**Supplement to "Selection on an extreme-yet-conserved larval life-history strategy in a tapeworm"**

# Contents



# Statistical models and decomposing trait variation

Over 4 generations of selection on the larval developmental rate of *Schistocephalus solidus*, eleven parasite and host traits were recorded. The factors shaping these traits – like starting conditions, environments, and parasite genes – were assessed with several generalized linear mixed models (Fig. S1). A 'base' model for each trait included starting conditions and shared environments. Starting conditions were fixed effects and encompassed mainly host characteristics like copepod stage or stickleback size and sex, but also parasite inbreeding coefficient (Hedrick and Kalinowski 2000). Shared spatiotemporal environments, like the same block of copepod infections or the same fish tank, were included as random effects. The base models for each trait are summarized in Table 1. For host traits like survival an additional base model was fitted that included infection and its interactions with other starting conditions. This tested whether the effects of starting conditions were dependent on *S. solidus* infection, e.g. if host size affected host survival differently for uninfected and infected hosts.

Genetic effects were assessed by adding three different terms to the base models. Adding the interaction between generation and selection line tested responses to selection (Figs. 3-5). However, this model does not account for the genetic variance within lines, so it was not considered when decomposing trait variances. By contrast, including either parasite sibship or the parasite pedigree as random effects explicitly estimated additive genetic variance and thus trait heritability. For four traits with sufficient sample sizes (copepod infection success, cercomer presence 9 dpi, procercoid size, and copepod survival 13 dpi), 'dam' was also added as a random effect to check whether offspring from the two mothers/fathers within a sibship differed. Thus, depending on the trait, variance could be explained by starting conditions (fixed effects like host size), shared environments (infection block or tank), genetics (sibship or pedigree), or parental effects (dam within a sibship). The fraction of variance explained by different terms was estimated according to Nakagawa and Schielzeth (2013) and compared among models 1, 2a, 2b, 3a, and 3b in Fig. S1.

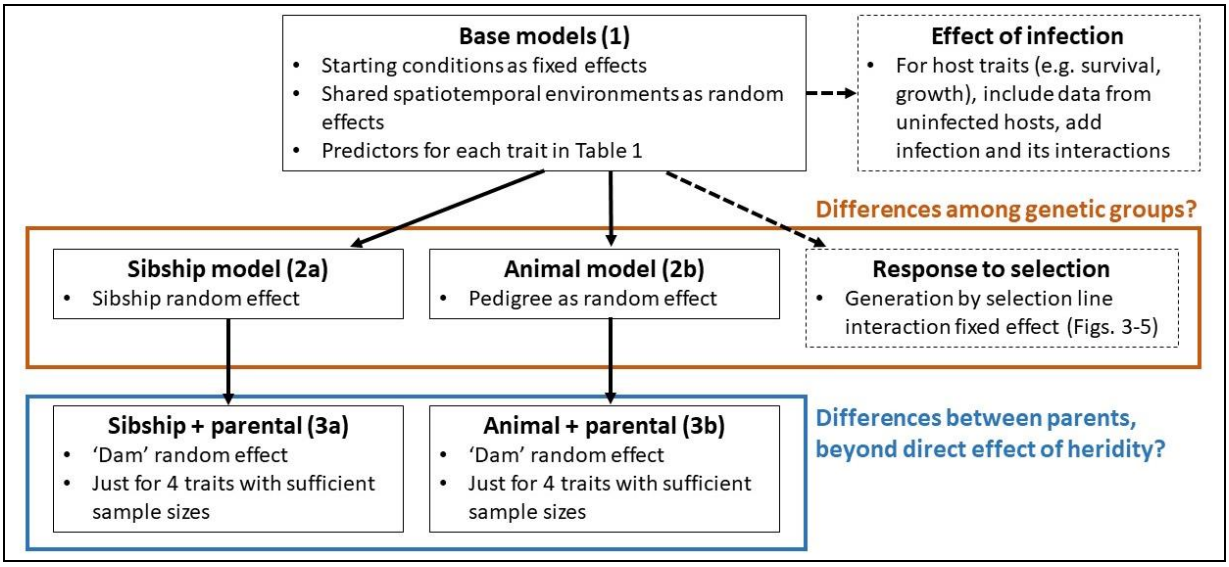


Figure S1 – Modelling approach. A series of mixed models were fitted to determine which factors shaped 11 parasite and host traits. For each trait, a 'base' model included starting conditions and environmental effects (Table 1). The impact of infection on host traits like survival was assessed with an additional model including uninfected hosts. Otherwise, models were restricted to infected hosts to quantify the effect of parasite genetics. Genetic effects were evaluated by adding parasite sibship, pedigree, or selection line to the base model, though the selection line model was not considered when partitioning trait variance (nor was the other model denoted with dashed lines). Finally, parental effects were assessed for four traits with sufficient sample sizes (>6 per dam).

Starting conditions explained at least some variation for all traits (i.e. the lower bound of the 95% credible interval (CI) was above 0.5% and thus considered non-zero), albeit very little for infection (~1%) and developmental success in copepods (2-4%) (Fig. S2A). Traits from the fish stage of the life cycle were more strongly affected by starting conditions. For example, over 75% of the variance in plerocercoid mass was explained by factors like initial fish size (Fig. S5) and age post infection (Fig. S6), whereas only 13% of the variance in procercoid size was explained by initial copepod stage (Fig. S3). Environmental variance components were non-zero for all traits except larval developmental success and fecundity, and they tended to be larger at the fish than copepod stage, e.g. fish survival differed more among tanks than copepod survival differed among infection blocks (Fig. S2B). The variance attributed to starting conditions and shared environments differed little among models, suggesting these effects were not confounded with genetic or parental effects.

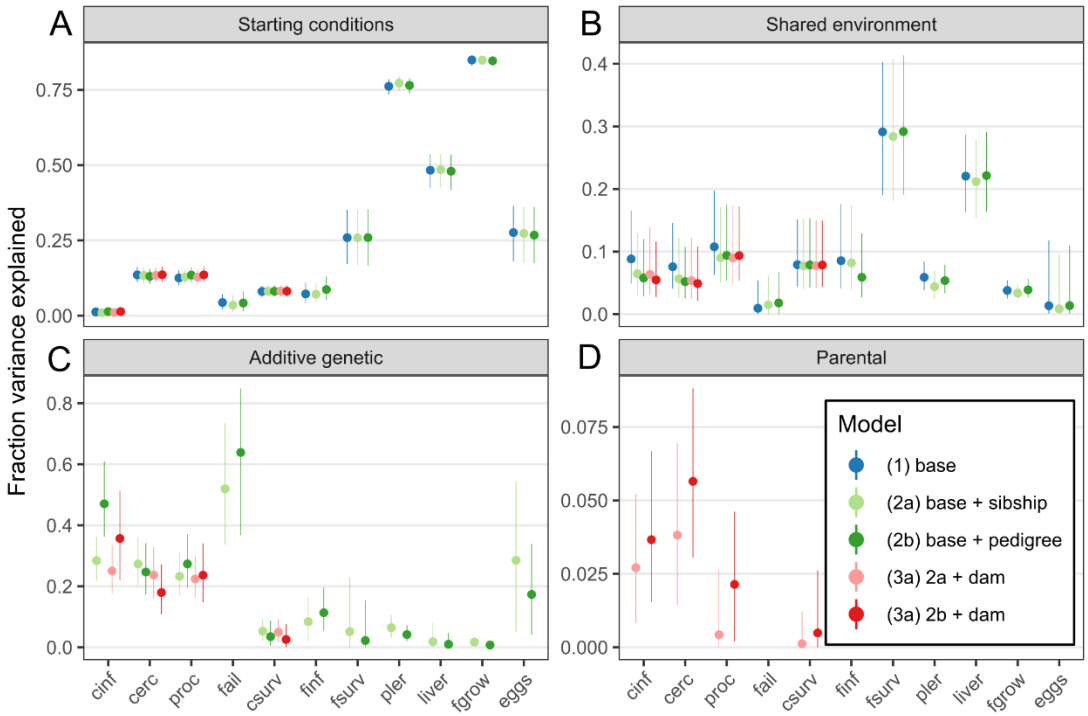


Fig. S2 – The fraction of trait variance explained by (A) starting conditions (fixed effects in Table 1), (B) shared spatiotemporal environments (copepod block, fish tank), (C) additive genetic variance, and (D) parental effects. Since full siblings share half their alleles, the among-sibship variance component was doubled to estimate additive genetic variance. The lower left panel gives heritabilities (i.e. the fraction of phenotypic variance explained by additive genetic effects). Eleven traits throughout the parasite life cycle were evaluated (cinf: copepod infectivity, cerc: cercomer presence 9 dpi, proc: procercoid size 13 dpi, fail: failed development 13 dpi, csurv: copepod survival 13 dpi, finf: stickleback infectivity, fsurv: fish survival 74 dpi, pler: plerocercoid mass, liver: fish liver mass, fgrow: fish growth, eggs: fecundity from days 2-7 days of in vitro breeding). Bars represent 95% credible intervals.

## The effects of starting conditions

Factors like host size or sex were assumed to affect parasite life-history traits, but they were not a main interest. Their effects are presented for completeness (Fig. S3-S7). They were estimated with base models, though other models yielded similar parameter estimates (Fig. S2A). The effect of a given variable was plotted holding other fixed model terms at these values: copepod stage – C2, inbreeding coefficient – 0, dpi in copepods – 14, fish mass (g) at exposure – 0.379, fish sex – 0 (males and females were scored as -0.5 and 0.5, respectively), dpi in sticklebacks – 86.

Larger copepods were characterized by lower infection rates (Fig. S3A) and higher survival rates (Fig. S3E). Parasites in larger copepods had slower development (Fig. S3B) but larger body sizes (Fig. S3C). Only the smallest infected copepods had higher mortality than uninfected copepods (Fig. S3E). These trends are in line with previous studies on *S. solidus* infectivity (Van Der Veen and Kurtz 2002; Van Der Veen 2003), growth and development (Wedekind 1997; Benesh 2010a), and virulence (Michaud et al. 2006; Benesh 2010b).

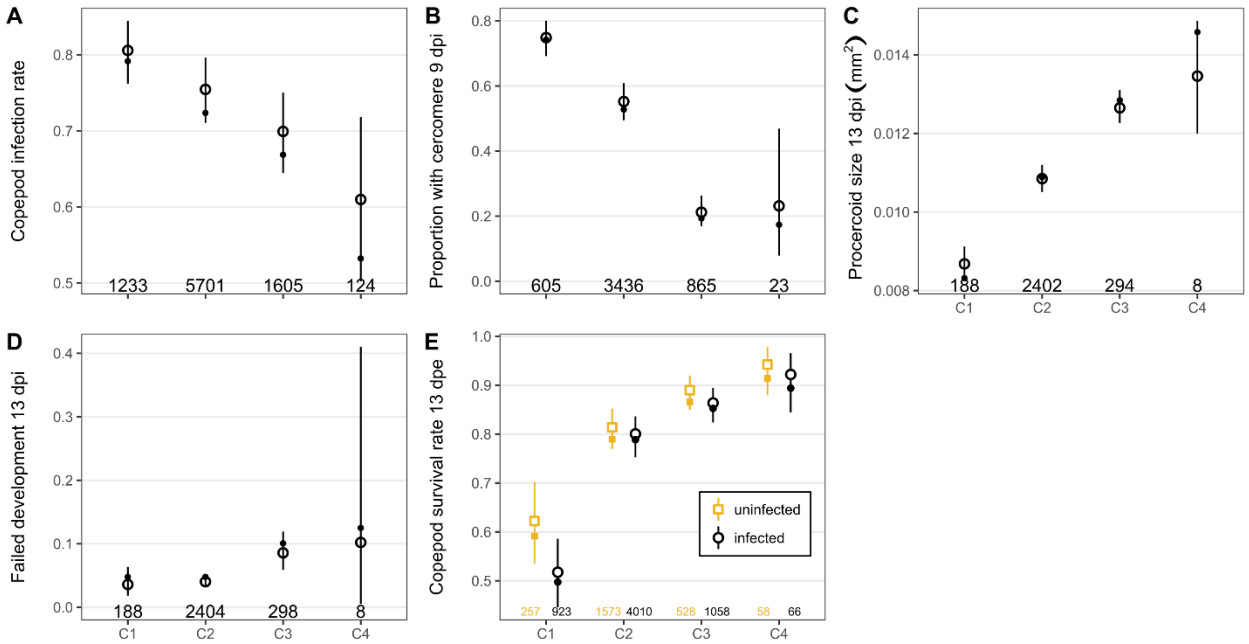


Figure S3 – The effect of copepod stage on (A) tapeworm infection rate, (B) development, (C) growth, and (D) developmental defects, as well as the (E) survival rate of infected and uninfected copepods 13 days post exposure (dpe). Solid points are unadjusted means, whereas the open points and bars (95% credible interval) are model-estimated means that account for starting conditions and shared environments (see Table 1). Sample sizes per group are given along the plots' bottom edge.

Two characteristics of sticklebacks were unstandardized in the experiment: sex and initial size. First, I consider sex. Male sticklebacks were slightly more susceptible to infection than females (Fig. S4A), but this should be interpreted cautiously, as other experimental infections have not yielded this sex difference (Barber and Svensson 2003; Benesh et al. 2012) and sex-biased infection rates are not observed in field samples (Pennycuick 1971). Male fish harbored marginally smaller worms (Fig. S4C) that then produced fewer eggs (~10,000 [95% CI: 2440 – -22714]), though the trend was not significant (Fig. S4E). Regardless of infection, male fish had smaller livers and grew less than females (Fig. S4D, E).

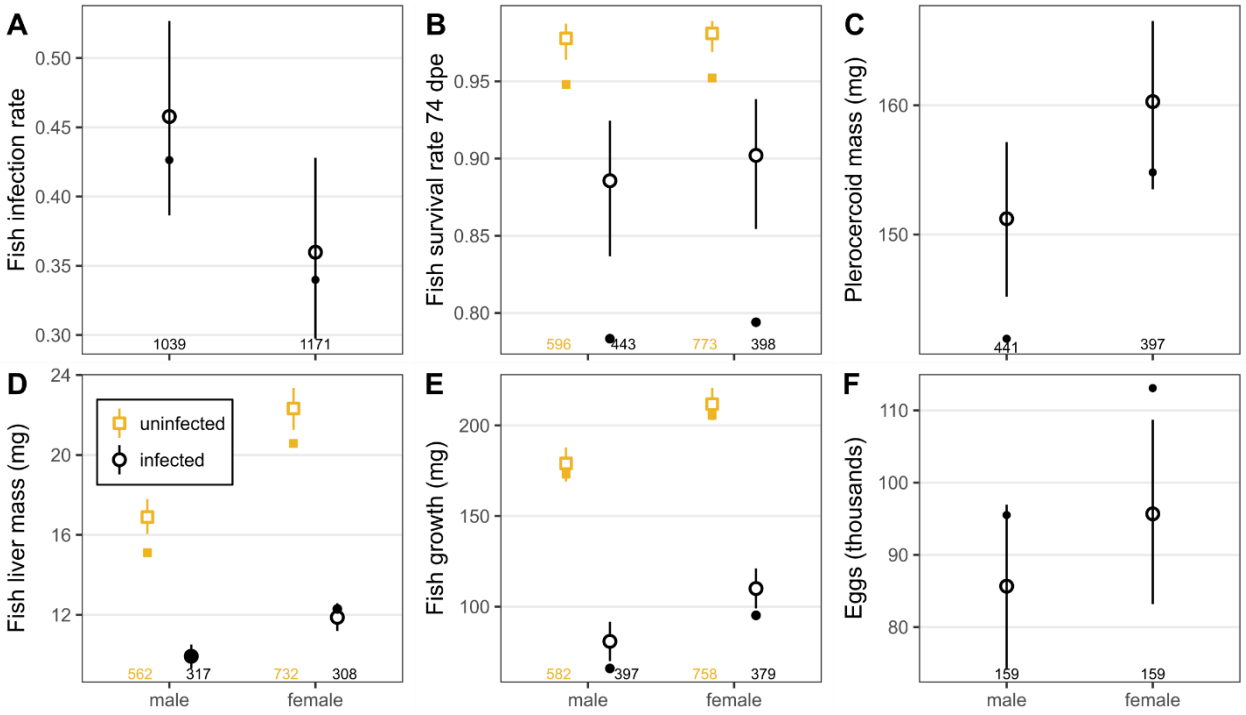


Figure S4 – The effect of stickleback sex on (A) tapeworm infection rate, (B) fish survival rate 74 days post exposure (dpe), (C) tapeworm body mass, (D) fish liver mass at dissection, (E) fish growth, and (F) tapeworm fecundity. Solid points are unadjusted means, whereas the open points and bars (95% credible interval) are model-estimated means that account for starting conditions and shared environments (see Table 1). Sample sizes per group are along the plots' bottom edge.

Smaller sticklebacks were not more susceptible to infection (Fig. S5A), but they had lower survival when infected (Fig. S5B). Parasites grew larger and produced more eggs when they infected larger sticklebacks (Fig. S5C, F), at least in part because larger fish were characterized by larger livers and more somatic growth (Fig. S5D,E) (Barber 2005). Also, as expected, both worm and fish growth increased with time post exposure (Fig. S6A, C).

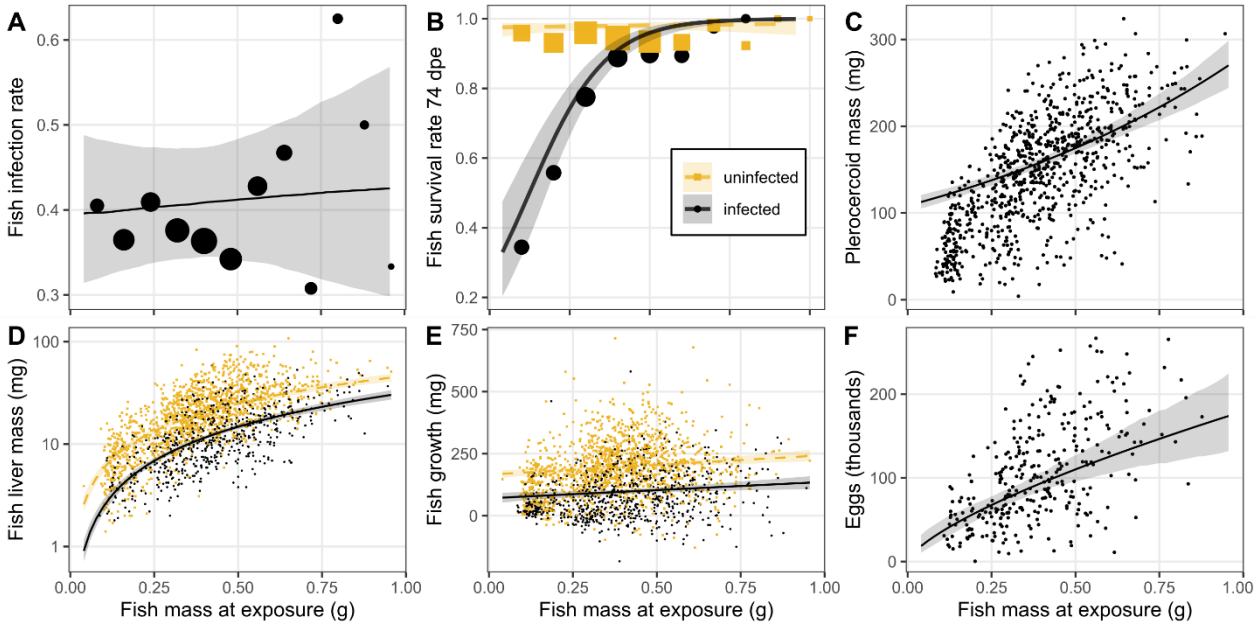


Figure S5 – The effect of stickleback mass at exposure on (A) tapeworm infection rate, (B) fish survival rate 74 days post exposure (dpe), (C) tapeworm body mass, (D) fish liver mass at dissection, (E) fish growth, and (F) tapeworm fecundity. In (A) and (B), fish masses were binned to calculation proportions, with point diameter proportional to the number of fish of a given size. Relationships and 95% credible intervals were estimated holding other variables, like fish sex and days post infection, constant.

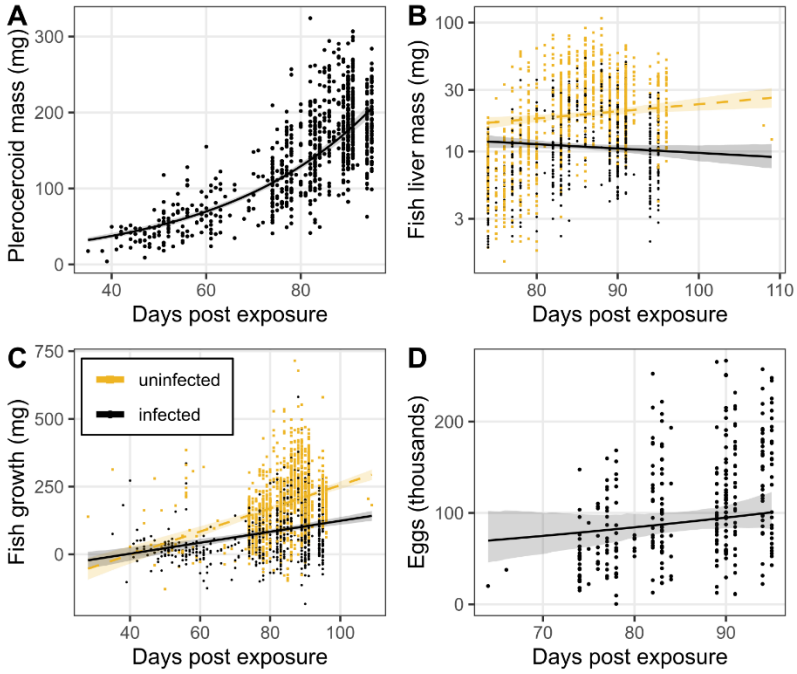


Figure S6 – The effect of time (days after fish exposure) on (A) tapeworm body mass, (B) fish liver mass at dissection, (C) fish growth, and (D) tapeworm fecundity. Relationships and 95% credible intervals were estimated holding other variables, like fish size and sex, constant.

Most parasites were outbred (Fig S7M), but some resulted from breeding full siblings (f = 0.25), in one case for consecutive generations (f = 0.375). The inbreeding coefficient was included in the base model. Inbred worms had significantly lower infection rates in both hosts (Fig. S7A, F), slower and more often failed development in copepods (Fig. S7B, D), less growth in both hosts (Fig. S7C, H), and lower fecundity and hatching (albeit non-significant, Fig. S7K, L). These trends are consistent with the inbreeding depression noted in previous studies (Christen et al. 2002; Christen and Milinski 2003), particularly in this *S. solidus* population (Benesh et al. 2014).

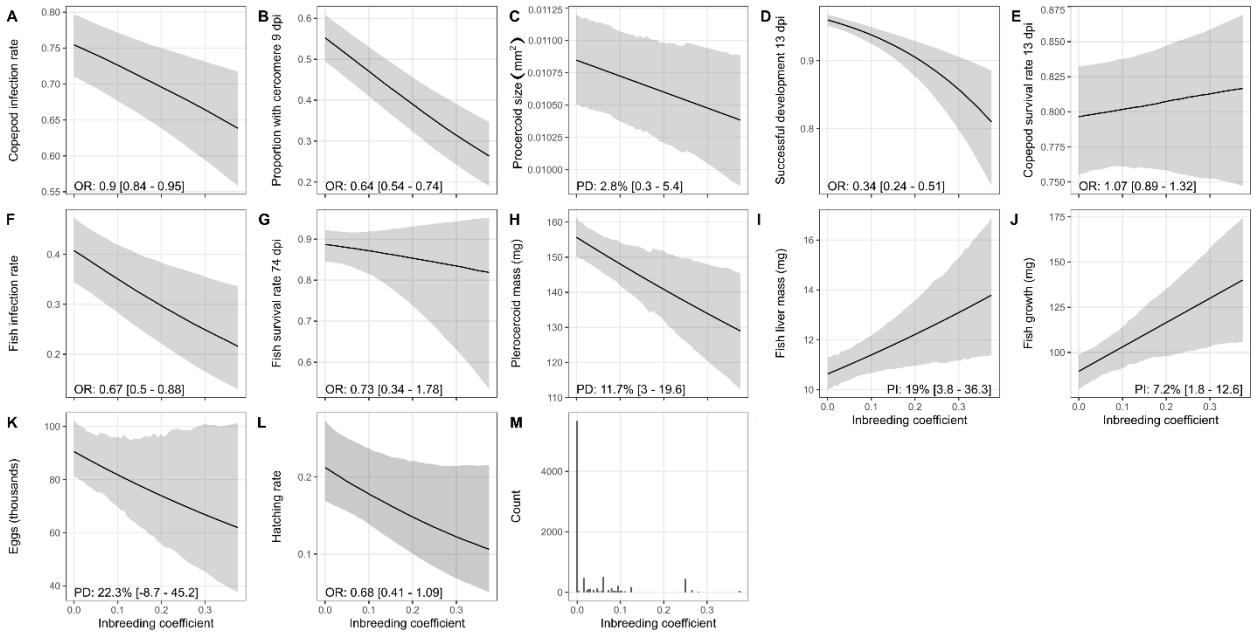


Figure S7 – The effect of parasite inbreeding on traits of parasites (A-D, F, H, K-L) and of infected hosts (E, G, I-J), as well as the distribution of inbreeding coefficients (M). Inbreeding coefficients, *f*, were calculated from the parasite pedigree assuming outbred founders. Lines and shading depict model-estimated relationships and 95% credible intervals when holding other variables constant (see Table 1). For binary traits, the odds ratio (OR) compares outbred parasites (*f* = 0) to offspring from breeding between full siblings (*f* = 0.25). For continuous variables, the percent decrease or increase (PD or PI) in the trait with full-sib mating is given.

## Genetic and parental effects

Inbreeding depression in *S. solidus* is probably caused by deleterious recessive alleles (Benesh et al. 2014). Such alleles (and nonadditive genetic architectures generally) are common for traits closely related to fitness because dominant deleterious alleles are quickly eliminated by selection, leaving recessive alleles (Crnokrak and Roff 1995; Roff and Emerson 2006). If alleles are only partially recessive, or if they increase in frequency due to inbreeding, drift, or selection, then they can contribute to additive genetic variance (Van Buskirk and Willi 2006). This is not obviously the case in *S. solidus*, though, as the phenotypes of related selfed and outbred tapeworms are not correlated (Benesh et al. 2014), emphasizing that inbred relatives share sources of genetic variation with each other (rare recessives in homozygous form) that they do not share with their outbred relatives (Cockerham and Weir 1984; Abney et al. 2000; Moorad and Wade 2005). Even if directional dominance 'hides' many detrimental alleles, life-history traits like growth and development represent large mutational targets (Houle 1998), and most new mutations are deleterious (Eyre-Walker and Keightley 2007). Thus, standing genetic variation may

often be skewed towards lower fitness, reducing adaptive potential (Frankham 1990; McGuigan and Blows 2009). To assess this, I investigated not only the magnitude of genetic variance, but also its distribution.

Table S1 – Heritabilities for host and parasite traits. Heritability is the fraction of phenotypic variation attributed to additive genetic effects ($V_g$), and $V_g$ was estimated with sibship or pedigree models (Fig. S1). Since full siblings share half their alleles, the among-sibship variance component was doubled to estimate $V_g$. Besides genetics, trait variance was, depending on the trait, partitioned into the effects of starting conditions (fixed effects like host size, $V_f$), shared environments (infection block or tank, $V_e$), and parent (dam within a sibship, $V_m$), as well as the unexplained residual variance ($V_r$). Heritability was calculated with and without the variance attributed to unstandardized starting conditions like host size (i.e. $V_g/(V_g + V_f + V_e + V_m + V_r)$ vs $V_g/(V_f + V_e + V_m + V_r)$), the latter of which can be thought of as the heritability when all parasites infect average hosts.

| | | Sibship model | | Animal model | |
|---|---|---|---|---|---|
| Trait | n | $h^2$, overall | $h^2$, average host | $h^2$, overall | $h^2$, average host |
| Infectivity to copepods | 8663 | 0.25 [0.18-0.33] | 0.25 [0.18-0.34] | 0.36 [0.22-0.51] | 0.36 [0.22-0.52] |
| Cercomere 9 dpi | 4929 | 0.24 [0.16-0.33] | 0.27 [0.19-0.38] | 0.18 [0.11-0.27] | 0.21 [0.13-0.31] |
| Procercoid size 13 dpi | 2892 | 0.22 [0.16-0.30] | 0.26 [0.19-0.34] | 0.24 [0.15-0.34] | 0.27 [0.17-0.39] |
| Fully developed 13 dpi | 2898 | 0.52 [0.34-0.73] | 0.54 [0.35-0.76] | 0.64 [0.37-0.85] | 0.67 [0.39-0.89] |
| Copepod survival 13 dpi | 6057 | 0.05 [0.02-0.09] | 0.05 [0.02-0.10] | 0.03 [0.01-0.09] | 0.03 [0-0.08] |
| Infectivity to sticklebacks | 2230 | 0.08 [0.02-0.17] | 0.09 [0.02-0.18] | 0.11 [0.05-0.20] | 0.12 [0.06-0.22] |
| Stickleback survival 74 dpi | 854 | 0.05 [0-0.23] | 0.07 [0-0.31] | 0.02 [0-0.15] | 0.03 [0-0.20] |
| Ln plerocercoid mass | 799 | 0.06 [0.03-0.11] | 0.28 [0.14-0.45] | 0.04 [0.02-0.07] | 0.18 [0.09-0.30] |
| Relative liver mass | 625 | 0.02 [0-0.08] | 0.04 [0-0.16] | 0.01 [0-0.05] | 0.02 [0-0.09] |
| Stickleback growth | 781 | 0.02 [0.01-0.03] | 0.11 [0.05-0.22] | 0.01 [0-0.02] | 0.05 [0.02-0.10] |
| Ln fecundity | 318 | 0.29 [0.05-0.54] | 0.39 [0.07-0.75] | 0.17 [0.04-0.34] | 0.24 [0.06-0.44] |

Host traits had low heritability; heritabilities were not significantly different from zero (i.e. the lower bound of the 95% CI exceeded 0.005) for stickleback survival and liver mass, and they were just ~0.05 for copepod survival and 0.02 for fish growth (Fig. S2C; Table S1). All other traits exhibited non-zero heritability, regardless of whether heritabilities were based on the total phenotypic variance or phenotypic variance after accounting for factors like host size (Table S1). Additive genetic variances were estimated from among-sibship variance or directly from animal models (Fig. S1). Heritabilities from sibship and animal models were similar with overlapping CIs (Fig. S2C; Table S1) with the exception that heritability for copepod infectivity was estimated to be ~11 percent points higher by the animal model

(0.25 vs 0.36). The similarity is notable because full-sib heritabilities can be inflated by dominance, maternal, or common environment effects. Bias through common environment effects should be negligible because parasites develop in separate hosts, related parasites were often in different environments (e.g. sibling tapeworms in fish kept in different tanks), and the effect of such environments was included in the models. Maternal effects, or more precisely parental effects (as maternal vs paternal effects cannot be distinguished), were also modelled. Consistent with a previous study (Benesh 2013), the two maternal clutches within sibships exhibited small but significant differences in copepod infectivity and cercomer presence (3-4% of the trait variance; Fig. S2D), which slightly reduced heritability estimates (Fig. S2C). Because factors inflating differences among sibships were explicitly modelled, and because sibship- and pedigree-based models yielded similar results, there is little reason to suspect that sibship-derived estimates of genetic variance were biased.

Consequently, model-estimated sibship means (i.e. random effects) were plotted to explore the distribution of genetic effects, specifically whether genetic variance was skewed towards lower fitness. The response of developmental rate to selection was symmetric (Fig. 3A,B). By contrast, the distribution of sibship means was significantly negatively skewed for copepod infectivity, developmental success in copepods, stickleback survival, and fecundity (Fig. S8). This suggests that most sibships had high fitness, but some had noticeably lower fitness, indicating that much of the genetic variance in these traits is for reduced fitness. By contrast, just one trait, fish growth, exhibited positively skewed sibship means (Fig. S8), which is probably irrelevant, as the skew was weak and parasite sibships differed little in their impact on fish growth (i.e. heritability was only 0.02; Fig. S2C). Skewed random effects violate the distributional assumptions of mixed models, but this need not invalidate conclusions, as parameter estimates are remarkably robust to such violations (Schielzeth et al. 2020).

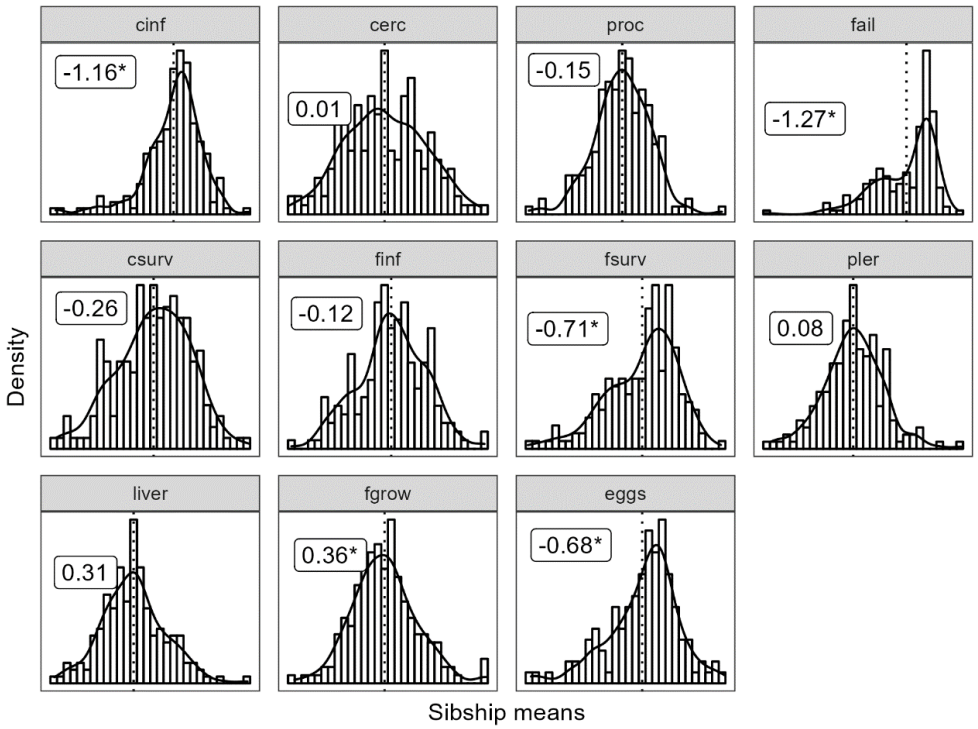


Fig. S8 – The distribution of model-estimated sibship means for 11 traits (trait abbreviations as in Fig. S2). The vertical dashed line is the expected value after accounting for starting conditions

and shared environments. Skewness is stated. For fecundity, one extreme outlier with low fecundity was removed before the skewness calculation.

When development is canalized, a phenotype is consistently produced despite environmental and genetic perturbations. If larval development is robust to mutations, e.g. due to directional dominance (Fig. S7) or other non-additive effects, then genetic variation could silently accumulate. This cryptic genetic variation could be released (decanalization) through strong directional selection. In this scenario, the distribution of sibship means would be broader and possibly more skewed in the selected lines. To test this, the sibship models (2a and 3a) were re-fit with separate among-sibship variance components for the three selection lines. Among-sibship variance was greater in the selected lines for copepod infectivity, developmental success in copepods, and fecundity (Fig. S9B). This was mostly attributable to strong negative skew in the selected lines, relative to the controls (Fig. S9A). These trends are consistent with the idea that directional selection on larval developmental rate releases cryptic genetic variation, much of it detrimental.

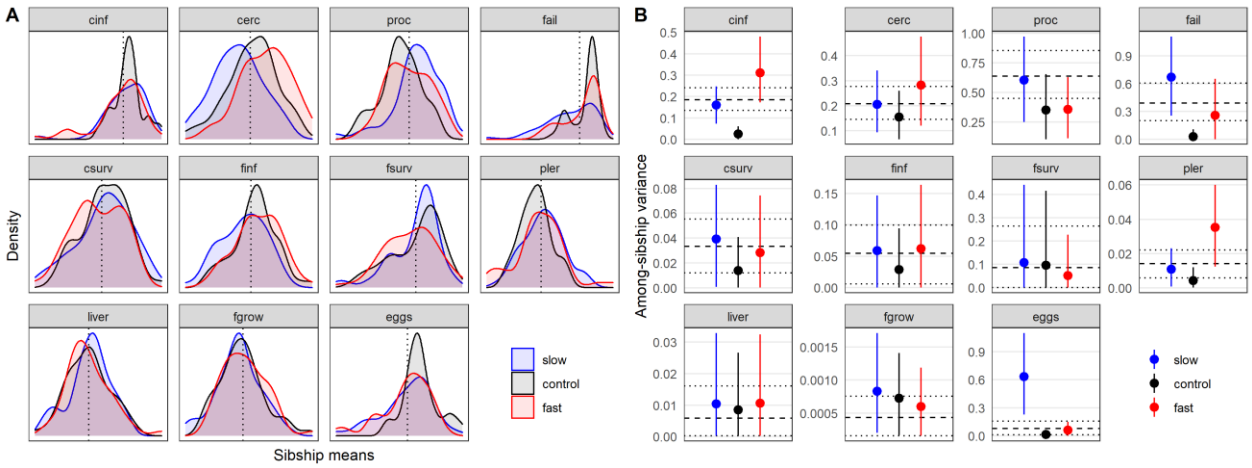


Fig. S9 – (A) Density plots for sibship averages and (B) among-sibship variances for three selection lines. In (A) the vertical dashed line is the expected value after accounting for starting conditions and shared environments. In (B) the horizontal dashed and dotted lines show the overall among-sibship variance irrespective of selection line (± 95% credible interval). Error bars represent 95% credible intervals. Trait abbreviations as in Fig. S2.

## What drives trait covariance?

Trait correlations were expected. For instance, cercomer presence and procercoid size are both related to infectivity to sticklebacks (Benesh and Hafer 2012; Benesh et al. 2012). Such correlations can be caused by shared environments (e.g. parasites in block A develop faster and are more infective to fish than those in block B), shared genes (e.g. genotypes that develop faster have higher infectivity), shared parents (e.g. offspring from one parent develop faster and have higher infectivity than those from the other parent in the sibship), or something else (e.g. parasites in “good” copepods may develop faster and have higher infectivity). These environmental, genetic, parental, and residual correlations were estimated with bivariate mixed models. Bivariate models with sibship were computationally faster and easier to fit than bivariate animal models, so they were used to estimate genetic correlations (i.e. bivariate versions of models 2a or 3a in Fig S2). Genetic correlations were presented in the main text

(Fig. 6). Here, the other correlations are given, specifically across environments (Fig. S10), parents (Fig. S11), and individuals (Fig. S12).

Correlations across copepod blocks were mostly positive – some blocks were characterized by increased rates of copepod infection, larval developmental, copepod survival, and fish infection – though they were not significant (Fig. S10A). Correlations across fish tanks were clearer; the tanks with larger plerocercoids were also those in which fish grew more and had bigger livers (better condition) (Fig. S10B).

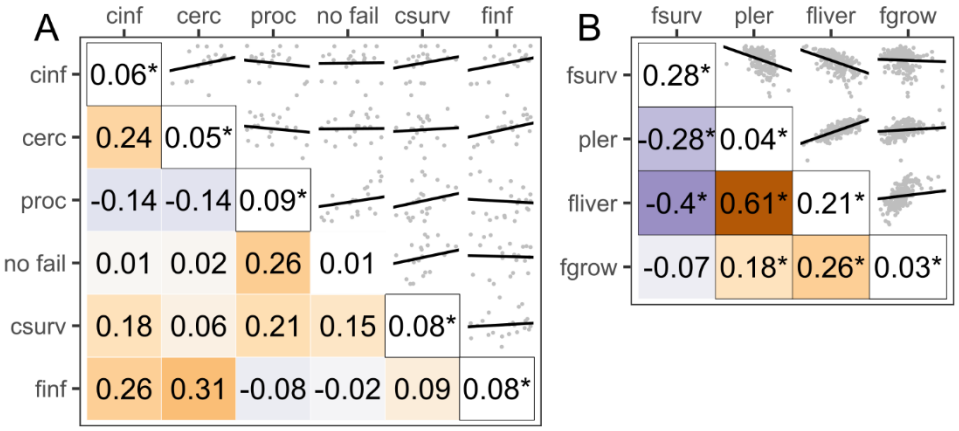


Fig. S10 – Environmental trait correlations (E-matrix) among (A) copepod infection blocks (n = 22) and (B) fish tanks (n = 199). Trait covariances were estimated with bivariate mixed models in which spatiotemporal environment was a random effect. The estimated random effects (i.e. block/tank averages after accounting for fixed effects like host size; Table 1) are plotted above the diagonal; solid lines are univariate regressions (i.e. covariance / variance in x). Values below the diagonal are correlation coefficients. Values along the diagonal are the fraction of trait variation explained by the environment (see also Fig. S2). Stars indicate variances and covariances that differed significantly from zero.

Just two traits exhibited small but significant 'parental' effects: copepod infectivity and cercomer presence (Fig. S11). When the eggs from one parent within a sibship had higher copepod infection rates, they also tended to have faster development in copepods (Fig. S11), though this correlation was not quite significant (0.21, 95% CI: -0.04 – 0.44).

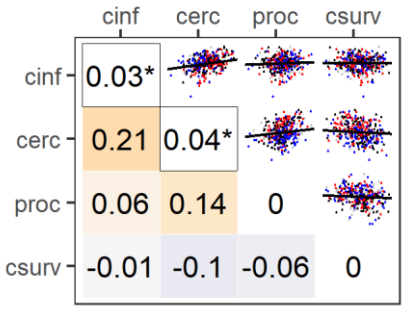


Fig. S11 – Parental trait correlations. Parental effects were only estimated for 4 traits in which sample sizes per parent were sufficient. Plot elements as in Fig. S10, except that point color in the scatterplots reflects selection line.

After accounting for all other effects, some traits were still correlated at the level of parasite individuals. Such residual correlations were undefined for many trait combinations. For instance, residual covariances involving copepod infectivity were not estimable because the other traits were only measured for individuals that successfully infected copepods (i.e. the residuals were not unique).

Several residual correlations at the copepod stage were positive; all else equal, parasites that developed faster also grew larger, were in copepods with higher survival, and were more likely to infect sticklebacks (Fig. S12). Similarly, plerocercoids that were larger than expected tended to be from fish that grew more and had larger relative liver masses. These worms then also produced more eggs (Fig. S12). Unexpectedly, tapeworms had lower residual fecundity when they developed faster than predicted, which is the reverse of the genetic correlation (Fig. 6). Discordant genetic and phenotypic correlations are relatively rare (Cheverud 1988; Roff 1996; Hadfield et al. 2007), and this trend was at least partly driven by low-fecundity outliers (Fig. S12). Nonetheless, it begs the question of whether something (besides copepod stage or parasite genes) accelerates parasite development in copepods but with a long-term cost (e.g. host immune attacks; Kurtz and Franz 2003)? Tradeoffs between tissue differentiation and growth; Arendt 2000)?).

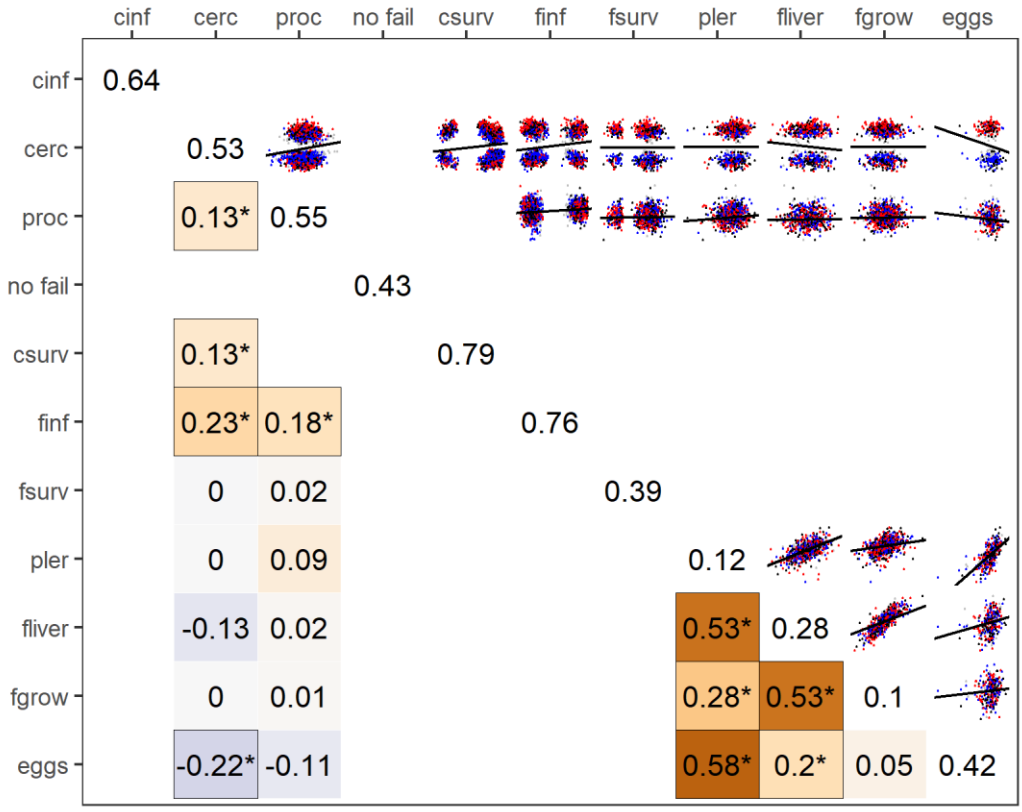


Fig. S12 – Residual trait correlations after accounting for starting conditions, shared environments, genes, and parents. Residual correlations could not be estimated for all trait combinations, especially for binary traits. Plot elements as in Fig. S11. The diagonal is the fraction of phenotypic variance that was not accounted for by other factors.

More broadly, variability among individual hosts may be an important determinant of parasite performance, given the high levels of residual variance in several traits (diagonal of Fig. S12). Higher quality hosts (e.g. copepods with higher survival or fish in better condition, relative to expectations)

tended to harbor bigger, more infectious, and more fecund parasites, yet host traits like survival and growth were little affected by parasite genotype (heritabilities were ≤ 0.05). Even for parasite traits with significant heritability, like cercomer presence and fish infection (Fig. 6), performance was impacted more by unexplained variation among individuals than parasite genotype. For instance, the bivariate model predicted that a fast-developing genotype ($10^{th}$ percentile) had a 6% points higher infection probability than an average genotype (47 vs 41% at 14 dpi), but a parasite that developed much faster than expected ($10^{th}$ percentile of the residuals) had an 11% higher infection rate (52 vs 41%). Perhaps parasites with fast residual development infected copepods that were in better condition or maybe the coracidia was consumed quickly before its resources were depleted. In any case, the contingency of fitness on non-genetic factors will slow parasite adaptation, even if fast-developing genotypes have a fitness advantage on average (see next section).

## Comparing lifetime reproductive success between larval genotypes

Parasites selected for rapid larval development had, on average, higher infectivity (Fig. 3C) but also lower fecundity (Fig. 4F). So, do fast-developing genotypes have a fitness advantage over average ones? Genotypes' lifetime reproductive success was compared two ways. First, I compared the fitness of parasites in the three selection lines in the final generation, which assumes the lines differ genetically by the end of the experiment. Although true on average, this ignores substantial genetic variation within lines (Fig. S9). So, after comparing lines, I estimated the phenotypes of fast- and slow-developing genotypes (specifically the top and bottom 10%) based on traits' genetic covariance with cercomer presence in bivariate models (Fig. 6). Trait estimates were used to calculate expected lifetime reproductive success.

As is typical in helminths (e.g. Skorping et al. 1991; Trouvé et al. 1998), the fecundity of *S. solidus* increases with body size (Schärer et al. 2001), so fitness (f) may take the form:

$$f(t)= P(t) \cdot b \cdot W(t) \qquad (1),$$

where t is time, P(t) is the probability of surviving until t, W(t) is growth in body size, and b relates fecundity to body size; $b \cdot W$ is thus expected egg production (Parker et al. 2003; Benesh et al. 2013). Both parasite survival and fecundity were measured in the experiment, enabling rough fitness estimates. Survival and reproduction are considered in turn.

Survival to maturity in *S. solidus* depends on encountering, infecting, and developing in three successive hosts. Encounter rates could not be estimated from the data, but if we assume that parasites always encounter (i.e. are consumed by) the next host, then we can calculate the probability that an individual coracidium survives to reproductive maturity as:

$$P(t) = p_{ci} \cdot p_{cs}(t_c) \cdot p_{fi}(t_c) \cdot p_{fs}(t_f) \cdot p_{bi}(t_f) \qquad (2),$$

where $p_{ci}$ is the copepod infection rate, $p_{cs}$ is the probability of copepod survival over time $t_c$, $p_{fi}$ is infection rate in fish as a function of time in copepods $t_c$, and $p_{fs}$ is the probability of stickleback survival over time $t_f$, and $p_{bi}$ is the probability of infecting birds after a given amount of time in fish $t_f$. Survival probabilities ($p_{cs}$ and $p_{fs}$) were calculated as $e^{-mt}$ where m is the mortality rate and t is time in the host; mortality rates were based on survival until 13 and 74 dpi in copepods and sticklebacks, respectively. The probability of infecting sticklebacks $p_{fi}$ was estimated by fitting an asymptotic curve to stickleback infection rates at 12, 14, and 16 dpi (see equation 5 in Hammerschmidt et al. 2009). Finally, the probability to infect birds $p_{bi}$ depends on growth in *S. solidus* (Tierney and Crompton 1992);

plerocercoids below 50 mg have very low infection rates, whereas those above this threshold have infection rates of ~70% or more (i.e. $p_{bi} = 0$ for $W < 50$ and $p_{bi} = 0.7$ for $W > 50$).

Besides governing infectivity to birds, plerocercoid growth also determines fecundity. In the experiment, plerocercoid mass was estimated to change with time as $\ln(W) = w_0 + g \cdot t_f$, where $w_0$ is initial size, g is relative growth rate, and $t_f$ is time in fish (Fig. S6A shows this relationship overall, without distinguishing genotypes). This equation is sufficient for our fitness calculations, though growth rates will eventually decelerate as plerocercoid size becomes constrained by fish size. The plerocercoid size-fecundity relationship was obtained by replacing host traits with ln-transformed plerocercoid weight in the fecundity model. Since both fecundity and body size were ln-transformed, egg output was estimated as a power law (similar to the trend in Fig. S5F): $\ln(e_0) + b \cdot \ln(W)$, where $e_0$ is base egg production and b relates fecundity to body size. Combining the growth-age and the growth-fecundity relationships yields an estimate for how fecundity varies with time in sticklebacks:

$$\ln(b \cdot W(t_f)) = \ln(e_0) + b \cdot (w_0 + g \cdot t_f) \qquad (3)$$

The parameters needed to calculate survival (eq. 2) and reproduction (eq. 3) were taken from the models' posterior distributions, so uncertainty was incorporated into fitness calculations. Furthermore, the trait models control for starting conditions (like copepod and fish size; Figs S3-S7) and environmental effects (blocks and tanks), so fitness comparisons were made for 'average' hosts. In addition to absolute fitness (expected egg production), relative fitness was calculated as $f_{fast}/f_{avg}$.

The component functions used to estimate fitness for the selection lines are shown in Fig. S13. In the final generation, parasites from the fast line tended to infect fish at higher rates (Fig. S13B) but had lower infection rates in copepods (Fig. S13A) and lower fecundity (Fig. S13F). To evaluate whether these trends cancel out, the expected fitness when spending different amounts of time in copepods or sticklebacks was examined.

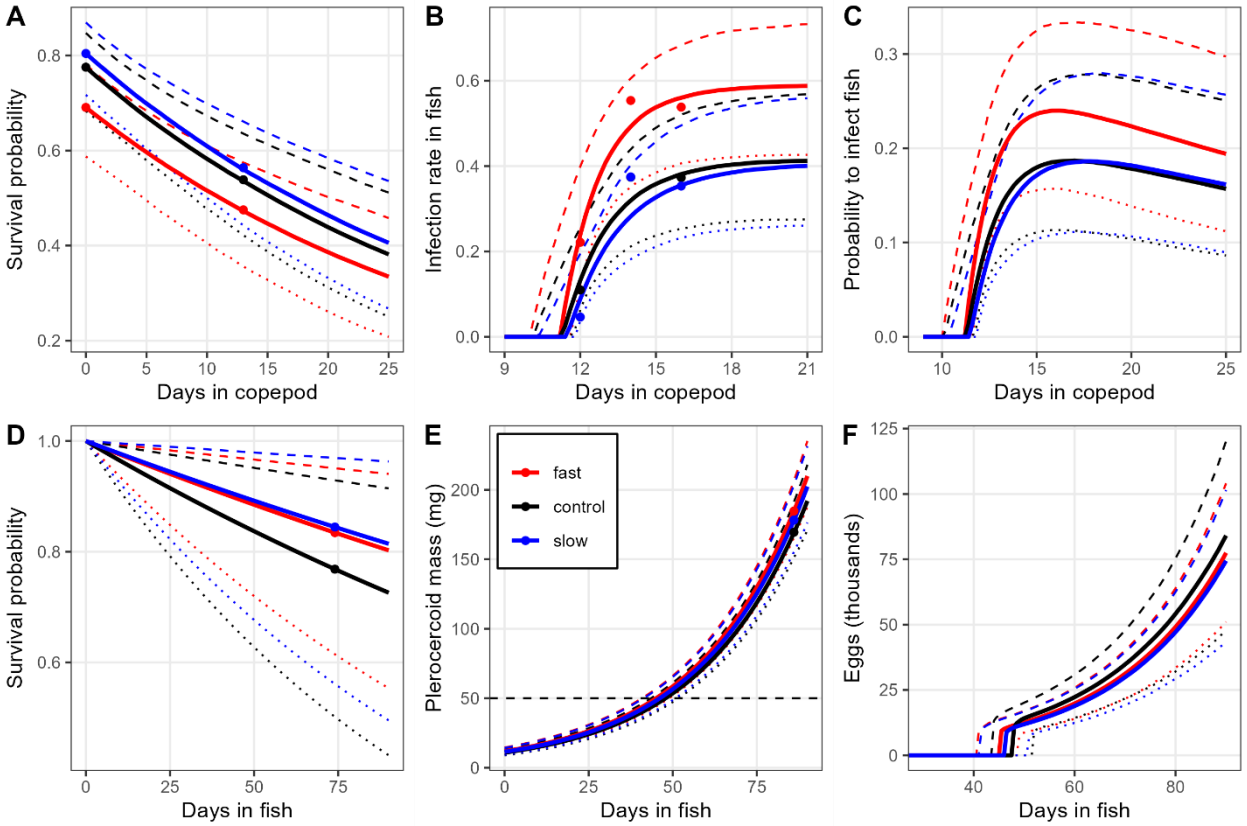


Fig. S13 – The functions used to estimate fitness for *S. solidus* selection lines. (A) The probability of infecting and surviving in copepods ($p_{ci} \cdot p_{cs}(t_c)$), (B) stickleback infection rates ($p_{fi}(t_c)$), (C) the probability that a coracidum survives until infecting a stickleback ($p_{ci} \cdot p_{cs}(t_c) \cdot p_{fi}(t_c)$), (D) stickleback survival over time ($p_{fs}(t_f)$), (E) plerocercoid growth (note the 50 mg threshold determining infectivity to birds, $p_{bi}$), and (F) average egg output given survival and growth in sticklebacks ($p_{fs}(t_f) \cdot p_{bi}(t_f) \cdot b \cdot W(t_f)$). Solid lines represent averages, whereas upper and lower credible intervals are given with dashed and dotted lines, respectively. Dots are means estimated for the selection lines in the final generation (see Figs. 3-5). Combining the functions in (C) and (F) yields the estimated lifetime reproductive success.

Across lines, estimates of lifetime reproductive success were highest after ~15 days in copepods (maximum infectivity; Fig. S14A,B) and after the maximum amount of time in fish (large, fecund sizes; Fig. S14C,D). The relative fitness difference between the fast and control lines was greatest when transmission to fish occurred earlier, e.g. the fast line had ~38% higher fitness than controls when transmitted 12 dpi vs ~12% when transmitted 16 dpi (Fig. S14C vs Fig. S14D). This confirms the expectation that rapid larval development is most advantageous when fish predation rates are high. The lines did not clearly differ in plerocercoid growth (Fig. S13E) or their impact on fish survival (Fig. S13D), so time spent in fish had little impact on relative fitness, e.g. the fast line had ~11-14% higher fitness than controls when reproducing after either 60 or 90 days in sticklebacks (Fig. S14A vs Fig. S14B).

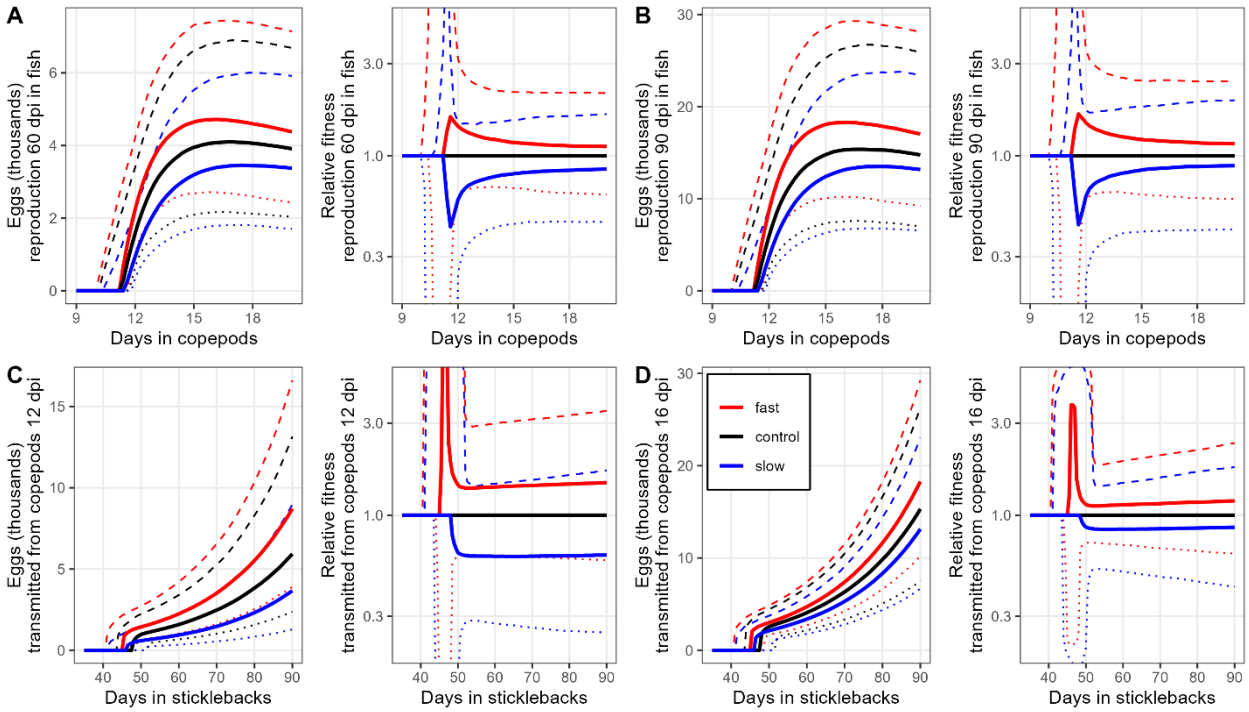


Fig. S14 – Absolute and relative fitness (left and right subplots) versus time in copepods (A and B) and time in sticklebacks (C and D). Early transmission from fish to birds (60 days post infection, dpi) and from copepods to fish (12 dpi) and from are shown in (A) and (C), whereas later transmission scenarios (90 and 16 dpi) are given in (B) and (D). Solid lines represent averages, whereas upper and lower credible intervals are given with dashed and dotted lines.

Some caution is warranted in interpreting the fitness differences among lines in Fig. S14 because, given the uncertainty in multiple parameter estimates, CIs for fitness were wide. Wide CIs may partly reflect error propagation; fitness calculations were based on univariate models (Table 1), so there was an implicit assumption that traits are independent (i.e. the posterior samples to calculate, say, infection rate were uncorrelated with those used to calculate, say, survival, which could inflate CIs). An additional caveat is that some differences between lines were unique to the last generation. For example, the fast line had lower mean copepod infection rates in the final generation, but not in the preceding generation (Fig. 4A). This result also contrasted with the estimated genetic correlation; genotypes characterized by fast larval development tended to have higher copepod infection rates (Fig. 6). Thus, fitness comparisons based on the genetic covariance among traits, regardless of the selection line, is an additional way to assess selection on fast-developing genotypes. I repeated the same fitness calculations but with trait values estimated from their genetic covariance with cercomer presence in bivariate models. Specifically, I calculated fitness for the top and bottom 10% of genotypes for larval developmental rate.

Fast-developing genotypes tended to have higher infection rates in copepods, higher infection rates in fish, larger plerocercoids, and higher fecundity; the only (non-significant) negative genetic correlation was with fish survival (Fig. 6). This negative correlation was not enough to reduce fitness below that of average genotypes. Like the line comparisons, the relative fitness advantage of fast genotypes was greatest when transmission to fish occurred earlier, e.g. ~36% higher fitness when transmitted 12 dpi vs ~21% higher when transmitted 16 dpi (Fig. S15C vs Fig. S15D). Unlike the line comparisons, the relative fitness of fast-developers also increased with time in sticklebacks (45% vs 23% higher when reproducing

90 vs 60 dpi in sticklebacks; Fig. S15A vs B), due to the positive covariance between larval development and fecundity.

Fitness effects appeared symmetric. For example, fast and slow genotypes had ~36% higher and ~30% lower fitness than average genotypes, respectively, when transmitted to fish 12 dpi (Fig. S15C). This symmetry is deceptive, though. Fitness calculations were made under the assumption that genetic variance was normally distributed, i.e. that the top and bottom 10% of genotypes differed from the average by the same amount (1.28 SD). However, genetic variances for several traits were negatively skewed (Fig. S8), implying that a high-fitness genotype is closer to the average than a low-fitness genotype. Thus, the fitness values in Fig. S15 probably overestimate the advantage of fast-developing genotypes and underestimate selection against slow-developing genotypes.

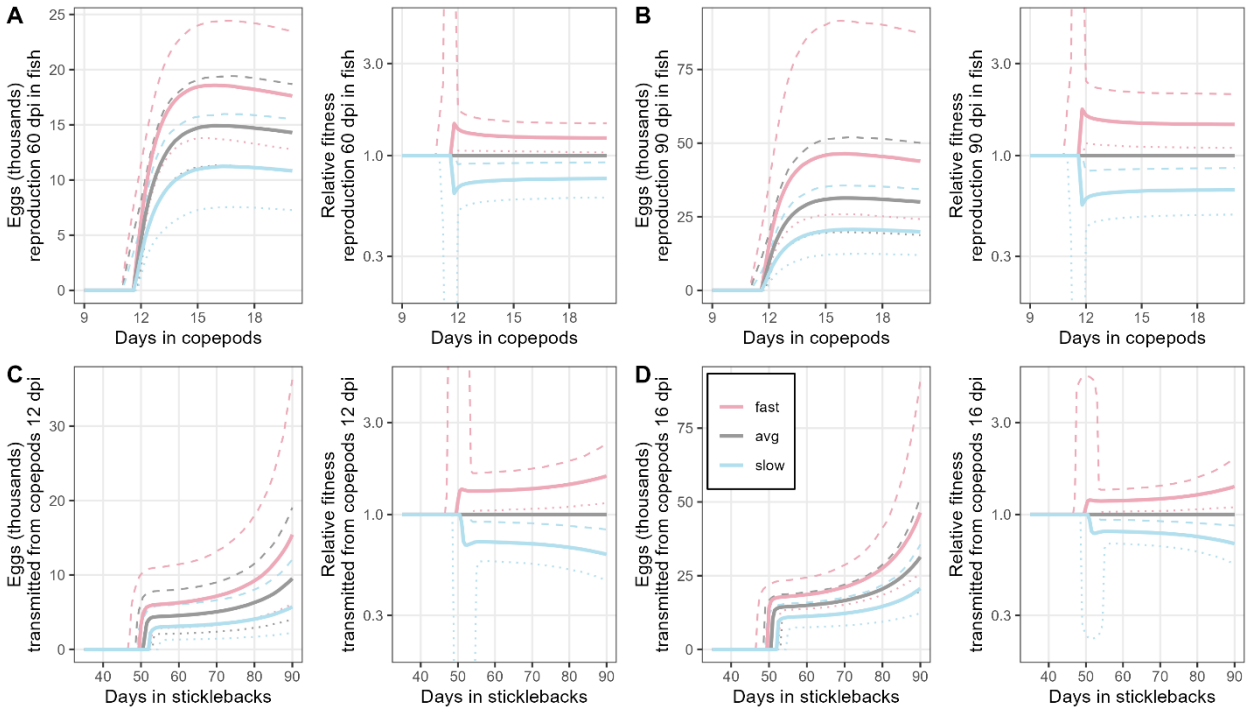


Fig. S15 – As in Fig. S14, except fitness is contrasted for fast- and slow-developing genotypes ($10^{th}$ percentile), as opposed to selection lines.

To roughly assess the bias caused by genetic asymmetry, I re-ran the calculations but assumed that fast-developing genotypes had similar copepod infectivity and fecundity to average genotypes (i.e. traits with negative skew in the selection lines; Fig. S9A). Accordingly, the relative fitness advantage of fast genotypes decreased, e.g. from 36 and 21% to 20 and 7% when transmitted to fish 12 and 16 dpi, respectively (Fig. S15C, D vs Fig. S16C, D), and it was not significant higher than average genotypes (i.e. the 95% CI for relative fitness overlapped 1). The modest advantage of fast genotypes depended almost entirely on their higher infectivity to fish; the genetic covariance between larval developmental rate and fish infectivity was positive but not quite significant, in part because there was comparatively little genetic variance for infectivity (Fig. 6).

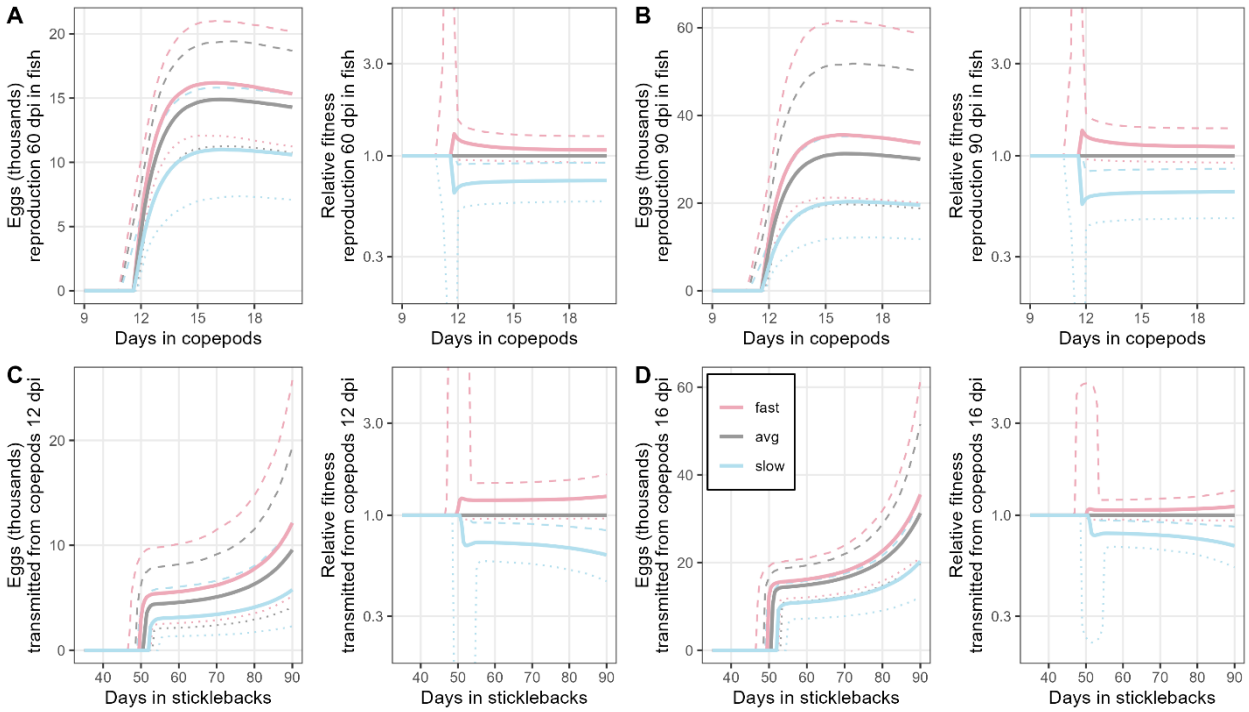


Fig. S16 – As in Fig. S15, except that fast-developing genotypes were assumed to have copepod infection rates and fecundities similar to average genotypes (see Fig. S9).

In sum, genotypes that developed faster in copepods seem to have a fitness advantage, especially when fish predation rates are high and parasite transmission from copepods occurs earlier. A constraint on evolving faster development, however, is genetic asymmetry. Directional selection exposed genetic variance for lower copepod infectivity, impaired development, and reduced fecundity (Fig. S8), all of which would slow down selection responses.

**Literature cited**


Abney, M., M. S. McPeek, and C. Ober. 2000. Estimation of variance components of quantitative traits in inbred populations. Am. J. Hum. Genet. 66:629–50.

Arendt, J. D. 2000. Allocation of cells to proliferation vs. differentiation and its consequences for growth and development. J. Exp. Zool. 288:219–34.

Barber, I. 2005. Parasites grow larger in faster growing fish hosts. Int. J. Parasitol. 35:137–43.

Barber, I., and P. A. Svensson. 2003. Effects of experimental *Schistocephalus solidus* infections on growth, morphology and sexual development of female three-spined sticklebacks, *Gasterosteus aculeatus*. Parasitology 126:359–367.

Benesh, D. P. 2010a. Developmental inflexibility of larval tapeworms in response to resource variation. Int. J. Parasitol. 40:487–97.

Benesh, D. P. 2010b. What are the evolutionary constraints on larval growth in a trophically transmitted parasite? Oecologia 162:599–608.

Benesh, D. P. 2013. Parental effects on the larval performance of a tapeworm in its copepod first host. J. Evol. Biol. 26:1625–1633.

Benesh, D. P., and N. Hafer. 2012. Growth and ontogeny of the tapeworm *Schistocephalus solidus* in its copepod first host affects performance in its stickleback second intermediate host. Parasit. Vectors 5:90.

Benesh, D. P., J. C. Chubb, and G. A. Parker. 2013. Complex life cycles: why refrain from growth before reproduction in the adult niche? Am. Nat. 181:39–51.

Benesh, D. P., F. Weinreich, and M. Kalbe. 2012. The relationship between larval size and fitness in the tapeworm *Schistocephalus solidus*: bigger is better? Oikos 121:1391–1399.

Benesh, D. P., F. Weinreich, M. Kalbe, and M. Milinski. 2014. Lifetime inbreeding depression, purging, and mating system evolution in a simultaneous hermaphrodite tapeworm. Evolution 68:1762–74.

Cheverud, J. M. 1988. A comparison of genetic and phenotypic correlations. Evolution 42:958–968.

Christen, M., J. Kurtz, and M. Milinski. 2002. Outcrossing increases infection success and competitive ability: experimental evidence from a hermaphrodite parasite. Evolution 56:2243–51.

Christen, M., and M. Milinski. 2003. The consequences of self-fertilization and outcrossing of the cestode *Schistocephalus solidus* in its second intermediate host. Parasitology 126:369–378.

Cockerham, C. C., and B. S. Weir. 1984. Covariances of relatives stemming from a population undergoing mixed self and random mating. Biometrics 40:157–64.

Crnokrak, P., and D. A. Roff. 1995. Dominance variance - associations with selection and fitness. Heredity 75:530–540.

Eyre-Walker, A., and P. D. Keightley. 2007. The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8:610–618.

Frankham, R. 1990. Are responses to artificial selection for reproductive fitness characters consistently asymmetrical? Genet. Res. 56:35–42.

Hadfield, J. D., A. Nutall, D. Osorio, and I. P. F. Owens. 2007. Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. J. Evol. Biol. 20:549–557.

Hammerschmidt, K., K. Koch, M. Milinski, J. C. Chubb, and G. A. Parker. 2009. When to go: optimization of host switching in parasites with complex life cycles. Evolution 63:1976–1986.

Hedrick, P., and S. Kalinowski. 2000. Inbreeding depression in conservation biology. Annu. Rev. Ecol. Syst. 31:139–162.

Houle, D. 1998. How should we explain variation in the genetic variance of traits? Genetica 102–103:241–53.

Kurtz, J., and K. Franz. 2003. Evidence for memory in invertebrate immunity. Nature 425:37–38.

McGuigan, K., and M. W. Blows. 2009. Asymmetry of genetic variation in fitness-related traits: apparent stabilizing selection on g(max). Evolution 63:2838–47.

Michaud, M., M. Milinski, G. A. Parker, and J. C. Chubb. 2006. Competitive growth strategies in intermediate hosts: experimental tests of a parasite life-history model using the cestode, *Schistocephalus solidus*. Evol. Ecol. 20:39–57.

Moorad, J. A., and M. J. Wade. 2005. A genetic interpretation of the variation in inbreeding depression. Genetics 170:1373–84.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining $R^2$ from generalized linear mixed-effects models. Methods Ecol. Evol. 4:133–142.

Parker, G. A., J. C. Chubb, M. A. Ball, and G. N. Roberts. 2003. Evolution of complex life cycles in helminth parasites. Nature 425:480–4.

Pennycuick, L. 1971. Differences in the parasite infections in three-spined sticklebacks (*Gasterosteus aculeatus* L.) of different sex, age and size. Parasitology 63:407–18.

Roff, D. 1996. The evolution of genetic correlations: an analysis of patterns. Evolution 50:1392–1403.

Roff, D. A., and K. Emerson. 2006. Epistasis and dominance: evidence for differential effects in life-history versus morphological traits. Evolution 60:1981–90.

Schärer, L., L. M. Karlsson, M. Christen, and C. Wedekind. 2001. Size-dependent sex allocation in a simultaneous hermaphrodite parasite. J. Evol. Biol. 14:55–67.

Schielzeth, H., N. J. Dingemanse, S. Nakagawa, D. F. Westneat, H. Allegue, C. Teplitsky, D. Réale, N. A. Dochtermann, L. Z. Garamszegi, and Y. G. Araya-Ajoy. 2020. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods Ecol. Evol. 2020:1141–1152.

Skorping, A., A. Read, and A. Keymer. 1991. Life history covariation in intestinal nematodes of mammals. Oikos 60:365–372.

Tierney, J., and D. Crompton. 1992. Infectivity of plerocercoids of *Schistocephalus solidus* (Cestoda: Ligulidae) and fecundity of the adults in an experimental definitive host, *Gallus gallus*. J. Parasitol. 78:1049–1054.

Trouvé, S., P. Sasal, J. Jourdane, F. Renaud, and S. Morand. 1998. The evolution of life-history traits in parasitic and free-living platyhelminthes: a new perspective. Oecologia 115:370–378.

Van Buskirk, J., and Y. Willi. 2006. The change in quantitative genetic variation with inbreeding. Evolution 60:2428–34.

Van Der Veen, I. T. 2003. Is body size or activity of copepods related to ingestion of parasite larvae? Parasitology 126:173–178.

Van Der Veen, I. T., and J. Kurtz. 2002. To avoid or eliminate: cestode infections in copepods. Parasitology 124:465–474.

Wedekind, C. 1997. The infectivity, growth, and virulence of the cestode *Schistocephalus solidus* in its first intermediate host, the copepod *Macrocyclops albidus*. Parasitology 115:317–324.