Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin

Hoadley, KA et al. Cell 158(4):929-944

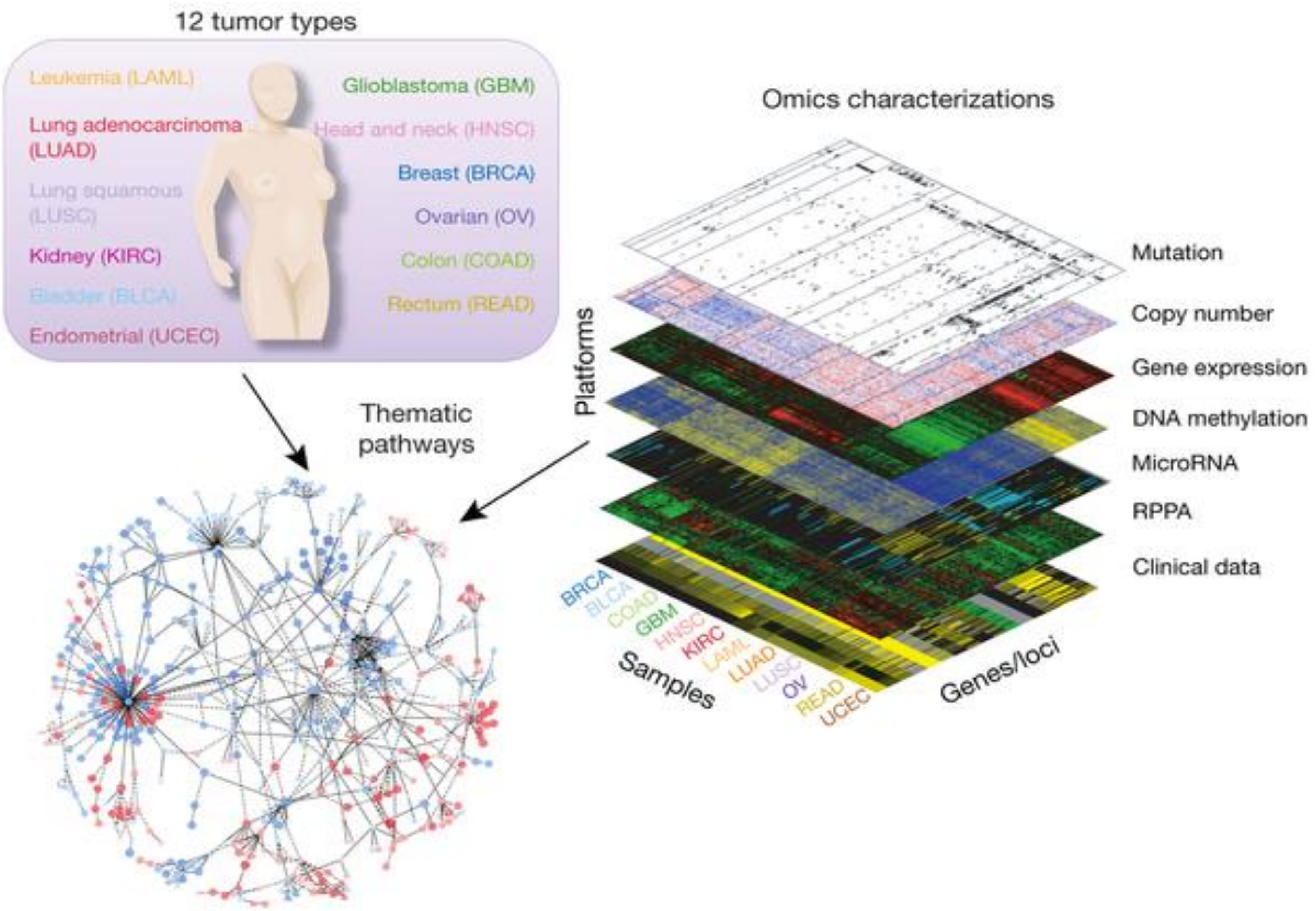
Background

- Cancers classified based on pathological criteria – tissue site of origin
- TCGA reported genome-wide studies of 10 malignancies
- Each single-tissue cancer type can be divided into 3-4 molecular subtypes

Motivation

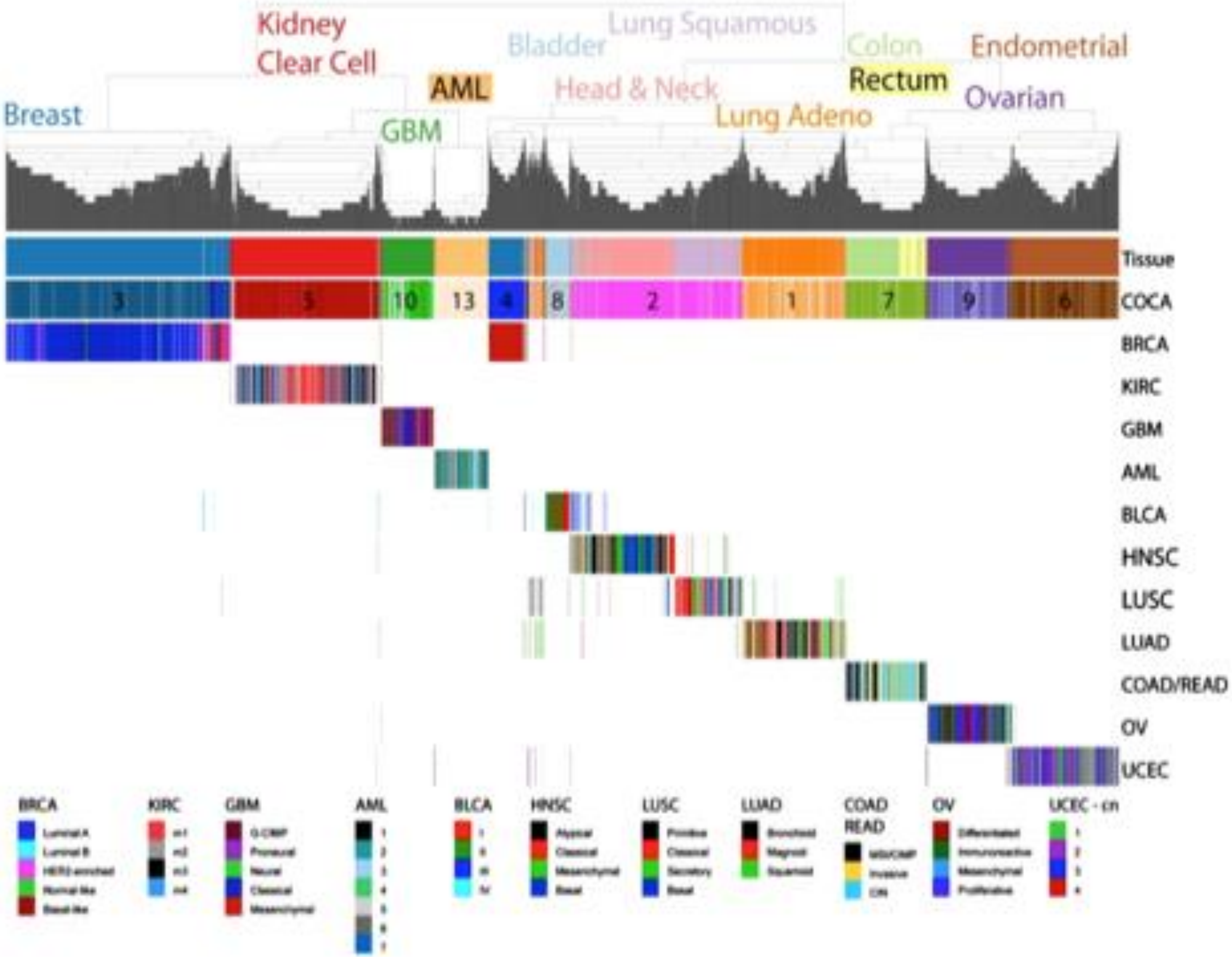
- Possible shared molecular alterations across cancers from different tissues - unclear
- Do disease subtypes span multiple tissues of origin?
- Molecular signatures may give distinct taxonomy relative to tissue-of-origin-based classification (current method)

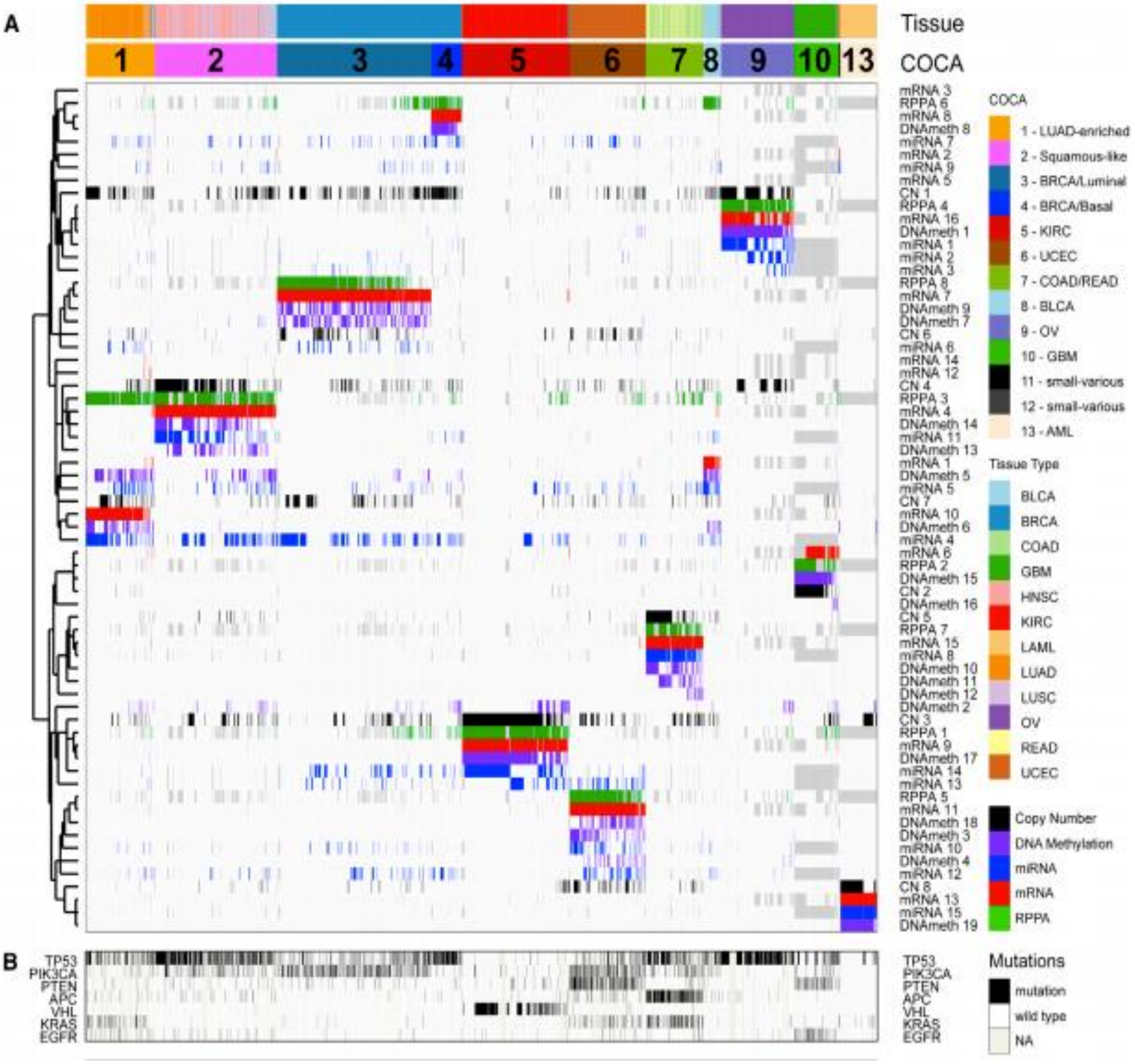
Pan-Cancer-12

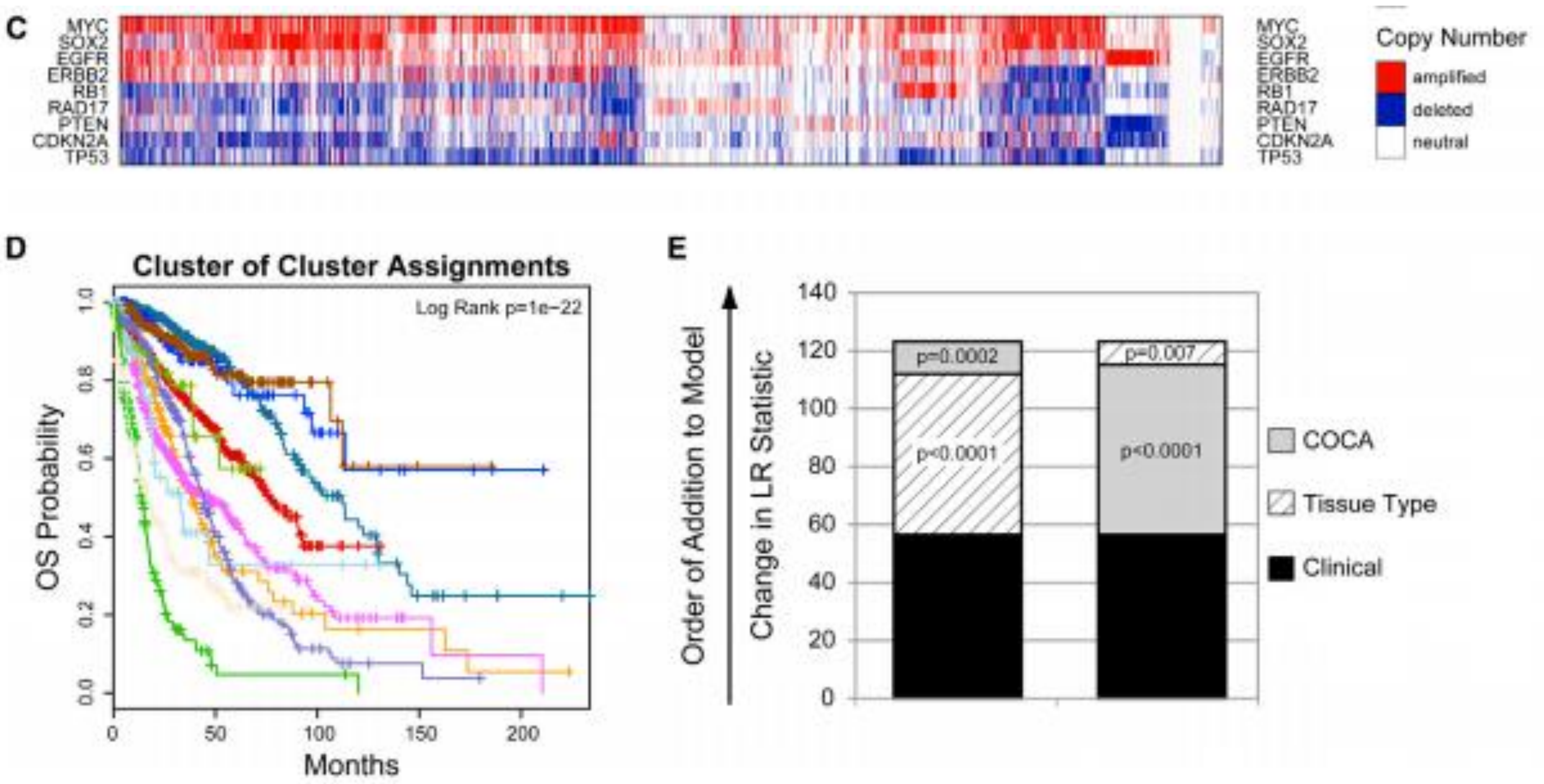


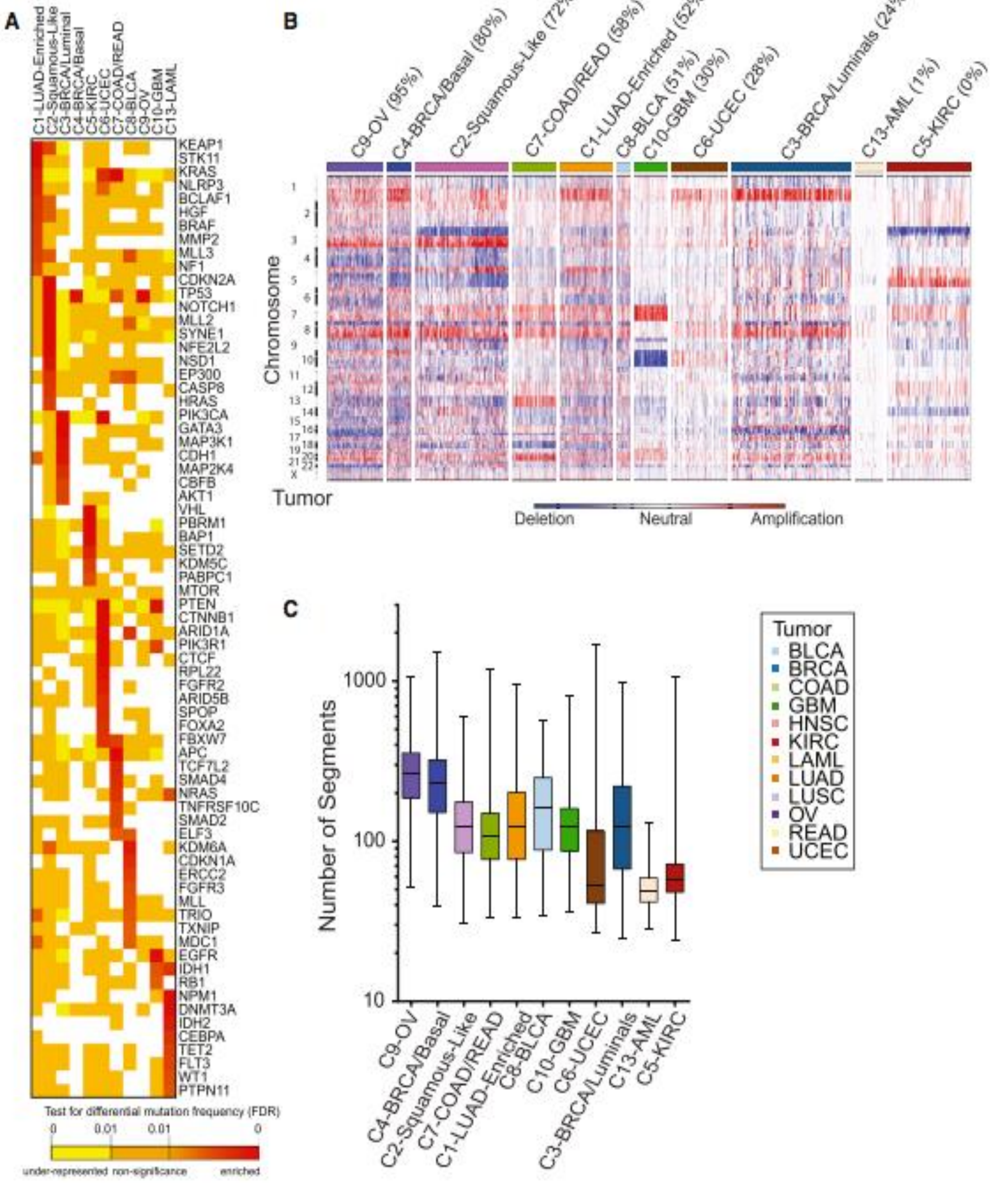
COCA

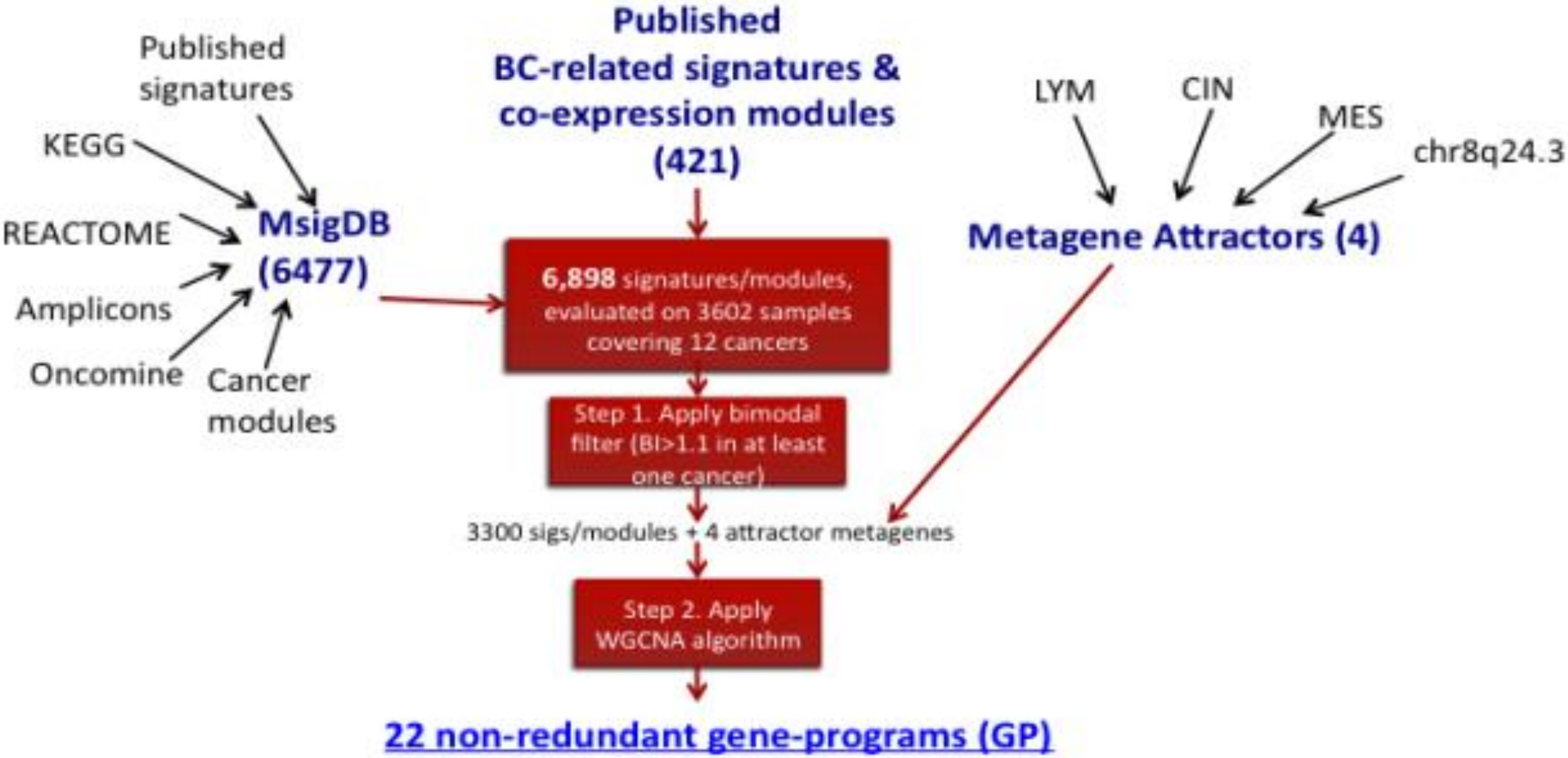
- “Cluster of cluster assignments”
- Input : binary vectors representing each platform-specific cluster group
- Each platform influences integrated result with weight proportional to # of distinct subtypes from consensus clustering
- Output : Reclusters samples according to vectors


Kidney
Clear Cell
Lung Squamous
Bladder
Colon
Endometrial
AML
Head & Neck
Rectum
Ovarian
Breast
GBM
Lung Adeno
3
5
10
13
4
8
2
1
7
9
6
Tissue
COCA
BRCA
KIRC
GBM
AML
BLCA
HNSC
LUSC
LUAD
COAD/READ
OV
UCEC
BRCA
Luminal A
Luminal B
HER2-enriched
Normal-like
Basal-like
KIRC
m1
m2
m3
m4
GBM
G-CIMP
Proneural
Neural
Classical
Mesenchymal
AML
1
2
3
4
5
6
7
BLCA
I
II
III
IV
HNSC
Atypical
Classical
Mesenchymal
Basal
LUSC
Primitive
Classical
Secretory
Basal
LUAD
Bronchioid
Magnoid
Squamoid
COAD
READ
MSI/CIMP
Invasive
CIN
OV
Differentiated
Immunoreactive
Mesenchymal
Proliferative
UCEC - cn
1
2
3
4


A
Tissue
COCA
1
2
3
4
5
6
7
8
9
10
13
mRNA 3
RPPA 6
miRNA 8
DNAmeth 8
miRNA 7
mRNA 2
miRNA 9
mRNA 5
CN 1
RPPA 4
mRNA 16
DNAmeth 1
miRNA 1
miRNA 2
miRNA 3
RPPA 8
mRNA 7
DNAmeth 9
DNAmeth 7
CN 6
miRNA 6
mRNA 14
mRNA 12
CN 4
RPPA 3
mRNA 4
DNAmeth 14
miRNA 11
DNAmeth 13
mRNA 1
DNAmeth 5
miRNA 5
CN 7
mRNA 10
DNAmeth 6
miRNA 4
mRNA 6
RPPA 2
DNAmeth 15
CN 2
DNAmeth 16
CN 5
RPPA 7
mRNA 15
miRNA 8
DNAmeth 10
DNAmeth 11
DNAmeth 12
DNAmeth 2
CN 3
RPPA 1
mRNA 9
DNAmeth 17
miRNA 14
miRNA 13
RPPA 5
mRNA 11
DNAmeth 18
DNAmeth 3
miRNA 10
DNAmeth 4
miRNA 12
CN 8
mRNA 13
miRNA 15
DNAmeth 19
COCA
1 - LUAD-enriched
2 - Squamous-like
3 - BRCA/Luminal
4 - BRCA/Basal
5 - KIRC
6 - UCEC
7 - COAD/READ
8 - BLCA
9 - OV
10 - GBM
11 - small-various
12 - small-various
13 - AML
Tissue Type
BLCA
BRCA
COAD
GBM
HNSC
KIRC
LAML
LUAD
LUSC
OV
READ
UCEC
Copy Number
DNA Methylation
miRNA
mRNA
RPPA
B
TP53
PIK3CA
PTEN
APC
VHL
KRAS
EGFR
Mutations
mutation
wild type
NA


C
MYC
SOX2
EGFR
ERBB2
RB1
RAD17
PTEN
CDKN2A
TP53
Copy Number
amplified
deleted
neutral
D
Cluster of Cluster Assignments
Log Rank p=1e−22
OS Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
Months
E
Order of Addition to Model
Change in LR Statistic
0
20
40
60
80
100
120
140
p=0.0002
p<0.0001
p=0.007
p<0.0001
COCA
Tissue Type
Clinical

A
C1-LUAD-Enriched
C2-Squamous-Like
C3-BRCA/Luminal
C4-BRCA/Basal
C5-KIRC
C6-UCEC
C7-COAD/READ
C8-BLCA
C9-OV
C10-GBM
C13-LAML
KEAP1
STK11
KRAS
NLRP3
BCLAF1
HGF
BRAF
MMP2
MLL3
NF1
CDKN2A
TP53
NOTCH1
MLL2
SYNE1
NFE2L2
NSD1
EP300
CASP8
HRAS
PIK3CA
GATA3
MAP3K1
CDH1
MAP2K4
CBFB
AKT1
VHL
PBRM1
BAP1
SETD2
KDM5C
PABPC1
MTOR
PTEN
CTNNB1
ARID1A
PIK3R1
CTCF
RPL22
FGFR2
ARID5B
SPOP
FOXA2
FBXW7
APC
TCF7L2
SMAD4
NRAS
TNFRSF10C
SMAD2
ELF3
KDM6A
CDKN1A
ERCC2
FGFR3
MLL
TRIO
TXNIP
MDC1
EGFR
IDH1
RB1
NPM1
DNMT3A
IDH2
CEBPA
TET2
FLT3
WT1
PTPN11
Test for differential mutation frequency (FDR)
0
0.01
0.01
0
under-represented
non-significance
enriched
B
C9-OV (95%)
C4-BRCA/Basal (80%)
C2-Squamous-Like (72%)
C7-COAD/READ (58%)
C1-LUAD-Enriched (52%)
C8-BLCA (51%)
C10-GBM (30%)
C6-UCEC (28%)
C3-BRCA/Luminals (24%)
C13-AML (1%)
C5-KIRC (0%)
Chromosome
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Tumor
Deletion
Neutral
Amplification
C
Number of Segments
1000
100
10
C9-OV
C4-BRCA/Basal
C2-Squamous-Like
C7-COAD/READ
C1-LUAD-Enriched
C8-BLCA
C10-GBM
C6-UCEC
C3-BRCA/Luminals
C13-AML
C5-KIRC
Tumor
BLCA
BRCA
COAD
GBM
HNSC
KIRC
LAML
LUAD
LUSC
OV
READ
UCEC



Published signatures

KEGG

REACTOME

Amplicons

Oncomine

Cancer modules

MsigDB (6477)

Published BC-related signatures & co-expression modules (421)

LYM

CIN

MES

chr8q24.3

Metagene Attractors (4)

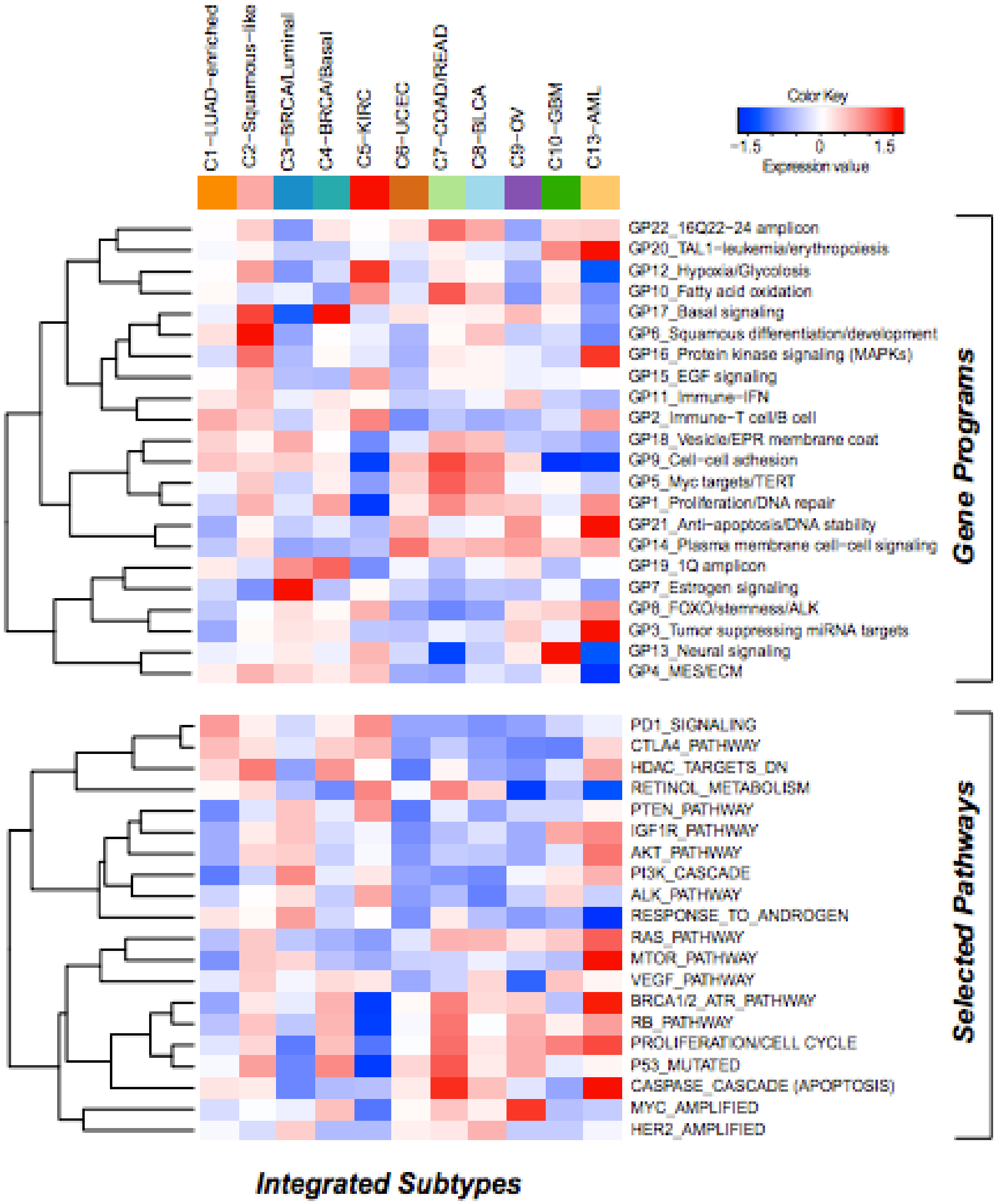
6,898 signatures/modules, evaluated on 3602 samples covering 12 cancers

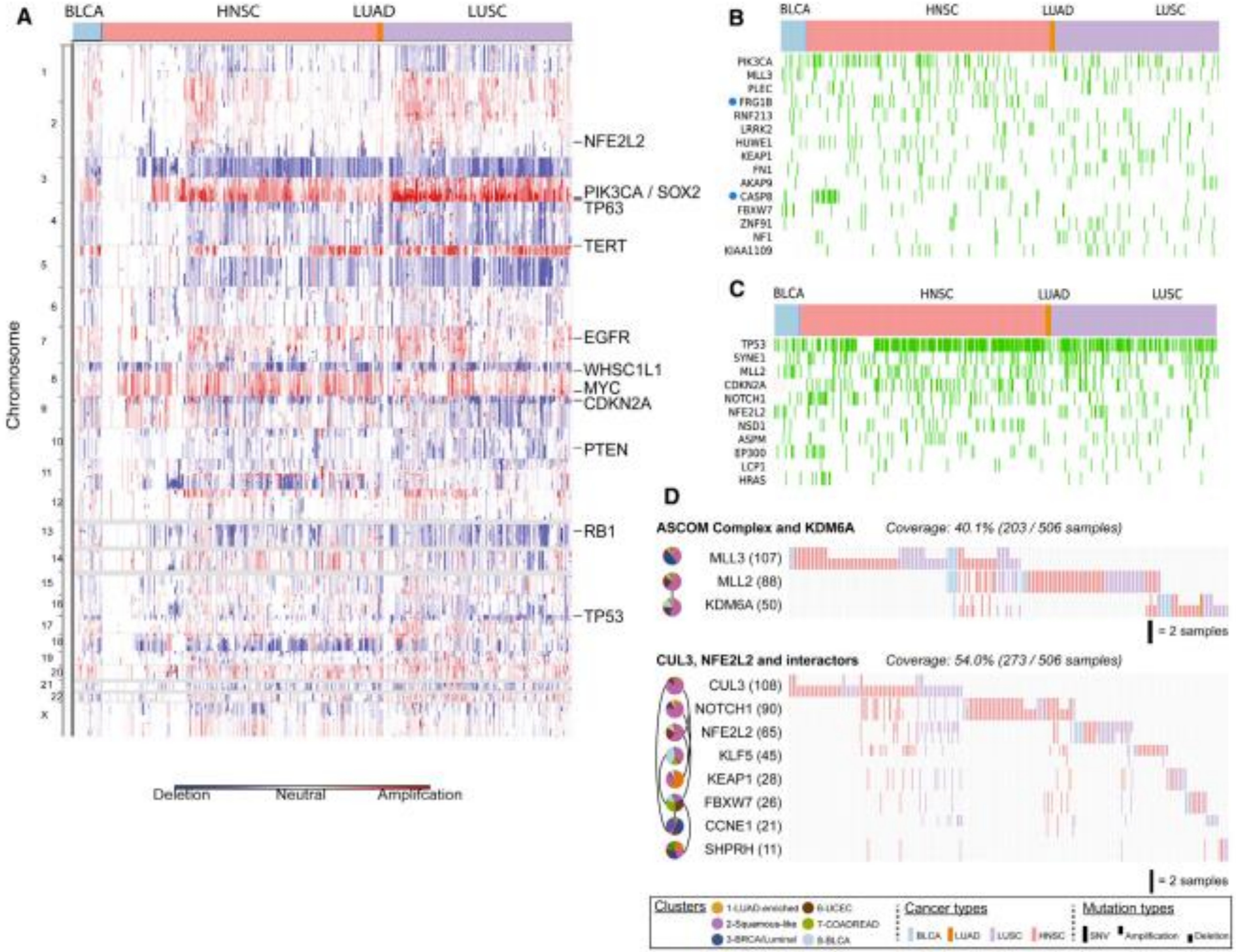
Step 1. Apply bimodal filter (BI>1.1 in at least one cancer)

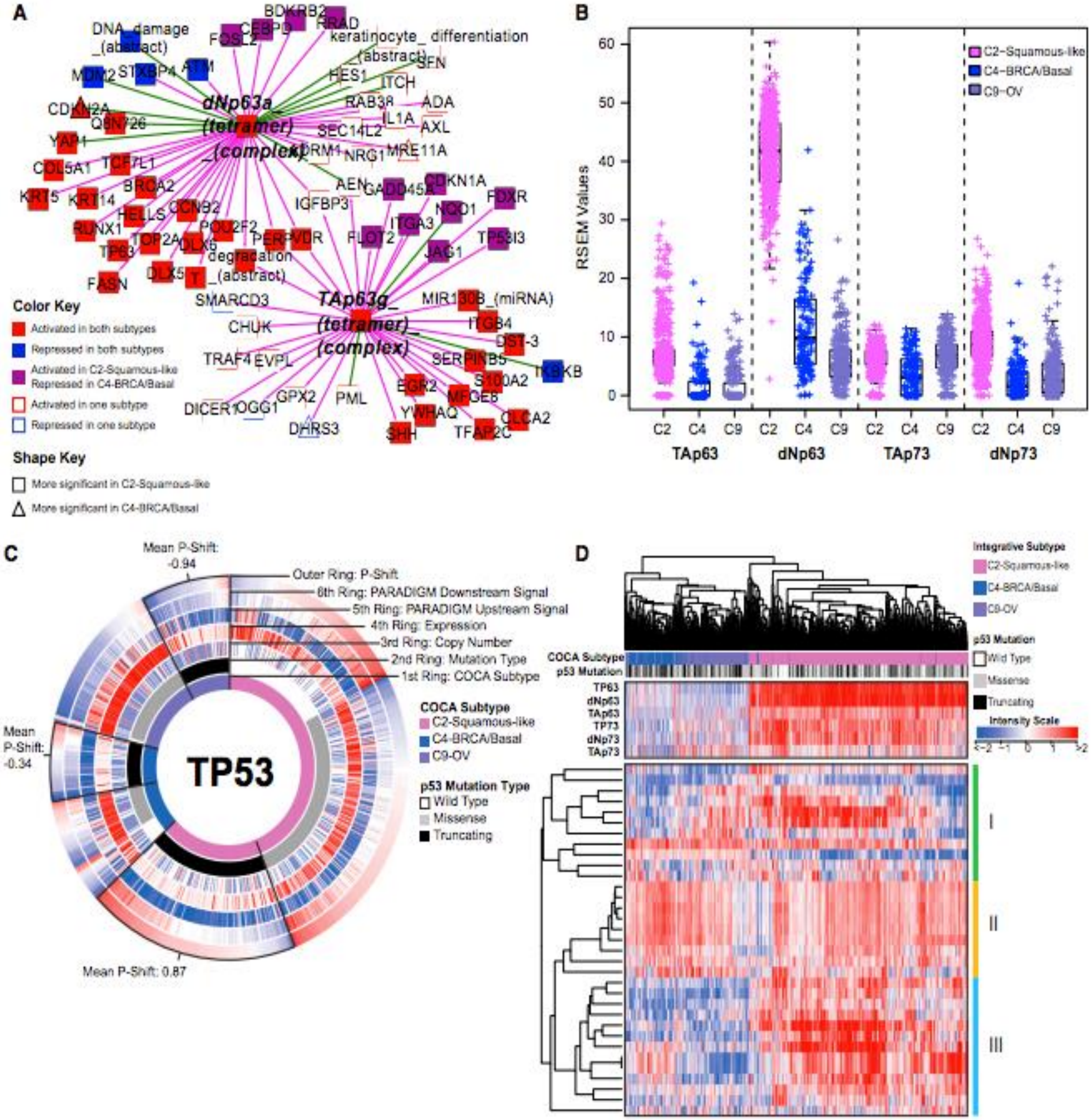
3300 sigs/modules + 4 attractor metagenes

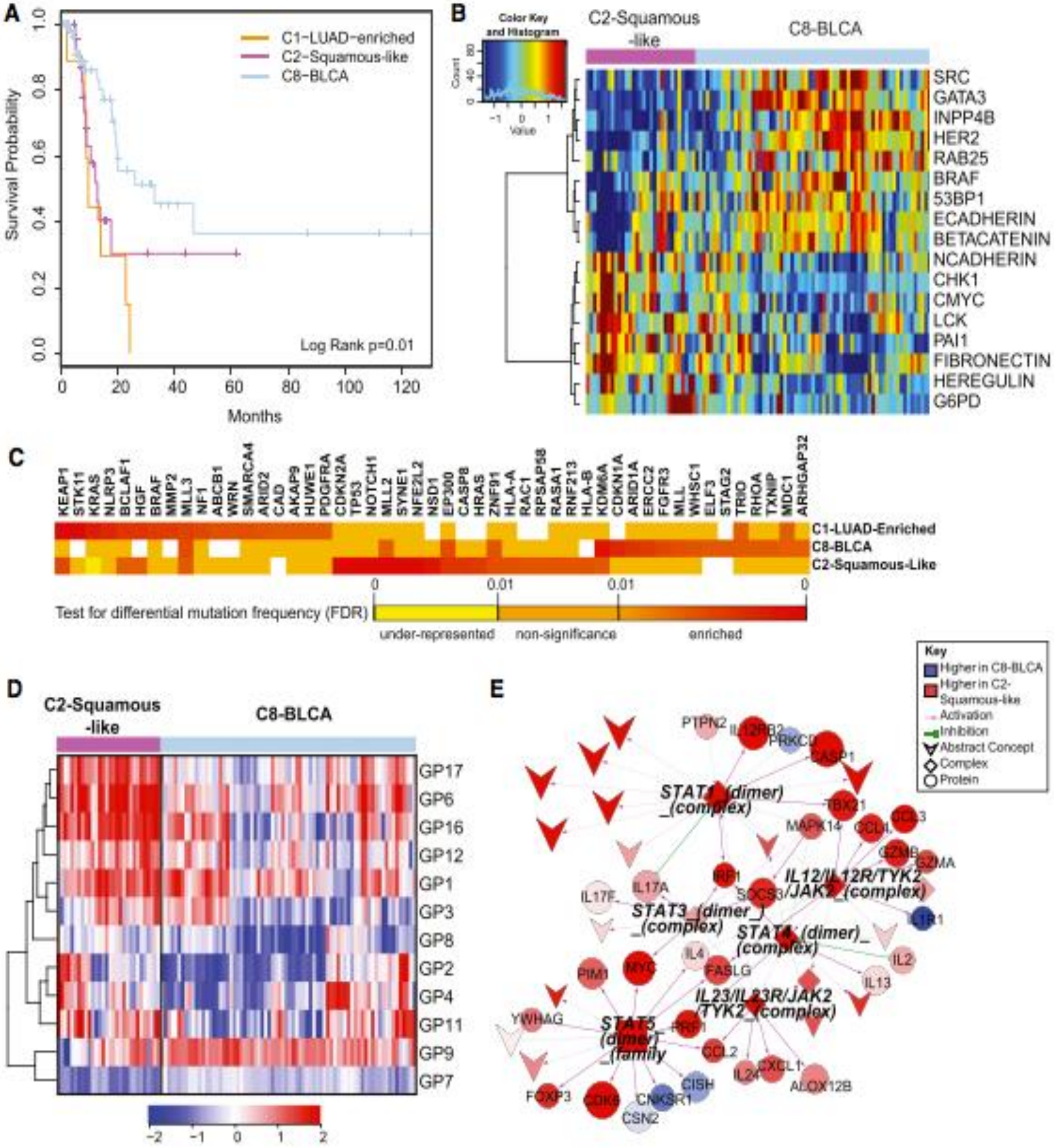
Step 2. Apply WGCNA algorithm

22 non-redundant gene-programs (GP)


C1−LUAD−enriched
C2−Squamous−like
C3−BRCA/Luminal
C4−BRCA/Basal
C5−KIRC
C6−UCEC
C7−COAD/READ
C8−BLCA
C9−OV
C10−GBM
C13−AML
Color Key
−1.5
0
1.5
Expression value
GP22_16Q22−24 amplicon
GP20_TAL1−leukemia/erythropoiesis
GP12_Hypoxia/Glycolosis
GP10_Fatty acid oxidation
GP17_Basal signaling
GP6_Squamous differentiation/development
GP16_Protein kinase signaling (MAPKs)
GP15_EGF signaling
GP11_Immune−IFN
GP2_Immune−T cell/B cell
GP18_Vesicle/EPR membrane coat
GP9_Cell−cell adhesion
GP5_Myc targets/TERT
GP1_Proliferation/DNA repair
GP21_Anti−apoptosis/DNA stability
GP14_Plasma membrane cell−cell signaling
GP19_1Q amplicon
GP7_Estrogen signaling
GP8_FOXO/stemness/ALK
GP3_Tumor suppressing miRNA targets
GP13_Neural signaling
GP4_MES/ECM
Gene Programs
PD1_SIGNALING
CTLA4_PATHWAY
HDAC_TARGETS_DN
RETINOL_METABOLISM
PTEN_PATHWAY
IGF1R_PATHWAY
AKT_PATHWAY
PI3K_CASCADE
ALK_PATHWAY
RESPONSE_TO_ANDROGEN
RAS_PATHWAY
MTOR_PATHWAY
VEGF_PATHWAY
BRCA1/2_ATR_PATHWAY
RB_PATHWAY
PROLIFERATION/CELL CYCLE
P53_MUTATED
CASPASE_CASCADE (APOPTOSIS)
MYC_AMPLIFIED
HER2_AMPLIFIED
Selected Pathways
Integrated Subtypes


A
BLCA
HNSC
LUAD
LUSC
Chromosome
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
NFE2L2
PIK3CA / SOX2
TP63
TERT
EGFR
WHSC1L1
MYC
CDKN2A
PTEN
RB1
TP53
Deletion
Neutral
Amplification
B
BLCA
HNSC
LUAD
LUSC
PIK3CA
MLL3
PLEC
FRG1B
RNF213
LRRK2
HUWE1
KEAP1
FN1
AKAP9
CASP8
FBXW7
ZNF91
NF1
KIAA1109
C
BLCA
HNSC
LUAD
LUSC
TP53
SYNE1
MLL2
CDKN2A
NOTCH1
NFE2L2
NSD1
ASPM
EP300
LCP1
HRAS
D
ASCOM Complex and KDM6A
Coverage: 40.1% (203 / 506 samples)
MLL3 (107)
MLL2 (88)
KDM6A (50)
= 2 samples
CUL3, NFE2L2 and interactors
Coverage: 54.0% (273 / 506 samples)
CUL3 (108)
NOTCH1 (90)
NFE2L2 (85)
KLF5 (45)
KEAP1 (28)
FBXW7 (26)
CCNE1 (21)
SHPRH (11)
= 2 samples
Clusters
1-LUAD-enriched
2-Squamous-like
3-BRCA/Luminal
6-UCEC
7-COADREAD
8-BLCA
Cancer types
BLCA
LUAD
LUSC
HNSC
Mutation types
SNV
Amplification
Deletion


A
dNp63a_(tetramer)_(complex)
TAp63g_(tetramer)_(complex)
DNA_damage_(abstract)
keratinocyte_differentiation_(abstract)
degradation_(abstract)
BDKRB2
RRAD
CEBPD
FOSL2
ATM
STXBP4
MDM2
CDKN2A
QSN726
YAP1
COL5A1
TCF7L1
BRCA2
KRT5
KRT14
HELLS
CCNB2
RUNX1
TP63
TOP2A
DLX6
DLX5
T
FASN
POU2F2
PERP
VDR
HES1
ITCH
SFN
RAB38
ADA
IL1A
AXL
SEC14L2
NRG1
MRE11A
PRM1
AEN
GADD45A
CDKN1A
IGFBP3
NQO1
FDXR
ITGA3
FLOT2
TP53I3
JAG1
SMARCD3
CHUK
TRAF4
EVPL
MIR130B_(miRNA)
ITGB4
DST-3
SERPINB5
IKBKB
S100A2
EGR2
MFGE8
GPX2
PML
DICER1
OGG1
DHRS3
YWHAQ
SHH
TFAP2C
CLCA2
Color Key
Activated in both subtypes
Repressed in both subtypes
Activated in C2-Squamous-like Repressed in C4-BRCA/Basal
Activated in one subtype
Repressed in one subtype
Shape Key
More significant in C2-Squamous-like
More significant in C4-BRCA/Basal
B
RSEM Values
0
10
20
30
40
50
60
C2
C4
C9
TAp63
dNp63
TAp73
dNp73
C2−Squamous-like
C4−BRCA/Basal
C9−OV
C
Mean P-Shift: -0.94
Mean P-Shift: -0.34
Mean P-Shift: 0.87
Outer Ring: P-Shift
6th Ring: PARADIGM Downstream Signal
5th Ring: PARADIGM Upstream Signal
4th Ring: Expression
3rd Ring: Copy Number
2nd Ring: Mutation Type
1st Ring: COCA Subtype
TP53
COCA Subtype
C2-Squamous-like
C4-BRCA/Basal
C9-OV
p53 Mutation Type
Wild Type
Missense
Truncating
D
COCA Subtype
p53 Mutation
TP63
dNp63
TAp63
TP73
dNp73
TAp73
Integrative Subtype
C2-Squamous-like
C4-BRCA/Basal
C9-OV
p53 Mutation
Wild Type
Missense
Truncating
Intensity Scale
<-2
-1
0
1
>2
I
II
III


A
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
C1−LUAD−enriched
C2−Squamous−like
C8−BLCA
Log Rank p=0.01
0
20
40
60
80
100
120
Months
B
Color Key and Histogram
Count
0
40
80
−1
0
1
Value
C2-Squamous -like
C8-BLCA
SRC
GATA3
INPP4B
HER2
RAB25
BRAF
53BP1
ECADHERIN
BETACATENIN
NCADHERIN
CHK1
CMYC
LCK
PAI1
FIBRONECTIN
HEREGULIN
G6PD
C
KEAP1
STK11
KRAS
NLRP3
BCLAF1
HGF
BRAF
MMP2
MLL3
NF1
ABCB1
WRN
SMARCA4
ARID2
CAD
AKAP9
HUWE1
PDGFRA
CDKN2A
TP53
NOTCH1
MLL2
SYNE1
NFE2L2
NSD1
EP300
CASP8
HRAS
ZNF91
HLA-A
RAC1
RPSAP58
RASA1
RNF213
HLA-B
KDM6A
CDKN1A
ARID1A
ERCC2
FGFR3
MLL
WHSC1
ELF3
STAG2
TRIO
RHOA
TXNIP
MDC1
ARHGAP32
C1-LUAD-Enriched
C8-BLCA
C2-Squamous-Like
0
0.01
0.01
0
Test for differential mutation frequency (FDR)
under-represented
non-significance
enriched
D
C2-Squamous -like
C8-BLCA
GP17
GP6
GP16
GP12
GP1
GP3
GP8
GP2
GP4
GP11
GP9
GP7
−2
−1
0
1
2
E
Key
Higher in C8-BLCA
Higher in C2-Squamous-like
Activation
Inhibition
Abstract Concept
Complex
Protein
PTPN2
IL12RB2
PRKCD
CASP1
STAT1_(dimer)_(complex)
TBX21
MAPK14
CCL3
CCL3L
GZMB
GZMA
IL12/IL12R/TYK2/JAK2_(complex)
IRF1
SOCS3
IL17A
IL17F
STAT3_(dimer)_(complex)
STAT4_(dimer)_(complex)
IL1R1
IL2
IL4
IL13
PIM1
MYC
FASLG
IL23/IL23R/JAK2/TYK2_(complex)
YWHAG
STAT5_(dimer)_(family)
PRF1
CCL2
IL24
CXCL1
ALOX12B
CISH
FOXP3
CDK6
CNKSR1
CSN2

Results

- 1/10 cancer patients classified differently based on new taxonomy

- Guide new therapeutic decisions

- Bladder cancers most heterogeneous

- “Cell-of-origin” features still dominate molecular taxonomy, but refined taxonomy provides independent info