Biological Statistical Mechanics: Anticipating Critical Transitions

Alessandro Giuliani
Istituto Superiore di Sanità, Roma

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

August 2005 | Volume 2 | Issue 8 | e124

Table 4. PPV of Research Findings for Various Combinations of Power (1 − β), Ratio of True to Not-True Relationships (*R*), and Bias (*u*)

1 − β	*R*	*u*	Practical Example	PPV
0.80	1:1	0.10	Adequately powered RCT with little bias and 1:1 pre-study odds	0.85
0.95	2:1	0.30	Confirmatory meta-analysis of good-quality RCTs	0.85
0.80	1:3	0.40	Meta-analysis of small inconclusive studies	0.41
0.20	1:5	0.20	Underpowered, but well-performed phase I/II RCT	0.23
0.20	1:5	0.80	Underpowered, poorly performed phase I/II RCT	0.17
0.80	1:10	0.30	Adequately powered exploratory epidemiological study	0.20
0.20	1:10	0.30	Underpowered exploratory epidemiological study	0.12
0.20	1:1,000	0.80	Discovery-oriented exploratory research with massive testing	0.0010
0.20	1:1,000	0.20	As in previous example, but with more limited bias (more standardized)	0.0015

http://www.nature.com/news/reproducibility-1.17552

Deming, data and observational studies

A process out of control and needing fixing

"Any claim coming from an observational study is most likely to be wrong." Startling, but true. Coffee causes pancreatic cancer. Type A personality causes heart attacks. Trans-fat is a killer. Women who eat breakfast cereal give birth to more boys. All these claims come from observational studies; yet when the studies are carefully examined, the claimed links appear to be incorrect. What is going wrong? Some have suggested that the scientific method is failing, that nature itself is playing tricks on us. But it is our way of studying nature that is broken and that urgently needs mending, say **S. Stanley Young** and **Alan Karr**; and they propose a strategy to fix it.

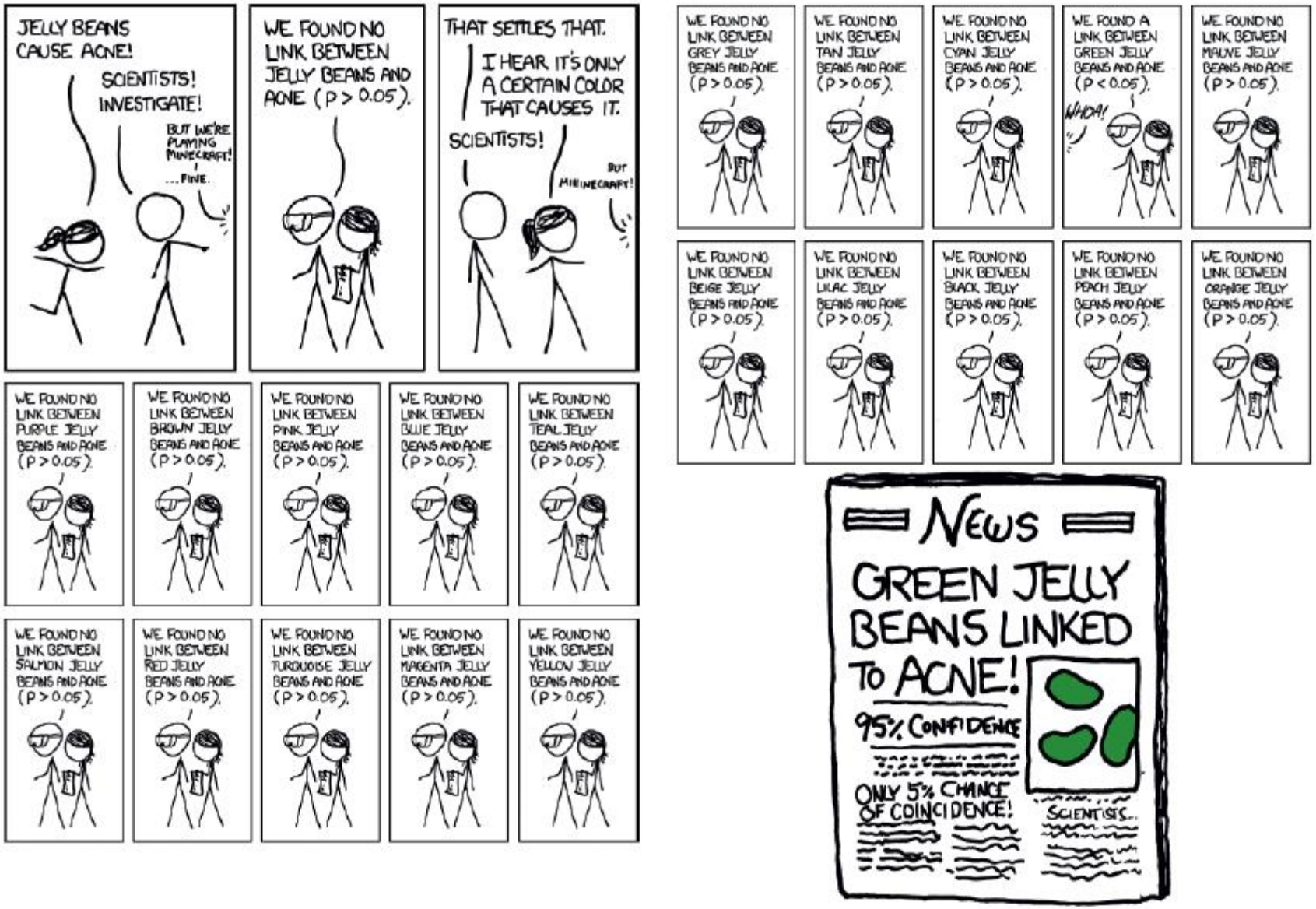
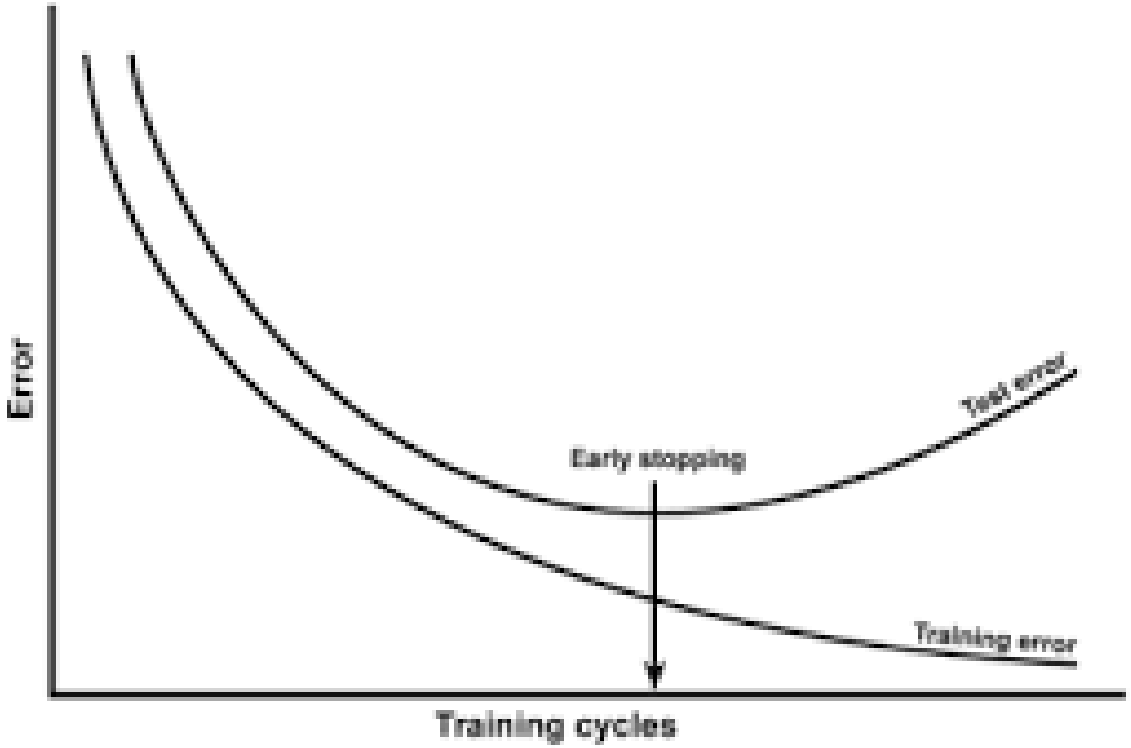


Figure 1. There is no overall effect of jelly beans on acne. Bummer. How about subgroups? Often subgroups are explored without alerting the reader to the number of questions at issue. Courtesy xkcd, http://xkcd.com/882/

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

Overfitting refers to a model that fits the training data too well. **Overfitting** happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data.



The overfitting phenomenon tells us that any data set has inside both general (applicable to an entire class of problems) and idyosincratic (specific of the data set) information.

Too precise fitting provokes the model to get stuck into idyosincratic properties: after a given complexity level we start to model noise.

The observables are not ‘the real thing’, the ‘real thing’ is latent.

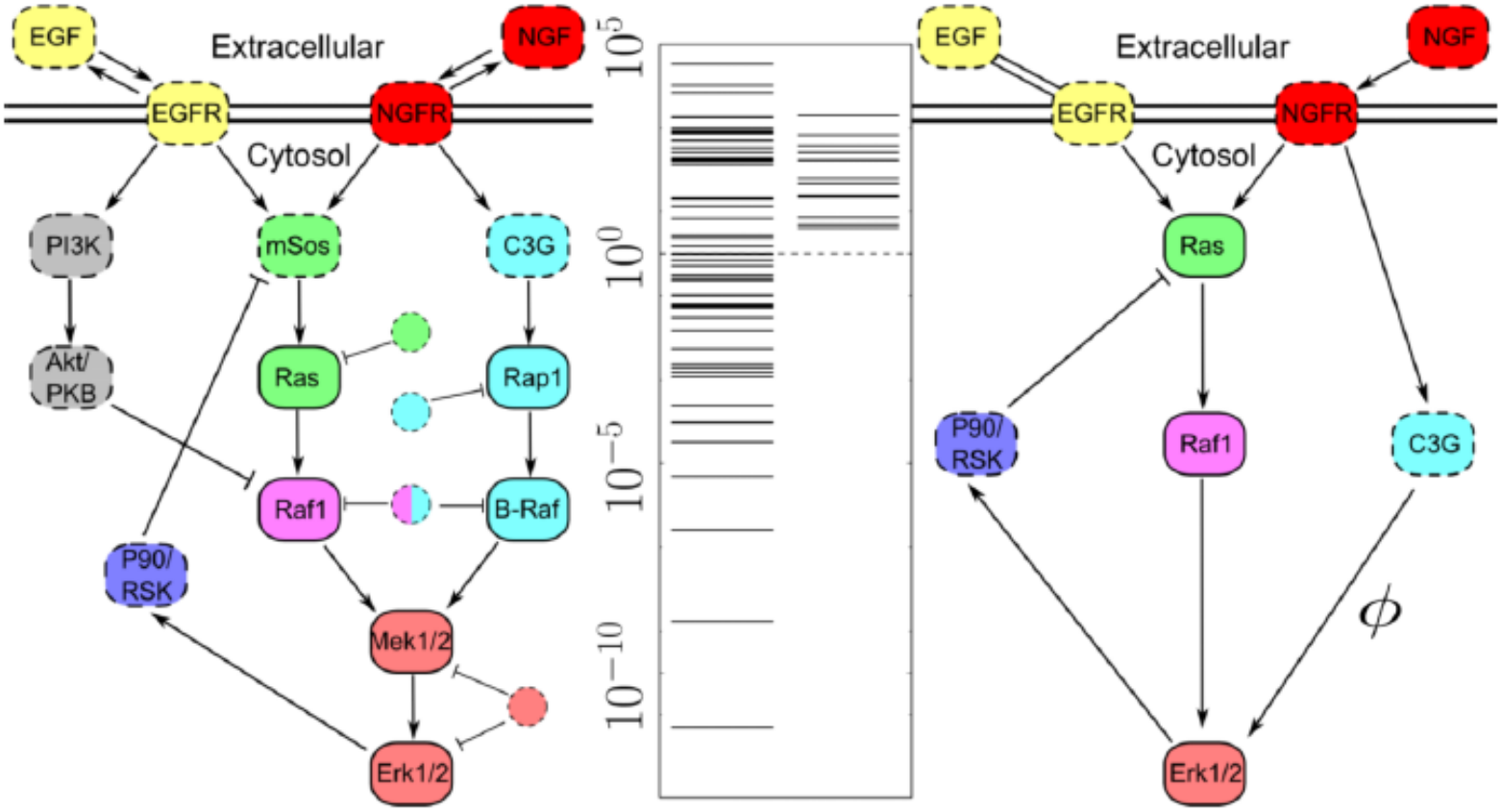
THE JOURNAL OF CHEMICAL PHYSICS **143**, 010901 (2015)

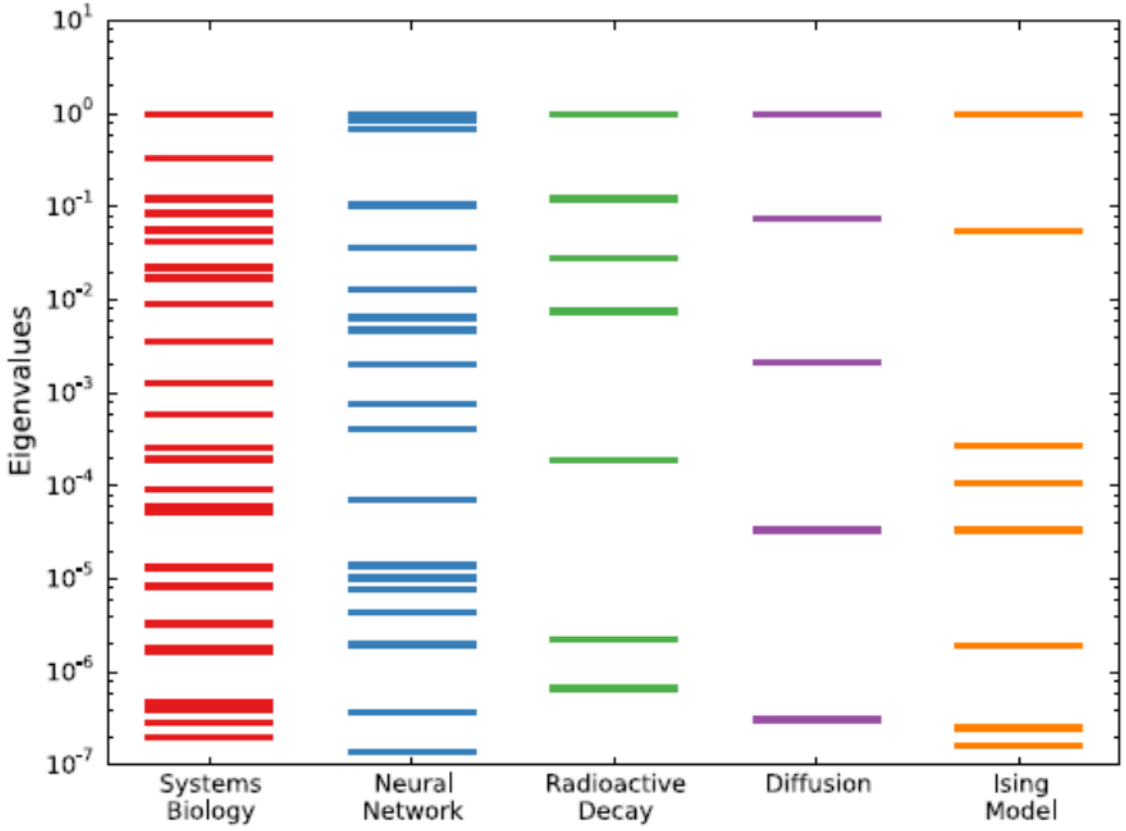
Perspective: Sloppiness and emergent theories in physics, biology, and beyond

Mark K. Transtrum,[1] Benjamin B. Machta,[2] Kevin S. Brown,[3,4] Bryan C. Daniels,[5] Christopher R. Myers,[6,7] and James P. Sethna[6]
[1]*Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA*
[2]*Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Departments of Biomedical Engineering, Physics, Chemical and Biomolecular Engineering, and Marine Sciences, University of Connecticut, Storrs, Connecticut 06269, USA*
[4]*Institute for Systems Genomics, University of Connecticut, Storrs, Connecticut 06030-1912, USA*
[5]*Center for Complexity and Collective Computation, Wisconsin Institute for Discovery, University of Wisconsin, Madison, Wisconsin 53715, USA*
[6]*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA*
[7]*Institute of Biotechnology, Cornell University, Ithaca, New York 14853, USA*

(Received 2 February 2015; accepted 4 June 2015; published online 1 July 2015)

As a young physicist, Dyson paid a visit to Enrico Fermi[1] (recounted in Ditley, Mayer, and Loew[2]). Dyson wanted to tell Fermi about a set of calculations that he was quite excited about. Fermi asked Dyson how many parameters needed to be tuned in the theory to match experimental data. When Dyson replied there were four, Fermi shared with Dyson a favorite adage of his that he had learned from Von Neumann: "with four parameters I can fit an elephant, and with five I can make him wiggle his trunk." Dejected, Dyson took the next bus back to Ithaca.

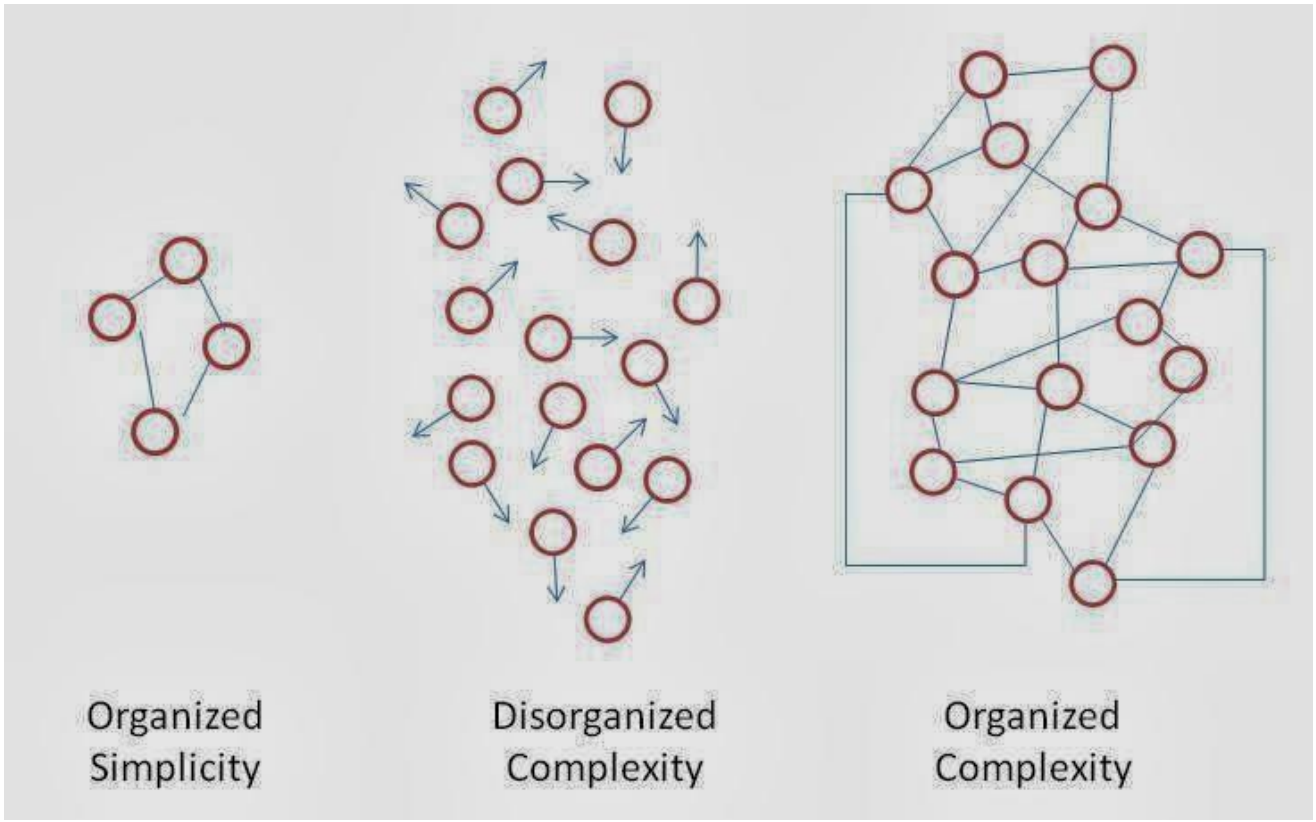

EGF
Extracellular
NGF
EGFR
NGFR
Cytosol
PI3K
mSos
C3G
Akt/
PKB
Ras
Rap1
Raf1
B-Raf
P90/
RSK
Mek1/2
Erk1/2
10^{-5}
10^{0}
10^{-5}
10^{-10}
EGF
Extracellular
NGF
EGFR
NGFR
Cytosol
Ras
P90/
RSK
Raf1
C3G
ϕ
Erk1/2


Eigenvalues
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Systems Biology
Neural Network
Radioactive Decay
Diffusion
Ising Model

SCIENCE AND COMPLEXITY

By WARREN WEAVER
Rockefeller Foundation, New York City

"Science and Complexity", American Scientist, 36: 536 (1948).

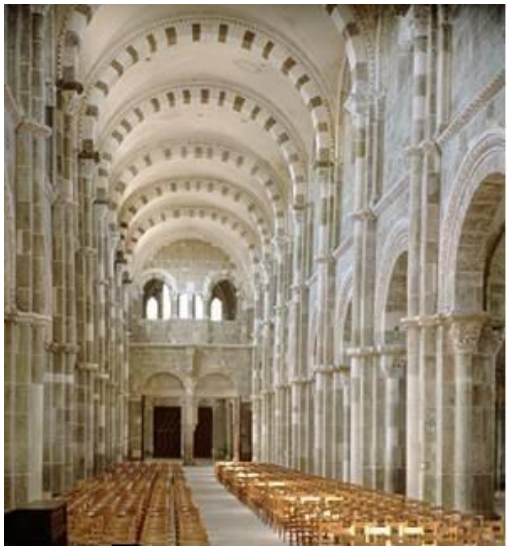


The middle way

R. B. Laughlin*, David Pines†‡§, Joerg Schmalian¶, Branko P. Stojković‖**, and Peter Wolynes††

32–37 | PNAS | January 4, 2000 | vol. 97 | no. 1

Mesoscopic organization in soft, hard, and biological matter is examined in the context of our present understanding of the principles responsible for emergent organized behavior (crystallinity, ferromagnetism, superconductivity, etc.) at long wavelengths in very large aggregations of particles. Particular attention is paid to the possibility that as-yet-undiscovered organizing principles might be at work at the mesoscopic scale, intermediate between atomic and macroscopic dimensions, and the implications of their discovery for biology and the physical sciences. The search for the existence and universality of such rules, the proof or disproof of organizing principles appropriate to the mesoscopic domain, is called the middle way.

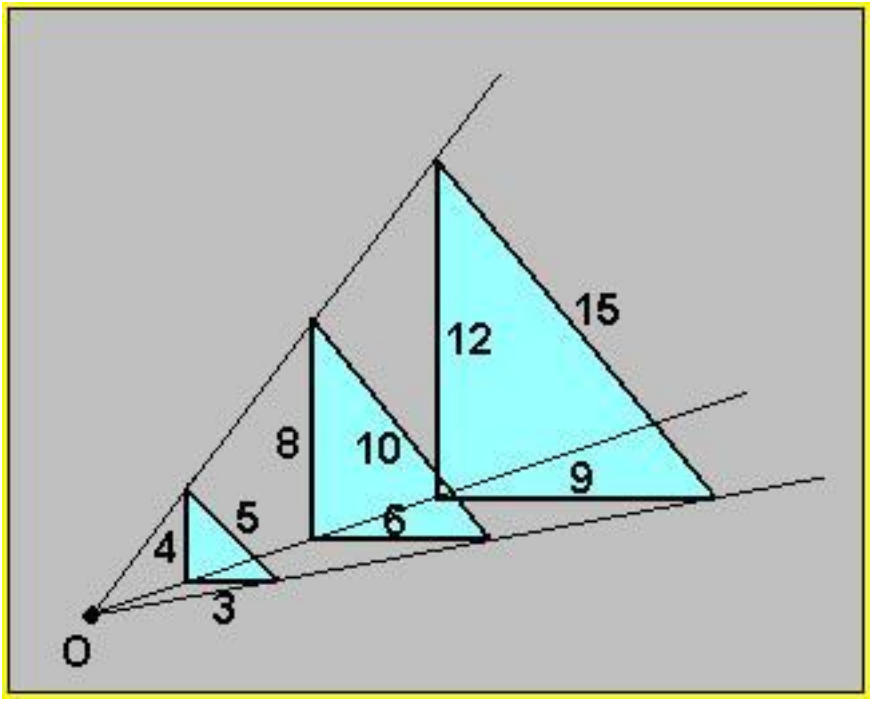


A common (even if often misunderstood) feature of biological structures in both space and time are the presence of few ‘priviliged’ forms.

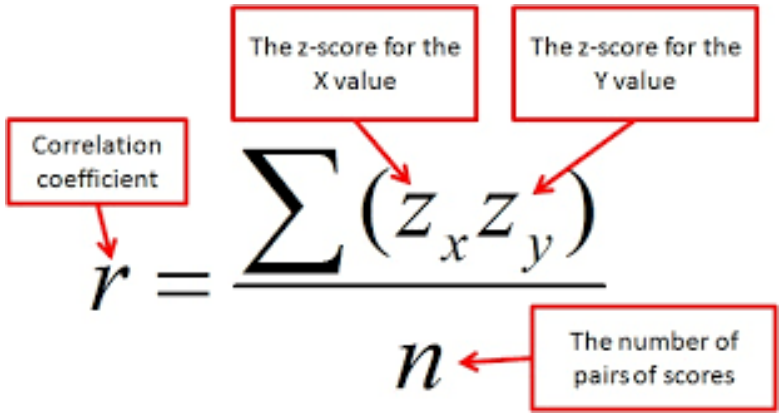
1) Around 1000 folds are sufficient to get rid of any protein structure
2) Any metazoan can be built by no more than 250 tissue types (with a very invariant gene expression profile).
3) Four basic ‘body-plans’ (bauplan) are at the basis of animal morphologies.
4) Four main rhytmic activities explain heartbeat dynamics.
……

The presence of few discrete priviliged forms has important consequences on data analysis strategies..

A shape is kept invariant if the relations between the mutual distances of a set of landmarks is kept invariant.



Here: 3/5 = 6/10 = 9/15…



$$r = r_{xy} = \frac{\mathrm{Cov}(x, y)}{S_x \times S_y}$$

Drug Discovery Today • Volume 22, Number 7 • July 2017 REVIEWS

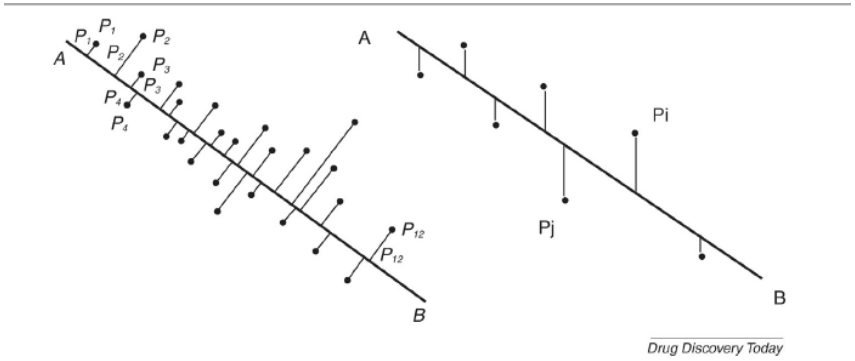
ELSEVIER

The application of principal component analysis to drug discovery and biomedical data

Alessandro Giuliani

Environment and Health Department, Istituto Superiore di Sanità, Roma, Italy

Reviews • INFORMATICS



LIII. *On Lines and Planes of Closest Fit to Systems of Points in Space.* By KARL PEARSON, *F.R.S., University College, London* *.

(1) IN many physical, statistical, and biological investigations it is desirable to represent a system of points in plane, three, or higher dimensioned space by the "best-fitting" straight line or plane. Analytically this consists in taking

$$y = a_0 + a_1 x, \quad \text{or} \quad z = a_0 + a_1 x + b_1 y,$$
$$\text{or} \quad z = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots + a_n x_n,$$

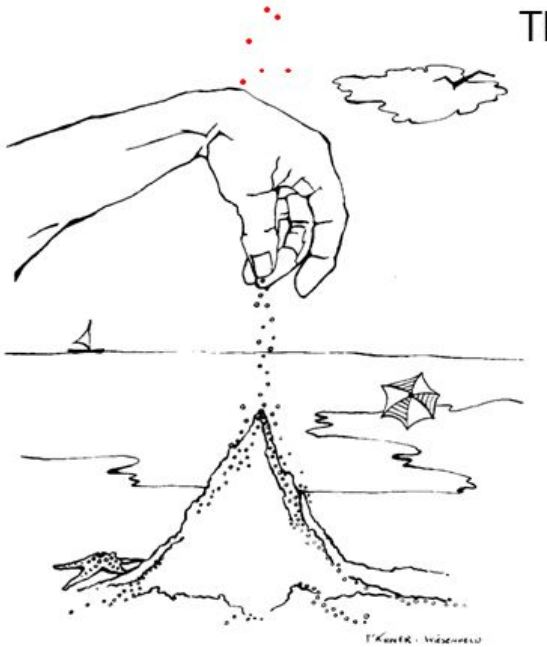
Laws do not exist as such
Forms do exist

Biology has to do with forms

Variables (gene expressions, metabolite concentrations, behavioral tests..) are relevant only if they allow to describe a form.

Self-Organizing-Criticality (SOC)

Avalanche Behavior



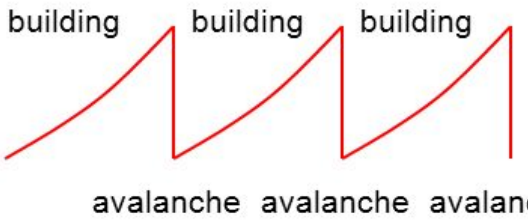
The sand pile builds . . . grain . . . by grain . . .

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.

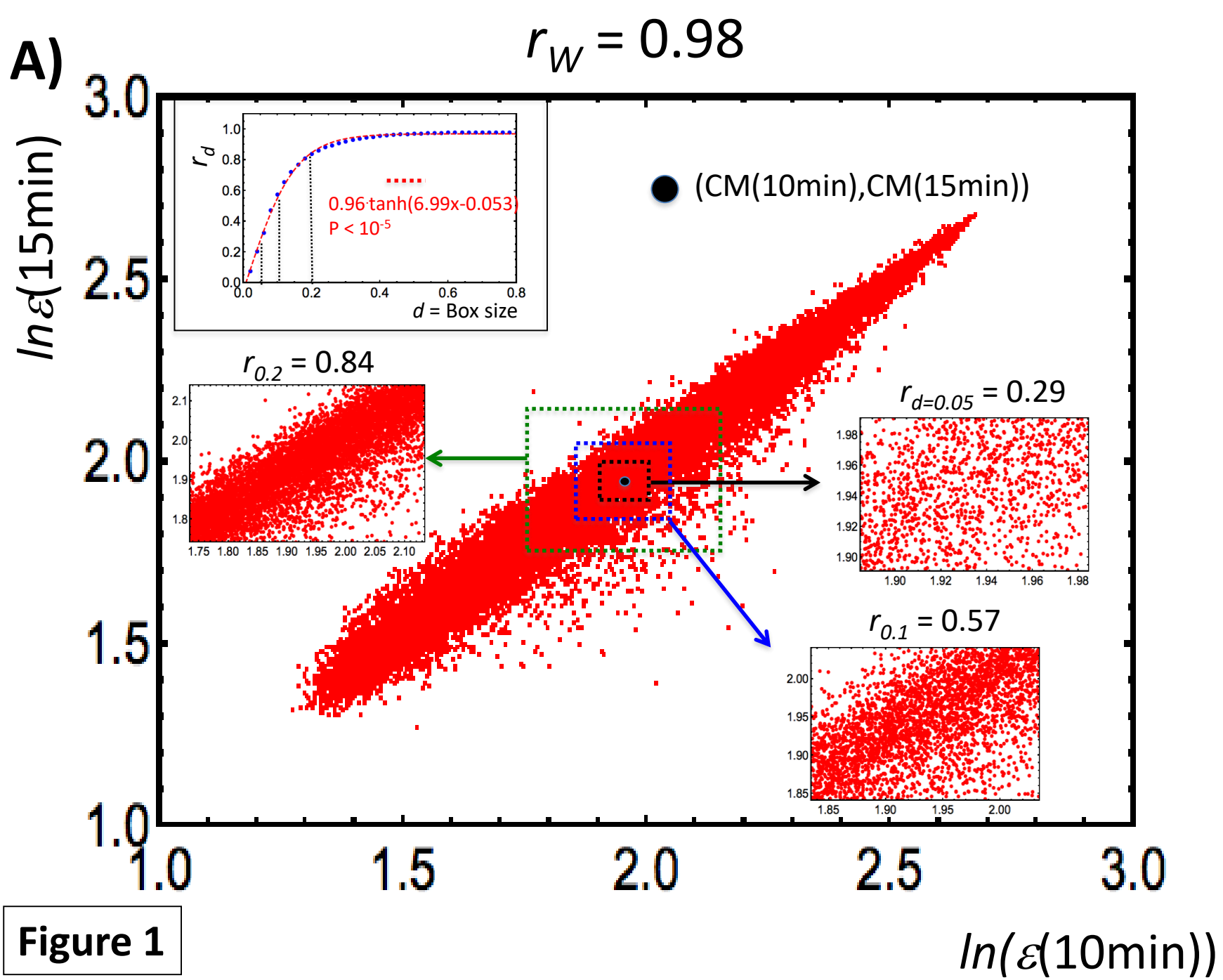
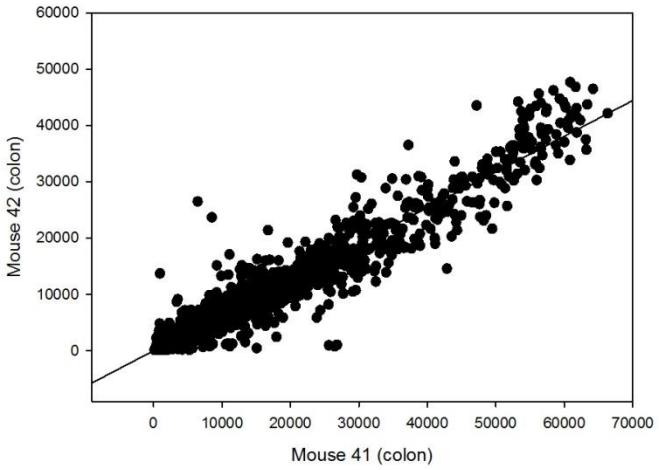
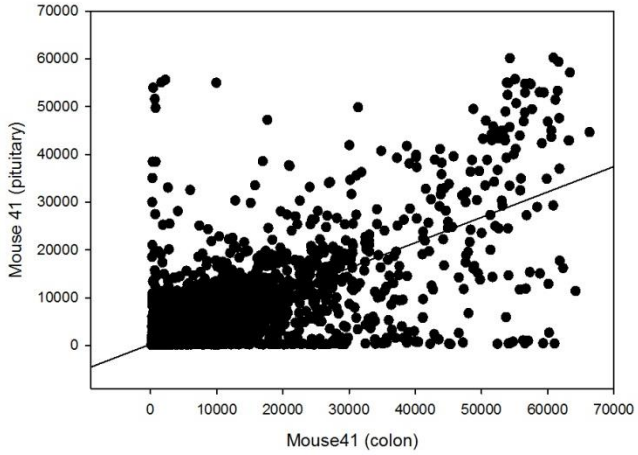

A)
r_W = 0.98
ln(ε(15min))
ln(ε(10min))
(CM(10min),CM(15min))
r_d
0.96·tanh(6.99x-0.053)
P < 10^-5
d = Box size
$r_{0.2}$ = 0.84
$r_{d=0.05}$ = 0.29
$r_{0.1}$ = 0.57

Figure 1



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

Detecting the optimal scale for the analysis is the most crucial problem in science.



Environmental Practice 16: 281–286 (2014)

Defining Appropriate Spatial and Temporal Scales for Ecological Impact Analysis[1]

Harriet L. Nash

Levin argues that scale is "the fundamental conceptual problem in ecology, if not in all of science" (Levin, 1992,

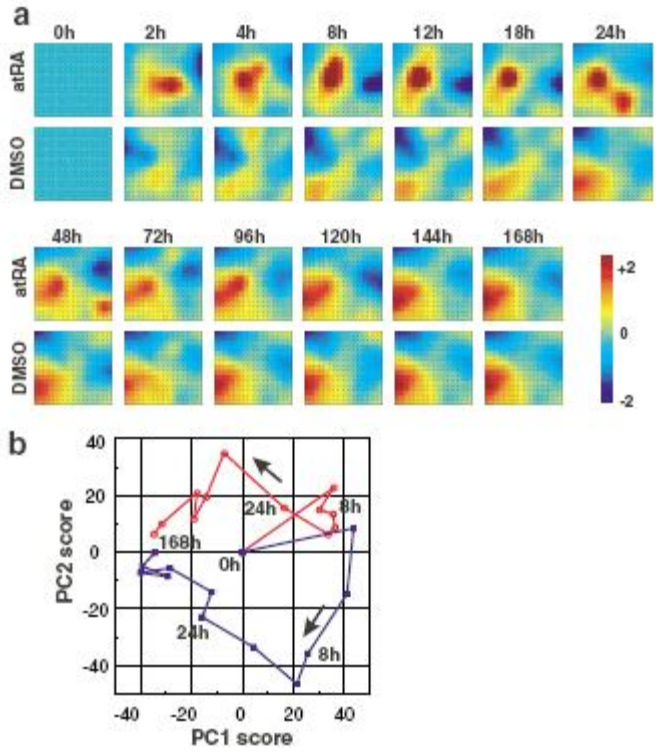
Cell Fates as High-Dimensional Attractor States of a Complex Gene Regulatory Network

Sui Huang,[1,*] Gabriel Eichler,[1] Yaneer Bar-Yam,[2] and Donald E. Ingber[1]

[1]*Vascular Biology Program, Departments of Pathology & Surgery, Children's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA*

[2]*New England Complex Systems Institute, Cambridge, Massachusetts 02138, USA*

(Received 13 September 2004; published 1 April 2005)



Contents lists available at ScienceDirect

Journal of Theoretical Biology

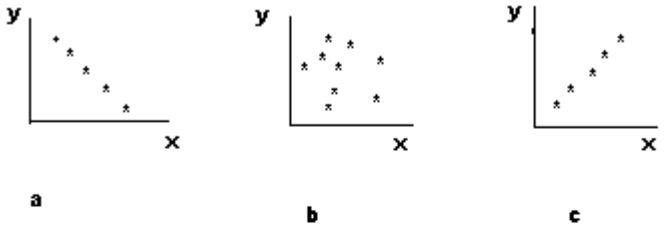
journal homepage: www.elsevier.com/locate/jtb

Ensembles, dynamics, and cell types: Revisiting the statistical mechanics perspective on cellular regulation☆

Stefan Bornholdt[a,*], Stuart Kauffman[b]

[a] Institute for Theoretical Physics, University of Bremen, 28359 Bremen, Germany
[b] Institute for Systems Biology, Seattle, WA 98109, USA

There is now, however, good evidence that cell types are high dimensional attractors. Huang and collaborators (Huang et al., 2005) took HL60, and induced differentiation to polymorphoneuterophil, PMN, using vitamin A and another substance. They followed gene expression of all 23,000 genes using gene arrays at three time points for both treatments. This shows that the gene expression pattern diverged for the two treatments at the temporal midpoint, then converged to the same new expression pattern corresponding to being a PMN. So trajectories converged on the same new pattern of expression from two different directions in high dimensional space, demonstrating that target pattern is an attractor of the dynamics.



$$r = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle (x_i - \langle x \rangle)^2 \rangle} \sqrt{\langle (y_i - \langle y \rangle)^2 \rangle}}$$

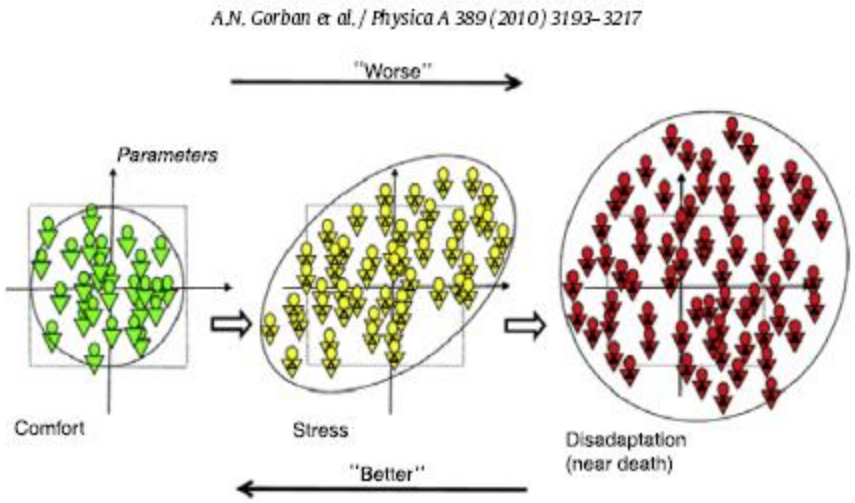
Pearson correlation coefficient (something we should learn at the introductory statistics courses)
Is the basic metrics for approaching organized complexity

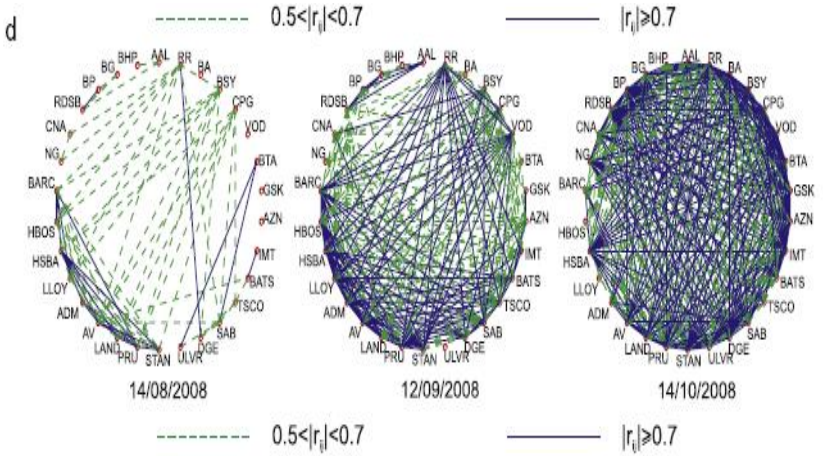
Physica A 389 (2010) 3193–3217

Correlations, risk and crisis: From physiology to finance

Alexander N. Gorban [a,*], Elena V. Smirnova [b], Tatiana A. Tyukina [a]

[a] University of Leicester, Leicester, LE1 7RH, UK
[b] Siberian Federal University, Krasnoyarsk, 660041, Russia





A Multi Scale Graph Theoretical Approach To Gene Regulation Networks: a Case Study In Atrial Fibrillation

Federica Censi, Alessandro Giuliani, Pietro Bartolini, Giovanni Calcagnini

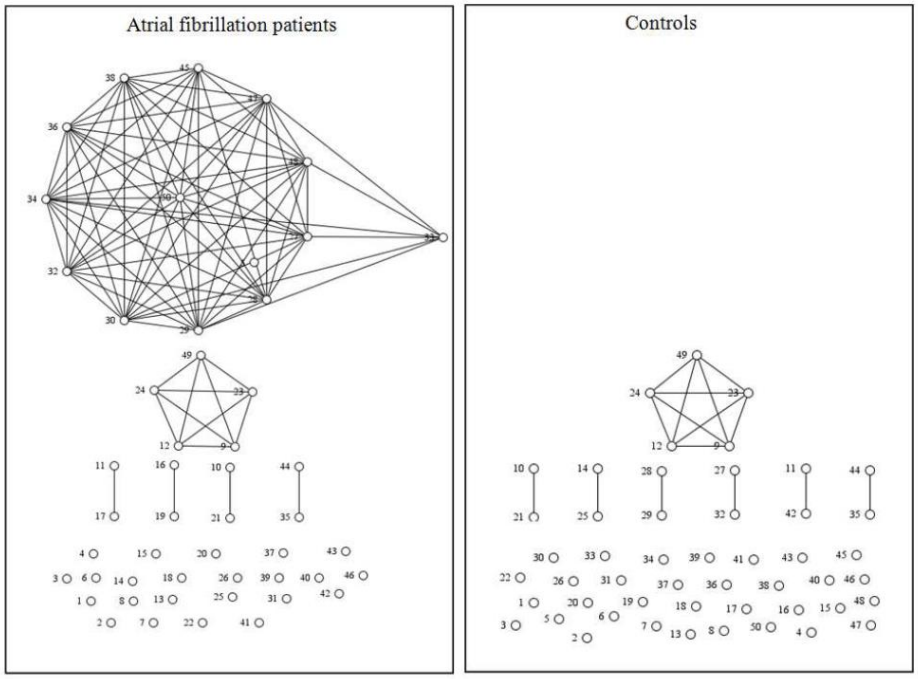
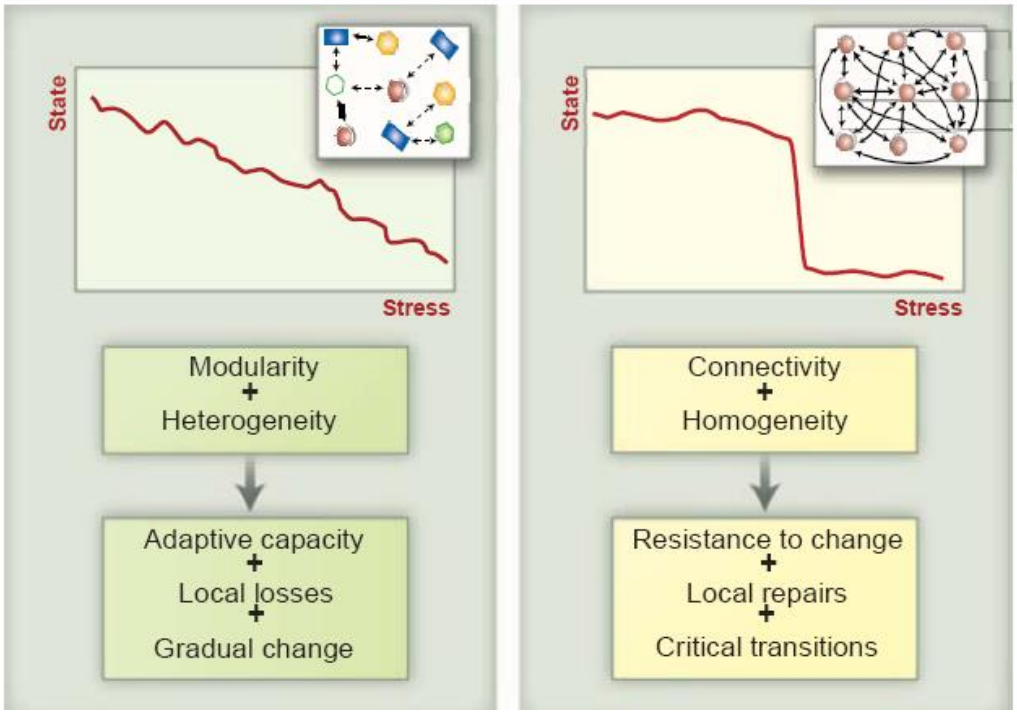


Figure 1. On the left the network relative to AF patients is reported, on the right the corresponding network for controls. It is immediate to note the invariance of the fully connected sub-network constituted by 5 genes involved in protein synthesis, and the presence in the AF patients of a strongly connected sub-network induced by the disease

Anticipating Critical Transitions

Marten Scheffer,[1,2*] Stephen R. Carpenter,[3] Timothy M. Lenton,[4] Jordi Bascompte,[5] William Brock,[6] Vasilis Dakos,[1,5] Johan van de Koppel,[7,8] Ingrid A. van de Leemput,[1] Simon A. Levin,[9] Egbert H. van Nes,[1] Mercedes Pascual,[10,11] John Vandermeer[10]

Tipping points in complex systems may imply risks of unwanted collapse, but also opportunities for positive change. Our capacity to navigate such risks and opportunities can be boosted by combining emerging insights from two unconnected fields of research. One line of work is revealing fundamental architectural features that may cause ecological networks, financial markets, and other complex systems to have tipping points. Another field of research is uncovering generic empirical indicators of the proximity to such critical thresholds. Although sudden shifts in complex systems will inevitably continue to surprise us, work at the crossroads of these emerging fields offers new approaches for anticipating critical transitions.



Predicting the transition from normal aging to Alzheimer's disease: A statistical mechanistic evaluation of FDG-PET data

Marco Pagani [a,b,*], Alessandro Giuliani [c], Johanna Öberg [d], Andrea Chincarini [e], Silvia Morbelli [f], Andrea Brugnolo [g], Dario Arnaldi [g], Agnese Picco [g], Matteo Bauckneht [f], Ambra Buschiazzo [f], Gianmario Sambuceti [f], Flavio Nobili [g]

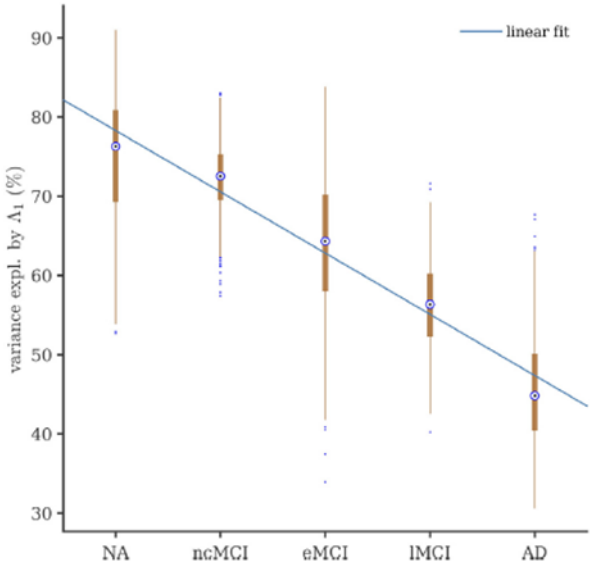
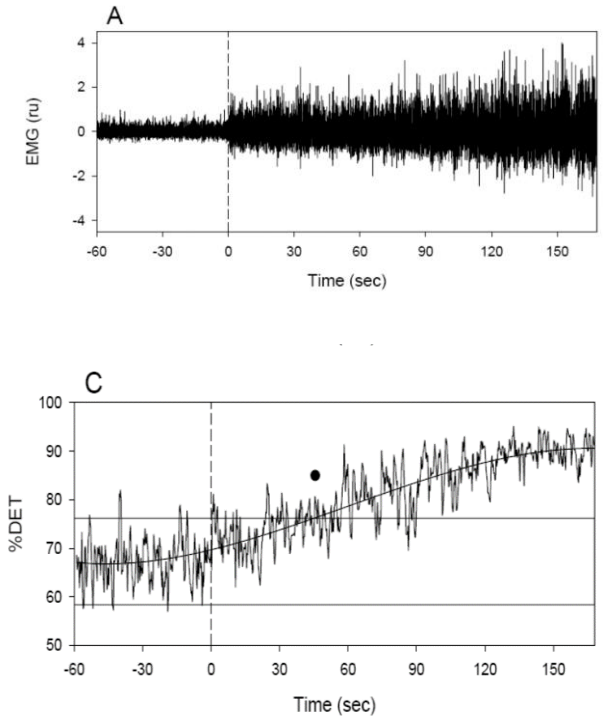


Fig. 1. The dynamics of the loss of order along the clinical status. Y-axis: variance explained by the first component; X-axis: disease severity. NA: normal aging; ncMCI: MCI patients not converting to AD at 5 years follow up; eMCI: MCI patients that converted to AD later than 2 years; lMCI: MCI patients that converted to AD within 2 years; AD: patients with mild AD dementia. The point distribution around the center of mass corresponds to bootstrap simulation.

Recurrence quantification analysis of surface electromyographic signal: Sensitivity to potentiation and neuromuscular fatigue

Claire Morana, Sofiane Ramdani, Stéphane Perrey, Alain Varray *

EA 2991 Motor Efficiency and Deficiency Laboratory, University of Montpellier 1, Faculty of Sport Sciences, 700 Avenue du Pic Saint Loup, 34090 Montpellier, France



Recurrence quantification analysis of the logistic equation with transients

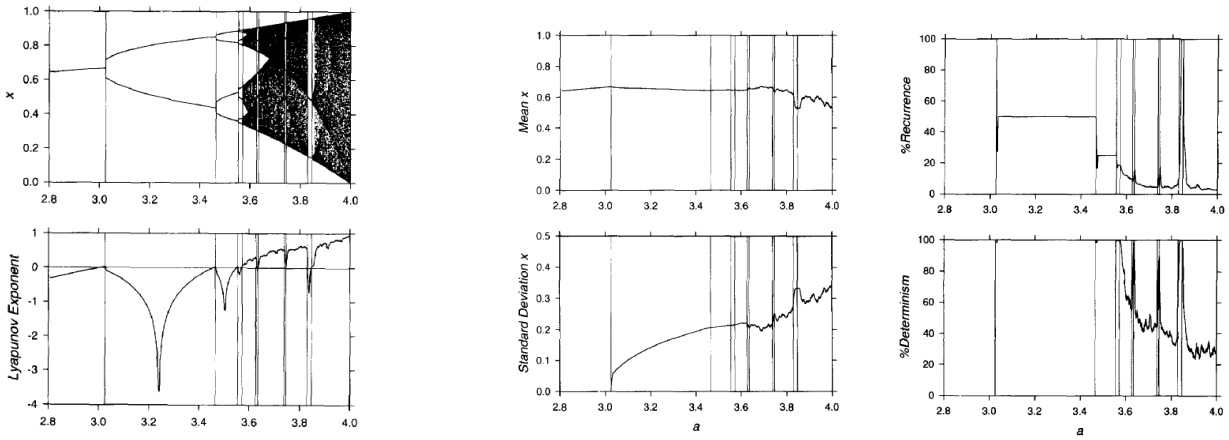
L.L. Trulla [a], A. Giuliani [a], J.P. Zbilut [b,1], C.L. Webber Jr. [c,2,*]

[a] Institute for Research on Senescence, Sigma Tau, Via Pontina Km 30400, 00040 Pomezia, Rome, Italy
[b] Department of Molecular Biophysics and Physiology, Rush Medical College, 1653 W. Congress Pkwy., Chicago, IL 60612, USA
[c] Department of Physiology, Loyola University Chicago, Stritch School of Medicine, 2160 S. First Ave., Maywood, IL 60153, USA

Physics Letters A 223 (1996) 255–260

$$x_{n+1} = a x_n (1 - x_n/k).$$

$$a = 2.8 \text{ to } 4.0$$

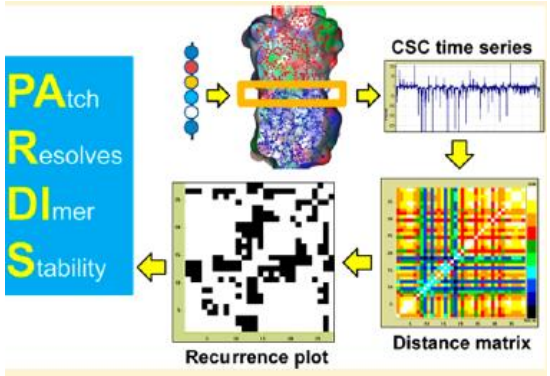


JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING
Article
Cite This: J. Chem. Inf. Model. 2018, 58, 733–745
pubs.acs.org/jcim

Partner-Specific Prediction of Protein-Dimer Stability from Unbound Structure of Monomer

Hamid Hadi-Alijanvand*,† and Maryam Rouhani†

†Department of Biological Sciences, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, 45137-66731, Iran



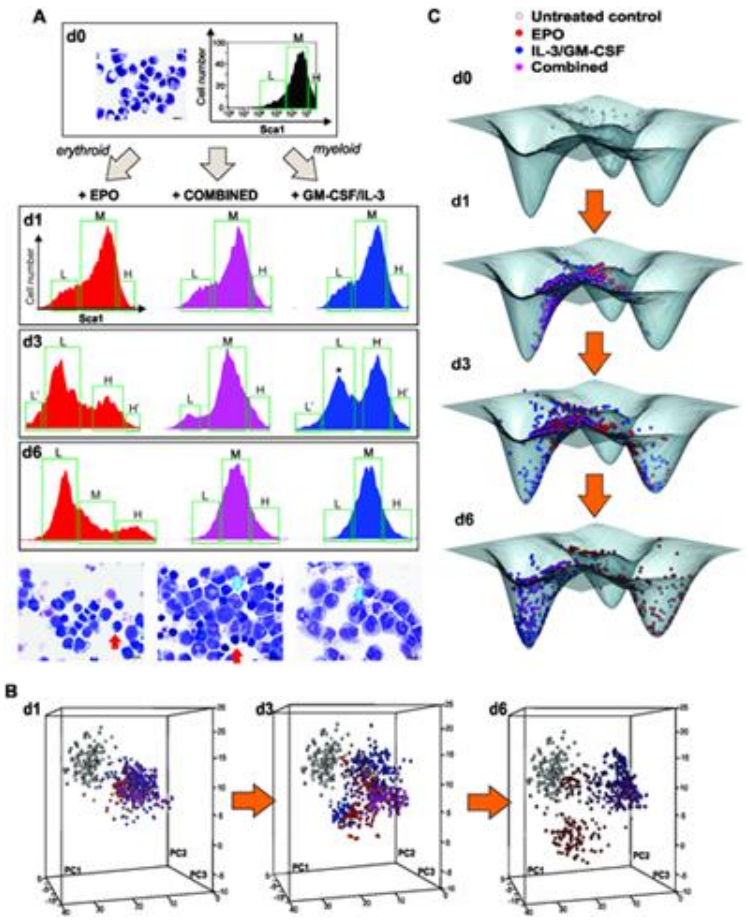
PLOS | BIOLOGY

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



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-Dap	2,8	2,8	6,2	8,6	1,3	2,5
AFFX-Dap	9,2	21,2	13,4	16,4	14,5	18,4
AFFX-Dap	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

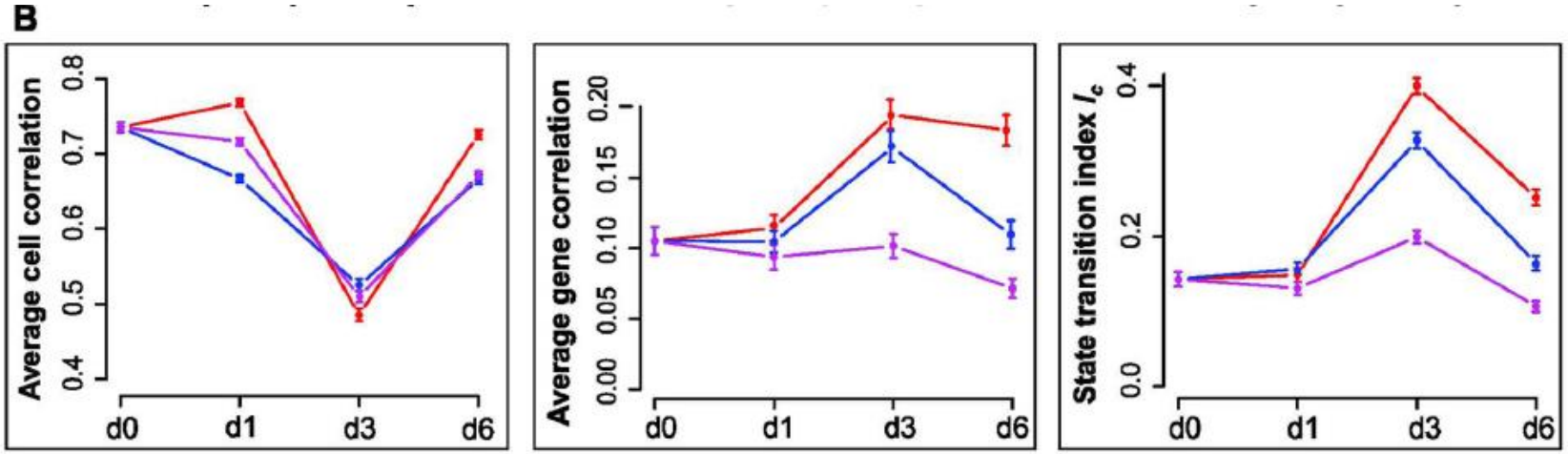
The attractor-like properties of cell kind implies (at the stable state) a near to unity positive Pearson correlation between expression profiles.

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
Myd88ko 0h	115,3	223,4	91,1	293,3	223,3	198,2	1276,3	2353,3	4019,1
MyD88ko 1h	133,7	239,3	105,4	321,2	225,9	231,3	1742,7	2740,8	4995
MyD88ko 4h	136,2	234,8	115,7	344,6	247,1	248	1813,5	3031	6266,2

On the contrary, gene-gene correlations are relatively low with
both positive and negative values (around 0.20-0.30) but greater than what expected by chance alone.

In the vicinity of a transition between cells correlation **decreases** because the previous order (driving the correlation) starts to fade away with different rates and trajectories in different cells.

In the vicinity of a transition between genes correlation **increases** because the non synchronized changes in gene expression increase variance and thus correlation.



$$I_C(t) = \frac{\langle |R(g_i, g_j)| \rangle}{\langle R(S^k, S^l) \rangle},$$

The energy landscape in non-biological and biological molecules

Hans Frauenfelder and Daan Thorn Leeson

The concept of energy landscapes promises to connect aspects of biology, chemistry and physics. A recent paper highlights the need for continuous exchange of information between fields to maximize the utility of this idea.

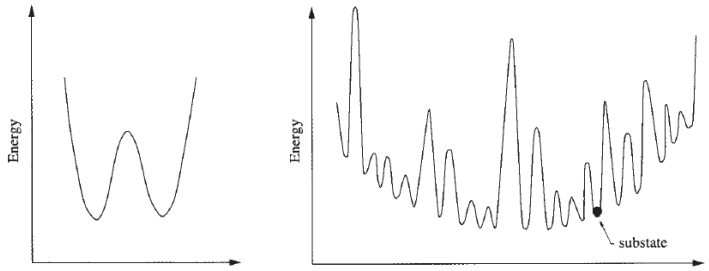


Fig. 1 A very simple and a very complex energy landscape. ***a***, The energy landscape of ammonia, NH_3. The conformational coordinate describes the distance of the nitrogen atom from the plane of the three hydrogen atoms. ***b***, A highly simplified energy landscape of a protein. In reality a landscape is a function of 3N coordinates, where N is very large.

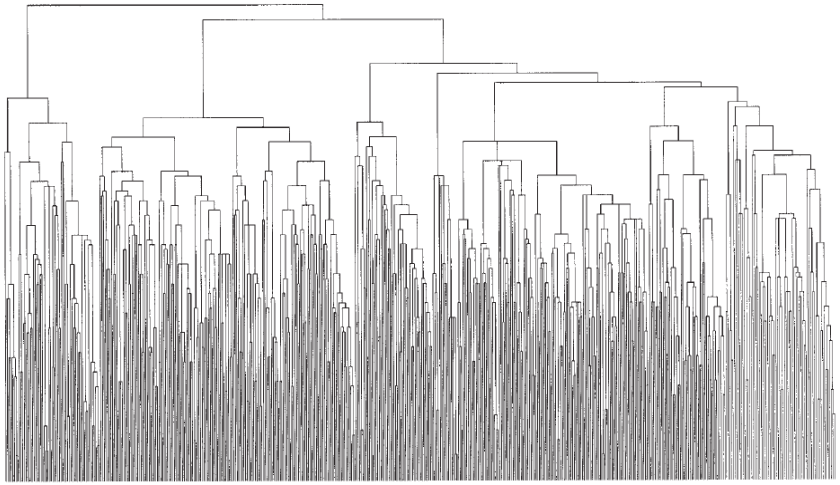


Fig. 2 Hierarchical tree representation of part of the energy landscape of crambin. Figure kindly provided by A.E. Garcia.

npj | Microgravity

www.nature.com/npjmgrav

ARTICLE OPEN

Phenotypic transitions enacted by simulated microgravity do not alter coherence in gene transcription profile

Agnese Po[1-3], Alessandro Giuliani[3], Maria Grazia Masiello[1], Alessandra Cucina[1,4], Angela Catizone[1], Giulia Ricci[1], Martina Chiacchiarini[1], Marco Tafani[1], Elisabetta Ferretti[1] and Mariano Bizzarri[1,2*]

Cells in simulated microgravity undergo a reversible morphology switch, causing the appearance of two distinct phenotypes. Despite the dramatic splitting into an adherent-fusiform and a floating-spherical population, when looking at the gene-expression phase space, cell transition ends up in a largely invariant gene transcription profile characterized by only mild modifications in the respective Pearson's correlation coefficients. Functional changes among the different phenotypes emerging in simulated microgravity using random positioning machine are adaptive modifications—as cells promptly recover their native phenotype when placed again into normal gravity—and do not alter the internal gene coherence. However, biophysical constraints are required to drive phenotypic commitment in an appropriate way, compatible with physiological requirements, given that absence of gravity foster cells to oscillate between different attractor states, thus preventing them to acquire a exclusive phenotype. This is a proof-of-concept of the adaptive properties of gene-expression networks supporting very different phenotypes by coordinated 'profile preserving' modifications.

npj Microgravity (2019)5:27; https://doi.org/10.1038/s41526-019-0088-x

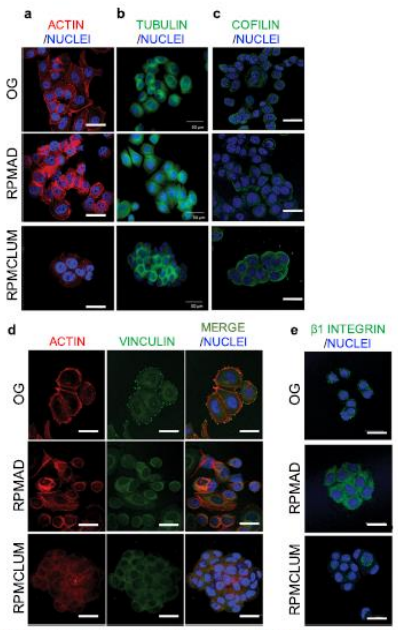
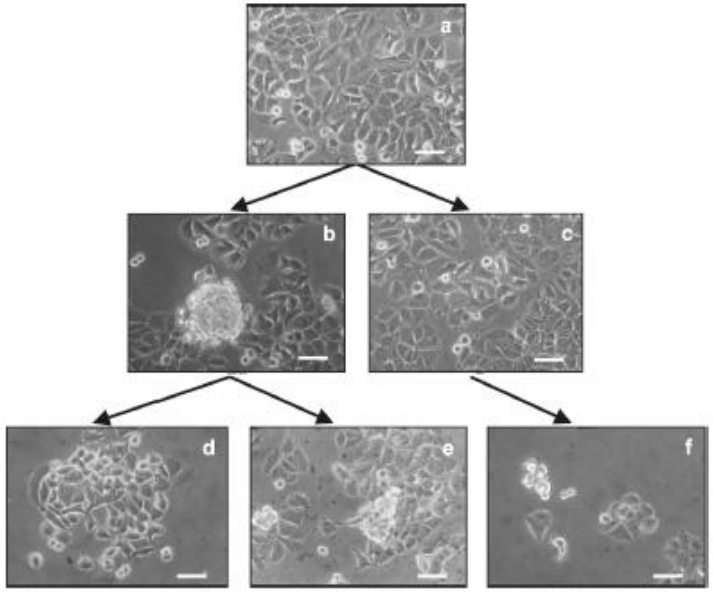


Fig. 3 Cytoskeleton proteins in cells exposed to microgravity. Panels **a**, **b** show F-actin **a** and tubulin **b** respectively in MCF7 cells on ground and in RPM. In OG cells the network of cytosolic F-actin appears well organized in bundles associated with the cell plasma membrane. In RPMAD cells, stress fibers are less evident and F-actin bundles appeared mostly localized at the cell border. In floating cell clumps, the actin meshwork loses its organization and actin filaments looked short, fragmented and spreading over the cytosol. About the tubulin organization, we observed that the microtubule-organizing centre (MTOC) near the nucleus in OG cultured cells, disappears in both RPMAD and RPMCLUM **b**. However, in RPMAD samples microtubules are still identifiable, while in RPMCLUM samples tubulin meshwork was completely disrupted, and tubulin appeared almost completely aggregated around the nucleus without any polarization. Panel **c** shows cofilin distribution. Cofilin was dispersed in the whole cell body with a visible accumulation in the cytosol of OG and RMPAD cells; instead, in RPMCLUM cells, an impressive, dense accumulation of cofilin was observed under the cortical ring of the cellular membrane. Vinculin distribution is reported in panel **d**. Vinculin decreases in RPMAD group, and especially in both the cytosol and the membrane of RPMCLUM cells. Reduction of vinculin is accompanied by a reduced amount of stress fibers, formation of fewer focal adhesions, and inhibition of lamellipodia extension. Integrin distribution is also significantly impaired by microgravity **e**. An intense deposition of β1-integrin at the cell membrane in RPMAD cells is recorded, while in RPMCLUM cells β1-integrin almost completely disappears. Scale bars: 30 μm.



ig. 1 Morphological changes in cells exposed to different gravity conditions. MCF7 cells cultured under static 1g-conditions grew as a ormal 2D monolayer **a**. MCF7 cells in microgravity resulted partitioned into two phenotypes. The first, represented by floating-clump RPMCLUM) cells, and the second constituted by adherent cells (RPMAD) (**b**, **c**). Both phenotypes revert to the native morphology when they re reseeded in normal gravity, independently from the time they have spent in microgravity. In panel **d** it is shown how cells growing in nicrogravity for 24 hours recover their native phenotype when replaced in normal gravity for 6 h. When the two cell clusters previously btained during a first-course culture in weightlessness are isolated, and then again reseeded in the same microgravity field, two distinct henotypes emerge once more from each cell phenotype (**e**, **f**). Scale bar: 50 μm

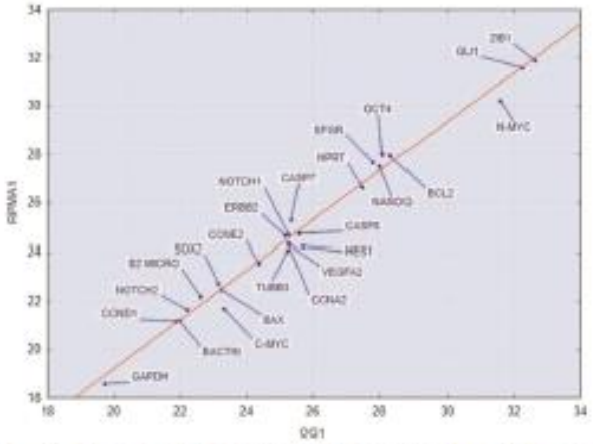


Fig. 4 Pairwise correlation between gene-expression profile in cells growing at OG and RPMAD at 1 h. The correlation coefficients are near to unity, pointing to a very strong invariance of gene-expression profiles despite the dramatic phenotypic changes. Vector points correspond to gene-expression values, and the axes refer to the different conditions (microgravity exposure and on ground, after 1 h of conditioning)

Table 2. Pearson correlation coefficients between gene-expression profiles

	OG1	OG2	OG6	OG24	RPMAD1	RPMAD2	RPMAD6	RPMAD24	RPMCL1	RPMCL2	RPMCL6	RPMCL24
OG1	1	0.99545	0.98874	0.98498	0.99269	0.98903	0.98821	0.95613	0.98509	0.97551	0.98656	0.99303
		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
OG2		1	0.99440	0.98840	0.99121	0.99255	0.99491	0.97022	0.98445	0.97318	0.98954	0.99486
			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
OG6			1	0.98094	0.98950	0.98593	0.99585	0.95731	0.98667	0.97756	0.99301	0.99213
				<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
OG24				1	0.98189	0.98619	0.97960	0.95514	0.96898	0.95926	0.97239	0.98260
					<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RPMAD1					1	0.98925	0.98623	0.95151	0.99304	0.98407	0.98698	0.99201
						<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RPMAD2						1	0.98654	0.97274	0.97285	0.95789	0.97839	0.98896
							<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RPMAD6							1	0.96505	0.98514	0.97574	0.99355	0.99097
								<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RPMAD24								1	0.93787	0.92226	0.95578	0.96736
									<0.0001	<0.0001	<0.0001	<0.0001
RPMCL1									1	0.99544	0.98875	0.98426
										<0.0001	<0.0001	<0.0001
RPMCL2										1	0.98121	0.97361
											<0.0001	<0.0001
RPMCL6											1	0.99361
												<0.0001
RPMCL24												1

The pairwise between profiles Pearson correlation coefficients are reported together with their statistical significance values

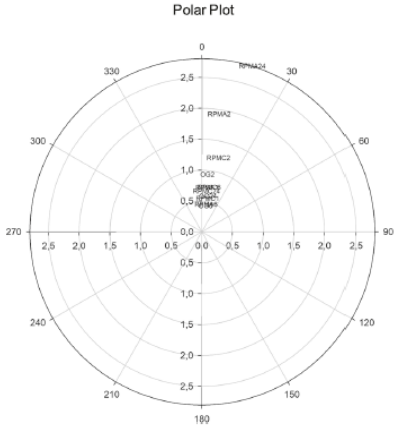
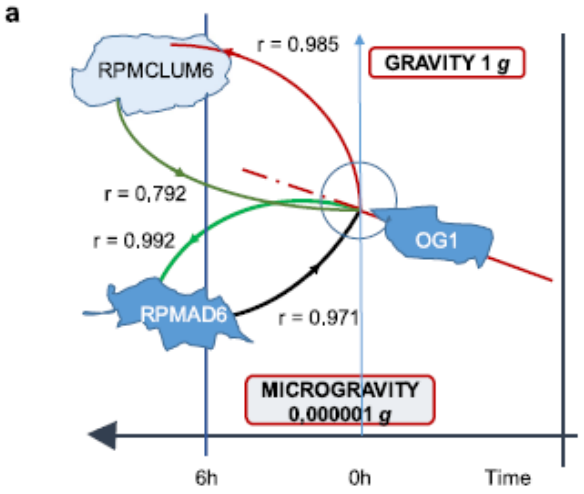
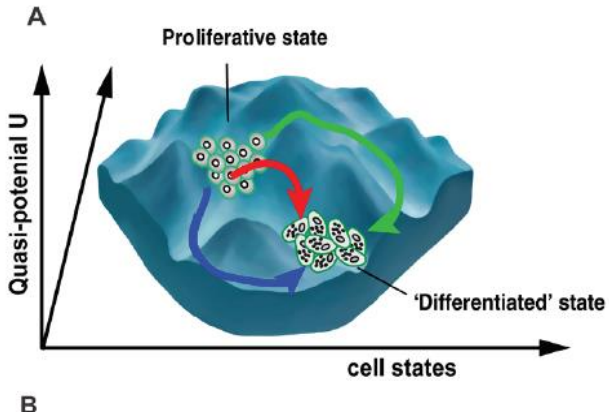


Fig. 6 Euclidean and Angular distances in gene-expression patterns. The plot reports the Euclidean distance (see also Supplementary Fig. 3) from the centre computed over the gene-expression values (concentric circles) and the angle of deviation (the angle between OG1 and different profiles having Pearson r with OG1 as cosine) from baseline (OG1) condition. Data are computed at different times for both RMPCLUM and RPMAD (1, 2, 6, 24 h). The overlapping or partially overlapping experimental points are the following: OG6, OG24, RPMA1, RPMA6, RPMC1, RPMC6, RPMC24



Systematic drug perturbations on cancer cells reveal diverse exit paths from proliferative state

Joseph X. Zhou[1,3,*], Zerrin Isik[2,4,*], Caide Xiao[3], Irit Rubin[1], Stuart A. Kauffman[1,3], Michael Schroeder[4] and Sui Huang[1,3]



We found that MCF7 cells exit proliferation states in several distinctive trajectories after being stimulated by different drugs. Among the genes which significantly changed expression levels, about 10% of them diverged at first and converge later (Figure 2D). It means that cells are destabilized by drug stress, then move to different directions and may fall into the same cell state which are defined by the gene-gene interactions from gene regulatory network. The destabilization mechanism also explains why many nonspecific drugs induced MCF7 cells differentiation in low efficiency (Figure 1F). These drugs destabilize the cancer cell state but lack of stimulus to guide the cells to differentiated state. The discovery implied a new direction of cancer drug development: rather than identifying one drug which cause cell transition in a well-defined pathway, we can use multiple drugs to destabilize the proliferation state of cancer and induce cells to exit in various ways.

PLOS ONE

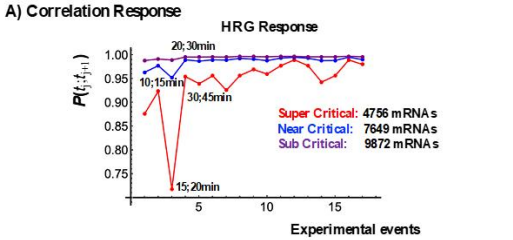
RESEARCH ARTICLE

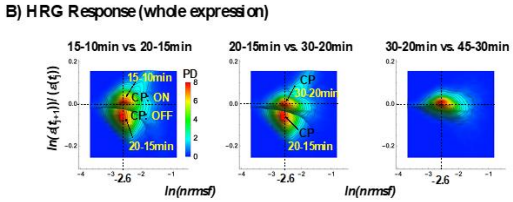
Emergent Self-Organized Criticality in Gene Expression Dynamics: Temporal Development of Global Phase Transition Revealed in a Cancer Cell Line

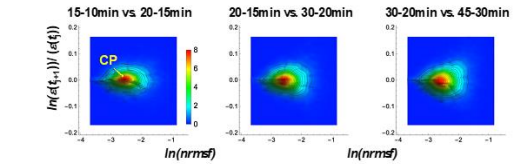
Masa Tsuchiya[1]*, Alessandro Giuliani[2]*, Midori Hashimoto[3], Jekaterina Erenpreisa[4], Kenichi Yoshikawa[5]*

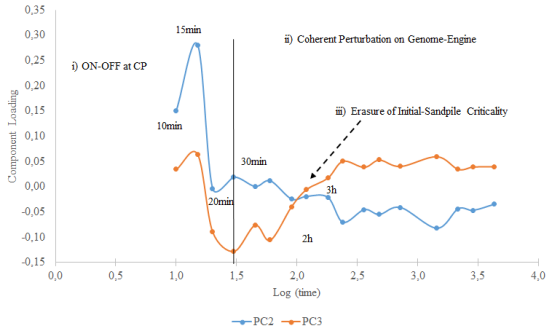
1 Systems Biology Program, School of Media and Governance, Keio University, Fujisawa, Japan, **2** Environment and Health Department, Istituto Superiore di Sanità, Rome, Italy, **3** Graduate School of Frontier Science, The University of Tokyo, Kashiwa, Japan, **4** Latvian Biomedical Research & Study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

* tsuchiya.masa@gmail.com (MT); alessandro.giuliani@iss.it (AG); keyoshik@mail.doshisha.ac.jp (KY)







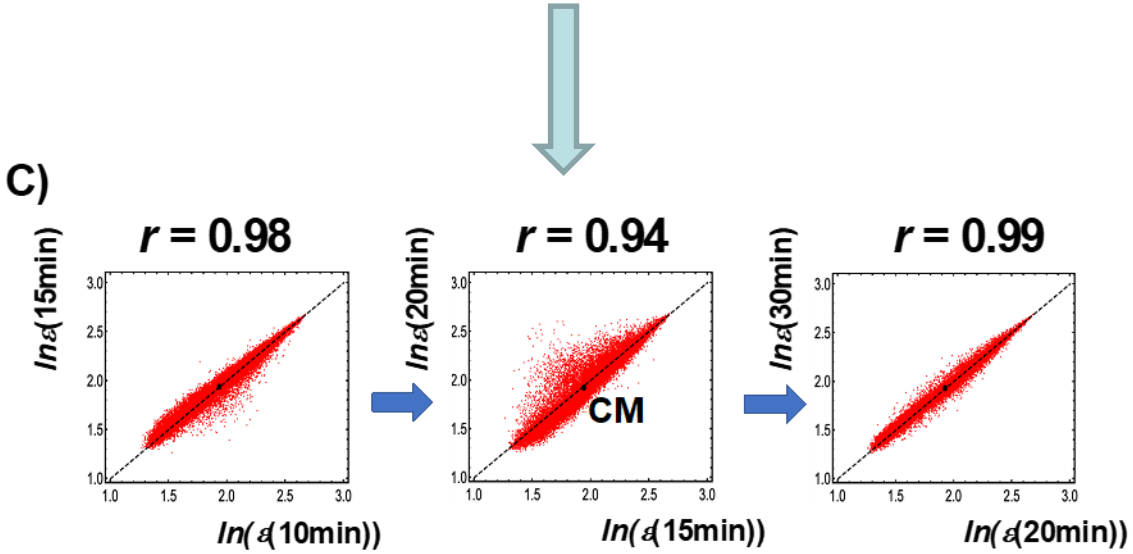


Component	Eigenvalue	Explained Variance (%)	Cumulative variance (%)
PC1	**17.63**	**97.95**	**97.95**
PC2	**0.129**	**0.72**	**98.67**
PC3	**0.065**	**0.36**	**99.04**
PC4	0.040	0.22	99.26
PC5	0.031	0.17	99.43

Table 2b Loading Pattern (HRG)

LABEL	TIME(min)	PC1	PC2	PC3
T0	0	0,984	0,046	-0,023
T10	10	0,983	0,150	0,034
T15	15	0,956	0,280	0,064
T20	20	0,992	-0,004	-0,089
T30	30	0,989	0,018	-0,128
T45	45	0,993	0,000	-0,076
T60	60	0,991	0,011	-0,106
T90	90	0,994	-0,025	-0,040
T2H	120	0,995	-0,020	-0,006
T3H	180	0,994	-0,022	0,017
T4H	240	0,992	-0,071	0,050
T6H	360	0,995	-0,046	0,038
T8H	480	0,994	-0,055	0,053
T12H	720	0,994	-0,042	0,040
T24H	1440	0,990	-0,083	0,060
T36H	2160	0,994	-0,044	0,034
T48H	2880	0,994	-0,047	0,039
T72H	4320	0,990	-0,035	0,040

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…



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.

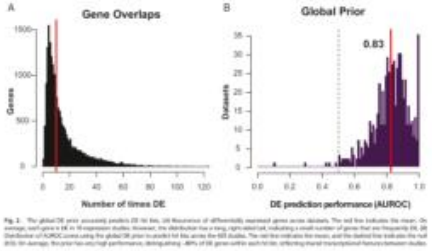
But we need something more:

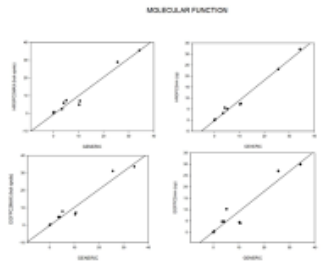
Predictability of human differential gene expression

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

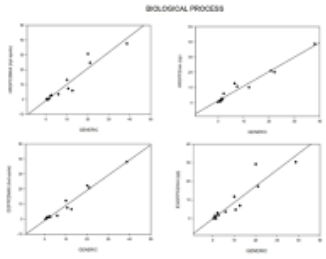
PNAS | March 26, 2019 | vol. 116 | no. 13 | 6491–6500

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.

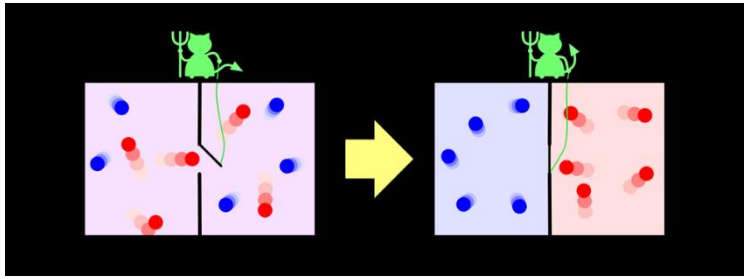




MOLECULAR FUNCTION	HRGPC1MAX	HRGPC1min	EGFPC1MAX	EGFPC1min	GENERIC
transporter	6,9	6,6	4,6	4,7	4,2
translation regulator	0,7	0,5	0,4	0,4	0,5
cargo receptor	0,3	0	0	0	0
transcription regulator	7,3	5	7,9	10,2	5,2
catalytic	28,9	23,1	31,1	27	25,6
molecular function regulator	7	7,5	7,1	4	10,6
molecular transducer	4,9	7,2	6,2	4,4	10,2
structural	2,4	3,1	4,8	4,7	3,5
binding	36,5	32,2	33,6	29,9	34,2



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.



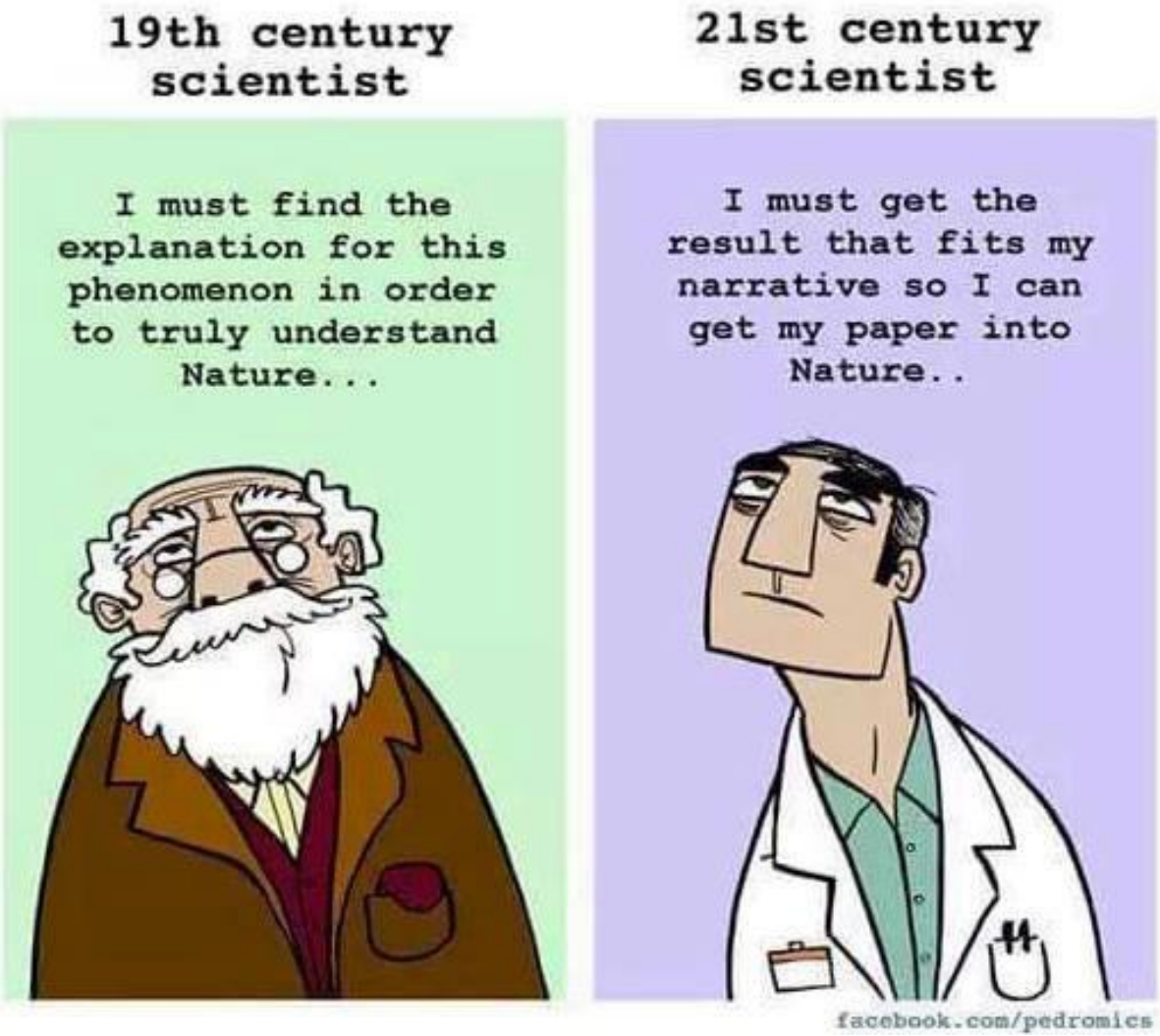
This mechanism implies the presence of an 'intelligent agent' (demon) that alters the natural fate of the system…the issue is much more serious than a scientific joke: almost totalityof biological explanations follow a 'Maxwell's demon' style.

Statistical Science
2010, Vol. 25, No. 3, 289–310
DOI: 10.1214/10-STS330

To Explain or to Predict?

Galit Shmueli

In terms of the data collection instrument, whereas in explanatory modeling the goal is to obtain a reliable and valid instrument such that the data obtained represent the underlying construct adequately (e.g., item response theory in psychometrics), for predictive purposes it is more important to focus on the measurement quality and its meaning in terms of the variable to be predicted.


19th century scientist
I must find the explanation for this phenomenon in order to truly understand Nature...
21st century scientist
I must get the result that fits my narrative so I can get my paper into Nature..
facebook.com/pedromics