CLEMSON UNIVERSITY
CENTER FOR HUMAN GENETICS

# The Transcript

Newsletter of the Center for Human Genetics at Clemson University
**Fall, 2020**

## From the Director

Welcome to the first issue of *The Transcript*, the semiannual newsletter of the Center for Human Genetics at Clemson University!

The Center for Human Genetics (CHG) formally opened its doors on July 1, 2018. We are housed in Self Regional Hall on the Innovation Partnership campus in Greenwood adjacent to the Greenwood Genetic Center, a facility dedicated to the diagnosis, treatment and research of pediatric genetic disorders.

The goals of the CHG are to leverage comprehensive systems genetic approaches and comparative genomics to elucidate fundamental principles of the genetic underpinnings of human complex traits, including disease risk, and to promote precision medicine by developing advanced mathematical models to predict disease risk and assess therapeutic benefits based on genetic and environmental factors.

We also seek to develop local, regional, national and international collaborations to advance human genetics and to generate genetic repositories and data bases as resources for the scientific community. Another aspect of our mission is to educate the next generation of human geneticists by providing educational opportunities for high school students and their teachers, for undergraduate and graduate students, for postdoctoral fellows and visiting scientists, and to promote public understanding of human genetics through community outreach.

Common and rare genetic diseases affect a large fraction of the world's population. Many advances in genomic medicine have occurred in the past 20 years. However, current human genetic research faces several challenges and opportunities. Only about 2% of the human genome codes for proteins. Genome-wide association studies for a common disease or trait identifies statistically significant variants, the majority of which are in non-coding regions of the genome (introns, upstream or downstream of the nearest gene, or intergenic) and thus have presumably regulatory effects on gene expression. However, these variants are not necessarily the causal variants because of the haplotype block structure of the human genome. A major challenge for the future is tackling the issue of causality and elucidating the role of non-coding elements of the genome.

In addition, we need to develop models to accurately predict individual disease risk or trait phenotype based on genetic information.

**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.**

Elucidating the mechanisms by which genetic variants affect function thus requires expertise from multiple disciplines ranging from quantitative, population, molecular, cellular and developmental genetics; statistics, bioinformatics and computational biology; to functional genomics. It also requires the ability to directly test hypotheses derived from human studies by functional analyses in cell and tissue culture and animal models and to perform high dimensional analyses of pleiotropic effects on molecular networks in these model systems that are not possible in humans.

To support such endeavors, we established an in-house high-throughput genomics facility to enable cost-effective whole genome sequencing on the Illumina platform using a Novaseq, long-read sequencing supported by PacBio Sequel II and Oxford Nanopore Systems and single-cell DNA and RNA sequencing using 10x Genomics microfluidics. In addition to South Carolina's Palmetto supercomputer, we also set up an in-house server for analysis of large data sets.

I am delighted that during the last two years we were able to recruit a cadre of stellar young scientists who will advance the frontiers of human genetics in a collaborative environment. Dr. Andrei Alexandrov, who is featured in this issue of The Transcript, joined us from Yale University and uses innovative technologies to study the role of non-coding RNAs in cancer. Dr. Lela Lackey came to us from the University of North Carolina at Chapel Hill and studies the relation between RNA structure and gene regulation in the context of human health and disease. Dr. Zhana Duren from Stanford University uses advanced computational technologies to disentangle the complex interrelationships between different 'omics' levels. Dr. Fabio Morgante develops methods for genomic prediction for cardiovascular disorders using large data sets of human populations. Ongoing studies in the Mackay-Anholt laboratory focus on alcohol use disorder and addiction to psychostimulants, as well as comparative genomics studies on rare human diseases in collaboration with research faculty and clinicians at the Greenwood Genetic Center.

In the last two years we have made great strides, but this is only the beginning. We are looking forward to developing the CHG as the primary Center for Human Genetics in the South-Eastern United States.

## A Warm Welcome to David Clayton and Julia George

The Center for Human Genetics extends a warm welcome to Dr. David Clayton, who will lead Clemson University's Department of Genetics and Biochemistry as its new Department Chair. David Clayton and his spouse Julia George have returned to the United States from the United Kingdom, where Clayton served as Chair of the Department of Biological Sciences at Queen Mary University in London.

Clayton is well-known for his contributions to the neurogenetics of birdsong. Songbirds develop their song during a critical developmental period by listening and emulating their parents' song. Birdsong serves as a model for the development of human language and vocal communication. After receiving his Ph.D. with Professor James E. Darnell Jr. at the Rockefeller University in New York, Clayton launched the first investigations of genes expressed in the brain of songbirds. He worked initially with Professor Fernando Nottebohm, also at the Rockefeller University, and later as a faculty member at the University of Illinois, where he worked for 22 years.

Clayton and his students discovered the Immediate Early Gene response in vocal communication and described the first patterns of developmental and sex-specific gene regulation in the vocal learning circuit. At the University of Illinois, Clayton organized a broad set of international collaborations which led to the complete sequencing of the songbird (zebra finch) genome at a time when still few whole genomes had been sequenced.

Julia George was an early pioneer in the study of the alpha-synuclein protein, which is associated with Parkinson's disease and which she identified in the developing songbird vocal control circuit. Since moving to London in 2012, her primary focus has been on the epigenetics of neural responses to social and environmental signals in songbird models. Dr. George is joining the Department of Biological Sciences at Clemson University.

At Clemson, Clayton and George will continue their studies on the role of the genome in brain plasticity – the ability of the brain to respond and adapt to signals involved in brain circuit development, sex differences, learning and memory, and social and environmental conditions. Measurements of gene regulation and epigenetic mechanisms are central to their approach.

"We are delighted to have David and Julia join the Clemson family," says Center Director Dr. Trudy Mackay, "They will contribute another valuable model system for comparative genomic studies to the Center for Human Genetics."

## Andrei Alexandrov Brings Fireworks to Clemson

Dr. Andrei Alexandrov is an innovator. After obtaining his undergraduate diploma from Lomonosov State University in Moscow, Alexandrov left his native Russia to further his graduate education at the University of Rochester School of Medicine, where he received his Ph.D. in biophysics. His expertise in biophysics placed him in an ideal position to develop new technologies in RNA biology during a postdoctoral period in Joan Steitz' laboratory at Yale University. There he developed an approach that enables ultra-high throughput forward genetic discovery of components that regulate human RNA biogenesis. Nuclear-localized vertebrate-specific long non-coding RNAs play critical roles in an impressive variety of human diseases, including numerous cancers, developmental and viral diseases. Alexandrov developed a new technology to study the functions of these long non-coding RNAs in human diseases. Since this technology makes sophisticated use of fluorescent reporter signals, he named it "Fireworks."

Using Fireworks, Alexandrov identified human nuclear components that are required for maturation of MALAT1 (metastasis associated lung adenocarcinoma transcript 1), a long non-coding RNA associated with various cancers. He also discovered the Rapid tRNA Decay pathway and a long elusive molecular link between splicing, deposition of the exon junction complex, and mRNA surveillance via nonsense-mediated decay.

In 2019, Alexandrov brought Fireworks to the Clemson Center for Human Genetics, where he is an Assistant Professor in the Department of Genetics and Biochemistry, and immediately developed a dynamic research laboratory. "My long-term research goals are to alleviate disease severity by modulating regulation of long non-coding RNA pathways, and to identify relationships between disease risk and regulation of non-coding elements in the human genome," says Alexandrov.

The technology developed by Alexandrov allows for the first time interrogation of pathways of biogenesis, regulation, and surveillance of human nuclear long non-coding RNAs, like MALAT1, which has long been considered impossible because of their strictly nuclear localization.

"We were fortunate to recruit Dr. Alexandrov," says Center Director Dr. Trudy Mackay, "Andrei is a brilliant scientist, who thinks out-of-the-box and uses innovative interdisciplinary approaches to explore critical problems at the forefront of human genetics."

## Flies on Cocaine

Europeans first discovered cocaine in the 16th century when Spanish conquistadors observed that indigenous people in Peru chewed on coca leaves and experienced its stimulating effects. For several centuries, cocaine was hailed as a beneficial therapeutic in part due to its anesthetic properties. Cocaine was first isolated from coca leaves in 1855 by the German chemist Friedrich Gaedcke. Sigmund Freud, the father of psycho-analysis, was an avid advocate for the use of cocaine and noted that "You perceive an increase of self-control and possess more vitality and capacity for work. In other words, you are simply normal, and it is soon hard to believe you are under the influence of any drug. Long intensive physical work is performed without any fatigue. This result is enjoyed without any of the unpleasant after-effects that follow exhilaration brought about by alcoholic beverages. No craving for the further use of cocaine appears after the first, or even after repeated taking of the drug."

Today Freud's enthusiastic endorsement of cocaine has vanished. Instead, it is now recognized that cocaine use presents a significant socioeconomic health problem. The Substance Abuse and Mental Health Services Administration reported that in the United States in 2018 an estimated 5.5 million people aged 12 or older were past users of cocaine, amounting to about 2.0% of the US population. The Centers for Disease Control reported 13,942 cocaine-related deaths in 2017 and a recent survey of 8th, 10th and 12th graders found that 2.2% of 12th graders had used cocaine. Although cocaine use leads to arousal and euphoria, there are serious side effects, including accelerated heart rate, mood swings, difficulty sleeping, loss of appetite and cognitive distortions. In contrast to Dr. Freud's assessment, addiction which results in escalated consumption of cocaine, can result in psychosis, cardiovascular disease and stroke.

Much has been learned about the neurological effects of cocaine but information about the underlying genetics that render individuals susceptible to cocaine use and addiction remains incomplete. Members of the Mackay and Anholt laboratory are using the fruit fly, Drosophila melanogaster, as an advantageous model system for systems genetic analyses of cocaine consumption. Why flies, one might wonder?

Flies provide a powerful model system for comparative genomics studies, because they can be reared rapidly in large numbers at low cost in defined genetic backgrounds and under controlled environmental conditions. Further-more, about 75% of disease-causing genes in humans have fly counterparts. Cocaine exerts its effects in people by blocking reuptake of the neurotransmitter dopamine in the central nervous system's reward circuit, thereby prolonging neural stimulation. The crystal structure of the Drosophila dopamine transporter has been obtained and its binding site can accommodate cocaine. Exposure of flies to cocaine elicits motor responses that resemble behaviors observed in rodents, and flies develop sensitization to repeated intermittent exposure to cocaine. Thus, fundamental neural mechanisms affected by exposure to psychostimulants are conserved across phyla, from flies to humans.

Graduate student Brandon Baker and postdoctoral fellow Chad Highfill in the Mackay-Anholt laboratory explored variation in susceptibility to the effects of cocaine among lines of the Drosophila melanogaster Genetic Reference Panel, a collection of fruit fly lines with fully sequenced genomes derived from a natural population. They correlated naturally occurring DNA variants with variation in consumption and development of preference for cocaine. After identifying candidate genes that harbor associated polymorphisms, they used RNA interference to inhibit expression of genes in specific neurons and glia. These studies implicated dopaminergic neurons and the mushroom bodies, central brain structures associated with experience-dependent modification of behavior, with consumption and development of preference for cocaine. In addition, reduction in expression of candidate genes in glia also affected cocaine consumption, suggesting that widespread brain regions contribute to cocaine-associated behaviors.

After Chad Highfill took a position as System's Biologist at Takeda Pharmaceuticals in Yokohama, graduate students Sneha Mokashi and Spencer Hatfield joined forces with Brandon Baker along with staff bioinformatician, Dr. Vijay Shankar, to explore the effects of acute cocaine exposure on gene expression at single cell resolution throughout the brain. These studies will provide insights in the extent to which cocaine exerts effects throughout the brain, identify targets of epigenetic modifications that can lay a blueprint for enhanced preference, a proxy for addiction, and identify potential targets for therapeutic intervention in humans at risk for cocaine addiction.

# A Conversation with Alex Feltus

***Could you briefly describe your research program?***

Our group uses software engineering and computational biology techniques to make useful molecular discoveries in human and plant biological systems. We also engineer elastic advanced compute systems and technologies to run robust genomics workflows to enable small labs to perform innovative petascale compu-tational biology. The lab also actively engages in traditional PhD training and the development of a scalable asynchronous training platform for data-intensive computing including but not limited to computational biology.

My lifetime research goal is to reveal the genomic mechanisms underlying phenotype expression. A core aspect of this approach to identify biomarkers that group interesting biological states (e.g. normal kidney verses renal tumor somatic mutation and/or transcriptome profiles). Given that most traits are under control by complex cellular control systems, we always seek to identify sets of functionally interacting genes (biomarker systems) that discriminate between biological states. My group focuses on the transcriptome layer (RNA) of gene expression but we are always seeking methods to integrate data from other genome information orbitals.

**Dr. F. Alex Feltus received a B.Sc. in Biochemistry from Auburn University in 1992, served two years in the Peace Corps in the Fiji Islands, and then completed advanced training in biomedical sciences at Vanderbilt University and Emory University. He has performed research in artificial intelligence, bioinformatics, cyberinfrastructure, high-performance computing, network biology, tumor biology, agrigenomics, genome assembly, systems genetics, paleogenomics, and bioenergy feedstock genetics. Feltus is Professor in Clemson University's Department of Genetics and Biochemistry and a member of the Center for Human Genetics and Biomedical Data Science & Informatics program. He is also co-founder of Praxis AI, LLC.**

***You have a diverse research program that combines software engineering, computational biology and molecular genetics to explore problems in human and plant systems. What are the challenges you face in recruiting and integrating a diverse group of collaborators that contribute to your research program?***

I am not sure how I got here, but my lab is truly interdisciplinary. There are geneticists, computer scientists, computer engineers, and bioengineers. We are all linked by a need and/or love of data intensive computing required to discover new knowledge about the life on Earth. I have not had any issues recruiting students due to my involvement in formal undergraduate and graduate training programs in the Genetics & Biochemistry department, decade long collaboration and co-training relationship with Dr. Melissa Smith's group in the Electrical and Computer Engineering department, membership in the Biomedical Data Science and Informatics (BDSI) program, and of course the Clemson Center for Human Genetics.

***Are you concerned about science literacy in the general population? How would you effectively communicate the importance and impact of your work to the general public?***

I am not sure what "science literacy" means. If it means a realistic understanding of the natural world as opposed to an opinion of the world should be, then I think more people would be excited about "science" if they focus on the beauty of our Universe as revealed through discovery. Regarding communication to the general public, I always try to put my work into perspective and focus on the impact to food security and how it might help heal you when you are sick.

***What do you consider the most pressing unresolved issue in Human Genetics?***

The most important tool we are missing in genetics are realistic quantitative genetic models that allow for perturbations of the genotype to predict the impact on phenotype. Simulation approaches will reduce the time from discovery to impact and allow folks to tweak biological systems thus creating a new era of true genetic engineering.

***After completing your B.Sc. degree from Auburn University you spent two years in the Peace Corps. How did this experience influence your subsequent academic career?***

There were two big takeaways from my time in the Fiji Islands. First, I learnt that there are geniuses everywhere and we need to provide access and training to reduce the barriers to learning. The Internet didn't exist then, so now this is possible. The second thing I continued to develop was a love for diverse cultures including my American culture. Putting diverse opinions together is a lot of fun and productive!

***What is your advice for young investigators?***

I remember when I started my PhD training in 1995, I met a couple of professors who thought molecular biology was a fad. I don't remember their names. I believe that computational biology skills are on the same order of importance as molecular biology. My advice is to learn how run gels AND use the command line in Linux, program in Python, and be able to run data-intensive containerized workflows in the Cloud. Also, learn basic data science skills and embrace the complexity of biological systems using computers. Genomics datasets are only going to get bigger and you must learn how to hurdle computational problems in order to make discoveries just like you do when you are performing wet bench experiments. You have to be able to run gels and computational jobs if you want people to remember your name.

# Fruit Flies and the Joy of Teaching

*by Rebecca MacPherson*

**Originally from Anchorage, Alaska, Rebecca MacPherson graduated from Clemson University in 2018 with a B.S. in genetics and minors in microbiology and psychology. She completed departmental and general honors from the Clemson University Honors College and worked on a variety of research projects during her undergraduate years, including canine genetics, eukaryotic pathogens, synthetic biology, and cancer biology. She started her Ph.D. in genetics in the fall of 2018 at the Clemson Center for Human Genetics under the mentorship of renowned geneticist Trudy Mackay.**

I knew from the look on their faces that they didn't believe me. "Fruit flies?" they asked incredulously. I had just told my grandfather, aunts, uncles and cousins that I would be spending the next five years of my life – the prime of my 20s, the pinnacle of my academic career thus far – working on the genetics of a pest that is brought home on bananas from the grocery store. I took their disbelief as a challenge; it was now my mission to tell them why a Ph.D. on fruit fly genetics was relevant to the study of human health.

Fruit flies, or *Drosophila melanogaster*, are incredibly important to biological research. Fruit flies share with humans about 75% of all genes that cause human disease. Additionally, scientists can use fruit flies to make discoveries that are ethically or financially impossible to make in humans. After discussing these topics with my family, including the six Nobel prizes awarded to scientists who had worked with fruit flies, my relatives realized that I was actually pursuing something impactful.

Our conversations about genetics reminded me of something important – I love to teach. Given my passion for outreach and teaching, I knew I needed to find a way to stay involved with genetics education while making it through graduate school.

A few months later, I was standing at the head of a classroom at the Clemson Center for Human Genetics. The room was filled with 7$^{th}$- through 12$^{th}$-grade students and their parents from a local homeschool group. A parent had recently reached out, wondering if I could talk to the students about genetics, and I had leapt at the opportunity. Initially I was worried about connecting with these students because I was short on time and had never taught in a formal setting. However, these fears quickly faded when I saw a transformation in one of the middle-schoolers that I'll never forget. She started out quiet at the beginning of the lesson, but as she learned more her eyes opened wide. The wheels in her head were turning rapidly, and she began to ask question after question until we ran out of time. She truly "got it," and nothing has encouraged me more.

As a graduate student at the Clemson Center for Human Genetics, I have also been able to foster research and outreach collaborations with the Greenwood Genetic Center. The GGC's education department offers excellent genetics education activities directly to 7$^{th}$- through 12$^{th}$-grade students across South Carolina. I have traveled with their mobile labs, partnered with GGC education instructors to create fruit fly activities for students, and worked with local science teachers to help bring genetics into their classrooms. I am so grateful to have the opportunity to teach genetics to students across multiple backgrounds, ages, and geographic locations.

Each summer, the GGC education department offers the Junior Genetics Scholars camp to high school students. Participants are introduced to GGC faculty and staff and are exposed to real-world laboratory activities, including some that involve fruit flies. If there's one thing I've learned in my time teaching about fruit flies, it is that I never know how a student will react to seeing a fruit fly in a laboratory environment. Perhaps the students' reactions stem from the contradiction between a sterile space and a "gross" insect, or perhaps the idea of studying an insect is as ludicrous for the students as it was initially for my family. Either way, many students are overcome by curiosity and excitement, and they often ask me humorous questions. Due to its frequency, a question about what I name my fruit flies (no, I don't name my flies) no longer startles me. However, I'm still surprised by questions about fruit fly dreams, the biggest fruit fly ever, and if a fruit fly can swim. Students' excitement about fruit flies also manifests into rapid shouting, where the students provide me updates about the movement of individual flies like sports commentators describing the action on the field. And although some students remain quiet throughout the activity, it is rare that I do not see a smile on their faces as they leave the classroom.

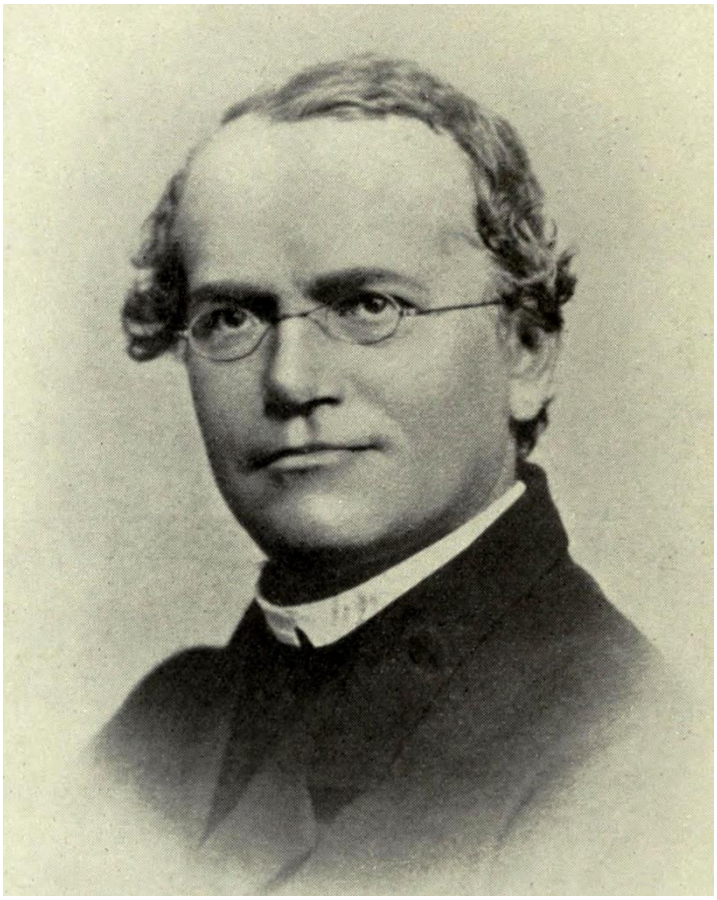
Teaching opportunities at the Center for Human Genetics have been a blessing, forcing me to go out and create opportunities for myself and work in the areas of education I most enjoy. I have successfully been able to teach across a range of ages and educational environments, something I would not have obtained as a teaching assistant in an undergraduate classroom. Not only have I found in-person education opportunities, but through volunteering at local science fairs, working with outreach programs like Skype a Scientist, and attending formal science communication workshops such as ComSciCon, I have further broadened my scientific outreach and education experiences.

Looking back over the first half of my graduate career, I've realized that some things never change. For example, I will always have to justify why I'm spending five years of my life studying the genetics of a pesky little insect. But hey, thanks to my outreach and education experiences, I'd like to think I'm at least a little better at justifying it than I used to be.

## Viewpoint: Are Mendelian Diseases Truly Monogenic?

*by Robert Anholt*

Gregor Mendel was both brilliant and lucky. The traits he studied in his peas were well-defined categorical traits, each of which could be accounted for by segregating alleles from a single gene. Mendel might not have been so successful if he had studied traits with a continuous quantitative distribution in the population, as is typical for complex traits that result from many interacting genes. Whereas most common diseases result from a complex genetic architecture, Mendel's rules of heredity can be applied to rare disorders. Such Mendelian diseases result from a genetic disruption with a highly penetrant effect that segregates in pedigrees as a monogenic polymorphism. Although the incidence of each disease is generally less than 1 in 200,000, a total of about 7,000 different rare diseases have been described. Rare disorders collectively affect up to 10% of live births, a significant pediatric patient population.



*Gregor Mendel*

But are Mendelian diseases truly monogenic? This is true for some disorders, like cystic fibrosis, where the same disrupted gene segregating in different families causes a disease phenotype with high penetrance. However, there are many examples of diseases where a different single pathogenic mutation in independent pedigrees results in similar disease manifestation. Often, these different defective genes sequestered in different families are associated with a common biological process. For example, Bardet-Biedl syndrome is a ciliopathy, which impacts multiple body systems. Several different mutations have been identified in different families with affected children, but they affect different gene products essential for ciliogenesis. Similarly, occurrence of Coffin-Siris syndrome can result from different genes harboring pathogenic mutations in different families, all associated with chromatin remodeling. While such diseases can be considered Mendelian within each pedigree, should they not be considered polygenic across the patient population?

Another observation to consider is variable penetrance of the same disease alleles in different pedigrees. Differences in penetrance could result from interactions with modifiers in the genetic background or interactions between the genome and environment. Indeed, such interactions are hallmarks of complex traits. Assessing effects of genetic modifiers or effects of environmental exposure on the manifestation of rare diseases is virtually impossible in human populations but could plausibly be addressed through comparative genomic approaches in model organisms.

Perhaps the widely held view that rare diseases are relatively simple Mendelian disorders should be revisited. Instead, the collection of those approximately 7,000 rare disorders might be viewed as a kaleidoscope that provides a continuum of complexity ranging from true monogenic disorders to disorders arising from various degrees of genetic complexity

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

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

# Publications

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


Anderson KL, Kearns R, **Lyman RL**, and Correa MT. 2019. Staphylococci in dairy goats and human milkers, and the relationship with herd management practices. *Small Ruminant Research* **171**: 13-22 doi: 10.1016/j.smallrumres.2018.11.021.

Ang CE, Ma Q, Wapinski OL, Fan S, Flynn RA, Lee QY, Coe B, Onoguchi M, Olmos VH, Do BT, Dukes-Rimsky L, Xu J, Tanabe K, **Wang L**, Elling U, Penninger JM, Zhao Y, Qu K, Eichler EE, Srivastava A, Wernig M and Chang HY. 2019. The novel lncRNA *lnc-NR2F1* is pro-neurogenic and mutated in human neurodevelopmental disorders. *Elife* **8**: 41770. doi: 10.7554/eLife.41770.

**Anholt RRH**. 2020. Chemosensation and evolution of Drosophila host plant selection. *iScience* **23:** 100799. doi: 10.1016/j.isci.2019.100799.

**Anholt RRH.** 2020. Evolution of epistatic networks and the genetic basis of innate behaviors. *Trends Genet* **36**: 24-29. doi: 10.1016/j.tig.2019.10.005.

**Anholt RRH**, Wolfner MF, O'Grady P and Harbison ST. 2020. Evolution of reproductive behavior. *Genetics* **214**: 49-73. doi: 10.1534/genetics.119.302263.

Bilder C, Tebbs J and **McMahan C**. 2019. Informative group testing for multiplex assays. *Biometrics* **75:** 278-288. doi: 10.1111/biom.12988.

Casanova EL and **Konkel MK**. 2020. The developmental gene hypothesis for punctuated equilibrium: Combined roles of developmental regulatory genes and transposable elements. *Bioessays* **42**: e1900173. doi: 10.1002/bies.201900173.

Casanova EL, Switala AE, Dandamudi S, Hickman AR, Vandenbrink J, Sharp JL, **Feltus FA** and Casanova MF. 2019. Autism risk genes are evolutionarily ancient and maintain a unique feature landscape that echoes their function. *Autism Res* **12**: 860-869. doi: 10.1002/aur.2112.

**Duren Z**, Chen X, Xin J, Wang Y and Wong WH. 2020. Time course regulatory analysis based on paired expression and chromatin accessibility data. *Genome Res* **30**: 622-634. doi: 10.1101/gr.257063.119.

Everett LJ, Huang W, Zhou S, Carbone MA, **Lyman RF**, Arya GH, Geisz MS, Ma J, **Morgante F**, St Armour G, Turlapati L, **Anholt RRH** and **Mackay TFC**. 2020. Gene expression networks in the *Drosophila* Genetic Reference Panel. *Genome Res* **30**: 485-496. doi: 10.1101/gr.257592.119.

Gabrawy MM, Campbell S, Carbone MA, Morozova TV, Arya GH, Turlapati LB, Walston JD, Starz-Gaiano M, Everett L, **Mackay TFC**, Leips J and Abadir PM. 2019. Lisinopril preserves physical resilience and extends life span in a genotype-specific manner in *Drosophila melanogaster*. *J Gerontol A Biol Sci Med Sci* **74**: 1844-1852. doi: 10.1093/gerona/glz152.

Gettings JR, Self SCW, **McMahan CS**, Brown DA, Nordone SK and Yabsley MJ. 2020. Local and regional temporal trends (2013-2019) of canine *Ehrlichia spp*. seroprevalence in the USA. *Parasites and Vectors* **13:** 153. doi: 10.1186/s13071-020-04022-4.

Greene MA, Britt JL, Powell RR, **Feltus FA**, Bridges WC, Bruce T, Klotz JL, Miller MF and Duckett SK. 2019. Ergot alkaloid exposure during gestation alters: 3. Fetal growth, muscle fiber development and miRNA transcriptome. *J Anim Sci* **97**: 3153-3168. doi:10.1093/jas/skz153.

Gregory K, Wang D and **McMahan C**. 2019. Adaptive elastic net regression with group testing data. *Biometrics* **75:** 13-23. doi: 10.1111/biom.12973.

Gudenas BL, Wang J, Kuang SZ, Wei AQ, Cogill SB and **Wang LJ**. 2019. Genomic data mining for functional annotation of human long noncoding RNAs. *J Zhejiang Univ Sci* **20**: 476-487. doi:10.1631/jzus.B1900162.

Harbison ST, Kumar S, Huang W, McCoy LJ, Smith KR and **Mackay TFC**. 2019. Genome wide association study of circadian behavior in *Drosophila melanogaster*. *Behav Genet* **49**: 60-82. doi: 10.1007/s10519-018-9932-0.

Highfill CA, **Baker BM**, Stevens SD, **Anholt RRH** and **Mackay TFC**. 2019. Genetics of cocaine and methamphetamine consumption and preference in *Drosophila melanogaster*. *PLoS Genet* **15**: e1007834. doi: 10.1371/journal.pgen.1007834.

Hou P, Tebbs J, Wang D, **McMahan C** and Bilder C. 2020. Array testing with multiplex assays. *Biostatistics* **21**: 417-431. doi: 10.1093/biostatistics/kxy058.

Huang W, Campbell T, Carbone MA, Jones WE, Unselt D, **Anholt RRH** and **Mackay TFC.** 2020. Context-dependent genetic architecture of *Drosophila* lifespan. *PLoS Biol* **18**: e3000645. doi: 10.1371/journal.pbio.3000645.

Joyner C, **McMahan C,** Baurley J and Pardamean B. 2020. A two-phase Bayesian methodology for the analysis of binary phenotypes in genome-wide association studies. *Biom J* **62**: 191-201. doi: 10.1002/bimj.201900050

Krombeen SK, **Shankar V**, Noorai RE, Saski CA, Sharp JL, Wilson ME and Wilmoth TA. 2019. The identification of differentially expressed genes between extremes of placental efficiency in maternal line gilts on day 95 of gestation. *BMC Genomics* **20**: 254. doi: 10.1186/s12864-019-5626-0.

Kuang S and **Wang L**. 2020. Identification and analysis of consensus RNA motifs binding to the genome regulator CTCF. *NAR Genomics and Bioinformatics* **2**: lqaa031. doi: 10.1093/nargab/lqaa031.

Lingjie L, Wang Y, Torkelson JL, Shankar G, Pattison JM, Zhen HH, Fang F, **Duren Z**, Xin J, Gaddam S, Melo SP, Piekos SN, Li J, Liaw EJ, Chen L, Li R, Wernig M, Wong WH, Chang HY and Oro AE. 2019. TFAP2C-and p63-dependent networks sequentially rearrange chromatin landscapes to drive human epidermal lineage commitment. *Cell Stem Cell* **24**: 271-284. doi: 10.1016/j.stem.2018.12.012.

Mateus RP, **Nazario-Yepiz NO**, Ibarra-Laclette E, Ramirez Loustalot-Laclette M, Markow TA. 2019. Developmental and transcriptomal responses to seasonal dietary shifts in the cactophilic *Drosophila mojavensis* of North America. *J Hered* **110**: 58-67. doi: 10.1093/jhered/esy056.

Matute DR, Comeault AA, Earley E, Serrato-Capuchina A, Peede D, Monroy-Eklund A, Huang W, Jones CD**, Mackay TFC** and Coyne JA. 2020. Rapid and predictable evolution of admixed populations between two *Drosophila* species pairs. *Genetics* **214**: 211-230. doi: 10.1534/genetics.119.302685.

Mills N, Bensman EM, Poehlman WL, Ligon WB 3rd and **Feltus FA**. 2019. Moving just enough deep sequencing data to get the job done. *Bioinform Biol Insights* **13**: 1177932219856359. doi: 10.1177/1177932219856359.

**Morgante F**, Huang W, Sørensen P, Maltecca C and **Mackay TFC**. 2019. Leveraging multiple layers of data to predict *Drosophila* complex traits. *bioRxiv* 824896. doi: 10.1101/824896.

Murphy SC, Recio A, de la Fuente C, Guo LT, Shelton GD and **Clark LA**. 2019. A glycine transporter SLC6A5 frameshift mutation causes startle disease in Spanish greyhounds. *Hum Genet* **138**: 509-513. doi: 10.1007/s00439-019-01986-x.

Nayduch D, **Shankar V**, Mills MK, Robl T, Drolet BS, Ruder MG, Scully ED and Saski CA. 2019. Transcriptome response of female *Culicoides sonorensis* biting midges (Diptera: Ceratopogonidae) to early infection with Epizootic Hemorrhagic Disease Virus (EHDV-2). *Viruses* **11**: 473. doi: 10.3390/v11050473.

Noorai RE, **Shankar V**, Freese NH, Gregorski CM and Chapman SC. 2019. Discovery of genomic variants by whole-genome resequencing of the North American Araucana chicken. *PLoS One* **14**: e0225834. doi: 10.1371/journal.pone.0225834.

Parker GA, Kohn N, Spirina A, McMillen A, Huang W and **Mackay TFC**. 2020. Genetic basis of increased lifespan and postponed senescence in *Drosophila melanogaster*. *G3 (Bethesda)* **10**: 1087-1098. doi: 10.1534/g3.120.401041.

Paterson AH, Kong W, Johnston RM, Nabukalu P, Wu G, Poehlman WL, Goff VH, Isaacs K, Lee TH, Guo H, Zhang D, Sezen UU, Kennedy M, Bauer D, **Feltus FA**, Weltzien E, Rattunde HF, Barney JN, Barry K, Cox TS and Scanlon MJ. 2020. The evolution of an invasive plant, *Sorghum halepense* L. ('Johnsongrass'). *Front Genet* **11**: 317. doi: 10.3389/fgene.2020.00317.

Pfeiler E and **Nazario-Yepiz NO**. 2020. DNA-based taxonomy and potential suppression of long-established names: the case of *Telegonus fulgerator* (Lepidoptera: Hesperiidae). *Syst Biodivers* **18**: 338–346. doi: 10.1080/14772000.2020.1758825.

Pfeiler E, **Nazario-Yepiz NO**, Hernández-Cervantes PL and Markow TA. 2020. Mitochondrial DNA barcodes provide insight into the phylogeography and subspecies controversy in the widespread neotropical White Peacock butterfly, *Anartia jatrophae* (Nymphalidae: Nymphalinae). *Biol J Linn Soc* **130**: 700–714. doi: 10.1093/biolinnean/blaa080.

Poehlman WL, Hsieh JJ and **Feltus FA**. 2019. Linking binary gene relationships to drivers of renal cell carcinoma reveals convergent function in alternate tumor progression paths. *Sci Rep* **9**: 2899. doi: 10.1038/s41598-019-39875-y.

Rogers J, Raveendran M, Harris RA, Mailund T, Leppälä K, Athanasiadis G, Schierup MH, Cheng J, Munch K, Walker JA, **Konkel MK**, Jordan V, Steely CJ, Beckstrom TO, Bergey C, Burrell A, Schrempf D, Noll A, Kothe M, Kopp GH, Liu Y, Murali S, Billis K, Martin FJ, Muffato M, Cox L, Else J, Disotell T, Muzny DM, Phillips-Conroy J, Aken B, Eichler EE, Marques-Bonet T, Kosiol C, Batzer MA, Hahn MW, Tung J, Zinner D, Roos C, Jolly CJ, Gibbs RA and Worley KC. 2019. The comparative genomics and complex population history of *Papio* baboons. *Sci Adv* **5**: eaau6947. doi: 10.1126/sciadv.aau6947.

Sass TN, **MacPherson RA, Mackay TFC** and **Anholt RRH**. 2020. A high-throughput method for measuring alcohol sedation time of individual *Drosophila melanogaster. J Vis Exp* **158**: e61108. doi: 10.3791/61108.

Sekhon R, Joyner C, Ackerman A, **McMahan C,** Cook D and Robertson D. 2020. Stalk bending strength is strongly associated with maize stalk lodging incidence across multiple Environments. *Field Crops* **249**: e107737. doi: 10.1016/j.fcr.2020.107737.

Self SCW, Liu Y, Nordone SK, Yabsley MJ, Walden HS, Lund RB, Bowman DD, Carpenter C, **McMahan CS** and Gettings JR. 2019. Canine vector-borne disease: mapping and the accuracy of forecasting using big data from the veterinary community. *Anim Health Res Rev* **20**: 47-60. doi: 10.1017/S1466252319000045.

Spoor S, Cheng CH, Sanderson LA, Condon B, Almsaeed A, Chen M, Bretaudeau A, Rasche H, Jung S, Main D, Bett K, Staton M, Wegrzyn JL, **Feltus FA** and Ficklin SP. 2019. Tripal v3: an ontology-based toolkit for construction of FAIR biological community databases. *Database (Oxford)* doi: 10.1093/database/baz077.

Storer JM, Mierl JR, Brantley SA, Threeton B, Sukharutski Y, Rewerts LC, St Romain CP, Foreman MM, Baker JN, Walker JA, Orkin JD, Melin AD, Phillips KA, **Konkel MK** and Batzer MA. 2019. Amplification dynamics of Platy-1 retrotransposons in the Cebidae Platyrrhine lineage. *Genome Biol Evol* **4**: 1105-1116. doi: 10.1093/gbe/evz062.

**Sunkara LT**, Ahmad SM and Heidari M. 2019. RNA-seq analysis of viral gene expression in the skin of Marek's disease virus infected chicken. *Vet Immunol Immunopathol* **213**: 109882. doi: 10.1016/j.vetimm.2019.109882.

Targonski C, Shearer CA, Shealy BT, Smith MC and **Feltus FA.** 2019. Uncovering biomarker genes with enriched classification potential from Hallmark gene sets. *Sci Rep* **9**: 9747. doi: 10.1038/s41598-019-46059-1.

Tsai KL, Evans JM, Noorai RN, Starr-Moss AN and **Clark LA**. 2019. Novel Y chromosome retrocopies in canids revealed through a genome-wide association study for sex. *Genes (Basel)* **10**: 320. doi: 10.3390/genes10040320.

Vanderlinde T, Guimarães DE, **Nazario-Yepiz NO** and Carvalho AB. 2019. An improved genome assembly for *Drosophila navojoa*, the basal species in the mojavensis cluster. *J Hered* **110**: 118-123. doi: 10.1093/jhered/esy059.

Walker JA, Jordan VE, Storer JM, Steely CJ, Gonzalez-Quiroga P, Beckstrom TO, Rewerts LC, St Romain CP, Rockwell CE, Rogers J, Jolly CJ and **Konkel MK**. 2019. *Alu* insertion polymorphisms shared by *Papio* baboons and *Theropithecus gelada* reveal an intertwined common ancestry. *Mob DNA* **10**: 46. doi: 10.1186/s13100-019-0187-y.

**Walters JD**, Gill SS and Mercuri JJ. 2019. Ethanol-mediated compaction and cross-linking enhance mechanical properties and degradation resistance while maintaining cytocompatibility of a nucleus pulposus scaffold. *J Biomed Mater Res B Appl Biomater* **8**: 2488-2499. doi: 10.1002/jbm.b.34339.

Wang J and **Wang L**. 2020. Deep analysis of RNA N6-adenosine methylation (m6A) patterns in human cells. *NAR Genomics and Bioinformatics* **2**: lqaa007. doi: 10.1093/nargab/lqaa007.

Wang J and **Wang L**. 2019. Deep learning of the back-splicing code for circular RNA formation. *Bioinformatics* **35**: 5235-5242. doi: 10.1093/bioinformatics/btz382.

Withana Gamage PW, Chaudari M, **McMahan CS**, Kim EH and Kosorok MR. 2020. An extended proportional hazards model for interval-censored data subject to instantaneous failures. *Lifetime Data Anal* **26**: 158-182. doi: 10.1007/s10985-019-09467-z.

Yang T, Gallagher CM and **McMahan CS.** 2019. A robust regression methodology via M-estimation. *Communications in Statistics* **48:** 1092-1107. doi: 10.1080/03610926.2018.1423698.

Zeng W, Chen X, **Duren Z**, Wang Y, Jiang R and Wong WH. 2019. DC3 is a method for deconvolution and coupled clustering from bulk and single-cell genomics data. *Nat Commun* **10**: 4613. doi: 10.1038/s41467-019-12547-1.

Zhang X, Hong D, Ma S, Ward T, Ho M, Pattni R, **Duren Z**, Stankov A, Bade Shrestha S, Hallmayer J, Wong WH, Reiss AL and Urban AE. 2020. Integrated functional genomic analyses of Klinefelter and Turner syndromes reveal global network effects of altered X chromosome dosage. *Proc Natl Acad Sci USA* **117**: 4864-4873. doi: 10.1073/pnas.1910003117.

Zhebentyayeva T, **Shankar V**, Scorza R, Callahan A, Ravelonandro M, Castro S, DeJong T, Saski CA, Dardick C. 2019. Genetic characterization of world-wide *Prunus domestica* (plum) germplasm using sequence-based genotyping. *Hortic Res* **6**: 12. doi: 10.1038/s41438-018-0090-6.

Zhou S, **Morgante F**, Geisz MS, Ma J, **Anholt RRH** and **Mackay TFC**. 2020. Systems genetics of the *Drosophila* metabolome. *Genome Res* **30**: 392-405. doi: 10.1101/gr.243030.118.

Zhu X, **Duren Z** and Wong WH. 2020. Modeling regulatory network topology improves genome-wide analyses of complex human traits. *BioRxiv* doi: 10.1101/2020.03.13.990010.