Animal Behaviour 215 (2024) 177–190

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Species differences in learning about gustatory and visual stimuli in two recently diverged species of *Drosophila*

Madeline P. Burns [a, b, *], Julia B. Saltz [a]

[a] Department of BioSciences, Rice University, Houston, TX, U.S.A.
[b] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, U.S.A.

ARTICLE INFO

Article history:
Received 7 December 2023
Initial acceptance 19 February 2024
Final acceptance 6 May 2024

MS. number: A23-00637R

Keywords:
context generality
Drosophila
evolution
genotypic variation
learning

Learning is central to our understanding of how behaviour is shaped by the environment. A key open question is whether learning across contexts evolves as an integrated process, or whether learning in each context is free to evolve separately. Here, we measured learning in two sensory contexts in multiple genotypes and both sexes of two closely related, but ecologically divergent, species of fruit flies, *Drosophila simulans* and *Drosophila sechellia*. These species are morphologically very similar but differ dramatically in ecology and population biology. We tested how flies from each genotype, sex and species responded to and learned about different gustatory and visual cues. This approach allowed us to test whether species differences in learning were independent or correlated across contexts. Surprisingly, we found no evidence that *D. simulans* learned in any of our treatments. In contrast, *D. sechellia* learned to avoid gustatory stimuli that were paired with an aversive stimulus, as predicted, but unexpectedly learned to approach visual stimuli that were paired with the aversive stimulus. At the genotype level, genotypes, but not species, differed in their naïve responses to stimuli, but genotypes did not differ in learning in either species. Our results demonstrate that *D. sechellia* indeed differs from *D. simulans* in both learning contexts, but in a stimulus-dependent way. We suggest that studies of additional species or population pairs that employ this framework will be critical for evaluating the dimensionality of learning and its evolution.

Learning is a form of phenotypic plasticity involved in many important behaviours impacting fitness, such as foraging, mate choice, predator avoidance and more (Shettleworth, 2001). While previous work has shown that learning can increase fitness (Dukas, 2005; Richardson et al., 2003; Snell-Rood, 2013), learning can also be costly (Burger et al., 2008; DeWitt et al., 1998; Dukas, 1999; Dunlap & Stephens, 2009; Kotrschal et al., 2013; Laughlin et al., 1998; Mery & Kawecki, 2003, 2004a, 2004b; Stephens, 1991). As such, there has been a growing interest in understanding how learning evolves, and in particular, understanding the evolutionary constraints or limits that may exist (Auld et al., 2009; Buchanan et al., 2013; Dingemanse & Wolf, 2013; Dunlap & Stephens, 2009; Johnson et al., 2013; Mery & Kawecki, 2002; Pravosudov & Clayton, 2002; Tello-Ramos et al., 2019).

Learning is often defined as lasting changes in behaviour resulting from informative prior experience (Shettleworth, 1993; Stephens, 1991; Thompson, 1986). These definitions usually do not mention what kinds of information can or cannot be learned by a particular individual at a particular time. While this level of abstraction is necessary to study learning across contexts and species, it may also lead to the implicit assumption that learning represents a generalized ability that can be deployed across contexts, a phenomenon we will refer to here as 'context generality', rather than a unique response to particular stimuli in particular contexts (Dingemanse & Wolf, 2013; Frost et al., 2015; Morand-Ferron et al., 2016). Many studies have found support for context generality; in particular, positive among-individual correlations in performance on various kinds of cognitive tasks have been discovered in a wide range of animals (Ashton et al., 2018; Galsworthy et al., 2002; Matzel et al., 2003, 2020; Plomin, 1999, 2001; Plomin & Spinath, 2002; Prentice et al., 2022). However, not all theory and data support context generality (e.g. DuBois et al., 2018; Ellis et al., 2024; Keagy et al., 2011; Tello-Ramos et al., 2019). For example, foundational research on learning has

* Correspondence and present address: Huffington Center on Aging, Baylor College of Medicine, Houston, TX, U.S.A.
E-mail addresses: Madeline.burns@bcm.edu (M. P. Burns), Julia.b.saltz@rice.edu (J. B. Saltz).

https://doi.org/10.1016/j.anbehav.2024.07.009

demonstrated that specific combinations of conditioned and unconditioned stimuli may be particularly easy or difficult for animals to associate (Dunlap & Stephens, 2014; Garcia & Koelling, 1966; Shettleworth, 1972). Additional theory has focused on the potential for trade-offs among behavioural plasticities (Sih & Del Giudice, 2012). Taken together, theory and data indicate that context generality is not universal and highlight key examples of the opposite phenomenon, which we will call 'context specificity' (Shettleworth, 1993, 2012).

When considering evolutionary changes, context generality and context specificity represent fundamentally different ideas about how learning is expected to evolve. Under context generality, theory predicts that evolutionary divergence in learning in one context should be associated with correlated changes in one or more other contexts (Dingemanse & Wolf, 2013; Lande & Arnold, 1983; Plomin, 1999, 2001). However, under context specificity, learning in each context would be expected to evolve independently. Distinguishing between these hypotheses is thus important for understanding learning and its ecological and evolutionary implications, such as the role learning may play in adaptation to novel environments (Ghalambor et al., 2007; Pfenning et al., 2010; Sih, 2013; Snell-Rood, 2013; Verzijden et al., 2012).

Understanding context generality or context specificity and its role in the evolution of learning, if any, requires measurement of genetic variation in learning within contexts and estimation of genetic covariation for learning across contexts, both within a generation and over longer evolutionary timescales. Indeed, it remains unclear whether variation in learning (within or across contexts) has a genetic basis in many cases (Branch et al., 2022; Croston et al., 2015). Furthermore, genetic correlations themselves may change over generations, meaning that evolutionary divergence cannot always be predicted from standing genetic (co-) variation (Lande & Arnold, 1983; Saltz, Hessel et al., 2017, Saltz, Lymer et al., 2017). Comparing learning between species and genotypes under controlled conditions may be particularly important for studying the context generality of learning (Lambert & Guillette, 2021), as learning itself may provide increased access to resources, potentially leading to positive correlations across learning contexts when measuring the same individual in multiple learning assays (Bell, 2012; Cauchoix et al., 2018; Morand-Ferron et al., 2016).

One promising approach to studying genetic variation across multiple contexts and evolutionary timescales is to study population genetic differences and species differences simultaneously. Here, we focus on a recently diverged pair of fruit fly species, *Drosophila sechellia* and *Drosophila simulans*. These species diverged only 250 000 to 413 000 years ago and are morphologically very similar (Garrigan et al., 2012; Kliman et al., 2000; Schrider et al., 2018), but differ markedly in their diets, ecological characteristics, demography and behaviour. Most prominently, *D. simulans* is a broadly distributed cosmopolitan species that inhabits a wide range of habitat (i.e. fruit) types (Behrman et al., 2015; Machado et al., 2015; R'Kha et al., 1990), while *D. sechellia* is a specialist restricted to the Seychelles archipelago and found to live almost exclusively on the toxic fruit *Morinda citrifolia* 'noni' fruit (Jones, 2005; Lachaise et al., 1986; Matute & Ayroles, 2014; R'Kha et al., 1990; Schrider et al., 2018).

We studied associative conditioning for two different sensory modalities, gustatory and visual, on a panel of inbred isofemale lines (hereafter 'genotypes') from *D. sechellia* and *D. simulans*. First, we confirmed that all genotypes avoid quinine, an aversive bitter tastant that has long been used to study learning in *Drosophila melanogaster* (see Appendix). To measure learning, we tested independent flies from the same genotypes with pairs of either gustatory stimuli (imitation noni food and plain food) or visual stimuli (stripes or zigzags) and measured which stimulus each fly preferred (i.e. spent more time on) before training. Next, we exposed the same individual flies to a training experience in which one of the stimuli was paired with quinine. Finally, we measured the fly's preferences again, in the absence of quinine, allowing us to measure how each individual's pretraining preference was modified by the training experience. Changes to a fly's response to a stimulus following the training stage would indicate learning.

We fitted hierarchical mixed models in a Bayesian framework that allowed us to evaluate species and genotype differences in behaviour across stimulus types. First, we investigated species differences and other differences in flies' pretraining preferences; since learning involves the modification of these pretraining preferences, this step was necessary to interpret the resulting learning scores. Next, we modelled the effects of species, genotype and stimulus type on learning scores to test the following three hypotheses: (1) species and genotypes differ in learning about visual and gustatory stimuli; (2) learning scores covary across stimulus types, i.e. one species is better at both visual and gustatory learning; (3) species differences reflect genetic covariation within species.

Species divergence in context generality learning would be evident if one species shows greater learning performance for both gustatory and visual stimuli, supporting hypothesis (2). Alternatively, if divergence in learning is limited to specific stimuli, then we would expect to observe greater differences in learning performance for the gustatory stimuli than the visual stimuli, or vice versa. To our knowledge, our study is one of the first to investigate divergence in learning over multiple evolutionary timescales and across contexts.

EXPERIMENTAL METHODS

Study System Background

As mentioned above, *D. sechellia* and *D. simulans* are recently diverged and are morphologically very similar (Garrigan et al., 2012; Kliman et al., 2000; Schrider et al., 2018), but differ markedly in important aspects of their biology. Notably, *D. sechellia* prefers noni fruit over other fruits, whereas *D. simulans* shows the opposite behaviour (Burns et al., 2020; Dworkin & Jones, 2009).

Although it is outside the scope of the current study to identify the evolutionary mechanisms producing differences in learning in these species, if any, it is reasonable to hypothesize that such dramatic differences in the species' ecologies (Kalan et al., 2020; Roth et al., 2010; Schuck-Paim et al., 2008; Sol et al., 2005) and in demographic factors may be accompanied by adaptive or nonadaptive divergence in learning (Hoedjes et al., 2011; Poolman Simons et al., 1992; Ratcliffe et al., 2003). Indeed, dietary differences are thought to be particularly important to individual and species differences in learning (Arien et al., 2018; Cordner & Tamashiro, 2015; Messier et al., 2007; Molteni et al., 2002).

Drosophila sechellia and *D. simulans* offer important experimental tools for testing species differences and genetic correlations for learning across contexts. First, learning in the closely related *D. melanogaster* has been studied for decades (Gerber & Stocker, 2007; Quinn et al., 1974), providing well-validated methods for measuring learning. In contrast, learning in *D. sechellia and D. simulans* is only beginning to be described (Ellis et al., 2024; Kacsoh et al., 2018; Pak & Murashov, 2021): recent work by Ellis et al. (2024) found that *D. melanogaster*, *D. simulans* and *D. sechellia* were all capable of learning to avoid odours that had previously been paired with an electric shock. Here, our focus is on measuring learning, quantifying genetic variation and species divergence in learning and testing whether these patterns covary at the genotype and/or species level across stimulus types.

Second, isofemale lines from each species are available (Burns et al., 2020; Matute et al., 2014; Schrider et al., 2018) and flies can be reared in the laboratory under controlled conditions, allowing for robust measurements of genetic variation and covariation across contexts.

D. sechellia *and* D. simulans *Genotypes*

We tested isofemale lines, hereafter 'genotypes', from each species. Each genotype was established by inbreeding a single wild-collected female; therefore, individuals of the same genotype are much more genetically similar to one another than they are to individuals of other genotypes. Genotypes were generously provided by D. Matute in 2016 (Burns et al., 2020). *Drosophila sechellia* genotypes (specifically: NF 18, NF 33, NF 74, NF 111) were collected from various locations across the Seychelles, while the *D. simulans* genotypes were collected across various locations in central and southern Africa (NMB-014 collected in Namibia, NS-39 collected in Nairobi, LNP 15-063 collected in the Luwanga National Park, Zambia, LZV 15-003 collected in the Lower Zambieze Valley, Zambia) (Matute et al., 2014; Schrider et al., 2018; D. R. Matute, personal communication).

Fly Rearing

To rear flies for the trials, we placed 10 unmated females with 10 males of the same species and genotype in vials containing a standardized rearing medium, which consisted of corn meal, corn syrup, malt sugar, dead yeast, soy flour, Tegosept (methyl paraben), propionic acid and phosphoric acid. Previous work has demonstrated that *D. sechellia* can be reared successfully on standard laboratory medium, with no adverse effects on survival (Lavista-Llanos et al., 2014). We chose to rear all flies on the same standardized food to reduce the possibility that differences in food type could produce differences in learning independently from any genetic differences between species; this consideration may be particularly important when studying genetic correlations (Houle, 1991).

We allowed the parents to mate and lay eggs for 14 days and collected newly eclosed virgin offspring under light CO_2 anaesthesia on day 15. Collected flies were housed individually in vials containing standardized rearing medium and allowed to recover from the CO_2 anaesthesia for 3 days prior to beginning trials.

Stimuli

Aversive conditioning using an aversive gustatory stimulus has been found to elicit strong learning responses in many animals (Gustavson et al., 1974; Ralphs & Provenza, 1999; Yamamoto, 1993). During aversive conditioning, a negative unconditioned stimulus is paired with a conditioned stimulus for a training period; learning is indicated if preference for the conditioned stimulus is subsequently altered, even when tested in the absence of the negative unconditioned stimulus (Ayestaran et al., 2010).

To identify an appropriate aversive stimulus, we first confirmed that flies from the genotypes studied here avoid quinine. For each genotype, 4–15 flies of each sex were tested, for a total of 160 individuals. Our analysis, detailed in the Appendix, revealed that flies showed robust quinine avoidance in all contexts tested, with no significant differences among genotypes, species or sexes.

Therefore, we tested whether flies could associate quinine-laced foods with (1) gustatory and (2) visual stimuli. *Drosophila melanogaster* flies have previously been shown to detect and learn about these stimuli in a foraging context (gustatory stimuli: Gerber & Stocker, 2007; visual stimuli: Liu et al., 1999). For gustatory stimuli, we used plain fly food substrate (consisting of a standard recipe of agar, malt sugar, inactive dry yeast and deionized water) and imitation *M. citrifolia* (noni) food substrate. We chose to measure learning about noni fruit because this fruit is fundamental to differences in *D. sechellia* and *D. simulans* ecology, and thus is most likely to show differences between species, if any differences exist (Khurana et al., 2012). Imitation noni food substrate was made by adding octanoic and hexanoic acids to the plain food substrate (as in Burns et al., 2020; Dworkin & Jones, 2009). Previous work on *D. sechellia* has demonstrated this imitation noni food mimics key properties of noni while being highly reproducible (Auer et al., 2020; Burns et al., 2020; Dekker et al., 2006; Ibba et al., 2010; Jones, 2005; Lavista-Llanos et al., 2014; Prieto-Godino et al., 2017).

For visual stimuli, we used plain substrate (same as mentioned above) surrounded by a black-and-white stripe pattern tape and plain substrate surrounded by a black-and-white zigzag pattern tape. Black-and-white patterns were chosen to control for any possible variation in colour perception between individuals and species (Chittka et al., 2014; Dyer & Arikawa, 2014; Toler et al., 2005).

Trial Overview

We used a well-established 'nonreciprocal' conditioning paradigm (Croteau-Chonka et al., 2022; Honjo & Furukubo-Tokunaga, 2009; Khurana et al., 2012; Saltz, Hessel et al., 2017, Saltz, Lymer et al., 2017; Widmann et al., 2018), in which each individual's response to a stimulus was compared before and after that stimulus was paired with an aversive unconditioned stimulus. In the somewhat more common 'reciprocal' conditioning design, naïve flies are instead exposed first to the stimulus and the aversive tastant (or other unconditioned stimulus), then their preference is tested (Khurana et al., 2012; Widmann et al., 2018). The reciprocal approach is suitable in cases where it can be safely assumed that (1) animals have no naïve preferences and respond similarly to all the gustatory and visual stimuli (this is often established through preliminary experiments), (2) all genotypes show this same lack of naïve preference and (3) animals have equal ability to learn about different stimuli (Khurana et al., 2009, 2012). None of these assumptions were appropriate for the current study because we were explicitly testing the hypotheses that learning abilities may or may not generalize across different stimuli (and stimulus types). Moreover, ecologically relevant stimuli are likely to be subject to naïve preferences, but also particularly important for studying species differences that have potentially evolved. One downside of the nonreciprocal approach is that pretraining exposure to stimuli may itself influence learning (Jacob et al., 2021; but see Khurana et al., 2012). With this consideration in mind, the nonreciprocal approach is best suited to addressing the questions of interest here (Saltz, Hessel et al., 2017, Saltz, Lymer et al., 2017).

Therefore, each trial tested a single fly's behaviour across three trial stages. In stage 1, each fly was allowed to choose between the two gustatory stimuli (i.e. plain food versus imitation noni), or between the two visual stimuli (i.e. stripes versus zigzags). In this stage, we measured each fly's pretraining preference for the relevant stimuli.

In stage 2, one of the stimuli (either plain food or imitation noni in gustatory stimulus trials; either stripes or zigzags in visual stimulus trials) was paired with quinine. This 'training stage' provided the opportunity for the fly to associate one of the stimuli with the aversive quinine experience. Notably, this training paradigm does not control the temporal order of stimulus presentation. Instead, we aimed for a more ecologically relevant learning opportunity, in which individuals can sample habitats of variable

qualities and potentially adjust their future behaviour based on this experience. This approach, where quinine is added to a habitat type and flies are allowed to interact with that habitat type on their own schedule, has been used extensively in the closely related *D. melanogaster* (Dunlap & Stephens, 2009, 2014; Mery & Kawecki, 2002, 2003).

Finally, in stage 3, the fly was again allowed to choose between the two gustatory stimuli, or the two visual stimuli, exactly as in stage 1. Importantly, no quinine was present in stage 3. Thus, to quantify learning, we compared each fly's pretraining preference for the relevant stimulus with its post-training preference for that stimulus. Learning would be indicated if flies changed their preference; in particular, we predicted that flies would reduce their preference for the stimulus that previously had been paired with quinine.

We also had control treatments in which no aversive stimulus was present during the training stage; in this treatment, we predicted that flies' preferences would remain the same.

Each individual fly was tested only in a single trial. Individual flies were randomly assigned to a treatment.

Treatments

Gustatory stimulus learning trials included three aversive conditioning treatments: plain substrate paired with quinine, imitation noni substrate paired with quinine or a control treatment containing both substrates and lacking quinine.

Similarly, we had three treatments for visual stimulus learning trials: stripe pattern paired with quinine, zigzag pattern paired with quinine or a control treatment lacking quinine. Importantly, in all visual learning trials, only the plain fly food substrate was present, never imitation noni. Thus, flies could choose between plain food associated with one visual stimulus or plain food associated with the other visual stimulus.

Trial Details

Stage 1: pretraining preference

In the first stage of the experiment, each fly was gently aspirated into a short pipette tip and allowed to emerge on their own accord into a preference arena.

Each preference arena consisted of two petri dishes taped together, containing the two substrate options cut into halves (Fig. 1a, b). The substrates available were either imitation noni and plain food in gustatory learning trials, or plain food with zigzag visual cues and plain food with striped visual cues in visual learning trials (Fig. 1). Before trials occurred, arenas were prepared such that half of the arenas had a particular stimulus (e.g. noni or stripes) on the left side and the other half had the opposite spatial orientation. During trials, each fly at each stage was randomly assigned to an arena, allowing us to measure and account for any side biases (Alves et al., 2007; Andrade et al., 2001; Castellano et al., 1987; Jackson et al., 1998; Kight et al., 2008; Letzkus et al., 2006).

To measure pretraining responses to stimuli, the fly's location was measured immediately upon entering the arena, and then every 10 min over the course of a 2 h observation period (Fig. 2). If the fly was located in the middle of the arena, equidistant to both stimuli, the observation was excluded. In total, 13 choices were recorded for each individual fly during stage 1.

Stage 2: training (or controls)

Following the pretraining preference stage, each fly was gently aspirated into a training arena (Fig. 1c, d). Training arenas were the same as in stage 1, except now one stimulus was spiked with quinine. Control treatments included both substrate options, neither of which were paired with quinine. As in the pretraining preference stage, the relative location of the stimuli within the arenas was varied.

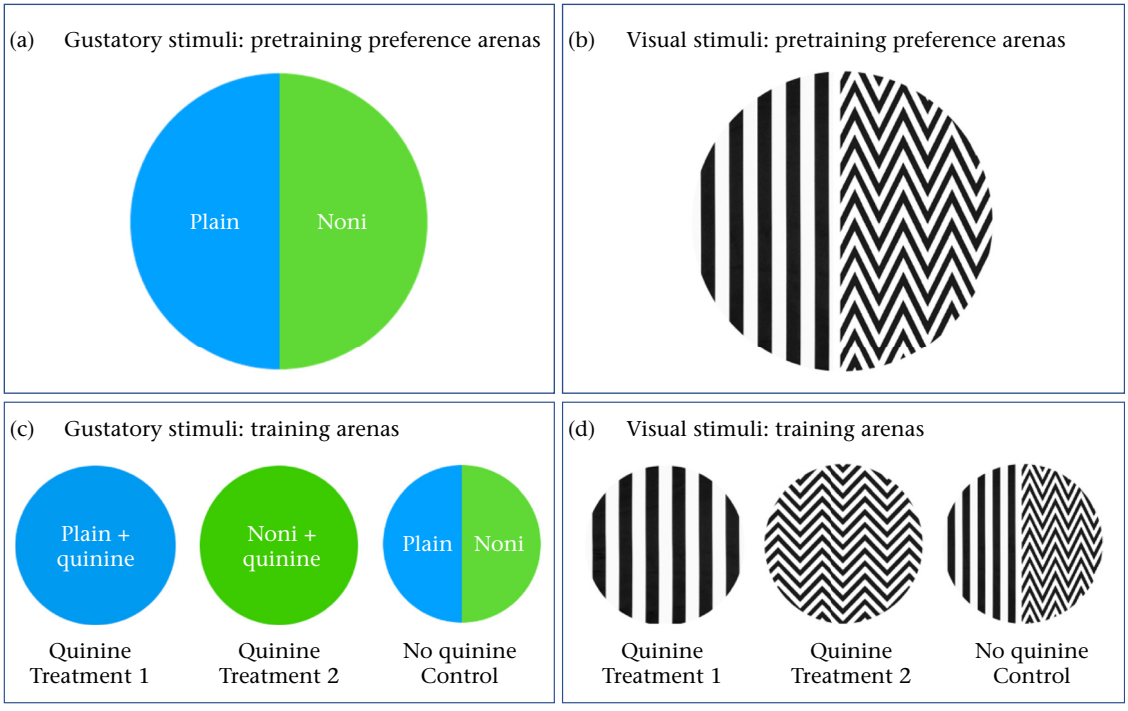

Figure 1. Pretraining preference arenas and training arenas. (a, b) Pretraining preference arenas consisted of one petri dish containing the two substrate options cut into halves, covered by a second petri dish (acting as a lid), and sealed together with tape. (c, d) Aversive conditioning training arenas consisted of a small petri dish containing either one substrate paired with quinine, or both substrates lacking quinine (acting as controls), depending on the assigned treatment. In all stages of the experiment, the relative location of the substrate options within the arenas was varied.

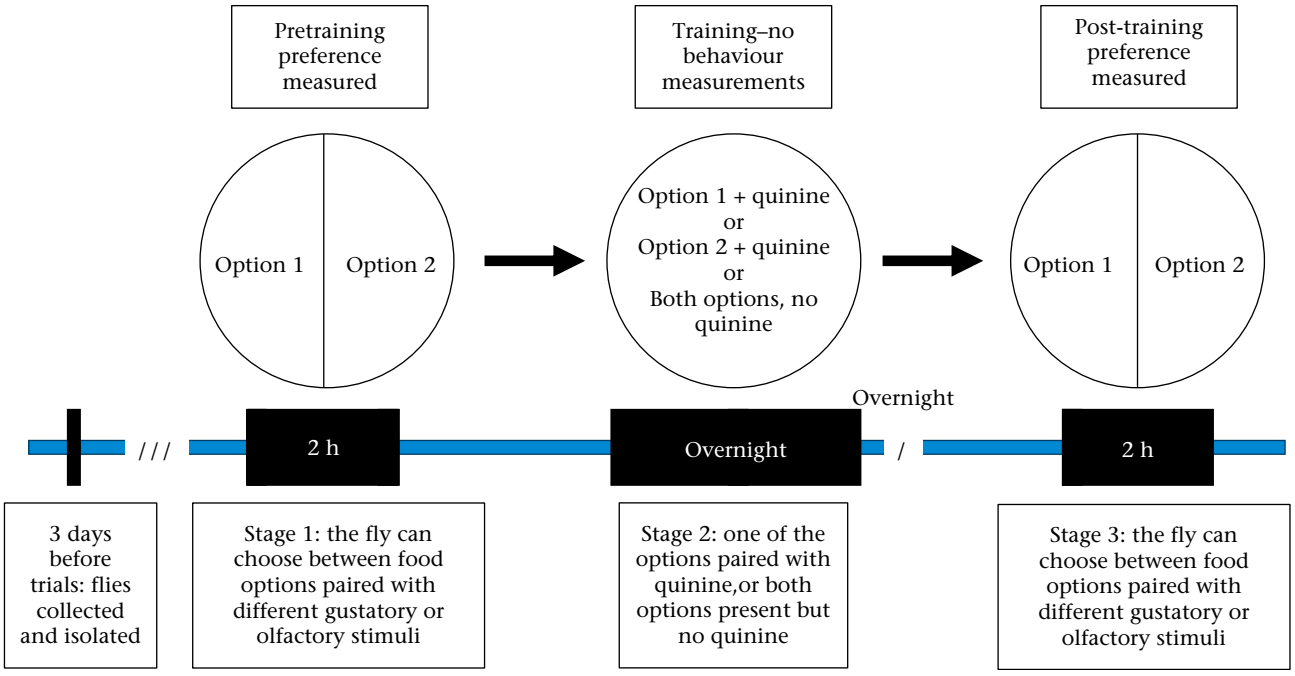

Figure 2. Overview of learning trials. In stage 1, each fly's pretraining preference for each of the two stimulus options was measured over the course of a 2 h period. Next, each fly was gently transferred into a stage 2 training arena and left overnight (for approximately 24 h). Finally, each fly's post-training preference was measured over a 2 h period in stage 3. After stage 3, the trial was complete; each individual fly was used in only one trial.

Flies were allowed to remain in the training arenas overnight; no behavioural measurements were obtained during this stage (Fig. 2). This 24 h overnight training period was used to account for the highly circadian behaviour of flies (Allada & Chung, 2010). In comparing pre- and post-training preferences, we aimed to ensure that changes in preference accurately reflected learning, and not circadian changes in activity levels or feeding motivation.

Stage 3: measuring learning

Following the training stage, flies were gently aspirated into a short pipette tip and allowed to emerge into the new preference arenas. The arenas, stimuli, protocols and behaviour measurements were exactly as in stage 1 (Fig. 2). In total, 13 choice observations were recorded for each fly in stage 3.

Replication

A total sample size of 452 individuals (198 *D. sechellia* and 254 *D. simulans*) were measured. A range of 6–16 flies of each sex–genotype–stimulus type combination were measured for the quinine learning trials. For no-quinine controls, a range of 2–9 flies of each sex–genotype–stimulus type combination were measured. Note that, for the no-quinine controls, our sample size for one genotype precluded us from estimating three-way interactions between genotype, sex and stimulus type for the controls, but this was not a goal of the current experiment; indeed, we did not identify any sex differences at all (see below). The number of replicates varied between genotypes because of variation in the availability of flies on the day of testing (Burns et al., 2020).

Ethical Note

For all aims included in this project, we worked exclusively with *Drosophila*, and thus did not require any licenses or permits. We adhered to all ASAB/ABS Guidelines. All flies were kept on ample food and housed in an experimental room with a 12:12 h light:dark cycle at approximately 26 °C (Stamps et al., 2013). Flies spent minimal time under CO_2 anaesthesia and were allowed to recover for 72 h before beginning experiments. Upon completion of each experiment, flies were sacrificed by being placed in a freezer for a minimum of 24 h.

ANALYSIS METHODS

Quantifying Pretraining Preference

To measure pretraining preference for each individual, we calculated the proportion of observed choices for each of the stimuli during stage 1. For gustatory stimuli, the plain substrate was arbitrarily given a value of 0, while the imitation noni substrate was given a value of 1 (Burns et al., 2020). For visual stimuli, the stripe pattern was arbitrarily given a value of 0, while the zigzag pattern was given a value of 1. Therefore, gustatory stimulus preference values near 0 indicated that the fly spent most of its time on the plain food, while values at or closer to 1 indicated that the fly spent most of its time on the imitation noni substrate. Similarly, visual stimulus preference values near 0 indicated that the fly spent most of its time on the striped side of the arena, while values at or closer to 1 indicated preference for the zigzag pattern. In all trials, pretraining preference values of 0.5 indicated no preference for either stimulus.

Quantifying Learning

To calculate learning scores, we recalculated pretraining preference scores (stage 1) based on the treatment received and compared these scores to the same fly's preference after training (stage 3), allowing us to compute learning scores that were directly comparable across treatments.

For each treatment, preferences were calculated such that the stimulus that was paired with quinine was assigned a value of 1 and the stimulus that was not paired with quinine was assigned a value

of 0. Then, we calculated stage 1 pretraining preferences and stage 3 preferences as described above (see Quantifying Pretraining Preference above). Learning scores were calculated as follows: learning score = stage 1 pretraining preference − stage 3 post-training preference. For example, imagine a fly in the treatment where the zigzag visual stimulus was paired with quinine. Imagine the fly showed no obvious preference for either stripes or zigzags in the pretraining preference test in stage 1 (i.e. the fly spent 50% of the time on the stripes–plain food side of the arena and 50% of the time on the zigzag-plain food side of the arena, resulting in an pretraining preference score of 0.5). During training, this fly learned that zigzag visual stimuli predict aversive quinine. Then, in stage 3, the fly avoided zigzags (even though quinine is not present during stage 3), spending only 10% of its time of the plain food associated with the zigzags visual cue (preference score = 0.1). This fly would receive a learning score of 0.5–0.1 = 0.4. Thus, positive learning scores indicate a decrease in preference for the stimulus that was paired with quinine, as expected for aversive associative conditioning. Negative learning scores indicate an unexpected increase in preference for the stimulus that was paired with quinine. Learning scores of 0 indicate no change in behaviour (Saltz, Hessel et al., 2017, Saltz, Lymer et al., 2017).

For the control treatments, the 'training' stage did not include quinine, so the directionality of the preference scores was computed using the same arbitrary conventions described in Quantifying Pretraining Preference above. Learning scores were calculated as: learning score = stage 1 pretraining preference − - stage 3 preference. However, in the control treatments, we predicted that learning scores would be indistinguishable from 0 on average, indicating no directional change in preference.

Modelling the Effects of Species, Sex, Genotype and Stimulus Type on Behaviour

Our goal was to estimate species, genotypic and sex differences in learning scores across the two stimulus types (gustatory and visual). Our first step was to investigate whether there were species, genotypic and sex differences in the flies' pretraining preferences (stage 1), which is important for interpreting the resulting learning scores (Mery & Burns, 2010; Stamps et al., 2018). Next, we applied the same analysis framework to the learning scores. All analyses were performed using R statistical software (version 4.1.0; R Core Team, 2021).

To investigate how differences among species, stimulus type, sex and genotype were associated with an individual's pretraining preference or learning, we fitted linear mixed models (LMMs) in a Bayesian framework using the 'brms' package in R (Bürkner, 2018), which is an interface to the MCMC sampler 'Stan' (Carpenter et al., 2017). This approach was chosen for several reasons. First, the Bayesian analysis framework provides, for each parameter, a summary of the range of parameter values that are reasonably likely given the data and the model (i.e. the credible interval estimated from the posterior distribution). Credible intervals can be straightforwardly compared to each other and to our null hypothesis of 0 (indicating no learning), which provides a quantitative overview of learning and its differences across species and treatments that accurately communicate uncertainty (for discussions of how Bayesian and frequentist inference differ, see Berry & Hochberg, 1999; Fornacon-Wood et al., 2021; Kruschke & Liddell, 2018; for a discussion of the benefits of combining these approaches, see Bayarri & Berger, 2004; Pick et al., 2023; for examples of using these approaches to study individual and genotypic differences in behaviour, see Araya-Ajoy & Dingemanse, 2017; Hutchins et al., 2024). The 'brms' package provides a robust toolkit for fitting Bayesian models and allows for the complex random effects needed in our analysis (see below).

We fitted three models to analyse variation in (1) pretraining preference, (2) no-quinine learning controls and (3) learning trials with quinine present. Control and quinine treatment learning scores were analysed in separate models because learning scores in the control trials were calculated with an arbitrary direction (see above). Each model included fixed effects of species, sex and stimulus type as well as a two-way interaction between species and stimulus type. For the learning with quinine model, we also included a parameter called 'quinine treatment' indicating which stimulus was associated with quinine during the training stage. Each quinine treatment was uniquely coded (as either quinine–noni or quinine–plain for gustatory stimulus learning, or quinine–zigzags or quinine–stripes for the visual stimulus learning), so this term was implicitly nested within stimulus type. In the pretraining preference model, a nonzero effect of species would indicate that species differ in their pretraining responses to gustatory and/or visual stimuli. In the learning with quinine present models, a nonzero effect of species would allow us to test hypothesis (1), that species differ in learning. The interaction term between species and stimulus type tests whether any differences between species were consistent or variable between gustatory and visual stimuli. A nonzero effect of this interaction term would provide evidence against hypothesis (2), that species differences were consistent across all stimulus types. Instead, a nonzero effect of the interaction term would indicate that the species with the highest learning scores when learning about visual stimuli did not also have the highest learning scores when tested with gustatory stimuli and vice versa.

In preliminary models, we also included a parameter describing the spatial arrangement of choices (to account for side bias) and another parameter to account for batch effects arising from differences among trial dates or observers. However, WAIC analysis indicated that models including these additional covariates were a poorer fit to the data (pretraining preference model: delta WAIC = 0.3; control model: delta WAIC = 6.5; learning with quinine model: delta WAIC = 8.6), so the terms were excluded from the final models.

Each model also included random intercepts for each genotype and a random effect two-way genotype * stimulus type interaction. The random effect of genotype provides an estimate of genetic variance; genotypes were implicitly nested within species. The random genotype * stimulus type interaction allowed us to test for genetic covariance; a nonzero effect of this term would indicate that genotype differences estimated for one stimulus type do not generalize to the other stimulus type, while an estimate for this term that is near zero would indicate consistency in genotype differences in learning across stimulus types. Furthermore, the model allowed for a correlation between genotype intercepts and slopes (i.e. the genotype-specific effect of stimulus type on learning scores). Thus, this correlation directly quantifies genetic correlations for learning scores across stimulus types. Therefore, these terms allowed us to test hypothesis (3), that species differences (if any) reflected correlations among genotypes within the species.

As a reminder, each individual fly was tested in only one trial with one stimulus type (visual or gustatory), so these random effects indicate whether different individuals from the same genotype are more similar to each other within and/or among stimulus types, compared to individuals from different genotypes.

Finally, because we previously observed differences in behavioural variation between these species (Burns et al., 2020) and because heterogeneity of variance was observed in preliminary models, we allowed the model to fit independent residual variances for each species.

In all models, for beta values of fixed effects, we specified weakly noninformative priors centred on zero. Error distributions were specified as Gaussian based on inspection of histograms of the raw data and model residuals. We visually assessed trace plots to ensure model fit and used $\widehat{R}$ criteria (less than 1.05) to assess convergence. All chains in all models converged properly and showed reasonable effective sample sizes.

Inference

We assessed the importance of fixed effects by inspecting the medians and credible intervals for the relevant posterior parameter distributions.

In the learning with quinine model, we computed posterior odds ratios for quinine treatment and for the species * stimulus type interaction, in order to interpret planned contrasts between each level of these parameter combinations. For each draw from the posterior, the difference in estimated mean response was calculated for each pairwise combination of ranks, resulting in one value per pair per draw. This allowed us to present the mean and credible intervals for each planned contrast, i.e. to compare each quinine treatment or each species–stimulus type combination.

Since variance components are bounded at zero (i.e. a variance cannot be negative), credible intervals are not a reliable indication of whether a particular variance component is meaningfully different from zero. Following recent recommendations (Pick et al., 2023), we computed *P* values (in the frequentist sense) for the random effects by using permutations (Araya-Ajoy & Dingemanse, 2017; Pick et al., 2023). For data pertaining to each model, we randomly reshuffled the values for genotype 1000 times. We restricted the reshuffling such that genotype information was permuted within species and quinine treatment. Each permuted data set was analysed using the relevant model, and we retained each model's variance parameter estimate for genotype intercepts, slopes (genotype differences in behaviour across stimulus types) and their correlation. Note that correlations are not bounded by zero, meaning that the permutations are not strictly necessary for this term; however, we included it for consistency with the other random effects of interest.

This approach produced a null distribution of 1000 medians of the posterior parameter estimate for each random effect of interest, describing the expected value of the posterior parameter estimate if a fly's behaviour was randomized with respect to its genotype. We then computed the proportion of these null estimates that were greater than or equal to the median of the posterior distributions from analysis of our real data to compute a *P* value (Pick et al., 2023). Variance parameter estimates from analysis of our real data were considered to be nonzero if their magnitude was greater than 97.5% of the corresponding estimates in the null distribution. This significance threshold was more stringent than the more typical 95%, because the pretraining preference data were used for the pretraining preference model and to compute learning scores. In other words, we implemented the Bonferroni correction. We report corrected *P* values.

Full model structures, posterior parameter estimates and *P* values (for random effects of interest) are reported in Table 1.

RESULTS

Pretraining Preference

For pretraining preference, we found no evidence for sex differences, species differences, differences between stimulus types or a species * stimulus type interaction (Table 1). The global intercept was estimated to be 0.50 (95% CI = [0.33, 0.66]), indicating that flies, on average, had no detectable preference for (or against) either stimulus in the visual or gustatory trials.

However, we did find strong support for genotypic differences in pretraining preference (estimate = 0.14, 95% CI = [0.07, 0.30], permuted $P < 0.001$) and an interaction between genotype and stimulus type (estimate = 0.20, 95% CI = [0.11, 0.43], permuted $P < 0.001$). Furthermore, genotypic differences were strongly negatively correlated across stimulus types (estimate = −0.92, 95% CI = [−1.00, −0.35], permuted $P = 0.002$; Table 1, Fig. 3a, b).

Learning Controls

In control trials where no quinine was present during any stage, all of the fixed effects in our models had credible intervals overlapping 0 (Table 1). The global intercept was estimated to be 0.04 (95% CI = [−0.18, 0.27]), indicating that no directional change in preference occurred. Similarly, we did not see evidence for differences among genotype intercepts (estimate = 0.11, 95% CI = [0.008, 0.35], permuted $P = 0.128$). However, we did see a significant genotype * stimulus type interaction (estimate = 0.24, 95% CI = [0.45, 0.65], corrected permuted $P = 0.028$). Genotypic differences in learning scores were not correlated across stimulus types (estimate = −0.45, 95% CI = [−0.97, 0.8], permuted $P = 0.20$).

Together, these results confirm that no overall directional change in preference occurred in either species in the absence of quinine. However, some genotypes may have shown nonzero changes in preferences for some stimulus types.

Learning with Quinine

In trials with quinine present in stage 2, the global intercept for learning scores was nonzero and positive, consistent with our prediction that flies learned in the expected direction, on average (estimate = 0.15, 95% CI = [0.013, 0.28]). We did not see main effects of species, sex or stimulus type (Table 1). Although most of the comparisons between quinine treatments were not different from zero (Table 1), we did see a difference between the two visual stimulus treatments, i.e. quinine associated with stripes compared to quinine associated with zigzags (estimate = −0.12, 95% CI = [−0.23, −0.01]). Thus, we saw variation in learning scores even within one of the stimulus types.

Most importantly, we saw evidence for a nonzero interaction between species and stimulus type (estimate = 0.26, 95% CI = [0.05, 0.46]; Table 1, Fig. 4a, b). Inspection of estimated marginal means and 95% CIs for each species–stimulus type combination revealed no evidence that *D. simulans*'s learning scores were different from zero for either stimulus type (simulans gustatory learning estimate = 0.0411, 95% CI = [−0.06, 0.15]; visual learning estimate = −0.02, 95% CI = [−0.16, 0.13]), indicating that *D. simulans* did not show evidence for learning. In contrast, *D. sechellia* estimated marginal means were nonzero for both stimulus types (gustatory learning estimate = 0.16, 95% CI = [0.04, 0.27]; visual learning estimate = −0.17, 95% CI = [−0.21, −0.022]). For gustatory learning, the positive estimated marginal mean indicates that *D. sechellia* learned to avoid the gustatory stimulus that was paired with quinine during stage 2 (adjusted for quinine treatment differences), as we expected (Fig. 4a). For visual learning, the negative estimated marginal mean indicates that *D. sechellia* unexpectedly increased their preference for the visual stimulus that was associated with quinine during stage 2, i.e. they learned in the 'wrong' direction (Fig. 4b).

We saw no evidence for genotype differences in learning (Table 1).

Table 1
Summary of model structures and results

Model name	Model structure	Parameter	Estimate	95% CI	*P* value (from permutations)
Pretraining preference model	Initial preference~1+species+sex+stimulus type+species*stimulus type+(1+stimulus type | genotype), sigma~0+species	**Intercept**	**0.50**	**[0.33, 0.66]**	–
		Species	−0.10	[−0.33, 0.13]	–
		Sex	−0.02	[−0.06, 0.03]	–
		Stimulus type	0.07	[−0.17, 0.30]	–
		Species*stimulus type	0.03	[−0.29, 0.37]	–
		Sechellia variance	−1.20	[−1.30, −1.1]	–
		Simulans variance	−1.50	[−1.59, −1.41]	–
		Genotype (intercept)	**0.14**	**[0.07, 0.30]**	**<0.001**
		Genotype*stimulus type	**0.20**	**[0.11, 0.43]**	**<0.001**
		Genetic correlation across stimulus types	**−0.92**	**[−1.00, −0.35]**	**0.002**
No quinine controls model	Learning score~1+species+sex+stimulus type+species*stimulus type+(1+stimulus type | genotype), sigma~0+species	Intercept	0.04	[−0.18, 0.27]	–
		Species	−0.09	[−0.36, 0.20]	–
		Sex	−0.04	[−0.15, 0.08]	–
		Stimulus type	0.03	[−0.36, 0.39]	–
		Species*stimulus type	0.02	[−0.45, 0.52]	–
		Sechellia variance	−0.74	[−0.90, −0.55]	–
		Simulans variance	−1.19	[−1.34, −1.02]	–
		Genotype (intercept)	0.11	[0.008, 0.35]	0.128
		Genotype*stimulus type	**0.24**	**[0.045, 0.65]**	**0.02**
		Genetic correlation across stimulus types	−0.45	[−0.97, 0.80]	0.2
Learning with quinine model	Learning score~1+species+sex+stimulus type+quinine treatment+species*stimulus type+(1+stimulus type | genotype), sigma~0+species	**Intercept**	**0.15**	**[0.01, 0.28]**	–
		Species	−0.11	[−0.27, 0.04]	–
		Sex	0.05	[−0.03, 0.13]	–
		Stimulus type	−0.23	[−1.36, 0.92]	–
		Quinine treatment: quinine noni−quinine plain	0.04	[−0.07, 0.14]	–
		Quinine treatment: quinine noni−quinine stripes	0.16	[−0.94, 1.34]	–
		Quinine treatment: quinine noni−quinine zigzags	0.04	[−1.11, 1.18]	–
		Quinine treatment: quinine plain−quinine stripes	0.12	[−0.96, 1.33]	–
		Quinine treatment: quinine plain−quinine zigzags	0.00	[−1.14, 1.15]	–
		Quinine treatment: quinine stripes−quinine zigzags	**−0.12**	**[−0.23, −0.01]**	–
		Species*stimulus type	**0.26**	**[0.05, 0.46]**	–
		Sechellia variance	−0.98	[−1.10, −0.85]	–
		Simulans variance	−1.16	[−1.26, −1.05]	–
		Genotype (intercept)	0.05	[0.002, 0.18]	0.63
		Genotype*stimulus type	0.05	[0.003, 0.23]	1
		Genetic correlation across stimulus types	−0.10	[−0.96, 0.94]	1

The structure of each model is reported along with results for each parameter estimate. Nonzero effects are bolded. Note that inference about fixed effects was based on inspection of 95% credible intervals (CI), while inference about genotypic variances was based on frequentist *P* values obtained through a permutation approach and corrected for multiple tests; we report corrected *P* values. No inference was conducted on species-specific residual variances. For details about the model fitting and inference approaches, please see main text.

DISCUSSION

Despite sustained interest in understanding how learning evolves, we still know little about when, whether and how evolutionary changes to learning are independent or linked across contexts and how these patterns arise from standing genetic covariation within species. In this study, we tested for differences in learning in two contexts between closely related, but ecologically divergent, species, *D. simulans* and *D. sechellia*. Our species comparisons also included tests for genetic variation and covariation within each species. Our most important findings were that (1) *D. simulans* flies showed no evidence for learning about either visual or gustatory cues, while (2) *D. sechellia* flies were able to learn to avoid gustatory stimuli that were previously paired with an aversive stimulus (Fig. 4a). Furthermore, (3) *D. sechellia* showed evidence for learning but in the unexpected direction in visual stimulus trials (Fig. 4b). (4) Neither species showed genetic variation in learning (Table 1); however, (5) there was strong evidence for genetic variation and covariation in pretraining preferences for both visual and gustatory stimuli in both species, despite the absence of overall species differences in pretraining preference (Fig. 3a, b).

None of our hypotheses were fully supported. *Drosophila simulans* showed no evidence for learning in either context, but *D. sechellia* showed nonzero learning in both contexts, supporting the basic idea of context generality (hypothesis 1). However, *D. sechellia*'s learning scores in the visual stimuli experiments were not in the predicted direction: *D. sechellia* flies unexpectedly increased their preference for whichever visual stimulus had previously been paired with quinine. These results suggest that gustatory responsiveness and learning may be under strong selection in *D. sechellia* as part of their evolutionary shift to specialization onto a novel food resource (Ellis et al., 2024), whereas this may not be the case for visual stimuli. This could be indicative of flies' inability to associate quinine, a noxious gustatory stimulus, with the visual pattern stimuli, possibly because visual stimuli have historically varied unreliably in relation to gastric consequences (Dunlap & Stephens, 2014; Garcia & Koelling, 1966). Instead, *D. sechellia* may have shown an increased preference for whichever pattern they had the most experience with. It is also possible that the opposite directionality of learning scores in *D. sechellia* between the visual and gustatory stimuli may be due to the speed, timing and sequence of stimulus perception, which our experimental design did not explicitly control for (Tanimoto et al., 2004). Notably, a recent study found that transgenic *D. sechellia* and *D. simulans* can learn to navigate a T-maze to avoid an aerosolized odour plume that had previously been paired with electric shock (Ellis et al., 2024). While there were many differences between this study and ours,

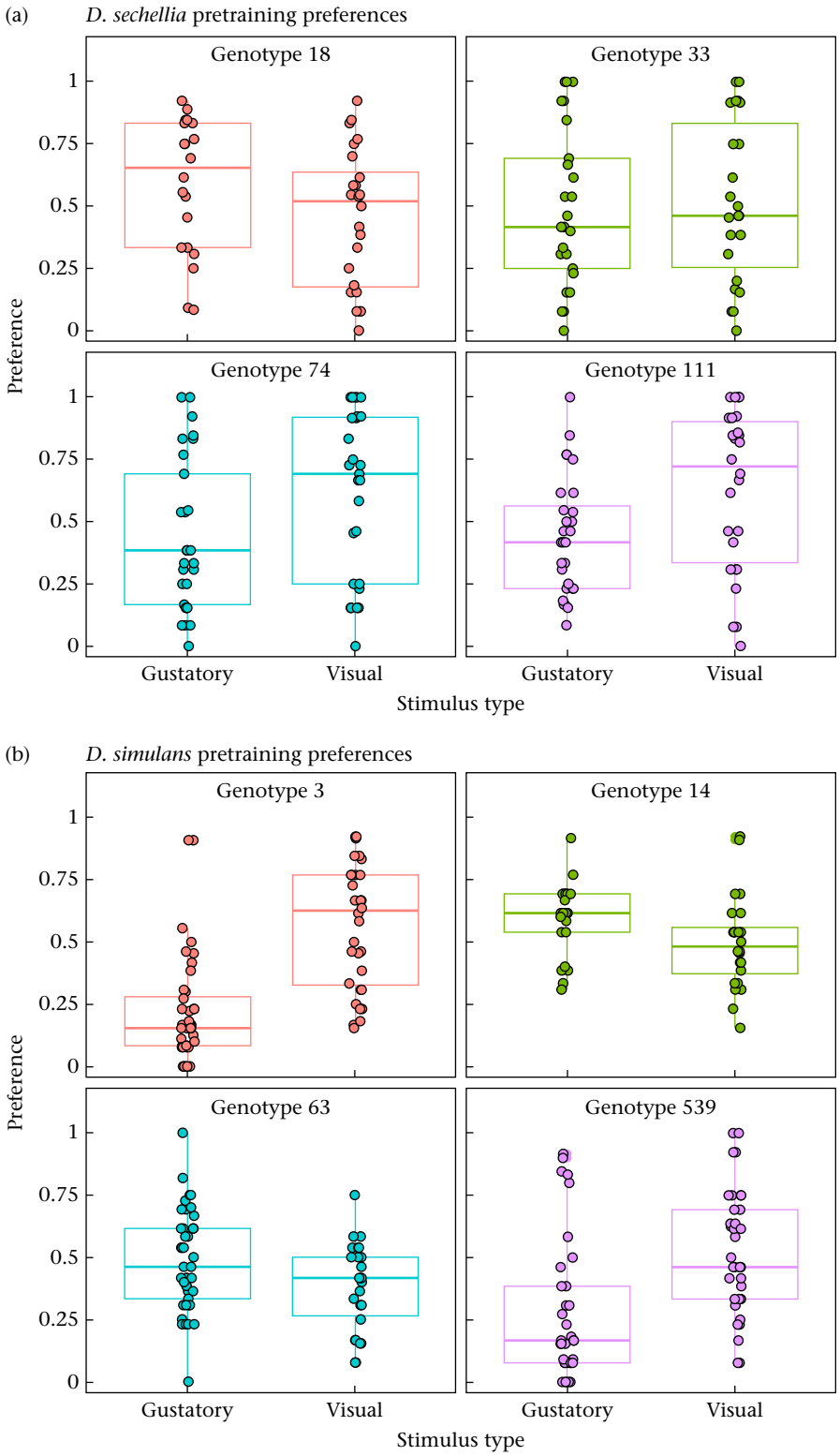

Figure 3. Pretraining preferences of (a) *D. sechellia* and (b) *D. simulans* for gustatory and visual stimuli. Higher values on the *Y* axes indicate a greater preference for noni over plain (for gustatory stimuli) or for zigzag over stripes (for visual stimuli). Box plots show medians (indicated by the line inside the boxes), the upper and lower quartiles (indicated by the box) and the outermost minimum and maximum values (indicated by the whiskers). Circles represent individual data points (i.e. the preference score of each individual fly).

taken together they suggest that at least some genotypes of each species can learn under some conditions, but that there are important differences in learning between these species that are highly context dependent. Overall, while it is clear that *D. sechellia* differed from the closely related *D. simulans* in behaviour on both the visual and gustatory learning tests, the specific cognitive mechanisms underlying these differences are likely complex.

The finding that the host specialist *D. sechellia* showed nonzero learning scores, but the opposite was true for the host generalist *D. simulans*, differs from previous findings in bats (Ratcliffe et al.,

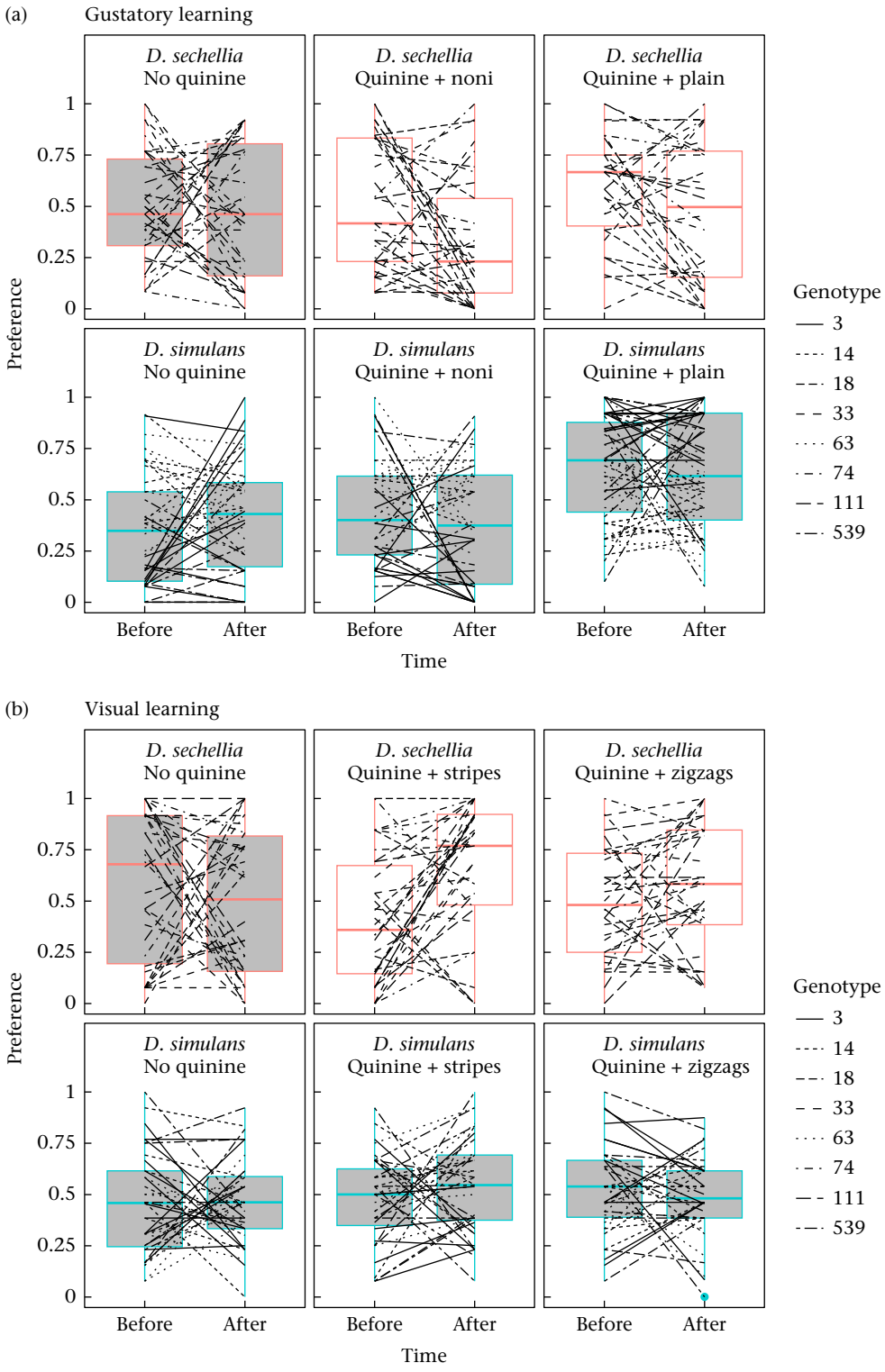

Figure 4. Pretraining (before) and post-training (after) preferences of *D. simulans* and *D. sechellia* during (a) gustatory learning and (b) visual learning. Higher values on the *Y* axes indicate a greater preference for whichever stimulus was paired with the quinine, or for an arbitrary stimulus in the no-quinine controls (see main text). Grey box plots indicate treatments that showed no evidence for learning, whereas white box plots indicate treatments where evidence for nonzero learning scores was discovered (see analysis in Methods and Results in main text). Lines show reaction norms for each individual, with line types corresponding to each genotype tested. Box plots show medians (indicated by the line inside the boxes), the upper and lower quartiles (indicated by the box) and the outermost minimum and maximum values (indicated by the whiskers). Circles represent individual data points (i.e. the preference score of each individual fly).

2003) and parasitoid wasps (Hoedjes et al., 2011; Poolman Simons et al., 1992). In these other examples, specialists showed weaker or absent learning compared to generalists. It is possible that the *D. simulans*–*D. sechellia* differences in learning evolved due to factors other than host specialization, such as the unique features of the noni fruit or demographic differences between these recently diverged species. Additional generalist–specialist comparisons, in flies and other organisms, will be needed to resolve the relationships between host specialization and learning ability.

At the genotype level, genotypes within each species differed in their pretraining preferences for the visual and gustatory stimuli but not in their learning scores. These results demonstrate that species differences in learning scores were due to species-specific tendencies to adjust their responses to each stimulus following training (for *D. sechellia*) or not (for *D. simulans*). Our study had only four genotypes per species; however, if additional sampling continues to show a greater magnitude of genetic variation for pretraining preferences than for learning, this would have interesting implications. First, our previous work in *D. melanogaster* found that pretraining preferences can be functionally related to learning scores (Stamps et al., 2018), which is seemingly not the case here. Perhaps these functional relationships vary across species and/or contexts. At the evolutionary level, we know that there must have been genetic variation within one or both species at some point, because without genetic variation, we could not observe species divergence. If studies continue to find little genetic variation in learning in these species, this would suggest that ancestral variation has since been lost through strong selection and/or drift.

It is somewhat surprising that we did not find an overall species differences in preference for the imitation noni substrate, as noni is the main host plant to *D. sechellia* and is toxic to *D. simulans*, and species differences in preferences have been found in previous work (Burns et al., 2020; Dworkin and Jones, 2009). The reason for this discrepancy is genetic variation: while two of our *D. simulans* strongly avoided noni as expected, the other two showed slight attraction to or weak avoidance of noni (Fig. 3a, b). Interestingly, in the wild, *D. simulans* have been collected from noni fruit, suggesting that some *D. simulans* individuals and genotypes are willing to explore noni fruit (Matute & Ayroles, 2014). Furthermore, in both species, we saw genetic covariance between preferences for visual and gustatory stimuli: genotypes that preferred the stripe visual stimulus over the zigzag visual stimulus also preferred the noni food over the plain food and vice versa. This result suggests that preferences for these stimuli may be functionally or evolutionarily linked in some unexpected way, perhaps contributing to the complex patterns we observed for pretraining preference. This finding underscores the utility of studying population genetic differences and species differences simultaneously.

This study, while powerful, was limited in several ways. First, our species comparison was limited to two species. This fact limits our ability to identify the relevant selection pressures, if any, that produced the species differences we observed. Despite this limitation, our design was suitable for our main goal of testing the context generality of genotypic and species differences in learning. In addition, our experimental design used relatively arbitrary time limits for each experimental stage, and the timing of stimulus presentation was shaped by the flies' own behaviours. As learning is inherently a process about detecting coincidences in time, allowing the flies to shape the timing of the learning process may have influenced our results as discussed above. Complementary approaches that increase ecological realism (e.g. by using wild-caught flies and/or heterogeneous fruits) and that increase control over stimulus presentation (e.g. by forcibly administering stimuli to immobile individuals) may be needed to fully understand the cognitive differences between these species and their evolutionary causes and consequences.

Learning is thought to be an intrinsic feature of brains (Hollis & Guillette, 2015), yet we still know little about whether learning across contexts evolves as an integrated process, or whether learning in each context is free to evolve separately. Here, we extend existing work focused on individual level differences to genetic variation and species differences. Additional empirical studies investigating divergence in learning across multiple contexts are required to determine whether these results are unique to our species comparison or indicative of a larger pattern.

Data Availability

Data and code are available via FigShare (https://doi.org/10.6084/m9.figshare.26153533).

Author Contributions

Madeline P. Burns: Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Formal analysis, Funding acquisition. **Julia B. Saltz:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Interest

The authors declare no conflict of interest.

Acknowledgments

We thank Daniel Matute for kindly providing the isogenic lines of *D. sechellia* and *D. simulans*. We thank Scott Egan, Kory Evans, Margaret Beier, Lisa O'Bryan, Marina Hutchins, Lea Pollack and Gihan Jayasinghe for providing helpful feedback on this manuscript. We thank Omar Moussa Pasha, Erin Harrison, Diana Alvarado, Jessica Nguyen and Anu Ayeni for assistance with data collection. This work was supported by U.S. National Science Foundation (NSF) IGERT Award 1250104 (PI: Raphael) and by NSF DEB Award 2217557 (PI: Saltz).

References

Allada, R., & Chung, B. Y. (2010). Circadian organization of behavior and physiology in *Drosophila*. *Annual Review of Physiology, 72*, 605–624.

Alves, C., Chichery, R., Boal, J. G., & Dickel, L. (2007). Orientation in the cuttlefish *Sepia officinalis*: Response versus place learning. *Animal Cognition, 10*, 29–36.

Andrade, C., Alwarshetty, M., Sudha, S., & Suresh Chandra, J. (2001). Effect of innate direction bias on T-maze learning in rats: Implications for research. *Journal of Neuroscience Methods, 110*, 31–35.

Araya-Ajoy, Y. G., & Dingemanse, N. J. (2017). Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. *Journal of Animal Ecology, 86*(2), 227–238.

Arien, Y., Dag, A., & Shafir, S. (2018). Omega-6:3 ratio more than absolute lipid level in diet affects associative learning in honey bees. *Frontiers in Psychology, 9*, Article 1001.

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature, 554*(7692), 364–367.

Auer, T. O., Khallaf, M. A., Silbering, A. F., Zappia, G., Ellis, K., Alvarez-Ocana, R., Arguello, J. R., Hansson, B. S., Jefferis, G. S. X. E., Caron, S. J. C., Knaden, M., & Benton, R. (2020). Olfactory receptor and circuit evolution promote host specialization. *Nature, 579*, 402–434.

Auld, J. R., Agrawal, A. A., & Relyea, R. A. (2009). Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society B: Biological Sciences, 277*(1681), 503–511.

Ayestaran, A., Giurfa, M., & Gabriela de Brito Sanchez, M. (2010). Toxic but drank: Gustatory aversive compounds induce post-ingestional malaise in harnessed honeybees. *PLoS One*, Article e150000.

Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal, 10*, 395–411.

Bayarri, M. J., & Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science, 19*(1), 58–80.

Behrman, E. L., Watson, S. S., O'Brien, K. R., Heschel, M. S., & Schmidt, P. S. (2015). Seasonal variation in life history traits in two *Drosophila* species. *Journal of Evolutionary Biology, 28*(9), 1691–1704.

Bell, A. (2012). Randomized or fixed order for studies of behavioral syndromes? *Behavioral Ecology, 24*(1), 16–20.

Berry, D. A., & Hochberg, Y. (1999). Bayesian perspectives on multiple comparisons. *Journal of Statistical Planning and Inference, 82*(1–2), 215–227.

Branch, C. L., Semenov, G. A., Wagner, D. N., Sonnenberg, B. R., Pitera, A. M., Bridge, E. S., Taylor, S. A., & Pravosudov, V. V. (2022). The genetic basis of spatial cognitive variation in a food-caching bird. *Current Biology, 32*, 210–219.

Buchanan, K. L., Grindstaff, J. L., & Pravosudov, V. V. (2013). Condition dependence, developmental plasticity, and cognition: Implications for ecology and evolution. *Trends in Ecology & Evolution, 28*, 290–296.

Burger, J. M. S., Kolss, M., Pont, J., & Kawecki, T. J. (2008). Learning ability and longevity: A symmetrical evolutionary trade-off in *Drosophila*. *Evolution, 62*(6), 1294–1304.

Burns, M. P., Cavallaro, F. D., & Saltz, J. B. (2020). Does divergence in habitat breadth associate with species differences in decision making in *Drosophila sechellia* and *Drosophila simulans*? *Genes, 11*(5), 528.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1), 1–32.

Castellano, M. A., Diaz-Palarea, M. D., Rodriguez, M., & Barroso, J. (1987). Lateralization in male rats and dopaminergic system: Evidence of right-side population bias. *Physiology & Behavior, 40*, 607–612.

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Caribol, A., Cauchard, L., Claidière, L., Dalesman, S., Devaud, J. M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., … Morand-Ferron, J. (2018). The repeatability of cognitive performance: A meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*, Article 20170281.

Chittka, L., Faruq, S., Skorupski, P., & Werner, A. (2014). Colour constancy in insects. *Journal of Comparative Physiology, 200*, 435–448.

Cordner, Z. A., & Tamashiro, K. L. (2015). Effects of high-fat diet exposure on learning & memory. *Physiology & Behavior, 152*(Pt B), 363–371.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology, 26*(6), 1447–1459.

Croteau-Chonka, E. C., Clayton, M. S., Venkatasubramanian, L., Harris, S. N., Jones, B. M. W., Narayan, L., Winding, M., Masson, J. B., Zlatic, M., & Klein, K. T. (2022). High-throughput automated methods for classical and operant conditioning of Drosophila larvae. *Elife, 11*, Article e70015.

Dekker, T., Ibba, I., Siju, K. P., Stensmyr, M. C., & Hansson, B. S. (2006). Olfactory shifts parallel superspecialism for toxic fruit in *Drosophila melanogaster* sibling, *D. sechellia*. *Current Biology, 16*, 101–109.

DeWitt, T. J., Sih, A., & Wilson, D. S. (1998). Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution, 13*(2), 77–81.

Dingemanse, N. J., & Wolf, M. (2013). Between-individual differences in behavioural plasticity within populations: Causes and consequences. *Animal Behaviour, 85*(5), 1031–1039.

DuBois, A. L., Nowicki, S., Peters, S., Rivera-Caceres, K. D., & Searcy, W. A. (2018). Song is not a reliable signal of general cognitive ability in a songbird. *Animal Behaviour, 137*, 205–213.

Dukas, R. (1999). Ecological relevance of associative learning in fruit fly larvae. *Behavioral Ecology and Sociobiology, 45*, 195–200.

Dukas, R. (2005). Learning affects mate choice in female fruit flies. *Behavioral Ecology, 16*(4), 800–804.

Dunlap, A. S., & Stephens, D. W. (2009). Components of change in the evolution of learning and unlearned preference. *Proceedings of the Royal Society B: Biological Sciences, 276*(1670), 3201–3208.

Dunlap, A. S., & Stephens, D. W. (2014). Experimental evolution of prepared learning. *Proceedings of the National Academy of Sciences of the United States of America, 111*(32), 11750–11755.

Dworkin, I., & Jones, C. D. (2009). Genetic changes accompanying the evolution of host specialization in *Drosophila sechellia*. *Genetics, 181*, 721–736.

Dyer, A. G., & Arikawa, K. (2014). A hundred years of color studies in insects: With thanks to Karl von Frisch and the workers he inspired. *Journal of Comparative Physiology, 200*, 409–410.

Ellis, K. E., Bervoets, S., Smihula, H., Ganguly, I., Vigato, E., Auer, T. O., Benton, R., Litwin-Kumar, A., & Caron, S. J. C. (2024). Evolution of connectivity architecture in the *Drosophila* mushroom body. *Nature Communications, 15*, 4872.

Fornacon-Wood, I., Mistry, H., Johnson-Hart, C., Faivre-Finn, C., O'Connor, J. P. B., & Price, G. J. (2021). Understanding the differences between Bayesian and frequentist statistics. *International Journal of Radiation Oncology, Biology, Physics, 112*(5), 1076–1082.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125.

Galsworthy, M. J., Paya-Cano, J. L., Monleon, S., & Plomin, R. (2002). Evidence for general cognitive ability (*g*) in heterogeneous stock mice and an analysis of potential confounds. *Genes, Brain and Behavior, 1*, 88–95.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*(1), 123–124.

Garrigan, D., Kingan, S. B., Geneva, A. J., Andolfatto, P., Clark, A. G., Thornton, K. R., & Presgraves, D. C. (2012). Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Research, 22*, 1499–1511.

Gerber, B., & Stocker, R. F. (2007). The *Drosophila* larva as a model for studying chemosensation and chemosensory learning: A review. *Chemical Senses, 32*(1), 65–89.

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology, 21*(3), 394–407.

Gustavson, C. R., Garcia, J., Hankins, W. G., & Rusiniak, K. W. (1974). Coyote predation control by aversive conditioning. *Science, 184*(4136), 581–583.

Hoedjes, K. M., Kruidhof, H. M., Huigens, M. E., Dicke, M., Vet, L. E., & Smid, H. M. (2011). Natural variation in learning rate and memory dynamics in parasitoid wasps: Opportunities for converging ecology and neuroscience. *Proceedings of the Royal Society B: Biological Sciences, 278*(1707), 889–897.

Hollis, K. L., & Guillette, L. M. (2015). What associative learning in insects tells us about the evolution of learned and fixed behavior. *International Journal of Comparative Psychology, 28*, 1–22. https://doi.org/10.46867/ijcp.2015.28.01.07

Honjo, K., & Furukubo-Tokunaga, K. (2009). Distinctive neuronal networks and biochemical pathways for appetitive and aversive memory in Drosophila larvae. *Journal of Neuroscience, 29*(3), 852–862.

Houle, D. (1991). Genetic covariance of fitness correlates: What genetic correlations are made of and why it matters. *Evolution, 45*(3), 630–648.

Hutchins, M., Douglas, T., Pollack, L., & Saltz, J. (2024). Genetic variation in male aggression is influenced by genotype of prior social partners in *Drosophila melanogaster*. *American Naturalist, 203*(5), 551–561. https://www.journals.uchicago.edu/doi/10.1086/729463.

Ibba, I., Angioy, A. M., Hansson, B. S., & Dekker, T. (2010). Macroglomeruli for fruit odors change blend preference in *Drosophila*. *Naturwissenschaften, 97*, 1059–1066.

Jackson, S., Nicolson, S. W., & Lotz, C. W. (1998). Sugar preferences and 'side bias' in Cape sugarbirds and lesser double-collared sunbirds. *Auk, 115*, 156–165.

Jacob, P. F., Vargas-Gutierrez, P., Okray, Z., Vietti-Michelina, S., Felsenberg, J., & Waddell, S. (2021). Prior experience conditionally inhibits the expression of new learning in *Drosophila*. *Current Biology, 31*(16), 3490–3503.

Johnson, D. D., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology & Evolution, 28*(8), 474–481.

Jones, C. D. (2005). The genetics of adaptation in *Drosophila sechellia*. *Genetica, 123*, 137–145.

Kacsoh, B. Z., Bozler, J., & Bosco, G. (2018). *Drosophila* species learn dialects through communal living. *PLoS Genetics, 14*(11), Article e1007825.

Kalan, A. K., Kulik, L., Arandjelovic, M., Boesch, C., Haas, F., Dieguez, P., Barratt, C. D., Abwe, E. E., Agbor, A., Angedakin, S., Aubert, F., Ayimisin, E. A., Bailey, E., Bessone, M., Brazzola, G., Buh, V. E., Chancellor, R., Cohen, H., Coupland, C., … Kühl, H. S. (2020). Environmental variability supports chimpanzee behavioural diversity. *Nature Communications, 11*, 4451.

Keagy, J., Savard, J. F., & Borgia, G. (2011). Complex relationship between multiple measures of cognitive ability and male mating success in satin bowerbirds, *Ptilonorhynchus violaceus*. *Animal Behaviour, 81*(5), 1063–1070.

Khurana, S., Abu Baker, M. B., & Siddiqi, O. (2009). Odour avoidance learning in the larva of *Drosophila melanogaster*. *Journal of Biosciences, 34*(4), 621–631.

Khurana, S., Robinson, B. G., Wang, Z., Shropshire, W. C., Zhong, A. C., Garcia, L. E., Corpuz, J., Chow, J., Hatch, M. M., Precise, E. F., Cady, A., Godinez, R. M., Pulpanyawong, T., Nguyen, A. T., Li, W. K., Seiter, M., Jahanian, K., Sun, J. C., Shah, R., … Atkinson, N. S. (2012). Olfactory conditioning in the third instar larvae of *Drosophila* melanogaster using heat shock reinforcement. *Behavior Genetics, 42*(1), 151–161.

Kight, S. L., Steelman, L., Coffey, G., Lucente, J., & Castillo, M. (2008). Evidence of population-level lateralized behaviour in giant water bugs, *Belostoma flumineum* Say (Heteroptera: Belostomatidae): T-maze turning is left biased. *Behavioural Processes, 79*, 66–69.

Kliman, R. M., Andolfatto, P., Coyne, J. A., Depaulis, F., Kreitman, M., Berry, A. J., McCarter, J., Wakeley, J., & Hey, J. (2000). The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics, 156*, 1913–1931.

Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., Immler, S., Makalov, A. A., & Kolm, N. (2013). Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Current Biology, 23*, 168–171.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review, 25*(1), 178–206.

Lachaise, D., David, J. R., Lemeunier, F., & Tsacas, L. (1986). The reproductive relationships of *Drosophila sechellia* with *D. mauritiana, D. simulans*, and *D. melanogaster* from the Afrotropical region. *Evolution, 402*, 262–271.

Lambert, C. T., & Guillette, L. M. (2021). The impact of environmental and social factors on learning abilities: A meta-analysis. *Biological Reviews, 96*(6), 2871–2889.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characteristics. *Evolution, 37*(6), 1210–1226.

Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nature, 1*(1), 36–41.
Lavista-Llanos, S., Svatos, A., Kai, M., Riemensperger, T., Birman, S., Stensmyr, M. C., & Hansson, B. S. (2014). Dopamine drives *Drosophila sechellia* adaptation to its toxic host. *eLife, 3*, 1–17.
Letzkus, P., Ribi, W. A., Wood, J. T., Zhu, H., Zhang, S. W., & Srinivasan, M. V. (2006). Lateralization of olfaction in the honeybee *Apis mellifera*. *Current Biology, 16*, 1471–1476.
Liu, L., Wolf, R., Ernst, R., & Heisenberg, M. (1999). Context generalization in *Drosophila* visual learning requires the mushroom bodies. *Nature, 400*, 753–756.
Machado, H. E., Bergland, A. O., O'Brien, K. R., Behrman, E. L., Schmidt, P. S., & Petrov, D. A. (2015). Comparative population genomics of latitudinal variation in *Drosophila simulans* and *Drosophila melanogaster*. *Molecular Ecology, 25*(3), 723–740.
Matute, D. R., & Ayroles, J. F. (2014). Hybridization occurs between *Drosophila simulans* and *D. sechellia* in the Seychelles archipelago. *Journal of Evolutionary Biology, 27*(6), 1057–1068.
Matute, D. R., Gavin-Smyth, J., & Liu, G. (2014). Variable post-zygotic isolation in *Drosophila melanogaster/D. simulans* hybrids. *Journal of Evolutionary Biology, 27*, 1691–1705.
Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., Specht, S. M., & Gandhi, C. C. (2003). Individual differences in the expression of a 'general' learning ability in mice. *Journal of Neuroscience, 23*(16), 6423–6433.
Matzel, L. D., Patel, H. M., Piela, M. C., Manzano, M. D., Tu, A., & Crawford, D. W. (2020). General cognitive ability predicts survival-readiness in genetically heterogeneous laboratory mice. *Frontiers in Ecology and Evolution, 8*, Article 531014. https://doi.org/10.3389/fevo.2020.531014
Mery, F., & Burns, J. G. (2010). Behavioural plasticity: An interaction between evolution and experience. *Evolutionary Ecology, 24*, 571–583.
Mery, F., & Kawecki, T. J. (2002). Experimental evolution of learning ability in fruit flies. *Proceedings of the National Academy of Sciences of the United States of America, 99*(22), 14274–14279.
Mery, F., & Kawecki, T. J. (2003). A fitness cost of learning ability in *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences, 270*(1532), 2465–2469.
Mery, F., & Kawecki, T. J. (2004a). The effect of learning on experimental evolution of resource preference in *Drosophila melanogaster*. *Evolution, 58*(4), 757–767.
Mery, F., & Kawecki, T. J. (2004b). An operating cost of learning in *Drosophila melanogaster*. *Animal Behaviour, 68*(3), 589–598.
Messier, C., Whately, K., Liang, J., Du, L., & Puissant, D. (2007). The effects of a high-fat, high-fructose, and combination diet on learning, weight, and glucose regulation in C57BL/6 mice. *Behavioural Brain Research, 178*(1), 139–145.
Molteni, R., Barnard, R. J., Ying, Z., Roberts, C. K., & Gómez-Pinilla, F. (2002). A high-fat, refined sugar diet reduces hippocampal brain-derived neurotrophic factor, neuronal plasticity, and learning. *Neuroscience, 112*(4), 803–814.
Morand-Ferron, J., Cole, E. E., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: A review of practical and conceptual challenges. *Biological Reviews, 91*, 367–389.
Pak, E. S., & Murashov, A. K. (2021). *Drosophila* passive avoidance behavior as a new paradigm to study associative aversive learning. *Journal of Visualized Experiments, 176*, Article e63163.
Pfenning, D. W., Wund, M. A., Snell-Rood, E. C., Cruickshank, T., & Schlichting, C. D. (2010). Phenotypic plasticity's impact on diversification and speciation. *Trends in Ecology & Evolution, 25*(8), 459–467.
Pick, J. L., Kasper, C., Allegue, H., Dingemanse, N. J., Dochtermann, N. A., Laskowski, K. L., Lima, M. R., Schielzeth, H., Westneat, D. F., Wright, J., & Araya-Ajoy, Y. G. (2023). Describing posterior distributions of variance components: Problems and the use of null distributions to aid interpretation. *Methods in Ecology and Evolution, 14*, 2557–2574.
Plomin, R. (1999). Genetics and general cognitive ability. *Nature, 402*, C25–C29.
Plomin, R. (2001). The genetics of *g* in human and mouse. *Nature Reviews Neuroscience, 2*, 136–141.
Plomin, R., & Spinath, F. M. (2002). Genetics and general cognitive ability (*g*). *Trends in Cognitive Sciences, 6*, 169–176.
Poolman Simons, M. T. T., Suverkropp, B. P., Vet, L. E. M., & de Moed, G. (1992). Comparison of learning in related generalist and specialist eucoilid parasitoids. *Entomologia Experimentalis et Applicata, 64*(2), 117–124.
Pravosudov, V. V., & Clayton, N. S. (2002). A test of the adaptive specialization hypothesis: Population differences in caching, memory, and the hippocampus in black-capped chickadees (*Poecile atricapilla*). *Behavioral Neuroscience, 116*(4), 515–522.
Prentice, P. M., Mnatzaganian, C., Houslay, T. M., Thornton, A., & Wilson, A. J. (2022). Individual differences in spatial learning are correlated across tasks but not with stress response behaviour in guppies. *Animal Behaviour, 188*, 133–146.
Prieto-Godino, L. L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A. F., Ruta, V., Peraro, M. D., & Benton, R. (2017). Evolution of acid-sensing olfactory circuits in drosophilids. *Neuron, 93*, 661–676.
Quinn, W. G., Harris, W. A., & Benzer, S. (1974). Conditioned behavior in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America, 71*(3), 708–712.
R'Kha, S., Capy, P., & David, J. R. (1990). Host-plant specialization in the *Drosophila melanogaster* species complex: A physiological, behavioral, and genetical analysis. *Proceedings of the National Academy of Sciences of the United States of America, 88*, 1835–1839.
Ralphs, M., & Provenza, F. (1999). Conditioned food aversions: Principles and practices, with special reference to social facilitation. *Proceedings of the Nutrition Society, 58*(4), 813–820.
Ratcliffe, J., Fenton, B., & Galef, B. (2003). An exception to the rule: Common vampire bats do not learn taste aversions. *Animal Behaviour, 65*, 385–389.
Richardson, D. S., Burke, T., & Komdeur, J. (2003). Sex-specific associative learning cues and inclusive fitness benefits in the Seychelles warbler. *Journal of Evolutionary Biology, 16*(5), 854–861.
Roth, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: A common garden approach. *Proceedings of the Royal Society B: Biological Sciences, 277*, 3187–3193.
Saltz, J. B., Hessel, F. C., & Kelly, M. W. (2017). Trait correlations in the genomics era. *Trends in Ecology & Evolution, 32*(4), 279–290.
Saltz, J. B., Lymer, S., Gabrielian, J., & Nuzhdin, S. V. (2017). Genetic correlations among developmental and contextual behavioral plasticity in *Drosophila melanogaster*. *American Naturalist, 190*(1), 61–72.
Schrider, D. R., Ayroles, J., Matute, D. R., & Kern, A. D. (2018). Supervised machine learning reveals introgressed loci in the genomes of *Drosophila simulans* and *D. sechellia*. *PLoS Genetics, 14*, Article e1007341.
Schuck-Paim, C., Alonso, W. J., & Ottoni, E. B. (2008). Cognition in an ever-changing world: Climatic variability is associated with brain size in Neotropical parrots. *Brain, Behavior and Evolution, 71*(3), 200–215.
Shettleworth, S. J. (1972). Constraints on learning. *Advances in the Study of Behavior, 4*, 1–68.
Shettleworth, S. J. (1993). Varieties of learning and memory in animals. *Journal of Experimental Psychology: Animal Behavior Processes, 19*(1), 5–14.
Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour, 61*(2), 277–286.
Shettleworth, S. J. (2012). Modularity, comparative cognition and human uniqueness. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1603), 2794–2802.
Sih, A. (2013). Understanding variation in behavioural responses to human-induced rapid environmental change: A conceptual overview. *Animal Behaviour, 85*(5), 1077–1088.
Sih, A., & Del Giudice, M. (2012). Linking behavioural syndromes and cognition: A behavioural ecology perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1603), 2762–2772.
Snell-Rood, E. (2013). An overview of the evolutionary causes and consequences of behavioural plasticity. *Animal Behaviour, 85*(5), 1004–1011.
Sol, D., Duncan, R. P., Blackburn, T. M., Cassey, P., & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 5460–5465.
Stamps, J. A., Biro, P. A., Mitchell, D. J., & Saltz, J. B. (2018). Bayesian updating during development predicts genotypic differences in plasticity. *Evolution, 72*(10), 2167–2180.
Stamps, J. A., Saltz, J. B., & Krishnan, V. V. (2013). Genotypic differences in behavioural entropy: Unpredictable genotypes are composed of unpredictable individuals. *Animal Behaviour, 86*(3), 641–649.
Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *International Society for Behavioral Ecology, 2*, 77–89.
Tanimoto, H., Heisenberg, M., & Gerber, B. (2004). Event timing turns punishment to reward. *Nature, 430*, 983.
Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., & Pravosudov, V. V. (2019). Spatial memory and cognitive flexibility trade-offs: To be or not to be flexible, that is the question. *Animal Behaviour, 147*, 129–136.
Thompson, R. F. (1986). The neurobiology of learning and memory. *Science, 233*(4767), 941–947.
Toler, T. R., Evans, E. W., & Tepedino, V. J. (2005). Pan-trapping for bees (Hymenoptera: Apiformes) in Utah's West Desert: The importance of color diversity. *Pan-Pacific Entomologist, 81*(3/4), 103–113.
Verzijden, M. N., ten Cate, C., Servedio, M. R., Kozak, G. M., Boughman, J. W., & Svensson, E. I. (2012). The impact of learning on sexual selection and speciation. *Trends in Ecology & Evolution, 27*(9), 511–519.
Widmann, A., Eichler, K., Selcho, M., Thum, A. S., & Pauls, D. (2018). Odor–taste learning in *Drosophila* larvae. *Journal of Insect Physiology, 106*(1), 47–54.
Yamamoto, T. (1993). Neural mechanisms of taste aversion learning. *Neuroscience Research, 16*(3), 181–185.

Appendix

Confirming Quinine Avoidance in D. sechellia *and* D. simulans *Genotypes*

Quinine hydrochloride has been used extensively as an aversive stimulus in learning experiments in *D. melanogaster*. In *D. melanogaster*, flies show strong avoidance of quinine that does not lessen over time (i.e. flies do not habituate to quinine) (Mery & Kawecki, 2002; Quinn et al., 1974). To confirm that quinine is also aversive to *D. sechellia* and *D. simulans*, we allowed flies from the same genotypes used in the learning trials to choose between quinine-laced (3.2 g/litre concentration) and quinine-free substrate in two treatments. Each fly was randomly assigned to experience an arena with either (1) plain fly food substrate (consisting of a

standard recipe of agar, malt sugar, inactive dry yeast and deionized water) laced with quinine and quinine-free imitation noni substrate, or (2) imitation noni substrate laced with quinine and quinine-free plain substrate. We also randomized the orientation of the substrate options. Flies were allowed to explore these options overnight, and their location (on plain substrate or on noni substrate) was recorded the following morning.

For each genotype, 4–15 flies of each sex were tested, for a total of 160 individuals. To investigate any differences among species, sexes, and genotypes in quinine avoidance, we fitted generalized linear mixed models (GLMMs) in a Bayesian framework using the brms package in R (Bürkner, 2018), which is an interface to the MCMC sampler 'Stan' (Carpenter et al., 2017).

Our response variable was substrate choice, where values of 1 were arbitrarily designated to indicate that the fly was on the quinine-laced food, and values of 0 indicated that the fly was on the quinine-free food. Because each fly was measured only once, after overnight experience with both substrates, we specified a Bernoulli error distribution.

Our model included fixed effects of species, sex and quinine treatment (i.e. whether quinine was added to the plain food or the imitation noni food), as well as a two-way interaction between species and quinine treatment.

In the preliminary model, we also included a parameter describing the spatial arrangement of choices (i.e. which was on the left and which was on the right, relative to the fly's entrance point into the arena), and another parameter to account for any batch effects arising from differences among trial dates or observers; however, WAIC analysis indicated that this models was a poorer fit to the data than a competing model without these additional covariate (WAIC 1.6) so the terms were excluded from the final model.

The model also included genotype and a two-way genotype∗quinine treatment interaction as random effects. The intercept of this term tests for differences among genotypes (adjusted for any overall species differences via the fixed effects in the same model), in quinine avoidance. Note that, because each genotype has a unique identifier, genotypes were implicitly nested within species. The interaction between genotype and quinine treatment tests whether genotype differences are consistent when quinine is added to the plain food and the noni food.

We assessed the importance of fixed effects by inspecting the medians and credible intervals for the relevant posterior parameter distributions. We assessed random effect significance using a permutation approach as detailed in the article main text.

For beta values of fixed effects, we specified weakly non-informative priors centred on zero. We visually assessed trace plots to ensure model fit and used $\widehat{R}$ criteria (less than 1.05) to assess convergence.

Seventy-five per cent (120/160) of our flies were observed on the quinine-free food. Consistent with this, we found that the model intercept was nonzero and negative (estimate = −2.37, 95% CI = [−3.55, −1.20]), confirming that flies overall avoided the quinine-laced food. Credible intervals for all the other fixed effects overlapped zero (species differences: estimate = 0.80, 95% CI = [−0.40, 2.0]; quinine treatment: estimate = 0.80, 95% CI = [−0.35, 1.9]; sex differences: estimate = 0.16, 95% CI = [−0.61, 0.95]; species∗quinine treatment: estimate = 0.01, 95% CI = [−0.13, 2.54]), demonstrating that there were no species differences, sex differences or effects of quinine treatment.

Similarly, we found no evidence that any of our variance components differed significantly from zero (genotype differences: estimate = 0.64, permuted $P = 0.125$; genotype∗quinine treatment: estimate = 0.66, permuted $P = 0.365$).

Based on this experiment, we conclude that quinine is an appropriate aversive conditioning stimulus for the tested genotypes of these species.

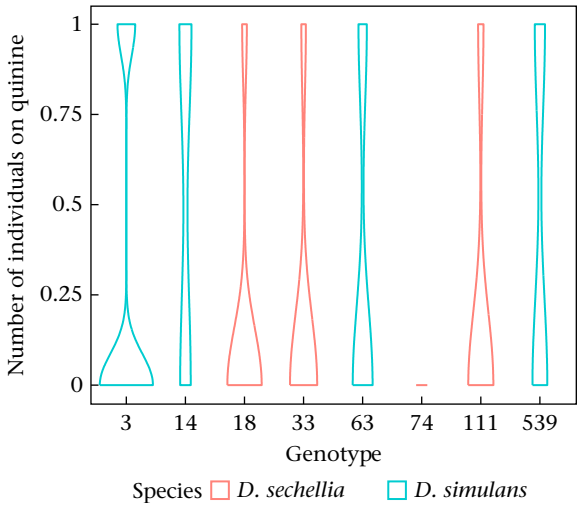

Figure A1. Number of individuals on quinine per genotype. For each genotype, 4–15 flies of each sex were tested, for a total of 160 individuals.