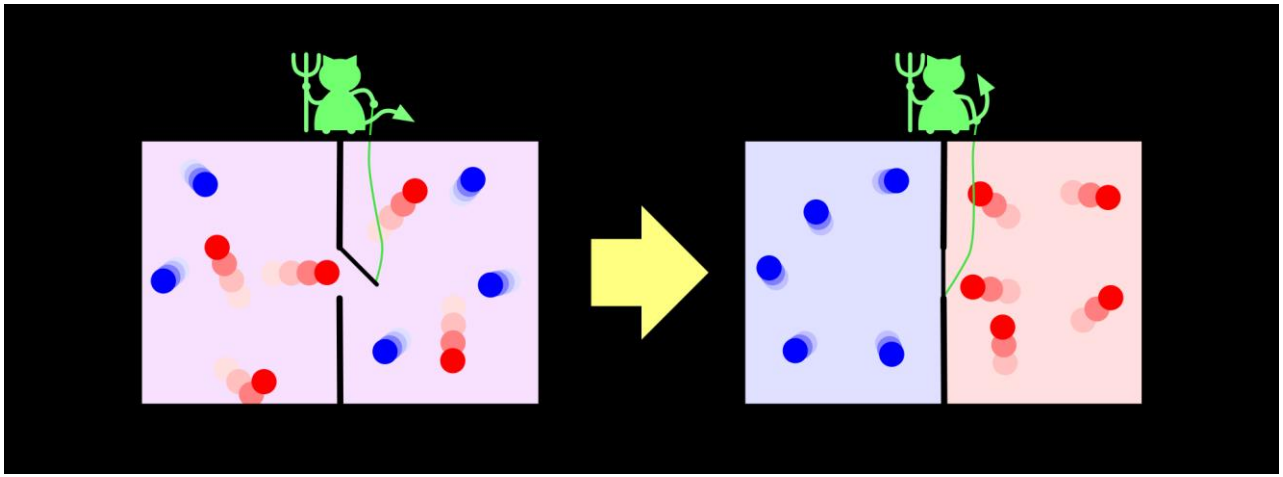
Principles of genome positional information, large scale chromatin organization, nuclear network and their consequences in gene expression regulation

Alessandro Giuliani

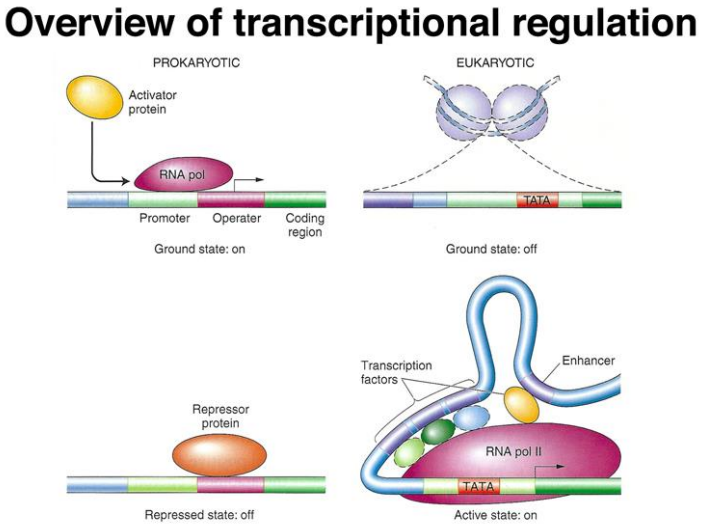
Istituto Superiore di Sanità, Roma, Italia

Maxwell's demon is a thought experiment created by James Clerk Maxwell in 1867 in which he suggested how the second law of thermodynamics might hypothetically be violated. In the thought experiment, a demon controls a small door between two chambers of gas. As individual gas molecules approach the door, the demon quickly opens and shuts the door so that only fast molecules are passed into one of the chambers, while only slow molecules are passed into the other. Because faster molecules are hotter, the demon's behaviour causes one chamber to warm up and the other to cool down, thereby decreasing entropy and violating the second law of thermodynamics.



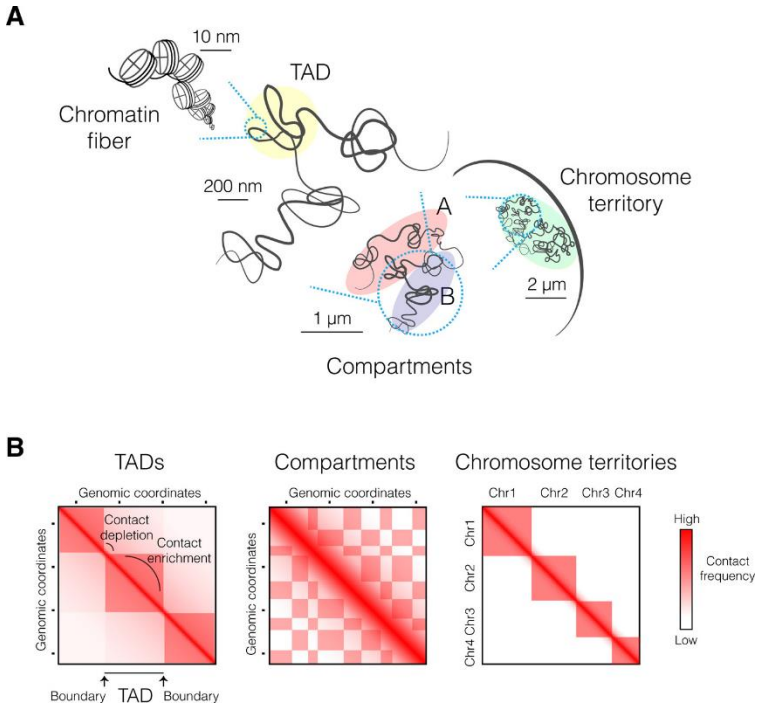
This mechanism implies the presence of an ‘intelligent agent’ (demon) that alters the natural fate of the system…the issue is much more serious than a scientific joke: almost totalityof biological explanations follow a ‘Maxwell’s demon’ style.

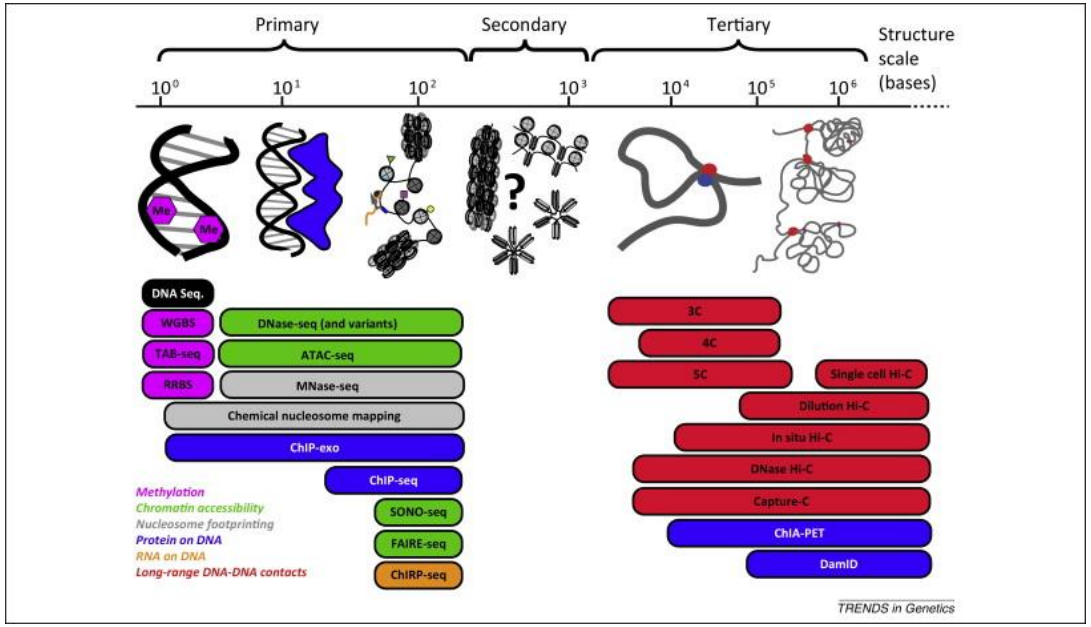
Regulatory elements (e.g transcription factors) **recognize specific** regions of DNA and enhance (or repress) transcription.
Recognition is an intelligent Maxwell's demon-like activity.



These facts happen (e.g.enzymatic activities are based on recognition) but do not constitute a satisfactory general explanation (as any Maxwell demon this kind of regulation asks for an 'extra added' energy that restores second principle) for biological regulation. The huge energy expenditure needed to take care simultaneously of tens of thousands genes and a simple ..matter of scale.. allow to understand why we need a different approach…..

A matter of scale *(why finding a specific gene in the nucleus is more difficult than finding a needle in an haystack)*

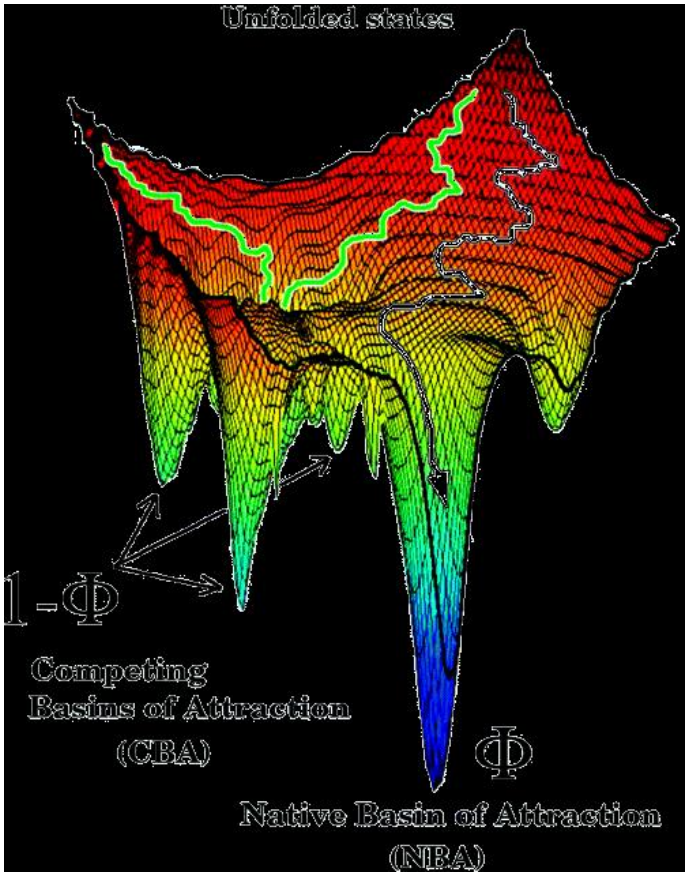




The possible ‘expression patterns’ are limited by the ‘allowed folding states’
of chromatin (with only fine tuning open to specific regulation).

The very limited (in the order of hundreds out of the transfinite number of configurations theoretically possible with tens of thousands gene each one potentially varying across four order of magnitude of expression level) number of tissues in metazoans,
each one corresponding to a very stable expression profile, mirrors this condition.

Genome level expression is driven by a ‘rugged landscape’:



Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest

Matthew K. Vickaryous and Brian K. Hall*

Biol. Rev. (2006), 81, pp. 425–455. f 2006 Cambridge Philosophical Society 425 doi:10.1017/S1464793106007068

The most complete (and detailed) list of cell types encompasses 411 different human cell kinds. Each cell kind has a very invariant gene expression profile. This 'invariant profile' (minimum of a rugged landscape) encompasses general structuring of higher order chromatin structure, 'local co-regulation' of neighboring sites relates to 'small changes' within a given 'cell-kind specific' frame (motions on the bottom of the hole).

Differences of Supranucleosomal Organization in Different Kinds of Chromatin: Cell Type-specific Globular Subunits Containing Different Numbers of Nucleosomes

HANSWALTER ZENTGRAF and WERNER W. FRANKE
Institute of Virus Research, and Department of Membrane Biology and Biochemistry, Institute of Cell and Tumor Biology, German Cancer Research Center, D-6900 Heidelberg, Federal Republic of Germany (FRG)

THE JOURNAL OF CELL BIOLOGY · VOLUME 99 JULY 1984 272–286

NUCLEUS
2018, VOL. 9, NO. 1, 171–181
https://doi.org/10.1080/19491034.2018.1431081

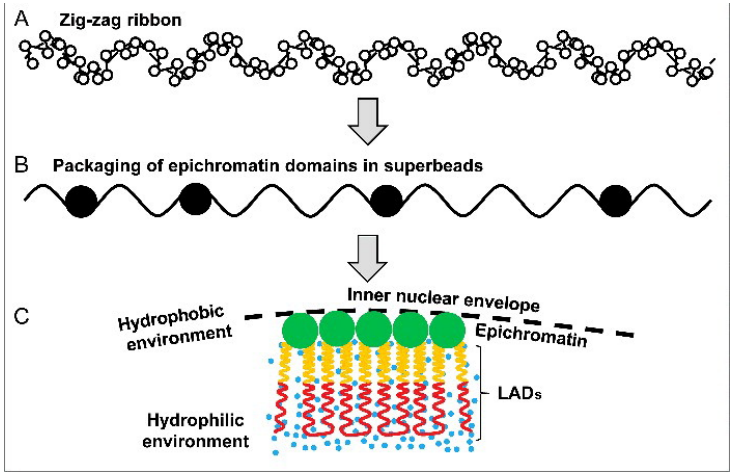
Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

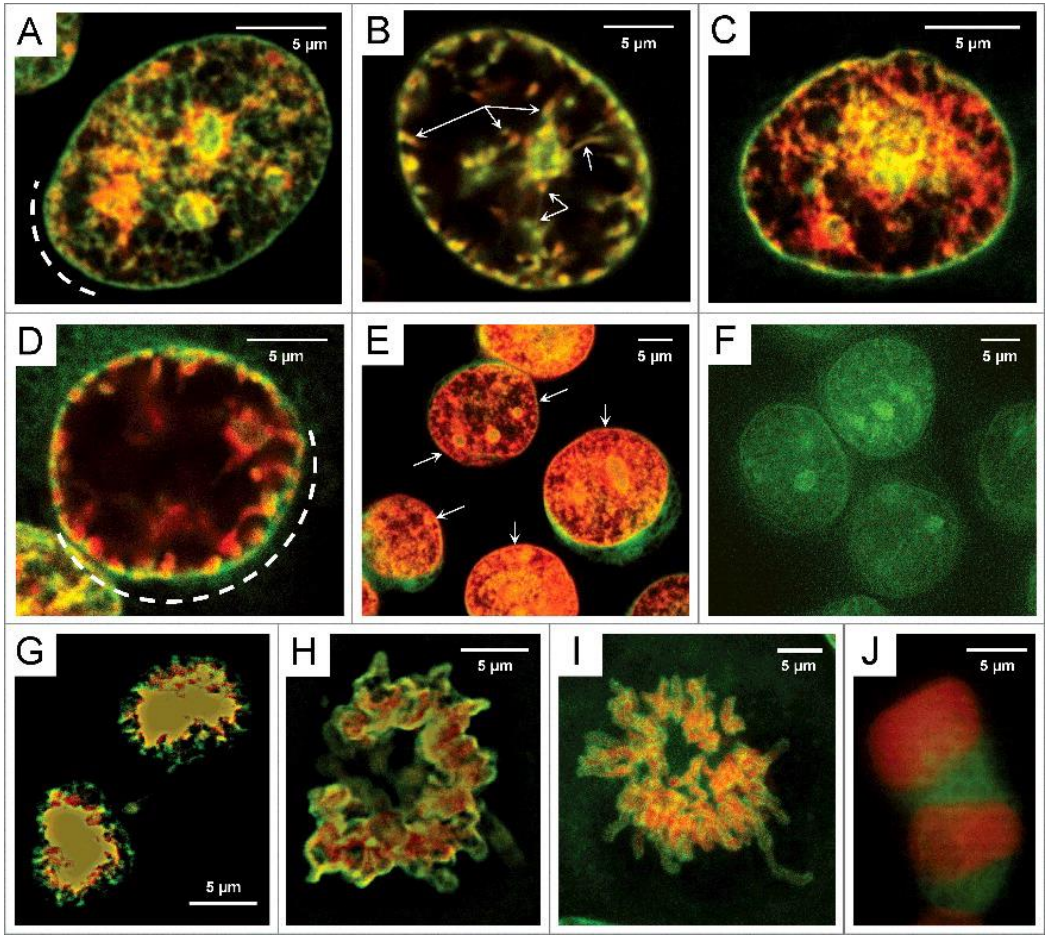
OPEN ACCESS Check for updates

Differential staining of peripheral nuclear chromatin with Acridine orange implies an A-form epichromatin conformation of the DNA

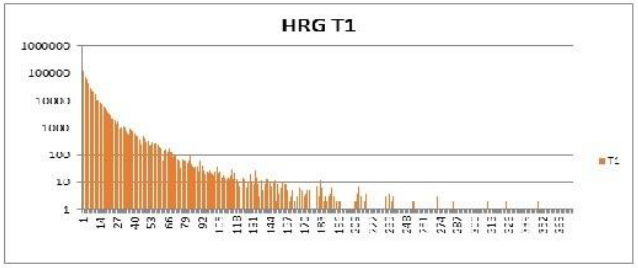
Jekaterina Erenpreisa [a,+], Jekabs Krigerts[a,b], Kristine Salmina[a], Turs Selga[c], Hermanis Sorokins[b] and Talivaldis Freivalds[d,+]

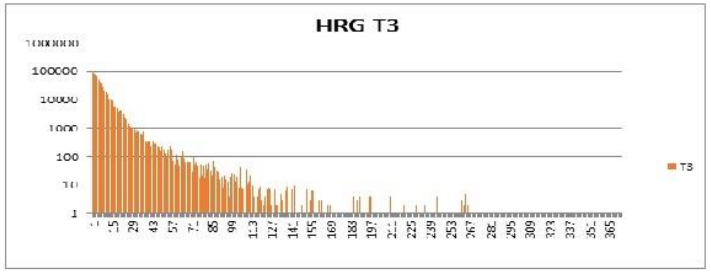


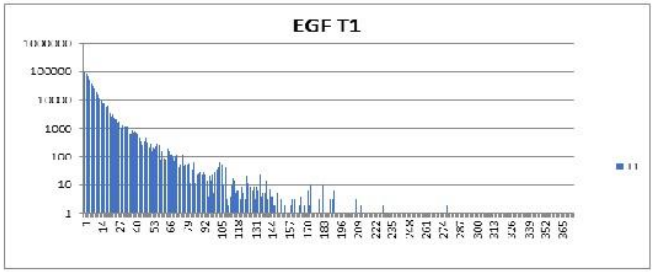
Peripheral nuclear chromatin transduces regulatory messages coming from cell microenvironment

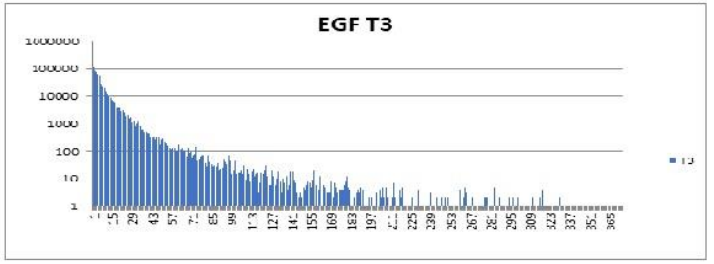


The ‘amount of co-regulation’ (fine tuning) decreases with increasing distance along the chromosome (independently of the entity and nature of the stimulus).









(within the same cell-kind)

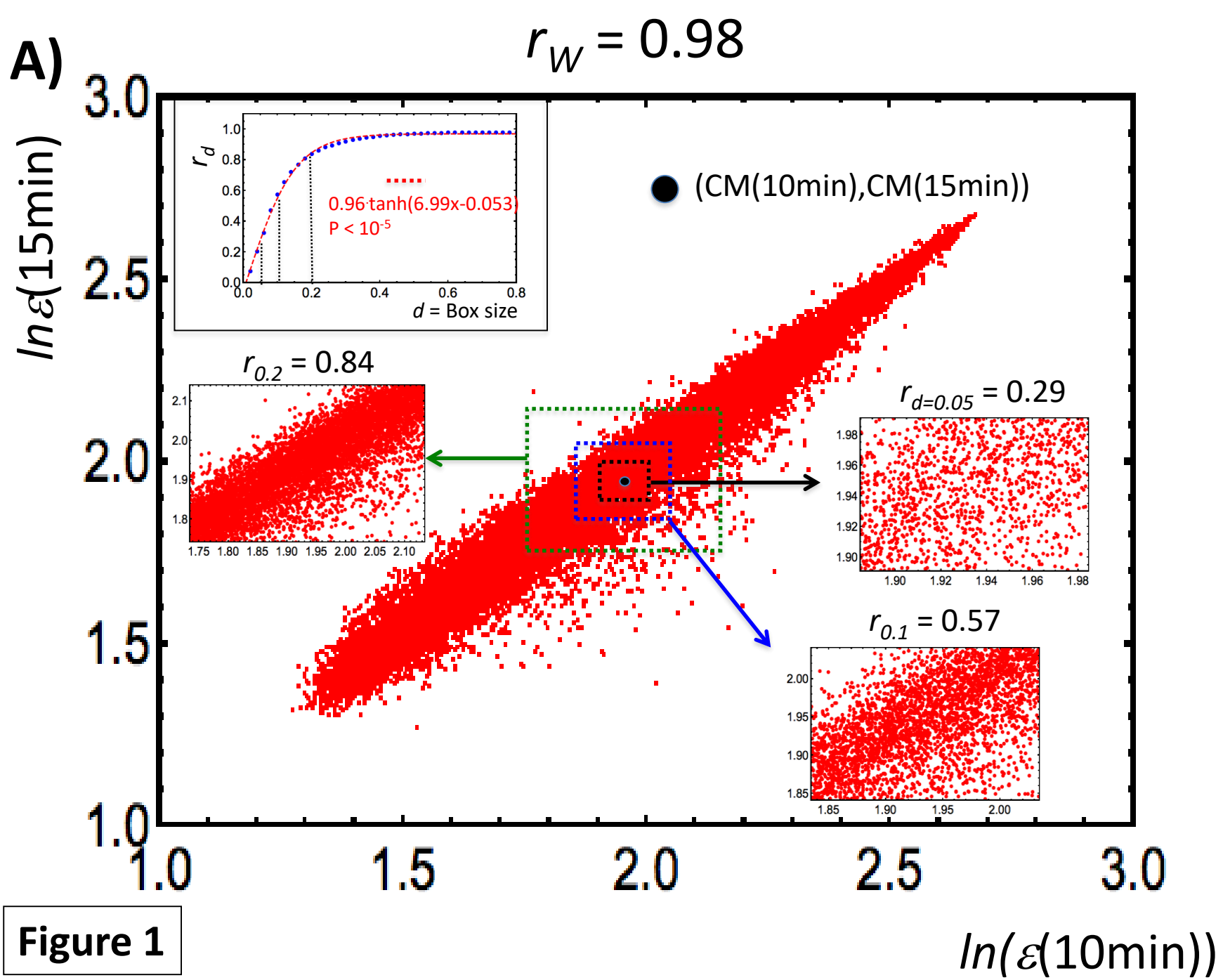
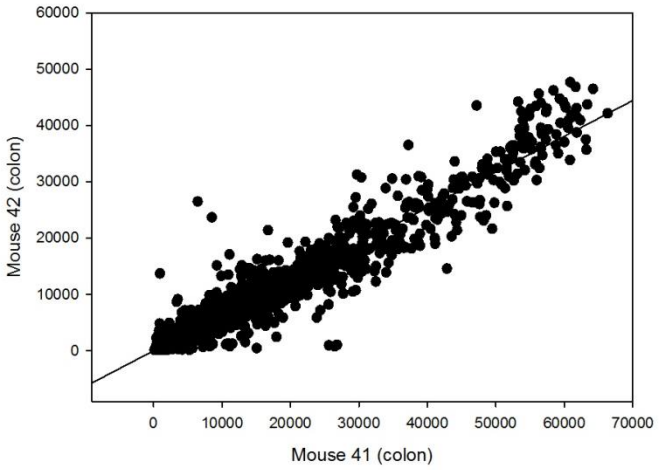
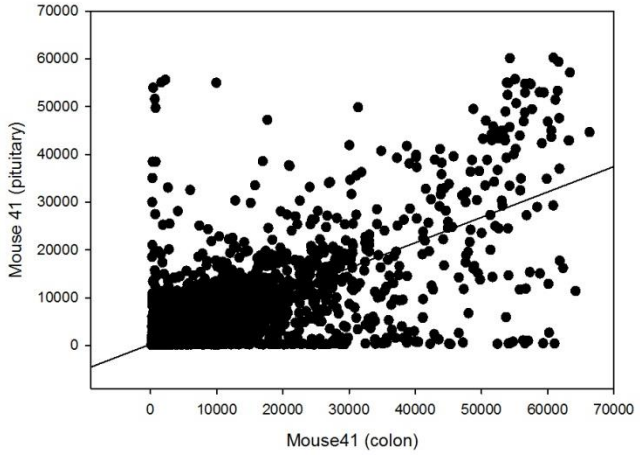

A)
r_W = 0.98
ln(ε(15min))
ln(ε(10min))
(CM(10min),CM(15min))
r_d
0.96·tanh(6.99x-0.053)
P < 10^-5
d = Box size
$r_{0.2}$ = 0.84
$r_{d=0.05}$ = 0.29
$r_{0.1}$ = 0.57

Figure 1



The ‘tissue attractor’ is much stronger than the organism individuality.



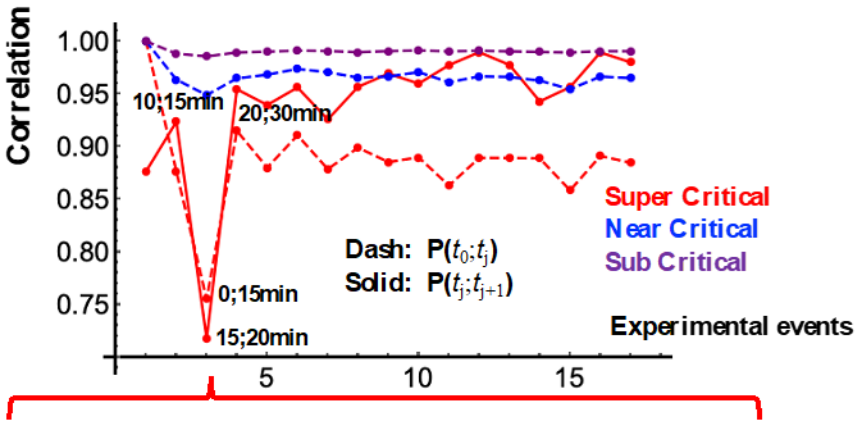
PLOS ONE

RESEARCH ARTICLE

Self-Organizing Global Gene Expression Regulated through Criticality: Mechanism of the Cell-Fate Change

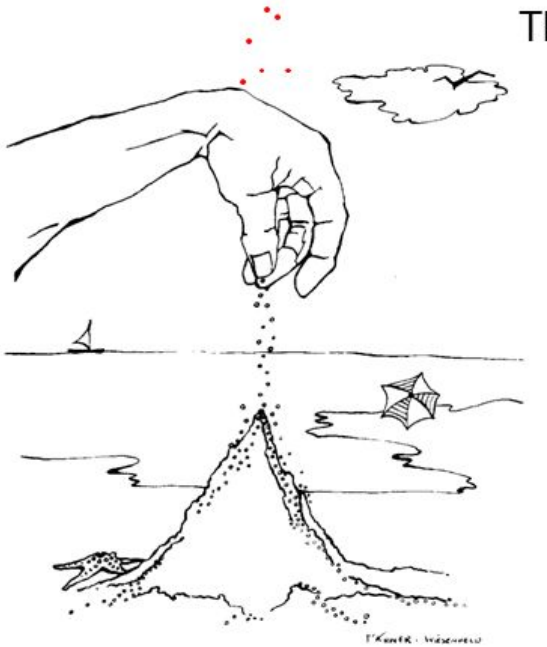
Masa Tsuchiya[1]*, Alessandro Giuliani[2], Midori Hashimoto[3], Jekaterina Erenpreisa[4], Kenichi Yoshikawa[5]

1 Systems Biology Program, School of Media and Governance, Keio University, Fujisawa, Japan, **2** Environment and Health Department, Istituto Superiore di Sanità, Rome, Italy, **3** Graduate School of Frontier Science, the University of Tokyo, Kashiwa, Japan, **4** Latvian Biomedical Research & Study Centre, Riga, Latvia, **5** Faculty of Life and Medical Sciences, Doshisha University, Kyotanabe, Japan



Self-Organizing-Criticality (SOC)

Avalanche Behavior



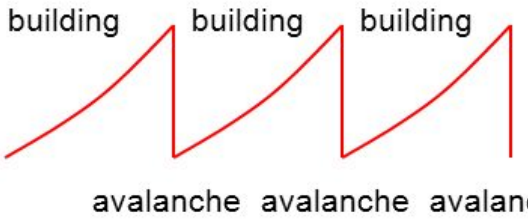
The sand pile builds . . . grain . . . by grain . . .

by grain . . . by grain . . . by grain . . .

by grain . . . by grain . . . by grain . . .

Building toward the critical state . . .

Where it avalanches



Avalanche- a large mass of snow, ice, etc., detached from a mountain slope and sliding or falling suddenly downward.

Avalanche- anything like an avalanche in suddenness and overwhelming quantity: an avalanche of misfortunes; an avalanche of fan mail.

CHEMICAL REVIEWS
Review
pubs.acs.org/CR

Emerging Computational Methods for the Rational Discovery of Allosteric Drugs

Jeffrey R. Wagner,[†] Christopher T. Lee,[†] Jacob D. Durrant,[†,‡] Robert D. Malmstrom,[†,‡] Victoria A. Feher,[†] and Rommie E. Amaro[*,†,‡]

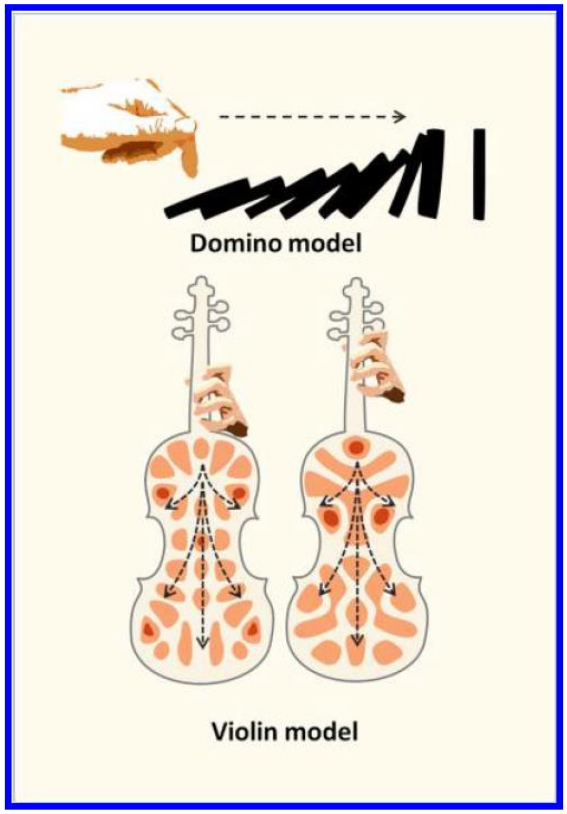


Figure 1. When a small molecule binds to the allosteric site of a protein, information is transferred through the protein molecule to its active site. Two different methods of transmission can be defined. The first mechanism, here defined as the "domino model", is a sequential set of events propagating linearly from the allosteric site to the active site. Binding of the effector triggers local structural changes that sequentially propagate via a single pathway to the active site. It was suggested that this mechanism is applicable for the PDZ domain family,[49] G protein-coupled receptors, the chymotrypsin class of serine proteases, and hemoglobin.[50] The second mechanism, defined here conceptually as a "violin model", is based on vibration pattern changes inside the protein. In a violin its pitch can be changed by a slight movement of the violin player's finger on the fingerboard. Information about the finger movement is, thus, transferred throughout the whole body of the violin with no specific pathway for the signal transduction. By analogy, protein allosteric site is a fingerboard of the protein and a small signaling molecule is the player's finger. If a protein is in a particular vibration mode, it is possible to suggest that binding a small effector molecule to a specific site can change this mode. The signal, thus, will be spread throughout the whole protein including its active site. The "domino model" is a reliable way to transfer information in a macro world, but on a molecular level, with significant thermal motions of the protein, this mechanism will be prone to random triggering of the domino chain reaction, creating noise in the signaling system. Thermal motions in the case of the "violin model" do not hinder the transduction. In fact, the permanent motion of the molecule is a prerequisite for this mechanism.

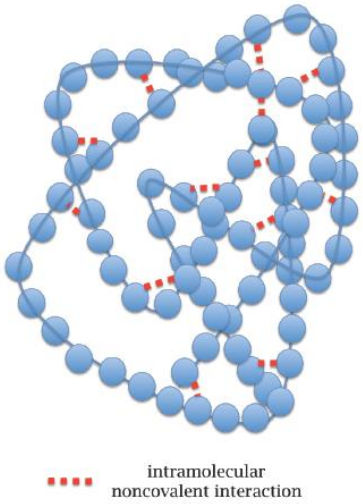
Update TRENDS in Biochemical Sciences Vol.29 No.7 July 2004

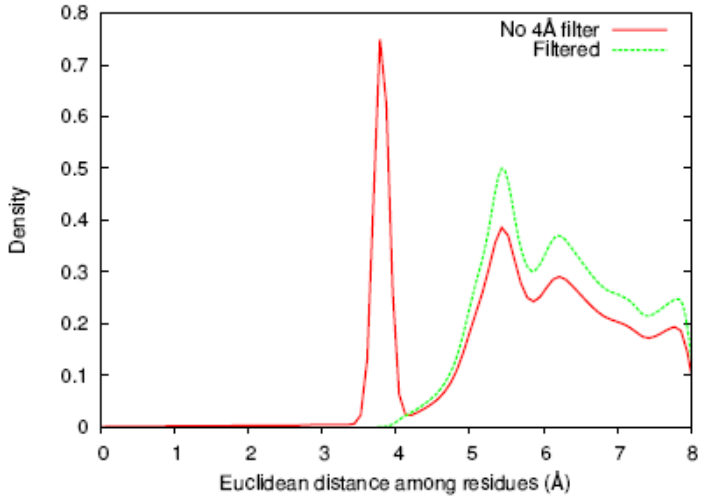
Research Focus

Strong links are important, but weak links stabilize them

Peter Csermely

Department of Medical Chemistry, Semmelweis University, H-1088 Budapest, Puskin str. 9, Hungary





Node degree = number of edges connected to a node

Average Shortest Path (ASP) = Average length of the shortest path connecting any two nodes of the graph.

Betweeness Centrality : Number of shortest paths passing by a node.

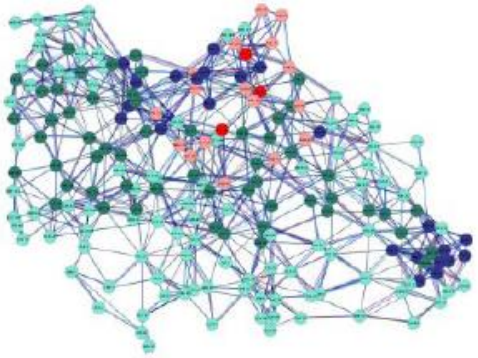
CHEMICAL REVIEWS

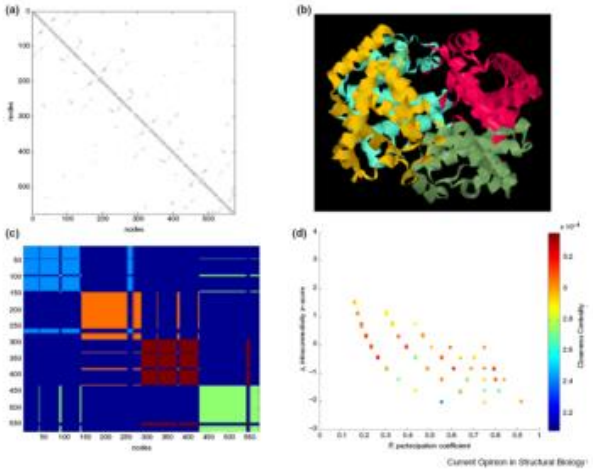
Review

pubs.acs.org/CR

Protein Contact Networks: An Emerging Paradigm in Chemistry

L. Di Paola,[†] M. De Ruvo,[‡] P. Paci,[‡] D. Santoni,[§] and A. Giuliani*[,∥]





Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Protein contact network topology: a natural language for allostery

Luisa Di Paola[1] and Alessandro Giuliani[2]

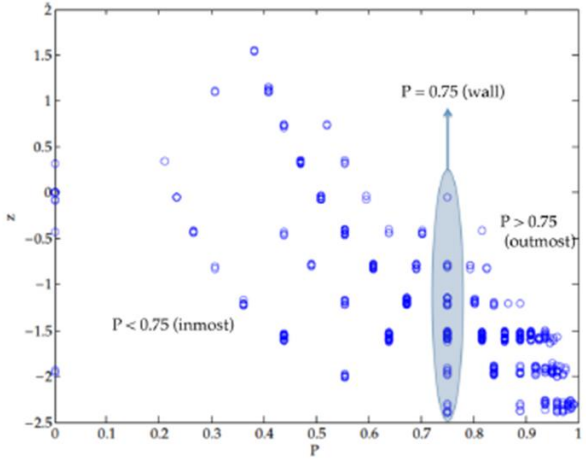


Figure 1. Dentist's chair map. The different roles of nodes on the basis of *P* values are shown in Table 2. The map refers to the hemoglobin structure (PDB code 1HBB) partitioned into 4 clusters.

The basic principles of topology-dynamics relations in networks: An empirical approach

Havva Kohestani [a], Mahbubeh Totonkuban [b], Luisa Di Paola [c,*], Virginia Todde [d], Alessandro Giuliani [d]

Physica A 508 (2018) 584–594

Table 1
Principal Components (PC) loading profile (Pearson correlation coefficients between original variables and components) of the complete data set considered as a whole (no distinction among different network classes).

	PC1	PC2
Extent	**−0.857**	−0.161
Time	**0.842**	−0.244
Maxdis	0.069	**0.977**

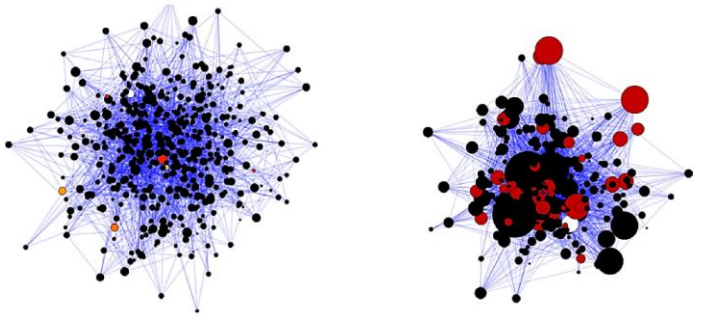
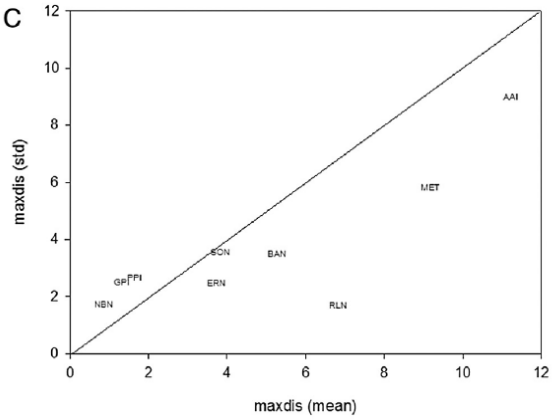


Fig. 4. Sample network perturbation in Turbine. The NBN (right) and AAI (left) networks undergo random single point perturbation. The white node is the source of perturbation that released dissipating energy effects by other nodes (in red) based on rules governing the dynamics of networks. Black nodes are the ones not affected by perturbation . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



⟸ AAI

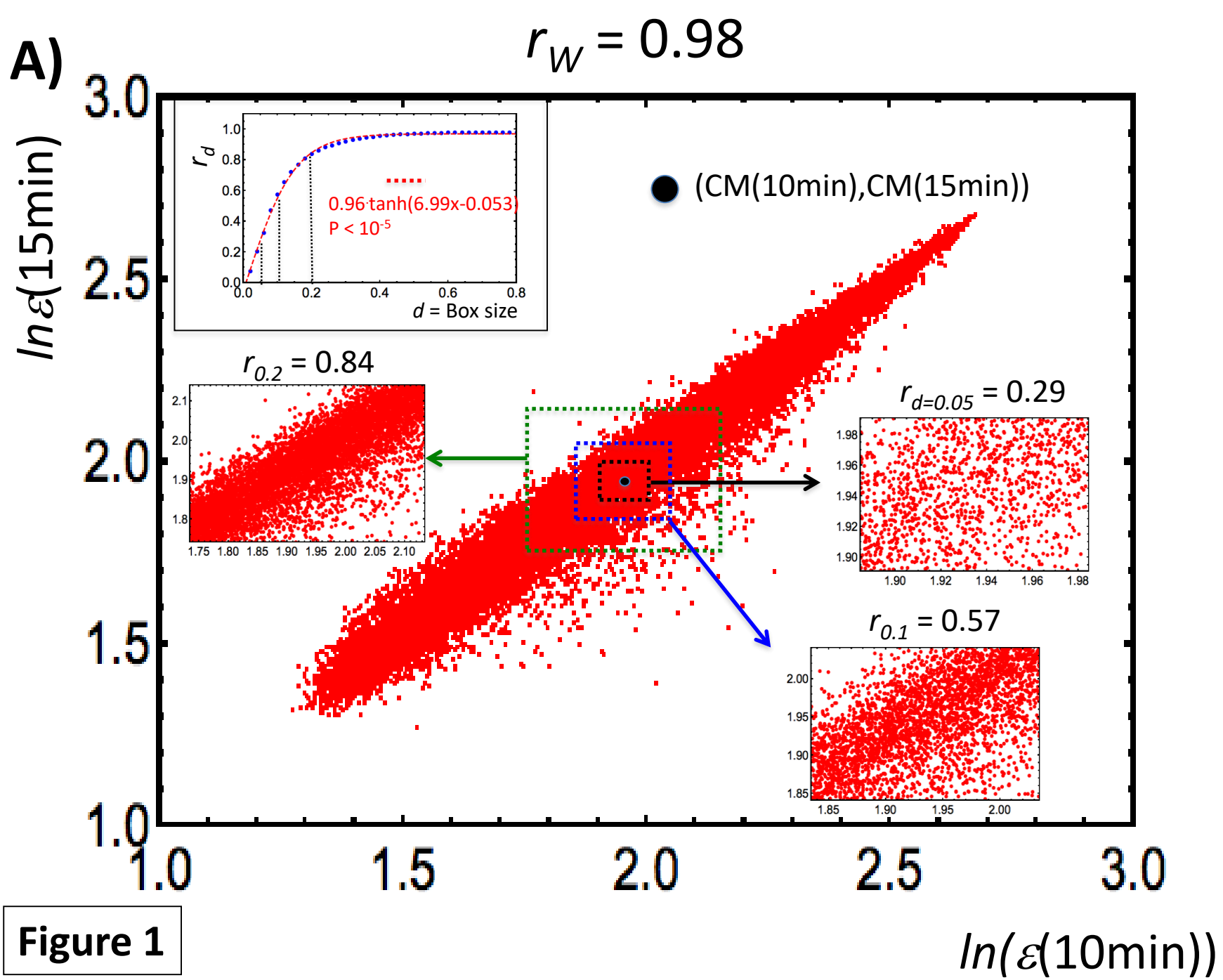

A)
r_W = 0.98
ln(ε(15min))
ln(ε(10min))
(CM(10min),CM(15min))
r_d
0.96·tanh(6.99x-0.053)
P < 10^-5
d = Box size
$r_{0.2}$ = 0.84
$r_{d=0.05}$ = 0.29
$r_{0.1}$ = 0.57

Figure 1

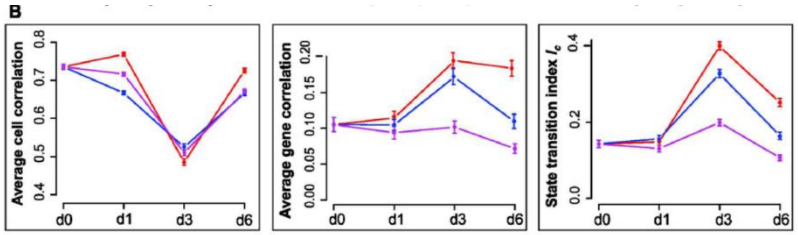
PLOS | BIOLOGY

RESEARCH ARTICLE

Cell Fate Decision as High-Dimensional Critical State Transition

Mitra Mojtahedi[1,2☯], Alexander Skupin[2,3☯], Joseph Zhou[2], Ivan G. Castaño[1,4], Rebecca Y. Y. Leong-Quong[1], Hannah Chang[5], Kalliopi Trachana[2], Alessandro Giuliani[6], Sui Huang[1,2]*

1 Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada, **2** Institute for Systems Biology, Seattle, Washington, United States of America, **3** Luxembourg Centre for Systems Biomedicine, Esch-sur Alzette, Luxembourg, **4** Corporación Parque Explora, Department of innovation and design, Medellin, Colombia, **5** 5AM Ventures, Menlo Park, California, United States of America, **6** Environment and Health Department, Istituto Superiore di Sanità, Roma, Italy



At tipping point (just before transition) cell-cell correlation decreases and gene-gene correlation increases.

nome	wild0h	wild1h	wild4h	Myd88ko0	MyD88ko1	MyD88ko4
AFFX-BioB	101,1	105,2	117	115,3	133,7	136,2
AFFX-BioB	188,3	185,7	168,4	223,4	239,3	234,8
AFFX-BioB	70,8	80,4	104,7	91,1	105,4	115,7
AFFX-BioC	289,2	265,1	321,8	293,3	321,2	344,6
AFFX-BioC	190,9	220,6	193	223,3	225,9	247,1
AFFX-BioD	167,8	178,8	186,8	198,2	231,3	248
AFFX-BioD	1295,9	1243,7	1404	1276,3	1742,7	1813,5
AFFX-CreX	2143,7	2484	2385,7	2353,3	2740,8	3031
AFFX-CreX	3532,9	4247,8	4606,1	4019,1	4995	6266,2
AFFX-DapX	2,8	2,8	6,2	8,6	1,3	2,5
AFFX-DapX	9,2	21,2	13,4	16,4	14,5	18,4
AFFX-DapX	7,1	1,5	3	2,9	2,7	1,4
AFFX-LysX	2,2	0,9	1,2	1,2	1,3	2,1
AFFX-LysX	2,9	10,6	4,7	2,4	2,7	2,3
AFFX-LysX	13,2	14,7	13,3	12,9	14,3	12,7
AFFX-Phe	2,9	2,1	1,3	1,7	1,7	2
AFFX-Phe	8	1,9	3,4	2,4	2,3	3
AFFX-Phe	9,2	22,3	9,8	4,9	18,7	7,9
AFFX-ThrX	1,6	4	7,8	2,9	4,2	2,3
AFFX-ThrX	16,9	10,1	8,3	19,7	14,3	13
AFFX-ThrX	3,5	8,6	6,2	14,4	11,6	2,7
AFFX-Trpn	5,7	3,8	14,4	5,5	2,4	3,3
AFFX-Trpn	3	5,5	2,3	5,2	2,1	2,6
AFFX-Trpn	1,2	1,6	1,3	1,8	1,4	0,6
AFFX-r2-E	106,3	126	128,2	116,1	173,8	179,5
AFFX-r2-E	237,7	233,8	228,1	232,6	276,9	312
AFFX-r2-E	198,1	163,4	156,8	155,9	203,4	227,3
AFFX-r2-E	423,4	387,3	333,4	362,5	473,4	487,7
AFFX-r2-E	414,1	441,5	385,9	430,7	515,5	554,6
AFFX-r2-E	1042,7	965,2	1038,7	916,2	1299,8	1480,9
AFFX-r2-E	1498,7	1630,5	1592,9	1575,2	2054,2	2177,2

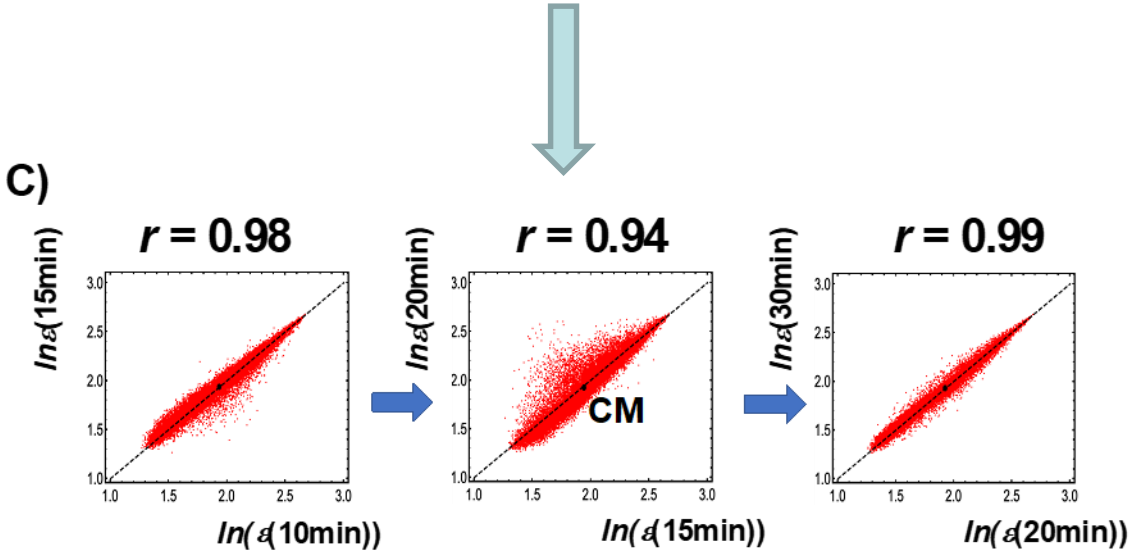
At equilibrium gene expression profile correlations (cell-cell) are near to unity

ROTATE

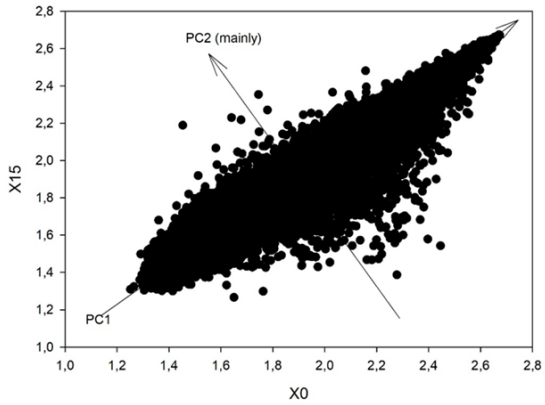
nome	AFFX-BioB-5_at	AFFX-BioB-M_at	AFFX-BioB-3_at	AFFX-BioC-5_at	AFFX-BioC-3_at	AFFX-BioDn-5_at	AFFX-BioDn-3_at	AFFX-CreX-5_at	AFFX-CreX-3_at
wild0h	101,1	188,3	70,8	289,2	190,9	167,8	1295,9	2143,7	3532,9
wild1h	105,2	185,7	80,4	265,1	220,6	178,8	1243,7	2484	4247,8
wild4h	117	168,4	104,7	321,8	193	186,8	1404	2385,7	4606,1
Myd88ko0h	115,3	223,4	91,1	293,3	223,3	198,2	1276,3	2353,3	4019,1
MyD88ko1h	133,7	239,3	105,4	321,2	225,9	231,3	1742,7	2740,8	4995
MyD88ko4h	136,2	234,8	115,7	344,6	247,1	248	1813,5	3031	6266,2

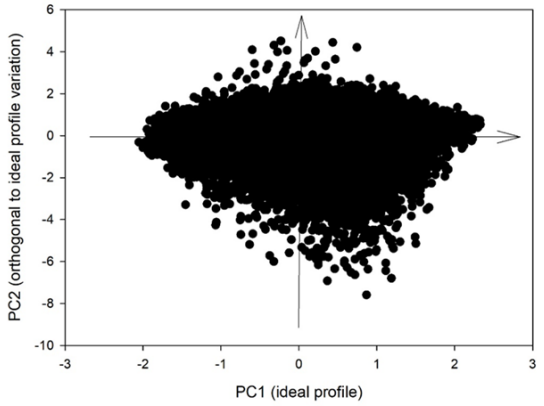
At equilibrium gene-gene correlations are relatively low (around 0.20-0.30 on average).

When cell kind transition happens, motion does not involve only ‘peripherical’ genes (sand grains) but invades (domino/violin effect) all the genome expression and provokes the motion of normally invariant (near the identitiy line) genes…

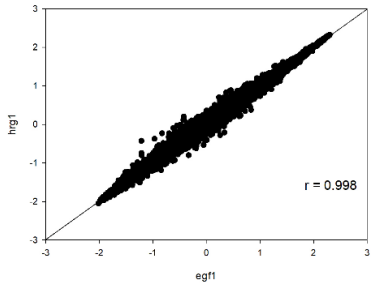


HRG case : 'directions of motion'

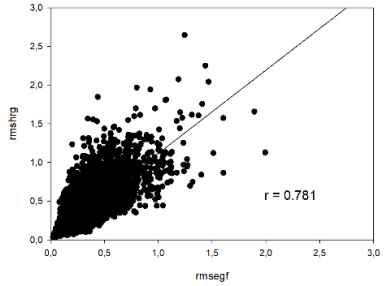




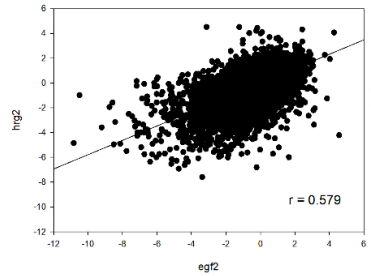
cell-kind attractor



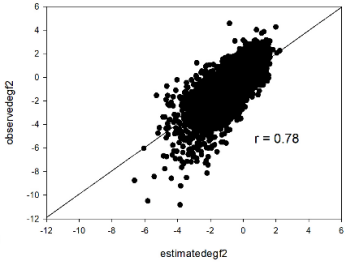
rms (egf-hrg)



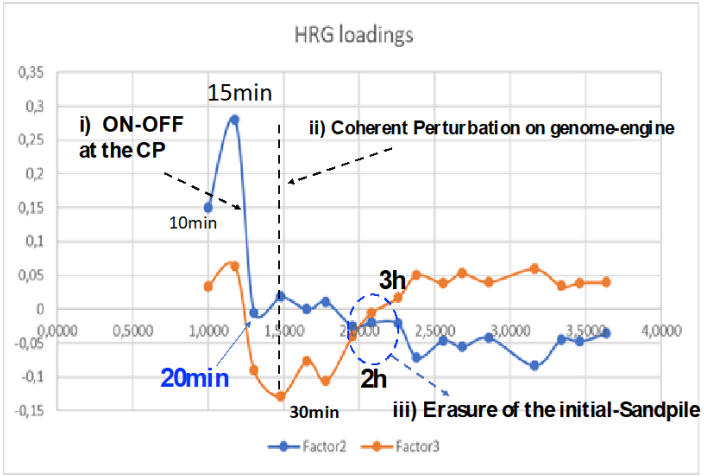
main oscillating mode (PC2)



egf2 splitted into hrg2-hrg3



egf2 = 0.58*hrg2 + 0.45hrg3



The independence of the phenomenology of the transition from the particular selected genes, suggests we can grasp the essential of the transition behavior by means of collective descriptors of the degree of order of gene expression pattern.

But we need something more:

Predictability of human differential gene expression

Megan Crow[a], Nathaniel Lim[b,c,d], Sara Ballouz[a], Paul Pavlidis[b,c], and Jesse Gillis[a,1]

PNAS | March 26, 2019 | vol. 116 | no. 13 | 6491–6500

The identification of genes that are differentially expressed provides a molecular foothold onto biological questions of interest. Whether some genes are more likely to be differentially expressed than others, and to what degree, has never been assessed on a global scale. Here, we reanalyze more than 600 studies and find that knowledge of a gene's prior probability of differential expression (DE) allows for accurate prediction of DE hit lists, regardless of the biological question. This result suggests redundancy in transcriptomics experiments that both informs gene set interpretation and highlights room for growth within the field.

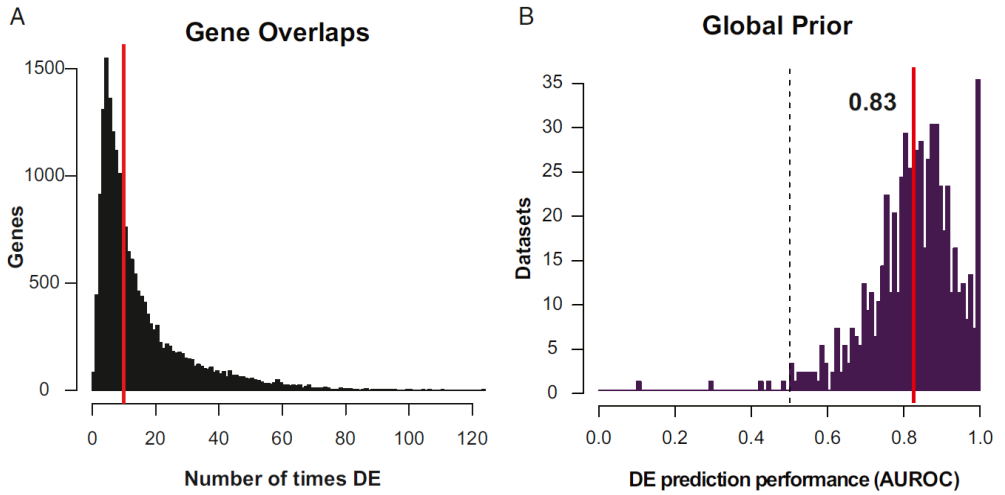
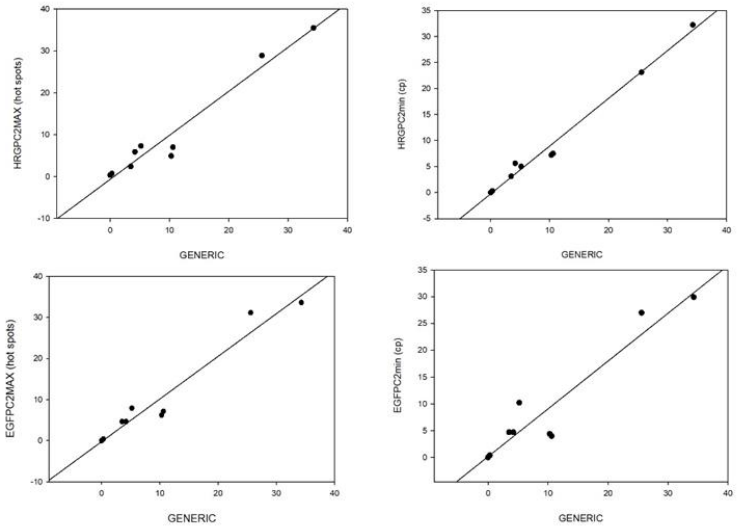


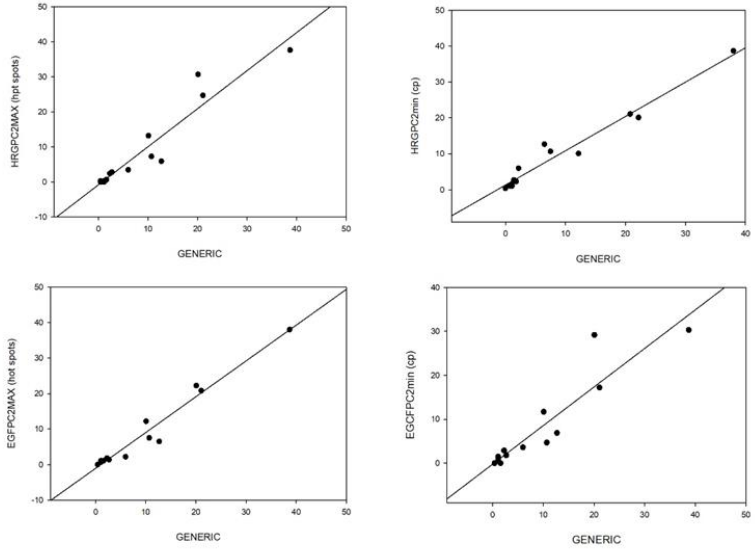
Fig. 2. The global DE prior accurately predicts DE hit lists. (*A*) Recurrence of differentially expressed genes across datasets. The red line indicates the mean. On average, each gene is DE in 10 expression studies. However, the distribution has a long, right-sided tail, indicating a small number of genes that are frequently DE. (*B*) Distribution of AUROC scores using the global DE prior to predict hit lists across the 635 studies. The red line indicates the mean, and the dashed line indicates the null (0.5). On average, the prior has very high performance, distinguishing ~80% of DE genes within each hit list, reflecting shared transcriptional features between studies.

MOLECULAR FUNCTION



MOLECULAR FUNCTION	HRGPC2MAX	HRGPC2min	EGFPC2MAX	EGFPC2min	GENERIC
transporter	5,9	5,6	4,6	4,7	4,2
translation regulator	0,7	0,3	0,4	0,4	0,3
cargo receptor	0,3	0	0	0	0
transcription regulator	7,3	5	7,9	10,2	5,2
catalytic	28,9	23,1	31,1	27	25,6
molecular function regulator	7	7,5	7,1	4	10,6
molecular transducer	4,9	7,2	6,2	4,4	10,3
structural	2,4	3,1	4,6	4,7	3,5
binding	35,5	32,2	33,6	29,9	34,3

BIOLOGICAL PROCESS



BIOLOGICAL PROCESS	HRGPC2MAXprocess	HRGPC2minprocess	EGFPC2MAXprocess	EGFPC2minprocess	GENERICprocess
cell comp biogenesis	0	0,7	0,8	1,5	1,1
cellular process	37,6	38	41,9	30,3	38,7
localization	13,2	12,2	13,3	11,7	10,1
reproduction	0,7	1,1	1,2	0	1,6
biological regulation	24,7	20,8	22,4	17,2	21,1
response stimulus	5,9	6,5	8,7	6,9	12,7
development	2,4	1,8	3,7	2,9	2,3
multi cellular organismal	7,3	7,5	6,6	4,7	10,7
adhesion	2,8	1,4	3,3	1,8	2,7
metabolism	30,7	22,2	25,3	29,2	20,1
cell proliferation	0,3	1,1	0,4	0,7	1,1
immune system	3,5	2,2	2,5	3,6	6
biological phase	0,3	0	0,4	0	0,4
rhytmic	0	0	0,4	0	0,4

Jorge Luis Borges: El Rigor de la Ciencia (On the Exactitude in Science)

.. In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.**”**
purportedly from Suárez Miranda, Travels of Prudent Men, Book Four, Ch. XLV, Lérida, 1658