

## Video Article

# Large-scale Reconstructions and Independent, Unbiased Clustering Based on Morphological Metrics to Classify Neurons in Selective Populations

Elise M. Bragg<sup>1</sup>, Farran Briggs<sup>1</sup><sup>1</sup>Physiology & Neurobiology, Geisel School of Medicine at DartmouthCorrespondence to: Farran Briggs at [Farran.Briggs@dartmouth.edu](mailto:Farran.Briggs@dartmouth.edu)URL: <https://www.jove.com/video/55133>DOI: [doi:10.3791/55133](https://doi.org/10.3791/55133)

Keywords: Neuroscience, Issue 120, neuroanatomy, retrograde tracer, modified rabies virus, neuronal reconstruction, independent clustering algorithms, cluster evaluation

Date Published: 2/15/2017

Citation: Bragg, E.M., Briggs, F. Large-scale Reconstructions and Independent, Unbiased Clustering Based on Morphological Metrics to Classify Neurons in Selective Populations. *J. Vis. Exp.* (120), e55133, doi:10.3791/55133 (2017).

## Abstract

This protocol outlines large-scale reconstructions of neurons combined with the use of independent and unbiased clustering analyses to create a comprehensive survey of the morphological characteristics observed among a selective neuronal population. Combination of these techniques constitutes a novel approach for the collection and analysis of neuroanatomical data. Together, these techniques enable large-scale, and therefore more comprehensive, sampling of selective neuronal populations and establish unbiased quantitative methods for describing morphologically unique neuronal classes within a population.

The protocol outlines the use of modified rabies virus to selectively label neurons. G-deleted rabies virus acts like a retrograde tracer following stereotaxic injection into a target brain structure of interest and serves as a vehicle for the delivery and expression of EGFP in neurons. Large numbers of neurons are infected using this technique and express GFP throughout their dendrites, producing "Golgi-like" complete fills of individual neurons. Accordingly, the virus-mediated retrograde tracing method improves upon traditional dye-based retrograde tracing techniques by producing complete intracellular fills.

Individual well-isolated neurons spanning all regions of the brain area under study are selected for reconstruction in order to obtain a representative sample of neurons. The protocol outlines procedures to reconstruct cell bodies and complete dendritic arborization patterns of labeled neurons spanning multiple tissue sections. Morphological data, including positions of each neuron within the brain structure, are extracted for further analysis. Standard programming functions were utilized to perform independent cluster analyses and cluster evaluations based on morphological metrics. To verify the utility of these analyses, statistical evaluation of a cluster analysis performed on 160 neurons reconstructed in the thalamic reticular nucleus of the thalamus (TRN) of the macaque monkey was made. Both the original cluster analysis and the statistical evaluations performed here indicate that TRN neurons are separated into three subpopulations, each with unique morphological characteristics.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/55133/>

## Introduction

Neuroanatomy is one of the foundations of neuroscience<sup>1</sup> and recent interest in "connectomics" has renewed enthusiasm for understanding the morphological diversity of neuronal populations and the connections between specific neurons<sup>2</sup>. Methods for labeling and reconstructing neurons have greatly improved with recent innovations, including genetic and virus-mediated circuit tracing approaches<sup>3,4</sup>, enabling more comprehensive morphological surveys of neuronal populations<sup>5</sup>. In addition to improvements in labeling individual neurons, quantitative data analysis techniques have also emerged that enable independent and unbiased classification of neurons into distinct subpopulations based on morphological data<sup>5,6</sup>. These unbiased techniques are an improvement upon more traditional qualitative classification methods that have been the standard in the field for over a century. The goal of this study is to outline, step-by-step, the combination of virus-mediated labeling of neurons within a selective population, large-scale reconstructions of a comprehensive sample of these neurons, and quantitative data analysis based on independent clustering with statistical evaluation. By combining these methods, we outline a novel approach toward the collection and analysis of neuroanatomical data to facilitate comprehensive sampling and unbiased classification of morphologically unique neuronal types within a selective neuronal population.

As an example of these methods, we describe our analysis of a large population of neurons within a single sector of the thalamic reticular nucleus (TRN) of the macaque monkey. These data are from a prior study<sup>7</sup>. Methods for selectively labeling TRN neurons projecting to the dorsal lateral geniculate nucleus of the thalamus (dLGN) using surgical injection of modified rabies virus encoding EGFP<sup>4,8</sup> (see **Table of Specific Materials/Equipment**, row 2) are outlined. This modified rabies virus lacks the gene encoding an essential coat protein, eliminating trans-synaptic movement of the virus. Once the virus enters axon terminals at the injection site, it acts like a traditional retrograde tracer with the

important benefit of driving EGFP expression throughout the full dendritic arborization of infected neurons<sup>5,9,10</sup>. Accordingly, this G-deleted rabies virus can be utilized to selectively infect and label any neuronal population following injection and retrograde transport.

In order to perform a comprehensive analysis of a specific neuronal population, it is important to sample from a broad distribution of neurons within the population. Because the virus-mediated labeling technique produces complete intracellular, "Golgi-like" fills of many neurons with axons at the virus injection site, it is possible to reconstruct a very large sample of neurons within the full extent of a brain structure. Additionally, because the modified rabies virus is so effective at infecting and labeling large numbers of neurons, it is possible to reconstruct hundreds of neurons per animal. Procedures for sampling 160 neurons throughout the visual sector of the TRN<sup>11</sup> in order to generate a comprehensive sample of dLGN-projecting TRN neurons are outlined. The process of reconstructing individual neurons using a neuron reconstruction system including a microscope, camera, and reconstruction software is described. Also described are methods to determine positions of individual neurons within a brain structure (in this case within the TRN) and to verify virus injection site volume and location within a structure (in this case within the dLGN) using volumetric contour reconstructions. Steps to export morphological data and perform independent cluster analyses based on morphological metrics measured for each neuron are outlined. There are limitations to clustering methods and there are also a variety of different clustering algorithms available. Accordingly, these options and the benefits of some of the more commonly used algorithms are described. The cluster analysis does not provide statistical verification of the uniqueness of clusters. Therefore, additional steps are outlined to verify optimal clustering as well as the relationships between morphological data within and across clusters. Statistical methods for evaluating clusters for the TRN dataset to confirm that TRN neurons are grouped into three unique clusters based on 10 independent morphological metrics are described.

Thus, by outlining steps for selectively labeling, reconstructing, and analyzing morphological data from a specific neuronal population, we describe methods for quantifying morphological differences among neurons within a population. Prior findings of distinct neuronal types within the visual sector of the macaque monkey TRN are confirmed with separate statistical evaluation methods. Together, we hope these techniques will be broadly applicable to neuroanatomical datasets and help establish quantitative classification of the diversity of neuronal populations through the brain.

## Protocol

Note: The tissue examined in this study was prepared as a part of a separate study<sup>5</sup>. Therefore, all of the experimental methods involving the use of animals have been described in detail in the Experimental Methods section of Briggs *et al.* (2016). All procedures involving animals conducted as a part of the prior study were approved by the Institutional Animal Care and Use Committees. The steps for injection of virus into the dLGN and histological processing of brain tissue are described briefly below in sections 1 - 2.

### 1. Stereotaxic Injection

1. Perform all surgical procedures in a sterile environment using aseptic techniques. Steam-sterilize all metal surgical instruments, gauze and drape material in autoclave. Sterilize all materials that could be damaged by steam using chemical sterilization (e.g. ethylene oxide).
2. Induce and maintain anesthesia according to animal- and protocol-specific requirements.
  1. For virus injection in monkeys, induce anesthesia with ketamine (10 mg/kg) and maintain animals under full surgical anesthesia with isoflurane (1 - 3% inhaled in oxygen) monitoring expired CO<sub>2</sub>, EKG, respiration rate, and temperature continuously and periodically checking for jaw tone to ensure proper anesthetic depth.
3. Place animal in a stereotaxic frame to stabilize head position and to allow use of stereotaxic coordinates to locate the injection structure of interest (e.g. dLGN). If measuring visual responses, place eye drops (1% atropine eye drops if pupil dilation is desired, otherwise saline eye drops) and then contact lenses in both eyes to prevent dryness. If not measuring visual responses, place ophthalmic ointment in eyes.
4. Make a midline scalp incision and retract the skin and muscles.
5. According to stereotaxic coordinates for the structure of interest (dLGN) in one hemisphere, make a small craniotomy.
6. Gradually lower a recording electrode (e.g. tungsten or platinum/iridium electrode (see **Table of Specific Materials/Equipment**, row 3) clamped to a micromanipulator that is manually positioned) through the craniotomy into the brain. Lower impedance (< 1 MΩ) and slightly thicker (~ 250 μm diameter) electrodes may be utilized to penetrate the dura, if desired.
7. Record the manipulator position at the cortical surface and at the point where visual responses are measured. Visual responses are audible, time-locked neuronal responses to light flashed in the eyes with an ophthalmoscope or small LED/flashlight.
8. Determine the location and thickness of the dLGN by noting manipulator positions at the start, middle, and end of visual responses and subtracting the cortical surface position from these values. Record the depths for each designed injection site within the dLGN.  
NOTE: Similar procedures can be used for locating non-visual structures by using appropriate sensory stimuli. Additionally, coupled recording/injection systems can also be used to target small brain structures.
9. Once optimum injection locations are noted, remove the recording electrode and place an injection pipette (glass pipette or injection syringe) at the same stereotaxic position and lower to the appropriate depth for each injection, starting with the deepest injection site and proceeding to the shallowest, waiting at least 1 min after injecting before changing electrode position. Glass pipettes with outer diameter of ~ 50 μm may reduce backflow of virus compared to syringe tips (~ 200 μm diameter).
10. Using an injection system (see **Table of Specific Materials/Equipment**, row 4), picospritzer, or syringe pump, inject small volumes (1 - 5 μL; refer to manufacturer instructions for injection procedures) of modified rabies virus over multiple (3 - 10) depths at a rate of ~ 100 nL/min within the target structure (dLGN). Note: Separate injection penetrations can be made in larger target structures (e.g. monkey dLGN). Adjust total injection volume according to target structure size.
11. Pause at least 5 min before retracting the injection pipette. When the injection pipette is removed, fill the craniotomy with protective material (glue bone piece in place, apply bone wax or gelfoam), then suture the muscle and skin back together.
12. Administer analgesics (e.g. ketoprofen) and antibiotics prior to the end of surgery and monitor animal continuously until animal is ambulatory.

- For at least 3 days and up to 10 days following surgery, monitor animal daily to ensure sutures are clean, dry, and intact and animal shows no signs of pain or discomfort. Administer daily analgesics and antibiotics as required. Begin social housing of post-surgical animal after animal has fully recovered from surgery.

## 2. Tissue Harvesting, Sectioning, and Staining

- Allow 7 - 14 days for retrograde transport of virus, then euthanize the animal by overdose of euthanasia solution (e.g. Euthasol) and perfuse the animal transcardially with 0.1 M phosphate buffered saline, 4% paraformaldehyde, then 4% paraformaldehyde with 10% sucrose.
- Remove the brain (see<sup>12</sup> for analogous steps in a rodent model) and place in 20 - 30% sucrose with 4% paraformaldehyde in 0.1 M phosphate buffer and refrigerate for 1 - 2 days.
- When the brain has sunk to the bottom of the container, remove it and cut the tissue into blocks containing the brain region with retrogradely labeled neurons (e.g. visual cortex) and the injection target structure (e.g. dLGN). Cut the same blocks of tissue from the non-injected hemisphere — these will serve as control sections. For best results, begin histological procedures immediately on freshly sectioned tissue.
- Section tissue coronally at a thickness of 50  $\mu$ m per section using a freezing microtome (see **Table of Specific Materials/Equipment**, row 5).
- Stain all sections for cytochrome oxidase activity (see **Table of Specific Materials/Equipment**, rows 5 - 8) to visualize cortical layers and subcortical structures<sup>13</sup>. Note: Sections can be stored overnight in a refrigerator in the final rinse following the cytochrome oxidase stain.
- Label all sections with a primary antibody against GFP (1:1,000 dilution, ~ 12 h incubation; see **Table of Specific Materials/Equipment**, row 9) followed by a secondary antibody matching the primary host and tagged with biotin<sup>5</sup> (1:500 dilution, 2 - 4 h incubation; see **Table of Specific Materials/Equipment**, row 10). It is recommended to incubate sections in the secondary antibody overnight.
- Label sections with an avidin/biotin complex then react with DAB and peroxide to permanently stain all labeled neurons<sup>5</sup>.
- Mount sections on subbed glass slides and let dry overnight.
- Defat sections using a series of alcohol and xylene rinses then cover slip<sup>14</sup>.

## 3. Neuronal Reconstruction

NOTE: All reconstructions for original experiments were made using a neuron reconstruction system made up of a microscope (see **Table of Specific Materials/Equipment**, row 14), attached camera (see **Table of Specific Materials/Equipment**, row 13), and reconstruction software package (see **Table of Specific Materials/Equipment**, rows 11 - 12). Software-assisted neuronal reconstruction enables visualization of tissue slides overlaid with computer-based drawings of neuronal processes. Importantly, the software digitizes morphological reconstruction data in three dimensions, enabling extraction of position-specific morphological information. The associated data extraction program (see **Table of Specific Materials/Equipment**, row 12) enables extraction of a rich set of morphological data from each saved reconstruction.

- Place a slide in the microscope slide holder and focus on a single section of interest using a low-magnification objective (e.g. 2X or 10X). Make sure the camera image is visible in the software view by properly positioning the camera shutter.
- Choose labeled neurons within the brain structure of interest (e.g. TRN) that are reasonably isolated so as to unambiguously reconstruct as much of the dendritic arborization as possible. Preferentially choose neurons if the cell body is fully within a single section in order to capture the largest extent of each cell body and accurately estimate its area and roundness. Use 10X microscope objective to identify cell body in home section.
- Once a well-stained, well-isolated neuron is identified, add the "home" section containing the cell body to the Section Manager by clicking the "New Section" button. Enter the number of sections to include (recommend 2 - 3 sections to begin). Assign a section thickness of 50  $\mu$ m when prompted (see Step 2.3).
- Trace the contours of relevant brain structures in the home section containing the cell body. Use the 2X objective/magnification for contour tracing.
  - First, select "Contour" from the drop-down menu on the top toolbar and then select the contour type from the drop-down menu (e.g. "LGN" or "TRN").
  - Trace contours of the target structure (e.g. TRN) as well as adjacent structures, if desired, by clicking on points along the contour using the mouse as a draw tool.
  - Select "Close Contour" by right-clicking the mouse and selecting this option from the menu to complete and enclose each traced contour.
- Place a marker at the center of the target structure by clicking on the desired marker symbol on the right toolbar and then clicking on the desired location in the reconstruction to place the marker at that location. Use the same type of marker to mark the center of the target structure in all reconstructions.
- Once the contours are complete, trace the outline of the cell body using 40 - 60X objectives/magnification. To do so, first select "Neuron" from the drop-down menu on the top toolbar and then select the neuron structure to trace, in this case "Cell Body". Next trace the cell body as in 3.4.
- Place a different style marker at the center of the cell body, following the steps outlined in 3.5.
- Trace all dendrites beginning at the cell body.
  - First, select "Dendrite" from the drop-down menu. Then trace each dendrite starting at the cell body and using the mouse as a draw tool. Be sure to adjust the z-depth throughout the tracing to accurately capture the angle and direction of the dendrite.
  - Place a node at each branch point along the dendrite by right-clicking the mouse and selecting "Bifurcating Node" or "Trifurcating Node" from the drop-down menu. Use 40 - 60X objective/magnification for all dendrite reconstructions.
  - At the end of each dendrite, right-click the mouse and select "Ending" from the drop-down menu.
- Identify dendritic endings that are likely to continue into the adjacent section and bring these into focus at the appropriate z-depth in the microscope image. Include major landmarks nearby such as blood vessels or easily recognizable dendrite patterns or bundles in the microscope image.

1. Reduce magnification to 20X or 10X on the microscope (choose the magnification that allows greatest landmark visualization), and take a picture of the microscope image using a digital camera (e.g. cell phone, tablet) hand-held to the computer screen.
10. Move to the adjacent section on the slide and line up the contours of the previously traced section with the boundaries of the dLGN and TRN in the new section.
11. Return the field of view to the general area of the cell body (based on contours and major landmarks) and bring the magnification back up to 10 - 20X.
12. Use the photo of the dendrite endings from the previously traced section to aid in aligning the endings with the beginnings of the dendrites in the new section.
  1. To rotate the tracing and move it to align dendrites, use the arrow tool to select the reconstruction, right-click and select "Move" from the drop-down menu.
  2. Alternatively, use the "Match" tool to match dendritic endings to beginnings. Select the Match tool from the top toolbar and when prompted, click on the ending in the reconstruction and the corresponding continuation point in the image. Repeat for 3 or more endings.
13. Once the tracing from the previous section is lined up with the dendrites in the new section, make sure the corresponding new section is selected in the section manager by clicking on the current section. Or, add a new section to the Section Manager, as described above in 3.3, adjusting the z-depth and position of each new section relative to the home section accordingly and setting each section thickness to 50 mm (see Step 2.3).
14. Increase the magnification to 40 - 60X and trace the dendrite continuations by right-clicking the mouse on the end of the dendrite from the reconstruction and selecting "Add to Ending" from the drop-down menu, then tracing the dendrite using the mouse as a draw tool. Follow prompt to specify if continuation is in the new section.
15. Follow dendrites through at least 3 adjacent sections (one on each side of the home section) following Steps 3.8 - 3.15. Add to the reconstruction until at least 3 total sections are traced for the neuron or until the dendrites can no longer be followed or found. Trace contours in neighboring sections following the steps above, if desired.

## 4. Independent Clustering

Note: Independent cluster analyses enable unbiased analyses of large, multi-dimensional datasets that might otherwise be difficult to visualize and, importantly, provide a quantitative assessment of morphological diversity. A matrix-based programming platform is quite useful for the analysis of multi-dimensional datasets and enables sophisticated data manipulations and statistical analyses. The functions listed in steps 4 - 6 are defined in the programming platform listed in the **Table of Specific Materials/Equipment**, row 15.

1. Extract morphological data from each individual neuronal reconstruction using an extraction program associated with the neuron reconstruction system (see **Table of Specific Materials/Equipment**, rows 11 - 12).
  1. First, choose the morphological data desired for each reconstruction including: contour information, marker information, morphological data of cell body and dendritic arborization, etc.
  2. From the Edit drop-down menu in the top toolbar, select "Select all objects". Then select "Branched Structure Analysis" from the Analysis drop-down menu in the top toolbar and click on each tab and select the desired analysis options in each tab.
  3. Extract all desired data by clicking the "OK" button in the Analysis window and save in a spreadsheet format by right-clicking on the output windows and selecting "Export to Excel".
2. Compile a Master spreadsheet (see **Table of Specific Materials/Equipment**, row 16) in which each morphological variable/metric is represented by a single numerical observation per neuron. In this Master spreadsheet, keep a record of the row identifier for each neuron.
3. Verify that all variables included in the cluster analysis (e.g. morphological metrics) are independent of one another. For example, the number of nodes will correlate with the number of branches in a dendritic or axonal arborization. Accordingly, only one of these two variables should be included in a cluster analysis.
4. Make sure each measurement contains an observation for every variable, i.e. each neuron (measurement), must have a data-point (observation) corresponding to each morphological metric (variable). If observations for a particular variable (for example, axon length) are lacking for a subset of neurons, this variable (axon length) cannot be included in the cluster analysis.
5. Organize the data on the Master Spreadsheet into a single matrix in which each row represents a neuron and each column contains numerical data for each morphological metric (**Figure 1**). For example, a dataset including 100 neurons for which there are observations for 5 independent variables would be represented by a 100 x 5 (row-by-column) matrix. It is not necessary to include any identification for each neuron within the matrix - the row index will serve as each neuron's unique identifier. There should also be no gaps or "NaN"s in the matrix. Save the matrix as a single variable (e.g. Data = [100 x 5 matrix]; see **Supplemental Code File** for sample code).
6. Choose an algorithm to calculate distances between points representing individual neurons in an n-dimensional space where n is defined by the number of variables. The 'pdist' function enables flexibility in calculating between-neuron distances. The default distance calculation is Euclidean distance and has been used previously for similar analyses of neuronal morphological data<sup>5,6</sup>.
7. With the between-neuron distances output of the 'pdist' function, use the 'linkage' function to form clusters. Again, multiple clustering options are available. Ward's method and centroid distance have yielded similar results in analyses of morphological datasets<sup>5,6</sup>.
8. If desired, use the 'kmeans' function as an alternative clustering approach to 'pdist' and 'linkage' functions. 'kmeans' employs squared Euclidean distance to calculate between-neuron distances and then centroid distance to assign clusters.
9. To visualize a hierarchical cluster tree, use the 'dendrogram' function on the output of the 'linkage' operation. This function has a default setting that limits visualization to 30 measurements, but it can be overridden by specifying the number of measurements (e.g. neurons) displayed. The dendrogram illustrates the linkage distances on the y-axis for each of the neurons, identified by their row index on the x-axis. NOTE: The outputs of 'linkage' and 'dendrogram' do not define the optimal number of clusters. The output of 'kmeans' does assign measurements to clusters, but it does not test whether these clusters are optimal. Separate statistical comparisons and verification of optimal clustering can be performed (see below).

## 5. Verification of Clustering

Note: As stated above, the cluster analysis itself does not directly provide a statistical assessment of whether the clusters illustrated in the cluster dendrogram are unique and representative of the sample. Methods for verifying clusters from the dendrogram have been proposed<sup>15</sup>, however these do not provide statistical verification of optimal clustering. There are multiple methods for verifying optimal clustering.

1. Method 1: Evaluating clustering:
  1. Use the 'evalclusters' function specifying the method used for calculating clusters, such as 'kmeans' or 'linkage'. A Gaussian mixture model output can also be evaluated (see Step 5.2 below; see **Supplemental Code File** for sample code).
  2. Specify the evaluation criterion from a number of options (e.g. 'CalinskiHarabasz' - for descriptions of criterion options, refer to the Help menu, searching for "evalclusters").
  3. Specify a range of optimal cluster numbers to test (e.g. [1:6] to test whether 1, 2, 3, 4, 5, or 6 clusters are optimal).  
Note: The output of 'evalclusters' is an optimal number of clusters given the clustering algorithm and criterion specified.
2. Method 2: Gaussian mixture model clustering:

NOTE: Gaussian mixture model clustering algorithms (GMMs) assume that observations for each variable come from a mixture of Gaussian distributions defining each putative cluster, each with their own mean and covariance. The GMM then uses an expectation maximization algorithm to assign posterior probabilities to each measurement, indicating the probability of belonging to a specific cluster<sup>16</sup>.

  1. Implement a GMM using the 'fitgmdist' function by inputting the putative number of clusters observable in the data. Alternatively, to avoid *a priori* assignment of the number of clusters, use principal components analysis (PCA) with a series of different optimal cluster numbers using the steps outlined here (see **Supplemental Code File** for sample code).
  2. Use the 'pca' function on the original data matrix and save the principal component scores as an output.
  3. In a loop starting at 1 and going through some number of putative optimal clusters, use the 'fitgmdist' function on the principal component scores and generate GMMs for each number of putative optimal clusters.
  4. Evaluate each GMM by examining the negative log likelihood, Akaike information criterion, and Bayes information criterion. The GMM with the lowest criteria is optimal.
  5. If desired, test whether the mean values for each cluster are significantly different when the number of clusters is optimal (means of each cluster are stored in the "mu" output of each GMM).

## 6. Statistical Analyses of Clustered Data

1. Once the optimal number of clusters is determined using hierarchical clustering and evaluated, return to the original data, separate neurons into clusters, and determine which morphological metrics contribute to the clustering of neurons into unique classes.
2. Sort the original morphological metric data such that neurons are grouped according to cluster assignment.
3. To examine relationships between morphological data across neurons (for all neurons or for neurons separated into clusters), perform linear regression fits to 2- and 3-way comparisons of morphological metric observations. These fits can be estimated using the 'fit' function and evaluated using the 'fitlm' function in which goodness of fit outputs include  $R^2$  and p values for regression fits.
4. To examine the statistical relationships between neurons in each cluster, use non-parametric two-sample (Wilcoxon rank-sum test) or multiple-sample comparisons tests (ANOVA), depending on the number of clusters. Importantly, the use of multiple-sample ANOVAs ensures that p values are corrected for multiple comparisons.

### Representative Results

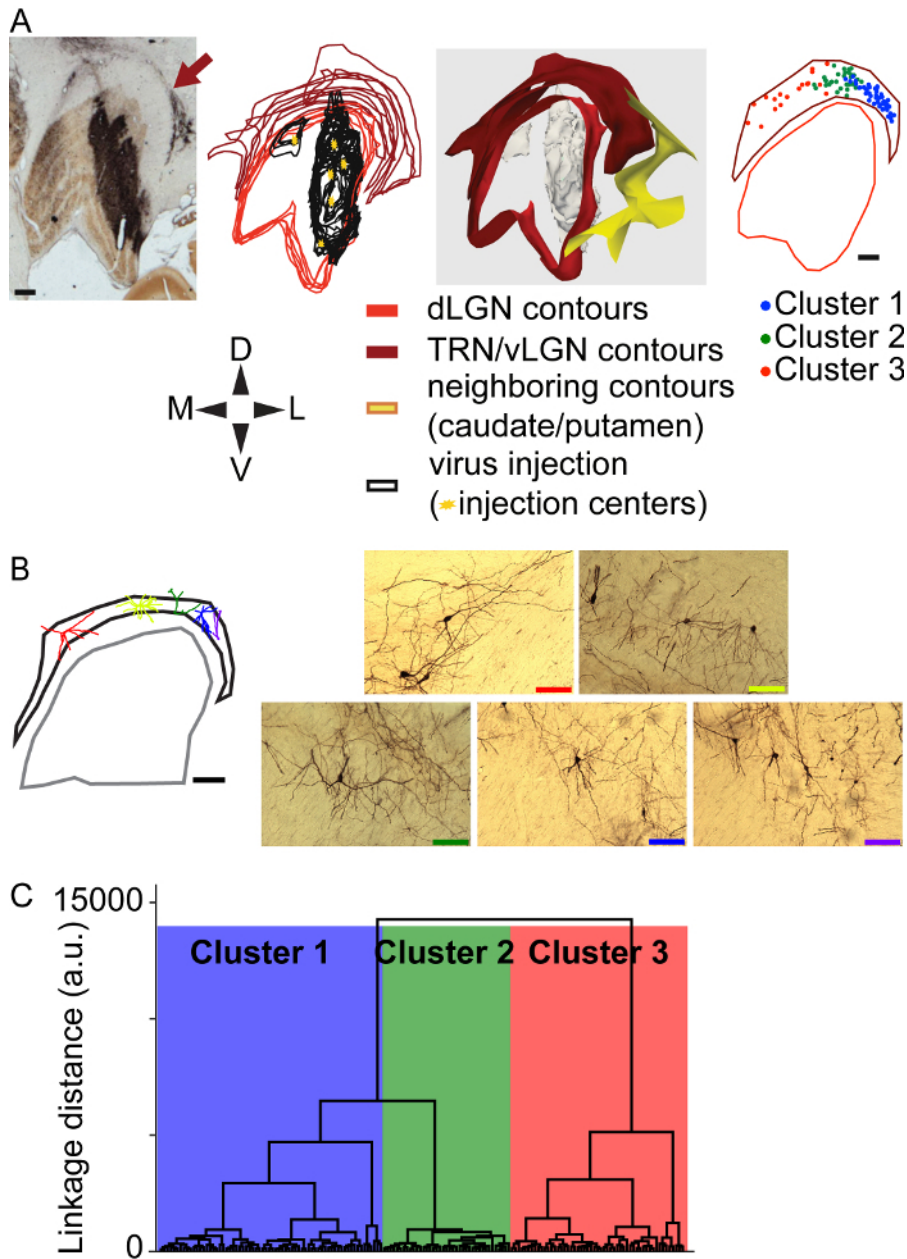
We have shown previously that large-scale reconstructions of neurons within a selective population is feasible following injection of modified rabies virus into the dLGN<sup>5</sup>. Recently, the same tissue was utilized to reconstruct 160 neurons in the visual sector of the TRN (Bragg *et al.*, in review; **Figure 2A-B**) following the detailed methodological steps described above. In the TRN study, three unique clusters of TRN neurons were identified based on independent cluster analysis of 10 morphological metrics: cell body area, cell body roundness, medial-lateral position relative to the center of the TRN, dorsal-ventral position within the TRN, number of dendritic trees, average dendritic distance to nodes, average length of 3<sup>rd</sup> and higher order dendrites, average angle of 1<sup>st</sup> order dendrites, average phase angle of the total dendritic arborization, and angular deviation of the total dendritic arborization. TRN neurons were equivalently distributed across these three clusters (**Figure 2C**). Separate statistical comparisons of all 10 morphological metrics across neurons in the three clusters yielded statistically significant differences across clusters for all but two of the morphological metrics included in the original cluster analysis. Thus, based on the hierarchical tree dendrogram generated from the cluster analysis (**Figure 2C**) and the separate statistical comparisons of morphological metrics across clusters, TRN neurons are classified into three unique groups.



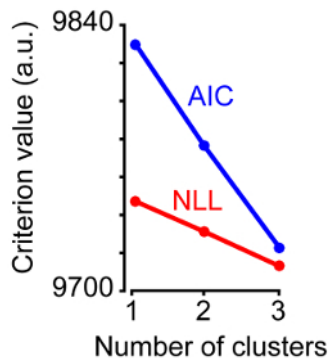
In this study, we specifically wanted to apply statistical methods to separately evaluate clusters determined with a cluster analysis. The TRN dataset from the prior study was evaluated to determine whether the three clusters observed were optimal. Separate additional cluster analyses, namely PCA and GMM clustering, were performed to verify the prior clustering methods (**Figure 1**). Three key outcomes separately verify the prior cluster analysis. Firstly, when the original cluster analysis method was evaluated using the 'evalclusters' function specifying the 'linkage' function to generate clusters, the optimal number of clusters was 3, matching the original conclusion based on the hierarchical tree dendrogram (**Figure 2C**). Secondly, PCA was performed on the original data matrix of 10 morphological metrics for 160 TRN neurons as an alternative means of grouping neurons based on contributions of morphological metrics. Three GMMs were generated, assuming TRN neurons were grouped into 1, 2, or 3 unique clusters. The negative log likelihood (NLL) and Akaike information criterion (AIC) yielded the lowest, and therefore most informative, values and both values were optimal when the GMM used 3 clusters to describe TRN morphological data (**Figure 3**). Thirdly, the GMM-based clustering was separately evaluated using the 'evalclusters' function and the optimal number of clusters was 3. Thus observations of the hierarchical tree dendrogram, statistical comparisons of morphological metrics across neurons separated into the three clusters, PCA and GMM-based clustering, and separate evaluations of each clustering method (linkage and GMM), all yielded the same outcome that TRN neurons are optimally separated into 3 clusters based on the 10 morphological metrics utilized.

Cell body area	Cell body roundness	Total length dendrite	% dendrite in layer 4	Total length 3 <sup>rd</sup> order dendrite
Value for neuron 1	Value for neuron 1	Value for neuron 1	Value for neuron 1	Value for neuron 1
Value for neuron 2	Value for neuron 2	Value for neuron 2	Value for neuron 2	Value for neuron 2
Value for neuron 3	Value for neuron 3	Value for neuron 3	Value for neuron 3	Value for neuron 3
...				

**Figure 1: Data Format.** Example Data matrix with n rows corresponding to the number of neurons and m = 5 columns corresponding to the total number of independent variables to include in the cluster analysis (do not include headers in Data matrix).



**Figure 2: Independent Clustering of 160 TRN Neurons.** **A.** Left, photograph of single coronal section through the dLGN of one animal, stained for cytochrome oxidase activity to visualize dLGN layers and against GFP to visualize virus injection. Arrow indicates regions of dense retrogradely labeled TRN neurons. Section orientation follows the dorsal-ventral/medial-lateral (DV/ML) compass below and scale bar represents 500 µm for all panels in **A**. Second from left, contour outlines of dLGN (red) and TRN (maroon) for all sections containing virus injection (black contours). Yellow stars indicate injection site centers. Orientation and scale bars as in **A**. Third from left, 3-d renderings of contours and injection site, orientation and scale bars as in **A**. Right, maps of locations of reconstructed TRN neurons, color-coded according to cluster assignment (**C**) within a single aggregate TRN contour (maroon). Orientation and scale bars as in **A**. **B.** Left, aggregate contours of TRN (black) and dLGN (grey) with 5 reconstructed TRN neurons colored warm to cool according to their medial-lateral position within the TRN. Orientation according to the DV/ML compass in **A**, scale bar represents 500 µm. Right, photographs of the same 5 TRN neurons with color-matched scale bars representing 100 µm. **C.** Hierarchical tree dendrogram following cluster analysis illustrating linkage distances between 160 reconstructed TRN neurons based on 10 independent morphological metrics. Three clusters illustrated in blue, green, and red. All portions of this figure are adapted from figures included in Bragg *et al.* (in review). [Please click here to view a larger version of this figure.](#)



**Figure 3: Cluster Evaluation.** Plot of two different criterion values, negative log likelihood (NLL, red) and Akaike information criterion (AIC, blue) generated from three GMMs assuming 1, 2, or 3 clusters. GMMs with 3 clusters provide the lowest criterion values.

**Supplemental Code File:** Example code, written in a Matlab-compatible language, to perform cluster analysis on an example Data matrix followed by cluster evaluation using PCA and GMM approaches. [Please click here to download this file.](#)

## Discussion

Neuroanatomical studies have remained a pillar of neuroscience and recent interest in connectomics and structure-function relationships has renewed enthusiasm for detailed morphological characterization of selective neuronal populations. Traditionally, neuroanatomical studies have relied on qualitative classifications of neurons into morphologically distinct classes of neurons defined by expert neuroanatomists. With advances in the techniques for reconstructing neurons and extracting morphological data, it is now possible to utilize more sophisticated and quantitative data analysis methods to classify morphologically distinct neuronal classes in an unbiased manner. In this study, step-by-step methods are outlined for 1) selectively labeling hundreds of individual neurons using virus-mediated retrograde labeling methods; 2) systematically reconstructing hundreds of individual neurons to form a comprehensive and therefore representative sample of neurons within a selective population; and 3) independent clustering analyses with statistical evaluation to determine optimal clusters of neurons within a selective population based on morphological metrics. These analysis methods are verified by applying different clustering methods to an existing dataset of TRN neurons and we demonstrate based on three separate evaluation methods that TRN neurons are optimally clustered into three morphologically distinct subpopulations.

A major advantage of the protocols outlined in this study is their flexibility. The G-deleted rabies virus has been effective in driving EGFP expression in hundreds of retrogradely labeled neurons following injection into a target structure. The same rabies virus strain is similarly effective in multiple species including macaque monkeys<sup>5,9</sup>, ferrets (Hasse & Briggs, in revision), and rodents<sup>7,8</sup>. Accordingly, modified rabies virus can be utilized to selectively label the complete dendritic arborization patterns of any population of neurons following injection into a target structure. Antibody staining against GFP followed by DAB/peroxide reaction is recommended because this causes permanent staining of labeled neurons and enables tissue sections to be visualized under a microscope without the aid of fluorescence microscopy (which bleaches fluorescent label with longer exposure). However, alternative staining methods are available and/or fluorescence can be directly visualized if desired. An advantage of measuring raw fluorescence is that different viruses driving expression of different fluorescent molecules can be combined, for example to determine whether neurons project to multiple target brain structures. Additionally, the steps outlined above may be applied to all available morphological data (for example axons and dendrites) to enable increasingly robust morphological classification of neurons. It is also important to note that smaller quantities of rabies virus can also be injected in order to generate sparser retrograde labeling, which could be advantageous for reconstructions of axons in addition to dendrites.

An additional advantage of the analysis methods outlined in this study is that multiple different clustering approaches can be utilized and compared. At each step in the cluster analysis, different options are available to calculate between-neuron distances in parameter space as well as different algorithms to generate clusters. Additionally, multiple evaluation methods are available. The latter is especially useful as a means to verify optimal clustering, as outlined in the Representative Results section. Undoubtedly, additional algorithms for generating clusters and for evaluating optimal clustering will become available as the field moves forward. Together, these analysis methods will continue to improve quantitative approaches in the field of neuroanatomy and enhance the robustness of findings.

There is growing interest in automation of neuronal reconstructions because by-hand reconstructions of individual labeled neurons are time consuming, require training and experience, and are still somewhat subjective. Because computer-based automatic neuronal reconstructions are still error-prone, electron microscopy-based neuroanatomical data are still analyzed by people (see for example<sup>17</sup>). However, it is likely that automated neuronal reconstruction will be available in the near future. Automated neuronal reconstruction will require even more scrutiny on the data analysis end, thus the clustering analyses and evaluation methods outlined here will become even more important as the field continues to advance technologically.

## Disclosures

The authors have nothing to disclose.



## Acknowledgements

We would like to thank Drs. Ed Callaway and Marty Usrey for allowing us to use the tissue prepared as a part of a prior study and Libby Fairless and Shiyuan Liu for help with neuronal reconstructions. This work was funded by the NIH (NEI: EY018683) and the Whitehall Foundation.

## References

1. Cajal, S. R. y. *Histologie du systeme nerveaux de l'homme et des vertebres*. Maloine (1911).
2. Seung, H. S. Toward functional connectomics. *Nature*. **471**, 170-172 (2011).
3. Callaway, E. M. Transneuronal circuit tracing with neurotropic viruses. *Current Opinion in Neurobiology*. **18**, 1-7 (2009).
4. Wickersham, I. R., Finke, S., Conzelmann, K. K., & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nature Methods*. **4**, 47-49 (2007).
5. Briggs, F., Kiley, C. W., Callaway, E. M., & Usrey, W. M. Morphological substrates for parallel streams of corticogeniculate feedback originating in both V1 and V2 of the macaque monkey. *Neuron*. **90**, 388-399 (2016).
6. Cauli, B. *et al.* Classification of fusiform neocortical interneurons based on unsupervised clustering. *PNAS*. **97**, 6144-6149 (2000).
7. Bragg, E.M., Fairless, E.A., Liu, S., Briggs, F. Morphology visual sector thalamic reticular neurons in the macaque monkey suggests retinotopically-specialized, parallel stream-mixed input to the lateral geniculate nucleus. *J. Comparative Neurology*. **525** (5) 1273-1290 (2017).
8. Osakada, F. *et al.* New rabies virus variants for monitoring and manipulating activity and gene expression in defined neural circuits. *Neuron*. **71**, 617-631 (2011).
9. Callaway, E. M., & Luo, L. Monosynaptic circuit tracing with glycoprotein-deleted rabies viruses. *Journal of Neuroscience*. **35**, 8979-8985 (2015).
10. Nhan, H. L., & Callaway, E. M. Morphology of superior colliculus- and middle temporal area-projecting neurons in primate primary visual cortex. *J. Comparative Neurology*. **520**, 52-80 (2012).
11. Pinault, D. The thalamic reticular nucleus: structure, function and concept. *Brain Research Reviews*. **46**, 1-31 (2004).
12. Gage, G. J., Kipke, D. R., & Shain, W. Whole Animal Perfusion Fixation for Rodents. *Journal of Visualized Experiments*. **65**, e3564 (2012).
13. Wong-Riley, M. Changes in the visual system of monocularly sutured or enucleated cats demonstrable with cytochrome oxidase histochemistry. *Brain Research*. **171**, 11-28 (1979).
14. Gerfen, C. R. Basic neuroanatomical methods. *Current Protocols in Neuroscience*. **Chapter 1**, Unit 1.1 (2003).
15. Thorndike, R. L. Who belongs in the family? *Psychometrika*. **18**, 267-276 (1953).
16. Talebi, V., & Baker, C. I. Categorically distinct types of receptive fields in early visual cortex. *Journal of Neurophysiology*. **115**, 2556-2576 (2016).
17. Helmstaedter, M. *et al.* Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*. **500**, 168-174 (2013).