
Key facts

- Computational biologist with 10+ years' industrial research experience
- Key contributions to drug candidates in development and marketed products
- Inventing and implementing proprietary technologies and workflows
- Establishing a new research group on computational toxicology and safety
- Author of patents, open-source software, and publications (*h*-index 29)
- Teaching two courses and supervising postdocs, students, and apprentices

Research and teaching

2011– **Expert 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, the Pathway-Annotated Chemical Ensemble (PACE) library, and new methods for computational toxicology and safety assessment.

2018– **Lecturer**, *Department of Mathematics and Computer Science, University of Basel*

Teaching the lecture series *Introduction to Applied Mathematics and Informatics In Drug Discovery* (<http://AMIDD.ch>) in autumn semesters, and *Mathematical and Computational Biology in Drug Discovery* (<http://MCBDD.ch>) in spring semesters.

Training, mentoring, and academic commitment

2022– Co-organizer of Roche PMDA (Predictive Modeling and Data Analytics) Summer School for PhD students.

2021– Certified *Lehrmeister* for informatics vocational apprenticeship.

2020– Certified Software and Data Carpentry instructor, hosting company internal carpentries.

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.

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

- [1] Optimization of the teratox assay for preclinical teratogenicity assessment. *Toxicological Sciences*, 2022.
- [2] Besca, a single-cell transcriptomics analysis toolkit to accelerate translational research. *NAR Genomics and Bioinformatics*, 2021.
- [3] Inducers of the endothelial cell barrier identified through chemogenomic screening in genome-edited hPSC-endothelial cells. *PNAS*, 2020.
- [4] Multiscale modelling of drug mechanism and safety. *Drug Discovery Today*, 2020.
- [5] Assessment of network module identification across complex diseases. *Nature Methods*, 2019.
- [6] Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics*, 2019.
→ See the full, manually curated list at Google Scholar (<http://goo.gl/CoeJu7>).

Selected invited talks

- 2021 *Optimization of the TeraTox assay for preclinical teratogenicity assessment*, co-presentation with Manuela Jaklin, OpenTox Virtual Conference, 2021
- 2019 *Bioinformatics and exploratory data analysis in drug discovery: an industrial perspective*, ISMB/ECCB, 2019, Basel, Switzerland
- 2018 *Mathematics in drug discovery: a practitioner's view*, Perlen-Kolloquium, University of Basel, Basel, Switzerland

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.

Supervised postdocs

Davide Bassani	2023- (co-supervision with Dr. Neil Parrott and Dr. Nenad Manevski), computational approaches to preclinical pharmacokinetic property prediction
Milad Adibi	2020-2022 (co-supervision with Dr. Ekaterina Breous-Nystrom), multiscale modeling of drug-induced liver toxicity, currently senior computational biologist at University of Zurich
Simon Gutbier	2018-2020 (co-supervision with Dr. Christoph Patsch and Dr. Markus Britschgi), immune pathway characterization with tool-compound screening for Parkinson's Disease, currently principal scientist at Roche

Supervised master and PhD students

Alexandra Valeanu	2023, internship on inferring causal kinases with phosphoproteomics data
Simon Crouzet	2022, internship on predicting protein-ligand interaction with graph neural networks, currently PhD student in Dal Peraro's group at EPFL
Anja Lieberherr	2022-2023 (co-supervision with Prof. Niko Beerenwinkel), master thesis on applying graph neural networks in drug discovery, currently junior technology consultant at BearingPoint, Zürich
Sarah Leonardo Morillo	2021-2022, internship on proteomics data and protein-half life, currently PhD student in Erik van Nimwegen's group at Biozentrum, University of Basel
Andreea Ciuprina	2020 (co-supervision with Prof. Niko Beerenwinkel), internship and master thesis on computational inference of immune-cell contribution to drug-induced liver toxicity, currently data engineer and scientist at Endress+Hauser Flowtec
Rudolf Biczok	2018-2019 (co-supervision with Prof. Alexandros Stamatakis), internship and master thesis on building a gene-expression database and evaluating gene-set comparison metrics, currently full-stack Java architect and Azure cloud specialist at Bank for International Settlements (BIS)
Moaraj Hasan	2017-2018, internship on automating gene-expression data analysis
Tao Fang	2017-2018 (co-supervision with Prof. Mark Robinson), internship and master thesis on histopathology prediction with deep neural networks, currently postdoctoral fellow in von Mering group at University of Zurich
Gregor Sturm	2016-2017, master intern working on gene signatures and the BioQC software, currently Scientist Clinical Bioinformatics, Boehringer Ingelheim

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 ^A T _E X

Personal information

Date of Birth	Sept. 28 th , 1983 in Tianjin, China
Marital status	Married, father of two daughters
Languages	English, German, Chinese