



New Study Explains Why Genetic Mutations Cause Disease in Some People but Not in Others

Large Genomic Datasets and CRISPR Offer New Insights into Disease Risk

NEW YORK, NY (August 20, 2018) – Researchers at the New York Genome Center (NYGC) and Columbia University have uncovered a molecular mechanism behind one of biology’s long-standing mysteries: why individuals carrying identical gene mutations for a disease end up having varying severity or symptoms of the disease. In this widely acknowledged but not well understood phenomenon, called variable penetrance, the severity of the effect of disease-causing variants differs among individuals who carry them.

[Reporting in the August 20 issue of *Nature Genetics*](#), the researchers provide evidence for modified penetrance, in which genetic variants that regulate gene activity modify the disease risk caused by protein-coding gene variants. The study links modified penetrance to specific diseases at the genome-wide level, which has exciting implications for future prediction of the severity of serious diseases such as cancer and autism spectrum disorder.

NYGC Core Faculty Member and Columbia University Department of Systems Biology Assistant Professor Dr. Tuuli Lappalainen led the study alongside post-doctoral research fellow Dr. Stephane Castel.

“Our findings suggest that a person’s disease risk is potentially determined by a combination of their regulatory and coding variants, and not just one or the other,” Dr. Lappalainen said. “Most previous studies have focused on either looking for coding variants or regulatory variants that affect disease in these individuals or potentially looking at common variants that could affect disease. We have merged these two fields into one clear hypothesis that uses data from both of them, which was fairly unheard of before.”

Variable penetrance has long posed a challenge for predicting the severity of a disease, even for diseases with a strong genetic association. Dr. Lappalainen and colleagues developed the modified penetrance hypothesis from their interest in the idea that gene variants that regulate the activation of genes could also play a role in modifying the penetrance of coding variants for the same gene.

As a first test of the modified penetrance hypothesis, the researchers conducted an analysis of data from the Genotype-Tissue Expression (GTEx) project, a large catalog of genetic variants that affect gene expression in humans, to evaluate the interactions of regulatory and coding variants in a human population without severe genetic diseases. They found an enrichment of combinations of regulatory and coding variants, called haplotypes, that act as protective against disease by decreasing the penetrance of coding variants associated with disease development. This finding was expected in the general population, Dr. Castel explained, as a result of natural selection removing damaging gene variants from the genome over time.

To test their hypothesis in a disease-specific population of patients, the researchers analyzed data from the National Institutes of Health’s The Cancer Genome Atlas (TCGA) and the Simons Simplex Collection, a permanent repository of genetic samples from 2,600 families, each of which has one child affected with an autism spectrum disorder, and unaffected parents and siblings. In the cancer patients and individuals with autism, they found an enrichment of haplotypes predicted to increase the penetrance of coding variants associated with cancer and autism spectrum disorder, respectively.

Finally, they designed an experiment using CRISPR/Cas9 genome editing technology to test the modified penetrance hypothesis with a coding variant that is known to be associated with a disease. They chose a coding variant associated with Birt-Hogg-Dubé Syndrome, a rare hereditary disease that increases the risk of certain types of tumors. They edited the SNP into a cell line on different haplotypes with a regulatory variant. The researchers were able to show that the regulatory variant indeed modified the effect of the coding disease-causing variant, consistent with expectations based on the large-scale data collections. This finding provides an important framework for scientists moving forward to experimentally test specific disease SNPs to determine if they could be affected by modified penetrance.

“Now that we have demonstrated a mechanism for modified penetrance, the long-term goal of the research is better prediction of whether an individual is going to have a disease using their genetic data by integrating the regulatory and coding variants,” Dr. Lappalainen said.

“In future, studies of the genetic causes of severe diseases should take into account this idea that regulatory variants need to be considered alongside coding variants,” Dr. Castel said. “This should eventually lead to a more fine-grained understanding of the risk of coding variants associated with disease.”

#

About the New York Genome Center

The New York Genome Center (NYGC) is an independent, nonprofit academic research institution at the forefront of transforming biomedical research and clinical care. Founded as a collaborative venture by the region’s premier academic, medical and industry leaders, NYGC’s goal is to translate genomic research into new diagnostic tools and treatments for human disease. NYGC member organizations and partners are united in this unprecedented collaboration of technology, science and medicine, designed to harness the power of innovation and discoveries to advance genomic services. Their shared objective is the acceleration of medical genomics and precision medicine to benefit patients around the world.

Institutional Founding Members include: Albert Einstein College of Medicine, Cold Spring Harbor Laboratory, Columbia University, Icahn School of Medicine at Mount Sinai, The Jackson Laboratory, Memorial Sloan Kettering Cancer Center, NewYork-Presbyterian Hospital, Northwell Health, New York University, The Rockefeller University, Stony Brook University and Weill Cornell Medicine. Institutional Associate Members include: American Museum of Natural History, Hospital for Special Surgery, The New York Stem Cell Foundation, Princeton University and Roswell Park Cancer Institute. IBM is a Founding Technology Member.

For more information on the NYGC, please visit www.nygenome.org.

About Columbia University Department of Systems Biology

The Columbia University Department of Systems Biology offers an interdisciplinary program aimed at advancing the integration of quantitative and experimental research methods in the biological and biomedical sciences. DSB faculty, students and post-docs focus on understanding the behavior of complex biological networks and how small changes in those networks produce the diversity of life around us. Our onsite research infrastructure includes high-performance computing, next-generation sequencing and high-throughput screening, and we also provide graduate education and postdoctoral training to increase understanding of systems biology and promote its use in biological research and disease-related areas, including cancer, infectious diseases, and metabolic disorders.

SCIENTIST CONTACT:

Tuuli Lappalainen, PhD

O: 646-977-7037

M: 917-973-2661

tlappalainen@nygenome.org

MEDIA CONTACT:

Karen Zipern

Director of Communications

O: 646-977-7065

M: 917-415-8134

kzipern@nygenome.org