

Data integration in systems biology: characterization of biological phenomena from structural and functional information

Ronaldo Fumio Hashimoto

Institute of Mathematics and Statistics
University of Sao Paulo
Brazil

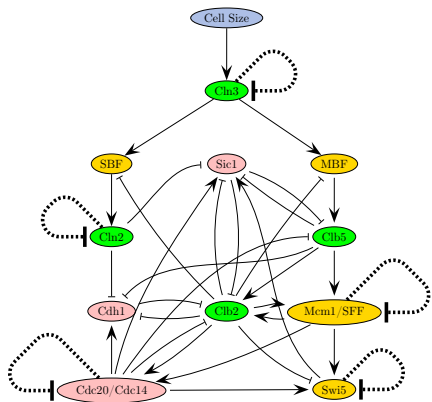
Project MSR-FAPESP 2011-2013

Summary

- Motivation
- Our Approach
- Domain Partners
- Research Team

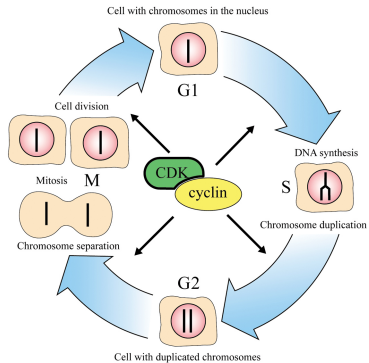
Motivation

Biological Molecules Interactions



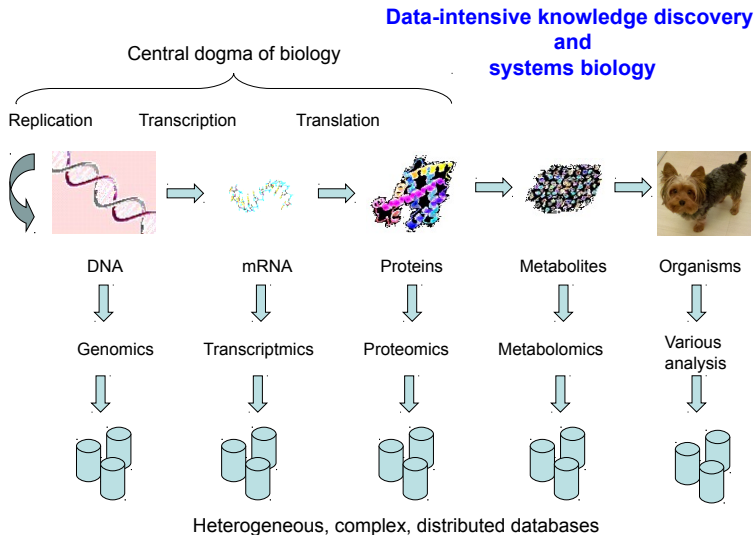
Li et al. PNAS 2004.

The Cell Cycle



http://nobelprize.org/nobel_prizes/medicine/laureates/2001/press.html

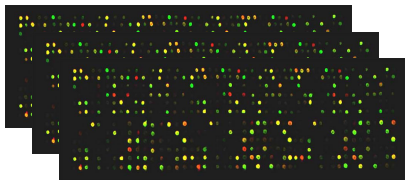
Cell Activity to Data Sets



Our Approach

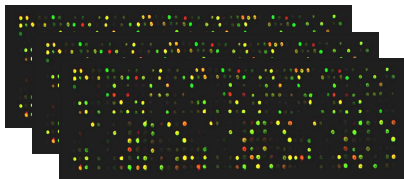
Steps for Gene Regulatory Network Inference

Set of Microarrays



Steps for Gene Regulatory Network Inference

Set of Microarrays



dados-raozes-frio-normal-root-exemplo.xlsx

Search in Sheet

Home Layout Tables Charts SmartArt Formulas Data Review

Font Alignment Number Format Cells Themes

General Conditional Formatting Styles Actions

A1	B	C	D	E	F	G	H	I	J
ProbeSetName	30n-R1	1h-R1	3h-R1	6h-R1	12h-R1	24h-R1	quorra	Oh2	
1 AFFX-Biob-D	0.690140845	0.725118463	0.673071199	0.749494949	0.470552777	0.227472527	-999		
3 AFFX-Biob-M	1.138895859	0.933742077	0.903314119	0.857146633	0.6624305	0.919819756	-999		
4 AFFX-Biob-S	1.238189533	0.898612022	0.701317716	0.894628621	0.708029222	0.849895174	-999		
5 AFFX-Biob-O	1.898989897	0.750231073	0.759294732	0.624627951	0.811894882	0.993952226	-999		
6 AFFX-Biob-C	1.020375	0.843875174	0.871128707	0.881009175	0.785345765	0.815737263	-999		
7 AFFX-Biob-I	1.106660667	0.8354095	0.809485095	0.861220289	0.754432824	0.921549637	-999		
8 AFFX-Biob-D	1.398139891	0.844349174	0.790486973	0.942788474	0.785262126	0.618499556	-999		
9 AFFX-Cmk-A	1.028552467	0.713284629	0.865202884	0.820282309	0.852888721	0.936330733	-999		
10 AFFX-Cmk-X	1.10037119	0.8489577	0.897592867	0.827453775	0.871155283	0.873750886	-999		
11 AFFX-DapK-3	1.181818182	0.152946515	0.386363636	0.685308505	1.125	0.888888889	-999		
12 AFFX-DapK-A	1.002090901	0.8075	1.15	0.847098654	0.235294116	0.871902321	-999		
13 AFFX-DapK-X	1.1	875	0.990309091	1.1	1125	1.5	-999		
14 AFFX-LysK-A	1.333333333	0.568695858	0.5	0.650583156	0.4	0.712727273	-999		
15 AFFX-LysK-X	1.815384615	0.75	0.388421053	0.333333333	0.823020412	2	-999		
16 AFFX-LysK-3	1.333333333	0.181818182	1.125	0.734920870	0.5	0.444444444	-999		
17 AFFX-Pmk-A	0.777777777	0.666666667	1.4	0.833333333	0.8125	0.828971429	-999		
18 AFFX-Pmk-X	1.1	1375	0.5	0.444444444	2	0.857142857	-999		
19 AFFX-Pmk-3	1.1	0.33461538	1.05205158	0.51724310	0.838263685	2.35	-999		
20 AFFX-Trk-A	1.357142857	0.333333333	1.10384154	0.79523769	0.333333333	0.29415094	-999		
21 AFFX-Trk-M	1.450740741	0.363836364	0.533333333	0.757575758	0.116949119	0.333333333	-999		
22 AFFX-Trk-X	0.81011236	0.20	0.747121854	0.610970740	0.712041865	0.333333333	-999		
23 AFFX-TypK-2	1.53271208	0.574833333	0.718622077	0.827678921	0.822222222	0.97217017	-999		
24 AFFX-TypK-4	1.0	0.8	0.75	0.833333333	0.4	0.333333333	-999		
25 AFFX-TypK-1	0.278333333	0.887359867	0.666666667	0.742887148	1.470788284	0.454545454	-999		
26 AFFX-Q-Eob	0.371845701	0.80514752	0.7250874	0.226702509	0.638157158	0.748957856	-999		
27 AFFX-Q-Eob	0.987848948	0.388259101	0.96529963	0.208634269	0.912902620	0.78749898	-999		
28 AFFX-Q-Eob	1.186289494	0.775757576	0.838611985	0.2862464	0.889174948	0.2838415	-999		
29 AFFX-Q-Eob	1.206835331	0.895523291	0.888368462	0.821561051	0.85695321	0.840088419	-999		
30 AFFX-Q-Eob	1.228303748	0.845075125	0.826247089	0.86416185	0.92371115	0.070088272	-999		
31 AFFX-Q-Eob	1.168586492	0.818181818	0.856448611	0.822881193	0.887733263	0.811336653	-999		
32 AFFX-Q-Eob	0.869844217	0.888495116	0.959631637	0.371420111	0.992725898	0.784777130	-999		
33 AFFX-Q-P1-c	1.15222414	0.888546113	0.87304207	0.95037896	0.846327280	0.02620945	-999		
34 AFFX-Q-P1-c	0.88202514	0.88734917	0.81700078	0.89649649	0.896208505	0.864487698	-999		
35 AFFX-Q-Bsd	1.333333333	0.125	0.333333333	0.89041096	0.426229506	1	-999		
36 AFFX-Q-Bsd	1.38683838	0.4	0.8	0.286208	0.11815	0.866666667	-999		
37 AFFX-Q-Bsd	1.127895974	0.857142857	0.47828087	0.733333333	1.22222222	2	-999		
38 AFFX-Q-Bsd	1.486333333	0.333333333	0.683333333	0.833333333	0.833333333	0.833333333	-999		

Steps for Gene Regulatory Network Inference

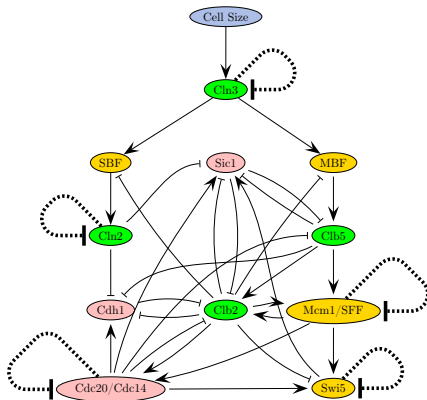
Microsoft Excel spreadsheet titled "dados-raozes-frio-normal-root-exemplo.xlsx". The spreadsheet displays a table with columns labeled A1 through Q2 and rows representing different gene sets. The data consists of numerical values and a final column of values ranging from -999 to 2.

ProbeSetName	30m-R1	1h-R1	2h-R1	6h-R1	12h-R1	24h-R1	quorra	Oh-R2
1 AFFX-BuOb-0	1.690140845	0.725118463	0.673671199	0.749494949	0.470057277	0.027472527	-999	-999
2 AFFX-BuOb-M	1.138885859	0.933742037	0.903141419	0.653174633	0.6624305	0.919619756	-999	-999
4 AFFX-BuOb-3	1.028189533	0.889612022	0.701317716	0.894628621	0.709602922	0.884665174	-999	-999
5 AFFX-BuOb-C	1.898989897	0.752037749	0.739294732	0.624627951	0.811894882	0.993952226	-999	-999
6 AFFX-BuOb-2	1.002375	0.843897114	0.671129707	0.891001975	0.785345765	0.917507263	-999	-999
7 AFFX-BuOb-n	1.106666667	0.9354095	0.809489595	0.891220269	0.754432824	0.921549637	-999	-999
8 AFFX-BuOb-n	1.981398901	0.84341914	0.790486073	0.940788474	0.785262216	0.916499556	-999	-999
9 AFFX-CraX-8	1.026525487	0.713384625	0.895262984	0.920928296	0.853888721	0.993330733	-999	-999
10 AFFX-CraX-3	1.1037119	1.0489577	0.897028267	0.927453775	0.971155263	0.873750386	-999	-999
11 AFFX-DapX-2	1.181818182	0.615284615	0.380363636	0.585308505	1.125	0.888888889	-999	-999
12 AFFX-DapX-A	1.002909091	0.8878	1.5	0.847086824	0.235294116	0.23709231	-999	-999
13 AFFX-DapX-3	1.1	0.7875	0.909309091	1.1	1.125	1.5	-999	-999
14 AFFX-Lyx-A	1.333333333	0.568495358	0.5	0.650528158	0.4	0.272727273	-999	-999
15 AFFX-Lyx-M	1.615384615	1.75	0.368421053	0.333333333	0.823029412	2	-999	-999
16 AFFX-Lyx-X	1.333333333	1.81818182	1.125	0.73463878	0.5	0.444444444	-999	-999
17 AFFX-PheX-5	1.077777778	0.666666667	1.4	0.633333333	0.8125	0.628611429	-999	-999
18 AFFX-PheX-A	1.1	1.375	0.8	0.444444444	0.2	0.857142857	-999	-999
19 AFFX-PheX-3	1.1	0.338461538	1.05293158	0.51724318	0.838263658	0.25	-999	-999
20 AFFX-TrkX-0	1.357142857	0.333333333	1.52848154	0.79523789	0.333333333	0.26415094	-999	-999
21 AFFX-TrkX-M	1.490740741	0.363836364	0.533333333	0.757575758	0.101694919	0.333333333	-999	-999
22 AFFX-TrkX-X	1.01011236	0.29	0.741721864	0.836970749	0.72047865	0.333333333	-999	-999
23 AFFX-TrpX-2	1.53671268	0.574837688	0.078629277	0.26278629	0.822222222	0.37037037	-999	-999
24 AFFX-TrpX-4	1.0	0.8	0.75	0.833333333	0.4	0.333333333	-999	-999
25 AFFX-TrpX-5	1.278333333	0.871350667	0.696666667	1.142857143	1.147058264	0.545454545	-999	-999
26 AFFX-Q-Eob	0.371845701	0.80514752	0.72520574	0.226702509	0.639157158	0.786957856	-999	-999
27 AFFX-Q-Eob	0.987848948	0.393229001	0.96529963	0.208834269	0.912906203	0.787449988	-999	-999
28 AFFX-Q-Eob	1.186289494	0.275175716	0.828691185	0.28662464	0.998917494	0.70338415	-999	-999
29 AFFX-Q-Eob	1.206835331	0.899552339	0.888368462	0.621561051	0.569605321	0.840068419	-999	-999
30 AFFX-Q-Eob	1.228302748	0.845075125	0.826247689	0.68416185	0.92637115	0.67008272	-999	-999
31 AFFX-Q-Eob	1.168584949	0.818178182	0.936404861	0.222846451	0.987733338	0.911335603	-999	-999
32 AFFX-Q-Eob	0.969644217	0.888895816	0.959851637	0.271420111	0.92725986	0.784777139	-999	-999
33 AFFX-Q-P1-c	1.152294418	0.888546153	0.857304267	0.950387966	0.843037289	0.20620945	-999	-999
34 AFFX-Q-P1-c	1.028620514	0.89734907	0.837290276	0.89649896	0.896289505	0.884487868	-999	-999
35 AFFX-Q-Bsd	1.333333333	1.25	0.333333333	0.89041096	0.426229508	1	-999	-999
36 AFFX-Q-Bsd	1.863838384	1.1	0.8	0.25862096	0.1875	0.866666687	-999	-999
37 AFFX-Q-Bsd	1.127859574	0.857142857	0.478262697	0.733333333	1.222222222	2	-999	-999
38 AFFX-Q-Bsd	1.1	0.8	0.8	0.8	0.8	0.8	-999	-999

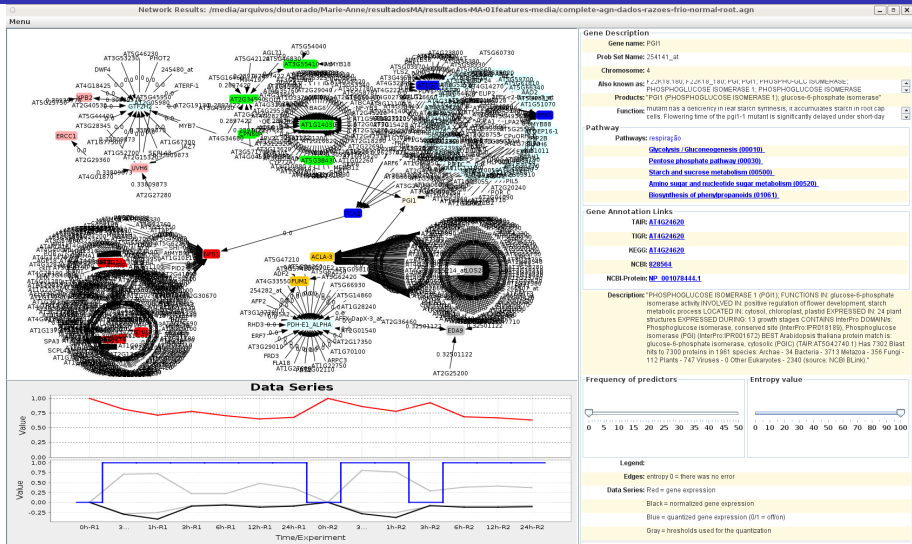
Inference Procedure

Steps for Gene Regulatory Network Inference

Inference Procedure




Tool for Knowledge Discovery



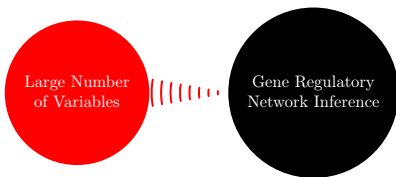
Lopes et al. Feature Selection Environment for Genomic Applications, BMC Informatics, 2008.

Challenges of Using Gene Expression Data

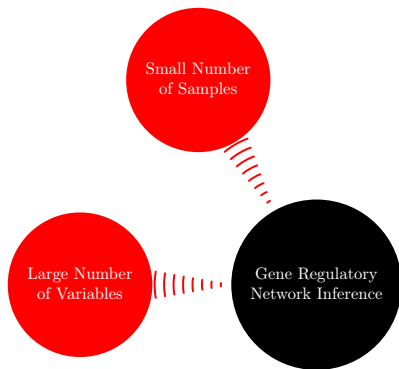


Gene Regulatory
Network Inference

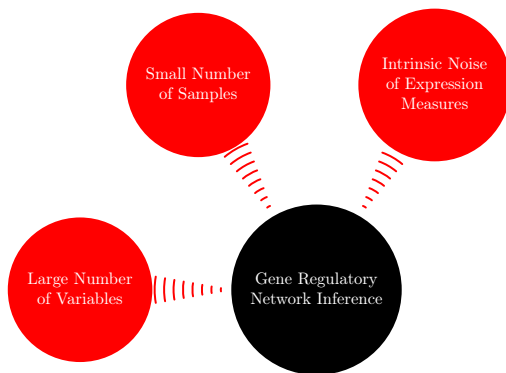
Challenges of Using Gene Expression Data



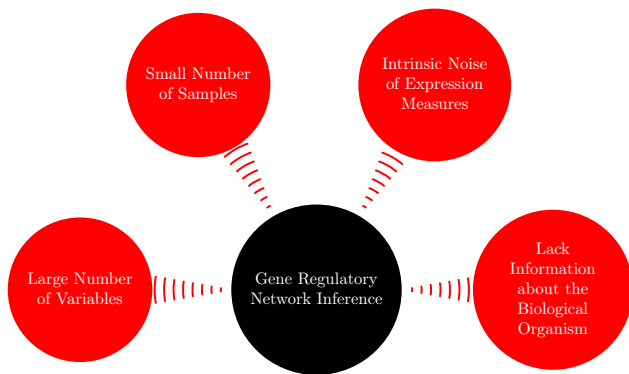
Challenges of Using Gene Expression Data



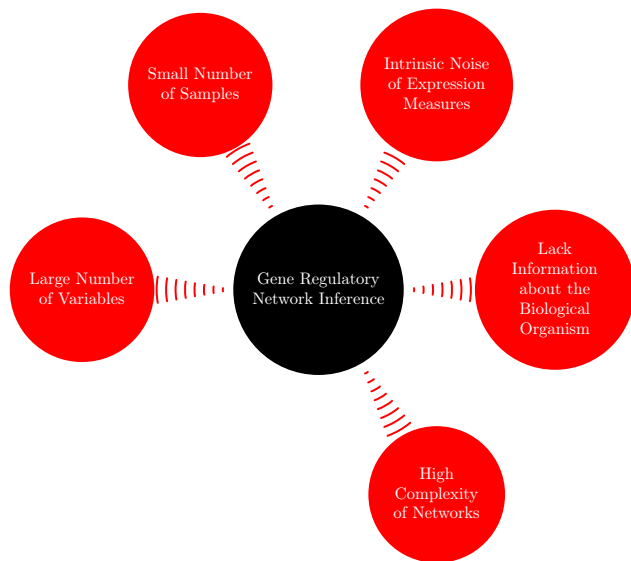
Challenges of Using Gene Expression Data



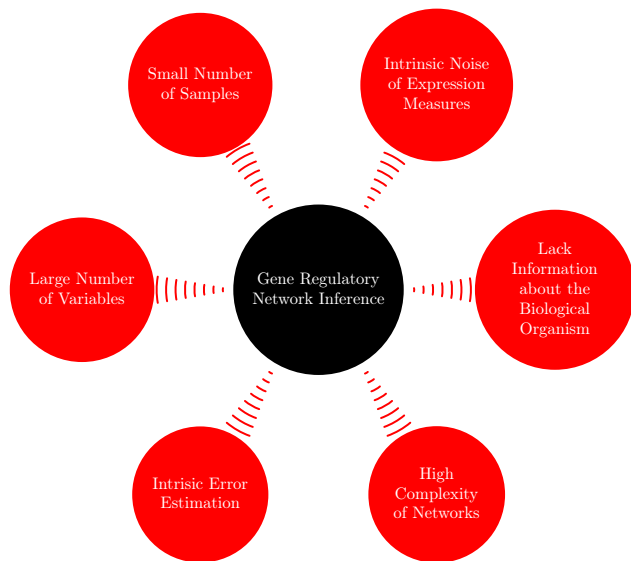
Challenges of Using Gene Expression Data



Challenges of Using Gene Expression Data



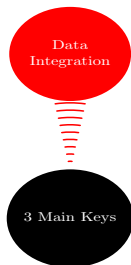
Challenges of Using Gene Expression Data



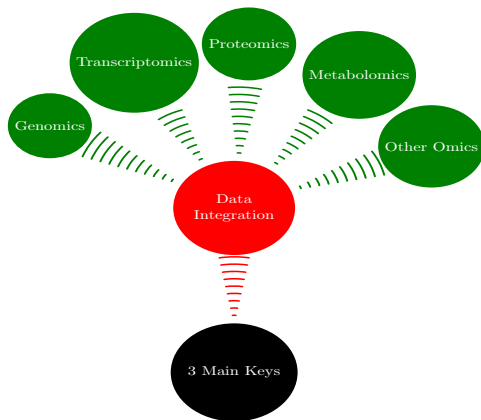
Our Proposal

3 Main Keys

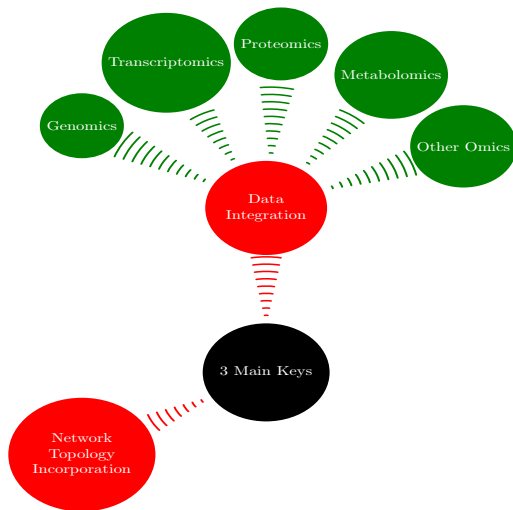
Our Proposal



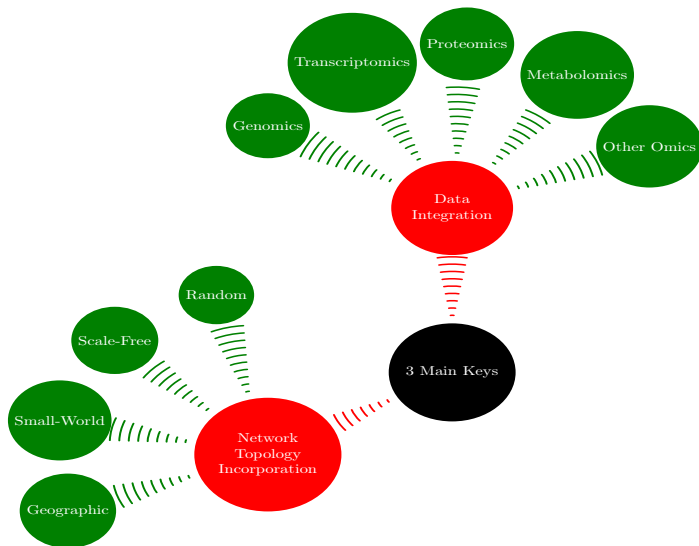
Our Proposal



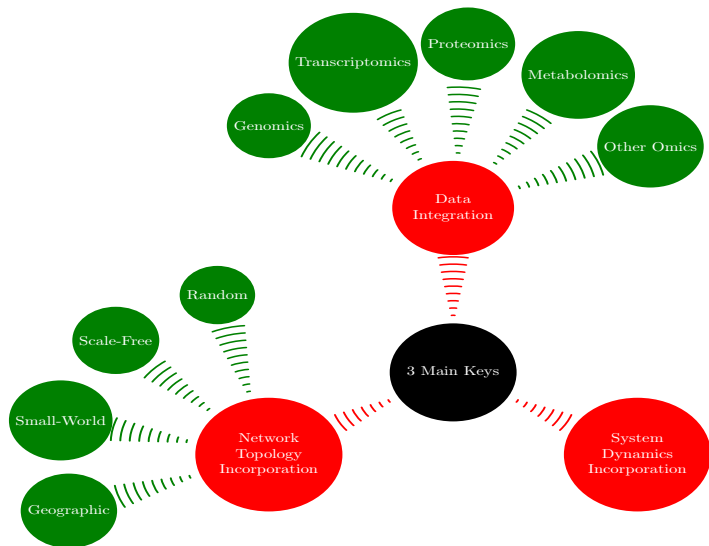
Our Proposal



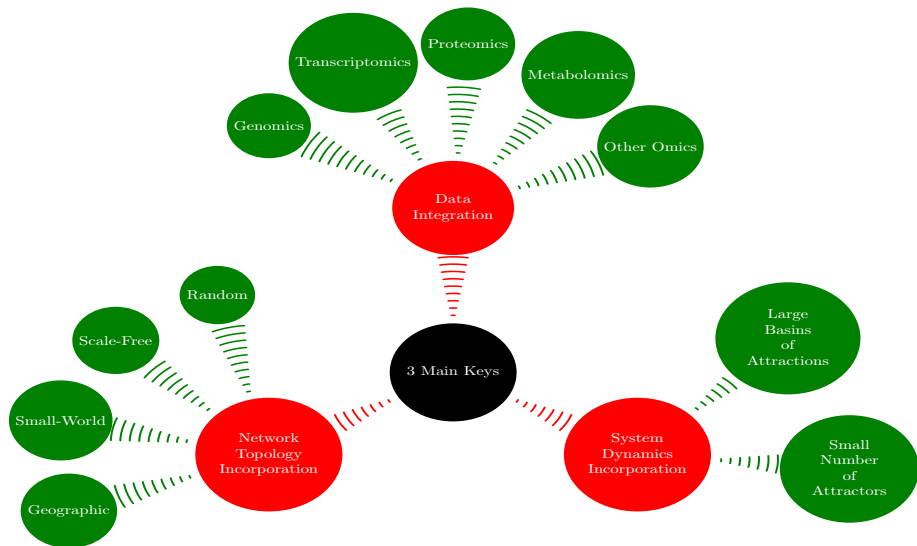
Our Proposal



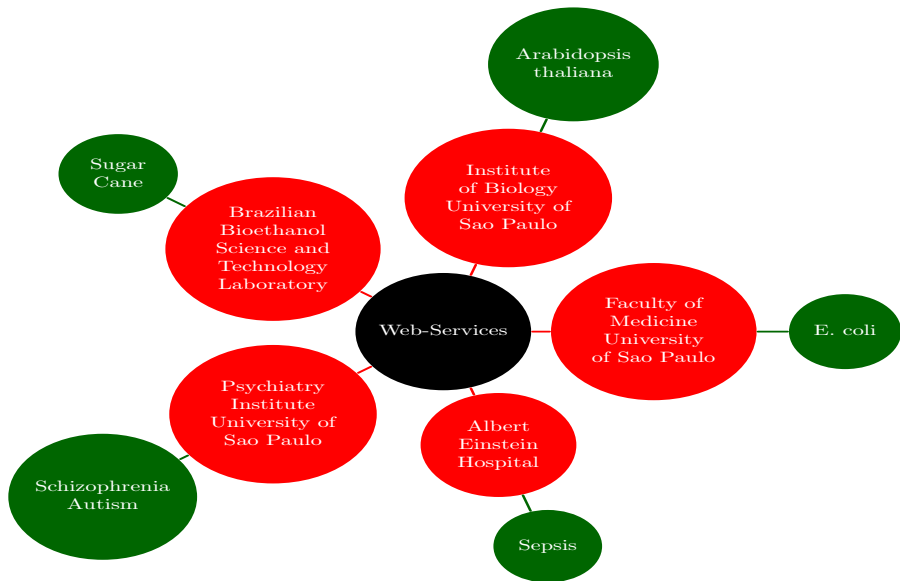
Our Proposal



Our Proposal

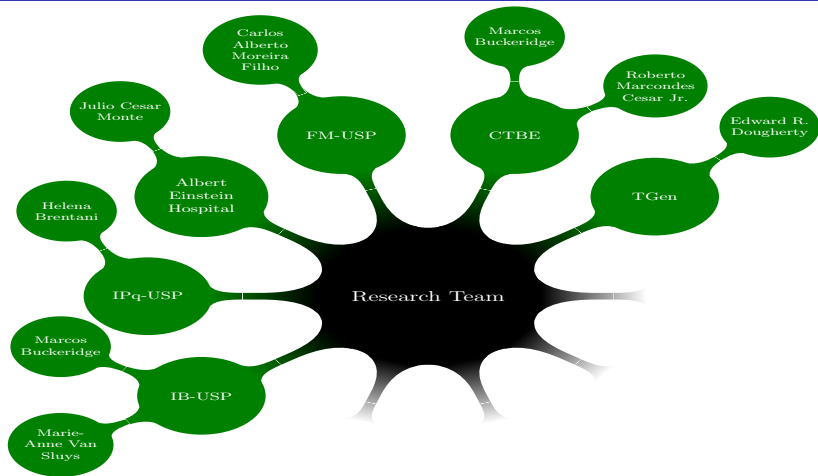


Domain Partners

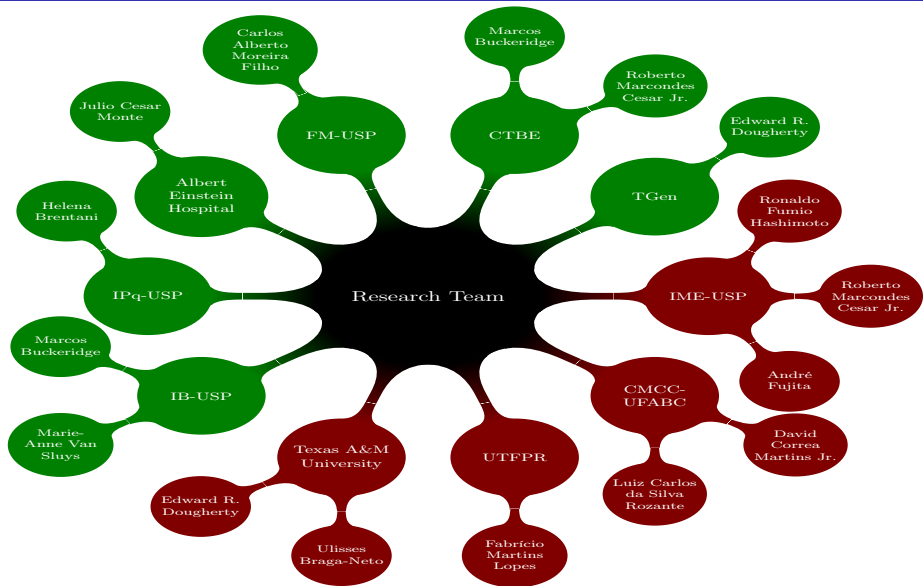




Research Team



Research Team



- F. M. Lopes, R. M. Cesar Jr., L. F. da Costa. Gene Expression Complex Networks: Synthesis, Identification and Analysis. *Journal of Computational Biology*, 2011.
- F. M. Lopes, E. A. de Oliveira, R. M. Cesar Jr. Inference of Gene Regulatory Networks from Time-Series by Tsallis Entropy. *BMC Systems Biology*, 2011.
- C. H. A. Higa, V. H. P. Louzada, T. P. de Andrade, R. F. Hashimoto. Constraint-Based Analysis of Gene Interactions using Restricted Boolean Networks and Time-Series Data. *BMC Proceedings*, 2011.



Microsoft®
Research