Genetic integration of behavioural and endocrine components of the stress response

T.M. Houslay*, R.L. Earley, S.J. White, W. Lammers, A.J. Grimmer, L.M. Travers, E.L. Johnson, A.J. Young and A.J. Wilson

1 Centre for Ecology and Conservation, University of Exeter (Penryn Campus), Cornwall, TR10 9FE, UK.
2 Department of Biological Sciences, University of Alabama, 300 Hackberry Lane, Box 870344, Tuscaloosa, AL 35487, USA.
3 Current address: School of Biological Sciences, University of East Anglia, Norwich, Norfolk, NR4 7TJ, UK.
4 Current address: Southern Research, 2000 Ninth Avenue South, Birmingham, AL 35205, USA.
* Corresponding author: houslay@gmail.com

Keywords: stress; coping styles; animal behaviour; quantitative genetics; physiology; endocrinology; genetic integration; heritability
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. Natural selection should favour functional integration, leading to a prediction of genetic integration of these 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 acute stress response components in Trinidadian guppies are both heritable and integrated on the major axis of genetic covariation. This integration could either facilitate or constrain evolutionary responses to selection, depending upon the alignment of selection with this axis. Such integration also suggests artificial selection on the genetically correlated behavioural responses to stress could offer a viable non-invasive route to the 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; Kasper et al. 2020). 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; Sadoul et al. 2021) 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). Evolutionary theory
predicts that correlational selection will shape the structure of multivariate quantitative
gene variance (as represented by the genetic covariance matrix G; Cheverud 1982;
Blows 2007; Roff & Fairbairn 2012), contingent on the distribution of new pleiotropic
mutations that generate multivariate genetic variance as well as the selection that
depletes it (Blows & Walsh 2009; Walsh & Blows 2009). In general, correlational
selection should (under certain assumptions) have direct and strong effects on genetic
covariances (Lande 1980; Jones, Arnold & Bürger 2003; Arnold et al. 2008). In the
context of the stress response we should therefore expect genetic – as well as
phenotypic – integration of behavioural and physiological traits (McGlothlin & Ketterson
2008; Ketterson, Atwell & McGlothlin 2009; Cox, McGlothlin & Bonier 2016). By genetic
integration, we mean genetic correlation structure among suites of traits. Genetic
integration of the stress response components is hypothesised to underpin adaptive
variation in multivariate phenotypes in many fields of evolutionary biology (e.g., life
history (Stearns 1989; Roff 1992), behavioural ecology (Sih & Bell 2008)), but has not
been tested for explicitly using quantitative genetic approaches.
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 Oncorhynchus mykiss (Pottinger and
Carrick, 1999), Japanese quail *Coturnix coturnix japonicus* (Jones, Satterlee & Ryder 1994), house mice *Mus musculus domesticus* (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), 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). In the same selection lines, there was also no link between 'boldness' (an aspect of personality related to risk-taking behaviour) and cortisol response 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). Insight into past selection 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 the 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). Insight into future responses to selection follows from the fact that the direction (in multivariate trait space) and magnitude of a response to selection is limited by the amount of variance in $G$ that is in alignment with the vector of (directional) selection $\beta$ (Blows & Walsh 2009; Walsh & Blows 2009). Integration in $G$ is manifest as genetic correlations between trait pairs, and also at the level of the whole matrix by an overdispersion of eigenvalues—indicating that there are fewer effective dimensions of genetic variation than there are traits measured (Blows 2007).

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 on the major axis of genetic (co)variation (i.e., the first eigen vector of $G$, $g_{max}$), but also (ii) whether the structure and orientation of this axis suggests variation in overall stress responsiveness and/or ‘coping style’ (explained further below; Koolhaas et al. 2010; Boulton et al. 2015). We use fish from a captive colony of guppies derived from wild ancestors sampled from the Aripo River, Trinidad in 2008 and subsequently maintained at high population size (with no deliberate inbreeding or selection). We have validated the use of standardised ‘open field trials’ (OFTs) for testing (acute) behavioural stress responses in this species (Houslay et al. 2018), and demonstrated significant additive genetic (co)variance underpinning variation in risk-taking, exploratory, and ‘flight’ type components of the behavioural stress response using this testing paradigm (White & Wilson 2018; White, Houslay & Wilson 2018). We have also demonstrated, using a non-invasive waterborne hormone sampling method, that individuals differ significantly in their GC (specifically, free circulating cortisol) response to an acute stressor (handling,
coupled with short term isolation and confinement; Houslay et al. 2019) and that, on average, this physiological response declines with repeated stressor exposure (consistent with habituation). Nothing is known about the genetic basis of variation in these physiological traits, or about their integration (phenotypically or genetically) with behavioural components of the stress response.

First, we combine OFT results with complementary ‘emergence trials’ (ET) and ‘shoaling trials’ (ST) to characterise among-individual and genetic variation in the behavioural stress response. Second, we characterise the physiological stress response and its rate of habituation by assaying GC levels following first and third exposure to a handling and confinement stressor (see methods). Utilising repeated behavioural and physiological testing of individual fish within a known pedigree structure, we are able to estimate the repeatable (among-individual) component of phenotypic (co)variance in these stress response traits, and then determine the additive genetic contribution to this (G, the genetic variance-covariance matrix for this suite of stress response traits). We predicted that individual traits will be heritable and that G will contain strong genetic correlation structure between behavioural and physiological components of the stress response consistent with genetic integration. We also predicted that both behavioural and endocrine components of the stress response would load on the major axis of genetic variation in multivariate trait space, $g_{\text{max}}$. The ‘stress coping style’ model (Koolhaas et al. 1999) predicts variation in the type of response to stress. Simplistically, as originally proposed this verbal model posits that individuals (or genotypes) perceive equal degrees of stress but differ in how this manifests phenotypically: genotypes at one end of the axis having ‘reactive’ behavioural phenotype (e.g., freezing behaviour) coupled to lower GC levels, while the ‘proactive’ end is characterised by more active ‘fight or flight’ behaviour coupled to higher GC levels. However, previous analyses of this population suggest variation may be more in terms of ‘stress responsiveness’ than stress coping style (Houslay et al. 2018; Prentice et al. 2020; White, Pascall & Wilson 2020).
That is to say, some individuals (or genotypes) perceive the trial as a more severe stimulus and exhibit more characteristic stress behaviours (e.g., flight and/or freezing, thigmotaxis) while others show more typical ‘unstressed’ behavioural profiles (e.g., exploration of the arena). In this scenario we predicted high GC levels to co-occur with characteristic stress behaviours, and low GC levels with ‘unstressed’ behavioural profiles.

Results

In total we obtained multivariate behavioural data from 5,966 trials (3,379 OFTs, 1,548 ETs and 1,039 STs) on 1,384 individual fish. The number of individuals phenotyped (OFTs = 1,365, ETs = 806, STs = 532) and the mean number of observations per fish (OFTs = 2.5, ETs = 1.9, STs = 2.0) varied across the behavioural data types. All fish were contained within a genetic pedigree structure comprising maternal full-sibships nested within paternal half-sibships. This structure was produced via multiple rounds of breeding work and has a maximum depth of 5 ‘generations’. Some of the OFT data have already been used in studies of the evolutionary genetics of personality (White & Wilson 2018; White et al. 2018), but here we extend that dataset and use it in conjunction with other behavioural and physiological measures for different purposes. We also obtained 1,238 waterborne assays of cortisol levels for 629 fish (almost all from the final generation). The handling and confinement stressor applied for this assay was performed 3 times (at 48h intervals) for all fish tested, but the holding water sample was only processed for GC content at two time points (the first and last confinement, subsequently Cortisol₁ and Cortisol₃). Full details of husbandry, phenotyping and analysis are provided in Materials and Methods.

Genetic variance in behavioural components of the stress response
Behavioural data were extracted from OFTs, ETs and STs using video tracking of fish (as described in White, Kells & Wilson 2016; Houslay et al. 2018). Time to emerge from the shelter (‘emergence time’) was extracted from ETs and natural log (ln) transformed for analysis, while shoaling tendency was calculated from STs as the time spent in the third of the tank closest to a same-sex shoal (which was visible but physically separated) minus the time spent in the third of the tank farthest from the shoal. The OFT, ET and ST testing paradigms are all considered to assay behavioural components of the stress response in the broad sense, as each test starts with the capture and transfer of the focal fish into a novel and brightly lit arena away from their home tank and familiar mates.

Four traits were defined from the OFT and measured by videotracking for 4m 30s after an initial 30 second acclimation period upon transfer into the arena: track length (distance swum), area covered (as a proportion of the arena floor area), the number of freezings (the number of times a fish’s speed dropped below the minimum velocity threshold of 4cm/s for at least 2.5s, Houslay et al. 2018; this trait was square root transformed for analysis) and time in the middle (i.e., in the central area of the open field arena away from the tank walls, which is assumed to be perceived as riskier, e.g., Houslay et al. 2018). Note low values of time in the middle imply thigmotaxis (i.e., tendency to avoid exposure to potential threats by hugging walls), and were very highly correlated with a measure of average distance from the wall at the observation level ($r = 0.94, t_{13368} = 159.2, P < 0.0001$; see Supplementary File 1 for further discussion of the selection of OFT behavioural traits). All of these OFT traits are repeatable and heritable in this population (White et al. 2016, 2018; Houslay et al. 2018; White & Wilson 2018).

The absence of a strong positive correlation between track length and area covered (Fig 1A) is notable and potentially biologically informative; if fish moved randomly with respect to direction in the arena then area covered would increase monotonically (to an asymptote at 100%) with track length. A possible explanation is
that a long track length arises sometimes from a (putatively) less stressed fish exploring the arena (fish 1 in Fig 1B) and sometimes from a (putatively) more stressed fish exhibiting a typical ‘flight’ response (fish 4 in Fig 1B). These two types of response can be discriminated based on whether, in a given trial, higher track length is associated with higher area covered and time in the middle (exploration) or the converse (flight response).

To quantitatively discriminate between exploratory behaviour and flight responses we derived a new trait, ‘relative area covered’. We used a simple simulation procedure (see Methods) to predict expected area covered for any given track length under a null ‘random swim’ within the arena (Fig 1C). Relative area covered is then calculated as observed area covered – expected area covered given the track length (Fig 1D) and will be high for fish engaging in exploration, and low for an obvious ‘flight’ response manifest as rapid swimming around the tank walls. Our subsequent analyses then used this derived relative area covered trait in place of area covered.

Pedigree-based ‘animal models’ (Wilson et al. 2009) were used to test for and estimate additive genetic variation in each of the six behavioural traits (emergence time, shoaling tendency, and the four OFT traits) while controlling statistically for social housing group and non-genetic sources of among-individual variance (as well as several fixed effects; see methods for full details). These confirmed the presence of significant additive genetic variation for the relative area covered trait, as well as for track length, time in the middle (as expected from previous findings; White & Wilson 2018; White et al. 2018), √freezeings and ln emergence time (Table 1). With the exception of shoaling tendency, heritabilities (conditional on fixed effects; see methods) are low to moderate (range of 9-20%; Table 1) and within the expected range for behaviours (Stirling, Réale & Roff 2002). We detected no additive genetic variance for shoaling tendency (Table 1), despite there being repeatable differences among individuals; $R = 0.19 \pm 0.04$; $\chi^2_{0.1} = 20.01, P < 0.001$).
Table 1: Estimated variance components, along with adjusted heritability, for each trait as estimated in a univariate model (± standard error). Chi-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 | 42.87 ± 12.89   | 56.16 ± 10.92     | 19.17 ± 4.92      | 203.17 ± 6.26       | 0.13 ± 0.04| 31.2           | <0.001|
| Time in the middle  | 554.78 ± 144.18 | 473.24 ± 114.77   | 196.94 ± 53.54    | 2002.06 ± 62.38     | 0.17 ± 0.04| 47.2           | <0.001|
| Track length        | 23584.54 ± 5534.90 | 26587.74 ± 4556.87| 3173.34 ± 1396.84| 76140.49 ± 2376.32  | 0.18 ± 0.04| 88.7           | <0.001|
| $\sqrt{\text{Freezings}}$ | 0.34 ± 0.08   | 0.18 ± 0.06       | 0.08 ± 0.02       | 1.13 ± 0.04         | 0.20 ± 0.05| 48.2           | <0.001|
| In Emergence time   | 0.12 ± 0.05    | 0.06 ± 0.06       | 0.05 ± 0.02       | 1.07 ± 0.05         | 0.09 ± 0.02| 22.5           | <0.001|
| Shoaling tendency   | 0 ± 0          | 2457.36 ± 570.96  | 708.87 ± 316.30   | 9900.95 ± 622.10    | 0 ± 0      | 0              | 0.5   |
| In Cortisol         | 0.08 ± 0.02    | 0.02 ± 0.02       | 0.01 ± 0.01       | 0.15 ± 0.01         | 0.31 ± 0.09| 24.4           | <0.001|
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.

Genetic variance in physiological components of the stress response

Using a series of nested bivariate animal models, we tested for additive genetic variance in cortisol levels (ln-transformed) following stressor exposure (handling and confinement) and for genotype-by-environment interaction (GxE). In this context, the environment (E) is the trial number in each fish’s stress trial series (i.e., cortisol level after stressor trial 1 or 3). Any GxE present can therefore be interpreted as genetic variance for habituation to the stressor, given that the average cortisol level was lower following exposure to the third stressor than the first (ln transformed ng/hr, mean ± SE; Cortisol$_1$ = 8.47 ± 0.05, Cortisol$_3$ = 8.02 ± 0.05, Wald $F_{1,38.8} = 110.0$, $P < 0.001$; see methods for explanation of units). We first modelled Cortisol$_1$ and Cortisol$_3$ as distinct response variables in a bivariate framework assuming no GxE (such that we constrain $V_{A-Cortisol1} = V_{A-Cortisol3}$ and the cross context additive genetic correlation $r_{A-Cortisol1,Cortisol3} = 1$). This model revealed a significant additive genetic component to variation among individuals in their cortisol levels following stressor exposure ($\chi^2_{0.1} = 18.2$, $P < 0.001$).

Expanding the model to allow GxE by estimating separate genetic variances for Cortisol$_1$ and Cortisol$_3$ provides a significantly better fit to the data ($\chi^2_{0.1} = 3.8$, $P = 0.03$), meaning GxE is present. This can be viewed as genetic variance for habituation to repeated stressor exposure, or as a change in genetic variance for cortisol from the first to the third sampling (Figure 2). These are two perspectives of the same phenomenon; a
reduction in additive genetic variance between the first stressor ($V_{A-Cortisol1} = 0.081 \pm 0.029$) and the third ($V_{A-Cortisol3} = 0.041 \pm 0.025$) arises because genotypes with higher-than-average levels for Cortisol habituate more rapidly (i.e., have more negative reaction norm slopes). Note however that allowing the cross-context genetic correlation to deviate from +1 does not significantly improve the model fit ($\chi^2_1 = 0.0, P = 1$): thus 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 ($r_{A-Cortisol1,Cortisol3} \pm SE = 0.83 \pm 0.21$).

In this model we also find that variance in cortisol explained by housing group effects is similar across contexts ($V_{\text{Group-Cortisol1}} = 0.038 \pm 0.014, V_{\text{Group-Cortisol3}} = 0.032 \pm 0.012$), but that residual (unexplained) variance is greater after the third stressor exposure ($V_{R-Cortisol1} = 0.164 \pm 0.022, V_{R-Cortisol3} = 0.237 \pm 0.023$). 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 ($h^2_{\text{Cortisol1}} = 0.285 \pm 0.094, h^2_{\text{Cortisol3}} = 0.131 \pm 0.078$).
Figure 2: Guppies habituate to the waterborne sampling procedure, as shown by a decline in average ln-transformed cortisol level (ng/hr) following stressor exposure between first and third exposures. Black circles and associated bars denote predicted population means (± standard error) from mixed model analysis. Grey lines depict the predicted genetic reaction norms across repeated stressor exposure for each individual. Weak, but statistically significant GxE is manifest as variance in the genetic reaction norm slopes (i.e. lack of parallelism) and results in a slight reduction of genetic variance for cortisol at the third exposure relative to the first.

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 six response traits: *relative area covered*, *time in the middle*, *track length*, (square root-transformed) *freezings*, (ln-transformed) *emergence time*, and (ln-transformed) *Cortisol* (now treated as two repeats of a single trait). We standardised all traits to standard deviation units (after transformation where applicable) to assist multivariate model fitting and to prevent eigenvectors of $G$ (see below) being dominated by traits with higher variance in observed units. To simplify interpretation of $G$ we also multiplied *emergence time* by -1 after transformation. Thus high values denote rapid emergence from the shelter.

The resultant estimate of $G$ (Table 2) contains significant additive genetic covariance structure overall (Likelihood Ratio Test of the full model vs. a reduced model $G$ that contains variances but not covariances: $\chi^2_{15} = 91.06, P < 0.001$). We found a large number of strong genetic covariance/correlation estimates between trait pairs that were deemed statistically significant (based on the bootstrapped 95% confidence intervals not crossing zero). The 4 OFT traits (*relative area covered*, *time in the middle*, *track length*, and $\sqrt{\text{freezings}}$) show strong patterns of pairwise genetic correlation. *Track length* shows significant negative genetic correlation with *relative area covered*, *time in the middle*, and $\sqrt{\text{freezings}}$ (such that genotypes predisposed to higher values of track length are also predisposed to lower values of *relative area covered*, *time in the middle*, and $\sqrt{\text{freezings}}$, and vice versa). *Relative area covered* and *time in the middle* show
significant positive genetic correlation with one another, and \( \sqrt{\text{freezings}} \) shows significant positive correlation with \textit{time in the middle}. The only genetic correlation among OFT traits that is not significant is between \( \sqrt{\text{freezings}} \) and \textit{relative area covered}. The speed at which individuals emerge from the shelter during the emergence trials, \( -\ln \text{ emergence time} \), shows significant positive genetic correlation with \textit{track length} (i.e., genotypes predisposed to higher values of \textit{track length} in the OFT are also predisposed to faster \textit{emergence} in the ET). There are no significant pairwise genetic correlations between \( -\ln \text{ emergence time} \) and the other OFT traits, but all correlation estimates are negative. Pairwise genetic correlations between \textit{ln cortisol} and all behavioural traits are illustrated in Figure 3. Ln-transformed \textit{cortisol} shows significant negative genetic correlations with both \textit{time in the middle} and \( \sqrt{\text{freezings}} \). Genetic correlations between \textit{ln cortisol} and other OFT traits were not significantly different from zero (based on 95% confidence intervals), with correlations estimated as negative with \textit{relative area covered} and positive with \textit{track length}. The genetic correlation between \( -\ln \text{ emergence time} \) and \textit{ln cortisol} is also not significant, and is estimated at very close to zero.

Eigen decomposition of \( G \) provides a more holistic view of the genetic covariance structure and the level of integration among traits. Here the major axis \( g_{\text{max}} \) (first principal component, PC1, with 95% confidence intervals from 5000 bootstrap replicates) explains 59.5\% (47.5\%, 75.7\%) of the genetic variance in multivariate phenotype. Subsequent axes necessarily capture declining proportions of the multivariate genetic variance (PC2 = 20.2\% [14.8, 32.1]; PC3 = 14.1\% [5.6, 19.6]; PC4 = 4.5\% [0.9, 8.2]; PC5 = 1.0\% [0, 2.8], PC6 = 0.7\% [0.0, 0.1]; see Table S6, Figure 4—figure supplement 1). All traits except \( -\ln \text{ emergence time} \) load significantly on \( g_{\text{max}} \) (Figure 4). \textit{Relative area covered}, \textit{time in the middle} and \( \sqrt{\text{freezings}} \) load in one direction, while \textit{track length} and \textit{ln Cortisol} load in the other direction. This structure indicates a major axis of genetic variation in integrated stress response (Figure 4—figure supplement 2), where genotypes at one end of this axis can be considered to have a ‘freeze’ (or possibly
‘weaker’; discussed later) type of behavioural stress response to the OFT assay (i.e., freezing more frequently, swimming shorter distances, spending more time in the central area of the tank, and exhibiting exploratory swimming patterns that cover greater areas relative to their distance swum). This behavioural profile is associated with ‘weaker’ physiological responses to stress (i.e., producing lower cortisol levels in response to the stressor). Meanwhile, genotypes at the other end of $g_{max}$ have a more ‘flight’ (or arguably ‘stronger’) type of behavioural stress response to the OFT assay (i.e., freezing less frequently, swimming further, spending more time close to the tank edges, and covering less area relative to their distance travelled in OFTs). This behavioural profile is associated with ‘stronger’ physiological responses to stress (i.e., producing higher cortisol levels in response to the stressor).
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 | √Freezings | -ln Emergence time | ln Cortisol |
|----------------------|-------------------|-------------|-----------|--------------------|------------|
| Relative area covered | 0.115 (0.050,0.182) | 0.795 (0.601,0.952) | -0.549 (-0.789,-0.239) | 0.139 (-0.260,0.511) | -0.438 (-1.305,0.239) | -0.376 (-0.959,0.272) |
| Time in the middle | 0.103 (0.044,0.165) | 0.145 (0.070,0.215) | -0.658 (-0.86,-0.414) | 0.363 (0.024,0.659) | -0.153 (-0.815,0.538) | -0.617 (-1.139,-0.155) |
| Track length | -0.071 (-0.127,-0.017) | -0.096 (-0.154,-0.035) | 0.147 (0.080,0.219) | -0.801 (-0.931,-0.647) | 0.61 (0.064,1.328) | 0.425 (-0.027,0.968) |
| √Freezings | 0.020 (-0.034,0.076) | 0.059 (-0.001,0.120) | -0.132 (-0.200,-0.060) | 0.185 (0.091,0.269) | -0.483 (-1.103,0.146) | -0.556 (-1.069,-0.067) |
| -ln Emergence time | -0.041 (-0.103,0.012) | -0.016 (-0.073,0.041) | 0.065 (0.008,0.121) | -0.057 (-0.126,0.009) | 0.076 (0.009,0.148) | -0.020 (-0.807,0.790) |
| ln Cortisol | -0.044 (-0.105,0.023) | -0.082 (-0.145,-0.018) | 0.057 (-0.002,0.119) | -0.083 (-0.155,-0.01) | -0.002 (-0.058,0.053) | 0.121 (0.036,0.206) |
A

Relative area covered (genetic deviation)

In Cortisol (genetic deviation)

B

Time in the middle (genetic deviation)

In Cortisol (genetic deviation)

C

Track length (genetic deviation)

In Cortisol (genetic deviation)

D

√Freezings (genetic deviation)

In Cortisol (genetic deviation)

E

In Emergence time (genetic deviation)

In Cortisol (genetic deviation)
**Figure 3**: The additive genetic relationship between ln-transformed cortisol (x-axis) and five behaviours (a, relative area covered; b, time in the middle; c, track length; d, \(\sqrt{\text{freezings}}\); e, \(\ln\) 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 4: Trait loadings from $g_{\text{max}}$, the first eigen vector (principal component) of $G$. This axis explains 59.5% 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).
In this study we sought to determine whether – and to what extent – there exists genetic variation for, and integration between, behavioural and physiological (endocrine) components of the stress response. Our results provide three main novel insights. First, we find that genetic variation does underpin individual differences in both behavioural and physiological components of the stress response. Second, we find genetic covariance structure among these behavioural and physiological traits, indicating that they are indeed genetically integrated. Thirdly, having identified the structure of the major axis $g_{\text{max}}$ of the genetic variance-covariance matrix $\mathbf{G}$ we suggest that it is more readily interpreted as an axis of genetic variation in stress coping style than in stress responsiveness, although we acknowledge that any distinction may be somewhat subjective. Overall, by estimating the genetic covariance structure among traits we find the first quantitative genetic support to date for the hypothesis of evolutionary integration between behavioural and endocrine components of the stress response.

We find heritable variation in, and covariation among, behaviours assayed in the open field trial (OFT), including the derived trait relative area covered. The latter trait, derived by considering an appropriate biological null model of the relationship between track length and (absolute) area covered serves as a useful proxy for exploratory behaviour. Here we demonstrate an axis of repeatable and heritable variation that spans from ‘freeze’ type responses characterised by less active but more exploratory swimming patterns (more freezings and lower track length, coupled with higher relative area covered and time in the middle) through to a ‘flight’ type response characterised by higher activity coupled to thigmotaxis and low exploration (fewer freezings and higher track length, coupled with lower relative area covered and time in the middle). Given the wide use of OFTs in biomedical research (e.g., Rex et al. 1998; von Horsten, Karl & Pabst 2003) as well as in animal behaviour, our phenotyping approach (including the derivation of relative area) may have broad applicability for discriminating between exploration and stress/anxiety-related patterns of behaviour. The OFT paradigm is also widely applied to studies of ‘shy—bold’ type personality variation in fishes (Toms, Echevarria & Jouandot 2010) and other vertebrates (Carter et al. 2013; Perals et al. 2017). The extent to which behavioural differences deemed characteristic of a ‘shy—bold’ personality axis (commonly, if not
universally, defined as repeatable variation in response to perceived risk; Wilson et al. 1994) should
be viewed as equivalent to variance in 'stress coping style' and/or 'stress responsiveness', is a matter
of debate (see Boulton et al. 2015). While we view these as overlapping concepts, our findings do
highlight reasons for caution in considering them interchangeable. For example, while the structure of
the major axis of $G$, $g_{\text{max}}$, is broadly consistent with it reflecting an axis of integrated genetic variation
in 'stress coping style' (see below), emergence time, an OFT trait that is commonly considered to reflect
'boldness' (Burns 2008), was found to be heritable but did not load significantly on $g_{\text{max}}$ (see Figure 4).
This could indicate that the shy—bold continuum does not align with the stress coping style
continuum in our model organism. However, it could also reflect complexities with interpreting
emergence time. For example, at least some genotypes (and individuals) could perceive the shelter
area as less safe than the open arena (i.e., counter to the assumption that fast emergence reflects a lack
of fear of the open arena, and thus greater boldness, e.g., Burns 2008). Indeed, in an earlier study some
guppies decreased (rather than increased) shelter use following simulated predation events (Houslay
et al. 2018).

Tendency to shoal varies among individuals but is not detectably heritable. Though not
generally considered a stress-response trait per se, shoaling is an anti-predator behaviour in guppies
(Herbert-Read et al. 2017). We had therefore predicted that heightened perception of risk in the open
field might also be associated with increased shoaling tendency. This was not the case at the among-
individual level (Table S5), while the absence of detectable genetic variance meant that we could not
test this prediction in $G$.

We find strong evidence of significant additive genetic variance in a key physiological
component of the stress response: waterborne cortisol concentrations following exposure of the fish
to a handling stressor. Our findings suggest that previously detected differences among individuals in
their cortisol response to a stressor (Houslay et al. 2019) are primarily attributable to genetic effects,
with the estimated heritability ($h^2 = 0.31$) constituting over 80% of the individual–level repeatability
($R = 0.37$) for ln-transformed Cortisol. In addition, by adopting a reaction norm approach to modelling
stress physiology, as recently advocated by ourselves (Houslay et al. 2019) and others (e.g., Fürtbauer
et al. 2015; Hau & Goymann 2015; Taff & Vitousek 2016; Guindre-Parker 2020; Malkoc et al. 2021), we
detect GxE reflecting genetic differences in the extent of habituation to the stressor over repeated exposures. This result is potentially important because poor habituation of the hypothalamic-pituitary-adrenal/interrenal (HPA/I) response to repeated or ongoing stressors can lead to well-documented health problems in human and animal populations (Segerstrom & Miller 2004; Koolhaas 2008; Romero, Dickens & Cyr 2009; Mason 2010). Our detection of heritable variation in the degree of habituation to stressors raises the possibility of developing targeted selection strategies to improve welfare in captive populations (e.g., Frankham et al. 1986; Muir & Craig 1995; Oltenacu & Algers 2009).

Our findings also highlight that there is greater additive genetic variance (and heritability) for cortisol levels following the first exposure to the stressor than following the third. This pattern, which occurs because genotypes that produce the highest cortisol response at first exposure also show the most marked habituation, is consistent with the idea of cryptic genetic variance being ‘released’ by exposure to novel, and so potentially stressful, environments (Ledón-Rettig, Pfennig & Crespi 2010; Ledón-Rettig et al. 2014; Berger et al. 2011; Paaby & Rockman 2014). All else being equal, it also means that selection on cortisol levels following stressor exposure should induce a stronger evolutionary response in naïve relative to habituated fish. However, the strong positive cross-environment correlation means that the ranking of genotypes with regard to their cortisol responses is consistent across repeated stressor exposures. Thus selection on the (average) GC response would result in a correlated evolutionary response of habituation rate, and vice versa. With regard to drawing inferences from these GC levels about the extent to which an organism is in a ‘stressed’ state, it is worth noting the potential complexity that may arise from GCs serving numerous functions. For example, in addition to mobilising energy and down-regulating non-essential processes in order to facilitate coping with an acute stressor (Romero & Beattie 2021), GCs such as cortisol are also thought to play a key role in reproductive processes (Wingfield & Sapolsky 2003; Breuner, Patterson & Hahn 2008; Bonier et al. 2009b; Ouyang et al. 2011; but see also Bonier et al. 2009a). Thus, there is a need for some caution when equating a high GC level with a highly stressed ‘state’. Notably, studies of GC levels alone will naturally not capture a complete picture of among-individual variation in stress physiology. The HPA/I axis (which culminates in GC secretion) reflects just one of several important
neuro-endocrine stress response pathways (see MacDougall-Shackleton et al. 2019) whose wider
integration, while logistically challenging to study at present, strongly merits attention going forward
(Gormally & Romero 2020).

Considering all traits together, G shows evidence of genetic integration between behavioural
and endocrine components of the stress response. The ends of the major axis of G (g_max) are largely
consistent with expectations given genetic variance in ‘stress coping styles’ (Koolhaas et al. 1999).
Genotypes showing (putatively) more proactive ‘flight’ type behaviour in the OFT (i.e. thigmotaxis,
high track length, low relative area covered, few freezings) produce higher levels of cortisol following
the handling and confinement stressor, while genotypes showing (putatively) more reactive ‘freezing’
type behaviour in the OFT (i.e., low track length and many freezings, in addition to higher relative area
covered and more time in the middle) produce lower levels of cortisol following the same stressor.
Were one to accept this view, the trait loadings on g_max would suggest that thigmotaxis and freezing
behaviour are strong indicators of stress coping style. Notably, here we find continuous variation
along this major axis of genetic variation (see Figure 4—figure supplement 2), rather than the bimodal
distribution suggested by some of the coping styles literature (Koolhaas et al. 1999, 2007, 2010). An
alternative interpretation, however, is that g_max represents variance in stress response ‘magnitude’
rather than coping style per se. This is because the putatively ‘reactive’ end of g_max might simply reflect
a low magnitude stress response rather than a particular style of ‘coping’ with a stressor. Such an
interpretation would help explain why fish (or genotypes) that produce lower levels of cortisol
following handling and confinement also have more ‘exploratory’ movement (higher relative area
covered) and reduced thigmotaxis (i.e., increased time in the middle) in the OFT. We note that the
distinction between ‘style’ and ‘responsiveness’ may be rather moot if, for example, fish with proactive
styles are also more responsive.

In fact, a subsequent ‘two-tier’ iteration of the coping style model proposed that variation in
the magnitude of the stress response (termed ‘stress reactivity’) could be viewed as a second
dimension of variation, distinct from differences in the ‘type’ (or style) of response (Koolhaas et al.
2010). It is not clear to us that these two dimensions, if both present, can be tested for and
disentangled empirically in the current data—for example, there is no a priori expectation that style
and magnitude should manifest as orthogonal axes in $G$ among traits analysed here (and so both could conceivably align on $g_{\text{max}}$). In principle, disentangling these dimensions should be possible by characterising stress response through within-subject comparison of behaviour and endocrine state between stressed and 'baseline' (unstressed) states. In practice, we found here that genetic variation in GC responsiveness (i.e., the slope across states, accepting that habituation represents a form of this) is very strongly correlated with the average (or intercept). We found similar results for behavioural responses to increased stressor severity in another study of this population, where there exist strong among-individual correlations between the intercept and slope of the plastic response (Houslay et al. 2018). While we cannot claim a true 'baseline' state in either case, these results suggest that stress coping style and the magnitude of response do both exist as axes of variation, but may also be tightly correlated.

The genetic integration of behaviour and physiology detected here is consistent with (but not proof of) the hypothesis that correlational selection in the past has led to the coevolution of these components of the stress response. This hypothesis assumes that correlational selection favours any combination of trait values that yield higher fitness, creating a genetic correlation as mutations that generate such combinations will be recruited into the population while those that do not will be selected out (Lande 1980; Roff & Fairbairn 2012). While reflections on the role of past correlational selection are necessarily speculative, we can say that the structure of $G$ should shape (and potentially constrain) future evolutionary responses to selection—whether natural or artificial. Here we have no direct knowledge of how contemporary selection is acting in the wild or whether it might be changing as a consequence of anthropogenic stressors. Nor do we know exactly how well our estimate of $G$ will match that which may be found in the wild (though our study animals are all recent descendants of wild-collected animals). Thus we cannot comment directly on how $G$ will shape future evolution of the guppy stress response beyond noting that selection on behaviour will cause correlated evolution of endocrine physiology (and vice versa). Nonetheless, while it seems reasonable to expect that current integration of stress response in natural populations should be broadly adaptive, this seems less likely in captive populations (at least for species without a long history of domestication and opportunity for adaptation to artificial environments). We know that prolonged, chronic activation of stress response
pathways (notably the HPA/I axis) frequently disrupts health and survival in captive animals (Huether 1996; Boonstra 2013). It may be that more stress-responsive genotypes are disadvantaged in novel artificial conditions (e.g., if acute stress responsiveness positively predicts susceptibility to chronic stress). However, even if true this would not imply high (acute) stress-responsiveness was also disadvantageous in the wild. Since natural selection should purge alleles that are universally detrimental, it seems more plausible that genetic variation along the major axis described here is maintained by some form of selective trade-off (as widely hypothesised for maintenance of personality variation; e.g., Stamps 2007; Wolf et al. 2007; Réale et al. 2010). For instance, genotypes susceptible to harm under chronic stressor exposure will likely persist in populations if they also confer advantages under an acute stress challenge. In natural populations not only is exposure to acute stressors more common than to chronic stressors, but also selection through chronic stress exposure may be conditional on (and subsequent to) surviving acute challenges (such as predator attacks).

Conclusions and future directions

Here we find evidence for genetic variation in – and integration of – behavioural and physiological (endocrine) components of the stress response. Overall we consider the structure of $G$ to be broadly consistent with the widely invoked ‘reactive—proactive’ model of variation in stress coping style (Koolhaas et al. 2007). This interpretation rests largely on the structure of behavioural variation revealed by the OFT, which is dominated by a major axis running from genotypes with more proactive ‘flight’ type stress responses to those with more reactive ‘freeze’ type stress responses. Endocrine traits align with this axis: genotypes exhibiting ‘flight’ type responses show higher cortisol levels (and exhibit faster habituation of GC physiology) when subjected to repeated handling and confinement stressors than those exhibiting more ‘freeze’ type responses. However, as these latter genotypes with more ‘freeze’ type responses and low cortisol responses to stressors also tend – on average – to display space use patterns characteristic of exploration and reduced thigmotaxis (potentially indicative of being relatively unstressed), an alternative interpretation is that $g_{\text{max}}$ primarily reflects differences in stress responsiveness. Further distinguishing between ‘style’ and ‘responsiveness’ may depend on their association, for example if fish with proactive styles are also more responsive.
Although future studies could certainly target separation of these dimensions further, we think greater insights may come from expanding the set of traits (in particular, to include other components of the physiological stress response, as noted above) and/or stress contexts (i.e. exposing the subjects to different types of stressor). With respect to the latter, here we observed behavioural and GC responses to two different stressors separated in time. We think that our demonstration of genetic correlation structure between behaviours tested in one context and physiology assayed in a different one adds at least some support to the idea that the integration characterised here may ultimately prove generalisable across stress contexts.

Our results suggest that continued evolution of stress-related behaviour will have consequences for glucocorticoid physiology and vice versa. Determining contemporary selection on the stress response, and testing the possibility that its underlying genetic variation is maintained by fitness trade-offs, is thus an obvious – if empirically challenging – next step to understanding the functional importance of genetic variation in the stress response in wild populations. In a more applied context, integration of behavioural and endocrine stress response components at the genetic level has potential utility for artificial selection to improve resilience to chronic stressors in managed populations (Gebauer et al. 2021). Specifically, it may be possible to identify non-invasive, high throughput, behavioural biomarkers and target them in selection schemes to reduce chronic activation of the HPA/I endocrine axis and its attendant deleterious effects.
Materials and methods

Husbandry and breeding

We used fish taken from our captive population housed at the University of Exeter’s Penryn campus, which is descended from wild fish collected in 2008 from the lower Aripo River in Trinidad. This population has been maintained at a population size of several thousand and has undergone no deliberate selection or inbreeding. All fish are fed to satiation twice daily (0800 – 1000h and again at 1600 – 1800h) using commercial flake food and live *Artemia nauplii*. Water temperature is maintained at 23-24°C in well-aerated closed system tank stacks that undergo 25% water changes each week and with weekly tests for ammonia, nitrate and nitrite levels. Lighting is kept at a 12:12 light/dark cycle.

The experiment described in this study was carried out in accordance with the UK Animals (Scientific Procedures) Act 1986 under licence from the Home Office (UK), and with local ethical approval from the University of Exeter.

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

Within 24h of a female producing a brood we recorded her weight (g) and brood size. We kept juvenile fish in full-sib family groups in 2.8L tanks before moving them to 15L 'growth' tanks at an average age of 56 days. At an average age of 133 days (range 59-268) we tagged individuals and placed them into mixed family groups of 16-20 adults (with an even mix of males and females), kept in 15L tanks. Note that variation in tagging age arose largely because groups were necessarily established sequentially as sufficient individuals from multiple families reached a large enough size that we deemed the procedure to be safe. Each adult group comprised a mix of fish from different families, reducing the potential for common environment effects to upwardly bias our genetic parameter estimation.

Overview of behavioural phenotyping

Behavioural phenotyping commenced at least one week after tagging. In all trials, we filmed movement behaviour of individual fish using a Sunkwang video camera equipped with a 6-60mm manual focus lens suspended over the tank. We used the tracking software Viewer II (BiObserve) to extract behavioural data from each recording (detailed below). The tank was lit from below using a light box and screened with a cardboard casing to prevent external visual disturbance. After each behavioural trial, the individual tested was weighed and then moved to a temporary 'holding tank'. Once a full group (as described above) had been tested, all were moved from the holding tank back to their home tank. We replaced the water in the testing and holding tanks between groups to reduce the build-up of hormones or other chemicals. The first offspring generation experienced 4 repeat open field trials (OFTs) over a 2-week period, with at least 48h between trials. Subsequent generations experienced 4 repeat behavioural trials, alternating 2 OFTs with 2 emergence trials (ETs). For the final 2 generations, we extended the OFTs by including a shoaling trial (ST) at the end of each OFT.
Open field trials (OFT) followed the methodology described by White et al (2016). Briefly, we assessed individual behaviour in a 20cm x 30cm tank, filled to a depth of 5cm with room-temperature water from the main supply. We caught fish individually from their home tank, examined them quickly for identification tags, then placed them immediately into the centre of the OFT tank. After allowing 30s for acclimation, we filmed behaviour for 4m30s. Behaviours characterised from the tracking software were track length (the total distance the fish moved during the trial; cm), area covered (the percentage of 1cm x 1cm grid squares through which the fish moved during the trial; %), time in middle (time spent in a rectangular inner zone which was defined as being the same size as an outer area; seconds), and number of freezings (defined in practice as the number of times that an individual's velocity dropped below 4cm/s for a minimum of 2.5 seconds; White & Wilson 2018).

Shoaling trials (ST) were appended to a subset of OFTs, by positioning a small tank containing 10 stock fish (of same sex as the test subject) next to one end of the OFT tank but with visual cues blocked by a cardboard divider. At the end of the normal OFT, we removed this divider slowly, allowing the focal animal to have visual contact with the shoal. We began recording the shoaling trial 30s after removing the divider in order to limit any artefacts of slight disturbance. (Note that we used a further cardboard casing around the shoaling tank to avoid any additional external visual stimulus). We then recorded behaviour of the test fish for an additional 3 minutes. We characterised shoaling tendency via the tracking software by subdividing the tank area into 3 equal-sized rectangular areas: one next to the tank holding the group of same-sex fish, one farthest from this group, and the central area. We then calculated shoaling tendency as the time spent in the 1/3 area closest to the same-sex group after subtracting the time spent in the 1/3 area farthest away. The decision to use a single-sex shoal aimed to reduce any effects of (potential) mate preference and/or avoidance, but also this necessitated replicate arena setups allowing male and female individuals from each group to be tested in the OFT/ST concurrently. We randomised which tank was used for each sex in each group and recorded this information.
Emergence trials (ET) followed the methodology described by White et al. (2016). Briefly, we tested individuals in a 20 cm x 40 cm tank, filled to a depth of 8 cm with room-temperature water from the main supply. A 10 cm section of the tank length was walled off creating a shelter area (20 cm x 10 cm), the walls and floor of which were painted black. The focal fish was placed into the shelter area and allowed to acclimate for 30 s, at which point we opened a sliding door to allow access to the rest of the tank, which was brightly lit from below and otherwise bare. Time to emerge (in seconds) was recorded by the tracking software automatically as the fish exited the shelter area and emerged into the open tank section. Trials were ended either at emergence or at 15 min if the fish had not emerged by that point; in the case of non-emergence, fish were given the maximum value (i.e., 900 s).

Derivation of ‘relative area’ from OFT trials

The ‘area covered’ variable assayed in the OFT is calculated in BiObserve by dividing the arena (i.e., the total area of the tank as viewed from the camera) into 1 cm x 1 cm grid squares. The path taken by the fish during observation is then used to determine what proportion of these grid squares the fish entered. However, we sought to derive a measure of ‘relative area’ that describes whether a fish covers a large, or small area relative to its observed track length.

To do this we simulated ‘random swims’ within the arena across the observed range of track lengths. We first selected 40 OFT results at random from our total data set and extracted the coordinates of the fish in each frame from the raw tracking file, creating a set of x and y movements and their associated distances. As original coordinates were recorded in pixels we used the calibration of the software to convert to cm units. We then use a ‘random walk’ algorithm to select a movement (i.e., step size and direction) from this observed distribution at random and calculate the new coordinates. If the movement keeps the ‘fish’ within the bounds of the ‘tank’ (i.e., defined as a 20 cm x 30 cm arena), the movement is accepted and coordinates added to a movement matrix; if not, a new movement is drawn from the distribution. If the movement is greater than 1 cm in distance, we break the movement into a number of smaller parts to be added to the matrix (such that we capture the coordinates of grid squares through which the ‘fish’ moved along the way). Once the total distance of the random walk reached or exceeded the track length set as the simulation objective, the path is
terminated and the area covered is calculated by counting the number of unique grid squares in the
matrix of coordinates and dividing by the total number possible.

After simulating random walks across 500 values of track length (using a vector of 100 values
evenly spaced across the range of true data, repeated 5 times), we modelled (simulated) area covered
as a fourth order polynomial function of track length. Using this regression model (which explained
97.8% of the variance in simulated data), we calculated the relative area for each actual OFT trial as
the observed area covered minus the area covered under a random swim, as predicted from our
regression model and the observed track length.

Waterborne hormone sampling

On completion of behavioural data collection, individuals entering the endocrine testing program were
left undisturbed for a minimum of two weeks. Waterborne hormone sampling was then conducted
over a 5-day period that included three handling and confinement stressor exposures with 48h
between each. We followed the method described by Houslay et al. (2019) to obtain repeated non-
invasive GC measures of individuals using holding water samples from the first and third
confinements. Note that only two samples per fish were analysed because the financial and time costs
of doing three was deemed prohibitive. We nonetheless applied the stressor stimulus three times as
our prior study showed this was sufficient to produce a strong habituation response, i.e., a significant
decrease in water-borne cortisol over the three sampling periods (Houslay et al. 2019).

We collected samples between 1200 – 1400h to control for diel fluctuations in GC levels. For
each sample, we netted an entire group from their home tank quickly using a large net, transferring
them to 2 holding tanks (containing water from the home tank supply) for moving to an adjacent quiet
room (performed within 20s of the net first hitting the water). We then transferred fish to individual
Pyrex beakers containing 300ml of clean water from the main supply (which serves the main housing
units), which had been warmed to the appropriate temperature (mean = 24.1°C, range 23-24.9°C).
Beakers were placed within cardboard ‘chambers’ to prevent fish from seeing each other or
experiencing outside disturbance. One fish was transferred every 30s, alternating across holding
tanks, such that all fish were in their beakers within 10min of the initial netting. After 60 mins in the
beaker, each fish was removed by pouring its sample through a clean net into a second beaker, with the fish then quickly checked to confirm ID and returned to the holding tank until the entire group could be returned to its home tank.

We immediately filtered each water sample using Grade 1 filter paper (Whatman), then passed them slowly through solid phase C18 extraction columns (Sep-Pak C18, 3cc, Waters) via high-purity tubing (Tygon 2474, Saint Gobain) under vacuum pressure (Earley et al. 2006). Columns were primed beforehand with 2 x 2ml HPLC-grade methanol followed by 2 x 2ml distilled water and were washed afterwards with a further 2 x 2ml distilled water to purge salts. We then covered both ends of each column with film (Parafilm M, Bemis) and stored them at -20°C for future analysis. We washed all beakers, tubes and funnels with 99% ethanol and rinsed them with distilled water prior to each sampling procedure. The remainder of the endocrine assay procedure involved elution, resuspension, separation and quantification of free cortisol by enzyme immunoassay (EIA) using Cayman Chemicals, Inc EIA kits. Detailed methods are exactly as described by Houslay et al. (2019) and so not repeated here (note that here we assayed the free fraction of cortisol only). To validate the cortisol kits, we examined whether the kit standard curve was parallel to a serial dilution curve derived from pooled guppy water-borne hormone extract. 20µl was taken from each of the male samples and pooled; 20µl was taken from each of the female samples and combined into a separate pool. 400µl of the pools was serially diluted from 1:1 to 1:128 and these samples were assayed alongside the kit standard curve on two occasions (June and December 2017, marking the start and finish of sample processing). All dilution curves were parallel to the standard curve (slope comparison test, Zar 1996, p.355; June, male: \( t_{12} = 0.029, P = 0.97 \); June, female: \( t_{12} = 0.343, P = 0.74 \); December, male: \( t_{12} = 0.119, P = 0.91 \); December, female: \( t_{12} = 0.224, P = 0.83 \)). The serial dilution curves also identified 1:32 as an appropriate dilution to ensure that all samples fell on the linear phase of the standard curve. A total of 37 96-well plates were used, and the pooled sample was included at the beginning and end of each plate. Intra-assay coefficients of variation ranged from 0.12-19.83% with a median of 3.08%; the inter-assay coefficient of variation was 19.22%. Cortisol is presented and modelled in (ln-transformed) units of ng/hr to reflect the 1 hour sampling duration.
Statistical methods

All data handling and analysis was performed in R version 3.6.3 (Team 2020). We used the 'tidyverse' set of packages for data handling and visualisation (Wickham 2017), and ASreml-R v4 (Butler 2021) for fitting linear mixed effects models (as described in full below). We also used 'nadiv' for pedigree preparation (Wolak 2012). All models fitted assumed (multivariate) Gaussian error structures, and we visually assessed residuals to verify this was reasonable (after data transformation in some cases). To test for significance of among individual and/or genetic (co)variance components, we fitted nested models with different random effects structures and compared them using likelihood ratio tests (LRTs). We calculated $\chi^2_{nDF}$ as twice the difference in model log likelihoods, with the number of degrees of freedom ($n$) equivalent to the number of additional parameters in the more complex model. When testing a single random effect (variance component), we assumed the difference to be asymptotically distributed as an equal mix of $\chi^2_0$ and $\chi^2_1$ (denoted $\chi^2_{0,1}$; Self and Liang, 1987; Visscher, 2006).

For each OFT, ST and ET behaviour in turn (relative area, time in middle, track length, freezings, shoaling tendency, and emergence time), we used the random effects specification to partition phenotypic variation ($V_p$, conditional on fixed effects as described below) into the effects of additive genetics ($V_a$), permanent environment defined as the non-(additive) genetic component of among-individual differences, $V_{pe}$, and housing group ($V_{group}$), as well as residual variation ($V_{residual}$). We natural log-transformed emergence time prior to analysis to meet assumptions of residual normality and homoscedasticity. For all behavioural traits, we included fixed effects of assay repeat, the order within each group in which the fish was trialled (mean-centred continuous predictor), temperature (mean-centred and scaled to standard deviation units), time (in minutes from midnight, mean-centred and scaled to standard deviation units), age (mean-centred and scaled to standard deviation units), sex, and the generation from the breeding population. The order caught predictor in particular was used to control statistically for variation in disturbance over the course of measuring a group (White & Wilson 2018). For shoaling tendency only, we incorporated an additional fixed effect of setup (as

36
detailed above). We tested the significance of genetic variance for each behaviour by LRT comparison of the corresponding full model to one in which the (additive) genetic random effect was excluded.

Cortisol data were also natural log (ln) transformed for analysis. We formulated a bivariate model to test for both additive genetic variation and genotype-by-environment interaction (GxE) in cortisol levels across the two ‘contexts’ (i.e. samples retained for each individual at first and third confinement, denoted Cortisol1, Cortisol3). Random effects were first used to partition phenotypic (co)variance (conditional on fixed effects) into among-group and residual components. Fixed effects included the context-specific means, and overall effects of the order in which the fish was caught from each group for assay (mean-centred continuous predictor), temperature (mean-centred and scaled to standard deviation units), time of day (mean-centred and scaled to standard deviation units), age (mean-centred and scaled to standard deviation units), and sex. In addition, we included fixed covariates of body mass (mean-centred and scaled to standard deviation units) and a sex by body mass interaction (see Houslay et al. 2019 for rationale of controlling for body size effects on waterborne hormone levels in this way). Note that modelled in this way each individual is sampled only once for each context-specific cortisol trait so no random effect of individual identity is included. To test for additive genetic variation (Va) we compared this first bivariate model to a second formulation that also included the (additive) genetic merit, but under the assumption that this is invariant with context within an individual (such that Va1=Va3 and ra1,3=1 and there is no GxE). We then test for the GxE by comparing the second model to a third in which we allow GxE (i.e., the context-specific genetic variances are free to differ and the cross-context genetic correlation can be <+1).

Lastly, we built a multivariate animal model to estimate G and to test the hypothesised genetic integration among behavioural and physiological stress components. We retained only response traits that harboured significant Va as shown in univariate models, and so the final model comprised response traits relative area, time in middle, track length, freezings (square root transformed), emergence time (ln transformed), and Cortisol (ln transformed). We multiplied (transformed) emergence time by -1 to simplify interpretation (higher values represent faster emergence). We also scaled all (transformed) response variables to standard deviation units. This was to facilitate model fitting, and also prevent scale effects complicating interpretation of eigenvectors of G. Fixed and
random effects were fitted on each trait as specified for the univariate models. Note that one exception
to this is that we elected to treat Cortisol as a single repeated-measures trait here (with two repeats,
one per context) such that a permanent environment effect was now included. Fixed effects estimates
are reported in the supplementary 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), $r_{a(x,y)} = \frac{COV_{a(x,y)}}{\sqrt{(V_{a(x)})} \times \sqrt{(V_{a(y)})}}$. 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 the major axis of variation (G_{max}) could indeed explain most of the
genetic variance in the multivariate phenotype. We estimated uncertainty on the trait loadings
associated with each principal component (eigenvector), as well as on the variance explained by each
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).

Acknowledgements

We are grateful to P. Sharman, D. Maskrey and C. Mnatzagian for help in the lab, and to J. Martin, J. Styga, P. Vullioud and all members of the Wilson group for discussion. We also thank J. McGlothlin and two anonymous reviewers for comments and suggestions that greatly improved this manuscript.
REFERENCES

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

Arnold, S.J., Bürger, R., Hohenlohe, P.A., Aije, B.C. & Jones, A.G. (2008) Understanding the evolution and stability of the G-matrix. Evolution, 62, 2451–61.

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

Berger, D., Bauerfeind, S.S., Blanckenborn, W.U. & Schäfer, M.A. (2011) HIGH TEMPERATURES REVEAL CRYPTOGENETIC VARIATION IN A POLYMORPHIC FEMALE SPERM STORAGE ORGAN. Evolution, 65, 2830–2842.

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

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

Bonier, F., Martin, P.R., Moore, I.T. & Wingfield, J.C. (2009a) Do baseline glucocorticoids predict fitness? Trends in Ecology and Evolution, 24, 634–642.

Bonier, F., Moore, I.T., Martin, P.R. & Robertson, R.J. (2009b) The relationship between fitness and baseline glucocorticoids in a passerine bird. General and Comparative Endocrinology, 163, 208–213.

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

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

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

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

Breuner, C.W., Patterson, S.H. & Hahn, T.P. (2008) In search of relationships between the acute adrenocortical response and fitness. General and Comparative Endocrinology, 157, 288–295.

Broom, D.M. & Johnson, K.G. (1993) Assessing welfare: short-term responses. Stress and Animal Welfare, pp. 87–110. Springer Netherlands, Dordrecht.

Busch, D.S. & Hayward, L.S. (2009) Stress in a conservation context: A discussion of glucocorticoid actions and how levels change with conservation-relevant variables. Biological Conservation, 142, 2844–2853.

Butler, D. (2021) asreml: Fits the Linear Mixed Model.

Carere, C., Groothuis, T.G.G., Möstl, E., Daan, S. & Koolhaas, J.M. (2003) Fecal corticosteroids in a territorial bird selected for different personalities: Daily rhythm and the response to social stress. Hormones and Behavior, 43, 540–548.

Carter, A.J., Feeney, W.E., Marshall, H.H., Cowlishaw, G. & Heinsohn, R. (2013) Animal personality: What are behavioural ecologists measuring? Biological Reviews, 88, 465–475.

Cheverud, J.M. (1982) Phenotypic, Genetic, and Environmental Morphological Integration in the Cranium. Evolution, 36, 499.

Clinchy, M., Sheriff, M.J. & Zanette, L.Y. (2013) Predator-induced stress and the ecology of fear. Functional Ecology, 27, 56–65.

Cox, R.M., McGlothlin, J.W., & Bonier, F. (2016) Hormones as Mediators of Phenotypic and Genetic Integration: An Evolutionary Genetics Approach. Integrative and Comparative Biology, 56, 126–137.

Earley, R.L., Edwards, J.T., Aseem, O., Felton, K., Blumer, L.S., Karom, M. & Grober, M.S. (2006) Social
interactions tune aggression and stress responsiveness in a territorial cichlid fish (Archoecentrus nigrofasciatus). *Physiology and Behavior, 88*, 353–363.

Frankham, R., Hemmer, H., Ryder, O.A., Cothran, E.G., Soulé, M.E., Murray, N.D. & Snyder, M. (1986) Selection in captive populations. *Zoo Biology, 5*, 127–138.

Fürtbauer, I., Pond, A., Heistermann, M. & King, A.J. (2015) Personality, plasticity and predation: linking endocrine and behavioural reaction norms in stickleback fish. *Functional Ecology, n/a–n/a.*

Gebauer, T., Gebauer, R., Palińska-Żarska, K., Cisář, P., Movchan, O., Tomášek, O., Prokešová, M., Matoušek, J., Hliwa, P., Król, J., Žarski, D., Rebl, A. & Stejskal, V. (2021) Assessment of behavioural and physiological traits as indicators of suitability for European perch aquaculture. *Aquaculture, 544.*

Gormally, B.M.G. & Romero, L.M. (2020) What are you actually measuring? A review of techniques that integrate the stress response on distinct time-scales. *Functional Ecology, 34*, 2030–2044.

Guindre-Parker, S. (2020) Individual variation in glucocorticoid plasticity: Considerations and future directions. *Integrative and Comparative Biology, 60*, 79–88.

Hau, M. & Goymann, W. (2015) Endocrine mechanisms, behavioral phenotypes and plasticity: known relationships and open questions. *Frontiers in Zoology, 12*, S7.

Herbert-Read, J.E., Rosén, E., Szorkovszky, A., Ioannou, C.C., Rogell, B., Perna, A., Rammarine, I.W., Kotrschal, A., Kolm, N., Krause, J. & Sumpter, D.J.T. (2017) How predation shapes the social interaction rules of shoaling fish. *Proceedings of the Royal Society B: Biological Sciences, 284*, 20171126.

Hine, E., Chenoweth, S.F. & Blows, M.W. (2004) Multivariate quantitative genetics and the lek paradox: genetic variance in male sexually selected traits of Drosophila serrata under field conditions. *Evolution, 58*, 2754–62.

von Horsten, S., Karl, T. & Pabst, R. (2003) Behavioral phenotyping of mice in pharmacological and toxicological research. *Experimental and Toxicologic Pathology, 55*, 69–83.

Houslay, T.M., Earley, R.L., Young, A.J. & Wilson, A.J. (2019) Habituation and individual variation in the endocrine stress response in the Trinidadian guppy (Poecilia reticulata). *General and Comparative Endocrinology, 270*, 113–122.

Houslay, T.M., Vierbuchen, M., Grimmer, A.J., Young, A.J. & Wilson, A.J. (2018) Testing the stability of behavioural coping style across stress contexts in the Trinidadian guppy. *Functional Ecology, 32*, 424–438.

Houslay, T.M. & Wilson, A.J. (2017) Avoiding the misuse of BLUP in behavioural ecology. *Behavioral Ecology, 00*, 1–5.

Huether, G. (1996) The central adaptation syndrome: Psychosocial stress as a trigger for adaptive modification of brain structure and brain function. *Progress in Neurobiology, 48*, 569–612.

Hunt, J., Blows, M.W., Zajitschek, F., Jennions, M.D. & Brooks, R.C. (2007) Reconciling strong stabilizing selection with the maintenance of genetic variation in a natural population of black field crickets (Telegryllus commodus). *Genetics, 177*, 875–80.

Jones, A.G., Arnold, S.J. & Bürger, R. (2003) Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution, 57*, 1747–1760.

Jones, R.B., Satterlee, D.G. & Ryder, F.H. (1994) Fear of Humans in Japanese-Quail Selected for Low or High Adrenocortical-Response. *Physiology & Behavior, 56*, 379–383.

Kasper, C., Ribeiro, D., Almeida, A.M. De, Larzul, C., Liaubet, L. & Murani, E. (2020) Emphasis on Stress Response and Damaging Behaviour in Pigs. *Genes, 11*, 1–24.

Ketterson, E.D., Atwell, J.W. & McGlothlin, J.W. (2009) Phenotypic integration and independence: Hormones, performance, and response to environmental change. *Integrative and Comparative Biology, 49*, 365–379.

Koolhaas, J.M. (2008) Coping style and immunity in animals: Making sense of individual variation. *Brain, Behavior, and Immunity, 22*, 662–667.

Koolhaas, J.M., de Boer, S.F., Buwalda, B. & Van Reenen, K. (2007) Individual variation in coping with stress: A multidimensional approach of ultimate and proximate mechanisms. *Brain, Behavior and Evolution, 70*, 218–226.

Koolhaas, J.M., de Boer, S.F., Coppens, C.M. & Buwalda, B. (2010) Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. *Frontiers in Neuroendocrinology, 31,*
Paaby, A.B. & Rockman, M.V. (2014) Cryptic genetic variation: Evolution's hidden substrate. *Nature Reviews Genetics, 15*, 247–258.

Perals, D., Griffin, A.S., Bartomeus, I. & Sol, D. (2017) Revisiting the open-field test: what does it really tell us about animal personality? *Animal Behaviour, 123*, 69–79.

Pottinger, T.G. & Carrick, T.R. (1999) Modification of the plasma cortisol response to stress in rainbow trout by selective breeding. *General and Comparative Endocrinology, 116*, 122–132.

Prentice, P.M., Houlsay, T.M., Martin, J.G.A. & Wilson, A.J. (2020) Genetic variance for behavioural ‘predictability’ of stress response. *Journal of Evolutionary Biology, 1–11.

Réale, D., Garant, D., Humphries, M.M., Bergeron, P., Careau, V. & Montgilio, P.-O. (2010) Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 365*, 4051–4063.

Rex, A., Voigt, J.P., Voits, M. & Fink, H. (1998) Pharmacological evaluation of a modified open-field test sensitive to anxiolytic drugs. *Pharmacology Biochemistry and Behavior, 59*, 677–683.

Roff, D.A. (1992) *The Evolution of Life Histories: Theory and Analysis.* Chapman & Hall.

Roff, D.A. & Fairbairn, D.J. (2012) A test of the hypothesis that correlational selection generates genetic correlations. *Evolution, 66*, 2953–60.

Romero, L.M. (2004) Physiological stress in ecology: Lessons from biomedical research. *Trends in Ecology and Evolution, 19*, 249–255.

Romero, L.M. & Beattie, U.K. (2021) Common myths of glucocorticoid function in ecology and conservation. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology, 1–8.*

Romero, L.M., Dickens, M.J. & Cyr, N.E. (2009) The reactive scope model - A new model integrating homeostasis, allostasis, and stress. *Hormones and Behavior, 55*, 375–389.

Sadoul, B., Blumstein, D.T., Alfonso, S. & Geoffroy, B. (2021) Human protection drives the emergence of a new coping style in animals. *PLoS Biology, 19*, 1–11.

Segerstrom, S.C. & Miller, G.E. (2004) Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological bulletin, 130*, 601–30.

Self, S.G. & Liang, K. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc., 82*, 605–610.

Sih, A. & Bell, A.M. (2008) Insights for behavioral ecology from behavioral syndromes. *Advances in the Study of Behavior*

Stamps, J.A. (2007) Growth-mortality tradeoffs and "personality traits" in animals. *Ecology Letters, 10*, 355–363.

Stearns, S.C. (1989) Trade-Offs in Life-History Evolution. *Functional Ecology, 3*, 259–268.

Stirling, D.G., Réale, D. & Roff, D.A. (2002) Selection, structure and the heritability of behaviour. *Journal of Evolutionary Biology, 15*, 277–289.

Taborsky, B., English, S., Fawcett, T.W., Kuiper, B., Leimar, O., McNamara, J.M., Ruuskanen, S. & Sandi, C. (2020) Towards an Evolutionary Theory of Stress Responses. *Trends in Ecology and Evolution, 1*, 1–10.

Taff, C.C. & Vitousek, M.N. (2016) Endocrine Flexibility: Optimizing Phenotypes in a Dynamic World? *Trends in Ecology and Evolution, 31*, 476–488.

Tarlow, E.M. & Blumstein, D.T. (2007) Evaluating methods to quantify anthropogenic stressors on wild animals. *Applied Animal Behaviour Science, 102*, 429–451.

Team, R.C. (2020) R: A Language and Environment for Statistical Computing.

Thomson, J.S., Watts, P.C., Pottinger, T.G. & Sneddon, L.U. (2011) Physiological and genetic correlates of boldness: Characterising the mechanisms of variation in rainbow trout. *Onchorhynchus mykiss. Hormones and Behavior, 59*, 67–74.

Toms, C.N., Echevarria, D.J. & Jouandot, D.J. (2010) A Methodology Review of Personality-Related Studies in Fish: Focus on the Shy-Bold Axis of Behaviour. *Berkeley Planning Journal, 26*, 12.

Trenzado, C.E., Carrick, T.R. & Pottinger, T.G. (2003) Divergence of endocrine and metabolic responses to stress in two rainbow trout lines selected for differing cortisol responsiveness to stress. *General and Comparative Endocrinology, 133*, 332–340.

Veenema, A.H., Meijer, O.C., De Kloet, E.R., Koolhaas, J.M. & Bohus, B.G. (2003) Differences in basal and stress-induced HPA regulation of wild house mice selected for high and low aggression. *Hormones and Behavior, 43*, 197–204.

Visscher, P.M. (2006) A Note on the Asymptotic Distribution of Likelihood Ratio Tests to Test Variance
Components. *Twin Research and Human Genetics*, 9, 490–495.

Walsh, B. & Blows, M.W. (2009) Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and Systematics*, 40, 41–59.

White, S.J., Houslay, T.M. & Wilson, A.J. (2018) Evolutionary genetics of personality in the Trinidadian guppy II: sexual dimorphism and genotype-by-sex interactions. *Heredity*.

White, S.J., Kells, T.J. & Wilson, A.J. (2016) Metabolism, personality and pace of life in the Trinidadian guppy, Poecilia reticulata. *Behaviour*, 153, 1517–1543.

White, S.J., Pascall, D.J. & Wilson, A.J. (2020) Towards a comparative approach to the structure of animal personality variation. *Behavioral Ecology*, 31, 340–351.

White, S.J. & Wilson, A.J. (2018) Evolutionary genetics of personality in the Trinidadian guppy I: maternal and additive genetic effects across ontogeny. *Heredity*.

Wickham, H. (2017) tidyverse: Easily Install and Load “Tidyverse” Packages.

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

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

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

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

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

Wingfield, J.C. & Sapolsky, R.M. (2003) Reproduction and resistance to stress: when and how. *Journal of neuroendocrinology*, 15, 711–724.

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

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

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

Table S1: Fixed effects estimates from the full multivariate animal model.

Table S2: Permanent environment (co)variance matrix from the full multivariate animal model.

Table S3: Group (co)variance matrix from the full multivariate animal model.

Table S4: Residual variance-correlation matrix from the full multivariate animal model.

Table S5: Among-individual (co)variance matrix from the multivariate model that excluded genetic effects.

Table S6. Table showing full description of eigen decomposition of the $G$ matrix.

Figure 4—figure supplement 1: Scree plot showing the proportion of total variance explained by each
eigen vector, with vertical bars indicating 95% confidence intervals as estimated from 5000
bootstrapped replicates. The horizontal dotted line indicates where equal variance across all
eigen vectors would lie.

Figure 4—figure supplement 2: Histogram of PC scores for $g_{\text{max}}$ calculated for each individual that was
assayed for both behaviour and endocrine responses. The distribution is skewed but clearly not
bimodal, as is predicted by some verbal models of ‘stress coping style’.

Supplementary File 1: Selection of OFT behaviours.
Table S1. Fixed effects estimates from the full multivariate animal model.

| Trait                          | Parameter                     | Estimate | SE   | z     |
|-------------------------------|-------------------------------|----------|------|-------|
| Relative area                 | (Intercept)                   | -0.198   | 0.153| -1.291|
|                               | Age                           | 0.04     | 0.04 | 1.003 |
|                               | Generation (1)                | -0.118   | 0.107| -1.035|
|                               | Generation (2)                | 0.268    | 0.115| 2.327 |
|                               | Generation (3)                | 0.657    | 0.282| 2.331 |
|                               | Generation (S F1)             | 0.868    | 0.274| 3.166 |
|                               | Order                         | -0.004   | 0.004| -1.067|
|                               | Repeat                        | -0.07    | 0.015| -4.584|
|                               | Sex (Male)                    | 0.142    | 0.039| 3.614 |
|                               | Temperature                   | -0.026   | 0.023| -1.161|
|                               | Time of day                   | 0.18     | 0.122| 1.47  |
| Time in the middle            | (Intercept)                   | -0.344   | 0.156| -2.209|
|                               | Age                           | 0.084    | 0.041| 2.047 |
|                               | Generation (1)                | 0.064    | 0.11  | 0.578 |
|                               | Generation (2)                | 0.359    | 0.12  | 3     |
|                               | Generation (3)                | 0.582    | 0.287| 2.028 |
|                               | Generation (S F1)             | 1.077    | 0.279| 3.865 |
|                               | Order                         | -0.004   | 0.004| -0.943|
|                               | Repeat                        | -0.098   | 0.015| -6.344|
|                               | Sex (Male)                    | 0.346    | 0.04  | 8.721 |
|                               | Temperature                   | 0.025    | 0.023| 1.109 |
|                               | Time of day                   | 0.271    | 0.124| 2.184 |
| Track length                  | (Intercept)                   | -0.694   | 0.143| -4.861|
|                               | Age                           | -0.167   | 0.038| -4.436|
|                               | Generation (1)                | 0.567    | 0.097| 5.871 |
|                               | Generation (2)                | 0.469    | 0.106| 4.41  |
|                               | Generation (3)                | 0.703    | 0.266| 2.646 |
|                               | Generation (S F1)             | 0.233    | 0.257| 0.907 |
|                               | Order                         | -0.009   | 0.004| -2.419|
|                               | Repeat                        | 0.147    | 0.015| 10.05 |
|                               | Sex (Male)                    | -0.155   | 0.04  | -3.862|
|                               | Temperature                   | 0.04     | 0.021| 1.894 |
|                               | Time of day                   | 0.028    | 0.115| 0.247 |
| √Freezings                    | (Intercept)                   | 0.431    | 0.158| 2.731 |
|                               | Age                           | 0.157    | 0.041| 3.862 |
|                               | Generation (1)                | -0.236   | 0.105| -2.236|
|                               | Generation (2)                | -0.252   | 0.116| -2.163|
|                               | Generation (3)                | -0.193   | 0.293| -0.66 |
|                               | Generation (S F1)             | 0.088    | 0.284| 0.311 |
|                               | Order                         | -0.002   | 0.004| -0.518|
|                               | Repeat                        | -0.138   | 0.016| -8.45 |
|                               | Sex (Male)                    | -0.022   | 0.04  | -0.535|
|                               | Temperature                   | 0.037    | 0.023| 1.616 |
|                               | Time of day                   | 0.058    | 0.127| 0.456 |
| -ln Emergence time            | (Intercept)                   | 0.087    | 0.334| 0.26  |
|                               | Age                           | 0.002    | 0.045| 0.046 |
|                               | Generation (3)                | -0.713   | 0.668| -1.068|
|                               | Generation (S F1)             | -0.593   | 0.668| -0.899|
|                               | Order                         | -0.004   | 0.005| -0.737|
|                               | Repeat                        | 0.073    | 0.045| 1.597 |
|                               | Sex (Male)                    | 0.259    | 0.052| 4.988 |
|                               | Temperature                   | 0.08     | 0.039| 2.048 |
|                               | Time of day                   | -0.265   | 0.326| -0.811|
| ln Cortisol                   | (Intercept)                   | 1.054    | 0.334| 3.154 |
|                               | Age                           | 0        | 0.029| 0.007 |
|                               | Order                         | -0.001   | 0.004| -0.318|
|                               | Sex (Male)                    | -0.846   | 0.079| -10.668|
|                               | Stressor exposure (first)      | 0.672    | 0.035| 19.316|
|                               | Temperature                   | -0.1     | 0.048| -2.1  |
|                               | Time of day                   | -1.337   | 0.465| -2.876|
|                               | Weight                        | 0.261    | 0.055| 4.74  |
|                               | Weight x Sex (Male)           | -0.150   | 0.102| -1.554|
**Table S2.** Permanent environment (co)variance matrix from the full multivariate animal model.

|                       | Relative area | Time in the middle | Track length | √Freezings | -ln(Emergence time) | In Cortisol |
|-----------------------|---------------|--------------------|--------------|------------|---------------------|-------------|
| Relative area         | 0.15 ± 0.03   |                    |              |            |                     |             |
| Time in the middle    | 0.10 ± 0.03   | 0.14 ± 0.03        |              |            |                     |             |
| Track length          | -0.14 ± 0.02  | -0.10 ± 0.02       | 0.18 ± 0.03  |            |                     |             |
| √Freezings            | 0.10 ± 0.02   | 0.10 ± 0.03        | -0.10 ± 0.03 | 0.10 ± 0.03|                     |             |
| -ln(Emergence time)   | 0.05 ± 0.03   | -0.01 ± 0.03       | -0.02 ± 0.03 | -0.01 ± 0.03| 0.07 ± 0.04        |             |
| In Cortisol           | 0.03 ± 0.03   | 0.09 ± 0.03        | -0.04 ± 0.03 | 0.05 ± 0.03 | -0.03 ± 0.03       | 0.08 ± 0.04 |
Table S3. Group (co)variance matrix from the full multivariate animal model.

| Relative area | Time in the middle | Track length | √Freezings | -ln Emergence time | ln Cortisol |
|---------------|-------------------|--------------|------------|--------------------|------------|
| Relative area | 0.05 ± 0.01       |              |            |                    |            |
| Time in the middle | 0.03 ± 0.01       | 0.06 ± 0.02  |            |                    |            |
| Track length  | -0.02 ± 0.01      | -0.03 ± 0.01 | 0.03 ± 0.01|                    |            |
| √Freezings    | 0.01 ± 0.01       | 0.03 ± 0.01  | -0.02 ± 0.01| 0.05 ± 0.01        |            |
| -ln Emergence time | -0.00 ± 0.01      | -0.02 ± 0.01 | 0.01 ± 0.01 | -0.03 ± 0.01       | 0.03 ± 0.01|
| ln Cortisol   | -0.01 ± 0.01      | 0.02 ± 0.01  | -0.03 ± 0.01| 0.02 ± 0.01        | 0.01 ± 0.01| 0.07 ± 0.03 |

Table S4. Residual variance-correlation matrix from the full multivariate animal model.

| Relative area | Time in the middle | Track length | √Freezings | -ln Emergence time | ln Cortisol |
|---------------|-------------------|--------------|------------|--------------------|------------|
| Relative area | 0.55 ± 0.02       | -0.64 ± 0.01 | 0.41 ± 0.02| -                  | -          |
| Time in the middle | 0.65 ± 0.01       | -0.59 ± 0.01 | 0.51 ± 0.02| -                  | -          |
| Track length  | 0.56 ± 0.02       | -0.59 ± 0.01 | 0.51 ± 0.02| -                  | -          |
| √Freezings    | 0.50 ± 0.02       | -0.68 ± 0.01 | -          | -                  | -          |
| -ln Emergence time | 0.63 ± 0.02       | 0.79 ± 0.04  | -          | 0.35 ± 0.02        | -          |
| ln Cortisol   |                   |              |            |                    |            |
Table S5. Among-individual (co)variance matrix from the multivariate model excluding genetic effects.

|                | Relative area | Time in the middle | Track length | √Freezings | -ln Emergence time | Shoaling tendency | ln Cortisol |
|----------------|---------------|--------------------|--------------|------------|--------------------|-------------------|------------|
| Relative area  | 0.25 (0.211,0.288) | 0.747 (0.692,0.801) | -0.728 (-0.78,-0.674) | 0.49 (0.392,0.576) | 0.032 (-0.219,0.273) | 0.198 (-0.054,0.436) | -0.002 (-0.186,0.16) |
| Time in the middle | 0.19 (0.154,0.223) | 0.258 (0.216,0.298) | -0.645 (-0.702,-0.577) | 0.601 (0.522,0.68) | -0.134 (-0.373,0.113) | 0.025 (-0.22,0.287) | 0.123 (-0.05,0.299) |
| Track length   | -0.203 (-0.237,-0.167) | -0.182 (-0.218,-0.148) | 0.31 (0.265,0.35) | -0.76 (-0.808,-0.709) | 0.168 (-0.046,0.399) | -0.053 (-0.286,0.168) | -0.003 (-0.159,0.156) |
| √Freezings     | 0.123 (0.09,0.155) | 0.153 (0.118,0.187) | -0.212 (-0.248,-0.173) | 0.251 (0.210,0.297) | -0.333 (-0.605,-0.101) | -0.066 (-0.307,0.185) | -0.038 (-0.231,0.131) |
| -ln Emergence time | 0.006 (-0.038,0.044) | -0.024 (-0.066,0.017) | 0.033 (-0.001,0.073) | -0.059 (-0.101,-0.02) | 0.127 (0.060,0.193) | 0.319 (0.001,0.666) | -0.236 (-0.502,0.015) |
| Shoaling tendency | 0.043 (-0.009,0.093) | 0.006 (-0.048,0.058) | -0.013 (-0.007,0.036) | -0.014 (-0.006,0.036) | 0.049 (0.004,0.098) | 0.185 (0.102,0.267) | -0.01 (-0.243,0.237) |
| ln Cortisol    | 0 (-0.038,0.033) | 0.026 (-0.011,0.063) | -0.001 (-0.037,0.035) | -0.008 (-0.046,0.028) | -0.035 (-0.069,0.002) | -0.002 (-0.045,0.037) | 0.176 (0.133,0.222) |

Shoaling tendency is included as there is significant among-individual variance in this trait. Among-individual variances provided on the shaded diagonal, with among-individual covariances below and among-individual correlations above. 95% confidence intervals in parentheses are estimated from 5000 bootstrapped replicates.
Table S6. Table showing full description of eigen decomposition of the \( G \) matrix.

| Eigen vectors | 1    | 2    | 3    | 4    | 5    | 6    |
|---------------|------|------|------|------|------|------|
| Eigen value   | 0.47 | 0.16 | 0.11 | 0.04 | 0.01 | 0.01 |
| Percentage of total | 59.5% | 20.2% | 14.1% | 4.5% | 1.0% | 0.7% |
| Trait loadings |      |      |      |      |      |      |
| Relative area covered | -0.33 | -0.53 | -0.33 | 0.32 | -0.24 | 0.58 |
| Time in the middle | -0.45 | -0.50 | 0.10 | -0.41 | 0.57 | -0.20 |
| Track length   | 0.52 | -0.17 | 0.26 | 0.40 | 0.61 | 0.32 |
| sqrt(Freezeings) | -0.50 | 0.62 | 0.17 | -0.03 | 0.28 | 0.51 |
| ln(Emergence time) | 0.21 | -0.20 | 0.58 | -0.54 | 0.35 | 0.42 |
| ln(Cortisol)   | 0.35 | 0.11 | -0.67 | -0.53 | 0.21 | 0.30 |
Figure 4—figure supplement 1: Scree plot showing the proportion of total variance explained by each eigen vector, with vertical bars indicating 95% confidence intervals as estimated from 5000 bootstrapped replicates. The horizontal dotted line indicates where equal variance across all eigen vectors would lie.
Figure 4—figure supplement 2: Histogram of PC scores for \( g_{\text{max}} \) calculated for the subset of individuals that were assayed for both behaviour and endocrine responses. The distribution is clearly not bimodal, as is predicted by some verbal models of 'stress coping style'.