北京大学定量生物学中心
CENTER FOR QUANTITATIVE BIOLOGY

Quantitative Biology 2023:

AI and Cell Fate

Saturday 14 October 2023 - Sunday 15 October 2023

Peking University, China

Host:
Center for Quantitative Biology (CQB) at Peking University（https://cqb.pku.edu.cn/）

Organization Committee:
Chao Tang (CQB, Peking University, China)
Jianhua Xing (University of Pittsburgh, USA)
Zexian Zeng (CQB, Peking University, China)
Zhiyuan Li (CQB, Peking University, China)
Yihan Lin (CQB, Peking University, China)

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

Conference Website:
https://qbio.pku.edu.cn/QB2023/

Supported by:

生命科学联合中心 CENTER FOR LIFE SCIENCES

成都高新 CDHT 北京大学成都前沿交叉生物技术研究院 Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies

北京大学 PEKING UNIVERSITY 国际合作部 港澳台办公室

北京大学 PEKING UNIVERSITY 研究生院

Sponsors：

PARATERA 并行®

捷通 CHATONE

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

Friday, October 13th **Registration 14: 00-18: 00** **Bldg 1＃，Zhongguanyuan Global Village PKU（中关新园 1 号楼大厅）**

Saturday, October 14th **Registration 07: 20-08: 20** **Bldg 1＃，Zhongguanyuan Global Village PKU（中关新园 1 号楼群英厅）**

2023/10/14（Saturday）

Time	Speaker	Presentation Title
08:30-08:40	Opening Remarks (**Chao Tang)**	
Session 1 (Chair: **Zexian Zeng**)		
08:40-09:05	**Jianhua Xing** University of Pittsburgh, USA	Dynamics of Cell State Transitions Emerges as A New Frontier of Studying Complex Systems
09:05-09:30	**Ting Wang** Washington University in St.Louis, USA	The Human Pangenome Project
09:30-09:55	**Wei Xie** Tsinghua University, China	Decoding The Transcription Circuitry When Life Begins
09:55-10:40	**Group photo/ Poster Session/ Coffee Break**	
Session 2 (Chair: **Jianhua Xing**)		
10:40-11:05	**Siyuan (Steven) Wang** Yale University, USA	A Genome-Wide Single-Cell 3D Genome Atlas of Lung Cancer Progression
11:05-11:30	**Matheus Palhares Viana** Allen Institute, USA	Developing Quantitative Frameworks for Intracellular Organization and Dynamics in Human Induced Pluripotent Stem Cells
11:30-11:55	**Mingwei Min** Guangzhou Laboratory, China	The Processual Perspective of Cell Fate Decisions: A Case Study in The Proliferation-Quiescence Decision
12:00-13:30	**Lunch** (Bldg 6#, Chenguang Cafeteria 辰光咖啡厅)	
Session 3 (Chair: **Zhiyuan Li**)		
13:30-13:55	**Qing Nie** University of California, Irvine, USA	Multiscale Spatiotemporal Reconstruction of Single-Cell Genomics Data
13:55-14:20	**Zheng Hu** SIAT, CAS, China	Mapping Single-Cell Velocity Fields Using Monotonically Expressed Genes
14:20-14:45	**Tatsuhisa Tsuboi** Tsinghua SIGS, China	Mitochondrial Protein Heterogeneity Stems From the Stochastic Nature of Co-Translational Protein Targeting in Cell Senescence
14:45-15:10	**Peijie Zhou** Peking University, China	Dissecting Cell-State Attractors and Transitions in Single-Cell Transcriptome Data
15:10-15:40	**Poster Session/ Coffee Break**	
Session 4 (Chair: **Jie Lin**)		
15:40-16:05	**Jian Ma** Carnegie Mellon University, USA	Single-Cell Spatial Modeling of Cell Identity
16:05-16:30	**Jialiang Huang** Xiamen University, China	Deciphering Spatial Regulation of Enhancer Networks Driving Cell Fate Determination at Single-Cell Level
16:30-16:55	**Zhiyuan Li** Peking University, China	Making the Gene Regulatory Networks More Logical
17:30-19:30	**Dinner** (Bldg 6#, Chenguang Cafeteria 辰光咖啡厅)	

Sunday, October 15th Bldg 1#，Zhongguanyuan Global Village PKU（中关新园 1 号楼群英厅）

2023/10/15（Sunday）		
Time	**Speaker**	**Presentation Title**
Session 1 (Chair: **Kaifu Chen**)		
08:30-08:55	**Luonan Chen** CEMCS, CAS, China	Dynamical Data Science and AI for Biology
08:55-09:20	**Lei Zhang** Peking University, China	Construction of Solution Landscapes of Complicated Biological Systems
09:20-09:45	**Xiaowo Wang** Tsinghua University, China	Flanking Sequence Engineering for Efficient Promoter Design
09:45-10:10	**David van Dijk** Yale University, USA	Understanding Cellular State Changes Through Causal Inference and Large Language Models
10:10-10:30	**Poster Session/Coffee Break**	
Session 2 (Chair: **Zhi Qi**)		
10:30-10:55	**Duanqing Pei** Westlake University, China	Cell Fate Decision by A Morphogen-Transcription Factor-Chromatin Modifier Axis
10:55-11:20	**Xiaohua Shen** Tsinghua University, China	From DNA to Life: Decode the Noncoding Genome
11:20-11:45	**Harinder Singh** University of Pittsburgh, USA	Gene Regulatory Networks Orchestrating Immune Cell Fate Dynamics
11:45-12:10	**Fan Zhou** Tsinghua University, China	Peri-Implantation Embryogenesis and Regulation
12:10-13:30	**Lunch** (Bldg 6#, Chenguang Cafeteria 辰光咖啡厅)	
Session 3 (Chair: **Ming Han**)		
13:30-13:55	**Yi Xing** Children's Hospital of Philadelphia, USA	Omics and AI Strategies to Study Splicing Defects in Rare Diseases
13:55-14:20	**Guoji Guo** Zhejiang University, China	Mapping Cell Landscapes at Single-Cell Level
14:20-14:45	**Kaifu Chen** Harvard University, USA	System Methods to Define Master Regulator of Cell Identity in Development and Diseases
14:45-15:10	**Shouwen Wang** Westlake University, China	Infer Cell Dynamics and Clonal Memory from Single-Cell Lineage Tracing
15:10-15:40	**Poster Session/ Coffee Break**	
Session 4 (Chair: **Yihan Lin**)		
15:40-16:05	**Ge Guo** University of Exeter, UK	Regulatory Mechanism of Cell Lineage Segregation in Early Human Embryogenesis
16:05-16:30	**Ming Han** Peking University, China	Statistical Learning of Multicellular Dynamics
16:30-16:55	**Tong Zhu** East China Normal University, China	Target and Drug Discovery Based on Covalent Interaction
16:55-17:20	**Jingdong (Jackie) Han** Peking University, China	AI and Aging
17:20-17:30	Closing Remarks (**Jianhua Xing**)	
Close		

Dynamics of Cell State Transitions Emerges as A New Frontier of Studying Complex Systems

Jianhua Xing

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

Email: xing1@pitt.edu

It has long been a perplexing question in the field of life sciences: How can a single cell develop into a complex organism comprised of various types of cells? The ability to reprogram terminally differentiated cells into pluripotent stem cells or other cell types has opened up new avenues for regenerative medicine. These advancements in biomedical research have also given rise to a nascent field focused on studying cell state transitions within the framework of dynamical systems. This field naturally extends the study of state transitions in physics and chemistry (Hanggi et al., Rev Mod Phys, 62:251, 1990) to complex cellular systems that exist out of thermodynamic equilibrium. Such systems exhibit a large number of highly interconnected degrees of freedom and operate across broad timescales.

The recent explosion of single-cell data presents both challenges and unprecedented opportunities for physicists to analyze the data and uncover underlying mechanistic principles. In this presentation, I aim to provide an introduction to this exciting and rapidly evolving field, with a particular focus on the research conducted in my laboratory.

Jianhua Xing

Dr. Xing received B.S. in Chemistry from Peking University, M.S. in Chemical Physics from University of Minnesota, and PhD in Theoretical Chemistry from UC Berkeley. After being a postdoc researcher in theoretical biophysics at UC Berkeley and an independent fellow at Lawrence Livermore National Laboratory, he assumed his first faculty position at Virginia Tech, then moved to University of Pittsburgh in 2015. Currently Dr Xing is a professor in the Computational and Systems Biology Department, School of Medicine, and an affiliated faculty member of Department of Physics and Astronomy, University of Pittsburgh. He is also an affiliated member of University of Pittsburgh Hillman Cancer Center. Dr. Xing's research uses statistical and chemical physics, dynamical systems theory, mathematical/computational modeling in combination with quantitative measurements to study the dynamics and mechanism of biological processes.

The Human Pangenome Project

Ting Wang

School of Medicine, Washington University in St.Louis, USA

Email: twang@wustl.edu

The Human Pangenome Project will create a human pangenome reference that represents genomic diversity across human populations, that improves technology for assembly, and that develops a robust tool ecosystem for pangenomic analyses. I will introduce the project's effort in generating high-quality genome assemblies with diverse genetic ancestries and building computational representations of this multi-genome reference (the pangenome). I will also discuss the impact of the human pangenome on genetic variation studies and functional genomics studies.

Ting Wang

Dr. Ting Wang is the inaugural Sanford C. and Karen P. Loewentheil Distinguished Professor of Medicine and Head of Genetics Department at Washington University School of Medicine in St. Louis. Dr. Wang studies the genetic and epigenetic impact of transposable element (TE) on gene regulation. His group is known for defining the widespread contribution of TEs to the evolution of gene regulatory networks as well as to the 3D genome architecture, and for revealing that epigenetic dysregulation of TEs is a major mechanism driving oncogenesis. His lab is home to the WashU Epigenome Browser, utilized by investigators around the world to access hundreds of thousands of genomic datasets generated by large Consortia including the NIH Roadmap Epigenome Project, ENCODE, 4D Nucleome, TaRGET, IGVF, and the Human Pangenome Project. Dr. Wang currently directs the NIEHS Environmental Epigenomics Data Center, the Human Pangenome Reference Consortium, the IGVF Data Administrative and Coordination Center, the SMaHT Network Organization Center and Genome Characterization Center.

Decoding The Transcription Circuitry When Life Begins

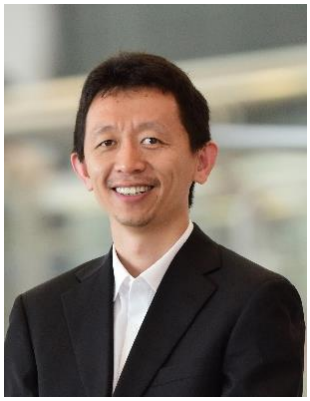
Wei Xie

School of Life Sciences, Tsinghua University, Beijing, China, 100084

Email: xiewei121@tsinghua.edu.cn

Zygotic genome activation (ZGA) marks the beginning of the embryonic program for a totipotent embryo, which later gives rise to inner cell mass (ICM) where pluripotent epiblast arises to form germ layers, and extraembryonic trophectoderm that subsequently contributes to placenta. Deciphering key transcription factors from ZGA to the first lineage segregation is crucial for understanding how life begins and how a totipotent embryo arises from terminally differentiated gametes. Probing these questions in mammals however was long hindered by the scarce experimental materials that are available from early embryos. By developing a set of ultra-sensitive chromatin analyses and Ribo-seq technologies, we recently identified key transcription factors TPRX and OBOX that act at the onset of ZGA in human and mouse, respectively. Nevertheless, how ZGA is connected to the first lineage segregation in early embryos remains elusive. In this talk, I will discuss how our recent findings, powered by these cutting-edge technologies, may help illuminate the core transcription circuitry underlying the beginning of life and the earliest cell fate determination.

Wei Xie



Dr. Wei Xie is a Professor and Vice Dean of School of Life Sciences, Tsinghua University, and also an HHMI International Research Scholar. He received his B.S. degree in Molecular Biology at Peking University in China in 2003. He pursued his Ph.D study at UCLA, where he joined the laboratory of Michael Grunstein to study the function of histones and histone modifications. He also obtained an M.S. double degree in statistics at UCLA with Ker-Chau Li. After completing his graduate studies in 2008, he continued research in epigenetics and transcription regulation as a postdoctoral fellow in Bing Ren's lab at the Ludwig Institute for Cancer Research, UCSD in 2009. After his postdoc training, he joined Tsinghua University, School of Life Sciences, in Beijing as a Principal Investigator in 2013. He is also a member of the Tsinghua-Peking Joint Center for Life Sciences. Using interdisciplinary approaches, Dr. Xie is dedicated to understanding how the life clock is reset after fertilization in mammalian embryos. His group established a series of ultra-sensitive technologies to analyze chromatin dynamics using hundreds of cells or fewer. By doing so, his team revealed how chromatin accessibility, histone modifications, and 3D chromatin architecture are reprogrammed during early mammalian development. His work also

demonstrated how the embryonic program is activated during zygotic genome activation. Such epigenetic reprogramming is essential for successful parental-to-zygotic transition and the ultimate generation of a totipotent embryo, which occurs through regulatory mechanisms that are often distinct from those in somatic cells and embryonic stem cells. He has authored over 80 publications with over 19,000 citations. He also received numerous awards including the HHMI International Research Scholar and New Cornerstone Investigator. Dr. Xie previously or currently served as the Review Editor of *Science*, the editorial board members of Stem *Cell Reports* and *Development*.

A Genome-Wide Single-Cell 3D Genome Atlas of Lung Cancer Progression

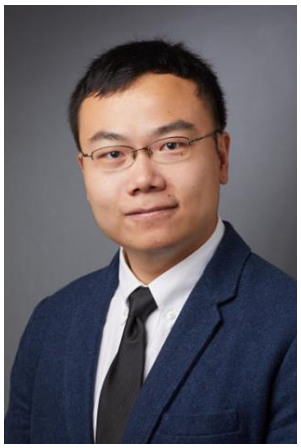
Siyuan (Steven) Wang

Yale School of Medicine, Yale University, USA

Email: siyuan.wang@yale.edu

Alterations in three-dimensional (3D) genome structures are associated with cancer. However, how genome folding evolves and diversifies during subclonal cancer progression in the native tissue environment remains unknown. Here, we leveraged a genome-wide chromatin tracing technology to directly visualize 3D genome folding *in situ* in a faithful *Kras*-driven mouse model of lung adenocarcinoma (LUAD), generating the first single-cell 3D genome atlas of any cancer. We discovered stereotypical 3D genome alterations during cancer development, including a striking structural bottleneck in preinvasive adenomas prior to progression to LUAD, indicating a stringent selection on the 3D genome early in cancer progression. We further showed that the 3D genome precisely encodes cancer states in single cells, despite considerable cell-to-cell heterogeneity. Finally, evolutionary changes in 3D genome compartmentalization – partially regulated by polycomb group protein Rnf2 through its ubiquitin ligase-independent activity – reveal novel genetic drivers and suppressors of LUAD progression. Our results demonstrate the importance of mapping the single-cell cancer 3D genome and the potential to identify new diagnostic and therapeutic biomarkers from 3D genomic architectures.

Siyuan (Steven) Wang



Siyuan (Steven) Wang, Ph.D., is an Associate Professor of Genetics and Cell Biology at Yale School of Medicine. Research in his lab focuses on the development and application of state-of-the-art image-based omics approaches to understand the spatial organization of mammalian genome and transcriptome, and how they impact cellular states. Originally from China, Dr. Wang received a Bachelor of Science degree in Physics from Peking University in 2007. He then moved to the US and received a Ph.D. in Molecular Biology from Princeton in 2011 and later his postdoctoral training at Harvard. In 2017 he was appointed Assistant Professor by Yale University, and was promoted to the rank of Associate Professor in 2023. Dr. Wang developed/co-developed multiple influential technologies in the spatial omics field including the first-in-kind image-based 3D genomics method termed "chromatin tracing" to trace the spatial folding of genome, "MERFISH" for spatial transcriptome

profiling (*Nature Methods* Method of the Year 2020), and "MINA" to map multiscale chromatin folding, copy numbers of numerous RNA species, and associations of numerous genomic regions with nuclear landmarks in the same, single cells in mammalian tissue. His recent work focused on generating the first single-cell 3D genome atlas in cancer and building the 3D genomic "regulatome". He received the 2011 American Physical Society Award for Outstanding Doctoral Thesis Research in Biological Physics (1-2 recipients per year worldwide), the 2012-2015 ane Coffin Childs Fellowship, the 2016 International Union of Pure and Applied Physics Young Scientist Prize in Biological Physics (one recipient per year worldwide), the 2017 Harvard Chinese Life Sciences Distinguished Research Award, the 2018 35 Innovators Under 35 of China by *MIT Technology Review*, the 2019 NIH Director's New Innovator Award, the 2022 Pershing Square Sohn Prize for Young Investigators in Cancer Research, and the 2023 Biophysical Society Early Career Award in Physical Cell Biology.

Developing Quantitative Frameworks for Intracellular Organization and Dynamics in Human Induced Pluripotent Stem Cells

Matheus Palhares Vian

Allen Institute, USA

Email: matheus.viana@alleninstitute.org

The Allen Institute for Cell Science aims to understand the principles by which human induced pluripotent stem cells (hiPSCs) establish and maintain robust dynamic localization of cellular structures, and how they transition between states during differentiation and disease. To do this, we take advantage of 3D microscopy images, molecular and cellular image-based assays, and machine learning tools to characterize cell states holistically, integrating measurements of cellular organization and behavior dynamics with the cell's molecular census and the cell-extrinsic environment. As an initial step towards this goal, we developed a computational analysis framework for integrated intracellular organization that embraces the vast cell-to-cell variability observed within a normal population of hiPSCs and permits quantitative analyses of distinct, separable aspects of organization within and across different cell populations. In building this framework, we found a combination of spherical harmonics expansion and PCA provides compact and interpretable representations for cell and nuclear shapes. We then extended our analysis by incorporating temporal dynamics to answer how hiPSCs control their nuclear growth and shape. We found that the nucleus undergoes two distinct phases of growth regulation, an early and late growth phase, with a transition point occurring in G1 at a consistent time and size. Although successful as cell and nuclear shape representations, spherical harmonics are not well suited to represent other morphologies, such as tubular networks, spatial distributions and multi pieces structures. To extend our framework to these more complex morphologies, we need to develop appropriate data representations to capture the biological variability of interest. We tested the use of point clouds as representations for spatial protein pattern like DNA replication foci (via PCNA) and signed distance functions as representations for intricated, multi pieces shapes like nucleoli (via nucleophosmin). In both cases, we obtained compact representations and high-fidelity reconstructions using these approaches. We use a variational autoencoder to learn, in an unsupervised way, latent variables from these data representations. We found the learned latent variables recapitulate known sources of variation for these structures, such as changes in number of pieces and volume for nucleoli, and characteristic protein pattern across cell cycle for PCNA.

Authors: Matheus P. Viana[1], Allen Institute for Cell Science, Julie A. Theriot[2], Susanne Rafelski[1]

1 Allen Institute for Cell Science Seattle, WA, USA
2 Department of Biology and Howard Hughes Medical Institute, University of Washington Seattle, WA USA

Matheus Palhares Vian

Bachelor's degree in physics from Federal University of Sao Carlos in 2004, Brazil. Master's and PhD degrees in Computational Physics from University of Sao Paulo, Brazil. Research focus on image analysis and network analysis applied to complex systems. Post-doc in University of California Irvine using graph theory to study mitochondrial networks in budding yeast. Worked at IBM Research in Brazil applying machine learning for medical image analysis. Working at the Allen Institute for Cell Science in the last 5 years. Currently leading a team that works with image analysis and machine learning for microscopy data representation.

The Processual Perspective of Cell Fate Decisions: A Case Study in The Proliferation-Quiescence Decision

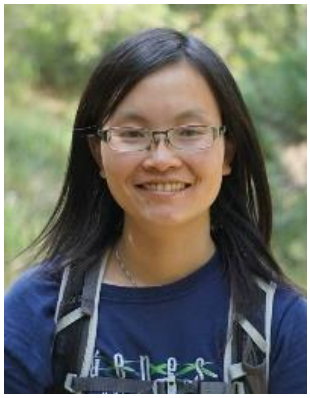
Mingwei Min

Guangzhou Laboratory, China

Email: min_mingwei@gzlab.ac.cn

While molecular biology has undoubtedly provided us with profound insights into molecular components of cellular systems, the crucial missing links often lie in understanding how these components orchestrate phenotypes over time. In this talk, I will share one such example. Multicellular organisms use mitogens to regulate cell proliferation, but how fluctuating mitogenic signals are converted into proliferation-quiescence decisions is poorly understood. We combined live-cell imaging with temporally controlled perturbations to determine the time scale and mechanisms underlying this system in human cells. Contrary to the textbook model that cells sense mitogen availability only in the G1 cell cycle phase, we find that mitogenic signaling is temporally integrated throughout the entire mother cell cycle and that even a 1-hour lapse in mitogen signaling can influence cell proliferation more than 12 hours later. This case study underscores the pivotal significance of the processual perspective in studying cell fate decisions.

Mingwei Min



Mingwei completed her PhD with Cath Lindon at the University of Cambridge, working on protein degradation in mitosis. She then moved across the pond to the US and did a postdoc with Dan Finley at the Harvard Medical School, and then with Sabrina Spencer at the University of Colorado-Boulder, to study how cells choose between proliferation and quiescence. She recently started her own lab at the Guangzhou Laboratory, exploring how internal cell states and environmental information jointly control cell fate in various systems.

Multiscale Spatiotemporal Reconstruction of Single-Cell Genomics Data

Qing Nie

Department of Mathematics
Department of Developmental and Cell Biology
NSF-Simons Center for Multiscale Cell Fate Research
University of California, Irvine

Email: qnieact@gmail.com

Cells make fate decisions in response to dynamic environments, and multicellular structures emerge from multiscale interplays among cells and genes in space and time. The recent single-cell genomics technology provides an unprecedented opportunity to profile cells. However, those measurements require fixing individual cells that lose many important spatiotemporal information. Is it possible to infer temporal relationships among cells from single or multiple snapshots? How to recover spatial interactions among cells, for example, cell-cell communication? In this talk I will present our newly developed computational tools that are mostly based on dynamical models and machine-learning methods, with a focus on inference and analysis of transitional properties of cells and cell-cell communication using both single-cell and spatial transcriptomics, as well as multi-omics data for some cases. Through their applications to various complex systems in development, regeneration, and diseases, we show the discovery power of such methods in addition to identifying areas for further method development for spatiotemporal analysis of single-cell data.

Qing Nie

Dr. Qing Nie is a Distinguished Professor of Mathematics and Developmental & Cell Biology at University of California, Irvine. Dr. Nie is the director of the *NSF-Simons Center for Multiscale Cell Fate Research* jointly funded by NSF and the Simons Foundation – one of the four national centers on mathematics of complex biological systems. In research, he uses systems biology and data-driven methods to study complex biological systems with focuses on single-cell analysis, multiscale modeling, cellular plasticity, stem cells, embryonic development, and their applications to diseases. Dr. Nie has published more than 200 research articles. In training, Dr. Nie has supervised more than 50 postdoctoral fellows and PhD students, with many of them working in academic institutions. Dr. Nie is a fellow of the *American Association for the Advancement of Science*, a fellow of *American Physical Society*, and a fellow of *Society for Industrial and Applied Mathematics*.

Mapping Single-Cell Velocity Fields Using Monotonically Expressed Genes

Zheng Hu

Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences(CAS), China

Email: zheng.hu@siat.ac.cn

Single-cell RNA sequencing (scRNA-seq) is a powerful approach for studying cellular differentiation, but accurately tracking cell fate transitions can be challenging, especially in disease conditions. Here we introduce PhyloVelo, a computational framework that estimates the velocity of transcriptomic dynamics by using monotonically expressed genes (MEGs) or genes with expression patterns that either increase or decrease, but do not cycle, through phylogenetic time. Through integration of scRNA-seq data with lineage information, PhyloVelo identifies MEGs and reconstructs a transcriptomic velocity field. We validate PhyloVelo using simulated data and *Caenorhabditis elegans* ground truth data, successfully recovering linear, bifurcated and convergent differentiations. Applying PhyloVelo to seven lineage-traced scRNA-seq datasets, generated using CRISPR–Cas9 editing, lentiviral barcoding or immune repertoire profiling, demonstrates its high accuracy and robustness in inferring complex lineage trajectories while outperforming RNA velocity. Additionally, we discovered that MEGs across tissues and organisms share similar functions in translation and ribosome biogenesis.

Zheng Hu

Dr. Zheng Hu is a Principal Investigator at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (CAS). Before joining SIAT, he was a postdoctoral fellow in Dr Chrsitina Curtis's laboratory at Stanford University School of Medicine. He received his Ph.D. in Evolutionary Genetics from Beijing Institute of Genomics CAS and B.S. in Biomedical Engineering from Huazhong University of Science and Technology. Dr Hu's research interests lie at the interface of evolutionary genetics and cancer biology. He has focused on developing and applying computational and evolutionary approaches to study the mechanisms and dynamics of somatic evolution in cancer and tissue homeostasis.

Mitochondrial Protein Heterogeneity Stems From the Stochastic Nature of Co-Translational Protein Targeting in Cell Senescence

Abdul Haseeb Khan[1], Aidan I. Brown[2], Matheus P. Viana[3], Susanne M. Rafelski[3], Brian M. Zid[4],* and Tatsuhisa Tsuboi[1,4,5,*]

[1]Institute of Biopharmaceutical and Health Engineering, Tsinghua Shenzhen International Graduate School, Shenzhen, China

[2]Department of Physics, Toronto Metropolitan University, Toronto, ON, Canada

[3]Allen Institute for Cell Science, Seattle, WA, USA

[4]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA, USA

Email: ttsuboi@sz.tsinghua.edu.cn

A decline in mitochondrial function is a hallmark of aging and many neurodegenerative diseases. It has been proposed that changes in mitochondrial morphology, including fragmentation of the tubular mitochondrial network, can lead to mitochondrial dysfunction, yet the mechanism of this loss of function is unclear. Most proteins contained within mitochondria are nuclear-encoded and must be properly targeted to the mitochondria. Here we report that sustained mRNA localization and co-translational protein import leads to heterogeneous protein distribution across fragmented mitochondria. We find that age-induced mitochondrial fragmentation drives an exponential increase in protein expression noise across fragments in yeast, mammalian cells, and nematodes. Using a translational kinetic and molecular diffusion model, we find that protein noise is explained by the nature of stochastic compartmentalization and that the co-translational protein import is the main contributor to the increased heterogeneity. We observed that cells repress heterogeneity of protein distribution mainly by mitochondrial fission-fusion reaction and not through the mitophagy pathway. Furthermore, we could reduce heterogeneous protein distribution by inhibiting co-translational protein targeting. This research lays the framework for a better understanding of the detrimental impact of mitochondrial fragmentation on the physiology of cells in aging and disease.

Tatsuhisa Tsuboi

Dr. Tatsuhisa Tsuboi is a cell biologist at the Tsinghua Shenzhen International Graduate School. He received his Ph.D. degree from the Tohoku University, Japan, in 2014. He studied quality control mechanism of translational regulation under the supervision of Dr. Toshifumi Inada. From 2014, he started his postdoctoral training in cell biology at the University of California, Irvine, with Dr. Susanne Rafelski and from 2017, at the University of California, San Diego, with Dr. Brian Zid. He had a unique training opportunity in Robert Singer's lab at the Albert Einstein School of Medicine, as a visiting researcher. Since 2021 he has been a principal investigator at the Institute of Biopharmaceutical and Health Engineering, Tsinghua-SIGS. His main research interests are in quantitative cell biology and microscopic technologies. His current research focuses on analysis of translational regulation and developing computer vision techniques for analyzing and understanding biological microscopy images and on using these methods for understanding the molecular and systems mechanisms of basic cellular processes such as dynamic organization of intracellular organelle networks.

Dissecting Cell-State Attractors and Transitions in Single-Cell Transcriptome Data

Peijie Zhou

Center for Machine Learning Research, Peking University, China

Email: peijiezhou@pku.edu.cn

Single-cell sequencing technologies provide unprecedented resolution for studying the dynamic process of cell-state transitions during development and complex disease. However, analyzing high-dimensional, static single-cell RNA-sequencing (scRNA-seq) data with dynamical systems models can be challenging due to the curse of dimensionality. In this talk, I will discuss how machine learning has enabled us to overcome this challenge and use dynamical systems techniques to analyze scRNA-seq data. I will introduce the low-dimensional dynamical manifold to identify attractor basins and transition probabilities in snapshot data. I will also present the usage of non-equilibrium dynamical systems theory to analyze attractor stability and identify transition-driving genes in gene expression and splicing processes. Overall, these approaches contribute to bridge the model-based and data-driven methods in the rationale analysis of single-cell biology.

Peijie Zhou

Peijie Zhou is a tenure-track Assistant Professor at the Center for Machine Learning Research at Peking University. He completed his B.S. and Ph.D. degrees in computational mathematics at Peking University in 2014 and 2019, respectively. From 2020 to 2023, he served as a Visiting Assistant Professor under the supervision of Professor Qing Nie in the Department of Mathematics at the University of California, Irvine. His research is anchored in computational systems biology, with recent interest in exploring single-cell data-driven modeling and computation of complex biological systems, merging machine learning methodologies with multiscale dynamical system approaches.

Single-Cell Spatial Modeling of Cell Identity

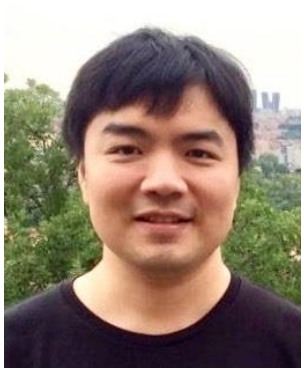
Jian Ma

Computational Biology, School of Computer Science, Carnegie Mellon University, USA

Email: jianma@cs.cmu.edu

Spatial transcriptomics can reveal spatially resolved gene expression of diverse cells in complex tissues. However, the development of computational methods that can use the unique properties of spatial transcriptome data to unveil cell identities remains a challenge. We recently developed SPICEMIX, an interpretable method based on probabilistic, latent variable modeling for joint analysis of spatial information and gene expression from spatial transcriptome data. Our findings indicate that SPICEMIX improves the detection of intricate cell identities, delineates interpretable spatial metagenes, and reveals differentiation trajectories. An enhanced version of SPICEMIX has also been recently developed to consider cohesively multiple samples. Collectively, these frameworks offer generalizable tools for analyzing spatial transcriptome data, enabling the exploration of cell-type composition and the spatial arrangement of cells within complex tissues.

Jian Ma



Jian Ma is the Ray and Stephanie Lane Professor of Computational Biology at the School of Computer Science at Carnegie Mellon University. His lab focuses on developing computational methods to study the structure and function of the human genome and cellular organization and their implications for evolution, health and disease. His group has recently pioneered a series of new machine learning methods for 3D epigenomics, comparative genomics, spatial genomics, and single-cell analysis. He currently leads a multi-disciplinary NIH Center, as part of the NIH 4D Nucleome Program. His recent work has been supported by NIH, NSF, CZI, Google, and the Mark Foundation. He has received several awards, including an NSF CAREER Award and a Guggenheim Fellowship (in Computer Science), and is an elected Fellow of AAAS.

Deciphering Spatial Regulation of Enhancer Networks Driving Cell Fate Determination at Single-Cell Level

Jialiang Huang

School of Life Sciences, Xiamen University, Xiamen, Fujian, China

Email: jhuang@xmu.edu.cn

Many enhancers exist as clusters in the genome and control cell identity and disease genes; however, the underlying mechanism remains largely unknown. The emergence of single-cell and spatial multi-omics has enabled mapping gene regulation in the tissue context at high resolution. Here, we first propose a model of enhancer networks through single-cell multi-omics analysis. We find cell identity and disease genes tend to be regulated by complex enhancer networks, where hub enhancers are the most functionally important. Next, we introduce Pesca, a computational method that predicts and enhances spatial chromatin accessibility profiling, hence enables deciphering spatial regulation of enhancer networks. The accuracy of Pesca is validated using various ground-truth datasets including the imaged H3K27ac MERFISH, in situ hybridization (ISH) data and enhancer activity by reporter assay in VISTA. Last, we conduct transgenic reporter assays and in vivo CRISPR/Cas9-mediated perturbation experiments, thereby confirming the spatiotemporal regulation of the Atoh1 enhancer network in mouse embryonic spinal cord and brain. Taken together, our work facilitates investigations of spatial regulation of enhancer networks in control of gene expression, thus cell fate decisions during development and diseases.

Jialiang Huang

Dr. Jialiang Huang is a Professor at School of Life Sciences, Xiamen University since 2018. He obtained Ph.D. in Bioinformatics at Chinese Academy of Science (2012), following by postdoc training at Harvard Medical School (2013-2018). His lab aims to understand enhancer function and regulation in development and disease, particularly in leukemia, using the cutting-edge and interdisciplinary technologies, such as bioinformatics, single-cell multi-omics, and CRISPR/Cas9 genome-editing assays.

Making the Gene Regulatory Networks More Logical

Zhiyuan Li

Center for Quantitative Biology, Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, China

Email: zhiyuanli@pku.edu.cn

A longstanding question in the field of gene regulatory networks is how these tangled interactions synergistically contribute to decision-making processes. For example, when multiple TFs co-regulate one target gene, should their combinatory regulatory logics be AND or OR, or something in between? To comprehensively understand the role of regulatory logic in cell fate decisions, we constructed a logic-incorporated GRN model and examined its behavior under two distinct driving forces (noise-driven and signal-driven). We distilled the opposite cell fate bias under different regulatory logics, uncovered distinctive trajectories of reprogramming influenced by logic motifs, and bridged regulatory logic and progression-accuracy trade-offs. Inspired by our successful logic-to-fate mappings in biological instances like hematopoiesis and embryogenesis, we developed machine learning algorithm to infer the combinatory logics and driving forces from single-cell omics data. Our work presents a logic-focused approach to the taxonomy of cell fate decisions.

Zhiyuan Li

Zhiyuan Li is an Assistant Professor at the Center for Quantitative Biology at Peking University. Her academic journey began at the School of Physics, Peking University where she completed her undergraduate studies. She later pursued graduate studies at the Department of Biophysics at UCSF and conducted postdoctoral research at Princeton.

Zhiyuan Lab concentrates on self-organization in biological systems (http://cqb.pku.edu.cn/zyli/). The team uses a variety of computational methods including data mining and network dynamics to explore two main areas: 1. Quantitative microbial ecology, specifically the evolution and ecology mediated by secondary metabolites; 2. The study of collective cell fate decisions and pattern formations.

Dynamical Data Science and AI for Biology

Luonan Chen

Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences (CAS), China

Email: lnchen@sibcb.ac.cn

I will present a new concept "dynamical systems biology" for quantifying dynamical processes, disease progressions, dynamical causality and various phenotypes, including dynamic network biomarkers (DNB) for early-warning signals of critical transitions, spatial-temporal information (STI) transformation for short-term time-series prediction, and partial cross-mapping (PCM) for causal inference among variables. These methods are all data-driven or model-free approaches but based on the theoretical frameworks of nonlinear dynamics. We show the principles and advantages of dynamics-based data-driven approaches for phenotype quantification as explicable, quantifiable, and generalizable. In particular, dynamics-based data science approaches exploit the essential features of dynamical systems in terms of data, e.g. strong fluctuations near a bifurcation point, low-dimensionality of a center manifold or an attractor, and phase-space reconstruction from a single variable by delay embedding theorem, and thus are able to provide different or additional information to the traditional approaches, i.e. statistics-based data science approaches. The dynamical-based data science approaches for the quantifications of various phenotypes will further play an important role in the systematical research of various fields in biology and medicine.

Luonan Chen

Luonan Chen received BS degree in the Electrical Engineering, from Huazhong University of Science and Technology, and the M.E. and Ph.D. degrees in the electrical engineering, from Tohoku University, Sendai, Japan, in 1988 and 1991, respectively. From 1997, he was an associate professor of the Osaka Sangyo University, Osaka, Japan, and then a full Professor. Since 2010, he has been a professor and executive director at Key Laboratory of Systems Biology, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences. He was elected as the founding president of Computational Systems Biology Society of OR China, and Chair of Technical Committee of Systems Biology at IEEE SMC Society. In recent years, he published over 400 journal papers and three monographs (books) in the area of bioinformatics, nonlinear dynamics and machine learning.

Construction of Solution Landscapes of Complicated Biological Systems

Lei Zhang

Beijing International Center for Mathematical Research (BICMR)

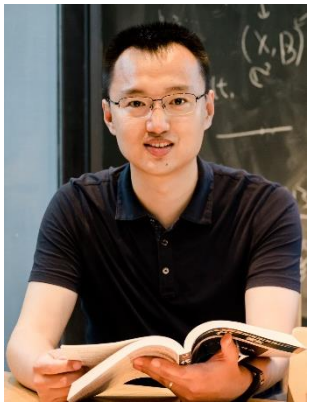
Center for Quantitative Biology (CQB)

Center for Machine Learning Research (CMLR), Peking University, China

Email: pkuzhangl@pku.edu.cn

Energy landscape has been widely applied to many physical and biological systems. A long standing problem in computational physics is how to search for the entire family tree of possible stationary states on the energy landscape without unwanted random guesses? Here we introduce a novel concept “Solution Landscape”, which is a pathway map consisting of all stationary points and their connections. We develop a generic and efficient saddle dynamics method to construct the solution landscape, which not only identifies all possible minima, but also advances our understanding of how a complex system moves on the energy landscape. As illustrations, we apply the solution landscape approach to study two problems: One is construction of the solution landscapes of gene regulatory networks in cell fate decisions, and the other one is to construct the solution landscape of reaction-diffusion systems, which reveals a nonlinear mechanism for pattern formation beyond Turing instability..

Lei Zhang



Lei Zhang is Boya Distinguished Professor at Beijing International Center for Mathematical Research, Peking University. He is also a Principle Investigator at Center for Quantitative Biology, Center for Machine Learning Research. He obtained his Ph.D in Mathematics at Penn State University in 2009. His research is in the area of computational and applied mathematics and interdisciplinary science in biology, materials, and machine learning. He has published the papers in Phys. Rev. Lett., PNAS, Acta Numerica, Science journals, Cell journals, SIAM journals. He was awarded/funded by NSFC Innovation Research Group, NSFC Outstanding Youth Award, National Key Research and Development Program of China, NSFC Excellent Youth Award, Royal Society Newton Advanced Fellowship, etc. He serves as an Associate Editor for SIAM J. Appl. Math, Science China Mathematics, CSIAM Trans. Appl. Math, DCDS-B, The Innovation, and Mathematica Numerica Sinica.

Flanking Sequence Engineering for Efficient Promoter Design

Xiaowo Wang

Department of Automation, Center for Synthetic and Systems Biology, Bioinformatics Div. BNRist, Tsinghua University, China

Email: xwwang@tsinghua.edu.cn

Designing promoters with desirable properties is essential in synthetic biology. Human experts are skilled at identifying strong explicit patterns in small samples, while deep learning models excel at detecting implicit weak patterns in large datasets. Biologists have described the sequence patterns of promoters via transcription factor binding sites (TFBSs). However, the flanking sequences of cis-regulatory elements, have long been overlooked and often arbitrarily decided in promoter design. To address this limitation, we introduced an AI-aided framework that efficiently designs synthetic promoters by combining expert knowledge with deep learning techniques. We demonstrated the power of this framework in improving the properties of *Escherichia coli* constitutive, IPTG-inducible, and mammalian cell doxycycline (Dox)-inducible promoters. Our work provided new insights into de novo gene regulatory element design, indicating the potential ability of AI to obtain new optimized genetic elements.

Xiaowo Wang

Dr. Xiaowo Wang is a Professor of bioinformatics at the Department of Automation Tsinghua University. He received his Bachelor's degree of Engineering and Ph.D. in Bioinformatics from Tsinghua University. He joined the faculty of Tsinghua University since 2008. He was a visiting student in Cold Spring Harbor laboratory in 2007-2008, and a Tang Distinguished Scholar in Synthetic Biology Institute of UC Berkeley in 2012-2013. His lab aims to bring AI and biology approaches together to understand gene regulation networks quantitatively and systematically, and guide the design of synthetic biological systems.

Understanding Cellular State Changes Through Causal Inference and Large Language Models

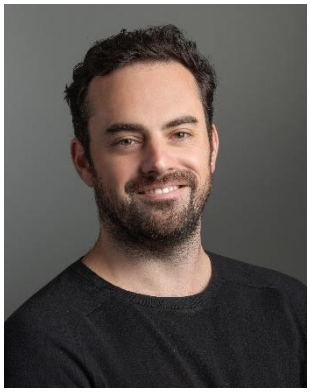
David van Dijk

Yale School of Medicine, Yale University, USA

Email: david.vandijk@yale.edu

The emergence of single-cell technologies has catalyzed a deeper understanding of cellular heterogeneity. At the van Dijk Lab, we are pioneering innovative computational frameworks to decode the complexities within cellular landscapes. This talk will introduce novel algorithms crafted to elucidate cellular state transitions. Cell2Sentence employs large language models, offering a natural language interface to biology, while CINEMA-OT harnesses causal inference in single-cell data, providing a pathway towards profound insights into cellular responses to perturbations. I will explain the core algorithms underpinning these methodologies and demonstrate their application in unveiling new biological phenomena.

David van Dijk



David completed his PhD at the University of Amsterdam and the Weizmann Institute of Science in Computer Science where he used machine learning to understand how gene regulation is encoded in DNA sequence. As a postdoctoral fellow at Yale Computer Science and Yale Medical School, he developed new machine learning and manifold learning methods for discovering hidden signatures in large biomedical data with an emphasis on single-cell data. David is currently an Assistant Professor in Medicine and in Computer Science at Yale, where he leads a research group in machine learning for biomedicine.

Cell Fate Decision by A Morphogen-Transcription Factor-Chromatin Modifier Axis

Duanqing Pei

School of Life Sciences, Westlake University, Hangzhou, China

Email: peiduanqing@westlake.edu.cn

Cell fate decision remains poorly understood at the molecular level. Embryogenesis provides a unique opportunity to analyse molecular details associated cell fate decisions. Works based on model organisms have provided a conceptual framework of genes that specify cell fate control, for example, transcription factors or TFs controlling processes from pluripotency to immunity. How TFs specify cell fate remains unknown. Here we report that Sall4 relies on NuRD to interpret BMP4 signal to decide cell fates in a well controlled system in vitro. While NuRD complex cooperates with SALL4 to convert mouse embryonic fibroblast or MEFs to pluripotency, BMP4 diverts the same process to an alternative fate, PrE or primitive endoderm. Mechanistically, BMP4 signals the dissociation of Sall4 from NuRD physically to establish a gene regulatory network for PrE. Our results provide a conceptual framework to explore the rich landscapes of cell fate choices intrinsic to development in higher organisms involving morphogen-TF-chromatin modifier pathways.

Duanqing Pei

Dr. Duanqing Pei is Chair Professor of Regenerative Biology at School of Life Sciences, Westlake University. Dr. Pei received his PhD from the University of Pennsylvania in 1991; trained as a postdoctoral fellow at University of Michigan from 1991 to 1996. Dr. Pei began his independent research career in 1996 as Assistant Professor at the University of Minnesota School of Medicine, and was promoted to Associate Professor with tenure in 2002. Then, he joined immediately the Medical Faculty at Tsinghua University. He moved to the Guangzhou Institutes of Biomedicine and Health (GIBH) in 2004. While pursuing his scholarly work, he has been active in building various institutions and organizations including GIBH, GDL(Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Hong Kong Center for Regenerative Medicine and Health. Dr. Pei was elected as Associate Member of EMBO in 2018.

From DNA to Life: Decode the Noncoding Genome

Xiaohua Shen

School of Medicine, Tsinghua University, China

Email: xshen@tsinghua.edu.cn

Life starts with a fertilized egg, guided by its DNA for intricate four-dimensional development. Our genome is remarkably versatile, generating hundreds of unique cell types. Yet, the specific algorithm governing cell-fate specification remains elusive. Interestingly, ~98% of mammalian genomes comprise noncoding sequences, producing RNA transcripts without encoding protein, primarily residing in the nucleus, challenging the traditional view of protein synthesis. In this talk, I will discuss the evolving paradigm that has fundamentally transformed our understanding of gene regulation consensus.

Xiaohua Shen

致力于非编码核酸调控染色质折叠、转录和细胞核稳态的基层规律，从独特的视角揭示细胞命运决定的普适性规律。近年成果包括：揭示基因组折叠的基层规律, L1 和 B1/Alu 转座子重复序列是染色质大尺度三维结构形成的遗传分子基础；非编码 RNA 顺式调控邻近转录和染色质状态是基因表达调控的一种普遍模式；RNA 和 RNA 结合蛋白通过相分离反馈调控转录和染色质结构。

Gene Regulatory Networks Orchestrating Immune Cell Fate Dynamics

Harinder Singh

Center for Systems Immunology, Department of Immunology,

University of Pittsburgh, USA

Email: HARINDER@pitt.edu

I have had a longstanding interest in analyzing transcription factors that regulate cell fate dynamics within the immune system. As an HHMI Investigator at the University of Chicago, with a focus on transcription factors that regulate immunoglobulin gene transcription and recombination, my lab discovered that the Ets family member PU.1 was required for the development of multiple innate and adaptive immune cell lineages (Scott et al., Science 1994). This genetic and accompanying phenotypic analysis served as the launching point for a series of studies that systematically unraveled the developmental and molecular functions of PU.1 in the generation of macrophages, granulocytes as well as B-lymphocytes. Of particular note was our demonstration that PU.1 functioned in a cell intrinsic manner to regulate the developmental potential of multipotential myeloid-lymphoid progenitors and that its graded levels were used to orchestrate myeloid versus lymphoid cell fates (Scott et al., Immunity 1997; DeKoter and Singh, Science 2000). These analyses led us to propose that innate and adaptive immune cells arise from common hematopoietic progenitors and the latter cells have co-opted regulatory factors utilized by their innate counterparts, which have preceded them in evolutionary time (Glimcher and Singh, Cell, 1999; Spooner et al., Immunity 2009). Around the year 2000, I realized that we needed to experimentally and computationally analyze coherent networks of transcription factors (Gene Regulatory Networks) rather than simply focusing on activities of individual regulators. This led us to the first comprehensive regulatory model for B cell development (Medina et al., Dev Cell, 2004, Singh et al., PNAS 2005) and also a combined experimental and theoretical model for the regulation of macrophage versus neutrophil cell fate choice (Laslo et al., Cell, 2006). My interests in fusing experimental and computational approaches have been sustained by collaborations with theoretical biologists including, Aaron Dinner at the University of Chicago (Sciammas et al., Molecular Systems Biology, 2011). I will focus my presentation on our ongoing analysis of the regulatory network orchestrating a functionally critical bifurcation that B cells undertake in response to pathogen encounter or vaccination. The partitioning of cells along distinct trajectories dictates the quality and durability of antibody responses.

Harinder Singh

Harinder Singh obtained his PhD at Northwestern University working with Prof. Lawrence Dumas (1979-84) and was a Jane Coffin Childs Postdoctoral fellow at MIT mentored by the Nobel laureates, Professors Phillip Sharp and David Baltimore (1984-88). He was a member of the faculty of the University of Chicago and an HHMI Investigator from 1989-2009 becoming a Louis Block Professor of Molecular Genetics and Cell Biology. While at the University of Chicago, he also served as Director of the graduate training program in Molecular and Cell Biology and also as the Chair of the Committee on Immunology. He joined Genentech in 2009 serving as Senior Director of Discovery Immunology and as a Staff Scientist. In 2013, he moved to the Cincinnati Children's Hospital Medical Center as Director of the Division of Immunobiology and the Center for Systems Immunology where he recruited and assembled a group of systems immunologists. He joined the University of Pittsburgh as Professor and Director of the Center for Systems Immunology in 2019. The newly established Center spans the faculties of Immunology and Computational and Systems Biology. Prof. Singh has served as an Editor of the Journal of Molecular and Cellular Biology. He has been a Chair of the Board of Scientific Counselors (Basic Research) of the National Cancer Institute and a member of an advisory group to the NIH Director Francis Collins to develop a 10-year framework for the NIH intramural research program. He is currently a member of the Scientific Advisory Board of the Allen Institute for Immunology.

Peri-Implantation Embryogenesis and Regulation

Qingyuan Zhu[1], Jitao Ge[1], Ying Liu[1], Fan Zhou[1*]

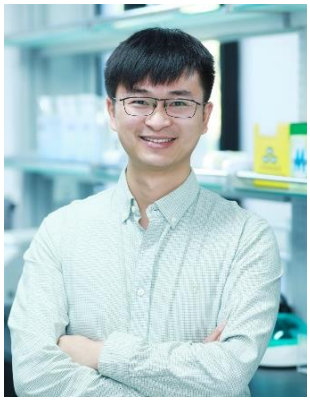
[1]Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing

Email: zhoufanlove@tsinghua.edu.cn

Connecting preceding blastocyst formation and following gastrulation respectively, peri-implantation embryogenesis is a key biological event during mammalian development. The embryo undergoes a series of cellular and molecular regulatory processes from pre- to post-implantation transition (PPT). In this presentation, we will discuss *in vitro* and *in vivo* models, omics measurement and molecular marker identification to explore the ingenious linkages among molecular program, lineage specialization, and polarity formation from a perspective of multidimensional molecular regulation. Relevant studies potentially provide clues to understand cell fate and regulation of embryo development, as well as the possible causes of habitual abortion and infertility.

Keywords: peri-implantation embryo; anterior-posterior axis; distal visceral endoderm (DVE); anterior visceral endoderm (AVE)

Fan Zhou



Dr. Fan Zhou is now an assistant professor and principal investigator at School of Life Sciences, Tsinghua University. Fan received his PhD from the Academy of Military Medical Sciences in 2016, and postdoctoral training at Peking University from 2016 to 2020. He joined Tsinghua as a faculty and set up his independent laboratory in August 2020. The current group now aims to integrate *in vivo* and *in vitro* functional identification, omics mining and genetic manipulation systems to study cell fate and peri-implantation embryo development. With newly-developed single-cell-initiated *in vivo* transplantation system, he revealed key signalling pathway during HSC emergence (*Nature*, 2016). Fan and his colleagues uncovered the gene networks and DNA methylome patterns of human implantation (*Nature*, 2019), the cellular and molecular dynamics of embryonic polarity formation across species (*Developmental Cell*, 2023), and the molecular characteristics of monkey blastoids (*Cell Stem Cell*, 2023). Fan has received the Ray Wu Prize (2016) and Young Elite Scientists Sponsorship Program from China Association for Science and Technology (2017). The research of human implantation was selected as the 2019 Top Ten Advances in Life Sciences in China.

Omics and AI Strategies to Study Splicing Defects in Rare Diseases

Yi Xing

Children's Hospital of Philadelphia, USA

Email: Yi.Xing@pennmedicine.upenn.edu

Genomic variants affecting pre-messenger RNA splicing and its regulation are known to underlie many rare genetic diseases. However, common workflows for genetic diagnosis and clinical variant interpretation frequently overlook splice-altering variants. To better serve patient populations and advance biomedical knowledge, it has become increasingly important to develop and refine approaches for detecting and interpreting pathogenic splicing variants. In this talk, I will summarize a few recent developments in using RNA sequencing technologies for rare disease investigation. Moreover, I will discuss how recent computational splicing prediction tools have emerged as complementary approaches for revealing disease-causing variants underlying splicing defects. We expect that continuous improvements to sequencing technologies and AI-based predictive modeling will not only expand our understanding of splicing regulation but also bring us closer to filling the diagnostic gap for rare disease patients.

Yi Xing

Dr. Yi Xing is the Director of the Center for Computational and Genomic Medicine and the Executive Director of the Department of Biomedical and Health Informatics at the Children's Hospital of Philadelphia. He is also a Professor of Pathology and Laboratory Medicine at the University of Pennsylvania. Dr. Xing holds the Francis West Lewis Chair in Computational and Genomic Medicine, an endowed chair that commemorates Dr. Francis West Lewis, the physician who founded the Children's Hospital of Philadelphia in 1855 as the first children's hospital in North America. His laboratory focuses on the computational biology and genomics of RNA processing and regulation, as well as their applications to human genetics and medicine. He has developed some of the foundational algorithms and several widely used tools for studying transcriptome variation using high-throughput technologies. His current research merges the fields of computational biology, RNA genomics, human genetics, precision medicine, and immuno-oncology.

Mapping Cell Landscapes at Single-Cell Level

Guoji Guo

Center for Stem Cell and Regenerative Medicine, Zhejiang University School of Medicine, Hangzhou 310058, China

Email: ggj@zju.edu.cn

Despite extensive efforts to sequence different genomes, genetic models to interpret regulatory sequences and cell fate decisions are lacking for most species. Here, we mapped single-cell landscapes covering different metazoan species using single-cell sequencing technologies. We determined the cell type composition and studied gene regulatory network across species. We developed deep learning–based models to predict landscapes and decipher regulatory sequences at the single-cell level. We have also built frameworks to decode genetic variation of regulatory sequences at cell type and single-nucleotide resolutions by integrating deep-learning-based variant predictions with population-based association analyses. Our work provides valuable resources for studying regulatory language in diverse biological systems.

Guoji Guo

Dr. Guoji Guo is Qiushi Distinguished Professor at Zhejiang University School of Medicine. He is also the deputy director for Institute of hematology and the deputy chair of Stem Cell Society at Zhejiang University. He obtained his Ph.D at National University of Singapore in 2010, and then moved to Harvard Medical School for postdoc training. In 2014, he was tenured at Zhejiang University. Professor Guoji Guo is interested in developing new single cell analysis technologies and applying these technologies to investigate complex biological systems.

System Methods to Define Master Regulator of Cell Identity in Development and Diseases

Kaifu Chen

Boston Children's Hospital, Harvard Medical School, Harvard University, USA

Email: Kaifu.Chen@childrens.harvard.edu

A cell type is determined by the network function of its cell identity genes, including master transcription factors that govern the expression status of this network. Cell identity transition is fundamental in normal differentiation and development, whereas a cell that loses its normal identity may cause disease including cancers. Targeting the driver genes for an abnormal cell identity holds great promise for new therapy. However, our understanding of cell identity regulators is incomplete. Integrating over >10,000 genomic and epigenomic profiles, we uncovered that cell identity genes as a unique group are distinct from other genes in the mechanisms to regulate their expression. These discoveries laid the foundation for us to develop novel machine learning techniques, which utilize expression regulation mechanisms for systematic identification of driver genes for normal cell differentiation and tumor development. These driver genes will lead to new therapeutic targets and diagnostic markers, as successfully verified in cell, xenograft, and PDX model for cancers, and thus, will benefit numerous patients.

Kaifu Chen

Dr. Chen is a computational biologist interested in computational modeling of cell identity regulation. He received PhD degree from the renowned Beijing Institute of Genomics. Thereafter, he joined the Dan L Duncan Cancer Center at Baylor College of Medicine as a postdoctoral fellow. Dr. Chen started his bioinformatics lab as an Assistant Professor in the Cornell University Weill Cornell Medical College at Houston Methodist Hospital Research Institute, where he further became an Associate Professor and was designated as the founding Director to develop a Center for Bioinformatics and Computational Biology. Dr. Chen is currently an Associate Professor in the Harvard Medical School and Director of Computational Biology Program at Boston Children's Hospital.

Infer Cell Dynamics and Clonal Memory from Single-Cell Lineage Tracing

Shou-Wen Wang[1,2,3]

[1]*Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China*

[2]*School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China*

[3]*School of Science, Westlake University, Hangzhou, Zhejiang 310024, China*

Email: wangshouwen@westlake.edu.cn

A goal of single cell genome-wide profiling is to reconstruct dynamic transitions during cell differentiation, disease onset, and drug response. Single cell assays have recently been integrated with lineage tracing, a set of methods that identify cells of common ancestry to establish bona fide dynamic relationships between cell states. These integrated methods have revealed unappreciated cell dynamics, but their analysis faces recurrent challenges arising from noisy, dispersed lineage data. Here, we develop coherent, sparse optimization (CoSpar) as a robust computational approach to infer cell dynamics from single-cell transcriptomics integrated with lineage tracing. Built on assumptions of coherence and sparsity of transition maps, CoSpar is robust to severe down-sampling and dispersion of lineage data, which enables simpler experimental designs and requires less calibration. In datasets representing hematopoiesis, reprogramming, and directed differentiation, CoSpar identifies early fate biases not previously detected, predicting transcription factors and receptors implicated in fate choice. Documentation and detailed examples for common experimental designs are available at https://cospar.readthedocs.io/.

I will also talk about our recent progress with single-cell multi-omic lineage tracing. Using this new tool, we revealed that cells within the same clone share more similar profile of DNA methylation, rather than gene expression or chromatin accessibility.

Reference

[1] L. Li, S. Bowling, Q. Yu, S.E. McGeary, K. Alcedo, B. Lemke, M. Ferreira, A. M. Klein, **S.-W. Wang**#, F.D. Camargo#, A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells, Cell (2023, in press)

[2] **S.-W. Wang**#, M. J. Herriges, K. Hurley, D. N. Kotton, A. M. Klein#, CoSpar identifies early cell fate biases from single-cell transcriptomic and lineage information, Nat. Biotechnol. (2022).

#: correspondence

Shou-Wen Wang

Dr. Shou-Wen Wang received his bachelor's degree in Engineering Physics from Tsinghua University in 2013. He went on to obtain a PhD in Physics from Tsinghua University in 2018. During 2018-2022, Dr. Wang worked as a postdoctoral researcher in the Department of Systems Biology at Harvard Medical School, where he developed methods for analyzing single-cell lineage tracing and multi-omic data to better understand the fundamental principles of cell differentiation and embryonic development. He was awarded the Quantitative Biology Award from the Damon Runyon Cancer Research Foundation during his postdoc. In 2023, he joined Westlake University as an assistant professor in the School of Life Sciences, with a joint appointment in the physics department at the School of Science.

Regulatory Mechanism of Cell Lineage Segregation in Early Human Embryogenesis

Ge Guo

The Living System Institute, University of Exeter, UK

Email: g.guo@exeter.ac.uk

Understanding how cell fates are specified and spatially organised to form a human embryo is a fascinating fundamental biology question. Studying human embryogenesis is inherently challenging, and significant efforts have been made to develop alternative stem cell-based models. We have established human naïve pluripotent stem cells (PSCs), which resemble the naïve epiblast in the blastocyst. These human naïve PSCs retain extraembryonic lineage potency, and they can self-organize into blastoids that closely resemble blastocysts in cellular composition, topology, and regulation of lineage specification. Therefore, human naïve PSCs and the derivative blastocyst model provide a valuable experimental system to study cell fate specification in early embryogenesis. In this talk, I will discuss our current research aimed at understanding the regulatory mechanism of the sequential trophectoderm and hypoblast lineage specification in human naïve PSCs, blastoids, and natural embryos.

Ge Guo

Dr. Ge Guo is a stem cell biologist. She obtained her PhD at the University of Cambridge with Prof. Allan Bradley. During her PhD, she developed approaches for recessive genetic screens using mouse embryonic stem cells. In 2006 she joined Prof. Austin Smith's laboratory and was awarded an MRC stem cell career development fellowship to investigate novel genes regulating mouse pluripotency. She then became intrigued by the difference between mouse and human pluripotent stem cell states and redirected her research towards human naïve pluripotency. She is one of the pioneers in the establishment and characterisation of human naïve pluripotent stem cells. In 2020, Ge joined the Living Systems Institute, University of Exeter, as a Principal Investigator. In her recent research, Ge discovered the trophectoderm differentiation potential of human naïve pluripotent stem cells, which led to the establishment of integrated human embryo models. Her current research is focused on mammalian pluripotent stem cells and the regulatory mechanisms of cell fate transition during early embryo development.

Statistical Learning of Multicellular Dynamics

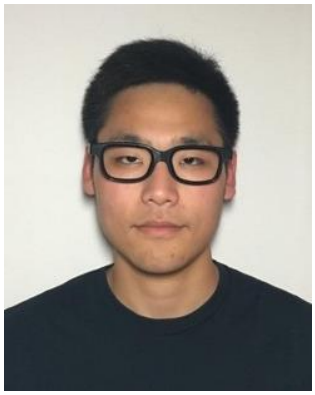
Ming Han

Center for Quantitative Biology, Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, China

Email: hanmingcr@pku.edu.cn

Cells in a tissue often form a disordered network. The geometry and topology of this network mediate complex intercellular interactions, which strongly affect the migration and division of tissue cells. Such multicellular dynamics play a crucial role in many biological processes ranging from wound healing to morphogenesis. However, given the complexity of the internal drive of individual cells and their interactions, it is extremely difficult to establish a theoretical model from the first principle. In this talk, I will introduce a generic machine-learning approach capable of learning various multicellular dynamics from recorded experimental videos. Instead of requiring the internal states of each living cell that are hard to access, our model relies solely on external geometric information such as cell shape, size, and interconnectivity that are easier to measure. To machine learn the cell interactions that can be both non-reciprocal and pathway-dependent, we represent a tissue system in terms of both cell and interaction graphs and apply advanced neural networks onto this dual graph representation. Taking epithelial tissues as an illustrative example, we show that our graph neural network not only captures the stochastic cell motion but also predicts the evolution of cell states in their division cycle. We demonstrate with experimental data that this method can be easily extended to forecast developmental systems, such as the fly wing, and cell signaling processes.

Ming Han



Ming Han is an assistant professor in the Center for Quantitative Biology at Peking University. He obtained his bachelor's degree in physics and applied math at Shanghai Jiao Tong University. Then he pursued his PhD in applied physics at Northwestern, working with Erik Luijten on the study of active and bio-inspired materials through the lens of molecular dynamics simulations. Before joining PKU, he was a Kadanoff-Rice postdoctoral fellow at the University of Chicago, where he collaborated with both Vincenzo Vitelli and Juan J. de Pablo on the fluctuating hydrodynamic theory of active fluids as well as the machine learning methodologies for living many-body systems especially multicellular tissues. At PKU, his research centers around the study of multi-component biological systems, in particular how their spatiotemporal behaviors and the underlying interactions among different components.

Target and Drug Discovery Based on Covalent Interaction

Tong Zhu[1,2*]

[1]*School of Chemistry and Molecular Engineering, NYU-ECNU Joint Research Institute of Computational Chemistry, Shanghai, 200062, China*

[2]*Shanghai Institute of Algorithmic Innovation, Shanghai, 200062, China*

Email: tzhu@lps.ecnu.edu.cn

Target discovery is a costly and low-success-rate process. In recent years, the activity-based protein profiling method has been increasingly validated as an efficient tool for discovering new targets. It not only significantly shortens the target discovery cycle but also provides starting points for drug discovery research. However, a major challenge in this technique is the lack of active molecules. To address this key scientific issue, we have combined computational chemistry, artificial intelligence, and organic synthesis to construct a covalent molecular library with diverse stereostructures, solving the long-standing selectivity problem of covalent active molecules. Utilizing these molecules, we elucidated the key apoptotic signaling pathways in multiple subtypes of osteosarcoma with unknown pathogenesis, identified and validated potential new drug targets. Building upon this foundation, we have also discovered novel candidate drug molecules targeting osteosarcoma.

Tong Zhu

Tong Zhu is a professor at the School of Chemistry and Molecular Engineering, East China Normal University, and a researcher at the Shanghai Institute of Algorithmic Innovation. He was awarded the Excellent Youth Foundation in 2022 and serves as a member of the Biomedical Informatics and Drug Discovery Committee. His research primarily focuses on the simulation of chemical reaction kinetics in biological systems, including enzyme catalytic mechanisms and covalent interaction-based drug and target discovery. His research achievements have been published in high-impact journals such as Nat. Mach. Intell., Nat. Commun., Nucleic Acids Res., and he holds two international patents for new drugs, as well as one FDA clinical trial approval.

AI and Aging

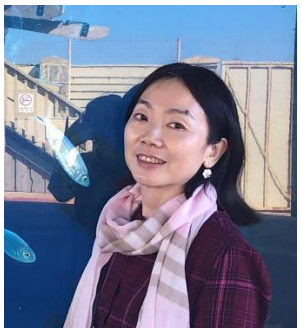
Jing-Dong Jackie Han

[1]*Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Center for Quantitative Biology (CQB), Peking University, Beijing 100871, China*

Email: jackie.han@pku.edu.cn

Aging is a systems level process and needs systems level models to quantify. We have recently focused our attention on phenotypic images and single cell clocks using a combination of experimental and computational approaches, most recently artificial intelligence (AI). Our deep learning AI models trained on either chronological age or perceived age of the 3D facial images can precisely estimate individuals' aging status, and infer the molecular regulators mediating the impact of lifestyles (Xia et al., 2020). Further analysis of human aging related lncRNAs find that they are preferentially involved in senescence associated secretory phenotype and inflammation (Cai and Han, 2021). We also found the highly abundant lncRNA KCNQ1OT1 through forming RNA-DNA Triplex targets and represses the evolutionarily young transposon elements in a sequence specific manner, thus guards the cells against genome instability and cellular senescence (Zhang et al., 2022). I will also discuss our recent results using AI to decipher aging and disease status (Zhu et al. 2023), and our recent Transformer-based single cell annotation tool, TOSICA and its application to aging clocks (Chen et al., 2023).

Jing-Dong Jackie Han



Prof. Jing-Dong Jackie Han obtained Ph.D. degree from Albert Einstein College of Medicine. She had her postdoctoral training at The Rockefeller University and Dana-Farber Cancer Institute. In 2004, she became an investigator/professor at the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences. In 2010-2019, she was a director of the CAS-Max Planck Partner Institute for Computational Biology. In 2019, she became Boya professor at Peking University. Her research focuses on the structure and dynamic inference of molecular networks, using a combination of large-scale experiments and computational analysis to explore the design principles of the networks and to find how the complex phenotypes, in particular aging and stem cell development are regulated through molecular networks. She was awarded the NSFC Outstanding Young Scientist Award in 2006, and the Hundred Talent Plan Outstanding Achievement Award in 2009, selected as a Max Planck Follow in 2011 and a MaxNetAging Fellow in 2014, F1000 faculty in developmental biology in 2016.

Efficient computation by molecular competition networks

Haoxiao Cai[†], Lei Wei[†,*], Xiaoran Zhang, Rong Qiao, Xiaowo Wang[*]
Ministry of Education Key Laboratory of Bioinformatics,
Center for Synthetic and Systems Biology,
Beijing National Research Center for Information Science and Technology,
Department of Automation, Tsinghua University, Beijing 100084, China

Abstract

Most biomolecular systems exhibit computation abilities, which are often achieved through complex networks such as signal transduction networks. Particularly, molecular competition in these networks can introduce crosstalk and serve as a hidden layer for cellular information processing. Despite the increasing evidence of competition contributing to efficient cellular computation, how this occurs and the extent of computational capacity it confers remain elusive. In this study, we introduced a mathematical model for Molecular Competition Networks (MCNs) and employed a machine learning-based optimization method to explore their computational capacity. Our findings revealed that MCNs, when compared to their non-competitive counterparts, demonstrate superior performance in both discrete decision-making and analog computation tasks. Furthermore, we examined how real biological constraints influence the computational capacity of MCNs, and highlighted the nonnegligible role of weak interactions. The study suggested the potential of MCNs as efficient computational structures in both *in vivo* and *in silico* scenarios, providing new insights into the understanding and application of cellular information processing.

[*]These authors contributed equally to this work.
[†]Corresponding author.
Email: chx21@mails.tsinghua.edu.cn

Discovery of NLRP3 inhibitors using machine learning: Identification of a hit compound to treat NLRP3 activation-driven diseases

Cheng Shi[1] , Xiangyu Zhang[1], Xiaowei Chi[1], Yi Ran zhou[a], Weiping Lyu[a], Tongfei Gao[a], Jiaxu Zhou[a], Yanming Chen[a], Chuxiao Yi[a], Xiaojiao Sun[a],*, Liangren Zhang[a,] * , Zhenming Liu[a,]*

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing, 100191, PR China

Abstract

NLRP3 is vital in developing many human diseases as one of the most critical inflammasomes. Developing related inhibitors has been instrumental in advancing the development of therapies for associated diseases. To date, there are no NLRP3 inhibitors on the market. This study identified a series of NLRP3 inhibitors using the self-developed machine learning model. Among them, CSC-6 was validated as the hit molecule with optimal activity and significantly inhibited IL-1β secreted by PMA-THP-1 cells (IC_{50} = 2.3 ±0.38 μM). The results show that CSC-6 specifically binds NLRP3 and inhibits NLRP3 activation by blocking ASC oligomerization during NLRP3 assembly. In vivo experiments have demonstrated that CSC-6 effectively reduces the symptoms of NLRP3 overactivation-mediated sepsis and Gout in mouse models. Importantly, CSC-6 has lower cytotoxicity and exhibits better stability in human-derived liver microsomes, which is more favorable for the drug to maintain its efficacy in vivo for longer. The discovery of CSC-6 may contribute to the design and discovery of related NLRP3 inhibitors.

Keywords: NLRP3 Inhibitors, Drug discovery, Machine learning, Inflammasome-driven diseases, Gout

References

[1] Kim J.Y., Hwang M., Choi N.Y., et al. Inhibition of the NLRP3 Inflammasome Activation/Assembly through the Activation of the PI3K Pathway by Naloxone Protects Neural Stem Cells from Ischemic Condition. Mol. Neurobiol., 2023, 60: 5330-5342.

[2] Fullard N.; O'Reilly, S., Role of innate immune system in systemic sclerosis. Semin. Immunopathol., 2015, 37 (5): 511-517.

[3] Sharif H., Wang L., Wang W. L., et al. Structural mechanism for NEK7-licensed activation of NLRP3 inflammasome. Nature., 2019, 570 (7761): 338-343.

[4] Wang X.Q., Jiang W., Yan Y.Q.; et al. RNA viruses promote activation of the NLRP3 inflammasome through a RIP1-RIP3-DRP1 signaling pathway. Nat. Immunol., 2014, 15 (12): 1126-1133.

Email: shicheng_pku@163.com

Inferring dynamical models from time-series data using an interpretable machine learning method based on weighted expression trees

Yu Zhou and Xiufen Zou
School of Mathematics and Statistics, Wuhan University,
Wuhan, Hubei, China

Abstract
The growing time-series data make it possible to glimpse the hidden dynamics in many different fields; however, developing a computational toolbox with high interpretability for unveiling the interaction dynamics from data remains a crucial challenge. Here, we propose a new computational approach called Automated Dynamical Model Inference based on Expression Tree (ADMIET), in which the machine learning algorithm, the numerical integration of ordinary differential equations and the interpretability from prior knowledge are embedded into the symbolic learning scheme to establish a general framework for revealing the hidden dynamics in time-series data. ADMIET takes full advantage of both machine learning algorithm and expression tree. We first translate the prior knowledge into constraints on the structure of expression tree, reducing the search space and improving the interpretability. Second, we use the proposed adaptive penalty function to ensure the convergence of gradient descent algorithm and the selection of the symbols. Compared to gene expression programming, the ADMIET exhibits its remarkable capability in function fitting with higher accuracy and broader applicability. We further apply ADMIET to two typical biological systems with different priori knowledge to infer the dynamical equations. The results indicate that ADMIET can not only discover the interaction relationships but also provide accurate estimate of the parameters in the equations. These results demonstrate the superiority of ADMIET in revealing interpretable dynamics from time-series data.

References

Email: zhouyu_math@whu.edu.cn

Principle and evolution of neuronal generation in the human hypothalamus

Zhen-Hua Chen

[1] *State Key Laboratory of Molecular Development Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, 100101, China.*

[2] *University of Chinese Academy of Sciences, Beijing 100101, China.*

Abstract

As an ancient brain region, the hypothalamus controls the body's instincts and homeostasis. In contrast to the laminar cortex, the hypothalamus has a more diffuse organization, with more than a dozen nuclei and 100 neuronal subtypes (*1*). Dysplasia of hypothalamic neurons can lead to a variety of disorders, including sleep, diet, energy metabolism, reproduction, and hormone secretion (*2*). However, the mechanisms of human hypothalamic neurogenesis, neuronal lineage establishment, neuronal fate determination and neuronal evolution remain poorly understood. Here, we collected the developmental hypothalamus of mouse, macaque and human and harvested over 350,000 cells, combining single-cell transcriptomics, single-cell lineage tracing and spatial transcriptomics data. **In the phase of pattern generation**, we found that conserved FOX family transcription factors and morphogen factors segment hypothalamic progenitor domain along the anterior-posterior axis. The interplay of these intrinsic and extrinsic factors shapes the heterogeneity of early hypothalamic progenitor domains. **During neurogenesis**, we developed a program of lineage inference to infer the potential hypothalamic lineage. We identified a series of conservative lineage-specific transcription factors guiding the generation of neuronal subtypes in different lineages across species. By integrating mouse Tbx3 lineage tracing and Tbx3 genetic knockout snRNA-seq data, we confirmed that hypothalamic lineage factors, such as TBX3, can influence the production of specific lineages of neurons. Additionally, we observed either functional diversification or redundancy in individual multigene families for lineage specification. **During the fate determination of neurons**, we identified some new lineage factors, such as TBX3, SIX6 and BARHL1, served as terminal selectors, determining and maintaining neuronal identity. In addition, we developed a cell type-specific regulatory network inference (CESNI) program and decoded the conserved neuronal subtype-specific gene regulatory networks across species. **By decoding the evolution path of neurons**, we found that the hypothalamus adopts multiple strategies such as changing the proportion of neuronal subtypes, redistribution of neurons, and recoupling expression of neurotransmitters and neuropeptides, which enables the human hypothalamus to develop more complex functions. **Together**, our study reveals a comprehensive view of hypothalamic neurogenesis, lineage construction, neuronal fate determination, and strategy of neuronal evolution.

References

1 Xie, Y. and R.I. Dorsky, Development of the hypothalamus: conservation, modification and innovation. Development, 2017. 144(9): p. 1588-1599.

2 Biran, J., et al., Role of developmental factors in hypothalamic function. Front Neuroanat, 2015. 9: p. 47.

Email: czh@genetics.ac.cn

NMR-based traits: The Next Generation of Precision Medicine

Weng Kung Peng[1*]

Songshan Lake Materials Laboratory, 523-808, Dongguan, China

Abstract

Genetic contribution to different diseases were found to be varies, with non-genetic factors (e.g., lifestyle, environmental hazards) having much greater attributable risks, producing a large phenotypic variation. As a result, the current one-size-fits-all medical practices are sub-optimal, leaving much room for improvement. Our research focuses on addressing the challenges (and unprecedented opportunities) by introducing a novel class of multiparametric time-domain NMR-based ´molecular signature´ of biological fluids (e.g., blood) with respect to its´ various pathophysiological states. Liquid biopsy holds great promises in clinical medicine as it provides multiple global snapshot information non-invasively for disease management (e.g., predictive treatment/ recurrence) in uniquely personalized manner. We demonstrated that highly unique and specific `molecular fingerprinting´ in single drop of blood can be rapidly typed for disease diagnosis (e.g., malaria, diabetes mellitus, hemoglobinopathies) using point-of-care NMR system.

Keywords: multidimensional NMR; manifold-based learning; genotypic-phenotypic relationship.

References

1. Peng, W.K. *et al. npj Science of Food* **6**:59 (2022); https://doi.org/10.1038/s41538-022-00173-z

2. Peng, W.K. *et al. Rev. Sci. Instrum.* **83**, 095115 (2012); doi:10.1063/1.4754296

3. Peng, W.K. *et al.*npj Aging and Mechanisms of Disease.6:11(2020);https://doi.org/10.1038/s415141--020-00049-0.

4. Peng, W.K. *et al. Communications Biology* **3**, 535 (2020); doi:10.1038/s42003-020-01262-z

5. Peng, W.K. *et al. Nature Medicine*, **20**, 1096 (2014); doi:10.1038/nm.3622

Email: pengwengkung@sslab.org.cn

Generative modelling of perturbation-induced development trajectories during directed differentiation

Weizhong Zheng, Junyi Chen, Mingze Gao,Yiming Chao, Ralf Jauch, Yuanhua Huang, Joshua Wing Kei Ho
School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong SAR, China

Abstract

Trajectory inference methods can reconstruct the cellular transition dynamics from thegene expression data of single-cell snapshots [1-3]. However, current approaches oftenoverlook the influence of extrinsic conditons on cell state transitions, limiting their applicability in studies involving complex designs such as genetic perturbations or drug treatments. Here we present *GenDiff,* a conditional generative model specifically designed for induced differentiation data. GenDiffeffectively incorporates the perturbation label to sample the condition-specific trajectories from the single-cell population. Subsequently, it learns to predict the short-term transcriptional changes occurring between any two neighboring cells along these trajectories. We demonstrated that our method outperforms other baselines in predicting cellular responses across various single-cell perturbation settings, including overexpression of Transcription Factors (TFs) and combinatorial drug treatments. During this process, *GenDiff* also learns a meaningful latent representation that can better discriminate cells based on the received perturbations. Moreover, *GenDiff* seamlessly integrates with other single-cell trajectory analysis tools [4], thereby facilitating the quantification and interpretation of perturbation effects on cell state transitions. To illustrate its capabilities, we applied *GenDiff* to model the TF Atlas [5], a huge single-cell dataset that captures the gene expression when overexpressing different TF isoforms, and identified novel TFs that can guide embryonic stem cells towards the neural lineage.

References

[1] La Manno, Gioele, et al. "RNA velocity of single cells." Nature 560.7719 (2018):494-498.
[2] Bergen, Volker, et al. "Generalizing RNA velocity to transient cell states through dynamical modeling." Nature biotechnology 38.12 (2020): 1408-1414.
[3] Cao, Junyue, et al. "The single-cell transcriptional landscape of mammalian organogenesis."Nature 566.7745 (2019): 496-502.
[4]Lange, Marius, et al. "CellRank for directed single-cell fate mapping." Nature methods 19.2(2022):159-170.
[5] Joung, Julia, et al."A transcription factor atlas of directed differentiation." CeIl 186.1 (2023):209-229.

Email: zhengwzh@connect.hku.hk, jwkho@hku.hk

Screening for biomarkers for progression from oral leukoplakia to oral squamous cell carcinoma and evaluation of diagnostic efficacy by twelve machine learning algorithms

Fengyang Jing, Heyu Zhang, Tiejun Li

Department of Oral Pathology, Peking University School and Hospital of Stomatology, Beijing 100871, China

Abstract

Objective: To identify key genes during the progression from oral leukoplakia (OL) to oral squamous cell carcinoma (OSCC) and predict effective diagnosis.

Methods: Weighted gene co-expression network analysis and differential expression analysis were used to screen for genes associated with progression from OL to OSCC. Twelve machine learning algorithms such as k-Nearest Neighbor (KNN), Neural Network (NNet), and extreme Gradient Boosting (XGBoost) were used to construct multi-gene models. The functional mechanism or the pathways associated with these genes were evaluated using enrichment analysis, subtype clustering. Finally, a nomogram and Kaplan-Meier survival analysis were used to predict the prognostic efficacy of key genes in OSCC patients.

Results: The WGCNA and differential expression analysis were performed to identify seven genes associated with progression from OL to OSCC. The analysis conducted using twelve machine learning algorithms revealed that each model had good diagnostic efficacy. The enrichment analysis revealed that the genes enriched were associated with the cell cycle and cell division. Further, these genes were also associated with the regulation of intracellular energy metabolism. The prognostic analysis showed that genes could predict the prognosis of the patients, and patients in the high-risk group had a poor prognosis.

Conclusion: Our study identified that a multi-gene model was associated with the progression from OL to OSCC. These genes had good diagnostic efficacy and could be used as potential biomarkers for the prognosis of OSCC patients.

References

Jing. F.; Zhang. J.; Cai. X.;Zhou. X.; Bai. J.; Zhang. H.; Li. T. Screening for Biomarkers for Progression from Oral Leukoplakia to Oral Squamous Cell Carcinoma and Evaluation of Diagnostic Efficacy by Multiple Machine Learning Algorithms. Cancers 2022, 14, 5808.

Email: jfyjfyjfy@bjmu.edu.cn

MSC-seq: a transcriptomic database of mesenchymal stem cells for human and mouse

Peibin Dai[1], Chenyu Yang[2], Ke Hu[3], Lu Xie[2], Zhengliang Gao[1,3]

1. *Yangzhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), Tongji University School of Medicine, Shanghai 201619, China*
2. *Shanghai-MOST Key Laboratory of Health and Disease Genomics, Shanghai Institute for Biomedical and Pharmaceutical Technologies, Shanghai 200037, China*
3. *Shanghai University School of Medicine, Shanghai 200444, China*

Abstract

Mesenchymal stem cells (MSCs) play a crucial role in the development, maintenance, function, and regeneration of majority tissues. Given their multi-lineage differentiation potential as well as distinct immunomodulatory and trophic effects, MSCs have been considered as a promising source for therapeutics in tissue repair and organ regeneration. Recent advances in high-throughput sequencing technologies offer an unprecedented opportunity to further unravel the heterogeneity of these cells. In this context, we systematically collected, curated and integrated 366 high-quality datasets of MSCs, finally developed MSC-seq (https://www.biosino.org/mscs/Index.php), the first transcriptomic database specifically focusing on MSCs research. MSC-seq provides online analysis modules for comparing developmental and functional differences among MSCs deriving from multiple tissues, different species and various donors. Moreover, MSC-seq presents a robust computational tool called De novo Screening and Scoring (DNSS), which not only allows identifying critical genes or pathways involving in cell fate transition, but also enables predicting and quantifying the potency or potential of a specific biological process such as adipogenesis, angiogenesis, immunomodulation and aging. Therefore, MSC-seq is a powerful platform that provides searching, visualizing, analyzing and downloading MSCs data online, and wound greatly help the understanding of the complex biology of MSCs.

References

Bernardo ME, Fibbe WE. Mesenchymal stromal cells: sensors and switchers of inflammation. *Cell Stem Cell.* 2013;13(4):392-402.

Lemos DR, Duffield JS. Tissue-resident mesenchymal stromal cells: Implications for tissue-specific antifibrotic therapies. *Sci Transl Med.* 2018;10(426):eaan5174.

Shi Y, Wang Y, Li Q, et al. Immunoregulatory mechanisms of mesenchymal stem and stromal cells in inflammatory diseases. *Nat Rev Nephrol.* 2018;14(8):493-507.

Levy O, Kuai R, Siren EMJ, et al. Shattering barriers toward clinically meaningful MSC therapies. *Sci Adv.* 2020;6(30):eaba6884.

Wang Y, Fang J, Liu B, Shao C, Shi Y. Reciprocal regulation of mesenchymal stem cells and immune responses. *Cell Stem Cell.* 2022;29(11):1515-1530.

Email: 2010588@tongji.edu.cn

Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma

Lingfei Wan[1], Xinlong Yan[1#]

[1] *College of Life Science and Bioengineering,Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China*

Abstract

Single-cell RNA sequencing (scRNA-seq) analysis emerged as a powerful tool for revealing cellular diversity and intercellular communication at single-cell resolution. Recently, scRNA seq has been applied to dissect the complex tumor and immune landscapes of several cancers, including glioblastoma, breast cancer, lung cancer, head and neck cancer, pancreatic ductal adenocarcinoma, and liver cancer. This technique has improved our understanding of cellular heterogeneity and facilitated the screening of promising molecular targets to guide anti-tumor therapies. However, tumor heterogeneity and the interplay between malignant cells and niche cells at single cell resolution in human ICC remains poorly understood.

In this study, we used a droplet-based scRNA-seq sequencing platform (10x Genomics) to profile single cells from human ICC and adjacent tissues. We identified high inter-tumor heterogeneity in ICC samples along with prominent immunosuppressive characteristics in CD4 regulatory T cells (Tregs). Furthermore, we defined 6 fibroblast subclusters in ICC and adjacent tissues, among which $CD146^{+}$ vCAFs comprised the majority of CAFs in ICC tissues and could significantly induce enhancer of zeste homolog 2 (EZH2) upregulation and enhance ICC malignancy via the interleukin (IL)-6/IL-6 receptor (IL-6R) axis. Furthermore, exosomal miR-9-5p derived from ICC cells strongly induced vCAFs to promote ICC progression. Together, our results provide a comprehensive transcriptomic overview and dissect the intercellular crosstalk between ICC cells and vCAFs, suggesting potential targets for ICC therapy.

References

Zhang, M., Yang, H., Wan, L., Wang, Z., Wang, H., Ge, C., Liu, Y., Hao, Y., Zhang, D., Shi, G., Gong, Y., Ni, Y., Wang, C., Zhang, Y., Xi, J., Wang, S., Shi, L., Zhang, L., Yue, W., Pei, X., … Yan, X. (2020). Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. Journal of hepatology, 73(5), 1118–1130.

Email: wan-lingfei@outlook.com
Yxlong2000@bjut.edu.cn

Development of multi-constraint cell models and their application in metabolic engineering

Zhitao Mao, Xue Yang, Xin Zhao, Jinhui Niu, Ke Wu and Hongwu Ma

Biodesign Center, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308 China

Abstract

Biomanufacturing, characterized by its green, low-carbon, and sustainable nature, holds significant potential for achieving strategic goals of "carbon neutrality" and for promoting the development of the bio-economy. Computation design of industrial strains capable of efficient bioconversion for the production of valuable chemicals holds the key for the success of biomanufacturing. In the last decades, genome-scale metabolic models (GEMs) have been employed to provide guidance for industrial strain development by identifying potential engineering targets using constraint optimization methods. However, the accuracy of GEM predictions is often limited by the lack of consideration of additional constraints beyond stoichiometric constraints, leading to unreliable predictions. As such, there is a need for cellular models that incorporate multiple levels of constraints to accurately simulate cellular behaviour.

This report outlines the progress made in constructing and analysing multi-constraint models (MCMs). A simplified Python-based workflow, ECMpy, was proposed and utilized to construct enzyme-constrained models for several model organisms, namely *Escherichia coli* (eciML1515), *Corynebacterium glutamicum* (ecCGL1), and *Bacillus subtilis* (ecBSU1). The enhanced ECMpy workflow entailed automatic acquisition of *kcat* values from databases and filling of missing values using AutoPACMEN. It also achieved semi-automated correction of gene-protein-reaction (GPR) relationships by leveraging the GPRuler tool and protein homology similarity, as well as automated acquisition of quantitative subunit composition data from UniProt. Novel algorithms were also developed to identify target genes that enhance the yield of commodity chemicals. Most results were consistent with experimental data, while others showed promise as potential novel targets for metabolic engineering.

Additionally, a novel method was proposed that integrates both enzymatic and thermodynamic constraints within a single Pyomo modeling framework (ETGEMs). Using this framework, the EcoETM (*E. coli* metabolic model with enzymatic and thermodynamic constraints) was constructed, and the optimal pathways for cellular growth and production of 22 metabolites were calculated. Comparing the results with the stoichiometric model iML1515 and models with only one constraint, it was observed that many thermodynamically unfavorable and/or high enzyme cost pathways were excluded from EcoETM, enabling better prediction of optimal biosynthesis pathways and engineering targets.

The multi-constraint models presented in this study provide an accurate design blueprint for strain engineering, and the established framework for model construction can be easily adopted to construct MCMs for other organisms.

Email: ma_hw@tib.cas.cn

CosGeneGate Selects Multifunctional Biomarkers for Single-cell Analysis

Wenxin Long
School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, China

Abstract
Selecting marker genes for distinct cell types is an important task in single-cell sequencing analysis. Current approaches either fail in capturing the reasonable patterns of marker genes, orselect genes limited in downstream applications.

To overcome these problems, we present a model based on interpretable neural networks with stochastic **gate** (STG) design [1] and **cosine similarity** regularization [2] for the marker **gene** selection task, known as **CosGeneGate**. We utilize STG to select candidate genes based on cell type prediction accuracy, and then utilize cosine similarity of candidate genes as a post selector, which can shrink the scope of candidate genes so that the final list contains marker genes which are insimilar expression patterns. The final marker list can not only achieve high cell classification accuracy, but also follow the binary expression profile patterns supported by [2] and [3].

Our model can select marker genes with reasonable patterns by utilizing large-scale annotated scRNA-seq data as training datasets. We offer a list of marker genes for different tissues including PBMC, Pancreas and Heart. Moreover, we prove the superiority of CosGeneGate in different downstream applications by comprehensive benchmarking analysis, for example, accurate cell-type annotation, performance improvement of bulk-seq deconvolution tools, dimension reduction of single-cell data, spatially informed pattern identification for spatial transcriptomic data, incorrect cell types refinement, and disease-specific marker identification in Alzheimer's Disease (AD).

CosGeneGate can select high-quality marker genes for multiple downstream applications and has acceptable running time and memory usage. As a biology-informed tool, the next step of CosGeneGate will shift to discovering similar genes across species in multi-species data analysis and discovering relation among genes, peaks and proteins in multi-omics data analysis.

References

1. Yamada, Y., et al. (2020). Feature Selection using Stochastic Gates. Proceedings of the 37th International Conference on Machine Learning. D. Hal, III and S. Aarti. Proceedings of Machine Learning Research, PMLR. 119: 10648--10659.
2. Dai, M., et al. (2022). Accurate and fast cell marker gene identification with COSG. Briefings in Bioinformatics 23(2): 1--12
3. Aevermann, B., et al. (2021) A machine learning method for the discovery of minimum marker gene combinations for cell-type identification from single-cell RNA sequencing. Genome Research 33(8)

Email: 2012193@mail.nankai.edu.cn

Quantification of Cell Phenotype Transition Manifold with Information Geometry

Miao Huang[1,2], Junda Ying[3], Yuxuan Wang[3], Haijun Zhou[1,2], Lei Zhang[3], Weikang Wang[1,2]

[1]*CAS Key Laboratory for Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

[2]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[3]*School of Mathematics, Peking University, Beijing, China*

Abstract

Cell phenotype transition (CPT) is a common occurrence in many biological processes, such as development, and the precise control of CPT is critical in these processes. Recent advancements in single-cell sequencing techniques have revealed that cell transition dynamics in development are confined to low-dimensional manifolds. However, no available method currently exists for quantifying the manifolds from experimental data. Current methods, such as manifold learning, can only preserve topology information and can only be used for visualization.

In this work, we present a novel method to quantify the cell phenotype transition manifold using information geometry. Specifically, we transform single cells' high-dimensional expression vectors into probability distributions with Gaussian embedding. The Fisher-Rao metric is then naturally defined in this embedding space. With Gaussian embedding, we calculate the Ollivier-Ricci curvature of each single cell. Our analysis revealed that cells with low curvature are related to critical transitions.

To further analyze the invariant characteristics of the manifold in CPT, we calculated the information velocity of each single cell based on RNA velocity. The high information velocity regions correspond to the low curvature region, indicating that the geometry can guide the dynamics of single cells on the manifold.

We applied this method to different single-cell RNA sequencing datasets and compared it with traditional methods. Our approach focuses on analyzing the invariant characteristics of the manifold in CPT, and the results illustrate that it can be a general approach for quantifying the CPT manifold.

References

1. François P, Jutras-Dubé L. Landscape, bifurcations, geometry for development. Current Opinion in Systems Biology. 2018;11:129-36.
2. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nature Biotechnology. 2020;38(12):1408-14.
3. Bojchevski A, Günnemann S. Deep Gaussian Embedding of Graphs: Unsupervised Inductive Learning via Ranking. arXiv: Machine Learning. 2017.
4. Costa SIR, Santos SA, Strapasson JE. Fisher information distance: A geometrical reading. Discrete Applied Mathematics. 2015;197:59-69.
5. Ollivier Y. Ricci curvature of Markov chains on metric spaces. Journal of Functional Analysis. 2009;256(3):810-64.

Email: huangmiao@itp.ac.cn

Deciphering sequence-informed regulatory patterns from single-cell multi-omics

Fangxin Cai, Yuanhua Huang
School of Biomedical Sciences, University of Hong Kong

Abstract

Single cell multi-omics presents parallel and asynchronous perspectives to biological processes. To disentangle these relations, RNA velocity methods learn the splicing dynamics from spliced and unspliced scRNA-seq data to model cell state transitions[1]. Amongst the dynamic parameters, transcription rate is crucially important for uncovering the driving factors for cell fate decision. We introduce a sequence-informed RNA velocity framework that embeds spliced and unspliced data into a low dimensional space, which, encoded by a convolutional neural network (CNN) from genes' transcription start site (TSS) sequences, may highlight the regulatory dynamics. Velocity is subsequently learned by neural ordinary differential equations (neural ODEs)[2] for each factor in the sequence space. Preliminary velocity results correctly recover challenging trajectories, including erythroid lineage[3] and embryonic mouse brain lineage[4]. The sequence model naturally extends to paired scATAC-seq and scRNA-seq data by including ATAC peak sequences. We plan to further decipher genes' regulatory mechanisms by the coupling between expression and chromatin accessibility factors as well as sequence similarity.

References

1. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, Fan J. RNA velocity of single cells. Nature. 2018 Aug;560(7719):494-8.
2. Chen RT, Rubanova Y, Bettencourt J, Duvenaud DK. Neural ordinary differential equations. Advances in neural information processing systems. 2018;31.
3. Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, Mulas C, Ibarra-Soria X, Tyser RC, Ho DL, Reik W. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature. 2019 Feb 28;566(7745):490-5.
4. Li C, Virgilio MC, Collins KL, Welch JD. Multi-omic single-cell velocity models epigenome–transcriptome interactions and improves cell fate prediction. Nature Biotechnology. 2023 Mar;41(3):387-98.

Email: u3553349@connect.hku.hk

Exploring the Potential Pharmacological Mechanism of Hesperidin and Glucosyl Hesperidin against COVID-19 Based on Bioinformatics Analyses and Antiviral Assays

Yun Huang, Wei Zhou, Jing Sun, Guoliang Ou, Nan-shan Zhong and Zhigang Liu

Department of Respirology & Allergy, Third Affiliated Hospital of Shenzhen University, Shenzhen 518020, P. R. China

Abstract

The development of anti-COVID-19 drugs has become the top priority since the outbreak of the epidemic and Traditional Chinese medicine plays an important role in reducing mortality. Here, hesperidin and its glycosylation product, glucosyl hesperidin were selected to determine their antiviral activity against SARS-CoV-2 due to their structural specificity as reported. To be specific, their binding ability with ACE2, M, S, RBD and N proteins were verified with both *in silico* and wet lab methods, i.e. molecular docking and binding affinity tests, including biolayer interferometry assay (BLI) and isothermal titration calorimetry assay (ITC). Moreover, systematic pharmacological analysis was conducted to reveal their pharmacological mechanism in treating COVID-19. Finally, their antiviral activity against SARS-CoV-2 was determined *in vitro* in a biosafety level 3 (BSL3) laboratory. The results demonstrated their outstanding binding affinity with ACE2, M, S and RBD proteins, while showed barely unobserved binding with N protein, indicating their key roles in influencing the invasion and early replication phase of SARS-CoV-2. In addition, both hesperidin and glucosyl hesperidin showed great impacts to immune, inflammation and virus infection induced by COVID-19 according to the systematic pharmacological analysis. Moreover, the IC50 of hesperidin and glucosyl hesperidin against SARS-CoV-2 was further determined (51.5 μM and 5.5 mM, respectively) with cell-based *in vitro* assay, suggesting their great anti-SARS-CoV-2 activity. All in all, present research was the first to verify the binding ability of hesperidin and glucosyl hesperidin with SARS-CoV-2 proteins with both *in silico* and wet-lab methods, and proposed the possibility of applying hesperidin and glucosyl hesperidin to treat COVID-19.

References

1. Basu, A., A. Sarkar, U. Maulik. Molecular docking study of potential phytochemicals and their effects on the complex of SARS-CoV2 spike protein and human ACE2. *Sci. Rep.* 10: 17699, 2020.

2. Bellavite, P. and A. Donzelli. Hesperidin and SARS-CoV-2: New Light on the Healthy Function of Citrus Fruits. *Antioxidants (Basel).* 9: 742, 2020.

Email: yhuangbr@connect.ust.hk

A live-cell image-based machine learning strategy for reducing variability in PSC differentiation systems

Xiaochun Yang[1†], Daichao Chen[†], Qiushi Sun[†], Yao Wang, Yu Xia, Jinyu Yang, Chang Lin, Xin Dang, Zimu Cen, Dongdong Liang, Rong Wei, Ze Xu, Guangyin Xi, Gang Xue, Can Ye, Li-Peng Wang, Peng Zou, Shi-Qiang Wang, Pablo Rivera-Fuentes, Salome Püntener, Zhixing Chen, Yi Liu[*], Jue Zhang[*], Yang Zhao[1*]

[1]*State Key Laboratory of Natural and Biomimetic Drugs, MOE Key Laboratory of Cell Proliferation and Differentiation, Beijing Key Laboratory of Cardiometabolic Molecular Medicine, Institute of Molecular Medicine, College of Future Technology, Peking University, China.*

Abstract

The differentiation of pluripotent stem cells (PSCs) into diverse functional cell types provides a promising solution to support drug discovery, disease modeling, and regenerative medicine. However, functional cell differentiation is currently limited by the substantial line-to-line and batch-to-batch variabilities, which severely impede the progress of scientific research and the manufacturing of cell products. For instance, PSC-to-cardiomyocyte (CM) differentiation is vulnerable to inappropriate doses of CHIR99021 (CHIR) that are applied in the initial stage of mesoderm differentiation. Here, by harnessing live-cell bright-field imaging and machine learning (ML), we realize real-time cell recognition in the entire differentiation process, e.g., CMs, cardiac progenitor cells (CPCs), PSC clones, and even misdifferentiated cells. This enables non-invasive prediction of differentiation efficiency, purification of ML-recognized CMs and CPCs for reducing cell contamination, early assessment of the CHIR dose for correcting the misdifferentiation trajectory, and evaluation of initial PSC colonies for controlling the start point of differentiation, all of which provide a more invulnerable differentiation method with resistance to variability. Moreover, with the established ML models as a readout for the chemical screen, we identify a CDK8 inhibitor that can further improve the cell resistance to the overdose of CHIR. Together, this study indicates that artificial intelligence is able to guide and iteratively optimize PSC differentiation to achieve consistently high efficiency across cell lines and batches, providing a better understanding and rational modulation of the differentiation process for functional cell manufacturing in biomedical applications.

References

Yang, X. et al. A live-cell image-based machine learning strategy for reducing variability in PSC differentiation systems. Cell Discov. doi: 10.1038/s41421-023-00543-1 (2023).

Email: yangxiaochun@pku.edu.cn

Linking the Mammalian Oocyte Proteome Asymmetry to Embryonic Cell Fate Patterning

John Xiao-He Li and Anming Meng
Centre for Life Sciences and School of Life Sciences, Tsinghua University, Beijing 100084, China

Abstract

Unlike other vertebrates, mammals are thought to lack a prepatterned fate map in their embryos[1]. In support of this dogma, various single cell RNA-seq have failed to identify asymmetrically located maternal transcripts that can determine cell fate in mammalian early embryos[2]. However, transplantation of bisected 2-cell mouse embryos proved their unequal potentials to develop to term[3]. Labelling the oocyte animal pole by the second polar body (pb), we showed that the animal pole-derived blastomere (pbB) of the 2-cell embryo contributes to the inner cell mass (ICM) more than the vegetal blastomere (npbB) does[4]. Whether there exists a maternal factor dictating such bias is unknown. Here we identified one such factor by proteomics of bisected 2-cell embryos. 13 maternal proteins were asymmetrically located between the two cells among 3319 identified proteins ($\geq \pm 2$ fold, $P \leq 0.05$). Among them Macrophage Migration Inhibitory Factor (Mif) was the only highly expressed protein (2.2-fold npbB enriched). We verified the asymmetry of Mif in oocytes and embryos by immunofluorescence. Maternal *Mif* knockout mutation affected cell fate at E4.5 specifically reducing the epiblast (EPI) cell number without affecting primitive endoderm (PE) or trophectoderm (TE). Unlike all known zygotic mutations affecting EPI/PE differentiation (e.g. *Fgf4*, *Nanog*, *Gata6*), *Mif* was a maternal effect mutation. Further, *Mif* mutation did not cause EPI/PE interconversion as the other mutations did, possibly regulating the second cell fate differentiation by a novel mechanism. Indeed, lineage tracing showed that the Mif gradient biased pbB towards ICM but did not sway it between the EPI and PE fate. The prevailing model where ICM cells stochastically acquire EPI or PE fate can explain the bistable outcome of ICM differentiation the but may not guarantee the robust cell number regulation of EPI or PE[5]. Our study thus links an asymmetrically located maternal factor to the robust outcome of cell fate differentiation in mammals.

References

[1] Q. Chen, et al., Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. *Nat. Commun.* **9**, 1819 (2018).
[2] E. Casser, et al., Differences in blastomere totipotency in 2-cell mouse embryos are a maternal trait mediated by asymmetric mRNA distribution. *Mol. Hum. Reprod.* **25**, 729–744 (2019).
[3] M. Boiani, et al., Totipotency continuity from zygote to early blastomeres: a model under revision. *Reproduction*. **158**, R49–R65 (2019).
[4] H. Jin, et al., The second polar body contributes to the fate asymmetry in the mouse embryo. *Natl. Sci. Rev.* **9**, nwac003 (2022).
[5] S. Chowdhary, A.-K. Hadjantonakis, Journey of the mouse primitive endoderm: from specification to maturation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **377**, 20210252 (2022).

Email: xiaohe-li@tsinghua.edu.cn; mengam@tsinghua.edu.cn

RoboFold: a robot swarm system mimicking protein folding

Yuping Gu and Yongliang Yang*
Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning, 110016, China

Abstract

In the past two decades, the behaviors of animal swarms have inspired many control algorithms for robot swarms (*1*). The capability of current robots, however, is far from that of the animals. The capability gap between robots and animals hinders the reliability of these robot swarm systems. Thus, few robot swarms have been applied to solve problems in real world (*2*). Toward reliable robot swarm systems, a promising solution is learning from the collective behaviors of low intelligent natural swarms, such as cell collectives and molecular assemblies. The protein molecules, for example, can fold into specific 3D structures. In addition, all the information for folding a protein is coded in the sequence of amino acids of this protein (*3*). The requirements of a sequence forming protein molecules are thermal dynamical stable, fast folding time, and robust to mutations (*4*), which are also the ones for reliable robot swarm systems. Thus, the designability of protein molecules proposed in late 90s of last century may provide guiding principles for designing robot swarm systems (*5*, *6*).

Here, we reported our homemade RoboFold platform for mimicking protein folding toward a designing principle for developing reliable robot swarm systems. Using this system, we proposed a protein-folding inspired programming method for robot swarms. By designing the sequence of a thread of protein-like robots, the robots self-organized into specific morphologies. The sequence of robots mimicked the tertiary structure of protein molecules. This results together with other three experimental results verified the feasibility of this programmable designing method for robot swarms. To summarize, this work successfully translated the principles of protein-folding into algorithms for controlling the morphology of robot swarms. Besides benefiting swarm robotics, the RoboFold may also serve as a platform investigating the principles of protein folding.

References

1. M. Dorigo, G. Theraulaz, V. Trianni, Swarm Robotics: Past, Present, and Future [Point of View]. Proc. IEEE. 109, 1152–1165 (2021).
2. M. Dorigo, G. Theraulaz, V. Trianni, Reflections on the future of swarm robotics. Sci. Robot. 5, eabe4385 (2020).
3. C. B. Anfinsen, Principles that Govern the Folding of Protein Chains. Science. 181, 223–230 (1973).
4. R. Mélin, H. Li, N. S. Wingreen, C. Tang, Designability, thermodynamic stability, and dynamics in protein folding: A lattice model study. The Journal of Chemical Physics. 110, 1252–1262 (1999).
5. H. Li, R. Helling, C. Tang, N. Wingreen, Emergence of Preferred Structures in a Simple Model of Protein Folding. Science. 273, 666–669 (1996).
6. R. Helling, H. Li, R. Mélin, J. Miller, N. Wingreen, C. Zeng, C. Tang, The designability of protein structures. Journal of Molecular Graphics and Modelling. 19, 157–167 (2001).

* Email: ylyang@sia.cn

Deep flanking sequence engineering for efficient promoter design using DeepSEED

Pengcheng Zhang[1,*], Haochen Wang[1,*], Hanwen Xu[1,*], Lei Wei[1], Liyang Liu[1], Zhirui Hu[2], Xiaowo Wang[1,†]

[1] Ministry of Education Key Laboratory of Bioinformatics; Center for Synthetic and Systems Biology; Bioinformatics Division, Beijing National Research Center for Information Science and Technology; Department of Automation, Tsinghua University,

[2] Center for Statistical Science, Tsinghua University, Beijing, China

* Equally contributed

† Corresponding author: Xiaowo Wang, Tsinghua University

Beijing, 100084, China

Abstract

Designing promoters with desirable properties is essential in synthetic biology. Human experts are skilled at identifying strong explicit patterns in small samples, while deep learning models excel at detecting implicit weak patterns in large datasets. Biologists have described the sequence patterns of promoters via transcription factor binding sites (TFBSs). However, the flanking sequences of cis-regulatory elements, have long been overlooked and often arbitrarily decided in promoter design. To address this limitation, we introduce DeepSEED, an AI-aided framework that efficiently designs synthetic promoters by combining expert knowledge with deep learning techniques. DeepSEED has demonstrated remarkable success in improving the properties of Escherichia coli constitutive, IPTG-inducible, and mammalian cell doxycycline (Dox)-inducible promoters. Furthermore, our results show that DeepSEED captures the implicit features in flanking sequences, such as k-mer frequencies and DNA shape features, which are crucial for determining promoter properties.

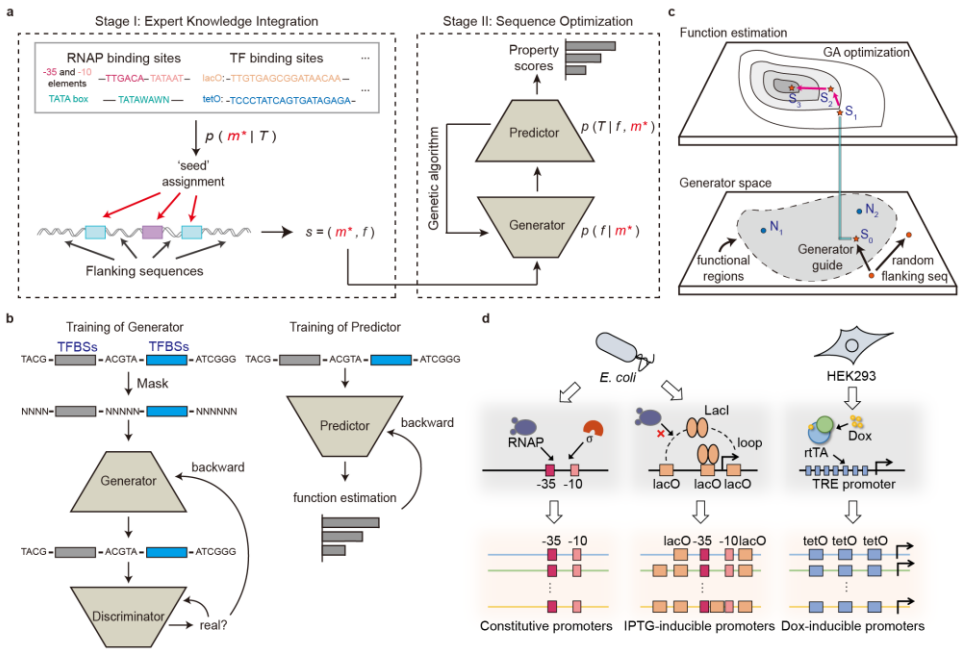


Fig. 1. An overview of the DeepSEED approach.

Email: wanghc17@mails.tsinghua.edu.cn

Deciphering principles in cell fate decisions via a logic-based gene regulatory network

Gang Xue[1], Lei Zhang[2,3,*] and Zhiyuan Li[1,2,*]

[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] *Beijing International Center for Mathematical Research, Center for Machine Learning Research, Peking University, Beijing 100871, China*

[*] *Correspondence: zhangl@math.pku.edu.cn; zhiyuanli@pku.edu.cn*

Abstract: How cells make fate decisions is a fundamental question in life science and has been the subject of extensive theoretical efforts from the perspective of systems biology. One notable example of such efforts is the seesaw model [1], where reprogramming into pluripotency can be achieved by restoring the balance between different lineages. However, despite decades of efforts, some essential questions have yet to be clarified. For example, in mathematical modeling of Gene Regulatory Networks (GRNs): 1) When multiple Transcription Factors (TFs) co-regulate one target gene, should their regulatory functions be multiplicative or additive? 2) Should cell fate transitions be modeled as switching between different "wells" under the same parameter setting or as a bifurcation driven by parameter changes? In experiments, these two questions correspond to: 1) What is the role of combinatory logic for a fate decision, and does it matter even under the same network topology? 2) Is a cell fate decision driven by noise or signal? Here, we constructed a logic-based GRN and examined its behaviors under two driving forces [2]. Under the noise-driven mode, we distilled the relationship between fate bias, regulatory logic, and noise. Under the signal-driven mode, we bridged regulatory logic with progression-accuracy trade-off. Additionally, we characterized a particular logic-dependent priming stage by the solution landscape during differentiation. Finally, we applied our findings to three real biological instances. Our work presents a generalizable framework for top-down fate-decision studies and a practical approach to the taxonomy of cell fate decisions.

This work was supported by the National Key Research and Development Program of China (No.2021YFF1200500, 2021YFA0910700), National Natural Science Foundation of China (No.12225102, 12050002, and 12226316). It is also supported by grants from Peking-Tsinghua Center for Life Sciences.

References

[1] J. Shu, C. Wu, Y. Wu, Z. Li et al., *Cell*, vol. 153, no. 5, pp. 963 – 75 (2013).

[2] G. Xue, X. Zhang, W. Li, L. Zhang et al., *bioRxiv* (2023).

Email: xuegang19@pku.edu.cn

Biocatalytic synthesis design with a graph neural network

Jiasheng Guo, Kangjie Lin, Luhua Lai
College of Chemistry and Molecular Engineering, Peking University, Beijing, 100871

Abstract
Biocatalysis, also known as enzyme catalysis, is a key complementary synthetic methodology to traditional chemical synthesis. Enzymatic reactions usually possess high efficiency and high selectivity, as well as special catalytic reactivities rarely found in synthetic chemistry. Moreover, enzymatic reactions are more environment-friendly and safer to operate. Enzymes have wide application in the synthesis of fine chemicals, especially for natural products and pharmaceuticals. According to the source of the enzyme, biocatalysis practices can be divided in to two classes, *in vivo* biocatalysis(enzymes are produced by living organisms) and *in vitro* biocatalysis(using purified enzymes). Using purified enzymes is more convenient for chemists to design and conduct enzymatic reactions in laboratories. Therefore, there are numerous synthetic routes in literature which includes one or more *in vitro* enzymatic reactions. However, it is not a common practice to combine enzymes with chemical synthesis, and the cases of enzymatic synthesis are still outnumbered by chemical synthesis.
Computer assisted synthesis planning has gain tremendous progress in the past decade, mainly driven by the development of deep learning techniques and the expansion of chemical reaction databases, and is able to provide innovative, chemically feasible and experimentally viable synthetic pathways for a given molecule. This paper proposes that similar methods used in chemical synthesis planning can be transferred to enzymatic synthesis design. Using a state-of-the-art graph neural network for chemical retrosynthesis prediction named LocalRetro and enzymatic reactions from literature and Reaxys database as training data, this paper presents an accurate single-step retrosynthetic predictive model for retrobiosynthesis. The prediction accuracy exceeds a dense neural network and a Transformer model trained with the same data. Then, the paper uses this model with multistep planning algorithms, MCTS and Retro*, to perform automatic enzymatic synthesis planning. This paper aims to assist chemists with designing more efficient synthesis empowered by enzymes.

Email: guojs@stu.pku.edu.cn

The mRNA concentration-dependent translational regulations of gap genes shape the gap protein patterns during *Drosophila* embryogenesis

Jiayi Chen[1], Feng Liu[1,2]

[1] *Center for Quantitative Biology, Peking University, Beijing, 100871, China*

[2] *Key Laboratory of Hebei Province for Molecular Biophysics, Institute of Biophysics, School of Health Science & Biomedical Engineering, Hebei University of Technology, Tianjin, 300130, China*

Abstract

Translational regulations play an important role in controlling the dynamic protein expressions[1] and features like noise attenuation[2]. The gap gene proteins exhibiting specific patterns on the anterior-posterior (AP) axis serve as an ideal model for investigating precise protein expression during *Drosophila* embryogenesis[3]. Due to the lack of the detailed mechanisms[4, 5] and the potential importance of the translational regulations[6-8], we focus on the translational regulations during the establishment of the gap gene protein patterns. Here we show that both the ratio of the translating mRNA (P_{on}) and the translation efficiency of these translating mRNA (α_{on}) exhibit spatial-temporal heterogeneity, which can be simply unified into the mRNA concentration-dependent ([mRNA]-dependent) manners. Specifically, P_{on} is non-monotonically correlated with the increasing [mRNA] and peaks at a typical concentration C_0, while the α_{on} decreases monotonically. These translational regulations are observed in both *hb* and *kni*, which represent the anterior and posterior patterns of the gap genes on the AP axis respectively. We further propose an electrostatic-based reentrant phase separation model to provide a potential mechanism of these [mRNA]-dependent translational regulations. As a test, we confirmed that this model can fit the published mRNA and protein pattern of *hb* during Drosophila embryogenesis. Our results demonstrate the complexity at the translation process and suggest the contributions of these regulations in shaping the gap protein patterns.

References

1. Kang, J.-Y., et al., Science, 2022.
2. Wu, H.-W., et al., Nature Plants, 2022.
3. Yang, Z., et al., Elife, 2020.
4. Verd, B., A. Crombach, and J. Jaeger, PLOS Computational Biology, 2017.
5. Little, S.C., M. Tikhonov, and T.J.C. Gregor, Cell, 2013.
6. Dufourt, J., et al., Science, 2021.
7. Vinter, D.J., et al., Development, 2021.
8. Livingston, N.M., et al., Molecular Cell, 2023.

Email: jiayi.chen@pku.edu.cn

A synthetic circuit capable of automatic division of labor within a homogeneous population

Zhi Sun[1], Ning Yang[1], Baiyi Jiang[1], Chunbo Lou[2], Chao Tang[1] and Xiaojing Yang[1]

[1]*Center for Quantitative Biology, Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China.*

[2]*Insititute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China.*

Abstract

Division of labor within populations is a game-changer for complex functions in synthetic biology, dramatically bolstering system efficiency, stability, and robustness. Conventional co-cultivation strategies, however, fall short in adjusting labor division in response to stimuli and often invite intra-species competition. We here designed and constructed a gene circuit that empowers homogeneous populations to auto-divide labor, with the subpopulation ratio flexibly responding to stimuli frequency. Remarkably, after stimuli response, the population can revert to a homogeneous state for labor re-division, greatly reducing intra-species competition. We enumerated two-node topologies for critical functional requirements. The results shows that the negative feedback coupled with positive feedback on the buffer node can achieve all functions, robustly. Further simulation revealed a saddle-node bifurcation controls all the designed functions. 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. Parameter swapping based on a detailed stochastic model of the genetic circuit showed that 32% perturbations were tolerable. Within the acceptable range, bistable and response property could be adjusted and further demonstrated by experimentally changing the genetic parts. Finally, we demonstrated that the heterogeneous subgroups could regenerate the differentiation process identically, indicating a memoryless division of labor. The synthetic system will keep repeating the differentiation and convergence process if the input stimulus switches periodically. Both the response strength and the functional subgroup proportion changed with the input frequency. Our work demonstrates that the population-level labor division control can be rationally designed by individual regulation, offering a new avenue for dynamic heterogeneous community design. The circuit can also be useful in the rational design of stable distributed functional control and differentiable multicellular life.

Email: sunz@pku.edu.cn

A phase diagram structure determines the optimal sensitivity-precision trade-off in signaling systems

Ka Kit Kong[1], Chunxiong Luo[1,2,3], Feng Liu[2,4]

[1] *The State Key Laboratory for Artificial Microstructures and Mesoscopic Physics, School of Physics, Peking University, Beijing, 100871, China*

[2] *Center for Quantitative Biology, Peking University, Beijing, 100871, China*

[3] *Wenzhou Institute University of Chinese Academy of Sciences, Wenzhou, Zhejiang, China*

[4] *Key Laboratory of Hebei Province for Molecular Biophysics, Institute of Biophysics, School of Health Science & Biomedical Engineering, Hebei University of Technology, Tianjin, 300130, China*

Abstract

Signal transduction is widely presented in many biological processes such as cell fate decision. Response noise and response sensitivity are two key features of signaling processes, but their quantitative relationship remains to be characterized. Furthermore, understanding the behavior of such processes often requires *ad hoc* modeling, which usually relies heavily on data fitting. In our study, we propose a top-down approach based on the phase diagram structure to study the quasi-steady-state properties of signaling systems. Using this modeling approach, we analytically derive that the lower limit of the response noise (represented by the variance of the response), is bounded by the sensitivity (defined as the square of the slope of the response curve) up to a scale. This is a general relationship for signaling processes under static encoding noise, which resembles the role of fluctuation-dissipation relation in statistical physics for connecting the response and fluctuation in equilibrium systems. Our results reveal that signaling networks with the optimal trade-off can be better characterized by a phase diagram structure, rather than the widely-used network topology, offering an alternative method for predictive modelling. Furthermore, we identify the predicted noise-sensitivity relation in the patterning gene expression data in early fly embryogenesis[1] and find that the patterning network appears to be optimized. Based on the optimal phase diagram structure, we can quantitatively predict the patterning response to the dosage of morphogen Bicoid[2] without parameter fitting. Our model provides a novel perspective in understanding noise propagation in biological signal transduction.

References

1. Petkova, M. D., Tkačik, G., Bialek, W., Wieschaus, E. F. & Gregor, T. Optimal Decoding of Cellular Identities in a Genetic Network. Cell 176, 844-855.e15 (2019).

1. Liu, F., Morrison, A. H. & Gregor, T. Dynamic interpretation of maternal inputs by the Drosophila segmaentation gene network. Proceedings of the National Academy of Sciences 110, 6724–6729 (2013).

Email: kakit@pku.edu.cn

Balancing Reaction-diffusion Network for Cell Polarization Pattern with Stability and Asymmetry

Yixuan Chen [1,2,3 †], Guoye Guan [1,4,5 †], Lei-Han Tang [1,5,6,7 *], Chao Tang [3,4,8 *]

1. South Bay Interdisciplinary Science Center, Songshan Lake Materials Laboratory, Guangdong 523808, China
2. Department of Physics, Sichuan University, Sichuan 610065, China
3. School of Physics, Peking University, Beijing 100871, China
4. Center for Quantitative Biology, Peking University, Beijing 100871, China
5. Department of Physics, Hong Kong Baptist University, Hong Kong 999077, China
6. Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Hong Kong 999077, China
7. State Key Laboratory of Environmental and Biological Analysis, Hong Kong Baptist University, Hong Kong, 999077, China
8. Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China
[†]These authors contributed equally to this work. [*]Corresponding authors.

Abstract

Cell polarization is a critical process that separates molecules into two distinct regions in prokaryotic and eukaryotic cells, guiding biological events like cell division and cell differentiation. Although some underlying antagonistic reaction-diffusion networks capable of setting up cell polarization have been identified experimentally and theoretically, it is still elusive how the pattern stability and asymmetry can be modulated. Here, we first numerically demonstrate that the polarized pattern generated by an antagonistic 2-node network would collapse into a homogeneous distribution when single-sided self-regulation, single-sided additional regulation, or unequal system parameters are added. Interestingly, the combination of two of those unbalanced modifications can stabilize the polarized pattern. The observed behavior can be attributed to front propagation in a bistable reaction-diffusion system whose speed can be tuned by kinetic processes at the interface. We conduct an elaborate literature search to reconstruct the cell polarization network in the nematode *Caenorhabditis elegans* zygote, where a 4-node network with full mutual inhibitions between anterior and posterior is modified by a mutual activation in the anterior and an additional mutual inhibition between the anterior and the posterior, constituting a 5-node network. Our computational exploration confirms the generic behavior mentioned above. Numerical simulation further reveals the balance between these two modifications, which jointly maintain pattern stability and enhance pattern asymmetry. Our computational framework successfully simulates the simple 2-node network and *C. elegans* 5-node network in wild-type and perturbed embryos, providing new insight into the design principles of both natural and artificial cell polarization systems. Last but not least, we build user-friendly software, *PolarSim*, to facilitate the exploration of networks with alternative node numbers, parameter values, and regulation pathways.

References

[1] Lim, et al. A balance between antagonizing PAR proteins specifies the pattern of asymmetric and symmetric divisions in *C. elegans* embryogenesis. Cell Rep 2021;**36**:109326.
[2] Lang et al. The PAR proteins: from molecular circuits to dynamic self-stabilizing cell polarity. Development 2017;**144**:3405-16.
[3] Seirin-Lee & Shibata. Self-organization and advective transport in the cell polarity formation for asymmetric cell division. J Theor Biol 2015;**382**:1-14.
[4] Guan, et al. Volume segregation programming in a nematode's early embryogenesis. Phys Rev E 2021;**104**:054409.
E-mail: lhtang@hkbu.edu.hk (L.-H. Tang), tangc@pku.edu.cn (C. Tang).

Coordinate-Wise Monotonic Transformations Enable Privacy-Preserving Facial Age Estimation

Xinyu Yang[1,2], Runhan Li[2], Xindi Yang, Yong Zhou, Yi Liu & Jing-Dong J. Han[2*]

[1] *School of Life Sciences, Peking University, Beijing 100871, China*
[2] *Center for Quantitative Biology (CQB), Peking University, Beijing 100871, China*

Abstract

The human face is a valuable biomarker of aging, but the collection and use of its image raise significant privacy concerns. Here we present an approach for facial data masking that preserves age-related features using coordinate-wise monotonic transformations. We first develop a deep learning model that estimates age directly from non-registered facial point clouds with high accuracy and generalizability. We show that the model learns a highly indistinguishable mapping using faces treated with coordinate-wise monotonic transformations, indicating that the relative positioning of facial information is a low-level biomarker of facial aging. Through visual perception tests and computational 3D face verification experiments, we demonstrate that transformed faces are significantly more difficult to perceive for human but not for machines, except when only the face shape information is accessible. Our study leads to a facial data protection guideline that has the potential to broaden public access to face datasets with minimized privacy risks.

References

Guo, M. *et al*. (2021). PCT: Point cloud transformer. Computational Visual Media 7, 187-199.
Xia, X., Chen, X., Han, J.J. *et al*. (2020). Three-dimensional facial-image analysis to predict heterogeneity of the human ageing rate and the impact of lifestyle. Nat Metab 2, 946-957.
Chen, W.Y., Han, J. *et al*. (2015). Three-dimensional human facial morphologies as robust aging markers. Cell Res 25, 574-587.

Email: xinyu.yang@pku.edu.cn; jackie.han@pku.edu.cn

Three-dimensional facial image analysis reveals cross-ethnic aging and developmental patterns

Yiyang Wang, Jing-Dong Jackie Han

Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences
Center for Quantitative Biology, Peking University

Abstract

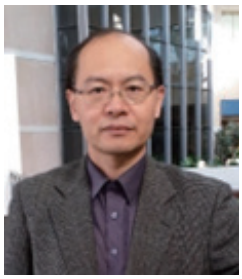
Considering the ethnic difference of aging and developmental process, we extend our previous aging clock of Asians to Africans. We combine the deep CNNs and 1200 Ghana samples to construct an African aging clock. The African aging clock reaches a decent performance with an MAR as 5.05 years. The aging heterogeneity of Africans also peaks at the middle age, similar as the Asian cohorts. We also have found numerous healthy parameters associated with chronological age and aging rate in African cohort, like the key one SBP. By causal inference test, we found that visceral fat can increase the SBP and BP difference thus directly elevate the aging rate in African cohorts, suggesting the importance of managing blood pressure for Africans. As the aging clock capturing aging patterns of face, we jointly analyze the whole face and landmark features landscape during aging for both African and Asian cohorts, and found that those areas (like eyes and mouth) with differential aging patterns potentially are under nature selection. As the aging patterns are due to face structure, we further detect the face shape related genes on transcriptome level of Asian peripheral blood mononuclear cells. By phenotype-driven face segmentation and canonical correlation analysis, we successfully identified 366 face shape related genes (FSGs) from global to local face level. Further enrichment analysis shows that those FSGs significantly enrich at morphogenesis pathways especially cranial structure development like skeletal, brain development and adipogenesis. Those FSGs also show higher expression level in cranial neural crest cells, suggesting they potentially regulate the 3D face structure variation of normal level. We then detect the African ancestry related FSGs (anc- FSGs) that play key roles in both African and Asian cohorts. 26 anc-FSGs are identified and manifest significantly optimized ability to synthesize the ethnic face difference than random selected genes. By jointly visualizing the face ancestry PCC and gene importance in pseudo-face model of the 26 genes, we found the most vital target genes ZNF81 may expressed differentially in embryogenesis stage of African and Asian cohorts thus regulate the face structure variation at a normal level.

Email: wangyiyang@picb.ac.cn

Quantitative Biology
Open Access

Call for paper: Understanding complex biological phenomena from mathematical and physical perspectives

Invited Editors



Prof. Luonan Chen
Center for Excellence in Molecular Cell Science, CAS

Prof. Qi Ouyang
Peking University

Prof. Ming Li
Institute of Physics, CAS

Background and Objective

Physical models and mathematical methods are exhibiting magic vigor to understand the complex biological phenomena. They accelerate the research of dynamics and thermodynamics of life systems at the molecular level, reveal the motion laws and quantitative principles in multi-scale life systems, explore the predictability and predictable design boundaries of artificial life systems and so on. It's no doubt that the widespread application of these physical models and/or mathematical methods in elucidating the mechanisms of complex biological phenomena will advance a new research paradigm in the field of life sciences.

QB is planning a special issue on the theme "Understanding complex biological phenomena from mathematical and physical perspectives". We welcome research articles, short research letters, reviews, mini-reviews, perspectives, and insights on any topics related to this theme.

Higher Education Press WILEY

About *Quantitative Biology*

Quantitative Biology (Editors-in-Chief: Dr. Chao Tang and Dr. Michael Q. Zhang, Executive Editor-in-Chief: Dr. Xuegong Zhang; http://journal.hep.com.cn/qb; https://onlinelibrary.wiley.com/journal/2095469) is an interdisciplinary journal that focuses on significant research that uses quantitative approaches and technologies to analyze and integrate biological systems, construct and model engineered life systems, and to gain a deeper understanding of life sciences. It aims to provide a platform for not only the analysis but also the integration and construction of biological systems.

Publication and Peer-review Process

All manuscripts are reviewed by the Editorial Board and qualified reviewers. The journal makes the decisions as rapidly as possible, and we strive to return reviewers comments to authors within 3 weeks. All accepted articles will be published online first after proofreading and formatting process. As an incentive, there is no page charges.

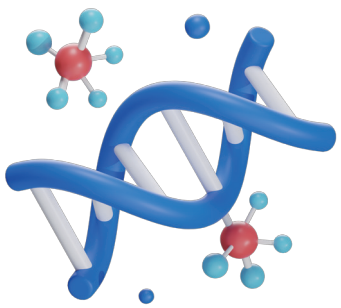
Indexing Body and Partners

QB has been indexed by ESCI, BIOSIS Previews, Biological Abstracts, Google Scholar, Scopus, Chemical Abstracts Service (CAS), Chinese Science Citation Database (CSCD), EBSCO Discovery Service, Institute of Scientific and Technical Information of China, Naver, OCLC WorldCat Discovery Service, ProQuest-ExLibris Primo, ProQuest-ExLibris Summon.

IF 2022
3.1

CiteScore2022
5.7

- **Important Date**
  Submission Deadline: Feb 28, 2024



- **Manuscripts Submission**
  https://mc.manuscriptcentral.com/qb

- **Content Available Online**
  http://journal.hep.com.cn/qb
  https://onlinelibrary.wiley.com/journal/2095469t