ALVARO FERRIZ | Cambridge, UK | alvaro.ferrizmejias@gmail.com | Linkedin

Bioinformatician with 7+ years of experience developing data analysis pipelines, automating high-throughput workflows, and leading data operations for large clinical and research studies. Proven expertise in Python, R, Bash, cloud platforms, workflow management, and multidisciplinary collaboration, with a strong commitment to robust data governance and delivery. I bring a detail-oriented approach, work well with diverse teams, stay focused on real-world impact, and proactively collaborate across groups and organisations to move projects forward.

PROFESSIONAL EXPERIENCE

Diagnostic Bioinformatician - AstraZeneca | 2022 - Present (3 years and 8 months)

Precision Medicine and Biosamples department, Oncology R&D

- Led Bioinformatics QC and Data Management for large-scale oncology clinical trials utilising NGS diagnostics, enabling precise cohort assignment for +20,000 patients through genomic biomarker analyses, supporting successful Companion Diagnostics submissions.
- Reduced bioinformatics data processing time in clinical trials by 50% through:
    - Establishing NGS FAIR data practices across studies by coordinating and negotiating with stakeholders.
    - Developing a cross-platform cloud data operations solution in AWS/Python in collaboration with IT.
    - Creating SOPs to streamline task delegation to DataOps teams.
- Delivered an automated end-to-end digital workflow infrastructure, for the department's lab to support the decentralised NGS diagnostic strategy, consisting of post-sequencing data annotation, transfer to external clouds, data retrieval and storage, and comprehensive evaluation of assays.
- Led biomarker performance evaluation of commercial NGS diagnostics to support clinical trial assay selection for both tissue and liquid biopsies, utilising personalised assays, targeted panels, WGS, and WES technologies.
- Mitigated patient misclassification risk through:
    - Evaluating buffy-coat sequencing in liquid biopsy assays to detect confounding CHIP variation.
    - Establishing pre-analytical sample swaps detection workflows.

Bioinformatics Research Engineer - Barcelona Supercomputing Center | 2018 - 2022 (3 years and 3 months)

Computational Genomics Group, Life Sciences department

- Researcher for international cancer genome projects (ICGC ARGO, EUCANCAN), advancing the standardization and benchmarking of variant calling methods across institutions.
- Developed, deployed, and validated automated pipelines for cross-institutional data processing in HPC environments, utilizing Python, R, and Nextflow to ensure scalability and reproducibility.
- Implemented and evaluated machine learning models such as Random Forests and Gradient Boosting for somatic variant analysis, supporting innovative approaches in oncology genomics.
- Ensured robust data governance and traceability maintaining full compliance with data privacy regulations.

Bioinformatics Intern - Sequentia Biotech S.L. | 2018 (6 months)

- Developed and benchmarked automated variant calling pipelines for plant genomics.

Bioinformatics Intern - Miguel Hernandez University of Elche | 2015 - 2017 (9 months)

- Performed assembly and annotation of plant mitochondrial genomes and assisted in molecular genetics laboratory procedures (DNA extraction, PCR, Sanger sequencing).

EDUCATION

- MSc in Bioinformatics - Autonomous University of Barcelona, Spain | 2017-2018
- Exchange student in Biomedical Sciences graduate program - University of New Mexico, USA | 2017
- BSc in Biotechnology - Miguel Hernandez University of Elche, Spain | 2013 - 2017

SKILLS

- **Bioinformatics Infrastructure:** Python, R, Bash, Nextflow, Docker, Singularity, HPC, AWS, Git, GitHub, Jira.
- **Data Operations & Pipeline Automation:** Developed, optimized, and supported cloud-based and high-performance computing workflows for omics data; enabled secure storage, retrieval, and transfer for clinical/research projects.
- **Algorithm & Pipeline Development:** Variant calling, biomarker discovery, data annotation, QC workflows.
- **Data Handling:** Managed large-scale biological and clinical datasets, implemented data governance, and followed FAIR principles for traceable, high-quality data.
- **Collaboration:** Extensive multidisciplinary teamwork with clinical, IT, and research groups.
- **Documentation & SOP Writing:** Produced clear SOPs and technical documentation for workflows and infrastructure.

CERTIFICATIONS

- AWS Cloud Practitioner | AWS | Ongoing
- Finance for Managers | IESE Coursera | 2020 | Certificate
- Principles of Financial Accounting | IESE Coursera | 2020 | Certificate
- Python Programming | Datacamp | 2020 | Certificate

LANGUAGES

- Spanish - Native
- Catalan - Full Working Proficiency
- English - Full Working Proficiency

PUBLICATIONS AND POSTERS

- Martín R, Gaitán N, Jarlier F, Feuerbach L, de Soyres H, Arbonés M, Gutman T, Puiggròs M, **Ferriz A**, et al. ONCOLINER: A new solution for monitoring, improving, and harmonizing somatic variant calling across genomic oncology centers. Cell Genom. 2024 Sep 11;4(9):100639. doi: 10.1016/j.xgen.2024.100639.
- Analytical comparison of tissue-based next-generation sequencing assays for the detection of PIK3CA, AKT1, and PTEN tumor alterations in breast cancer. San Antonio Breast Cancer Symposium® – December 10–13, 2024 Poster number: P1-05-27 (in Acknowledgments).
- Tools report on best practices for genome analysis within the EUCANCAN project for the European Commission. Dec 2019.