Aaron N. Brooks Ph.D.

AI and agentic systems · Synthetic biology · Genomics

Albuquerque, NM

scalefreegan.github.io aaron-n-brooks scalefreegan Google Scholar

Scientific leader building AI-native tools that transform and accelerate the life sciences. I design and ship enterprise-scale agentic AI systems, drawing on a research career in synthetic genomics, gene regulation and machine learning. My long-term goal is to design genomes from scratch, pairing AI with tight design–build–test–learn cycles.

EXPERIENCE

Principal Scientist, Applied AI Nov 2024 – present
Generate:Biomedicines · Boston, MA (remote)

- **Shipped agentic products that reshape scientific workflows across the company.**
- Architected and deployed enterprise-scale agentic infrastructure: an agent SDK, MCP gateway, evaluation framework, CLI tooling, and a memory system that enables iterative self-improvement.
- Delivered human-in-the-loop agent workflows that automate time-consuming, business-critical tasks, such as generating on-brand executive presentations from raw scientific results.
- Built an extensible digital-twin simulation framework that analyzes hundreds of thousands of DNA variants in parallel, compressing research timelines from weeks to hours and surfacing high-value design candidates.

Associate Director, Genetic Design Jan 2023 – Jun 2024
Ginkgo Bioworks · Boston, MA (remote)

- **Founded a high-performing team and a new business function that streamlined DNA asset management, enabling enterprise-wide collaboration and faster innovation cycles.**
- Prototyped AI-assisted knowledge management systems using retrieval-augmented generation (RAG) and other embedding approaches.
- Predicted DNA synthesizability with machine learning (gradient boosting, random forest, logistic regression).
- Fine-tuned a DNA foundation model (HyenaDNA) for application-specific regression and classification tasks.

Staff Scientist, Data Science Dec 2020 – Dec 2022
Inscripta · Boulder, CO (remote)

- **Enabled and contributed to 4 customer publications and 3 internal patent applications.**
- **Led a team that designed and implemented a biological knowledge graph for target selection.**
- Developed statistical methods to interpret highly multiplexed selection experiments in CRISPR-engineered cell libraries.
- Built an interactive web service (Dash) for submitting deep mutational scanning data for analysis and visualization.
- Designed promoter-insertion DNA libraries using a pre-trained convolutional neural network (CNN).

Research Scientist, Steinmetz Lab Jun 2015 – Nov 2020
European Molecular Biology Laboratory (EMBL) · Heidelberg, Germany

- **Secured €1.4M in funding; published 4 papers in top journals, including *Science* and *Cell*, and 1 patent application.**
- **Conceived and led the funding and regulatory approval of a clinical study that used wearable technologies to evaluate protracted symptoms in patients with mitochondrial disorders outside clinical settings.**
- Processed and analyzed terabytes of third-generation sequencing reads with a custom end-to-end pipeline (Snakemake).
- Applied machine learning (gradient boosting) to disentangle factors influencing transcript start and end sites in *S. cerevisiae*.

SKILLS

AI and agentic systems	MCP, agent frameworks, PyTorch, scikit-learn, Hydra; Claude Code, Cursor, Codex
Computational biology	DNA/RNA-seq, multi-omics, genome assembly, variant analysis
Software engineering	Python, R, Git, Docker; Snakemake, Nextflow; Dash, Django, Shiny
Data and cloud	SQL/NoSQL, Neo4j, data lakes; AWS, GCP
Languages	German (A2), Spanish

SELECTED PUBLICATIONS

13 publications · 852 citations · h-index 12 · * equal contribution

Transcriptional neighborhoods regulate transcript isoform lengths and expression levels 2022
Science · **AN Brooks***, AL Hughes*, S Clauder-Münster, LA Mitchell, JD Boeke, LM Steinmetz

A system-level model for the microbial regulatory genome 2014
Mol Syst Biol · **AN Brooks***, DJ Reiss*, A Allard, W Wu, DM Salvanha, CL Plaisier, S Chandrasekaran, M Pan, A Kaur, NS Baliga

Design, construction, and functional characterization of a tRNA neochromosome in yeast 2023
Cell · D Schindler, RSK Walker, S Jiang, **AN Brooks**, …, LM Steinmetz, JD Boeke, Y Cai

Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions 2023
Cell · Y Zhao, C Coelho, AL Hughes, L Lazar-Stefanita, S Yang, **AN Brooks**, …, LM Steinmetz, YP Cai, JD Boeke

Deep mutational scanning of essential bacterial proteins can guide antibiotic development 2023
Nat Commun · L Dewachter, **AN Brooks**, K Noon, N Krishnamurthy, W Versées, W Vranken, J Michiels

PATENTS

Methods for isolating and/or obtaining captured polynucleotide fragments · US Patent App. 18/501,530 2024

Genome-wide rationally-designed mutations leading to enhanced cellobiohydrolase I production in *S. cerevisiae* · WO2023028521A1 2023

System and method for gene editing cassette design · US20220246235A1 2022

EDUCATION

Ph.D., Molecular and Cellular Biology 2008 – 2014
University of Washington · Seattle, WA

- Advisor: Nitin Baliga. Dissertation: *Data-driven inference of dynamic transcriptional regulatory mechanisms in prokaryotes: a systems perspective*

B.S., Biochemistry · B.A., Political Science · Minor, Philosophy 2002 – 2007
University of New Mexico · Albuquerque, NM

HONORS AND AWARDS

EIPOD Postdoctoral Fellowship
Marie Skłodowska-Curie, EMBL · 2015

Graduate Fellowship
DOE Office of Science · 2010

Goldwater Scholar
Barry M. Goldwater Foundation · 2006

PROJECTS AND CONTINUING EDUCATION

Personal projects. I build AI-native apps end to end with coding agents, including tools for learning formal proof with Lean 4, music theory and spoken languages. Details at scalefreegan.github.io.

Certifications (2024). Deep Learning Specialization (DeepLearning.AI), Deep Learning Fundamentals (Lightning AI), and Neo4j application and LLM courses. Full list on LinkedIn.