ADRIEN OLIVA

Senior Bioinformatician · Data Scientist · Agentic AI Builder

PhD · 8+ Years · 6 First-Author Q1 Publications · $3.2M+ Funded Research

📍 Brisbane, QLD, Australia ☎ +61 433 098 609 ✉ adrioliva@hotmail.fr
🌐 adrien-oliva.com 🔗 linkedin.com/in/adrienoliva ✨ github.com/AdrienOliva
🎓 DOB: 23 April 1993 NZ AU Nationality: French & Australian (dual) 🗣 Languages: French (native) · English (fluent) · Spanish (beginner)

PROFESSIONAL SUMMARY

I build production-grade ML and data pipelines that transform complex biological and high-dimensional data into decisions — across genomics, precision medicine, and increasingly agentic AI systems. With a PhD in Computational Biology, 8+ years of hands-on experience, and 6 first-author Q1 publications, I operate in the messy middle between research-grade methods and production infrastructure: the place where raw sequencing data, statistical assumptions, cloud systems, stakeholder needs, and deadlines all collide.

At CSIRO I co-led GeneGuardian — an AI-assisted dynamic consent platform integrating FHIR, blockchain, and decentralised storage — and served as lead data scientist on over $3.2M in NIH, UNDP, and DFAT-funded research, collaborating with Mayo Clinic, MIT, Microsoft, NVIDIA, and clinical partners across Indonesia, Korea, and Australia. I also optimised Google DeepVariant for ancient DNA, deployed pharmacogenomics and federated genomic exchange pipelines internationally, and represented CSIRO in South Australia's federated-learning consortium.

On the quantitative side, I applied ML signal validation and anomaly detection on multi-asset time-series at AE Capital Hedge Fund — an experience that sharpened my intuition for high-dimensional noisy data and decisions that have to hold up under scrutiny. The pattern across genomics, pathogens, and markets is the same.

Outside of work, I am genuinely passionate about agentic AI. I have built OpenClaw + Hermes, my own open-source multi-agent orchestration framework, routing tasks across Claude API, OpenAI GPT, and local Ollama/Qwen models. This is not a side project — it is a live system running crypto trading research, business-development automation, personal AI assistance, and automated webpage generation. I am drawn to roles where I can bring all of this together: bioinformatics depth, data science rigour, cloud engineering, and an AI-native mindset.

CORE SKILLS SNAPSHOT

Programming: Python · R · Bash · SQL · C · MATLAB · JavaScript · PHP

Bioinformatics & Genomics: NGS/WGS · Nanopore · Pangenome Graphs · Variant Calling (GATK, DeepVariant) · SNP Arrays · Microbiome (16S) · Pharmacogenomics · Pathogen Detection · Population Genetics · GWAS · Ancient DNA (aDNA)

Cloud & Infrastructure: AWS (EC2 · S3 · Lambda) · Terraform · Docker · Singularity · HPC/SLURM · Git · CI/CD · Snakemake · Nextflow

Data Science & ML: Scikit-learn · PyTorch · Bayesian Inference (MCMC) · Random Forest · Anomaly Detection · Time-Series Modelling · Feature Engineering · EDA · Biostatistics

Agentic AI & LLMs: Claude API · OpenAI GPT · Ollama · Qwen · Multi-Agent Orchestration · Prompt Engineering · Workflow Automation · LLM-Assisted Coding

Data Governance & Security: FHIR · GA4GH · HIPAA · GDPR · Blockchain · Decentralised Storage (IPFS · FileCoin · StorJ) · Federated Learning · Consent Management

Leadership & Management: Team Lead (3–10) · Student Supervision (15+) · Agile/Scrum · Jira · Grant & Project Management ($3.2M+) · SOP Authoring · Stakeholder Communication (C-suite to farmers) · Scientific Outreach

PROFESSIONAL EXPERIENCE

Lead Bioinformatician

Genics Pty Ltd | January 2026 – Present | Brisbane, QLD, Australia

Leading bioinformatics strategy and end-to-end pipeline development for a commercial genomics company serving agricultural and industrial clients.

- **Nanopore 16S Microbiome Pipeline:** Architected the company's full Oxford Nanopore 16S microbiome analytical stack from scratch — from raw electrical signal through basecalling, mapping, variant calling, pathway analysis, statistical testing, SQL database integration, and automated client report generation. Validated and reliability-tested end to end.
- **Commercial SNP Array Pipeline:** Sole developer of a commercial-grade SNP Array pipeline for a new tree species: project scoping, budget and resource estimation, pricing, development, validation, delivery, and client-owned IP handover.
- **Breeding Program Infrastructure:** Designed a relational SQL database (DbSchema) for genomic selection analysis and optimised existing analytical scripts, reducing runtime and compute cost on large-scale client datasets.
- **Strategy & Standardisation:** Authored SOPs and centralised analytical workflows for reproducibility, quality assurance, and team onboarding.
- **Stakeholder Communication & Outreach:** Translated complex genomic outputs into actionable reports for non-technical stakeholders (farmers, producers, C-suite). Authoring peer-reviewed publications, industry magazine articles, and blog posts to grow scientific engagement.

Tech Stack: Python · R · Bash · SQL · DbSchema · Jira · Git · Agile · Nanopore · Microbiome · Pathogen Detection · SNP Array · Population Genetics · Machine Learning · LLM-Assisted Coding

Cloud Genomic Data Scientist

CSIRO (Commonwealth Scientific and Industrial Research Organisation) | October 2022 – October 2025 (3 years) | Remote, Australia

Led and delivered 4+ concurrent national and international research programs at CSIRO eHealth, spanning cloud genomics, AI-driven bioinformatics, and federated learning in collaboration with Mayo Clinic (USA), Université Paris-Saclay (France), MIT, Microsoft, NVIDIA, and clinical partners across Indonesia, Korea, and Australia. Winner of the CSIRO Health & Biosecurity Entrepreneurship Award (2024).

- **GeneGuardian — Data Governance Platform:** Co-led the full design and deployment of GeneGuardian — an AI-assisted dynamic consent platform for genomic data governance. Personally owned FHIR integration, blockchain architecture, and decentralised cloud storage (IPFS, FileCoin, StorJ) on AWS with Terraform. HIPAA, GDPR, and GA4GH compliant. Managed a 10+ cross-functional team (IT, medical informatics, bioinformatics) under SCRUM methodology. Deployed to live customers at geneguardian.online. First-author publication in GigaScience (Q1, DOI: 10.1093/gigascience/giae021). Presented and hosted a workshop at ABACBS 2024.
- **Pangenome & Precision Medicine — Mayo Clinic:** Totally led a multi-institution pangenome project with Mayo Clinic and Université Paris-Saclay, improving precision medicine for underrepresented Middle Eastern and Horn of Africa populations. Conducted all biostatistics analysis quantifying reference bias effects on PCA, D-statistics, and F3/F4-statistics. Built and ran scalable Python/Bash workflows on HPC. Supervised 2 postgraduate students throughout. First-author publication in BioMolecules (Q1, DOI: 10.3390/biom15040582). Poster at GA4GH Planetary 2024.
- **Cloud Pipelines & International Deployment:** Deployed PharmCAT (pharmacogenomics), sVEP (variant effect prediction), VariantSpark (Random Forest GWAS), and sBeacon (federated genomic data exchange) pipelines to partner AWS accounts in Indonesia, Korea, and Australian

clinical sites. Targeted pharmacogenomics, COVID-19 and TB surveillance, and population genomics — all with data sovereignty maintained in-country.

- **PrecisionGo — Pharmacogenomics:** Led PrecisionGo in collaboration with Sydney University and Westmead Hospital: a pharmacogenomics pipeline linking donor/recipient genetic variants to liver transplant outcomes and drug reactions, incorporating time-series ML regression. Served as main contact and project manager across all collaborating institutes.
- **Federated Learning Consortium:** Represented CSIRO in the South Australian Federated Learning Consortium alongside partners including MIT, Microsoft, NVIDIA, and the South Australian Government, implementing federated EHR sharing across SA hospitals. Federated learning architecture moved algorithms to the data, not the reverse.
- **Cancer Genomics Analysis:** Analysed WGS data from breast cancer rat models (treated vs untreated chemotherapy), performing variant calling, VCF analysis, and statistical testing (Wilcoxon rank-sum, t-tests) to identify significant genomic differences.
- **Leadership & Awards:** Mentored 15+ team members across bioinformatics, IT, and medical disciplines. Winner of the CSIRO Health and Biosecurity Entrepreneurship Award (2024) for expanding a grant into an international data exchange initiative for disease surveillance, diagnostics, and pharmacogenomics in Indonesia.
- **Scientific Outreach:** Speaker and workshop host at EY Digital Health 2024, GA4GH Planetary 2024, ABACBS 2024, APIdays 2023, and CNRS France. Published three CSIRO bioinformatics blog posts reaching a broad scientific and industry audience.

Tech Stack:
AWS · Terraform · Git · Python · Bash · R · Docker · Singularity · SQL · HPC · SLURM · Pharmacogenomics · Variant Calling · GWAS · Pangenome · Federated Learning · FHIR · Blockchain · IPFS · GA4GH · HIPAA · GDPR · Agile · Scrum · Jira · LLM-Assisted Coding

Quantitative Analyst

AE Capital Hedge Fund | March 2022 – August 2022 (6 months) | Melbourne, VIC, Australia

Applied ML and advanced statistical modelling to high-dimensional multi-asset financial time-series in a fast-paced, high-stakes hedge fund environment.

- **Signal Validation:** Validated ML model signals for predictive analytics, pattern recognition, and anomaly detection across equity, bond, and macro market datasets.
- **Predictive Modelling:** Built predictive models for bid/ask price movement using feature engineering, EDA, and regression on complex multi-dimensional financial datasets.
- **Statistical Modelling & Risk:** Applied statistical modelling and ML signal validation for systematic trading research and risk management across diverse asset classes.

Tech Stack: Python · MATLAB · Time-Series Analysis · ML Signal Validation · Anomaly Detection · Feature Engineering · EDA · Risk Management

Research Software Developer & Computational Modeller

CNRS — Centre National de la Recherche Scientifique | February 2017 – April 2018 (14 months) | Montpellier, France

- **MPEE Model Development:** Designed and implemented the MPEE (Minimum Posterior Expected Error) statistical model for ancestral sequence reconstruction in PhyML — a widely used phylogenetic inference software package — using Python and C.
- **Comprehensive Benchmarking:** Benchmarked Maximum Parsimony, Maximum Likelihood, and Bayesian Inference approaches across large simulated and empirical datasets; applied MCMC methods to quantify evolutionary uncertainty.
- **Publication:** First-author publication in Bioinformatics, Oxford University Press (Q1, 2019).

Tech Stack: Python · C · Bayesian Statistics · MCMC · Phylogenetics · PhyML · Bioinformatics Pipeline Development

Research Intern — Phylogenetics

Australian National University (ANU) | March 2016 – September 2016 (6 months) | Canberra, ACT, Australia

- **Phyloland Package Development:** Implemented MCMC-based evolutionary history simulation (dispersal rate, competition rate, migration event rate) in C for the Phyloland R package for phylogeographic analysis, incorporating Bayesian statistics and ancestral sequence reconstruction.

Tech Stack: R · C · Bayesian Statistics · MCMC · Phylogenetics

Web Development Intern

Comité Département Olympique et Sportif | May 2014 – August 2014 (3 months) | Poitiers, France

- **Full-Stack Development:** Independently maintained and developed new features for BasiCompta — a full-stack financial tracking web application for sports associations — including SQL database management and front-end development.

Tech Stack: JavaScript · jQuery · PHP · SQL · Web Development

Research Intern — Genomics

University of Auckland — Prof. Cristin Print's Research Group | April 2013 – July 2013 (3 months) | Auckland, New Zealand

- **Endometriosis Gene Database:** Developed a comprehensive genomics database for endometriosis research, annotating SNPs via the UCSC Genome Browser in collaboration with PhD researchers and the ENCODE consortium.

Tech Stack: SQL · UCSC Genome Browser · Variant Annotation · Scientific Communication

EDUCATION

PhD in Bioinformatics & Computational Biology

University of Adelaide | August 2018 – April 2022 | Adelaide, SA, Australia

Thesis: "Quantifying and Reducing Biases in Genomic Research Using Pangenome"

- Benchmarked alignment methods (BWA, Bowtie2, NovoAlign) for ancient DNA — published in Briefings in Bioinformatics (Q1, IF 9.5, DOI: 10.1093/bib/bbab076).
- Quantified reference bias impact on D-statistics, F3/F4-statistics, and PCA — second Q1 publication in Ecology and Evolution (DOI: 10.1002/ece3.8297).
- Built pangenome graph pipelines for variant calling and population analyses demonstrating superior accuracy over linear references on degraded short reads.
- Trained and optimised Google DeepVariant for ancient DNA — designed custom training data capturing aDNA-specific damage and fragmentation patterns.
- Performed population structure analysis, admixture modelling, and phylogenetic inference on East Asian ancient DNA cohorts.
- Co-authored FAIR/CARE community standards for reproducible paleogenomic analyses. Led engagement with international labs for consensus guidelines.
- Designed and taught 'Bioinformatics 101' curriculum (Python, Jupyter, command-line, computational thinking) to interdisciplinary lab teams.

Master of Engineering — Biotechnology, Bioinformatics & Data Analysis

Université de Poitiers | 2014 – 2016 | Poitiers, France

Specialisation in Computational Biology, Statistical Analysis, NGS, and Genomic Data Processing. Thesis project: developed the R package 'Phyloland' for phylogeographic analysis incorporating MCMC algorithms.

Bachelor / Licence — Biotechnology & Genomics

France | 2012 – 2013 |

AGENTIC AI — PERSONAL PROJECT & PASSION

OpenClaw + Hermes: Open-Source Multi-Agent Orchestration Framework

Personal Project | February 2026 – Present | Open Source

Agentic AI is not just a professional interest — it is a genuine passion project I pour time and energy into outside of work. OpenClaw + Hermes is an open-source multi-agent orchestration framework I built from scratch, combining a routing layer (OpenClaw) with a coordinating agent (Hermes) to autonomously design and execute multi-step workflows.

- **Architecture:** Multi-LLM architecture: Claude API handles advanced reasoning and code generation, OpenAI GPT models cover general-purpose task execution, and locally hosted Ollama/Qwen models run as lightweight heartbeat agents — each model specialised for its strengths within a single coordinated environment.
- **Live Use Case 1:** Crypto algorithmic trading research pipelines — signal generation, anomaly detection, and strategy backtesting on live market data.
- **Live Use Case 2:** Personal AI assistant for day-to-day life management and research workflows.
- **Live Use Case 3:** Automated webpage generation and business-development automation scripts deployed at company level.
- **Engineering Focus:** Token optimisation, version control integration (Claude Skills, OpenAI Codex), and local LLM infrastructure management.

Technologies: Agentic AI · Multi-Agent Systems · Claude API · OpenAI GPT · Ollama · Qwen · Local LLMs · Workflow Automation · Algorithmic Trading · LLM-Assisted Coding · Open Source

GRANTS & FUNDED RESEARCH

- NIH (USA): USD $850,000/year — Pangenome & Precision Medicine, in collaboration with Mayo Clinic
- UNDP: USD $1,250,000 — Genetic Analysis Support Platform, Indonesia
- AIHSP / DFAT: AUD $1,100,000 — SpreadStopper infectious disease surveillance project

Total funded research managed/co-led: $3.2M+ (NIH · UNDP · DFAT)

PUBLICATIONS — 6 FIRST-AUTHOR Q1 PEER-REVIEWED PAPERS

1. Oliva, A. et al. (2025). **A Pangenomic Approach to Improve Population Genetics Analysis and Reference Bias in Underrepresented Middle Eastern and Horn of Africa Populations.** *BioMolecules (Q1).* DOI: 10.3390/biom15040582
2. Oliva, A., Souilmi, Y. et al. (2025). **Lessons Learned: Recommendations for Reproducible Paleogenomic Data Analyses.** *Q1 journal (in press).*
3. Oliva, A. et al. (2024). **Future-Proofing Genomic Data and Consent Management: A Comprehensive Review of Technology Innovations.** *GigaScience (Q1).* DOI: 10.1093/gigascience/giae021
4. Oliva, A. et al. (2021). **Systematic Benchmark of Ancient DNA Read Mapping.** *Briefings in Bioinformatics (Q1, IF 9.5).* DOI: 10.1093/bib/bbab076
5. Oliva, A., Tobler, R., Llamas, B., Souilmi, Y. (2021). **Additional Evaluations Show that Specific BWA-aln Settings Still Outperform BWA-mem for Ancient DNA Data Alignment.** *Ecology & Evolution (Q1).* DOI: 10.1002/ece3.8297

6. Oliva, A. et al. (2019). **Accounting for Ambiguity in Ancestral Sequence Reconstruction.** *Bioinformatics, Oxford University Press (Q1).*

Co-authored & Under Review:

- Dikeocha, I.J., Oliva, A., Bateman, E. et al. (2026). Exploring the Molecular and Genomic Landscape of the Dark Agouti Rat Mammary Adenocarcinoma. J Mammary Gland Biol Neoplasia 31, 15. DOI: 10.1007/s10911-026-09606-x
- Denis Bauer, Adrien Oliva, Thomas Hardjono. Digital Data Exchange for Genomics in Australia. (Under Review)

CONFERENCE TALKS, WORKSHOPS & PRESENTATIONS

- **AI Beyond OMICS: Navigating Data Governance Challenges** — ABACBS, 2024
- **GeneGuardian: Secure Consent and Data Management (Talk + Workshop)** — ABACBS, 2024
- **AI Impact on Digital Health in Australia** — EY Conference, 2024
- **Precision Medicine for Coronary Heart Disease in Somali Population (Poster)** — GA4GH Planetary Session, Melbourne, 2024
- **The Power of APIs in Genomic Research** — APIdays, 2023
- **Cloud in Genomics and Personalised Medicine (Invited)** — CNRS, France, 2023
- **Systematic Benchmark of aDNA Mapping Bias** — ABACBS, 2020

Selected Blog Posts (CSIRO Bioinformatics):

- Future-Proofing Genomic Data and Consent Management — bioinformatics.csiro.au/blog/managing-genomic-data-in-the-future/
- GeneGuardian: Informing and Protecting Genomic Data Decisions — bioinformatics.csiro.au/blog/gene-guardian/
- Case Study: Indonesia's PathsBeacon — bioinformatics.csiro.au/blog/case-study-indonesias-pathsbeacon/

AWARDS & RECOGNITION

- CSIRO Health and Biosecurity Entrepreneurship Award (2024) — awarded for expanding a grant into an international data exchange initiative for disease surveillance, diagnostics, and pharmacogenomics in Indonesia.

PERSONAL

Date of Birth: 23 April 1993 **Nationality:** French & Australian (dual) **Languages:** French (native) · English (fluent) · Spanish (beginner)

Hobbies & Interests: Paragliding · Basketball (volunteer coach, Under-10s) · Agentic AI & algorithmic trading research · Coding prototypes · Board games · Video games

Open to: Australia · Europe (Switzerland, France) · Remote **Available:** Q2 2026