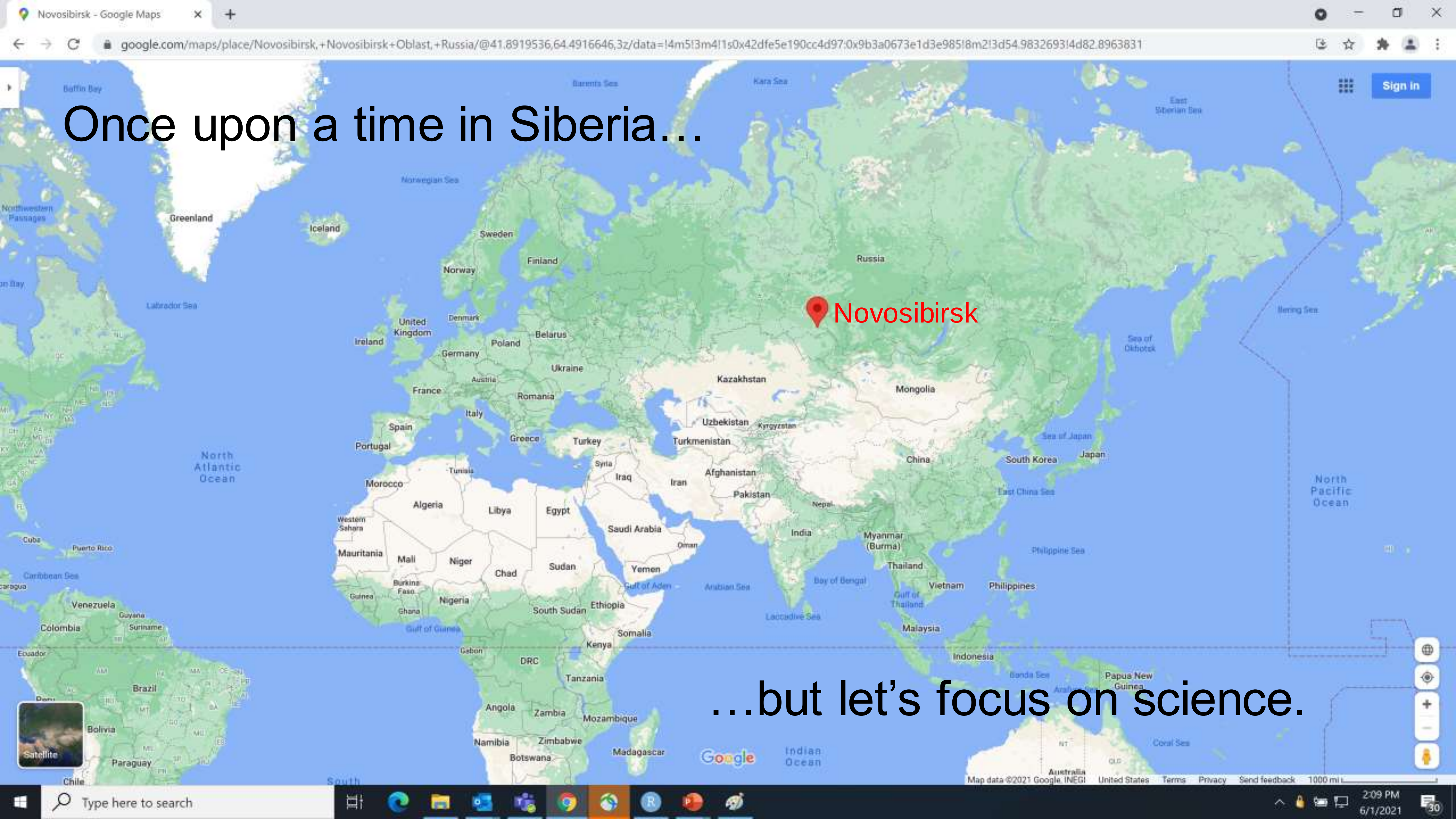
My Journey Through Computational Biology

Alexander (Alex) Mitrophanov, PhD
Senior Statistician
FNLCR/NIH/DMS/BRM*i*

ABCS “Meet and Greet”
4 June 2021


Novosibirsk - Google Maps
google.com/maps/place/Novosibirsk,+Novosibirsk+Oblast,+Russia/@41.8919536,64.4916646,3z/data=!4m5!3m4!1s0x42dfe5e190cc4d97:0x9b3a0673e1d3e985!8m2!3d54.9832693!4d82.8963831
Sign in
Once upon a time in Siberia…
Novosibirsk
Russia
Greenland
Iceland
Sweden
Finland
Norway
United Kingdom
Ireland
Denmark
Belarus
Poland
Germany
Ukraine
France
Romania
Italy
Spain
Portugal
Greece
Turkey
Kazakhstan
Mongolia
Uzbekistan
Turkmenistan
China
South Korea
Japan
North Atlantic Ocean
North Pacific Ocean
Morocco
Algeria
Libya
Egypt
Iraq
Iran
Afghanistan
Pakistan
Saudi Arabia
India
Myanmar (Burma)
Thailand
Vietnam
Philippines
Malaysia
Indonesia
Papua New Guinea
Mali
Niger
Chad
Sudan
Yemen
Mauritania
Nigeria
Ethiopia
South Sudan
Somalia
Kenya
DRC
Tanzania
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
Venezuela
Colombia
Brazil
Bolivia
Paraguay
Chile
Indian Ocean
…but let’s focus on science.
Map data ©2021 Google, INEGI
Type here to search
2:09 PM
6/1/2021

The Purifier

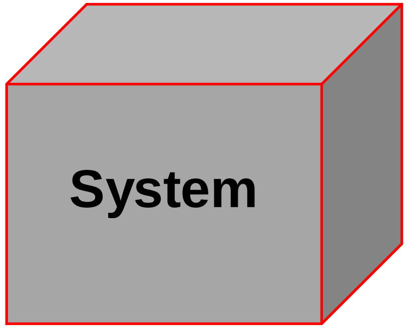
“We all began as something else…”
(*The Chronicles of Riddick*)

“We all began as something else…”

(*The Chronicles of Riddick*)

The Purifier

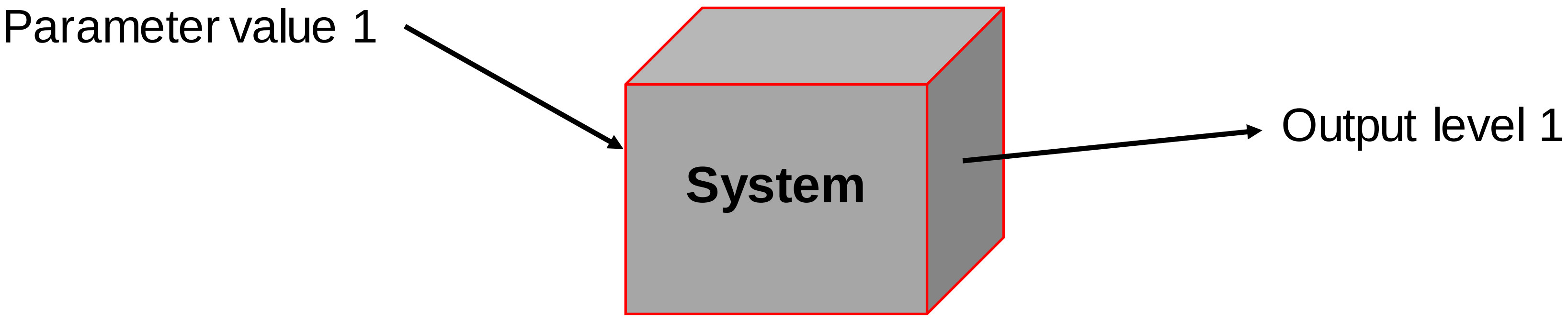
Mathematics of system stability



The Purifier

“We all began as something else…”
(*The Chronicles of Riddick*)

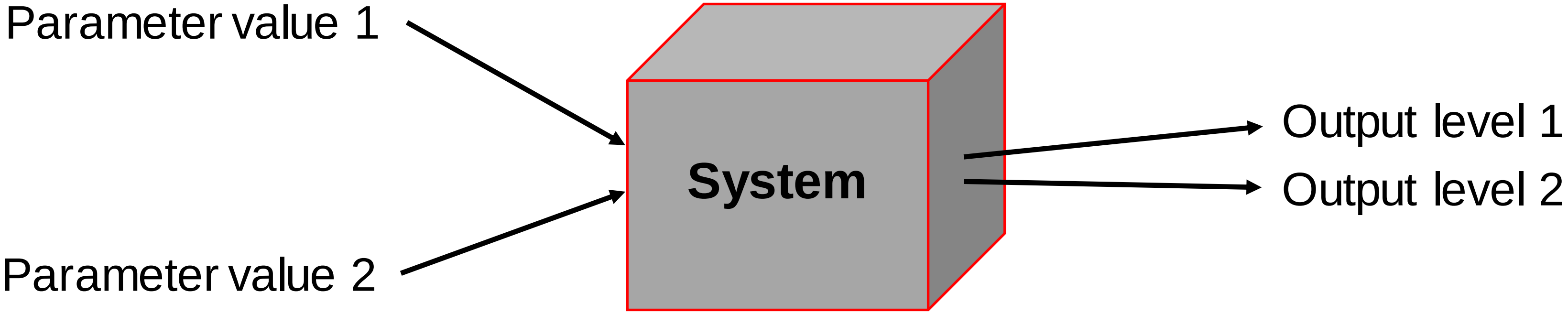
Mathematics of system stability



The Purifier

“We all began as something else…”
(*The Chronicles of Riddick*)

Mathematics of system stability



How can we quantify/predict stability?

Stability, chemical kinetics, and Markov chains

$$A + B \underset{k_-}{\overset{k_+}{\longleftrightarrow}} AB$$

Parameters

- initial concentrations ("Type 1")
- rate constants k_+, k_- ("Type 2")

Stability, chemical kinetics, and Markov chains

$$A + B \underset{k_-}{\overset{k_+}{\longleftrightarrow}} AB$$

Parameters

- initial concentrations (“Type 1”)
- rate constants k_+, k_- (“Type 2”)

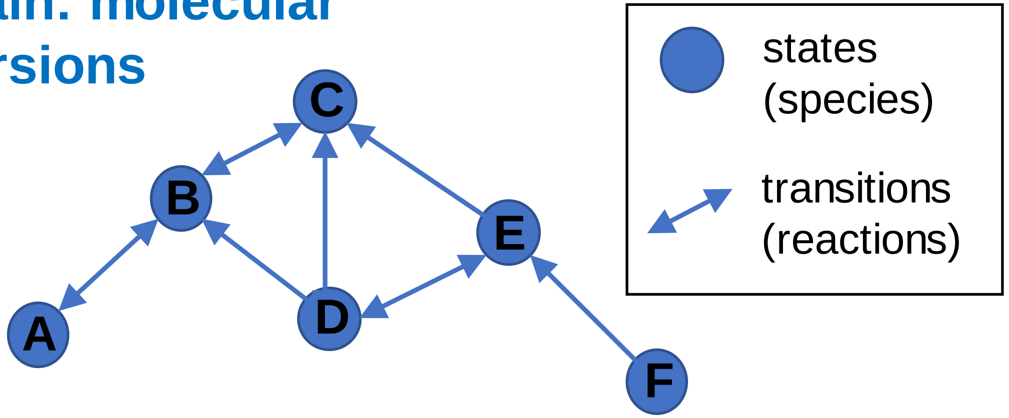
Stability, chemical kinetics, and Markov chains

$$A + B \underset{k_-}{\overset{k_+}{\longleftrightarrow}} AB$$

Parameters

- initial concentrations ("Type 1")
- rate constants k_+, k_- ("Type 2")

Markov chain: molecular interconversions



Stability, chemical kinetics, and Markov chains

$$A + B \underset{k_-}{\overset{k_+}{\longleftrightarrow}} AB$$

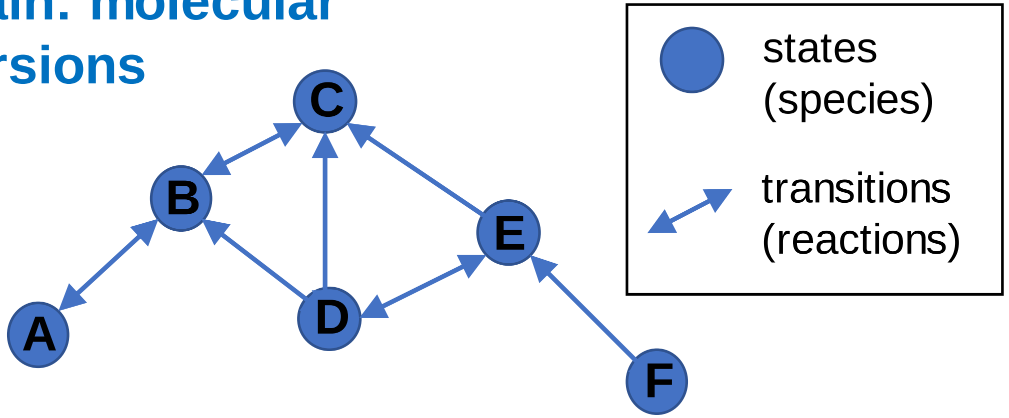
Parameters

- initial concentrations ("Type 1")
- rate constants k_+, k_- ("Type 2")

Type-1 stability → *quantification* → Type-2 stability

Main result!

Markov chain: molecular interconversions



$$\mathbf{Q} = (q_{ij})$$

transition rate matrix (rate constants)

$$\mathbf{p}(t) = (p_i(t))$$

state probability vector

$$d\mathbf{p}(t)/dt = \mathbf{p}(t)\mathbf{Q}$$

governing equation

(Kolmogorov **differential** equation)

Recent applications of the theory

Stat Comput (2016) 26:29–47
DOI 10.1007/s11222-014-9521-x

CrossMark

Noisy Monte Carlo:
with approximate tr

P. Alquier · N. Friel · R. Everit

Received

Abstract Monte Carlo algorithm
distribution π by simulating a Ma
kernel P such that π is invariant
are many situations for which it i
ble to draw from the transition ke
is the case with massive datasets,
expensive to calculate the likelih
for intractable likelihood models a
Gibbs random fields, such as those

Bernoulli **24**(4A), 2018, 2610–
https://doi.org/10.3150/17-BEJ

Perturbation th
via Wassersteir

DANIEL RUDOLF[1] and N

Perturbation theory for Markov c
probabilities of Markov chains ar
ful and flexible bounds on the dis
them satisfies a Wasserstein ergod
mate Markov chain Monte Carlo (
based on Lyapunov functions, we
assumptions. In an autoregressive
ory by showing quantitative estim
Metropolis–Hastings and stochasti

Keywords: big data; Markov chain

Statistics and algorithms

Statistics and Computing
https://doi.org/10.1007/s11222-018-9817-3

Informed sub-sampling MCMC: a
datasets

Florian Maire[1,2] · Nial Friel[1,2] · Pierre Alquier[3]

Received: 26 June 2017 / Accepted: 4 June 2018

Abstract

This paper introduces a framework for speeding up

Recent applications of the theory

JOURNAL OF MATHEMATICAL PHYSICS **54**, 032203 (2013)

CrossMark

Perturbation bounds for quantum Markov processes and their fi

Oleg Sze
Departmen

(Received 2

We investi
tions of th
measures
establish u
eigenvalue
processes
the evolut
mixing tin
respect to
[http://dx.

I. INTRODUCTIO

Quantum Ma
quantum statistica
evolution of some
as a quantum Ma

Artificial quantum

temperature of a
quantum bath tha
state $\frac{e^{-H/T}}{\mathrm{Tr}(e^{-H/T})}$ wit
algorithm. For a
an engineered th
thermodynamica
we discuss how a

DOI: 10.1103/Ph

I. INTR

PRL **116,** 020502 (2016) PHYSICAL REVIEW

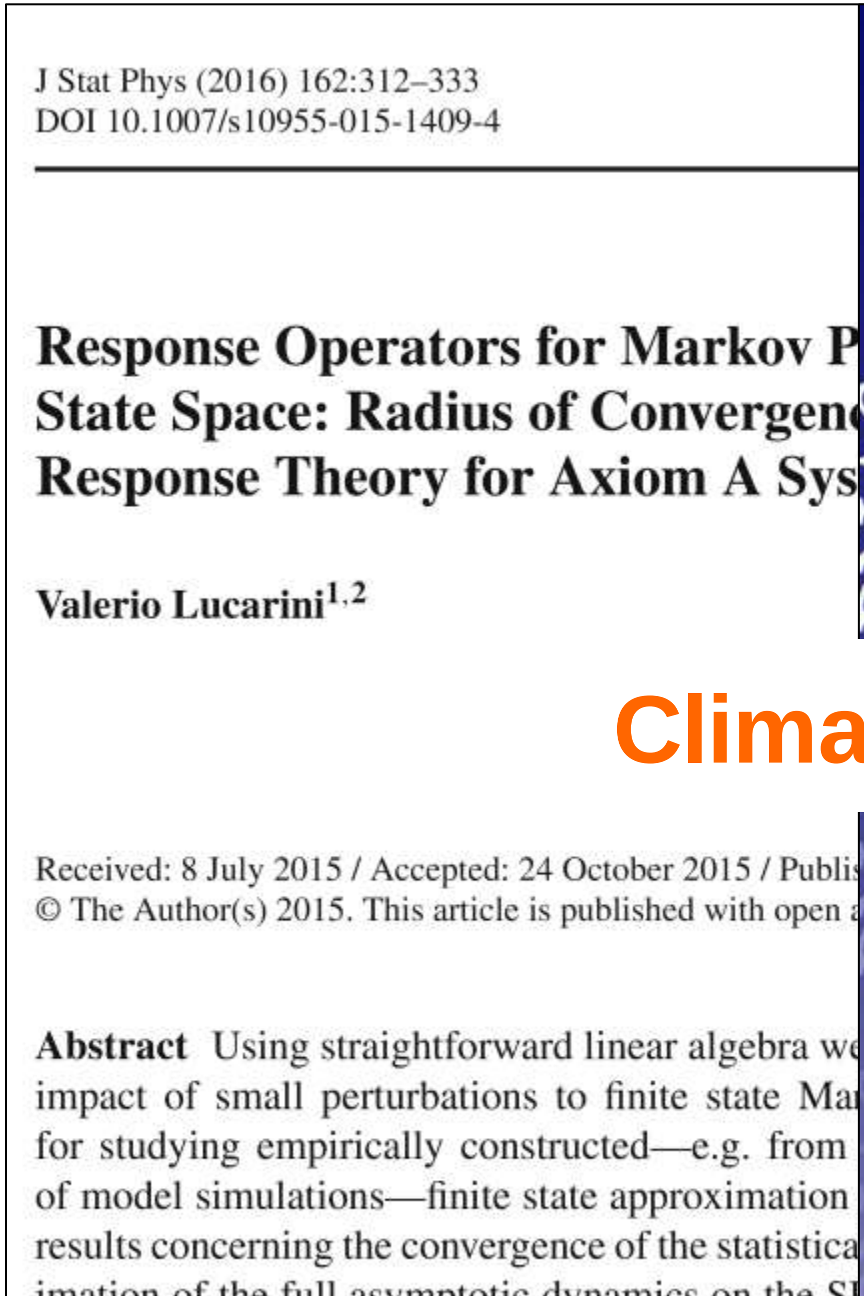
Renormalizing Entanglem

Stephan Waeldchen,[1] Janina Gertis,[1] Earl
Center for Complex Quantum Systems, Freie
ent of Physics and Astronomy, University of Sh
(Received 2 May 2015; publishe

Entanglement distillation refers to the task of transformi
fewer highly entangled ones. It is a core ingredient in qua
transmit entanglement over arbitrary distances in order t
Usually, it is assumed that the initial entangled pairs are ide
uncorrelated with each other, an assumption that might n
generation process involving memory channels. Here, we i
ment distillation in the presence of natural correlations aris
bring together ideas from condensed-matter physics—ideas f
and operators—with those of local entanglement manipulat
correction. We identify meaningful parameter regions for

Quantum physics

Recent applications of the theory


J Stat Phys (2016) 162:312–333
DOI 10.1007/s10955-015-1409-4

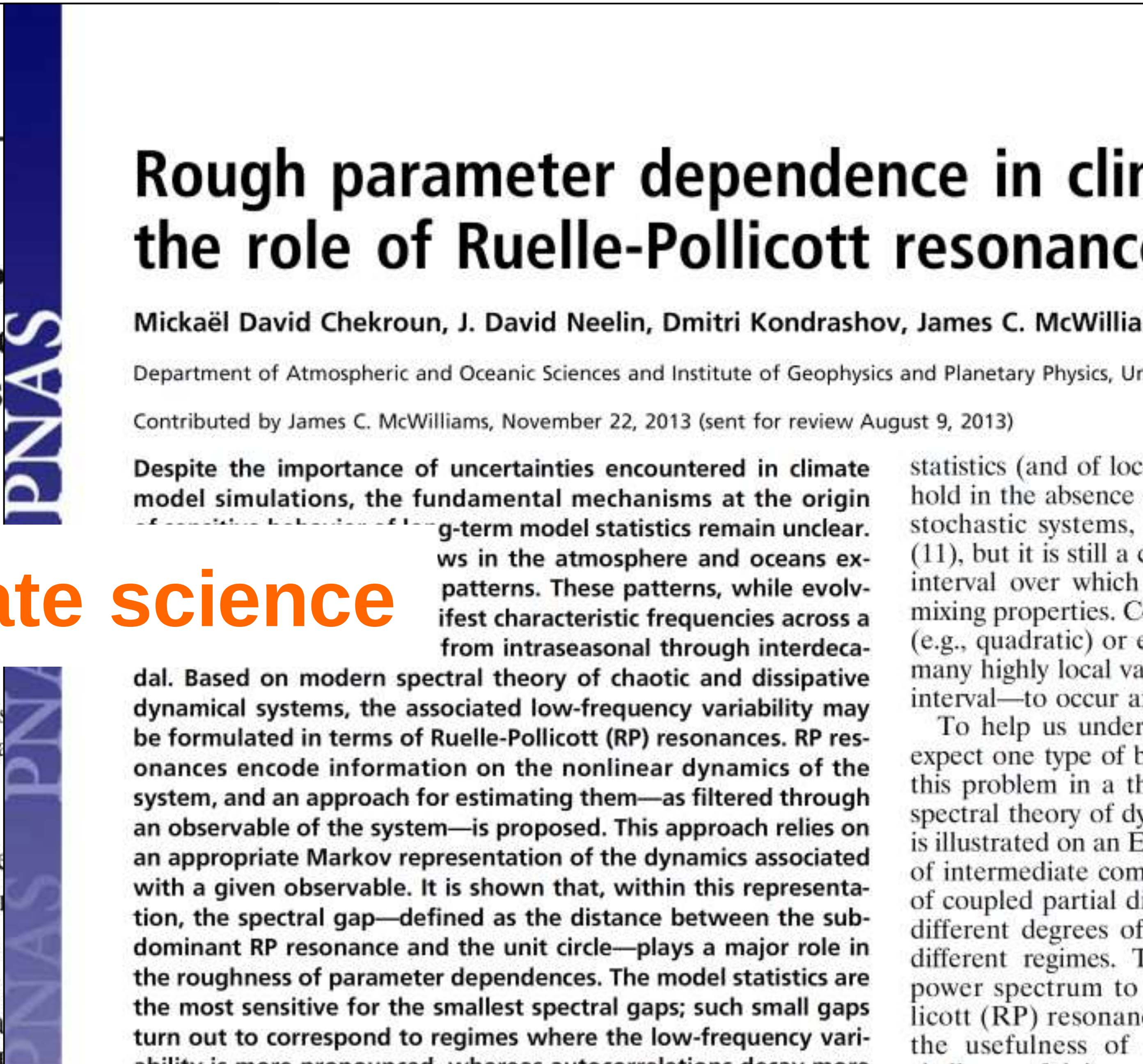
Response Operators for Markov P
State Space: Radius of Convergen
Response Theory for Axiom A Sys

Valerio Lucarini[1,2]

Received: 8 July 2015 / Accepted: 24 October 2015 / Publis
© The Author(s) 2015. This article is published with open a

Abstract Using straightforward linear algebra we
impact of small perturbations to finite state Ma
for studying empirically constructed—e.g. from
of model simulations—finite state approximation
results concerning the convergence of the statistica

Climate science


PNAS

Rough parameter dependence in cli
the role of Ruelle-Pollicott resonanc

Mickaël David Chekroun, J. David Neelin, Dmitri Kondrashov, James C. McWillia

Department of Atmospheric and Oceanic Sciences and Institute of Geophysics and Planetary Physics, Ur

Contributed by James C. McWilliams, November 22, 2013 (sent for review August 9, 2013)

Despite the importance of uncertainties encountered in climate model simulations, the fundamental mechanisms at the origin ...g-term model statistics remain unclear. ...ws in the atmosphere and oceans ex- ...patterns. These patterns, while evolv- ...ifest characteristic frequencies across a ...from intraseasonal through interdeca- dal. Based on modern spectral theory of chaotic and dissipative dynamical systems, the associated low-frequency variability may be formulated in terms of Ruelle-Pollicott (RP) resonances. RP resonances encode information on the nonlinear dynamics of the system, and an approach for estimating them—as filtered through an observable of the system—is proposed. This approach relies on an appropriate Markov representation of the dynamics associated with a given observable. It is shown that, within this representation, the spectral gap—defined as the distance between the subdominant RP resonance and the unit circle—plays a major role in the roughness of parameter dependences. The model statistics are the most sensitive for the smallest spectral gaps; such small gaps turn out to correspond to regimes where the low-frequency vari-

statistics (and of loc
hold in the absence
stochastic systems,
(11), but it is still a
interval over which
mixing properties. C
(e.g., quadratic) or
many highly local va
interval—to occur a
To help us under
expect one type of b
this problem in a th
spectral theory of dy
is illustrated on an E
of intermediate com
of coupled partial d
different degrees of
different regimes. T
power spectrum to
licott (RP) resonan
the usefulness of

Recent applications of the theory

J Stat Phys (2016) 162:312–333
DOI 10.1007/s10955-015-1409-4

Response
State Spac
Response

Valerio Lucar

Received: 8 July 2015 / Accepted: 24 October 2015 / Publis

Abstract Using straightforward linear algebra we
impact of small perturbations to finite state Ma
for studying empirically constructed—e.g. from
of model simulations—finite state approximation
results concerning the convergence of the statistica

Biology??

- Molecular dynamics simulations
- Ion channels
- General biochemical kinetics

…

dependence in cli
-Pollicott resonanc

n, Dmitri Kondrashov, James C. McWillia
and Institute of Geophysics and Planetary Physics, Ur

, 2013 (sent for review August 9, 2013)

countered in climate
anisms at the origin
tistics remain unclear.
phere and oceans ex-
patterns, while evolv-
c frequencies across a
al through interdeca-
haotic and dissipative
dynamical systems, the associated low-frequency variability may be formulated in terms of Ruelle-Pollicott (RP) resonances. RP resonances encode information on the nonlinear dynamics of the system, and an approach for estimating them—as filtered through an observable of the system—is proposed. This approach relies on an appropriate Markov representation of the dynamics associated with a given observable. It is shown that, within this representation, the spectral gap—defined as the distance between the subdominant RP resonance and the unit circle—plays a major role in the roughness of parameter dependences. The model statistics are the most sensitive for the smallest spectral gaps; such small gaps turn out to correspond to regimes where the low-frequency vari-

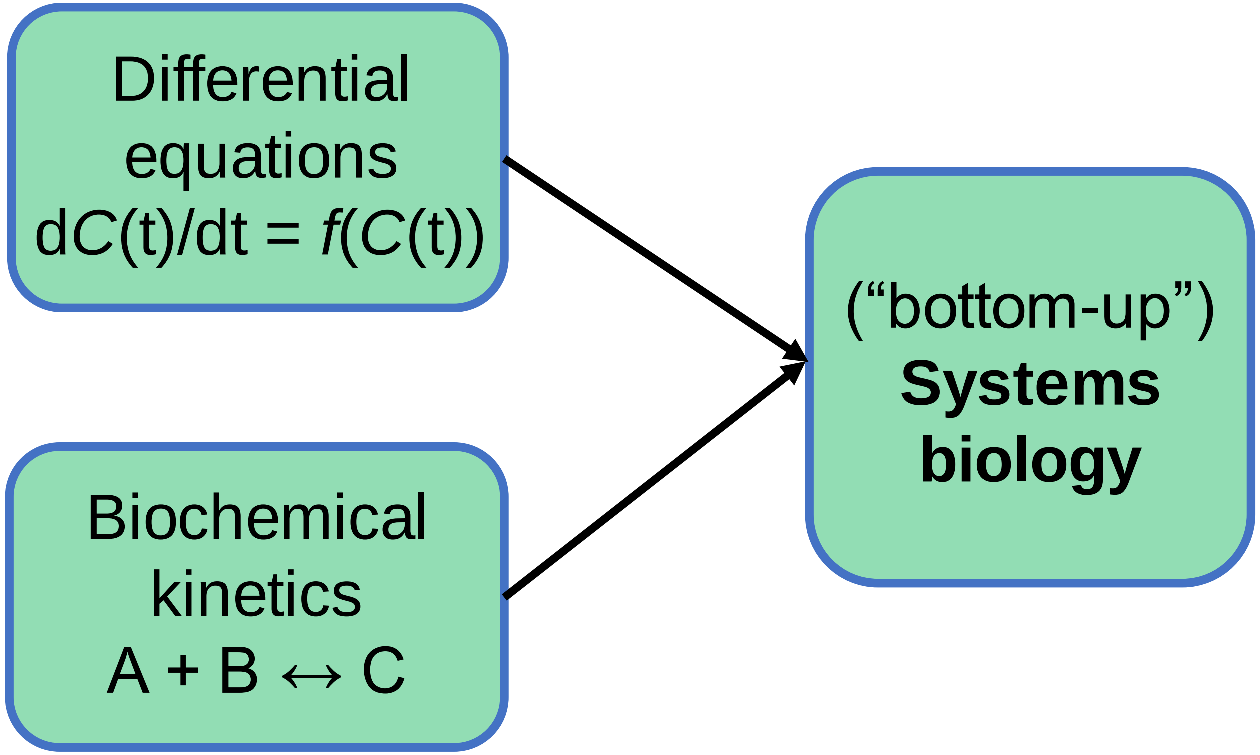
statistics (and of loc
hold in the absence
stochastic systems,
(11), but it is still a c
interval over which
mixing properties. C
(e.g., quadratic) or e
many highly local va
interval—to occur a
To help us under
expect one type of b
this problem in a th
spectral theory of dy
is illustrated on an E
of intermediate com
of coupled partial d
different degrees of
different regimes. T
power spectrum to
licott (RP) resonan
the usefulness of

Next chapter: systems biology

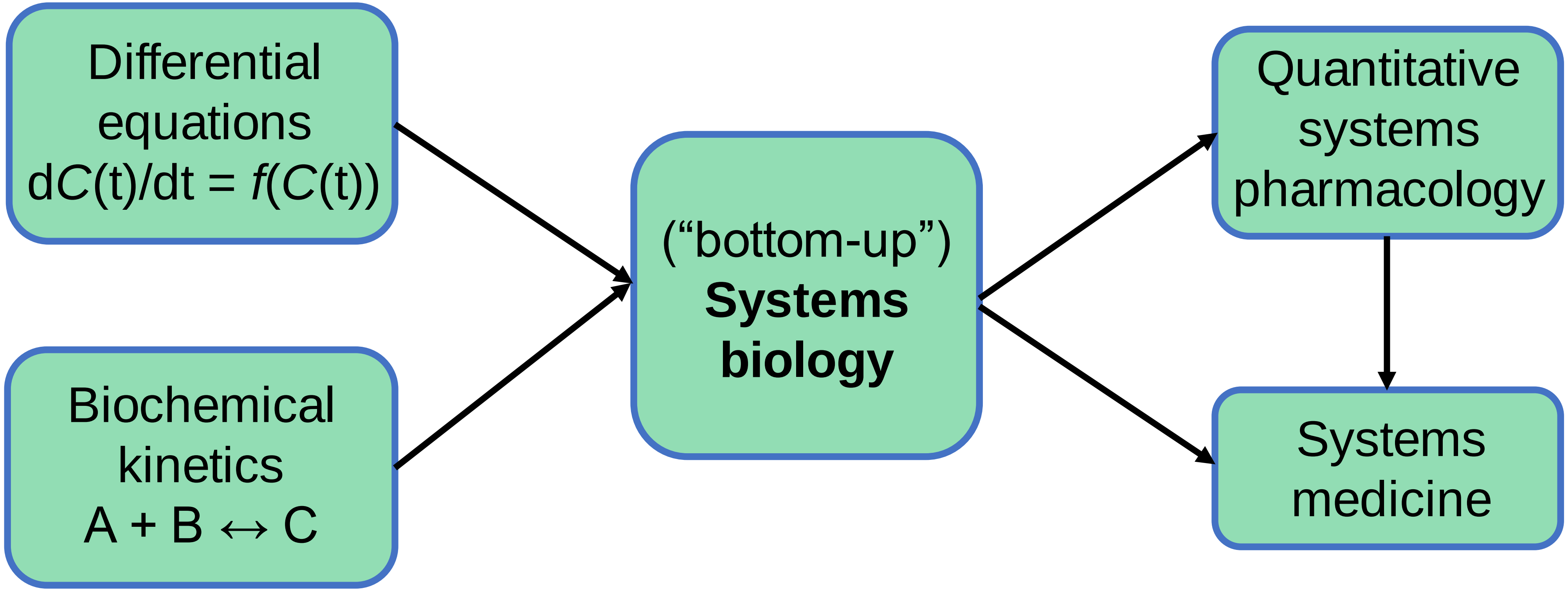
Differential equations
$dC(t)/dt = f(C(t))$

Biochemical kinetics
$A + B \leftrightarrow C$

Next chapter: systems biology



Next chapter: systems biology



Bacterial signal transduction: two-component systems

Signal

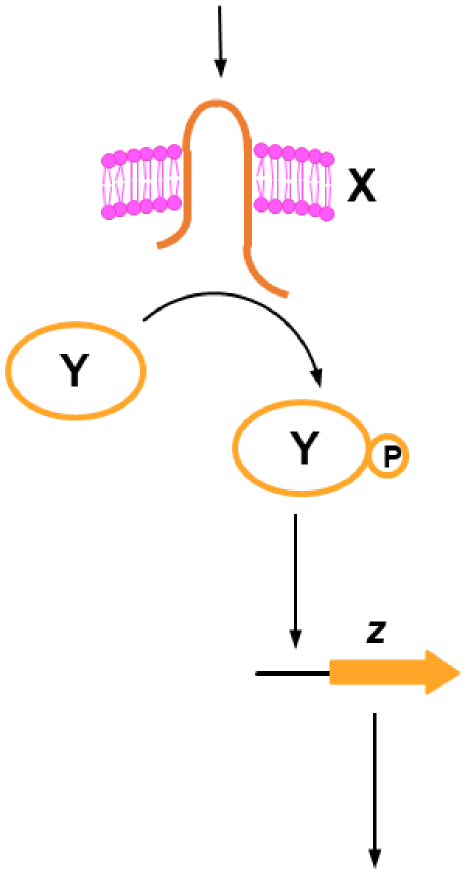
Sensor

Regulatory circuit

Effectors

Response

Metal ions, small molecules, pH, etc.

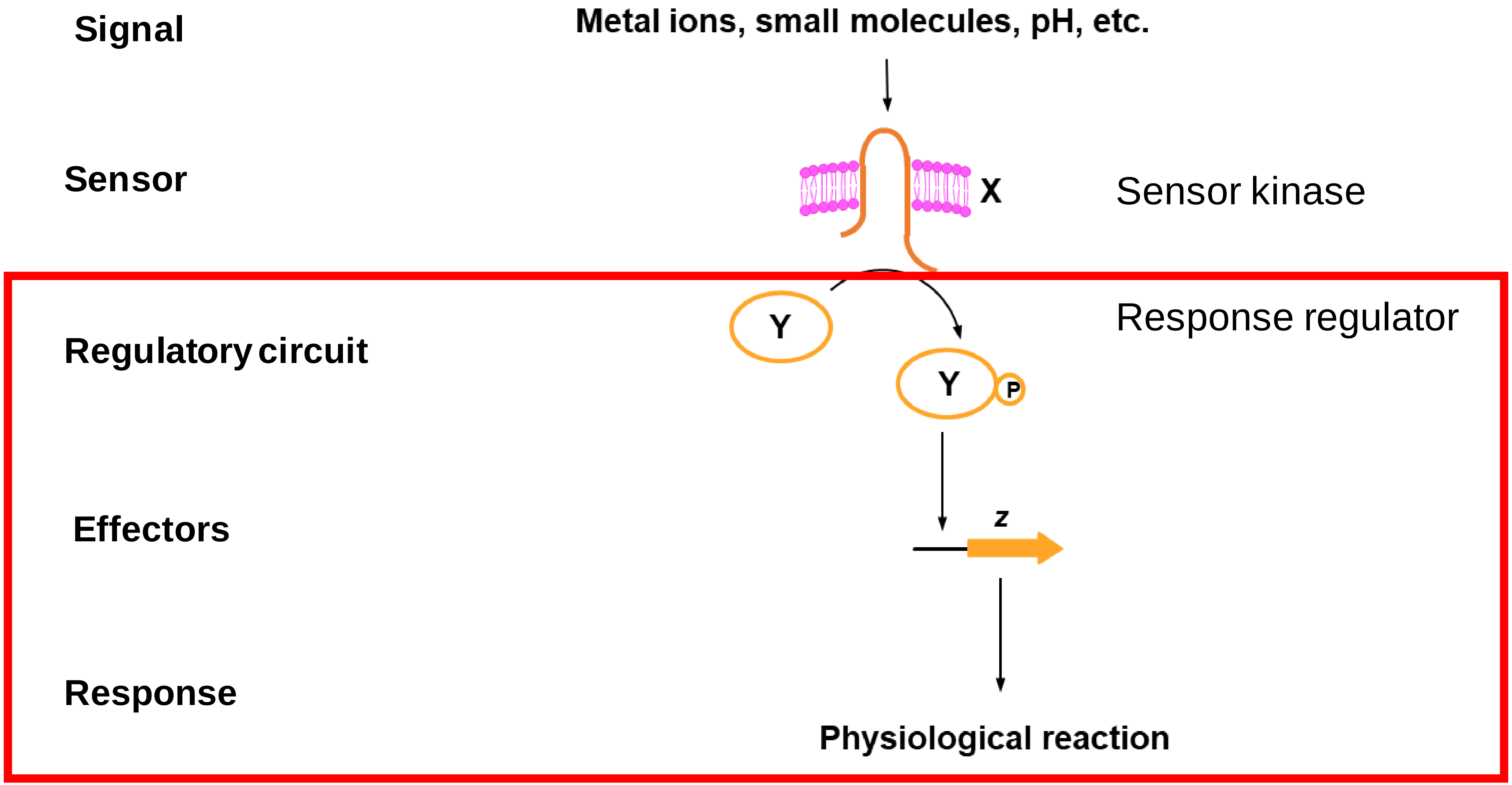


Sensor kinase

Response regulator

Physiological reaction

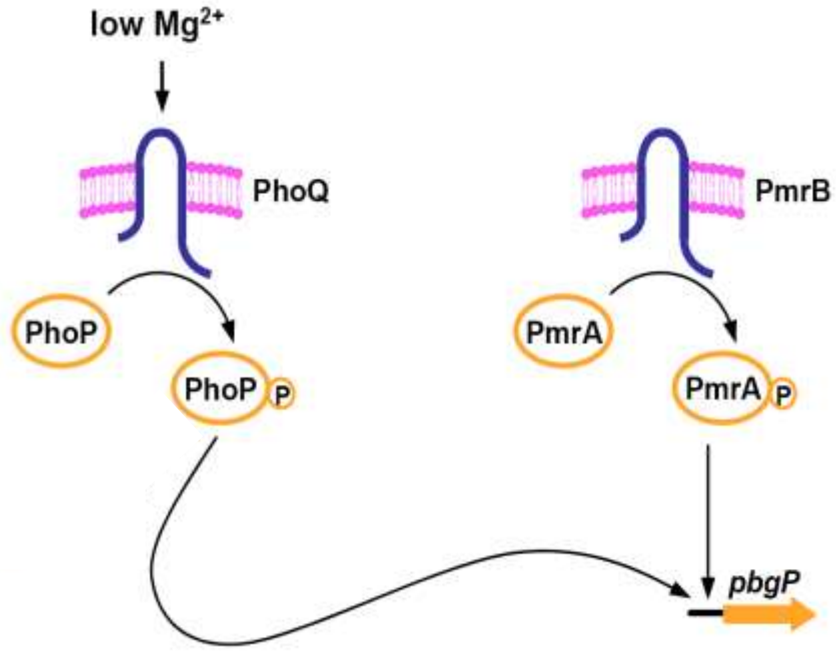
Bacterial signal transduction: two-component systems



Different species of enteric bacteria use distinct architectures to activate *pbgP* by low Mg^{2+}

Direct pathway

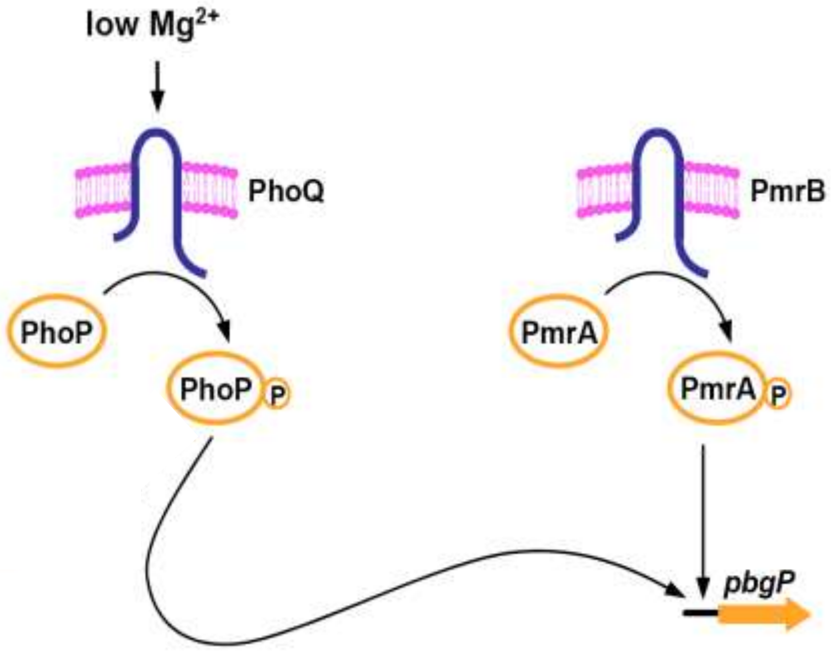
Yersinia pestis



Different species of enteric bacteria use distinct architectures to activate *pbgP* by low Mg^{2+}

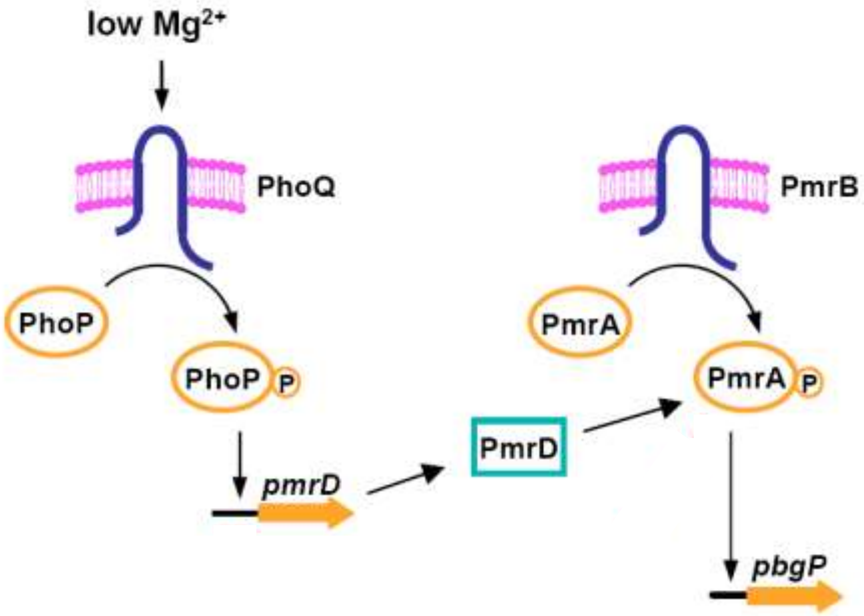
Direct pathway

Yersinia pestis

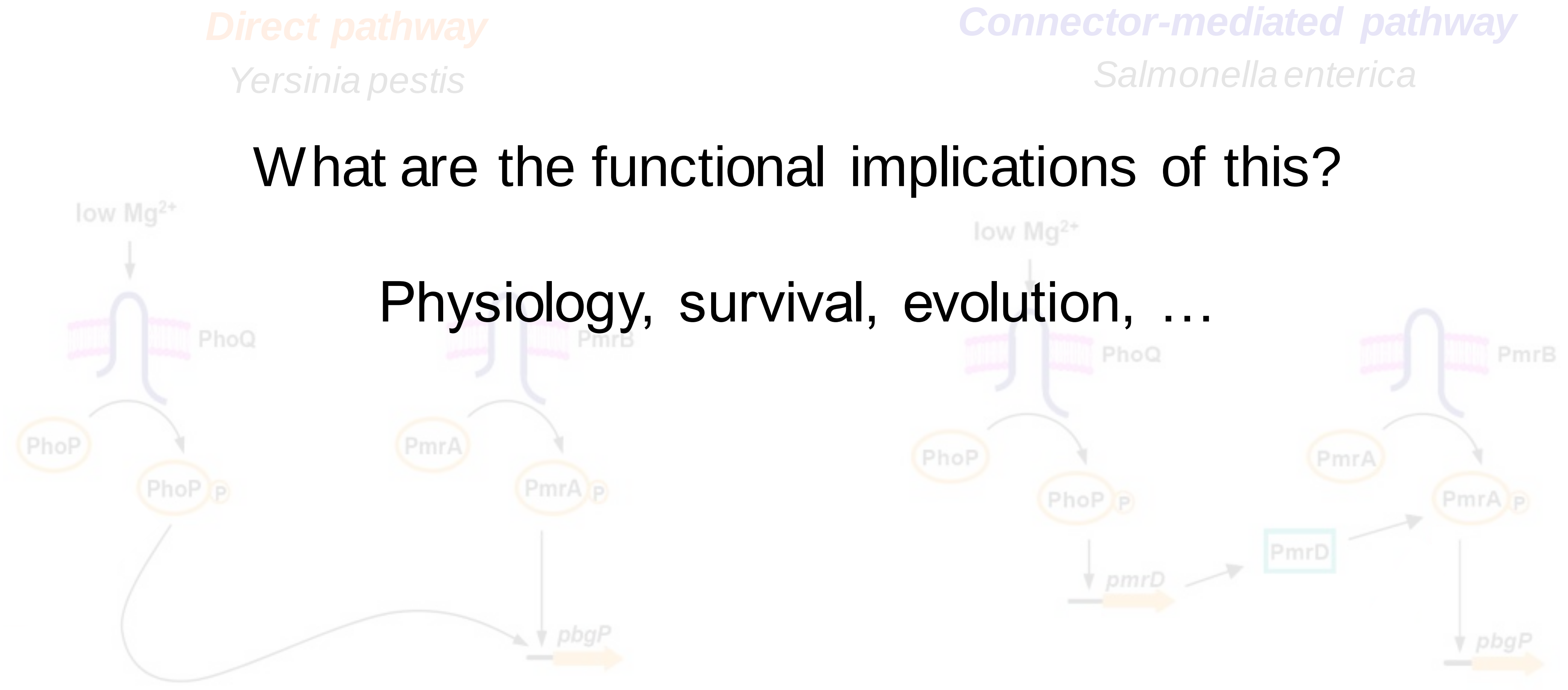


Connector-mediated pathway

Salmonella enterica



Different species of enteric bacteria use distinct architectures to activate *pbgP* by low Mg^{2+}

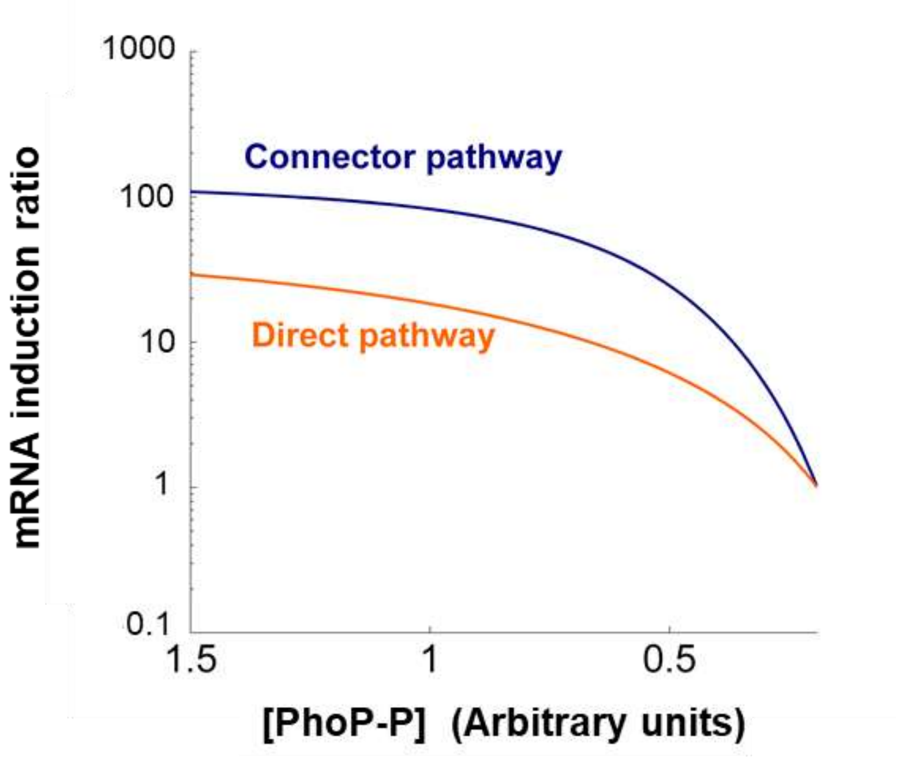


What are the functional implications of this?

Physiology, survival, evolution, …

Connector-mediated pathway promotes signal amplification

Computation

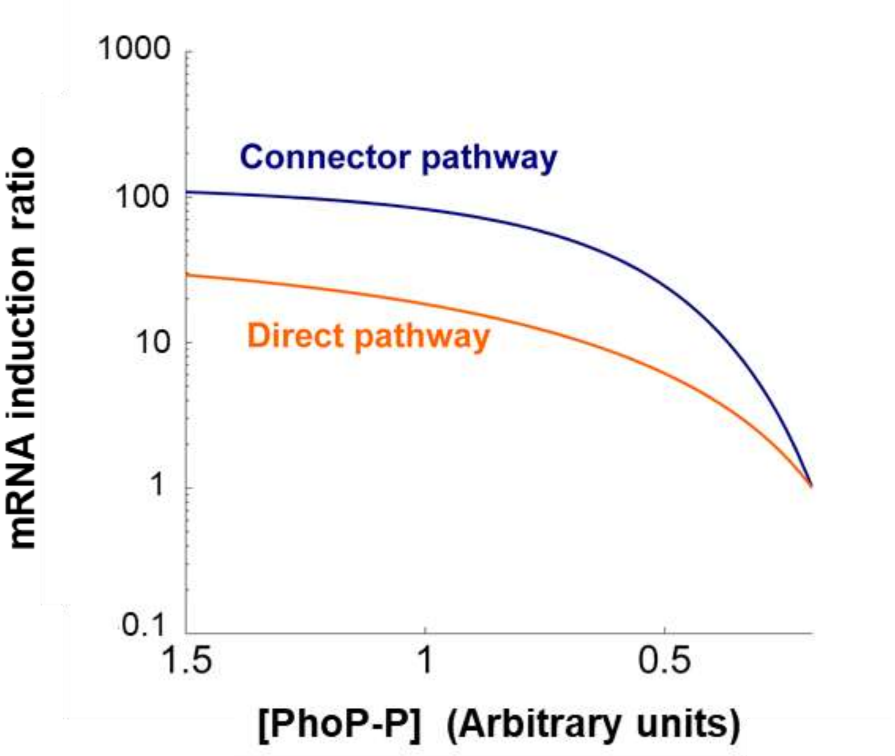


Induction ratio = mRNA level (inducing conditions) / mRNA level (repressing conditions)

Kato, Mitrophanov, Groisman, *PNAS* (2007)

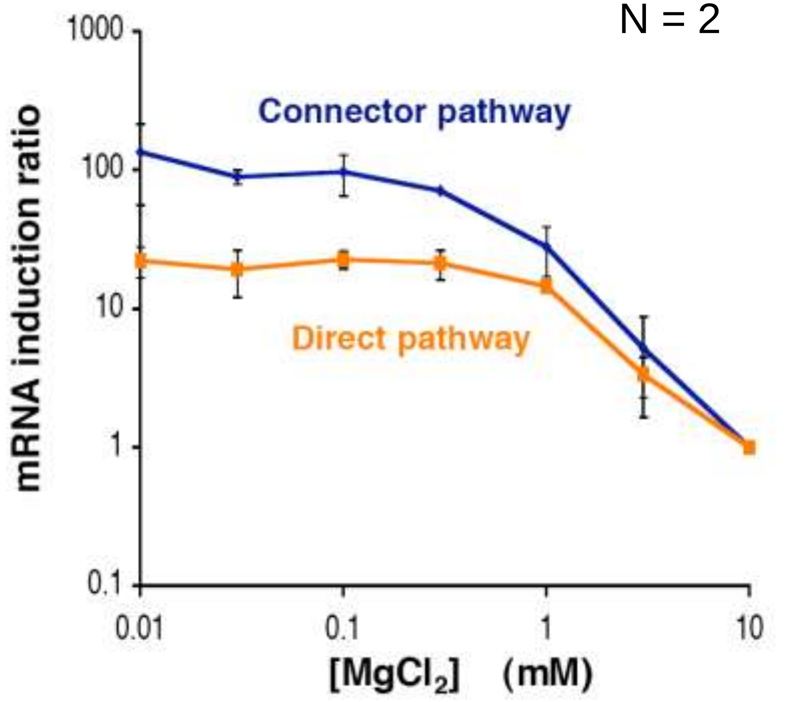
Connector-mediated pathway promotes signal amplification

Computation



Experiment

N = 2

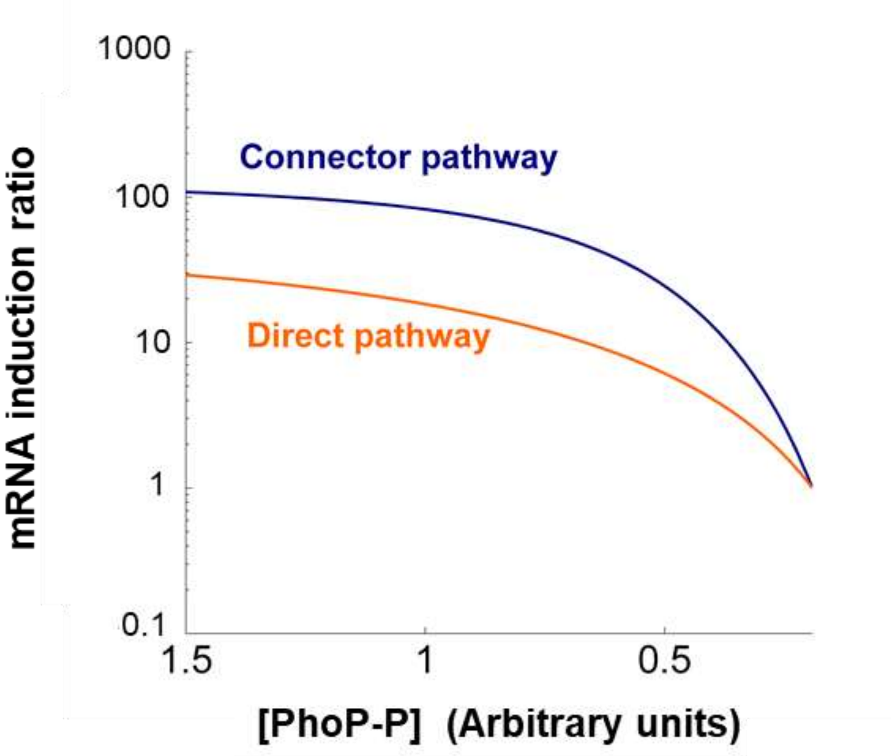


Induction ratio = mRNA level (inducing conditions) / mRNA level (repressing conditions)

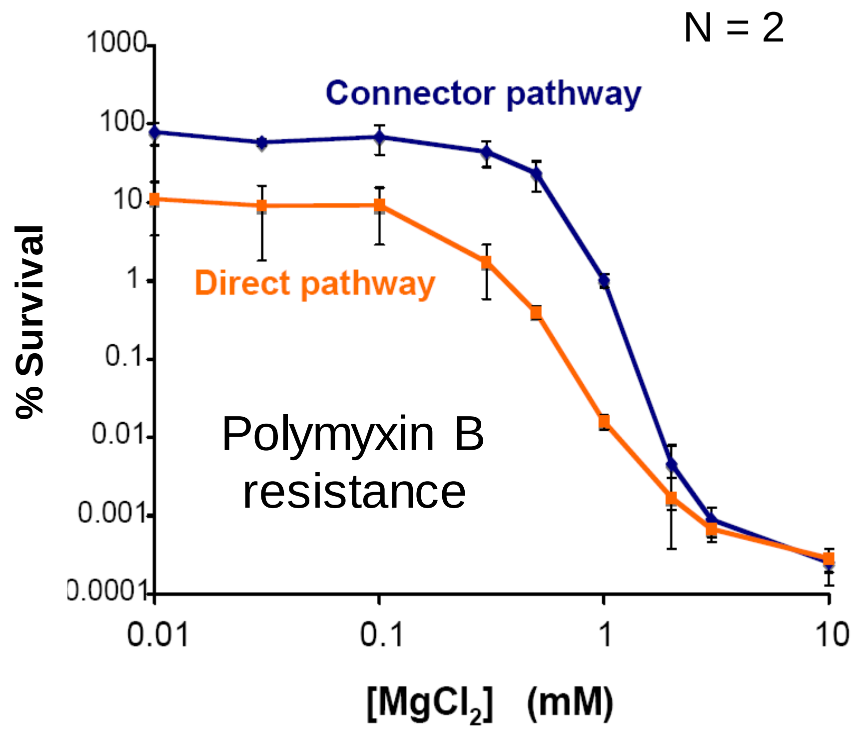
Kato, Mitrophanov, Groisman, *PNAS* (2007)

Connector-mediated pathway promotes signal amplification

Computation



Experiment

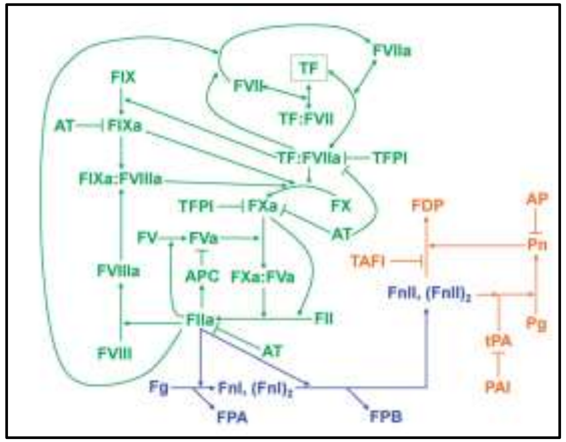


Induction ratio = mRNA level (inducing conditions) / mRNA level (repressing conditions)

Kato, Mitrophanov, Groisman, *PNAS* (2007)

Blood coagulation and traumatic coagulopathy

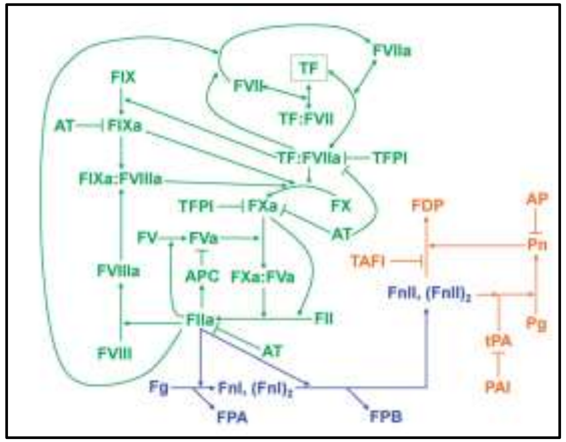
Blood coagulation system



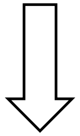
Mitrophanov et al., *Mol Biosyst* (2014)

Blood coagulation and traumatic coagulopathy

Blood coagulation system



Mitrophanov et al., *Mol Biosyst* (2014)



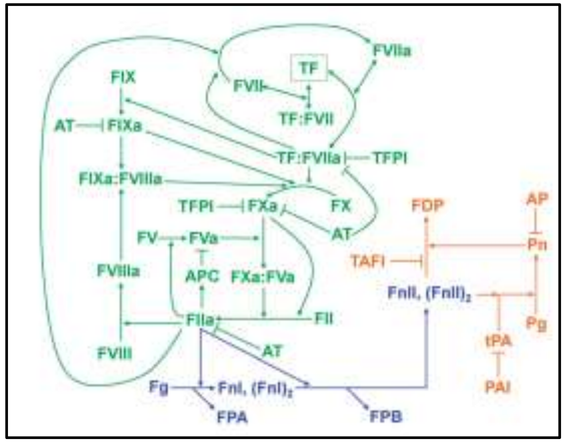
Computational kinetic modeling

C(t) = species concentration;

***d*C(t)/*d*t = (production rate)**
− (depletion rate)

Blood coagulation and traumatic coagulopathy

Blood coagulation system



Mitrophanov et al., *Mol Biosyst* (2014)

In vitro experiments (blood plasma)

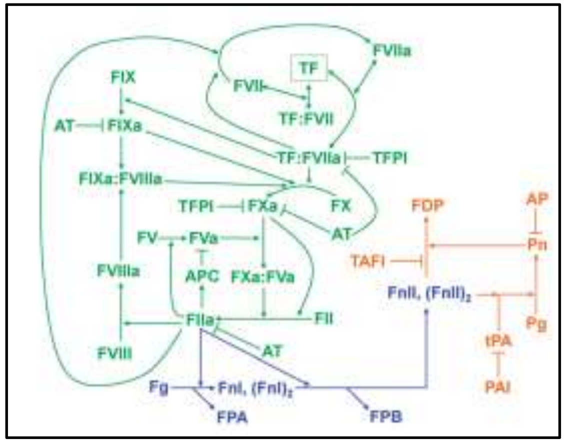
Computational kinetic modeling

C(t) = species concentration;

$dC(t)/dt$ = (production rate) − (depletion rate)

Blood coagulation and traumatic coagulopathy

Blood coagulation system



Mitrophanov et al., *Mol Biosyst* (2014)

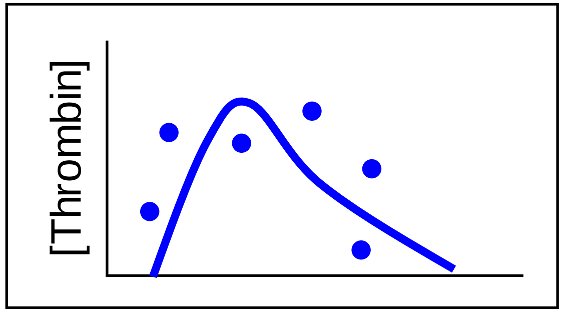
In vitro experiments (blood plasma)

Computational kinetic modeling

C(t) = species concentration;

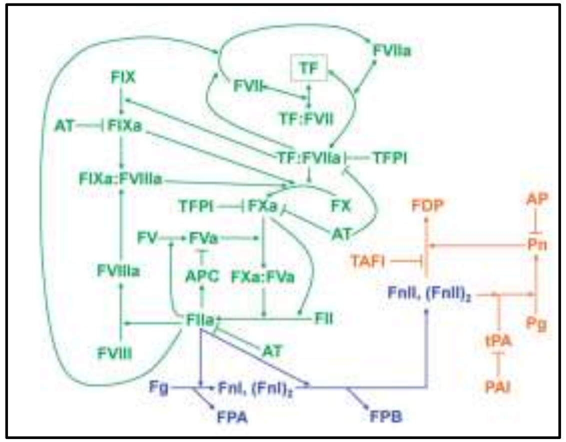
***d*C(t)/*d*t = (production rate) − (depletion rate)**

Comparison: validation



Blood coagulation and traumatic coagulopathy

Blood coagulation system



Mitrophanov et al., *Mol Biosyst* (2014)

In vitro experiments (blood plasma)

coronapathology.org

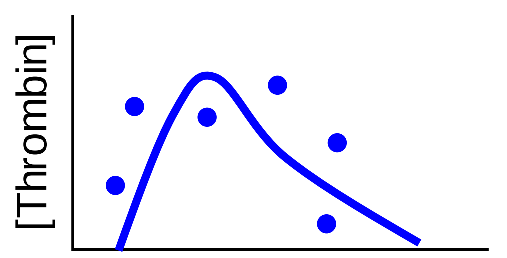
Learning

Computational kinetic modeling

C(t) = species concentration;

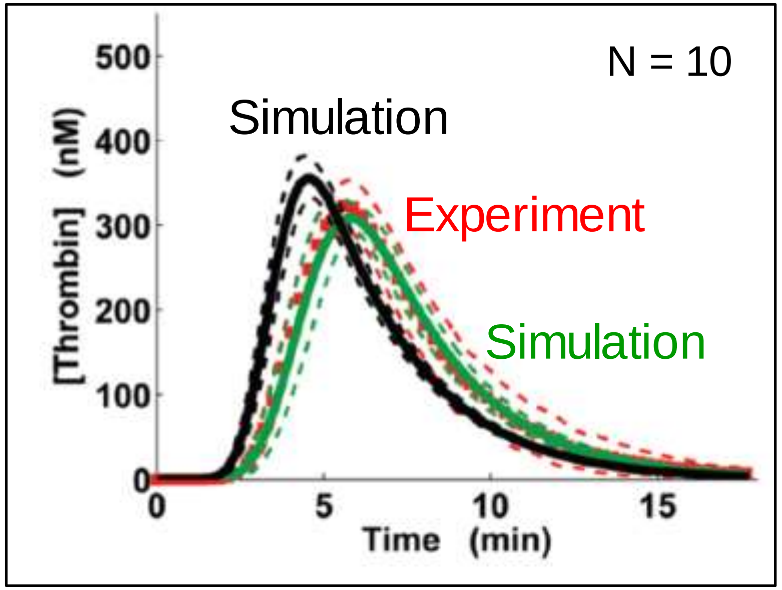
***d*C(t)/*d*t = (production rate) − (depletion rate)**

Comparison: validation



Thrombin generation in blood plasma

Normal plasma

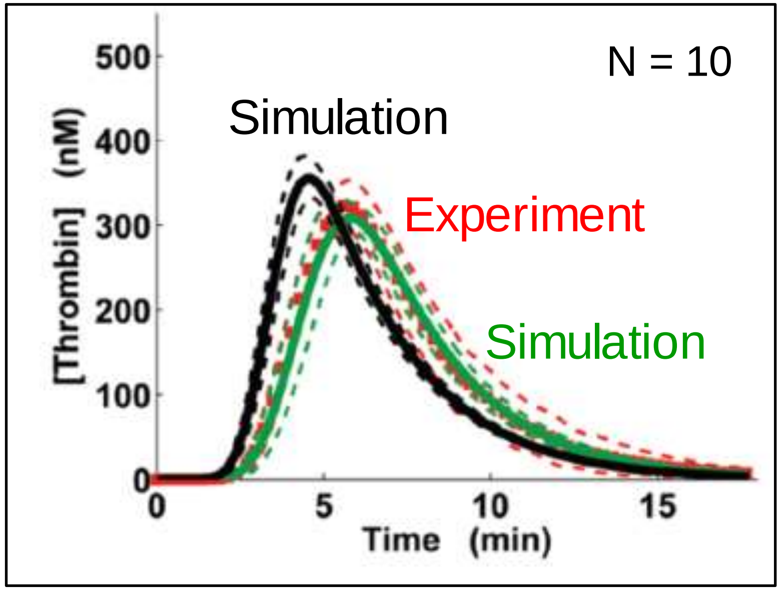


prothrombin → thrombin ↷ fibrinogen
fibrin → blood clot

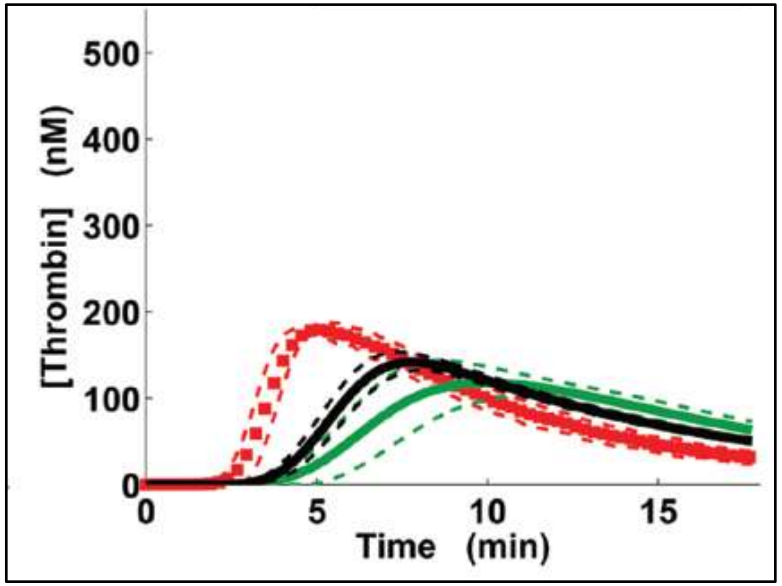
Mitrophanov et al., *Anesth Analg* (2016)

Thrombin generation in blood plasma

Normal plasma



3-fold dilution



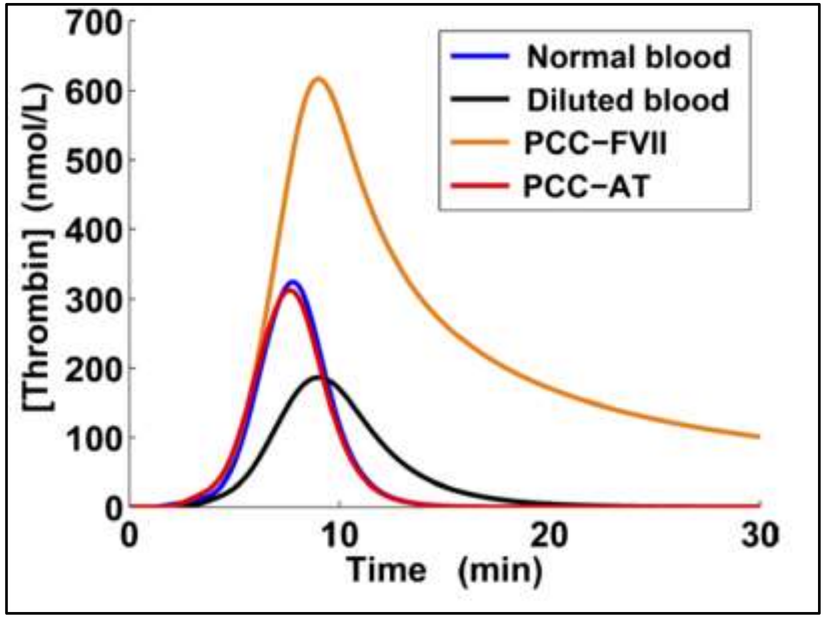
Dilution reduces peak height

prothrombin → thrombin ↷ fibrinogen
fibrin → blood clot

Mitrophanov et al., *Anesth Analg* (2016)

Restoring reduced thrombin generation in plasma

Simulations*

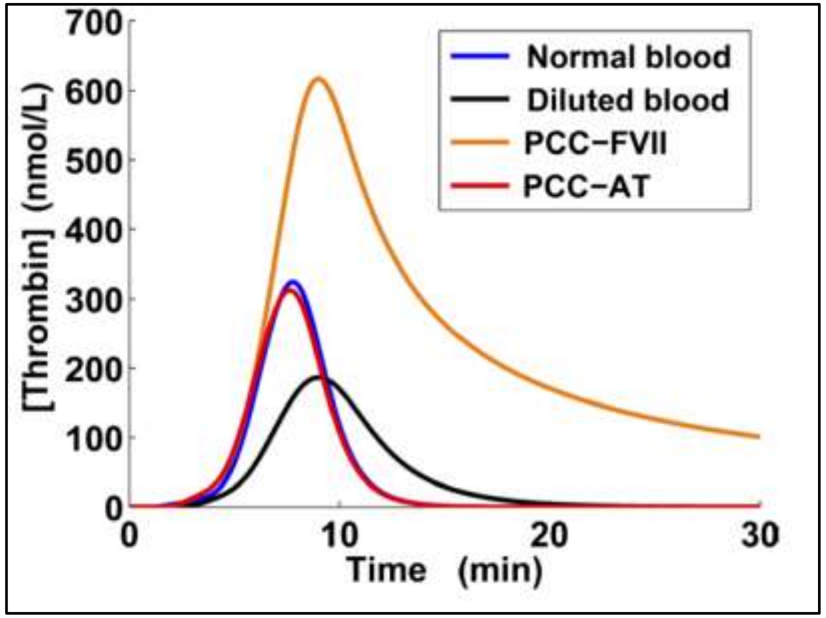


PCC-FVII = FII + FIX + FX + FVII	strong procoagulants
PCC-AT = FII + FIX + FX + antithrombin	procoagulants + anticoagulant

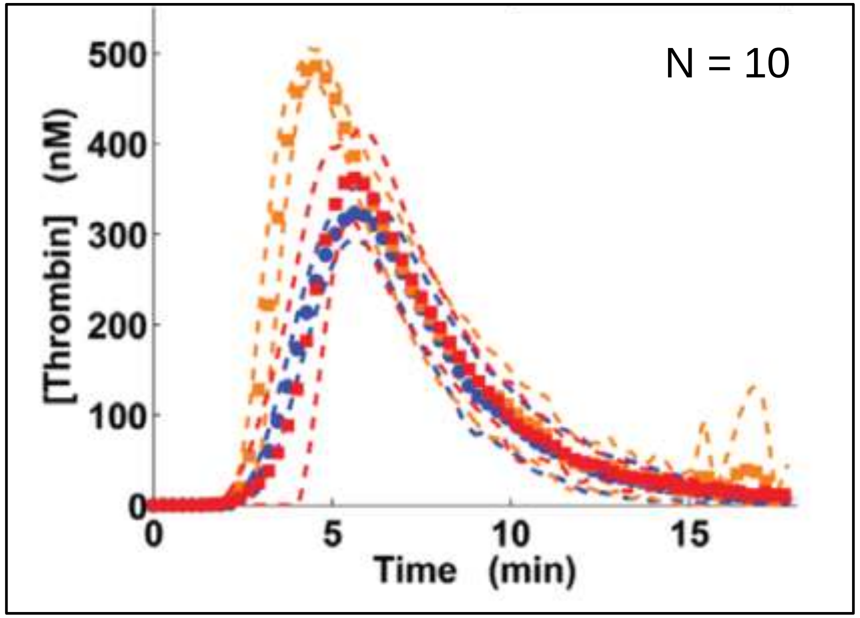
*Mitrophanov et al., *J Trauma* (2012)

Restoring reduced thrombin generation in plasma

Simulations*



Experimental data**

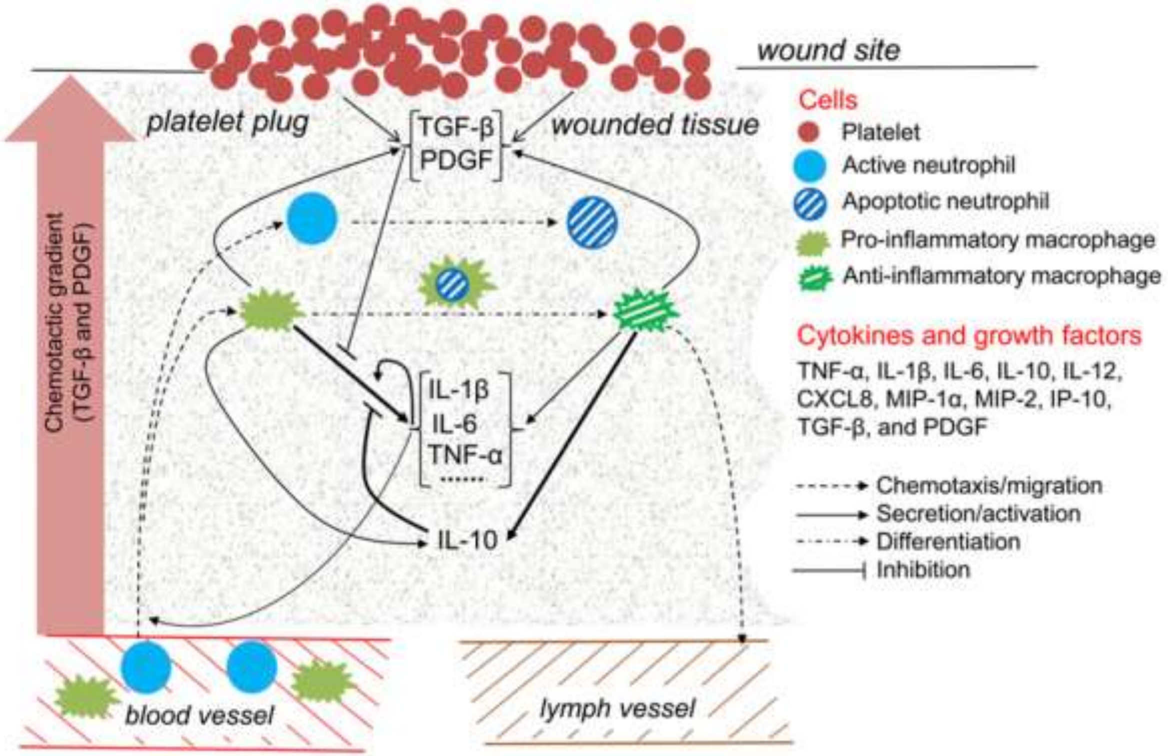


PCC-FVII	= FII + FIX + FX + FVII	strong procoagulants
PCC-AT	= FII + FIX + FX + antithrombin	procoagulants + anticoagulant

*Mitrophanov et al., *J Trauma* (2012)

**Mitrophanov et al., *Anesth Analg* (2016)

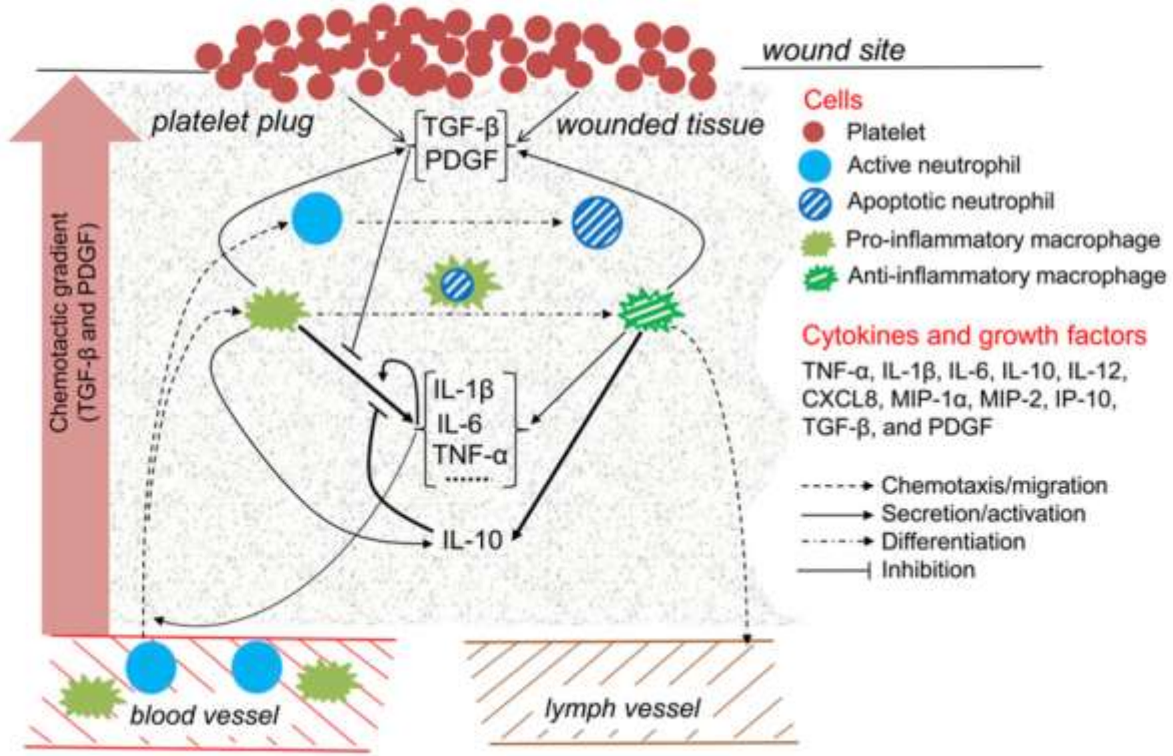
Wound-healing research



Nagaraja et al., *J Immunol* (2014)

- Inflammation
- Proliferation
- Angiogenesis

Wound-healing research



Nagaraja et al., *J Immunol* (2014)

1000s of model simulations

Predicted **biomarkers** and **drug targets** (e.g., for chronic inflammation)

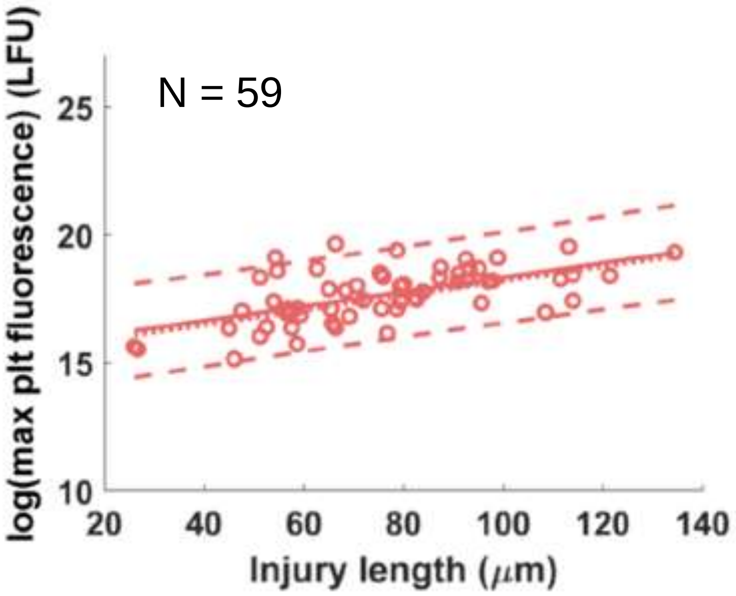
- Inflammation
- Proliferation
- Angiogenesis

Some findings

- **IL-6** as biomarker of chronic inflammation
- Unique influence of **TGF-β** throughout wound healing

From systems biology to (biomedical) data science

NOT just a simple regression!

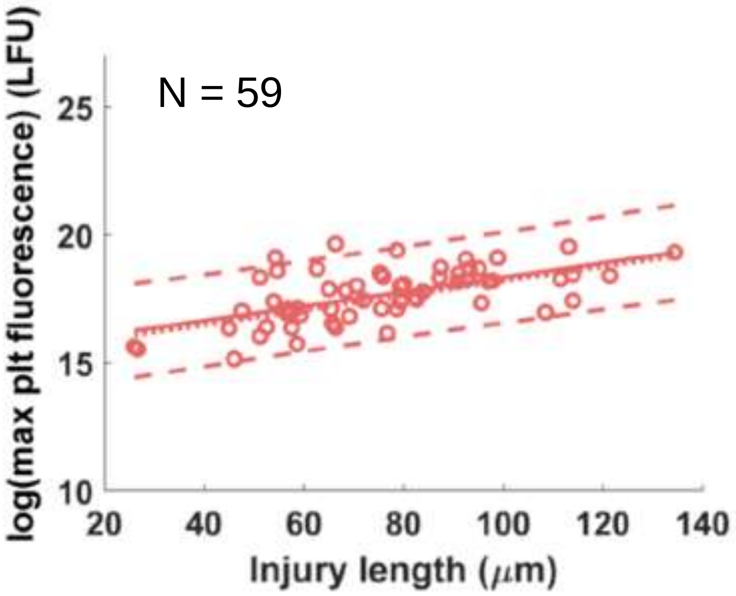


- Statistical model selection
- Variable selection

Mitrophanov et al., *Arterioscler Thromb Vasc Biol* (2020)

From systems biology to (biomedical) data science

NOT just a simple regression!



My data science:

The knowledge and understanding of robust patterns and relationships between variables in data sets.

- Statistical model selection
- Variable selection

Mitrophanov et al., *Arterioscler Thromb Vasc Biol* (2020)

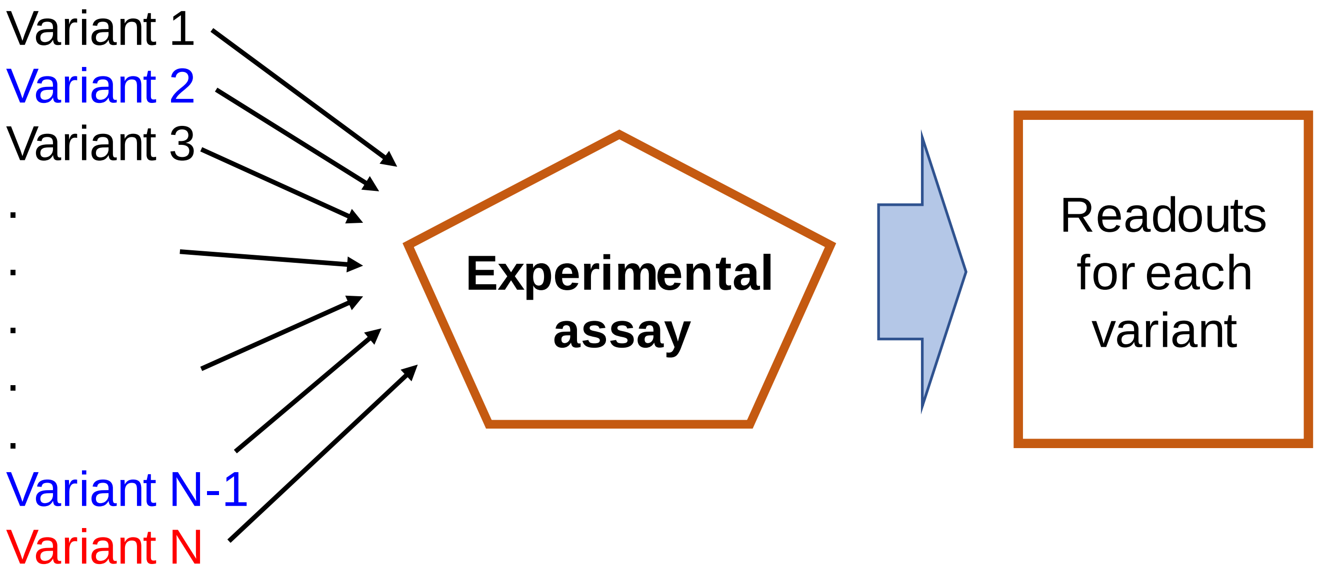
Current work

Mutation pathogenicity annotation in BRCA2-oncogene variants

Joint work with Kajal Biswas, Shyam Sharan, and Tyler Malys

Current work

Mutation pathogenicity annotation in BRCA2-oncogene variants

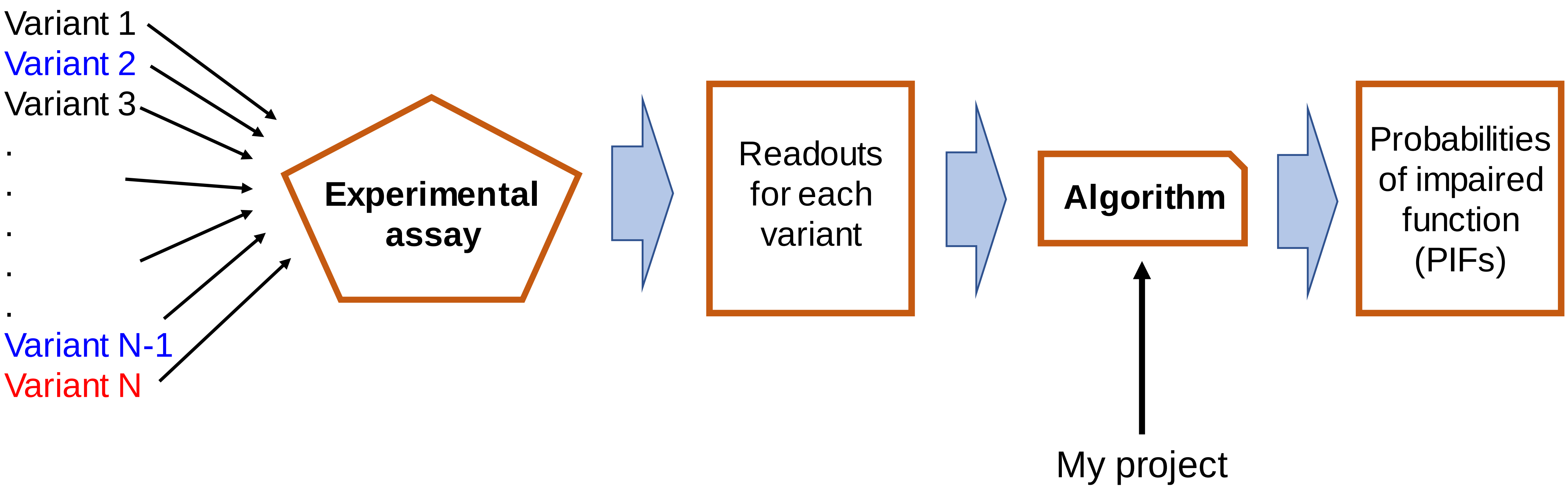


Red: pathogenic variant
Blue: neutral variant
Black: status unknown

Joint work with Kajal Biswas, Shyam Sharan, and Tyler Malys

Current work

Mutation pathogenicity annotation in BRCA2-oncogene variants



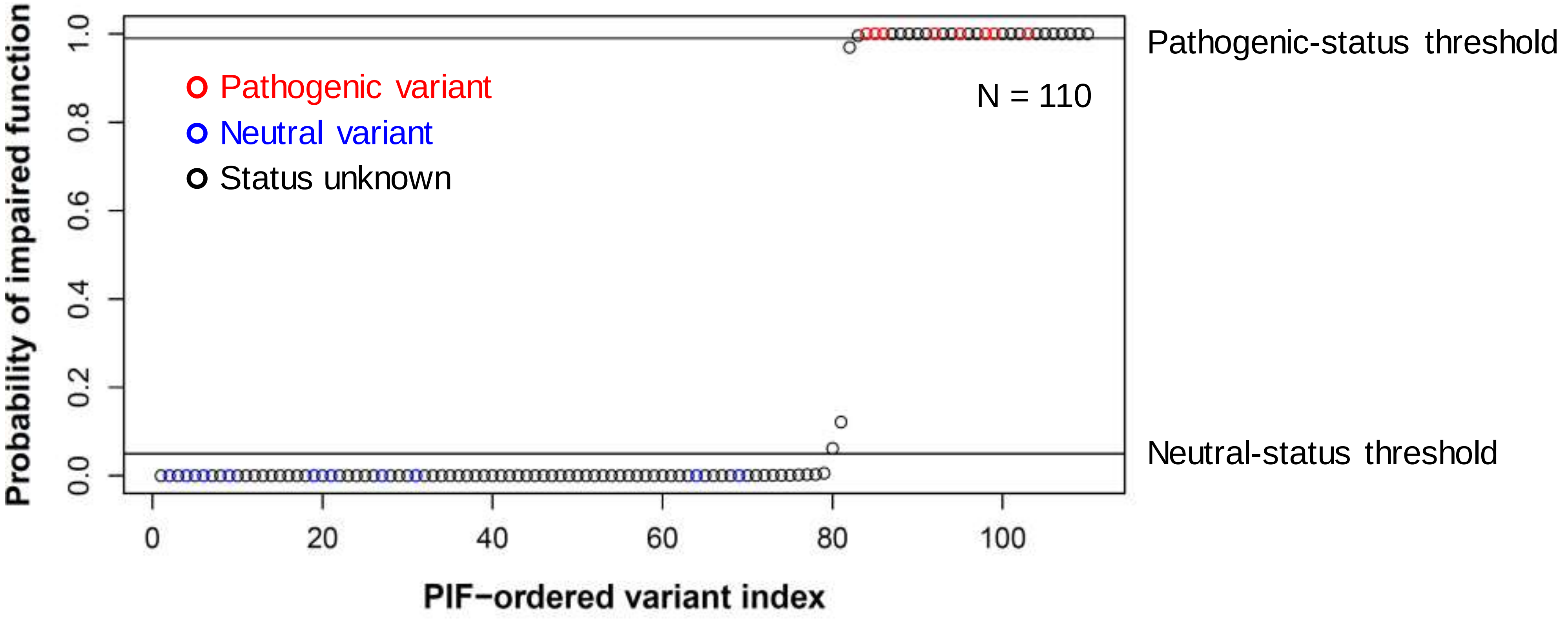
Red: pathogenic variant
Blue: neutral variant
Black: status unknown

Joint work with Kajal Biswas, Shyam Sharan, and Tyler Malys

Current work

Mutation pathogenicity annotation in BRCA2-oncogene variants

Approach: statistical mixture modeling; semi-supervised learning



Joint work with Kajal Biswas, Shyam Sharan, and Tyler Malys

THANKS!!!