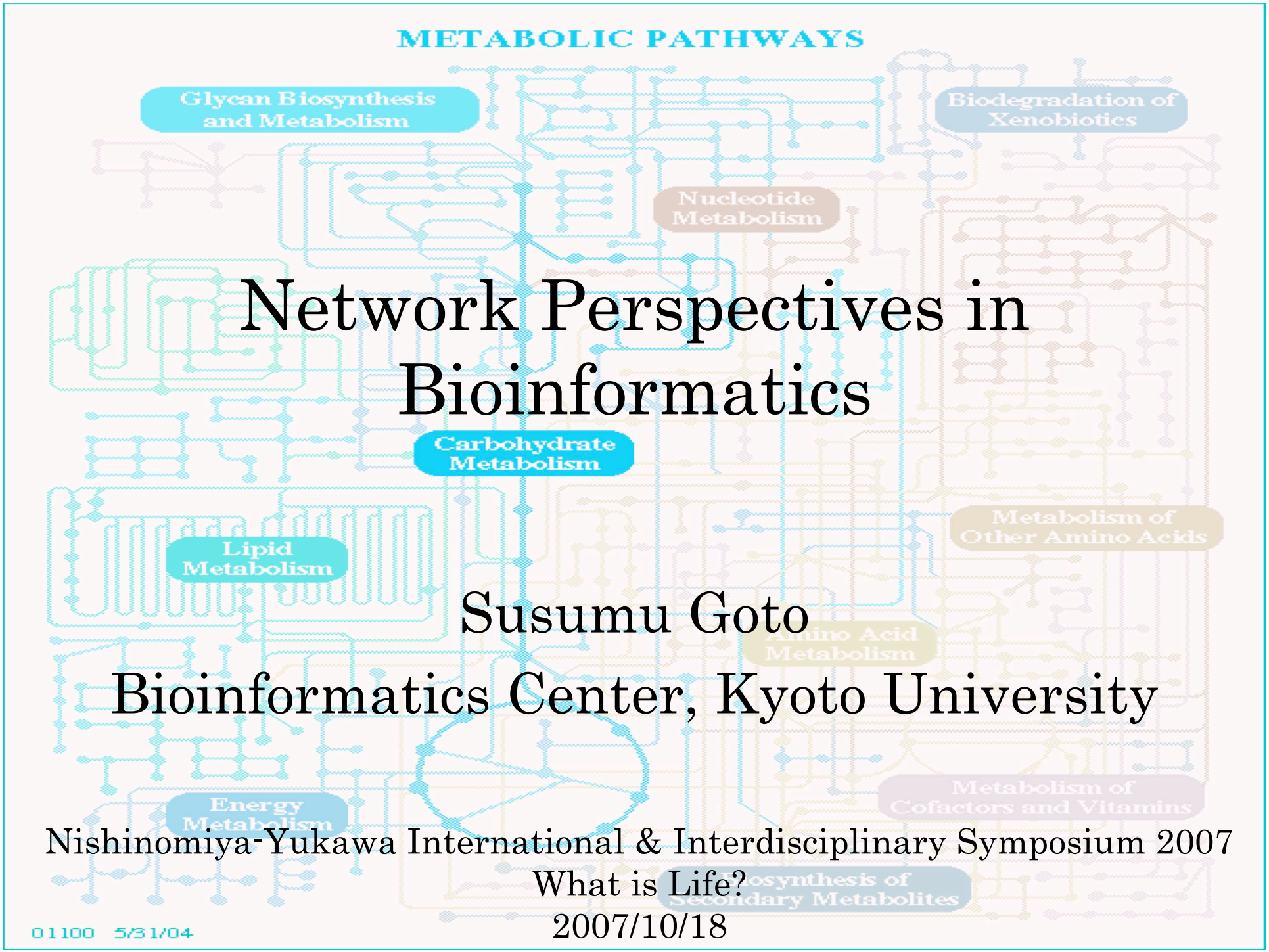


Network Perspectives in Bioinformatics

Susumu Goto

Bioinformatics Center, Kyoto University

Nishinomiya-Yukawa International & Interdisciplinary Symposium 2007
What is Life?
2007/10/18

Bioinformatics

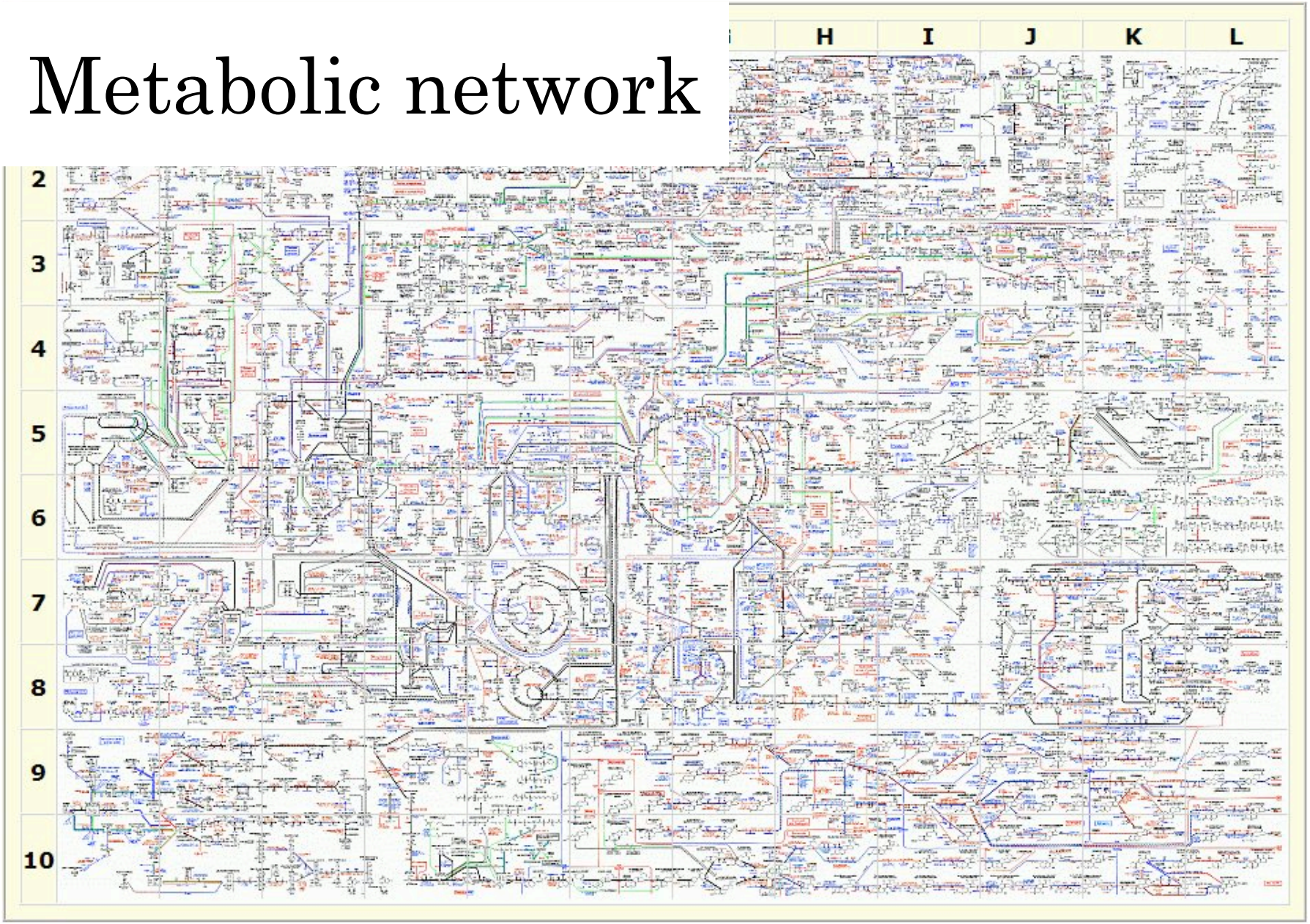
- Large amount of omics data
  - Genomics, Transcriptomics, Proteomics, Interactomics, Metabolomics, etc.
- Development of various techniques for interpreting the data and solving biological problem, mainly on the molecular level

What is Life?

- Autonomy?
- Interaction with others?

- Uptake of nutrients outside
  - Metabolism
- Sensing signals from outside
  - Signal transduction
- Replication

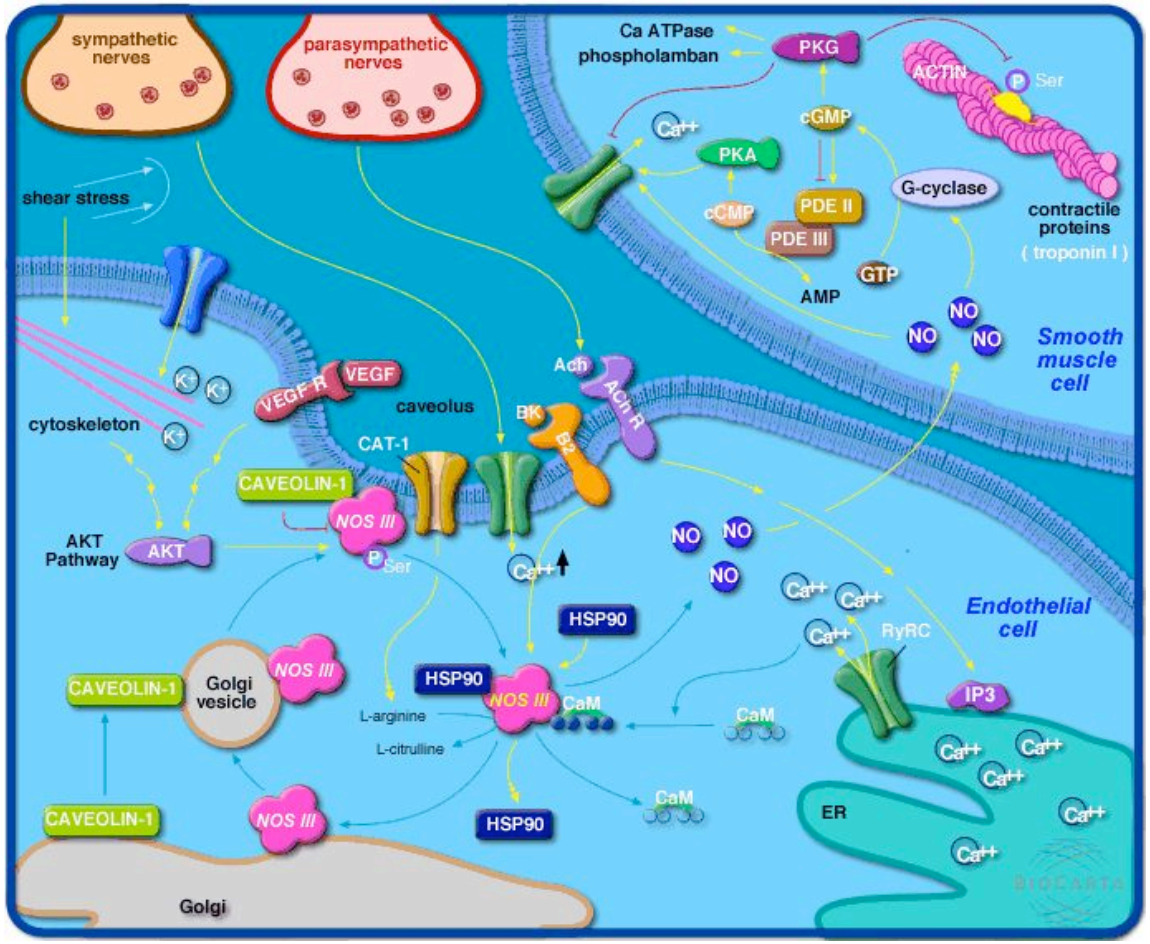
Metabolic network



http://www.expasy.org/cgi-bin/show_thumbnails.pl

Signaling network

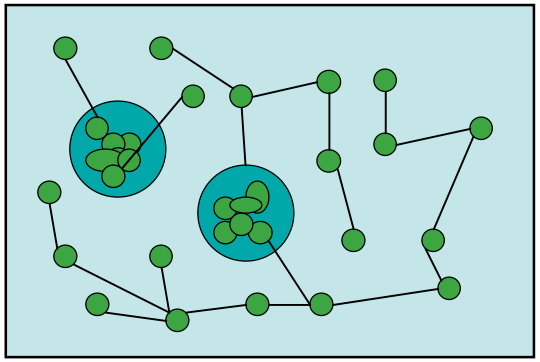
Protein interaction network working in response to the signals from outside and within cells



http://www.biocarta.com/pathfiles/h_no1Pathway.asp

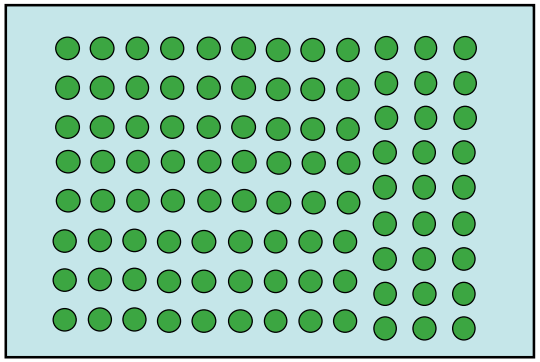
Network for understanding biological systems

Biological system as a network of genes or gene products



Reductive approach

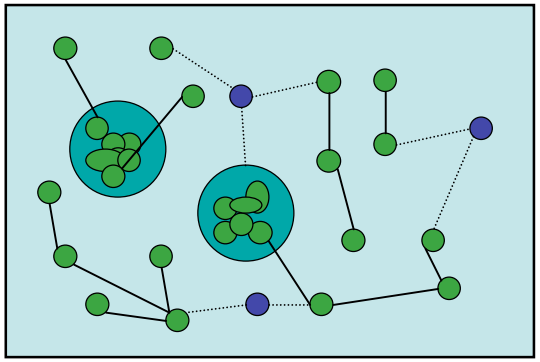
Synthetic approach
(Bioinformatics)



Gene catalogs from genome projects

Current Status of Functional Reconstruction

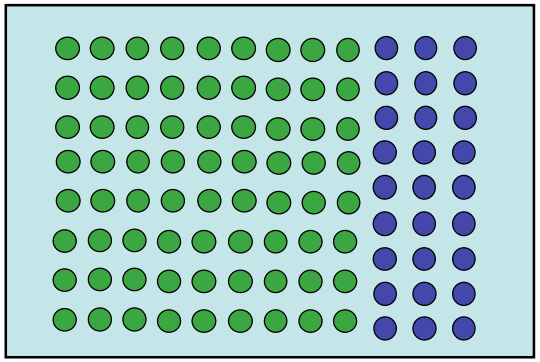
Biological system as a network of genes or gene products



Missing enzymes in metabolic pathway

Reductive approach

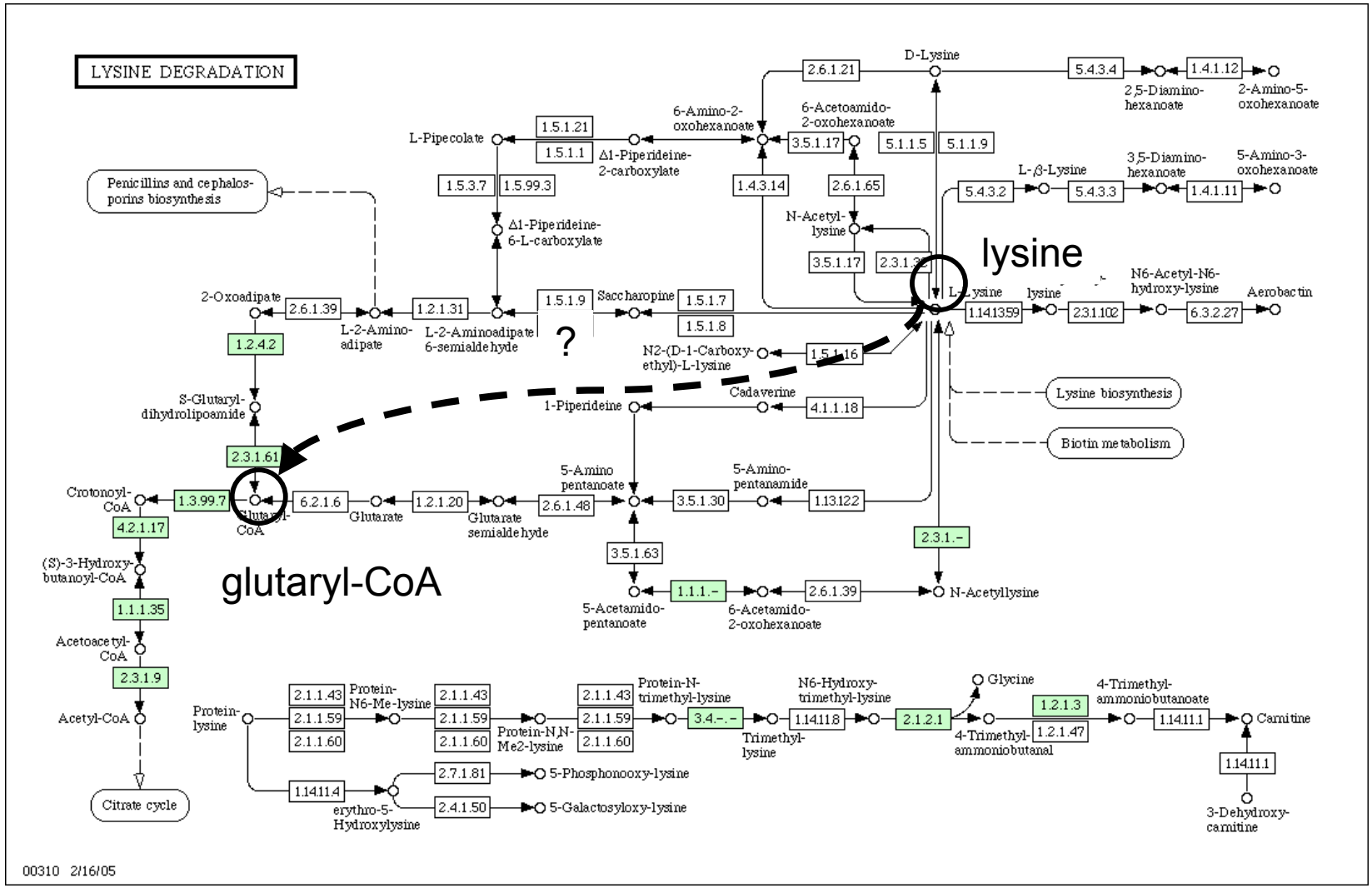
Synthetic approach (Bioinformatics)



Genes with unknown function

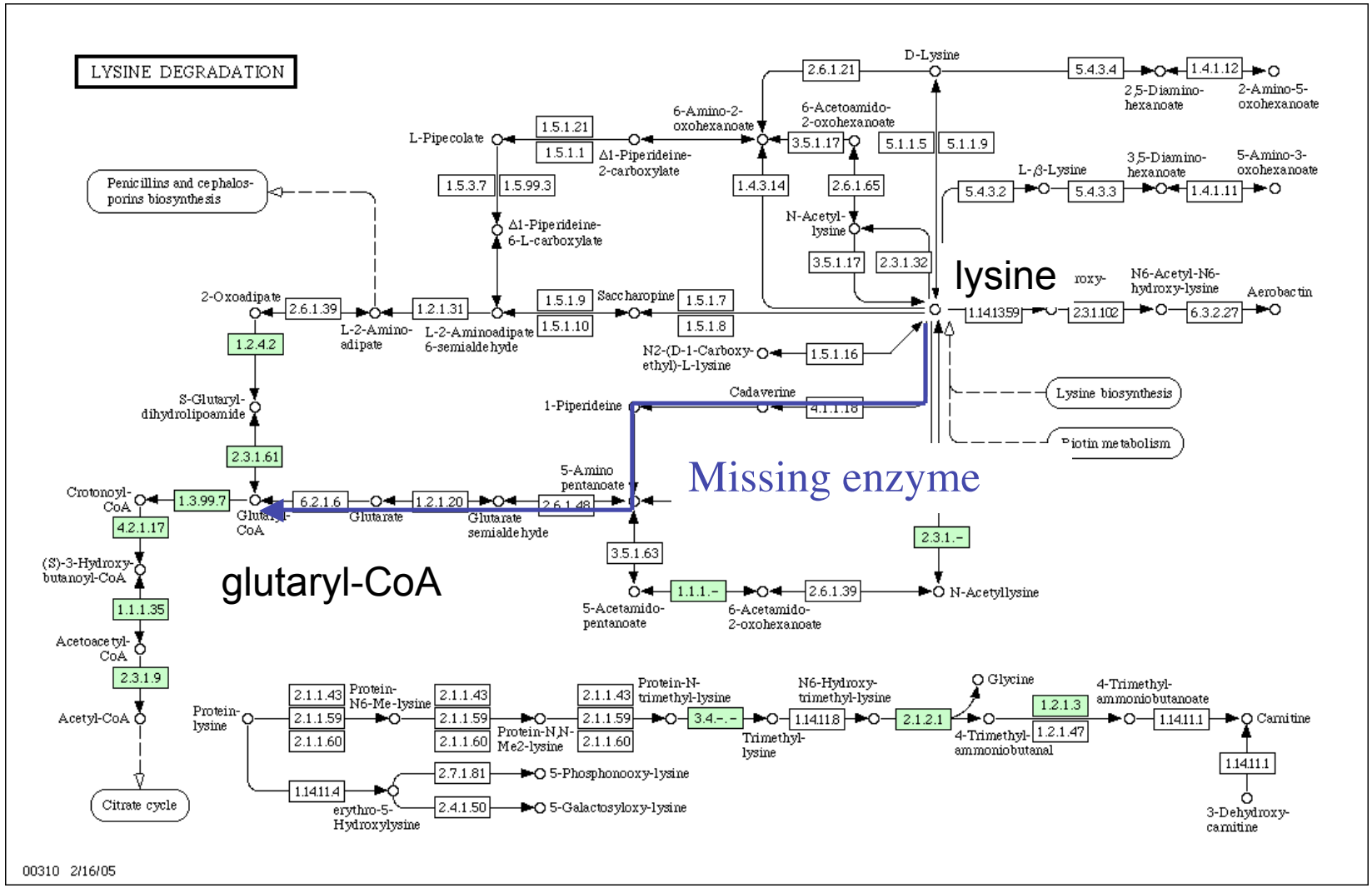
Gene catalogs from genome projects

More specific example



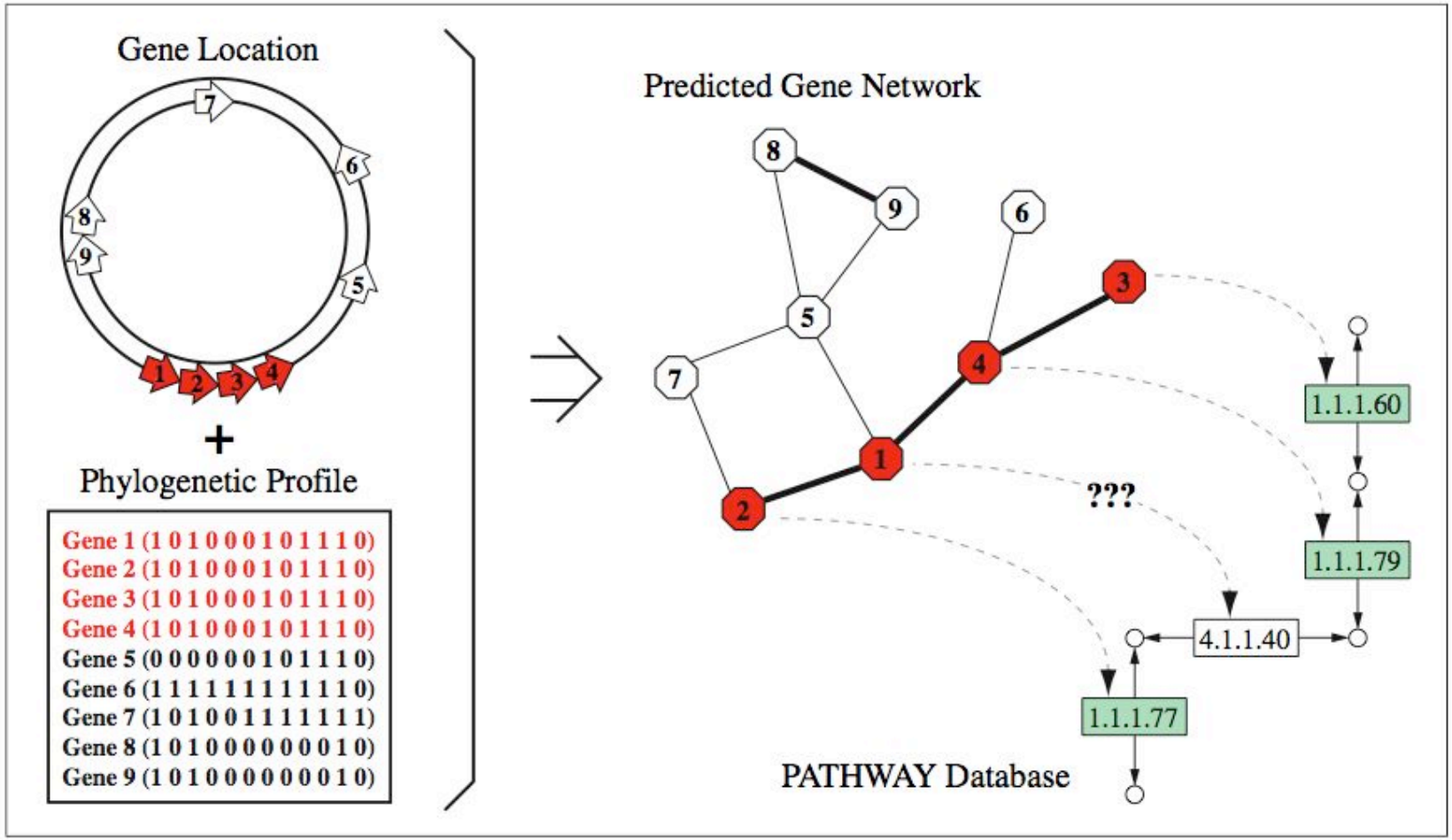
Lysine degradation pathway of *Pseudomonas aeruginosa*
reconstructed from genomic information

More specific example



Lysine degradation pathway of *Pseudomonas aeruginosa* by biochemical knowledge

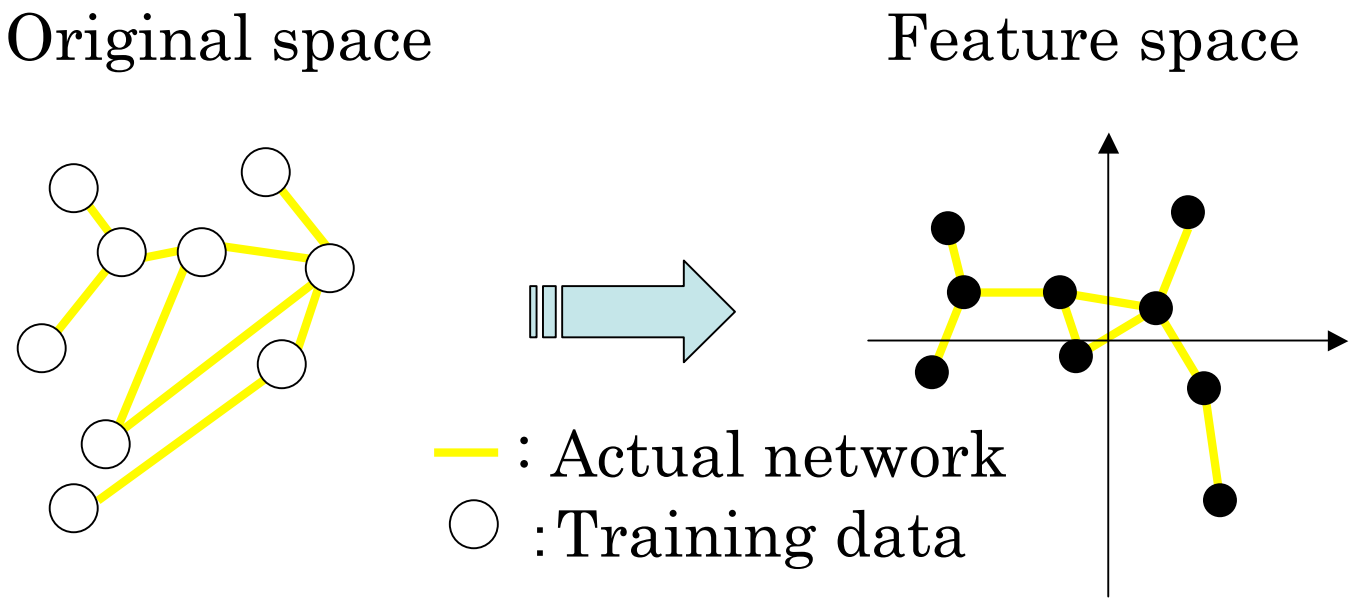
Prediction of missing enzyme genes



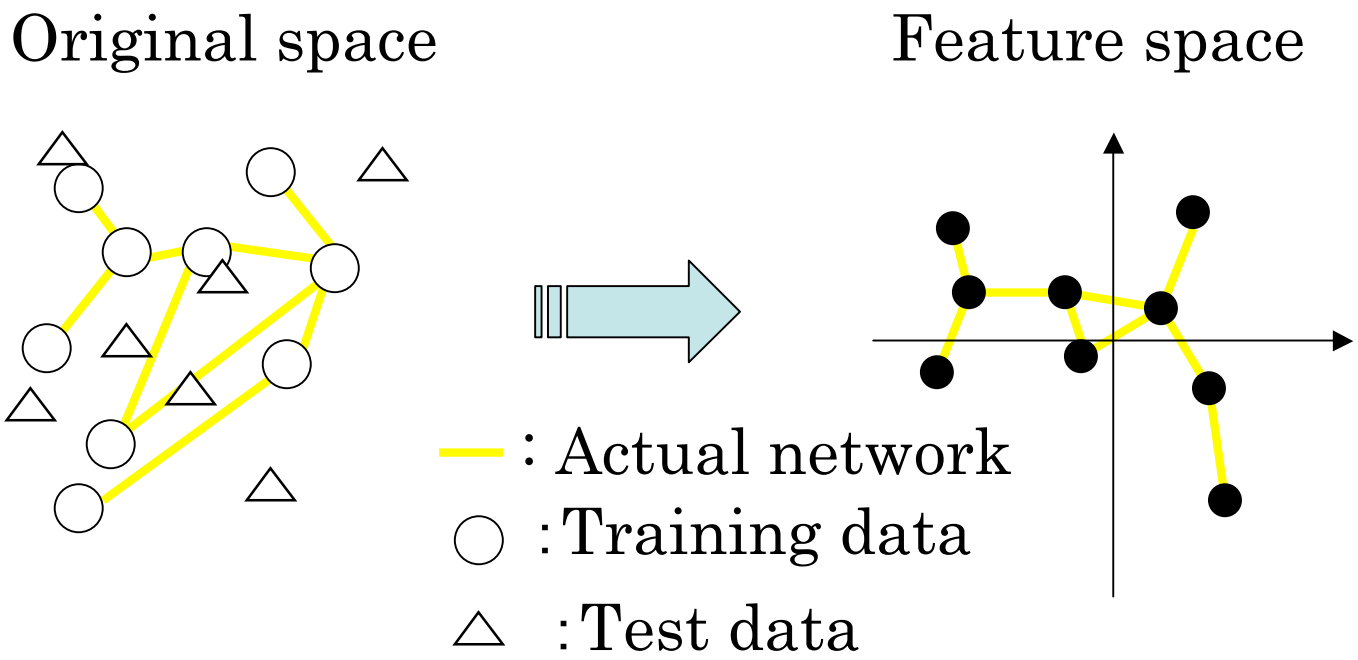
Yamanishi, et al. (2007) FEBS J. 274:2262

Supervised approach 1/2

Step 1: Project the data into the feature space that explains the actual network

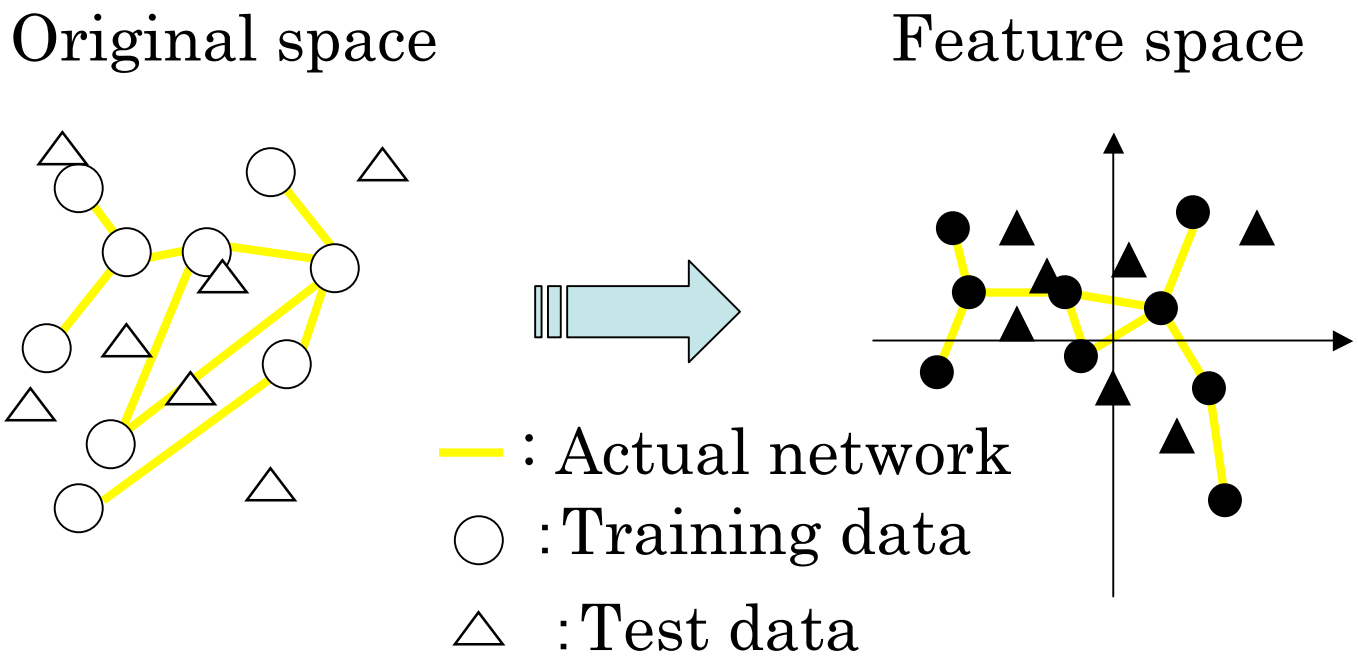


Supervised approach 2/2



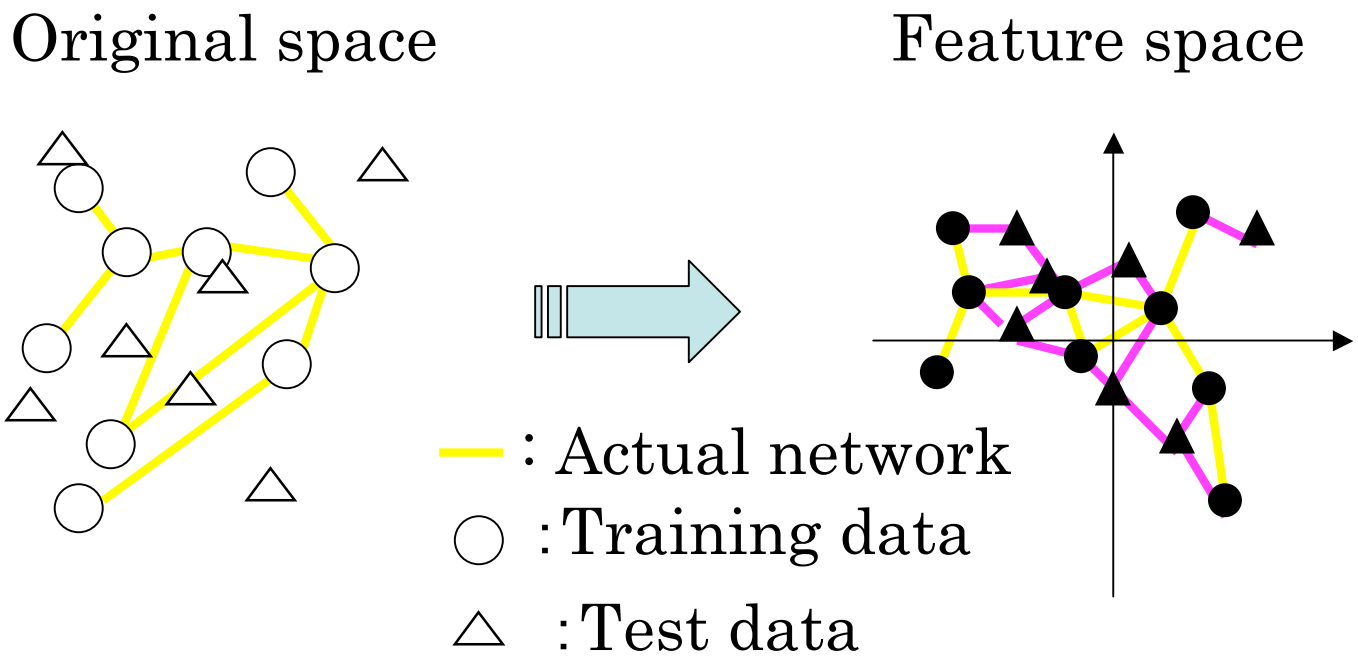
Supervised approach 2/2

Step 2: Prediction of the interactions related to the test data

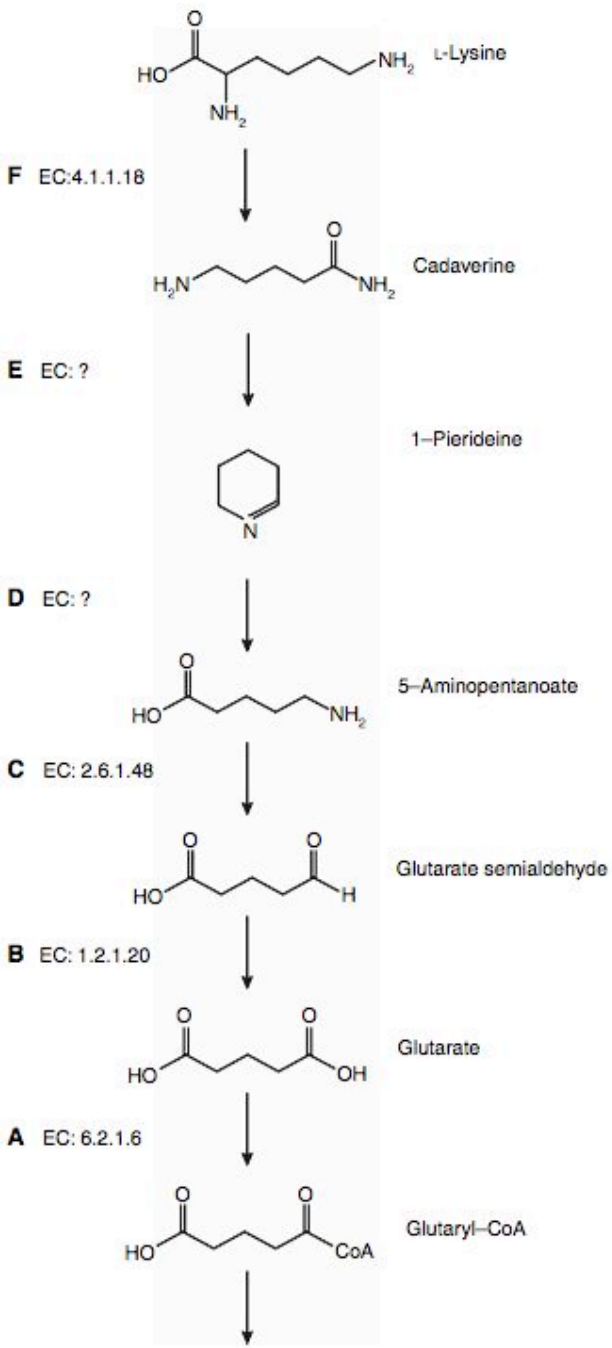


Supervised approach 2/2

Step 2: Prediction of the interactions related to the test data



Prediction of missing enzyme genes



Reaction	Candidate gene
A. EC:6.2.1.6	PA1589 (succinyl-CoA synthetase; EC 6.2.1.5)
B. EC:1.2.1.20	PA0265 (dehydrogenase; EC 1.2.1.16)
C. EC:2.6.1.48	PA0266 (amino-transferase; EC 2.6.1.19)
D. 5ami.1-pip	PA1576 (dehydrogenase; EC 1.1.1.31)
E. Cadav.Delta	not specified
F. EC:4.1.1.18	not specified

Yamanishi, et al. (2007) FEBS J. 274:2262

Network Biology

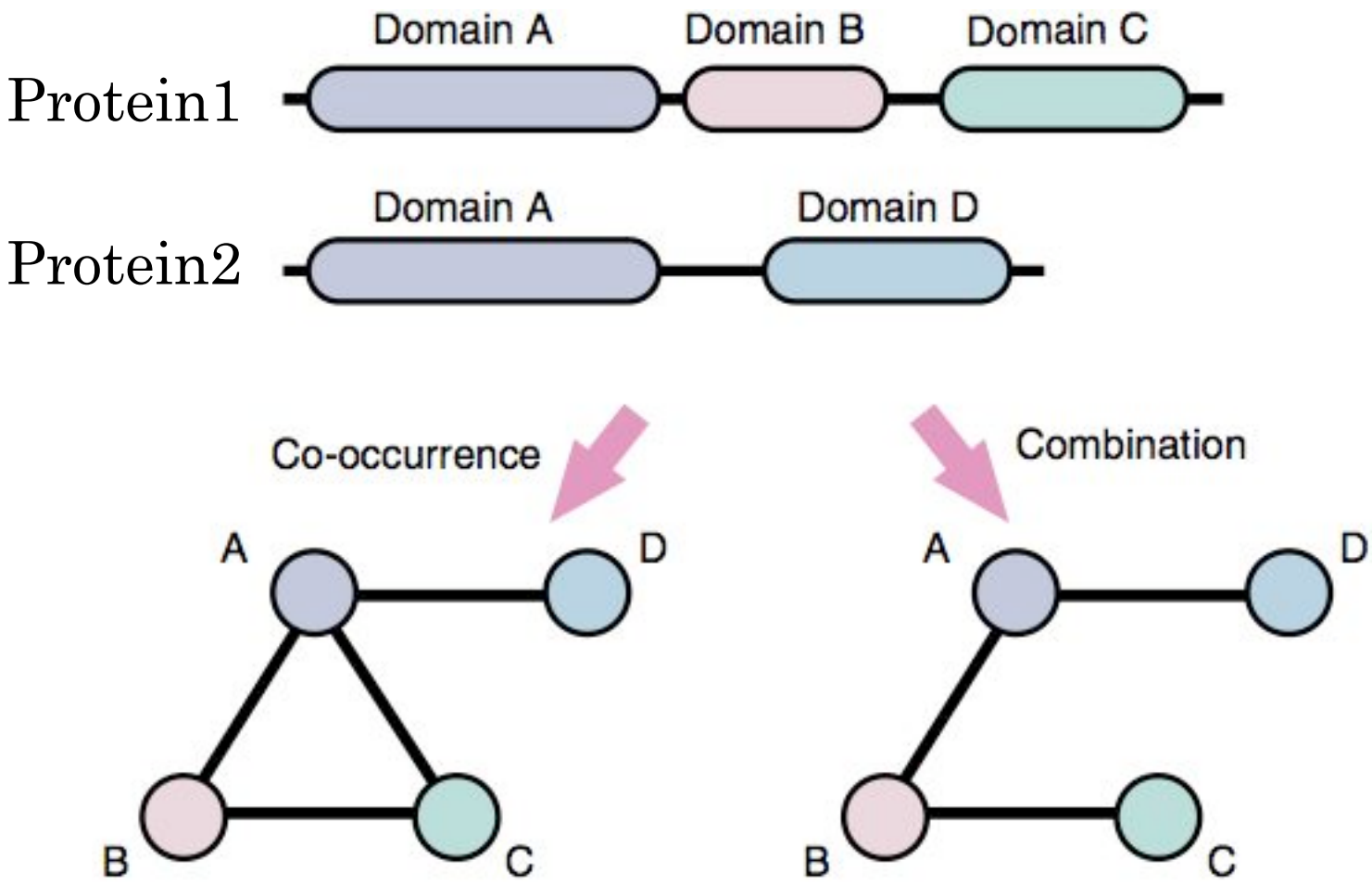
- Power-law and scale-free distribution of node degree
- Robustness of the network
- Preferential or rich-get-richer model of evolution

Barabasi & Oltvai (2004) Nat Rev Genet. 5:101

Topological analysis of biological network

- To elucidate how biological network has emerged
  - Metabolic network including activation and inhibition regulation -> network module
  - Signaling and transcription network
  - Domain combination network
- To analyze function of genes in terms of network

Domain combination network

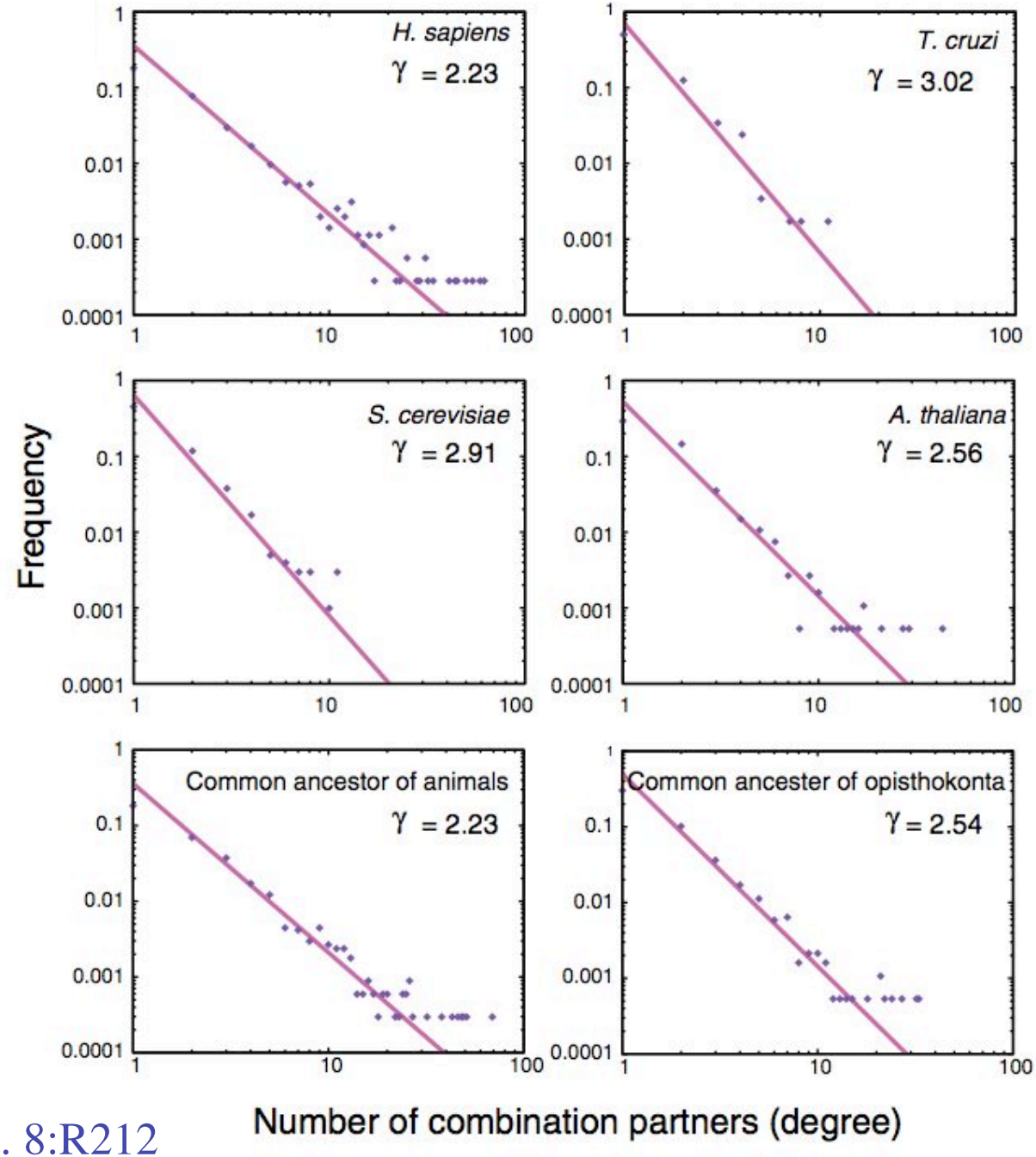


Domain and its combination assignment to eukaryote proteomes

Category				Species	Specific domains (combinations)				Domains shared by prokaryotes
Eukaryotes	Animals	Vertebrates	Mammals	*H. sapiens* *P. troglodytes* *M. musculus* *R. norvegicus* *C. familiaris*	116 (185)	* 235 (610)	407 (875)	** 83 (70)	1439 (715)	1211 (225)
			Bird	*G. gallus*						
			Fishes	*D. rerio* *F. rubripes* *T. nigroviridis*	2 (40)					
		Ascidian		*C. intestinalis*	0 (188)					
		Insects		*D. melanogaster* *D. pseudoobscura* *A. gambiae* *A. mellifera* *B. mori*	22 (40)					
		Nematoda		*C. elegans* *C. briggsae*	73 (70)					
	Fungi	Basidiomycetes		*C. neoformans* B-3501A *C. neoformans* JEC21	1 (10)		34 (55)			
		Ascomycetes		*N. crassa* *M. grisea* *S. bayanus* *S. cerevisiae* *S. mikatae* *S. paradoxus* *K. lactis* *Y. lipolytica* *D. hansenii* *A. gossypii* *C. albicans* *C. glabrata* *S. pombe*	40 (46)					
		Microsporidian		*E. cuniculi*	1 (0)					
	Amoebozoa			*D. discoideum* *E. histolytica*	5 (9)					
	Alveolata			*C. hominis* *C. parvum* *P. falciparum* *P. yoelii* *T. annulata* *T. parva*	5 (9)					
	Euglenozoa			*L. major* *T. brucei* *T. cruzi*	4 (5)					
	Plants	Land plants		*A. thaliana* *O. sativa*	240 (178)		31 (30)			
		Red algae		*C. merolae*	8 (33)					
Prokaryotes										

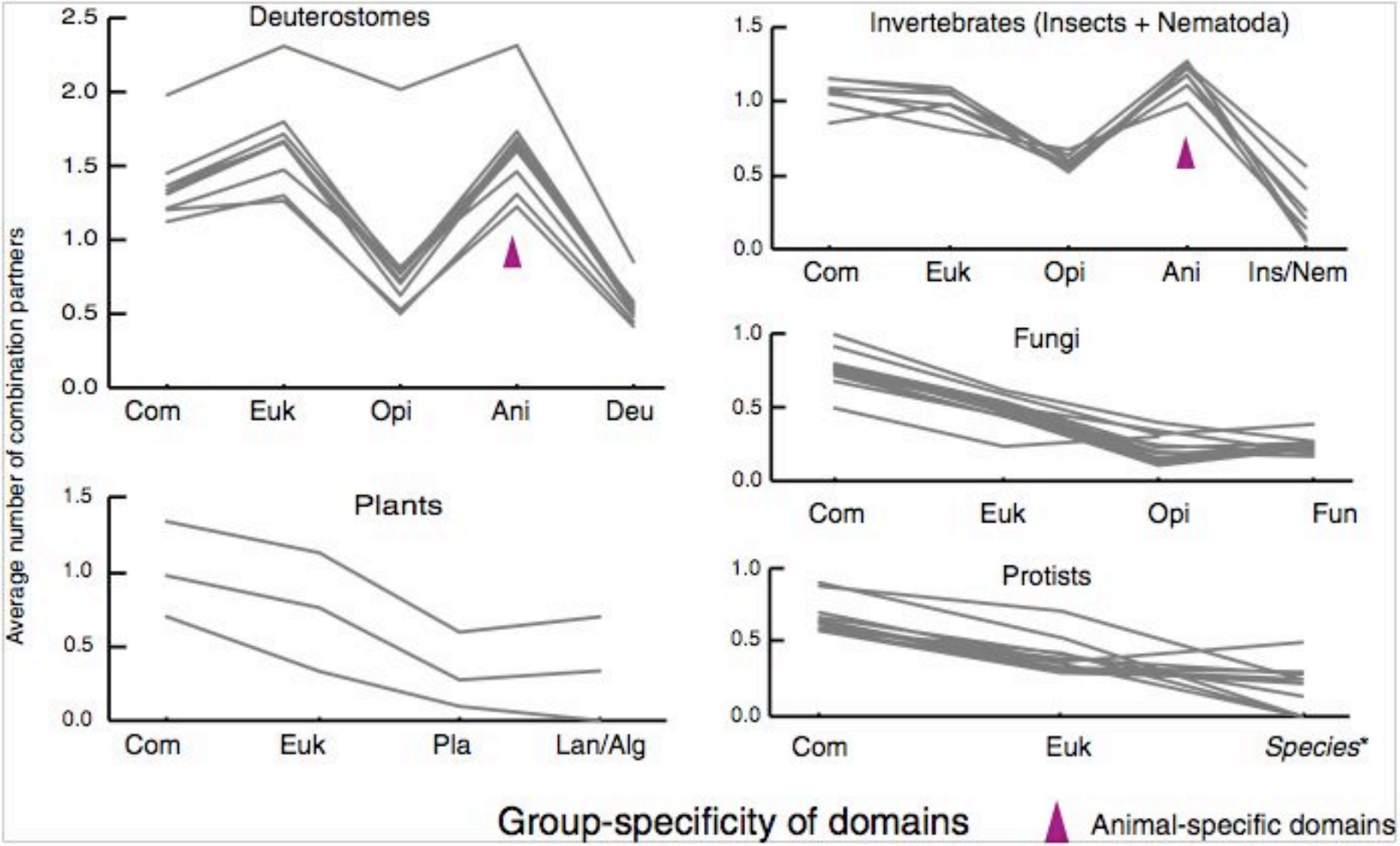
Itoh, et al. (2007) Genome Biol. 8:R212

Domain combination network shows a scale-free property



Itoh, et al. (2007) Genome Biol. 8:R212

Older domains have more partners with some exceptions

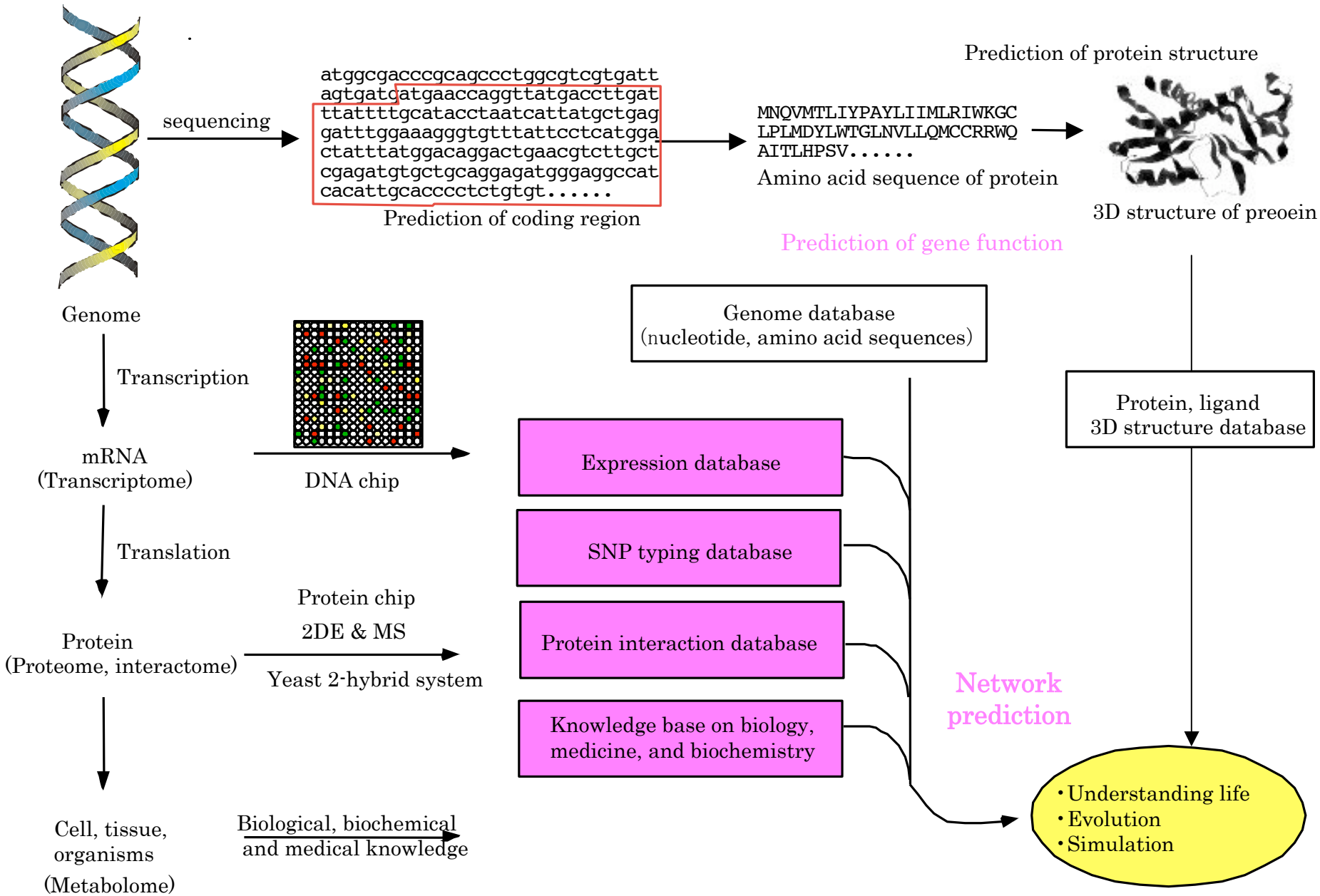


Itoh, et al. (2007) Genome Biol. 8:R212

Summary

- Importance of network in bioinformatics analysis
  - Function prediction
  - Domain network analysis
  - Functional modules in metabolic network
    - Yamada, et al. (2006) BMC Bioinformatics 7:130
    - Gutteridge, et al. (2007) BMC Bioinformatics 8:88

- What is Life?
  - What is not life?
    - What is or is not used in biological system in terms of protein network, domain network?


Genome
sequencing
atggcgacccgcagccctggcgtcgtgatt
agtgatcatgaaccaggttatgaccttgat
ttattttgcatacctaatcattatgctgag
gatttggaaagggtgtttattcctcatgga
ctatttatggacaggactgaacgtcttgct
cgagatgtgctgcaggagatgggaggccat
cacattgcacccctctgtgt.....
Prediction of coding region
MNQVMTLIYPAYLIIMLRIWKGC
LPLMDYLWTGLNVLLQMCCRRWQ
AITLHPSV.....
Amino acid sequence of protein
Prediction of protein structure
3D structure of preoein
Prediction of gene function
Transcription
mRNA
(Transcriptome)
DNA chip
Translation
Protein
(Proteome, interactome)
Protein chip
2DE & MS
Yeast 2-hybrid system
Cell, tissue,
organisms
(Metabolome)
Biological, biochemical
and medical knowledge
Genome database
(nucleotide, amino acid sequences)
Expression database
SNP typing database
Protein interaction database
Knowledge base on biology,
medicine, and biochemistry
Protein, ligand
3D structure database
Network
prediction
• Understanding life
• Evolution
• Simulation

Acknowledgement

Funtion prediction

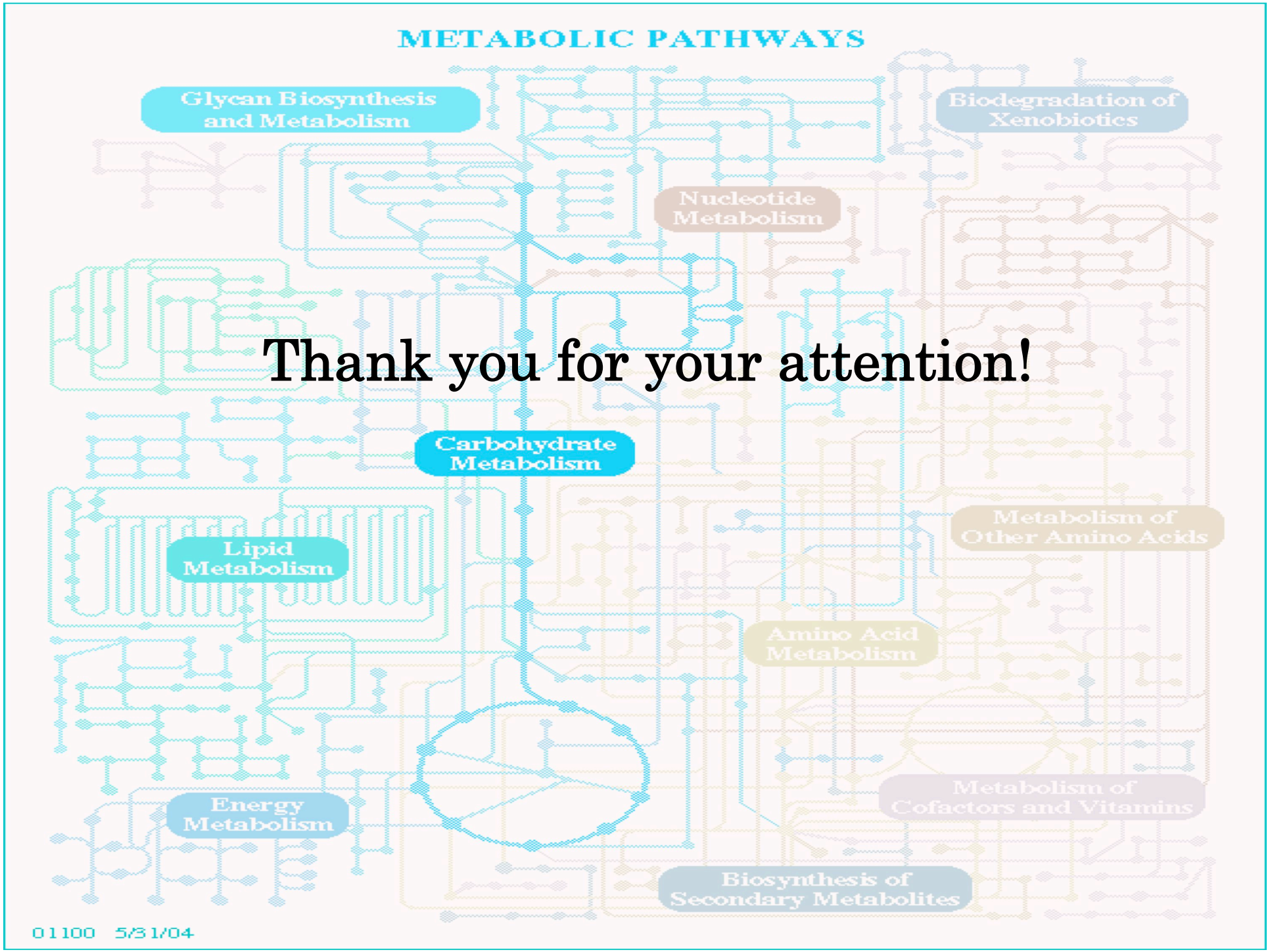
Yoshihiro Yamanishi
Hisaaki Mihara
Motoharu Osaki
Hisashi Muramatsu
Nobuyoshi Esaki
Tetsuya Sato
Yoshiyuki Hizukuri
Minoru Kanehisa

Domain network analysis

Masumi Itoh
Jose C Nacher
Kei-ichi Kuma
Minoru Kanehisa

Database

Minoru Kanehisa
KEGG development team



Thank you for your attention!