

# Michael Cariaso

Principal Bioinformatics Software Engineer

Bethesda, MD · cariaso@gmail.com · [github.com/cariaso](https://github.com/cariaso) · [linkedin.com/in/cariaso](https://linkedin.com/in/cariaso)

## PROFESSIONAL SUMMARY

---

Bioinformatics software engineer with 20+ years of experience designing and delivering production genomics software, cloud infrastructure, and large-scale DNA analysis platforms. Began on the Human Genome Project at Lawrence Livermore National Laboratory and has since developed software spanning consumer genetics, DNA synthesis, CAR-T therapeutics, agricultural genomics, and forensic science.

Cofounder of River Road Bio and sole software developer behind SNPedia and Promethease, creating widely used personal genome interpretation platforms acquired by MyHeritage. Career highlights include coauthoring a cover-featured *Nature* paper, being featured by the AWS Blog for pioneering cloud-based BLAST analysis, and building production software for virtually every major DNA sequencing platform, including Illumina, PacBio, Oxford Nanopore, ABI 3730, SOLiD, Ion Torrent, Helicos, and 454 sequencing systems.

## CORE COMPETENCIES

---

|                        |                                                                      |
|------------------------|----------------------------------------------------------------------|
| <b>Languages:</b>      | Python, Bash, JavaScript, Perl, SQL, C/C++, Rust, Lua, PHP           |
| <b>AWS Cloud:</b>      | EC2, Batch, S3, Lambda, DynamoDB, Bedrock, HealthOmics               |
| <b>Workflows:</b>      | Nextflow, Docker, SGE, PBS, MPI, Celery, Linux processes & pipes     |
| <b>Sequencing:</b>     | Illumina, PacBio, Oxford Nanopore, Affymetrix & Illumina microarrays |
| <b>Bioinformatics:</b> | Variant calling & annotation, genome assembly, NCBI & EMBL databases |
| <b>Databases:</b>      | PostgreSQL, MySQL/MariaDB, Oracle, SQLite, SQLAlchemy, Redis         |
| <b>AI/ML:</b>          | scikit-learn, NLP, LLMs, RAGs                                        |

## PROFESSIONAL EXPERIENCE

---

**Elegen Bio** — Staff Bioinformatics Developer *San Carlos, CA* 4/2024–6/2026

- Developed production sequencing and laboratory automation pipelines for DNA synthesis validation.
- Created novel Python & JavaScript tooling supporting Oxford Nanopore–based plasmid workflows.
- Sole developer of software for a new PacBio platform, enabling diagnosis of technical issues invisible on ONT and helping preserve a vital strategic collaboration with GSK.

**Othram** — Bioinformatics Consultant *Houston, TX* 1/2024–4/2024

- Human forensic cold-case criminal investigations with degraded samples. [PMID 40772272](https://pubmed.ncbi.nlm.nih.gov/40772272/)

**Serotiny** — Senior Bioinformatics Scientist *South San Francisco, CA* 9/2022–7/2023

- Developed sequencing analysis pipelines supporting CAR-T and recombinase engineering.
- Built scalable analysis workflows for Illumina, PacBio, and Oxford Nanopore data.

**MyHeritage** — VP, Genetic Health R&D *Israel* 7/2017–2/2020

- After acqui-hire of River Road Bio (see below), led development of consumer genetic health products.
- Directed engineering of production-scale genomic interpretation software.

**River Road Bio** — Founder & Sole Developer *Maryland* 7/2006–7/2017

- Created SNPedia, a public genotype–phenotype knowledge base, enabling individuals to interpret their personal genomic data.

- Created Promethease a personal genomics interpretation tool, generating millions of reports across dozens of platforms.
- Profiled in [MIT Technology Review](#) for creating widely used human DNA tools.

**KeyGene** — Bioinformatician

*Netherlands & Washington, DC* 7/2009–6/2017

- Bioinformatic analysis of numerous plant genomes; experience with polyploids, large genomes, quantitative phenotypes, and selfing.
- Coauthor on the Tomato Genome ([Nature cover, PMID 22660326](#)).
- Created commercially successful species-specific databases integrating multi-omics data (WGS, WES, markers, annotations, publications).
- Introduced and administered AWS cloud for ~150 employees; bootstrapped the USA office.

**Kailos Genetics** — Bioinformatics Consultant

*Huntsville, AL* 9/2014–8/2015

- Built scikit-learn–based tools for DOD forensic ancestry and phenotype prediction.

**Bioteam** — Bioinformatician

*Boston, MA* 7/2007–7/2009

- Became the company's most billable employee within the first year.
- Contracted to a federal client to address the unique data challenges of the Anthrax investigation and early access to the Solexa/Illumina Genome Analyzer II — the first NGS platform to produce data at a scale that made existing LIMS software obsolete.
- Adapted the MediaWiki software stack into a custom LIMS; enthusiastically adopted, then extended with quality visualizations and support for Helicos and ABI SOLiD platforms.
- Introduced AWS cloud to the BioTeam; wrote about it for [Bio-IT World](#).

**RunBlast** — Personal Project

*Bethesda, MD* 5/2006–7/2007

- Ran large-scale mpiBLAST on EC2 at a time when cloud computing was novel; featured on the [AWS Blog](#) ( [https://aws.amazon.com/blogs/aws/amazon\\_ec2\\_for\\_/](https://aws.amazon.com/blogs/aws/amazon_ec2_for_/) )

**Celera** — Bioinformatician

*Rockville, MD* 4/2004–1/2005

- Inherited 100k lines of Perl sequence-handling pipelines after the departure of Craig Venter and most of the original 10 person team; kept everything running smoothly.
- Added detection of indels and bisulfite-treated methylation.
- Leveraged Primer3 for automated design of a minimal primer set for heavily alternatively spliced genes.

**SAIC** — Bioinformatician

*Rockville, MD* 1/2003–1/2005

- Built an in-house cluster from scratch (ordered servers, cut CAT5 cables, installed OS) for a DOD-funded project to identify unique genomic regions of bacterial and viral pathogens.
- Designed the algorithm, wrote the code, and created the visualizations.

**Gene Logic** — Bioinformatician

*Gaithersburg, MD* 7/1998–1/2003

- ABI 3730–based sequence identification and analysis pipelines.
- Affymetrix microarray–based database collaboration with GSK.

**Lawrence Livermore National Laboratory** — Bioinformatician

*Livermore, CA* 3/1997–6/1998

- Human Genome Project: expressed sequence tag–based minimal catalog of all known human genes via massive all-vs-all BLAST
- [PMID 10745985](#) Clustering and ranking of I.M.A.G.E. cDNA clones corresponding to known genes

**EDUCATION**

**University of Florida** — B.S. Computer Engineering

1998