VIEWPOINT

Tangled web of proteins holds clues to autism's complexity

BY **LILIA IAKOUCHEVA**

28 FEBRUARY 2017

Mutations in hundreds of genes have been linked to autism, thanks to dozens of studies over the past decade. But many families, funders and researchers are wondering why this progress has **yet to yield treatments** for the condition.

The answer is simple: We still don't know precisely how these genetic changes lead to autism.

Before we can design effective autism treatments, we need to better understand the impact of autism-linked mutations on neurons and the brain circuits they form. First, we need to figure out where and when the mutated genes are expressed in the brain. Then we need to investigate how mutations affect proteins and the interactions among them.

Researchers have laid the groundwork for this project by mapping gene expression levels throughout the brain and cataloging the interactions between proteins. The **BrainSpan Atlas**, for instance, charts the expression of messenger RNAs (mRNAs), the intermediaries between genes and proteins, in 16 brain regions across the lifespan[1]. And the **Human Interactome** details some 14,000 protein-protein interactions[2].

Researchers have used these databases to explore whether certain autism-linked genes are switched on in the same brain regions or at the same time. In one study, this approach revealed a vulnerability of neurons in certain regions of the brain **during mid-fetal development**[3].

Among the top candidates for autism is a stretch on chromosome 16 called **16p11.2**, which is missing or duplicated in some people with autism. My colleagues and I mapped when and where in the brain genes on this segment work together[4]. Our analysis highlighted a molecular pathway that regulates neuronal migration during fetal development and that **may go awry in autism**.

Alternative reality:

The path from gene to protein is complex. A single gene can give rise to many different versions of a protein, called 'isoforms,' that differ only slightly from one another. The roughly 20,000 protein-coding genes in the human genome may generate as many as 100,000 distinct mRNAs and more than a million different proteins.

Proteins that are subtly different can have different properties and binding partners. But existing databases only detail the interactions of a single 'reference' isoform for each protein. Similarly, gene expression atlases tend to include an average level of gene expression based on multiple mRNAs. By focusing on only one version of a gene or protein, researchers may miss some important connections.

In a study published in 2016, we investigated how 1,423 isoforms corresponding to 506 genes interact with a set of 15,000 other proteins[5]. We were surprised to find that most isoform pairs encoded by the same gene share less than half of their binding partners. About 16 percent of isoform pairs share no protein partners at all. This finding suggests that different isoforms can behave like completely different proteins.

We know that different isoforms are expressed in different tissues. But most protein interaction databases do not capture this tissue specificity. They may even describe interactions among protein isoforms that are not produced in the same tissue.

To address this shortcoming, we **mapped the interactions** among 422 isoforms from 168 genes implicated in autism[6]. The resulting network — a sort of autism interactome — doubled the number of previously known interactions for this set of genes. More than 90 percent of these interactions were previously unknown. The new network connected genes from 27 chromosomal regions that are deleted or duplicated in people with autism.

New connections:

We have surveyed only a small proportion of all the isoforms that are expressed in the brain. Based on our estimates from BrainSpan, the developing human brain could express as many as 70,000 unique isoforms.

Assessing the interactions for this many isoforms would be challenging and expensive. But we can begin to build networks relevant to autism by combining sequencing technologies that can detect isoforms using **computational methods** that predict interactions among them.

Identifying the isoforms that autism-linked mutations affect will enhance our understanding of the condition. If a mutation falls in a segment of mRNA that is skipped in one of the isoforms, it may have no effect. Similarly, if an isoform with a mutation is not expressed in a brain region or

developmental period that is important in autism, the mutation may be silent.

Alternatively, if an autism mutation maps to an isoform that is highly expressed in a key brain region or during a critical developmental period, it could have profound effects.

Knowing how mutations affect isoforms and isoform networks is highly relevant to treatments. We are just setting out on the path to constructing these networks. We have a long way to go, but it's a trip worth taking.

Lilia Iakoucheva is associate professor of psychiatry at the University of California, San Diego.

REFERENCES:

1. Kang H.J. *et al. Nature* **478**, 483-489 (2011) **PubMed**
2. Rolland T.M. *et al. Cell* **159**, 1212-1226 (2014) **PubMed**
3. Willsey A.J. *et al. Cell* **155**, 997-1007 (2013) **PubMed**
4. Lin G.N. *et al. Neuron* **85**, 742-754 (2015) **PubMed**
5. Yang X. *et al. Cell* **164**, 805-817 (2016) **PubMed**
6. Corominas R. *et al. Nat. Commun.* **5**, 3650 (2014) **PubMed**