An open-source computational tool to automatically quantify immunolabeled retinal ganglion cells

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>10.1016/j.exer.2016.04.012</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:34216313">http://nrs.harvard.edu/urn-3:HUL.InstRepos:34216313</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Methods in eye research

An open-source computational tool to automatically quantify immunolabeled retinal ganglion cells

Ana C. Dordea a, Mark-Anthony Bray b, Kaitlin Allen a, David J. Logan b, Fei Fei c, Rajeev Malhotra d, Meredith S. Gregory c, Anne E. Carpenter b, Emmanuel S. Buys a, b, *  

a Anesthesia Center for Critical Care Research, Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital Research Institute, Harvard Medical School, Boston, MA, USA  
b Broad Institute of MIT and Harvard, Cambridge, MA, USA  
c Department of Ophthalmology, Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA  
d Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

**Abstract**

A fully automated and robust method was developed to quantify β-III-tubulin-stained retinal ganglion cells, combining computational recognition of individual cells by CellProfiler and a machine-learning tool to teach phenotypic classification of the retinal ganglion cells by CellProfiler Analyst. In animal models of glaucoma, quantification of immunolabeled retinal ganglion cells is currently performed manually and remains time-consuming. Using this automated method, quantifications of retinal ganglion cell images were accelerated tenfold: 1800 images were counted in 3 h using our automated method, while manual counting of the same images took 72 h. This new method was validated in an established murine model of microbead-induced optic neuropathy. The use of the publicly available software and the method’s user-friendly design allows this technique to be easily implemented in any laboratory.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Glaucoma is the leading cause of irreversible blindness worldwide, affecting an estimated 60 million people (Tham et al., 2014). While there are many forms of glaucoma, all are associated with an optic neuropathy characterized by the loss of retinal ganglion cells (RGCs) and their axons, resulting in optic nerve degeneration and irreversible vision loss. Animal models of glaucoma simulate the optic neuropathy observed in human disease facilitate the elucidation of possible mechanisms of RGC loss and enable researchers to develop and evaluate neuroprotective therapies. The ability to specifically identify and accurately count RGCs is essential to assess the death or survival of RGCs in models of the disease.

Various techniques for visualizing and quantifying RGCs have been reported, including retrograde labeling and immunolabeling (Buckingham et al., 2008; Huihui et al., 2011). In a healthy, non-diseased retina, fluorogold retrograde labeling is a specific and accurate method to label and automatically quantify RGCs (Buckingham et al., 2008; Danias et al., 2003). However, this method is technically challenging, requiring intracranial surgery and, in a glaucomatous retina, retrograde labeling can result in the labeling of retinal microglia in addition to RGCs, due to the phagocytosis of degenerating RGCs (Peinado-Ramon et al., 1996; Thanos, 1991a,b). An alternative quantification method of RGC viability may be carried out with immunolabeling of RGCs in retinal flatmounts using an RGC-specific antibody directed against β-III-tubulin, which labels the RGC somata and nerve fiber extensions. The specificity of the β-III-tubulin antibody has been confirmed by colocalization of β-III-tubulin staining with fluorogold-labeled RGCs (Huihui et al., 2011).

While the immunolabeling method is technically simpler than fluorogold labeling, manual quantification of immunolabeled murine RGCs is onerously time-consuming. In a small-scale pre-clinical project with only 2 experimental groups, an investigator would be required to manually quantify RGCs in approximately 1000 images, necessitating at least 48 h of effort. In addition, manual counting can be inconsistent, with significant inter- and even intra-observer variability, due to the use of differing quantification techniques and/or inhomogeneous staining.
2. Materials and supplies

2.1. Animals

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institute of Health. Wild-type mice bred on a 129S6 background (129S6/SvEvTac, Taconic Farms Inc.) were used in this study. All animals were treated in accordance with the Institutional Animal Care and Use Committees (IACUC) of Massachusetts General Hospital (Subcommittee on Research Animal Care), and the Schepens Eye Research Institute.

2.2. Processing and imaging of retinas

Animals were sacrificed under CO2. Eyes were enucleated and retinas dissected from the anterior segments at the conclusion of the study: day 32 post-injection. Resultant retinal cups were incised to create four quadrants of similar size and were fixed in 4% paraformaldehyde at 4 °C overnight. Retinas were then treated with 1% Triton-X-100 and 5% fetal BSA in PBS for 1 h, followed by a 2 h incubation with DAPI (1:500) and the primary antibody against an RGC marker, β-III-tubulin (anti-TUJ1®, Millipore, Billerica, MA, 1:500), and 1 h incubation with the Alexa Fluor® 594 Goat Anti-Mouse IgG secondary antibody (Life Technologies, Carlsbad, CA, 1:500) at room temperature. Retinal whole-mounts were then flattened on SuperFrost Plus slides (VWR, Batavia, IL), coverslipped with mounting medium for fluorescence (VectaShield®, Vector Laboratories, Burlingame, CA) and imaged under the Leica TSC SP5 confocal microscope at ×63 magnification. Imaging was performed on the mid-peripheral area of the retina (around 0.5 mm distal from the optic nerve head) divided into 4–5 distinct areas across all four quadrants. Each retina was imaged in 20–25 frames of 0.0696 mm².

2.3. Microbead injections

Mice were anesthetized by intraperitoneal injection of a mixture of ketamine (100 mg/kg; Ketaset; Fort Dodge Animal Health, Fort Dodge, IA) and xylazine (9 mg/kg; TranquiVed; Vedco, Inc., St. Joseph, MO) and eyes were dilated by topical application of proparacaine (0.5%; Bausch & Lomb, Tampa, FL). Elevation of IOP was induced unilaterally in adult 129S6 mice by injection of polystyrene microbeads (FluoSpheres; Invitrogen, Carlsbad, CA; 15 μm diameter) into the anterior chamber of the right eye of each animal using the micropipette. Mice were placed on a heating pad for recovery after the injection, and antibiotic Vetropolycin ointment (Dechra Veterinary Products, Overland Park, KS) was applied topically onto the injected eye to prevent infection.

2.4. IOP measurement

Mice were anesthetized with isoflurane inhalation (2%), which was delivered in 95% oxygen with a precision vaporizer. IOP measurement was initiated within 2 min after animal lost toe pinch and blink reflex. IOPs were acquired with a TonoLab rebound tonometer (iCare, Franconia, NH, USA). Five TonoLab readings were averaged to obtain a single IOP value per eye. IOP measurements were carried out at days 4, 10, 15, 18, 22 and 28 of the 32-day study.

2.5. Statistical analyses

All statistical analyses were performed using GraphPad Prism (version 6.0). Statistical tests included the Student’s t-test, Pearson correlation, or multivariate linear regression. In order to compare correlation coefficients, the Fisher method of r-to-z transformation was employed, using the following equation:

\[ z_n = \frac{1}{2} \ln \left( \frac{1 + r}{1 - r} \right) \]

Pairwise correlations with their respective samples sizes (\(n_1\) and \(n_2\)) were compared with the following test statistic:

\[ Z = \frac{(Z_1 - Z_2)}{\sqrt{\frac{1}{(n_1-3)} + \frac{1}{(n_2-3)}}} \]

Correlation dispersion was used to quantify differences between automated and manual counts for each retina. Correlation dispersion was determined using the following equation:

\[ CD = \sqrt{(y_{\text{manual}} - y_{\text{automated}})^2} \]

where \(y\) represents the cell count in RGC/mm². 

\(P < 0.05\) was considered significant throughout the study. Data are presented as mean ± s.d.

3. Detailed methods

3.1. Generation of a rapid and robust automated quantification technique

To minimize variability and enable faster and more efficient quantification of immunolabeled RGCs, we sought to develop and validate a reproducible automated quantification method. The approach uses the freely-available softwares, CellProfiler and CellProfiler Analyst (CPA) (Carpenter et al., 2006; Kamentsky et al., 2011; Lamprecht et al., 2007), and can be used for high-throughput image analysis. The CellProfiler image analysis template was designed to recognize all DAPI-positive (nuclear marker) and β-III-tubulin-labeled cylindrical RGC somata (Box 1).

The workflow was as follows: the CellProfiler template was loaded and images from experimental set-ups were fetched into the “File list”. Automated processing of each image was initiated (“Analyze Images”): the morphological features for each cell, including the nuclei and cellular shapes, as well as intensity-based and textural features from both the DAPI and β-III-tubulin channels were recognized and measured by CellProfiler; the complete set of these measurements for each cell is defined as the “cytoprofile”

Box 1

CellProfiler software.

The CellProfiler software is freely available to download at www.cellprofiler.org. The image analysis template used for the automated quantification presented in the manuscript, the latest software updates and the source code are available to download at http://www.cellprofiler.org/published_pipelines.shtml under the “β3-tubulin_RetinalGanglionCell_Assay.cpproj” nomenclature.
of the original color (red-green-blue (RGB), Supplementary Fig. 2A and D) to an 8-bit grayscale, set to a threshold intensity of 89 and inverted so that fluorescent cells would appear as black objects on a white background (Supplementary Fig. 2B and E). Particles were then analyzed to assess circular areas of pixels (with thresholds set to include Circularity 0.05–1.00 and Size in pixel units 200–Infinity), denoting circular RGCs within same-sized frames in “high-quality/ included” (Supplementary Fig. 2C) and “low-quality/excluded” images (Supplementary Fig. 2F). This method highlighted the significant differences in circular pixilation areas between high-quality and low-quality images (Supplementary Fig. 2C and F and Supplementary Table 1), irrespective of the RGC counts obtained per image. Images with a circularity score <0.13 (selected cut-off value; Supplementary Fig. 2 and Supplementary Table 1) were deemed of insufficient quality.

The impact of image quality on the reproducibility of both manual and automated methods was studied in 18 selected retinas (21 images per retina): nine retinas representing the highest quartile of correlation dispersions ($\Delta$398 ± 80.8 RGC/mm$^2$) and nine retinas representing the lowest quartile of correlation dispersions ($\Delta$701 ± 86.8 RGC/mm$^2$) (Fig. 2A, see “Statistical analyses” for a description of correlation dispersions calculations). Using the unbiased ImageJ quality assessment platform, 272 images of the total 378 images from 18 retinas were categorized as included images (Fig. 3A and B), while 28% (106 images) were labeled as excluded images (Fig. 3C and D). The linear regression fit between manual and automated quantifications improved markedly for included images ($r^2 = 0.64$, Fig. 3E) as compared to excluded images ($r^2 = 0.22$, Fig. 3F), highlighting the advantage of a quality control step prior to RGC quantifications, whether performed manually or automatically.

Additionally, to investigate whether observer bias may have affected the supervised classification quality, two researchers independently counted images manually, and carried out automated quantification using CellProfiler and the machine-learning tool in CPA. The two observers’ counts were strongly correlated for included ($r^2 = 0.87$) and excluded ($r^2 = 0.83$) image groups using automated counting ($r = 1.25$, Fig. 3G and I, respectively), while correlations were lower for manual quantification, with $r = 0.70$ for included images and $r = 0.47$ for the excluded images ($r = 3.08$, Fig. 3H and J, respectively). Consistent quantification of RGCs was achieved using both techniques. Importantly, the automated counting method provided greater reproducibility for images of lower quality as compared to the manual quantification method (Fig. 3I and J), highlighting the robustness of our automated quantification method. Particularly advantageous is the fact that the automated approach was considerably faster than the manual method; manual counting of 1800 images took one observer 72 h to complete, while automated quantification was achieved in 3 h.

### 3.3. In vivo validation of the computational quantification method

To validate our approach, the automated counting method was applied to study RGC loss in an established mouse model of optic neuropathy, the microbead occlusion model (Huihui et al., 2011; Sappington et al., 2010). Injection of 15 μm-diameter microbeads into the anterior chamber blocks the outflow of aqueous humor, resulting in a significant increase in intraocular pressure (IOP) for up to three weeks post-injection and a 25–40% loss of RGCs (Huihui et al., 2011). The microbead-injected eyes exhibited a significant increase in IOP over a period of 18 days compared to un.injected contra-lateral eyes (average changes in IOP over 18 days: $14.0 \pm 8.8$ mmHg to $22.1 \pm 11.1$ mmHg, $n = 6$, $P < 0.05$). Both our automated approach and manual quantification were able to identify significant neuropathy at the conclusion of the 32-day...
Fig. 1. Automated scoring of cell morphologies using the CellProfiler Analyst machine learning and iterative feedback system. (A) The software system presents the researcher with individual cells to classify from segments of images, sampled randomly from the set of images. After classification of a few cells, the researcher begins the iterative machine-learning phase (“Train Classifier”), in which the software generates a tentative rule based on the classified cells and presents the researcher with cells classified according to that rule. (B) The accuracy of the CPA classifier rule may be checked after each training set (“Check Progress”) and accuracy above 80% is considered suitable. (C) Following 4 training
Box 2
(Fiji is Just) ImageJ.

The (Fiji is Just) ImageJ software is freely available to download at http://fiji.sc/Fiji. The Cell Counter plugin information used for manual quantifications of RGCs in this study is available at http://fiji.sc/Cell_Counter.

Fig. 2. (A) Correlation between manual and automated quantifications of 72 murine retinas. Each count represents the averages of 20–25 images per retina. Counts are presented as number of RGC/mm². Green square symbols: 9 retinas with highest quartiles of correlation dispersion (well-correlated manual and automated counts). Red round symbols: 9 retinas with lowest quartiles of correlation dispersion (poorly-correlated manual and automated counts). (B) “High-quality” image with clean RGC body delineations. (C) Image with discordant/low staining or underexposure to highlight the RGCs. (D) Image of a fiber-rich uneven exposure, possibly due to the retinal structure not being fully flattened onto the imaging slide.

study, with comparable RGC loss (29.2 ± 9.1% vs. 29.8 ± 10.3% loss of RGC/mm², Fig. 4A and B, respectively). Manual and automated quantifications were also correlated ($r^2 = 0.61$, Fig. 4C), highlighting the efficacy and translatability of our fully automated quantification method.

4. Potential pitfalls and troubleshooting

The CellProfiler software provides a platform to recognize cells according to texture, size and color; the training set in CPA is the tool enabling reproducible quantifications. Since CPA is initialized and overseen by the researcher, setting up a template to accurately recognize and measure β-III-tubulin-positive RGC somata and training for recognition of “positive” and “negative” cells to “include” or “exclude” in quantifications remain relatively subjective. However, subjectivity is also inherent to manual quantifications, in which the researcher arbitrarily decides which cells to count. The CPA training set will consistently quantify cells assigned as “positive” throughout images of experiments, while, during manual quantifications, human error may inadvertently occur and subjectivity may fluctuate between 2 researchers and/or from one set of images to another. The greater reproducibility of the CPA was highlighted with greater correlations between automated quantifications from 2 independent researchers than those between manual quantifications from the same 2 investigators.

Quantifications were also influenced by the quality of the image. In an experimental setting, heterogeneous immunolabeling of β-III-tubulin-positive cells occurs, affecting the ability of the CPA tool to accurately count cells. An ImageJ-based quality control was established to filter high-quality images for quantification. It is important to note that, while performing the ImageJ-based quality control improved the correlation between manual and automated counts, foregoing the use of ImageJ prior to quantifying images using CellProfiler and CPA would not corrupt results obtained, as automated quantifications of cells using poor quality images remained more reliable than those from manual counting. Moreover, the number of training sets within CPA may be increased to refine the “RGC recognition rule” for image sets that present with higher numbers of lower quality images.

In the validation cohort (Fig. 4), absolute RGC numbers were higher when counted using CellProfiler and CPA than when counted manually. This discrepancy may be the result of a combination of subjectivity in manual counts throughout images and definition of inclusion and exclusion criteria in the CPA training phase. Previous studies reported RGC counts similar to the average of 3,019 RGC/mm² we obtained using the automated method, with various staining methods in various strains of mice (Table 1). Our data, together with published data, suggest that our manual count may have underestimated the number of RGCs present. In general, it remains challenging to determine absolute numbers of RGCs: manual quantifications are affected by inter-observer variability due to subjective observer error. Importantly, the reproducibility of the automated method was validated in an established model of glaucoma, in which relative differences in RGC counts (a primary endpoint in several studies (Chen et al., 2011; Huihui et al., 2011; Yang et al., 2012)) between microbead-injected and uninjected eyes were comparable using the presented automated and the manual methods. Taken together, our data suggest that automated counting is at least as reliable as, but much less time consuming and prone to inter-observer variability, than manual counting.

Finally, while the presented automated counting platform was designed for β-III-tubulin immunolabeled RGC recognition, the CellProfiler template features may recognize and process additional cell sizes and shapes or be modified to do so as necessary. The use of the machine-learning tool in CPA may then be taught to implement the new cell features. For instance, staining with Brn3a, an RGC-
specific protein and alternative to β-III-tubulin staining, has been highlighted as a dependable identification method of RGC somata alone (Leahy et al., 2004; Schlamp et al., 2013). The current pipeline set up in CellProfiler was altered to recognize Brn3a-stained RGCs and cells were successfully recognized and quantified by CellProfiler (Supplementary Fig. 3), with RGC numbers averaging...
2,546 ± 425 RGC/mm² (172.1 ± 27.9 RGC/image) and 2,586 ± 413 RGC/mm² (174.8 ± 28.7 RGC/image) for manual and automated quantifications, respectively, which is similar to previous reports of Brn3a-stained RGC counts (2,889 RGC/mm²) (Nadal-Nicolas et al., 2015).

In summary, a reliable investigator-supervised counting tool was developed for automatically quantifying immunolabeled RGC somata in a highly time-efficient manner. This freely available protocol enabled rigorous quantification of β-III-tubulin-stained RGCs for unbiased assessments of relative changes in RGC numbers in pathological conditions (e.g. between control animals and animals treated with an experimental therapy). This approach may also serve to reconcile intra- and inter-observer variability.

Author contributions

A.C.D. and E.S.B. conceived the project. M.A.B., D.L., A.C.D. and A.E.C. developed the cell recognition pipeline. A.C.D. and K.A. carried out manual and automated quantifications. F.F. and M.G.K. carried out microbead injections. R.M. advised on statistical analyses and data interpretation. A.C.D. analyzed all the data. A.C.D. wrote the manuscript. E.S.B., M.G.K and A.E.C. reviewed the manuscript.

Acknowledgments

This work was supported by NIH grants: [F.F. and M.G.K.] EY021543, [A.C.D., K.A. and E.S.B.] EY022746-01 and [M.A.B. and A.C.] GM089652. Jie Ma provided assistance with Brn3a staining.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jexer.2016.04.012.

References


