



ASHG 2023 Plenary Abstracts

As of October 10, 2023

Instructions

1. View the Table of Contents on the next page.
 - a. The Table of Contents is sorted chronologically by the Session Number.
2. Click on the Abstract Title of the abstract you would like to view, and you will be taken to the page the abstract is on.

Table of Contents

Session 007: Featured Plenary Abstract Session I	2
Title: Innovative non-viral delivery of CRISPR/Cas9 gene editing for neurogenetic disorder of Angelman syndrome	2
Title: Protein-truncating variants in <i>BSN</i> are associated with severe obesity, type 2 diabetes and fatty liver disease.....	3
Title: Implicit bias in genetics education and healthcare: Traditional textbook photographs elicit negative student responses while natural photographs are educational and stimulate positive quality of life perceptions.	4
Title: Reconstructing the <i>cis</i> -regulatory landscape of archaic hominids using deep learning.	5
Title: Modeling and characterizing development of transposition of the great arteries using patient-derived iPSC cardiac organoids	6
Session 047: Featured Plenary Abstract Session II	7
Title: In vivo correction of human phenylketonuria variants via prime editing and base editing: A gateway to equitable treatment of inborn errors of metabolism.....	7
Title: A single cell eQTL atlas for cell type specific regulatory effects.....	8
Title: Assessment of human duplicated genes in a complete telomere-to-telomere genome implicates novel paralogs in brain evolution.....	9
Title: Somatic DNA repeat expansion underlies Huntington’s Disease neuropathology.....	10
Title: Perturb-tracing enables scalable high-content discovery of 3D genome regulators.....	11
Session 127: Featured Plenary Abstract Session III	12
Title: Identifying modifier genes in a PIGA-CDG pedigree with reduced penetrance	12
Title: Integration of genetics and proteomics at scale yields novel insights on the underpinnings of asthma risk and heterogeneity.....	13
Title: The NIH IDENTIFY study: A prospective evaluation of pregnant women with prenatal cfDNA sequencing results that suggest maternal malignancy.	14
Title: Deciphering the current landscape of PGS prediction in diverse populations - phenome-wide evaluations of genetic risk in the biobank at the Colorado Center for Personalized Medicine	15
Title: Antagonistic selection of infectious and non-infectious immune-mediated disorders revealed by vertical pleiotropy.	16

ASHG 2023 Annual Meeting Plenary Abstracts

Session 007: Featured Plenary Abstract Session I

Location: Conv Ctr/Ballroom ABC/Level 3

Session Time: Wednesday, November 1, 2023, 5:30 pm - 7:10 pm

Title: Innovative non-viral delivery of CRISPR/Cas9 gene editing for neurogenetic disorder of Angelman syndrome

Authors: X. Lu, Y. Xie, Y. Bao, G. Long, W. Sheu, J. Zhou, Y-H. Jiang; Yale Univ. Sch. of Med., New Haven, CT

Abstract:

Gene editing using CRISPR/Cas9 for neurogenetic disorders is challenged by difficulty in crossing blood-brain barrier, limited penetration, and narrow therapeutic windows. While modified adeno-associated virus (AAVs) overcome some of these barriers, their immunogenicity and higher risk of off-target effects due to the long-term presence of Cas9 protein limits their translational value for human neurogenetic disorders. To address this, we developed an innovative non-viral delivery tool using chemically modified ribonucleoprotein (RNP) conjugated with Cas9 protein and sgRNA (cRNP-Cas9/sgRNA, cRNPcg). Due to its small size (12um), cRNPcg achieved effective penetration of neuronal cells in the brain and the transient Cas9 protein substantially reduced the risk of off-target effects. We tested the cRNPcg in vitro and in vivo for Angelman syndrome (AS), a neurodevelopmental disorder caused by the deficiency of neuronal and maternal-specific *UBE3A* gene expression. The repressed expression of *UBE3A* in the paternal chromosome is mediated by the paternally expressed non-coding *UBE3A* antisense transcripts (*UBE3A-ATS*). Inactivation of *UBE3A-ATS* via antisense oligo (ASO) has shown positive clinical efficacy in ongoing phase 1/2 clinical trials. However, the transient effect of ASO requires monthly intrathecal deliveries poses a challenge as a standard clinical treatment. We designed cRNPcg system selectively inactivates *Ube3a-ATS* expression and potentially achieves permanent therapeutic effects via a single treatment. Using *Ube3a-YFP* reporter mice, we observed high gene editing efficiency (>75% targeting cells) with widespread brain penetration. We intrathecally administered cRNPcg to newborn (P1-2) and P21 AS *Ube3a^{mat-p}* model and observed significant reduction of *Ube3a-ATS* and reactivation of *Ube3a* up to 30% of normal levels across the cortex, hippocampus, and cerebellum. Accordingly, this treatment led to significant improvement in multiple behavioral domains, including locomotor function, anxiety-like behaviors, learning and memory, and also extended the latency for chemical-induced myoclonic and tonic seizures in adult AS *Ube3a^{mat-p}* mice. Importantly, we have not observed any acute or chronic toxicity associated with cRNPcg. Moreover, we have found cRNPcg effectively reactivates *UBE3A* expression from the paternal chromosome in AS patients' iPSC-derived neural progenitor cells with large maternal deletions of 15q11-q13. Together, our results demonstrated cRNPcg is an innovative platform to deliver CRISPR/Cas9 gene editing to the brain that has broad applications and potential to treat many other neurogenetic disorders.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Protein-truncating variants in *BSN* are associated with severe obesity, type 2 diabetes and fatty liver disease

Authors: Y. Zhao¹, M. Chukanova², K. Kentistou¹, A. M. Siegert², R. Jia¹, G. Dowsett², E. Gardner¹, F. R. Day¹, L. R. Kaisinger¹, Y-C. L. Tung², B. Y. H. Lam², H-J. C. Chen², F. Merkle², N. J. Wareham¹, S. O'Rahilly², K. K. Ong¹, G. S. H. Yeo², J. R. B. Perry^{1,2}; ¹MRC Epidemiology Unit, Wellcome-MRC Inst. of Metabolic Sci., Univ. of Cambridge Sch. of Clinical Med., Cambridge, United Kingdom, ²Metabolic Res. Lab., Wellcome-MRC Inst. of Metabolic Sci., Univ. of Cambridge Sch. of Clinical Med., Cambridge, United Kingdom

Abstract:

Obesity is a major risk factor for many common diseases and has a significant heritable component. While clinical and large-scale population studies have identified several genes harbouring rare alleles with large effects on obesity risk, there are likely many unknown genes with highly penetrant effects remaining. To this end, we performed whole exome-sequence analyses for adult body mass index (BMI) in up to 587,027 individuals. We identified rare, loss of function variants in two genes - *BSN* and *APBA1* - with effects on BMI substantially larger than well-established obesity genes such as *MC4R*. One in ~6500 individuals carry a heterozygous protein truncating variant (PTV) in *BSN*, which confers a 6.6, 3.7 and 3-fold higher risk of severe obesity (BMI >40kg/m²), non-alcoholic fatty liver disease and type 2 diabetes, respectively. In contrast to most other obesity-related genes, rare variants in *BSN* and *APBA1* had no apparent effect on childhood adiposity. Furthermore, *BSN* PTVs magnified the influence of common genetic variants associated with BMI, with a common polygenic score exhibiting an effect on BMI twice as large in *BSN* PTV carriers than non-carriers. Finally, we explored the plasma proteomic signatures of *BSN* PTV carriers as well as the functional consequences of *BSN* deletion in human iPSC-derived hypothalamic neurons. These approaches highlighted a network of differentially expressed genes that were collectively enriched for genomic regions associated with BMI, and suggest a role for degenerative neuronal synaptic function and neurotransmitter release in the etiology of obesity.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Implicit bias in genetics education and healthcare: Traditional textbook photographs elicit negative student responses while natural photographs are educational and stimulate positive quality of life perceptions.

Authors: C. Falugi^{1,2}, S. Hashmi³, N. Wood⁴, C. Hornbuckle⁵, R. Guidotti⁶, C. Singletary^{2,3,7}, A. D'Angelo⁸, J. Nguyen⁹, L. Vanner-Nicely¹⁰, S. Elsea¹¹; ¹Univ. of Texas Med. Branch, Galveston, TX, ²Genetic Counseling Program, Univ. of Texas MD Anderson Cancer Ctr. UTHlth.Houston Graduate Sch. of BioMed. Sci., Houston, TX, ³Dept. of Pediatrics, McGovern Med. Sch. at UTHlth.Houston, Houston, TX, ⁴Sarasota Mem. Hosp. Cancer Inst., Sarasota, FL, ⁵Dept. of Genetics, Children's Hlth.System of Texas, Dallas, TX, ⁶Positive Exposure, New York, NY, ⁷Dept. of Obstetrics, Gynecology and Reproductive Sci., McGovern Med. Sch. at UTHlth., Houston, TX, ⁸Dept of Genetics, LSU Hlth.Sci. Ctr., New Orleans, LA, ⁹Clinical Genetics, Cook Children's Hosp., Fort Worth, TX, ¹⁰Dept. of Obstetrics and Gynecology, Virginia Commonwealth Univ., Richmond, VA, ¹¹Dept. of Molecular and Human Genetics, Baylor Coll. of Med., Houston, TX

Abstract:

Traditional textbook images historically used in healthcare education often depict individuals with visible genetic conditions in the nude and in stark, clinical settings to emphasize dysmorphic features. Images taken in natural surroundings may offer a different perspective and educational opportunity. To assess the effectiveness of natural images in comparison to traditional textbook photos for educational purposes and their impact on viewers' perceptions, we surveyed medical, dental, nursing, and allied healthcare students in a two-phase analysis. In phase one, 516 students were randomly assigned to online knowledge-based surveys modeled after textbook entries on Marfan, Cornelia de Lange, and Noonan syndromes that included either traditional textbook photographs or natural portraits. Comprehensive analysis of the presented information revealed no significant difference in the readers' understanding of the subject material when viewing natural or traditional images. Phase two, a crossover study, assessed the influence of these photographs on students' (n=549) perceptions of the depicted individual with the genetic condition and their quality of life using a Characteristic Assessment (CA) composed of 7 Likert scale questions, a Perception Assessment (PA) consisting of 18 yes/no adjective-association questions, and open-ended response questions. Natural images positively influenced the perception of the depicted individuals with visible genetic conditions and were more often associated with positive characteristics (e.g., beautiful, respectful, intelligence, higher quality of life) ($p < 0.05$), while traditional textbook photographs consistently elicited negative perceptions and were associated with negative characteristics (e.g., degrading, institutionalized, humiliating, neglected) ($p < 0.05$). These associations were evident regardless of which photograph was viewed first. Unsolicited comments from participants centered on students finding natural photos educational and beneficial, positive changes in their perception after viewing natural photographs, and their utility for patient and their family members' education. Photos are a common media utilized in healthcare student education and have long-lasting impact on learners' views, perceptions, and knowledge. Preconceived or implicit bias in healthcare must be addressed early in education. Utilizing patient-centered images will support change by reducing the biases associated with genetic conditions that are reinforced by more traditional approaches to healthcare student education.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Reconstructing the *cis*-regulatory landscape of archaic hominids using deep learning.

Authors: A. Patel, G. K. Marinov, A. Kundaje; Stanford Univ., Stanford, CA

Abstract:

Sequencing ancient genomes has revolutionized our understanding of human evolution. However, the potential of these sequences remains untapped, as the impact of genetic variation in these genomes, particularly in *cis*-regulatory elements, is poorly understood. In this study, we employ interpretable deep learning techniques to investigate the functional consequences of over 50,000 unique variants found in the genomes of Neanderthals and Denisovans, the closest archaic hominid species to humans. We trained deep learning models to accurately map DNA sequences to base-resolution chromatin accessibility profiles across over 1500 diverse cellular contexts from the ENCODE project. Leveraging these models, we conduct *in silico* mutagenesis experiments to quantify and score the regulatory effects of each variant within specific cell contexts, shedding light on disrupted sequence features.

Our analysis reveals enriched variant effect sizes in specific organs and tissue systems, including the nervous system (with a focus on brain development), metabolism, and skeletal morphology. By clustering variant effects across cell types, we identify distinct clusters associated with heightened activity in brain tissue, immune cells, and other important cell types. These findings are consistent with known differences between archaic and modern humans, as corroborated by previous studies and the skeletal record.

To gain insights into the mechanisms underlying these regulatory effects, we utilize feature attribution methods to predict the disruption of transcription factor sequence motifs by the variants. Notably, we discover instances of motif creation and disruption near functionally relevant genes due to hominid alleles. For example, an archaic allele near brain-related genes *SPRY3* and *TMLHE* generates a predicted accessibility peak by creating a leucine zipper motif, suggesting its potential involvement in nervous system regulation in archaic humans. Furthermore, we observe compelling cases of synergistic variant effects, where high accessibility impacts are predicted only when a series of archaic variants in close proximity are substituted together, highlighting the importance of considering variant combinations.

This work represents the first quantitative analysis of cell context-specific regulatory differences in chromatin activity between modern humans and our closest archaic hominid relatives. Our study provides a comprehensive dataset of archaic variant effects specific to cell types, highlighting the utility of deep learning approaches in unraveling the intricacies of archaic genomes.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Modeling and characterizing development of transposition of the great arteries using patient-derived iPSC cardiac organoids

Authors: M. Keever-Keigher¹, D. A. Louiselle¹, J. C. Means¹, S. T. Younger², C. Smail³; ¹Children's Mercy Res. Inst., Kansas City, MO, ²Children's Mercy Kansas City, Kansas City, MO, ³Children's Mercy Kansas City, Leawood, KS

Abstract:

Transposition of the great arteries (TGA) is a congenital heart defect where the normal connections of the ventricles and arteries are reversed. Specifically, in dextro-transposition of the great arteries (d-TGA), the attachment of the aorta to the right ventricle and the pulmonary artery to the left ventricle results in parallel circulation of deoxygenated and oxygenated blood. Thus, deoxygenated blood is unable to travel back to the lungs for reoxygenation. Much of our current understanding of the development of d-TGA is derived from genome-wide association studies, whole exome sequencing, whole genome sequencing, and animal studies. While these studies have aided in our understanding of the genetics of d-TGA, the complex nature and timing of onset of d-TGA in utero contribute to the difficulty of precisely identifying genes and pathways involved in this phenotype. Identification of d-TGA cases before birth is a critical unsolved challenge as undetected cases require emergency surgical intervention. We hypothesized that single-cell functional profiling of patient-derived induced pluripotent stem cells (iPSCs) across differentiation timepoints would recapitulate causal gene regulatory landscapes and help pinpoint the timing of d-TGA onset. To test this hypothesis, we combined clinical whole genome sequencing with single-cell RNA-seq and ATAC-seq of iPSCs across eight differentiation timepoints - from embryonic bodies to cardiac organoids - from a d-TGA patient case and this patient's morphologically normal parents. As proof-of-concept of our iPSC approach, we compared bulk RNA-seq expression of candidate d-TGA genes (HPO ID: HP0031348; N genes = 7) in 1,441 iPSC lines from the i2QTL consortium resource with disease-relevant tissue samples from GTEx (Atrial Appendage and Left Ventricle), reprocessing GTEx to match the i2QTL processing pipeline. Six of the seven genes were identified in the i2QTL resource: *SMG9*, *ACVR2B*, *ZNF462*, *NODAL*, *ZIC3*, *SMAD2*. Five of these genes had significantly increased expression (mean logFC = 3.15; FDR < 0.05) in iPSCs compared to the selected GTEx data. We performed an additional analysis using d-TGA candidate genes prioritized by state-of-the-art variant interpretation software (Illumina Emedgene). Our proof-of-concept findings support the rationale for using iPSCs to understand the complex causal mechanisms of d-TGA candidate genes. Ultimately, we expect single-cell iPSC functional profiling across developmental timepoints from this rare disease family to yield important insights into the role of candidate genes in the development of d-TGA, as well as pathways and gene networks in which they participate.

ASHG 2023 Annual Meeting Plenary Abstracts

Session 047: Featured Plenary Abstract Session II

Location: Conv Ctr/Ballroom ABC/Level 3

Session Time: Thursday, November 2, 2023, 5:30 pm - 7:10 pm

Title: In vivo correction of human phenylketonuria variants via prime editing and base editing: A gateway to equitable treatment of inborn errors of metabolism.

Authors: D. Brooks¹, M. N. Whittaker¹, S. A. Grandinette¹, X. Wang¹, R. C. Ahrens-Nicklas², K. Musunuru¹; ¹Univ. of Pennsylvania, Philadelphia, PA, ²Children's Hosp. of Philadelphia, Philadelphia, PA

Abstract:

Phenylketonuria (PKU), an autosomal recessive disorder caused by pathogenic variants in the phenylalanine hydroxylase (*PAH*) gene, results in the buildup of blood phenylalanine (Phe) to neurotoxic levels. Current dietary and medical treatments are chronic and reduce, rather than normalize, blood Phe levels. The most frequently occurring *PAH* variants in PKU patients are c.1222C>T (R408W), c.1066-11G>A, c.782G>A (R261Q), c.728G>A (R243Q), c.1315+1G>A, and c.842C>T (P281L), varying widely across populations. R408W is the most prevalent variant in the U.S. and parts of Europe. Using real-world data from a PKU cohort managed at a Metabolic Specialty Clinic, we found that most patients with at least one R408W allele (n = 36) experience chronic, severe Phe elevations and do not meet Phe monitoring guidelines—highlighting the high unmet medical need arising from the challenges of lifelong adherence to dietary and medical therapy.

Motivated by these findings, we are developing “one-and-done” CRISPR editing therapeutics for PKU. We have established a pipeline to rapidly screen and optimize prime editing or base editing strategies in vitro and deploy them in vivo with either lipid nanoparticles (LNPs) or adeno-associated viral (AAV) vectors on a timescale of months. In initial work, we created six humanized PKU mouse models with the aforementioned *PAH* variants and, in parallel, prime-edited variant-bearing hepatocyte cell lines. In homozygous or compound heterozygous PKU mice, we observe complete and long-term durable normalization of blood Phe levels (>90% reduction) as soon as 48 hours after treatment, with whole-liver corrective *PAH* editing as high as >50% with either prime editing or base editing. We have developed therapeutic leads for R408W and P281L to permanently normalize blood Phe levels and definitively treat PKU in these patients—on track for early-phase clinical trials within a few years—and have promising candidates for the other *PAH* variants.

Despite these encouraging results, we are cognizant of the risk of “mutational discrimination” in unduly focusing on high-prevalence variants skewed towards specific ancestry groups. Accordingly, we have extended our pipeline to develop and validate a corrective therapeutic for any variant in any patient with a hepatic inborn error of metabolism. We have begun to apply this workflow in real time to patients with devastating, ultra-rare inborn errors of metabolism such as urea cycle disorders and organic acidemias, tackling any variant—even n-of-1 variants—identified via universal newborn screening. Our goal is to be able to rapidly devise and deploy a personalized editing treatment for any patient in need.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: A single cell eQTL atlas for cell type specific regulatory effects

Authors: L. Wang, D. Liu; Penn State Coll. of Med., Hershey, PA

Abstract:

GWAS has discovered many associations between human complex traits and diseases. Most associated variants are non-coding and regulate gene expression. Identifying the target genes requires integrating expression quantitative trait locus (eQTL) data from TWAS or colocalization analysis. However, many GWAS hits fail to colocalize with eQTL variants, possibly because regulatory effect differences between cell types are not captured by bulk RNASeq data.

Single-cell RNASeq (scRNASeq) can systematically characterize cellular heterogeneity of gene expression. By grouping scRNASeq into pseudo-bulks of similar cell types, cell type specific eQTLs (sc-eQTLs) can be calculated. As in sc-RNASeq data, the success of TWAS and colocalization critically depends on the power of sc-eQTL studies. The largest sc-eQTL datasets contains 1000s of individuals, which are still much smaller than bulk RNASeq data. To improve the power of sc-eQTL datasets, we develop a new method JOBS (Joint analysis Of Bulk and Sc-eQTL). JOBS relies on the key insight that bulk eQTL datasets can be viewed as a weighted average of sc-eQTLs from constituting cell types. Jointly analyzing bulk- and sc-eQTL datasets with a compositional model will borrow strength from the large sample size of bulk eQTL datasets and substantially improve sc-eQTL effect estimates from different cell types.

Applying JOBS, we jointly analyze OneK1K (a large sc-eQTL data with $N=966$) and eQTLGen (a large bulk eQTL data with $N=31684$) and validate the results in DICE and sc-eQTLGen. JOBS identifies 41.5% and 67.6% more eGenes, i.e., genes with at least one significant eQTL ($p < 0.05$ with Bonferroni correction), that are replicated in DICE and sc-eQTLGen. To illustrate the utility of JOBS, we compiled GWAS summary statistics from 14 autoimmune diseases with largest $n=740000$. Using COLOC, we colocalize 271 loci, which is 29.86% and 33.21% more than using OneK1K and eQTLGen respectively. We also employed the summary statistics-based TWAS method EXPRESSO with OneK1K, eQTLGen, JOBS-eQTLs, and mashR-eQTLs, and individual level data based TWAS method UTMOST and PrediXcan. EXPRESSO-JOBS yields 25.26%, 26.54%, 22.57%, 39.22%, and 79.72% more significant prediction models compared to alternative methods. Applying our results to 14 immune-related traits, JOBS identified 86.9% and 50.2% more cell type-specific significant gene-trait associations than EXPRESSO-SC and UTMOST. It also nominated novel risk genes including *IRF1* for systemic lupus erythematosus and *PTGER4* for Crohn's disease. The sc-eQTL atlas we created will be a valuable resource for the research community to investigate cell type specific effects for complex traits.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Assessment of human duplicated genes in a complete telomere-to-telomere genome implicates novel paralogs in brain evolution

Authors: J. Uribe-Salazar¹, D. Soto¹, G. Kaya¹, A. Sekar¹, E. Green², S. Simó¹, G. Quon¹, A. Andres³, M. Dennis¹; ¹Univ. of California, Davis, Davis, CA, ²Natl. Human Genome Res. Inst. (NIH), Bethesda, MD, ³Univ. Coll. London, London, United Kingdom

Abstract:

Genomic drivers of human-specific neurological traits remain largely undiscovered. Duplicated genes represent an understudied source of species innovation that are historically missing/error-prone in reference assemblies. Using the new telomere-to-telomere assembly (T2T-CHM13), we identified 1,834 recently-duplicated genes (>98% paralog identity) of which 1,121 can be found in all modern humans in the 1000 Genomes Project (1KGP; n=2,504) with 276 fixed in copy-number. Assaying transcriptome data of human fetal brain and embryonic stem-cell derived cortical neurons, we identified 288 protein-coding paralogs expressed during early brain development, with a significant enrichment of copy-number fixed genes (92% vs. 62% total) suggesting likely functional genes; this list includes *SRGAP2C*, *ARHGAP11B*, and *NOTCH2NLC*, all previously implicated in neocortex function, as well as exciting new candidates, such as autism-associated *CNTNAP3*. Calculating Tajima's D using 1KGP variant data, we identified 22 protein-coding genes showing remarkable signatures of selection, including the *SPDYE3* locus; however, ~90% of duplicated loci were inaccessible to short reads. Narrowing in on a set of 15 duplicate gene families, we sequenced via targeted PacBio HiFi ~200 modern humans of diverse ancestries and, coupled with variants discovered from the Human Pangenome Reference Consortium, show that many paralogs retain conservation at the amino-acid level, likely as a result of purifying selection. To characterize functions of ten duplicate gene families in zebrafish, we performed CRISPR "knockout" of orthologs and transiently introduced mRNA encoding paralogs, effectively "humanizing" larvae. Phenotyping >8,500 zebrafish larvae using morphological and behavioral assessments, and single-cell RNA-seq of >167,000 cells across models narrowed in on duplicate gene families with neural functions. This list included *GPR89*, a gene encoding a voltage-gated chloride channel that regulates Golgi pH, leading to a 9% reduction in *gpr89*-knockouts ($p=7 \times 10^{-7}$) and 3% increase ($p=0.022$) of *GPR89B*-humanized larvae in head area compared to batch-sibling controls. Using a transgenic line with GFP-tagged neurons corroborated these results in the brain, with 30% reduction in knockouts ($p=0.006$) and 26% enlargement in humanized larvae ($p=4 \times 10^{-4}$) of the forebrain area. Ongoing work in mouse neocortex suggests the gene may be important in cell morphology and neurite outgrowth. Our holistic approach provides important insights into what it means to be human as well as a methodological roadmap to characterize complex genes implicated in neurological function and disease.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Somatic DNA repeat expansion underlies Huntington's Disease neuropathology

Authors: R. Handsaker^{1,2}, S. Kashin^{1,2}, N. Reed^{1,2}, W-S. Lee^{1,2}, T. M. McDonald^{1,2}, S. Tan^{1,2}, K. French^{3,4}, N. Kamitaki^{1,2}, S. Berretta^{3,4}, S. McCarroll^{1,2}; ¹Harvard Med. Sch., Dept. of Genetics, Boston, MA, ²Broad Inst., Cambridge, MA, ³Lab. for Structural NeuroSci., McLean Hosp., Belmont, MA, ⁴Harvard Med. Sch., Dept. of Psychiatry, Boston, MA

Abstract:

Huntington's Disease (HD) is a fatal genetic neurodegenerative disorder involving loss of specific types of neurons. It is not known why these specific neurons are vulnerable, nor why they live for decades before degenerating.

Persons with HD (pwHD) have inherited alleles of the Huntingtin (*HTT*) gene in which a DNA sequence repeat (CAG_n, encoding polyglutamine) is repeated 36 or more times; this same DNA repeat also exhibits length mosaicism in their tissues.

We developed a way to measure the length of the *HTT* CAG repeat together with genome-wide RNA expression in individual cells. We found that inherited HD-causing alleles undergo profound somatic expansion, up to 1,000 CAGs, in striatal spiny projection neurons (SPNs or MSNs) and cortical projection neurons - which degenerate in HD - but modest-to-no expansion in other neuronal and glial cells in the same brain areas. By identifying allelic series of SPNs (with varying *HTT*-CAG lengths) within the same person's brain, we uncovered the cell-autonomous effect of the *HTT*-CAG repeat upon the biological states of SPNs. Surprisingly, *HTT*-CAG expansions from 36 to 180 CAGs had no apparent effect on single-neuron gene expression patterns. However, SPNs with *HTT* CAG repeats longer than 180 exhibited profound gene-expression changes, systematically losing the gene-expression features that distinguish SPNs from other types of cells. Such transformed neurons appeared in all pwHD we assayed, with frequencies that matched the kinetics of SPN loss at their disease stage.

These results reveal that individual neurons pass asynchronously, after decades of somatic *HTT*-CAG expansion, through a cell-biological toxicity phase that is brief and rapidly progressive. An important implication is that the pathogenesis of HD is a DNA process for almost all of a neuron's life. These results suggest promising directions for development of therapeutics targeting the somatic expansion process in HD and strategies for the design of clinical trials.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Perturb-tracing enables scalable high-content discovery of 3D genome regulators

Authors: S. Wang¹, M. Hu¹, B. Yang¹, T. Jensen¹, Y. Cheng¹, R. Yu¹, Z. Ma², J. Radda¹, S. Jin¹, C. Zang²; ¹Yale Univ., New Haven, CT, ²Univ. of Virginia, Charlottesville, VA

Abstract:

Three-dimensional (3D) genome organization becomes altered during development, aging, and disease, but the factors regulating chromatin topology are incompletely understood and currently no technology can efficiently screen for new regulators of multiscale chromatin organization. Here, we developed an image-based high-content genetic screening platform (Perturb-tracing) that combines pooled CRISPR screen, a new cellular barcode readout method (BARC-FISH), and 3D tracing of the super-resolved folding conformation of an entire chromosome. We performed a loss-of-function screen in human cells, and visualized alterations to their genome organization from 13,000 imaging target-perturbation combinations and 1.4 million 3D positions along chromosome traces, alongside perturbation-paired barcode readout in the same single cells. We discovered tens of new regulators of chromatin folding at different length scales, ranging from adjacent topologically associating domains (TADs) to A-B compartments to chromosome territories. Subsets of the regulators showed 3D genome effects correlated with loop-extrusion and A-B compartmentalization mechanisms, while others appeared to be largely unrelated to these known 3D genome regulatory mechanisms. We found that the ATP-dependent helicase CHD7, previously known to promote local chromatin openness, counter-intuitively compacts chromatin over long range in a CTCF-dependent manner, and facilitates a multi-TAD compaction effect of cohesin. Finally, we identified new regulators of nuclear architectures and found a mechanistic link between chromatin compaction and nuclear shape. Altogether, our method enables scalable, high-content identification of chromatin and nuclear topology regulators that will stimulate new insights into the 3D genome and nucleome.

ASHG 2023 Annual Meeting Plenary Abstracts

Session 127: Featured Plenary Abstract Session III

Location: Conv Ctr/Ballroom ABC/Level 3

Session Time: Saturday, November 4, 2023, 5:00 pm - 6:40 pm

Title: Identifying modifier genes in a PIGA-CDG pedigree with reduced penetrance

Authors: H. Thorpe, B. Pedersen, J. Bonkowsky, A. Quinlan, C. Chow; Univ. of Utah, Salt Lake City, UT

Abstract:

Phosphatidylinositol glycan class A (PIGA) encodes the catalytic component of the enzyme in the first step of GPI-anchor biosynthesis. GPI-anchors act as membrane anchors for over 150 proteins involved in signal transduction, immune response, and cellular communication among other functions. Loss of function mutations in *PIGA* lead to PIGA deficiency (PIGA-CDG), an ultra-rare disease, typically presenting with seizures, hypotonia, and neurodevelopmental delays. PIGA deficiency is an X-linked recessive congenital disorder of glycosylation (CDG). We identified two brothers (proband) with PIGA-CDG, presenting with mild developmental delay, epilepsy, and autism. Both probands carry the novel, rare *PIGA*^{S132C} variant, a predicted damaging variant not found in the gnomAD database. Confirming this diagnosis, both probands show a 50% decrease in GPI-anchor proteins on the cell surface. Strikingly, the maternal grandfather and a great uncle both also carry *PIGA*^{S132C}, but neither presents with symptoms associated with PIGA-CDG. We hypothesized that there might be a modifier segregating in the family that contributes to this reduced penetrance. Using whole genome sequencing and pedigree analysis, we identified all the possible susceptibility variants found in the probands and not in carriers and all the possible protective variants found in the carriers and not in the probands. This list of potential candidates included heterozygous, damaging variants in each of three other genes also involved directly in GPI-anchor biosynthesis, *PIGS*, *PGAP5*, and *DPM1*, and a small number of genes involved in other glycosylation pathways or encoding GPI-anchored proteins. To functionally test our predicted modifiers, we used a *Drosophila* eye-based model of PIGA-CDG. We created double knockdowns (KD) of *PIGA* and the candidate modifiers and compared the eye sizes of the double KD to the PIGA eye model and single KD of the candidate modifiers. We observed a genetic interaction between *PIGA* and *PIGS*, a predicted susceptibility gene, in the fly that mimics what we predict in the family. We also found that loss of *CNTN2*, a predicted protective gene, can rescue *PIGA* in the fly, similar to what we predict in the family. Further testing of top candidates was also performed in a *Drosophila* neurological model of PIGA-CDG displaying seizures. Modification of the neurological model indicates that proposed susceptibility and protective modifier genes are plausible human modifiers and could explain the incomplete penetrance in this family. The identification and study of rare disease modifier genes in human pedigrees may lead to pathways and targets that may be developed into therapies.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Integration of genetics and proteomics at scale yields novel insights on the underpinnings of asthma risk and heterogeneity

Authors: L. Donoghue, C. Benner, D. Chang, R. Pendergrass, B. Yaspan, M. McCarthy; Genentech, South San Francisco, CA

Abstract:

Numerous genetic and environmental risk factors for asthma have been identified, yet harnessing these findings to explain the significant heterogeneity and unmet clinical needs for asthma remains challenging. We aimed to identify putative causal regulators of asthma risk and explore mechanisms underlying asthma heterogeneity by integrating genetic and proteomic data at scale from the UK Biobank Pharma Proteomics Project (UKB-PPP). Specifically, we leveraged levels of 2,923 plasma proteins (measured by the Olink multiplexed platform) related to diverse biological contexts from 34,490 European-ancestry participants, 4,342 with doctor-diagnosed or self-reported asthma. We identified 616 proteins associated with asthma status (Bonferroni-corrected $P < 1.7 \times 10^{-5}$), the majority of which are previously unreported. To gain insights on whether protein level variation may be causally related to disease risk or reflect the consequence of disease-state (i.e. downstream biomarker), we used fine-mapped pQTL from the UKB-PPP dataset to generate genetically-predicted protein levels (GPPLs) for 1,126 proteins that had ≥ 0.05 pQTL heritability. Association testing between these GPPLs and asthma status, asthma subtype, or asthma-related endophenotypes (e.g. blood eosinophil count, age-of-onset) in 307K independent UKB participants yielded evidence for putative causal roles of multiple proteins related to asthma drug targets (e.g. sIL4R, sIL1RL1), proteins with functional evidence in asthma biology (e.g. TNFRSF8, PD-1), and proteins with poorly understood roles asthma or asthma-endophenotypes (e.g. TSPAN8). Mendelian randomization using cis-pQTL from the UKB-PPP as instruments provided additional evidence of causal roles for proteins in asthma risk, including TLR1. Lastly, to examine how joint effects of asthma risk variants may converge on biological processes underlying etiological and clinical heterogeneity in asthma, we used patterns of genetic risk across asthma-endophenotypes, co-morbidities, and genetically-correlated traits to generate partitioned polygenic scores. These scores differentially associated with disease markers (e.g. blood eosinophils), plasma proteins (e.g. MMP10, CCL26, KIT), and single-cell RNAseq profiles of immune or structural cell types including T cells and endothelial cells, demonstrating the ability to link components of individualized genetic risk to heterogenous axes of disease and biological contexts. These findings showcase the utility of integrating genetics and proteomics toward identifying candidate drug targets, novel biomarkers, and process-specific genetic scores for complex diseases.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: The NIH IDENTIFY study: A prospective evaluation of pregnant women with prenatal cfDNA sequencing results that suggest maternal malignancy.

Authors: D. Bianchi¹, A. Turiff¹, C. Annunziata², A. Malayeri³, I. Goldlust², M. Pavelova², B. Redd³; ¹Natl. Human Genome Res. Inst., Bethesda, MD, ²Natl. Cancer Inst., Bethesda, MD, ³NIH Clinical Ctr, Bethesda, MD

Abstract:

Background Maternal malignancy is a potential explanation for discordant or nonreportable non-invasive prenatal testing (NIPT) results using cell-free (cf) DNA sequencing of maternal plasma. IDENTIFY (NCT04049604) is an ongoing prospective study at the NIH Clinical Center that aims to identify biological causes underlying discordant or nonreportable NIPT results and to generate evidence to inform reporting guidelines and the subsequent diagnostic workup of pregnant people. **Methods** Eligible participants are adult women who have received abnormal or nonreportable NIPT results, with either a sonographically normal fetus, or a euploid fetal or neonatal karyotype or chromosome microarray. At enrollment, individuals may be pregnant or up to two years postpartum and perceive themselves to be asymptomatic. Participants undergo a standardized cancer evaluation, including history and physical examination, whole-body magnetic resonance imaging (MRI), repeat research NIPT using a common genome-wide platform, fecal occult blood test, pap/HPV test, and hematological and biochemical laboratory tests. When available, placentas are analyzed for confined placental mosaicism. Participants are followed for up to five years. **Results** Between December 2019 and June 2023, over 80 of 190 individuals referred to the IDENTIFY study were enrolled. Participants' initial cfDNA sequencing was performed by one of 11 clinical laboratories offering NIPT in North America. In nearly 75% of the cases, participants received either "non-reportable" results or results that suggested an "atypical finding, likely maternal in origin." The other 25% had one or more abnormalities reported. Over 50% of participants had a malignancy diagnosed at the initial NIH evaluation. Lymphomas were the most common, followed by colorectal cancer, but other rare tumors were also found. Malignancies initially detected using SNP-based NIPT platforms tended to be more advanced. **Conclusions** In this prospective cohort, there is a high rate of cancer detection in asymptomatic women who receive nonreportable or discordant NIPT results. This highlights the importance of a timely and thorough clinical evaluation. Many non-evidence-based tests that have been suggested in the literature to evaluate maternal health following malignancy suspicious NIPT results have low yield and are likely unnecessary. Partial workups that do not include whole body imaging are insufficient. MRI was critical for detecting malignancy in this asymptomatic cohort. We hypothesize that detecting and treating maternal malignancies before development of symptoms will improve health and extend life.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Deciphering the current landscape of PGS prediction in diverse populations - phenome-wide evaluations of genetic risk in the biobank at the Colorado Center for Personalized Medicine

Authors: M. Lin¹, C. H. Arehart^{1,2}, N. Rafaels¹, K. R. Crooks¹, M. Fisher¹, S. Chavan¹, H. A. Smith¹, J. Shortt¹, N. Pozdeyev¹, A. Hendricks^{3,1}, S. Raghavan^{1,4}, C. R. Gignoux¹; ¹Univ. of Colorado Anschutz Med. Campus, Aurora, CO, ²Univ. of Colorado Boulder, Boulder, CO, ³Univ. of Colorado Denver, Denver, CO, ⁴Rocky Mountain Regional VA Med. Ctr., Aurora, CO

Abstract:

The resource of polygenic risk scores (PGS) has been rapidly growing to keep up with the abundance of genome-wide association studies. By leveraging ~3,000 scores from the PGS catalog, we comprehensively evaluated genetic risk across the phenome of more than 1,300 conditions based on electronic health records (EHR) in the biobank at the Colorado Center for Personalized Medicine (CCPM) of 73,346 participants, yielding 2,670,117 associations in total. We observed ~100 phenotypes in which individuals at the top decile of the respective PGS showed greater than two-fold increased risk of developing the condition. In addition, over 1,500 predictions of various outcomes achieved highly significant signals with $P < 5 \times 10^{-100}$, including some common diseases that were predicted especially well by PGSs together with demographic covariates, such as type 2 diabetes (T2D) and hypertension (both $P < 5 \times 10^{-324}$, AUC=0.82 and 0.80). For both of these phenotypes, individuals scoring in the top 1-3% in the CCPM biobank had OR>3, with a consistent prediction effect across genetic ancestry groups ($I^2=0$). However, most predictions showed considerable inter-group heterogeneity in performance (average $I^2=0.2$ in phecode~PGS pairs with FDR<0.1), which can greatly affect the potential use of PGS in personalized medicine. Furthermore, we observed bias in downstream applications such as Mendelian Randomization, where we found opposite directions between instrument and outcome in different ancestry groups for known correlated trait pairs (e.g. T2D and chronic kidney disease). We also noticed that the choice of PGS unit of measure, whether per SD or stratified at the top decile against the remainder, yielded weakly concordant estimates of heterogeneity ($r=0.3$). We then characterized both the accuracy and portability of PGSs in general across phenotypes using multiple predictors from both the training and test sets via multilevel nested mixed models. Our results demonstrate that there remain unaddressed challenges hindering downstream PGS applications in multi-ancestry settings, with significant associations (hierarchical LMM $p < 0.05$) including disease domain, PRS construction type, as well as population characteristics of both the training and CCPM test populations. This quantifies and describes the hopes and pitfalls underlying ongoing efforts to apply PGS resources in diverse populations.

ASHG 2023 Annual Meeting Plenary Abstracts

Title: Antagonistic selection of infectious and non-infectious immune-mediated disorders revealed by vertical pleiotropy.

Authors: G. Kerner¹, E. Patin¹, L. Quintana-Murci²; ¹Inst. Pasteur, Paris, France, ²Pasteur Inst. and Coll. de France, Paris, France

Abstract:

Pleiotropy - the process by which a locus affects multiple traits - is thought to be widespread in the human genome (Watanabe *et al.* 2019, Nat Genet). Inferring how evolution has affected genetically correlated traits can shed light on the evolutionary causes of common disease, specifically when the pleiotropic effects are antagonistic. By using an evolutionary approach and archeogenetic data, we have recently shown that the genetic risk for infectious and NIIM diseases have decreased and increased, respectively, since the Neolithic period in Europe, suggesting antagonistic selection (Kerner *et al.* 2023, Cell Genomics). However, it remains unclear whether genetic adaptations to pathogens have causally mediated the increased risk of NIIM disorders by vertical pleiotropy.

Here we integrate 200 GWASs of infectious and NIIM diseases conducted in large national biobanks and delineate 1,968 independent genetic components for both types of diseases, using SuSiE, a widely used fine-mapping method (Wang *et al.* 2020, J Royal Stat Soc). Colocalization analyses on the identified SuSiE genetic components (Wallace 2021, PLoS Genet), revealed a strongly significant excess of causal variants associated with both infectious and NIIM diseases, with 44% of the variants showing opposite direction of effects, consistent with the occurrence of antagonistic pleiotropy. By leveraging three families of neutrality tests, deriving from local tree genealogies, extended haplotype homozygosity and temporal frequency trajectories from ancient DNA data, we provide evidence of positive selection targeting pleiotropic loci, with selection signals affecting differently risk and protective derived alleles. Specifically, positively-selected risk alleles for NIIM traits are twice as many as positively-selected protective alleles. Furthermore, using selection coefficients estimated from archeogenetic data, we find NIIM-risk alleles to be under significantly stronger positive selection than NIIM-protective alleles. These results suggest that the prevalence of NIIM traits has increased in Europe as a consequence, rather than the cause, of past selection. Accordingly, we show that the effect sizes of all causal, colocalizing variants between infectious and NIIM diseases are negatively correlated ($r = -0.27$; $p = 0.016$) and the fitness of NIIM diseases is null when jointly estimating the fitness gradient of infectious and NIIM traits. Collectively, our study provides evidence that past genetic adaptation to pathogen pressures has increased the risk of NIIM diseases by antagonistic selection.