

RESEARCH ARTICLE

Interocular symmetry, intraobserver repeatability, and interobserver reliability of cone density measurements in the 13-lined ground squirrel

Benjamin S. Sajdak^{1,2,3}, Alexander E. Salmon², Rachel E. Linderman², Jenna A. Cava¹, Heather Heitkotter², Joseph Carroll^{1,2,4*}

1 Ophthalmology & Visual Sciences, Medical College of Wisconsin, Milwaukee, WI, United States of America, **2** Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI, United States of America, **3** Morgridge Institute of Research, Madison, WI, United States of America, **4** Biophysics, Medical College of Wisconsin, Milwaukee, WI, United States of America

* jcarroll@mcw.edu



OPEN ACCESS

Citation: Sajdak BS, Salmon AE, Linderman RE, Cava JA, Heitkotter H, Carroll J (2019) Interocular symmetry, intraobserver repeatability, and interobserver reliability of cone density measurements in the 13-lined ground squirrel. *PLoS ONE* 14(9): e0223110. <https://doi.org/10.1371/journal.pone.0223110>

Editor: Tudor C Badea, National Eye Centre, UNITED STATES

Received: February 8, 2019

Accepted: September 15, 2019

Published: September 26, 2019

Copyright: © 2019 Sajdak et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its supporting files.

Funding: Research reported in this manuscript was supported by the Foundation Fighting Blindness consortium PPA-0641-0718-UCSF <<https://www.blindness.org/>>, and by the National Eye Institute of the National Institutes of Health (NIH, Bethesda, MD, USA) under award numbers U24EY029891, U01EY025477, P30EY001931 and T32EY014537 to JC <<https://nei.nih.gov/>>. The content is solely

Abstract

Background

The 13-lined ground squirrel (13-LGS) possesses a cone-dominant retina that is highly amenable to non-invasive high-resolution retinal imaging. The ability for longitudinal assessment of a cone-dominant photoreceptor mosaic with an adaptive optics scanning light ophthalmoscope (AOSLO) has positioned the 13-LGS to become an accessible model for vision research. Here, we examine the interocular symmetry, repeatability, and reliability of cone density measurements in the 13-LGS.

Methods

Thirteen 13-LGS (18 eyes) were imaged along the vertical meridian with a custom AOSLO. Regions of interest were selected superior and inferior to the optic nerve head, including the cone-rich visual streak. Non-confocal split-detection was used to capture images of the cone mosaic. Five masked observers each manually identified photoreceptors for 26 images three times and corrected an algorithm's cell identification outputs for all 214 images three times. Intraobserver repeatability and interobserver reliability of cone density were characterized using data collected from all five observers, while interocular symmetry was assessed in five animals using the average values of all observers. The distribution of image quality for all images in this study was assessed with open-sourced software.

Results

Manual identification was less repeatable than semi-automated correction for four of the five observers. Excellent repeatability was seen from all observers (ICC = 0.997–0.999), and there was good agreement between repeat cell identification corrections in all five observers (range: 9.43–25.71 cells/degree²). Reliability of cell identification was significantly different

the responsibility of the authors and does not necessarily represent the official views of the NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

in two of the five observers, and worst in images taken from hibernating 13-LGS. Interocular symmetry of cone density was seen in the five 13-LGS assessed. Image quality was variable between blur- and pixel intensity-based metrics.

Conclusions

Interocular symmetry with repeatable cone density measurements suggest that the 13-LGS is well-suited for longitudinal examination of the cone mosaic using split-detection AOSLO. Differences in reliability highlight the importance of observer training and automation of AOSLO cell detection. Cone density measurements from hibernating 13-LGS are not repeatable. Additional studies are warranted to assess other metrics of cone health to detect deviations from normal 13-LGS in future models of cone disorder in this species.

Introduction

Development and translation of treatment strategies for diseases involving cone degeneration has been limited by accessibility of animal models that mimic human pathophysiology. There are several cone photoreceptor disorder models involving nocturnal mice and rats [1], despite these animals having sparse cone mosaics. As they possess a cone-exclusive fovea, non-human primates like the macaque and marmoset may be more appropriate models [2], but the cost and logistics of maintaining non-human primates is too burdensome for most research institutions. We have been examining a supplemental strategy of studying small, diurnal, cone-dominant mammalian models for investigation of retinal health and disease. For example, the 13-lined ground squirrel (13-LGS) is a cone-dominant mammal (~85% cones) that is highly amenable to non-invasive imaging of the cone mosaic using adaptive optics scanning light ophthalmoscopy (AOSLO) [3, 4]. Central to determining the suitability of these animals as models for studying cone structure in health and disease is assessing the repeatability and reliability of quantitative measurements of the cone mosaic.

The repeatability and reliability of AOSLO cone density measurements has been described in visually normal subjects [5–7] and patients with inherited retinal diseases [8–11]. Cone density measurement validation has been lacking in studies using animal models, with the exception of validating OCT-derived cone density measurements in zebrafish [12]. Density is a common metric for assessing the cone mosaic and will be critical for longitudinal assessment and interpretation of pathology in animal models of cone disorder. Reliable identification of cones is vital for accurate density measurements, which can be challenging with confocal AOSLO due to varying reflectance intensity of cones [13–15]. The advent of non-confocal split-detection has enabled more reliable detection of cones in the human periphery [16], and in the 13-LGS retina [4]. Automated algorithms to identify cones in split-detection AOSLO images have been validated in visually normal subjects [17–19], patients with achromatopsia [20], and patients with Stargardt disease [18, 21], but have yet to be validated in animal models.

The present study evaluates the intraobserver repeatability and interobserver reliability of cone density measurements from split-detection AOSLO images captured in the 13-LGS. We also evaluated the interocular symmetry of cone density using the average of all observer measurements. These data will be useful for evaluating future 13-LGS models and the effect of subsequent experimental treatments.

Methods

Animals

Thirteen 13-LGS (*Ictidomys tridecemlineatus*; 9 female, 4 male) were obtained from the University of Wisconsin-Oshkosh Squirrel Colony for use in this study at the Medical College of Wisconsin. Animal husbandry and dietary protocols were provided according to the Oshkosh Squirrel Colony guidelines [22]. The experimental procedures described were approved by the Institutional Animal Care and Use Committee of the Medical College of Wisconsin and were in accordance with the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research. None of the animals were sacrificed upon completion of this study.

Seasonal set. Five 13-LGS were used for longitudinal assessment of the cone mosaic throughout the hibernation cycle from a previous study [23]. These 13-LGS were included to encompass the range of image quality that results from natural seasonal variability. Imaging for these animals was performed longitudinally at distinct physiological states of the 13-LGS seasonal cycle (pre-hibernation, torpor, and post-hibernation euthermia), between the hours of 10 AM and 3 PM from October to March. Additional information regarding the husbandry and monitoring of these animals imaged throughout hibernation is detailed elsewhere [23]. These 5 squirrels were 5–10 months old.

Euthermic set. Eight 13-LGS were in a euthermic state when imaged to avoid potential seasonal confounds from this species' annual hibernation cycle, which has an effect on non-invasive retinal imaging procedures [23, 24]. Imaging for these animals was performed once, between the hours of 10 AM and 3 PM from June to September. At the time of retinal imaging 3 squirrels were 4 months old, 2 squirrels were 15 months old, and 3 squirrels were 27 months old.

Adaptive optics scanning light ophthalmoscopy (AOSLO)

The 13-LGS photoreceptor mosaic was imaged with non-confocal split-detection AOSLO [16] using a custom instrument optimized for 13-LGS imaging (4.5 mm subject pupil). Animals were imaged under inhaled isoflurane anesthesia (5% induction, 4–5% maintenance in 1L/min O₂ flow; torpid animals did not require induction but were maintained on 2% isoflurane in 0.5L/min O₂ flow). Pupil dilation and cycloplegia were induced with 1% tropicamide and 2.5% phenylephrine, and the eyelids were held open with a pediatric ocular speculum. Saline or artificial tears were used to maintain corneal hydration throughout imaging. The imaging protocol started at the optic nerve head (ONH); the horizontal ONH stretches across ~8mm of the posterior pole of the 13-LGS eye and serves as a landmark that divides superior from inferior retina. Image sequences were captured up to 10 degrees superiorly and 20 degrees inferiorly relative to the ONH. The scale of the AOSLO images was calculated using Ronchi gratings to determine the pixels per degree for the field of view captured during each imaging session.

Analyzing the 13-LGS photoreceptor mosaic

Reference frames were automatically selected from image sequences [25], then 80–150 frames were registered and averaged (S1 Video ([26])). The resulting images were automatically montaged (github.com/BrainardLab/AOAutomontaging; Version 1.5; [27]), with manual correction to the resulting alignment performed in Photoshop CS6 (Adobe, San Jose, CA). Regions of interest (ROIs) were extracted from the montages at 2-degree intervals up to 10 degrees superior and 20 degrees inferior from the ONH using custom software (Translational Imaging Innovations, Inc., Hickory, NC; Fig 1A). The ROIs from the seasonal set were taken from the same 2-degree superior location longitudinally. Each 0.55 x 0.55° ROI was cropped from a

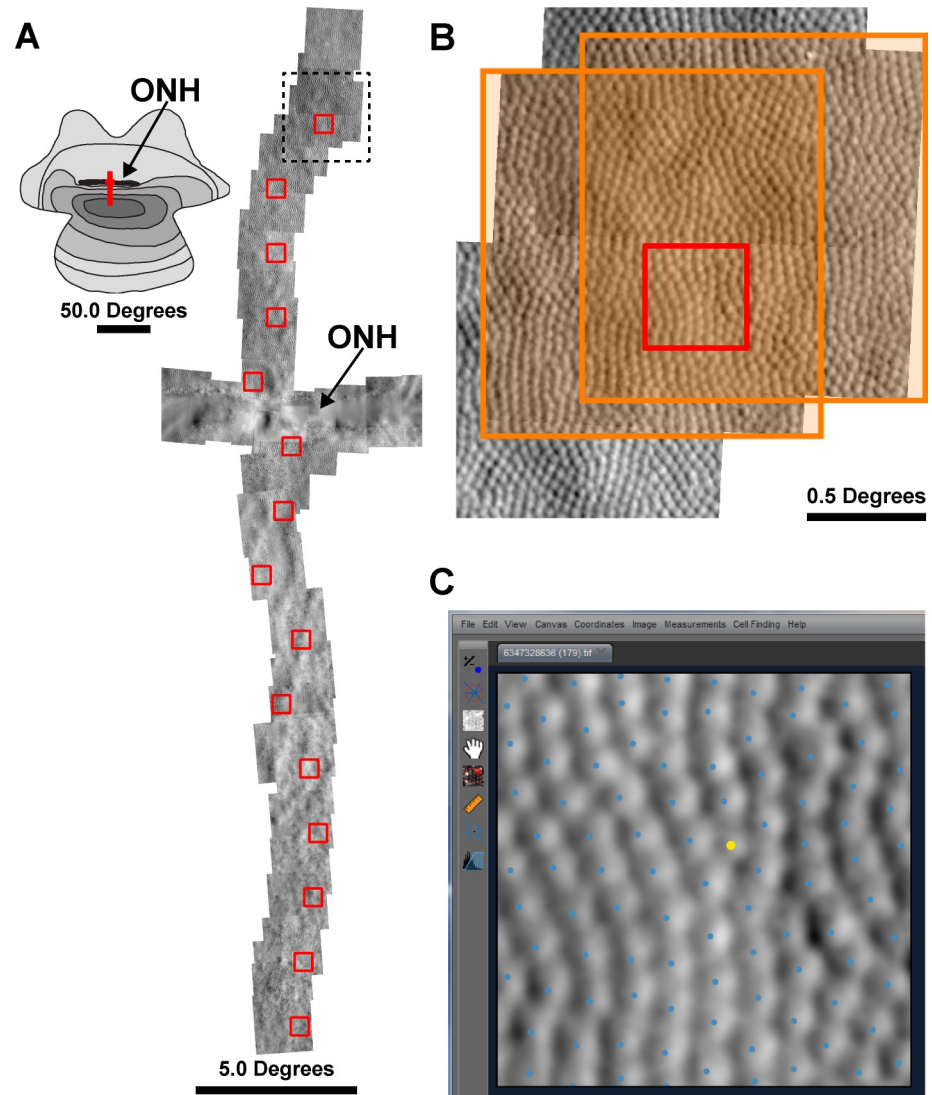


Fig 1. Region of Interest (ROI) and cell identification. (A) Schematic (photoreceptor density map reproduced from Long & Fisher [1983]) showing approximate retinal location (red line) of split-detection AOSLO montage relative to the optic nerve head (ONH). For each montage, ROIs (red boxes) were selected at 2-degree intervals. (B) A zoomed-in section from the montage in (A), outlined by the dashed line. The orange boxes indicate two available images that contain the entire ROI. Once the ROI is cropped from one of the images, the cone segmentation algorithm is applied to the ROI. (C) Example of an ROI (from panel [B]) with cone segmentation results in the observer interface for photoreceptor identification correction. In some cases, rods may be distinguished based on their smaller size relative to cones (rod marked as a yellow dot); rods were included in the analysis.

<https://doi.org/10.1371/journal.pone.0223110.g001>

single image at each retinal location (Fig 1B), and the ROI selector was moved away from blood vessels that obscure the photoreceptor mosaic. The seasonal image set consisted of 26 images (five animals, left eye, 5–6 images per animal), and the euthermic image set consisted of 188 images (eight animals, one or both eyes, 13–30 images per animal). Five observers with varying levels of expertise in analyzing AOSLO photoreceptor mosaics were selected to review and identify 13-LGS photoreceptors. Each observer was introduced to the tools of the cone counting program, which includes a brightness and contrast histogram adjustment, a “flag missing cells” algorithm, and a Voronoi diagram overlay. The 26 images from the seasonal set were presented in a random order and the observer was masked from any information

regarding animal, physiological state, and retinal location. Photoreceptors were then manually identified in these 26 images by each observer three times using the cone counting program interface (Fig 1C). Photoreceptors were automatically detected in all 214 images using an adaptive filtering and local detection algorithm [17]. After automated detection of 13-LGS photoreceptors, the five observers were presented the masked images in random order and identified photoreceptors missed by the program. All 214 images were corrected by each observer three times; once the set of images was corrected to the best of the observer's ability, they moved on to the next copy of the set. No time restrictions were given for this experiment. Infrequent rod photoreceptors have slightly smaller inner segment diameters compared to cones in the 13-LGS [28], but cannot dependably be distinguished from cones in the central retina (Fig 1C), so they were unavoidably included in the analysis. Density was calculated by dividing the number of bound Voronoi cells by their summed area [29].

Image quality analysis

All 214 images were imported into the CellProfiler software [30] and analyzed using the 'Measure Image Quality' module [31]. 'Focus Score' measures pixel intensity variance across the image using a normalized variance algorithm [30, 32, 33]. 'Power Log-Log Slope' measures the slope of the image log-log power spectrum. 'Std Intensity' measures the standard deviation of pixel intensity values. 'MAD Intensity' measures the median absolute deviation of pixel intensity values.

Statistical methods

Fully manual segmentations were performed on the 26 images in the seasonal set three times by each of the five observers to compare to algorithm performance on 13-LGS images ($26 \times 3 \times 5 = 390$). The total image set ($n = 214$ images) was corrected after algorithm segmentation three times by each of the five observers, resulting in 3,210 observations for analysis ($214 \times 3 \times 5 = 3,210$). Within-observer standard deviation (S_w) was used to calculate the repeatability coefficient ($2.77 \cdot S_w$) and measurement error ($1.96 \cdot S_w$) [34]. Repeatability is reported both with the cone density measurements in cells/degree² and as a percentage of the mean cone density for each observer. Photoreceptor density was compared between sexes and ages from the euthermic image set using a Sidak's multiple comparisons test. Interocular symmetry bias, limits of agreement, and 95% confidence intervals were calculated using methods described by Bland and Altman [35–37]. Image quality metric correlations to variance in density measures were tested with Spearman's R Test. Data were tested for normality using the Shapiro-Wilk test. Calculations were completed using Microsoft Excel (2016, Version 1803) and Prism version 8.1.1 (GraphPad, La Jolla, CA).

Results

Fully manual photoreceptor identification compared to semi-automated correction

To evaluate the performance of the segmentation algorithm compared to manual cell selection, we compared the repeatability of fully manual observer photoreceptor identification to the segmentation algorithm output using the seasonal set of 26 images (Table 1). The fully manual photoreceptor identifications were less repeatable overall, with repeatability coefficients ranging from 25.92 to 75.08 cells/degree² (or 8.09–26.14%), compared to the repeatability coefficients ranging from 18.01 to 59.09 cells/degree² (or 4.21–15.72%) after semi-automated photoreceptor identification correction. Observer performance varied between the

Table 1. Repeatability of 13-LGS photoreceptor density measurements in fully manual photoreceptor identification compared to semi-automated correction.

		Observer 1	Observer 2	Observer 3	Observer 4	Observer 5
Manual	Repeatability Coefficient ^{#,*} (95% CI)	25.92 (24.06–27.77)	61.99 (57.56–66.42)	75.08 (69.72–80.45)	45.36 (42.11–48.60)	39.03 (36.24–41.82)
	Measurement Error ^{^,*}	18.34	43.86	53.13	32.09	27.61
Semi-Auto	Repeatability Coefficient ^{#,*} (95% CI)	15.58 (14.46–16.69)	18.01 (16.72–19.29)	28.03 (26.03–30.03)	59.09 (54.87–63.31)	15.94 (14.80–17.08)
	Measurement Error ^{^,*}	11.02	12.74	19.83	41.81	11.28

*Density values in cells/degrees²

[#]The difference between two measurements by that observer for the same image is expected to be less than this for 95% of paired observations.

[^]The difference between a measurement and the true value is expected to be less than this for 95% of observations.

<https://doi.org/10.1371/journal.pone.0223110.t001>

photoreceptor identification techniques, with only observer 4 having better repeatability in the manual photoreceptor identification task.

Repeatability and reliability of 13-LGS photoreceptor density measurements

Intraobserver repeatability of photoreceptor density measurements was excellent for all observers in this study (ICC = 0.997–0.999, [Table 2](#)), suggesting that 99.7% of variance is due to differences in photoreceptor density across different animals, physiological states, and/or retinal locations. However, to assess the magnitude of variance of repeated measurements for this study, we analyzed within-observer standard deviation (S_w). Based on the three readings on 214 images by each observer, the repeatability coefficients (as absolute differences) ranged from 9.43 to 25.71 cells/degree² (or 1.99–5.55%), which provides an estimate of what the differences would be between two photoreceptor density measurements for 95% of occasions ([Table 2](#)). These data suggest that the detectable magnitude of cell loss in these animals would be around 28 cells/degree² depending on the observer. While all observers’ measurements can be considered highly repeatable, observer 1 had the best repeatability, followed by observers 5, 3, 2, and 4 ([Table 2](#)).

We wanted to assess any density-dependent trends since the dataset contained images throughout the vertical meridian, which varies substantially in photoreceptor density (Range of means = 166–885 cells/degree²; [Table 2](#)). Despite the range of photoreceptor densities, there was not a significant trend when comparing the total mean photoreceptor density (derived from 15 observations across all observers) and standard deviation for all 214 images

Table 2. Repeatability and reliability of 13-LGS photoreceptor density measurements.

	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5
Range of Means*	272–885	271–876	166–867	265–881	263–877
Repeatability Coefficient ^{#,*} (95% CI)	9.43 (8.76–10.11)	13.49 (12.52–14.45)	13.23 (12.28–14.17)	25.71 (23.87–27.55)	10.61 (9.85–11.37)
Measurement Error ^{^,*}	6.68	9.54	9.36	18.19	7.51
ICC	0.9996	0.9991	0.9988	0.9967	0.9994
(95% CI)	(0.9994–0.9997)	(0.9989–0.9993)	(0.9985–0.9990)	(0.9959–0.9974)	(0.9993–0.9995)

*Density values in cells/degrees²

[#]The difference between two measurements by that observer for the same image is expected to be less than this for 95% of paired observations.

[^]The difference between a measurement and the true value is expected to be less than this for 95% of observations.

<https://doi.org/10.1371/journal.pone.0223110.t002>

($p = 0.456$, linear regression; **S1 Fig**), suggesting there was no density-dependent bias when analyzing the observers as a group. The standard deviation of each observer's measurements corresponded to the repeatability values in **Table 2**, in that more images had a larger standard deviation in observers with worse repeatability (most notably with observer 4; **S1 Fig**).

Interobserver reliability was assessed by comparing the average values for each image from each of the 5 observers. These values did not pass normality testing for any observer ($p < 0.0001$, Shapiro-Wilk), so the Friedman test was used as a nonparametric alternative to a one-way ANOVA. The observers were significantly different from one another ($p < 0.0001$, Friedman test), and multiple comparison post testing revealed any pair involving observers 3 or 5 were significantly different ($p < 0.0001$, Dunn's multiple comparisons test). There was no obvious difference in the training or level of experience of observers 3 and 5.

For each image, we examined the addition and removal of cones by each observer following automated processing by the automated algorithm. The mean \pm SD number of cells added was 5.15 ± 2.80 across all 3,210 measurements and the mean \pm SD number of cells removed was 4.56 ± 3.61 . On average, observers shifted the location of the cell center estimated by the algorithm for 5% - 22% of cells. For several images, the number of cells added was similar to the number of cells removed (**Fig 2A and 2C**), offsetting any major effect on cone density. In high quality images (e.g. **Fig 2A**), cells added or removed by the observer were mostly around the edges of the image, and even then, such differences are unlikely to affect the computed density as we are using bound Voronoi cells to compute density (**Fig 2C**). In low quality images (e.g. **Fig 2B**), automated cell identification was unable to find the necessary contrast features of the image and more subjective and variable observer cell addition/removal was needed (**Fig 2D**).

Distribution of image quality

In order to objectively assess the distribution of image quality independent of cell identification, we measured image qualities through CellProfiler, which measures a suite of unique image features. **S2 Fig** shows the distribution of image quality for all 214 images for these 4 image quality metrics. Image quality was variable between blur- and pixel intensity-based metrics. There was a negative association between variance (from all 15 observations for all 214 images) and 'Focus Score', 'Power Log-Log Slope', 'Std Intensity', and 'MAD Intensity' ($r_s = -0.319, -0.320, -0.349, \text{ and } -0.3879$, respectively; $p < 0.0001$; Spearman Correlation). This demonstrates that the agreement between observers was worse for images of lower image quality.

Interocular symmetry

Interocular symmetry of density measurements from AOSLO images of the 13-LGS photoreceptor mosaics was assessed using the average values of all observers. Vertical strip montages (i.e. **Fig 1A**) from both eyes were available for five of eight animals. **Fig 3A** shows the mean differences between eyes, and the bias was close to zero ($2.74 \text{ cells/degree}^2$, 95% CI = 9.602 to $-4.122 \text{ cells/degree}^2$), which is consistent with interocular symmetry. However, as there was a significant relationship between the density difference for the pairs of eyes and the average density for those pairs of eyes ($r = 0.42$, $p < 0.0003$, Pearson correlation), we plotted the ratio of the right eye (OD) to the left eye (OS) values on the y-axis instead of the OD-OS difference (**Fig 3B**). This is effectively the same as doing a log transformation of the individual values and subtracting them [36]. In this case, the 95% limits of agreement are 0.88 to 1.12 (*dashed lines*, **Fig 3B**). This means that we would expect the OD and OS values to differ by less than 12% for 95% of animals examined. **Fig 3C** shows a consistent density gradient (in degrees^2) relative to the optic nerve across these five animals. Similar to a previous study [4], the lowest photoreceptor densities ($270\text{--}409 \text{ cells/degree}^2$) were found superior to the ONH, and peak

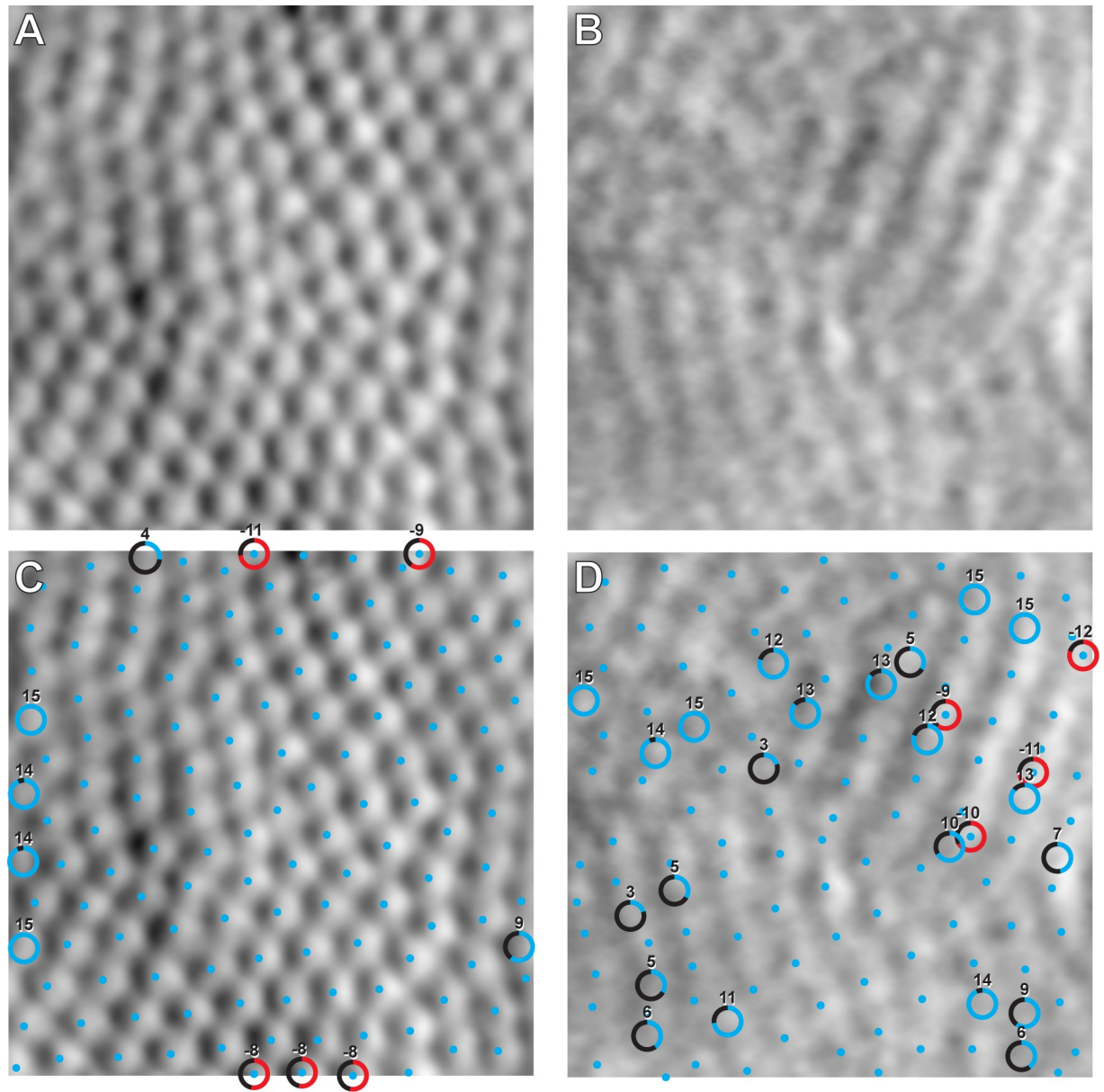


Fig 2. Extremes of interobserver agreement. (A) The image with the highest agreement (lowest variance) in photoreceptor identification. (B) The image with the lowest agreement (highest variance) in photoreceptor identification from the euthermic image set. (C) and (D) show the results of the cone segmentation algorithm (blue dots) with observer corrections. Circles indicate photoreceptors that were added (blue gradient) or removed (red gradient) by more than one observer, and the number outside the circle denotes how many times the photoreceptor was added or removed (15 being the maximum). In panel (D), three out of four commonly removed cells overlapped with added cells, suggesting that these were cells simply shifted to a new location rather than truly being added or removed.

<https://doi.org/10.1371/journal.pone.0223110.g002>

photoreceptor densities in the 13-LGS visual streak (656–881 cells/degree²) are found 11 degrees (~1.1mm) inferior from the ONH. Thus, while the range of densities between animals varies, the gradient of densities superior and inferior from the ONH are consistent across animals (Fig 3). There was not a significant effect of sex on photoreceptor density at any of the 15 regions analyzed ($p > 0.05$, 2-way ANOVA, Sidak's multiple comparisons test, $n = 3$ males, 5 females). There was a significant effect of age on photoreceptor density only in the comparison of 4-month old 13-LGS compared to 15-month old 13-LGS, and only around the visual streak

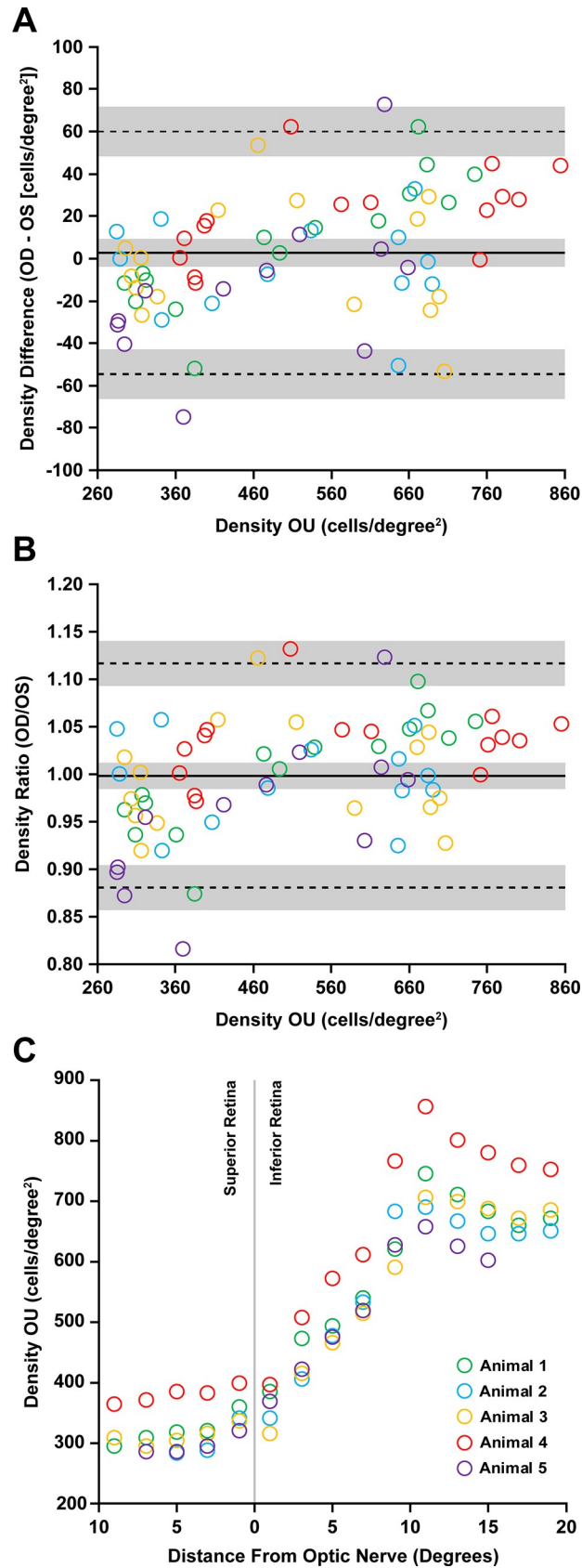


Fig 3. Interocular comparisons of photoreceptor density. (A-B) Bland-Altman plots showing photoreceptor density has interocular symmetry between left (OS) and right (OD) eyes of five 13-LGS (Animals 1–5; legend in (C)) used in this study. Density measurements from all observers were averaged for this comparison. (A) The mean difference between eyes is 2.74 cells/degree² with 95% limits of agreement 60.15 and -54.57 cells/degree². Pearson correlation coefficient (r) calculations reveal a significant proportional bias to the magnitude of density ($r = 0.419$, 95% CI = 0.20 to 0.60, $P = 0.0003$). (B) The data were then transformed to highlight the relative interocular symmetry unrelated to mean density differences. The mean ratio of density measurements was 0.99 with 95% limits of agreement 1.12 and 0.88. Thus, the OD and OS values differ by less than 12% for 95% of animals examined. *Solid lines* represent the average difference (A) or ratio (B) between eyes, while *dotted lines* represent 95% limits of agreement. The *gray shading* represents the 95% confidence intervals. (C) Photoreceptor density at distances from the ONH. Left and right eyes were averaged for each data point. Density varies between animals, but the gradient is consistent in these five animals—with low photoreceptor density in the superior retina, and peak density (at the visual streak) 11 degrees inferior to the ONH.

<https://doi.org/10.1371/journal.pone.0223110.g003>

(9 degrees inferior, $p = 0.021$; 11 degrees inferior, $p = 0.001$; 13 degrees inferior, $p = 0.020$; 2-way ANOVA, Sidak's multiple comparisons test, $n = 3$ at the age of 4 months, 2 at the age of 15 months, 3 at the age of 27 months).

Discussion

The ability to reliably assess the cone-dominant photoreceptor mosaic of the 13-LGS *in vivo* is important as models of cone disorder are developed in this species. Confocal AOSLO is limited in the detection of single 13-LGS photoreceptors because of their variable multimodal reflective appearance [4], which is also seen in perifoveal human cones [38]). However, split-detection AOSLO can more clearly resolve all photoreceptors imaged in the 13-LGS compared to confocal AOSLO [4]), and limited eye movement of anesthetized animals helps register high quality images for analysis (S1 Video). Automated detection of photoreceptors (via the adaptive filtering and local detection algorithm [17]) works well on 13-LGS split-detection images of high quality (Fig 2). However, manual correction by observers is still needed to add missed photoreceptors or remove erroneously selected photoreceptors. Defining repeatability, reliability, and interocular symmetry of 13-LGS photoreceptor density measurements is critical to effectively assess any photoreceptor phenotypes in future transgenic models in this species, or experimental models that use one eye as a control.

As expected, interocular symmetry was observed with photoreceptor density measurements, with mean difference between eyes being approximately zero (Fig 3A) and the ratio of both eyes being approximately 1:1 (Fig 3B). The Bland-Altman analysis indicates that a density difference between eyes of 12% or more would be considered significantly different. The magnitude of density measurements showed proportional bias (Fig 3A), which highlights a methodological limitation that interocular measurements were not collected at the same position along the horizontal meridian. These results suggest that cell density may not be as uniform across the ground squirrel visual streak as we assumed in this study (based on California ground squirrel topography maps [39, 40]). Precise topographical distribution of photoreceptors is unknown in the 13-LGS. Interocular symmetry of axial distance measurements between hyper-reflective bands that correspond to the photoreceptors and RPE in this species was seen with OCT [24]), suggesting that an experimental paradigm involving a contralateral control eye to assess photoreceptor changes in the 13-LGS is a valid approach. Photoreceptor density in this study ranged from 166–881 cells/degree², which corresponds to 16,562–87,902 cells/mm² if a retinal magnification factor of 100 $\mu\text{m}/\text{degree}$ is used (based on the 7.90 mm axial length of the European GS [41])). This density range is similar to the California ground squirrel [39, 40] and our initial 13-LGS AOSLO report [4]). A limitation of the present study is that axial length was not measured in these animals. As axial length is known to affect lateral scale of *in vivo* retinal images [42, 43], we reported cell density in cells/degree². Human AOSLO

images are often scaled linearly by the subject's axial length using an assumed 291 $\mu\text{m}/\text{degree}$ retinal magnification factor derived from the 24 mm Gullstrand model eye [44]. An established model eye does not exist for the 13-LGS but should be derived to reduce some assumptions in scaling of non-invasive retinal images. Once lateral scaling can be more accurately approximated in individual 13-LGS, it will be possible to derive a more complete understanding of the range of photoreceptor densities in 13-LGS as a function of age and sex. For example, while no differences in cone density were detected between sexes in our study, the extent to which axial length and photoreceptor density varies between animals and sexes is unknown. As such, we cannot rule out the possibility of sex-based differences in areal cone density measurements. It is possible that observed effects of age on photoreceptors are due to eye growth rather than a real change in photoreceptor density. In addition, similar comparisons will be needed to examine possible differences between wild-caught and captive-bred animals.

Intraobserver repeatability ranged from 9.43 to 25.71 cells/degree² (1.99–5.55%) in this study (Table 2), and these data suggest that the difference between two density measurements from 13-LGS images captured from our AOSLO should be less than 28 cells/degree² for 95% of observations (Table 2). While we used images the span the range of typical 13-LGS image quality, increased image quality results from the stability of anesthetized 13-LGS subjects compared to actively fixating human volunteers with varying levels of fixational stability. However, this study used only non-confocal split-detection AOSLO images, whereas many of the repeatability studies of human cone density used confocal AOSLO in which the appearance of individual cones can be more variable [5, 6, 11]. Cone density measurements using confocal AOSLO images captured from a population of healthy individuals using a similar semi-automated method were similar in intraobserver repeatability (2.7%) [5] and interobserver reliability (ICC of 0.957) [6] to this study.

Our assessment of interobserver reliability suggests that not all observers perform alike. Observers 3 and 5 were significantly different in their measurements than the observers 1, 2, and 4, but not different from each other. Since observers 3 and 5 had repeatable measurements (Table 2), this result suggests that these observers corrected cell identification outputs differently than the other 3 observers. Previous work has showed poor interobserver reliability related to experience in cell identification [9]. Additionally, observer experience was not considered in the data analysis but may be a source of variance in these results. For example, observer 1 was the most experienced 13-LGS photoreceptor identifier, whereas observer 2 was the least experienced in identifying 13-LGS photoreceptors. The number of years in an imaging lab performing these tasks may be related to the repeatability and reliability performance in this study; Observer 1 was the most senior (5 years), followed by observer 5 (3 years), observer 2 (2 years), observer 3 (1 year), and observer 4 (<1 year). This level of experience is nearly identical to the performance in repeatability (Table 2). Our results highlight the importance of observer training and accurate automation for AOSLO cell identification.

Interobserver agreement between cone density measurements was worse in lower quality images (Fig 2). When quantitative assessments of image quality were performed in this study, the images one might qualitatively determine as high or low quality often matched the scores in the CellProfiler metrics (S2 Fig). Since this method scores images based on blur and pixel intensity metrics, it may not be appropriate for all images of damaged or disease photoreceptors, where the images can have high contrast but photoreceptor mosaic structure is changed or unrecognizable. This is particularly relevant to 13-LGS retinal imaging where some photoreceptor images from torpid animals have high contrast and low blur but are unrecognizable as photoreceptors [45]. Therefore, when counting photoreceptors, computational methods involving training datasets and machine learning should be developed for the appropriate animal model or retinal disorder to accurately automate cell identification. This approach has

been recently applied to finding cones in subjects with achromatopsia [20] and Stargardt disease [21]. Image quality variability with *in vivo* imaging is common if the tear film of the subject is not meticulously maintained for wavefront correction using adaptive optics. Methods to reject blurry image collection may be warranted to increase repeatability of density and other photoreceptor mosaic metrics. In addition, future studies may determine a threshold image quality above which reliable cone density estimates can be obtained, which could significantly enhance longitudinal studies in these animals.

In conclusion, we have characterized the interocular symmetry (Fig 3), intraobserver repeatability (Table 2, S1 Fig), and interobserver reliability of 13-LGS photoreceptor density measurements from images collected with our AOSLO system using the adaptive filtering and local detection algorithm [17]. This algorithm performed well on 13-LGS split-detection AOSLO images and led to excellent repeatability of observers to correct the cell identification output of the algorithm. Our analysis was limited to 30 degrees in the posterior pole, which includes the cone-rich visual streak. True peak cone density could be found elsewhere along the visual streak horizontal meridian, but we have found peak cone density to be consistently located 11 degrees inferior from the ONH (~1.1 mm). While this is a promising start for assessing repeatability and reliability of photoreceptor metrics in animal models, our work is limited to one measurement metric from images collected by one AOSLO system in one animal species. Validation of this method in 13-LGS will come with additional studies using custom or commercial AO systems in additional lab environments working with additional species. Overall, this method could be used to monitor longitudinal changes in photoreceptor density in the 13-LGS, though attention to image quality will be required for reliable data collection.

Supporting information

S1 Dataset. Observer measurements.

(XLSX)

S1 Fig. Standard deviation of observer measurements. Mean and standard deviations from all 15 measurements from all five observers (top left panel), and mean and standard deviations from the three measurements from each observer (remaining five panels). Despite the range of photoreceptor densities, there was not a significant trend when comparing the total 15-observation (from all observers) mean photoreceptor density and standard deviation for all 214 images ($p = 0.456$, linear regression).

(PDF)

S2 Fig. Distribution of image quality. Results from CellProfiler analysis of image quality from all 214 images, and the images with Min, Median, and Max score for each metric. (A) 'Focus Score' measures pixel intensity variance across the image using a normalized variance algorithm. (B) 'Power Log-Log Slope' measures the slope of the image log-log power spectrum. (C) 'Std Intensity' measures the standard deviation of pixel intensity values. (D) 'MAD Intensity' measures the median absolute deviation of pixel intensity values.

(PDF)

S1 Video. Raw and registered video of 13-LGS photoreceptors.

(AVI)

Acknowledgments

The authors thank Christine Skumatz, Lisa King, and the Biomedical Resource Center at MCW for their contributions to 13-LGS care, Dana Merriman for supplying the 13-LGS used

for this study, Robert Cooper for developing and supporting the cone counting software, David Cunefare and Sina Farsiu for developing and supporting the adaptive filtering and local detection software, Alfredo Dubra for developing and supporting the AOSLO and image registration software used in this work, and Sergey Tarima and Alexis Visotcky for the helpful discussions related to statistical analysis.

Author Contributions

Conceptualization: Benjamin S. Sajdak, Joseph Carroll.

Data curation: Benjamin S. Sajdak, Alexander E. Salmon, Rachel E. Linderman, Jenna A. Cava, Heather Heitkotter, Joseph Carroll.

Formal analysis: Benjamin S. Sajdak, Alexander E. Salmon, Rachel E. Linderman, Jenna A. Cava, Heather Heitkotter, Joseph Carroll.

Investigation: Benjamin S. Sajdak.

Methodology: Benjamin S. Sajdak, Joseph Carroll.

Project administration: Benjamin S. Sajdak.

Resources: Benjamin S. Sajdak, Joseph Carroll.

Software: Joseph Carroll.

Supervision: Joseph Carroll.

Validation: Benjamin S. Sajdak, Joseph Carroll.

Visualization: Benjamin S. Sajdak, Alexander E. Salmon, Joseph Carroll.

Writing – original draft: Benjamin S. Sajdak.

Writing – review & editing: Benjamin S. Sajdak, Alexander E. Salmon, Rachel E. Linderman, Jenna A. Cava, Heather Heitkotter, Joseph Carroll.

References

1. Slijkerman RW, Song F, Astuti GD, Huynen MA, van Wijk E, Stieger K, et al. The pros and cons of vertebrate animal models for functional and therapeutic research on inherited retinal dystrophies. *Progress in Retinal and Eye Research*. 2015; 48:137–59. <https://doi.org/10.1016/j.preteyeres.2015.04.004> PMID: 25936606
2. Walters S, Schwarz C, Sharma R, Rossi EA, Fischer WS, DiLoreto DA, et al. Cellular-scale evaluation of induced photoreceptor degeneration in the living primate eye. *Biomedical Optics Express*. 2019; 10(1):66–82. <https://doi.org/10.1364/BOE.10.000066> PMID: 30775083
3. Merriman DK, Sajdak BS, Li W, Jones BW. Seasonal and post-trauma remodeling in cone-dominant ground squirrel retina. *Experimental Eye Research*. 2016; 150:90–105. <https://doi.org/10.1016/j.exer.2016.01.011> PMID: 26808487
4. Sajdak B, Sulai YN, Langlo CS, Luna G, Fisher SK, Merriman DK, et al. Noninvasive imaging of the thirteen-lined ground squirrel photoreceptor mosaic. *Visual Neuroscience*. 2016; 33:e003. <https://doi.org/10.1017/S0952523815000346> PMID: 26923645
5. Garrioch R, Langlo C, Dubis AM, Cooper RF, Dubra A, Carroll J. Repeatability of in vivo parafoveal cone density and spacing measurements. *Optometry and Vision Science*. 2012; 89(5):632–43. <https://doi.org/10.1097/OPX.0b013e3182540562> PMID: 22504330
6. Liu BS, Tarima S, Visotcky A, Pechauer A, Cooper RF, Landsem L, et al. The reliability of parafoveal cone density measurements. *British Journal of Ophthalmology*. 2014; 98(8):1126–31. <https://doi.org/10.1136/bjophthalmol-2013-304823> PMID: 24855115
7. Morgan JIW, Vergilio GK, Hsu J, Dubra A, Cooper RF. The reliability of cone density measurements in the presence of rods. *Translational Vision Science & Technology*. 2018; 7(3):21.

8. Langlo CS, Erker LR, Parker M, Patterson EJ, Higgins BP, Summerfelt P, et al. Repeatability and longitudinal assessment of foveal cone structure in *CNGB3*-associated achromatopsia. *Retina*. 2017;1956–66. <https://doi.org/10.1097/IAE.0000000000001434> PMID: 28145975
9. Abozaid MA, Langlo CS, Dubis AM, Michaelides M, Tarima S, Carroll J. Reliability and repeatability of cone density measurements in patients with congenital achromatopsia. *Advances in Experimental Medicine and Biology*. 2016; 854:277–83. https://doi.org/10.1007/978-3-319-17121-0_37 PMID: 26427422
10. Tanna P, Kasilian M, Strauss R, Tee J, Kalitzeos A, Tarima S, et al. Reliability and repeatability of cone density measurements in patients with Stargardt disease and RPGR-associated retinopathy. *Investigative Ophthalmology & Visual Science*. 2017; 58(9):3608–15.
11. Zayit-Soudry S, Sippl-Swezey N, Porco T, Lynch SK, Syed R, Ratnam K, et al. Repeatability of cone spacing measures in eye with inherited retinal degenerations. *Investigative Ophthalmology & Visual Science*. 2015; 56(10):6179–89.
12. Huckenpahler A, Wilk M, Link B, Carroll J, Collery R. Repeatability and reproducibility of in vivo cone density measurements in the adult zebrafish retina. *Advances in Experimental Medicine and Biology*. 2018; 1074:151–6. https://doi.org/10.1007/978-3-319-75402-4_19 PMID: 29721939
13. Grieve K, Roorda A. Intrinsic signals from human cone photoreceptors. *Investigative Ophthalmology & Visual Science*. 2008; 49(2):713–9.
14. Jonnal R, Kocaoglu OP, Wang Q, Lee S, Miller DT. Phase-sensitive imaging of the outer retina using optical coherence tomography and adaptive optics. *Biomedical Optics Express*. 2012; 3(1):104–24. <https://doi.org/10.1364/BOE.3.000104> PMID: 22254172
15. Meadway A, Sincich LC. Light propagation and capture in cone photoreceptors. *Biomedical Optics Express*. 2018; 9(11):5543–65. <https://doi.org/10.1364/BOE.9.005543> PMID: 30460146
16. Scoles D, Sulai YN, Langlo CS, Fishman GA, Curcio CA, Carroll J, et al. In vivo imaging of human cone photoreceptor inner segments. *Investigative Ophthalmology & Visual Science*. 2014; 55(7):4244–51.
17. Cunefare D, Cooper RF, Higgins B, Katz DF, Dubra A, Carroll J, et al. Automatic detection of cone photoreceptors in split detector adaptive optics scanning light ophthalmoscope images. *Biomedical Optics Express*. 2016; 7(5):2036–50. <https://doi.org/10.1364/BOE.7.002036> PMID: 27231641
18. Bergeles C, Dubis AM, Davidson B, Kasilian M, Kalitzeos A, Carroll J, et al. Unsupervised identification of cone photoreceptors in non-confocal adaptive optics scanning light ophthalmoscope images. *Biomedical Optics Express*. 2017; 8(6):3081–94. <https://doi.org/10.1364/BOE.8.003081> PMID: 28663928
19. Liu J, Jung H, Dubra A, Tam J. Automated photoreceptor cell identification on nonconfocal adaptive optics images using multiscale circular voting. *Investigative Ophthalmology & Visual Science*. 2017; 58(11):4477–89.
20. Cunefare D, Langlo CS, Patterson EJ, Blau S, Dubra A, Carroll J, et al. Deep learning based detection of cone photoreceptors with multimodal adaptive optics scanning light ophthalmoscope images of achromatopsia. *Biomedical Optics Express*. 2018; 9(8):3740–56. <https://doi.org/10.1364/BOE.9.003740> PMID: 30338152
21. Davidson B, Kalitzeos A, Carroll J, Dubra A, Ourselin S, Michaelides M, et al. Automatic cone photoreceptor localisation in healthy and Stargardt afflicted retinas using deep learning. *Scientific Reports*. 2018; 8(1):7911. <https://doi.org/10.1038/s41598-018-26350-3> PMID: 29784939
22. Merriman DK, Lahvis G, Jooss M, Gesicki JA, Schill K. Current practices in a captive breeding colony of 13-lined ground squirrels (*Ictidomys tridecemlineatus*). *Lab Animal*. 2012; 41(11):315–25. <https://doi.org/10.1038/labon.150> PMID: 23079915
23. Sajdak BS, Salmon AE, Litts KM, Wells C, Allen KP, Dubra A, et al. Evaluating seasonal changes of cone photoreceptor structure in the 13-lined ground squirrel. *Vision Research*. 2019; 158:90–9. <https://doi.org/10.1016/j.visres.2019.02.009> PMID: 30826354
24. Sajdak BS, Bell BA, Lewis TR, Luna G, Cornwell GS, Fisher SK, et al. Assessment of outer retinal remodeling in the hibernating 13-lined ground squirrel. *Investigative Ophthalmology & Visual Science*. 2018; 59(6):2538–47.
25. Salmon AE, Cooper RF, Langlo CS, Baghaie A, Dubra A, Carroll J. An automated reference frame selection (ARFS) algorithm for cone imaging with adaptive optics scanning light ophthalmoscopy. *Translational Vision Science & Technology*. 2017; 6(2):9.
26. Dubra A, Harvey Z. Registration of 2D images from fast scanning ophthalmic instruments. In: Fischer B, Dawant B, Lorenz C, editors. *Biomedical Image Registration*. Lecture Notes in Computer Science. 1 ed. Berlin: Springer-Verlag; 2010. p. 60–71.
27. Chen M, Cooper RF, Han GK, Gee J, Brainard DH, Morgan JI. Multi-modal automatic montaging of adaptive optics retinal images. *Biomedical Optics Express*. 2016; 7(12):4899–918. <https://doi.org/10.1364/BOE.7.004899> PMID: 28018714

28. von Schantz M, Szél A, van Veen T, Farber DB. Expression of phototransduction cascade genes in the ground squirrel retina. *Investigative Ophthalmology & Visual Science*. 1994; 35(5):2558–66.
29. Cooper RF, Wilk MA, Tarima S, Carroll J. Evaluating descriptive metrics of the human cone mosaic. *Investigative Ophthalmology & Visual Science*. 2016; 57(7):2992–3001.
30. Lamprecht MR, Sabatini DM, Carpenter AE. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques*. 2007; 42(1):71–5. <https://doi.org/10.2144/000112257> PMID: 17269487
31. Bray MA, Fraser AN, Hasaka TP, Carpenter AE. Workflow and metrics for image quality control in large-scale high-content screens. *Journal of Biomolecular Screening*. 2012; 17(2):266–74. <https://doi.org/10.1177/1087057111420292> PMID: 21956170
32. Groen FC, Young IT, Lighthart G. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*. 1985; 6(2):81–91. <https://doi.org/10.1002/cyto.990060202> PMID: 3979220
33. Yeo TTE, Ong SH, Jayasooriah T, Sinniah R. Autofocusing for tissue microscopy. *Image and Vision Computing*. 1993; 11(10):629–39.
34. Bland JM, Altman DG. Statistics notes: Measurement error proportional to the mean. *British Medical Journal*. 1996; 313(7049):106. <https://doi.org/10.1136/bmj.313.7049.106> PMID: 8688716
35. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound in Obstetrics & Gynecology*. 2003; 22:85–93.
36. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*. 1999; 8(2):135–60. <https://doi.org/10.1177/096228029900800204> PMID: 10501650
37. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–10. PMID: 2868172
38. Dubra A, Sulai Y, Norris JL, Cooper RF, Dubis AM, Williams DR, et al. Noninvasive imaging of the human rod photoreceptor mosaic using a confocal adaptive optics scanning ophthalmoscope. *Biomedical Optics Express*. 2011; 2(7):1864–76. <https://doi.org/10.1364/BOE.2.001864> PMID: 21750765
39. Kryger Z, Galli-Resta L, Jacobs GH, Reese BE. The topography of rod and cone photoreceptors in the retina of the ground squirrel. *Visual Neuroscience*. 1998; 15(4):685–91. <https://doi.org/10.1017/s0952523898154081> PMID: 9682870
40. Long KO, Fisher SK. The distributions of photoreceptors and ganglion cells in the California ground squirrel, *Spermophilus beecheyi*. *Journal of Comparative Neurology*. 1983; 221(3):329–40. <https://doi.org/10.1002/cne.902210308> PMID: 6655087
41. Hughes A. The topography of vision in animals with contrasting life styles. In: Crescitelli F, editor. *Handbook of Sensory Physiology*. VII/5. Berlin: Springer; 1977. p. 614–42.
42. Izatt JA, Choma MA. Theory of optical coherence tomography. In: Drexler W, Fujimoto JG, editors. *Optical Coherence Tomography: Technology and Applications*. 1 ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 47–72.
43. Linderman R, Salmon AE, Strampe M, Russillo M, Khan J, Carroll J. Assessing the accuracy of foveal avascular zone measurements using optical coherence tomography angiography: segmentation and scaling. *Translational Vision Science & Technology*. 2017; 6(3):16.
44. Hirsch J, Curcio CA. The spatial resolution capacity of human foveal retina. *Vision Research*. 1989; 29(9):1095–102. [https://doi.org/10.1016/0042-6989\(89\)90058-8](https://doi.org/10.1016/0042-6989(89)90058-8) PMID: 2617858
45. Sajdak BS, Salmon AE, Litts KM, Wells C, Allen KP, Dubra A, et al. Evaluating seasonal changes of cone photoreceptor structure in the 13-lined ground squirrel. *Vision Research*. 2019; 158:90–9. <https://doi.org/10.1016/j.visres.2019.02.009> PMID: 30826354