NETWORK-BASED INTEGRATION OF OMICS DATA

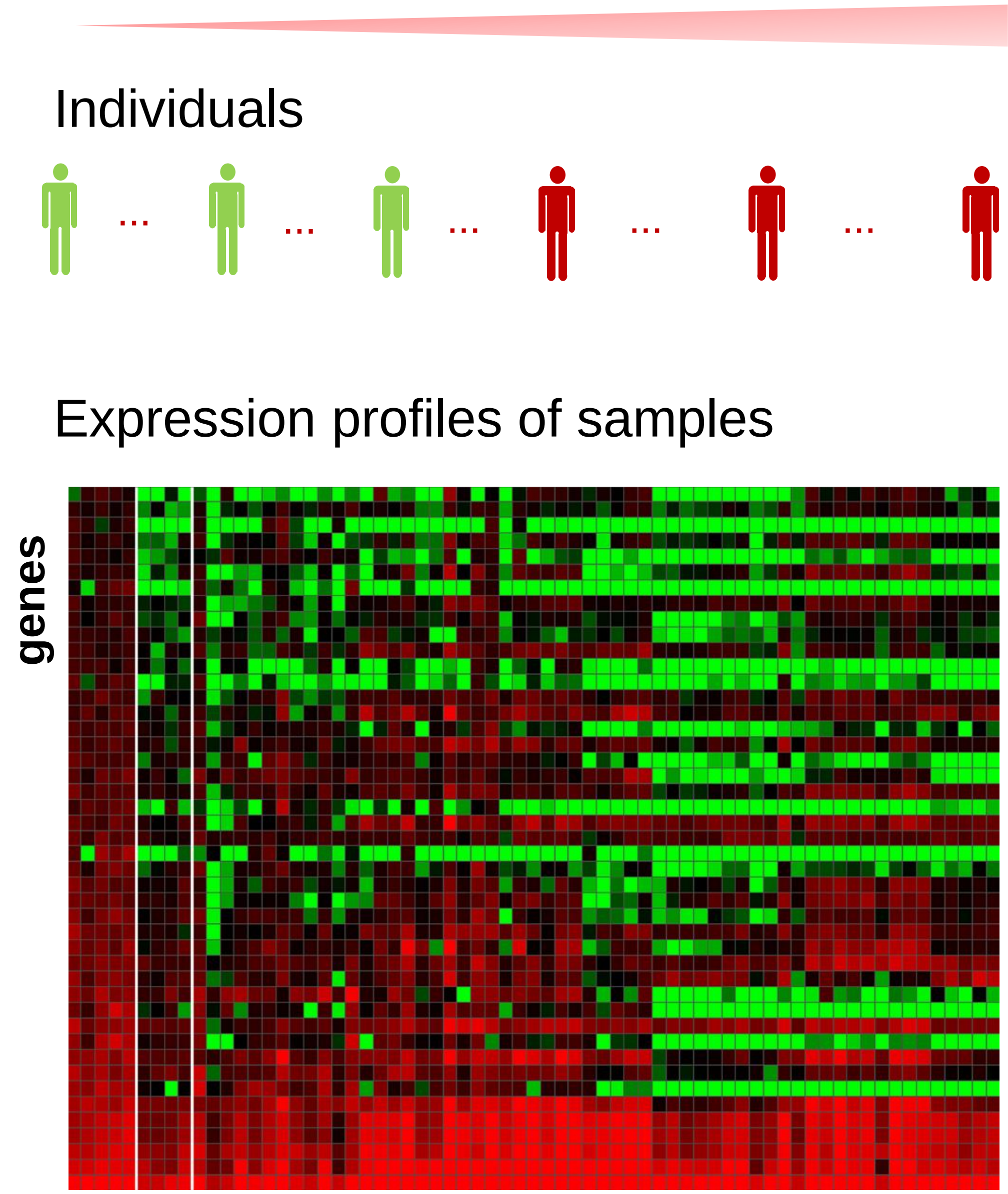
Kathleen Marchal
Dept of Information Technology, IMEC
Dept of Plant Biotechnology and Bioinformatics

GHENT UNIVERSITY

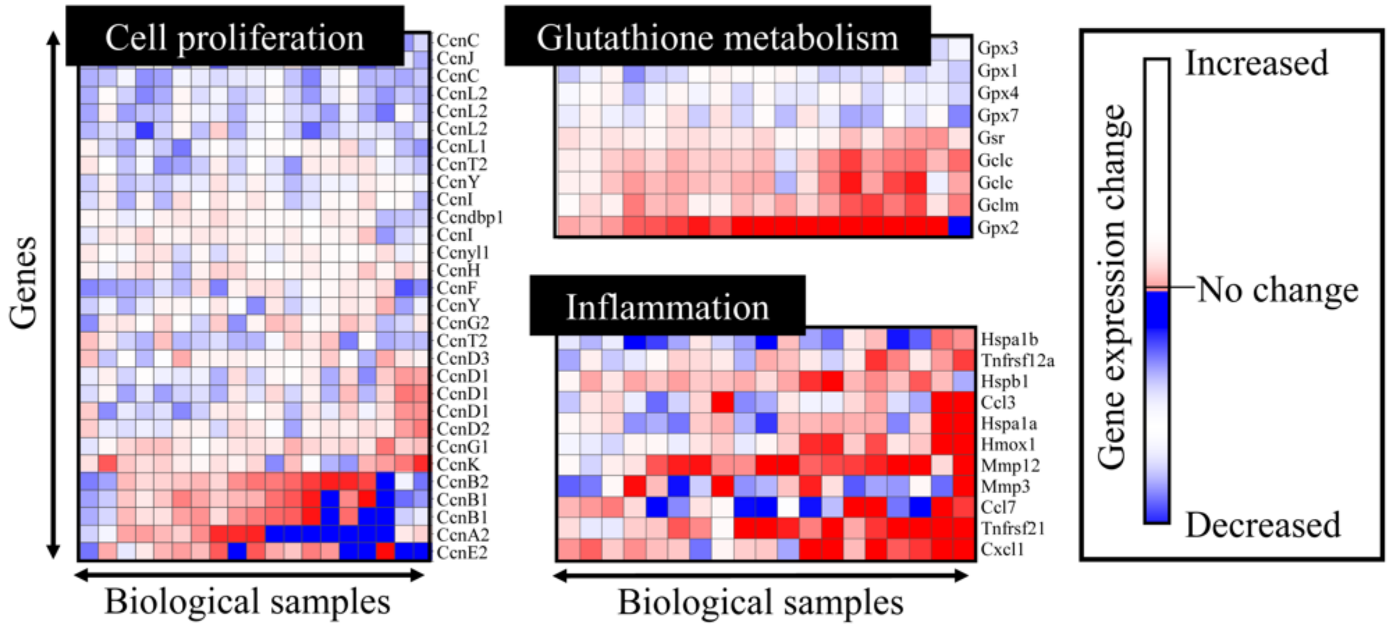
EFSA SCIENCE · FOOD · SOCIETY
THE PROGRAMME IS ONLINE!

Omics data for food safety assessment

Assessing the molecular signature of a toxic compound

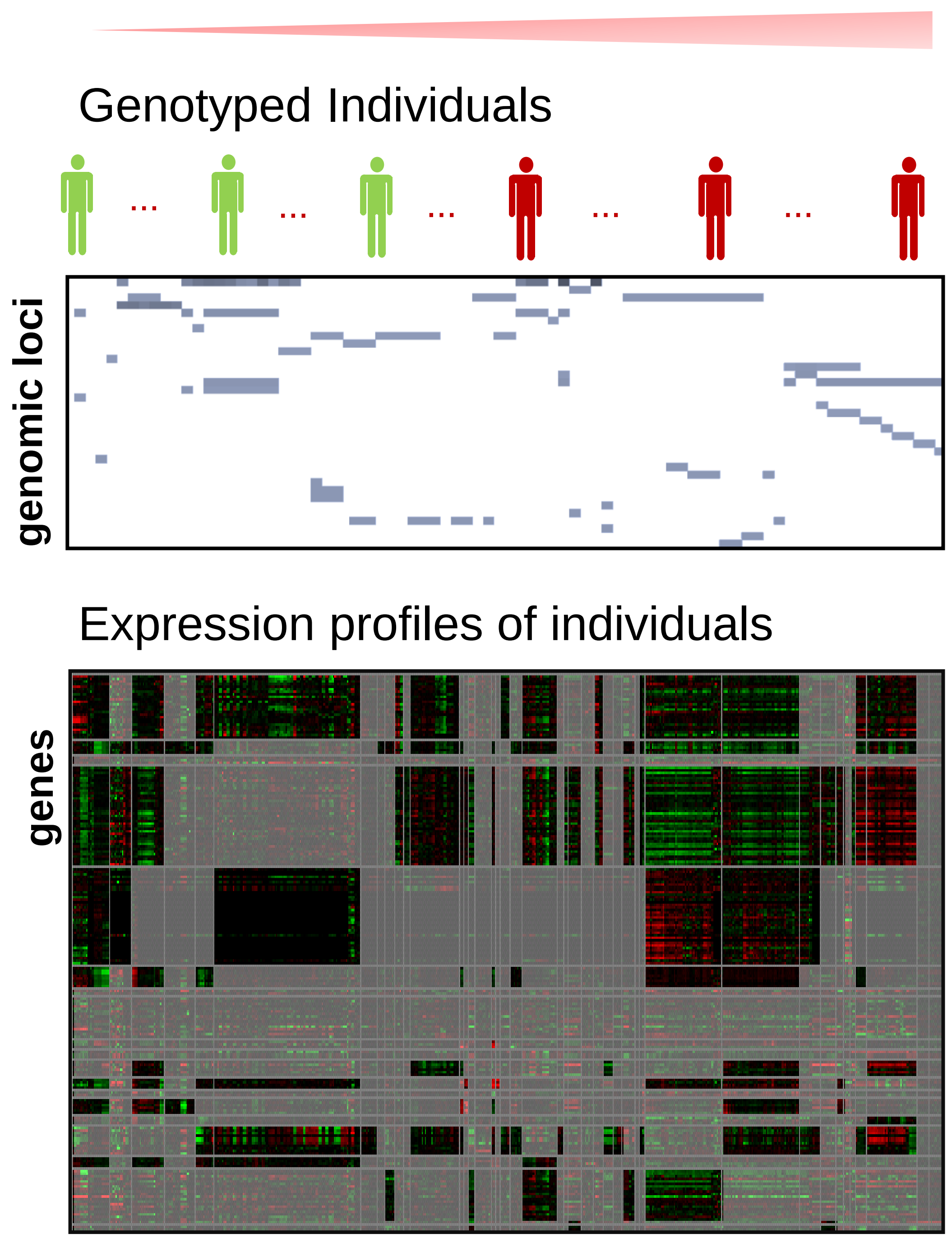


Use expression as a molecular phenotype to identify molecular signatures that can be used as biomarkers & ideally unveil the mode of action of the toxicity



Omics data for food safety assessment

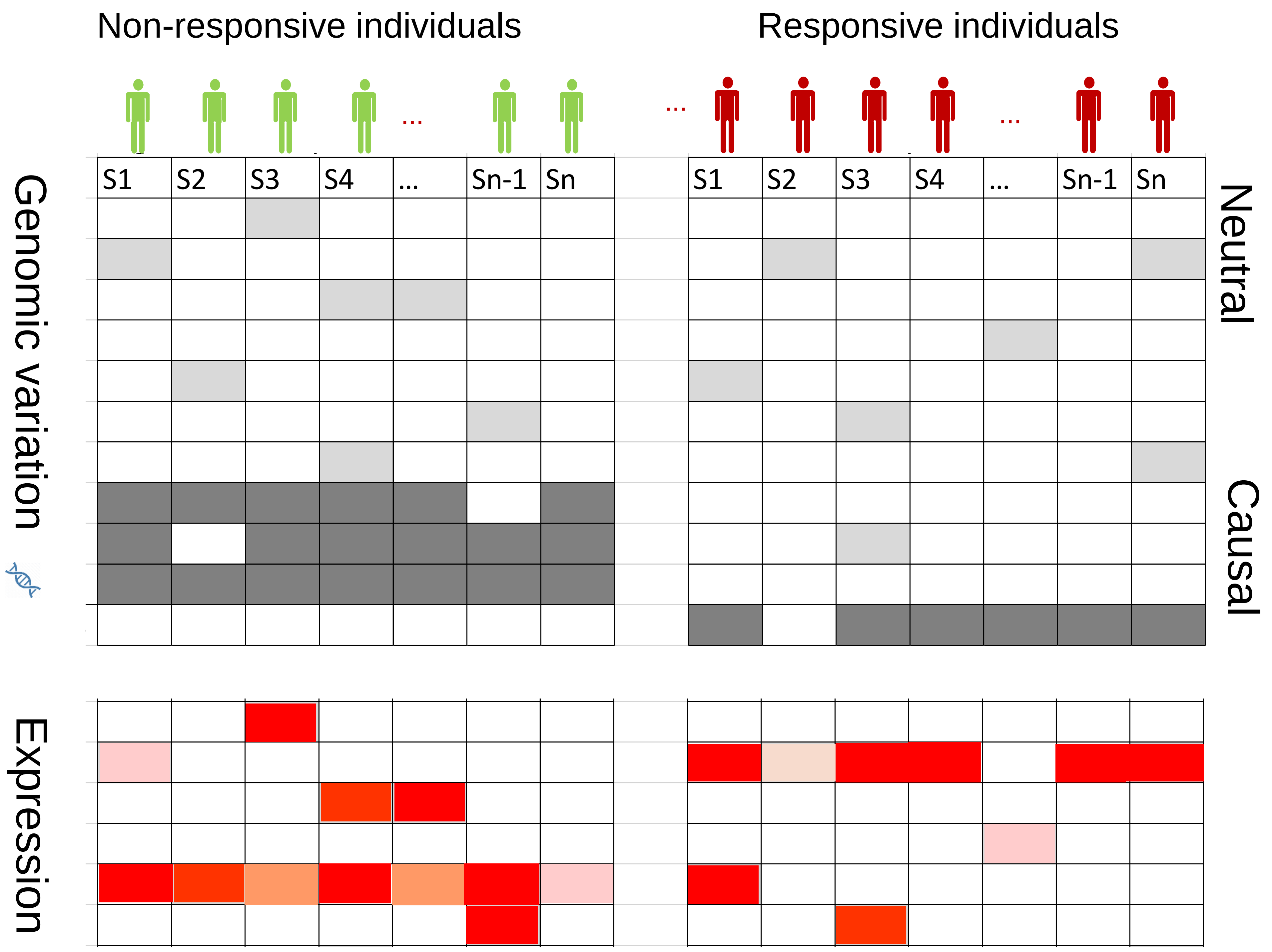
Personalized nutrigenomics



Identifying genomic biomarkers that determine whether or not an individual will elicit a toxic response boils down to omics-based cohort analysis

Omics data integration is statistically ill defined

Multivariate approaches: ANOVA (QTL, eQTL)



Many variables

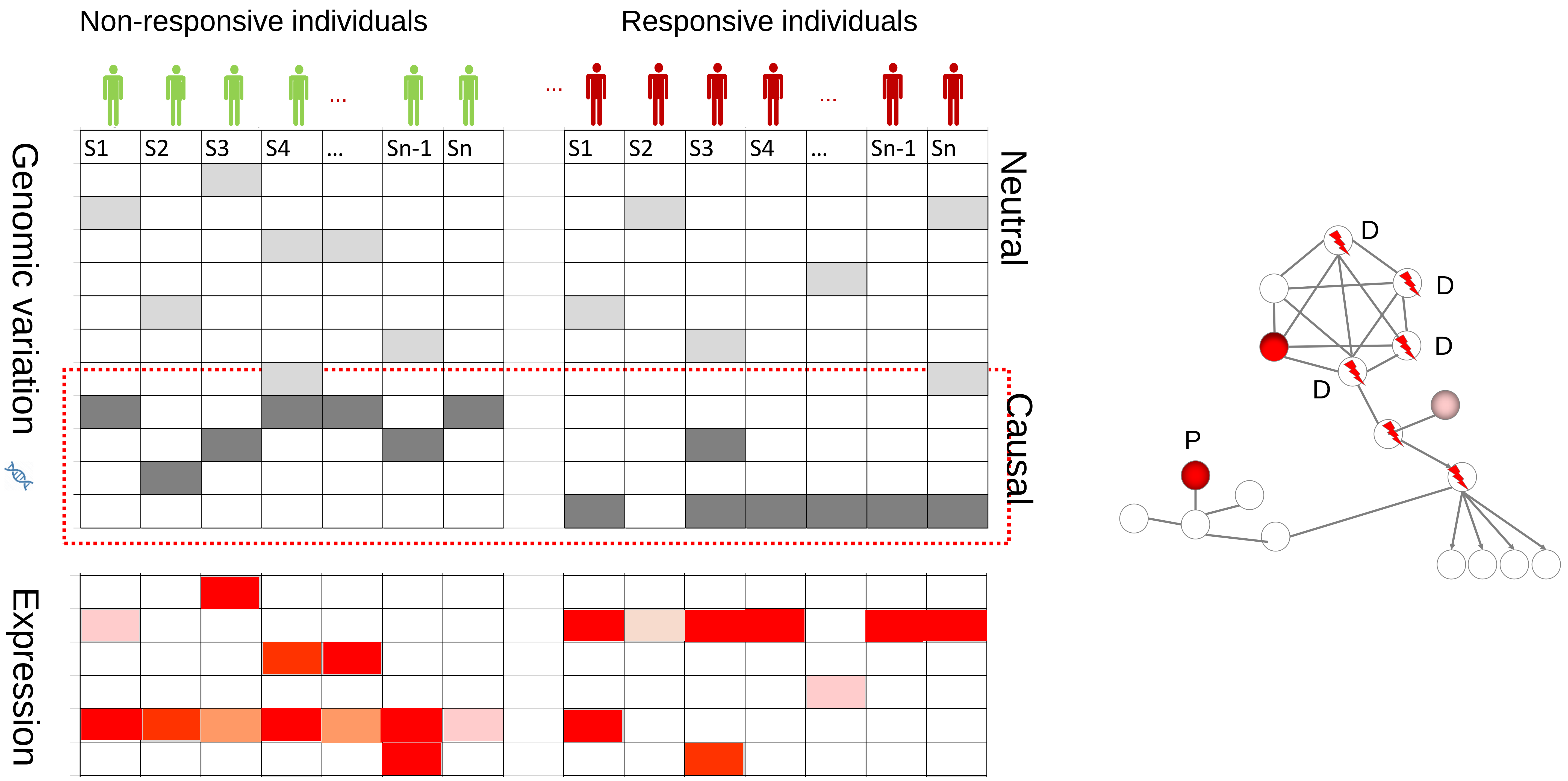
Small cohort size

Noisy data

⇨ Many spurious associations

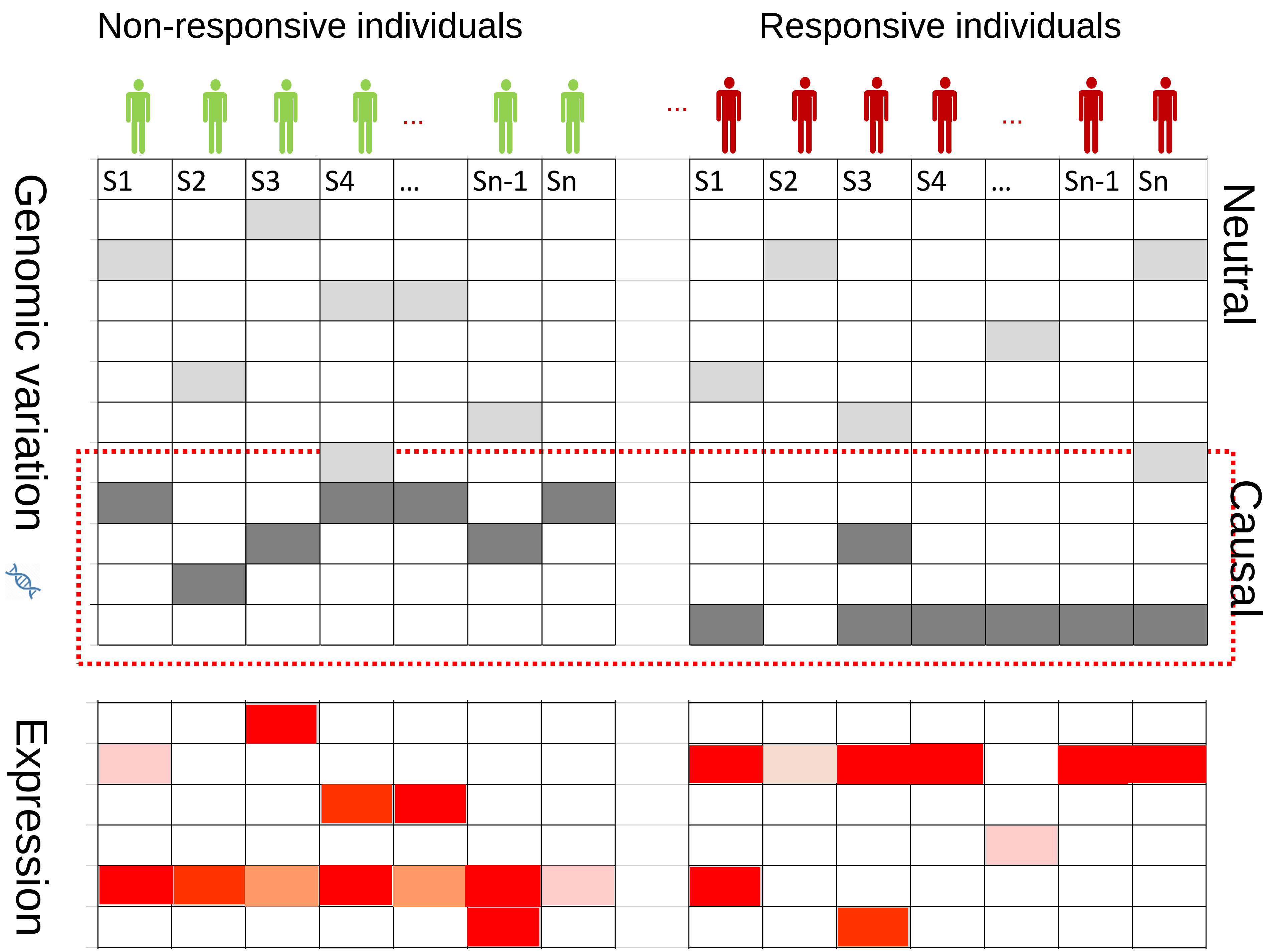
Omics data integration is statistically ill defined

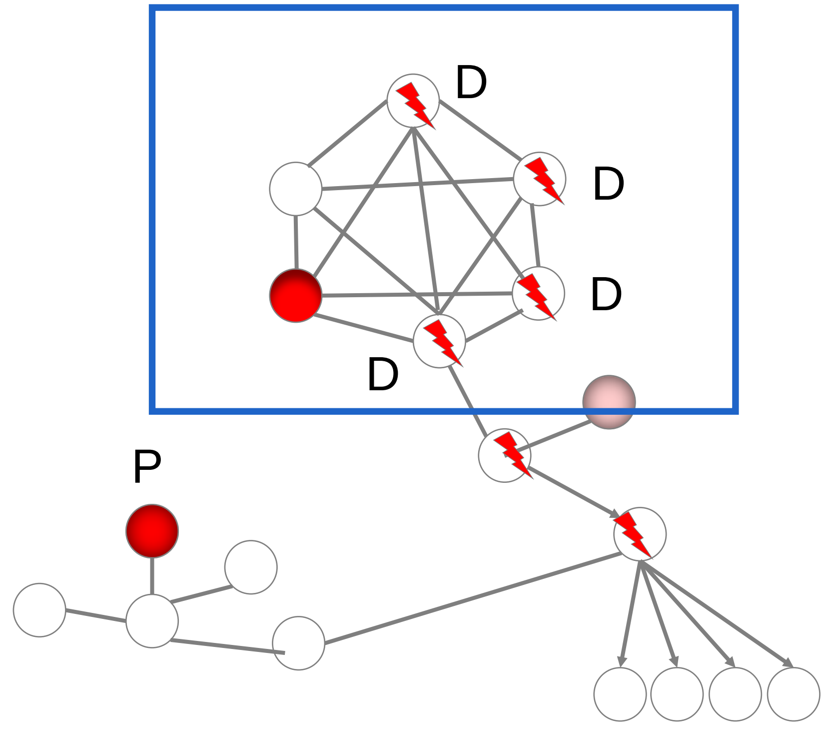
Multivariate approaches: ANOVA (QTL, eQTL)



Omics data integration is statistically ill defined

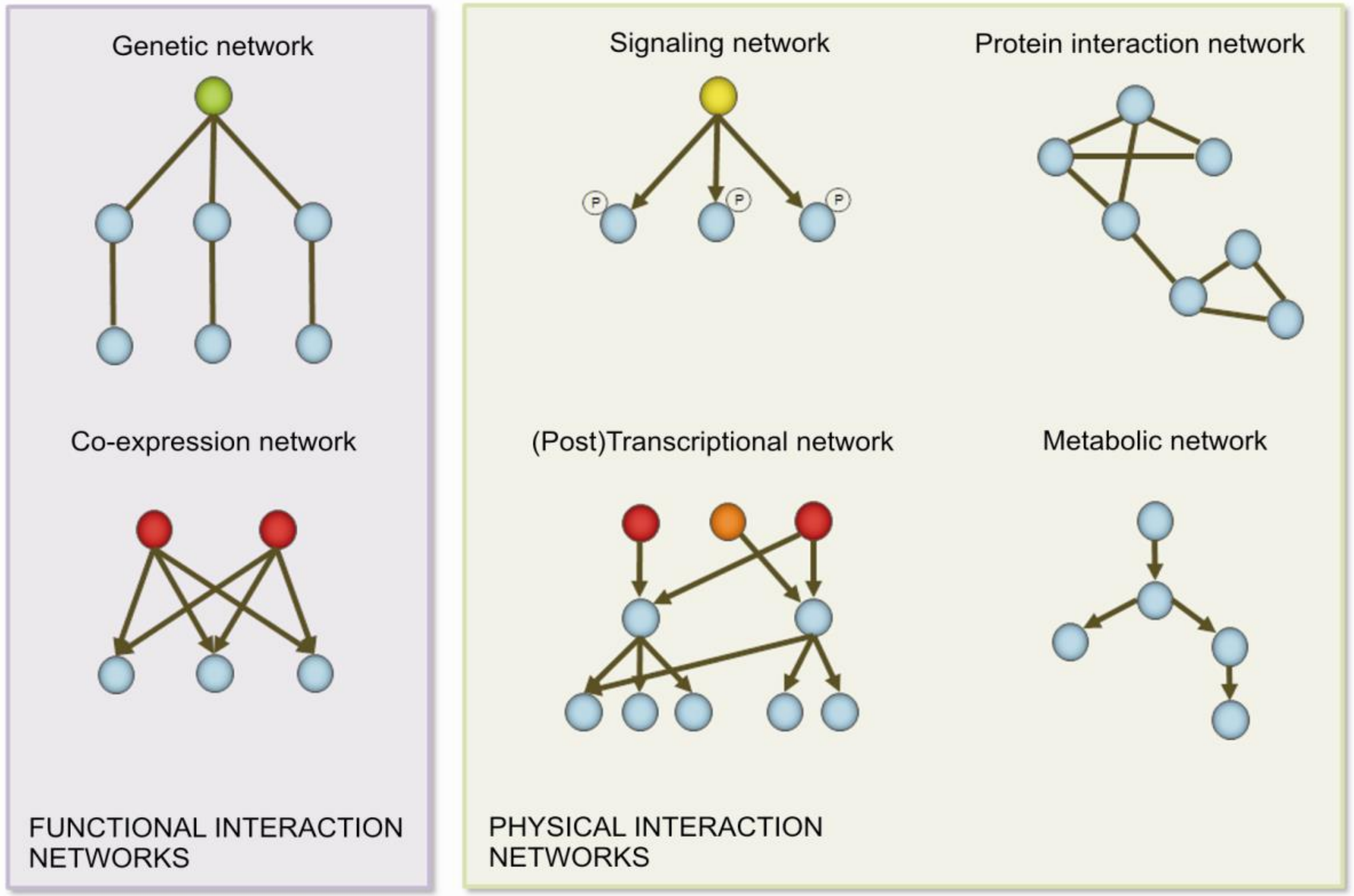
Multivariate approaches: ANOVA (QTL, eQTL)

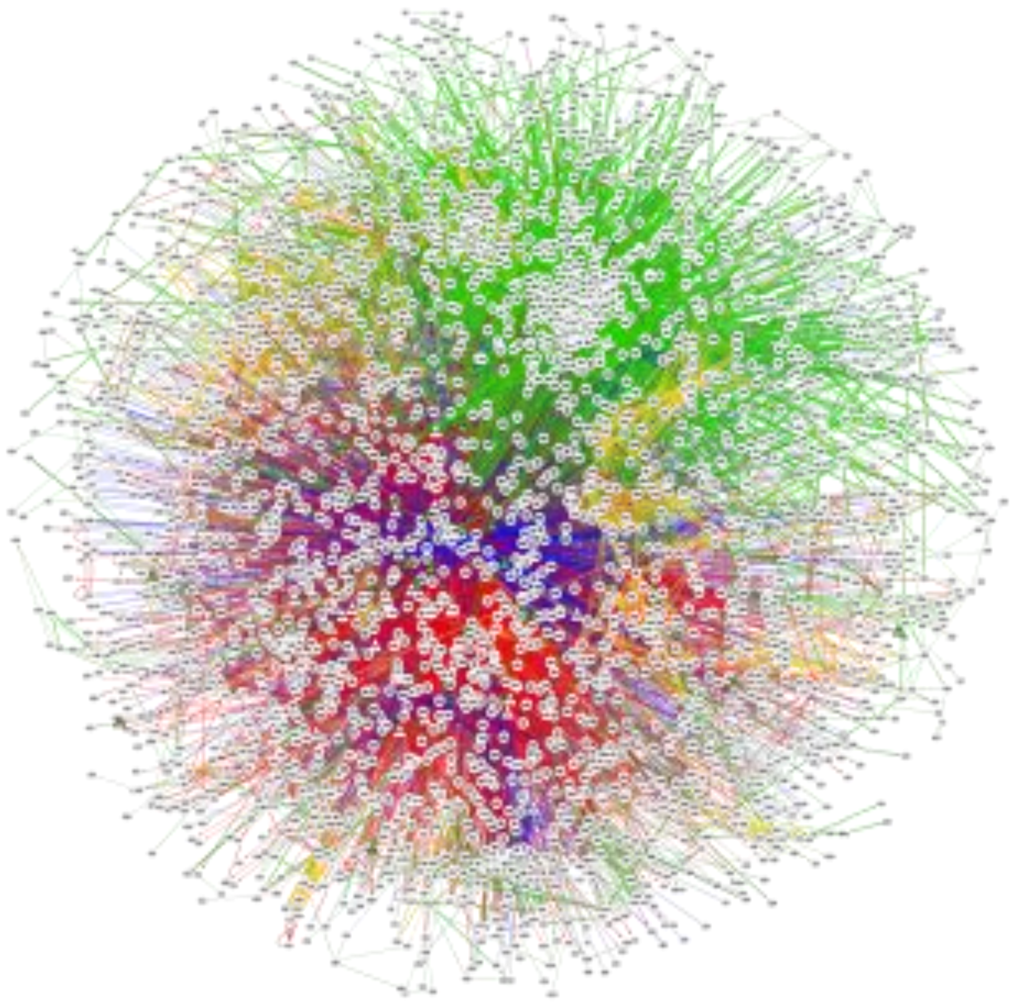




Network-based analysis of omics data

Networks increase the power of the analysis





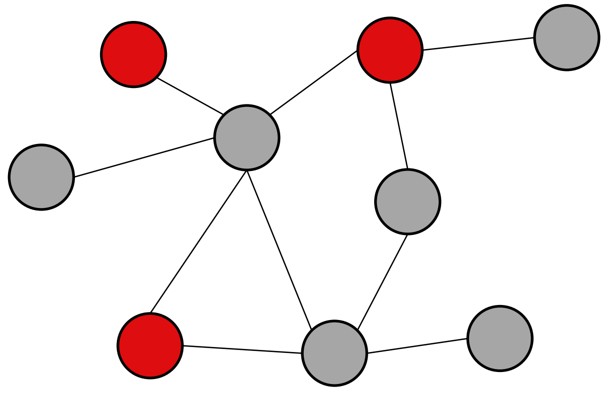
Exploit prior information on molecular interactions as a scaffold to drive the analysis

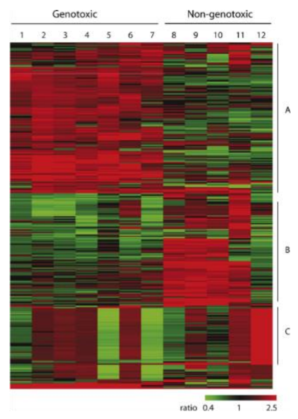
- Steers the solution of the data-integration problem to the most biologically relevant one

Network-based analysis of omics data

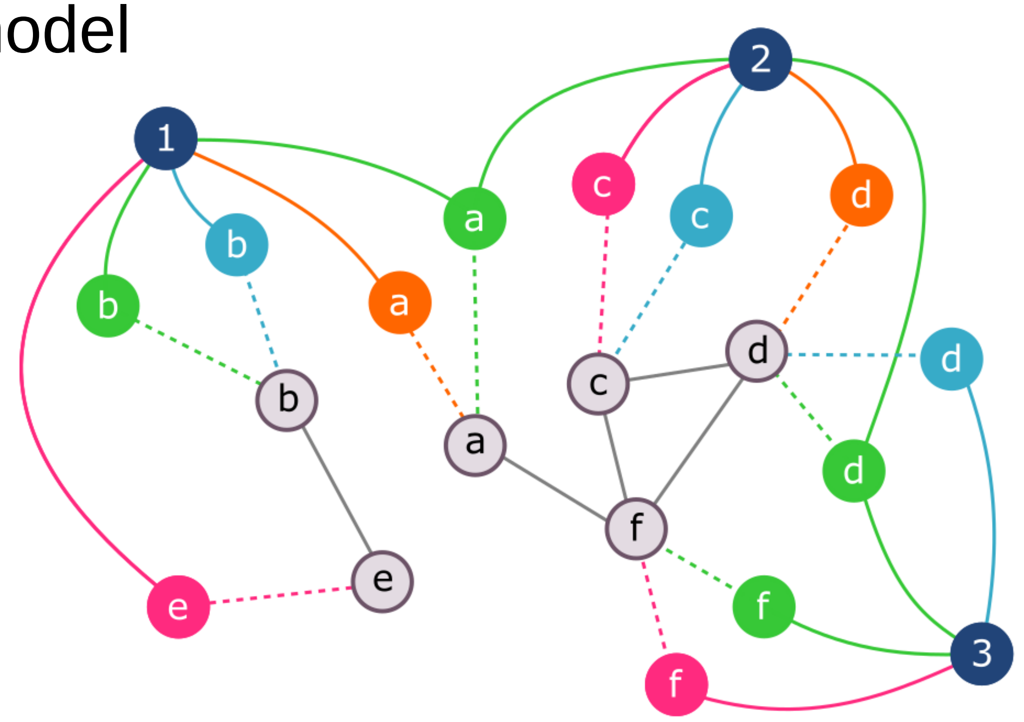
Networks provide an intuitive scaffold to integrate data

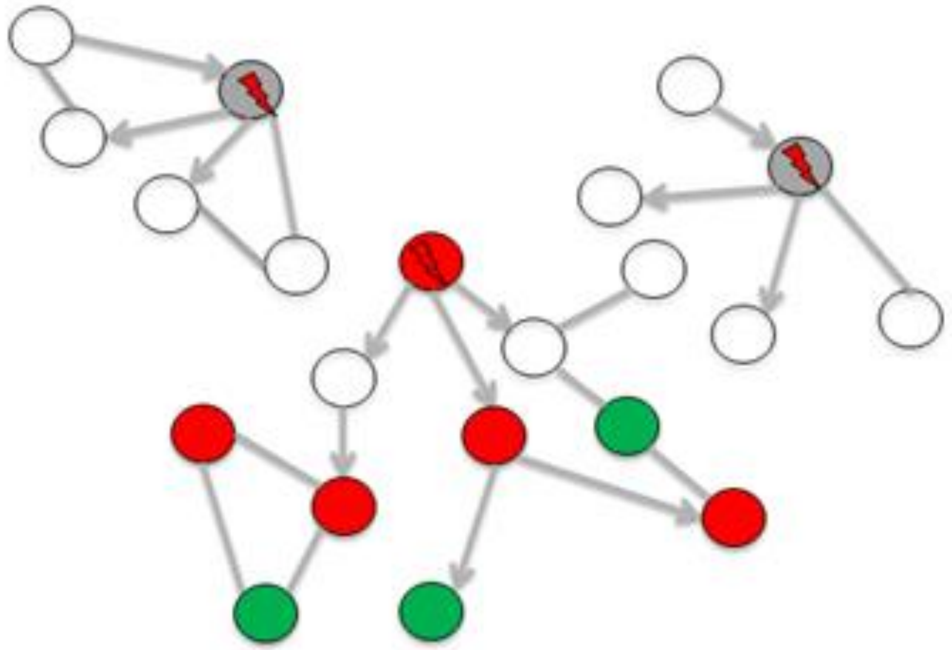
Simple network model

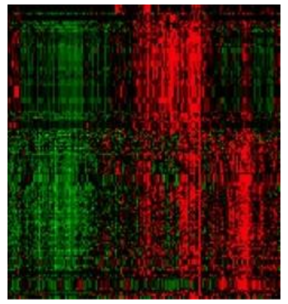


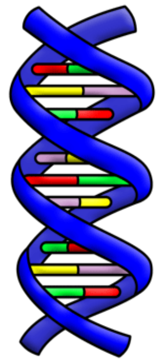


Complex network model



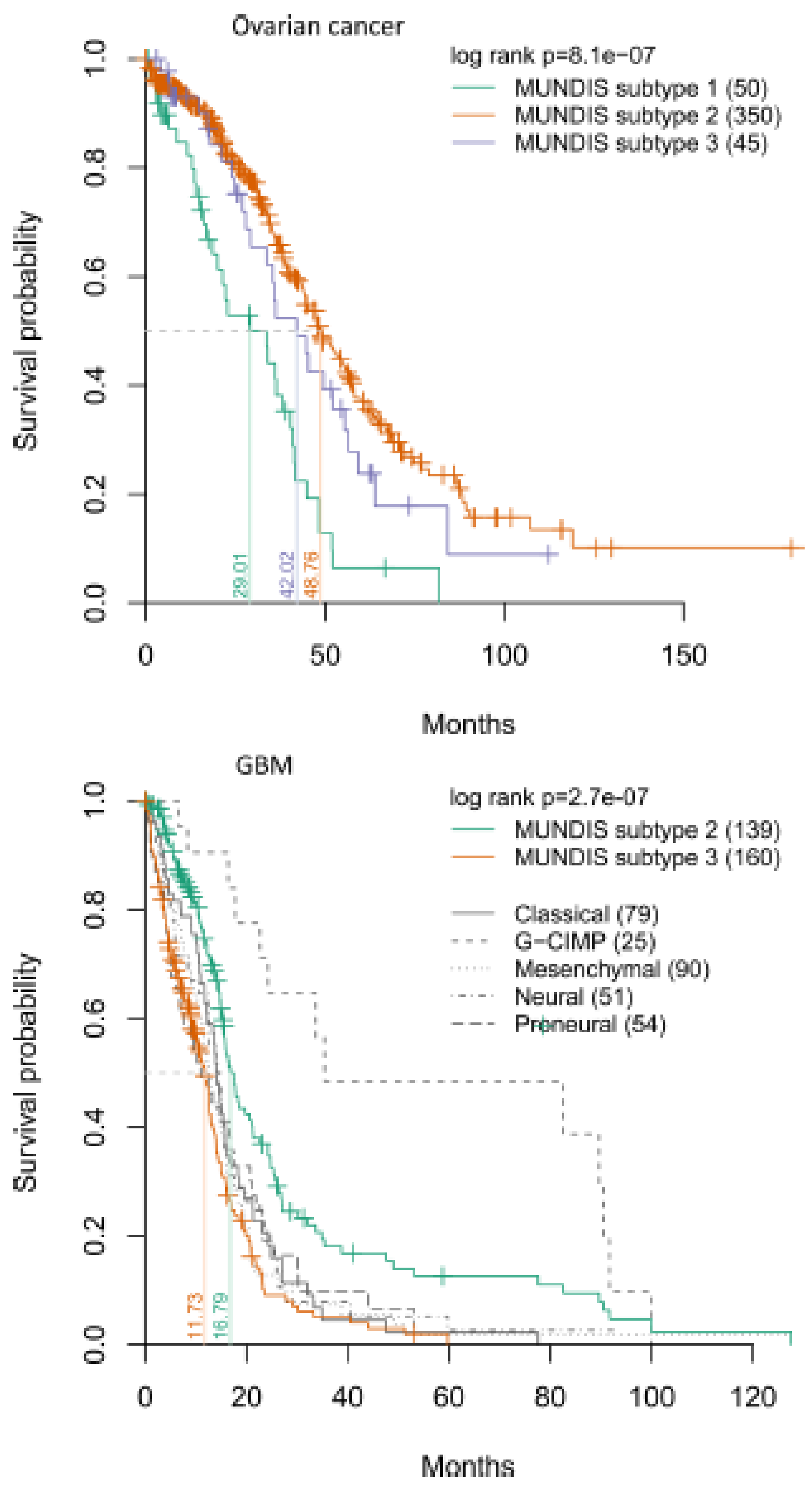


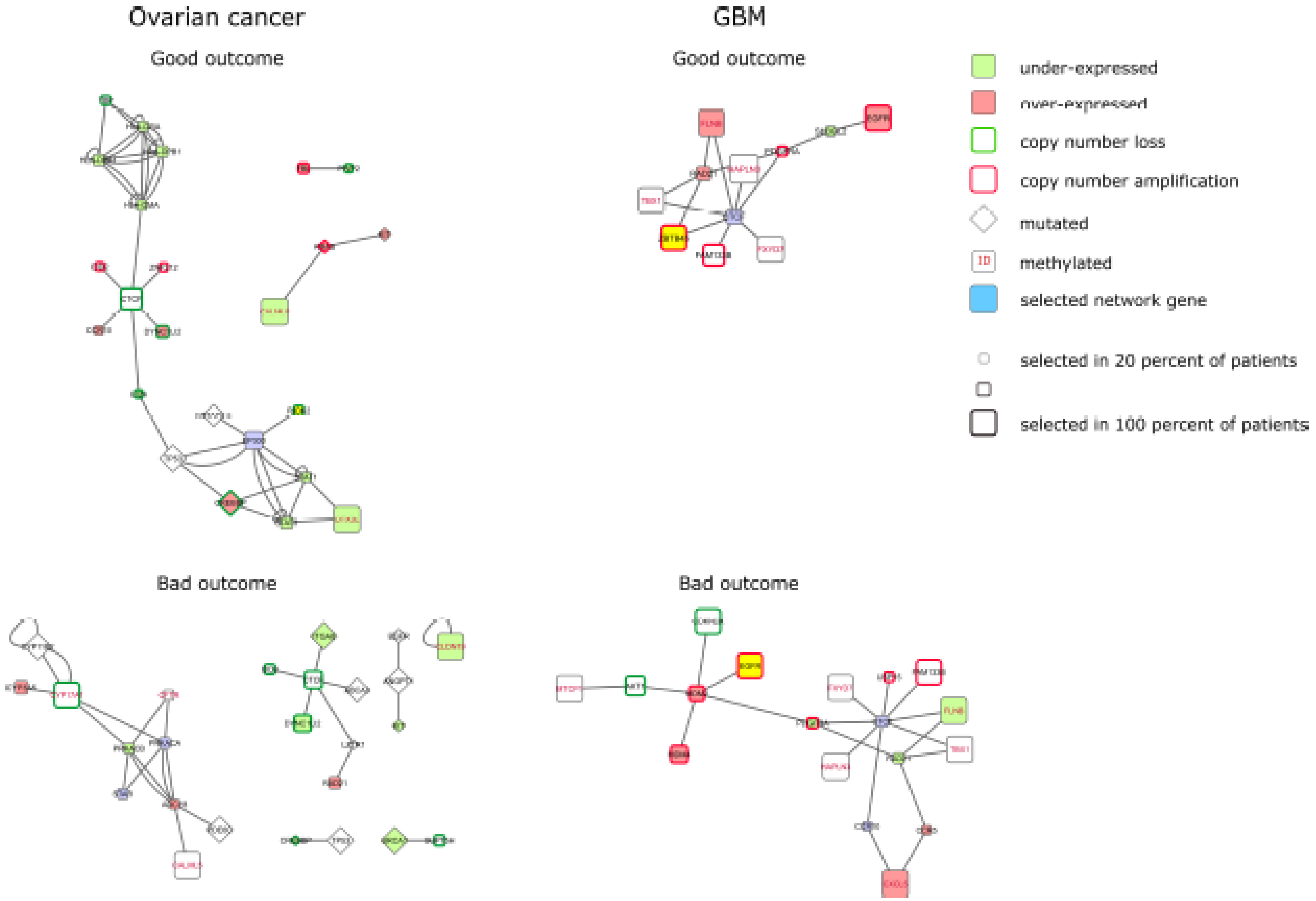




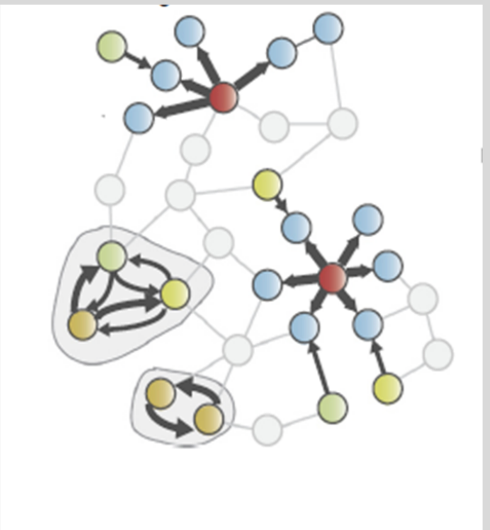
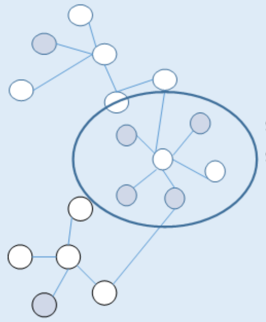
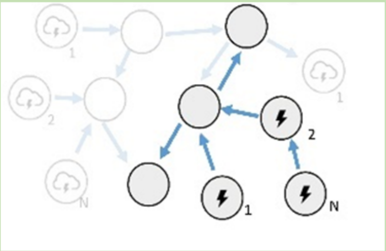
Network-based analysis of omics data

Results provide insight in the mechanism of the disease

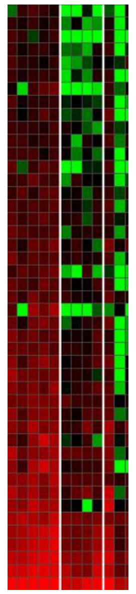




Network-based analysis of omics data

Network model		Gene centric information: Mutation	Gene centric information: Epigenetic	Gene centric information: Expression	Gene set information: Mutual exclusivity	Network properties
Network propagation 	Hotnet[2]	✓				undirected, simple
	Hofree[1]	✓		(✓)		undirected, simple
	Le Van[1]	✓		(✓)		undirected, simple
	NBDI[1,3]	✓	✓	✓		undirected, complex
	NetICS[1,3]	✓	✓	✓		undirected, complex
Significance area extraction 	DriverNet[2, 3]	✓		✓		undirected, complex
	NetSig[2]	✓				undirected, simple
	SSA.ME[2]				✓	undirected, simple
	MEMO[2]				✓	undirected, simple
	MUTEX[2]				✓	undirected, simple
	MEMCOVER[2]				✓	undirected, simple
(Probabilisitic) path-finding 	PARADIGM*	✓		✓		directed, complex
	FAME	✓			✓	undirected, simple
	Phenetic	✓		✓	✓	directed, simple

Case 1: Network-guided data interpretation



Gene list

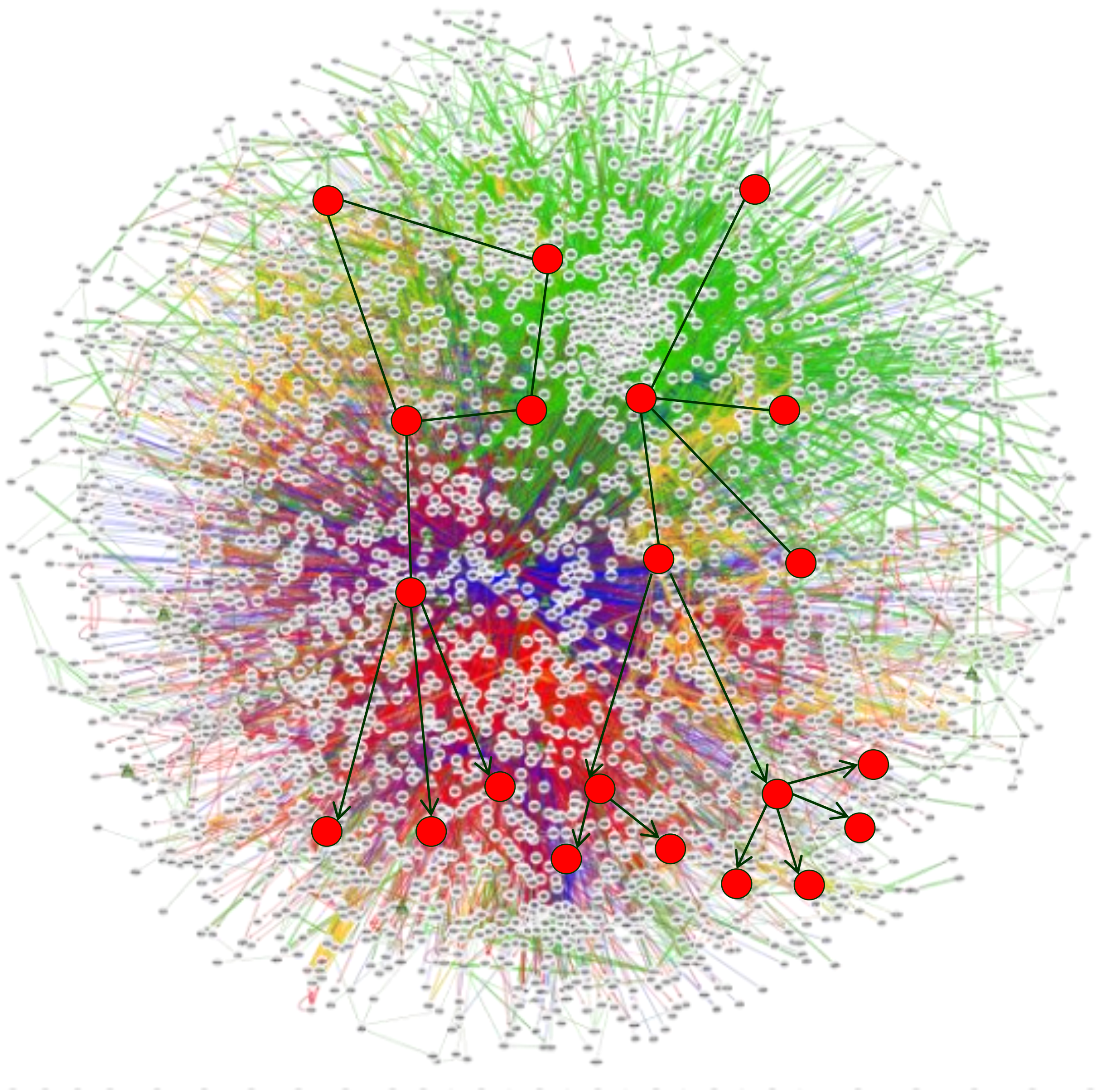
…

Gene 1

Gene 4

Gene 57

Gene n

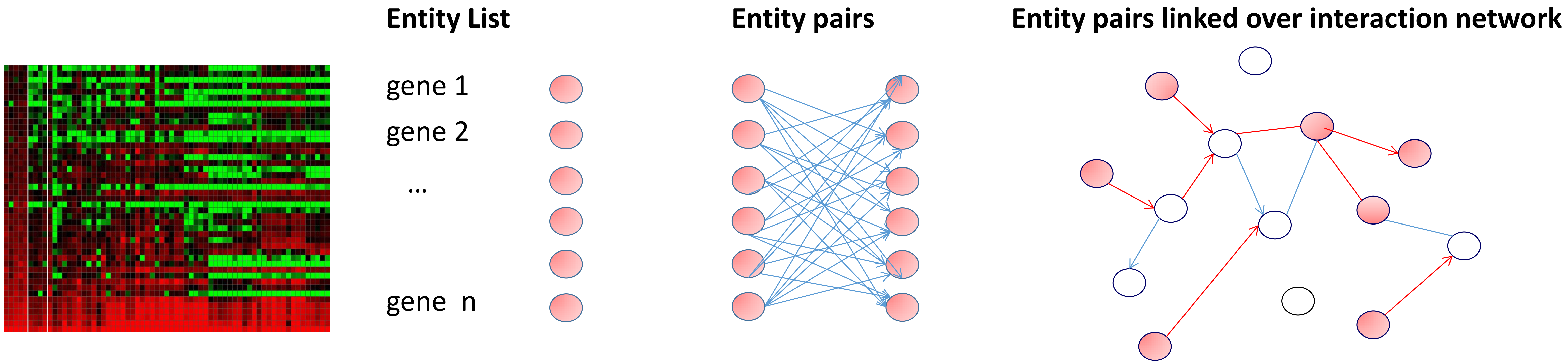


- Map the genes on the interaction network
- Connect the genes on the interaction network and extract relevant edges

Non trivial because the network is overconnected

Case 1: Network-guided data-interpretation

Phenetic: probabilistic subnetwork selection



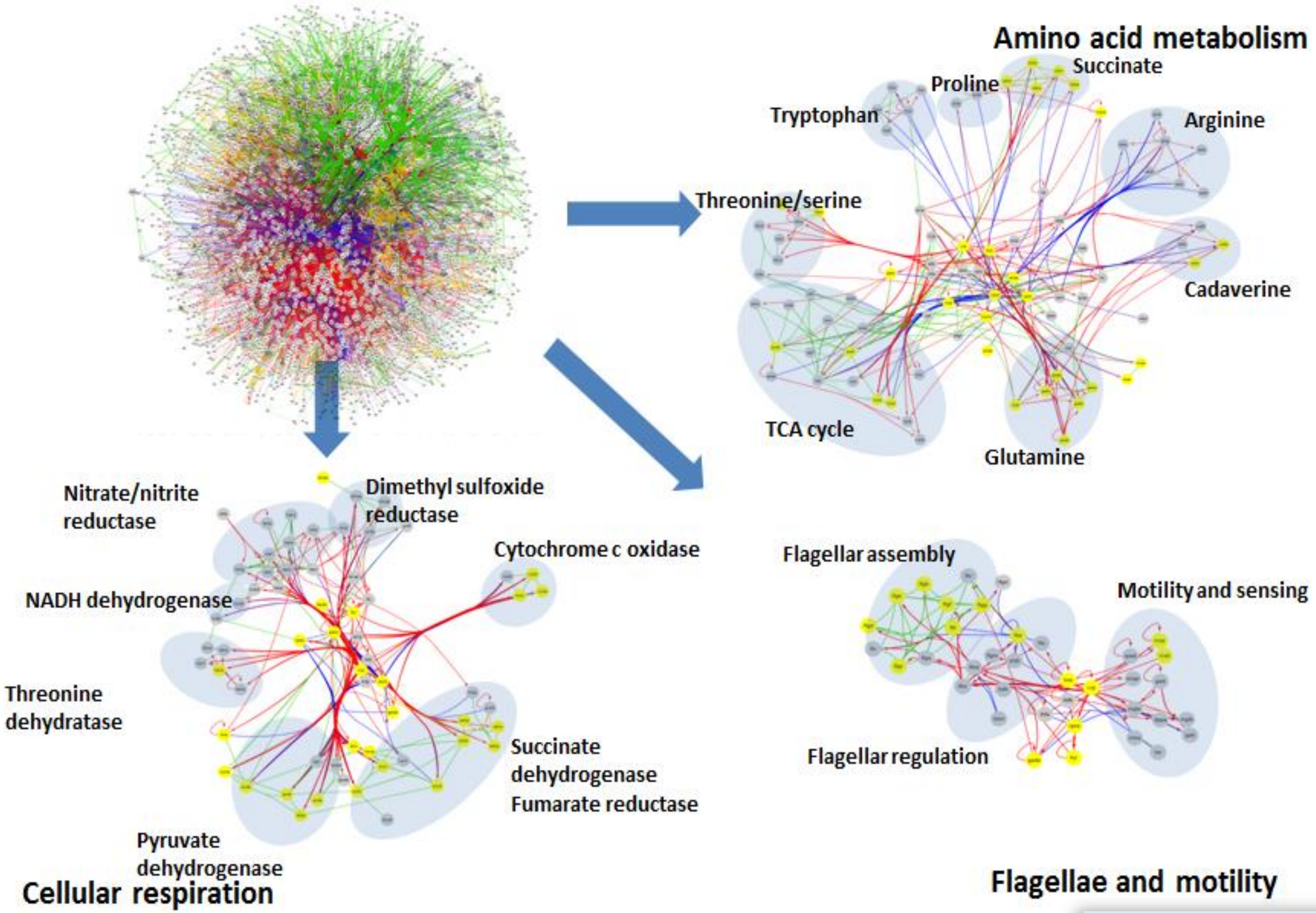
$$S(K) = \sum_{i}^{n} \left(\sum_{j}^{l} \left(P\left(path(C_{i,j}, A_i) \middle| Q_i, K\right) \right) \right) - |K| * x_e$$

De Maeyer et al., Mol Biosyst. 2013;
De Maeyer et al. NAR 2015
De Maeyer. Genome Biol Evol. 2016
Swing et et al. Mol Biol Evol. 2017

Webserver: http://bioinformatics.intec.ugent.be/phenetic/#/index

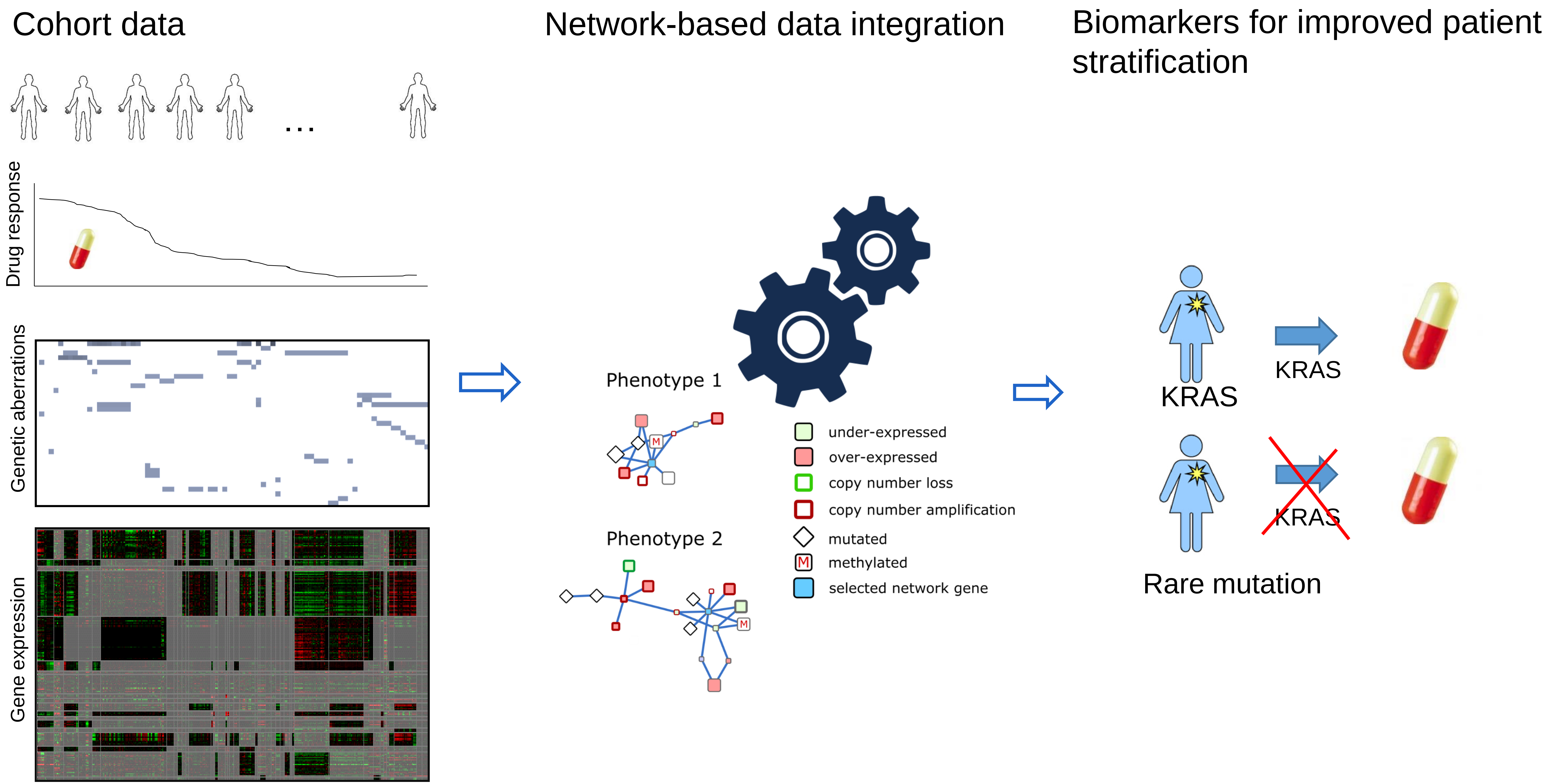
Case 1: Network-guided data-interpretation

Phenetic: probabilistic subnetwork selection



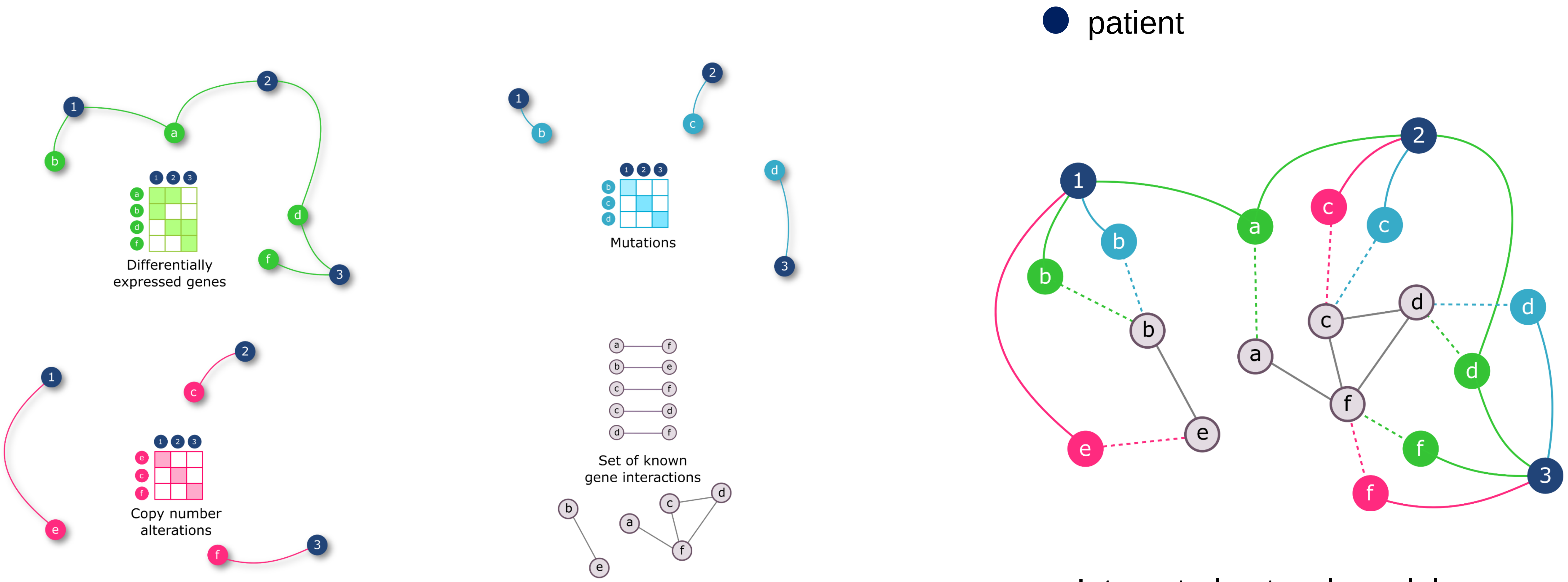
De Maeyer et al., Mol Biosyst. 2013; NAR 2015

Case 2: Identifying pathway signatures



Iorio F et al. Cell. 2016.166(3):740-54.

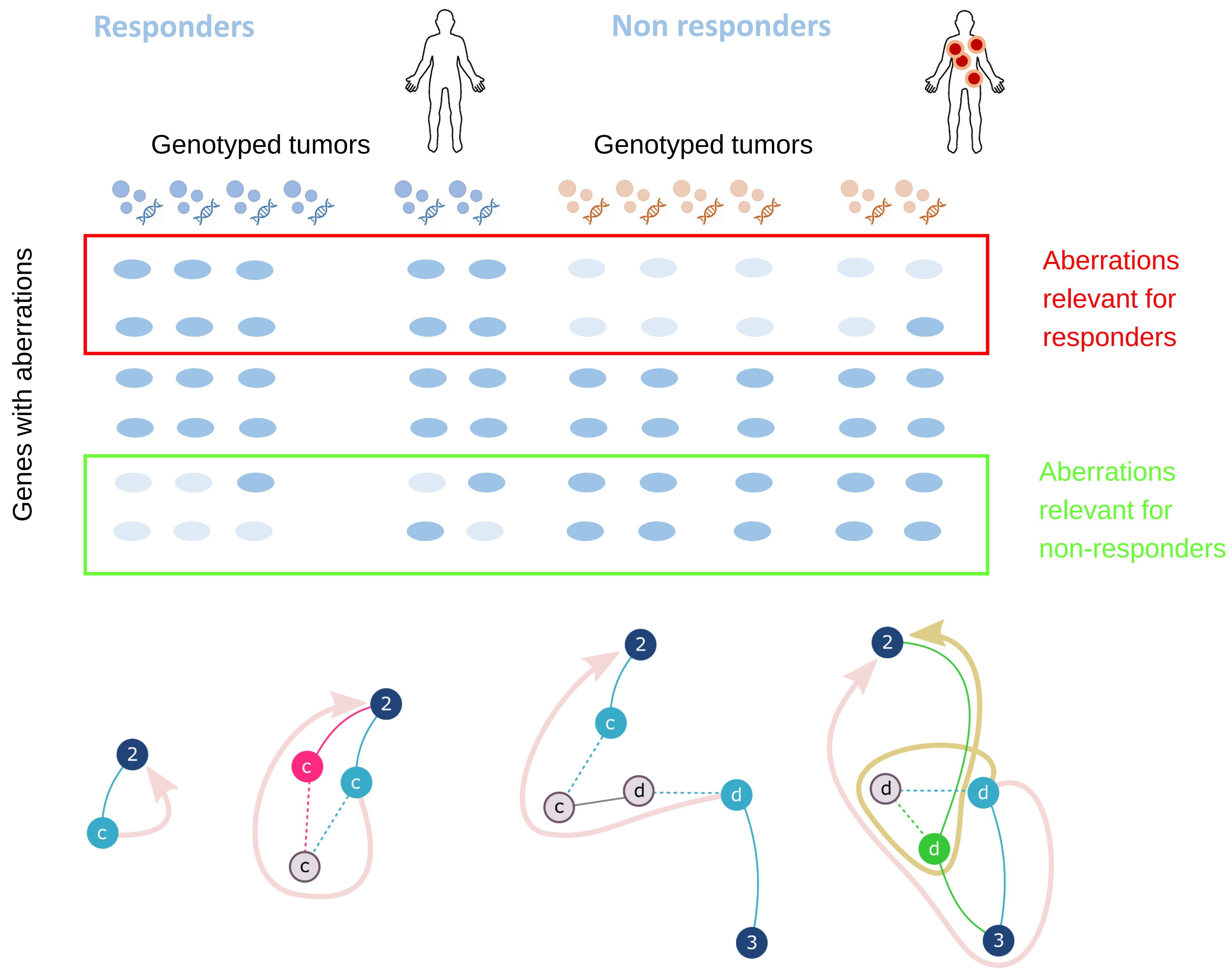
Case 2: Identifying pathway signatures



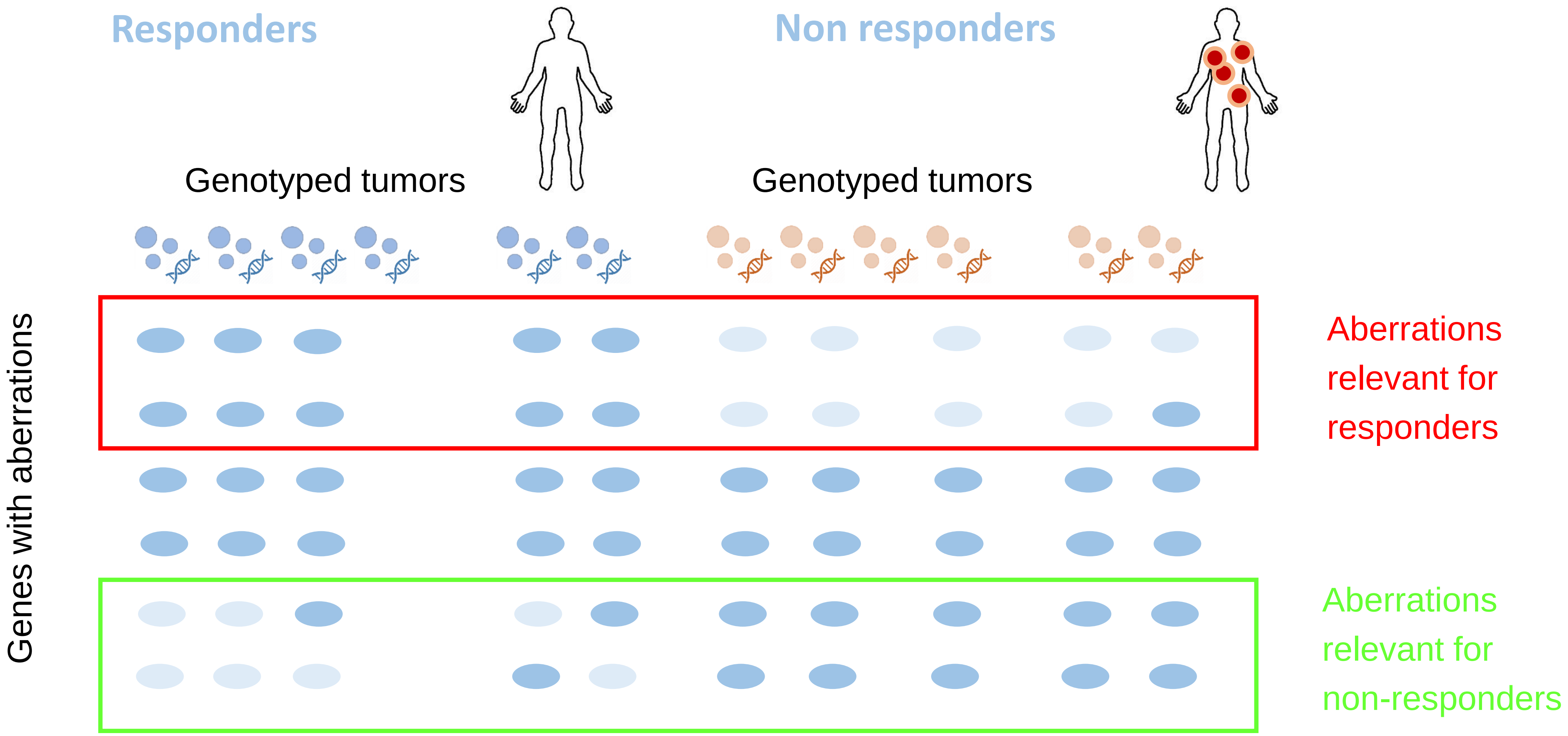
Integrated network model

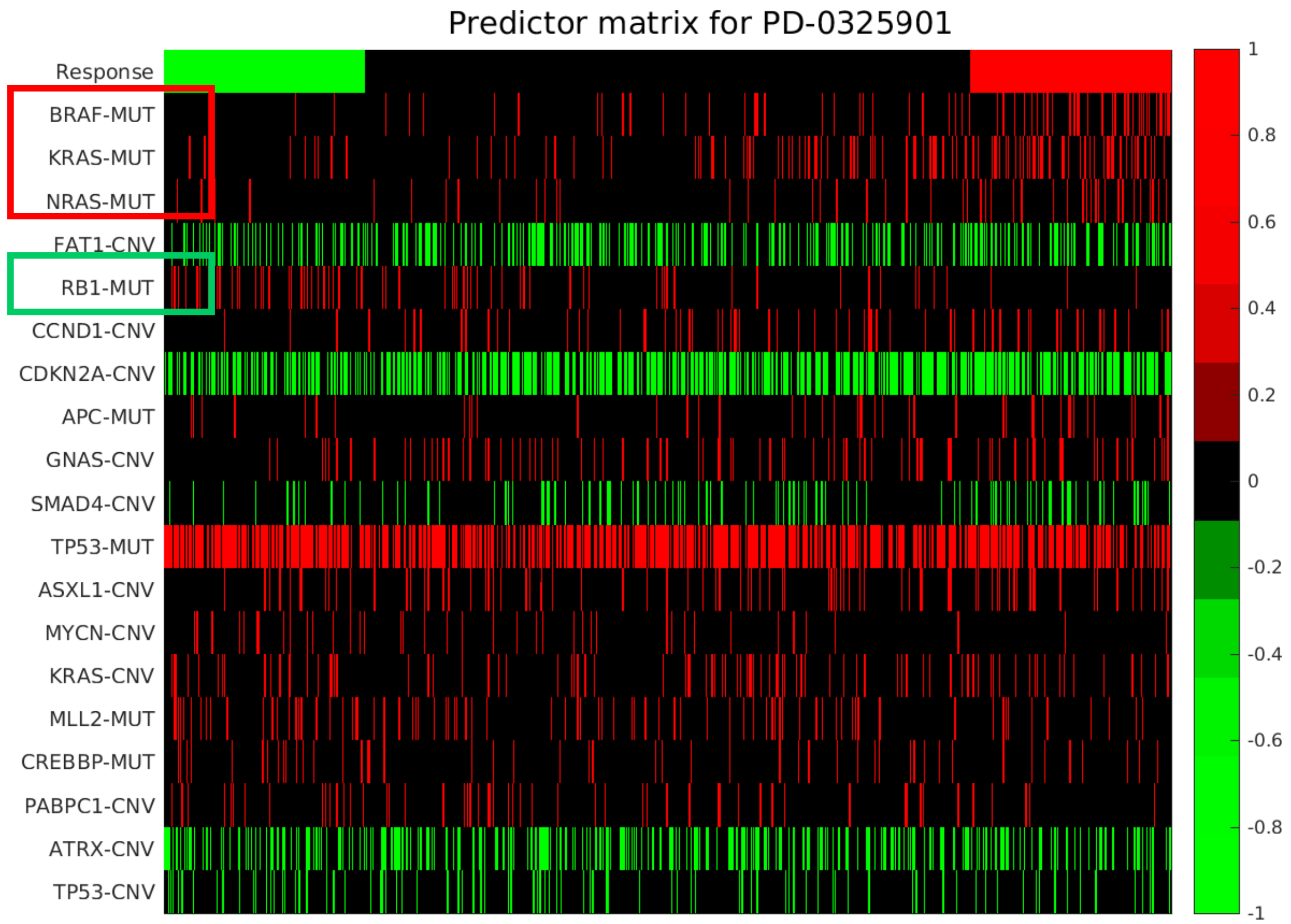
Verbeke et al. Plos One 2015
Used in ICGC PAWG network

Case 2: Identifying pathway signatures



Case 2: Identifying pathway signatures

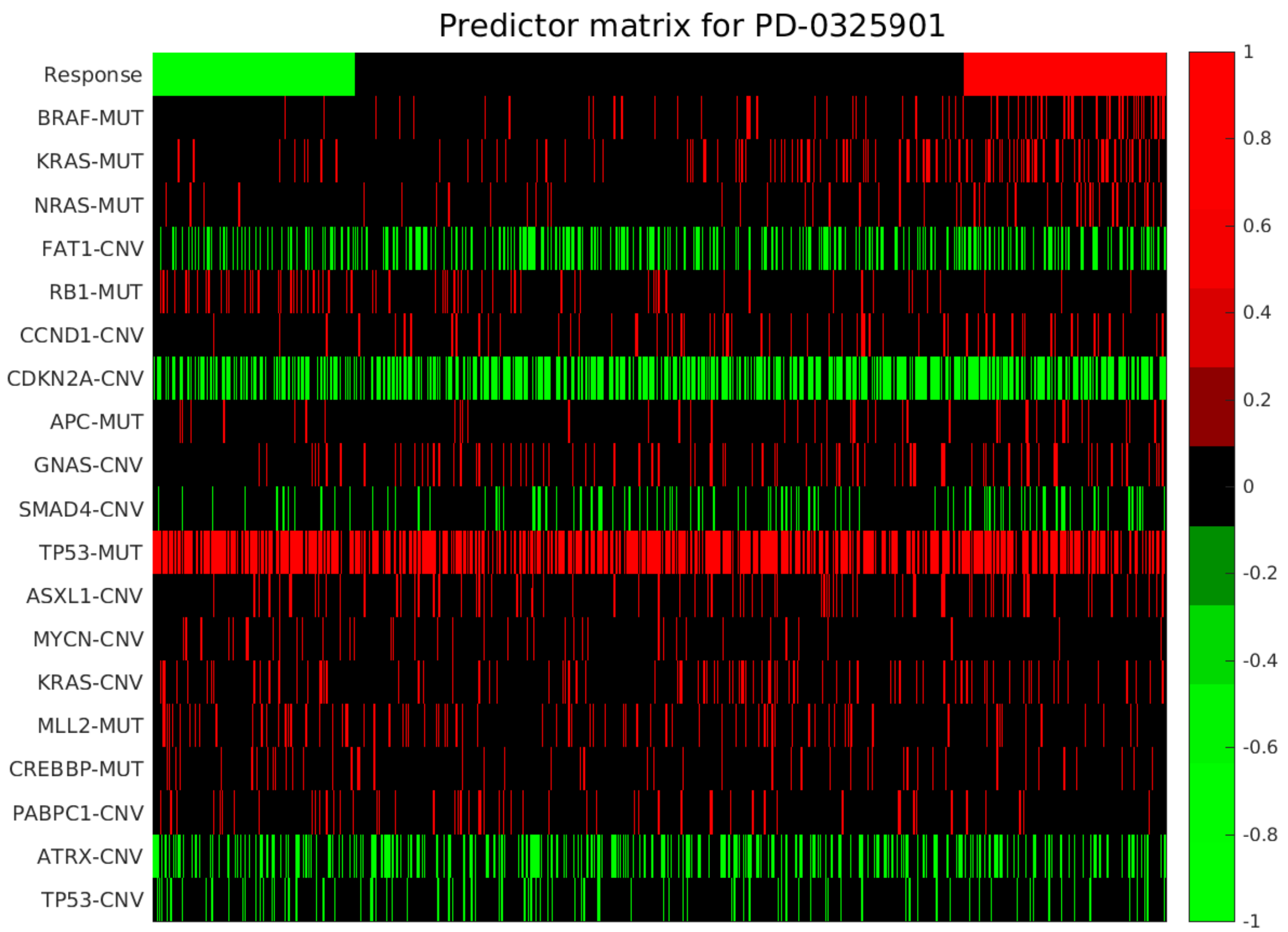


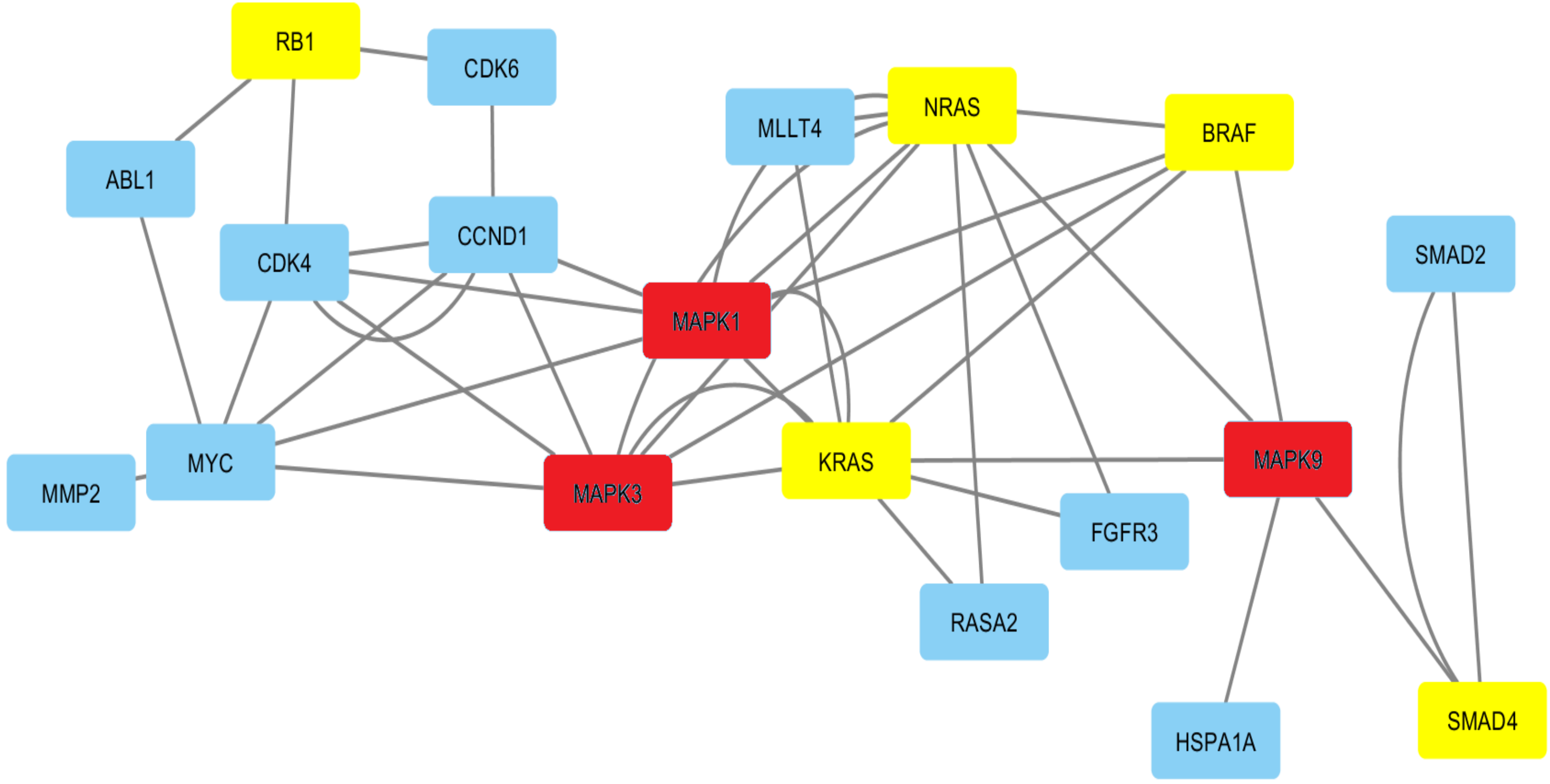


Mushthofa et al. in prep

Case 2: Identifying pathway signatures

Pathway-based biomarkers: network relevant to drug response

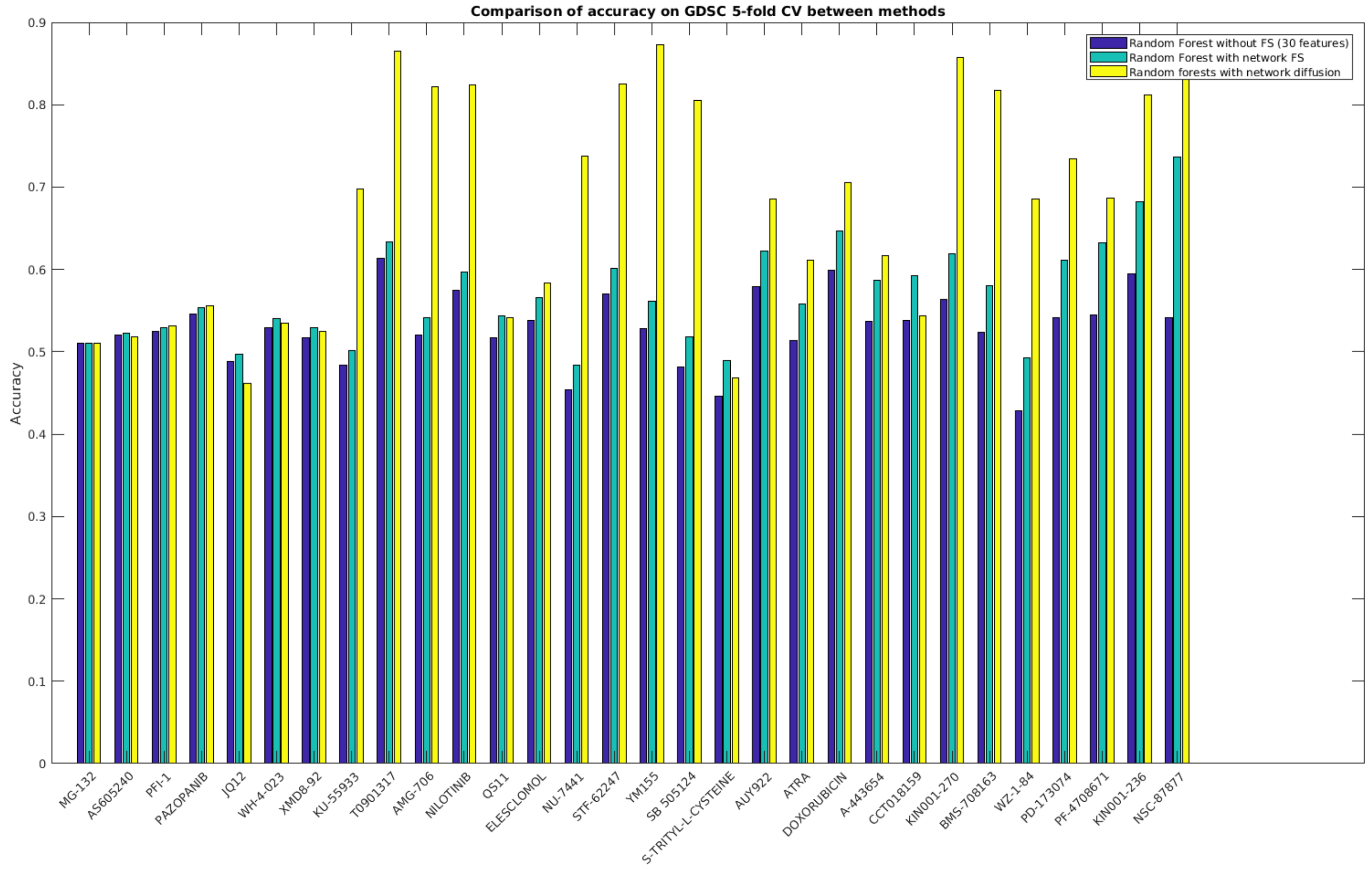




Mushthofa et al. in prep

Case 2: Identifying pathway signatures

Drug Name	Target	Pathway	Feature	Sign	Network ranking	Predictor Ranking
AC220	FLT3	RTK signaling	FLT3-MUT	+	9 (MUTP)	3
AFATINIB	ERBB2, EGFR	EGFR signaling	EGFR-AMP	+	5 (CNVP)	4
AMG-706	VEGFR, RET, c-KIT, PDGFR	RTK signaling	PDGFRA-AMP	+	2 (CNVP)	2
BOSUTINIB	SRC, ABL, TEC	ABL signaling	BCR-ABL-MUT	+	2 (MUTP)	1
DABRAFENIB	BRAF	ERK MAPK signaling	BRAF-MUT	+	1 (MUTP)	1
DASATINIB	ABL, SRC, KIT, PDGFR	ABL signaling	BCR-ABL-MUT	+	3 (MUTP)	2
GDC0941	PI3K (class 1)	PI3K signaling	PIK3CA-MUT	+	1 (NETP)	2
GEFINITIB	EGFR	EGFR signaling	EGFR-MUT	+	1 (NETP)	4
GEFINITIB	EGFR	EGFR signaling	EGFR-AMP	+	1 (NETP)	5
GSK690693	AKT	PI3K signaling	PIK3CA-MUT	+	2 (MUTP)	4
GSK690693	AKT	PI3K signaling	PTEN-MUT	+	1 (MUTP)	6
IMATINIB	ABL, KIT, PDGFR	ABL signaling	BCR-ABL-MUT	+	1 (MUTP)	1
LAPATINIB	ERBB2, EGFR	EGFR signaling	ERBB2-AMP	+	3 (CNVP)	-
MITOMYCIN C	DNA crosslinker	DNA replication	TP53-MUT	+	1 (NETP)	3
NILOTINIB	ABL	ABL signaling	BCR-ABL-MUT	+	1 (MUTP)	1
NUTLIN-3	MDM2	p53 pathway	TP53-MUT	-	1 (NETN)	1
OLAPARIB	PARP1, PARP2	Genome integrity	EWSR1-FLI1-MUT	+	1 (MUTP)	1
PD-0332991	CDK4, CDK6	cell cycle	RB1-DEL	-	1 (CNVN)	3
PLX4720	BRAF	ERK MAPK signaling	BRAF-MUT	+	1 (MUTP)	1
TRAMETINIB	MAP2K1 (MEK1), MAP2K2 (MEK2)	ERK MAPK signaling	BRAF-MUT	+	2 (MUTP)	1
TRAMETINIB	MAP2K1 (MEK1), MAP2K2 (MEK2)	ERK MAPK signaling	KRAS-MUT	+	1 (MUTP)	2
TRAMETINIB	MAP2K1 (MEK1), MAP2K2 (MEK2)	ERK MAPK signaling	NRAS-MUT	+	8 (NETP)	3



Mushthofa et al. in prep

Conclusion

- Omics data integration offers a huge potential for phenotypic analysis, but data-analysis is not trivial

- Network-based data integration methods are tuned towards the integration of molecular data by using biologically relevant assumptions
  - Increases the power of the analysis
  - Insight into the mode of action
  - Intuitive scaffold to integrate omics data

Acknowledgements

UGENT/Network group

- *Dries De Maeyer*
- *Sergio Pulido*
- Bram Weytjens
- Lieven Verbeke
- Mushthofa Mushthofa
- Maarten Larmuseau
- *Jimmy Van den Eynden*

KUL/Computer science

- Prof. Luc De Raedt
- *Thanh Levan*

UGent/INTEC

- Prof. Piet Demeester
- Prof. Jan Fostier

ICGC PAWG network group

- Prof. Ben Raphael - Princeton University
- Dr. Matthew Reyna - Princeton University
- Prof. Josh Stuart - UC Santa Cruz

International Cancer Genome Consortium

N2N FROM NUCLEOTIDES TO NETWORKS

IWT

fwo

INTEC

http://bioinformatics.psb.ugent.be/DBN/