JINZHUANG DOU, Ph.D

MD Anderson Cancer Center
1400 Pressler St., Houston,77030
jdou1@mdanderson.org
+1(512) 6969846

RESEARCH INTERESTS

Statistical Genomics, Computational Biology, AI in Biomedicine

EDUCATION

Ph.D in Genetics August 2012 - June 2015
Ocean University of China, QingDao, China (Supervisor Zhenmin Bao)
M.S. in Applied Mathematics August 2009 - June 2012
Ocean University of China, QingDao, China
B.S. in Applied Mathematics August 2005 - June 2009
Ocean University of China, QingDao, China

RESEARCH EXPERIENCE

Associate Data Scientist Jan 2021 - Now, Supervisor: Ken Chen
MD Anderson Cancer Center, Houston, TX, USA
Postdoctoral research fellow July 2019 - Jan 2021, Superviosr: Ken Chen
MD Anderson Cancer Center, Houston, TX, USA
Postdoctoral research fellow July 2015 - July 2018, Supervisor: Chaolong Wang
Genome Institute of Singapore, A*STAR,Singapore

TEACHING

Instructor: "Multiomics integration and feature selection", Gulf Coast Consortia (GCC) Single Cell Omics Workshop 2022

AWARDS

"Wall of Science", MD Anderson Cancer Center (2023)

PUBLICATIONS

First/Co-first author

J Dou , Y Tan, KH Kock, J Wang , X Cheng, LM Tan, KY Han, CC Hon, WY Park, JW Shin, H Jin, H Chen, L Ding, S Prabhakar, N Navin. K Chen (**2023**). Monopogen: single nucleotide variant calling from single cell sequencing. ***Nature Biotechnology***. 2023 Aug 17:1-0

[Research Highlight in ***Nature Genetics***. https://www.nature.com/articles/s41588-023-01544-2]

D Wu*, **J Dou***, X Chai*, C Bellis, A Wilm, C.C. Shih, W.W.J Soon, N Bertin, C.B Lin, C.C Khor, M DeGiorgio, S.S Cheng, L Bao, N Karnani, W Hwang, S Davila, P Tan, A Shabbir, A Moh, E Tan, J.N Foo, L.L Goh, K.P Leong, R.S.Y Foo, C Lam, A Richards, C.Y Cheng, T Aung, T Wong, H Ng, M Ackers-Johnson, E Aliwarga, K Kim Ban, D Bertrand, J Chambers, D Hui Chan, C Li Chan, M Chee, M Chee, P Chen, Y Chen, E Chew, W Chew, L Chiam, J Chong, I Chua, S Cook, W Dai, R Dorajoo, C Foo, R Goh, A Hillmer, I Irwan, F Jaufeerally, A Javed, J Jeyakani, J Koh, J Koh, P Krishnaswamy, J Kuan, N Kumari, A Lee, S Lee, S Lee, Y Lee, S Leong, Z Li, P Li, J Liew, O Liew, S Lim, W Lim, C Lim, T Lim, C Lim, S Loh, A Lok, C Chin, S Majithia, S Maurer-Stroh, W Meah, S Mok, N Nargarajan, P Ng, S Ng, Z Ng, J Ng, E Ng, S Ng, S Nusinovici, C Ong, B Pan, V Pedergnana, S Poh, S Prabhakar, K Prakash, I Quek, C Sabanayagam, W See, Y Sia, X Sim, W Sim, J So, D Soon, E Tai, N Tan, L Tan, H Tan, W Tan, M Tandiono, A Tay, S Thakur, Y Tham, Z Tiang, G Toh, P Tsai, L Veeravalli, C Verma, L Wang, M Wang, W Wong, Z Xie, K Yeo, L Zhang, W Zhai, Y Zhao, J Liu, C Wang (**2019**) Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. ***Cell*** 179(3):736-749 (* joint first author) [**Cover story**]

L Li*, V Mohanty*, **J Dou***, Y Huang*, P P. Banerjee, Q Miao, J Lohr, T Vijaykumar, B Knoechel, R Basar, S Liang, M Kaplan, M Daher, E Liu, L Muniz-Feliciano, T J. Laskowski, D Marin, S Mielke, R E. C, Elizabeth J. Shpall, K Chen, K Rezvani (**2023**). Metabolic competition is an important driver of tumor resistance after CAR NK cell therapy and can be overcome by cytokine engineering. ***Science Advance*** 9(30), eadd6997. (* joint first author)

J Dou, S Liang, V Mohanty, X Cheng, S Kim, J Choi, Y Li, K Rezvani, R Chen, K Chen (**2022**). Unbiased integration of single cell multi-omics data. ***Genome Biology*** 23(1), pp.1-25

J Dou, D Wu, L Ding, K Wang, M Jiang, X Chai, D. Reilly, E Tai, J Liu, X Sim, S Cheng, C Wang (**2020**). Using off-target data from whole-exome sequencing to improve genotyping accuracy, association analysis, and polygenic risk prediction. ***Brief in Bioinformatics*** bbaa084

J Dou*, B Sun*, X Sim, I Irwan, J D Hughes, D F Reilly, E Tai, J Liu, C Wang (**2017**). Estimation of kinship coefficient in structured and admixed populations using sparse sequencing data. ***PloS Genetics*** 13: e1007021s.

J Dou*, H Dou*, C Mu, L Zhang, Y Li, J, T Li, X Hu, S Wang, Z Bao (**2017**). Whole-genome restriction mapping by "subhaploid"-based RAD sequencing: an efficient and flexible approach for physical mapping and genome scaffolding. ***Genetics*** 206:1237-1250. (* joint first author) [**Cover story**]

J Dou, X Zhao, X Fu, W Jiao, N Wang, L Zhang, X Hu, S Wang, Z Bao (**2012**). Reference-free SNP calling: improved accuracy by preventing incorrect calls from repetitive genomic regions. ***Bioloy Direct*** 7:17.

J Dou, X Li, Q Fu, W Jiao, Y Li, T Li, Y Wang, X Hu, S Wang, Z Bao (**2016**). Evaluation of the 2b-RAD method for genomic selection in scallop breeding. ***Scientific Report*** 6:19244.

W Jiao*, X Fu*, **J Dou***, H Li, H Su, J Mao, Q Yu, L Zhang, X Hu, X Huang, Y Wang, S Wang, Z Bao (**2013**).High-resolution linkage and quantitative trait locus mapping aided by genome survey sequencing: building up an integrative genomic framework for a bivalve mollusc. ***DNA Research*** 21 (1): 85-101. (* joint first author) [**Editor's choice**]

X Fu*, **J Dou***, J Mao, H Su, W Jiao, L Zhang, X Hu, X Huang, S Wang, Z Bao (**2013**).RADtyping: an integrated package for accurate de novo codominant and dominant RAD genotyping in mapping populations. ***PloS One*** 8:11. (* joint first author)

Contributing author

(selected)

D Marin, Y Li, R Basar, H Rafei, M Daher, **J Dou**, V Mohanty, M Dede, Y Nieto,N Uprety,S Acharya,E Liu ,J Wilson, P Banerjee, H Macapinlac, C Ganesh, P Thall, R Bassett, M Ammari, S Rao, Cao,M Shanley, C Hosing,P Kebriaei, L Nastoupil, C Flowers, S Moseley,P Lin, S Ang, U Popat, M Qazilbash,R Champlin, K Chen, E Shpall, K Rezvani. Cord blood characteristics determine the efficacy of allogeneic IL-15 armored CAR-transduced natural killer cells against CD19-positive B cell tumors. (2024: 1-13) ***Nature Medicine***)

M Shanley, **J Dou**, M Daher, R Basar, J Gumin,M Dede ,Q Miao, V Mohanty, H Shaim, S Li, S Acharya,..., K Chen, K Rezvani (2023). IL-21-induced CEBP is a critical regulator of NK memory against glioblastoma. (to be submitted)

Y Lyu, Y Guan, L Deliu, E Humphrey, J Frontera, Y Yang, **J Dou**,...,K Chen,P Nagarajan, Y Ge. (2022).

KLF5 governs sphingolipid metabolism and barrier function of the skin. ***Genes Development***, 36(13-14), 822-842.

R Fu*, W He*, **J Dou**, O D. Villarreal, E Bedford, H Wang, C Hou, L Zhang, Y Wang, Y Chen, X Gao, M Depken, H Xu. (2022) Systematic decomposition of sequence determinants governing CRISPR/Cas9 specificity. ***Nature Communications*** 13(1), 474.

S Liang, V Mohanty, **J Dou**, Q Miao, Y Huang, M Muftuoglu, L Ding, W Peng, K Chen (2021).Single-cell manifold-preserving feature selection for detecting rare cell populations. ***Nature Computational Science*** 5: 374-384

F Wang, Q Wang, V Mohanty, S Liang, **J Dou**, J Han, D Conterno Minussi, R Gao, L Ding, N Navin, K Chen (2021). Single-cell copy number lineage tracing enabling gene discovery. ***Genome Biology*** 22: 1-22

D Wu, P Li, B Pan, Z Tiang, **J Dou**, I Williantarra, A Pribowo, R Nurdiansyah, SG Peranakan Project, R SY Foo, C Wang (2021). ***Molecular Biology and Evolution*** 22: 1-22

R Basar, N Uprety, E Ensley, M Daher, K Klein,**J Dou**,K Rezvani (2021). Generation of glucocorticoid-resistant SARS-CoV-2 T cells for adoptive cell therapy.***Cell Reports*** 3:109432

H Shaim, M Shanley, R Basar, M Daher,..., **J Dou**,...,K Chen, F Lang, E Shpall, A Heimberger, K Rezvan. (2021) Targeting the v integrin-TGF-beta axis improves natural killer cell function against glioblastoma stem cells. ***The Journal of Clinical Investigation***

W He, L Zhang, O Villarreal, R Fu, E Bedford, **J Dou**, A Patel, M Bedford, X Shi, T Chen, B Bartholomew, H Xu (2019). De novo identification of essential protein domains from CRISPR-Cas9 tiling-sgRNA knockout screens. ***Nature Communications*** 10: 150130.

Y Li, X Sun, X Hu, X Xun, J Zhang, X Guo,..., **J Dou**,..., Z Jiang, R Li, S Wang, Z Bao (2017). Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins. ***Nature Communications*** 8:1721.

S Wang, J Zhang, W Jiao, J Li, X Xun, Y Sun,...,**J Dou**,..., Z Jiang, D Chourrout, R Li, Z Bao (2017). Scallop genome provides insights into evolution of bilaterian karyotype and development. ***Nature Ecology and Evolution*** 1:0120

OTHERS

BOOK: S Wang, J Lv, **J Dou**, Q Lu, L Zhang, Z Bao (**2017**).Genotyping by Sequencing and Data Analysis: RAD and 2b-RAD Sequencing. **Bioinformatics in Aquaculture**.

Performed analysis of CRISPR-engineered CAR-NK/T trial in glioblastoma for acquiring FDA approval (NY-ESO-1 TCR/IL-15 NK).

PATENT: K.Rezvani, M Shanley, K Chen, **J Dou**, D Merve, M Vakul. ENGINEERED NATURAL KILLER CELLS WITH ENHANCED ANTITUMOR MEMORY RESPONSES. NRF: MDAC.P1355US.P1/1001246688 (2 percentage contribution)

SOFTWARE

CCAdge (CCA-based DEGs analysis in single cell data)
This tool is developed to identify biologically varied genes in population level single cell data. (https://github.com/KChen-lab/CCAdge).

Monopogen (Single nucleotide variant calling from single cell sequencing)

This tool is developed to achive SNV calling from single cell RNA sequencing, single cell ATAC sequencing. (https://github.com/KChen-lab/Monopogen).

bindSC (Bi-order integration of single cell multi-omics data)
This tool is developed to achive bi-order integration (in silico multi-omics data) of single cell RNA sequencing, single cell ATAC sequencing, spacial transcriptomics and CyTOF data.
(https://github.com/KChen-lab/bindSC).

WEScall (A genotype calling pipeline for whole exome sequencing (WES) data)
WESCall can analyze both target and off-target data, and is a useful tool to facilitate WES sutides with decent amount of off-target data.
(https://github.com/dwuab/WEScall).

SEEKIN (Sequencing-based estimation of kniship and inbreeding coefficients)
This program estimates the kinship and inbreeding coefficients based on extremely low-sequencing coverage datasets (typically lower than 0.5X).
(https://github.com/chaolongwang/SEEKIN).

AMMO (An integrated package for whole-genome restriction mapping and genome scaffolding).
This is an integrated program that generates 2b-RAD-based restriction maps for scaffolding draft genome assemblies produced by using short Illumina reads or long PacBio reads.
(https://github.com/jinzhuangdou/AMMO).

RADtyping (De novo RAD Genotyping in Mapping Populations)
This is an integrated package for accurate De Novo codominant and dominant RAD genotyping in mapping populations.
(https://github.com/jinzhuangdou/RADtyping)

CONFERENCE PRESENTATION

Estimation of kinship coefficient using sparse sequencing data. The American Society of Human Genetics 66th Annual Meeting, Vancouver, Canada, Oct 2016. (**poster**)

Estimation of kinship coefficient using sparse sequencing data. Dana-Farber Brigham Cancer Center, Boston, USA, May 2019. (**talk**)

Single-cell characterization of CAR-NK cell exhaustion in transplantation immunotherapy. Texas Single Cell Seminar, Houston, USA, Feb 2020. (**talk**)

Unbiased integration of single cell multi-omics data. CZI Seed Networks 2020 Annual Meeting, Zoom, Nov 2020. (**talk**)

Unbiased integration of single cell multi-omics data. Texas Single Cell Seminar, Zoom, May 2021. (**talk**)

Single-cell characterization of CAR-NK cell exhaustion in transplantation immunotherapy. 2rd Annual GCC Single Cell Omics Symposium, Houston, USA, Oct 2021. (**poster**)

Monopogen: single nucleotide variant calling from single cell sequencing. 3rd Annual GCC Single Cell Omics Symposium, Houston, USA, Oct 2022. (**poster**)

Single nucleotide variant calling from single cell sequencing using Monopogen. NCI Informatics Technology for Cancer Research (ITCR) Meeting, Chicago, USA, Sep 2023. (**poster, the second prize**)

Single nucleotide variant calling from single cell sequencing using Monopogen. Texas Single Cell Seminar, Zoom, Oct 2023 (**talk**)

PROFESSIONAL ACTIVITIES

Peer Reviewer Cell, Nature Biotechnology, Genome Biology (assist with PIs)

Genome Research, Bioinformatics, Biology Direct, PLOS One

Editors Guest editor for Research Topic: Advances in Mathematical and Computational Oncology, Volume III in Frontiers (2022)

REFERENCES

1. Ken Chen, Ph.D. - Current supervisor. Professor of Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center. kchen3@mdanderson.org

2. Rui Chen, Ph.D. - Collaborator. Professor of Molecular and Human Genetics, Baylor College of Medicine. ruichen@bcm.edu

3. Chaolong Wang, Ph.D. - Postoctoral Mentor. Professor of Tongji Medical College, Huazhong University of Science and Technology. chaolong@hust.edu.cn

4. Zhenmin Bao, Ph.D. - Ph.D Mentor. Professor of Marine Biology, Ocean University of China. Academician of the Chinese Academy of Engineering. zmbao@ouc.edu.cn