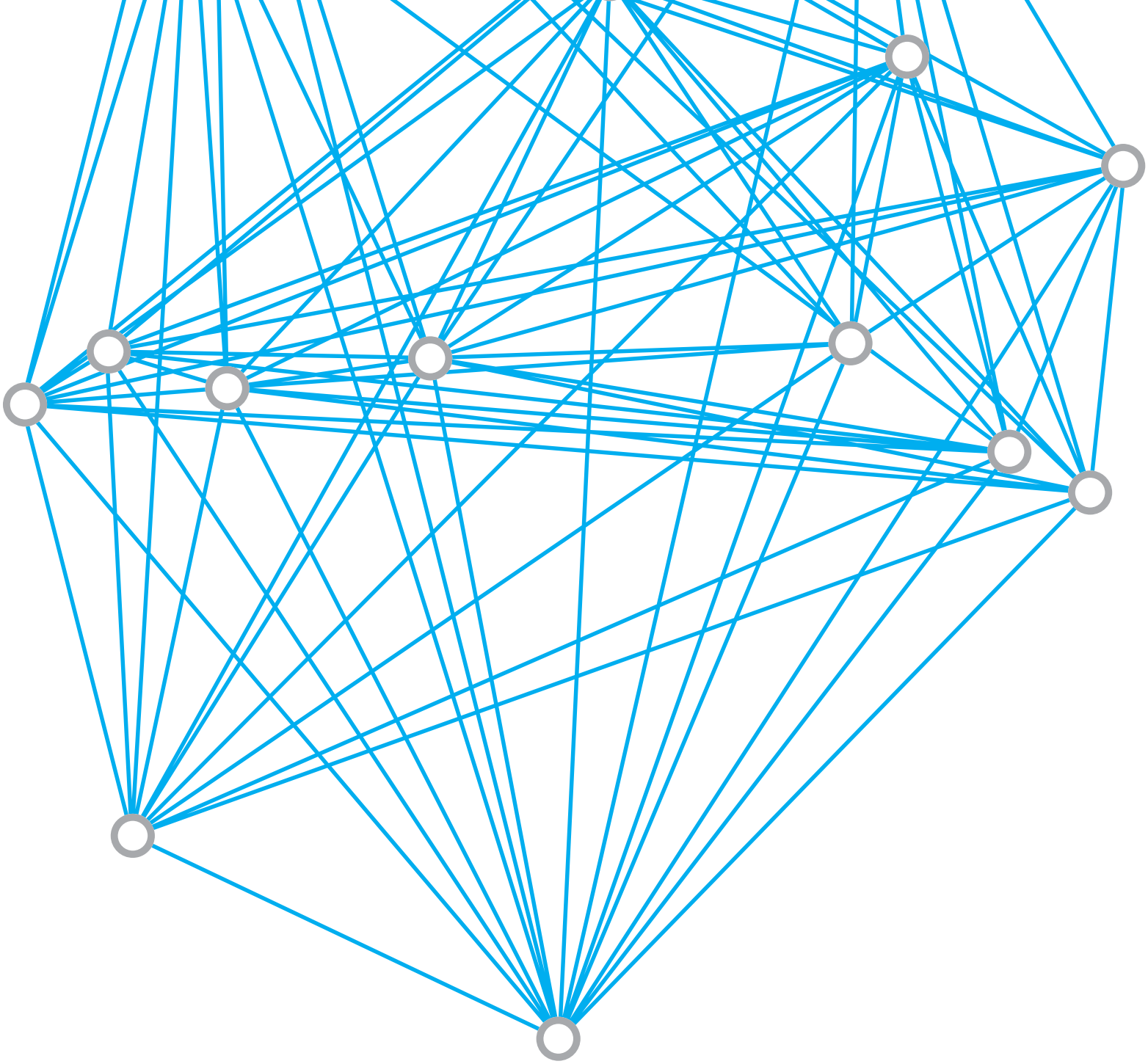


SIMONS FOUNDATION

Annual Report

2014 Edition

Navigation

The navigation tag system below appears throughout this report to assist in identifying which Simons Foundation division(s) an article is associated with. An article may also relate to a Simons Collaboration, a new programmatic model that brings investigators together to address complex scientific questions.

Divisions

SFARI Simons Foundation Autism Research Initiative

LS Life Sciences

SCDA Simons Center for Data Analysis

MPS Mathematics and Physical Sciences

EO Education & Outreach

Simons Collaborations

GB The Global Brain

AG Algorithms and Geometry

MEP The Many Electron Problem

OPE Ocean Processes and Ecology

Table of Contents

2014

Board Members

Marilyn Hawrys Simons, Ph.D.
President, Simons Foundation

David Eisenbud, Ph.D.
Director,
Mathematical Sciences
Research Institute

Gerald D. Fischbach, M.D.
Chief Scientist and Fellow
of the Simons Foundation

Mark Silber, J.D., M.B.A.
Executive Vice President and
Chief Financial Officer,
Renaissance Technologies

James H. Simons, Ph.D.
Chair, Simons Foundation

Staff

Maria Adler
Leyla Ahari
Alpha Amatya
Kate Augenblick
Agnes Barszcz
Marta Benedetti
Lawrence Bianco
Aaron Biscombe
Alexandra Bolter
Richard Bonneau
Greg Boustead
Alexandra Bowe
Elizabeth Brooks
Martin Butler
Jacob Cappell
Marian Carlson
Nick Carriero
Mani Cavalieri
Ahmad Chatha
Dmitri "Mitya" Chklovskii
Andrew Choi
Salim Chowdhury
Wendy Chung
Justin Creveling
Amy Daniels
Donna DeJesus-Ortiz
Lisa Dicce
Meghan Fazzi
Pamela Feliciano
Gerald Fischbach
Ian Fisk
Chris Fleisch
Neil Flood
Patrick Flood
Steven Ford
Calissia Franklyn
Annaliese Gaeta
Laura Geggel
Daniel Goodwin
Katie Goodwin
Lee Anne Green Snyder
Adrienne Greenberg
Anastasia Greenebaum
Leslie Greengard
Marion Greenup
Fang Han
Rebecca Hawkins
Jessica Holthouser
Marian Jakubiak
Rachel Jansen
Lydia Jung
Rachel Jurd
Timothy Kane
Deborah Kenyon
Patricia Kim
Jack Kimura
Timothy Kissane
Boyana Konforti
Abraham Lackman
Alex Lash
Monika Lenard
David Leyden
Thomas Lin
Alice Luo Clayton
Apoorva Mandavilli
Julie Manoharan
Richard Marini
Richard McFarland
Andrew Millis
Emily Miraldi
Katie Moisse
Michael Moyer
Tanja Mrdjen
Elizabeth Mrozinska
Christian L. Müller
Layla Naficy
Jessa Netting
Camille Norrell
Debra Olchick
Joette Orsini
Joanna Pacholarz
Alan Packer
Mary Paliouras
Andras Pataki
Danielle Patch
Cengiz Pehlevan
Alyssa Picchini Schaffer
Eftychios Pnevmatikakis
Louis Reichardt
P. Douglas Renfrew
Alexander Rives
Mariah Roda
Elizabeth Roy
Kathleen Savarese
Kim Scobie
Krista Selasky
Alexandra Shaheen
Olena Shmahalo
Emily Singer
Kori Smith
Julia Sommer
John Spiro
Marina Spivak
Jennifer Tjernagel
John Tracey
MaryKate Trawinski
Olga Troyanskaya
Yuri Tschinkel
M. Amedeo Tumolillo
Evan Tziazas
Hope Vanderberg
Benjamin VanderSluis
Kathryn VanderWoude
Jan Varghese
Natalia Volfovsky
Karen Walton-Bowen
Patricia Weisenfeld
James Whalley
Casey White Lehman
John Whitfield
Ingrid Wickelgren
Natalie Wolchover
Aaron Wong
Jessica Wright
Shirley Ying
Nicholette Zeliadt
Stephen R. Zukin

Letter From the President and the Chair

In 2014, *The New York Times* featured 7,029 articles that included the word 'data.' One hundred years earlier, in 1914, there were just 467 articles mentioning 'data,' and 50 years before that, in 1864, a mere 63 articles. Percentagewise, these figures represent 8.5 percent, 0.63 percent and 0.31 percent of the newspaper's stories for those years, respectively. As these statistics from The *New York Times*' language usage tool Chronicle illustrate, the role of data is increasingly important in our world.

Why is data the new frontier? Recent technology is facilitating our ability to capture, store and process massive sets of information. With the potential to accumulate such troves of numerical observations across fields of expertise, researchers see many opportunities for querying large resources to ask the big questions in science. What are the ultimate constituents of the universe? What are the origins of life? Is there life on other planets? How do neural circuits integrate information to form thoughts and memories? What can we learn from our genome about disease, evolution and the diversity of life?

While large-scale data collection holds the promise of advancing our knowledge, there are also many challenges to surmount in the handling of these vast datasets. We need improved techniques for analysis, advanced filtering algorithms, larger storage capacity, better transporting capability and faster processing hardware.

At the Simons Foundation, we see the potential gains to be garnered from the analysis of datasets, and we see the complexities inherent in processing immense stores of information. As you will see in the following pages, we are supporting both theoretical and applied efforts in big data.

The foundation is interested not only in big data, however, but in data generally. We support the development of datasets and aim to provide them to investigators as a no-cost, collective resource. Such shared resources facilitate the cross-pollination of ideas among scientists who share information across disciplines and organizations. The foundation also fosters collaboration between outside investigators and in-house working groups.

At the nexus of this interchange is our new internal data research division, the Simons Center for Data Analysis. SCDA seeks to study datasets of great scientific interest and, in the process, develop new mathematical tools for their study. With an initial focus on neuroscience, genomics and systems biology, the modus operandi of SCDA is to collect, analyze, innovate and share.

With data taking on an increasingly important role in our society and in decision-making, mathematical skills and scientific literacy are becoming ever more essential. We need these skills not only to process information, but also to weigh the validity of its purported conclusions. As John Ewing cautions in his opinion piece, intelligence and insight must always be applied to truly gain insights from a set of numbers.

In this annual report, our goal is to show you some of our efforts around big data, and around data in general. The word 'data' will appear in 64.7 percent of our stories, or 76.57 percent if you include the words 'database' and 'dataset.' We hope you enjoy reading about our work.

Marilyn Hawrys Simons, Ph.D.
President, Simons Foundation

James H. Simons, Ph.D.
Chair, Simons Foundation

Simons Simplex Collection

8,660

Average number of data points yielded by testing an SSC family

Fifteen years ago, autism genetics research was at something of an impasse. Twin and family studies suggested that the disorder had a strong genetic component. But even though geneticists believed that there were likely dozens of different mutations responsible for autism, when researchers went looking for the mutated genes that were presumably passed down from parent to child, they largely came up dry.

Michael Wigler, a geneticist at Cold Spring Harbor Laboratory in New York, suspected that researchers were going about their search in the wrong way. Most autism genetics studies were being done on 'multiplex' families, in which more than one family member has the disorder, because those are the cases with the highest chance of having been caused by inherited mutations. But autism, in fact, appears more often in 'simplex' families, in which only one individual is affected, suggesting that the disorder may frequently result not from an inherited mutation but from a spontaneous, or de novo, mutation in a sperm or egg cell. "I had the hypothesis that people were failing [in their search for autism genes] because they were using the wrong tool," Wigler says.

Together with Jonathan Sebat, then in Wigler's lab and now at the University of California, San Diego, Wigler analyzed DNA from the Autism Genetic Resource Exchange, a gene bank consisting mostly of genetic material from multiplex autism families: a valuable dataset from many perspectives, but the opposite of what is needed to best isolate de novo mutations. Indeed, as Wigler had predicted, the simplex families the team studied from that collection had a higher proportion of de novo mutations than the multiplex families did.

In 2003, Wigler broached to Jim Simons the idea of creating a large collection of simplex families. The Simons Foundation was already looking for ways to invigorate the field of autism research, and earlier that year it had convened a large meeting of autism experts who had concluded that it was imperative to lure more talented researchers into the field. A large, carefully curated collection of data from simplex families, freely accessible to all researchers, would be an ideal way to jump-start research, the foundation decided.

Today, the Simons Simplex Collection (SSC), which holds genetic, phenotypic and biological data from more than 2,600 simplex families, has helped lead the revolution in autism research. "I don't know of another autism collection that compares to it," says Evan Eichler, an autism researcher at the University of Washington in Seattle.

The logic behind the SSC has proved to be "profoundly right," says Matthew State, a geneticist at the University of California, San Francisco. "It has made the field."

The Verga family of Queens, New York, participants in the Simons Simplex Collection.

229

Projects using SSC resources

"It's hard to imagine where the genomics of autism would be without the SSC," says State, who carried out the exome study together with Wigler, Eichler and Jay Shendure, Eichler's colleague at the University of Washington. "It has absolutely transformed autism research."

From the earliest days of the SSC, the emphasis was on creating a resource that could be used in many different ways by many different kinds of scientists. "No other research group has put so much effort into making sure their dataset would be widely usable," says Catherine Lord of Weill Cornell Medical College in New York City, who oversaw the formation of the collection together with Gerald Fischbach, then director of the Simons Foundation Autism Research Initiative (SFARI) and now the foundation's chief scientist. "Normally, people collect data for themselves, but we were thinking from the start about what data outside researchers would want."

To be most useful to the autism research community, Lord and Fischbach decided, the collection must be not only large but also deep, with detailed phenotypes and biospecimens. "We wanted people to be able to use it to test their hypotheses even if they didn't have access to anyone with autism," says Lord. The meticulous data the collection acquired, Lord says, "cut out millions of steps for many researchers who otherwise wouldn't have even gotten started."

Sequencing studies of the collection over the past five years are bringing the genetic landscape of autism into sharp focus. Instead of dozens of autism mutations, researchers now believe that 300 to 1,000 genes will eventually be implicated in the disorder. A recent analysis of the exomes — the protein-coding regions of the genome — of most of the SSC families, published November 13, 2014, in *Nature*, has identified 27 autism genes with high confidence, as well as hundreds more candidate genes worthy of further study.

The collection has helped to spark a new era in autism research, and its accessibility has attracted many scientists into the field who had not previously focused on autism.

"The fact that the SSC was out there without any strings attached — that it wasn't wrapped up in someone's empire — is a big part of the reason I moved into autism research," Eichler says. "The collection allows an entirely new dimension of researchers to explore autism."

157,880

SSC biospecimens shipped

No one institution would have been able to collect data on as many simplex families as the collection needed, so Fischbach and Lord enlisted the aid of 12 clinics and universities in the U.S. and Canada: Baylor College of Medicine; the University of California, Los Angeles; Columbia University; Emory University; Harvard University/Boston Children's Hospital; the University of Illinois at Chicago; the University of Michigan; McGill University; the University of Missouri; Vanderbilt University; the University of Washington; and Yale University. The different clinics brought in simplex families — not just the affected child, but also the parents and siblings — for a full day of diagnostic tests and blood draws.

"It's a significant commitment of time and energy for the families," says Casey White Lehman, the collection's project manager at the Simons Foundation. "And they did it mainly out of a desire to contribute to research, which I've always found inspiring."

Silvia Verga of New York City, whose family participated in the Simons Simplex Collection, says, "We wanted to be part of something that could be the beginning of discoveries in the future." The family reaped some immediate benefit from the detailed assessment process: "The evaluation we received cleared up a lot of questions I had with regard to my son's diagnosis, and helped us figure out what kinds of services would be best for him," Verga says.

With 12 different sites collecting data, it was imperative that the tests be carried out consistently from one site to another. "If you're going to produce a number, it should mean something," Lord says. "We had to make sure the numbers meant the same thing at different places."

Lord brought the various clinicians to the University of Michigan (where she was based at the time) for rigorous training on the diagnostic tests to be performed, some of which she herself had pioneered. The clinicians were later videotaped as they assessed families, to make sure they were carrying out the tests uniformly, and the teams also participated in site visits, monthly phone calls and biannual group meetings with the staff at the Simons Foundation.

This attention to consistency paid off. "When the scores were tallied at these sites scattered around the country, they all came out very close on a wide range of measures — of social cognition, or repetitive movements, or the severity of the disease," Fischbach says. "Given what a heterogeneous disorder autism is, that seems like a miracle."

Somewhat to her surprise, Lord found that the area of greatest variability among the different clinics was the name the clinicians assigned to each diagnosis — autistic disorder, Asperger syndrome or pervasive developmental disorder-not otherwise specified (PDD-NOS).

Each clinic seemed to assign these names according to its own internal logic, which varied greatly from site to site. Lord's statistical analysis of this variation, with Eva Petkova of New York University, played an influential role in the decision of the American Psychiatric Association in 2012 to replace the three labels with the umbrella term "Autism Spectrum Disorder" in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*.

By 2011, the collection had completed its data accumulation phase, with a total of 2,659 families — well beyond its initial target of 2,000. In the process, the teams at the 12 sites had become a community. "There has been a lot of back and forth since then, with people working together who met through the SSC," Lord says.

Data from the collection are available to qualified researchers through a central database called SFARI Base. The foundation has also recently federated with the National Database for Autism Research (NDAR), maintained by the National Institutes of Mental Health, so that a researcher who types a query into NDAR will receive results from both databases. To date, more than 200 projects have used data from the SSC, and more than 60 published papers have resulted from their analyses, appearing in publications such as *Nature*, *Science*, *Neuron*, *Cell* and *Nature Genetics*.

2,659

Total SSC families

120

Researchers using SSC resources

The first clear indication of the collection's potential for elucidating autism's genetic architecture came in 2010, when Wigler and State completed an analysis of de novo copy number variants (CNVs) — genetic aberrations in which a chunk of DNA is duplicated or deleted — in more than 1,000 SSC families. The study highlighted six genomic regions that appeared to be strongly linked to autism, and about 70 other candidate autism CNVs.

These findings supported the creation in 2010 of the Simons Variation in Individuals Project (Simons VIP), which has collected clinical information and blood samples from more than 200 carriers of 16p11.2 CNVs, to home in on the shared neurological and behavioral features of this group. The project's long-term goal is to identify the features of different genetic subtypes of autism, which might respond to different therapeutic approaches.

In 2012, Wigler, State and Eichler's labs sequenced the whole exomes of nearly 800 SSC families, identifying several high-confidence autism risk genes. More recently, the 2014 exome study of nearly the entire SSC by Wigler, State, Eichler and Shendure has suggested hundreds of candidate autism risk genes. "Many of these candidates will be confirmed in the coming years, by additional deep sequencing of autism collections," predicts Alan Packer, a senior scientist at SFARI.

Seven of the genes identified in the 2014 study had mutations in three or more children with autism, establishing them unassailably as autism risk genes. Another 20 genes had mutations in two children, which translates into more than a 90 percent likelihood of their being genuine autism genes.

"Before the SSC was created, I was waiting for the one rare kid to walk into my lab that had a de novo mutation in a gene, and then we would still have to figure out how to prove that the mutation was related to the disorder," State says. "It used to take us a decade to find one autism gene, so to publish a paper with 27 is amazing."

While the studies so far have illuminated the incredible genetic diversity of autism, they also strongly suggest that the hundreds of autism genes likely converge on a much smaller set of biological pathways. Many of the candidate genes to emerge from the exome study, for instance, interact with targets of the gene causing fragile X syndrome, which causes intellectual disability in boys. Other candidate genes are involved in the regulation of chromatin, a DNA-protein complex that helps package DNA in the cell nucleus and controls gene expression. Understanding these biological pathways, researchers hope, will eventually lead to targeted therapies for the different genetic types of autism.

The genetic studies have given rise to a burst of animal studies to try to decode the biological mechanisms of the strongest autism candidate genes.

"In the long run, the neurobiology is going to be even more important than the genetics for understanding mechanisms," Fischbach says. "But without the genes, we wouldn't even know which animal models to create and study."

Although genetics has been the primary thrust of SSC research so far, researchers

15
New projects with SSC families

57
Projects using SSC biospecimens

69
SSC projects using genetic data

are examining its data from a host of other perspectives as well. For example, a paper in the November 2013 issue of *Molecular Psychiatry* reported that mothers of children in the SSC were four times more likely than controls to harbor anti-brain antibodies, which might be pathogenic to the developing brain, and they also had an increased prevalence of autoimmune disorders such as rheumatoid arthritis and lupus. The SSC has also enabled scientists to study, for example, repetitive behaviors, the relationship between head circumference and IQ, and even the stigma of autism.

In addition, the collection has offered researchers a variety of ways to tackle the puzzling question of why autism in girls seems to be so different from autism in boys — simultaneously rarer and more severe. State, Wigler and Eichler's genetic studies of the SSC indicate that girls with autism typically have more damaging mutations than boys do.

The Simons Foundation has created a way for researchers to engage many of the SSC families in future studies. SSC@IAN — administered by the foundation in partnership with the Interactive Autism Network of the Kennedy Krieger Institute in Baltimore — is an online platform that connects researchers whose projects have been approved by SFARI with a pool of SSC families. More than 1,500 of the original families in the SSC have agreed to take part. "Time and time again the families have shown us how engaged they are," White Lehman says.

The foundation itself is conducting the SSC@IAN Family Update Study, a set of online questionnaires to find out how the families have fared in the years since the original data collection.

The project may well become a multi-year study, White Lehman says. "It's important to get information about people's lives over the long term," she says. "There hasn't been a lot of research on what happens as individuals with autism transition to the adult world."

Much more remains to be mined from the SSC's genetic data, as the exome accounts for only 1.5 percent of the human genome. The Simons Foundation has launched a pilot study, to be carried out by the nonprofit New York Genome Center, to sequence the entire genomes of 40 families from the collection. If that goes well, larger studies will follow.

The exome studies of the SSC suggest that ultimately, at least 30 percent of simplex autism will be traceable to de novo mutations. "I still read in newspapers that we don't know what causes autism, or that autism is thought to involve gene mutations, but that's not really just a hypothesis anymore," Packer says. "Autism is less mysterious than it used to be."

SFARI Research Roundup

The Simons Foundation Autism Research Initiative (SFARI) Investigator program supports nearly 200 researchers who are carrying out bold, innovative autism research. In 2014, these researchers published dozens of papers, encompassing genetics research, imaging studies of individuals with autism, behavioral studies and a host of other approaches to understanding the complex disorder. The following are some highlights from SFARI Investigators' research activities in the past year.

259
Active grants in 2014

244
Active investigators in 2014

46
New investigators in 2014

Insufficient Pruning

As children with autism transition into adolescence, their neurons prune far fewer dendritic spines — the protrusions that receive messages from other neurons — than in children without the disorder, according to a study in the September 3, 2014, issue of Neuron.

The study's researchers, led by SFARI Investigator David Sulzer of Columbia University in New York City, examined postmortem brain samples from 16 children and teenagers with autism and 12 controls. The control teenagers had 45 percent fewer dendritic spines than their child counterparts did, whereas the teenagers with autism had only 15 percent fewer dendritic spines than the children with autism did. The adolescents with autism also had unusually high levels of mTOR, a protein that inhibits autophagy, the process by which cells recycle unneeded parts.

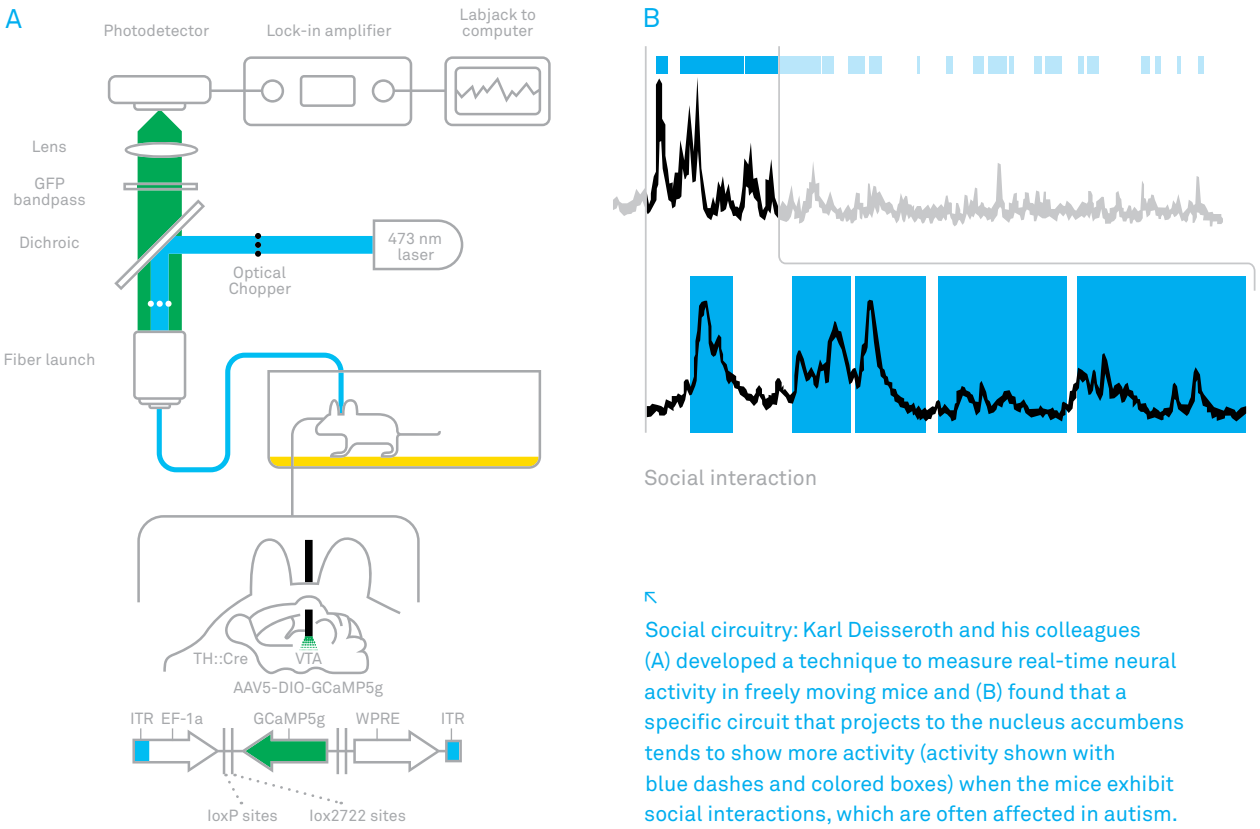
The team also examined a mouse model that has unusually high levels of mTOR activity and exhibits autism-like social behavior. As with the humans, the adolescent mice had a higher density of dendritic spines than did control mice, as well as a lower level of LC3-II, a marker for autophagy. When the mice were treated with rapamycin, a drug that inhibits mTOR activity, their spine density and LC3-II levels became normal, and their social behavior became similar to that of controls, suggesting that drugs that restore autophagy could be a promising approach for treating autism.

A Clear Autism Subtype

Individuals with mutations in the gene CHD8 — the autism candidate gene with by far the strongest evidence — have a consistent and recognizable phenotype, a new study shows. The study is one of the first to examine people with autism who all have the same mutation, an approach that may eventually help researchers zero in on personalized treatments for the highly heterogeneous disorder, says SFARI Investigator Evan Eichler of the University of Washington, who led the study.

The researchers collected detailed observations from 15 individuals with CHD8 mutations, 13 of whom have autism diagnoses (the other two have been diagnosed with intellectual disability and may well have undiagnosed autism, the researchers say). Most of these individuals have wide-set eyes, large ears, and broad foreheads and noses. Twelve have enlarged heads, 12 have digestive problems, especially constipation, and 10 have sleep problems. To gain insight into the biological mechanism underlying this phenotype, the researchers blocked CHD8 expression in zebrafish embryos. Like the people with CHD8 mutations, the zebrafish developed wide-set eyes and a sluggish digestive tract, and had only half as many neurons in their gut as controls did.

Attempts to identify subtypes of autism by looking at people with similar symptoms have not been very successful, the researchers wrote in the July 17, 2014 issue of *Cell*. The study, which indicates that CHD8 disruption is a distinct autism subtype, suggests that a 'gene-first' approach may be more fruitful, Eichler says.



Social circuitry: Karl Deisseroth and his colleagues (A) developed a technique to measure real-time neural activity in freely moving mice and (B) found that a specific circuit that projects to the nucleus accumbens tends to show more activity (activity shown with blue dashes and colored boxes) when the mice exhibit social interactions, which are often affected in autism.

Pinpointing Brain Circuits

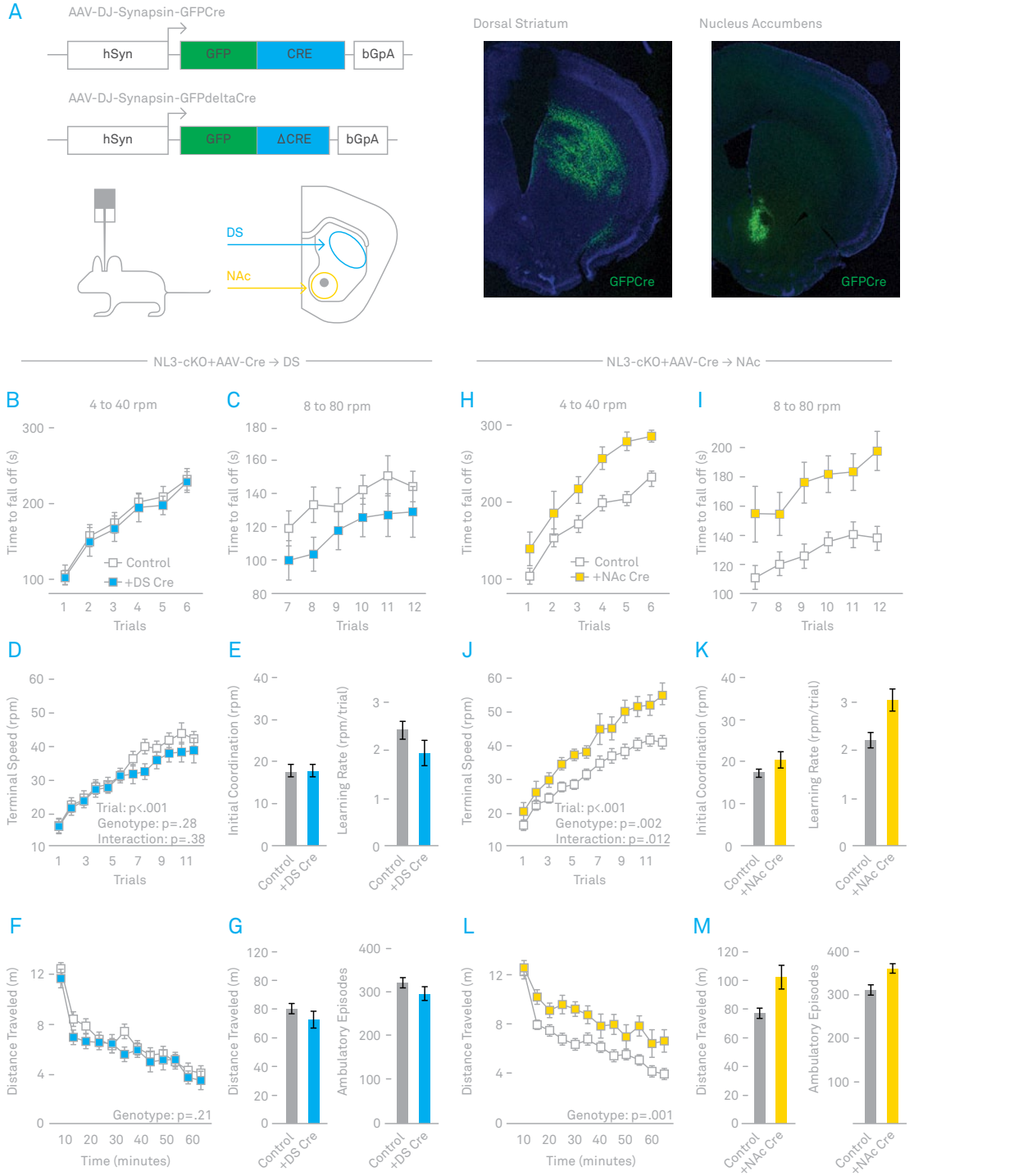
Two Stanford University studies have linked autism with a brain region called the nucleus accumbens, which is involved in goal-related behavior. Combined, the studies identify a particular circuit that seems to be involved in two core domains of autism, social and repetitive behaviors, giving researchers a toehold on the largely unsolved problem of understanding the neural circuitry of autism.

One study, led by SFARI Investigators Thomas Südhof and Robert Malenka, suggests that the region plays a role in repetitive behaviors in mice. The researchers looked at mice with mutations that either knock out or greatly lower the level of neuroligin-3 (NLGN3), a protein involved in synapse formation that has been strongly linked to autism. The team reported in the July 3, 2014, issue of *Cell* that both kinds of mutations cause mice to perform unusually well in tests of learned repetitive routines akin to log-rolling contests.

The researchers next removed NLGN3 from specific brain cells, to try to home in on the part of the brain underlying this repetitive behavior. Their study implicated a set of neurons in the nucleus accumbens that express a receptor for the chemical messenger dopamine, which is involved in reward signals. When the researchers restored NLGN3 to these neurons, the behavior of the mice returned to normal, suggesting that this type of repetitive behavior might be amenable to drug therapies.

In the second study, researchers developed a new technique that provides scientists with a window into the dynamics of real-time circuit activity during social interactions in mice, and points to a specific circuit whose activity encodes such interactions.

The technique, created by a team led by SFARI Investigator Karl Deisseroth and Robert Malenka, involves inserting a tiny fiber-optic probe into the brains of mice that have been engineered to express a fluorescent molecule during the firing of neurons that make dopamine. When the probe was placed in the nucleus accumbens, it measured that significantly more dopamine spikes occurred when the mice interacted with new mice than when they interacted with novel objects, the researchers reported June 19, 2014, in *Cell*. The results suggest that a circuit connecting a region called the ventral tegmental area, a major source of dopamine neurons, to the nucleus accumbens specifically modulates social behavior.



↗

Pinpointing brain circuits: Thomas Südhof's lab used a motor-learning task akin to log-rolling to model a cardinal feature of autism: acquired repetitive behavior. The researchers found (A) that absence of the autism-associated gene NL3 (B-G) in the dorsal striatum, a brain region known for its role in motor function, surprisingly did not change performance on the task. (H-M) Rather, deletion of NL3 in the nucleus accumbens, an area associated with learning and motivation, seemed to enhance learned repetitive behavior.

Regulating Brain Size

A chromosomal region called 16p11.2, which is strongly linked to autism, may control brain size, researchers reported August 20, 2014, in *The Journal of Neuroscience*.

Deletions and duplications in 16p11.2, which contains 29 genes, are among the genetic variations that have been most frequently associated with autism. To understand the implications of these deletions and duplications, in 2010 the Simons Foundation created the Simons Variation in Individuals Project (Simons VIP, *see* Simons Simplex Collection, *pg.* 8), which, like Eichler's study of individuals with CHD8 mutations, takes a 'gene-first' approach by studying the phenotypes of several hundred individuals with 16p11.2 deletions or duplications and their families.

In the August study, a research team led by SFARI Investigator Randy Buckner of Harvard University used magnetic resonance imaging to scan the brains of a cohort of individuals in the Simons VIP collection, along with a control group. Deletions in 16p11.2 increased brain size by about 9 percent, the researchers found, while duplications reduced brain size by a similar amount. Changes in brain size were observed across all brain regions, and within each region the deletion and the duplication produced comparable changes in opposite directions, suggesting that the effect of 16p11.2 on brain size is dose dependent. Within the cortex, surface area, which is determined early in development, was affected more strongly than thickness, which is determined later on, suggesting that the mechanisms underlying the mutations' effects may come into play in early embryonic development.

Incorrect Splicing

Depending on which cell type it is in, the same gene can give rise to many different proteins, as different portions of the gene's sequence get spliced out while the gene is being transcribed to RNA. One of the RNA-binding proteins that regulates this splicing in the nervous system, RBFOX, has been implicated in autism, but until now it has been difficult to identify just which genes RBFOX regulates.

In a study published March 27, 2014 in *Cell Reports*, a team led by SFARI Investigators Robert Darnell of Rockefeller University and Chaolin Zhang of Columbia University made a complete map, at single-nucleotide resolution, of all RBFOX's RNA interaction sites in the mouse brain. Of the 1,059 RBFOX splicing events the team identified, more than 10 percent involve autism risk genes, a far greater proportion than chance would predict. The work suggests that RBFOX might be a 'hub' that regulates many autism risk genes.

RBFOX is actually a family of three proteins, the products of a single gene, that play similar roles in the regulation of mRNA splicing: In postmortem brain tissue of some individuals with autism, the researchers found that all three RBFOX proteins were present at lower-than-normal levels, suggesting that mRNA splicing critical for normal brain function was altered in these individuals.

105

New papers published by SFARI investigators in 2014

80

New grants in 2014

Autism BrainNet

650

Registrants between project launch in May and December 31, 2014

Within the last five years, human genetic sequencing studies, animal models and other approaches have led to dramatic advances in our understanding of autism spectrum disorders. Yet, in contrast to many other disorders, autism research is hampered by an almost total lack of access to the affected organ: the human brain. Current magnetic resonance imaging technology does not allow the level of resolution needed to study individual circuits and cells, so the only way to study them is through the use of postmortem brain tissue, which is in short supply.

"There are big unanswered questions that we just don't have the material to answer," says David Amaral, research director of the University of California, Davis, MIND Institute, "such as when different autism genes get expressed in developing and mature brains, and why some young children with autism have disproportionately enlarged brains. If we could answer them, it could start us on a trajectory to more effective interventions. You have to know what the pathology is, first and foremost, before you can consider how to treat it."

"Currently, fewer than 30 high-quality brains of individuals with autism are available in brain banks — nowhere near enough," Amaral says, especially given the heterogeneity of the disorder.

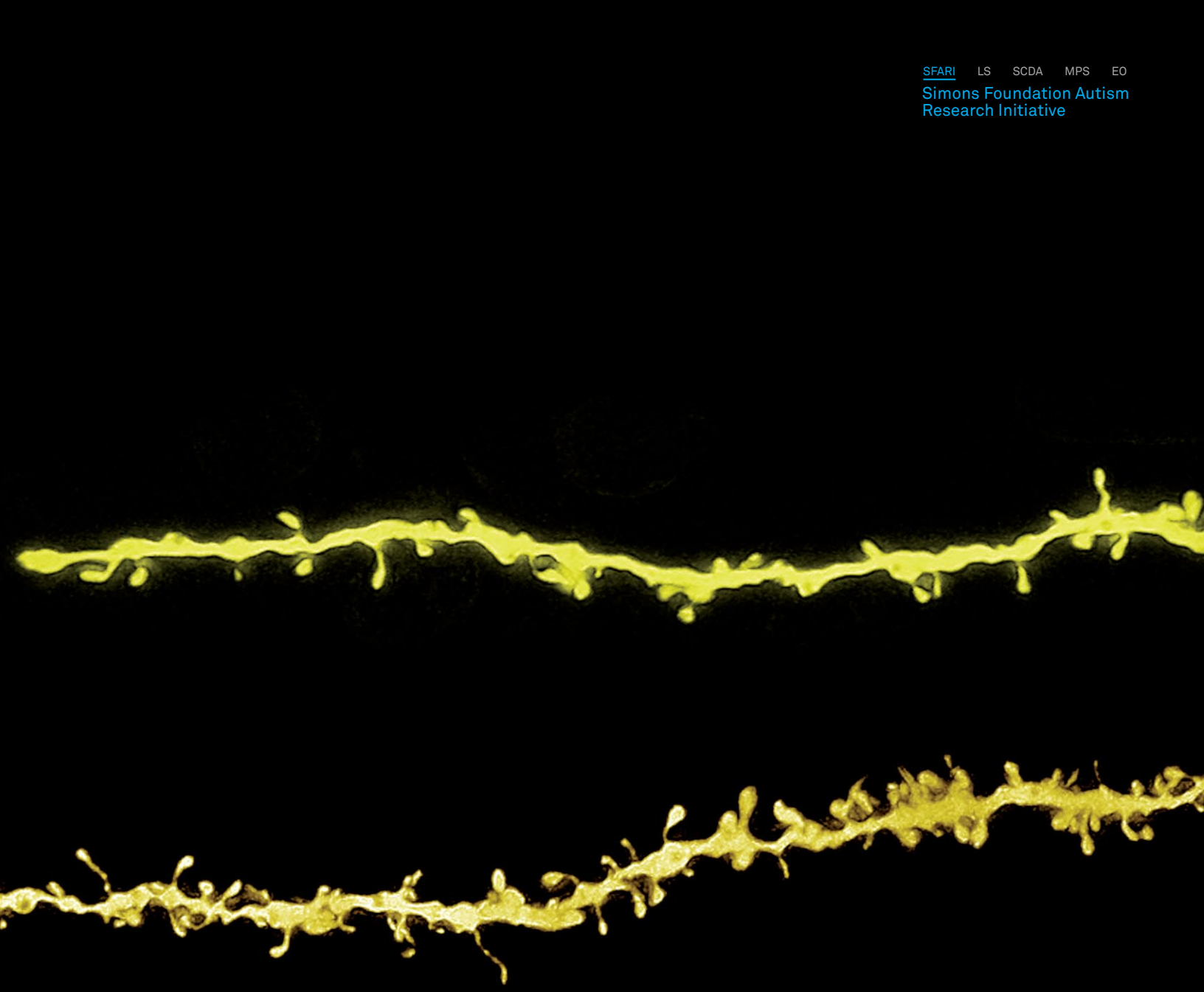
To address this gap, the Simons Foundation has partnered with the science and advocacy organization Autism Speaks to create Autism BrainNet, a network of brain banks launched under Amaral's direction in May 2014. The project, which begins its collection with brain tissue from Autism Speaks' Autism Tissue Program, aims to acquire at least 24 new autism brains and 24 control brains in 2015, roughly doubling the number of brains available to researchers. Once the project is in full swing, Amaral says, it might be possible to collect as many as 100 new brains each year, creating an unparalleled resource for autism researchers. Autism BrainNet is also in discussion with the National Institutes of Health about a potential future partnership that could allow the autism brain bank to grow even more.

Initially, tissue collection and storage will be carried out at four sites across the United States: the MIND Institute, the University of Texas Southwestern Medical Center in Dallas, Beth Israel Deaconess Medical Center in Boston and the Icahn School of Medicine at Mount Sinai in New York City. Information from the sites' banks will be consolidated into a single database, and standardized collection and storage protocols will ensure consistency from site to site.

Previous tissue collection efforts were hamstrung by ineffective outreach campaigns and typically acquired only two to four new brains each year. "You can set up the most beautiful bank in the world, but it will only be as good as the brains it has," says Alison Singer, president of the New York-based Autism Science Foundation. Autism BrainNet has enlisted the Autism Science Foundation to get the message out to autism families, and the resulting It Takes Brains campaign (see TakesBrains.org) has already spurred about 100 families per month to register as potential donors — about 10 times the rate for previous autism brain banks, Singer says.

The campaign emphasizes the heroism of autism families who register as donors, and the potential for their donations to ultimately lead to new therapies and treatments. "Our hope is that this profound gift to science that these families make will increase our understanding of the underlying causes of autism, and eventually improve the quality of life for people on the spectrum," says Marta Benedetti, senior scientist at the Simons Foundation Autism Research Initiative and president of Foundation Associates, LLC, which sponsors Autism BrainNet.

→
The brains of individuals with autism have more abundant synapses than the brains of controls, as cortical pyramidal neurons in autism brains do not undergo normal pruning during childhood and adolescence. These images show representative dendrites seen in unaffected brains (upper) and in autism brains (lower). Spines on the dendrites indicate the location of synapses. Images courtesy of the Sulzer Lab (Guomei Tang and Mark Sonders), Columbia University.



"Currently, fewer than 30 high-quality brains of individuals with autism are available in brain banks — nowhere near enough."

David Amaral

Project on Scientific Transparency

64

SciTran research groups

"We're trying to create a model in which people are surrounded by the tools of sharing, reproducing and computation from the first moment they get data."

Brian Wandell

Like all areas of science, neuroscience has to emphasize reproducibility. Datasets from imaging studies are expensive to obtain and cumbersome to share, and published papers often provide only the vaguest details about how the data were analyzed.

"Papers often claim, 'We used custom in-house software' — as if that's a description of anything," says Brian Wandell, a neuroscientist at Stanford University in California. "It's very hard to make sure the implementation the author wrote about is right."

The Project on Scientific Transparency, directed by Wandell at Stanford, aims to change all that. "The current model is that you do your work and publish your paper, and then if you're not too exhausted at the end, you might put your data into a centralized repository," Wandell says. "We're trying to create a model in which people are surrounded by the tools of sharing, reproducing and computation from the first moment they get data."

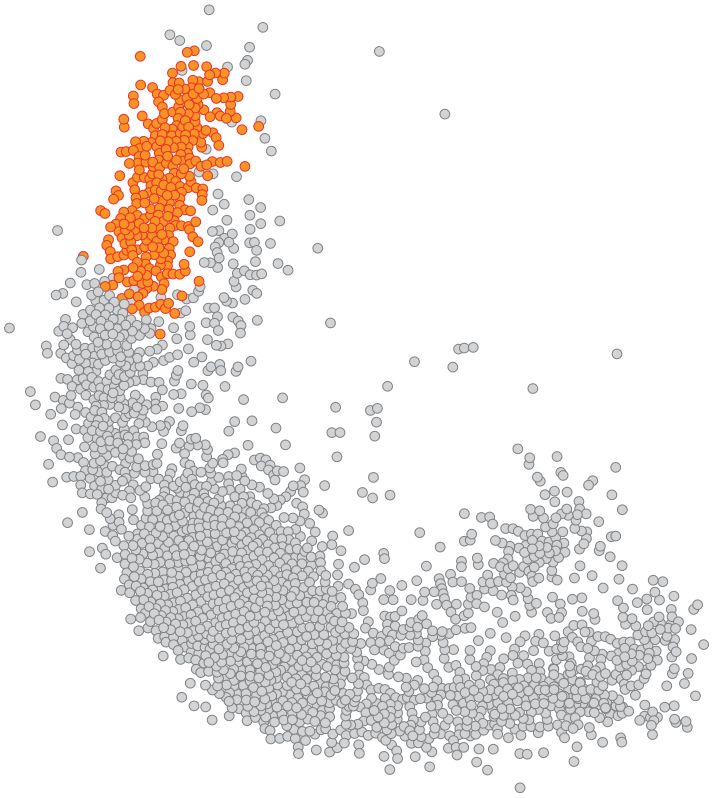
The project has created a data-sharing platform called Scientific Transparency (SciTran) that instantly archives all data obtained by the approximately 500 users of Stanford's shared magnetic resonance imaging (MRI) machines. "There are tools that start from the very first time you push the button on the MRI scanner," Wandell says. Although researchers control access to their own data, sharing with others is easy. Scientists can search through the data — there are already more than 70,000 scans from 9,000 individuals — to identify potential collaborators, for example, or to find control data for an experiment. Wandell is working to install SciTran at other sites, including Indiana University and the University of California, Davis, so that scientists can share data across institutions.

Wandell and his collaborators are now in the process of integrating validated computational tools into SciTran — for instance, a Web-based interactive computational environment called the Jupyter Notebook (formerly the IPython Notebook), created by Fernando Perez of the University of California, Berkeley. This platform, which allows researchers to combine code, data and plain English explanations in a single dynamic website, is like a "living, breathing notebook," says Alex Lash, the Simons Foundation's chief informatics officer — one that can be shared with collaborators simply by sending them a link. The ultimate goal, Lash says, is for researchers to start including links to notebooks in their published papers, so that interested readers can see the detailed protocols that were used to analyze the data.

As a proof of concept, the Simons Foundation is working to upload imaging data to SciTran from about 200 families in the Simons Variation in Individuals Project, a database of families in which at least one individual has a copy number variant in the 16p11.2 chromosomal region, which is strongly linked to autism and other developmental disorders. The foundation is in discussion with SFARI Investigator Randy Buckner of Harvard University, whose laboratory performed one of the analyses of the neuroimaging dataset, about the possibility of representing the study's methods in a Jupyter notebook. "This is a difficult-to-acquire dataset about families with a very specific genetic deletion or duplication; there's probably a lot more we can learn from it if other investigators were to analyze it in different ways," Wandell says.

←
Signals between brain hemispheres are carried by millions of long-range neural fibers called axons, which form a visible and large structure called the corpus callosum. This magnetic resonance image (MRI) shows some of the largest groups of axons in the corpus callosum, colored according to the cortical regions that they connect. In the last decade, researchers found that individual differences in the yellow-colored tract correlate with a child's reading ability. This observation is a topic of study and a source of hypotheses for how we might help children who have difficulty learning to read. MRIs are one type of data shared on SciTran.

Simons Collaboration on the Global Brain



When British physiologists Alan Lloyd Hodgkin and Andrew Huxley first quantitatively described the action potential from a squid neuron in 1952, they initiated decades of work by researchers to uncover the significance of these action potentials for the function of the nervous system.

Nearly half a century later, scientists have mastered the ability to trigger action potentials in single cells. While data from single cells are valuable, much remains unknown about how the brain's neurons are connected into networks, how those networks are connected to one another, and how interactions between neurons and networks lead to thoughts and behaviors. Between triggering a single neuron and observing a discrete behavior, neurons and neural networks produce internal brain states — interacting with each other, performing computations and generating activity — processes that researchers have just begun to decipher.

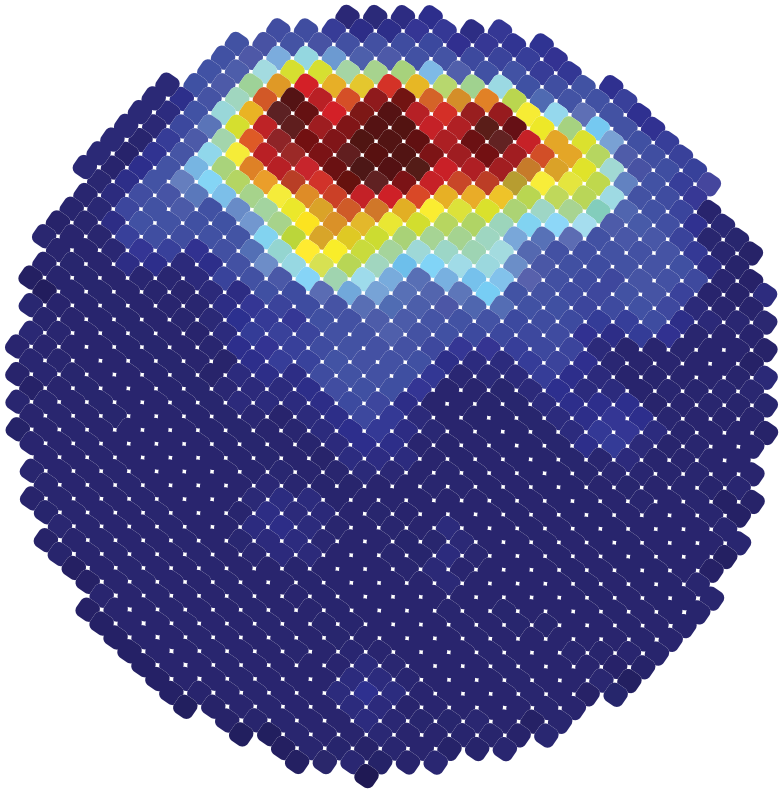
The Simons Collaboration on the Global Brain (SCGB) aims to achieve a comprehensive and mechanistic understanding of brain processes. Using new technologies that allow researchers to record activity from thousands of neurons simultaneously in the brains of awake, behaving animals, 64 investigators and 93 postdoctoral fellows are studying a variety of brain regions, a range of animals and a diverse set of questions in labs around the world.

By collecting data from the activity of many neurons concurrently, SCGB investigators are getting closer to understanding the codes used by relevant neural circuits and how they produce behavior. This deeper understanding of how the brain produces cognition will shed light not only on observable behaviors like movement, but eventually, based on a comprehensive understanding of how neuronal networks function and interact, on the brain's most elusive and unobservable activities: planning, motivation, judgment, memory, emotion and other aspects of reverie that may function even in the absence of obvious external stimuli.

"We know a lot about individual neurons and we know a lot about behavior, but we don't know what happens in the middle," says Alyssa Picchini Schaffer, scientific officer for the SCGB. "Understanding how neural circuits take inputs, make sense of them and then create outputs primes us for a deeper understanding of how the brain works."

Once SCGB investigators achieve a more mechanistic understanding of the brain, they will begin examining the neuronal processes that drive externally unobservable activities, hoping that general principles will emerge. For example, neuronal computations for planning may be similar regardless of what is being planned.

"The ability to study the idea of movement, the planning of movement, before a subject actually moves has been elusive and mysterious," says Gerald D. Fischbach, chief scientist and fellow of the Simons Foundation. "Because issues of internal



40

Experimentalists in the collaboration

30

Theorists in the collaboration

19

Brain regions studied by SCGB investigators

These images represent the firing of a 'place' cell as an animal explores its environment. ↗ Above: A place cell fires more frequently when the animal reaches a particular location in its round environment; warmer colors indicate increased firing. ↖ Left: While the animal explores the environment, spikes are recorded from many different cells, with each circle representing the firing of a cell in the brain. Advanced sorting algorithms identify which spikes came from the place cell, shown in red. Techniques used to capture these data are at the cutting edge of neuroscience research and are central to the work of many SCGB investigators. Images adapted from Aronov D. and D. W. Tank, *Neuron* **84**, 442-456 (2014).

states are no longer simply a matter of philosophy, scientists will be able to study what is going on in the brain that is responsible for these ideas and thoughts."

The collaboration brings together experimentalists — who study various aspects of brain activity in a range of brain regions and in a variety of species, collecting vast amounts of data — and theorists — who create and apply statistical analyses to those massive datasets in an effort to decipher and guide the work of experimentalists. By incorporating biophysical properties of neurons into the data, theorists can create mathematical models of how neurons may assemble into circuits in the brain. Models created by theorists may drive experimentalists' questions, helping frame the way researchers approach the data.

Together, the experimentalists and theorists hope to gain a comprehensive picture of healthy brain function. And once researchers have a solid mechanistic understanding of healthy brain function, members of the SCGB hope to better understand disordered brain activity and its potential treatment.

"These are big questions ... and the brain is a big place," says David Tank, director of the SCGB. "The challenges of understanding the brain are never going to be solved by individual labs. They're just too hard. It's through cross-fertilization of different individuals' ideas from different perspectives that we'll figure these things out."

Simons Center for Data Analysis

One might think that the biggest challenge of founding a research laboratory dedicated to "examining data whose scale and complexity have historically resisted analysis" would lie in unraveling that very complexity. Instead, says Leslie Greengard, founding director of the Simons Center for Data Analysis (SCDA), the challenge lay in recruiting the "initial core" of scientists to lead these research efforts. "It was not easy to find them," Greengard says, "but the right people fell into place faster than I thought."

At one year, SCDA now comprises six research groups focusing on computing, software development, algorithms, systems biology, neuroscience and genomics. "What unifies the application areas is that they're all parts of biomedical and biophysical research, where there are opportunities for discovery that cannot be achieved by classical techniques," Greengard explains. "It's not just about doing a better experiment. It's about combining experiment with theory, simulation and data analysis to learn things that are beyond the capacity of a single measurement."

Part of Dmitri Chklovskii's neuroscience group, for example, is developing techniques for reconstructing the three-dimensional architecture of neural networks from data generated by high-resolution electron microscopy — a computational task that is currently intractable, according to Greengard. Members of Olga Troyanskaya's team conduct meta-analyses of genomic data from different organs (and even different species) to better predict gene expressions associated with specific cell types or specific diseases and developmental disorders, including autism. The systems biology group, led by Richard Bonneau, aims in part to understand the regulatory controls of the immune system and the interaction between the human genome and the 'microbiome' — the microorganisms that live on and in our bodies and outnumber our own cells by an order of magnitude.

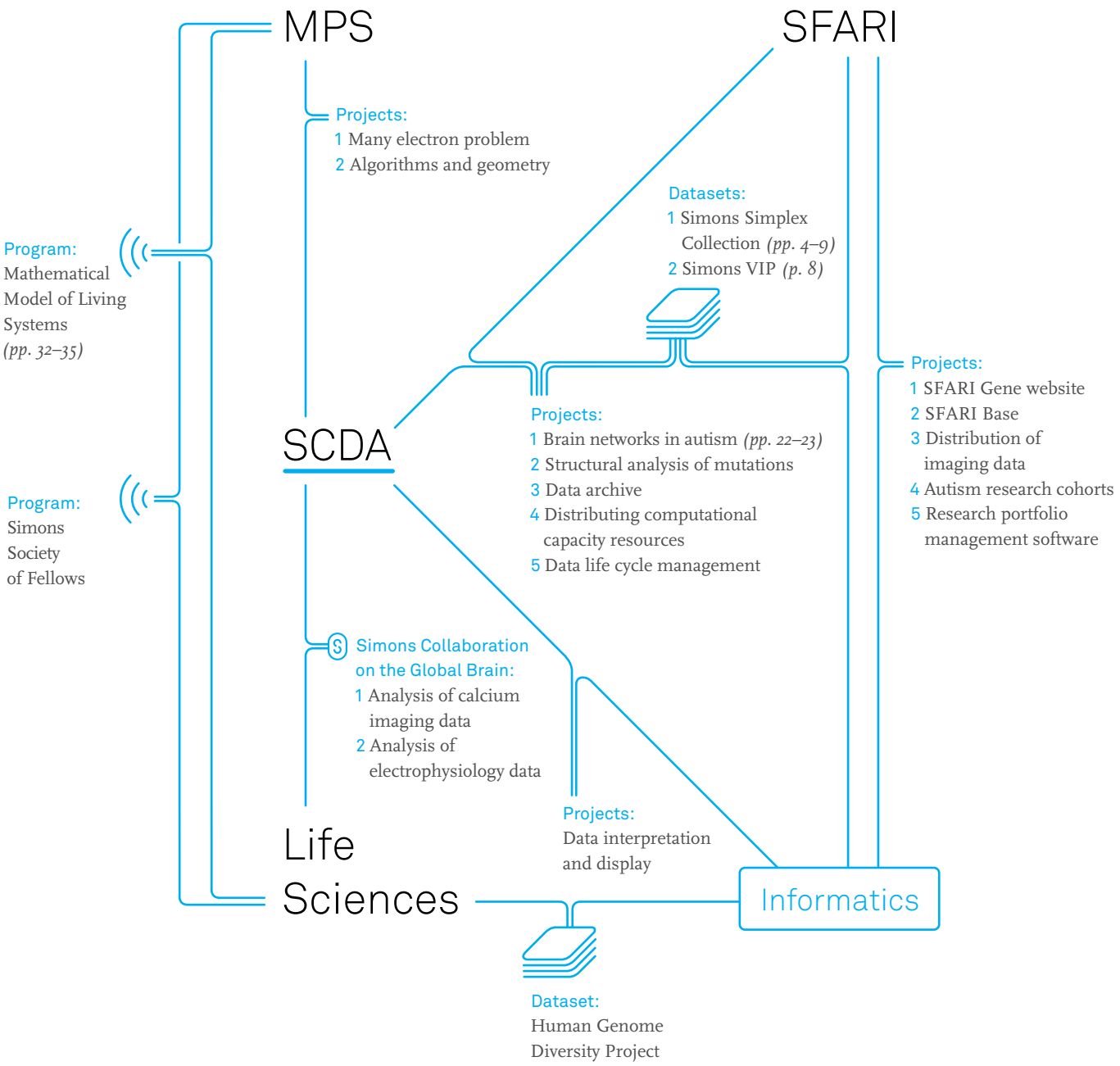
According to Greengard, each of these scientific endeavors requires new analytic approaches, and the development of methods that make sense of the data ('data science') is becoming a discipline in its own right. SCDA requires leaders who have both a deep understanding of the underlying science and the ability to create the mathematical and computational frameworks that will permit scientists to ask previously inaccessible research questions. Going forward, each group at SCDA will typically work closely with external experimental collaborators to analyze their results, which will also drive the center's thinking about new questions to investigate. Greengard intends to be selective about what problems SCDA takes on, because each one could easily consume all of the center's scientific and technical resources. "The reason SCDA exists," he says, "is to develop new ways of thinking about big data in biology and to develop tools that will enhance the productivity of individual researchers."

14
New team members in 2014

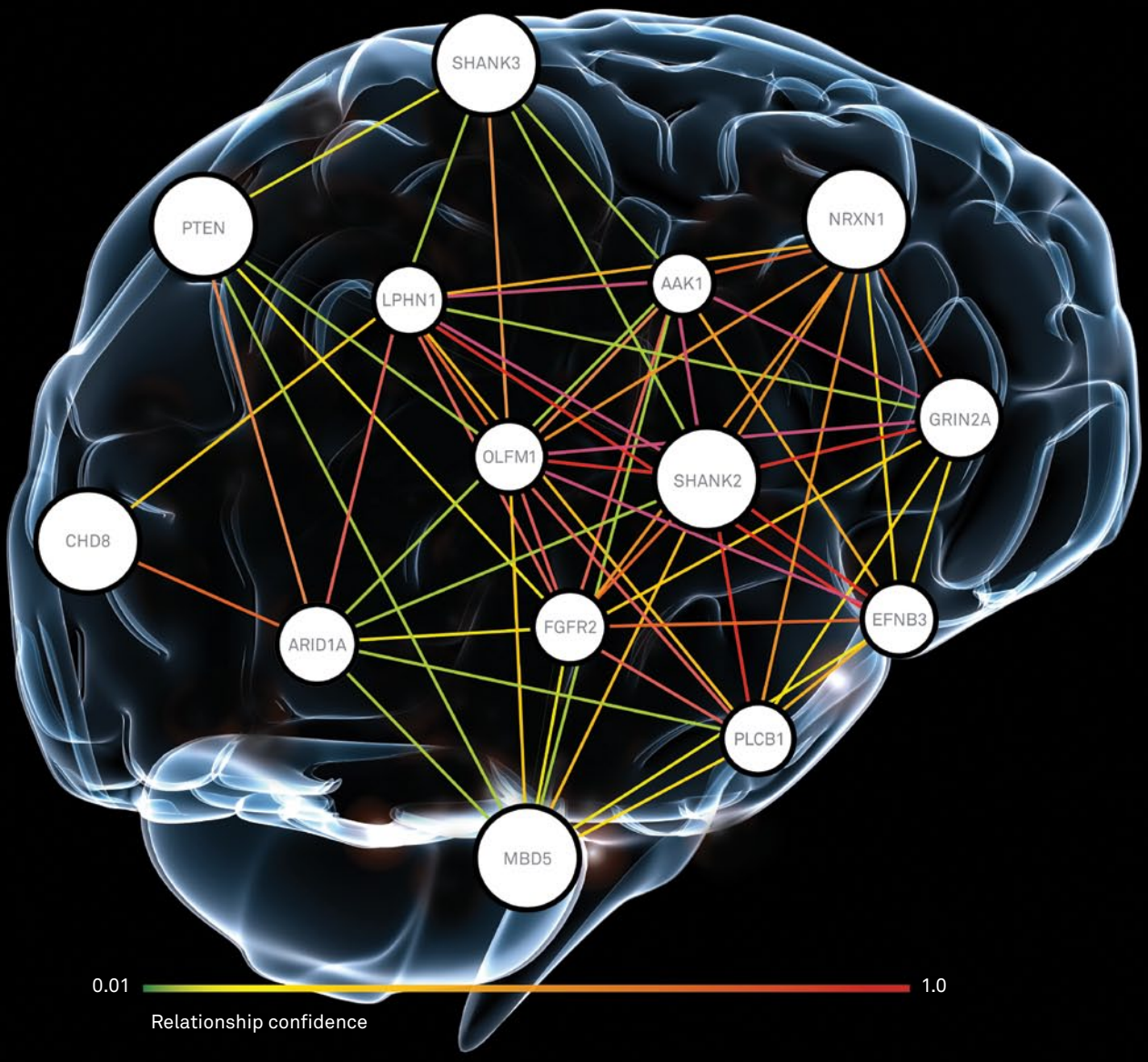
6
Research groups within SCDA

49
External collaborators

In-House Collaborations



↗ This chart shows collaborative projects between working groups at the Simons Foundation in 2014. Some of these efforts are formal grants programs ("Programs"), but many are ongoing collaborations between scientists on staff ("Projects"). SCDA and the Informatics group often assist other foundation groups in the planning, storage, analysis and dissemination of large datasets.



↖

Brain-specific functional gene networks can illuminate the molecular basis of neurodevelopmental disorders. Above, a section of a brain network thought to be relevant to autism.

SCDA: Genomics Group

"Our networked approach is really well suited to unraveling this puzzle, because autism is a networked disorder ..."

Olga Troyanskaya

To Olga Troyanskaya, leader of the genomics group at SCDA, figuring out how to use big data to study complex disorders such as autism or cancer is like trying to develop a Google for genomics. "Before Google and other 'smart' search engines, the Internet was a collection of directories with no clear assessment of quality or relevance, and written in different languages," she says. "Genomic data are even more complicated: They represent hundreds of diseases, tissues and clinical treatments, and are made by more than 50 different technologies. How can one identify and integrate relevant information across all these datasets?"

Troyanskaya's team develops algorithms that can spot similar patterns in gene expression across many different kinds of tissue and disease, regardless of the technology used to gather the data. For example, the same genetic pathways that are important in neurons in the brain also exist in kidney cells, so kidney disease data might actually teach us something about brain disorders such as autism. "It's very counterintuitive," she admits. "It's not based on symptoms or single-gene mutations. It's only algorithmically that you can systematically identify such signals."

This "messy gold mine" of gene expression data, as Troyanskaya calls it, could unlock new understanding of complex disorders by uncovering genetic links that were previously invisible. Any biological experiment inevitably perturbs many different aspects of a cell's function, and the Troyanskaya group's methods put those inevitable extra perturbations to good scientific use by first identifying patterns in these 'noisy' datasets that are useful outside of the original experimental context in which they occurred, and then aggregating these datasets together.

According to Troyanskaya, autism research lends itself especially well to this approach precisely because there is no 'autism gene' to pinpoint in isolation. Instead, autism is a networked disorder whose symptoms are associated with the coordinated behavior of multiple genes. While damage to a single gene can have major impact, this impact is most likely modified by small differences between individuals in expression and function of other genes in the network. Troyanskaya's computational analysis allows every human gene to be ranked based on how likely it is to be associated with autism based on its functional role in the brain's molecular networks.

As her team uncovers these associations in collaboration with SFARI, they also build software that lets other researchers apply the same algorithmic methods to explore other open questions in cell biology and medicine. "We're working across diverse tissues and cell types — looking at large collections of biological data, and figuring out algorithms that are able to isolate the relevant signals in a very accurate way," she says. "The philosophy of my group is that with smart algorithms, more data is always better."

Approximately

38,000

genome-wide experiments integrated

SCDA: Systems Biology Group

> "We don't just want to tell a story about a single particular organism or tissue — we want to tell a story that can generalize to the cell and to the whole organism."
>
> Richard Bonneau

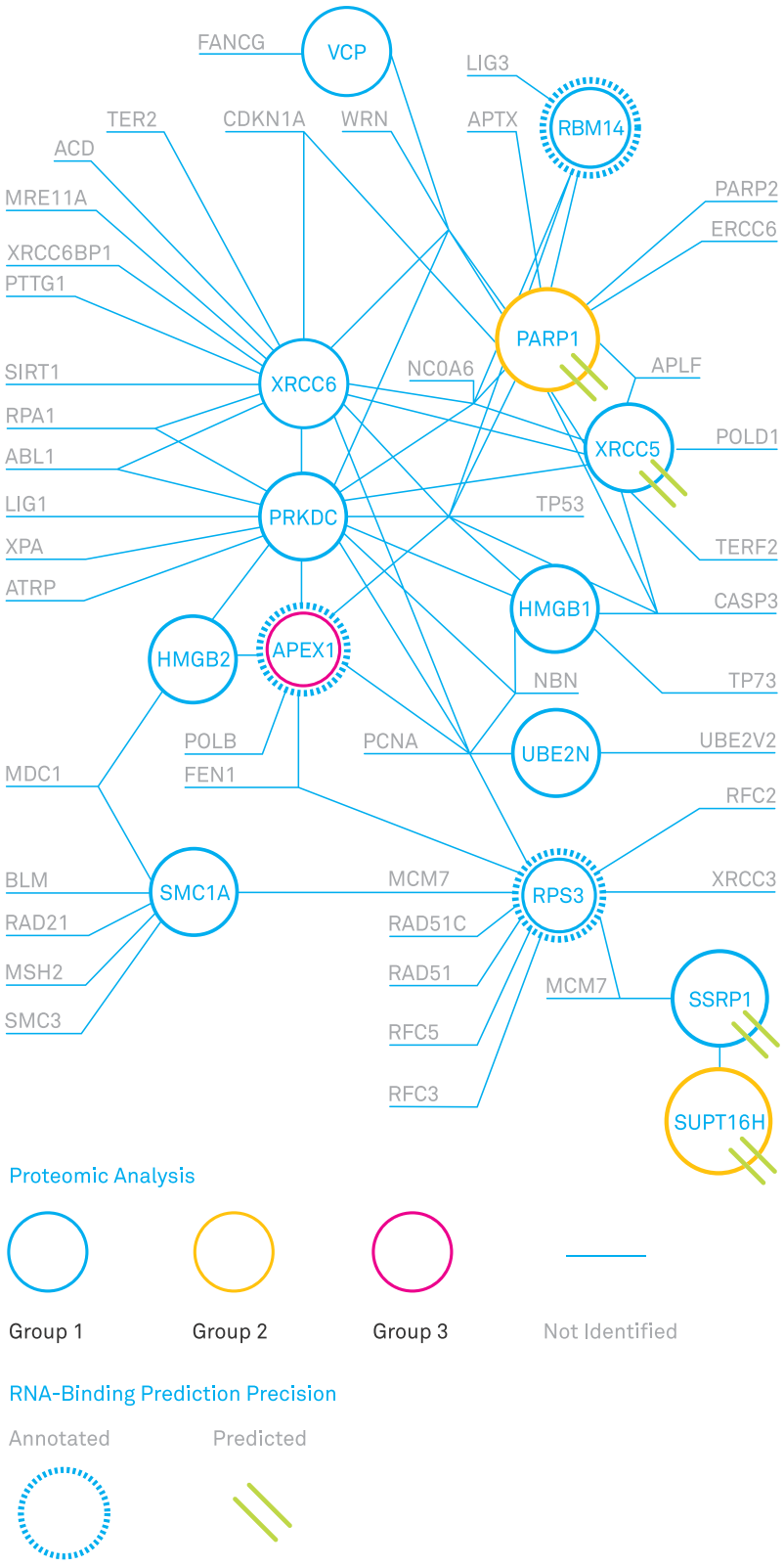
What differentiates 'systems biology' — the discipline that Richard Bonneau's research group at SCDA specializes in — from other types of biological study? All biologists consider the cell to be the basic unit of life, but every living cell also contains an immense amount of molecular machinery of its own, which biologists were unable to model in detail before the advent of modern genomic technology. "Genomics is like getting the parts list and seeing what the parts are doing at any given moment," Bonneau explains, "and systems biology is like putting that information together into a circuit or picture of the machine."

What that view enables is a way to examine the links between the intricate machinery inside cells and the larger dynamics of the human body, such as the immune system and the microbiome — both of which are objects of Bonneau's interest at SCDA. His group aggregates experimental data on the immune system from multiple human, animal and bacterial sources, collected by diverse collaborators such as Dan Littman and Jane Skok at New York University School of Medicine. "When we go after big biological systems with thousands of moving parts, we really need to see what other data is out there and integrate it with our own," Bonneau says. "We don't just want to tell a story about a single particular organism or tissue — we want to tell a story that can generalize to the cell and to the whole organism."

Bonneau collaborates closely with Olga Troyanskaya's genomics group at SCDA. "I'm hoping Olga can teach me how to take some of our reconstruction algorithms and scale them up to 'all data for the immune system' or 'all human data, period,'" he says, only half joking.

"I think our groups are going to come up with some algorithms that are going to be extremely important for the very big and growing community of people who are trying to figure out how the immune system interacts with the microbiome."

Bonneau considers SCDA's interdisciplinary scientific team to be one of the most exciting — and crucial — aspects of his current research. Bringing biologists together with experts in data visualization, applied mathematicians and software engineers allows for a more integrated and efficient approach that is not always available in traditional academic settings. "At SCDA we can approach problems closely as a team without having to market little pieces of what we're doing one project at a time, and getting pulled off in many different directions," he says.



68

Microbiome datasets compiled and analyzed

←

This figure depicts interactions between mRNAs (gray text) and proteins (circles). A central goal of systems biology is to integrate computational and experimental techniques in order to understand how interactions between biological molecules contribute to cellular function. Here, computational methods were used to predict which proteins bind RNA molecules, and those results were validated with mass spectrometry. Double green lines on proteins indicate RNA-binding proteins confirmed through both methodologies. Previously identified RNA-binding proteins are also shown ('haloed'); their presence increases confidence in the network.

SCDA: Neuroscience Group

Ask computational neuroscientist Dmitri Chklovskii what his research goals are at SCDA, and he doesn't mince words. "Our goal is to understand how the brain computes," he says simply. Of course, it's anything but simple. And for Chklovskii's neuroscience group at SCDA, the question is doubly complex: Not only do they intend to crack the brain's computational code by reconstructing the three-dimensional (3-D) 'wiring diagram' of its network of neurons, but they also need to invent novel data-processing techniques just to make sense of the deluge of experimental data generated by contemporary brain imaging technology. The current method for mapping neural connections requires collecting millions of extremely high-resolution two-dimensional images, and then "stacking the slices" together so that the 3-D connections are revealed. "Even for a small animal like a fruit fly, the datasets are humongous," Chklovskii says.

Prior to joining SCDA, Chklovskii's group developed software that significantly accelerated the creation of these neuronal wiring diagrams, or connectomes. The neuroscience group intends to assemble a connectome of the entire nervous system of *Megaphragma mymaripenne*, a microscopic wasp. "It's only 200 microns long, but it can see, smell, fly and find food with a small number of neurons," Chklovskii says. "We hope that by reconstructing the complete nervous system of this insect, we can start to understand how it all works together."

Of course, as Chklovskii says, human beings actually have powerful computers between their ears that can accomplish extremely complex, real-time data analysis. "At any given time, we are bombarded by millions of different signals from the outside world, and we have to react without waiting for the whole dataset, because by that time we would be dead," he says. Chklovskii believes that our brains are biologically implementing some sort of highly efficient algorithm to process the constant stream of 'big data' from our senses. If the insights gained from constructing a connectome reveal how biological brains — even tiny ones like that of *Megaphragma mymaripenne* — accomplish this feat, they might also shed light on how to design more efficient computers that are capable of analyzing the torrent of real-time streaming data soon expected to come from networked sensors embedded in everything from home appliances to civil infrastructure: the so-called 'Internet of things.'

Chklovskii acknowledges that both mysteries may not be solved at once, but he predicts significant growth and value in the synergy between neuroscience and computer science — a synergy that needs a multidisciplinary research center like SCDA in order to flourish. "It's very hard to teach computers to do what humans seem to do effortlessly, and vice versa," he says. "But we can look for some sort of guidance in the data."

→
These fully functional mini-insects are so small that researchers can image their entire nervous systems with unprecedented detail. Mini-insects and single-cell organisms are shown at right using a 200-micron scale. (A single human hair is 100 microns in thickness.) Clockwise from top left: *Nanosella sp.*, *Amoeba proteus*, *Paramecium caudatum*, *Megaphragma mymaripenne* and *Dicopomorpha echmepterygis*.

Credit: Alexei Polilov


200 µm

Simons Collaboration on Algorithms and Geometry

14

Principal investigators

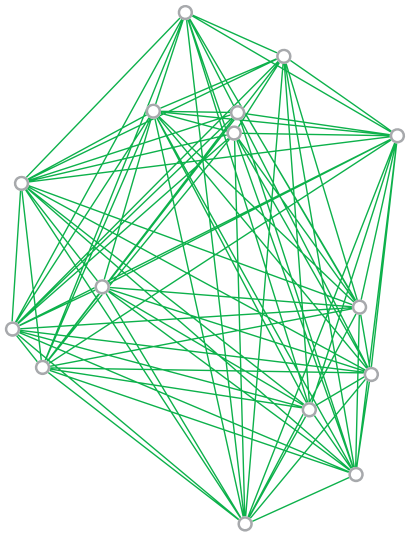
The line between mathematics and theoretical computer science is blurry, but for some the distinction doesn't exist at all. The Simons Collaboration on Algorithms and Geometry, which launched in September 2014, brings together mathematicians and computer scientists who work on a variety of questions at the interface of the two disciplines. Despite the fact that this collaboration has only been up and running for a few months, it has already helped some researchers draw unexpected connections between seemingly disparate problems.

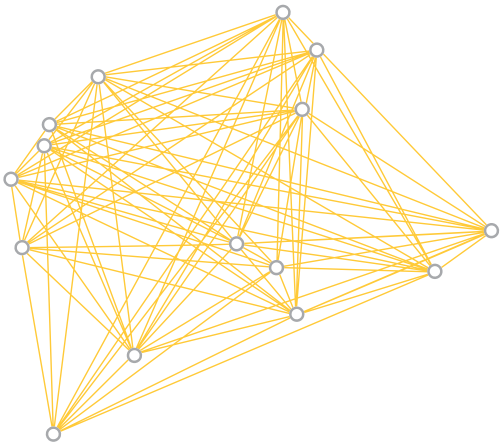
Amit Singer, an applied mathematician at Princeton University, is using cryo-electron microscopy (cryo-EM) to determine the structure of various biologically important molecules. The pictures produced by cryo-EM are noisy, two-dimensional projections taken from random directions. The computational challenge is to determine, from these two-dimensional images, their three-dimensional orientations and the configuration of the molecule pictured.

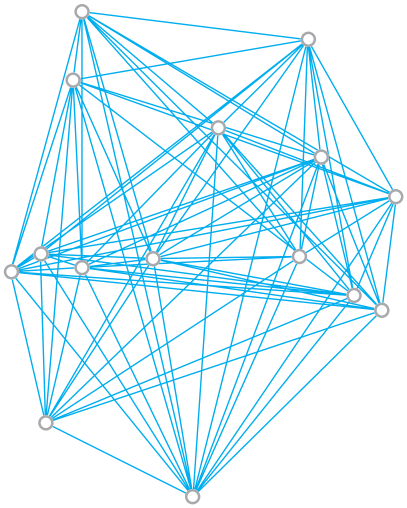
At first glance, this question would not seem to have much to do with, for example, the unique games conjecture, a problem in theoretical computer science originally formulated by fellow collaboration member Subhash Khot, winner of the 2014 Rolf Nevanlinna Prize. The unique games conjecture says that a certain problem in graph theory is difficult not only to solve but even to approximate efficiently, and has deep implications for other questions in theoretical computer science.

Singer reports that with the help of collaborator Moses Charikar, he realized that one of the algorithms that came from his research in cryo-EM was related to the unique games conjecture. "It's a very interesting connection between a very applied problem in structural biology all the way to a very abstract problem in theoretical computer science."

The collaboration is focused on problems and algorithms that transcend specific applications — and even the constraints of modern computing power. For example,





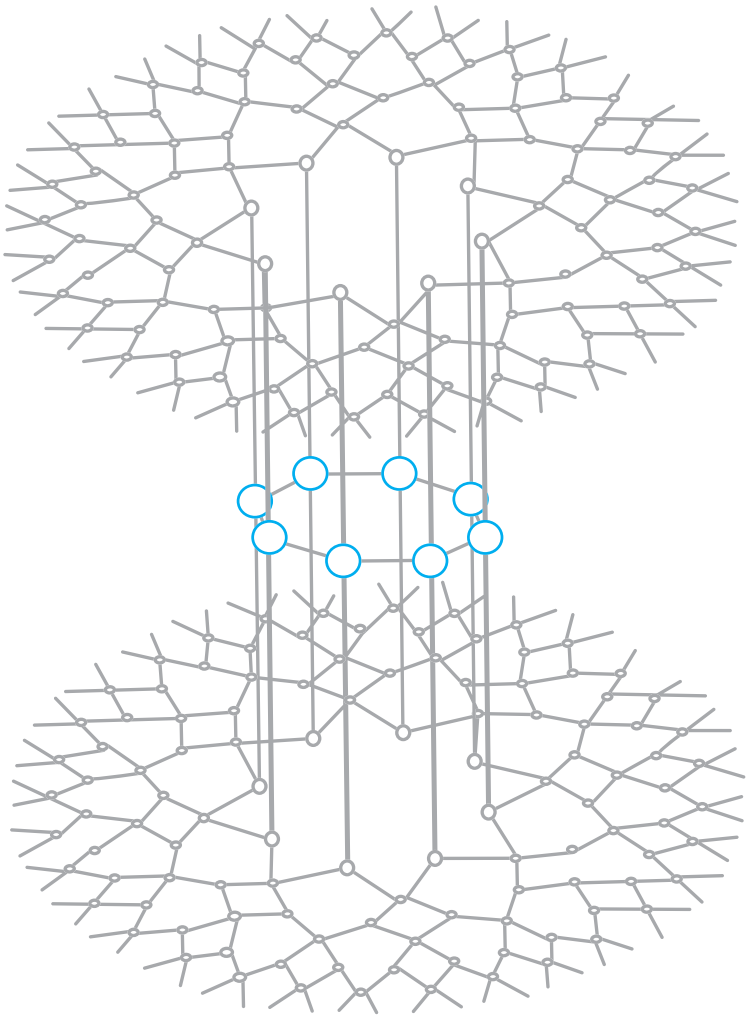


←
This new method for clustering data points into k clusters produces values between 0 and 1 for all pairs of points. These values express the likelihood that both points in the pair are in the same cluster. When data can be partitioned into k clusters with a value of 1 for all pairs inside the same cluster, and a value of 0 for all pairs of points across different clusters, the solution is optimal.

given a collection of objects, with an inherent notion of 'distance' between any two of them, can one store them efficiently so that one could quickly determine which objects are the most similar to others? This question, called the nearest neighbor search, comes up in many guises and applications, from storing images and music to studying how different proteins work. "We're interested in the foundations," collaboration director Assaf Naor says. "We want to come up with general mathematical theorems that will apply to certain specific datasets, but also, when something new comes up, we want to have an arsenal to attack it with."

Naor says the collaboration is currently in the "get to know you" phase, with researchers ramping up with in-depth presentations on their work. After that phase concludes, meetings will become increasingly focused on making progress on specific topics.

Simons Collaboration on the Many Electron Problem



8 Theoretical methods benchmarked

54 Students and postdoctoral fellows registered for 2015 summer school

This drawing depicts a member of a class of mathematical objects called tensor networks, which are used to represent many-body wave functions. This particular tensor network has a dual interpretation: It describes a thermal state of a many-body system, but it may also be interpreted as representing the space-time geometry of a black hole. ←

Why are some things hard and others soft? Why do some materials conduct electricity? And can we make materials that conduct even better? The properties of any material are determined by the interactions between its electrons, and understanding the way such properties arise at the electron level is both a fundamental scientific problem and key to creating new materials with desirable properties. The Simons Collaboration on the Many Electron Problem, which launched in March 2014, brings together physicists and chemists to work to refine existing approaches to this problem, and to invent new approaches.

A great deal of the challenge boils down to big numbers. The number of states available to a many-particle system grows exponentially as the number of electrons grows. The sheer size of the vector space that represents those states means that conventional computational approaches to managing electron-electron interactions fail. "A lot of our work consists of finding clever ways around that exponential barrier," says Emanuel Gull of the University of Michigan, who is a member of the collaboration.

The collaboration brings together four main threads of research: cluster embedding theory, Monte Carlo methods, real material methods and tensor networks. The program hosts several conferences each year in addition to a summer school for graduate and postdoctoral students.

A variety of numerical methods have been developed to attack different aspects of the many electron problem, and the early stages of the collaboration have focused on taking stock of the strengths and weaknesses of those methods by using them on simplified model systems. "The 'fruit fly' of this type of physics is the so-called Hubbard model. That's a model that has all of the physics stripped down, except for one local interaction term and one nearest-neighbor term," says Gull. "It's sort of the minimum model that gives you the physics of electron correlation." In comparing various approaches, researchers try to identify the ways that the different methods might fail in a simple system, in order to understand how to get around the same problems in a more complicated one.

One promising approach is based on tensor networks, a reformulation of the wave function description of quantum mechanics. A breakthrough in that field about ten years ago radically changed the way researchers can attack the many electron problem; in essence, tensor networks provide a way to home in on a small subset of states characterized by certain quantum entanglement conditions. Guifre Vidal, a researcher at the Perimeter Institute for Theoretical Physics and one of the researchers behind the breakthrough, says, "What we have identified is that in this huge vector space, there are many vectors we should not care about, that are not relevant to the many electron problem, and a very small subset — a subset of measure zero — is actually the one we care about." This means that in certain situations, the computational cost has been vastly reduced — it grows as a power of the number of electrons rather than exponentially.

The tensor network breakthrough, when coupled with increases in computing power, means that an understanding of larger and larger systems becomes more feasible. "We are greedy," says Vidal. "There is never a computer that is big enough. We used to complain that we couldn't go beyond 20 electrons. Now we can go to hundreds. Once we go to hundreds, we try to go to thousands. It is not unreasonable to be so greedy. We'd like to study millions. That's what real systems are made of."

The other collaboration research areas have made similar progress in recent years. The logical next steps will be to compare, contrast and combine the approaches, to determine which work best in which situations, and to use insights in one area to improve results in another. For example, the collaboration is exploring the implications of the recent discovery that it is possible to use stochastic ('Monte Carlo') methods to evaluate Feynman diagram series, in the hope that these methods can be used to improve the standard techniques used to provide first order approximations to the properties of molecules and solids.

"A couple of theoretical breakthroughs allowed us to access a new world of physics," says Gull. "That's the spirit of the collaboration."

Mathematical Modeling of Living Systems and Conference on Theory and Biology

Ned Wingreen of Princeton University delivers "Keeping It Together: Organizing the Bacterial Chromosome for Division" at the MMLS Theory and Biology meeting in April 2014.

The tremendous advances in technology over the past few decades allow scientists to collect and analyze much more information than they could in the past. In biology, this means researchers can do things a little differently. Where biologists used to look at fairly sparse data and make general descriptive statements, they can now combine the results of measurements with sophisticated theoretical analysis, uncovering new hypotheses and quantitatively testing them. That's the rationale behind the Simons Foundation's Mathematical Modeling of Living Systems (MMLS) program: to use theoretical and mathematical concepts to develop a quantitative understanding of all aspects of life. To this end, the foundation supports individual researchers through its Investigators in MMLS program, and funds targeted grants to research groups studying particular projects.

132

Participants in 2014 Conference on Theory and Biology

302

Targeted grant proposals received in 2014

The foundation also organizes an annual Conference on Theory and Biology, providing a venue for scientists in the greater New York area to meet, exchange insights and hear talks by leading scientists in the field.

Olga Zhaxybayeva is assistant professor of biological sciences at Dartmouth College, and a Simons Investigator. She sifts through bacterial genomes in an effort to understand their evolution and speciation. Modeling and data analysis go hand in hand in her work. Her group uses sophisticated mathematical models to understand how different traits and conditions influence bacterial populations, and they analyze genome databases to inform those models.

One of the focuses of Zhaxybayeva's research group is to study the history and distribution of genes in microbial communities to determine which ones might effectively differentiate between bacterial species. "In the eukaryotic world, Linnaeus created this beautiful taxonomy where every organism is classified from species to genera to larger groups," she says. "In bacteria, we'd like to think that organisms have some features that allow us to put them into categories that we can call species."

However, unlike plants and animals, bacteria propagate asexually, so the 'can produce fertile offspring' definition of species doesn't work. The obvious choice is to consider genetic similarity instead. Back when there were only three *Escherichia coli* genomes available for study, researchers compared them to determine which traits were present in all three. "We think of them as the same *E. coli*, but really they had only 40 percent of their genes in common," says Zhaxybayeva. "When the study was expanded to about 60 genomes, it turned out that they only shared on the order of 5 percent of their DNA. Now that there are thousands of genomes, it's leveled off, but it's a very small fraction of genes." That means it's difficult to find genes that can effectively categorize the organisms.

Zhaxybayeva's investigations also include research into the evolution of two bacterial traits: tolerance of extreme temperatures and what she describes as cooperation in microbial communities.

The modus operandi of many common viruses is simple: Infect a host, make copies of virus DNA or RNA, kill the host, repeat. But some bacterial populations contain entities called gene transfer agents that behave like viruses but contain bacterial genes instead of their own genome. Like viruses, they kill their hosts, but for some reason, the population continues to host the gene transfer agents. "The hypothesis is that it's some sort of form of cooperation in bacterial populations," says Zhaxybayeva. The question is how the gene transfer agents persist in the population. Researchers think that the bacterial DNA they contain could help 'patch' DNA damaged by ultraviolet radiation or other stressors, or that it could enable the exchange of beneficial genes within the population. "The bacteria population designates a small fraction of cells to 'sacrifice' themselves to produce gene transfer agents," she says.

She and her colleagues model bacterial populations with gene transfer agents and determine how they might maintain the trait, given that it's quite detrimental to individual hosts. On the data-analysis side, they find organisms with particular traits and use sequence comparison techniques to reconstruct the traits' evolutionary histories. While Zhaxybayeva doesn't do experiments herself, she collaborates with

labs that grow bacteria that produce gene transfer agents, to study them on a molecular level. Both that information and her analysis of genome databases then inform later simulations and modeling.

Terry Hwa, professor of physics and biology at the University of California, San Diego, is the recipient of a targeted grant in the MMLS program. Whereas Zhaxybayeva's doctorate is in molecular biology, Hwa's is in theoretical physics. As a physicist, he was most interested in complexity: nonlinear dynamics, turbulence, spin glass, and so on. Over the course of his career, he started moving more toward biology. "You can't close your eyes to the amazing complexity of life," he says. "I was intrigued by living systems and started to poke into it. One thing led to another ..." Today he heads a lab that employs physicists, biologists and applied mathematicians to study the way interactions between molecules in a cell determine the cell's overall behavior.

Earlier in his career, Hwa showed that organisms' responses to environmental and genetic perturbations tend to obey simple mathematical rules, a counterintuitive finding, given how many components a cell has and how complex intracellular interactions are known to be. For example, the number of proteins a cell devotes to nutrient uptake has a negative relationship to the cell's growth rate: The poorer the nutrients, the more uptake proteins are used to retrieve them. "We simplify the description of bacteria to 'a bag of proteins and other small molecules,'" Hwa says. "We then see simple patterns in the macroscopic cellular behavior in selective environments, and can use these patterns to predict behaviors in other environments. Ultimately, these simple behaviors come from the interaction of molecules. Something takes us from the complexity at molecular scale to the simplicity at cellular scale." Hwa is trying to find that something.

Hwa sees a parallel between today's research into the molecular foundations of cellular behavior and the 19th century development of thermodynamics and statistical mechanics. At the time, the ideal gas laws and features of phase transitions were known, but they hadn't been understood in quantitative and molecular terms. Major breakthroughs came when scientists discovered how the behavior of a gas can be described by a number of simple mathematical rules — the laws of thermodynamics — and subsequently how these laws arise mathematically from averaging over many complex molecular collisions. Hwa thinks research such as his that bridges the gap between biology and physics might lead to a similar breakthrough in our understanding of behaviors of cells on a molecular level.

As with Zhaxybayeva, real data inform Hwa's models. His lab grows cells in different environments and takes 'snapshots' of their contents at various points in time. Those snapshots are analyzed to determine how much of different molecules has been produced by the cells or how the cells have grown. The data give Hwa's team a starting point for creating their models, and later give them a way to see whether their models are on track.

"The ultimate goal is to make our model predictive," says Hwa. "We've had some success at the cellular level — we want to make it predictive at the molecular level."

4

Investigators appointed in 2014

Simons Collaboration on Ocean Processes and Ecology

The ocean is alive, through and through, and not just with plants and animals that we can see. Every teaspoon of seawater contains millions of microorganisms that we are just beginning to understand.

Launched in July 2014, the Simons Collaboration on Ocean Processes and Ecology (SCOPE) aims to advance our understanding of the biology, ecology and biogeochemistry of the microbial processes that dominate the global ocean. The collaboration, now composed of 16 investigators working at Station ALOHA, about 60 miles north of Oahu, Hawai'i, in the North Pacific Subtropical Gyre, studies an ecosystem representative of a broad swath of the North Pacific Ocean.

The ocean's billions of microorganisms, which together compose the ocean's microbiome, depend on one another and on the ecosystem as a whole for healthy functioning. Therefore, study of the ocean's microbiome on site is essential. And a combination of recent technological advances is making in situ study of the ocean's microbiome possible.

New robotic submarines, controlled remotely by SCOPE investigators, will enable more in-depth examination of the open ocean and collection of new types of data. Novel genomics technologies will allow faster and more accurate genome sequencing of microorganisms. And advances in computer modeling facilitate management of new and massive amounts of data, which can then be analyzed to create models of how microorganisms may biochemically and ecologically function together in the ecosystem.

"We're trying to extend, with the aid of these new technologies, the development of the field of ocean systems biology," says Edward F. DeLong, who co-directs the collaboration with David M. Karl. "And we're doing this by studying the whole ecosystem of complex microbial communities that drive ocean processes."

Using these new technologies, SCOPE investigators intend to shed light on some fundamental questions of microbial oceanography — how gases are absorbed into the ocean from the atmosphere, how matter is transported to the deep sea, and the roles of microorganisms in fisheries and energy transfer.

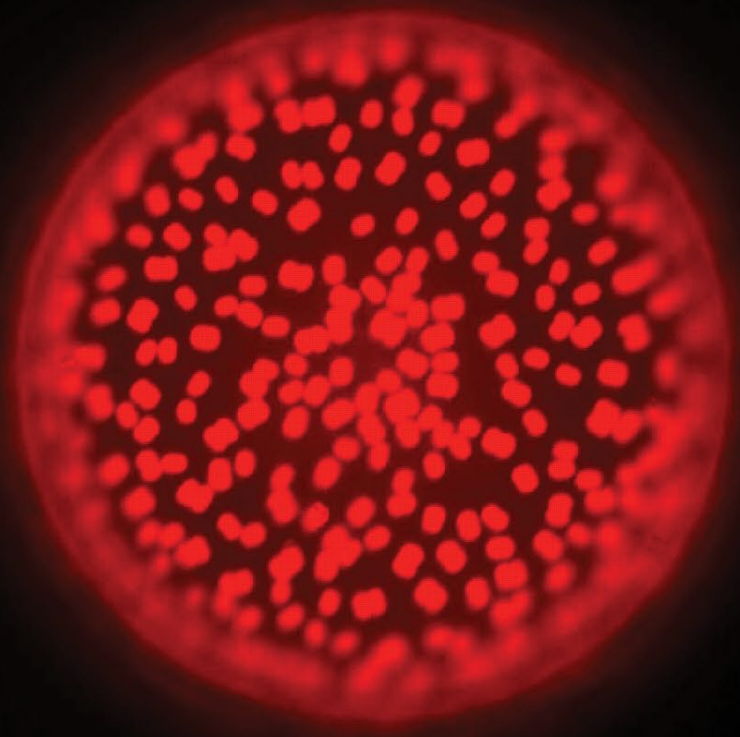
"Having the right scientific tools, being able to gather the right sort of data, and being able to model how these complex systems work is becoming possible for the first time," says DeLong. "Often it's at the interfaces of disciplines that discoveries are made."

To this end, the collaboration brings together experts from diverse disciplines — from scientists who study specific microorganisms to researchers who create mathematical models of organisms' complex interactions — to shed light on fundamental questions of microbial oceanography. SCOPE held its inaugural annual meeting in New York City in December 2014, allowing investigators to meet and learn about one another's areas of expertise. In shaping the collaboration, DeLong, Karl and the SCOPE steering committee selected investigators not only from a variety of disciplines, but also at different career stages. "We wanted to have a structure that is conducive of a long-term approach," says Karl. "These problems are multigenerational, so we wanted a collaboration that is multigenerational as well."

It is hoped that by working in concert, SCOPE investigators will create accurate models of microbial processes. Once these models are developed, the collaboration aims to conduct perturbation experiments of whole ecosystems using robots, ocean sensors and remote sampling. These experiments could enhance researchers' understanding of how human activity affects ocean nutrient levels and many other ecosystem factors.

"Nobody is looking at the microbial processes of the ocean as closely, as intensively or as collaboratively as we are, so we believe the collaboration will be a model for other programs in the future," says Karl. "After a couple of years, we might be rewriting the ocean textbooks."

DeLong adds, "Once you understand that the ocean — Earth's largest biome — is a complex living thing, then it is, from a fundamental knowledge perspective, essential to understand how it works."



↗
Image of the marine diatom of the genus *Coscinodiscus*, found in coastal waters. Diatoms are microscopic organisms that carry out photosynthesis and form the base of most marine food webs.

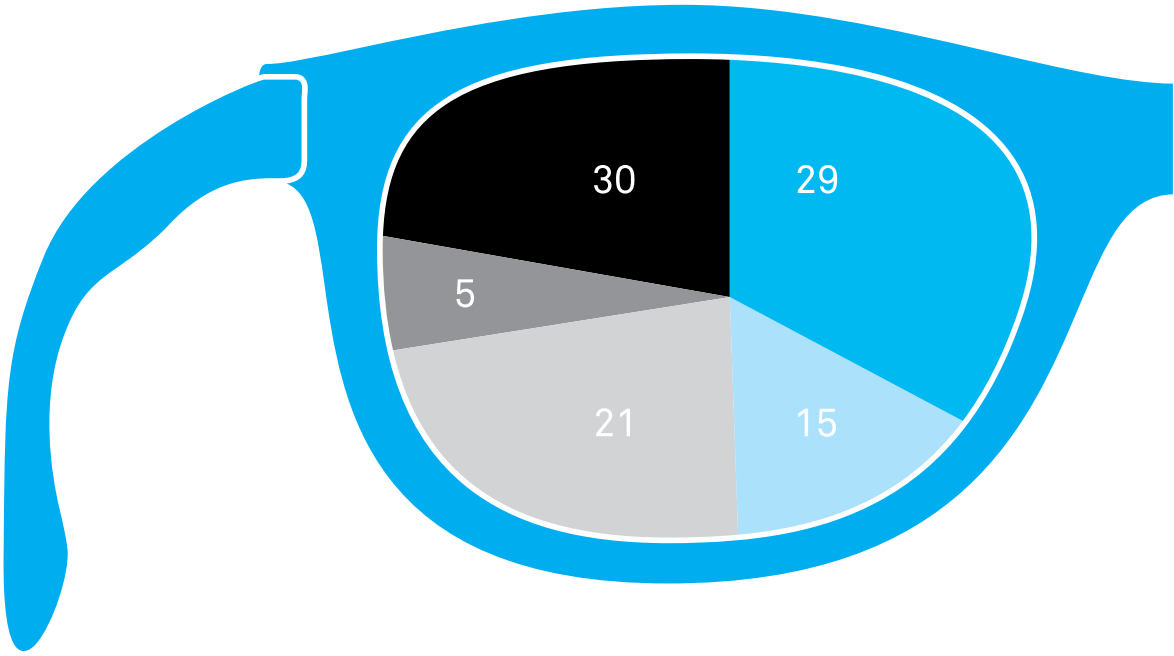
Credit: Armbrust Laboratory

As many as

5 million

microbes per teaspoon of seawater

Essay: Driven by Data



Data play an increasingly important role in our lives these days. Partly the cause is vastly improved technology: We have new tools to collect massive amounts of data, ever-expanding storage capability, and much faster computers (along with sophisticated algorithms) for analysis. But partly the cause is cultural: Our society has chosen to emphasize data (and its fruits, usually some derived statistics) as evidence that is superior to 'mere' observation or judgment. This bias is apparent in the language we use to describe good decision making — it should be data driven, involve metrics, and focus on 'measurable outcomes.' This last phrase was even codified into law by the Government Performance and Results Act of 1993 (GPRA).

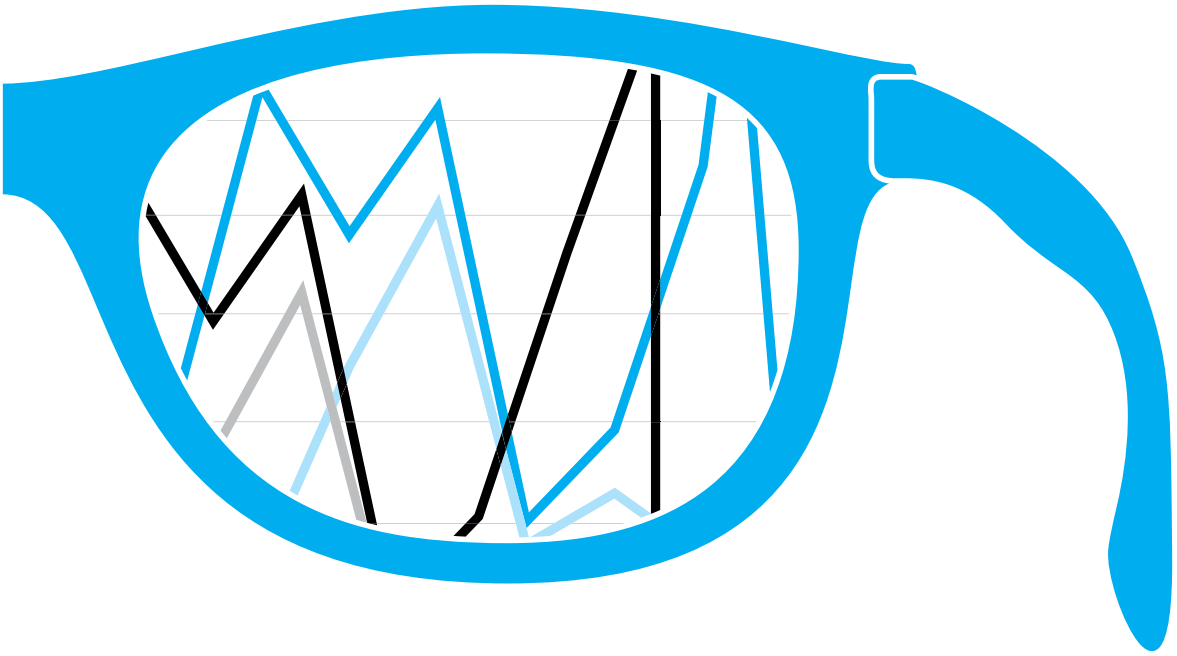
But while society is fascinated by the power of data mining, some people worry about its unanticipated consequences. In a famous 1976 paper, the social scientist Donald Campbell made the observation that the very act of collecting data about human activities may change the data itself: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."[1] A prominent recent example comes from data on college acceptance rates, which *U.S. News & World Report* uses in ranking colleges. While a lower acceptance rate usually indicates a more prestigious institution, colleges can game the system by enticing large numbers of unqualified students to apply, which then artificially lowers their acceptance rate. Other examples of corruption abound, from crime statistics to bestseller lists and television sweeps. When we collect data on human activities, we frequently change the way humans behave.

Around the same time, the Nobel Prize-winning economist Herbert A. Simon noted that data and quantitative measures often do "not even remotely describe the processes that human beings use for making decisions in complex situations."[2] Simon's initial examples concerned public policy, where data and statistics alone cannot answer basic questions (for example, whether it is better to invest in new facilities or in additional staff for public parks). There are many examples of complex social policy decisions that cannot be settled by data, no matter how massive.

These two concerns — the corruption of quantitative measures and the inability of even massive amounts of data to capture social complexity — are particularly worrisome in our current obsession with data driven education. We capture large amounts of data about standardized test scores for many students over many

"Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom."

Clifford Stoll and Gary Schubert



years, and we calculate various statistics based on that data. We use those statistics to measure the quality of teachers, the quality of schools, and efficacy of education policies. But these measures are easily corrupted when educators focus only on the tests or even when they engage in outright cheating. And any measure based on test scores alone captures only a small part of the actual goal of education. Education has suffered because of this. The lesson is not that tests are bad or useless, but rather that education is (and should be) more complex than test score data.

These concerns predate the recent explosion of 'big data,' but they are amplified by that explosion. The sophisticated mathematics we use to analyze vast collections of data sometimes disguises weak data. The elaborate algorithms we use to derive statistics are sometimes based on faulty assumptions or hidden biases. When policy makers uncritically rely on data, but do not understand the mathematics used to analyze it, they can make decisions that appear to be objective and scientific without actually being so. While this is a particular problem in social science, it affects science itself when, for example, the practice of medicine is governed by doctor 'report cards' or the value of research is measured by mysterious numbers derived from citation counts.

Data is indispensable, both in science and social science, but no matter how alluring, numbers are not a priori superior to other forms of evidence. Knowledge, understanding, and wisdom are indeed much more.

John Ewing
President, Math for America

[1]Donald T. Campbell, 1975. Assessing the impact of planned social change. In *Social Research and Public Policies* (G. M. Lyons, Ed.). Hanover, NH: Dartmouth College, Public Affairs Center.

[2]Herbert A. Simon, 1978. Rational decision-making in business organizations. Nobel Memorial Lecture, December 8. Pittsburgh, PA: Carnegie Mellon University.

Math for America

High school algebra teacher Zach Korzyk knew that Math for America (M*f*A) was enhancing his impact as a mathematics educator when, one afternoon in Manhattan last December, his students literally told him so. "Kids were coming up to me after school and saying, 'Mr. Korzyk, that lesson was so fun, but also so interesting,'" he says. "I don't hear that very much."

M*f*A, a nonprofit organization founded in 2004 whose mission is to "make teaching a viable, rewarding and respected career choice for the best minds in science and mathematics," provides fellowships and professional workshops for both early-career and established science, technology, engineering and math (STEM) educators in public schools. Zach Korzyk, who completed M*f*A's entry-level fellowship four years ago and is now an M*f*A Master Teacher, had recently presented at an M*f*A TED-style event focusing on STEM teaching and technological innovation. It was another Master Teacher's presentation during the evening event that inspired Korzyk to design the lesson that so piqued the interest of his students.

"That's the great thing about M*f*A — no other organization exists for teachers to share awesomeness," says Sarah Prendergast, an M*f*A Master Teacher who presented at the event on the topic of paper engineering. "We really are a community of teachers who are not only working together to deepen our knowledge and better our own practices, but also to better the profession in general."

"That's the great thing about M*f*A — no other organization exists for teachers to share awesomeness."

Sarah Prendergast

M*f*A brings expert teachers together into a professional community. Once admitted to the fellowship program, teachers receive a 26-page catalog of evening professional events that offer wide-ranging opportunities for them to stay at the cutting edge, through high-level workshops with research scientists and mathematicians as well as workshops led by M*f*A teachers on best practices in their content area, emerging technologies and ways to deepen teachers' knowledge of their students.

Having opened its Master Teacher fellowships to science teachers in 2013, M*f*A will again expand its reach in 2015 by opening all its fellowships to science teachers — and to elementary school teachers with demonstrated expertise in mathematics or science. According to M*f*A president John Ewing, the programmatic expansion has already greatly increased the intellectual cross-pollination that M*f*A teachers enjoy. "The science teachers want to investigate topics in mathematics and the math teachers are joining the Astronomy Professional Learning Team," Ewing says. "Not only that, the high-school teachers want to talk to the middle-school teachers, and vice versa. These teachers are very accomplished, but they almost never get a chance in their normal workday to talk to teachers from other fields or grade levels in a professional way."

M*f*A offers its fellowships and professional programs — which cover topics ranging from debate techniques in the STEM classroom to programming Arduino microcontrollers — to approximately 800 teachers in New York City and more than 1,000 across the country. In New York City, Ewing expects that group to grow to 1,000 teachers in 2015 — representing about 10 percent of the math and science teachers in New York City.

Sarah Prendergast, for one, is leading the charge. "I want to tell everybody about what I'm doing and help them bring it into their own classrooms," she says. "Where else but M*f*A could we do this?"

↖ Co-facilitated by Master Teachers Eyal Wallenberg and Kate Belin, the Algebra I Modeling Professional Learning Team collaborates to develop methods for exploration and analysis of algebraic functions using models. Photo taken October 2014.


MATH DOESN'T SUCK
DANICA MCKELLAR
KISS MATH

Mathematical Sciences Research Institute

When the Mathematical Sciences Research Institute (MSRI) put on its first exhibition for the public in 1992, few in the surrounding San Francisco Bay Area believed that anyone would pay to come to a theater to learn about math. But William Thurston, who spearheaded the outreach effort as director of MSRI at the time, proved the naysayers wrong. "The venue sold out completely," recalls David Eisenbud, professor of mathematics at the University of California, Berkeley, and current director of MSRI. "They were scalping tickets outside to meet demand."

MSRI's public educational activities have only increased in ambition since that night. As part of an ongoing commitment to expanding public understanding and appreciation of mathematics, MSRI's past events in San Francisco have included appearances by playwright Tom Stoppard and comedians Robin Williams and Steve Martin. Now, a three-year outreach grant from the Simons Foundation has allowed the organization to expand its efforts onto the national and online stages, including the first-ever National Mathematics Festival, held in Washington, D.C., in April 2015; a prize for children's literature related to mathematics; support for *Numberphile*, the most watched informal math channel on YouTube; and support for the documentary *Counting from Infinity: Yitang Zhang and the Twin Prime Conjecture*.

The three-day National Mathematics Festival, produced in partnership with the Institute for Advanced Study, featured discussions of educational policy at the state and federal levels, an event devoted to public and private support of basic research and a day-long public festival, involving interactive math displays, math buskers, and art installations at many of the Smithsonian museums.

The Mathical Prize: Books for Kids from Tots to Teens was created because "we wanted to do for math what *Harry Potter* did for reading," Eisenbud says. The prizes will not be awarded to textbooks or explicitly pedagogical materials, but rather to engaging and playful narrative works. The first winners of the prize were announced at the festival.

MSRI also provides support for the *Numberphile* YouTube series, created and hosted by former BBC reporter Brady Haran. The series surpassed 1 million subscribers and 100 million views in 2014; the popular episode "Mile of Pi," in which Haran printed out the first million digits of pi on a single sheet of paper, was filmed to celebrate this accomplishment. With MSRI's backing, Haran has been able to interview 'heavy hitters' such as John Conway, Barry Mazur and Donald Knuth. "I am just a layman exploring mathematics — and MSRI keeps me in tune with what is happening at the mathematical cutting edge," Haran says.

Eisenbud views MSRI's expanded outreach efforts as a necessary part of improving science, technology, engineering and math (STEM) education in the United States. "Math is much bigger than numbers," he asserts. "It's about structure and pattern. We want to counter the tendency of people to say, 'I always hated math in school,' and feel instead that it's an interesting and necessary part of the modern world and its culture."

←
Claire Huskins enjoys Danica McKellar's *Kiss My Math* in MSRI's library.

↗
Quanta illustrators whimsically depict physicists' search for evidence of an ancient collision with another universe, described in the two-part series "Infinity and Beyond: The Ultimate Test."

Quanta Magazine

Quanta Magazine, the editorially independent online publication the Simons Foundation launched in 2013 to 'illuminate science,' continues to expand in size, readership and ambition, covering advances in fundamental research in mathematics and the physical and life sciences. Maintaining its mission to offset what editor-in-chief Thomas Lin calls "a serious gap in mainstream media coverage of math and basic science research," *Quanta* produces in-depth feature stories on subjects such as the mysterious universal statistical law called the Tracy-Widom distribution, recent breakthroughs in understanding prime number gaps and tantalizing new ideas about the underlying physics that could drive the origin and evolution of life.

The magazine's most ambitious editorial project of 2014 also turned out to be among its most popular: a series of long-form features profiling the year's four Fields medalists and the Rolf Nevanlinna Prize winner. The Fields Medal is informally considered to be the "Nobel Prize of mathematics," Lin says, "but I was aware that none of the major publications would cover it in a substantive way. That left things wide open for us — not only to portray the mathematicians themselves, who are some of the top minds in the field today, but also to describe their work and bring it to a larger audience." *Quanta* spent months planning its coverage, and it paid off: As the most in-depth profiles of this Fields Medal class and their work, the series was widely shared on social media and referenced by the popular websites of NBC News, *The New York Times*, *Business Insider*, *Der Spiegel* and the *Financial Times*. *Quanta* contributor and former math professor Erica Klarreich's profile of topologist Maryam Mirzakhani — the first woman to ever win the Fields Medal — proved especially popular.

Quanta also built on 2013's five-part big data series with ongoing reporting intended to expand mainstream understanding of the topic. "Scientists know that so-called 'big data' is not based purely on the size of a dataset, but also on its complexity or dimensionality," Lin explains. "There are smaller datasets that are highly complex." Whereas mainstream media tends to cover the topic from a technological point of view, Lin says that an ongoing theme of *Quanta*'s coverage is to show how meeting the challenge of big data requires collaboration between multiple scientific disciplines — from the experimentalists gathering field data in unprecedented quantities, to the mathematicians modeling novel patterns in them and the computer scientists building efficient algorithms to process them. "We tried to explain that just having more data is not helpful if you can't make sense of it," Lin says.

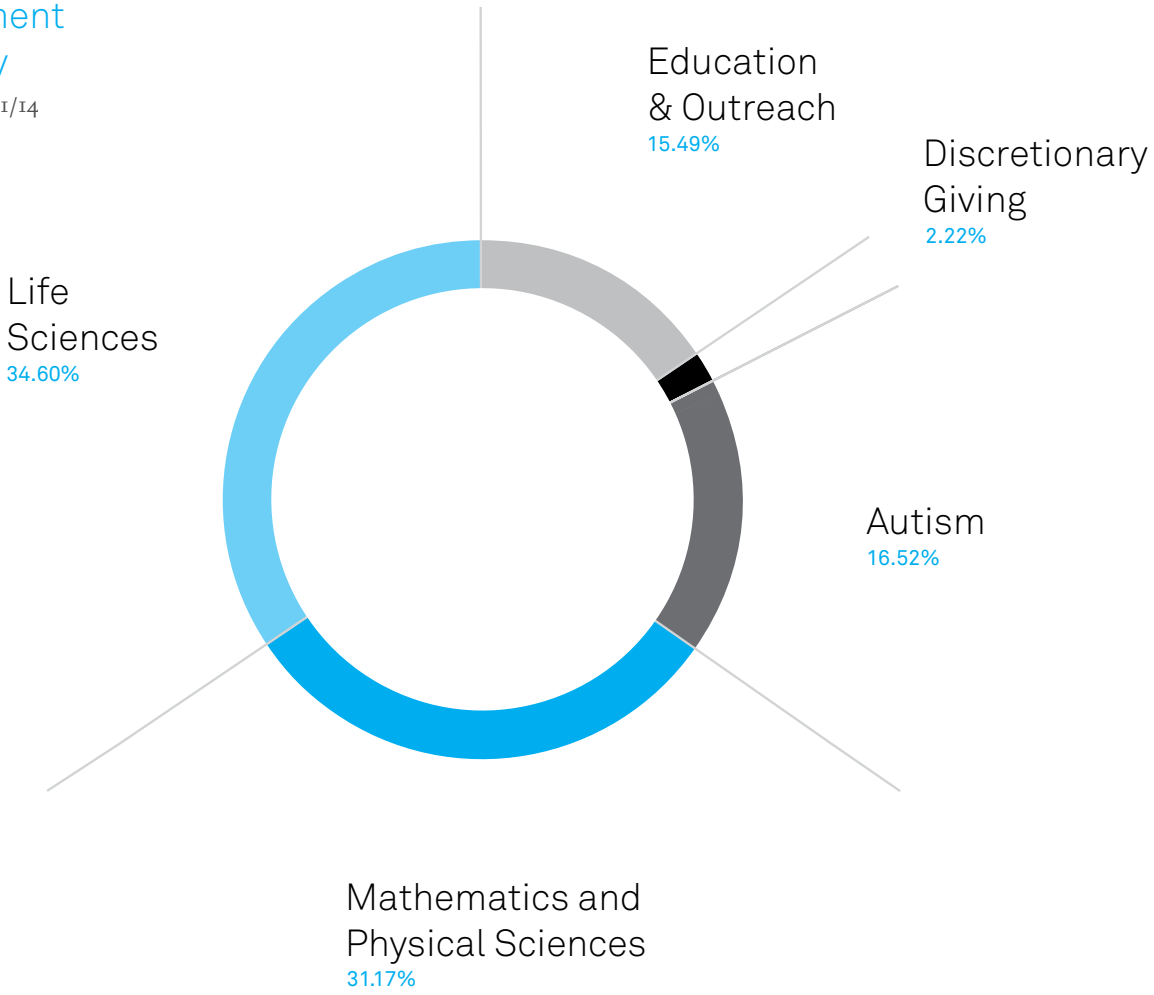
Quanta expanded its staff with two new hires: Deputy editor Michael Moyer, who previously oversaw award-winning physics and space coverage at *Scientific American*, joined this year, as did Olena Shmahalo, a former advertising art director. Moyer now spearheads *Quanta*'s expanding life science coverage. Shmahalo leads *Quanta*'s efforts to enhance the publication's photography and illustrations; she also commissions and designs custom infographics that provide another way for readers to understand difficult theoretical concepts.

Lin's plans for the coming year include video content to enhance engagement with its written reporting. "This is why *Quanta* exists," Lin says. "We have the expertise to do this kind of coverage well, and it does make an impact."

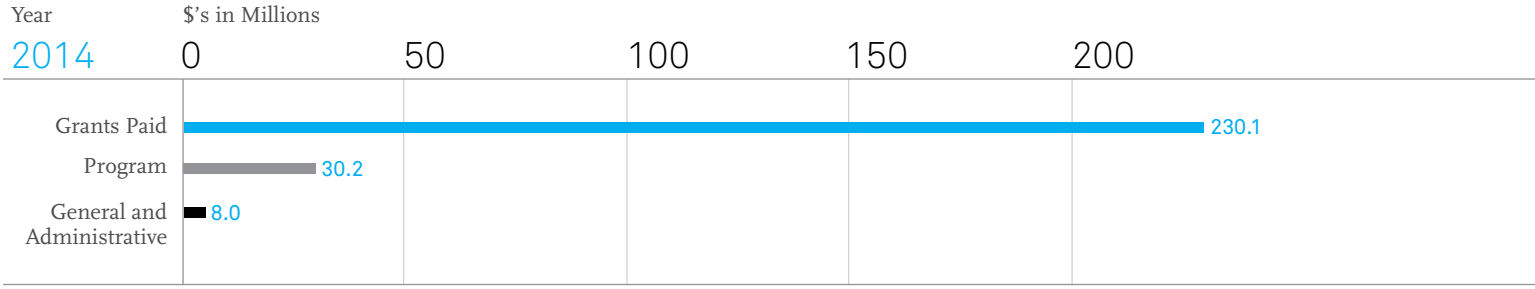
Financials

Grant Payment by Category

For year ended 12/31/14



Proportions of Expenses

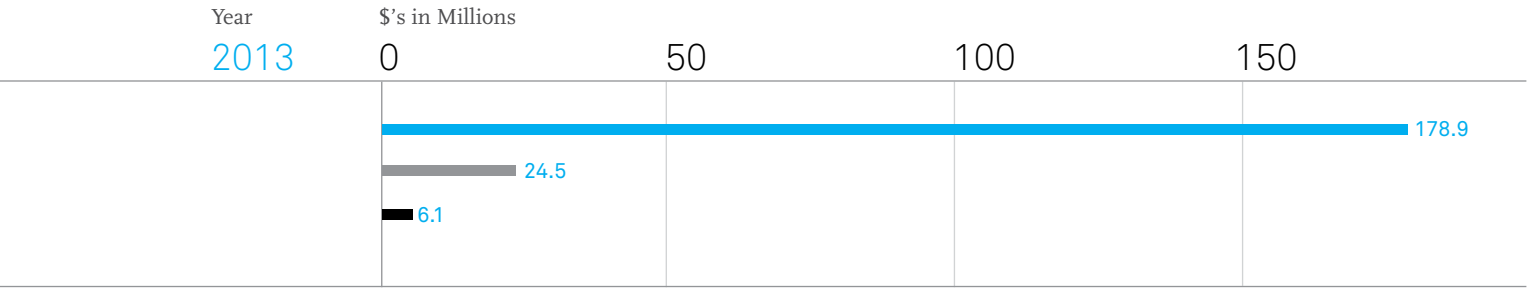


Balance Sheet

Assets	12/31/14	12/31/13
Cash and Cash Equivalents	80,565,539	72,433,180
Investment Portfolio	2,232,380,477	2,079,173,547
Property and Equipment, Net	20,702,904	22,673,961
Prepaid Excise Taxes	76,177	-
Other	3,440,582	3,243,297
Total	2,337,165,679	2,177,523,985
Liabilities		
Accounts Payable	5,221,969	4,115,822
Deferred Rent Liability	4,651,675	5,059,369
Grants Payable	387,213,114	345,316,452
Deferred Excise Tax Liability	14,435,381	12,656,470
Total	411,522,139	367,148,113
Net Assets	1,925,643,540	1,810,375,872

Income Statement

Revenue	For 12 months ended 12/31/14	For 12 months ended 12/31/13
Contributions	23,587,305	84,000,000
Investment Income	404,352,556	227,998,199
Total	427,939,861	311,998,199
Expenses		
Grants Paid	230,069,768	178,889,844
Change in Grants Payable	40,463,200	(69,469,677)
In-Kind Donation	683,877	750,538
Program	30,185,970	24,479,936
General and Administrative	5,763,238	3,660,475
Depreciation and Amortization	2,233,780	2,459,763
Taxes	3,406,303	2,698,249
Other (Income) / Expense	(133,945)	(48,352)
Total	312,672,191	143,420,776
Net Income	115,267,670	168,577,423



Simons Investigators

Life Sciences Investigators

Larry Abbott
Misha Ahrens
David Amaral
David Anderson
Dora Angelaki
E. Virginia Armbrust
William Bialek
Donna Blackmond
Tanja Bosak
David Brainard
Dieter Braun
Carlos Brody
Matteo Carandini
David A. Caron
Irene Chen
E. J. Chichilnisky
Penny Chisholm
Matthew Church
Anne Churchland
Marlene Cohen
John Cunningham
Edward DeLong
James DiCarlo
Brent Doiron
Shaul Druckmann
Jason Dworkin
Sonya Dyhrman
Uri Eden
Michael Follows
Loren Frank
Jeremy Freeman
Stefano Fusi
Surya Ganguli
Lisa Giocomo
Mark Goldman
Wayne Goodman
Kenneth Harris
Fritz Henn
Anitra Ingalls
Rudolf Jaenisch
Mehrdad Jazayeri
Seth John
Gerald Joyce
Lisa Kaltenegger
David Karl
Alla Karpova
Roozbeh Kiani
Ramanarayanan Krishnamurthy
Brian Lau
Andrew Leifer
Debbie Lindell
Michael Long
Zachary Mainen
Sheref Mansy
Valerio Mante
Markus Meister
J. Anthony Movshon
Dianne Newman
William Newsome
Karin Öberg
Svante Pääbo
Liam Paninski
Joseph Paton
Nick Patterson
Fernando Pérez
Bijan Pesaran
Jonathan Pillow
Xaq Pitkow
Alexandre Pouget
Matthew Powner
John Pringle
Stephen Quake
Didier Queloz
David Reich
Daniel Repeta
Fred Rieke
Gene Robinson
Bernardo Sabatini
Nita Sahai
Maneesh Sahani
Dimitar Sasselov
Spencer Smith
Haim Sompolinsky
Michael Stryker
Lisa Stubbs
Roger Summons
John Sutherland
Karel Svoboda
Jack Szostak
David Tank
Doris Tsao
Benjamin Van Mooy
Brian Wandell
Joshua Weitz
Angelicque White
George Whitesides
Jonathan Zehr
Manuel Zimmer
Steven Zucker

SFARI Investigators

Ted Abel
Ralph Adolphs
John Allman
Dora Angelaki
Manuel Ascano
Naama Barnea-Goraly
Peter Barrett
Michiel Basson
Helen Bateup
Mark Bear
Carrie Bearden
Esther Becker
Marlene Behrmann
Yehezkel Ben-Ari
Raphael Bernier
Randy Blakely
Thomas Bourgeron
Randy L. Buckner
Jessica Cardin
William Catterall
Aravinda Chakravarti
Steven Chance
Chinfei Chen
Anjen Chenn
Benjamin Cheyette
Yun-Beom Choi
Wendy Chung
Edwin H. Cook, Jr.
Giovanni Coppola
Joseph Corbo
Rui Costa
Eric Courchesne
Christopher Cowan
Gerald Crabtree
Michael Crair
Jacqueline Crawley
Joseph Cubells
Mark Daly
Robert Darnell
Sandeep Datta
Karl Deisseroth
Orrin Devinsky
Scott Dindot
Catherine Dulac
Robert Edwards
Evan Eichler
Ype Elgersma
Mayada Elsabbagh
Michela Fagiolini
Jin Fan
W. Andrew Faucett
Daniel Feldman
Guoping Feng
Andre Fenton
Gordon Fishell
Loren Frank
William Gaetz
Daniel Geschwind
Charles Gilbert
Cecilia Giulivi
Joseph Gleeson
Robin Goin-Kochel
Mitchell Goldfarb
Matthew Goodwin
Alessandro Gozzi
Ann Graybiel
Christopher Gregg
Adam Guastella
Abha Gupta
James F. Gusella
Joachim Hallmayer
Christian Hansel
Ellen Hanson
Christopher Harvey
David Heeger
Michael Higley
Mady Hornig
Z. Josh Huang
Kimberly Huber
Richard L. Huganir
Rudolf Jaenisch
Daoyun Ji
Peng Jin
Zsuzsanna Kaldy
Joshua Kaplan
Nicholas Katsanis
Raymond Kelleher
Tal Kenet
Jonathan Kipnis
Eric Klann
Abba Krieger
Kenneth Kwan
Anthony Lamantia
Gary Landreth
David Ledbetter
Charles Lee
Christa Lese Martin
Pat Levitt
Ellen Li
Paul Lipkin
W. Ian Lipkin
Dan Littman
Catherine Lord
Liqun Luo
Robert C. Malenka
Dara Manoach
Oscar Marin
Sarkis Mazmanian
A. Kimberley McAllister
Steven McCarroll
Mollie Meffert
Sunil Mehta
Vinod Menon
Carolyn Mervis
Daniel Messinger
Judith Miles
Alea Mills
Partha Mitra
Eric Morrow
Scott Murray
Mor Nahum
James Noonan
Brian O'Roak
Lucy Osborne
Sally Ozonoff
Theo Palmer
Alex Parker
Karen Parker
Elior Peles
Kevin Pelphrey
Bradley Peterson
Deirdre Phillips
Ben Philpot
Joseph Piven
Michael Platt
Carlos Portera-Cailliau
Ning Qian
Indira Raman
Irving Reti
Alexandre Reymond
Timothy Roberts
Kathryn Roeder
J. Amiel Rosenkranz
John Rubenstein
Uwe Rudolph
Shasta Sabo
Mustafa Sahin
Stephan Sanders
Celine Saulnier
Elad Schneidman
Robert Schultz
Ethan Scott
Marco Seandel
Jonathan Sebat
Nenad Sestan
Carla Shatz
Stephen Shea
Elliott Sherr
Songhai Shi
Sagiv Shifman
Lawrence Shriberg
Matthew Siegel
Steven Siegel
James Sikela
Alison Singer
Jeffrey Singer
Pawan Sinha
Jesse Snedeker
Vikaas Sohal
Hongjun Song
Sarah Spence
Matthew State
Beth Stevens
Thomas Südhof
David Sulzer
Mriganka Sur
James Sutcliffe
Francis Szele
Michael Talkowski
Nien-Pei Tsai
Thomas Tuschl
Nathan Urban
Roger Vaughan
Dennis Vitkup
Dennis Wall
Mark Wallace
Christopher Walsh
Fan Wang
James Waschek
Rachel Wevrick
Michael Wigler
Mark Wu
Ofer Yizhar
Xiaobing Yuan
Konstantinos Zarbalis
Mark Zervas
Chaolin Zhang
Larry Zipursky
Huda Zoghbi
Lonnie Zwaigenbaum

Simons Fellows

Mathematics & Physical Sciences Investigators

Igor Aleiner
Rajeev Alur
Sanjeev Arora
Tim Austin
François Baccelli
Manjul Bhargava
Mark Braverman
Michael Brenner
Emmanuel Candès
Garnet Chan
Moses Charikar
Ngô Bảo Châu
Jeff Cheeger
Alex Eskin
Tony Ezome
Paul François
Victor Galitski
Antoine Georges
Sharon Glotzer
Shafi Goldwasser
Daniel Gottschling
Alice Guionnet
Emanuel Gull
Larry Guth
Christopher Derek Hacon
Oskar Hallatschek
Patrick Hayden
Christopher Hirata
Terence Hwa
Russell Impagliazzo
Piotr Indyk
Randall Kamien
Marc Kamionkowski
Charles Kane
Brian Keating
Richard Kenyon
Subhash Khot
Jon Kleinberg
Bruce Kleiner
Gabriel Kotliar
Evgeny Kozik
Adrian Lee
Edward Lungu
Pankaj Mehta
Maryam Mirzakhani
Joel Moore
Elchanan Mossel
Assaf Naor
Andrei Okounkov
Hirosi Ooguri
Frans Pretorius
Nikolai Prokofev
Eliot Quataert
Leo Radzihovsky
Ran Raz
Oded Regev
Michael Saks
Diaraf Seck
Paul Seidel
Boris Shraiman
Amit Singer
Rachel Somerville
Dam Thanh Son
Yun S. Song
Kannan Soundararajan
Dan Spielman
Anatoly Spitkovsky
David Steurer
Iain Stewart
Terence Tao
Daniel Tataru
Shang-Hua Teng
Senthil Todadri
Salil Vadhan
Mark van Schilfgaarde
Guifre Vidal
Steven White
Horng-Tzer Yau
Xi Yin
Shiwei Zhang
Olga Zhaxybayeva

Mathematics & Physical Sciences Fellows

Theoretical Physics

Zvi Bern
Eric G. Blackman
Piers Coleman
Keshav Dasgupta
Benjamin Davidovitch
Andrei Derevianko
Charles Doering
Glennys Farrar
Paul Ginsparg
Michael Hermele
Andrea Liu
Liam McAllister
Sonia Paban
Christopher Reynolds
Anders Sandvik
David Vanderbilt
Thomas Weiler

Mathematics

Denis Auroux
Jason Behrstock
Roman Bezrukavnikov
Francis Bonahon
Samuel R. Buss
Daniela Calvetti
Mark Andrea de Cataldo
Guang Cheng
Tim Cochran
Donatella Danielli
Giovanni Forni
Dan Freed
Alexander Furman
William Goldman
Shelly Harvey
Lizhen Ji
Svetlana Jitomirskaya
Joel Kamnitzer
Martin Kassabov
Ludmil Katzarkov
Dmitry Kleinbock
Bruce Kleiner
Elena Kosygina
Sándor Kovács
Slava Krushkal
George Lusztig
Ivan Mirkovic
Kartik Prasanna
Firas Rassoul-Agha
Romyar Sharifi
Eric Shea-Brown
Scott Sheffield
Robert Sims
Pham Huu Tiep
Kari Vilonen
Daqing Wan
Mu-Tao Wang
Lauren Williams
Christopher Woodward
Michael Yampolsky

Life Sciences Fellows

Simons Collaboration on the Origins of Life Postdoctoral Fellows

Elizabeth Bell
Ziwei Liu
Irena Mamajanov
Kartik Temburnikar
Falk Wachowius
Dmitry Zubarev

Simons Fellows of The Jane Coffin Childs Memorial Fund for Medical Research

Joseph Castellano
Tina Han
Duncan Leitch
Joshua Modell

Simons Fellows of the Helen Hay Whitney Foundation

Eleanor Clowney
Tania Lupoli
Jeremy Rock

Simons Fellows of the Life Sciences Research Foundation

Aakash Basu
Rogier Braakman
Robert Jinkerson
Wei-Ting Liu
James Marshel
Constance Richter
Caroline Runyan
Ophelia Venturelli
Joshua Weinstein
Rayka Yokoo

Education & Outreach Support

100kin10
BioBus & Base Harlem
Center for Talented Youth, Johns Hopkins University
Charlie Rose Brain Series, Thirteen WNET
Joe's Big Idea, NPR
Mathematical Association of America
Mathematical Sciences Research Institute
MIT Science Museum, Science Festival Alliance
National Academy of Sciences
NautilusThink, Inc.
Posse Foundation, Inc.
Science Friday Inc.
STEM Summit Junior Achievement
STEM Teachers NYC
Stony Brook Simons Summer Research
Stony Brook Undergraduate Research
Strategic Education Research Partnership
Two Turtle Productions, Inc.
WNYC Health Project

Supported Institutions

Cold Spring Harbor Laboratory
Columbia University
Institut des Hautes Études Scientifiques
Institute for Advanced Study
Marine Biological Laboratory
Massachusetts Institute of Technology
Math for America
Mathematical Sciences Research Institute
Museum of Mathematics
New York Genome Center
New York Structural Biology Center
New York University School of Medicine
Rockefeller University
Science Festival Foundation
Stony Brook University
University of California, Berkeley

Advisory Boards

Mathematics and Physical Sciences Scientific Advisory Board

Gérard Ben Arous, Ph.D.
New York University

Susan Coppersmith, Ph.D.
University of Wisconsin-Madison

Charles L. Epstein, Ph.D.
University of Pennsylvania

Igor Klebanov, Ph.D.
Princeton University

Barry Mazur, Ph.D.
Harvard University

Dusa McDuff, Ph.D.
Barnard College

Michael Sipser, Ph.D.
Massachusetts Institute of Technology

Éva Tardos, Ph.D.
Cornell University

Scott Tremaine, Ph.D.
Institute for Advanced Study

Avi Wigderson, Ph.D.
Institute for Advanced Study

SFARI Scientific Advisory Board

Cori Bargmann, Ph.D.
Rockefeller University

Tobias Bonhoeffer, Ph.D.
Max Planck Institute for Neurobiology

David Lewis, M.D.
University of Pittsburgh

Richard P. Lifton, M.D.
Yale University School of Medicine

J. Anthony Movshon, Ph.D.
New York University

Eric Nestler, M.D., Ph.D.
Mount Sinai School of Medicine

Martin Raff, M.D.
University College London

Arnon Rosenthal, Ph.D.
Alector LLC

Elizabeth Spelke, Ph.D.
Harvard University

Huntington F. Willard, Ph.D.
Marine Biological Laboratory

SFARI.org Advisory Board

Monica Coenraads, M.B.A.
Rett Syndrome Research Trust

Michael E. Goldberg, M.D.
Columbia University

Laura Helmuth, Ph.D.
Slate

Robin Marantz Henig, M.S.
The New York Times Magazine

Gary Marcus, Ph.D.
New York University

Ivan Oransky, M.D.
MedPage Today

Isabelle Rapin, M.D.
Albert Einstein College of Medicine

David Sassoon, M.S.
InsideClimate News

Will Talbot, Ph.D.
Stanford University

Quanta Magazine Advisory Board

Sean B. Carroll, Ph.D.
University of Wisconsin-Madison

Laura Chang
The New York Times

Benedict H. Gross, Ph.D.
Harvard University

Hopi E. Hoekstra, Ph.D.
Harvard University

Vincent Racaniello, Ph.D.
Columbia University

Howard Schneider, M.S.
Stony Brook University School of Journalism

Steven Strogatz, Ph.D.
Cornell University

Michael S. Turner, Ph.D.
University of Chicago

Leslie B. Vosshall, Ph.D.
Rockefeller University

Board of Directors

President, Simons Foundation
Marilyn Hawrys Simons, Ph.D.

Marilyn Hawrys Simons is president of the Simons Foundation. Under her leadership the foundation has grown to become one of the country's leading private funders of basic science, and she is an advocate nationally for the increased involvement of philanthropy in funding basic science. Simons is also vice president of the board of Cold Spring Harbor Laboratory, a board member of the Turkana Basin Institute, treasurer of the board of the Learning Spring School and a member of the board of trustees of the East Harlem Scholars Academy. She received a B.A. and a Ph.D. in economics from Stony Brook University.

Director,
Mathematical Sciences
Research Institute
David Eisenbud, Ph.D.

David Eisenbud is director of the Mathematical Sciences Research Institute in Berkeley, California. Previously, Eisenbud was director of the Mathematics and Physical Sciences division at the Simons Foundation. A former president of the American Mathematical Society, Eisenbud serves on the board of Math for America and is a member of the U.S. National Committee of the International Mathematical Union. In 2006, he was elected a fellow of the American Academy of Arts and Sciences. Eisenbud holds a Ph.D. in mathematics from the University of Chicago and has been on the faculty at the University of California, Berkeley, since 1997.

Chief Scientist and
Fellow of the Simons
Foundation
Gerald D. Fischbach, M.D.

Gerald D. Fischbach joined the foundation in 2006 to oversee SFARI and is now the foundation's chief scientist and fellow. He was formerly dean of the faculty of health sciences at Columbia University and director of the National Institute of Neurological Disorders and Stroke at the National Institutes of Health (NIH). Fischbach began his research career at the NIH and later served on the faculty of Harvard Medical School, where he became chair of the neurobiology department, a position he also held at Massachusetts General Hospital. Fischbach was also head of the department of anatomy and neurobiology at Washington University School of Medicine. He was a nonresident fellow of the Salk Institute for more than 20 years. Fischbach's research has focused on trophic interactions between nerve cells and the targets they innervate.

Executive Vice President and
Chief Financial Officer,
Renaissance Technologies
Mark Silber, J.D., M.B.A.

Mark Silber is executive vice president and chief financial officer of Renaissance Technologies LLC. Silber joined Renaissance in 1983 and is responsible for the overall operations of its finance, administration and compliance departments. He was formerly a certified public accountant with the firm of Seidman & Seidman, now BDO USA. He holds a B.A. from Brooklyn College, a J.D. and an L.L.M. in tax law from New York University School of Law and an M.B.A. in finance from New York University Graduate School of Business Administration.

Chair, Simons Foundation
James H. Simons Ph.D.

James Simons is chairman of the Simons Foundation and board chair and founder of Renaissance Technologies. Prior to his financial career, Simons was chairman of the mathematics department at Stony Brook University, taught mathematics at the Massachusetts Institute of Technology (MIT) and Harvard University, and was a cryptanalyst at the Institute for Defense Analyses. Simons holds a B.S. from MIT and a Ph.D. from the University of California, Berkeley. In 1976, he won the Veblen Prize of the American Mathematics Society for his work in geometry. He is a trustee of the Stony Brook Foundation, Rockefeller University, MIT, Brookhaven National Laboratory, the Mathematical Sciences Research Institute, New York Genome Center and the Institute for Advanced Study, and is a member of the National Academy of Sciences, the American Academy of Arts and Sciences and the American Philosophical Society.

SIMONS FOUNDATION

160 Fifth Avenue New York, NY
10010 Tel. 646 654 0066

simonsfoundation.org
sfari.org