Heredity **87** (2001) 410–420 *Received 31 October 2000, accepted 8 May 2001*

Lack of response to artificial selection on the slope of reaction norms for seasonal polyphenism in the butterfly *Bicyclus anynana*

PIETER J. WIJNGAARDEN* & PAUL M. BRAKEFIELD

Institute of Evolutionary and Ecological Sciences, Leiden University, PO Box 9516, NL-2300 RA Leiden, The Netherlands

The tropical butterfly *Bicyclus anynana* shows adaptive phenotypic plasticity in response to wet–dry seasonality. The wet season form (WSF) has a conspicuous wing pattern with large eyespots, whereas the dry season form (DSF) lacks eyespots and therefore has a more cryptic appearance. Temperature is the main factor controlling this difference: rearing larvae at a low (<19°C) temperature in the laboratory results in the DSF, whereas rearing at a high (>23°C) temperature induces the WSF. We applied truncation selection in opposite directions in successive generations reared at two alternating temperatures (18.5°C and 23.5°C) to increase (for two High Plasticity (HP) lines), and decrease (for two Low Plasticity (LP) lines) wing pattern plasticity. Plasticity was assessed by partitioning full-sib families over four rearing temperatures (18.5°C, 20.5°C, 21.5°C and 23.5°C). Several wing pattern elements were measured for which the first principal component (PC1) provides a useful summary. The slopes of reaction norms for PC1 were significantly steeper in the HP lines than in the LP lines; however, the selection lines did not always differ significantly from the unselected stock. The results of crosses between the replicates of the selection lines gave no indication for effects of inbreeding. We argue that high, positive genetic correlations across temperatures retard a response to selection in opposite directions in different environments. This is discussed with respect to potential evolutionary constraints in natural populations in these butterflies.

Keywords: artificial selection, *Bicyclus anynana*, phenotypic plasticity, reaction norms, seasonal polyphenism, wing pattern.

Introduction

Seasonal changes in their environment can have a profound impact on organisms. Habitats may change drastically and selection pressures are then likely to change as well. The cyclic nature of these changes makes adaptation by genetic differentiation a 'clumsy way to deal with an environment which is both seasonal and predictable' (Shapiro, 1984; p. 297) because the genetic composition of a population reflects the selective regime of the previous season (Levins, 1968; Shapiro, 1984). This 'lag' or 'load' can be reduced when '… individuals are equally competent to make correct developmental decisions and that the cue(s) they respond to in the environment is (are) trustworthy.' (Shapiro, 1984; p. 297). Phenotypic plasticity (the ability of a single genotype to express different phenotypes in different environments) may thus evolve as an adaptive response to seasonality, especially when a change of seasons can be anticipated (Moran, 1992). Plasticity is a consequence of a genotype coding not for a fixed phenotype, but for a reaction norm, i.e. a set of phenotypes across an environmental gradient (Schlichting & Pigliucci, 1998).

The butterfly *Bicyclus anynana* shows phenotypic plasticity in its wing pattern in response to wet–dry seasonality. The wet season form (WSF) has large eyespots and a broad, white median band on the ventral side of its wings; in the dry season form (DSF) the eyespots and the band are greatly reduced in size. The large eyespots of the WSF are believed to act as deflection devices, whereas DSF individuals rely on crypsis. Results from field experiments (including reciprocal transplant experiments and manipulations of eyespot sizes) support this adaptive explanation (N. Reitsma & P. Brakefield,

*Correspondence and present address: Laboratory of Genetics, Wageningen University, Dreijenlaan 2, NL-6703 HA Wageningen, The Netherlands. E-mail: pieter.wijngaarden@genetics.dpw.wau.nl

unpublished; see Brakefield & French (1999)). Temperature is the main factor inducing the difference between the two forms: rearing larvae at $>23°C$ in the laboratory results in WSF butterflies, whereas rearing at $<19°C$ yields DSF adults. Intermediate phenotypes are rare in the field, but can be obtained in the laboratory by rearing at intermediate temperatures. Only the outer parts of the (continuous) reaction norms are thus exposed to selection in the field. Temperature has its effect mainly (but not solely, see Windig (1993)) through lengthening or shortening development time.

The experiments described here examined the extent to which this sensitivity is malleable by selection. We may expect a response to selection on the degree of sensitivity (i.e. slope of reaction norms) when sufficient genetic variation is available, when genetic correlations across temperatures are smaller than $+1$, and when plasticity has little or no costs (DeWitt *et al*., 1998; Van Tienderen, 1991; Via & Lande, 1985). Windig (1994a,b) conducted a split-family experiment to assess wing pattern plasticity in unselected *Bicyclus* butterflies across four temperatures (17°C, 20°C, 23°C and 28°C). He found significant (broad-sense) heritabilities for plasticity and weak genetic correlations across temperatures. These results indicate that we may expect a response to selection. Holloway & Brakefield (1995) arrived at a different conclusion, however. They selected at a single temperature (28°C) on the size of one of the eyespots on the ventral forewing and examined the correlated response on reaction norms for eyespot size using the same temperatures as Windig. Heritabilities of eyespot size at a particular temperature are substantial (0.5–0.6) leading to rapid responses to selection and highly divergent phenotypes (see also Brakefield *et al*., 1996). Because of the small amount of interaction variance Holloway & Brakefield (1995, p. 97) concluded that 'the population studied here has relatively little genetic variation available that would enable a given genome to produce a very large eyespot in the wet season and a very small eyespot in the dry season or vice versa'.

By selecting in opposite directions in alternate generations at high and low temperatures, we tried to establish High Plasticity (HP) lines, with a higher degree of plasticity than the unselected stock, and Low Plasticity (LP) lines, with average phenotypes across temperatures (i.e. no plasticity). Following selection, we assessed wing pattern plasticity by partitioning the offspring of single-pair matings from each line across four temperatures. In an additional experiment we crossed the replicates of the selection lines to eliminate effects of inbreeding. The plasticity of these crosses was assessed across two temperatures. Split-brood designs enable one to investigate both the genetic and environmental causes of (co)variation within and across temperatures. We will mainly focus on mean eyespot sizes within temperatures, however. Given the relatively small numbers of families we used differences between means will be more reliably estimated than differences between (co)variances.

Materials and methods

Study organism

The stock (S) was established from more than 80 gravid females collected in 1988 from near Nkhata Bay in Malawi (the families and selection lines used by Windig (1994a,b) and Holloway & Brakefield (1995) also originated from this stock). Each generation consisted of several hundred adult individuals so that high levels of heterozygosity are maintained (Saccheri & Bruford, 1993). Larvae were raised on young maize plants in climate rooms or cabinets with controlled temperature, high relative humidity, and a 12 h:12 h light:dark cycle. Adult butterflies were fed on mashed banana.

Selection procedure

The aim of selection is shown in Fig. 1. We set up High and Low Plasticity lines by applying synergistic and antagonistic selection (cf. Falconer, 1990), respectively, in an F_1 population (derived from the stock) at 23.5°C. Subsequent selection took place at two alternating temperatures, 18.5°C and 23.5°C. We chose these temperatures for two reasons: (1) they are comparable to the average temperatures in the field during the periods of larval development in the dry and in the wet season (Brakefield & Mazzotta, 1995), and (2) these temperatures allow sufficient phenotypic variation for selection (i.e. there are some individuals resembling the WSF at 18.5°C and some individuals resembling the DSF at 23.5°C).

Selection was done by eye, thus precluding a precise assessment of the selection intensity and the subsequent response to selection (although crude estimates can be made, see below). At 18.5°C extreme DSF individuals were selected as parents for the next generation in the High Plasticity (HP) line, whereas for the Low Plasticity (LP) line individuals resembling the WSF were selected. The offspring of parents selected at 18.5°C were reared at 23.5°C. At 23.5°C extreme WSF individuals were selected for the HP line and individuals resembling the DSF for the LP line. Each generation consisted of at least 200 individuals in each line; 30–40 individuals of each sex were selected as parents of the next generation. When phenotypic values are normally distributed this proportion amounts to a selection intensity of about 1.45 (cf. Table 2 in Becker (1984)). Two years (i.e. six generations of selection) after the first pair of lines (HP1

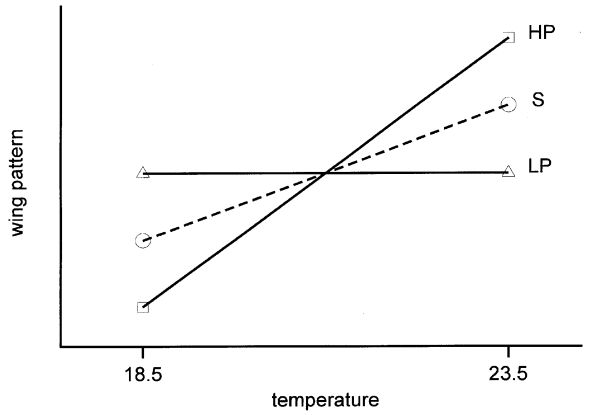

Fig. 1 The target reaction norms selected for in the experiment. Dashed line: unselected stock, squares, High Plasticity; triangles, Low Plasticity.

and LP1) had been set up a replicate pair (HP2 and LP2) was started.

Plasticity experiments

After 24 generations of selection for HP1 and LP1 (and 18 generations for HP2 and LP2) we assessed plasticity in a split-family experiment. In this experiment ('expt. 1a') we reared 16 families of the unselected stock and 10 families of each of the four selection lines (mean family sizes: 5.8–11.7 for each sex). We repeated the experiment after one more 18.5°C–23.5°C cycle ('expt. 1b'). In expt. 1b we reared 13 families of the HP2 line and 11 families of the other lines and the stock (mean family sizes: 6.1–12.1 for each sex). The offspring of single-pair matings were divided over four temperatures: 18.5°C, 20.5°C, 21.5°C and 23.5°C.

Differences in wing pattern are to some extent associated with differences in development time (high temperature leading to rapid growth); inbreeding might therefore affect wing pattern indirectly when development time shows inbreeding depression. After 36 generations of selection we crossed the replicate lines to eliminate effects of inbreeding. Reciprocal crosses were made to enable the detection of sex-specific effects. When referring to these crosses in the text the female parents will be given first (e.g. HP1 × 2 refers to a cross between HP1 females and HP2 males). These crosses were used in a plasticity experiment ('expt. 2') with 18.5°C and 23.5°C as experimental temperatures; 14–27 families were used for each line with mean family sizes ranging from 3.9 to 11.5 for each sex.

Measurements

Several measurements were made on the ventral side of both the forewing and the hindwing of the adult butterfly using a stereomicroscope fitted with a micrometer (Fig. 2). On the hindwing the diameters of the gold outer ring, the black disc, and white centre of the largest (fifth) eyespot were measured. Furthermore, the distance between the junction of the veins separating the fourth and fifth distal hindwing cells (a wing cell is an area of the wing that is bounded by veins) and the proximal edge of the median band was taken as a measure of the width of that band (which is difficult to measure directly because of its fuzzy distal edge); unlike the eyespot diameters this distance becomes smaller at

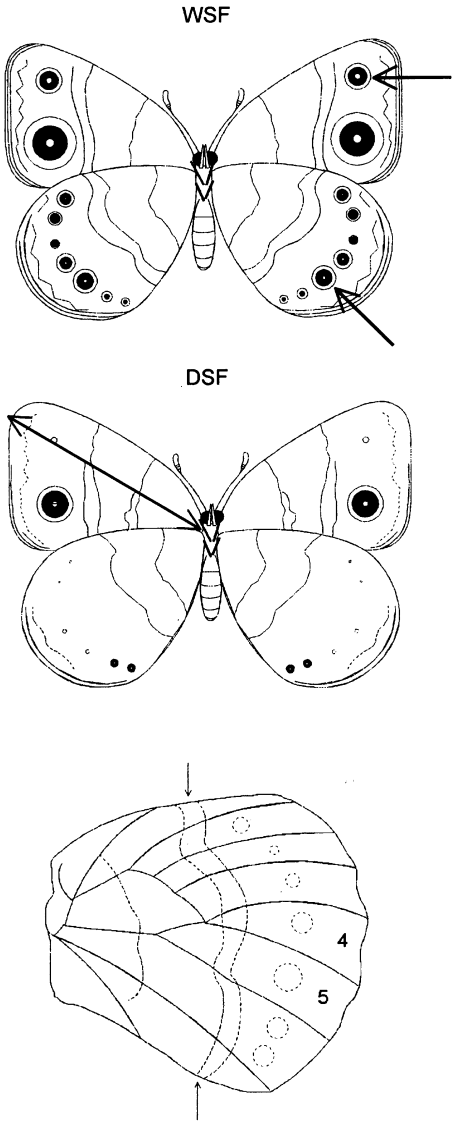

Fig. 2 Ventral view of fore- and hindwings of *Bicyclus anynana*. The second eyespot on the forewing and the fifth eyespot on the ventral hindwing are indicated by arrows in the wet season (WSF) individual. The forewing length is shown in the dry season (DSF) individual. In the hindwing the veins separating wing cells 4 and 5 and the band are indicated.

higher temperatures (implying that the band itself becomes broader). On the forewing the total diameter of the second eyespot was measured. These traits will be referred to as gold, black, white, band and FWS, respectively (note that gold includes both black and white and that black includes white).

The length of the forewing was taken as an indicator of overall size in expts 1a and 1b; the length of the vein separating the fourth and fifth distal hindwing cells was used for individuals from expt. 2 (both size measures are highly correlated).

Statistical analyses

General. We used principal component analysis to summarize the six different wing measures. The first component (PC1, accounting for 60–65% of the variation in both sexes) can be considered a measure of wing pattern since the first eigenvector shows moderately high (0.3–0.5) and approximately equal, positive loadings on gold, black, white, band and FWS. The second eigenvector has a high (0.8–1.0) positive loading on size and this component (PC2, accounting for 15–20% of the variation in both sexes) seems therefore a measure of size. Only PC1 will be used in our analyses.

Mixed models were analysed using the MIXED procedure of the SAS package. The effects of the selection lines and temperatures were treated as fixed effects; the effect of families (nested within lines) was treated as a random effect. Approximate *F*-tests of fixed effects are based on Type III sums of squares; the appropriate degrees of freedom were determined using Satterthwaite's approximation. Random effects were tested using the log-likelihood ratio test in which full and reduced models are compared. The resulting test statistic is χ^2 distributed with 1 d.f.; the *P*-value has to be halved because variances are constrained to be non-negative (Littell *et al*., 1996, p. 44).

Genetic correlations. We used two methods to estimate genetic correlations (r_g) across temperatures. Pearson's product–moment correlation of family means is both conceptually and computationally simple, but it may underestimate the absolute value of r_g for small family sizes because the between-family (co)variances may contain within-family error (co)variance (Via, 1984). Confidence intervals for r_g were computed by using the *z*-transformation (Sokal & Rohlf, 1995).

Another way of estimating r_g is based on mixed-model analysis of variance (Fry, 1992). In this approach the covariance of family means across two environments is divided by the geometric mean of the between-family variance components obtained from separate one-way ANOVAS in each environment. As the family mean correlations were all positive, we used restricted maximum likelihood (REML) to estimate these variance components (the REML algorithm sets negative estimates to zero). Significant family main effects in ANOVAS for each line and each pair of temperatures were taken as evidence for significant correlations (this test is exact only when the heritabilities in both environments are the same).

Heritabilities. Estimates of heritabilities were obtained by full-sib analysis and calculated as two times the intraclass correlation coefficient. This yields estimates that are potentially biased due to dominance, additive × additive interaction, and common environment effects (Falconer & Mackay, 1996). REML was used to estimate the between- and within-family variance components. Significant variance among families (obtained from analyses for each line and each temperature) was taken as evidence for significant heritabilities.

Results

Selection intensity

As indicated earlier, an estimate of the selection intensity is 1.45 for each generation. In the unselected stock estimates of the phenotypic standard deviations are 1.15 (females) and 0.95 (males) at both 18.5°C and 23.5°C, yielding estimates of selection differentials of 1.67 (females) and 1.38 (males) per generation. The cumulative selection differentials after 12 18.5°C–23.5°C cycles (for HP1 and LP1) are therefore 20.0 (females) and 16.5 (males) at each temperature; for HP2 and LP2 the values are 15.0 (females) and 12.4 (males).

Response to selection

In expts 1a, 1b and 2, differences among lines, families, temperatures, and interactions between temperature and lines/families were found (Table 1). At 18.5°C we selected for the following order in mean eyespot size: LP1 > LP2 > S > HP2 > HP1. We did not obtain this order in any of the experiments (Table 2, see also Fig. 3); the discrepancy is mainly due to the LP lines having lower (instead of higher) means than the unselected stock at this temperature. The selected order in mean eyespot sizes at 23.5°C is HP1 > HP2 > S > LP2 > LP1. We did not find this order (Table 2, see also Fig. 3); the LP lines always showed lower means than the stock, but the means of the HP lines were not consistently higher than the stock mean.

Selection was applied at 18.5°C and 23.5°C, so that changes at 20.5°C and 21.5°C constitute correlated responses (barring environmental effects). In Fig. 4

Table 1 Mixed model ANOVA for PC1 per experiment in *Bicyclus anynana*. Sources of variation: lines (L), families (F(L)), temperatures (T), and their interactions. The test statistic for fixed effects is the Type III *F*-value; for random effects the likelihood ratio statistic (LRS) is given (see text). Note that in expt. 2 (the crosses between replicates of the selection lines) only two temperatures were used

Expt.	Source	Type	F/LRS_f	d.f.	*P*	F/LRS_m	d.f.	*P*
1a	L	Fixed	5.07	4, 51.8	**	6.16	4, 51.6	***
	T	Fixed	708.91	3, 159	***	605.58	3, 156	***
	L × T	Fixed	5.81	12, 153	***	4.18	12, 152	***
	F(L)	Random	101.69	1	***	99.54	1	***
	F × T(L)	Random	13.02	1	***	14.89	1	***
1b	L	Fixed	11.28	4, 51.7	***	10.62	4, 51.2	***
	T	Fixed	250.26	3, 147	***	264.32	3, 152	***
	L × T	Fixed	6.11	12, 145	***	4.91	12, 150	***
	F(L)	Random	47.35	1	***	58.11	1	***
	F × T(L)	Random	14.58	1	***	15.29	1	***
2	L	Fixed	7.31	4, 84.8	***	10.42	4, 89.9	***
	T	Fixed	834.13	1, 80.9	***	684.65	1, 84.1	***
	L × T	Fixed	10.77	4, 80.1	***	5.66	4, 84.2	***
	F(L)	Random	12.16	1	***	12.04	1	***
	F × T(L)	Random	25.17	1	***	50.60	1	***

$^{**}P < 0.01$; $^{***}P < 0.001$.

mean eyespot sizes at 20.5°C and 21.5°C have been added for expts 1a and 1b. A striking feature of the LP lines is that in some cases they have lower mean PC1 values at 20.5°C than at 18.5°C; these differences are even significantly different from zero for females of both LP1 and LP2 in expt. 1b ($t_{140} = -2.65$, $P < 0.01$ and $t_{170} = -2.36$, $P < 0.05$, respectively), and for LP1 males in expt. 1b ($t_{145} = -2.20$, $P < 0.05$).

A response to selection on reaction norms requires genetic variation both within and across temperatures. Significant heritabilities within temperatures were found in most cases in both the stock and the selection lines (Table 3). In all experiments, significant family × temperature interactions (indicative of genetic variation in slopes) were found in males ($F_{15,281} = 2.12$, $P < 0.01$; $F_{10,215} = 3.29$, $P < 0.001$; $F_{26,302} = 1.83$, $P < 0.01$, for expts 1a, 1b, and 2, respectively), but not in females, of the stock. Other cases of significant family × temperature interactions are found in LP2 males (expt. 1a: $F_{9,135} = 3.33$, $P < 0.01$), HP2 females (expt. 1a: $F_{9,157} = 2.98$, $P < 0.01$), LP1 × 2 males (expt. 2: $F_{14,231} = 3.02$; $P < 0.001$), LP2 × 1 males (expt. 2: $F_{15,217} = 2.64$; $P < 0.01$), HP1 × 2 females (expt. 2: $F_{19,316} = 2.29$; $P < 0.01$), and in HP2 × 1 (expt. 2, males: $F_{17,346} = 4.28$, $P < 0.001$; females: $F_{17,294} = 2.85$, $P < 0.001$).

The cumulative selection differentials given earlier can be used to estimate realized heritabilities of slopes for PC1. However, the calculation of the cumulative selection differentials is loaded with assumptions so that the estimates (they fall between 0 and 0.1) are likely to be of little value.

Estimates of genetic correlations between values of PC1 at 18.5°C and at 23.5°C are all positive, but only half of them are significantly different from zero (Table 4). In expt. 1a the estimates are similar (0.6–0.8) among lines and between sexes; estimates based on mixed-model ANOVAS are slightly higher than the product–moment correlations. This consistency is not found in expts 1b and 2; estimates are generally smaller and more variable in these experiments than in the first experiment. Correlations among other temperatures (not shown) are often of similar magnitude. Only in two out of 20 cases is there an increase in estimates with increasing similarity of temperatures.

Had inbreeding in our selection lines resulted in longer development times and smaller eyespots, we would expect this effect to disappear in crosses between the replicate lines (expt. 2). Inspection of Fig. 3 gives no indication for effects of inbreeding (they may be overridden by environmental effects, see the shifts in the stock relative to expts 1a and 1b).

We found only weak evidence for differences between HP2 × 1 and HP1 × 2 in expt. 2. In males HP2 × 1 has larger eyespots than HP1 × 2 at both 18.5°C ($t_{83} = -2.38$, $P = 0.0194$) and 23.5°C ($t_{82.5} = -2.56$, $P = 0.0124$). In females HP2 × 1 has larger eyespots at 18.5°C ($t_{75} = -2.60$, $P = 0.0112$); no significant differences were found at 23.5°C. LP2 × 1 and

Table 2 Differences in least square means between selection lines and the unselected stock of *Bicyclus anynana* at 18.5°C (Δ18) and 23.5°C (Δ23). For expt. 2 (the crosses between the replicates) only the female parent is mentioned (e.g. HP1 is actually HP1 × 2, see text). The *t*-tests are two-sided

Expt.	Line	Δ18	SE	d.f.	*t*	*P*	Δ23	SE	d.f.	*t*	*P*
Females											
1a	HP1	1.06	0.298	49.4	3.56	**	−0.73	0.348	47.8	−2.09	NS
	HP2	0.23	0.299	50.7	0.78	NS	−0.03	0.348	47.3	−0.09	NS
	LP1	1.02	0.292	46.2	3.48	*	1.07	0.343	46	3.11	*
	LP2	0.74	0.305	54.9	2.43	NS	1.25	0.358	52.7	3.51	**
1b	HP1	0.87	0.322	53.9	2.69	NS	−0.18	0.328	47.1	−0.56	NS
	HP2	−0.08	0.296	46.4	−0.27	NS	0.93	0.309	44.1	3.00	*
	LP1	0.56	0.311	46.8	1.81	NS	2.04	0.332	48.2	6.14	***
	LP2	0.09	0.318	51.4	0.30	NS	1.32	0.322	44.7	4.09	**
2	HP1	1.38	0.278	101	4.98	***	−0.18	0.231	86.8	0.76	NS
	HP2	0.68	0.280	94.9	2.42	NS	−0.54	0.237	88.2	2.26	NS
	LP1	1.25	0.295	89.2	4.25	***	0.48	0.257	86.2	1.85	NS
	LP2	0.96	0.295	102	3.26	*	0.86	0.243	85.6	3.55	***
Males											
1a	HP1	0.72	0.224	56.9	3.23	*	−0.12	0.305	53.1	−0.39	NS
	HP2	0.78	0.218	51.6	3.57	**	0.48	0.302	51.1	1.59	NS
	LP1	0.84	0.216	49.8	3.91	**	1.11	0.300	49.5	3.69	**
	LP2	0.53	0.217	50.7	2.44	NS	1.11	0.306	52.7	3.63	**
1b	HP1	0.91	0.232	45.4	3.94	**	0.06	0.324	48.6	0.18	NS
	HP2	0.78	0.226	49.4	3.44	*	0.32	0.303	44.4	1.05	NS
	LP1	0.87	0.230	45.9	3.76	**	1.98	0.322	47.4	6.15	***
	LP2	0.63	0.239	52	2.62	NS	1.06	0.324	48.1	3.27	*
2	HP1	1.00	0.209	90.3	4.77	***	0.31	0.192	91.3	1.63	NS
	HP2	0.46	0.213	88.1	2.16	NS	−0.22	0.195	85.2	−1.11	NS
	LP1	1.09	0.225	87.5	4.84	NS	0.80	0.214	97.9	3.74	***
	LP2	0.60	0.221	89.7	2.69	***	0.83	0.208	97.6	4.01	***

NS, not significant.
*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; *P*-values are Bonferroni adjusted (Rice, 1989).

LP1 × 2 do not differ significantly in both sexes and at both temperatures.

Discussion

The results of the three experiments to detect responses to selection were often not consistent. Although great care was taken to keep experimental conditions the same, small differences in temperature or food-plant quality may have affected our results. Probably more important as a source of error is variation caused by the sampling of families. Our estimates of quantitative genetic parameters have almost certainly suffered from the small numbers of families we used and should therefore treated with caution. Windig (1993) also performed a second plasticity experiment to raise the number of *Bicyclus* families and he too found that the experiments differed significantly.

We applied synergistic selection to our HP lines and these lines showed some divergence in their mean PC1 values (i.e. steeper average reaction norms), though the differences with the unselected stock were not always significant. Our results therefore support the conclusion of Holloway & Brakefield (1995) rather than Windig's (1994b) that there is little scope in terms of standing genetic variation for producing novel slopes of reaction norms in *B. anynana*, that is slopes which are not represented in the phenotypic variation found in our unselected stock. Females responded more strongly than males, which is in accordance with the higher estimates of the selection differentials in females. It conflicts, however, with the nonsignificant estimates of family × temperature interaction in females of the unselected stock and with the estimates of heritabilities at 18.5°C and 23.5°C being higher for males than for females in expts 1a and 1b (they were very similar in expt. 2).

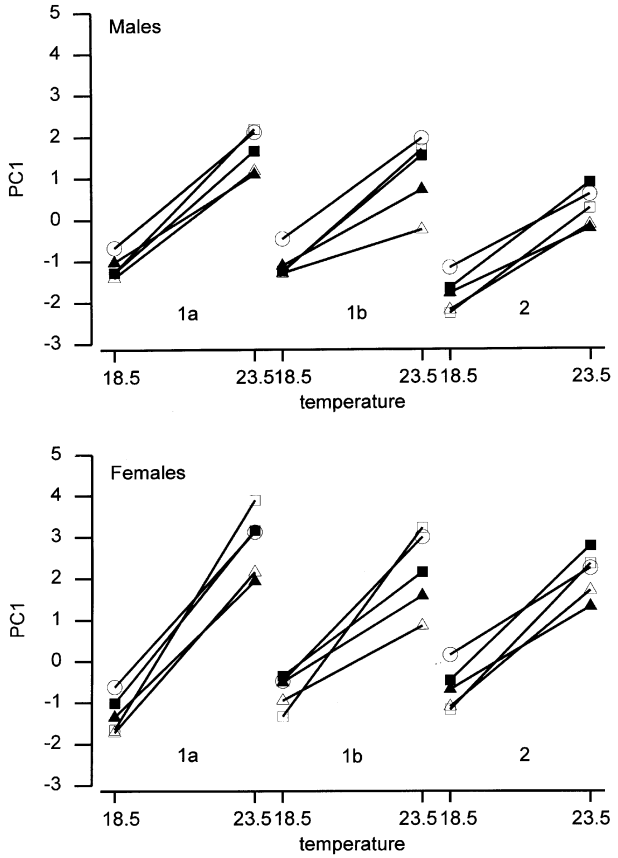

Fig. 3 Reaction norms for mean PC1 in the three experiments using pooled butterflies of each sex from each line. For expt. 2 (the crosses between the replicates) only the female parent is mentioned (e.g. HP1 is actually HP1 × 2). Circles, stock; squares, HP1; solid squares, HP2; triangles, LP1; solid triangles, LP2.

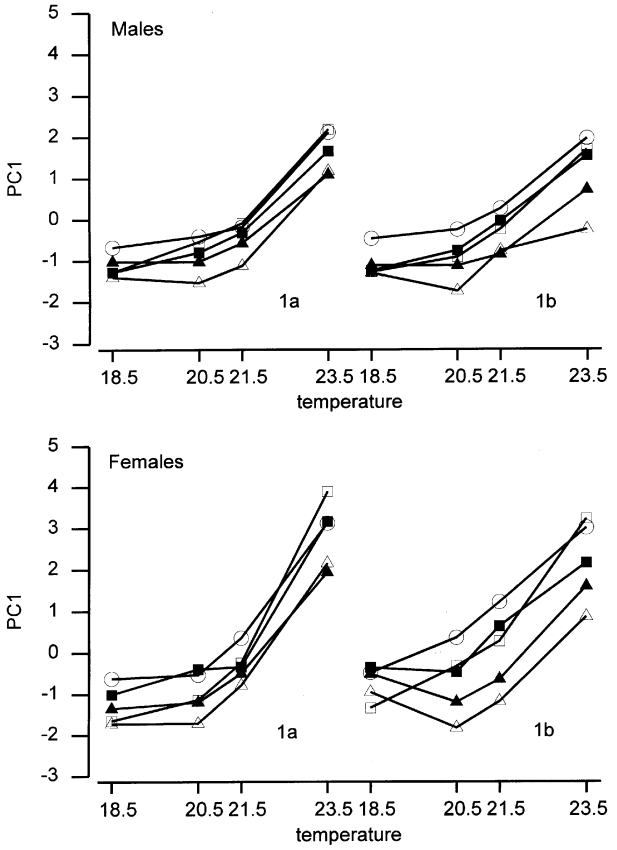

Fig. 4 Reaction norms for mean PC1 for expts 1a and 1b in *Bicyclus anynana*. Circles, stock; squares, HP1; solid squares, HP2; triangles, LP1; solid triangles, LP2.

Estimates of genetic correlations between 18.5°C and 23.5°C were (very) similar for both sexes.

That the selection lines show some divergence indicates that these genetic correlations are smaller than +1. Further evidence for positive genetic correlations across temperatures has been provided by Brakefield *et al.* (1996) and Holloway & Brakefield (1995). They selected on the size of eyespots on the ventral hindwing and forewing, respectively, at a single temperature. They found that eyespot sizes diverged not only at that particular temperature but also across a range of temperatures, i.e. the elevations of the reaction norms diverged due to correlated responses. The slopes of the reaction norms were affected to a lesser extent, however, suggesting that genetic correlations across temperatures were positive and high, but smaller than unity. Genetic correlations across environments are thought to reflect shared genetic control across temperatures (see Wijngaarden & Brakefield (2000) for details on the genetic basis of eyespot size within temperatures).

As correlated responses are generally expected to be less than direct responses (Falconer, 1990; Falconer & Mackay, 1996) it has been suggested that plasticity can be reduced by selecting upwards in a 'bad' environment and downwards in a 'good' environment (Jinks & Connolly, 1973; Jinks & Pooni, 1988). Falconer (1990) compared the results of seven published experiments where selection was applied in opposite directions in two different environments and found this 'Jinks–Conolly rule' to be more often right than wrong. Roff (1997) pointed out that the Jinks–Conolly rule will especially apply to fitness-related traits because of their expected negative correlation between slopes and elevations of reaction norms.

We did not succeed in establishing a line with medium-sized eyespots across temperatures (or even one approaching this condition). The LP lines did show smaller eyespots than the stock at 23.5°C, but the eyespots were also smaller at 18.5°C (where selection was for larger eyespots). The heritability estimates for PC1 at both temperatures were often similar, suggesting

Table 3 Full-sib heritabilities of PC1 per line, sex, experiment, and temperature (°C) in *Bicyclus anynana*. For expt. 2 (the crosses between the replicates) only the female parent is mentioned (e.g. HP1 is actually HP1 × 2, see text); 18.5°C and 23.5°C were the only experimental temperatures in this experiment

Line	Sex	Expt.	18.5	*P*	20.5	*P*	21.5	*P*	23.5	*P*
S	F	1a	0.66	***	0.71	***	0.41	***	0.54	***
		1b	0.10	NS	0.13	NS	0.29	**	0.32	**
		2	0.53	**	—	—	—	—	0.53	***
	M	1a	0.83	***	0.63	***	0.53	***	0.88	***
		1b	0.11	NS	0.19	NS	0.58	***	0.65	***
		2	0.51	***	—	—	—	—	0.53	***
LP1	F	1a	0.49	***	0.06	NS	0.18	NS	0.28	*
		1b	0.45	***	0.41	***	0.71	***	0.35	NS
		2	0.16	NS	—	—	—	—	0.48	***
	M	1a	0.18	NS	0.48	**	0.50	***	0.23	NS
		1b	0.33	**	0.04	NS	0.35	*	0.70	***
		2	0.89	***	—	—	—	—	0.61	***
LP2	F	1a	0.86	***	0.81	***	0.56	***	0.54	**
		1b	0.44	**	0.59	***	0.91	***	0.21	NS
		2	0.39	**	—	—	—	—	0.07	NS
	M	1a	0.84	***	0.80	***	0.15	NS	0.65	***
		1b	0.45	**	0.70	***	0.33	*	0.33	NS
		2	0.90	***	—	—	—	—	0.24	NS
HP1	F	1a	0.43	**	0.18	NS	0.13	NS	0.46	**
		1b	0.64	***	0.18	NS	0.27	**	0.12	NS
		2	0.53	***	—	—	—	—	0.48	***
	M	1a	0.03	NS	0.38	*	0.51	***	0.50	**
		1b	0.57	***	0.48	**	0.19	NS	0.53	***
		2	0.52	***	—	—	—	—	0.21	**
HP2	F	1a	0.58	***	0.67	***	0.47	***	1.01	***
		1b	0.24	NS	0.16	NS	0.37	***	0.45	***
		2	0.84	***	—	—	—	—	0.45	***
	M	1a	0.45	***	0.49	**	0.38	**	0.54	***
		1b	0.15	NS	0.20	NS	0.46	***	0.49	***
		2	0.73	***	—	—	—	—	0.65	***

F, female; M, male; NS, not significant.
*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; *P*-values are Bonferroni adjusted (Rice, 1989).

that selection intensities were different at the two temperatures. Due to our crude way of selecting, the selection intensity may have been higher at 23.5°C than at 18.5°C (for the calculation of the selection intensities given earlier we assumed that the selected individuals really were the most extreme), resulting in a net response towards smaller eyespots (or more DSF phenotypes in general). These results lend no support to Falconer's (1990) suggestion that in addition to reducing the environmental sensitivity, antagonistic selection will result in an improved mean performance in the two environments.

An odd feature of both LP lines is that they have lower mean PC1 values at 20.5°C than at 18.5°C (especially females in expt. 1b). A temperature effect seems unlikely because the stock and the HP lines do not show this drop in their mean PC1. Furthermore, the development times of the LP lines show a normal pattern (i.e. lower values at 20.5°C than at 18.5°C) and estimates of family mean correlations between PC1 and development time were not significantly different from zero (Wijngaarden, 2000). Reversed responses to directional selection are observed occasionally, usually in the initial stage of a selection experiment (Gimelfarb, 1986). Gimelfarb (1986) has suggested that, in addition to multiplicative genotype–environment interaction, other types of interaction as well as other nonadditive effects are possible causes of reversed responses to selection.

The results of crosses between the replicate selection lines suggest that inbreeding has not affected our results.

Expt.	Line	Sex	PMC	95% CI	MMA	95% CI
1a	S	F	0.67	0.26, 0.88	0.73	0.37, 0.90
1a	S	M	0.80	0.50, 0.93	0.74	0.39, 0.90
1a	LP1	F	0.62		0.67	
1a	LP1	M	0.63		1.03	—
1a	LP2	F	0.81	0.37, 0.95	0.86	0.50, 0.97
1a	LP2	M	0.28		0.30	
1a	HP1	F	0.63		0.70	
1a	HP1	M	0.62		1.76	—
1a	HP2	F	0.63		0.58	
1a	HP2	M	0.63		0.78	0.30, 0.92
1b	S	F	0.26		0.49	
1b	S	M	0.24		0.47	
1b	LP1	F	0.33		0.46	
1b	LP1	M	0.51		0.55	
1b	LP2	F	0.61		1.04	—
1b	LP2	M	0.38		0.87	0.56, 0.97
1b	HP1	F	0.13		0.28	
1b	HP1	M	0.44		0.69	
1b	HP2	F	0.22		0.44	
1b	HP2	M	0.32		0.65	
2	S	F	0.30		1.27	—
2	S	M	0.31		0.37	
2	LP1	F	0.50		0.30	
2	LP1	M	0.33		1.11	—
2	LP2	F	0.29		0.36	
2	LP2	M	0.32		0.86	0.63, 0.95
2	HP1	F	0.34		1.23	—
2	HP1	M	0.33		0.28	
2	HP2	F	0.41		0.83	0.59, 0.93
2	HP2	M	0.35		0.22	

Table 4 Genetic correlations between values of PC1 at 18.5°C and at 23.5°C in *Bicyclus anynana*. For expt. 2 (the crosses between the replicates) only the female parent is mentioned (e.g. HP1 is actually HP1 × 2, see text). PMC, product—moment correlations of family means; MMA, estimates based on mixed-model ANOVA (see text). 95% confidence intervals are given when $P < 0.05$ after Bonferroni adjustment

This was to be expected as only females that were able to produce four batches of hatched larvae (for each of the four temperatures) were included in the experiment. Since egg hatching is very sensitive to inbreeding (Van Oosterhout *et al*., 2000) we will have selected against inbreeding.

It is intriguing that a reversed slope of the population reaction norm was produced for a segment of the environmental range in one of our selected lines even though a shallower slope representing the production of intermediate phenotypes over the whole range is not possible. There are species of *Bicyclus* from the equatorial primary forest of West Africa which do not show individuals of the dry season form that are characteristic of all those species from regions with wet–dry seasonal environments (Condamin, 1973; Windig *et al*., 1994). We do not yet know whether these equatorial forest species if raised at low temperatures in the laboratory would produce a dry season form lacking ventral eyespots (see Roskam & Brakefield, 1996, 1999). However, they do show quantitative variation in ventral eyespot size within the broad range of the wet season form phenotype. This parallels the variation found in an old selected line of *B. anynana* which only produces the wet season form across all rearing temperatures (from 17°C to 27°C) but in which the eyespots are, on average, larger at higher temperatures; in other words, phenotypic plasticity is retained although seasonal polyphenism is no longer found (Brakefield *et al*., 1996). The apparent constraint which we have found in our present attempt to select a line with intermediate phenotypes across all temperatures may be relevant to the evolution of such species of *Bicyclus* in non-seasonal environments of equatorial Africa. The phenotypic plasticity found in *Bicyclus* species of seasonal environments in Africa appears to be characteristic of all satyrine species occurring in such environments throughout tropical and subtropical regions (see Brakefield & Larsen, 1984; K. Brown, personal communication for the neo-tropics). It has been particularly well described in *Melanitis leda* (Brakefield, 1987). It is also noteworthy that similar phenotypic plasticity may also occur in many satyrine

species in temperate regions. In at least in one well studied species, *Maniola jurtina*, individuals with longer developmental times tend to show smaller eyespots (Brakefield, 1984). Thus the underlying hormonal mechanism linking development rate and ventral eyespot expression (Brakefield *et al*., 1998) may reflect a long-term historical constraint which is very difficult to break.

The results of the present study also indicate that the standing genetic variation in our laboratory stock does not provide the basis for the short-term evolution of a bundle of reaction norms with substantially steeper slopes in response to rearing temperature. It may thus be difficult or even impossible to evolve anything approaching a sharp developmental switch in which the alternative phenotypes can be produced by single genotypes across a very narrow range of rearing temperatures. Selection experiments on *Bicyclus* in the field in Malawi and measurements of phenotypic variation in field-collected *Melanitis leda* indicate that some butterflies of the generation which must survive the dry season to reproduce nevertheless have small ventral eyespots, which decrease their chance of survival (N. Reitsma & P. M. Brakefield, unpublished). The production of such individuals with apparently suboptimal phenotypes might then represent a constraint in terms of an inability of a single genotype to produce equally favoured phenotypes with the largest, and smallest possible eyespots in the wet and dry season environments, respectively. Although such a scenario is speculative it does illustrate the potential relevance of our results from artificial selection in the laboratory to patterns of evolution in natural seasonal environments.

Acknowledgements

We are grateful to Gerdien de Jong for her suggestions concerning the design of the experiment. For their help with the rearing and measuring of the butterflies of expt. 2 we would like to thank Marcel Biesenbeek, Joep Bovenlander and Sander van der Werf. Els Schlatmann and colleagues secured a steady supply of maize for insatiable caterpillars. Peter van Tienderen and Sara Via made valuable comments on an earlier version of the manuscript.

References

BECKER, W. A. 1984. *Manual of Quantitative Genetics*, 4th edn. Academic Enterprises, Pullman, WA.

BRAKEFIELD, P. M. 1984. The ecological genetics of quantitative characters of *Maniola jurtina* and other butterflies. In: Vane-Wright, R. I. & Ackery, P. R. (eds) *The Biology of Butterflies*, pp. 167–190. Princeton University Press, Princeton, NJ, USA.

BRAKEFIELD, P. M. 1987. Tropical dry and wet season polyphenism in the butterfly *Melanitis leda* (Satyrinae): phenotypic plasticity and climatic correlates. *Biol. J. Linn. Soc*., **31**, 175–191.

BRAKEFIELD, P. M. AND FRENCH, V. 1999. Butterfly wings: the evolution of development of colour patterns. *BioEssays*, **21**, 391–401.

BRAKEFIELD, P. M. AND LARSEN, T. B. 1984. The evolutionary significance of dry and wet season forms in some tropical butterflies. *Biol. J. Linn. Soc*., **22**, 1–12.

BRAKEFIELD, P. M. AND MAZZOTTA, V. 1995. Matching field and laboratory experiments: effects of neglecting daily temperature variation on insect reaction norms. *J. Evol. Biol*., **8**, 559–573.

BRAKEFIELD, P. M., GATES, J., KEYS, D., KESBEKE, F. *ET AL*. 1996. Development, plasticity and evolution of butterfly eyespot patterns. *Nature*, **384**, 236–242.

BRAKEFIELD, P. M., KESBEKE, F. AND KOCH, P. B. 1998. The regulation of phenotypic plasticity of eyespots in the butterfly *Bicyclus anynana*. *Am. Nat*., **152**, 853–860.

CONDAMIN, M. 1973. Monographie du Genre *Bicyclus* (Lepidoptera, Satyridae). *Mem. Inst. Fond. Afr. Noire*, **88**, 1–324.

DEWITT, T. J., SIH, A. AND WILSON, D. S. 1998. Costs and limits of phenotypic plasticity. *Trends Ecol. Evol*., **13**, 77–81.

FALCONER, D. S. 1990. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genet. Res*., **56**, 57–70.

FALCONER, D. S. AND MACKAY, T. F. C. 1996. *An Introduction to Quantitative Genetics*, 4th edn. Longman, Harlow.

GIMELFARB, A. 1986. Multiplicative genotype–environment interaction as a cause of reversed response to directional selection. *Genetics*, **114**, 333–343.

HOLLOWAY, G. J. AND BRAKEFIELD, P. M. 1995. Artificial selection of reaction norms of wing pattern elements in *Bicyclus anynana*. *Heredity*, **74**, 91–99.

JINKS, J. L. AND CONNOLLY, V. 1973. Selection for specific and general response to environmental differences. *Heredity*, **30**, 33–40.

JINKS, J. L. AND POONI, H. S. 1988. The genetic basis of environmental sensitivity. In: Weir, B. S., Eisen, E. J., Goodman, M. M. and Namkoong, G. (eds) *Proceedings of the 2nd International Conference on Quantitative Genetics*, pp. 505–522. Sinauer, Sunderland, MA.

LEVINS, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton, NJ.

LITTELL, R. C., MILLIKEN, G. A., STROUP, W. W. AND WOLFINGER, R. D. 1996. *SAS System for Mixed Models*. SAS Institute Inc., Cary, NC.

MORAN, N. A. 1992. The evolutionary maintenance of alternative phenotypes. *Am. Nat*., **139**, 971–989.

VAN OOSTERHOUT, C., ZIJLSTRA, W. G., VAN HEUVEN, M. K. AND BRAKEFIELD, P. M. 2000. Inbreeding depression and genetic load in laboratory metapopulations of the butterfly *Bicyclus anynana*. *Evolution*, **54**, 218–225.

RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

ROFF, D. A. 1997. *Evolutionary Quantitative Genetics*. Chapman & Hall, New York.

ROSKAM, J. C. AND BRAKEFIELD, P. M. 1996. A comparison of temperature-induced polyphenism in African *Bicyclus* butterflies from a seasonal savannah–rainforest ecotone. *Evolution*, **50**, 2360–2372.

ROSKAM, J. C. AND BRAKEFIELD, P. M. 1999. Seasonal polyphenism in *Bicyclus* (Lepidoptera: Satyridae) butterflies: different climates need different cues. *Biol. J. Linn. Soc.*, **66**, 345–356.

SACCHERI, I. J. AND BRUFORD, M. W. 1993. DNA fingerprinting in a butterfly, *Bicyclus anynana* (Satyridae). *J. Hered.*, **84**, 195–200.

SCHLICHTING, C. D. AND PIGLIUCCI, M. 1998. *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer, Sunderland, MA.

SHAPIRO, A. M. 1984. Experimental studies on the evolution of seasonal polyphenism. In: Vane-Wright, R. I. and Ackery, P. R. (eds) *The Biology of Butterflies*, pp. 297–308. Princeton University Press, Princeton, NJ.

SOKAL, R. R. AND ROHLF, F. J. 1995. *Biometry*, 3rd edn. W. H. Freeman, New York.

VAN TIENDEREN, P. H. 1991. Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution*, **45**, 1317–1331.

VIA, S. 1984. The quantitative genetics of polyphagy in an insect herbivore. II. Genetic correlations in larval performance within and among host plants. *Evolution*, **38**, 896–905.

VIA, S. AND LANDE, R. 1985. Genotype–environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**, 505–522.

WIJNGAARDEN, P. J. 2000. *Quantitative and Endocrine Genetics of Reaction Norms for Wing Pattern in the Butterfly* Bicyclus anynana. PhD Thesis, Leiden University.

WIJNGAARDEN, P. J. AND BRAKEFIELD, P. M. 2000. The genetic basis of eyespot size in the butterfly *Bicyclus anynana*: an analysis of line crosses. *Heredity*, **85**, 471–479.

WINDIG, J. J. 1993. *The Genetic Background of Plasticity in Wing Pattern of* Bicyclus *Butterflies*. PhD Thesis, Leiden University.

WINDIG, J. J. 1994a. Reaction norms and the genetic basis of plasticity in the wing pattern of the butterfly *Bicyclus anynana*. *J. Evol. Biol.*, **7**, 665–695.

WINDIG, J. J. 1994b. Genetic correlations and reaction norms in wing pattern of the tropical butterfly *Bicyclus anynana*. *Heredity*, **73**, 459–470.

WINDIG, J. J., BRAKEFIELD, P. M., REITSMA, N. AND WILSON, J. G. M. 1994. Seasonal polyphenism in the wild: survey of wing patterns in five species of *Bicyclus* butterflies in Malawi. *Ecol. Entomol.*, **19**, 285–298.