Evolutionary genetic integration of behavioural and endocrine components of the stress response

T.M. Houslay*[1,2], R.L. Earley[3], S.J. White[1,4], W. Lammers[1], A.J. Grimmer[1], L.M. Travers[1,5], E.L. Johnson[3,6], A.J. Young[1] and A.J. Wilson[1]

* Corresponding author: houslay@gmail.com

[1] Centre for Ecology and Conservation, University of Exeter (Penryn Campus), Cornwall, TR10 9FE, UK.

[2] Current address: Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK.

[3] Department of Biological Sciences, University of Alabama, Biology Building 211-213, Box 870344, Tuscaloosa, AL 35487, USA.

[4] Current address: Department of Ecology and Genetics, Center for Evolutionary Biology, Uppsala University, Uppsala, Sweden.[CHECK THIS]

[5] Current address: School of Biological Sciences, University of East Anglia, Norwich, Norfolk, NR4 7TJ, UK.

[6] Current address: Southern Research, 2000 Ninth Avenue South, Birmingham, AL 35205, USA.

Abstract

The vertebrate stress response comprises a suite of behavioural and physiological traits that must be functionally integrated to ensure organisms cope adaptively with acute stressors. The expectation that natural selection has favoured functional integration leads to a prediction of genetic integration: genetic variation in the stress response should include covariation between its component behavioural and physiological traits. Despite the implications of such genetic integration for our understanding of human and animal health, as well as evolutionary responses to natural and anthropogenic stressors, formal quantitative genetic tests of this prediction are lacking. Here we demonstrate that Trinidadian guppies (*Poecilia reticulata*) show genetic variation in a suite of behavioural and physiological components of the acute stress response, and that these are indeed integrated into a single major axis of genetic variation. This axis appears to reflect continuous variation in the magnitude of integrated stress responsiveness, rather than variation in 'coping style' (a verbal model that postulates equal levels of stress responsiveness will manifest differently across individuals). The genetic integration we find here could either facilitate or constrain evolutionary responses to selection, depending upon the extent to which the direction of selection aligns with this single major axis of genetic covariation among stress response traits. Such integration also suggests that, while stress-related disease typically arises from physiological components of the stress response, selection on the genetically correlated behavioural responses to stress could offer a viable non-invasive route to the genetic improvement of health and welfare in captive animal populations.

Introduction

Stress responses comprise suites of physiological and behavioural traits that enable individuals to cope with adverse conditions (Romero 2004; Øverli *et al.* 2007; McEwen & Wingfield 2010; Taborsky *et al.* 2020). Some individuals are likely better at coping with adverse conditions than others, and understanding the role played by underlying genetic variation could have important implications for managing stress-related disease in captive populations and predicting the evolutionary responses of free-living populations to both natural and anthropogenic stressors (Barton & Iwama 1991; Koolhaas *et al.* 1999; McEwen & Wingfield 2003; Romero 2004; Koolhaas 2008). For instance, if among-individual differences in stress response traits are a product of genetic variation (Koolhaas *et al.* 1999, 2007), then they may be a viable target for artificial selection strategies (Mignon-Grasteau *et al.* 2005). This could be used to reduce stress-related welfare issues in captive populations (e.g. in livestock; Broom & Johnson 1993; von Borell 1995; Möstl & Palme 2002). In free-living populations, variation in stress response traits is expected to cause fitness variation under stressful conditions (Wingfield 2003; Koolhaas 2008). Thus, exposure to stressors could lead to evolutionary changes in the distributions of traits that contribute to a population's long-term resilience in the face of natural (e.g. predation risk; Clinchy, Sheriff & Zanette 2013) and/or anthropogenic (Tarlow & Blumstein 2007; Busch & Hayward 2009; Angelier & Wingfield 2012) challenges. However, as the evolutionary response of any trait to selection is determined in large part by its genetic variation, our ability to predict – and potentially harness – evolutionary changes in the stress response is currently hampered by limited understanding of the underlying genetics.

Natural selection does not act on single traits in isolation, but rather on multivariate phenotypes (Lande & Arnold 1983; Blows 2007). This is likely to be an important consideration for understanding the evolution of the stress response. For instance, while glucocorticoid (GC) levels are frequently used to measure the stress

response (McEwen & Wingfield 2003; Korte *et al.* 2005), an individual's first line of defence against acute environmental challenges will typically be behavioural (Moberg 2000). This may include risk avoidance strategies as well as the widely known 'fight-or-flight' responses. Subsequent GC release then serves to mediate physiological (and further behavioural) responses (Wingfield *et al.* 1998; Wingfield & Kitaysky 2002). Natural selection is therefore expected to favour combinations of behavioural and physiological stress response traits that act synergistically to maintain fitness under stressful conditions (Koolhaas *et al.* 1999; Øverli *et al.* 2007). Since evolutionary theory predicts that correlational selection will shape the structure of multivariate quantitative genetic variance (as represented by the genetic covariance matrix **G**; Blows 2007), we should expect genetic – as well as phenotypic – integration of behavioural and physiological stress response traits (McGlothlin & Ketterson 2008; Ketterson, Atwell & McGlothlin 2009; Cox, McGlothlin & Bonier 2016). However, there exists relatively little supporting empirical work for this prediction.

The most compelling evidence for genetic integration of behavioural and physiological stress response traits to date comes from artificial selection experiments on domestic animal populations (e.g., rainbow trout (Pottinger and Carrick, 1999), Japanese quail (Jones, Satterlee & Ryder 1994), house mice, Veenema *et al.* 2003b). For example, lines of rainbow trout selected for stress-induced plasma cortisol levels (Pottinger and Carrick, 1999) experienced correlated evolutionary changes in behaviour (Øverli, Winberg & Pottinger 2005). In a rare study of a wild-type population (albeit under captive conditions) species, cortisol levels were found to evolve in response to selection on behavioural 'personality' in great tits (*Parus major*; Carere *et al.* 2003). In these examples, the correlated responses of behavioural and physiological stress response traits to selection are consistent with some degree of genetic integration of these behavioural and physiological traits. However, the extent and 'structure' of this integration remains unclear, and some results were inconsistent with a hypothesised

simple axis of genetic (co)variation among behavioural and physiological components of the stress response. For example, in the trout study, the 'low-cortisol response' selected lines actually showed a higher metabolic stress response under confinement (suggestive of opposing responses to selection by different physiological components of the stress response; Trenzado, Carrick & Pottinger 2003). Surprisingly (given the correlated selection response), in the great tit study cortisol levels were found to be phenotypically unrelated to behavioural 'boldness' (the trait selected on) under standardised testing (Thomson *et al.* 2011).

While selection experiments illustrate that genetic integration of behaviour and physiology can occur, estimation of the genetic variance-covariance matrix (**G**) through quantitative genetic modelling provides a complementary strategy that also allows investigation of exactly how (multivariate) genetic variation is structured within populations. In the context of the stress response, this should provide insights into both how selection has acted in the past (Ketterson *et al.* 2009), and whether responses to future selection are likely to be constrained (Blows & Walsh 2009; Walsh & Blows 2009). The former follows from the fact that strong correlational selection should lead to integration of traits in **G** over the long term, a phenomenon explored most commonly for suites of morphological traits (following, e.g., Cheverud 1982), but that is equally applicable to any aspect of phenotype (see, e.g., Hine, Chenoweth & Blows 2004; Hunt *et al.* 2007; Oswald, Singer & Robison 2013 for examples pertaining to behavioural evolution and mate choice). The latter follows from the fact that the direction (in multivariate trait space) and magnitude of a response to contemporary selection is limited by the amount of variance in **G** alignment with the vector of (directional) selection $\boldsymbol{\beta}$ (Blows & Walsh 2009; Walsh & Blows 2009).

Here we estimate **G** for behavioural and physiological components of the acute stress response in Trinidadian guppies (*Poecilia reticulata*). This enables us to determine not only (i) whether these components are genetically integrated into a

126 single major axis of genetic (co)variation, but also (ii) whether the structure and
127 orientation of this axis suggests variation in overall stress responsiveness and/or
128 'coping style' (explained further below; Koolhaas *et al.* 2010; Boulton *et al.* 2015). We
129 use fish from a captive colony of guppies derived from wild ancestors sampled from the
130 Aripo River, Trinidad in 2008 and subsequently maintained at high population size
131 (with no deliberate inbreeding or selection). We have validated the use of standardised
132 'open field trials' (OFTs) for testing (acute) behavioural stress responses in this species
133 (Houslay *et al.* 2018), and demonstrated significant additive genetic (co)variance
134 underpinning variation in risk-taking, exploratory, and 'flight' type components of the
135 behavioural stress response using this testing paradigm (White & Wilson 2018; White,
136 Houslay & Wilson 2018). We have also demonstrated, using a non-invasive waterborne
137 hormone sampling method, that individuals differ significantly in their GC (specifically,
138 free circulating cortisol) response to an acute stressor (handling, coupled to short term
139 isolation and confinement; Houslay *et al.* 2019) and that, on average, this physiological
140 response declines with repeated stressor exposure (consistent with habituation).
141 Nothing is known about the genetic basis of variation in these physiological traits, or
142 about their integration (phenotypically or genetically) with behavioural components of
143 the stress response.
144 First, we combine OFT results with complementary 'emergence trials' (ET) and
145 'shoaling trials' (ST) to characterise among-individual and genetic variation in the
146 behavioural stress response. Second, we characterise the physiological stress response
147 and its rate of habituation by assaying GC levels following first and third exposure to a
148 handling and confinement stressor (see methods). Utilising repeated behavioural and
149 physiological testing of individual fish within a known pedigree structure, we are able to
150 estimate the repeatable (among-individual) component of phenotypic (co)variance in
151 these stress response traits, and then determine the additive genetic contribution to this
152 (**G**; the genetic variance-covariance matrix for this suite of stress response traits). We

153 predict that individual traits will be heritable and that **G** will contain strong genetic
154 correlation structure between behavioural and physiological components of the stress
155 response consistent with genetic integration. We also predict that **G** will be dominated
156 by a single major axis of genetic variation in multivariate trait space, but are more
157 circumspect about how that might look. The 'stress coping style' model (Koolhaas *et al.*
158 1999) predicts variation in the type of response to stress. Simplistically, as originally
159 proposed this verbal model posits that individuals (or genotypes) perceive equal
160 degrees of stress but differ in how this manifests phenotypically: genotypes at one end
161 of the axis having 'reactive' behavioural phenotype (e.g., freezing behaviour) coupled to
162 lower GC levels, while the 'proactive' end is characterised by more active 'fight or flight'
163 behaviour coupled to higher GC levels. However, previous analyses of this population
164 suggest variation may be more in the stress responsiveness than coping style (Houslay
165 *et al.* 2018; Prentice *et al.* 2020; White, Pascall & Wilson 2020). That is to say, some
166 individuals (or genotypes) perceive the trial as a more severe stimulus and exhibit more
167 characteristic stress behaviours (e.g., flight and/or freezing, thigmotaxis) while others
168 show more typical 'unstressed' behavioural profiles (e.g., exploration of the arena). In
169 this scenario we predict high GC levels to co-occur with characteristic stress behaviours.
170

171 **Results**

172 In total we obtained (multivariate) behavioural data from 7,637 trials (3,812 OFTs,
173 1,548 ETs and 1,039 STs) on 1,518 individual fish. The number of individuals
174 phenotyped (OFTs = 1,487, ETs = 806, STs = 532) and the mean number of observations
175 per fish (OFTs = 2.6, ETs = 1.9, STs = 2.0) varied across the behavioural data types. All
176 fish were contained within a genetic pedigree structure comprising maternal full-
177 sibships nested within paternal half-sibships. This structure was produced via multiple
178 round of breeding work and has a maximum depth of 5 'generations'. Some of the OFT
179 data have already been used in studies of the evolutionary genetics of personality

180 (White & Wilson 2018; White *et al.* 2018), but here we extend that dataset and use it in
181 conjunction with other behavioural and physiological measures for different purposes.
182 We also obtained 1,238 waterborne assays of cortisol levels for 629 fish (almost all from
183 the final generation). The handling and confinement stressor applied for this assay was
184 performed 3 times (at 48h intervals) for all fish tested, but the holding water sample
185 was only processed for GC content at two time points (the first and last confinement,
186 subsequently $Cortisol_1$ and $Cortisol_3$). Full details of husbandry, phenotyping and
187 analysis are provided in Materials and Methods.
188
189 **Genetic variance in behavioural components of the stress response**
190 Behavioural data were extracted from OFTs, ETs and STs using video tracking of fish (as
191 described in White, Kells & Wilson 2016; Houslay *et al.* 2018). Time to emerge from the
192 shelter ('*emergence time*') was extracted from ETs and natural log (ln) transformed for
193 analysis, while *shoaling tendency* was calculated from STs as the time spent in the third
194 of the tank closest to a same-sex shoal (which was visible but physically separated)
195 minus the time spent in the third of the tank farthest from the shoal. The OFT, ET and ST
196 testing paradigms are all considered to assay behavioural components of the stress
197 response in the broad sense, as each test starts with the capture and transfer of the focal
198 fish into a novel, brightly lit, arena away from their home tank and familiar tank mates.
199 Three traits were defined from the OFT and measured by videotracking for 4m 30s after
200 an initial 30 second acclimation period on transfer into arena. *Track length* (distance
201 swum), *area covered* (as a proportion of the arena floor area), and *time in the middle*
202 (i.e., in the central area of the open field arena away from the tank walls, which is
203 assumed to be perceived as riskier; e.g., Houslay *et al.* 2018) were recorded. Note low
204 values of *time in the middle* imply thigmotaxis (i.e., tendency to avoid exposure to
205 potential threats by hugging walls). All three of these OFT traits are repeatable and
206 heritable in this population (White *et al.* 2016, 2018; Houslay *et al.* 2018; White &

207 Wilson 2018). However, the absence of a strong positive correlation between *track*
208 *length* and *area covered* (Fig 1A) is notable and potentially biologically informative; if
209 fish moved randomly with respect to direction in the arena then *area covered* would
210 increase monotonically (to an asymptote at 100%) with *track length*. A possible
211 explanation is that a long *track length* can sometimes arise from a (putatively) less
212 stressed fish exploring the arena (fish 1 in Fig 1B) and sometimes a (putatively) more
213 stressed fish exhibiting a typical 'flight' response (Fish 4 in Fig 1B). These can be
214 discriminated based on whether, in a given trial, high track length is associated with
215 high *area covered* and *time in the middle* (exploration) or the converse (flight response).
216 To quantitatively discriminate between exploratory behaviour and flight
217 responses we derived a new trait, '*relative area covered*'. We used a simple simulation
218 procedure (see Methods) to predict *expected area covered* for any given *track length*
219 under a null 'random swim' within the arena (Fig 1C). *Relative area covered* is then
220 calculated as *observed area covered – expected area covered* given the *track length* (Fig
221 1D) and will be high for fish engaging in exploration, and low for an obvious 'flight'
222 response manifest as rapid swimming around the tank walls.
223 Pedigree-based 'animal models' (Wilson *et al.* 2009) were used to test for and
224 estimate additive genetic variation in each of the five behavioural traits while
225 controlling statistically for social housing group and non-genetic sources of among-
226 individual variance (as well as several fixed effects; see methods for full details). These
227 confirmed the presence of significant additive genetic variation for the *relative area*
228 *covered* trait, as well as for *track length* and *time in the middle* (as expected from
229 previous findings; White & Wilson 2018; White *et al.* 2018) as well as *emergence time*
230 (Table 1). With the exception of *shoaling tendency*, heritabilities (conditional on fixed
231 effects; see methods) are low to moderate (range of 8-17%; Table 1) but within the
232 expected range for behaviours (Stirling, Réale & Roff 2002). We detected no additive

233 genetic variance for *shoaling tendency* (Table 1), despite there being repeatable
234 differences among individuals; R = 0.19 ± 0.04; $\chi^2_{0,1}$ = 20.01, P < 0.001).
235
236 **Genetic variance in physiological components of the stress response**
237 Using a series of nested bivariate animal models, we tested for additive genetic variance
238 in cortisol levels (ln-transformed) following stressor exposure (handling and
239 confinement) and for genotype-by-environment interaction (GxE). In this context, the
240 environment (E) is the trial number in each fish's stress trial series (i.e. cortisol level
241 following stressor trial 1 or 3). Any GxE present can therefore be interpreted as genetic
242 variance for habituation to the stressor, given that the average cortisol level was lower
243 following exposure to the third stressor than the first (ln transformed ng/hr, mean ± SE;
244 $Cortisol_1$ = 8.50 ± 0.05, $Cortisol_3$ = 8.05 ± 0.06, Wald $F_{1,12.9}$ = 120.5, P < 0.001; see
245 methods for explanation of units). We first modelled $Cortisol_1$ and $Cortisol_3$ as distinct
246 response variables in a bivariate framework assuming no GxE (such that we constrain
247 $V_{A\text{-}Cortisol1} = V_{A\text{-}Cortisol3}$ and the cross context additive genetic correlation $r_{A\text{-}Cortisol1,Cortisol3}$ =
248 1). This model revealed a significant additive genetic component to variation among
249 individuals in their cortisol levels following stressor exposure ($\chi^2_{0,1}$ = 6.58, P = 0.005).
250 Expanding the model to allow GxE (by estimating separate genetic variances for
251 $Cortisol_1$ and $Cortisol_3$ and allowing the cross-context genetic correlation to deviate
252 from +1) provides a significantly better fit to the data (χ^2_2 = 9.65, P = 0.008), meaning
253 GxE is present. This can be viewed as genetic variance for habituation to repeated
254 stressor exposure, or as a change in genetic variance for cortisol from the first to the
255 third sampling (Figure 2). These are two perspectives of the same phenomenon; a
256 reduction in additive genetic variance between the first stressor ($V_{A\text{-}Cortisol1}$ = 0.076 ±
257 0.028) and the third ($V_{A\text{-}Cortisol3}$ = 0.047 ± 0.029) arises because genotypes with higher
258 than average levels for *$Cortisol_1$* habituate more rapidly (i.e. have more negative reaction
259 norm slopes). Note however that the rank order of the genotypes does not appreciably

change across the two contexts (i.e. genetic reaction norms show little crossing; Figure 2), so there is a strong positive cross-context genetic correlation ($r_{A\text{-}Cortisol1,Cortisol3}$ ± SE = 0.74 ± 0.25).

In this model we also find that variance in cortisol explained by housing group effects is similar across contexts ($V_{Group\text{-}Cortisol1}$ = 0.034 ± 0.013, $V_{Group\text{-}Cortisol3}$= 0.045 ± 0.016), but that residual (unexplained) variance is greater after the third stressor exposure ($V_{R\text{-}Cortisol1}$ = 0.166 ± 0.021, $V_{R\text{-}Cortisol3}$ = 0.229 ± 0.025). In combination, the changes in both additive genetic and residual variance between the two contexts lead to appreciably higher heritability for cortisol levels following the first stressor exposure relative to the third ($h^2_{Cortisol1}$= 0.275 ± 0.093, $h^2_{Cortisol3}$= 0.146 ± 0.088).

Testing for genetic integration and identifying the major axis of genetic (co)variance

There is strong evidence for phenotypic integration of *Cortisol* with behaviour at the among-individual levels (see Table S5). To test for and characterise the hypothesised genetic integration between behavioural and physiological components of the stress response, we built a multivariate animal model to estimate **G**. We excluded *shoaling tendency* given the absence of detectable genetic variance in the univariate model. We also elected to treat cortisol as a single trait (allowing for a fixed effect of stressor exposure number (1 vs 3) on the mean). Although the above analysis demonstrates GxE for cortisol, the strong positive cross-context genetic correlation justifies collapsing $Cortisol_1$ and $Cortisol_3$ into a single trait to maximise statistical power to detect any genetic covariance with behaviour.

Our final model contained 5 response traits: *relative area covered*, *time in the middle* and *track length*, (ln-transformed) *emergence time*, and (ln-transformed) *Cortisol* (now treated as two repeats of a single trait). We standardised all (transformed) traits to standard deviation units, to assist multivariate model fitting and to prevent

287 eigenvectors of **G** (see below) being dominated by traits with higher variance in
288 observed units. To simplify interpretation of **G** we also multiplied *emergence time* by -1
289 after transformation. Thus high values denote rapid emergence from the shelter.
290 The resultant estimate of **G** (Table 2) contains significant additive genetic
291 covariance structure overall (Likelihood Ratio Test of the full model vs. a reduced model
292 **G** that contains variances but not covariances: $\chi^2_{10} = 36.79$, $P < 0.001$). Strong genetic
293 covariance/correlation estimates, both positive and negative, were found for a number
294 of trait pairs and were deemed statistically significant (based on the bootstrapped 95%
295 confidence intervals not crossing zero). We find strong genetic covariance among all 3
296 OFT traits: *track length* shows significant negative genetic covariance with both *relative*
297 *area covered* and *time in the middle*, and *relative area covered* and *time in the middle*
298 show significant positive genetic covariance with one another. *Track length* also has a
299 significant positive genetic covariance with –*ln(emergence time)*. Ln-transformed
300 *Cortisol* shows significant negative genetic covariance with *time in the middle* (Figure 3).
301 Ln-transformed *Cortisol* also covaries negatively with both *relative area covered* and –
302 *ln(emergence time)*, and positively with *track length*, although these covariances were
303 not significantly different from zero (based on 95% confidence interval).
304 Eigen decomposition of **G** provides a more holistic view of the genetic
305 covariance structure and the level of integration among traits. Here the major axis (first
306 principal component, PC1, with 95% confidence intervals from 5000 bootstrap
307 replicates) explains 57.8% (47.3%, 79.6%) of the genetic variance in multivariate
308 phenotype (PC2 = 25.7% [14.3, 37.6]; PC3 = 9.3% [4.7, 16.5]; PC4 = 5.6% [0, 9.2]; PC5 =
309 1.6% [0, 3.6]). All traits except *emergence time* load significantly on this axis (Figure 4).
310 *Relative area covered* and *time in the middle* load in one direction, while *track length* and
311 *Cortisol* load in the other direction. This structure is suggestive of a single major axis of
312 genetic variation in integrated stress response, where genotypes at one end of this axis
313 can be considered to have 'weaker' behavioural stress responses to the OFT assay (i.e.,

314 swim shorter distances, spend more time in the central area of the tank, and exhibit
315 exploratory swimming patterns that cover greater areas relative to their distance
316 swum) and 'weaker' physiological responses to stress (i.e., produce lower cortisol levels
317 in response to the stressor). Meanwhile, genotypes at the other end of this axis can be
318 considered to have 'stronger' behavioural responses to stress (i.e., swimming further,
319 while spending more time close to the tank edges, and covering less area relative to
320 their distance travelled in OFTs) and 'stronger' physiological responses to stress (i.e.,
321 produce higher cortisol levels in response to the stressor).
322
323 **Discussion**
324 In this study we sought to determine whether – and to what extent – there exists
325 genetic variation for, and integration between, behavioural and physiological
326 (endocrine) components of the stress response. Our results provide three main novel
327 insights. First, we find that genetic variation does underpin individual differences in
328 both behavioural and physiological components of the stress response. Second, we find
329 genetic covariance structure among these behavioural and physiological traits,
330 indicating that they are indeed genetically integrated. Thirdly, having identified the
331 structure of the major axis of **G** we suggest that it is more readily interpreted as an axis
332 of genetic variation in stress responsiveness than in stress coping style (although we
333 acknowledge the distinction may be somewhat subjective). Overall, by estimating the
334 genetic covariance structure among traits we find the first quantitative genetic support
335 to date for the hypothesis of evolutionary integration between behavioural and
336 endocrine components of the stress response.
337 We find heritable (co)variation in and among behaviours assayed in the open
338 field trial (OFT), including the derived trait *relative area covered*. The latter trait, derived
339 by considering an appropriate biological null model of the relationship between *track*
340 *length* and (absolute) *area covered* serves as a useful proxy for exploratory behaviour.

341 Here we demonstrate an axis of repeatable and heritable variation that spans from less
342 active but more exploratory swimming patterns (lower *track length* but higher *relative*
343 *area covered* and *time in the middle*) through to a 'flight' type response, characterised by
344 higher activity coupled to thigmotaxis and low exploration (higher *track length*, lower
345 *relative area covered* and *time in the middle*). Although the distinction is both subjective
346 and arguable, we interpret this axis as more consistent with variation in the magnitude
347 of stress responsiveness than in coping style. This is because, while low *track lengths*
348 could arise from stressed fish exhibiting 'reactive' freezing behaviour, we would not
349 necessarily expect it to be associated with high *time in the middle* (i.e. reduced
350 thigmotaxis). Given the wide use of OFTs in biomedical research (e.g., Rex *et al.* 1998;
351 von Horsten, Karl & Pabst 2003) as well as in animal behaviour, our phenotyping
352 approach may have broad applicability for discriminating between exploration and
353 stress/anxiety-related behaviours. At the slight risk of introducing semantic confusion,
354 we also note that the OFT paradigm is widely applied to studies of 'shy-bold' type
355 personality variation in fishes (Toms, Echevarria & Jouandot 2010) and other
356 vertebrates (Carter *et al.* 2013; Perals *et al.* 2017). The extent to which behavioural
357 differences deemed characteristic of a 'shy—bold' personality axis (commonly, if not
358 universally defined as repeatable variation in response to perceived risk; Wilson *et al.*
359 1994) should be viewed as equivalent to variance in behavioural stress responsiveness,
360 or coping style, is a matter of debate (see Boulton *et al.* 2015). We view these as
361 overlapping – if not necessarily identical – concepts, and suggest it will generally be
362 prudent to empirically assess the structure of OFT variation for any system rather than
363 assume a priori that it will match a preferred verbal model (see also White *et al.* 2016;
364 Houslay *et al.* 2018; White, Pascall & Wilson 2020).
365 We also find significant heritable variation in *emergence time*, although this is
366 not tightly integrated with OFT traits in **G** in the manner we had expected. Faster
367 emergence in typically interpreted as reflecting a lack of fear of the open arena (i.e.,

368 greater 'boldness'; see Burns 2008). In our study the qualitative pattern in **G** runs
369 counter to this, if we consider that boldness and behavioural stress responsiveness are
370 broadly analogous. Thus genotypes predisposed to shorter *emergence time* are
371 associated with greater *track length* and, albeit non-significantly, lower *relative area*
372 *covered* and *time in the middle* during OFT trials. That is, shorter *emergence time* is
373 associated with the putatively more stressed 'flight' end of the behavioural axis revealed
374 in the OFT. Our interpretation of this result is that at least some genotypes (and
375 individuals) likely perceive the shelter area as less safe than the open arena. This
376 possibility was also suggested by an earlier finding that some individual guppies
377 decrease (rather than increase) shelter use following simulated predation events
378 (Houslay *et al.* 2018). If general to other systems, unexpected results such as these may
379 have important consequences for interpretation of personality tests, and correlations
380 (or lack thereof) between behaviours assayed across test paradigms (Carter *et al.* 2013).
381 Tendency to shoal varies among individuals but is not detectably heritable.
382 Though not generally considered a stress-response trait *per se*, shoaling is an anti-
383 predator behaviour in guppies (Herbert-Read *et al.* 2017). We had therefore predicted
384 that heightened perception of risk in the open field might also be associated with
385 increased shoaling tendency. This was not the case at the among-individual level (Table
386 S5), while the absence of detectable genetic variance meant that we could not test this
387 prediction in **G**.
388 We find strong evidence of significant additive genetic variance in a key
389 physiological component of the stress response: waterborne cortisol concentrations
390 following exposure of the fish to a handling stressor. Our findings suggest that
391 previously detected differences among individuals in cortisol response to a stressor
392 (Houslay *et al.* 2019) are primarily attributable to genetic effects, with the estimated
393 heritability ($h^2 = 0.26$) being almost 75% of the individual–level repeatability ($R = 0.35$)
394 for ln-transformed *Cortisol*. In addition, by adopting a reaction norm approach to

395 modelling stress physiology, as recently advocated by ourselves (Houslay *et al.* 2019)
396 and others (e.g., Fürtbauer *et al.* 2015; Hau & Goymann 2015; Taff & Vitousek 2016), we
397 detect GxE reflecting genetic differences in the extent of habituation to the stressor over
398 repeated exposures. This result is potentially important since poor habituation of the
399 hypothalamic-pituitary-adrenal/interrenal (HPA/I) response to repeated or ongoing
400 stressors can lead to well documented health problems in human and animal
401 populations (Segerstrom & Miller 2004; Koolhaas 2008; Romero, Dickens & Cyr 2009;
402 Mason 2010). Our detection of heritable variation in the degree of habituation to
403 stressors raises the possibility of developing targeted selection strategies to improve
404 welfare in captive populations (e.g., Frankham *et al.* 1986; Muir & Craig 1995; Oltenacu
405 & Algers 2009).
406 Our findings also highlight that there is greater additive genetic variance (and
407 heritability) for cortisol levels following the first exposure to the stressor than following
408 the third. This pattern, which occurs because genotypes that produce the highest
409 cortisol response at first exposure also show the most marked habituation, is consistent
410 with the idea of cryptic genetic variance (Paaby & Rockman 2014) being 'released' by
411 exposure to novel, and so potentially stressful, environments (Ledón-Rettig, Pfennig &
412 Crespi 2010; Ledón-Rettig *et al.* 2014; Berger *et al.* 2011). All else being equal, it also
413 means that selection on cortisol levels following stressor exposure should induce a
414 stronger evolutionary response in naïve relative to habituated fish. However, the strong
415 positive cross-environment correlation means that the ranking of genotypes with regard
416 to their cortisol responses is consistent across repeated stressor exposures. Thus
417 selection on the (average) GC response would result in a correlated evolutionary
418 response of habituation rate, and *vice versa*.
419 Considering all traits together, **G** shows evidence of genetic integration between
420 behavioural and endocrine components of the stress response. As noted with respect to
421 OFT behaviours, we consider that the major axis of **G** is best interpreted as genetic

422 variance in the magnitude of stress responsiveness. Accordingly, genotypes tending to
423 show (putatively) more stressed 'flight' type behaviour in the OFT (i.e. thigmotaxis, high
424 *track length*, low *relative area covered*) also produce higher levels of cortisol following
425 the handling and confinement stressor. The fact that the only significant bivariate
426 correlation including cortisol is with *time in the middle* (negative correlation) suggests
427 that thigmotaxis is a particularly strong indicator of high stress responsiveness. We note
428 that our interpretation that **G** is dominated by variation in the magnitude of stress
429 responsiveness does not mean the 'coping style' model has no merit. Indeed, a
430 subsequent 'two-tier' iteration of the coping style model proposed that variation in the
431 magnitude of the stress response (termed 'stress reactivity') could be viewed as a
432 second axis, distinct from variation in the 'type' (or style) of response (Koolhaas *et al.*
433 2010).

434 Here **G** certainly contains substantial variation not explained by its leading axis
435 and this may point towards interesting avenues for further study. For instance, while
436 *relative area* and *tracklength* are negatively correlated and both load significantly on the
437 major axis of **G**, there is also much variation in the former at low to moderate *track*
438 *lengths* (illustrated by e.g., fish 1 and 4 in Figure 1). It seems plausible, if speculative,
439 that this could be caused by a tendency of some individuals to show initial immobility
440 (i.e. freezing) in response to the stressor but then recover quickly and begin explorative
441 swimming. Others may show an initial flight response before relatively rapid recovery
442 to explorative swimming. Such differences, if present, would not be detectable by simple
443 principal components analysis of covariance structures based on whole trial data (as
444 here and, for example, Van Reenen *et al.* 2005). Rather this would require more detailed
445 modelling of within-trial behavioural dynamics.

446 The genetic integration of behaviour and physiology detected here is consistent
447 with the idea that correlational selection in the past has led to the coevolution of these
448 stress response components. Covariance structure in **G** will modify, and potentially

449 constrain, evolutionary responses to selection - whether natural or artificial. Here we
450 have no direct knowledge of how contemporary selection is acting in the wild. Nor do
451 we know whether it might be changing as a consequence of anthropogenic stressors.
452 Thus we cannot comment directly on how **G** will shape future evolution of the guppy
453 stress response beyond noting that selection on behaviour will cause correlated
454 evolution of endocrine physiology (and vice versa). Nonetheless, while it seems
455 reasonable to expect that current integration of stress response in natural populations
456 should be broadly adaptive, this seems less likely in captive populations (at least for
457 species without a long history of domestication and opportunity for adaptation to
458 artificial environments). We know that prolonged, chronic activation of stress response
459 pathways (notably the HPA(I) axis) frequently disrupts health and survival in captive
460 animals (Huether 1996; Boonstra 2013). It may be that more stress-responsive
461 genotypes are disadvantaged in novel artificial conditions (e.g., if acute stress
462 responsiveness positively predicts susceptibility to chronic stress). However, even if
463 true this would not imply high (acute) stress-responsiveness was also disadvantageous
464 in the wild. Since natural selection should purge alleles that are universally detrimental,
465 it seems more plausible that genetic variation along the major axis described here is
466 maintained by some form of selective trade-off (as widely hypothesised for maintenance
467 of personality variation; e.g., Stamps 2007; Wolf *et al.* 2007; Réale *et al.* 2010). For
468 instance, genotypes susceptible to harm under chronic stressor exposure will likely
469 persist in populations if they also confer advantages under an acute stress challenge. In
470 natural populations not only is exposure to acute stressors more common than to
471 chronic stressors, but also selection through chronic stress exposure may be conditional
472 on (and subsequent to) surviving acute challenges (such as predator attacks).
473
474 **Conclusions and future directions**

475 Here we find evidence for genetic variation in – and integration of – behavioural
476 and physiological (endocrine) components of the stress response. Overall we consider
477 the structure of **G** to be more consistent with a continuous axis of variation in acute
478 stress responsiveness than with the widely invoked 'reactive – proactive' model of
479 variation in stress coping style (Koolhaas *et al.* 2007). This interpretation rests largely
480 on the structure of behavioural variation revealed by the OFT, which is dominated by an
481 axis running from genotypes that are more stress responsive (rapid 'flight' behaviour
482 coupled to thigmotaxis) to those that show more exploratory behaviour expected from a
483 (putatively) unstressed fish. Endocrine traits align with this axis: genotypes exhibiting
484 'flight' behaviour show higher cortisol levels (and exhibit faster habituation of GC
485 physiology) when subject to repeated handling and confinement stressors.
486 Our results suggest that correlational selection in the past has likely shaped the
487 multivariate stress response, and that continued evolution of stress-related behaviour
488 will have consequences for glucocorticoid physiology and *vice versa*. Determining
489 contemporary selection on the stress response, and testing the possibility that genetic
490 variation is maintained by fitness trade-offs, is thus an obvious – if empirically
491 challenging – next step to understanding the functional importance of genetic variation
492 in wild populations. In a more applied context, integration of behavioural and endocrine
493 stress-response components at the genetic level has potential utility for genetic
494 improvement of managed populations. Specifically, it may be possible to identify non-
495 invasive, high throughput, behavioural biomarkers and target them in selection schemes
496 to reduce chronic activation of the HPA/I axis and its attendant deleterious effects.

497 Table 1: Estimated variance components, along with adjusted heritability, for each trait as estimated in a univariate model (± standard error). Chi-
498 square test statistics and p-values are provided for the pedigree term, testing for the presence of significant additive genetic variance (V_a).

Trait	V_a	V_{pe}	V_{group}	$V_{residual}$	h^2	$\chi^2_{0,1}$	P
Relative area covered	26.35 ± 9.37	72.42 ± 9.28	33.33 ± 6.86	205.22 ± 5.90	0.08 ± 0.03	20.7	<0.001
Time in the middle	588.48 ± 139.61	554.58 ± 109.57	203.29 ± 53.41	2070.06 ± 60.08	0.17 ± 0.04	53.7	<0.001
Track length	26832.64 ± 5925.25	32204.05 ± 4868.88	9956.4 ± 2626.41	93921.54 ± 2711.28	0.16 ± 0.03	86.3	<0.001
log Emergence time	0.12 ± 0.05	0.06 ± 0.06	0.05 ± 0.02	1.07 ± 0.05	0.09 ± 0.04	23.2	<0.001
Shoaling tendency	0 ± 0	2457.36 ± 570.96	708.87 ± 316.30	9900.95 ± 622.10	0 ± 0	0	0.5
log Cortisol	0.07 ± 0.02	0.02 ± 0.02	0.01 ± 0.01	0.15 ± 0.01	0.26 ± 0.08	22.0	<0.001

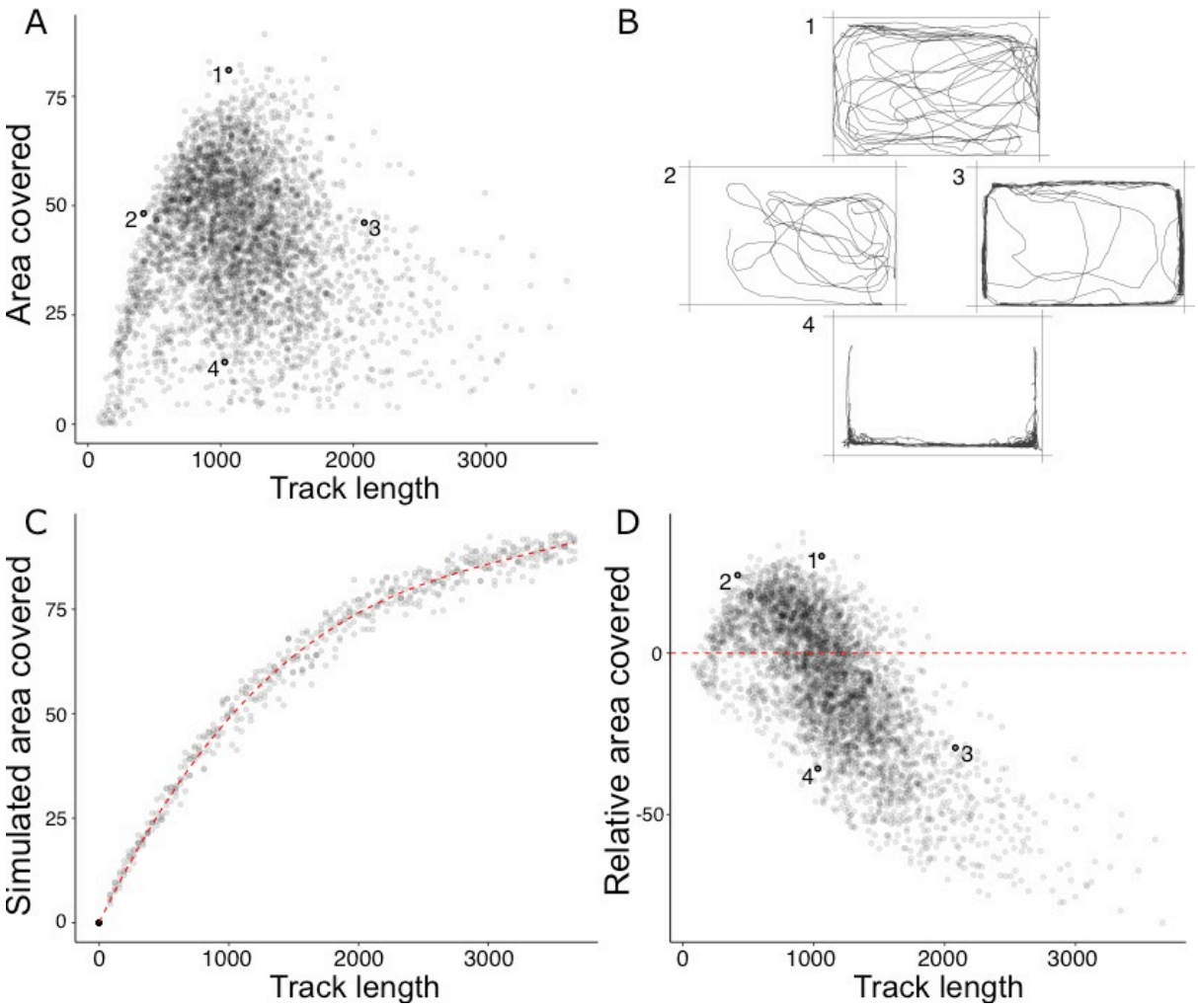
499

500

Table 2: Additive genetic covariance-correlation matrix (G) from the full multivariate animal model. Genetic variances provided on the shaded diagonal, with genetic covariances below and genetic correlations above. 95% confidence intervals in parentheses are estimated from 5000 bootstrapped replicates. Where the confidence intervals for any estimate do not cross zero the estimate is considered statistically significant (at the 0.05 alpha level) and are shown in bold.

	Relative area covered	*Time in the middle*	*Track length*	-log *Emergence time*	log Cortisol
Relative area covered	**0.074 (0.029,0.122)**	**0.761 (0.549,0.955)**	**-0.506 (-0.758,-0.184)**	-0.503 (-1.394,0.256)	-0.414 (-1.035,0.225)
Time in the middle	**0.075 (0.031,0.124)**	**0.130 (0.062,0.191)**	**-0.554 (-0.774,-0.295)**	-0.117 (-0.791,0.531)	**-0.686 (-1.165,-0.220)**
Track length	**-0.048 (-0.086,-0.008)**	**-0.070 (-0.116,-0.022)**	**0.121 (0.067,0.171)**	0.559 (-0.026,1.256)	0.279 (-0.238,0.823)
-log *Emergence time*	-0.038 (-0.083,0.014)	-0.012 (-0.070,0.040)	**0.055 (0.001,0.106)**	**0.079 (0.011,0.149)**	-0.177 (-0.910,0.560)
log *Cortisol*	-0.038 (-0.091,0.011)	**-0.082 (-0.138,-0.021)**	0.032 (-0.026,0.085)	-0.017 (-0.073,0.038)	**0.111 (0.036,0.191)**

Figure 1: The lack of a strong positive relationship between observed *track length* and *area covered* (panel A), is initially puzzling given expected autocorrelation and that both are used as positive indicators of exploratory (or 'bold') behaviour. Inset examples of OFT tracks from 4 individuals (panel B) shed light on this. Fish 1 and 2 appear to be exploring the tank, while 3 and 4 are engaging in stereotypical 'flight' behaviour characterised by strong thigmotaxis (remaining close to tank walls) and/or rapid movement along tank walls. As a consequence, individuals 2 and 3 have similar *area covered* during the OFT, but very different *track lengths*. We simulated random movements to define an expected null relationship between *area covered* and *track length* (panel C; dashed red line shows the fourth order polynomial model fit; see Appendix A). The polynomial regression was then used to predict the expected area covered under random movement for each trial's observed *track length*, and the '*relative area covered*' was calculated as the observation minus this prediction. Panel D shows the resultant *relative area covered* plotted against *track length* for all trials (dashed red line at *relative area covered* = 0, shows where individuals of any *track length* are expected to lie if they move randomly with respect to direction). From this it is apparent that fish 1 and 2 have high *relative area covered*, while 3 and 4 do not.



523

524

525

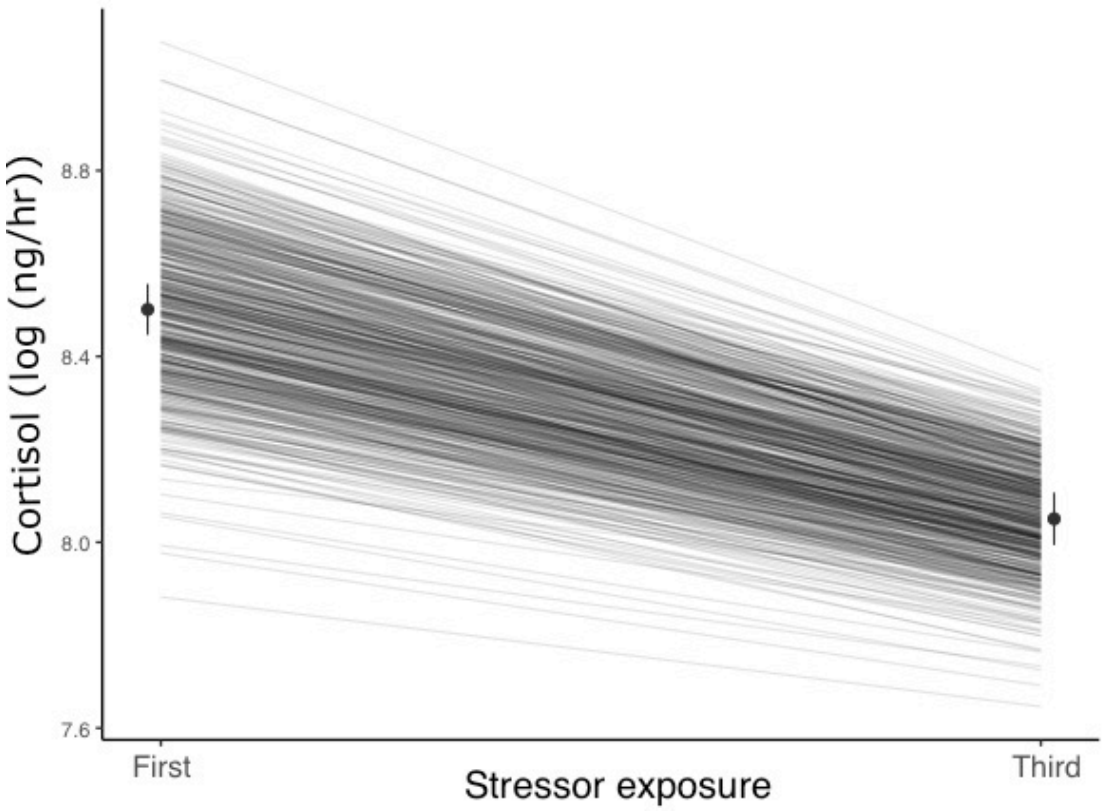
526

527

528

529

530 Figure 2: Guppies habituate to the waterborne sampling procedure, as shown by a decline in average
531 log-transformed cortisol level (ng/hr) following stressor exposure between first and third exposures.
532 Black circles and associated bars denote predicted population means (± standard error) from mixed
533 model analysis. Grey lines depict the predicted genetic reaction norms across repeated stressor
534 exposure for each individual. Weak, but statistically significant GxE is manifest as variance in the
535 genetic reaction norm slopes (i.e. lack of parallelism) and results in a slight reduction of genetic
536 variance for cortisol at the third exposure relative to the first.



537

538

Figure 1: The additive genetic relationship between log-transformed cortisol (y-axis) and four behaviours (a, *relative area covered*; b, *time in the middle*; c, *track length*; d, *-log emergence time*). Points show (predicted) bivariate genetic deviations from the population means, plotted for those individuals in the pedigree with cortisol data. In each case the black ellipse depicts the 'shape' of the relationship as given by the point estimate of **G**. Specifically it encompasses the area expected to contain 95% of the bivariate genetic distribution for the population. Grey ellipses denote the corresponding areas defined from 5000 bootstrapped replicates of **G**, and so highlight the uncertainty around these bivariate distributions.

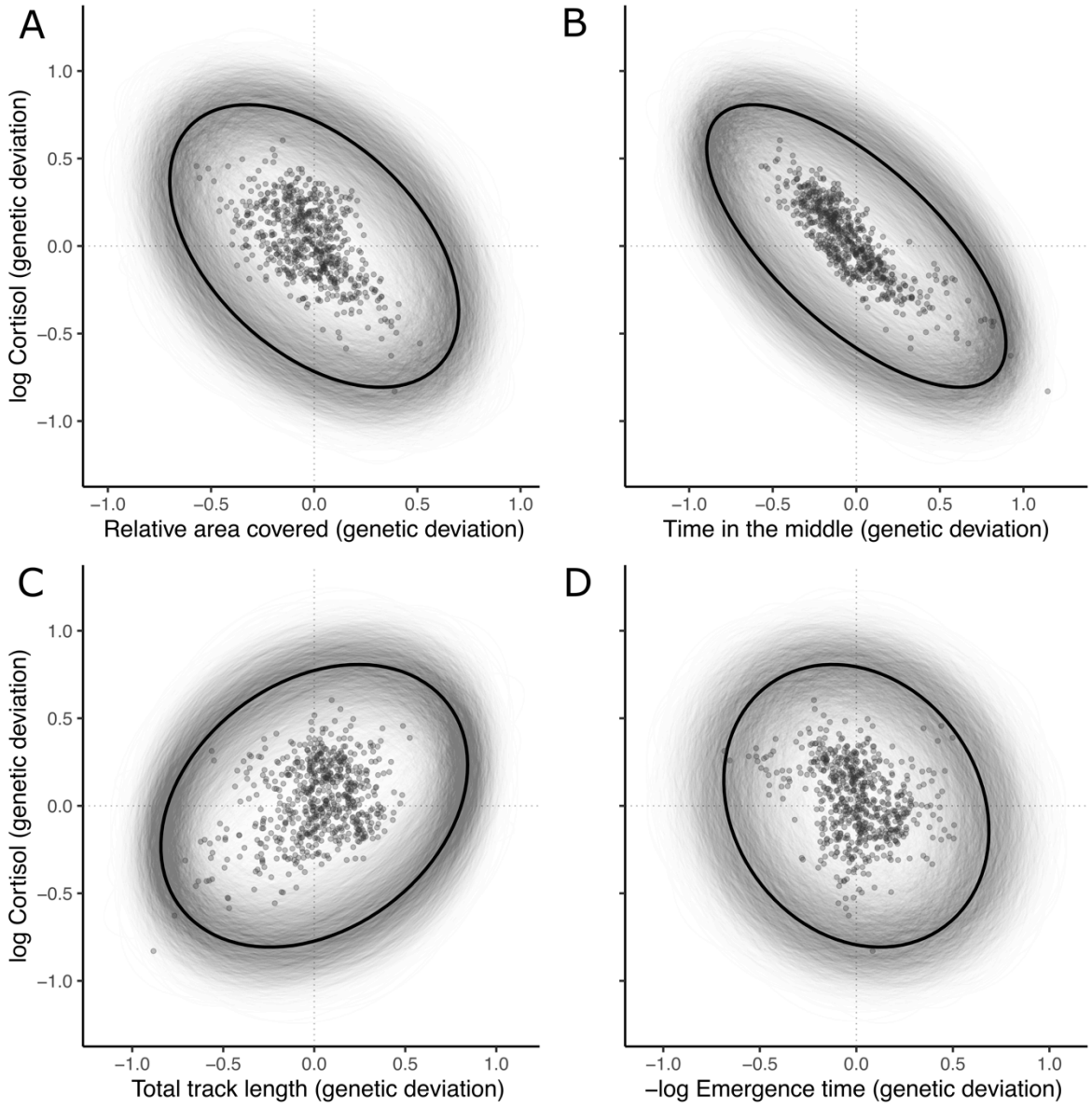
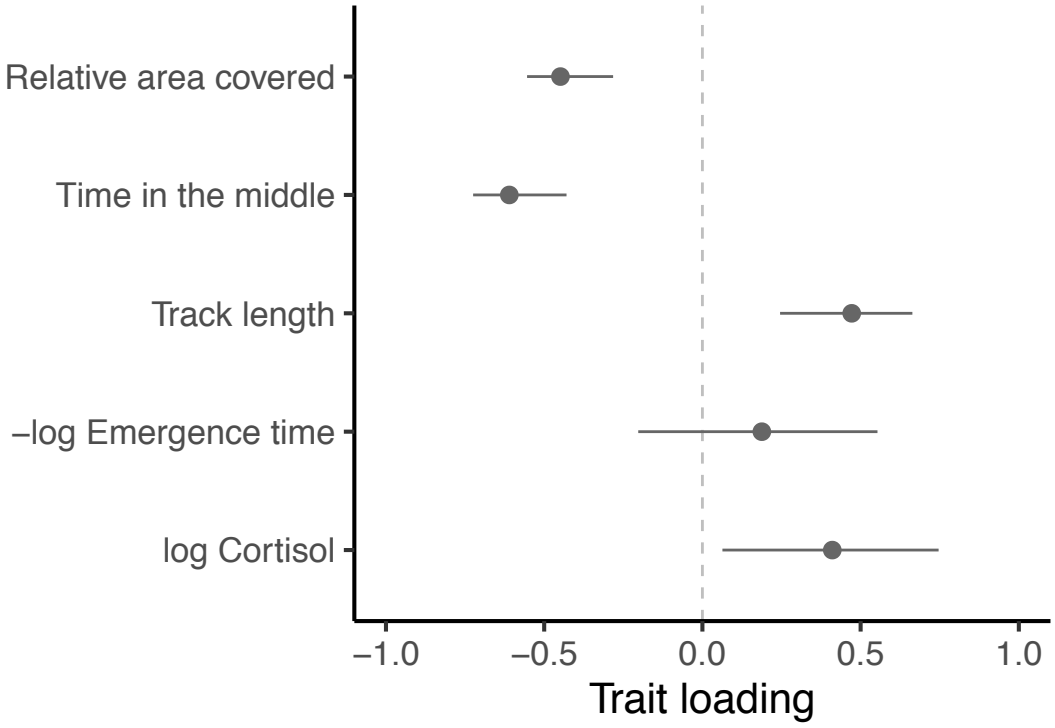


Figure 2: Trait loadings from the first eigen vector (principal component) of **G**. This axis explains 57.8% of the genetic (co)variation found in the focal behavioural and physiological components of the stress response in our guppy population. Points show trait loadings from the first eigen vector of our estimate of G, with bars representing 95% confidence intervals on each loading (calculated from 5000 bootstrapped replicates of the model).



555 **Materials and methods**

556

557 *Husbandry and breeding*

558 We used fish taken from our captive population housed at the University of Exeter's Penryn campus,
559 which is descended from wild fish collected in 2008 from the lower Aripo River in Trinidad. This
560 population has been maintained at a population size of several thousand, and has undergone no
561 deliberate selection or inbreeding. All fish are fed to satiation twice daily (0800 – 1000h and again at
562 1600 – 1800h) using commercial flake food and live *Artemia nauplii*. Water temperature is maintained
563 at 23-24°C in well-aerated closed system tank stacks that undergo 25% water changes each week and
564 with weekly tests for ammonia, nitrate and nitrite levels. Lighting is kept at a 12:12 light/dark cycle.
565 The experiment described in this study was carried out in accordance with the UK Animals (Scientific
566 Procedures) Act 1986 under licence from the Home Office (UK), and with local ethical approval from
567 the University of Exeter.

568 To create our pedigreed sub-population, female fish were sampled haphazardly from the stock
569 tanks and kept in female-only groups for 3 months. Isolation from male contact minimised the chance
570 of females carrying viable sperm from previous matings. For the first generation of offspring, we used
571 a group breeding design (as detailed in White and Wilson, 2018); briefly, females were tagged under
572 anaesthetic (buffered MS222 solution) using visible implant elastomer (VIE) to allow individual
573 identification. We then assigned groups of 4 females to 1 male in 15L breeding tanks (18.5cm x 37cm x
574 22cm), and inspected females daily for high gravidity (swollen abdomens and enlarged 'gravid spots').
575 Heavily gravid females were then isolated in 2.8L brood tanks to give birth (and were returned to the
576 breeding tanks either after producing a brood or two weeks of isolation). Any offspring produced in
577 the breeding tanks were excluded from the experiment as maternal identity could not be positively
578 identified. For the following generations, after 3 months of isolation from males we moved females
579 into individual 2.8L tanks, with 1 male then circulated among 3 females. Males were moved between
580 females every 5-8 days. In this way, females did not have to be moved to brood tanks, and any
581 offspring could be assigned to mothers definitively. In this setup, offspring were moved to a separate
582 brood tank on the day of birth. Note that as the gestation period for guppies is approximately 1 month,

583 any brood produced by a female less than one month after exposure to their designated male was
584 recorded in the pedigree as having unknown paternity.
585 Within 24h of a female producing a brood we recorded her weight (g) and brood size. We kept
586 juvenile fish in full-sib family groups in 2.8L tanks before moving them to 15L 'growth' tanks at an
587 average age of 56 days. At an average age of 133 days (range 59-268) we tagged individuals and
588 placed them into mixed family groups of 16-20 adults (with an even mix of males and females), kept in
589 15L tanks. Note that variation in tagging age arose largely because groups were necessarily
590 established sequentially as sufficient individuals from multiple families reached a large enough size
591 that we deemed the procedure to be safe. Each adult group comprised a mix of fish from different
592 families, reducing the potential for common environment effects to upwardly bias our genetic
593 parameter estimation.
594
595 *Overview of behavioural phenotyping*
596 Behavioural phenotyping commenced at least one week after tagging. In all trials, we filmed
597 movement behaviour of individual fish using a Sunkwang video camera equipped with a 6-60mm
598 manual focus lens suspended over the tank. We used the tracking software Viewer II (BiObserve) to
599 extract behavioural data from each recording (detailed below). The tank was lit from below using a
600 light box, and screened with a cardboard casing to prevent external visual disturbance. After each
601 behavioural trial, the individual tested was weighed and then moved to a temporary 'holding tank'.
602 Once a full group (as described above) had been tested, all were moved from the holding tank back to
603 their home tank. We replaced the water in the testing and holding tanks between groups to reduce the
604 build-up of hormones or other chemicals. The first offspring generation experienced 4 repeat open
605 field trials (OFTs) over a 2-week period, with at least 48h between trials. Subsequent generations
606 experienced 4 repeat behavioural trials, alternating 2 OFTs with 2 emergence trials (ETs). For the final
607 2 generations, we extended the OFTs by including a shoaling trial (ST) at the end of each OFT.
608
609 *Open field trials (OFT)* followed the methodology described by White et al (2016). Briefly, we assessed
610 individual behaviour in a 20cm x 30cm tank, filled to a depth of 5cm with room-temperature water

611 from the main supply. We caught fish individually from their home tank, examined them quickly for
612 identification tags, then placed them immediately into the centre of the OFT tank. After allowing 30s
613 for acclimation, we filmed behaviour for 4m30s. Behaviours characterised from the tracking software
614 were *track length* (the total distance the fish moved during the trial; cm), *area covered* (the percentage
615 of 1cm x 1cm grid squares through which the fish moved during the trial; %), and *time in middle* (time
616 spent in a rectangular inner zone which was defined as being the same size as an outer area; seconds).
617
618 *Shoaling trials (ST)* were appended to a subset of OFTs, by positioning a small tank containing 10 stock
619 fish (of same sex as the test subject) next to one end of the OFT tank but with visual cues blocked by a
620 cardboard divider. At the end of the normal OFT, we removed this divider slowly, allowing the focal
621 animal to have visual contact with the shoal. We began recording the shoaling trial 30s after removing
622 the divider in order to limit any artefacts of slight disturbance. (Note that we used a further cardboard
623 casing around the shoaling tank to avoid any additional external visual stimulus). We then recorded
624 behaviour of the test fish for an additional 3 minutes. We characterised *shoaling tendency* via the
625 tracking software by subdividing the tank area into 3 equal-sized rectangular areas: one next to the
626 tank holding the group of same-sex fish, one farthest from this group, and the central area. We then
627 calculated *shoaling tendency* as the time spent in the 1/3 area closest to the same-sex group after
628 subtracting the time spent in the 1/3 area farthest away. The decision to use a single-sex shoal aimed
629 to reduce any effects of (potential) mate preference and/or avoidance, but also this necessitated
630 replicate arena setups allowing male and female individuals from each group to be tested in the
631 OFT/ST concurrently. We randomised which tank was used for each sex in each group and recorded
632 this information.
633
634 *Emergence trials (ET)* followed the methodology described by White *et al.* (2016). Briefly, we tested
635 individuals in a 20cm x 40cm tank, filled to a depth of 8cm with room-temperature water from the
636 main supply. A 10cm section of the tank length was walled off creating a shelter area (20cm x10cm),
637 the walls and floor of which were painted black. The focal fish was placed into the shelter area and
638 allowed to acclimate for 30s, at which point we opened a sliding door to allow access to the rest of the

639 tank, which was brightly lit from below and otherwise bare. *Time to emerge* (in seconds) was recorded
640 by the tracking software automatically as the fish exited the shelter area and emerged into the open
641 tank section. Trials were ended either at emergence or at 15 min if the fish had not emerged by that
642 point; in the case of non-emergence, fish were given the maximum value (i.e., 900s).
643
644 *Derivation of 'relative area' from OFT trials*
645 The '*area covered'* variable assayed in the OFT is calculated in BiObserve by dividing the arena (i.e., the
646 total area of the tank as viewed from the camera) into 1cm x 1cm grid squares. The path taken by the
647 fish during observation is then used to determine what proportion of these grid squares the fish
648 entered. However, we sought to derive a measure of '*relative area*' that describes whether a fish covers
649 a large, or small area relative to its observed *track length.*
650 To do this we simulated 'random swims' within the arena across the observed range of *track*
651 *lengths*. We first selected 40 OFT results at random from our total data set and extracted the
652 coordinates of the fish in each frame from the raw tracking file, creating a set of x and y movements
653 and their associated distances. As original coordinates were recorded in pixels we used the calibration
654 of the software to convert to cm units. We then use a 'random walk' algorithm to select a movement
655 (i.e., step size and direction) from this observed distribution at random, and calculate the new
656 coordinates. If the movement keeps the 'fish' within the bounds of the 'tank' (i.e., defined as a 20cm x
657 30cm arena), the movement is accepted and coordinates added to a movement matrix; if not, a new
658 movement is drawn from the distribution. If the movement is greater than 1cm in distance, we break
659 the movement into a number of smaller parts to be added to the matrix (such that we capture the
660 coordinates of grid squares through which the 'fish' moved along the way). Once the total distance of
661 the random walk reached or exceeded the *track length* set as the simulation objective, the path is
662 terminated and the area covered is calculated by counting the number of unique grid squares in the
663 matrix of coordinates and dividing by the total number possible.
664 After simulating random walks across 500 values of *track length* (using a vector of 100 values
665 evenly spaced across the range of true data, repeated 5 times), we modelled (simulated) area covered
666 as a fourth order polynomial function of *track length*. Using this regression model (which explained

667 97.8% of the variance in simulated data), we calculated the *relative area* for each actual OFT trial as
668 the observed area covered minus the area covered under a random swim, as predicted from our
669 regression model and the observed *track length*.
670
671 *Waterborne hormone sampling*
672 On completion of behavioural data collection, individuals entering the endocrine testing program were
673 left undisturbed for a minimum of two weeks. Waterborne hormone sampling was then conducted
674 over a 5-day period that included three handling and confinement stressor exposures with 48h
675 between each. We followed the method described by Houslay et al (2019) to obtain repeated non-
676 invasive GC measures of individuals using holding water samples from the first and third
677 confinements. Note that only two samples per fish were analysed because the financial and time costs
678 of doing three was deemed prohibitive. We nonetheless applied the stressor stimulus three times as
679 our prior study showed this was sufficient to produce a strong habituation response, i.e., a significant
680 decrease in water-borne cortisol over the three sampling periods (Houslay *et al.* 2019).
681 We collected samples between 1200 – 1400h to control for diel fluctuations in GC levels. For
682 each sample, we netted an entire group from their home tank quickly using a large net, transferring
683 them to 2 holding tanks (containing water from the home tank supply) for moving to an adjacent quiet
684 room (performed within 20s of the net first hitting the water). We then transferred fish to individual
685 Pyrex beakers containing 300ml of clean water from the main supply (which serves the main housing
686 units), which has been warmed to the appropriate temperature (mean = 24.1°C, range 23-24.9°C).
687 Beakers were placed within cardboard 'chambers' to prevent fish from seeing each other or
688 experiencing outside disturbance. One fish was transferred every 30s, alternating across holding
689 tanks, such that all fish were in their beakers within 10min of the initial netting. After 60 mins in the
690 beaker, each fish was removed by pouring its sample through a clean net into a second beaker, with
691 the fish then quickly checked to confirm ID and returned to the holding tank until the entire group
692 could be returned to its home tank.
693 We immediately filtered each water sample using Grade 1 filter paper (Whatman), then passed
694 them slowly through solid phase C18 extraction columns (Sep-Pak C18, 3cc, Waters) via high-purity

695 tubing (Tygon 2474, Saint Gobain) under vacuum pressure (Earley *et al.* 2006). Columns were primed
696 beforehand with 2 x 2ml HPLC-grade methanol followed by 2 x 2ml distilled water, and were washed
697 afterwards with a further 2 x 2ml distilled water to purge salts. We then covered both ends of each
698 column with film (Parafilm M, Bemis) and stored them at -20C for future analysis. We washed all
699 beakers, tubes and funnels with 99% ethanol and rinsed them with distilled water prior to each
700 sampling procedure. The remainder of the endocrine assay procedure involved elution, resuspension,
701 separation and quantification of free cortisol by enzyme immunoassay (EIA) using Cayman Chemicals,
702 Inc EIA kits. Detailed methods are exactly as described by Houslay et al (2019) and so not repeated
703 here (note that here we assayed the free fraction of cortisol only). To validate the cortisol kits, we
704 examined whether the kit standard curve was parallel to a serial dilution curve derived from pooled
705 guppy water-borne hormone extract. 20µl was taken from each of the male samples and pooled; 20µl
706 was taken from each of the female samples and combined into a separate pool. 400µl of the pools was
707 serially diluted from 1:1 to 1:128 and these samples were assayed alongside the kit standard curve on
708 two occasions (June and December 2017, marking the start and finish of sample processing). All
709 dilution curves were parallel to the standard curve (slope comparison test, Zar 1996, p.355; June,
710 male: t_{12} = 0.029, P = 0.97; June, female: : t_{12} = 0.343, P = 0.74; December, male: : t_{12} = 0.119, P = 0.91;
711 December, female: : t_{12} = 0.224, P = 0.83). The serial dilution curves also identified 1:32 as an
712 appropriate dilution to ensure that all samples fell on the linear phase of the standard curve. A total of
713 37 96-well plates were used and the pooled sample was included at the beginning and end of each
714 plate. Intra-assay coefficients of variation ranged from 0.12-19.83% with a median of 3.08%; the inter-
715 assay coefficient of variation was 19.22%. Cortisol is presented and modelled in (ln-transformed)
716 units of ng/hr to reflect the 1 hour sampling duration.
717
718 **Statistical methods**
719
720 All data handling and analysis was performed in R version 3.4.1 (R Core Team 2017). We used the
721 'tidyverse' set of packages for data handling and visualisation (Wickham 2017), and ASreml-R 3.0
722 (Butler 2009) for fitting linear mixed effects models (as described in full below). We also used 'nadiv'

723 for pedigree preparation and to estimate (approximate) standard errors on linear functions of
724 variance components as estimated from the mixed models (Wolak 2012). All models fitted assumed
725 (multivariate) Gaussian error structures, and we visually assessed residuals to verify this was
726 reasonable (after data transformation in some cases). To test for significance of among individual
727 and/or genetic (co)variance components, we fitted nested models with different random effects
728 structures and compared them using likelihood ratio tests (LRTs). We calculated $\chi^2_{n\mathrm{DF}}$ as twice the
729 difference in model log likelihoods, with the number of degrees of freedom (n) equivalent to the
730 number of additional parameters in the more complex model. When testing a single random effect
731 (variance component), we assumed the difference to be asymptotically distributed as an equal mix of
732 χ^2_0 and χ^2_1 (denoted $\chi^2_{0,1}$; Self and Liang, 1987; Visscher, 2006).

733 For each OFT and ST behaviour in turn (*relative area, time in middle, track length, shoaling*
734 *tendency,* and *emergence time*), we used the random effects specification to partition phenotypic
735 variation (V_p, conditional on fixed effects as described below) into the effects of additive genetics (V_a),
736 permanent environment defined as the non-(additive) genetic component of among-individual
737 differences, V_pe), and housing group (V_group), as well as residual variation (V_residual). We natural log-
738 transformed *emergence time* prior to analysis to meet assumptions of residual normality and
739 homoscedasticity. For all behavioural traits, we included fixed effects of assay *repeat*, the *order within*
740 *each group* in which the fish was trialled (mean-centred continuous predictor), *temperature* (mean-
741 centred and scaled to standard deviation units), *time* (in minutes from midnight, mean-centred and
742 scaled to standard deviation units), *age* (mean-centred and scaled to standard deviation units), *sex*,
743 and the *generation* from the breeding population. For *shoaling tendency* only, we incorporated an
744 additional fixed effect of *setup* (as detailed above). We tested the significance of genetic variance for
745 each behaviour by LRT comparison of the corresponding full model to one in which the (additive)
746 genetic random effect was excluded.

747 Cortisol data were also natural log (ln) transformed for analysis. We formulated a bivariate
748 model to test for both additive genetic variation and genotype-by-environment interaction (GxE) in
749 cortisol levels across the two 'contexts' (i.e. samples retained for each individual at first and third
750 confinement, denoted $\mathrm{Cortisol}_1$, $\mathrm{Cortisol}_3$). Random effects were first used to partition phenotypic

751 (co)variance (conditional on fixed effects) into among-group and residual components. Fixed effects
752 included the context-specific means, and overall effects of the *order* in which the fish was caught from
753 each group for assay (mean-centred continuous predictor), *temperature* (mean-centred and scaled to
754 standard deviation units), *time of day* (mean-centred and scaled to standard deviation units), *age*
755 (mean-centred and scaled to standard deviation units), and *sex*. In addition, we included fixed
756 covariates of *body mass* (mean-centred and scaled to standard deviation units) and a *sex* by *body mass*
757 interaction (see Houslay *et al.* 2019 for rationale of controlling for body size effects on waterborne
758 hormone levels in this way). Note that modelled in this way each individual is sampled only once for
759 each context-specific cortisol trait so no random effect of individual identity is included. To test for
760 additive genetic variation (V_a) we compared this first bivariate model to a second formulation that also
761 included the (additive) genetic merit, but under the assumption that this is invariant with context
762 within an individual (such that V_{a1}=V_{a3} and $r_{a1,3}$=1 and there is no GxE). We then test for the GxE by
763 comparing the second model to a third in which we allow GxE (i.e., the context-specific genetic
764 variances are free to differ and the cross-context genetic correlation can be <+1).
765 Lastly, we built a multivariate animal model to estimate **G** and to test the hypothesised genetic
766 integration among behavioural and physiological stress components. We retained only response traits
767 that harboured significant V_a as shown in univariate models, and so the final model comprised
768 response traits *relative area, time in middle, track length, emergence time* (log transformed), and
769 *Cortisol* (log transformed). We multiplied (transformed) *emergence time* by -1 to simplify
770 interpretation of estimated correlation structures (i.e., higher values for all behavioural traits then
771 represent nominally 'bolder' behaviours). We also scaled all (transformed) response variables to
772 standard deviation units. This was to facilitate model fitting, and also prevent scale effects
773 complicating interpretation of eigenvectors of **G**. Fixed and random effects were fitted on each trait as
774 specified for the univariate models. Note that one exception to this is that we elected to treat *Cortisol*
775 as a single repeated-measures trait here (with two repeats, one per context) such that a permanent
776 environment effect was now included. Fixed effects estimates are reported in the supplementary
777 information (Table S1).

We specified additive genetic (**G**), permanent environment (**PE**), group (**GROUP**), and residual (**R**) covariance structures as unstructured matrices to be estimated. Note that **R** partitions observation-level covariances (as opposed to individual-level in **PE**) that are not definable or statistically identifiable if traits are not measured at the same time (i.e., all covariances relating to *emergence time* or *Cortisol*). Where this was the case we constrained specific covariance terms in **R** to equal zero. Estimates of **PE**, **GROUP** and **R** are provided in the supplementary information (Tables S2-S4). We tested for overall additive genetic covariance among the traits by comparing this model against a reduced one in which **G** was specified as a diagonal matrix (i.e., additive variances are estimated but covariances are assumed to equal zero). To aid the interpretation of covariance terms contained in **G,** we calculated the corresponding genetic correlations r_a from the full model. For any pair of traits (x,y), $r_{a(x,y)} = COV_{a(x,y)} / (\sqrt{(V_{a(x)})} \times \sqrt{(V_{a(y)})})$. We also subjected our estimate of **G** to eigen decomposition to determine the proportion of additive genetic variation captured by each principal component and assess whether a single major axis of variation could indeed explain most of the genetic variance in the multivariate phenotype (consistent with a simple proactive-reactive coping style model). We estimated uncertainty on the trait loadings associated with each principal component (eigenvector) using a parametric bootstrap approach as described by Boulton et al (2014).

For visualisation of bivariate relationships at the additive genetic level, we used the R package 'ellipse' (Murdoch & Chow 2018) to determine the coordinates of an ellipse representing the approximate 95% confidence region of deviations based on the point estimate of **G**. We repeated this procedure for the corresponding regions defined from 5000 bootstrapped values of **G** (i.e., to indicate uncertainty arising from estimation of the genetic covariance structure itself). Best linear unbiased predictors (BLUPs) are used for visualisation only, not for any statistical analysis (Houslay & Wilson 2017).

To test for associations between all traits (i.e., including *shoaling tendency*) at the among-individual level, we also built a multivariate model as above with the addition of *shoaling tendency* and without estimating additive genetic effects. The estimates of all among-individual (co)variances are provided in the supplementary information (Table S5).

806

807 **SUPPLEMENTARY MATERIAL**

808

809 Table S1: Fixed effects estimates from the full multivariate animal model.
810 Table S2: Permanent environment (co)variance matrix from the full multivariate animal
811 model.
812 Table S3: Group (co)variance matrix from the full multivariate animal model.
813 Table S4: Residual variance-correlation matrix from the full multivariate animal model.
814 Table S5: Among-individual (co)variance matrix from the multivariate model that excluded
815 genetic effects.
816

817 **REFERENCES**

818

819 Angelier, F. & Wingfield, J.C. (2012) Importance of the glucocorticoid stress response in a changing
820 world: Theory, hypotheses and perspectives. *General and Comparative Endocrinology*, **190**, 118–
821 128.

822 Barton, B.A. & Iwama, G.K. (1991) Physiological changes in fish from stress in aquaculture with
823 emphasis on the response and effects of corticosteroids. *Annual Review of Fish Diseases*, **1**, 3–26.

824 Berger, D., Bauerfeind, S.S., Blanckenhorn, W.U. & Schäfer, M.A. (2011) HIGH TEMPERATURES REVEAL
825 CRYPTIC GENETIC VARIATION IN A POLYMORPHIC FEMALE SPERM STORAGE ORGAN.
826 *Evolution*, **65**, 2830–2842.

827 Bijlsma, R. & Loeschcke, V. (2005) Environmental stress, adaptation and evolution: an overview.
828 *Journal of evolutionary biology*, **18**, 744–9.

829 Blows, M.W. (2007) A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of*
830 *evolutionary biology*, **20**, 1–8.

831 Blows, M.W. & Walsh, B. (2009) Spherical cows grazing in Flatland: Constraints to selection and
832 adaptation. *Adaptation and Fitness in Animal Populations*, pp. 83–101. Springer, Dordrecht.

833 Boonstra, R. (2013) Reality as the leading cause of stress: Rethinking the impact of chronic stress in
834 nature. *Functional Ecology*, **27**, 11–23.

835 von Borell, E. (1995) Neuroendocrine integration of stress and significance of stress for the
836 performance of farm animals. *Applied Animal Behaviour Science*, **44**, 219–227.

837 Boulton, K., Couto, E., Grimmer, A.J., Earley, R.L., Canario, A.V.M., Wilson, A.J. & Walling, C.A. (2015)
838 How integrated are behavioral and endocrine stress response traits? A repeated measures
839 approach to testing the stress-coping style model. *Ecology and Evolution*, **5**, 618–633.

840 Boulton, K., Grimmer, A.J., Rosenthal, G.G., Walling, C.A. & Wilson, A.J. (2014) How stable are
841 personalities? A multivariate view of behavioural variation over long and short timescales in the
842 sheepshead swordtail, Xiphophorus birchmanni. *Behavioral Ecology and Sociobiology*, **68**, 791–
843 803.

844 Broom, D.M. & Johnson, K.G. (1993) Assessing welfare: short-term responses. *Stress and Animal*

845 *Welfare*, pp. 87–110. Springer Netherlands, Dordrecht.
846 Burns, J.G. (2008) The validity of three tests of temperament in guppies (Poecilia reticulata). *Journal of*
847 *comparative psychology*, **122**, 344–356.
848 Busch, D.S. & Hayward, L.S. (2009) Stress in a conservation context: A discussion of glucocorticoid
849 actions and how levels change with conservation-relevant variables. *Biological Conservation*, **142**,
850 2844–2853.
851 Butler, D. (2009) asreml: asreml() fits the linear mixed model.
852 Carere, C., Groothuis, T.G.G., Möstl, E., Daan, S. & Koolhaas, J.M. (2003) Fecal corticosteroids in a
853 territorial bird selected for different personalities: Daily rhythm and the response to social stress.
854 *Hormones and Behavior*, **43**, 540–548.
855 Carter, A.J., Feeney, W.E., Marshall, H.H., Cowlishaw, G. & Heinsohn, R. (2013) Animal personality: What
856 are behavioural ecologists measuring? *Biological Reviews*, **88**, 465–475.
857 Cheverud, J.M. (1982) Phenotypic, Genetic, and Environmental Morphological Integration in the
858 Cranium. *Evolution*, **36**, 499.
859 Clinchy, M., Sheriff, M.J. & Zanette, L.Y. (2013) Predator-induced stress and the ecology of fear.
860 *Functional Ecology*, **27**, 56–65.
861 Cox, R.M., McGlothlin, J.W. & Bonier, F. (2016) Hormones as Mediators of Phenotypic and Genetic
862 Integration: An Evolutionary Genetics Approach. *Integrative and Comparative Biology*, **56**, 126–
863 137.
864 Earley, R.L., Edwards, J.T., Aseem, O., Felton, K., Blumer, L.S., Karom, M. & Grober, M.S. (2006) Social
865 interactions tune aggression and stress responsiveness in a territorial cichlid fish (Archocentrus
866 nigrofasciatus). *Physiology and Behavior*, **88**, 353–363.
867 Frankham, R., Hemmer, H., Ryder, O.A., Cothran, E.G., Soulé, M.E., Murray, N.D. & Snyder, M. (1986)
868 Selection in captive populations. *Zoo Biology*, **5**, 127–138.
869 Fürtbauer, I., Pond, A., Heistermann, M. & King, A.J. (2015) Personality, plasticity and predation:
870 linking endocrine and behavioural reaction norms in stickleback fish. *Functional Ecology*, n/a-
871 n/a.
872 Hau, M. & Goymann, W. (2015) Endocrine mechanisms, behavioral phenotypes and plasticity: known

873 relationships and open questions. *Frontiers in Zoology*, **12**, S7.
874 Herbert-Read, J.E., Rosén, E., Szorkovszky, A., Ioannou, C.C., Rogell, B., Perna, A., Ramnarine, I.W.,
875 Kotrschal, A., Kolm, N., Krause, J. & Sumpter, D.J.T. (2017) How predation shapes the social
876 interaction rules of shoaling fish. *Proceedings of the Royal Society B: Biological Sciences*, **284**,
877 20171126.
878 Hine, E., Chenoweth, S.F. & Blows, M.W. (2004) Multivariate quantitative genetics and the lek paradox:
879 genetic variance in male sexually selected traits of Drosophila serrata under field conditions.
880 *Evolution*, **58**, 2754–62.
881 von Horsten, S., Karl, T. & Pabst, R. (2003) Behavioral phenotyping of mice in pharmacological and
882 toxicological research. *Experimental and Toxicologic Pathology*, **55**, 69–83.
883 Houslay, T.M., Earley, R.L., Young, A.J. & Wilson, A.J. (2019) Habituation and individual variation in the
884 endocrine stress response in the Trinidadian guppy (Poecilia reticulata). *General and*
885 *Comparative Endocrinology*, **270**, 113–122.
886 Houslay, T.M., Vierbuchen, M., Grimmer, A.J., Young, A.J. & Wilson, A.J. (2018) Testing the stability of
887 behavioural coping style across stress contexts in the Trinidadian guppy. *Functional Ecology*, **32**,
888 424–438.
889 Houslay, T.M. & Wilson, A.J. (2017) Avoiding the misuse of BLUP in behavioural ecology. *Behavioral*
890 *Ecology*, **00**, 1–5.
891 Huether, G. (1996) The central adaptation syndrome: Psychosocial stress as a trigger for adaptive
892 modification of brain structure and brain function. *Progress in Neurobiology*, **48**, 569–612.
893 Hunt, J., Blows, M.W., Zajitschek, F., Jennions, M.D. & Brooks, R.C. (2007) Reconciling strong stabilizing
894 selection with the maintenance of genetic variation in a natural population of black field crickets
895 (Teleogryllus commodus). *Genetics*, **177**, 875–80.
896 Jones, R.B., Satterlee, D.G. & Ryder, F.H. (1994) Fear of Humans in Japanese-Quail Selected for Low or
897 High Adrenocortical-Response. *Physiology & Behavior*, **56**, 379–383.
898 Ketterson, E.D., Atwell, J.W. & McGlothlin, J.W. (2009) Phenotypic integration and independence:
899 Hormones, performance, and response to environmental change. *Integrative and Comparative*
900 *Biology*, **49**, 365–379.

901 Koolhaas, J.M. (2008) Coping style and immunity in animals: Making sense of individual variation.
902 *Brain, Behavior, and Immunity*, **22**, 662–667.
903 Koolhaas, J.M., de Boer, S.F., Buwalda, B. & Van Reenen, K. (2007) Individual variation in coping with
904 stress: A multidimensional approach of ultimate and proximate mechanisms. *Brain, Behavior and*
905 *Evolution*, **70**, 218–226.
906 Koolhaas, J.M., de Boer, S.F., Coppens, C.M. & Buwalda, B. (2010) Neuroendocrinology of coping styles:
907 Towards understanding the biology of individual variation. *Frontiers in Neuroendocrinology*, **31**,
908 307–321.
909 Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C.,
910 Ruis, M.A.W. & Blokhuis, H.J. (1999) Coping styles in animals: Current status in behavior and
911 stress- physiology. *Neuroscience and Biobehavioral Reviews*, **23**, 925–935.
912 Korte, S.M., Koolhaas, J.M., Wingfield, J.C. & McEwen, B.S. (2005) The Darwinian concept of stress:
913 Benefits of allostasis and costs of allostatic load and the trade-offs in health and disease.
914 *Neuroscience and Biobehavioral Reviews*, **29**, 3–38.
915 Lande, R. & Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution*, **37**,
916 1210–1226.
917 Ledón-Rettig, C.C., Pfennig, D.W., Chunco, A.J. & Dworkin, I. (2014) Cryptic genetic variation in natural
918 populations: a predictive framework. *Integrative and comparative biology*, **54**, 783–793.
919 Ledón-Rettig, C.C., Pfennig, D.W. & Crespi, E.J. (2010) Diet and hormonal manipulation reveal cryptic
920 genetic variation: Implications for the evolution of novel feeding strategies. *Proceedings of the*
921 *Royal Society B: Biological Sciences*, **277**, 3569–3578.
922 Mason, G.J. (2010) Species differences in responses to captivity: Stress, welfare and the comparative
923 method. *Trends in Ecology and Evolution*, **25**, 713–721.
924 McEwen, B.S. & Wingfield, J.C. (2003) The concept of allostasis in biology and biomedicine. *Hormones*
925 *and Behavior*, **43**, 2–15.
926 McEwen, B.S. & Wingfield, J.C. (2010) What is in a name? Integrating homeostasis, allostasis and stress.
927 *Hormones and Behavior*, **57**, 105–111.
928 McGlothlin, J.W. & Ketterson, E.D. (2008) Hormone-mediated suites as adaptations and evolutionary

929 constraints. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*,
930 **363**, 1611–1620.
931 Mignon-Grasteau, S., Boissy, A., Bouix, J., Faure, J.M., Fisher, A.D., Hinch, G.N., Jensen, P., Le Neindre, P.,
932 Mormède, P., Prunet, P., Vandeputte, M. & Beaumont, C. (2005) Genetics of adaptation and
933 domestication in livestock. *Livestock Production Science*, **93**, 3–14.
934 Moberg, G.P. (2000) Biological response to stress: implications for animal welfare. *The biology of*
935 *animal stress: basic principles and implications for animal welfare*, pp. 1–21.
936 Möstl, E. & Palme, R. (2002) Hormones as indicator of stress. *Domestic animal and endocrinology*, **23**,
937 67–74.
938 Muir, W.M. & Craig, J. V. (1998) Improving animal well-being through genetic selection. *Poultry Science*,
939 **77**, 1781–1788.
940 Murdoch, D. & Chow, E.D. (2018) ellipse: Functions for Drawing Ellipses and Ellipse-like Confidence
941 Regions.
942 Oltenacu, P.A. & Algers, B. (2009) Selection for Increased Production and the Welfare of Dairy Cows:
943 Are New Breeding Goals Needed? *AMBIO: A Journal of the Human Environment*, **34**, 311–315.
944 Oswald, M.E., Singer, M. & Robison, B.D. (2013) The Quantitative Genetic Architecture of the Bold-Shy
945 Continuum in Zebrafish, Danio rerio. *PLoS ONE*, **8**, 1–10.
946 Øverli, Ø., Sørensen, C., Pulman, K.G.T., Pottinger, T.G., Korzan, W., Summers, C.H. & Nilsson, G.E. (2007)
947 Evolutionary background for stress-coping styles: Relationships between physiological,
948 behavioral, and cognitive traits in non-mammalian vertebrates. *Neuroscience and Biobehavioral*
949 *Reviews*, **31**, 396–412.
950 Øverli, Ø., Winberg, S. & Pottinger, T.G. (2005) Behavioral and Neuroendocrine Correlates of Selection
951 for Stress Responsiveness in Rainbow Trout—a Review. *Integrative and Comparative Biology*, **45**,
952 463–474.
953 Paaby, A.B. & Rockman, M. V. (2014) Cryptic genetic variation: Evolution's hidden substrate. *Nature*
954 *Reviews Genetics*, **15**, 247–258.
955 Perals, D., Griffin, A.S., Bartomeus, I. & Sol, D. (2017) Revisiting the open-field test: what does it really
956 tell us about animal personality? *Animal Behaviour*, **123**, 69–79.

957 Pottinger, T.G. & Carrick, T.R. (1999) Modification of the plasma cortisol response to stress in rainbow
958 trout by selective breeding. *General and Comparative Endocrinology*, **116**, 122–132.
959 Prentice, P.M., Houslay, T.M., Martin, J.G.A. & Wilson, A.J. (2020) Genetic variance for behavioural
960 'predictability' of stress response. *Journal of Evolutionary Biology*, 1–11.
961 R Core Team. (2017) R: A language and environment for statistical computing.
962 Réale, D., Garant, D., Humphries, M.M., Bergeron, P., Careau, V. & Montiglio, P.-O. (2010) Personality
963 and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical*
964 *transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 4051–4063.
965 Van Reenen, C.G., O'Connell, N.E., Van der Werf, J.T.N., Korte, S.M., Hopster, H., Jones, R.B. & Blokhuis,
966 H.J. (2005) Responses of calves to acute stress: Individual consistency and relations between
967 behavioral and physiological measures. *Physiology & Behavior*, **85**, 557–570.
968 Rex, A., Voigt, J.P., Voits, M. & Fink, H. (1998) Pharmacological evaluation of a modified open-field test
969 sensitive to anxiolytic drugs. *Pharmacology Biochemistry and Behavior*, **59**, 677–683.
970 Romero, L.M. (2004) Physiological stress in ecology: Lessons from biomedical research. *Trends in*
971 *Ecology and Evolution*, **19**, 249–255.
972 Romero, L.M., Dickens, M.J. & Cyr, N.E. (2009) The reactive scope model - A new model integrating
973 homeostasis, allostasis, and stress. *Hormones and Behavior*, **55**, 375–389.
974 Segerstrom, S.C. & Miller, G.E. (2004) Psychological stress and the human immune system: a meta-
975 analytic study of 30 years of inquiry. *Psychological bulletin*, **130**, 601–30.
976 Self, S.G. & Liang, K. (1987) Asymptotic properties of maximum likelihood estimators and likelihood
977 ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.*, **82**, 605–610.
978 Stamps, J.A. (2007) Growth-mortality tradeoffs and "personality traits" in animals. *Ecology Letters*, **10**,
979 355–363.
980 Stirling, D.G., Réale, D. & Roff, D.A. (2002) Selection, structure and the heritability of behaviour. *Journal*
981 *of Evolutionary Biology*, **15**, 277–289.
982 Taborsky, B., English, S., Fawcett, T.W., Kuijper, B., Leimar, O., McNamara, J.M., Ruuskanen, S. & Sandi,
983 C. (2020) Towards an Evolutionary Theory of Stress Responses. *Trends in Ecology and Evolution*,
984 1–10.

985 Taff, C.C. & Vitousek, M.N. (2016) Endocrine Flexibility: Optimizing Phenotypes in a Dynamic World?
986 *Trends in Ecology and Evolution*, **31**, 476–488.
987 Tarlow, E.M. & Blumstein, D.T. (2007) Evaluating methods to quantify anthropogenic stressors on wild
988 animals. *Applied Animal Behaviour Science*, **102**, 429–451.
989 Thomson, J.S., Watts, P.C., Pottinger, T.G. & Sneddon, L.U. (2011) Physiological and genetic correlates of
990 boldness: Characterising the mechanisms of behavioural variation in rainbow trout,
991 Oncorhynchus mykiss. *Hormones and Behavior*, **59**, 67–74.
992 Toms, C.N., Echevarria, D.J. & Jouandot, D.J. (2010) A Methodology Review of Personality-Related
993 Studies in Fish: Focus on the Shy-Bold Axis of Behaviour. *Berkeley Planning Journal*, **26**, 12.
994 Trenzado, C.E., Carrick, T.R. & Pottinger, T.G. (2003) Divergence of endocrine and metabolic responses
995 to stress in two rainbow trout lines selected for differing cortisol responsiveness to stress.
996 *General and Comparative Endocrinology*, **133**, 332–340.
997 Veenema, A.H., Meijer, O.C., De Kloet, E.R., Koolhaas, J.M. & Bohus, B.G. (2003) Differences in basal and
998 stress-induced HPA regulation of wild house mice selected for high and low aggression.
999 *Hormones and Behavior*, **43**, 197–204.
1000 Visscher, P.M. (2006) A Note on the Asymptotic Distribution of Likelihood Ratio Tests to Test Variance
1001 Components. *Twin Research and Human Genetics*, **9**, 490–495.
1002 Walsh, B. & Blows, M.W. (2009) Abundant Genetic Variation + Strong Selection = Multivariate Genetic
1003 Constraints: A Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and*
1004 *Systematics*, **40**, 41–59.
1005 White, S.J., Houslay, T.M. & Wilson, A.J. (2018) Evolutionary genetics of personality in the Trinidadian
1006 guppy II: sexual dimorphism and genotype-by-sex interactions. *Heredity*.
1007 White, S.J., Kells, T.J. & Wilson, A.J. (2016) Metabolism, personality and pace of life in the Trinidadian
1008 guppy, Poecilia reticulata. *Behaviour*, **153**, 1517–1543.
1009 White, S.J., Pascall, D.J. & Wilson, A.J. (2020) Towards a comparative approach to the structure of
1010 animal personality variation. *Behavioral Ecology*, **31**, 340–351.
1011 White, S.J. & Wilson, A.J. (2018) Evolutionary genetics of personality in the Trinidadian guppy I:
1012 maternal and additive genetic effects across ontogeny. *Heredity*.

1013 Wickham, H. (2017) tidyverse: Easily Install and Load "Tidyverse" Packages.

1014 Wilson, D.S., Clark, A.B., Coleman, K. & Dearstyne, T. (1994) Shyness and boldness in humans and other
1015 animals. *Trends in Ecology and Evolution*, **9**, 442–446.

1016 Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., Kruuk, L.E.B. & Nussey,
1017 D.H. (2009) An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13–26.

1018 Wingfield, J.C. (2003) Control of behavioural strategies for capricious environments. *Animal Behaviour*,
1019 **66**, 807–816.

1020 Wingfield, J.C. & Kitaysky, A.S. (2002) Endocrine responses to unpredictable environmental events:
1021 Stress or anti-stress hormones? *Integr Comp Biol*, **42**, 600–609.

1022 Wingfield, J.C., Maney, D.L., Breuner, C.W., Jacobs, J.D., Lynn, S., Ramenofsky, M. & Richardson, R.D.
1023 (1998) Ecological Bases of Hormone-Behavior Interactions: The "Emergency Life History Stage."
1024 *American Zoologist*, **38**, 191–206.

1025 Wolak, M.E. (2012) Nadiv: An R package to create relatedness matrices for estimating non-additive
1026 genetic variances in animal models. *Methods in Ecology and Evolution*, **3**, 792–796.

1027 Wolf, M., van Doorn, G.S., Leimar, O. & Weissing, F.J. (2007) Life-history trade-offs favour the evolution
1028 of animal personalities. *Nature*, **447**, 581–4.

1029 Zar, J.H. (1996) *Biostatistical Analysis*. Prentice-Hall International, Inc., London.

1030