COLUMBIA PRECISION MEDICINE INITIATIVE

ADVANCES IN PRECISION MEDICINE

FRONTIERS IN HUMAN GENETICS

OCTOBER 10 TO 11, 2024

COLUMBIA | PRECISION MEDICINE

Welcome Letter

Welcome to the Eighth Annual Columbia Precision Medicine Initiative conference, on the Frontiers in Human Genetics. We are grateful to Dr. Vagelos and for the Vagelos funds, as well as for Dr. McKiernan and Dr. Armstrong's support, which made this gathering possible.

Over the past couple of decades, the field of human genetics has increasingly turned to the study of variation, from "the" human genome to millions of genomes. This shift in perspective has posed new and fundamental questions, challenging us to understand how thousands of genetic and environmental effects together give rise to human difference. This challenge is central to precision medicine, undergirding accurate disease risk prediction and tailored individual treatments, and requires the integration of multidisciplinary expertise.

To highlight fundamental work in these areas, this conference brings together an exceptional set of interdisciplinary researchers studying the causes and consequences of genetic perturbations on human phenotypes. Over the course of two days, we will hear about their different perspectives on these fundamental questions, from the processes that generate genetic variation and their cellular consequences, to the impact of regulatory variation and environmental exposures on phenotypic differences, and back from trait variation to its underpinnings.

We hope you will find the meeting inspiring.

Molly Przeworski, PhD
Alan H. Kempner Professor of Biological Sciences and of Systems Biology

Christina K. Garcia, MD, PhD
Director of the Columbia Precision Medicine Initiative (CPMI); Frode-Jensen Professor of Medicine and Chief of the Division of Pulmonary, Allergy, and Critical Care Medicine, Vagelos College of Physicians and Surgeons

Conference Schedule

FRONTIERS IN HUMAN GENETICS

Columbia University, New York

October 10, 2024	**9:00–9:10 a.m.**	**Molly Przeworski, PhD:** Introductory Remarks
	9:10–9:50 a.m.	**Inigo Martincorena, PhD**
	9:50–10:30 a.m.	**Sarah Aitken, MD, PhD**
	10:30–11:00 a.m.	Break
	11:00–11:40 a.m.	**Anjali Hinch, PhD**
	11:40 a.m.–12:20 p.m.	**Jennifer Phillips-Cremins, PhD**
	12:20–1:30 p.m.	Lunch
	1:30–2:10 p.m.	**Vijay G. Sankaran, MD, PhD**
	2:10–2:50 p.m.	**Adrianna San Roman, PhD**
	2:50–3:20 p.m.	Break
	3:20–4:00 p.m.	**Yun Song, PhD**
	4:00–4:15 p.m.	**Provost Angela Olinto, PhD:** Appreciation for Dr. Roy Vagelos
	4:15–5:00 p.m.	Reception
October 11, 2024	**9:00–9:10 a.m.**	**Christine K. Garcia, MD, PhD:** Introductory Remarks
	9:10–9:50 a.m.	**Oliver Stegle, PhD**
	9:50–10:30 a.m.	**Jonathan Pritchard, PhD**
	10:30–11:00 a.m.	Break
	11:00–11:40 a.m.	**Michel Georges, DVM**
	11:40 a.m.–12:20 p.m.	**Nasa Sinnott-Armstrong, PhD**
	12:20–1:30 p.m.	Lunch
	1:30–2:10 p.m.	**Miriam Udler, MD, PhD**
	2:10–2:50 p.m.	**Gil McVean, PhD**
	2:50–3:20 p.m.	Break
	3:20–4:00 p.m.	**Nancy Cox, PhD**
	4:00–4:40 p.m.	**Simon Myers, PhD**
	4:40–4:50 p.m.	**Molly Przeworski, PhD:** Closing Remarks

Event Chair

Christine K. Garcia, MD, PhD

Director of the Columbia Precision Medicine Initiative (CPMI); Frode-Jensen Professor of Medicine and Chief of the Division of Pulmonary, Allergy, and Critical Care Medicine in the Vagelos College of Physicians and Surgeons

Christine Kim Garcia is the Frode-Jensen Professor of Medicine, director of the Columbia Precision Medicine Initiative, and chief of the Division of Pulmonary, Allergy, and Critical Medicine within the Department of Medicine at Columbia University Irving Medical Center. Her laboratory studies the genetic basis of monogenic lung disease, with a specific focus on familial pulmonary fibrosis. Her group has identified several rare variants in genes belonging to the telomere, surfactant, and spindle pathways. She received her MD and PhD from the University of Texas Southwestern Medical Center, where she completed her residency in internal medicine and a fellowship in pulmonary and critical care medicine. In 2019 she moved to Columbia and has been a member of the Center for Precision Medicine and Genomics and an affiliate of the Institute for Genomic Medicine. Dr. Garcia has received a number of awards and honors, including the Irene and Arthur Fishberg Prize from the Vagelos College of Physicians and Surgeons (2022), American Society for Clinical Investigation (2012), Doris Duke Charitable Foundation Clinical Scientist Development Award (2008), President's Research Council Distinguished Young Investigator Award from University of Texas Southwestern Medical Center (2006), Charles E. Culpeper Foundation Medical Scholar Award (2004), Parker B. Francis Fellowship Award in Pulmonary Research (2003), and Alpha Omega Alpha (1991). She currently co-chairs the NIH Clinical Genetic (ClinGen) Pulmonary Domain Executive Committee.

Event Chair

Molly Przeworski, PhD

Professor of Biological Sciences and of Systems Biology, Columbia University

Molly Przeworski is currently the Alan H. Kempner Professor of Biological Sciences and of Systems Biology. The focus of her work is in population genetics. Dr. Przeworski earned her BA in mathematics from Princeton University and her PhD in evolutionary biology from the University of Chicago; she completed a postdoc in the statistics department at the University of Oxford. Prior to joining Columbia in 2014, she was a researcher at the Max Planck Institute for Evolutionary Anthropology and on the faculty at Brown and the University of Chicago. She was elected a member of the National Academy of Sciences and the American Academy of Arts and Sciences in 2020, and was the recipient of the 2023 Scientific Achievement Award from the American Society of Human Genetics. Dr. Przeworski's research has contributed to a better understanding of mutation and meiotic recombination processes in vertebrates, as well as of natural selection in humans.

Event Chair

Simon Tavaré, PhD

Herbert and Florence Irving Director of the Irving Institute for Cancer Dynamics (IICD); Professor, Departments of Statistics and Biological Sciences, Columbia University

Simon Tavaré obtained his PhD in probability and statistics in 1979 from the University of Sheffield and began his research career in the US. After a postdoc with Sam Karlin in Stanford, he held positions in the Department of Mathematics at the University of Utah, Statistics at Colorado State University, and Mathematics at the University of Southern California. He held the Kawamoto Chair in Biological Sciences at USC from 1998 to 2014. His research there included work in computational statistics, bioinformatics, probabilistic combinatorics, and inference for stochastic processes.

In 2003, Dr. Tavaré moved to the University of Cambridge as professor of cancer research in the Department of Oncology, a group leader in the Cambridge Research Institute from 2006, and a professor in the Department of Applied Mathematics and Theoretical Physics, where he was director of the Wellcome Trust PhD Programme in Mathematical Genomics and Medicine, and director of its MPhil Program in Computational Biology.

In February 2013, he became director of the (renamed) Cancer Research UK Cambridge Institute, a department of the University of Cambridge since January 2013. His research group focuses on statistical bioinformatics and computational biology, particularly evolutionary approaches to understanding cancer biology. In 2009, Dr. Tavaré was elected a Fellow of the Academy of Medical Sciences (FMedSci), in 2011 a Fellow of the Royal Society (FRS) and in 2015 a member of EMBO. He gave the American Mathematical Society's Einstein Lecture in 2015, and was one of the invited speakers at ICIAM2015 in Beijing. He was president of the London Mathematical Society from 2015 to 2017 and was elected a Fellow of the American Mathematical Society in 2018.

In 2018, Dr. Tavaré moved to Columbia University in New York, where he is a professor in the Departments of Statistics and Biological Sciences, and director of the new Irving Institute for Cancer Dynamics. He was elected as a foreign associate of the U.S. National Academy of Sciences in 2018.

Event Chair

Elham Azizi, PhD

Herbert & Florence Irving Assistant Professor of Cancer Data Research, Irving Institute for Cancer Dynamics; Assistant Professor of Biomedical Engineering; Affiliated Faculty of Computer Science; Affiliated Member of Data Science Institute, Columbia University

Elham Azizi's multidisciplinary research utilizes novel machine learning techniques and cutting-edge genomic technologies to study the composition and circuitry of cells in tumors. Characterizing various interacting cell types in the tumor microenvironment, and unraveling their underlying mechanisms, can guide the development of improved and personalized cancer treatments. Azizi's approach involves leveraging genomic profiling at single-cell resolution and developing machine learning and statistical methods to analyze and integrate high-dimensional genomic data.

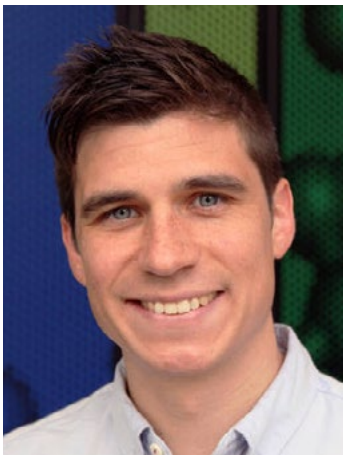
Dr. Azizi holds a BSc in electrical engineering from Sharif University of Technology (2008), and an MSc in electrical engineering (2010) and PhD in bioinformatics (2014) from Boston University. She was a postdoctoral fellow at Columbia University and Memorial Sloan Kettering Cancer Center (2014–2019). She joined the faculty of Columbia Biomedical Engineering and Irving Institute of Cancer Dynamics in 2020. She is also affiliated with the Department of Computer Science, Data Science Institute, and the Herbert Irving Comprehensive Cancer Center.

Event Chair

Guy Sella, PhD

Affiliate Member, Program for Mathematical Genomics; Professor, Department of Biological Sciences, Columbia University

Guy Sella received a BA in mathematics and physics from Tel Aviv University, an MSc in neural computation from Hebrew University, and a PhD in mathematics from Tel Aviv University (while a visitor at Stanford University). He then pursued postdoctoral research at Hebrew University and the Weizmann Institute. He is currently a professor in the Columbia University Department of Biological Sciences and a member of the Center for Computational Biology and Bioinformatics. His lab studies the genetic and evolutionary processes that give rise to differences among individuals, populations, and closely related species. Current work focuses primarily on the evolutionary roots of human adaptations and variation in disease susceptibility. Before joining Columbia in 2013, Dr. Sella was a faculty member at Hebrew University in Israel and a visiting professor at the University of Chicago and Stanford University.



Inigo Martincorena, PhD

Group Leader, Martincorena Group, Sanger Institute

Inigo Martincorena completed his studies in biology and biochemistry in Spain in 2007, followed by a PhD at the University of Cambridge in 2008–2012 in evolutionary genomics, and a postdoc at the Sanger Institute in 2013–2016. He has been a group leader at the Sanger Institute (Cambridge) since 2016. His work over the last 10 years has focused on the study of somatic mutation and selection in normal tissues.

Somatic Mutation and Clonal Selection in Normal Tissues

ABSTRACT

Somatic mutations underpin cancer development and have been speculated to contribute to aging and a range of diseases. Over the past 15 years, the ability to sequence cancer genomes has transformed our understanding of the genetics and evolution of a wide range of cancers. However, owing to technical limitations, much less is known about the earliest steps of cancer and how cells in our tissues accumulate mutations during normal aging. In this talk, I will summarize our work over the past few years and present unpublished results on the extent and consequences of somatic mutation and clonal selection in normal tissues. I will also discuss how recent technological developments are opening new frontiers in this nascent and exciting field.

Sarah Aitken, MD, PhD

Group Leader and Consultant Histopathologist, University of Cambridge, UK

Sarah Aitken trained in medicine at the University of Edinburgh (UK), with an intercalated BMedSci in experimental pathology and a postgraduate MSc in translational medicine. She undertook a mixed experimental-computational PhD at the CRUK Cambridge Institute (Cambridge, UK), followed by an EMBO Fellowship in bioinformatics at the IRB Barcelona (Spain) with Professor Nuria Lopez-Bigas. She completed her clinical residency in histopathology in Cambridge as an NIHR Academic Clinical Fellow and subsequently NIHR Clinical Lecturer. Dr. Aitken started her independent research group at the University of Cambridge (UK) in September 2021 and is also a consultant pathologist (attending, board certified) at Addenbrooke's Hospital, Cambridge (UK).

Disentangling Strand Interactions in DNA Damage and Repair

ABSTRACT

Cancers arise through the acquisition of oncogenic mutations and grow by clonal expansion. Here we reveal that most mutagenic DNA lesions are not resolved into a mutated DNA base pair within a single cell cycle. Instead, DNA lesions segregate, unrepaired, into daughter cells for multiple cell generations, resulting in the chromosome-scale phasing of subsequent mutations. Furthermore, DNA replication across these persisting lesions can produce multiple alternative alleles in successive cell divisions, thereby generating both multiallelic and combinatorial genetic diversity. We exploit our discovery of lesion segregation to reveal how strand-asymmetric processes such as replication, transcription, and protein interactions shape DNA damage and repair. We demonstrate that replication over lesions produces almost identical collateral mutagenesis on the leading and lagging strands. In transcription, the triggering of repair is stochastic with frequent lesion bypass by RNA-polymerase. Finally, by defining multiallelic variation patterns we demonstrate that, under certain circumstances, nucleotide excision repair is mutagenic. These results provide insight into how strand asymmetric mechanisms underlie the formation, tolerance, and repair of DNA damage and thus shape cancer genome evolution.

Anjali Hinch, PhD

Group Leader, Dunn School of Pathology, University of Oxford, UK

Anjali Hinch is a Wellcome Trust and Royal Society Sir Henry Dale Fellow and a winner of the Wellcome-Beit Prize. She worked as a strategist at Goldman Sachs in New York prior to doing genetics research and has graduate and undergraduate degrees from the University of Oxford (Oxford, UK), Massachusetts Institute of Technology (Cambridge, USA), and IIT Bombay (Mumbai, India).

Dr. Hinch's lab uses a combination of computational and experimental approaches toward understanding how genetic changes take place in human eggs and sperm leading to modifications that may be passed down to future generations. Her lab studies the impacts of these changes on human health and the evolution of species.

How Our Genomic Landscape Is Reshaped in the Making of Eggs and Sperm

ABSTRACT

The DNA sequence in each egg and sperm is unique, a mosaic created by the shuffling of chromosomes by the process of recombination. At a cellular scale, recombination is essential for producing healthy eggs and sperm. At the scale of populations, recombination generates genetic diversity that underpins the evolution of species. In this talk, I will show how our understanding of the biological mechanisms underlying recombination has been powered by combining human genetic data and genomic experiments from these opposing scales.

Jennifer Phillips-Cremins, PhD

Associate Professor and Dean's Faculty Fellow in Engineering and Medicine; Department of Genetics–Perelman School of Medicine; Department of Bioengineering–School of Engineering and Applied Sciences, University of Pennsylvania

Jennifer Phillips-Cremins is an associate professor and Dean's Faculty Fellow in Engineering and Medicine at the University of Pennsylvania (UPenn) with primary appointments in the Departments of Genetics and Bioengineering. Dr. Cremins obtained her PhD in biomedical engineering from the Georgia Institute of Technology, in the laboratory of Andres Garcia. She conducted a multidisciplinary postdoc in the laboratories of Job Dekker and Victor Corces. Dr. Cremins now runs the Laboratory of Chromatin and Spatial Neurobiology at UPenn. Her primary research interests lie in understanding the long-range chromatin architecture mechanisms that govern neural specification and synaptic plasticity in healthy neurons and how chromatin-synapse communication is dysregulated in neurodevelopmental and neurodegenerative diseases. She has been selected as a 2014 New York Stem Cell Foundation Robertson Investigator, a 2015 Albert P. Sloan Foundation Fellow, a 2016 and 2018 Kavli Frontiers of Science Fellow, 2015, an NIH Director's New Innovator Awardee, a 2020 NSF CAREER Awardee, a 2020 CZI Neurodegenerative Disease Pairs Awardee, and the 2022 ISSCR Susan B. Lim Outstanding New Investigator Award and as a recipient of the 2021 NIH Pioneer Award.

Dissecting the 3D Genome's Structure-Function Relationship in the Mammalian Brain

ABSTRACT

The Cremins lab aims to understand how chromatin works through long-range physical folding mechanisms to influence neuronal specification and enduring changes in synaptic plasticity in normal neurophysiology and in neurological disorders. We pursue a multidisciplinary approach integrating data across biological scales in the brain, including molecular Chromosome-Conformation-Capture sequencing technologies, single-cell imaging, optogenetics, genome engineering, and induced pluripotent stem cell differentiation to neurons/organoids. At the lab's inception, it was unclear how genomes were folded in the mammalian neurodevelopment below the resolution of a Megabase, and whether and how higher-order structure could deterministically influence genome function. We have developed and applied new molecular and computational technologies to elucidate chromatin folding patterns at kilobase-resolution genome-wide, thus discovering that long-range looping interactions in cis and inter-chromosomal interactions in trans change substantially during neural lineage commitment, somatic cell reprogramming, and activation of post-mitotic neural circuits and in neurological disorders. We have demonstrated that cohesin-mediated loops are necessary for the establishment of new gene expression programs in post-mitotic neurons, including the upregulation of genes encoding axon guidance, dendritic spine morphology, and synaptic plasticity during neuron maturation in vivo as well as activity-dependent transcription during neural stimulation in vitro. We have also identified cohesin-mediated loops anchored by

divergently oriented CTCF binding sites that are necessary and sufficient for the firing efficiency and localization of human replication origins during S phase reentry after mitosis. Using fragile X syndrome as a natural perturbation, we have uncovered BREACHes (Beacons of Repeat Expansion Anchored by Contacting Heterochromatin)—rare inter-chromosomal interactions connecting heterochromatinized synaptic genes susceptible to repeat instability. Our work provides early insights into the genome's structure-function relationship during the establishment of new transcription, repeat instability, and replication patterns when mammalian cells transition states in the healthy and diseased brain.

Vijay G. Sankaran, MD, PhD

Jan Ellis Paradise, MD Professor of Pediatrics at Harvard Medical School; Investigator, Howard Hughes Medical Institute; Attending Physician, Dana-Farber/Boston Children's Cancer and Blood Disorders Center; Associate Member, Broad Institute

Vijay Sankaran and his lab seek to understand the influence of human genetic variation on blood and immune cell production in health and disease. Their work has resulted in a number of therapies for blood diseases, including work that led to the development of Casgevy for sickle cell disease and beta-thalassemia. Dr. Sankaran has received a number of awards for his work, including the 2019 Seldin-Smith Award for Pioneering Research from the American Society of Clinical Investigation, the 2022 E. Mead Johnson Award from the Society for Pediatric Research, and the 2024 Trailblazer Prize from the Foundation for the National Institutes of Health. He is an elected member of the Association of American Physicians and the American Society of Clinical Investigation.

Genetic Influences on Human Hematopoiesis

ABSTRACT

In this presentation, I will discuss work from our laboratory exploring how naturally occurring genetic variation in humans can alter the process of blood and immune cell production, or hematopoiesis, in health and disease. I will illustrate how integrating human genetic studies with diverse functional methods offers profound mechanistic insights, enhancing our understanding of hematopoiesis. Additionally, I will discuss recent work exploring how variable expressivity can arise in blood disorders at both a genetic and mechanistic level.

Adrianna San Roman, PhD

Assistant Professor of Molecular Genetics and Microbiology, Duke University

Adrianna San Roman is an assistant professor of molecular genetics and microbiology at Duke University. In 2009, she received a BA in biology from Williams College, followed by a PhD in developmental and regenerative biology at Harvard University in 2015. For her doctoral research with Dr. Ramesh Shivdasani, she defined key transcription factors that partner to regulate gene expression during stem cell self-renewal and differentiation in the mouse intestinal epithelium. Dr. San Roman subsequently pursued postdoctoral training in human genetics at the Whitehead Institute. Working with Dr. David Page, she gained novel insights into the regulation of the active and inactive X chromosomes and quantified the contributions of the sex chromosomes across the genome. Dr. San Roman established her lab at Duke University in January 2024 with the goal of elucidating molecular mechanisms of sex differences in human biology.

Genomic Contributions to Sex Differences in Human Biology

ABSTRACT

Biomedical research has historically ignored the role that sex differences play in health and disease, resulting in disparities in both healthcare and our understanding of the basic human biology that drives medical advancements. However, there is an increasing appreciation for differences between sexes that manifest across the lifespan and in disease. The biological origins of these sex differences are poorly understood, even though sex chromosome constitution—the number of X and Y chromosomes—is the largest source of genetic variation in the human population. To understand the cellular impacts of this variation, we sampled blood and skin from individuals with a range of sex chromosome constitutions—from one to four copies of the X and zero to four copies of the Y. Quantitative analyses of gene expression across these cells identified X chromosome genes that may mediate phenotypic effects of X chromosome copy number and revealed a novel gene regulatory mechanism through which the inactive X modulates transcription from the active X. More broadly, we found that thousands of autosomal genes cell-autonomously respond to X and/or Y chromosome copy number, and that a pair of transcription factors on the sex chromosomes is a key driver of these gene expression programs. These results provide a foundation for future work to elucidate molecular mechanisms linking genomic and phenotypic sex differences in humans.

Yun Song, PhD

Professor, University of California, Berkeley

Yun Song was originally trained in mathematics and theoretical physics, with degrees from MIT and Stanford University. He transitioned into population genetics and computational biology during his postdoc at the University of Oxford and the University of California, Davis; and he has been carrying out interdisciplinary research on diverse bio-related problems over the past two decades. He has been on the faculty at the University of California, Berkeley, since 2007—in the Department of Electrical Engineering and Computer Sciences and the Department of Statistics. His awards and honors include the NIH Pathway to Independence Award (K99/R00), NSF CAREER Award, Alfred P. Sloan Research Fellowship, Packard Fellowship for Science and Engineering, and Chan Zuckerberg Biohub Investigator Award.

Learning Complex Functional Constraints in Proteins and Non-coding DNA

ABSTRACT

Accurately predicting the impact of genetic variants in the human genome has several important applications, including improving disease diagnosis, advancing our understanding of gene regulation, and elucidating the genetic architecture of complex traits. In this talk, I will present my lab's recent work on developing powerful probabilistic models of biological sequences to better understand the functional and evolutionary constraints in proteins and non-coding DNA.

Oliver Stegle, PhD

Group Leader, European Molecular Biology Laboratories; Head, Computational Genomics and Systems Genetics Division, German Cancer Research Center

Oliver Stegle is head of the Computational Genomics and Systems Genetics Division at the German Cancer Research Center (DKFZ) and group leader at EMBL in Heidelberg, Germany. His laboratory is developing and applying statistical and machine learning methods for deciphering molecular variation across individuals, space, and time. He coordinates the German Human-Genome-Phenome Archive, the health program of the European Laboratory for Learning and Intelligent Systems and is an ERC investigator.

Advancing Human Genetics through Machine Learning and Single-Cell Sequencing

ABSTRACT

Understanding the impact of genetic variants on human traits remains a hallmark question of modern human genetics. In this talk I will discuss how computational innovations combined with new molecular readouts can help address this question. In the first part of my talk, I will describe new approaches based on machine learning to harness rare and private genetic variants that can be assayed in cohorts such as the UK Biobank, providing access to whole-genome sequencing data in hundreds of thousands of individuals. The model I describe allows to effectively aggregate genetic information across rare and private variants by tying together genetic variants with information on variant annotation, thereby boosting power for discovery and phenotype prediction. In the second part of the talk, I will discuss how molecular traits and in particular single-cell readouts can be leveraged to assay the consequences of genetic factors on gene expression, cellular programs, and cell states. I will describe applications of these methods to population-scale resources from iPSC-derived models and provide an outlook on upcoming single-cell data from thousands of individuals in the UK Biobank.

Jonathan Pritchard, PhD

Bing Professor of Population Studies, Department of Genetics and Biology, Stanford University

Jonathan Pritchard is a professor of biology and genetics at Stanford University. He grew up mainly in England and studied at Penn State, Stanford, and Oxford before joining the faculty of the University of Chicago in 2001. He returned to Stanford to take his current position in 2013. His lab has done wide-ranging research on using genetics to study human population structure, history, and adaptation, as well as on understanding the mechanisms that link genetic variation to variation in gene regulation and complex traits. One of his early contributions was the Structure algorithm for using genetic data to infer population structure and personal ancestry. His current work focuses on the genetic basis of complex traits in humans and on genetic approaches to studying human population history and adaptation.

Integrating Genetic Association Data with Cellular Perturbations to Study the Molecular Pathways of Human Trait Variation

ABSTRACT

In this talk I will argue that a major current challenge in human genetics is to move from a paradigm that is mainly focused on gene mapping toward a paradigm that is focused on developing causal mechanistic models that link genes to molecular mechanisms to disease. We are starting to see exciting work in this area from a number of groups around the world, including at this conference. Here I will describe our work on new methods for integrating genetic association data with experimental perturbations to infer causal pathways. As an sample model system, we combine UK Biobank data from red blood cell traits, such as cellular hemoglobin, with data from a genome-scale Perturb-seq study in a leukemia cell line. I will describe how integrated analysis of these data sets allows us to infer causal pathways from genetic effects to gene regulatory programs to organism-level traits. More broadly, we propose that this type of data integration will be crucial for developing genome-scale causal models of human phenotypic variation.

Michel Georges, DVM

Full Professor, Faculty of Veterinary Medicine, Université of Liège; Laboratory Head, Unit of Animal Genomics, GIGA Institute

Michel Georges is professor in genetics and genomics at the University of Liège, and the head of the Unit of Animal Genomics. He played an instrumental role in establishing the GIGA Institute and as its director from 2016 to 2023. He devoted his career to the development of genomic tools for the identification of genes and mutations underlying complex traits of agronomic and medical importance. With his collaborators, he cloned the double-muscling gene, discovered polar overdominance, identified the first mammalian phenotype due to perturbed microRNA regulation, discovered a novel CNV generating mechanism underlying color-sidedness, identified several QTN at single-base pair resolution, and demonstrated that ABO genotype alters the microbiome by affecting GalNAC concentrations. His laboratory has contributed to multiple GWAS and post-GWAS studies for inflammatory bowel disease. Dr. Georges was awarded the Wolf Prize in Agriculture in 2007, the Francqui Prize in Biomedical Sciences in 2008, and an ERC advanced grant in 2013. He has been a member of NAS since 2013. He obtained his DVM from the University of Liège (1983), an MSc in molecular biology from the Free University of Brussels (1985), and his habilitation from the University of Liège (1991). From 1989 to 1993 he was senior scientist at Genmark and adjunct professor in human genetics in Salt Lake City, Utah.

eQTL Matching IBD Risk Loci: Paucity or Plethora?

ABSTRACT

The majority of GWAS-identified risk loci for common complex diseases are driven by regulatory variants. Cis-eQTL analyses in disease relevant cell types is a common approach to identify the underlying causal genes. However, with the available data sets, “matching eQTL” have remained elusive for a majority of risk loci (~2/3). We will describe the identification of colocalized eQTL for 483 protein coding genes in 140 of 206 risk loci for inflammatory bowel disease (IBD) with a newly generated eQTL dataset from 26 FACS/MACS-sorted circulating immune cell populations; and 43 scRNA-Seq-sorted intestinal cell populations in the ileum, colon, and rectum. Several risk loci likely encompass multiple causative genes extending the notion of multigenicity within risk loci. The list of positional candidate genes includes over 200 genes connected to innate/adaptive immunity or epithelial homeostasis (including CFTR); four genes that are associated with monogenic forms of IBD; 69 that—upon perturbation—affect colitis susceptibility in mice; and 10 that are or have been used to treat IBD. We combine eQTL information with expression data in affected samples to identify new repurposing candidates.

Nasa Sinnott-Armstrong, PhD

Assistant Professor of Computational Biology, Fred Hutchinson Cancer Center;
Assistant Professor of Genome Sciences, University of Washington

Nasa Sinnott-Armstrong is an assistant professor of computational biology at Fred Hutchinson Cancer Center and genome sciences at University of Washington. Their research focuses on understanding genotype-phenotype mapping, including modeling of gene-environment interactions. Dr. Sinnott-Armstrong uses both computational and experimental methods to address questions around causal mechanism and heterogeneity between individuals, with a focus on Lyme disease and hormone replacement therapy. Before joining the Hutch, they earned a PhD in population and statistical genetics at Stanford University, researching methods to understand the molecular and cellular basis of complex traits, as well as gene-environment interactions of cardiometabolic and pulmonary traits. Finally, Dr. Sinnott-Armstrong is passionate about science-art collaboration and community centered science as ways to inform, disseminate, and ground scientific inquiry.

Complex Dynamics Underpin Gene-Environment Interactions

ABSTRACT

Human disease risk is conferred primarily through a complex set of environmental and behavioral exposures throughout life. These exposures are shaped by individual genetic factors, as well as by large-scale environmental forces like climate change, in a highly structured fashion. One common exposure of interest is to pathogenic agents. Here, we focus on the tick-borne illness Lyme disease. We previously conducted a genome-wide association study (GWAS) for Lyme disease and discovered a single strong association with a common missense variant in the gene SCGB1D2. SCGB1D2 is part of the mammal-specific secretoglobin family of short, secreted proteins with no known function. In vitro experiments suggest that SCGB1D2, and possibly other secretoglobins, form a novel mammalian bacterial defense system. We also discovered novel associations between Lyme disease and diseases of the female reproductive tract, informed by community partners, patients, and evidence of tissue-specific function of SCGB1D2. Together, these findings highlight the relationship between genetic and environmental contributions to disease risk. Overall, combined use of statistical and experimental methods, community-engaged research approaches, and computational modeling can improve our understanding of human disease.

Miriam Udler, MD, PhD

Assistant Professor, Harvard Medical School; Attending in Endocrinology, Massachusetts General Hospital; Director, MGH Diabetes Genetics Clinic; Investigator, MGH Center for Genomic Medicine; Associate Member, Broad Institute

Miriam S. Udler is an endocrinologist at Massachusetts General Hospital (MGH) and an assistant professor at Harvard Medical School. Dr. Udler is the founding director of the MGH Diabetes Genetics Clinic, which provides genetic testing, counseling, and management to patients with monogenic forms of diabetes. She is also an associate member of the Broad Institute of MIT and Harvard.

Dr. Udler's lab at MGH and the Broad Institute studies the genetic contribution to metabolic diseases and clinical applications of genomic data, including using genetics to identify atypical forms of diabetes and to dissect disease heterogeneity. Dr. Udler is a principal investigator of the NIDDK U54 Rare and Atypical Diabetes Network (RADIANT) study and the NIDDK U01 Heterogeneity of Diabetes study and a recipient of the Doris Duke Foundation Clinical Scientist Development Award. She is also an investigator in the Accelerating Medicines Partnership Program Type 2 Diabetes (AMP T2D) and the Polygenic RIsk Methods in Diverse Populations (PRIMED) Consortium. Additionally, she is a member of the Clinical Genome (ClinGen) Expert Panel for Monogenic Diabetes and lectures regularly on clinical genetics of diabetes.

Dr. Udler received an AB degree cum laude in applied mathematics from Harvard College, an MPhil and PhD in genetic epidemiology from University of Cambridge, and an MD degree from University of Massachusetts Medical School. She completed a residency in internal medicine at Mount Sinai Medical Center and an endocrinology fellowship at MGH.

Process-Specific Polygenic Scores for Insights into Disease Mechanisms

ABSTRACT

The vast discovery of genetic contribution to type 2 diabetes offers an important opportunity to gain insight into disease mechanisms and potentially identify disease endotypes. Dr. Udler and her team have developed approaches to perform physiologically informed clustering of disease genetic variation. These genetic clusters can be used to generate process-specific polygenic scores, which inform on disease heterogeneity as well as cellular and molecular phenotypes. Dr. Udler will describe current progress in the development of type 2 diabetes genetic clusters and process-specific polygenic scores, with a focus on potential clinical applications.

Gil McVean, PhD

Professor of Statistical Genetics, University of Oxford, UK; Fellow, Linacre College, Oxford, UK; Co-founder and President, Genomics plc

Gil McVean is professor of statistics genetics at the University of Oxford and president and co-founder of Genomics plc. He has a background in statistical and population genetics and played key roles in the International HapMap and 1000 Genomes Projects, as well as leading research into human genetic diversity; the processes of mutation, selection, and recombination; and the genetic basis of complex traits. He was the founding director of Oxford University's Big Data Institute, is a fellow of both the Royal Society and the Academy of Medical Sciences and is the incoming president of the Genetics Society.

Longitudinal Modeling of Human Disease Comorbidities and the Implications for Our Understanding of Risk and Pathology

ABSTRACT

The analysis of longitudinal data from electronic health records (EHRs) has the potential to improve clinical diagnoses and enable personalized medicine, motivating efforts to identify disease subtypes from patient comorbidity information. Recently, we developed an age-dependent topic-modeling (ATM) method that provides a low-rank representation of longitudinal records of hundreds of distinct diseases in large EHR datasets and applied it to about 300,000 individuals from UK Biobank and over 200,000 individuals from the All of Us program. We defined subtypes of the 52 heterogeneous diseases (i.e., those that occur in more than one topic) based on their comorbidity profiles, finding differential genome-wide and locus-specific genetic risk profiles for at least 18 of these. Such stratification improves understanding of patient heterogeneity, leading to better identification of genetic risk, characterization of pathological pathways, and the discovery of new targets for medicines.

Nancy Cox, PhD

Mary Phillips Edmonds Gray Professor and Director, Vanderbilt Genetics Institute, Vanderbilt University Medical Center

Nancy Cox is a quantitative human geneticist with a long-standing research program on the genetic basis of human diseases and related quantitative traits. She earned a BS in biology at the University of Notre Dame in 1978 and a PhD in human genetics at Yale University in 1982; and she was a postdoc at Washington University (1982–85) and the University of Pennsylvania (1985–87) before joining the faculty at the University of Chicago, where she worked until 2015, when she joined Vanderbilt as the inaugural director of the Vanderbilt Genetics Institute and the Mary Phillips Edmonds Gray Professor. Her current research is focused on integrating information on genome variation and genome function and electronic health records in the biobank at Vanderbilt, BioVU, with DNA samples on over 335,000 people.

How the Genetic Component to Laboratory Measurements in Medicine Can Degrade Outcomes and Add to Health Inequities: What Should We Do When the Genetics Adds Noise, Not Signal?

ABSTRACT

Like many quantitative biomarkers, about two-thirds of all commonly ordered laboratory values are significantly heritable. Lab values that have come into use more recently may be relatively specific indicators of disease risk, and these more recent labs may also be in the direct causal pathways to disease. LDL cholesterol is an example of this newer type of biomarker. The genetic component of LDL is predictive of atherosclerosis and cardiovascular disease; and the non-genetic factors that influence LDL levels are also predictive of these outcomes, because high LDL levels are causal related to the development of atherosclerosis and cardiovascular disease. Many of the older, long-established biomarkers are also quite heritable, but these biomarkers tend to be more non-specific indicators of a departure from healthy homeostasis. White blood cell counts may rise when people have infections and may fall when people take drugs that are toxic to white cells. But the genetic variation that influences interindividual variability in white cell counts does not predict whose counts will fall the most when they have chemotherapy, or whose counts will rise the most when they have appendicitis. For white cell counts and many other of the more traditional biomarkers, genetic variation contributing to the variability in these everyday lab values will sometimes be misleading for clinical decision-making. We will discuss some of the medical consequences of these challenges and show how removing the genetic component to the lab value can improve the bias and precision of the risk to disease that the lab values are intended to measure.

Simon Myers, PhD

Professor of Mathematical Genomics, University of Oxford

Simon Myers is a member of the Department of Statistics, and St John's College, in Oxford, where he leads a group studying genetic variation in humans and other species, using a combination of statistical analysis and experimental studies of meiosis. Having obtained a degree in mathematics, he undertook his graduate studies in Oxford, moving to Harvard and the Broad Institute of MIT/Harvard for two years as a research fellow, before his return to Oxford. He has extensively studied the process of recombination, in early work finding the first sequence motif positioning mammalian recombination hotspots. Later, he identified a gene, PRDM9, that binds this motif, in work that has since broadened to encompass speciation, fertility, and how chromosomes pair during sexual reproduction. Alongside this, his group has led the development of approaches to understand population structure at fine geographic scales, and more recently to build genome-wide genealogies from DNA sequences, and this strand has identified the timings and impacts on our DNA of many human migration events. Present projects include work leveraging single-cell data to study how chromatin modifications and meiotic events such as synapsis control gene expression, as well as seeking to understand human trait evolution using genealogies. Dr. Myers previously received the Genetics Society Balfour Prize, the Royal Society Francis Crick Medal and Lecture,, and the Weldon Prize for his work.

Why Does Genomic Trait Prediction Translate Poorly across Human Ancestries?

ABSTRACT

Many thousands of mutations in humans across the genome have been associated with hundreds of traits as a result of modern genome-wide association studies (GWAS), with the majority so far conducted in populations of European ancestry. These have revealed that quantitative phenotypes and disease risk are influenced by the additive, small effects of large numbers of mutations genome-wide. GWAS have been widely used to generate polygenic scores for trait predictions. However, they typically perform poorly in populations distinct from those used for the GWAS: for example, prediction from European-based polygenic scores is typically >4-fold less accurate in African-ancestry individuals. This is likely due to differing variation patterns across populations, and differing underlying causal effect sizes due to gene-gene or gene-environment interactions, but estimates regarding the relative importance of these factors have varied. Using a general, robust, and conceptually simple approach leveraging ancestry along the genome, we infer that true effect sizes across 53 traits within UK Biobank do not vary significantly between European and African-ancestry individuals (correlation 0.98+/-0.07), ruling out differing effect sizes as a significant non-portability driver. Instead, I will show that genealogies we have constructed across the genome offer a powerful tool that reveals non-portability as one consequence of widespread natural selection acting to eliminate older functional variants that are shared among groups, alongside clustering of causal variants. As a

conservative estimate, the variance in—for example—human height is at least 3.7-fold reduced by this force. These results imply that even very weak genetic risk factors operate similarly at the biological level among human populations, so that findings from one human population are directly applicable, for most phenotypes, across all groups. In future, portability can be optimized by fine-mapping using tools that leverage young ages and shared effect sizes among clustered causal mutations.