Using cutting-edge statistical models to analyze data from nearly 2,000 families with an autistic child, a multi-institute research team discovered tens of thousands of rare mutations in noncoding DNA sequences and assessed if these contribute to autism spectrum disorder.

Published December 14 [2] in the journal Science, the study is the largest to date for whole-genome sequencing in autism. It included 1,902 families comprised of both biological parents, a child affected with autism and an unaffected sibling.

Scientists representing Carnegie Mellon University, UC San Francisco, University of Pittsburgh School of Medicine, Massachusetts General Hospital, Harvard Medical School and the Broad Institute led the research team.

The study is one of 13 being released Dec. 14 as part of the first round of results to emerge from the National Institute of Mental Health's PsychENCODE consortium [3]?a nationwide research effort that seeks to decipher how noncoding DNA, often referred to as the ?dark
matter of the human genome, contributes to psychiatric diseases such as autism, bipolar disorder and schizophrenia.

Over the past decade, scientists have identified dozens of genes associated with autism by studying so-called "de novo" mutations—newly arising changes to the genome found in children but not their parents. To date, most de novo mutations linked to autism have been found in protein-coding genes. It has proven far more difficult for scientists to identify autism-associated mutations in noncoding regions of the genome.

"Protein-coding genes clearly play an important role in human disorders like autism, yet their expression is regulated by the "noncoding" genome, which covers the remaining 98.5 percent of the genome and remains somewhat mysterious," said Carnegie Mellon's Kathryn Roeder, PhD, corresponding author and UPMC Professor of Statistics and Life Sciences. "Because the genome comprises 3 billion nucleotides, identifying which portions of the noncoding genome, when mutated, enhance the risk of autism is as challenging as looking for a needle in a haystack."

Using a novel bioinformatics framework, the researchers were able to compress the search from billions of nucleotides to tens of thousands of functional categories that potentially contribute to autism. Working with these categories, they used machine learning tools to build statistical models to predict autism risk from a subset of the families in the study. They then applied this model to an independent set of families and successfully predicted patterns of risk in the noncoding genome.

Though rare de novo mutations were found in many noncoding regions of the genome, the strongest signals arose from promoters—noncoding DNA sequences that control gene transcription. These risk-conferring promoters were most often located far from the genes...
under their control. They were also found to be largely conserved across species, suggesting that any rare mutations that might arise in these promoters are more likely to disrupt normal biology.

For years, scientists have used genome-wide studies to find common variants that confer disease risk. Our group has now focused on creating a computational framework that’s capable of finding rare, high-impact variants associated with a human disorder, looking across all the noncoding regions of the genome, said Stephan Sanders, BMBS, PhD, corresponding author and professor of psychiatry at the UCSF Weill Institute for Neurosciences and Institute for Human Genetics.

The team’s findings have practical implications for future research on model organisms, like mice, as attempts are made to move toward genetically informed therapies for autism. But the value of studying the noncoding genome extends well beyond autism.

We were particularly interested in the elements of the genome that regulate when, where and to what degree genes are transcribed. Understanding this noncoding sequence could provide insights into a variety of human disorders, said Bernie Devlin, PhD, corresponding author and professor of psychiatry at the University of Pittsburgh School of Medicine.

We are just scratching the surface of what there is to learn about noncoding regulatory variation in human disease, and the new methods this team has developed will catalyze an important step forward into larger and more comprehensive studies, said Michael Talkowski, PhD, of Massachusetts General Hospital, Harvard Medical School and the Broad Institute, who also served as corresponding author on the study.

Authors: Lead authors on the paper are Joon-Yong An, PhD, and Donna Werling, PhD, of the UCSF Department of Psychiatry, and Kevin Lin and Lingxue Zhu, PhD, of CMU’s Department of Statistics and Data Science. Additional authors on the paper include Shan Dong, PhD; Grace B. Schwartz; Claudia Dastmalchi; Jeanselle Dea; Clif Duhn; Michael C. Gilson; Lindsay Liang; Eirene Markenscoff-Papadimitriou, PhD; Nadav Ahituv, PhD; Young Shin Kim, MD, PhD; John L. Rubenstein MD, PhD; Matthew W. State, MD, PhD; and A. Jeremy Willsey, PhD, of UCSF; Harrison Brand, PhD, MPH, Harold Z. Wang, Xuefang Zhao, PhD, Ryan L. Collins, Benjamin B. Curroll, PhD, Mark J. Daly, PhD, Benjamin M. Neale, PhD, of Massachusetts General Hospital, Harvard Medical School and the Broad Institute; Lambertus Klei, PhD, of the University of Pittsburgh School of Medicine; Sirisha Pochareddy, PhD, Nenad Sestan, MD, PhD, of the Yale School of Medicine; Joseph D. Buxbaum, PhD, of Icahn School of Medicine at Mount Sinai; Hilary Coon, PhD, Gabor T. Marth, DSc, Aaron R. Quinlan, PhD, of the University of Utah School of Medicine.

Funding: Research was supported by National Institute of Mental Health PsychENDODE Consortium; Simons Foundation for Autism Research Initiative grants 40228, 385110, 574598, 385027, 346042, 575097, 573206, 513631 and 388196; NIH grants U01 MH105575, U01 MH100239-03S1, R01 MH110928, R01 MH109901, U01 MH111662, U01 MH111658, U01 MH111660, U01 MH111661, R37 MH057881, R01 HD081256, R01 MH115957, R01 MH049428, R01 MH107649-03 and R01 MH094400; the Stanley Center for Psychiatric Genetics.
UCSF Disclosures: Matthew W. State, MD, PhD, is on the scientific advisory boards for ArRett Pharmaceuticals and BlackThorn Therapeutics and holds stock options in ArRett Pharmaceuticals.

Read the study

- **Science**: Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder [2]

Further coverage

- **New York Times**: Mapping the brain's genetic landscape [22]
- **Simons Foundation**: Noncoding mutations contribute to autism risk [23]

---

About UCSF Psychiatry

The UCSF Department of Psychiatry [24], UCSF Langley Porter Psychiatric Hospital and Clinics [25], and the Langley Porter Psychiatric Institute are among the nation's foremost resources in the fields of child, adolescent, adult, and geriatric mental health. Together they constitute one of the largest departments in the UCSF School of Medicine and the UCSF Weill Institute for Neurosciences, with a mission focused on research (basic, translational, clinical), teaching, patient care, and public service.

UCSF Psychiatry conducts its clinical, educational and research efforts at a variety of locations in Northern California, including UCSF campuses at Parnassus Heights, Mission Bay and Laurel Heights, UCSF Medical Center, UCSF Benioff Children’s Hospitals, Zuckerberg San Francisco General Hospital and Trauma Center, the San Francisco VA Health Care System, and UCSF Fresno.

About the UCSF Weill Institute for Neurosciences

The UCSF Weill Institute for Neurosciences [26], established by the extraordinary generosity of Joan and Sanford I. "Sandy" Weill, brings together world-class researchers with top-ranked physicians to solve some of the most complex challenges in the human brain.

The UCSF Weill Institute leverages UCSF’s unrivaled bench-to-bedside excellence in the neurosciences. It unites three UCSF departments?Neurology, Psychiatry, and Neurological Surgery?that are highly esteemed for both patient care and research, as well as the Neuroscience Graduate Program, a cross-disciplinary alliance of nearly 100 UCSF faculty members from 15 basic-science departments, as well as the UCSF Institute for Neurodegenerative Diseases, a multidisciplinary research center focused on finding effective treatments for Alzheimer’s disease, frontotemporal dementia, Parkinson’s disease, and other neurodegenerative disorders.

About UCSF

UC San Francisco (UCSF) [27] is a leading university dedicated to promoting health worldwide
through advanced biomedical research, graduate-level education in the life sciences and health professions, and excellence in patient care. It includes top-ranked graduate schools of dentistry, medicine, nursing and pharmacy; a graduate division with nationally renowned programs in basic, biomedical, translational and population sciences; and a preeminent biomedical research enterprise.

It also includes UCSF Health [28], which comprises three top-ranked hospitals ? UCSF Medical Center and UCSF Benioff Children?s Hospitals in San Francisco [29] and Oakland [30] ? as well as Langley Porter Psychiatric Hospital and Clinics, UCSF Benioff Children?s Physicians, and the UCSF Faculty Practice. UCSF Health has affiliations with hospitals and health organizations throughout the Bay Area. UCSF faculty also provide all physician care at the public Zuckerberg San Francisco General Hospital and Trauma Center, and the San Francisco VA Medical Center. The UCSF Fresno Medical Education Program is a major branch of the University of California, San Francisco?s School of Medicine.
[28] https://www.ucsfhealth.org/
[29] https://www.ucsfbenioffchildrens.org/