PAPER

The iFly tracking system for an automated locomotor and behavioural analysis of *Drosophila melanogaster*

Kai J. Kohlhoff,[a] **Thomas R. Jahn,**[ab] **David A. Lomas,**[cd] **Christopher M. Dobson,**[a] **Damian C. Crowther**[bcd] **and Michele Vendruscolo***[a]

The use of animal models in medical research provides insights into molecular and cellular mechanisms of human disease, and helps identify and test novel therapeutic strategies. *Drosophila melanogaster*—the common fruit fly—is one of the most well-established model organisms, as its study can be performed more readily and with far less expense than for other model animal systems, such as mice, fish, or primates. In the case of fruit flies, standard assays are based on the analysis of longevity and basic locomotor functions. Here we present the iFly tracking system, which enables to increase the amount of quantitative information that can be extracted from these studies, and to reduce significantly the duration and costs associated with them. The iFly system uses a single camera to simultaneously track the trajectories of up to 20 individual flies with about 100 μm spatial and 33 ms temporal resolution. The statistical analysis of fly movements recorded with such accuracy makes it possible to perform a rapid and fully automated quantitative analysis of locomotor changes in response to a range of different stimuli. We anticipate that the iFly method will reduce very considerably the costs and the duration of the testing of genetic and pharmacological interventions in *Drosophila* models, including an earlier detection of behavioural changes and a large increase in throughput compared to current longevity and locomotor assays.

Introduction

Since *Drosophila melanogaster* has been for over a century one of the most commonly used animal models, a vast array of physiological, cytological and genetic tools is available for its study.[1–5] One particularly attractive feature of the use of *Drosophila* as a model organism is that it has a brain composed of functionally specialised cell types, which is capable of learning and memory, and underpins a range of complex behaviours. Furthermore, it has been established that more than two-thirds of genes associated with human disease have orthologs in *Drosophila*,[6,7] which is particularly relevant considering that biochemical pathways are conserved across eukaryotes.[6,8]

The detailed study of fruit flies allows the observation of the various developmental stages of this organism, and of the progression of abnormal behavioural patterns associated to disease. Measurements on various disease-related phenotypes are performed by a variety of behavioural, locomotor and cognitive assays.[9–24] Considerable advances have been made during the last ten years to improve the sensitivity and accuracy

[a] *Department of Chemistry, University of Cambridge, Cambridge, UK. E-mail: mv245@cam.ac.uk*
[b] *Department of Genetics, University of Cambridge, Cambridge, UK*
[c] *Department of Medicine, University of Cambridge, Cambridge, UK*
[d] *Cambridge Institute for Medical Research, Cambridge, UK*

Insight, innovation, integration

We introduce a highly quantitative method of analysing the motion of *Drosophila melanogaster* (the fruit fly), which is one of the most widely used animal models of human disease. Our approach, called iFly, is based on the fully automated combination of a three-dimensional tracking system with a computational analysis of the statistical properties of the trajectories of the fruit flies, which enables the rapid and accurate detection of their specific patterns of behaviour. Since the iFly method is designed to reduce very significantly the costs and the duration of assays to test genetic and pharmacological interventions in *Drosophila* models, while increasing their quantitative significance, we expect it to help increase even further the scope of *Drosophila* studies.

of these assays, with the goal of increasing the reliability of the assessment of the efficacy of genetic or pharmacological interventions. In one of the first locomotor assays that have been introduced to measure a phenotype in *Drosophila* models,[9] the fraction of a population of flies capable of climbing past a particular height in a tube within a set time was considered. This assay provides an overall performance index, but it does not give a detailed description of the features of the locomotor behaviour that the fly employed in climbing, or failing to climb, up the tube. More recently, computer vision technology has been employed to track flies both while moving along surfaces[18,20,21,23,25] and in flight.[10] These studies have allowed the investigation of questions such as the analysis of social behaviour[20,21] or the rate of gene expression with respect to fly activity using GFP fluorescence.[17] While the tracking of locomotion on a flat surface only requires a single camera,[18,20,21] three-dimensional tracking is commonly performed by the use of at least two cameras. For example, the tracking of fly locomotion over a cylindrical surface has been achieved using a three-camera approach and a sophisticated software to calibrate and synchronize the video streams.[18]

We here present a method, called iFly, for the three-dimensional tracking of fly trajectories that works with a single camera. This approach delivers three different time-synchronous perspectives by using two appropriately placed mirrors in place of additional cameras (Fig. 1). As an example of the application of the iFly method, we demonstrate that this approach greatly facilitates quantitative studies of the ageing process in *Drosophila*.

Materials and methods

Design of the tracking hardware

The iFly basic hardware is composed of an inexpensive off-the-shelf fixed-focus digital webcam linked to a standard Microsoft Windows personal computer by a USB interface. The tube of flies is placed in a chamber that is lit either through a frosted-plastic lid or by internal illumination. To enable three-dimensional tracking, two mirrors are placed at equal angles behind the tube of flies that allow side images of the flies to be captured by the camera (Fig. 1a). In the configuration used in this study, the mirrors were placed at an angle of δ = 39 degrees to the wall of the chamber. A webcam with a horizontal field of view of 50 degrees and VGA resolution was placed at 95 mm distance from the centre of the tube and at equal distance from the mirrors. The three captured images are thus taken from perspectives at close to 120 degree angles of one another. To allow placement on any lab bench and easy portability, the chamber is kept light and compact in size, taking up less area than a small laptop (Fig. 1b).

Design of the tracking software

The iFly software in its current implementation is a Java application that uses the Java Media Framework to capture individual frames either from a pre-recorded video file or a video stream from the webcam. An image segmentation algorithm is used for background subtraction and to identify pixels that can be assigned to flies. To rapidly cluster these pixels into fly representations, a multi-step hierarchical

clustering algorithm was developed. The first step consists in a row-based single-pass traversing of the image data and joining pairs of adjacent 'fly-pixels' into clusters. In the second step, spatial decomposition, these initial clusters are then assigned to quadratic tiles spread across the image with a side length that corresponds to the adjustable fly size in pixels. The third step is bottom-up hierarchical clustering with distances calculated only between pixels in adjacent tiles, thus greatly reducing computational cost. Finally, on-screen positions of flies are derived from clusters fulfilling minimum requirements such as the required number of pixels per fly (Fig. 1d). For simplicity, we treat the cluster centre, computed as the average x- and y-coordinates of all pixels assigned to a qualifying cluster, as the location of the fly's projection onto the screen.

A virtual representation of the fly chamber, camera, mirrors, and tube is generated as a module of the iFly software. Geometric parameters are defined in a configuration file and can easily be adjusted to fit a variety of different layouts for the fly chamber and objects therein. This virtual representation is used by an image registration algorithm to spatially integrate the three different perspectives captured by the camera. Triangulation is used to compute the fly positions as 3D coordinates. To achieve this result, all on-screen positions of flies are converted into rays by assuming that the image seen on screen is equivalent to the image projected onto the lens of the camera. Rays are then defined as running from the pixel-equivalent position on the lens of the camera to the camera's fixed focal point, as defined by its technical specifications. The rays are then traced back in the virtual representation, where they either intersect directly with the position of the fly in the virtual tube, or after reflection on one of the mirrors. To allow easy conversion, Cartesian coordinates in the virtual representation are defined to correspond in scale 1 : 1 to the physical world, with the focal point of the camera set as the origin. Depending on which part of the screen they map to, the rays are split into three different regions (regions I, II, and III in Fig. 1a, which are computed from the virtual geometry). The software then tests for intersections between all possible combinations of three rays, where each of the three rays is taken from a different region. Intersection in this case is defined as rays passing within a chosen cut-off distance of one another.

We used two parameters to control size and velocity of the tracked insects. The fly size is a user-defined parameter that defines the maximal intra-cluster distance between any two pixels in the individual clusters formed during hierarchical clustering. This parameter can be used to define an arbitrary size of the object to be tracked. In the case that not just the size of the objects tracked, but the layout of the chamber changes, a range of parameters such as the distances between camera and tube and between tube and mirrors, mirror angles, camera tilt, and fly tube radius are available and can easily be adjusted. Tracking of animals of different velocity is a parameter of the temporal resolution of the camera hardware used—in off-line mode iFly is capable of analyzing a pre-recorded video stream frame by frame regardless of the number of frames per second. The time-stamp (in milliseconds since frame 0) associated with each frame as provided by the
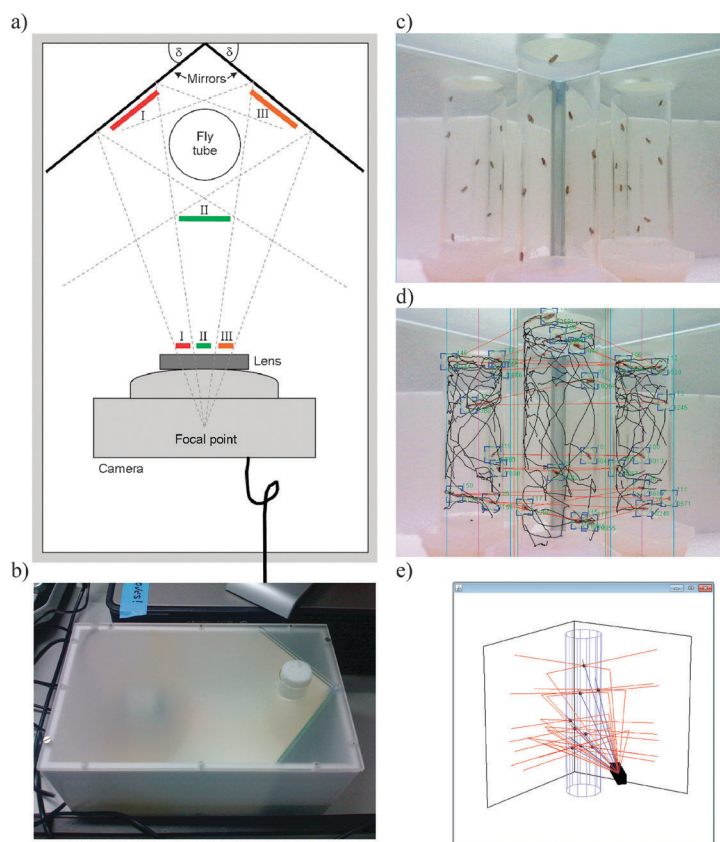
Fig. 1 **Scheme of the iFly apparatus**. (a) Design of the fly chamber with test tube, camera, and mirrors placed at an angle δ; the three perspectives (I, II, and III) are projected onto the camera lens. (b) Image of the fly chamber, showing how a test tube with flies can be inserted through a hole in the frosted plastic lid. (c) Snapshot from the graphical user interface (GUI) showing live video stream. The GUI contains components that allow the user to adjust parameters for the iFly chamber, camera, and image segmentation algorithm. (d) Example of a captured frame with results from the image segmentation algorithm overlaid; trajectories (black lines) and current fly positions (green circles and blue corners), from which the Cartesian coordinates of the trajectories of the flies are extracted, are shown together with red lines that connect fly images predicted to be projections originating from the same fly. (e) Real-time three-dimensional reconstruction of the ray tracing calculations for visual control during image analysis. Images reflected by two mirrors are triangulated with the direct image to accurately locate the flies in the tube; camera (black box), tube (blue cylinder), mirrors, and fly positions (black dots) are shown, with direct rays drawn in blue, and those reflected on mirrors in red.

Java Media Framework automatically translates into correct computation of all derived parameters including velocity.

Typical performance permits 10–15 frames per second (fps) for a group of 10 flies on a PC with a 2.0 GHz Intel Dual Core processor. Faster analysis is possible with fewer flies or a faster CPU. The software can be used in one of two modes: in real-time mode, which allows processing a live video stream at a frame rate determined by the speed of the hardware, or in offline-mode, which allows frame by frame processing of a previously acquired video file at a typical time resolution of 30 fps or above. This latter option allows the acquisition of detailed fly trajectories even on slow machines.

Calibration of the virtual representation to the physical world

To adjust for discrepancies between the ideal virtual representation of the fly chamber and the real-world geometry, as for example slight inaccuracies during manufacturing of the chamber, tilted mirrors, or inaccurate placement of the camera, parameters such as normal vectors to the mirror planes and distances between mirrors, tube, and camera are

optimized by a steepest descend algorithm as follows. First, either a short video of a well-occupied fly tube (*e.g.* with 10 flies) in its current configuration is taken, or a video from a previous experiment is selected. The user is then given the choice of a number of sample frames from the video stream. Based on the frame chosen, ideally one where flies occupy various parts of the tube, the software then iterates over computing fly positions and adjusting geometric parameters for a preset number of times. Visual clues are provided in the GUI to allow an observer to monitor the progress (Fig. 1e). A scoring function that is based on the minimum distance between traced rays, *i.e.* the accuracy with which flies are detected, serves to automatically decide which set of parameters to use for the next iteration. Initial parameters and parameters for the steepest descend algorithm are taken from the configuration file in the directory of the experiment. The set of optimized parameters is appended to the file and henceforth applies to all video files stored in that directory. The optimization can be refined by repetition with additional frames and only has to be performed once for a given experiment. Configuration files can be shared among

experiments unless the physical configuration of the fly chamber changes, *e.g.* by replacing the camera with a different model.

To reduce the occurrence of false positives in the detection of flies, the distance between passing rays in the virtual representation is subjected to a cut-off value. This maximal permitted distance is set by the user and was kept at 1 mm for this study. Furthermore, flies had to occur in at least five consecutive frames without having travelled more than 7 mm between any two consecutive frames to be included in velocity computations. This speed may sometimes be exceeded by a fly while flying of falling, but the use of a faster camera enables to reduce this effect to the required level.

Application to Drosophila senescence

To demonstrate the potential of the iFly method in assisting locomotor analysis of *Drosophila* we examined ageing flies at various stages of their lives. Sets of 10 flies were placed in test tubes of 1 cm radius and 10 cm height. Videos were taken on days 2, 9, 16, 23, 30, and 35. Test tubes were dropped into the iFly apparatus at time t_0 and videos recorded for 90s, with the tube briefly lifted and dropped back into the apparatus after $t_1 = t_0 + 30s$ and $t_2 = t_0 + 60s$. All videos were processed with the iFly software. A pre-processing step was added to the software to automatically detect t_0, t_1, and t_2 by analysing changes in image brightness levels caused by shadows cast by the hand of the experimenter. Dropping of the test tubes resulted in flies dropping to the bottom of the tube on impact and immediate initiation of upwards movement triggered by innate negative geotaxis. Three trajectories were recorded on hard disk for 15 s each in form of time-stamped Cartesian coordinates of the derived fly positions in space.
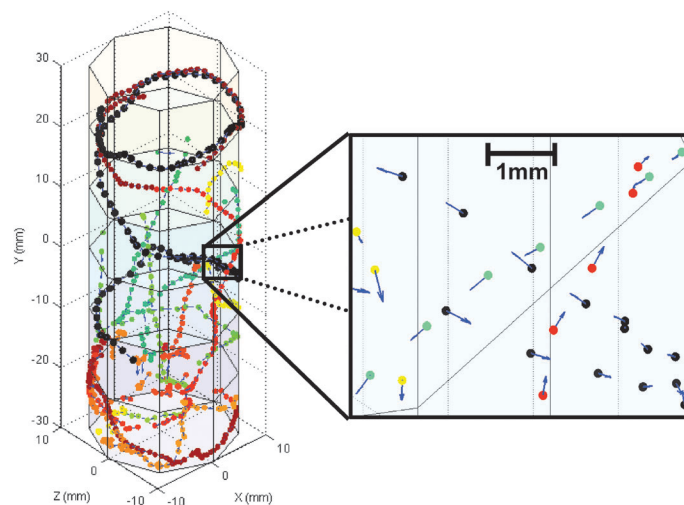
Reconstruction of fly trajectories

Fly trajectories were reconstructed in MatLab from the data provided by the iFly software (Fig. 2). The combined use of time-stamp information, Cartesian coordinates, and fly ID, which is provided by iFly as a first approximation based on the distance between fly locations in subsequent frames, allows a clear separation of location data into individual trajectories with high spatial resolution, as indicated in the magnified section in Fig. 2. From the individual locations and their trajectories over time, velocity vectors were interpolated and included in the plot as three-dimensional vectors. The resulting resolution allows detecting even minute directional changes and variations in the locomotion behaviour of each individual fly.

In the 3D tracking strategy implemented in the iFly method a particular attention is given to lost and gained fly tracks. The built-in iFly software is currently capable of merging effectively these tracks on the basis of the temporal and spatial proximity between end and start points of tracks. This approach fails rarely, for example when there is a sudden dropping of multiple flies to the bottom of the tube. For a more sophisticated analysis, iFly does not relying on this built-in module, but imports the recorded location data into MatLab, where analysis and plotting are facilitated by the already available functionality and new detection algorithms can be developed much more efficiently. In the velocity analysis presented in this work, interrupted fly tracks were treated as if originating from different flies and new trajectory IDs were assigned, which had no implication on the validity of the statistical analysis (velocity/time).

Statistical analysis of the trajectories

The Cartesian coordinates for the locomotion behaviour are analysed to extract statistical descriptors of the fly populations (Fig. 3). The statistical properties of these parameters are useful to discriminate flies at various stages of their lives with a very high level of confidence provided by the large amount of data made available through the automated data acquisition and analysis procedure.

Fig. 2 **Demonstration of the spatial resolution of the iFly method**. Trajectories are from healthy flies at day 1 after hatching and reconstructed from fly locations determined by the iFly software over a short 7.5 s video segment. The fly tube has a radius of 10 mm with flies traveling up to 55 mm vertically. A section is magnified to provide a close-up on various intersecting fly trajectories. Velocity vectors indicate direction and amplitude of fly locomotion. Fly locations are shown as small colored dots indicating clear separation and a sub-millimetre spatial resolution. For comparison, the length of a fly body is approximately 1.5–2 mm.
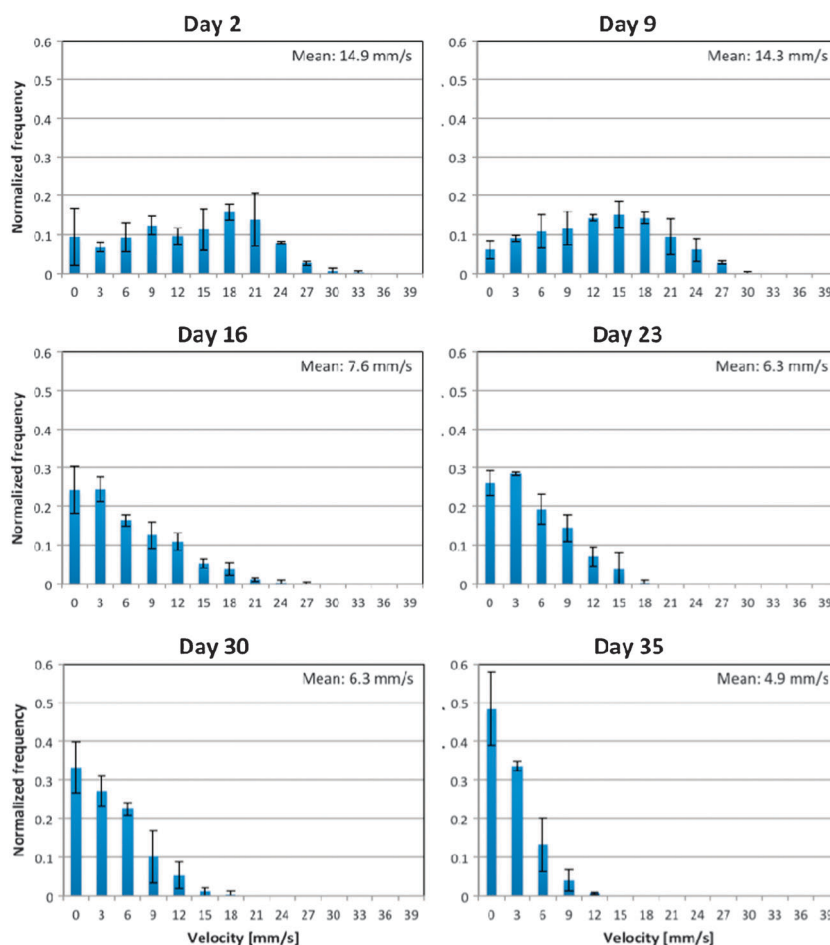
Fig. 3 Change in distribution of fly velocities over ageing. The six panels show velocity distributions for 10 flies in a test tube at different stages of their lives. At each measurement, the test tube was dropped three times during continuous video capture followed by analysis of the first 15 s of each of the three segments for a total of 45 s. Velocities were assigned to bins of width 3 mm s^{-1} with the first bin starting at 0 mm/s. A decline in motility in flies is quantitatively expressed by a progressive left shift of the distributions and a decline in the mean velocity. Error bars indicate the bin-wise standard deviation over three replicas of 10 flies each.

Discussion

There has recently been a great interest in the development of automated methods for detecting behavioural changes in *Drosophila* models in order to decrease the duration and costs of the current standard assays, which require the manual assessment of the number, speed and lifespan of flies.[9–24] One of the most important aspects of the iFly approach is that it enables the detection of subtle changes in phenotypes very early in the lifespan of flies, thus reducing the experimental time required for these studies by an order of magnitude, *i.e.* from weeks to just days. The panels in Fig. 3 indicate that clear changes in fly behaviour can be detected as early as in the first week, whilst lifespan data require about four weeks to be collected.

By restricting the flies to a confined space the method achieves high spatial resolution with a relatively low-resolution camera. Despite the reduction in camera resolution by dividing the field of view into three distinct regions we are still able to obtain sub-millimetre resolution as evidenced by contiguous trajectories as shown in Fig. 2. The current layout of the system makes it highly portable, with a simple setup consisting

in connecting the hardware to a suitable computer *via* the camera's USB cable and installing the software. The flexible design of the software also allows analysis of videos captured by cameras with high spatial and temporal resolution. By design, the software is suitable for use with a wide-variety of different fly chamber layouts, including much larger-scale environments necessary for the study of free-flying flies. While not required for this study, the tree graph created by the iFly software during hierarchical clustering with individual pixels represented as leaves, and flies identified at intermediate nodes, contains all the information necessary to compute more elaborate descriptors such as fly contours or visual hulls.

A major benefit of the iFly system over related approaches is its low cost and portability. Where typically multiple cameras and computers are employed to capture multiple perspectives, a single camera in combination with two mirrors achieves the same effect, using a single computer for storage and analysis of the video data. For high-throughput fly video analysis, the iFly software accepts command line arguments and can be controlled in batch mode without the need of additional user input and thus enables the use in a fully automated environment, which would enable data acquisition

over prolonged period of times, if required to answer specific questions about fly behaviour. While the current JMF-based version has a bottleneck in the extraction of individual pixels from a video stream, early tests with a version of the software using Visual C++ and Intel's OpenCV computer vision library have shown that a significant speed up is possible, to the point of real-time analysis of video streams at much higher temporal and spatial resolution. Our recently completed prototype of a Visual C++/OpenCV-based version of the software is capable of performing the most time-consuming task of the process, the grabbing of 640 by 480-pixel resolution video frames and extraction of pixel colors as byte arrays, at a rate of 100 frames per second on a low-end Intel E5200 Dual Core PC at 2.5 GHz, leading to a factor of five speed-up when compared to the original Java version. Furthermore, the iFly method allows quantitative comparisons to be made between experiments carried out at different times, as long as the original video and configuration files are retained. In particular, this method promises to reveal sophisticated manifestations of responses to tested novel drug compounds, and can be extended to a variety of organisms.

Conclusions

We have introduced the iFly method, which is designed to increase the sensitivity and throughput of *Drosophila* locomotor assays. We have demonstrated that this method allows a quantitative analysis of fly locomotion over time, which makes it possible to detect and describe behavioural changes typical of ageing flies. These results indicate that the iFly method has the potential to increase very considerably the accuracy and reproducibility of locomotor and behavioural *Drosophila* studies.

We expect that the availability of the iFly method will enable identification of novel behavioural patterns, through which it will be possible to cut down significantly the duration and cost of fly assays. This result can be achieved by replacing time consuming survival assays through detailed early-stage analysis with the iFly method. We thus anticipate that the automated and quantitative nature of the iFly method will facilitate large-scale screening studies of genetic or pharmacological interventions *Drosophila* models by enabling detection of locomotor and behavioural anomalies at early stages of development.

iFly availability

The iFly apparatus (hardware and software) is available from the authors upon request.

References

1 M. A. Ashburner, K. G. Golic and R. S. Hawley, *Drosophila: A Laboratory Handbook*, Cold Spring Harbor Laboratory Press, 2005.
2 M. D. Adams, S. E. Celniker, R. A. Holt, C. A. Evans and J. D. Gocayne, *et al.*, *Science*, 2000, **287**, 2185–2195.
3 E. Bier, *Nat. Rev. Genet.*, 2005, **6**, 9–23.
4 S. Tweedie, M. Ashburner, K. Falls, P. Leyland and P. McQuilton, *et al.*, *Nucleic Acids Res.*, 2009, **37**, D555–D559.
5 L. Gilboa and M. Boutros, *EMBO Rep.*, 2010, **11**, 724–726.
6 G. M. Rubin, M. D. Yandell, J. R. Wortman, G. L. G. Miklos and C. R. Nelson, *et al.*, *Science*, 2000, **287**, 2204–2215.
7 L. T. Reiter, L. Potocki, S. Chien, M. Gribskov and E. Bier, *Genome Res.*, 2001, **11**, 1114–1125.
8 M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein and H. Butler, *et al.*, *Nat. Genet.*, 2000, **25**, 25–29.
9 M. B. Feany and W. W. Bender, *Nature*, 2000, **404**, 394–398.
10 S. N. Fry, R. Sayaman and M. H. Dickinson, *Science*, 2003, **300**, 495–498.
11 N. M. Bonini and M. E. Fortini, *Annu. Rev. Neurosci.*, 2003, **26**, 627–656.
12 J. Bilen and N. M. Bonini, *Annu. Rev. Genet.*, 2005, **39**, 153–171.
13 B. W. Lu and H. Vogel, *Annu. Rev. Pathol.: Mech. Dis.*, 2009, **4**, 315–342.
14 T. E. C. Mackay and R. R. H. Anholt, *Annu. Rev. Genomics Hum. Genet.*, 2006, **7**, 339–367.
15 J. L. Marsh and L. M. Thompson, *Neuron*, 2006, **52**, 169–178.
16 A. Moloney, D. B. Sattelle, D. A. Lomas and D. C. Crowther, *Trends Biochem. Sci.*, 2010, **35**, 228–235.
17 D. Grover, J. S. Yang, S. Tavare and J. Tower, *BMC Biotechnol.*, 2008, **8**, 93.
18 D. Grover, J. Tower and S. Tavare, *J. R. Soc. Interface*, 2008, **5**, 1181–1191.
19 B. R. Bettencourt, B. W. Drohan, A. T. Ireland, M. Santhanam, M. B. Smrtic and E. M. Sullivan, *Behavior Genetics*, 2009, **39**, 306–320.
20 K. Branson, A. A. Robie, J. Bender, P. Perona and M. H. Dickinson, *Nat. Methods*, 2009, **6**, 451–457.
21 H. Dankert, L. M. Wang, E. D. Hoopfer, D. J. Anderson and P. Perona, *Nat. Methods*, 2009, **6**, 297–303.
22 D. Lessing and N. M. Bonini, *Nat. Rev. Genet.*, 2009, **10**, 359–370.
23 J. D. Seelig, M. E. Chiappe, G. K. Lott, A. Dutta, J. E. Osborne, M. B. Reiser and V. Jayaraman, *Nat. Methods*, 2010, **7**, 535–540.
24 G. Maimon, A. D. Straw and M. H. Dickinson, *Nat. Neurosci.*, 2010, **13**, 393–399.
25 J. C. Simon and M. H. Dickinson, *PLoS One*, 2010, **5**, e8793.