A Conversation with Tony Zador

INTERVIEWER: GARY STIX

Editor, Scientific American

Tony Zador is a Professor of Biology and Program Chair in Neuroscience at Cold Spring Harbor Laboratory.

Gary Stix: You are giving a talk at this conference entitled "Sequencing the Connectome." What is a connectome and why is it important?

Dr. Zador: The brain has a lot of neurons. The human brain has 100 billion, the mouse brain maybe 100 million. What we'd really like to understand is how we go from a bunch of neurons to thought, feelings, behavior, all those things. We think that the key to understanding is finding out how the neurons are connected to one another. There have been a lot of traditional techniques for studying connectivity but at a fairly gross level. We can tell that a bunch of neurons here tend to be connected to a bunch of neurons there. There are also techniques for looking at how single neurons are connected but only one at a time. What we would love to be able to do is to tell how every single neuron in the brain is connected to every single other neuron. If you wanted to navigate through the United States, one of the most useful things you could have is a road map. It wouldn't tell you everything about the United States but, without it, it would be very hard to get around.

Gary Stix: You're proposing that sequencing the connectome might be a novel way to understand the connectome.

Dr. Zador: Traditionally, connectivity has been studied as a branch of microscopy. People talk about mapping the brain. Typically, people use one method or another to label a neuron, then observe that neuron at some level of resolution.

The challenge at the core of all microscopy techniques is that neurons can extend long distances—millimeters in a mouse brain, or in a giraffe brain, all the way from the brain to its foot, which must be 10 feet . . . so very long distances.

They're connected with one another at structures called synapses which are below the resolution of light microscopy. So if you really want to understand how one neuron is connected to another, you have to be able to resolve the synapse, which requires electron microscopy. And that means taking incredibly thin sections of brain and imaging them using electron microscopy.

People are doing this impressively well. A big success was in *C. elegans,* an organism with 302 neurons and 7000 synapses. With over 50 person years of work, they were able to reconstruct the entire wiring diagram of this tiny creature. Since then, they've scaled it up. It's working pretty well but it's still extremely challenging.

A few years ago, largely because I'm at Cold Spring Harbor and steeped in sequencing technology, it occurred to me that sequencing has the right scale to handle billions of synapses. If you have a brain that has 100 million neurons and each neuron makes, let's say, 1000 synapses, that's 100 billion synapses.

A sequencing run these days costs about $1000 for one billion reads. The way we propose to do it, one read equals one synapse. That price is coming down. Fifteen years ago, the cost of sequencing the first human genome was approximately $1 billion. Now you can get your genome sequenced for $1000 and within a couple years, it will be well under that.

Sequencing is getting better faster at a rate that's even faster than the rate at which computers get better. My iPhone has more computing power than a computer 20 years ago. That follows something called Moore's Law, and sequencing technology has been improving at a rate substantially faster than Moore's Law, since 2008.

There's every expectation that sequencing will continue to get faster. There's a huge potential benefit to converting the problem of connectivity into a problem of sequencing. Basically, the benefit is cost and speed. Sequencing takes a week to sequence a billion reads.

Gary Stix: A circuit is not a gene. How would you sequence a circuit?

Dr. Zador: The idea is that, first of all, we endow every neuron with a unique, random sequence of DNA. We call it a barcode.

This sounds fanciful but the immune system has solved this problem. B cells and T cells generate novel receptors

 doi: 10.1101/sqb.2014.79.20

through somatic recombination. They scramble pieces of their chromosome to produce, for example, novel antibodies.

We're not using the particular collection of enzymes used in the immune system because they're not really convenient. But we're hijacking similar proteins from other organisms to try to do the same things in neurons.

We're going to put in every neuron a cassette which, when we express the protein, will scramble the pieces and generate a novel sequence in every neuron in the brain. "Every" neuron sounds like a lot, but combinatorics works in our favor.

A sequence of 30 random nucleotides has a potential diversity of 4 to the power of 30 because there are four nucleotides. This is way more than the number of neurons in the brain. If we just put this random cassette in and can cause enough scrambling to occur, then by chance, the probability that two neurons will have the same bar code, is infinitesimally small. If we can do that, the next step is to express little pieces of RNA that encode that little random barcode.

We've already engineered proteins that will drag those RNAs to the synapse. In each synapse, there will be a presynaptic barcode and postsynaptic barcode. With biochemistry, we can join and link them together so you have a single piece of DNA that has the presynaptic barcode and the postsynaptic barcode.

Then all we have to do is read that out and we get, in principle, a huge connectivity matrix that has neurons numbered 1 through 100 million here, 1 through 100 million here, and a dot where neuron 22 is connected to neuron 53. Then another dot where neuron 27 is connected to neuron 68.

Gary Stix: There are lots of different types of neurons and lots of different structures in the brain.

How far do you think you could go with this? Could you look at all of the brain?

Dr. Zador: There's no reason we couldn't do an entire mouse brain. In fact, there's no reason we couldn't do many mouse brains once we get the transgenic mice working. The sequencing costs will be not negligible, but will be well within the range that would make this project worthwhile. We envision getting the connectivity of many individuals, not just a portion of the circuit but the entire circuit.

I've given you the bare-bones version, but there are bells and whistles that we can add that we think will allow us not only to know the connectivity but also the precise location of those cells and their gene expression pattern which, in turn, will tell us about their cell type.

Although it's ambitious, but not beyond what we think is reasonable, we'd like to have the complete connectivity matrix, then associated with each one of those elements, what cell type it is, what genes it expresses. and therefore what cell type it is.

Gary Stix: As with gene sequencing, in principle, this could be done very quickly.

Dr. Zador: Basically, in this form, once we have a transgenic animal, extracting the DNA takes a few days, sequencing takes 2 weeks or less. Per individual, it could take a month. Making the animals we can do that with has taken us a few years and will probably take us a few more years. Once we have them, it won't take much time.

Gary Stix: There's a lot of emphasis in the U.S. and elsewhere on developing new technologies for understanding the brain better. Do you see sequencing as one of them?

Dr. Zador: I've applied for grants. I should be hearing in a few months. What I hope to have soon is a proof of principle that I hope will convince people that this is something worth pursuing. To scale it up to the level I'm talking about is beyond what one lab could do.

After we've shown the proof of principle, I'd love lots of labs to get involved and come up with better ideas for doing a lot of the things we're doing, in the quickest possible way.

I would be thrilled if resources were devoted to this. I think having the complete connectivity of organisms, would be incredibly useful and transform how we do neuroscience.

Gary Stix: Are there competing technologies?

Dr. Zador: There are two other main approaches to figuring out connectivity.

Gary Stix: Those are traditional approaches, aren't they?

Dr. Zador: The people who practice them would say they've extended them to the point that they're not traditional anymore. The main competition, although I see it as complementary, comes from various forms of electron microscopy. That technology has improved so much that for certain questions, it's going to be super useful.

Gary Stix: Could these approaches be combined?

Dr. Zador: It's hard for me to envision how a combination could scale up to the speed and cost required.

Gary Stix: In your recent paper in *PLoS Biology*, you mentioned that this might be a good technique to test out hypotheses such as how the brain circuits go awry in a disorder like autism.

Dr. Zador: Autism is one of my research interests. There's been a lot of progress recently in identifying the genes involved in causing autism. There are dozens and perhaps hundreds of genes which when perturbed, can cause autism. Yet, autism, although it's heterogeneous, is still a meaningful diagnosis. There must be something in common amid the heterogeneity.

There's an appealing hypothesis what is shared is disruption of circuitry. It's possible that the genetic lesions we know cause autism in people could be recapitulated in mice. Then we can ask the question, what goes awry in the circuits of mice which express the same genetic lesions that people have?

The hope is that we could take 20 mouse models of autism, look at their brains, look at their connectivity, and say, "Hey, we noticed that in 17 out of these 20, there's a

disruption in this circuit compared with how it is in non-autistic mice." That would lead us to look in the direction of that circuit.

That circuit could be long-range connectivity, let's say from the front of the brain to the back of the brain. Or far more specific things like the inability of one subset of neurons in this part of the brain to connect a particular subset of inhibitory interneurons in some distant region. Or even some local region. There are an enormous number of possible hypotheses. We can test and validate any one of them, but testing them all, at this point, is not possible unless you have a method like the one I describe.