CLEMSON UNIVERSITY
CENTER FOR
HUMAN GENETICS

The Transcript
Newsletter of the Center for Human Genetics at Clemson University
Spring, 2021

From the Director

Winston Churchill's famous words "If you are going through hell, keep going" certainly ring true in the age of COVID. Clemson University has weathered the storm well, thanks to early establishment of containment policies and aggressive testing. The Center for Human Genetics has thus far navigated the COVID crisis well without any incidences. I am grateful to our faculty, students, and staff members for their unwavering compliance with our safety protocols, their dedication and good humor, which included this year a socially distanced zoom mediated holiday party, which turned out to be fun and enjoyable.

Despite the COVID restrictions, the CHG has managed to operate effectively and to retain an active academic environment. We have established a stimulating series of Monday afternoon discussion sessions to explore Advances in Human Genetics, which are conducted on-line and are thus easily accessible for colleagues at the main campus and colleagues at the Greenwood Genetic Center. In the spring semester Advances in Human Genetics will incorporate a graduate student journal club and monthly meetings focusing on RNA biology with support from the RNA society obtained by Lela Lackey.

We are also looking forward to our inaugural series of Distinguished Seminars in Human Genetics in conjunction with the Greenwood Genetic Center, which will feature outstanding speakers. This semester the series will feature three prominent women in science, Dr. Arpana Agrawal from Washington University in St. Louis, Dr. Carole Ober from the University of Chicago, and Dr. Teri Manolio from the National Human Genome Research Institute. Again, since these seminars will be conducted by zoom they are easily accessible to faculty and students at both the main campus and the Greenwood Partnership Innovation campus as well as research staff and clinicians at the Greenwood Genetic Center.

The CHG has expanded its participation in national and international collaborations. Of special notice is inclusion of the Center for Human Genetics at Clemson University in

Dr. Trudy F. C. Mackay, FRS, is the Self Family Endowed Chair of Human Genetics. She is a Fellow of the Royal Society of London, a member of the National Academy of Sciences of the USA, and recipient of the 2016 Wolf Prize.

an international consortium of 15 institutions sponsored by the European Commission Horizon 2020 program. This five-year project anchored at the University of Birmingham in the UK, is titled "Toward Precision Toxicology: New Approach Methodologies for Chemical Safety" and is supported by €19,305,583.75, of which Clemson University will receive €998,936.25. Investigators at the Center for Human Genetics will use the Drosophila model to identify genes and genetic networks that underlie susceptibility to some of the world's most prevalent environmental toxicants.

This coming year we anticipate additional growth of the Center for Human Genetics. We will be recruiting another assistant professor to complete the faculty of CHG at the Greenwood campus. Please, see the advertisement on page 10 of this issue of The Transcript.

This year's graduate student recruitment in Genetics and Biochemistry has been very successful with an outstanding group of applicants of excellent academic quality. Many prospective students have indicated their interest in working with faculty of the Center for Human Genetics and we are looking forward to welcoming a new cohort of bright young scientists to our Clemson family. We also welcome Dr. Xinji Li, who joined the Center as an affiliate member from the School of Mathematics and Statistics. Dr. Li will further enhance our considerable depth in computational biology and bioinformatics. Finally, we welcome our new staff member, research technician Yu-Chen (Janet) Pan, who will be joining the Mackay-Anholt laboratory.

Successful science depends on a supportive and collaborative environment and I am grateful to my colleagues, our staff members, and students, who collectively make sure that the Center for Human Genetics is a pleasant and productive working environment.

Fish, Flies, and Friendship: Collaborative Efforts Between the GGC and CHG

by Heather Flanagan-Steet

In the summer of 2018 we packed the house and the lab, and moved thirteen years of research, 5,000 zebrafish and 2 teenagers two hours due east. Not fully knowing what the future held we took the road less traveled, leaving the hustle and bustle of a major university, to set up shop at the Greenwood Genetic Center. A few hours north Trudy and Robert were also packing their house, their lab, thousands of fruit flies and moving south, where our paths would converge in the quiet community of Greenwood. Now two and half years later we have not only found good friends and great colleagues in our neighbors across the street, together we are constructing a new road map that holds the potential to bring much needed answers and hope to GGC patients and their families.

Once the boxes were unpacked and the labs set up, the real work began. With our respective research programs rolling again, we held monthly meetings to share our respective interests and expertise. Through these conversations the path to collaboration began to emerge. Trudy and Robert's knowledge of complex genetic traits and ours in disease pathogenesis provided the foundation to pursue decades of unsolved questions surrounding common and rare genetic disorders. In the first several months our commitment to collaborate was realized when two CHG graduate students, Rebecca MacPherson and Kristin Bussey, initiated their Ph.D. thesis projects on Coffin-Siris syndrome and MED12-related disorders - two syndromes with a longstanding history at GGC. Patients with these disorders exhibit a spectrum of features whose severity is highly variable even among individuals in the same family. As part of their theses research, Rebecca and Kristin are generating fly models of these disorders and using them to define the disease mechanisms. These promising young scientists also intend to probe the genome for variants that may modify disease severity. Through this approach, important new insight into disease processes, and targets for therapeutic intervention, will be gained.

Our collaboration was further solidified in 2020 by the submission of the first COBRE grant in Human Genetics. The COBRE combines the CHG's expertise in cutting edge sequencing and bioinformatics with two model systems (the fish and fly). The grant will support the innovative molecular and cellular platforms of four different investigators, along with outstanding genomic resources. The investigators (CHG's Fabio Morgante and Andrei Alexandrov, Clemson's Miriam Konkel, and GGC's Heather Flanagan-Steet) will utilize human cell and animal systems, as well as large genomic datasets to define the genetic elements and molecular mechanisms driving common and rare disease. The synergy between this diverse team of investigators will unlock many of the genome's undiscovered secrets.

Our collective efforts were recently expanded when a group of the GGC's clinicians, genetic counselors, and diagnostic directors joined forces with the CHG's bioinformaticians, Dr. Vijay Shankar and Maria Adonay, to determine the genetic cause for four GGC patients who remain undiagnosed following exome analyses. Using a combination of long and short read technologies, CHG researchers sequenced the genome of each patient and their parents. Employing a novel bioinformatic pipeline, Dr. Shankar's team has now identified several candidate variants, whose functional analyses are expected to provide each patient a lon-awaited diagnosis. As genomic sequencing becomes a standard of care for patients with genetic disorders, novel genetic alterations with unclear significance are often identified. The addition of new CHG faculty members, including Dr. Lela Lackey, further enhances the GGC's ability to functionally characterize these variants and provide patients with answers. Dr. Lackey's expertise with the regulatory elements present in mRNA untranslated regions (UTRs) provides a valuable opportunity to functionally analyze genetic variants present in RNA modifying proteins.

Although 2020 brought many unexpected challenges, culminating in lab closures, suspended projects and delayed time lines, members of the CHG and GGC collaborative teams continue to exhibit tremendous resilience. With the number of exciting projects shared between the CHG and GGC, 2021 looks very bright for our growing collaboration.

Heather Flanagan_Steet is the Director of Functional Studies at the Greenwood Genetic Center.

Clemson Center for Human Genetics Initiates RNA Society of South Carolina Meetings

by Lela Lackey

The protein coding portion of the human genome only comprises ~1% of the total number of base-pairs in the genome. Study of non-coding variation is an essential aspect of modern genetics. The non-coding genome includes intergenic regions that are not transcribed into RNA but regulate many aspects of DNA accessibility and transcription. In addition, the non-coding genome includes transcribed regions that produce RNAs that are not translated into protein. This includes ribosomal RNAs, long non-coding RNAs (lncRNAs), small RNAs and non-coding regions of messenger RNAs (mRNAs) like 5' and 3' untranslated regions and introns, which are removed from the mature mRNA transcripts. Historically, it has been difficult to identify pathogenic variants within these regions as they do not have a simple amino acid code that can be used to detect disruptive changes. However, the field of RNA biology has exploded with the discovery of new species of long and small non-coding RNAs and their myriad functions. It is clear that deregulation of RNA pathways can lead to disease. The RNA Society of South Carolina is a group of forward-looking researchers interested in understanding the roles of non-coding RNAs and how genetic variants within these regions drive disease.

The first official meeting of the RNA Society of South Carolina was an introductory session with 22 participants followed by small group discussion of emerging topics in RNA biology. There were several main topics that came up in discussion. LncRNAs were one major focus of conversation. Although lncRNAs are prevalent and many are important for essential biological processes, the majority are poorly functionally described. We see characterization of lncRNAs as an important avenue for future research. Specifically, we would also like to learn more about the relationship between lncRNA structure and function. For example, the XIST lncRNA binds to the X chromosome and regulated X-linked silencing by influencing chromatin structure. RNA secondary structure elements are thought to play a role in XIST-mediated gene silencing. Are RNA structural elements a common regulatory feature of lncRNAs? Do structure elements also play a role in mRNA regulation? How frequent are long range structural interactions and what are their roles in biology? These questions address cutting-edge aspects of lncRNA and RNA structural biology.

Another major focus of discussion was on network regulation of RNAs. A common thread of RNA regulation is multi-tasking. RNA binding proteins and microRNAs generally bind and regulate hundreds of RNA transcripts and work in concert to co-regulate systems. We do not understand the underlying processes that control this complicated system. Can we identify co-regulated transcript pathways? How can we harness advances in technology, like single cell sequencing, to identify co-regulated RNA networks? How does network regulation of RNAs contribute to larger biological processes? For example, the RNA binding proteins Regnase and Roquin recognize 3'UTR hairpin structures leading to their degradation and subsequent repression of the innate immune response. Can we use similar information to expand upon other even more complex networks of regulation? As geneticists, we are interested in how genetic variation affects RNA at the network level. How do many small variants influence the final phenotype? This is part of the omnigenetic model of complex traits and fits well with coordinated action of transcripts to subtly affect transcripts as a population. We look forward to experiments designed to test the ability of non-coding variants to contribute to omnigenetic control.

A third focus area was a continued interest in the intersection between RNA biology and disease. In addition to the impact of genetic variants on network level RNA regulation, (likely through upstream RNA binding proteins), we are also interested in how genetic variation within an RNA might influence its co-factor binding motifs, stability, localization and translational efficiency. Further research is important for linking these molecular mechanisms with human disease. Do non-coding variants commonly drive serious human disease? How can we identify which non-coding variants that disease and reclassify these variants as pathogenic? We talked about the need for further work on defects of RNA processing, such as disruption of RNA binding proteins like the DEAD-box helicases, and how they generally relate to disease. Finally, we look forward to additional research on how to take new knowledge surrounding RNA processing and function and use this information to design new therapeutics. This includes targeting newly identified RNA-based molecular mechanisms with conventional drugs and harnessing the power of RNA as a therapeutic itself. The power of RNA therapeutics is evident in the SARS-CoV-2 mRNA-based vaccines currently being distributed across the world.

RNA Society of South Carolina meetings will continue into 2021 as part of our Advances in Human Genetics seminar series, funded by the International RNA Society and Lexogen. We hope to explore many of the issues we discussed as emerging topics in the field of RNA biology over the next year. There are many exciting advances to pursue, especially at the intersection of genetics and disease.

Lela Lackey is Assistant Professor of Genetics and Biochemistry
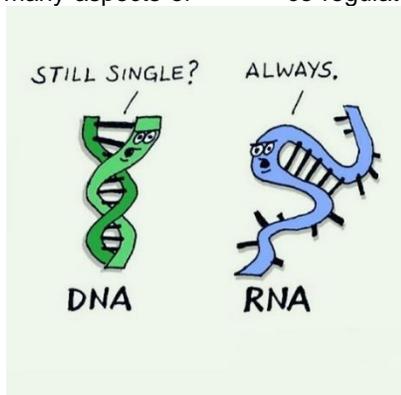
Mobile Elements: Sources of Genetic Variation

Only about 2-3% of the human genome encodes for proteins. The remaining 98% was once considered "junk" DNA: evolutionary remnants that are neither functional nor deleterious, and hence are continuing to segregate and pile up in the genome, where they represent a growing repository of genetic ballast. It has become increasingly clear, however, that the non-coding portion of our DNA is anything but junk and contains regulatory elements that contribute to genetic variation in gene expression and are relevant to health and disease.

One category of the non-coding portion of the genome encompasses transposable elements, initially identified in maize by Barbara McClintock as "jumping genes", a discovery for which she received the Nobel Prize. The presence of transposable elements is widespread among both vertebrate and invertebrate genomes and with these transposons in aggregate is referred to as the "mobilome."

In humans, transposable elements make up more than half of the human genome. Some continue to create new insertions, which can disrupt codons, splice site junctions, and regulatory sequences that affect the structure and function of mRNAs. It has been estimated that 0.2% of human genetic diseases may result from transposable element insertions. Despite the fact that transposable elements represent a major source of genetic instability in mammalian genomes, they remain relatively understudied.

Dr. Miriam Konkel, an Assistant Professor in the Department of Genetics and Biochemistry and the Center for Human Genetics is leading the way toward deciphering the evolution of transposable elements in primates and their effects on gene regulation. Konkel received an M.D. from the Charité, Humboldt University and completed residencies in Pediatrics and Internal Medicine at the Charité, Universtätsmedizin Berlin in Germany. In 2005, she joined Louisiana State University as a postdoctoral researcher and in 2011 became a research assistant professor at LSU. She joined Clemson University in 2017. In 2018, she received the University Research, Scholarship and Artistic Achievement Award.

Dr. Miriam Konkel

Konkel's laboratory explores how genomes evolve with respect to mobile elements, and how mobile elements evolve over time. Much of her work is done as part of large international consortia. The analysis of transposable elements requires novel custom-tailored computational approaches. Konkel's work focuses especially on a category of transposable elements, designated *Alu* elements and LINE1s. There are more than one million *Alu* elements and 500,000 LINE1s sprinkled across the human genome. They are highly polymorphic and inserted in genomic regions with repetitive sequences. This poses unique challenges for their annotation and for identifying *Alu* elements that are capable of transposition. Identifying such elements is particularly important for making the link between the mobilome and human disease.

Movement of transposable elements occurs not only in the germ line; it also may occur in somatic cells, during (for example) aging when activation of transposable elements may contribute to neurodegenerative disorders. According to Konkel:"A better annotation and understanding of transposable element expansion dynamics will be the foundation for future regulatory studies and allow for the investigation of effects of somatic transposition during an individual's lifespan."

The differently colored kernels in corn are the result of "jumping genes"

Predicting Complex Phenotypes

Computational approaches are becoming increasingly important in understanding the relationship between genotype and phenotype. Rare disorders that arise from genetic perturbations with large effects, such as trisomy 21 that results in Down's syndrome, can be readily predicted with high accuracy. However, most common diseases, including cardiovascular diseases, cancer and diabetes, emerge from the interplay of multiple segregating alleles, each contributing a portion of the risk. Summing all those fractions of risk allows us to calculate a "polygenic risk score" as an indication of disease predisposition. However, those scores still do not approach the high accuracy needed to confidently predict disease risk, which is one prerequisite for personalized medicine. To complicate things further, disease risk is affected by gene-gene and gene-environment interactions. Thus, predicting complex phenotypes requires developing sophisticated statistical methods, which take into account both genetic and environmental information from large populations. This is the focus of Dr. Fabio Morgante's research program.

Dr. Morgante started his scientific career in Italy in the field of animal breeding. Although one may wonder what the connection is between breeding cows and predicting human disease, animal breeders have led the field of quantitative genetics for most of the 20th century developing fundamental principles that govern the expression of complex traits. Thus, it was a logical step for Morgante to apply those principles to developing prediction algorithms for human phenotypes. After earning a Master of Science in Animal Breeding and Genetics at the University of Edinburgh in Scotland, he pursued his doctoral research in Genetics at North Carolina State University with Trudy Mackay, where he simultaneously completed a Master's degree in Statistics.

Dr. Fabio Morgante

In the Mackay laboratory, in collaboration with Drs. Peter Sørensen and Daniel Sorensen at Aarhus University in Denmark, he took advantage of natural genetic variation in a panel of several hundred inbred lines of fruit flies, Drosophila, which were derived from a natural fly population at the farmers' market in Raleigh and which had fully sequenced and well characterized genomes. Using this model system, he developed statistical methods and analytical strategies for genomic prediction and showed that prediction accuracy could improve significantly when gene-gene interactions and biological information are taken into account. Following his discoveries in the Mackay laboratory he joined the Section of Genetic Medicine in the Department of Medicine at the University of Chicago as a postdoctoral fellow advised by Drs. Matthew Stephens and Yang Li, where he continued to develop advanced computational models to increase accuracy of phenotypic prediction. This work was selected for an oral presentation at the recent International Conference of Quantitative Genetics.

In 2020, Morgante joined the faculty of the Clemson University Center for Human Genetics, where he is establishing a research program, which investigates the interaction between genetic make-up and environmental influences for risk of cardiovascular disease. These studies use massive data sets from hundreds of thousands of individuals with both documented genetic and environmental information from large data collections, such as the UK Biobank. "Fabio Morgante works at the forefront of genomic prediction," says Center Director Dr. Trudy Mackay, "Disentangling genetic factors from environmental factors to predict complex traits is an enormous challenge, which will have huge impact on human welfare. We are fortunate to have the talent of Dr. Fabio Morgante at our Center to lead this courageous endeavor."

"One brain's blueprint may promote joy more readily than most; in another, pessimism reigns. Whether happiness infuses or eludes a person depends, in part, on the DNA he has chanced to receive."

— Thomas Lewis, A General Theory of Love

A Conversation with Leigh Anne Clark

Could you briefly describe your research program?

My laboratory studies inherited disorders of domestic dogs. Our primary goal is to identify deleterious variants and develop genetic assays that can be used to reduce the frequency of genetic diseases in dog populations. With genetic tools, dog breeders can predict individual disease risk, select mate pairs that will yield healthy litters, and preserve genetic variation in their breed. Dogs have hundreds of naturally occurring diseases that closely resemble genetic diseases of humans, so our research also aims to identify new candidate genes, regulatory regions, and pathways that may underlie human disorders.

What inspired you to study canine inherited diseases?

I became passionate about all-things dog when I was about five years old. As a child, I worked for a Shetland sheepdog breeder and learned my first genetics lessons predicting coat colors and patterns. I started on a path to veterinary medicine as an undergraduate at Texas A&M University, which at the time seemed like the best way to turn my passion into a career. Chance led me to a laboratory in the veterinary school that studied canine inherited diseases. It was an exciting time to be in canine genetics because it was a small field and there was a lot of work left to firmly establish the dog as a model system for the study of human inherited diseases. I join the laboratory as a doctoral student and remained a member through my post-doctoral studies.

Do you study specific breeds of dogs? If so, which ones and why?

I have studied many different dog breeds over the years. At Clemson University, my laboratory has developed genetic tests for nine Mendelian disorders of diverse breeds, from Jack Russell Terriers to Great Danes. We also have developed risk assessments for genetically complex diseases of Shetland sheepdog and collies, breeds that I am particularly fond of and have owned for most of my life.

You completed your undergraduate and graduate education at Texas A&M University. What is it about Clemson that brought you here?

In 2008, the laboratory in which I was trained at Texas A&M University relocated to Clemson University, which was starting a new focus area in animal genetics. I started my own laboratory at Clemson a year later. I am excited to now be part of the Center for Human Genetics and have many colleagues with common goals yet diverse areas of expertise.

Can you explain how canines serve as an appropriate model for studying other mammalian hereditary diseases?

Mammals are genetically very similar to one another. One characteristic of purebred dogs that makes them a great model system is their unique population structure. Each dog breed is an isolated population with long linkage disequilibrium and limited genetic diversity, genetic features that facilitate the mapping and identification of causal variants. A characteristic of dogs that sets them apart from rodent models is that they share our environment. Complex diseases often have an environmental trigger that dogs, as our household companions, are more likely to be exposed to than a laboratory mammal.

What is your advice for young genetics investigators?

I would say that the one thing I consistently regret is not knowing more! Be open-minded and learn everything you can from each opportunity that arises. Especially computation skills! Learn those as soon as you can.

Now for an easy question. Because you study canines, does that mean you are more of a dog person than a cat person?

I am definitely a dog person, but my two children prefer cats, so it must skip a generation!

Leigh Anne Clark is Associate Professor of Genetics and Biochemistry.

Viewpoint: The Legacy of Bonferroni

by Robert Anholt

Carlo Emilio Bonferroni, an Italian mathematician who worked on probability theory, could not have imagined that his remarkably simple insight would have a distinct impact on the practice of genomics, a science that did not yet exist during his lifetime. In simple terms, the Bonferroni correction solves the multiple testing problem, simply by dividing the nominal probability threshold by the number of tests. This simple correction became the gold standard for declaring genome-wide significance in genome-wide association studies. In fact, many journals would not publish results unless they would reach genome-wide statistical significance based on the Bonferroni correction for multiple tests.

But there is a problem. The Bonferroni correction assumes that the tests are independent, an assumption that is at odds with the realities of gene regulation. There is considerable evidence that the functional genome is highly interconnected and gene expression is characterized by dynamic coregulated ensembles of genes which can be modified during development, influenced by environmental effects and, when

Carlo Bonferroni (1892-1960)

analyzed in populations, subject to pervasive gene-gene interactions. This raises the question whether the Bonferroni correction should be used to correct for multiple tests when analyzing data from genome-wide association studies. Clearly, when applied in this context, associations that survive the Bonferroni test are established with high confidence. However, strict reliance on the Bonferroni correction results in the loss of huge data sets that fail to meet this excessively strict criterium but are in fact biologically relevant. As a further consequence, those DNA variants that meet genome-wide significance (often in genes of unknown function or in non-coding regions of the genome) are detached from their functional context, complicating interpretations of their biological significance.

So, how can we establish an appropriate significance threshold that takes into account the interdependence of the genome? Alternatives to the Bonferroni correction is establishing a permutation threshold, which is however

computationally intensive or applying a false discovery rate criterium. There are, however two relatively simple solutions to the problem of establishing an appropriate significance threshold. One is to base the cut-off for a significant P-value on deviation from linearity of quantile-quantile plots. Using this criterium studies on Drosophila found a threshold of $P < 10^{-5}$ to be appropriate for most studies. In contrast to people, however, flies enable associations derived from genome-wide association studies to be functionally validated through mutational analyses using the huge public resources available for this organism and the ease with which flies can be genetically manipulated. Such studies showed that about 70-80% of associations at $P < 10^{-5}$ could be functionally validated.

There is, however, a much simpler method for establishing a proper significance threshold in human studies, which will elicit considerable skepsis from statistical afficionados. Simply look at the Manhattan plots and draw a horizontal cut-off line where single points clearly emerge above the otherwise densely colored background. It seems simple and makes common sense. There is no statistical justification for this eyeball method, except to remember that establishing a statistical threshold is itself arbitrary and up to the judgment of the researcher. We must make a balance between Carlo Bonferroni's strict statistical formula to limit the identification of false positives and the "eyeball" approach which may identify false negatives. At the end of the day, causality of associations must be established through independent means. Until then, it is better to catch a wide net than to sacrifice potential transformative insights by discarding whole ensembles of genes.

Robert Anholt is the Provost Distinguished Professor of Genetics and Biochemistry and Director of Faculty Excellence in the College of Science at Clemson University.

A Warm Welcome to Zhana Duren

In August 2020, Dr. Zhana Duren joined the Clemson Center for Human Genetics as an Assistant Professor in the Department of Genetics and Biochemistry. Duren received his Ph.D. in Operational Research and Cybernetics from the Chinese Academy of Sciences in 2017 and completed postdoctoral training with Professor Wing Hung Wong's at Stanford University.

His research is computational modeling of genomics data and interpretation of genetic variants. Advances in genomics technology have generated, and continue to generate, "big data" by measuring DNA, RNA, proteins, and organismal function alongside clinical features including measures of disease activity, progression and related metadata. How to use advanced artificial intelligence (AI) and other computational methods to analyze these big data to facilitate precision medicine is currently of great interest. A person's genome typically contains millions of variants which represent the differences between this personal genome and the reference human genome. It is challenging to understand the mechanism of how these genetic variants contribute to disease because over 90% of trait-associated genetic variants are located in non-coding regions which do not encode protein-coding genes but may have regulatory functions. Duren's research focuses on a key question: How do non-coding genetic variants act through cellular context specific gene regulatory networks to influence disease?

To approach this challenge, Duren developed novel machine learning methods and bioinformatics tools including an inference method of mechanistic gene regulatory networks by integrating different types of genomics data, single-cell genomics data analysis, and identification of causal variants for complex diseases by using cellular context specific gene regulatory networks. Using these computational models, he discovered the key transcription factors in epidermal development and identification of the non-coding causal genetic variants of high-altitude hypoxia adaptation of Tibetans.

"Zhana will add considerable strength to our computational genomics group," says Center Director Dr. Trudy Mackay, "We are fortunate to have him as a member of our Center."

Seminars

The Center for Human Genetics presents a series of Distinguished Seminars in Human Genetics.

On Friday, **February 19** at 2:30 pm **Dr. Arpana Agrawal**, Professor of Psychiatry at Washington University at St. Louis and Co-PI of the Substance Use Disorders Working Group of the Psychiatric Genomics Consortium, will present a seminar titled "Recent discoveries in genomic studies of substance use and use disorders."

On Friday, **March 12** at 3:30pm, **Dr. Carole Ober**, Bluem-Riese Professor and Chair of the Department of Human Genetics at the University of Chicago, will present a seminar titled "Fine-mapping genetic risks in the HLA region for childhood onset and adult onset asthma".

On Friday, **April 2** at 2:30pm, **Dr. Teri Manolio**, Director of the Division of Genomic Medicine at the National Human Genome Research Institute, will present a seminar titled "Implementing genomics in clinical care: Genomic medicine programs of the NHGRI."

All seminars will be conducted via zoom. Zoom links will be announced one week prior to each seminar.

Publications

(affiliates of the Center for Human Genetics are in bold font)

Abrams SR, Hawks AL, Evans JM, Famula TR, Mahaffey M, Johnson GS, **Mason JM** and **Clark LA**. 2020. Variants in *FtsJ RNA 2′-O-Methyltransferase 3* and *Growth Hormone 1* are associated with small body size and a dental anomaly in dogs. *Proc. Natl. Acad. Sci. U. S. A.* **117:** 24929-24935. doi: 10.1073/pnas.2009500117.

Bell ZW, Lovell P, Mello CV, Yip P, **George JM** and Clayton DF. 2019. Urotensin-related gene transcripts mark developmental emergence of the male forebrain vocal control system in songbirds. *Sci. Rep.* **9:** 816. doi 10.1038/s41598-018-37057-w.

Cho H, Zitkovsky J, **Li X**, Lu M, Shah K, Sperger J, Tsilimigras MCB, and Kosorok MR. 2020. Comment: Diagnostics and kernel-based extensions for linear mixed effects models with endogenous covariates. *Stat. Sci.* **35**: 396-399. doi:10.1214/20-STS782.

George JM, Bell ZW, Condliffe D, Dohrer K, Abaurrea T, Spencer K, Leitao A, Gahr M, Hurd PJ and Clayton DF. 2020. Acute social isolation alters neurogenomic state in songbird forebrain. *Proc. Natl. Acad. Sci. U.S.A.* **117:** 23311-23316. doi 10.1073/pnas.1820841116.

Gettings JR, Self SCW, **McMahan CS**, Brown DA, Nordone SK and Yabsley MJ. 2020. Regional and local temporal trends of *Borrelia burgdorferi* and *Anaplasma* spp. seroprevalence in domestic dogs: Contiguous United States 2013-2019. *Front. Vet. Sci.* **7:** 561592. doi: 10.3389/fvets.2020.561592.

Hang Y, Aburidi M, Husain B, Hickman AR, Poehlman WL and **Feltus FA**. 2020. Exploration into biomarker potential of region-specific brain gene co-expression networks. *Sci. Rep.* **10:** 17089. doi: 10.1038/s41598-020-73611-1.

Hou P, Tebbs JM, Wang D, **McMahan CS** and Bilder CR. 2020. Array testing for multiplex assays. *Biostatistics* **21:** 417-431. doi: 10.1093/biostatistics/kxy058.

Huang W, Carbone MA, Lyman RF, **Anholt RRH** and **Mackay TFC**. 2020. Genotype by environment interaction for gene expression in *Drosophila melanogaster. Nat. Commun.* **11**: 5451. doi: 10.1038/s41467-020-19131-y.

Husain B, Hickman AR, Hang Y, Shealy BT, Sapra K and **Feltus FA.** 2020. NetExtractor: Extracting a cerebellar tissue gene regulatory network using differentially expressed high mutual information binary RNA profiles. *G3 (Bethesda)* **10:** 2953-2963. doi: 10.1534/g3.120.401067.

Joyner CN, **McMahan CS**, Tebbs JM and Bilder CR. From mixed effects modeling to spike and slab variable selection: A Bayesian regression model for group testing data. *Biometrics* **76:** 913-923. doi: 10.1111/biom.13176.

Li W, **Duren Z**, Jiang R and Wong WH. 2020. A method for scoring the cell type-specific impacts of noncoding variants in personal genomes. *Proc. Natl. Acad. Sci.* **35**: 21364-21372. doi: 10.1073/pnas.1922703117.

Morgante F, Huang W, Sørensen P, Maltecca C and **Mackay TFC**. 2020. Leveraging multiple layers of data to predict *Drosophila* complex traits. G3 *(Bethesda)* **10:** 4599-4613. doi: 10.1534/g3.120.401847.

Pfeiler E, **Nazario-Yepiz NO** and Markow TA. 2020. Additional species records and nomenclature updates of butterflies from a threatened coastal habitat in southern Sonora, Mexico. *J. Lepid. Soc.* **74**:197-200. doi: 10.18473/lepi.74i3.a8.

Spoor S, Wytko C, Soto B, Chen M, Almsaeed A, Condon B, Herndon N, Hough H, Staton M, Wegrzyn J, Main D, **Feltus FA** and Ficklin S. 2020. Tripal and Galaxy: supporting reproducible scientific workflows for community biological databases. *Database.* doi: 10.1093/database/baaa032.

Targonski C, Bender MR, Shealy BT, Husain B, Smith MC, Paseman B and **Feltus FA**. 2020. Cellular state transformations using deep learning for precision medicine applications. *Patterns* **1**: 6. doi: 10.1016/j.patter.2020.100087.

Tsai KL, Vernau KM, Winger K, Zwueste DM, Sturges BK, Knipe M, Williams DC, Anderson KJ, Evans JM, Guo LT, **Clark LA** and Shelton GD. 2020. Congenital myasthenic syndrome in Golden Retrievers is associated with a novel *COLQ* mutation. *J. Vet. Intern. Med.* **34**: 258-265.

Wang J and **Wang L**. 2020. Prediction and prioritization of autism-associated long non-coding RNAs using gene expression and sequence features. *BMC Bioinformatics* **21:** 505. doi: 10.1186/s12859-020-03843-5.

Xin J, Zhang H, He Y, **Duren Z**, Bai C, Chen L, Luo X, Yan DS, Zhang C, Zhu X, Yuan Q, Feng Z, Cui C, Qi X, Ouzhuluobu, Wong WH, Wang Y and Su B. 2020. Chromatin accessibility landscape and regulatory network of high-altitude hypoxia adaptation. *Nat. Commun.* 11: 4928. doi: 10.1038/s41467-020-18638-8.

Yanagawa A, Huang W, Yamamoto A, Wada-Katsumata A, Schal C and **Mackay TFC**. 2020. Genetic basis of natural variation in spontaneous grooming in *Drosophila melanogaster. G3 (Bethesda)* **10:** 3453-3460. doi: 10.1534/g3.120.401360.

Yip PK, Schmitzberger M, Al-Hasan M, **George J**, Tripoliti E, Michael-Titus AT, Clayton D, and Priestley JV. 2020. Serotonin expression in the song circuitry of adult male zebra finches. Neuroscience 444:170-182.
doi: 10.1016/j.neuroscience.2020.06.018.

Visit our website at https://scienceweb.clemson.edu/chg/

Clemson University

Assistant Professor in Human Genetics

Clemson University invites applications for a tenure-track Assistant Professor as part of a Cluster Hire at the Center for Human Genetics, with an expected start date of August 2021. Clemson University offers competitive salaries, benefits and start-up funds.

The successful applicant will have an accomplished research record at the forefront of human genetics/genomics. Applicants whose research combines experimental laboratory work and computational approaches are especially desirable.

The Center for Human Genetics (https://scienceweb.clemson.edu/chg/) is housed in Self Regional Hall, a new 17,000-square-foot building located in Greenwood, South Carolina on Greenwood Genetic Center Partnership Campus. The Center for Human Genetics provides a vibrant interactive research environment with state-of-the-art genomic and computational resources, and is ideally configured for collaborative research with excellent bioinformatics facilities and state-of-the-art molecular laboratories. The Center for Human Genetics and the Greenwood Genetic Center are well-equipped for genomics, proteomics and metabolomics research, including NovaSeq 6000 and PacBio Sequel II sequencers. Successful applicants will be part of a collaborative and interdisciplinary environment that includes the research, diagnostic and clinical geneticists at the Greenwood Genetic Center, the genetics, genomics, statistics and bioinformatics faculty at Clemson University, the USC School of Medicine in Greenville and the Greenville Health System. The home department will be determined by the fit of the applicant's research interests with the mission of one of the departments in the College of Science (www.clemson.edu/science), including the Department of Genetics and Biochemistry (www.clemson.edu/science/departments/genetics-biochemistry/index.html), the Department of Mathematical Sciences (www.clemson.edu/science/department s/mathematical-sciences/index.html) and the Department of Biological Sciences (www.clemson.edu/science/departments/biosci/index.html).

The University and Center for Human Genetics are committed to building a diverse and inclusive community of faculty scholars dedicated to working and teaching in a multi-cultural environment (http://www.clemson.edu/inclusion/). We encourage applications from women, minorities and individuals with a commitment to mentoring colleagues and students from demographic groups underrepresented in the sciences. We are also supportive of the needs of dual-career couples.

Successful candidates must hold a doctoral degree and have postdoctoral experience. Competitive candidates will demonstrate an ability to develop a vigorous and independent, externally funded and nationally recognized research program; demonstrate teaching excellence and a commitment to diversity inclusion; and participate in relevant undergraduate and graduate education programs.

Applicants should submit the following items via Interfolio at http://apply.interfolio.com/82505: (1) cover letter; (2) Curriculum Vitae; (3) statement of research interests including future plans; (4) statement of teaching interests and experience; (5) statement describing past experience and/or future plans to promote diversity and inclusion; and (6) up to three reprints in one PDF. Applicants should also arrange, through Interfolio, the submission of three confidential letters of recommendation on their behalf. Review of applications will begin on February 15, 2021 and continue until the position is filled.

Inquiries should be directed to Dr. Trudy Mackay (tmackay@clemson.edu)

Clemson University is an AA/EEO employer and does not discriminate against any person or group on the basis of age, color, disability, gender, pregnancy, national origin, race, religion, sexual orientation, veteran status or genetic information. Clemson University is building a culturally diverse faculty and staff committed to working in a multicultural environment.