# Inter- and intra – tumor heterogeneity

Ulrich Mansmann
IBE, LMU
Ulrich.mansmann@lmu.de

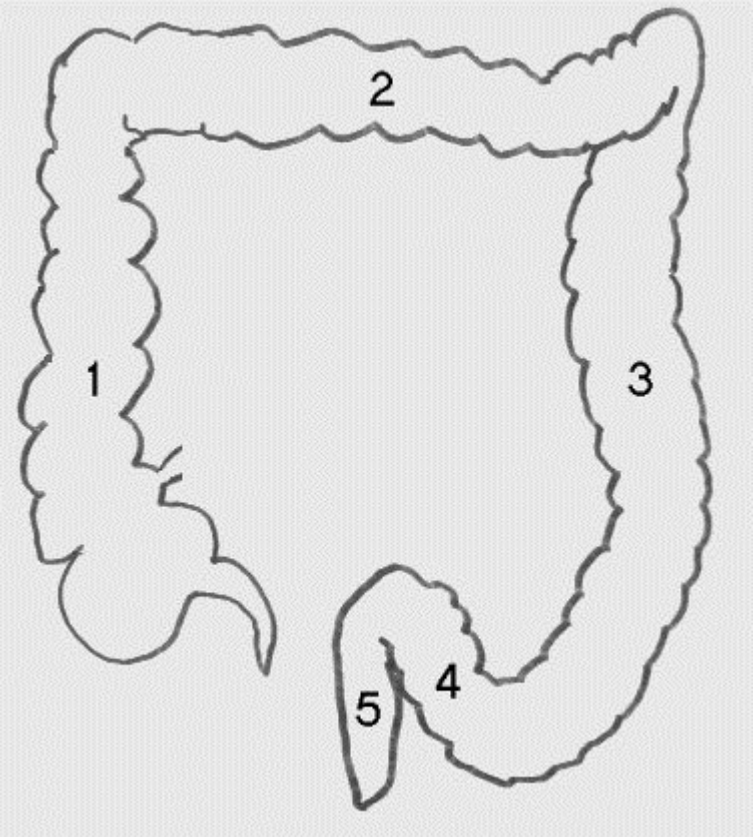
Published in final edited form as:
*Nature*. ; 487(7407): 330–337. doi:10.1038/nature11252.

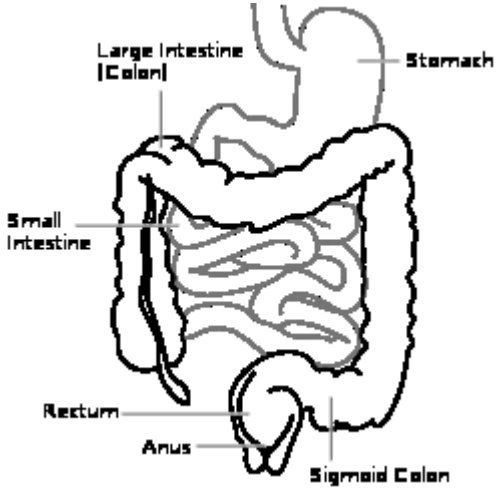
# Comprehensive Molecular Characterization of Human Colon and Rectal Cancer

**The Cancer Genome Atlas Network**

## Summary

To characterize somatic alterations in colorectal carcinoma (CRC), we conducted genome-scale analysis of 276 samples, analyzing exome sequence, DNA copy number, promoter methylation, mRNA and microRNA expression. A subset (97) underwent low-depth-of-coverage whole-genome sequencing. 16% of CRC have hypermutation, three quarters of which have the expected high microsatellite instability (MSI), usually with hypermethylation and *MLH1* silencing, but one quarter has somatic mismatch repair gene mutations. Excluding hypermutated cancers, colon and rectum cancers have remarkably similar patterns of genomic alteration. Twenty-four genes are significantly mutated. In addition to the expected *APC*, *TP53*, *SMAD4*, *PIK3CA* and *KRAS* mutations, we found frequent mutations in *ARID1A*, *SOX9*, and *FAM123B/WTX*. Recurrent copy number alterations include potentially drug-targetable amplifications of *ERBB2* and newly discovered amplification of *IGF2*. Recurrent chromosomal translocations include fusion of *NAV2* and WNT pathway member *TCF7L1*. Integrative analyses suggest new markers for aggressive CRC and important role for *MYC*-directed transcriptional activation and repression.

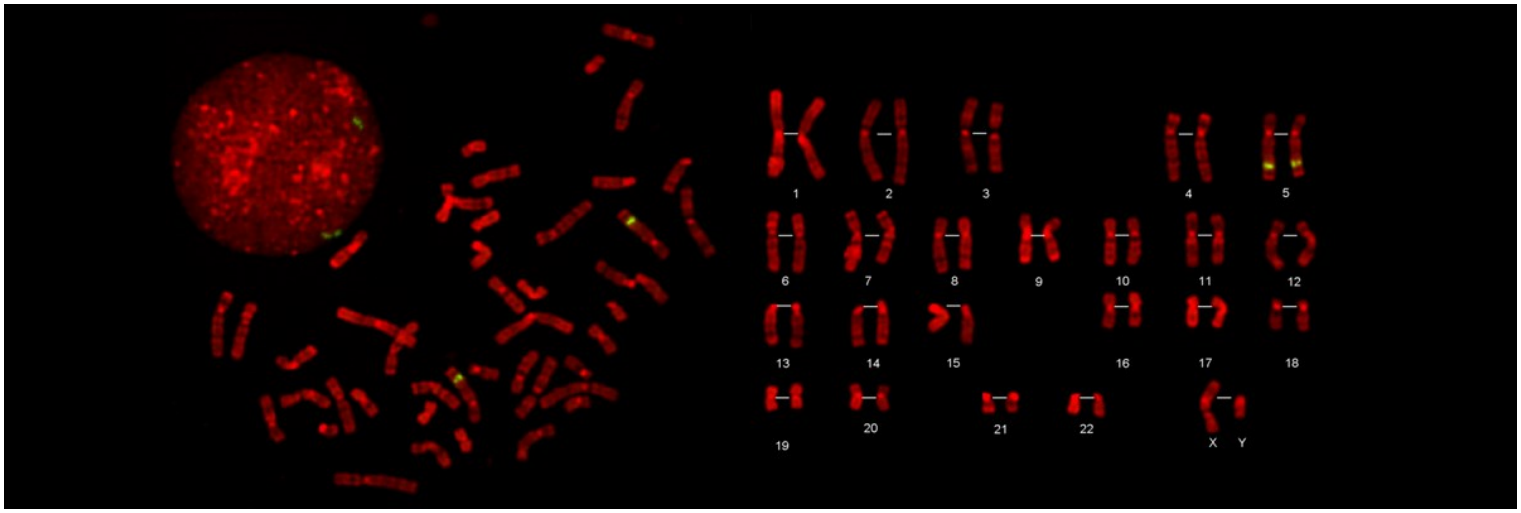



1. *Colon ascendens*
2. *Colon transversum*
3. *Colon descendens*
4. *Colon sigmoideum*
5. *Rectum*

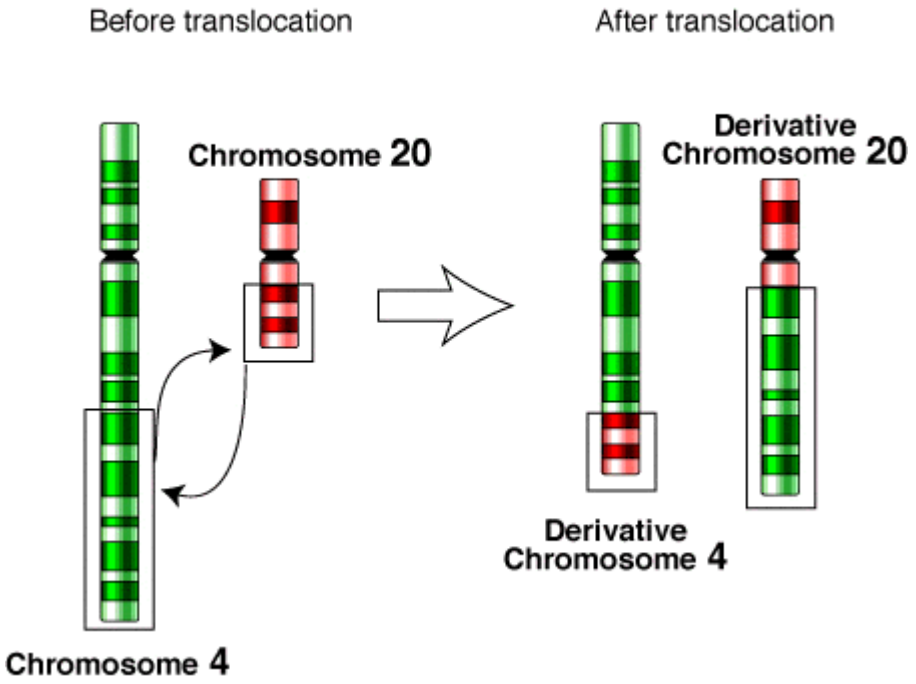
Chromosomal Stability

Chromosome instability (CIN) involves the unequal distribution of DNA to daughter cells upon mitosis, resulting in a failure to maintain euploidy (the correct number of chromosomes) leading to aneuploidy (incorrect number of chromosomes). In other words, the daughter cells do not have the same number of chromosomes as the cell they originated from. These changes have been studied in solid tumors, which may or may not be cancerous.



Chromosomal Translocation

A chromosome translocation is a chromosome abnormality caused by rearrangement of parts between nonhomologous chromosomes. A gene fusion may be created when the translocation joins two otherwise separated genes, the occurrence of which is common in cancer. It is detected on cytogenetics or a karyotype of affected cells. Translocations can be balanced (in an even exchange of material with no genetic information extra or missing, and ideally full functionality) or unbalanced (where the exchange of chromosome material is unequal resulting in extra or missing genes).




Translocation 4-20

# Exome Sequence

Exome sequencing (also known as targeted exome capture) is an efficient strategy to selectively sequence the coding regions of the genome as a cheaper but still effective alternative to whole genome sequencing.

Exons are short, functionally important sequences of DNA which represent the regions in genes that are translated into protein and the untranslated region (UTR) flanking them.

UTRs are usually not included in exome studies. In the human genome there are about 180,000 exons: these constitute about 1% of the human genome, which translates to about 30 megabases (Mb) in length. It is estimated that the protein coding regions of the human genome constitute about 85% of the disease-causing mutations.

Figure 1

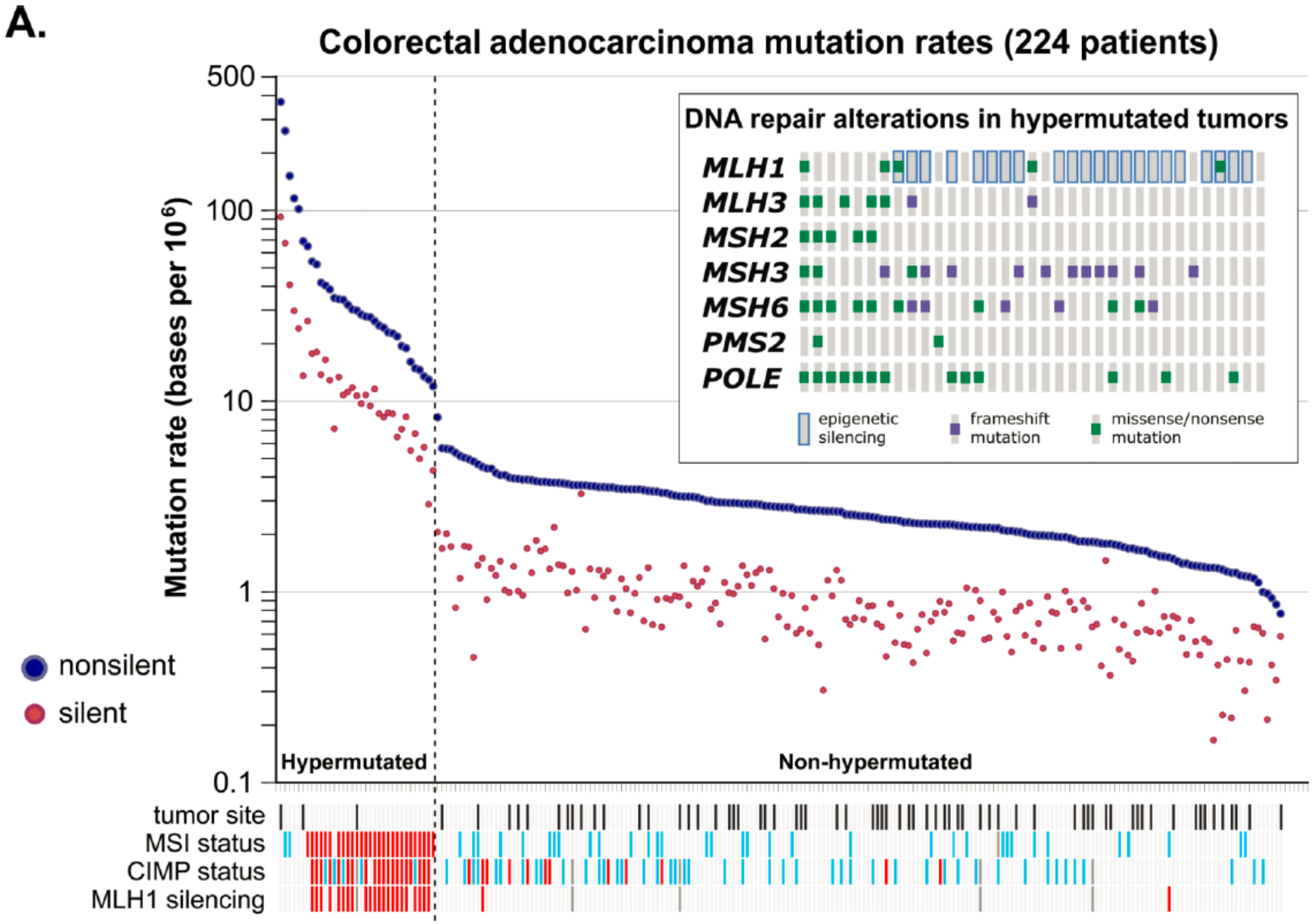
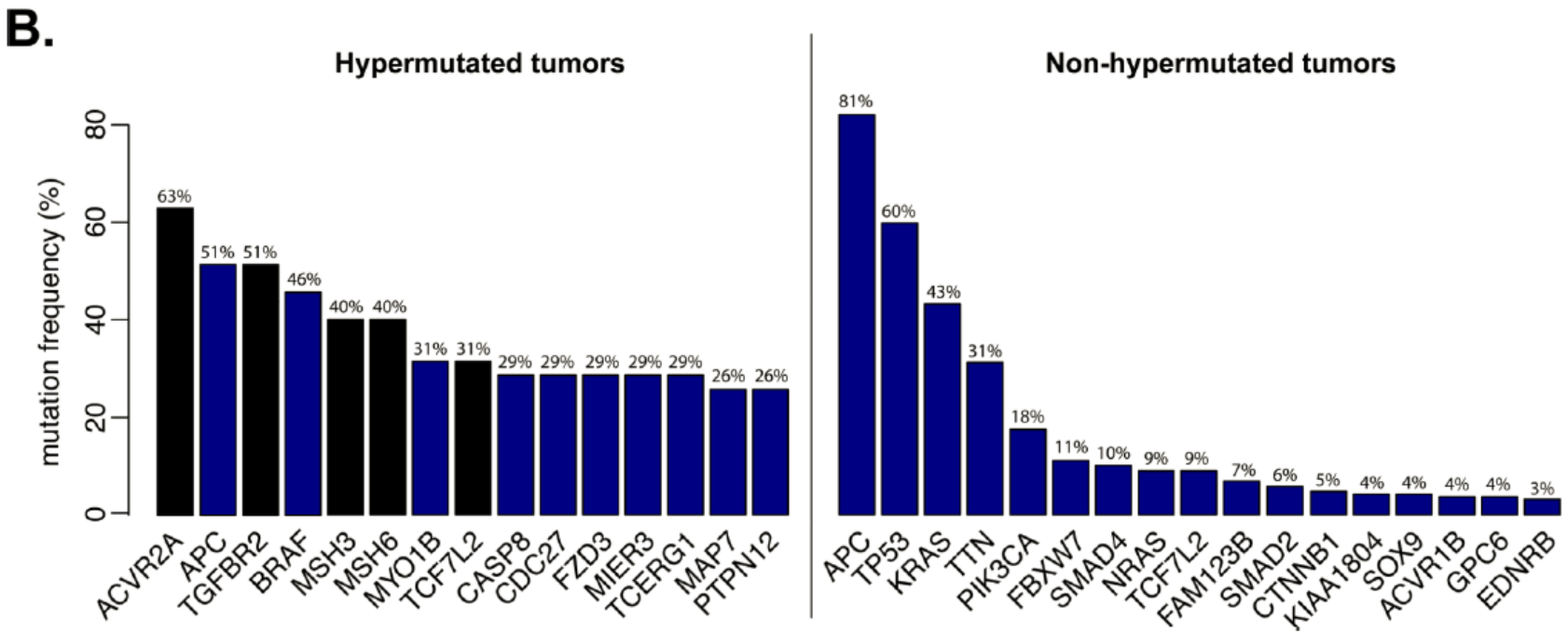
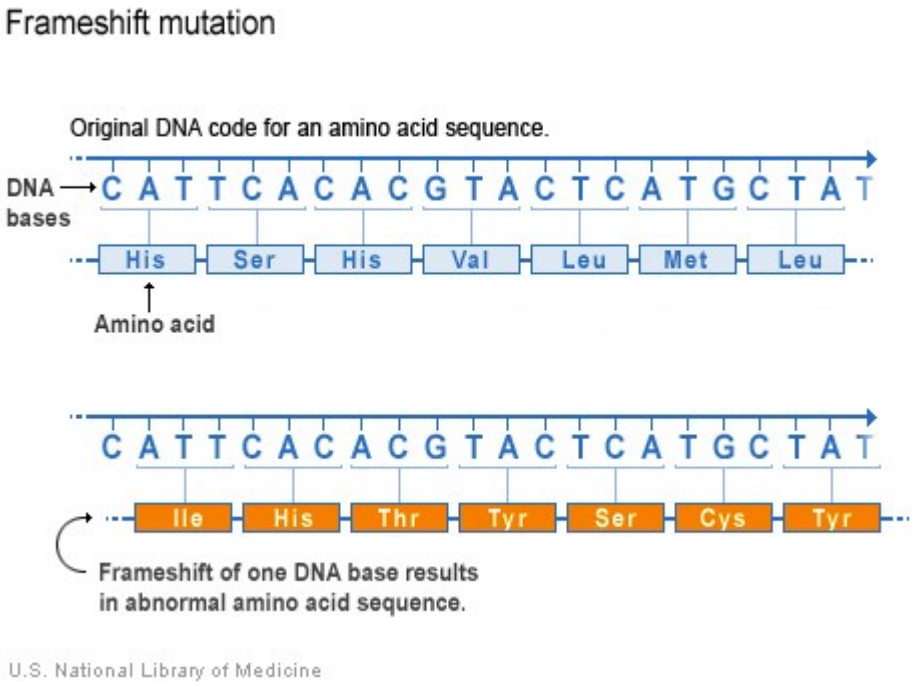

A.
Colorectal adenocarcinoma mutation rates (224 patients)
Mutation rate (bases per 10^6)
500
100
10
1
0.1
DNA repair alterations in hypermutated tumors
MLH1
MLH3
MSH2
MSH3
MSH6
PMS2
POLE
epigenetic silencing
frameshift mutation
missense/nonsense mutation
nonsilent
silent
Hypermutated
Non-hypermutated
tumor site
MSI status
CIMP status
MLH1 silencing

Figure 1


B.
Hypermutated tumors
Non-hypermutated tumors
mutation frequency (%)
0
20
40
60
80
63%
51%
51%
46%
40%
40%
31%
31%
29%
29%
29%
29%
29%
26%
26%
ACVR2A
APC
TGFBR2
BRAF
MSH3
MSH6
MYO1B
TCF7L2
CASP8
CDC27
FZD3
MIER3
TCERG1
MAP7
PTPN12
81%
60%
43%
31%
18%
11%
10%
9%
9%
7%
6%
5%
4%
4%
4%
4%
3%
APC
TP53
KRAS
TTN
PIK3CA
FBXW7
SMAD4
NRAS
TCF7L2
FAM123B
SMAD2
CTNNB1
KIAA1804
SOX9
ACVR1B
GPC6
EDNRB

# Frameshift Mutation
# Nonsense Mutation




U.S. National Library of Medicine

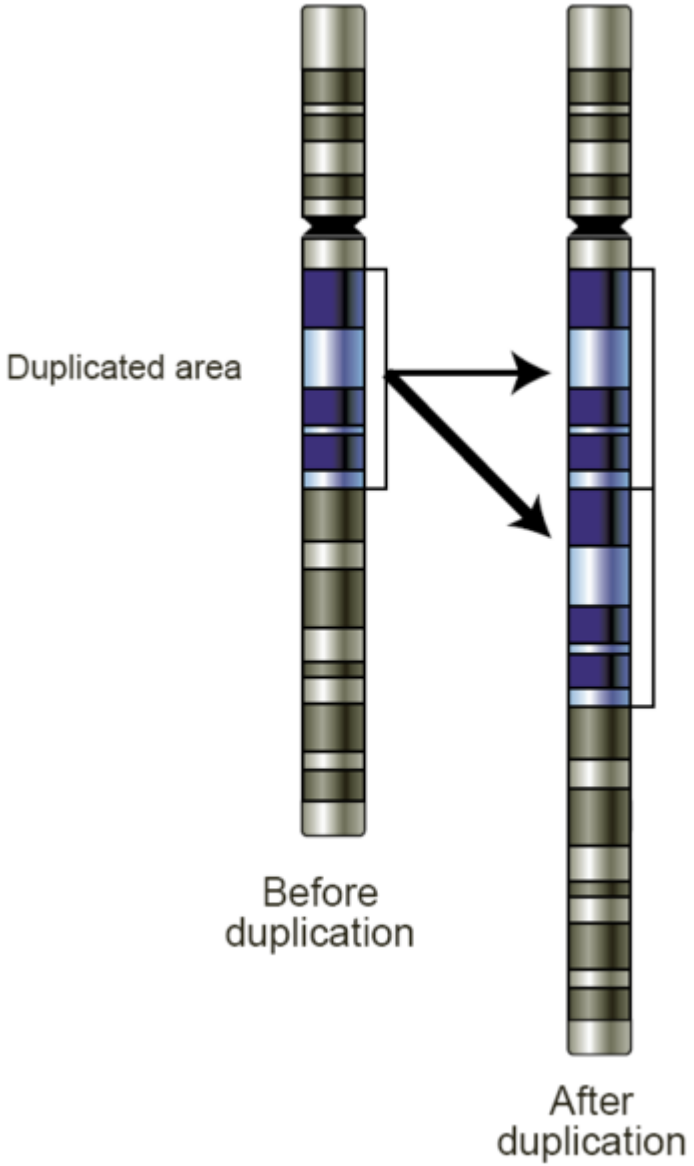
A frameshift mutation is a genetic mutation caused by a deletion or insertion in a DNA sequence that shifts the way the sequence is read.

Frameshift mutations arise when the normal sequence of codons is disrupted by the insertion or deletion of one or more nucleotides, provided that the number of nucleotides added or removed is not a multiple of three. For instance, if just one nucleotide is deleted from the sequence, then all of the codons including and after the mutation will have a disrupted reading frame. This can result in the incorporation of many incorrect amino acids into the protein. In contrast, if three nucleotides are inserted or deleted, there will be no shift in the codon reading frame; however, there will be either one extra or one missing amino acid in the final protein. Therefore, frameshift mutations result in abnormal protein products with an incorrect amino acid sequence that can be either longer or shorter than the normal protein.

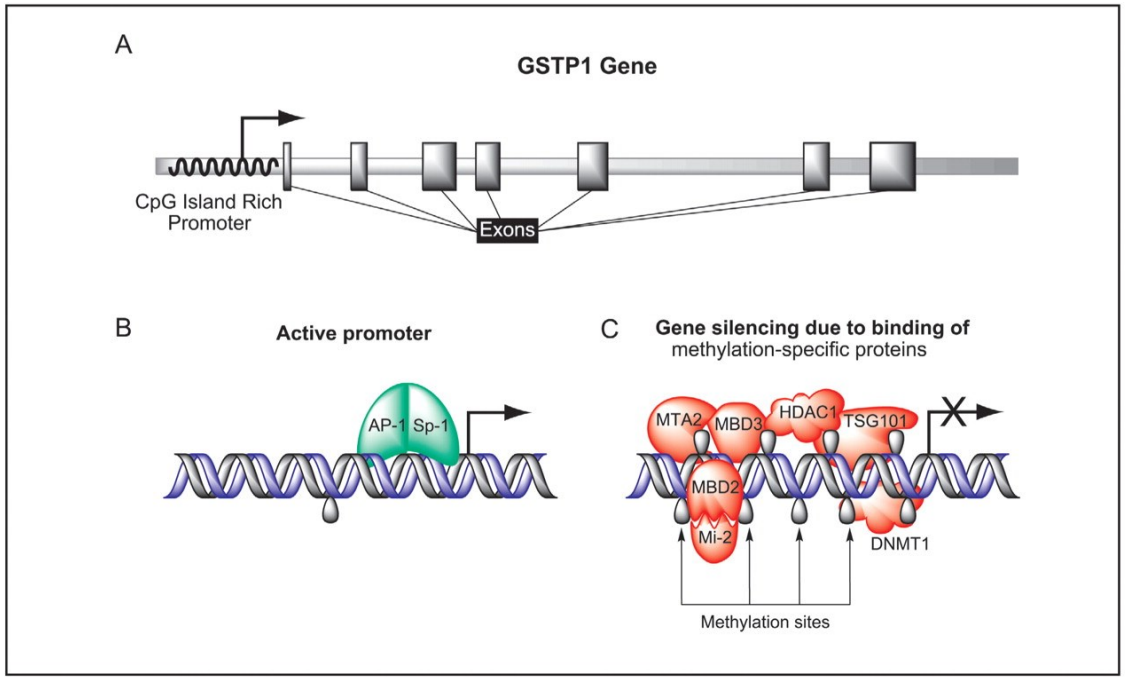
# DNA copy number

Copy-number variations (CNVs)—a form of structural variation—are alterations of the DNA of a genome that results in the cell having an abnormal number of copies of one or more sections of the DNA.

CNVs correspond to relatively large regions of the genome that have been deleted (fewer than the normal number) or duplicated (more than the normal number) on certain chromosomes. For example, the chromosome that normally has sections in order as A-B-C-D might instead have sections A-B-C-C-D (a duplication of "C") or A-B-D (a deletion of "C").



# Promoter Methylation




GSTP1 gene silencing by CpG island methylation.

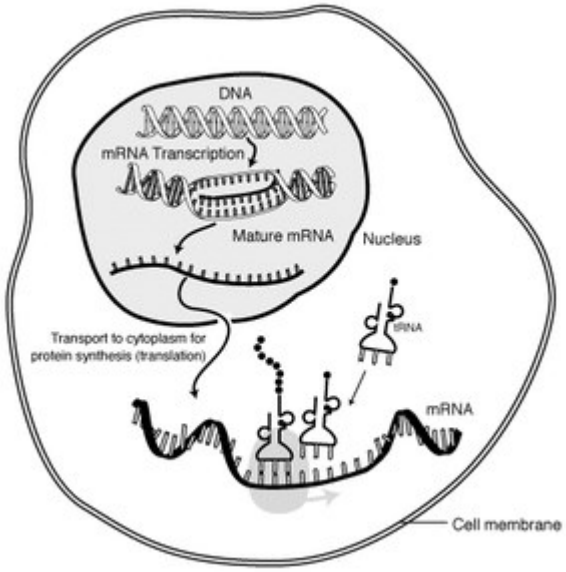
(A) The glutathione S-transferase Pi (GSTP1) gene has an abundance of CpG sites in the promoter region.

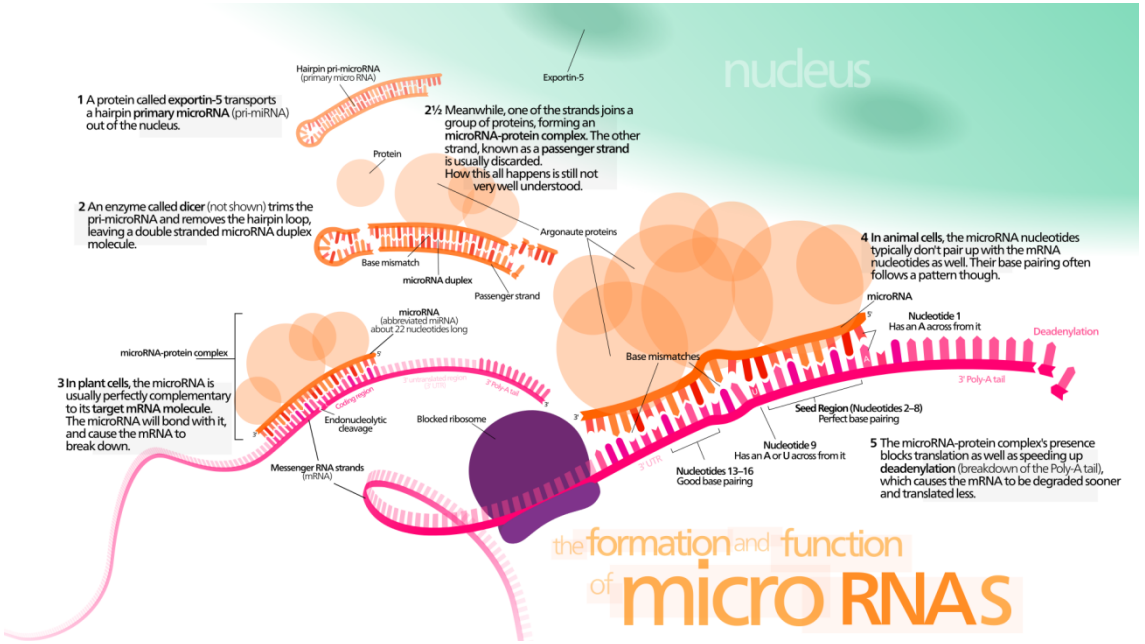
(B) In healthy prostate tissue, transcription factors such as AP-1 and Sp-1 bind to the promoter and stimulate transcription.

(C) In prostate cancer, GTSP1 CpG islands are hypermethylated. Methylation-specific proteins compete with transcription factors for binding to methylated CpG sites. In this way, hypermethylated genes, such as GSTP1, are effectively silenced.

# mRNA and miRNA

A microRNA (abbr. miRNA) is a small non-coding RNA molecule (ca. 22 nucleotides) which functions in transcriptional and post-transcriptional regulation of gene expression. miRNAs function via base-pairing with complementary sequences within mRNA molecules, usually resulting in gene silencing via translational repression or target degradation. The human genome may encode over 1000 miRNAs, which may target about 60% of mammalian genes and are abundant in many human cell types.



MSI – Microsatellite Instability

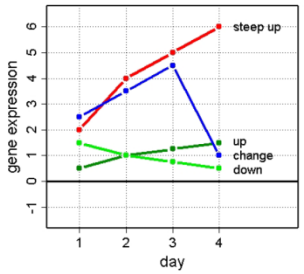
Microsatellites are repeated sequences of DNA.

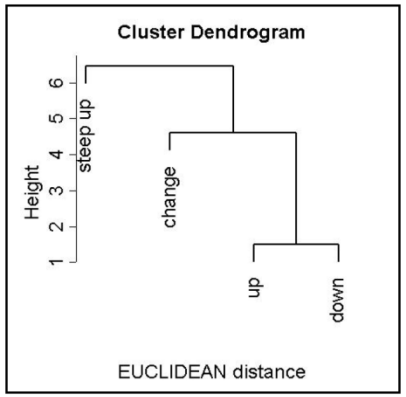
Although the length of these microsatellites is highly variable from person to person, each individual has microsatellites of a set length. These repeated sequences are common, and normal. The most common microsatellite in humans is a dinucleotide repeat of CA, which occurs tens of thousands of times across the genome.

In cells with mutations in DNA repair genes, however, some of these sequences accumulate errors and become longer or shorter. The appearance of abnormally long or short microsatellites in an individual's DNA is referred to as microsatellite instability.

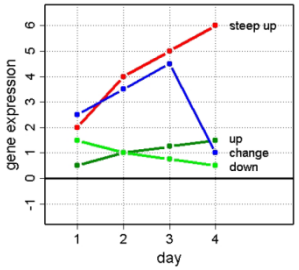
Microsatellite instability (MSI) is a condition manifested by damaged DNA due to defects in the normal DNA repair process. Sections of DNA called microsatellites, which consist of a sequence of repeating units of 1-6 base pairs in length, become unstable and can shorten or lengthen. Microsatellites are also known as simple sequence repeats (SSRs).

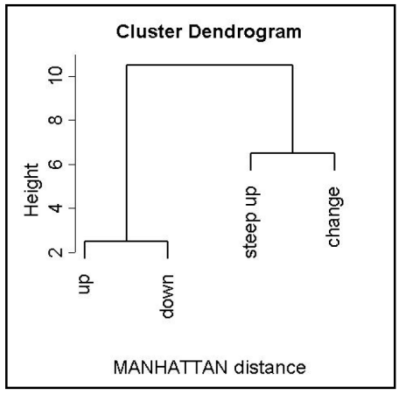
- **Euclidean distance**
  **Similar values are clustered together**



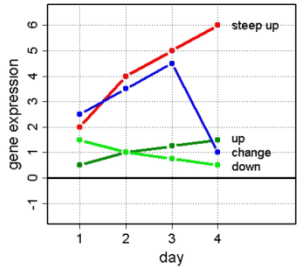



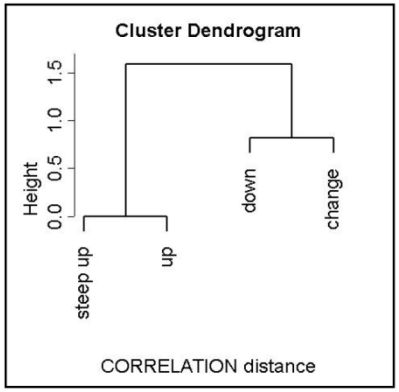

- **Manhattan distance**
  **Similar values are clustered together (robust)**



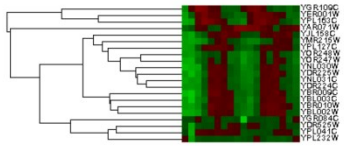

- **Correlation distance**
  **Similar trends are clustered together**



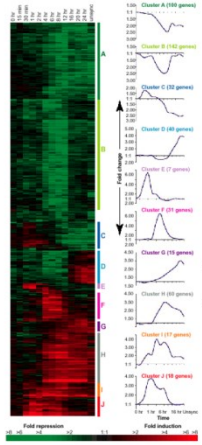




**Iyer et al., Science, Jan 1999:**

**Genes from functinal classes are clustered together.**



**Cluster dendrogram**

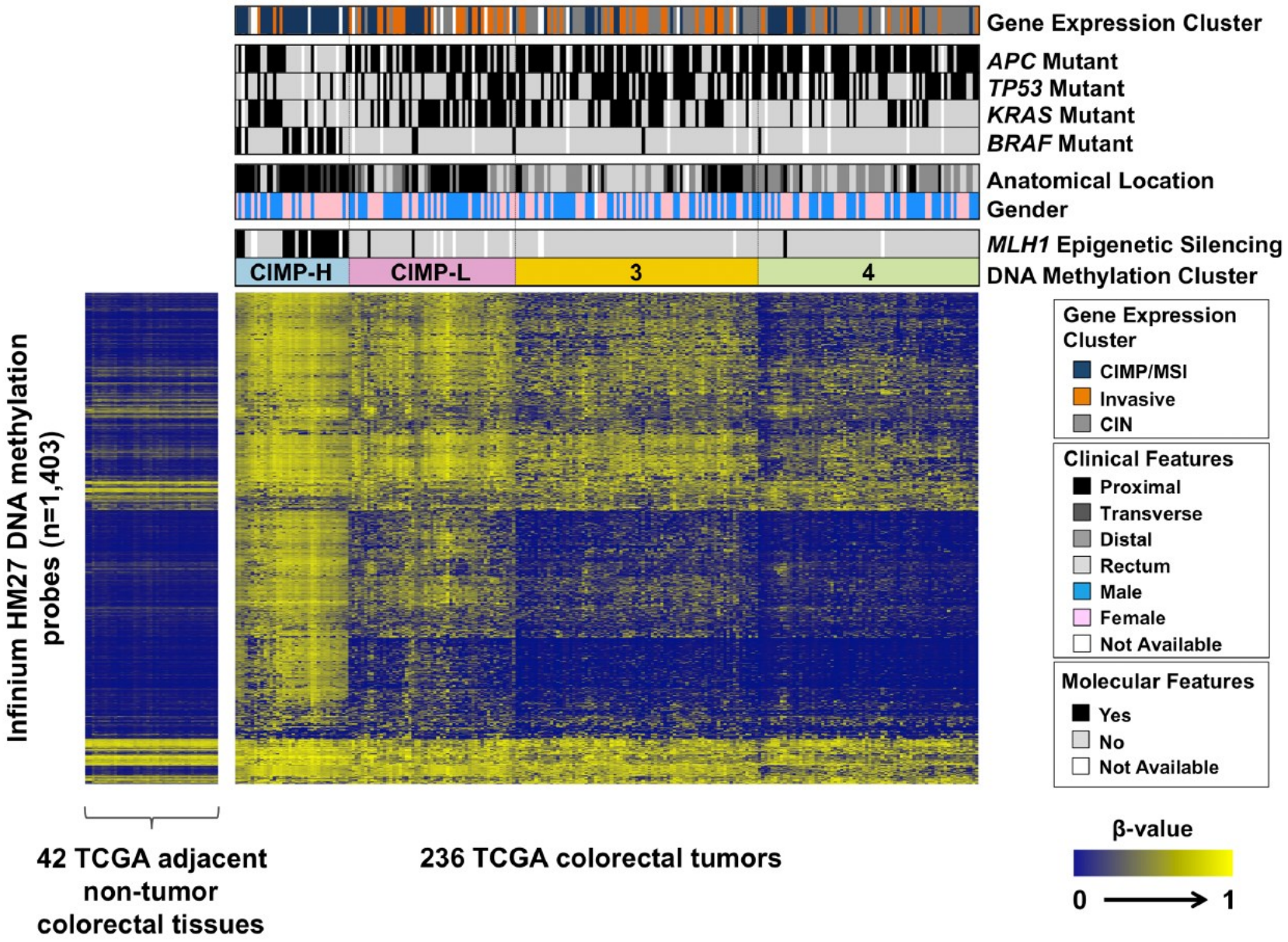


**Cluster average patterns**

Cluster Analysis – Hierarchical clustering

1. Each object is a cluster of its own
2. Calculate the distance matrix
3. Combine the two objects which are closest to a new cluster
4. Give a value to the new cluster by calculating the mean of the two combined objects
5. Calculate the distance matrix for the new set of clusters
6. Go back to 3
7. Stop if one cluster is left.

Supplementary Figure 1



Supplementary Figure 2

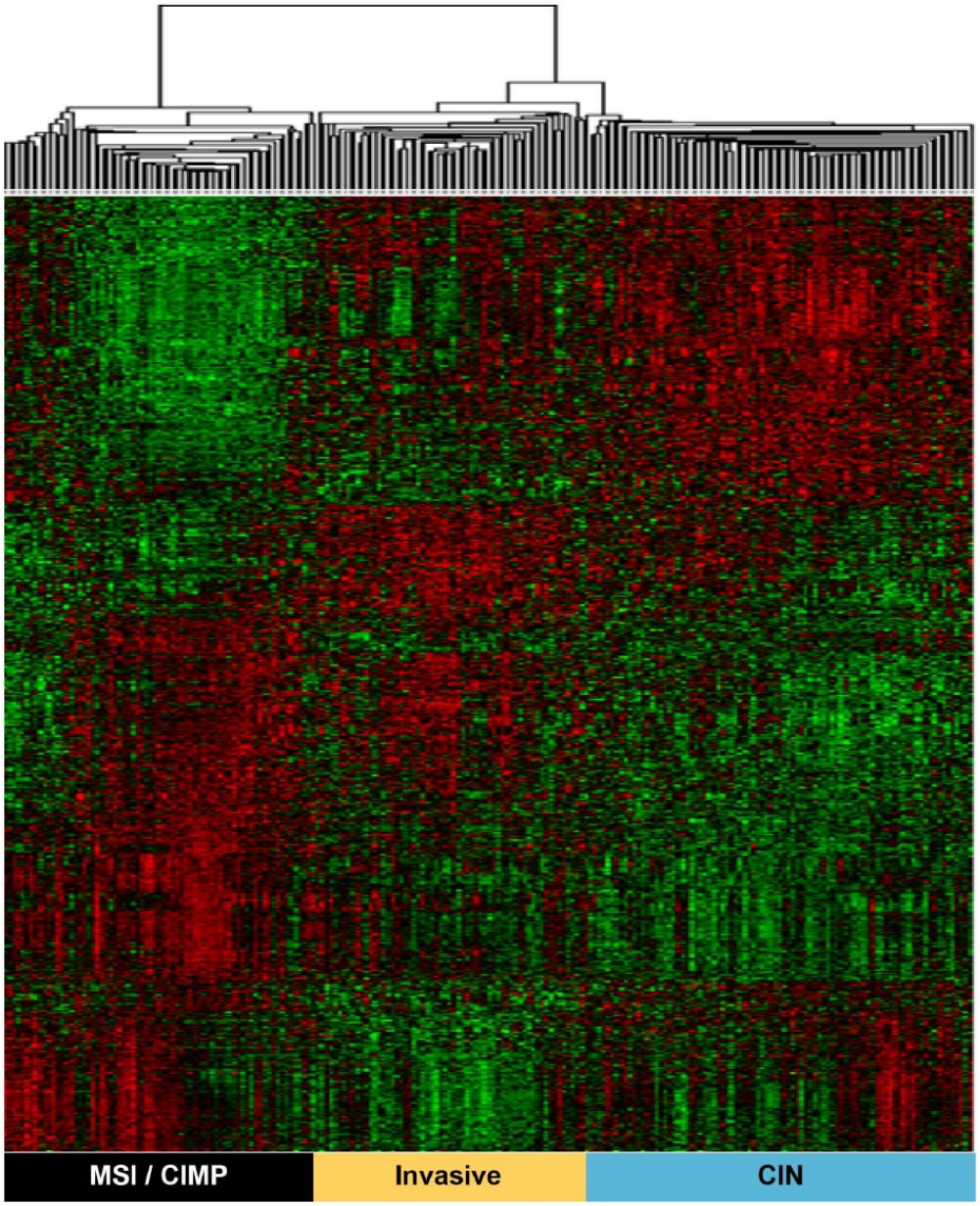
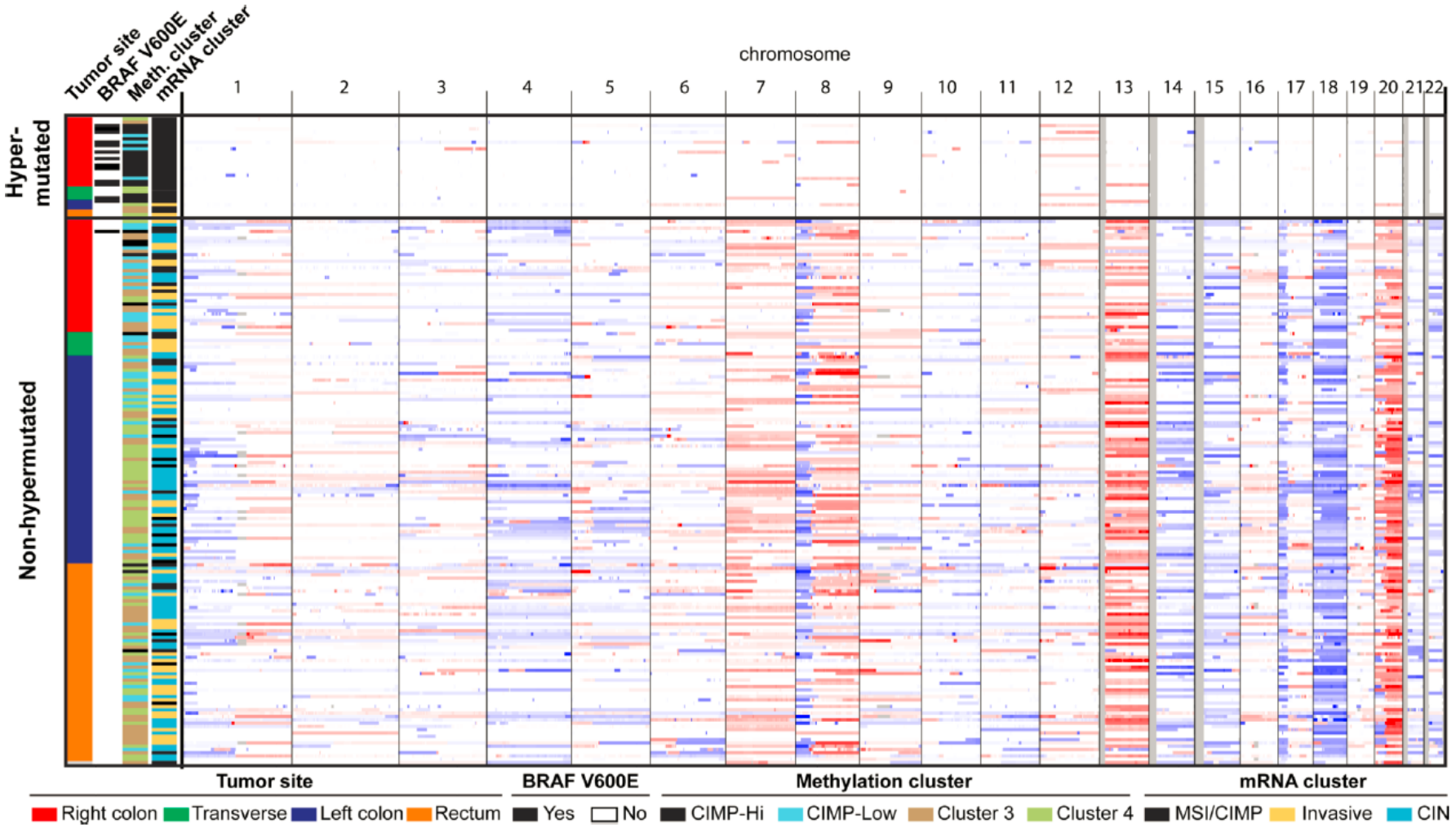
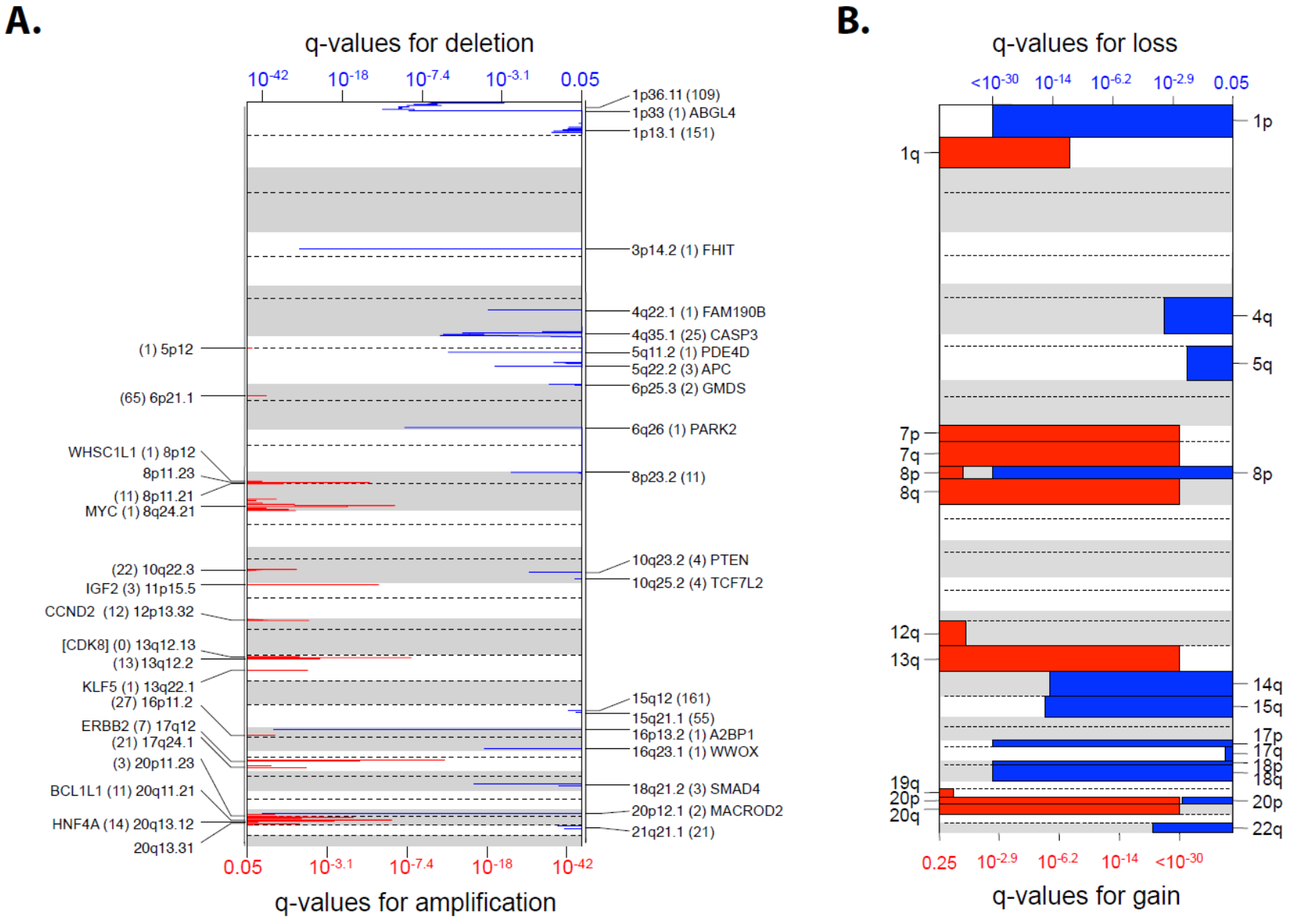


Figure 2

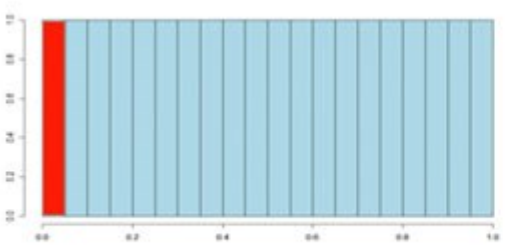


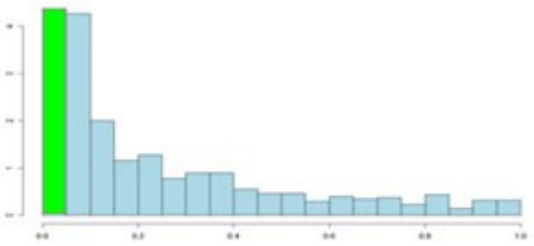

Copy-number gains in shades of red, copy-number losses in shades of blue

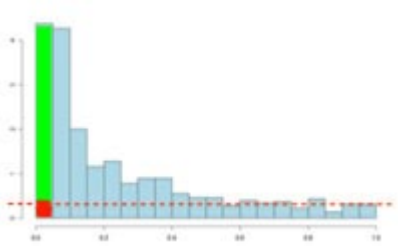

A.
q-values for deletion
10-42
10-18
10-7.4
10-3.1
0.05
1p36.11 (109)
1p33 (1) ABGL4
1p13.1 (151)
3p14.2 (1) FHIT
4q22.1 (1) FAM190B
4q35.1 (25) CASP3
(1) 5p12
5q11.2 (1) PDE4D
5q22.2 (3) APC
(65) 6p21.1
6p25.3 (2) GMDS
6q26 (1) PARK2
WHSC1L1 (1) 8p12
8p11.23
8p23.2 (11)
(11) 8p11.21
MYC (1) 8q24.21
10q23.2 (4) PTEN
(22) 10q22.3
10q25.2 (4) TCF7L2
IGF2 (3) 11p15.5
CCND2 (12) 12p13.32
[CDK8] (0) 13q12.13
(13) 13q12.2
KLF5 (1) 13q22.1
15q12 (161)
(27) 16p11.2
15q21.1 (55)
ERBB2 (7) 17q12
16p13.2 (1) A2BP1
(21) 17q24.1
16q23.1 (1) WWOX
(3) 20p11.23
18q21.2 (3) SMAD4
BCL1L1 (11) 20q11.21
20p12.1 (2) MACROD2
HNF4A (14) 20q13.12
21q21.1 (21)
20q13.31
0.05
10-3.1
10-7.4
10-18
10-42
q-values for amplification
B.
q-values for loss
<10-30
10-14
10-6.2
10-2.9
0.05
1p
1q
4q
5q
7p
7q
8p
8q
8p
12q
13q
14q
15q
17p
17q
18p
18q
19q
20p
20q
20p
22q
0.25
10-2.9
10-6.2
10-14
<10-30
q-values for gain

q-values

q-values are the name given to the adjusted p-values found using an optimised FDR approach. The FDR approach is optimized by using characteristics of the p-value distribution to produce a list of q-values.







# Statistical significance for genomewide studies

**John D. Storey*† and Robert Tibshirani‡**

*Department of Biostatistics, University of Washington, Seattle, WA 98195; and ‡Departments of Health Research and Policy and Statistics, Stanford University, Stanford, CA 94305

9440–9445 | PNAS | August 5, 2003 | vol. 100 | no. 16

Chromosomal Stability

Chromosome instability (CIN) involves the unequal distribution of DNA to daughter cells upon mitosis, resulting in a failure to maintain euploidy (the correct number of chromosomes) leading to aneuploidy (incorrect number of chromosomes). In other words, the daughter cells do not have the same number of chromosomes as the cell they originated from. These changes have been studied in solid tumors, which may or may not be cancerous.

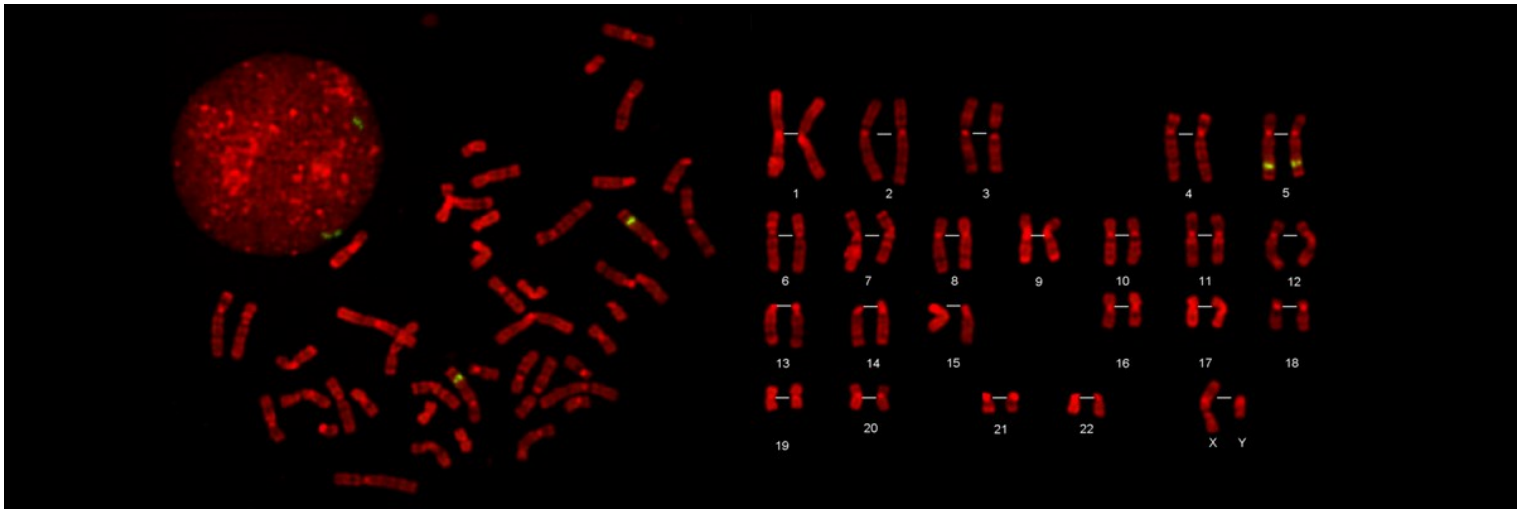


Figure 3

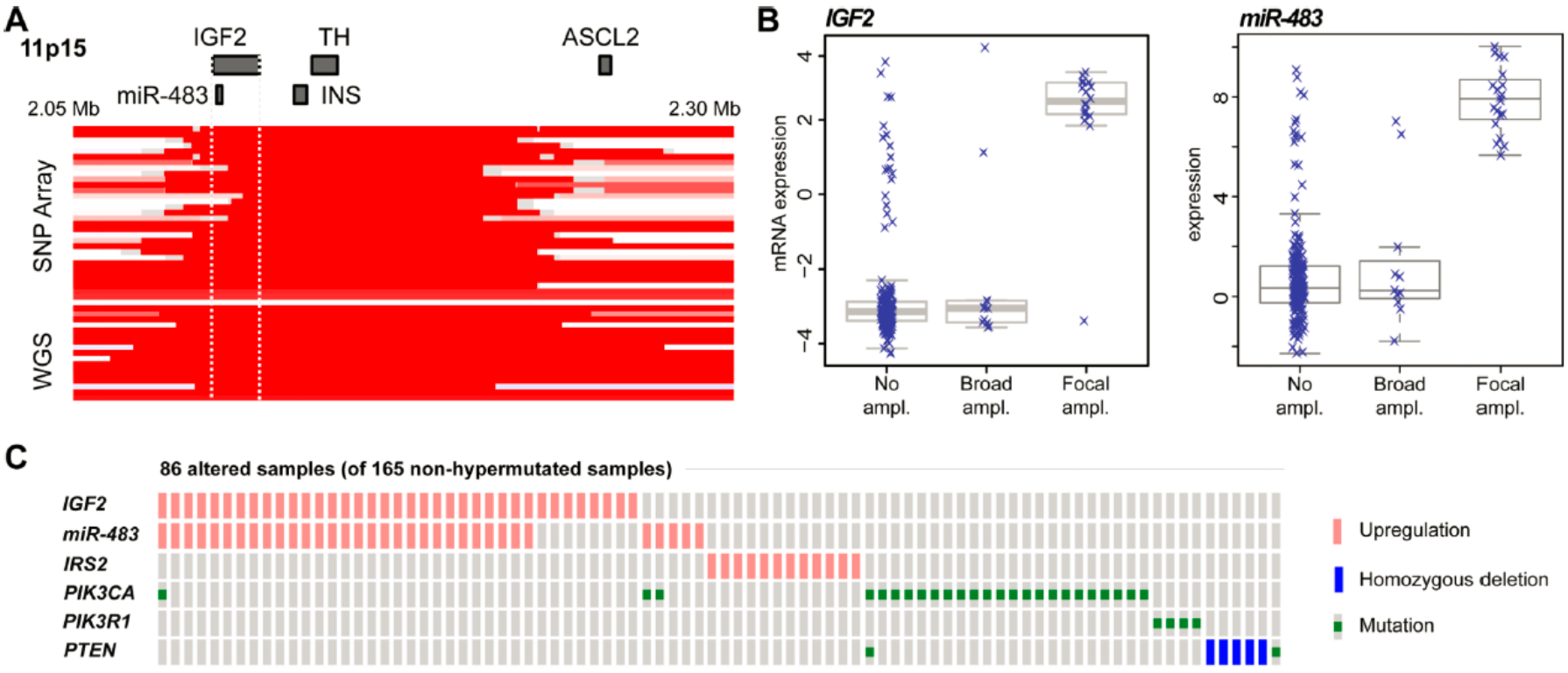

A
11p15
IGF2
TH
ASCL2
miR-483
INS
2.05 Mb
2.30 Mb
SNP Array
WGS
B
IGF2
mRNA expression
4
2
0
−2
−4
No ampl.
Broad ampl.
Focal ampl.
miR-483
expression
8
4
0
No ampl.
Broad ampl.
Focal ampl.
C
86 altered samples (of 165 non-hypermutated samples)
IGF2
miR-483
IRS2
PIK3CA
PIK3R1
PTEN
Upregulation
Homozygous deletion
Mutation

Figure 3

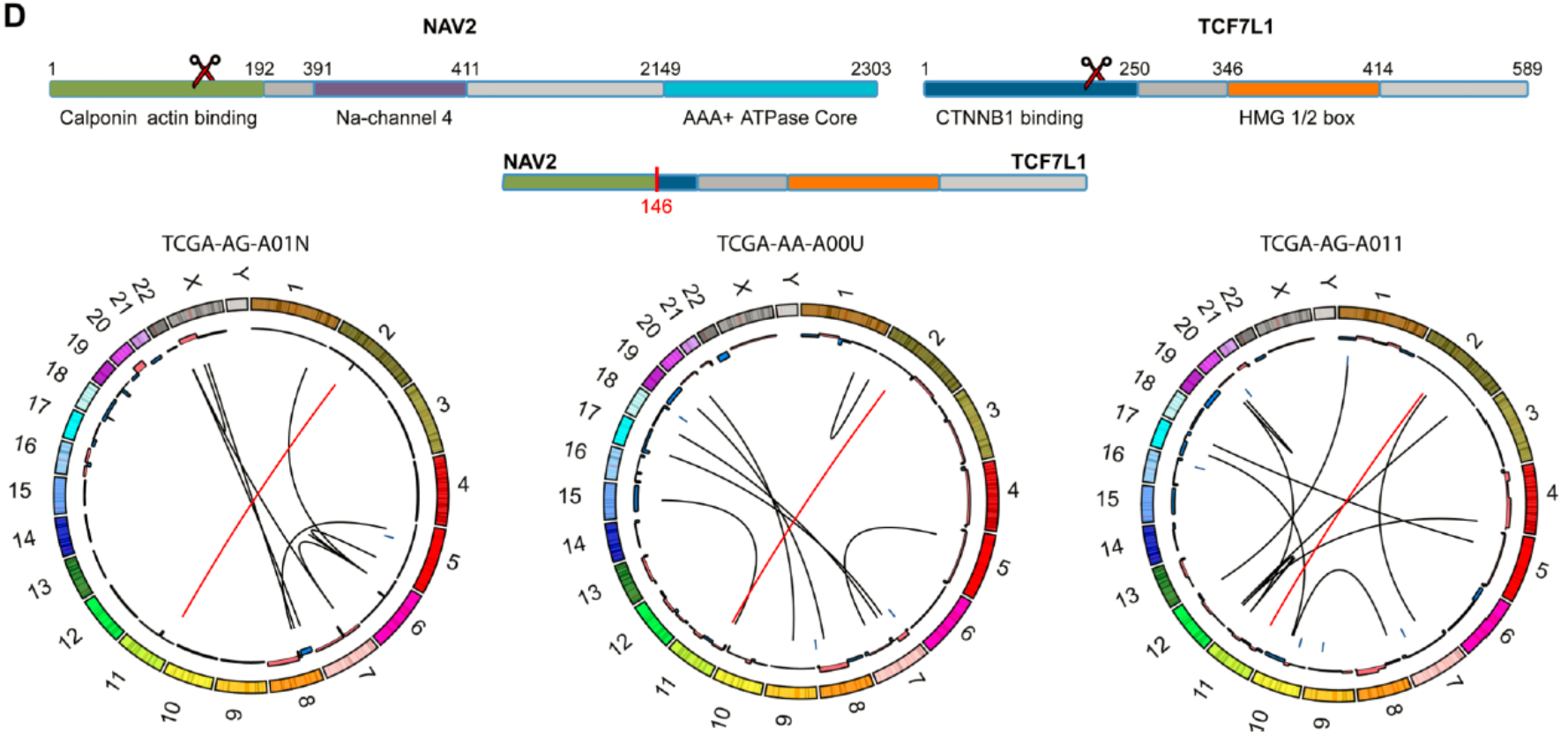
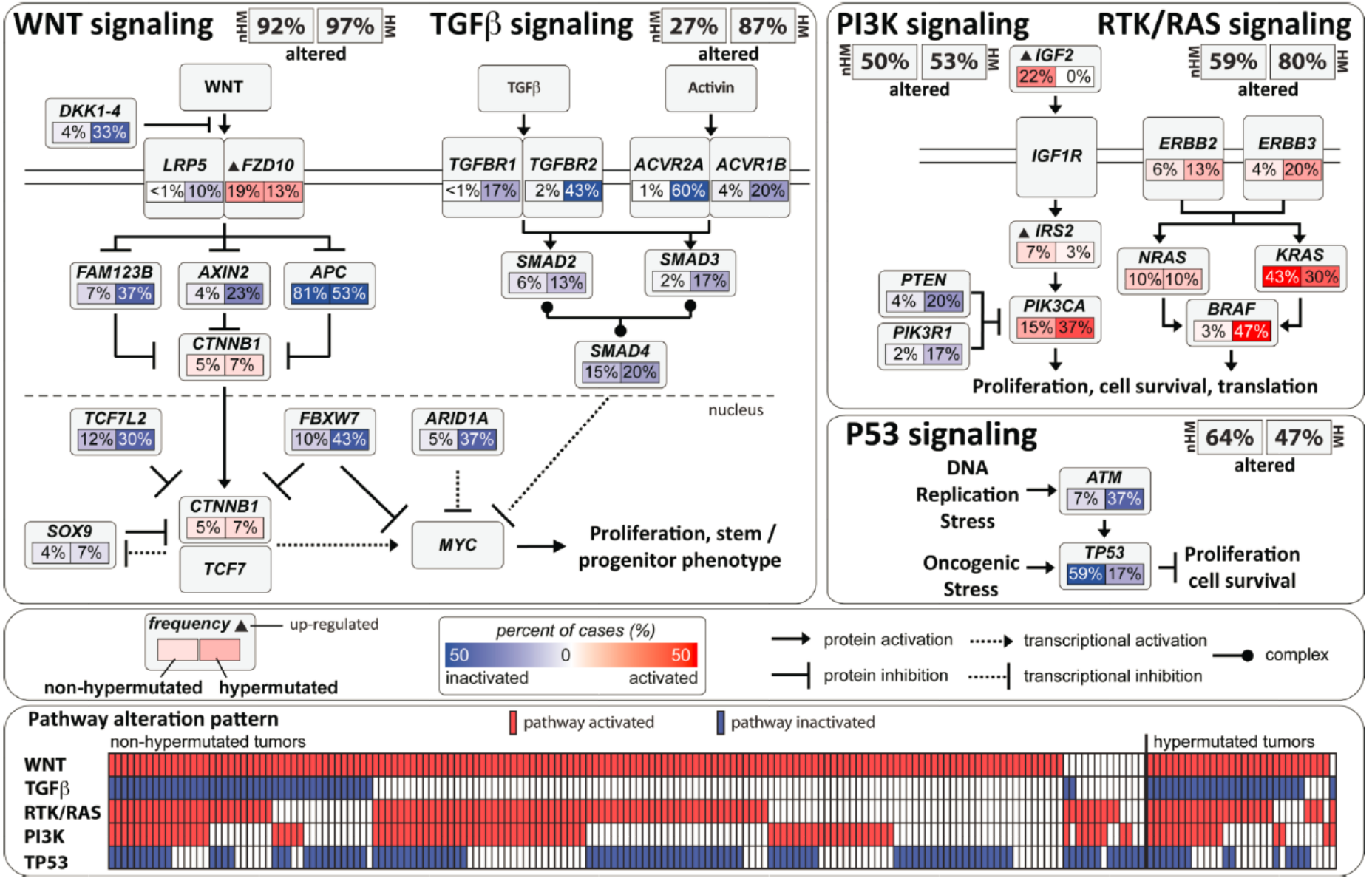
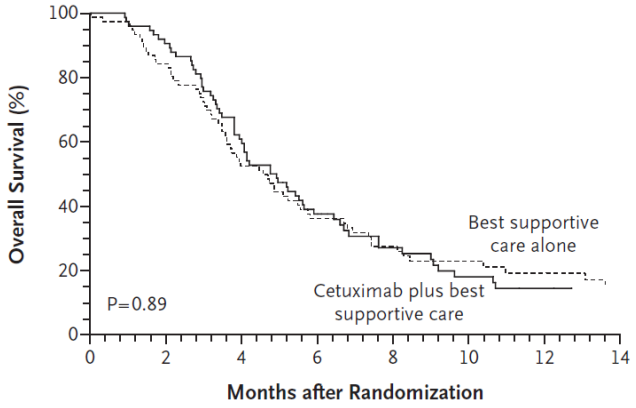

D
NAV2
1
192
391
411
2149
2303
Calponin actin binding
Na-channel 4
AAA+ ATPase Core
TCF7L1
1
250
346
414
589
CTNNB1 binding
HMG 1/2 box
NAV2
TCF7L1
146
TCGA-AG-A01N
TCGA-AA-A00U
TCGA-AG-A011

Figure 4


WNT signaling
nHM 92% 97% HM
altered
WNT
DKK1-4 4% 33%
LRP5 <1% 10%
▲FZD10 19% 13%
FAM123B 7% 37%
AXIN2 4% 23%
APC 81% 53%
CTNNB1 5% 7%
TCF7L2 12% 30%
FBXW7 10% 43%
ARID1A 5% 37%
CTNNB1 5% 7%
TCF7
SOX9 4% 7%
MYC
Proliferation, stem / progenitor phenotype
nucleus
TGFβ signaling
nHM 27% 87% HM
altered
TGFβ
Activin
TGFBR1 <1% 17%
TGFBR2 2% 43%
ACVR2A 1% 60%
ACVR1B 4% 20%
SMAD2 6% 13%
SMAD3 2% 17%
SMAD4 15% 20%
PI3K signaling
nHM 50% 53% HM
altered
▲IGF2 22% 0%
IGF1R
▲IRS2 7% 3%
PTEN 4% 20%
PIK3R1 2% 17%
PIK3CA 15% 37%
RTK/RAS signaling
nHM 59% 80% HM
altered
ERBB2 6% 13%
ERBB3 4% 20%
NRAS 10% 10%
KRAS 43% 30%
BRAF 3% 47%
Proliferation, cell survival, translation
P53 signaling
nHM 64% 47% HM
altered
DNA Replication Stress
Oncogenic Stress
ATM 7% 37%
TP53 59% 17%
Proliferation cell survival
frequency ▲ up-regulated
non-hypermutated
hypermutated
percent of cases (%)
50 0 50
inactivated activated
protein activation
protein inhibition
transcriptional activation
transcriptional inhibition
complex
Pathway alteration pattern
pathway activated
pathway inactivated
non-hypermutated tumors
hypermutated tumors
WNT
TGFβ
RTK/RAS
PI3K
TP53

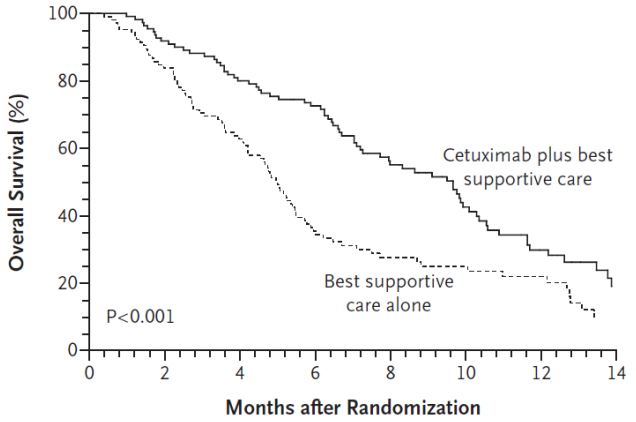
# Kras mutational status and cetuximab by mCRC

**A** Mutated *K-ras*



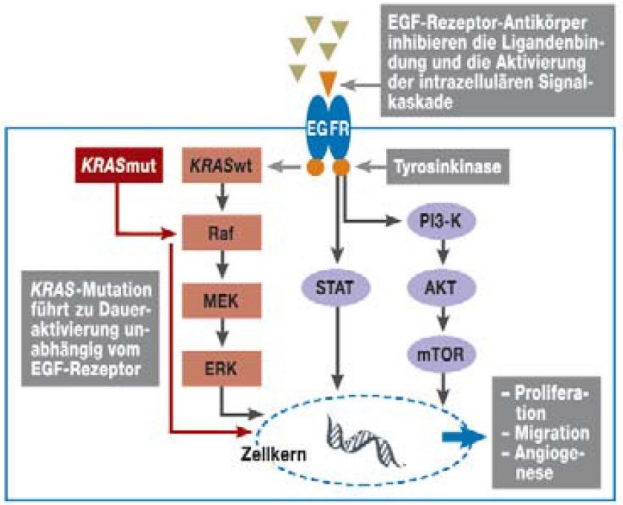

| No. at Risk | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cetuximab plus best supportive care | 75 | 67 | 45 | 26 | 15 | 10 | 7 | 4 |
| Best supportive care alone | 76 | 64 | 39 | 26 | 19 | 12 | 10 | 7 |

**B** Wild-type *K-ras*




| No. at Risk | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cetuximab plus best supportive care | 110 | 101 | 88 | 75 | 48 | 31 | 19 | 8 |
| Best supportive care alone | 105 | 88 | 65 | 34 | 23 | 17 | 12 | 5 |

Treatment with cetuximab, a monoclonal antibody directed against the epidermal growth factor receptor, improves overall and progression-free survival and preserves the quality of life in patients with colorectal cancer that has not responded to chemotherapy. The mutation status of the *K-ras* gene in the tumor may affect the response to cetuximab and have treatment-independent prognostic value.




Deutsches Ärzteblatt, 2009

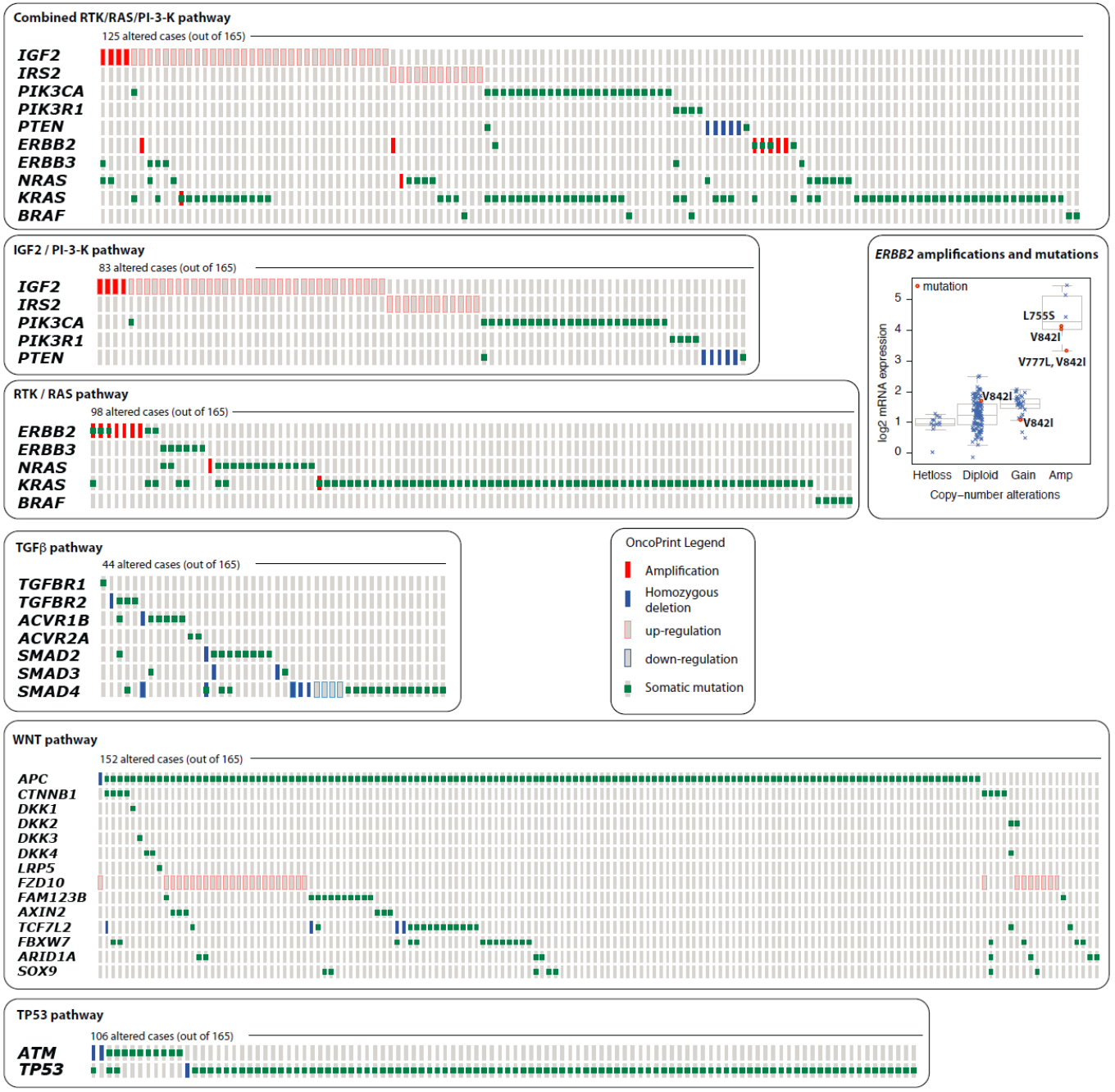
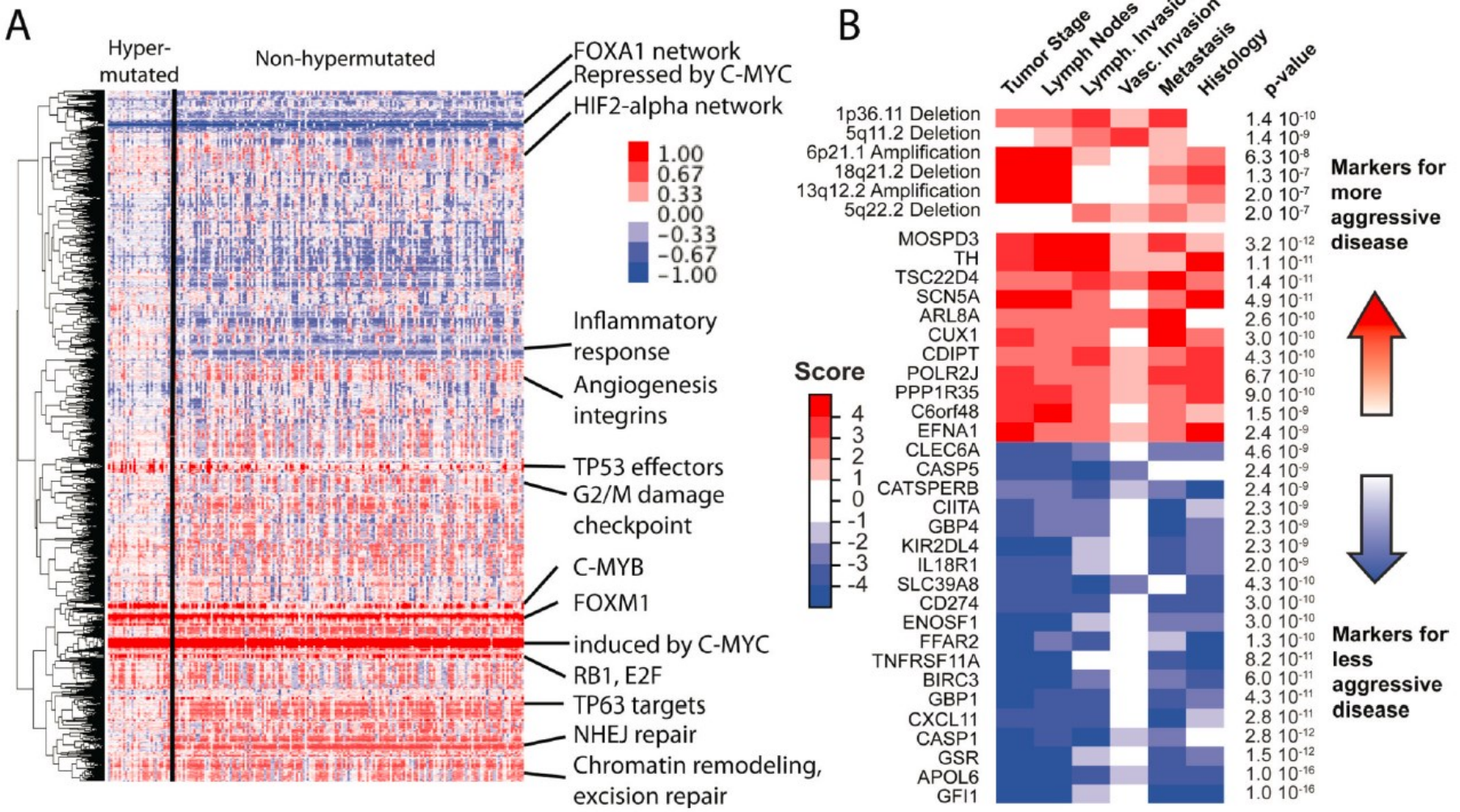

Combined RTK/RAS/PI-3-K pathway
125 altered cases (out of 165)
IGF2
IRS2
PIK3CA
PIK3R1
PTEN
ERBB2
ERBB3
NRAS
KRAS
BRAF
IGF2 / PI-3-K pathway
83 altered cases (out of 165)
IGF2
IRS2
PIK3CA
PIK3R1
PTEN
RTK / RAS pathway
98 altered cases (out of 165)
ERBB2
ERBB3
NRAS
KRAS
BRAF
ERBB2 amplifications and mutations
mutation
L755S
V842I
V777L, V842I
V842I
V842I
log2 mRNA expression
5
4
3
2
1
0
Hetloss
Diploid
Gain
Amp
Copy–number alterations
TGFβ pathway
44 altered cases (out of 165)
TGFBR1
TGFBR2
ACVR1B
ACVR2A
SMAD2
SMAD3
SMAD4
OncoPrint Legend
Amplification
Homozygous deletion
up-regulation
down-regulation
Somatic mutation
WNT pathway
152 altered cases (out of 165)
APC
CTNNB1
DKK1
DKK2
DKK3
DKK4
LRP5
FZD10
FAM123B
AXIN2
TCF7L2
FBXW7
ARID1A
SOX9
TP53 pathway
106 altered cases (out of 165)
ATM
TP53

Figure 5


A
Hyper-mutated
Non-hypermutated
FOXA1 network
Repressed by C-MYC
HIF2-alpha network
1.00
0.67
0.33
0.00
-0.33
-0.67
-1.00
Inflammatory response
Angiogenesis integrins
TP53 effectors
G2/M damage checkpoint
C-MYB
FOXM1
induced by C-MYC
RB1, E2F
TP63 targets
NHEJ repair
Chromatin remodeling, excision repair
B
Tumor Stage
Lymph Nodes
Lymph. Invasion
Vasc. Invasion
Metastasis
Histology
p-value
1p36.11 Deletion 1.4 10-10
5q11.2 Deletion 1.4 10-9
6p21.1 Amplification 6.3 10-8
18q21.2 Deletion 1.3 10-7
13q12.2 Amplification 2.0 10-7
5q22.2 Deletion 2.0 10-7
MOSPD3 3.2 10-12
TH 1.1 10-11
TSC22D4 1.4 10-11
SCN5A 4.9 10-11
ARL8A 2.6 10-10
CUX1 3.0 10-10
CDIPT 4.3 10-10
POLR2J 6.7 10-10
PPP1R35 9.0 10-10
C6orf48 1.5 10-9
EFNA1 2.4 10-9
CLEC6A 4.6 10-9
CASP5 2.4 10-9
CATSPERB 2.4 10-9
CIITA 2.3 10-9
GBP4 2.3 10-9
KIR2DL4 2.3 10-9
IL18R1 2.0 10-9
SLC39A8 4.3 10-10
CD274 3.0 10-10
ENOSF1 3.0 10-10
FFAR2 1.3 10-10
TNFRSF11A 8.2 10-11
BIRC3 6.0 10-11
GBP1 4.3 10-11
CXCL11 2.8 10-11
CASP1 2.8 10-12
GSR 1.5 10-12
APOL6 1.0 10-16
GFI1 1.0 10-16
Score
4
3
2
1
0
-1
-2
-3
-4
Markers for more aggressive disease
Markers for less aggressive disease

# Understanding cancer cell function or cell function in general

… is contingent upon our ability to acquire a better appreciation of the complexity of molecular interactions.

Technological advances allow to monitor cell function from a holistic perspective.
Improved computational capabilities

→ Cancer System Biology

# Advances in ordering the system

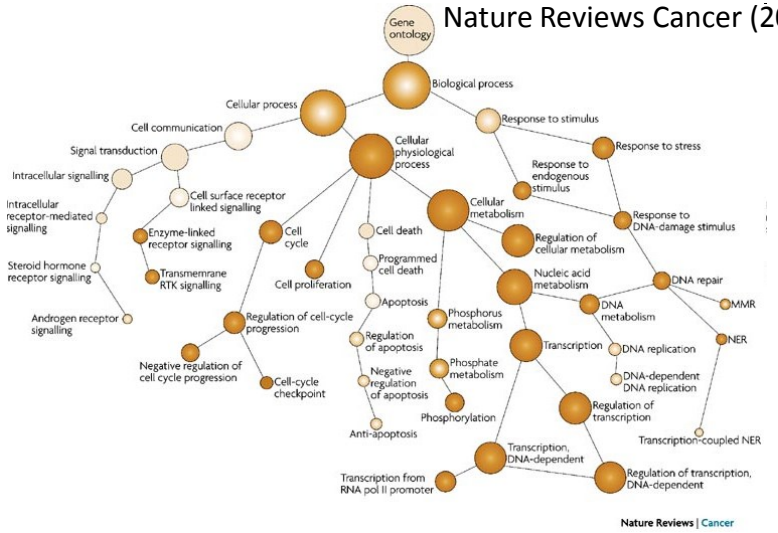
Gene Ontology
- Biological Process
- Cellular Component
- Molecular Function

Open Biomedical Ontology (OBO)

Transcription factors

Functional gene sets

Nature Reviews Cancer (2007) 7, 23-34




TRANSFAC® Public 版は、啓蒙活動用で、研究用ではありません。有償版 TRANSFAC® Professional には、多くの機能が搭載されており、PWMsに代表されるデータ量は、Public 版の約3倍です。最新のデータを利用できるTRANSFAC® Professional を是非ご検討ください。
両者の比較は、こちら。

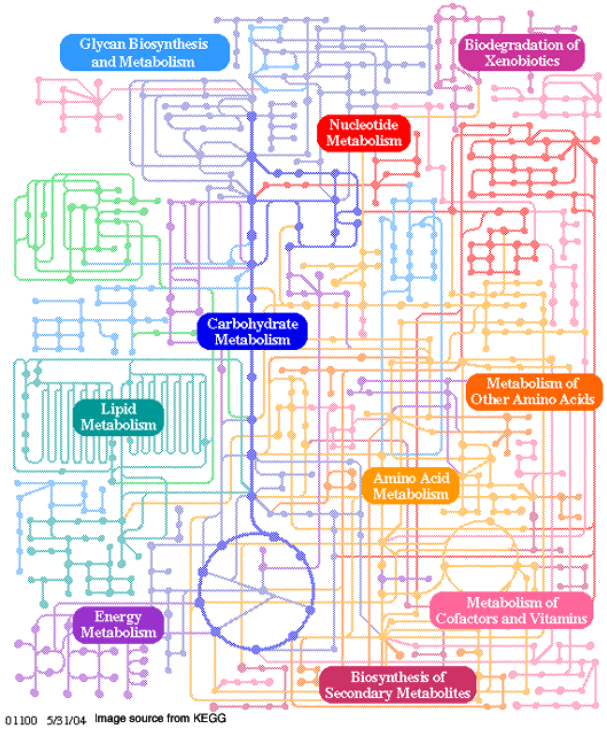
MSigDB
Molecular Signatures Database

THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY

# Strategies to identify cellular interactions

- Bottom-up approach
- Top – down approach
- Complementary merging of both

# Bottom-Up




A bottom-up approach is the piecing together of systems to give rise to grander systems.

It makes the original systems sub-systems of the emergent system.

In a bottom-up approach the individual base elements of the system are first specified in great detail.

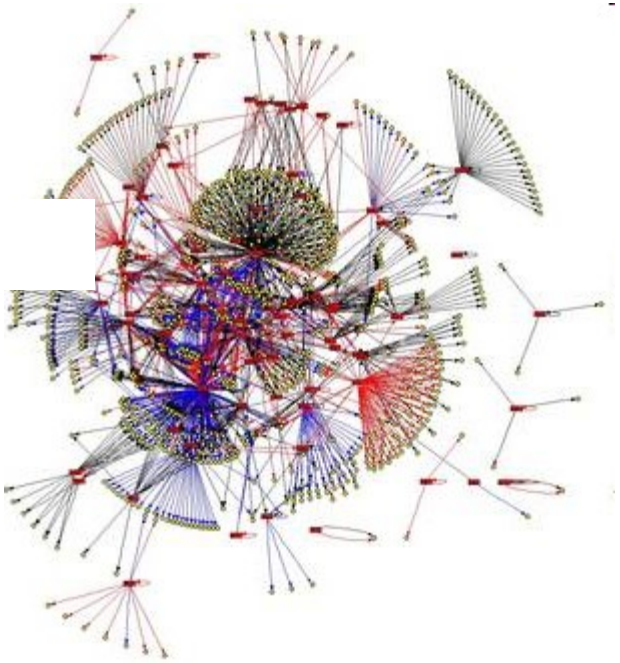
These elements are then linked together to form larger subsystems, which then in turn are linked, sometimes in many levels, until a complete top-level system is formed.

# Top-down

A top-down approach breaks a system down to gain insight into its compositional sub-systems.

In a top-down approach an overview of the system is formulated, specifying but not detailing any first-level subsystems.

Each subsystem is then refined in yet greater detail, sometimes in many additional subsystem levels, until the entire specification is reduced to base elements.



# Cancer Systems Biology: Bottom - Up

Reductionist in nature – focus is on known components, based on known and experimentally validated structures.

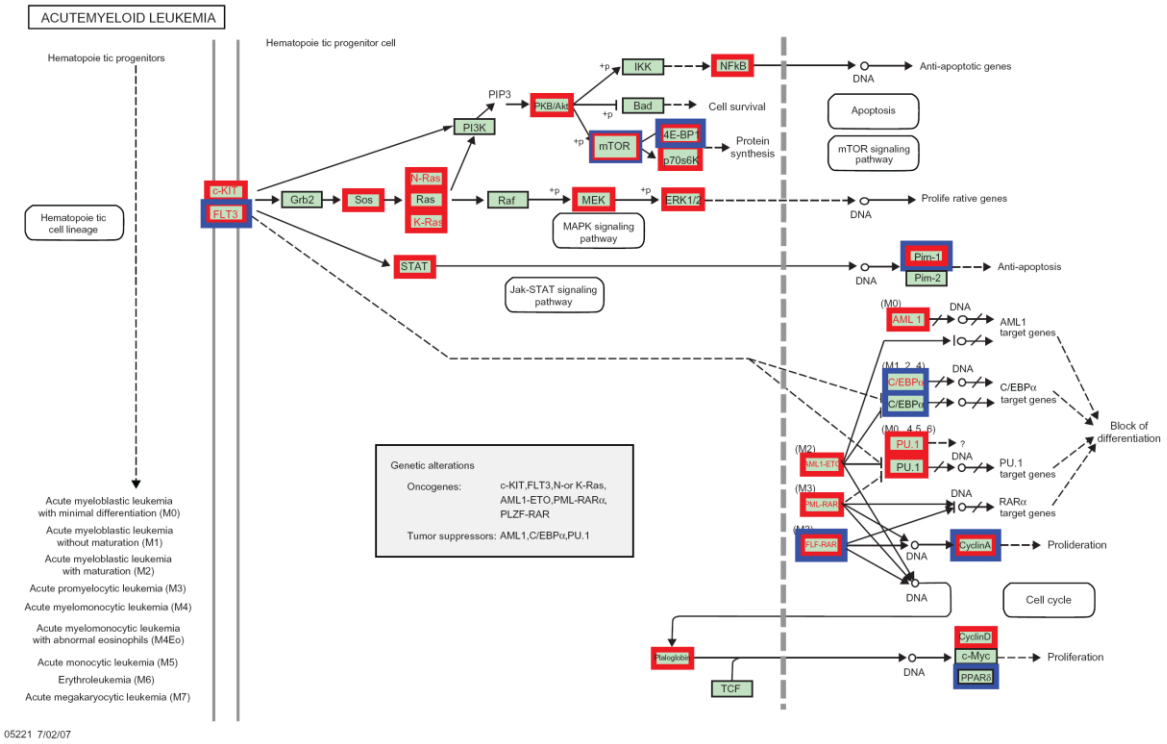
Map oncogenes / suppressor genes with known interacting genes (regulatory, developmental) previously characterized in cell lines or experimental organisms.

(Wnt, notch, hedgehog pathways) → identified as sets of interacting genes involved in Drosophila development.

Interactive networks are inferred from prior knowledge about the molecular interactions of their components.

# Integrating the AML signature into pathways

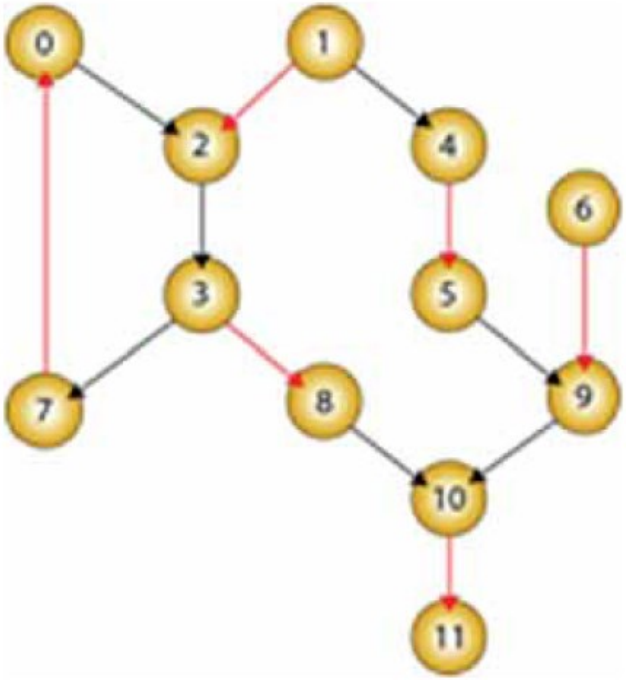
Hummel et al. *BBI* 2008:2 329-341




KEGG pathway ‘acute myeloid leukemia’ (hsa05221). Red boxes mark involved genes that correlate significantly with at least one of the signature genes. Blue boxes mark genes that show a significant partial correlation (in the gene association network) to at least one of the signature genes.

# Correlation versus Partial correlation

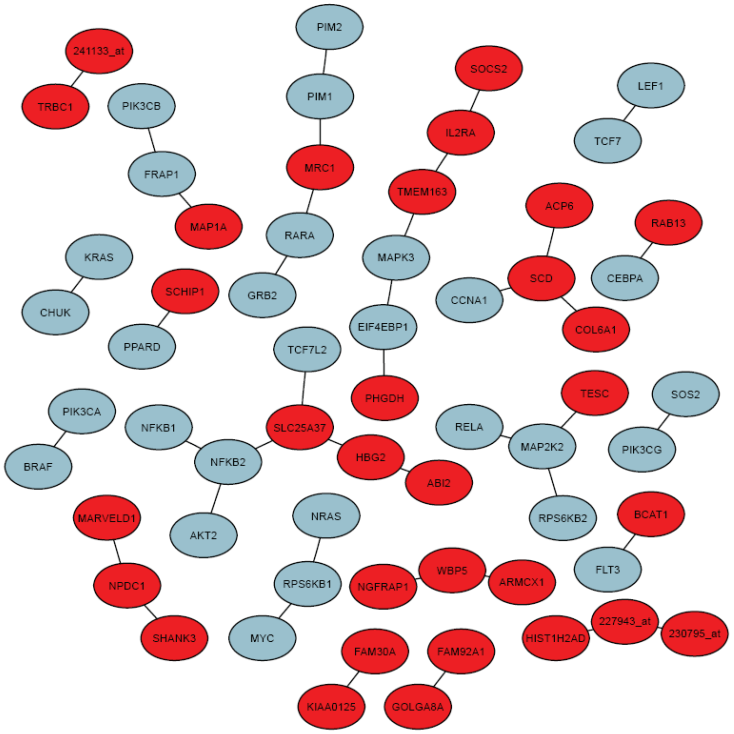
How to detect and quantify direct interaction?




Partial correlation is the basic statistical principle for the construction of statistically motivated network inference algorithms

# Integrating the AML signature into pathways

Hummel et al. *BBI* 2008:2 329-341

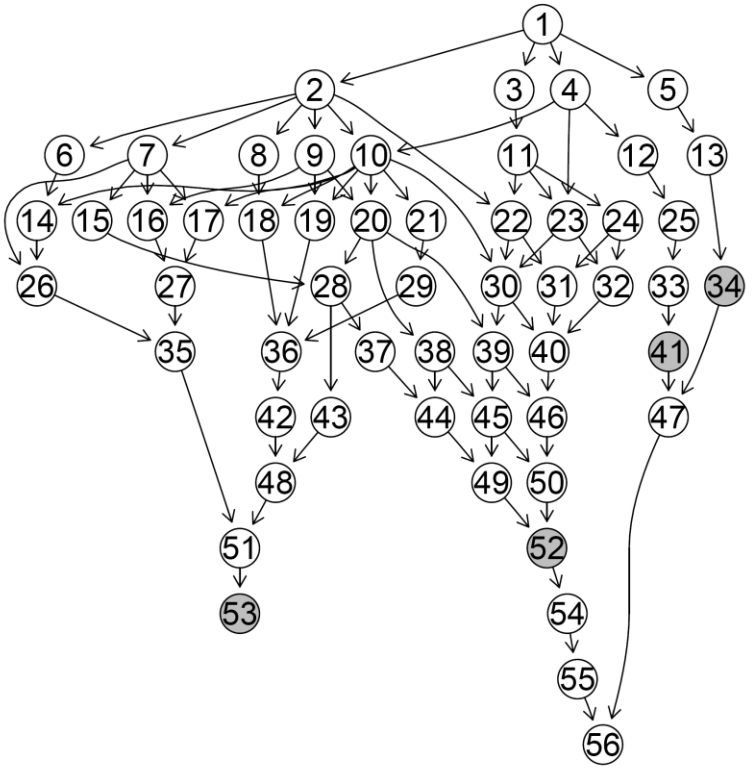



A gene association network between signature and AML pathway genes. Blue nodes correspond to genes in the pathway ‘acute myeloid leukemia’ (hsa05221), and red nodes correspond to genes in the prognostic signature.

# Functional relevance of mutations in AML

238 AML patients with normal karyotype, Recently, mutations in two genes, namely the fms-like tyrosine kinase 3 (FLT3) and nucleophosmin 1 (NPM1) were shown to be relevant for prognosis for AML patients with normal karyotype.

Results on the pure FLT3 effect. The pure effect of one mutation is defined by adjusting the analysis for the effect of the second mutation.




Node 53 (GO:0006427, histidyl-tRNA aminoacylation)
Node 56 (GO:0006994, sterol depletion response, SREBP target gene transcriptional activtion)

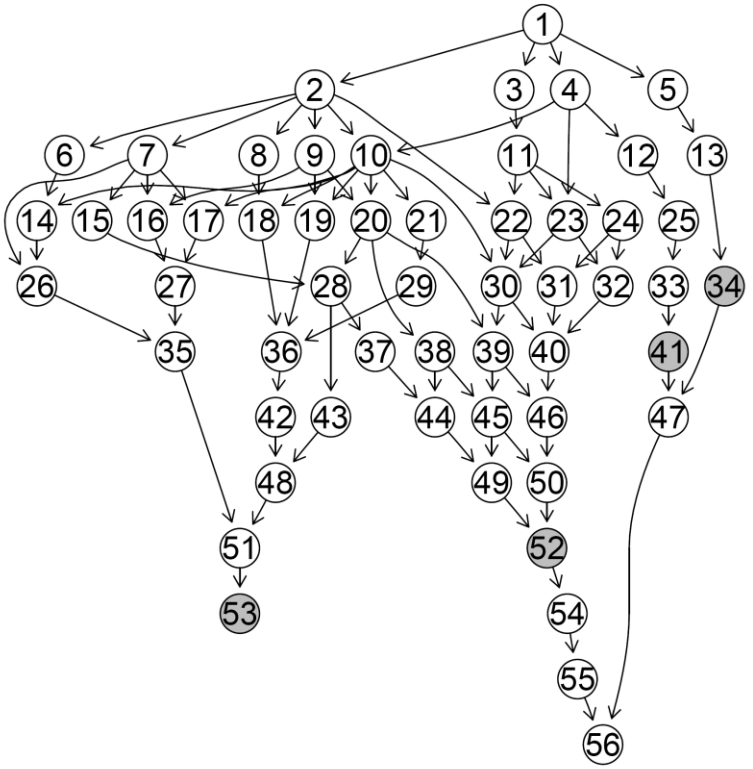
Goeman, Mansmann (2008) Bioninformatics, 537–544

# Mapping the relevance of a hypothesis to a structure of instances

A statistical model is used to attack a complex question like:

*Is there clear evidence for a pure systematic mutation effect on the transcription activities within a certain set of genes.*

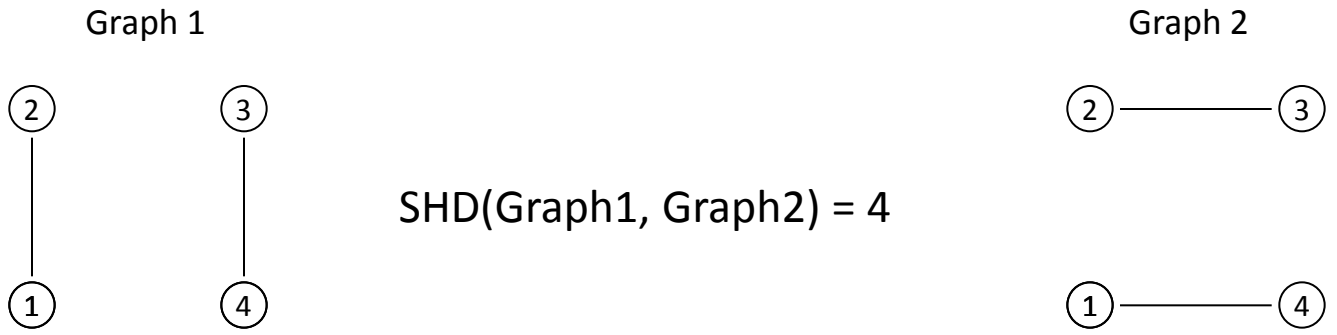
Available evidence can be located within a larger structure.




Hummel, Meister, Mansmann (2008) Bioinformatics, 24:78-85

# Differential structure between networks on the same set of nodes
# Structural Hamming Distance - SHD

The Hamming distance between two strings of equal length is the number of positions at which the corresponding symbols are different. Put another way, it measures the minimum number of substitutions required to change one string into the other, or the number of errors that transformed one string into the other.



# Meta-Analyses of Cancer Pathways

Schmitberger, Lennert, Mansmann *BBI* 2011:5 13–39

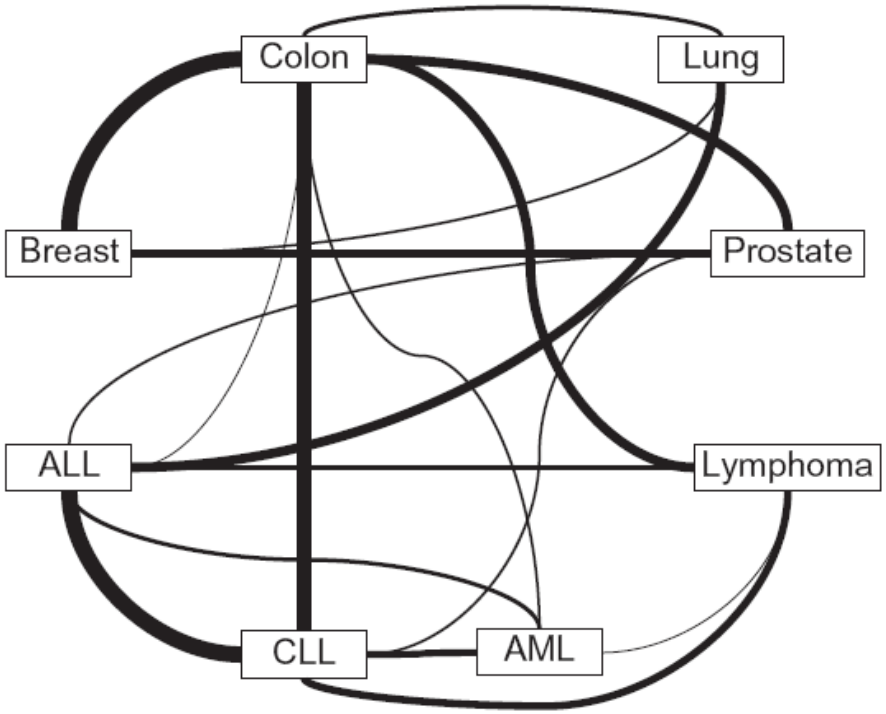
Structural difference in terms of SHD within a pathway between two cancer types

| | | Solid entities | | | | | |
|---|---|---|---|---|---|---|---|
| | | BRE-COL | BRE-LUN | BRE-PRO | COL-LUN | COL-PRO | LUN-PRO |
| CC | 04110 | 577 (0.302) | 642 (<0.002) | 627 (<0.002) | 435 (<0.002) | 420 (<0.002) | 477 (<0.002) |
| p53 | 04115 | 319 (0.994) | 373 (<0.002) | 350 (0.046) | 234 (<0.002) | 223 (0.016) | 279 (<0.002) |
| Apop | 04210 | 475 (0.054) | 540 (<0.002) | 475 (<0.002) | 335 (<0.002) | 298 (0.02) | 375 (<0.002) |
| Wnt | 04310 | 834 (<0.002) | 919 (<0.002) | 920 (<0.002) | 617 (<0.002) | 568 (<0.002) | 701 (<0.002) |
| ECM | 04512 | 435 (0.993) | 477 (<0.002) | 453 (0.614) | 314 (0.686) | 310 (0.592) | 354 (0.802) |
| Disease specific | 05210 | 506 (0.524) | 569 (<0.002) | 549 (<0.002) | 393 (<0.002) | 351 (<0.002) | 428 (<0.002) |
| | 05215 | 527 (0.998) | 607 (<0.002) | 596 (0.002) | 416 (<0.002) | 387 (0.046) | 457 (0.002) |
| | 05221 | 279 (0.993) | 322 (0.044) | 312 (0.84) | 235 (0.136) | 215 (0.374) | 272 (0.008) |
| | 05223 | 314 (0.644) | 336 (<0.002) | 313 (0.162) | 222 (0.038) | 213 (0.028) | 241 (0.022) |
| DNA repair | 04150 | 234 (0.418) | 242 (<0.002) | 268 (<0.002) | 172 (<0.002) | 170 (<0.002) | 180 (0.004) |
| | 03410 | 89 (0.744) | 107 (<0.002) | 107 (<0.002) | 68 (<0.002) | 68 (<0.002) | 76 (<0.002) |
| | 03420 | 145 (0.35) | 163 (<0.002) | 146 (<0.002) | 114 (<0.002) | 109 (<0.002) | 107 (<0.002) |
| | 03430 | 60 (0.993) | 73 (0.054) | 61 (0.918) | 63 (0.006) | 45 (0.366) | 62 (0.188) |
| | | **Hemic entities** | | | | | |
| | | ALL-AML | ALL-CLL | ALL-LYM | AML-CLL | AML-LYM | CLL-LYM |
| CC | 04110 | 581 (<0.002) | 542 (<0.002) | 570 (<0.002) | 447 (<0.002) | 453 (<0.002) | 436 (<0.002) |
| p53 | 04115 | 300 (<0.002) | 283 (0.142) | 266 (0.066) | 241 (<0.002) | 240 (<0.002) | 193 (0.128) |
| Apop | 04210 | 438 (<0.002) | 373 (0.118) | 393 (<0.002) | 345 (<0.002) | 383 (<0.002) | 296 (0.004) |
| Wnt | 04310 | 822 (<0.002) | 697 (0.05) | 761 (<0.002) | 585 (<0.002) | 675 (<0.002) | 526 (<0.002) |
| ECM | 04512 | 468 (0.004) | 435 (0.993) | 443 (0.544) | 373 (0.658) | 381 (<0.002) | 338 (0.52) |
| Disease specific | 05210 | 503 (<0.002) | 424 (0.644) | 472 (<0.002) | 389 (<0.002) | 413 (<0.002) | 336 (<0.002) |
| | 05215 | 516 (<0.002) | 470 (0.994) | 506 (0.006) | 410 (0.23) | 440 (<0.002) | 344 (0.562) |
| | 05221 | 265 (0.002) | 242 (0.994) | 261 (0.008) | 233 (0.386) | 248 (<0.002) | 195 (0.232) |
| | 05223 | 307 (<0.002) | 255 (0.868) | 258 (0.02) | 234 (0.014) | 249 (<0.002) | 193 (0.038) |
| DNA repair | 04150 | 241 (<0.002) | 223 (0.06) | 225 (<0.002) | 192 (<0.002) | 200 (<0.002) | 162 (0.002) |
| | 03410 | 104 (<0.002) | 85 (0.012) | 91 (<0.002) | 83 (<0.002) | 91 (<0.002) | 76 (<0.002) |
| | 03420 | 132 (<0.002) | 123 (0.034) | 131 (<0.002) | 125 (<0.002) | 133 (<0.002) | 100 (<0.002) |
| | 03430 | 62 (0.002) | 63 (0.758) | 63 (<0.002) | 61 (0.062) | 59 (0.042) | 48 (0.628) |

# Meta-Analyses of Cancer Pathways

Schmitberger, Lennert, Mansmann *BBI* 2011:5 13–39

Structural difference in terms of SHD within a pathway between two cancer types



# Cancer Systems Biology: Top - Down

- Computationally based
- Not dependent of a priory knowledge
- Infers network relationships
- Establishes a network of inferred interactions between all cellular components whether or not these interactions are currently understood on the molecular level.

# Computational advances to infer *network* or interaction between genes

Networks inferred from gene expression data

based on statistical reasoning: glasso, PC, GeneNet

Bayesian networks

based on bioinformatics techniques: ARACNE, C3NET

Networks inferred from gene silencing data (interventional data)

Nested effects models

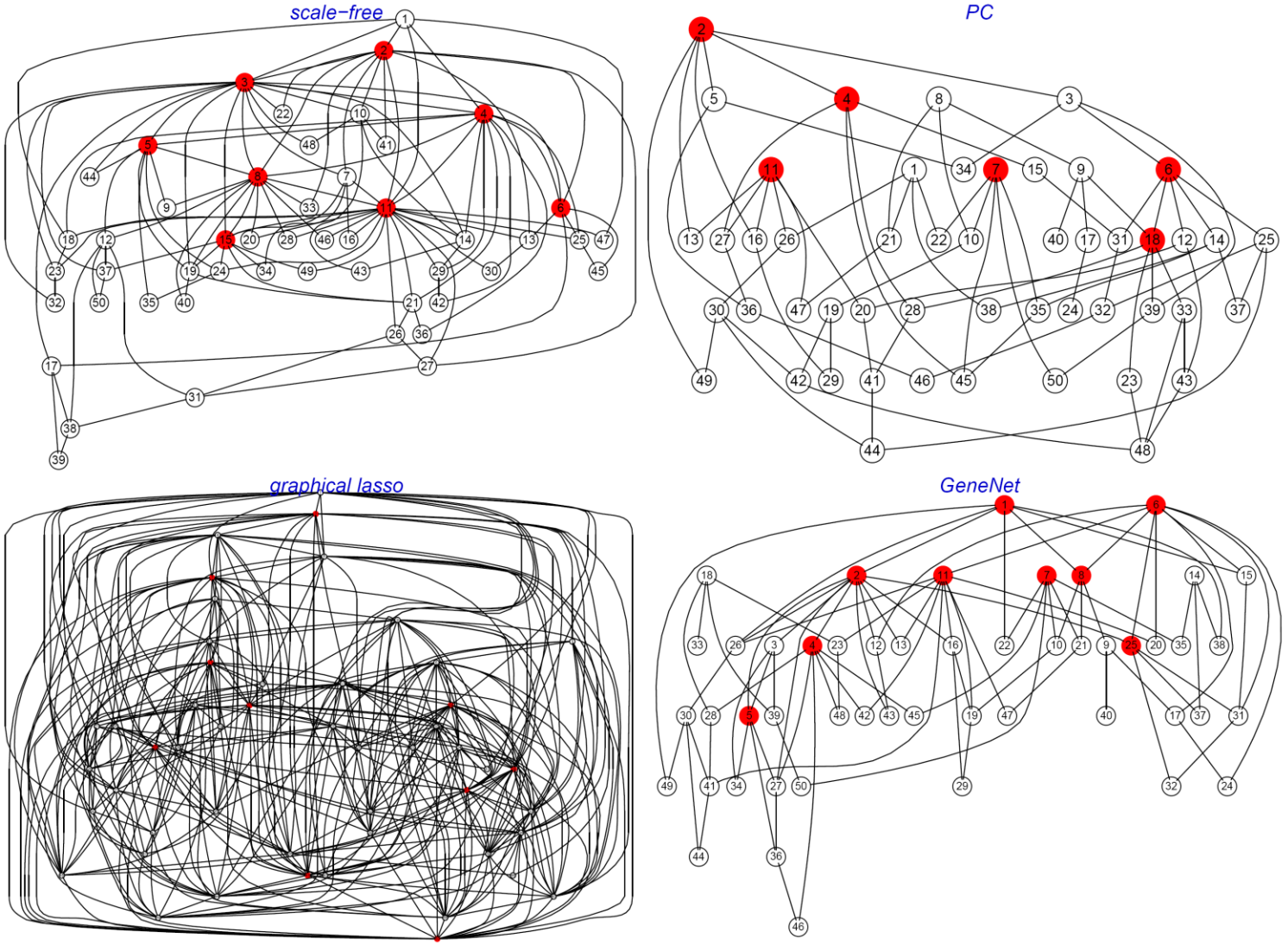
Networks inferred from time series

Empirical Bayes Dynamic Bayesian Networks

Differential interaction between networks

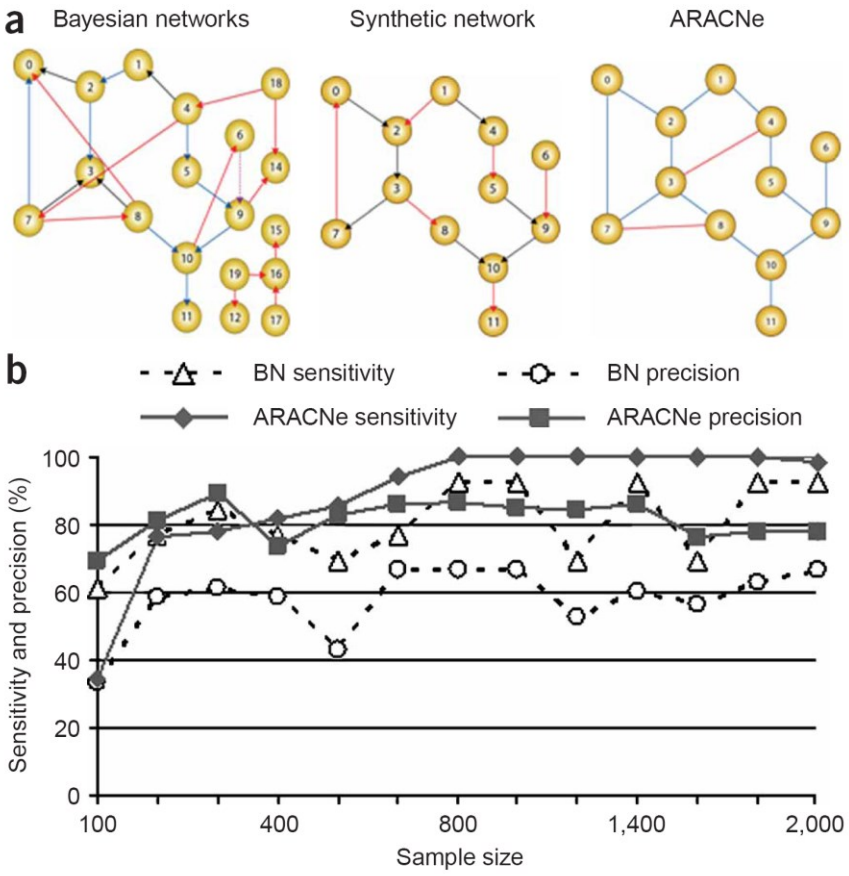
Relevant interactions can be identified by comparison

# Validation of network algorithms



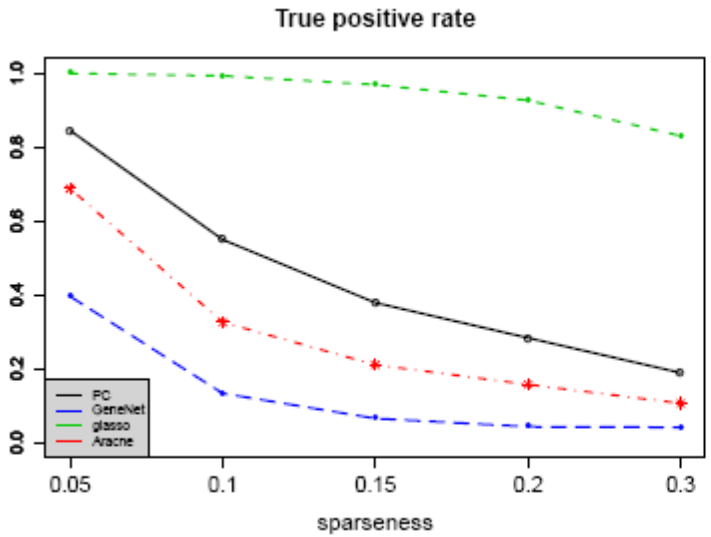
# Validating ARACNE

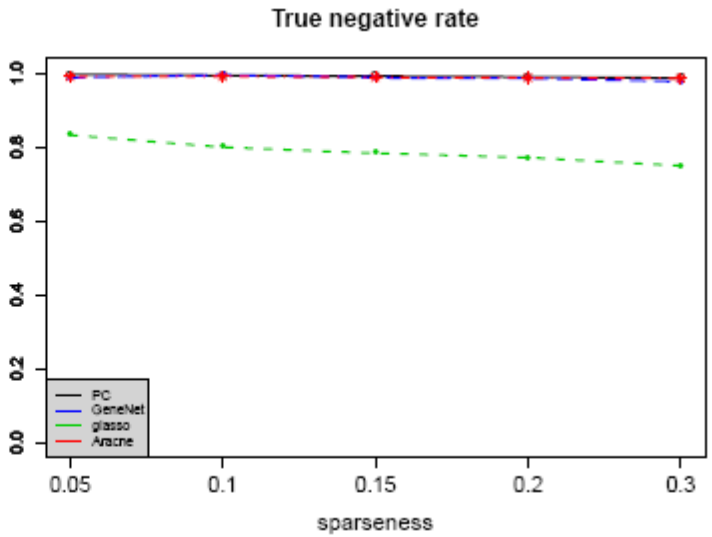
Basso et al. (2005) Nature Genetics, 37: 382-390

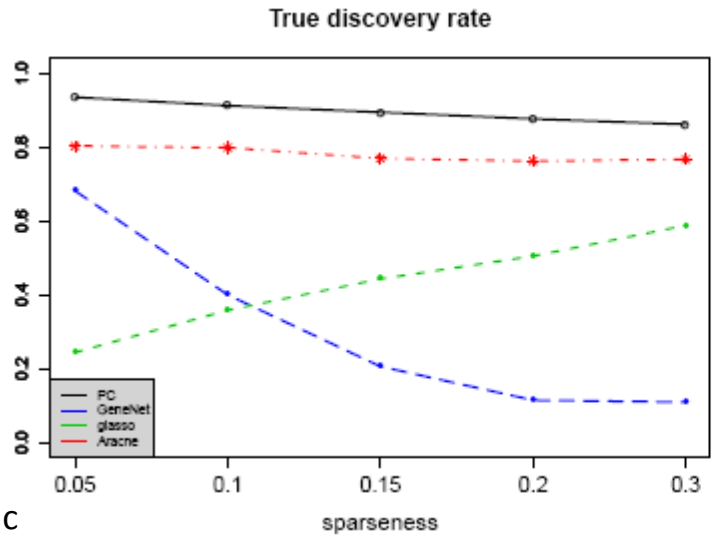


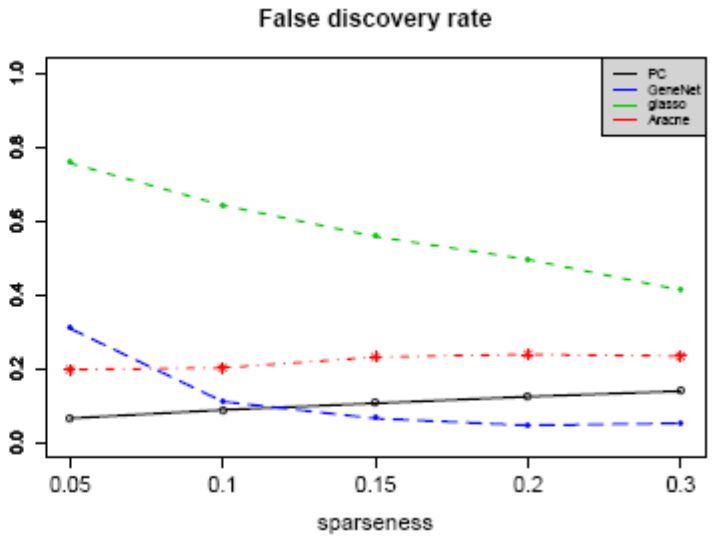
# Validation of network algorithms

50 nodes, Sample Size 200



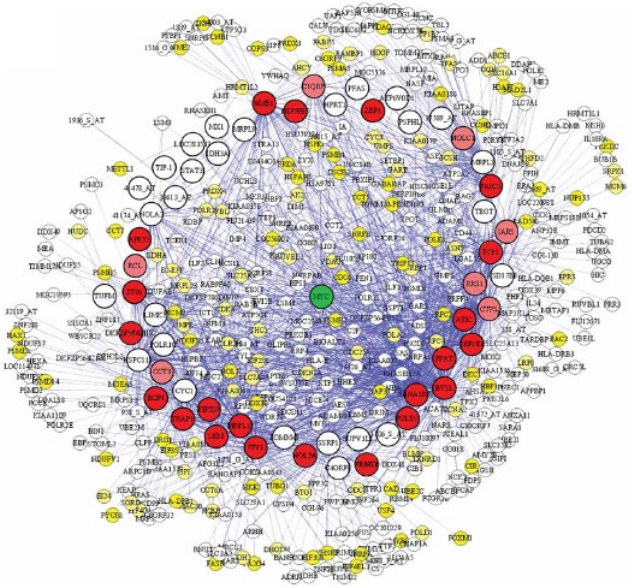



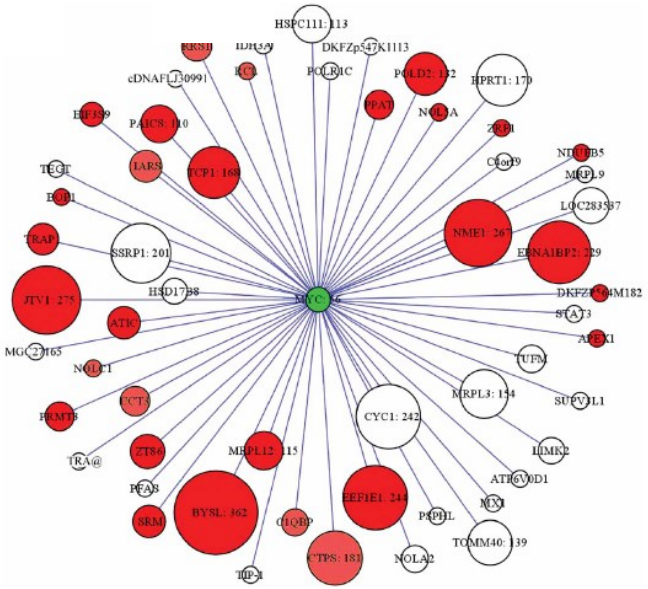



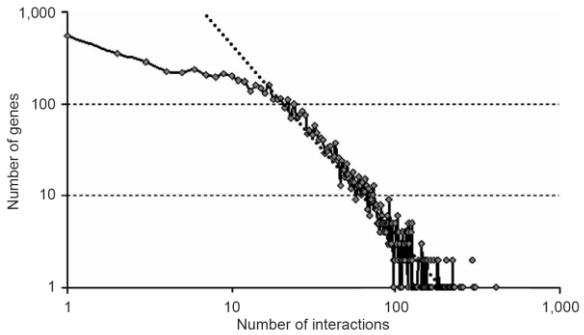



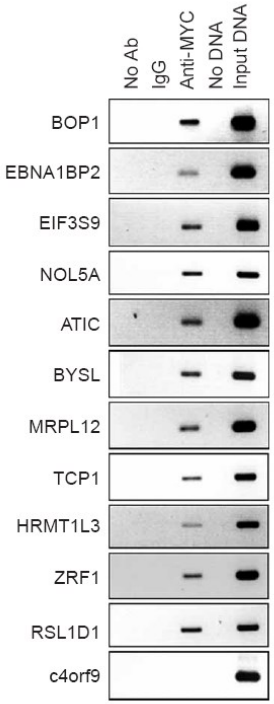
# Exploring interaction within a system

Basso et al. (2005) Nature Genetics, 37: 382-390



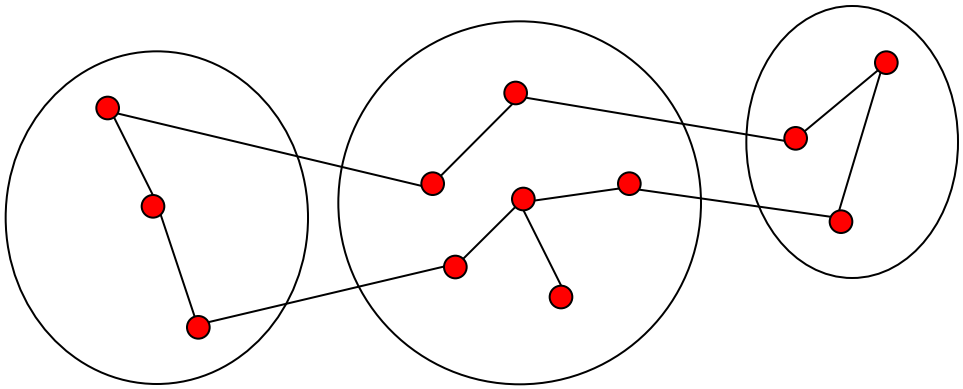


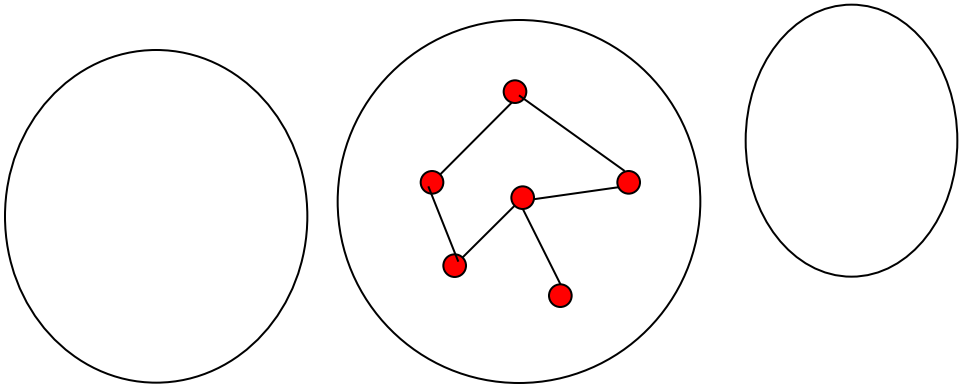



# Confounding

How does an artificial restriction to a bounded region within a wider system influences the observed interactions?



extended analysis

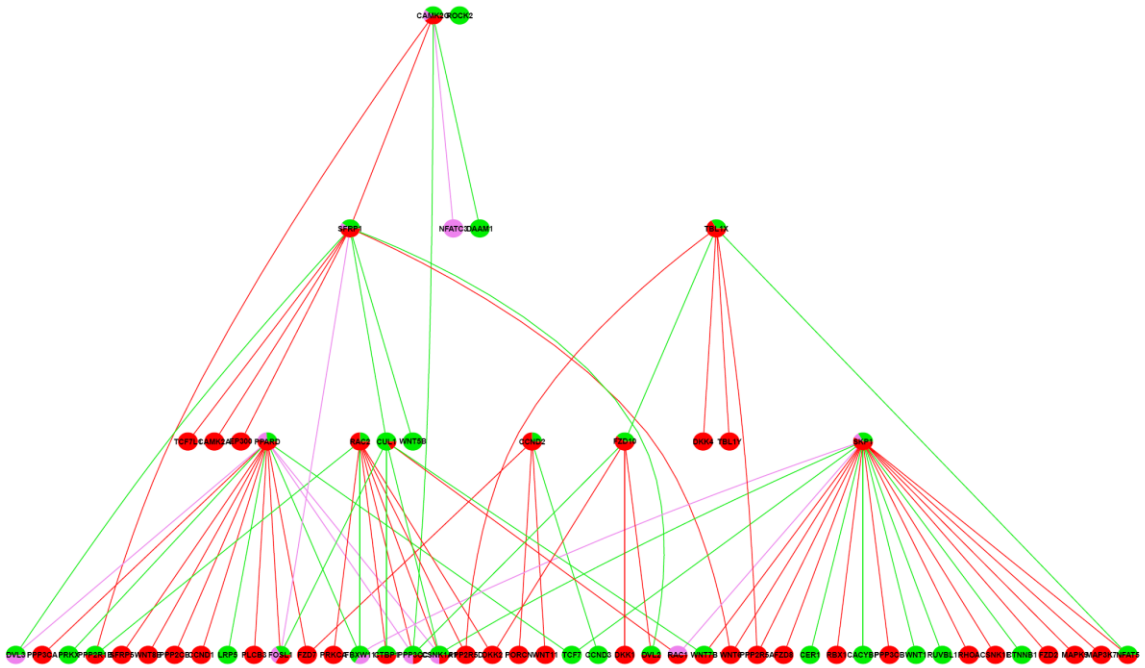


restricted analysis

# Confounding

Effect of MYC translocation of the network structure within specific pathways

wnt



Comparing differential interaction structure between the restricted and full inference.

**Green edges**: Differential interaction present for the inference restricted to the wnt pathway.

**Red edges**: Differential interaction which is observed within the restricted and the extended setting simultaneously.

**Purple edges**: Differential interaction only present for the extended inference.



Mansmann, Jurinovic (2011) Brief Bioinform. 12:230-44

## Bottom-Up

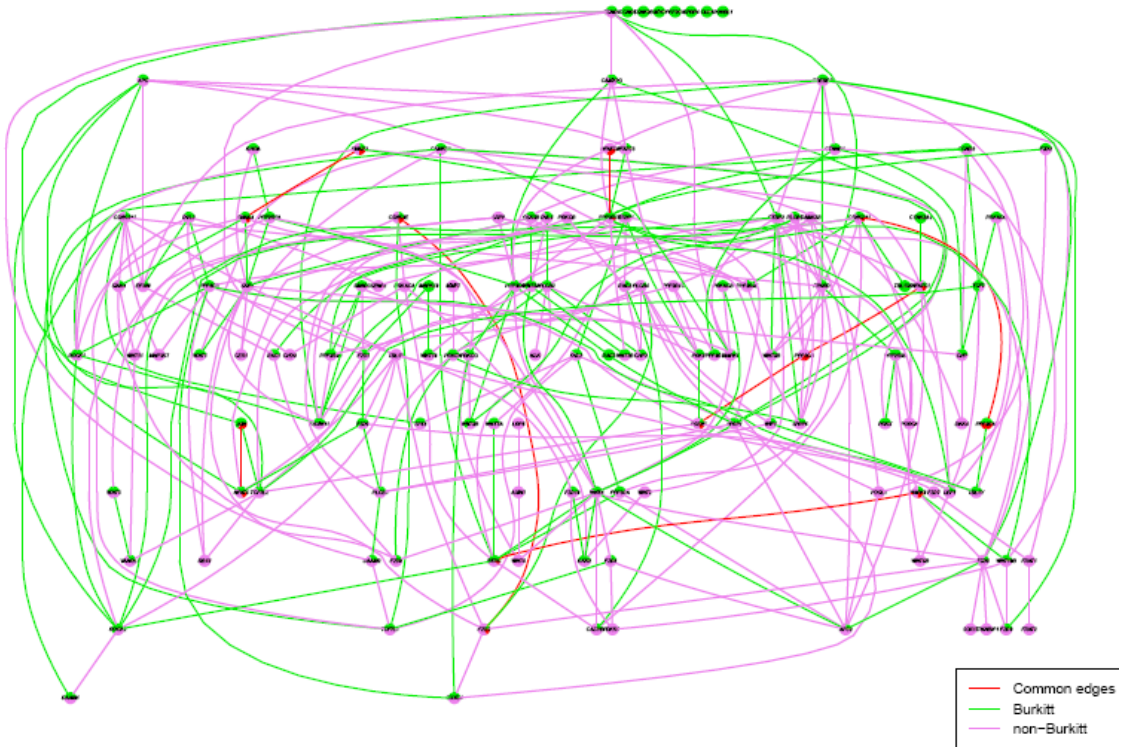
- intuitively more desirable;
- Grounded on experimental validation;
- Limitation in scope;
- The gap will widen;
- Is affected by confounding

## Top - Down

- Primary method to establish new molecular networks;
- Networks will probably stimulate new directions in laboratory studies to validate molecular interactions;
- Avoids confounding

# Complementary Merging

Complementary merging of the top-down and bottom-up approaches promises to be productive in helping unravel the functional complexity underlying cancer cell functions.



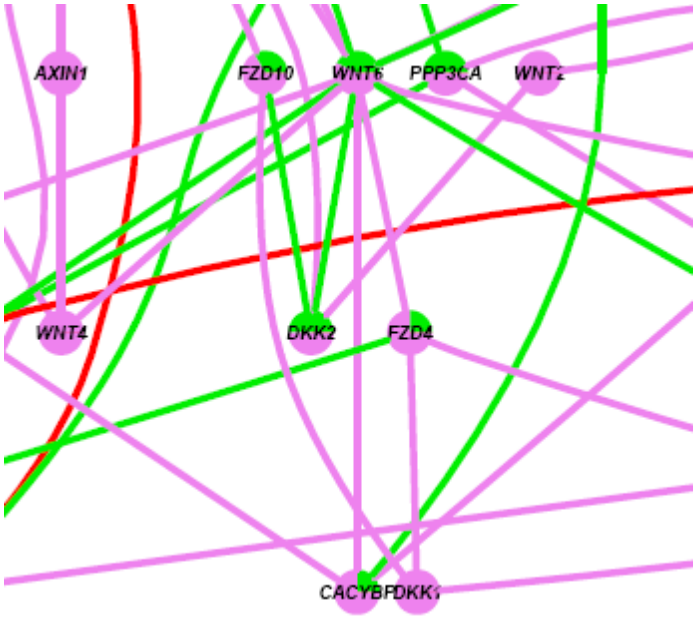

Wnt Pathway

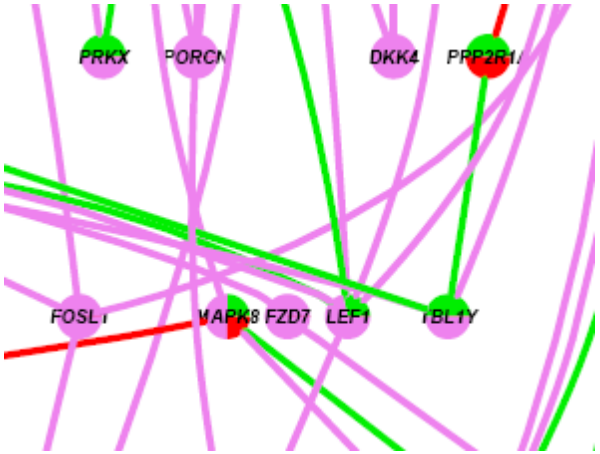


Mansmann, Jurinovic (2011) Brief Bioinform. 12:230-44

# Complementary Merging

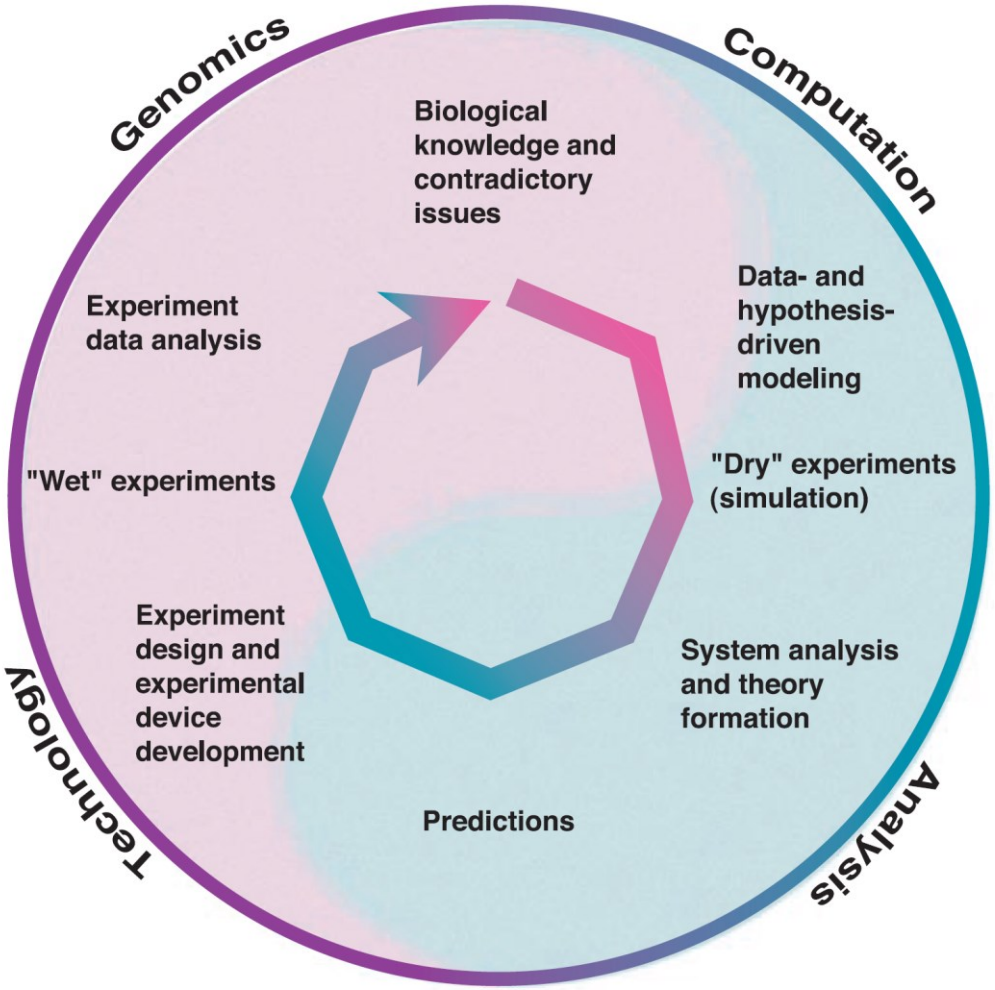
Complementary merging of the top-down and bottom-up approaches promises to be productive in helping unravel the functional complexity underlying cancer cell functions.




Mansmann, Jurinovic (2011) Brief Bioinform. 12:230-44

# HaematoSYS working plan



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 MARCH 8, 2012 VOL. 366 NO. 10

# Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing

Marco Gerlinger, M.D., Andrew J. Rowan, B.Sc., Stuart Horswell, M.Math., James Larkin, M.D., Ph.D., David Endesfelder, Dip.Math., Eva Gronroos, Ph.D., Pierre Martinez, Ph.D., Nicholas Matthews, B.Sc., Aengus Stewart, M.Sc., Patrick Tarpey, Ph.D., Ignacio Varela, Ph.D., Benjamin Phillimore, B.Sc., Sharmin Begum, M.Sc., Neil Q. McDonald, Ph.D., Adam Butler, B.Sc., David Jones, M.Sc., Keiran Raine, M.Sc., Calli Latimer, B.Sc., Claudio R. Santos, Ph.D., Mahrokh Nohadani, H.N.C., Aron C. Eklund, Ph.D., Bradley Spencer-Dene, Ph.D., Graham Clark, B.Sc., Lisa Pickering, M.D., Ph.D., Gordon Stamp, M.D., Martin Gore, M.D., Ph.D., Zoltan Szallasi, M.D., Julian Downward, Ph.D., P. Andrew Futreal, Ph.D., and Charles Swanton, M.D., Ph.D.

# Everolimus

Everolimus (RAD-001) is the 40-O-(2-hydroxyethyl) derivative of sirolimus and works similarly to sirolimus as an inhibitor of mammalian target of rapamycin (mTOR).

It is currently used as an immunosuppressant to prevent rejection of organ transplants and treatment of renal cell cancer and other tumours. Much research has also been conducted on everolimus and other mTOR inhibitors for use in a number of cancers.

# mTOR

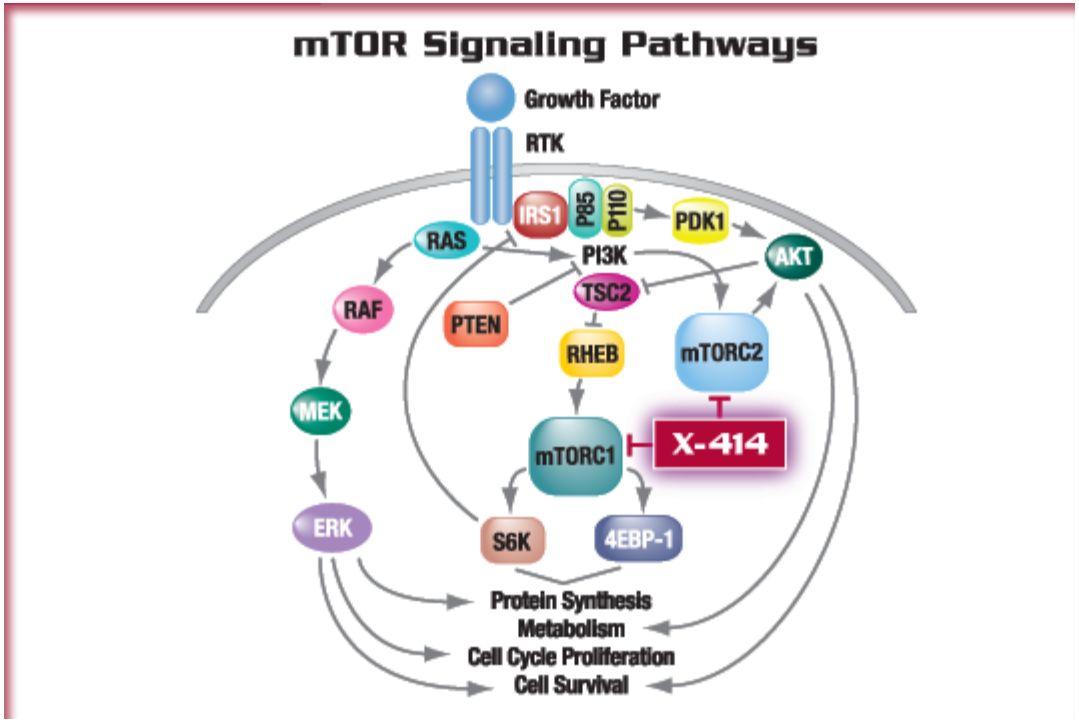
The mammalian target of rapamycin (mTOR), also known as mechanistic target of rapamycin or FK506-binding protein 12-rapamycin-associated protein 1 (FRAP1), is a protein that in humans is encoded by the MTOR gene.

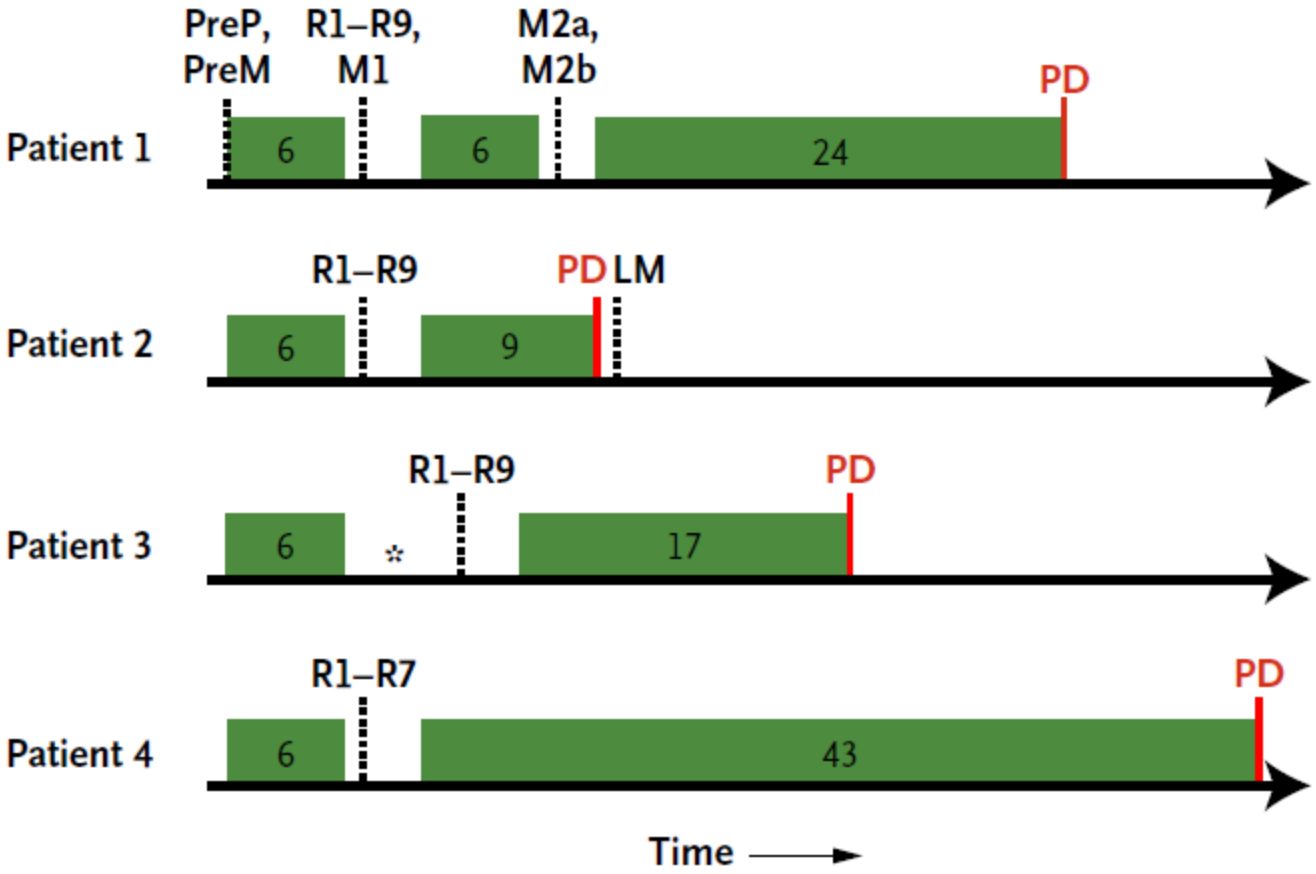
Rapamycin was discovered in a soil sample from Easter Island, known locally as Rapa Nui, in the 1970s. The bacterium Streptomyces hygroscopicus, isolated from that sample, produces an antifungal that researchers named rapamycin after the island.

Rapamycin arrests fungal activity at the G1 phase of the cell cycle. In mammals, it suppresses the immune system by blocking the G1 to S phase transition in T-lymphocytes. Thus, it is used as an immunosuppressant following organ transplantation.

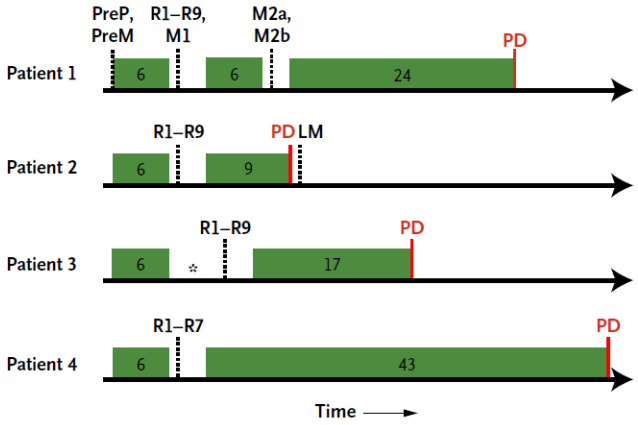
Saccharomyces cerevisiae can develop resistance to rapamycin, when one of three genes is knocked out. Two of the genes were called targets of rapamycin (TOR). The third gene was already known as Fpr1. Fpr1 is the ortholog of FKBP12 binding protein in the TOR complexes. In 1994, the mammalian target of rapamycin (mTOR) was identified as the rapamycin target in mammals by David M. Sabatini and Solomon H. Snyder (Johns Hopkins University) and Stuart L. Schreiber (Harvard University).

mTOR


mTOR Signaling Pathways
Growth Factor
RTK
IRS1
P85
P110
PDK1
RAS
PI3K
AKT
TSC2
RAF
PTEN
RHEB
mTORC2
MEK
mTORC1
X-414
ERK
S6K
4EBP-1
Protein Synthesis
Metabolism
Cell Cycle Proliferation
Cell Survival

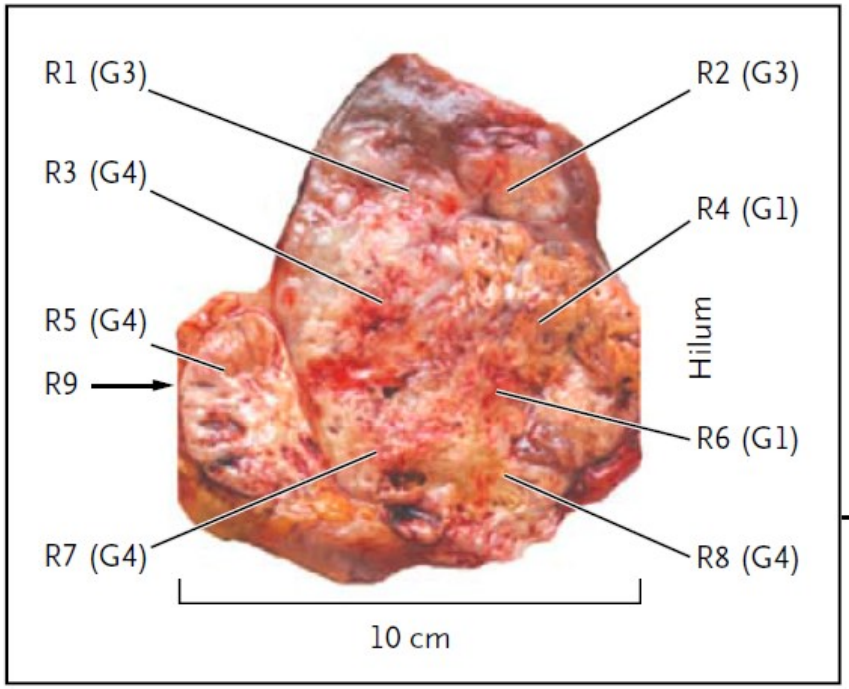

PreP,
PreM
R1–R9,
M1
M2a,
M2b
PD
Patient 1
6
6
24
R1–R9
PD
LM
Patient 2
6
9
R1–R9
PD
Patient 3
6
17
R1–R7
PD
Patient 4
6
43
Time

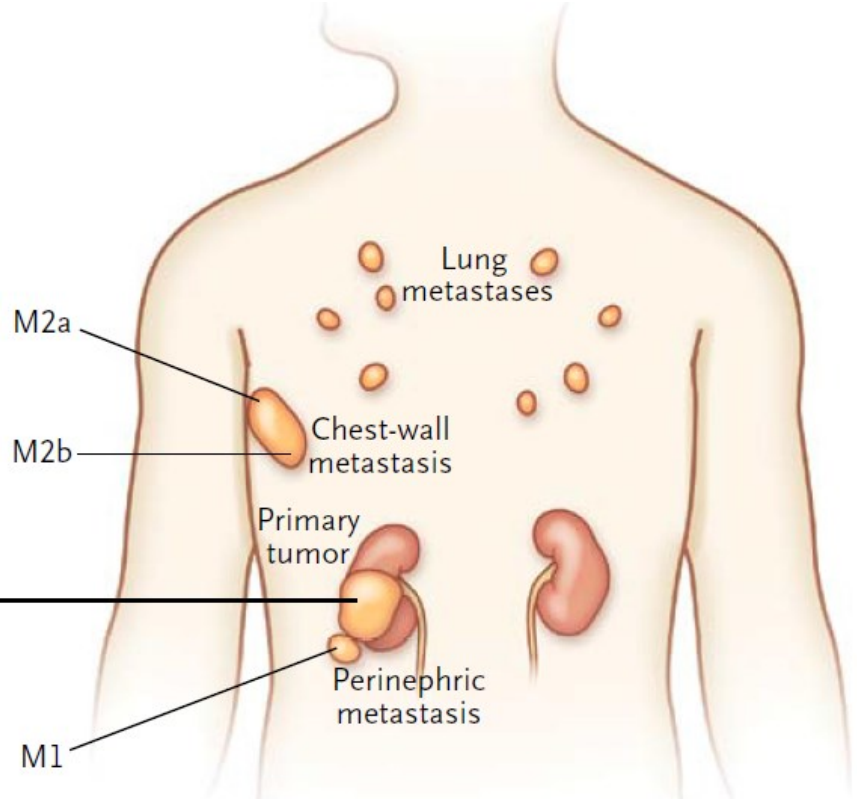
Patient 1 had a clear-cell carcinoma, pulmonary metastases, and a chest-wall metastasis. Sequencing detected a 2-bp deletion in the von Hippel–Lindau tumor-suppressor gene (*VHL*) leading to mutational inactivation, which is characteristic of clear-cell carcinoma. After 6 weeks of everolimus treatment and a 1-week washout period, a nephrectomy was performed. The patient restarted everolimus for 6 weeks and after another 1-week washout period proceeded to surgery of the chest-wall mass (Fig. 1). Computed tomography (CT) did not reveal any change in the dimensions of the primary tumor or chest-wall metastasis during everolimus treatment.



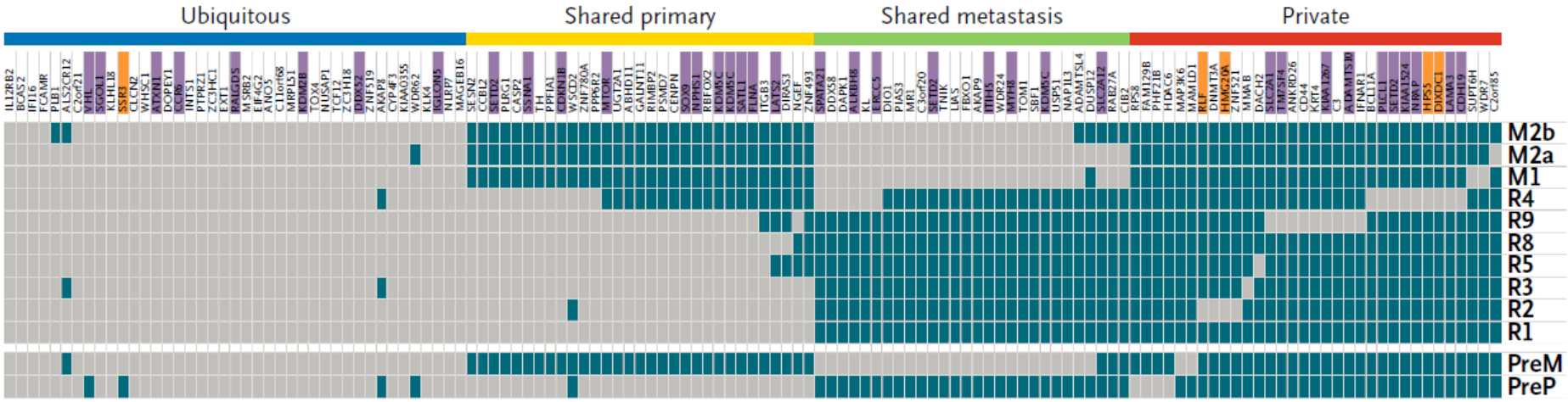

For Patient 1, we performed exon-capture multiregion sequencing on DNA from pretreatment biopsy samples of the primary tumor (PreP) and chest-wall metastasis (PreM), nine primary-tumor regions of the nephrectomy specimen (R1 to R9), a metastasis in the perinephric fat of the nephrectomy specimen (M1), two regions of the excised chest-wall metastasis (M2a and M2b), and germline DNA[19] (Fig. 2A). This sequencing resulted in a me-

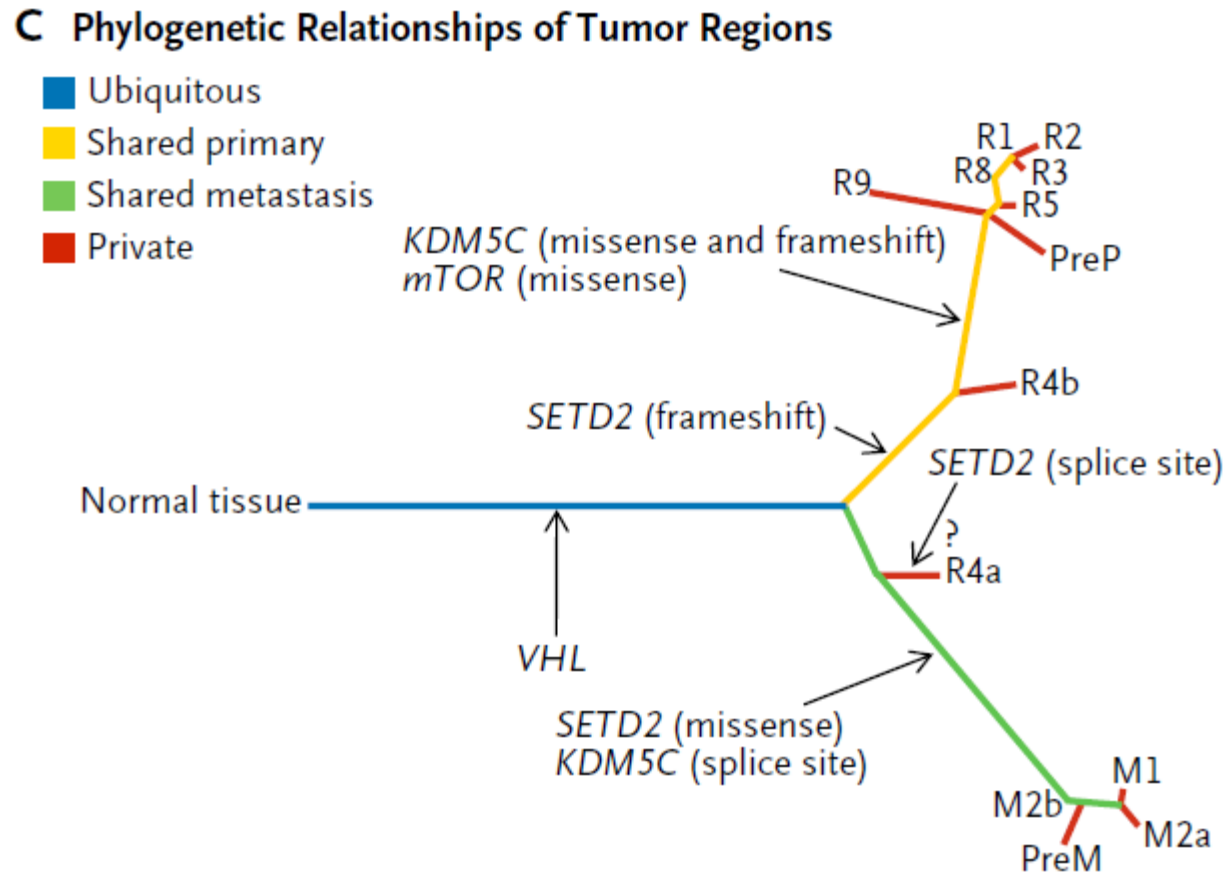
**A** **Biopsy Sites**



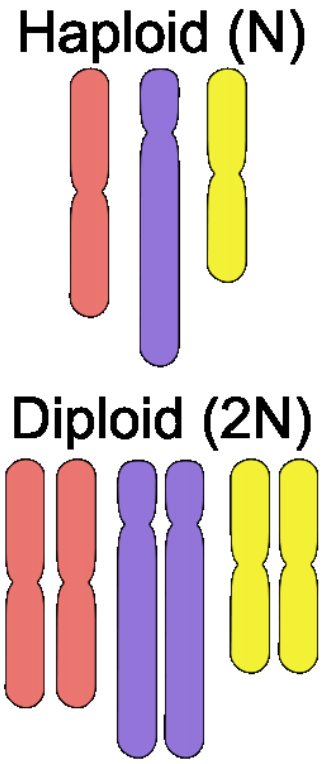
**B** **Regional Distribution of Mutations**



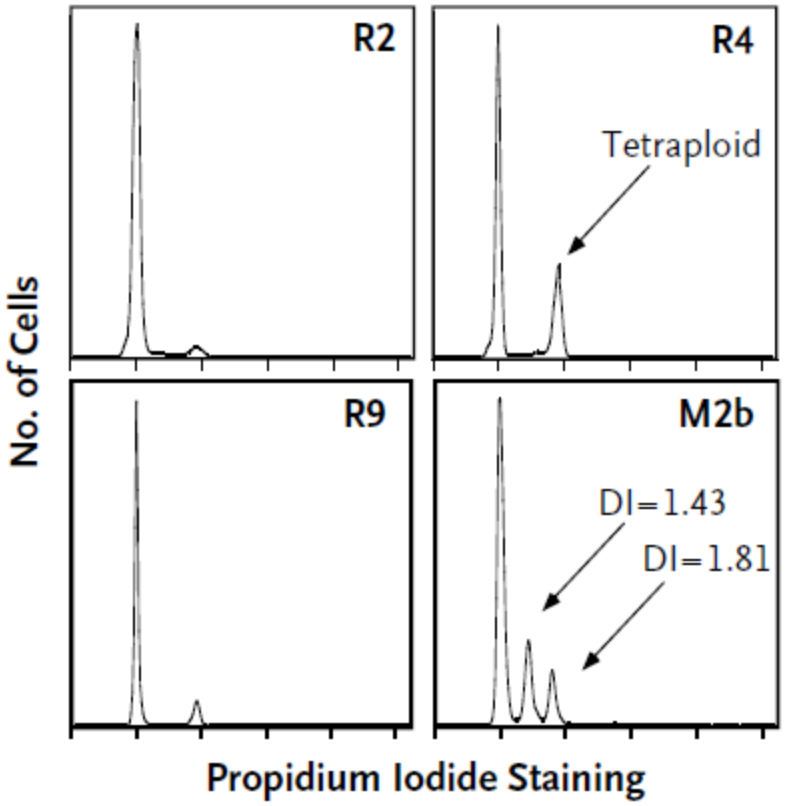

C Phylogenetic Relationships of Tumor Regions
Ubiquitous
Shared primary
Shared metastasis
Private
R1
R2
R8
R3
R9
R5
PreP
KDM5C (missense and frameshift)
mTOR (missense)
R4b
SETD2 (frameshift)
SETD2 (splice site)
Normal tissue
?
R4a
VHL
SETD2 (missense)
KDM5C (splice site)
M1
M2b
M2a
PreM

# DNA ploidy analysis

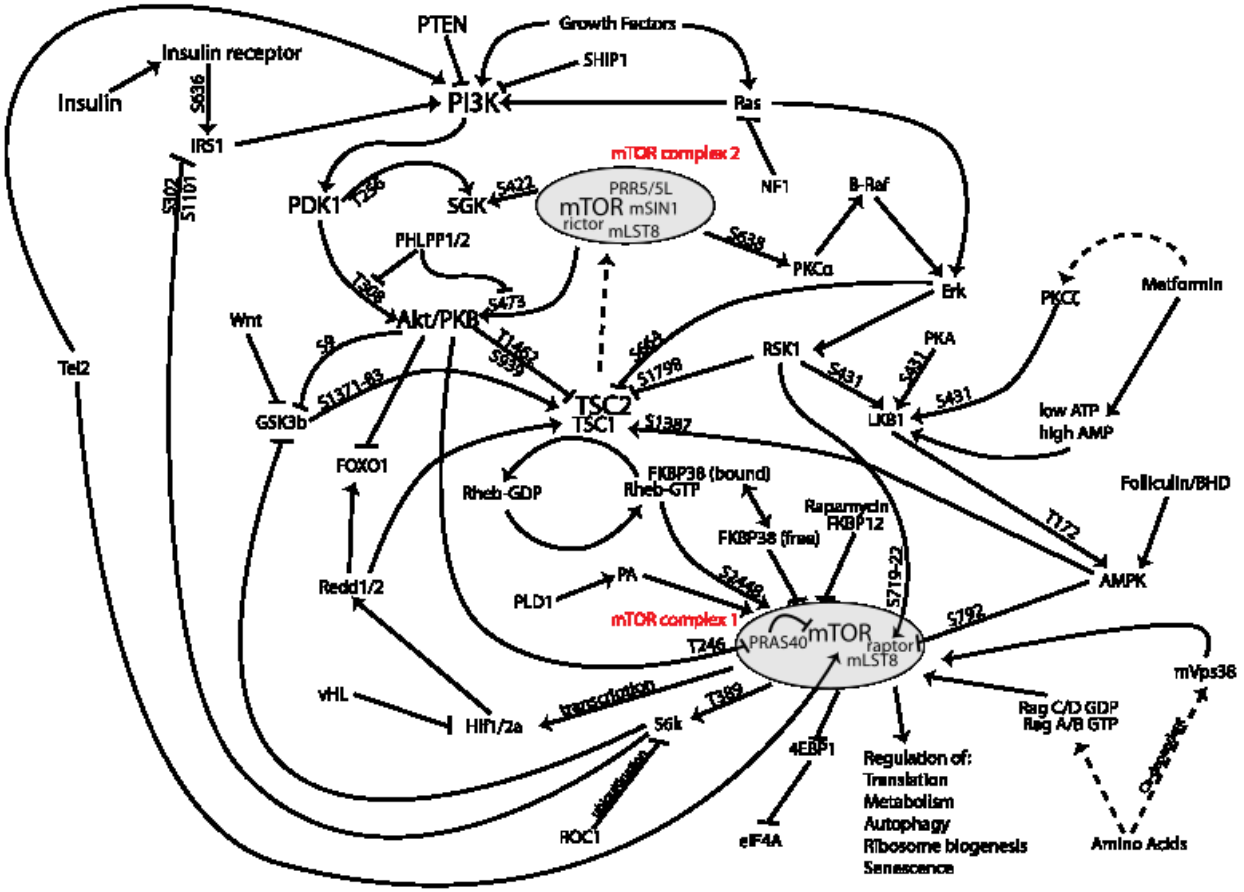
Ploidy is the number of sets of chromosomes in the nucleus of a biological cell. Normally a gamete (sperm or egg) carries a full set of chromosomes that includes a single copy of each chromosome, as aneuploidy generally leads to severe genetic disease in the offspring. The haploid number (n) is the number of chromosomes in a gamete. Two gametes form a diploid zygote with twice this number (2n) i.e. two copies of autosomal chromosomes. However, the sex chromosomes of diploid cells (excluding pseudoautosomal regions), which are subject to sex linkage, may be considered as haploid chromosomes, since haploid is also the term used to define a set of chromosomes with only one copy in the cell.

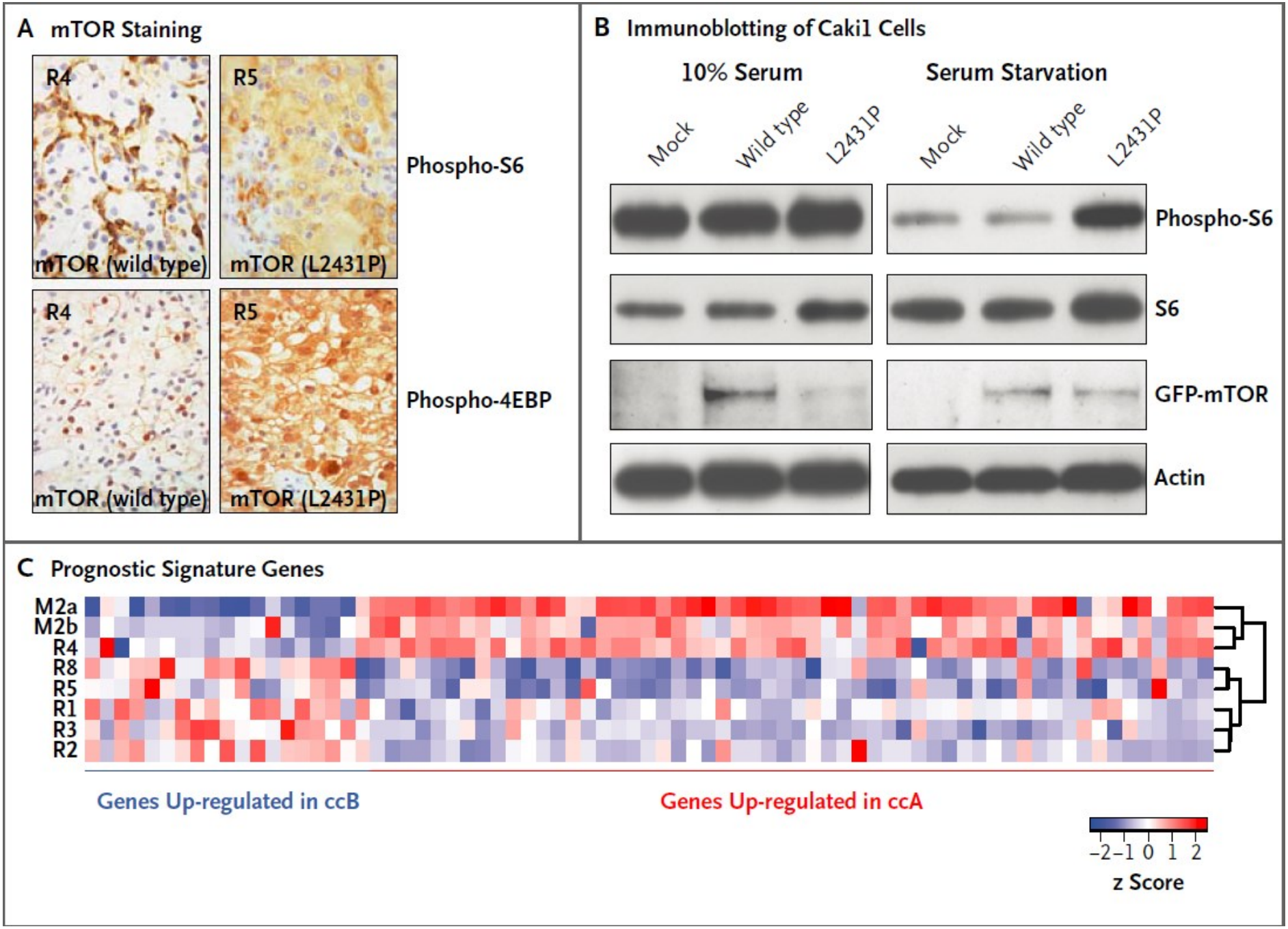


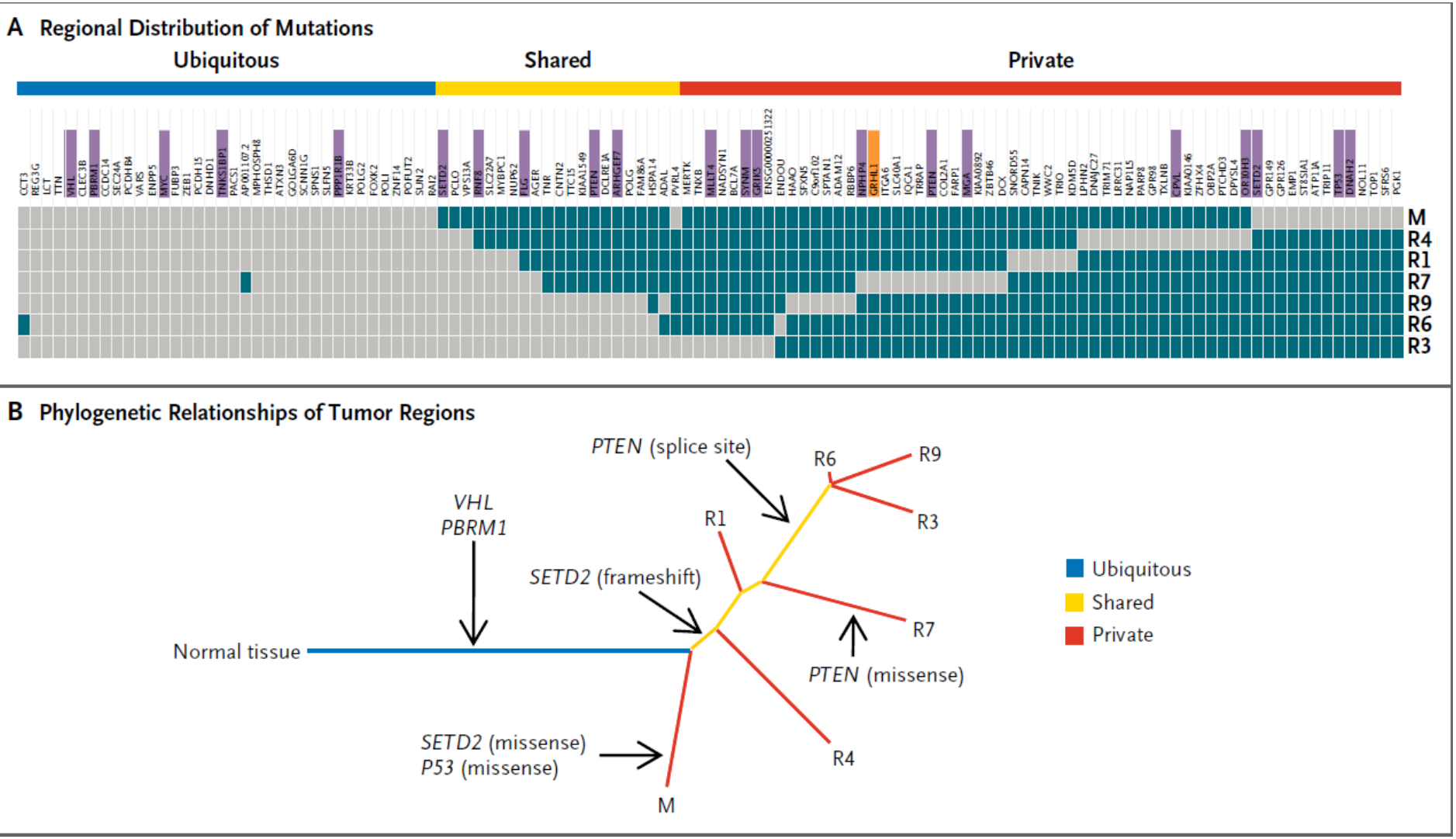
**D** Ploidy Profiling



mTOR integrates the input from upstream pathways, including insulin, growth factors (such as IGF-1 and IGF-2), and amino acids. mTOR also senses cellular nutrient and energy levels and redox status. The mTOR pathway is dysregulated in human diseases, especially certain cancers. Rapamycin is a bacterial product that can inhibit mTOR by associating with its intracellular receptor FKBP12.




A mTOR Staining
R4
R5
mTOR (wild type)
mTOR (L2431P)
Phospho-S6
R4
R5
mTOR (wild type)
mTOR (L2431P)
Phospho-4EBP
B Immunoblotting of Caki1 Cells
10% Serum
Serum Starvation
Mock
Wild type
L2431P
Mock
Wild type
L2431P
Phospho-S6
S6
GFP-mTOR
Actin
C Prognostic Signature Genes
M2a
M2b
R4
R8
R5
R1
R3
R2
Genes Up-regulated in ccB
Genes Up-regulated in ccA
−2 −1 0 1 2
z Score


A Regional Distribution of Mutations
Ubiquitous
Shared
Private
M
R4
R1
R7
R9
R6
R3
B Phylogenetic Relationships of Tumor Regions
PTEN (splice site)
VHL
PBRM1
SETD2 (frameshift)
Normal tissue
SETD2 (missense)
P53 (missense)
PTEN (missense)
Ubiquitous
Shared
Private

OPEN ACCESS Freely available online

PLOS | ONE

# TARGETgene: A Tool for Identification of Potential Therapeutic Targets in Cancer

Chia-Chin Wu[1]*, David D'Argenio[2], Shahab Asgharzadeh[3], Timothy Triche[3]

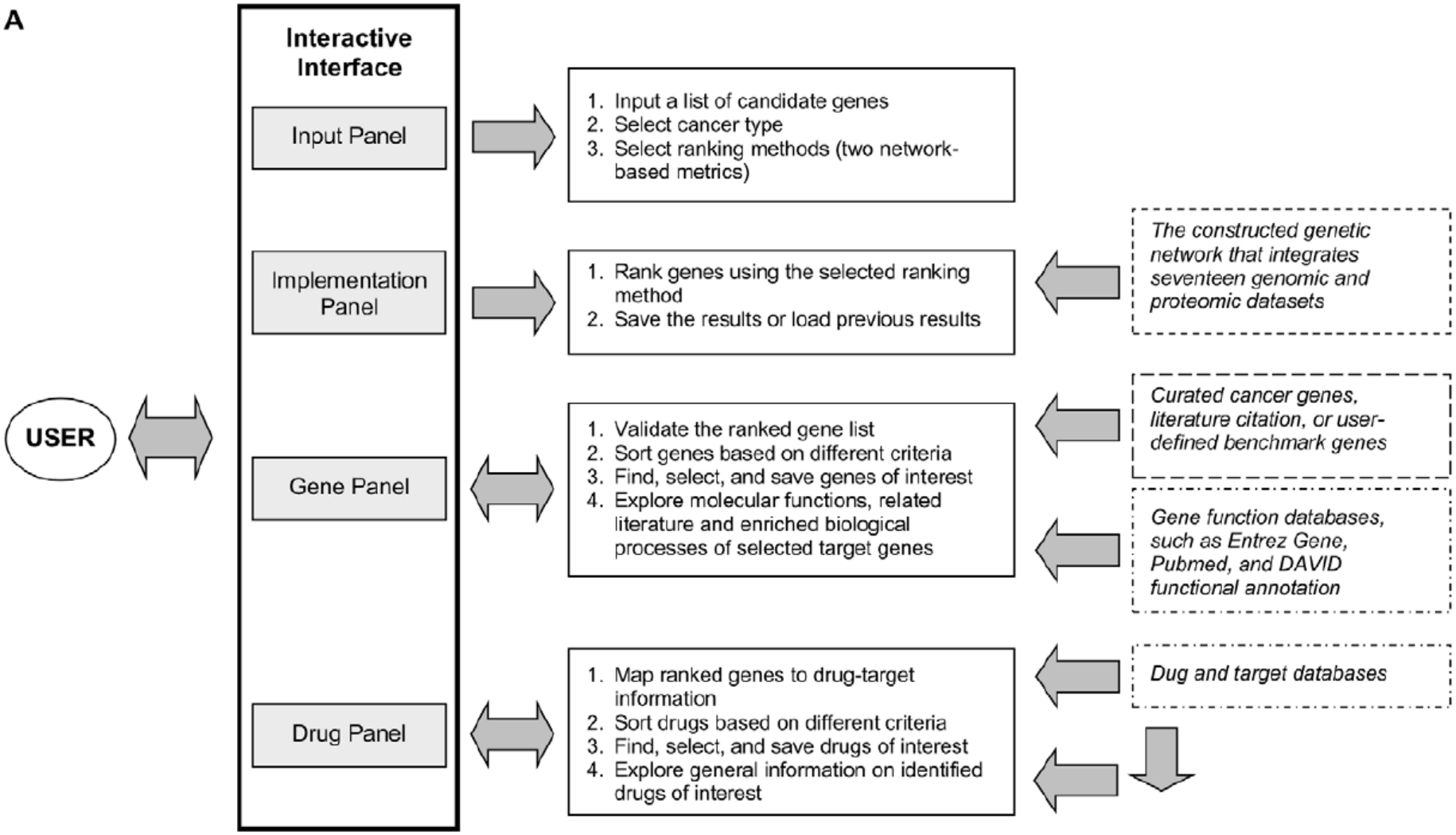
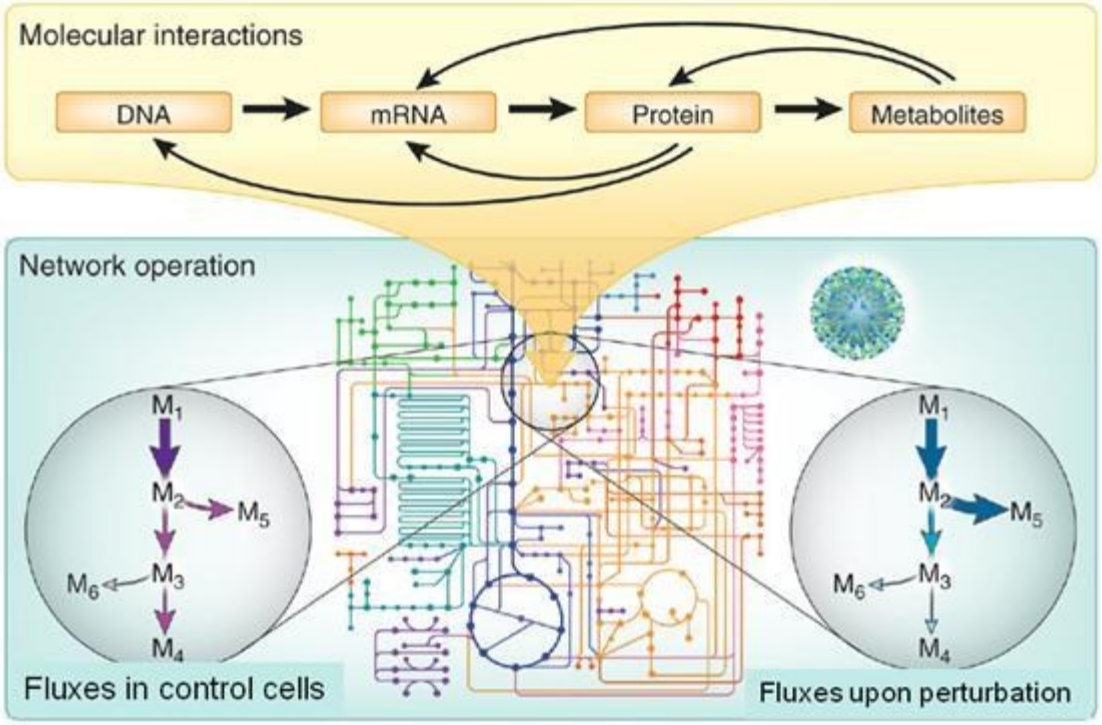

A
Interactive Interface
Input Panel
Implementation Panel
Gene Panel
Drug Panel
USER
1. Input a list of candidate genes
2. Select cancer type
3. Select ranking methods (two network-based metrics)
1. Rank genes using the selected ranking method
2. Save the results or load previous results
1. Validate the ranked gene list
2. Sort genes based on different criteria
3. Find, select, and save genes of interest
4. Explore molecular functions, related literature and enriched biological processes of selected target genes
1. Map ranked genes to drug-target information
2. Sort drugs based on different criteria
3. Find, select, and save drugs of interest
4. Explore general information on identified drugs of interest
The constructed genetic network that integrates seventeen genomic and proteomic datasets
Curated cancer genes, literature citation, or user-defined benchmark genes
Gene function databases, such as Entrez Gene, Pubmed, and DAVID functional annotation
Dug and target databases

**Table 1.** Selected Drugs Whose Targets Are Highly-Ranked (the case of Breast Cancer).

| Drugs/Compounds | Target Genes (Their Ranks and Fold Changes in Cancer) | Literatures Of Breast Cancer Treatment |
|---|---|---|
| **Dasatinib (E)*** | SRC(#10; 2.623) | [45],[46] |
| **Celecoxib (A)*** | PDPK1 (#14; 2.917) | [47] |
| **Flavopiridol (E)*** | CDK5 (#41; 4.640); CDC2 (#108; 4.382); CDK4 (#50; 2.092) | [48].[49] |
| **Staurosporine(UCN-01) (E)*** | PDPK1 (#14; 2.917); MAPKAPK2 (#62; 2.138); CSK (#19; 3.724); GSK3B (#84; 2.130) | [50] |
| **Alsterpaullone (E)** | CDK5 (#41; 4.640); GSK3B (#84; 2.130); CDC2 (#108; 4.382) | [51] |
| **Olomoucine (E)** | CDK5 (#41; 4.640); CDC2 (#108; 4.382) | [52] |
| **Trastuzumab (A)***** | ERBB2 (#25; 46.856) | [53],[54] |
| **Lapatinib (A)***** | ERBB2 (#25; 46.856) | [55],[56] |
| **Dexrazoxane (A)***** | TOP2A(#302; 10.965) | [57] |
| **Lithium (A)** | GSK3B (#84) | [58] |
| **Melatonin (A)** | CALR(#651) | [59] |
| **Calcidiol (A)** | VDR (#241) | [44] |
| **Vorinostat (A)*** | HDAC3 (#307; 2.336); HDAC1 (#497; 2.286); HDAC2 (#564; 2.520) | [60] |
| **Geldanamycin (17-AAG) (E)*** | HSP90B1 (#258; 1.779); HSP90AA1 (#275; 1.920) | [61],[62] |
| **Arsenic trioxide (A)*** | AKT1 (#1; 4.566); CCND1 (#418; 3.663) | [63],[64] |

**Note**: 1.Approved drugs are denoted as 'A'.
2.Experimental compounds are denoted as 'E'.
3.Drugs have been approved for the treatment of Breast Cancer are marked with ***.
4.Drugs in clinical trials for Breast Cancer are marked with *.
doi:10.1371/journal.pone.0043305.t001


Molecular interactions
DNA
mRNA
Protein
Metabolites
Network operation
M1
M2
M5
M6
M3
M4
Fluxes in control cells
M1
M2
M5
M6
M3
M4
Fluxes upon perturbation