Canalisation in the wild

Boonekamp, Jelle J.; Mulder, Ellis; Verhulst, Simon

LETTER

Canalisation in the wild: effects of developmental conditions on physiological traits are inversely linked to their association with fitness

Jelle J. Boonekamp,* Ellis Mulder and Simon Verhulst*

Groningen Institute for Evolutionary Life Sciences, University of Groningen, P.O.Box 11103, 9700CC Groningen, the Netherlands

*Correspondence: E-mails: jjboonekamp@gmail.com; s.verhulst@rug.nl

Abstract
Ecological conditions affect fitness, but mechanisms causing such effects are not well known, while evolved responses to environmental variation may depend on the underlying mechanisms. Consequences of environmental conditions vary strongly between traits, but a framework to interpret such variation is lacking. We propose that variation in trait response may be explained by differential canalisation, with traits with larger fitness effects showing weaker responses to environmental perturbations due to preferential resource allocation to such traits. We tested the canalisation hypothesis using brood size manipulation in wild jackdaw nestlings in which we measured eight physiological traits (mainly oxidative stress markers), and two feather traits. For each trait, we estimated manipulation response and association with fitness (over-winter survival). As predicted, a strong negative correlation emerged between manipulation response and association with fitness ($r = -0.76$). We discuss the consequences of differential trait canalisation for the study of mechanisms mediating environmental effects on fitness.

Keywords
Developmental stress, growth, life history, oxidative stress, phenotypic plasticity, somatic damage.

Ecology Letters (2018) **21**: 857–864

INTRODUCTION

The mechanistic underpinnings of key processes in evolutionary ecology, such as the fitness consequences of developmental conditions, are not well known despite their importance for understanding how ecology affects evolution (McNamara & Houston 2009). The general approach to the study of such mechanisms is to test responses of (physiological) traits to experimental or observational variation in the environment, growth or reproductive effort. However, this approach generally finds heterogeneous responses across traits, even when they are indicative of the same physiological mechanisms such as, for example, the immune system (Horrocks *et al.* 2012; Hegemann *et al.* 2013) or oxidative stress (Cohen *et al.* 2009; Smith *et al.* 2016). This heterogeneity is not well understood, perhaps because predictive frameworks to interpret such variation are not well developed (Cohen *et al.* 2012). This impedes attempts to integrate the heterogeneous effects on different physiological traits in composite traits using, for example, principal component analysis (e.g. (Cohen & McGraw 2009; Fowler & Williams 2017) or Pythagorean-based metrics like the Mahalanobis distance (Cohen *et al.* 2014; Milot *et al.* 2014). To contribute to the development of such a framework, we here test the hypothesis that variation in physiological responses to variation in offspring environment can be explained by the canalisation hypothesis, which states that traits should be more robust against environmental variation when they are more important for fitness (Waddington 1942, 1959). This hypothesis was originally formulated with an emphasis on robustness with respect to genetic variation, but is equally applicable to robustness with respect to environmental variation (Waddington 1942; Stearns *et al.* 1995;

Meiklejohn & Hartl 2002; Flatt 2005; Boonekamp *et al.* 2017; Vedder *et al.* 2017). It is in the latter way that we consider the hypothesis here. We assume that the fitness loss due to deviations from the optimal trait value varies between traits, being stronger in some traits than others. Deviations from the optimal trait values are induced by internal or environmental perturbations, but the magnitude of the induced deviation will depend on the allocation priorities of the organism. These priorities will be higher with increasing fitness costs of deviations from the optimal trait value. Consequentially, the effect of an environmental perturbation will be stronger on traits where deviations from the optimal trait value are less costly in fitness terms, that is, they are less well canalised (Fig. 1). In conceptual terms, a lower level of canalisation can be equated with a narrower homeostatic plateau in the framework developed by Nijhout *et al.* (2017). In this study, we do not consider the underlying mechanisms of differential allocation, because they are likely to be highly complex (e.g. Cheng *et al.* 2011) and beyond the scope of this study. Instead we use the predictive framework of the canalisation hypothesis (Fig. 1) in the attempt to explain variation in trait responses to environmental effects.

Oxidative stress (OS), that is, the damage resulting from an imbalance between reactive oxygen species production vs. the capacity to buffer such damage, has often been implicated as a candidate mechanism mediating long-term developmental effects and the costs of reproduction (Metcalfe & Monaghan 2013). However, this hypothesis has received mixed support, as illustrated, for example, by a recent meta-analysis of the association between growth and oxidative stress that revealed substantial heterogeneity in the direction of the observed relationships (Smith *et al.* 2016). A suggested explanation for this
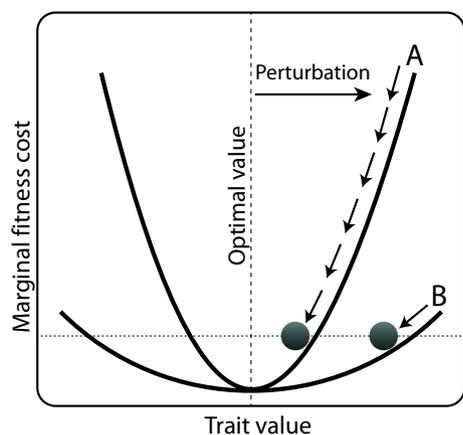
Figure 1 Marginal fitness cost as a function of deviation from the optimal trait value. The marginal fitness cost is the net cost considering the fitness effect of the trait value and the resource investment required to reach or maintain the trait value. We consider two traits, (a) and (b), that differ in the extent to which a deviation from the optimal trait value induces a fitness cost. Given a perturbation, for example, reduced resource availability, investment in trait (a) will be higher than in trait (b), to the extent that the marginal fitness loss of the perturbation is equal for the two traits (indicated by the two markers). The resulting difference between the realised and optimal trait values will be larger for trait B, that is, trait A is more canalised, which is the hypothesis we test in this paper.

heterogeneity is that most studies used a low number of OS markers, and correlations among OS markers are often highly variable (Monaghan *et al.* 2009; Cohen *et al.* 2010; Selman *et al.* 2012). The heterogeneous relationships among OS markers may be due to different OS markers providing different information about physiological state (Romero-Haro & Alonso-Alvarez 2014). As such, they are likely to be differentially maintained in the face of challenging conditions, in particular, when there is variation in fitness returns of investment in the different traits. For example, if some OS variables are more important for fitness than others, they are likely to be better canalised, increasing their robustness to variation in environmental conditions or growth (Fig. 1). Unfortunately, the fitness consequences of variation in different OS markers is not well known (Speakman & Selman 2011; Speakman *et al.* 2015), impeding testing functional explanations of OS variation. Moreover, information on fitness consequences is required to evaluate the extent to which OS traits mechanistically underpin effects of developmental conditions on fitness prospects.

First, to investigate the role of OS in mediating effects of developmental conditions on fitness prospects, we manipulated brood size in free-living jackdaws (*Coloeus monedula*), and studied its effect on a panel of physiological traits, including six markers of oxidative stress, triglyceride and protein level, all measured in blood. As argued elsewhere, an experimental approach is critical to reveal effects of developmental conditions on physiological state (Metcalfe & Monaghan 2013; Speakman *et al.* 2015). We determined three commonly used markers of oxidative damage (Thiobarbituric acid reactive substances – TBARS; hydroperoxides – dROMs; and oxidised glutathione – GSSG) and antioxidant protection (uric acid;

reduced glutathione – GSH; and glutathione redox state – GSSG relative to total glutathione). We subsequently studied the association between these traits and first winter survival of fledglings, a major fitness component in birds (van de Pol *et al.* 2006). We previously showed in this population that nestlings in enlarged broods have impaired development: they grow less well, have more feather fault bars and accelerated telomere attrition (Boonekamp *et al.* 2014, 2017). In the light of these previous findings, we predicted that nestlings in enlarged broods would have increased oxidative damage and/or decreased antioxidant protection [effects on antioxidants and oxidative damage are not necessarily in parallel, (Costantini & Verhulst 2009)], and reduced triglyceride and protein level. This pattern would support the hypothesis that OS physiology underlies the association between the number of offspring and their fitness prospects. Second, to investigate whether the canalisation hypothesis explains the heterogeneity of physiological responses to environmental variation, we tested the prediction that trait responses negatively co-vary with their associations to fitness (Fig. 1). In support of the canalisation hypothesis, we found that physiological traits that were significantly related to survival were not affected by the brood size manipulation and vice-versa. We previously reported associations between developmental conditions and feather fault bars and fitness and here we show that these morphological traits also fit the canalisation pattern of physiological traits.

METHODS

Study system

We studied a natural population of jackdaws in 2014 consisting of five nest box colonies in the vicinity of Groningen (53.1708 °N, 6.6064 °E, the Netherlands). General field procedures were as previously described (Boonekamp *et al.* 2017). We visited nest boxes once every 3 days to monitor the initiation of nest building, egg laying and hatching. Once eggs started hatching, we conducted daily nest visits, in order to determine the exact hatching dates of nestlings. Hatchlings were weighed and individually marked by clipping the nail tips, which additionally facilitated the collection of small blood droplets (5 μL) to be used for molecular sex determination (for details see (Salomons *et al.* 2008). Nestlings were counted, measured and weighed on days 5, 10, 20 and 30 (hatching of the oldest chick = day 1). We collected a blood sample on day 20 (≤ 1 mL) from the brachial vein for oxidative stress measurements. Blood samples were stored at ± 4 °C during field work, centrifuged within 2–3 h after sampling, and plasma was stored in aliquots at −80 °C. Nestlings were ringed prior to fledging (day 30) with a numbered metal ring, an RFID-tagged ring (Radio-frequency identification), and a colour ring with engraved markings. We monitored post-fledging survival in the study area by their colour rings in the periods February – May in 2015 and 2016 (for details see (Boonekamp *et al.* 2017). Survival is challenging to estimate, because the disappearance of individuals reflects a combination of mortality and dispersal. We previously established that the dispersal rate between breeding colonies is low and independent of brood size manipulation and fledging location

and hence we assume that our estimates of state-dependent survival are not biased by state-dependent dispersal (Boonekamp *et al.* 2014, 2017).

Brood size manipulations

Brood size manipulations were carried out as previously described (Boonekamp *et al.* 2014). In brief, age-matched broods were manipulated with net ± 2 nestlings. In most cases, manipulated broods contained both resident and cross-fostered offspring after manipulation, which was achieved by first transferring three nestlings from reduced to enlarged broods after which we repositioned one nestling of the original offspring from the enlarged to the reduced brood. In case broods contained insufficient number of nestlings we either moved 2 (22%) or 1 (16%) nestling(s) from reduced to enlarged broods, in which cases no nestlings were returned. We randomly assigned which nestlings were to be relocated using a randomizer smartphone app ('Dice').

Oxidative stress measurements and correlated blood variables

We measured TBARS, reactive oxygen metabolites (dROMs), uric acid and triglyceride and protein concentrations in blood plasma, and reduced and oxidised glutathione concentrations in whole blood of day 20 nestlings using commercially available kits and factory provided protocols. Laboratory assays are described in detail in the online supporting information (Supporting Information 1).

Statistics

To analyse the effect of brood size manipulation on our panel of physiological variables, we fitted linear mixed effects models using the lme4 package (Bates *et al.* 2015) in *R* with restricted maximum log-likelihood estimation. Models of these physiological variables included birth nest ID as random effect to take into account the dependence of related nestlings and to allow estimation of the birth nest variance component relative to the residual variance. Furthermore, models of the effect of brood size manipulation on OS markers included two 'background' variables (plasma protein and triglyceride level) because these variables are known to be potential substrates for ROS (reactive oxygen species) and hence can be expected to interact with any variable of oxidative stress (Pérez-Rodríguez *et al.* 2015). We tested the fixed effects of brood size manipulation (enlarged = 1; reduced = 0), sex (male = 1; female = 0) and their interaction on each physiological variable independently. We subsequently omitted the sex and interaction terms after having confirmed that they did not significantly affect any of the oxidative stress variables (Table S1). Visual inspection confirmed a normal distribution of the residuals of all models.

We performed mixed effects logistic regression models to analyse the association between survival and each physiological marker individually. We included fledging colony ID as random effect because survival and/or observation probabilities may differ between fledging locations. We did not include birth nest ID as random effect because too few numbers of

surviving nestlings originated from the same birth nest limiting model convergence. We included sex and the interaction between sex and the focal physiological marker, but since neither sex ($P > 0.45$) nor any interaction with sex were significant ($P > 0.13$) in any of these models, final models only included the oxidative stress variable as fixed effect.

We subsequently tested for a canalisation pattern by assessing the negative co-variation among manipulation and survival effect sizes using the nonparametric Spearman rank correlation test. To obtain comparable effect sizes for the brood size manipulation and survival, we computed Cohen's d effect sizes and 95% confidence intervals based on the *t*- and *z*-values of the mixed effects models (Nakagawa & Cuthill 2007). We used absolute effect sizes to test the canalisation pattern (i.e. negative Cohen's d values were made positive) because for canalisation only the strength of environmental effects and fitness associations are important, regardless of their direction.

RESULTS

In total 235 nestlings of 79 manipulated broods were included in this study. Pre-manipulation brood sizes were similar on average (enlarged – reduced = -0.14 ± 0.10; $P = 0.16$). After the manipulation, the enlarged broods contained significantly more nestlings than reduced broods (enlarged – reduced = 3.36 ± 0.08; $P < 0.001$), and nestlings growing up in enlarged broods had significantly reduced body mass at day 20 when we took blood samples for physiological analyses (enlarged – reduced = -16.14 ± 3.13 grams; $P < 0.001$). Thus, as evidence by the effect on growth, the brood size manipulation successfully modified developmental conditions.

Brood size manipulation

Brood size manipulation had heterogeneous effects on the six OS markers (Fig. 2a). GSH and GSSG were very similar in nestlings in reduced and enlarged broods (Table 1; Fig. 2a), and so were glutathione redox state (Table 1; Fig. 2a) and plasma uric acid level (Table 1; Fig. 2a). Thus, brood size manipulation did not significantly affect any of the antioxidant markers. In contrast, brood size manipulation significantly affected two markers of oxidative damage, but in opposite direction. Nestlings in enlarged broods had significantly higher dROMs, but lower TBARS levels (Table 1; Fig. 2a). These two OS markers were not significantly correlated ($r = -0.033$; $P = 0.64$). The 95% confidence intervals of the Cohen's d values of the manipulation effect did not overlap between dROMs and TBARS (Fig. 2a), indicating a significant interaction between brood size manipulation and the type of oxidative damage marker. We found no significant effects of the manipulation on plasma protein and triglyceride level (Table 1), and there were no effects of sex, or interactions between sex and brood size manipulation with respect to any of the physiological markers (Table S1). In conclusion, brood size manipulation had heterogeneous effects on the six oxidative stress variables that we measured, including opposite effects on two oxidative damage markers.
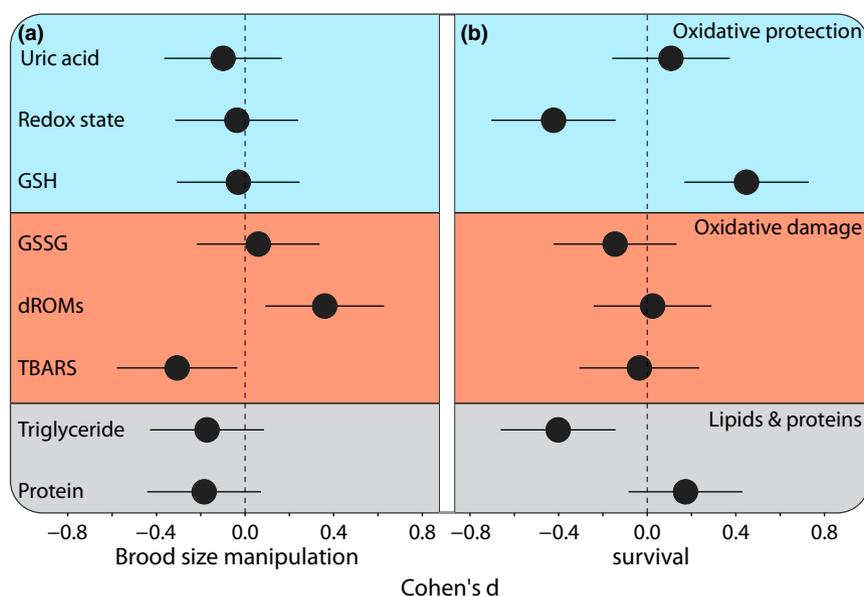
Figure 2 Oxidative stress in relation to brood size manipulation (a) and post-fledging survival prospects (b). Data points reflect Cohen's d estimates of Tables 1 & 2 and error bars denote the 95% confidence intervals. Manipulation effect sizes reflect the enlarged – reduced broods' value, that is, a positive difference reflects a higher value in enlarged broods. In panel B, a positive value indicates higher survival prospects of individuals with higher trait values.

Table 1 Brood size manipulation effects on six markers of oxidative stress and plasma protein and triglyceride level

Variable	n	Birth nest/residual σ^2	Manipulation (SE)	t	P	Cohen's d ($\pm$ 95% CI)
uric acid	220	< 0.001/0.621	−0.068 (0.093)	−0.730	0.466	−0.099 (−0.365; 0.166)
redox state	202	0.574/0.660	−0.027 (0.118)	−0.229	0.819	−0.037 (−0.314; 0.240)
GSH	202	0.582/0.687	−0.023 (0.123)	−0.189	0.851	−0.031 (−0.308; 0.247)
GSSG	202	0.520/0.673	0.046 (0.120)	0.380	0.704	0.060 (−0.217; 0.337)
dROMs	219	0.204/0.618	0.262 (0.104)	2.516	0.013*	0.360 (0.091; 0.628)
TBARS	212	0.058/0.961	−0.326 (0.151)	−2.159	0.032*	−0.307 (−0.579; −0.035)
Triglyceride	235	0.511/0.455	−0.135 (0.112)	−1.202	0.231	−0.172 (−0.429; 0.085)
Protein	235	0.605/0.324	−0.122 (0.096)	−1.266	0.207	−0.185 (−0.442; 0.073)

Sample size (n) denotes the number of nestlings in the analyses (note that it was not always possible to acquire sufficiently large blood samples to carry out all six oxidative stress assays). The random effect 'birth nest' reflects the variance partitioned into the birth nest factor relative to the residual variance. 'Manipulation' reflects the average difference (and standard error) of nestlings in the enlarged minus reduced broods. Cohen's d values (and the 95% confidence intervals) were determined using the t-test statistic of the manipulation effect obtained from the mixed effects model. Significant P-values are highlighted with an asterisk (*). Blue and red colours indicate markers of oxidative protection and damage, respectively.

Survival

We identified 77 out of 235 fledglings as survivors within the follow-up period (*c.* 33%; males: 33%; females: 34%). We found a highly significant negative association between glutathione redox state (GSSG/(GSSG + GSH/2)) and survival (Table 2; Fig. 2b), with a survival decline from 45 to 20% over the observed range of glutathione redox state (Fig. 3a). To investigate whether this relationship was primarily driven by variation in the GSH or the GSSG glutathione component, we quantified the survival association in two independent models and found a strong positive and significant association between GSH and survival, while the (negative) relationship was weak and non-significant for GSSG (Table 2; Fig. 2b). Thus, variation in the GSH component, reflecting the capacity to neutralise reactive oxygen species, was the primary component driving the observed relationship between survival and redox state.

The other markers of oxidative stress and plasma protein level were not significantly associated with survival, with Cohens'd values ranging between −0.15 and 0.17 (Table 2; Fig 1b). However, plasma triglyceride level was significantly negatively related to survival (Table 2; Fig. 3b). Nestlings that had lower triglyceride level survived substantially better (45%) than nestlings with high triglyceride level (20%) and this effect was independent of nestling mass or sex.

Testing for canalisation

When comparing physiological traits qualitatively, we find that variables that were associated with survival (Fig. 2b) were not significantly affected by the brood size manipulation (Fig. 2a). At the same time, physiological variables that were significantly affected by the brood size manipulation were not associated with survival (Fig. 2). These findings are in qualitative agreement with the hypothesis that physiological variables

Table 2 Post-fledging survival in relation to six markers of oxidative stress and plasma protein and triglyceride level estimated by logistic regression

Variable	n	Colony σ^2	Estimate (SE)	z	P	Cohen's d ($\pm$ 95% CI)
uric acid	220	0.149	0.172 (0.218)	0.788	0.431	0.107 (−0.158; 0.373)
redox state	202	0.237	−0.557 (0.188)	−2.967	0.003*	−0.423 (−0.703; −0.142)
GSH	202	0.347	0.930 (0.296)	3.138	0.002*	0.448 (0.168; 0.729)
GSSG	202	0.162	−0.158 (0.155)	−1.015	0.310	−0.145 (−0.423; 0.133)
dROMs	219	0.099	0.028 (0.155)	0.180	0.857	0.025 (−0.242; 0.291)
TBARS	212	0.152	−0.038 (0.148)	−0.255	0.799	−0.035 (−0.306; 0.235)
Triglyceride	235	< 0.001	−0.453 (0.148)	−3.049	0.002*	−0.402 (−0.661; −0.143)
Protein	235	0.203	0.203 (0.155)	1.312	0.189	0.173 (−0.084; 0.430)

Sample size (n) denotes the number of measured fledglings. 'Colony σ^2' reflects the variance partitioned into the random factor of fledging colony (note that the residual variance is 1 in logistic regression). 'Estimate' reflects the estimated association with survival (and standard error) of the blood variable. Cohen's d values (and confidence intervals) were determined using the z-test statistic obtained from the logistic regression model. Asterisks (*) indicate significant P-values. Blue and red colours indicate markers of oxidative protection and damage, respectively.
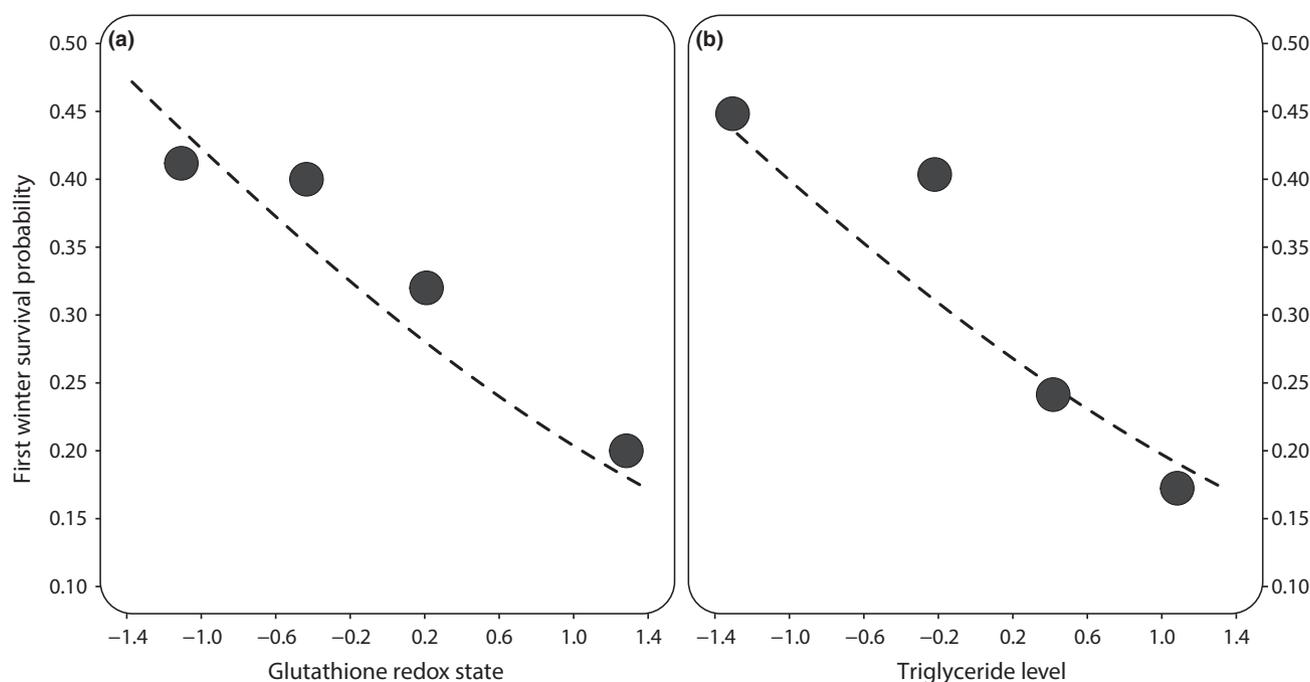
Figure 3 First winter survival in relation to glutathione redox state (a) and triglyceride level (b). Dashed lines reflect the logistic regression lines from the models presented in Table 2. Data points reflect the mean values of the data grouped by quartile bins (n = 50–51 per data point).

are more canalised when they are significantly related to fitness prospects.

To verify the negative co-variation between manipulation and survival effect sizes among physiological traits in a more quantitative way, we were limited by the fact that we had only four OS markers available for this purpose (including GSH and GSSG would induce pseudo-replication because they are both components of glutathione redox state) and two other physiological markers. With such a small sample size, correlations need to be very high to reach statistical significance (i.e. with n = 6, correlations are significant when r_s > 0.85). We therefore extended the data set with two variables related to the physiology of growth, the number of fault bars on tail and wing feathers, to which the canalisation hypothesis equally applies. We recently reported associations with brood size manipulation and survival for the number of fault bars in tail and wing feathers of jackdaw nestlings in the same cohort (Boonekamp *et al.* 2017) and here we

included these two variables in our test of the canalisation hypothesis. Among this set of eight variables, there was a significant negative correlation between manipulation and survival effect sizes (Fig. 4; r_s = −0.76; n = 8; P = 0.038), in agreement with the canalisation hypothesis. This correlation did not depend on the inclusion of the two morphological variables, because when excluded, the association became slightly stronger (r_s = −0.83; n = 6; P = 0.058). We also checked whether our finding was robust when taking into account the uncertainty of the effect sizes, that is, their standard deviations, because such uncertainty could potentially lead to a weaker association. Bayesian analysis including the standard deviations of the effect sizes confirmed our findings (Supporting Information 3). The median and mode of the posterior correlation distribution were even more negative (r_{median} = 0.80, r_{mode} = 0.89) supporting the strong negative covariance among survival and manipulation effect sizes.
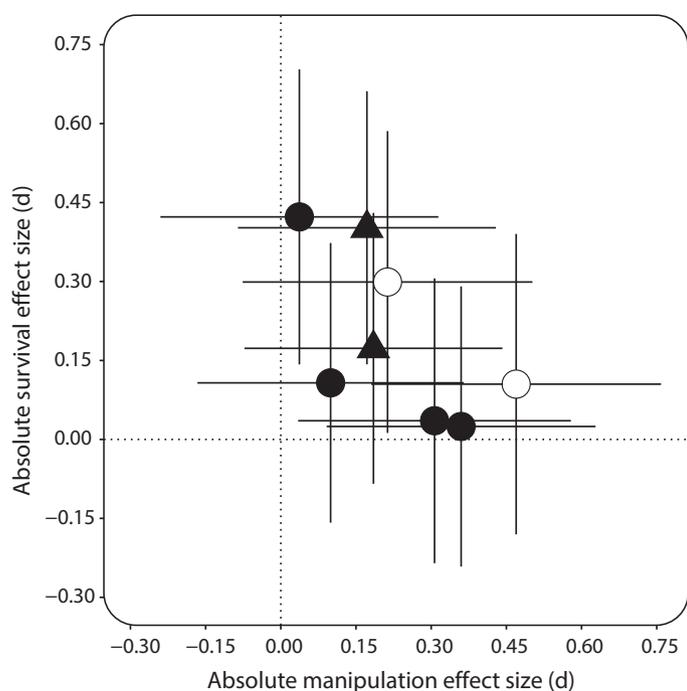
Figure 4 Canalisation of physiological and morphological variables. Data points reflect the absolute Cohen's d estimates and error bars denote the 95% confidence intervals (see Methods for details). Closed dots reflect effect sizes based on data presented in this paper, with circles representing oxidative stress markers, and triangles plasma protein and triglyceride level. Open dots are estimates of tail and wing fault bar numbers obtained from our recent publication (Boonekamp *et al.* 2017). Note that fault bar estimates were based on the same study population, including the 2014 cohort of fledglings that we used for our oxidative stress analyses.

DISCUSSION

Environmental variation has heterogeneous effects on physiological traits, and this variation is not well understood, which impedes the unravelling of mechanisms that mediate environmental effects on physiology, behaviour and fitness. We tested the hypothesis that canalisation provides a framework to understand heterogeneous responses of different traits, with an emphasis on physiological markers of OS. Below we discuss our findings in the context of the canalisation hypothesis to investigate the extent to which canalisation explains this heterogeneity.

Development

Previous studies found on average little effect of developmental conditions on markers of antioxidant protection (Smith *et al.* 2016, but see (Bourgeon *et al.* 2011; Cornell & Williams 2017), and our findings confirm this result (Fig. 2a). In contrast, we found significant manipulation effects on two markers of oxidative damage, TBARS and dROMs (Fig. 2a), but these effects were in opposite direction. We did not detect an effect on GSSG, the level of oxidised glutathione. That TBARS and dROMs were affected in opposite directions indicates that these OS markers reflect different physiological

processes, underlining the view that single OS markers are not sufficient to reliably assay oxidative stress (Monaghan *et al.* 2009; Cohen *et al.* 2010; Selman *et al.* 2012). TBARS are derivatives of MDA (Malondialdehyde) lipid peroxidation end-products (Halliwell & Gutteridge 2015), whereas the dROMs test quantifies organic hydroperoxides (Costantini 2016). Our results therefore indicate that adverse developmental conditions reduce the rate of lipid peroxidation (TBARS) while simultaneously increasing oxidative damage revealed by organic hydroperoxides (dROMs). Previous studies reported positive correlations between plasma lipid concentration and MDA (e.g. (Romero-Haro & Alonso-Alvarez 2014; Pérez-Rodríguez *et al.* 2015), but there was no such association in our data (Table S1), and the non-significant effect of brood size on triglyceride levels was negative, opposite to what would be expected when an effect on lipid levels caused the brood size effect on TBARS. Hence, the causes of the observed opposite manipulation effects on TBARS and dROMs are not easily explained by differences in the underlying physiological or biochemical pathways. This further underlines the pressing need for a theoretical framework that allows us to unravel the causes underpinning the observed heterogeneity in physiological responses to variation in developmental conditions from a functional perspective.

Survival

We observed substantial heterogeneity among OS markers in the extent to which they predicted post-fledgling survival (Fig. 2b). High glutathione redox state was associated with low survival (Fig. 3a), and this effect was primarily caused by an association between GSH and survival, but not GSSG. None of the other OS markers were significantly related to survival (Fig. 2b). There can be different reasons explaining the absence of an association between a physiological trait and survival, in addition to the association truly being weak. For example, a physiological measurement may not be very informative, when for technical or biological reasons the repeatability of the measurement is very low. However, we consider it unlikely that such methodological factors confounded our findings, because our test of the canalisation hypothesis showed that it was, in particular, the traits that were not associated with survival that were the most sensitive to the effects of brood size manipulation (Fig. 4). A more important factor in our view may be that we did not manipulate oxidative stress directly, but instead relied on natural variation in oxidative stress markers. Such variation may not only reflect the outcome of resource allocation between the maintenance of oxidative stress variables but also many other traits (Fig. 1), impeding the development of simple hypotheses with respect to their associations to survival.

Canalisation

The canalisation hypothesis has, to the best of our knowledge, not previously been considered as explanation for heterogeneity in physiological responses to environmental perturbations. We did, however, find two other studies that investigated both environmental and fitness correlations for multiple OS

markers, and their findings are in general agreement with the canalisation hypothesis. In St Kilda sheep, lamb growth can be taken as indicator of the state of developmental conditions, and lamb growth was correlated with MDA, but not with total antioxidant capacity, activity of superoxide dismutase (an antioxidant enzyme) or protein carbonyls, while only the latter two markers were associated with first winter survival (Christensen *et al.* 2016). A study using captive zebra finches reported slightly stronger effects of an environmental manipulation on plasma 8-OHdG, a DNA repair product, while SOD showed a stronger association with survival (Marasco *et al.* 2017). Clearly, more studies are needed to assess the extent to which the canalisation hypothesis can serve as a framework to understand the heterogeneous responses to environmental perturbations, but we consider the available results encouraging. We tested the canalisation hypothesis in the context of early development, but we note that the principle equally applies to changes in old age, where also a large heterogeneity among traits is found (Hayward *et al.* 2015; Briga 2016). For example, markers of antioxidant protection appeared to be more stable with age than markers of oxidative damage (Martin & Grotewiel 2006), which is in line with our finding that markers of antioxidant protection appeared to be more robust against the influences of environmental perturbation.

We assumed that the canalisation principle would emerge from 'physiological regulatory networks' (Cohen *et al.* 2012) as a consequence of differential resource allocation with respect to the fitness returns on investment, but whether this is indeed the process that caused the observed pattern remains to be tested. Regardless of the exact mechanism giving rise to the observed pattern, the canalisation principle has general implications for the study of the mechanistic link between environmental variation and its fitness consequences, which we consider one of the major challenges for the life sciences. This is so because the canalisation principle impedes the search for biomarker traits that link the environment and its fitness consequences, simply because the most important traits with respect to fitness will be least affected by the environment. The canalisation principle also raises questions with respect to the practice to integrate multiple traits into fewer composite variables using, for example, principal components analyses [e.g. (Cohen *et al.* 2015)]. Instead of utilising how traits correlate with each other, it may be more informative to integrate variables on the basis of their relationship with fitness when the purpose is to unravel the mechanistic link between the environmental variation or ageing and its fitness consequences.

AUTHOR CONTRIBUTIONS

JJB and SV designed the study, carried out the field work, conducted the data analyses and wrote the first draft of the manuscript. JJB and EM carried out the physiological measurements. All authors contributed to revising the manuscript and consented with the publication of its final version.

REFERENCES

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Soft.*, 67, 1–48.

Boonekamp, J.J., Salomons, M., Bouwhuis, S., Dijkstra, C. & Verhulst, S. (2014). Reproductive effort accelerates actuarial senescence in wild birds: an experimental study. *Ecol. Lett.*, 17, 599–605.

Boonekamp, J.J., Dijkstra, R., Dijkstra, C. & Verhulst, S. (2017). Canalization of development reduces the utility of traits as fitness biomarkers: feather fault bars in nestling birds. *Funct. Ecol.*, 31, 719–727.

Bourgeon, S., Guindre-Parker, S. & Williams, T.D. (2011). Effects of sibling competition on growth, oxidative stress, and humoral immunity: a two-year brood-size manipulation. *Physiol. Biochem. Zool.*, 84, 429–437.

Briga, M. (2016). Growing up and growing old. PhD Thesis, University of Groningen, the Netherlands. ISBN: 978-94-6233-475-5.

Cheng, L.Y., Bailey, A.P., Leevers, S.J., Ragan, T.J., Driscoll, P.C. & Gould, A.P. (2011). Anaplastic lymphoma kinase spares organ growth during nutrient restriction in drosophila. *Cell*, 146, 435–447.

Christensen, L.L., Selman, C., Blount, J.D., Pilkington, J.G., Watt, K.A., Pemberton, J.M., *et al.* (2016). Marker-dependent associations among oxidative stress, growth and survival during early life in a wild mammal. *Proc. R. Soc. B*, 283, 20161407–20161408.

Cohen, A.A. & McGraw, K.J. (2009). No simple measures for antioxidant status in birds: complexity in inter- and intraspecific correlations among circulating antioxidant types. *Funct. Ecol.*, 23, 310–320.

Cohen, A.A., McGraw, K.J. & Robinson, W.D. (2009). Serum antioxidant levels in wild birds vary in relation to diet, season, life history strategy, and species. *Oecologia*, 161, 673–683.

Cohen, A.A., de Magalhães, J.P. & Gohil, K. (2010). Ecological, biomedical and epidemiological approaches to understanding oxidative balance and ageing: what they can teach each other. *Funct. Ecol.*, 24, 997–1006.

Cohen, A.A., Martin, L.B., Wingfield, J.C., McWilliams, S.R. & Dunne, J.A. (2012). Physiological regulatory networks: ecological roles and evolutionary constraints. *Trends Ecol. Evol.*, 27, 428–435.

Cohen, A.A., Milot, E., Li, Q., Legault, V., Fried, L.P. & Ferrucci, L. (2014). Cross-population validation of statistical distance as a measure of physiological dysregulation during aging. *Exp. Gerontol.*, 57, 203–210.

Cohen, A.A., Milot, E., Li, Q., Bergeron, P., Poirier, R., Dusseault-Bélanger, F., *et al.* (2015). Detection of a novel, integrative aging process suggests complex physiological integration. *PLoS ONE*, 10, e0116489–26.

Cornell, A. & Williams, T.D. (2017). Variation in developmental trajectories of physiological and somatic traits in a common songbird approaching fledging. *J. Exp. Biol.* 220, 4060–4067.

Costantini, D. (2016). Oxidative stress ecology and the d-ROMs test: facts, misfacts and an appraisal of a decade's work. *Behav. Ecol. Sociobiol.*, 70, 809–820.

Costantini, D. & Verhulst, S. (2009). Does high antioxidant capacity indicate low oxidative stress? *Funct. Ecol.*, 23, 506–509.

Flatt, T. (2005). The evolutionary genetics of canalization. *Q. Rev. Biol.*, 80, 287–316.

Fowler, M.A. & Williams, T.D. (2017). A physiological signature of the cost of reproduction associated with parental care. *Am. Nat.*, 190, 762–773.

Halliwell, B. & Gutteridge, J.M.C. (2015). *Free Radicals in Biology and Medicine*. 5th ed. Oxford University Press, Oxford, England.

Hayward, A.D., Moorad, J., Regan, C.E., Berenos, C., Pilkington, J.G., Pemberton, J.M., et al. (2015). Asynchrony of senescence among phenotypic traits in a wild mammal population. *Exp. Gerontol.*, 71, 56–68.

Hegemann, A., Matson, K.D., Versteegh, M.A., Villegas, A. & Tieleman, B.I. (2013). Immune response to an endotoxin challenge involves multiple immune parameters and is consistent among the annual-cycle stages of a free-living temperate zone bird. *J. Exp. Biol.*, 216, 2573–2580.

Horrocks, N.P.C., Hegemann, A., Matson, K.D., Hine, K., Jaquier, S., Shobrak, M., et al. (2012). Immune indexes of larks from desert and temperate regions show weak associations with life history but stronger links to environmental variation in microbial abundance. *Physiol. Biochem. Zool.*, 85, 504–515.

Marasco, V., Stier, A., Boner, W., Griffiths, K., Heidinger, B. & Monaghan, P. (2017). Environmental conditions can modulate the links among oxidative stress, age, and longevity. *Mech. Ageing Dev.*, 164, 100–107.

Martin, I. & Grotewiel, M.S. (2006). Oxidative damage and age-related functional declines. *Mech. Ageing Dev.*, 127, 411–423.

McNamara, J.M. & Houston, A.I. (2009). Integrating function and mechanism. *Trends Ecol. Evol.*, 24, 670–675.

Meiklejohn, C.D. & Hartl, D.L. (2002). A single mode of canalization. *Trends Ecol. Evol.*, 17, 468–473.

Metcalfe, N.B. & Monaghan, P. (2013). Does reproduction cause oxidative stress? An open question *Trends Ecol. Evol.*, 28, 347–350.

Milot, E., Cohen, A.A., Vézina, F., Buehler, D.M., Matson, K.D. & Piersma, T. (2014). A novel integrative method for measuring body condition in ecological studies based on physiological dysregulation. *Methods Ecol. Evol.*, 5, 146–155.

Monaghan, P., Metcalfe, N.B. & Torres, R. (2009). Oxidative stress as a mediator of life history trade-offs: mechanisms, measurements and interpretation. *Ecol. Lett.*, 12, 75–92.

Nakagawa, S. & Cuthill, I.C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.*, 82, 591–605.

Nijhout, H.F., Sadre-Marandi, F., Best, J. & Reed, M.C. (2017). Systems Biology of Phenotypic Robustness and Plasticity. *Integr. Comp. Biol.*, 57, 171–184.

Pérez-Rodríguez, L., Romero-Haro, A.A., Sternalski, A., Muriel, J., Mougeot, F., Gil, D., et al. (2015). Measuring oxidative stress: the confounding effect of lipid concentration in measures of lipid peroxidation. *Physiol. Biochem. Zool.*, 88, 345–351.

van de Pol, M., Bruinzeel, L.W., Heg, D., van der Jeugd, H.P. & Verhulst, S. (2006). A silver spoon for a golden future: long-term effects of natal origin on fitness prospects of oystercatchers (Haematopus ostralegus). *J. Anim. Ecol.*, 75, 616–626.

Romero-Haro, A.A. & Alonso-Alvarez, C. (2014). Covariation in oxidative stress markers in the blood of nestling and adult birds. *Physiol. Biochem. Zool.*, 87, 353–362.

Salomons, H.M., Dijkstra, C. & Verhulst, S. (2008). Strong but variable associations between social dominance and clutch sex ratio in a colonial corvid. *Behav. Ecol.*, 19, 417–424.

Selman, C., Blount, J.D., Nussey, D.H. & Speakman, J.R. (2012). Oxidative damage, ageing, and life-history evolution: where now? Trends Ecol. Evol., 27, 570–577.

Smith, S.M., Nager, R.G. & Costantini, D. (2016). Meta-analysis indicates that oxidative stress is both a constraint on and a cost of growth. *Ecol. Evol.*, 6, 2833–2842.

Speakman, J.R. & Selman, C. (2011). The free-radical damage theory: accumulating evidence against a simple link of oxidative stress to ageing and lifespan. *BioEssays*, 33, 255–259.

Speakman, J.R., Blount, J.D., Bronikowski, A.M., Buffenstein, R., Isaksson, C., Kirkwood, T.B.L., et al. (2015). Oxidative stress and life histories: unresolved issues and current needs. *Ecol. Evol.*, 5, 5745–5757.

Stearns, S.C., Kaiser, M. & Kawecki, T.J. (1995). The differential genetic and environmental canalization of fitness components in Drosophila melanogaster. *J. Evol. Biol.*, 8, 539–557.

Vedder, O., Verhulst, S., Bauch, C. & Bouwhuis, S. (2017). Telomere attrition and growth: a life-history framework and case study in common terns. *J. Evol. Biol.*, 30, 1409–1419.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150, 563–565.

Waddington, C.H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature*, 183, 1654–1655.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.