Mutual exclusivity analysis identifies oncogenic network modules

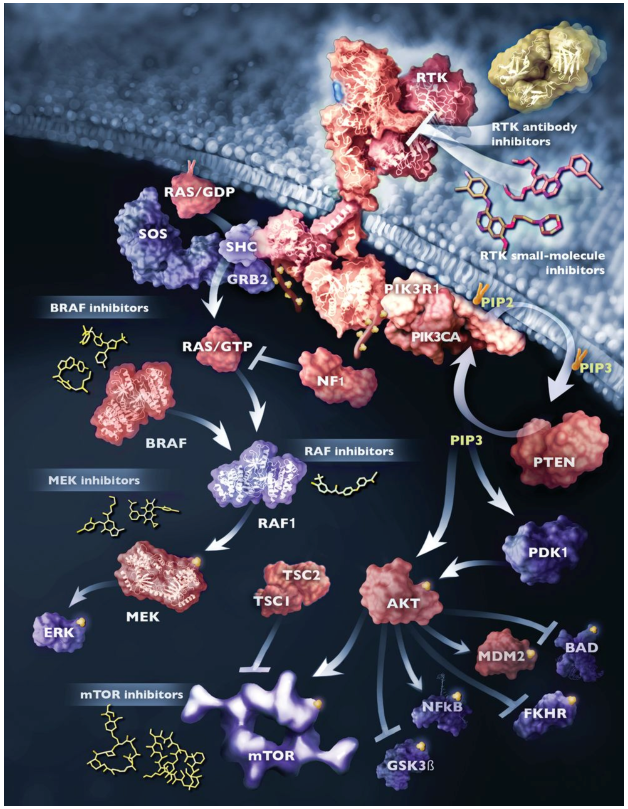
Giovanni Ciriello,1,3,4 Ethan Cerami,1,2,3 Chris Sander,1 and Nikolaus Schultz1

- Background
- Motivation
- Method
- Result

Background

- Pathway

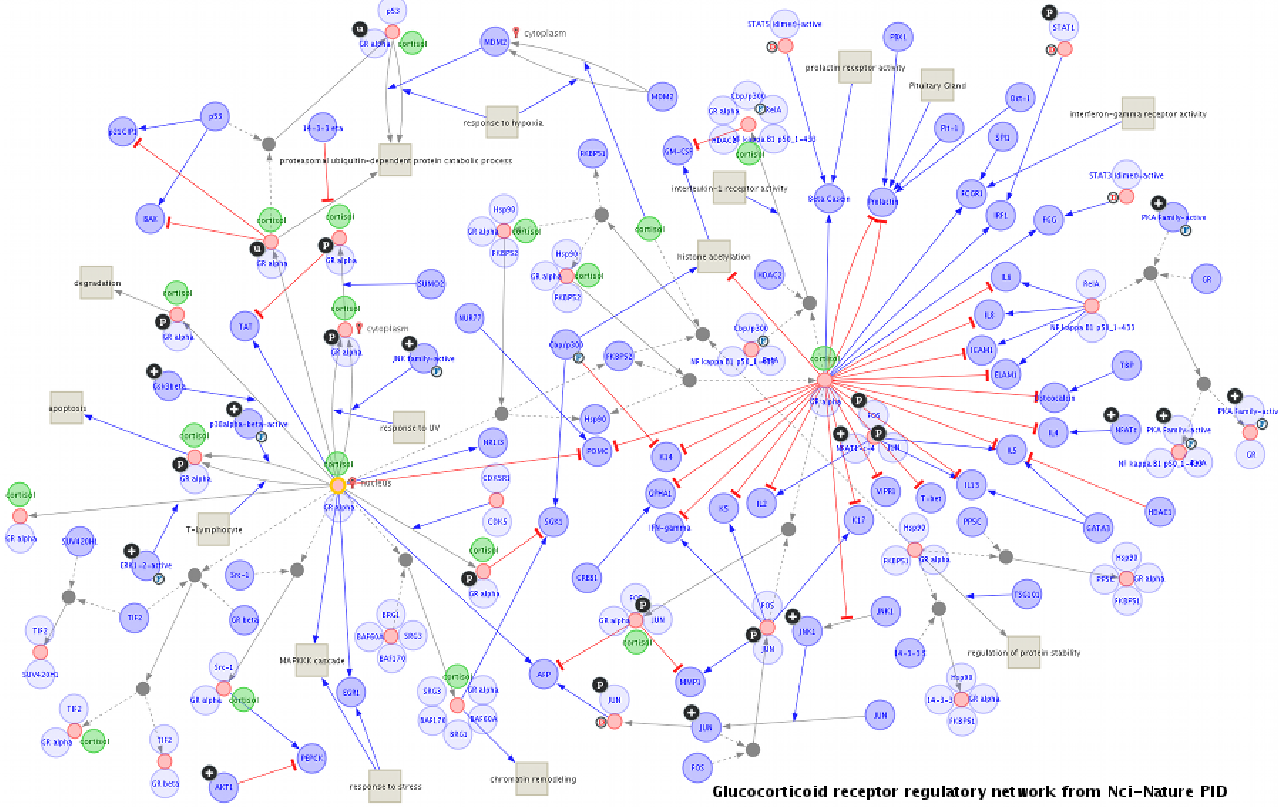
Smallest functional unit of a network of proteins that interacts to performs a single task.



Background

- Network

  Union of all pathways



Glucocorticoid receptor regulatory network from Nci-Nature PID

Motivation

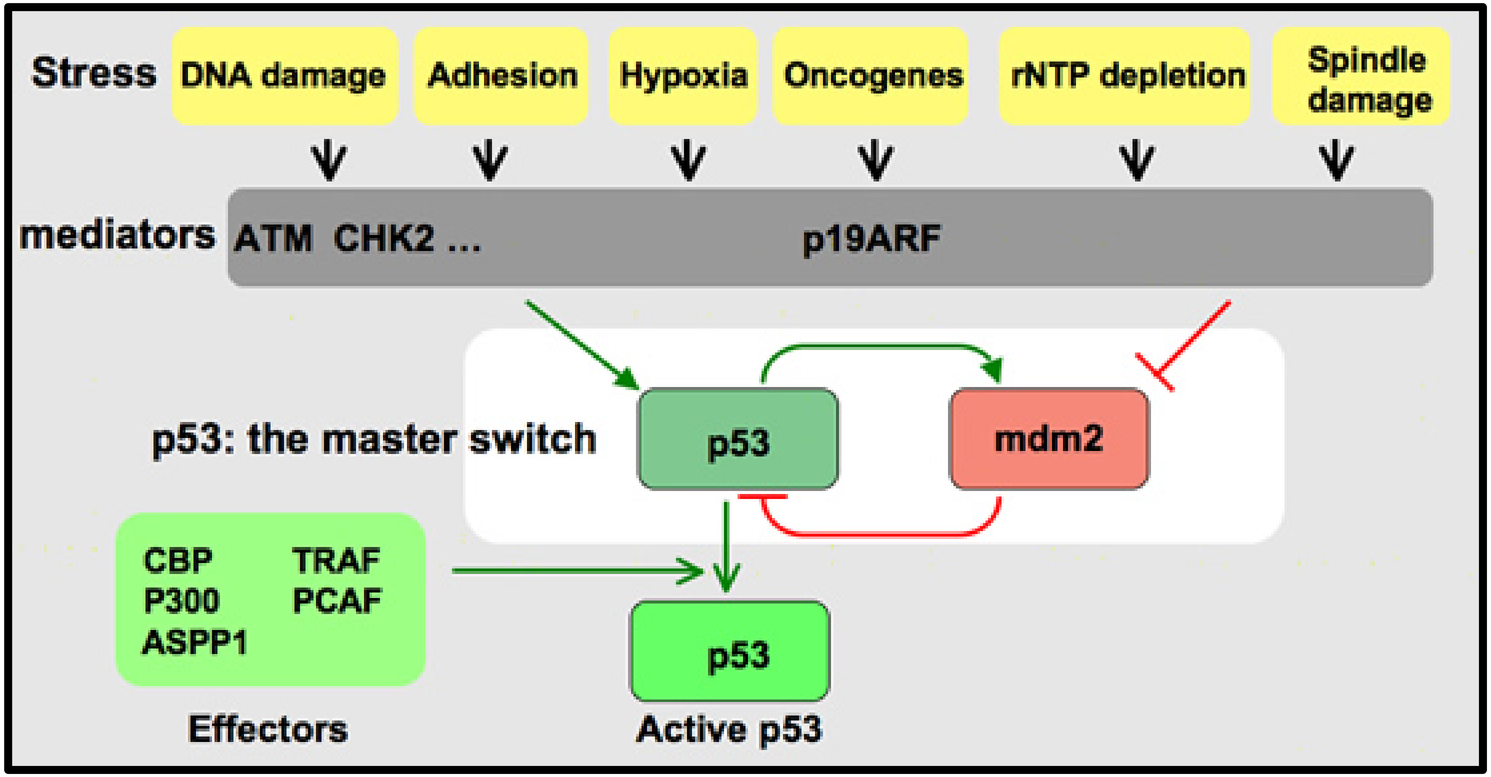
- Basic motivation: to identify oncogenic pathway modules.

Why Mutual exclusivity analysis?

- Many oncogenic events effect a limited number of biological pathways
- Mutually exclusive genomic alteration observed

Example

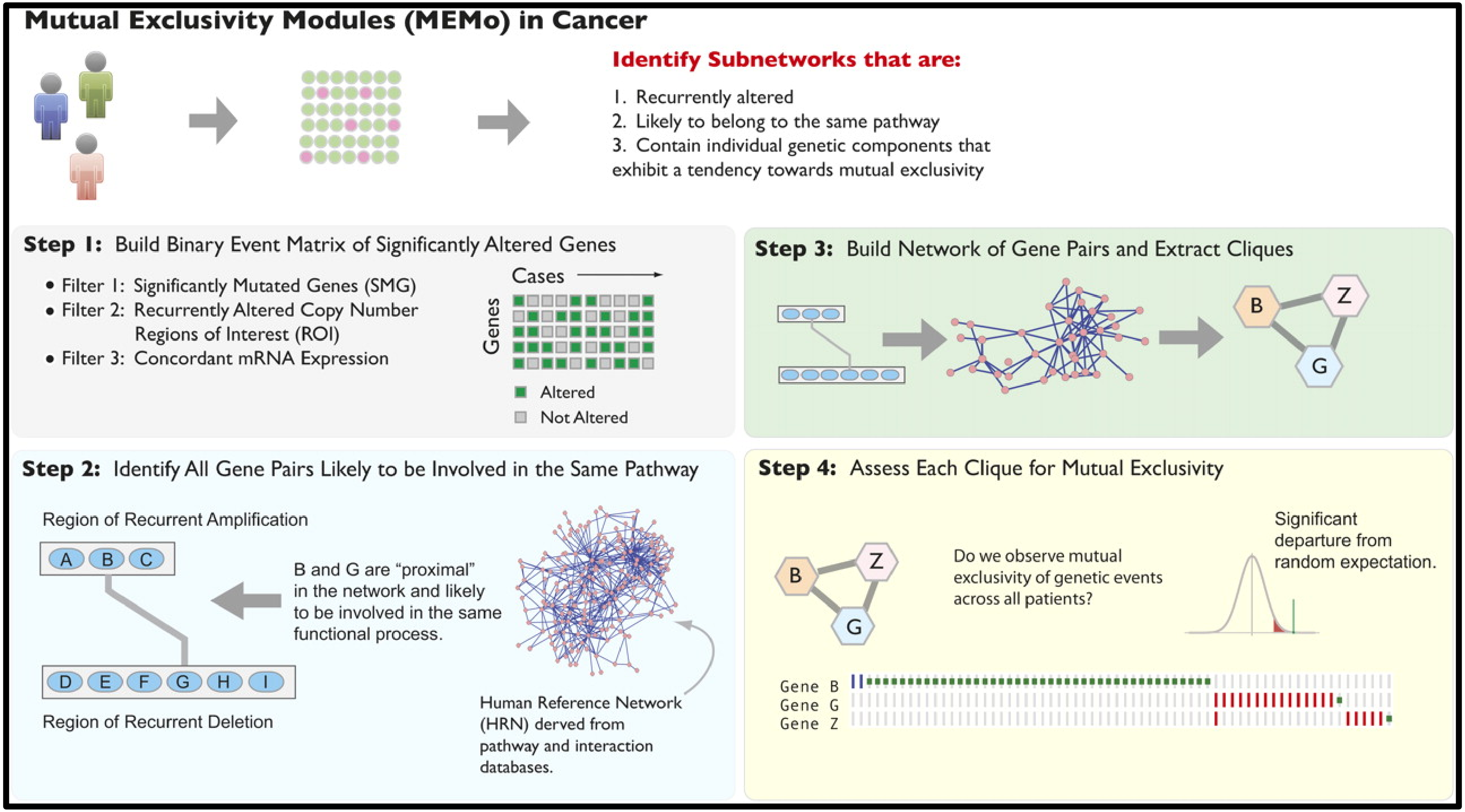
P53 VS MDM2



Goal of MEMo

Identify sets of connected genes that are recurrently altered, likely belongs to the same pathway or biological process, and exhibit patterns of mutually exclusive generic alteration across multiple patients.

Method



Mutual Exclusivity Modules (MEMo) in Cancer

Identify Subnetworks that are:

1. Recurrently altered
2. Likely to belong to the same pathway
3. Contain individual genetic components that exhibit a tendency towards mutual exclusivity

Step 1: Build Binary Event Matrix of Significantly Altered Genes

- Filter 1: Significantly Mutated Genes (SMG)
- Filter 2: Recurrently Altered Copy Number Regions of Interest (ROI)
- Filter 3: Concordant mRNA Expression

Step 2: Identify All Gene Pairs Likely to be Involved in the Same Pathway

B and G are "proximal" in the network and likely to be involved in the same functional process.

Human Reference Network (HRN) derved from pathway and interaction databases.

Step 3: Build Network of Gene Pairs and Extract Cliques

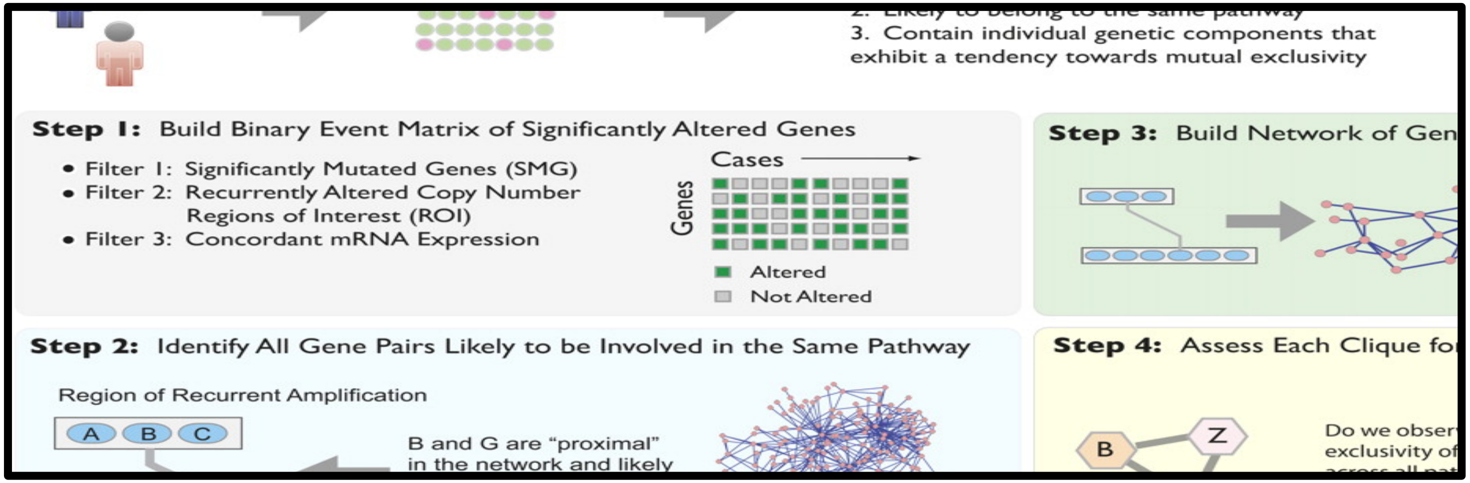
Step 4: Assess Each Clique for Mutual Exclusivity

Do we observe mutual exclusivity of genetic events across all patients?

Significant departure from random expectation.

Step 1: Build Binary Event Matrix of Significant Altered Genes

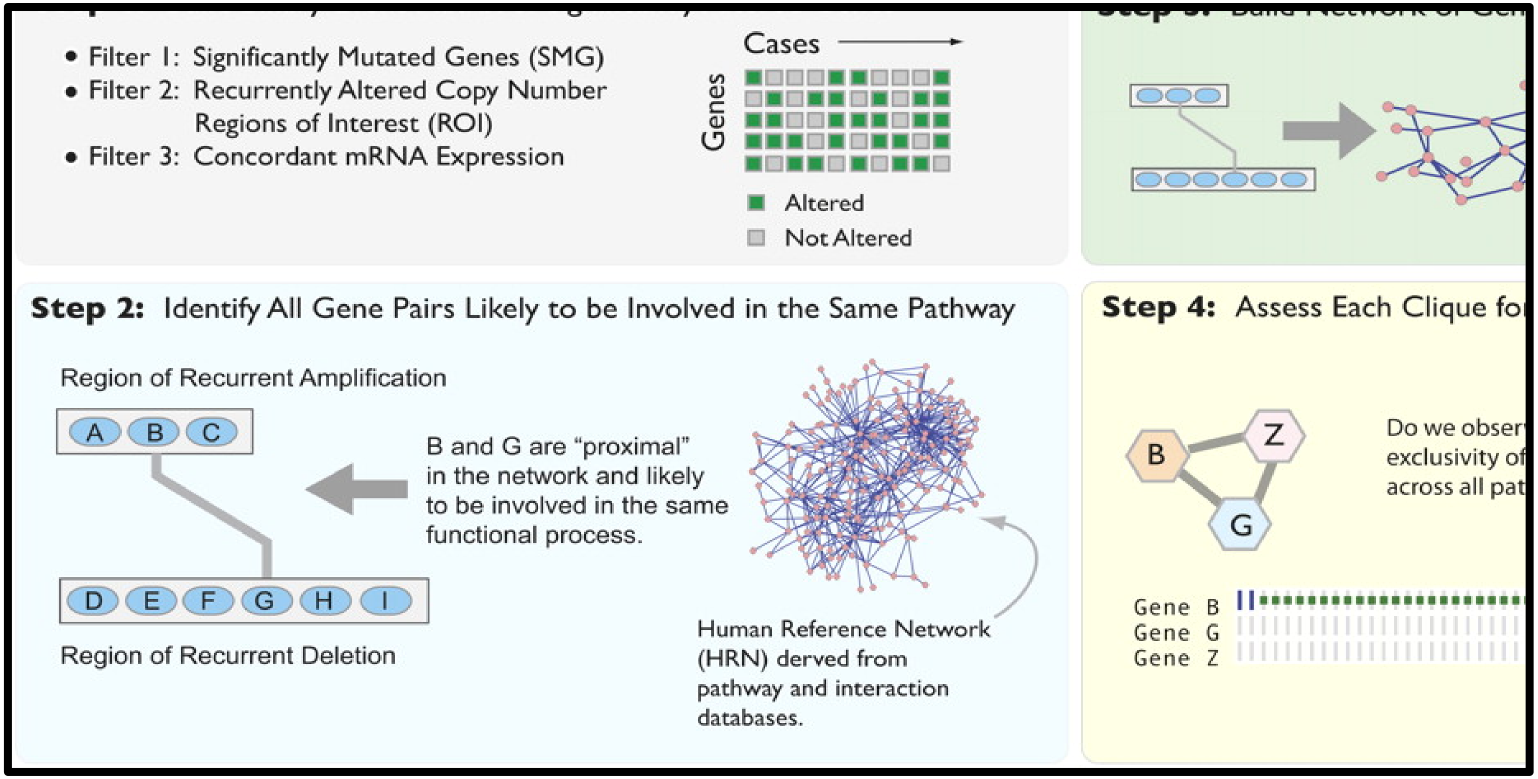
- The first filter identifies genes that are mutated significantly above the background mutation rate (BMR).
- The second filter identifies genes that are targets of recurrent copy number amplification or deletion.
- The third filter identifies copy number altered genes that have concordant mRNA expression



Note

- Genes that does not have a concordant mRNA expression would not likely to change the pathway function and therefore unlikely to be drivers.
- The binary matrix built does not take into account for the multiply mutation within a gene/case, nor does it not account for varying allelic frequency

Step 2: Identifying all gene pairs likely to be involved in the same pathway



- Filter 1: Significantly Mutated Genes (SMG)
- Filter 2: Recurrently Altered Copy Number Regions of Interest (ROI)
- Filter 3: Concordant mRNA Expression

Step 2: Identify All Gene Pairs Likely to be Involved in the Same Pathway

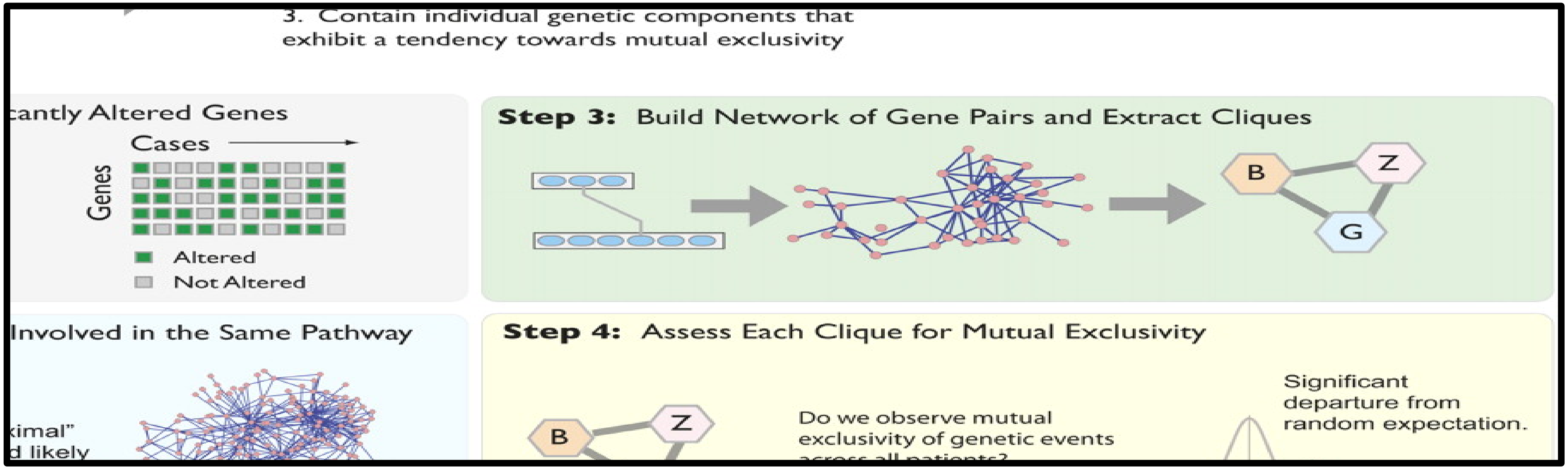
B and G are "proximal" in the network and likely to be involved in the same functional process.

Human Reference Network (HRN) derved from pathway and interaction databases.

Step 4: Assess Each Clique fo

Do we obser
exclusivity of
across all pat

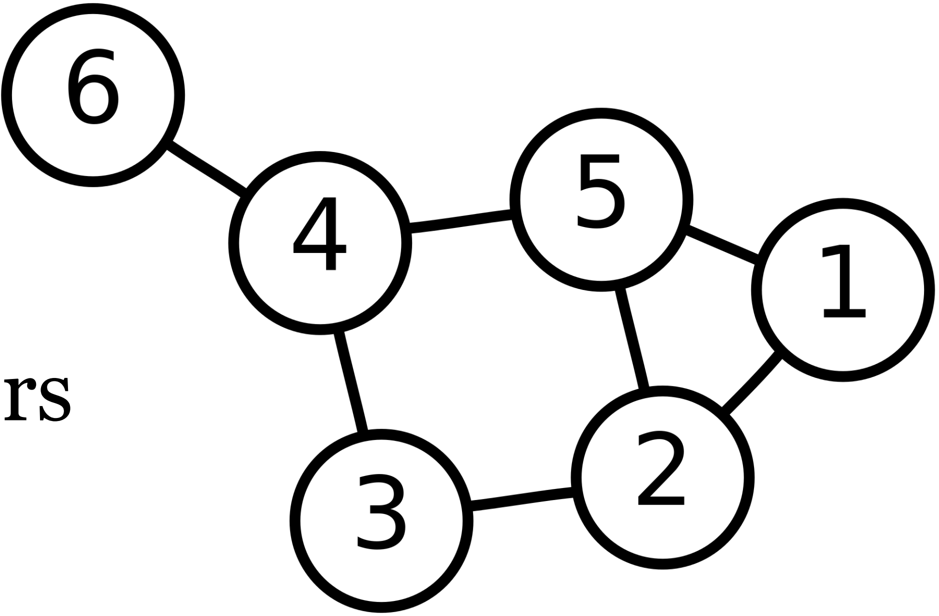
Step 3: Build graph of gene pairs and extract clique



Similarity metric between genes

$$J(u,v) = \left| \frac{N(u) \cap N(v)}{N(u) \cup N(v)} \right|$$

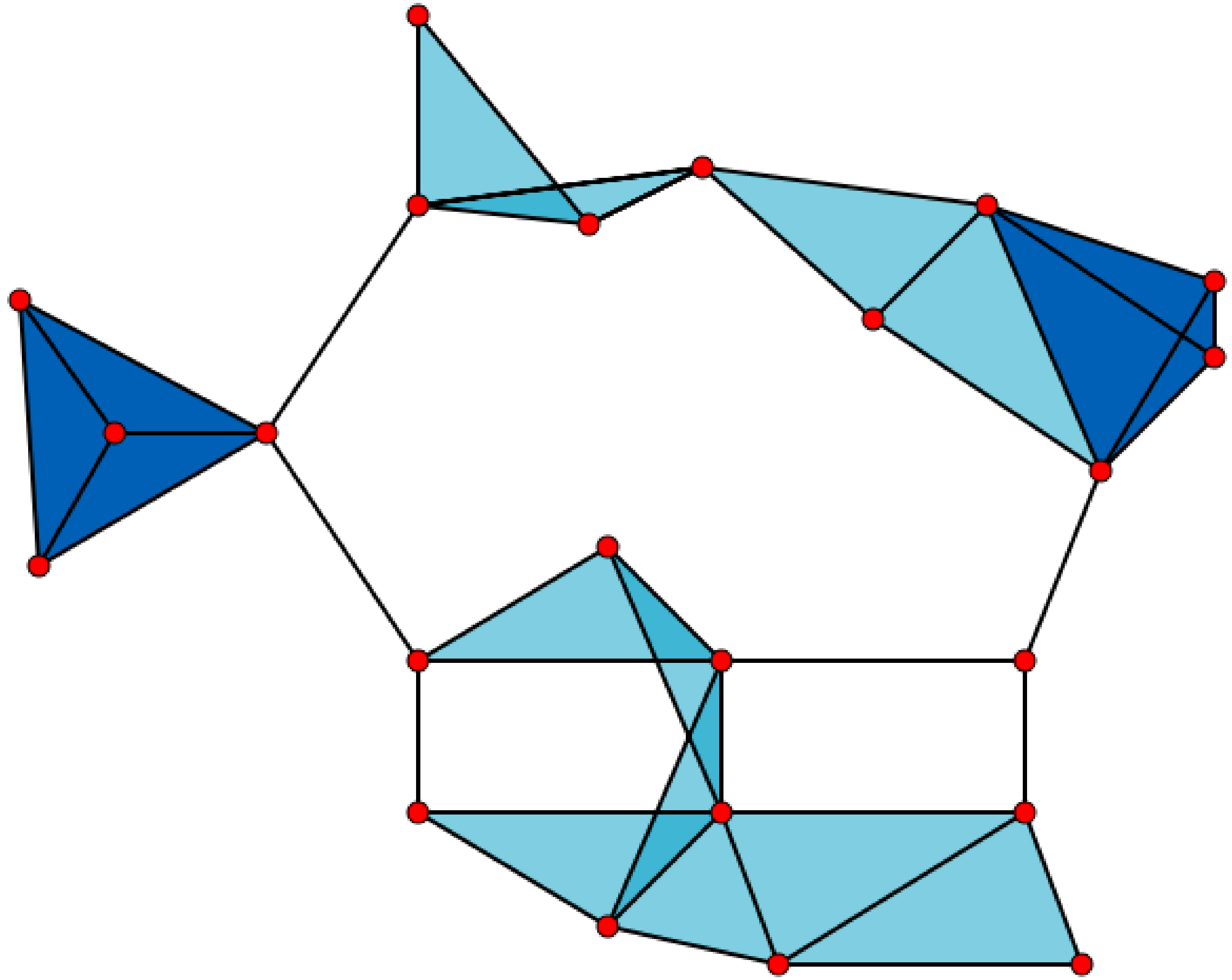
J_{avg} is 4% to 7% for known gene pairs that have similar functions



Connecting similar genes

If a pair of genes has a high J value, marked them as functional similar and connect them.

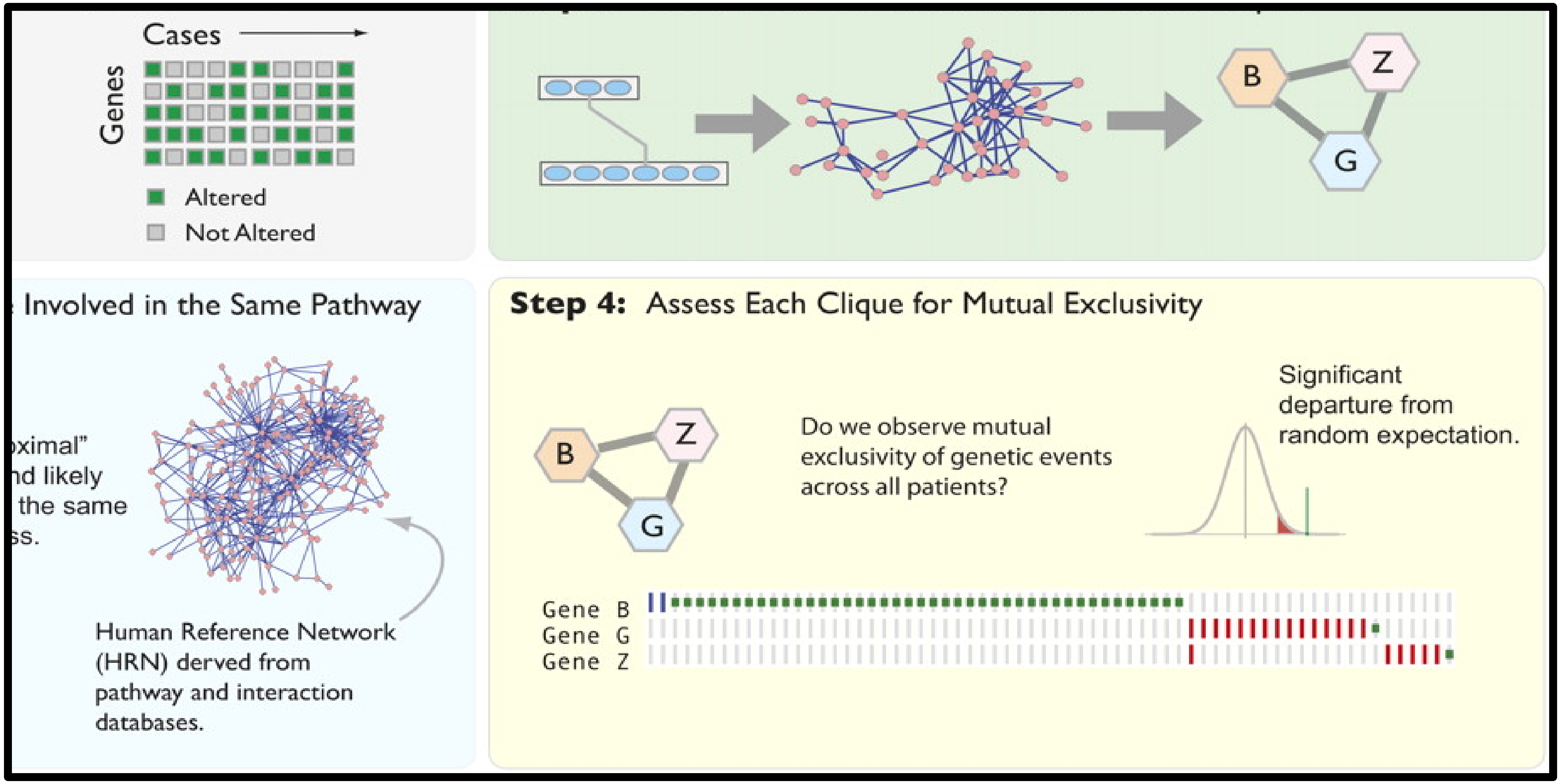
Clique extraction



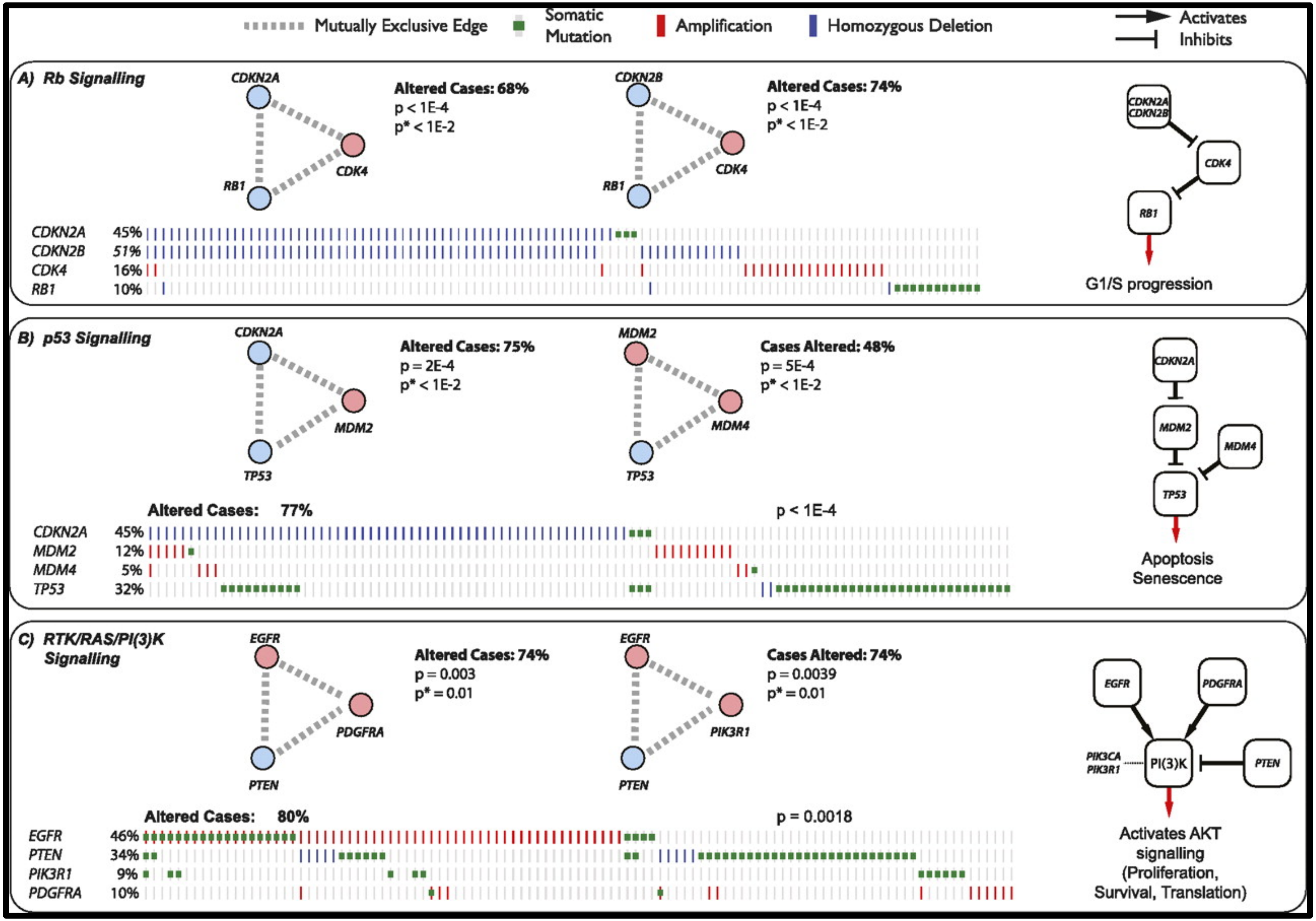
Non informative clique deletion

A clique is said to be informative if number of times the corresponding gene is altered concurrently with other genes in the clique is smaller than the number of unique alterations

Step 4: Mutual exclusivity test



Result


Mutually Exclusive Edge
Somatic Mutation
Amplification
Homozygous Deletion
Activates
Inhibits
A) Rb Signalling
CDKN2A
CDK4
RB1
Altered Cases: 68%
p < 1E-4
p* < 1E-2
CDKN2B
CDK4
RB1
Altered Cases: 74%
p < 1E-4
p* < 1E-2
CDKN2A 45%
CDKN2B 51%
CDK4 16%
RB1 10%
CDKN2A CDKN2B
CDK4
RB1
G1/S progression
B) p53 Signalling
CDKN2A
MDM2
TP53
Altered Cases: 75%
p = 2E-4
p* < 1E-2
MDM2
MDM4
TP53
Cases Altered: 48%
p = 5E-4
p* < 1E-2
Altered Cases: 77%
p < 1E-4
CDKN2A 45%
MDM2 12%
MDM4 5%
TP53 32%
CDKN2A
MDM2
MDM4
TP53
Apoptosis
Senescence
C) RTK/RAS/PI(3)K Signalling
EGFR
PDGFRA
PTEN
Altered Cases: 74%
p = 0.003
p* = 0.01
EGFR
PIK3R1
PTEN
Cases Altered: 74%
p = 0.0039
p* = 0.01
Altered Cases: 80%
p = 0.0018
EGFR 46%
PTEN 34%
PIK3R1 9%
PDGFRA 10%
EGFR
PDGFRA
PIK3CA
PIK3R1
PI(3)K
PTEN
Activates AKT signalling (Proliferation, Survival, Translation)