

BRIAN HIE

brianhie@stanford.edu | github.com/brianhie | brianhie.com

Interests: Machine learning, computational biology, protein engineering and evolution, systems biology.

CURRENT POSITIONS

- | | |
|---|--------------|
| Stanford University School of Medicine , Palo Alto, CA | |
| <i>Stanford Science Fellow</i> | 2021-Present |
| ♦ Machine learning at the host-pathogen interface. | |
| Meta Platforms, Inc. , Menlo Park, CA | 2022-Present |
| <i>Visiting Researcher, Meta AI</i> | |
| ♦ Machine learning for protein biology. | |

EDUCATION

- | | |
|--|-----------|
| Massachusetts Institute of Technology , Cambridge, MA | |
| <i>Electrical Engineering and Computer Science, Doctor of Philosophy</i> | 2017-2021 |
| <i>Electrical Engineering and Computer Science, Master of Science</i> | |
| ♦ Areas of concentration: Computational biology, machine learning, statistics. | |
| Stanford University , Palo Alto, CA | |
| <i>Computer Science, Bachelor of Science with Honors and Distinction</i> | 2012-2016 |
| <i>English Literature, Academic Minor</i> | |
| ♦ Areas of concentration: Computational biology, computer systems, machine learning. | |

PROFESSIONAL EXPERIENCE

- | | |
|--|-----------|
| Massachusetts Institute of Technology , Cambridge, MA | |
| <i>Graduate Researcher, Computer Science and Artificial Intelligence Laboratory</i> | 2017-2021 |
| ♦ Neural language modeling of viral evolution. | |
| ♦ Machine learning for biological discovery and design under uncertainty. | |
| ♦ Efficient geometric algorithms for single-cell biology. | |
| ♦ Cryptographically secure neural network training. | |
| Google LLC , Mountain View, CA | |
| <i>Artificial Intelligence/Machine Learning Resident, X – The Moonshot Factory</i> | 2019 |
| ♦ Machine learning for early-pipeline moonshots. | |
| Illumina, Inc. , San Diego, CA | |
| <i>Machine Learning Intern, Bioinformatics</i> | 2018 |
| ♦ Statistical signal processing for genomics-based health monitoring. | |
| Salesforce.com, Inc. , San Francisco, CA | |
| <i>Software Engineer, Cloud Infrastructure</i> | 2016-2017 |
| ♦ Robust performance monitoring of globally distributed core application infrastructure. | |
| Stanford University , Palo Alto, CA | |
| <i>Undergraduate Researcher, Biology</i> | 2013-2016 |
| ♦ Statistics and machine learning for computational genomics. | |

Stanford University, Palo Alto, CA

Undergraduate Researcher, Digital Humanities, Stanford Humanities Center 2016

- ♦ Graph-theoretic analysis of the social network of early modern authors and publishers.

Microsoft Corporation, Redmond, WA

Software Engineering Intern, Azure Compute and Microsoft Research 2015

- ♦ Distributed scheduling algorithms and their impact on data center utilization and availability.

Synaptics, Inc., San Jose, CA

Systems Architecture/Algorithms Intern 2014

- ♦ Algorithm design and implementation for embedded touchscreen firmware.

PUBLICATIONS

*Equal contribution. †Corresponding author.

Journal articles

B. Hie[†], D. Xu, V. Shanker, T. Bruun, P. Weidenbacher, S. Tang, and P.S. Kim[†].

"Efficient evolution of human antibodies from general protein language models and sequence information alone."

bioRxiv (in press at *Nature Biotechnology*), DOI: 10.1101/2022.04.10.487811. 2023

Z. Lin*, H. Akin*, R. Rao*, **B. Hie***, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y.

Shmueli, A. dos Santos Costa, M. Fazel-Zarandi, T. Sercu, S. Candido, and A. Rives.

"Evolutionary-scale prediction of atomic level protein structure with a language model."

Science, 379:6637. 2023

B. Hie[†], K.K. Yang, and P.S. Kim[†].

"Evolutionary velocity with protein language models predicts evolutionary dynamics of diverse proteins."

Cell Systems, 13:4 (featured article). 2022

M.C. Maher, I. Bartha, S. Weaver, J. di Iulio, E. Ferri, L. Soriaga, F.A. Lempp, **B. Hie**, B.

Bryson, B. Berger, D.L. Robertson, G. Snell, D. Corti, H.W. Virgin, S.L. Kosakovsky Pond,

and A. Telenti.

"Predicting the mutational drivers of future SARS-CoV-2 variants of concern."

Science Translational Medicine, 14:633. 2022

B. Hie and K.K. Yang.

"Adaptive machine learning for protein engineering."

Current Opinion in Structural Biology, 72: February 2022. 2022

R. Singh*, **B. Hie***, A. Narayan, and B. Berger.

"Schema: Metric learning enables interpretable synthesis of heterogeneous single-cell modalities."

Genome Biology, 22:131. 2021

B. Hie, E. Zhong, B. Berger, and B. Bryson.

"Learning the language of viral evolution and escape."

Science, 371:6526 (featured article). 2021

B. Hie, B. Bryson, and B. Berger.
"Leveraging uncertainty in machine learning accelerates biological discovery and design."
Cell Systems, 11:5. 2020

B. Hie*, J. Peters*, S. Nyquist*, A. Shalek, B. Berger, and B. Bryson.
"Computational methods for single-cell RNA sequencing."
Annual Review of Biomedical Data Science, 3:1. 2020

B. Hie*, H. Cho*, B. DeMeo, B. Bryson, and B. Berger.
"Geometric sketching compactly summarizes the single-cell transcriptomic landscape."
Cell Systems, 8:6 (cover article). 2019

B. Hie, B. Bryson, and B. Berger.
"Efficient integration of heterogeneous single-cell transcriptomes using Scanorama."
Nature Biotechnology, 37:6. 2019

A. Tehranchi, **B. Hie**, M. Dacre, I. Kaplow, K. Pettie, P. Combs, and H. Fraser.
"Fine-mapping cis-regulatory variants in diverse human populations."
eLife, 8:e39595. 2019

B. Hie*, H. Cho*, and B. Berger.
"Realizing private and practical pharmacological collaboration."
Science, 362:6417. 2018

A. Tehranchi, M. Myrthil, T. Martin, **B. Hie**, D. Golan, and H. Fraser.
"Pooled ChIP-seq links variation in transcription factor binding to complex disease risk."
Cell, 165:3. 2016

Conference papers

C. Hsu, R. Verkuil, J. Liu, Z. Lin, **B. Hie**, T. Sercu, A. Lerer, and A. Rives.
"Learning inverse folding from millions of predicted structures."
International Conference on Machine Learning (ICML). 2022

B. Hie, E. Zhong, B. Bryson, and B. Berger.
"Learning mutational semantics."
Neural Information Processing Systems (NeurIPS). 2020

B. Hie, H. Cho, B. DeMeo, B. Bryson, and B. Berger.
"Geometric sketching of single-cell data preserves transcriptional structure."
Research in Computational Molecular Biology (RECOMB). 2019

Preprints

B. Hie*, S. Candido*, Z. Lin, O. Kabeli, R. Rao, N. Smetanin, T. Sercu, and A. Rives.
"A high-level programming language for generative protein design."
bioRxiv, DOI: 10.1101/2022.12.21.521526. 2022

C. Itoh, C. Gunnarson, G. Babunovic, A. Nibasumba, Ngomu A., M. Wadsworth III, T. Hughes II, S. Solomon, **B. Hie**, B. Berger, A. Shalek, S. Fortune, and B. Bryson.

"GM-CSF differentiation of human monocytes stabilizes macrophage state via oxidative signaling."

bioRxiv, DOI: 10.1101/2020.09.29.318352.

2020

SOFTWARE

Scanorama, Primary Developer

<https://github.com/brianhie/scanorama>, 90k+ PyPI downloads

Geosketch, Primary Developer

<https://github.com/brianhie/geosketch>, 63k+ PyPI downloads

Evolocity, Primary Developer

<https://github.com/brianhie/evolocity>, 10k+ PyPI downloads

esm, Contributor

<https://github.com/facebookresearch/esm>

scverse, Contributor

<https://github.com/scverse>

TEACHING

Massachusetts Institute of Technology, Cambridge, MA

Teaching Assistant, Algorithms for Inference (6.438)

2019

Graduate-level course on statistical inference with probabilistic graphical models. Responsible for preparing exams/assignments, leading discussion sections, and holding office hours.

ACADEMIC SERVICE

Mentorship

Stanford University, Doctoral thesis research

2021-Present

University of Toronto, iGEM competition team

2021

Massachusetts Institute of Technology, Masters of Engineering thesis research

2020-2021

Public Engagement and Science Communication

Press interviews for *Al Jazeera*, *Freethink*, *Ideas Roadshow*, *IEEE Spectrum*, *Inverse*, *MIT Technology Review*, *NSF The Discovery Files*, *Nautilus*, *Quanta Magazine*, *Swiss Radio*, *Wall Street Journal*, and *Wired*

Peer Review

Contributed reviews to *Bioinformatics*, *BMC Bioinformatics*, *Cell*, *Cell Systems*, *Frontiers in Genetics*, *GigaScience*, *Intelligent Systems for Molecular Biology (ISMB)*, *Journal of Molecular Biology*, *Nature*, *Nature Biotechnology*, *Nature Communications*, *Nature Methods*, *NeurIPS Workshop on Machine Learning in Structural Biology (MLSB)*, *Nucleic Acids Research*, *PLOS Computational Biology*, *PLOS ONE*, *PNAS*, *PNAS Nexus*, *Research in Computational Molecular Biology (RECOMB)*, and *Science Translational Medicine*

PATENTS

A. Thugabere Jagadeesh and **B. Hie**.

"Hyperspectral scanning to determine skin health."

US Patent No. 11,532,400.

2022

B. Hie, B. Berger, and H. Cho.

"Realizing private and practical pharmacological collaboration using a neural network architecture configured for reduced computation overhead."

US Patent No. 11,450,439. 2022

B. Hie, B. Bryson, and B. Berger.

"Escape profiling for therapeutic and vaccine development."

US Patent No. 11,011,253. 2021

H. Ma, **B. Hie**, and B. Ni.

"Quality control in electronic nose sensing."

US Patent App. 16/738,586. 2020

H. Ma, **B. Hie**, and B. Ni.

"Analyte classification using electronic noses."

US Patent App. 16/737,648. 2020

AWARDS AND FELLOWSHIPS

STAT Wunderkinds 2022

Stanford Science Fellows Program 2021-Present

Department of Defense (DoD) National Defense Science and Engineering Graduate (NDSEG) Fellowship 2019-2021

RECOMB/National Science Foundation Travel Fellowship Award 2019

Hoefer Prize for Writing in the Major, Nominated, *Stanford University* 2016

Tau Beta Pi Engineering Honors Society 2015-Present

Lunsford Award for Oral Presentation, Nominated, *Stanford University* 2014

Boothe Prize for Excellence in Writing, Honorable Mention, *Stanford University* 2013

President's Award for Academic Achievement, *Stanford University* 2013

National Merit Scholarship Finalist 2012