

Experimental evolution for generalists and specialists reveals multivariate genetic constraints on thermal reaction norms

D. BERGER^{*†}, R. J. WALTERS^{†‡} & W. U. BLANCKENHORN[†]

^{*}Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

[†]Institute for Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[‡]School of Biological Sciences, University of Reading, Reading, UK

Keywords:

adaptation;
canalization;
generalist;
multivariate trade-off;
performance curve;
phenotypic plasticity;
reaction norm;
specialist;
temperature tolerance.

Abstract

Theory predicts the emergence of generalists in variable environments and antagonistic pleiotropy to favour specialists in constant environments, but empirical data seldom support such generalist–specialist trade-offs. We selected for generalists and specialists in the dung fly *Sepsis punctum* (Diptera: Sepsidae) under conditions that we predicted would reveal antagonistic pleiotropy and multivariate trade-offs underlying thermal reaction norms for juvenile development. We performed replicated laboratory evolution using four treatments: adaptation at a hot (31 °C) or a cold (15 °C) temperature, or under regimes fluctuating between these temperatures, either within or between generations. After 20 generations, we assessed parental effects and genetic responses of thermal reaction norms for three correlated life-history traits: size at maturity, juvenile growth rate and juvenile survival. We find evidence for antagonistic pleiotropy for performance at hot and cold temperatures, and a temperature-mediated trade-off between juvenile survival and size at maturity, suggesting that trade-offs associated with environmental tolerance can arise via intensified evolutionary compromises between genetically correlated traits. However, despite this antagonistic pleiotropy, we found no support for the evolution of increased thermal tolerance breadth at the expense of reduced maximal performance, suggesting low genetic variance in the generalist–specialist dimension.

Introduction

Phenotypic plasticity plays a fundamental role in evolutionary theory (West- Eberhard, 2003); it is central to understanding trade-offs (Levins, 1968; Via & Lande, 1985; Lynch & Gabriel, 1987; DeWitt *et al.*, 1998; Lande, 2014), it determines the strength of selection and rates of adaptation (Huey & Kingsolver, 1993; Whitlock, 1996; Snell- Rodd *et al.*, 2010; Walters *et al.*, 2012; Chevin *et al.*, 2013), and it also governs rates of migration and habitat colonization that ultimately shape species distributions (Van Tienderen, 1991; Sultan & Spencer, 2002; Crispo, 2008). Organisms experiencing heteroge-

neous environments need to express some level of plasticity, and the more variable the environment, the more plastic the organism needs to be. This is manifested as greater environmental tolerance and is typically measured, and modelled, as the breadth of the reaction norm for fitness or traits closely associated with it (Levins, 1968; Lynch & Gabriel, 1987; Gilchrist, 1995). However, if there is a cost of maintaining redundant physiological machinery, resulting in a trade-off between tolerance breadth and maximal performance at the optimal conditions (a *jack-of-all-trades is a master of none* trade-off: Huey & Kingsolver, 1989), stable environments will favour locally specialized genotypes.

The spatiotemporal pattern of environmental variation is predicted to affect the evolution of plasticity. When environmental variation is fine-grained, so that each individual experiences fluctuating conditions, evolution may favour either specialists or generalists depending on whether performance in each microenvironment

Correspondence: David Berger, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden.
Tel.: +46 18 471 2634; fax: +46 18 471 0000;
e-mail: david.berger@ebc.uu.se

contributes additively or multiplicatively to fitness (Lynch & Gabriel, 1987; Gabriel & Lynch, 1992; Gilchrist, 1995). In the case when fitness is equal to the sum of contributions, if the physiological costs of maintaining a broad environmental tolerance are high, an organism should specialize to the most common and/or productive conditions, forsaking performance elsewhere (Gilchrist, 1995). However, if fitness accrues multiplicatively across environments, fitness is largely determined by the performance in the least favourable environment (Lynch & Gabriel, 1987). A generalist strategy is therefore favoured, even when there is a trade-off between breadth and maximal performance. Similarly, in coarse-grained environments, where single genotypes need the ability to propagate at multiple discrete conditions, a generalist strategy is necessary. This is often the case for species with multiple generations per year and which are subject to seasonal variation, or where there is gene flow between different environments in a heterogeneous landscape (Van Tienderen, 1991; Sultan & Spencer, 2002). Although these theoretical predictions are well founded, they all assume that there is a cost to phenotypic plasticity, for which there is scarce empirical evidence. Instead, superior generalist genotypes are often observed to outcompete specialists across the entire environmental gradient measured in both experimental evolution and comparative studies (e.g. DeWitt *et al.*, 1998; Kassen, 2002; Relyea, 2002; Pigliucci, 2005; Callahan *et al.*, 2008; Van Buskirk & Steiner, 2009). Either trade-offs are unimportant (see Whitlock, 1996) or the costs associated with trade-offs are not manifesting under the conditions or in the traits being assayed.

Thermal tolerance is a classic example of phenotypic plasticity. Thermal tolerance should be governed by generalist–specialist trade-offs as it depends on upholding homeostasis across temperatures, which in turn relies on various costly physiological processes (Huey & Kingsolver, 1989; Angilletta, 2009). Thus, conserved performance across temperature gradients corresponds to high levels of plasticity at the physiological level (Hochachka & Somero, 2002; Clarke, 2003; Grether, 2005). Thermal generalist–specialist trade-offs may arise from the conflicting demands of expressing temperature-sensitive enzymes with efficient catalytic properties working well at cold temperatures, and more stable isoforms working better at warm temperatures (Huey & Kingsolver, 1989). Alleles encoding synthesis of different enzyme variants are thus predicted to show antagonistic pleiotropy for performance across temperatures, and trade-offs between performance maximum and breadth should result from reduced expression of each specific enzyme as a consequence of the need to simultaneously express multiple variants (Huey & Kingsolver, 1989) (Fig. 1a).

Superior performance across a broad range of temperatures may nevertheless be permitted by an overall increase in protein synthesis and parallel or sequential

expression of multiple, diverse protein configurations. For example, faster growth rates across the entire temperature range can be achieved by increasing food acquisition rates together with increased production of catalysing enzymes. However, such a strategy is likely to either divert resources away from other important physiological functions, resulting in an allocation trade-off, or render elevated predation risk associated with increased foraging rates, resulting in an acquisition trade-off (Houle, 1991; de Jong & van Noordwijk, 1992; Dmitriev, 2011). As a consequence, trade-offs may not be apparent at the level of single-trait reaction norms because the costs are incurred by other traits sharing the same functional and/or developmental basis (Angilletta *et al.*, 2003; Clarke, 2003; Ghalambor *et al.*, 2007; Callahan *et al.*, 2008). Evidence for antagonistic pleiotropy for fitness across thermal gradients may thus materialize as temperature-specific solutions to such multivariate trade-offs (Fig. 1b–d).

Thermal reaction norms are ideal for studying the evolution of phenotypic plasticity because temperature has pivotal and relatively well-understood effects on physiology, and because thermal performance is assessable in the laboratory. Evolutionary shifts in thermal reaction norms often incorporate compensatory allocation and acquisition trade-offs (Huey & Kingsolver, 1989; Angilletta *et al.*, 2003) because typically genetic variance for temperature specificity of enzymatic reactions is low (Hochachka & Somero, 2002; Latimer *et al.*, 2014). This may explain why ‘superior’ generalists are often observed in empirical studies of thermal performance curves and why generalist–specialist trade-offs seldom come to dominate standing genetic variation in single-trait reaction norms. We selected for generalists and specialists in the black scavenger fly *Sepsis punctum* (Diptera: Sepsidae) via experimental evolution over 20 generations using four treatments: adaptation to constant 31 °C (‘Hot’ selection) or constant 15 °C (‘Cold’), or under thermal regimes fluctuating between these temperatures either within (‘Intra’) or between (‘Inter’) generations. Juveniles were raised on a restricted food resource, conditions that should maximize the strength of selection at the juvenile stage and have potential to reveal trade-offs at the level of the reaction norm, or multivariate allocation trade-offs between different traits (Houle, 1991; de Jong & van Noordwijk, 1992). We assessed thermal reaction norms, and genetically correlated responses that could signify trade-offs, in three life-history traits closely related to fitness: size at maturity, juvenile growth rate and juvenile survival. Our experimental design also allowed assessment of possible parental acclimation effects, that is, putative plasticity in offspring reaction norms induced by temperature cues received by the parents. Such plasticity may confer selective benefits whenever the parental environment acts as a reliable cue for offspring conditions (Gabriel & Lynch, 1992; Fox & Mousseau, 1998;

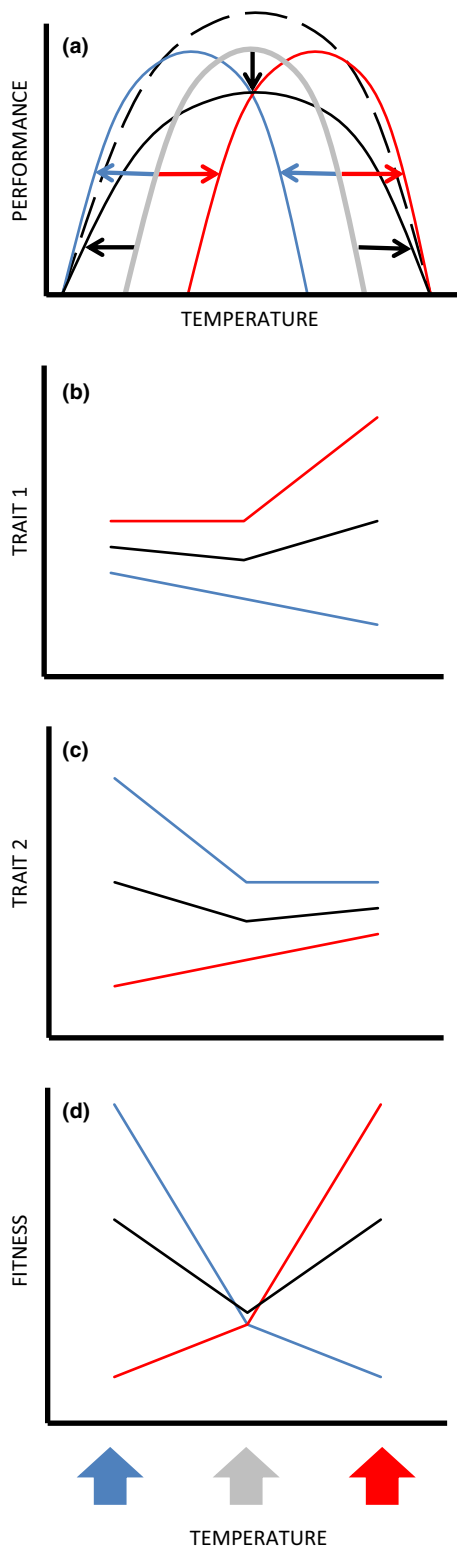


Fig. 1 Evolution of thermal generalists and specialists. (a) Theoretical predictions of thermal reaction norm evolution from the wild type (grey) adapted to an intermediate temperature; specialization to cold (blue) or hot (red) temperature, or adaptation to an environment varying between cold and hot temperatures leading to the evolution of a thermal generalist (black) with broader thermal tolerance at the expense of reduced maximal performance. If costs of plasticity are high and cumulative fitness is approximated by the arithmetic mean calculated across temperatures, variation may select for specialization to warm temperature (the *hotter-is-better hypothesis*: Kingsolver & Huey, 2008; Angilletta *et al.*, 2010), resulting in any intermediate form of the red and black reaction norm. If costs of plasticity are low, such as when performance at suboptimal temperatures can be enhanced by compensatory resource acquisition to low costs, a superior thermal generalist evolves (hatched line). (b–d) Evolution of specialization due to antagonistic pleiotropy for fitness via correlated responses in two traits involved in a trade-off. Both traits 1 and 2 individually show positive cross-temperature genetic (co)variances but are involved in a trade-off and are thus negatively genetically correlated with each other at each temperature ($r = -0.5$). At hot temperature (red arrow), directional selection (β) is stronger on trait 1 ($\beta = 1$) than on trait 2 ($\beta = 0.5$), whereas at cold temperature (blue arrow) the situation is exactly reversed. Due to positive cross-temperature covariance for each trait, but negative covariance between the two traits, evolution exclusively at hot temperature leads to a genotype with overall high values for trait 1 (b) but low values for trait 2 (c), whereas evolution at cold temperature leads to a genotype with low values for trait 1 (b) and high values for trait 2 (c). The thermal generalist shows intermediate values. Under the assumption that these two traits contribute multiplicatively to fitness, thermal specialization to cold and hot temperatures, respectively, evolves as a result of different temperature-specific optima for the trade-off between traits 1 and 2 (d).

Gabriel, 2005), and could thereby reduce demographic costs under an initial phase of adaptation (Chevin *et al.*, 2013; Lande, 2014).

Predictions

We expected that our experimental design would expose antagonistic pleiotropy for performance at hot and cold temperatures (Angilletta, 2009), with the possibility that this antagonistic pleiotropy could be driven by temperature-specific responses in single-trait reaction norms and/or temperature-specific solutions to a trade-off involving these traits (Fig. 1b–d). As a result, Hot and Cold populations should evolve to become more specialized to warm and cold temperatures, respectively (Fig. 1a). We expected that the Inter-populations, experiencing coarse-grained between-generation variation, would evolve to become generalists and exhibit relatively broader thermal performance curves at the expense of reduced maximal performance (*generalist–specialist trade-off*: Fig. 1a). Finally, for the Intra-populations adapting to

fine-grained within-generation variation, two outcomes are plausible. Flies could evolve generalist strategies as predicted for the Inter-treatment. Alternatively, they could become adapted to a specific thermal range, which would imply (i) high costs associated with being a thermal generalist and (ii) unequal fitness pay-offs associated with adaptation to hot vs. cold temperature, which has been suggested to be a general pattern (the *hotter-is-better* hypothesis: Kingsolver & Huey, 2008; Angilletta *et al.*, 2010) (Fig. 1a).

Results reveal that thermal specialization and antagonistic pleiotropy for performance at hot and cold temperature can emerge from temperature-specific solutions to a trade-off between size at maturity and juvenile survival (Fig 1b–d). Despite finding a pivotal role for antagonistic pleiotropy in shaping reaction norms, we observed no response in the generalist–specialist dimension, implying that the evolution of reaction norm breadth is subject to constraints.

Materials and methods

Study species

Sepsidae are a fly family with 36 known genera and ca. 300+ species worldwide (Pont & Meier, 2002). The genus *Sepsis* comprises 35 palaearctic species. *Sepsis punctum* has a widespread distribution, ranging from North America to Europe, North Africa and parts of Asia. It can be found on a range of vertebrate excrements, particularly on cow dung. *Sepsis punctum* is multivoltine throughout most of its range, except at very high latitudes (Pont & Meier, 2002), and thus experiences very variable temperatures both within and between generations in nature (Berger *et al.*, 2013). Development rates are fast with individuals completing development in ca. 2 weeks at 23 °C, which suggests that this species maximizes juvenile growth rate, as adult life at the same temperature can span over more than 2 months (Puniamoorthy *et al.*, 2012a,b).

Creating the replicated experimental populations

The source flies were collected from a multivoltine population with overlapping generations (Berger *et al.*, 2013) in Zürich, Switzerland, in August 2009. Ninety traps containing cow dung for wild females to lay their eggs in were put out at sites approximately 10 metres apart around the university campus. Traps were brought to the laboratory after 1 week and placed singly in a transparent 1-L rearing container (8 × 8 × 15 cm³) with sugar, water and fresh cow dung, at 23 °C, 60% humidity and a 16:8 LD cycle, which are optimal conditions for this species (e.g. Puniamoorthy *et al.*, 2012a,b). F0 offspring emerged from 45 traps (typically 30–100 flies per trap). These flies were allowed to mate freely within each container,

expanding population sizes to 200–400 adults in the F1 generation.

The 45 lines were then randomly mixed in groups of five and placed into nine larger 3-L containers. After an additional generation of mating and population expansion, 50 newly emerged F2 flies were transferred from each of the nine population containers to a new 3-L container (12 × 12 × 22 cm³) to create an experimental population comprising 450 flies. In total, we created 32 such replicate populations. To attain estimates of reaction norms for wild-type flies acclimated to 23 °C, we at this point randomly sampled 11 populations for newly hatched F3 larvae that were set up in Performance assays. Subsequently, each experimental population was assigned haphazardly to one of four selection treatments: constant 31 °C (Hot populations), constant 15 °C (Cold populations), a thermal regime fluctuating between 15 and 31 °C on a 24-h basis (Intra-populations) or a thermal regime fluctuating between 15 and 31 °C between generations (Inter-populations), yielding eight replicate populations per treatment. In the Inter-treatment, four replicates were started at 31 °C and the other four at 15 °C. This controlled for the order of thermal regime on the evolutionary response and allowed assessment of bidirectional acclimation effects from parents developing either at 15 °C or at 31 °C (see Performance assays).

Experimental evolution

Containers were supplied with (organic) cow dung, sugar and water, and we exchanged containers and refreshed resources every generation. Dung was supplied in four containers at 50 g, which supported ca. 400–500 *S. punctum* but represents a limited food resource for offspring because a single female may produce hundreds of eggs (Pont & Meier, 2002; Puniamoorthy *et al.*, 2012a). In accordance with this, we observed several hundred, if not thousands of dead larvae in the dung at each generation of evolution. As is standard procedure, dung was homogenized and deep-frozen (–80 °C) to minimize random phenotypic variation attributable to food quality and to eliminate competing species and predators (e.g. Blanckenhorn *et al.*, 2010; Berger *et al.*, 2011a). To reduce direct selection on early reproduction, we supplied two containers of dung 2 (4) and 4 (8) days after emergence of the first adults at 31 °C (15 °C). Generation times were approximately 12 days for the Hot, 20 days for the Intra and 35 days for the Cold treatment. The Inter-treatment varied between 12 and 35 days every other generation as populations were moved between the Hot and Cold selection regimes. Inter-populations were moved between temperatures when adults had emerged and were given new fresh dung for feeding and ovipositing, assuring that juveniles experienced the same constant temperature (15 or 31 °C) throughout development.

Performance assays

One problem arising when comparing selection treatments adapting to different thermal regimes resulting in different generation times is that a common garden design at one point in time would compare selection treatments that have adapted to laboratory conditions for different numbers of generations. Alternatively, selection treatments can be scored at different times, then having adapted for equal numbers of generations. We chose the second alternative, because allowing adaptation to laboratory conditions for an unequal number of generations would likely have revealed strong differences between selection treatments not associated with the thermal regimes *per se* (compare evolved and wild-type populations in Fig. 2 and Fig. S3). To minimize potential temporal block effects, we used the same climate cabinets and dung sampled at a single occasion, homogenized and deep-frozen at -80°C .

After 20 generations of evolution, we scored thermal reaction norms of two types of offspring for all populations. Offspring ‘acclimated to selection regime’

were derived from parents moved directly to the intermediate and benign 23°C after adult emergence from their respective selection treatment. Offspring ‘acclimated to 23°C ’ were derived from parents that had been reared for one additional generation at 23°C . The comparison between these two types should thus have the potential to reveal parental effects. Importantly, differences between offspring from parents ‘acclimated to selection regime’ should include parental in addition to genetic effects, whereas differences between offspring from parents ‘acclimated to 23°C ’ should primarily reflect genetic differences.

For our assays, populations kept at 23°C were supplied with a thin layer of dung for egg laying. After 24 h we removed the dung samples. Newly hatched larvae were transferred under microscope to fresh dung provided *ad libitum*. We note that although our thermal regimes were selected for juvenile performance under food limitation and larval competition, our objective in the performance assays was to assess resulting differences in reaction norms for maximal genetic values in respect of only temperature, so larval competition was eliminated. Due to their transparent bodies, it was

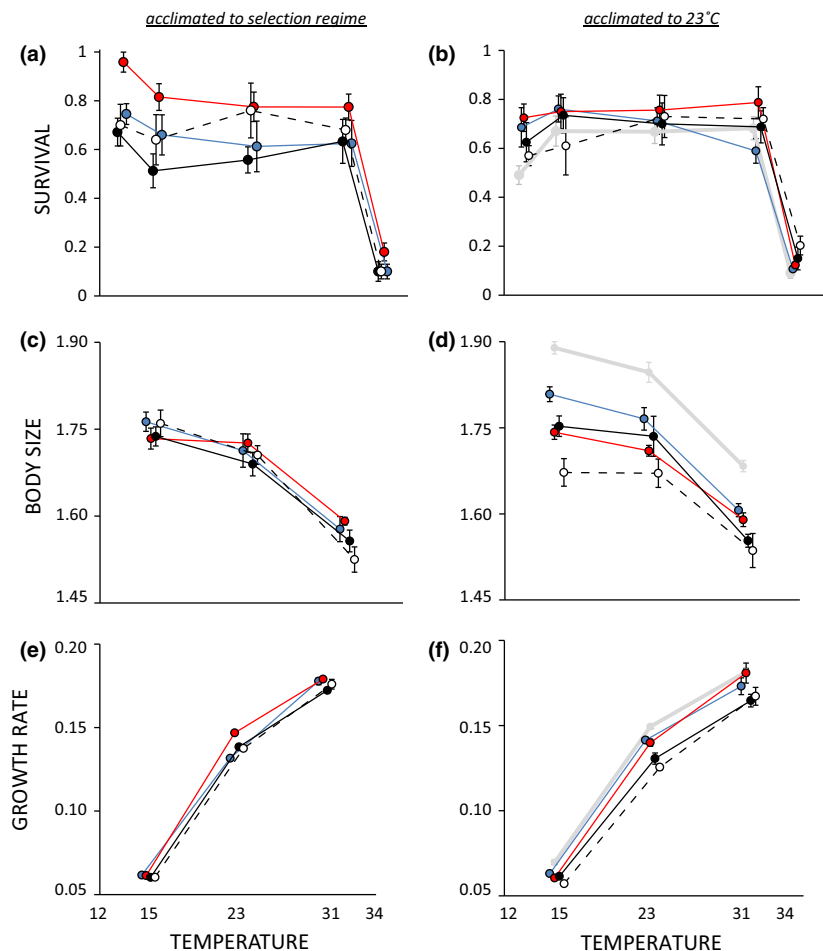


Fig. 2 Thermal reaction norms for survival (a–b), body size (c–d) and growth rate (e–f) following 20 generations of evolution at 15°C (blue), 31°C (red), inter-generational variation (black) and intra-generational variation (white). Left-hand panels (a, c, e) show results for offspring of parents moved directly from respective selection treatment to 23°C for egg laying. Right-hand panels (b, d, f) show results for offspring whose parents had been allowed to acclimate to 23°C for an additional full generation. Grey circles in right-hand panels represent wild-type flies scored prior to laboratory evolution. Means ± 1 SE. Body size was measured as hind tibia length in millimetres, and growth rate was calculated as hind tibia length/development time (mm day^{-1}). Figure is based on raw data whereas analyses were performed on standardized data, for which figures can be found in the supplementary material.

possible to reject larvae that had fed, ensuring that possible variation in timing of egg laying did not contribute to differences between populations. A replicate of 10 larvae was set up on *ad libitum* dung contained within a plastic container that was put into a glass vial and sealed with a foam stopper, ensuring humid conditions that kept the dung from drying. In total, we set up one to four replicates per population at each of five experimental temperatures.

Thermal reaction norms were assessed by measuring body size (hind tibia length in mm) and growth rate (hind tibia length/larva-to-adult development time) at 15, 23 and 31 °C, and larva-to-adult survival at 12, 15, 23, 31 and 34 °C. Test temperatures 12 and 34 °C represent extremes experienced by larvae living within cow dung pats in our population (Berger *et al.*, 2013). Measurements of body size and growth rate at 12 °C were abandoned, however, for logistic reasons, and the few flies emerging at 34 °C were severely crippled and not possible to measure accurately (see also Berger *et al.*, 2013). Because adult reproductive success at 15 and 31 °C was only ca. 20–30% of that at 23 °C in this population (D. Berger & R.J. Walters, unpublished), this 16 °C temperature range remains ecologically relevant.

Statistics

For all traits, we calculated replicate means that were used as statistical units in the analyses. Survival was recorded per replicate as the number of emerging flies divided by the number set-up. Body size was measured for two males and two females from each replicate. Growth rate was calculated as the ratio between each replicate's mean body size and development time (based on all emerging individuals). On a few occasions, fewer than two individuals from one sex emerged in a replicate. We therefore calculated sex-specific estimates of body size and growth rate per replicate and used sex as a factor in statistical models. For body size, these estimates supply additional interesting information on the sex specificity of evolutionary responses (see Results). However, as we did not have sex-specific estimates of development times, for growth rate we merely added this factor to the model to control for occasions where there were an unequal number of males and females measured for body size (note that males are bigger, which results in higher growth rates), and we do not report sex differences in this trait.

Both growth rate and body size were mean-standardized prior to analysis. This homogenized variances across experimental temperatures and ensures that significant interactions between selection treatment and test temperature are due to changes in relative differences between selection treatments, and not merely a result of overall effects of temperature on trait means. We also performed analyses on logged data for these variables.

However, as the results from these analyses were qualitatively the same as the results of the main analyses, we do not comment further on this. Survival was square-root transformed to improve model fit. Main figures are presented using raw data, but figures on transformed data, on which the statistical analyses were performed, are available in the supplementary material.

We analysed differences in thermal reaction norms between selection treatments using linear mixed models with maximum likelihood available in the package lme4 v. 0.99999-2 (Bates *et al.*, 2011) for the statistical software R (R Development Core Team, 2011). For models on growth rate and body size, we included selection treatment, test temperature, parental acclimation treatment, sex and their interactions as fixed effects. We included population identity crossed by test temperature as random effects, ensuring correct level of replication (populations) for the fixed effects. The model on survival was identical except that sex was not included as we did not note the sex for all emerging flies and did not know the sex of the original larvae. Significance was evaluated by comparing likelihood ratios of models with and without the specific effect of interest using a type-II sums-of-squares approach.

Expected fitness

Reproductive success increases with adult body size in most species (Blanckenhorn 2000), and life-history theory assumes trade-offs between size at maturity and survival to adulthood. We expect antagonistic pleiotropy for alleles regulating growth and survival because, even in the absence of increased predation risk on juveniles, high growth rates should trade off against allocation to somatic maintenance, leading to increased intrinsic mortality (Blanckenhorn 2000, Dmitriew, 2011). In the Swiss population of *S. punctum*, both sexes do indeed experience positive selection on adult size under laboratory conditions similar to those used here (Puniamoorthy *et al.*, 2012a). We therefore calculated 'expected fitness' at each temperature as a function of each population's relative adult body size and survival to that size. Adult reproductive success (RS) at each temperature was predicted from our adult body size data and the previously estimated selection differentials on adult body size ($\beta_i = 0.30 \pm 0.05$, $\beta_m = 0.31$ (0.22, 0.39, 0.34 at different sex ratios); Puniamoorthy *et al.*, 2012a):

$$RS_{iT} = 1 + \left[\beta \frac{B_{iT} - \overline{B_T}}{SD_T} \right] \quad (1)$$

where we set $\beta = 0.3$, B_{iT} represents the mean body size of population i at temperature T , and $\overline{B_T}$ and SD_T represent, respectively, the mean and standard deviation of body size calculated across all population means at the same temperature. Expected fitness at each temperature is then:

$$\text{Expected fitness} = RS_{iT} \frac{S_{iT}}{\bar{S}_T} \quad (2)$$

where S_{iT} is the survival of population i at temperature T , and \bar{S}_T is the mean survival calculated across all population means at the same temperature. We then compared fitness across test temperatures and selection treatments using ANOVA with experimental populations as statistical replicates.

Partitioning evolutionary responses in reaction norms into dimensions of biological interest

We used the data from offspring of parents acclimated to 23 °C to offer a complementary view on the evolutionary responses in the four selection treatments by partitioning variation into three dimensions: reaction norm elevation (overall reaction norm height), slope (specialization to either hot or cold temperature) and curvature (reaction norm breadth). For each of the experimental populations, we calculated a mean for each of the four traits (survival, body size, growth rate and expected fitness) at each temperature, again using data mean-standardized per temperature. For each experimental population, values for elevation, slope and curvature for trait Z could then be obtained (see also Murren *et al.*, 2014):

$$\text{elevation} = \frac{\sum_1^n Z_i}{n} \quad (3.1)$$

$$\text{slope} = \frac{\sum_1^{n-1} S_i}{n-1}; S_i = Z_{i+1} - Z_i \quad (3.2)$$

$$\text{curvature} = \frac{\sum_1^{n-2} C_i}{n-2}; C_i = S_{i+1} - S_i \quad (3.3)$$

where n equals the number of test temperatures and i the focal temperature. These scores were then analysed with ANOVAs, using experimental populations as statistical replicates, to look for differences across selection treatments. We made three meaningful comparisons by

making subsets of the data. We compared Hot and Cold populations to investigate differences between populations adapting to constant 31 and 15 °C, and Intra- and Inter-populations to investigate whether the mode of temporal variation (within or between generations) affected genetic responses. Finally, we pooled Hot and Cold populations and compared them with pooled data from Intra- and Inter-populations to investigate whether evolution at constant vs. variable temperatures affected genetic responses in the three dimensions.

Results

Four of our 32 populations went extinct during the course of experimental evolution, all of which were observed as having low adult counts (<200) for a number of generations prior to extinction. One population belonged to the Cold and three to the Intra-treatment. Our selection regimes were harsh, and analysed differences are the result of selection at both the population and individual levels.

Because at any one time half of the Inter-treatment populations were developing at 15 °C and the other half at 31 °C, we tested whether the temperature at which the parental generation had been raised had any effect on offspring performance in these populations. Although statistical power for detecting moderate effects of parental acclimation was low, there appeared to be no significant influence of parental temperature on any of the three traits measured in offspring (S1); we therefore pooled these two categories of Inter-populations for all subsequent analyses.

Evolution of thermal performance curves

Survival

The selection treatments differed in larva-to-adult survival ($P < 0.001$; Table 1), with Hot populations tending to show highest survival. There was no significant interaction between selection treatment and test temperature to indicate temperature-specific adaptation,

Table 1 Effects of selection treatment (sel), parental acclimation (acc) and test temperature (temp) on survival, body size and growth rate. P -values are based on χ^2 -statistics from likelihood-ratio tests. Note that although the main effect of temperature was retained in all models, the effect is not reported as all analyses were performed on data mean-standardized per test temperature, effectively nullifying the main effect of temperature.

Effect	Survival			Body size			Growth rate		
	χ^2	d.f.	P	χ^2	d.f.	P	χ^2	d.f.	P
Acc	0.04	1	0.84	0.13	1	0.72	4.91	1	0.026
Sel	32.1	3	< 0.001	12.8	3	0.005	32.8	3	< 0.001
Acc:temp	11.0	3	0.012	0.17	2	0.92	1.61	2	0.45
Acc:sel	6.71	3	0.082	12.0	3	0.007	9.65	3	0.022
Temp:sel	12.7	9	0.18	9.46	6	0.15	31.8	6	< 0.001
Acc:temp:sel	3.72	9	0.93	10.9	6	0.09	23.4	6	< 0.001

Significant P -values shown in bold.

although statistical power was too weak to pick up potential minor differences (Fig. 2a,b). Parental acclimation at 23 °C affected temperature-specific survival ($P = 0.012$; Table 1), evident primarily as a decrease in survival at 12 °C, but there was no corresponding increase at the hot temperatures as would be predicted under a trade-off scenario (Fig. 2a,b). Survival in the evolved populations tended to be higher than wild-type survival (Fig. 2b and Fig. S3).

Body size

Selection treatments differed in body size ($P = 0.005$; Table 1). Cold populations tended to be largest, but sizes had decreased compared to wild-type flies in all selection treatments (Fig. 2d and Fig. S3). There was no interaction between selection treatment and test temperature when analysing all selection treatments. However, the interaction was significant when comparing Hot and Cold populations only ($P < 0.008$; S6), demonstrating temperature-specific genetic responses to the selection regimes (Fig. 2c,d and Fig. S5). The effect of parental acclimation differed between selection treatments ($P = 0.007$; Table 1), resulting in greater differences between treatments after acclimation to 23 °C (compare Fig. 2c,d). Male and female thermal reaction norms showed strong differences ($P < 0.001$, Fig. S2), and the evolutionary response in these reaction norms was sex specific, signified by a significant three-way interaction between selection treatment, sex and test temperature ($P = 0.006$; S2). Specifically, in the Hot and Cold populations, sexual size dimorphism was greatest at the temperature at which the respective populations had evolved (Fig. S2).

Growth rate

There were large differences between selection treatments ($P < 0.001$; Table 1), with populations evolving under the Cold or Hot regime growing faster than populations evolving under variable regimes (Fig. 2e,f). Notably, there was a strong interaction between test temperature and selection treatment ($P < 0.001$; Table 1). As predicted, Hot populations grew faster than Cold populations at 31 °C, and Cold populations grew faster than Hot populations at 15 °C (post hoc comparison of only Hot and Cold populations, $P < 0.001$; Table S6), indicating temperature-specific allelic effects on growth rate (Fig. 2e,f and Fig. S5). There was also a difference between selection treatments in how reaction norms were affected by parental acclimation, indicated by a significant three-way interaction between acclimation treatment, test temperature and selection treatment ($P < 0.001$; Table 1).

Correlated responses and genetic trade-offs

Across the wild-type plus the four types of evolved populations, there was a tendency for populations with

large body size and fast growth rates to have relatively lower survival (Fig. S3). We explored this result further by calculating correlations between juvenile survival on the one hand and adult body size and growth rate on the other. The calculations were based on means for each of the 28 evolved populations measured plus the 11 wild-type populations and thus should estimate broad-sense genetic correlations. The use of population means as opposed to replicate means reduces the problem of the 'missing fraction' and should give more robust estimates of trait correlations (Hadfield, 2008). The inclusion of the wild-type populations left correlations virtually unchanged, but improved statistical power through increased sample size. In general, these population-mean correlations were negative, but much more so when calculated across hot and cold test temperatures. The negative correlations with survival were similar for both growth rate and body size, although slightly more negative for the latter (Table 2), suggesting that somatic growth trades off with juvenile survival under harsh conditions. Cross-temperature correlations for each separate trait were positive or nil (S4).

Expected fitness

'Expected fitness' (the product of relative juvenile survival and predicted adult reproductive success associated with the evolved adult body size: eqns 1 and 2) differed between selection treatments ($\chi^2 = 11.0$, d.f. = 3, $P = 0.012$). The Intra-generalists had lower expected fitness, whereas the Inter-generalists had average, but certainly not superior, expected fitness compared with specialists (Fig. 3). When considering all populations together, there was no significant

Table 2 Broad-sense genetic correlations between survival to adulthood and size at maturity based on population means, including wild types (total $n = 39$). Large body size at relatively warm temperature is negatively correlated with survival at the coldest temperature, and vice versa, large body size at relatively cold temperature is negatively correlated with survival at the warmest temperature.

	Survival 12	Survival 15	Survival 23	Survival 31	Survival 34
Correlations					
Size 15	-0.10	0.22	-0.11	-0.09	-0.31
Size 23	-0.36	-0.07	-0.05	-0.04	-0.44
Size 31	-0.31	0.01	-0.15	-0.01	-0.25
<i>P</i> -values					
Size 15	0.535	0.185	0.490	0.603	0.061
Size 23	0.023	0.666	0.768	0.795	0.007
Size 31	0.057	0.973	0.347	0.957	0.140

Significant and marginally significant correlations are highlighted in bold.

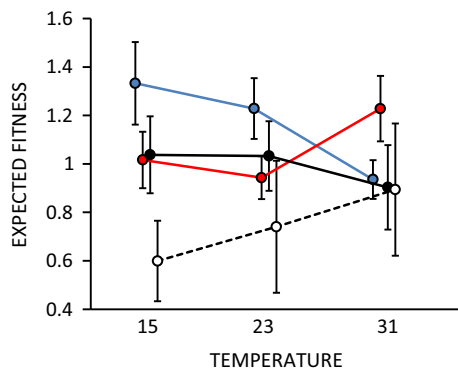


Fig. 3 'Expected fitness', estimated as the product of each population's relative juvenile survival and predicted adult reproductive success based on size at maturity, following 20 generations of evolution and one full generation of parental acclimation at 23 °C, for Cold (blue), Hot (red), Inter (black)- and Intra (white)-populations. Means \pm 1 SE.

interaction between selection treatment and test temperature ($\chi^2 = 11.3$, d.f. = 6, $P = 0.080$); however, the expected pattern of antagonistic pleiotropy was observed when only the Hot and Cold populations were compared (selection treatment: $\chi^2 = 1.41$, d.f. = 1, $P = 0.23$, selection*test temperature: $\chi^2 = 10.4$, d.f. = 2, $P = 0.005$) (Fig. 3 and Fig. S5). These results illustrate that antagonistic pleiotropy between temperatures can become apparent when considering trade-offs between correlated traits (here juvenile survival and adult size) (Fig. 1b–d). Intra-lines, experiencing fine-grained within-generation variation, tended to show augmented expected fitness at the hot temperature (Fig. 3) in accordance with the 'hotter-is-better' hypothesis. However, there was no significant difference between Intra-populations and the Inter-populations that had experienced coarse-grained between-generation variation and were

predicted to evolve generalist reaction norms (selection treatment: $\chi^2 = 2.49$, d.f. = 1, $P = 0.11$, selection*temperature: $\chi^2 = 2.11$, d.f. = 2, $P = 0.35$), likely at least in part due to the low statistical power offered by the few (5) Intra-populations that could be assayed.

Partitioning evolutionary responses in reaction norms into dimensions of biological interest

The alternative analyses partitioning the evolutionary responses of reaction norms into elevation (overall), slope (hot-cold) and curvature (breadth) supported the former analyses. Hot populations tended to have higher survival but had smaller body size compared with Cold populations. For growth rate and expected fitness, there were significant differences in slope, indicating antagonistic pleiotropy for performance at hot and cold temperatures (Table 3; Fig. S5).

There were no significant differences between Intra- and Inter-populations to indicate that the temporal pattern of temperature variation affected evolutionary responses. However, sample size was low, and there were trends of Intra-populations having relatively higher trait values at 31 °C but overall smaller body size compared with Inter-populations, similar to the former comparison between Hot and Cold populations (Table 3; Fig. S5).

Pooling Hot and Cold populations, both evolving at constant temperature, and comparing them to the pooled Intra- and Inter-populations, both evolving at variable temperatures, showed that evolution at constant temperature had led to increased body size and growth rate but no differences in survival compared to evolution at variable temperatures. Evolution at constant temperature also led to overall higher fitness compared with evolution at variable temperature, opposite to what would be expected if superior thermal generalists would have evolved in the variable selection treatments (Table 3; Fig. S5).

Table 3 Genetic differences between selection treatments in thermal reaction norms partitioned into elevation (overall differences), slope (hot–cold) and curvature (breadth), in the range of 15–31 °C for body size, growth rate and expected fitness, and 12–34 °C for survival.

Comparison	Dimension	Survival			Body size			Growth rate			Expected fitness		
		<i>F</i>	d.f.	<i>P</i>	<i>F</i>	d.f.	<i>P</i>	<i>F</i>	d.f.	<i>P</i>	<i>F</i>	d.f.	<i>P</i>
Hot–cold	Elevation	3.69	13	<i>0.077</i>	23.1	13	< 0.001	0.09	13	0.77	1.25	13	0.28
	Slope	0.20	13	0.66	2.34	13	0.15	7.25	13	0.019	5.69	13	0.033
	Curvature	0.13	13	0.72	0.52	13	0.48	0.13	13	0.73	2.48	13	0.14
Intra–inter	Elevation	1.02	11	0.33	3.27	11	<i>0.098</i>	1.80	11	0.21	2.11	11	0.17
	Slope	1.69	11	0.22	1.33	11	0.27	2.66	11	0.13	1.81	11	0.21
	Curvature	2.92	11	0.12	0.10	11	0.76	0.09	11	0.77	0.06	11	0.82
Constant–variable	Elevation	0.004	26	0.95	6.62	26	0.016	24.6	26	< 0.001	5.15	26	0.032
	Slope	1.80	26	0.19	0.01	26	0.93	1.28	26	0.27	0.23	26	0.64
	Curvature	0.60	26	0.45	1.18	26	0.29	3.12	26	0.09	0.30	26	0.59

P-values below 0.10 in italic and below 0.05 in bold.

Discussion

The evolution of generalists and specialists

Previous experimental evolution studies on environmental tolerance, the majority performed on various microbes studying the effects of new mutations, similar to our experiment, found evidence for negative genetic correlations between performance in different environments but to a much lesser extent between performance breadth and maximum signifying generalist–specialist trade-offs (reviewed in: Kassen, 2002; Scheiner, 2002; Chevin *et al.*, 2013). Negative cross-environment correlations may in principle result from three mechanisms that are not mutually exclusive. The first is antagonistic pleiotropy, which occurs when certain alleles act to increase fitness in one environment at the cost of decreasing it in others, which is a prerequisite for generalist–specialist trade-offs. The second is differential selection of alleles with conditional effects apparent in the focal environment but neutral in others, which can lead to specialists with highest fitness in their respective selective regimes. The third mechanism is DNA decay, whereby long-term evolution in stable environments could lead to specialization through the accumulation of conditionally deleterious alleles with effects seen only at rare conditions (Whitlock, 1996; Kassen, 2002; Elena & Lenski, 2003; Martin & Lenormand, 2006; Hoffmann, 2010). If mutation accumulation and selection of alleles with conditional effects are the major forces leading to specialization, evolution in variable environments should create superior generalists by purging conditionally deleterious alleles, as often seen in experimental evolution studies (e.g. Bennett & Lenski, 1993, 2007; Reboud & Bell, 1997; Kassen & Bell, 1998; Hughes *et al.*, 2007). In contrast, the results from our study, focusing on standing genetic variation in *S. punctum*, did not lead to the evolution of superior generalist genotypes and are instead more consistent with the first proposed mechanism: antagonistic pleiotropy. Cold and Hot populations generally performed best in their respective environments, and the Inter- and Intra (generalist)-populations fell in between or below these genotypes (Figs 2 and 3), which would suggest that there are constraints on the evolution of increased thermal tolerance and/or costs to being a generalist.

Based on the different temperature sensitivities of enzyme variants, antagonistic pleiotropy is predicted to play a pivotal role in thermal reaction norm evolution (Huey & Kingsolver, 1989; Hochachka & Somero, 2002; Angilletta, 2009). Surprisingly however, trade-offs between performance at hot and cold temperature, or between performance breadth and maximum, are not often found (Angilletta *et al.*, 2003). Studies of genetic variance within and between natural

populations instead indicate that thermal performance often is shaped by variation in resource acquisition, leading to positive cross-temperature genetic correlations and superior generalist genotypes (e.g. Conover & Schultz, 1995; Yamahira *et al.*, 2007; Berger *et al.*, 2011b; Klepsatel *et al.*, 2013). This view is supported by outcomes of experimental evolution studies on organisms as diverse as *E. coli* (Bennett & Lenski, 2007) and *Drosophila* (Cooper *et al.*, 2012; Condon *et al.*, 2013). Further support comes from a recent study using *Drosophila* mutation accumulation lines to investigate mutational variance in thermal reaction norms for locomotor activity. The study found that most new mutations affect reaction norm elevation (>70% of the variation), whereas antagonistic effects between hot and cold temperatures explained only ca 20% of the mutational variance, and generalist–specialist variation < 6% (Latimer *et al.*, 2014). Thus, superior generalists may evolve because the polygenic basis of resource acquisition potential constitutes a relatively larger mutational target compared with the genes regulating the temperature specificity of enzymatic reaction rates (Houle, 1991; Angilletta *et al.*, 2003; Clarke, 2003). However, if genetic variance in reaction norm elevation is mostly due to deleterious mutations reducing resource acquisition potential (i.e. via mutation load), these alleles should quickly be removed by natural selection and may therefore contribute relatively little to standing genetic variation in thermal reaction norms, an assertion that has found some empirical support (e.g. Izem & Kingsolver, 2005; Latimer *et al.*, 2011). Therefore, given time, allelic variation with antagonistic effects on performance at hot and cold temperatures may come to shape reaction norms (Angilletta *et al.*, 2003; Yamahira *et al.*, 2007; Berger *et al.*, 2013; Murren *et al.*, 2014).

Our finding of apparent antagonistic pleiotropy guiding thermal adaptation and an absence of superior generalists is in contrast to results from many experimental evolution studies and patterns of population differentiation across latitude, but is in line with high levels of standing genetic variation in the hot–cold dimension compared to variation for reaction norm elevation. Our experimental evolution settings restricted juvenile food supply, evidenced by the documented high larval mortality during evolution and reductions in body size and growth rate in all selection treatments compared to the wild type (Fig. 2d,f). Although effects of food resources on thermal reaction norms can sometimes be complex (e.g. Diamond & Kingsolver, 2010; Nilsson-Örtman *et al.*, 2014), it seems plausible that the low food supply and high competition may have contributed further to our results by limiting possibilities for the evolution of increased thermal tolerance via increases in resource acquisition levels.

Interestingly however, we did not find an increase in performance breadth at the cost of reduced maximal

performance as predicted under the variable conditions in the generalist selection treatments (Intra and Inter), indicating that there are genetic constraints on reaction norm evolution in the generalist–specialist dimension. This would explain why the Intra-populations suffered low fitness, ultimately resulting in the extinction of three of eight populations. It also seems likely that the overall low trait values in the remaining populations may have partly occurred due to drift or inbreeding brought about by high demographic costs. In populations of larger size, the typical case for small invertebrates, rare mutations needed for adaptation might have been available (Bell & Gonzales, 2009; Chevin *et al.*, 2013). Nevertheless, our results indicate relatively severe genetic constraints on evolution in the generalist–specialist dimension compared to specialization to hot or cold temperature that evolved readily from standing genetic variation.

Multivariate trade-offs and temperature adaptation

A trade-off between juvenile growth and survival was supported here by negative cross-temperature genetic correlations (Table 2). The Cold populations had a comparatively larger body size but lower survival than the Hot populations, suggesting that the optimal solution to the growth–survival trade-off varies with the thermal regime (Fig. 1b–d). Temperature-specific trade-offs can result from differences in the strength of selection on individual fitness components across temperatures. Body size has been shown to be under strong positive selection in both sexes in this population (Puniamoorthy *et al.*, 2012a), and reduced size in favour of increased juvenile survival is likely to come at a cost. However, we do not know whether selection on adult size differs across temperatures in this species, or in general (Stillwell *et al.*, 2010). Our data are nevertheless indicative of selection for increased juvenile survival at the expense of reduced size at variable and hot temperatures.

A possible explanation for the larger body size but lower survival in the Cold populations could be that cold temperatures usually slow growth substantially, which might increase the relative strength of selection on growth efficiency. This hypothesis is supported by a detailed study quantifying the strength of selection on growth rate in larvae of the butterfly *Pieris rapae* across a broad thermal range (Kingsolver *et al.*, 2007), and more generally by the observation that northern populations often evolve faster growth rates compared with their southern conspecifics (*counter-gradient variation*: Conover & Schultz, 1995; Clarke, 2003; Grether, 2005; Blanckenhorn *et al.*, 2006; Yamahira *et al.*, 2007; Berger *et al.*, 2013). The larger body size of Cold populations is also in line with the observation that genotypes having evolved in cold conditions often have larger body size (Angilletta & Dunham,

2003) and that a majority of ectotherms express larger body sizes as direct plastic responses to cold temperature (*the temperature-size rule*: Atkinson & Sibly, 1997). It is thus possible that the major selective force brought about by the Cold selection regime was on growth and large size rather than on survival, whereas the situation might have been reversed in the populations that experienced 31 °C, which lies close to the physiological tolerance limit (34 °C) but allows growth rate maximization (see Berger *et al.*, 2013; Fig. 2f).

Our results are in agreement with the assertion that the evolution of plasticity in general (Ghalambor *et al.*, 2007; Callahan *et al.*, 2008), and thermal adaptation in particular (Angilletta *et al.*, 2003; Clarke, 2003; Foray *et al.*, 2014), often encompasses multiple fitness components, thus potentially masking trade-offs at the level of single-trait reaction norms. Furthermore, our study presents direct evidence for how antagonistic pleiotropy for fitness across temperatures can become apparent as a result of temperature-specific experimental evolution of its underlying components (here juvenile survival and size at maturity). We quantified a trade-off signified by cross-temperature negative broad-sense genetic correlations involving only two traits (Table 2). However, bivariate genetic correlations often do not present insurmountable genetic constraints on adaptation (Agrawal & Stinchcombe, 2009), and therefore, given widespread pleiotropy of life-history genes, it seems likely that if more traits had been measured, more phenotypic dimensions underlying the antagonistic pleiotropy and maintenance of the cross-temperature correlation between juvenile survival and size at maturity may have been revealed (Walsh & Blows, 2009; Conner, 2012).

The sex-specific responses in body size, observable in our study as phenotypic plasticity and genetic changes in the selection regimes, further suggest that temperature can induce changes in the relative strength of natural and sexual selection on different fitness components (Ketola *et al.*, 2012; Berger *et al.*, 2014). This may lead to sex-specific solutions to temperature-dependent multivariate trade-offs (Rogell *et al.*, 2013), ultimately resulting in evolutionary changes in sexual dimorphism (Blanckenhorn *et al.*, 2006; Stillwell *et al.*, 2010). The sex that experiences stronger sexual selection usually shows more condition dependence and environmental sensitivity in those characters selected (Rowe & Houle, 1996; Bonduriansky, 2007). Male-biased sexual size dimorphism in European *S. punctum* has likely evolved through intense intra- and inter-sexual selection in males (Dmitriev & Blanckenhorn, 2012; Puniamoorthy *et al.*, 2012a). The condition-dependence hypothesis thus fits with our findings showing that male body size is more sensitive (plastic) to temperature and also shows more temperature-specific evolution (Fig. S2).

Fine-grained variation and the hotter-is-better hypothesis

If costs of plasticity are nontrivial as suggested by our experiment, environments composed of fine-grained temporal variation like our Intra-treatment may select for specialists (Gilchrist, 1995). In such cases, selection should favour specialization to the most productive conditions (Van Tienderen, 1991; Kingsolver & Gomulkiewicz, 2003). The *hotter-is-better* hypothesis states that, due to evolutionary histories in warm environments, ectotherm physiology works best at relatively warm temperatures and fitness should be higher overall in species adapted to warm environments (Frazier *et al.*, 2006; Knies *et al.*, 2009; Walters *et al.*, 2012). Interestingly, the Intra-selection regime led to apparent biased adaptation to 31 °C (Fig. 3), consistent with the 'hotter-is-better' hypothesis. However, due to the low replication of this selection treatment, statistical power was low, and with the data at hand, we are not able to fully evaluate this hypothesis. Interestingly, experimental evolution for generalists and specialists in *E. coli* led to biased adaptation to hot temperatures (Bennett *et al.*, 1992), and unicellular chlorophytes have shown biased adaptation to light (the more productive environment) relative to dark (Kassen & Bell, 1998), in agreement with the pattern reported here, supporting the prediction that fine-grained environmental variation may select for specialization to the most productive conditions.

Non-beneficial transgenerational acclimation

Acclimation may incur benefits whenever the current environment acts as a reliable cue for future conditions (Gabriel & Lynch, 1992; Fox & Mousseau, 1998; Gabriel, 2005) and can increase population persistence by reducing demographic costs under an initial phase of adaptation (Chevin *et al.*, 2013; Lande, 2014). Although our experimental design allowed assessment of parental acclimation effects, these were idiosyncratic and overall not consistent with beneficial acclimation. For example, acclimation of the parental generation at 23 °C seemed to reduce differences in offspring survival between selection treatments compared to when survival was measured of offspring originating from parents raised in their respective selection regime (Fig. 2a vs. b), implying that genetic adaptation and parental effects acted in the same direction to induce phenotypic change. Contrary to this result, however, parental acclimation at 23 °C instead increased differences in body size between selection treatments as compared to when offspring originated from parents raised in their respective selection regimes (Fig. 2c vs. d), indicating that parental effects acted to canalize body size across test temperatures.

Studies investigating transgenerational acclimation often fail to demonstrate beneficial effects (but for a con-

vincing example, see Salinas & Munch, 2012). In fact, parents acclimated to ancestral (benign) temperatures usually produce offspring that have higher fitness at temperature extremes than offspring produced by parents acclimated to the extreme temperature (Leroi *et al.*, 1994; Hoffmann, 1995; Bennett & Lenski, 1997; Woods & Harrison, 2002). Acclimation responses in the offspring in situations where the temperature cues used by the parents to predict future conditions are in themselves directly detrimental are thus often the result of both adaptive shifts in offspring phenotypes in response to the cue, and detrimental effects associated with poor provisioning by low-condition parents. Our experimental treatments were designed to be stressful to the parents, so the above explanation probably applies to our results.

Conclusions

Theory predicts that antagonistic pleiotropy should play a leading role in shaping thermal performance curves. Nevertheless, antagonistic pleiotropy and generalist–specialist trade-offs are seldom confirmed in experimental evolution and comparative studies. We have demonstrated standing allelic variation at loci with antagonistic effects on juvenile performance at hot and cold temperatures, by obtaining corresponding evolutionary responses to experimental evolution for a natural population of black scavenger flies under conditions that limited the opportunity to increase larval resource acquisition. Our study illustrates that under these conditions genetic constraints on reaction norms may manifest in multivariate allocation trade-offs across temperatures. However, despite antagonistic pleiotropy dominating the evolutionary responses by driving specialization to hot and cold temperatures, respectively, we found no evidence for the evolution of thermal performance breadth at the expense of reductions in maximal performance, suggesting low genetic variance in the generalist–specialist dimension.

Acknowledgments

We thank N. Puniamoorthy and M. Schäfer for scientific discussions and advice on experimental design and P. Wadewitz, J. Roy, U. Briegel and Y. Choffat for assistance in the laboratory. This work was supported by a grant from The Swedish Research Council (VR) to DB, a mobility grant from the European Science Foundation (ESF) Programme 'ThermAdapt' to RW and a grant from the Swiss National Fund (SNF) to WUB.

References

- Agrawal, A.F. & Stinchcombe, J.R. 2009. How much do genetic covariances alter the rate of adaptation? *Proc. R. Soc. B* **276**: 1183–1191.

- Angilletta, M.J. 2009. *Thermal Adaptation: A Theoretical and Empirical Synthesis*. Oxford University Press, Oxford.
- Angilletta, M.J. & Dunham, A.E. 2003. The temperature-size rule in ectotherms: simple evolutionary explanations may not be general. *Am. Nat.* **162**: 332–342.
- Angilletta, M.J., Wilson, R.S., Navas, C.A. & James, R.S. 2003. Tradeoffs and the evolution of thermal reaction norms. *Trends Ecol. Evol.* **18**: 234–240.
- Angilletta, M.J., Huey, R.B. & Frazier, M.R. 2010. Thermodynamic effects on organismal performance: is hotter better? *Physiol. Biochem. Zool.* **83**: 197–206.
- Atkinson, D. & Sibly, R.M. 1997. Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends Ecol. Evol.* **12**: 235–239.
- Bates, D., Maechler, M. & Bolker, B. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42. <http://CRAN.R-project.org/package=lme4>
- Bell, G. & Gonzales, A. 2009. Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.* **12**: 942–948.
- Bennett, A.F. & Lenski, R.E. 1993. Evolutionary adaptation to temperature II. Thermal niches of experimental lines of *Escherichia coli*. *Evolution* **47**: 1–12.
- Bennett, A.F. & Lenski, R.E. 1997. Evolutionary adaptation to temperature. VI. Phenotypic acclimation and its evolution in *Escherichia coli*. *Evolution* **51**: 36–44.
- Bennett, A.F. & Lenski, R.E. 2007. An experimental test of evolutionary trade-offs during temperature adaptation. *Proc. Natl. Acad. Sci. USA* **104**: 8649–8654.
- Bennett, A.F., Lenski, R.E. & Mittler, J.E. 1992. Evolutionary adaptation to temperature I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution* **46**: 16–30.
- Berger, D., Bauerfeind, S.S., Blanckenhorn, W.U. & Schäfer, M.A. 2011a. High temperatures reveal cryptic genetic variation in a polymorphic female sperm storage organ. *Evolution* **65**: 2830–2842.
- Berger, D., Friberg, M. & Gotthard, K. 2011b. Divergence and ontogenetic coupling of larva behavior and thermal reaction norms in three closely related butterflies. *Proc. R. Soc. Lond. B* **278**: 313–320.
- Berger, D., Postma, E., Blanckenhorn, W.U. & Walters, R.J. 2013. Quantitative genetic divergence and standing genetic (co)variance in thermal reaction norms along latitude. *Evolution* **67**: 2385–2399.
- Berger, D., Grieshop, K., Lind, M.I., Goenaga, J., Maklakov, A.A. & Arnqvist, G. 2014. Intralocus sexual conflict and environmental tolerance. *Evolution*. doi/10.1111/evo.12439.
- Blanckenhorn, W.U. 2000. The evolution of body size: What keeps organisms small? *Q. Rev. Biol.* **75**: 388–407.
- Blanckenhorn, W.U., Stillwell, R.C., Young, K.A., Fox, C.W. & Ashton, K.G. 2006. When Rensch meets Bergmann: does sexual size dimorphism change systematically with latitude? *Evolution* **60**: 2004–2011.
- Blanckenhorn, W.U., Pemberton, A.J., Bussière, L.F., Roemke, J. & Floate, K.D. 2010. Natural history and laboratory culture of the yellow dung fly, *Scathophaga stercoraria* (L.; Diptera: Scathophagidae). *J. Insect Sci.* **10**: 1–17.
- Bonduriansky, R. 2007. The evolution of condition-dependent sexual size dimorphism. *Am. Nat.* **169**: 9–19.
- Callahan, H.S., Maughan, H. & Steiner, U.K. 2008. Phenotypic plasticity, costs of phenotypes, and costs of plasticity. Toward an integrative view. *Ann. N.Y. Acad. Sci.* **1133**: 44–66.
- Chevin, L., Gallet, R., Gomulkiewicz, R., Holt, R.D. & Fellous, S. 2013. Phenotypic plasticity in evolutionary rescue experiments. *Philos. Trans. R. Soc. B* **368**: 20120089.
- Clarke, A. 2003. Costs and consequences of evolutionary temperature adaptation. *Trends Ecol. Evol.* **18**: 573–581.
- Condon, C., Cooper, B.S., Yeaman, S. & Angilletta, M.J. 2013. Temporal variation favors the evolution of generalists in experimental populations of *Drosophila melanogaster*. *Evolution* **68**: 720–728.
- Conner, J.K. 2012. Quantitative genetic approaches to evolutionary constraint: how useful? *Evolution* **66**: 3313–3320.
- Conover, D.O. & Schultz, E.T. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends Ecol. Evol.* **10**: 248–252.
- Cooper, B.S., Hammad, L.A., Fisher, N.P., Karty, J.A. & Montooth, K.L. 2012. In a variable thermal environment selection favors greater plasticity of cell membranes in *Drosophila melanogaster*. *Evolution* **66**: 1976–1984.
- Crispo, E. 2008. Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *J. Evol. Biol.* **21**: 1460–1469.
- DeWitt, T.J., Sih, A. & Wilson, D.S. 1998. Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* **13**: 77–81.
- Diamond, S.E. & Kingsolver, J.G. 2010. Environmental dependence of thermal reaction norms: host plant quality can reverse the temperature-size rule. *Am. Nat.* **175**: 1–10.
- Dmitriew, C.M. 2011. The evolution of growth trajectories: what limits growth rate? *Biol. Rev.* **86**: 97–116.
- Dmitriew, C.M. & Blanckenhorn, W.U. 2012. The role of sexual selection and conflict in mediating among-population variation in mating strategies and sexually dimorphic traits in *Sepsis punctum*. *PLoS ONE* **7**: e49511.
- Elena, S.F. & Lenski, R.E. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4**: 457–469.
- Foray, V., Desouhant, E. & Gibert, P. 2014. The impact of thermal fluctuations on reaction norms in specialist and generalist parasitic wasps. *Funct. Ecol.* **28**: 411–423.
- Fox, C.W. & Mousseau, T.A. 1998. Maternal effects as adaptation for transgenerational phenotypic plasticity in insects. In: *Maternal Effects as Adaptations*, Ch 10 (T.A. Mousseau & C.W. Fox, eds), pp. 159–177. Oxford University Press, New York.
- Frazier, M.R., Huey, R.B. & Berrigan, D. 2006. Thermodynamics constrain the evolution of insect population growth rates: ‘warmer is better’. *Am. Nat.* **168**: 512–520.
- Gabriel, W. 2005. How stress selects for reversible phenotypic plasticity. *J. Evol. Biol.* **18**: 873–883.
- Gabriel, W. & Lynch, M. 1992. The selective advantage of reaction norms for environmental tolerance. *J. Evol. Biol.* **5**: 41–59.
- Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* **21**: 394–407.
- Gilchrist, G.W. 1995. Specialists and generalists in changing environments. I. Fitness landscapes of thermal sensitivity. *Am. Nat.* **146**: 252–270.

- Grether, G.F. 2005. Environmental change, phenotypic plasticity, and genetic compensation. *Am. Nat.* **166**: E115–E123.
- Hadfield, J.D. 2008. Estimating evolutionary parameters when viability selection is operating. *Proc. R. Soc. B* **275**: 723–734.
- Hochachka, P.W. & Somero, G.N. 2002. *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. Oxford University Press, Oxford.
- Hoffmann, A.A. 1995. Acclimation: increasing survival at a cost. *Trends Ecol. Evol.* **10**: 1–2.
- Hoffmann, A.A. 2010. Physiological climatic limits in *Drosophila*: patterns and implications. *J. Exp. Biol.* **213**: 870–880.
- Houle, D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution* **45**: 630–648.
- Huey, R.B. & Kingsolver, J.G. 1989. Evolution of thermal sensitivity of ectotherm performance. *Trends Ecol. Evol.* **4**: 131–135.
- Huey, R.B. & Kingsolver, J.G. 1993. Evolution of resistance to high temperature in ectotherms. *Am. Nat.* **142**: S21–S46.
- Hughes, B.S., Cullum, A.J. & Bennett, A.F. 2007. An experimental evolutionary study on adaptation to temporally fluctuating pH in *Escherichia coli*. *Physiol. Biochem. Zool.* **80**: 406–421.
- Izem, R. & Kingsolver, J.G. 2005. Variation in continuous reaction norms: quantifying directions of biological interest. *Am. Nat.* **166**: 277–289.
- de Jong, G. & van Noordwijk, A.J. 1992. Acquisition and allocation of resources: genetic (co)variances, selection, and life histories. *Am. Nat.* **139**: 749–770.
- Kassen, R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. *J. Evol. Biol.* **15**: 173–190.
- Kassen, R. & Bell, G. 1998. Experimental evolution in *Chlamydomonas*. IV. Selection in environments that vary through time and space. *Heredity* **80**: 732–741.
- Ketola, T., Kristensen, T.N., Kellermann, V.M. & Loeschcke, V. 2012. Can evolution of sexual dimorphism be triggered by developmental temperatures? *J. Evol. Biol.* **25**: 847–855.
- Kingsolver, J.G. & Gomulkiewicz, R. 2003. Environmental variation and selection on performance curves. *Integr. Comp. Biol.* **43**: 470–477.
- Kingsolver, J.G. & Huey, R.B. 2008. Size, temperature, and fitness: three rules. *Evol. Ecol. Res.* **10**: 251–268.
- Kingsolver, J.G., Massie, K.R., Schlichta, J.G., Smith, M.H., Rangeland, G.J. & Gomulkiewicz, R. 2007. Relating environmental variation to selection on reaction norms: and experimental test. *Am. Nat.* **169**: 163–174.
- Klepsatel, P., Galikova, M., De Maio, N., Huber, C.D., Schlötterer, C. & Flatt, T. 2013. Variation in thermal performance and reaction norms among populations of *Drosophila melanogaster*. *Evolution* **67**: 3573–3587.
- Knies, J.L., Kingsolver, J.G. & Burch, C.L. 2009. Hotter is better and broader: thermal sensitivity of fitness in a population of bacteriophages. *Am. Nat.* **173**: 419–430.
- Lande, R. 2014. Evolution of phenotypic plasticity and environmental tolerance of a labile quantitative character in a fluctuating environment. *J. Evol. Biol.* **27**: 866–875.
- Latimer, C.A.L., Wilson, R.S. & Chenoweth, S.F. 2011. Quantitative genetic variation for thermal performance curves within and among natural populations of *Drosophila serrata*. *J. Evol. Biol.* **24**: 965–975.
- Latimer, C.A., McGuigan, K., Wilson, R.S., Blows, M.W. & Chenoweth, S.F. 2014. The contribution of spontaneous mutations to thermal sensitivity curve variation in *Drosophila serrata*. *Evolution* **68**: 1824–1837.
- Leroi, A.M., Bennett, A.F. & Lenski, R.E. 1994. Temperature acclimation and competitive fitness: an experimental test of the beneficial acclimation hypothesis. *Proc. Natl. Acad. Sci. USA* **91**: 1917–1921.
- Levins, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton, NJ.
- Lynch, M. & Gabriel, W. 1987. Environmental tolerance. *Am. Nat.* **129**: 283–303.
- Martin, G. & Lenormand, T. 2006. The fitness effect of mutations across environments: a survey in light of fitness landscape models. *Evolution* **60**: 2413–2427.
- Murren, C.J., Maclean, H.J., Diamond, S.E., Steiner, U.K., Heskell, M.A., Handelsman, C.A. et al. 2014. Evolutionary change in continuous reaction norms. *Am. Nat.* **183**: 453–467.
- Nilsson-Örtman, V., Stoks, R. & Johansson, F. 2014. Competitive interactions modify the temperature dependence of damselfly growth rates. *Ecology* **95**: 1394–1406.
- Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going? *Trends Ecol. Evol.* **20**: 481–486.
- Pont, A.C. & Meier, R. 2002. The Sepsidae (Diptera) of Europe. *Fauna Ecol. Scand.* **37**: 1–222.
- Puniamoorthy, N., Schäfer, M.A. & Blanckenhorn, W.U. 2012a. Sexual selection accounts for the geographic reversal of sexual size dimorphism in the dung fly *Sepsis punctum* (Diptera: Sepsidae). *Evolution* **66**: 2117–2126.
- Puniamoorthy, N., Blanckenhorn, W.U. & Schaefer, M.A. 2012b. Differential investment in pre- versus post-copulatory sexual selection reinforces a cross-continental reversal of sexual size dimorphism in *Sepsis punctum* (Diptera: Sepsidae). *J. Evol. Biol.* **25**: 2253–2263.
- R Development Core Team. 2011. *R Foundation for Statistical Computation*. R Development Core Team, Vienna, Austria.
- Reboud, X. & Bell, G. 1997. Experimental evolution in *Chlamydomonas*. III. Evolution of specialist and generalist types in environments that vary in space and time. *Heredity* **78**: 507–514.
- Relyea, R.A. 2002. Costs of phenotypic plasticity. *Am. Nat.* **159**: 272–282.
- Rogell, B., Widegren, W., Hallsson, L., Berger, D., Björklund, M. & Maklakov, A.A. 2013. Sex-dependent evolution of life-history traits following adaptation to climate warming. *Funct. Ecol.* **28**: 469–478.
- Rowe, L. & Houle, D. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. *Proc. R. Soc. Lond. B* **263**: 1415–1421.
- Salinas, S. & Munch, S.B. 2012. Thermal legacies: transgenerational effects of temperature on growth in a vertebrate. *Ecol. Lett.* **15**: 159–163.
- Scheiner, S.M. 2002. Selection experiments and the study of phenotypic plasticity. *J. Evol. Biol.* **15**: 889–898.
- Snell-Rodd, E.C., Van Dyken, J.D., Cruickshank, T., Wade, M.J. & Moczek, A.P. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays* **32**: 71–81.
- Stillwell, R.C., Blanckenhorn, W.U., Teder, T., Davidowitz, G. & Fox, C.W. 2010. Sex differences in phenotypic plasticity

- affect variation in sexual size dimorphism in insects: from physiology to evolution. *Annu. Rev. Entomol.* **55**: 227–245.
- Sultan, S.E. & Spencer, H.G. 2002. Metapopulation structure favors plasticity over local adaptation. *Am. Nat.* **160**: 271–283.
- Van Buskirk, J. & Steiner, U.K. 2009. The fitness costs of developmental canalization and plasticity. *J. Evol. Biol.* **22**: 852–860.
- Van Tienderen, P.H. 1991. Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution* **45**: 1317–1331.
- Via, S. & Lande, R. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**: 505–522.
- Walsh, B. & Blows, M.W. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Ann. Rev. Ecol. Evol. Syst.* **40**: 41–59.
- Walters, R.J., Blanckenhorn, W.U. & Berger, D. 2012. Forecasting extinction risk of ectotherms under climate warming: an evolutionary perspective. *Funct. Ecol.* **26**: 1324–1338.
- West- Eberhard, M.J. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, New York, NY.
- Whitlock, M.C. 1996. The red queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. *Am. Nat.* **148**: S65–S77.
- Woods, H.A. & Harrison, J.F. 2002. Interpreting rejections of the beneficial acclimation hypothesis: when is physiological plasticity adaptive? *Evolution* **56**: 1863–1866.
- Yamahira, K., Kawajiri, M., Takeshi, K. & Irie, T. 2007. Inter- and intrapopulation variation in thermal reaction norms for growth rate: evolution of latitudinal compensation in ectotherms with a genetic constraint. *Evolution* **61**: 1577–1589.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Effect of parental acclimation on offspring survival, body size and growth rate for populations in the selection regime experiencing inter-generational temperature variation.

Figure S1 Effect of parental acclimation on offspring survival, body size (mm hind tibia length) and growth rate for populations in the selection regime experiencing inter-generational temperature variation.

Table S2 Sex-specific effects on body size.

Figure S2 Sex-specific effects of selection regime and test temperature on body size (mm hind tibia length) in Hot and Cold populations.

S3 Differences in mean trait values.

Figure S3 Arithmetic means of relative scores at each temperature, for survival at 12–34 °C, and for body size and growth rate at 15–31 °C, following 20 generations of laboratory evolution and one full generation of acclimation at 23 °C.

S4 Correlation matrix for all traits (development rate, growth rate, body size and survival) across temperatures based on population means.

S5 Relative differences (temperature-specific mean-standardized data) between selection treatments.

S6 Pairwise comparisons of Hot & Cold, or Inter & Intra populations for the three traits. Analyses were performed in the same way as the main analyses described in the main article but for the subsets of data.

S7 Response in development rate (1/development time) ± 1 SE.

Data deposited at Dryad: doi:10.5061/dryad.77m77

Received 1 February 2014; revised 17 June 2014; accepted 23 June 2014