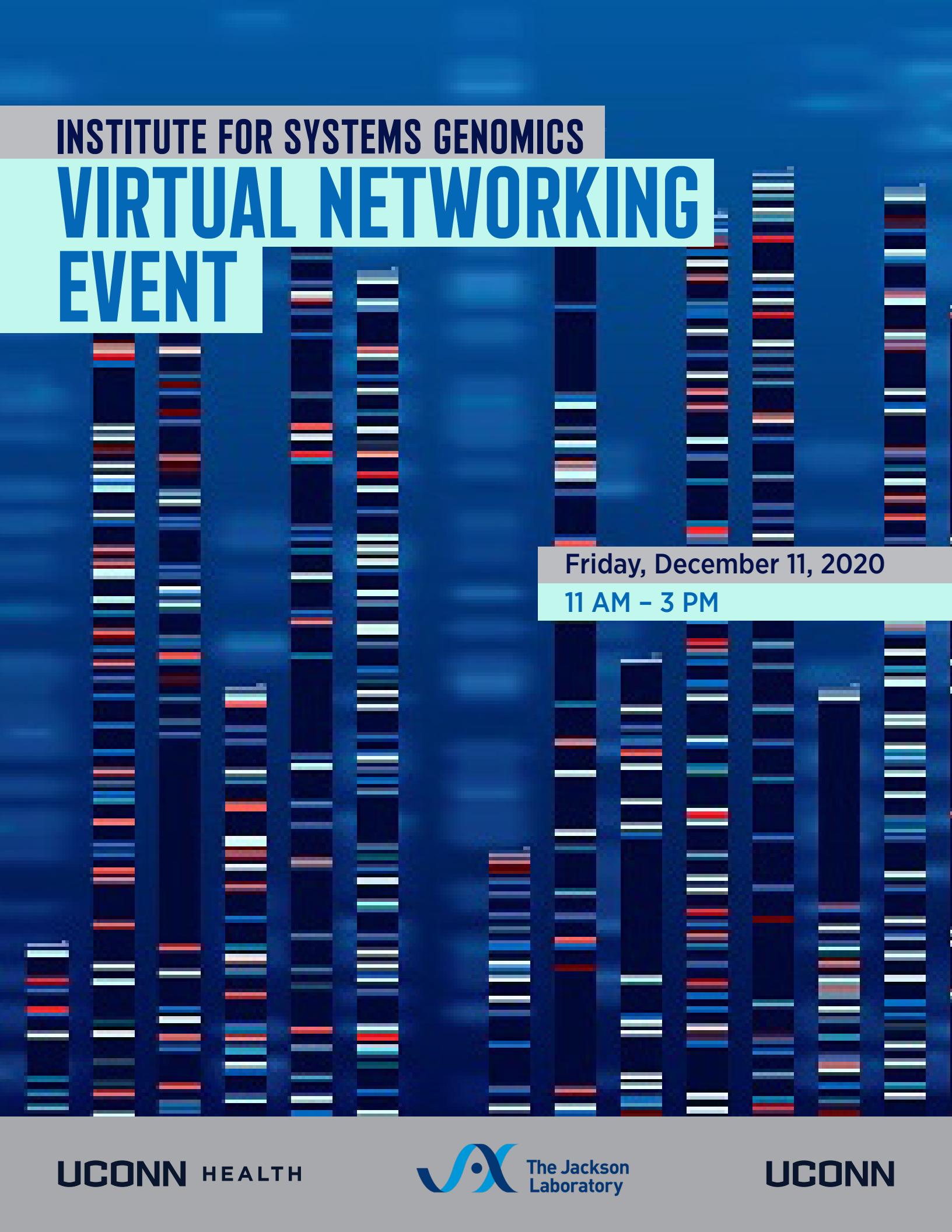

INSTITUTE FOR SYSTEMS GENOMICS

VIRTUAL NETWORKING EVENT

Friday, December 11, 2020

11 AM – 3 PM

UCONN HEALTH

The Jackson Laboratory

UCONN

Agenda:

Much appreciation for your patience if the timing on our agenda becomes slightly off as we move through this new virtual format.

11:00AM	Welcome **Rachel O'Neill, Ph.D., University of Connecticut**
11:05	Introductory Remarks **Provost Carl Lejuez, University of Connecticut**
11:10	*Finishing the complete sequence of a human genome* **Adam Phillipy, Ph.D., National Institutes of Health**
12:00PM	Center for Genome Innovation **Bo Reese, Ph.D. University of Connecticut**
12:05	Computational Biology Core **Jill Wegrzyn, Ph.D., University of Connecticut**
12:10	Single Cell Genomics Facility **Paul Robson, Ph.D., The Jackson Laboratory for Genomic Medicine**
12:15	Microbial Analysis, Resources and Services Facility **Kendra Maas, Ph.D., University of Connecticut**
12:21	Concurrent breakout sessions with each core facility director
12:42	Poster Session
1:30	*Characterizing the regulatory activity of 5'-UTR elements on translation initiation in vertebrates* **Jean-Denis Beaudoin, Ph.D.**, **University of Connecticut School of Medicine**
1:43	*A dynamical systems and control theory approach to cell reprogramming* **Paola Vera Licona, Ph.D., University of Connecticut School of Medicine**
1:56	*Alterations of host-gut microbiome interactions in patients with multiple sclerosis* **Yanjiao Zhou, M.D., Ph.D.**, **University of Connecticut School of Medicine**
2:09	*Exploring B chromosomes at the genomic level* **Stacey Hanlon, Ph.D.**, **University of Connecticut**
2:22	*Migration, microbes & metatranscriptomics* **Sarah Hird, Ph.D.**, **University of Connecticut**
2:35	*Chemistry between genes: Metabolomics for precision medicine and precision immunology* **Shuzhao Li, Ph.D., The Jackson Laboratory for Genomic Medicine**
2:50	Closing Remarks **Charles Lee, Ph.D., FACMG, The Jackson Laboratory for Genomic Medicine**
3:00	Virtual Happy Hour

Speaker Bios:

Finishing the complete sequence of a human genome

Adam Phillipy, Ph.D.
Head of the Genome Informatics Section, National Human Genome Research Institute
National Institutes of Health

Dr. Adam Phillippy is a Senior Investigator and head of the Genome Informatics Section at the National Human Genome Research Institute. His lab develops foundational methods for genomics, focusing specifically on the problems of genome sequencing, assembly, and comparative genomics. As a co-founder of the Telomere-to-Telomere consortium, he is currently working towards finishing the remaining gaps in the human reference genome using long-read sequencing technologies. His lab homepage can be found at https://genomeinformatics.github.io/

Characterizing the regulatory activity of 5'-UTR elements on translation initiation in vertebrates

Jean-Denis Beaudoin, Ph.D.
Assistant Professor, Department of Genetics and Genome Sciences
University of Connecticut School of Medicine

Dr. Beaudoin earned a B.Sc. in Biotechnology from Université de Sherbrooke in 2006. In 2009, he completed his Master degree, and started his PhD both in the laboratory of Dr. Jean-Pierre Perreault at Université de Sherbrooke in the Department of Biochemistry. He earned his PhD in 2013, where he worked to characterize the ability of the G-quadruplex structure to modulate RNA activities. He was a postdoc in the Giraldez Lab from 2013 to 2019, investigating how RNA structures regulate post-transcriptional events during the maternal-to-zygotic transitions in zebrafish. Dr. Beaudoin has been awarded the NIH Pathway to Independence Award (K99/R00). In January 2020, he became a principal investigator at UConn Health in the Department of Genetics and Genome Sciences. His lab studies how RNA structure functions control gene expression during vertebrate development. For more information please visit beaudoinlab.org and follow him on twitter @JeanDenisB35.

Exploring B chromosomes at the genomic level

Stacey Hanlon, Ph.D.
Assistant Professor, Department of Molecular and Cell Biology
University of Connecticut

After graduating from Texas A&M University with a B.S. in Biology, Dr. Hanlon attended the University of California, San Francisco to obtain a Ph.D. in Biochemistry and Molecular Biology. Her extensive molecular training in league with her interests in chromosome biology led

Dr. Hanlon to pursue the unknown biology of B chromosomes in the fruit fly *Drosophila melanogaster* during her postdoctoral studies at the Stowers Institute for Medical Research. Dr. Hanlon's work has launched the field of B chromosome biology in *D. melanogaster* and exposed exciting new directions about how these chromosomes move, form, and evolve. Starting in January, the Hanlon Lab will continue to explore the B chromosomes using both classic genetic and modern genomic approaches.

Migration, microbes & metatranscriptomics

Sarah Hird, Ph.D.
Assistant Professor, Department of Molecular and Cell Biology
University of Connecticut

Dr. Hird's lab researches the ecology and evolution of host-associated microbiomes, with a particular focus on the avian microbiome. She got her B.S. from the University of Idaho and Ph.D. from Louisiana State University. She was a postdoctoral fellow at the University of California, Davis prior to joining UConn's Department of Molecular and Cell Biology in 2016.

Chemistry between genes: Metabolomics for precision medicine and precision immunology

Shuzhao Li, Ph.D.
Associate Professor, The Jackson Laboratory for Genomic Medicine

Dr. Li's research focus is metabolomics and immunometabolomics. He got his BA from Sichuan University and PHD from University of Southern Mississippi. He was an assistant professor at Emory University prior to his joining of JAX in 2020. He is part of the NIH Metabolomics Consortium, and studies human immunology via vaccine projects.

A dynamical systems and control theory approach to cell reprogramming

Paola Vera Licona, Ph.D.
Assistant Professor, Center for Quantitative Medicine
University of Connecticut School of Medicine

Dr. Vera-Licona is an assistant professor at the Center for Quantitative Medicine at UConn Health, the Cell Biology and Pediatrics departments and a joint faculty member in the Mathematics department at UConn Storrs. She is an interdisciplinary scientist with a background in mathematics and computational systems biology. She has extensive experience in the design, implementation, and application of mathematical and computational tools for data-driven modeling, analysis and control of biological networks. Dr. Vera-Licona's research group works at the intersection of computational systems medicine, mathematical biology and bioinformatics with a particular focus on cancer biology and immunology. Biological systems of interest include gene regulatory

networks and intracellular signaling networks where the aim is to understand and control the cells' intricate regulatory programs.

Alterations of host-gut microbiome interactions in patients with multiple sclerosis

Yanjiao Zhou, M.D., Ph.D.
Assistant Professor, Medicine
University of Connecticut School of Medicine

Dr. Zhou is a computational biologist. She was one of the lead analysts in the first and second phase of Human Microbiome Project. She has been involved in over 40 microbiome-related projects that study the role of microbiome in health and disease. The major research interest of her laboratory is to develop microbiome-based diagnostics and therapeutics for multiple sclerosis as well as other diseases, with a focus on how nutrition and microbial metabolites influence the gut-brain axis.

Core Facilities:

Center for Genome Innovation (CGI) cgi.uconn.edu

The CGI offers a variety of laboratory training opportunities as well as sample QC, NextGen library preparation/sequencing and genotyping services. These services are available to UConn-affiliated researchers across all campuses and range from single run instrument access through full-service NextGen library preparation and sequencing. All services are also available to external users and for-profit companies.

The CGI also offers laboratory-based workshops for a variety of NextGen sequencing applications, including RNA-Seq, amplicon sequencing and ChIP-Seq. Free consultations are also available for experimental design, budgeting and troubleshooting.

Computational Biology Core (CBC) bioinformatics.uconn.edu

The CBC provides computational power and technical support to both academia and industry. These services are available to faculty and students within the University system. Services provided: research collaboration; project design and data analysis consultation; bioinformatics support for NextGen sequencing; software development; and access to computational resources.

Upcoming Events: Visit a complete list at https://bioinformatics.uconn.edu/cbc-workshops/

Workshop: Introduction to HPC and Linux
Date: January 11, 2021
Time: 8:30am-12:00pm
Location: online

Workshop: Genomes: Sequencing and Assembly
Date: January 11-14, 2021
Time: 9:00am-12:00pm
Location: online

Single Cell Genomics Center core.uconn.edu/resources/JAX

The Jackson Laboratory (JAX) Single Cell Biology Laboratory (SCBL) develops and offers single cell capabilities to JAX & UConn investigators. We provide numerous capabilities in this area including single cell RNA-seq, single nucleus ATAC-seq, spatial transcriptomics, imaging mass cytometry, and, newly added, an organoid screening platform. We can advise on design of experiments through to interpretation of results in addition to our data generation capabilities, and our happy to assist in grant applications by providing Letters of Support.

Microbial Analysis, Resources, and Services (MARS) mars.uconn.edu

The Microbial Analysis, Resources and Services (MARS) facility supports research specializing in the analysis of microbial samples and high-throughput processing of nucleic acids. Examples include the characterization of microbiomes, sequencing of small genomes, 96-well and 384-well PCR setup or DNA quantification and other automated liquid handling applications. Services are available a la carte, ranging from fee-for-service to unassisted use of the equipment by trained and certified users.

Posters:

Salp in Bloom: Genome dynamics provide insight into *Salpa thompsoni's* reproductive success

Presenter: Kate Castellano, Ph.D. Candidate
Affiliated Lab: Dr. Rachel O'Neill, Department of Molecular and Cell Biology, UConn

Warming trends in the Southern Ocean have altered trophic dynamics in favor of the pelagic tunicate, *Salpa thompsoni*, which is showing rapid population expansion (blooms), that displace other key species. Salp blooms are facilitated by their reproductive life history, which consists of seasonal alternations between sexual and asexual stages. However, we lack a foundational understanding of the genomic features that define this species and support salp bloom formation. Both paired-end short read and linked-read sequencing approaches have failed to generate a quality genome assembly for any salp species, likely the result of the high repeat content of the salp genome. Herein, using Oxford Nanopore long read sequencing, *de novo* assembly and comprehensive transcriptomics, we have derived a new assembly for *S. thompsoni*, consisting of 9,029 contigs, an N50 of 188 kb, and genome coverage of 78%. Through this work, we have discovered strong secondary structures within the *S. thompsoni* genome that dramatically affect sequencing efficiency. Our analyses of these secondary structures led to the discovery of abundant G-quadruplex sequences distributed throughout the *S. thompsoni* genome at a significantly higher frequency compared to other tunicate species, suggesting such structures are a defining feature of this salp genome. The link between these G-quadruplex sequences, *de novo* gene and repeat annotations and transcriptional profiles across life stages will be discussed. Collectively, out results provide novel insights into the function of unique genomic features in the regulation of genome stability between asexual to sexual reproduction.

From CartograTree to CartograPlant: Cyberinfrastructure to improve plant health and productivity in the context of a changing climate

Presenter: Irene Cobo Simón, Postdoctoral Fellow
Affiliated Lab: Dr. Jill Wegrzyn, Ecology and Evolutionary Biology, UConn

Climate change is threatening plant health and productivity at all spatial scales. To date, it remains largely unknown whether plant breeding can keep pace with the rate and direction of environmental change, as well as species' adaptive potential. In addition, the frequency and impact of invasive pests and pathogens is increasing as a consequence of globalization and is exacerbated by climate change. Both forest trees and agricultural crop species are threatened. Hence, the identification of genes controlling traits which provide tolerance to both biotic and abiotic stresses constitutes one of the most important research objectives in evolutionary ecology. However, progress may be limited as these analyses require the integration of traditionally disparate data sources: genotypic, phenotypic and environmental.

CartograTree is the first web-based application which integrates genotypic, phenotypic and environmental data, and associated meta-data, from georeferenced forest tree individuals, connected by analytic workflows. The development of this robust and integrative resource is relevant and timely since nowadays the applications which collect these data are not accessible. In specific, the analytic workflows are designed to facilitate meta-analysis for association studies (with measured traits and/or environmental variables). CartograTree is now being extended as CartograPlant to include more plant species, to address issues related to wild type variation and invasives. Here, we describe the recent updates in data sources, functionalities, and workflows offered by the web-based application CartograPlant.

Evolution and spread of rabies virus in Connecticut and its neighboring states

Presenter: Julia Desiato, Senior Undergraduate Student
Affiliated Lab: Dr. Dong-Hun Lee, Department of Pathobiology and Veterinary Science, UConn

Understanding repression of maternal 15q11-q13

Presenter: Rachel Gilmore, Ph.D. Candidate
Affiliated Lab: Dr. Stormy Chamberlain, Department of Genetics & Genome Sciences, UConn Health

Three neurodevelopmental disorders are caused by genetic aberrations on chromosome 15q11-q13: Angelman syndrome (AS), Prader-Willi syndrome (PWS), and 15q duplication syndrome (Dup15q). These disorders are related because the genes responsible for them are subject to regulation by genomic imprinting and share a common imprinting control region, the PWS imprinting center. Currently, little is known about the mechanism behind repression of the maternal chromosome 15q11-q13 region. Using AS and PWS induced pluripotent stem cells (iPSCs) in combination with isogenic human embryonic stem cells (hESCs) engineered with large deletions, it is our goal to determine the chromatin states of both alleles and mechanisms underlying repression of the maternal allele. We used a single CRISPR to target chromosome 15 specific GOLGA8 repeats, which comprise common breakpoints for all three disorders, in H9 hESCs. Cell lines harboring ~5-8Mb deletions on both maternal or paternal alleles of chromosome 15q were created, generating models of AS and PWS, respectively.

We previously showed that by knocking out the KRAB-domain zinc-finger protein, ZNF274, the silent maternal allele can be activated, leading to a restoration of gene expression from the 15q11.2-q13 region. Using our engineered H9 hESCs described above, we have generated an inducible ZNF274 PWS hESC line in which ZNF274 can be expressed from ZNF274 cDNA under the control of a dox-inducible promoter to ask questions related to the mechanisms of maternal allele repression. These studies are important as they have implications for PWS therapeutics as

well as the mechanisms underlying maintenance of genomic imprinting in human neurons.

GABRA2 genetic variants and chromatin accessibility in induced pluripotent stem cell-derived neural cells and postmortem brain samples in the context of alcohol use disorder

Presenter: Alexandra Goetjen, Ph.D./MD Candidate
Affiliated Lab: Dr. Jonathan Covault, Department of Psychiatry, UConn Health

Approximately 8.5% of adults in the United States are afflicted by moderate or severe alcohol use disorder (AUD). Twin studies indicate a significant genetic contribution to the risk for developing AUD. Although several candidate loci have been identified, with the exception of coding variants in the ADH1B and ALDH2 genes, little is known about the molecular effects of loci associated with AUD. We have observed a correlation between genotype of AUD-associated synonymous SNP rs279858 located in *GABRA2* and transcription of *GABRA2* and *GABRB1* on chromosome 4p12. Publicly-available virtual circular chromatin conformation capture (4C) data on dorsolateral prefrontal cortex samples supports a hypothesis that chr4p12 $GABA_A$ gene expression is regulated in *cis* by a shared regulatory element. Allele-specific differences in chromatin accessibility and transcription factor binding may predispose individuals to complex polygenic disease. The assay for transposase-accessible chromatin followed by high-throughput sequencing (ATAC-seq) allows for identification of sites of allele-specific open chromatin. Initial sequencing of one rs279858 C/C neural culture and one rs279858 T/T neural culture was supplemented with analysis of publicly-available analysis on five post-mortem subjects through the Brain Open Chromatin Atlas (BOCA). There is a neural-specific peak of open chromatin that overlaps rs1442059, a SNP in strong linkage disequilibrium with rs279858. The Genome Editing Core at UConn Health has generated homozygote iPSCs of the opposite genotype at rs1442059, and while editing of this single base pair does not significantly alter chromatin accessibility in *GABRA2*, transcription of chr4p12 GABA-A receptor subunit genes increased when rs1442059 was changed from CC to TT.

Regulation of *UBE3A-ATS* Expression in Human Neurons

Presenter: Dea Gorka, Ph.D. Candidate
Affiliated Lab: Dr. Stormy Chamberlain, Department of Genetics and Genome Sciences, UConn Health

Angelman Syndrome (AS) is a neurodevelopmental disorder characterized by motor dysfunction, intellectual disability, severe seizures, and absent speech. AS is caused by loss-of- function from the maternally inherited allele of *UBE3A*. In most cell types, *UBE3A* is expressed from both the maternal and paternal alleles. In mature neurons, *UBE3A* is only expressed from the maternally-inherited allele. Thus, loss-of-function from the maternal allele leads to nearly complete loss of *UBE3A* RNA and protein in the brain. Imprinted (maternal-only) expression of *UBE3A* occurs because the paternal allele of *UBE3A* is silenced by a long non-coding antisense RNA, termed *UBE3A-ATS*. In non-neuronal cell types, transcription of *UBE3A-ATS* terminates before it silences *UBE3A*. However, in mature neurons, the transcription extends to the *UBE3A* locus and silences paternal *UBE3A*. The mechanism restricting *UBE3A* imprinting to neurons is not well understood. Our goal is to identify the underlying mechanism of how the neuron-specific

UBE3A-ATS transcript is regulated in a cell-type specific manner. This is important because activation of paternal *UBE3A* through the suppression of *UBE3A-ATS* transcription is a promising therapeutic strategy for AS.

We have recently reported a bipartite chromatin boundary that stops transcription of *UBE3A-ATS* in iPSCs and therefore restricts *UBE3A* imprinting to mature human iPSC-derived neurons. This discovery revealed that *UBE3A* imprinting requires both the removal of boundary function as well as increased expression of *UBE3A-ATS*. We hypothesize that the expression level of *UBE3A-ATS* are regulated by an increased usage of alternative *SNRPN* promoters in neurons. Here, we test this hypothesis using forebrain cortical neurons derived from AS patient iPSCs. This data will provide important insights into AS therapeutics and the underlying physiological deficits in AS neurons.

Examining Hypotheses of Centromere and CENP Co-Evolution in Macropods at a Genomic Scale

Presenter: Patrick Grady, Ph.D. Candidate
Affiliated Lab: Dr. Rachel O'Neill, Department of Molecular and Cell Biology, UConn

Macropods present a unique opportunity to study the role of epigenetics, especially of the centromere, in genome structural dynamics. Macropods have undergone huge karyotypic change throughout their evolutionary history, varying between 2n=22 and 2n=14. Previous work has shown that karyotypic changes are associated with breakpoint reuse at the centromere, resulting in recombination of syntenic blocks to form differing chromosome composition between species. This makes macropods one of the most dynamic examples to study chromosome biology. Macropod centromeres also represent a unique centromeric biology study opportunity. The dominant hypothesis in centromere biology is that the sequence of the centromere is highly variable due to mechanisms such as centromere drive, in which the sequence of the centromere and related binding proteins evolve at high rates due the unequal nature of female meiosis. However, macropod centromeres are composed of three highly conserved satellite sequences. How this unique centromere changes with the macropod genome architecture may change our understanding of how centromeres are determined and how they function. These questions can be addressed by studying the centromere composition with ChIPseq technology and comparing the evolutionary rates of the centromeric binding proteins. We are currently working on bringing this project to a comparative genomics scale with long read and Hi-C hybrid assemblies to directly study centromere function, starting with a chromosome level assembly of the wallaby model organism, *Macropus eugenii*.

The highly reorganized genome of the threatened eastern hoolock gibbon unlocks unique insights into centromere and chromosome evolution

Presenter: Gabrielle Hartley, Ph.D. Candidate
Affiliated Lab: Dr. Rachel O'Neill, Department of Molecular and Cell Biology, UConn

Whole Genome Analysis of Salmonella enterica subspecies diarizonae serovar 61:k:1,5 Isolates Associated with Abortions in Sheep

Presenter: Ji-Yeon Hyeon, Postdoctoral Fellow
Affiliated Lab: Dr. Guillermo Risatti, Department of Pathobiology and Veterinary Science, UConn

Telomere-2-Telomere cenTE analysis

Presenter: Savannah Hoyt, Ph.D. Candidate
Affiliated Lab: Dr. Rachel O'Neill, Department of Molecular and Cell Biology, UConn

Evolution of Mexican-lineage low pathogenic avian influenza (H5N2) viruses in Dominican Republic

Presenter: David Hyunjung Chung, Ph.D. Candidate
Affiliated Lab: Dr. Dong-Hun Lee, Department of Pathobiology and Veterinary Science, UConn

Coordination of SR protein expression through alternative splicing of poison exons

Presenter: Nathan Leclair, Ph.D. Candidate
Affiliated Lab: Dr. Olga Anczukow, JAX-GM

Alternative RNA splicing (AS) provides a layer of functional diversity to gene expression by creating multiple RNA isoforms from the same expressed gene. Proper regulation of AS by splicing factor (SF) proteins is key for normal development, and dysregulation of SFs is at the center of many human diseases including cancer. SR proteins are a family of SFs that are frequently altered in human tumors and promote the formation of pro-oncogenic RNA isoforms. SR protein expression is auto-regulated, whereby high expression of an SR protein triggers AS of highly conserved poison-exons in their own pre-mRNA that induce RNA degradation. However, the extent to which these poison-exons are cross-regulated by other SFs or how they function during disease states, such as cancer, is not well understood. Here, we show that SR protein poison-exons are differentially spliced in human tumors compared to normal tissue. We further uncover an extensive cross-regulatory network that coordinates the expression of evolutionarily related SR proteins. Focusing on TRA2β, a highly oncogenic SR protein that is upregulated in multiple cancer types, we utilize deletion mutagenesis and CRISPR/Cas13d gRNA screens to uncover regulatory sequences within its ultra-conserved poison-exon and flanking intron regions that dictate poison-exon inclusion. Additionally, we create CRISPR Artificial Splicing Factors for each SR protein (CASFx-SR) and demonstrate their utility in directing SF activity to different transcripts. Finally, we develop splice switching antisense oligonucleotides (ASOs), a class of FDA approved therapeutics, that promote inclusion of the *TRA2β* poison-exon and cause anti-tumor effects in breast cancer cell lines.

Reversion of Cisplatin Chemoresistance in a Triple Negative Breast Cancer Subtype

Presenter: Lauren Marazzi, Ph.D. Candidate
Affiliated Lab: Dr. Paola Vera-Licona, Center for Quantitative Medicine, UConn Health

Triple Negative Breast Cancer (TNBC) represents a diverse group of cancers with a high prevalence among African American, Hispanic ethnic groups, and a younger age of onset compared to other breast cancer subtypes. The Tandem Duplicator Phenotype (TDP) is a genomic configuration characterized by numerous distributed tandem duplications found in 40% of TNBCs. A subset of TDP tumors are initially highly sensitive to the chemotherapy drug cisplatin, but they eventually develop resistance. Development of acquired drug resistance can be viewed as a systems-level process driven by a core resistance intracellular signaling network, where corresponding long-term dynamic behaviors (attractors) can be associated to the resistant and sensitive phenotypes of the tumor. The reprogramming of a cancer cell's long-term behavior away from the resistance phenotype may enhance a tumor's sensitivity to treatment. This reprogramming can be achieved through interventions on the control targets of resistance-associated phenotypes. This project aims to take a dynamical systems approach to identifying ombinations of therapeutic targets for cisplatin resistance reversion in a TDP tumor in silico.

We have developed a pipeline for constructing a static intracellular signaling network with multi-omics data from a TDP TNBC Patient Derived Xenograft (PDX) model treated with cisplatin. The resultant network captures known TNBC dysregulated genes and cisplatin resistance-associated genes. Control targets of the network were identified using structure-based control theory. Combinations of resistance-reversion control target perturbations were estimated using signal flow analysis. These intervention target predictions will be validated experimentally to identify chemotherapy adjuvants for the TNBC TDP subtype.

Genomic characterization of two maples highlights genes involved in the stress response to acidic soils

Presenter: Susan McEvoy, Ph.D. Candidate
Affiliated Lab: Dr. Jill Wegrzyn, Ecology and Evolutionary Biology, UConn

Acer saccharum (sugar maple) is a key ecological hardwood native to Northeastern forests, declining in recent decades due to a variety of stressors brought on by climate change that are underlied by systemic susceptibility to particular nutrient deficiencies exacerbated by the acidification of soils. *Acer negundo* (box elder) provides a counterpoint, with its broad native range and tolerance to a variety of these abiotic stressors. Newly created reference genomes and annotations for both species allow for investigation into the adaptive potential of these diploid, highly heterozygous trees. The sequencing design consists of deep long read coverage of Pacific Biosciences SEQUEL and Hi-C data (*A. negundo*, 100x; *A. saccharum*, 65x) resulting in moderate-sized plant genomes, estimated at 590Mbp and 440Mbp, respectively. Gene annotation combined existing and novel approaches to evaluate gene prediction methods, leveraging RNA-Seq data generated for both species. Genomic comparisons among the three existing *Acer* and other land plants were used to identify putative expansions and contractions of gene families underlying the

unique and shared biology of these species. The new *A. saccharum* genome was used with stem tissue RNA-seq in a differential expression analysis comparing aluminum and calcium treatment plots at the Hubbard Brook Experimental Forest. This Long Term Ecological Research forest in the mountains of NH provides an ideal system to examine the impact of past, current, and future soil conditions. Integrated genomic and expression comparisons revealed insights into tolerance mechanisms, both in response to aluminum directly, and at the broader systemic level.

Alternatively Spliced Transposons Influence Isoform Diversity and Patient Survival in Breast Cancer

Presenter: Alex Nesta, Predoctoral Associate
Affiliated Labs: Drs. Christine Beck and Jacques Banchereau, JAX-GM

Investigating the Enrichment of Non-B-DNA at Drosophila melanogaster Centromeres

Presenter: Venkata Patchigolla, Undergraduate Student
Affiliated Lab: Dr. Barbara Mellone, Department of Molecular and Cell Biology, UConn

16S rRNA and Shotgun Metagenomic Sequencing Analysis of the Microbial Response to 2-Deoxy-Glucose Treatment

Presenters: Farzaneh Rastegari and Nickolina Doran, Ph.D. Candidates
Affiliated Lab: Dr. George Weinstock, JAX-GM

2-deoxy-glucose (2DG), a metabolic inhibitor which blocks the first reaction of glycolysis, is being evaluated as a potential therapeutic for autoimmune disease in mouse models. Our research was to determine the effects of 2DG treatment in the gut microbiome. Using both 16S rRNA sequencing and shotgun metagenomic sequencing, we were able to characterize the differences in microbiome composition between cohorts of 2DG treated and control mice. Not only did we establish that 2DG treated mice exhibit distinctly different microbiome compositions than control mice, but we were also able to identify gene and pathway abundance variations between the two cohorts. This analysis allows us to identify bacterial strains that appear to be resistant to 2DG, as well as the metabolic pathways these bacteria use to evade the glycolysis inhibition of 2DG. The influence of the microbiome in the body is robust, and understanding how bacteria respond to inhibition, such as that from 2DG, is a crucial step in the development and advancement of therapeutics.

Identification of a non-LTR retrotransposon at *Drosophila* centromeres

Presenter: Bryce Santinello, Ph.D. Candidate
Affiliated Lab: Dr. Barbara Mellone, Department of Molecular and Cell Biology, UConn

Testing models of centromere specification using genome editing

Presenter: Prachi Tandale, Ph.D. Candidate
Affiliated Lab: Dr. Barbara Mellone, Department of Molecular and Cell Biology, UConn

Inference of transcriptional regulators reveals genome-wide regulatory relationships in Symbiodinaceaen dinoflagellates

Presenter: Felipe Wendt Porto, Ph.D. Candidate
Affiliated Lab: Dr. Senjie Lin, Department of Marine Sciences, UConn

Understanding dinoflagellate mechanistic responses to changing ocean conditions is essential to expand our understanding of coral symbiosis. Transcriptional regulation ubiquitously controls responses in eukaryotic organisms, yet it is poorly understood in dinoflagellates, particularly regarding transcription factors (TFs). Here we used an integrative genomic approach to identify and characterize TFs in model symbiodiniacean dinoflagellates, and TF regulatory relationships in these dinoflagellates were inferred by combining the discovered information with transcriptomic data. A total of 1189 and 882 putative TFs and transcriptional regulators were found in *Fugacium kawagutii* and *Brevolium minutum*, respectively. Among the abundant TFs in dinoflagellates, 64 appear to be unique to the Symbiodinaceae family. Transcriptomic data provided evidence that *F. kawagutii* and *B. minutum*respectively utilize 38 and 44 TFs and transcriptional regulators, some of which plausibly regulate processes necessary for the maintenance of coral symbiosis like transmembrane lipid transport, transmembrane ion transport, photosynthesis, and heat stress response. Given the widely held notion that dinoflagellate gene regulation largely lies at posttranscriptional levels, it is striking to find that post-translational modifications like glycosylation and phosphorylation appear to be heavily transcriptionally regulated through TFs. We further discuss how evolution of TFs has possibly altered transcriptional regulation. Furthermore, the data shows that dinoflagellates seem to favor transcriptional regulation for stress responses or the regulation of metabolic processes, as previously postulated. This study lays a foundation for understanding what processes can be transcriptionally regulated in dinoflagellates and how symbiodiniacean dinoflagellates regulate gene expression in symbiosis with corals.

A new graph-based clustering method with application to single-cell RNA-seq data from human pancreatic islets

Presenter: Hao Wu, Ph.D. Candidate
Affiliated Lab: Dr. Yuping Zhang, Department of Statistics, UConn

Traditional bulk RNA-sequencing of human pancreatic islets mainly reflects the transcriptional response of major cell types. Single-cell RNA sequencing technology enables transcriptional characterization of individual cells, and thus makes it possible to detect cell types and subtypes.

To tackle the heterogeneity of single-cell RNA-seq data, powerful and appropriate clustering is required to facilitate the discovery of cell types. In this paper, we propose a new clustering framework based on a graph-based model with various types of distances. We take the compositional nature of single-cell RNA-seq data into account and employ log-ratio transformations. The practical merit of the proposed method is demonstrated through the application to the centered log-ratio transformed single-cell RNA-seq data for human pancreatic islets. The practical merit is also demonstrated through comparisons with existing single-cell clustering methods. The R-package for the proposed method can be found at https://github.com/Zhang-Data-Science-Research-Lab/LrSClust.

Easy-to-synthesize antimicrobial nanomaterials as sanitizers under challenging environments

Presenter: Roberto Vazquez-Munoz, M.Sc., Ph.D., Postdoctoral Fellow
Affiliated Lab: Dr. Anna Dongari-Bagtzoglou, Department of Oral Health and Diagnostic Sciences, UConn Health

Communicable diseases are one of the major threats to public health worldwide. Emergent multidrug-resistant pathogens are a particular challenge, especially in non-favorable environments –rural areas, mobile clinical units, developing countries, etc. Developing potent antimicrobial agents with a broad antimicrobial activity that can be applied under challenging conditions is critical. Nanoantibiotics –antimicrobial nanomaterials- are effective against emergent, multidrug-resistant pathogens, such as viruses, microorganisms, and protozoa. However, as some nanomaterials are expensive or difficult to synthesize or obtain, our goal was designing facile and non-expensive protocols for producing nanomaterials with potent, broad-range antimicrobial activity. We developed Easy-to-synthesize nanomaterials (E2s-nano) via simple 1-pot chemical reduction methods. Then, we characterized the obtained nanomaterials and assessed their antimicrobial activity. Our results confirm that our protocols produced silver and bismuth nanoparticles (AgNPs and BiNPs, respectively) with the expected nanostructured organization. Both nanoantibiotics displayed potent antimicrobial activity. The AgNPs MIC were 4, 2, and <0.5 µg mL^{-1}, for *S. aureus*, *C. albicans*, and *C. auris*, respectively, whereas for BiNPs, the MIC were 1, 8, and 2 µg mL^{-1}, respectively; regardless of their biological differences or the drug-susceptibility profile. Also, our nanomaterials are stable for weeks when properly stored. Moreover, these methods can be easily replicated and used under non-favorable conditions and environments. In summary, we designed two E2s-nano with wide-range antimicrobial activity against bacteria and fungi for sanitizing facilities and instrumentation under challenging conditions.

Time series transcriptomics reveal networks of co-expression and candidate disease genes during embryonic craniofacial development

Presenter: Tara N. Yankee, Ph.D. Candidate
Affiliated Lab: Dr. Justin Cotney, Department of Genetics and Gemone Sciences, UConn Health

Craniofacial disorders are among the most common of all congenital defects and have a clear genetic component as indicated by the over 500 Mendelian syndromes that exhibit orofacial clefts and sequencing studies that have identified several patient gene mutations.

Recent studies of Autism Spectrum Disorder and Congenital Heart Development have revealed that genes which have similar trends of gene expression, or co-expression, during development are organized in networks of genes with coherent biological functions. In particular, strong disease candidates are genes that are co-expressed with many other genes, serving as regulatory hubs within these networks. Identifying such networks of genes during craniofacial development could reveal new disease genes and potential targets of previously identified disease genes.

A majority of craniofacial development occurs early in pregnancy and to fully understand how craniofacial defects arise, it is essential to observe gene expression during this critical time period. To address this, we performed RNA-seq on human craniofacial tissue from embryonic development -4 to 8 weeks post conception. We constructed co-expression networks using weighted gene co-expression network analysis (WGCNA).
We identified ~30 modules of co-expressed genes and find enrichment of craniofacial relevant biology. We leveraged large functional genomics databases including GTEx and GnomAD to reveal genes that are specifically expressed in embryonic craniofacial tissue and regulatory hub genes which are resistant to mutation in the normal healthy population. Our analysis revealed dozens of novel disease candidate genes that warrant further study in human populations and mouse models.

EASEL: Efficient, Accurate, Scalable Eukaryotic ModeLs for *De-novo*Genome Annotation

Presenter: Sumaira Zaman, Ph.D. Candidate
Affiliated Lab: Dr. Jill Wegrzyn, Ecology and Evolutionary Biology, UConn

The increase in reference genomes and their associated contiguity is lagged by intelligent workflows for accurate detection of protein coding genes. The majority of pipelines utilize RNA-seq reads or pre-assembled transcripts for training (supervised or semi-supervised) Hidden Markov Models (HMM) to predict gene structures for species with minimal or substantial existing genomic resources. However, such pipelines struggle to predict uncommon gene structures (long introns, micro-exons), locate preferred translation initiation start site (TiS), distinguish pseudogenes, and inflate the gene space with respect to genome size. We present EASEL (Efficient, Accurate, Scalable Eukaryotic modeLs), a genome annotation tool that leverages deep learning, RNA folding, and functional annotation to enhance gene prediction accuracy. EASEL features a deep LSTM network that can learn species-specific patterns, predict non-canonical gene structures, and train in reasonable time. Existing high-quality alignments from BUSCO and related protein sources train the network to find patterns in TiS detection. The implicated genomic regions are further refined via supervised training with RNA folding and primary sequence conservation. Highly probable TiS are fed as hints into HMM for the extension of the gene models. Finally, predicted proteins are subject to functional and structural criteria. The pipeline is benchmarked for efficiency, completeness, and accuracy.