Computational developments in microRNA-regulated protein-protein interactions and pathways

Professor Phoebe Chen, PhD
Department of Computer Science and Information Technology,
La Trobe University,
Melbourne, Australia

http://homepage.cs.latrobe.edu.au/ypchen/

LA TROBE UNIVERSITY AUSTRALIA

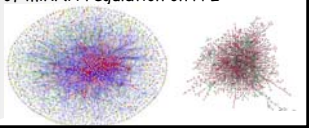
Contents

Brief introduction to MicroRNA and protein-protein interaction networks

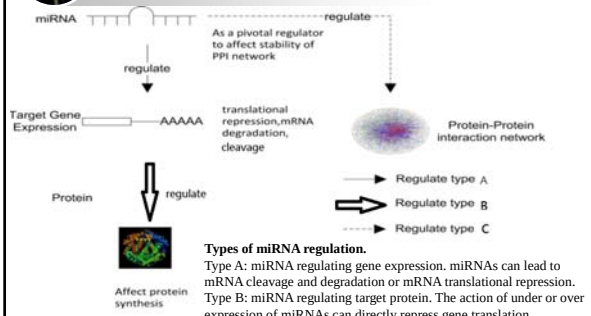
- MicroRNAs (miRNAs) are a subset of small (~22 nucleotide – structural units of RNA and DNA - in length) non-coding RNA molecules, which comprise 1% of genes in animal genomes.
- The interaction between miRNA and mRNA provides a new way to determine gene functions. In studies on miRNA-mRNA interaction, how to accurately find the target mRNA genes for miRNAs is the most important question.
- In recent years, researchers have studied miRNA-regulated networks, including miRNA co-regulated networks, miRNA-mRNA networks, miRNA-transcription factor networks and miRNA-protein interaction networks.

Brief introduction to MicroRNA and protein-protein interaction networks

- Protein-Protein Interaction (PPI) is one of the most important tasks required for a living cell to carry out its biological functions such as DNA replication, transcription, translation, signal transduction.
- Research on miRNA-regulated PPI networks can be divided into two main areas:
  a) basic studies on the correlation between miRNAs and general PPI networks.
  b) identification of the impact of miRNA regulation on PPI networks in diseases.

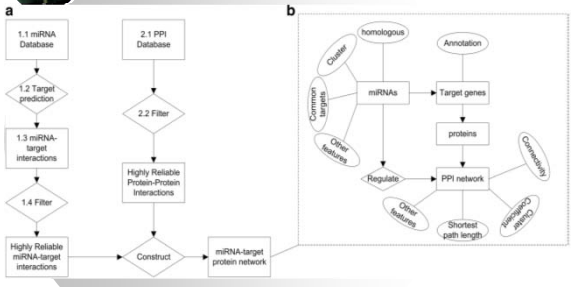


Understanding the role of miRNA in regulating target genes and proteins



Types of miRNA regulation.
Type A: miRNA regulating gene expression. miRNAs can lead to mRNA cleavage and degradation or mRNA translational repression.
Type B: miRNA regulating target protein. The action of under or over expression of miRNAs can directly repress gene translation.
Type C: miRNA regulating on PPI networks. miRNA shows as an indirect regulator to affect dynamic PPI network stability.

Resources to construct a reliable miRNA-protein network



Constructing a miRNA-target protein network.
a) Process to construct a miRNA-target protein network
b) Relationships within miRNA-target protein network and its features.

Resources to construct a reliable miRNA-protein network - database

- The construction of a miRNA-protein network using highly reliable resources is important for the commencement of a miRNA-regulated PPI network study.
- A miRNA–target protein network is constructed by miRNA-target interactions and PPI network data. During the process, the selection of miRNA target predicting approaches and filtering approaches is important to obtain highly reliable data.

Name	Process	Main feature	URL	Ref.
BioGrid	2.1	Protein-protein interaction database	http://thebiogrid.org/	[46]
Cytoscape	Platform	Social or molecular networks analysis and visualization.	http://www.cytoscape.org/	[47]
Database of Interacting Proteins(DIP)	2.1	Protein-protein interaction	http://dip.doe-mbi.ucla.edu/dip/	[48]
Ingenuity system	Platform	signalling and metabolic pathways analysis; molecular network Analysis etc.	http://www.ingenuity.com/	
Human Protein Reference Database(HPRD)	2.1	protein-protein interaction	http://www.hprd.org/	[49]
String	2.1	protein-protein interaction	http://string-db.org/	[50]
The MIPS Mammalian Protein-Protein Interaction Database	2.1	protein-protein interaction	http://mips.helmholtz-muenchen.de/proj/ppi/	[51]
Protein Interaction Network Analysis (PINA)	2.2	PPI network construction, filtering, analysis, visualization and management	http://cbg.garvan.unsw.edu.au/pina/	[52]
HitPredict	2.1	integrated PPI database	http://hintdb.hgc.jp/htp/	[53]
iRefIndex	2.1	integrated PPI database	http://irefindex.uio.no/wiki/iRefIndex	[54]

Resources to construct a reliable miRNA-target protein network

Name	Process	Main feature	URL	Ref.
SynechoNET	1.1	integrated PPI database	http://bioportal.kobic.re.kr/SynechoNET/	[55]
PMRD	1.1	plant miRNA database	http://bioinformatics.cau.edu.cn/PMRD/	[56]
Gene Ontology	Platform	gene annotation, develop controlled vocabulary of genes	http://www.geneontology.org/	[57]
MiRTarBase	Platform	miRNA-target interactions	http://mirtarbase.mbc.nctu.edu.tw/	[58]
PicTar	1.2	miRNA target prediction	http://pictar.mdc-berlin.de/	[59]
RNAhybrid	1.2	miRNA target prediction	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/	[60]
TargetScan	1.2	miRNA target prediction	http://www.targetscan.org/	[61]
GeneSet-2miRNA	1.2	miRNA target predicting with mRNA expression profile	http://mips.helmholtz-muenchen.de/proj/gene2mir/	[62]
MMIA	1.2	miRNA target predicting with mRNA expression profile	http://129.79.244.122/~MMIA/	[63]
miRanda	1.2	miRNA target predicting & miRNA expression profiles	http://www.microrna.org/	[64]
MiRTif	1.4	miRNA target interaction filter	http://mirtif.bii.a-star.edu.sg/	[65]
miRBase	1.1	miRNA sequences and annotations	http://www.mirbase.org/	[66]
The human microRNA disease database	1.1	miRNA sequences and annotations	http://202.38.126.151/hmdd/mirna/md/	[49]
miRExpress	1.1	extract miRNA expression profiles based on HTS results	http://mirexpress.mbc.nctu.edu.tw/	[67]
TarBase	1.2	experimental supported miRNA target	http://diana.cslab.ece.ntua.gr/tarbase/	[68]
miRDeep	1.1	detect novel miRNA based on HTS	http://www.mdc-berlin.de/en/research/research_teams/systems_biology_of_gene_regulatory_elements/projects/miRDeep/index.html	[69]
miRTools	1.1	detect novel miRNA based on HTS	http://centre.bioinformatics.zj.cn/mirtools/	[70]
starBase	1.3	decoding microRNA-target and protein-RNA interaction	http://starbase.sysu.edu.cn/	[71]
IPA	1.4	comprehensive software on biological analysis. Support miRNA target filtering	http://www.ingenuity.com/products/training.html	

Resources to construct a reliable miRNA-protein network

miRNA target identification

- Current experimental approaches to miRNA target identification mainly focus on the use of large scale mRNA expression profiling.
- The common way to uncover miRNA targets is to directly test miRNA expression levels on different mRNA profiling or to use different phenotypes to test expression levels based on microarray .
- The most frequently used applications are TargetScan, PicTar, miRanda, RNAHybrid [60] etc.
- The combination of computational approaches with mRNA expression profiles have proved efficient in recent years, and it has been shown that they can effectively minimize the false positives of miRNA target prediction.

miRNA target filtering

- Considering conservation in strains combined with performing seed matches or evaluating the accessibility of binding sites can facilitate miRNA target predictions.
- Recent studies show that combining both conservation and accessibility can achieve better results in filtering miRNA targets.

Features of miRNA-regulated PPI networks

- The features of PPI networks are defined with a focus on topological characteristics, such as connectivity, cluster coefficients, shortest path length and so on. PPI network features are commonly used to predict unannotated protein functions combined with Gene Ontology.
- In miRNA-regulated PPI networks, miRNAs are classified by several properties:
  a) the miRNA family or homologous miRNAs, which denote the miRNA group whose conserved seed regions are common;
  b) clustered miRNAs, which are miRNAs whose pairwise chromosomal distances are no more than 3000nt;
  c) miRNAs with common targets. miRNAs can directly or indirectly down-regulate 100–200 genes on average , and each gene can be targeted by multi-miRNAs.

Current findings in the study of miRNA-regulated PPI networks

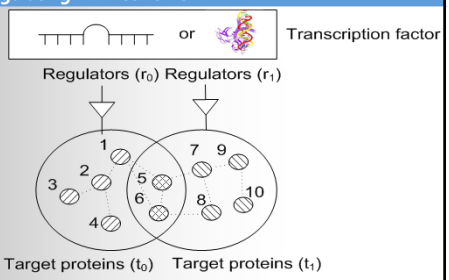
Computational Developments in the Study of miRNA-regulated PPI networks

Current findings in the study of miRNA-regulated PPI networks

Research area	Description
Correlation between protein connectivity and miRNA regulation complexity	A. There is positive correlation between miRNA target site types and its regulated protein connectivity. B. MiRNA target propensity may be due to high protein connectivity. C. MiRNA regulation propensity changes due to different hub proteins [72].
	miRNA targeted proteins have short distance and higher modularity than randomly selected proteins [73].
miRNA-regulated specific proteins in PPI networks	A. MiRNAs that target a lower number genes have the propensity to regulate commonly expressed proteins rather than tissue-specific proteins. B. Commonly expressed proteins and tissue-specific proteins are always regulated together by a miRNA, and the numbers of protein expressed are close in both proteins [44].
The coordination role of MicroRNAs: miRNA clusters regulate PPI networks	miRNAs in the same clusters have the tendency to coordinate to regulate protein functions in protein-protein interaction networks [28].
The coordination role of MicroRNAs: miRNAs coordinate to regulate protein complex	A. MiRNAs coordinate to regulate protein complexes in posttranscriptional level. B. Correlations between the proteins exist in the same complex regulated by miRNAs[25].
The coordination role of MicroRNAs: miRNA crossingtalking with transcription factors	Crosstalk motifs between miRNAs and transcription factors motif demonstrate higher network properties in miRNA-regulated PPI networks [7]
Identifying miRNA-regulated PPI networks in special diseases	A. In gastric cancer [74]: six miRNA-regulated protein networks are identified in gastric cancer based on the human PPI network; it is suggested that miR-148a may resist tumor extension. B. In human ovarian cancer [77]: six miRNAs (hsa-mir-20a, hsa-mir-24-2, hsa-mir-34a, hsa-mir-21, hsamir-17 and hsa-mir-hsa-mir-155) and six TFs (BRCA1, SP1, ESR1, SMAD3, POZF1 and TFE2) play key roles in ovarian cancer progression. C. In aging-related diseases [78]: 35 genes related to diseases associated with aging were identified.

The coordination role of miRNAs in regulating PPI networks

- clustered miRNAs may coordinate to regulate the PPI network.
- miRNAs coordinate to regulate protein complex.
- miRNAs may coordinate with other regulators such as transcription factors (TFs).



Regulation motifs of miRNA and transcription factors

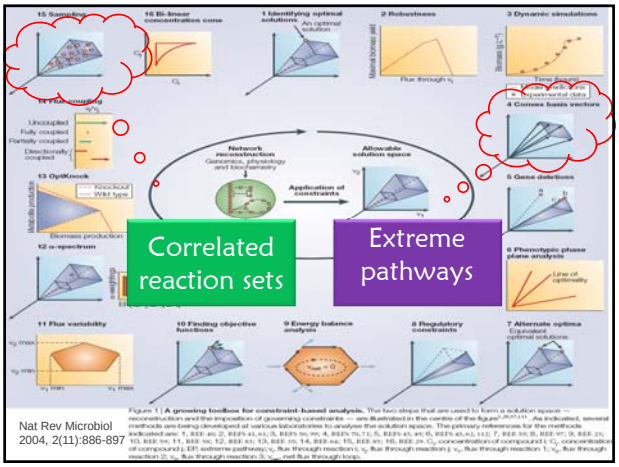
Four Regulator Motifs: single regulation motif = {r_0, 1,2,3,4,5,6} or {r_1, 5,6,7,8,9,10}; co-regulation motif = {r_0, r_1, 5,6} crosstalk motif = {r_0, r_1,1,2,3,4,7,8,9,10}; If {5} or {6} does not exist, independent motif = {r_0, 1,2,3,4,5,6} and { r_1, 5,6,7,8,9,10}

Conclusion

- Proteins in PPI networks are different in miRNA regulation. miRNAs have the propensity to regulate higher connected proteins.
- Additionally, the target proteins have shorter and higher modularity than other random proteins. Especially for hub proteins, miRNAs seem to play more important roles in higher cluster coefficient proteins than lower ones.
- For proteins in special tissues, miRNAs that target a lower number of genes have the propensity to regulate commonly expressed proteins rather than tissue-specific proteins. Additionally, protein complexes tend to be regulated by miRNAs;
- miRNAs may cooperate to regulate target proteins with others. The collaborator can be other miRNAs in a cluster or it can also be TFs. Interacting proteins have the tendency to be regulated by miRNAs in the same cluster.
- On the other hand, the crosstalk between miRNAs and TFs plays a more important role in regulating proteins than other types, such as single-regulation, co-regulation and independent regulation.
- Finally, through analysing miRNA-regulated PPI networks in cancer, several key miRNAs associated with cancers can be identified.
- These findings suggest that by combining miRNAs and PPI network features, it is possible to obtain useful information or rules on both miRNA regulation and protein interaction.

Analysis on relationship between extreme pathways and correlated reaction sets

- A framework for genome-scale metabolic network modeling
- Concepts and Methods
  - Extreme pathway(ExPa)
  - Correlated reaction set(CoSet)
  - Relationship between ExPas and CoSets
- Results
- Conclusion



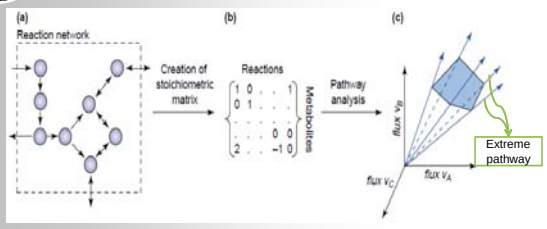
Nat Rev Microbiol 2004, 2(11):886-897

Extreme Pathway(1)

- The internal reversible reactions are decoupled to two separate reactions for the forward and reverse direction separately. Each internal flux can take non-negative value only.
- All available flux vectors lie in a convex polyhedral cone in a high-dimensional space.
- The edges of this convex polyhedral cone make up the set of extreme pathways($\mathbf{p}^i, i = 1, \cdots, k$).

$$v = \sum_{i=1}^{k} \alpha_i \mathbf{p}^i, \ \alpha_i \geq 0, \forall i$$

Extreme pathway(2)



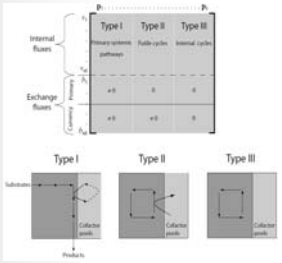
[J.A. Papin, N.D. Price, S.J. Wiback, D.A. Fell, and B.O. Palsson : **Metabolic pathways in the post-genome era.** *Trends in Biochemical Science*, 28:250-258, 2003.]

Extreme pathway(3)

- Properties of extreme pathways:
  - The extreme pathway set of a given network is unique.
  - Each extreme pathway uses least reactions to be a functional unit.
  - The extreme pathway set is systemically independent.
  - The number of extreme pathways grows exponentially with the size of metabolic network.

Extreme pathway(4)

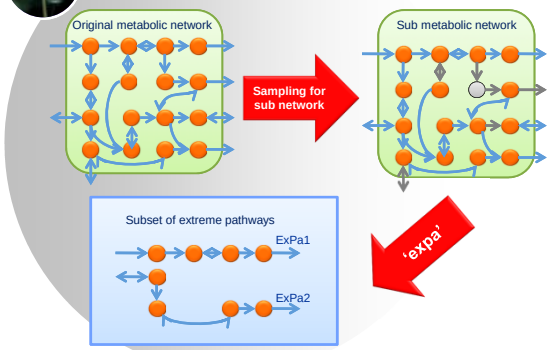
- Classification of extreme pathways
  - Type I
  - Type II
  - Type III: neglected (thermodynamically infeasible)



[B.O. Palsson. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, 1st edition, 2006]

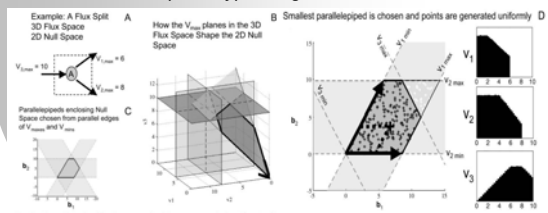
Extreme pathway(5)

- Computation of extreme pathways
  - Network with relatively smaller size: 'expa', an open source tool [ref: Bell SL, Palsson BO: **expa: a program for calculating extreme pathways in biochemical reaction networks**. *Bioinformatics* 2005, 21(8):1739-1740]
  - Network with large size: Sampling + 'expa'
    - Sampling: A sub network generated by randomly deleting a few reactions in the original metabolic network
    - 'expa': Enumerating all the extreme pathways of the sub network.



Correlated reaction set(1)

- Uniform random sampling
  - A rapid and scalable method to quantitatively characterize all allowable phenotypes of genome-scale networks.



[N.D. Price, J. Schellenberger, and Palsson BO. **Uniform sampling of steady state flux spaces: Means to design experiments and to interpret enzymopathies.** *Biophysical Journal*, 87:2172-2186, 2004]

Correlated reaction set(2)

- Correlated reaction set can be defined based on the pair wise correlation coefficients between all reaction fluxes of the sample phenotypes.
- Correlated reaction sets are unbiased, condition-dependent definitions of modules.
- Reactions of the same correlated reaction set have to be co-utilized in precise ratios.
- Correlated reaction sets provide clues about regulated procedures of metabolic networks.

Correlated reaction set(3)

- Computation of correlated reaction set:
  - Tool: COBRA Toolbox.[ref: Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.** *Nat. Protoc.* 2007,2(3):727-738]
  - Condition: optimum growth
  - 1000,000 unique sample flux distributions
  - 10,000 samples are randomly selected to measure correlation coefficients
  - Threshold: $1\text{-}1e^{-8}$
  - Repeating 20 times. The results are stable.

Relationship between ExPa & CoSet

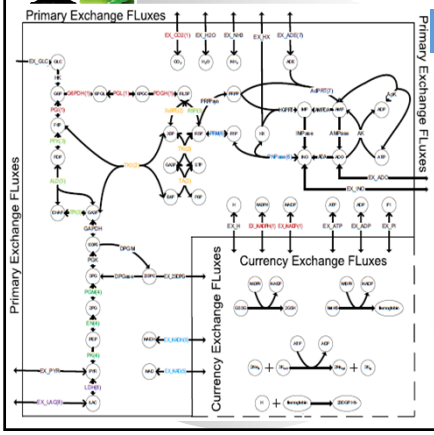
- Extreme pathways and correlated reaction sets are both determined by the topology of a metabolic network.
- What is the relationship between them?
  - How many extreme pathways that use k reactions of a CoSet?
  - What about the CoSets coverage rate of ExPas?

$$\text{CoSets Coverage Rate} = \frac{\sum_{\text{CoSets } C \text{ that fully covered by ExPa } \mathbf{P}} \text{number of reactions in CoSet } C}{\text{number of reactions on ExPa } \mathbf{P}}$$

Result

- Results on three different metabolic networks
  - *Escherichia coli* core metabolic network
  - Human red blood cell (RBC) metabolic network
  - *Saccharomyce cerevisiae* metabolic network

RBC metabolic network(1)



- 39 metabolites, 32 internal metabolic reactions, 19 exchange fluxes
- Method for ExPas computation: 'expa'
- ExPas: 55 (type I,II: 39, type III: 16)
- CoSets: 8

RBC metabolic network(2)

Co-Set ID	Co-Set Size	Number of ExPas using k reactions of a CoSet						
		0	1	2	3	4	5	6	7
1	7	18	6	0	0	0	0	9	6
2	4	21	0	0	0	18	-	-	-
3	4	18	6	0	6	9	-	-	-
4	3	27	0	0	12	-	-	-	-
5	2	19	0	20	-	-	-	-	-
6	2	24	0	15	-	-	-	-	-
7	2	30	0	9	-	-	-	-	-
8	2	37	0	2	-	-	-	-	-

Table 4 - Relationship between ExPas and CoSets for RBC metabolic network.

CoSet ID	CoSets Size	Reactions
1	7	PDGH, Ex_CO2, Ex_NADPH, PGI, PGL, G6PDH, Ex_NADP
2	4	Xu5PE, TKI, TKII, TA
3	4	PFK, ALD, TPI, R5PI
4	3	PGM, EN, PK
5	2	Ex_NAD, Ex_NADH
6	2	PNPase, PRM
7	2	AdPRT, Ex_ADE
8	2	LDH, Ex_LAC

Table 3 - CoSets of RBC metabolic network .

RBC metabolic network(3)

- The CoSets of RBC metabolic network show agreement with its regulatory structure.

CoSet ID	Regulated reactions	Notes
1	G6PDH, PFGH	Balance the ratio: ADPH/NADP
2	TKI, TA, TKII	-
3	RPI,PFK	-
4	EN, PK	-
5	-	Balance the ratio: NAD/NADH
6	-	Indirectly regulated
7	AdPRT	-
8	-	Trivial CoSet

RBC metabolic network(4)

- Each CoSet is covered by an ExPa in an 'all or none' manner, except the CoSets 1 and 3.
- CoSets 1 and 3: 'all or none' + 'one or all but one'
  - CoSet 1: PGI belongs to EMP while other reactions are in PPP.
  - CoSet 3: R5PI belongs to PPP while other reactions are in EMP.
  - PPP and EMP share metabolite 'G6P' and 'GA3P' and have to work together so that the cell could arrive optimal growth.
  - Splitting brings redundancy and flexibility.

CoSet ID	Expas using only one reaction in the CoSet	Expas using n-1 reactions in the coSet	Notes
1	34, 35, 36, 37, 38, 39	1, 2, 3, 19, 20, 21, 22, 23, 24	Reaction: 'PGI'
3	1, 2, 3, 25, 26, 27	19, 20, 21, 34, 35, 36	Reaction: 'R5PI'

RBC metabolic network(5)

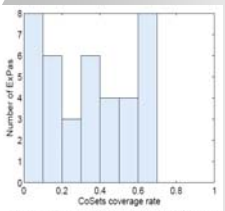


Figure 3 - CoSets coverage rate of ExPas of RBC metabolic network

- Nearly 1/3 ExPas of RBC model has a CoSets coverage rate higher than 20%.
- Seven ExPas' CoSets coverage rate is 0 probably because they are relatively short (1-3 internal reactions as well as the corresponding exchange reactions).
- ExPas are under control of the corresponding CoSets.

Conclusion

- Extreme pathways show strong complementary relationship on the usage of reactions in the same CoSets.
- The relationship may result from the topological constraints composed on metabolic networks.
- The strong relationship suggests a possible mechanism of how a metabolic network transferring its phenotype from one steady state to another.

All papers can be found at:
http://homepage.cs.latrobe.edu.au/ypchen/index.htm

LA TROBE UNIVERSITY AUSTRALIA