Tech News

Unraveling autism one de novo mutation at a time

Affecting 1 in 88 children in the US, autism spectrum disorders represent a challenge for children and their families, as well as researchers struggling to understand the genetic basis of autism. But as genome sequencing improves and costs plummet, this technology is beginning to offer opportunities to probe the genetics of autism at a deeper level. In fact, earlier this year, four groups published studies (1-4) using exome sequencing to hunt for rare de novo mutations in children with autism. Identifying these rare mutations is important as they may occur in as many as 20% of individuals with autism. Though new treatments and diagnostic tools are still a long way off, "[these studies] show a clear path forward for gene discovery and autism, a systematic way to identify specific genes that are contributing to risks," notes Matthew State of the Yale Child Study Center, a principal investigator on one of the studies published in *Nature* (1). Even as this article went to press, another paper in *Nature* (5) connected the number of de novo mutations in children with the father's age at conception, a finding that could have implications for the risk for autism in children.

The modern genetics of autism

Researchers have long recognized a link between genes and autism. Individuals with certain inherited single gene mutations, such as those that cause Fragile X syndrome and tuberous sclerosis, have highly increased incidences of autism as well. But the majority of children with autism do not have autistic parents, and the combination of rare genes and the heterogeneity of the disorder parallels other neurodevelopmental and neuropsychiatric disorders, says State.

One approach to understanding autism genetics has been through the use of genome-wide association studies (GWAS) that search for changes in common genes among larger numbers of individuals that might confer risk for autism. The trouble is that searching through GWAS data is like looking for a needle in a haystack. If you imagine genetic variation as a pyramid,

Ben Neale of Massachusetts General Hospital and his colleagues performed exome sequencing on samples from 175 individuals with autism and their parents outside the SSC. Credit: Maria Nemchuk, Broad Communications

GWAS mostly looks at the wide bottom, sifting through the largest numbers of genes, says Ben Neale of Massachusetts General Hospital and the Broad Institute and first author of the second study in *Nature* (2). The end result is that GWAS can often miss de novo mutations.

De novo mutations are not present in the genomes of parents, but rather arise within egg or sperm cells, leading to genetic changes in their children. The vast majority of these rare mutations have no pathological effect. Limiting searches to this much rarer set of mutations provides a kind of filter that reduces hit numbers in a screen, a way of looking at just the tip-top of that genetic variation pyramid. In 2007, a team led by Cold Spring Harbor Laboratory scientists Jonathan Sebat and Mike Wigler showed for the first time that de novo copy number variations could have implications in autism spectrum disorders (6). Ten percent of autism subjects had de novo mutations compared with only 1% of controls. "That's where the interest in de novo mutations really started," says Dan Geschwind of UCLA, and a co-author on State's *Nature* paper.

If researchers locate a de novo mutation in a functional gene of a child with autism that doesn't occur in the parents or in an unaffected sibling, there exists the

Michael Ronemus of Cold Spring Harbor Laboratory and his team used mouse studies to identify a number of mutated genes connected to Fragile X mental retardation protein. Courtesy of Michael Ronemus

possibility that this gene could confer a greater risk of autism. But if that same mutation is identified in two unrelated individuals, both with autism, that confidence level increases nearly to 99.9%, says Michael Ronemus of Cold Spring Harbor Laboratory and a co-author on the paper published in *Neuron* (4).

Initially, researchers used microarrays in their de novo mutation studies to examine large-scale mutations on the order tens of thousands of base pairs. But even though identifying these copy number variations represented progress, such large segments often include multiple genes, says State. To really understand the neurobiology, researchers would like to examine the single gene level to investigate the connection between a specific mutation and the resulting pathology of a complex neuropsychiatric disorder. Today microarrays have a partner in the hunt for de novo mutations as advances in high-throughput sequencing and the decreasing cost associated with this technology now allow researchers to take a higher-resolution look at these genetic variants, providing the ability to identify smaller scale mutations that change protein function. In addition, it turns out that large-effect mutations are also incredibly rare, which means that researchers can use targeted sequencing to sample particular

genomic regions in tens of thousands of cases and controls to identify differences.

“It’s the right disorder and the right technology kind of coming together,” says State. Not to mention the fact that researchers have increasing access to samples from children with autism and their family members, which is also helping to speed research (See sidebar: The importance of biological samples).

These exome sequencing studies are a natural extension of the earlier work on copy number variations according to State, with researchers now looking across the coding genome with single-base resolution for loss-of-function mutations. “We would not have had the capacity to do that previously.”

A path forward with a few hints

The critical finding emerging from these autism exome studies has less to do with specific genes and more to do with validating this particular approach as a tool for understanding the genetic underpinnings of these disorders. “What our data and all of the papers showed was that if you move to a higher level of resolution, essentially you’re seeing about the same picture, which is that there’s an overrepresentation of these highly deleterious mutations,” acknowledges State.

Across the four studies, researchers observe *de novo* mutations in the genomes of children with autism and, where tested, their unaffected siblings. However, children with autism were more likely to have mutations that disrupted protein function—mutations that halted gene transcription early or splice site mutations.

Three of the four studies looked at samples from the Simons Simplex Collection, a collection of genetic samples from 2700 families with a single child with an autism spectrum disorder and unaffected parents and an unaffected sibling. Even when grouping all of the Simons Simplex Collection (SSC) families together whose genes were sequenced in these three studies, they total less than 800 out of 2700 families (Neale and his colleagues performed exome sequencing on samples from 175 individuals with autism and their parents outside the SSC). With only a handful of genes, did researchers find disruptive *de novo* mutations in the same gene in different individuals with autism? “We almost never saw lightning strike twice,” says Evan Eichler of the University of Washington, the principal investigator on the third *Nature* paper. If a proportional number of recurrent genes pop up in the

Evan Eichler of the University of Washington and his colleagues found recurrent mutations in *CHD8*, a gene encoding a chromodomain helicase that can modify transcription of a whole host of other genes. Credit: Ron Wurzer/AP, HHMI

rest of the SSC, exome sequencing of this collection might point to a total of 25 to 30 new genetic mutations that confer autism risk, says Geschwind.

It might be too early to make predictions about specific genes and what they mean for the biology autism. “I think most people who are cautious in this field were thrilled by those papers, but they really want to see what the whole [SSC] is going to look like,” says John Spiro, deputy scientific director for the Simons Foundation Autism Research Initiative in New York, NY.

Still, some of the findings offer hints that may develop into wider patterns. In their *Nature* study, State and his colleagues pulled out one gene where disruptive *de novo* mutations occurred in two autism probands (Neale and his colleagues also

pulled out this gene while their *Nature* paper was under review). The gene, *SCN2A*, encodes a sodium channel, that is highly expressed in the brain. In addition, *SCN2A* mutations have been connected with epilepsy, and a variety of genetic data has pointed to overlaps between genetic risk for epilepsy and the risk for autism spectrum disorders.

Two groups, one led by Eichler at the University of Washington and the other by Neale, found recurrent mutations in *CHD8*, a gene encoding a chromodomain helicase which could modify transcription of a whole host of other genes. Eichler’s group also found a recurrent disruptive mutation in netrin G1 (*NTNG1*), a gene important in axonal guidance, while Neale and his colleagues found recurrent mutations in *KATNAL2*, a gene whose function is not well characterized.

Researchers also see potential connections between the different observed *de novo* mutations. Forty percent of the mutations Eichler’s team observed were in genes making up a highly interconnected pathway surrounding *CHD8*, and Ronemus and his colleagues used mouse studies to identify a number of mutated genes connected to Fragile X mental retardation protein (FMRP), an important regulatory of Fragile X syndrome suggesting the pathway involved in Fragile X syndrome could be central in a wider number of autism cases.

Remaining questions

Whole exome sequencing data should be available for the entire SSC within two years, according to Ronemus. But those mutations only represent a fraction of genetic changes

The importance of biological samples

In addition to sequencing technology, these types of studies require access to samples from individuals with autism and their unaffected family members. Three of these studies used samples from family groups in the Simon Simplex Collection (SSC), a permanent repository of 2700 families affected by autism. To simplify analysis, this collection was limited to families where the parents do not have autism and only one child has an autism spectrum disorder. They could not have other neurodevelopmental disorders, and, where possible, they include genetic material from an unaffected sibling.

This particular collection, operated by the Simons Foundation Autism Research Initiative (SFARI) and collected at 12 academic medical research centers, is unique because of its size, its focus on simplex families, and its inclusion of unaffected siblings, explains John Spiro, deputy research director at SFARI.

“The thing that cannot be underestimated is putting together a really good set of families. That’s tough, and it requires a ton of money,” says Michael Ronemus of Cold Spring Harbor Laboratory. The sequencing and the molecular analysis are not the most expensive part of this research. “The expense is in finding and diagnosing people, and collecting biological samples from them in the first place.” -S.W.

Think you need an RGB imager
and a Near IR scanner?

Why buy two when one will do?

Get it all with the
Gel Logic 6000 PRO,
the ultimate **"all in one"**
blot & gel imaging system

- Fluorescence with no restrictions-UV to RGB to NIR
- Rapid chemiluminescence imaging
- Superior resolution and sensitivity
- Multiplex and overlay up to 4 labels
- Fast, easy, and intuitive software

Carestream
Molecular Imaging

Toll-free 1-877-747-4357
+1 203-786-5600

micarestream.com/GL6000

See us at WMIC booth 624
and SFN booth 2841

that could boost autism risk. Researchers suspect that hundreds of genes might be contributing to autism spectrum disorders and uncovering and confirming all those genes could require sequencing the genomes of tens of thousands of families.

In addition, simply finding the gene isn't enough—researchers need to understand the function of the gene, the protein it encodes for, and what other proteins it might interact with in biological pathways. Even with just 20 to 30 new genes, that's a tremendous amount of functional, molecular, and animal model research. The major question for Geschwind is whether or not there are convergent pathways. And then there are the clinical questions: As with cancer, will researchers be able to define molecular subtypes of autism that are treated with similar drugs or behavioral methods? Or will there be hundreds of subtypes? Answering those questions will be critical for identifying potential targets that pharmaceutical companies could use to develop drug treatments.

At the moment, the genes that have been discovered and linked to autism may be difficult to target with therapies. "They're tough genes, at least the ones that we're finding because they're so highly interconnected at the protein level and early in development," says Eichler.

Mutations found in the exome will be easier to interpret than those found in other parts of the whole genome. "But there are plenty of people who are now throwing these samples onto whole genome sequencing platforms, and hoping that they'll find mutations in conserved extra-genomic regions," says Spiro. After that, researchers might start to look for epigenetic changes in the same samples, he adds. But moving into whole genome sequencing may present a data interpretation bottleneck, Ronemus says. Even with the exome, it's not that simple. "I'm sure it will be done eventually, and I think it should be done, but I think the interpretive capacity, the computational capacity, if you analyze whole genome data it's really not anywhere near as advanced," says Ronemus.

Eventually, as researchers learn more about the connections between the identified genes and the biology of autism, they may be able to use sequencing as a part of diagnosis, as a way to predict prognosis, and make treatment choices. Someday, families might be able to use sequencing data to determine whether inherited genes or de novo mutations contributed to their child's autism, which could help them understand the risk of autism for other children within their family, says Geschwind. If subtypes are identified, a genetic profile might

eventually steer families toward effective therapies and away from ineffective ones, says Ronemus.

"The landscape is sort of coming into focus," says Spiro, noting that the genetic work will be critical for the understanding of autism. But at this point, he adds, researchers are far from saying: "Aha, you know, now we understand it, and we're ready to just hand it off to a pharmaceutical company." I think everyone wishes that were the case, but we've got plenty more time to go."

References:

1. Sanders, S.J., M.T. Murtha, A.R. Gupta, J.D. Murdoch, M.J. Raubeson, A.J. Willsey, et al. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485:237-241.
2. Neale, B.M., Y. Kou, L. Liu, A. Ma'ayan, K.E. Samocha, A. Sabo, et al. 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485:242-245.
3. O'Roak, B.J., L. Vives, S. Girirajan, E. Karakoc, N. Krumm, B.P. Coe, et al. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485:246-250.
4. Iossifov, I., M. Ronemus, D. Levy, Z. Wang, I. Hakker, J. Rosenbaum, et al. 2012. De novo gene disruptions in children on the autism spectrum. *Neuron* 74:285-299.
5. Kong, A., M.L. Frigge, G. Masson, S. Besenbacher, P. Sulem, G. Magnusson, et al. 2012. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 488:471-475.
6. Sebat, J., B. Lakshmi, D. Malhotra, J. Troge, C. Lese-Martin, T. Walsh, et al. 2007. Strong association of de novo copy number mutations with autism. *Science* 316:445-449.

Written by Sarah A. Webb, Ph.D.

BioTechniques 53:133-136 (September 2012)
doi: 10.2144/000113920