Alejandro Gonzales-Irribarren

https://alejandrogzi.github.io   Lima, Peru   +51 934-090-994

Education	**National University of San Marcos**, Lima, Peru	March 2018-Present
	Doctor of Veterinary Medicine (DVM) with Highest Honours	Graduation: December 2023

Specialization: Bioinformatics and Medicine Informatics
GPA: 4.0/4.0, AM: 16.0/20.0, Hons: *Summa Cum Laude* (1/82)

Publications

Castillo G, Barrios-Arpi L, Ramos-Gonzalez M, Vidal P, **Gonzales-Irribarren A**, Ramos-Cevallos N, Rodríguez JL. Neurotoxicity associated with oxidative stress and inflammasome gene expression induced by allethrin in SH-SY5Y cells. Toxicology and Industrial Health. 2022 Dec;38(12):777-88.

Gonzales-Irribarren A. gtfsort: a tool to efficiently sort GTF files. bioRxiv. 2023 Oct 25:2023-10.

Software

gxf2bed: Fastest GTF/GFF-to-BED converter chilling around. **Gonzales-Irribarren A**. 2024. [Rust]. Available at: https://github.com/alejandrogzi/gxf2bed

gtfsort: A chr/pos/feature GTF sorter that uses a lexicographically-based index ordering algorithm. **Gonzales-Irribarren A**. 2023. [Rust]. Available at: https://github.com/alejandrogzi/gtfsort

bed2gtf: A high-performance BED-to-GTF converter written in Rust. **Gonzales-Irribarren A**. 2023. [Rust]. Available at: https://github.com/alejandrogzi/bed2gtf

bed2gff: BED-to-GFF3 converter that runs in parallel. **Gonzales-Irribarren A**. 2023. [Rust]. Available at: https://github.com/alejandrogzi/bed2gff

noel: GTF/GFF per gene non-overlapping exon length calculator. **Gonzales-Irribarren A**. 2023. [Rust]. Available at: https://github.com/alejandrogzi/noel

to-trans: A high-performance exon/CDS spliced transcriptome builder from fasta + GTF/GFF. **Gonzales-Irribarren A**. 2023. [Rust]. Available at: https://github.com/alejandrogzi/to-trans

stoseq: A portable storage-optimized RNA-seq processing pipeline. **Gonzales-Irribarren A**. 2023. [Nextflow, Python]. Available at: https://github.com/alejandrogzi/stoseq

postoga: The post-TOGA processing pipeline. **Gonzales-Irribarren A**, Hiller M, Kirilenko B. 2023. [Python]. Available at: https://github.com/alejandrogzi/postoga

Research Experience

UNAM International Laboratory for Human Genome Research
with Ph.D. Daniela Robles-Espinoza   May 2023-January 2024
Worked on human/dog transcriptional parallelisms in prostate cancer as the lead researcher.

- Conceptualized and initiated the project, proposing a novel methodology to compare transcriptomic profiles in malignancies across two distinct species through genome alignments.
- Designed a HPC-compatible pipeline using Python to post-process gene models inferred from genome alignments between the dog and human assemblies.
- Created a set of high-performance tools using Rust to address limitations encountered during the post-processing phase, including parallel converters, transcriptome builders and a sorter tool.
- Significantly enhanced the current understanding of the *Canis lupus familiaris* genome, correcting previously unreported inaccuracies in gene models, such as the omission of oncogenes (e.g., RSF1) and the masking of cancer-related genes by duplicated coordinates (e.g., BCL7B-TBL2).

UPCH Bioinformatics and Molecular Biology Lab
with Ph.D. Mirko Zimic   October 2023-December 2023
Worked on *The role of the thermodynamic pressure in Lensky's long-term evolution experiment* and initiated the identification of potential platelet-related molecular markers for accurate tuberculosis diagnosis.

- Led the ideation and implementation of the code framework for the thermodynamic pressure project. Contributed significantly to the statistical model development, enabling the re-analysis of over 1,200 bacterial genomes using Python, C++ and Nextflow, available at https://github.com/alejandrogzi/thermut.

- Collaborated with two Biology undergraduate students to conduct in-depth analysis on mRNA sequencing data obtained from TB-positive and TB-negative patients in Peru using a storage-optimized RNA-seq pipeline I developed, available at https://github.com/alejandrogzi/stoseq.
- Implemented additional parallel programming pipelines in Python and Nextflow to automate protein dynamics in three different *M. tuberculosis* kinases related to pyrazinamide resistance.

UNMSM Biology and Molecular Genetics Lab

with Dr. Lenin Maturrano   June 2019-March 2023

Worked on *The genomic epidemiology and spatiotemporal spread of SARS-CoV-2 in Peru* as the second author and performed wet-lab procedures.

- Partnered with MSc. Dennis Carhuaricra from the University of Sao Paulo (USP) to establish the research problem and goals of the project, including the computational modeling requirements, data collection strategies, and the integration of advanced algorithms for analysis and interpretation.
- Analyzed the history of the SARS-CoV-2 pandemic in Peru from a genomic epidemiology perspective, highlighting significant temporal changes and generating insightful visualizations to illustrate the findings.
- Standarized several qPCR protocols for gram-negative bacterial isolates, performed short-read Illumina sequencing, conducted Kirby-Bauer tests and performed bacterial characterization assays to ensure rigorous experimental procedures.

Other Projects	**Cross-species analysis of human and dog prostate cancer transcriptome**

with Ph.D. Daniela Robles-Espinoza and Dr. Lenin Maturrano   June 2023-Present

- Conceptualized and initiated the project, proposing a novel methodology to compare transcriptomic profiles in malignancies across two distinct species through genome alignments.
- Identified and documented previously unrecognized errors within the dog gene models, enhancing assembly completeness and refining cancer-related analyses.
- Applied machine learning to re-classify orthologs using TOGA and several analytic techniques to interpret the data, fostering a deeper understanding of the cross-species transcriptomic landscape in prostate cancer.
- Established a complete analysis pipeline to replicate all procedures systematically, using Python and Nextflow, that includes pre and post-processing phases.

Unraveling the complexity of gene fusions in canine cancers

with Ph.D. Daniela Robles-Espinoza   October 2023-Present

- Defined the research problem from scratch, addressing the lack of information concerning gene fusions and their impact within the spectrum of canine malignancies.
- Examined over 1,300 whole transcriptomes spanning 12 diverse malignancy types and 16 distinct normal tissues. This extensive analysis aimed to construct an updated catalogue of normal gene fusions and establish the first public database of gene fusions within the *Canis lupus familiaris* species.
- Identified cancer-specific gene-fusions prevalent across various canine malignancies, discerning distinct gene-fusion events unique to specific cancer types, cross-referencing extensive datasets of cancer transcriptomes against normal tissue samples.

Neurotoxicity markers induced by allethrin in SH-SY5Y cells

with Dr. Jose-Luis Rodriguez   June 2022-October 2022

- Analyzed the mRNA levels from neuroblastoma cells (SH-SY5Y) inflammatory markers, performed the statistical analysis, wrote part of the paper and made all the figures.

Other Experience	UNMSM MVPPS007 Molecular Biology, *Teaching Assistant*	August 2023-November 2023
	UNMSM Faculty Board Representative *Student Representative*	September 2022-November 2023
	UNMSM MVOS0003 Animal Genetics, *Teaching Assistant*	March 2023-May 2023

Awards	2023 - *Summa Cum Laude* honors for holding the 1st place during six consecutive academic years. National University of San Marcos, School of Veterinary Medicine.
	2021 - VI International Scientific Conference on Biosciences. Hamutay - Young Peruvian Scientists

Network for Bioscience Research. Best Undergraduate Research Presentation - 1st Place.
2021 - XXVI Pan American Congress of Veterinary Sciences. Featured Congress Poster.

Community Involvement	Lima MatLab Mathematics Program, *Math Tutor Volunteer*	June 2018-November 2023
	UNMSM Circle of Veterinary Studies in Epidemiology, *President*	March 2022-September 2023
	UNMSM DVM Mentoring Program, *Mentor*	March 2021-June 2021
	UNMSM EEGG Public Health for Schools, *Speaker*	March 2018-June 2018
	UNMSM EEGG Health Projects Fair, *Speaker* on *About language: the animals?*	May 2018

References

Ph.D. Daniela Robles-Espinoza Associate Professor of Genomics at LIIGH-UNAM,
Email: drobles@liigh.unam.mx,
Tel: 55 56 23 43 31 (208).

Ph.D. Mirko Zimic Professor of Bionformatics and Molecular Biology at UPCH,
Email: mirko.zimic@upch.pe,
Tel: (51) 01-319-0000 (2604)

Dr. Lenin Maturrano Professor of Genetics and Molecular Biology at UNMSM,
Email: amaturranoh@unmsm.edu.pe,
Tel: (51) 01-619-7000 (5006).