Ana Hernandez-Lopez

ana.ang.herlop@gmail.com https://orcid.org/0000-0001-6009-7092
+41 78 838 3673 ana.ang.herlop hlana 1015 Lausanne, Switzerland

CAREER OBJECTIVE

To contribute my expertise in high-throughput sequence data processing, and actively collaborate with interdisciplinary teams, to the discovery, identification, investigation and analysis of pressing problems in omics studies to devise judicious solutions by leveraging computational methods in predictive analytics, toward causal inference.

EDUCATION

Switzerland

PhD candidate in Computer Science 2014 – Oct. 2019
Swiss Federal Institute of Technology, Lausanne (EPFL)

Research on new approaches to genomic information representation and processing

Dissertation: On the relevance of quality score metadata in genomic sequence data for omics applications

Mexico

Master in Computer Science 2011 – 2013
Center for Research and Advances Studies of the National Polytechnic Institute of Mexico (CINVESTAV-LTI)

Research on computer vision and embedded systems

Thesis title: Design of a versatile FPGA-based architecture for local feature detection under time constraints

Bachelor in Electrical and Electronic Engineering 2002 – 2009
National Autonomous University of Mexico, Faculty of Engineering (UNAM-FI), Institute of Engineering (IIUNAM)

Major in Electronics

Thesis title: Data recording on SD card memories implementing FAT16 and FAT32 file systems

PROFESSIONAL EXPERIENCE

Switzerland

SCI-STI-MM group | **EPFL** 2015 – present

Doctoral assistant

- Systematically applied the workflow of data science to process high-throughput sequence data (whole genome and RNA-seq) for variant calling, differential gene expression, and sequence alignment
- Studied the impact of lossy compression of genome sequence metadata in the context of variant calling and differential gene expression pipelines, and sequence alignment

GenomSys/GenomSoft | *Systems for Efficient Genomic Information Processing* 2016 – present

Research assistant

- Surveyed downstream applications and tools for genome sequence analysis, and proposed pipelines to explore the impact of lossy compression on genomic sequences
- Examined and assessed the effect of lossy compression of sequence metadata and published findings

MPEG-G | *MPEG Genomic Compression Group* 2015 – 2018

Swiss delegate to the ISO/IEC JTC 1, Genomic Data Compression Working Group

- Investigated genomic information storage and compression to identify problems in the representation and manipulation of sequence data

- Contributed to the documentation of findings and dissemination activities of the group

Swiss Institute of Bioinformatics (SIB/Vital-IT) | *PoSeNoGap Project* 2015 – 2016

Research assistant and IT technical support

- Inspected, analyzed and computed preprocessing of genomic sequence data files
- Researched and evaluated genomic data compressors to define a benchmark strategy that provided quantitative evidence of the storage footprint of genomic sequence metadata

Mexico

CINVESTAV-LTI | *Department of Embedded Systems and Reconfigurable Computing* 2014

Research assistant

- Published results of my master thesis in a scientific journal (IJARS Women in Robotics winning paper, 2015)

Nextel | *Network Operation Center* 2010

Monitoring engineer

Siemens | *Business Learning Program* 2009

Trainee in the Electrical Substations division, Energy sector

IIUNAM | *Department of Instrumentation* 2006 – 2009

Research and project assistant

RELEVANT PUBLICATIONS

Hernandez-Lopez, A., Alberti, C. and Mattavelli, M. "Toward a Dynamic Threshold for Quality-Score Distortion in Reference-Based Alignment". **Journal of Computational Biology** [Accepted]

Hernandez-Lopez, A. and Mattavelli, M. "Relevance of QVS Information for Analysis Applications of Genomic Sequencing Data". **IEEE Data Science Workshop**, Minneapolis, MN, USA, June 2–5, 2019.

Hernandez-Lopez, A., Voges, J., Alberti, C., Mattavelli, M. and Ostermann, J. "Lossy Compression of Quality Scores in Differential Gene Expression: A first Assessment and Impact Analysis". **IEEE 2018 Data Compression Conference**, Snowbird, Utah, USA, March 27–29, 2018.

TECHNICAL SKILLS

Shell scripting: Bash, AWK, Perl

Processing and Visualization: R, tidyverse (tidyr, dplyr, ggplot2, etc), R shiny, Python, SciPy, NumPy, Scikit-learn, Matlab, C, genome browsers and viewers

Others: SAMtools, bcftools, GATK, Illumina BaseSpace, Galaxy, TeX/LaTeX, Git

LANGUAGES

Spanish Native language

English Fluent

French Intermediate level spoken and written (B2)

EXTRACURRICULAR ACTIVITIES

Avid reader of general psychology and existential philosophy. Also reader of The Economist, Quanta magazine and Le Monde

Coffee enthusiast: roasting specialty coffee beans, pour-over brewing and lever espresso extraction

Plant-based cooking: exploring flavor profiles with minimal ingredients

Long-distance running: 10 km and half marathon