Continuous tracking of startled *Drosophila* as an alternative to the negative geotaxis climbing assay

Matthew J. Taylor and Richard I. Tuxworth

Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

**ABSTRACT**

The fruit fly, *Drosophila*, is commonly used to study late-onset neurodegenerative diseases due to the combination of powerful genetic tools, cheap and simple husbandry and short lifespan. One widely-used measure of disease progression is the age-dependent decline in motor performance that manifests in most *Drosophila* neurodegeneration models. This is usually quantified using a simple climbing assay. However, the standard climbing assay lacks sensitivity and suffers from high variability meaning large numbers of flies are needed or bespoke apparatus and software solutions. Here, we present a modification of the open-source, MATLAB-based, DART software to measure the decline in “startle response” with age. We demonstrate that the DART setup is more sensitive to the motor performance decline induced by adult-onset neuronal expression of amyloid beta (Aβ) peptides than a traditional climbing assay despite using smaller cohorts of flies. DART also has the potential to generate multiple metrics of motor behaviour during the startle response. The software requires no coding skills to operate and the required apparatus can be purchased commercially. Therefore, DART is a more useful method than the climbing assay for longitudinal assays of motor performance and will enable higher-throughput screen for genetic and pharmacological modifiers of neurodegeneration. In our proof-of-concept screen for modifiers of Aβ-dependent phenotypes, we identified that in *vivo* knock-down of p53 in adult neurons is neuroprotective. This supports recent work targeting p53 in *vivo* and demonstrates the potential for DART to be used to screen for targets that ameliorate neurodegeneration.

**Introduction**

The short lifespan of *Drosophila* coupled with powerful tools for genetic manipulation and genome-wide screening combine to make it an attractive model system to study aspects of human neurodegenerative disorders. *Drosophila* models have been particularly widely used for types of dementia, including Alzheimer’s disease and frontotemporal dementia, and for various movement disorders, including Parkinson’s disease and amyotrophic lateral sclerosis (reviewed in: Bouleau & Tricoire, 2015; Casci & Pandey, 2015; Hewitt & Whitworth, 2017; McGurk, Berson, & Bonini, 2015). In many studies, the progressive decline in neural output associated with the disease needs to be measured. While neurodegenerative disease-relevant phenotypes such as tests of memory (Saitoe, Horiuchi, Tamura, & Ito, 2011) or changes to circadian behaviour (Rosato & Kyriacou, 2006) can be quantified using flies, these assays are complex and ill-suited to large-scale screens. Consequently, the progressive decline in motor function is very commonly used as a measure of neural output. Expression of most of the aggregative proteins associated with human neurodegenerative diseases in the *Drosophila* CNS, including amyloid beta (Aβ) (Beharry, Alaniz, & Alonso, 2013), Tau (Kerr et al., 2011), TDP-43 (Voigt et al., 2010, p. 43), or expanded Htt (Romero et al., 2008), results in an accelerated decline in motor function with age. Where tested, this correlates with a decline in the efficiency of neural transmission (Kerr et al., 2011).

The method most commonly used for measuring motor output takes advantage of the innate negative geotaxis response displayed by *Drosophila* when startled. When tapped to the bottom of a vial, adult flies will climb back upwards towards the top. This escape reflex has been used for more than 3 decades by *Drosophila* researchers (Ganetzky & Flanagan, 1978) and is the basis of the Rapid Iterative Negative Geotaxis (RING) assay (Gargano, Martin, Bhandari, & Grotewiel, 2005) (more commonly referred to as the climbing assay). In this assay, groups of flies are transferred into empty vials that are then tapped so that the flies fall down to the base. The flies climbing back up the sides of the vial are recorded by camera and the number reaching a set distance above the base within a time limit is noted to generate a simple percentage. Alternatively, a performance index is calculated.

The climbing assay, as performed in most laboratories, is simple to perform but has some significant disadvantages. The assay requires relatively large numbers of flies –
commonly 100 flies are used per genotype or treatment – and there is considerable performance variability between vials of identical flies. There are several possible reasons for the variability, but variations in the “tapping” force applied to the vials and the binary scoring system are likely to be two of the more important. The assay also requires significant “hands-on” time, particularly if manual scoring is employed, and the simple scoring method also misses more nuanced behavioural metrics, especially the speed of movement. Various solutions to streamline the climbing assay and improve its reproducibility have been described (Kohlhoff et al., 2011; Liu et al., 2015; Podratz et al., 2013; Willenbrink et al., 2016). Each of these requires bespoke apparatus and software to track the movement of flies and are not designed for tracking movement over extended periods with the possibility of differential stimulation events during that period.

Here, we present an alternative method for monitoring the locomotor performance of Drosophila over time that addresses many of the issues with the climbing assay. We have adapted the MATLAB-based, open-source Drosophila ARousal Tracking (DART) system (Faville, Kottler, Goodhill, Shaw, & van Swinderen, 2015) – initially developed to study long-term circadian behaviour – to continuously monitor the horizontal movement of flies before, during and after a startle response. We have used the DART software suite to apply vibrational stimulus events multiple times over a 1 h period to elicit repeated “startle” reflexes in the flies. Both the baseline movement of flies and the increase in movement immediately after the stimulus can be quantified. Here, we have used an adult-onset Drosophila Alzheimer’s disease model in which aggregative Aβ 1–42 peptides are secreted from neurons (Speretta et al., 2012) to demonstrate the improved performance of the DART system in comparison to the climbing assay. We also highlight how the DART system can be used to test compounds or in genetic screens to identify modifiers of the Aβ-induced phenotypes.

Results

Setup of the DART system

Our use of the DART system setup was a slight modification from its original use for monitoring circadian behaviour (Faville et al., 2015). The setup is described fully in the methods but, briefly, flies are housed individually in Trikinetics activity vials. Prior to the assay, each vial is clamped horizontally to a platform in cohorts of 20. Movement is recorded from above both before, during and after the stimulation events in which the DART software delivers a train of short vibrational stimuli to small motors glued to the underside of each platform. Pilot experiments enabled us to settle on an experimental assay paradigm in which 4 groups of 20 flies were stimulated 5 times per assay. The stimulus comprised 5 × 0.5 s bursts of vibration over 3 s with each stimulus event separated by 10 min, a period long enough for fly movement to return to the baseline pre-stimulus speed. Post-recording, the DART software tracks the positions of each fly throughout the 1 h recording and quantifies non-stimulated movement and the response of the population to each of 5 stimulation events. An example of the paradigm with representative startle responses is shown in Supplementary Figure 2.

Drosophila Alzheimer’s model

To compare the DART system with the conventional climbing assay, we expressed a dimer of the human amyloid beta 1–42 peptide separated by a flexible 12-amino acid linker (tAβ1–42) (Speretta et al., 2012). The linker facilitates oligomerisation and increases toxicity and a N-terminal secretion motif ensures extracellular aggregation (Speretta et al., 2012). We expressed tAβ1–42 only in adult neurons through use of the standard Elav-Gal4C19S driver combined with the temperature-sensitive Gal80Ts inhibitor. Flies were reared at low temperature to prevent expression of tAβ1–42 and avoid confounding effects on neurodevelopment then shifted to higher temperatures once eclosed as adults.

Comparison of the climbing assay with DART tracking

Initially, we used a standard climbing assay to confirm that expression of tAβ1–42 only in adult neurons would produce an accelerated decline in motor performance. Flies were tested in cohorts of 20 and allowed to climb for 10 s after tapping to the base of the vial. A performance index was calculated using a binary cut-off for climbing success 3 cm above the base (see "Methods"). As expected, adult flies shifted to 29°C to maximise expression of tAβ1–42 displayed a very rapid decline in climbing ability and were essentially unable to climb above the 3 cm line after 10 days (Figure 1(a)); control flies performed significantly better. Flies reared at 25°C to reduce expression of tAβ1–42 maintained their climbing ability for significantly longer (Figure 1(a)). Interestingly, the control flies also performed substantially better when reared at 25°C compared to controls reared at 29°C (Figure 1(a)). This is potentially due to toxicity caused by expressing large quantities of GAL4 in neurons (Kramer & Staveley, 2003, p. 4; Rezával, Werbajh, & Ceriani, 2007) together with the effects of more rapid ageing when flies are reared at higher temperatures (Miquel, Lundgren, Bensch, & Atlan, 1976). However, we did not detect a significant difference in the gradient of the decline between the control and tAβ1–42 lines at 25°C.

We used the DART system to quantify the startle response in genetically identical flies reared simultaneously with those used for climbing assays. We plotted the startle response as a performance index (see “Methods”) and saw an age-dependent decline in response (Figure 1(b)). The day-to-day responses were more variable than for the climbing assay but showed the same pattern of responses at 29°C: flies expressing tAβ1–42 performed significantly worse than control flies; and control flies raised at 29°C performed significantly worse than control flies raised at 25°C. However, using DART, we were also able to see a significant difference in the gradient of decline between the control and tAβ1–42
flies at 25°C, which suggests that DART is more sensitive to subtle differences in phenotype than climbing assays.

Given the extreme toxicity of tAβ1–42 at 29°C, we used data from the flies reared at 25°C and asked whether the DART system was capable of extracting further movement information that might give further insight into the behaviour of the flies. We compared tAβ1–42 expressing and control flies at 0- and 19- days post-temperature shift to simulate a 2-point method that might be used in a simple genetic screen. On day 19, we can detect a significant decrease in the performance index of the tAβ1–42-expressing flies compared to the control. However, there is no difference in the mean walking speed between the groups (Figure 1(c)). This suggests a difference in response to a stimulus without there being a change to baseline locomotor activity. A difference such as this might not be observed using other systems, such as the Hillary Climber (Willenbrink et al., 2016), which rely on measuring the mean speed of flies. Therefore, DART has the potential to quantify different facets of motor behaviour during the startle response that are differentially affected by tAβ1–42-expression.

Consistency of the DART tracking system

To be useful as a screening platform, we needed to ensure that the DART system produces reproducible data across multiple experiments. We used a breeding paradigm to look at the effect of Aβ expression on the mature adult CNS after the critical period of neuronal plasticity had ended (Sachse et al., 2007; Sugie, Marchetti, & Tavosanis, 2018) and compared the performance of the tAβ1–42-expressing flies across 4 independent experiments. Adult flies were reared at 18°C for 1–2 weeks after eclosion and tracked 2–3 times to get a baseline measurement during this period. Then, the flies
were shifted to 27°C to induce expression of tAβ₁₋₄₂. As expected, at 27°C Aβ toxicity is intermediate to the 25 and 29°C phenotypes. We found that the slopes of the decline in motor performance were almost identical across all 4 experiments (Figure 2(a)), indicating that motor performance can be quantified reproducibly by the DART setup.

**Using DART as a screening tool: candidate neuroprotective compounds**

The consistency, reproducibility and speed at which data are acquired using DART suggests that the system could be used as part of a higher-throughput screen for pharmacological or genetic modifiers of the neurodegenerative phenotype than would be possible with climbing assays. We tested this idea initially with Congo Red, a diazo dye that has the ability to bind to Aβ monomers and inhibit fibril formation (Lorenzo & Yankner, 1994). It has previously been shown that feeding a Drosophila Aβ model Congo Red-enriched food reduces plaque formation and increases lifespan (Crowther et al., 2005). However, the authors in that study were unable to test whether motor performance was rescued because control flies raised on Congo Red-enriched food also displayed a significant motor deficit (Crowther et al., 2005). Since our assay does not rely on climbing ability and is able to detect more subtle differences in performance between groups, we used to DART to ask whether Congo Red can rescue the motor performance of the tAβ₁₋₄₂ flies. We found that while Congo Red does not have a negative effect on control flies, it also cannot rescue the motor performance decline of our tAβ₁₋₄₂ model (Figure 2(b)).

**Using DART as a screening tool: RNAi screen**

Next, we postulated that DART has potential as a genetic screening tool. As part of a wider project investigating the role of the DNA damage response (DDR) in mediating Aβ toxicity, we expressed UAS-RNAi to various components of the DDR alongside tAβ₁₋₄₂. As before, Gal80ts was used to restrict expression of tAβ₁₋₄₂ to adult neurons and after the critical period for plasticity. This also prevented the UAS-RNAi construct from being expressed and ensured no knockdown of the target gene expression during development or early adulthood.

A RNAi screen must be able to distinguish true positives from potential effects of expressing the RNAi construct alone. Expression of the UAS-RNAi alone should lead to one of three outcomes: no effect; enhancement of the starting phenotype by the RNAi; or suppression (or reversal) of the starting phenotype by the RNAi. We searched for each of these outcomes in our pilot screen. First, we tested spn-A (the homolog of Rad51), which is involved in homologous recombination (HR), where it binds to processed double-strand breaks (DSBs) (Sung & Klein, 2006). We observed that knockdown of spn-A in neurons has no effect on the decline in motor performance in control or tAβ₁₋₄₂ flies (Figure 3(a)). Importantly, these data show that simply adding an additional UAS-transgene does not affect Aβ toxicity by diluting out the Gal4.

---

Figure 2. Utilization of the DART setup for compound screening. (A) Comparison of the performance of the tAβ₁₋₄₂-expressing flies across 4 biological repeats of the same longitudinal experiment. No significant differences between the slopes were found (p = .3396). (B) Feeding flies Congo Red shows no significant improvement in motor performance of control or tAβ₁₋₄₂-expressing flies (p = .978 and p = .9709 respectively). All comparisons by one-way ANOVA with Tukey’s multiple comparisons test *p < .05.
Bre1 is involved in the recruitment of Rad6 (Wood et al., 2003, p. 1), which is required for post-replication repair through ubiquitination of PCNA (Hoege, Pfander, Moldovan, Pyrowolakis, & Jentsch, 2002). Knockdown of Bre1 had no effect on the motor performance decline of our tAβ1–42 model but expression of UAS-Bre1 RNAi alone was sufficient to produce a decline in motor performance comparable to tAβ1–42-expression alone (Figure 3(b)).

Both spn-A and Bre1 function in HR, which operates primarily in M- and S-phases of the cell cycle. Given post-mitotic neurons are in G0, HR is not likely to be a major mechanism of repair of DSB in adult neurons. Instead, repair is likely to be via non-homologous end joining (NHEJ), which does not require a homologous template for repair (Beucher et al., 2009). Next, we targeted Ku80, a component of the heterodimeric Ku complex (with Ku70) which

---

**Figure 3.** RNAi screen of DNA damage response components using DART to identify neuroprotective targets. UAS-RNAi to spn-A (A), Bre-1 (B), Ku80 (C) and p53 (D) was expressed in adult neurons alone or alongside tAβ1–42. (A) Knockdown of spn-A in neurons has no effect on the motor performance of control (p = .978) or tAβ1–42 flies (p = .872). (B) Knockdown of Bre-1 has no effect on the motor performance of tAβ1–42 flies (p = .878) but Bre-1 RNAi/+ flies perform significantly worse than controls (p = .0099). (C) Knockdown of Ku80 significantly attenuates the motor performance decline of the tAβ1–42 flies (p < .0001) but expression of the RNAi alone increases motor performance compared to control flies (p > .0001). (D) Neural-specific knockdown of p53 significantly attenuates the decline of the tAβ1–42 group (p = .0086). There is no significant difference between control and p53-RNAi groups (p = .5425). All comparisons by one-way ANOVA with Tukey’s multiple comparisons test. *p < .05; **p < .01; ***p < .001
threads onto the DNA at the site of the DSB in the first step of NHEJ (Pannunzio, Watanabe, & Lieber, 2018). Knockdown of Ku80 in neurons significantly attenuated the decline in motor performance of Aβ-expressing flies. However, expression of UAS-Ku80 RNAi alone dramatically increased the response of the control flies to the vibrational stimuli over time (Figure 3(c)). We eliminated each of these genes from consideration and concluded that our screening platform would be capable of identifying RNAi constructs truly suppressing the decline in Aβ-induced decline in motor performance.

Finally, we looked at the effects of knockdown of p53 in neurons. In response to DSBs, p53 is activated by ATM and mediates cell-cycle arrest or apoptosis (Amaral, Xavier, Steer, & Rodrigues, 2010). p53 is upregulated in a number of neurodegenerative disorders, including Alzheimer’s disease (Cenini, Sultana, Memo, & Butterfield, 2008), but it is not clear whether this is a cause or consequence of pathology (Szybińska & Leśniak, 2017). Knockdown of p53 in microglia prevents Aβ-induced activation and the neurotoxic effects of microglial apoptosis (Davenport, Sevastou, Hooper, & Pocock, 2010). Consistent with this, in our screen knockdown of p53 in adult Drosophila neurons significantly reduced the Aβ-induced decline in motor performance whilst having no effect on the control (Figure 3(d)).

Thus, we have shown that DART can be used as a rapid, reproducible and sensitive motor performance assay to identify age-dependent phenotypes with the potential to be used for large-scale screening of pharmacological or genetic modifiers of neurodegeneration.

Discussion

One advantage of using Drosophila as an animal model of neurodegenerative disorders is the disease course is much shorter than with vertebrate models. A widely-used method to follow the disease course in Drosophila neurodegeneration models is to measure the age-dependent decline in motor function that manifests in almost every case (Bouleau & Tricmoire, 2015; Casci & Pandey, 2015; Hewitt & Whitworth, 2017; McGurk et al., 2015). By far the most commonly used assay for assessing motor function is the negative geotaxis climbing assay, first described in the 1970s (Ganetzky & Flanagan, 1978). However, this requires large numbers of flies and data points and extensive hands-on experiment time. Solutions to automate the assay to increase throughput or to reduce variability have generally involved bespoke engineering and/or software (e.g. Kohlhoff et al., 2011; Liu et al., 2015; Podratz et al., 2013; Willenbrink et al., 2016).

Here, we have presented an alternative setup which does not rely on climbing but instead on startling the flies in horizontal vials through vibration. The setup is based on the open-source, MATLAB-based DART software package (Faville et al., 2015) which is GUI-based and does not require coding skills to use. The apparatus is relatively inexpensive and commercially available, avoiding the need for bespoke engineering. The setup requires far fewer flies than the climbing assay and requires only a small number of recordings (~10) to give a high power to detect differences in motor performance between groups. In addition, the software controls the delivery of vibrational stimuli to the flies and these can be varied in duration and intensity to suit different experimental paradigms. We have demonstrated in this study that the DART system produces reproducible data with a well-characterised Drosophila neurodegeneration model using cohorts of 20 flies only; that it has greater sensitivity to detect differences between groups than the standard climbing assay; and that it has the potential to be used in large-scale screens for modifiers of neurodegeneration.

One of the key advantages of the DART setup over climbing assays is the reduction of the experimenter hands-on time. We have routinely run assays with only 20 flies per genotype vs. 100 flies in our climbing assays. Given most experiments require age-matched flies of multiple genotypes, using DART results in a large reduction in fly handling at each stage of the breeding process. Moreover, the food in the mini-vials used in this system needs to be replaced only on a fortnightly basis, provided that the incubator is maintained at ~70% humidity. In our case, the duration of each assay was 1 h, but since the vibrational stimuli are controlled by the software, there is no need for the experimenter to be present (unlike for the manual tapping of vials in the standard climbing assays). Furthermore, tracking of the fly positions and quantification of locomotion behaviours are handled by the DART software, in contrast to the climbing assay, which is routinely scored manually by the experimenter from video files.

A weakness of the DART setup is the day-to-day variation in the startle response of the flies. We saw variable responses for genetically identical flies, housed identically. The reasons for this are not clear. However, we were able to reduce the effect of this variation by recording the startle responses of the flies on 2–3 occasions in the period post-emergence but before the shift to a higher temperature i.e. before induction of tAβ;42 expression. This provided a baseline response for each cohort used for normalization and generation of a PI (see “Methods”). Regression lines were then fitted to the longitudinal data for comparison between groups. This method produces highly reproducible data with only 20 flies per genotype (Figure 2(a)) but does mean the DART setup is more suitable for measuring longitudinal performance, such as the age-dependent decline in motility in neurodegeneration, rather than to compare motility in one-off, single time point experiments. Other metrics quantified by DART may prove useful measures in the future, such as pre-stimulus speed, which cannot be measured by the RING assay (Gargano et al., 2005), or sleep duration. The shape of the response curves of cohorts of flies may also provide information about the motor performance to accompany the amplitude of the response: we have seen examples where two genotypes decline at a similar rate but the flies of one cohort fail to respond coherently to the stimulation at older ages while the controls do still respond similarly within the group; this data manifests itself as a difference in goodness-of-fit to the exponential model DART uses to calculate mean amplitude of response. We have also
been able to demonstrate significant differences in this response to vibration without accompanying changes to the mean walking speed over the duration of the experiment (Figure 1(c)). This would not be possible to observe in a setup such as iFly, that calculates average velocity in unstimulated conditions (Kohlhoff et al., 2011) or with the Trikinetics DAM activity monitors used in a previous study of the neuroprotective effects of curcumin on Aβ toxicity (Caesar, Jonson, Nilsson, Thor, & Hammarström, 2012).

During our testing of the application of DART as a screening tool for modifiers of the neurodegeneration phenotype, we identified a neuroprotective effect of knocking down p53 expression in neurons. p53 is known to have a pro-apoptotic role: its overexpression leads to widespread apoptosis in cultured hippocampal neurons (Jordán et al., 1997). In addition, under neurodegenerative conditions, p53 increases oxidative stress by activation of pro-oxidant genes and repression of antioxidant genes (Chatoo, Abdouh, & Bernier, 2010). There is growing evidence for cooperation and repression of antioxidant genes (Chatoo, Abdouh, & Bernier, 2010). There is growing evidence for cooperation and repression of antioxidant genes (Chatoo, Abdouh, & Bernier, 2010).

Therefore, p53 is a potential target for treating various neurodegenerative disorders (Chang et al., 2012; Culmsee & Mattson, 2005), but little work has been done in vivo. Our results demonstrate that knockdown of p53 in neurons in vivo reduces the decline of motor function in our tAβ1-42 model, strengthening the case for future work targeting p53 in neurons to ameliorate pathology.

In conclusion, we have demonstrated a novel method for quantifying the progression of motor performance decline in Drosophila neurodegeneration models that is reproducible, sensitive and rapid compared to other assays and which reduces hands-on time considerably. We have highlighted how it is best suited to tracking longitudinal performance such as the age-dependent decline of motility in neurodegeneration, and how different aspects of motor performance can be affected differently in these models. Finally, we have shown the utility of DART as a high-throughput screen for modifiers of neurodegeneration, with our screen providing in vivo evidence that p53 may be an effective target for neuroprotection in Alzheimer’s disease.

**Methods**

**Drosophila husbandry**

Fly crosses were maintained on a standard yeast–sugar–agar food mix (50 g/L yeast, 50 g/L glucose, 0.8% agar, 1% soy flour). Crosses were maintained at 18°C in all cases and shifted to 25, 27 or 29°C 1–2 days post-eclosion. Relative humidity was kept constant at 70%.

**Drosophila genetics**

The driver line used in all experiments was the pan-neuronal driver Elav-GAL4C155 combined with a tubP-GAL80ts insertion on chromosome II. The temperature-sensitive GAL80 represses GAL4 activity at 18°C, allowing us to limit neuronal Aβ expression to adulthood. The control line used was an isogenic w1118 (BL5905). Virgin females of the driver lines were crossed to UAS-α42 (12-linker) males (Speretta et al., 2012), or to w1118/Y control males. Drosophila lines were supplied by the Bloomington stock centre except for the UAS-α1-42 which was a kind gift of Dr Damien Crowther (University of Cambridge).

**Climbing assays**

For each genotype, flies were separated into 10 cohorts of 10 flies in vials and shifted to either 25 or 29°C after an initial “Day 0” climbing assay. 5 empty vials were marked with a line 3 cm from the base to be used as the climbing apparatus. Flies were allowed 1 h to acclimatise to laboratory temperature and humidity, before being transferred into the test vials. Climbing was tested within the same 2 h time window each time. For testing, flies were tapped to the bottom of the vial, and the number climbing above the line after 10 s recorded using a Logitech C920 HD webcam at 480p. The tap was repeated twice more after a 20 s interval. Flies climbing above the lines was determined manually from recordings. During the analysis, results from the 2nd and 3rd taps only were used since climbing performance was inconsistent on the first tap. Finally, flies were returned to food-containing vials after each experiment. The food was changed once per week.

**DART movement tracking**

**Fly handling**

1–2 days post-eclosion, male flies were collected under CO2 anaesthesia, sorted for genotype and 20 flies transferred to individual 5 × 65 mm Trikinetics monitor tubes (Trikinetics.com) with a small amount of food inserted at one end. The food end was covered with a small rubber cap to prevent drying out and cotton wool was used to seal the other end. Tubes were housed at constant temperature and 70% rel. humidity. Tubes were moved to the testing room 1 h before the start to acclimatise the flies and replaced in the incubator after testing. Movement was tested within the same 2 h time window each day.

**Equipment setup**

The equipment setup was essentially as described (Faville et al., 2015) with minor modifications. Briefly, tubes were housed on horizontal platforms supplied by BFKlabs (www.bfklab.com) with each tube held in place by clips. The platforms were immobilised on the base of a photography copy stand with a Logitech webcam positioned above. The camera was set to record at a resolution of 960 × 540 at 5 fps. Each platform had two coin motors attached to the underside to deliver vibrational stimuli under the control of the DART software. The apparatus was housed in a room with constant temperature and a stable lighting setup arranged to minimise intensity variation across the apparatus and to eliminate glare or shadows on the glass vials. An image of the setup is included in Supplementary Figure S1. The DART software ran under MATLAB (2017a) on a standard PC.
Experimental paradigm

The experimental paradigm was controlled by DART. It comprised recording for 5 min to establish baseline speed of movement followed by a 3.5 V (2.4 × g) vibrational stimulus event delivered in 5 × 0.5 s bursts with 0.1 s between each. The stimulus events were repeated a further four times with 10 min between events. After the final stimulus, the recording was continued for a further 15 min. The total recording time was 1 h. The experimental paradigm, an example dataset and diagrammatic explanation of the performance index calculation is shown in Supplementary Figure 2.

Movement tracking

Within the DART suite, each video was loaded into the Fly Position Tracking program and each platform identified and individual vials delineated. The background image was calculated and then the position of each fly within its vial determined at 5 Hz.

Quantification of movement

The mean walking speed of the cohort was calculated within the DART Quantification program for each tracking position i.e. at 5 Hz. Pre-stimulation speed was set as the mean speed for the 120 s prior to delivery of the vibration. The response of each fly within the cohort to the stimulus was aggregated to calculate the maximum speed of the population and hence the amplitude of the startle response for the cohort for each of the five stimulation events. Speed calculations were exported from DART as .csv files and processed within Excel and Prism.

Statistics and analysis

All linear regression and statistical tests were performed in GraphPad Prism 7.

For the climbing assays, the mean proportion of flies climbing above the line in the 2nd and 3rd repeats were calculated and tabulated in Excel as a performance index (PI), where:

\[ \text{PI} = \frac{\text{number of flies crossing the line}}{\text{total number of flies}}. \]

Flies touching the line were adjudged to have crossed it. The PI, SD and n of each genotype for each day was transferred into an XY graph in Prism.

To calculate a PI for the DART data, the mean amplitude of response (defined at the maximum speed post-vibration minus the pre-stimulus speed) on the days prior to the temperature shift was calculated and used as the baseline (such that the "day 0" PI = 1). All subsequent recordings were then normalised to this baseline in Excel and then transferred into Prism (see Supplementary Figure S2).

For all experiments, linear regression lines were fit to the PI data with the line constrained to go through the point \( X = 0, Y = 1 \). The values of the slope, associated standard error and N were then compared with ordinary one-way ANOVA with Tukey’s multiple comparisons test. The comparison of the day 0 and day 19 PI and mean speed was performed using two-way ANOVA with Sidak’s multiple comparisons test.

Power calculations

All power calculations were performed in G*Power. We calculated the sample size required \textit{a priori} based on the effect size and standard deviation determined from preliminary experiments. If we use DART to compare the PI of two different groups on any given day, a difference in PI of 0.5 can be detected at 90% power at \( \alpha = 0.05 \) with \( n = 5 \) stimulations per group. Effect size in this case = difference/SD = 0.5 mm s\(^{-1}\)/0.2 mm s\(^{-1}\) = 2.5.

If we use DART to compare slopes of regression lines (here \( n = \) number of days of recording), a representative effect size of 1.37 can be detected at 90% power at \( \alpha = 0.05 \) with \( n = 10 \) stimulations per group (equivalent to 2–3 times per week over the course of an experiment).

Acknowledgments

The authors thank Dr Damien Crowther and the Bloomington Stock centre for providing \textit{Drosophila} lines and Dr Niki Anthey for comments on the manuscript. We thank Drs Benjamin Kottler and Richard Faville for help in the setting up of the DART system and troubleshooting the software.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

MJT is funded by UK Biotechnology and Biological Sciences Research Council award BB/N008472/1 to RIT.

ORCID

Matthew J. Taylor  http://orcid.org/0000-0003-3055-4453
Richard I. Tuxworth  http://orcid.org/0000-0001-5697-6254

References

Amaral, J. D., Xavier, J. M., Steer, C. J., & Rodrigues, C. M. (2010). The role of p53 in apoptosis. \textit{Discovery Medicine}, 9, 145–152.

Beharry, C., Alaniz, M. E., & Alonso, A. d. C. (2013). Expression of Alzheimer-like pathological human tau induces a behavioral motor and olfactory learning deficit in \textit{Drosophila} melanogaster. \textit{Journal of Alzheimer’s Disease}, 37, 539–550. doi:10.3233/JAD-130617

Beucher, A., Birraux, J., Tchouandong, L., Barton, O., Shibata, A., Conrad, S., ... Löbrich, M. (2009). ATM and artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. \textit{The EMBO Journal}, 28, 3413–3427. doi:10.1038/emboj.2009.276

Bouleau, S., & Tricoire, H. (2015). Drosophila models of Alzheimer’s disease: Advances, limits, and perspectives. \textit{Journal of Alzheimer’s Disease}, 45, 1015–1038. doi:10.3233/JAD-142802

Caesar, I., Jonson, M., Nilsson, K. P. R., Thor, S., & Hammarström, P. (2012). Curcumin promotes A-beta fibrillation and reduces neurotoxicity in transgenic Drosophila. \textit{PLoS One}, 7, e31424. doi:10.1371/journal.pone.0031424
