doi:10.1111/evo.12711

Local divergence of thermal reaction norms among amphibian populations is affected by pond temperature variation

Alex Richter-Boix,[1,2] Marco Katzenberger,[3] Helder Duarte,[3] María Quintela,[4,5] Miguel Tejedo,[3] and Anssi Laurila[1]

[1]*Animal Ecology/Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236 Uppsala, Sweden*

[2]*E-mail: richterboixalex@gmail.com*

[3]*Department of Evolutionary Ecology, Estación Biológica de Doñana, CSIC, Avda. Américo Vespucio s/n, 41092 Sevilla, Spain*

[4]*Grupo de Investigación BIOCOST, University of Coruña, 15071, Coruña, Spain*

[5]*Current Address: Population Genetics, Institute of Marine Research, 5817, Bergen, Norway*

Received December 9, 2014
Accepted June 15, 2015

Although temperature variation is known to cause large-scale adaptive divergence, its potential role as a selective factor over microgeographic scales is less well-understood. Here, we investigated how variation in breeding pond temperature affects divergence in multiple physiological (thermal performance curve and critical thermal maximum [CT_{max}]) and life-history (thermal developmental reaction norms) traits in a network of *Rana arvalis* populations. The results supported adaptive responses to face two main constraints limiting the evolution of thermal adaptation. First, we found support for the faster–slower model, indicating an adaptive response to compensate for the thermodynamic constraint of low temperatures in colder environments. Second, we found evidence for the generalist–specialist trade-off with populations from colder and less thermally variable environments exhibiting a specialist phenotype performing at higher rates but over a narrower range of temperatures. By contrast, the local optimal temperature for locomotor performance and CT_{max} did not match either mean or maximum pond temperatures. These results highlight the complexity of the adaptive multiple-trait thermal responses in natural populations, and the role of local thermal variation as a selective force driving diversity in life-history and physiological traits in the presence of gene flow.

KEY WORDS: **Amphibians, countergradient variation, CT_{max}, ectotherms, F_{ST}–Q_{ST}, generalist–specialist trade-off, hotter–colder, thermal adaptation, thermal fluctuations, thermal performance curves.**

Temperature is an important physical environmental parameter acting directly on the whole-organism physiology and having a pivotal role in ecology and evolution (Ward and Stanford 1982; Clarke 2003; Angilletta 2009). Ambient temperature varies in both space and time, and identifying the geographical patterns of thermal adaptation may help to predict the scale of the evolutionary responses and thus the mitigation potential of organisms coping with environmental and climate shifts (Williams et al. 2008). Adaptation to large-scale thermal variation—often latitudinal or altitudinal—is well documented in a range of organisms (Endler 1986; Angilletta 2009; Conover et al. 2009; Huey et al. 2009; Berger et al. 2013). However, thermal conditions show marked small-scale spatial and temporal variation, which can lead to adaptive divergence among local populations in thermal performance (Blanckenhorn 1991; Kawecki and Ebert 2004; Keller and Seehausen 2012; Blanquart et al. 2013). There is little information on how upper thermal resistance and thermal performance traits diverge in response to local variation in either average or

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

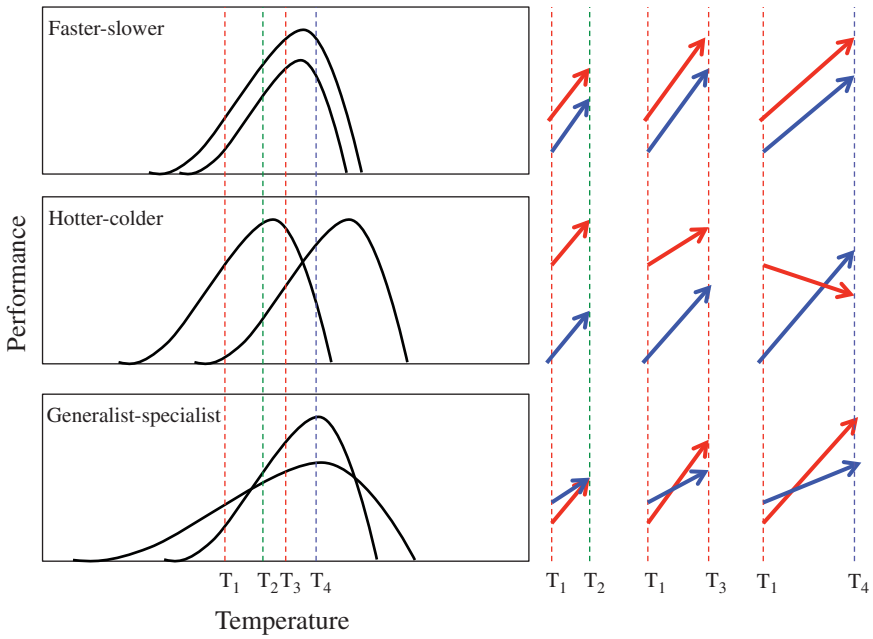


Figure 1. **A graphical illustration of the difficulty to understand evolutionary change in reaction norms and the patterns of thermal plasticity in a two-temperature experiment when thermal reaction norms are concave in shape. The left panel shows the three thermal models of optimal performance curves: vertical shift in average fitness (faster–slower), horizontal shift in optimal temperature (hotter–colder), and the generalist–specialist variation (niche width). The arrows in the right show the observed reaction norms in hypothetical two-temperature experiments considering different temperature ranges (T1–T2, T1–T3, or T1–T4). The faster–slower model is easy to identify because this mode represents a simple linear translation around the mean curve of the different population, with parallel reaction norms. However, both hotter–colder and generalist–specialist models show differences in slope and/or crossing reaction norms and, consequently, it is hard to discriminate between the models with just linear models obtained from a two-temperature experiment.**

peak temperatures (Skelly and Freidenburg 2000; Eliason et al. 2011; Yampolsky et al. 2014). Information on these evolutionary responses is important both for understanding the present-day diversity in thermal performance and for predicting evolutionary responses to future environmental change (Knies et al. 2009; Chevin et al. 2010; Keller and Seehausen 2012).

Thermal performance curve (TPC), a continuous reaction norm in which an organism's performance (i.e., growth, development, locomotor ability) is described as a function of temperature, is a highly useful tool to examine the divergence in thermal performance (Huey and Stevenson 1979). TPCs are typically concave in shape, with a central or right-skewed maximum and lower trait values at both low and high temperatures (Huey and Stevenson 1979; Knies et al. 2006; Angilletta 2009). The shape of the curve is described by several parameters: T_{opt}, the temperature at which performance is maximal (Z_{max}); B_i, the temperature over which performance is above some arbitrary level (i); and the critical thermal limits that permit performance (CT_{min} and CT_{max}, Angilletta et al. 2002).

Alterations in these parameters (especially in T_{opt}, B_i, and Z_{max}) modify the shape of the TPC, and evolutionary thermal physiologists have proposed three modes of variation in the curve shape that allow unraveling evolutionary mechanisms and possible trade-offs underlying thermal adaptations (Fig. 1; Huey and Kingsolver 1989; Kingsolver et al. 2001; Knies et al. 2006). (1) Vertical shifts (faster–slower) produce changes in average performance. This mode of thermal adaptation arises from selection opposing the influence of the environmental gradient on the trait (countergradient selection; Conover and Schultz 1995; Conover et al. 2009). Although the resulting genotypes (masters of all temperatures) are not predicted by optimality models of thermal evolution (Lynch and Gabriel 1987; Gilchrist 1995), there is ample empirical evidence for them across ectotherm taxa (Angilletta 2009; Conover et al. 2009). This mode of adaptation is often considered an evolutionary response to seasonality, in which variation in length of the growing season constrains the time available for biological processes (Conover and Schultz 1995; Yamahira et al. 2007; Richter-Boix et al. 2010). The predicted adaptive response is that time-constrained populations, basically those living in colder environments (high latitude, altitude or colder microhabitat), will perform better than populations living in less time-constrained environments. (2) Horizontal shifts in optimal temperature (hotter–colder). The "hotter–colder" model predicts a shift in thermal reaction norms toward a higher or lower temperature range in accordance with the temperatures most often experienced in nature (Angilletta 2009), and is supported by

both theoretical models and empirical data (Lynch and Gabriel 1987; Gilchrist 1995). This mode inherently imposes a thermodynamic constraint in thermal adaptation, the "hotter is better," which implies that cold-adapted populations will be outperformed by the warm-adapted counterparts. Considering a continuous norm of reaction, hotter is better predicts a positive correlation between an organism's optimal temperature and its maximum performance (Huey and Kingsolver 1989; Frazier et al. 2006). (3) The "generalist–specialist" variation produces changes in niche width, indicating a trade-off between maximal performance and performance breadth (Huey and Hertz 1984) based on structural constraints in the thermal flexibility and stability of enzymes (Huey and Kingsolver 1989). In this case, temporal variation in environmental conditions is assumed to be the major determinant for the evolution of thermal performance breadth, and temporal temperature variation within generations is considered more important than thermal variation between generations (Lynch and Gabriel 1987; Huey and Kingsolver 1989).

Empirical studies on thermal adaptation are largely based on studies examining variation among populations across latitudinal or altitudinal gradients, in which both mean temperature and seasonality vary concomitantly, especially in higher latitudes (Baumann and Conover 2011; Higgins et al. 2014). The confounding effects of seasonality on mean and range of temperature, and the effect of temporal temperature variation within versus between generations have yet to be disentangled in studies of thermal adaptation (Baumann and Conover 2011). Previous studies on thermal reaction norms in anuran tadpoles have detected countergradient variation over large (e.g., Berven 1982; Laugen et al. 2003; Orizaola et al. 2010) and small (e.g., Skelly 2004; Richter-Boix et al. 2010; Muir et al. 2014a) spatial scales, suggesting that at large spatial scale time constraints rather than mean temperature are the main force driving thermal adaptation, whereas at smaller scales differences in temperature also create time constraints that may explain the observed countergradient variation. However, most of the previous studies considered only growth and development (but see Skelly and Freidenburg 2000; Freidenburg and Skelly 2004; Lindgren and Laurila 2005, 2009; Muir et al. 2014b for physiological studies) and used simple linear reaction norms, which lack the precision to distinguish between the alternative modes of variation (Fig. 1; e.g., Izem and Kingsolver 2005; Murren et al. 2014).

In this study, we asked how variation in temperature mean and range affects thermal performance in a network of closely located moor frog *Rana arvalis* populations. All populations included in this study face a similar seasonal time constraint in terms of latitudinal position; however, the breeding ponds vary in temperature mean and variance due to microgeographic variation in forest canopy cover among the ponds. Ponds with high canopy cover have lower and less variable water temperatures than ponds with less forest canopy (Richter-Boix et al. 2010, 2011a, 2013). This fine spatial scale allowed us to study variation in thermal reaction norms that is predominantly related to microclimatic temperature variation beyond the seasonal effects expected across a large-scale latitudinal gradient (S1). Our previous studies found pronounced genetic divergence in larval developmental and growth rates among the populations in the presence of ample gene flow, and higher developmental and growth rates in populations with more canopy cover and lower water temperature, suggesting adaptive countergradient variation (Richter-Boix et al. 2011a, 2013). As adaptation occurred at spatial scales at which populations should experience high gene flow (Richter-Boix et al. 2013), the present system fulfills the criteria for microgeographic adaptation as defined by Richardson et al. (2014). However, in the previous studies we only examined this divergence under only a single thermal regime.

We investigated population divergence in thermal performance using a combination of experimental approaches by determining (1) two-point thermal developmental reaction norms (TDRN) for larval life-history traits, (2) continuous thermal locomotor performance curves (TPCs; six temperatures between 10 and 32ºC) for maximum tadpole swimming speed, and (3) upper thermal tolerance (CT_{max}) in seven *R. arvalis* populations. The seven populations were selected from the network of 17 sites (Richter-Boix et al. 2013) to capture variation across the forest canopy cover gradient from ponds with completely open canopy to wetlands with almost closed (80%) canopy. Moreover, the selected wetlands had similar breeding times to avoid the possible maternal effects on development and growth mediated by breeding time (Richter-Boix et al. 2014).

We test the following, not mutually exclusive, hypotheses under a scenario of thermal adaptation to local conditions: (1) optimal temperature and the upper thermal tolerance will match the prevailing environmental mean and maximum temperatures (Lynch and Gabriel 1987; Gilchrist 1995). We expect to find positive correlations between T_{opt} and pond mean temperature, and between CT_{max} and maximum pond temperature. In this case, we expect that hotter–colder mode of variation explains a considerable amount of variation IN TPC curves (Knies et al. 2006). (2) As the ponds differ in the amount of thermal variation, we expect more generalist phenotypes in open ponds and more specialist phenotypes in less thermally variable ponds with more canopy cover. We test this hypothesis by correlating thermal variation with locomotor performance breadth (B_{80} and B_{95}) and examining the variation explained by the generalist–specialist mode of variation. (3) Based on hotter is better thermodynamic constraint, we could expect poorer performers at cold ponds in the case that compensatory adaptive responses do not counteract the constraint (Frazier et al. 2006). We tested this prediction by correlating T_{opt} with Z_{max}. Here a positive trend will support the

Table 1. Breeding time (collection day of the freshly laid clutches in 2009), temperature variables, and forest canopy cover for the seven breeding sites studied.

	s1	s2	s3	s4	s5	s6	s7
Latitude	59°50'N	59°36'N	59°43'N	59°42'N	59°47'N	59°44'N	59°37'N
Longitude	17°21'E	16°55'E	16°59'E	17° 4'E	16°52'E	16°50'E	17°13'E
Breeding time	16/4	14/4	10/4	13/4	16/4	10/4	14/4
Mean temperature	9.44	11.14	10.74	10.41	10.61	13.24	12.31
Mean minimum temperature	2.81	5.95	3.44	0.47	2.14	5.46	0.92
Mean maximum temperature	16.14	17.99	22.58	22.12	23.72	25.78	25.41
Mean range temperature	13.34	12.04	19.14	21.64	21.57	20.32	24.48
Forest canopy (%)	80	80	60	40	20	0	0

Temperature variables correspond to the mean values of local temperature during the spring in 2008–2010.

hotter is better hypothesis (Huey and Kingsolver 1989), whereas no or a negative correlation suggests that the thermodynamic constraint is circumvented (Clarke 2003; Latimer et al. 2011). (4) Finally from (2) we could expect that lower variability in colder ponds would select for individuals with higher performance (Z_{max}) as a correlated response to the trade-off between thermal breadth and maximal performance. We tested this prediction by examining the amount of variance explained in the "faster–slower" mode of variation for the locomotor TPC, and expecting to find better performance in growth and developmental traits for cold populations at both temperatures in the TDRN experiment. To summarize, this multiple-trait experimental approach allows us to test and demonstrate divergent thermal performance in spatially and temporally heterogeneous thermal environments in the presence of gene flow, and to identify some of the evolutionary constraints and trade-offs associated with thermal variation.

Materials and Methods

STUDY AREA AND FIELD METHODS

We sampled about 30 eggs from each of the 10 freshly laid clutches (families) at seven *R. arvalis* breeding sites in Uppsala county (central Sweden) between 10 and 16 April 2009. These breeding sites are a subsample of a 17-site network located within a 40 × 40 km area used in previous studies (see Richter-Boix et al. 2011a, 2013). The mean distance between the seven selected sites was 15.58 (±8.33 [SD]) km (see S2 for whole distance matrix). The eggs were transported to the laboratory in Uppsala University, in which each family was kept in a separate 3-l plastic vial in a walk-in climate-controlled room (15ºC, 16L:8D photoperiod) until hatching.

To analyze intra- and interannual temperature variation within and between the ponds, we collected surface water temperature data with data loggers (HOBO Water Temp Pro v2 Data Logger) in each wetland. Temperature loggers were placed near the sites where the egg clutches were laid. The data were collected automatically at 15-min intervals each year in 2008–2010 from the start of breeding season in April until mid-June (Table 1). Data on forest canopy cover and predator densities for each pond were collected in 2008 (see Richter-Boix et al. 2011a for details).

In 2008, 30 presumably unrelated individuals were collected at each of the seven breeding sites and genotyped using six neutral microsatellite markers (see Richter-Boix et al. 2011a for details). Pairwise and overall genetic differentiation between the sites (F_{ST}) was assessed according to Weir and Cockerham (1984), and significance tested with 10,000 permutations using GenAlEx (Peakall and Smouse 2006).

MEASURING POPULATION DIVERGENCE IN THERMAL PLASTICITY OF LARVAL LIFE HISTORY, LOCOMOTOR PERFORMANCE, AND TOLERANCE TRAITS

Three different experiments were conducted to examine divergence in tadpole thermal traits among the populations. First, we examined the effects of temperature on larval developmental and growth rates by measuring TDRN. Second, we studied divergence in thermal sensitivity of locomotor performance by estimating TPCs (Huey and Stevenson 1979; Kingsolver et al. 2004) in tadpole burst swimming speed for each population. In ectotherms, locomotor performance has been employed as a good proxy for thermal sensitivity and optimal temperatures because it is often correlated with fitness (Le Galliard et al. 2004; Husak 2006; Logan et al. 2014). Third, we examined variation in upper thermal resistance among the populations by measuring the CT_{max}.

Thermal developmental reaction norms

We conducted a laboratory common garden experiment to compare life-history traits plasticity among the seven populations. To control for possible maternal effects, manifested as initial size differences at family level (Loman 2003), hatchling (developmental stage 21; Gosner 1960) body length was measured from dorsal photographs. Alternative maternal effects such as maternal provisioning of nutrients and/or hormones were not controlled for, but

earlier studies on maternal effect influences on tadpole size, shape, and performance have shown strong relationships with initial offspring size (Pakkasmaa et al. 2003; Kaplan and Phillips 2006).

At Gosner stage 25, 10 tadpoles from each family were haphazardly selected and randomly divided into either 15 or 20ºC temperature treatments. We chose 15ºC as experimental temperature because a previous laboratory experiment at 12ºC, which is closer to mean temperature in the field, showed high mortality rates during the larval stage (A. Laurila, unpubl. data). Constant 20ºC reflects the average of pond mean maximum temperatures encountered in the ponds (Table 1). The experiment was conducted in two walk-in climate-controlled rooms with 16L:8D photoperiod. Tadpoles were raised individually in 0.75-l opaque plastic vials. Tadpoles from the different populations were distributed evenly across the vertical temperature gradient in the laboratory rooms. Throughout the experiment we used reconstituted soft water (RSW; APHA 1985) changed every three days. The tadpoles were fed chopped spinach ad libitum in conjunction with water change, and the amount of food was increased as the tadpoles grew. When the first tadpoles reached metamorphosis (emergence of at least one forelimb; Gosner stage 42), the vials were checked twice daily for metamorphosing individuals. For each metamorph, we measured larval period (days) and wet mass at metamorphosis (to the nearest 0.1 mg), and a metric related to growth rate. The nonlinear nature of tadpole growth requires repeated periodic measurements, which strongly increases handling stress. To avoid handling stress effects on tadpole growth and survival, we used the ratio between mass at metamorphosis and larval period. This metric (referred in the text as "growth rate") calculates the average mass gain throughout the entire developmental period. The use of this metric is justified by the fact that we found no significant differences in initial size among the populations (see results below), and no differences in results for growth rate when initial size at the family level was used as a covariate. The experiment was ended when all the tadpoles metamorphosed.

Thermal performance curves

At Gosner stage 25, 30 tadpoles per breeding site (three tadpoles per family) were randomly selected and individually raised in 0.75-l opaque plastic containers at 20ºC under the conditions described above. The swimming performance trials were started after three weeks of rearing at Gosner stages 28–33. Burst swimming speed was measured at six temperatures tested randomly during six consecutive days in the following order: 20, 32, 15, 10, 28, and 24 ºC. Tadpoles were acclimated to the test temperature for 45–60 minutes before the swimming performance test. To determine burst speed, tadpoles were placed individually in a portable thermal bath (patent license ES 2372085), which consisted of an opened cross section methacrylate tube (100 cm long × 6 cm wide × 3 cm deep) filled with water of a given temperature (accuracy ±0.4ºC). We then gently prodded the tadpole with a thin stick to stimulate swimming. At least 10 startle burst-swimming sequences per tadpole were videotaped using a video camera (JVC Everio GZ-MG505) at 30 frames/sec installed 2 m above the tube. The videos were analyzed with the software Measurement in Motion version 3.1 (http://www.motion.com/products/measurement/index.html). Burst speed of the tadpoles was determined by calculating the distance swum in the first five frames after tadpole's first reaction to stimulus, when burst speed is highest (e.g., Katzenberger et al. 2014). Startle responses were measured for each individual at each temperature and the fastest speed measured was used in the analyses. When all performance tests were finished at a temperature, six individuals from each population were randomly selected to repeat the swimming performance test in that temperature. Because swimming speed may scale with body size (Gvoždík and van Damme 2008), we used size-corrected velocity to construct TPCs. After the swimming experiment, each individual was laterally photographed from a standardized distance. The image processing software ImageJ (Abramoff et al. 2004) was used to measure the total length (body and tail) of each tadpole, which was used to obtain a size-corrected speed.

Upper thermal tolerance

To determine the CT_{max}, 24 tadpoles per breeding site (three tadpoles from each of eight families in every population: 168 individuals in total) were raised at 20ºC under the conditions described above. CT_{max} measurements were started after three weeks of rearing at Gosner stages 28–35. We used the Hutchinson's dynamic method (Lutterschmidt and Hutchison 1997), which has been used successfully in tadpoles (Duarte et al. 2012). Each tadpole was placed individually in 400 ml of RSW in a 700-ml glass bowl that was partially submerged within a 2-l beaker. The water in the beaker was heated at a rate of $\Delta T = 1.0$ºC/min. A magnetic stirrer (Agimatic-N, JP Selecta S.A., Barcelona, Spain) ensured uniform heating of the water. Each trial started at the tadpoles' acclimation temperature (20ºC). Loss of righting response and the onset of spasms were recorded, and the temperature at which the onset of spasms was detected was considered as the CT_{max} (Lutterschmidt and Hutchison 1997). Water temperature was monitored with a digital thermometer (0.1ºC). Once CT_{max} was reached, the tadpole was removed and placed in 20ºC water to recover. Tested individuals were weighed to the nearest 0.1 mg and the developmental stage was determined according to Gosner (1960).

STATISTICAL ANALYSES

Variation in initial size among the breeding sites was analyzed with a generalized mixed effect model, in which breeding site

was included as fixed effect and individuals from the same family were nested within family and breeding site. In the TDRN experiment, the effects of temperature, breeding site, and their interactions on the three larval-history traits measured were analyzed with a Generalized Least Squares (GLS) model. Temperature treatment, breeding site, and their interaction were included as fixed effects. Considering the nonindependence of tadpoles from the same family, we defined a general correlation matrix assuming that the residuals of the same family and location are not independent of each other (Schielzeth and Nakagawa 2013). Analyses were performed with restricted maximum likelihood (REML) estimation in the *nlme* package in R version 2.10.1 using the *corCompSymm* argument in the *gls* function.

We analyzed the TPCs for burst swimming speed by using the Template Mode of Variation (TMV; Izem and Kingsolver 2005) method. TMV is a statistical method developed to characterize and quantify nonlinear modes of variation for reaction norms in which standard linear methods may be inadequate or difficult to interpret (Izem 2004). It employs a polynomial function, which decomposes variation among TPCs into the three predetermined modes of variation: vertical shift (faster–slower), horizontal shift (hotter–colder), and specialist–generalist trade-off (Huey and Kingsolver 1989). Because burst speed measured at 10ºC showed low values of repeatability, this temperature was not included in the analyses. Because we tested tadpole performance at five temperatures, we assumed that the common template curve was a fourth-degree polynomial (e.g., Izem and Kingsolver 2005; Gvoždík and van Damme 2008; Katzenberger et al. 2014). The fitted curves allowed us to calculate the maximum performance (Z_{max}); optimum temperature (T_{opt}) where the maximum swimming velocity was reached; and two measurements of performance breadth B_{80} and B_{95}, representing the range of temperatures at which performance values exceeded 80 and 95% of the maximum, respectively (Huey and Stevenson 1979). All computations on TPCs, except B_{80} and B_{95}, were done using the Matlab (The MathWorks Inc. 2009) code by Izem and Kingsolver (2005). We also confirmed the fit of the curve for each population and calculated standard errors for the curve parameters using *nlinfit* and *nlparci* functions in Matlab. We estimated the repeatability of swimming performance to ensure the accuracy of our conclusions (see S3 for details and data).

As in the TDRN experiment, differences in CT_{max} among sites were analyzed with a GLS method, with breeding site as a fixed effect and the nonindependence of tadpoles from the same family was defined with a general correlation matrix assuming that the residuals of the same family and location are not independent of each other using the *corCompSymm* argument in the *gls* function in the *nlme* package in R. CT_{max} was analyzed both without a covariate and by including either developmental stage or mass as a covariate in the model.

INDIRECT EVIDENCE OF DIVERGENT SELECTION: Q_{ST}–F_{ST}

To test if the differences among the breeding sites were a consequence of neutral process or adaptive divergence, we compared the estimated quantitative divergence (Q_{ST}) to divergence in putatively neutral molecular markers (F_{ST}; Spitze 1993; Merilä and Crnokrak 2001). Assuming random mating and an inbreeding coefficient of zero, we computed Q_{ST} according to Spitze (1993). We used a Bayesian approach to calculate point estimates with a nested random effects linear model using the MCMCglmm function in the R package *MCMglmm* (Hadfield 2010; see details in S2).

Q_{ST}–F_{ST} differentiation was tested using the approach of Whitlock and Guillaume (2009) with the R script provided by Lind et al. (2011). We estimated the predicted sample variance for Q_{ST} of a neutral trait by simulating it using information on mean F_{ST} and the within-population genetic variance of the quantitative trait in question. The difference between both statistics ($Q_{ST} - F_{ST}$) was simulated 50,000 times to create the distribution of a neutral quantitative trait with which we compared the observed Q_{ST}–F_{ST}. Q_{ST}–F_{ST} analyses were performed for the three life-history traits in each temperature treatment, for phenotypic plasticity in each trait by using the plasticity index (mean low temperature – individuals at high temperature)/mean low temperature (Valladares et al. 2006), and for CT_{max}.

RELATIONSHIPS BETWEEN PHENOTYPIC TRAITS AND THERMAL ENVIRONMENTAL PARAMETERS

We used Pearson correlations at the population level to analyze the relationships between the parameters obtained from TDRN, TPC, and CT_{max} experiments and the pond-specific thermal parameters. The thermal parameters used were the mean and maximum pond temperatures and the temperature range (maximum minus minimum temperature registered in the pond) during larval development. As we had temperature data for three consecutive years, we first analyzed the repeatability of water temperature parameters among years (period 2008–2010). We used the intraclass correlation coefficient (ICC) and considered variance among the ponds and variance within the ponds over time (Bell et al. 2009). We used functions *ICCbare* and *ICCest* in the ICC package (Wolak et al. 2012) in R to estimate the ICC and the ICC confidence intervals for alpha 0.05.

Given the multiple tests performed between thermal traits and environmental parameters, we applied (Benjamini and Hochenberg) false discovery rate controlling procedure (Benjamini and Hochenberg 1995). Due to the relatively small sample of seven sites we also included power analyses to allow for an unbiased interpretation of the results obtained. The BH analyses were done with library *sgof* (Carvajal-Rodríguez et al. 2009) and the power analyses with package *pwr*, both in R.

Results

STUDY AREA AND THERMAL ENVIRONMENTAL PARAMETERS

There was no significant relationship between mean temperature and temperature range among the wetlands ($r = 0.581$, $P = 0.171$). Repeatability coefficient for mean temperature was low (ICC = 0.39, $_{\text{Lower}}$IC = –0.05, $_{\text{Upper}}$IC = 0.82), suggesting that mean pond temperature varied among years. Temperature range was consistent through time, with higher variance among (20.47) than within ponds (1.52; ICC = 0.93, $_{\text{Lower}}$IC = 0.78, $_{\text{Upper}}$IC = 0.98). However, considering that the mean temperature in 2009 and mean temperatures in 2008 and 2010 were highly correlated ($r = 0.976$, $P < 0.001$), as were temperature range in 2009 and the mean temperature range in 2008 and 2010 ($r = 0.972$, $P < 0.001$), only 2009 temperatures were used in correlations with phenotypic traits. Both mean temperature ($r = –0.794$, P = 0.033) and temperature range ($r = –0.913$, $P = 0.004$) were negatively correlated with forest canopy cover. Sites with higher forest canopy cover had lower water temperatures and narrower temperature variation than open-canopy sites (S4).

THERMAL DEVELOPMENTAL REACTION NORMS

We did not find differences in hatchling size among the sites ($F_{6, 273} = 0.172$, $P = 0.531$), and initial body size was not included as a covariate in the subsequent analyses. We found differences among sites and temperature treatments in all life-history traits measured. Larval period was longer at 15ºC than at 20ºC ($F_{1, 546} = 1046.3$, $P < 0.0001$), with significant differences among sites ($F_{6, 546} = 13.17$, $P < 0.0001$) and a significant site × temperature interaction indicating differences in the slopes of reaction norms among the sites ($F_{6, 546} = 4.08$, P < 0.0005; Fig. 2A, E). The effects on mass at metamorphosis were also statistically significant for all factors (site: $F_{6, 546} = 7.36$, $P < 0.0001$; temperature treatment: $F_{1, 546} = 6.22$, $P = 0.0151$; site × temperature: $F_{6, 546} = 8.07$, $P < 0.0001$). Some sites had larger metamorphs at 20ºC, whereas in other sites the metamorphs were larger at 15ºC (Fig. 2B, F). Growth rate was higher at 20ºC for all sites ($F_{1, 546} = 355.74$, $P < 0.0001$), and there were differences among the sites ($F_{6, 546} = 14.31$, $P < 0.0001$) and in the slopes of reaction norms ($F_{6, 546} = 10.18$, $P < 0.0001$; Fig. 2C, G).

MODES OF TPC VARIATION AND CT_{max}

In the TPC experiment, the decomposition of variation in burst swimming speed into the three biological modes by TMV method explained 68.62% of the total variation in the data (S5). Two modes, the generalist–specialist and the vertical shift (faster–slower), accounted for the largest and almost equal proportions of the explained variation (30.93 and 29.03%, respectively), whereas the horizontal shift, the hotter–colder mode, explained only 8.66% of the total variation. Optimal temperature (T_{opt}) was negatively correlated with maximum performance (Z_{max}; $r = –0.803$, $P = 0.030$, not significant after BH multiple-test correction), but positively correlated with performance breadth even after BH correction (B_{80}: $r = 0.946$, B_{95}: $r = 0.953$, $P < 0.001$ in both cases). These correlations indicated that individuals from sites with higher T_{opt} had a broader performance breadth but lower maximum performance, whereas individuals with lower T_{opt} and narrower performance breadth showed higher maximum performance (Fig. 3). This pattern does not support the thermodynamic hotter is better constraint and suggests a pattern in which colder is better but narrower. Additionally, although there was no significant relation between performance breadth and maximum performance ($r = –0.63$, $P = 0.129$), the high variance explained by the generalist–specialist mode suggests that some trade-off may exist.

Maximum performance was negatively correlated with CT_{max} ($r = –0.847$, $P = 0.016$, not significant after BH correction). Mean CT_{max} varied from 35.5 to 36.1ºC among the populations. There was significant variation among the populations when analyzed without a covariate ($F_{6, 156} = 3.93$, $P = 0.001$; Fig. 2D), and when Gosner stage ($F_{6, 156} = 7.44$, P < 0.001) or body mass ($F_{6, 156} = 5.93$, $P < 0.001$) was used as a covariate.

INDIRECT EVIDENCE OF DIVERGENCE SELECTION: Q_{ST}–F_{ST}

The overall F_{ST} among the sites was very low (mean [SD]: 0.016 ± 0.012) and nonsignificant. A Mantel test between pairwise F_{ST} and geographic distances between the sites did not find significant Isolation by Distance (IBD; P-obs = 0.335, P-simulated = 0.123). Both analyses suggest a lack of population structure and high level of gene flow across the geographic area included in the study (Table S2.1; Fig. S2.1, S2). The Q_{ST}–F_{ST} analyses showed adaptive divergence among the breeding sites in larval development (15ºC: $Q_{ST} = 0.049$, $P_{QST>FST} = 0.0088$; 20ºC: $Q_{ST} = 0.058$, $P_{QST>FST} = 0.0032$) and mass at metamorphosis (15ºC: $Q_{ST} = 0.205$, $P_{QST>FST} < 0.0001$; 20ºC: $Q_{ST} = 0.507$; $P_{QST>FST} < 0.0001$), but not in plasticity of these traits. The differences in growth rate among sites did not exceed the neutral variation (15ºC: $Q_{ST} = 0.024$, $P_{QST>FST} = 0.1941$; 20ºC: $Q_{ST} = 0.022$, $P_{QST>FST} < 0.1893$; Table S2.2; Fig. S2.2, S2). All estimates of Q_{ST} for CT_{max} were significantly higher than F_{ST} (noncorrected: $Q_{ST} = 0.059$, $P_{QST>FST} = 0.002$; developmental-stage-corrected: $Q_{ST} = 0.153$, $P_{QST>FST} < 0.001$; body-mass-corrected $Q_{ST} = 0.176$, $P_{QST>FST} < 0.001$; Fig. S6).

RELATIONSHIP BETWEEN LIFE-HISTORY TRAITS AND TPC PARAMETERS

Life-history traits measured at 20ºC showed significant relationships with TPC parameters. Populations with higher maximum

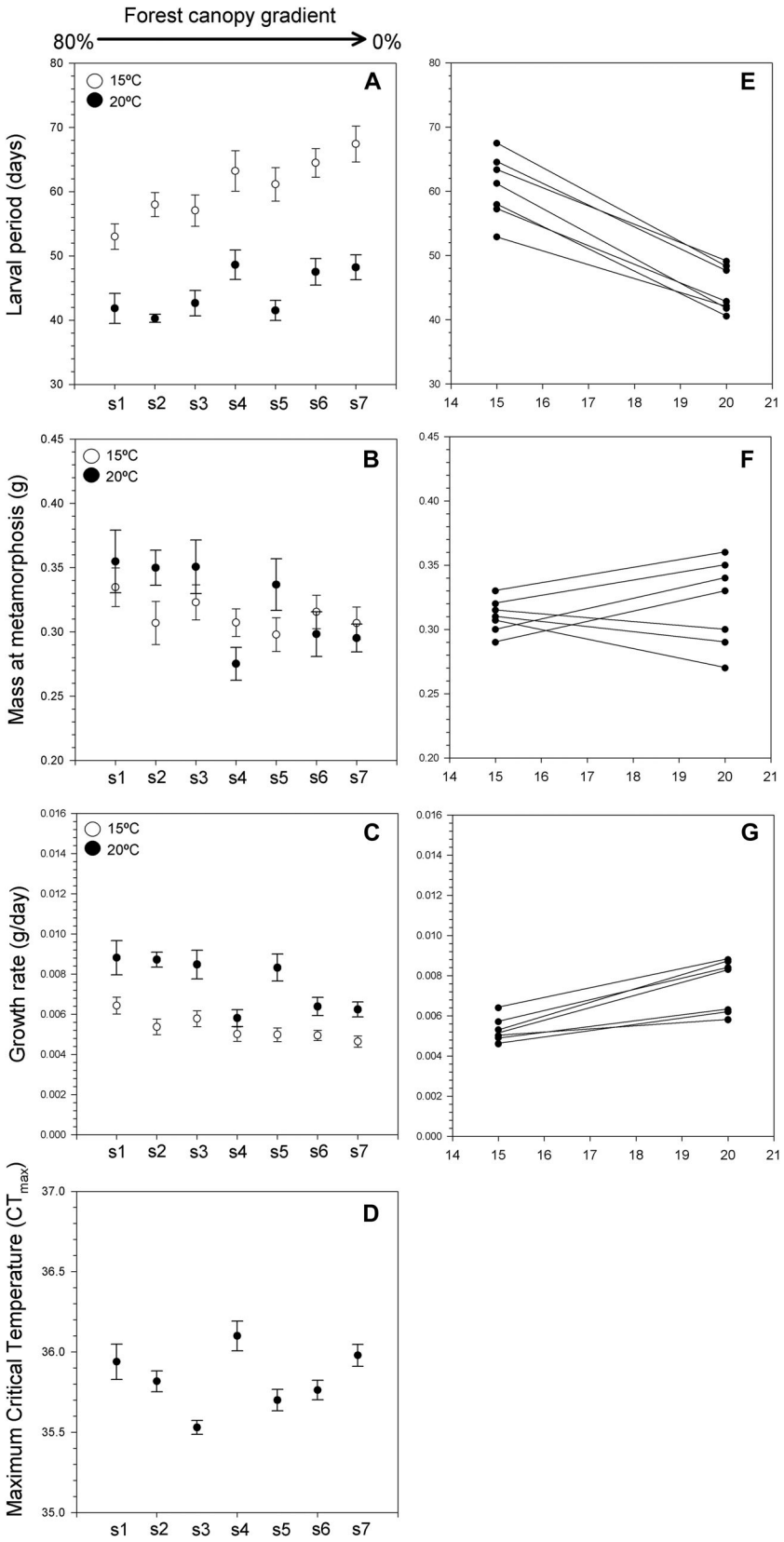


Figure 2. Mean (±2 SE) of (A) larval period, (B) mass at metamorphosis, (C) growth rate, and (D) CT_{max} (mean ± 1SE) and the population reaction norms across the two temperatures for larval period (E), mass at metamorphosis (F), and growth rate (G) in the thermal developmental reaction norms and CT_{max} experiment.

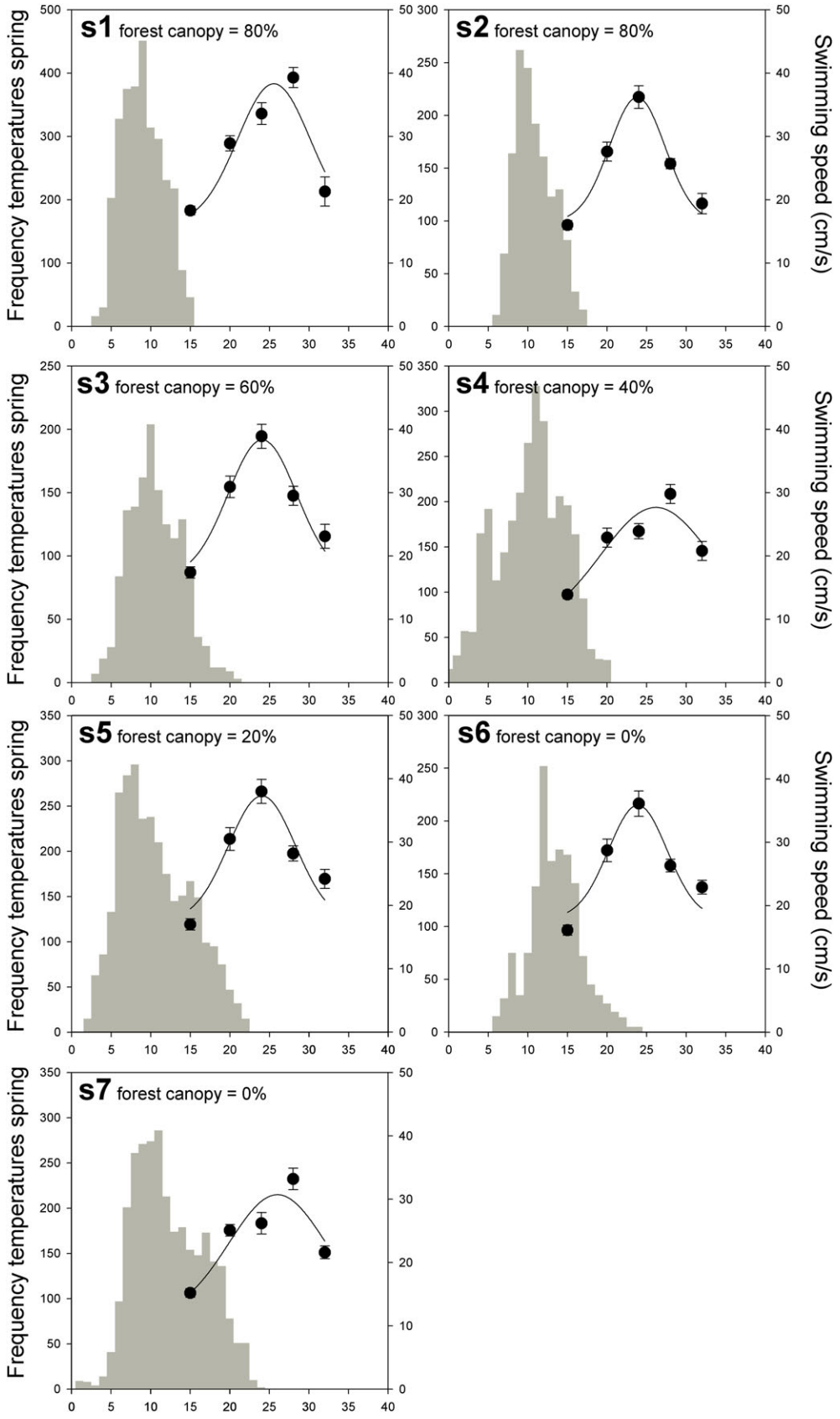


Figure 3. S1–S7: TPC curves (±1 SE) for swimming speed and frequency distribution of field water temperatures measured in spring 2009 for each population.

Table 2. Correlations between breeding-site temperature parameters and the ecophysiological traits measured in the TPC and CT_{max} experiments, and the life-history traits measured in the thermal developmental reaction norms experiment.

	Trait	T_{mean}	T_{range}	Z_{max}	T_{opt}	B_{95}	CT_{max}
Environmental parameters	T_{mean}	1					
	T_{range}	0.58	1				
	DT_{range}	0.14	0.74				
Ecophysiological	Z_{max}	−0.42	−0.46	1			
	T_{opt}	0.28	**0.79**	**−0.80**	1		
	B_{95}	0.38	**0.92**	−0.63	**0.95**	1	
	CT_{max}	−0.02	0.23	**−0.84**	0.74	0.55	1
Life-history	LP_{20}	0.53	**0.78**	**−0.82**	**0.89**	**0.83**	0.61
	LP_{15}	**0.78**	**0.85**	−0.72	**0.77**	**0.83**	0.41
	MM_{20}	−0.52	**−0.76**	**0.90**	**−0.91**	**−0.85**	−0.70
	MM_{15}	−0.18	−0.21	−0.16	−0.15	−0.23	−0.07
	GR_{20}	−0.60	**−0.80**	**0.86**	**−0.89**	**−0.84**	−0.62
	GR_{15}	−0.74	−0.73	0.61	−0.62	−0.69	−0.29

Values represent the Pearson's correlation coefficient, bold indicates significance at $P < 0.05$; gray indicates significance after BH multiple-test correction at $P = 0.05$.

T_{mean} = mean temperature in spring 2009, T_{range} = temperature range in spring 2009, DT_{range} = mean daily temperature range in spring 2009, Z_{max} = maximum performance, T_{opt} = optimal temperature, B_{95} = performance breadth, CT_{max} = maximum critical temperature, LP_{20} = larval period at 20ºC, LP_{15} = larval period at 15ºC, MM_{20} = mass at metamorphosis at 20ºC, MM_{15} = mass at metamorphosis at 15ºC, GR_{20} = growth rate at 20ºC, GR_{15} = growth rate at 15ºC.

Table 3. Correlations between breeding-site temperature parameters and the ecophysiological traits measured in the TPC and CT_{max} experiments, and the life-history traits measured in the thermal developmental reaction norms experiment.

	Trait	T_{mean}	T_{range}	Z_{max}	T_{opt}	B_{95}	CT_{max}
Environmental parameters	T_{mean}	1					
	T_{range}	0.58	1				
	DT_{range}	0.14	0.74				
Ecophysiological	Z_{max}	0.132	1.153	1			
	T_{opt}	0.075	**0.587**	**0.602**	1		
	B_{95}	0.111	**0.897**	0.292	**0.949**	1	
	CT_{max}	0.036	0.060	**0.705**	0.476	0.221	1
Life-history	LP_{20}	0.211	**0.598**	**0.642**	**0.818**	**0.678**	0.286
	LP_{15}	**0.561**	**0.712**	0.442	**0.540**	**0.663**	0.128
	MM_{20}	0.199	**0.520**	**0.864**	**0.887**	**0.720**	0.406
	MM_{15}	0.052	0.058	0.047	0.047	0.061	0.038
	GR_{20}	0.272	**0.601**	**0.758**	**0.831**	**0.697**	0.304
	GR_{15}	0.471	0.463	0.285	0.297	0.401	0.079

Values represent the power of the Pearson's product-moment correlation test, with the seven sites and significance level of 0.05. Values highlighted in gray indicate a power >80.

T_{mean} = mean temperature in spring 2009, T_{range} = temperature range in spring 2009, DT_{range} = mean daily temperature range in spring 2009, Z_{max} = maximum performance, T_{opt} = optimal temperature, B_{95} = performance breadth, CT_{max} = maximum critical temperature, LP_{20} = larval period at 20ºC, LP_{15} = larval period at 15ºC, MM_{20} = mass at metamorphosis at 20ºC, MM_{15} = mass at metamorphosis at 15ºC, GR_{20} = growth rate at 20ºC, GR_{15} = growth rate at 15ºC.

performance also developed and grew faster, resulting in larger size at metamorphosis than in populations with lower maximum performance. On the other hand, populations with slower growth and development had higher T_{opt} and wider performance breadth (Table 2, Fig. S7). The results at 15ºC were generally similar for development time (Tables 2 and 3), but not significant after BH correction. Correlations between TPC parameters and mass at metamorphosis at 15ºC were very low, and this was in turn reflected to correlations between TPC parameters and growth rate (Tables 2 and 3).

RELATIONSHIPS BETWEEN PHENOTYPIC TRAITS AND THERMAL ENVIRONMENTAL PARAMETERS

T_{opt} and B_{95} were positively correlated with temperature range, but not with mean temperature, highlighting the importance of temporal temperature variation in defining the thermal performance differences among breeding sites (Tables 2 and 3, Fig. 4). Life-history traits were, especially at the higher temperature, correlated with temperature range, with faster development, higher growth, and larger mass at metamorphosis in ponds with narrower temperature range (Tables 2 and 3, Fig. S7). CT_{max} was not correlated with any of the thermal environmental parameters, not even with absolute pond temperature maximum ($r = 0.141$, $P = 0.327$).

Discussion

LOCAL THERMAL ADAPTATION

This study suggests adaptive divergence in larval life history and thermal locomotor performance among local *R. arvalis* populations in the presence of high gene flow, suggesting that natural selection by local thermal environment is driving the population divergence. These results are in accordance with recent studies demonstrating that thermal environment may exert strong selection on thermal physiology in ectotherms (Bradshaw and Holzapfel 2008; Leal and Gunderson 2012; Higgins et al. 2014; Logan et al. 2014). The fine-grained scale of divergence in our system seems to reflect the level of the underlying thermal environmental heterogeneity due to variation in shading forest canopy cover around the breeding ponds, creating predictable temperature conditions over time. Similar divergence was found in embryonic development and larval thermal performance in *Lithobates sylvaticus* in a comparable forest pond system, suggesting microevolution in life-history and thermal traits in response to microclimatic variation (Skelly and Freidenburg 2000; Freidenburg and Skelly 2004; Skelly 2004).

We found adaptive responses to two main constraints that may limit the evolution of thermal sensitivity, the generalist–specialist trade-off (Levins 1968; Huey and Slatkin 1976) and the hotter is better thermodynamic constraint (Huey and Kingsolver 1989; Savage et al. 2004). Forest pond populations experiencing smaller thermal variation exhibited narrower performance curves, lower optimum temperatures, and superior swimming performance as compared to populations originating from thermally more-variable open ponds, the latter functioning over a wider range of temperatures. This pattern, compatible with the generalist–specialist trade-off, is in accordance with previous studies on other ectotherms (e.g., Izem and Kingsolver 2005; Latimer et al. 2011).

The negative correlation between maximum performance (Z_{max}) and temperature of maximal performance (T_{opt}) contrasts with the hotter is better hypothesis predicting that Z_{max} and T_{opt} are positively correlated due to thermodynamic depression of biochemical reactions at low temperatures (Bennett 1987; Gillooly et al. 2002; Hochachka and Somero 2002; Savage et al. 2004). However, organisms may undergo biochemical adaptations to compensate for this depression in performance, and as both processes can be acting simultaneously, the result will depend on the relative weights of constraint versus adaptive response (Clarke 2003; Angiletta et al. 2010). In our case, the results do not support the thermodynamic constraint hypothesis, but indicate a compensatory response in colder closed-canopy pond populations (see below). These results contradict the general pattern found in ectothermic organisms from interspecific comparisons (e.g., Frazier et al. 2006; Knies et al. 2009; Angilletta et al. 2010), but are in accordance with the few intraspecific studies, which have not found support for the thermodynamic constraint hypothesis (Gilchrist 1995; Izem and Kingsolver 2005; Latimer et al. 2011; Klepsatel et al. 2013, but see Knies et al. 2009).

HOW DO OUR EMPIRICAL DATA FIT THE EVOLUTIONARY PHYSIOLOGICAL MODES OF VARIATION?

The pattern of population divergence in the thermal physiology and life-history traits analyzed here was complex and did not fully match with the predictions of optimality models of thermal sensitivity evolution (Lynch and Gabriel 1987; Gilchrist 1995). These optimality models predict—in accordance with the horizontal-shift model—that the maximal performance for each population should fit the temperatures that tadpoles experience most frequently in their local environments. In addition, the performance breadth in thermal sensitivity curve should depend on the amount of thermal variability within and among generations. Our results fit these predictions only partially. Mean pond temperatures were not correlated with thermal responses of *R. arvalis* populations for most traits, and the horizontal mode of variation was the least successful in explaining the variance in the thermal locomotor performance curves.

The poor fit of thermal responses with mean environmental conditions could be explained by the relatively small number of sites we used, but the general tendency observed between phenotypic traits, mean temperature and temperature range, supported by high statistical power in temperature range but not in mean temperature, suggest that temperature variation is more important in this system than mean temperature. The probable reason for this is the fact that the gradient in mean temperature was not as wide (from 9.44 to 13.24ºC) as the gradient in temperature range (from 12.04 to 24.48 ºC) and had higher variability between years. An alternative could be that water surface temperatures were not representative of thermal conditions in the whole pond. Although we find this unlikely, further analyses of field

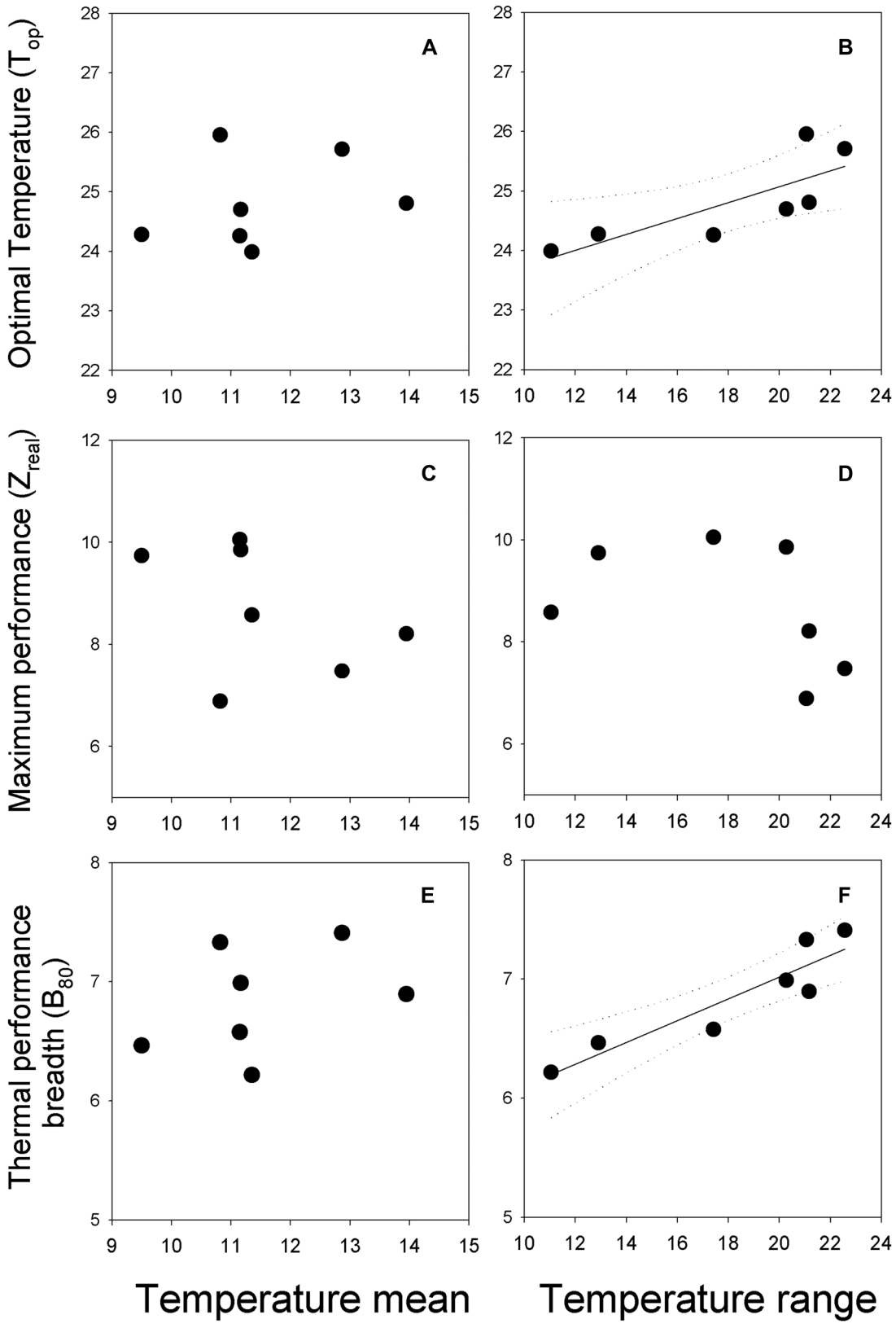


Figure 4. Relationships between the TPC parameters and the mean and range of pond temperatures in spring 2009 for the seven populations. The regression line is included only when the relationship is significant.

operative temperatures and larval thermoregulatory scope could complement our laboratory observations.

However, the poor fit of thermal responses with mean temperature contrasts with the high correlations between thermal responses and local thermal variation, especially when considering seasonal rather than daily thermal ranges. We found that populations experiencing higher thermal oscillations, typically originating from open-canopy ponds, had lower developmental and growth rates, and smaller size at metamorphosis at constant temperatures, possibly because they were adapted to thermally fluctuating conditions. Low development and growth rates may reduce fitness, for instance, by increasing tadpole mortality through pond desiccation (Richter-Boix et al. 2011b), or by lower juvenile survival through small size at metamorphosis (e.g., Smith 1987; Semlitsch et al. 1988; Altwegg and Reyer 2003). On the other hand, increased thermal heterogeneity may enhance fitness through wider performance breadth and adjustable optimum temperatures, as has been reported for latitudinal clines in both locomotor performance (van Berkum 1986; John-Alder et al. 1988) and growth rates (Eggert et al. 2003; Stefansson et al. 2013, but see de Boer et al. 2005; Condon et al. 2013).

Previous studies on ectotherms in general, and larval amphibians in particular, have described the vertical shift as the most common pattern of life-history variation along latitudinal and altitudinal gradients (Conover et al. 2009; see Berven 1982; Laugen et al. 2003; Orizaola et al. 2010; Muir et al. 2014a for amphibian examples) indicating an adaptive compensation mechanism for the short period when environmental conditions are favorable (the metabolic cold adaptation hypothesis; Scholander et al. 1953). Empirical studies have demonstrated the compensation also in small-scale variation among local habitats, even when differences in growing season length among localities are small like in the present study (Skelly 2004; Richter-Boix et al. 2010). Our TDRN experiment partially supported these studies showing faster development and growth rate in some of the populations across both temperatures. These populations tended to experience lower thermal oscillations, typical for more-shaded ponds with lower mean temperatures. Adaptation in terms of faster growth implies the existence of trade-offs, which are often suggested to occur between growth and defense (e.g., Munch and Conover 2003; Laurila et al. 2008). In our system, the numbers of predators are higher in warmer open-canopy ponds, and time constraints imposed by relatively low temperatures in closed-canopy ponds and higher predation risk in open-canopy ponds have been suggested as key factors driving the fast–slow gradient (Richter-Boix et al. 2013).

In addition, a third environmental component may act as a complementary source of selection across the mosaic of forest and open-canopy ponds. As frog populations from cooler forest ponds breed later than populations from open-canopy wetlands (Richter-Boix et al. 2013), it appears that faster development and growth in these populations may also have evolved to compensate for late breeding. A recent experimental study conducted in this system showed that breeding time directly influences larval developmental and growth rates through maternal effects (Richter-Boix et al. 2014). However, the breeding time differences among the present populations were not large (max. six days), and including breeding time as a covariate in the analyses did not change the results, indicating that breeding time related maternal effects were of minor importance in the present study. However, we cannot exclude the possibility that maternal effects other than initial size and breeding time, and not analyzed here, could influence larval growth and development rates (Pakkasmaa et al. 2003; Kaplan and Phillips 2006).

We found evidence for population × temperature treatment interactions in the TDRN experiment suggesting genetic variation in plasticity in accordance with the generalist–specialist model (Fig. 1, Kingsolver et al. 2004). On the other hand, we found no adaptive divergence in reaction norm slopes in the Q_{ST}–F_{ST} comparisons suggesting no differences in plasticity among the populations, hence supporting a vertical-shift pattern. However, our TDRN experiment considered only a two-point reaction norm (Fig. 1), which may be insufficient for analyzing concave reaction norms. A recent review showed that two environment studies often underestimate the extent of evolutionary change in reaction norms and fail to capture biologically relevant patterns of plasticity (Murren et al. 2014). Moreover, the two forms of plasticity analyzed here target different traits and operate at different time scales: TDRN is a product of temperature effects during the whole larval period, whereas TPC considers short-term or instantaneous effects of temperature on an individual. We suggest that both TPC and TDRN analyses showed a combination of modes of variation reflecting shifts in both thermal breadth (generalist–specialist) and height (faster–slower).

WHY IS MAXIMAL PERFORMANCE FOUND AT TEMPERATURES HIGHER THAN THOSE IN NATURE?

In accordance with previous studies (Knies et al. 2009; Nilsson-Örtman et al. 2013), all populations had much higher thermal optima in the TPC analyses (24–26ºC) than modal or mean temperatures experienced during the larval stage in nature (modal: 8.5–12ºC, mean: 9.5–14ºC). Considering that TPCs are highly asymmetric with performance with fitness dropping rapidly at temperatures above the optimum (Izem and Kingsolver 2005), and as body temperatures above the performance optimum reduce fitness much more than temperatures below the optimum, a large thermal safety margin is expected (Deutsch et al. 2008). As body temperature in aquatic ectotherms is highly dependent on ambient temperature (Feder and Hofmann 1999), populations should select and occupy environments with temperatures somewhat lower than the temperature at which fitness is maximal ("the

suboptimal is optimal"; Martin and Huey 2008). In the present study, this safety margin (the difference between T_{opt} and mean maximal [sensu Deutsch et al. 2008] environmental temperature; Huey et al. 2009; Chown et al. 2010) increases with forest canopy, from 1 ± 2ºC (SD) in open-canopy ponds to 5.3 ± 3.7ºC in closed-canopy ponds. However, the large deviation between the thermal optima and the mean or modal temperatures in natural environments can be the result of derived thermal optima under constant laboratory conditions. Empirical studies have demonstrated that fluctuating temperature conditions can reduce both T_{opt} and CT_{max} measures (Paaijmans et al. 2013), consistent with the nonlinear effects of temperature according to Jensen's inequality (Ruel and Ayres 1999). However, other studies have shown an increase in growth rate and CT_{max} provided that the fluctuating temperature does not exceed the optimal temperature (Niehaus et al. 2012; M. Tejedo, L. M. Gutiérrez Pesquera, M. J. Piñero, and A. G. Nicieza, unpubl. data). The expected consequences of temperature fluctuations will depend on the curvature of the nonlinear thermal reaction norm (Ragland and Kingsolver 2008; Kingsolver et al. 2009). Jensen's inequality states that the average result of a nonlinear function does not equal the function evaluated at the average value, with variance in temperature lowering the response variable when the function is decelerating, and raising the response variable when the function is accelerating (Ruel and Ayres 1999).

We also found significant divergence in CT_{max}, indicating potential for adaptive microgeographic evolutionary changes in this key trait. Previous studies have demonstrated fast rates of divergence in thermal tolerance especially in insects (Hoffmann et al. 2003), but also in fish (Barrett et al. 2011) and lizards (Leal and Gunderson 2012). In tadpoles of *Rana sylvatica* inhabiting ponds of contrasting thermal regimes, CT_{max} had diverged in less than 40 years (Skelly and Freidenburg 2000). Our results agree with the view that thermal physiological traits of ectotherms are a target of microevolution. However, although the among-population divergence in CT_{max} appears to be the response to directional selection, the upper resistance was not related to environmental peak temperature variation between sites. In addition, maximum pond temperatures rarely reached high values, suggesting that tadpoles do not experience acute heat stress in the natural ponds. The warming tolerance, defined as the difference between CT_{max} and peak maximum temperature in the field, and used to estimate the vulnerability to heat stress (Somero 2005; Deutsch et al. 2008; Duarte et al. 2012), was very high (12–20ºC). These results imply that direct selection from lethal environmental temperatures is rare in the present populations.

ACKNOWLEDGMENTS

We thank L. Gvoždík, G. Orizaola, editors W. Blackenhorn and R. Shaw, and two anonymous reviewers for constructive comments on the manuscript. We also thank O. Massana for his help in analyzing the video recordings. The animals were collected with a permit from Uppsala County Board (521-3019-09) and the experiment was conducted with a permit from the Ethical Committee for Animal Experiments in Uppsala (C92/9). Our research was supported by Spanish postdoctoral grants MEC2007-0944 and 2008-BP-A-00032 to ARB, MICINN grant (CGL2009-12767-C02-02) and MINECO award (CGL2012-40246-C02-01) to MT, and Formas grant (2007-903) to AL.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.6d09b.

LITERATURE CITED

Abramoff, M. D., P. J. Magalhaes, and S. J. Ram. 2004. Image processing with Image. J. Biophot. Int. 11:36–42.

Altwegg, R., and H. U. Reyer. 2003. Patterns of natural selection on size at metamorphosis in water frogs. Evolution 57:872–882.

Angilletta, M. J. 2009. Thermal adaptation. A theoretical and empirical synthesis. Oxford Univ. Press, New York.

Angilletta, M. J., P. H. Niewiarowski, and C. A. Navas. 2002. The evolution of thermal physiology in ectotherms. J. Therm. Biol. 27:249–268.

Angilletta, M. J., R. B. Huey, and M. R. Frazier. 2010. Thermodynamic effects on organismal performance: is hotter better? Physiol. Biochem. Zool. 83:197–206.

APHA. 1985. Standard methods for the examination of water and wastewater. 16th ed. American Public Health Association, Washington, DC.

Barrett, R. D. H., A. Paccard, T. M. Healy, S. Bergek, P. M. Schutte, D. Schluter, and S. M. Rogers. 2011. Rapid evolution of cold tolerance in stickleback. Proc. R. Soc. Lond. B Biol. Sci. 278:233–238.

Baumann, H, and D. O. Conover. 2011. Adaptation to climate change: contrasting patterns of thermal-reaction-norms evolution in Pacific versus Atlantic silversides. Proc. R. Soc. Lond. B Biol. Sci. 278:2265–2273.

Bell, A. M., S. J. Hankison, and K. L. Laskowski. 2009. The repeatability of behavior: a meta-analysis. Anim. Behav. 77:771–783.

Benjamini, Y., and Y. Hochenberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. 57:289–300.

Bennett, A. F. 1987. Evolution of the control of body temperature: is warmer better? Pp. 421–431 *in* P. Dejours, L. Bolis, C. R. Taylor, and E. R. Weibel, eds. Comparative physiology: life in water and on land. Springer-Verlang New York, and Liviana Press, Padova, Italy.

Berger, D., E. Postma, W. U. Blanckenhorn, and R. J. Walters. 2013. Quantitative genetic divergence and standing genetic (co)variance in thermal reaction norms along latitude. Evolution 67:2385–2399.

Berven, K. A. 1982. The genetic basis of altitudinal variation in the wood frog *Rana sylvatica*. II. An experimental analysis of larval development. Oecologia 52:360–369.

Blanckenhorn, W. U. 1991. Life-history difference in adjacent water strider populations: phenotypic plasticity or heritable responses to stream temperature? Evolution 45:1520–1525.

Blanquart, F., O. Kaltz, S. L. Nuismer, and S. Gandon. 2013. A practical guide to measuring local adaptation. Ecol. Lett. 16:1195–1205.

Bradshaw, W. E., and C. M. Holzapfel. 2008. Genetic response to rapid climate change: it's seasonal timing that matters. Mol. Ecol. 17:157–166.

Carvajal-Rodríguez, A., J. deUña-Álvarez, and E. Rolán-Álvarez. 2009. A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests. BMC Bioinform. 10:1–14.

Chevin, L. M., R. Lande, and G. M. Mace. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biol. 8:e1000357.

Chown, S. L., A. A. Hoffmann, T. N. Kristensen, M. J. Angilletta, N. C. Stenseth, and C. Pertoldi. 2010. Adapting to climate change: a perspective from evolutionary physiology. Climate Res. 43:3–15.

Clarke, A. 2003. Costs and consequences of evolutionary temperature adaptation. Trends Ecol. Evol. 18:573–581.

Condon, C., B. S. Cooper, S. Yeaman, and M. J. Angilletta. 2013. Temporal variation favors the evolution of generalists in experimental populations of *Drosophila melanogaster*. Evolution 68:720–728.

Conover, D. O., and E. T. Schultz. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. Trends Ecol. Evol. 10:248–252.

Conover, D. O., T. A. Duffy, and L. A. Hice. 2009. The covariance between genetic and environmental influences across ecological gradients. Ann. NY Acad. Sci. 1168:100–129.

de Boer, M. K., E. M. Koolmees, E. G. Vrieling, A. M. Breeman, and M. vanRijssel. 2005. Temperature responses of three *Fibrocapsa japonica* strains (Raphidophyceae) from different climate regions. J. Plankton Res. 27:47–60.

Deutsch, C. A., J. J. Tewksbury, R. B. Huey, K. S. Sheldon, C. K. Ghalambor, D. C. Haak, and P. R. Martin. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. Proc. Natl. Acad. Sci. USA 105:6668–6672.

Duarte, H., M. Tejedo, M. Katzenberger, F. Marangoni, D. Baldo, J. F. Beltrán, D. A. Martí, A. Richter-Boix, and A. Gonzalez-Voyer. 2012. Can amphibians take the heat? Vulnerability to climate warming in subtropical and temperate larval amphibian communities. Global Change Biol. 18:412–421.

Eggert, A., E. M. Burger, and A. M. Breeman. 2003. Ecotypic differentiation in thermal traits in the tropical to warm-temperate green macrophyte *Valonia utricularis*. Bot. Mar. 46:69–81.

Eliason, E. J., T. D. Clark, M. J. Hague, L. M. Hanson, Z. S. Gallagher, K. M. Jeffries, M. K. Gale, D. A. Patterson, S. G. Hinch, and A. P. Farrell. 2011. Differences in thermal tolerance among sockeye salmon populations. Science 332:109–112.

Endler, J. A. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

Feder, M. E., and G. E. Hofmann. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu. Rev. Physiol. 61:243–282.

Frazier, M. R., R. B. Huey, and D. Berrigan. 2006. Thermodynamics constrains the evolution of insect population growth rates: "warmer is better." Am. Nat. 168:512–520.

Freidenburg, L. K., and D. K. Skelly. 2004. Microgeographical variation in thermal preference by an amphibian. Ecol. Lett. 7:369–373.

Gilchrist, G. W. 1995. Specialist and generalist in changing environments. 1. Fitness landscapes of thermal sensitivity. Am. Nat. 146:252–270.

Gillooly, J. F., E. L. Charnov, G. B. West, V. M. Savage, and J. H. Brown. 2002. Effects of size and temperature on developmental time. Nature 417:70–73.

Gosner, K. L. 1960. A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16:183–190.

Gvoždík, L., and R. van Damme. 2008. The evolution of thermal performance curves in semi-aquatic newts: thermal specialists on land and thermal generalists in water? J. Therm. Biol. 33:339–403.

Hadfield, J. D. 2010. MCMC methods for multi-response generalised linear mixed models: the MCMCglmm R package. J. Stat. Softw. 33:1–22.

Higgins, J. K., H. J. MacLean, L. B. Buckley, and J. G. Kingsolver. 2014. Geographic differences and microevolutionary changes in thermal sensitivity of butterfly larvae in response to climate. Funct. Ecol. 28:982–989.

Hochachka, P., and G. Somero. 2002. Biochemical adaptation: mechanism and process in physiological evolution. Academic Press, New York.

Hoffmann, A. A., J. G. Sorensen, and V. Loeschke. 2003. Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. J. Therm. Biol. 28:175–216.

Huey, R. B., and P. E. Hertz. 1984. Is a jack-of-all-temperatures a master of none? Evolution 38:441–444.

Huey, R. B., and J. G. Kingsolver. 1989. Evolution of thermal sensitivity of ectotherm performance. Trends Ecol. Evol. 4:131–135.

Huey, R. B., and M. Slatkin. 1976. Cost and benefits of lizard thermoregulation. Q. Rev. Biol. 51:363–384.

Huey, R. B., and R. D. Stevenson. 1979. Integrating thermal physiology and ecology of ectotherms: a discussion of approaches. Am. Zool. 19:357–366.

Huey, R. B., C. A. Deutsch, J. J. Tewksbury, L. J. Vitt, P. E. Hertz, H. J. Álvarez Pérez, and T. Garland Jr. 2009. Why tropical forest lizards are vulnerable to climate warming. Proc. R. Soc. Lond. B Biol. Sci. 276:1939–1948.

Husak, J. F. 2006. Does speed help to survive? A test with collared lizards of different ages. Funct. Ecol. 20:174–179.

Izem, R. 2004. Analyzing variation in curves of common shape. Ph.D. dissertation, University of North Carolina, Chapel Hill, NC. Available at http://www.unc.edu/~rizem/fda/data1pres.html.

Izem, R, and J. Kingsolver. 2005. Variation in continuous reaction norms: quantifying directions of biological interest. Am. Nat. 166:277–289.

John-Alder, H. B., P. J. Morin, and S. P. Lawler. 1988. Thermal physiology, phenology, and distribution of treefrogs. Am. Nat. 132:506–520.

Kaplan, R. H, and P. C. Phillips. 2006. Ecological and developmental context of natural selection: maternal effects and thermally induced plasticity in the frog *Bombina orientalis*. Evolution 60:142–156.

Katzenberger, M., J. Hammond, H. Duarte, M. Tejedo, C. Calabuig, and R. A. Relyea. 2014. Swimming with predators and pesticides: how environmental stressors affect the thermal physiology of tadpoles. PLoS ONE 9:e98265.

Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecol. Lett. 7:1225–1241.

Keller, I., and O. Seehausen. 2012. Thermal adaptation and ecological speciation. Mol. Ecol. 21:782–799.

Kingsolver, J. G., R. Gomulkiewicz, and P. A. Carter PA. 2001. Variation, selection and evolution of function-valued traits. Genetica 112: 87–104.

Kingsolver, J. G., R. Izem, and G. J. Ragland. 2004. Plasticity of size and growth in fluctuating thermal environments: comparing reaction norms and performance curves. Integr. Comp. Biol. 44:450–460.

Kingsolver, J. G., G. J. Ragland, and S. E. Diamond. 2009. Evolution in a constant environment: thermal fluctuations and thermal sensitivity of laboratory and field populations of *Manduca sexta*. Evolution 63:537–541.

Klepsatel, P., M. Gáliková, N. De Maio, C. D. Huber, C. Schlötterer, and T. Flatt. 2013. Variation in thermal performance and reaction norms among populations of *Drosophila melanogaster*. Evolution 67:3573–3587.

Knies, J. L., R. Izem, K. L. Supler, J. G. Kingsolver, and C. L Burch. 2006. The genetic basis of thermal reaction norm evolution in lab and natural phage populations. PloS Biol. 4:e201.

Knies, J. L., J. G. Kingsolver, and C. L. Burch. 2009. Hotter is better and broader: thermal sensitivity of fitness in a population of bacteriophages. Am. Nat. 173:419–430.

Laugen, A. T., A. Laurila, K. Räsänen, and J. Merilä. 2003. Latitudinal countergradient variation in the common frog (*Rana temporaria*) development rates – evidence for local adaptation. J. Evol. Biol. 16:996–1005.

Laurila, A., B. Lindgren, and A. T. Laugen. 2008. Antipredator defenses along a latitudinal gradient in *Rana temporaria*. Ecology 89:1399–1413.

Latimer, C. A. L., R. S. Wilson, and S. F. Chenoweth. 2011. Quantitative genetic variation for thermal performance curves within and among natural populations of *Drosophila serrata*. J. Evol. Biol. 24: 965–975.

Le Galliard, J. F., J. Clobert, and R. Ferrière. 2004. Physical performance and Darwinian fitness in lizards. Nature 432:502–505.

Leal, M., and A. R. Gunderson. 2012. Rapid change in the thermal tolerance of a tropical lizard. Am. Nat. 180:815–822.

Levins, R. 1968. Evolution in changing environments: some theoretical explorations. Princeton Univ. Press, Princeton, NJ.

Lind, M. I., P. K. Ingvarsson, H. Johansson, D. Hall, and F. Johansson. 2011. Gene flow and phenotypic in an island system of *Rana temporaria*. Evolution 65:684–697.

Lindgren, B., and A. Laurila. 2005. Proximate causes of adaptive growth rates: growth efficiency variation among latitudinal populations of *Rana temporaria*. J. Evol. Biol. 18:820–828.

———. 2009. Physiological variation along latitudinal gradients: standard metabolic rate in *Rana temporaria* tadpoles. Biol. J. Linn. Soc. Lond. 98:217–224.

Logan, M. L., R. M. Cox, and R. Calsbeek. 2014. Natural selection on thermal performance in a novel thermal environment. Proc. Natl. Acad. Sci. USA 111:14165–14169.

Loman, J. 2003. Growth and development of larval *Rana temporaria*: local variation and countergradient selection. J. Herpetol. 37:595–602.

Lutterschmidt, W. I., and V. H. Hutchison. 1997. The critical thermal maximum: history and critique. Can. J. Zool. 75:1561–1574.

Lynch, M., and W. Gabriel. 1987. Environmental tolerance. Am. Nat. 129:283–303.

Martin, T. L., and R. B. Huey. 2008. Why "suboptimal" is optimal: Jensen's inequality and ectotherm thermal preferences. Am. Nat. 171:E102–E118.

Merilä, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. J. Evol. Biol. 14:892–903.

Muir, A. P., R. Biek, R. Thomas, and B. K. Mable. 2014a. Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria)*. Mol. Ecol. 23:561–574.

Muir, A. P., R. Biek, and B. K. Mable. 2014b. Behavioural and physiological adaptations to low-temperature environments in the common frog, *Rana temporaria*. BMC Evol. Biol. 14:110.

Munch, S. B., and D. O. Conover. 2003. Rapid growth results in increased susceptibility to predation in *Menidia menidia*. Evolution 57:2119–2127.

Murren, C. J., H. J. Maclean, S. E. Diamond, U. K. Steiner, M. A. Heskel, C. A. Handelsman, C. K. Ghalambor, J. R. Auld, H. S. Callahan, D. W. Pfennig, et al. 2014. Evolutionary change in continuous reaction norms. Am. Nat. 183:453–467.

Niehaus, A. C., M. J. Angilletta, M. W. Sears, C. E. Franklin, and R. S. Wilson. 2012. Predicting the physiology performance of ectotherms in fluctuating thermal environments. J. Exp. Biol. 215:694–701.

Nilsson-Örtman, V. N., R. Stoks, M. deBlock, and F. Johansson. 2013. Latitudinal patterns of phenology and age-specific thermal performance across six *Coenagrion* damselfly species. Ecol. Monogr. 83:491–510.

Orizaola, G., M. Quintela, and A. Laurila. 2010. Climatic adaptation in an isolated and genetically impoverished amphibian population. Ecography 33:730–737.

Paaijmans, K. P., R. L. Heinig, R. A. Seliga, J. I. Blanford, S. Blanford, C. C. Murdock, and M. B. Thomas. 2013. Temperature variation makes ectotherms more sensitive to climate change. Glob. Chang. Biol. 19:2373–2380.

Pakkasmaa, S., J. Merilä, and R. B. O'Hara. 2003. Genetic and maternal effect influences on viability of common frog tadpoles under different environmental conditions. Heredity 91:117–124.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetics software for teaching and research. Mol. Ecol. Notes 6:288–295.

Ragland, G. J, and Kingsolver. J. G. 2008. The effect of fluctuating temperatures on ectotherm life-history traits:comparisons among geographic populations of *Wyeomyia smithii*. Evolutionary Ecology Research 10:29–44.

Richardson, J. L., M. C. Urban, D. I. Bolnick, and D. K. Skelly. 2014. Microgeographic adaptation and the spatial scale of evolution. Trends. Ecol. Evol. 29:165–176.

Richter-Boix, A., C. Teplitsky, B. Rogell, and A. Laurila. 2010. Local selection modifies phenotypic divergence among *Rana temporaria* populations in the presence of gene flow. Mol. Ecol. 19:716–731.

Richter-Boix, A., M. Quintela, G. Segelbacher, and A. Laurila. 2011a. Genetic analysis of differentiation among breeding ponds reveals a candidate gene for local adaptation in *Rana arvalis*. Mol. Ecol. 20:1582–1600.

Richter-Boix, A., M. Tejedo, and E. L. Rezende. 2011b. Evolution and plasticity of anuran larval development in response to desiccation. A comparative analysis. Ecol. Evol. 1:15–25.

Richter-Boix, A., M. Quintela, M. Kierczak, M. Franc, and A. Laurila. 2013. Fine-grained adaptive divergence in an amphibian: genetic basis of phenotypic divergence and the role of non-random gene flow in restricting effective migration among wetlands. Mol. Ecol. 22:1322–1340.

Richter-Boix, A., G. Orizaola, and A. Laurila. 2014. Transgeneration phenotypic plasticity links breeding phenology with offspring life-history. Ecology 95:2715–2722.

Ruel J. J., and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. Trends Ecol. Evol. 14:361–366.

Savage, V. M., J. F. Gillooly, J. H. Brown, G. B. West, and E. L. Charnov. 2004. Effects of body size and temperature on population growth. Am. Nat. 163:E429–E441.

Schielzeth, H., and S. Nakagawa. 2013. Nested by design: model fitting and interpretation in a mixed model era. Methods Ecol. Evol. 4:14–24.

Scholander, P. F., W. Flagg, V. Walters, and L. Irving. 1953. Climatic adaptation in Arctic and tropical poikilotherms. Physiol. Zool. 26:67–92.

Semlitsch, R. D., D. E. Scott, and J. H. K. Pechmann. 1988. Time and size at metamorphosis related to adult fitness in *Ambystoma talpoideum*. Ecology 69:184–192.

Skelly, D. K. 2004. Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. Evolution 58:160–165.

Skelly, D. K., and L. K. Freidenburg. 2000. Effects of beaver on the thermal biology of an amphibian. Ecol. Lett. 3:483–486.

Smith, D. C. 1987. Adult recruitment in chorus frogs: effects of size and date at metamorphosis. Ecology 68:344–350.

Somero, G. N. 2005. Linking biogeography to physiology: evolutionary and acclamatory adjustments of thermal limits. Front. Zool. 2:16.

Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135:367–374.

Stefansson, T. S., B. A. McDonald, and Y. Willi. 2013. Local adaptation and evolutionary potential along a temperature gradient in the fungal pathogen *Rhynchosporium commune*. Evol. Appl. 6:524–534.

Valladares, F, D. Sanchez-Gomez, and M. A. Zavala. 2006. Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications. J. Ecol. 94:1103–1116.

van Berkum, F. H. 1986. Evolutionary patterns of thermal sensitivity of sprint speed in *Anolis* lizards. Evolution 40:596–604.

Ward, J. V., and Stanford, J. A. 1982. Thermal responses in the evolutionary ecology of aquatic insects. Annual Review of Entomology 27:97–117.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Whitlock, M. C., and F. Guillaume. 2009. Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. Genetics 183:1055–1063.

Williams, S. E., L. P. Shoo, J. L. Isaac, A. A. Hoffmann, and G. Langham. 2008. Towards an integrated framework for assessing the vulnerability of species to climate change. PLoS Biol. 6:2621–2626. doi:10.1371/journal.pbio.0060325.

Wolak, M. E., D. J. Fairbairn, and Y. R. Paulsen. 2012. Guidelines for estimating repeatability. Methods Ecol. Evol. 3:129–137.

Yamahira, K., M. Kawajiri, K. Takeshi, and T. Irie. 2007. Inter- and intrapopulation variation in thermal reaction norms for growth rate: evolution of latitudinal compensation in ectothems with genetic constraints. Evolution 61:1577–1589.

Yampolsky, L. Y., T. M. M. Schaer, and E. D. Ebert. 2014. Adaptive phenotypic plasticity and local adaptation for temperature tolerance in freshwater zooplankton. Proc. R. Soc. Lond. B Bol. Sci. 281:20132744.

Associate Editor: W. Blanckenhorn
Handling Editor: R. Shaw

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

S1. Seasonality or/and temperature, how to isolate both factors?
S2. TDRN experiments: Q_{ST}–F_{ST} statistics and results (F_{ST} table, Q_{ST}–F_{ST} figures).
S3. Repeatibility of swimming performance.
S4. Correlation figures between environmental parameters and thermal environments.
S5. TPCs analyses: TMV parameters.
S6. Q_{ST}–F_{ST} figures CT_{max}.
S7. Correlation figures between phenotypic and environmental parameters.

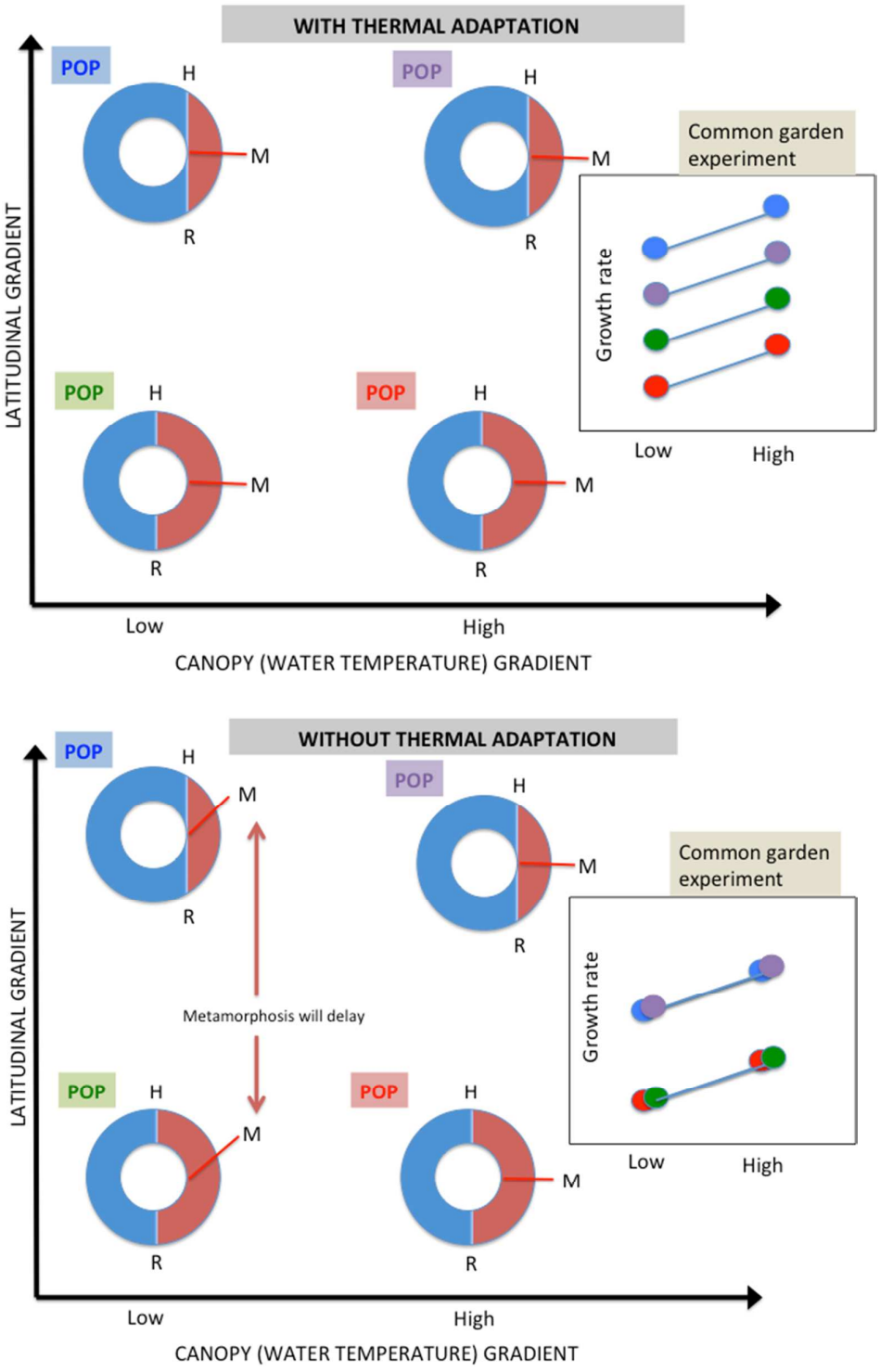
Supplementary Materials

Local temperature variation determines divergence of thermal reaction norms among local amphibian populations

Alex Richter-Boix; Marco Katzenberger; Helder Duarte; María Quintela; Miguel Tejedo; Anssi Laurila

S1 – Seasonality or/and Temperature, how to isolate both factors?

Time-constraint is always interacting with temperature in ectotherms, but here we want to make a distinction between "seasonality" and "temperature", considering that "seasonality" represents the effective growth season limited by temperatures which are outside the species' temperature tolerance, and where species cannot adjust the phenotype. We illustrate this with the figures below. They represent four populations (2 from north, 2 from south; and populations within a latitude differ in temperature environments: low and high temperature). The red area defines the effective growth season (starting with reproduction [R] and finishing with hibernation [H], with metamorphosis [M] in the middle) according to the species' temperature tolerance (blue color shows the time when temperature is below the temperature allowing for growth and development). The point is that when comparing northern populations against southern ones it is impossible to remove the "season effect". Even without temperature adaptation northern populations will grow faster to compensate for the time-constraints of the shorter growth season (see reactions norm in the common garden experiment). However, if we focus on populations at the same latitude (with similar breeding time) but from different temperature environments, population from the cold environment with temperature adaptation will grow faster to compensate for lower temperature and metamorphose at same time as the warm population. Without thermal adaptation (same values and reaction norms in the common garden experiment), metamorphosis will be delayed in the nature with less time to grow as a juvenile until the hibernation period. Obviously, in this case the "force" to grow faster is again the time-constraint to prevent a strong delay in metamorphosis, but the time-constraint here is only due to thermal environment and adaptation can be more readily identified.



S2 – TDRN EXPERIMENTS: Q_{ST}-F_{ST} statistics and results (FST table; QST – FST figures).

To test whether life-history traits are under local directional or homogenizing selection, we used the common approach to compare population differentiation estimates at quantitative traits (Q_{ST}) and neutral molecular markers (F_{ST}). Q_{ST} measures quantitative genetic variation in a manner, which is expected to be equal to F_{ST} for neutral traits (Whitlock 1999). A correct estimation of Q_{ST} requires one to disentangle genetic variation among populations from environmental variation; thus, a common garden system is a good research design to study Q_{ST}. However, common-garden experiments are not perfect, because it is possible that plastic responses to a habitat are part of the mechanism for local adaptation (Räsänen et al. 2003). If there is plasticity for the trait under study, the genetic divergence among populations may not be properly expressed in a common-garden setting because unpredictability respond to the common-garden environment (Withlock 2008). It has been demonstrated experimentally that Q_{ST} could be influenced by genotype × environment interactions, suggesting that such an impact of phenotypic plasticity on Q_{ST} and F_{ST} should be tested more often (Cano et al. 2004). Recently, the practise of comparing F_{ST} and Q_{ST} has been criticised on several ground (Pujol et al., 2008; Whitlock, 2008; Edelaar et al. 2011). Most importantly, individuals must be raised in a common garden in order to estimate the genetic component of the phenotypic differences (Pujol et al., 2008) and maternal effects need also to be controlled for (Whitlock, 2008). This was done in the analyses, by using a common garden experiment with two food treatments and controlling for maternal effects transferred through egg size. Assuming random mating and that the inbreeding coefficient was zero, we computed Q_{ST} according to the formula by Spitze (1993).

$$Q_{ST} = \frac{\sigma_B^2}{(\sigma_B^2 + 2\sigma_W^2)} \qquad \text{[eq. 1]}$$

, where σ_B^2 is the component of additive genetic variance between populations and σ_W^2 the component within populations. σ_B^2 was estimated by V_{pop}, the observed component of variance between populations and σ_W^2 was estimated by V_{fam}, the observed component of variance between families assuming full-sib families: $\sigma_W^2 = 2V_{fam}$.(Lynch and Walsh 1998). As a work revealed that multiple paternities in *R. arvalis* are relatively high (around 20% of the egg clutches; Knopp and Merillä 2009), we used an alternative conservation approach where σ_W^2 was estimated by $3V_{fam}$. Q_{ST} were thus estimated using the following term:

$$Q_{ST} = \frac{V_{pop}}{(V_{pop} + 6V_{fam})} \qquad \text{[eq.2]}$$

As recent studies have shown that Bayesian methods yield the most precise estimates of the confidence intervals of Q_{ST} (O'Hara and Merilä 2005), we used a Bayesian approach to calculate point estimates and confidence intervals of Q_{ST}. Considering the following nested random effects linear model

$$Y_{ijk} = \mu + p_i + f_{ij} + e_{ijk} \qquad \text{[eq. 3]}$$

, where Y_{ijk} is the observed quantitative trait measurement of individual k belonging to family j at population i, μ the overall mean, p_i the population effect at population i, f_{ij} the family effect of family j in population i, and e_{ijk} the residual. The model was fitted using the *MCMCglmm* function in the R package *MCMCglmm* (Hadfield 2010).

This function was used to generate a sample from the posterior distribution of the parameters of the fitted model using Markov Chain Monte Carlo (MCMC) methods, and the *coda* package in R to perform output analysis and convergence diagnostics for the MCMC simulations. With Bayesian analysis the full posterior distribution of a statistic of interest is obtained, we therefore calculated Q_{ST} mode, mean and the 95% highest posterior density (HPD).

We tested the null hypothesis that there is no difference between the estimated Q_{ST} from each trait for expected neutrally evolving traits using the approach of Whitlock and Guillaume (2009), using the R-code script providing by Lind and co-workers (2011). To estimate the significance we calculated the expected among-population variance component for a neutral trait using the observed values of F_{ST} and the within-population variance. The difference between both statistics (Q_{ST} - F_{ST}) was simulated 50,000 times to create the distribution of a neutral trait, with which compare the observed Q_{ST} - F_{ST}.

Table S2.1 with the pairwise Euclidean geographic distance in km in the upper part of the diagonal, and genetic differentiation estimated among breeding sites (F_{ST}) for the neutral microsatellites used in the molecular analyses below the diagonal. The average of the F_{ST} was 0.016 with a standard deviation of 0.012. The method of Weir and Cockerham (1984) was used, because genetic variance among populations is calculated in the same way than Q_{ST} values (Whitlock 2008).

s1	s1	s2	s3	s4	s5	s6	s7
s1	0	35.64	24.35	21.77	27.69	31.08	25.27
s2	0.01908	0	13.53	13.98	20.61	15.58	10.04
s3	0.01951	0.01430	0	5.04	9.89	8.64	17.23
s4	0.03333	0.00941	0.02330	0	14.58	13.64	12.55
s5	0.01464	0.01127	0.02435	0.01840	0	5.87	27.07
s6	0.00709	0.00000	0.00593	0.00579	0.00369	0	25.19
s7	0.05001	0.01403	0.01775	0.01003	0.03715	0.01811	0

The capacity for our relatively low number of loci to detect fine-scale patterns of population structure was evaluated by using the software POWSIM (Ryman and Palm 2006) in a previous study including the whole 17 ponds network (Richter-Boix et al. 2013). We simulated population drift to F_{ST} levels of 0.00, 0.01 and 0.02 among localities with 6 loci and our data set characteristics (17 localities and respective number of individuals), two effective population sizes (N_e 200 and 500), and three divergence times (0, 5 and 10 generations). Simulations indicated that our data set had sufficient statistical power to detect the level of genetic differentiation observed (observed global mean F_{ST} = 0.0374) with a probability that varied from 88% to 93%. Population structure would also be detected when F_{ST} values were as low as 0.01(>95% for chi-squared test and Fisher's exact test) or 0.02 (>97%).

Figure S2.1 Mantel test result between FST and geographic distances: Isolation By Distance (IBD). Histogram represents the distribution of the 10,000 simulated correlations.

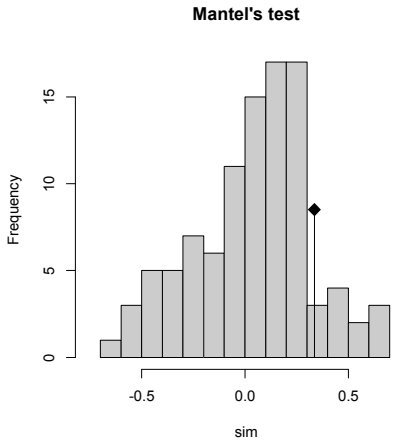


Figure S2.2 with the Q_{ST}-F_{ST} results for the TDRN experiment. The High Temperature corresponds to the 20ºC treatment, and the Low Temperature to the 15ºC treatment. Plasticity was estimated using one the phenotypic plasticity index:

([*mean LowT – individuals at HighT]/mean LowTemp*).

Histograms represent the distribution of the 50,000 simulated (Q_{ST} - F_{ST}) values for a neutral trait. Red arrow represents the difference between the obtained two statistics (Q_{ST} - F_{ST}) (Whitlock and Guillaume, 2009), and blue arrows represent the observed QST value.

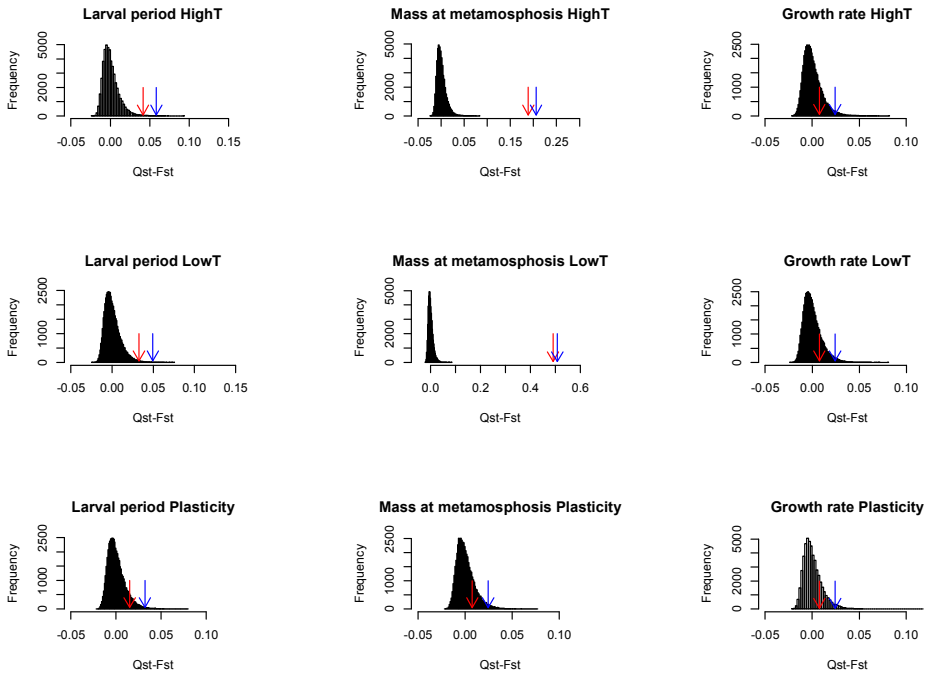


Table S2.2 with the Q_{ST}-F_{ST} results for the TDRN experiment.

Trait	20ºC		15ºC		Plasticity	
	Q_{ST}-*value*	$P_{QST>FST}$	Q_{ST}-*value*	$P_{QST>FST}$	Q_{ST}-*value*	$P_{QST>FST}$
Larval period	0.058	**0.0032**	0.049	**0.0088**	0.018	0.0792
Mass at metamorphosis	0.507	**<0.0001**	0.205	**<0.0001**	0.024	0.1964
Growth rate	0.022	0.1893	0.024	0.1941	0.023	0.1958

REFERENCES S2

Cano JM, Laurila A, Palo JU, Merilä J. 2004. Population differentiation in G matrix structure in response to natural selection in *Rana temporaria*. Evolution 58:2013–2020.

Edelaar P, Burraco P, Gomez-Mestre, I. 2011. Comparison between Qst and Fst – how wrong have we been? Molecular Ecology 20: 4830–4839.

Hadfield JD. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software 33: 1–22.

Knopp T, Merilä J. 2009. Multiple paternity in the moor frog, *Rana arvalis*. Amphibia-Reptilia 30: 515–521.

Lind MI, Ingvarsson PK, Johansson H, Hall D, Johansson F. 2011. Gene flow and selection on phenotypic plasticity in an island system of *Rana temporaria*. Evolution 65: 684–697.

Lynch M, Walsh JB. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associations.

O'Hara RB, Merilä J. 2005. Bias and precision in Qst estimates: problems and some solutions. Genetics 171: 1331–1339.

Pujol B, Wilson AJ, Ross RIC, Pannell R. 2008. Are Qst–Fst comparisons for natural populations meaningful? Molecular Ecology 17: 4782–4785.

Räsänen K, Laurila A, Merilä J. 2003. Geographic variation in acid stress tolerance of the moor frog Rana arvalis. II. Adaptive maternal effects. Evolution 57: 363–371.

Richter-Boix A, Quintela M, Kierczak M, Franch M, Laurila A. 2013. Fine-grained adaptive divergence in an amphibian: genetic basis of phenotypic divergence and the role of nonrandom gene flow in restricting effective migration among wetlands. Molecular Ecology 22: 1322–1340.

Ryman N, Palm S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6: 600–602.

Spitze K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135:367–374.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Whitlock MC. 2008. Evolutionary inference from Q_{ST}. Molecular Ecology 17: 1885-1896.

Whitlock MC, Guillaume F. 2009. Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. Genetics 183: 1055-1063.

S3 – REPEATIBILITY OF SWIMMING PERFORMANCE

Repeatability of Swimming performance measure for the Thermal performance curves (TPCs): we estimated repeatability like the fraction of thermal swimming performance variation that is due to differences among individuals. We used the most broadly used magnitude know as Interclass Correlation Coefficient (ICC):

$$ICC = \frac{S_A^2}{S_A^2 + S_W^2}$$

where S_A^2 is the variance among individuals and S_W^2 is the variance within individuals over time (Sokal & Rohlf 1995; Bell et al. 2009). In our case, the ICC was estimated from three independent measurements per individual. A limited number of individuals per population were selected to repeat the swimming performance test across all temperatures, once all individuals' performance was finish at each temperature. Six individuals per population, representing 20% of individuals, were selected to estimate swimming performance repeatability. Mortality during the experiment reduced the 42 individuals used in the repeatability test in the beginning to 36 and the end (17% of the sample). For the 36 remaining individuals we obtained three different swimming performances at different times in all temperature conditions used in TPCs estimation.

For the analyses of repeatability we used the R package *ICC* (Wolak et al. 2012; http://r-forge.r-project.org/projects/icc.), with the function *ICCbare* to estimate the ICC, and the function *ICCest* to estimate the confidence intervals for an alpha = 0.05.

We obtained an ICC coefficient of 0.852 with a LowerIC = 0.813 and an UpperIC = 0.886, demonstrating that repeatability was good in our study. Among individual variance was 2.95, much higher than the within individual variance estimated of 0.51. We obtained a high ICC value, which suggested that the individual's swimming performance measured is partly canalized, giving more accuracy in our discussion about differences among populations. The low within individual variance estimated and the higher among individual variance, suggested that individual swimming performance at different temperatures is a behavior consistently through time.

REFERENCES S3

Bell AM, Hankison SJ, Laskowski KL. 2009. The repeatability of behavior: a meta-analysis. Animal Behaviour 77: 771–783.

Sokal RR, Rohlf FJ. 1995. Biometry the Principles and Practice of Statistics in Biological Research, 3rd edn. WH Freeman and Company, New York.

Wolak ME, Fairbairn DJ, Paulsen YR. 2012. Guidelines for estimating repeatability. Methods in Ecology and Evolution 3: 129–137.

S4 – CORRELATION FIGURES BETWEEN ENVIRONMENTAL PARAMETERS AND THERMAL ENVIRONMENTS.

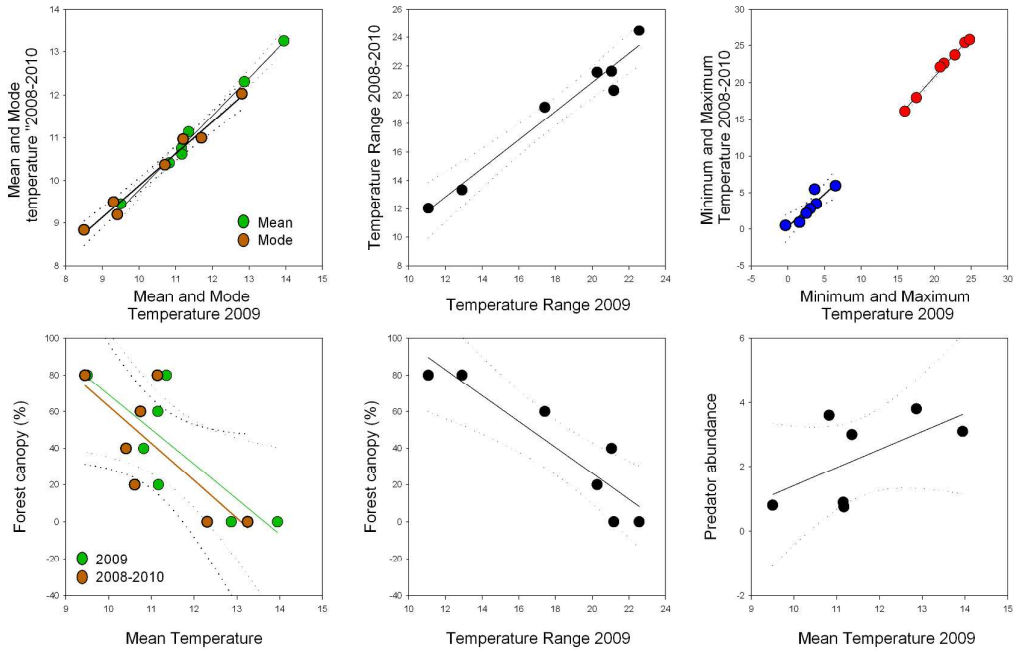


Figure S4 correlations between environmental parameters like mean and range water temperature with percentage of forest canopy cover or predator abundance. First row include the correlations between temperatures measured in 2009 with temperatures during the period 2008–2010 to demonstrate the high correlation between both parameters and that it is indifferent to use only data from 2009 or the three years series. Warmer ponds in 2009 were warmer also in 2008 and 2010 than the 2009s colder ponds.

S5 – TPCs ANALYSES: TMV PARAMETERS

Table S5a with the swimming speed measurement by population and temperature

			Swimming Speed (cm/s ± SE)				
Population	N	TTL (mm ± SE)	15°C	20°C	24°C	28°C	32°C
s1	30	37.1 ± 0.4	18.3 ± 0.7	28.9 ± 1.2	33.6 ± 1.7	39.3 ± 1.6	21.3 ± 2.3
s2	30	37.0 ± 0.4	16.0 ± 0.7	27.6 ± 1.5	36.2 ± 1.8	25.7 ± 0.8	19.4 ± 1.6
s3	30	35.5 ± 0.5	17.4 ± 0.9	30.9 ± 1.7	38.9 ± 1.9	29.5 ± 1.5	23.1 ± 1.9
s4	31	38.6 ± 0.3	13.9 ± 0.7	22.9 ± 1.5	23.9 ± 1.2	29.8 ± 1.5	20.8 ± 1.5
s5	31	34.9 ± 0.4	17.0 ± 0.9	30.5 ± 1.8	38.0 ± 1.9	28.2 ± 1.2	24.2 ± 1.5
s6	30	39.3 ± 0.5	16.1 ± 0.8	28.7 ± 1.8	36.1 ± 2.0	26.3 ± 1.0	22.9 ± 1.1
s7	30	38.9 ± 0.4	15.2 ± 0.7	25.1 ± 0.9	26.2 ± 1.7	33.2 ± 1.7	21.6 ± 1.0

Table S5b with the TMV parameters obtained from the TPCs experiment. W = width (dimensionless); m(T_{opt}) = optimal temperature in ºC; h = height (log TTL/s); $Z_{max(log)}$ = maximum performance (log TTL/s); Z_{max} = maximum performance (TTL/s); B_{95} = thermal performance breadth (ºC) at 95% of performance; B_{80} = thermal performance breadth (ºC) at 95% of performance.

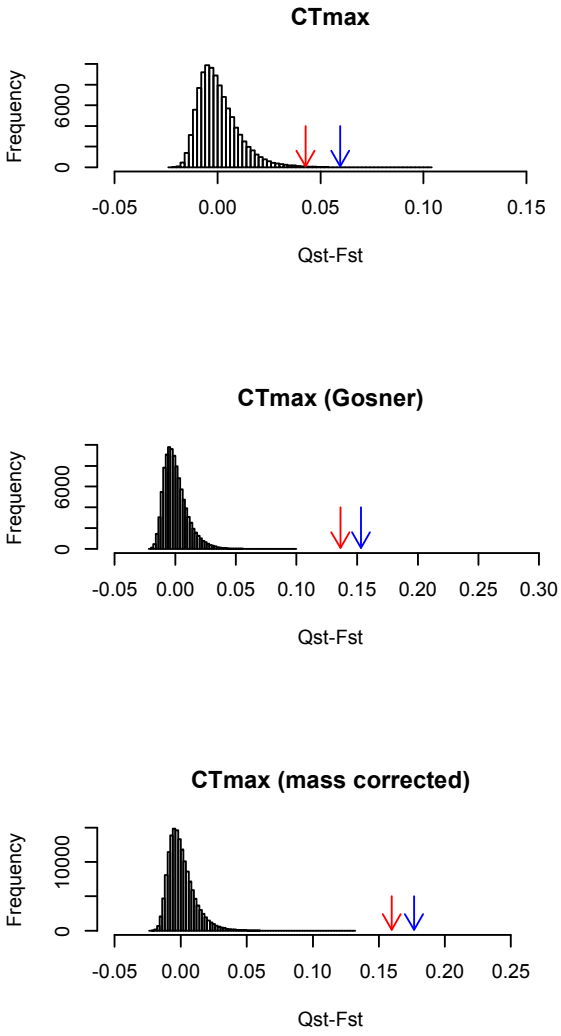
	TMV parameters						
Population	w	m (T_{opt} °C)	h	$z_{max(log)}$	Z_{max}	B_{95}	B_{80}
s1	0.9489	24.2793	0.0014	0.9885	9.7408	6.4612	13.4174
s2	0.9431	23.9903	-0.0597	0.9335	8.5803	6.2163	12.8740
s3	0.9557	24.2584	0.0220	1.0022	10.0511	6.5770	13.6710
s4	1.0907	25.9512	-0.0208	0.8379	6.8853	7.3289	15.2016
s5	0.9976	24.6977	0.0546	0.9936	9.8551	6.9869	14.5471
s6	1.0170	24.8062	-0.0065	0.9144	8.2124	6.8944	14.3115
s7	1.0832	25.7098	0.0088	0.8736	7.4751	7.4087	15.3911

The results of the decomposition of variation were:

Generalist-specialist	30.93
Hotter-colder	8.66
Faster-slower	29.03
Model Total	**68.62**

S6 – Q_{ST}-F_{ST} FIGURES CT_{max}

Figure S6 with the Q_{ST}-F_{ST} results of the CT_{max} experiment and the different corrections used to the measured CT_{max}: Gosner stage and mass. Histograms represent the distribution of the 50,000 simulated (Q_{ST} - F_{ST}) values for a neutral trait. Red arrow represents the difference between the obtained two statistics (Q_{ST} - F_{ST}) (Whitlock and Guillaume, 2009), and blue arrows represent the observed QST value.



S7 – CORRELATION FIGURES BETWEEN PHENOTYPIC AND ENVIRONMENTAL PARAMETERS.

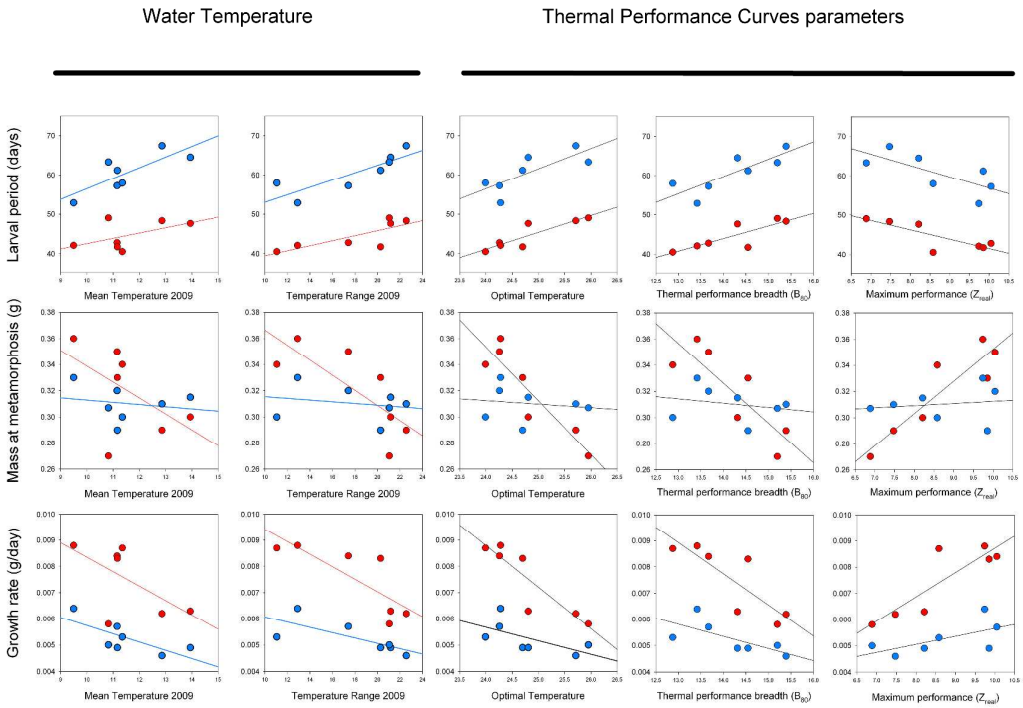


Figure 7 with the relationship between the life-history traits and water temperature parameters and thermal performance curves parameters. Blue circles correspond to traits at the 15 ºC treatment, and red circles at the 20 ºC treatment.