# 2022 g2p2pop Workshop

*Accessing Approaches to Study Epigenetic Regulation*
University of Nevada Las Vegas
June 6th & 7th, 2022

# Workshop Agenda

**Monday, June 6th**

*Morning Session:*
**Gather at: Juanita Greer White Life Sciences Building (WHI) – Central Atrium**
8:00 to 8:50 – Coffee and Breakfast (provided in WHI Atrium)
**Head to Science and Engineering Building (SEB) – Auditorium # 1311**
9:00 to 9:15 – Welcome Remarks: Loren Buck, Northern Arizona University & Frank van Breukelen, University of Nevada Las Vegas
9:15 to 10:15 – Plenary 1: Dr. Jason Podrabsky, Portland State University
*'Epigenetic regulation of development, dormancy, and stress tolerance in a vertebrate extremophile'*
10:15 to 10:45 – Coffee Break and Networking
10:45 to 11:45 – Plenary 2: Dr. Kathleen Keough, Fauna Bio
*'Genome evolution over multiple timescales and dimensions'*
11:45 to 12:00 – Q&A with Speakers 1 & 2 / Networking
**Return to: Juanita Greer White Life Sciences Building (WHI) – Central Atrium**
12:00 to 1:00 – Lunch (provided in WHI Atrium)

*Afternoon Session:*
Breakout discussion and rotations through video interviews with Education Team

The afternoon breakout sessions will have a science as well as an education goal. Our breakout science exercise is to form integrative grant proposal ideas and determine what expertise, resources, and analytics are required to answer a question that spans genomes to phenomes to populations. Our education goal is to record mini-interviews with participants to showcase diversity of paths within and paths to science.
1:00 to 1:15 – Dr. Jenifer Utz, Intro to interviews & logistics, Dr. Frank van Breukelen, Intro to breakout sessions
1:15 to 2:00 – Breakout / video - Rotation 1
2:00 to 2:45 – Breakout / video - Rotation 2
2:45 to 3:00 – Break
3:00 to 3:45 – Breakout / video - Rotation 3
3:45 to 4:30 – Breakout / video - Rotation 4
**Gather at: Juanita Greer White Life Sciences Building (WHI) – Room 105**
4:30 to 5:00 – Day 1 wrap-up & dinner group organization (dinner is optional and will be participant-funded)

**Tuesday, June 7th**

*Morning Session:*
**Gather at: Juanita Greer White Life Sciences Building (WHI) – Central Atrium**
8:00 to 8:50 – Coffee and Breakfast (provided in WHI Atrium)
**Head to Science and Engineering Building (SEB) – Auditorium # 1311**
9:00 to 9:15 – Education and Outreach update: Jenifer Utz, University of Nevada Las Vegas
9:15 to 10:15 – Plenary 3: Dr. Mira Han, University of Nevada Las Vegas
*'Learning the effect of 3D genome organization on gene expression'*
10:15 to 10:45 – Coffee Break and Networking
10:45 to 11:45 – Plenary 4: Dr. Irene Kaplow, Carnegie Mellon
*'Relating enhancer genetic variation across mammals to complex phenotypes using machine learning'*
11:45 to 12:00 – Q&A with Speakers 3 & 4 / Networking
**Return to: Juanita Greer White Life Sciences Building (WHI) – Central Atrium**
12:00 to 1:00 – Lunch (provided in WHI Atrium)

*Afternoon Session:*
1:00 to 4:00 – Afternoon breakout discussions, prepare and present proposal ideas
Education Team will be editing mini-interview videos – Room FDH 235
**Gather at: Juanita Greer White Life Sciences Building (WHI) – Room 105**
4:00 to 4:45 – Presentation and discussion about 2023 Brazil Workshop (Ramiro Melinski)
4:45 to 5:00 – Day 2 wrap-up & dinner group organization (dinner is optional and will be participant-funded)
Evening – RCN Steering Committee Meeting

**Speaker Info and Abstracts:**

**Jason Podrabsky, Portland State University**

**Title:** Epigenetic regulation of development, dormancy, and stress tolerance in a vertebrate extremophile

**Abstract:** It has been long appreciated that the diversity of form and function observed in the natural world cannot be explained by differences in genomic sequences alone. Large portions of the genome are highly conserved, in sequence and organization, across incredibly diverse and divergent species. It is differences in the timing and location of genomic expression, during development and otherwise, that generates a great deal of the observed diversity in form and function. Epigenetics is the study of heritable changes in gene expression without an underlying change in DNA sequence. These changes can be mediated by a variety of factors including DNA methylation, histone modifications, changes in chromatin architecture, small noncoding RNAs, and provisioning of small molecules such as hormones or morphogens. The annual killifish *Austrofundulus limnaeus* is a vertebrate extremophile that produces embryos that can survive for months without oxygen and years without access to liquid water. Survival in their ephemeral and often unpredictable habitat is dependent on epigenetic mechanisms that regulate development, entrance into dormancy, and induce tolerance to environmental stress. In this talk we will review some of the major mechanisms for epigenetic regulation in the context of annual killifish development.

**Biography:** Jason Podrabsky is a Professor of Biology at Portland State University interested in how the environment affects embryonic development, and how these effects can persist in an individual and their offspring. His work focuses on mechanisms that support metabolic dormancy, and the importance of oxygen for supporting metabolism during development and in adult organs such as the heart and brain. He studies embryos of annual killifish from Venezuela that can survive for years without water and months in the complete absence of oxygen by arresting their development, stopping their heart, and entering a coma-like state of dormancy. The mechanisms these embryos use to survive extreme conditions may help us understand how to devise interventions to improve human performance in extreme environments or provide mechanisms to engineer human cells and tissues to function in novel and stressful environments. Website: https://www.pdx.edu/podrabsky-lab/

**Kathleen Keough, Fauna Bio**

**Title:** Genome evolution over multiple timescales and dimensions

**Abstract:** The majority of the vertebrate genome is noncoding, however, unlike in genes where we can predict which sequences will generate which amino acids, functional characterization of the noncoding genome remains elusive. Characterizing functionality of the noncoding genome is important because that is where many of the most interesting features are located, including variants linked to disease via GWAS, and Human Accelerated Regions (HARs), regions conserved in vertebrates that evolved fast specifically in humans. Epigenomics help us functionally characterize the noncoding genome, identifying regulatory features such as enhancers which control gene expression. The genome also folds in organized patterns in 3D, bringing elements separated by millions of nucleotides in close proximity. In this talk I will discuss how 3D genome organization may have contributed to the human-specific evolutionary patterns of HARs. I will then discuss how Fauna Bio is translating extraordinary genomics in animals to human therapeutics, with a focus on how we're leveraging epigenomic data.

**Biography:** Kathleen earned a PhD in Pharmacogenomics and Pharmaceutical Chemistry from UCSF. She next trained as a postdoctoral scholar in Katherine Pollard's lab at Gladstone Institutes, studying comparative genomics in the context of the 3D genome. Prior to her graduate studies, she worked at the Centers for Disease Control based in Los Angeles. Her research has focused on studying genomics in 2D and 3D, encompassing CRISPR gRNA design, investigating how the 3D genome influences human evolution and vice versa, and predicting risk of COVID transmission to non-human species. She was motivated to join Fauna in order to learn more about how similarities and differences between humans and other species can be leveraged for innovation. Website: https://www.faunabio.com/our-team

**Mira Han, University of Nevada Las Vegas**

**Title:** Learning the effect of 3D genome organization on gene expression.

**Abstract:** The development of the Hi-C technology that can identify long range chromatin interactions in a high throughput, genome-wide manner brought us new insights on how the genome is organized in the nucleus. Hi-C data showed us that there are defined regions in the genome that interact more frequently than others that we now call Topological Associating Domains (TADs). In this talk, I will summarize the literature that supports the prevailing idea that these TADs affect gene expression, and also review some puzzling observations that seem to be inconsistent with the idea. After reviewing the literature on TADs, I will introduce our work

that utilizes published massively parallel enhancer perturbation data to learn the factors that affect gene expression. One of the questions we aim to answer with our machine learning approach is whether the long-range chromatin interactions between the enhancers and the target gene affects the outcome of the enhancer perturbation on the target gene.

**Biography:** I am an assistant professor in the School of Life Sciences at University of Nevada, Las Vegas. I obtained my Ph.D. in Informatics at Indiana University Bloomington, studying genome evolution through gene duplication, loss, and transposition with Dr. Matthew Hahn. Afterwards, I moved to Durham, NC to work on gene transposition at the National Evolutionary Synthesis Center (NESCent). I joined UNLV in 2013. The Han lab studies the evolution of genome structure. We use bioinformatics to investigate how genomes change through gene duplication, loss, and gene transpositions. We also are interested in the phenotypic effects of Copy Number Variations (CNVs), indels and transposable element polymorphisms. Website: https://hanlabunlv.github.io/

**Irene Kaplow, Carnegie Mellon**

**Title:** Relating enhancer genetic variation across mammals to complex phenotypes using machine learning

**Abstract:** Advances in the genome sequencing have provided a comprehensive view of cross-species conservation across small segments of nucleotides. These conservation measures have proven invaluable for associating phenotypic variation, both within and across species, to variation in genotype at protein-coding genes or very highly conserved enhancers. However, these approaches cannot be applied to the vast majority of enhancers, where the conservation levels of individual nucleotides are often low even when enhancer function is conserved and where activity is tissue- or cell type-specific. To overcome these limitations, we developed the Tissue-Aware Conservation Inference Toolkit (TACIT), in which convolutional neural network models learn the regulatory code connecting genome sequence to open chromatin in a tissue of interest, allowing us to accurately predict cases where differences in genotype are associated with differences in open chromatin in that tissue at enhancer regions. We established a new set of evaluation criteria for machine learning models developed for this task and used these criteria to compare our models to models trained using different negative sets and to conservation scores. We then developed a framework for connecting these predictions to phenotypes in a way that accounts for the phylogenetic tree. When applying our framework to motor cortex, we identified dozens of new enhancers associated with the evolution of brain size and vocal learning.

**Biography:** Irene Kaplow received her B.S. in Mathematics with a minor in Biology from the Massachusetts Institute of Technology in 2010. There, she began her career as a computational biologist while doing research with Bonnie Berger. She then went to graduate school at Stanford University, where she received her Ph.D. in Computer Science in 2017. At Stanford, she worked in the Hunter Fraser and Anshul Kundaje's labs to develop methods to analyze novel high-throughput sequencing datasets to better understand the roles of DNA methylation and Cys2-His2 zinc finger transcription factor binding in transcriptional regulation. Irene is now a Lane Postdoctoral Fellow in Andreas Pfenning's lab in the Computational Biology Department at Carnegie Mellon University, where she is developing methods to identify regulatory elements involved in the evolution of neurological phenotypes that have evolved through gene expression. Website: https://www.researchgate.net/profile/Irene-Kaplow