A high throughput method for egg size measurement in *Drosophila*

Authors

Neda Barghi[1*] and Claudia Ramirez-Lanzas[1,2]

Affiliations

[1] Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria

[2] Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Vienna, Austria

*Corresponding author: Neda Barghi, barghi.neda@gmail.com

Address: Institut für Populationsgenetik, Veterinärmedizinische Universität Wien,

Veterinärplatz 1, 1210 Wien

Keywords: *Drosophila* egg size, Large particle flow cytometry, Genetic variation

Abstract

Many life-history traits are used as proxies of fitness in *Drosophila*. Egg size is an adaptive trait that can be used for fitness measurement. However, the low throughput of manual measurement of egg size has hampered the widespread use of this trait in evolutionary biology. We established a method for accurate and high throughout measurement of *Drosophila* egg size using large particle flow cytometry which allows the automated analysis of large particles based on size and optical density. The measurement of egg size is high throughput (average 238 eggs analyzed per minute) and sorting eggs of specific size is efficient (37-75 % efficiency). Using this approach, we showed the natural variation in egg size among strains of *Drosophila* species. The egg size distribution determined by this approach can be used as a measure of fitness and the reproductive investment in *Drosophila*. Measurement and sorting by flow cytometer do not decrease the egg viability making it a suitable approach for sorting viable eggs to be used for downstream analyses such as selection experiments.

Introduction

Drosophila is one of the most widely used sexual multicellular organisms in evolutionary biology. This popularity is partly due to *Drosophila*'s short generation time and ease of maintenance in the laboratory conditions. Life-history traits such as survival, longevity and female fecundity are usually measured as proxies of fitness in *Drosophila*. Another important life history trait is egg size, that in combination with egg number, i.e. female fecundity, is a proxy of reproduction investment (Avelar, 1993). In *Drosophila*, egg size affects other life-history traits, e.g. embryonic viability, hatching rate and embryonic development (Azevedo et al., 1997, p.). Environmental factors such as temperature influence the egg size (Avelar, 1993; Imai, 1932). *Drosophila* egg size increases with latitude in latitudinal clines in Australia and South America (Azevedo et al., 1996) suggesting that it is an adaptive trait under stabilizing selection (Azevedo et al., 1997).

The current methods for measuring *Drosophila* egg size involve the manual measurement of alive eggs under a microscope (Azevedo et al., 1996, 1997; Jha et al., 2015) or their images (Lott et al., 2007). Manual measurement is a tedious and error prone (i.e., experimenter's bias) task with low throughput that cannot ensure recovery of viable individuals. Despite the ecological importance of egg size (Azevedo et al., 1996; Markow et al., 2009) and its importance as a life history trait the low throughput of manual measurement has hampered the use of *Drosophila* egg size in evolutionary and population biology studies.

Flow cytometry is a well-established method for the analysis and sorting of small objects. However, the average egg size in *Drosophila* species ranges from 300 to 600 µm (Imai, 1932; Lott et al., 2007; Markow et al., 2009; Warren, 1924) which cannot be processed with the traditional flow cytometers. Advent of large particle flow cytometers has made the automated analysis and sorting of large particles up to 1500 µm possible. Analysis and sorting of nematodes (Pillai & Dandurand, n.d.), *D. melanogaster* (Furlong et al., 2001) and *Anopheles gambiae* eggs (Marois et al., 2012), and coral larvae (Randall et al., 2020) have been performed by measuring the fluorescence intensity of the expressed protein reporters. However, the fluorescence -based analysis requires fluorescence

labelling or transgenic modification of individuals. In addition to fluorescence measurement, the analysis and sorting of particles can also be done based on the size and optical density of objects (Morales et al., 2020).

We developed a method to accurately measure the size of *Drosophila* eggs with high throughput using large particle flow cytometry. Our protocol involves in-silico filtering steps to remove small objects and misaligned eggs. Using this approach, we show the natural variation in egg size distribution in different strains of *Drosophila* which can be exploited for investigating the genetic basis of egg size using GWAS and selection experiments. Moreover, the high throughput of egg size measurement in *Drosophila* using this approach allows the use of egg size as a life history trait in evolutionary and ecological studies.

Materials and Methods

Maintenance of *Drosophila* strains

Drosophila strains were maintained on standard *Drosophila* medium at 25 °C and 50-60 % humidity. Larval crowding (Venkitachalam et al., 2022) and temperature (Avelar, 1993; Azevedo et al., 1996; Imai, 1932) affect *Drosophila* egg size. To reduce the environmental effects we maintained the cultures at low density (400 eggs in a bottle) using a pipetting method (Nouhaud et al., 2018) for at least one generation before egg size measurement.

Preparation of samples for flow cytometry

For each strain, around 500 4-6 days old individuals were transferred to an embryo collection cage (Cat. #59-101, Genesee Scientific) with an agar plate (4% agar and 4% sucrose) topped with yeast paste to facilitate oviposition (Becher et al., 2012). Although the age of parents doesn't affect the egg size (Imai, 1932; Lott et al., 2007; Warren, 1924) we restricted the measurements to eggs laid by 4-6 days old flies. We limited oviposition period to 17-19 hours to prevent hatching of eggs. Few drops of lukewarm water were added to the agar plate to dissolve the yeast paste, and the solution

of yeast and water together with the eggs were transferred to a fine net using a brush. Eggs were washed with water until all yeast grains were dissolved, the eggs were then transferred to a 1.5 ml tube containing 600 µl 1x PBS.

Size measurement and sorting with large particle flow cytometer

Large particle flow cytometers analyze and sort objects based on size, optical density and fluorescence intensity. These systems operate at slower flow rate and lower pressures compared to traditional flow cytometers to avoid disruptive shear forces. Samples are stirred continuously using a mechanical stirrer in a sample cup and travel to the flow cell surrounded by a sheath solution that focuses them into the center of the stream where they are interrogated by multiple lasers. If samples are to be sorted, a droplet containing fluid and the sorted object is generated which drops to the collecting container.

We ran all our samples on a Biosorter® (Union Biometrica) machine with fluidics and optics core assembly units (FOCA) 1000 µm using 1x PBS as sheath solution. To prevent clogging the flow cell, we re-suspended the eggs in 1x PBS such that the final density of ~400 eggs/ml was reached. We ran Biosorter® using FlowPilot™ software using pressure sample cup 0.65-0.7 and pressure diverter 1.2 and sorted samples using the following parameters: drop width 22 ms and sort delay 28 ms. We cleaned the mechanical stirrer with distilled water and the system with 1x PBS after running each sample to avoid cross contamination.

Biosorter® measures two optical characteristics of objects: time of flight (TOF) that is the time each object blocks the light, and extinction (EXT) which is determined by the total integrated signal of the blocked light (Fig. S1). We measured TOF of polystyrene reference beads (200 and 430 µm) and fitted a linear regression model using TOF as the response variable and size as the explanatory variable. Using the estimated intercept and slope, TOF measurements of all samples were converted to size (size = 5.94 TOF - 491.78). FlowPilot™ software allows real time visualization of EXT and TOF of objects as scatter plots and regions of interest, i.e. gates, can be chosen on this

plot for sorting objects (Fig. S1). Based on the estimated size using TOF measurements, multiple gates were set up which corresponded to different size ranges (Fig. S1). These gates were used for sorting eggs which were dispensed in 1x PBS and developed to adults on standard *Drosophila* medium.

In silico filtering of small objects and misaligned eggs

The distribution of the object sizes in each Biosorter® run was bimodal with the first mode corresponding to the egg debris and the second one to the eggs. We fitted kernel density estimation that is controlled by a smoothing parameter, i.e. bandwidth, to the size distribution. To identify a suitable bandwidth for smoothing the density distribution we performed an exhaustive search over a range of bandwidth by fitting kernel density function with Gaussian shape. We performed K-Folds cross validation by splitting dataset into train and test sets. We split data into five folds; four folds formed the training set and one fold was used as a validation set. The best bandwidth giving the highest score was chosen as the bandwidth to smooth the kernel density distribution. We then computed the log-likelihood of each dataset under the fitted kernel density distribution and identified the border between the two modes as the first valley in the log-likelihood. The border between the two modes was used as the empirical size threshold for filtering small objects for each sample.

When an object passes through the laser path along its long axis the recorded TOF corresponds to its size (Fig. 1A). But if the object is misaligned, TOF and subsequently size will be underestimated (Fig. 1B). Using a more viscous sheath solution, e.g. 1x PBS with 4% methyl cellulose, will reduce the flow rate and facilitate the alignment of objects along their long axis when passing through the laser path (Fig. S2). However, the viscosity of sheath solution dramatically decreases the speed of analysis. To circumvent this issue, we used the optical density information to filter the objects that are not aligned in their long axis. The optical density which measures the extinction of objects varies as a function of the structure of objects (Fig. 1A-B). We extracted the maximum

optical density recorded for an object (W) and the total number of extinction measurements along each object's time of flight (L) and computed the ellipticalness index ($EI = \frac{Area}{\pi/4LW}$) which is the ratio of the area under the optical density profile of an object to the area of the closest elliptical profile (Li et al., 2022; Morales et al., 2020). Due to the elliptical shape of *Drosophila* egg, an egg aligned on the long axis has smaller W/L ratio than eggs aligned on the short axis (Fig. 1A-B). For each sample, we computed EI and W/L for all objects and retained objects distributed within half of standard deviation around median of EI and W/L.

Egg-to-adult viability assay

Large particle flow cytometers employ slower flow rate and lower pressure to ensure that the sorted objects are viable. However, eggs might be immersed in PBS up to several hours during size measurement and sorting. We performed egg-to-adult viability assay to determine the effect of submersion in PBS and passing through the flow cell on the development of eggs to adults. Around 2000 individuals from an outbred *D. simulans* population were transferred to an embryo collection cage with agar plates topped with yeast paste and laid eggs for 17-19 hours. We washed the yeast paste and transferred around 30 µl eggs to four 1.5 ml tubes containing 600 µl 1x PBS; 3 tubes corresponded to '4-hour', '6-hour,' and '8-hour' treatments and 1 tube to 'sorted 6-hour' treatment. The eggs in 4-hour, 6-hour, and 8-hour treatments were maintained at room temperature (25 °C) for 4, 6, and 8 hours, respectively. Then, 3-5 independent sets of 50 eggs (except for 6-hour treatment which had 14 sets) were manually counted under a stereoscope and each set was transferred to a vial containing standard *Drosophila* medium. We ran the eggs in sorted 6-hour treatment on Biosorter® and sorted 15 independent sets of 50 eggs from gates corresponding to 200-900 µm. Each set was transferred to a vial. The number of eclosed adults from each vial was counted on 2 consecutive days.

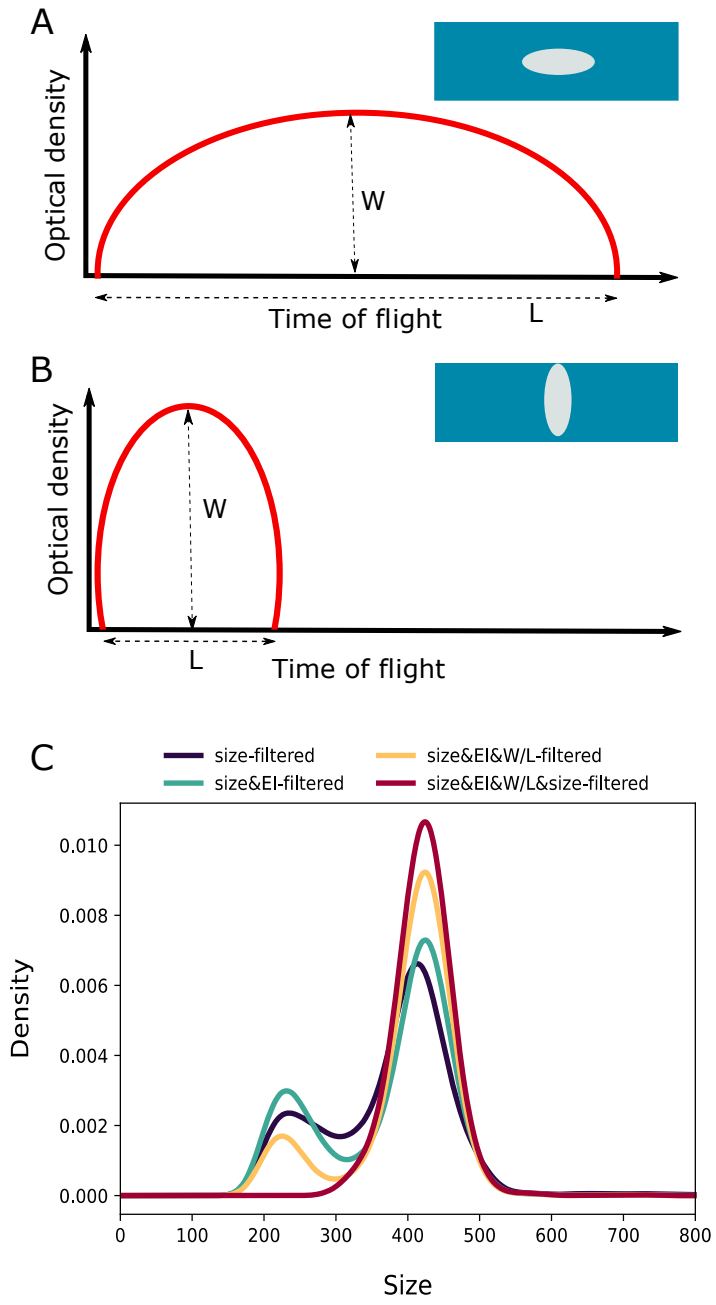

Fig. 1 In silico filtering of objects using the optical density profiles. A) *Drosophila* eggs are elliptical and accurate size measurement based on time of flight (TOF) is obtained when an egg is aligned on its long axis. B) The size of misaligned egg, i.e. aligned on its short axis, are underestimated. The misaligned eggs can be identified based on ellipticalness index (EI) and W/L ratio where W is the maximum optical density of an object and L is the number of extinction measurements along an object's time of flight. C) The initial size distribution is bimodal with the 1st mode around 92 µm (Table S1) corresponding to egg debris (not shown here). Removal of these small objects (size-filtered) reduces the number of small objects. Further filtering steps based on EI (size&EI-filtered) and W/L parameters (size&EI&W/L-filtered) reduces the number of small objects which correspond to misaligned eggs. A final size filtering step (size&EI&W/L&size-filtered) results in a unimodal size distribution. Size is in µm unit.

To test for the effect of PBS submersion on egg to adult viability we fitted a generalized linear mixed model (using a binomial distribution) with submersion time as a fixed categorical effect with 3 levels (4-hour, 6-hour, and 8-hour). We also tested the influence of sorting on Biosorter® on egg to adult viability by fitting a generalized linear mixed model (using a binomial distribution) and used the sort status with 2 levels (sorted and not sorted: sorted 6-hour and 6-hour treatments) as a fixed categorical effect. In both models, replicates were included as a random effect. The significance of the fixed effect was tested using ANOVA F-tests and we used Tukey's HSD to correct for multiple testing.

Manual egg size measurement

To validate the estimated egg size using Biosorter®, we sorted eggs from a gate corresponding to 400-500 µm in a 96-well plate and manually measured the eggs. We transferred the eggs to a petri-dish containing a dark medium (4% agar and 0.5% charcoal), took high resolution photos using a Leica stereomicroscope (M205 FA) and measured the egg sizes manually using ImageJ (Schneider et al., 2012). The scale of mages was transformed to µm and the egg size was measured by drawing a line along the longitudinal axis of each egg. The concordance between manual and Biosorter® size estimates was tested using Spearman's rank correlation.

Results

Egg size distribution after in-silico filtering

The size distribution of eggs was bimodal (Fig. 1C). The average size of the 1st mode was 92.38 µm across 16 samples (Table S1). We filtered datasets using sample-specific threshold, ranging from 92.9 to 212.21 µm (Table S1) to remove the objects in the 1st mode that mostly contained egg debris. However, the 2nd mode that corresponds to the size of *Drosophila* eggs (Imai, 1932; Lott et al., 2007; Markow et al., 2009; Warren, 1924), contained both properly aligned and misaligned eggs (Fig. 1A-B). Thus, we computed ellipticalness index (EI) and the ratio of the maximum optical density

of an object to the number of extinction measurements along an object's time of flight (W/L) values using the optical density data for each sample. Each dataset was further filtered by removing eggs that were not aligned on the long axis (Fig. 1B) by retaining objects with EI and W/L values within half of standard deviation around median. These filtering steps removed objects with small size but not large size showing that EI and W/L parameters properly identify misaligned eggs with lower estimated size (Fig. 1C). We anticipate that the average maximum height of optical density (W) of retained objects after filtering can be used in combination with other sorting criteria, e.g. size, to enhance the efficiency of sorting eggs of desired size. As the final filtering step, we removed the objects corresponding to the 1st peak of the egg distribution (Fig. 1C). The computed sample-specific empirical threshold for filtering these objects ranged from 163.36 to 317.12 µm (Table S2). We also filtered small objects using a constant threshold, i.e. the average sample-specific threshold across 16 samples (184.3 and 252.56 µm for the first and last filtering of 1st peak of the bimodal distribution), and observed a high correlation in the number and average size of the filtered small objects (Fig. S3A-B) between both filtered datasets. Moreover, the number of eggs and mean, and median of their distribution were highly corelated between the datasets filtered using sample-specific and constant thresholds (Fig. S3C-E). Therefore, for all the subsequent analysis sample-specific thresholds were used. Interestingly, we observed a moderate correlation between the number of eggs and small objects across samples (Spearman's rank correlation, $\rho = 0.506$, p-value = 0.046). This suggests that small objects likely correspond to egg debris and yeast that may increase with the total number of eggs in a sample.

Throughput and accuracy of egg size measurement and sorting

We analyzed average of 1681 eggs/min which corresponds to 238 eggs/min after filtering small and misaligned objects (Fig. S4A). For sorting eggs, we set up several gates that corresponded to 200 to 900 µm with intervals of 100 µm. We evaluated the efficiency of sorting as the sort recovery which is the percentage of objects that were sorted and dispensed as a fraction of all the objects fulfilling

sorting criteria, i.e. size. The median sort recovery across gates of different size ranges varied from 37.02 to 75.36% (Fig. S4C). The sort recovery partially depends on the number of available eggs in specified gate (Fig. S4B) as there is a slight significant correlation between the number of objects that match the size criteria and the sort recovery (Spearman's rank correlation ρ = 0.43, *p*-value = 0.014). In our hands, the speed of sorting ranged from 29.07 to 187.14 eggs/min depending on the available number of eggs in the specified gate (Fig. S4D). We manually measured the size of 49 eggs sorted from the gate corresponding to 400-500 µm and observed a high correlation between the manual and biosorter® size measurements (Fig. 2).

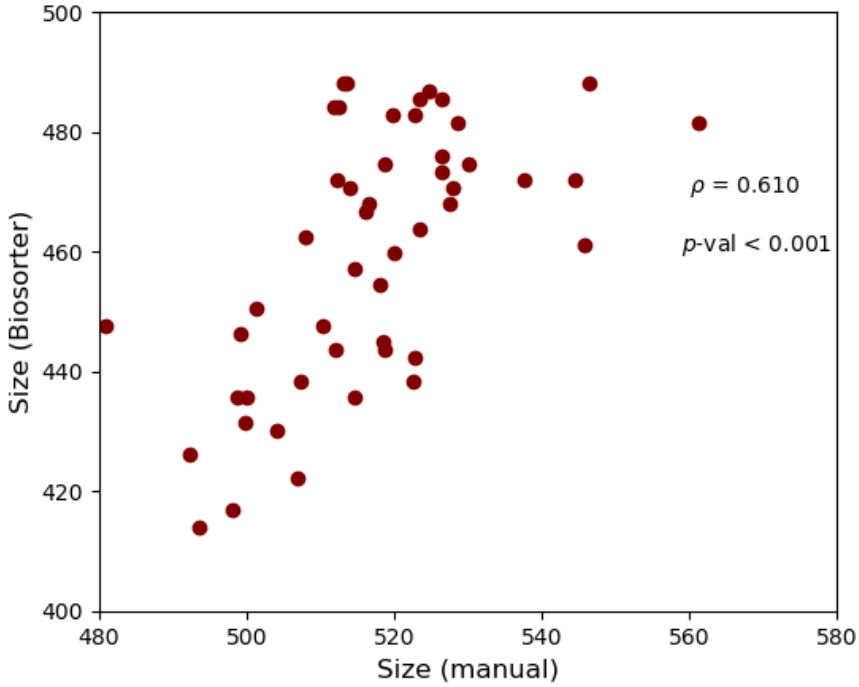

Fig. 2 Accuracy of estimated egg sizes using large particle flow cytometry. A total of 49 eggs were sorted from a gate corresponding to 400-500 µm by large particle flow cytometry. The size of these eggs was manually measured using ImageJ. The estimated sizes through manual measurement and large particle flow cytometry are highly correlated (Spearmans rank correlation ρ = 0.61, *p*-value < 0.001).

The effect of handling and sort by large particle flow cytometer on egg viability

To demonstrate the suitability of using Biosorter® for measurement and sorting of viable eggs, we tested whether the sorted eggs develop normally to adults. We observed a slight reduction in egg-

to-adult viability after 8 hours of PBS submersion (38.33±2.31 out of 50) compared to 4 (42.8±3.42 out of 50) and 6 (42.5±3.48 out of 50) hours of submersion (Fig. S5A, generalized linear mixed model followed by Tukey's HSD test with correction for multiple testing, $P_{6hour\text{-}8hour}$ = 0.0355, $P_{4hour\text{-}6hour}$ = 0.9715, $P_{4hour\text{-}8hour}$ = 0.0638). Moreover, eggs that passed through Biosorter® had a slight decrease in the egg-to-adult viability (36±4.64 out of 50) compared to the eggs submerged in 1x PBS for the same duration of time (42.5±3.48 out of 50) (Fig. S5B, generalized linear mixed model P <.0001). However, the reduction in egg-to-adult viability is minimal and can be compensated by increasing the number of sorted eggs.

Intra- and inter-specific variation in egg size distribution

Similar to previous studies (Imai, 1932; Lott et al., 2007; Warren, 1924), we did not observe differences in egg size laid by females of different age (Fig. S6). Thus, when available, we pooled the data from separate Biosorter® runs for a sample. We observed variation in mean (417 to 461.06 µm) and distribution of the egg size across 12 inbred strains of *D. simulans* (Signor et al., 2018) (Fig. 3A, Table S2). We also observed differences in mean and distribution of egg size between other *Drosophila* species (*D. melanogaster*, *D. erecta*, *D. santomea* and *D. mauritiana*) (Fig. 3B). While few strains were analyzed for *Drosophila* species, measurement of more strains will provide the opportunity to discern inter-specific egg size differences.

Discussion

Here, we present a method for fast and accurate measurement of *Drosophila* egg size using large particle flow cytometry. Taking advantage of this high throughput method we showed intra- (Fig. 3A) and inter-specific (Fig. 3B) variation in egg size distribution in *Drosophila* species. Egg size variation has been observed in *D. melanogaster* strains (Imai, 1932; Lott et al., 2007; Warren, 1924). However, the strains were mutants (Warren, 1924) or experimentally selected for large or small egg size (Lott et al., 2007) so the observed variation did not reflect the natural variation. Moreover, due to the low

throughput of manual measurement very few eggs, e.g. 8 (Lott et al., 2007) to 24 eggs (Markow et al., 2009), have been measured for each strain. The distribution of egg size in different strains (Fig. 3) suggests that measuring only few eggs will only provide information about the average but not the distribution of egg size. Moreover, the analysis of *Drosophila* egg size using large particle flow cytometry provides information about the total number of laid eggs, i.e. female fecundity, making it a high throughput approach for the measurement of egg size and number as life history traits contributing to the total reproductive investment in *Drosophila*.

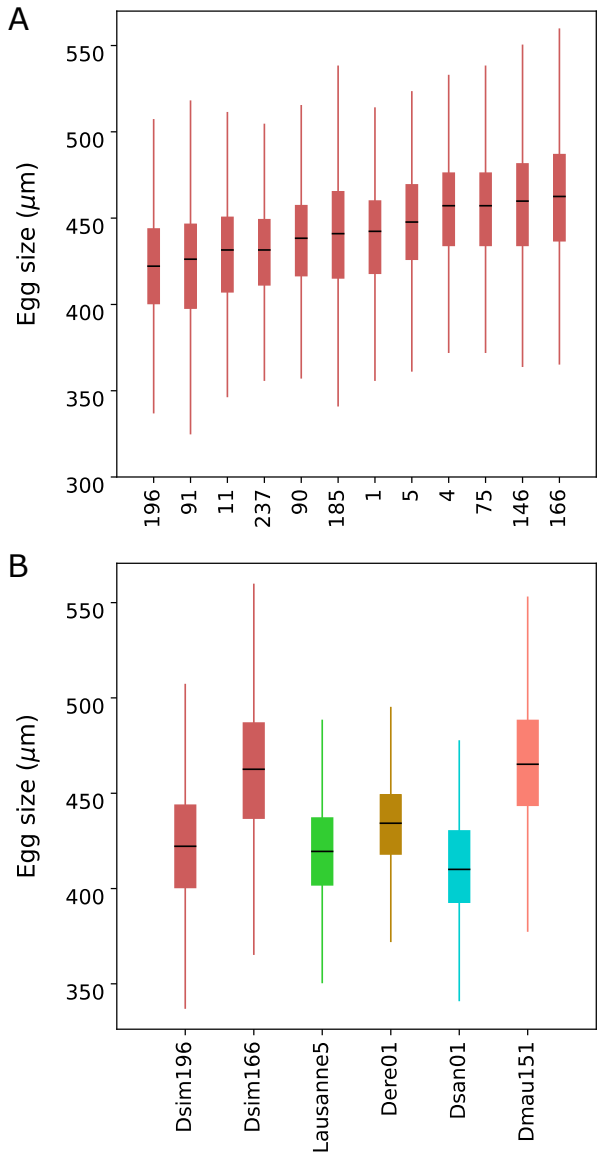

Fig. 3 Variation in egg size distribution among (A) different strains of *D. simulans* and (B) *Drosophila* species. Filtering steps described in Fig. 1 have been implemented for all samples. Except for line 1 and 4 (Bloomington stock center) all *D. simulans* lines are from (Signor et al., 2018). Lausanne5 (*D. melanogaster*), Dere01 (D. *erecta*), Dsan01 (D. *santomea*) and Dmau151 (*D. mauritiana*) are obtained from Bloomington stock center.

Despite the extensive use of *Drosophila* eggs for the study of oogenesis, little is known about the genetic basis of egg size mostly due to the low throughput of manual egg measurement. The availability of reference panels of *D. melanogaster* (Mackay et al., 2012) and *D. simulans* (Signor et al., 2018) coupled with the high throughput and accuracy of egg size measurement using large particle flow cytometry provide a great opportunity to investigate the genetic architecture of egg size variation (Fig. 3) through genome-wide association studies (GWAS).

The analyzed eggs using large particle flow cytometry are viable and can develop to adults or be used for subsequent molecular work. Viability of the analyzed eggs, coupled with the heritability and rapid response of egg size to selection (Jha et al., 2015) allows implementing selection experiments to understand the genetic basis of egg size. The lack of information about the distribution of egg size limits the experiments to truncating selection where selection is imposed on the individuals with extreme phenotypes (Jha et al., 2015; Schwarzkopf et al., 1999). However, Mmost quantitative traits are under stabilizing selection (Sanjak et al., 2018) and theoretical studies typically investigate the adaptive response of quantitative traits under stabilizing selection after shifts in the optima (Hayward & Sella, 2019; Jain & Stephan, 2017; Thornton, 2019). Thus, the possibility of sorting of viable *Drosophila* eggs from the range of the distribution using large particle flow cytometry allows performing more realistic selection experiments where the trait optimum can be shifted.

The opportunities provided by large particle flow cytometry are not limited to measuring *Drosophila* eggs but can be applied to any organism within the detectable range of this instrument (10-1500 µm) given that preliminary analysis are required to adjust the measurement and filtering parameters.

Author Contributions

Neda Barghi conceived the ideas and designed methodology; Neda Barghi and Claudia Ramirez-Lanzas collected the data; Neda Barghi analyzed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data Availability

All the data and code to reproduce the results are deposited in https://github.com/NedaBarghi/EggSizeBiosorterProtocol.

Acknowledgement

The BioSorter® Large Particle Flow Cytometer (Union Biometrica, Holliston, MA) of the Max Perutz Labs BioOptics FACS Facility (https://www.maxperutzlabs.ac.at/research/facilities/biooptics-facs) was used for sorting. We thank Johanna Stranner and Kitti Dora Csalyi for their help in optimizing the sorting protocol. We thank Paula Marconi for her help in egg-to-adult viability assay. We thank Emmanouil Lirakis for assistance with using Leica stereomicroscope. This work was supported by the Austrian Science Fund (FWF, P 32672). We thank Sara Signor (North Dakota State University) for providing *D. simulans* inbred lines. Lausanne5 (*D. melanogaster*), Dere01 (D. *erecta*), Dsan01 (D. *santomea*) and Dmau151 (*D. mauritiana*) are obtained from Bloomington stock center by Robert Kofler.

Conflict of Interest statement

The authors have no conflicts of interest to declare.

References

Avelar, T. (1993). Egg size in Drosophila: Standard Unit of investment of variable response to environment? The effect of temperature. *Journal of Insect Physiology*, *39*(4), 283–289. https://doi.org/10.1016/0022-1910(93)90058-Y

Azevedo, R. B. R., French, V., & Partridge, L. (1996). Thermal Evolution of Egg Size in Drosophila melanogaster. *Evolution*, *50*(6), 2338–2345. https://doi.org/10.2307/2410702

Azevedo, R. B. R., French, V., & Partridge, L. (1997). Life-History Consequences of Egg Size in *Drosophila Melanogaster*. *The American Naturalist*, *150*(2), 250–282. https://doi.org/10.1086/286065

Becher, P. G., Flick, G., Rozpędowska, E., Schmidt, A., Hagman, A., Lebreton, S., Larsson, M. C., Hansson, B. S., Piškur, J., Witzgall, P., & Bengtsson, M. (2012). Yeast, not fruit volatiles mediate Drosophila melanogaster attraction, oviposition and development. *Functional Ecology*, *26*(4), 822–828. https://doi.org/10.1111/j.1365-2435.2012.02006.x

Furlong, E. E. M., Profitt, D., & Scott, M. P. (2001). Automated sorting of live transgenic embryos. *Nature Biotechnology*, *19*(2), 153–156. https://doi.org/10.1038/84422

Hayward, L. K., & Sella, G. (2019). Polygenic adaptation after a sudden change in environment. *BioRxiv*, 792952. https://doi.org/10.1101/792952

Imai, T. (1932). The influence of temperature on egg size and variation in Drosophila melanogaster. *Wilhelm Roux' Archiv Für Entwicklungsmechanik Der Organismen*, *132*(1), 206–219. https://doi.org/10.1007/BF00579868

Jain, K., & Stephan, W. (2017). Modes of Rapid Polygenic Adaptation. *Molecular Biology and Evolution*, *34*(12), 3169–3175. https://doi.org/10.1093/molbev/msx240

Jha, A. R., Miles, C. M., Lippert, N. R., Brown, C. D., White, K. P., & Kreitman, M. (2015). Whole-Genome Resequencing of Experimental Populations Reveals Polygenic Basis of Egg-Size Variation in Drosophila melanogaster. *Molecular Biology and Evolution*, *32*(10), 2616–2632. https://doi.org/10.1093/molbev/msv136

Li, Y., Quinn, B. K., Niinemets, Ü., Schrader, J., Gielis, J., Liu, M., & Shi, P. (2022). Ellipticalness index – a simple measure of the complexity of oval leaf shape. *Pakistan Journal of Botany*, *54*(6). https://doi.org/10.30848/PJB2022-6(44)

Lott, S. E., Kreitman, M., Palsson, A., Alekseeva, E., & Ludwig, M. Z. (2007). Canalization of segmentation and its evolution in Drosophila. *Proceedings of the National Academy of Sciences*, *104*(26), 10926–10931. https://doi.org/10.1073/pnas.0701359104

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas, S., Han, Y., Magwire, M. M., Cridland, J. M., Richardson, M. F., Anholt, R. R. H., Barrón, M., Bess, C., Blankenburg, K. P., Carbone, M. A., Castellano, D., Chaboub, L., Duncan, L., … Gibbs, R. A. (2012). The Drosophila melanogaster Genetic Reference Panel. *Nature*, *482*(7384), 173–178. https://doi.org/10.1038/nature10811

Markow, T. A., Beall, S., & Matzkin, L. M. (2009). Egg size, embryonic development time and ovoviviparity in Drosophila species. *Journal of Evolutionary Biology*, *22*(2), 430–434. https://doi.org/10.1111/j.1420-9101.2008.01649.x

Marois, E., Scali, C., Soichot, J., Kappler, C., Levashina, E. A., & Catteruccia, F. (2012). High-throughput sorting of mosquito larvae for laboratory studies and for future vector control interventions. *Malaria Journal*, *11*(1), 302. https://doi.org/10.1186/1475-2875-11-302

Morales, A., Teapal, J., Ammerlaan, J. M. H., Yin, X., Evers, J. B., Anten, N. P. R., Sasidharan, R., & van Zanten, M. (2020). A high throughput method for quantifying number and size distribution of Arabidopsis seeds using large particle flow cytometry. *Plant Methods*, *16*(1), 27. https://doi.org/10.1186/s13007-020-00572-x

Nouhaud, P., Mallard, F., Poupardin, R., Barghi, N., & Schlötterer, C. (2018). High-throughput fecundity measurements in Drosophila. *Scientific Reports*, *8*(1), 1–6. https://doi.org/10.1038/s41598-018-22777-w

Pillai, D. S. S., & Dandurand, D. L.-M. (n.d.). *Potato cyst nematode egg viability assessment and pre-parasitic juvenile screening using large particle flow cytometer and sorter*. 28.

Randall, C. J., Speaks, J. E., Lager, C., Hagedorn, M., Llewellyn, L., Pulak, R., Thompson, J., Bay, L. K., Mead, D., Heyward, A. J., & Negri, A. P. (2020). Rapid counting and spectral sorting of live coral larvae using large-particle flow cytometry. *Scientific Reports*, *10*(1), 12919. https://doi.org/10.1038/s41598-020-69491-0

Sanjak, J. S., Sidorenko, J., Robinson, M. R., Thornton, K. R., & Visscher, P. M. (2018). Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences*, *115*(1), 151–156. https://doi.org/10.1073/pnas.1707227114

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schwarzkopf, L., Blows, M. W., & Caley, M. J. (1999). Life-History Consequences of Divergent Selection on Egg Size in *Drosophila melanogaster*. *The American Naturalist*, *154*(3), 333–340. https://doi.org/10.1086/303242

Signor, S. A., New, F. N., & Nuzhdin, S. (2018). A Large Panel of Drosophila simulans Reveals an Abundance of Common Variants. *Genome Biology and Evolution*, *10*(1), 189–206. https://doi.org/10.1093/gbe/evx262

Thornton, K. R. (2019). Polygenic Adaptation to an Environmental Shift: Temporal Dynamics of Variation Under Gaussian Stabilizing Selection and Additive Effects on a Single Trait. *Genetics*, *213*(4), 1513–1530. https://doi.org/10.1534/genetics.119.302662

Venkitachalam, S., Das, S., Deep, A., & Joshi, A. (2022). Density-dependent selection in Drosophila: Evolution of egg size and hatching time. *Journal of Genetics*, *101*(1), 13. https://doi.org/10.1007/s12041-021-01355-6

Warren, D. C. (1924). INHERITANCE OF EGG SIZE IN *DROSOPHILA MELANOGASTER*. *Genetics*, *9*(1), 41–69. https://doi.org/10.1093/genetics/9.1.41

Supplementary figures and Tables

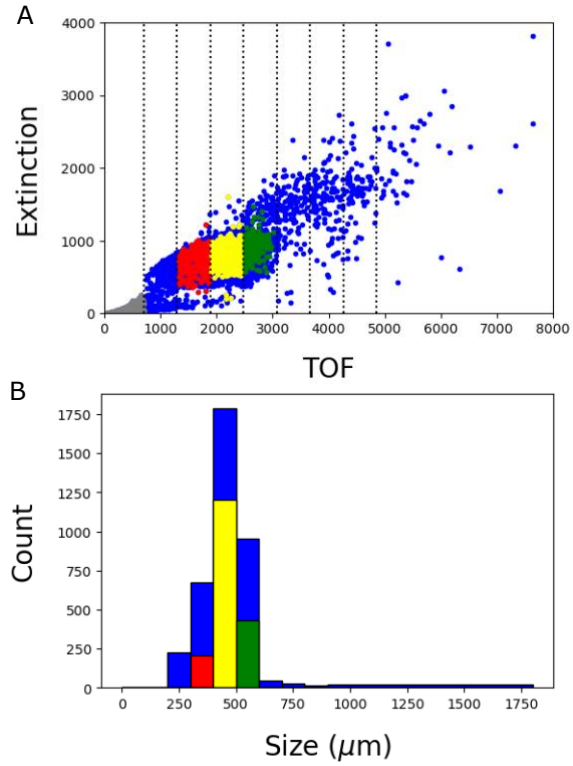

Fig. S1 The measured time an object blocks the light, time of flight (TOF), is used to estimate the object's size. *Drosophila* eggs are washed and dispensed in 1x PBS buffer and run on Biosorter®. TOF and extinction are measured (A) and TOF is converted to size (µm) using the TOF data of reference beads. Based on the size distribution, user-defined sorting gates are set up (depicted as dotted lines in (A), these gates correspond to size 200-900 µm with intervals of 100µm), and the required number of eggs from each gate can be sorted. Red, yellow and green bars in (B) show the number of sorted eggs in these 3 size gates in (A).

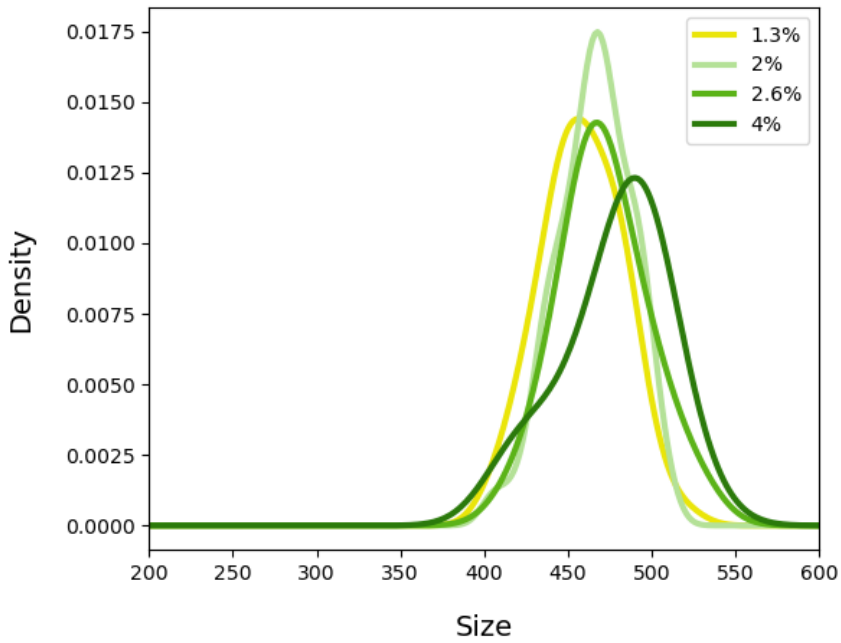

Fig. S2 The distribution of egg size estimated using different concentrations of methyl cellulose (MC) in 1x PBS as sheath buffer. As the concentration of MC increases the sheath buffer becomes more viscous which will reduce the flow rate. The slower flow rate facilitates the alignment of objects along the long axis while passing through the laser path. However, the increased viscosity of the sheath buffer dramatically reduces the speed of analysis. The same sample has been used for all runs shown in the plot.

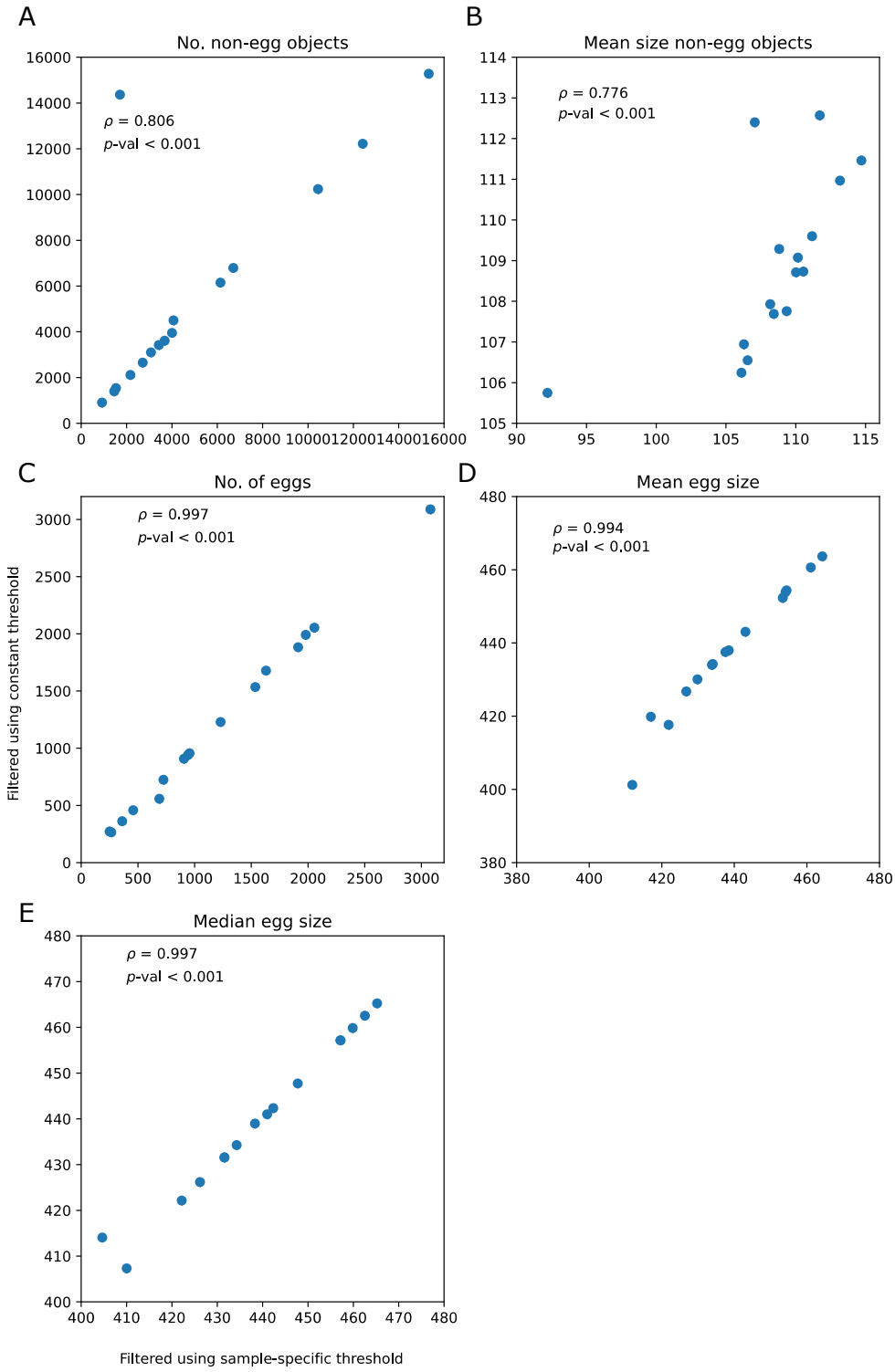

Fig. S3 The filtered egg size datasets using sample-specific threshold or a constant threshold, i.e. average sample-specific, to remove egg debris and misaligned eggs are highly correlated. The number (A) and average size (B) of the objects in the 1st mode of the distribution of each sample are highly correlated regardless of the threshold used for filtering. Moreover, the number of eggs (C), mean (D), and median (E) of the distribution after filtering misaligned eggs were also highly correlated between the datasets filtered using sample-specific and constant thresholds. The datasets are the same as those in Fig. 3.

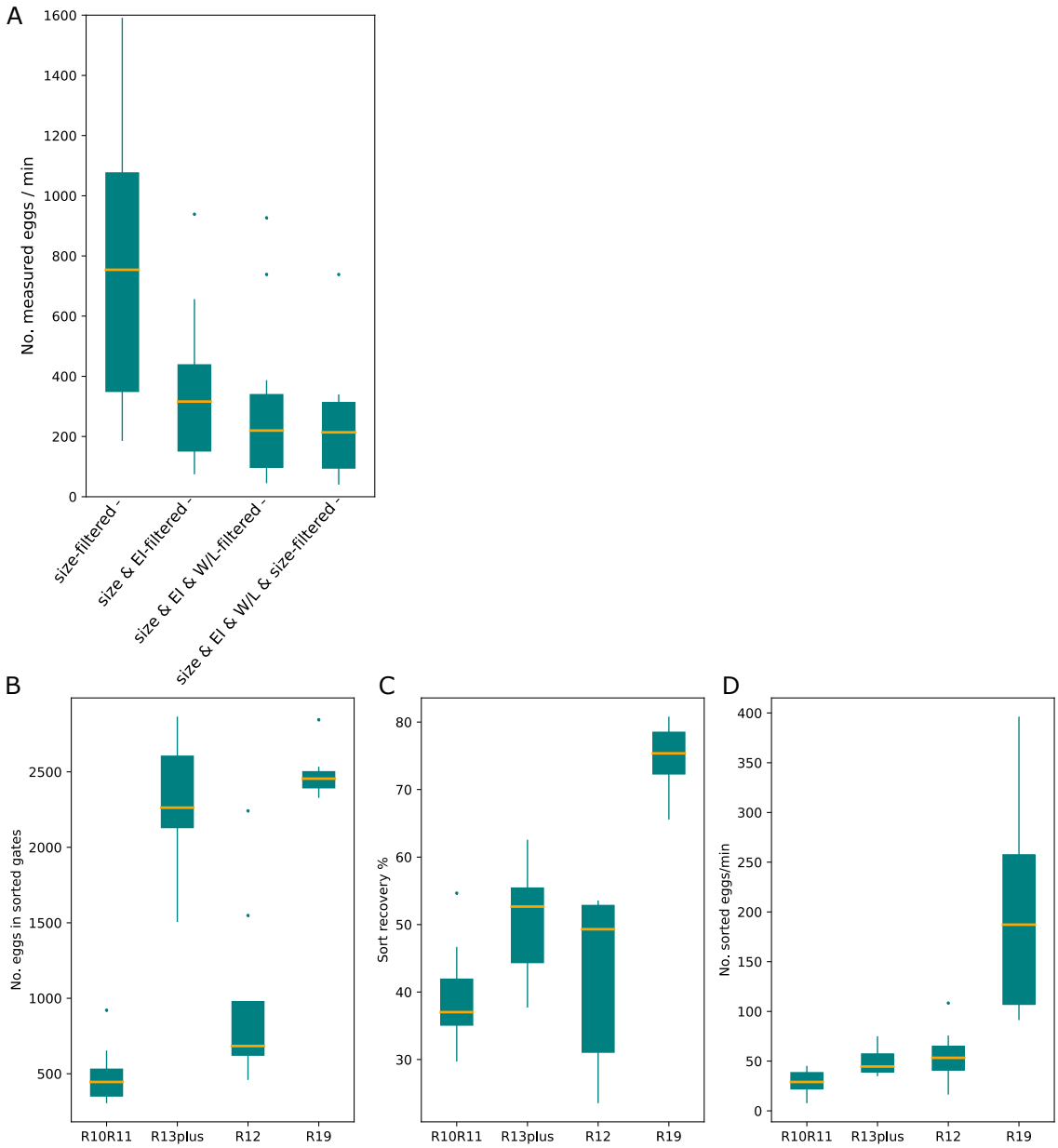

Fig. S4 Throughput and accuracy of egg size measurement and sorting. A) The speed of egg size measurement after implementing each filtering step decreases from average of 1681 eggs/min (size-filtered) to 238 eggs/min after filtering small and misaligned objects (sie&EI&W/L&size-filtered). Different filtering steps in A are the same as described in Fig. 1. The efficiency of sorting is computed as the sort recovery which is the percentage of objects that were sorted and dispensed as a fraction of all the objects fulfilling sorting criteria, i.e. size. The median sort recovery (C) and the speed of sorting (D) varies across gates of different size depending on number of available eggs in each gate (B). A total of 8 datasets have been used for each gate category: R10R11 (200-400 µm), R12 (400-500 µm), R13plus (> 500 µm) and R19 (200 to >500 µm).

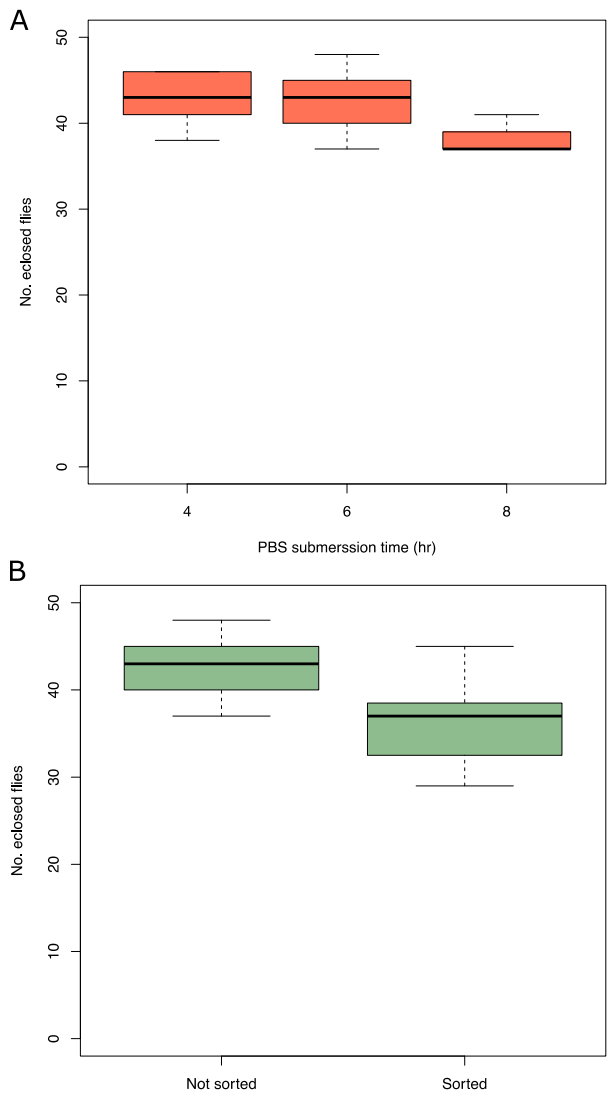

Fig. S5 The effect of handling and sort by large particle flow cytometer on egg viability. A) Egg-to-adult viability after 8 hours of PBS submersion (38.33±2.31 out of 50) is significantly less compared to 4 (42.8±3.42 out of 50) and 6 (42.5±3.48 out of 50) hours of submersion (generalized linear mixed model followed by Tukey's HSD test with correction for multiple testing, $P_{6hour\text{-}8hour} = 0.0355$, $P_{4hour\text{-}6hour} = 0.9715$, $P_{4hour\text{-}8hour} = 0.0638$). B) Passing through Biosorter® slightly (generalized linear mixed model $P < .0001$) reduces the egg-to-adult viability (36±4.64 out of 50) compared to the eggs that were only submerged in 1x PBS for the same duration of time (42.5±3.48 out of 50).

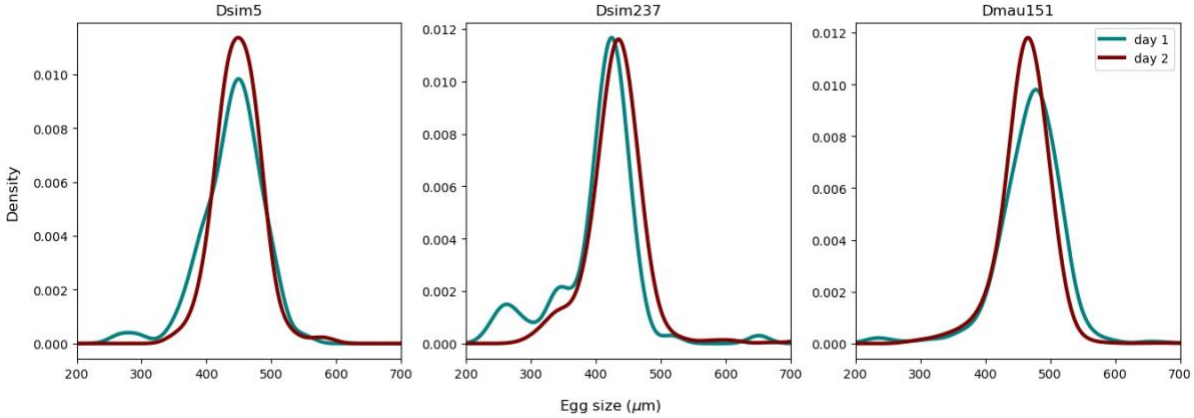

Fig. S6 The size distribution of eggs laid by females of different age, on two consecutive days, are highly similar. Thus, we pooled the data from these separate runs for each sample: D.sim5 and Dsim237 (*D. simulans*), D.mau151 (*D. mauritiana*). Each dataset is filtered as described in Fig. 1.

Table S1 The statistics of egg size distribution. Size threshold is used to remove 1st peak containing egg debris from the 2nd mode in the first filtering step. Dsim#, #: stock number (*D. simulans*), Lausanne5 (*D. melanogaster*), Dere01 (D. *erecta*), Dsan01 (D. *santomea*) and Dmau151 (*D. mauritiana*).

Sample	1st peak mode (µm)	size threshold (µm)	2nd peak mode (µm)
Dsim1	94.903	212.212	449.078
Dsim4	92.209	176.977	459.851
Dsim5	92.209	196.196	427.531
Dsim11	92.209	177.778	426.184
Dsim75	92.209	207.407	450.425
Dsim90	92.209	173.774	427.531
Dsim91	92.209	92.893	93.556
Dsim146	92.209	180.981	450.425
Dsim166	92.209	198.599	436.958
Dsim185	92.209	201.802	426.184
Dsim196	92.209	194.595	423.491
Dsim237	92.209	182.583	431.571
Lausanne5	92.209	146.547	420.798
Dere01	92.209	207.407	431.571
Dsan01	92.209	211.411	384.438
Dmau151	92.209	187.387	455.811

Table S2 The statistics of egg size distribution filtered using sample-specific threshold. Each dataset is filtered as described in Fig. 1. The 1st filtering step is based on the size threshold in Table S1. The filtered objects at this step are mostly egg debris. # nonegg: the number of small objects filtered in the first filtering stage. The mode, mean and median of these objects are Mode nonegg, Mean nonegg, and Median nonegg respectively. The second filtering step involved removing the misaligned eggs. The number, mean and median of these misaligned eggs are shown as #misaligned egg, Mean misaligned egg and Median misaligned egg. The last filtering step is to remove the objects in the first peak of distribution after removal of the misaligned eggs. Size threshold (μm) is used for this filtering step. #egg, Mode egg, Mean egg and Median egg are the number, mode, mean and median of the eggs in the final dataset. Dsim#, #: stock number (*D. simulans*), Lausanne5 (*D. melanogaster*), Dere01 (D. *erecta*), Dsan01 (D. *santomea*) and Dmau151 (*D. mauritiana*).

Sample	size threshold (μm)	#egg	#nonegg	#misaligned egg	Mode nonegg	Mean nonegg	Median nonegg	Mode egg	Mean egg	Median egg	Mean misaligned egg	Median misaligned egg
Dsim1	269.069	361	2710	849	94.903	110.546	102.983	440.998	438.457	442.345	382.667	391.171
Dsim4	283.483	1980	6707	4572	92.209	111.721	105.676	458.505	453.336	457.158	373.373	387.131
Dsim5	248.248	265	917	699	92.209	108.428	102.983	427.531	443.051	447.731	384.705	389.824
Dsim11	241.041	458	1530	1211	92.209	108.814	102.983	420.798	426.752	431.571	377.862	385.784
Dsim75	263.463	955	3678	2396	92.209	109.350	101.636	457.158	454.088	457.158	393.318	399.251
Dsim90	221.822	941	3075	2119	92.209	106.286	100.289	427.531	434.046	438.304	384.049	397.904
Dsim91	217.818	1913	1707	14055	92.209	92.209	92.209	438.304	416.989	426.184	128.336	102.983
Dsim146	256.256	1229	6138	2957	92.209	106.108	100.289	455.811	454.377	459.851	389.055	404.638
Dsim166	274.675	906	4005	2263	92.209	110.152	102.983	453.118	461.060	462.545	391.385	389.824
Dsim185	249.049	1535	12413	3878	92.209	110.019	101.636	443.691	437.533	440.998	376.696	385.784
Dsim196	298.699	1630	10442	6186	92.209	111.171	101.636	423.491	421.893	422.144	362.492	380.398
Dsim237	237.838	725	3422	2147	92.209	106.550	100.289	427.531	429.827	431.571	396.969	388.478
Lausanne5	163.363	689	4069	2401	92.209	107.065	102.983	420.798	363.069	404.638	312.924	315.757
Dere01	221.021	2056	1453	4757	92.209	114.704	104.329	431.571	433.829	434.264	392.111	403.291
Dsan01	317.117	250	2169	929	92.209	113.170	104.329	403.291	411.872	410.024	364.281	365.584
Dmau151	277.878	3079	15325	6214	92.209	108.169	101.636	466.585	464.235	465.238	381.830	388.478