

Genetic variation in variability: phenotypic variability of fledging weight and its evolution in a songbird population

Running title: Great tit families differ in variability

H. A. Mulder¹, P. Gienapp^{1,2}, M. E. Visser^{1,2}

1 Animal Breeding and Genomics Centre, Wageningen University & Research, PO Box 338, 6700 AH Wageningen, The Netherlands

2 Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), PO Box 50, 6700 AB Wageningen, The Netherlands

Corresponding author: H.A. Mulder

Animal Breeding and Genomics Centre, Wageningen University & Research, PO Box 338, 6700 AH Wageningen, The Netherlands

Han.mulder@wur.nl

Key words: residual variance, songbird, natural population, fitness, evolution, body mass

Figures: 4

Tables: 2

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/evo.13008](#).

This article is protected by copyright. All rights reserved.

The data used was archived in Dryad: <http://dx.doi.org/10.5061/dryad.2qv8n>.

Abstract

Variation in traits is essential for natural selection to operate and genetic and environmental effects can contribute to this phenotypic variation. From domesticated populations we know that families can differ in their level of within-family variance, which leads to the intriguing situation that within-family variance can be heritable. For offspring traits, such as birth weight, this implies that within-family variance in traits can vary among families and can thus be shaped by natural selection. Empirical evidence for this in wild populations is however lacking. We investigated whether within-family variance in fledging weight is heritable in a wild great tit (*Parus major*) population and whether these differences are associated with fitness. We found significant evidence for genetic variance in within-family variance. The genetic coefficient of variation was 0.18 and 0.25, when considering fledging weight a parental or offspring trait, respectively. We found a significant quadratic relationship between within-family variance and fitness: families with low or high within-family variance had lower fitness than families with intermediate within-family variance. Our results show that within-family variance can respond to selection and provide evidence for stabilizing selection on within-family variance.

Introduction

Phenotypes are determined by genetic and environmental components and using standard quantitative genetic models we can partition the phenotypic variance (V_P) into additive genetic variance (V_A) and residual or environmental variance (V_E) (Falconer and Mackay 1996; Lynch and Walsh 1998) in absence of dominance and epistasis. Laboratory studies

have shown that genotypes do not differ only in phenotype, but also in their environmental variation (V_E) (Waddington 1942, 1960; Rendel et al. 1966; Scharloo et al. 1972; Kaufman et al. 1977) and, more recently, that in outbred domestic populations this V_E can be partly heritable (Hill and Mulder 2010). The fact that V_E can be heritable opens up the exciting possibility that the variation itself can evolve when under selection. Such selection can be caused by, for example, non-linear fitness functions (Falconer and Mackay 1996; Dall 2010; Hamilton 2010). Interestingly, also selection on the trait itself can lead to evolution of V_E , because directional (truncation) selection will select for increased trait variation and hence could lead to increase in V_E if this variation is heritable (Hill and Zhang 2004; Mulder et al. 2007). Such an increase in trait variation could hinder ‘evolutionary rescue’ as, despite the population mean phenotype responding adaptively, more individuals will become maladapted, which would negatively impact on population mean growth rate, a key parameter in the context of evolutionary rescue (Lynch and Lande 1993; Gomulkiewicz and Holt 1995).

Although many studies found genetic variation in plasticity in natural systems (Scheiner 1993; Brommer et al. 2005; Nussey et al. 2005), currently we do not have direct information about the occurrence and magnitude of genetic variation in environmental variation (V_E) in natural systems as current examples are all from laboratory studies or domestic populations. Mackay and Lyman (2005) showed highly significant genetic variance in V_E using iso-female lines of *Drosophila*. Morgante et al. (2015) and Sorensen et al. (2015) found substantial genetic variation in V_E using inbred lines in *Drosophila*. In different livestock species the median genetic coefficient of variation in V_E was 0.3 (Hill and Mulder 2010). Other evidence comes from genome-wide association studies that found QTLs affecting phenotypic variability of traits in humans (Yang et al. 2012), *Arabidopsis* (Shen et al. 2012), *Drosophila* (Morgante et al. 2015) or livestock (Mulder et al. 2013a; Sell-Kubiak et al. 2015).

Although empirical evidence for genetic differences in V_E in laboratory studies and animal breeding is growing, the biological understanding of these differences in V_E is limited. Differences in V_E between genotypes could be due to differential buffering of disturbing external or internal environmental effects, which has been discussed as canalization (Waddington 1942), developmental (in)stability and plasticity (Gibson and Wagner 2000; Debat and David 2001; De Visser et al. 2003; Flatt 2005; Masel and Siegal 2009; Masel and Trotter 2010). Canalization consists of genetic and environmental canalization (Wagner et al. 1997), in which the latter is related to buffering internal environmental perturbations and would lead to a reduction in V_E , while the former is related to buffering effects of genetic variation, e.g. mutation (Gibson and Dworkin 2004; Gibson 2009). Developmental (in)stability is usually measured as differences in within-individual variance such as fluctuating asymmetry (Palmer and Strobeck 1986; Clarke 1998), while plasticity is the variation among genotypes in response to (quantifiable) external environmental factors such as temperature or diet and studied using reaction norms (Scheiner 1993; Lynch and Walsh 1998; Schlichting and Pigliucci 1998; Nussey et al. 2005).

The biological basis of differences in V_E can also be that within a single macro-environment unknown micro-environmental differences still exist. Therefore, each individual experiences a slightly different environment. A direct mathematical link between plasticity and differences in V_E is given by Gavrillets and Hastings (1994), using a model for differences in V_E due to a linear reaction norm to an unknown environmental factor. Genetic differences in canalization or plasticity would be reflected in differences in variance among lines, e.g. clones, inbred lines or in differences in within-family variance in outbred populations (Hill and Mulder 2010).

So far, we mostly considered traits that are only affected by direct genetic effects on the phenotype and their V_E . However, many offspring traits, such as weaning weight in mammals or fledging weight in birds, are not only affected by direct effects of the offspring's genes for body weight, but also by the genes for parental care of the parents, which determine the amount and quality of food provided to the offspring (i.e. brood environment) (Figure 1). Thus, the variation in offspring's fledging weight is due to (1) the variation in the 'individual environment' among chicks, which is due to both the overall provisioning by the parents and the allocation rules among chicks the parents use, and (2) the inherited reaction norm of the chicks in how 'individual environment' influences fledging weight. The parental genes affect both these routes: the first via genetic variation in provisioning behaviour, the second via inherited reaction norms.

Direct effects of offspring genotype on plasticity as well as parental effects on plasticity could thus contribute to heterogeneity in within-family variance. The direct genetic and indirect parental genetic effects on V_E cannot be easily partitioned unless cross-fostering is applied. However, studying the net effect of both types of genes on within-family variance offers exciting opportunities to study the evolution of V_E in natural populations.

As pointed out above, variation in V_E can lead to variation in individual fitness and V_E could hence be under selection. Stabilizing selection on the mean is expected to reduce V_E as demonstrated by Mulder et al. (2015), whereas intense directional selection is expected to increase V_E (Hill and Zhang 2004; Mulder et al. 2007). Under stabilizing selection, Zhang and Hill (2005) showed that small homogeneity costs or large temporal variation in optimal phenotype and strength of stabilizing selection can maintain V_E . Theoretical studies have focused mainly on stabilizing selection, while empirical studies provide more evidence for directional rather than stabilizing selection acting on the phenotype (Kingsolver et al. 2001;

Kingsolver et al. 2012). Furthermore, there is no conclusive pattern whether there is a positive relationship between developmental stability and fitness (Clarke 1998).

Fledging weight and its V_E in great tits is an interesting case to study the selective forces on V_E , because parents could maximize their fitness by either having few similar offspring of high quality or many offspring of more variable quality, i.e. bet-hedging strategies or adaptive coin-flipping (Kaplan and Cooper 1984; Seger and Brockmann 1987). A crucial condition for these strategies to be adaptive is a relationship between V_E of fledging weight and parental fitness caused by a non-linear relationship between the trait itself and fitness (Fig. 2A, B). If this relationship is linear, broods with high and low V_E will have equal fitness, as the fitness increase caused by more extreme offspring on one side is exactly balanced by the fitness reductions by more extreme offspring on the other side. However, if this relationship is non-linear, broods with high and low V_E do not have equal fitness, as decreasing and increasing the trait by the same value does not yield symmetrical changes in fitness, resulting in a net effect on fitness. As this relationship is sigmoidal in great tits (Tinbergen and Boerlijst 1990; Verboven and Visser 1998; Both et al. 1999; Monros et al. 2002), the relationship between V_E and fitness could be non-linear (Fig. 2) and V_E of fledging weight could be under selection.

We here use avian fledging weight as our model trait and make use of the unique opportunity to quantify genetic variance in and selection on V_E offered by the pedigreed long-term study population on wild great tits in the Hoge Veluwe (the Netherlands). The development of the double hierarchical generalized linear model (DHGLM) enables modelling of genetic and non-genetic effects on the trait level and its V_E in pedigreed populations (Rönnegård et al. 2010; Felleki et al. 2012). Studying V_E of avian fledging weight as a model trait will not only increase our understanding of parental behaviour but will

generally improve our understanding of evolutionary processes in natural populations, as e.g. the adaption to environmental disturbances.

Material and Methods

Data collection

We used data from a natural population of Great tits (*Parus major*) at our long-term (1955 – present) study population at the Hoge Veluwe (the Netherlands). In an area of 171 ha of mixed woodland, about 400 nest boxes were checked weekly for egg-laying date and clutch size. Chicks were weighed when they were 14-16 days old, because this is just prior to fledging (day 18) and chicks at this age have reached their fledging weight. All chicks were ringed with standard aluminium rings and breeding adults were captured at the nest, which allows a pedigree to be constructed. More details on the data can be found in Gienapp et al. (2006). We used only chicks from first clutches, where at least five chicks were weighed and with known father and mother. We excluded data before 1975, because of small numbers of nests per year in which chicks were weighted in those early years. In total, 17,535 phenotypic records were used and the pedigree contained 65,280 animals. Summary statistics and the distribution of fledging weight suggest that fledging weight is approximately normally distributed, but with some skewness (-1.20) and kurtosis (6.19).

Statistical model to estimate genetic variance in V_E

To estimate simultaneously genetic variance in phenotype and its V_E the Double Hierarchical Generalized Linear Model (DHGLM) was used in ASReml 2.0 (Gilmour et al. 2006;

Rönnegård et al. 2010; Felleki et al. 2012). We used here the within-brood variance (V_W) instead of V_E ; V_W contains in addition to V_E the Mendelian sampling variance and unexplained non-additive genetic variance such as dominance and epistasis. Ideally, one would use the animal model to analyse both phenotypes and V_E rather than V_W . However, the DHGLM gives biased variance components when using an animal model when there is only a single observation per animal (Mulder et al. 2013b; Sonesson et al. 2013). Therefore, we used a sire-dam model following Sonesson et al. (2013) and assessed evidence for genetic variation in V_E as variation in the variance within fullsib families (V_W). The DHGLM can be represented as:

$$\begin{bmatrix} \mathbf{y} \\ \boldsymbol{\Psi} \end{bmatrix} = \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_v \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_v \end{bmatrix} + \begin{bmatrix} \mathbf{Z} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_v \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{u}_v \end{bmatrix} + \begin{bmatrix} \mathbf{V} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_v \end{bmatrix} \begin{bmatrix} \mathbf{n} \\ \mathbf{n}_v \end{bmatrix} + \begin{bmatrix} \mathbf{e} \\ \mathbf{e}_v \end{bmatrix} \quad (1)$$

where \mathbf{y} is a vector with fledging weight, $\boldsymbol{\Psi}$ is the response variable for V_W , the matrix \mathbf{X} (\mathbf{X}_v) is the incidence matrix for fixed effects, \mathbf{Z} (\mathbf{Z}_v) is the incidence matrix for sire-dam genetic effects, \mathbf{V} (\mathbf{V}_v) is the incidence matrix for common brood effects. The vectors with solutions for fixed effects were \mathbf{b} and \mathbf{b}_v . To correct for effects of clutch size, year and within-year seasonal trend on fledging weight and its variance, year and annual 5-day period of hatching starting from 1 April were fitted as fixed effect factors and clutch size was fitted as a covariate with linear and quadratic regressions. Years and the seasonal trend within year are affecting mean fledging weight and its variance due to differences in weather and food

availability. To be able to disentangle brood effects and these periodic effects, annual 5-day periods with fewer than five nests were merged to the closest 5-day period before or after. These fixed effects were tested for fledging weight with a Wald test using ASReml and were significant ($P < 0.05$); for V_W the same fixed effects were used. The effect of observer was tested as random effect, but was not significant for fledging weight (proportion explained phenotypic variance was 0.003; likelihood ratio test (LRT) with one degree of freedom: $\chi^2 = 0.86$; $P = 0.35$) nor for V_W (estimate on log scale: $2.57 * 10^{-3}$; LRT with a mixture of 1 and 2 degrees of freedom (Stram and Lee 1994; Visscher 2006): $\chi^2 = 0.52$; $P = 0.62$) when only observers with records in three or more years were used. We restricted the analysis of the observer effect in this way to avoid confounding annual and observer effects on fledging weight and its V_W .

The response variable Ψ for V_W was calculated for one observation as $\psi_i = \log(\widehat{\sigma}_{e_i}^2) + \frac{e_i^2 / (1 - h_i) - \widehat{\sigma}_{e_i}^2}{\widehat{\sigma}_{e_i}^2}$, where e_i^2 is the squared residual for observation i , h_i is the leverage, i.e. the diagonal element of the hat-matrix of \mathbf{y} for observation i (Hoaglin and Welsh 1978), and $\widehat{\sigma}_{e_i}^2$ is the predicted residual variance for observation i based on the estimated fixed and random effects for ψ_i (Felleki et al. 2012). The leverage h_i is a measure of the uncertainty in the residual and as a consequence the variance of e_i is smaller than the true residual variance. Dividing e_i^2 by $(1 - h_i)$ scales all estimates towards the true residual variance. Because $e_i^2 / (1 - h_i)$ follows a χ^2 -distribution a log-link function needs to be used or as an alternative the first-order Taylor series approximation as derived by Felleki et al. (2012), which is

$$\psi_i = \log(\widehat{\sigma}_{e_i}^2) + \frac{e_i^2 / (1 - h_i) - \widehat{\sigma}_{e_i}^2}{\widehat{\sigma}_{e_i}^2}.$$

The vectors \mathbf{u} and \mathbf{u}_v contained the sire-dam effects for fledging weight and its V_W and were assumed bivariate normally distributed $\begin{bmatrix} \mathbf{u} \\ \mathbf{u}_v \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \frac{1}{4} \begin{bmatrix} \sigma_a^2 & cov_{aa_v} \\ symmetric & \sigma_{a_v}^2 \end{bmatrix} \otimes \mathbf{A} \right)$, where σ_a^2 ($\sigma_{a_v}^2$) is the additive genetic variance in fledging weight (V_W), cov_{aa_v} is the additive genetic covariance between fledging weight and its within-family deviation and \mathbf{A} is the additive genetic or numerator relationship matrix. Note that the sire and dam effects were assumed to have the same additive genetic variance. The vectors \mathbf{n} and \mathbf{n}_v contained the common brood or nest effects $\begin{bmatrix} \mathbf{n} \\ \mathbf{n}_v \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_n^2 & cov_{nn_v} \\ symmetric & \sigma_{n_v}^2 \end{bmatrix} \otimes \mathbf{I} \right)$, where σ_n^2 ($\sigma_{n_v}^2$) is the brood variance in fledging weight (V_W), cov_{nn_v} is the brood covariance between fledging weight and its within-family deviation and \mathbf{I} is the identity matrix. The vector with residuals \mathbf{e} and \mathbf{e}_v were assumed to be independent and normally distributed $\begin{bmatrix} \mathbf{e} \\ \mathbf{e}_v \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{W}^{-1}\sigma_\varepsilon^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_v^{-1}\sigma_{\varepsilon_v}^2 \end{bmatrix} \right)$, where $\mathbf{W} = \text{diag}(\exp(\hat{\boldsymbol{\psi}}))^{-1}$ and $\mathbf{W}_v = \text{diag}(\frac{1-h}{2})$, and σ_ε^2 and $\sigma_{\varepsilon_v}^2$ are scaling variances, which are expected to be unity, because \mathbf{W} and \mathbf{W}_v contain the reciprocals of the residual variances. Estimation of σ_ε^2 and $\sigma_{\varepsilon_v}^2$ allows for departures from the expectation of unity. The algorithm was run for 100 runs in which the matrices \mathbf{W} and \mathbf{W}_v and the vector $\boldsymbol{\psi}$ were updated after each full ASReml-run. Estimates of variance components oscillated, but the sum of the squared deviations between estimates in the current and previous ASReml-run relative to the current estimates were smaller than $2.1 * 10^{-2}$ in the last 50 runs, indicating only small oscillations relative to the parameter estimates. Therefore, we presented the results of the last run. More details on the algorithm are in Felleki et al. (2012), and Mulder et al. (2013b).

Model comparison

Model 1 is the full model with fixed and random effects for fledging weights and its V_W . In addition, we analysed fledging weight and its V_W with a model containing either only fixed effects for V_W (Model 2) or a model with fixed effects and a random brood effect for V_W (Model 3). For model comparison, we used a LRT and Akaike's Information Criterion (AIC) in combination with the adjusted profile h-likelihood (see Mulder et al. (2013b) for details). LRT was assumed to follow a mixture of χ^2 -test 1 and 2 degrees of freedom (Stram and Lee 1994; Visscher 2006).

Quantitative genetic parameters

A sire-dam model was used for statistical reasons (see above). This raises the question whether V_W is a trait of the parent(s) or a trait of the offspring (see also discussion). If V_W was assumed a trait of the offspring, then the sire-dam variance was multiplied with four to obtain the additive genetic variance (Sonesson et al. 2013). If V_W was assumed a trait of both parents, then the sire-dam variance was multiplied with two to obtain the total genetic variance due to both parents assuming that the maternal and paternal genetic variance are equal. In order to compare results to literature, we calculated the heritability of V_E (h_v^2) and its genetic coefficient of variation (GCV_{V_e}) for both considering V_W as a trait of the offspring and as a trait of the parents assuming that all genetic heterogeneity in V_W is due to genetic heterogeneity in V_E . The h_v^2 can be used to calculate the accuracy of selection and GCV_{V_e} indicates how much V_E could be changed by selection (Mulder et al. 2007). The h_v^2 is defined as the regression of the breeding value for V_E in an additive model for V_E on the squared phenotypic deviation as an analogy of the normal heritability (Mulder et al. 2007). The h_v^2

was calculated following Mulder et al. (2007) and Felleki and Lundeheim (2012); the GCV_{Ve} was the square root of $\sigma_{a_v}^2$ (see Appendix). Standard errors for h_v^2 and GCV_{Ve} were approximated according to Mulder et al. (2016).

The effect of environment on V_W

To estimate effects of year on fledging weight and its V_W from model 1, we obtained a simple linear regression of estimates of V_W on estimates of fledging weight. To investigate whether genetic variation in V_W is higher under bad conditions than under good conditions, we estimated the genetic variance and covariance for V_W in years with high and low mean fledging weight using a bivariate model on Ψ . The dataset was divided in approximately equal halves based on the estimated year effect for fledging weight in equation 1. We used the response variable Ψ and the weights \mathbf{W}_v from the last iteration of DHGLM. We used LRT to test whether the genetic variance in V_W was the same in the two groups of years and whether the genetic correlation was significantly different from one.

Relationships between V_W and fitness

When a relationship between fledging weight and fitness at individual chick level is non-linear, it will create a relationship between V_W of fledging weight and fitness at the brood level (Figure 2 A, B). Therefore, the relationship between V_W and fitness was investigated at two levels: (1) the relationship between fledging weight and recruit survival at individual chick level and (2) the relationship between variance and the number of recruits at brood level. First at individual chick level, we calculated for each chick whether it appeared as parent in the pedigree, i.e. returned as a recruit to the population. We excluded the year 2013,

because at the time of creating the dataset and pedigree it was unknown whether chicks hatched in 2013 had recruited. The logistic regression model implemented in ASReml contained as fixed effects fledging weight, its squared value, a year effect and a random brood effect. These were all highly significant. Significance was based on the Wald-test using the working variable, which is an approximate test.

Secondly, the number of recruits per brood was analyzed using year always as a fixed class effect, which is similar to analyzing relative fitness, i.e. the number of recruits divided by the average number of recruits per brood per year (Lande and Arnold 1983). We used the glm-package in R because the number of recruits is Poisson distributed (R Development Core team, 2015). We regressed the number of recruits on the mean fledging weight, brood size and the log-variance of fledging weight of each brood and the quadratic term of the log-variance to investigate potential non-linearity. Note that we used here raw means and variances and not those estimated from model 1, because pre-adjustment can lead to over-correction and can remove valuable information. These effects were tested using a χ^2 -test.

Results

Genetic variance in fledging weight and its V_W

To investigate the importance of genetic variation in V_W we compared different models for V_W . Comparing a model with a random brood effect for V_W (Model 3) to a model without any random effects (Model 2), indicated large and highly significant differences between broods in V_W ($P < 0.001$, LRT, χ^2 distribution with 1 degree and 2 degrees of freedom) (Table 1). Subsequently, the model with sire-dam genetic effects in addition to brood effects (Model 1) was tested. It was found that sire-dam genetic effects were significant ($P = 0.028$, LRT, χ^2 -

distribution with mixture of 1 degree and 2 degrees of freedom), indicating the presence of significant genetic variance in V_W . Furthermore, the AIC indicated that the model for V_W with sire-dam genetic effects and brood effects was the best fitting model. This model testing shows significant evidence for genetic variation in V_W for fledging weight.

Using the full model, variance components and its derived parameters for fledging weight and its V_W were obtained (Table 2). The heritability for fledging weight was 0.21 (se = 0.05). Brood effects explained almost half of the phenotypic variance (44%), indicating large differences between broods in mean fledging weight. The other almost half of the phenotypic variance was within broods (V_W) contributing 46% to the phenotypic variance. For V_W of fledging weight, we estimated a large genetic variance as well as large brood effects. When considering V_W as a trait of the offspring, the genetic variance on the exponential scale was 0.065 (se = 0.039) and the GCV_{V_e} was 0.25 (se = 0.077). This indicates that V_E would change by 25% when the directional selection differential would be one genetic standard deviation and assuming that all genetic heterogeneity in V_W is due to genetic heterogeneity in V_E . The heritability h_p^2 was $4.8 * 10^{-3}$ (se = $2.8 * 10^{-3}$). When considering V_W as a trait of the parents, the genetic variance was halved, 0.032 (se = 0.020), the GCV_{V_e} was 0.18 (se = 0.054) and the heritability was $3.8 * 10^{-3}$ (se = $2.1 * 10^{-3}$). The estimated genetic correlation between mean fledging weight and V_W was 0.10 (se = 0.29). The estimated genetic correlation was not significantly different from zero ($P = 0.45$; LRT with one degree of freedom comparing reduced model with zero genetic correlation and full model with estimated genetic correlation), indicating no significant genetic relationship between mean fledging weight and its V_W . Besides genetic differences in V_W , we observed large non-genetic differences in V_W indicating large differences between broods in V_W , i.e. the brood effect. The correlation between brood effects for fledging weight and its V_W was -0.55 (se = 0.05) and highly

significantly deviating from zero ($P < 0.001$, LRT with one degree of freedom comparing reduced model with zero brood correlation and full model with estimated brood correlation), indicating that broods with higher mean fledging weight had lower V_W . The analyses show existence of genetic variation in V_W as well as large non-genetic differences between broods in V_W .

The effect of environmental quality and clutch size on V_W

To investigate the impact of environmental effects on V_W and fledging weight, we investigated the estimated effects of year and clutch size on fledging weight and its V_W . The V_W of fledging weight increased when the annual average fledging weight decreased ($r = -0.70$, $se = 0.12$), thus in years with poor environmental conditions, i.e. food shortage, V_W was large (Figure 3). The difference in V_W was four-fold between years with low fledging weight compared to years with high fledging weight. With increasing clutch size, independent of year, fledging weight decreased and V_W increased (Figure 4). For fledging weight, the linear and quadratic regression coefficient were significantly different from zero ($\beta_1 = -0.60$, $se_{\beta_1} = 0.16$; $P < 0.001$; $\beta_2 = 0.32$, $se_{\beta_2} = 0.15$; $P = 0.03$; Student t-test two sided with 15357 degrees of freedom), whereas for V_W only the linear regression coefficient was significantly different from zero ($\beta_1 = 0.27$, $se_{\beta_1} = 0.10$; $P = 0.006$; $\beta_2 = 0.05$, $se_{\beta_2} = 0.10$; $P = 0.57$; Student t-test two sided with 15357 degrees of freedom). The opposite effects of clutch size on mean fledging weight and its V_W indicated again a strong negative correlation between mean fledging weight and its V_W . In summary, high clutch sizes and years with poor environmental conditions increased V_W and decreased mean fledging weight.

To investigate whether genetic variation in V_W is higher under bad conditions than under good conditions, we performed a bivariate analysis on V_W in bad and good years. The analysis showed no significant evidence that genetic variance was different in bad or good years, although estimates were numerically different (genetic variance in V_W : 0.16 (se = 0.08) in bad years and 0.04 (se = 0.06) in good years) (null model: equal genetic variance; alternative model: different genetic variance; LRT one degree of freedom: $P = 0.20$). Furthermore, the genetic correlation between V_W in bad and good years was estimated to be 0.32 (se = 0.78), but not significantly different from either zero (null model: genetic correlation = 0; alternative model: genetic correlation is estimated; LRT one degree of freedom: $P = 0.86$) or one (null model: genetic correlation = 1.0; alternative model: genetic correlation is estimated; LRT one degree of freedom: $P = 1.00$). It seems that we do not have sufficient power to answer the question whether V_W is genetically more expressed in bad years than in good years and whether there is evidence for genotype by environment interactions.

Relationship between fledging weight and fitness

When investigating the relationship between recruitment survival and fledging weight at the chick level (level 1), a highly significant quadratic relationship was found using logistic regression ($P = 0.005$; logistic scale: linear effect = 1.26 (se = 0.36); quadratic effect = -0.029 (se = 0.01)) (Figure 2C). Furthermore, year effects were highly significant ($P < 0.001$) indicating substantial differences between years in survival.

Relationships between V_W and fitness

When investigating the relationship between V_W and the number of recruits at brood level (level 2), we found no association between the number of recruits and log-variance in fledging weight (χ^2 -test, $P = 0.21$). We found significant associations between number of recruits and mean fledging weight (χ^2 -test; $P < 0.001$; log-scale: effect = 0.21 (se = 0.03)) and brood size (χ^2 -test; $P < 0.001$; as factor). This shows that broods with a higher mean fledging weight have a higher number of recruits, but there is no linear effect of V_W on fitness. However, when testing the quadratic regression of number of recruits on log-variance of fledging weight, it was highly significant (χ^2 -test; $P < 0.001$; log-scale: linear effect = -0.14 (se = 0.05); quadratic effect = -0.11 (se = 0.04)). Figure 2D shows that families with either low or high log-variance of fledging weight have lower fitness than families with intermediate log-variance. This quadratic relationship indicates that there is stabilizing selection on V_W .

Discussion

Genetic variance in V_E

The aims of this study were to quantify the genetic variation in V_E for fledging weight and to explore the relationships between V_E and fitness in a wild great tit population. We analysed the within-brood variance (V_W) instead of V_E , because a sire-dam model was used. We found genetic variance in V_W and a significant quadratic relationship between V_W and fitness. To interpret the size of the additive genetic variance in V_W , we assume that genetic variance in V_W is entirely due to genetic variance in V_E . Analogous to the genetic coefficient of variation for traits (GCV , Houle 1992), the genetic variance in V_E can also be expressed as a genetic coefficient of variation for V_E (GCV_{V_e}) (see Mulder et al., 2007). The GCV_{V_e} was 0.18 when

considering V_E as a trait of the parents and 0.25 when considering it as a trait of the offspring. The estimate of GCV_{V_e} at the level of the offspring was lower than the median value of 0.34 but well within the range ($SD = 0.21$) found in domestic species (Hill and Mulder 2010). The heritability estimate either at parental level or offspring level was smaller than in most domestic populations due to the substantial variance attributed to brood, which leads to a relatively small proportion of V_E , i.e. the within-brood variance, to the phenotypic variance after correcting for systematic environmental effects ($V_E/V_P = 0.37$). Here we used a sire-dam model, which means that V_W contains the Mendelian sampling variance if we assume that V_E is a trait of the offspring and in absence of dominance or epistasis. In Mulder et al. (2013b) we used an adjusted Double Hierarchical Generalized Linear Model (DHGLM) to get unbiased estimates of genetic variance in V_E when using a sire model. Preliminary analysis showed that an adjusted DHGLM did not give stable variance components and therefore the standard DHGLM was used here. However, this leads to underestimation of the genetic variance in V_E when using a sire(-dam) model. If the adjustment would be applied afterwards, the GCV_{V_e} at offspring level would increase from 0.25 to 0.33.

Besides being genetically caused, heterogeneity of within-family variance can be caused by: (1) scale of measurement, (2) differences in measurement error, (3) differences in developmental homeostasis, (4) mortality selection against extremes, (5) segregation of major genes, (6) differences in levels of inbreeding between parents of broods (Mitchell-Olds and Bergelson 1990), or (7) genotype by environment interactions (Mulder et al. 2013b). We will discuss in the following why we do not think that these alternative explanations apply here, except the last one (7).

(1) Scale of measurement typically leads to higher variances with higher means for biological traits (Falconer and Mackay 1996; Lynch and Walsh 1998). This would suggest a

positive correlation between fledging weight and its V_W . We observed, however, negative relationships for brood and year effect (Figure 3), although we observed a small, but not significant positive genetic correlation between fledging weight and its V_W .

(2) Measurement error contributes to V_W and different depths of pedigree affect estimation of genetic variance in V_W . The data did not offer the possibility to quantify the contribution of measurement error to V_W as this would require repeated observations on the same individual, which obviously is impossible for fledging weight. The measurement error is, however, expected to be random and its variance homogeneous and not confounded with genetic effects as fledging weight is measured by simply reading off spring balances, which does not require special training or experience. As expected, we hence found no evidence that observer effects explained any variance in fledging weight and its V_W . Any possible changes in equipment over time are accounted for by fixed year effects on V_W , although no significant trend was observed in year estimates (ANOVA, F-test, $P=0.24$). Furthermore, different depths of pedigrees result in variation in precision of estimated random effects and their residuals. The DHGLM accounts for this by using leverages and weights. Simulations show that estimates of genetic variance in V_W are unbiased with DHGLM (Felleki et al. 2012; Mulder et al. 2013b). Therefore, estimates of genetic variance in V_W are not expected to be affected by measurement error or differences in precision of random effects due to differences in pedigree depth. However, confounding of pedigree with environmental effects could still have inflated the genetic variance.

(3) Differences in developmental homeostasis are expected to give higher variance for families that deviate more from the mean, but with lower variance close to the centre of the distribution (Mitchell-Olds and Bergelson 1990). Here we found smaller variance with increasing fledging weight, rather than higher variance for families with extreme means.

(4) Mortality against extremes could have played a role here in the form of adaptive brood reduction. At the day of measuring fledging weight, small chicks could have died already and therefore reducing V_W . However, the scope for adaptive brood reduction is small in great tits as we found that, on average, the number of weighed chicks was 1.4 smaller than the clutch size.

(5) Segregation of major genes causing heterogeneity of within-family variance is unexpected, because traits like fledging weight are highly polygenic, although the expected polygenic nature of the trait does not rule out existence of one or a few major genes (Rowe et al. 2006; Manolio et al. 2009).

(6) Differences in levels of inbreeding between parents of broods are expected to be small because the average inbreeding coefficient in the used pedigree was $9.22 * 10^{-4}$ and only 297 animals had an inbreeding coefficient larger than 0.06.

(7) Genotype by environment interactions in response to micro-environmental variables may contribute to differences in within-family variance. In dairy cattle, it was shown that macro-environmental sensitivity and V_E have largely a common genetic background as indicated by a genetic correlation of 0.77 (Mulder et al. 2013b). Differences in within-family variance may arise as a consequence of response to multiple environmental factors, i.e. multidimensional phenotypic plasticity (Westneat et al. 2014).

Are fledging weight and its V_W traits of the parents of the offspring?

One of the key questions in the genetic analysis of fledging weight and its V_W is whether fledging weight is an offspring trait or a parental trait. Preliminary results showed that for

fledging weight a model with direct genetic and maternal genetic effects had equal log-likelihood compared to a model with only direct genetic effects, indicating either absence of maternal genetic effects or complete confounding between direct genetic and maternal genetic effects. From a biological point of view (see Figure 1), fledging weight is affected by the genes of the individual chick, but also by the genes for parental care of the parents, indicating how good they are in provisioning food to their chicks (i.e. brood environment). Some evidence for the latter comes from the finding that heritability estimates from parent-offspring regression either on own parent or on foster parent when chicks were cross-fostered were similar (Gebhardt-Henrich and Van Noordwijk 1991), which indicates that parental care, which consists of genetic and non-genetic effects (MacColl and Hatchwell 2003; Hadfield 2012), can play an important role compared to direct genetic effects. Hadfield et al. (2013) however showed in blue tits using cross-fostering that the heritability of body mass is smaller than previously reported, i.e. 0.06 - 0.08.

For V_W of fledging weight, it becomes even more the question whether V_W is genetically a trait of the offspring or a trait of the parents or both. It is likely that both direct genetic effects as well as parental effects affect V_W of fledging weight. Parental effects relate to food allocation strategies of the parents whether they feed all chicks equally well or favour the larger ones compared to the small ones (Figure 1). Because of differences in allocation of food to their chicks (Dickens et al. 2008; Tanner et al. 2008), each chick has a different individual environment and the chicks respond differently to their individual environment, for instance in begging behaviour (Dickens et al. 2008) or plasticity of growth (Kunz and Ekman 2000). If both direct genetic and parental genetic effects affect V_W of fledging weight, then the sire-dam model captures part of the total genetic variance. Depending on the sizes of the parental genetic variance, the direct genetic variance and the covariance between them, the

total genetic variance is between two and four times the estimate of the sire-dam additive genetic variance, when the genetic correlation between direct and parental genetic effects is zero or positive.

Behavioural studies (Tanner et al. 2008) have shown that differential food allocation among nestlings is partly influenced by parental choice and by the nestlings by adjusting positioning and begging behaviour. Experimental work such as done in Tanner et al. (2008) could reveal whether parental choice decisions or differential nestling behaviour would contribute to genetic differences in V_W , e.g. by comparing parents with high or low breeding values for V_W . Other recent studies have found evidence for differences in within-individual variance of food provisioning behaviour (Westneat et al. 2013). Studying within-individual variance of behavioural traits related to parental behaviour using for instance double hierarchical generalized linear models (Westneat et al. 2013; Westneat et al. 2014; Cleasby et al. 2015) would open an avenue to better understand the genetics of fledging weight and its V_W .

Heritability versus genetic coefficient of variation: why bother about both measures?

We here presented both the heritability (h_v^2) and the genetic coefficient of variation (GCV_{V_e}) as measures of the ‘evolutionary potential’ to reflect the accuracy of mass selection and the potential response to selection, respectively (Houle 1992; Hansen et al. 2011). We here observed a low h_v^2 and a high GCV_{V_e} of V_E , similar to fitness traits (Houle 1992). When considering the extremely low heritability of 0.005, it could be concluded that heritable variation in V_E can be neglected. However, that is only part of the story because the GCV_{V_W} (0.25 as offspring trait; 0.18 as a parental trait) is much higher than the GCV for fledging

weight itself, which is 0.040. This indicates that there is substantial room for natural selection to shape V_W , even more than to shape fledging weight itself. The low heritability is highly related to the nature of the trait and the definition of the heritability. The heritability is here defined as the regression of the breeding value for V_E on the squared phenotypic deviation as an analogy of the heritability of traits means (Mulder et al. 2007). Therefore, the heritability is an indication of the accuracy of mass selection. To estimate variances with the same precision as means one needs at least five times as many observations according to Tukey's rule (Lee and Nelder 2006). Therefore, the accuracy of mass selection for V_E is low due to high uncertainty, i.e. a high sampling variance with one phenotypic record. In conclusion, considering both the heritability and the GCV shows that there is ample room for natural selection to shape V_E , though the accuracy of mass selection would be low.

Relationships between within-family variance and fitness: consequences for natural selection

We found strong evidence for stabilizing selection on V_W indicated by the significant negative quadratic regression (Figure 2D). Stabilizing selection on V_W (Figure 2D) implies that parents with intermediate variance have the highest fitness and parents with either uniform offspring or very variable offspring have a lower fitness. The estimated parabolic fitness function peaked around a log-variance of -0.54, which is lower than the current mean log-variance (mean = -0.28). Stabilizing selection on the variance can explain the maintenance of the abundant phenotypic variability among individuals within animal populations. The strong evidence for stabilizing selection is in agreement with the sigmoidal curve of recruitment survival as a function of fledging weight (Figure 2C) as we hypothesized

in Figure 1A and 1C. Mathematically, the optimum of the parabola of number of recruits on V_W is at the inflection point of the sigmoidal curve of recruitment survival on fledging weight. The inflection point of the sigmoid was 17.31 g, very close to the mean fledging weight of 17.35g. The effect of sigmoidal trait-fitness curves on selection for or against variance in trait values is an example of Jensen's inequality (Ruel and Ayres 1999; Martin and Huey 2008; Dall 2010; Hamilton 2010).

The presented framework opens an avenue to better understand the maintenance of V_E and increase in V_E in unfavourable conditions (Hoffmann and Merila 1999; Merila et al. 1999; Charmantier and Garant 2005). The increase in V_E in unfavourable conditions suggests that the optimal V_E depends on the environment and unfavourable conditions lead to uncovering hidden genetic variation in V_E . Our data was not suitable to detect a statistically significant genotype by environment interaction for V_E . We found, however, a strong indication that years with high food abundance show lower V_E than years with food shortage. Genomic data could help to reveal the genetics of V_E in favourable and unfavourable conditions. It is known that for instance heat-shock proteins (HSP90) buffer effects of stress in laboratory species such as *Drosophila* and *Arabidopsis* (Rutherford and Lindquist 1998; Queitsch et al. 2002; Sangster et al. 2008). Therefore, genome-wide association studies could help to unravel the biological control of V_E in stressed and unstressed conditions.

In conclusion, we find evidence for genetic differences in V_E of fledging weight and presence of stabilizing selection on V_E , indicating that V_E is maintained. However, understanding the levels of V_E in different environments and the role of genotype by environment interactions is largely lacking. This research provides some first insights and is a stepping-stone for further research to understand the evolution of phenotypic variability in natural populations.

Acknowledgement

Arthur Gilmour is acknowledged for his help with logistic regression in ASReml, the board of the National Park de Hoge Veluwe for their permission to work in the park, all the field workers over the past decades for collecting the data and Louis Vernooij for database management. Johan van Arendonk, two associate editors, the editor in chief and two anonymous reviewers are acknowledged for suggestions to improve the manuscript.

Literature cited

- Both, C., M. E. Visser, and N. Verboven. 1999. Density-dependent recruitment rates in great tits: the importance of being heavier. *Proc. R. Soc. B-Biol. Sci.* 266:465-469.
- Brommer, J. E., J. Merila, B. C. Sheldon, and L. Gustafsson. 2005. Natural selection and genetic variation for reproductive reaction norms in a wild bird population. *Evolution* 59:1362-1371.
- Charmantier, A. and D. Garant. 2005. Environmental quality and evolutionary potential: lessons from wild populations. *Proc. R. Soc. B-Biol. Sci.* 272:1415-1425.
- Clarke, G. M. 1998. Developmental stability and fitness: The evidence is not quite so clear. *Am. Nat.* 152:762-766.
- Cleasby, I. R., S. Nakagawa, and H. Schielzeth. 2015. Quantifying the predictability of behaviour: statistical approaches for the study of between-individual variation in the within-individual variance. *Methods Ecol. Evol.* 6:27-37.
- Dall, S. R. X. 2010. Managing risk: the perils of uncertainty *in* D. F. Westneat, and C. W. Fox, eds. *Evolutionary behavioral ecology*. Oxford University Press, New York, US.

- De Visser, J. A. G. M., J. Hermisson, G. P. Wagner, L. Ancel Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, G. Gibson, T. F. Hansen, D. Krakauer, R. C. Lewontin, C. Ofria, S. H. Rice, G. Von Dassow, A. Wagner, and M. C. Whitlock. 2003. Perspective: Evolution and detection of genetic robustness. *Evolution* 57:1959-1972.
- Debat, V. and P. David. 2001. Mapping phenotypes: canalization, plasticity and developmental stability. *Trends Ecol. Evol.* 16:555-561.
- Dickens, M., D. Berridge, and I. R. Hartley. 2008. Biparental care and offspring begging strategies: hungry nestling blue tits move towards the father. *Anim. Behav.* 75:167-174.
- Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Pearson Education Limited, Essex, UK.
- Felleki, M., D. Lee, Y. Lee, A. R. Gilmour, and L. Rönnegård. 2012. Estimation of breeding values for mean and dispersion, their variance and correlation using double hierarchical generalized linear models. *Genet. Res.* 94:307-317.
- Felleki, M. and N. Lundeheim. 2012. Genetic control of residual variance in teat number in pigs. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 20:538-541.
- Flatt, T. 2005. The evolutionary genetics of canalization. *Q. Rev. Biol.* 80:287-316.
- Gavrilets, S. and A. Hastings. 1994. A quantitative-genetic model for selection on developmental noise. *Evolution* 48:1478-1486.
- Gebhardt-Henrich, S. G. and A. J. Van Noordwijk. 1991. Nestling growth in the Great Tit I. Heritability estimates under different environmental conditions. *J. Evol. Biol.* 4:341-362.
- Gibson, G. 2009. Decanalization and the origin of complex diseases. *Nat. Rev. Genet.* 10:134-140.
- Gibson, G. and I. Dworkin. 2004. Uncovering cryptic genetic variation. *Nat. Rev. Genet.* 5:681-690.

- Gibson, G. and G. Wagner. 2000. Canalization in evolutionary genetics: a stabilizing theory? *BioEssays* 22:372-380.
- Gienapp, P., E. Postma, and M. E. Visser. 2006. Why breeding time has not responded to selection for earlier breeding in a songbird population. *Evolution* 60:2381-2388.
- Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2006. ASReml User Guide Release 2.0. VSN International Ltd, Hemel Hempstead.
- Gomulkiewicz, R. and R. D. Holt. 1995. When does evolution by natural selection prevent extinction? *Evolution* 49:201-207.
- Hadfield, J. D. 2012. The quantitative genetic theory of parental effects. *in* N. J. Royle, P. T. Smiseth, and M. Kölliker, eds. *The evolution of parental care*. Oxford University Press, Oxford, UK.
- Hadfield, J. D., E. A. Heap, F. Bayer, E. A. Mittell, and N. M. A. Crouch. 2013. Disentangling genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine. *Evolution* 67:2701-2713.
- Hamilton, I. M. 2010. Foraging theory *in* D. F. Westneat, and C. W. Fox, eds. *Evolutionary behavioral ecology*. Oxford University Press, New York, US.
- Hansen, T. F., C. Pelabon, and D. Houle. 2011. Heritability is not evolvability. *Evol. Biol.* 38:258-277.
- Hill, W. G. and H. A. Mulder. 2010. Genetic analysis of environmental variation. *Genet. Res.* 92:381-395.
- Hill, W. G. and X.-S. Zhang. 2004. Effects of phenotypic variability of directional selection arising through genetic differences in residual variability. *Genet. Res. Camb.* 83:121-132.
- Hoaglin, D. C. and R. E. Welsh. 1978. The hat matrix in regression and ANOVA. *Am. Stat.* 32:17-22.

- Hoffmann, A. A. and J. Merila. 1999. Heritable variation and evolution under favourable and unfavourable conditions. *Trends Ecol. Evol.* 14:96-101.
- Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130:195-204.
- Kaplan, R. H. and W. S. Cooper. 1984. The evolution of developmental plasticity in reproductive characteristics - an application of the adaptive coin-flipping principle. *Am. Nat.* 123:393-410.
- Kaufman, P. K., F. D. Enfield, and R. E. Comstock. 1977. Stabilizing selection for pupa weight in *Tribolium castaneum*. *Genetics* 87:327-341.
- Kingsolver, J. G., S. E. Diamond, A. M. Siepielski, and S. M. Carlson. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* 26:1101-1118.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245-261.
- Kunz, C. and J. Ekman. 2000. Genetic and environmental components of growth in nestling blue tits (*Parus caeruleus*). *J. Evol. Biol.* 13:199-212.
- Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210-1226.
- Lee, Y. and J. A. Nelder. 2006. Double hierarchical generalized linear models. *Appl. Stat.* 55:139-185.
- Lynch, M. and R. Lande. 1993. Evolution and extinction in response to environmental change. Pp. 234–250 *in* P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. *In Biotic interactions and global change*. Sinauer, Sunderland, MA.

- Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, US.
- MacColl, A. D. C. and B. J. Hatchwell. 2003. Heritability of parental effort in a passerine bird. *Evolution* 57:2191-2195.
- Mackay, T. F. C. and R. F. Lyman. 2005. *Drosophila* bristles and the nature of quantitative genetic variation. *Philo. Trans. R. Soc. B.* 360:1513-1527.
- Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. C. Mackay, S. A. McCarroll, and P. M. Visscher. 2009. Finding the missing heritability of complex diseases. *Nature* 461:747-753.
- Martin, T. L. and R. B. Huey. 2008. Why "Suboptimal" is optimal: Jensen's inequality and ectotherm thermal preferences. *Am. Nat.* 171:E102-E118.
- Masel, J. and M. L. Siegal. 2009. Robustness: mechanisms and consequences. *Trends Genet.* 25:395-403.
- Masel, J. and M. V. Trotter. 2010. Robustness and evolvability. *Trends Genet.* 26:406-414.
- Merila, J., R. Przybylo, and B. C. Sheldon. 1999. Genetic variation and natural selection on blue tit body condition in different environments. *Genet. Res. Camb.* 73:165-176.
- Mitchell-Olds, T. and J. Bergelson. 1990. Statistical genetics of annual plant, *Impatiens Capensis*. 1. Genetic basis of quantitative variation. *Genetics* 124:407-415.
- Monros, J. S., E. J. Belda, and E. Barba. 2002. Post-fledging survival of individual great tits: the effect of hatching date and fledging mass. *Oikos* 99:481-488.

- Morgante, F., P. Sorensen, D. A. Sorensen, C. Maltecca, and T. F. C. Mackay. 2015. Genetic architecture of micro-environmental plasticity in *Drosophila melanogaster*. *Sci. Rep.* 5:09785.
- Mulder, H. A., P. Bijma, and W. G. Hill. 2007. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics* 175:1895-1910.
- Mulder, H. A., R. E. Crump, M. P. L. Calus, and R. F. Veerkamp. 2013a. Unraveling the genetic architecture of environmental variance of somatic cell score using high density SNP and cow data from experimental farms. *J. Dairy Sci.* 96:7306-7317.
- Mulder, H. A., W. G. Hill, and E. F. Knol. 2015. Heritable environmental variance causes non-linear relationships between traits: application to birth weight and stillbirth of pigs. *Genetics* 199:1255-1269.
- Mulder, H. A., L. Rönnegård, W. F. Fikse, R. F. Veerkamp, and E. Strandberg. 2013b. Estimation of genetic variance for macro- and micro-environmental sensitivity using double hierarchical generalized linear models. *Genet. Sel. Evol.* 45:23.
- Mulder, H. A., J. Visscher, and J. Fablet. 2016. Estimating the purebred-crossbred genetic correlation for uniformity of eggshell color in laying hens. *Genet. Sel. Evol.* 48:39.
- Nussey, D. H., E. Postma, P. Gienapp, and M. E. Visser. 2005. Selection on heritable phenotypic plasticity in a wild bird population. *Science* 310:304-306.
- Palmer, A. R. and C. Strobeck. 1986. Fluctuating asymmetry - measurement, analysis, patterns. *Annu. Rev. Ecol. Syst.* 17:391-421.
- Queitsch, C., T. A. Sangster, and S. Lindquist. 2002. Hsp90 as a capacitor of phenotypic variation. *Nature* 417:618-624.
- R. 2015. R Development Core Team, R: A language and environment for statistical computing [<http://www.R-project.org>].

- Rendel, J. M., B. L. Sheldon, and D. E. Finlay. 1966. Selection for canalization of the scute phenotype. II. *Am. Nat.* 100:13-31.
- Rönnegård, L., M. Felleki, F. Fikse, H. A. Mulder, and E. Strandberg. 2010. Genetic heterogeneity of residual variance: estimation of variance components using double hierarchical generalized linear models. *Genet. Sel. Evol.* 42:8.
- Rowe, S. J., I. M. S. White, S. Avendano, and W. G. Hill. 2006. Genetic heterogeneity of residual variance in broiler chickens. *Genet. Sel. Evol.* 38:617-635.
- Ruel, J. J. and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. *Trends in Ecology & Evolution* 14:361-366.
- Rutherford, S. L. and S. Lindquist. 1998. Hsp90 as a capacitor for morphological evolution. *Nature* 396:336-342.
- Sae-Lim, P., A. Kaune, M. Janhunen, H. Vehvilainen, H. Koskinen, B. Gjerde, M. Lillehammer, and H. A. Mulder. 2015. Genetic (co)variance of rainbow trout (*Oncorhynchus mykiss*) body weight and its uniformity across production environments. *Genet. Sel. Evol.* 47:46.
- Sangster, T. A., N. Salathia, S. Undurraga, R. Milo, K. Schelienberg, S. Lindquist, and C. Queitsch. 2008. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. *Proc. Natl. Acad. Sci. U. S. A.* 105:2963-2968.
- Scharloo, W., A. Zweep, K. A. Schuitema, and J. G. Wijnstra. 1972. Stabilizing and disruptive selection on a mutant character in *Drosophila*. IV. Selection on sensitivity to temperature. *Genetics* 71:551-566.
- Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. *Annu. Rev. Ecol. Syst.* 24:35-68.
- Schlichting, C. D. and M. Pigliucci. 1998. Phenotypic evolution: a reaction norm perspective. Sinauer Associates, Inc. Publishers, Sunderland, MA, US.

- Seger, J. and H. J. Brockmann. 1987. What is bet-hedging? *Oxf. Surv. Evol. Biol.* 4:182-211.
- Sell-Kubiak, E., N. Duijvesteijn, M. S. Lopes, L. L. G. Janss, E. F. Knol, P. Bijma, and H. A. Mulder. 2015. Genome-wide association study reveals novel loci for litter size and its variability in a Large White pig population. *BMC Genomics* 16:1049.
- Shen, X., M. Pettersson, L. Rönnegård, and O. Carlborg. 2012. Inheritance beyond plain heritability: variance-controlling genes in *Arabidopsis thaliana*. *PLoS Genetics* 8:8.
- Sonesson, A. K., J. Odegard, and L. Ronnegard. 2013. Genetic heterogeneity of within-family variance of body weight in Atlantic salmon (*Salmo salar*). *Genet. Sel. Evol.* 45:41.
- Sorensen, P., G. de los Campos, F. Morgante, T. F. C. Mackay, and D. Sorensen. 2015. Genetic control of environmental variation of two quantitative traits of *Drosophila melanogaster* revealed by whole-genome sequencing. *Genetics* 201:487-497.
- Stram, D. O. and J. W. Lee. 1994. Variance components testing in the longitudinal mixed effects model. *Biometrics* 50:1171-1177.
- Tanner, M., M. Kolliker, and H. Richner. 2008. Differential food allocation by male and female great tit, *Parus major*, parents: are parents or offspring in control? *Anim. Behav.* 75:1563-1569.
- Tinbergen, J. M. and M. C. Boerlijst. 1990. Nestling weight and survival in individual great tits (*Parus major*). *J. Anim. Ecol.* 59:1113-1127.
- Verboven, N. and M. E. Visser. 1998. Seasonal variation in local recruitment of great tits: the importance of being early. *Oikos* 81:511-524.
- Visscher, P. M. 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Res. Hum. Genet.* 9:490-495.
- Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* 150:563-565.
- Waddington, C. H. 1960. Experiments on canalizing selection. *Genet. Res. Camb.* 1:140-150.

- Wagner, G. P., G. Booth, and H. Bagheri-Chaichian. 1997. A population genetic theory of canalization. *Evolution* 51:329-347.
- Westneat, D. F., M. Schofield, and J. Wright. 2013. Parental behavior exhibits among-individual variance, plasticity, and heterogeneous residual variance. *Behav. Ecol.* 24:598-604.
- Westneat, D. F., J. Wright, and N. J. Dingemanse. 2014. The biology hidden inside residual within-individual phenotypic variation. *Biol Rev*:doi: 10.1111/brv.12131.
- Yang, J., R. J. F. Loos, J. E. Powell, S. E. Medland, E. K. Speliotes, D. I. Chasman, L. M. Rose, G. Thorleifsson, V. Steinthorsdottir, R. Maegi, L. Waite, A. V. Smith, L. M. Yerges-Armstrong, K. L. Monda, D. Hadley, A. Mahajan, G. Li, K. Kapur, V. Vitart, J. E. Huffman, S. R. Wang, C. Palmer, T. Esko, K. Fischer, J. H. Zhao, A. Demirkan, A. Isaacs, M. F. Feitosa, J. Luan, N. L. Heard-Costa, C. White, A. U. Jackson, M. Preuss, A. Ziegler, J. Eriksson, Z. Kutalik, F. Frau, I. M. Nolte, J. V. Van Vliet-Ostaptchouk, J. J. Hottenga, K. B. Jacobs, N. Verweij, A. Goel, C. Medina-Gomez, K. Estrada, J. L. Bragg-Gresham, S. Sanna, C. Sidore, J. Tyrer, A. Teumer, I. Prokopenko, M. Mangino, C. M. Lindgren, T. L. Assimes, A. R. Shuldiner, J. Hui, J. P. Beilby, W. L. McArdle, P. Hall, T. Haritunians, L. Zgaga, I. Kolcic, O. Polasek, T. Zemunik, B. A. Oostra, M. J. Junttila, H. Groenberg, S. Schreiber, A. Peters, A. A. Hicks, J. Stephens, N. S. Foad, J. Laitinen, A. Pouta, M. Kaakinen, G. Willemsen, J. M. Vink, S. H. Wild, G. Navis, F. W. Asselbergs, G. Homuth, U. John, C. Iribarren, T. Harris, L. Launer, V. Gudnason, J. R. O'Connell, E. Boerwinkle, G. Cadby, L. J. Palmer, A. L. James, A. W. Musk, E. Ingelsson, B. M. Psaty, J. S. Beckmann, G. Waeber, P. Vollenweider, C. Hayward, A. F. Wright, I. Rudan, L. C. Groop, A. Metspalu, K. T. Khaw, C. M. van Duijn, I. B. Borecki, M. A. Province, N. J. Wareham, J. C. Tardif, H. V. Huikuri, L. A. Cupples, L. D. Atwood, C. S. Fox, M. Boehnke, F. S. Collins, K. L. Mohlke, J. Erdmann, H. Schunkert, C. Hengstenberg, K.

Stark, M. Lorentzon, C. Ohlsson, D. Cusi, J. A. Staessen, M. M. Van der Klauw, P. P. Pramstaller, S. Kathiresan, J. D. Jolley, S. Ripatti, M. R. Jarvelin, E. J. C. de Geus, D. I. Boomsma, B. Penninx, J. F. Wilson, H. Campbell, S. J. Chanock, P. van der Harst, A. Hamsten, H. Watkins, A. Hofman, J. C. Witteman, M. C. Zillikens, A. G. Uitterlinden, F. Rivadeneira, L. A. Kiemeny, S. H. Vermeulen, G. R. Abecasis, D. Schlessinger, S. Schipf, M. Stumvoll, A. Toenjes, T. D. Spector, K. E. North, G. Lettre, M. I. McCarthy, S. I. Berndt, A. C. Heath, P. A. F. Madden, D. R. Nyholt, G. W. Montgomery, N. G. Martin, B. McKnight, D. P. Strachan, W. G. Hill, H. Snieder, P. M. Ridker, U. Thorsteinsdottir, K. Stefansson, T. M. Frayling, J. N. Hirschhorn, M. E. Goddard and P. M. Visscher. 2012. FTO genotype is associated with phenotypic variability of body mass index. *Nature* 490:267-273.

Zhang, X.-S. and W. G. Hill. 2005. Evolution of the environmental component of the phenotypic variance: stabilizing selection in changing environments and the cost of homogeneity. *Evolution* 59:1237-1244.

Appendix

Calculation of heritability and genetic coefficient of variation of V_E

Although equations were derived elsewhere, we present them here for readers interested without having the need to go back to other articles. The h_v^2 was defined using an additive model for V_E rather than an exponential model as used in the DHGLM. Therefore, the genetic variance in V_E in an additive model ($\sigma_{av,add}^2$) was calculated as (Mulder et al. 2007; Felleki and Lundeheim 2012; Sae-Lim et al. 2015):

$$\sigma_{a_v,add}^2 = (\sigma_{a_v,add}^2 + \sigma_{n_v,add}^2) \left(\frac{\sigma_{a_v}^2}{\sigma_{a_v}^2 + \sigma_{n_v}^2} \right) \quad (A1)$$

where

$$(\sigma_{a_v,add}^2 + \sigma_{n_v,add}^2) = \sigma_{E_{exp}}^4 \exp(2\sigma_{a_v}^2) \exp(2\sigma_{n_v}^2) - \sigma_{E_{exp}}^2 \exp(\frac{1}{2}\sigma_{a_v}^2) \exp(\frac{1}{2}\sigma_{n_v}^2) \quad (A2)$$

where

$$\sigma_{E_{exp}}^2 = \sigma_E^2 / \exp(\frac{1}{2}\sigma_{a_v}^2) \exp(\frac{1}{2}\sigma_{n_v}^2) \quad (A3)$$

where σ_E^2 is the average V_E estimated with the model with homogeneous residual variance in step 1 of the algorithm. Note that in this case V_E is considered a trait of the offspring $\sigma_E^2 = \sigma_e^2 - \frac{1}{2}\sigma_a^2$, because the residual variance contains half of the additive genetic variance.

Subsequently, h_v^2 is calculated as:

$$h_v^2 = \frac{\sigma_{a_v,add}^2}{2\sigma_P^4 + 3(\sigma_{a_v,add}^2 + \sigma_{n_v,add}^2)} \quad (A4)$$

where σ_P^4 is the squared phenotypic variance ($\sigma_P^2 = \frac{1}{2}\sigma_a^2 + \sigma_n^2 + \sigma_e^2$) and $\sigma_{n_v,add}^2$ is the variance due to brood, similar to equation A1:

$$\sigma_{n_{v,add}}^2 = (\sigma_{a_{v,add}}^2 + \sigma_{n_{v,add}}^2) \left(\frac{\sigma_{n_v}^2}{\sigma_{a_v}^2 + \sigma_{n_v}^2} \right) \quad (\text{A5})$$

The GCV_{V_e} is calculated as:

$$GCV_{V_e} = \sigma_{a_v} \quad (\text{A6})$$

Table 1. Adjusted profile h-likelihood (APHL) and Akaike's information criterion (AIC) comparing different variance models within DHGLM while mean model was kept the same.

The full model with a random sire-dam genetic and random brood effect was set to a APHL = 0 and AIC = 0 for comparison; lower values are better.

Name	Model fledging weight (y)	Model on V_W (ψ)	APHL	AIC
Model 1	Fixed + sire-dam + brood	Fixed + sire-dam + brood	0.00 ¹	0.00
Model 2	Fixed + sire-dam + brood	Fixed + brood	6.27	2.27
Model 3	Fixed + sire-dam + brood	Fixed + no random effects	882.45	874.45

¹ P = 0.028 (Variance model with fixed, random sire-dam and random brood tested against reduced variance model with only fixed and random brood effects; Likelihood Ratio Test: χ^2 distribution with a mixture of one and two degrees of freedom (Stram and Lee 1994; Visscher 2006).

Table 2. Variance components (and their approximated se) for phenotype and V_W of fledging weight using a sire-dam model in DHGLM (assumes it is an offspring trait; genetic variance $V_W =$ genetic variance V_E).

Variance component	Phenotype	V_W
Genetic variance	0.488 (0.115)	0.065 (0.039)
Brood variance	0.966 (0.060)	0.258 (0.025)
Residual variance ¹	0.854 (0.062)	
Phenotypic variance	2.308 (0.047)	
h^2 ²	0.211 (0.050)	0.005 (0.003)
GCV ³	0.040 (0.005)	0.254 (0.077)
r_{aa_v} ⁴	0.100 (0.293)	
r_{nn_v} ⁴	-0.548 (0.047)	

¹ Residual variance is based on estimate from homogeneous model.

² Se h^2 fledging weight is based upon model with homogeneous residual variance; se h_v^2 is approximated following Mulder et al. (2016).

³ $seGCV \cong se\sigma_a^2/2\sigma_a\mu$; $seGCV_{Ve} \cong se\sigma_{a_v}^2/2\sigma_{a_v}$ (Lynch and Walsh 1998).

⁴ r_{aa_v} is the genetic correlation between fledging weight and its V_W ; r_{nn_v} is the correlation between brood effects for fledging weight and its V_W .

Figure 1. Genetic and environmental effects shaping the phenotype of the offspring and the within-family variance. Differences between offspring within a brood can be caused by variation in the genes of the offspring, by variation in individual environment or both. Within-family variance in offspring phenotype can thus be the result from genetic differences

in the allocation rules of the parents as well as genetic differences between families in plasticity to the individual environment leading to heritable variation in within-family variance.

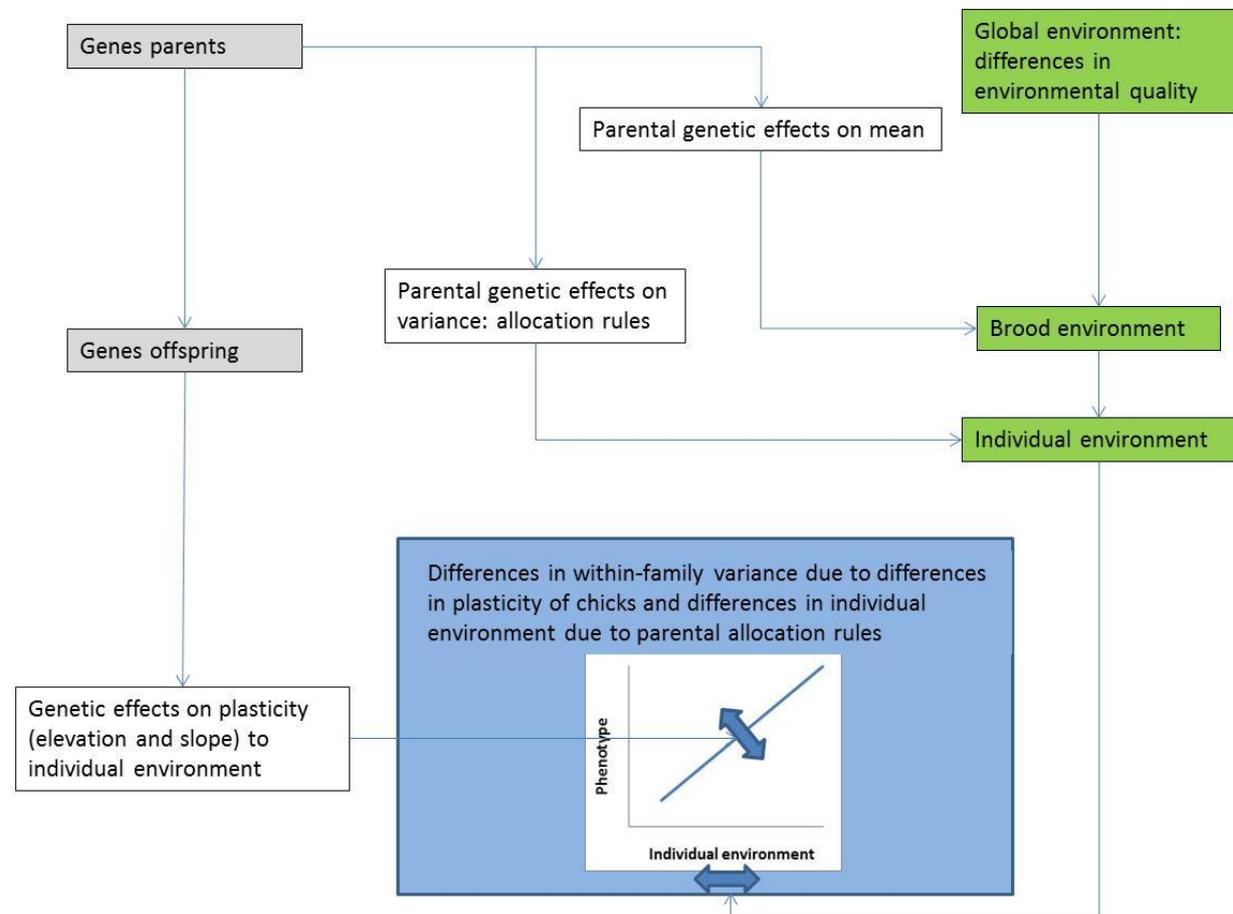


Figure 2. Theoretical and observed relationships between recruitment probability and fledging weight (A and C) and between number of recruits and V_W of fledging weight (B and D). The Figures A and B show how theoretical relationships (linear, quadratic convex/concave or sigmoid) between fledging weight and recruitment probability at the individual chick level (panel A) translate into theoretical relationships between the number of recruits and the log within-brood variance at the brood level. Figures C and D show the same relationships observed in the studied population; obs consists of raw means when dividing

data in 20 bins, pred is the predicted line based on the statistical analyses outlined in ‘Relationships between V_W and fitness’.

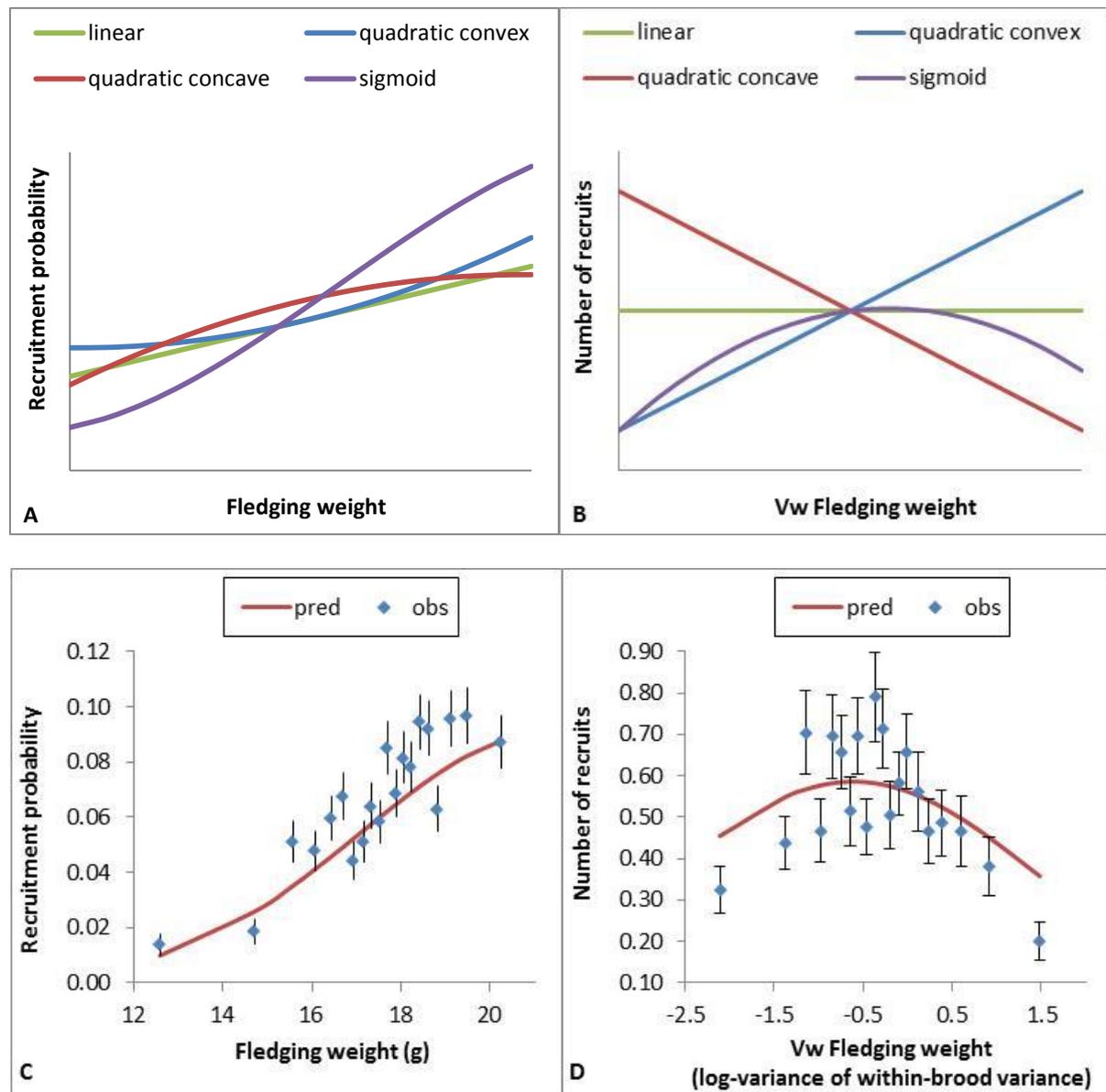


Figure 3. Relative within-brood V_W per year as a function of the mean annual fledging weight (g) relative to 1975 using estimates and their standard errors of model 1 (mean fledging weight = 15.5g). A simple linear regression is fitted through the estimates.

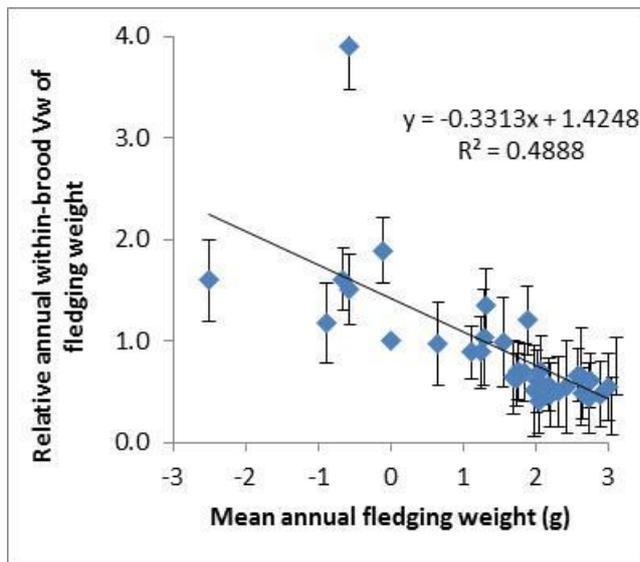


Figure 4. The predicted effect of clutch size on fledging weight (panel A) and its within-brood V_w across years using estimates of model 1 (panel B).

