



HAL
open science

Topological and semantic Web based method for analyzing TGF- β signaling pathways

Jean Coquet, Geoffroy Andrieux, Jacques Nicolas, Olivier Dameron, Nathalie
Théret

► **To cite this version:**

Jean Coquet, Geoffroy Andrieux, Jacques Nicolas, Olivier Dameron, Nathalie Théret. Topological and semantic Web based method for analyzing TGF- β signaling pathways. JOBIM 2015, Dec 2015, Clermont-Ferrand, France. JOBIM: Journées Ouvertes en Biologie, Informatique & Mathématiques, 2015. hal-01242893

HAL Id: hal-01242893

<https://hal.science/hal-01242893v1>

Submitted on 14 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSIS OF TGF- β SIGNALIZATION PATHWAYS THANKS TO TOPOLOGICAL AND SEMANTIC WEB METHOD

Jean COQUET ^{1,2}, Geoffroy ANDRIEUX ³, Jacques NICOLAS ¹, Olivier DAMERON ¹ and Nathalie THERET ²

¹ Université de Rennes 1 - IRISA/INRIA, UMR6074, 263 avenue du Général Leclerc, 35042 Rennes, Cedex, France

² Université de Rennes 1 - IRSET EA 4427 SeRAIC, IFR140, 2 avenue Pr Léon Bernard, Rennes, Cedex, France

³ German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

CONTEXT

- TGF- β = Transforming Growth Factor β
 - necessary in healthy situation
 - deleterious role in fibrosis and cancer
- Pleiotropic effects of TGF- β are linked to the complex nature of its activation and signaling networks

BIOLOGICAL QUESTIONS:

Which pathways are cancer-specific?

Which mechanism is involved in the healthy \rightarrow pathological transition?

SIGNALIZATION MODEL

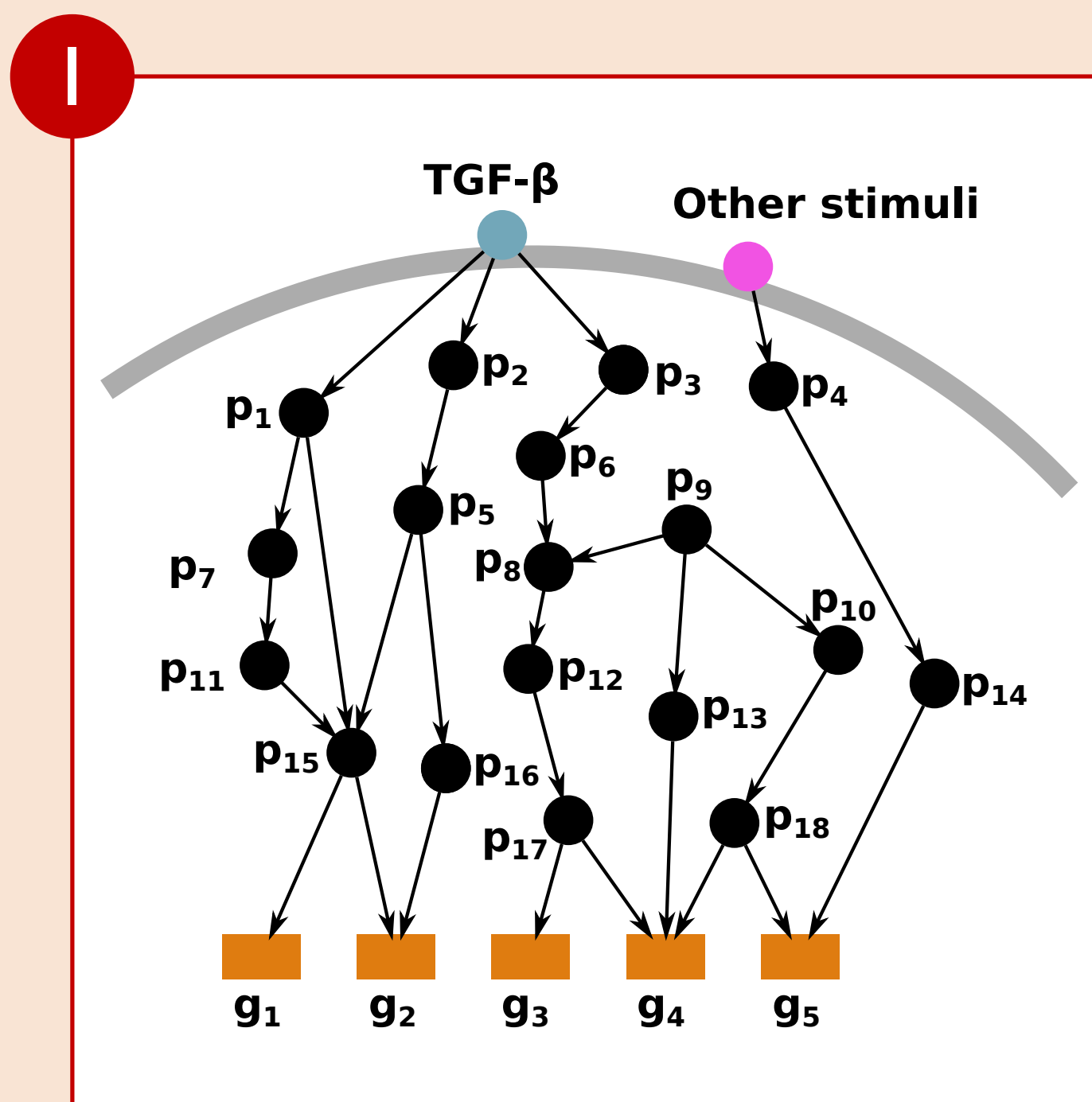
- Model of TGF- β signal propagation based on guarded transitions [Andrieux et al, 2014] represented in CADBIOM.
- Pathway Interaction Database (PID) transform into a single unified model of signal transduction

DATA ANALYSIS PROBLEM:

16,000 chains of reactions linking TGF- β to at least one of 159 target genes.

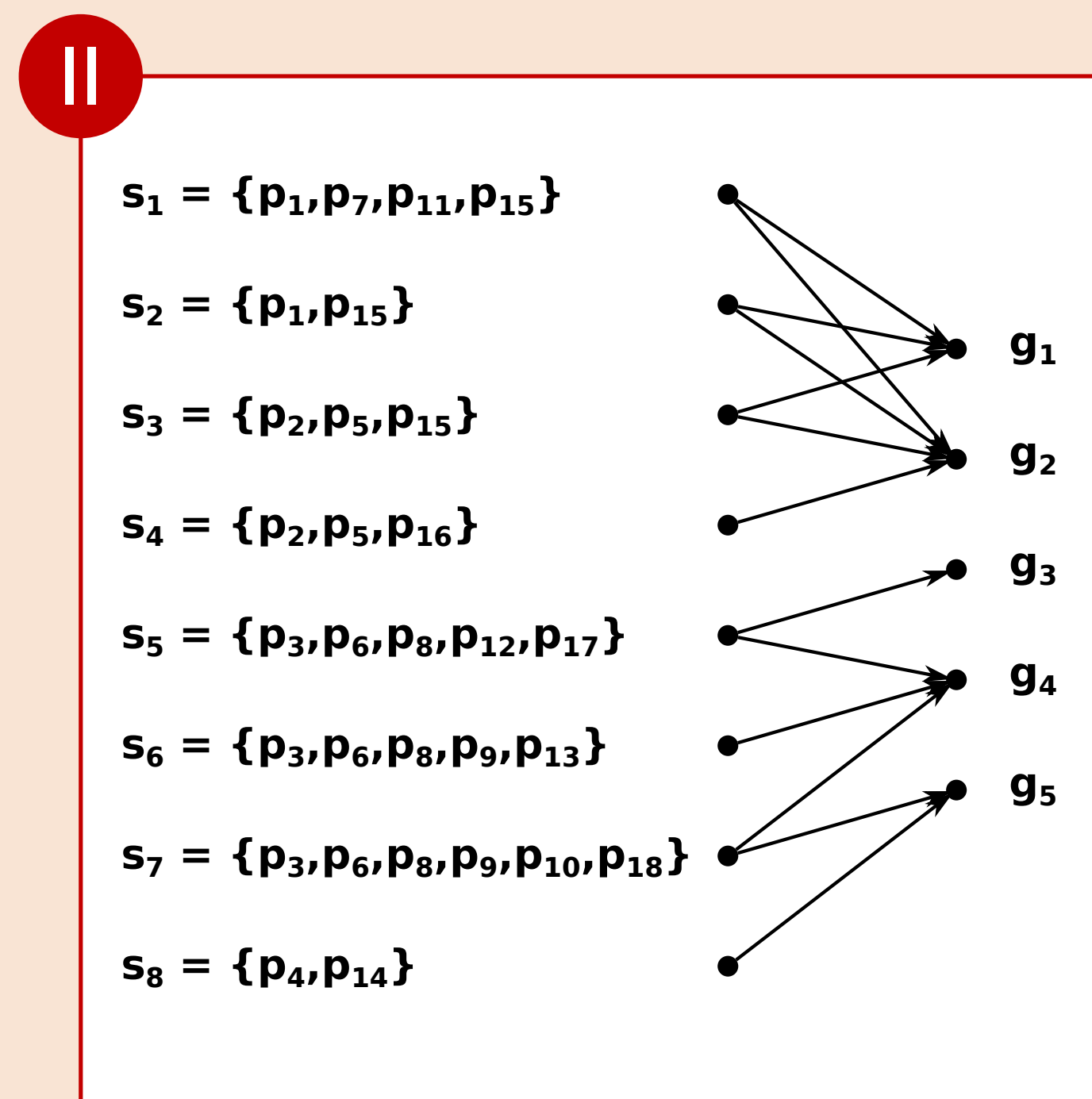
How to explore the combinatorial complexity of cell signaling?

ANALYSIS OF SIGNALING PATHWAYS : STEPS METHOD



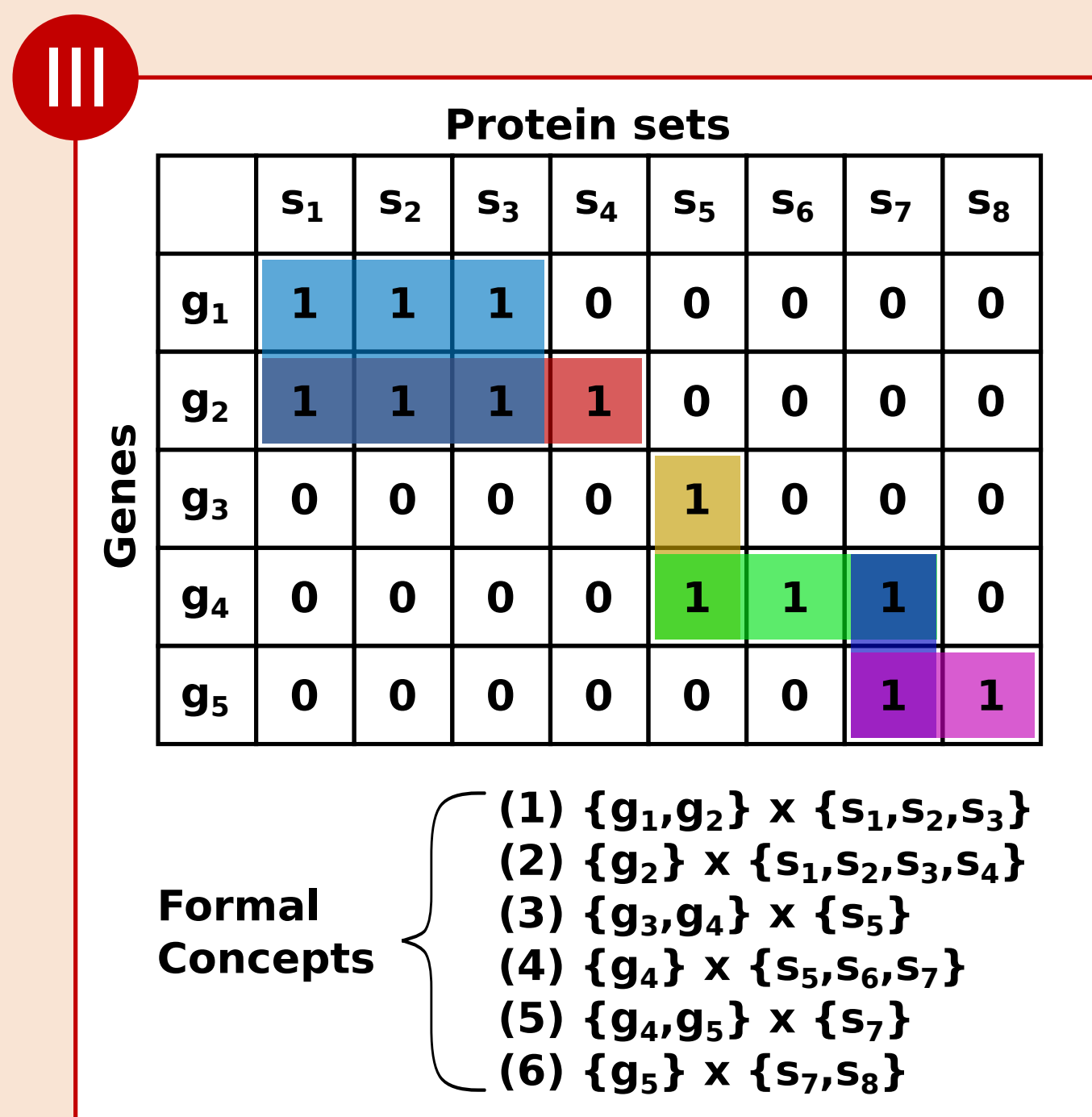
RAW DATA

- TGF- β signalizations influence some genes
- Other stimuli in addition to TGF- β



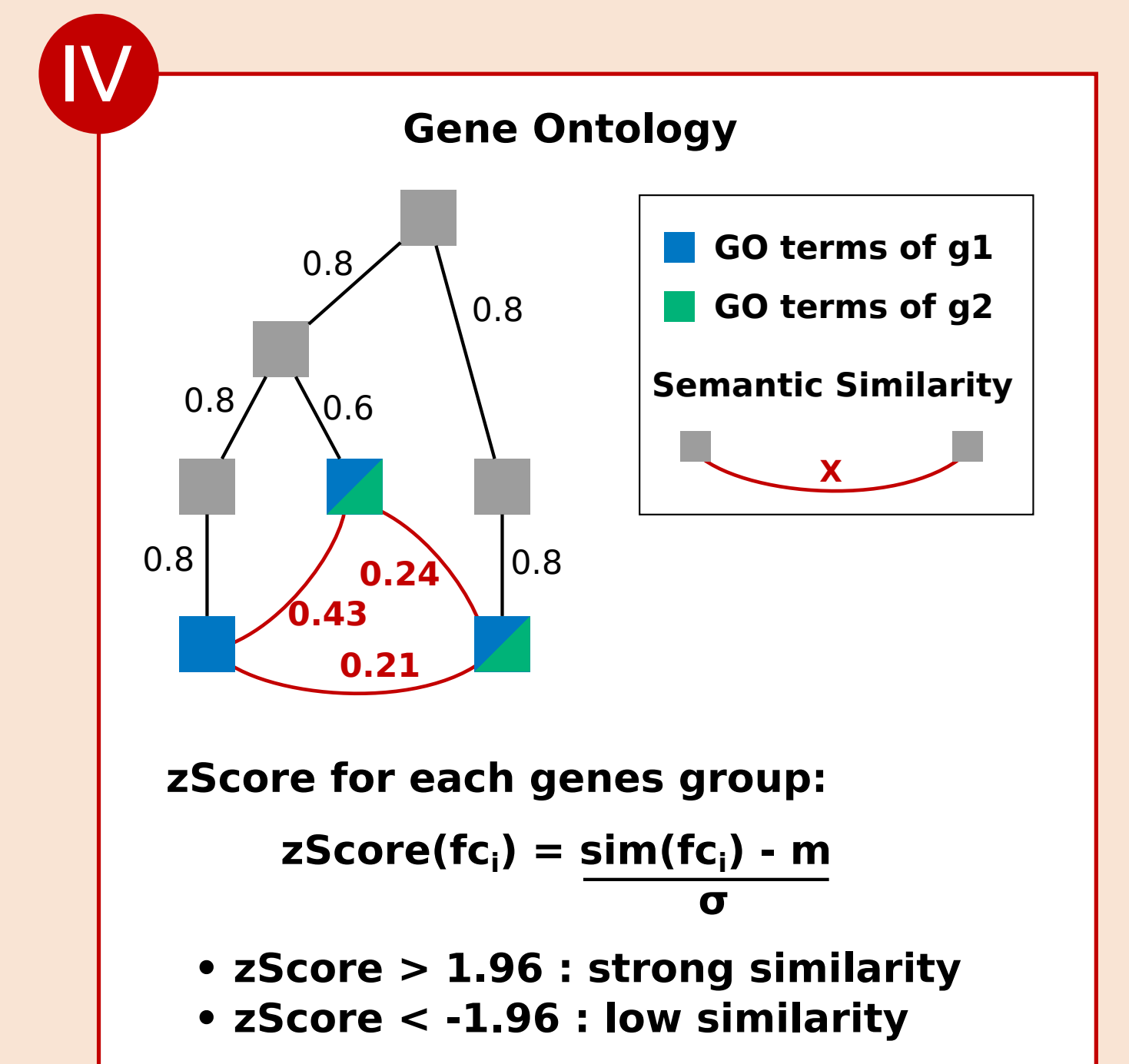
GENES INFLUENCED BY PROTEIN SETS

- Signaling pathway = proteins set
- Proteins set influence at least one gene



CLUSTERS OF SIMILAR GENES AND PROTEIN SETS

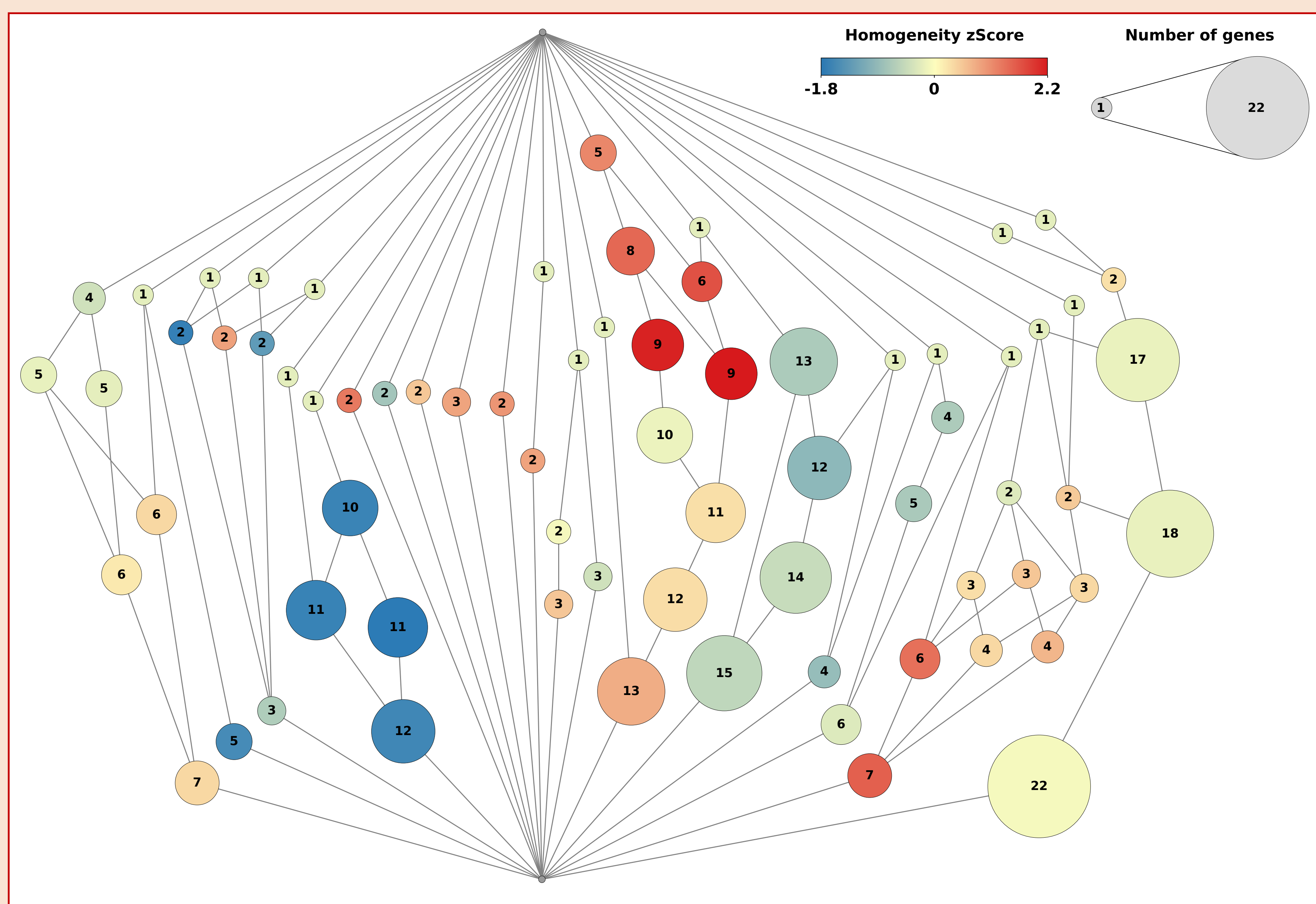
- Identify clusters of similar genes and protein sets with formal concept analysis
- Binary matrix
- Formal concepts = maximal bi-cliques



ASSESS CLUSTERS' BIOLOGICAL RELEVANCE WITH SEMANTIC HOMOGENEITY

- Formal concept homogeneity = zScore of similarity gene group
- Semantic similarity based on Wang measure [Wang et al, 2007]

RESULTS



Concept lattice. Nodes are formal concepts, two nodes are linked if one is included into the other.

DISCUSSION

- Some formal concepts are highly connected
- Gene clusters can have:
 - High homogeneity = genes similar influenced by same pathways
 - Low homogeneity = genes dissimilar influenced by same pathways
- Chains of gene clusters included into one-another indicates that some influence sets are specific to some gene subsets

IN PROGRESS

- Validate the method on small synthetic models
 - Improve the method to find better gene clusters
 - Integrate the overlap between proteins sets
 - Try different methods of clustering
- Compute the semantic homogeneity of protein sets
- Take into account the structure underlying gene sets
- Compute the stability of formal concepts

References:

- [Andrieux et al, 2014] Andrieux, G., Le Borgne, M., & Thérêt, N. (2014). An integrative modeling framework reveals plasticity of TGF-beta signaling. BMC systems biology, 8(1), 30.
 [Wang et al, 2007] Wang, J. Z., Du, Z., Payattakool, R., Philip, S. Y., & Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. Bioinformatics, 23(10), 1274-1281.

CONCLUSION

Model of TGF- β signal propagation

New method to analyze pathways:

- Topological analysis
- Clustering
- Biological relevance

Gene clusters influenced by same pathways with a biological relevance