

# Supplementary Figures and Tables

Selcen Ari Alper Yilmaz

2020-10-30

## Contents

<b>1. Functions defined for ceRNA models and workflow of method</b>	<b>1</b>
<b>2. Supplementary Figures</b>	<b>2</b>
2.1 Sample dataset analysis in absence of interaction factors.	4
2.2 Calculations with interaction factors	7
2.3 Sample+ dataset analysis with interaction factors.	8
2.4 Common target perturbation in Sample+ dataset.	10
<b>3 Supplementary Tables</b>	<b>12</b>
3.1 <i>Sample+</i> dataset	12
3.2 Significant factors in miRNA:target interactions	12
3.3 Content of High-throughput experimental studies	13
3.4 Variables of network object during simulation	14
<b>4. Notes</b>	<b>14</b>
<b>REFERENCES</b>	<b>15</b>

## 1. Functions defined for ceRNA models and workflow of method

We defined the functions that can be used with R programming. Briefly, these functions process a given miRNA:gene dataset and convert to graph object. All values that are significant in miRNA:target interactions are stored in edge variables and processed with formulations that are given in previous section. The functions and steps of approach are explained as following (Figure S1) :

**Conversion of dataset:** `priming_graph()` function processes the given dataset that includes competing elements in first variable and repressive element in second variable. If the affinity and/or degradation factors are specified in the function, factors are taken into account, are processed with defaults in vice versa. The formulations that are given in equations (1-4) are performed in this function. This step gives the graph object which contains efficiency values of miRNA:competing target pairs in steady-state in terms of amount. It is assumed that the initial target amounts in the dataset is observed after the repressive activity of miRNAs in steady-state.

**Transition of variables in graph:** In the previous step, the calculations are performed in the edge variables of the graph object. However, the graph object allows to use node variables, while the node features are handled to the graph. In this direction, `update_nodes` function carries the amount values to node variables. This step must be applied with “once” option because it is primary process.

**Trigger change in graph:** The dataset are assumed as steady-state in previous step and the efficiency coefficients are calculated according to this acceptance. In the network that is found in steady-state conditions, the change is applied to the graph object for perturbation of steady-state. To provide the perturbation in the network the work-flow offer two methods: `update_variables` and `update_how`. The first, a new dataset that is contained competing and repressive element names and current values of these can be processed with `update_variables`. The second option, the amount of the given node name in `update_how` function can be changed according to “how” argument.

**Simulation of competing behavior of targets:** After the change in the steady-state conditions, the network elements try to gain steady-state again. This process progresses as repeating of regulations after the spreading the changes in the network. In this step, simulation of regulations according to given cycle count in `simulate` function is applied. After each simulation cycle, the miRNA repression values are re-calculated and the current values of competing elements are found and saved. The process is performed in the edge data and at the same time outputs of the calculations are carried from edge to node data.

The node elements in the dataset are handled as two type; repressive (miRNAs) and competing (targets). It is assumed in approach that while targets are degrading or inhibiting by miRNAs continuously, miRNAs reversibly used. If the trigger of the network is a miRNA, it maintains the current value of amount that provides by user. On the contrary, it tries to help this process to provide steady-state through the regulations on its amount, if a competing element is used as a trigger. The functions that are used in the this study are developed with R and are available in [Bioconductor](#).

## 2. Supplementary Figures

The codes used to obtain networks in this file can be found on the github page. Images that produced have been referred with codes.

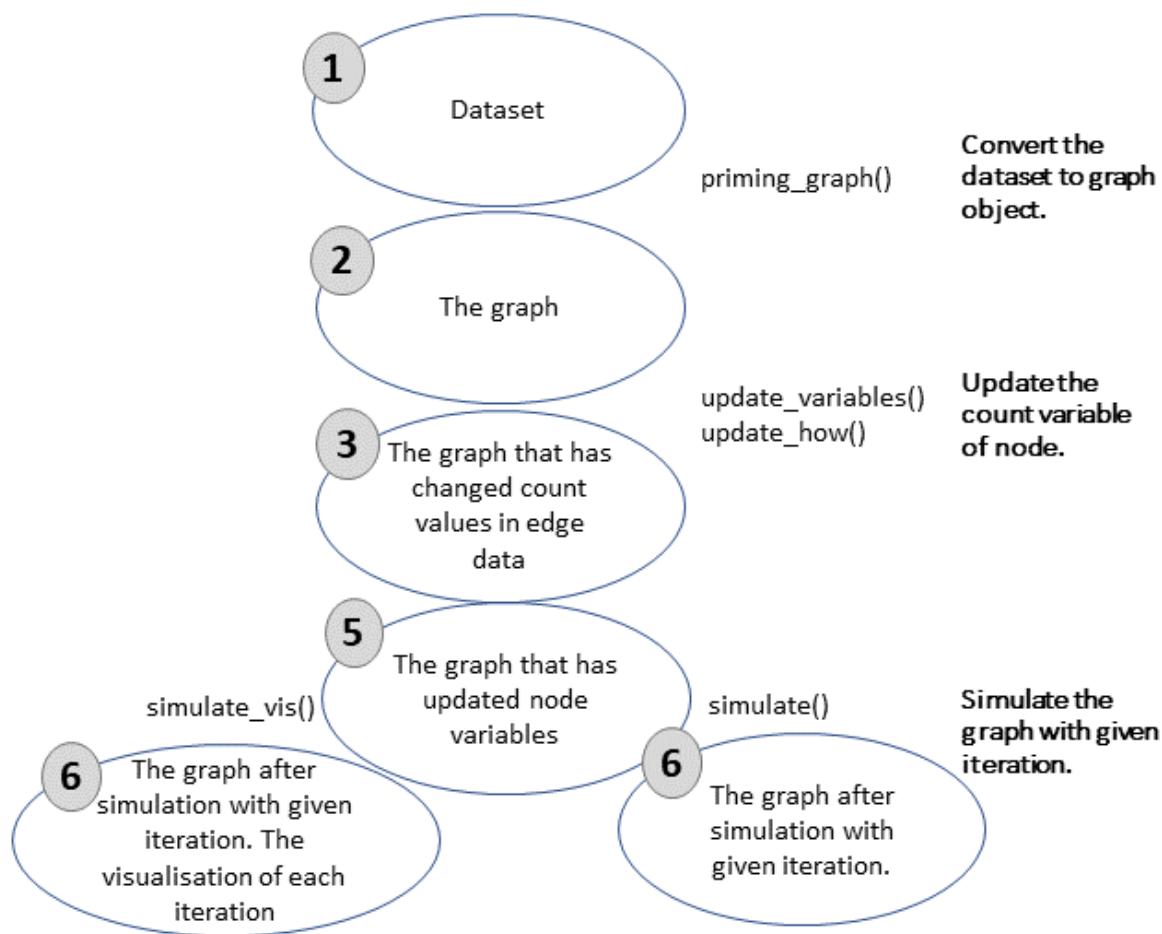


Figure S1: Workflow for simulation of competing endogenous RNA regulations. Graph object in steps 2-6 is saved and updated continuously.

## 2.1 Sample dataset analysis in absence of interaction factors.

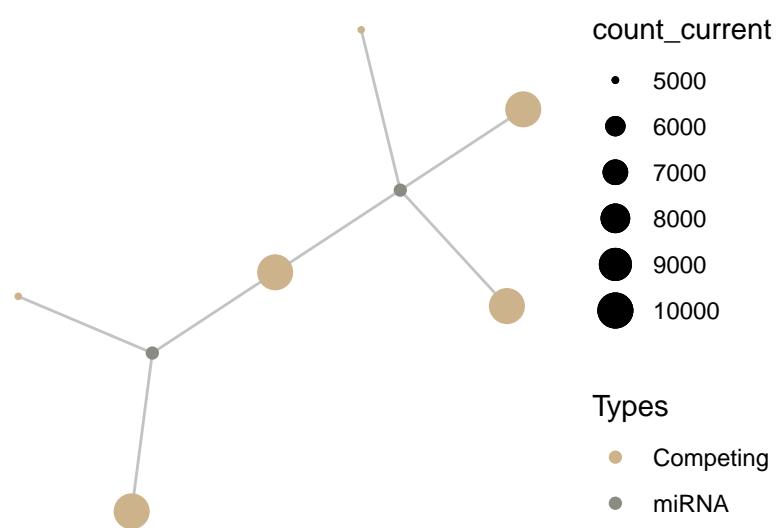


Figure S2: Sample Dataset in Steady-state

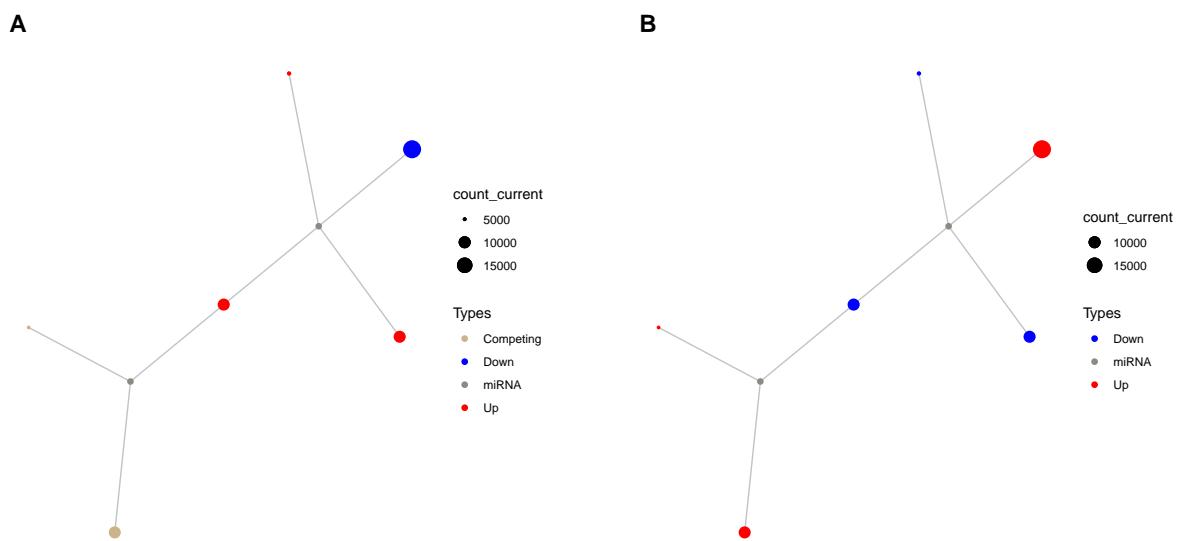
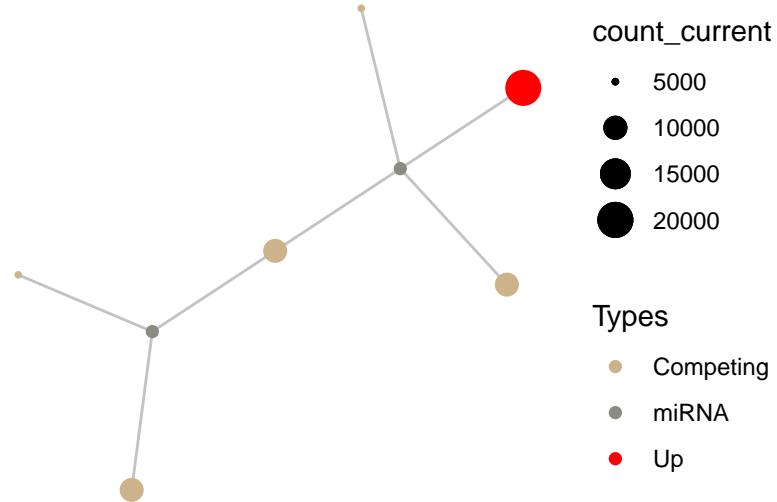


Figure S4: Sequential iteration of Sample data. A) First response of system to Gene2 upregulation (2nd iteration). B) Spreading of perturbation on system (3rd iteration)



## 2.2 Calculations with interaction factors

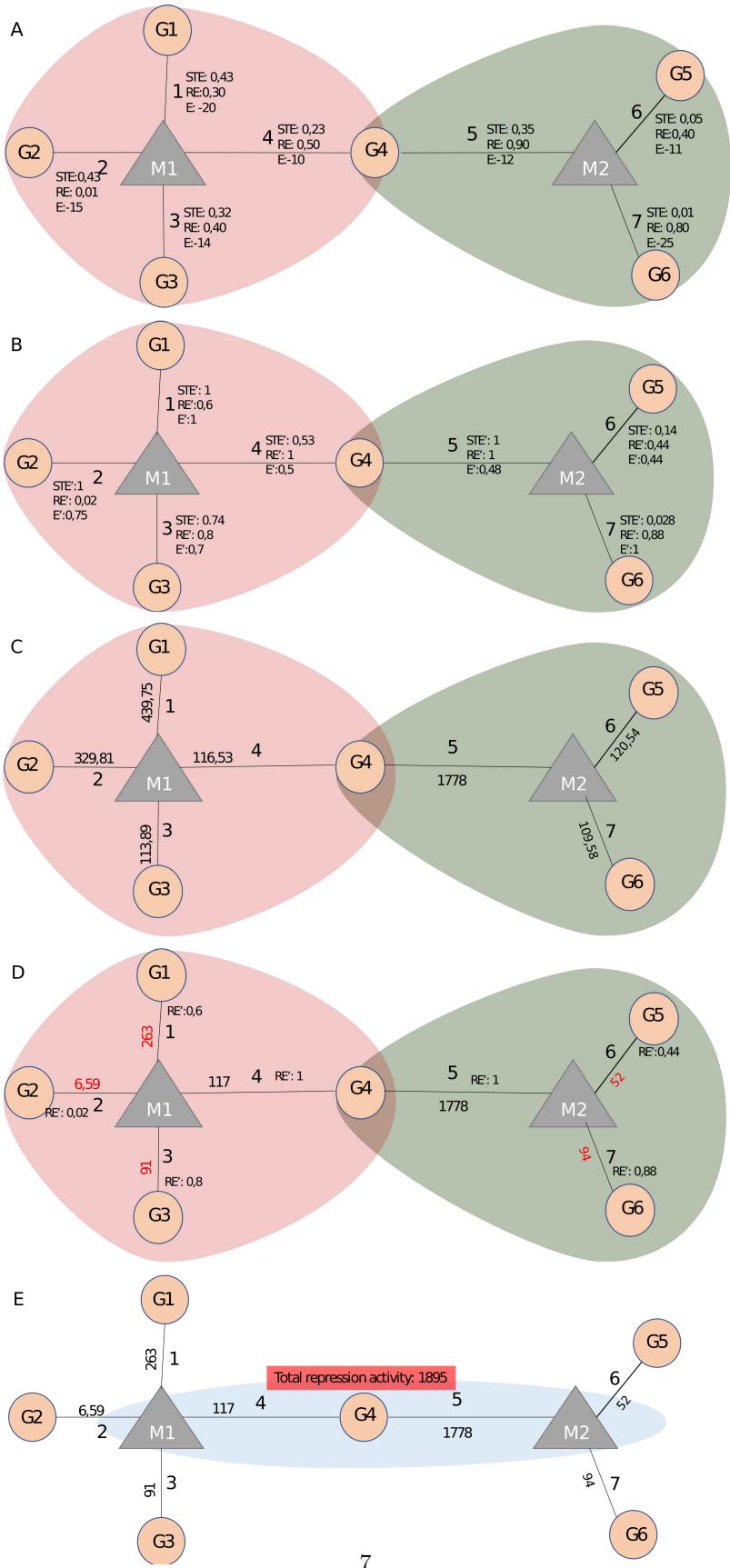


Figure S5: Calculations to determine of miRNA binding and repression efficiency.

$G$ , Gene;  $M$ , miRNA;  $STE$ , seed type effect;  $RE$ , Region Effect;  $E$ , Energy;  $STE'$ , normalized values of seed type efficiency coefficient;  $RE'$ , normalized values of region efficiency coefficient;  $E'$ , normalized values of energy coefficient. Numbering on edges match the pair order in Table S1.

### 2.3 Sample+ dataset analysis with interaction factors.

Interaction factors of Sample+ network are available in Table S1.

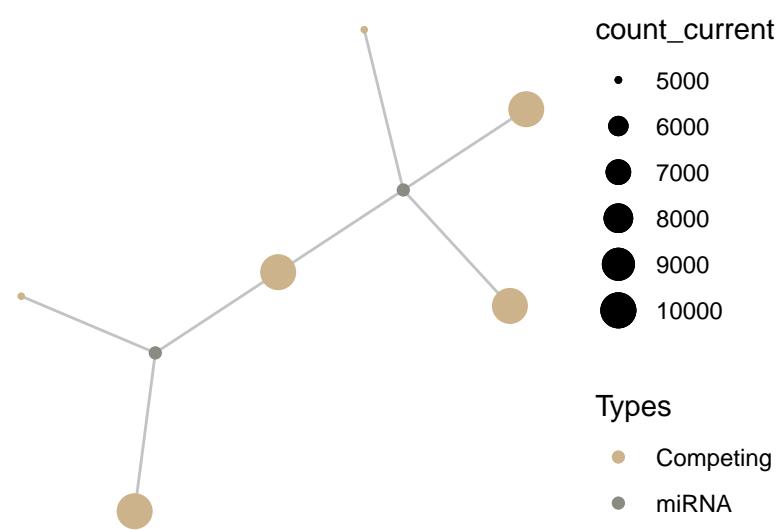


Figure S6: Sample+ in Steady-state

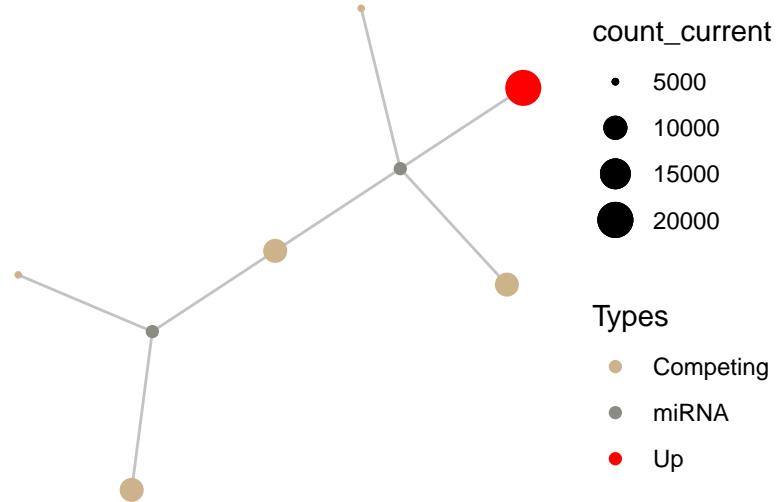


Figure S7: When Gene2 is upregulated on Sample+

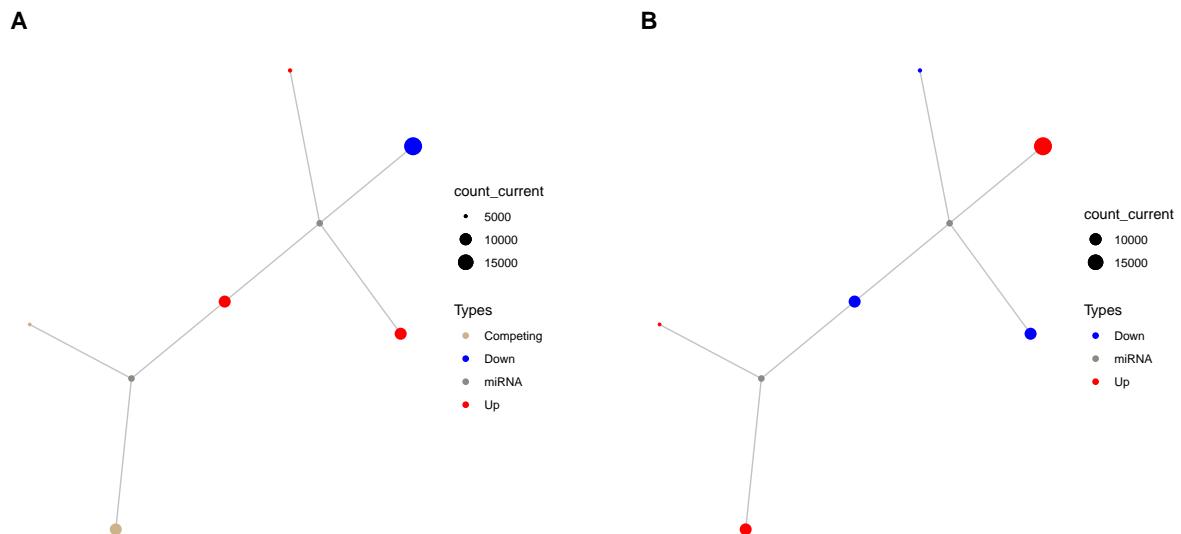


Figure S8: Sequential iteration of Sample+ A)First response of system to Gene2 upregulation (2nd iteration). B)Spreading of perturbation on system (3th iteration)

## 2.4 Common target perturbation in Sample+ dataset.

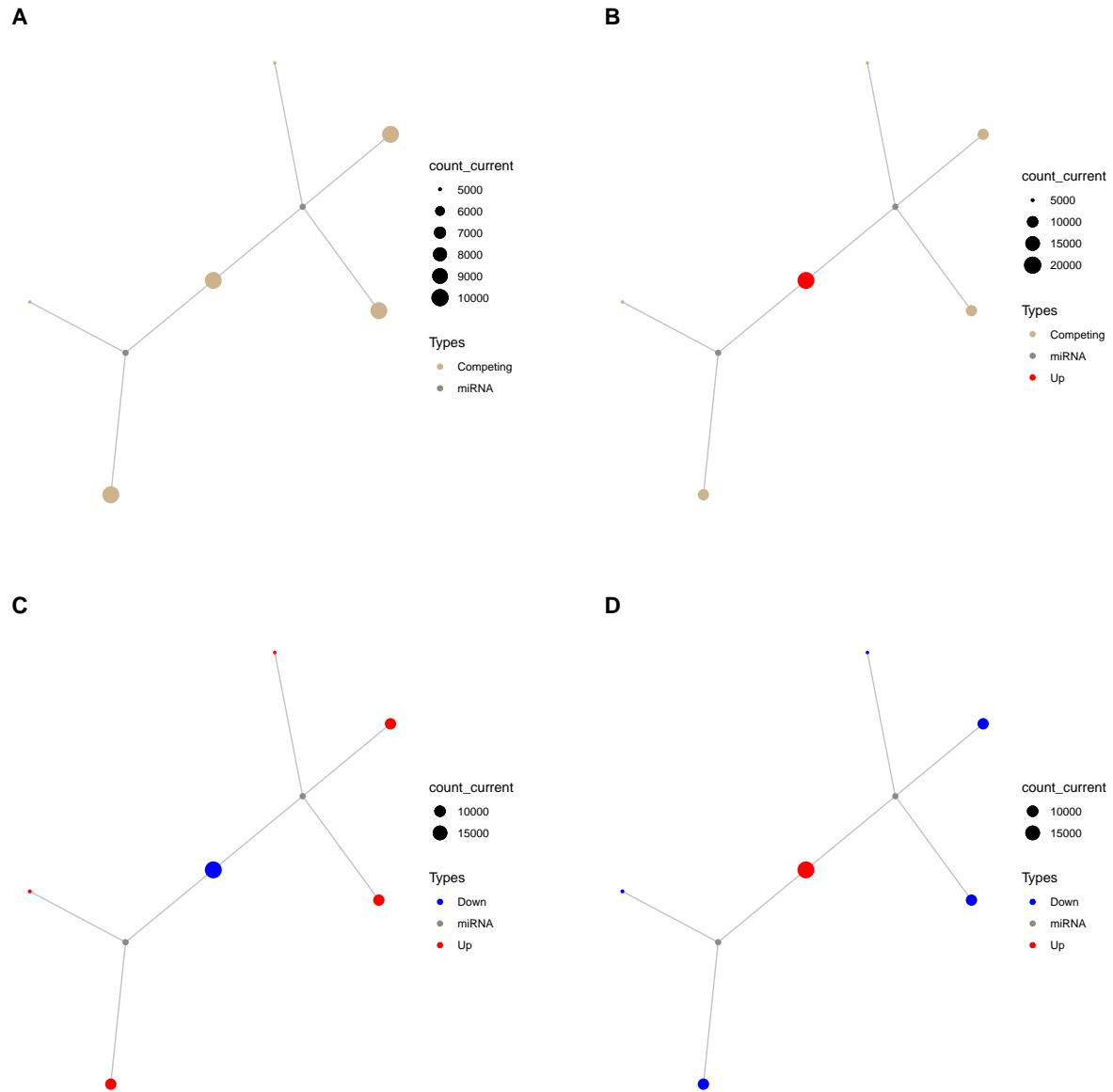


Figure S9: Perturbation of Gene4 on Sample+. A) Network at steady-state. B) Upregulation of Gene4. C) Primary response of network to upregulation of Gene4. D) Re-regulation of whole nodes on system (3th iteration)

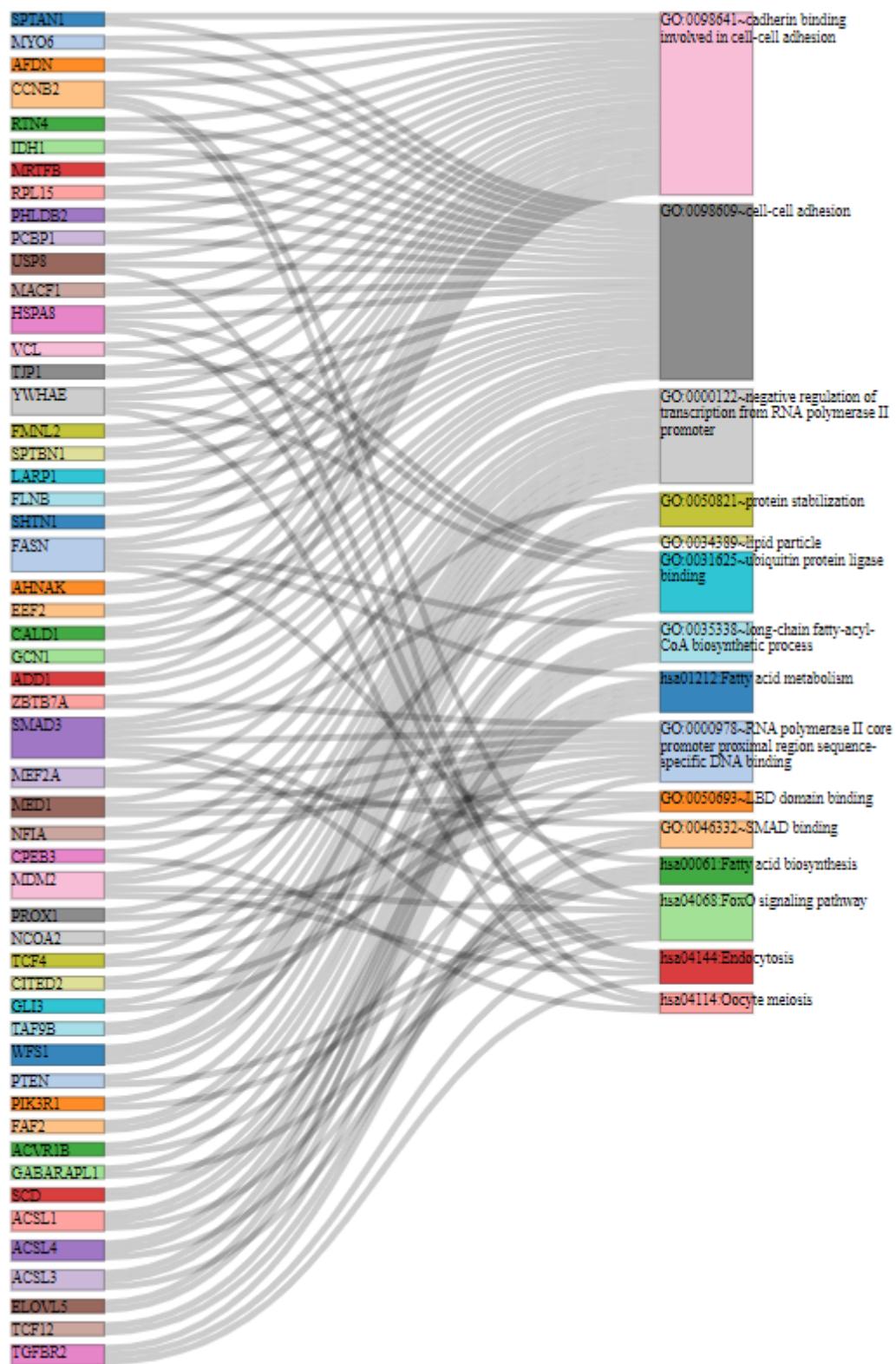


Figure S10: Sankey diagram represents top five KEGG and GO (molecular function and biological process) terms and genes enriched on these terms. Genes with single edge are not shown.

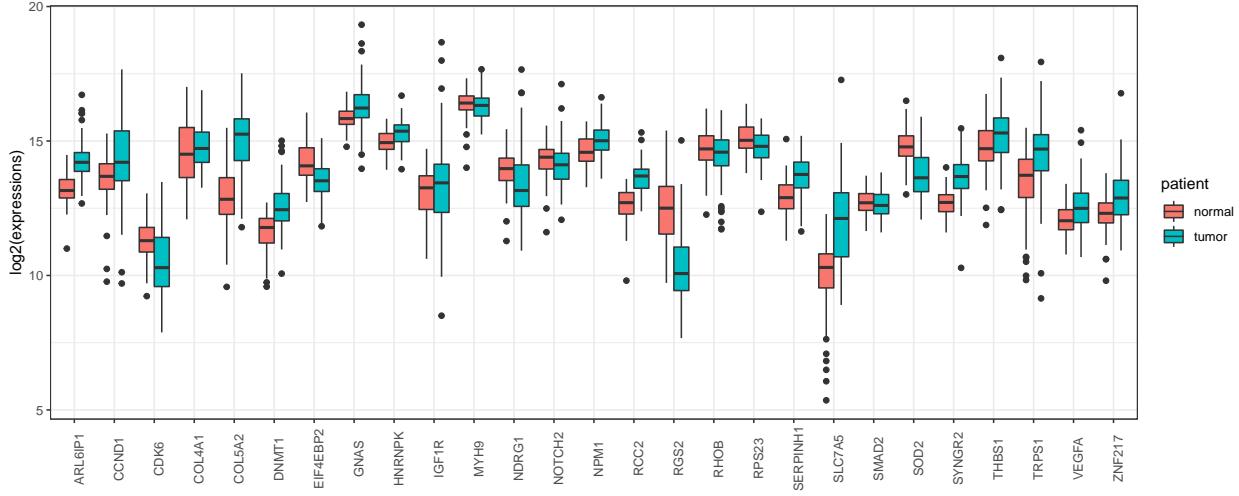


Figure S11: Tumor-specific perturbing nodes log2 transformed expression levels in tumor and normal tissue samples of 87 patients.

### 3 Supplementary Tables

#### 3.1 Sample+ dataset

Table S1: The parameters which affect miRNA:target interactions (i.e. seed type, region, energy) are provided in Sample+ dataset, while these factors are not utilized in simulation of Sample dataset.

Competing	miRNA	Competing Expression	miRNA Expression	Seed Type Effect	Region Effect	Energy
Gene1	Mir1	10000	1000	0.43	0.30	-20
Gene2	Mir1	10000	1000	0.43	0.01	-15
Gene3	Mir1	5000	1000	0.32	0.40	-14
Gene4	Mir1	10000	1000	0.23	0.50	-10
Gene4	Mir2	10000	2000	0.35	0.90	-12
Gene5	Mir2	5000	2000	0.05	0.40	-11
Gene6	Mir2	10000	2000	0.01	0.80	-25

#### 3.2 Significant factors in miRNA:target interactions

Some of information about miRNA:target interactions were exhibited directly by high-throughput studies. On the other hand, we were examined other interacion factors based on different studies.

- (Helwak et al. 2013; Moore et al. 2015) reported the energy values in miRNA:target interactions.
- Comparisons of canonical seed types were evaluated by study of (Grimson et al. 2007), while functional and non-functional seed interactions were studied by (Bartel 2009) and (Betel et al. 2010).
- Numeric definition of target region location effect was performed based on studies of (Hausser et al. 2013) and (Helwak et al. 2013)

Table S2: Efficiency factors for seed types.

seed type	seed type effect
6-mer_noncanonical	0.05
9-mer	0.43
6-mer	0.07
8-mer	0.43
7-mer	0.23
none	0.01
5-mer_noncanonical	0.04
5-mer	0.05
6-merA1_noncanonical	0.05
7-mer-8m_noncanonical	0.21
7-mer-8m	0.25
8-mer_noncanonical	0.35
7-merA1_noncanonical	0.16
7-merA1	0.19
6-merA1	0.07

Table S3: Efficiency factors for binding regions on targets

region	region effect
3UTR	0.84
CDS	0.42
3UTRCDS	0.93
5UTR	0.01
5UTRCDS	0.42
none	0.01
intron	0.01
CDS3UTR	0.93
CDS5UTR	0.42
exon_unclassified	0.20
CDS3UTRintron	0.93
3UTRintron	0.84
CDSintron	0.42
5UTRintron	0.01
5UTR3UTR	0.93
CDS5UTR3UTR	0.93

### 3.3 Content of High-throughput experimental studies

Table S4: Context of miRNA:target pairs supported by High-throughput Experiments. CLEAR-CLIP and CLASH datasets were integrated as described in Section 2 of Supplementary Material and Method.

Variable	Structure	Definition
cluster	character	Barcode from experimentally method
chromosome	character	Chromosome of Target gene from raw data

Variable	Structure	Definition
start_position	numeric	Gene start position from raw data
end_position	numeric	Gene end position from raw data
strand	character	Gene strand
hgnc_symbol	character	Gene name (Symbol)
Ensembl_Gene_Id	character	Ensembl Gene Id of gene
Ensembl_Transcript_Id	character	Ensembl transcript id of mRNA of Target gene
target_seq	character	mRNA sequences targeted by miRNA
miRNA	character	miRNA id (from miRBase version 21 )
miR_seq	character	miRNA sequence
seed_type	character	seed type of miRNA:target interaction
Energy	numeric	Energy of miRNA:target binding
HG38build_loc	character	Recent chromosomal location of Gene
Genome_build	character	Genome build of given chromosome, start and end positions
region	character	interaction location on target
region_effect	numeric	Coefficient of location on target
seed_type_effect	numeric	Coefficient for seed sequence of miRNA:target interaction

### 3.4 Variables of network object during simulation

As a result of simulation a dataset, a graph object is obtained that includes various variables in edge and node data. A graph object includes variables at following.

## 4. Notes

To evaluate significant nodes through parallel processing in breast cancer patient network, perturbation on all nodes were triggered by 3-fold up-regulation with 10 iteration (cycle) due to small diameter of Real and Real+ networks. We analyzed perturbation efficiencies of all nodes in situation that accepts nodes with more than one percent change in expression as re-regulated (i.e, limit=1 argument in *find\_node\_perturbation()* function).

All codes are available at [github repo](#).

Table S5: The context graph object during the process.

Variables	Structure	Description
<i>Node Variables</i>		
name	character	node name
type	character	Competing or miRNA
node_id	numeric	in on graph object
initial_count	numeric	Initial Expression value of node
count_pre	numeric	Expression value of node at previous regulation
count_current	numeric	Existing expression value of node
changes_variable	character	Regulation of node (Up, down or steady)
<i>Edge Variables</i>		
Competing name	character	name of genes
miRNA name	character	name of miRNAs
Competing expression	numeric	Expression values of competing elements at steady-state
miRNA expression	numeric	Expression values of miRNA elements at steady-state
energy	numeric	coefficient of miRNA:target interactions (binding affinity)
seed type	numeric	coefficient of miRNA:target interactions (binding affinity)
region	numeric	coefficient of miRNA:target interactions (degradation efficiency)
afff factor	numeric	coefficient scaled and combined affinity factor
degg factor	numeric	coefficient scaled and combined degradation factor
comp_count_list	list	list of competing expression for each iteration
comp_count	numeric	pre: competing expression at previous iteration; current: competing expression at present iteration
mirna_count_list	list	list of miRNA expression for each iteration
mirna_count	numeric	pre: miRNA expression at previous iteration; current: miRNA expression at present iteration
effect	numeric	pre: total miRNA repressive effect on individual target at previous iteration ; current: miRNA repressive effect on individual target at present iteration
effect_list	list	list of miRNA repressive effect on individual target for each iteration

## REFERENCES

- Bartel, David P. 2009. “MicroRNAs: Target Recognition and Regulatory Functions.” *Cell* 136 (2): 215–33. <https://doi.org/10.1016/j.cell.2009.01.002>.
- Betel, Doron, Anjali Koppal, Phaedra Agius, Chris Sander, and Christina Leslie. 2010. “Comprehensive Modeling of microRNA Targets Predicts Functional Non-Conserved and Non-Canonical Sites.” *Genome Biology* 11 (8): R90.
- Grimson, Andrew, Kyle Kai-How Farh, Wendy K. Johnston, Philip Garrett-Engele, Lee P. Lim, and David P. Bartel. 2007. “MicroRNA Targeting Specificity in Mammals: Determinants Beyond Seed Pairing.” *Molecular Cell* 27 (1): 91–105. <https://doi.org/10.1016/j.molcel.2007.06.017>.
- Hausser, J., A. P. Syed, B. Bilen, and M. Zavolan. 2013. “Analysis of CDS-Located miRNA Target Sites Suggests That They Can Effectively Inhibit Translation.” *Genome Research* 23 (4): 604–15. <https://doi.org/10.1101/gr.139758.112>.
- Helwak, Aleksandra, Grzegorz Kudla, Tatiana Dudnakova, and David Tollervey. 2013. “Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding.” *Cell* 153 (3): 654–65. <https://doi.org/10.1016/j.cell.2013.03.043>.
- Moore, Michael J., Troels K. H. Scheel, Joseph M. Luna, Christopher Y. Park, John J. Fak, Eiko Nishiuchi, Charles M. Rice, and Robert B. Darnell. 2015. “miRNA-Target Chimeras Reveal miRNA 3’-End Pairing

as a Major Determinant of Argonaute Target Specificity.” *Nature Communications* 6 (November): 8864.  
<https://doi.org/10.1038/ncomms9864>.