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 perform a single task.
Background

- Network

Union of all pathways
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

Step 3: Build Network of Gene Pairs and Extract Cliques

Step 4: Assess Each Clique for Mutual Exclusivity

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

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

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

Gene B
Gene G
Gene Z

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

Region of Recurrent Amplification

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

Region of Recurrent Deletion

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

Step 4: Assess Each Clique for...
**Step 3:** Build graph of gene pairs and extract clique
Similarity metric between genes

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

\( J_{\text{avg}} \) 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

A) Rb Signalling

CDKN2A 45%
CDKN2B 51%
CDK4 16%
RB1 10%

Altered Cases: 68%
p < 1E-4
p* < 1E-2

G1/S progression

B) p53 Signalling

CDKN2A 45%
MDM2 12%
MDM4 5%
TP53 32%

Altered Cases: 77%
p < 1E-4

Apoptosis
Senescence

C) RTK/RAS/PIK(3)K Signalling

EGFR 46%
PTEN 34%
PIK3R1 9%
PDGFRα 10%

Altered Cases: 80%
p = 0.003
p* = 0.01

Activates AKT signalling
(Proliferation, Survival, Translation)