Statistical properties of the MetaCore network of protein-protein interactions

Ekaterina Kotelnikova

Clarivate Analytics, Barcelona, Spain

<u>José Lages</u>

jose.lages@utinam.cnrs.fr Équipe de Physique Théorique, Institut UTINAM, CNRS, Université Bourgogne Franche-Comté, Besançon

Klaus Frahm, Dima L. Shepelyansky

Laboratoire de Physique Théorique, CNRS, Université Paul Sabatier, Toulouse

French Regional Conference on Complex Systems FRCCS 2021 Dijon, France







Projects



ApliGoogle project (2016-2018) funded by MASTODONS CNRS Mission interdisciplinarité Partners : LPT, CNRS, UPS, Toulouse / UTINAM, CNRS, UBFC, Besançon / I. Curie, Inserm, PSL, Paris / IRIT, CNRS, UPS, Toulouse



APEX project (2017-2020) funded by the Bourgogne Franche-Comté region council.





GNETWORKS project (2018-2021) funded by ISITE-UBFC (PIA ANR).

REpTILs project (2020-2023) funded by the Bourgogne Franche-Comté region council. Partners: UTINAM, CNRS, UBFC, Besançon / IHGT, Inserm, UBFC, Besançon / PEPITE, UFC

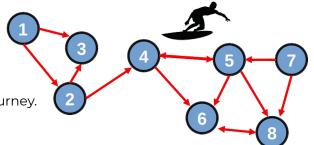
Projects devoted to the physical analysis of complex networks and the application of Google matrix based analysis to complex systems.

From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the random surfer jumps randomly on an adjacent node and continue its journey.

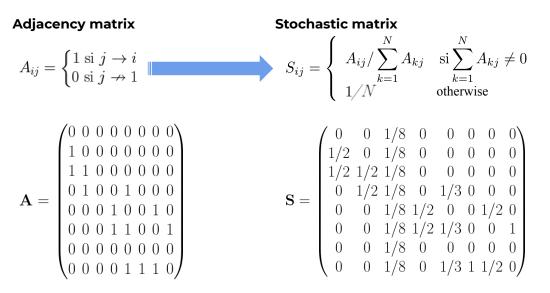
Adjacency matrix

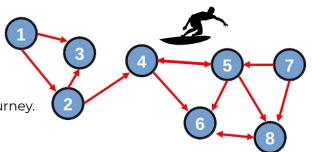
$$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \not\to 1 \end{cases}$$



From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the random surfer jumps randomly on an adjacent node and continue its journey.





From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the random surfer jumps randomly on an adjacent node and continue its journey.

Adjacency matrix	Stochastic matrix	Google matrix
$A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \not\to 1 \end{cases}$	$S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si} \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1 / N & \text{otherwise} \end{cases}$	$G_{ij} = \alpha S_{ij} + (1-\alpha)/N$ with $0.5 < \alpha < 1$ Perron-Frobenius operator
$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 &$	$\mathbf{S} = \begin{pmatrix} 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/8 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 1/8 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/8 & 1/2 & 1/3 & 0 & 0 & 1 \\ 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/8 & 0 & 1/3 & 1 & 1/2 & 0 \end{pmatrix}$	$\mathbf{G} = \begin{pmatrix} 1/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 17/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 17/40 & 17/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 & 1/40 \\ 1/40 & 17/40 & 1/8 & 17/40 & 7/24 & 1/40 & 1/40 & 1/40 \\ 1/40 & 1/40 & 1/8 & 17/40 & 7/24 & 1/40 & 1/40 & 33/40 \\ 1/40 & 1/40 & 1/8 & 17/40 & 7/24 & 1/40 & 1/40 & 33/40 \\ 1/40 & 1/40 & 1/8 & 1/40 & 1/40 & 1/40 & 1/40 \\ 1/40 & 1/40 & 1/8 & 1/40 & 7/24 & 33/40 & 17/40 & 1/40 \\ \end{pmatrix}$ $\alpha = 0.8$

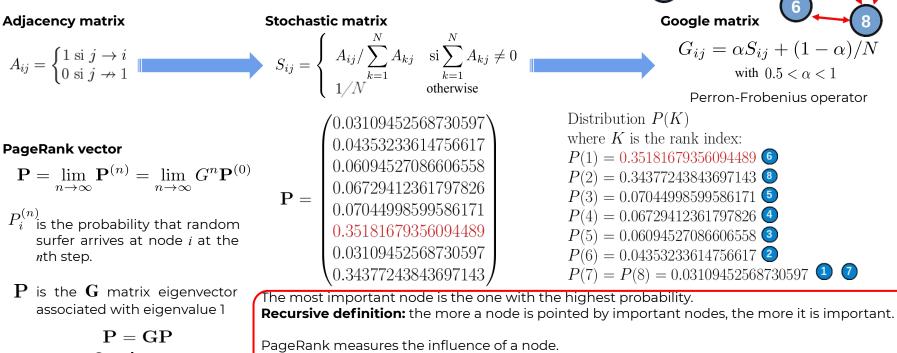
3

2

5

From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the random surfer jumps randomly on an adjacent node and continue its journey.



Steady-state

PageRank measures the influence of a hode. PageRank was (is ?) at the heart of Google search engine (Brin, Page '98).

From Markov (1906) to Brin & Page (1998)

Markovian process : a random surfer probe the structure of a directed network. A each step, the random surfer jumps randomly on an adjacent node and continue its journey.

Adjacency matrix

Stochastic matrix $A_{ij} = \begin{cases} 1 \text{ si } j \to i \\ 0 \text{ si } j \to 1 \end{cases} \qquad \qquad S_{ij} = \begin{cases} A_{ij} / \sum_{k=1}^{N} A_{kj} & \text{si } \sum_{k=1}^{N} A_{kj} \neq 0 \\ 1 / N & \text{otherwise} \end{cases}$

Google matrix

$$G_{ij} = \alpha S_{ij} + (1 - \alpha) / N$$

with $0.5 < \alpha < 1$

6

Perron-Frobenius operator

PageRank vector

$$\mathbf{P} = \lim_{n \to \infty} \mathbf{P}^{(n)} = \lim_{n \to \infty} G^n \mathbf{P}^{(0)}$$

 $P_i^{(n)}$ is the probability that random surfer arrives at node *i* at the *n*th step.

\mathbf{P} is the \mathbf{G} matrix eigenvector associated with eigenvalue 1

 $\mathbf{P} = \mathbf{G}\mathbf{P}$ Steady-state The most important node is the one with the highest probability. **Recursive definition:** the more a node is pointed by important nodes, the more it is important.

PageRank measures the influence of a node. PageRank was (is ?) at the heart of Googe search engine (Brin, Page '98).

CheiRank vector $P^* = G^*P^*$

Similar to the PageRank vector for the network with inverted links. With inverted adjacency matrix elements $A_{ij}^* = A_{ji}$ it is possible to define the stochastic matrix elements $S_{ij}^* \neq S_{ji}$, and the Google matrix elements $G_{ij}^* \neq G_{ji}$ associated to the inverted network (Fogaras '03, Chepelianskii '10).

Recursive definition: the more a node points toward important nodes, the more it is important.

The CheiRank measures the diffusion/the communication of a node.

Statistical properties of the multi-omic MetaCore network

Top 40 PageRank "Process outcomes"

Top 40 CheiRank "Process triggers"

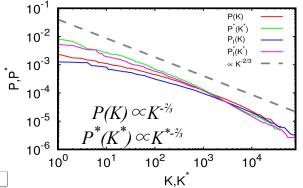
K	P(K) (10 ⁻²)	k	M(K)	Name	Class	Localization	<i>K</i> *	$P^*(K^*)$ (10 ⁻²)	k*	M(K*)	Name	Class	Localization
1	0.2506		0	H ⁺ cytosol	Inorganic ion	Cytosol	1	1.1464	1	0.006332	c-Mvc	Transcription factor	Nucleus
2	0.2376		0	Na ⁺ cvtosol	Inorganic ion	Cytosol	2	0.8172		0.035667	elF2C2 (Argonaute-2)	Generic enzyme	Cytoplasm
3	0.1741		-0.045970	Beta-catenin	Generic binding protein	Cytoplasm	3	0.6722		-0.174071	IGF2BP3	Generic binding protein	Cytoplasm
4	0.1701	1	-0.028308	p53	Transcription factor	Nucleus	4	0.4890		0.680968	Ubiquitin	Generic binding protein	Cytoplasm
	0.1469		0.256018	c-Src	Protein kinase	Cytoplasm	5	0.3719		0.110759	SOX9	Transcription factor	Nucleus
6	0.1435		0.708154	mRNA intracellular	RNA	Intracellular	6	0.3529	2	-0.028308	p53	Transcription factor	Nucleus
7	0.1352		0	H ⁺ extracellular region	Inorganic ion	Extracellular region	7	0.3373	-	0.228978	c-Fos	Transcription factor	Nucleus
8	0.1189	2	0.105603	EGER	Receptor with enzyme activity	Plasma membrane	8	0.3276		0.2203/0	CUX1 (p110)	Transcription factor	Nucleus
9	0.1180	-	-0.014278	DNA	DNA	Nucleus	9	0.2989		-0.057557	SP1	Transcription factor	Nucleus
10	0.1125	3	-0.004135	ESR1 (nuclear)	Transcription factor	Nucleus					ESR1 (nuclear)		Nucleus
11	0.1125	-	0	K ⁺ extracellular region	Inorganic ion	Extracellular region	10	0.2770		-0.004135 0.002825		Transcription factor	Nucleus
12	0.1056		0	ADP cytoplasm	Compound	Cytoplasm	11	0.2769	4	-0.010911	RelA (p65 NF-kB subunit)	Transcription factor	
13	0.1023	4	0.250910	STAT3	Transcription factor	Nucleus	12				elF2C1 (Argonaute-1)	Generic binding protein	Cytoplasm
14	0.0997		0.062046	Androgen receptor	Transcription factor	Nucleus	13	0.2354		0.062046	Androgen receptor	Transcription factor	Nucleus
15	0.0947		0.287801	Rac1	RAS superfamily	Cytoplasm	14	0.2350		-0.045970	Beta-catenin	Generic binding protein	Cytoplasm
16	0.0946		0.207.001	PO ₃ ⁻ cytoplasm	Compound	Cytoplasm	15	0.2330		-0.075622	BRD4	Generic binding protein	Nucleus
17	0.0940	5	0.006332	c-Myc	Transcription factor	Nucleus	16	0.2308		0.153950	Oct-3/4	Transcription factor	Nucleus
18	0.0940	6	0.360271	FAK1	Protein kinase	Cytoplasm	17	0.2259		-0.001577	PUM2	Generic binding protein	Cytoplasm
19	0.0899	0	0.962815	cytosol K ⁺ → extracellular re-	Reaction	NA	18	0.2239		0.188479	EZH2	Generic enzyme	Nucleus
19	0.0099		0.902015	gion K ⁺	neaction	IN/A	19	0.2193		0.208146	p300	Generic enzyme	Nucleus
20	0.0889	7	0.003377	ESR2 (nuclear)	Transcription factor	Nucleus	20	0.2072		-0.407833	TUG1	RNA	Cytoplasm
20	0.0884		0.003377	K ⁺ cytosol	Inorganic ion	Cytosol	21	0.2072		-0.118501	E2F1	Transcription factor	Nucleus
22	0.0849	8	0.002825	RelA (p65 NF-kB subunit)	Transcription factor	Nucleus	22	0.2062		0	ASCC2	Generic binding protein	Nucleus
23	0.0849	9	0.002825	ARX	Transcription factor	Cytoplasm	23	0.2005		0	LIMR	Generic receptor	Plasma membrane
24	0.0828	10	0.208984	ITGB1	Generic receptor	Plasma membrane	24	0.1903		0.148471	BRG1	Generic enzyme	Nucleus
25	0.0828	11	0.548888	SHP-2	Protein phosphatase	Cytoplasm	25	0.1871	5	0.250910	STAT3	Transcription factor	Nucleus
26	0.0776	12	0.364614	GRB2	Generic binding protein	Cytoplasm	26	0.1811		0.381258	RBM24	Generic binding protein	Cytoplasm
27	0.0760	12	0.479956	PI3K reg class IA (p85)	Generic binding protein	Cytoplasm	27	0.1789		0.746981	SUMO-1	Generic binding protein	Nucleus
28	0.0759		-0.114311	E-cadherin	Generic binding protein	Plasma membrane	28	0.1728		0.140357	c-IAP2	Generic binding protein	Cytoplasm
20	0.0759		0.757892	$CO_2 + H_2O \rightarrow H^+ + HCO_2^-$	Reaction	NA	29	0.1699		0.038221	HIF1A	Transcription factor	Nucleus
30	0.0753		-0.098664	p21	Generic binding protein	Nucleus	30	0.1677		0	Zn ²⁺ cytosol	Inorganic ion	Cytosol
31	0.0752		0.148707	Caveolin-1	Generic binding protein	Cytoplasm	31	0.1623		-0.013644	CDK9	Protein kinase	Cytoplasm
32	0.0752		0.007470	Caveolin-1 Ca ²⁺ cvtosol	Inorganic ion	Cytoplasm	32	0.1587		-0.223816	MeCP2	Generic binding protein	Nucleus
32	0.0749		0.381345	PI3K reg class IA (p85-alpha)	Generic binding protein	Cytoplasm	33	0.1533		-0.053592	ELAVL1 (HuR)	Generic binding protein	Nucleus
33	0.0744		-0.220751	Bcl-2		Mitochondrion	34	0.1497		0.120649	HDAC1	Generic enzyme	Nucleus
34	0.0727		-0.220751	CI ⁻ intracellular	Generic binding protein Inorganic ion	Intracellular	35	0.1473		-0.034082	BRD7	Generic binding protein	Nucleus
				MDM2		Nucleus	36	0.1473		0.131956	CREB1	Transcription factor	Nucleus
36 37	0.0712		-0.208082 -0.169004	PTEN	Generic enzyme		37	0.1449		0.131350	Zn ²⁺ nucleus	Inorganic ion	Nucleus
	0.0707		-0.169004 0.391984	PTEN PPAR-gamma	Lipid phosphatase Transcription factor	Cytoplasm Nucleus	37	0.1449		0.096830	SUMO-2	Generic binding protein	
38											SUMO-2 BRD2	Protein kinase	Cytoplasm
39	0.0698		0.031543	ACTB Acetyl-CoA intracellular	Generic binding protein	Cytoplasm Intracellular	39	0.1400		-0.051730			Cytoplasm
40	0.0679		0	Acetyl-GoA intracellular	Compound	intraceitular	40	0.1343		0.228824	C/EBPbeta	Transcription factor	Nucleus

Statistical properties of the multi-omic MetaCore network

The PageRank P(K) and CheiRank $P^*(K^*)$ probabilities have similar distributions $P(K) \propto K^{-\frac{3}{2}}$ and $P(K^*) \propto K^{*-\frac{1}{3}}$. On average, equal importance of ingoing and outgoing links in the MetaCore network.

Top 40 PageRank "Process outcomes"

Top 40 CheiRank
"Process triggers"



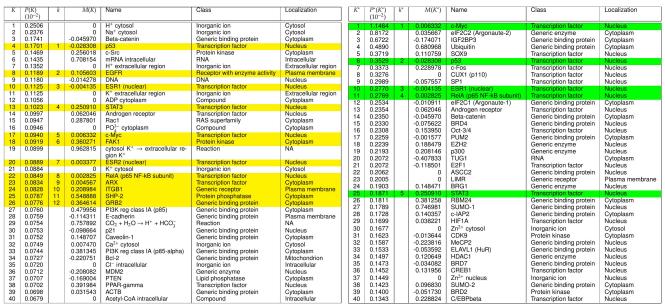
1 0.25 2 0.23 3 0.17 4 0.17 5 0.14 6 0.14 7 0.13 8 0.11 9 0.11	376 741 701 469 435 352 189		0 0 -0.045970 -0.028308 0.256018 0.708154	H ⁺ cytosol Na ⁺ cytosol Beta-catenin p53	Inorganic ion Inorganic ion Generic binding protein	Cytosol Cytosol	1	1,1464					
5 0.14 6 0.14 7 0.13 8 0.11 9 0.11	469 435 352 189	1	0.256018			Cytoplasm	2	0.8172		0.006332 0.035667 -0.174071	c-Myc elF2C2 (Argonaute-2) IGF2BP3	Transcription factor Generic enzyme Generic binding protein	Nucleus Cytoplasm Cytoplasm
6 0.14 7 0.13 8 0.11 9 0.11	435 352 189				Transcription factor	Nucleus	4	0.4890		0.680968	Ubiquitin	Generic binding protein	Cytoplasm
7 0.13 8 0.11 9 0.11	352 189		0 709154	c-Src	Protein kinase	Cytoplasm	5	0.3719		0.110759	SOX9	Transcription factor	Nucleus
8 0.11 9 0.11	189			mRNA intracellular	RNA	Intracellular	6	0.3529	2	-0.028308	p53	Transcription factor	Nucleus
9 0.11			0	H ⁺ extracellular region	Inorganic ion	Extracellular region	7	0.3373	_	0.228978	c-Fos	Transcription factor	Nucleus
		2	0.105603	EGFR	Receptor with enzyme activity	Plasma membrane	8	0.3276		0	CUX1 (p110)	Transcription factor	Nucleus
10 0.11	180		-0.014278	DNA	DNA	Nucleus	9	0.2989		-0.057557	SP1	Transcription factor	Nucleus
		3	-0.004135	ESR1 (nuclear)	Transcription factor	Nucleus	10	0.2770	3	-0.004135	ESR1 (nuclear)	Transcription factor	Nucleus
11 0.11			0	K ⁺ extracellular region	Inorganic ion	Extracellular region	11	0.2769	4	0.002825	RelA (p65 NF-kB subunit)	Transcription factor	Nucleus
12 0.10			0	ADP cytoplasm	Compound	Cytoplasm	12	0.2534		-0.010911	elF2C1 (Argonaute-1)	Generic binding protein	Cytoplasm
13 0.10		4	0.250910	STAT3	Transcription factor	Nucleus	13	0.2354		0.062046	Androgen receptor	Transcription factor	Nucleus
14 0.09	997		0.062046	Androgen receptor	Transcription factor	Nucleus	14	0.2350		-0.045970	Beta-catenin	Generic binding protein	Cytoplasm
15 0.09	947		0.287801	Rac1	RAS superfamily	Cytoplasm	15	0.2330		-0.075622	BRD4	Generic binding protein	Nucleus
16 0.09	946		0	PO ₄ ³⁻ cytoplasm	Compound	Cytoplasm	16	0.2308		0.153950	Oct-3/4	Transcription factor	Nucleus
17 0.09	940	5	0.006332	c-Mvc	Transcription factor	Nucleus	17	0.2259		-0.001577	PUM2	Generic binding protein	Cytoplasm
18 0.09	919	6	0.360271	FAK1	Protein kinase	Cytoplasm	18	0.2239		0.188479	EZH2	Generic enzyme	Nucleus
19 0.08	899		0.962815	cvtosol K ⁺ → extracellular re-	Reaction	NA	19	0.2239		0.208146	p300		Nucleus
				gion K ⁺				0.2193		-0.407833	TUG1	Generic enzyme RNA	
20 0.08	889	7	0.003377	ESB2 (nuclear)	Transcription factor	Nucleus	20	0.2072			E2F1		Cytoplasm Nucleus
21 0.08	884		0	K ⁺ cytosol	Inorganic ion	Cytosol				-0.118501		Transcription factor	
22 0.08	849	8	0.002825	RelA (p65 NF-kB subunit)	Transcription factor	Nucleus	22	0.2062		0	ASCC2	Generic binding protein	Nucleus
23 0.08	834	9	0.004567	ABX	Transcription factor	Cytoplasm	23	0.2005		~	LIMR	Generic receptor	Plasma membrane
24 0.08	828	10	0.208984	ITGB1	Generic receptor	Plasma membrane	24	0.1903		0.148471	BRG1	Generic enzyme	Nucleus
25 0.07	787	11	0.548888	SHP-2	Protein phosphatase	Cytoplasm	25	0.1871	5	0.250910	STAT3	Transcription factor	Nucleus
26 0.07	776	12	0.364614	GRB2	Generic binding protein	Cytoplasm	26	0.1811		0.381258	RBM24	Generic binding protein	Cytoplasm
27 0.07	760		0.479956	PI3K reg class IA (p85)	Generic binding protein	Cvtoplasm	27	0.1789		0.746981	SUMO-1	Generic binding protein	Nucleus
28 0.07	759		-0.114311	E-cadherin	Generic binding protein	Plasma membrane	28	0.1728		0.140357	c-IAP2	Generic binding protein	Cytoplasm
29 0.07	754		0.757892	$CO_2 + H_2O \rightarrow H^+ + HCO_2^-$	Reaction	NA	29	0.1699		0.038221	HIF1A	Transcription factor	Nucleus
30 0.07	753		-0.098664	p21	Generic binding protein	Nucleus	30	0.1677		0	Zn ²⁺ cytosol	Inorganic ion	Cytosol
31 0.07	752		0.148707	Caveolin-1	Generic binding protein	Cytoplasm	31	0.1623		-0.013644	CDK9	Protein kinase	Cytoplasm
32 0.07			0.007470	Ca ²⁺ cytosol	Inorganic ion	Cytosol	32	0.1587		-0.223816	MeCP2	Generic binding protein	Nucleus
33 0.07			0.381345	PI3K reg class IA (p85-alpha)	Generic binding protein	Cytoplasm	33	0.1533		-0.053592	ELAVL1 (HuR)	Generic binding protein	Nucleus
34 0.07	727		-0.220751	Bcl-2	Generic binding protein	Mitochondrion	34	0.1497		0.120649	HDAC1	Generic enzyme	Nucleus
35 0.07			0	CI ⁻ intracellular	Inorganic ion	Intracellular	35	0.1473		-0.034082	BRD7	Generic binding protein	Nucleus
36 0.07			-0.208082	MDM2	Generic enzyme	Nucleus	36	0.1452		0.131956	CREB1	Transcription factor	Nucleus
37 0.07			-0.169004	PTEN	Lipid phosphatase	Cytoplasm	37	0.1449		0	Zn ²⁺ nucleus	Inorganic ion	Nucleus
38 0.07			0.391984	PPAR-gamma	Transcription factor	Nucleus	38	0.1423		0.096830	SUMO-2	Generic binding protein	Cytoplasm
39 0.06			0.031543	ACTB	Generic binding protein	Cytoplasm	39	0.1400		-0.051730	BRD2	Protein kinase	Cytoplasm
40 0.06			0	Acetyl-CoA intracellular	Compound	Intracellular	40	0.1343		0.228824	C/EBPbeta	Transcription factor	Nucleus

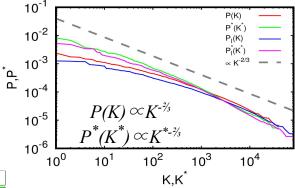
Statistical properties of the multi-omic MetaCore network

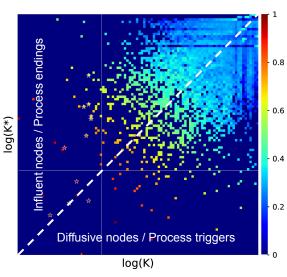
The PageRank P(K) and CheiRank $P^*(K^*)$ probabilities have similar distributions $P(K) \propto K^{\frac{3}{2}}$ and $P(K^*) \propto K^{*-\frac{3}{2}}$. On average, **equal importance of ingoing and outgoing links** in the MetaCore network.

Top 40 PageRank "Process outcomes"

Top 40 CheiRank "Process triggers"







Ising Metacore network: bi-functionality of the interactions

We take into account the nature of the interaction:

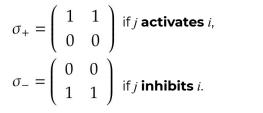
• either,

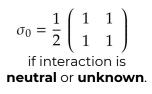
j activates i

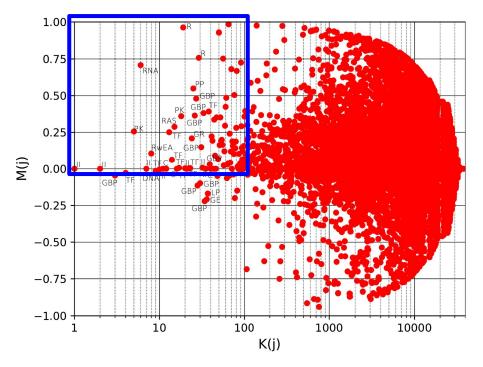
• or, *j* inhibits *i*.

Nodes are doubled, the node *i* is now *i*+ and *i*-.

The matrix element G_{ij} , associated to the link $j \rightarrow i$, is replaced by one of the following 2x2 matrices:







Consequently, $P_{+}(i)$ is the probability that node *i* is activated, and

and

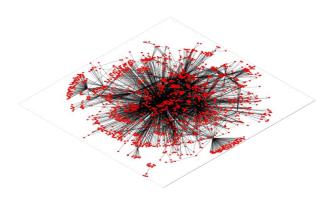
 $P_{i}(i)$ is the probability that node *i* is inhibited.

PageRank Magnetization

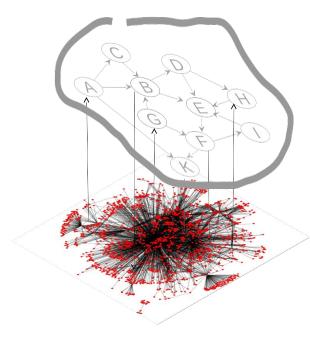
$$M(j) = \frac{P_{+}(j) - P_{-}(j)}{P_{+}(j) + P_{-}(j)}$$
$$-1 \le M(j) \le 1$$

Most of the omics in **top 100 PageRank (K≤100) are more activated** (M>0) than inhibited.

Let us consider a very large network with $N \gg 1$.

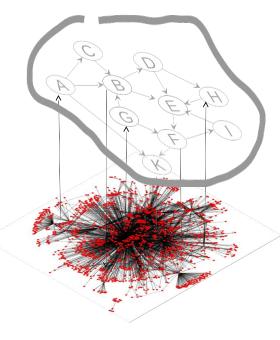


Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest.



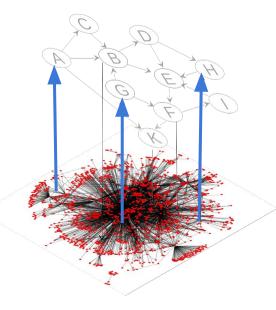
Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$



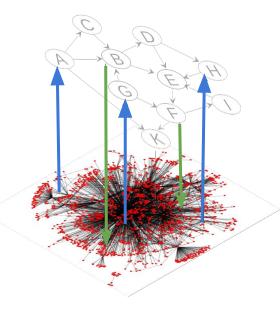
Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$



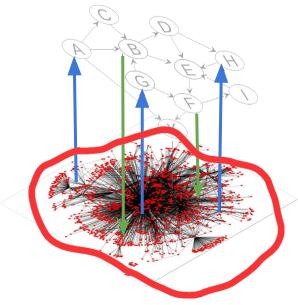
Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$



Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$



Let us consider a very large network with $N \gg I$.

Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

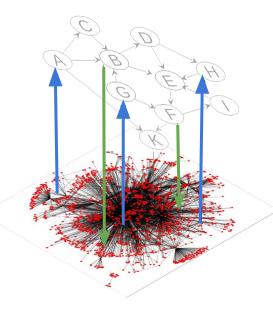
$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

For the global matrix, we have

$$\mathbf{GP} = \mathbf{P}$$

We define the reduced Google matrix G_{R} associated to the N_{r} -size subset of interest such as

$$\mathbf{G}_R \mathbf{P}_r = \mathbf{P}_r$$



Let us consider a very large network with $N \gg I$.

Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

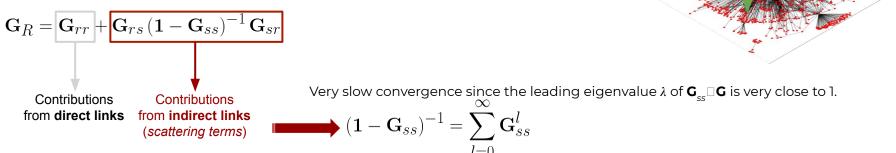
For the global matrix, we have

 $\mathbf{GP} = \mathbf{P}$

We define the reduced Google matrix G_{R} associated to the N_{r} -size subset of interest such as

$$\mathbf{G}_R \mathbf{P}_r = \mathbf{P}_r$$

The reduced Google matrix can be written as



J. Lages, D. Shepelyansky, A. Zinovyev, Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks, PLoS ONE 13(1): e0190812 (2018)

Let us consider a very large network with $N \gg 1$.

Consider a sub-network of $N_r \ll N$ nodes of interest.

The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

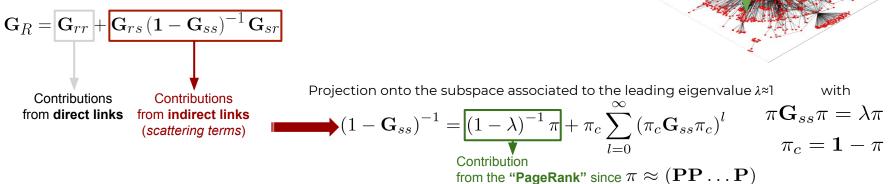
For the global matrix, we have

 $\mathbf{GP}=\mathbf{P}$

We define the reduced Google matrix G_{R} associated to the N_{r} -size subset of interest such as

$$\mathbf{G}_R \mathbf{P}_r = \mathbf{P}_r$$

The reduced Google matrix can be written as



J. Lages, D. Shepelyansky, A. Zinovyev, Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks, PLoS ONE 13(1): e0190812 (2018)

Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The Google matrix of the size N network and the associated PageRank vector can be written as

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{rr} & \mathbf{G}_{rs} \\ \mathbf{G}_{sr} & \mathbf{G}_{ss} \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} \mathbf{P}_r \\ \mathbf{P}_s \end{pmatrix}$$

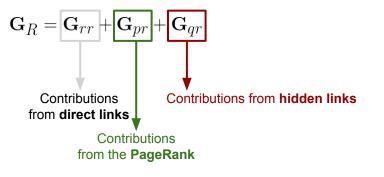
For the global matrix, we have

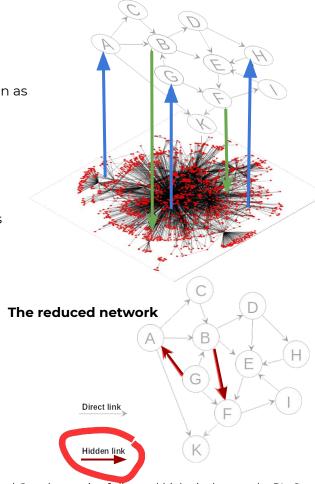
 $\mathbf{GP}=\mathbf{P}$

We define the reduced Google matrix G_{R} associated to the N_{r} -size subset of interest such as

 $\mathbf{G}_R\mathbf{P}_r=\mathbf{P}_r$

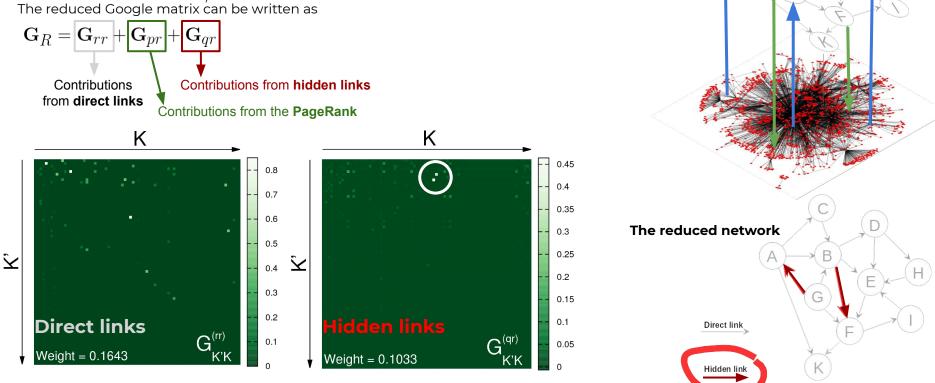
The reduced Google matrix can be written as





J. Lages, D. Shepelyansky, A. Zinovyev, Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks, PLoS ONE 13(1): e0190812 (2018)

Let us consider a very large network with $N \gg I$. Consider a sub-network of $N_r \ll N$ nodes of interest. The reduced Google matrix can be written as

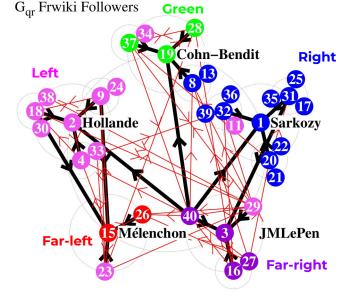


J. Lages, D. Shepelyansky, A. Zinovyev, Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks, PLoS ONE 13(1): e0190812 (2018)

Proof of concept with Wikipedia as a complex network

Hidden links between political leaders Gar Enwiki G20 EN Grr Enwiki G20 EN Obama Obama Putin Putin Cameron Cameron Hidden links Harper Harper Merkel Merkel Singh Singh Jintao Jintao Holland Holland Gillard Gillard Barroso Barroso Zuma Zuma Monti Monti Kirchne Kirchne Erdogan Erdogan Calderó Calderó Abdulla Abdulla Yudhoyo Yudhoyo Roussef Roussef Myung-b Myung-b Noda Noda ard and 10yo 111a 1eró 1eró

Analysis of hidden links between 2012 **G20 leaders** from the English edition Wikipedia (extracted in 2013)

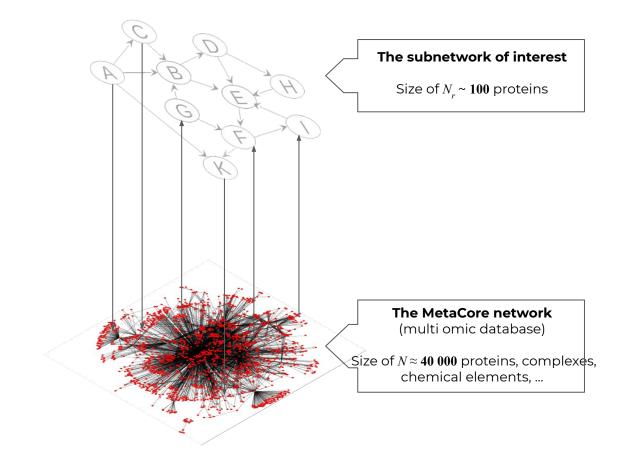


Analysis of hidden links between 2013 **French politics** from the French edition Wikipedia (extracted in 2013)

We retrieve knowledge about known political acquaintances (not trivially stated in Wikipedia).

The reduced Google matrix approach was also used for the **network analysis** of:

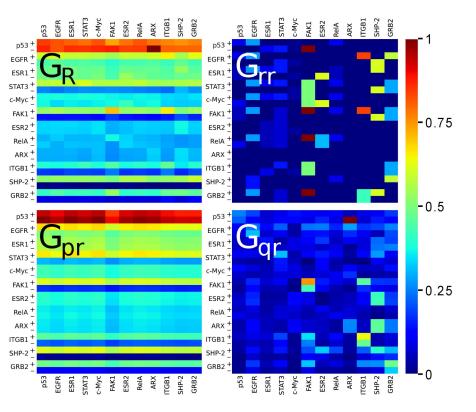
terrorist groups,	pharmaceutical groups,	infectious diseases,	bitcoin transactions,	the world trade,
	(within Wikipedia)		(within corresp. ec	onomical networks)
<u></u>	Ŷ)	<u> </u>	



Top 40 PageRank "Process outcomes"

	K	P(K) (10 ⁻²)	k	M(K)	Name	Class	Localization
	2 3	0.2506 0.2376 0.1741		0 0 -0.045970	H ⁺ cytosol Na ⁺ cytosol Beta-catenin	Inorganic ion Inorganic ion Generic binding protein	Cytosol Cytosol Cytoplasm
		0.1701	1	-0.028308	p53	Transcription factor	Nucleus
		0.1469		0.256018	c-Src	Protein kinase	Cytoplasm
		0.1435		0.708154	mRNA intracellular	RNA	Intracellular
		0.1352		0	H ⁺ extracellular region	Inorganic ion	Extracellular region
	8	0.1189	2	0.105603	EGFR	Receptor with enzyme activity	Plasma membrane
	9	0.1180		-0.014278	DNA	DNA	Nucleus
1	0	0.1125	3	-0.004135	ESR1 (nuclear)	Transcription factor	Nucleus
1	1	0.1125		0	K ⁺ extracellular region	Inorganic ion	Extracellular region
1	2	0.1056		0	ADP cytoplasm	Compound	Cytoplasm
1	3	0.1023	4	0.250910	STAT3	Transcription factor	Nucleus
1.	4	0.0997		0.062046	Androgen receptor	Transcription factor	Nucleus
1		0.0947		0.287801	Rac1	RAS superfamily	Cytoplasm
1		0.0946		0	PO ₄ ³⁻ cytoplasm	Compound	Cytoplasm
1		0.0940	5	0.006332	c-Myc	Transcription factor	Nucleus
1		0.0919	6	0.360271	FAK1	Protein kinase	Cytoplasm
1		0.0899	0	0.962815	cytosol K ⁺ \rightarrow extracellular re-	Reaction	NA
	-				gion K ⁺		
2		0.0889	7	0.003377	ESR2 (nuclear)	Transcription factor	Nucleus
2		0.0884		0	K ⁺ cytosol	Inorganic ion	Cytosol
2		0.0849	8	0.002825	RelA (p65 NF-kB subunit)	Transcription factor	Nucleus
2		0.0834	9	0.004567	ARX	Transcription factor	Cytoplasm
2		0.0828	10	0.208984	ITGB1	Generic receptor	
2							Plasma membrane
2		0.0787	11	0.548888	SHP-2	Protein phosphatase	Cytoplasm
		0.0776	11 12	0.548888 0.364614	SHP-2 GRB2	Protein phosphatase Generic binding protein	Cytoplasm Cytoplasm
	7	0.0776 0.0760		0.548888 0.364614 0.479956	SHP-2 GRB2 PI3K reg class IA (p85)	Protein phosphatase Generic binding protein Generic binding protein	Cytoplasm Cytoplasm Cytoplasm
2	7 8	0.0776 0.0760 0.0759		0.548888 0.364614 0.479956 -0.114311	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein	Cytoplasm Cytoplasm Cytoplasm Plasma membrane
2	7 8 9	0.0776 0.0760 0.0759 0.0754		0.548888 0.364614 0.479956	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction	Cytoplasm Cytoplasm Cytoplasm
2 2 3	7 8 9 0	0.0776 0.0760 0.0759 0.0754 0.0753		0.548888 0.364614 0.479956 -0.114311	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein	Cytoplasm Cytoplasm Cytoplasm Plasma membrane
2	7 8 9 0	0.0776 0.0760 0.0759 0.0754		0.548888 0.364614 0.479956 -0.114311 0.757892	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction	Cytoplasm Cytoplasm Cytoplasm Plasma membrane NA
2	7 8 9 0 1	0.0776 0.0760 0.0759 0.0754 0.0753		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein	Cytoplasm Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm
2 2 3 3 3	7 8 9 0 1 2	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707	SHP-2 GBB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21 Caveolin-1	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein	Cytoplasm Cytoplasm Cytoplasm Plasma membrane NA Nucleus
2 2 3 3 3	7 8 9 0 1 2 3	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21 Caveolin-1 Ca ²⁺ cytosol	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion	Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasol
2 2 3 3 3 3	7 8 9 0 1 2 3 4	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345	SHP-2 GRB2 PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21 Caveolin-1 Ca^{2+} cytosol PI3K reg class IA (p85-alpha)	Protein phosphatase Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion Generic binding protein Generic binding protein Generic binding protein	Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytosol Cytoplasm
2233333333	7 8 9 0 1 2 3 4 5	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751	$\begin{array}{c} SHP-2\\ GPB2\\ Pl3K \text{ reg class IA (p85)}\\ E\text{-cadherin}\\ CO_2 + H_2O \rightarrow H^+ + HCO_3^-\\ p21\\ Caveolin-1\\ Ca^{2^+} \text{ cytosol}\\ Pl3K \text{ reg class IA (p85-alpha)}\\ Bcl-2 \end{array}$	Protein phosphatase Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion Generic binding protein	Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Cytoplasm Mitochondrion
2 2 3 3 3 3 3 3 3 3	7 8 9 0 1 2 3 4 5 6	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0	$\begin{array}{c} SHP-2\\ GRB2\\ Pl3K reg\ class\ IA\ (p85)\\ Ecadherin\\ Co_2 + H_2O \to H^+ + HCO_3^-\\ p21\\ Caveolin-1\\ Ca^{2+}\ cytosol\\ Pl3K\ reg\ class\ IA\ (p85-alpha)\\ Bcl-2\\ Cl^-\ intracellular \end{array}$	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Generic binding protein Generic binding protein Generic binding protein Inorganic ion Generic may me	Cytoplasm Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Cytoplasm Mitochondrion Intracellular Nucleus
2 2 3 3 3 3 3 3 3 3 3 3	7 8 9 0 1 2 3 4 5 6 7	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720 0.0712		0.548888 0.364614 0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0 -0.208082	$\begin{array}{c} SHP-2\\ GRB2\\ Pl3K reg\ class\ IA\ (p85)\\ E\text{-acherin}\\ Codevenin\\ Caveolin-1\\ Ca^{a^{+}}\ cytosol\\ Ca^{+}\ cytosol\\ Pl3K\ reg\ class\ IA\ (p85\text{-alpha})\\ Bcl-2\\ Cl^{-}\ intracellular\\ MDM2\\ PTEN \end{array}$	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein Inorganic binding protein Generic binding protein Generic binding protein Inorganic ion Generic binding protein Inorganic ion Generic binding protein Inorganic ion	Cytoplasm Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytosol Cytoplasm Mitochondrion Intracellular
2 2 3 3 3 3 3 3 3 3 3 3 3 3 3	7 8 9 0 1 2 3 4 5 6 7 8	0.0776 0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720 0.0712 0.0707		0.548888 0.364614 0.479956 0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0 0.208082 -0.169004	$\begin{array}{c} SHP-2\\ GRB2\\ Pl3K reg\ class\ lA\ (p85)\\ E\ cadherin\\ CO_2\ H_2O \to H^*\ +\ HCO_3^-\\ p21\\ Caveolin-1\\ Caveolin-1\\ Ca^{2+}\ cytosol\\ Pl3K\ reg\ class\ lA\ (p85\ alpha)\\ Bcl-2\\ Cl^-\ intracellular\\ MDM2 \end{array}$	Protein phosphatase Generic binding protein Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Generic binding protein Generic binding protein Generic binding protein Inorganic ion Generic may me	Cytoplasm Cytoplasm Votplasm Plasma membrane NA Nucleus Cytosol Cytosol Cytosol Cytoplasm Mitochondrion Intracellular Nucleus Cytoplasm

We select **12 proteins of interest** (in yellow) among the top 40 PageRank

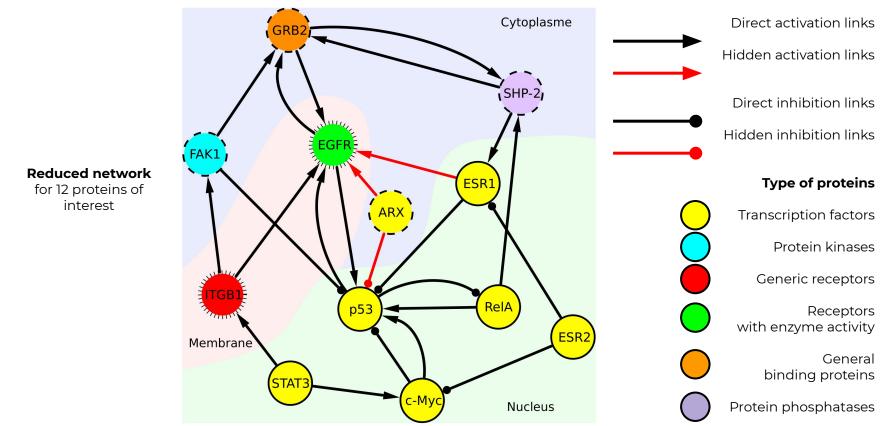


Top 40 PageRank "Process outcomes"

K	P(K) (10 ⁻²)	k	M(K)	Name	Class	Localization
1	0.2506		0	H ⁺ cytosol	Inorganic ion	Cytosol
2	0.2376		ŏ	Na ⁺ cytosol	Inorganic ion	Cytosol
3	0.1741		-0.045970	Beta-catenin	Generic binding protein	Cytoplasm
4	0.1701	1	-0.028308	p53	Transcription factor	Nucleus
5	0.1469		0.256018	c-Src	Protein kinase	Cytoplasm
6	0.1435		0.708154	mBNA intracellular	RNA	Intracellular
7	0.1352		0.700101	H ⁺ extracellular region	Inorganic ion	Extracellular region
8	0.1189	2	0.105603	EGFR	Receptor with enzyme activity	Plasma membrane
9	0.1180	-	-0.014278	DNA	DNA	Nucleus
10	0.1125	3	-0.004135	ESR1 (nuclear)	Transcription factor	Nucleus
11	0.1125	9	0.004100	K ⁺ extracellular region	Inorganic ion	Extracellular region
12	0.1056		ő	ADP cytoplasm	Compound	Cytoplasm
13	0.1023	4	0.250910	STAT3	Transcription factor	Nucleus
14	0.0997	4	0.062046	Androgen receptor	Transcription factor	Nucleus
15	0.0947		0.287801	Rac1	RAS superfamily	Cytoplasm
16	0.0946		0.207001	PO ₄ ⁻ cytoplasm	Compound	Cytoplasm
17	0.0940	5	0.006332		Transcription factor	Nucleus
18	0.0940	6	0.360271	c-Myc FAK1	Protein kinase	
19	0.0919	0	0.962815		Reaction	Cytoplasm NA
				cytosol K ⁺ → extracellular re- gion K ⁺		
20	0.0889	7	0.003377	ESR2 (nuclear)	Transcription factor	Nucleus
21	0.0884		0	K ⁺ cytosol	Inorganic ion	Cytosol
22	0.0849	8	0.002825	ReIA (p65 NF-kB subunit)	Transcription factor	Nucleus
23	0.0834	9	0.004567	ARX	Transcription factor	Cytoplasm
24	0.0828	10	0.208984	ITGB1	Generic receptor	Plasma membrane
25	0.0787	11	0.548888	SHP-2	Protein phosphatase	Cytoplasm
26	0.0776					
		12	0.364614	GRB2	Generic binding protein	Cytoplasm
27	0.0760	12	0.479956	PI3K reg class IA (p85)	Generic binding protein	Cytoplasm
28	0.0760 0.0759	12	0.479956 -0.114311	PI3K reg class IA (p85) E-cadherin	Generic binding protein Generic binding protein	Cytoplasm Plasma membrane
28 29	0.0760 0.0759 0.0754	12	0.479956 -0.114311 0.757892	PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$	Generic binding protein Generic binding protein Reaction	Cytoplasm Plasma membrane NA
28 29 30	0.0760 0.0759 0.0754 0.0753	12	0.479956 -0.114311 0.757892 -0.098664	PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21	Generic binding protein Generic binding protein Reaction Generic binding protein	Cytoplasm Plasma membrane NA Nucleus
28 29 30 31	0.0760 0.0759 0.0754 0.0753 0.0752	12	0.479956 -0.114311 0.757892 -0.098664 0.148707	PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21 Caveolin-1	Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein	Cytoplasm Plasma membrane NA Nucleus Cytoplasm
28 29 30 31 32	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470	PI3K reg class IA (p85) E-cadherin $CO_2 + H_2O \rightarrow H^+ + HCO_3^-$ p21 Caveolin-1 Ca ²⁺ cytosol	Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytosol
28 29 30 31 32 33	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345	PI3K reg class IA (p85) E-cadherin CO ₂ + H ₂ O → H ⁺ + HCO ₃ ⁻ p21 Caveolin-1 Ca ²⁺ cytosol PI3K reg class IA (p85-alpha)	Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion Generic binding protein	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytosol Cytoplasm
28 29 30 31 32 33 34	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470	$\begin{array}{l} \text{PI3K reg class IA (p85)} \\ \text{E-cadherin} \\ \text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}^+ + \text{HCO}_3^- \\ \text{p21} \\ \text{Caveolin-1} \\ \text{Caveolin-1} \\ \text{Ca^{2+} cytosol} \\ \text{PI3K reg class IA (p85-alpha)} \\ \text{Bcl-2} \end{array}$	Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion Generic binding protein Generic binding protein	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Mitochondrion
28 29 30 31 32 33 34 35	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0	$\begin{array}{l} \text{PI3K reg class IA (p85)} \\ \text{E-cadherin} \\ \text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}^+ + \text{HCO}_3^- \\ \text{p21} \\ \text{Caveolin-1} \\ \text{Caveolin-1} \\ \text{Ca}^{2^-} \text{cytosol} \\ \text{PI3K reg class IA (p85-alpha)} \\ \text{Bcl-2} \\ \text{CI^- intracellular} \end{array}$	Generic binding protein Generic binding protein Reaction Generic binding protein Inorganic ion Generic binding protein Generic binding protein Inorganic ion	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Mitochondrion Intracellular
28 29 30 31 32 33 34 35 36	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751	$\begin{array}{l} \text{PI3K reg class A (p85)} \\ \text{E-cadherin} \\ \text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}^+ + \text{HCO}_3^- \\ \text{p21} \\ \text{Caveolin-1} \\ \text{Ca}^{2,2} \text{ cytosol} \\ \text{PI3K reg class A (p85-alpha) } \\ \text{Bcl-2} \\ \text{CI^- intracellular} \\ \text{MDM2} \end{array}$	Generic binding protein Generic binding protein Reaction Generic binding protein Generic binding protein Inorganic ion Generic binding protein Generic binding protein	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Mitochondrion Intracellular Nucleus
28 29 30 31 32 33 34 35 36 37	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720 0.0712 0.0707	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0 -0.208082 -0.169004	$\begin{array}{l} Pl3K reg \ class \ IA \ (p85) \\ E^{cadherin} \\ CO_2 + H_2 O \to H^+ + HCO_3^- \\ p21 \\ Caveolin-1 \\ Ca^{c2} \ cytosol \\ Ca^{c2} \ cytosol \\ Pl3K \ reg \ class \ IA \ (p85\text{-slpha}) \\ Bcl-2 \\ Bcl-2 \\ Cl^- \ intracellular \\ MDM2 \\ PTEN \end{array}$	Generic binding protein Generic binding protein Reaction Generic binding protein Inorganic ion Generic binding protein Generic binding protein Inorganic ion Generic binding protein Inorganic ion Generic enzyme Lipid phosphatase	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Mitochondrion Intracellular Nucleus Cytoplasm
28 29 30 31 32 33 34 35 36 37 38	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720 0.0712 0.0707 0.0702	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0 -0.208082 -0.169004 0.391984	PI3K reg class IA (p85) E-cadherin CO ₂ + $h_2O \rightarrow H^* + HCO_3^-$ p21 Caveolin-1 Ca ²⁺ e cytosol PI3K reg class IA (p85-alpha) Bcl-2 Cl ⁻ intracellular MDM2 PTEN PPAR-gamma	Generic binding protein Generic binding protein Reaction Generic binding protein Inorganic ion Generic binding protein Generic binding protein Generic binding protein Inorganic ion Generic enzyme Lipid phosphatase Transcription factor	Cytoplasm Plasma membrane NA Vucleus Cytoplasm Cytoplasm Mitochondrion Intracellular Nucleus Cytoplasm Nucleus Sucleus
28 29 30 31 32 33 34 35 36 37	0.0760 0.0759 0.0754 0.0753 0.0752 0.0749 0.0744 0.0727 0.0720 0.0712 0.0707	12	0.479956 -0.114311 0.757892 -0.098664 0.148707 0.007470 0.381345 -0.220751 0 -0.208082 -0.169004	$\begin{array}{l} Pl3K reg \ class \ IA \ (p85) \\ E^{cadherin} \\ CO_2 + H_2 O \to H^+ + HCO_3^- \\ p21 \\ Caveolin-1 \\ Ca^{c2} \ cytosol \\ Ca^{c2} \ cytosol \\ Pl3K \ reg \ class \ IA \ (p85\text{-slpha}) \\ Bcl-2 \\ Bcl-2 \\ Cl^- \ intracellular \\ MDM2 \\ PTEN \end{array}$	Generic binding protein Generic binding protein Reaction Generic binding protein Inorganic ion Generic binding protein Generic binding protein Inorganic ion Generic binding protein Inorganic ion Generic enzyme Lipid phosphatase	Cytoplasm Plasma membrane NA Nucleus Cytoplasm Cytoplasm Mitochondrion Intracellular Nucleus Cytoplasm

We select **12 proteins of interest** (in yellow) among the top 40 PageRank

Types of links



Take home messages

- We have presented a detailed description of the statistical properties of the multi-omic interactions MetaCore network obtained with extensive Google matrix analysis.
- Proteins and molecules which are at the top PageRank are at "outcomes" of multiple cascades of causal interactions.
- Proteins and molecules which are at the top CheiRank are "triggers" of multiple cascades of causal interactions.
- The **reduced Google Matrix** is an analytical approach for **inferring hidden indirect connections** within a set of nodes embedded in a very large network.
- In the case of the MetaCore database, hidden signaling pathways can be detected.
- The reduced Google matrix analysis allows to determine the effective interactions between proteins of interest taking into account all the indirect pathways between these proteins through the global MetaCore network.
- The reduced Ising Google matrix algorithm, based on Ising spin description, allows to take into account the bi-functional nature of the protein-protein interactions (activation or inhibition) and to determine the average action type (or magnetization) of each protein.
- Further studies will address biological effects inferred from the reduced Google matrix analysis of the MetaCore network.

Thank you for your attention !!