Lois Randolph, M.S.

randolphl@uthscsa.edu | www.linkedin.com/in/lois-randolph | www.github.com/Low-is | https://low-is.github.io/Lois_portfolio/

SUMMARY

Bioinformatician with expertise in statistical analysis, experimental design, model building, and multi-omics data analysis. Proficient in data visualization using R, Python, and Power BI. Experienced using Linux for file management, version control, and container orchestration. Skilled in managing large-scale databases and implementing bioinformatics workflows to perform tertiary analyses for clinically relevant variant interpretations. Active in web development, including a portfolio site, fitness/lifestyle platforms, and academic resource hubs.

SKILLS

Programming Languages: R, Python, JavaScript, SQL, SAS
Machine Learning: Scikit-Learn, caret, randomForest, XGBoost, nnet
Deep Learning: TensorFlow, Keras, NNs & CNNs
Analysis & Visualization Tools: R, Python, Power BI
Database Management Systems: MySQL, RSQLite
Version Control & Command Line: Git**,** Unix/Linux, WSL2 (Ubuntu on Windows), Docker Desktop, Bash scripting
Web Development: Flask, HTML, CSS
Workflow Management: Nextflow
Development Environments: Jupyter Notebook, RStudio, VS Code, PyCharm

WORK EXPERIENCE

UT Health Science Center **Dec. 2023 - Present**
Bioinformatician *San Antonio, TX*

- Created unsupervised and supervised learning models to identify biomarkers for neonatal care. Applied standardized and unstandardized meta-analyses for dimensionality reduction and feature selection, accounting for inter-study variability and covariates. Improved model performance by ~15% achieving AUCs of 0.80-0.90.
- Used linear mixed-effects models to quantify temporal gene expression dynamics associated with BPD progression during the first month of neonatal life. Integrated Gene Ontology to track evolving biological processes over time.
- Developed dockerized bioinformatics workflows focused on providing robust analyses and predictions.
- Utilized Git for code reproducibility and the management of complex bioinformatics pipelines, enhancing workflow generalizability and improving overall efficiency.
- Built and trained convolutional neural networks (CNNs) in keras to classify brain MRI images (tumor vs. normal); trained on 3,200+ images and validated on 1,400+ images. Applied data augmentation and optimized CNN architecture, improving validation accuracy by ~ 10%.

UT Health Science Center **Aug. 2021 – Dec. 2023**
Graduate Research Assistant *San Antonio, TX*

- Utilize RNA-Seq bioinformatic tools from Bioconda to perform end-to-end processing of raw sequencing data, including quality control, alignment to the UCSC mm9 mouse reference genome using TopHat2, and gene quantification with HTSeq.
- Conducted differential expression analysis on 12 triple-negative breast cancer (TNBC) xenograft and syngeneic mice models, uncovering suppression of oncogenic signaling pathways in high-fat diet (HFD) groups treated with the small-molecule compound EC359.
- Identified 179 genes uniquely associated with the HFD/obesity phenotype, providing insights into the molecular underpinnings of obesity-driven TNBC progression.

- Applied a Python-adapted version of the Hypermut2 algorithm to identify APOBEC-mediated mutational signatures across ~ 1 million base pairs from HIV-infected human and non-human primate samples.

Baylor College of Medicine **June 2021 – Aug. 2021**
Intern *Houston, TX*

- Linked and organized data from over 10 entity types, including Allele Molecular Consequence, Variant, and Population Allele Frequency enhancing the curation process for ClinGen.
- Maintained and updated the LDH, ensuring availability of 1+ million data points in a highly accessible format for researchers and curators.
- Translated 1000+ lines of Ruby code into JavaScript, enabling smooth data parsing and interaction with APIs across multiple domains.
- Developed and optimized data pipelines to transform and aggregate gene and variant information, improving the efficiency of data retrieval and supporting 500+ ClinGen curation projects.

UT Health Science Center **June 2020 – June 2021**
PREP Scholar *San Antonio, TX*

- Designed custom code from ImageJ-processed microscopy images to identify regions of interest and detect colocalized proteins across 100+ images, analyzing two color channels to pinpoint and quantify protein interactions.
- Applied image processing tools in Python to enhance microscopy images, improving quality and contrast, identifying samples, annotating labels, and tailoring visuals for professional reporting.
- Employed an automated co-localization analysis on microscopy images, applying Mander's colocalization coefficient to identify and quantify 20+ co-localized regions of proteins CTD, SRSF2, and EWS across various experimental conditions, revealing significant interactions between these proteins.

PROJECTS

Multiclass 3D CNN for Lung Cancer Subtype Classification Using CT Scans

- Developed a multiclass classification model using a 3D convolutional neural network (3D CNN) in TensorFlow/Keras to predict histopathological subtypes of lung cancer from volumetric CT scan data.
- Parsed and preprocessed raw DICOM files across nested directory structures representing cancer subtypes and individual patients, converting axial slices into 3D image volumes.
- Constructed a deep 3D CNN with multiple convolutional, max-pooling, batch normalization, and dense layers, culminating in a softmax output for multiclass prediction.
- Tracked model performance with training and validation metrics, and visualized training history (accuracy/loss) to assess model convergence and generalization.

Generalized Statistical Modeling & Multi-Group Analysis Pipeline

- Designed a meta-analysis workflow to compute effect sizes and apply Cochran's Q test for heterogeneity, enabling robust feature selection by filtering ~20% of features.
- Initiated effect-size based feature selection method that reduces dimensionality of features by identifying features with consistent directional effects across studies and strong pooled estimates ($|ES| > 0.5$, 1, etc.), improving model interpretability and robustness.
- Containerized the pipeline using Docker to ensure reproducibility, promote quicker analyses, and have a general workflow for use with various datasets.

Multi-Omics Analysis Pipeline Development

- Developed a scalable gene annotation and expression analysis pipeline supporting over 50 multi-omics datasets across multiple species using biomaRt, RSQLite, and Shiny.

- Built a custom SQLite-backed gene annotation database to enable automated, species-wide ID conversion across 250+ organisms using filtered Ensembl attributes, significantly improving performance, reducing redundant API calls, and enhancing reproducibility of analyses.

Dose-Dependent Disruption of The Gut Microbiome and Intestinal Pathophysiology Following Acute Radiation Exposure

- Processed raw 16S rRNA FASTQ files by performing quality assessment, filtering, trimming, and denoising using DADA2 R package.
- Constructed sequence tables and assigned taxonomic classifications to ASVs to enable downstream ecological and statistical analyses.
- Developed a reproducible and automated R-based pipeline to streamline microbiome data processing, diversity analyses, and visualization across multiple timepoints and treatment groups.
- Visualized microbial community structure via Principal Coordinates Analysis (PCoA), and quantified group-level differences using PERMANOVA and pairwise PERMANOVA, with adjustments for repeated measures and multiple comparisons.
- Modeled longitudinal trends and individual variability using linear mixed-effects models (LMMs) implemented in lme4, lmerTest, and nlme.

PUBLICATIONS /ABSTRACTS

- Dose-Dependent Disruption of The Gut Microbiome and Intestinal Pathophysiology Following Acute Radiation Exposure. (2025). AAST and Clinical Congress of Acute Care Surgery.
- Insights into bronchopulmonary dysplasia: A meta-analysis of transcriptomic data. (2025). Pediatric Academic Societies (PAS) Meeting.
- McOmber BG, **Randolph L**, et al. (2025). Predicting Future Respiratory Hospitalizations in Extremely Premature Neonates Using Transcriptomic Data and Machine Learning. [*Manuscript in progress*].
- Romo B, Fuentes Z, **Randolph L**, et al. (2025). Targeting the LIF/LIFR axis reduces the progression of inflammatory breast cancer by promoting ferroptosis. *Cancers*. https://doi.org/10.3390/cancers17050790.
- **Randolph L**, Joshi J, et al. (2024). Significance of LIF/LIFR Signaling in the Progression of Obesity-Driven Triple-Negative Breast Cancer. *Cancers*, 16, 3630. https://doi.org/10.3390/cancers16213630 .
- Spencer N, Sanchez Rodriguez AL, et al. (2023). The LIFR Inhibitor EC359 Effectively Targets Type II Endometrial Cancer by Blocking LIF/LIFR Oncogenic Signaling. *Int. J. Mol. Sci*, 24(24), 17426. https://doi.org/10.3390/ijms242417426 .
- LeBaron RG, Perez L, **Randolph L** & Phelix C. (2020). A TGFβ-BIGH3-Apoptosis Axis Comprising Peptidase and Integrin α3β1 Promotes Renal Cell Death. *The FASEB Journal*, 34(S1), 03078. https://doi.org/10.1096/fasebj.2020.34.s1.03078.

EDUCATION

M.S. in Cancer Biology — **May 2023**
University of Texas Health Science Center San Antonio — *San Antonio, TX*

B.S. in Biology | Minor in Mathematics — **May 2020**
University of Texas at San Antonio — *San Antonio, TX*