ANATOLY BUCHIN, PhD

Computational Biology | Machine Learning | Drug Discovery

+1 (206) 488 23 43
anat.buchin@gmail.com
Seattle, United States
www.buchin.info
www.github.com/abuchin

TECHNICAL SKILLS

Machine Learning & AI
Deep learning, Large Language Models (LLMs), Generative models, Bayesian inference, Statistical learning, Variational Autoencoders

Machine Learning Frameworks
PyTorch, TensorFlow, Keras, scikit-learn, scvi-tools, Hugging Face, Transformers

Computational Biology & Bioinformatics
Single-cell & bulk RNA-seq, Omics integration, Sequence analysis, Genomic alignment, Proteomics, Multi-omics data integration

Programming languages
Python, R, Matlab, Bash

Model development
MLOps, CI/CD, Data pipelines, Unit testing, Model deployment, Feature stores (Feast)

Big Data & Cloud Computing
AWS (S3, EC2, Lambda, SageMaker), Google Cloud, Distributed computing, Parallel processing, High-performance computing (HPC), GPU acceleration

Unix/Linux & DevOps
Shell scripting, Git/GitHub, Docker, Singularity, Workflow automation

Distributed and cloud computing
AWS Cloud, SageMaker, Google Cloud, HPC, GPU

Workflow Managers & Pipelines
Snakemake, Nextflow, Airflow

Database & Data Engineering
SQLalchemy, PostgreSQL, Athena, Pandas, Spark

Visualization & Analysis
Matplotlib, Seaborn, Plotly, Scanpy, IGV-browser

Development Tools
VS Code, Cursor, Anaconda, JupyterLab

PROFILE

Computational Biologist and Machine Learning Scientist with 10+ years of experience developing AI-driven solutions for drug discovery and multi-omics data integration. Proven track record of building scalable cloud pipelines, generative models, and predictive frameworks applied to large-scale single-cell datasets (>80M profiles). Experienced in translating advanced machine learning research into production systems supporting precision medicine, oncology, and neurodegeneration programs. Strong collaborator across biology, data science, and engineering teams, with publications in Nature, Neuron, and Cell Reports.

INDUSTRIES: Drug discovery, Biotechnology, GenAI

PRINCIPAL SCIENTIST - Machine Learning and Bioinformatics

SandboxAQ | Seattle | October 2025 - Present

- Architected cloud-native **multi-omics data pipeline** integrating RNA-seq and WGS datasets to enable scalable precision oncology discovery and biomarker development.
- Designed and deployed an **AI-driven antigen prioritization** framework to optimize immunogenic payload selection for glioblastoma vaccine programs.
- Developed a scalable **Knowledge Graph-powered RAG** platform that integrates heterogeneous biomedical datasets enabling evidence-grounded natural-language interrogation of Parkinson's disease biology.
- Provided **technical leadership** across cross-functional teams (computational biology, ML, and translational research) to define AI strategy for immuno-oncology discovery initiatives.

SENIOR COMPUTATIONAL BIOLOGIST - AI/ML for Computational Biology

Synthesize Bio | Seattle | August 2024 - June 2025

- Co-developed **foundational AI models** for generating synthetic gene expression data across human tissues, achieving >0.63 Pearson correlation with observed values.
- Led design and scaling of cloud-based transcriptomics pipelines processing over **80M** single-cell profiles using AWS, EC2, and Athena.
- Developed benchmarking framework across **100M** cells for model generalization over drug, cell line, and tissue domains.
- Collaborated with biologists, ML engineers, and software teams to deploy production-ready models on internal platforms **https://product.synthesize.bio/**

SENIOR SCIENTIST - Computational Biology and Drug Discovery

Cajal Neuroscience | Seattle | May 2021 - June 2024

- Built ML models to predict cellular states from transcriptomic features, supporting drug target validation with 83% classification accuracy.
- Developed and maintained RNA-seq **data pipelines** (bulk/single-cell) across 20+ datasets; improved preprocessing efficiency by 13%.
- Integrated **6TB**+ of transcriptomics, proteomics, and genomics data to inform neurodegenerative disease programs.
- Created a pseudotime inference algorithm correlating gene dynamics with disease progression (r=0.34), presented at **ADPD2022**, **ADPD2023**

EDUCATION

PhD in Computational Neuroscience, École Normale Supérieure, Paris
2012 – 2015

Master of Physics,
Peter the Great Polytechnic University, St Petersburg
2010 – 2011

Master of Biomedical Research,
Descartes University, Paris
2009 – 2010

Bachelor of Physics,
Peter the Great Polytechnic University, St Petersburg,
2005 – 2009

AWARDS

2017 Assistant professor in neuroscience (France)

2016 NSF Travel grant

2016 Swartz Foundation fellowship

2014 Foundation of Medical Research grant

LANGUAGES

ENGLISH – Full proficiency

FRENCH – Proficient

GERMAN – Proficient

RUSSIAN – Native speaker

MENTORSHIP

2024, Jayasree Peri,
Cedars-Sinai Medical Center,
Master student

2022, Temitope Adeoye,
University of South Florida, PhD student

2021, Sasha Batoukova, Tesla high school student

REFERENCES

Ben Logsdon
Cajal Neuroscience

Costas Anastassiou
Cedars-Sinai Medical Center

Adrienne Fairhall
University of Washington

Boris Gutkin
École Normale Supérieure

RESEARCH SCIENTIST - Data Science & Computational Modeling

Allen Institute for Brain Science | Seattle | April 2017 - March 2021

- Designed ML and deep learning models (CNNs, ensemble classifiers) for cell type prediction using transcriptomic and morphometric features (90%+ accuracy) (**Nature**, **Cell Reports**, **Neuron**).
- Scaled **biological simulations** and model **parameter optimization** across HPC clusters and GPU compute environments.
- Developed bio-realistic computer simulations using HPC to characterize epileptic oscillations in biological neural networks (>**100K** neurons).

POST-DOCTORAL RESEARCHER - Mathematical modeling and Data analysis

University of Washington, | Seattle | February 2016 – April 2017

- Analyzed **time series** electrical recordings from epileptic human brain slices using computational modeling and non-linear dynamics (**eNeuro**).
- Developed novel **neural network model** to simulate animal behavior of Hydra vulgaris.
- Implemented **data analysis pipelines** for in vivo calcium imaging data and applied motion tracking algorithms for video behavioral data.

DOCTORAL RESEARCHER - Computational Neuroscience

École Normale Supérieure, | Paris | January 2012 – November 2015

- Developed biophysical **computational models** for single neurons and neural networks in the human brain.
- Applied dynamical systems theory to explain brain dynamics in human epilepsy and rodent motor systems (**Journal of Neuroscience**, **PLOS Computational Biology**).
- Developed a **multi-scale model** combining reaction-diffusion dynamics with spiking neural network to explain the mechanisms of human epilepsy.

SELECTED PUBLICATIONS

Buchin A., et al. (2022). Multi-modal characterization and simulation of human epileptic circuitry. Cell Reports. **(Paper | Code | Web product)**

Wei Y., **Buchin A.**, et al. (2023). Associations between in vitro, in vivo and in silico cell classes in mouse primary visual cortex. Nature communications.
(Paper | Code)

Nandi A., **Buchin A.** et al. (2022). Single-neuron models linking electrophysiology, morphology and transcriptomics across cortical cell types. Cell Reports.
(Paper | Code | Web Product)

Berg, J., **Buchin A.** et al. (2021). Human cortical expansion involves diversification and specialization of supragranular intratelencephalic-projecting neurons. Nature. **(Paper | Code)**

Kalmbach K. E., **Buchin A.** et al. (2018). H-channels contribute to divergent electrophysiological properties of supragranular pyramidal neurons in human versus mouse cerebral cortex. Neuron. (**Paper | Code)**

Buchin A., et al. (2018). Adaptation and inhibition control pathological synchronization in a model of focal epileptic seizure. eNeuro, 0019-18.2018. (**Paper | Code)**

Buchin A. et al. (2016). Reduced Efficacy of the KCC2 Cotransporter promotes epileptic oscillations in a subiculum network model. Journal of Neuroscience. (**Paper | Code)**

Buchin A. et al. (2016). Inverse stochastic resonance in cerebellar Purkinje cells. PLOS Computational Biology. (**Paper | Code)**

SELECTED CONFERENCE PROCEEDINGS

Buchin A., et al. (2023). Lineage tracing and differential expression analysis of single nuclei RNA-seq data from human neocortex identifies novel genes and pathways involved into progression of Alzheimer's disease. Alzheimer's and Parkinson's disease conference. Gothenburg, Sweden.

Buchin A., et al. (2022). Lineage tracing and differential expression analysis of single cell RNA-sequencing data from human midbrain identifies novel genes affected by idiopathic Parkinson's disease. Alzheimer's and Parkinson's disease conference. Barcelona, Spain.

Buchin A., et al. (2019). Conserved and divergent electrophysiological and morphological properties of mouse and human transcriptomically-defined cell types. Chicago, United States.

Buchin A., et al. (2018). Morpho-electric properties and computational simulation of human dentate gyrus granule cells from the epileptogenic hippocampus. Society for Neuroscience. San Diego, United States.