February 2025

Dear Colleagues,

I'm delighted to share the Columbia Precision Medicine Initiative (CPMI) progress over the last year. This is my first annual newsletter following a year of transitions, after assuming the Directorship from CPMI's inaugural leader, Dr. Tom Maniatis. In the enclosed newsletter, we outline two new objectives: to promote the clinical implementation of precision medicine and to build bioinformatics infrastructure to support genomic research. In addition to these new objectives, CPMI has continued its commitments to education and early-stage research through its continued support of courses, lectures, and symposia.

I hope you enjoy reading about these activities. None of these enterprises would be possible without Roy and Diana Vagelos' generous support and vision of scientific progress and clinical improvements that leverage precision medicine.

Warm wishes,

Christine Kim Garcia, MD, PhD
Frode Jensen Professor of Medicine,
Vagelos College of Physicians and Surgeons
Chief, Pulmonary, Allergy, and Critical Care Medicine
Director, Columbia Precision Medicine Initiative

Overview
As charged by Dr. Katrina Armstrong, EVP of CUIMC and Interim President of Columbia University, the Columbia Precision Medicine Initiative (CPMI) has focused on creating a cohesive program for implementing genomic medicine, producing infrastructure to support genetic bioinformatic research, and continuing its commitment to education and early-stage research. The overarching objective of CPMI is to prepare for the eventuality of routine genetic sequencing of patients. Thus, the following plan prioritizes activities to support discovery science and precision medicine implementation in clinical practice.

A Plan for Prioritized Activities
A plan for prioritization of activities was developed over the past year through meetings with a newly established CPMI Advisory committee (Appendix 1), which includes broad representation from across the Columbia campus. The committee formed two working groups to focus on clinical implementation and research infrastructure working groups, led by **Drs. Benjamin Herzberg** and **Joshua Milner**, respectively. Here, we report on the progress made in two key areas, Clinical Implementation of Precision Medicine and Research Genomic Infrastructure to Support Precision Medicine, as recommended by these groups.

Clinical Implementation of Precision Medicine
All patient care activities that involve the testing of inherited and somatic genetic variants, the return of results to patients, and using these results to provide personalized medical care, encompass the clinical implementation of precision medicine. As described in this announcement from Dr. Armstrong, a key recommendation of a Columbia genomic medicine taskforce was the creation of a new leadership role, a Clinical Genomics Officer (CGO), to coordinate and oversee the implementation of clinical genomics across all clinical services. CPMI assisted in identifying candidates and recruiting the top candidate, **Jennifer Posey, MD, PhD**. More information about Dr. Posey is presented below and additional announcements regarding her arrival will be forthcoming.

The CPMI clinical implementation working group has also identified **genetic testing in cancer as a key priority**, as highly effective precision medicine therapies can attack specific targets or vulnerabilities of cancer cells with minimal effect on normal cells. Over the last two decades, genomic testing in oncology has experienced a revolution. Now, most assays are based on next-generation sequencing (NGS). Targeted panels (up to about 50 genes) report only a limited set of actionable genes. In contrast, comprehensive panels of up to 500-600 genes report variants that suggest effective therapies, support new drug development, identify novel targets, allow for patient enrollment in clinical trials, and suggest a host of prognostic markers. Whole-exome and other advanced tests are still largely beyond standard-of-care for most oncology testing, although they may be useful for immunotherapy programs focused on neoantigens and novel targets.

The CPMI clinical implementation group found that the percentage of the Columbia patient population who undergoes oncologic genetic testing is small. Genomic data, when it is

ordered and returned, typically comes from external vendors whose reports are uploaded into the EPIC electronic medical record system in PDF format and are not easily searchable. The group reviewed the Columbia Pathology molecular workflows, offerings, and cost structure and compared these to external vendor abilities, projections, and potential uses. In collaboration with NYP partners, they identified opportunities to input genetic testing ordering through EPIC in a vendor-agnostic fashion and generate workflows to more readily access resulting genomic data.

To date, the CPMI clinical implementation group and **Dr. Herzberg** have worked with the Columbia Pathology molecular diagnostics team to explore the use of its testing platform, the Columbia Combined Cancer Panel (CCCP) that uses next-generation sequencing to query 586 genes. They are working with NYP partners to facilitate ease of ordering genetic testing, disseminating results in EPIC, and storing clinical genetic data to be integrated into future research platforms.

Research Genomic Infrastructure to support Precision Medicine
A second key objective identified by the CPMI Advisory Committee was to develop the infrastructure to support discovery science by creating a Columbia genomic data sharing platform. Over the last year, CPMI has assumed oversight of the Genomic & Bioinformatics Analysis Resource (GenBAR), which manages genomics data that was previously generated by the Institute for Genomic Medicine (IGM). In 2024, GenBAR migrated petabytes of genomic data from an on-premises data center to Amazon Web Services (AWS) cloud computing. GenBAR has also migrated to the cloud the bioinformatic tool Analysis Tool for Annotated variants (ATAV), which has been used by many Columbia researchers to interrogate research genetic data and perform case/control studies.

By working with the Office of Research, CPMI has created protocols by which genomic data from patients who consented to participate in the Columbia University Biobank (CUB) can be aggregated and combined with >50,000 existing whole exome and whole genome sequencing data. With the leadership of **Dr. Joshua Motelow**, GenBAR has designed harmonized exome and genome pipelines that utilize AWS HealthOmics and WARP (WDL Analysis Research Pipelines) tools developed by the Broad Institute Data Sciences Platform. These open-source genomic analysis pipeline tools have been used by large consortia. As of January 2025, >1800 archived whole exome sequencing datasets have been reprocessed, aligned to an updated reference sequence (GRCh38), and aggregated in a new joint called file. Similar work in reprocessing archived whole genome sequencing data is in progress.

The CPMI research implementation working group discussed the status and critical needs for state-of-the-art genomic infrastructure. Major strengths include Columbia investigator expertise, large collections of legacy data, strong partnership with the Columbia University Biobank (CUB), availability of the Pakistani Genomic Resource, partnership with the New York Genome Center (NYGC), ongoing large-scale genomics projects, and broad interest in centralizing resources and access. Key challenges identified include inadequate legacy

infrastructure, limited technical and developer staffing, the persistence of siloed data, and the need for coordination across NYP/Columbia during system development. Major opportunities noted were the accessibility to substantial tissue banks from normal, diseased, transplant, and rare disease cohorts, as well as opportunities to use AI-driven analytic tools.

Overall, there was a high level of enthusiasm for creating a platform to centralize access to and integrate clinical and research genomics information across the institution. Such a platform could ultimately merge different types of 'omic data (genomic, transcriptomic, proteomic, radiomic, exposomic) with clinical phenotypic data to produce a cutting-edge toolbox to promote real-time patient-relevant discoveries, advance phenotype/genotype correlations, and explore novel methods for discovery. CPMI will continue to work with key partners, including NYGC, NYP, Columbia IT, CUB leadership, and the Dean's office.

Education:
CPMI has supported education at multiple levels over the last year. It supported an undergraduate course entitled "Precision Medicine: Ethical, biological and societal implications" in the Fall of 2024, directed by **Drs. Sam Sternberg** (Biochemistry), **Rachel Adams** (English), and **Maya Sabatello** (Medicine and Ethics). CPMI and the Irving Institute for Clinical and Translational Research hosted a Precision Medicine Scholars' Day, with invited speakers and poster presentations to foster collaborative science. In addition, CPMI supported its 8th academic conference, Advances in Precision Medicine: Frontiers in Human Genetics, chaired by **Dr. Molly Przeworksi** (Biological Sciences and Systems Biology), to bring together international and national leaders of genomic science. The conference was a great success. Videos of the talks can be found on the CPMI website.

Welcome Dr. Jennifer Posey!
Jennifer Posey, MD, PhD, FACMG, will join Columbia University Irving Medical Center (CGO-CUIMC) as the chief of the division of clinical genetics in the department of pediatrics and the inaugural chief genomics officer. As CGO-CUIMC, Dr. Posey will develop a new medical genetics and genomics vision and lead the clinical implementation arm of the Precision Medicine Initiative.

Dr. Posey joins us from Baylor College of Medicine in the Department of Molecular and Human Genetics. She has served as faculty and an attending physician since 2014 following her completion of fellowships in medical genetics and research genetics (with Dr. James Lupski) at Baylor and Residency in Medicine at Columbia. She is a graduate of the University of Texas, Austin and received her MD and PhD degrees at Baylor College of Medicine. She is the recipient of several prestigious awards, including the American Society for Human Genetics Early Career Award and the American Society for Clinical Investigation Young Physician-Scientist Award. Her impact has grown at an inspiring pace, as evidenced by her 257 listed Scopus publications and her current leadership roles as the Principal Investigator of the NIH Mendelian Genomics Research Center U01 and GREGoR

(Genomics Research to Elucidate the Genetics of Rare Diseases) consortium. She is widely recognized for her 2017 New England Journal of Medicine paper, Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation, which opened the door to the next wave of genetic disease explanation and etiology.

At Columbia, Dr. Posey will assume the leadership of the longstanding Division of Clinical Genetics in the Department of Pediatrics while broadening its reach by establishing a linked service in the Department of Medicine and serving as the inaugural CGO-CUIMC. Thus, Dr. Posey will be a leader in both the Departments of Pediatrics and Medicine and, in the role of CGO-CUIMC, will link to Columbia's Precision Medicine Initiative, The Department of Pathology and Cell Biology, the VP&S Dean's office, as well as New York-Presbyterian's Genomics Laboratory and Genomic Data Infrastructure.

Precision Medicine Awards:

Columbia Precision Medicine Initiative Roy and Diana Vagelos Pilot Awards
Two research teams at Columbia University have been awarded a 2024 Precision Medicine Pilot Grant. The Precision Medicine Pilot Grants underscore Columbia University's commitment to supporting diverse, cross-disciplinary research targeting the promise of precision medicine. Each group received a one-year $100,000 grant to support their research. The two projects are being led by principal investigators: Chaolin Zhang, PhD, Associate Professor of Systems Biology and Biochemistry and Molecular Biophysics; and Gamze Gursoy, PhD, Assistant Professor of Biomedical Informatics.

Harnessing pentatricopeptide repeat proteins for programmable RNA targeting
Investigators: Chaolin Zhang, PhD (Principal Investigator); Harris Wang, PhD
Numerous genetic diseases are caused by mutations that disrupt individual genes and could potentially be treated by correcting disease-causing mutations or modulating gene expression to restore the production of the functional protein. This research project explores the use of pentatricopeptide repeat proteins (PPRs) as a versatile tool for manipulating RNA, aiming to address genetic diseases. PPRs are like molecular machines that can be engineered to target specific RNA sequences in a programmable manner. We plan to decipher the "PPR code," a set of rules that govern how PPRs interact with RNA. By understanding this code, we hope to design custom PPRs (designer PPRs or dPPRs) capable of precisely targeting and modifying RNA, which could be a breakthrough for treating genetic disorders. We will conduct innovative and high-throughput experiments to unravel the PPR code and then test the engineered PPRs in real-life scenarios, particularly focusing on modifying RNA splicing to address a genetic condition known as spinal muscular atrophy. Success in this study could open new avenues for RNA-based therapies and contribute to advancements in precision medicine.

A deep learning-based approach for phenome-wide association studies

Investigators: Gamze Gursoy, PhD (Principal Investigator); David Knowles, PhD
*P*henome-wide Association Studies (PheWAS) is an approach examining the connections between genetic variants and a wide range of health-related traits and diseases. Unlike traditional genome-wide association studies (GWAS), which focus on single traits, PheWAS takes a broader view, exploring how a single genetic variant can impact multiple aspects of health—a phenomenon known as pleiotropy. PheWAS utilizes Electronic Health Records (EHR) data to streamline research, but it faces challenges such as phenotype classification, managing high-dimensional data, computational complexity, and sensitivity to rare variants. To overcome these hurdles, we propose innovative methods, including machine learning to create compact patient phenotype representations and advanced language models, particularly attention-based Transformers, to enhance the PheWAS methodology. By combining clinical knowledge with real-world data and leveraging embeddings and attention mechanisms, our research aims to provide profound insights into genetic influences on health and extend its applications beyond PheWAS to areas like disease risk assessment and early detection. We plan to train and test our approach using data from the UK Biobank and All of Us. Looking ahead, we aim to extend our methodology to include data from Columbia Medical Center by establishing a partnership with the biobank to conduct additional sequencing on the identified patients.

Columbia Precision Medicine Initiative Models of Human Disease Awards
Three teams of researchers have been awarded pilot grants to fund a diverse set of models of human disease for precision medicine research. The proposals reflect the high standard and the broad base of precision medicine research being conducted and conceived at Columbia. They include an innovative lagomorph model with the potential to become a standard gene therapy tool; a zebrafish Alzheimer disease model to study the role of the blood brain barrier; and a mouse model to investigate the molecular mechanisms contributing to mitochondrial eye diseases.

Precision Single Nucleotide Variant (SNV)-Mediated Ablation in an FDA-Compliant Humanized Lagomorph Model of Autosomal Dominant Disorder
Lead Investigator: Stephen Tsang, MD, PhD
The development of an FDA-compliant humanized lagomorph model, engineered with patient-specific DNA, marks a significant advancement in preclinical testing for CRISPR-based therapeutic editing. This large animal model provides an alternative to non-human primates for Investigational New Drug-enabling studies. Specifically designed with patient-specific genetic mutations, the humanized lagomorph is ideal for evaluating CRISPR therapies that precisely target patient DNA sequences. The model adheres to FDA guidelines for pharmacokinetic, toxicological, and biodistribution studies, ensuring regulatory compliance while delivering critical data on systemic toxicity and therapeutic efficacy. By accelerating the transition from preclinical studies to human trials, this engineered lagomorph model improves the translatability of therapeutic findings, creating a streamlined path toward the implementation of precision medicine.

Functional analyses of ADAMTS1 missense gene variant segregating in families with AD identified in the National Institute on Aging Alzheimer's disease family-based study (NIA-AD FBS)

Lead Investigator: Caghan Kizil, PhD

Alzheimer's disease (AD) is a debilitating neurodegenerative disorder characterized by cognitive decline and memory loss. Recent studies in large multiethnic AD cohorts in families have identified a rare missense variant in the ADAMTS1 gene, which segregates within families affected by AD in the National Institute on Aging Alzheimer's Disease Family Based Study. ADAMTS1 is crucial for extracellular matrix organization, angiogenesis, and tissue remodeling, and has been implicated in amyloid-beta degradation and blood-brain barrier (BBB) integrity—key factors in AD pathology. This proposal aims to investigate the functional impact of the ADAMTS1 variant using a zebrafish model, given their genetic similarity to humans and suitability for in vivo studies. The gene is 94% conserved in humans and zebrafish, the mutated residue is identical. We will generate zebrafish carrying the precise ADAMTS1 point mutation identified in humans to explore its effects on brain function, BBB integrity, and amyloid-beta regulation through molecular, histological, and omics analyses. By leveraging zebrafish's transparency and ease of genetic manipulation, we will conduct comprehensive in vivo imaging, behavioral assays, and transcriptomic, proteomic, and metabolomic profiling. Our research addresses the need for precision medicine approaches by focusing on ethnicity-dependent genetic variations and their contributions to AD. This study will provide insights into the mechanistic roles of ADAMTS1 in AD, identify potential therapeutic targets, and advance the development of tailored treatments for diverse populations. The anticipated outcomes include a deeper understanding of AD pathology and the establishment of zebrafish as a valuable model for studying genetic variants associated with neurodegenerative diseases.

A mouse model with defective mtDNA maintenance

Lead Investigator: Nan-Kai Wang, MD, PhD

This research focuses on the SSBP1 gene, essential for mitochondrial DNA (mtDNA) synthesis and protection. Mutations in this gene cause Optic atrophy-13 (OPA13), leading to vision loss due to optic atrophy and retinal degeneration. These mutations disrupt SSBP1 function, affecting mtDNA replication and causing mtDNA depletion. The study aims to determine if the disease mechanism is due to haploinsufficiency or a dominant-negative effect. Previous zebrafish studies were inconclusive, so mouse models are preferred for their similar eye anatomy to humans and genetic tool availability. Germline knockout of Ssbp1 results in embryonic lethality and reduced mtDNA in heterozygous mice, highlighting SSBP1's critical role. Therefore, knock-in mouse models are needed to understand OPA13 mechanisms. A mouse model with a specific Ssbp1 mutation will be created and studied for pre-clinical treatment exploration. This research aligns with precision medicine by tailoring treatments to the genetic makeup of the disease, potentially leading to therapies that improve mitochondrial function and prevent vision loss.

Appendix 1. Members of the Columbia Precision Medicine Initiative Advisory Committee

Katrina Armstrong, MD; Co-Chair, Dean and Interim President
Christine Kim Garcia, MD, PhD; Co-Chair, Director of CPMI

Benjamin Herzberg, MD; Oncology (Lead for the Clinical Implementation Working Group)
Joshua Milner, MD; Pediatric Allergy and Immunology (Lead for Research Working Group)

Additional Members:
Aliaa Abdelhakim, MD, PhD; Ophthalmology
Alejandra Aguirre, DrPH, MPH; Community Health Initiative in Medical Humanities and Ethics
Paul Appelbaum; Psychiatry; Medicine and Law; Division of Law, Ethics, and Psychiatry
Amanda Bergner, MS, LCGC; Genetic Counseling
Barry Fine, MD; Cardiology
Peter Fleischut, MD; NewYork-Presbyterian Hospital
Matthew Harms, MD; Neurology
Eldad Hod, MD; Pathology
Steven Kernie, MD; NewYork-Presbyterian Hospital
Krzyzstof Kiryluk, MD; Nephrology
Steven Kushner, MD; Psychiatry
Aimee Payne, MD, PhD; Dermatology
Elaine Pereira, MD; Pediatric Genetics
Maya Sabatello, LLB, PhD; Medical Ethics, Medicine
Soumitra Sengupta, PhD; Department of Biomedical Informatics
Mary Beth Terry, PhD; Epidemiology
Meghna Trivedi, MD; Oncology
Badri Vardarajan, PhD; Neurology
Neil Vasan, MD, PhD; Oncology
Ronald Wapner, MD; Obstetrics-Gynecology
Jennifer Williamson; Office of Research