

# Jitao David Zhang

*A Computational Biologist in Drug Discovery*

F. Hoffmann-La Roche AG  
4070 Basel, Switzerland

✉ [jitao\\_david.zhang@roche.com](mailto:jitao_david.zhang@roche.com)

📧 [jdzhang.me](http://jdzhang.me)



## Key facts

- Interdisciplinary researcher with 10+ years' experience in drug discovery.
- Supporting drug discovery projects between target identification and phase I.
- Proponent of multiscale modeling and causal inference of drug mechanism and safety.
- Open-source software developer and open-access author (40+ publications, *h*-index 25).

## Research career

- 2011– **Senior Principal Computational Biologist**, *F. Hoffmann-La Roche AG*, Basel, Switzerland.  
I develop algorithms and software to mine, interpret, model and integrate heterogeneous data, and apply mathematical and computational tools to support preclinical drug discovery projects. I am responsible for target assessment and validation, multiscale modeling and causal inference of drug mechanism and safety, and MoA characterisation and de-risking of drug candidates. With colleagues I co-develop novel platforms and resources to support drug discovery, for instance the molecular phenotyping platform and the Small-molecule PATHway Research Kit (SPARK) library.

## Teaching, mentoring, and academic commitment

- 2018– Teaching lecture series *Introduction to Applied Mathematics and Informatics In Drug Discovery* (<http://AMIDD.ch>) and *Mathematical and Computational Biology in Drug Discovery* (<http://MCBDD.ch>), Department of Mathematics and Informatics, University of Basel.
- 2018– Participation in the lecture series *From Novel Targets To Novel Therapeutic Modalities*, Master Programme Drug Sciences, University of Basel.
- 2016– Supervising master students in bioinformatics and computational biology, and co-supervise PostDocs in collaboration with academic and industrial collaboration partners.
- 2012– Reviewing research proposals for funding agencies and manuscripts for journals including *Bioinformatics*, *PLOS Comp. Biol.*, *NAR Genom. and Bioinform.*, etc.
- 2012– Leading courses, seminars and workshops about programming, bioinformatics, and computational biology in drug discovery. Certificated as a Carpentry Instructor in 2021.

## Education

- 2008–2011 **Dr.rer.nat. Bioinformatics**, *German Cancer Research Center/ Universität Heidelberg*.  
Computational and statistical approaches to study gene networks, supervised by Dr. Stefan Wiemann
- 2007–2007 **Marie-Curie Fellow**, *Huber Group, European Institute of Bioinformatics*, Cambridge, UK.
- 2006–2008 **M.Sc. Bioinformatics**, *Universität Heidelberg*, Germany.
- 2002–2006 **B.Sc. Biology, hon.**, *Peking University*, Beijing, China.

---

## Selected peer-reviewed publications

Besca, a single-cell transcriptomics analysis toolkit to accelerate translational research. *NAR Genomics and Bioinformatics*, 2021.

Inducers of the endothelial cell barrier identified through chemogenomic screening in genome-edited hPSC-endothelial cells. *PNAS*, 2020.

Multiscale modelling of drug mechanism and safety. *Drug Discovery Today*, 2020.

Assessment of network module identification across complex diseases. *Nature Methods*, 2019.

Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics*, 2019.

A novel orally available small molecule that inhibits hepatitis B virus expression. *Journal of Hepatology*, 2018.

Detect tissue heterogeneity in gene expression data with BioQC. *BMC Genomics*, 2017.

Molecular phenotyping combines molecular information, biological relevance, and patient data to improve productivity of early drug discovery. *Cell Chemical Biology*, 2017.

14-3-3 $\zeta$  turns TGF- $\beta$ 's function from tumor suppressor to metastasis promoter in breast cancer by contextual changes of smad partners from p53 to Gli2. *Cancer Cell*, 2015.

White-to-brown metabolic conversion of human adipocytes by JAK inhibition. *Nature Cell Biology*, 2015.

Data mining reveals a network of early-response genes as a consensus signature of drug-induced *in vitro* and *in vivo* toxicity. *The Pharmacogenomics Journal*, 2013.

KEGGgraph: a graph approach to KEGG PATHWAY in R and Bioconductor. *Bioinformatics*, 2009.

→ See the full, manually curated list at Google Scholar (<http://goo.gl/CoeJu7>).

---

## Selected open-source software

BESCA Single-cell omics data analysis pipeline, built together with BEDA colleagues.

BioQC Detecting tissue heterogeneity in high-throughput expression data.

KEGGgraph Data mining and network analysis of biological pathways as graphs in R and Bioconductor.

---

## Computer skills

Programming	R/Bioconductor, C/C++, Java, Python	Database	PostgreSQL, SQLite, MongoDB
Scripting	Python, Bash, Erlang	Web	HTML, CSS, JavaScript, Flask, FastAPI
OS	Debian Linux, Windows	Others	d3js, OpenCV, ImageMagick, L <sup>A</sup> T <sub>E</sub> X

---

## Personal information

Date of Birth	Sept. 28 <sup>th</sup> , 1983 in Tianjin, China
Marital status	Married, father of two daughters
Hobbies	Family, run & bike, reading, mathematics, programming, music
Languages	English, German, Chinese