Principles of systems biology and Dmitri Belyaev's co-selection of traits

Hans V. Westerhoff and friends

MCISB, MIB, SCEAS, University of Manchester

SysBA, Universities of Amsterdam

Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

Bioinformatics:
From biological data to information

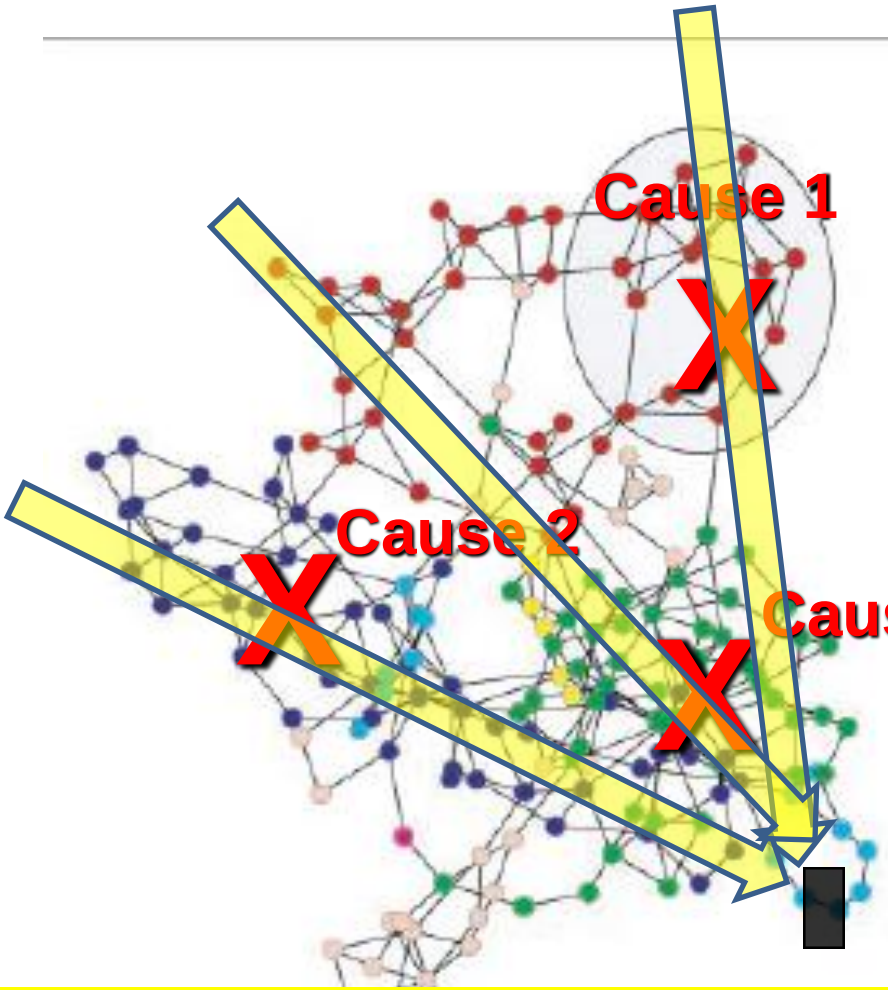
Systems Biology:

From that information to understanding

Systems Biology:
From data to understanding: why is this such an issue?

- **Because the mapping from genome to function is extremely nonlinear**
- E.g.:
  - *P.D. Polly*. Morphometrics and evolution: the challenge of crossing rugged phenotypic landscapes with straight paths
  - The DNA in all our cells is the same, but:
    a heart cell is essentially different from a brain cell
  - Self organization, bistability: Belousov, Zhabotinsky, Waddington, Ilya Prigogine, Boris Kholodenko

Why systems biology?

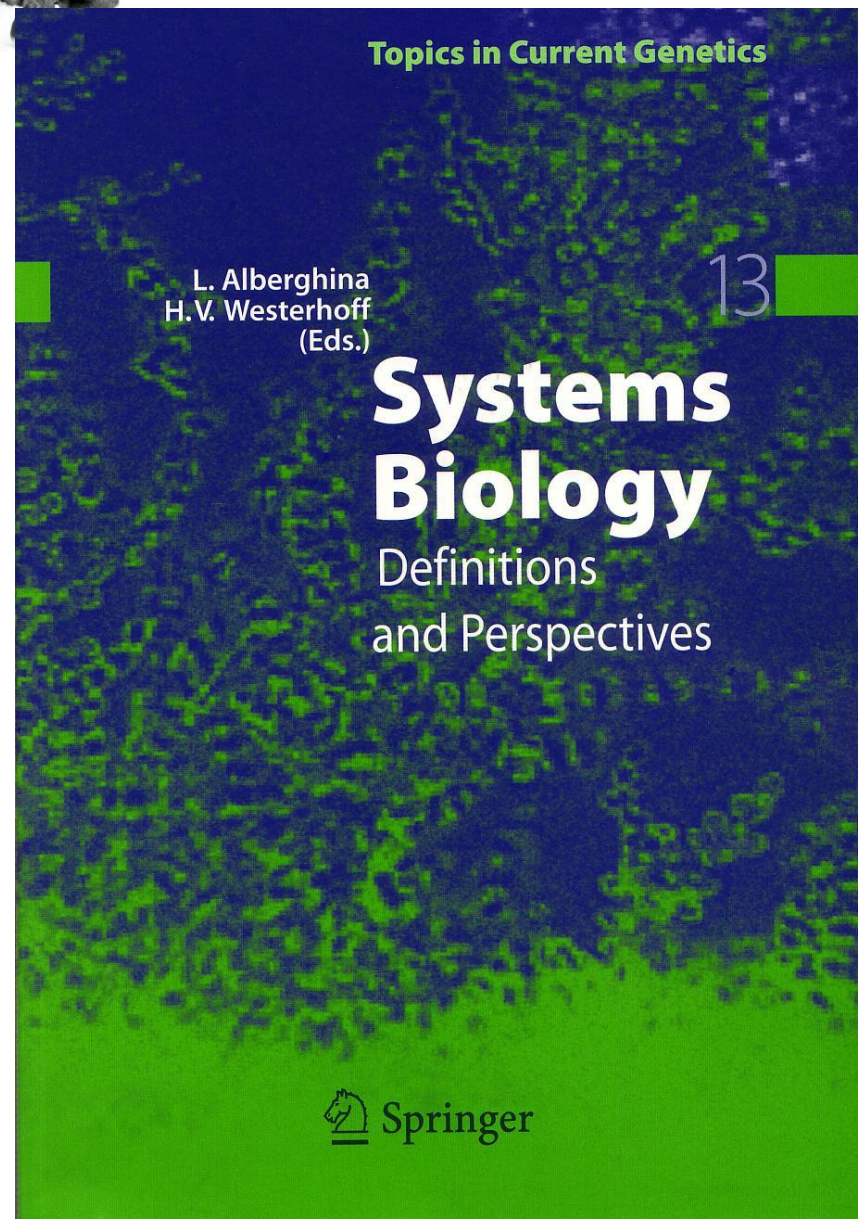


~all functions are network functions

Multiple causality

Multifactorial disease

2006 Hornberg et al: 'Cancer: a systems biology disease'. Now: 'virtually all disease are Systems Biology diseases.' This causes the 'missing heritability problem (Baranov; Stepanov)'



Systems Biology=

- **The Science that**
- **aims to understand**
- **principles governing**
- **how *the* biological functions**
- **arise from *the* interactions = from *the* networking**

This leads to precision, personalized, 4P medicine, PPP4M
And to precision biotechnology

Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

Henrik Kacser (Student of Waddell)

Henrik Kacser

Recessivity of most lack-of-function mutations

Lack of dominance:

No loss of function in heterozygote

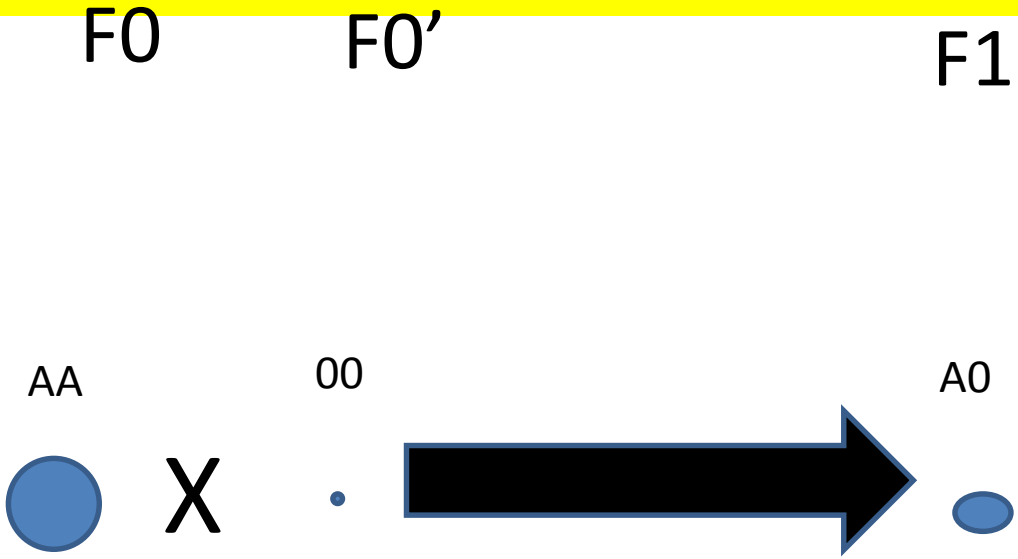
Lack of dominance: observation

F0 F0' (knock out) F1

X 

Function= Flux J **100%** **0%** **95%**

Lack of dominance: single molecule explanation fails



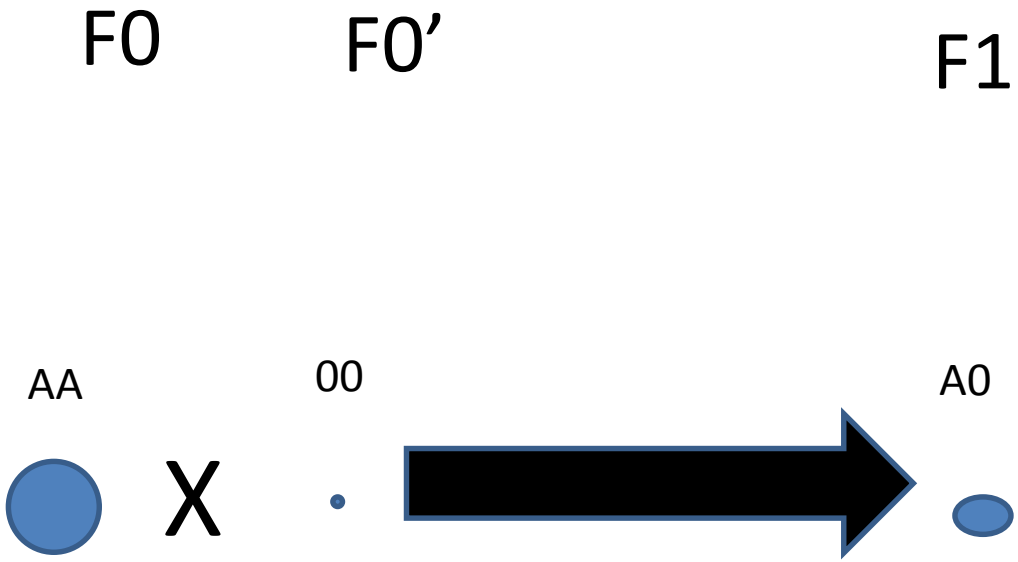
This is almost always incorrect

Function= Flux J

100% **100%** **50%**

Cf. talk by Prof Baranov

Lack of dominance: single molecule explanation fails



Function=
Flux J

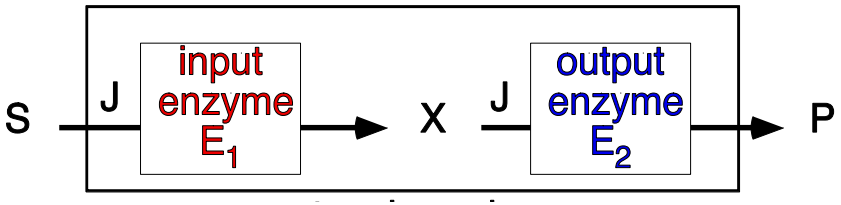
100% **100%**

>90%

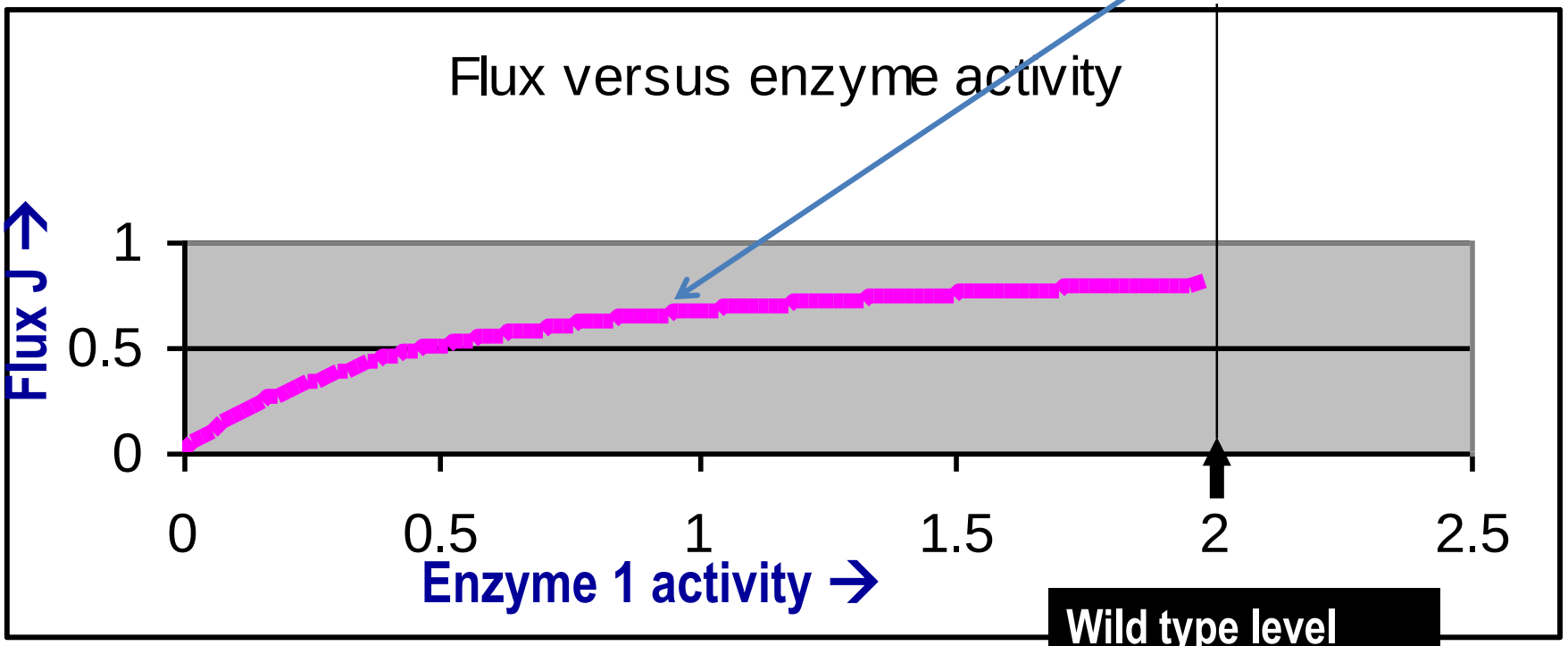
The systems biology explanation

(Henrik Kacser)

Dependence of pathway flux on the concentration of an enzyme

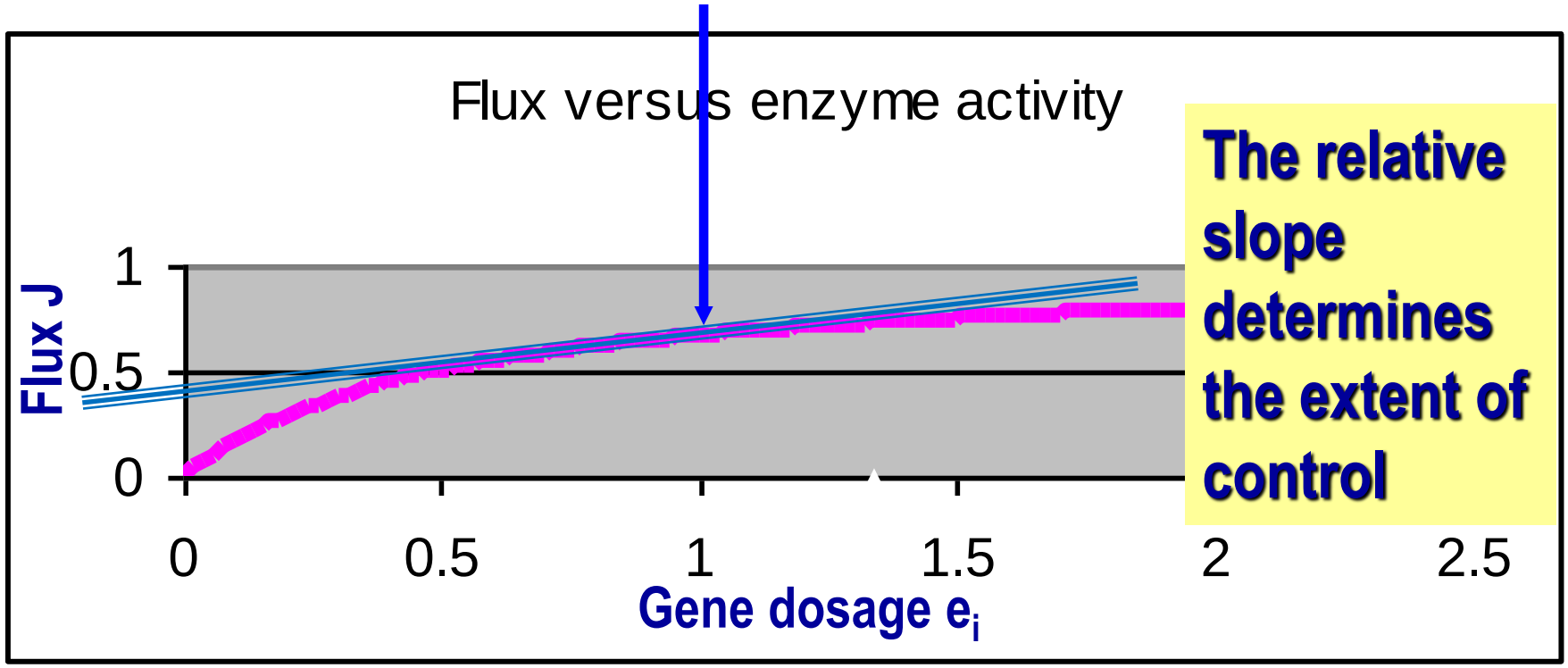


Nonlinear:
Remember the
Polly presentation



How to measure whether a component is limiting: take the relative slope: the flux control coefficient ≈ extent of dominance

Flux Control Coefficient C^J_i = $relative\ slope = \left(\frac{dJ/de_i}{J/e_i} \right)_{steady\ states\ and\ all\ other\ e_j\ constant}$

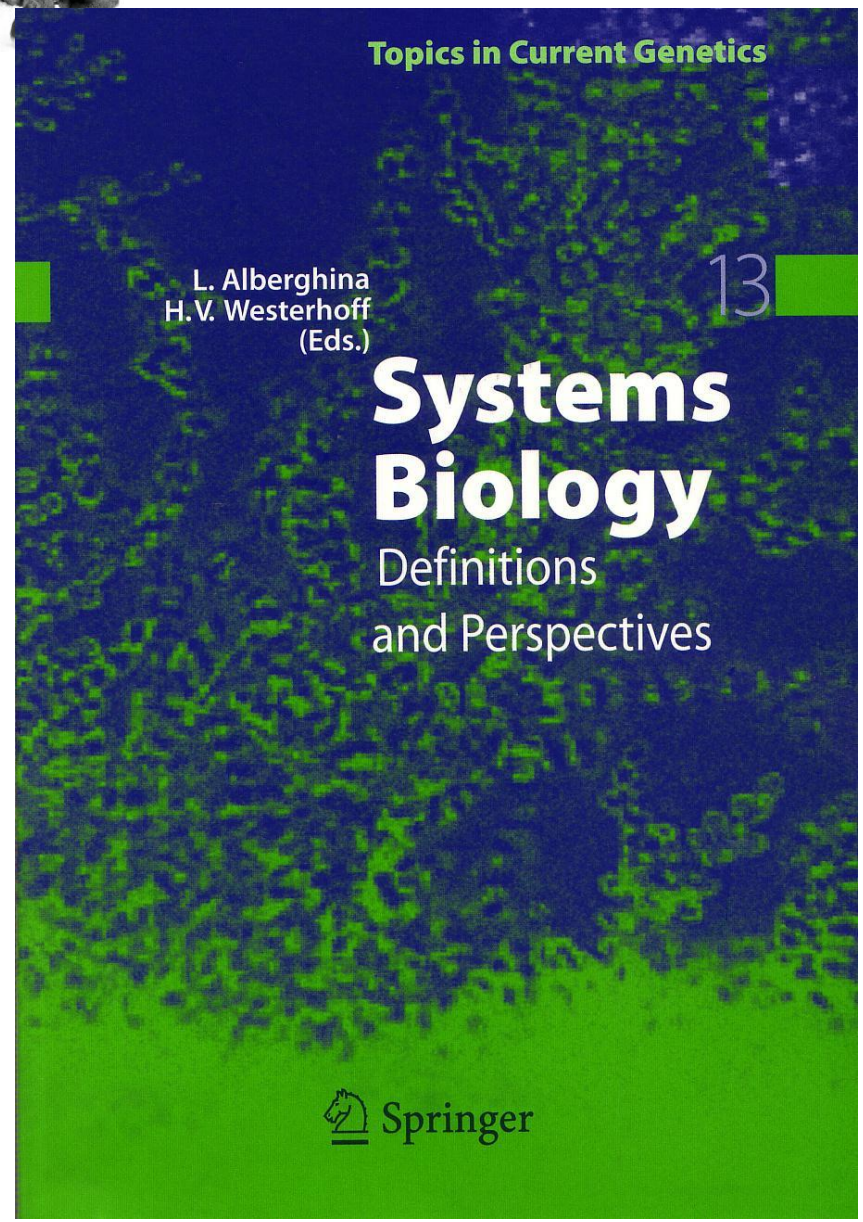


Example

Reaction	4 mM glucose C_i^J	4 mM glucose Γ/K_{eq}	5 mM glucose C_i^J	5 mM glucose Γ/K_{eq}	8 mM glucose C_i^J	8 mM glucose Γ/K_{eq}
Glucose transport					**0.63**	$9.2 \cdot 10^{-3}$
HK					0.04	$<< 10^{-3}$
PFK					0.01	$<< 10^{-3}$
ALD					**0.10**	0.17
GAPDH					**0.09**	0.20
PGK					**0.06**	$3.4 \cdot 10^{-3}$
PYK					0.01	$<< 10^{-3}$
Pyruvate transport					0.00	$<< 10^{-3}$
GDH					**0.06**	$9.1 \cdot 10^{-3}$
GPO					0.01	$<< 10^{-3}$
ATP utilization					0.00	
					1.01	

C is not confined to the single first irreversible step in the pathway

$$C_i^J = \left(\frac{d \ln |J|}{d \ln e_i} \right)_{steadystate} = \frac{dJ/J}{de_i/e_i} = \frac{\%dJ}{\%de_i}$$



Systems Biology=

- The Science
- That aims to understand
- **principles** governing
- how *the* biological functions
- arise from *the* interactions

Flux Control Summation Principle

$$C_1^J + C_2^J + C_3^J + \cdots + C_n^J \equiv 1$$

- J: steady state flux
- 1, 2, 3, : number of the enzyme (gene product)
- **Consequence:**

$$C_i^J \approx \frac{\%reduction\ in\ function}{for\ a\ 50\%\ reduction\ in\ gene\ dosage} = 1/n$$

Flux Control Summation Principle and recessivity

Henrik Kacser
(Student of
Waddell)

$$\overline{C_i^J} \approx \frac{\%reduction\ in\ function}{for\ a\ 50\%\ reduction\ in\ gene\ dosage} \cong 1/10,$$

hence 5 % reduction in function for

a pathway of 10 genes

Lack of dominance: observation

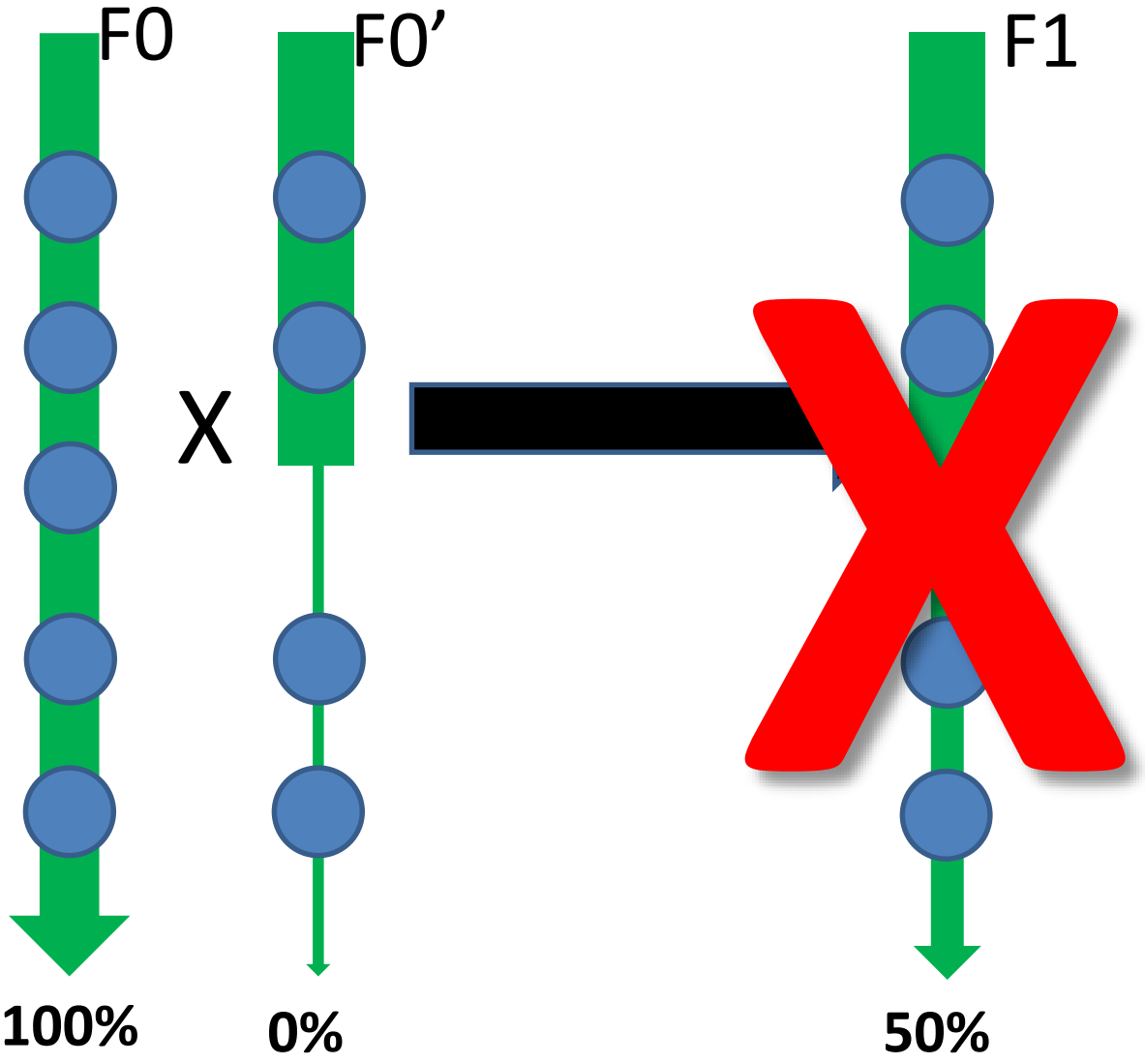
F0 F0’ F1

X

Function= Flux J

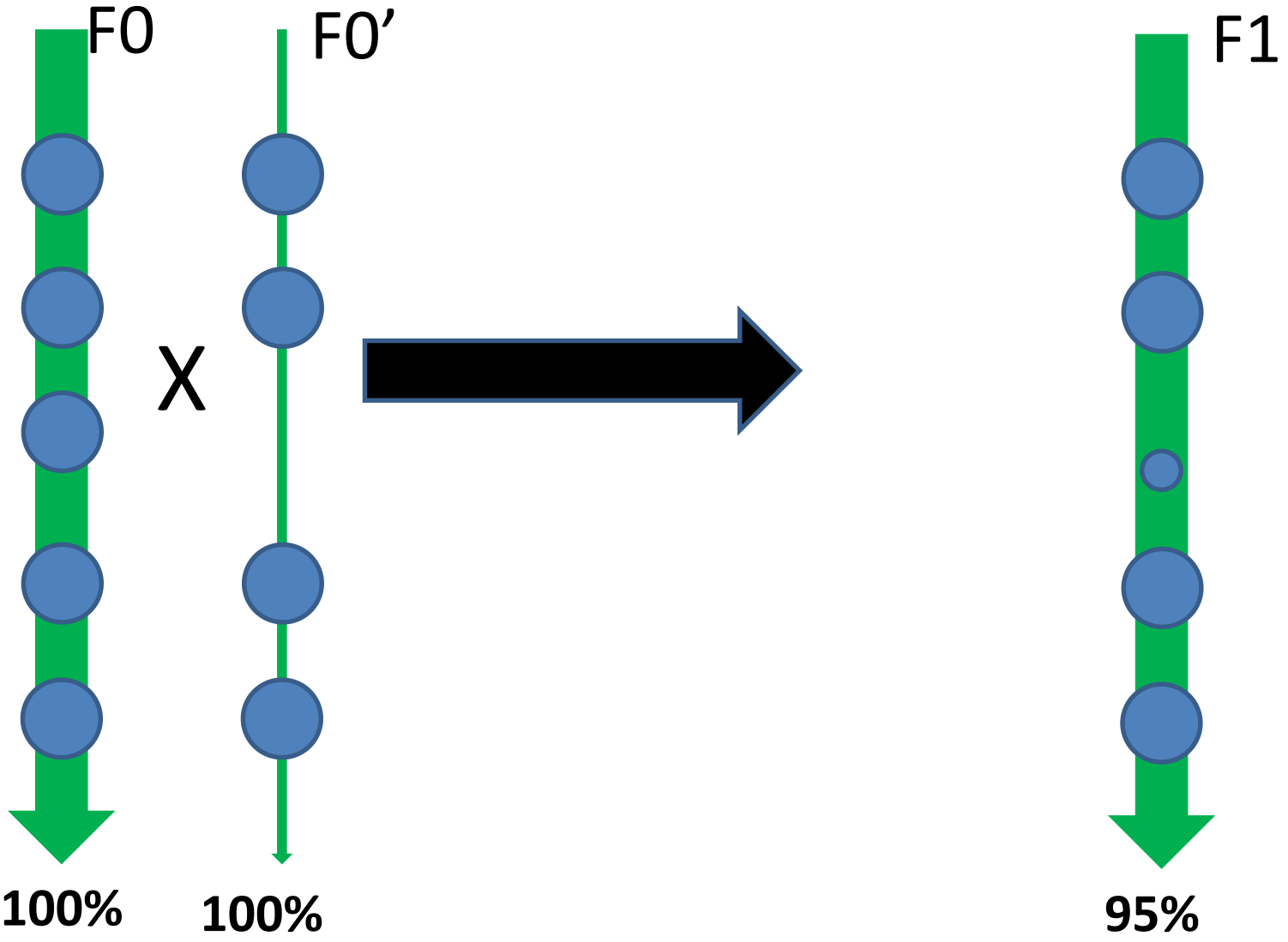
100% **0%** **95%**

Lack of dominance (expectation?)



Function= Flux J

Lack of dominance (network explanation; Kacser)



Function= Flux J

Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

Principles

The Novosibirk example of co-selection:

Selection for domestication also brings drooping ears

Belyaev's observation

By selection

for lack of aggressivity

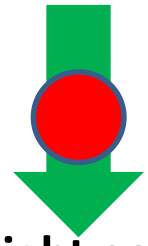
Function: Aggressivity | Upright ears | Domestified | Drooping ears

Singe molecule explanation fails

By selection

for lack of aggressivity

Function: Aggressivity

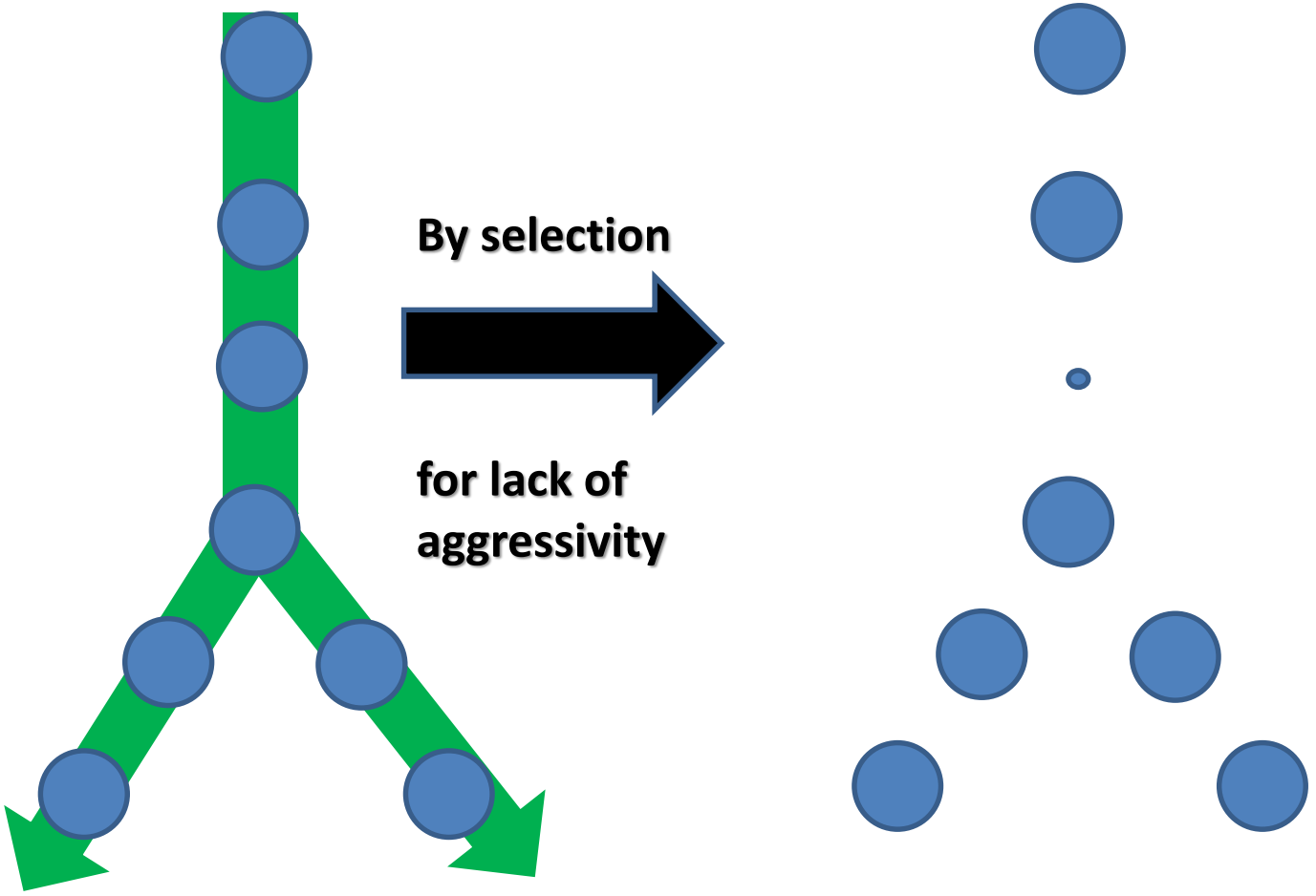


Upright ears

Domesticated

Upright ears

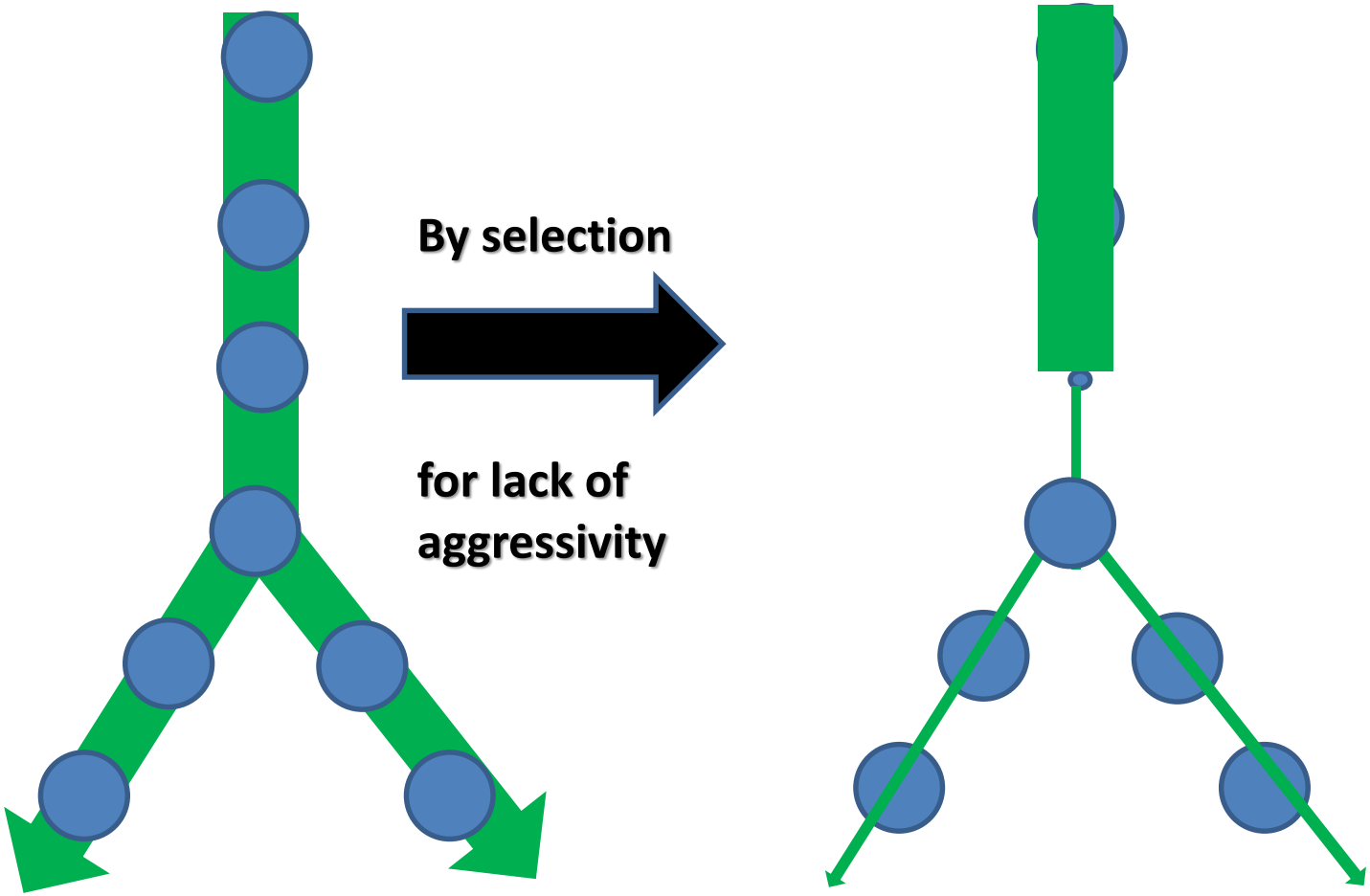
Network explanation



Function: Aggressivity

Upright ears

Network explanation



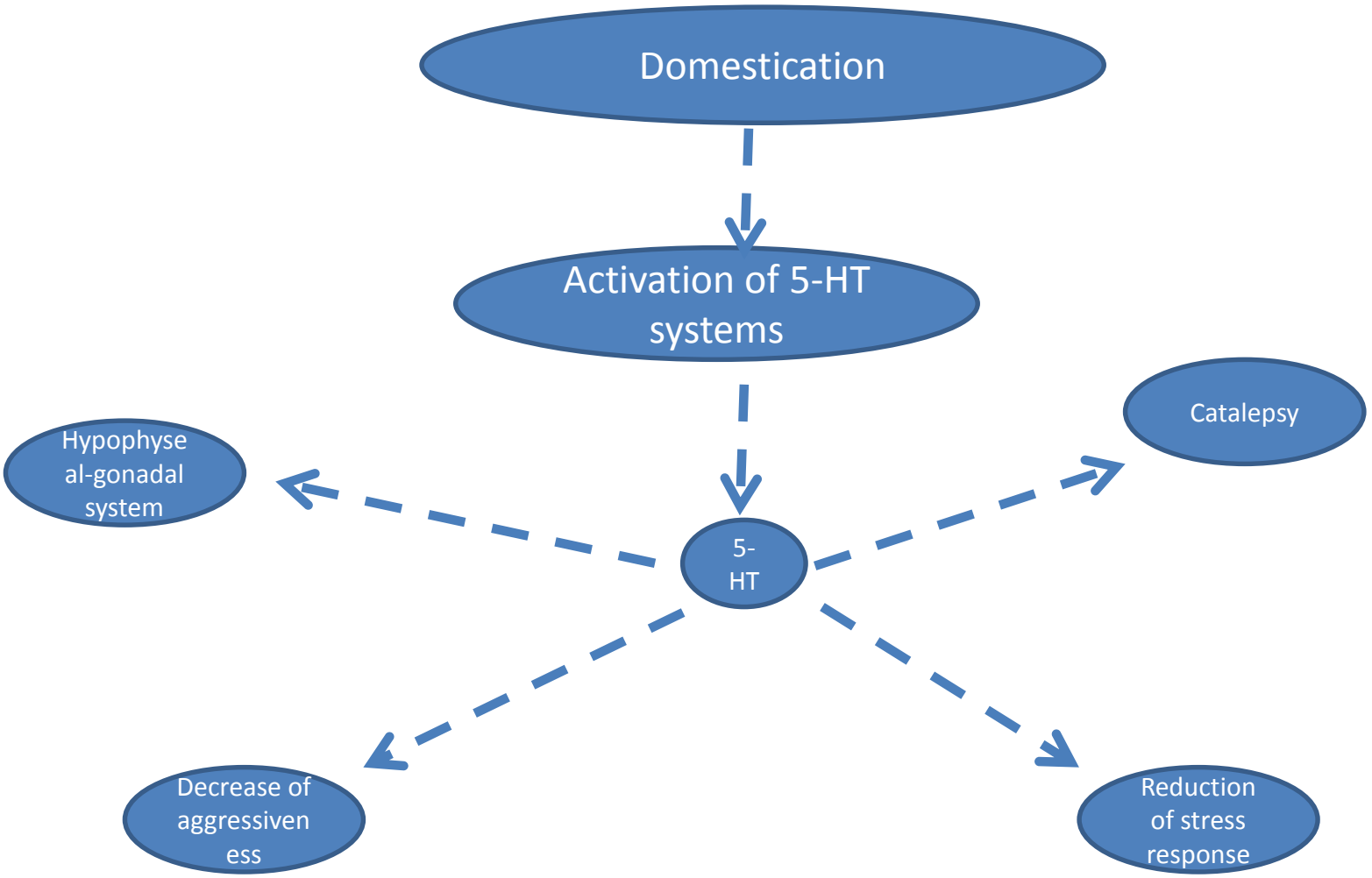
Function: Aggressivity

Upright ears

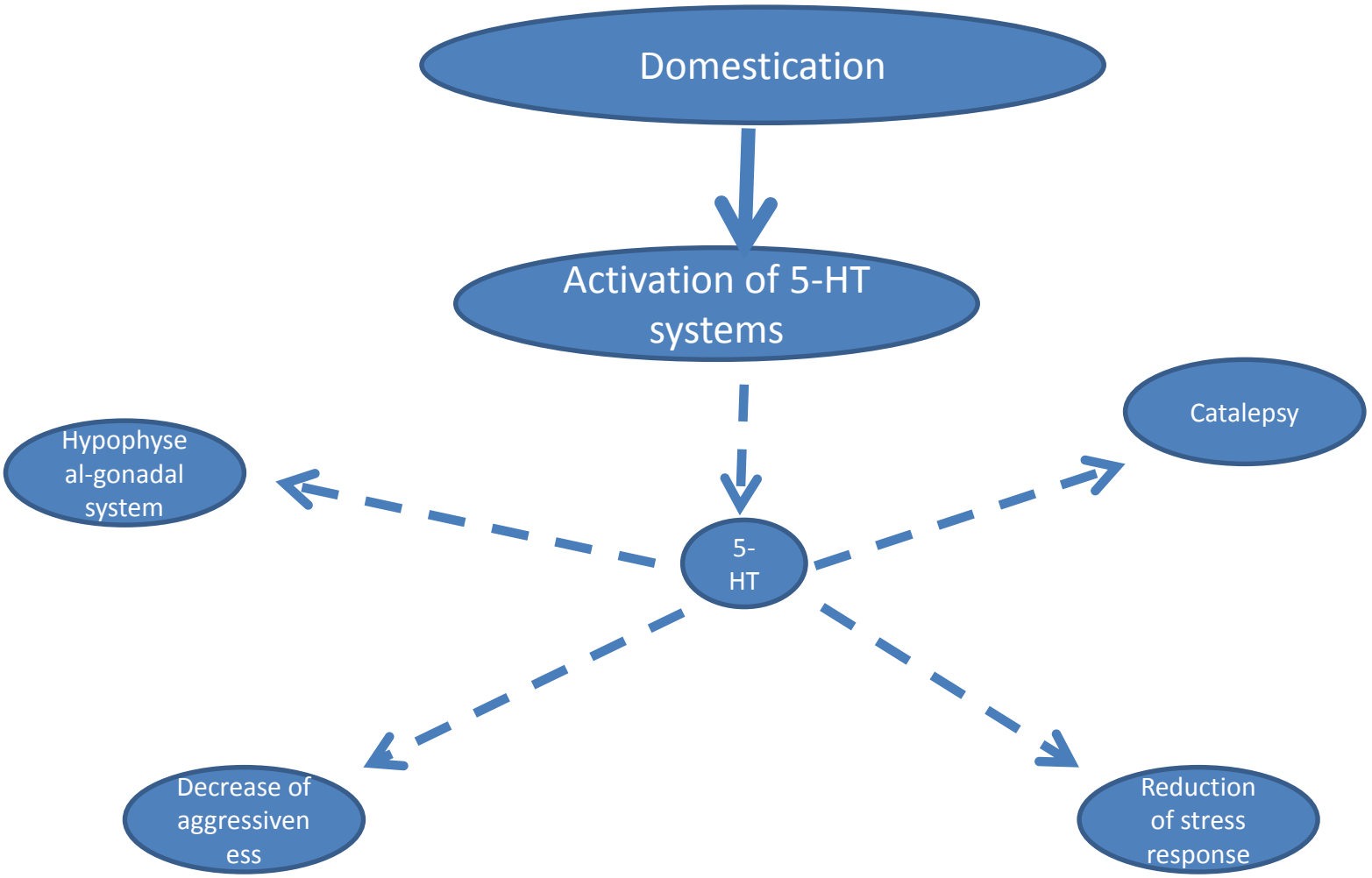
Domestified

Drooping ears

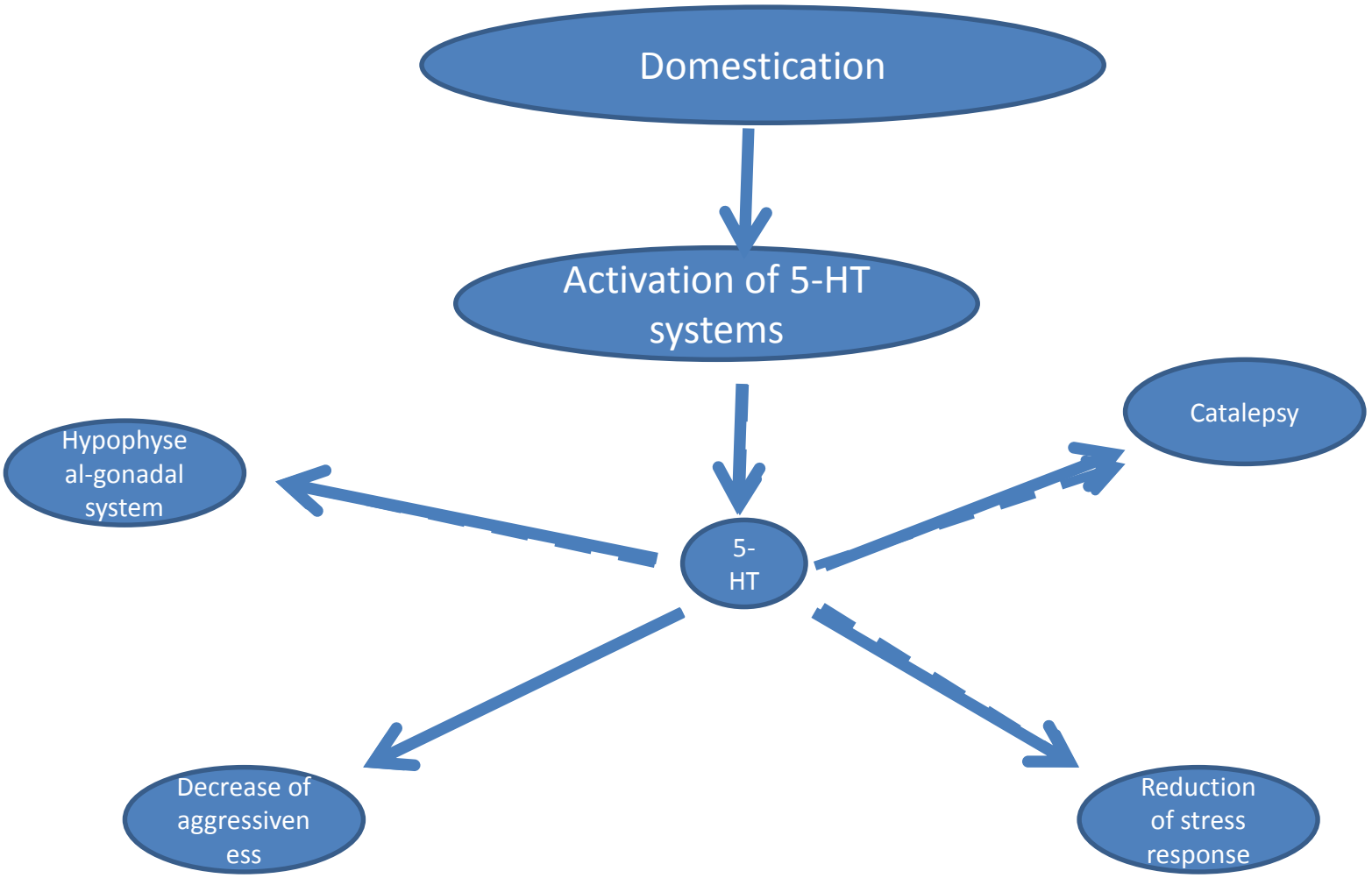
N.K. Popova' lecture



Selection during domestication



After selection



Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

M⁴ @ ISBE.NL team

The Netherlands' arm of ISBE-Light Infrastructure for Systems Biology Europe

ISBE.NL

Alexey Kolodkin, Matteo Barberis, Ablikim Abdukerim, Zahid Hassan, Thierry Mondeel, Samrina Rehman and Hans V. Westerhoff

executive director

founding director

Stefania Astrologo, Ewelina Weglarz-Tomczak, YanFei Zhang

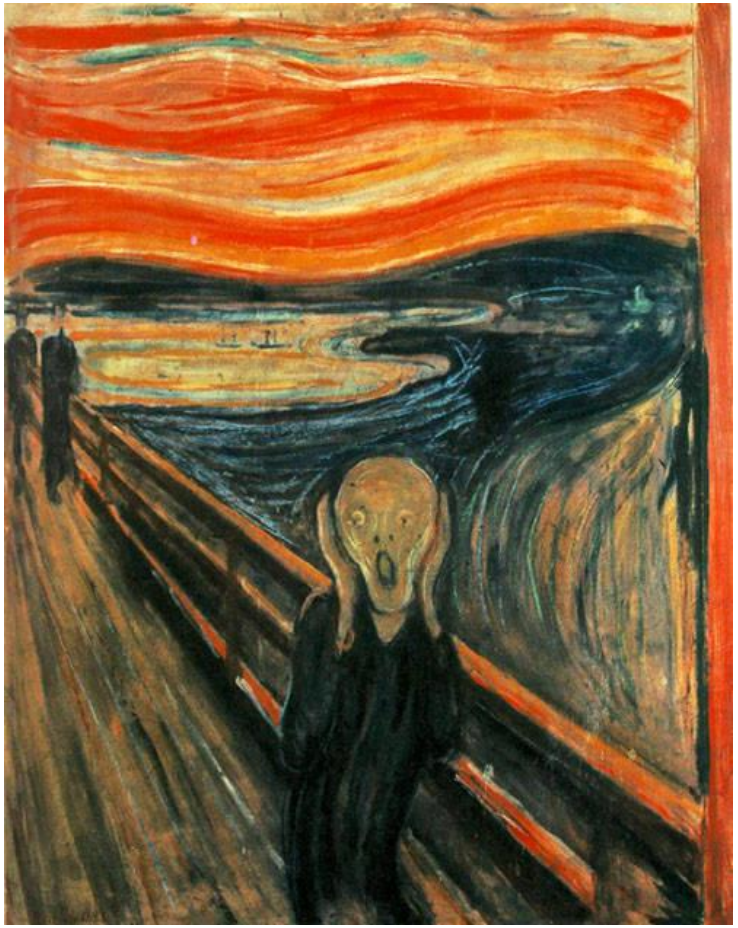
Make Me My Model service: Something 4U?

Systems Biology

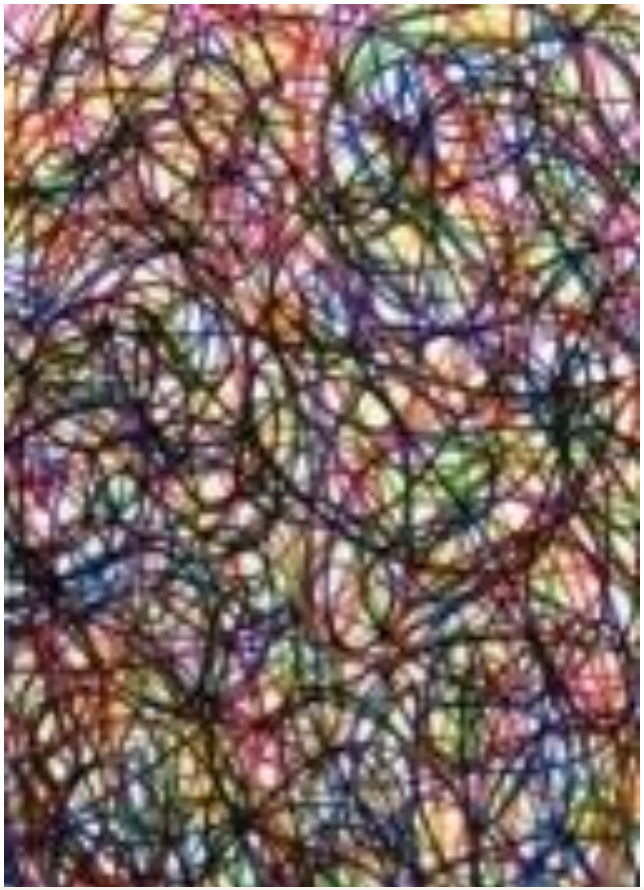
- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

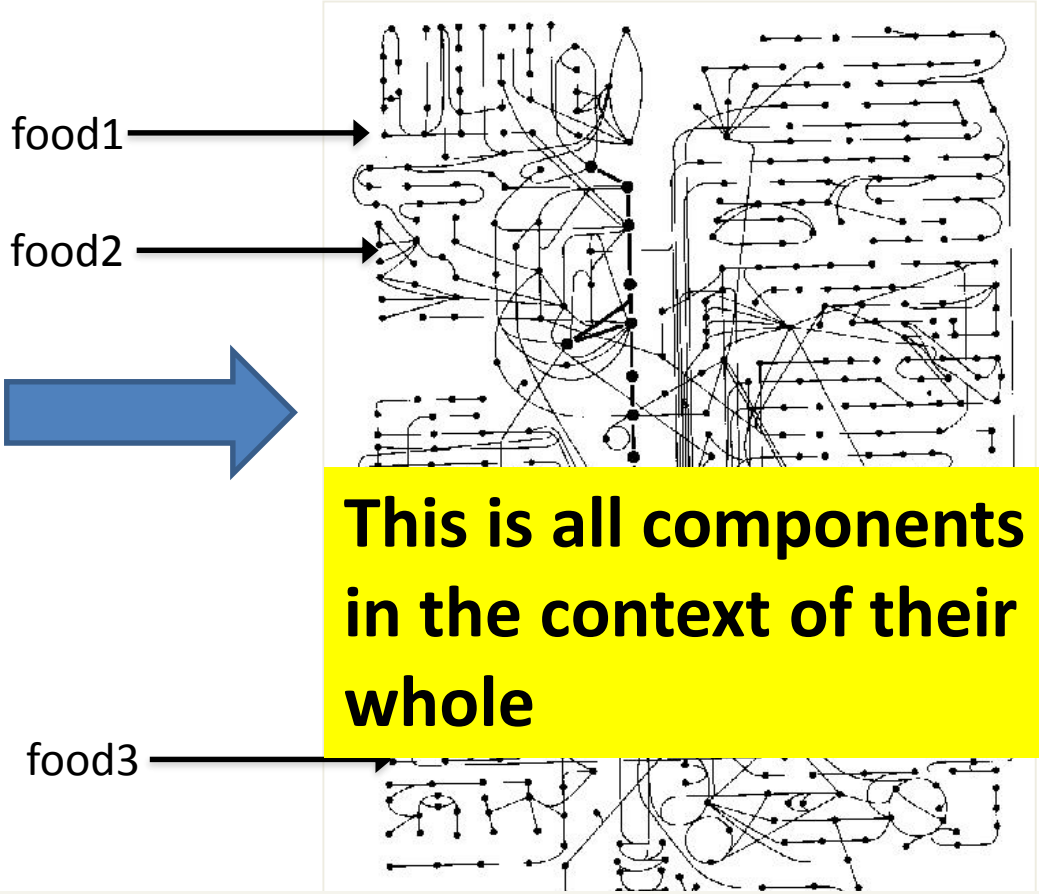
The human: A jungle of 25 000 genes and gene products and various nutrition, life style and ambition factors This must be impossible to deal with.

Where to start……?



We obtained the consensus **genome wide metabolic map (Recon2)**, i.e.
all the human network can make from any nutrition





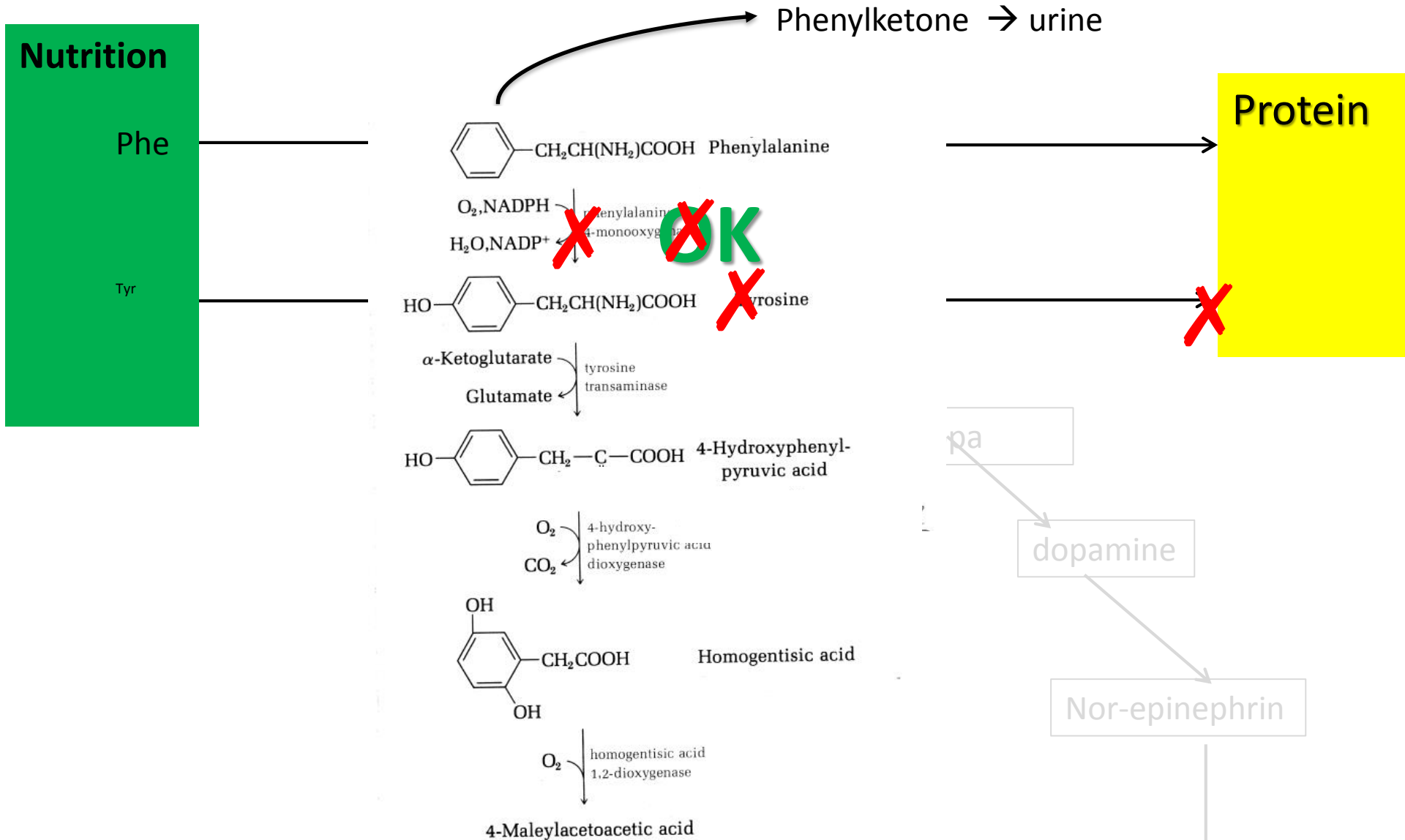
This is all components in the context of their whole

Data concerning all metabolic genes have hereby been integrated into a predictive format.
Predicting how every molecule in our body is made by our body

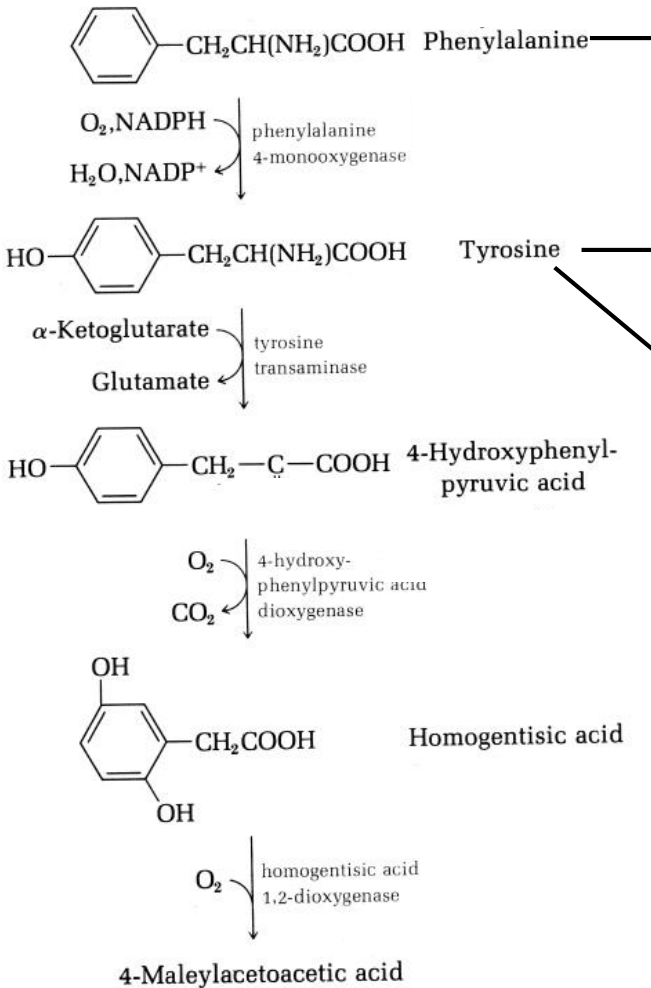
Does this help?

Could it lead to cures?

Example of map utilization tyrosine metabolism: nature Phenylketonuria (PKU) = IEM



Example of map utilization tyrosine metabolism: Nurture and brain



Protein

dopa

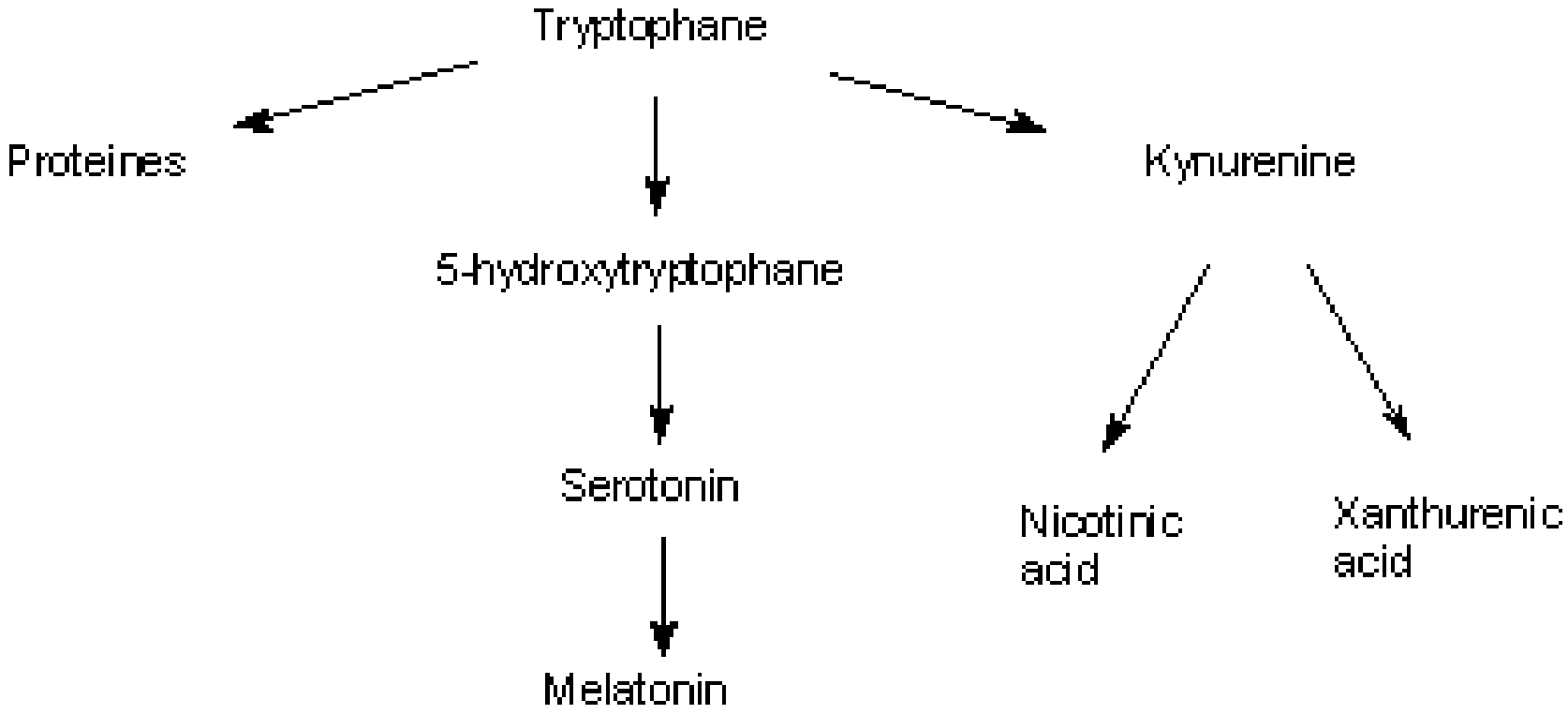
dopamine

(Herbeck presentation)

Nor-epinephrin

Epinephrin=adrenaline

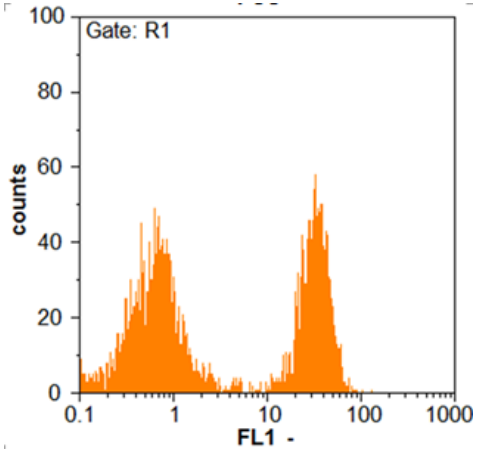
Now: after hearing Prof Popova: Mapping serotonin (?)



Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

How can we understand clonal heterogeneity (such as in these glucose transporter activities)?



Yeast uptake of fluorescently labelled glucose analogue

Explanations
Phenotypic heterogeneity

- **Variations in external conditions (extrinsic noise)**
- **Genetic diversity**
- **Epigenetic diversity:**
  - **Intrinsic Noise**
  - **Bistability**

Explanations
Phenotypic heterogeneity

- Variations in external conditions (extrinsic noise)
- Genetic diversity
- **Epigenetic diversity:**
  - **Intrinsic Noise**
  - **Bistability**

What is noise?

Due to temporarily increased/decreased activities of molecular processes

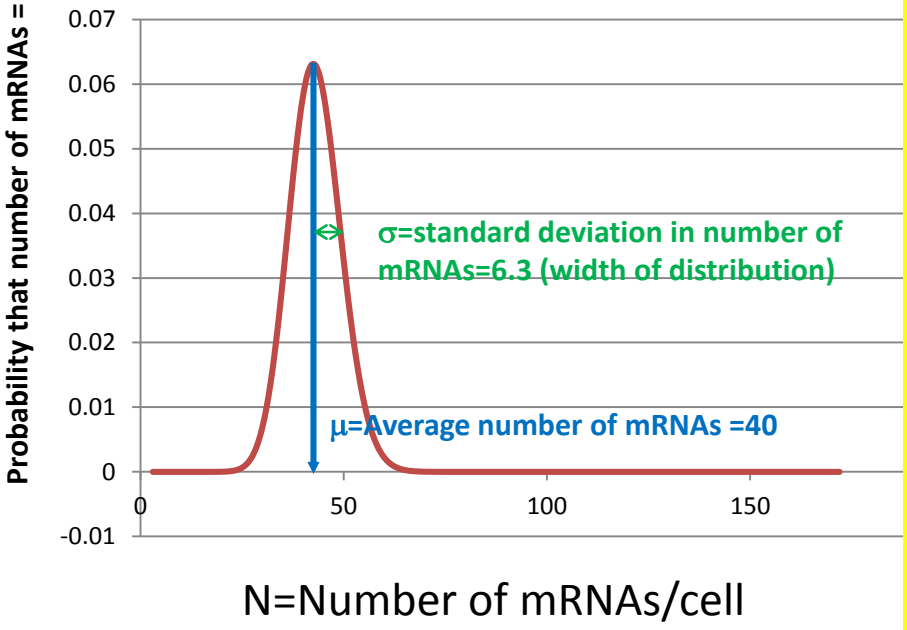
At different moments in different individual cells,

Hence noise →cell-cell heterogeneity

From statistical mechanics:

In a flat (bio)chemical network at steady state, noisy molecule numbers should be (approximately) Poisson distributed

Probability distribution: characteristics



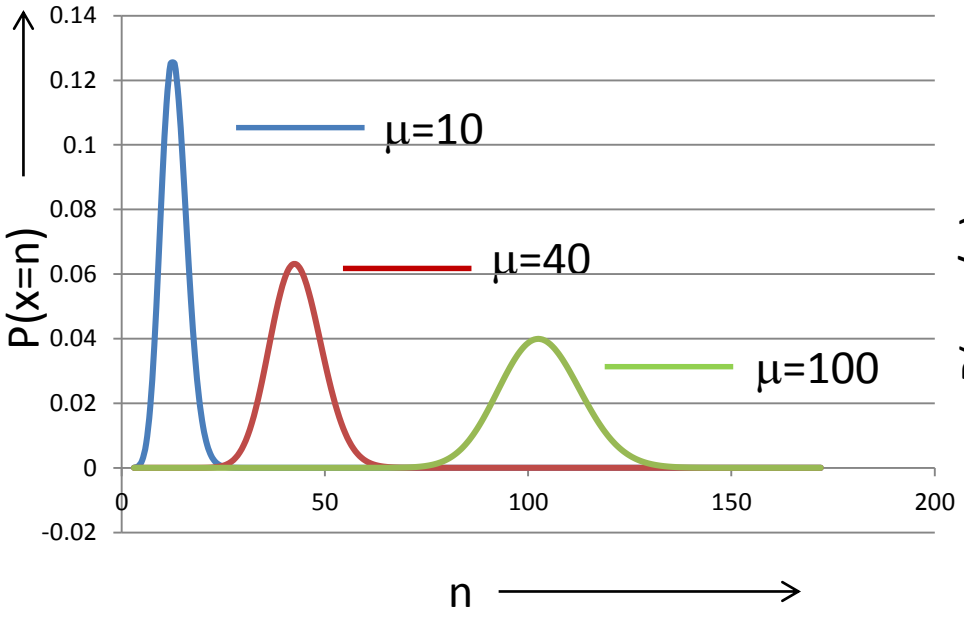
- $\mu = \bar{x}_i$: the average of x
- $\sigma^2 = \overline{(x_i - \mu)^2} = \overline{(x_i)^2} - \mu^2$: Noise:
- $\sigma = standard\ deviation = \sqrt{\sigma^2}$: the width of the distribution
- Relative noise = Coefficient of variation: $cv = {}^{\sigma}/_{\mu}$
- Fano factor: $F = {}^{\sigma^2}/_{\mu}$: deviation from trivial noise

The equilibrium case: Poisson distributed molecule numbers

F=1, i.e. $\sigma^2 = \mu$

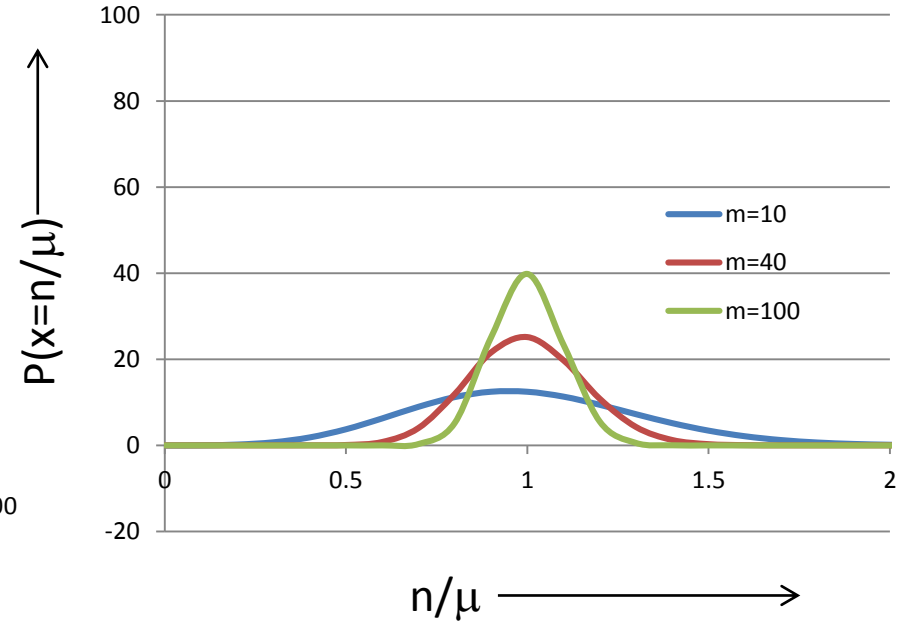
Absolute noise (σ) **in**creases with average number of molecules (μ)

$\sigma = \sqrt{\mu}$

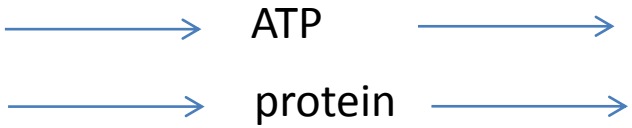


Relative noise ($\frac{\sigma}{\mu}$) **de**creases with average number of molecules (μ)

cv=$\frac{\sigma}{\mu} = \frac{1}{\sqrt{\mu}}$



If Poisson, then for 'normal' reactions and 'usual' molecule numbers: **relative noise < 3%**

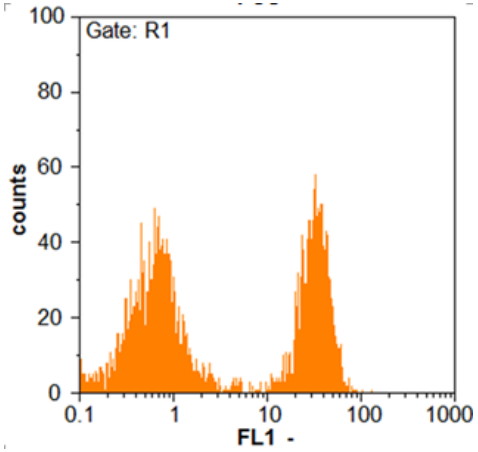


E. coli:
1 μm^3, 5 mM ATP: 3 million molecules; $1/\sqrt{3000000}=0.1$ %
Total protein: 3 million; on average of each type; 1000; **3 % relative noise?**

S. cerevisiae:
40 μm^3, 5 mM ATP: 120 million molecules; <0.01%
Total protein: 100 million; on average of each type; 16 000; **<1 % relative noise?**

Mammalian cell:
2 pL, 5 mM ATP: 6 billion molecules; << 0.01%
Total protein: 10 billion; on average of each type; 400 000**; <0.2 % relative noise?**

Therefore: Poisson noise can not explain our (and others') observations.

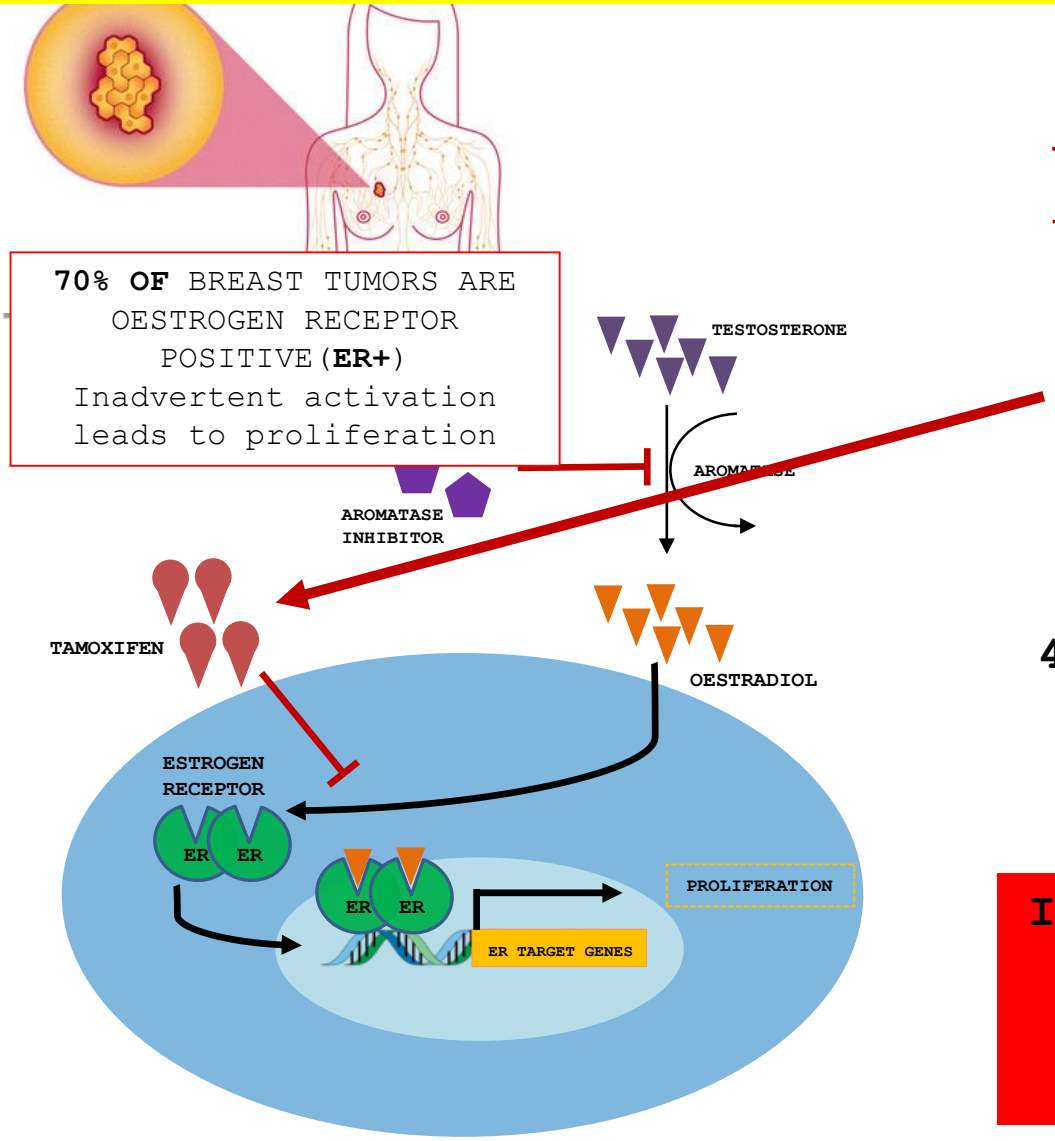


But what can?

And:

Is such noise important?

We now study this vis-à-vis oestrogen receptor positive breast cancer



Therefore these cases are treated with tamoxifen that binds to the Estrogen receptor without activating it

However

40-50% OF ER+ PATIENTS **DEVELOP RESISTANCE**

Do all cancer cells become resistant,

INTRATUMOR HETEROGENEITY REPRESENTS A **MAJOR OBSTACLE TO** EFFECTIVE **CANCER TREATMENT.**

But these outgrow the others

Noise in MCF-7 (clonal) cells; noisy CD44 promotor

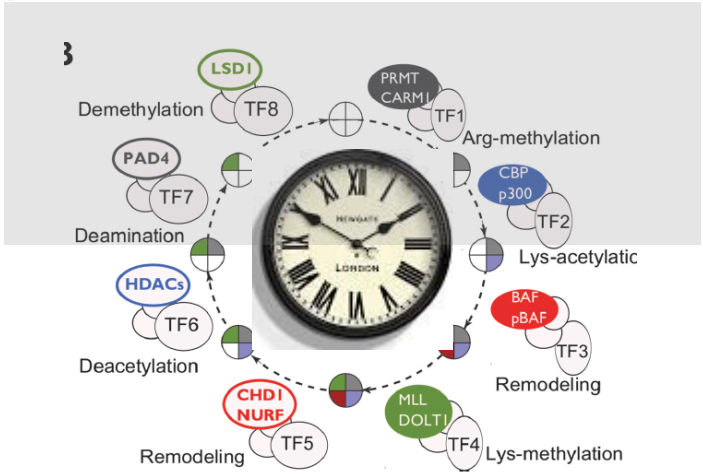
- 4 MCF-7 **sister cells, DAPI stain, GFP that reports CD44 promoter activity(possibly resistance related)** and **GAPDH** FISH mRNA probe.

And as shown by Moshkin this morning it may be important for IVF

Multiplex Eukaryotic Transcription (In) activation: Timing, Bursting and Cycling of a Ratchet Clock Mechanism

(2015) PLoS Comput Biol 11(4): e1004236.
doi:10.1371/journal.pcbi.1004236

Katja N. Rybakova[1‡¤], Frank J. Bruggeman[2‡], Aleksandra Tomaszewska[3], Martijn J. Moné[1], Carsten Carlberg[3], Hans V. Westerhoff[1,4,5*]



Between 3 and 9 h precisely b mRNAs molecules synthesized:
b = burst size

Could such epigenetics lead to transcription bursting and to Non-Poisson distributed mRNA?

$$\textbf{The Fano factor} = F \overset{\text{def}}{=} \frac{\sigma^2}{\overline{n}} = \left(\sigma^2/\overline{n}\right) \Big/ \left(\sigma^2/\overline{n}\right)_{Poisson}$$

is a measure of the deviation from Poisson distribution

So, the issue is whether F >>1 due to bursting

Modelling backed up by Statistical Mechanics

Advance in Chemical Physics, Volume XXXIV
Edited by I. Prigogine, Stuart A. Rice

THE EXPANSION OF THE MASTER EQUATION

N. G. VAN KAMPEN

Institute of Theoretical Physics of the University, Utrecht, Netherlands

Probability to have n' mRNAs at time t

Probability to have n mRNAs at time t

$$\dot{P} \stackrel{\text{def}}{=} \frac{dP(n,t)}{dt} = \sum_{n'=0}^{\infty} W(n|n') \cdot P(n',t) - W(n'|n) \cdot P(n,t)dt$$

Increase in the probability that a cell contains n RNA molecules

Probability for cell with n' RNA molecules to become a cell with n RNA molecules

Probability for cell with n RNA molecules to become a cell with n' RNA molecules

For bursting transcription and linear degradation of mRNAs we derived for the noise in the number of mRNAs:

variance

$$F_{steady\ state} \overset{\text{def}}{=} \frac{\sigma^2}{\overline{N}} = \frac{burst\ size + 1}{2}$$

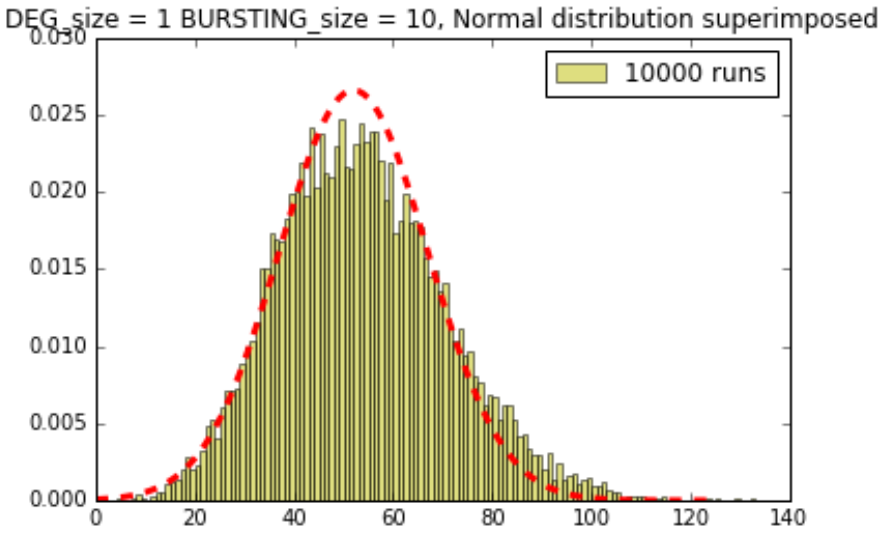
Mean number of mRNA molecules

Implications:

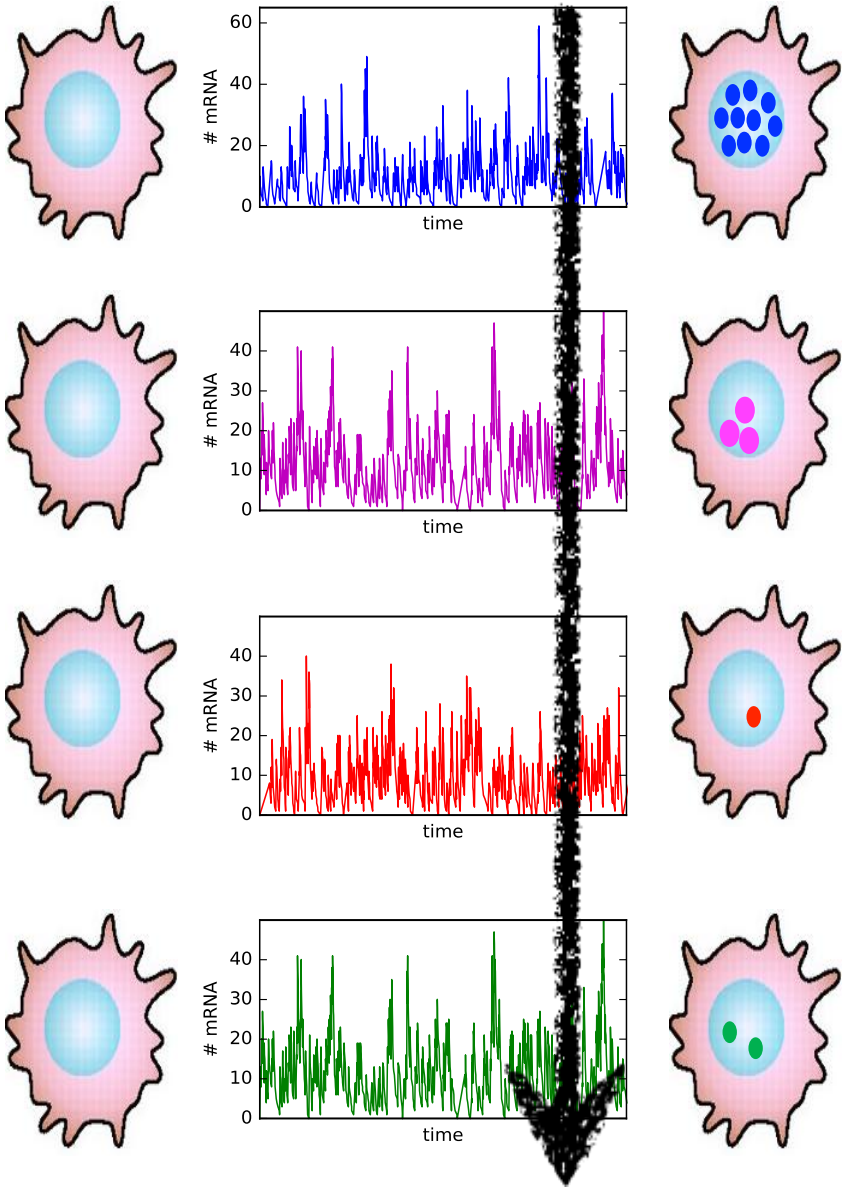
- For burst size of 1: F=1, distribution is Poisson, variance equals the average
- F is ONLY a function of burst size, not of burst kinetics
- For burst size =100: F=50.5 and distribution far from Poisson, variance 50 times the average, so **Yes, bursting can cause high F's**

Confirmations by Gillespie modelling

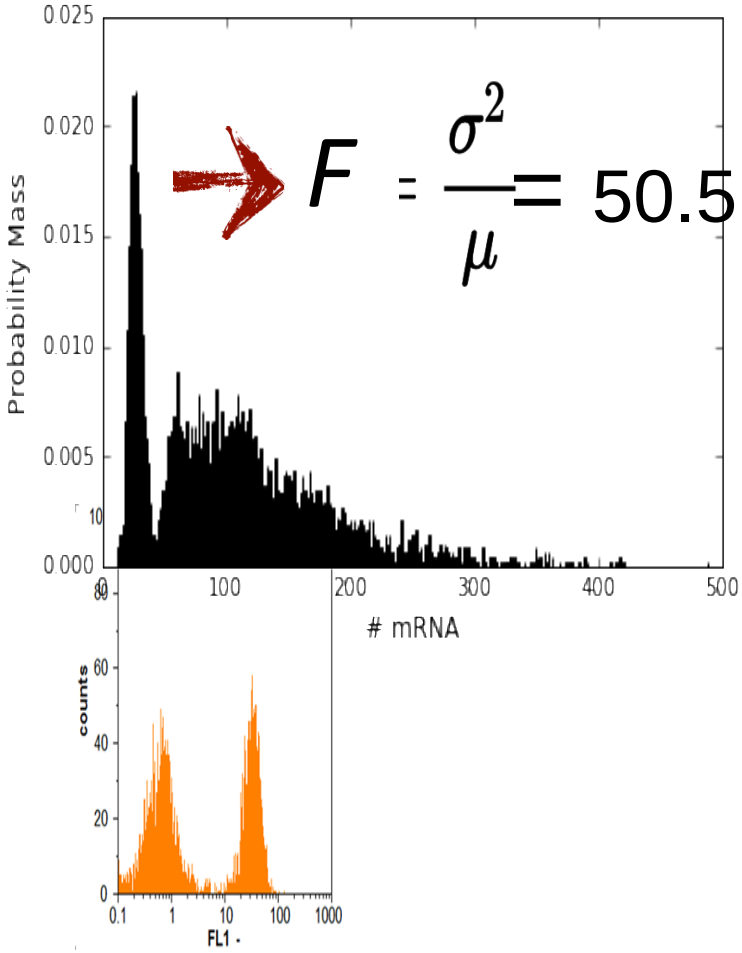
Bursting: indeed a distribution much broader than Poisson



This can even produce a bimodal distribution

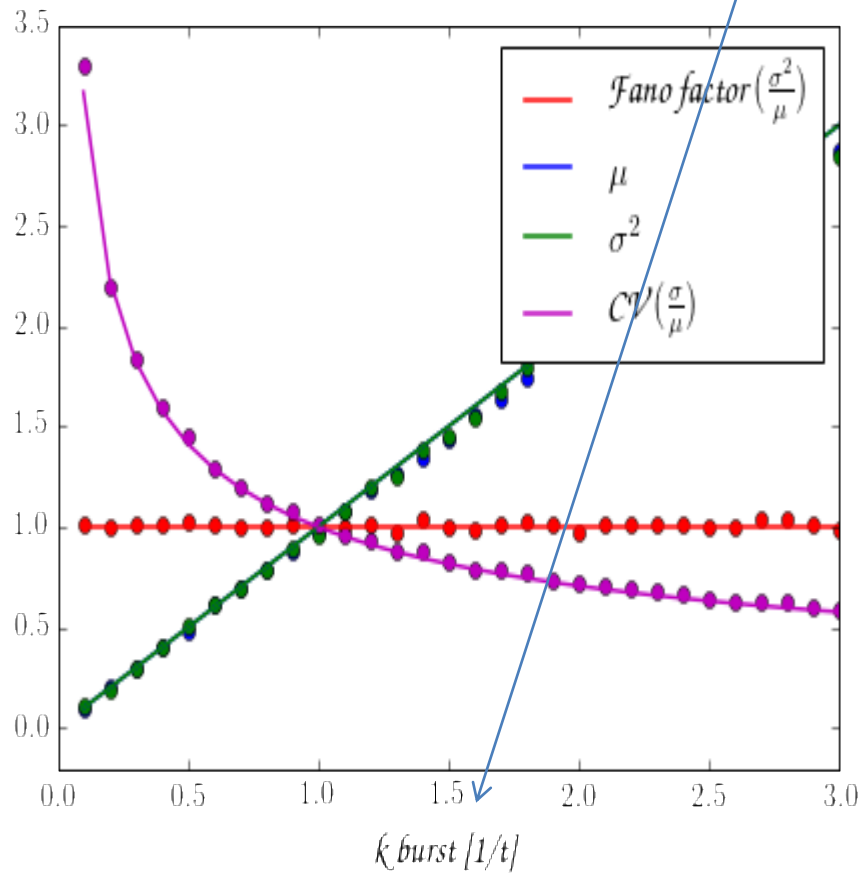


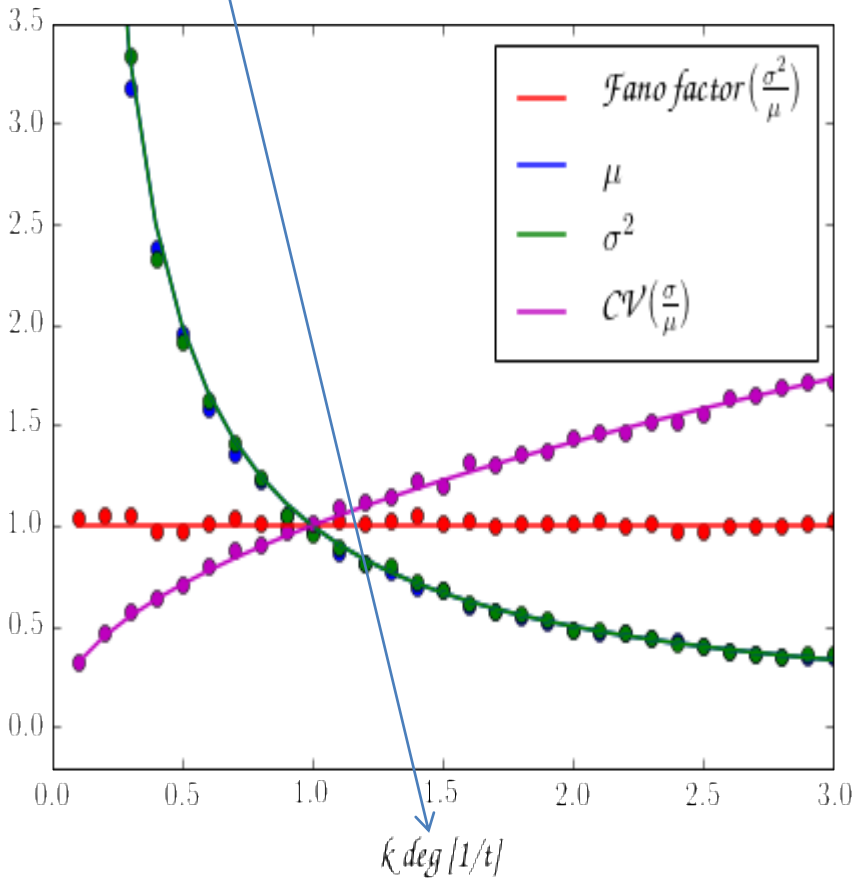
$k_{deg} = 0.1$
$k_{burst} = 0.1$
Burst size $b = 100$



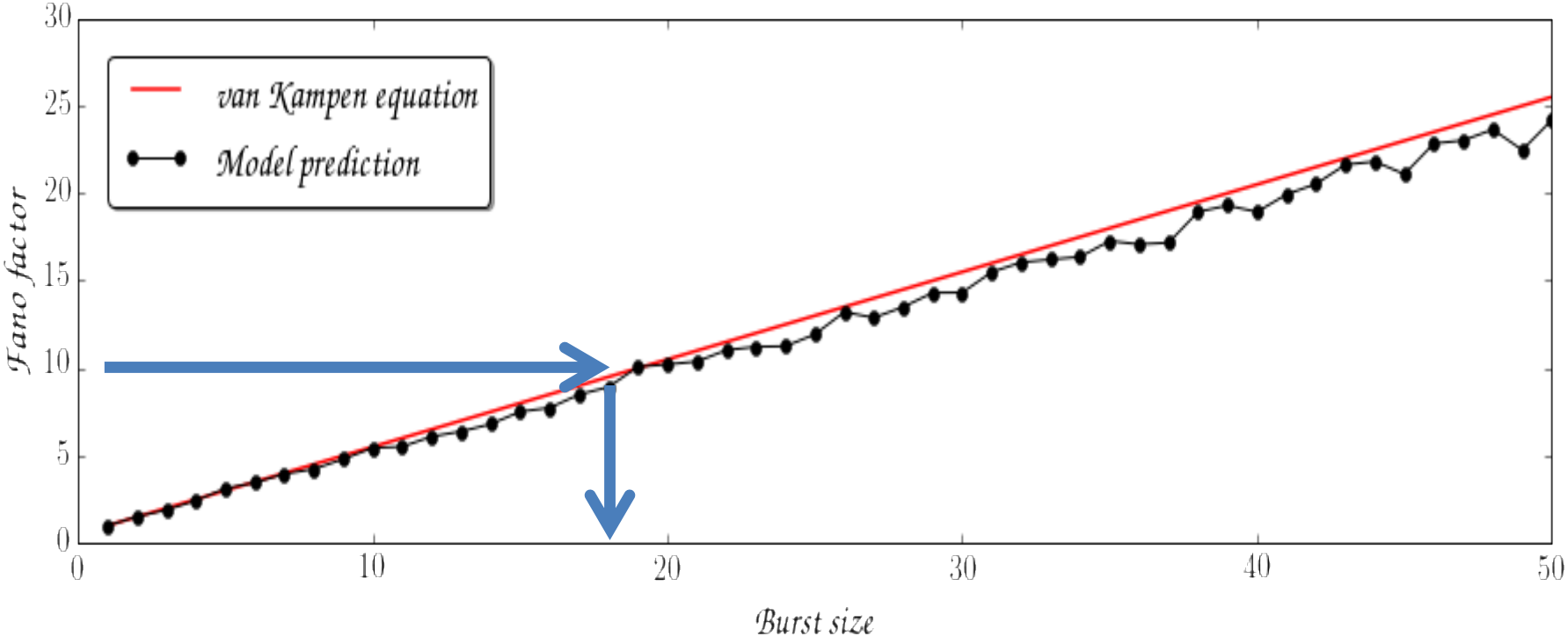
$$F = \frac{\sigma^2}{\mu} = 50.5$$

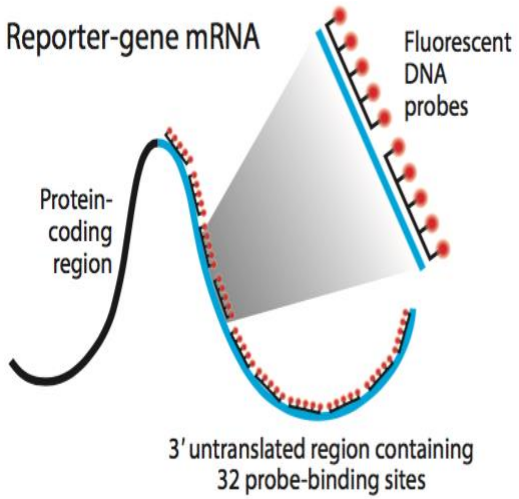
Indeed, the Fano factor is independent of k_{burst} and k_{deg} , whereas the other noise factors are not



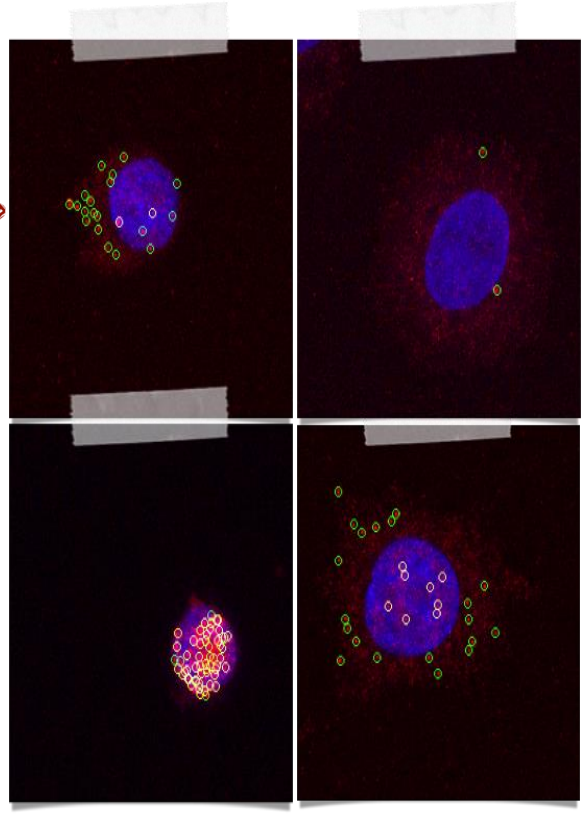


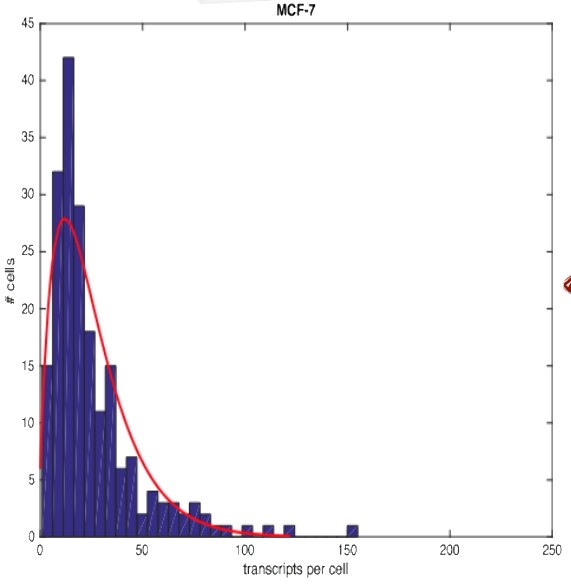
Hence: one may infer the burst size from the Fano factor precisely because of the non identifiability



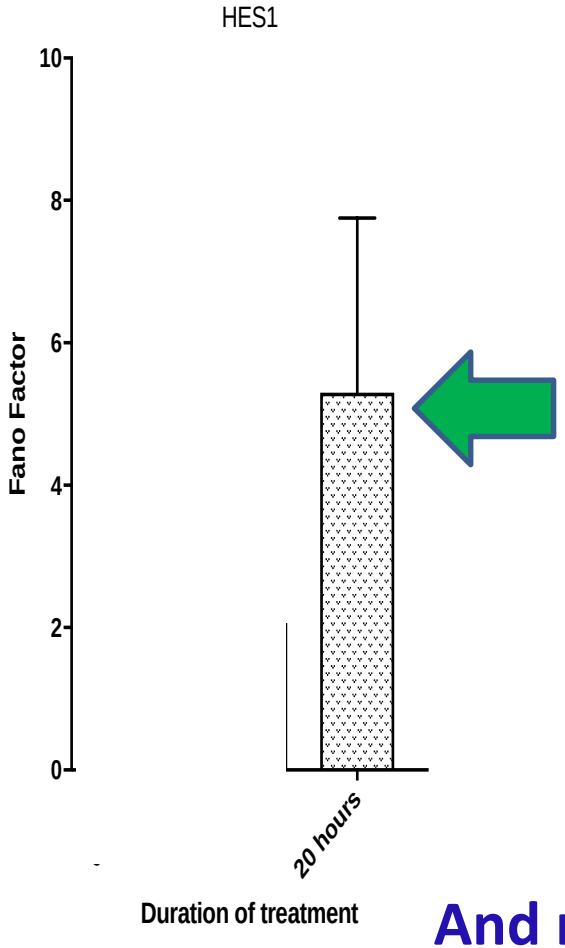


MCF-7 cell, DNA stained with DAPI (blue) and mRNA stained with fluorescent ssDNA probe (red). Green and yellow circles enclose single mRNA molecules.





Distributions of the number of transcripts obtained through smFISH experiments, in MCF-7 cell lines.

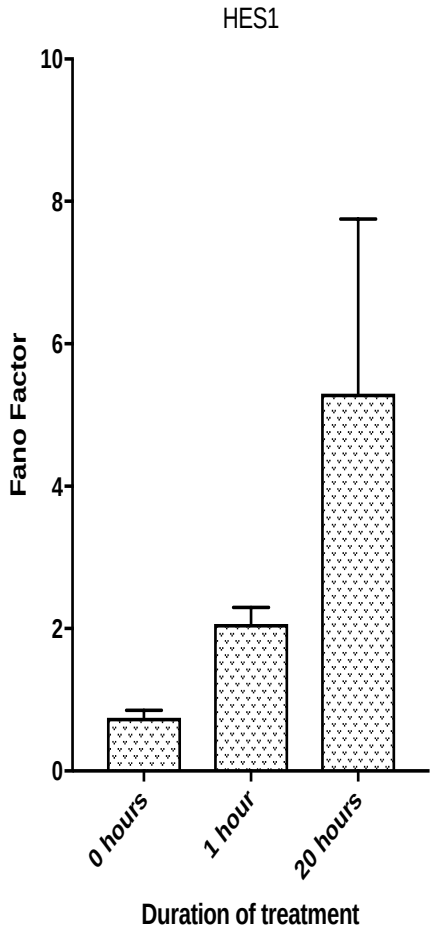


i.e. burst size = 9

So Yes, the observed diversity in mRNA could be explained by our epigenetic transcription clock!

And now for an intriguing discovery (??)

When starting from resting cells F increases as if cells begin to 'exploring'



Statistical mechanics:

$$\frac{dF}{dt} \sim \frac{b \cdot k_s}{\bar{n}} \cdot (b - F) \sim \frac{b - F}{t} > 0$$

Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - **Epigenetics and noise/cell diversity**

Systems Biology

- What is it?
- Principles
  - Lack of dominance (Kacser)
  - Co-selection (Belyaev)
- Progress
  - Make Me My Model
  - The genome wide metabolic maps
  - Epigenetics and noise/cell diversity

Acknowledgements

Pernette Verschure

EpiPredict

Thierry Mondeel

Will Beckman

Stefania Astrologo

and many other teachers, colleagues and students

Before the tea break

14:40	**Plenary lecture:** *Hans V. Westerhoff*. Principles of systems biology and Dmitri Belyaev's co-selection of traits
15:10	*N.A. Kolchanov, S.A. Lashin*. Regulatory circuits in gene networks: organization and evolution
15:30	*N. Sahin, H.V. Westerhoff, A. Kolodkin*. Designing the emergence of progressive (PROP) and regressive (REP) preconditioning responses: from intelligent intracellular networks to the domestication of animals
15:50	*J.G. Koster*. Systems forensics: systems biology and the inference of crime
16:10	*M.V. Sharakhova, A.A. Yurchenko, A.N. Naumenko, G.N. Artemov, V.N. Stegniy, I.V. Sharakhov*. Evolutionary history of the malaria mosquitoes from maculipennis group in Eurasia