HARRISON WANG
PH.D. STUDENT

✉ harrison.c.wang@gmail.com
🌐 https://harrisonized.github.io
in https://www.linkedin.com/in/harrisonized/
⌂ https://www.github.com/harrisonized

PROGRAMMING: Python, R, SQL, Matlab, Datalog
DATABASES & WEB FRAMEWORKS: Postgresql, Datomic, MongoDB, Flask, RShiny
CLOUD COMPUTING & DASHBOARDING: AWS, EC2, S3, Sagemaker, Redshift, Athena, Quicksight, Tableau, Power BI
WET LAB: qPCR, Flow Cytometry, Cloning, Mouse Studies, RNA FISH, ELISA, Tissue Culture, Protein Purification, Assay Development

EXPERIENCE

UNIVERSITY OF PENNSYLVANIA
Ph.D. Student
Aug. 2022 to Current
Bethesda, MD

Laboratory of Dr. Roxane Tussiwand
Ph.D. Student in the NIH-Penn Immunology Partnership Program
Rotation Work:

- Fourth rotation in Roxane Tussiwand's lab at the NIH, studying the role of DNase enzymes in **autoimmunity**. Optimized a **PCR-based assay** to quantify the relative RNA expression of in 14 mouse tissues. Designed a 12-color flow panel for identifying myeloid cells from mouse spleen, bone marrow, and lymph nodes.
  Wet lab: PCR, flow cytometry, mouse work, Dry lab: R, Flowjo
- Third rotation in Mariana Kaplan's lab at the NIH, studying sex differences in neutrophils. Optimized an assay for using qPCR to examine transcriptional differences between male and female **human neutrophils** after **in vitro treatment** with sex hormones.
  Wet lab: human neutrophil isolation, RNA extraction, qPCR, Dry lab: Python
- Second rotation in Montserrat Anguera's lab at Penn, studying how X-inactivation contributes to female bias in autoimmune diseases. Designed a **Python image analysis** pipeline using CellProfiler and a custom algorithm based on PCA to quantify nuclear localization of Xist clouds. Used R to correct and productionize an **RNA-seq data analysis** pipeline to identify and evaluate escape genes.
  Wet lab: RNA FISH and H3K27me3 immunofluorescence. Dry lab: R, Python
- First rotation in Michela Locci's lab at Penn, exploring the role Activin A in rheumatoid arthritis. Treated 25 mice with an anti-Activin A antibody to study the disease progression of **collagen-induced arthritis**. Designed a custom **Python** plot to show average disease severity scores for each toe and paw on the mouse.
  Wet lab: tissue culture, mouse IP and tail dermal injections, anesthesia, euthanasia, flow cytometry. Dry lab: Flowjo, Python

THERMO FISHER SCIENTIFIC
R&D Data Scientist
Feb. 2021 to Aug. 2022
South San Francisco, CA

- Built a **scalable, end-to-end ETL process** to **monitor SARS-CoV-2 RT-PCR test data** from labs and universities using **Amazon AWS** and **Power BI**. Listed as co-author on a **symposium abstract** based on this work
- Led the development of the **data pipeline**. Heavily **refactored** and **optimized** the existing data pipeline for **scalability** through clever use of **parallelization** and **batch processing**, reducing the runtime by 40x. Added new KPIs and support for backward compatibility
- **Analyzed** performance of sample retests using **Python** and **SQL**, enabling stakeholders to understand the limit-of-detection of our SARS-CoV-2 RT-PCR tests
- Designed an **Amazon Quicksight** dashboard to monitor data ingestion and filtering, eliminating up-front time for scoping new projects
- Created new **Plotly interactive visualizations** for 96- and 384-well plates, enabling the Data Science team to rapidly **detect anomalies**
- Rescued a legacy **Matlab** codebase from disrepair through **heavy refactoring**

INVITAE
Dry Lab Operations Data Scientist
Jan. 2020 to Sept. 2020
San Francisco, CA

- Discovered the **root cause** for a high-visibility. time-sensitive issue on a new assay that resulted in batch failures impacting 2000+ patients
- Analyzed **trends** using **Python** and **SQL** to monitor **next-generation sequencing (NGS)** production-line data for anomalies
- Efficiently diagnosed **customer issues** by writing a **Python script** to automate generation of standardized **visualizations** and **statistics**
- Wrote maintainable **Datalog** and **SQL** queries for **Flask** apps used to migrate data from **Datomic** to **Amazon Redshift** for consumption by **Tableau**
- Built an 8-page **Tableau dashboard** to enable our operations team to carefully monitor a new product launch

SANOFI, FORMERLY TRUE NORTH THERAPEUTICS
Research Associate 2
Jan. 2017 to Mar. 2019
South San Francisco, CA

- Led **protein-engineering project** of lead antibody drug (TNT009). Handled all **cloning and expression** of antibody variants and most **protein purification** for the lab. **Collected binding and efficacy (KD and IC50) data** using **biolayer interferometry** and **ELISAs**. Performed **regression analysis**, revealing a log-linear relationship between KD and IC50, enabling our team to select variants to use in downstream experiments.
- Led a **research project** to make a protein complex of TNT009 and its target (complement C1s) for crystallography study
- Developed **ELISAs**, **hemolysis assays**, and **flow cytometry** assays to measure classical and alternative complement activation
- Wrote **Python** script to automate design of short DNA oligos, reducing a 2 hours/week task to a 3 minute script.

Wet lab: tissue culture, protein expression, protein purification, column chromatography, ELISA, biolayer-interferometry kinetics assays

UC SAN DIEGO, LABORATORY OF DR. GENE YEO
Staff Research Associate 1
Aug. 2015 to Nov. 2016
La Jolla, CA

- Used **CRISPR/Cas9** to create, validate, and maintain 36 stable HEK293XT cell lines with end-tagged RNA-Binding Proteins (RBPs) for CLIP-seq as an extension of the Encyclopedia of DNA Elements (ENCODE) project
- Constructed and tested catalytically dead Cas9 (dCas9) fused to split proteins (GFP, Venus, thymidine kinase) as tools to detect RNA abundance in cultured **HEK293XT** and **U2OS cells**. Performed **flow cytometry** and **Western blot** to quantify results.
- Designed and **Gibson assembled** plasmids encoding human codon-optimized Cas-protein orthologs
- **Co-authored a Cell paper** on using CRISPR/Cas9 to track RNA in live cells

Wet lab: cloning, next-generation sequencing, high-throughput sequencing, CRISPR-Cas9, flow cytometry

HARRISON WANG
PH.D. STUDENT

EXPERIENCE (CONTINUED)

GENE YEO LAB, UCSD
Undergraduate Research Assistant

Aug. 2015 to Nov. 2016
La Jolla, CA

- Optimized a **fluorescence-coupled activity assay** to measure D- and L-serine levels with up to 5 nM sensitivity on in-vitro synthesized DAO and in-vivo DAO-overexpression HEK293 cell lines
- Used **Flp-In system** to create, validate, and maintain stable D-amino Acid Oxidase (DAO) overexpression HEK293 cell lines
- Took high-resolution **confocal microscopy** images of stained iPSC-derived motor neurons
- Validated via **qPCR** splicing changes of 36 genes related to familial amyotrophic lateral sclerosis (fALS) found via RNA-seq
- **Co-authored a Neuron paper** that included my experiments on investigating the mechanism for how a mutation in an RNA-binding protein (hnRNPA2B1) can cause ALS

Wet lab: cloning, tissue-culture, qPCR, Western blot, confocal microscopy, fluorescent cell imaging, assay development

TEACHING EXPERIENCE

OFFICE OF ACADEMIC SUPPORT AND INSTRUCTIONAL SERVICES, UCSD
Physics Tutor

Apr. 2013 to Mar. 2015

- Independently led 2-hour workshops of 10 to 20 students twice a week
- Created 18 original 2-hour physics lectures with 4 practice problems each
- Fostered a collaborative learning environment free of discrimination and judgment

PUBLICATIONS

ELIMINATION OF TOXIC MICROSATELLITE REPEAT EXPANSION RNA BY RNA-TARGETING CAS9

Ranjan Batra, David A. Nelles, Elaine Pirie, Steven M. Blue, Ryan J. Marina, **Harrison Wang**, Isaac A. Chaim, James D. Thomas, Nigel Zhang, Vu Nguyen, Stefan Aigner, Sebastian Markmiller, Guangbin Xia, Kevin D. Corbett, Maurice S. Swanson, and Gene W. Yeo. **Cell.** 2017 Aug 24;170(5):899-912.e10. doi: 10.1016/j.cell.2017.07.010. Epub 2017 Aug 10. PMID: 28803727; PMCID: PMC5873302.

PROTEIN-RNA NETWORKS REGULATED BY NORMAL AND ALS-ASSOCIATED MUTANT HNRNPA2B1 IN THE NERVOUS SYSTEM

Fernando J. Martinez, Gabriel A. Pratt, Eric Van Nostrand, Ron Batra, Stephanie C., Huelga, Katannya Kapeli, Peter Freese, Seung J. Chun, Karen Ling, Chelsea Gelboin-Burkhart, Layla Fijany, **Harrison C Wang**, Julia K. Nussbacher, Hong-joo Kim, Rea Lardelli, Balaji Sundararaman, John P. Donohue, Jens Lykke-Andersen, Frank Bennett, Manuel Ares Jr., Christopher B. Burge, J. Paul Taylor, Frank Rigo, Gene W. Yeo. **Neuron.** 2016 Nov 23;92(4):780-795. doi: 10.1016/j.neuron.2016.09.050. Epub 2016 Oct 20. PMID: 27773581; PMCID: PMC5123850.

EDUCATION

Metis Data Science Bootcamp

2019

University of California, San Diego
Double Major: B.S. Physics, B.S. Physiology & Neuroscience

2015