NEUROSCIENCE

NBLAST: a similarity search for neurons

By analogy to protein and DNA similarity searches, NBLAST provides a fast and efficient way of finding morphological similarities between neurons.

High-throughput imaging approaches generate a wealth of information on the many different neurons that make up the brains of model organisms, but few tools can mine this diversity. Greg Jefferis from the Laboratory of Molecular Biology in Cambridge experienced this challenge firsthand as his work on neural circuits in the fly moved “from having images of hundreds of neurons to thousands, tens of thousands and eventually millions of neurons,” he says. To cope with these large data sets, Jefferis and his team developed NBLAST, a tool that compares neuron structures and enables researchers to find similar neurons in an image database.

Jefferis had toyed with the idea of comparing neurons for a long time, but the lack of a suitable test data set prevented him from developing a general approach. This changed with the publication of a FlyCircuit data set of more than 16,000 individually imaged *Drosophila* neurons by the group of Ann-Shyn Chiang in Taiwan. It was crucial that Chiang made the data freely available. “Fly labs have been pretty good about sharing the image data from these larger-scale studies,” he says.

Jefferis and his colleagues, including first author Marta Costa, started by registering each image in the data set to a common template. They then represented each neuron with a set of vectors corresponding to small neuronal segments, producing a ‘vector cloud representation’. The representation is useful because it is flexible and easy to generate—“it is something that you can easily do in an automatic fashion from an image volume without manual tracing,” says Jefferis.

Using the vector cloud representations, the researchers were able to compute similarity scores based on neuron shape and position. “This had analogy with, for example, the BLAST for protein similarity, where there is a scoring matrix...based on looking at a lot of aligned protein sequences and looking at the substitution rates of

different residues,” explains Jefferis. The researchers took a similar idea and applied it to neurons.

NBLAST is not the first tool that allows neurons to be compared, but existing tools are computationally demanding and are best suited for comparing pairs of neurons. Pre-registration to a standard template and the use of vector clouds make neuronal comparison a computationally efficient process. “It’s only one or two milliseconds a comparison, typically,” says Jefferis. This makes it possible to compare all the neurons in the FlyCircuit data set with one another, even though this requires more than 250 million comparisons. “With our tool, that takes a day, and if we had used one of these other tools, we would have been waiting for centuries, even with quite a lot of computational power,” says Jefferis.

Not only can NBLAST be used to organize large image data sets of neurons and to discover new subtypes of cells, but the researchers have also extended it to allow comparisons with expression patterns in the popular *Drosophila* GAL4 lines. They have generated vector cloud representations for a collection of 3,500 GAL4 lines produced by the Rubin lab at Janelia Research Campus, which can be queried with single-neuron representations. This can help researchers find a GAL4 line that will allow them to eventually manipulate a neuron of interest. Jefferis is also extending NBLAST to electron microscopy data. With colleagues at Janelia, he has already shown that it is possible to query the neuron database with partially traced neurons from electron microscopy reconstructions and to assign cell types on the basis of similarities to neurons in the light microscopy data sets.

One limitation of the NBLAST approach is that it requires spatially registered neurons. But Jefferis hopes that this work might enable researchers to use such data “and maybe dream up new ways of measuring neuron similarity, possibly ones that don’t require image registration,” he says.

Nina Vogt

RESEARCH PAPERS

Costa, M. *et al.* NBLAST: rapid, sensitive comparison of neuronal structure and construction of neuron family databases. *Neuron* **91**, 293–311 (2016).