

# Genetic architecture underlying host choice differentiation in the sympatric host races of *Lochmaea capreae* leaf beetles

Shaghayegh Soudi<sup>1</sup> · Klaus Reinhold<sup>1</sup> · Leif Engqvist<sup>1,2</sup>

Received: 6 August 2015 / Accepted: 25 January 2016 / Published online: 9 February 2016  
© Springer International Publishing Switzerland 2016

**Abstract** Speciation in herbivorous insects has received considerable attention during the last few decades. Much of this group's diversity originates from adaptive population divergence onto different host plants, which often involves the evolution of specialized patterns of host choice behaviour. Differences in host choice often translates directly into divergence in mating sites, and therefore positive assortative mating will be created which will act as a strong barrier to gene flow. In this study, we first explored whether host choice is a genetically determined trait in the sympatric willow and birch host races of the leaf feeding beetle *Lochmaea capreae*, or whether larval experience influences adult host choice. Once we had established that host choice is a genetically based trait we determined its genetic architecture. To achieve this, we employed a reciprocal transplant design in which offspring from pure willow and birch cross-types, F1, F2 and backcrosses were raised on each host plant and their preference was determined upon reaching adulthood. We then applied *joint-scaling* analysis to uncover the genetic architecture of host preference. Our results suggest that rearing host does not have a pronounced effect on adult's host choice; rather the segregation pattern implies the existence of genetic loci affecting host choice in these host races. The joint-scaling analysis revealed that population differences in host choice are mainly influenced by the contribution of additive genetic effects and also maternally inherited cytoplasmic

effects. We explore the implications of our findings for evolutionary dynamics of sympatric host race formation and speciation.

**Keywords** Larval conditioning · Gene action · Joint-scaling analysis · Plant–herbivore interaction · Sympatric speciation

## Introduction

Since Darwin's publication of the origin of species the study of speciation has generated considerable enthusiasm and passionate debate especially regarding radiation in herbivorous insects (Erwin 1982; Thompson 1988; May and Beverton 1990; Sobel et al. 2009). Simpson (1955) claimed that organismal diversification can be partly explained by ecological opportunities such as shifting onto new habitats or resources. Simpson's adaptive zone hypothesis has been recently recast in terms of "ecological speciation", which unifies different modes of speciation proceeding in allopatry, parapatry or in sympatry by the primary action of divergent of natural selection on traits between populations adapted to different resources and leads directly or indirectly to the evolution of reproductive isolation barriers (Schluter 2000; Funk et al. 2006; Feder and Forbes 2007; Peccoud and Simon 2010; Yoder et al. 2010; Funk et al. 2011; Karpinski et al. 2014).

Divergence in sympatry via host shift is perhaps one of the best examples of ecological speciation, as it starts in the absence of extrinsic barriers to gene flow and initial differentiation evolves solely by adaptation to new ecological niches. Host shifts in phytophagous insects are often initiated by evolutionary changes in host preference: when a certain host species is preferred over others as food and oviposition

✉ Shaghayegh Soudi  
Shaghayegh.soudi@uni-bielefeld.de

<sup>1</sup> Evolutionary Biology, Bielefeld University, Morgenbreede 45, 33615 Bielefeld, Germany

<sup>2</sup> Behavioural Ecology, Institute of Ecology and Evolution, University of Bern, 3032 Hinterkappelen, Switzerland

substrate (Futuyma et al. 1993; Gassmann et al. 2006). Behavioural shifts thus often precede and subsequently drive divergent selection on physiological traits related to the processing of different plant secondary metabolites and morphological traits (Weislo 1989; Gassmann et al. 2006). The preferred host is often the place where most of life activities like mating and ovipositing occur, and therefore habitat isolation caused by divergence in host preference is often associated with divergence in mating sites (Via 1999; Malausa et al. 2005; Matsubayashi et al. 2010). This establishes a system of positive assortative mating that can accelerate selection for habitat-specific performance traits. If alternative resources are sufficiently different and compel strong enough fitness trade-offs on populations, a feedback loop can be created favouring increased habitat fidelity. This will act to reduce selection-recombination antagonism (Servedio et al. 2011)—as interbreeding and recombination would randomize gene associations under divergent selection—making it potentially possible for divergence to occur without geographic isolation (Craig et al. 2001; Linn et al. 2003; Feder and Forbes 2007).

Different elements can affect host preferences in herbivorous insects, including genetic inheritance, experience, and the physiological state of insects (Courtney et al. 1989; Craig et al. 2001; Oppenheim et al. 2012). The genetic basis of host preference is fundamental for several basic evolutionary questions such as sympatric speciation (Rice 1987; Dieckmann and Doebeli 1999), host shifts (Groman and Pellmyr 2000) and plant–herbivore co-evolution (Futuyma and Mitter 1996). However, in spite of its substantial role, the genetic basis of host choice is relatively poorly understood in phytophagous insects (Craig et al. 2001; Hawthorne and Via 2001; Berlocher and Feder 2002). The limited number of studies examining the genetic basis of host preferences have demonstrated that these are usually determined by few loci (1–5) (Matsubayashi et al. 2010; Caillaud and Via 2012; Henniges-Janssen et al. 2011; Oppenheim et al. 2012; Karpinski et al. 2014). Those loci were not characterized further, except for two genes coding for odorant binding proteins in *Drosophila melanogaster* (Matsuo et al. 2007). Another finding is that preference genes are often located on autosomes. Only in two butterfly systems Z-linked inheritance was reported (Thompson 1988; Nygren et al. 2006). Furthermore, in about half of the reviewed studies a pattern of dominance in the inheritance of host preferences has been revealed (Forister 2005; Matsuo et al. 2007; Xue et al. 2009). Yet, we still have too few studies to draw clear conclusions regarding the genetic architecture of population differences and its implications for population differentiation and adaptation.

Host preference in herbivorous insects can also be influenced by the effect of larval or early adult experience (Craig et al. 2001; Coyne and Orr 2004; Henniges-Janssen et al. 2011). Hence it is feasible that the reported host fidelity among taxa could be partially explained by larval

conditioning, rather than by heritable preference differences (Coyne and Orr 2004; Dambroski et al. 2005). Huettel and Bush (1972) showed that larval conditioning had an important effect on adults host preference in hybridization studies of two tephritid gall-inducers (Johnson et al. 1996; Craig et al. 2001). Although conditioning may potentially accelerate initial host shifts, it cannot lead to sympatric speciation, which is established as the genetically based divergence in host preference between populations (Craig et al. 2001). Detailed analysis of F2 and backcrosses are fundamental to confirm the inheritance of host choice and its genetic organization.

*Lochmaea capreae* is a chrysomelid beetle that mainly feeds on willow (*Salix capreae*, Salicaceae), birch (*Betula pendula*, Betulaceae) and occasionally on poplar (*Populus* sp. Salicaceae). This species exhibits sympatric host races (sympatric populations that are incompletely reproductively isolated but remain ecologically differentiated in the face of gene flow due to divergent selection on populations using alternative hosts, (Funk 2012)) on willow and birch with about 2 % gene flow between them. Previous studies have demonstrated that populations are highly sympatric in Western Russia, near Lake Baikal and further east (Kreslavsky and Mikheev 1993); yet sympatric populations also seem to extend westward at least to our study populations in the western parts of Germany (K.R. pers. obs.). Reciprocal transplant experiments have previously revealed that each race is locally adapted on each native host plant and has achieved different adaptive sets of preference and performance traits through close association with its own host plant (Soudi et al. 2015). The pronounced host-associated ecological divergence and local adaptation is accompanied by multiple forms of premating isolation barriers, specifically habitat isolation and immigrant inviability. Furthermore, crossing experiments have revealed signs of substantial hybrid inviability via both intrinsic and extrinsic post-zygotic isolation (Soudi et al. 2016) indicating that in our populations, genetic divergence is already at a quite advanced stage.

Host choice behaviour is of special importance to *L. capreae* for two main reasons. First, these sympatric host races can be a good candidate for studying speciation in the absence of geographic barriers to gene flow (sympatric speciation) via host plant shift. These races are cited as one among few examples of ongoing host races formation (the hypothesized initial stage of sympatric speciation (Drès and Mallet 2002), besides the well-studied cases such as apple maggot fly (Bush 1969a, b), larch budmoth (Emelianov et al. 1995), ecotypes of *Neochlamisus bebbianae* leaf beetles (Funk 1998) and *Timema* walking sticks (Nosil 2007). Second, assortative mating is relatively weak when beetles from different populations are confined together in the lab (Soudi et al. in prep.), but is greatly enhanced in the

field and in the presence of hosts (Kreslavsky and Mikheev 1993). Hence, differences in host choice can translate into mate choice and prezygotic reproductive isolation barriers between sympatric beetle populations. Investigating the genetic architecture of differences between these host races therefore has important consequences for identifying and understanding how ecological specialization and population divergence can occur in the face of gene flow.

We herein investigate the genetic architecture of host choice between the host races of *L. capreae* to elucidate the relative contribution of additive and non-additive genetic effects to the differences in host choice between these host races. We examined whether host choice is a genetically based trait and/or whether larval experience affects it. To achieve our goals, we applied a joint-scaling analysis (Lynch and Walsh 1998) in which we analysed the host choice behaviour of pure parental cross-types, hybrids and backcrosses between two host races of *L. capreae* reared on willow and birch.

## Materials and methods

### Natural history of *L. capreae* host races

*Lochmaea capreae* is a chrysomelid beetle mainly feeding on willow and birch (and occasionally on poplar). In the field, it has only one generation per year. Adults emerge from winter diapause in May and begin feeding and mating. Females move among host plants in May and June, laying eggs at the base of twigs or below the surface of the soil near the host plants. Newly hatched larvae migrate on to the leaves to feed and later pupate in the soil. Adult beetles of the new generation appear in August and feed until entering diapause in late summer.

Our previous study demonstrated that willow and birch races are ecologically divergent, females from the willow and birch host races differ considerably in their oviposition preferences (Soudi et al. 2015). However, these preferences seem to be asymmetrical, in that there are more distinct preferences within the willow race compared to the birch race. This study also revealed that these races differ significantly in their food preferences and each host race exhibited very clear preferences for their natal host plant. Willow host race had a complete disposition to feed on willow and not a single individual accepted to feed on birch; thus all willow offspring died on birch. Some individuals from the birch race accepted to feed on willow but these individuals consumed less than those on their native host plant, in addition they displayed lower fitness when they were reared on willow than on birch (Soudi et al. 2015). Previous work by Kreslavsky and Mikheev (1993) have suggested that some of these physiological differences between the willow and birch races of *L. capreae* may be

associated with a monogenic recessive mutation restricting individuals from the willow race to develop on birch.

### Rearing and experimental crosses

In May 2013, soon after the emergence of beetles from their overwintering sites, we collected overwintered adults of *L. capreae* on birch and willow on two locations in Kottenforst (50.715°N, 7.002°E; 50.671°N, 7.009°E) near Bonn, Germany. Field collected females of each population were transferred to the laboratory (Bielefeld University, Germany) and kept individually in single petridishes (Ø 9 cm) and were fed with leaves of their natural host plants to lay eggs. To obtain foliage for use throughout the experiments, willow (*Salix caprea*) and birch (*Betula pendula*) were collected in the vicinity of the university. Oviposition was checked daily and eggs were transferred into clean petri dishes lined with moistened filter paper. Abiotic conditions were maintained at 23–25 °C, 18L:6D light cycle, and 70 % relative humidity. Larvae from these eggs were reared to adulthood on cutting of their native host plants, and from these adults, F1-hybrids and conspecific crosses were made, to generate a second generation. Not one single egg hatched from one of the heterospecific crosses (birch females with willow males), and therefore only one F1-hybrid type was obtained from crossing willow females with birch males. F1-hybrids from the second generation was subsequently crossed with other F1-hybrids and backcrossed with each of the pure parental cross-types (in both male–female combinations) to generate a third generation used in the experiments. In total we created 8 crosses over three generations: two pure parental cross-types (WW = WW ♀ × WW ♂ and BB = BB ♀ × BB ♂), one F1-hybrid (WB = WW ♀ × BB ♂), one F2-hybrid (F2 = WB ♀ × WB ♂), two birch backcrosses (BC1b = BB ♀ × WB ♂, BC2b = WB ♀ × BB ♂) and two willow backcrosses (BC1w = WW ♀ × WB ♂, BC2w = WB ♀ × WW ♂). The reciprocal crosses allowed us to test for the presence of different genetic effects on line cross means. In order to test whether rearing host has any effect on host recognition, we established two independent sets of crosses. In one set, all larvae from different crosses described above were reared on willow, in the second set, all larvae were reared on birch. These produced 130 families (13–15 families per cross-type) and 648 offspring (401 females and 247 males) that were used to test the effects of rearing hosts on the relative importance of different genetic effects.

### Assays of host choice

Host choice was evaluated through laboratory choice experiments. We used individuals that developed until adulthood on either host plant used for the host preference

assay. Once adults enclosed, each individual was placed into a clean petridish ( $\varnothing$  8 cm) lined with moistened filter paper. Two rectangular leaf discs, one from willow and the other from birch, were simultaneously offered to each beetle and they were allowed to choose and feed freely on both host plants for 48 h. Initially, we aimed at analysing the difference in amount eaten from birch and willow leaves, respectively. However, almost all individuals (more than 96 %) chose exclusively from one of the host plants only. We therefore regarded the chosen host plant as the one each particular individual had fed most from and then compared the different cross-types regarding the proportion of individuals that accepted willow or birch.

### Statistical analysis

Since we observed that nearly all test individuals almost always chose one of host over the other for feeding, we analysed data with generalized linear mixed effects model (GLMMs) using logit as a link factor (beetles showed preference for willow being scored as 1, and those preferred birch being scored as 0). The full model for the present experiment was analysed using cross-type, rearing host (diet) and sex as fixed effects. To detect whether the effect of diet on host choice is dependent on cross-types, and also whether the effect of diet on host choice differs between males and females, we added cross-type  $\times$  diet and sex  $\times$  diet interactions as further fixed effects in the model. Furthermore, family was entered as a random effect. All models including random effect were fit using restricted maximum likelihood. Subsequently, we performed post hoc tests to examine differences among cross-types. All post hoc comparisons were conducted using a Tukey HSD test, which accounts for multiple comparisons. To assess whether lines had a significant preference for either of the host plants, we tested whether the estimated logit value for each line was significantly different from zero (corresponding to a 50:50 preference).

### Joint-scaling analysis

We estimated the additive, dominance, epistatic, cytoplasmic and sex chromosome effects using a joint-scaling analysis as described in Lynch and Walsh (1998). Joint-scaling is a quantitative genetic method by which the net additive and non-additive genetic effects can be inferred from phenotypic differences among lines of known pedigree. We followed the method described by Bieri and Kawecki (2003). Briefly, we optimized a linear model describing the composite genetic effects contributing to our eight line means. We used the expected phenotype of F2-hybrid offspring as the point of reference; hence, its mean preference defines the intercept

parameter ( $m$ ). The composite genetic effects represent the mean and the effect of additive genetic effect ( $a$ ), dominance ( $d$ ), ( $aa$ ), ( $ad$ ), and ( $dd$ ) describing additive–additive, additive–dominance, and dominance–dominance epistatic interactions, respectively, X-linked additive ( $xa$ ) and X-linked dominance ( $xd$ ) and cytoplasmic effect ( $c$ ). The analysis involved assigning the coefficients of determination to each cross-type that describe the expected contribution from each genetic effect (Demuth and Wade 2007). The coefficients of determination are listed in Table 1. Our full model that was fitted to the data to scrutinize genetic divergence in host preference was thus as follow:

$$\bar{Y}_i = m + x_a(a) + x_d(d) + x_{aa}(aa) + x_{ad}(ad) + x_{dd}(dd) + x_{xa}(xa) + x_{xd}(xd) + x_c(c) + e_i,$$

where  $\bar{Y}_i$  is the mean of the  $i$ th genetic class (parental species or hybrids),  $m$  is the intercept parameter, ( $a$ ), ( $d$ ), ( $aa$ ), ( $ad$ ), ( $dd$ ), ( $xa$ ), ( $xd$ ) and ( $c$ ) are the various genetic composite effects,  $x_a$ ,  $x_d$ ,  $x_{aa}$ ,  $x_{ad}$ ,  $x_{dd}$ ,  $x_{xa}$ ,  $x_{xd}$ ,  $x_c$  are the regression coefficients listed in Table 1, and  $e_i$  is the residual error.

Line cross techniques are commonly used to sequentially add parameters into the model (first additive effects are added, then dominance, epistasis, maternal effects and cytoplasmic effects) until predicted and observed means do not significantly differ (Mather and Jinks 1971). Yet, the order in which parameters are added into the model affects the ability to detect significant parameters that are added later. For this reason we followed the suggestion of Bieri and Kawecki (2003) and used Akaike's Information Criterion (AIC) to select the most parsimonious model, which is the best compromise between the amount of variance explained and the number of parameters in the model. Nevertheless, to decrease the risk of overfitting due to a massive number of candidate models, we treated the three epistasis parameters terms ( $aa$ ), ( $ad$ ), and ( $dd$ ) as one group ( $e$ ), as suggested by Bieri and Kawecki (2003). These terms were thus either all included or excluded from the model. In this way the number of candidate models was reduced to  $2^6 = 64$ . The significance of each parameter included in the most parsimonious model was tested with a likelihood-ratio test. The likelihood ratio is Chi square distributed with degrees of freedom equal to the difference in the number of parameters in the two models (Lynch and Walsh 1998; Bieri and Kawecki 2003).

All statistical analyses were conducted using the statistical software R 3.0.3 (Team 2012). The generalized linear models were performed using the command *lmer* in the *lme4* package (Bates et al. 2013), and the AIC model selection was done using the function *dredge* provided in the *MuMIn*-package (Barton 2014). All tests were two-tailed and the null hypotheses were rejected at  $P < 0.05$ .

**Table 1** The regression coefficients used for the eight candidate parameters assessing composite genetic effects

Line	Cross (♀ × ♂)		$x_a$	$x_d$	$x_{aa}$	$x_{ad}$	$x_{dd}$	$x_{xa}$	$x_{xd}$	$x_c$
P1	Willow pure race	Daughters	1	0	1	0	0	1	1	1
		Sons	1	0	1	0	0	1	1	1
P2	Birch pure race	Daughters	-1	0	1	0	0	-1	-1	-1
		Sons	-1	0	1	0	0	-1	-1	-1
F1	P1 × P2	Daughters	0	1	0	0	1	0	1	1
		Sons	0	1	0	0	1	-1	1	1
F2	F1 × F1	Daughters	0	0.5	0	0	0.25	-0.5	1	1
		Sons	0	0.5	0	0	0.25	0	1	1
BC1w	P1 × (P1 × P2)	Daughters	0.5	0.5	0.25	0.25	0.25	0.5	0.5	1
		Sons	0.5	0.5	0.25	0.25	0.25	0	0	1
BC2w	(P1 × P2) × P1	Daughters	0.5	0.5	0.25	0.25	0.25	0	1	1
		Sons	0.5	0.5	0.25	0.25	0.25	1	0	1
BC1b	P2 × (P1 × P2)	Daughters	-0.5	0.5	0.25	-0.25	0.25	-0.5	0.5	-1
		Sons	-0.5	0.5	0.25	-0.25	0.25	0	0	-1
BC2b	(P1 × P2) × P2	Daughters	-0.5	0.5	0.25	-0.25	0.25	0	1	1
		Sons	-0.5	0.5	0.25	-0.25	0.25	-1	0	1

$x_a$  additive,  $x_d$  dominance,  $x_{aa}$  additive–additive epistasis,  $x_{ad}$  additive–dominance epistasis,  $x_{dd}$  dominance–dominance epistasis,  $x_{xa}$  X-linked additive,  $x_{xd}$  X-linked dominance and  $x_c$  cytoplasmic effect (Table modified from Demuth and Wade 2007)

**Table 2** Generalized mixed effect model analysis to compare variation in host preference between different cross-types of *L. capreae*

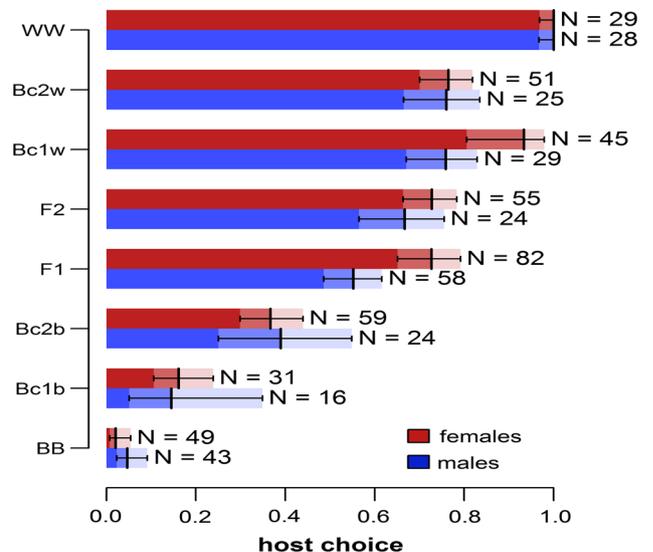
Source of variation	df	$\chi^2$	P
<i>Host preference</i>			
Host environment (diet)	1	0.65	0.42
Cross-type	7	146.1	<0.001
Sex	1	1.45	0.22
Diet × cross-type	6	8.18	0.22
Diet × sex	1	0.85	0.36

**Results**

**Feeding preference of parents and crosses**

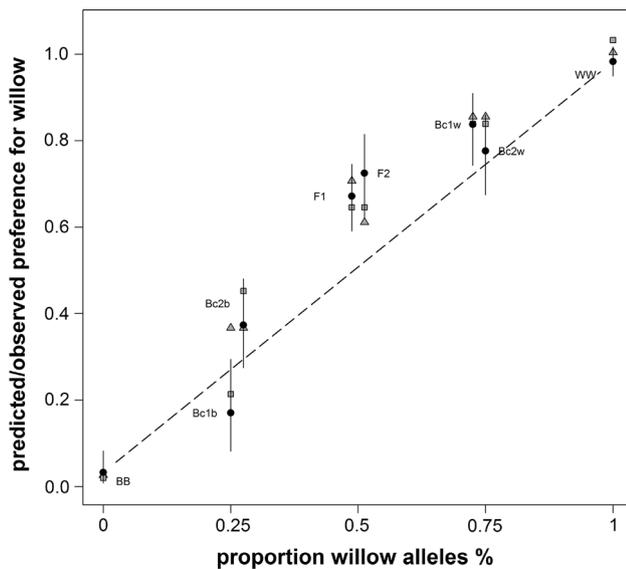
In total, host choice (in the dual choice test) was scored for 650 individuals, 401 females and 247 males that developed on either host plant until adulthood. Mixed effects model analyses revealed that host choice was overwhelmingly affected by cross-type, but not the rearing host (diet), sex and the interaction between them (Table 2; Fig. 1). Individuals from both pure willow and pure birch cross-types exhibited a very strong host preference for their native host plant regardless of the host plant that they had been reared during larval development.

A Tukey post hoc comparison showed that the host choice of the cross-types fell roughly into four homogenous groups based on the proportion of willow genes in different cross-types (Fig. 2). The first group contained only willow



**Fig. 1** Mean (+SE) host choice of willow and birch races of *L. capreae* and their hybrids separated by sex. Bars indicate the proportion of males and females from each cross-type that chose willow for feeding. Sample sizes are shown on the top of bars. BB represents pure birch parental type, WW presents pure willow parental type, Bc1w and Bc2w represent two willow-like backcrosses, and Bcb1 and Bcb2 represent two birch-like backcrosses

adults which exhibited a complete disposition to choose and feed on their native host plant and not a single one of the adults from 15 families chose to feed on birch. The second group consisted of the willow backcrosses, as well as the F1- and F2-hybrids, which all had a significant



**Fig. 2** Host choice of willow and birch races of *L. capreae* and hybrids and backcrosses between them (without separating males and females). Points and bars indicate mean  $\pm$  SE. The line indicates the expected hybrid means if the inheritance of population differs in host choice was purely additive. Grey squares shows predicted means for the most parsimonious model including additive and cytoplasmic effects. Grey triangles shows predicted means for the most intuitive model including additive and dominance effects. The former model is considerably more likely than the first given the data ( $\Delta$ AIC = 6.42, evidence ratio: ca. 25)

preference for willow (see Fig. 2; GLMM: F1:  $Z = 2.99$ ,  $P = 0.003$ ; F2:  $Z = 3.40$ ,  $P < 0.001$ ; BC1w:  $Z = 4.59$ ,  $P < 0.001$ ; BC2w:  $Z = 4.13$ ,  $P < 0.001$ ). The third group contained the birch backcrosses which avoided willow and chose to feed on birch instead (Fig. 2; GLMM: BC1b:  $Z = -1.98$ ,  $P = 0.048$ ; BC2b:  $Z = -3.66$ ,  $P < 0.001$ ). Pure birch race was categorized as the last group with the strongest avoidance against willow (Fig. 2; GLMM:  $Z = -5.61$ ,  $P < 0.001$ ) where only 2 % chose willow for feeding in the dual host-choice assays. Yet, the birch race was not completely distinct from the birch backcrosses as the Tukey comparisons revealed that its host acceptance was only significantly different from BC2b ( $Z = 4.44$ ,  $P < 0.001$ ) but not from BC1b ( $Z = 2.41$ ,  $P = 0.19$ ).

### Composite genetic effects and patterns of inheritance

The AIC most parsimonious model for genetic difference in host choice between the willow and birch host races of *L. capreae* included only additive (a) and maternally inherited cytoplasmic effects (c), regardless of the rearing host (Fig. 2; Tables 3, 4). No further contributions of dominance, epistatic and sex-chromosome effects to the difference between the lines were detected. The strong

**Table 3** Composite genetic effects contributing to differences in food preference between the host races of *L. capreae*

Parameters	df	$\chi^2$	P
<i>Included terms</i>			
Additive (a)	1	62.9	<0.001
cytoplasmic (c)	1	13.7	<0.001
<i>Excluded terms</i>			
Dominance (d)	1	0.05	0.81
Epistasis (aa, ad, dd)	3	0.44	0.93
Sex-chromosome (xa, xd)	2	2.00	0.37

The values represent the composite genetic effects from the Akaike's information criterion (AIC) most parsimonious genetic model

contribution of additive effects is obvious from the pronounced higher preference for willow in cross-types with higher proportion of willow alleles. Host choice was also influenced by cytoplasmic effects—birch backcross with F1-hybrid mother and birch father had higher preference for willow than the other birch backcross with birch mother and F1-hybrid father (Fig. 2).

### Discussion

Understanding how herbivorous insects evolve to traverse resource boundaries and specialize on new host plants is critical for determining the sources of biodiversity on earth (Dambroski et al. 2005). Although much progress has been made in recent years, the genetics of host choice in phytophagous insects is relatively poorly understood (Berlocher and Feder 2002; Drès and Mallet 2002). The aim of the current study was to investigate the environmental and genetic basis of host choice in the ecological host races of *L. capreae* specialized on willow and birch. Investigating the genetic basis of host choice in these specialized host races can have profound ramifications for understanding how ecological specialization and population divergence can rapidly occur in the face of gene flow.

The influence of alleles on phenotypes depends on the environment in which those alleles are expressed (Falconer et al. 1996). Environmental effects on genetic architecture are of substantial importance when examining traits that have evolved due to environment specific selection. Host selection in herbivorous insects involves several steps, including habitat location, host location, host recognition and host acceptance. All these phases may be under the control of both genetic factors and also environmental conditioning (Jaenike 1990; Desjardins et al. 2010). Our results suggest that larval conditioning does not have any pronounced effect on the subsequent behavioural response of adult beetles to willow and birch. The non-significant

**Table 4** Five candidate models that made the best compromise between the amount of variance explained and the number of parameters (most parsimonious models) selected by Akaike's information criterion (AIC)

Model parameters													
<i>m</i>	<i>a</i>	<i>d</i>	<i>aa</i>	<i>ad</i>	<i>dd</i>	<i>xa</i>	<i>xd</i>	<i>c</i>	<i>df</i>	log(L)	AICc	$\Delta_i$	$w_i$
0.510	0.372	–	–	–	–	–	–	0.119	5	–3.069	16.5	0.00	0.753
0.500	0.384	–	–	–	–	–	0.057	0.109	6	–4.385	21.2	4.76	0.070
0.505	0.377	0.017	–	–	–	–	–	0.113	6	–4.793	22.0	5.58	0.046
0.582	0.484	–	–	–	–	–	–	–	4	–7.242	22.7	6.24	0.034
0.512	0.394	–	–	–	–	–0.026	–	0.123	6	–5.125	22.7	6.24	0.034
0.518	0.478	0.139	–	–	–	–	–	–	5	–4.793	22.0	5.58	0.031

Model parameters including *m* represents the intercept parameter, *a*, *d* represent additive and dominance genetic effects respectively, *aa*, *ad*, *dd* represent additive–additive, additive–dominance and dominance–dominance epistasis effects respectively, *xa* and *xd* represent X-linked additive and X-linked dominance effects and *c* represents a cytoplasmic effect

host environment effect in our analysis rejects the hypothesis of larval experience which emphasizes that individuals from each cross-type would choose the host plant on which they have already experienced. Rather, the consistent pattern of host recognition implies the existence of discrete genetic loci affecting host choice and avoidance which satisfies the assumption of speciation in the face of gene flow. A large role for genetic factors has been indicated in the majority of studies that have investigated the genetic basis of host choice in herbivorous insects like sympatric host races of pea aphid, *Acyrtosiphon pisum* (Hawthorne and Via 2001), gall flies, *Eurosta solidagenis* (Craig et al. 2001), apple and hawthorn races of *Rhagoletis* flies (Linn et al. 2003), and *Altica* flea beetle (Xue et al. 2009). However future studies are required to make a clear inference about the influence of genetic and non-genetic factors on host preference in herbivorous insects (Matsubayashi et al. 2010).

The host choice pattern was not explained by sex-linked effects (Table 3), and this indicates that the host choice genes are probably located on autosomes in the host races of *L. capreae*. This further supported by the joint-scaling analysis which did not detect any contribution of sex-linked effect to divergence in host choice between these races. According to the percentage of willow-derived autosomal genes, the two parental species and six hybrid crosses can be classified into five groups: pure-willow race with 100 % of willow autosomal loci, willow backcrosses with 25 % difference from pure willow in the origin of autosomal loci; F1 and F2 hybrids with 50 % difference; birch backcrosses with 75 % difference; and pure-birch line with 100 % difference. These groups varied considerably in host choice; on the other hand host choice of crosses belonging to the same group did not differ significantly. With the percentage of willow-derived autosomal genes decreasing from pure-willow cross-type to F1 and F2 generations, backcross generations, and to pure-birch cross-type, the mean proportional of those individuals that accepted willow for

feeding decreased accordingly. Overall our results are therefore consistent with most other studies showing that preference genes are often located on autosomes (Matsubayashi et al. 2010).

Besides indicating that the difference in food preferences between the host races of *L. capreae* is likely to be determined by autosomal loci, our results also propose that these loci do not act completely in an additive manner. On first glance, the observed pattern seems to indicate slight genetic dominance. However, the observed deviation from additive genetic effects can most parsimoniously be explained by cytoplasmic effects inherited through the maternal line. The absence of dominance and epistatic effect was surprising, because their effects on phenotypic variation in behavioural traits are expected to be more influential than additive effects (Meffert et al. 2002; Mackay 2009), and may serve as a basic point for adaptive evolutionary divergence. Instead, the observed bias for all of the backcrosses to choose the natal host plant of their maternal grandmother (birch for Bc1b, and willow for the rest), strongly implies the existence of cytonuclear gene interactions, which is detected as a strong cytoplasmic effect. Cytoplasmically inherited effects are not uncommon between populations where some degree of differentiation exists between forms (Hutter and Rand 1995; Burton et al. 1999; Cruzan and Arnold 1999; Rand et al. 2001; Sackton et al. 2003; Dambroski et al. 2005). Such a pattern of maternal inheritance increases the resemblance between offspring and mother compared to offspring and father, and therefore may appear statistically along with other true maternal effects in quantitative genetics or other analysis (Wolf and Wade 2009). However, the mechanism responsible for it effect remains to be determined but must involve a means for host recognition to be inherited through the maternal lineage for multiple generations. Genomic imprinting due to retention of DNA methylation patterns in maternally inherited chromosomes, cytoplasmically inherited factors (e.g., microorganisms or maternal gene products), and mtDNA-nuclear gene interactions in which

specific maternally inherited preference and avoidance genes are up-regulated in certain mtDNA (cytoplasmic) genetic backgrounds could contribute to the cytoplasmic maternally inherited effects (Willett and Burton 2004; Dambroski et al. 2005; for review see Wolf and Wade 2009).

In herbivorous insects divergence in host choice and host recognition can act as habitat isolation by limiting inter-host migration, and if mating sites are restricted to on or very close vicinity of the preferred host plant, a system of positive assortative mating will be created even without any active assessment of potential mates (Rundle and Nosil 2005; Matsubayashi et al. 2010). In our study each host race exhibited a very strong preference for its native host plant; this pattern of host choice indicates that gene flow would be primarily reduced between these sympatric host races as it will reduce heterospecific encounters and subsequently heterospecific matings. However, if some individuals migrate between host plants complete reproductive isolation between the host races would be expected only if there is direct selection against migrants before mating or if there is selection against hybrid progeny. In the laboratory we have observed that assortative mating is relatively weak when beetles are confined together (Soudi et al. in revision), and it seems that mate choice experiments provide no support for the idea that individuals from migrant lineages are discriminated against as mates if they are encountered. The resulting F1-hybrid would have higher preference for willow, so they would be expected to mate primarily with pure willow race. These backcrosses to the willow race do survive on willow indicating that introgression of birch alleles into the willow population is possible. Introgression from the willow race to birch race is also possible, but to a lesser degree, as hybrids tend to dislike birch. Nevertheless, male F1-hybrids had relatively intermediate preferences, therefore some of them may move on birch and could also backcross to birch race. Birch backcrosses exhibited a high survival rate (Soudi et al. under review) which would facilitate gene flow also in this direction.

In conclusion, investigating the genetics of host choice behaviour in herbivorous insects is fundamental for understanding the origin of biodiversity and adaptive radiation on a grand scale, since there are more host specific phytophagous insects than any other organism on earth (Dambroski et al. 2005). Our results provide insights into the genetics of host choice for *L. capreae*, a key trait involved in host specialization that generates reproductive isolation in a promising model system for sympatric speciation. Our results indicate that host preference in the willow and birch races of *L. capreae* is a genetically based trait and preference genes are probably located on autosomes. Yet the mode of gene action on host preference could not be explained only by the additive gene effect; instead a model including additive and cytoplasmic effects

was required to explain divergence in host preference. One future promising approach is the development of genomic maps and the search for linkage groups that are associated with food preference phenotypes (Caillaud and Via 2012).

**Acknowledgments** We thank two anonymous reviewers for their helpful comments on previous versions.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no competing interests.

## References

- Barton K (2014) Package ‘MuMIn’: multi-model inference. R package version 1.9. 13
- Bates D, Maechler M, Bolker B, Walker S (2013) lme4: linear mixed-effects models using Eigen and S4. R package version 1.0-5
- Berlacher SH, Feder JL (2002) Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annu Rev Entomol* 47:773–815
- Bieri J, Kawecki TJ (2003) Genetic architecture of differences between populations of cowpea weevil (*Callosobruchus maculatus*) evolved in the same environment. *Evolution* 57:274–287
- Burton RS, Rawson PD, Edmands S (1999) Genetic architecture of physiological phenotypes: empirical evidence for coadapted gene complexes. *Am Zool* 39:451–462
- Bush GL (1969a) Mating behavior, host specificity, and the ecological significance of sibling species in frugivorous flies of the genus *Rhagoletis* (Diptera-Tephritidae). *Am Nat* 103:669–672
- Bush GL (1969b) Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution* 23:237–251
- Caillaud M, Via S (2012) Quantitative genetics of feeding behavior in two ecological races of the pea aphid, *Acyrtosiphon pisum*. *Heredity* 108:211–218
- Courtney SP, Chen GK, Gardner A (1989) A general model for individual host selection. *Oikos* 55:55–65
- Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland
- Craig TP, Horner JD, Itami JK (2001) Genetics, experience, and host plant preference in *Eurosta solidaginis*: implications for host shift and speciation. *Evolution* 55:773–782
- Cruzan MB, Arnold ML (1999) Consequences of cytonuclear epistasis and assortative mating for the genetic structure of hybrid populations. *Heredity* 82:36–45
- Dambroski HR, Linn C, Berlacher SH, Forbes AA, Roelofs W, Feder JL (2005) The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution* 59:1953–1964
- Demuth JP, Wade MJ (2007) Population differentiation in the beetle *Tribolium castaneum*. I. Genetic architecture. *Evolution* 61:494–509
- Desjardins CA, Perfectti F, Bartos JD, Enders LS, Werren JH (2010) The genetic basis of interspecies host preference differences in the model parasitoid *Nasonia*. *Heredity* 104:270–277
- Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature* 400:354–357
- Drès M, Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Philos Trans R Soc B* 357:471–492
- Emelianov I, Mallet J, Baltensweiler W (1995) Genetic differentiation in *Zeiraphera diniana* (Lepidoptera: Tortricidae, the larch budmoth): polymorphism, host races or. *Heredity* 75:16–424

- Erwin TL (1982) Tropical forests: their richness in Coleoptera and other arthropod species. *Coleop Bull* 36:74–75
- Falconer DS, Mackay TF, Frankham R (1996) Introduction to quantitative genetics (4th edn). *Trends Genet* 12:280
- Feder JL, Forbes AA (2007) Habitat avoidance and speciation for phytophagous insect specialists. *Funct Ecol* 21:585–597
- Forister ML (2005) Independent inheritance of preference and performance in hybrids between host races of *Mitoura* butterflies (Lepidoptera: Lycaenidae). *Evolution* 59:1149–1155
- Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution* 52:1744–1759
- Funk DJ (2012) Of “host forms” and host races: terminological issues in ecological speciation. *Int J Ecol* 2012:1–8
- Funk DJ, Nosil P, Etges WJ (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proc Natl Acad Sci USA* 103:3209–3213
- Funk DJ, Egan SP, Nosil P (2011) Isolation by adaptation in *Neochlamisus* leaf beetles: host-related selection promotes neutral genomic divergence. *Mol Ecol* 20:4671–4682
- Futuyma DJ, Mitter C (1996) Insect-plant interactions: the evolution of component communities. *Philos Trans R Soc B* 351:1361–1366
- Futuyma DJ, Keese MC, Scheffer SJ (1993) Genetic constraints and the phylogeny of insect-plant associations: responses of *Ophraella communa* (Coleoptera: Chrysomelidae) to host plants of its congeners. *Evolution* 47:888–905
- Gassmann A, Levy A, Tran T, Futuyma D (2006) Adaptations of an insect to a novel host plant: a phylogenetic approach. *Funct Ecol* 20:478–485
- Groman J, Pellmyr O (2000) Rapid evolution and specialization following host colonization in a *yucca* moth. *J Evol Biol* 13:223–236
- Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412:904–907
- Henniges-Janssen K, Reineke A, Heckel DG, Groot AT (2011) Complex inheritance of larval adaptation in *Plutella xylostella* to a novel host plant. *Heredity* 107:421–432
- Huettel MD, Bush GL (1972) The genetics of host selection and its bearing on sympatric speciation in *Procecidochares* (Diptera: Tephritidae). *Entomol Exp Appl* 15:465–480
- Hutter CM, Rand DM (1995) Competition between mitochondrial haplotypes in distinct nuclear genetic environments: *Drosophila pseudoobscura* vs *D. persimilis*. *Genetics* 140:537–548
- Jaenike J (1990) Host specialization in phytophagous insects. *Annu Rev Ecol Syst* 21:243–273
- Johnson PA, Hoppensteadt F, Smith JJ, Bush GL (1996) Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. *Evol Ecol* 10:187–205
- Karpinski A, Haenniger S, Schöfl G, Heckel DG, Groot AT (2014) Host plant specialization in the generalist moth *Heliothis virescens* and the role of egg imprinting. *Evol Ecol* 28:1075–1093
- Kreslavsky A, Mikheev A (1993) Gene geography of racial differences in *Lochmaea capreae* L (Coleoptera; Chrysomelidae), and the problem of sympatric speciation. *Entomol Rev* 72:50–58
- Linn C, Feder JL, Nojima S, Dambroski HR, Berlocher SH, Roelofs W (2003) Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proc Natl Acad Sci USA* 100:11490–11493
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer, Sunderland
- Mackay TF (2009) The genetic architecture of complex behaviors: lessons from *Drosophila*. *Genetica* 136:295–302
- Malausa T, Bethenod M-T, Bontemps A, Bourguet D, Cornuet J-M, Ponsard S (2005) Assortative mating in sympatric host races of the European corn borer. *Science* 308:258–260
- Mather K, Jinks JL (1971) Biometrical genetics. In: The study of continuous variation. Cambridge University Press, Cambridge
- Matsubayashi KW, Ohshima I, Nosil P (2010) Ecological speciation in phytophagous insects. *Entomol Exp Appl* 134:1–27
- Matsuo T, Sugaya S, Yasukawa J, Aigaki T, Fuyama Y (2007) Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. *PLoS Biol* 5:e118
- May RM, Beverton R (1990) How many species?[and discussion]. *Philos Trans R Soc B* 330:293–304
- Meffert LM, Hicks SK, Regan JL (2002) Nonadditive genetic effects in animal behavior. *Am Nat* 160:S198–S213
- Nosil P (2007) Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. *Am Nat* 169:151–162
- Nyren G, Nylin S, Stefanescu C (2006) Genetics of host plant use and life history in the comma butterfly across Europe: varying modes of inheritance as a potential reproductive barrier. *J Evol Biol* 19:1882–1893
- Oppenheim SJ, Gould F, Hopper KR (2012) The genetic architecture of a complex ecological trait: host plant use in the specialist moth, *Heliothis subflexa*. *Evolution* 66:3336–3351
- Peccoud J, Simon J (2010) The pea aphid complex as a model of ecological speciation. *Ecol Entomol* 35:119–130
- Rand DM, Clark AG, Kann LM (2001) Sexually antagonistic cytonuclear fitness interactions in *Drosophila melanogaster*. *Genetics* 159:173–187
- Rice WR (1987) Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evol Ecol* 1:301–314
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecol Lett* 8:336–352
- Sackton TB, Haney RA, Rand DM (2003) Cytonuclear coadaptation in *Drosophila*: disruption of cytochrome c oxidase activity in backcross genotypes. *Evolution* 57:2315–2325
- Schluter D (2000) The ecology of adaptive radiation. Oxford University Press, Oxford
- Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation: ‘magic’ but not rare? *Trends Ecol Evol* 26:389–397
- Simpson GG (1955) Major features of evolution. Columbia University Press, New York
- Sobel MS, Grace FC, Watt LR, Schemske DW (2009) The biology of speciation. *Evolution* 64:295–315
- Soudi S, Reinhold K, Engqvist L (2015) Host associated divergence in sympatric host races of the leaf beetle *Lochmaea capreae*: implications for local adaptation and reproductive isolation. *Biol J Linn Soc* 116:169–182
- Soudi S, Reinhold K, Engqvist L (2016) Ecologically dependent and intrinsic genetic signatures of postzygotic isolation between sympatric host races of the leaf beetle *Lochmaea capreae*. *Evolution*. doi:10.1111/evo.12846
- R Development Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria
- Thompson JN (1988) Coevolution and alternative hypotheses on insect/plant interactions. *Ecology* 69:893–895
- Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution* 53:1446–1457

- Wcislo WT (1989) Behavioral environments and evolutionary change. *Annu Rev Ecol Syst* 20:137–169
- Willett CS, Burton RS (2004) Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. *Mol Biol Evol* 21:443–453
- Wolf JB, Wade MJ (2009) What are maternal effects (and what are they not)? *Philos Trans R Soc B* 364:1107–1115
- Xue H-J, Magalhães S, Li W-Z, Yang X-K (2009) Reproductive barriers between two sympatric beetle species specialized on different host plants. *J Evol Biol* 22:2258–2266
- Yoder J, Clancey E, Des Roches S, Eastman J, Gentry L, Godsoe W, Hagey T, Jochimsen D, Oswald B, Robertson J (2010) Ecological opportunity and the origin of adaptive radiations. *J Evol Biol* 23:1581–1596