

Artificial Intelligence in Biology and Bioinformatics

Quantification of variegated *Drosophila* ommatidia with high-resolution image analysis and machine learning

Hunter J. Hill ^{1,*}, William Sullivan ² and Brandon S. Cooper¹

¹Division of Biological Sciences, University of Montana, Missoula, MT 59812, United States

²Department of Molecular, Cell, and Developmental Biology, University of California Santa Cruz, Santa Cruz, CA 95064, United States

*Corresponding author. Division of Biological Sciences, University of Montana, Missoula, MT 59812, United States. Email: hunter.hill@umt.edu

Abstract

A longstanding challenge in biology is accurately analyzing images acquired using microscopy. Recently, machine learning (ML) approaches have facilitated detailed quantification of images that were refractile to traditional computation methods. Here, we detail a method for measuring pigments in the complex-mosaic adult *Drosophila* eye using high-resolution photographs and the pixel classifier *ilastik* [1]. We compare our results to analyses focused on pigment biochemistry and subjective interpretation, demonstrating general overlap, while highlighting the inverse relationship between accuracy and high-throughput capability of each approach. Notably, no coding experience is necessary for image analysis and pigment quantification. When considering time, resolution, and accuracy, our view is that ML-based image analysis is the preferred method.

Keywords: quantification; variegating; *drosophila*; ommatidia; machine learning; heterochromatin

Introduction

In *Drosophila melanogaster*, the *white* gene resides at the distal end of the X chromosome in euchromatin and encodes an ABC-type guanine transporter responsible for the import of drospterines and ommochromes that pigment the eyes [2] (Fig. 1A). A mutation in *white* results in the white eye phenotype due to lack of pigment import to developing ommatidia [3] (Fig. 1B). The first example of eye variegation in *Drosophila* was recovered by Hermann Muller during an X-ray mutagenesis screen in 1930 [4]. He recovered an inversion in which the euchromatic *white*⁺ was juxtaposed with pericentric heterochromatin resulting in Position Effect Variegation (PEV, Fig. 1C) [5]. An abnormal boundary between heterochromatin and euchromatin generally exhibits PEV, whereby heterochromatin spreads into the nearby gene and deactivates expression in some cells but not others. PEV is best demonstrated with cell-autonomous phenotypes that can be viewed with large cell numbers in essentially two dimensions—the *Drosophila* eye is a textbook example with ideal features for studying PEV [6–9] (Fig. 1D).

In addition to chromosome rearrangements, transposon movement into heterochromatin may produce mosaic gene expression because any genes within the transposon are subject to PEV [10, 11]. The observed patterns of mosaic expression are unique depending on the heterochromatin causing the PEV [12]. Generally, Y and pericentric heterochromatin cause patched (or sectorized/clonal) inactivation while telomeric and chromosome 4 heterochromatin cause salt-and-pepper variegation. (The latter two types are also resistant to enhancer and suppressor modifiers.) These patterns may reflect the developmental timing of

inactivation and maintenance in different heterochromatin types [13, 14], but the topic remains an active area of research [15].

Mutant alleles which either enhance or suppress the variegation phenotype have generated considerable insight into how chromatin states are defined and maintained. By 1969, two mutations had been discovered which altered the PEV phenotype [16, 17]. Since then, ~500 dominant second-site mutations have been identified in ~150 genes by genetic screening, and ~30 genes have been studied in detail [18] (reviewed in the literature [6–8, 12]). Mutations changing the eye from variegating to white are enhancers of variegation, *E(var)*, and contribute to euchromatin formation and active gene expression (i.e. when an *E(var)* is dysfunctional, heterochromatin predominates and *white*⁺ is inactivated). Alternatively, mutants turning the eye from variegating to red are suppressors of variegation and normally contribute to heterochromatin formation (e.g. *HP1 = Su(var)2-5*) [19].

When paired with a suppressor or enhancer mutation, the shift in eye color is often dramatic, nearly complete, and there is little need for accurate quantification [18] (Fig. 1 D vs E&F). However, some modifiers have more subtle effects on eye color and require higher-resolution photos and improved accuracy for quantification methods. Furthermore, even within a PEV expressing stock isolate eyes can be highly variable, and a broad phenotypic range of eye color may occur in a single vial (Fig. 2). This highlights the critical need for improved techniques related to data generation, analysis, and interpretation. Researchers have recently developed computational methods to quantify high-resolution images of fly eyes [20, 21], focusing on the ommatidia organization that results from a tightly regulated developmental process. However, these methods do not measure *Drosophila*

Received: 5 November 2024; Revised: 24 December 2024; Editorial decision: 27 December 2024; Accepted: 7 January 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

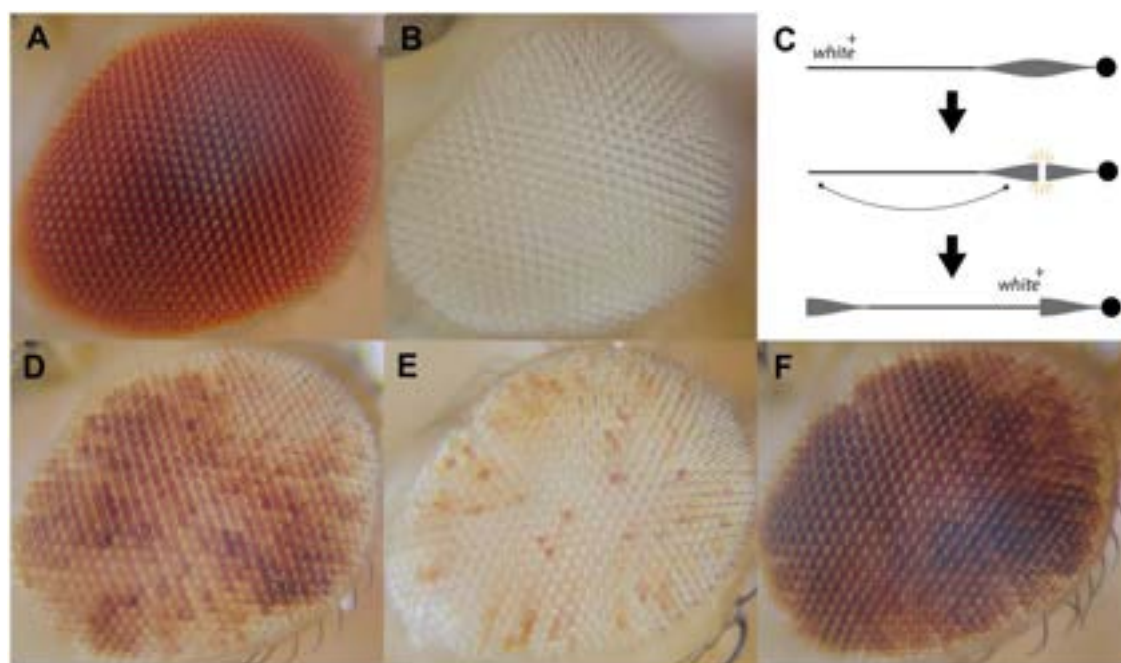


Figure 1. **A:** Wildtype ($white^+$) eye from $C(1)DX$ stock; **B:** white mutant from $white^{1118}$ stock; **C:** illustration of X chromosome rearrangement forming $In(1)w^{M4}$ chromosome, with the centromere (black circle), heterochromatin (thicker grey bar), and euchromatin (thin black line). A break in heterochromatin occurs, followed by an inversion bringing the euchromatic tip near pericentric heterochromatin; **D:** PEV eye from $In(1)w^{M4}$ stock; **E:** enhancement of variegation; and **F:** suppression of variegation.



Figure 2. Phenotypic variation within a single vial of the $In(1)w^{M4}$ stock. Eyes are from 16 isogenic male siblings.

ocular pigment, which is required for PEV screens and other analyses (e.g. FLP-induced mitotic recombination screens [22, 23] and RNAi controls [24]).

There are three methods currently used to quantify ocular pigment (or proxies) in *Drosophila*: RT-qPCR for the *white* gene [25], image grayscale conversion [26], and pigment extraction [26, 27]. While RT-qPCR quantification of *white* transcription does not directly assay the ocular pigment phenotype, it enables absolute quantification and the ability to assay timepoints pre-pupation when pigment is first imported. Challenges associated with this method include complex experimental designs, the potential for relatively low reproducibility because each amplification cycle contributes to the data, and the inability to assess single cells. Grayscale conversion and measurement in imageJ [28] enable relatively fast and accurate quantification. However, this process typically involves low-resolution images for 8-bit conversion, providing less information and introducing error due to lighting shifts and shadows. Finally, quantification of chemically

extracted pigments using a spectrophotometer has been widely used [2, 25–27], including by us here. Extraction of ocular pigments in acidified ethanol typically requires a pool of 5–25 (or more) *Drosophila* heads per absorbance reading, although this could be plausibly optimized for single flies or even individual eyes in the future. Pigments can be easily extracted and quantified, making this a useful approach for over 80 years [27].

To our knowledge, researchers have not developed computational software to specifically quantify *Drosophila* eye pigment from images. Here, we present a guide for using the freely available ML software “ilastik” in conjunction with Fiji (imageJ) to quantify eye pigments, which provides a straightforward and easily accessible approach for quantification. We utilize pixel classification wherein the ilastik user first imports images to the software for training. Training consists of the user identifying classes (or labels), followed by labeling using a mouse and cursor. This applies label-specific brush strokes directly to the image, with no coding experience required. ilastik classifies pixels using

an output of image filters (pixel color, intensity, edge-ness, and texture), and a Random Forest with 100 trees as a classifier [1]. A Random Forest [29] is preferred over non-linear classifiers since it has minimal parameters and is more robust. By training the pixel classification tool within ilastik on 10 high-resolution images, ilastik generates probability maps for hundreds of previously unseen and untrained images, which can be used to quantify pigment areas in imageJ. We compare the results of ilastik quantification to qualitative binning of flies into red/white categories and to results obtained through quantification of chemically extracted pigments demonstrating ilastik quantification is consistent with observations gained using historic methods on modern instruments (pigment extraction + nanodrop).

Materials and methods

Fly crosses

Fly stocks were maintained at room temperature on standard fly food. Young males were mated with virgin females at 26°C for 24 hours. Adult offspring eclosed by day 10 and were imaged on day 14 (Fig. 3).

Microscopy and imaging

Image stacks were taken on an Olympus BH2 upright microscope with 10x objective and Sony DSLR mirrorless camera. Flies were dissected and imaged 4 days after eclosing (Supplementary Fig. 1). One eye was imaged per fly. Image acquisition was manual (with shutter remote) including Z-step size for a series of 9-17 images per eye. The image series was then stacked together in HeliconFocus to produce a singular pseudo-3D image file with most ommatidia in focus.

Image analysis—training ilastik

Ten images were imported to ilastik for pixel classification training. Labels were created for red, white, and hair pixels. Each training image was manually painted with the mouse and cursor to assign pixels to respective labels, taking extra care to adjust the software prediction when pixel classes were incorrectly predicted (Fig. 4, and Supplementary Fig. 2). After training, prediction maps (three-channel HDF5 files) were batch exported for all eye images (267 total in our dataset).

Image analysis—measurements in imageJ

HDF5 files were individually imported to imageJ (Supplementary Fig. 3, import macro available). The “threshold” tool was used to threshold the prediction map for each label at a minimum score of 0.5, eliminating low-scoring data. This step generated a binary image that we used to measure the area fraction for each label (Fig. 5, threshold macro available). Using the “oval brush selection” tool, a region of interest (ROI) was identified that

included only ommatidia. Using the “measure” tool, the area was calculated for each channel or pixel class. We plot the normalized area fractions for the red label of each image in Fig. 6B (detailed in protocol section). We have successfully carried out the methods described here using a 2020 Macbook Air, with 8GB of RAM and an Apple M1 chip.

Pigment extraction and absorbance for comparison

Fly heads were removed with forceps and placed in a 1.5ml Eppendorf tube (25 heads/tube) with 100µl acidified ethanol (pH=2) for 72 hours. After 72 hours, 2µl of extract was used to measure absorbance at 485 nm on a Nanodrop spectrophotometer.

Results

A key strength of ML-based image analysis is the power to distinguish subtle eye pigment differences within individuals and between closely related stocks. This analysis also avoids investigator bias and enables multiple individuals to contribute to the same data set.

A typical observation when working with PEV flies is that they naturally acquire suppressor mutations that revert the eye color to wildtype. Due to the accumulation of multiple trans-acting suppressor mutations [26], geneticists select for variegation enhancement (whiter eyes) every few generations to maintain their stocks. Thus, PEV stocks tend to have different levels of variegation without selective inbreeding. We examined eyes from two isolated *Drosophila melanogaster* PEV stocks carrying the same chromosome rearrangement $In(1)w^{m4}$ A&B, starved $In(1)w^{m4}$ flies, and white eyes from the w^{1118} stock. Starved flies were aged beyond 3 weeks and have severely enhanced variegation due to malnutrition and developmental delays [30]. w^{1118} flies were used as a control.

We followed the above approach (protocol available, S1), including quantification using the three methods listed in Fig. 3. Importantly, our results for genotype comparisons maintain congruency between the methods: $In(1)w^{m4}$ -A is more red than $In(1)w^{m4}$ -B for binning ($P = .0268$), imaging ($P = .00193$), and extraction ($P = .03005$). $In(1)w^{m4}$ -starved is also more red than w^{1118} using imaging ($P < .00001$). Each method has inherent strengths and weaknesses. The binning strategy (Fig. 6A) is rapid and high throughput, since whole flies are scored based on human interpretation of eye pigment. In this method, flies are grouped into three categories, bins of <33% red, 33–66% red, and >66% red. Individual data points represent the number of >66% red males counted in a single day. This approach is largely subjective and the least accurate of the three strategies. Adding bins to this method will increase precision at the expense of throughput or assay time, although coarse bins may preclude statistical analysis as is the case for $In(1)w^{m4}$ -starved versus w^{1118} (Fig. 6A). Image analysis

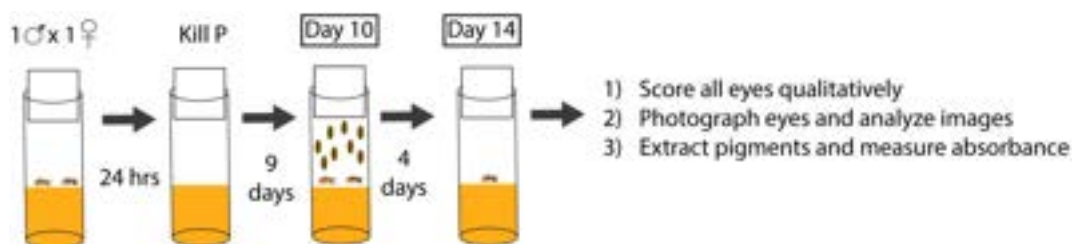


Figure 3. Rearing PEV flies for ommatidia imaging. Temperature and mating times are chosen to optimize offspring emergence on day 10 and to ensure flies are the same age during imaging. F₁ progeny are scored on day 10, sons are saved in a new vial and aged 4 more days to allow for white expression and pigment import. Only males are assayed because $In(1)w^{m4}$ is the stock X chromosome.

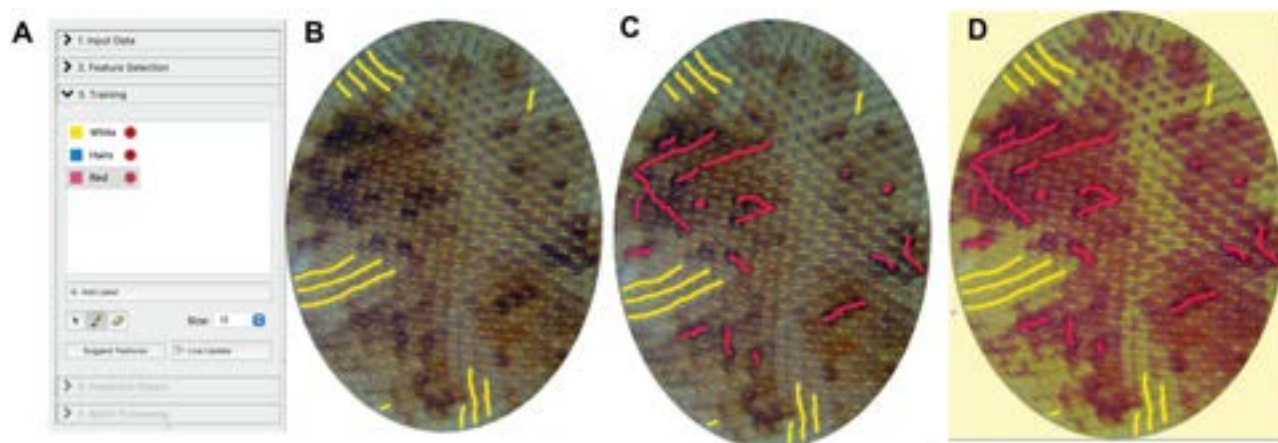


Figure 4. ilastik training. **A:** control panel for ilastik GUI; **B:** training ilastik to identify white ommatidia; **C:** training ilastik to identify pigmented ommatidia; **D:** Live Update shows how the entire image will be classified.

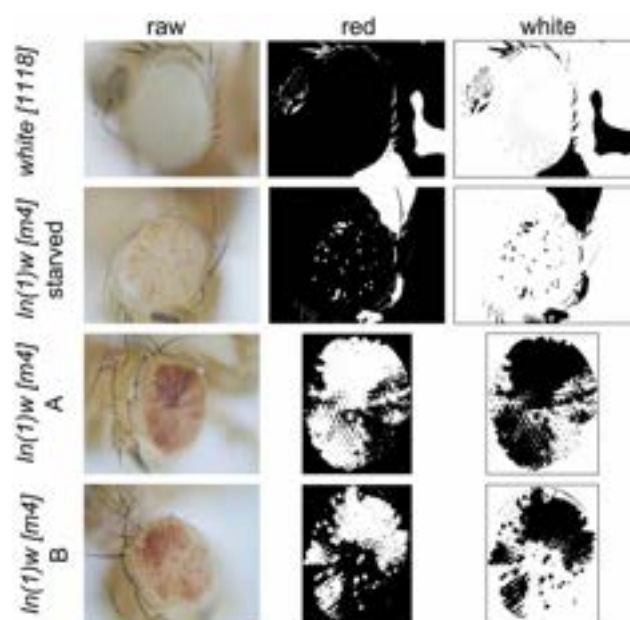


Figure 5. Examples of binary masks generated after thresholding HDF5 files to a minimum of 0.5 in imageJ. White = positive pixel for that label. Black = negative pixel for that label. White eyes have no ommatidia that are positive in the red channel and all are positive in the white channel. Starved eyes have several red-positive cells, and the well-fed PEV lines have many red ommatidia. ilastik can make predictions of entire raw images (top two rows) or image cutouts (bottom two rows).

(Fig. 6B) is the most accurate method, as it examines individual eyes and can detect minute changes in pigment between ommatidia. It is also the slowest, and most labor intensive since it involves photography, image preparation, software training, and computing. Pigment extraction (Fig. 6C) is quantitative, fast, and high throughput, but variation among individuals is undetectable due to the pooled approach. Pooling also requires many more individual flies, requiring stock expansion to achieve adequate sample sizes.

Discussion

Our results demonstrate that ilastik provides a useful and accessible method for quantification of eye pigmentation. While differences between genotypes acquired by binning flies and extracting/measuring pigments align with results of our method,

ilastik is quantitative and enables analysis of individual-level variation. The method we outline can be immediately leveraged by the community for PEV quantification. We discuss relevant considerations, constraints, and benefits of this method below.

The results of ML-based approaches like ilastik depend on the quality and quantity of images in the training set [1]. Ultimately, users must define the adequate number of training images for their project given their processing ability and research focus. The number of images required for a training set is not fixed and depends on image complexity and the desired classification accuracy. We trained the pixel classifier on a set of 10 images for this work, which was adequate for our needs (Supplementary Fig. 4). Accuracy of pixel classification has two variables: the number of photos and the number of teaching stroke pixels. Increasing photo numbers will improve errors caused by lighting reflections and shadow differences, and increasing strokes can improve and potentially correct predictions. Therefore, it is best to choose strokes carefully to train the software based on the prediction errors observed. As noted by the ilastik developers, ensuring that training images represent the complete range of features in the data set is more important than the inclusion of a large numbers of images. For our data, quantification improved as training images were added but the returns diminished after approximately seven to eight images, with us approaching computing limits on a typical machine at ~10 images. Clear tradeoffs exist between computing power and training strength with larger training sets—and many large-brushed training strengths—requiring more computing power.

Computing relative areas of a pseudo-3d image is challenging since the ommatidia at the boundaries of the eye have significantly less surface area in the image. An ideal solution is to assign each ommatidia a datapoint which can be accomplished with instance segmentation. If each ommatidia is a discrete object, those RGB data can then be exported independently, resulting in quantification focused on cell counts rather than surface area. We investigated this concept in ilastik using a control eye and assigning labels to “cell” and “cell boundary” during pixel classification. We then importing the HDF5 file to ilastik for object classification (Supplementary Fig. 5). While we have not fully applied this method to work for PEV images, it highlights that the accuracy of using ML to quantify complex phenotypes depends on the user-defined question, data quality, and the end-goal.

ML-based image analysis provides several benefits over alternative methods. First, researchers can tailor the process to their available resources, while maintaining reproducible and consistent results of individual-level variation not possible using other methods.

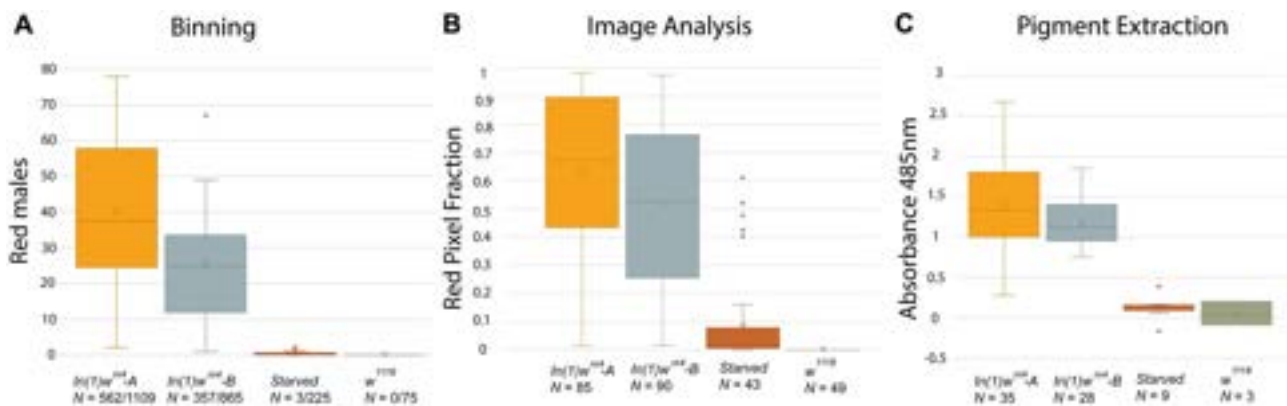


Figure 6. Results for three eye pigment quantification methods. P-values are from Mann–Whitney U tests: P_1 = comparison of stocks A & B, P_2 = comparison of stocks starved & *w¹¹¹⁸*. **A:** qualitative binning of whole flies into bins of <33%red, 33–66%red, and >66%red, N = red flies/total, P_1 = .0268, P_2 = NA (coarse binning precludes statistical analysis); **B:** results from ilastik-based image analysis, N = individual eye images, P_1 = .00193, P_2 < .00001; **C:** distribution of absorbance values after pigment extraction N = vials with 25 heads each ($N \times 25$ = # of eyes) P_1 = .03005, P_2 = NA (pooling requires many more flies for sampling to enable statistical analysis).

Second, classification of pixels in ommatidia that are either orange or pink (or reducing errors such as hairs or reflections) is possible by adding additional labels. Analysis can be applied to other *Drosophila* screens that require red: white quantification in the eye such as FLP/FRT mitotic recombination assays. Finally, similar ilastik-based pixel classification workflows could be implemented to quantify many other complex mosaic phenotypes such as GFP in living tissues, yellow on cuticles, mosaic patches of *multiple-wing-hairs*, etc [15, 31]. While other ML-based approaches could be leveraged for these analyses [32, 33], ilastik is accessible to non-experts with no previous coding proficiency and is designed for interactive training with manual user guiding of the classifier. ilastik also estimates which region of the raw data requires processing at any given moment, reducing RAM requirements and allowing everyday computers to handle large datasets. Comparing the results of ilastik to coding-based ML approaches [33]—with emphasis on the sensitivity of results to the quality and quantity of data in the training set given computational constraints—is an obvious direction for future work.

We recommend using multiple methods to determine PEV levels in any given variegating stock. When considering resolution, and accuracy, our view is that ML-based imaging is a preferred method, and ilastik is an accessible option for users. We hope this protocol is useful to the community.

Acknowledgements

We would like to give special thanks to Timothy B. Wheeler for assistance with the imaging setup, Alea Van Sickle for imaging efforts, and ML image analysis expert Dr. Drew Bischel for a conversation that led to the rapid development of this method. Additional thanks to the community of image analysis developers who host the site <https://forum.image.sc/> and actively provide feedback and answer questions.

Supplementary data

Supplementary data are available at *Biology Methods and Protocols* online.

Author contributions

Hunter Hill (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Investigation [lead], Methodology [lead], Validation

[lead], Visualization [lead], Writing—original draft [lead]), William Sullivan (Conceptualization [supporting], Investigation [supporting], Methodology [supporting], Project administration [supporting], Resources [supporting], Supervision [supporting], Writing—review & editing [equal]), and Brandon S. Cooper (Conceptualization [supporting], Data curation [supporting], Formal analysis [supporting], Funding acquisition [lead], Investigation [supporting], Methodology [supporting], Project administration [supporting], Resources [lead], Supervision [lead], Validation [supporting], Visualization [supporting], Writing—review & editing [supporting]).

Conflict of interest statement. None declared.

Funding

This work was supported by National Institutes of Health MIRA (R35GM139595, WS; R35GM124701, BSC) and National Science Foundation (NSF) CAREER (2145195, BSC) Awards. Funding from the M.J. Murdock Charitable Trust (202324717) toward the Cellular-Genetic Ecosystem in the University of Montana Genomics Core also contributed to this work.

Data availability

A full dataset for each method is available in [supplementary file S2](#). HDF5 files and raw image files are available upon request.

References

1. Berg S, Kutra D, Kroeger T et al. ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* 2019;**16**:1226–32.
2. Mackenzie SM, Brooker MR, Gill TR et al. Mutations in the white gene of *Drosophila melanogaster* affecting ABC transporters that determine eye colouration. *Biochim Biophys Acta* 1999;**1419**:173–85.
3. Morgan TH. Sex limited inheritance in *Drosophila*. *Science* 1910;**32**:120–2.
4. Muller H. Types of visible variations induced by X-rays in *Drosophila*. *Journ of Gen* 1930;**22**:299–334.
5. Gowen JW, Gay EH. Chromosome Constitution and Behavior in Eversporting and Mottling in *Drosophila Melanogaster*. *Genetics* 1934;**19**:189–208. <https://doi.org/10.1093/genetics/19.3.18917246719>

6. Ashburner M, Golic KG, Hawley RS. *Drosophila: A Laboratory Handbook*. New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2005.
7. Spofford JB. In: Ashburner, M & Novitski, E (eds) *The Genetics and Biology of Drosophila*, Vol. **1C**. New York, New York: Academic Press, 1976, 955–1018.
8. Elgin SCR, Reuter G. Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. *Cold Spring Harb Perspect Biol* 2013;**5**:a017780.
9. Hartl DL, Ruvolo M. *Genetics: Analysis of Genes and Genomes*. Burlington, MA: Jones & Bartlett, 2012.
10. Maggert KA, Golic KG. The Y chromosome of *Drosophila melanogaster* exhibits chromosome-wide imprinting. *Genetics* 2002;**162**:1245–58.
11. Wallrath LL, Elgin SCR. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev* 1995;**9**:1263–77.
12. Weiler KS, Wakimoto BT. Heterochromatin and gene expression in *Drosophila*. *Annu Rev Genet* 1995;**29**:577–605.
13. Lu BY, Ma J, Eissenberg JC. Developmental regulation of heterochromatin-mediated gene silencing in *Drosophila*. *Development* 1998;**125**:2223–34.
14. Lu BY, Bishop CP, Eissenberg JC. Developmental timing and tissue specificity of heterochromatin-mediated silencing. *Embo J* 1996;**15**:1323–32.
15. Bughio F, Huckell GR, Maggert KA. Monitoring of switches in heterochromatin-induced silencing shows incomplete establishment and developmental instabilities. *Proc Natl Acad Sci U S A* 2019;**116**:20043–53.
16. Shultz J. Interrelations of factors affecting heterochromatin induced variegation in *Drosophila*. *Genetics* 1950;**35**:134.
17. Spofford JB. Single-locus modification of position-effect variegation in *Drosophila melanogaster*. II. Region 3C Loci. *Genetics* 1969;**62**:555–71.
18. Schotta G, Ebert A, Dorn R et al. Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin Cell Dev Biol* 2003;**14**:67–75.
19. Ebert A, Schotta G, Lein S et al. Su(var) genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* 2004;**18**:2973–83.
20. Iyer J, Wang Q, Le T et al. Quantitative assessment of eye phenotypes for functional genetic studies using *Drosophila melanogaster*. *G3 (Bethesda)* 2016;**6**:1427–37.
21. Tran H, Dostatni N, Ramaekers A. EyeHex toolbox for complete segmentation of ommatidia in fruit fly eyes. *bioRxiv* 2024.03.04.583398; <https://doi.org/10.1101/2024.03.04.583398>
22. Golic KG, Lindquist S. The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell* 1989;**59**:499–509.
23. Golic MM, Rong YS, Petersen RB et al. FLP-mediated DNA mobilization to specific target sites in *Drosophila* chromosomes. *Nucleic Acids Res* 1997;**25**:3665–71.
24. Bladen J, Cooper JC, Ridges JT et al. A new hybrid incompatibility locus between *Drosophila melanogaster* and *Drosophila sechellia*. *Genetics* 2024;**226**:1–9.
25. Hoyer-Fender S. Transgenerational effect of drug-mediated inhibition of LSD1 on eye pigment expression in *Drosophila*. *BMC Ecol* 2020;**20**:62.
26. Wang SH, Elgin SCR. The impact of genetic background and cell lineage on the level and pattern of gene expression in position effect variegation. *Epigenetics Chromatin* 2019;**12**:70.
27. Ephrussi B, Herold JL. Studies of eye pigments of *Drosophila*. I. Methods of extraction and quantitative estimation of the pigment components. *Genetics* 1944;**29**:148–75.
28. Pipkin HJJ, Lindsay HL, Smiley AT et al. An accessible digital imaging workflow for multiplexed quantitative analysis of adult eye phenotypes in *Drosophila melanogaster*. *bioRxiv* 2024; <https://doi.org/10.1101/2024.01.26.577286>
29. Breiman L. *Machine Learning* 2001;**45**:5–32. <https://doi.org/10.1023/A:1010933404324>
30. Gloor H, Van Breugel FMA, Volkers WS. Some observations on variegation in *Drosophila*. *Genetica* 1967;**38**:95–114.
31. Germani F, Bergantinos C, Johnston LA. Mosaic analysis in *Drosophila*. *Genetics* 2018;**208**:473–90.
32. Bischel DM. Dissertation: *Classification of the Sex of Drosophila Suzukii with Pre-Trained Networks*. UC Santa Cruz. escholarship.org/uc/item/1j1173ms. 2023.
33. Radford A, Kim JW, Hallacy C et al. Learning transferable visual models from natural language supervision. In: *International Conference on Machine Learning* 2021, 8748–36.