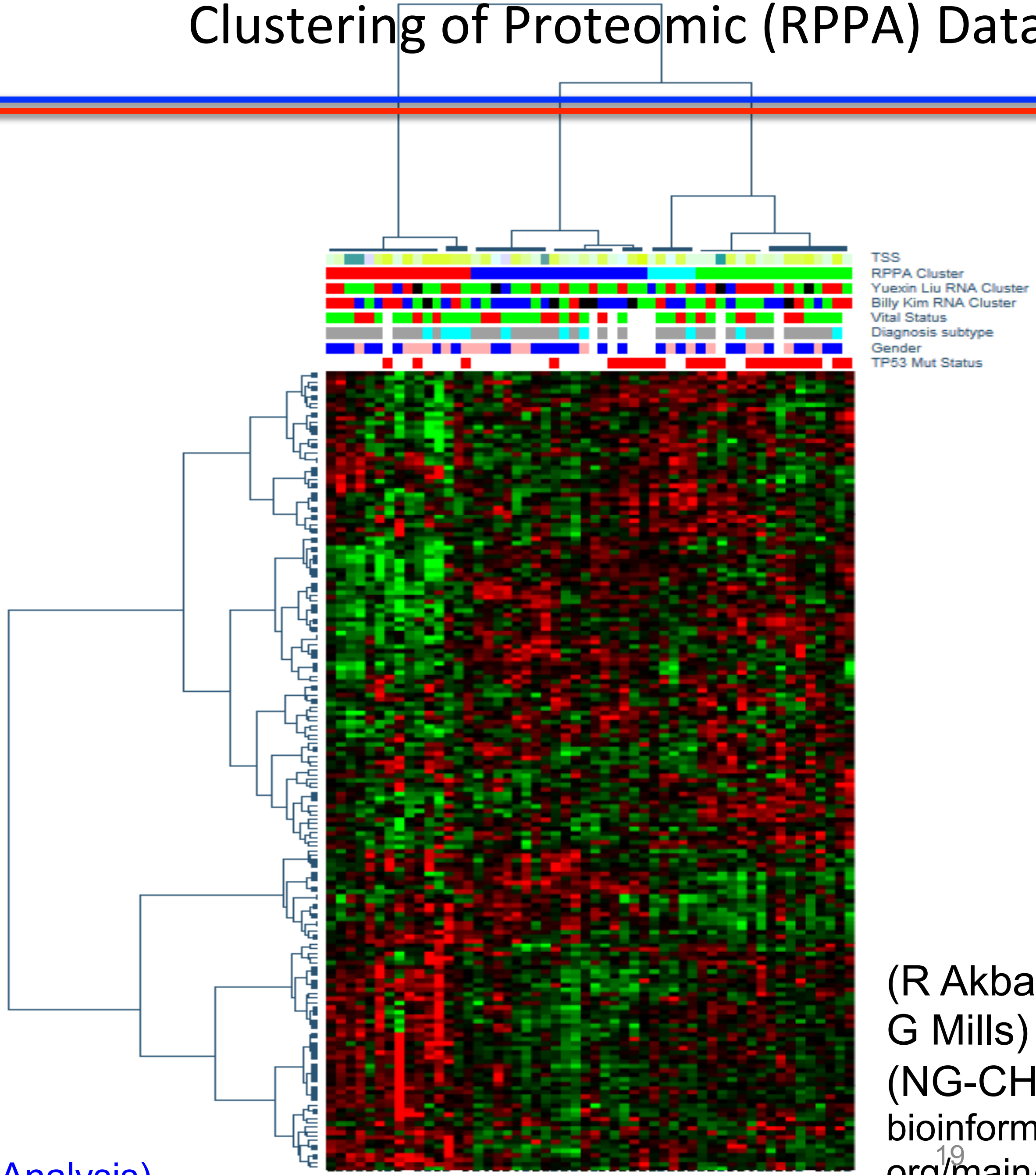
grp	2	3	1	4	all
n	41	16	22	42	121
sil w	0.97	0.93	0.97	0.87	0.93

N = 121

(Preliminary Analysis)

(G Robertson, A Chu, A Mungall, M Marra, BCGSC)

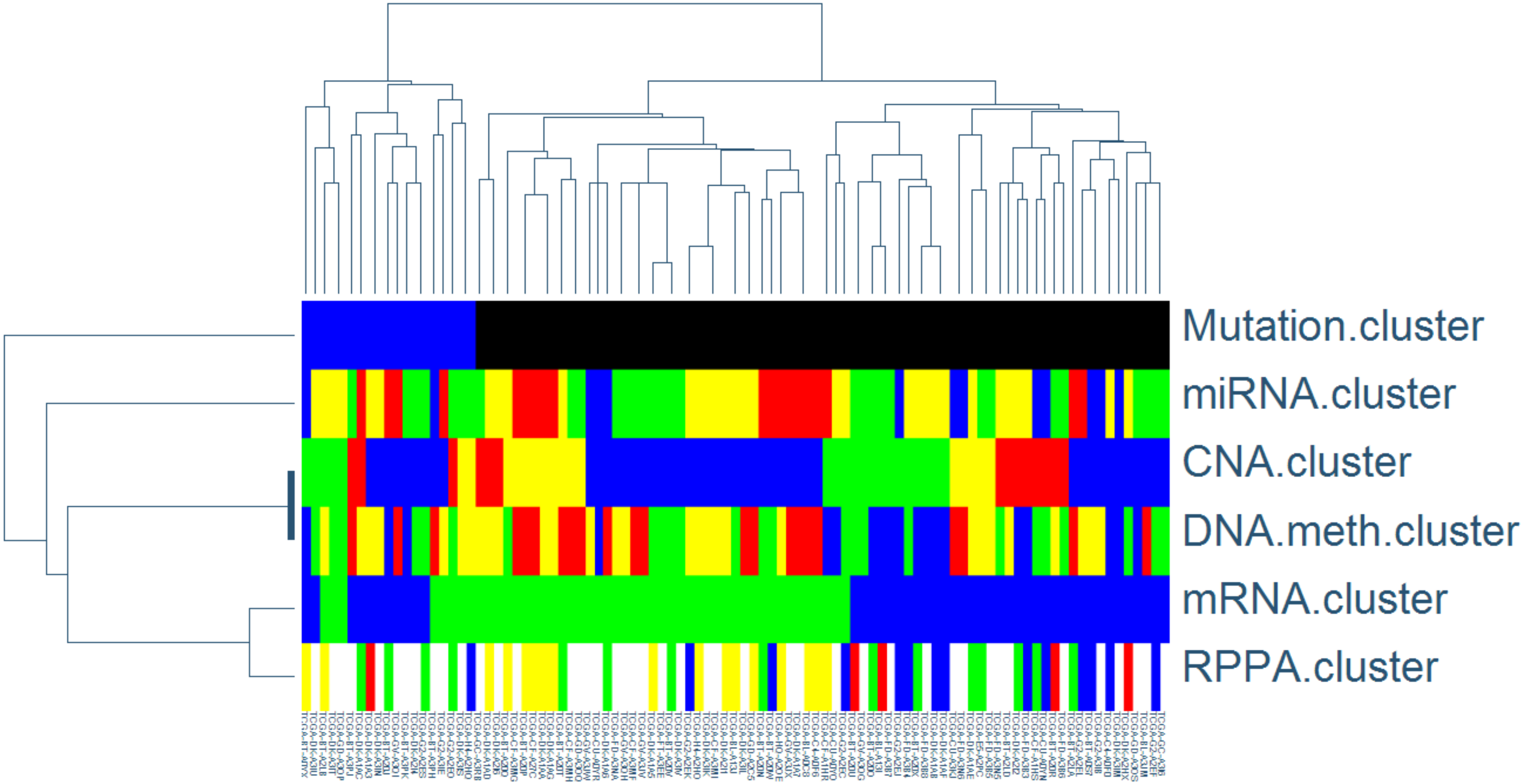
Clustering of Proteomic (RPPA) Data



(R Akbani, Z Liu, J Weinstein, G Mills)
(NG-CHM: http://bioinformatics.mdanderson.org/main/TCGA/NGCHM)

(Preliminary Analysis)

SuperCluster Results



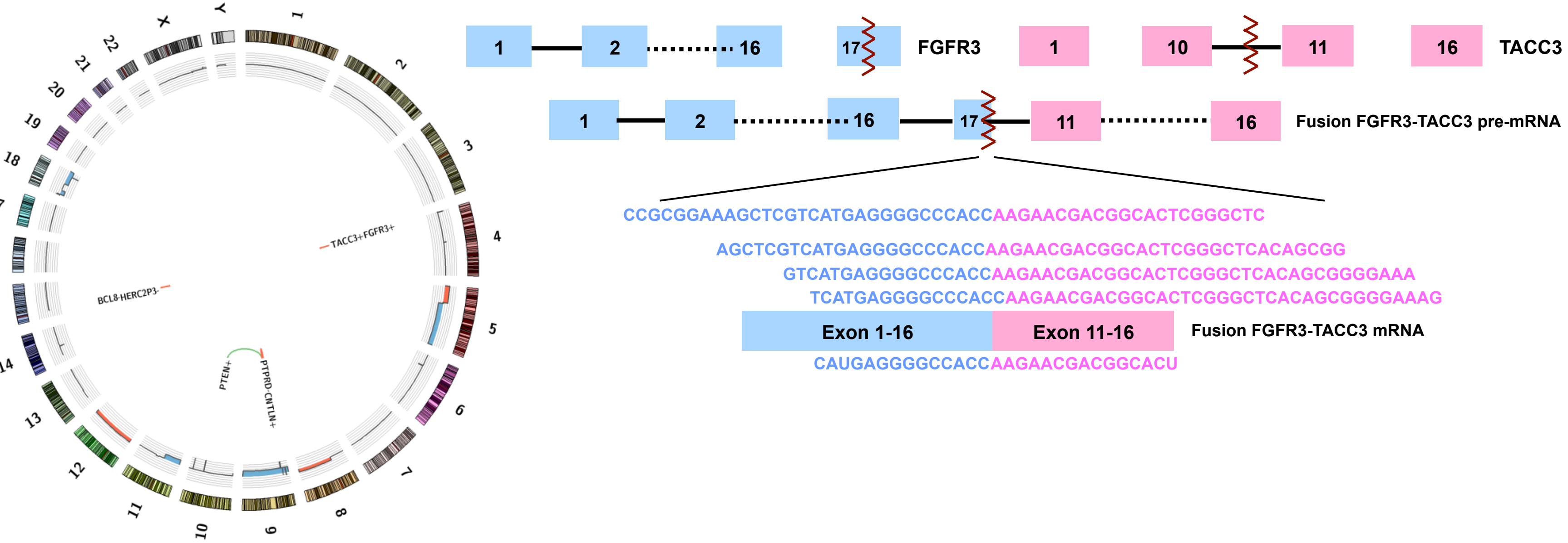
- Two to Three clusters found

(R Akbani) (NG-CHM: http://bioinformatics.mdanderson.org/main/TCGA/NGCHM)

(Preliminary Analysis)

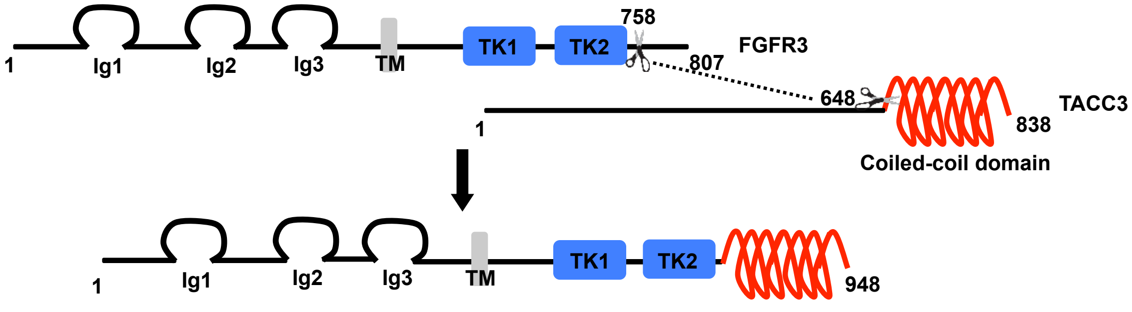
Detection of Fusion protein FGFR3-TACC3

A.



TCGA-CF-A3MH

B.



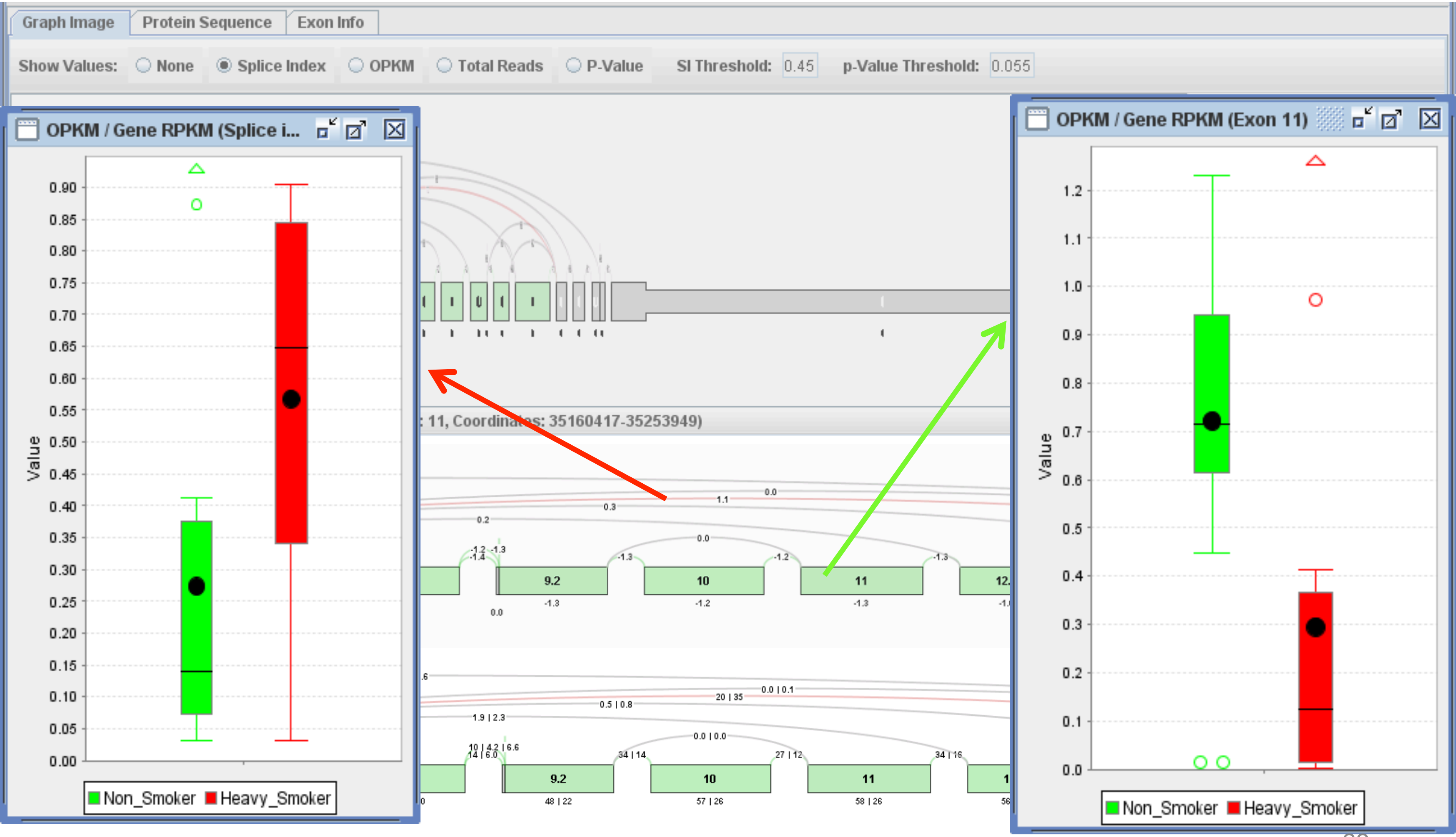
Fusion protein FGFR3-TACC3

(Preliminary Analysis)

(Kucherlapati lab)

CD44 Exon Skip Non-Smoker vs Heavy Smoker

(SpliceSeq Analysis)



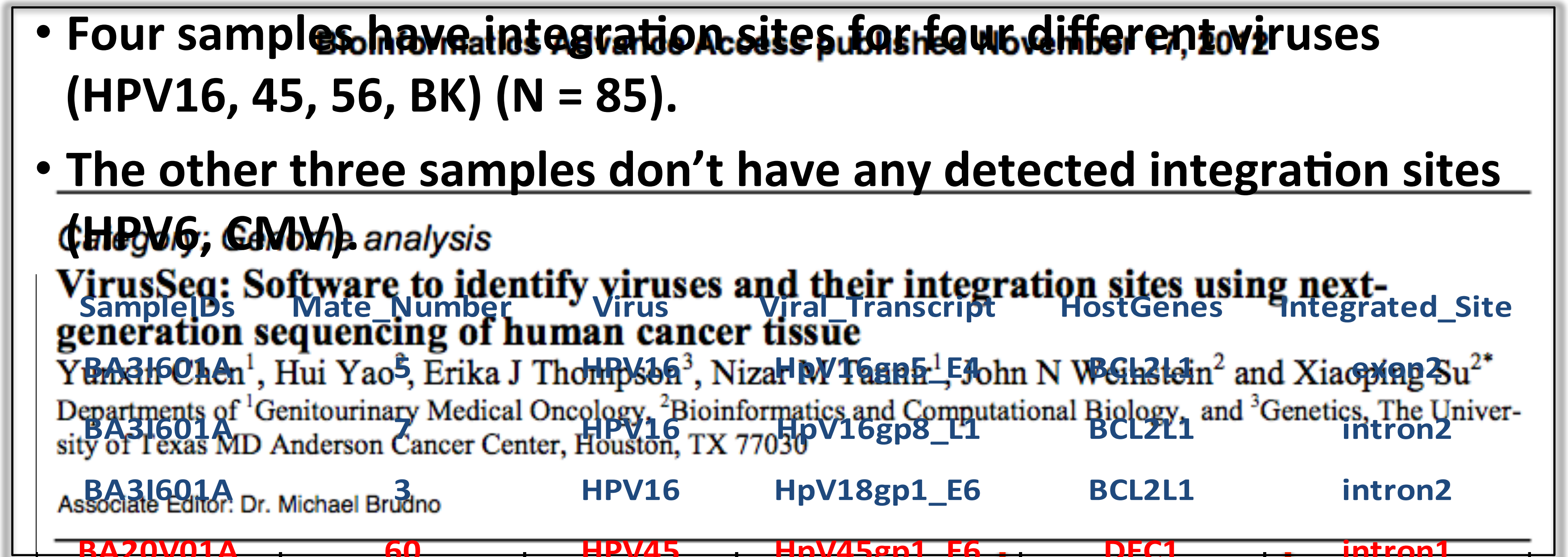
Show Values: ○ None ● Splice Index ○ OPKM ○ Total Reads ○ P-Value SI Threshold: 0.45 p-Value Threshold: 0.055

(Preliminary Analysis)

(M Ryan, J Weinstein)

Virus integration sites in 7 Samples (VirusSeq Analysis)

- Four samples have integration sites for four different viruses (HPV16, 45, 56, BK) (N = 85).
- The other three samples don't have any detected integration sites (HPV6, CMV).



Bioinformatics Advance Access published November 17, 2012

Category: Genome analysis

VirusSeq: Software to identify viruses and their integration sites using next-generation sequencing of human cancer tissue

Yunxin Chen[1], Hui Yao[2], Erika J Thompson[3], Nizar M Tannir[1], John N Weinstein[2] and Xiaoping Su[2*]

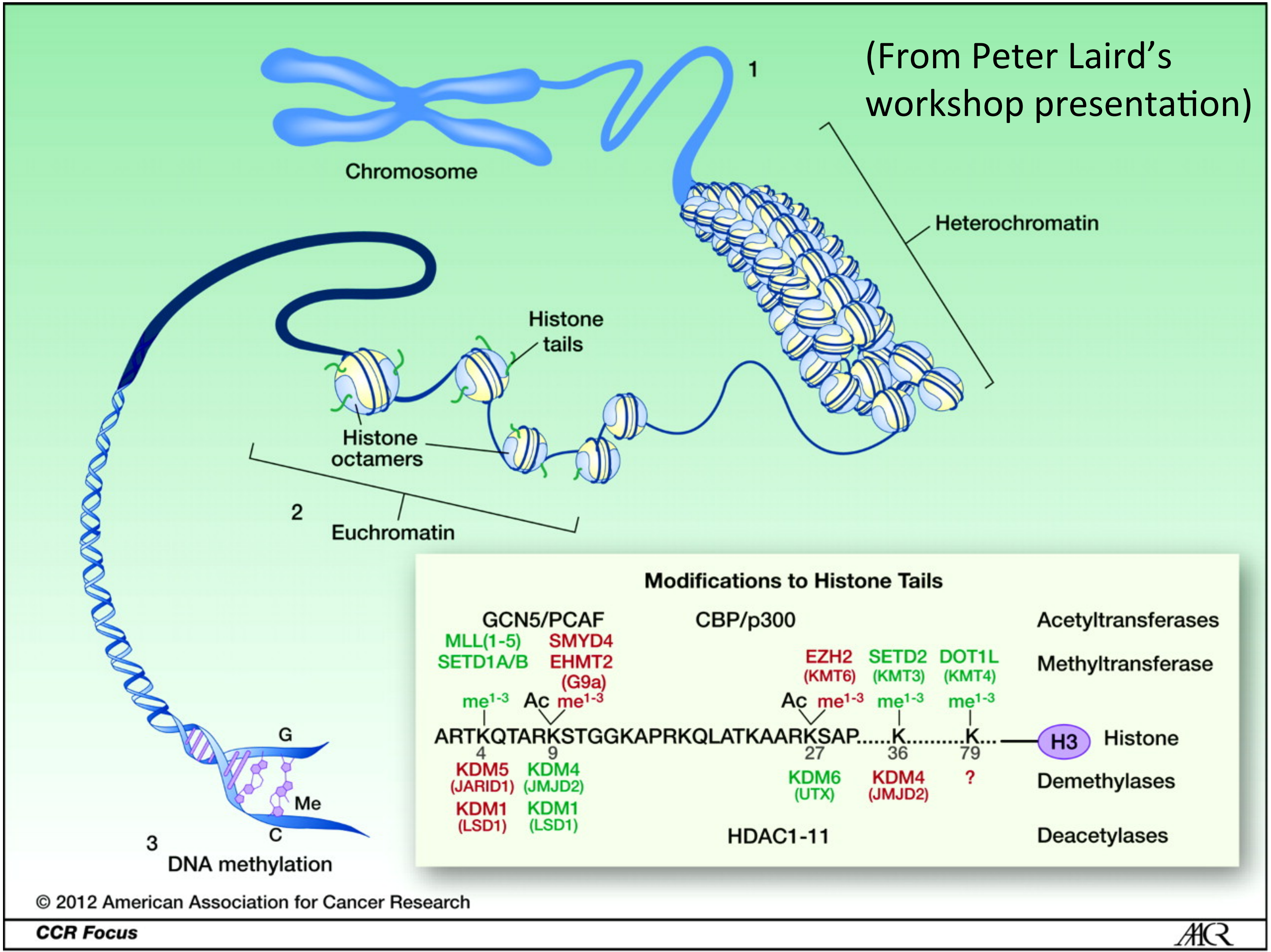
Departments of [1]Genitourinary Medical Oncology, [2]Bioinformatics and Computational Biology, and [3]Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030

Associate Editor: Dr. Michael Brudno

SampleIDs	Mate_Number	Virus	Viral_Transcript	HostGenes	Integrated_Site
BA3I601A	5	HPV16	HpV16gp5_E4	BCL2L1	exon2
BA3I601A	7	HPV16	HpV16gp8_L1	BCL2L1	intron2
BA3I601A	3	HPV16	HpV18gp1_E6	BCL2L1	intron2
BA20V01A	60	HPV45	HpV45gp1_E6	DEC1	intron1
BA20V01A	351	HPV45	HpV45gp2_E7	DEC1	intron1
BA20V01A	62	HPV45	HpV45gp3_E1	DEC1	intron1
BA3B401A	82	HPV56	HpV56gp1_E6	NOTCH1	exon27
BA3B401A	11	HPV56	HpV56gp2_E7	NOTCH1	exon27
BA3B401A	10	HPV56	HpV56gp1_E6	SEC16A	exon2
BA3B401A	227	HPV56	HpV56gp2_E7	SEC16A	exon3
BA3B401A	19	HPV56	HpV56gp3_E1	SEC16A	exon3
BA3B401A	36	HPV56	HpV56gp3_E1	SEC16A	intron1

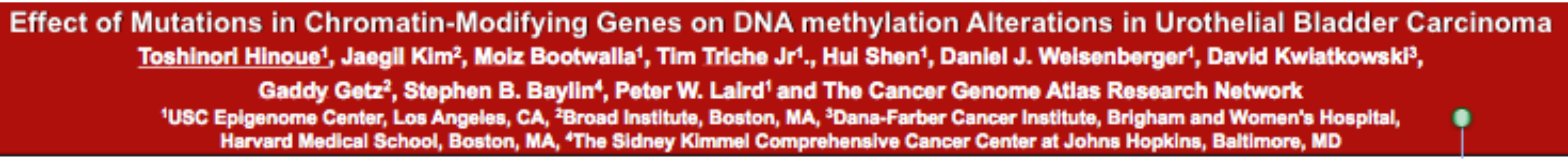
(X Su, H Yao)

Chromatin Remodeling Theme



(From Peter Laird's workshop presentation)

The Chromatin Remodeling Story

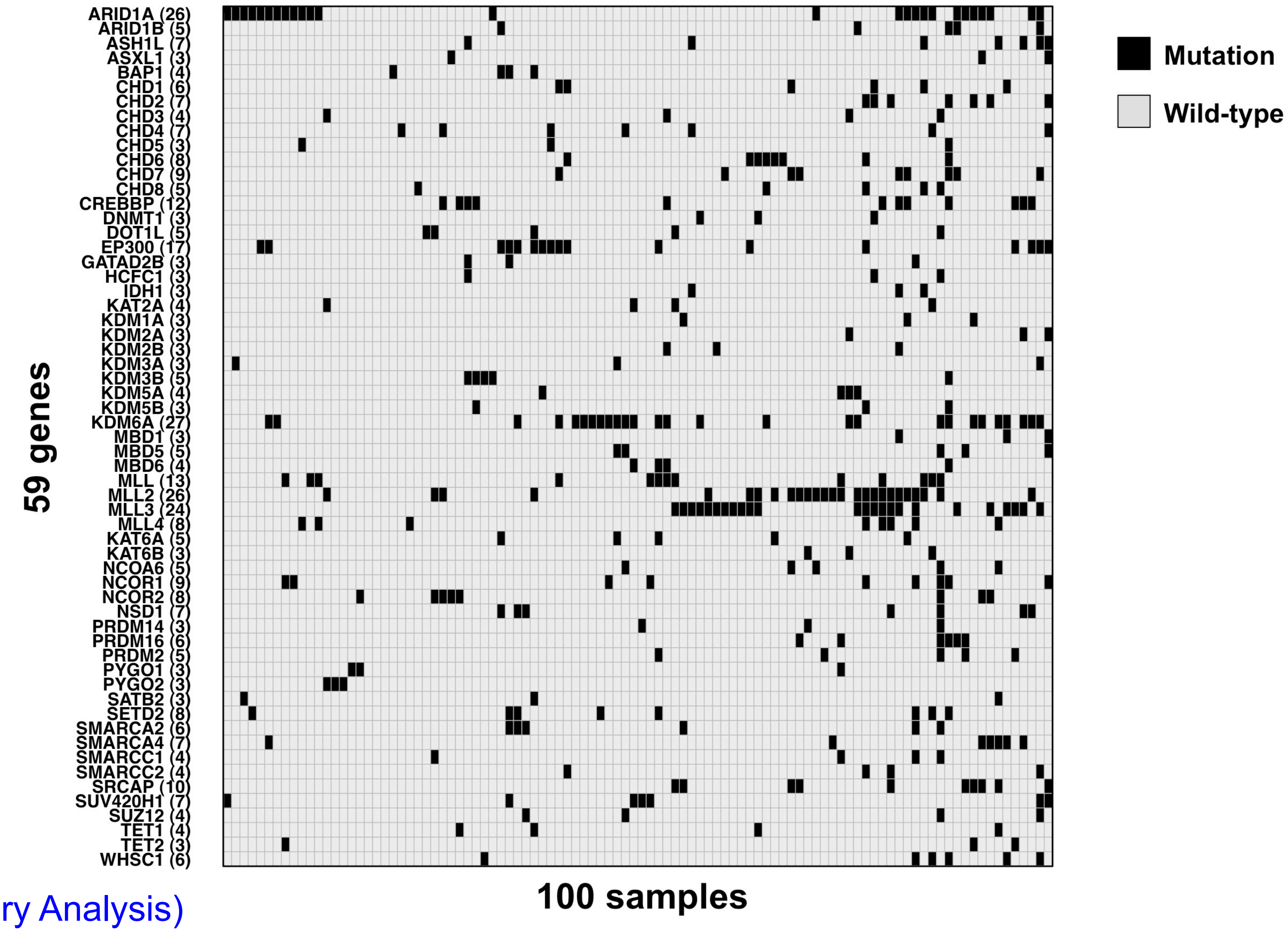


Effect of Mutations in Chromatin-Modifying Genes on DNA methylation Alterations in Urothelial Bladder Carcinoma

Toshinori Hinoue[1], Jaegil Kim[2], Moiz Bootwalla[1], Tim Triche Jr[1]., Hui Shen[1], Daniel J. Weisenberger[1], David Kwiatkowski[3],
Gaddy Getz[2], Stephen B. Baylin[4], Peter W. Laird[1] and The Cancer Genome Atlas Research Network
[1]USC Epigenome Center, Los Angeles, CA, [2]Broad Institute, Boston, MA, [3]Dana-Farber Cancer Institute, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA, [4]The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD

Toshinori Hinoue, Jaegil Kim, Moiz Bootwalla, Tim Triche Jr, Hui Shen, Daniel J. Weisenberger, David Kwiatkowski, Gaddy Getz, Stephen B. Baylin, Peter W. Laird

Epigenetic Modifiers Mutated in More than 3 Samples out of 100 Analyzed



(Preliminary Analysis)

Significantly Mutated Genes

(Preliminary Analysis)

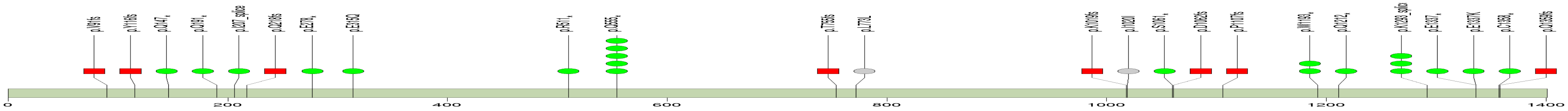
○ **Silent**

● **Nonsense**
Missense
Splice site

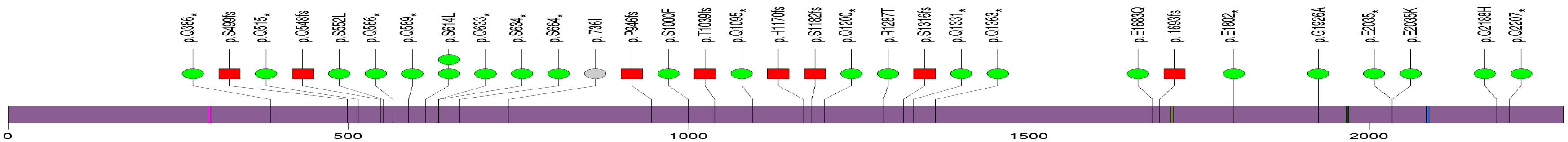
🟥 **Frame shift ins**
Frame shift del

● **In frame del**

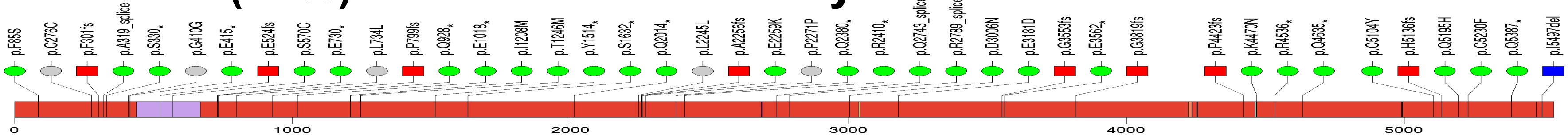
1. *KDM6A* (27%) - Histone H3K27 demethylase



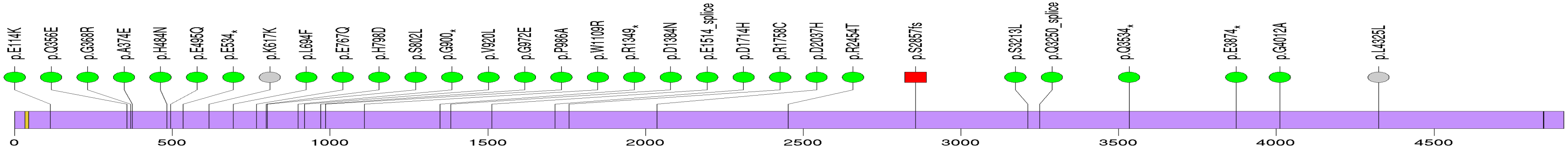
2. *ARID1A* (26%) - SWI/SNF complex



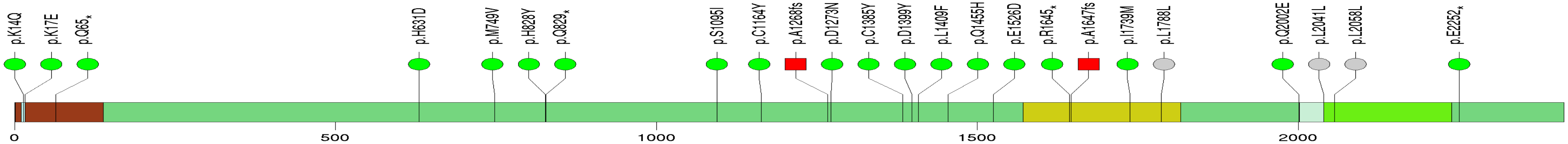
3. *MLL2* (26%) - Histone H3K4 methyltransferase

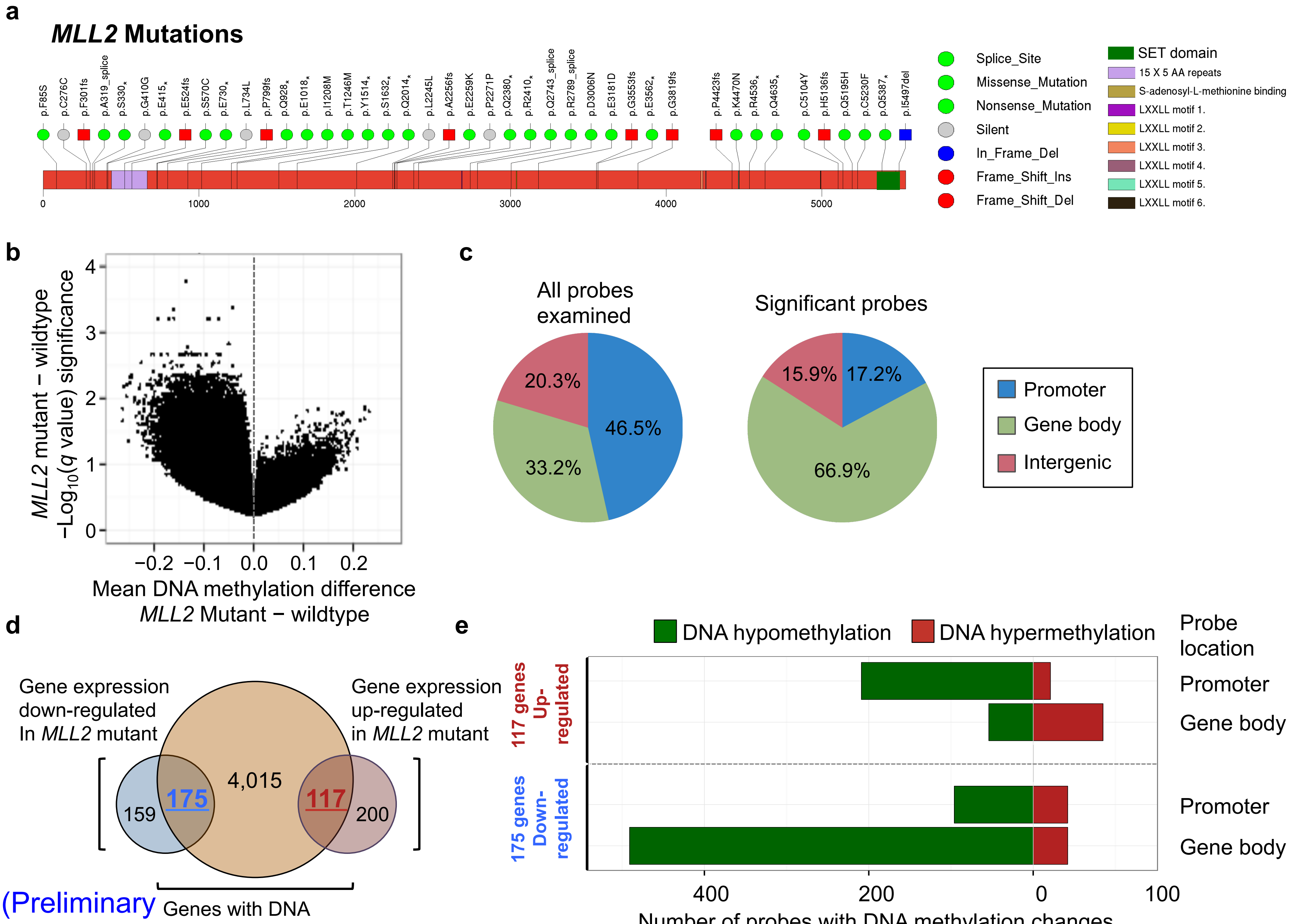


4. *MLL3* (24%) - Histone H3K4 methyltransferase



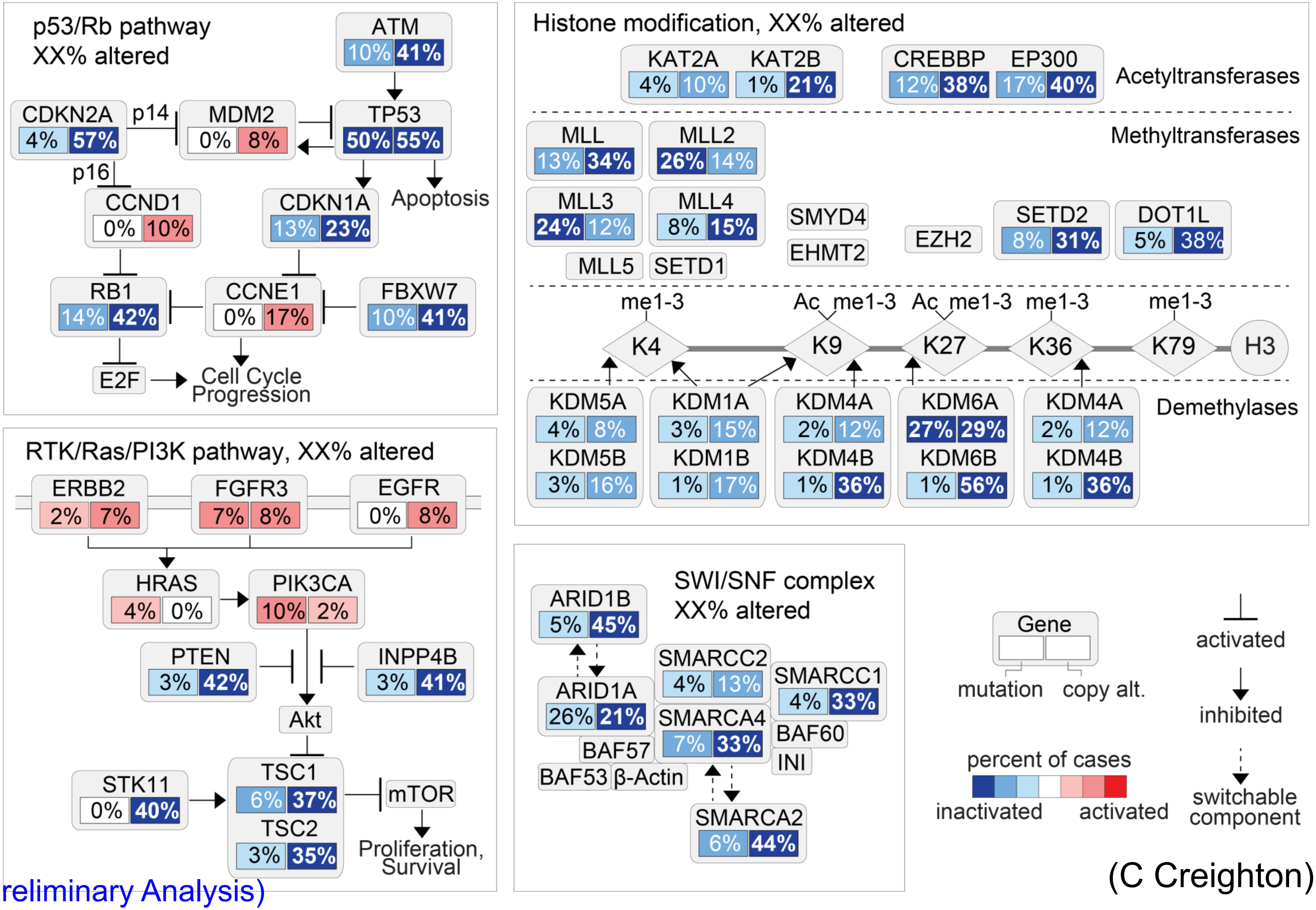
5. *EP300* (17%) - Histone acethyltransferase





(Preliminary Analysis)

Draft Pathway Figure



p53/Rb pathway
XX% altered

ATM 10% 41%
TP53 50% 55%
CDKN2A 4% 57%
p14
MDM2 0% 8%
p16
CCND1 0% 10%
CDKN1A 13% 23%
Apoptosis
RB1 14% 42%
CCNE1 0% 17%
FBXW7 10% 41%
E2F
Cell Cycle Progression

RTK/Ras/PI3K pathway, XX% altered

ERBB2 2% 7%
FGFR3 7% 8%
EGFR 0% 8%
HRAS 4% 0%
PIK3CA 10% 2%
PTEN 3% 42%
INPP4B 3% 41%
Akt
STK11 0% 40%
TSC1 6% 37%
TSC2 3% 35%
mTOR
Proliferation, Survival

Histone modification, XX% altered

Acetyltransferases
KAT2A 4% 10%
KAT2B 1% 21%
CREBBP 12% 38%
EP300 17% 40%

Methyltransferases
MLL 13% 34%
MLL2 26% 14%
MLL3 24% 12%
MLL4 8% 15%
MLL5
SETD1
SMYD4
EHMT2
EZH2
SETD2 8% 31%
DOT1L 5% 38%

me1-3 Ac me1-3 Ac me1-3 me1-3 me1-3
K4 K9 K27 K36 K79 H3

Demethylases
KDM5A 4% 8%
KDM5B 3% 16%
KDM1A 3% 15%
KDM1B 1% 17%
KDM4A 2% 12%
KDM4B 1% 36%
KDM6A 27% 29%
KDM6B 1% 56%
KDM4A 2% 12%
KDM4B 1% 36%

SWI/SNF complex
XX% altered

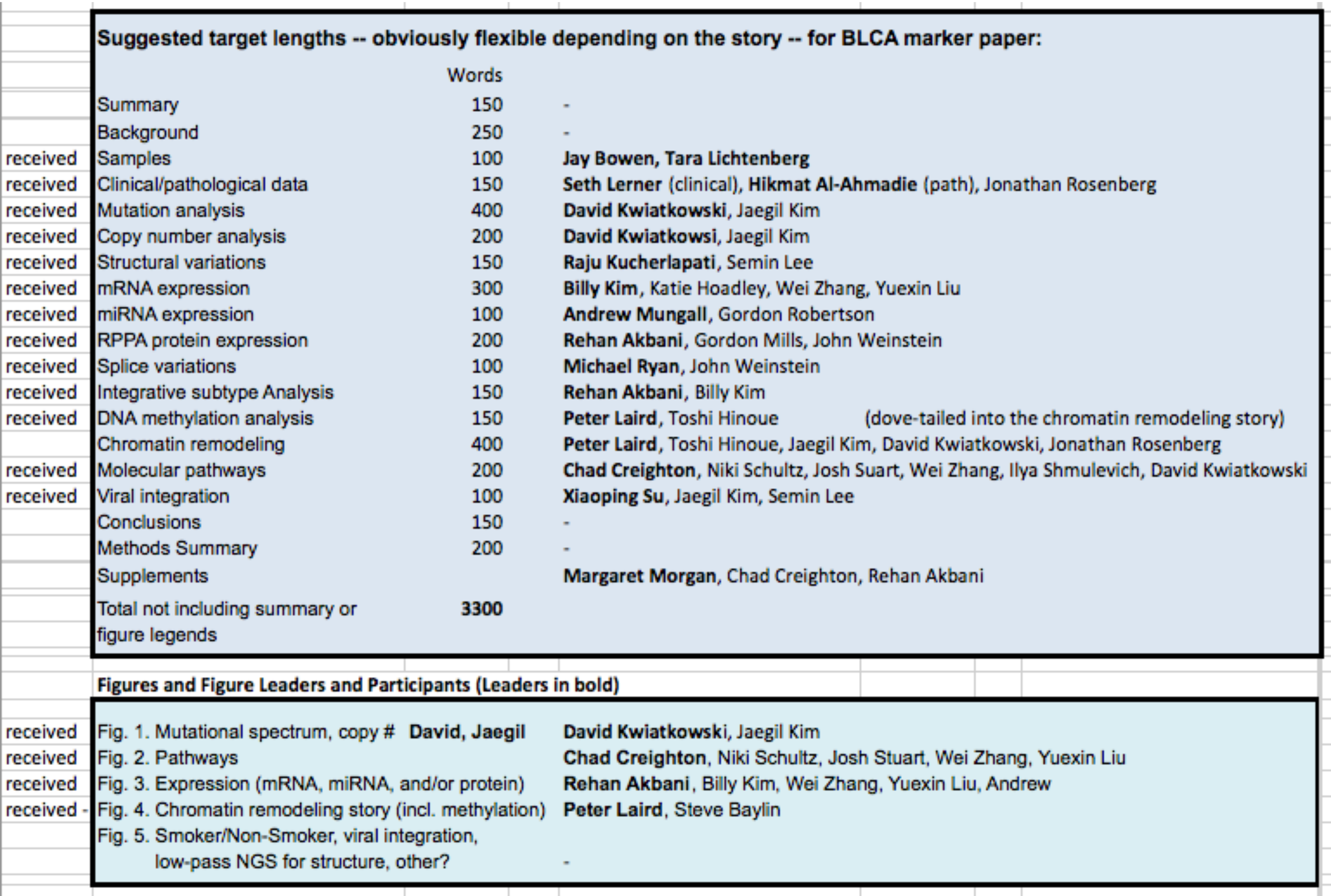
ARID1B 5% 45%
ARID1A 26% 21%
SMARCC2 4% 13%
SMARCC1 4% 33%
SMARCA4 7% 33%
BAF60
INI
BAF57
BAF53
β-Actin
SMARCA2 6% 44%

Gene
mutation copy alt.
percent of cases
inactivated activated
activated
inhibited
switchable component

(Preliminary Analysis)

(C Creighton)

Roles and Word Counts for writing BLCA Marker Paper



Suggested target lengths -- obviously flexible depending on the story -- for BLCA marker paper:

		Words		
	Summary	150	-	
	Background	250	-	
received	Samples	100	**Jay Bowen, Tara Lichtenberg**	
received	Clinical/pathological data	150	**Seth Lerner** (clinical), **Hikmat Al-Ahmadie** (path), Jonathan Rosenberg	
received	Mutation analysis	400	**David Kwiatkowski**, Jaegil Kim	
received	Copy number analysis	200	**David Kwiatkowsi**, Jaegil Kim	
received	Structural variations	150	**Raju Kucherlapati**, Semin Lee	
received	mRNA expression	300	**Billy Kim**, Katie Hoadley, Wei Zhang, Yuexin Liu	
received	miRNA expression	100	**Andrew Mungall**, Gordon Robertson	
received	RPPA protein expression	200	**Rehan Akbani**, Gordon Mills, John Weinstein	
received	Splice variations	100	**Michael Ryan**, John Weinstein	
received	Integrative subtype Analysis	150	**Rehan Akbani**, Billy Kim	
received	DNA methylation analysis	150	**Peter Laird**, Toshi Hinoue	(dove-tailed into the chromatin remodeling story)
	Chromatin remodeling	400	**Peter Laird**, Toshi Hinoue, Jaegil Kim, David Kwiatkowski, Jonathan Rosenberg	
received	Molecular pathways	200	**Chad Creighton**, Niki Schultz, Josh Suart, Wei Zhang, Ilya Shmulevich, David Kwiatkowski	
received	Viral integration	100	**Xiaoping Su**, Jaegil Kim, Semin Lee	
	Conclusions	150	-	
	Methods Summary	200	-	
	Supplements		**Margaret Morgan**, Chad Creighton, Rehan Akbani	
	Total not including summary or figure legends	**3300**		

Figures and Figure Leaders and Participants (Leaders in bold)

received	Fig. 1. Mutational spectrum, copy #	**David, Jaegil**	**David Kwiatkowski**, Jaegil Kim
received	Fig. 2. Pathways		**Chad Creighton**, Niki Schultz, Josh Stuart, Wei Zhang, Yuexin Liu
received	Fig. 3. Expression (mRNA, miRNA, and/or protein)		**Rehan Akbani**, Billy Kim, Wei Zhang, Yuexin Liu, Andrew
received -	Fig. 4. Chromatin remodeling story (incl. methylation)		**Peter Laird**, Steve Baylin
	Fig. 5. Smoker/Non-Smoker, viral integration, low-pass NGS for structure, other?		-