Mathematical Models of RNA Expression Profiles: Potential Applications to Drug Discovery Research and Personalized Medicine

Sungchul Ji, Ph.D.
Department of Pharmacology and Toxicology
Ernest Mario School of Pharmacy
Rutgers University
Piscataway, NJ

Paracelsus (1493-1541)

3 Main Approaches to Drug Discovery

Prigogine (1917-1541)

1. *I. Prigogine (1917-2003) divides structures into two classes – Equilibrium (e.g.,* chair, DNA sequences) and Dissipative Structures (e.g., flame of a candle, concentration gradients). *"Life is dissipative structure."*
2. Paracelsus (1493-1541): *"The dose makes the poison."*
3. The Paracelus–Prigogine Principle of Medical Science: *"Dissipative structures make medicines or poisons."*

Approach	Drug Target	
	Equilibrium Structures *(e.g., receptors)*	Dissipative Structures *(e.g., action potentials)*
Top-Down (e.g., herbal medicine)	-	+
Bottom-Up (e.g., molecular pharmacology, receptor pharmacology)	+	-
Hybrid (or Complementary) (e.g., '*ribonoscopic theragnotics*' [S. Ji, conformon.net])	+	+

Sample Preparation

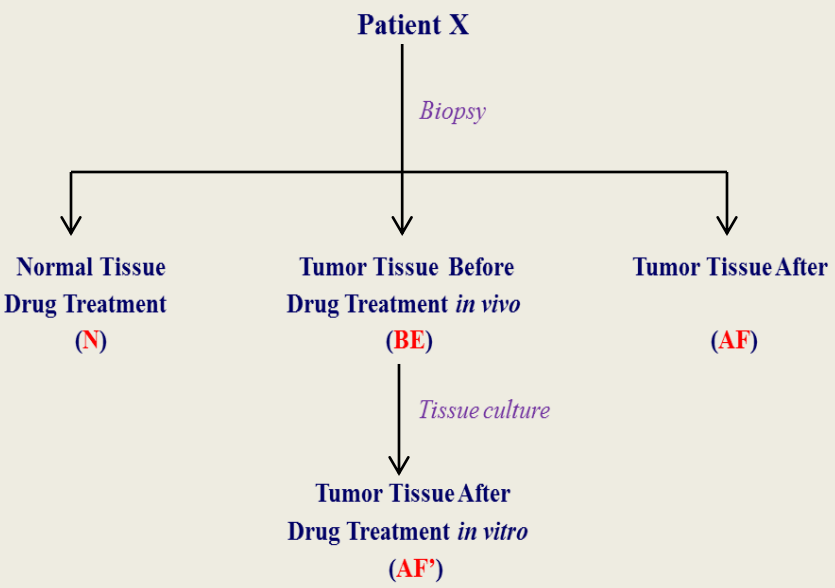
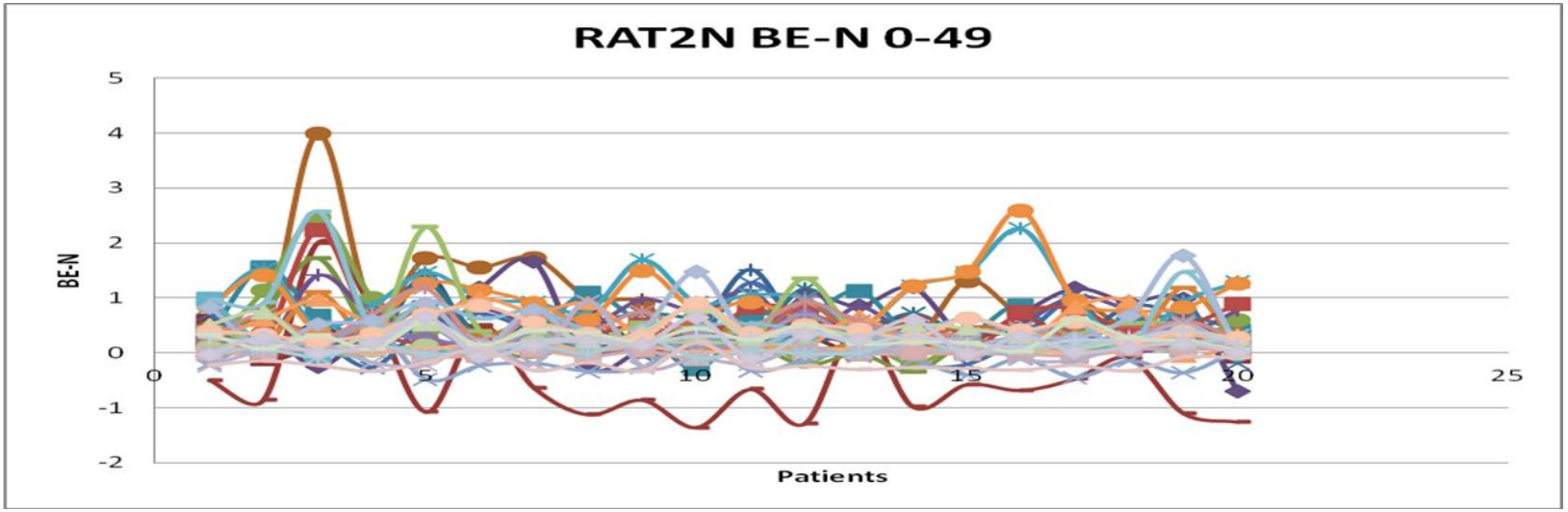
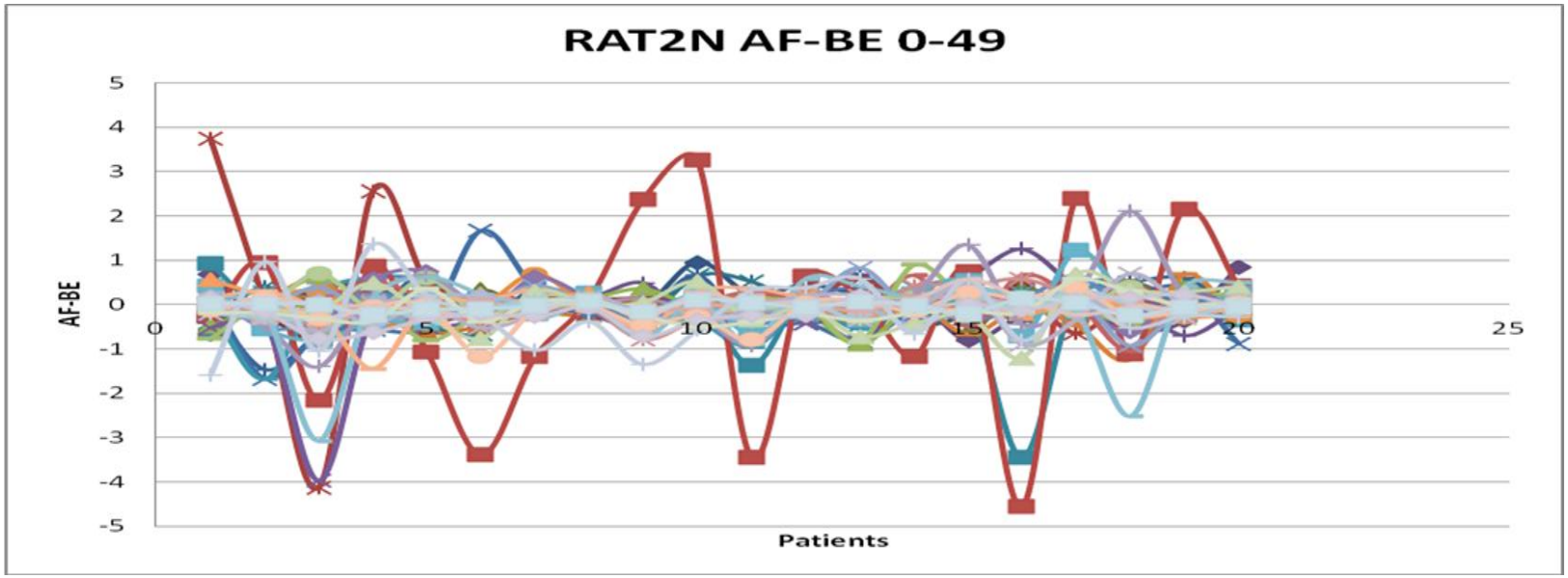


Figure 1. The four types of tissues or cells that are required to generate the molecular data, e.g., RNA sequences and differential expression patterns measured with microarrays or equivalent next-generation sequencing techniques (collectively called ribonoscopy). N, BE and AF are needed to generate the molecular data (in the form of the *mechanism tables* described in **[8]** and **[9]**) for *theragnostics*, while N, BE, and AF' are needed to generate the molecular data for *personalized therapy* which is not discussed in this poster. For the sake of simplicity, the symbol A F may be used to indicate either AF or AF', whenever no confusion can arise under the context of a given discussion. This poster will analyze the microarray data measured by Perou et al. [3] from i) normal breast tissues (N), ii) tumor before (BE) treating with doxorubicin, and iii) tumor after (AF) the drug therapy. The tumor samples were obtained from 65 surgical specimens of human breast tumors and microarrays were used to measure the RNA levels encoded by 8,102 genes., of which 4,740 genes and their transcripts have been analyzed in this poster.

The mRNA fold changes in breast tumor tissues of 20 patients before (BE) and after drug treatment (AF)

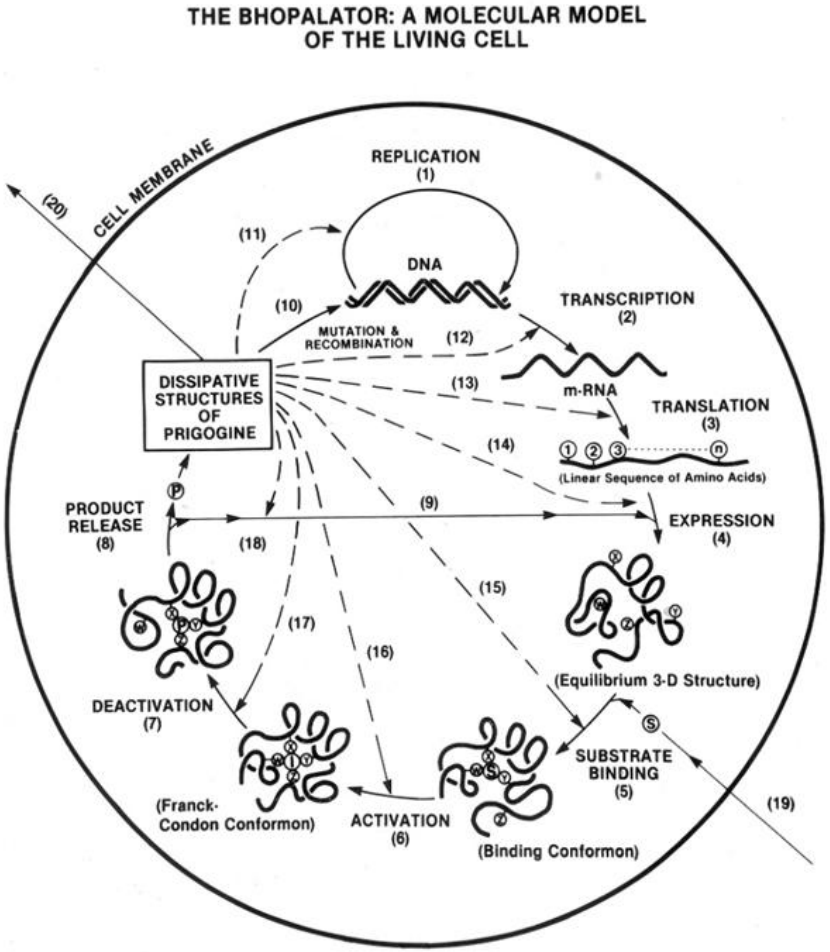




The Theoretical Model of the Living Cell, the *Bhopalator,*

proposed in Bhopal in 1983,

as a Self-Organizing Chemical Reaction-Diffusion System



9 Mechanisms of Responses of Tumor Cells to Anti-Cancer Drugs

ΔT = mRNA changes due to tumor; ΔD = mRNA changes due to drug treatment

α° = arcTan (ΔD/ΔT)

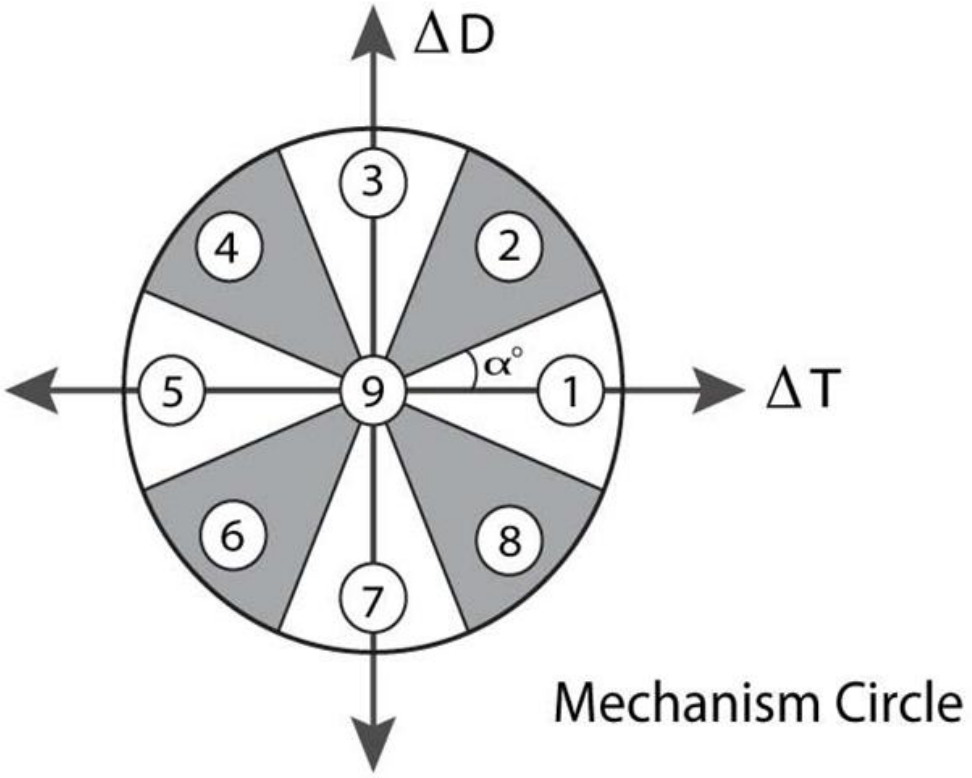


Table 1. The definition of the *mechanism numbers* and their meanings. The symbols are defined thus: + = increase; - = decrease; 0 = no change.

Mechanism Number	Angle (α) from the mechanism circle (°)	Effects on RNA levels due to	
		Tumor (ΔT)	Drug (ΔD)
1	-22.5 ~ 22.5	+	0
2	22.6 ~ 67.5	+	+
3	67.6 ~ 112.5	0	+
4	112.6 ~ 157.5	-	+
5	157.6 ~ 202.5	-	0
6	202.6 ~ 249.5	-	-
7	249.6 ~ 292.5	0	-
8	292.6 ~ -22.5	+	-
9	Defined as the mean +/- 5% of the range of angles excluding those lying outside of the mean +/- 2 σ's.	0	0

Table 2. The"unfiltered mechanism table". N = the number of patients; n = the number of ORFs; SM = survival months; imTI = individual micro-therapeutic index defined by Eq. (3); ITI = individualized therapeutic index (see below). N = normal, BE before drug treatment; AF = after drug treatment; M = mechanism number defined in Table 1 and Figure 3. The numbers in the interior of the table are arbitrary ones selected for an illustrative purpose only.

ORF	Patient 1				Patient 2				. . .	Patient N			
	N	BE	AF	M	N	BE	AF	M	. . .	N	BE	AF	M
1				2				1	. . .				6
2				3				6	. . .				2
3				8				5	. . .				7
4				1				4	. . .				9
5				4				2	. . .				1
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
n				8				6	. . .				9
imTI	4.5				8.5					5.6			
SM	12				80				. . .	23			
ITI	2.6				5.1				. . .	1.2			

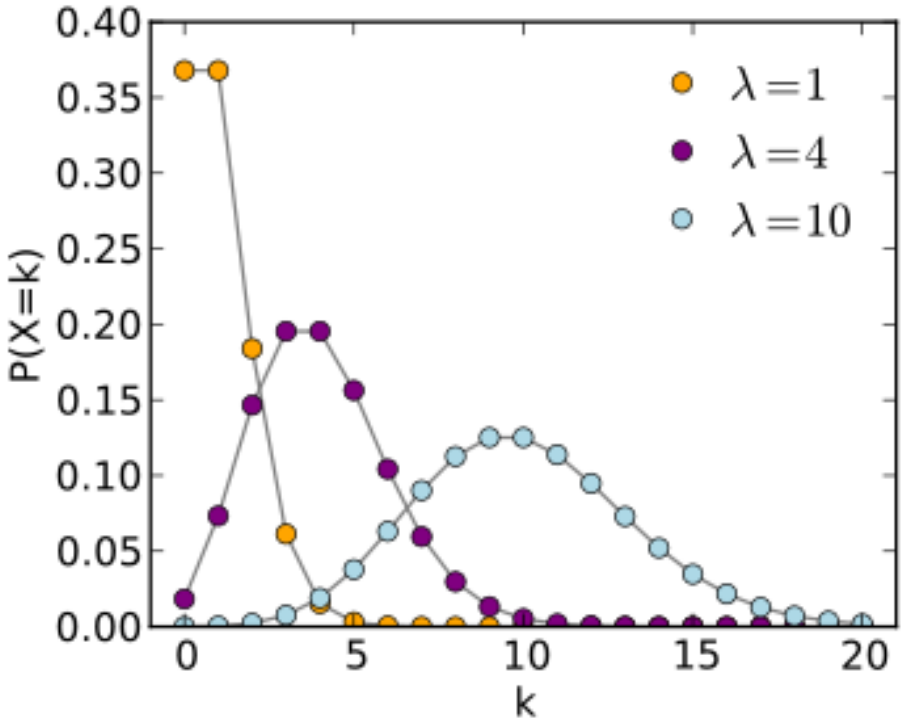
mRNA Level Data Processed to Reveal the Therapeutic Effects of Doxorubicin on 20 Breast Cancer Patients

ORF	Symbol	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
1	ZFX	9	8	2	8	3	2	2	2	8	4	9	5	8	9	1	1	9	8	2	1
2	CDC34	9	9	4	8	9	8	7	1	8	2	8	5	9	8	2	8	9	1	3	9
3	UQCRH	9	8	2	8	7	8	8	9	9	2	1	5	8	8	2	8	3	8	1	1
4	TIMP1	8	9	3	9	8	9	1	9	9	1	9	5	9	1	9	9	8	9	1	9
5	RELA	9	1	9	8	9	1	8	8	8	3	9	5	9	8	3	8	4	2	3	9
6	ACAT2	8	9	4	9	8	8	8	1	8	9	8	5	9	9	9	9	7	9	9	9
7	TRA@	8	2	8	3	3	1	8	8	1	1	1	7	8	8	1	1	1	1	3	1
8	RBM5	7	1	4	8	9	2	7	1	8	4	8	5	6	7	2	7	4	8	3	1
9	SFRS10	8	9	3	8	8	9	9	1	8	9	9	5	9	8	2	9	8	8	9	9
10	RBM3	1	2	2	8	8	2	1	1	8	3	2	3	9	8	2	8	1	1	2	9
11	PXN	9	7	3	7	9	5	4	7	5	4	4	5	8	7	3	9	5	5	5	6
12	TM9SF2	1	9	8	1	1	1	1	1	8	1	8	8	8	1	2	3	8	1	1	9
13	MLF2	8	1	1	8	8	1	8	1	9	1	8	5	8	9	2	9	8	9	1	2
14	ABCC5	8	1	1	9	1	2	1	1	8	2	2	5	1	2	2	1	8	1	2	2
15	DECR1	1	3	4	8	8	1	1	1	8	2	3	3	8	9	2	3	8	8	2	3
16	LOC55977	8	2	1	8	9	9	1	1	8	3	8	5	1	8	9	9	1	8	2	1
17	PPFIA1	1	2	2	8	8	1	1	1	8	3	1	4	8	1	2	8	1	1	2	1
18	ELAVL1	9	9	3	8	8	9	9	1	8	2	8	5	8	8	2	8	9	8	2	9
19	RBM4	8	2	2	8	1	2	1	1	8	9	8	5	8	8	2	8	1	1	1	2
20	FKBP8	9	7	3	5	9	4	5	1	5	3	3	5	7	7	3	7	4	5	3	4
21	TRAF1	8	1	1	8	7	2	1	8	8	2	1	5	8	8	2	8	1	1	2	2
22	DRAP1	8	1	2	8	8	1	1	1	8	1	9	5	8	8	2	8	1	1	1	1
23	ZNF148	1	1	1	8	1	1	1	1	1	8	1	7	1	2	1	1	8	8	1	1
24	TP53BP2	1	8	2	8	2	9	1	8	9	1	8	3	8	9	2	1	9	8	1	2
25	H326	6	4	8	9	9	1	8	4	7	7	7	5	6	5	9	9	5	7	3	4
26	SCAMP3	1	1	1	8	8	1	1	1	8	2	8	1	9	9	1	8	1	1	1	2
27	PDK2	5	5	5	5	4	7	5	9	5	4	5	5	5	5	4	5	7	5	5	4
28	ELF1	9	3	9	8	9	1	2	1	8	2	1	5	8	2	1	3	1	1	9	1
29	DCK	8	2	3	8	8	9	8	1	8	9	1	5	8	1	1	9	8	8	1	8
30	SSR1	8	9	3	8	1	9	1	9	9	8	1	5	9	9	2	9	1	8	1	9
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4740	EIF3S5	2	1	8	7	3	8	8	1	3	1	3	5	3	3	3	4	1	1	3	1

Poisson Distribution

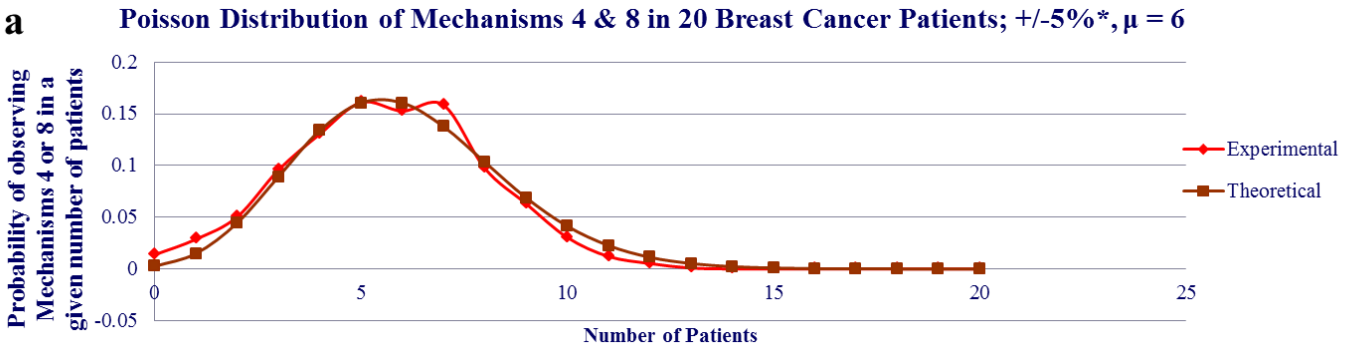
A discrete probability distribution that expresses the probability of a given number of events, k, occurring in a fixed interval of time and/or space if these events occur with a known average rate, μ, and independently of the time since the last event.

$$\mathbf{f(k; \lambda) = ((\lambda^k/k!)e^{-\lambda}}$$

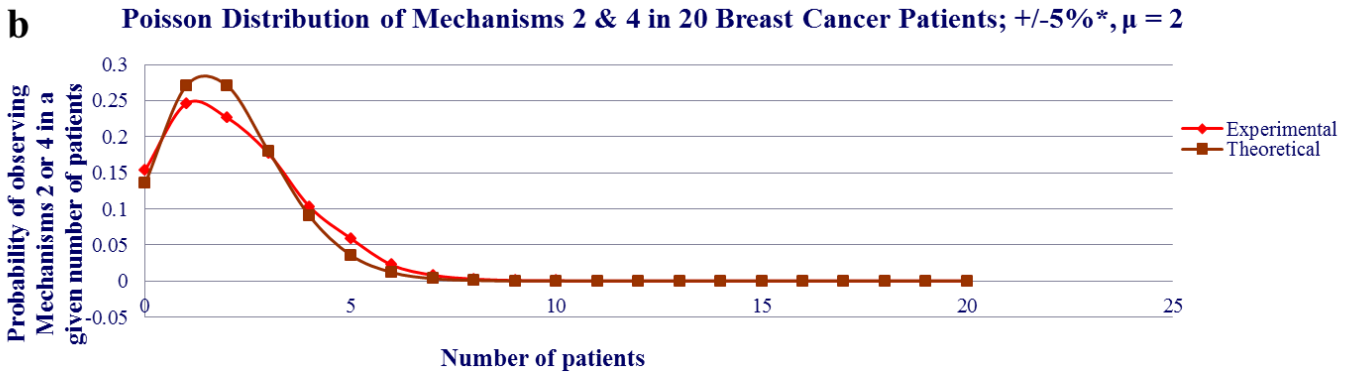


Poisson distributions of beneficial and harmful mechanisms

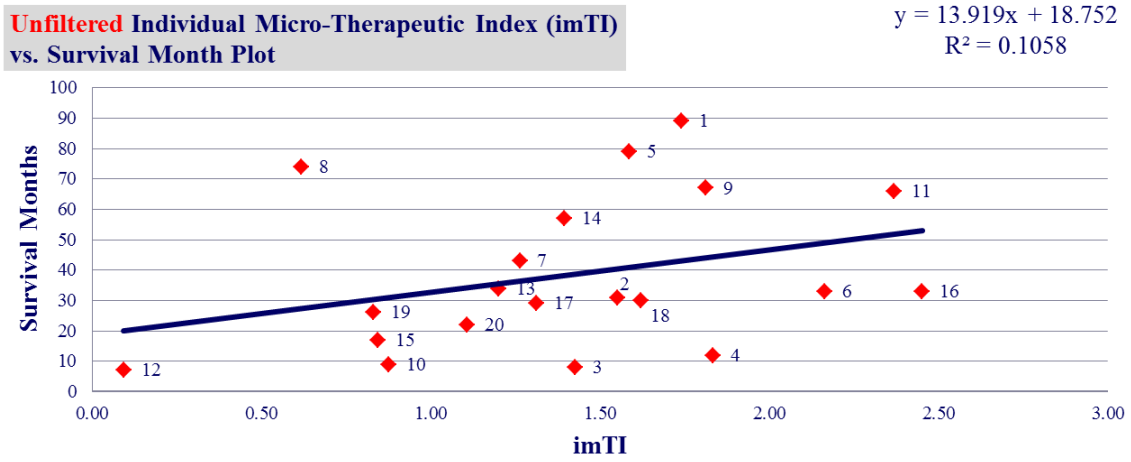
a

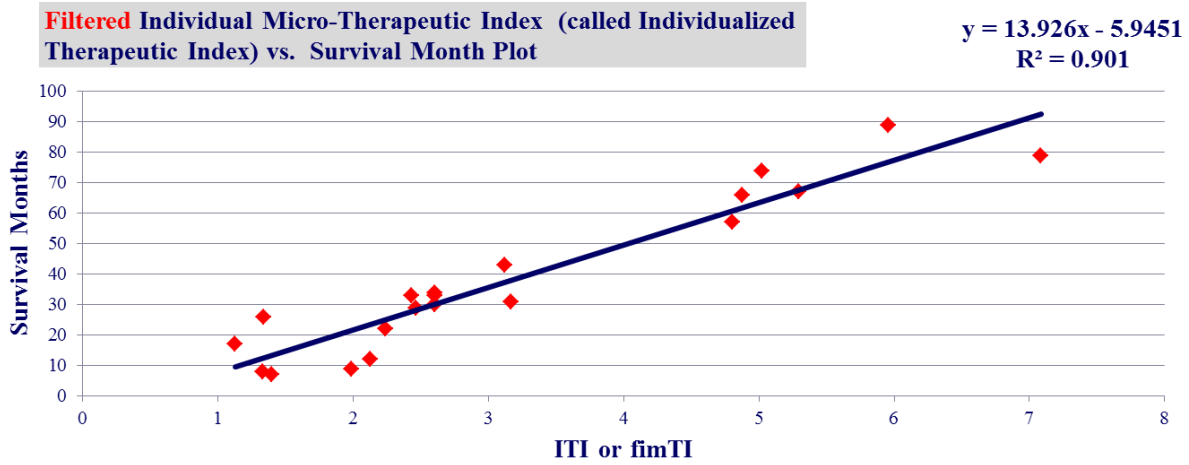


b

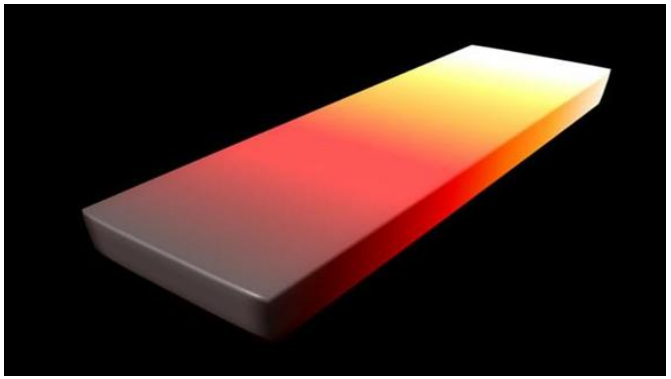


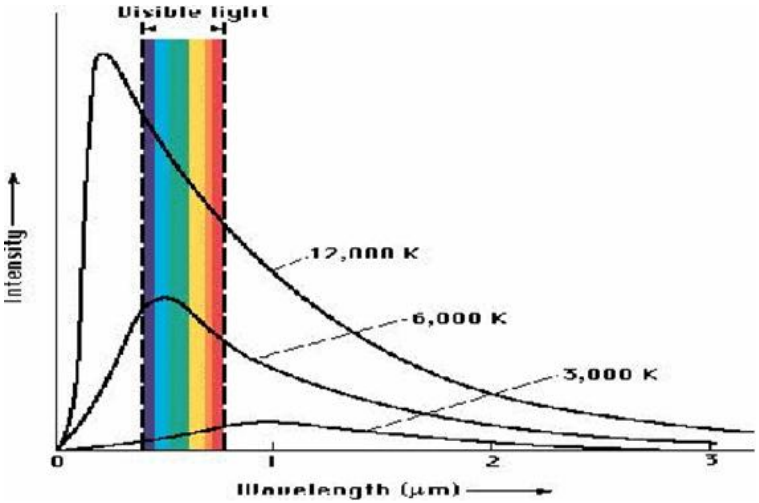
The Micro-Therapeutic Index vs. Survival Month Plot





The Derivation of the Planck Distribution Law





Plank's Equation

$$B_\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\frac{hc}{\lambda kT} - 1}$$

Where:

B_λ = Magnitude of Radiation per Wavelength.

λ = Wavelength.

h = Plank's Conastant ($6.6238 * 10^{-34}$ Js).

c = Speed of Light ($3.0 * 10^{8}$ m/s).

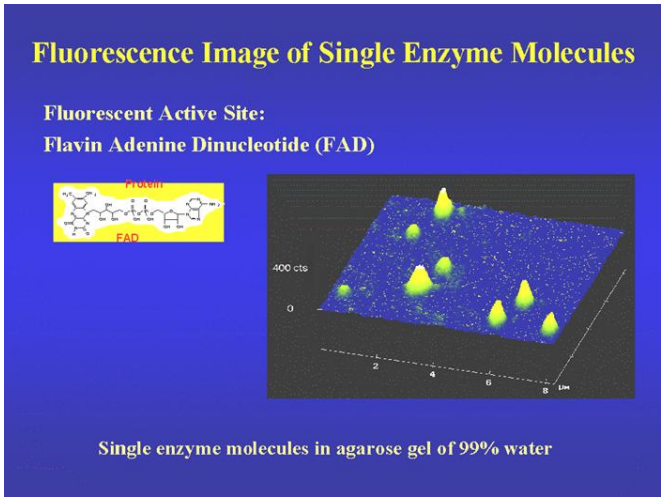
k = Boltzmann Constant ($1.3807 * 10^{-23}$ J/K).

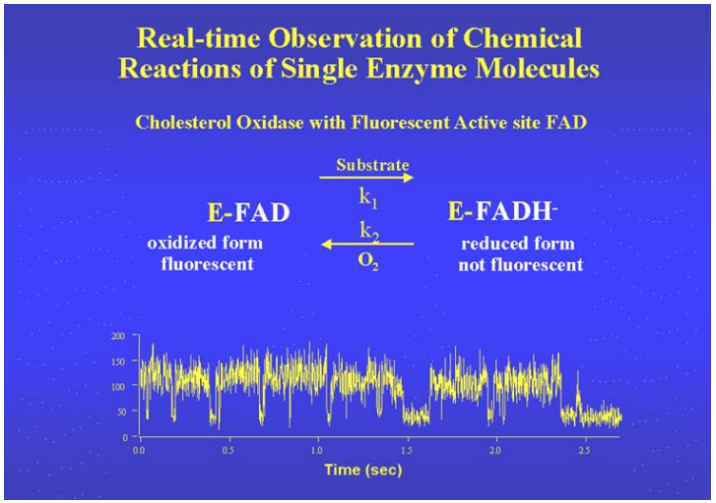
Blackbody radiation equation: $u(\lambda, T) = (8\pi hc/\lambda^5)/(e^{hc/\lambda kT} - 1)$ **(1)**

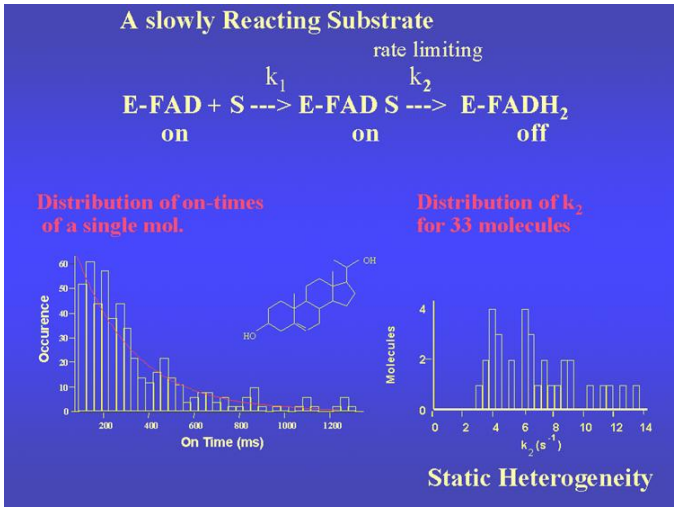
Blackbody radiation-like equation (BRE): $y = (a/x^5)/(e^{b/x} - 1)$ **(2)**

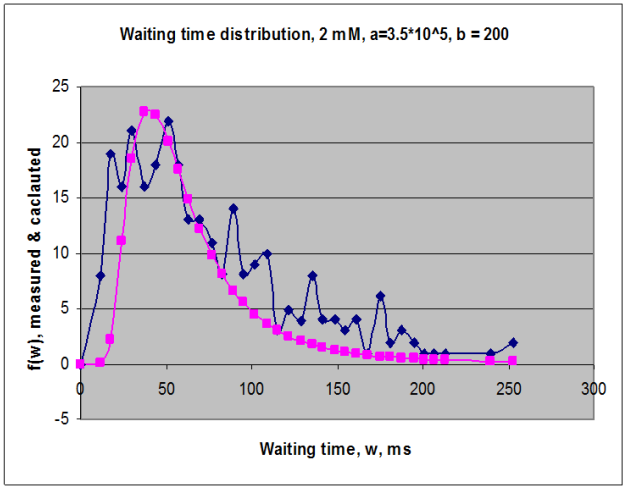
$y = (a(Ax + B)^{-5})/(e^{b/(Ax+B)} - 1)$ **(3)**

Single-Molecule Enzyme Turnover Time Histogram fits the Planck Distribution

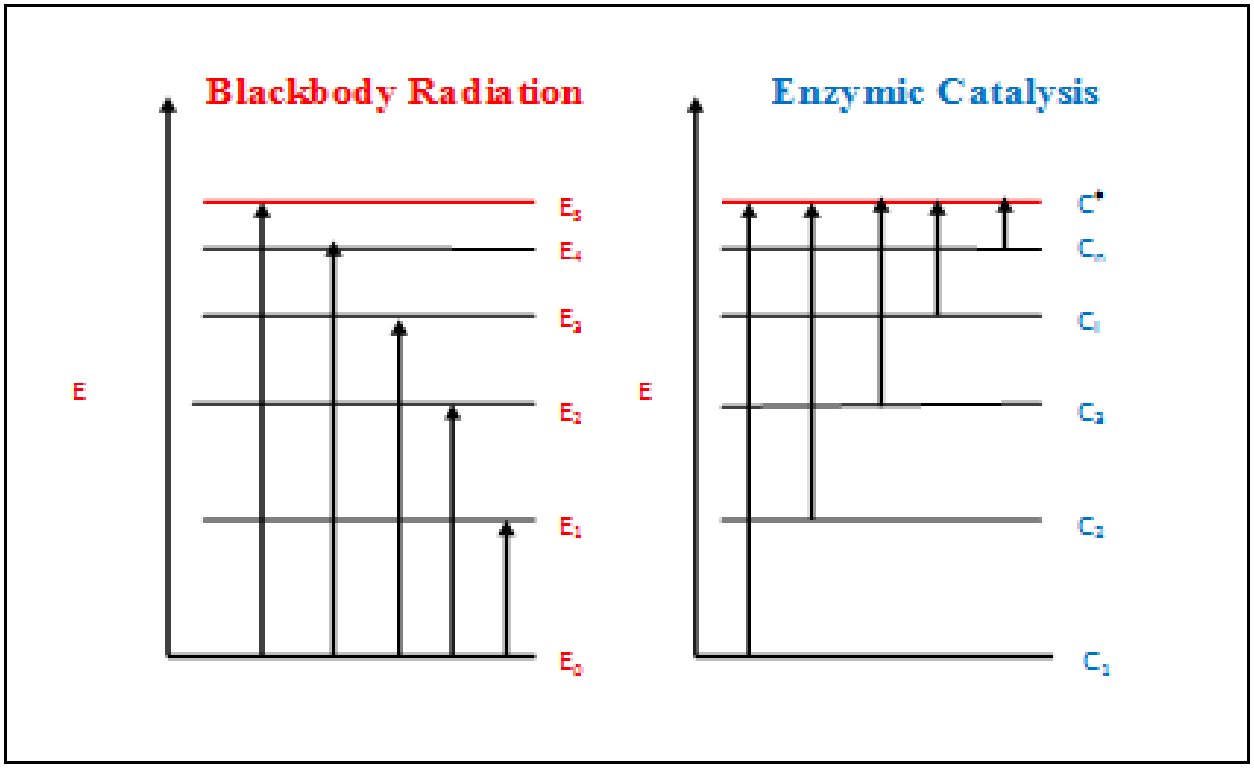






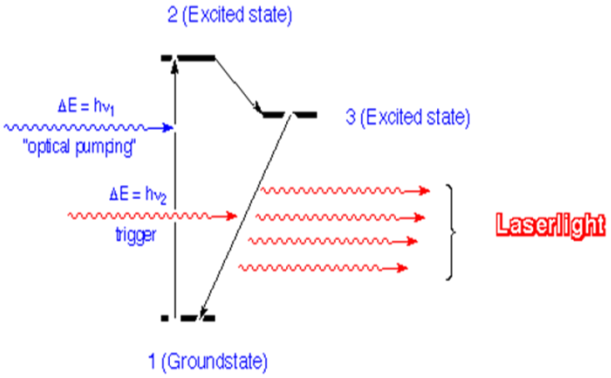


The Quantization of Energy Levels in Atoms and Enzymes



Laser vs. Raser

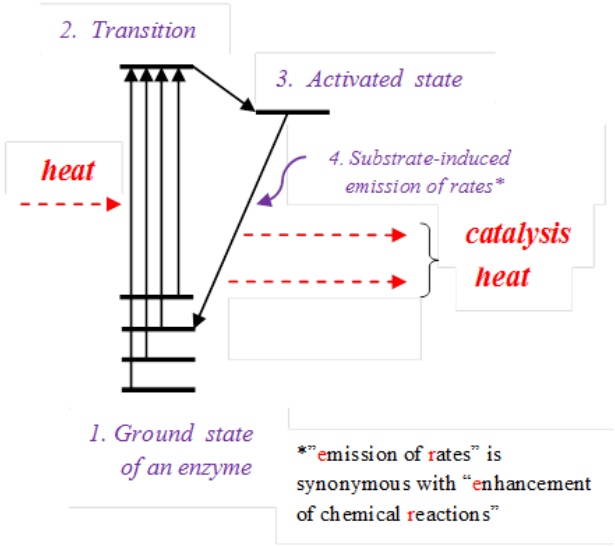
Laser = Light Amplification by Stimulated Emission of Radiation
Raser = Rate Amplification by Substrate-Enhancement of reaction Rates)



1 - > 2: Photochemical, $\Delta E = h\nu_1$

2 - > 3: Thermal (energy transmitted into vibrational modes)

3 - > 1: Stimulated emission gives laserlight



The Planck Distribution as a *Universal Pattern Recognizer*

$$y = (a/(Ax + B)^5)/(e^{b/(Ax + B)} - 1)$$

classifying each pattern in terms of the numerical values of a, b, A and B.

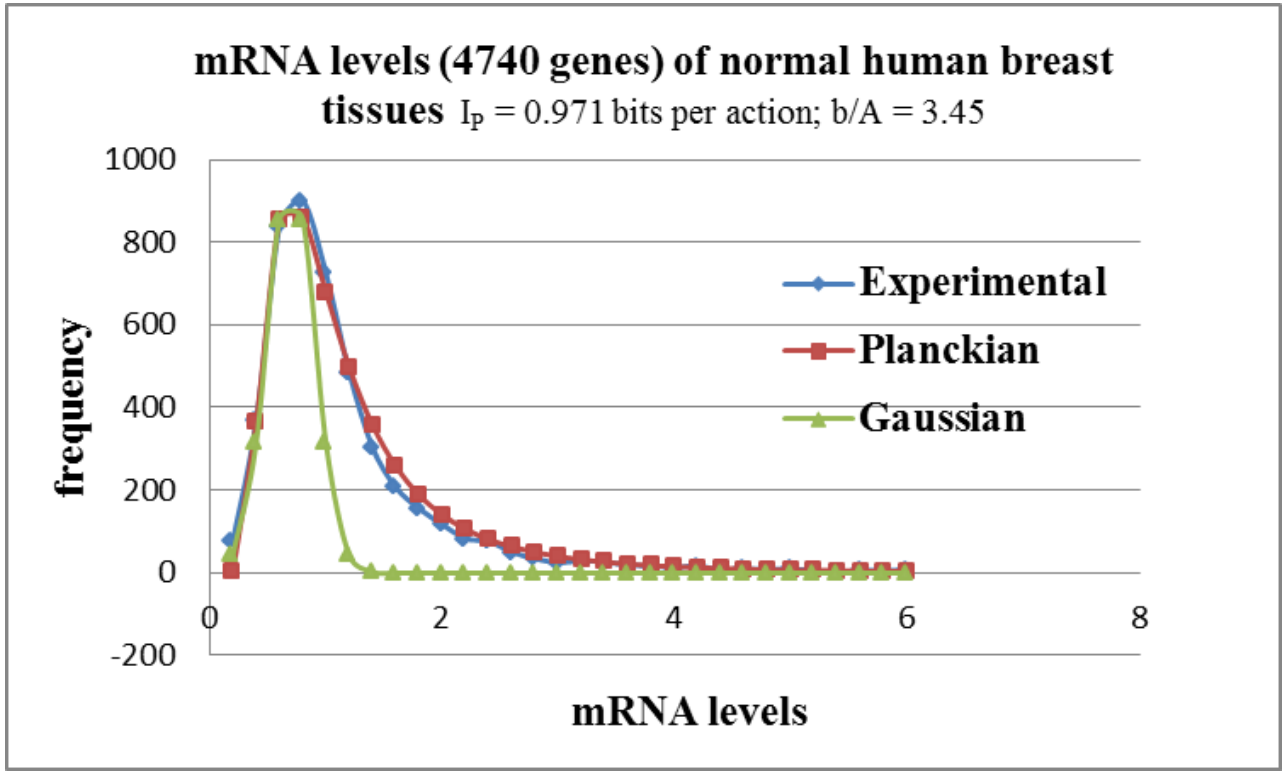


Table 1. Protein families studied and their associated cellular functions.	
Protein Family	**Cellular Function**
Cluster of Differentiation (CD)	Cell Signaling & Adhesion
Kinase-Binding Protein (CGI)	Telomere Uncapping & Elongation
Electron-Transferring Flavoprotein (ETF)	Fatty Acid Oxidation
Heat Shock Protein (HSP)	Stress Response
Interferons (INFR)	Immune System Activation
Integrins (IPA)	Signal Transduction
Unknown Proteins (KIAA)	Unknown
Mitogen-Activated Protein Kinase	Cell Proliferation & Survival
Sterol Carrier Protein (SCP)	Fatty Acid Oxidation
Zinc Finger Protein (ZFP)	DNA Transcription

The *Planck Distribution* as a Classifier of Metabolic Patterns in Tumor Tissues Before and After Drug Treatment

	CGI		MAPK		ZFP		CD		ETF		Whole Genome	
	BE	AF	BE	AF	BE	AF	BE	AF	BE	AF	BE	AF
b/A	3.539	3.427	2.688	2.781	3.273	3.219	6.215	6.571	2.434	2.539	3.188	3.135
	P = 0.14		P = 0.02		P = 0.08		P = 0.007		P = 0.09		P = 0.06	
B/b	-0.0278	-0.0282	-0.1211	-0.0948	-0.0454	-0.0372	0.0499	0.0702	-0.0329	-0.0101	-0.0624	-0.0537
	P = 0.48		P = 0.34		P = 0.15		P = 0.04		P = 0.1		P = 0.23	
B/A	-0.0984	-0.0969	-0.3254	-0.2636	-0.1487	-0.1199	0.3104	0.4609	-0.0803	-0.0258	-0.1869	-0.1603
	P = 0.33		P = 0.09		P = 0.21		P = 0.02		P = 0.03		P = 0.12	

	P-Values (AF, b/A)					
	CGI	MAPK	ZFP	CD	ETF	Whole Genome
CGI	-	1.3E-3	0.04	2.62E-6	7.65E-4	0.02
MAPK	1.3E-3	-	0.01	4.38E-8	0.03	0.01
ZFP	0.04	0.01	-	7.92E-7	6.61E-2	0.04
CD	2.62E-6	4.38E-8	7.92E-7	-	2.54E-9	5.84E-6
ETF	7.65E-4	0.03	6.61E-2	2.54E-9	-	0.03
Whole Genome	0.02	0.01	0.04	5.84E-6	0.03	-

Conclusions

- The microarray technique or its equivalent, when used in combination with mathematical tools such as Poisson and Planckian distribution laws, will enable biomedical scientists to discover anti-cancer drugs without knowing detailed underlying molecular mechanisms.
- The same microarray-based method can be utilized to identify the most efficacious anti-cancer drugs for individual patients.
- There are no genes uniquely responsible for tumorigenesis, hence no single anti-cancer drug applicable to all cancer patients: ***Personalized medicine is inevitable.***