QBIO 2025 CONFERENCE:

Emergent Orders in Living Systems Across Scales

Monday 21 July 2025 – Thursday 24 July 2025

Peking University, China

Hosts

- Center for Quantitative Biology (CQB) at Peking University (https://cqb.pku.edu.cn/)
- Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Sciences (CAS) (https://www.siat.ac.cn/)
- Institute of Theoretical Physics (ITP), Chinese Academy of Sciences (CAS) (https://itp.cas.cn/)
- Westlake University (https://westlake.edu.cn/)

Local Organizers:

Chao Tang (Peking University, China)
Jianhua Xing (University of Pittsburgh, USA)
Zhiyuan Li (Peking University, China)
Jie Lin (Peking University, China)
Weikang Wang (Institute of Theoretical Physics, CAS, China)
Lei Dai (Shenzhen Institutes of Advanced Technology, CAS, China)
Ping Wei (Shenzhen Institutes of Advanced Technology, CAS, China)
Yifan Yang (Westlake University)

Contact:

Email: qbio@pku.edu.cn
Tel: +86-10-62759599
Fax: +86-10-62759595

Conference Website:

https://qbio.pku.edu.cn/QB2025/

Supported by:

北京大学 PEKING UNIVERSITY 研究生院

Venue Address:

Bldg 1#，Zhongguanyuan Global Village PKU（北京大学中关新园 1 号楼群英厅）

Address: No.126, Zhongguancun North Street, Haidian District, Beijing, China（北京市海淀区中关村北大街 126 号）

Hotel website: https://pkugv.pku.edu.cn/

Hotel Tel: +86-10-62752288

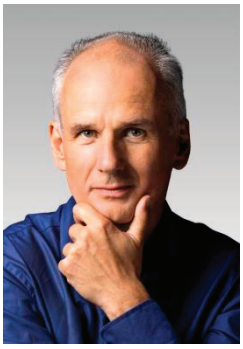
July 21 Registration 13: 30-18: 00 Bldg 1#, Zhongguanyuan Global Village PKU（中关新园1号楼大厅）
July 22 Registration 08: 20-08: 50 Bldg 1#, Zhongguanyuan Global Village PKU（中关新园1号楼群英厅）

2025/7/22 (Tuesday)		
Time	**Speaker**	**Title**
9:00-9:10	**Opening**	
Session 1 (Chair: Jing-Dong Jackie Han)		
9:10-9:55	**Nikolaus Rajewsky** Max Delbrück Center for Molecular Medicine - Berlin, Germany	Spatial Omics to Predict Disease Trajectories
9:55-10:30	**Jin Wang** Stony Brook University, USA	Uncovering Underlying Physical Principles and Driving Forces of Cancer from Single-cell Transcriptomics
10:30-11:00	**Group Photo/Coffee Break**	
Session 2 (Chair: Weikang Wang)		
11:00-11:45	**Jianhua Xing** University of Pittsburgh, USA	Dynamical Systems Theory Modeling in the Big Data Era
11:45-12:00	**Jin Yu** University of California, Irvine, USA	Modeling, Simulation, and Learning on Transcription Machinery
12:00-12:15	**Zhuoyi Song** Fudan University, China	When Stochastic Timing Matters: Molecular Regulation of Signal Refractoriness Shapes Vision in the Fly Eye
12:15-13:30	**Lunch (Bldg 1#, Conference Venue, 群英厅)**	
Session 3 (Chair: Jianhua Xing)		
13:30-14:15	**Thomas A. Rando** University of California, Los Angeles, USA	Emergent Properties of Muscle Stem Cells During Muscle Regeneration
14:15-14:50	**Yifan Yang** Westlake University, China	A Mathematical Theory of Aging and Its Implications for Healthspan Extension
14:50-15:05	**Zachary Gao Sun** Yale University, USA	Feedback between F-actin Organization and Active Stress Govern Criticality and Energy Localization in the Cell Cytoskeleton
15:05-15:20	**Yuansheng Cao** Tsinghua University, China	Quantifying Subunit Exchange between Protein Complexes in Cyanobacterial Circadian Clock
15:20-15:50	**Coffee Break**	
Session 4 (Chair: Zhiyuan Li)		
15:50-16:35	**Kevin Foster** University of Oxford, UK	Competition and Warfare in Bacteria and the Human Microbiome
16:35-17:10	**Lei Dai** Shenzhen Institutes of Advanced Technology, CAS, China	Quantitative Ecology of Host-Associated Microbiomes
17:10-17:25	**Po-Yi Ho** Westlake University, China	Emergent Simplicity in Bacterial Growth under Complex Nutrient Environments
17:25-17:40	**Fajia Sun** Zhejiang University, China	Dynamic Adaptive Sampling for Enhanced Recovery and Characterization of Rare Species
18:00-20:00	**Dinner /Poster Session (Bldg 1#, Conference Venue, 群英厅)**	

2025/7/23 (Wednesday)		
Time	**Speaker**	**Title**
Session 1 (Chair: Ming Han)		
9:00-9:45	**David Weitz** Harvard University, USA	Cell Rheology, Revisited
9:45-10:20	**Adrienne Roeder** Cornell University, USA	Stay Smooth or Buckle: Coordination of Cell Growth Across the Arabidopsis Sepal
10:20-10:50	**Coffee Break**	
Session 2 (Chair: Lei Dai)		
10:50-11:35	**Robert H. Austin** Princeton University, USA	Inter-Agent Information Entropy Transfer from Bacteria to Cyborgs
11:35-11:50	**Dongliang Zhang** Max Planck Institute for the Physics of Complex Systems, Germany	An Altruistic Resource-Sharing Mechanism for Synchronization: The Energy-Speed-Accuracy Tradeoff
11:50-12:05	**Shiling Liang** Max Planck Institute for the Physics of Complex Systems, Germany	Thermodynamic Space: Operational Limits of Nonequilibrium Chemical Reaction Networks
12:05-13:30	**Lunch (Heyuan Restaurant, 和园餐厅)**	

2025/7/24 (Thursday)		
Time	**Speaker**	**Title**
Session 1 (Chair: Xiaojing Yang)		
9:00-9:45	**Wallace Marshall** University of California, San Francisco, USA	Learning without Neurons in a Single Cell
9:45-10:20	**Ke Li** University of Exeter, UK	Foundation Models for Genomics: From Predictive Modeling to Biological Design, towards AI Scientists and Future
10:20-10:50	**Coffee Break**	
Session 2 (Chair: Jie Lin)		
10:50-11:35	**Michael Shelley** Flatiron Institute & Courant Institute, USA	Self-Organizing Dynamics in Cellular and Multi-Cellular Systems
11:35-12:10	**Sean Megason** Harvard University, USA	Algorithms for Creating Form
12:10-13:30	**Lunch (Bldg 1#, Conference Venue, 群英厅)**	
Session 3 (Chair: Yifan Yang)		
13:30-14:05	**Jie Lin** Peking University, China	Universal Laws Emerging from Competition between Genes for Limiting Resources
14:05-14:20	**Hanqing Guo** Westlake University, China	Inhibition of Cellular Contractility Reveals Compression-enabled Mechanical Bistability during Epithelial Folding
14:20-14:35	**Xiaojing Yang** Peking University, China	Synthetic Conscription: Achieving Stable Labor Division with Rational Design of Gene Circuitry
14:35-14:50	**Zitong Jerry Wang** Westlake University, China	A Cellular Solution to a Robotics Problem
14:50-15:35	**Yuhai Tu** Flatiron Institute, USA	Nonequilibrium Thermodynamics in Living Systems: Towards Answering Schrödinger's Question
15:35-15:45	**Closing Speech**	

Spatial Omics to Predict Disease Trajectories



Nikolaus Rajewsky

Max Delbrück Center for Molecular Medicine - Berlin, Germany

Abstract

Nikolaus Rajewsky will give an overview of the advances and challenges in spatial transcriptomics. He will discuss how these data can be used to predict molecular mechanisms which are driving phenotypes in space and time, for example in tumorigenesis. He will also present analyses of micro/nanoplastics (which is contaminating our environment at exponential rate) and its uptake and molecular consequences of this uptake in human cells and tissues.

Uncovering Underlying Physical Principles and Driving Forces of Cancer from Single-Cell Transcriptomics

Jin Wang

Stony Brook University, USA

Abstract

Cancer has been a serious disease for human health. Genetic mutations have often been thought to be mainly responsible for the cancer formation. More evidences have been accumulated that cancer emergence is not just caused by individual gene perturbation but also from the whole network or state of the system. This shift of thinking demands global quantification and physical understanding of the underlying mechanisms for cancer formation. Here, we develop cancer models from the underlying gene regulatory networks either based on the available low throughput or the recent high throughput sequence experimental studies [1][2][3]. Cancer landscape can be quantified and cancer can be revealed as attractors representing cancer states. The landscape barrier and switching time become the measures of how difficult to transform from normal to cancer state. Furthermore, the cancer formation process can be quantified by the optimal paths between the normal state and cancer state. Due to the presence of the curl flux as the nonequilibrium driving force, the forward path and backward paths for cancer formation and normal state restoration are distinctly different. The global sensitivity analysis based on the landscape topography and nonequilibrium driving forces identify key genes and regulations responsible for the cancer formation. We further identify the nonequilibrium indicators through the curl flux, entropy production and time irreversibility as the early warning signals for the cancer formation. This helps to design practical strategy for cancer prevention and treatment.

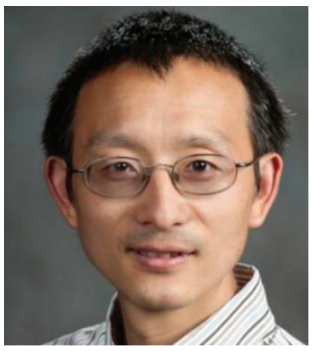
References:

[1] Ligang Zhu and Jin Wang*, Quantifying Landscape-Flux via Single-Cell Transcriptomics Uncovers the Underlying Mechanism of Cell Cycle. Adv. Sci. 2024, 11, 2308879 (2024)

[2] Ligang Zhu, Songlin Yang, Kun Zhang, Hong Wang, Xiaona Fang*, and Jin Wang*. Uncovering underlying physical principles and driving forces of cell differentiation and reprogramming from single-cell transcriptomics. Proc. Natl. Acad. Sci. USA 121: 34, e2401540121 (2024)

[3] Ligang Zhu and Jin Wang*, Deciphering physical mechanism of cancer from single cell transcriptomics and lieage-tracing. Submitted.

Dynamical Systems Theory Modeling in the Big Data Era



Jianhua Xing

Department of Computational and Systems Biology,
Department of Physics and Astronomy,
University of Pittsburgh, Pittsburgh, PA, USA

Abstract

Recent advances in single-cell experimental techniques and computational analysis have transformed biomedical research, shifting it from traditional question-driven, mechanistic approaches toward more data-driven paradigms. There is growing interest in integrating these distinct research styles to achieve deeper insights into biological processes. My lab is actively working to bridge this gap by placing high-throughput single-cell data analysis within the framework of dynamical systems theory — a foundational tool in traditional mathematical biology.

In this talk, I will present our recent work on how aging influences fate decisions in muscle stem cells. Genome-wide data offer an unbiased, systematic view of how aging alters gene expression and regulatory mechanisms at the single-cell level. Through mathematical analysis of the core gene regulatory network, we examine how these age-related dysregulations of individual genes impact cell fate decisions and contribute to the depletion of the stem cell pool at the population level. Our findings reveal key design principles underlying muscle stem cell function and highlight how aging compromises the robustness of this regulatory program.

Modeling, Simulation, and Learning on Transcription Machinery

Carmen AI Masri[1], Biao Wan[2], Liqiang Dai[2], Shannon J McElhenny[3], Chao E[4], and Jin Yu [1,3]

***Short Abstract* — Combining minimal model construction and molecular dynamics simulations from atomic to coarse-grained level, along with stochastic dynamics/kinetic approaches and machine learning methods, we have derived transcription factor (TF) target search free energy landscape around DNA binding sites with flanking sequences, and zoomed into structural dynamics of eukaryotic TF domain and dimeric TFs for stepwise 1D search along DNA with sequence recognition. We also investigated mechanochemical & fidelity control of single-subunit RNA polymerases (RNAPs). We particularly explored a bacteriophage RNAP on DNA promoter recognition and switching from one promoter to the other on the path of lab directed evolution.**

***Keywords* — Transcription Factor (TF), RNA Polymerase (RNAP), Physical Modeling, Molecular Dynamics (MD) Simulation, Machine Learning (ML)**

I. Introduction

We study physical mechanisms of protein machinery regulating gene transcription, including transcription factors or TFs that search for target DNA sequences and RNA polymerases or RNAPs that are responsible for initiating and elongating RNA synthesis. The main purpose is to reveal structure dynamics and energetics of genetic regulation from microscopic to mesoscopic scale, complementary to experimental challenges on high-resolution dynamics tracking, and hopefully to connect single molecule protein and DNA/RNA operation with cellular regulation function.

II. Results

We have demonstrated atomic resolution molecular dynamics (MD) of 1D diffusion of a representative TF DNA binding domain (DBD), e.g., WRKY domain from plant, in 1-bp stepping spontaneously on DNA [1]. We also probed DBD coordination of a dimeric TF, e.g. Myc-Max from human and related TFs on DNA, and analyzed TF-DNA interfacial stabilities comparatively [2,3]. To determine TF-DNA binding affinities and specificities from structure-based simulations, we mechanically steered protein-DNA dissociation [4], alchemically induced DNA sequence variation [5], and combined physically based simulations and energetic calculations with machine learning (ML), trained upon dataset of genomic context protein binding microarray (gcPBM) [6]. Our minimal structural model of TF-DNA stochastic dynamics and construction of binding & diffusion free energy landscape [7], however, suggests experimental data limitation of TF binding profiling acquiring DNA sequence information.

We also investigated single-subunit RNAP mechanochemical and fidelity control during transcription elongation, comparatively among bacteriophage T7 RNAP [8,9], viral RNA dependent RNA polymerase (RdRp) from SARS-CoV-2 [10], and structurally similar human mitochondrial RNAP or POLRMT [11]. In particular, we have learned mechanism of T7 RNAP-DNA promoter recognition and its switching to T3 promoter recognition following a lab directed evolution path with our MD computational microscope [12].

III. Conclusion

Protein-DNA structural dynamics and energetics for genetic information acquisition and fidelity control are focuses of our above studies. We identified key structural dynamics features such as collective hydrogen bond dynamics at the TF-DNA interface, interfacial surface complementarity, and protein-DNA electrostatic interactions in some hierarchical order facilitating the protein-DNA sequence readout. While for template-based RNA synthesis driven by chemical free energy, the fidelity control of RNAP is primarily achieved on top of protein diffusion or Brownian motion poised highly sensitively to energetics of incoming cognate or non-cognate nucleotide substrate initial binding, insertion, and incoporation.

Acknowledgements: Studies were funded by UC-CRCC, NSF, &NSFC.
[1]Department of Physics, University of California-Irvine, CA, USA.
Corresponding E-mail: Jin Yu jin.yu@uci.edu
[2]Wenzhou Institute, UCAS, Wenzhou, Zhejiang, China
[3]Department of Chemistry, University of California-Irvine, CA,USA
[4]Beijing Computational Science Research Center, Beijing, China

References

[1] L Dai, Y Xu, Z Du, XD Su, and J Yu (2021) *PNAS* **118** (23): e2102621118
[2] L Dai and J Yu (2020) *BBRC* **533**, 97-103.
[3] CA Masri and J Yu (2025) *preprint* 10.26434/chemrxiv-2025-04q7d
[4] CA Masri, B Wan, and J Yu (2023). *Biophys J* **122**: 4476-4487
[5] CA Masri, JZ Vilseck, J Yu, RL Hayes (2025). *J Chem Theory Comput* **21**: 3536–3544
[6] CA Masri and J Yu (2025) *preprint* 10.26434/chemrxiv-2025-mc5q4
[7] B Wan and J Yu (2025) *Biophys J*, 124(4):677-692
[8] LT Da, C E, Y Shuai, S Wu, XD Su, and J Yu (2017). *Nucleic Acids Res* **45**(13) 7909-7921.
[9] C Long, C E, LT Da, and J Yu (2019). *Nucl Acids Res* **47**(9), 4721-4735.
[10] ME Romero, SJ McElhenney and J Yu (2024). *Phys Chem Chem Phys*, **26**, 1792-1808
[11] SJ McElhenney and J Yu (2025). *J Chem Theory Comput* **21**, 4815-4829
[12] C E, L Dai, and J Yu (2022). *Biophys J* **121**, 582-595

When Stochastic Timing Matters: Molecular Regulation of Signal Refractoriness Shapes Vision in the Fly Eye

Changqian Rao[1,2], David Waxman[3,4,5,6], Wei Lin[1,2,3,4], Zhuoyi Song[3,4,5,6*]

[1]*School of Mathematical Sciences, Fudan University, Shanghai, 200433, China*
[2]*Research Institute of Intelligent Complex Systems, Fudan University, Shanghai, 200433, China*
[3]*Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, 200433, China*
[4]*Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, China*
[5]*MOE Frontiers Center for Brain Science, Fudan University, Shanghai 200433, China*
[6]*Zhangjiang Fudan International Innovation Center, Shanghai, China*

Abstract

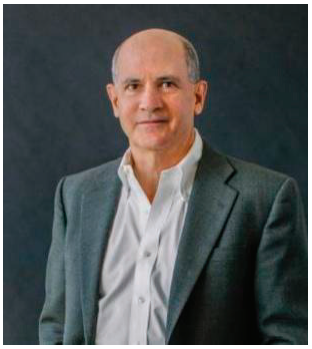
Cells don't merely respond to signals—they regulate when and how they respond, often through precisely timed yet stochastic mechanisms. Beyond amplification and attenuation, a third and often overlooked phase—the refractory period (RP)—introduces transient unresponsiveness after stimulation, shaping pulsatile dynamics across signaling networks[1]. Yet how RP stochasticity is molecularly regulated—and how it shapes system behavior—remains unclear. Here, we develop an analytical framework to model RP distributions in the *Drosophila* phototransduction cascade[2, 3], revealing how input statistics, feedback motifs, and spontaneous ion channel noise govern RP timing. These molecular controls reshape photoreceptor filtering from low-pass to band-pass, enhancing the temporal precision and motion sensitivity of the fly compound eye. Our framework links molecular mechanisms to system-level function across scales—from molecules to cells to the whole eye—revealing a molecular logic for speed and sharpness in visual perception. Because RPs and pulsatile dynamics are widespread, this work offers a general framework for understanding stochastic timing in cell signaling[4].

References

1. Amit, I., et al., *A module of negative feedback regulators defines growth factor signaling.* Nature Genetics, 2007. **39**(4): p. 503-512.
2. Song, Z., et al., *Stochastic, adaptive sampling of information by microvilli in fly photoreceptors.* Current Biology, 2012. **22**(15): p. 1371-80.
3. Rao, C., Lin, W., and Song, Z., *Analytical refractory period distribution for a class of time-variant biochemical systems with second-order reactions.* Journal of Chemical Physics, 2023. **159**(12).
4. Sumit, M., Takayama, S., and Linderman, J.J., *New insights into mammalian signaling pathways using microfluidic pulsatile inputs and mathematical modeling.* Integrative Biology, 2017. **9**(1): p. 6-21.

*Email: songzhuoyi@fudan.edu.cn

Emergent Properties of Muscle Stem Cells During Muscle Regeneration



Thomas A. Rando

Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, USA

Biography

Thomas A. Rando is a stem cell biologist and neurologist, best known for his research on basic mechanisms of stem cell biology and the biology of aging. He is the Director of the Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research and a professor of Neurology and Molecular, Cell and Developmental Biology at the University of California, Los Angeles. He is an elected member of the American Academy of Arts and Sciences and the National Academy of Medicine.

The main focus of Rando's laboratory is on the molecular regulation of stem cell function. Their primary goals are to understand how stem cells orchestrate tissue homeostasis and repair and to harness their remarkable therapeutic potential in the treatment of tissue damage and degenerative diseases. With a primary focus on skeletal muscle and their resident stem cells ("satellite cells"), they also examine how stem cells age and how to restore youthful properties to aged stem cells. Their long-terms goals are to unravel the mysteries of stem cells and to apply that knowledge to the future of cell therapies for diseases for which no treatments or cures are currently available.

As the director of the Broad Stem Cell Research Center, he promotes a culture of scientific excellence, academic integrity, and interdisciplinary collaboration. He also builds partnerships with the larger scientific community — including other institutions and private companies — to address the most urgent challenges in regenerative medicine with the shared goal of improving human health.

A Mathematical Theory of Aging and Its Implications for Healthspan Extension

Yifan Yang

Westlake University, China

Abstract

A century of work has catalogued myriad molecules, pathways, and physiological states that modulate aging, yet an integrative quantitative theory of why organisms age is still missing. Such quantitative theories require coarse-graining biological complexity into a small set of universal, predictive, and interpretable variables. We have now realized this crucial step. Longitudinal measurements in mice and single *E. coli* cells reveal that the dynamics of physiological damage in both organisms obey the same low-dimensional stochastic differential equation. This model—with no retraining—predict new experimental data quantitatively, and helps to uncover a previously unknown relationship between healthspan and lifespan inequality.

Feedback between F-actin organization and active stress govern criticality and energy localization in the cell cytoskeleton

Zachary Gao Sun[1,2,3], Nathan Zimmerberg[4,5], Patrick Kelly[4,6], Carlos Floyd[7], Garegin Papoian[8], and Michael Murrell[1,2,3,9,*]

[1] *Department of Physics, Yale University, 217 Prospect Street, New Haven, Connecticut 06511, USA*
[2] *Systems Biology Institute, Yale University, 850 West Campus Drive, West Haven, Connecticut, 06516, USA*
[3] *Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, Connecticut 06520, USA*
[4] *Institute for Physical Science and Technology, University of Maryland, College Park, MD, USA*
[5] *Biophysics Program, University of Maryland, College Park, MD, 20742, USA*
[6] *Chemical Physics Program, University of Maryland, College Park, MD, USA*
[7] *The James Franck Institute, The University of Chicago, Chicago, Illinois, 60637, USA*
[8] *Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, USA*
[9] *Department of Biomedical Engineering, Yale University, 55 Prospect Street, New Haven, Connecticut 06511, USA*
[*]*Corresponding author. Email address: michael.murrell@yale.edu* .

Abstract

Self-organized criticality can occur in earthquakes, avalanches and biological processes and is characterized by intermittent, scale-free energy dissipation. In living cells, the actin cytoskeleton undergoes dynamic structural reorganization, particularly during migration and division, where molecular motors generate mechanical stresses that drive large dissipative events. However, the mechanisms governing these critical transitions remain unclear. Here, we show that cytoskeletal criticality emerges from the interplay between F-actin organization and active stress generation. Our study focuses on a minimal actomyosin system *in vitro*, which is composed of F-actin filaments, myosin II motors and nucleating promoting factors. By systematically varying actin connectivity and nematic order, we demonstrate that ordered and sparsely connected networks exhibit exponential stress dissipation, whereas disordered and highly connected networks show heavy-tailed distributions of energy release and 1/f noise characteristic of self-organized criticality. Increased disorder leads to stress localization, shifting force propagation into stiffer mechanical modes, reminiscent of Anderson localization in condensed matter systems. Furthermore, we show that network architecture directly regulates myosin II filament size, establishing a chemical-mechanical feedback loop that modulates criticality. Our findings provide insights into collective cytoskeletal dynamics, energy localization and cellular self-organization.

References

Sun et al., *Nature Physics* (Accepted, April 15th, 2025)

Email: zachary.sun@yale.edu

Quantifying subunit exchange between protein complexes in cyanobacterial circadian clock

Yuansheng Cao
Department of Physics, Tsinghua University
Tsinghua University, Beijing 100084, China

Abstract

Protein complexes are essential for cellular functions, with their dynamics increasingly recognized as key regulatory elements. In the cyanobacterial circadian clock, the KaiC hexamer, the core clock protein complex, are highly dynamic. Depending on their phosphorylation levels, the KaiC hexamers can exchange subunits in a phenomenon called monomer-shuffling, which is crucial for synchronizing individual hexamers vital to the circadian rhythm. However, the molecular mechanism of monomer-shuffling is not fully understood. In this study, we developed a chromatography-based method to monitor the dynamics of the hexamerization domain of KaiC (KaiC-CI). Our findings indicate that ATPase activity is essential for monomer-shuffling. By analyzing experimental data with quantitative models, we found that ATP hydrolysis at two independent monomer-monomer interfaces disassembles a KaiC-CI hexamer into two oligomers. Two complementary oligomers from two different hexamers can reform a mixed hexamer rapidly. Mechanistically, ATP hydrolysis induces unstable, shuffling-competent states in hexamers, while nucleotide exchange stabilizes them, suppressing shuffling. The balance between these two processes quantitatively explains the relationship between shuffling rate and ATPase activity across KaiC mutants observed in our experiments. Applying this model to full-length KaiC revealed that phosphorylation of the KaiC-CII domain modulates shuffling by affecting the nucleotide exchange rate in the CI domain. This establishes a feedback loop where CII phosphorylation alters CI ATPase activity, which controls subunit exchange to synchronize CII phosphorylation across the KaiC hexamer population. Our findings provide a mechanistic framework linking ATPase-driven subunit exchange to the robustness of circadian clocks, with implications for activation-coupled dynamics in other molecular machines.

Email: yscao@tsinghua.edu.cn

Competition and Warfare in Bacteria and the Human Microbiome

Kevin R Foster

University of Oxford, UK

Abstract

Microbial communities contain many evolving and interacting bacteria, which makes them difficult to understand and predict. Using a combination of theory and experiment, we study what it takes for bacteria to succeed in diverse communities. One way is to actively kill and inhibit competitors and we study the strategies that bacteria use in toxin-mediated warfare. We are now also using our understanding of bacterial competition to try to manipulate gut communities for better health. Our ultimate goal is to both stabilise microbiome communities and remove problem species without the use of antibiotics.

Quantitative Ecology of Host-Associated Microbiomes

Lei Dai

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

Abstract:

The realization that microbiomes, associated with virtually all multicellular organisms, have tremendous impact on their host health is considered as one of the most important scientific discoveries in the last decade. The host-associated microbiomes, composed of tens to hundreds of co-existing microbial species, are highly heterogenous at multiple scales (e.g. between different hosts and within a host). In this talk, I will share our recent works on understanding the heterogeneity of complex microbial communities, and illustrate that conceptual and technological advances in microbial ecology are critical for microbiome engineering.

Emergent simplicity in bacterial growth under complex nutrient environments

Aoyu Zhu[1] and Po-Yi Ho[1]

***Short Abstract* — Bacterial metabolism is dependent on the nutrient environment and variable across species. To what extent can this complexity be explained by simple quantitative principles? Here, we show that in complex nutrient environments, the growth of a species is proportional to the number of nutrients it utilizes, both experimentally and in genome-scale metabolic models. To uncover the network features underlying this proportionality, we coarse-grained the models using representation learning and identified efficient pathways between reaction clusters that enable high growth at low proteome cost. These efficient coarse-grained pathways underlie synergistic nutrient interactions and emerge as key determinants of growth across species and environments.**

I. Purpose

BACTERIA regulate their metabolism to grow in diverse nutrient environments. In simple environments such as defined laboratory conditions with only one or two carbon sources, growth regulation strategies like co-utilization and diauxie have been extensively studied and can be quantitatively understood through mathematical models of enzyme kinetics and gene regulation. Less understood is growth in complex environments with many nutrients such as human guts and other natural contexts. A central challenge is that it is unclear how to generalize the principles for growth in simple environments to diverse metabolic networks with thousands of metabolites and reactions. We recently made a curious experimental observation that provides a hint: For phylogenetically diverse bacterial species grown individually in a complex medium, biomass yield is proportional to the number of nutrients utilized. This proportionality suggests that there exist coarse-grained network features shared across species that dictate growth in complex environments. Here, we search for these emergent principles in silico.

II. Results

A. Optimization under proteome constraints gives rise to proportionality between growth and nutrient count

We hypothesized that the proportionality arises because every species implements a universal regulation strategy to maximize growth rate while minimizing proteome cost. While the molecular mechanisms of this regulation likely varies across species, the end result on growth rate and metabolic fluxes can be predicted from network structure alone. This framework is known as flux balance analysis (FBA) and is widely used for metabolic circuit design. We applied FBA to a collection high-quality metabolic networks for more than two hundred phylogenetically diverse species. When all possible nutrients are provided, the predicted growth rate is indeed proportional to nutrient count, recapitulating the experimental observation.

B. Environment complexity promotes nutrient synergy

To further quantify how nutrients contribute to growth, we recasted FBA-derived fluxes into reaction-based graphs in which an edge between two reactions represent the mass flow between them mediated by a shared metabolite. These mass flow graphs (MFGs) highlight major metabolic routes, in part by reducing the confounding influence of highly connected hub metabolites. By tracing nutrient-to-biomass paths on MFGs, we quantified the mass flow mediated by specific nutrients through metabolic subsystems. The resulting decomposition revealed that nutrients are either multifunctional (involved in many subsystems) or specialized (direct biomass contributors) and that the proportion of specialized nutrients increases with environment complexity.

C. Efficient coarse-grained pathways enable nutrient synergy and are universal across species and environments

If the effects of environment complexity are driven by some network features that are shared across species, then these features are likely coarse-grained because metabolic networks are highly dissimilar across species. To obtain a coarse-grained representation, we embedded MFGs based on both the local context and the global position of reactions using node2vec and clustered the learned embeddings into a small number of reaction groups, amounting to an unsupervised learning of subsystems. By using this coarse-grained representation to perform systematic analyses across environments and species, we identified links between reactions groups that emerge at high environment complexity. These efficient coarse-grained pathways can also be activated by specific nutrient combinations, indicating that they represent network features that enable nutrient synergy. Remarkably, the number of activated pathways is strongly linearly correlated with growth.

[1]Center for Interdisciplinary Studies, Westlake University, Hangzhou, China. Email: poyiho@westlake.edu.cn

Dynamic Adaptive Sampling for Enhanced Recovery and Characterization of Rare Species

Fajia Sun[1,2,3] Qi Ouyang[1,*], Long Qian[2,*] and Nick Goldman[3,*]

***Short Abstract* — Real-time nanopore sequencing allows selective rejection of reads to enrich targets. We present a dynamic adaptive sampling (DAS) method that combines real-time database querying with de Bruijn graph-based decision-making to prioritize sequencing of rare and novel species in metagenomic samples. DAS dynamically evaluates the expected benefit of each read and reallocates sequencing efforts accordingly. Our strategy achieved 48–135% enrichment of rare species and significantly accelerated their detection in both simulated and real-world experiments, including hospital wastewater and standardized gut microbiome samples.**

***Keywords* — dynamic adaptive sampling, nanopore sequencing, metagenomics, de Bruijn graph**

I. BACKGROUND

Nanopore sequencing is a transformative technology for genomics due to its real-time data acquisition and capacity for long-read sequencing with minimal sample preparation [1–3]. Unlike traditional sequencing-by-synthesis platforms, nanopore sequencing permits immediate read analysis and decision-making, which has led to the development of adaptive sampling strategies that dynamically determine whether to continue sequencing individual DNA molecules [4–6]. This functionality enables efficient use of sequencing capacity by enriching informative targets and reducing redundant data.

Dynamic adaptive sampling (DAS) builds upon this concept by incorporating real-time feedback loops and algorithmic decision-making to iteratively refine read selection during sequencing [7]. While DAS has demonstrated success in controlled microbial mock communities [7,8], its application to natural metagenomic samples remains challenging. Real metagenomes are composed of thousands of species with highly variable abundance and complex genomic architectures [9–12]. These complexities are exacerbated by taxonomic similarities and horizontal gene transfer among microbes, which cause substantial sequence sharing even between distantly related species.

[1]Institute for Advanced Study in Physics, Zhejiang University, Hangzhou, 310027, China.
[2]Center for Quantitative Biology, Peking University, Beijing, 100871, China
[3]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI); Hinxton, CB10 1SD, UK.
*Corresponding author. Email: qoy@zju.edu.cn; long.qian@pku.edu.cn; goldman@ebi.ac.uk

II. SUMMARY

In this study, we address these challenges by developing a novel DAS framework that combines database-based species recognition with a de Bruijn graph-based strategy for read evaluation. Our method prioritizes sequencing reads based on their estimated contribution to resolving rare or novel species, suppresses overrepresented taxa, and enables dynamic reallocation of sequencing resources as the experiment progresses. This strategy provides a scalable, real-time solution for improving resolution in metagenomic nanopore sequencing and lays the foundation for enhanced detection and characterization of low-abundance or previously uncharacterized organisms.

REFERENCES

[1] Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17(6), 333–351.
[2] Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. Nature Biotechnology, 39(11), 1348–1365.
[3] Ying, Y.-L., Hu, Z.-L., Zhang, S., Qing, Y., Fragasso, A., Maglia, G., Meller, A., Bayley, H., Dekker, C., & Long, Y.-T. (2022). Nanopore-based technologies beyond DNA sequencing. Nature Nanotechnology, 17(11), 1136–1146.
[4] Martin, S., Heavens, D., Lan, Y., Horsfield, S., Clark, M. D., & Leggett, R. M. (2022). Nanopore adaptive sampling: a tool for enrichment of low abundance species in metagenomic samples. Genome Biology, 23(1), 11.
[5] Ulrich, J. U., Lutfi, A., Rutzen, K., & Renard, B. Y. (2022). ReadBouncer: precise and scalable adaptive sampling for nanopore sequencing. Bioinformatics, 38(Supplement_1), i153–i160.
[6] Munro, R., Wibowo, S., Payne, A., & Loose, M. (2024). Icarust, a real-time simulator for Oxford Nanopore adaptive sampling. Bioinformatics, 40(4), btae141.
[7] Weilguny, L., De Maio, N., Munro, R., Manser, C., Birney, E., Loose, M., & Goldman, N. (2023). Dynamic, adaptive sampling during nanopore sequencing using Bayesian experimental design. Nature Biotechnology, 41(7), 1018–1025.
[8] Weilguny, L. (2023). Computational optimisation strategies for targeted DNA sequencing using nanopores [Apollo - University of Cambridge Repository]. https://doi.org/10.17863/CAM.104549.
[9] Hugenholtz, P., & Tyson, G. W. (2008). Metagenomics. Nature, 455(7212), 481–483.
[10] Sharon, I., & Banfield, J. F. (2013). Genomes from metagenomics. Science, 342(6162), 1057–1058.
[11] Chiu, C. Y., & Miller, S. A. (2019). Clinical metagenomics. Nature Reviews Genetics, 20(6), 341–355.
[12] Miller, S., & Chiu, C. (2022). The role of metagenomics and next-generation sequencing in infectious disease diagnosis. Clinical Chemistry, 68(1), 115–124.

Cell Rheology, Revisited

David Weitz

Dept. of Physics and SEAS, Harvard University, Cambridge MA USA

Abstract

The mechanical properties of a cell are determined primarily by an interpenetrating network of biopolymers. This talk will revisit several features of the mechanical properties of a cell. By using magnetic tweezers to pull a magnetic particle through the cytoplasm of a cell, we show that the particle exhibits unusual behavior: Its velocity is independent of the force pulling the particle. This velocity can be used as a probe of the mechanics within the cell and the contribution of the different filament networks. We suggest that this behavior requires a different constitutive equation to describe the rheology of the cell. We also reexamine the properties of vimentin intermediate filaments and suggest that their behavior is reminiscent of a self-assembled structure, a worm-like micelle, formed by surfactants. This perspective accounts for many properties that are observed for vimentin intermediate filament networks. Vimentin also forms phase-separated liquid droplets that are a precursor to formation of the filament network. These properties demonstrate how soft-matter physics can be used to describe the mechanical properties of a cell.

Stay Smooth or Buckle: Coordination of Cell Growth across the *Arabidopsis* Sepal

Adrienne Roeder

Cornell University, USA

Abstract

Development is remarkably reproducible, producing plant organs such as leaves and sepals with the same size and shape repeatedly from individual to individual. Yet, these reproducible organs are composed of cells that are highly variable in size, shape, growth, and division. My laboratory uses a computational morphodynamics approach, including genetics, live imaging, image processing, mechanical assays, and computational modeling, to elucidate the mechanisms that produce robust organ size and shape from cellular heterogeneity. We use *Arabidopsis* sepals as a model organ because they are accessible for imaging and manipulation. We identified a mutant (*as2-7D*) with ectopic expression of the transcription factor *ASYMMETRIC LEAVES 2* (*AS2*) on the outer epidermis. Our analysis reveals that ectopic *AS2* expression causes the outer epidermis of *as2-7D* sepals to buckle during early stages of development. We show that buckling of the outer epidermis occurs due to conflicting cell growth directions and unequal tissue stiffness across the epidermal layers. These studies reveal that coordination of growth across the organ is required for smooth organ morphogenesis.

Inter-Agent Information Entropy Transfer from Bacteria to Cyborgs

Robert H. Austin[1], Shengkai Liu[1], Trung Phan[2] and Liyu Liu[3]

[1]Department of Physics, Princeton University
[2]Dept. of Chemical Engineering, Johns Hopkins University
[3]Institute of Physics, Chinese Academy of Sciences

The early realization that quorum sensing information transfer in bacteria was a key way that bacteria collectively communicated with each other has expanded our understanding of information entropy exchange between biological agents as a over-arching concept in biology. This idea of inter-agent information entropy can be expanded to recent attempts where we embed synthetic systems with biological properties (cyborgs) and utilize information entropy as the primary means of collective computation.

I. PURPOSE The emergence of the importance of quorum sensing in bacterial communities [1] transformed our understanding of the complexity of the communication between even "simple" organisms such as bacteria. At first quorum sensing was considered to be confined to bacteria where they communicated via signaling molecules their local presence, but this idea has been expanded in the realization that many other organisms exhibit collective population density dynamics [2]. While it is interesting to model these collective dynamics from a purely physics perspective, presumably these organisms have a fitness reason for their collective behavior, and this can be addressed from a information entropy perspective.

Perhaps more intriguingly, we now are at the point technologically where we can begin to embed in non-biological entities (originally robots) more of the attributes normally considered to be exclusively in the realm of biology [3], that is, synthetic life. But biological constructs, such as brain organoids, still have potential capabilities in terms of information processing that eludes synthetic life. By embedding brain organoids in synthetic life constructs, which we here call cyborgs, we will see the merging of biology information processing, which seems to be primarily non-bit related as we see in quorum sensing, with the digital world of the robot.

We will present 3 examples from our own work which span this progression of information entropy transfer from bacteria to cyborgs.

II. EXAMPLES 1) We introduce an experiment using a hydrodynamic information sink but metabolic source linked to a semicircular arrangement of confined bacteria to

probe bacterial information exchange. The sink removes auto-inducer signaling molecules from an enclosed colony of *E. coli* but provides fresh nutrients. We show that a collective spatial population inversion, that is, fewer bacteria where there is more nutrient but less auto-inducer signal, emerges at the information sink, even with the pumping action of funnels to the information sink. We reproduce this population inversion using both a set of partial differential equations and an agent-based simulation of the depletion of the auto-inducer signaling fields and their motility response [1]. **2)** We are conducting an experimental and theoretical investigation of intelligent active matter, where robotic agents, equipped with multimodal sensors, artificial intelligence chips, and reinforcement learning algorithms, exhibit adaptive behavior models driven by the complex and dynamic environment. These agents undergo allosteric swarm strategy transitions and exhibit goal-oriented swarm intelligence evolutions. This work demonstrates that the coupling between internal information processing and energy dissipation can drive the emergence of collective structures and dynamics, thereby constructing evolvable and adaptive swarm intelligence [2]. **3)** We record electrophysiological signals from cerebral organoids using microelectrode arrays and associate them with the kinematic behavior of robots to control robotic actions. Interactions between robots are then fed back to the corresponding organoids via electrical stimulation, thus forming a closed-loop system with one-to-one mapping. The collective behavior of the robot swarm can characterize pathological features of the cerebral organoids and can further be used to test the therapeutic effects of drugs on the organoids.

CONCLUSION The general phenomena and analysis presented here showing the importance of quantitive aspects of information entropy transfer can be modified more generally to many aspects of the social dynamics of populations, including the world of cyborgs which is being rapidly developed.

REFERENCE

[1] S. Ro, B. Guo, A. Shih, T. V. Phan, R. H. Austin, D. Levine, P. M. Chaikin, and S. Martiniani, Model-free measurement of local entropy production and extractable work in active matter, Physical review letters 129, 220601 (2022).

[2] G. Wang, T. V. Phan, S. Li, M. Wombacher, J. Qu, Y. Peng, G. Chen, D. I. Goldman, S. A. Levin, R. H.Austin, et al., Emergent field-driven robot swarm states, Physical review letters 126, 108002 (2021).

[3] Smith A, Jones B, Simpson C "Tentative title of manuscript in preparation," in preparation.

[4] Smith A, Jones B (in press) Title of a paper in press. J Mol Biol. optional doi:code and/or link

[5] Who RU, Smith A, Jones B (2005) Please abbreviate author list only when there are four or more authors. PNAS 100, 5403-5407.

[6] Big MR, Rocky Mountain Town, Santa Fe, NM, private communication.

[7] CNLS web site, optional date, http://cnls.lanl.gov/

An altruistic resource-sharing mechanism for synchronization: The energy-speed-accuracy tradeoff

Dongliang Zhang[1], Yuansheng Cao[2], Qi Ouyang[3] and Yuhai Tu[4]

***Short Abstract* — Synchronization among a group of active agents is ubiquitous in nature. Inspired by KaiA differential binding in the KaiABC system, we propose a new model for synchronization based on an altruistic resource-sharing (ARS) mechanism wherein resources are needed for individual agent to advance but a more advanced agent has a lower competence to obtain resources. By solving the model analytically, our study reveals a general tradeoff relation between the total energy dissipation rate and the two key performance measures of the system: average speed and synchronization accuracy. Implications of our work to realistic biological systems are discussed.**

***Keywords* — Synchronization, energetics, tradeoff**

I. Introduction

Synchronization of molecular systems is a collective phenomenon critical for maintaining biological functions. Mechanistically, synchronization can be achieved by either direct or indirect interactions between individual molecular systems. Although synchronization based on direct interactions between agents is relatively well understood (e.g. described by Kuramoto model), the other general mechanism based on indirect interactions among agents sharing limited resources is less known.

Here, inspired by KaiA differential binding in the KaiABC system, we propose a minimal thermodynamically consistent model for the altruistic resource-sharing (ARS) mechanism wherein resources are needed for an individual agent to advance but a more advanced agent has a lower competence to obtain resources. We find an energy-speed-accuracy (ESA) tradeoff relation [1], similar to the relation we found in the direct interaction mechanism in a previous work [2], suggesting a universal relation between energy cost and synchronization performance independent of detailed mechanisms.

II. Results

We show that while differential competence in ARS mechanism provides a negative feedback leading to synchronization it also breaks detailed balance and thus requires additional energy dissipation besides the cost of driving individual agents.

By solving the model analytically, we derive the steady state distribution of a traveling solution, from which we can calculate the speed, accuracy, and the free energy dissipation rate. Our results reveal that, besides the importance of differential competence, synchronization also depends critically on the scarcity of the resource: ARS mechanism only works when the resource is relatively scarce and synchronization improves as resource abundance decreases.

On top of that, our results reveal a general tradeoff relation between the total energy dissipation rate and the two key performance measures of the system: average speed and synchronization accuracy. For a fixed dissipation rate, there is a distinct speed-accuracy Pareto front traversed by the scarcity of resources: scarcer resources lead to slower speed but more accurate synchronization. Increasing energy dissipation eases this tradeoff by pushing the speed-accuracy Pareto front outwards, leading to potentially higher speed or synchronization accuracy.

III. Outlook and application

We summarize the similar points and the key differences between the synchronization generated by direct pairwise interaction mechanism and indirect ARS mechanism. We also discuss how our results can have realistic impact, e.g. the scarcity requirement of the resource. We examine how both the direct and indirect interactions work together in realistic biological systems such as KaiABC system, and briefly explain why they work synergistically (with a high efficiency).

[1]Max Planck Institute for the Physics of Complex Systems, Dresden, Germany. E-mail: dlzhang@pks.mpg.de

[2]Department of Physics, Tsinghua University, Beijing, China. Email: yscao@mail.tsinghua.edu.cn

[3]Institute for Advanced Study in Physics, School of Physics, Zhejiang University, Hangzhou, China

[4]Center for Computational Biology & Center for Computational Neuroscience, Flatiron Institute, New York, USA. Email: ytu@flatironinstitute.org

References

[1] J. S. van Zon, D. K. Lubensky, P. R. Altena, and P. R.ten Wolde, Proceedings of the National Academy of Sciences 104, 7420 (2007).

[2] Zhang, D., Cao, Y., Ouyang, Q., & Tu, Y. (2025). An altruistic resource-sharing mechanism for synchronization: The energy-speed-accuracy tradeoff. arXiv preprint arXiv:2502.02242.

[3] Zhang, D., Cao, Y., Ouyang, Q., & Tu, Y. (2020). The energy cost and optimal design for synchronization of coupled molecular oscillators. *Nature physics*, 16(1), 95-10.

Thermodynamic Space: Operational Limits of Nonequilibrium Chemical Reaction Networks

Shiling Liang[1,2,3,4,*], Paolo De Los Rios[1], and Daniel Maria Busiello[2,5]

***Short Abstract* — Living systems rely on Chemical Reaction Networks (CRN) operate far from equilibrium to sustain essential functions and maintain complex organization. Energy availability fundamentally limits their operational possibilities. We introduce the *Thermodynamic Space* – the accessible concentration range dictated by the energy budget. We derive universal thermodynamic bounds for *generic* CRNs, linking global energy inputs and network topology to local concentrations and reaction affinities. This framework constrains emergent behaviors like bistability, self-assembly, and pattern formation, providing a tool to understand the physical limits of biological complexity across scales**

***Keywords* —Chemical Reaction Networks, Non-equilibrium Thermodynamics.**

Living systems operate far from equilibrium, harnessing energy to maintain complex states via Chemical Reaction Networks (CRNs). Understanding the constraints imposed by thermodynamics on these networks, especially those involving multi-molecular interactions common in biology, is fundamental yet challenging [1-3]. This work aims to establish universal thermodynamic bounds for any CRN operating with a given energy budget, bridging the gap between global energetic driving and local molecular behaviour [4-5]. We introduce the *Thermodynamic Space* concept to delineate the accessible concentration range for species under non-equilibrium conditions, providing a general tool to predict limits and understand the prerequisites for complex phenomena like bistability, self-assembly, and pattern formation.

We derived two main sets of thermodynamic bounds for generic CRNs at steady state. First, we established limits on the affinity (thermodynamic force) of any reaction, showing it is constrained by the maximum and minimum cycle affinities of the Elementary Flux Modes (EFMs) it participates in. Using a novel "chemical probe" technique, we extended these bounds to effective affinities between any set of species compatible with the system's conservation laws.

Second, by translating these affinity bounds, we defined the *Thermodynamic Space*, which sets rigorous upper and lower limits on the stationary-state concentrations of chemical species. This accessible region quantifies how global non-equilibrium driving determines the possible range of chemical concentrations. At equilibrium, this space collapses to a single point; its non-zero volume is a signature of non-equilibrium conditions and a prerequisite for emergent complexity.

The framework's power is demonstrated by its ability to constrain diverse emergent phenomena in complex chemical systems. It quantifies how energy dissipation enables **bistability**, bounding the range of stable fixed points in biochemical systems like the Schlögl model. It limits the non-equilibrium **amplification** of complex structures beyond their equilibrium yields in dissipative **self-assembly** processes. It defines the achievable **degree of chiral symmetry breaking**, providing rigorous thermodynamic bounds on chiral imbalance. Furthermore, it establishes a direct relationship between thermodynamic driving and the characteristics of **spatial patterns**, constraining pattern contrast in reaction-diffusion systems. These examples illustrate how a non-zero Thermodynamic Space is a prerequisite for complex behaviors

By providing a general tool for analyzing CRNs based solely on underlying thermodynamic properties, this framework constitutes a stepping stone to deepen our ability to predict complex out-of-equilibrium behaviors and design artificial chemical systems.

Acknowledgements: This work was funded by wiss National Science Foundation under grant 200020 178763

[1]Ecole Polytechnique Federale de Lausanne (EPFL), 1015 Lausanne, Switzerland

[2]Center for Systems Biology Dresden, 01307 Dresden, Germany

[3]Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany

[4]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

[5]University of Padova, 35131 Padova, Italy

[*]Email: shiling@pks.mpg.de

References

[1] Fang, X., Kruse, K., Lu, T., & Wang, J. (2019). Nonequilibrium physics in biology. *Reviews of Modern Physics*, *91*(4), 045004.

[2] Palsson BO (2015) Systems Biology: Constraint-based Reconstruction and Analysis. Cambridge University Press, Cambridge.

[3] Qian, H. (2007). Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.*, *58*(1), 113-142.

[4] Liang, S., De Los Rios, P., & Busiello, D. M. (2024). Thermodynamic bounds on symmetry breaking in linear and catalytic biochemical systems. *Physical Review Letters*, *132*(22), 228402.

[5] Liang S., De Los Rios P., Busiello D. M. (2024). Thermodynamic Space of Chemical Reaction Networks. *arXiv:2407.11498*.

Learning without Neurons in a Single Cell

Wallace Marshall

University of California, San Francisco, USA

Abstract

Cells show a range of complex behaviors which might normally associated with cognition, including basic forms of learning such as habituation. The giant single-celled ciliate Stentor coeruleus is a pond-dwelling organism that attaches to substrates like pond plants and filter feeds by means of a ring of cilia at one end of the cell. Stentor cells can contract in response to mechanical stimulation, an effective escape mechanism but energetically costly. To avoid contracting due to random collision with non-threatening things like algae or pond plants, but still allow contraction when a large predator attacks, Stentor cells use a simple form of learning: when a cell is mechanically stimulated over and over again, it will gradually stop responding, representing a case of habituation. Our goal is to understand how a single cell learns, without a nervous system. We have developed automated devices for applying series of mechanical stimuli under computer control, which we are now using to test predictions of computational models for learning based on known or plausible biochemical mechanisms. Initial experiments showed that learning occurs in a stepwise manner, which suggested a simple two-state Markov model, however, subsequent experiments rule out this model. We next formulated a simplified biochemical model invoking known or suspected processes, which can account for all prior observations as well as new experiments designed to test the model. Based on this framework, we have investigated the molecular pathways involved in learning in Stentor, which suggests a role for CamKII kinase signaling in the learning process.

Foundation Models for Genomics: From Predictive Modeling to Biological Design, towards AI Scientists and Future

Ke Li

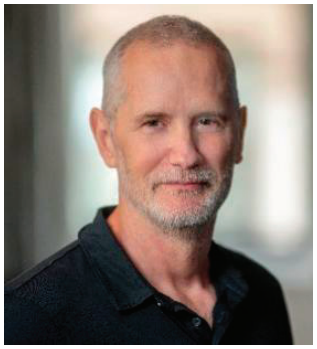
University of Exeter, UK

Abstract

What if we could build an intelligent machine capable of understanding and designing biological functions across the full diversity of life? Achieving this vision requires capturing the deep complexity of genomic information, far surpassing straightforward human intuition. Genomic language models (gLMs), analogous to large language models trained on human languages, offer a powerful, scalable paradigm for learning generalist representation of biological complexity directly from vast genomic sequences through self-supervised pre-training.

In this talk, I will first share our recent pioneer work on training large-scale foundations model, called PlantRNA-FM, to decode RNA regulatory codes of sequence-structure relationships across diverse plant species. PlantRNA-FM is interpretable, providing mechanistic insights into how nucleotide-level changes and structural motifs control key RNA functions, particularly translation efficiency. Building on this foundation, we recently developed OmniGenome, a novel model employing a structure-contextualized training strategy. OmniGenome uniquely enables accurate, bidirectional mappings between RNA sequences and their secondary structures, achieving state-of-the-art predictive performance. Interestingly, OmniGenome also demonstrates powerful capabilities in generative RNA design and reliable zero-shot structure prediction. To foster open science and reproducibility in gLM research and innovation, we recently released OmniGenBench, a modular benchmarking platform providing over 31 pre-trained gLMs, 123 curated genomic datasets (~42GB in total size), diverse accuracy metrics, and integrated interpretability tools. At the end, we will share our latest progress of building a virtual plant scientists through LLM-empowered AI agentic framework augmented by automated knowledge graphs.

Self-Organizing Dynamics in Cellular and Multi-Cellular Systems



Michael Shelley

Center for Computational Biology, Flatiron Institute & Courant Institute, NYU, USA

Abstract

A lot of fascinating and important biophysics involves the multiscale interaction of many active subunits with the result being dynamical, sometimes functional, large-scale structures. Canonical examples are the self-assembly of cellular spindles from cytoskeletal components, and the formation of system-scale flow structures in active suspensions. I'll discuss these and some non-canonical examples such as the collective dynamics of ultra-long sperm cells and the appearance of cell-spanning active flows in developing egg cells. All involve coarse-grained micro-macro models, active material stresses, transport, large-scale simulation, and very close interaction with quantitative experiment.

Algorithms for Creating Form

Sean Megason

Department of Systems Biology,
Harvard University, USA

Biography

Dr. Sean Megason is a professor and vice chair in the department of systems biology at Harvard University. He received a BS in Molecular Biology at the University of Texas at Austin in 1997. He performed his PhD in the laboratory of Andrew McMahon at Harvard on the control of organ size in the neural tube. He then went to Caltech for postdoctoral research in the laboratory of Scott Fraser on microscopy where he developed “in toto imaging”. Dr. Megason started his laboratory at Harvard Medical School in the Department of Systems Biology in 2008.

The Megason Lab uses imaging-based systems biology to elucidate the systems level principles of animal development. They are particularly interested in long standing problems in embryology whose understanding has defied molecular reduction namely patterning, morphogenesis, and size control.

The Megason Lab pursues these questions in the inner ear and spinal cord of zebrafish using a variety of techniques including microscopy, mathematical modeling, and molecular and mechanical perturbation.

Universal Laws Emerging from Competition between Genes for Limiting Resources

Jie Lin

Center for Quantitative Biology,
Peking University, China

Abstract

The expression levels of various genes across the genome are inherently interconnected due to limiting resources, such as ribosomes and RNA polymerase. Understanding the effects of competition among genes for these resources is essential for developing a systems-level understanding of gene expression. My presentation will explore some universal effects caused by gene competition and their implications for cell biology, including how cell volume influences protein concentrations, the buffering of RNA polymerase noise, and the fitness cost associated with expressing exogenous genes. Lastly, I will show that balancing the costs and benefits of gene expression results in a genome-wide scaling relationship between mRNA and protein levels.

Inhibition of Cellular Contractility Reveals Compression-enabled Mechanical Bistability during Epithelial Folding

Hanqing Guo[1,2,3], Michael Swan[4], Bing He[3]

***Short Abstract* — Apical constriction, driven by actomyosin contractility, generates in-plane contractile stress in epithelial tissues. How this leads to out-of-plane deep folding remains unclear. Combining optogenetics, laser ablation, tissue reconstruction and computational simulations, our results suggest that in-plane compression from the surrounding ectoderm facilitates mesoderm invagination by triggering buckling of the mesoderm epithelium.**

Keywords* — Contractility, Epithelium folding, *Drosophila

I. Purpose

THE epithelial folding is a fundamental morphogenetic process during development, driving tissue and organ formation through coordinated mechanical changes [1]. In *Drosophila melanogaster*, mesoderm invagination during gastrulation serves as a model system for studying how localized cellular forces, particularly actomyosin-driven apical constriction, lead to large-scale deformations [1-2]. However, recent studies suggest that epithelial folding is not solely the product of localized force generation but may also depend on mechanical contributions from surrounding tissues and emergent mechanical properties of the epithelium itself [3]. Understanding these dynamics requires an interdisciplinary approach that combines developmental biology with the principles of biophysics, soft matter mechanics, and systems-level modeling [4].

II. Results

How contractile stress generated near the apical surface of an epithelial sheet drives out-of-the-plane deep tissue folding remains unclear. To decipher the role of actomyosin contractility during epithelial folding, we combined experimental and computational approaches and suggested the essential function of surrounding ectoderm in driving buckling-like mesoderm folding.

A. Experimental Observation

By optogenetically inhibit actomyosin at precise developmental stages, we found that contractility is essential during an early "priming" phase but becomes dispensable once a critical shape transition is reached. This binary tissue response suggesting mesoderm is mechanical bistable during folding process [5]. Through 3D reconstruction, cell tracking, and laser ablation, we revealed that ectoderm is compressive and may therefore facilitate the buckling-like folding behavior of mesoderm [5-6].

B. Model Simulation

We adapted a previously published vertex model to recapitulate the experimental optogenetic inhibition observation. Our results showed that ectoderm compression is needed in generating the binary tissue response [5]. In order to simulate the cellular dynamics during the folding process, we also developed a 2D finite element model. The model simulation showed that including ectoderm compression, in addition to apical constriction, fully reproduced the mesoderm folding process observed *in vivo* [6].

III. Conclusion

Our results demonstrate that *Drosophila* mesoderm invagination arises from a composite mechanical mechanism integrating localized apical constriction within the mesoderm and global compression from the surrounding ectoderm, highlighting the synergy between local and global forces in epithelial folding.

References

[1] Martin, A. C. & Goldstein, B. Apical constriction: Themes and variations on a cellular mechanism driving morphogenesis. *Development* 141, 1987–1998 (2014).

[2] Martin, A. C., Kaschube, M. & Wieschaus, E. F. Pulsed contractions of an actin–myosin network drive apical constriction. *Nature* 457, 495–499 (2009).

[3] He, B., Khalil, A. & Martin, A. C. Apical constriction initiates new bud formation during monopodial branching of the *Drosophila* tracheal system. *Development* 141, 2467–2474 (2014).

[4] Lecuit, T. & Lenne, P.-F. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat. Rev. Mol. Cell Biol*. 8, 633–644 (2007).

[5] Guo, H., Swan, M. & He, B. Optogenetic inhibition of actomyosin reveals mechanical bistability of the mesoderm epithelium during *Drosophila* mesoderm invagination. *eLife* 11, e69082 (2022).

[6] Guo, H., Huang, S. & He, B. Evidence for a role of the lateral ectoderm in Drosophila mesoderm invagination. *Front. Cell Dev. Biol.* 10, 867438 (2022).

[1]School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. E-mail: guohanqing12@westlake.edu.cn

[2]Center for Interdisciplinary Studies, Westlake University, China.

[3]Department of Biological Sciences, Dartmouth College, U.S.

[4]Department of Molecular Biology, Princeton University, U.S.

Synthetic Conscription: Achieving Stable Labor Division with Rational Design of Gene Circuitry

Zhi Sun[1#], Ning Yang[2#], Baiyi Jiang[1], Kang Xia[2], Lei Zhang[2], Weijie Li[2], Shanwen Chen[3], Chunbo Lou[4*], Chao Tang[1,2*], Xiaojing Yang[1,2*]

Division of labor within populations is a game-changer for complex functions in synthetic biology, dramatically bolstering system efficiency, stability, and robustness. However, different functions inevitably introduce different burden, resulting different fitness of different functional subgroups, which destabilizes the system. Inspired by the real-world "conscription" systems, we propose a "synthetic conscription" strategy: by resetting the population to a homogeneous state for labor re-division after stimulus response, the whole population would collectively share the burden that division of labor might bring to subgroups, thereby ensuring long-term stability. To achieve this goal, we first formalized the steady-state and kinetic requirements for "synthetic conscription". We then exhaustively enumerated all feasible two-node transcriptional regulatory topologies satisfying these criteria. Leveraging parameter configurations enabling essential dynamics identified through this systematic search, we engineered a gene circuit implementing the optimal topology. The resulting circuit enables robust, high-fidelity division of labor across repeated stimulus cycles. Stochastic simulations coupled with global sensitivity analysis identified critical parameters governing the functional details of the conscription strategy, with experimental validation. Finally, integrating this circuit into probiotic Escherichia coli Nissle 1917 enabled targeted drug delivery and effective mitigation of inflammatory bowel disease (IBD). This proof-of-concept demonstration underscores the potential of synthetic conscription for chronic disease intervention.

I. Introduction

ADVANCING synthetic biology necessitates division of labor within cellular populations. However, functional heterogeneity inherently creates fitness disparities among specialized subpopulations, destabilizing long-time labor partitioning. This fitness-driven instability critically constrains executing complex functions synthetically[1-4].Here, inspired by real-world conscription systems, we propose a 'synthetic conscription' strategy. This approach resets the cellular population to homogeneity after each functional execution and re-establishes division of labor, ensuring equitable burden-sharing among all individuals to achieve sustained and stable division of labor.

II RESULTS

A. Topology enumeration

To realize the core mechanism of the "conscription" strategy, we enumerated all 81 possible transcription topologies that satisfy the following functional requirements: 1) exhibit two distinct behaviors: monostability and bistability. 2) one state hould be responsive to external stimuli, while the other remains unresponsive. 3) The responsive state ultimately revert to the homogeneous state after prolonged.

B.Genetic Circuit Construction and Characterization

A synthetic transcriptional circuit based on well-tested genetic parts was constructed in *E. coli*. The dynamic behavior of the circuit fully satisfies the desired goals.

B. Autonomous Therapeutic Systems for IBD

We applied our division-of-labor strategy to drug delivery by integrating the circuit into a probiotic chassis, adding lysis modules and therapeutic agents. This approach successfully achieved drug delivery, which we validated in a mouse IBD model.

Acknowledgements: This work was supported NSFC 12090053 and NSFC 32088101.

[1]Center for Quantitative Biology, Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China. Emails: tangc@pku.edu.cn, xiaojing_yang@pku.edu.cn

2 Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies, Chengdu 610095, China.

3 Gastro-Intestinal Surgery, Peking University First Hospital, Beijing 100034, China.

4 State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. Email: cb.lou@siat.ac.cn

References

[1] Kong, W., Meldgin, D. R., Collins, J. J. & Lu, T. Designing microbial consortia with defined social interactions. Nat. Chem. Biol. 14, 821–829 (2018).

[2] Gutiérrez Mena, J., Kumar, S. & Khammash, M. Dynamic cybergenetic control of bacterial co-culture composition via optogenetic feedback. Nat. Commun. 13, 4808 (2022).

[3] Glass, D. S., Bren, A., Vaisbourd, E., Mayo, A. & Alon, U. A synthetic differentiation circuit in Escherichia coli for suppressing mutant takeover. Cell 187, 931-944.e12 (2024).

A cellular solution to a robotics problem

Zitong Jerry Wang[1], Matt Thomson[2]

***Short Abstract* — Localizing and navigating toward a signal source in noisy environments remains a fundamental challenge in robotics. Similarly, cells in tissues must navigate noisy, fragmented molecular gradients, shaped by fluid flow and extracellular matrix interactions. We show that cells can perform source localization using a biophysical implementation of a computational algorithm called Bayes filtering. Specifically, the spatial distribution of molecules encodes a probability distribution over source location, and intracellular transport processes update this distribution. Unlike conventional Bayes filtering, the cellular implementation can adjust the weight of past observations based on current environmental signals. When translated to traditional robotics algorithms, this signal-aware learning rate significantly improves navigation performance in high-noise conditions, revealing how biological mechanisms can advance engineered systems.**

***Keywords* — Cell navigation, bayesian filtering, robotics**

I. Background

Localizing and navigating toward a target in a complex, noisy environment is a fundamental challenge in robotics. Over the past several decades, Bayes filtering and its variants Kalman filters, particle filters, and smoothing-based methods have become standard tools for robotic state estimation and navigation tasks ranging from self-driving cars to drone flight control [1]. These Bayesian approaches iteratively update a belief distribution over the target location using sensor measurements. Bayes filtering leads to optimal estimates of the true target location in the sense of having the smallest mean squared error under standard assumptions. Despite the success of Bayes filtering, high-noise conditions can still degrade performance, especially in unstructured or rapidly changing environments [2].

Cells can efficiently localize to ligand sources in complex tissue environments, but the mechanisms remain unclear. In tissues, extracellular matrix (ECM) binding and interstitial fluid flow break ligand gradients into irregular, fragmented patches [3]. For example, CCL21 secreted from lymphatic vessels in the dermal interstitium is transported by fluid flow and captured by a non-uniform ECM network, forming a stable, reticulated pattern with local concentration peaks [3]. As a result, imaging of intact tissues reveals that cells below 40μm in diameter often experience local gradients that do not reliably point toward the true source. Yet, despite this misalignment, cells can still navigate efficiently. This suggests cells use strategies beyond simple gradient-following, which risks trapping them at local signal peaks. Meanwhile, random walks are similarly inefficient over long distances due to the slow speed of cell migration. Understanding how cells navigate such noisy, complex environments may suggest new strategies to improve robotic algorithms.

Recent observations suggest dynamic spatial rearrangement of surface receptors may be important for effective navigation. For instance, receptors such as TrkB, Robo1, and PlxnA1 in neuronal growth cones reorganize according to local ligand distributions, and inhibiting their rearrangement impairs directional guidance [4, 5]. Similarly, blocking CCR2 receptor redistribution on mesenchymal stem cells, without changing its overall expression, severely disrupts targeted migration to injured muscle tissues [6]. These observations suggest that receptor dynamics, and not just expression, can be pivotal for robust navigation.

II. Conclusion

In this work, we show that dynamic receptor rearrangement can function as a biophysical implementation of the Bayes filtering algorithm optimized for tissue navigation. The spatial distribution of receptors encodes a probability distribution over source location and intracellular transport processes can update this distribution, thereby implementing key steps of the algorithm. Unlike conventional Bayes filtering, the cellular Bayes filter can adjust the weight of past observations in response to current signals. We demonstrate, through simulations, that translating this adaptive mechanism back into traditional robotics algorithms yields significantly improved source localization under high-noise conditions. Our results illustrate how biological mechanisms can inspire more robust navigation solutions in engineered systems.

References

[1] V. Fox, et al., Bayesian filtering for location estimation. IEEE Pervasive Computing 2 (3), 24–33 (2003).
[2] J. Cao et al., A review of object tracking methods: From general field to autonomous vehicles. Neurocomputing p. 127635 (2024).
[3] M.Weber et al., Interstitial dendritic cell guidanceby haptotactic chemokine gradients. Science 339 (61 17), 328–332 (2013).
[4] C. Guirland et al., Lipid rafts mediate chemotropic guidance of nerve growth cones. Neuron 42 (1), 51–62 (2004).
[5] A. Pignata et al., A spatiotemporal sequence of sensitization to slits and semaphorins orchestrates commissural axon navigation. Cell reports 29 (2), 347–362 (2019).
[6] F. Belema-Bedada et al., Efficient homing of multipotent adult mesenchymal stem cells depends on FROUNT-mediated clustering of CCR2. Cell stem cell 2 (6), 566–575 (2008).

[1]Center for Interdisciplinary Studies, School of Science, Westlake University, Hangzhou, P. R. China. E-mail: jerry@westlake.edu.cn
[2]Division of Biology and Biological Engineering, Caltech, Pasadena, USA. E-mail: mthomson@caltech.edu

Nonequilibrium Thermodynamics in Living Systems: Towards Aanswering Schrödinger's Question

Yuhai Tu

Center for Computational Biology &
Center for Computational Neuroscience,
Flatiron Institute,
New York, NY10010, USA

Abstract

In his influential book "What is Life?" published in 1944, Schrödinger, one of the founding fathers of quantum mechanics, made the simple statement that life feeds on 'negative entropy', which he meant free energy. Schrödinger's general statement that life costs free energy is certainly true but almost trivial – we all know we need to eat to survive. However, one of the real puzzle about life is how living systems use the free energy (negative entropy) they harvest from their environment to carry out vital biological functions to survive, to reproduce, and simply to be alive. Of course, this puzzle is impossible to solve at the time of Schrödinger's book and has thus been largely forgotten or ignored.

Now, nearly a century later, molecular mechanisms underlying many cellular functions have been revealed, time is ripe to revisit Schrödinger's question. In this talk, by using specific examples in cellular biology ranging from John Hopfield's famous kinetic proofreading mechanism to our own recent works on ultrasensitive switch of bacterial flagellar motor and synchronization of molecular clocks, we will try to address a few related general questions inspired by Schrödinger's original query: 1) How do biology systems uses free energy to enhance their performance? 2) How does the thermodynamic energy cost limit/constrain the performance of the underlying biochemical networks? 3) What are the design principles for biochemical networks to achieve highly efficient functions?

Overflow metabolism originates from growth optimization and cell heterogeneity

Xin Wang

School of Physics, Sun Yat-sen University, Guangzhou, China

Abstract

A classic problem in metabolism is that fast-proliferating cells use seemingly wasteful fermentation for energy biogenesis in the presence of sufficient oxygen. This counterintuitive phenomenon, known as overflow metabolism or the Warburg effect, is universal across various organisms. Despite extensive research, its origin and function remain unclear. Here, we show that overflow metabolism can be understood through growth optimization combined with cell heterogeneity. A model of optimal protein allocation, coupled with heterogeneity in enzyme catalytic rates among cells, quantitatively explains why and how cells choose between respiration and fermentation under different nutrient conditions. Our model quantitatively illustrates the growth rate dependence of fermentation flux and enzyme allocation under various perturbations and is fully validated by experimental results in Escherichia coli. Our work provides a quantitative explanation for the Crabtree effect in yeast and the Warburg effect in cancer cells and can be broadly used to address heterogeneity-related challenges in metabolism.

References

[1] Warburg et al., Biochem. Z. 152, 309 (1924)..

[2] Vander Heiden et al., Science 324, 1029 (2009).

[3] Basan et al., Nature 528, 99 (2015).

[4] Liberti and Locasale, Trends in Biochemical Sciences 41, 211 (2016).

[5] Shen et al., Nat. Chem. Biol. (2024).

[6] Wang, X., eLife 13, RP94586 (2024)

Email: wangxin36@mail.sysu.edu.cn

Discovery of highly bioactive peptides through hierarchical structural information and molecular dynamics simulations

Shu Li[1,#], Lu Peng[2,#], Liuqing Chen[3,#], Zengru Di[2], Yu Liu[2,*]

***Abstract*— We developed PepHiRe, an innovative computational framework that uses hierarchical structural information to design peptides with high bioactivity. Our method leverages the principles of Ladderpath Approach (which is under the umbrella of Algorithmic Information Theory) to generate novel peptides targeting Myeloid Cell Leukemia-1 (MCL-1), a protein crucial in various cancers. Using a limited dataset of eight wild-type BH3 peptides, PepHiRe efficiently produces peptides with substantially improved binding affinities, achieving IC50 values between 28.13 and 167.42 nM. This study underscores the potential of integrating advanced computational tools like multi-conformational docking and molecular dynamics simulations in drug design, enhancing both the accuracy and efficiency of therapeutic peptide development.**

***Index Terms*— Peptide inhibitors, Ladderpath theory, Molecular dynamics simulations, Bioactive peptides**

I. Introduction

This study addresses the challenges of designing peptide inhibitors for protein-protein interactions (PPIs), crucial in cellular functions and disease states when dysregulated[1]. We develop PepHiRe, a novel approach using the Ladderpath Approach (which is under the umbrella of Algorithmic Information Theory) to generate peptides from minimal data efficiently[2], [3].

PepHiRe was used to create BH3-like peptides targeting Myeloid Cell Leukemia-1 (MCL-1), an anti-apoptotic protein involved in various cancers[4]. Using a small dataset, we developed peptides that demonstrated high affinity for MCL-1 with IC50 values from 28.13 to 167.42 nM, suggesting significant therapeutic potential.

This method advances peptide-based drug design, providing a scalable approach to generate therapeutically relevant peptides.

[#]These authors contributed equally: Shu Li, Lu Peng, Liuqing Chen

[*]Email: yu.ernest.liu@bnu.edu.cn

We thank Dr. Weiyi Qiu from Swarma Research for her fruitful discussions and insightful comments.

[1]Centre of Artificial Intelligence driven Drug Discovery, Faculty of Applied Science, Macao Polytechnic University, Macao SAR, 999078, China.

[2]Department of Systems Science, Faculty of Arts and Sciences, Beijing Normal University, Zhuhai, 519087, China.

[3]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China.

II. Methods and Materials

A. Peptide Generation Using PepHiRe and Laddergraph

The PepHiRe algorithm, integrated with laddergraph calculations from reference [5], was utilized for generating and screening peptides. This process involved visualizing hierarchical structures of peptide sequences to inform design choices.

B. Molecular Dynamics and Energy Calculations

Molecular dynamics simulations were conducted using GROMACS 2022 with the AMBER99SB-ILDN force field, providing insights into peptide-MCL-1 complex stability. Binding free energies were calculated using gmx-MMPBSA, integrating molecular mechanics with solvation effects.

C. Protein Handling and Competitive Assay

Mcl-1 was expressed in E. coli, purified using a Ni-NTA column, and employed in competitive binding assays using a TR-FRET-based technique to measure peptide affinities and determine IC50 values.

III. Conclusion

Our study develops PepHiRe, a novel peptide design strategy using the Ladderpath approach to generate peptides with high affinity for Myeloid Cell Leukemia-1 (MCL-1), showcasing potential therapeutic benefits. PepHiRe operates independently of large datasets, enhancing its flexibility for designing peptides for lesser-studied targets. This method combines evolutionary insights with practical drug design, promising new avenues for developing peptide-based therapeutics.

References

[1] K. Titeca, I. Lemmens, and J. Tavernier, "Discovering cellular protein-protein interactions: Technological strategies and opportunities," *Mass Spectrom Rev*, vol. 38, pp. 79–111, 2019.

[2] Y. Liu, C. Mathis, and M. D. Bajczyk, "Exploring and mapping chemical space with molecular assembly trees," *Sci Adv*, vol. 7, p. eabj2465, 2021.

[3] Y. Liu, Z. R. Di, and P. Gerlee, "Ladderpath approach: How tinkering and reuse increase complexity and information." *Entropy*, vol. 24, p. 1082, 2022.

[4] L. Gambini, P. Udompholkul, and C. Baggio, "Design, synthesis, and structural characterization of lysine covalent bh3 peptides targeting mcl-1," *J Med Chem*, vol. 64, pp. 4903–4912, 2021.

[5] Z. Zhang, C. Liu, and Y. Zhu, "Evolutionary tinkering enriches the hierarchical and interlaced structures in amino acid sequences." *Preprint at Research Square*, pp. 1–26, 2024.

Quantifying system-environment synergistic information by effective information decomposition

Mingzhe Yang [1], Linli Pan [1], and Jiang Zhang [1,2]

Short Abstract — **Living systems maintain their structural and functional stability while adapting to environmental changes. Existing frameworks such as self-organization theory can not measure system-environment interaction at the causal level. In this article, we propose a new causal indicator to measure the ability of a system to flexibly respond to the environment. We showed this indicator satisfies the partial information decomposition framework's axiom system. It depends on entanglement between system-environment variables and noise magnitude. Through experiments on cellular automata(CA), random Boolean networks, and real gene regulatory networks(GRNs), we validate the indicator's relationship with CA types and Langton parameters, while demonstrating that feedback loops in GRNs exhibit exceptional flexibility. We also combined machine learning technology to prove that this framework can be applied in the case of unknown dynamics.**

Keywords — **synergy, flexibilty, effective information, partial information decomposition, gene regulatory networks**

I. Background

WHAT distinguishes living systems from ordinary systems? The self-organization theory first address this question [1]. People believe a living system must have the ability of self-organization. However, life is more than just an open system with self-organization. More importantly, they can maintain the stability of their own structure and function when faced with diverse and changing environments. Many papers have proposed this kind of adaptive systems in various fields [2-3].

To quantify this unique property of life, information theory serves as a vital tool for quantifying living systems [4]. However, few indicators currently exist to measure a system's capacity to handle environmental changes. The information theory of individuality [5] aims to assess the individuality of living systems through an information dynamics perspective. Yet, calculating this metric (mutual information) depends on the state distribution of observed data, which hinges on the system's initial conditions, environment, and duration of the dynamic process. Tononi and Hoel et al. proposed effective information (EI) measured at the causal level [6]. Causal mechanisms here are defined by Markovian transition probability matrices (TPM), with EI serving as a function of TPM [7]. EI measures the strength of causal effects in dynamics, and differences in EI between macro and micro dynamics quantify emergence in complex systems [6].

II. Results

We define flexibility at the causal level by decomposing EI, characterizing a system's ability to maintain its structure and function when facing environmental changes. Through mathematical proofs, we show that flexibility equals the system's synergistic information when the system and environment are coupled. Our numerical experiments demonstrate the correlation between flexibility and the edge of chaos in CA. We also apply this indicator to GRNs [8], revealing the significance of feedback loops (FBLs) and the responses of highly synergistic systems' steady states to environmental changes. FBLs are instrumental for biological systems to react to environmental changes [8]. Additionally, our machine learning experiments show that causal mechanisms can be identified using machine learning when only data is available.

III. Conclusion

Our results reveal that flexibility peaks at the edge of chaos in CA and identifies FBLs as biological flexibility hubs.

References

[1] S. A. Kauffman, The origins of order: Self-organization and selection in evolution (Oxford University Press, 1993).
[2] J. H. Holland and J. H. Miller, Artificial adaptive agents in economic theory, The American economic review 81, 365 (1991).
[3] J. H. Holland, Complex adaptive systems, Daedalus 121, 17 (1992).
[4] S. Bartlett, A. W. Eckford, M. Egbert, M. Lingam, A. Kolchinsky, A. Frank, and G. Ghoshal, The physics of life: Exploring information as a distinctive feature of living systems, arXiv preprint arXiv:2501.08683 (2025).
[5] D. Krakauer, N. Bertschinger, E. Olbrich, J. C. Flack, and N. Ay, The information theory of individuality, Theory in Biosciences 139, 209 (2020).
[6] E. P. Hoel, L. Albantakis, and G. Tononi, Quantifying causal emergence shows that macro can beat micro, Proceedings of the National Academy of Sciences 110, 19790 (2013).
[7] B. Yuan, J. Zhang, A. Lyu, J. Wu, Z. Wang, M. Yang, K. Liu, M. Mou, and P. Cui, Emergence and causality in complex systems: A survey of causal emergence and related quantitative studies, Entropy 26, 108 (2024).
[8] U. Alon, An introduction to systems biology: design principles of biological circuits (Chapman and Hall/CRC, 2019).

[1] School of Systems Science, Beijing Normal University, Beijing 100871, China
[2] Swarma Research, Beijing 102300, China

Transcriptional Condensates Encode a Golden Mean to Optimize "Enhancer-Promoter" Communication across Genomic Distances

Tao Zhu[1], Chunhe Li[1], and Xiakun Chu[2]

Short Abstract — **Enhancer-promoter (E-P) communication is essential for gene activation, yet its effectiveness over large genomic separation is increasingly attributed to transcription factor (TF) phase-separated condensates. Using polymer modeling, we systematically vary both the TF condensation strength and the E-P genomic distance to quantify their combined impact on E-P interactions within chromatin domains. We find that E-P contact efficiency peaks at a moderate degree of TF condensation, termed as the "golden mean" where TF clusters are large enough to promote bridging yet not so dense as to induce competitive interference. Acting as a tunable "rheostat", TF condensates buffer E-P contact against increasing genomic separation, in agreement with experimental findings, and can even equalize contact efficiency across distant loci. Our study resolves how stochastic chromatin dynamics achieve precise transcriptional control, highlighting the necessity of balanced TF condensation for robust long-range genomic communication.**

Keywords — **Genome folding, 4D genome, Coarse-grained model, Molecular dynamics simulation.**

I. Introduction

Enhancers play a central role in the mammalian genome by directing the spatiotemporal activation of genes through interactions with promoters, thereby ensuring precise control over developmental processes and cell-type-specific expression programs. Although enhancers can reside several kilobases to over a megabase away from their target promoters along the linear genome, productive transcription often requires that these distantly located elements come into close proximity and form physical contacts within the three-dimensional space of the nucleus. Yet how long-range enhancer-promoter (E-P) communication is accurately established and maintained remains poorly understood.

Over the past decade, transcription factors (TFs) have been well-recognized as key regulators of gene expression not only through DNA binding, but also by compartmentalizing biochemical reactions into high-concentration hubs, referred to as phase-separated condensates and formed via multivalent interactions among themselves. Owing to the inherent capacity to bind and cluster cooperatively, TFs are naturally equipped to promote E-P communication across large genomic distances. Indeed, a recent study revealed a weaker-than-expected dependence of E-P contact efficiency on linear genomic separation, which may be attributed to the organizing influence of TF condensates. However, to what extent TF phase separation modulates the sensitivity of E-P interactions to genomic distance remains to be elucidated.

Although mounting evidence has confirmed the positive role of TF condensates in mediating E-P interactions and regulating gene activation, emerging studies have revealed that excessive TF phase separation can suppress transcriptional output. This duality highlights that the strength of TF condensation must be finely tuned to support optimal transcriptional regulation. An imbalance in this process may compromise gene expression dynamics and cellular function. Despite these insights, the quantitative relationship between the size of TF condensates and their influence on target E-P interactions remains poorly characterized.

II. Conclusion

Here, we develop a polymer-based chromatin model and perform molecular dynamics simulations to dissect how TF phase separation and E-P genomic separation jointly govern E-P communication. We find that E-P contact probability depends non-monotonically on both TF abundance and TF-TF interaction, with intermediate levels of condensation maximizing contact efficiency across genomic distances. Increasing TF abundance and interaction strength progressively reduces and ultimately eliminates the disparity in contact probability among all genomic separations. By quantifying TF clustering around E-P loci, we establish a quantitative link between condensate size and its influence on E-P contact, identifying two distinct regimes: under weak TF-TF interactions, contact probability rises monotonically with condensate size, whereas under strong interactions, a non-monotonic "golden mean" relationship emerges. Finally, analysis based on the free energy landscape reveals that optimal phase separation reflects a balance between TF-mediated bridging that stabilizes the E-P complex and competitive occupancy that disrupts it. Together, these findings provide a mechanistic framework for understanding how finely tuned TF phase separation consistently ensure effective transcription across genomic distances.

[1]Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China.

[2]Advanced Materials Thrust, Function Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, Guangdong 511400, China. E-mail: xiakunchu@hkust-gz.edu.cn

Non-reciprocity across scales in bacterial mixtures

Yongfeng Zhao[1]

***Short Abstract* — Bacteria often experience mediated interactions that are not constrained by Newton's third law and are therefore inherently non-reciprocal. Here, we consider bacterial mixtures interacting via quorum sensing. By explicitly connecting microscopic and macroscopic dynamics, we show that non-reciprocity may fade upon coarse graining, leading to large-scale equilibrium descriptions. This allows us to account quantitatively for the rich behaviors observed in microscopic simulations including phase separation, demixing, and multi-phase coexistence. We also derive the condition under which non-reciprocity survives coarse graining, leading to a wealth of dynamical patterns. Our work demonstrates that the fate of non-reciprocity across scales is an important question.**

***Keywords* — Active matter, quorum sensing, pattern formation, non-reciprocal interactions**

I. Introduction

MEDIATED *N*-body interactions like chemotaxis, quorum sensing, or hydrodynamic interactions are often non-reciprocal: the effective "force" from particle *i* to *j* may not equal the opposite of the force from *j* to *i*. Non-reciprocal interactions are prevalent in biological systems and have generated significant interest due to the rich and complex phenomena they produce [1].

Quorum sensing refers to the mechanism by which cells sense the population density of their neighbors, typically through the release and detection of short-lived diffusive signaling molecules. It is a widely used module in synthetic biology and genetic engineering. When quorum sensing is coupled with motility control in bacteria, the effective interactions become non-reciprocal. However, this non-reciprocity at the particle scale does not necessarily persist at the macroscopic level.

It is known that motility regulation in self-propelled particles via quorum sensing can lead to motility-induced phase separation [2], which evolves into pattern formation in growing bacterial systems [3]. For instance, ring-like patterns have been observed in genetically engineered *E. coli* colonies [4–5]. Reciprocal motility regulation between two populations of engineered *E. coli* can result in either segregation or mixing, as seen in experiments [6].

These findings offer new and versatile mechanisms for self-organization in single- and multi-component bacterial systems, extending beyond classical Turing-type pattern formation. Interestingly, under certain conditions, these systems can be quantitatively mapped onto equilibrium models, which are inherently reciprocal at the macroscopic level [7–8]. When such an equilibrium mapping exists, the microscopic non-reciprocity does not survive the coarse-graining process.

[1]Center for Soft Condensed Matter Physics and Interdisciplinary Research, School of Physical Science and Technology, Soochow University, 215006, Suzhou, China. E-mail: yfzhao2021@suda.edu.cn

II. Outline of the Talk

In this talk, I will present the explicit conditions under which a generic microscopic multi-component system can be mapped onto an equilibrium theory at the macroscopic scale. This mapping is closely related to the preservation of macroscopic time-reversal symmetry, which can be quantified via the entropy production rate. Applying this criterion to multi-component quorum-sensing active particles allows us to quantitatively predict the phase diagram of phase-separating systems using the common tangent construction, without fitting parameters.

Furthermore, I will show how a rich variety of dynamic patterns emerge when motility regulations strongly violate reciprocity [8–9]. Our work illustrates that one can maintain an analytical framework while ascending the complexity ladder, allowing for a principled understanding of the emergent behavior in self-propelled bacterial systems.

References

[1] Suropriya Saha, Jaime Agudo-Canalejo, Ramin Golestanian (2020) Scalar Active Mixtures: The Nonreciprocal Cahn-Hilliard Model. *Phys. Rev. X* **10**, 041009.

[2] Jérémy O'Byrne, Alexandre P Solon, Julien Tailleur, Yongfeng Zhao (2023) *Out-of-Equilibrium Soft Matter: Active Fluids*. Royal Society of Chemistry, Ch. 4.

[3] Michael E. Cates, Davide Marenduzzo, Ignacio Pagonabarraga, Julien Tailleur (2010) Arrested phase separation in reproducing bacteria: a generic route to pattern formation? *PNAS* **107**, 11715-11720.

[4] Chenli Liu, et al. (2011) Sequential Establishment of Stripe Patterns in an Expanding Cell Population. *Science* **334**, 238-241.

[5] Xiongfei Fu, Lei-Han Tang, Chenli Liu, Jian-Dong Huang, Terence Hwa, Peter Lenz (2012) Stripe Formation in Bacterial Systems with Density-Suppressed Motility. *Phys. Rev. Lett.* **108**, 198102.

[6] Agnese I. Curatolo, Nan Zhou, Yongfeng Zhao, Chenli Liu, Adrian Daerr, Julien Tailleur, Jian-Dong Huang (2020) Cooperative pattern formation in multi-component bacterial systems through reciprocal motility regulation, *Nat. Phys.* **16**, 1152–1157.

[7] Alexandre P Solon, Joakim Stenhammar, Michael E Cates, Yariv Kafri, Julien Tailleur (2018) Generalized thermodynamics of motility-induced phase separation: phase equilibria, Laplace pressure, and change of ensembles. *New J. Phys.* **20** 07500.

[8] Alberto Dinelli, Jérémy O'Byrne, Agnese Curatolo, Yongfeng Zhao, Peter Sollich, Julien Tailleur (2023) Non-reciprocity across scales in active mixtures. *Nat. Commun.* **14**, 7035.

[9] Yu Duan, Jaime Agudo-Canalejo, Ramin Golestanian, Benoît Mahault (2023) Dynamical Pattern Formation without Self-Attraction in Quorum-Sensing Active Matter: The Interplay between Nonreciprocity and Motility. *Phys. Rev. Lett.* **131**, 148301.

Quorum sensing-based interactions within microbial communities

Changchang Liang[1], Danlei Chen[2], Chao Zou[1], Aidong Yang[3], Jianjun Qiao[1,2,*], Shengbo Wu[1,2,*]

***Short Abstract* — Quorum Sensing (QS) has emerged as a pivotal research domain in microbial ecology. Extensive investigations have elucidated the nature of QS systems across microbial communities, necessitating a systematic summary and profound exploration of their ecological frameworks. QS-based higher-order interactions, which are crucial for deciphering microbial community dynamics, have remained largely unexplored. Therefore, in this review, we present the first comprehensive summary of QS-based interactions in one-way, two-way, and higher-order modes to shed light on how these interactions shape the structures and regulate the functions of various microbial communities, which contribute to various consortia-based applications in different fields.**

***Keywords* — quorum sensing; microbial interactions; cell communication; microbial social behavior; microbial ecosystem; higher-order interactions.**

I. Purpose

In recent years, researchers have discovered that quorum sensing (QS) plays a significant role in microbial interactions (1). QS, as a widespread communication mechanism both within and between bacterial colonies, enables bacteria to coordinate community behaviors, such as biofilm formation and toxin production, by sensing the concentration of specific signal molecules (2). Studies on *Pseudomonas aeruginosa* and *Burkholderia thailandensis* showed that these bacteria display deceptive behaviors within their communities. Additionally, the research on communities containing Bifidobacteria and *Escherichia coli* or *Citrobacter rodentium* identified the phenomenon of interference in iron acquisition regulated by QS (3). Therefore, QS is crucial not only for the ecological stability of individual bacteria but also for the ecology and evolution of microbial communities comprising different strains. This highlights the capacity of bacteria to function as organized groups rather than as solitary individuals, revealing complex social behaviors among microbes (4).Given the diversity of ecological relationships within colonies, a more comprehensive and up-to-date summary is required to analyze the ecological interactions of QS within and between microbial colonies.

Acknowledgements: This work was funded by National Key Research and Development Program of China (2024YFA0920200).

[1]School of Synthetic Biology and Biomanufacturing,Tianjin University, Tianjin 300072, China. E-mail: wushengbo@tju.edu.cn

[2]Zhejiang Institute of Tianjin University, Shaoxing, 312300, China

[3]Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

II. Summary

In the review, to give a more comprehensive understanding on the QS-based interactions, we will systematically summarize and discuss various QS-based mutualisms, including neutrality, symbiosis (commensalism and amensalism), competition, predation (parasitism), cooperation, and higher-order interactions to explore the social interactions regulated by the QS system. Specifically, we first examine the QS-based interactions by investigating neutral mutualisms, which primarily involve regulatory behaviors within their own population. Then, we review the one-way interactions between cheaters or destroyers and mutualists in the context of symbiosis within microbial communities. Furthermore, within the frameworks of competition, predation, and cooperation, we systematize the two-way interaction relationships between different bacterial species. We also address QS-based higher-order interactions among microbes to improve our understanding of the complex interaction patterns in microbial communities and their roles in ecosystem function and stability. Finally, we discuss the current challenges and opportunities for the further development of QS-based communications..

III. Conclusion

This review provides a comprehensive summary of QS-based microbial interactions, offering an in-depth analysis of one-way, two-way, and higher-order interaction modes. Through the diversity and complexity of the QS system, bacteria can efficiently utilize resources and co-evolve, securing a competitive advantage in dynamic environments. QS-based higher-order interactions were also found to be prevalent, influencing both intra- and interspecies dynamics by shaping the stability and function of microbial communities through complex signal transduction networks.

References

[1] Whiteley M, Diggle SP, Greenberg EP. Progress in and promise of bacterial quorum sensing research. Nature 551, 313-320 (2017).

[2] Wu S, Qiao J. Vertical and horizontal quorum-sensing-based multicellular communications. Trends Microbiol. 29, 1130-1142 (2021).

[3] Wu S, et al. Design and analysis of quorum sensing language “interpreter” ecosystem for microbial community. Chem. Eng. J. 496, 153148 (2024).

[4] Wu S, Zhou Y, Dai L, Yang A, Qiao J. Assembly of functional microbial ecosystems: From molecular circuits to communities. FEMS Microbiol. Rev. 48, fuae026 (2024).

Cytoplasmic mRNA localization defines protein distribution in a mitochondrial graph-like network

Haohang Lin[1], Abdul Haseeb Khan[1], Chenxi Liu[1], Qinghe Wang[1], Ercan Kuruoglu[1], Tatsuhisa Tsuboi[1]

ATP synthase exhibits a striped distribution within mitochondria, corresponding to the cristae structure. It remains unclear whether this pattern arises primarily from the multimerization of ATP synthase or from the formation of cristae themselves. We employed quantitative image analysis, molecular kinetics modeling, and molecular dynamics to elucidate the regulatory mechanism underlying this spatial organization. We found that the pattern generated by protein import from randomly localized cytoplasmic mRNAs closely matches the distribution patterns observed in live-cell imaging. This study demonstrates how ATP synthase subunits become spatially organized and evenly distributed throughout the mitochondrial network in a way that supports their functional roles.

***Keywords* — mitochondria, cristae, ATP, molecular kinetics, molecular dynamics**

I. Purpose

ATP synthase, a pivotal enzyme in cellular energy production, is a rotary machine anchored in the inner mitochondrial membrane. While its structural components are well characterized, the processes of expression, transport, and assembly remain incompletely understood. Moreover, inconsistencies in the molecular abundance of ATP synthase subunits suggest the possible formation of sub-complexes with distinct stoichiometries.

II. Preparation of Abstracts

To investigate this, we analyzed the distribution of ATP synthase components within the mitochondrial tubular network using fluorescence intensity data and evaluated the patterns with a graph-based model. Surprisingly, we found that ATP3p, ATP4p, and ATP5p show distinct localization at the edge center and ends of mitochondrial tubules, whereas ATP1p, ATP2p, and ATP7p are evenly distributed throughout the network. Additionally, ATP3p and ATP4p were more concentrated at three-way junctions than at edge centers when the network formed such nodes. These findings challenge the traditional view that ATP synthase complexes accumulate primarily in cristae regions. Instead, they suggest that protein distribution is regulated by the physiological characteristics of the mitochondrial network. Live-cell imaging further revealed that mRNAs encoding ATP synthase components are randomly distributed along the mitochondrial tubules, supporting a model of co-translational protein import. To test this, we developed a stochastic computational model to simulate protein assembly within the mitochondrial network.

III. Conclusion

The model showed that the pattern formed by protein import driven by randomly localized cytoplasmic mRNAs closely matches the distribution patterns observed in live-cell experiments. Our study demonstrates how ATP synthase components are spatially organized and broadly distributed across the mitochondrial network in a way that supports their functional roles. These findings enhance our understanding of the assembly and distribution mechanisms of mitochondrial ATP synthase and propose a general framework for protein localization within graph-like mitochondrial structures.

[1] Shenzhen International Graduate School, Tsinghua University. E-mail: ttsuboi@sz.tsinghua.edu.cn

Exact First-Passage Time Distributions from Time-Dependent Solutions of the Chemical Master Equation

Changqian Rao[1,2], David Waxman[3,4,5,6], Wei Lin[1,2,3], and Zhuoyi Song[4,5,6*]

***Short Abstract* — This work presents the first exact solution for the full first-passage time distribution (FPTD) in nonlinear biochemical networks with second-order reactions like A+B→C. FPTD is crucial for understanding the timing and reliability of biochemical processes but has been difficult to compute due to its nonlinearity. Our method overcomes this challenge, surpassing existing approximations with an over two-fold improvement in accuracy, marking a significant advance in the theoretical study of stochastic biochemical dynamics.**

***Keywords* — nonlinear biochemical networks; chemical master equations; arbitrary initial conditions**

I. Extend Abstract

The first passage time (FPT) is a key measure in stochastic biochemical networks, representing the time it takes for a specific reaction or state to occur for the first time. While the mean FPT has traditionally been used to summarize system behavior, the full FPT distribution (FPTD) provides more detailed insight into the timing, reliability, and efficiency of biological processes such as gene activation, molecular binding, and signal transduction[1].

Computing exact FPTDs in nonlinear biochemical systems is highly challenging due to the stochastic and discrete nature of molecular interactions. The chemical master equation (CME) is the standard framework used to model these systems, but its time-dependent solutions are generally unavailable for nonlinear networks, especially those involving second-order reactions like A+B→C[2].

Existing approaches often rely on time-consuming stochastic simulations or approximate analytical methods[3-5]. These either obscure the causal effects of system parameters or introduce significant errors, particularly under high reaction rates or complex network structures.

Previously, we introduced an exact method to derive FPTDs in nonlinear networks with second-order reactions, assuming initial conditions followed a Poisson-product distribution[6]. Next, we extend our framework to arbitrary initial discrete distributions, significantly broadening its applicability to real-world biological systems, where molecular states are often non-Poissonian[7].

We validate our method through simulations and applications to biological models. The results show high accuracy across various reaction regimes, including time-varying rates. In example applications, such as gene regulation and Ras activation pathways, our method captures complex FPTDs more accurately than traditional approximations, confirming its robustness and biological relevance.

In summary, we present a general and exact analytical framework for computing FPT distributions in nonlinear biochemical systems with second-order reactions and arbitrary initial conditions. This method enhances both theoretical understanding and practical modeling of stochastic dynamics in biology, with potential applications in systems biology and synthetic circuit design.

[1] School of Mathematical Sciences, Fudan University, Shanghai, 200433, China
[2] Research Institute of Intelligent Complex Systems, Fudan University, Shanghai, 200433, China
[3] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, 200433, China
[4] Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, China
[5] MOE Frontiers Center for Brain Science, Fudan University, Shanghai 200433, China
[6] Zhangjiang Fudan International Innovation Center, Shanghai, China

References

[1] L. Ham, M. A. Coomer, K. Öcal, R. Grima, and M. P. H. Stumpf, "A stochastic vs deterministic perspective on the timing of cellular events," *Nat Commun*, vol. 15, no. 1, p. 5286, Jun. 2024.
[2] D. F. Anderson, "Time-dependent product-form Poisson distributions for reaction networks with higher order complexes".
[3] D. T. Gillespie, "Perspective: Stochastic algorithms for chemical kinetics," *J. Chem. Phys*, 2013.
[4] Z. Cao and R. Grima, "Linear mapping approximation of gene regulatory networks with stochastic dynamics," *Nat Commun*, vol. 9, no. 1, p. 3305, Aug. 2018.
[5] P. Smadbeck and Y. N. Kaznessis, "A closure scheme for chemical master equations," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 35, pp. 14261–14265, Aug. 2013.
[6] C. Rao, D. Waxman, W. Lin, and Z. Song, "Exact First-Passage Time Distributions from Time-Dependent Solutions of the Chemical Master Equation I: Nonlinear Networks with Bimolecular Reactions and Poisson-Product Initial Conditions," *J. Chem. Phys*, accepted for publication (2025).
[7] C. Rao, D. Waxman, W. Lin, and Z. Song, "Exact First-Passage Time Distributions from Time-Dependent Solutions of the Chemical Master Equation II: Nonlinear Networks with Bimolecular Reactions and Arbitrary Initial Conditions," *J. Chem. Phys*, accepted for publication (2025).

Microfluidic Sperm Trap Array for Single-Cell Flagellar Analysis with Unrestricted 2D Flagellar Movement

Kaiyu Wang[1], Antai Tao[2], Rongjing Zhang[3], and Junhua Yuan[4]

***Short Abstract* — Targeting challenges in quantitative sperm research, we present a microfluidic device that achieves high-throughput, non-destructive sperm capture with preservation of unrestricted 2D flagellar movement—critical for replicating free-swimming behavior and enabling whole-flagellum analysis. Utilizing this device, we investigate stimulus-induced modifications in bovine sperm flagellar dynamics at single-cell resolution. Our whole-flagellum methodology resolves discrepancies in prior partial-flagellum studies and enables precise quantification of flagellum-length integrated behavioral parameters like flagellar power dissipation. We develop an analytical model for calculating flagellar power dissipation with enhanced accuracy and comprehensiveness compared to previous methods.**

***Keywords* — microfluidic sperm capture, unrestricted flagellum, sperm flagellar dynamics, stimulus response.**

I. Background

SPERM capture techniques that immobilize sperm to halt motility are essential for analyzing flagellar motion—the primary determinant of sperm motility—and enable extended observation and individual analysis. These techniques are beneficial in assisted reproductive technologies like intracytoplasmic sperm injection (ICSI), facilitating selective sperm screening. However, current techniques either suffer from low efficiency [1], cause sperm damage [2,3], or restrict flagellar movement [4,5], highlighting a significant gap in high-throughput, non-destructive sperm capture methods that allow the flagellum to beat freely, crucial for accurately reflecting the behavior of unfettered, swimming sperm.

II. Summary of Results

We developed a microfluidic device that achieves orderly, non-destructive capture of individual sperm while maintaining unrestricted flagellar movement. By exploiting sperm's distinct boundary-following behaviors in 3D and 2D swimming modes, the device guides sperm into an array of capture structures. Geometric confinement immobilizes sperm heads without physical contact that restricts flagellar movement, enabling unrestricted beating patterns.

Using this device, we demonstrated that increasing hyperactivation agent concentration and lowering temperature have analogous effects on bovine sperm flagella—both reduce beat frequency while increasing amplitude and asymmetric harmonics, suggesting convergent modulation of flagellar dynamics. A custom laser-heating system revealed sperm's dual capacity for thermal sensitivity and regulatory robustness during stepwise temperature changes. By implementing whole-flagellum kinematic analysis, we reconciled measurement discrepancies inherent to legacy discrete-point approaches and quantified precisely length-integrated behavioral parameters, such as flagellar power dissipation. We numerically calculated flagellar power dissipation and demonstrated its positive correlation with temperature. We further established an analytical model incorporating nonuniform harmonic distributions, yielding (1) local power dissipation density profiles along the flagellum, and (2) a more accurate formula for total power dissipation than previous dimensionally based estimation methods.

III. Conclusions

We developed a non-destructive, high-throughput microfluidic device that enables unrestricted flagellar beating analysis. It uncovered convergent modulation of bovine sperm flagellar dynamics by a hyperactivation agent and temperature. Our whole-flagellum methodology addressed inconsistencies in previous partial-flagellum studies and allows precise quantification of length-integrated behavioral properties, alongside an analytical model, demonstrating its potential for fundamental quantitative sperm research and practical applications like ICSI sperm evaluation.

[1234]Hefei National Research Center for Physical Sciences at the Microscale and Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, China.
[1]Email: sa20002024@mail.ustc.edu.cn
[2]Email: taoantai@ustc.edu.cn
[3]Email: rjzhang@ustc.edu.cn
[4]Email: jhyuan@ustc.edu.cn

References

[1] Ishijima S, Oshio S, Mohri H (1986) Flagellar movement of human spermatozoa. *Gamete Res.* **13**, 185-197.
[2] Chow KW, Preece D, Berns MW (2017) Effect of red light on optically trapped spermatozoa. *Biomed. Opt. Express* **8**, 4200.
[3] Fuhr G, et al. (1998) High-frequency electric field trapping of individual human spermatozoa. *Hum. Reprod.* **13**, 136–141.
[4] de Wagenaar B, et al. (2015) Microfluidic single sperm entrapment and analysis. *Lab Chip* **15**, 1294–1301.
[5] You JB, et al. (2019) Live sperm trap microarray for high throughput imaging and analysis. *Lab Chip* **19**, 185.

Long-range genomic loci stochastically assemble into combinatorial forms of chromosome skeleton

Jingyu Zhang[1,*], Siyuan Wang[2], Simon C Watkins[3], and Jianhua Xing[1,*]

***Short Abstract* — One fundamental yet open question is how eukaryotic chromosomes fold into segregated territories, a process essential for gene transcription and cell fate. Through analyzing Hi-C and chromatin-tracing DNA-FISH data, and CRISPR-dCas9 live-cell imaging, we identify stable long-range chromosome skeleton structures formed by locus pairs with a genomic distance > 100 Mb. Biophysical model analyses reveal a multivalent binding mechanism. Our findings suggest a redundant, distributed cluster mechanism that ensures robustness across cell types and against mutations, guiding both chromosome compaction and the formation of smaller-scale chromosomal structures.**

***Keywords* — chromosome folding, multivalence binding.**

I. Background

A central problem in structural biology and biological physics is how macromolecules fold into three-dimensional structures. A human DNA polymer contains ~ 40 – 250 million base pairs (Mbs). Remarkably, eukaryotic cells of any cell type, even cancer cells, need to robustly fold multiple (i.e., 46 for normal human cells) copies of DNA molecules into segregated chromatin territories within a cell nucleus, typically only about 10 μm in diameter, in a few minutes to survive. Researchers have estimated that it would take 1000 years for a random search. To get an intuitive idea, let's scale the width of a DNA molecule to the width of a thin spaghetti of 1 *mm*, then the length of a typical human DNA molecule would be ~ 25 *km* (or 15.5 miles). That is, the task a cell faces is equivalent to folding 46 (or even more) pieces of such long and soft spaghetti spontaneously into a disk of 5 *m* in diameter and 1 *m* in height. Even more challenging, these folded polymers segregate well with minimal entanglement, not as what one would expect for a bowl of spaghetti. Such mechanism needs to work robust for different cell types (e.g., stem cells, epithelial cells, and even cancer cells), against various mutations from point mutations to loss or gain of part or whole chromosomes.

II. Results

Through analyzing existing hi-C data and high-resolution chromatin tracing structures through MERFISH (Multiplexed Error-Robust Fluorescence in situ Hybridization)[1], we identified prevalent multiple-way chromatin co-localization of segments with long genomic distance (l) in human chromosomes[2]. We identified a library of genomic locus pairs with l >100 Mbs, spatial distance d < 299 nm, and appearance frequency >8%, denoted as stable long-range colocalization (sLRC). A subset of the genomic loci act as nucleation centers (NCs) for organizing formation of the complexes.

Our hour-long live-cell imaging with CRISPR-dCas9 guided sequence-specific fluorescent labeling reveals that the structures are stable. We hypothesized that neighboring sLRCs stabilize each other via a cooperative multivalent binding mechanism. Analyses of the MERFISH data support the model, showing an effective two-state distribution.

III. Conclusion

We suggest that one of the functional roles of these structures is facilitating chromosome folding through a robust "divide-and-conquer" mechanism with redundant, distributed components. The genome harbors an extensive repertoire of locus pairs separated by long genomic distances that may potentially colocalize, from which only a subset may be available for a given cell type. At the individual chromosome level, subsets of these pairs of remote loci stochastically form segregated globally compact chromosome configurations allowing simultaneous crumpled folding (effective at < 5 Mb) at multiple regions, and enhancing contact frequencies for forming smaller (< 1 Mb) loops. The redundancy and stochasticity ensure a robust chromosome folding for different cell types against various mutations.

Acknowledgements: This work was funded by NIH grant R01GM148525, and the Charles E. Kaufman Fund of the Pittsburgh Foundation KA2018-98550.

[1]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15232, USA.

[2]Department of Cell Biology, Yale School of Medicine, New Haven, CT 06510, USA.

[3.] Department of Cell Biology, University of Pittsburgh; Pittsburgh, PA 15232, USA.

*To whom correspondence should be addressed. Email: xing1@pitt.edu, zhangjy4316@gmail.com

References

[1] J.-H. Su, P. Zheng, S. S. Kinrot, B. Bintu, X. Zhuang, Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin. *Cell* **182**, 1641-1659.e1626 (2020).

[2] J. Zhang, S. Wang, S. C. Watkins, J. Xing, Long-range genomic loci stochastically assemble into combinatorial forms of chromosome skeleton, *bioRxiv, 2025.02.10.637328* (2025).

Prolonged heat exposure decreases yeast thermal tolerance

Qian Li[1] and Yuping Chen[2]

***Short Abstract* — What defines the upper temperature limit for organisms to survive? Ken Dill and others conjectured that the denaturing temperature for essential proteins defines this limit. Previous studies showed that activation of heatshock proteins and synthesis of glycogen and trehalose protect proteins from denaturing, expanding the thermal range a cell can tolerate. Hence, a possible hypothesis is that with gradually increasing temperatures prepares cells for a harsher environment. Surprisingly, we found that prolonged culture of yeast cells at high temperature reduces their ability to survive at a previously permissive temperature. Transcriptomic analyses revealed the molecular reason behind this phenotype.**

***Keywords* — Thermal tolerance, protein denaturation, mitochondria, transcriptome.**

I. Purpose

What sets the thermal upper limit for life? Ken Dill and others have hypothesized that protein denaturation—one of the most temperature-sensitive biological processes—may define this boundary [1]. However, direct experimental evidence linking protein denaturation to thermal intolerance remains elusive, raising the question: Is protein denaturation the primary limiting factor, or do other mechanisms play a dominant role?

This question is further complicated by cells' dynamic adaptations to thermal stress. For example, cells counteract temperature shifts by producing stabilizing molecules like glycogen and trehalose and modulating cellular viscosity—a protective strategy termed viscoadaptation [2]. While such acute adaptations are well-documented, their long-term effectiveness in sustaining viability under persistent heat, as well as their interplay with protein denaturation, remains unclear.

Beyond physical adaptations, cells also rely on transcriptional reprogramming to survive thermal stress. Prior work has focused on stress-response pathways during acute heat shocks [3], but how the transcriptome enables—or fails to enable—survival under prolonged heat exposure has not been systematically investigated.

Acknowledgements: This work was funded by National Key Research and Development Program of China (2024YFA0919600).

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology,Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. E-mail: qian.li2@siat.ac.cn

[2]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology,Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. E-mail: cyp7cn@gmail.com

II. Results

A. Prolonged heat exposure reduces yeast thermal tolerance

When budding yeast cells were incubated at gradually increasing temperatures, their growth rates declined compared to cells transferred directly from lower temperatures. Notably, cultures inoculated from low-temperature-adapted cells exhibited a bimodal growth rate distribution, whereas growth rates decreased monotonically with the prior incubation temperature.

B. Impaired thermal tolerance correlates with mitochondrial dysfunction

After prolonged incubation at temperatures above 39°C, yeast cells consistently developed a petite phenotype. Further characterization revealed that these cells were [*rho*$^{-}$], indicating mitochondrial DNA loss and respiratory deficiency. While [*rho*$^{-}$] cells grew slowly at high temperatures, their growth was unaffected at lower temperatures, suggesting temperature-dependent fitness costs of mitochondrial dysfunction.

C. High temperature alters mitochondrial and vacuolar gene expression

RNA-seq analysis of log-phase cells grown at varying temperatures revealed two key trends: 1. Global RNA abundance dropped by ~50% between 30°C and 40°C, with pronounced declines in cytoplasmic translation machinery. 2. Mitochondrial genes were downregulated at high temperatures, meanwhile vacuolar genes were upregulated—a pattern resembling aging cells or stress-induced metabolic remodeling..

III. Conclusion

Cells dynamically reprogram their transcriptome in response to temperature, but this adaptation fails to sustain viability under prolonged heat stress. The observed transcriptional shifts—particularly in mitochondrial and vacuolar pathways—explain the loss of thermal tolerance after chronic exposure.

References

[1] Dill KA, Ghosh K, Schmit JD (2011) Physical limits of cells and proteomes. *PNAS* **108** (44), 17876-17882.

[2] Persson LB, Ambati VS, Brandman O (2020) Cellular Control of Viscosity Counters Changes in Temperature and Energy Availability. *Cell* **183** (6), 1572-1585.

[3] Causton HC, et al. (2001) Remodeling of Yeast Genome Expression in Response to Environmental Changes. *Mol Biol Cell* **12** (2), 323-337.

Multimolecular proofreading overcomes the activity-fidelity trade-off

Zhuo Mao[1], Yuanqi Jia[2], Yuxuan Yan[3], Bingze Wu[4] and Zibo Chen[5]

***Short Abstract* — Accurate signal processing is essential for proper cell functions, and can be achieved through kinetic proofreading, where an enzyme undergoes sequential state transitions and irreversible deactivation to enable high fidelity. Here, we designed a protein circuit that combines diffusion and endocytosis to enable kinetic proofreading at multimolecular level. Simulation revealed an experimentally crucial but overlooked trade-off between circuit activity and fidelity, and theoretical analysis confirmed its fundamentalness. By integrating self-activation and mutual inhibition mechanisms, the circuit overcomes this activity-fidelity trade-off within biologically plausible parameter regimes. Our results represent a practical multimolecular strategy for constructing high-fidelity synthetic biological circuits.**

***Keywords* — kinetic proofreading, spatial proofreading; activity-fidelity trade-off, speed-accuracy trade-off, protein circuit design, multicellular circuits**

I. Introduction

ACCURATE and precise signal processing is a hallmark of living systems. DNA polymerases replicate DNA with an error rate of 10^{-8} per nucleotide, while ribosomes translate proteins at an error rate of 10^{-4} per codon. Kinetic proofreading (KPR) explains this high fidelity [1-3], using time delay (the discrete intermediate chemical states of the proteins before they function) and irreversible deactivation (irreversible deactivation at the expense of energy) to exponentially amplify discrimination between cognate and noncognate molecules.

The high fidelity of KPR circuits is an attractive feature of biological circuit engineering, but its reliance on complex signe-protein conformational changes [1,4,5] poses design challenges. Galstyan et al. proposed an intracellular spatial proofreading scheme at the bulk level, using diffusion for time delay in space and dephosphorylation for irreversible deactivation [6]. In contrast to classic KPR, the decoupling of time delay and irreversible dissociation in spatial proofreading suggested a plausible way to build synthetic high-fidelity circuits. Here, to maximize experimental implementability and tunability, we designed an extracellular proofreading circuit.

[1]School of Life Science, Westlake University, Hangzhou, China. E-mail: maozhuo@westlake.edu.cn
[2] School of Life Science, Westlake University, Hangzhou, China. E-mail: jiayuanqi@westlake.edu.cn
[3] School of Life Science, Westlake University, Hangzhou, China. E-mail: yanyuxuan@westlake.edu.cn
[4] School of Life Science, Westlake University, Hangzhou, China. E-mail: w604631987@gmail.com
[5]School of Life Science, Westlake University, Hangzhou, China. E-mail: chenzibo@westlake.edu.cn

II. Results

A. Activity-fidelity trade-off limits the implementability of multimolecular proofreading

We calculated circuit activity (protein concentrations after diffusing across a certain distance) and fidelity of the extracellular proofreading system. Parameter sweeps revealed a Pareto-optimal front of the activity-fidelity trade-off, constraining the implementability of multimolecular proofreading.

B. Self-activation coupled with mutual inhibition (SAMI) gives rise to near-perfect proofreading

Our multimolecular design enabled flexible regulatory layers-SAMI, which eliminates the activity-fidelity trade-off.

C. Activity-fidelity trade-off is fundamental in KPR

Analysis of generalized KPR circuits involving enzymes and their cognate and noncognate substrates verified this trade-off also applies to single-molecular proofreading systems.

III. Conclusion

We designed a multimolecular proofreading circuit coupling with self-activation and mutual inhibition, to overcome the activity-fidelity trade-off. Theoretical analysis confirmed this activity-fidelity trade-off as fundamental to all KPR schemes. Our results provide a biologically plausible multi-molecular proofreading circuit design and principles for proofreading circuits to overcome the activity-fidelity trade-off.

References

[1] Boeger H (2022) Kinetic proofreading. *Annu Rev Biochem* **91**, 423–447. doi:10.1146/annurev-biochem-040320-103630
[2] Hopfield JJ (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci USA* **71**, 4135–4139. doi:10.1073/pnas.71.10.4135
[3] Ninio J (1975) Kinetic amplification of enzyme discrimination. *Biochimie* **57**, 587–595. doi:10.1016/s0300-9084(75)80139-8
[4] McKeithan TW (1995) Kinetic proofreading in T-cell receptor signal transduction. *Proc Natl Acad Sci USA* **92**, 5042–5046. doi:10.1073/pnas.92.11.5042
[5] Rodnina MV, Wintermeyer W (2001) Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem* **70**, 415–435. doi:10.1146/annurev.biochem.70.1.415
[6] Galstyan V, Amiri H, Chu D, Mugler A, Nemenman I (2020) Proofreading through spatial gradients. *eLife* **9**, e60415. doi:10.7554/eLife.60415

Genome-wide screening for mutants defective in biomass volume coordination

Qian Li [1, 2], Xiaoting Xu[1], Xiaofang Zhong[1], Yuping Chen[1, 3]

***Short Abstract* — Total macromolecular concentration is highly conserved and tightly controlled within cells. Yet, under certain physiological or pathological conditions – including cellular senescence – macromolecular concentration varies. Little is known about the controlling mechanism for macromolecular concentration and how concentration links to cellular senescence. To address these questions, we developed high-throughput methods to measure protein concentrations in a genome-wide screen. We aimed at identifying genetic regulators that affect cellular macromolecular concentration to uncover the genetic basis that connects biomass/volume growth with cellular aging.**

***Keywords* — Macromolecular concentration, homeostasis, cellular growth, biomass-volume coordination.**

I. Purpose

THE concentrations of macromolecules in cells are tightly controlled through the coordination of biomass production and cell volume[1]. Loss of macromolecular concentration control has been observed in senescent cells[2], yet the mechanisms underlying this coordination remain poorly understood.

Classical methods to alter macromolecular concentration in living cells include altering osmosis[3], applying turgor pressure[4], and mechanical compression[5]. Previously, we demonstrated that protein concentration homeostasis may arise from viscosity-dependent feedback regulation in an *in vitro Xenopus* egg extract system[6]. However, studying this process in cells has been challenging, as no regulatory factors for protein concentration homeostasis have been identified.

II. Results

Here, we utilized the Yeast Deletion Collection (YKO) library, where every non-essential gene is disrupted. Combining with liquid-handling robotics and flow cytometry, we are able to systematically measure protein concentrations in each mutant. This approach allowed us to comprehensively identify potential regulators of cytoplasmic macromolecular concentration.We developed an automated pipeline to quantify protein concentrations in single cells on a mutant-by-mutant basis.

Among the mutants analyzed, we observed a negative correlation between protein concentration and cell volume, suggesting that protein synthesis generally subscales with cell size. We found little correlation between total protein mass and cell volume, suggesting that these two physiological parameters are under independent controls. Additionally, protein concentration negatively correlated with doubling time, implying a link between growth rate and macromolecular density. Notably, many mitochondrial genes appear to influence cytoplasmic protein concentration, highlighting a potential role for metabolic regulation in biomass-volume coordination.

Consistent with previous findings, perturbations of cellular growth using small-molecules also led to a subscaling of protein relative to cell volume, further supporting the idea that growth dynamics modulate intracellular protein concentration.

III. Conclusion

We developed an automatic pipeline to measure the protein concentrations in yeast deletion mutants, enabling us to identify mutants defective in biomass-cell volume coordination. These mutants will help us understand the molecular basis of cellular senescence.

Acknowledgements: This work was funded by National Key Research and Development Program of China (2024YFA0919600) and SIAT startup funds. We thank Shenzhen Biofoundry, Dr. Tong Si, and Mr. Erpeng Guo for tremendous help and support in this project.

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology,Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

[2]E-mail: shlyecho@gmail.com

[3]E-mail: cyp7cn@gmail.com.

References

[1] OH S, LEE C, YANG W, et al. Protein and lipid mass concentration measurement in tissues by stimulated Raman scattering microscopy [J]. Proc Natl Acad Sci U S A, 2022, 119(17): e2117938119.

[2] NEUROHR G E, TERRY R L, LENGEFELD J, et al. Excessive Cell Growth Causes Cytoplasm Dilution And Contributes to Senescence [J]. Cell, 2019, 176(5): 1083-97 e18.

[3] KNAPP B D, ODERMATT P, ROJAS E R, et al. Decoupling of Rates of Protein Synthesis from Cell Expansion Leads to Supergrowth [J]. Cell Syst, 2019, 9(5): 434-45 e6.

[4] LEMIERE J, REAL-CALDERON P, HOLT L J, et al. Control of nuclear size by osmotic forces in Schizosaccharomyces pombe [J]. Elife, 2022, 11.

[5] ALRIC B, FORMOSA-DAGUE C, DAGUE E, et al. Macromolecular crowding limits growth under pressure [J]. Nat Phys, 2022, 18(4): 411-6.

[6] CHEN Y, HUANG J H, PHONG C, et al. Viscosity-dependent control of protein synthesis and degradation [J]. Nat Commun, 2024, 15(1): 2149.

Universal scaling laws linking protein folding and native dynamics revealed by AlphaFold Database

Zecheng Zhang[1], Weitong Ren[2] Jun Wang[3] and Qian-Yuan Tang[1*]

***Abstract*— The interplay between protein folding and functioning remains a central question in biophysics. Analyzing an extensive dataset of AlphaFold-predicted structures, we discover a fundamental universal scaling law linking contact order (protein folding) and fluctuation entropy (functional dynamics). This law persists across protein lengths and taxonomic boundaries, revealing a robust negative correlation: long-range contacts stabilize structures while constraining flexibility. Cross-organism comparisons show proteins from complex organisms exhibit lower contact order and higher entropy at similar chain lengths. These findings establish a quantitative framework connecting folding and functional dynamics, and reveal physical principles of protein organization implicitly captured by artificial intelligence — extending beyond static structure prediction toward insights into proteome evolution and rational protein design.**

***Keywords*— Protein structure, Protein folding, Dynamics.**

I. Purpose

Folding governs the acquisition of a polypeptide's functional three-dimensional structure (microseconds to seconds) , while native dynamics — the ensemble of postfolding conformational fluctuations (nanoseconds to microseconds) enable catalysis, binding, and allosteric regulation. Quantifying this relationship between structural organization and dynamic behavior is essential for decoding the physical principles driving protein evolution. Here, we give two key matrices, quantifies the protein folding and dynamics:

A. Folding: Contact Order (CO)

Contact Order quantifies the non-locality of residue contacts in a protein's native structure [1]. CO is defined as the normalized average sequence separation across contacts of a structure. Proteins with higher CO exhibit a greater proportion of long-range contacts, which typically results in longer folding times. Conversely, proteins with lower CO experience reduced entropic costs and shorter conformational search times, leading to faster folding, as supported by theoretical, computational, and experimental evidence [2].

This research was supported by Natural Science Foundation of China (No. 12305052), Research Grants Council of Hong Kong (No. 22302723), Hong Kong Baptist University's funding support (RC-FNRA-IG/22-23/SCI/03), and the grant of Wenzhou Institute, University of Chinese Academy of Sciences (WIUCASQD2023015).

[1]Department of Physics, Hong Kong Baptist University, Hong Kong SAR, China. Email: tangqy@hkbu.edu.hk

[2]Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, Zhejiang 325000, China.

[3]School of Physics, Nanjing University, Nanjing 210093, China.

B. Dynamics: Fluctuation Entropy (S)

Fluctuation Entropy S, quantifies the extent of dynamics around the native structure by measuring the effective volume of accessible conformational space. We focus on native-state dynamics dominated by global (slow) modes, effectively captured by Elastic Network Models (ENM). The covariance matrix C, defined as $C_{ij} = \langle \Delta\vec{r}_i \cdot \Delta\vec{r}_j \rangle$, captures the correlated displacements of residues around the native structure. It is noteworthy that this matrix is inversely related to the Hessian matrix. The fluctuation entropy directly relates to the configurational volume, which is given by determinant of C. Furthermore, this entropy can be decomposed into contributions from specific mode subsets, providing detailed insights into mode-resolved dynamics. [3]

II. Results

In summary,we have identified a robust negative correlation between contact order and fluctuation entropy, based on AlphaFold-predicted structures from 45 taxonomically diverse organisms available in AlphaFold Database (AFDB) [4]. This relationship follows a power law across varying protein lengths and taxonomic groups, revealing fundamental physical constraints that govern protein folding mechanisms and dynamic functionality. By comparing proteins of equivalent chain lengths across organisms, we observe that proteins in more complex organisms tend to fold with fewer long-range contacts and exhibit greater native-state fluctuations, thereby providing a structural signature of proteome organization that aligns with dynamic functional requirements. Scaling analysis further reveals how protein size influences this relationship and elucidates the universality of this scaling law. The observed scale invariance is characteristic of critical phenomena, supporting the view that a physical—rather than purely biological—mechanism shapes protein architecture and motions. Integrating this framework with models of sequence variation, protein-protein interactions, and functional constraints—such as allosteric transitions and ligand binding—has the potential to enable proteome-scale predictive biology and advance the field of rational protein design.

References

[1] K. W. Plaxco, K. T. Simons, and D. Baker, *J. Mol. Biol.*, vol. 277, no. 4, pp. 985–994, 1998.

[2] H. Kaya and H. S. Chan, *Proteins*, vol. 52, no. 4, pp. 524–533, 2003.

[3] A. T. Fenley, B. J. Killian, V. Hnizdo, A. Fedorowicz, D. S. Sharp, and M. K. Gilson, *J. Phys. Chem. B*, vol. 118, no. 24, pp. 6447–6455, 2014.

[4] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon *et al.*, *Nucleic Acids Res.*, vol. 50, no. D1, pp. D439–D444, 2022.

Decoding genetic and morphological regulatory architecture in development

Guoye Guan[1,†,a,b], Yixuan Chen[2,†], and Hongli Wang[1,2,3,*]

***Short Abstract* — Cell morphology is the manifestation of the physiological states, arising from a combined interplay of the intrinsic gene regulatory network and external mechanical forces. In this study, we integrate 12 shape descriptors with 412 gene expression profiles, mapping onto publicly available datasets of fluorescently labeled cell membranes from 29 independent *Caenorhabditis elegans* embryos. It's revealed that co-asymmetry of morphological and genetic features effectively distinguishes key differentiatial events, including body axis patterning, germ layer specification and cell fate decision. These insights provide a systematic approach not only for identifying cell identities but also diagnosing early disease.**

***Keywords* — 3D shape descriptor, gene expression profile, *Caenorhabditis elegans*, co-asymmetry, fate differentiation, body axis patterning, germ layer specification**

I. Introduction

CELL morphology dynamics encapsulate critical information about cellular physiological states, forming a cornerstone of developmental biology across diverse organism [1]. Understanding whether fate determination can be recognized by cellular morphological and genetic changes during early embryogenesis is not only critical for elucidating developmental mechanism, but also opens new avenues for early disease detection and therapeutic intervention [2].

II. Results

Benefited from the transparent body and reproducible cell lineage of the nematode *Caenorhabditis elegans* [3], two previously publised datasets generated through confocal microscopy imaging were used, including 8 wild-type embryos, 2 *lag-1* knockout mutants, 2 *pop-1* knockout mutants and 17 mechanically-compressed embryos, where nuclei and cell membranes were labeled with green fluorescent protein (GFP) and red fluorescent protein (mCherry), respectively [4-5].

An integrative computational pipeline is established, which combines 12 3D shape descriptors with 412 gene expression profiles to systematically identify sublineages exhibiting co-asymmetric pattern in both morphology and gene expression, enabling the prediction of their differentiation trajectories [6-7].

A. Body axis differentiation

The body axis differentiation involves coordinated changes in cell morphology and gene expression. (i) The sister sublineages of Cxp and MSx construct body wall muscles, located in the ventral and dorsal quadrant respectively. (ii) ABarpp produces left-right hypodermal seam cells with subtile chirality, while MS, C and D generate bilaterally symmetrical muscle cells.

B. Germ layer specification

A large number of genes asymmetrically express in the offspring of progenitors intended for germ layers segregation, accompanied by morphological differences. (i) MS, the anterior progeny of EMS, contributes to the mesoderm, while the posterior progeny E develops into endoderm and undergo early gastrulation invagination. (ii) Successive transdermal layer differentiation occurs in C progenies. The lineal anterior Cxa and posterior Cxp predominatly contributes to ectodermal skin and mesodermal muscle, repectively.

C. Fate decision

Each fate can be labled by multidimensional phenotype markers, aiding in early cell identity distinction.

Acknowledgements: This work was financially supported by the Major Program of the National Natural Science Foundation of China (No. 12090051) and the National Key Research and Development Program of China (No. 2018YFA0900200).

[1]Center for Quantitative Biology, Peking University, Beijing, China; E-mail: guanguoye@gmail.com (G.G.)

[2] The State Key Laboratory for Artificial Microstructures and Mesoscopic Physics, School of Physics, Peking University, Beijing, China; E-mail: yixuanchen@stu.pku.edu.cn (Y.C.)

[§]Present address: Department of Systems Biology, Harvard Medical School, Boston USA

[#]Present address: Department of Data Science, Dana-Farber Cancer Institute, Boston USA

[†] These authors contributed equally to this work.

[*]Correspondence: hlwang@pku.edu.cn (H.W.)

References

[1] Uriu K, Morelli LG. (2023) Orchestration of tissue shape changes and gene expression patterns in development. *Semin Cell Dev Biol* **147**, 24–33.

[2] Barbier de Reuille P, et al. (2015) MorphoGraphX: A platform for quantifying morphogenesis in 4D. *eLife* **4**, 05864.

[3] Sulston JE, et al. (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* **100**, 64-119.

[4] Cao J, et al. (2020) Establishment of a morphological atlas of the *Caenorhabditis elegans* embryo using deep-learning-based 4D segmentation. *Nat Commun* **11**, 6254.

[5] Guan G, et al. (2025) Cell lineage-resolved embryonic morphological map reveals signaling associated with cell fate and size asymmetry. *Nat Commun* **16**, 3700.

[6] Guan G, et al. (2024) Characterizing cellular physiological states with three-dimensional shape descriptors for cell membranes. *Membranes (Basel)* **14**, 137.

[7] Packer JS, et al. (2019) A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **365**: eaax1971.

Constructing a holistic map of cell fate decision by hyper solution landscape

Xiaoyi Zhang[1], Zhiyuan Li[1,2], and Lei Zhang[1,3,4,]

***Short Abstract* — The Waddington landscape metaphor has inspired extensive quantitative studies of cell fate decisions using dynamical systems. While these approaches provide valuable insights, the intrinsic nonlinear complexity and the parameter dependence limits systematic analysis of fate transitions. Here, we introduce the Hyper Solution Landscape (HSL), a minimally parameter-dependent methodology showing a comprehensive structure of all possible landscape configurations for gene regulatory networks. Building on the concept of solution landscape that primarily captures the complete stationary points in static landscape, HSL connects different solution landscapes to reflect dynamic changes in these landscapes associated with bifurcations. Applied to the Cross-Inhibition with Self-activation motif, HSL analysis identifies key hyperparameters driving distinct directional changes in cell fate propensity. Importantly, different routes through the HSL between the same initial and final states can produce markedly different fate distributions. This enables rational design of transition strategies. We validate HSL's utility in the seesaw model of cellular reprogramming, establishing a powerful framework for understanding and engineering cell fate decisions.**

***Keywords* — cell fate decision, Waddington landscape, gene regulatory network, solution landscape, hyper solution landscape**

I. Purpose

The Waddington landscape has served as a powerful metaphor for understanding cell fate decisions since its inception. Several attempts have been made to quantify cell fate decisions through dynamical systems. While nonlinear dynamics models have proven valuable in characterizing cell fate decisions, they face a fundamental limitation: parameter dependency. Different parameter sets can lead to distinct landscapes with varying predictions on cell fate behaviors.

The solution landscape concept offers a promising alternative approach. Building on this foundation, we propose a novel framework: Hyper Solution Landscape (HSL), which is the ensemble of all solution landscapes and their bifurcation relationships for a dynamic system. HSL is a minimally parameter-dependent methodology that shows a comprehensive structure of all possible landscape configurations for a given GRN topology.

[1] Center for Quantitative Biology, Peking University, Beijing, 100871, China
[2] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China
[3] Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China
[4] Center for Machine Learning Research, Peking University, Beijing 100871, China
Emails: xyzhang@stu.pku.edu.cn (X.Z.) zhiyuanli@pku.edu.cn (Z.L.) zhangl@math.pku.edu.cn (L.Z.)

II. Results

We apply this approach to both the CIS motif of differentiation and the seesaw model of reprogramming, establishing clear connections between HSL and cell fate decisions:

1. The alteration of solution landscapes corresponds to distinct geometry of cell fate decision.
2. HSL reveals intrinsic nonlinear complexity of cell fate decisions, including the most probable fate distributions and their responses to varying hyperparameters.
3. Different bifurcation routes in HSL induce distinct cell fate decision behaviors.
4. HSL enables the rational design of cell fate decision routes.

III. Conclusion

By constructing a holistic map, HSL provides a useful tool for both noise-driven and signal-driven cell fate decisions. This approach systematically identifies the ensemble of solution landscapes and their bifurcations, providing a novel theoretical framework and computational tool for understanding complex cell differentiation and reprogramming.

References

[1] Xiaoyi Zhang, Zhiyuan Li*, Lei Zhang*. Constructing a holistic map of cell fate decision by hyper solution landscape, *bioRxiv 2024.11.28.625944*.

Studying ecosystem of tumor microenvironment in colorectal cancer using spatial transcriptomics

Zhiyu Lv[1], Kaiwen Luo[1], Wenfei Jin[2#], Ziwei Dai[1#]

***Short Abstract* — The tumor microenvironment is similar to an ecosystem, a mixture composed of different types and quantities of cells. The behaviors of populations and the structure of the tumor microenvironment can both be explained by ecological theories, but more high-dimensional data are needed. Here, we analyzed 48 spatial transcriptomic datasets of colorectal cancer, defined 16 tumor spatial microniches, and quantified their neighborhood relationships. At the same time, we discovered unique spatial periphery distribution of two malignant cell subtypes, i2 and i3, and explained this phenomenon using the reaction-diffusion model. Finally, by analyzing the metabolic interaction between i2 and i3, the conclusion of the model was further verified.**

***Keywords*—Spatial transcriptomic,Tumor microenvironment, Reaction-diffusion model**

I. Background

The tumor microenvironment is similar to an ecosystem[1], a mixture composed of different cell types and quantities. The occurrence and growth of tumors are strikingly similar to an evolving ecosystem. For instance, species richness (also known as intratumoral heterogeneity) may be related to the robustness of immunotherapy and the long-term prognosis of patients[2][3]. The metabolic competition between immune cells and cancer cells, also known as interspecies competition[4], is also a key determinant of cancer progression.

The behavior of microenvironmental populations and the structure of the tumor microenvironment can be explained by ecological theories, but more high-dimensional data are needed. Therefore, one challenge in decoding the spatial structure of the tumor microenvironment is how to capture the high-throughput spatial profiles of the tumor microenvironment at the whole-genome level. Solving this problem requires the ability to simultaneously record transcriptional information and spatial coordinates.

II. Summary of the Results

By analyzing 48 spatial transcriptomic slices of colorectal cancer, we defined 16 tumor spatial microniches based on the proportion of cell types and their spatial positions, and quantified their neighbor relationships. At the same time, we discovered 3 Tumor Radiating Eco-Axes (TREA), which reflects the ordered spatial arrangement of cell types and tissue layers within the tumor.

In the first TREA, we found a special periphery distribution relationship between two malignant cell subtypes, i2 and i3. By simulating their different interaction relationships during proliferation using a reaction-diffusion model, we ultimately discovered that when they mutually promoted and the diffusion rate of i3 was faster, the most likely periphery distribution would form. We found that the expression of cell division-related genes was significantly higher after the coexistence of the two subtypes, and the expression of genes related to the mobility of i3 was higher than that of i2, verifying the conclusion of the model.

Finally, through cell communication analysis, we discovered that i2 and i3 may exchange metabolites related to nucleic acid metabolism. Specifically, i3 supplies various raw materials for nucleic acid de novo synthesis to i2, and both sides also exchange many raw materials for nucleic acid salvage pathways. This further explains the mechanism by which the two promote each other.

III. Conclusion

In our research, we conducted an in-depth analysis of the spatial ecological landscape of the tumor microenvironment in colorectal cancer using spatial transcriptomics data. We quantified the neighbor relationships in different microniches and defined three Tumor Radiating Eco-Axes. Additionally, we discovered the unique periphery distribution of two tumor subtypes. Through mathematical models, we inferred their interaction relationships during evolution and verified them using single-cell transcriptomics data. Furthermore, we found that they might promote each other's growth by exchanging metabolites related to nucleic acid metabolism. This provides important evidence for understanding the cooperative behaviors and mechanisms in the tumor microenvironment.

[1]College of Life Sciences, Southern University of Science and Technology. E-mail: 12331384@mail.sustech.edu.cn

[2]Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences.

References

[1] Reynolds, B. A., Oli, M. W. & Oli, M. K. Eco-oncology: Applying ecological principles to understand and manage cancer. Ecol Evol 10, 8538-8553, doi:10.1002/ece3.6590 (2020).

[2] Vitale, I., Shema, E., Loi, S. & Galluzzi, L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. Nat Med 27, 212-224, doi:10.1038/s41591-021-01233-9 (2021).

[3] Ma, L. et al. Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer. Cancer Cell 36, 418-430 e416, doi:10.1016/j.ccell.2019.08.007 (2019).

[4] Chang, C. H. et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. Cell 162, 1229-1241, doi:10.1016/j.cell.2015.08.016 (2015).

Mapping Stem Cell Dynamics in the Intestinal Crypt with Spatiotemporal Modeling

Xingrui Wang[1], Hao Dai[1], and Bin Zhou[1,2,3]

***Short Abstract* — The intestinal epithelium exhibits rapid turnover, fueled by stem cell proliferation.[1] While crypt-base Lgr5^{hi} cells classically follow a neutral drift model,[2,3] recent studies propose upper crypt Fgfbp1+/Lgr4+ cells as stem cells expanding bidirectionally.[4,5] To decipher stem cell positioning and dynamics, we employed novel *in vivo* fluorescence reporters for exclusive labeling at distinct locations.[6] Using a suite of genetically engineered mouse models, we are assessing clonogenic potential and developing enhanced models with expanded experimental data. Ultimately, we aim to establish a spatiotemporal atlas and mathematical framework of homeostatic stem cell proliferation within the intestinal crypt.**

***Keywords* — Intestinal stem cells, Crypt dynamics, Lineage tracing, Computational modeling**

REFERENCES

[1] Gehart, H., & Clevers, H. (2019). Tales from the crypt: new insights into intestinal stem cells. *Nature reviews. Gastroenterology & hepatology*, *16*(1), 19–34.

[2] Barker, N., van Es, J. H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P. J., & Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, *449*(7165), 1003–1007.

[3] Lopez-Garcia, C., Klein, A. M., Simons, B. D., & Winton, D. J. (2010). Intestinal stem cell replacement follows a pattern of neutral drift. *Science (New York, N.Y.)*, *330*(6005), 822–825.

[4] Capdevila, C., Miller, J., Cheng, L., Kornberg, A., George, J. J., Lee, H., Botella, T., Moon, C. S., Murray, J. W., Lam, S., Calderon, R. I., Malagola, E., Whelan, G., Lin, C. S., Han, A., Wang, T. C., Sims, P. A., & Yan, K. S. (2024). Time-resolved fate mapping identifies the intestinal upper crypt zone as an origin of Lgr5+ crypt base columnar cells. *Cell*, *187*(12), 3039–3055.e14.

[5] Malagola, E., Vasciaveo, A., Ochiai, Y., Kim, W., Zheng, B., Zanella, L., Wang, A. L. E., Middelhoff, M., Nienhüser, H., Deng, L., Wu, F., Waterbury, Q. T., Belin, B., LaBella, J., Zamechek, L. B., Wong, M. H., Li, L., Guha, C., Cheng, C. W., Yan, K. S., … Wang, T. C. (2024). Isthmus progenitor cells contribute to homeostatic cellular turnover and support regeneration following intestinal injury. *Cell*, *187*(12), 3056–3071.e17.

[6] He, L., Li, Y., Li, Y., Pu, W., Huang, X., Tian, X., Wang, Y., Zhang, H., Liu, Q., Zhang, L., Zhao, H., Tang, J., Ji, H., Cai, D., Han, Z., Han, Z., Nie, Y., Hu, S., Wang, Q. D., Sun, R., … Zhou, B. (2017). Enhancing the precision of genetic lineage tracing using dual recombinases. *Nature medicine*, *23*(12), 1488–1498.

[1]New Cornerstone Science Laboratory, State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China

[2]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China.

[3]School of Life Science and Technology, ShanghaiTech University, Shanghai, China.

The name of the author presenting is underlined. E-mail: wangxingrui2023@sibcb.ac.cn

Integrated microbial approaches for enhancing maize protein accumulation: from germplasm isolation to synthetic community engineering

Amina Hadayat[1], Cong Lan[1], Xin-Yu Wang[1], Rui Chen[1], Zhuo Wang[1], Wei-Hang Wang[1], Xuan Ming[1], Jun-Hao Huang[1], Jia-Yi Li[1], Zhao-Shuo Zhang[1], Rui Zhang[1], Xin-Hao Liu[1], Chenliao Wu[1], Peng Cai[1], Chun-Hui Gao[1]*

***Short Abstract* — We implemented two large-scale microbiota-based interventions in maize to explore their effects on protein accumulation and rhizosphere microbiome assembly. Using whole seed microbiota transplantation (n = 179) and variety-specific synthetic microbial communities (n = 186), we studied B73 maize under controlled conditions. Plant phenotyping and 16S rRNA V4 sequencing are complete; in silico analysis for phylogenetic diversity and microbial clustering is underway. Ongoing work evaluates nitrogen-use efficiency and root-microbe interactions using microfluidics and metabolic modeling.**

***Keywords* — maize microbiota, protein accumulation, synthetic microbial communities, seed microbiota transplantation, nitrogen-use efficiency**

I. PURPOSE

Modern maize hybrids typically contain only 5–10% seed protein due to past selection for yield and starch content [1, 2], while ancestral forms like teosinte contained over 20% [1]. Overuse of nitrogen fertilizer has improved yield but also caused environmental damage [3]. As global demand grows for plant-based protein and sustainable agriculture, improving nitrogen-use efficiency (NUE) and protein traits becomes a priority [4]. Microbiota, especially seed-derived communities, offer promising avenues to influence plant nutrient acquisition, stress resilience, and productivity [5]. This project investigates whether native seed microbiota and designed SynComs can modulate maize protein accumulation and rhizosphere microbiome assembly, with the ultimate goal of informing microbiome-assisted breeding and predictive plant-microbe models.

II. METHOD

We conducted two parallel microbiome intervention trials on the maize inbred line B73:

1. Whole Seed Microbiota Transplantation (n = 179): Seeds from 179 donor varieties were soaked in sterile water for 48 h to extract native microbiota. Surface-sterilized B73 seeds were then soaked in the donor microbiota solution for 2–3 h and grown in a potting soil mixture.

2. Synthetic Microbial Community (SynCom) Application (n = 186): From over 2,000 isolated epiphytic and endophytic strains, variety-specific SynComs were assembled and applied to B73 seeds via microbial priming.

Plants were grown in a randomized greenhouse design for 30 days. We measured root/shoot length, fresh biomass, and total protein content (Bradford assay). Rhizosphere samples were subjected to 16S rRNA V4 amplicon sequencing. Ongoing in silico processing includes ASV generation, phylogenetic tree construction, and distance-based community clustering to link microbial structure with host traits.

III. CONCLUSION

This study provides a scalable platform to dissect functional plant-microbe interactions in maize. By comparing whole microbiota versus strain-informed SynCom interventions, we aim to identify microbial features that enhance protein content and influence rhizosphere composition. Future phases will assess NUE, root exudate-driven recruitment, and genome-scale metabolic modeling via CarveMe and microfluidics. These findings contribute to microbiota-guided crop improvement and predictive models for host-microbe-environment interactions.

REFERNCES

1. Flint-Garcia SA, Bodnar AL, Scott MP. Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte. Theor Appl Genet. 2009;119:1129–42.
2. Duvick DN. Advances in Agronomy, Volume 86 - 1st Edition. 2005. https://shop.elsevier.com/books/advances-in-agronomy/sparks/978-0-12-000784-4. Accessed 16 Jan 2024.
3. Wani SH, Vijayan R, Choudhary M, Kumar A, Zaid A, Singh V, et al. Nitrogen use efficiency (NUE): elucidated mechanisms, mapped genes and gene networks in maize (Zea mays L.). Physiol Mol Biol Plants. 2021;27:2875–91.
4. Ciampitti I, Lemaire G. From use efficiency to effective use of nitrogen: A dilemma for maize breeding improvement. Science of The Total Environment. 2022;826:154125.
5. Ganugi P, Fiorini A, Ardenti F, Caffi T, Bonini P, Taskin E, et al. Nitrogen use efficiency, rhizosphere bacterial community, and root metabolome reprogramming due to maize seed treatment with microbial biostimulants. Physiol Plant. 2022;174:e13679.

Acknowledgements: This work is supported by the National Natural Science Foundation of China (32370127), and Fundamental Research Funds for the Central Universities (2662025ZHPY001).

[1]National Key Laboratory of Agricultural Microbiology, College of Resources and Environment, Huazhong Agricultural University, Wuhan, China. E-mail: gaoch@mail.hzau.edu.cn

Hierarchical Superposition Pharmacodynamic Model for Natural Medicine Formulas

Jiajun Huang[1,2],Weifeng Liang[2], Dingsi Xiang[2], Shengli Mi[1,2]

***Short Abstract* — This work introduces the Hierarchical Superposition Pharmacodynamic Model (HSPM), a novel method to analyze the complex effects of natural medicine formulas, particularly in traditional Chinese medicine (TCM). The HSPM leverages human protein-protein interaction networks to quantify interactions between drug targets and disease proteins, evaluating the impact of drug combinations on diseases. The model calculates pharmacodynamic action scores to account for drug effect transmission across hierarchical levels, considering decay rates and node degree values. Validated against extensive datasets, HSPM differentiates between known effective and unknown or ineffective drug-disease pairs, highlighting its predictive modeling potential. This research offering a promising framework for comprehending drug combination effects and guiding new therapeutic strategies for complex diseases.**

***Keywords* — Hierarchical Superposition Pharmacodynamic Model, Drug Combinations, Complex Disease Treatment.**

Complexity accumulation over time is crucial. Any object, physical or not, that exists over time and is part of causal chains can generate new, more complex objects. [1-3] Interactions between matter are often structured hierarchically. [1] Each level has unique interaction rules that aren't just a simple summation of lower - level rules. Biological systems, as typical complex systems, carry historical evolutionary information in their components at various levels. This information records past states and is a key part of the current system's material properties.

The development of drug combinations and pharmacological research mostly uses a strong reductionist approach. It focuses on the direct effects of interventions on single or multiple targets but ignores the cumulative impacts of secondary reactions and hierarchical effects. [4] This method fitting linear causality is good for understanding drug-target binding and direct biological effects, and suits developing drugs for single - disease mechanisms. But for complex - disease combination therapy, it can't fully grasp the multi-level biological responses of drugs in the body. Biological molecule networks at different levels interact complexly to create therapeutic effects. Intervening multiple targets at once may affect many regulatory networks, leading to new targets being intervened. Also, even the known pharmacological actions of a single drug may not be the end of efficacy due to secondary effects. Traditional chinese medicine, especially multi-component herbal combinations, shows this complexity in natural drug application. These combinations work via multiple targets and pathways, and their multi-point interventions and synergistic effects create complex biological features in the body.

The core concept of the Hierarchical Superposition Pharmacodynamic Model (HSPM) is to ascertain the hierarchical relationship between disease targets and drug targets, and evaluate drug efficacy by calculating the hierarchical transmission of drugs along chosen pathways. This method leverages Protein - Protein Interaction Networks (PPI), quantifies network interactions between drug and disease targets, and assesses the disease - fighting effects of single or combined drugs. The first step in calculating efficacy is to define the hierarchical relationship (distance in the network) between disease and drug targets. As efficacy can be transmitted via multiple pathways from drug to disease targets, the research team aims to find the most representative shortest path with the highest weight for hierarchical transmission. Specifically, they seek the path with the shortest length (hierarchical distance) and the smallest product of node degrees along the path. When calculating the pharmacodynamic effect score, we assumes that drug efficacy decays as it is transmitted across different hierarchies. Each transmission step leads to a reduction in drug effectiveness, with the extent of decay determined by the inter - hierarchy distance (the hierarchical distance between two nodes). Also, within the same hierarchy, efficacy transmission is influenced by protein node degree. Proteins with low node degree tend to pass effects to other hierarchies, while those with high node degree retain more efficacy, causing effects to mainly spread within the same hierarchy. This study pioneers the application of hierarchical structure principles from physics to pharmacology, introducing a novel framework to decipher the complex effects of natural medicine formulas.

Acknowledgements: This work is funded by the National Natural Science Foundation of China under grant no. 82304451.

[1] Bionic Design and Intelligent Manufacturing R&D Center, Research Institute of Tsinghua University in Shenzhen, Shenzhen

[2] Bio-manufacturing Engineering Laboratory, Institute of Data and Information, Tsinghua Shenzhen International Graduate School, Tsinghua University

E-mail: huangjiajun.2019@tsinghua.org.cn

REFERENCES

[1] Anderson PW. More Is Different: Broken symmetry and the nature of the hierarchical structure of science. Science 1972;177:393–6.

[2] Hébert-Dufresne L, Allard A, Garland J, Hobson EA, Zaman L. The path of complexity. Npj Complexity 2024;1:4.

[3] Sharma A, Czégel D, etc. Assembly theory explains and quantifies selection and evolution. Nature 2023;622:321–8.

[4] Sutherland JJ, etc. A preclinical secondary pharmacology resource illuminates target-adverse drug reaction associations of marketed drugs. Nature Communications 2023;14:4323.

Hollow condensates emerge from gelation-induced spinodal decomposition

Cheng Li[1, †], Lingyu Meng[2, †], Yongxin Tong[1, 3, †], Jie Lin[1, 2, *], and Zhi Qi[1, 4, *]

***Short Abstract* — Recent studies have identified diverse hollow biomolecular condensates, characterized by biomolecule-depleted interiors surrounded by biomolecule-rich shells. Although several formation mechanisms have been proposed, a general thermodynamic driving force remains elusive. Here, we investigate a well-defined system in which the human transcription factor p53 and non-specific double-stranded DNA (dsDNA) form biomolecule-rich condensates. Introduction of dsDNA containing p53-binding motifs induces a morphological transition to hollow structures, accompanied by a material state transition from liquid-like to gel-like. *In vitro* assays indicate that the formation of hollow condensates is driven by p21 DNA-induced localized gelation at the condensate periphery. Guided by these findings, we developed a three-component phase-field model that quantitatively recapitulates the formation of hollow condensates. Simulations show that peripheral gelation leads to gradual depletion of protein and Random DNA from the condensate core, triggering spinodal decomposition and lumen formation inside condensates. Together, these results offer mechanistic insights into multi-component hollow condensates.**

***Keywords* — Biomolecular condensate, hollow condensate, dsDNA-protein interactive co-condensate (DPIC), gelation, spinodal decomposition**

[†]Equal contribution

[1]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[3]The Integrated Science Program, Yuanpei College, Peking University, Beijing 100871, China

[4]School of Physics, Peking University, Beijing 100871, China

**To whom correspondence should be addressed. Email: zhiqi7@pku.edu.cn; linjie@pku.edu.cn*

Contact to Cheng Li:
Tel: 15650704509, Email: lichengowo@pku.edu.cn

References

[1] Alberti, S., Gladfelter, A., andMittag, T. (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates Cell **176**, 419-434 10.1016/j.cell.2018.12.035

[2] Alshareedah, I., Moosa, M. M., Raju, M., Potoyan, D. A., andBanerjee, P. R. (2020) Phase transition of RNA-protein complexes into ordered hollow condensates Proc Natl Acad Sci U S A **117**, 15650-15658 10.1073/pnas.1922365117

[3] Erkamp, N. A., Sneideris, T., Ausserwoger, H., Qian, D., Qamar, S., Nixon-Abell, J. *et al.* (2023) Spatially non-uniform condensates emerge from dynamically arrested phase separation Nat Commun **14**, 684 10.1038/s41467-023-36059-1

[4] Li, C., Bian, Y., Tang, Y., Meng, L., Yin, P., Hong, Y. *et al.* (2025) Deciphering the molecular mechanism underlying morphology transition in two-component DNA-protein cophase separation Structure **33**, 62-77 e68 10.1016/j.str.2024.10.02

Mitochondrial protein heterogeneity stems from the stochastic nature of co-translational protein targeting in cell senescence

Abdul Haseeb Khan[1], Xuefang Gu[1], Rutvik J. Patel[2], Prabha Chuphal[2], Matheus P. Viana[3], Aidan I. Brown[2], Brian M. Zid[4], Tatsuhisa Tsuboi[1, 4]

***Short Abstract* — Mitochondrial dysfunction is a hallmark of aging, yet the link between morphological changes and loss of function remains unclear. We demonstrate that sustained mRNA localization and co-translational protein delivery result in heterogeneous protein distribution across fragmented mitochondria. We find that age-induced mitochondrial fragmentation substantially increases protein expression noise. Translational kinetic–diffusion modeling revealed that stochastic compartmentalization explains protein expression noise, with co-translational delivery as the primary contributor to increased heterogeneity. Notably, fission-fusion dynamics suppress protein heterogeneity, while mitophagy does not. Increasing mRNA abundance and inhibiting co-translational targeting reduces heterogeneity, laying a framework for understanding and mitigating age-related mitochondrial dysfunction.**

***Keywords* — mitochondrial fragmentation, protein heterogeneity, co-translational targeting, mitochondrial dynamics, cell senescence**

I. Purpose

Mitochondria are crucial hubs for metabolites and energy generation, essential for diverse cellular functions, yet encode only a small subset of their required proteins themselves. The majority are nuclear-encoded and require intricate post-synthesis delivery mechanisms, including co-translational import, a process crucial for maintaining mitochondrial health and cellular homeostasis. Although the molecular pathways guiding mRNA localization and protein import have been characterized, it remains unclear how these processes shape protein composition within fragmented mitochondrial networks.

II. Preparation of Abstracts

To investigate this, we quantitatively analyzed the heterogeneity of nuclear-encoded protein distribution across individual mitochondrial fragments in yeast using MitoGraph-based 3D segmentation. We found that age-associated mitochondrial fragmentation notably increased the heterogeneity of nuclear-encoded Tim50 and Tim23. Using mother-enrichment-program yeast strains to enrich aged cells, we observed an elevated number of smaller mitochondrial fragments that correlated with increased protein expression noise. Live-cell imaging revealed TIM50 mRNAs remained localized to the mitochondrial surface, restricting their movement for extended periods. Stochastic simulations of dynamic mitochondrial networks supported that decreased fusion rates lead to increased fragmentation, elevating expression noise, particularly in smaller fragments. Experimentally, fusion-deficient (*fzo1Δ*) cells exhibited fragmented mitochondria and increased protein heterogeneity that inversely correlated with fragment size. In contrast, mitophagy-deficient (*atg32Δ*) cells showed no such changes, indicating primary reliance on fission-fusion dynamics over mitophagy to limit heterogeneity. Consistently, co-translationally imported proteins (Tim50, Mia40, Dld1, and Cox15) exhibited higher heterogeneity compared to post-translationally targeted proteins. We successfully reduced protein heterogeneity either by increasing Tim50 mRNA copy number to reduce stochasticity in protein levels or by tethering transcripts to the plasma membrane via MCP-CAAX, thereby hindering co-translational import.

III. Conclusion

Our study shows that sustained mitochondrial fragmentation significantly increases heterogeneity in the distribution of nuclear-encoded proteins, mainly due to stochastic co-translational targeting. Mitochondrial fusion dynamics, rather than mitophagy, are critical for maintaining protein homeostasis. We propose that this fragmentation-driven heterogeneity disrupts proteostasis, thereby accelerating cell senescence and contributing to age-related diseases. This research establishes a fundamental framework for understanding how morphological changes impact protein organization and function in the context of aging and disease.

REFERENCES

[1] Khan, AH, et al. (2024) Mitochondrial protein heterogeneity stems from the stochastic nature of co-translational protein targeting in cell senescence. *Nat Commun* **15**, 8274.

[1]Shenzhen International Graduate School, Tsinghua University, Shenzhen, China. E-mail: ttsuboi@sz.tsinghua.edu.cn
[2]Toronto Metropolitan University, Toronto, ON, Canada
[3]Allen Institute for Cell Science, Seattle, WA, USA
[4]University of California San Diego, La Jolla, CA, USA

Rational design of highly efficient nitrogen-fixing synthetic microbial communities

Cong Lan[1], Amina Hadayat[1], Rui Zhang[1], Jun-Hao Huang[1], Peng Cai[1], Chun-Hui Gao[1]*

***Short Abstract* — To address the need for efficient nitrogen utilization in maize, this study aims to construct high-performance biological nitrogen-fixing synthetic consortia. We're isolating both diazotrophs and engineered "helper strains" directly from maize seeds. Our goal is to enhance crop biological nitrogen fixation and reduce chemical fertilizer application. Leveraging single-strain to community-level metabolic network models, this research accelerates the Design-Build-Test-Learn (DBTL) pipeline, integrating strain isolation, genome analysis, model reconstruction, and intelligent design of community metabolic flux. This approach endeavors to overcome bottlenecks in traditional single-strain screening.**

***Keywords* — synthetic microbial consortia, nitrogen-fixing bacteria, maize microbiota, genome-scale metabolic model (GEM)**

I. Purpose

Developing high-protein maize cultivars is crucial, requiring enhanced nitrogen uptake and assimilation efficiency to support protein biosynthesis [1]. Recent studies demonstrate that maize nitrogen fixation systems, leveraging plant symbiotic microbiomes and designed synthetic microbial communities, can significantly improve nitrogen availability and crop yield while reducing disease incidence. Furthermore, consortia comprising diazotrophs and functional helper strains exhibit substantial potential in boosting biological nitrogen fixation efficiency [2]. Research also indicates that computational methods, such as genome-scale metabolic model (GEM) reconstruction and analysis, can aid in microbial community design, potentially optimizing inter-microbial ecological functions and interactive performance [3–5]. Through this approach, we aim to overcome the limitations of traditional single-nitrogen-fixing strain screening and performance bottlenecks, enhance the stability of synthetic communities, improve the rational design capabilities for nitrogen-fixing consortia, accelerate biofertilizer development, and support the future of sustainable agriculture.

II. METHOD

Design steps for nitrogen-fixing synthetic consortia. Developing effective nitrogen-fixing synthetic consortia requires a systematic and rational approach to ensure optimal performance and stability. Our design strategy involves the following key steps:

1. Maize seed microbiota acquisition

Maize seed microbiota isolation: We germinated B73 seeds from 179 donor varieties in sterile water. We then extracted both endophytic and epiphytic strains, which we subsequently cultured on plates. Individual colonies were picked from these plates to obtain the desired seed-associated strains.

2. Nitrogen-fixing strain screening

We screened for nitrogen-fixing strains by performing sequence alignment of the *nifH* nitrogenase gene against strains in the RefSeq database.

3. Nitrogen-fixing synthetic consortia design

We propose to construct community-level genome-scale metabolic models (GEMs) for consortia comprising nitrogen-fixing strains and helper strains, and then simulate their nitrogen fixation capacity.

II. Conclusion

We successfully identified dominant maize seed-associated bacterial genera, including *Pantoea* and *Burkholderia*. Simultaneously, we screened and confirmed nitrogen-fixing bacteria by matching them against the RefSeq database. Our next step involves utilizing genome-scale metabolic models (GEMs) to simulate their nitrogen fixation efficiency.

REFERNCES

[1]. Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota | PLOS Biology.

[2]. Zhang, L. *et al.* A highly conserved core bacterial microbiota with nitrogen-fixation capacity inhabits the xylem sap in maize plants. *Nat Commun* **13**, 3361 (2022).

[3]. Quinn-Bohmann, N., Carr, A. V., Diener, C. & Gibbons, S. M. Moving from genome-scale to community-scale metabolic models for the human gut microbiome. *Nat Microbiol* **10**, 1055–1066 (2025).

[4]. Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res* **46**, 7542–7553 (2018).

[5]. Xu, X. & Jiang, J. Engineering microbiomes for enhanced bioremediation. *PLOS Biology* **22**, e3002951 (2024).

Acknowledgements: This work is supported by the National Natural Science Foundation of China (32370127), and Fundamental Research Funds for the Central Universities (2662025ZHPY001).

[1]National Key Laboratory of Agricultural Microbiology, College of Resources and Environment, Huazhong Agricultural University, Wuhan, China. E-mail: GAOCH@MAIL.HZAU.EDU.CN

Experimental evolution of cytoplasmic density mutants

Xiaoting Xu[1,2], Qian Li[1], Xiaofang Zhong, and Yuping Chen[1,3]

***Short Abstract* — Cellular mass density is tightly regulated within a specific cell-type, variation of which affects many aspects of cellular behavior. Yet, the specific mechanism from which such tight regulation arises is largely unknown. To address this question, we performed experimental evolution to generate high-density and low-density yeast mutants. Using whole genome sequencing, we identified mutations that causes cellular mass to vary, which were then used to study the causes and effects of cellular mass density change.**

***Keywords* — Cytoplasmic concentration, cellular growth, experimental evolution, cell volume, cell biomass.**

I. Purpose

CELL mass density is tightly regulated through the coordination of biomass production and cell volume expansion [1]. However, this coordination is not always maintained. Studies have shown that osmotic stress[2], cell differentiation[3], and cellular aging can alter cell mass density[1,4]. Yet, no genetic element regulating this process has been identified so far[5]. Our previous research using the Xenopus egg extract system revealed that elevated cytoplasmic density has biphasic effects on protein translation[6]: it boosts translation when cytoplasmic density is low but slows it down when density is high, suggesting a negative feedback homeostatic mechanism for cell mass density.

Several methods can determine cell density, including suspended microchannel resonators [7], optical methods using quantitative phase microscopy [8] and stimulated Raman spectroscopy [1], and particle tracking [3]. While these methods allow measurements in single cells, it is challenging to apply them to select cells with varied densities due to their limitations in coupling with cell sorting methods or throughput.

II. Result

A. Experimental Evolution for Density Mutants

To investigate the cell density maintenance system, we designed an experimental evolution study in yeast cells to obtain mutants with varying cytoplasmic densities. Using density gradient centrifugation, we separated and selected cells with the highest and lowest densities, applying pulsatile selection pressure to drive divergence in cytoplasmic density. After over a hundred generations, we generated two distinct populations: a high-density population ("dense") and a low-density population ("light").

B. Quantification of Protein Concentration

We aimed to measure biomass and cell volume correlations in these mutants. Using the Bradford assay coupled with Coulter Counter measurements of cell volume, we quantified the protein concentration in these cells Surprisingly, we found that protein concentration negatively correlated with cellular buoyant density.

C. Mutation Identification Using Whole Genome Sequencing

We performed next-generation sequencing on the genomes of these density mutants to identify genes that play key roles in density regulation. This analysis will help us investigate which genes participate in maintaining cytoplasmic density and its regulation. Intriguingly, dense cells frequently carried mutations in sugar transporters and genes related to metabolism, suggesting that metabolism, particularly, carbohydrate metabolism, may influence cellular density. Additionally, several density mutants exhibited extra copies of chromosomes, implying that ploidy may also play a role in regulating cell volume and biomass.

III. Conclusion

We developed a method for screening and isolating yeast mutants with altered cytoplasmic density. This approach enables us to study the extensive functional consequences of changes in cytoplasmic density and to understand how cells maintain their cytoplasmic densities.

References

[1] Liu X, Oh S, and Kirschner MW (2022) The uniformity and stability of cellular mass density in mammalian cell culture. *Front. Cell Dev. Biol*. 10, 1017499.

[2] Alric, B et al. (2022) Macromolecular crowding limits growth under pressure. *Nat. Phys*. 18, 411–416.

[3] Delarue et al. (2018) mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding, *Cell*.

[4] Neurohr, GE et al. (2019) Excessive Cell Growth Causes Cytoplasm Dilution And Contributes to Senescence. *Cell* vol. 176,5: 1083-1097.e18.

[5] Joyner et al. (2016) A glucose-starvation response regulates the diffusion of macromolecules *eLife* 5: e09376.

[6] Chen, Y., Huang, JH., et al. (2024) Viscosity-dependent control of protein synthesis and degradation. *Nat Commun* 15, 2149.

[7] Burg TP et al. (2007) Weighing of biomolecules, single cells and single nanoparticles in fluid. *Nature* 446, 1066–1069.

[8] Park YK et al. (2018) Quantitative phase imaging in biomedicine. *Nature Photonics* 12, 578–589.

Acknowledgements: This work is funded by National Key Research and Development Program of China (2024YFA0919600) and SIAT startup funds.

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology,Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

[2]E-mail:yinxingci17@gmail.com

[3]E-mail: cyp7cn@gmail.com

Site-Specific Binding Protein Design via Reinforcement Learning

Yunxiao Dai[1,2], Xiaofang Zhong[1], Xiaoting Xu[1], Qian Li[1], Huanlun Li[1], and Yuping Chen[1,3]

***Short Abstract* — De novo protein design constitutes a fundamental challenge within protein engineering. Despite significant progress in prediction of protein-protein interactions (PPIs) between pairs of target and partner protein of known sequences, contemporary deep learning methodologies remain limited in generating binder protein sequences targeting specific sites on protein of interest, due to the immense, computationally intensive exploration of large sequence space. To address this, we propose a reinforcement learning-based generative model that designs potential binding partners tailored to interact with precise locations on target proteins. Model predictions will be experimentally validated using both purified recombinant protein systems and cell-based interaction assays.**

***Keywords* — De novo protein design, Protein-protein interaction, Deep learning, Reinforcement learning**

I. Purpose

Protein-protein interactions play a fundamental role in almost all biological processes and are central to protein design. Understanding PPIs provides insights into disease mechanisms, facilitates drug targets identification, and advances the development of therapeutic strategies[1-2]. Designing interacting partners also allows quick detection of proteins in research settings, and the detection of pathogens during pandemics, bypassing the need for labor-intensive and time-consuming antibody generation[3].

Despite significant progress, contemporary deep learning methodologies remain limited in their scope, particularly concerning the design of binders targeting specific sites on proteins[4-5].

II. Envisioned processes

A. Model construction and training

We developed a hybrid architecture model combining Graph Neural Networks (GNNs)[6] and Transformers[7] to serve as the policy model for reinforcement learning, using data derived from STRING[8] and PDB[9]. For the reward function, multiple dedicated models were trained to compute distinct functional components, such as protein-protein interaction (PPI) binding score and site-specificity score, which are then integrated to generate the total reward signal for policy model to generate binder amino acid sequence.

B. Validation

After obtaining potential candidates, we first tested protein interactions *in silico* via tools like PISA[10] and PRODIGY[11]. We plan to test the potential interactions using *in vitro* assays with recombinant proteins.

Drawing inspiration from the yeast mating system, where the α-factor peptide interacts with the Ste2 GPCR to halt the cell cycle creating a macroscopic phenotype, we will design modified GPCRs and express these modified receptors in yeast cells and test their ability to detect query proteins.

III. Conclusion

This project develops an efficient system for generative design of novel protein interacting partners. Our yeast-based detection method enables urgent protein detection without requiring costly antibody production/storage or specialized equipment, offering a practical alternative in resource-limited settings.

Acknowledgements: This work was funded by National Key Research and Development Program of China (2024YFA0919600) and SIAT startup funds.

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

[2] E-mail: yx.dai@siat.ac.cn

[3] E-mail: cyp7cn@gmail.com

References

[1] Nooren IMA., et al.(2003) New EMBO member's review: Diversity of protein-protein interactions. EMBO J, 22(14): 3486–3492.

[2] Marsh JA, Teichmann SA.(2015) Structure, dynamics, assembly, and evolution of protein complexes. Annu Rev Biochem. 84(1): 551 – 575.

[3] Mingda Xu, et al., 16 Apr 2024, Graph Neural Networks for Protein-Protein Interactions -- A Short Survey, arXiv:2404.10450

[4] BioGeometry Team, 6 May 2025, GeoFlow-V2: A Unified Atomic Diffusion Model for Protein Structure Prediction and De Novo Design, bioRxiv 2025.05.06.652551.

[5] Fengyuan Dai, et al., 2 Apr 2025, Toward De Novo Protein Design from Natural Language, bioRxiv 2024.08.01.606258.

[6] Zhou, Jie, et al.(2020) Graph neural networks: A review of methods and applications. AI open 1 : 57-81.

[7] Ashish Vaswani, et al.(2017) Attention is all you need. In Advances in neural information processing systems, pages 5998-6008.

[8] Szklarczyk et al. (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. ;51(D1):D638-D646.

[9] H.M. Berman, et al. The Protein Data Bank (2000) Nucleic Acids Research 28: 235-242.

[10] E. Krissinel and K. Henrick (2007). 'Inference of macromolecular assemblies from crystalline state.'. J. Mol. Biol. 372, 774--797.

[11] Li C. Xue, et al.(2016)PRODIGY: a web server for predicting the binding affinity of protein – protein complexes, Bioinformatics, Volume 32, Issue 23, December 2016, Pages 3676 – 3678

Regulation of mass density in mammalian cells

Xiaofang Zhong[1], Jiankun Liang[1], Xiaoting Xu[1], Qian Li[1], Yunxiao Dai[1], Huanlun Li[1], Wanze Chen[1], and Yuping Chen[1,2]

***Short Abstract* — Both cell mass and cell volume are regulated through complex mechanisms, yet cellular mass density - defined as the ratio of cell mass to cell volume - remains remarkably constant. The mechanism governing the coordination between the growth laws of cell mass and cell volume is unknown. To investigate the genetic basis of density regulation in mammalian cells, we aimed to separate multipotent stem cells, expressing a collection of over 100 transcription factors, into lighter or denser fractions from a density gradient centrifugation. Using barcode-sequencing, we identified transcription factors that influence cell mass density. We plan to conduct differentiation assays, both *in vitro* and *in vivo*, assessing the differentiation potential of cells with varying densities.**

***Keywords* — Cell mass density, transcription factors, differentiation potential.**

I. Purpose

CELL mass density is remarkably tightly regulated in mammalian cells of the same cell type (Liu, Oh et al. 2022). As cells differentiate into different cell types or as they age and senesce, their density changes (Neurohr, Terry et al. 2019, Oh, Lee et al. 2022). The mechanism of density maintenance and the cause and effect of density variation are unknown. Here, we set out to investigate the genetic basis for intracellular density maintenance in cycling cells and the mechanisms of mass density-influenced cellular decision-making.

It has been reported that cell mass density affects cell mass growth rate by changing the synthesis and degradation rates of molecules (Liu, Oh et al. 2020, Chen, Huang et al. 2024). Raghunath and Wohland found that macromolecules had an influence on the *ex vivo* differentiation of human mesenchymal stem cells (hMSCs) into the adipocytes (Rashid, Raghunath et al. 2011). To explore how cell mass density influences the differentiation abilities of cells, we will make use of *in vitro* and *in vivo* differentiation assays on cells of various densities.

II. Results

A. Next-Generation Sequencing (NGS)-based transcription factors screening of altered cell mass density

We recently developed an inducible murine transcription factor library covering over 100 transcription factors, enabling a genome-wide transcription factor screen for factors associated with cell density regulation.

We separated multipotent stem cells (C3H10T1/2 cells) expressing transcription factor combinations by Percoll gradient centrifugation according to their mass density. We sequenced DNA barcodes for each transcription factor in the separated cells with different mass density using NGS technology. Furthermore, we will perform functional genomic analysis to understand the regulatory network and potential mechanisms governing cell mass density.

B. Differentiation assays both in vitro and in vivo

After obtaining dense and light cells, we are investigating whether mass density of a cell influences cellular differentiation potential both *in vitro* and *in vivo*. Cells with different mass densities will undergo directed differentiation assays *in vitro*. Additionally, we plan to transplant purified stem cells with different mass densities into mice and analyze cellular differentiation *in vivo*.

III. Conclusion

This study proposes an experimental approach to study cell mass density in mammalian cells using a transcription factor library. We aim to identify transcription factors that alter cell mass density and determine the differentiation potential of cells with varying mass density using both *in vitro* and *in vivo* experiments. This study will shed light on understanding a key physical parameter of cells during their developmental process.

References

Chen, Y., et al. (2024). "Viscosity-dependent control of protein synthesis and degradation." Nature Communications **15**(1): 2149.

Liu, X., et al. (2022). "The uniformity and stability of cellular mass density in mammalian cell culture." Frontiers in Cell and Developmental Biology **10**: 1017499.

Liu, X., et al. (2020). "Computationally enhanced quantitative phase microscopy reveals autonomous oscillations in mammalian cell growth." Proceedings of the National Academy of Sciences **117**(44): 27388-27399.

Neurohr, G. E., et al. (2019). "Excessive cell growth causes cytoplasm dilution and contributes to senescence." cell **176**(5): 1083-1097. e1018.

Oh, S., et al. (2022). "Protein and lipid mass concentration measurement in tissues by stimulated Raman scattering microscopy." Proceedings of the National Academy of Sciences **119**(17): e2117938119.

Rashid, R., et al. (2011). "Macromolecular Crowding and Stem Cell Differentiation." Biophysical Journal **100**(3): 142a-143a.

Acknowledgements: This work was funded by National Key Research and Development Program of China (2024YFA0919600) and SIAT startup funds.

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China.

[2]Email: cyp7cn@gmail.com.

Transition from global stability to multiple attractors driven by strong species interactions in bacterial microcosms

You He[1*], Jiliang Hu[2*], Jinyeop Song[3] and Jeff Gore[4]

***Short Abstract* — Ecological communities with a fixed set of species can reach alternative stable states with distinct properties of the community. However, the factors determining whether a community will display multistability, as well as the shape of the landscape connecting these multiple states, remain poorly understood. Theory predicts that both stable and fluctuating attractors can coexist within the same community, but direct observations of such multiple dynamical attractors remain rare. Here we assemble ~100 bacterial communities to show that increasing species interaction strength leads to a transition from global equilibrium to alternative stable states, as predicted by Lotka-Volterra model. Moreover, our experiments revealed two community-level regimes with distinct properties on biomass and environmental pH. Interestingly, some multi-stable communities reach both regimes and with compositional multi-stability within a given property regime, thus exhibiting "hierarchical multi-stability". We also found that a single community can exhibit multiple stable and fluctuating dynamical attractors, as predicted by theory. Our work demonstrates that increasing species interactions leads to a transition from global stable state to multiple attractors in microbial communities, revealing a stable landscape driven synergistically by high-dimensional species interaction networks and low-dimensional environmental coupling.**

***Keywords* — alternative stable states, global stability, environmental coupling, multiple attractors**

I. Introduction

ECOLOGICAL communities composed of a fixed set of species can reach alternative stable states, characterized by distinct species abundances and community properties[1]. These different stable states can be reached depending on the timing of species arrival and environmental perturbations [2]. The phenomenon of multi-stability has gained prominence in the study of species-rich microbial communities, spanning ecosystems such as soil, aquatic environments, and the human microbiome [3-5]. In human-associated microbiomes, multi-stability can profoundly influence community properties and health outcomes. For example, in the vaginal microbiome, a healthy state dominated by lactic acid bacteria maintains a low pH that protects against infections, while a dysbiotic state characterized by higher pH increases susceptibility to bacterial vaginosis and other complications [3].

Despite significant advances in studying multi-stability, the conditions under which it arises and the factors that drive multi-stability in complex communities remain poorly understood [6]. Recent theoretical work has predicted that complex interaction networks in species-rich models can give rise to multi-stability [7-8], yet experimental validation is lacking. Moreover, the multi-stability landscape remains poorly understood, particularly whether it resembles a simple bi-stable system or exhibits a more complex structure [9]. Furthermore, it is unclear whether multi-stability is driven synergistically by low-dimensional environmental variables and high-dimensional interaction networks and whether these mechanisms act independently.

II. Results

A. Transition from global stability to multi-stability

We discovered that increasing species interaction strength induces a transition from global equilibrium to alternative stable states, as predicted by the Lotka-Volterra model and validated experimentally.

B. Community regimes driven by environmental coupling

Our experiments revealed two community-level regimes with bi-modal properties: an acidic regime with low biomass, an alkaline regime with high biomass. We observed both functional bi-stability, governed by a low-dimensional environmental variable (pH), and compositional multi-stability, driven by the complex interspecies interaction network. Some communities reach both regimes while also exhibiting multistability within a given regime, a phenomenon we term "hierarchical multistability".

C. Multiple dynamical attractors in microbial communities

Our experiment revealed that a single community can exhibit multiple stable and fluctuating dynamical attractors, as predicted by theory.

References

[1] Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. Nature 413, 591–596 (2001).
[2] Amor, D. R., Ratzke, C. & Gore, J. Transient invaders can induce shifts between alternative stable states of microbial communities. Sci. Adv. 6, eaay8676 (2020).
[3] Ravel, J. et al. Vaginal microbiome of reproductive-age women. Proc. Natl. Acad. Sci. 108, 4680–4687 (2011).
[4] Tropini, C. et al. Transient Osmotic Perturbation Causes Long-Term Alteration to the Gut Microbiota. Cell 173, 1742-1754.e17 (2018).
[5] Sonnenburg, E. D. et al. Diet-induced extinctions in the gut microbiota compound over generations. Nature 529, 212–215 (2016).
[6] Gonze, D., Lahti, L., Raes, J. & Faust, K. Multi-stability and the origin of microbial community types. ISME J. 11, 2159–2166 (2017).
[7] Biroli, G., Bunin, G. & Cammarota, C. Marginally stable equilibria in critical ecosystems. New J. Phys. 20, 083051 (2018).
[8] Ros, V., Roy, F., Biroli, G., Bunin, G. & Turner, A. M. Generalized Lotka-Volterra Equations with Random, Nonreciprocal Interactions: The Typical Number of Equilibria. Phys. Rev. Lett. 130, 257401 (2023).
[9] Estrela, S. et al. Functional attractors in microbial community assembly. Cell Syst. 13, 29-42.e7 (2022).

[1]Xingjian College, Tsinghua University, China. E-mail: youhe@mit.edu
[2]Department of Physics, Massechusetts Institute of Technology, USA. E-mail: jiliang@mit.edu
[3]Department of Physics, Massechusetts Institute of Technology, USA. E-mail: yeopjin@mit.edu
[4]Department of Physics, Massechusetts Institute of Technology, USA. E-mail: gore@mit.edu
*These authors contributed to the work equally

Predicting function of evolutionarily implausible DNA sequences

Shiyu Jiang[1,2], Xuyin Liu[1], and Zitong Jerry Wang[1]

***Short Abstract* — Genomic language models (gLMs) show potential for generating novel, functional DNA sequences for synthetic biology, but doing so requires them to learn not just evolutionary plausibility, but also sequence-to-function relationships. We introduce a set of prediction tasks called NULLSETTES, which assesses a model's ability to predict loss-of-function mutations created by translocating key control elements in synthetic expression cassettes. Across 12 state-of-the-art models, we find that mutation effect prediction performance strongly correlates with the predicted likelihood of the nonmutant. Furthermore, the range of likelihood values predictive of strong model performance is highly dependent on sequence length. Our work highlights the importance of considering both sequence likelihood and sequence length when using gLMs for mutation effect prediction.**

***Keywords* — Genomic language models, gene expression, mutation prediction, synthetic biology**

I. Background

GENOMIC language models (gLMs) learn a probability distribution over DNA sequences, representing the evolutionary plausibility of genomic sequences, which can be a useful proxy for biological function [1, 2]. However, current evaluations for gLM on mutant effet prediction focus primarily on natural sequences. In contrast, synthetic biology often requires the design of functional sequences with little or no evolutionary precedent, such as sequences that confer novel function, avoid crosstalk with native machinery, or push expression levels beyond natural limits [3-5]. It is unclear whether gLMs trained on natural genomes can generalize to synthetic constructs, especially in the absence of deep mutational scanning data for DNA. If successful, gLMs could support more systematic and scalable approaches to genetic design.

II. Methodology

NULLSETTES are synthetic expression cassettes created by permuting six regulatory elements in expression cassettes to disrupt transcription or translation, which includes 11 eukaryotic and 19 prokaryotic variants. We curated expression cassettes with high expression but low gLM likelihood, reflecting low evolutionary plausibility. We label datasets with random promoters as "Low" and those with natural motifs as "High", benchmarked 12 state-of-the-art genomic foundation models spanning diverse tokenization schemes, pretraining objectives, training corpora, and architectures. Model performance was evaluated by comparing log-likelihood (LL) distributions between original cassettes and NULLSETTES.

III. Results

In our benchmarking, Evo series consistently achieves competitive performance across four Massive Parallel Reporter Assay datasets. Overall, Evo-2-7B performs the best, especially on low-likelihood, synthetic sequences. We also find a strong correlation between model-assigned sequence likelihood and ability to detect LOF mutations, highlighting evolutionary plausibility as a proxy for function. Furthermore, we show that the optimal likelihood thresholds for mutation prediction vary with sequence length, stressing the need to account for both likelihood and length in synthetic design.

IV. Conclusion

In this study, we introduce NULLSETTES, the first systematic benchmark for assessing the ability of gLMs to predict loss-of-function mutations in synthetic expression cassettes. Beyond general performance benchmarking, our results underscore the importance of considering both likelihood and length when applying gLMs to synthetic sequence design. We hope this benchmark and its findings will inspire the development of gLMs that move beyond evolutionary priors to support the rational design of functional, out-of-distribution genetic sequences.

References

1. Consens, M.E., et al., *Transformers and genome language models.* Nature Machine Intelligence, 2025: p. 1-17.
2. Benegas, G., et al., *Genomic language models: opportunities and challenges.* Trends in Genetics, 2025.
3. Yan, X., et al., *Applications of synthetic biology in medical and pharmaceutical fields.* Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 199.
4. Schlabach, M.R., et al., *Synthetic design of strong promoters.* Proceedings of the National Academy of Sciences, 2010. **107**(6): p. 2538-2543.
5. Du, R., et al., *miRNA circuit modules for precise, tunable control of gene expression.* bioRxiv, 2024.

[1]Center for Interdisciplinary Studies, School of Science, Westlake University, Hangzhou, China.
[2]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA.
Correspondence to: Shiyu Jiang <shiyujia@usc.edu>, Zitong Jerry Wang <jerry@westlake.edu.cn>

Biophysical Constraints on mRNA Diffusion and Translation

Huanlun Li[1, 2], Xiaofang Zhong[1], Qian Li[1], Xiaoting Xu[1], Yunxiao Dai[1] , and Yuping Chen[1, 3]

Short Abstract — **The spatial organization and translation efficiency of mRNA are intricately regulated by the biophysical properties of the cellular environment. Yet, the relationship between mRNA spatial dynamics and translation dynamics remains poorly understood. In this study, we systematically measure mRNA diffusion and translation efficiency across a range of molecular sizes and cytoplasmic conditions. Our findings reveal how mRNA size and cytoplasmic physical properties jointly regulate mRNA translation, providing a quantitative framework for understanding the spatiotemporal mechanisms of protein synthesis.**

***Keywords* — Protein translation, mRNA, molecular diffusion, cytoplasmic environment**

I. Purpose

THE cytoplasm is a highly crowded environment densely packed with macromolecules [1], exhibiting pronounced heterogeneity and viscoelastic properties. Previous studies have demonstrated that intracellular protein concentration is optimized to maximize biochemical reaction rates [2], and that protein synthesis and degradation exhibit differential sensitivity to cytoplasmic concentration[3]. This disparity may be attributed to the size-dependent diffusion behavior of biomacromolecule assemblies, such as polyribosomes, under varying physical properties of the cytoplasm. However, this idea has not been previously tested.

In this study, we design mRNAs of different lengths and manipulate the physical properties of mammalian cells that expresses these mRNAs. Using single particle tracking, we systematically investigate the diffusion dynamics of these mRNAs in cytoplasmic environments with distinct viscosity and molecular crowding. In parallel, we quantify translation rates under the same conditions to elucidate how mRNA size and the physical properties of the cytoplasm jointly shape protein translation rate in the innate environment of cells.

II. Experimental progress

We have designed a series of mRNA constructs with varying lengths by adjusting the lengths of the coding sequence (CDS) and the 3' untranslated region (3'UTR). Mango RNA aptamers were incorporated into the 3'UTR to enable real-time imaging of mRNA transcriptional status and spatial distribution through specific binding with fluorogenic dyes[4-5]. Additionally, a miniGFP2 coding sequence was included to visualize translation products and monitor dynamic changes in translation rates.

We are currently generating multiple stable cell lines, each expressing mRNAs of defined sizes. Lentiviral-mediated genomic integration combined with clonal selection is employed to ensure consistent fluorescence signals and stable expression across cells. Subsequent experiments will utilize confocal fluorescence microscopy to perform quantitative imaging of mRNA behavior under varying cytoplasmic conditions, such as changes in viscosity and macromolecular crowding induced by osmotic treatments. We will analyze diffusion trajectories, mean square displacement (MSD), and diffusion coefficients, and apply time-resolved fluorescence measurements to assess transcriptional and translational activity. These data will form the foundation for developing physical models of mRNA dynamics in different cytoplasmic environments.

III. Conclusion

We anticipate that mRNAs of different sizes will exhibit distinct diffusion behaviors within the cytoplasm, with larger mRNAs showing more restricted mobility under conditions of high viscosity or molecular crowding. In parallel, mRNA size may influence transcription and translation rates across varying cytoplasmic environments, reflecting a coupling between the biophysical properties of the cellular environment and the regulation of gene expression.

Acknowledgements: This work is funded by National Key Research and Development Program of China (2024YFA0919600) and SIAT startup funds.

[1]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology,Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

[2] E-mail: hl.li2@siat.ac.cn

[3] E-mail:cyp7cn@gmail.com

References

[1] Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. International Review of Cytology, 192, 189–22

[2] Dill, K. A., Ghosh, K., & Schmit, J. D. (2011). Physical limits of cells and proteomes. Proceedings of the National Academy of Sciences of the United States of America, 108(44), 17876-17882

[3] Chen, Y., Huang, J.H., Phong, C., & Ferrell, J. E., Jr. (2024). Viscosity-dependent control of protein synthesis and degradation. Nature Communications, 15(1), 2149.

[4] Autour, A., Jeng, S. C. Y., Cawte, A. D., Abdolahzadeh, A., Galli, A., Panchapakesan, S. S. S., Rueda, D., Ryckelynck, M., & Unrau, P. J. (2018). Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. Nature Communications, 9(1), 656.

[5] Cawte, A. D., Unrau, P. J., & Rueda, D. S. (2020). Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. Nature Communications, 11(1), 1283.CNLS web site, optional date, http://cnls.lanl.gov/

Decoding stochastic dynamics from single-cell snapshot data: towards therapeutic target identification in type 2 diabetes

Jinchao Lv[1,2], Yunhui Kong[3], Gaowei Wang[3], Chunhe Li[2,1,4]

***Abstract*— Quantitative analysis of stochastic biological systems requires high-resolution data, but single-cell sequencing lacks temporal resolution for reconstructing underlying landscapes. Here we present LAPIS (Landscape reconstruction And Parameter Inference from Single-snapshot data), a method integrating single-cell snapshot data with stochastic differential equation modeling to estimate gene-associated parameters and noise intensities. LAPIS preserves both data fidelity and biophysical realism, validated on in-silico 2-dimensional mutual inhibition self-activation and 52-dimensional human embryonic stem cell systems. Applied to in-vivo single-cell snapshot data in type 2 diabetes, LAPIS reconstructs landscapes for healthy and diseased populations, identifying NEUROD1 and ATF6 as therapeutic targets. Overall, LAPIS enables biologically meaningful predictions for exploring therapeutic strategies in complex diseases.**

***Index Terms*— energy landscape, dynamics inference, stochastic dynamics, gene regulatory networks, type 2 diabetes**

I. Purpose

Integrating model- and data-driven methods to quantitatively characterize the stochastic dynamics of gene regulatory networks is a challenging problem. Our aim is to utilize existing snapshot experimental data to better estimate the unknown parameters in the model and noise intensities (rather than arbitrarily assigning them). Existing methods employ linear models to model regulatory interactions [1], [2], despite the strong nonlinearity inherent in gene regulation. This allows for the reconstruction of landscapes through differential equation simulations, in order to quantify the roles of individual genes and provide potential biological predictions.

II. Model and Methods

The evolution of gene expression is described by a refined additive model:

$$\frac{dx_i}{dt} = g_i - \frac{g_i}{\max_{1\le i\le N}\left\{\sum_{j=1}^{N} D_{ji}\right\}} \cdot \sum_{j=1}^{N} \frac{D_{ji}\times x_j^n}{S_j^n + x_j^n} - k_i x_i + \frac{G_i - g_i}{\max_{1\le i\le N}\left\{\sum_{j=1}^{N} C_{ji}\right\}} \cdot \sum_{j=1}^{N} \frac{C_{ji}\times x_j^n}{S_j^n + x_j^n} + \zeta_i(t).$$

[1]Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China. Email: 21110850008@m.fudan.edu.cn

[2]Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China. Email: chunheli@fudan.edu.cn

[3]Institute of Modern Biology, Nanjing University, Nanjing, China. Email: wanggaowei@nju.edu.cn

[4]School of Mathematical Sciences, Fudan University, Shanghai, China.

g_i: basal production rate, G_i: maximum production rate, k_i: degradation rate, n: Hill coefficient, S_j: regulation threshold, C_{ji}/D_{ji}: activation/inhibition strength (correlation coefficient), $\left\langle \zeta_i(t), \zeta_j(t') \right\rangle = 2d\delta_{ij}\delta(t-t')$ (d: diffusion coefficient).

We aim for the model-predicted steady-state distribution to fit the observed data distribution as accurately as possible:

$$\min \text{Loss} = \min \sum_{j=1}^{M} \beta_j \text{KL}\left(p_j(x) \| q_j(x)\right) + \frac{\lambda}{2}\|\alpha - \beta\|^2,$$

α_j/β_j: the j^{th} state weight obtained by the model/data, p_j/q_j: the j^{th} state density function estimated from the model/data, λ: trade-off parameter.

Under the assumption that the steady-state distribution can be approximated as Gaussian using truncated moment equations, the gradient of the KL divergence with respect to all parameters can be computed explicitly:

$$\frac{\partial}{\partial\theta}\text{KL}\left(p^j(x)\|q^j(x)\right) = \left(\mu_p - \mu_q\right)^T \Sigma_q^{-1} \frac{\partial}{\partial\theta}\mu_p - \frac{1}{2}\frac{1}{\det(\Sigma_p)}\frac{\partial}{\partial\theta}\det(\Sigma_p) + \frac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left(\Sigma_q^{-1}\right)_{kj} \cdot \frac{\partial}{\partial\theta}\left(\Sigma_p\right)_{jk},$$

The step size in LAPIS is determined using the Barzilai-Borwein method, while simulated annealing (Metropolis criterion) is incorporated to mitigate the issue of local minima.

III. Results and Conclusion

- With appropriate initial parameter settings, LAPIS is capable of achieving accurate parameter inference, almost fully reconstructing the system's stochastic dynamics and landscape from in-silico data.
- Based on optimal control, it is possible to achieve the transition from diseased to healthy state by continuously modulating all production rate-related parameters.
- By integrating single-parameter sensitivity analysis, we effectively identified critical therapeutic targets and potential combination strategies. Among them, HNF4A, HNF1A, and NEUROD1 are well supported by experimental evidence, while NFIA represents a novel prediction generated by LAPIS.

References

[1] L. Wang, N. Trasanidis, T. Wu, G. Dong, M. Hu, D. E. Bauer, and L. Pinello, "Dictys: dynamic gene regulatory network dissects developmental continuum with single-cell multiomics," *Nature Methods*, vol. 20, no. 9, pp. 1368–1378, 2023.

[2] F. Bocci, P. Zhou, and Q. Nie, "splicejac: transition genes and state-specific gene regulation from single-cell transcriptome data," *Molecular systems biology*, vol. 18, no. 11, p. e11176, 2022.

Emergent simplicity in bacterial growth under complex nutrient envrionments

Aoyu Zhu[1], and Po-Yi Ho[1]

***Short Abstract* — Bacterial metabolism is dependent on the nutrient environment and variable across species. However, to what extent this complexity can be explained by simple, quantitative principles remains unclear. Here, we show that the number of nutrients utilized emerges as a key determinant of growth under complex environments, both experimentally and in silico. By optimizing growth under proteome constraints, we obtained flux solutions for genome-scale metabolic networks across diverse species. Notably, we find that with increasing environmental complexity, more nutrients directly support biomass precursor synthesis. To uncover the network features underlying this trend, we coarse-grained metabolic networks using representation learning and found efficient pathways between reaction clusters that support high growth at low proteome cost. The activation of these pathways is highly correlated to the number of nutrients utilized, not specific metabolites, and predicts growth across environments and species.**

I. Purpose

BACTERIAL growth is fundamental to biology. In simple defined environments, diverse nutrient utilization strategies—such as co-utilization and diauxie—have been extensively studied and quantitatively explained through enzyme kinetics, gene regulation, and mathematical models[1]. However, for bacterial growth in complex environments—far more representative of nature—the underlying quantitative principles remain elusive.

Recent studies in gut bacteria reveal a surprising phenomenon that biomass yield is not proportional to nutrient concentration, but to the number of nutrients utilized[2]. This emergent top-down parameter challenges traditional views and suggests a new dimension for quantifying bacterial growth. However, the metabolic mechanisms behind this proportionality remain unclear. Here, we aim to uncover its origins using a systems biology framework.

II. Results

A. Bacterial growth across diverse species is proportional to the number of nutrients utilized in complex environments

Using genome-scale metabolic models (GEMs) from AGORA2 and a simplified proteome-constrained FBA (spcFBA), we reproduced this growth-nutrient proportionality in silico. To assess generality, we simulated randomized complex environments with equal nutrient counts, confirming the robustness of the trend. However, in simpler environments, the proportionality weakened, suggesting that environmental complexity stabilizes nutrient-growth relationships.

B. Environmental complexity promotes nutrient specialization and synergistic interactions

To clarify how nutrients drive this phenomenon, we developed a mass flow graph (MFG) based on GEMs and spcFBA. We systematically quantified nutrient function by tracing probabilities of nutrient-to-biomass paths. In *E.coli*, our method correctly identified ammonia and sulfate as precursors with specific roles in cysteine and methionine synthesis. In full media, nutrients clustered into two types: *omnipotent* (involved in many subsystems) and *specialized* (direct biomass contributors). The proportion of specialized nutrients increases in complex environments, explaining the emergence of proportionality via uniform nutrient contribution.

C. Nutrient synergy drives activation of efficient pathways supporting proportional growth in complex environments

To uncover the underlying mechanism, we developed a coarse-graining method for mass flow graphs (MFGs), allowing simplified and comparable analysis despite structural variability. Despite this abstraction, the coarse-grained MFGs retained key metabolic features, such as nutrient synergy and functional diversity. Using heuristic nutrient dropout, we systematically identified efficient and inefficient metabolic pathways. We found that in complex environments, proportionality arises because a greater number of utilized nutrients enhances synergy and activates more metabolic efficient pathways. Crucially, the number of active metaboli efficient pathways scales linearly with bacterial growth rate.

III. Conclusion

Bacterial growth proportionality with the number of nutrients utilized in complex environments emerges from enhanced nutrient specialization and synergistic effects, which collectively activate more efficient metabolic pathways to support faster growth.

References

[1] Erickson, D, et al. (2017) A global resource allocation strategy governs growth transition kinetics of Escherichia coli. Nature **551**, 119–123.

[2] Ho, PY, et al. (2024) Resource competition predicts assembly of gut bacterial communities in vitro. Nat Microbiol **9**, 1036–1048.

[1]Center for Interdisciplinary Studies (CIS), Westlake University, Hangzhou, China. E-mail: zhuaoyu@westlake.edu.cn

Modeling the Spatial Dynamics of Mitochondrial Proteins in a Branched Organelle Network

Chenxi Liu[1], Haohang Lin[1], Abdul Haseeb Khan[1], Qinghe Wang[1], Tatsuhisa Tsuboi[1]*

Short Abstract — **Mitochondria in Saccharomyces cerevisiae form a dynamic and highly branched network, where the spatial distribution of surface proteins is closely associated with mitochondrial function. To investigate how proteins are organized within this complex structure, we developed a multi-scale modeling framework that combines high-resolution imaging, graph-theoretical reconstruction, and physical simulations. Our findings demonstrate that protein distribution patterns emerge from constrained Brownian motion within a geometrically and topologically realistic mitochondrial scaffold. This framework reveals the coupling between protein localization and mitochondrial architecture, offering new insights into organelle-level regulation.**

Keywords — **mitochondria, protein distribution, 3D modeling, Brownian motion, graph theory, molecular simulation**

I. Purpose

Mitochondria in S. cerevisiae exhibit a highly dynamic and reticulated tubular structure, playing essential roles in energy production, cell signaling, and apoptosis. While individual mitochondrial proteins have been well characterized, the mechanisms underlying their spatial organization within the 3D mitochondrial network remain poorly understood. The distribution of proteins along the mitochondrial surface may reflect the organelle's functional and structural state, yet quantitative models to study this relationship are lacking. Our study aims to explore the spatial organization of mitochondrial surface proteins using a comprehensive framework that integrates experimental imaging with physical modeling.

II. Preparation of Abstracts

To systematically analyze protein localization, we first reconstructed the 3D mitochondrial architecture from high-resolution microscopy data of S. cerevisiae. By applying graph-theoretical methods, we extracted key topological features such as branching patterns and connectivity, and restored the spatial geometry of the network. This 3D scaffold served as the structural basis for simulating protein motion and interaction.

We developed a constrained Brownian motion model to simulate protein dynamics on the mitochondrial surface. This model incorporates local geometric constraints, topological boundaries, and interactions between proteins, enabling the capture of stochastic diffusion behavior within a complex subcellular environment. This framework provides a platform for quantitatively exploring how protein mobility and localization are influenced by the structural features of the mitochondrial network.

III. Conclusion

Our multi-scale modeling and simulation approach successfully recapitulates experimentally observed patterns of protein distribution on mitochondrial networks. The simulation results revealed that protein spatial distributions are non-uniform and are closely associated with structural features of the network, such as curvature and branching junctions. This finding suggests that protein localization is not solely governed by random diffusion or uniform translation but is instead modulated by the geometric and topological properties of the mitochondria. The results highlight the importance of mitochondrial geometry and topology in regulating protein localization. This study provides theoretical insights and computational tools for understanding the organization of mitochondrial proteins.

[1] Shenzhen International Graduate School, Tsinghua University. E-mail: ttsuboi@sz.tsinghua.edu.cn

Quantitative Investigation of Metabolic Treatments for Alzheimer's Disease

Ruiqin Liu[1] and Shiyu Liu[2]

***Short Abstract* —**

Beyond traditional strategies targeting Aβ (Amyloid β) and Tau proteins, metabolic treatments for Alzheimer's disease (AD)—including ketogenic diets, semaglutide, and low-protein diets—have demonstrated notable effectiveness in animal models. These treatments present promising benefits due to their safety, cost-effectiveness, and accessibility. However, the specific metabolic pathways targeted by these therapies are not well understood, hindering the optimization of their therapeutic effects. To address this, we utilize stable isotope metabolic flux analysis (MFA) in an AD mouse model to explore body-level metabolic responses to these treatments. By analyzing the flux interactions, we aim to uncover systematic metabolic alterations across various organs, providing valuable insights to improve existing metabolic treatments and guiding the development of new therapeutic strategies.

***Keywords* — Alzheimer's disease, Metabolic Flux Analysis (MFA), ketogenic diet, semaglutide, low-protein diet, Multi-organ metabolic network .**

I. Background

ALZHEIMER'S disease (AD) is a leading cause of cognitive impairment, accounting for 60% to 70% of dementia cases worldwide. As of 2020, approximately 50 million people are living with dementia, and this number is projected to rise to around 130 million by 2050. Traditional therapies for AD primarily target Aβ (amyloid-beta) plaques and Tau protein tangles, which are central to the pathogenesis of the disease. Nevertheless, an increasing number of alternative strategies are explored to slow down cognitive decline in AD.

In recent years, metabolic factors have gained significant attention in AD research. PET-CT studies have revealed a 21%-30% reduction in cerebral glucose metabolism in AD patients [1], which has led to the development of new treatment strategies targeting metabolic pathways. Approaches such as ketogenic diets, semaglutide, and low-protein diets have shown promising results in AD animal models[2-7]. However, the specific targets and mechanisms of these therapies remain poorly understood. This lack of clarity impedes the optimization of existing treatments and hinders the development of new therapeutic strategies.

To address this gap, we use stable isotope-based metabolic flux analysis (MFA) to construct a multi-organ metabolic flux map in the context of AD metabolic therapies. Our goal is to identify the core metabolic targets of these treatments, uncover new mechanisms of intervention, and ultimately improve therapeutic outcomes.

[1]Chinese Institutes for Medical Research, Beijing, China. E-mail: liuruiqin@cimrbj.ac.cn

[2]Chinese Institutes for Medical Research, Beijing, China. E-mail: shiyu.liu@cimrbj.ac.cn

II. Research Methods:

We used APP/PS1 transgenic mice as AD mice model and C57BL/6J mice as the control group. The mice were subjected to different treatments, including a ketogenic diet, drug-based ketogenic therapy, semaglutide, and a low-protein diet. After 3, 6, and 20 weeks of treatment, behavioral assessments and ^{13}C isotope perfusion experiments were conducted. Additionally, mass spectrometry analyses were performed on serum, brain, heart, liver, kidney, adipose tissue, and muscle samples from the mice.

We constructed a metabolic reaction pathway network using computational programs and quantified the metabolic fluxes. By analyzing the quantitative metabolic response pathways, we identified the pathways with the most significant changes. To further explore potential therapeutic targets, we employed transgenic mice and RNA-Seq technology. These methods provided insights into the underlying mechanisms of metabolic treatments and identified new targets for intervention.

III. Conclusion

This study demonstrates that fluxes in the tricarboxylic acid (TCA) cycle in the brain show significant differences between wildtype (WT) and AD model mice under normal and ketogenic diets, suggesting that TCA cycle reactions may be key points in slowing the progression of AD. We will extend this approach to investigate the effects of semaglutide and low-protein diets. Compared to static metabolomic measurements, metabolic flux analysis (MFA) offers a distinct advantage in quantifying *in vivo* pathway activity, enabling the identification of key pathways driving therapeutic benefits. Our findings aim to uncover clear therapeutic targets for intervention in Alzheimer's disease.

References

[1] Han W, Zhang B, Zhao W,*et al*. Ketogenic β-hydroxybutyrate regulates β-hydroxybutyrylation of TCA cycle-associated enzymes and attenuates disease-associated pathologies in Alzheimer's mice. Aging Cell. 2025 Jan;24(1):e14368.

[2] Phillips MCL, Deprez LM, Mortimer GMN, *et al*. Randomized crossover trial of a modified ketogenic diet in Alzheimer's disease. Alzheimers Res Ther. 2021 Feb 23;13(1):51.

[3] Di Lucente J, Persico G, Zhou Z, *et al*. Ketogenic diet and BHB rescue the fall of long-term potentiation in an Alzheimer's mouse model and stimulates synaptic plasticity pathway enzymes. Commun Biol. 2024 Feb 16;7(1):195.

[4] Qin Y, Bai D, Tang M, *et al*. Ketogenic diet alleviates brain iron deposition and cognitive dysfunction via Nrf2-mediated ferroptosis pathway in APP/PS1 mouse. Brain Res. 2023 Aug 1;1812:148404.

[5] Wang ZJ, Li XR, Chai SF, *et al*. Semaglutide ameliorates cognition and glucose metabolism dysfunction in the 3xTg mouse model of Alzheimer's disease via the GLP-1R/SIRT1/GLUT4 pathway. Neuropharmacology. 2023 Dec 1;240:109716.

[6] Zhang Y, Tang C, He Y, et al. Semaglutide ameliorates Alzheimer's disease and restores oxytocin in APP/PS1 mice and human brain organoid models. Biomed Pharmacother. 2024 Nov;180:117540.

[7] Zhai Y, Lu K, Yuan Y, et al. Semaglutide improves cognitive function and neuroinflammation in APP/PS1 transgenic mice by activating AMPK and inhibiting TLR4/NF-κB pathway. J Alzheimers Dis. 2025 May;105(2):416-432.

A simple cross-feeding structure explains coexistence and multistability in gut microbiomes *in vitro*

Zhenyu Wang[1], Po-yi Ho[1]

***Short Abstract —Emergent coexistence*, where some species survive in communities even though they cannot grow alone, was recently discovered. Understanding the mechanisms behind this phenomenon could help uncover the principles governing microbiome stability and diversity. In this study, using a controlled *in vitro* system, we found that pairwise cross-feeding is key to enabling emergent coexistence. We showed that non-growing species can coexist in communities by occupying unique metabolic niches. We also observed one species showing bi-stability in co-culture but not in community, suggesting that complex interactions stabilize community dynamics.**

***Keywords* — Cross-feeding, Coexistence, Multistability, Synthetic community**

I. Introduction

THE gut microbiome is highly diverse and stable over time. Understanding the ecological rules behind this diversity and stability is crucial for advancing gut microbiome research and navigate microbiome engineering.

One recent study have reported a phenomenon known as *emergent coexistence*, where certain gut bacterial species can survive in community settings despite being unable to grow in monoculture. However, the mechanisms driving this phenomenon remain unclear.

In this study, using a controlled *in vitro* system, we found that pairwise cross-feeding is a key mechanism underlying emergent coexistence. By integrating metabolomics and a coarse-grained ecological framework, we revealed non-growing species depended on occupying unique metabolic niches to coexist in community. Notably, All but one species exhibited bi-stability in co-culture with model strain *B. thetaiotaomicron* (Bt), but not in the full community—highlighting that an simple cross-feeding structures drives emergent coexistence and community dynamics.

II. Results

A. Emergent coexistence is widespread across different nutrients landscape

We established a representative collection of 45 gut bacterial strains and a chemically defined medium . Monoculture growth and community assembly revealed that 38 of 45 strains could grow in at least one medium in community settings, while only 24 of 45 could grow in monoculture. This gap indicates that emergent coexistence is widespread, consistent with previous reports.

B. Pairwise cross-feeding reproduced emergent coexsitence

Using glucose as the sole carbon source, we classified the 45 strains as growers (24/45) and non-growers (21/45). Pairwise co-culture and spent medium assays showed that non-growers could grow when co-cultured with other strains or exposed to their supernatants, indicating that pairwise cross-feeding enables their survival. These results suggest that pairwise metabolic interactions are central to emergent coexistence.

C. Bacteria needs unique niche to support its emergent coexistence

Using spent medium assays with the model strain *Bt*, we found that only non-growers with unique metabolic niches—those not overlapping with others—were able to persist in community contexts.

D. Cross-feeding rare leads to bistability

Theoretical models have predicted that cross-feeding can lead to bi-stability. To test this, we performed co-cultures of selected 16 non-growers with *Bt* under varying initial ratios. Only one species, *Mi*, exhibited bi-stability in co-culture but not in the full community.

III. Conclusion

In present study, we found that pairwise cross-feeding is a key mechanism enabling emergent coexistence. Further, we revealed an simple cross-feeding structures drives emergent coexistence and community dynamics.

References

[1] Naomi (2024) Emergent survival and extinction of species within gut bacterial communities. ***Biorxiv***, 2024.04.29.591619.

[1]Center for Interdisciplinary Science (CIS), Westlake University, Hangzhou, China. E-mail: wanzhenyu20@westlake.edu.cn

Quantifying Body-level Metabolic Rewiring in Obesity Development and Treatment Using In Vivo Metabolic Flux Analysis

Jianyun Zhang[1], Chong Zhou[2], and Shiyu Liu[3]

***Short Abstract* — Obesity leads to significant metabolic rewiring of carbohydrates, lipids, and amino acids across various organs. Despite the commercial success of treatments such as GLP-1 receptor agonists (GLP-1RAs) and bariatric surgery, the underlying mechanisms remain poorly understood. This lack of clarity hinders improvements in treatment efficacy and the mitigation of side effects, such as muscle loss. To address this, we utilize in vivo stable isotope-based metabolic flux analysis (MFA) in mice to construct a quantitative map of metabolic alterations in organs throughout obesity development and treatment. This approach will reveal systemic metabolic rewiring and identify specific flux changes in pathways such as central carbon metabolism, amino acid metabolism, ketogenesis, and lipid metabolism. Our goal is to provide a detailed mechanistic understanding of obesity and its treatments, paving the way for the development of combination therapies that maximize weight loss while minimizing side effects.**

***Keywords* — Obesity, Metabolic Flux Analysis (MFA), GLP-1, Bariatric Surgery, Sleeve Gastrectomy, Combination Therapy**

I. Background

THE obesity involves profound metabolic rewiring across the body, impacting the pathways of carbohydrates, lipids, and amino acids in key organs [1]. While some details of these changes are known, the specific metabolic fluxes that are altered during the development of obesity and in response to treatment remain poorly quantified [2]. Current leading treatments, including GLP-1RAs and bariatric surgery, are highly effective for weight loss. However, their therapeutic mechanisms are also not fully understood, and they come with significant side effects, such as the loss of metabolically important skeletal muscle mass, which can negatively impact long-term outcomes [3]. We utilize stable isotope-based MFA for quantifying the in vivo rates of metabolic pathways and elucidate the core metabolic changes that drive both the efficacy and the side effects of obesity treatments to identifying opportunities for new therapeutic strategies [4].

II. Experiment Design & Hypothesize Result

We established a diet-induced obesity (DIO) mouse model. Mice were subjected to one of four interventions: (1) Sleeve Gastrectomy (SG), (2) GLP-1RA treatment (Semaglutide), (3) Sham surgery with pair-feeding to the SG group (Sham-PF) to control for caloric restriction effects, and (4) Sham surgery with ad libitum feeding (Control). Following a 4-week intervention period, we performed in vivo metabolic labeling via continuous infusion of [U-^{13}C]-Glucose. Key metabolic organs (liver, kidney, heart, skeletal muscle, adipose tissue) and plasma were harvested. Mass spectrometry was used to measure isotope enrichment in central metabolites, and the data was integrated into a computational model to calculate in vivo metabolic fluxes [5].

Our hypothesis is that MFA will uncover intervention specific metabolic rewiring and anticipate identifying distinct flux signatures for GLP-1RA and SG across multiple organs. These flux-level insights will guide further mechanistic studies and new therapeutics development [6].

III. Conclusion

Based on this study, we will understand how systemic metabolism is remodeled during the development of obesity and in response to premier treatments. By moving beyond static measurements to quantify in vivo pathway activities, our work will provide clear mechanistic targets for intervention, with significant implications for mitigating side effects especially in muscle loss. Ultimately, our research will provide a systems-level view of obesity therapies, creating a rational basis for developing novel strategies aimed at improving efficacy of weight management.

[1]Chinese Institutes for Medical Research, Beijing, China. E-mail: zhangjianyun@cimrbj.ac.cn
[2]Chinese Institutes for Medical Research, Beijing, China. E-mail: zhouchong@cimrbj.ac.cn
[3]Chinese Institutes for Medical Research, Beijing, China. E-mail: shiyu.liu@cimrbj.ac.cn

References

[1] Kopelman, P. (2000) Obesity as a medical problem. *Nature* 404, 635–643.
[2] Liu, H., Wang, S., Wang, J. et al. (2025) Energy metabolism in health and diseases. *Sig Transduct Target Ther* 10, 69.
[3] Sargeant JA, Henson J, King JA, et al. (2019) A review of the effects of GLP-1 receptor agonists and SGLT-2 inhibitors on lean body mass in humans. *Endocrinology and Metabolism* 34:247-262.
[4] Kim, IY., Park, S., Kim, Y. et al. (2022) Tracing metabolic flux in vivo: basic model structures of tracer methodology. *Exp Mol Med* 54, 1311–1322.
[5] Liu, S., Liu, X. & Locasale, J.W. (2024) Quantitation of metabolic activity from isotope tracing data using automated methodology. *Nat Metab* 6, 2207–2209.
[6] Linge J, Birkenfeld AL, Neeland IJ. (2024) Muscle Mass and Glucagon-Like Peptide-1 Receptor Agonists: Adaptive or Maladaptive Response to Weight Loss? *Circulation* 150(16):1288-1298.

Modeling Diverse Modes of Cholesterol Regulation in Human Tissues

Jinyi Liu [1] , Jinling Cao [2] , Jinxuan Wen [3] , Pei Jiang [2] , Xiaojing Yang[3] , Fuchu He [2] , Chao Tang [3]

***Short Abstract* — We investigated the topological properties of the cellular cholesterol regulatory network using Boolean and RACIPE modeling. Both approaches reveal the existence of stable states. By analyzing transcriptome data from the HPA database, we inferred SREBP2 and LXR activity and performed clustering with key cholesterol-related genes. The results suggest three distinct regulatory modes of cholesterol homeostasis, consistent with RACIPE-derived steady states.**

***Keywords* — Cholesterol homeostasis, Boolean modeling, RACIPE, regulatory network, transcriptome**

I. Purpose

CHOLESTEROL is an essential lipid that plays structural and regulatory roles in eukaryotic cells. To maintain intracellular cholesterol within a narrow range, cells rely on a complex regulatory network involving synthesis, uptake, efflux, esterification, and sensing mechanisms [1]. While the biochemical components of this network have been extensively studied, the system-level logic that ensures cholesterol homeostasis remains poorly understood.

We hypothesize that the topology of the cholesterol regulatory network inherently supports robust homeostasis. To test this, we analyzed its dynamical properties using Boolean modeling and RACIPE [2]. In parallel, we used transcriptomic data from the Human Protein Atlas [3] to infer transcription factor activity and compare cholesterol-related gene expression across human organs, aiming to identify distinct homeostatic strategies in vivo.

Together, our modeling and data-driven approaches aim to systematically identify and characterize the distinct regulatory strategies that cells employ to maintain cholesterol homeostasis.

II. Preparation of Abstracts

To explore whether the structure of the cholesterol regulatory network itself supports stable regulatory behavior, we performed two types of simulations: Boolean logic modeling and RACIPE-based ODE modeling. These complementary approaches provide both qualitative and quantitative insights into the intrinsic dynamics of the network.

A. Boolean modeling reveals a global cyclic attractor

We built a Boolean model of the cholesterol network and found a single global cyclic attractor across all initial states. In this attractor, the cholesterol node oscillates between 0 and 1, stabilizing at ~0.5 when mapped to ternary states. This suggests the network topology alone supports robust, self-sustained regulation.

B. RACIPE modeling identifies multiple regulatory modes

To evaluate how the network behaves under diverse kinetic conditions, we employed the RACIPE framework, which converts the regulatory network into a system of ODEs and samples parameters from biologically plausible ranges. For each parameter set, we computed the steady states of the system. Most parameter sets yielded a unique steady state, and clustering analysis of these single steady states revealed three dominant phenotypes. Each phenotype represents a distinct regulatory strategy: one with high synthetic activity, one with high efflux, and one combining both. This diversity implies the network structure enables flexible homeostatic solutions.

C. Transcriptome analysis supports model predictions

Using RNA-seq data from the Human Protein Atlas (HPA), we inferred the activity of SREBP2 and LXR from target gene expression. Joint clustering with other cholesterol-related genes (e.g., HMGCR, ABCA1) revealed three organ groups: synthesis-dominant, efflux-dominant, and mixed. This pattern mirrors the phenotypes predicted by RACIPE simulations.

III. Conclusion

Our results show that network topology enables multiple stable regulatory strategies. The agreement between modeling and transcriptome data supports diverse, structure-driven modes of cholesterol homeostasis.

Acknowledgements: This work was funded by National Natural Science Foundation of China (NSFC 32088101and NSFC 12090053).

[1] Peking-Tsinghua Center for Life Sciences, Peking University. E-mail: liujinyi@stu.pku.edu.cn

[2] State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Beijing 102206, China

[3] Peking-Tsinghua Center for Life Sciences and Center for Quantitative Biology, Peking University, Beijing 100871, China

References

[1] Luo, J., Yang, H. & Song, B. L. (2020) Mechanisms and regulation of cholesterol homeostasis. Nat Rev Mol Cell Biol 21, 225-245.

[2] Huang, B., Jia, D., Feng, J. et al. (2018) RACIPE: a computational tool for modeling gene regulatory circuits using randomization. BMC Syst Biol 12, 74.

[3] Mathias Uhlén et al. (2015) ,Tissue-based map of the human proteome.Science347,1260419.

Dynamic Condensate Formation Enhances Circadian Clock Robustness

Renyu Wang[1], Yi Lin[2], Yuansheng Cao[3], and Chao Tang[4]

***Short Abstract* —Circadian rhythms rely on transcription-translation feedback loops (TTFLs) to generate robust 24-hour oscillations. Phase-separated condensates have recently emerged as important regulators of cellular biochemistry, but their role in circadian timing remains unclear. Using mathematical modeling, we show that static condensates, while enhancing translation, suppress transcription and limit clock coherence. In contrast, dynamically oscillating condensates can amplify protein levels while preserving mRNA expression, thereby maximizing clock robustness. This requires condensates to oscillate autonomously while remaining phase-aligned with the TTFL. Our results reveal a design principle for circadian control: dynamic condensates enhance rhythmic precision through self-sustained, clock-coupled regulation.**

***Keywords* — phase separation, dynamic condensates, translation, circadian clock, oscillating system, amplitude enhancement, robustness.**

I. Introduction

In mammals, circadian rhythms are generated by a network of core genes that are regulated by positive and negative transcription-translation feedback loops (TTFLs) [1]. A key component is the negative feedback between BMAL1, CLOCK, PER1/2/3, and CRY1/2. PER and CRY proteins form heterodimers that, after phosphorylation, translocate into the nucleus to inhibit their own transcription by repressing the CLOCK-BMAL1 complex. This transcriptional and post-translational cascade drives ~24-hour rhythmicity [2].

Recent work suggests a new player in this system: phase-separated condensates [3]. Condensates formed by the RNA-binding proteins ATXN2 and ATXN2L have been shown to enhance translation of key circadian factors. These membraneless subcompartments concentrate biomolecules such as ribosomes and translation factors, organizing biochemical reactions in space and time. Loss of these condensates severely compromises clock robustness. More broadly, phase-separated compartments are known to participate in a range of biological functions. Yet how condensates contribute to circadian stability—and what principles govern their effect—remains unclear.

[1]Peking-Tsinghua Center for Life Science, Peking University, Beijing 100871, China. E-mail: wryu@stu.pku.edu.cn

[2]State Key Laboratory of Membrane Biology, IDG/McGovern Institute for Brain Research, Tsinghua-Peking Joint Centre for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China. E-mail: linyi@mail.tsinghua.edu.cn

[3]Department of Physics, Tsinghua University, Beijing 100084, China. E-mail: yscao@tsinghua.edu.cn

[4]School of Physics, Center for Quantitative Biology, Peking University, Beijing 100871, China. E-mail: tangc@pku.edu.cn

II. Results

Here, we combine mathematical modeling with experimental perturbations to uncover design principles by which dynamic condensates support circadian robustness. We first show that static condensates, which constantly enhance translation, suppress transcription due to negative feedback, ultimately reducing clock coherence. This behavior is inconsistent with experimental observations, suggesting that static condensates alone are insufficient to sustain robust rhythms. In contrast, dynamically oscillating condensates —when properly phase-aligned—can enhance protein amplitude while leaving mRNA levels largely unaffected.

Our model predicts that optimal coherence is achieved when condensates exhibit autonomous oscillations while being modulated by the TTFL to fine-tune their phase. This prediction is supported by knock-out experiments showing that ATXN2 retains oscillations even when the core TTFL is disrupted. Together, these results suggest that dynamic, clock-coupled condensates are essential for maintaining stable circadian rhythms. Our study highlights how phase-separated condensates contribute to the precision of the circadian system, offering a new perspective on the regulation of biological timekeeping.

III. Conclusion

Phase-separated condensates are increasingly recognized across diverse biological processes. While many studies have focused on their equilibrium and static properties, growing evidence points to a rich spectrum of dynamic and nonequilibrium behaviors [4]. Our results highlight the importance of this dynamic behavior. Understanding how nonequilibrium condensate dynamics contribute to temporal regulation—and how cells tune them to optimize function—remains an exciting frontier in cell biology.

References

[1] Koike, N. et al. Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. *Science* 338, 349–354 (2012).

[2] Yao, X., Heidebrecht, B. L., Chen, J. & Tyson, J. J. Mathematical analysis of robustness of oscillations in models of the mammalian circadian clock. PLOS Comput. Biol. 18, e1008340 (2022).

[3] Zhuang, Y. et al. Circadian clocks are modulated by compartmentalized oscillating translation. Cell 186, 3245–3260 (2023).

[4] Heltberg, M. S. et al. Enhanced DNA repair through droplet formation and p53 oscillations. Cell 185, 4394-4408.e10 (2022).

Construction and Analysis of Genome-scale Metabolic Network Model for *Cryptococcus neoformans*

Chun Feng[1], Pengjie Hu[2], Xindi Gao[3], Bing Zhai[4]*, Ziwei Dai[1]*

***Short Abstract* — *Cryptococcus neoformans* is a life-threatening fungal pathogen, particularly in immunocompromised individuals. To characterize its metabolic potential and identify antifungal targets, we reconstructed and validated a genome-scale metabolic model (GEM) for strain H99, named iCNG99. The model accurately captured nutrient utilization and simulated the essentiality of known drug target. Integration with transcriptomic and metabolomic data enabled iCNG99 to capture condition-specific metabolic adaptations, including resistance-associated drug targets, and lipid metabolic reprogramming. This work provides a systems-level platform for investigating *C. neoformans* metabolism and supporting antifungal development.**

***Keywords* — genome-scale metabolic model, *Cryptococcus neoformans*, drug targets, condition-specific metabolic adaptations**

I. Background

Cryptococcus neoformans is a globally distributed opportunistic fungal pathogen that causes life-threatening infections, most notably cryptococcal meningitis[1]. It poses a particular threat to immunocompromised individuals, such as those with HIV/AIDS, and is responsible for over 100000 deaths annually[2]. Its remarkable environmental resilience and ability to colonize various host tissues make it a formidable clinical challenge.

The pathogenicity of *C. neoformans* is tightly linked to its metabolic capabilities, which underlie essential phenotypes such as proliferation, virulence, and antifungal resistance[3]. However, the absence of a high-quality, publicly available genome-scale metabolic model (GEM) has hindered systematic investigations of its metabolism and limited progress in identifying novel therapeutic targets. In this study, we reconstructed a robust GEM for *C. neoformans* var. *grubii* H99, named iCNG99, and used it to explore condition-specific metabolic shifts and to predict antifungal drug targets by integrating omics data with constraint-based modeling approaches.

[1] Department of Biology, School of Life Sciences, Southern University of Science and Technology. E-mail: 11930166@mail.sustech.edu.cn

[2] Institute of Microbiology, Chinese Academy of Sciences

[3] Department of Emergency Medicine, Xinqiao Hospital of Army Medical University

[4] CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

II. Summary of the Results

We constructed a high-quality genome-scale metabolic model of *Cryptococcus neoformans* var. *grubii* H99, named iCNG99, comprising 3278 reactions, 3432 metabolites, 1026 genes, and 7 compartments. The model achieved a MEMOTE score of 85%, outperforming most fungal GEMs. It accurately predicted nutrient utilization and known drug target essentiality, including growth inhibition upon deletion of the antifungal target Erg11.

Using this validated model, we integrated transcriptomic and metabolomic data from resistant and non-resistant isolates via a modified REMI approach. This analysis revealed several flux alterations associated with drug resistance and identified potential therapeutic targets in sterol and purine metabolism, prioritized by gene essentiality and lack of human homologs.

We further explored condition-specific reprogramming by modeling metabolic shifts across in vivo vs. in vitro and heat-tolerant vs. non-tolerant states. Lipid metabolism consistently emerged as a key adaptive module, accompanied by altered amino acid accumulation under stress.

III. Conclusion

We present iCNG99, the first well-curated genome-scale metabolic model for *Cryptococcus neoformans* var. *grubii* H99. Through multi-level validation and omics integration, the model supports accurate simulation of physiological and pathological states. Using this platform, we identified context-dependent metabolic reprogramming and potential antifungal targets under drug resistance and stress. This work enhances our understanding of *C. neoformans* metabolism and provides a reproducible framework for modeling non-model fungal pathogens.

References

[1] Rathore S, Sathiyamoorthy J, Lalitha C, et al. A holistic review on Cryptococcus neoformans[J]. Microbial Pathogenesis, 2022, 166: 105521.

[2] Rajasingham R, Smith R M, Park B J, et al. Global burden of disease of HIV associated cryptococcal meningitis: an updated analysis[J]. The Lancet Infectious Diseases, 2017, 17(8): 873-881.

[3] Zaragoza O. Basic principles of the virulence of Cryptococcus[J]. Virulence, 2019, 10(1): 490-501.

Construct the standard model of central carbon metabolism based on ^{13}C -MFA

Yonghong Hao[1], Wenchao Fan[2], and Ziwei Dai[1*]

***Short Abstract* — ^{13}C -MFA (Metabolic Flux Analysis), the gold standard for quantifying metabolic fluxes, suffers from model-dependent variability affecting reliability. We systematically evaluated this by cross-validating 168 MFA models across 8 metabolic networks and 21 datasets. This revealed substantial variability in flux estimates for most reactions when different networks were used. Critically, many biological conclusions changed with the model structure, raising concerns about reproducibility. Our study systematically deconstructs inherent MFA limitations and highlights the urgent need for standardized models to ensure robust and interpretable metabolic findings.**

***Keywords* — Meatabolic flux analysis, corss-validation, flux variability, biological conclusion divergency.**

I. Background

METABOLIC fluxomics dynamically monitors physiological phenotypes across biological scales, playing an indispensable role in metabolic engineering, disease research, and cancer metabolism [1]. It enables precise identification of metabolic pathways and quantification of flux variations, yet most intracellular fluxes must be computationally inferred from limited extracellular measurements. ^{13}C-MFA remains the gold standard due to precise isotopic data and minimal system assumptions, despite technical complexity.

Central carbon metabolism represents the most critical metabolic network across all organisms. To unravel the dynamic mechanisms underlying its complexity, researchers have developed numerous ^{13}C-MFA models for flux quantification. This approach has yielded paradigm-shifting biological insights [2,3]. However, ^{13}C-MFA results exhibit significant model dependency [4,5].

To systematically evaluate flux variability, we constructed 168 cross-validated MFA models by combining 21 experimental datasets with 8 distinct metabolic networks, enabling quantitative assessment of flux variability across architectures.

II. Results

We have developed an automated modeling platform that generates MFA models by inputting metabolic networks and target datasets. Using this platform, we constructed 168 metabolic flux analysis models, computed 168 sets of flux distributions, and rigorously validated the results through statistical analysis. Our systematic investigation has now identified two key findings.

We compiled the complete set of reactions across all metabolic networks and systematically traversed each reaction to construct a cross-model reaction flux estimate matrix. By calculating row-wise coefficients of variation (CV) for this matrix, we quantified the variability of flux estimates for each reaction across different network architectures under identical datasets. Our analysis reveals substantial variability in flux estimates for most reactions when metabolic networks are altered, demonstrating that differences in network structure significantly impact flux estimation outcomes.

To investigate whether consistent biological conclusions persist across different metabolic networks, we established a quantitative framework for evaluating biological interpretations. Our analysis reveals that most reactions fail to maintain consistent biological conclusions across networks, demonstrating that ^{13}C-MFA-derived biological insights are heavily dependent on simplified metabolic network design.

III. Conclusion

Through systematic construction of 168 MFA models, we quantified significant variability in metabolic fluxes and biological interpretations across distinct metabolic networks. This model-induced inconsistency in biological conclusions demonstrates that network selection critically affects the reproducibility of ^{13}C-MFA results.

In subsequent work, we will further analyze the robustness of metabolic networks in constructing MFA models, reduce flux uncertainty by introducing additional constraints, and ultimately attempt to establish a standardized MFA model for central carbon metabolism.

References

[1] Hui, S. et al. Quantitative fluxomics of circulating metabolites. Cell Metab 32, 676-688. e674 (2020).

[2] Metallo, C. M. et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature 481, 380-384 (2012).

[3] Bartman, C. R. et al. Slow TCA flux and ATP production in primary solid tumours but not metastases. Nature 614, 349-357 (2023).

[4] Hui, S. et al. Glucose feeds the TCA cycle via circulating lactate. Nature 551, 115-118 (2017).

[5] Liu, S. et al. Quantitative analysis of the physiological contributions of glucose to the TCA cycle. Cell Metab 32, 619-628. e621 (2020).

*Correspondance: daizw@sustech.edu.cn

[1] School of Life Science, Southern University of Science and Technology, Shenzhen.

[2]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing.

How Does *Bacillus subtilis* 'Smell' via State Transitions?

Zihan Li[1], Junfeng Lv, Jingxiang Shen , Xiaojing Yang[2]and Chao Tang[3]

***Short Abstract* — *Bacillus subtilis* employs probabilistic switching between motile and sessile states to colonize root patches optimally under fluctuating nutrient conditions. Using microfluidic assays, we measured state durations and developed a stochastic survival-game model demonstrating that simultaneous tuning of both state-switch rates markedly improves colonization efficiency versus single-parameter strategies. Model predictions closely match experimental switching patterns, revealing an adaptive bet-hedging mechanism. Our findings elucidate how dynamic two-state regulation enables robust biofilm initiation and precise root targeting in variable soil environments.**

***Keywords* — *Bacillus subtilis*, two-state model, bet-hedging**

REFERENCES

[1] Norman, T. M., et al. (2013). "Memory and modularity in cell-fate decision making." Nature **503**(7477): 481-486.
[2] Allard-Massicotte, R., et al. (2016). "Bacillus subtilis Early Colonization of Arabidopsis thaliana Roots Involves Multiple Chemotaxis Receptors." mBio **7**(6).
[3] Lord, N. D., et al. (2019). "Stochastic antagonism between two proteins governs a bacterial cell fate switch." Science **366**(6461): 116-120.
[4] Arnaouteli, S., et al. (2021). "Bacillus subtilis biofilm formation and social interactions." Nat Rev Microbiol **19**(9): 600-614.

[1] Academy for Advanced Interdisciplinary Studies, Peking University. E-mail: 1901111457@pku.edu.cn
[2] Academy for Advanced Interdisciplinary Studies, Peking University. E-mail: xiaojing_cqb@163.com
[3] Academy for Advanced Interdisciplinary Studies, Peking University. E-mail: tangc@pku.edu.cn

Proteome-wide amino-acid cost biases in enzymes and their potential relationship to human dietary profiles

Jie Wei[1], Ziwei Dai[1]

***Short Abstract* — Amino acids serve not only as protein building blocks but also as key players in metabolism and epigenetics. Their biosynthesis incurs varying metabolic costs, shaping proteome composition through evolutionary pressures. This study explores whether human metabolic enzymes preferentially use lower-cost amino acids compared to non-metabolic enzymes, and whether these patterns align with dietary amino-acid profiles. By introducing the concept of metabolic economics, we aim to reveal how cells balance resource efficiency and functional demands. Our findings may provide new insights into metabolic optimization, dietary design, and therapeutic strategies.**

***Keywords* — Metabolic economics; Proteome; Human dietary; Amino-acid cost**

I. Background

AMINO acids are the fundamental building blocks of proteins, but they also serve as metabolic precursors for numerous biomolecules and play critical roles in epigenetic modifications [1]. Different classes of amino acids have distinct roles and levels of importance in specific biological processes. Moreover, the composition and relative abundance of amino acids fluctuate dynamically across organisms and environmental conditions in nature, indicating adaptive regulation to meet changing physiological demands and living environments.

The spectrum and relative proportions of amino acids in the human diet fluctuate far more dynamically than those of fats or carbohydrates [2], making targeted amino-acid intake especially important. Deliberate modulation of specific amino-acid consumption could provide diverse and effective strategies for addressing metabolic disorders, slowing ageing, and suppressing tumour progression [3-6].

Optimizing the amino-acid profile of a diet is far from straightforward, because it requires balancing the synergistic and antagonistic interactions among all 20 amino acids. Scientifically regulating both the types and proportions of dietary amino acids could be critical for extending lifespan, preventing chronic diseases, and improving overall health [2,7].

[1]School of Life Sciences, Southern University of Science and Technology. E-mail: 12433083@mail.sustech.edu.cn

In order to maximize metabolic efficiency, organisms tend to optimize their amino acid usage, preserving protein functionality while reducing resource consumption. Selection for lower metabolic cost shapes proteome amino-acid composition broadly (though to different degrees) across many forms of life [8,9].

We introduce the concept of "metabolic economics", which views the cell from a resource-limited perspective. According to metabolic economics, a cell's metabolic circuitry aims for a dynamic, optimal balance between cost and benefit when resources are scarce. This study aims to uncover the economic logic underlying metabolic networks and to explore how organisms deftly balance metabolic costs against benefits.

II. Results

Our research compares the amino-acid composition of human metabolic enzymes and non-metabolic enzymes, examine how they relate to the amino-acid distribution typical of everyday diets. This approach promises to deepen our understanding of metabolic evolution and lay a solid foundation for future advancements in drug discovery, and innovative disease-treatment strategies.

References

[1] Chandel, N. S (2021). Amino Acid Metabolism. *Cold Spring Harb Perspect Biol* 13, a040584.

[2] Dai, Z., Zheng, W. & Locasale, J. W (2022). Amino acid variability, tradeoffs and optimality in human diet. *Nat Commun* 13, 6683 .

[3] Thandapani, P. et al (2022). Valine tRNA levels and availability regulate complex I assembly in leukaemia. *Nature* 601, 428–433.

[4] Taya, Y. et al (2016). Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. *Science* 354, 1152–1155.

[5] Knott, S. R. V. et al (2018). Asparagine bioavailability governs metastasis in a model of breast cancer. *Nature* 554, 378–381.

[6] Krall, A. S. et al (2021). Asparagine couples mitochondrial respiration to ATF4 activity and tumor growth. *Cell Metabolism* 33, 1013-1026.e6.

[7] Piper, M. D. W. et al (2017). Matching Dietary Amino Acid Balance to the In Silico Translated Exome Optimizes Growth and Reproduction without Cost to Lifespan. *Cell Metabolism* 25, 610–621.

[8] Chen, Y. & Nielsen, J (2022). Yeast has evolved to minimize protein resource cost for synthesizing amino acids. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2114622119.

[9] Zhang, H. et al (2018). Biosynthetic energy cost for amino acids decreases in cancer evolution. *Nat Commun* 9, 4124.

Toward a comprehensive understanding of *C. elegans* early embryogenesis

Wei Tian[1], Ming Han[1,2,*], and Chao Tang[1,2,3,*]

***Short Abstract* —The nematode *Caenorhabditis elegans (C. elegans)* is well-known for its invariant cell lineage and spatiotemporal pattern during embryogenesis, providing an excellent opportunity to delve deep into the design principles of development. We now have a detailed mechanical tool: a phase field model, to precisely simulate its early embryogenesis. Yet we still don't understand how gene expression regulates morphogenesis through mechanics. To bridge the gap, we use an adjoint method, which enables efficient parameter optimization for complex PDE systems, to estimate the main mechanical parameters from experimental images of wild-type and mutant embryos. Combining scRNA-seq data, we will gain a deeper understanding of the process of embryogenesis.**

***Keywords* — *Caenorhabditis elegans*, embryogenesis, phase field model, parameter optimization**

THE nematode *Caenorhabditis elegans* has long been served as a distinguished model organism for biological research, particularly in the study of early embryonic development. *C. elegans* is well-known for its invariant cell lineage and spatiotemporal pattern during embryogenesis, providing an excellent opportunity to delve deep into the design principles of development[1]. Great efforts have been made to study the detailed regulations during the early stage of embryogenesis, including single-cell RNA sequencing, live imaging of wild-type a large number of mutants, and image-based quantification of key transcription factors[2-5]. At the mechanical level, researchers have established several elegant models to describe the dynamical process of cell shape changing[6,7]. However, precisely determine the mechanical parameters in biological systems has always been a difficult task.

In this work, we employ an efficient parameter optimization method, the adjoint method, to determine the mechanical parameters in the phase field model[7], which we use to describe the dynamic process of morphogenesis during *C. elegans* early embryonic development. Combining the high-resolution in vivo 3D imaging data which have been collected, we are able to analyze the mechanical properties and regulations during this important dynamical process.

[1]Center for Quantitative Biology, Peking University, Beijing 100871, P. R. China.
[2]Peking-Tsinghua Center for Life Science, Peking University, Beijing 100871, P. R. China.
[3]School of Physics, Peking University, Beijing 100871, P. R. China.
[*] Corresponding author

REFERENCES

[1] Sulston, John E., et al. "The embryonic cell lineage of the nematode Caenorhabditis elegans." *Developmental biology* 100.1 (1983): 64-119.
[2] Cao, Jianfeng, et al. "Establishment of a morphological atlas of the Caenorhabditis elegans embryo using deep-learning-based 4D segmentation." *Nature communications* 11.1 (2020): 6254.
[3] Ma, Xuehua et al. "A 4D single-cell protein atlas of transcription factors delineates spatiotemporal patterning during embryogenesis." *Nature methods* vol. 18,8 (2021): 893-902.
[4] Cole, Alison G et al. "Gene regulatory patterning codes in early cell fate specification of the *C. elegans* embryo." *eLife* vol. 12 RP87099 (2024).
[5] Green, Rebecca A et al. "Automated profiling of gene function during embryonic development." *Cell* vol. 187,12 (2024): 3141-3160.e23.
[6] Berry, Joel er al. "Physical principles of intracellular organization via active and passive phase transitions." *Reports on progress in physics* vol. 81,4 (2018): 046601.
[7] Kuang, Xiangyu, et al. "Computable early *Caenorhabditis elegans* embryo with a phase field model." *PLoS Computational Biology* 18.1 (2022): e1009755.

The interplay between ecology and evolution of gut microbiomes

Kaling Cheung[1] and Po-Yi Ho[2]

Short Abstract —Human gut microbiomes are ecologically stable over decades, a timescale over which dozens of mutations can fix in a species. How evolution affects inter-species or inter-strain interactions and how the community remains ecologically stable remain unclear. We experimentally evolved 82 replicates of a 9-species in vitro microbiome for ~500 generations. Communities diverged into 4 types with distinct coexisting species. Within-species de novo mutations led to divergent monoculture growth. Evolved strains exhibited variable growth on the spent-medium of ancestral and evolved strains from other species. These interaction changes increased cross-feeding strength in most cases and produced emerged cross-feeding in some cases.

I. Introduction

HUMAN gut microbiomes are ecologically stable over decades, a timescale over which dozens of mutations can fix in a species [1]. How evolution affects inter-species or inter-strain interactions and how the community remains ecologically stable despite these changes remain unclear. Here, we experimentally evolved 82 replicates of a 9-species in vitro gut microbiome via serial dilution for ~500 generations and found that they diverged into 4 types of communities with different sets of coexisting species, suggesting the presence of strong selection guiding evolutionary outcomes. Regardless of community type, multiple strains with different de novo mutations emerged within species. These strains grew to substantially different amounts in monoculture, indicating that evolution rapidly led to phenotypic changes. Moreover, evolved strains also exhibited variable growth on the spent media of ancestral and evolved strains from other species. These interaction changes led to increased cross-feeding strength in most cases, and in newly emerged cross-feeding interactions in some cases. We hypothesize that evolution tends to strengthen positive interactions and will continue to combine genomics, high-throughput microbiology experiments, and mathematical modeling to uncover the underlying eco-evolutionary principles of complex microbial communities.

II. Results

A. Community compositions after experimental evolution diverged to 4 typese

Using a synthetic gut community composed of 9 representative and metabolically diverse species from different genera commonly found in the human gut, we evolved 82 independent replicate communities under constant environmental conditions for approximately 500 generations. The evolved communities diverged into 4 distinct compositions, suggesting the presence of strong selection guiding evolutionary outcomes.

B. Monoculture growth dynamics and metabolic interactions changed after evolution

By comparing the growth dynamics of representative evolved strains to those of ancestral strains in fresh BHI medium, and through pairwise spent-medium experiments using both ancestral and evolved isolates, we found that monoculture growth dynamics and metabolic interactions changed after evolution. Furthermore, a specific cross-feeding interaction from *Cs* to *Ba* evolved *de novo*.

C. Cross-feeding interactions are ecologically important and evolutionarily stable in communities

In over 93% of ancestral 9-species communities, the dominant species *Ai* showed significantly higher abundance than in monoculture, suggesting the presence of cross-feeding from other community members. This interaction was maintained in the evolved communities with *Ai*.

D. Cross-feeding interactions are prealent and both stable and evoluable in pairwise conbinations

Through quantifying pairwise cross-feeding interactions by coarse-graining resources, cross-feeding interactions were found to be widespread among ancestors. Moreover, different evolved *Cs* strains consistently promoted the growth of *Ba* through newly evolved cross-feeding across different communities. Mutual cross-feeding between *Ai* and *Ba* also intensified over evolution.

III. Conclusion

Our findings show that large-scale experimental evolution drives divergence in community composition and reshapes both monoculture dynamics and metabolic interactions, which can be quantified using a coarse-grained resources framework. Cross-feeding interactions not only maintain evolutionary stability but also can emerge *de novo.*

References

[1] Zhao et al. (2019) Adaptive Evolution within Gut Microbiomes of Healthy People. *Cell Host & Microbe* **25**, 656–667.

[1]School of Life Sciences, Westlake University, Hangzhou, China. E-mail: zhangjialing@westlake.edu.cn

[2]Center for Interdisciplinary Science (CIS), Westlake University, Hangzhou, China. E-mail: poyiho@westlake.edu.cn

Discovering *C. elegans* Aging Phenotypic Signatures Using NemaNet

Yan Pan[1,†] , Dan Wu[1,†] , Hongxia Cai[1], Li Fan[1] and Bo Xian[1,*]

***Short Abstract* — Assessing whether organisms exhibit premature aging is critical for anti-aging research and drug discovery. We developed a high-precision anatomical structure recognition model for *C. elegans* and established a standardized age-image dataset. By segmenting and reconstructing nematode bodies based on their anatomical segments, we eliminated posture-induced variability. The extracted features enabled an image-to-biological-age regression model, which further served as a framework to quantify premature aging phenotypes. Our analysis revealed distinct correlations between multitask phenotypic signatures and biological aging.**

***Keywords* — Aging, *C. elegans*, CNN**

I. Purpose

Aging is a complex biological process with profound implications for health and disease. The ability to accurately assess premature aging (progeria) is crucial for understanding its mechanisms and developing interventions. The nematode *Caenorhabditis elegans* (*C. elegans*) serves as a premier model organism in aging research due to its short lifespan and conserved aging pathways. However, current methods for *C. elegans* anatomical recognition often lack precision and offer only rudimentary functionality. To address these limitations, we 1)developed a segmentation and tracking model for *C. elegans* anatomical structures; 2) constructed a standardized age-image dataset to enable biological age prediction.

II. Methods

Under synchronized standard cultivation conditions, we collected a dataset comprising >30,000 annotated *C. elegans* images. Each specimen was meticulously labeled for head, tail, pharyngeal pumping, and egg presence, with chronological cultivation age serving as ground-truth biological age. For age regression modeling, we implemented a three-stage computational pipeline[1]: 1) Whole-worm segmentation, 2) Anatomical substructure partitioning based on body segments, and 3) Parallel CNN training using segment-derived image frames as distinct feature channels.

[†]Authors contributed equally to this work.
[1]Laboratory of Aging Research, School of Medicine, University of Electronic Science and Technology of China, Chengdu 611731, China. E-mail: yanpan@zohomail.com (Y.P.)
[*]Correspondence: xianbo@uestc.edu.cn (B.X.);

III. Results

The NemaNet model achieved high accuracy in predicting *C. elegans* biological age ($R^2 = 0.85$). To validate its progeria identification capability, we performed single-worm quantitative RT-PCR (qt-qPCR) and lipofuscin quantification on synchronized, size-matched worms classified by NemaNet as progeric, normal, or delayed-aging. Progeric worms exhibited a 20% upregulation of *daf-2* expression and a ~50% increase in lipofuscin accumulation compared to controls, confirming that NemaNet surpasses human visual assessment in detecting aging states and aligns with established molecular biomarkers.

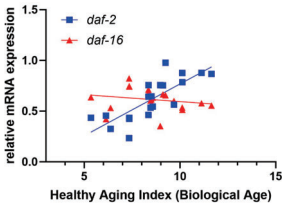


Fig. 1. Regression analysis between NemaNet-predicted biological age and single-worm qRT-PCR measurements in *C. elegans*.

While lipofuscin has been conventionally regarded as a passive biomarker of cumulative cellular damage in aging organisms[2], our findings provocatively suggest its potential role as an early indicator of progeria. The 50% elevation in lipofuscin in NemaNet-classified progeric worms, prior to detectable morphological aging, implies its accumulation may reflect specific dysregulation of autophagy-lysosomal pathways (ALP) rather than generalized aging.

Furthermore, we analyzed age-related motility patterns by computing real-time head/tail acceleration. Even after removing temporal information, trajectory-derived morphological features (e.g., head/tail movement sequences) retained significant aging signatures, suggesting that locomotion dynamics encode latent biological age information.

References

[1] Yan Pan; Zhihang Huang; et. al. *International Journal of Molecular Sciences* 2024, 25, 9675 .
[2] Yan Pan; Hongxia Cai; Fang Ye; et. al. *Aging Cell* 2025, e70116 .

12 Major-Term Classification System for Cellular Processes

Jiaming Hu[1], Yirui Liu[1], Jingfang Zhang[1]

Abstract—The complexity of Gene Ontology (GO) poses challenges for cellular process comprehension. Here, we developed a classification strategy by analyzing shared genes among terms, reducing cellular process terms to 12 major-terms. The 54 GO Slim terms can be embedded within these 12 major-terms, demonstrating their effectiveness in interpretating core cellular process. Term-associated proteins interaction networks were further analyzed and revealed small-world topology and scale-free properties. This work provides a simplified GO annotation system and demonstrats its complex network properties.

Keywords—Gene Ontology, STRING, complex network

THE Gene Ontology (GO) is a robust knowledge base that provides precisely defined ontology terms and associated annotations for a wide range of organisms. The GO is organized into three main categories: molecular function (MF), cellular component (CC), and biological process (BP). To reduce redundancy, the GO Consortium recently introduced a manually trimmed database known as 'Generic GO Slim', containing only 75 BP terms, 40 MF terms, and 29 CC terms [1]. However, no systematic attempt has been made to cluster GO terms into broader categories that would provide a comprehensive overview of cellular processes.

I. Generation of 12 major-terms of cellular process

Firstly, we obtained all 66 direct child terms of the cellular process (CP) branch from the AmiGO database. To analyze the relationships of these child terms, we constructed a shared gene network by calculating the overlapping genes between any term pairs. Notably, many small-scale terms were juxtaposed with large-scale terms, which sacrifice the general view of the whole cellular process. To reduce the complex structure of CP, three categorization strategies including exclusion, coverage and combination were applied manually. Finally, a simplified set of only 12 major-terms was obtained, which encompass most of the genes in CP, are relatively independent but allow interconnections.

Then, to test the interpretation ability of the 12 major-terms, the 75 BP terms from 'Generic GO Slim' were further analyzed. Among them, we mainly focused on 54 terms that are specifically classified under CP. By quantifying gene overlap between the 54 CP terms and the 12 major-terms, we found that the 54 CP terms were successfully embedded within the 12 major-terms, belonging to one or more of the 12 major-terms. This demonstrates their effectiveness in representing core cellular processes.

II. Structural Properties of Term-Associated Protein Interaction Networks

To further investigate the complex network of each term or the whole CP at the protein-protein interaction (PPI) level, we utilized the STRING database and Cytoscape software to generate term-associated protein interaction networks. Network properties including degree distribution, clustering coefficient, and network diameter were analyzed. Network topology analysis revealed three fundamental architectural features: First, the gene degree distribution followed a power-law pattern ($P(k) \sim k^{-\gamma}$), indicating scale-free properties; Second, the networks exhibited extreme sparsity, with connections representing only a tiny fraction of possible edges; Third, small-world characteristics were observed through significantly elevated clustering coefficients and reduced network diameters compared to random networks.

III. Conclusion

This study developed a systematic classification approach for GO cellular process terms, streamlining 66 direct child terms into 12 representative major-terms. The 54 GO Slim terms were successfully embedded within these core groups. PPI network analysis further revealed both small-world characteristics and scale-free properties in these biological networks.

References

[1] Aleksander, S.A., Balhoff, J., Carbon, S., Cherry, J.M., Drabkin, H.J., Ebert, D., et al. (2023). The GeneOntology knowledgebase in 2023. Genetics 224(1). doi: 10.1093/genetics/iyad031.

[1] School of Life Sciences, Beijing University of Chinese Medicine.

E-mail: zhangjingfang@bucm.edu.cn

Is the active suspension in a complex viscoelastic fluid more chaotic or more ordered?

Yuan Zhou[1], Ignacio Pagonabarraga[2], and Kai Qi[1]

***Short Abstract* — The habitat of microorganisms is typically complex and viscoelastic. A natural question arises: do polymers in a complex swimmer suspension increase chaos or promote order? To address this, we investigate squirmer suspensions in polymer solutions using lattice Boltzmann simulations. Our results show that at an intermediate swimmer volume fraction, polymers can enhance squirmer polarization by up to 26 times for neutral swimmers and 5 times for pullers, thereby increasing the system's order. This counterintuitive phenomenon is attributed to a hydrodynamic feedback mechanism arising from the interplay between swimmer activity and polymers. Our findings establish a framework for understanding collective motion in complex fluids and suggest strategies for controlling active systems through polymer-mediated hydrodynamic interactions.**

***Keywords* — Squirmer suspension, polymer, hydrodymic simulation, polarization order, hydrodynamic feedback.**

I. INTRUDUCTION

MICROSWIMMER locomotion in viscoelastic fluids is fundamental to processes like fertilization, infection, and biofilm formation. Complex fluids significantly alter swimming kinematics and transport, depending on swimmer type and fluid rheology. For example, *E. coli* swims faster in polymer solutions due to flagellar-induced polymer stretching that suppresses unbundling [1], while *C. reinhardtii* experiences restricted flagellar motion, reducing speed [2]. Theoretical and computational studies show that bacterial speeds can increase in polymeric fluids due to non-uniform polymer distributions and flagellar chirality [3]. Experiments show self-propelled Janus particles have enhanced rotation and transition to persistent circular motion above a critical Deborah number [4], explained by polymer adsorption and asymmetric encounters [5]. Beyond individual swimmers, viscoelasticity crucially impacts collective behavior by destabilizing vortex structures in pushers but not pullers. Rod-like pushers in films show viscoelasticity enhances aggregation via polymer stresses [6]. However, how polymer deformations modulate collective squirmer dynamics and influence the system's order remains underexplored.

[1]2020 X-Lab, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China. E-mail: kqi@mail.sim.ac.cn

[2]Departament de Física de la Matèria Condensada, Universitat de Barcelona, C. Martí Franquès 1, 08028 Barcelona, Spain. E-mail: ipagonabarraga@ub.edu

II. RESULTS

In this work, we perform hydrodynamic simulations of multiple squirmers in polymer solutions using the lattice Boltzmann method to investigate their clustering behavior. We quantify collective orientation through polarization and find that at a sufficiently high polymer concentration (φ_p = 0.16), polarization remains high and stable over a broad range of squirmer concentrations. Specifically, at squirmer concentration φ_s = 0.15, polarization increases by a factor of 26 for neutral swimmers and 5 for pullers compared to the polymer-free system. To elucidate the mechanism behind this enhancement, we analyze the cluster-size distribution and polymer stretching characteristics, including the radius of gyration and the anisotropy parameter. Our results indicate a hydrodynamic feedback mechanism in which squirmers stretch the polymers, causing their deformation and alignment with the swimming direction. This in turn reinforces swimmer orientation and amplifies polarization. To further quantify this effect, we introduce an alignment parameter that provides direct evidence of the interplay between squirmers and polymer deformation. This conclusion is supported by the coherent variation between the polarization vector and the principal eigenvector of the gyration tensor.

III. CONCLUSION

Our work illustrates that introducing polymers into the swimmer suspension does not necessarily increase the system's complexity. On the contrary, the hydrodynamic feedback mechanism arising from the interplay between swimmer activity and polymers can enhance the polarity of the suspension, thereby increasing its overall order. These findings provide new insights into how polymers can be harnessed to control the structure and dynamics of active matter systems.

REFERENCES

[1] Patteson A, et al. (2015) Running and tumbling with e. coli in polymeric solutions, *Sci. Rep.* **5**, 15761.

[2] Qin B, et al. (2015) Flagellar kinematics and swimming of algal cells in viscoelastic fluids, *Sci. Rep.* **5**, 9190.

[3] Zötol A, et al. (2019) Enhanced bacterial swim- ming speeds in macromolecular polymer solutions, *Nat. Phys.* **15**, 554.

[4] Gomez-Solano JR, et al. (2016) Dynamics of self-propelled janus particles in viscoelastic fluids, *Phys. Rev. Lett.* **116**, 138301.

[5] Qi K, E. et al. (2020) Enhanced rotational motion of spherical squirmer in polymer solutions, *Phys. Rev. Lett.* **124**, 068001.

[6] Li G, et al. (2016) Collective motion of microorganisms in a viscoelastic fluid, *Phys. Rev. Lett.* **117**, 118001.

Restricting cystine intake inhibits colorectal cancer

Lei Zhang[1], Kang Xia[1], Weijie Li[1], Mei Xiao[1],Xiaojing Yang[1,2], Yiping Wang[3], Chao Tang[1,2]

***Short Abstract* — Cysteine serves as the primary intracellular antioxidant provider. Depletion of cysteine triggers ferroptosis in cancer cells, rendering tumors more susceptible to diverse therapeutic interventions. In this study, we demonstrate that cysteine depletion amplifies the efficacy of immunotherapies in colorectal cancer. Employing a combination of quantitative metabolomics and single-cell transcriptomics, we observed that cysteine depletion rewires the metabolism of cancer cells and modifies the nutrient composition of the tumor microenvironment. This alteration impedes the activation and function of immune-suppressive regulatory T (Treg) cells, concurrently promoting the infiltration of cytotoxic CD8+ T cells.**

***Keywords* —Colorectal cancer, Cysteine metabolism, Immune microenvironmen.**

I. INTRUDUCTION

Colorectal cancer is the third most prevalent and second most deadly cancer worldwide. The development, metastasis and treatment of colorectal cancer are highly related to diet and intestinal nutrition. We found that colorectal cancer tumor immunity can be activated by restricting dietary cysteine. On this basis, we developed monoclonal antibodies targeting cysteine transport complexes and engineered bacteria targeting intestinal cysteine as novel therapeutic approaches for colorectal cancer.

II. PESULTS

Colorectal cancer is the third most prevalent and second most deadly cancer worldwide. The development, metastasis and treatment of colorectal cancer are highly related to diet and intestinal nutrition. We found that colorectal cancer tumor immunity can be activated by restricting dietary cysteine. re. On this basis, we developed monoclonal antibodies targeting cysteine transport complexes and engineered bacteria targeting intestinal cysteine as novel therapeutic approaches for colorectal cancer.

In vitro, we found that cysteine restriction did not affect the growth of intestinal cancer cells SW480, but altered their metabolism. The apparent Cysteine restriction significantly reduced the uptake of arginine and excretion of glutamate by SW480. On the other hand, cysteine restriction had no significant effect on the proliferation and activation of tumor-killing CD8 T cells, but inhibited the differentiation of CD4 T cells to Treg cells. In vivo, we found that cysteine restriction increased the arginine content and decreased the glutamate content in the tumor microenvironment, which may help to shape the environment favorable for immune activation. At the same time, we found that cysteine restriction increased the number of tumor-infiltrating CD8 T cells, B cells, and NK cells and decreased the number of Treg cells and macrophages. Monoclonal antibodies targeting the cysteine transport complex inhibited tumor development well. Engineered bacteria targeting intestinal cysteine significantly reduced cysteine levels in the gut and inhibited colorectal cancer tumor development.

III. CONCLUSION

Cysteine restriction activates tumor immunity through multiple pathways, and both the monoclonal antibody we developed to target the cysteine transport complex and the engineered bacteria targeting intestinal cysteine could have a therapeutic effect on tumors by perturbing cysteine metabolis.

[1]Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies, Chengdu, China

[2]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

How Isozymes Emerge: Mechanistic Insights from Metabolic Network Evolution

Xiangyu Hou[1], Ziwei Dai[1]

Isozymes are distinct enzymes that catalyze the same chemical reaction, often sharing a similar catalytic core while differing in sequence. The presence of isozymes implies a degree of functional redundancy. To investigate their evolutionary origin, we analyzed 5,587 genome-scale metabolic models (GEMs) constructed using the semi automated CarveMe pipeline. We found that certain reactions were catalyzed by multiple isozymes in some models, whereas the same reactions tended to be catalyzed by a single enzyme in others. Only a small subset of isozymes displayed high conservation across all models, which was significantly different from expectations under random simulation. This observation prompted us to explore the underlying mechanisms of isozyme emergence from the perspective of metabolic network evolution.

Our analyses revealed that isozymes tend to occur in reactions involving larger metabolites. Moreover, the sequences of isozyme-encoding genes are significantly longer than those of non-isozyme counterparts. These findings suggest that the emergence of isozymes may be an inevitable outcome of molecular evolution driven by genetic mutations: when the size of the catalytic core remains constant, longer sequences are more prone to mutation, thus increasing the likelihood of isozyme formation. A crucial assumption here is the frequent occurrence of gene duplication events in prokaryotes; however, previous work by Csaba Pál indicates that horizontal gene transfer, rather than duplication, is the primary driver of metabolic network evolution in prokaryotes. Therefore, the mechanistic origins of isozymes in metabolic network evolution warrant further investigation.

In addition to their mutational origin, isozyme distribution appears to be non-random. For example, studies by Simran Kaur Aulakh and colleagues have shown that isozymes are more likely to bind metal ions, and that approximately half of the proteome and most major signaling pathways are responsive to changes in metal availability. We further applied concepts from control theory to metabolic networks and found that isozymes are more likely to serve as control points. Collectively, these lines of evidence suggest that the seemingly redundant presence of isozymes in biological systems may follow an evolutionary directionality and be closely linked to the regulatory architecture of metabolic networks.

[1]School of Life Sciences, Southern University of Science and Technology, ShenZhen, China

Active learning with genomic language models accelerates genetic circuit design

Zelun Li[1], Zitong Jerry Wang[1]

Short Abstract — Genetic circuit design faces significant challenges due to vast combinatorial search spaces and time-intensive design-build-test-learn (DBTL) cycles, particularly for *in vivo* studies. We present DIAL (Deep learning informed active learning), a novel sequence selection methodology combining gLM embeddings with machine learning regressors for efficient genetic circuit optimisation. Using a dataset of 121k genetic circuit sequences, DIAL achieved a three-fold reduction in required samples compared to baseline strategies. Nonlinear regressors (random forest, K-nearest neighbors) demonstrated two-fold higher fold changes than linear models. DIAL offers substantial promise for accelerating synthetic biology applications across diverse research contexts.

***Keywords* — genetic circuits, active learning, genomic language model**

I. Introduction

Synthetic gene circuits enable cells to perform novel sense-and-response behaviours, but optimising circuit performance remains challenging. Tuning circuit behaviour usually involves changing multiple gene regulatory components, such as promoters, terminators, and ribosomal binding sites, simultaneously. Each component can have tens to hundreds of variants, resulting in a large combinatorial search space with millions of possible combinations [1], making the identification of near-optimal designs exceptionally difficult.

Traditionally, multiple rounds of design-build-test-learn (DBTL) are carried out in exploring the large combinatorial search space before reaching an optimal output that is required of the particular cellular context. However, testing genetic circuits *in vivo* can quickly become labour-intensive, since genetic circuits often behave unpredictably when introduced into living cells [2]. Computational models try to address the issue of long DBTL cycle by providing *in silico* approximations. While recent deep learning models have made notable improvements in sequence-to-expression predictions, they still fall short at synthetic sequence predictions absent from the training data [3].

The advancement of gLMs with robust few-shot prediction abilities creates new opportunities. gLMs trained on large-scale DNA sequence data are capable of encoding sequence data into representations that enhance learnability for predictive tasks [4]. By including the sequence representations extracted from gLMs, the learnability of these sequences can be increased effectively. This is important since a well-fitted selection strategy for picking genetic circuits to test would significantly reduce the number of experiments needed, thereby lowering time and costs.

[1]Westlake University, Center for Interdisciplinary Studies, Hangzhou China. E-mail: lizelun@westlake.edu.cn wangzitong@westlake.edu.cn

II. Method

We evaluated selection strategies using a dataset of 121k genetic circuit sequences with corresponding fold change measurements [5]. We propose DIAL (Deep learning informed active learning), a novel methodology combining gLM embeddings with regressors (random forest, K-nearest neighbors, linear regression) to select sequences based on predicted expression values. Performance was compared against two baselines: random selection and gLM-guided zero-shot selection using log likelihood scores.

III. result

Through preliminary testing of DIAL on the aforementioned genetic circuit dataset, we observe that DIAL can successfully identify high-performing sequence candidates with an average 3 times fold reduction in the sample needed compared to random or gLM-guided zero-shot selection. Notably, nonlinear regressors such as random forest and K-nearest neighbors yield selected genetic circuits that achieve two-fold higher fold changes compared to linear models.

IV. Conclusion

These results offer promising benefits for genetic circuit development, particularly in time and resource-intensive experimental regimes, where single experiments can take up to weeks or even months. By enabling more efficient exploration of combinatorial design spaces, DIAL could accelerate synthetic biology applications and make genetic circuit optimisation more accessible across diverse research contexts.

References

[1] Wong ASL, Choi GCG, Cheng AA, et al. (2015) Massively parallel high-order combinatorial genetics in human cells. Nat Biotechnol 33:952–61

[2] Cameron DE, Bashor CJ, Collins JJ. (2014) A brief history of synthetic biology. Nat. Rev. Microbiol. 12:381–90.

[3] Barbadilla-Martínez L, Klaassen N, Steensel B van, et al. (2025) Predicting gene expression from DNA sequence using deep learning models. Nat Rev Genet, 1–15.

[4] Brixi G, Durrant MG, et al. (2025) Genome modeling and design across all domains of life with Evo 2. bioRxiv 2025.02.18.638918. https://doi.org/10.1101/2025.02.18.638918.

[5] Rai K, O'Connell RW, Piepergerdes TC, et al. (2023) Ultra-high throughput mapping of genetic design space. bioRxiv 2023.03.16.532704. https://doi.org/10.1101/2023.03.16.532704

Mapping Biological Data into Low-Dimensional Space With Prior Distribution

Zhiyuan Cai[1], Chao Tang[1]

***Short Abstract* — It is always an essential issue to build mathematical models with both interpretability and predictive ability for high-dimensional biological data. We process input data with variational autoencoders to obtain a low-dimensional representation. Then we make a bridge between the data distribution on the latent space and any given prior distribution by reversible transformations, so that we can apply simple low-dimensional models to high-dimensional data.**

***Keywords* — Dimension Reduction, Dynamical Modelling, Variational Autoencoder, Flow Matching, Single Cell RNA-seq**

I. Background

When we encounter a new biological problem, being able to quickly build a mathematical model from observational data can play a significant role in analyzing the primary factors and key processes of the system. While it is feasible to treat all observations as variables in the mathematical model, this approach can lead to an excessively high dimensionality of the system when faced with the growing volume of high-throughput data, resulting in computational complexity, overfitting, and poor interpretability. Another approach is to assume that the system is governed by low-dimensional latent variables. By identifying the latent variables corresponding to each data point, the analysis process can be transferred to a low-dimensional latent space, which is called dimensionality reduction.

To address different needs and adapt to various application scenarios, numerous dimensionality reduction algorithms have already been developed. However, most of these are designed for purely data-driven analysis, and the latent variables they produce are difficult to use for further mathematical modeling. To combine the advantages of data-driven methods and model-based methods, the dimensionality reduction process must be continuous and capable of mapping points in the latent space back into the original space. In this regard, variational autoencoders (VAEs), as a deep learning framework based on latent variable models, can effectively perform dimensionality reduction while simultaneously reconstructing data.

II. Method

We made a series of modifications to the vanilla VAE framework to better model specific problems.

A. Dimension scanning

The dimension of the latent space has a decisive impact on the result of dimensionality reduction. By manually adding masks of different lengths, we can calculate the reconstruction results from latent spaces of different dimensions, so that obtain the scanning results of the latent space dimension within a single training.

B. Prior distribution

The prior distribution of vanilla VAE is standard normal distribution, which is chosen for the convenience and does not necessarily align with the distribution of actual data. By employing hierarchical structures, we can shape the prior distribution of the latent space to match our desired form. An appropriate prior distribution not only reduces reconstruction error and minimizes posterior collapse, but also enhances the interpretability of the model.

C. Inversible Transformation

The process of combining the dimension scanning results with the prior distribution is called a reversible transformation. We used the method of flow matching to make the points obtained from dimension reduction without prior distribution correspond one-to-one with the points in the prior distribution, so that the model with prior knowledge can be applied to the space of the data itself.

III. Result

We applied the current method to two types of problems. The first type involves data where the control variables exhibit Bernoulli distribution characteristics, such as between experimental and control groups, or between different batches of data. By specifying a Gaussian mixture model as a prior, we can concentrate the influence of control variables on specific dimensions while separating out intrinsic system information on other dimensions. This enables batch effect removal or prediction of the results of a new cell type under experimental conditions. The second type involves data where the control variables exhibit periodicity, such as cell cycles and circadian rhythms. By specifying a uniform distribution on a circle, we can isolate the characteristics of periodic transformations, thereby enabling further analysis of factors that do not vary with the cycle.

[1]Center for Quantitative Biology, Peking University

MotY modulates proton-driven flagellar motor output in *Pseudomonas aeruginosa*

Sanyuan Fu[1†], Maojin Tian[3†], Min Chen[1], Zhengyu Wu[2*], Rongjing Zhang[1*], and Junhua Yuan[1*]

***Short Abstract* — MotY homologs are present in a variety of monotrichous bacterial strains and are thought to form an additional structural T ring in flagellar motors. In this study, we investigate the role of MotY in *P. aeruginosa*, elucidating its interactions with the two sets of stator units (MotAB and MotCD) using Förster resonance energy transfer (FRET) assays. Employing a newly developed bead assay, we characterize the dynamic behavior of flagellar motors in *motY* mutants, identifying MotY as the key functional protein to affect the clockwise bias of naturally unbiased motors in *P. aeruginosa*. Our findings reveal that MotY enhances stator assembly efficiency without affecting the overall assembly of the flagellar structure. Additionally, we demonstrate that MotY is essential for maintaining motor torque and regulating switching rates. Our study highlights the physiological significance of MotY in fine-tuning flagellar motor function in complex environments.**

***Keywords* — Stator, FRET, Bead assay, Flagellum, Bacterial motility.**

I. Introduction

FLAGELLA widely exist in bacterial systems as locomotor organs. The flagellar rotary motor comprises two main parts: the rotor and the stator. The rotor controls motor directional switching [1], while the stator converts the energy of ion flux into mechanical power for torque generation [2]. *P. aeruginosa* possesses two sets of stator units, MotAB and MotCD [3]. Stator-associated proteins MotY can generate a ring-like structure, the T ring [4]. Research on MotY's actual physiological function has been lacking, with motility studies primarily limited to phenotypic observations. This study aims to address this gap by employing multiple technical approaches.

II. Summary of Results

A. MotY interacts with both stator proteins MotB and MotD

We used the Förster resonance energy transfer (FRET) assay to elucidate the interaction between MotY and the stator units. In both measurements, we found that CFP emission increased upon bleaching of YFP, indicating energy transfer and protein interaction.

B. Absence of MotY results in lower motor speed and higher motor switching rate

We attached micron-sized beads to the flagellar filament stubs of the *motY* knockout strain and used a high-speed camera to accurately record their trajectory, depicting single-motor level dynamic output. The loss of MotY results in decreased motor speed, along with an increased switching rate.

C. T ring is the structural basis for ensuring the unbiased motor characteristics

Through measurement of 142 Δ*motY* cells, we found that the average CW bias dropped to around 0.3, a decrease of nearly 40%. The Δ*motY* strain no longer maintains the unbiased motor characteristics of *P. aeruginosa*, thus revealing a key functional protein that affects the switching preference of this type of motor.

D. Stator assembly efficiency, but not overall flagellar assembly efficiency, is affected by MotY deletion

We used flagellar filament labeling technology to determine the proportion of individuals with flagellar filaments in the Δ*motY* strain and found that it did not differ from the wild-type. Gene editing techniques helped us discover that the unipolar assembly efficiency of stator proteins was reduced in the above mutants.

III. Conclusion

In this study, we investigate the role of MotY in *P. aeruginosa* by employing multiple technical approaches, identifying MotY as the key functional protein to affect the clockwise bias of naturally unbiased motors in *P. aeruginosa*.

[1] Hefei National Research Center for Physical Sciences at the Microscale and Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, China.
[2]Research Center of Translational Medicine, Central Hospital Affiliated to Shandong First Medical University, Jinan, Shandong 250013, China.
[3] Center of Translational Medicine, Zibo Central Hospital Affiliated to Binzhou Medical University, Zibo, Shandong 255036, China.
Correspondence:
Zhengyu Wu, wuzhy@mail.ustc.edu.cn
Rongjing Zhang, rjzhang@mail.ustc.edu.cn
Junhua Yuan, jhyuan@mail.ustc.edu.cn

References

[1] Blair DF (2006) Fine structure of a fine machine. *J Bacteriol* **188**, 7033-5.
[2] Hu H, Santiveri M, et al. (2022) Structural basis of torque generation in the bi-directional bacterial flagellar motor. *Trends Biochem Sci*. **47**, 160-72.
[3] Harshey RM (2003) Bacterial motility on a surface: many ways to a common goal. *Annu Rev Microbiol.* **57**, 249-73.
[4] Terashima H, et al. (2006) The Vibrio motor proteins, MotX and MotY, are associated with the basal body of Na-driven flagella and required for stator formation. *Mol Microbiol*. **62**, 1170-80.

Evaluating the association between cancer-cell EMT and cytotoxic T lymphocyte infiltration in solid tumors

Yuewen Wang[1,#], Dandan Pu[1,#], Xin Xiong[1,#], Yerong Liu[1,#], Susmita Mandal[2], Peng Wang[2], Jack Chon Lok Lei[2], Mohit Kumar Jolly[3], Herbert Levine[4], Xuefei Li[1,5,*]

***Short Abstract* — This research investigates the relationship between epithelial-to-mesenchymal transition (EMT) of cancer cells and the infiltration of CD8+ T lymphocytes across multiple solid tumor types. Though both EMT and CD8+ T cell infiltration are established prognostic indicators, their interaction remains unclear. We performed gene set enrichment analysis (GSEA) on The Cancer Genome Atlas (TCGA) data, revealing that EMT features are associated with lower CD8A expression in various cancers. Furthermore, we plan to explore tumor microenvironment factors influencing this association through cell-cell communication analysis, aiming to enhance therapeutic strategies targeting cancer plasticity and immune evasion.**

***Keywords* — epithelial-to-mesenchymal transition, CD8⁺ T cell infiltration, gene set enrichment analysis, tumor microenvironment**

I. Purpose

THE primary goal of this extended abstract is to provide background information, summarize the findings, and outline key conclusions of our research on the relationship between epithelial-to-mesenchymal transition (EMT) of cancer cells and CD8+ cytotoxic T lymphocyte infiltration in solid tumors.

Both EMT and CD8+ T cell infiltration are independent prognostic factors for cancer outcomes [1,2]. However, their interplay remains controversial [3,4]. Utilizing data from The Cancer Genome Atlas (TCGA), we performed gene set enrichment analysis (GSEA) to examine the relationship between EMT characteristics and CD8A expression across 21 cancer types [5,6]. Preliminary findings indicate that epithelial features correlate with lower CD8A expression, while mesenchymal features are linked to higher CD8A levels [7]. Tumor purity significantly influences these relationships [8].

Our future research will focus on tumor microenvironmental factors that govern the interaction between CD8+ T cell infiltration and EMT. We plan to use cell communication analysis to uncover involved pathways [9]. This research aims to deepen our understanding of cancer biology and inform tailored immunotherapy strategies.s.

II. Conclusion

This study explores the intricate relationship between epithelial-to-mesenchymal transition (EMT) in cancer cells and CD8+ T lymphocyte infiltration in solid tumors. Our preliminary findings indicate that EMT features may correlate with decreased CD8A expression in some cancers, while in others, the reverse is true. This suggests that tumor purity and the microenvironment significantly influence these interactions.

Future research will utilize single-cell RNA sequencing and tools like CellChat to clarify the underlying mechanisms. Ultimately, understanding these dynamics is essential for developing targeted immunotherapies that enhance treatment efficacy and improve patient outcomes in cancer.

References

[1] Kalluri R, Weinberg RA (2009) The basics of epithelial-mesenchymal transition. In: The Journal of Clinical Investigation. McGraw-Hill, New York, pp. 1420–1428.
[2] Hanahan D, Weinberg RA (2011) Hallmarks of cancer: The next generation. In: Cell. Wadsworth, Belmont, CA, pp. 646–674.
[3] Yang J, Weinberg RA (2008) Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. In: Developmental Cell. Springer-Verlag, New York, Ch. 6.
[4] Dongre A, Weinberg RA (2019) New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. In: Nature Reviews Molecular Cell Biology, in preparation.
[5] Li H, Zhang Z (2016) TCGA data analysis for correlation of CD8+ T cell infiltration and transformation features in human cancers. J Cancer Research 76, 3536–3549.
[6] Jiang Y, et al. (2020) Tumor purity and the immune microenvironment affect outcomes in cancer patients. Nature Communications 11, 1034.
[7] Yamaguchi T, et al. (2021) The role of epithelial-mesenchymal transition in cancer development and progression. J Cancer Research and Clinical Oncology 147, 2219–2229.
[8] Parsa AT, et al. (2018) Tumor microenvironmental factors and the immune response to cancers. Nature Reviews Cancer 18, 781–793.
[9] Zhou H, et al. (2021) CellChat: Inference and analysis of cell-cell communication. Nature Protocols 16, 1880–1906.

Acknowledgements: This work was funded by the National Natural Science Foundation of China (32170672 and 32000886).

[1]Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. E-mail: xuefei.li@siat.ac.cn

[2] Faculty of Health Sciences, University of Macau, Macau SAR, China. E-mail: chonloklei@um.edu.mo

A Coevolution-Based Interaction Model for Metabolite–Recognizer Pairing and Microbial Interaction Network Inference

Guanyue Xiong[1, 2], Zhiyuan Li[1]

***Short Abstract* — Microbes produce a wide range of metabolites that support both fundamental physiological processes and specialized ecological functions. We define co-functional metabolites (CFMs) as a class of compounds whose biological function requires specific interactions with cognate recognizers (Recs)—such as selective transport, targeted modification, or self-protection—which further leads to coordinated molecular evolution. Such highly specific pair-wise recognitions are essential to precise interspecies interactions, holding great potential for targeted control of microbiomes. However, predicting CFM–Rec pairing relationships directly from genomic data remains challenging. Here, we present a general computational framework to solve this problem by the co-evolution between CFMs and recognizers. We introduced a statistically grounded index that robustly quantifies the evolutionary association between different CFM–Rec pairs. Based on this metric, we designed a computational framework to cluster CFMs and recognizers into functionally equivalent groups that specifically interact with each other. We applied our algorithm across three bacterial families to the siderophore-receptor system, a key mechanism of microbial iron acquisition. Our method can automatically decode the iron-mediated interaction networks.**

***Keywords* —metabolite, recognizer, co-evolution**

OVERVIEW

Deciphering microbial interactions is essential for understanding and precisely manipulating community dynamics. However, the detailed mechanisms of these interactions, those mediated by secondary metabolites, remain largely unexplored. In many cases, interaction specificity relies on pairs of co-functional metabolites (CFMs) and their cognate recognizers (Recs). For example, siderophores require specific membrane receptors for uptake [1]; polymorphic toxins depend on matching antitoxins for self-protection [2]; and cobamides selectively bind to cobamide-dependent enzymes [3].

The functional coupling between CFMs and recognizers is reflected at both molecular and sequence levels: different CFM variants are recognized by recognizers with corresponding structural differences at the binding interface, and gene families synthesizing CFMs and their recognizers often follow similar sequence variation patterns. Such coordinated evolution is widely hypothesized to mirror a molecular arms race, driving the diversification of active metabolites. This evolving molecular arsenal plays a key role in shaping complex microbial ecosystems. However, the complexity of diversified variants poses a major challenge for predicting interaction networks, as sequence divergence in CFMs or recognizers may either derived by phylogenetic artifacts, or true functional division. Accurately clustering CFMs and recognizers into functionally equivalent families that specifically interact with each other has remained an important challenge in microbiomics [4].

In this study, we proposed an unsupervised algorithm to cluster CFMs and recognizers into functionally equivalent groups, and to match CFMs with their cognate recognizers. The core assumption of our method is that a CFM group and its corresponding recognizer group should exhibit co-evolution in related sequences, resulting in a strong correlation in their occurrence across microbes. Based on this mutual anchoring between CFMs and recognizers, we define a co-evolution index that enables accurate identification of whether a given recognizer group functions as the self-recognizer of a specific CFM group. This index allows us to dynamically tune clustering boundaries to precisely delineate CFM and recognizer families, enabling accurate inference of microbial interaction networks. We chose the siderophore system as an ideal model to test our algorithm, for its high diversity and receptor-binding specificity. Altogether, our method offers a novel perspective on the detection of gene cluster families, and serves as a general tool for constructing microbial interaction networks based on recognizer-related metabolites.

REFERENCES

[1] Hider R C, Kong X. Chemistry and biology of siderophores[J]. Natural product reports, 2010, 27(5): 637-657.

[2] Ruhe Z C, Low D A, Hayes C S. Polymorphic toxins and their immunity proteins: diversity, evolution, and mechanisms of delivery[J]. Annual review of microbiology, 2020, 74(1): 497-520.

[3] Sokolovskaya O M, Shelton A N, Taga M E. Sharing vitamins: Cobamides unveil microbial interactions[J]. Science, 2020, 369(6499): eaba0165.

[4] Caesar L K, Butun F A, Robey M T, et al. Correlative metabologenomics of 110 fungi reveals metabolite–gene cluster pairs[J]. Nature Chemical Biology, 2023, 19(7): 846-854.

[1] Center for Quantitative Biology, Peking University.
[2] E-mail: guanyue.xiong@stu.pku.edu.cn

Graph-based tracing of dynamically functioning gene circuits in development with isoform resolution

Wanqi Li[1,3,4], Gang Xue[1,3], and Zhiyuan Li[1,2]

***Short Abstract* — Cell fate decisions are dynamically orchestrated by transcription factors (TFs) and corresponding gene regulatory networks (GRNs). However, TFs are lowly expressed with their functions beyond the transcriptional level and varied across tissues/lineages and its multiple isoforms. To this end, we showcased that the metric of transcription factor activity (TFA) renders high-fidelity regulatory relationships compared to expression. We integrated a consensus regulome database derived from human embryonic stem cell (ESC) with isoform resolution, and leveraged TFA to characterize functional switching between isoforms, and enrich regulatory modules during human preimplantation. Together, our work highlights the utility of TFA for distilling regulatory networks, and further facilitates the isoform-level investigation of TFs.**

***Keywords* — Gene Regulatory Network, Cell Fate Decision, Transcription Factor Activity, Master Regulator, Alternative Splicing, Isoform Switching, Hematopoiesis, Regulon Database**

I. Background

TRANSCRIPTION factor (TF) and its corrsponding gene regulatory networks (GRNs) orchestrate complex cellular behaviors in cell fate decisions [1]. However, current studies aim to reconstruct GRN faced three major challenges: (1) TFs function at protein level that is limited to assess with transcriptomic data; (2) TFs possess low abundance and high noise in single-cell data; (3) snapshot of most transcriptomic assays fails to capture transient intermediates in cell fate transitions.

To overcome the disconnect between TF expression and function, transcription factor activity (TFA) metric have been developed to estimate TF functional states from the expression of their downstream targets, or regulons along with several regulon databases [2, 3]. Applying TFA to infer transcription factor (TF) regulatory networks faces several critical challenges: (1) lack of systematic benchmarking; (2) mismatch with tissue-specific TF functions that current regulon databases predominantly aggregate data across diverse cell types; (3) overlooking the functional complexity introduced by alternative splicing.

[1]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, 100871, China.

[2]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, 100871, China.

[3]These authors contributed equally to this work.

[4]E-mail: wanqili@pku.edu.cn

II. Result

A. TFA enables high-fidelity inference of TF regulatory relationships

We systematically evaluated the feasibility of using TFA to infer regulatory relationships. First, we demonstrated that TFA offers a more faithful readout of regulatory interactions than expression profiles, both in simulated networks and real-world spatial transcriptome with dropout events.

We next applied TFA to a 11-node hematopoietic network and a single-cell hematopoiesis atlas, indicating that TFA retains traces of upstream unctioning TF circuits in endpoint populations even when intermediate states are undersampled.

Futher, we compared the use of direct and transited regulons in TFA analysis. Our analysis showed that transited regulons facilitated the identification of dynamically functioning TFs more effectively in complex developmental systems.

B. Integration of comprehensive regulome specific to embgryonmic development with isoform resolution

To assess the influrence of cell-type specificity in regulon and address the limitations of current regulon resources, we constructed an embryonic stem cell (ESC)-specific transited regulon database (CRED) with isoform resolution. Based on that, we quantified functional heterogeneity among TF isoforms, highlighting their orthogonality in regulations and the necessity of isoform-level resolution in accurately inferring gene regulatory relationships.

C. Application of isoform-level TFA reveals germ layer‐specific regulators in human early development

We applied our isoform-resolved regulome (CRED) and TFA pipeline human preimplantation development. TFA enables tracing of isoform-level TF regulatory modules in single-cell datasets with severe dropouts and proposes new candicate regulators.

References

[1] Lambert, S. A. et al. (2018) The Human Transcription Factors. *Cell* **172**, 650‐665.

[2] Han, H. et al. (2015) TRRUST: a reference database of human transcriptional regulatory interactions. *Sci. Rep*. **5**, 11432.

[3] Garcia-Alonso, L. et al. (2019) Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res.* **29**, 1363‐1375.

Low-cost intercellular communication system in mammalian cells enabled by *de novo* design of super-sensitive intercellular communication signals

Zhi Sun[1#], Yanhui Xiang[2#], Yukui Pan[2], Min Yu[2], Long Qian[1], Qi Ouyang[1,3,4] and Chunbo Lou[2,5]*

***Short Abstract* — Mammalian cells rely on various intercellular signals to coordinate essential physiological and developmental processes; however, previous reported synthetic small molecule signals often suffer from lower sensitivity and impose significant metabolic burdens due to the biosynthesis of their precursors. We here developed a super-sensitive (EC_{50} ~ 10^{-9} mol/L) and low-burden cell-cell communication system, consisting of *de novo* designed sender and receiver modules. We also introduced a degradation module by systematically screening degradation enzymes. We finally applied the engineered intercellular communication system for precise and robust morphogen gradient formation over long distances.**

I. Background

IN multicellular systems, diverse and efficient intercellular signals are critical for developmental patterning, immune responses, and the maintenance of homeostasis. A wide array of biochemical molecules function as signal carriers in these processes, including small molecules, proteins, and RNAs. Small molecules offer several advantages over proteins, including faster diffusion, easier modulation, and reduced immunogenicity. Several small molecules have been developed as advanced signaling tools to enable sophisticated dynamic and spatial control in mammalian genetic circuits (e.g., auxins, HSLs). However, the threshold concentrations (EC_{50}) required to activate their receptors are often too high to be efficiently produced by mammalian cells without causing a metabolic burden, which would disrupt normal growth and physiological functions of mammalian cells.

[1]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[2]Center for Cell and Genetic Circuit Design, State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[3]The State Key Laboratory for Artificial Microstructures and Mesoscopic Physics, School of Physics, Peking University, Beijing 100871, China
[4]Institute for Advanced Study in Physics, School of Physics, Zhejiang University, Hangzhou 310058, China
[5]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, 100149, China
[#]These authors contributed equally: Zhi Sun and Yanhui Xiang.
*To whom correspondence may be addressed: Chunbo Lou (cb.lou@siat.ac.cn).

II. Results

In this work, we created a new intercellular communication channel in mammalian cells with the signal molecule cinnamoyl-HSL(Cinn), a novel aromatic acyl-HSL with an ultra-high sensitivity (EC_{50} ~ 10^{-12} mol/L) in bacteria. After phylogenetic analysis and sequence alignments, we identified a potential *braO* region upstream of the putative -35 box of *braI*'s promoter.

We designed the Cinn-receiver module and tested it in HEK 293T human cells. The titration experiments revealed that the EC_{50} of the designed receiver module is around nM-level, which is much lower than previously reported pC or IV receiver modules.

We minimized the braO sequence by truncating from both ends, and finally acquired a compact 18-bp braO-specific binding sequence for BraR. AI-assisted domain swapping strategy further reduces the receiver's leakage expressions.

We de novo designed a metabolic pathway to synthesize Cinn from phenylalanine, an essential amino acid, with three enzymes (yPAL, 4CL & BraI) from different species. The sender module enables intercellular signaling without any specific precursors.

We collected a diverse set of lactonases from different species for potential degrader modules. Two bacterial -derived lactonases and one human-derived paraoxonase exhibited significant activity.

We developed a synthetic "Source-Sink" patterning circuit by constructing synthetic developmental fields, where the sender and degrader defined "source" and " sink" field boundaries, and the receiver within the field decoded the gradient information. Experimental validations supported the key role of the "Source-Sink" mechanism in establishing and maintaining of robust sharp gradient patterns.

III. Conclusion

We constructed a novel, low-cost, super-sensitive communication system in mammalian cells. This efficient and user-friendly intercellular communication system provides a powerful tool for exploring design principles for multicellular organisms, enabling rational control of complex mammalian patterning, and offering potential therapeutic intervention for developmental pattern abnormalities.

Quantitative Biology

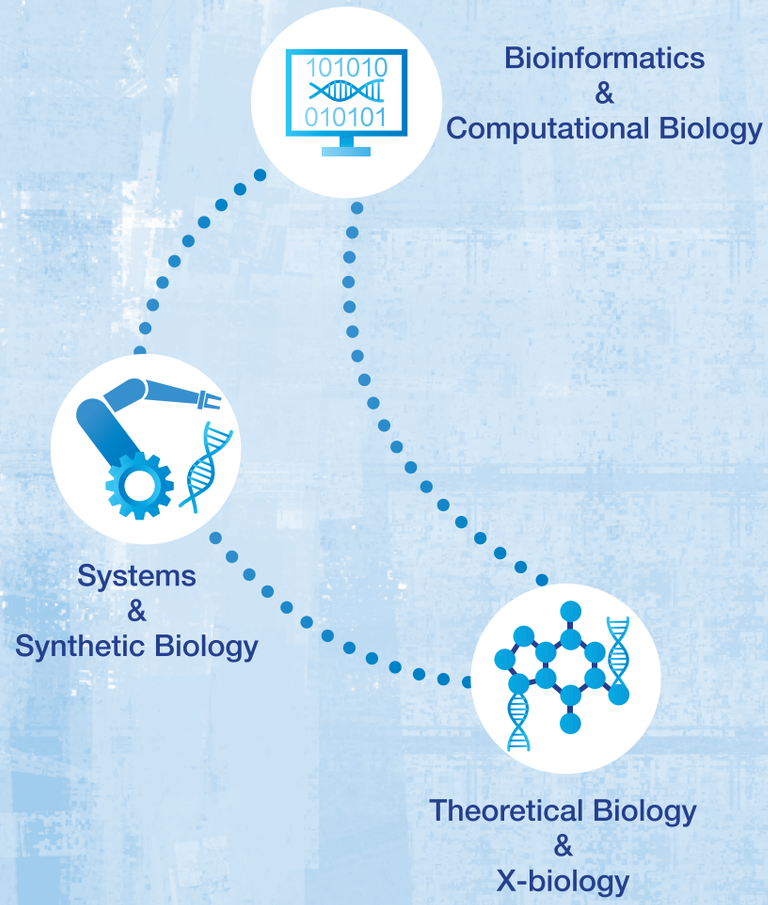
Open Access

An X-disciplinary Journal Leading to Future Biology

Aims

Quantitative Biology is an interdisciplinary journal that focuses on cutting-edge progresses, in-depth reviews as well as future perspectives of research that uses all quantitative approaches to analyze, model and engineer biological systems, and to gain quantitative understanding of the mystery of life. The quarterly journal aims to provide a platform for sharing not only new methods, technologies and results, but also novel or even sometimes bold ideas, visions and perspectives. “All rivers run into sea.” The journal covers all disciplines related to experimental biology, theoretical biology, computational biology, systems biology, synthetic biology, digital and AI biology.

Scopes



Journal Metrics

Time for the first decision: 5 days
Time for peer-review decision: 10 days
Time from submission to accept: 90 days
Time for fast track papers: 7 days

Indexing Information

Emerging Sources Citation Index
SCOPUS
Biological Abstracts
BIOSIS
CAS: Chemical Abstracts Service
COPE

Content Available Online

https://wileyonlinelibrary.com/journal/qb
https://journal.hep.com.cn/qb

Online Submission

https://mc.manuscriptcentral.com/qb

Contact

Miaomiao Tian tianmm@hep.com.cn

清華大學 Tsinghua University

北京大學 PEKING UNIVERSITY

Higher Education Press

WILEY

HEP website

WOL website

QB Wechat