RESEARCH ARTICLE

Repeatability of locomotor performance and morphology–locomotor performance relationships

Cara Conradsen[1], Jeffrey A. Walker[2], Catherine Perna[1] and Katrina McGuigan[1],*

ABSTRACT

There is good evidence that natural selection drives the evolution of locomotor performance, but the processes that generate the among-individual variation for selection to act on are relatively poorly understood. We measured prolonged swimming performance, U_{crit}, and morphology in a large cohort (n=461) of wild-type zebrafish (*Danio rerio*) at ~6 months and again at ~9 months. Using mixed-model analyses to estimate repeatability as the intraclass correlation coefficient, we determined that U_{crit} was significantly repeatable (r=0.55; 95% CI: 0.45–0.64). Performance differences between the sexes (males 12% faster than females) and changes with age (decreasing 0.07% per day) both contributed to variation in U_{crit} and, therefore, the repeatability estimate. Accounting for mean differences between sexes within the model decreased the estimate of U_{crit} repeatability to 21% below the naïve estimate, while fitting age in the models increased the estimate to 14% above the naïve estimate. Greater consideration of factors such as age and sex is therefore necessary for the interpretation of performance repeatability in wild populations. Body shape significantly predicted U_{crit} in both sexes in both assays, with the morphology–performance relationship significantly repeatable at the population level. However, morphology was more strongly predicative of performance in older fish, suggesting a change in the contribution of morphology relative to other factors such as physiology and behaviour. The morphology–performance relationship changed with age to a greater extent in males than females.

KEY WORDS: *Danio rerio*, U_{crit}, Prolonged swimming, Performance landscape, Ontogeny, Mixed models

INTRODUCTION

Locomotor performance impacts an individual's fitness through effects on activities such as feeding, migration, mating and predator evasion (Jayne and Bennett, 1990; Irschick and Garland, 2001; Husak and Fox, 2008; Irschick et al., 2008; Careau and Garland, 2012). Common patterns of repeated (parallel or convergent) evolution of the same performance–environment relationships emphasise the role that selection plays in generating among-taxa variation in locomotor performance, and in the physiological, morphological and behavioural traits that determine performance (Taylor and McPhail, 1985; McGuigan et al., 2003; Langerhans and DeWitt, 2004; Langerhans et al., 2006; Dalziel et al., 2012;

Franssen et al., 2013; Fu et al., 2013; da Silva et al., 2014; Haas et al., 2015; Nelson et al., 2015). Despite the adaptive significance of locomotion, how the variation in locomotion is generated among individuals within a population, which is what natural selection acts upon, is relatively poorly understood.

Repeatability is a useful concept for investigating among-individual variation (Falconer, 1981; Boake, 1989; Hayes and Jenkins, 1997). Laboratory assays of individual performance will reflect both ecologically relevant variation in the physiological capacity of individuals to move fast or for long periods of time (Bennett and Huey, 1990; Irschick and Losos, 1998; Plaut, 2001) and short-term variation in physiological state and behaviour (Belke and Garland, 2007; Bahrndorff et al., 2012; Astley et al., 2013). Repeatability analyses can determine the extent to which performance assays measure physiological capacity versus short-term, transient variation.

Repeatability analyses can also be used to assess the potential for within-generation selection to cause between-generation evolution. Evolution by natural selection depends on the sorting of individuals based on their relative fitness and transmission of high-fitness phenotypes from one generation to the next (Robertson, 1966; Price, 1970; Lande, 1979; Falconer, 1981; Lande and Arnold, 1983; Lynch and Walsh, 1998). Repeatability can give insight into both of these parameters (selection and heritability), providing evolutionarily relevant ways of characterising among-individual variation (Falconer, 1981; Lessells and Boag, 1987; Boake, 1989; Dohm, 2002; Davy et al., 2014).

Repeatable maximum sprint or endurance locomotor performances have been reported for a variety of taxa, particularly reptiles, amphibians and fish (Huey and Dunham, 1987; Kolok, 1999; Adolph and Pickering, 2008; Oufiero and Garland, 2009). Although these studies strongly suggest that standard performance assay methods are likely to be measuring physiological capacity (but see Losos et al., 2002; Irschick et al., 2008), several common aspects of the experimental conditions under which repeatability has been estimated limit conclusions about the evolutionary potential of locomotor performance (Hayes and Jenkins, 1997; Wolak et al., 2012; Biro and Stamps, 2015).

First, reflecting issues both with removing large numbers of individuals from natural populations and the time-consuming nature of performance measures, studies have typically involved small sample sizes. Published estimates of repeatability will therefore typically be imprecise, with large confidence limits (Bonett, 2002; Wolak et al., 2012). For example, published estimates of repeatability of swimming performance (both prolonged and sprint) in fish, reviewed by Oufiero and Garland (2009), depended on sample sizes of 6–86 individuals, with two measures per individual. Using the intraclass correlation coefficient R package (ICC) published by Wolak et al. (2012), parameterised by the sample size and repeatability estimate of each study in table 1 of Oufiero and Garland (2009), the confidence interval width in

[1]School of Biological Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia. [2]Department of Biological Sciences, University of Southern Maine, Portland, Maine 04103, USA.

*Author for correspondence (k.mcguigan1@uq.edu.au)

K.M., 0000-0002-0525-2202

these studies ranges from 0.15 up to 0.72 with a median confidence interval of 0.46. Such large confidence intervals indicate that, although locomotor performance is likely to be repeatable, the magnitude of repeatability remains an open question. Furthermore, whether repeatability varies with the time between measures, among different taxonomic groups or among performance types cannot be inferred with such wide confidence intervals around estimates.

Second, variation among individuals in the environment they have been exposed to (including maternal effects), as well as in age and sex will affect estimates of repeatability (Hayes and Jenkins, 1997; Dohm, 2002; Biro and Stamps, 2015). However, the effects of such variables on repeatability have not typically been considered. Estimates of repeatability often come from wild-caught animals where a detailed environmental history is unavailable. Many environmental variables have been shown to contribute to variation in locomotor performance (Garland and Carter, 1994; Hammer, 1995; Elphick and Shine, 1998) and can have persistent effects across an individual's life (Metcalfe and Monaghan, 2001). Although it might seem intuitive that environmental variation will reduce repeatability, differences among individuals in their environmental experiences can generate high among-individual variance relative to within-individual (among repeated measures) variance, leading to high estimates of repeatability. Although environmentally generated variation can be sorted by selection, without a genetic basis this variation does not contribute to evolution.

Third, variation in locomotor capacity with age could confound estimates of repeatability. Few studies have taken into account age-based changes in phenotype between repeated measures (Biro and Stamps, 2015). Locomotor speed and the energy demands of locomotion can vary dramatically across the life cycle as a result of ontogeny, reproduction and senescence (Schmidt-Nielsen, 1972; Fisher et al., 2000; Gibb et al., 2006; Katija et al., 2015). Without a detailed understanding of how performance changes with age, it is not clear how variation in age among assayed individuals might affect performance repeatability measures.

Locomotor performance is determined by morphology, physiology and behaviour, and selection acting on performance is expected to indirectly drive evolution of these contributing traits (Arnold, 1983, 2003; Walker, 2007). Although this conceptualisation as hierarchical form–function–fitness relationships has proved useful in studying divergent evolution among taxa (Wainwright, 2007; Langerhans, 2008; Moen et al., 2013), again, relatively little attention has been paid to within-population variation in these relationships (for examples, see Herrel et al., 2012). Morphology–performance relationships might differ between sexes, even in the absence of sexually selected morphological ornaments (Conradsen and McGuigan, 2015). Ontogenetic variation in locomotor performance and in contributing traits (Carrier, 1996; Herrel and Gibb, 2006) suggest that morphology–performance relationships are unlikely to be static across ontogeny, but the extent to which individual morphological trait contributions are repeatable is generally not known. Several studies have also demonstrated that morphology–performance relationships are variable through time (seasonally) for locomotor (Kolok, 1992a) and feeding (Irschick et al., 2006) performances. If morphology–performance relationships are not repeatable, the potential for selection on performance to drive the evolution of morphology might be limited, even if selection on performance levels remains consistent across the life cycle.

Here, we investigate the repeatability of prolonged (critical) swimming speed, U_{crit} (Brett, 1964; Beamish, 1978), in the freshwater fish, *Danio rerio* Hamilton 1822, and provide the first population-level estimate of the repeatability of morphology–

performance relationships. U_{crit} provides an ecologically relevant measure of prolonged locomotor performance (Kolok, 1999; Plaut, 2001; Wolter and Arlinghaus, 2003; Roche et al., 2013). Population mean U_{crit} is positively correlated with water flow rate (McGuigan et al., 2003; Langerhans, 2008; Haas et al., 2010, 2015), revealing the historical role of flow in shaping freshwater fish diversity. Both local (within catchment) and global environmental changes are expected to alter flow rates in freshwater systems, and particularly to increase the amplitude of flow rate variation (van Vliet et al., 2013). These expected changes in flow rates highlight the need to better understand the processes generating variation in swimming performance within populations in order to be able to predict how they might respond to environment change.

In a large sample of zebrafish, we directly investigate the effect of sources of variation in our experimental design, particularly sex and age, on the estimate of repeatability. We employ mixed-model ANOVA to estimate the repeatability of U_{crit} as the intraclass correlation coefficient, demonstrating the utility of mixed models in improving the understanding of variables affecting repeatability. We further determine the consistency between repeated measures of the morphology–performance relationships within each sex. Finally, we determine whether changes in swimming speed with age can be explained by growth-related changes in body size and shape.

MATERIALS AND METHODS
Experimental animals
This work was conducted following approval of The University of Queensland's Animal Welfare Unit (approval number SBS/107/12/ARC). Wild-type zebrafish (*Danio rerio*) of the WIK strain (Rauch et al., 1997), were bred at The University of Queensland between 10 September and 23 October 2013. Embryos were incubated in Petri dishes at 28°C until they had inflated their swimbladder (~4–5 days post fertilization, dpf), at which point they were transferred to 1 litre no-flow tanks with a polyculture of type-L marine rotifers (*Brachionus plicatilus*) and *Tetraselmis* sp. algae (Best et al., 2010). At ~10 dpf, flow was increased to 5 ml min^{-1}, and fry were fed on rotifers and dry fry food (a mix of INVE O.range Start-S and INVE Lansy Shrimp spirulina+ powder). At ~30 dpf, fish were transferred to 3.5 litre tanks, at which time they were density controlled, randomly culled down to a maximum of 30 fish per tank. Fish were maintained on a recirculating water system at 28°C, fed to satiation three times daily on a dry food diet [O.range Wean-S (200–400 µm) and NRD3/5 (300–500 µm) mixed 2:1, plus 100 g Lansy Shrimp spirulina+ per 1 kg of feed mix]. All feed components were manufactured by INVE Aquaculture, Amphur Pakkred, Nonthaburi, Thailand.

At 3 months old, fish were tagged with visible implant elastomer (VIE) tags (Northwest Marine Technology, Shaw Island, WA, USA) under anaesthesia (AQUI-S New Zealand, Lower Hutt, New Zealand). VIE tags are used extensively to identify individuals or groups of aquatic organisms. VIE tags have been demonstrated not to negatively impact growth or survival of small fish (<40 mm; Leblanc and Noakes, 2012), including zebrafish (Hohn and Petrie-Hanson, 2013). Available evidence from the Sacramento splittail (*Pogonichthys macrolepidotus*, mean±s.e. total length=9.9±0.6 cm; Sutphin et al., 2007) and rainbow trout (*Oncorhynchus mykiss*, 348–563 g; Davidson et al., 1999) suggests that VIE tags do not affect critical swimming speed. Each fish was tagged once with a single colour (a total of six colours were used) on either the left or the right dorsal surface, anterior to the dorsal fin. Fish were allowed at least 2 weeks to recover from tagging before further handling.

Individual fish were moved between tanks (and groups) several times during the experiment. Accurate tracking of each uniquely identified individual fish across this large cohort was confirmed through comparison of photographs, the first taken when fish were assigned an identity and the second taken for the repeated characterisation of morphology. Individually unique pigmentation patterns on body and fins make it possible to identify individual zebrafish. Of the 484 fish photographed twice, 17 could not be identified (and were excluded from all analyses) and a further 16 were initially mis-identified but their true identity was determined through comparison with photographs of candidate fish from the initial assay. The remaining 93% of fish were accurately tracked during the experiment. Data are available in Dryad (doi:10.5061/dryad.b91d1).

Locomotor performance

As described in Conradsen and McGuigan (2015), we used a stepped velocity test to estimate critical swimming speed, U_{crit} (Brett, 1964), following the protocol developed for zebrafish by Plaut (2000). Swimming trials were conducted using a Loligo Systems (Hontzsch, Bondby, Denmark) swimming flume (L×W×H, 40×10×10 cm swim chamber) at 28°C (200 W submersible heater, Hydor THEO, Bassano del Grappa, Vicenza, Italy). Fish were fasted for 24 h prior to swimming. Fish were introduced to the swim chamber with 4 cm s^{-1} flow and allowed to acclimate for 15 min; flow was then increased by 4 cm s^{-1} at 5 min intervals until fish were no longer able to maintain station (Plaut, 2000). U_{crit} was determined as:

$$U_{crit} = u_i + ([t_i/t] x u), \qquad (1)$$

where u_i was the highest velocity at which the fish swam for the full time interval, u is the velocity step increment (4 cm s^{-1}), t_i is the time for which the fish swam at the final (fatigue) velocity and t is the time each water velocity was imposed for (5 min) (Brett, 1964). Experiments were performed between 16 February and 16 April 2014, and again between 17 June and 18 July 2014.

Fish were swum in groups of six, consisting of three males and three females per trial. Zebrafish are a shoaling species (Pritchard et al., 2001; Miller and Gerlai, 2011), found in mixed-sex shoals and can exhibit signs of stress when held individually (Wright et al., 2003). By allowing fish the opportunity to swim in schools, mean U_{crit} might be inflated because of stress mitigation or kinematic benefits of schooling (e.g. Killen et al., 2012; Burgerhout et al., 2013). Fish were swum with a different group of five other fish in their first and second assay. A total of 104 and 79 swimming trials (each with six fish) were conducted in the first and second measurement period, respectively.

Estimating repeatability

Repeatability of an individual's performance can be measured in two ways: as the rank correlation coefficient and as the intraclass correlation coefficient. If the performance rank of individuals is consistent, it suggests that persistent directional selection could effectively sort individuals based on their performance (Boake, 1989; Davy et al., 2014). The intraclass correlation coefficient, the proportion of the total variation that is due to among-individual variation, sets an upper limit on the heritability of a trait (Falconer, 1981; Lessells and Boag, 1987; Boake, 1989; Dohm, 2002), providing insight into the potential for the trait (performance) to respond to selection.

A total of 435 fish (213 males and 222 females) with two measures of U_{crit} were available for analysis. The presence of outliers was assessed using critical values of Mahalanobis distance

(Stevens, 1984) at $P<0.0001$, and outliers were removed from the data prior to analyses (Osborne and Overbay, 2004). For U_{crit}, no outlier values were identified using this criterion. Based on the simulations designed by Wolak et al. (2012) and implemented in 'ICC' R package, this sample size ($n=435$ individuals and $k=2$ repeated measures per individual) is sufficient to estimate the intraclass correlation with a confidence interval of less than 0.2 for values as low as 0.1. Fish ranged in age from 111 days post fertilisation (dpf) for the youngest fish in the first assay period up to 311 dpf for the oldest fish in the second assay period. There was no overlap in age between the repeated-measures datasets. The average time between U_{crit} assays for an individual was 102 days (median 103 days), but ranged from 62 to 150 days.

We determined the rank repeatability of U_{crit} as the Spearman's rank correlation coefficient, ρ, using RStudio (v.0.98.953). We estimated the intraclass correlation, r, of U_{crit} using mixed-model analyses implemented in PROC MIXED in SAS (v.9.4, SAS, Cary, NC, USA). When estimating r, we explored the effect of known sources of variation within the data by including different parameters in the model. The most inclusive model was:

$$y_{ijklm} = \mu + Sex_i + Age_j + Sex \times Age_{ij} + Trial_{k(l)} + Ind_m + \varepsilon_{ijkl}, \qquad (2)$$

where y was the U_{crit} value of the mth individual (Ind) of the ith sex and jth age, assayed in the kth swimming trial within the lth measurement period. The continuous variable of age is confounded with the temporal difference between replicate measures, capturing any differences (whether due to age or other sources) between the two assays of an individual's performance, as well as variation in age among individuals within each measurement period.

We directly estimated r via a repeated-measures form of model (1), fitting a first-order autoregressive covariance structure, AR(1) (Littell et al., 1996), to the random effect of Individual. This estimate of r is the same as that calculated from the estimates of variance (s^2) among (Ind) and within (ε) individuals using the equation: $r=s^2_{Ind}/(s^2_{Ind}+s^2\varepsilon)$ (Lessells and Boag, 1987). Directly estimating the correlation between repeated measures of individuals (i.e. r) within the statistical model allows a more straightforward approach to hypothesis testing. We tested the hypothesis that $r\neq0$ by applying a log-likelihood ratio test to determine whether holding r (the correlation between repeated measures within an individual) to zero (using the PARMS statement in PROC MIXED) significantly increased the log-likelihood ratio compared with the model in which r was estimated from the data. This test statistic follows a χ-square distribution with one degree of freedom (Self and Liang, 1987). Estimates of r and significance tests of this repeatability measure were conducted under restricted maximum likelihood to ensure unbiased estimates of the random effects of interest (within and among individual estimates) (Shaw, 1987; Lynch and Walsh, 1998).

The effects of sex (a categorical effect for males versus females), age (in days, a continuous effect) and their interaction were fitted as fixed effects (Wilson, 2008), whereas the swimming trial (nested within measurement replicate) in which an individual's U_{crit} was measured was fitted as a random effect. We also considered models in which the random effects (individual, trial and residual) were constrained to be homogenous across males and females, or where sex-specific estimates were allowed (using the GROUP statement in PROC MIXED). This allowed us to determine whether males and females differed in the repeatability of swimming performance, and to gain insights into the potential causes of any difference. Models were fitted using maximum likelihood (ML) to allow us to compare

model fit (using Akaike information criterion, AIC) among models with different fixed effects.

Morphology

Body shape was characterised as in Conradsen and McGuigan (2015). Briefly, fish were anaesthetised using isoeugenol (AQUI-S) and photographed in lateral view over a 1 mm scale grid (Cannon PowerShot S110, Sydney, Australia). TpsDig2 (Rohlf, 2005) was used to record the positions of a total of 12 landmarks (Fig. 1). From this landmark data, standard length (SL) was determined as the distance between landmark 1 and landmark 6 (Fig. 1). Landmarks were then aligned by generalised Procrustes alignment, implemented in MorphoJ (Klingenberg, 2011). Landmark alignment generates collinearity (violating assumptions of multivariate analyses), a problem that can be addressed by, for example transforming the data to principal warps (Bookstein, 1989; Rohlf, 1996). However, principal warps are difficult to interpret from a functional perspective since something like 'fish depth' might increase on one warp and decrease on another. Alternatively, inter-landmark distances (ILDs) among aligned landmarks can be used to characterise shape, overcoming the multicollinearity problem while being relatively straightforward to interpret in functional analyses. Here, 10 ILDs among aligned landmarks were used to characterise body shape (Fig. 1). The 10 traits have been used previously to describe body shape variation in zebrafish (Conradsen and McGuigan, 2015) and were chosen because they have high coverage of body regions without repetition and low multi-collinearity, ensuring they are suitable for multivariate analyses.

Based on the critical value of Mahalanobis distance (Stevens, 1984) ($P<0.001$ and d.f.=11), nine outliers (three males within each assay time, one and two females in the first and second assay, respectively), representing 1.0% of the data, were identified and removed prior to further analysis (Osborne and Overbay, 2004). Although these samples had very little effect on the results of the regression analyses, and led to the same conclusions, here, we present only the analyses where the outliers were excluded. Repeated measures of morphology were available for 461 fish (230 males and 231 females); a subset of 426 (207 males and 219 females) fish with repeated measures of both performance and morphology was analysed here.

To determine whether fish grew as they aged, we analysed SL using model (1), without swimming trial, under ML. A multivariate version of model (1) (again, without trial) was applied to test for a change in body shape with age. For the multivariate body shape data, an unconstrained covariance structure was specified for the random effects (individual and residual), allowing shape traits to covary in differences between individuals or between repeated measures of the same individual. We specified a Sattherwaite correction for denominator degrees of freedom in this multivariate model.

Morphology–performance relationships were estimated through regression analyses. Repeatability of these relationships can be determined by estimation of the within- versus among-individual variation in slope, for example, using random regression analyses. However, such an approach requires more than two measures of the slope per individual (Martin et al., 2011). Here, we address the question of consistency in the morphology–performance relationship at a population level, determining whether the population average estimate in the first assay was correlated with the population average estimate from the second assay. These analyses were implemented separately in males and females because
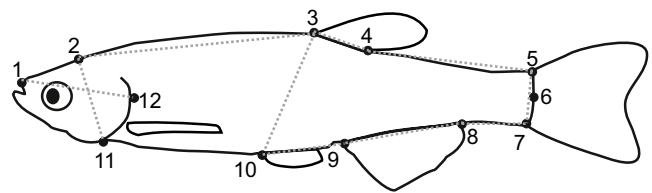

Fig. 1. Schematic illustrating landmark positions (12 numbered dots) and inter-landmark distances (10 dotted grey lines) used to characterise body shape. Landmarks were: anterior tip of snout at upper jaw (1), nape (2), dorsal fin origin (3), dorsal fin insertion (4), dorsal insertion of caudal fin (5), median caudal fin insertion (6), ventral insertion of caudal fin (7), the insertion of the 3rd most posterior ray of the anal fin (8), anal fin origin (9), pelvic fin origin (10), ventral posterior point of the operculum (11) and the posterior most point of the operculum (12). Standard length was taken as the inter-landmark distance from the anterior tip of snout at upper jaw (1) to the median caudal insertion (6).

we have previously determined that the morphology–performance relationship is sexually dimorphic in zebrafish (Conradsen and McGuigan, 2015).

For each of the four partitions of the data (males versus females in first versus second assay), we used the package 'leaps' in RStudio (v.0.98.953) to regress morphology (body size and shape) on U_{crit} for all models containing P=1 to 11 traits, and used Mallow's C_p criterion to identify the 40 best-fit models for each P. We identified the subset of these retained models with AIC within 2.0 of the best-fit model, and used the R package 'MuMIn' to calculate model averaged regression coefficients from this subset of models. Averages were taken over the full set of retained models, including any zeroes for excluded predictors. We compared the morphology–performance relationships between datasets using these model-averaged vectors of standardised regression coefficients, β. The vector correlation ($r=\beta_{rep1}^T\beta_{rep2}$, where T indicates vector transpose) was calculated between the normalised vectors (unit length: $\beta^T\beta=1$). We applied the model-averaged intercept and trait regression coefficients to predict swimming speed for each individual and compared predicted and observed values to determine the model-averaged R^2.

Finally, we asked whether the between-measures change in performance could be predicted by the changes in morphology. We first calculated the rate of change for each trait for each individual as:

$$\text{Rate of change} = \left(\frac{\text{Trait value}_{rep2} - \text{Trait value}_{rep1}}{\text{Age}_{rep2} - \text{Age}_{rep1}}\right). \quad (3)$$

We re-assessed the presence of outliers following this transformation, excluding a further 13 individuals with Mahalanobis distance above the critical threshold, resulting in 201 males and 212 females in this growth rate dataset. Again, the *a priori* identified outliers had very little effect on the regression results, with the same conclusions reached whether the outliers were or were not included; here, we present only the analyses where the outliers were excluded. For males and females separately, we regressed the morphology change traits on the change in U_{crit} and followed the same model selection approach as outlined above to identify the best-fit models, and calculate model-averaged parameters.

RESULTS
Repeatability and change in mean critical swimming speed
Critical swimming speed was significantly repeatable (Fig. 2), with a Spearman's ρ of 0.636 (d.f.=434, $P<0.001$) and an intraclass
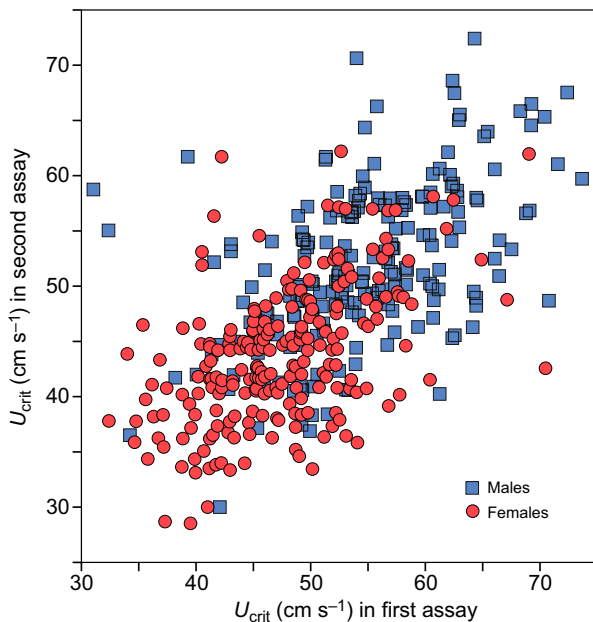
Fig. 2. The relationship between an individual's swimming speed in the first (*x*-axis) and second (*y*-axis) assay of swimming speed. Data from males (blue squares, *N*=213) and females (red circles, *N*=222) are plotted separately.

correlation of *r*=0.565 (model 1 in Table 1; 95% CI: 0.501–0.629). Model fit revealed that the identified predictors (sex, age and swimming trial) did affect swimming speed (see below), and the support for sexually heterogeneous variance components suggests that known (swimming trial) variables affected male and female performances differently (decreasing AIC with increasing model complexity in Table 1).

As expected, mean differences in performance between the sexes resulted in estimates of repeatability decreasing when the effect of sex was considered (e.g. model 2 in Table 1, *r*=0.48). In our population, U_{crit} changed with age (we address this in detail below) and variation among individuals in age when they were assayed therefore increased the within-individual component of variance; including age in the analysis increased the estimate of repeatability (e.g. model 3 in Table 1, *r*=0.65). Including swimming trial improved model fit (model 6 in Table 1), and allowing a sex-specific effect of swimming trial indicated that swimming trial contributed variance in male speed, but had little affect on females (model 7 in Table 1, among trial variance 10.7 and 2.6 in males and females respectively). Fitting sex-specific among-individual variance (model 8), and therefore sex-specific U_{crit} repeatability, did not improve model fit over a model with sex-specific among-trial variance (model 7) (ΔAIC=0.2). Thus, males and females appear to differ in their response to a component of measurement error (associated with swimming trial), but otherwise did not differ in performance repeatability.

Fitting both sex (compare AIC of models 2 and 5 with model 1 in Table 1) and age (compare AIC of models 3 and 5 with model 1 in Table 1) in the analysis of swimming speed improved model fit, indicating statistical support for affects of age and sex on U_{crit} (Fig. 3). However, there was no statistical support for a difference between the sexes in the way that age affected swimming speed (compare model 4 with model 5 in Table 1: ΔAIC=0.4). On average, males swam ~7 cm s⁻¹ (12% relative to population mean) faster than females, and swimming speed declined by ~0.040 cm s⁻¹ (0.07%) per day, resulting in an average swimming speed in the second assay ~3.7 cm s⁻¹ (7.5%) slower than in the first assay (Fig. 3).

Fish were swum in the same swim chamber, following the same protocol, implemented by the same researchers in both assays. Nonetheless, lower mean speed in the second assays could reflect systematic changes in measurement, or habituation due to prior exposure to the flume. In the initial U_{crit} assay, variation among individuals in age at which they were swum significantly predicted variation in U_{crit} (regression analysis of data from the first assay, centred on mean for each sex: $F_{1,433}$=26.34, P<0.001, r^2=0.057, standardised β=−0.239). Fish were not, on average, older in swimming trials conducted later (correlation between age and swim date within first assay data: Spearman's ρ=0.083, P=0.072, d.f.=463). This suggests that the observed decrease in mean performance between the two assays reflects a change in swimming capacity, rather than long-lasting habituation or changes in experimental protocol.

Repeatability of morphology–performance relationships
Regression analyses indicated a statistically significant relationship between body shape and swimming performance in each of the four partitions of the data (for all best-fit models, the probability of

Table 1. Intraclass correlation coefficient, *r*, for U_{crit} estimated under eight ANOVA models including different fixed and random effects

Model	Fixed	Random	Group	N	AIC	r	s.e.	ΔLL	Individual	Trial
1	–	Ind	–	2	5928.6	0.565	0.033	166.8	64.25	–
2	Sex	Ind	–	2	5792.0	0.447	0.038	96.9	50.62	–
3	Age	Ind	–	2	5813.1	0.649	0.028	237.0	60.96	–
4	Sex|Age	Ind	–	2	5675.3	0.547	0.034	154.0	47.19	–
5	Sex Age	Ind	–	2	5674.9	0.546	0.034	153.7	47.19	–
6	Sex Age	Ind;Tr	–	3	5654.2	0.621	0.033	175.1	43.38	4.71
7	Sex Age	Ind;Tr	M	4	5645.2	0.645	0.035	175.8	41.21	10.72
			F	–	–	–	–	–	–	2.62
8	Sex Age	Ind;Tr	M	6	5645.0	0.721	0.044	98.9	43.20	11.97
			F	–	–	0.577	0.051	79.8	39.48	1.72

Fixed indicates the fixed effects included in the model, where | denotes a full factorial model of main effect and interactions; Random indicates the random effects included in the model (Ind, individual; Tr, swimming trial); Group indicates whether a random effect was fitted with sex-specific estimates; N, the number of parameters in model; AIC, Akaike information criterion from maximum-likelihood (ML) fit of each model, where the smallest value indicates the model that best fits the data; *r*, intraclass correlation coefficient, with its corresponding standard error (s.e.); ΔLL, change in log-likelihood ratio (from the REML fit) between the reported model and a null model in which *r* was constrained to be zero, a test statistic that follows a χ-square distribution, with 1 degree of freedom; Individual, the model estimate of variance among individuals within the population; Trial, the model estimate of variance among swimming trials, nested within replicate measure.
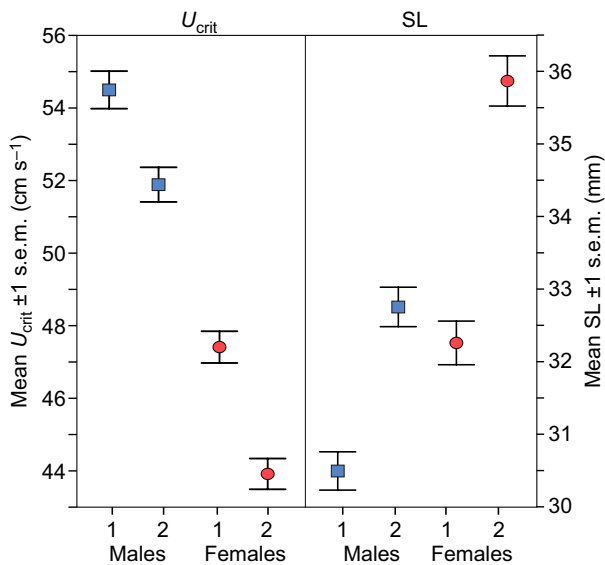
Fig. 3. Repeated measures of U_{crit} and size (standard length, SL) in each sex. Left panel, mean±s.e.m. U_{crit} in males (left, N=213) and females (right, N=222) in replicates 1 and 2. Right panel, mean±s.e.m. SL in males (left, N=230) and females (right, N=231) in replicates 1 and 2.

Table 2. The full model-averaged standardised regression coefficients (β) from the set of regression models within 2 AIC of the best fit model, for the regression of morphology on U_{crit} in males and females within each replicate, and for the regression of the rate of morphological change on the rate of change in U_{crit} in males and females

Trait	Males		Females		Growth	
	Assay 1	Assay 2	Assay 1	Assay 2	Males	Females
ILD 1–12	−0.143	−0.039	0.076	−0.040	0.004	*0.035*
ILD 2–11	−0.087	**−0.385**	−0.098	**−0.333**	−0.007	0.006
ILD 2–3	0.000	**−0.292**	−0.053	−0.034	0.020	0.002
ILD 3–4	0.070	**−0.238**	−0.052	0.001	0.001	0.000
ILD 3–10	0.001	0.016	−0.021	−0.057	0.064	−0.159
ILD 4–5	−0.015	**−0.391**	0.057	0.080	−0.011	0.000
ILD 5–7	**0.288**	**0.231**	**0.199**	**0.189**	0.089	−0.069
ILD 7–8	*−0.119*	−0.012	−0.044	**−0.257**	−0.070	**−0.246**
ILD 8–9	0.022	0.000	0.104	0.014	0.004	0.000
ILD 9–10	−0.006	*−0.091*	0.020	0.086	−0.012	**0.217**
SL	0.006	−0.001	0.003	0.018	0.000	0.113

For the regression of morphology on U_{crit} in males, assay 1 and 2, and in females, assay 1 and 2, the number of models underlying the average estimates presented were: 26, 10, 36 and 50 respectively. For the regression of change in morphology on change in U_{crit} in males and females, 37 and 14 models were averaged, respectively. Entries shown in bold were significant (P<0.01), and also correspond to the traits that were included in ≥99% of best-fit models. Entries in italics were included in most (>80%) of models for the data partition, but were not statistically significant (P>0.05). Traits are defined as inter-landmark distances (ILDs) as shown in Fig. 1. SL, standard length.

observing a larger F-value by chance was <0.01; Table S1). Repeatability of the morphology–performance relationship was statistically supported by the observation that the best-fit model from one time point significantly predicted swimming speed when applied to the data from the other time point (males, assay 1 best-fit model applied to assay 2 data: $F_{5,201}$=10.35, P<0.0001; males, assay 2 best-fit model applied to assay 1 data: $F_{6,200}$=3.93, P=0.0010; females, assay 1 best-fit model applied to assay 2 data: $F_{6,212}$=4.64, P=0.0002; females, assay 2 best-fit model applied to assay 1 data: $F_{7,211}$=2.93, P=0.0061). Repeatability was particularly apparent for one trait, which was included in all best-fit models in both males and females at both measurement times, and contributed to performance variation in the same way in each data partition. Fish with deeper caudal peduncles (ILD 5–7) swam faster (Table 2; Table S1).

Although repeatability was statistically supported, differences in morphology–performance maps were also apparent between the two assays. When the best-fit model from one measurement period was applied to the data from the other measurement period, a change in AIC of greater than 2.0 was observed for all such paired comparisons (ΔAIC=6.1, 10.2, 4.6 and 14.4 for male assay 1 model applied to assay 2 data, male assay 2 model applied to assay 1 data, female assay 1 model applied to assay 2 data, female assay 2 model applied to assay 1 data, respectively). This indicates that the best-fit model from one measurement period was not included in the best-fit set of models from the other measurement period.

Comparison of morphology–performance relationship between replicate measures of males suggested relatively low repeatability (vector correlation between normalised vectors in Table 2 was 0.38). Relatively low repeatability reflected differences in whether or not a trait contributed to performance variation, rather than opposing contributions of the same trait at different times (Table 2; Table S1). In the first assay, caudal depth was the only trait included in all models, with head length (ILD 1–12) and caudal peduncle length (ILD7–8) included in 80% of the best-fit model set (Table S1); faster males had shorter heads and shorter, deeper caudal peduncles (ILD 5–7) (Table 2). In the second assay, fast males had shallower heads (ILD 2–11), a shorter (ILD 3–4) and more anterior positioned

(ILD 2–3) dorsal fin, and a shorter (ILD 4–5), deeper (ILD 5–7) caudal peduncle (Table 2). In females, the body shape with the best performance was more similar between repeated assays (vector correlation=0.66). As with males, in the first assay, only caudal peduncle depth significantly predicted performance, whereas in the second assay, faster females had shallower heads (ILD 2–11) and deeper (ILD 5–7), shorter (ILD 7–8) caudal peduncles (Table 2).

For both males and females, body shape was more predictive of U_{crit} in the second measure. In males, body shape explained 12.0% of the variation in swimming speed in the first measure and 25.2% in the repeated measure (Table S1). Similarly, in females, shape explained 3.7% and 17.3% of variation in U_{crit} in the first and second measures, respectively (Table S1). The partial correlation coefficients of ILD 5–7, the only trait consistently implicated in all data partitions, were relatively similar across analyses. This suggests that it is the increasing contribution from other traits (e.g. head length and caudal length; Table 2) that accounts for this difference between assays.

Finally, we considered whether the observed decrease in U_{crit} between repeated measures could be attributed to a concomitant change in morphology. First, we determined that fish grew larger and changed shape between repeated measures. Growth rate was sex-specific (sex by age interaction for SL: $F_{1,463}$=33.36, P<0.001), with females growing to a greater extent than males (Fig. 3). On average, females grew 0.040 mm day^{-1}, with an increase of 12.8% in standard length between replicate measures, while males grew 0.026 mm day^{-1} or 8.8%. Body shape also changed with age in a sex-specific manner (sex by age interaction: $F_{10,496}$=17.29, P<0.0001; Fig. 4). Inspection of each shape trait suggested that, while males and females differed in the direction of change for some traits, much of the difference between sexes was in the magnitude of change, with females changing shape more than males (Fig. 4).

We used regression analyses to determine whether the observed changes in morphology predicted the observed changes in swimming speed. In males, change in size and shape significantly
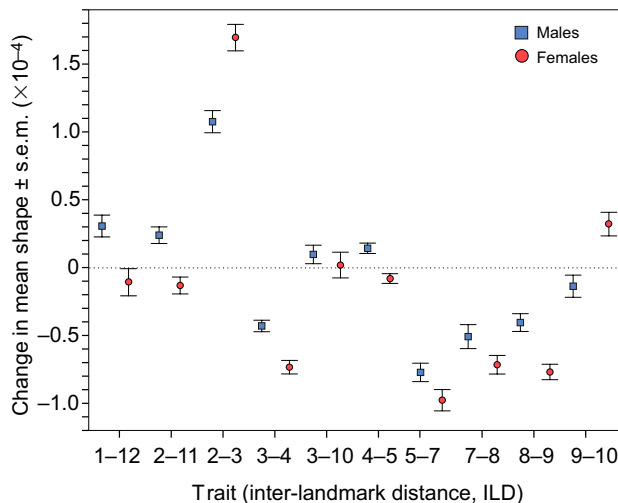
Fig. 4. Growth rate (mean±s.e.) for each of the 10 shape traits in males and females. Relative change in body measurement (in units of size) determined by the difference between repeated measures divided by the number of days between measures. Traits above dotted zero line increased (relative to size), and traits below zero decreased (relative to size) between measures. Traits defined as in Fig. 1. Mahalanobis distance between mean body shape in each replicate was 1.71 and 3.04 in males and females, respectively. Mahalanobis distance is the difference between group means scaled by the inverse of the covariance matrix (Mahalanobis, 1936; De Maesschalck et al., 2000).

($P<0.05$) predicted change in U_{crit} in the two best-fit models, but not in the 35 other models within 2 AIC of the best-fit model ($P>0.005$: Table S2). Furthermore, change in morphology explained <2.6% of the change in U_{crit}, suggesting that size and shape changes with age were relatively unimportant determinants of the observed change in swimming speed between repeated measures. In females, morphological growth did statistically significantly predict the change in performance ($P<0.001$ in all 14 models within 2 AIC of best-fit model; Table S2). Morphological change predicted 8% of the change in U_{crit}, with relative shortening of ventral caudal peduncle length (ILD 7–8) and lengthening of the distance between ventral and pelvic fins (ILD 9–10) contributing most to this model prediction (Table 2).

DISCUSSION

Over a time interval of approximately 3 months, critical swimming speed in WIK strain zebrafish was significantly repeatable, with an intraclass correlation coefficient of 0.57 (0.50–0.63). This result is consistent with several published reports of significant repeatability of U_{crit} across time periods from several days up to 6 months, in a range of fish taxa (Kolok, 1992b; Kolok and Farrell, 1994; Reidy et al., 2000; Nelson et al., 2003, 2008; Claireaux et al., 2007; Oufiero and Garland, 2009). With a sample size an order of magnitude larger than previous studies, our estimate of repeatability is more precise (Wolak et al., 2012), giving greater confidence in a relatively high upper limit on the heritability of swimming performance. The estimate of repeatability reported here sits at the lower end of the range of repeatability reported previously (>0.5; see Oufiero and Garland, 2009). Repeatability estimated from individuals reared under controlled environmental conditions, such as in this study, might be expected to be higher than for individuals experiencing natural micro-environmental variation in the wild (e.g. Auer et al., 2016). Further work, with larger sample sizes, is required to determine whether the 95% CI of repeatability of U_{crit} estimated in the wild are typically above 0.5.

The effect of sex on repeatability of U_{crit}

Although several studies have considered the effects of intra-sex variation on fish locomotion, for example, investigating the effects of sexual signalling in males (Basolo and Alcaraz, 2003; Karino et al., 2006; Baumgartner et al., 2011; Trappett et al., 2013; Oufiero et al., 2014) or reproduction (pregnancy or gravidity) in females (Plaut, 2002; Ghalambor et al., 2004; Svendsen et al., 2009; Belk and Tuckfield, 2010), surprisingly little attention has been paid to inter-sex differences of locomotor performance (Conradsen and McGuigan, 2015). Sexual dimorphism could have an impact on conclusions about repeatability in two ways. First, differences between the sexes in average performance could inflate among individual variance, upwardly biasing estimates of repeatability. Male WIK zebrafish are, on average, faster than females and consideration of this inter-sex difference in mean performance reduced repeatability to 21% below a naïve estimate (i.e. from a one-way ANOVA estimate). Some studies of the repeatability of critical swimming performance have considered only one sex (e.g. Oufiero and Garland, 2009), but other studies do not report the sex. Repeatability of performance across an experimental group (regardless of, for example, sex or population) validates the methodology, but the evolutionary relevance of such estimates needs careful consideration (Visscher et al., 2008; Wilson, 2008). Although males are consistently faster than females, positive directional selection on speed is not expected to drive the evolution of a male-only population. On the other hand, selection should sort individuals based on their phenotype, regardless of the cause of their phenotypic value (Wilson, 2008). There is no clear consensus on whether known sources of variance (such as sex) should be accounted for in estimates of repeatability (or heritability) (Visscher et al., 2008; Wilson, 2008). However, understanding how different variables contribute to repeatability estimates is crucial to their biological interpretation.

Second, repeatability might differ between the sexes. Here, we anticipated that repeatability would be lower in females because of variation in their reproductive status (i.e. how gravid they were) between repeated measures. However, there was no statistical support for a difference in repeatability of U_{crit} between sexes. A greater proportion of the performance variation among males than among females could be attributed to the effect of swimming trial (~20% of the variation among males in U_{crit} compared with ~5% in females: Table 1). Swimming trial captured both temporal variation (due, for instance, to diurnal performance rhythms) and variation generated by social interactions of specific groups of fish. It is likely that the greater among-trial variation in males (where all trials consisted of equal numbers of males and females) reflects sexual differences in social interaction, rather than in diurnal rhythms. Zebrafish are a highly social fish, with some evidence of greater aggressive interaction among males than among females (Filby et al., 2010; Hutter et al., 2010; Paull et al., 2010). Imposing two alternative artificial social contexts on fish in this experiment could have inflated within-individual variance, thus reducing repeatability relative to that estimated for fish swum individually. However, individually assayed fish might also exhibit greater variability between measures because of short-term physiological and behavioural effects of stress, or an absence of motivation. Mixed-model analyses could be further used to explore the effects on repeatability of group versus solo swimming in zebrafish.

Age affects repeatability of U_{crit} and morphology–performance relationships

Variation in age among zebrafish in this study significantly affected U_{crit} repeatability, with inclusion of age in the mixed model

increasing the estimate of repeatability by 15% over a naïve estimate. All fish in the current experiment appeared sexually mature, and were big enough (>25 mm SL) and old enough (>111 dpf) to be sexually mature (Maack and Segner, 2003; Parichy et al., 2009; Lawrence et al., 2012a,b). Therefore, although there was variation in the age at which phenotype was assessed, all fish were considered to be at the same, sexually mature, adult stage. In experiments of wild-caught animals, age is not typically precisely known. Our observation of an effect of variation in age on repeatability estimates suggests that greater consideration needs to be given to the effect of age differences among individuals.

On average, U_{crit} decreased between measures by ∼6% of the population mean initial speed. We anticipated that female performance might decline as a result of increased breeding activity. Locomotor speed is typically lower in reproductively active (pregnant or gravid) females (Plaut, 2002; Ghalambor et al., 2004; Iraeta et al., 2010) and in zebrafish, body shape associated with gravidity is negatively correlated with U_{crit} (Conradsen and McGuigan, 2015). However, males and females declined in performance to the same (statistically indistinguishable) extent, an observation at odds with reproductive investment in females as the sole cause.

Decreased performance with age is expected because of senescence of the neuromusculature system (Delbono, 2003) and has been demonstrated in guppies (Reznick et al., 2004; Oufiero and Garland, 2007). However, the maximum age of fish assayed here was ∼10 months, considerably younger than the average (42 months) lifespan reported for wild-type zebrafish (Gerhard et al., 2002). We also observed a negative age–speed relationship within the first assay period, where fish ranged from ∼4 to ∼7 months. No such change was observed within the second assay, where fish ranged in age from ∼8 to ∼10 months. Therefore, it is unlikely that the observed decline in performance was due to senescence. Rather, it seems likely that, despite the age and apparent sexual maturity of the fish in this experiment, we captured variation due to ontogenetic changes associated with the juvenile to adult transition.

Across a wide range of taxa and locomotor systems, juveniles have higher relative (but lower absolute) performance than adults (Herrel and Gibb, 2006). Here, absolute performance declined with age. Although allometric changes in morphology might underlie the relative differences in juvenile and adult performances (Herrel and Gibb, 2006; Brecko et al., 2008), behavioural compensation has also been implicated (Irschick, 2000; Herrel and Gibb, 2006). The average change in body shape between ∼6 and ∼9 month old fish did account for some of the variation in U_{crit} between measures, but relatively little (∼8% in females and <3% in males), suggesting that performance changes were driven by unknown ontogenetic changes in behaviour or physiology, rather than growth (size) itself.

Repeatability of morphology–performance relationships

The relationship between morphology and performance was significantly repeatable across the ∼3 months between assays. In particular, caudal peduncle depth had a similar (and statistically significant) relationship with critical swimming speed in both males and females at both time points. Although significantly repeatable, the morphology–performance map did diverge between repeated measures. Intriguingly, morphology was more strongly predictive of swimming performance in the second measure in our cohort of zebrafish. This observation suggests that physiological and behavioural traits contribute more to U_{crit} variation among young fish. Few studies have investigated the repeatability of morphology–performance relationships. Both Kolok (1992a) and Irschick et al.

(2006) found seasonal variation in how morphology predicted performance. A more nuanced understanding of these relationships and their variation will be necessary to improve our understanding of the function, development, genetic basis and evolution of these interacting, hierarchical phenotypes (Houle et al., 2010).

Using mixed models to estimate r

We emphasise the flexibility of mixed models for estimating phenotype repeatability and for investigating the effects on repeatability of variables affecting the trait of interest. We particularly note that covariance structures, such as the autoregressive covariance function fit here, would provide a powerful approach for determining if (how) repeatability changes with increasing lag time between assays. There is some evidence that repeatability of performances is negatively correlated with the lag between measures – closer measures are more strongly correlated than more distant measures (Oufiero and Garland, 2009; White et al., 2013). If such decays in repeatability are a common pattern, is it caused by environmental variance or does it reflect ontogenetic (senescent) physiological changes? Such questions can be addressed within a mixed-model framework, where covariance structures allowing decreasing correlation between more temporally distant measures can be fitted (Littell et al., 1996), while at the same time accounting for a mean effect of age (Biro and Stamps, 2015). Comparisons among covariance structure models (e.g. between a compound symmetry structure assuming constant repeatability and an autoregressive structure assuming repeatability decreases with increasing time between measures) constitute a powerful and parameter-efficient approach to statistical testing in comparison to applying many pairwise tests.

Conclusions

In conclusion, our data show that critical swimming speed in zebrafish is repeatable, to a similar extent in both males and females, and the population mean relationship between morphology and U_{crit} was also repeatable, although the data suggest a change in this relationship with ontogeny. The zebrafish population used in this study has been maintained through captive breeding for many generations and was reared under controlled laboratory conditions. This population history is expected to reduce repeatability by reducing the among-individual variation. Further work in other taxa is required to determine the generality of conclusions about the effects of variables such as age and sex on repeatability. Greater consideration of fine-scale population structure and the ontogeny of performance has the potential to improve understanding of the causes and consequences of among-individual variation in locomotor performance.

Data availability
Tracking data are available from the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.b91d1

Supplementary information
Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.141259.supplemental

References

Adolph, S. C. and Pickering, T. (2008). Estimating maximum performance: effects of intraindividual variation. *J. Exp. Biol.* **211**, 1336-1343.

Arnold, S. J. (1983). Morphology, performance and fitness. *Am. Zool.* **23**, 347-361.

Arnold, S. J. (2003). Performance surfaces and adaptive landscapes. *Integr. Comp. Biol.* **43**, 367-375.

Astley, H. C., Abbott, E. M., Azizi, E., Marsh, R. L. and Roberts, T. J. (2013). Chasing maximal performance: a cautionary tale from the celebrated jumping frogs of Calaveras County. *J. Exp. Biol.* **216**, 3947-3953.

Auer, S. K., Bassar, R. D., Salin, K. and Metcalfe, N. B. (2016). Repeatability of metabolic rate is lower for animals living under field versus laboratory conditions. *J. Exp. Biol.* **219**, 631-634.

Bahrndorff, S., Kjaersgaard, A., Pertoldi, C., Loeschcke, V., Schou, T. M. and Skovgård, H. and Hald, B. (2012). The effects of sex-ratio and density on locomotor activity in the house fly, *Musca domestica*. *J. Insect Sci.* **12**, 1-12.

Basolo, A. L. and Alcaraz, G. (2003). The turn of the sword: length increases male swimming costs in swordtails. *Proc. R. Soc. B Biol. Sci.* **270**, 1631-1636.

Baumgartner, A., Coleman, S. and Swanson, B. (2011). The cost of the sword: Escape performance in male swordtails. *PLoS ONE* **6**, e15837.

Beamish, F. W. H. (1978). Swimming capacity. In *Fish Physiology*, Vol. VII (ed. W. S. Hoar and D. J. Randall), pp. 101-187. New York: Academic Press.

Belk, M. C. and Tuckfield, R. C. (2010). Changing costs of reproduction: age-based differences in reproductive allocation and escape performance in a livebearing fish. *Oikos* **119**, 163-169.

Belke, T. W. and Garland, T., Jr. (2007). A brief opportunity to run does not function as a reinforcer for mice selected for high daily wheel-running rates. *J. Exp. Anal. Behav.* **88**, 199-213.

Bennett, A. F. and Huey, R. B. (1990). Studying the evolution of physiological performance. *Oxf. Surv. Evol. Biol.* **7**, 251-284.

Best, J., Adatto, I., Cockington, J., James, A. and Lawrence, C. (2010). A novel method for rearing first-feeding larval zebrafish: polyculture with Type L saltwater rotifers (*Brachionus plicatilis*). *Zebrafish* **7**, 289-295.

Biro, P. A. and Stamps, J. A. (2015). Using repeatability to study physiological and behavioural traits: ignore time-related change at your peril. *Anim. Behav.* **105**, 223-230.

Boake, C. R. B. (1989). Repeatability: Its role in evolutionary studies of mating behavior. *Evol. Ecol.* **3**, 173-182.

Bonett, D. G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Stat. Med.* **21**, 1331-1335.

Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567-585.

Brecko, J., Huyghe, K., Vanhooydonck, B., Herrel, A., Grbac, I. and Van Damme, R. (2008). Functional and ecological relevance of intraspecific variation in body size and shape in the lizard *Podarcis melisellensis* (Lacertidae). *Biol. J. Linn. Soc.* **94**, 251-264.

Brett, J. R. (1964). The respiratory metabolism and swimming performance of young sockeye salmon. *J. Fish. Res. Board Can* **21**, 1183-1226.

Burgerhout, E., Tudorache, C., Brittijn, S. A., Palstra, A. P., Dirks, R. P. and van den Thillart, G. E. E. J. M. (2013). Schooling reduces energy consumption in swimming male European eels, *Anguilla anguilla* L. *J. Exp. Mar. Biol. Ecol.* **448**, 66-71.

Careau, V. and Garland, T., Jr. (2012). Performance, personality, and energetics: Correlation, causation, and mechanism. *Physiol. Biochem. Zool.* **85**, 543-571.

Carrier, D. R. (1996). Ontogenetic limits on locomotor performance. *Physiol. Zool.* **69**, 467-488.

Claireaux, G., Handelsman, C., Standen, E. and Nelson, J. A. (2007). Thermal and temporal stability of swimming performance in the European sea bass. *Physiol. Biochem. Zool.* **80**, 186-196.

Conradsen, C. and McGuigan, K. (2015). Sexually dimorphic morphology and swimming performance relationships in wild-type zebrafish *Danio rerio*. *J. Fish Biol.* **87**, 1219-1233.

da Silva, J. M., Herrel, A., Measey, G. J., Vanhooydonck, B. and Tolley, K. A. (2014). Linking microhabitat structure, morphology and locomotor performance traits in a recent radiation of dwarf chameleons. *Funct. Ecol.* **28**, 702-713.

Dalziel, A. C., Vines, T. H. and Schulte, P. M. (2012). Reductions in prolonged swimming capacity following freshwater colonization in multiple threespine stickleback populations. *Evolution* **66**, 1226-1239.

Davidson, G. W., Sheehan, M. K. and Davie, P. S. (1999). The effect of tagging on the swimming performance of rainbow trout as a surrogate for kahawai *Arripis trutta*. *Trans. Am. Fish. Soc.* **128**, 971-973.

Davy, C. M., Paterson, J. E. and Leifso, A. E. (2014). When righting is wrong: performance measures require rank repeatability for estimates of individual fitness. *Anim. Behav.* **93**, 15-23.

De Maesschalck, R., Jouan-Rimbaud, D. and Massart, D. L. (2000). The Mahalanobis distance. *Chemometr. Intell. Lab. Syst.* **50**, 1-18.

Delbono, O. (2003). Neural control of aging skeletal muscle. *Aging Cell* **2**, 21-29.

Dohm, M. R. (2002). Repeatability estimates do not always set an upper limit to heritability. *Funct. Ecol.* **16**, 273-280.

Elphick, M. J. and Shine, R. (1998). Longterm effects of incubation temperatures on the morphology and locomotor performance of hatchling lizards (*Bassiana duperreyi*, Scincidae). *Biol. J. Linn. Soc.* **63**, 429-447.

Falconer, D. S. (1981). *Introduction to Quantitative Genetics*. London; New York: Longmans Green.

Filby, A. L., Paull, G. C., Bartlett, E. J., Van Look, K. J. W. and Tyler, C. R. (2010). Physiological and health consequences of social status in zebrafish (*Danio rerio*). *Physiol. Behav.* **101**, 576-587.

Fisher, R., Bellwood, D. R. and Job, S. D. (2000). Development of swimming abilities in reef fish larvae. *Mar. Ecol. Prog. Ser.* **202**, 163-173.

Franssen, N. R., Harris, J., Clark, S. R., Schaefer, J. F. and Stewart, L. K. (2013). Shared and unique morphological responses of stream fishes to anthropogenic habitat alteration. *Proc. R. Soc. B Biol. Sci.* **280**, 20122715.

Fu, S.-J., Cao, Z.-D., Yan, G.-J., Fu, C. and Pang, X. (2013). Integrating environmental variation, predation pressure, phenotypic plasticity and locomotor performance. *Oecologia* **173**, 343-354.

Garland, T. and Carter, P. A. (1994). Evolutionary physiology. *Annu. Rev. Physiol.* **56**, 579-621.

Gerhard, G. S., Kauffman, E. J., Wang, X. J., Stewart, R., Moore, J. L., Kasales, C. J., Demidenko, E. and Cheng, K. C. (2002). Life spans and senescent phenotypes in two strains of zebrafish (*Danio rerio*). *Exp. Gerontol.* **37**, 1055-1068.

Ghalambor, C. K., Reznick, D. N. and Walker, J. A. (2004). Constraints on adaptive evolution: The functional trade-off between reproduction and fast-start swimming performance in the Trinidadian guppy (*Poecilia reticulata*). *Am. Nat.* **164**, 38-50.

Gibb, A. C., Swanson, B. O., Wesp, H., Landels, C. and Liu, C. (2006). Development of the escape response in teleost fishes: Do ontogenetic changes enable improved performance? *Physiol. Biochem. Zool* **79**, 7-19.

Haas, T. C., Blum, M. J. and Heins, D. C. (2010). Morphological responses of a stream fish to water impoundment. *Biol. Lett.* **6**, 803-806.

Haas, T. C., Heins, D. C. and Blum, M. J. (2015). Predictors of body shape among populations of a stream fish (*Cyprinella venusta*, Cypriniformes: Cyprinidae). *Biol. J. Linn. Soc.* **115**, 842-858.

Hammer, C. (1995). Fatigue and exercise tests with fish. *Comp. Biochem. Physiol. A Physiol.* **112**, 1-20.

Hayes, J. P. and Jenkins, S. H. (1997). Individual variation in mammals. *J. Mammal.* **78**, 274-293.

Herrel, A. and Gibb, A. C. (2006). Ontogeny of performance in vertebrates. *Physiol. Biochem. Zool.* **79**, 1-6.

Herrel, A., Gonwouo, L. N., Fokam, E. B., Ngundu, W. I. and Bonneaud, C. (2012). Intersexual differences in body shape and locomotor performance in the aquatic frog, *Xenopus tropicalis*. *J. Zool.* **287**, 311-316.

Hohn, C. and Petrie-Hanson, L. (2013). Evaluation of visible implant elastomer tags in zebrafish (*Danio rerio*). *Biol. Open* **2**, 1397-1401.

Houle, D., Govindaraju, D. R. and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* **11**, 855-866.

Huey, R. B. and Dunham, A. E. (1987). Repeatability of locomotor performance in natural populations of the lizard *Sceloporus merriami*. *Evolution* **41**, 1116-1120.

Husak, J. F. and Fox, S. F. (2008). Sexual selection on locomotor performance. *Evol. Ecol. Res.* **10**, 213-228.

Hutter, S., Penn, D. J., Magee, S. and Zala, S. M. (2010). Reproductive behaviour of wild zebrafish (*Danio rerio*) in large tanks. *Behaviour* **147**, 641-660.

Iraeta, P., Salvador, A., Monasterio, C. and Díaz, J. A. (2010). Effects of gravidity on the locomotor performance and escape behaviour of two lizard populations: the importance of habitat structure. *Behaviour* **147**, 133-150.

Irschick, D. J. (2000). Effects of behaviour and ontogeny on the locomotor performance of a West Indian lizard, *Anolis lineatopus*. *Funct. Ecol.* **14**, 438-444.

Irschick, D. J. and Garland, T. (2001). Integrating function and ecology in studies of adaptation: Investigations of locomotor capacity as a model system. *Annu. Rev. Ecol. Syst.* **32**, 367-396.

Irschick, D. J. and Losos, J. B. (1998). A comparative analysis of the ecological significance of maximal locomotor performance in Caribbean Anolis lizards. *Evolution* **52**, 219-226.

Irschick, D. J., Ramos, M., Buckley, C., Elstrott, J., Carlisle, E., Lailvaux, S. P., Bloch, N., Herrel, A. and Vanhooydonck, B. (2006). Are morphology-performance relationships invariant across different seasons? A test with the green anole lizard (*Anolis carolinensis*). *Oikos* **114**, 49-59.

Irschick, D. J., Meyers, J. J., Husak, J. F. and Le Galliard, J. F. (2008). How does selection operate on whole-organism functional performance capacities? A review and synthesis. *Evol. Ecol. Res.* **10**, 177-196.

Jayne, B. C. and Bennett, A. F. (1990). Selection on locomotor performance capacity in a natural population of garter snakes. *Evolution* **44**, 1204-1229.

Karino, K., Orita, K. and Sato, A. (2006). Long tails affect swimming performance and habitat choice in the male guppy. *Zool. Sci.* 23, 255-260.

Katija, K., Colin, S. P., Costello, J. H. and Jiang, H. S. (2015). Ontogenetic propulsive transitions by *Sarsia tubulosa* medusae. *J. Exp. Biol.* 218, 2333-2343.

Killen, S. S., Marras, S., Steffensen, J. F. and McKenzie, D. J. (2012). Aerobic capacity influences the spatial position of individuals within fish schools. *Proc. R. Soc. B Biol. Sci.* 279, 357-364.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Mol. Ecol. Resour.* 11, 353-357.

Kolok, A. S. (1992a). Morphological and physiological correlates with swimming performance in juvenile largemouth bass. *Am. J. Physiol.* 263, R1042-R1048.

Kolok, A. S. (1992b). The swimming performance of individual largemouth bass (*Micropterus salmoides*) are repeatable. *J. Exp. Biol.* 170, 265-270.

Kolok, A. S. (1999). Interindividual variation in the prolonged locomotor performance of ectothermic vertebrates: a comparison of fish and herpetofaunal methodologies and a brief review of the recent fish literature. *Can. J. Fish. Aquat. Sci.* 56, 700-710.

Kolok, A. S. and Farrell, A. P. (1994). Individual variation in the swimming performance and cardiac performance of northern squawfish, *Ptychocheilus oregonensis*. *Physiol. Zool.* 67, 706-722.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33, 402-416.

Lande, R. and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* 37, 1210-1226.

Langerhans, R. B. (2008). Predictability of phenotypic differentiation across flow regimes in fishes. *Integr. Comp. Biol.* 48, 750-768.

Langerhans, R. B. and DeWitt, T. J. (2004). Shared and unique features of evolutionary diversification. *Am. Nat.* 164, 335-349.

Langerhans, R. B., Knouft, J. H. and Losos, J. B. (2006). Shared and unique features of diversification in greater Antillean *Anolis* ecomorphs. *Evolution* 60, 362-369.

Lawrence, C., Adatto, I., Best, J., James, A. and Maloney, K. (2012a). Generation time of zebrafish (*Danio rerio*) and medakas (*Oryzias latipes*) housed in the same aquaculture facility. *Lab Animal* 41, 158-165.

Lawrence, C., Best, J., James, A. and Maloney, K. (2012b). The effects of feeding frequency on growth and reproduction in zebrafish (*Danio rerio*). *Aquaculture* 368-369, 103-108.

Leblanc, C. A. and Noakes, D. L. (2012). Visible implant elastomer (VIE) tags for marking small rainbow trout. *N. Am. J. Fish. Manag.* 32, 716-719.

Lessells, C. M. and Boag, P. T. (1987). Unrepeatable repeatabilities: a common mistake. *Auk* 104, 116-121.

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996). *SAS System for Mixed Models*. Cary, NC, USA: SAS Institute, Inc.

Losos, J. B., Creer, D. A. and Schulte, J. A.II (2002). Cautionary comments on the measurement of maximum locomotor capabilities. *J. Zool.* 258, 57-61.

Lynch, M. and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates Inc.

Maack, G. and Segner, H. (2003). Morphological development of the gonads in zebrafish. *J. Fish Biol.* 62, 895-906.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proc. Natl. Inst. Sci. India* 12, 49-55.

Martin, J. G. A., Nussey, D. H., Wilson, A. J. and Réale, D. (2011). Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Methods Ecol. Evol.* 2, 362-374.

McGuigan, K., Franklin, C. E., Moritz, C. and Blows, M. W. (2003). Adaptation of rainbow fish to lake and stream habitats. *Evolution* 57, 104-118.

Metcalfe, N. B. and Monaghan, P. (2001). Compensation for a bad start: Grow now, pay later? *Trends Ecol. Evol.* 16, 254-260.

Miller, N. Y. and Gerlai, R. (2011). Shoaling in zebrafish: what we don't know. *Rev. Neurosci.* 22, 17-25.

Moen, D. S., Irschick, D. J. and Wiens, J. J. (2013). Evolutionary conservatism and convergence both lead to striking similarity in ecology, morphology and performance across continents in frogs. *Proc. R. Soc. B Biol. Sci.* 280, 20132156.

Nelson, J. A., Gotwalt, P. S. and Snodgrass, J. W. (2003). Swimming performance of blacknose dace (*Rhinichthys atratulus*) mirrors home-stream current velocity. *Can. J. Fish. Aquat. Sci.* 60, 301-308.

Nelson, J. A., Gotwalt, P. S., Simonetti, C. A. and Snodgrass, J. W. (2008). Environmental correlates, plasticity, and repeatability of differences in performance among blacknose dace (*Rhinichthys atratulus*) populations across a gradient of urbanization. *Physiol. Biochem. Zool.* 81, 25-42.

Nelson, J. A., Atzori, F. and Gastrich, K. R. (2015). Repeatability and phenotypic plasticity of fish swimming performance across a gradient of urbanization. *Environ. Biol. Fish.* 98, 1431-1447.

Osborne, J. W. and Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Pract. Assess. Res. Eval.* 9, 1-12.

Oufiero, C. E. and Garland, T. (2007). Evaluating performance costs of sexually selected traits. *Funct. Ecol.* 21, 676-689.

Oufiero, C. E. and Garland, T. (2009). Repeatability and correlation of swimming performances and size over varying time-scales in the guppy (*Poecilia reticulata*). *Funct. Ecol.* 23, 969-978.

Oufiero, C. E., Meredith, R. W., Jugo, K. N., Tran, P., Chappell, M. A., Springer, M. S., Reznick, D. N. and Garland, T. (2014). The evolution of the sexually selected sword in *Xiphophorus* does not compromise aerobic locomotor performance. *Evolution* 68, 1806-1823.

Parichy, D. M., Elizondo, M. R., Mills, M. G., Gordon, T. N. and Engeszer, R. E. (2009). Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev. Dyn.* 238, 2975-3015.

Paull, G. C., Filby, A. L., Giddins, H. G., Coe, T. S., Hamilton, P. B. and Tyler, C. R. (2010). Dominance hierarchies in zebrafish (*Danio rerio*) and their relationship with reproductive success. *Zebrafish* 7, 109-117.

Plaut, I. (2000). Effects of fin size on swimming performance, swimming behaviour and routine activity of zebrafish *Danio rerio*. *J. Exp. Biol.* 203, 813-820.

Plaut, I. (2001). Critical swimming speed: its ecological relevance. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 131, 41-50.

Plaut, I. (2002). Does pregnancy affect swimming performance of female Mosquitofish, *Gambusia affinis*? *Funct. Ecol.* 16, 290-295.

Price, G. (1970). Selection and covariance. *Nature* 227, 520-521.

Pritchard, V. L., Lawrence, J., Butlin, R. K. and Krause, J. (2001). Shoal choice in zebrafish, *Danio rerio*: the influence of shoal size and activity. *Anim. Behav.* 62, 1085-1088.

Rauch, G.-J., Granato, M. and Haffter, P. (1997). A polymorphic zebrafish line for genetic mapping using SSLPs on high-percentage agarose gels. *Tech. Tips Online* 2, 148-150.

Reidy, S. P., Kerr, S. R. and Nelson, J. A. (2000). Aerobic and anaerobic swimming performance of individual Atlantic cod. *J. Exp. Biol.* 203, 347-357.

Reznick, D. N., Bryant, M. J., Roff, D., Ghalambor, C. K. and Ghalambor, D. E. (2004). Effect of extrinsic mortality on the evolution of senescence in guppies. *Nature* 431, 1095-1099.

Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. *Anim. Prod.* 8, 95-108.

Roche, D. G., Binning, S. A., Bosiger, Y., Johansen, J. L. and Rummer, J. L. (2013). Finding the best estimates of metabolic rates in a coral reef fish. *J. Exp. Biol.* 216, 2103-2110.

Rohlf, F. J. (1996). Morphometric spaces, shape components and the effects of linear transformations. In *Advances in Morphometrics*, Vol. 284 (ed. L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor and D. E. Slice), pp. 117-129. New York: Plenum Press.

Rohlf, F. J. (2005). *tpsDig2. Stony Brook*. New York: Department of Ecology and Evolution, State University of New York.

Schmidt-Nielsen, K. (1972). Locomotion: energy cost of swimming, flying, and running. *Science* 177, 222-228.

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

Shaw, R. G. (1987). Maximum-likelihood approaches applied to quantitative genetics of natural populations. *Evolution* 41, 812-826.

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychol. Bull.* 95, 334-344.

Sutphin, Z. A., Myrick, C. A. and Brandt, M. M. (2007). Swimming performance of sacramento splittail injected with subcutaneous marking agents. *N. Am. J. Fish. Manage.* 27, 1378-1382.

Svendsen, J. C., Aarestrup, K., Dolby, J., Svendsen, T. C. and Christensen, R. H. B. (2009). The volitional travel speed varies with reproductive state in mature female brown trout *Salmo trutta*. *J. Fish Biol.* 75, 901-907.

Taylor, E. B. and McPhail, J. D. (1985). Burst swimming and size-related predation of newly emerged coho salmon *Oncorhynchus kisutch*. *Trans. Am. Fish. Soc.* 114, 546-551.

Trappett, A., Condon, C. H., White, C., Matthews, P. and Wilson, R. S. (2013). Extravagant ornaments of male threadfin rainbowfish (*Iriatherina werneri*) are not costly for swimming. *Funct. Ecol.* 27, 1034-1041.

van Vliet, M. T. H., Franssen, W. H. P., Yearsley, J. R., Ludwig, F., Haddeland, I., Lettenmaier, D. P. and Kabat, P. (2013). Global river discharge and water temperature under climate change. *Glob. Environ. Change* 23, 450-464.

Visscher, P. M., Hill, W. G. and Wray, N. R. (2008). Heritability in the genomics era — concepts and misconceptions. *Nat. Rev. Genet.* 9, 255-266.

Wainwright, P. C. (2007). Functional versus morphological diversity in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 38, 381-401.

Walker, J. A. (2007). A general model of functional constraints on phenotypic evolution. *Am. Nat.* 170, 681-689.

White, C. R., Schimpf, N. G. and Cassey, P. (2013). The repeatability of metabolic rate declines with time. *J. Exp. Biol.* 216, 1763-1765.

Wilson, A. J. (2008). Why h(2) does not always equal VA/VP? *J. Evol. Biol.* 21, 647-650.

Wolak, M. E., Fairbairn, D. J. and Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods Ecol. Evol.* 3, 129-137.

Wolter, C. and Arlinghaus, R. (2003). Navigation impacts on freshwater fish assemblages: the ecological relevance of swimming performance. *Rev. Fish Biol. Fish.* 13, 63-89.

Wright, D., Rimmer, L. B., Pritchard, V. L., Krause, J. and Butlin, R. K. (2003). Inter and intra-population variation in shoaling and boldness in the zebrafish (*Danio rerio*). *Naturwissenschaften* 90, 374-377.