Translational tidbits

By Lev Osherovich and Kai-Jye Lou, Senior Writers

Autism Speaks to Google

Patient advocacy group **Autism Speaks** and **Google Inc.** have teamed up to sequence and make accessible the complete genomes of at least 10,000 patients with autism-spectrum disorder and controls. The resulting data set will be the world's largest collection of whole-genome sequence data and will be open to industry and academic researchers searching for clues about the origins of autism-spectrum disorder and possibly other diseases.

Under the Autism Speaks Ten Thousand Genomes Program (AUT10K), Autism Speaks will sequence samples from the Autism Genetics Resource Exchange (AGRE), a repository of biological samples and clinical records from patients with ASD and their unaffected immediate relatives.

About one-third of the sequences will be from patients, and the rest will be from healthy controls.

Google will organize and host the data on its Google Cloud platform and will provide customized bioinformatic analysis tools to academic and industry researchers interested in probing the data set. Financial terms of the deal are undisclosed.

"We're aiming to assemble the most complete set of genomic data in ASD patients, their siblings and parents," said Autism Speaks CSO Robert Ring. "The database will contain medical information and biomarker analyses as well as sequence data."

Ring said that AUT10K will give researchers an opportunity to test hypotheses about the genetic factors behind ASD, a diverse set of neurodevelopmental disorders caused by a multitude of rare mutations and chromosomal abnormalities.

Previous microarray studies and exome sequencing efforts by academics have yielded a long list of inherited and spontaneous mutations and copy number variations that account for up to 30% of ASD cases. According to Ring, unlocking the rare sequence variants behind the remaining 70% of cases requires a deeper dive into the genome.

"Understanding the full complexity behind autism has not been possible just with microarrays and exome sequencing, which miss a lot of the action in the wider genome," said Ring.

For example, he said, noncoding DNA and epigenetic control of gene expression could be potential players in ASD that may have been missed by prior genomic studies.

"Whether through inherited *de novo* mutations or through epigenetic interactions with the environment, genetics is the fundamental driver of disease," said Ring. "If you want to move this field forward, mapping the complete genomic variation in ASD is the best investment you can make. The immediate goal of the program is to get through 10,000 genomes, which is considered by statistical geneticists to be the right number to distinguish the many different forms of ASD," Ring added.

Ring said that in the last few years Autism Speaks ran a pilot project to sequence 100 whole genomes from the AGRE collection. The work was done through contracts with Chinese sequencing company **BGI** and its subsidiary Complete Genomics Inc.

Although obtaining genomic sequences proved relatively cheap and easy, making sense of the sequences proved technically challenging because of the vast size of the accumulated data set.

"We realized we needed partners to address the storage and computational problems with these data," said Ring. Thus, Autism Speaks partnered with Google Genomics, part of Google's cloud computing business unit, to manage the mountain of data.

David Glazer, engineering director at Google Genomics, said that the company's data-handling technology will allow researchers "to store, process, explore and collaborate on data analysis" without clogging up their computers with raw data.

Glazer said that his team has built tools for genomic data analysis in the past, but the scale of data from AUT10K—Google Genomics' first publically disclosed project—dwarfs previous sequencing efforts. "I don't know of a similarly sized genomic database," said Glazer.

He said that the computational challenges "are a lot more than trivial but are a lot less than groundbreaking" for Google's computers, which routinely handle considerably more data for its search engine services.

Ring said that AUT10K has already sequenced 1,000 genomes, with another 2,000 in the pipeline. The project aims to open a web portal to allow researchers to access the data set by early 2015. All data will be anonymized and accessible to qualified researchers.

Ring noted that because AUT10K will include full sequences from thousands of controls and patients with ASD, the data could be useful to researchers interested in diseases besides ASD.

"If you look at cancer, diabetes or Alzheimer's disease, you won't find a program that is trying to achieve this scale of whole-genome sequencing," said Ring. "Every field will be watching this program."

Getting BRAIN under way

Researchers at the **University of California, San Francisco** have launched one of the first projects in support of the U.S.'s Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative.

The five-year study involves determining how neural circuits become dysregulated in neuropsychiatric disorders and developing miniature implantable stimulation devices that strengthen alternative circuits to correct or bypass the dysregulation.

The multi-institutional study will include researchers at **Cornell University**, **Lawrence Livermore National Laboratory**, **New York University** and the **University of California, Berkeley**.

The project's initial focus will be on anxiety and depression in patients with Parkinson's disease (PD) or intractable epilepsy. The first step is studies to record activity from various brain regions.

Table 1. Selected public-private partnerships for May 2014. Last month, the **University of California, San Francisco** launched a project to identify brain signaling pathways associated with anxiety and depression in epilepsy and Parkinson's disease (PD) and design implantable devices that could correct the abnormal brain patterns. The project is one of the first to be launched in support of the U.S.'s Brain Research through Advancing Innovative Neurotechnologies initiative and is funded through the Systems-Based Neurotechnology for Emerging Therapies program from the **U.S. Department of Defense**'s **Defense Advanced Research Projects Agency**. On the pharma side, **Daiichi Sankyo Co. Ltd.** (Tokyo:4568) announced a pair of discovery deals focused on cancer and cardiovascular and metabolic diseases.
Source: BioCentury Archives

Companies	Institutions	Business area	Disclosed value	Purpose
Cubist Pharmaceuticals Inc. (NASDAQ:CBST)	**NIH**; **Rutgers University**; **The Rockefeller University**	Infectious disease	Up to $26 million	Develop antibiotics against drug-resistant bacteria
None	**Cornell University**; **Lawrence Livermore National Laboratory**; **New York University**; **University of California, Berkeley**; University of California, San Francisco; U.S. Department of Defense	Neurology	Up to $26 million	Understand and repair disrupted brain circuitry to treat mental illnesses
AstraZeneca plc (LSE:AZN; NYSE:AZN)	**MRC Laboratory of Molecular Biology**	Pharmaceuticals	Up to £9 million ($15.1 million)	Fund preclinical projects at the two organizations to better understand disease biology
Berg Pharma LLC	**Medical University of South Carolina**	Autoimmune disease	Undisclosed	Identify new therapeutic pathways and biomarkers to treat lupus
Crown Bioscience Inc.	**National Resource Center of Mutant Mice**	Supply/service	Unavailable	Develop mouse models for cancer research
Daiichi Sankyo	**MRC Technology**	Cancer; cardiovascular disease; endocrine/ metabolic disease	Undisclosed	Identify and select drug targets sourced from academia for cancer and cardiovascular and metabolic diseases
	Sanford-Burnham Medical Research Institute	Cardiovascular disease; endocrine/metabolic disease	Undisclosed	Discover treatments for cardiovascular-metabolic diseases
Debiopharm Group	**Yale University**	Autoimmune disease; inflammation	Undisclosed	Discover and develop oral inhibitors of macrophage migration inhibitory factor (MIF)
Enigma Diagnostics Ltd.	**Chinese Center for Disease Control and Prevention**	Infectious disease; diagnostics	Undisclosed	Develop point-of-care molecular diagnostic technology for infectious disease pathogens
None	**Cell Therapy Catapult**; **Great Ormond Street Hospital**; **Royal Free Hospital**; **University College London**	Transplantation	Undisclosed	Develop a tissue repair product for babies with esophageal atresia using stem cells from amniotic fluid and a donor scaffold
Nuvilex Inc. (OTCQB:NVLX)	**University of Northern Colorado**	Cancer	Undisclosed	Develop cannabis-based cancer treatments that use Nuvilex's Cell-in-a-Box, cellulose-based, live-cell encapsulation technology
RaQualia Pharma Inc. (JASDAQ:4579)	**Kyoto University**	Gene/cell therapy	Unavailable	Identify a small molecule that could induce differentiation and proliferation of induced pluripotent stem cells into immune cells
Sanofi (Euronext:SAN; NYSE:SNY)	**Foundation Fighting Blindness**	Ophthalmic disease	Unavailable	Share information and expertise that could help identify and advance development of potential therapies for inherited retinal diseases

The **Defense Advanced Research Projects Agency** of the **U.S. Department of Defense** will provide $12 million in initial funding for the project through its Systems-Based Neurotechnology for Emerging Therapies program, with up to an additional $14 million based on undisclosed milestones.

The BRAIN initiative was announced by the Obama administration in 2013 as a public-private partnership to give "scientists the tools they need to get a dynamic picture of the brain in action and better understand how we think and how we learn and how we remember."[1]

The **NIH** committed $40 million to the BRAIN initiative in FY14. The White House requested $110 million for the initiative in its FY15 budget request.

Earlier this month, an NIH working group proposal called for $4.5 billion in federal funding over 10 years starting in FY16 to achieve the initiative's goals, with $400 million annually from FY16 to FY20 and then $500 million annually from FY21 to FY25.

Public-private partnership roundup

Another quiet month on the public-private partnership front includes several new deals in Asia. **Daiichi Sankyo Co. Ltd.** announced a pair of discovery deals focused on cancer and cardiovascular and metabolic diseases; **RaQualia Pharma Inc.** announced a partnership with **Kyoto University** in the areas of gene and cell therapy; U.K.-based **Enigma Diagnostics Ltd.** announced a deal with the **Chinese Center for Disease Control and Prevention** for infectious disease and diagnostics (*see* **Table 1, "Selected public-private partnerships for May 2014"**).

Osherovich, L. & Lou, K.-J. SciBX 7(24); doi:10.1038/scibx.2014.692
Published online June 19, 2014

REFERENCES
1. Insel, T.R. *et al. Science* **340**, 687–688 (2013)

COMPANIES AND INSTITUTIONS MENTIONED

Autism Speaks, New York, N.Y.
BGI, Shenzen, China
Chinese Center for Disease Control and Prevention, Beijing, China
Cornell University, Ithaca, N.Y.
Daiichi Sankyo Co. Ltd. (Tokyo:4568), Tokyo, Japan
Defense Advanced Research Projects Agency, Washington, D.C.
Enigma Diagnostics Ltd., Salisbury, U.K.
Google Inc. (NASDAQ:GOOG), Mountain View, Calif.
Kyoto University, Kyoto, Japan
Lawrence Livermore National Laboratory, Livermore, Calif.
National Institutes of Health, Bethesda, Md.
New York University, New York, N.Y.
RaQualia Pharma Inc. (JASDAQ:4579), Nagoya, Japan
University of California, Berkeley, Calif.
University of California, San Francisco, Calif.
U.S. Department of Defense, Washington, D.C.