Ecological Entomology (2017), DOI: 10.1111/een.12380

Hybridisation and lack of prezygotic barriers between *Phymata pennsylvanica and americana*

DAVID PUNZALAN[1,2] and LOCKE ROWE[2] [1]Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada and [2]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

Abstract. 1. The persistence of both geographical and reproductive boundaries between related species poses a fundamental puzzle in biology. Reproductive interactions between species can have a substantial impact on the maintenance of a boundary, potentially contributing to its collapse (e.g. via hybridisation) or facilitating reproductive isolation (e.g. via reinforcement).

2. The degree to which two parapatric insect species in the genus *Phymata* are reproductively isolated was evaluated and several mechanisms that could contribute to the maintenance of species boundaries were assessed.

3. Behavioural assays showed no indication of species-assortative mating, nor any fecundity costs associated with heterospecific mating. Thus, there was no evidence of prezygotic mechanisms of reproductive isolation between the two species.

4.In laboratory crosses, it was found that the two species were indeed capable of producing viable F1 hybrids. Morphologically, these hybrids were phenotypically intermediate to the two parental species, and similar to the phenotypes seen in natural populations thought to occur in a hybrid zone. F1 hybrids did not show reduced viability, although there was some suggestion of 'hybrid breakdown', evident from the lower viability observed for progeny of 'natural hybrids'.

5. Collectively, we show that despite genetically based morphological differences between species, *P. americana* and *pennsylvanica* can, and probably do hybridise. More studies are needed to understand the mechanisms that maintain the distinct phenotypes and geographical ranges of these species, despite the considerable potential for introgression.

Key words. Introgression, Phymatinae, range limit, reproductive isolation, tension zone.

Introduction

A suite of ecological and evolutionary processes are thought to determine species geographical ranges, invoking various combinations of dispersal ability, the steepness of environmental gradients, availability of (adaptive) genetic variance at range margins and interspecific interactions (reviewed in Bridle & Vines, 2006, also see Price & Kirkpatrick, 2009). The problem of species range limits becomes even more intriguing when considering the maintenance of boundaries between closely related species (e.g. occurring in parapatry or sympatry) because, in the the absence of diversifying forces, closely related species are expected to exhibit a high degree of similarity in their ecological niche as well as their phenotypes. Thus, shared ancestry may predispose these towards intense competition for resources (i.e. limiting phylogenetic similarity, Violle *et al.*, 2011).

When closely related species come into contact, there is also potential for natural selection against heterospecific matings via reduced offspring fitness and hybrid breakdown (Dobzhansky, 1936; Muller, 1942), as well as reproductive interference due to the direct costs of mating. For example, harassment during misdirected mating attempts is relatively widespread and can result in non-trivial energetic costs and physical damage, not to mention lost feeding and mating opportunities (Lima & Dill, 1990; Svensson, 2013). Moreover, some studies have cited direct negative consequences of heterospecific matings on

Correspondence: David Punzalan, Department of Natural History, Royal Ontario Museum, Toronto, Ontario M5S 2C6, Canada. E-mail: dpunzalan@mac.com

parental viability and fertility (e.g. Ribeiro & Spielman, 1986; Bargielowski *et al.*, 2013; Ting *et al.*, 2014). In fact, imperfect reproductive isolation has been invoked as a potentially important mechanism for enforcing parapatric distributions and costly heterospecific matings (including inviable hybrids) broadens the conditions under which a stable geographical border between species is expected to evolve (Goldberg & Lande, 2006). In contrast, if viable hybrid offspring can be produced, the introgression of maladaptive alleles can hinder local adaptation of one or both species (Groening & Hochkirk, 2008; Abbott *et al.*, 2013, but also see Seehausen, 2004 for an opposing view). Ultimately, both types of reproductive interference may have profound implications for the process of speciation, possibly facilitating the evolution of (species) assortative mating and reproductive character displacement (i.e. 'reinforcement'; Dobzhansky, 1937; Servedio & Noor, 2003). Despite hybrid zones being relatively common in nature (reviews in Arnold, 1997) and enthusiasm for the idea of reinforcement, good evidence of the latter is rather sparse (Servedio, 2004).

Here, we evaluate several mechanisms that might influence the parapatric distribution of two species in a genus of a true bug (*Phymata*). This pair offers a curious example of species that are considered to have a distinct morphology and roughly separate geographic ranges, despite the absence of an obvious geographic barrier, while sharing very similar habits or ecological 'niche'. Some previous observations, however, suggests biogeographical variation and potential for hybridisation between these two species (see Methods: Background and Source Populations) underscoring the need to consider reproductive and geographic boundaries as related problems. Specifically, the present study (i) assessed the degree of pre-mating reproductive isolation, evident as the propensity to mate assortatively by species, (ii) evaluated potential costs of heterospecific mating, assayed as negative effects on female fecundity/fertility, (iii) experimentally confirmed that viable *americana-pennyslvanica* hybrids could be formed in the laboratory, (iv) looked for evidence of post-zygotic isolation, inferred from the relative viability of these hybrids, and (v) evaluated morphological evidence of hybridisation in the wild.

Methods

Background and source populations

Phymata americana Melin and *P. pennsylvanica* Handlirsch (Heteroptera: Reduviidae) exhibit a roughly parapatric distribution, with the latter having a more southerly distribution in temperate North America. In Ontario, Canada, these are the only two representatives of the genus (Maw *et al.*, 2000). In the literature, nomenclature is a bit confused with some references also to *P. americana americana* but also *P. pennsylvanica americana* (e.g. Evans, 1931; Balduf, 1939; Kormilev, 1962). Both are generalist predators that use a wide range of flower species (Balduf, 1939; Yong, 2005; Punzalan *et al.*, 2008b). For brevity, we simply use *P. americana* and *P. pennsylvanica* to refer to the two recognised forms, which are distinguished primarily by male morphology (Evans, 1931; Kormilev, 1962). There is some uncertainty whether the two should be considered separate species or merely subspecies (Kormilev, 1962) but the issue of species concept is not crucial for the present study. For simplicity we refer to these two forms as 'species'. The two species sometimes co-occur sympatrically in 'contact zones' but both museum material and field collections often reveal individuals from such localities to be ambiguous or intermediate in morphology (Swanson, 2013, Figure S1), suggesting potentially frequent hybridisation events and, possibly, sites comprised of hybrid swarms. In fact, such hybridisation may be the main contributor to the aforementioned confusion in identification and nomenclature.

We conducted a series of studies (Parts 1–5) using individuals collected from three sites in Ontario, Canada, consisting of what appear to be temporally stable populations: Koffler Scientific Reserve, King, where individuals are all morphologically considered *P. americana* (hereafter referred to as population A); Stevensville Conservation Area, Fort Erie, approximately 126 km southeast from population A, where individuals are considered to represent *P. pennsylvanica* (population P) and from a site in the Short Hills Provincial Park (population M) geographically located roughly between sites A and P (see Figure S1). M is located on a site that is suspected of being populated by 'natural hybrids'; individuals exhibit a wide range of phenotypes with various characters appearing to resemble both *P. americana* and *P. pennsylvanica*. Note, that the distinctiveness between A and P in genetically-based morphology is corroborated in the present paper (Part 5). Field collection of adults and penultimate nymphs was conducted haphazardly by hand collection of bugs as encountered. Additional details corresponding to the collection methods used in Parts 1–5 are provided in the sections that follow.

Laboratory rearing and maintenance

All rearing and behavioural assays were conducted in an environmental chamber maintained at $26 \pm 1\,^{\circ}C$, 30% RH, 14L:10D (light–dark cycle). Except where noted, bugs were reared in groups in colony cages corresponding to separate populations and, upon reaching adulthood, moved to individual cages. Colony cages consisted of a transparent plastic container, roughly $30.5 \times 20.0 \times 28.0\,cm^3$ (L × W × D) with a ceiling made of fine mesh. The floor of each cage was perforated with several holes and supported a 4 cm layer of potting soil and a single planted *Kalanchoe blossfeldiana*. Cages were seeded with approximately 100 Collembola (*Folsomia* sp.) and 4 isopods (*Armadillium vulgare*) to act as detritivores and to inhibit the growth of mold and fungi. Every 2 days, these cages were provisioned with approximately 200 adult *Drosophila melanogaster* and misted with water, using an atomiser. Plants were watered (by immersing the cage bottom in a sink filled with water for 20 s). Individual cages consisted of a 19.5-ml clear plastic floral watering tube (Aquatube © #53), provided with a 4 cm × 1 cm segment of a wooden coffee stir stick as an oviposition substrate. Individual cages were provisioned with approximately 16 live *Drosophila* sp. per day and provisioned with two 5–7-day-old calliphorid fly (*Lucillia* or *Calliphora* spp.) pupae

every 6 days. Where applicable, eggs were 'overwintered' naturally, i.e. stored outdoors in a cotton mesh bag, between 9 October 2013 and 29 April 2014 and 15 October 2014 until 1 May 2015 in Toronto, Ontario, Canada. Egg diapause was terminated by placing eggs/oviposition substrate in 900-ml glass jars nested in a 60 -litre transparent bin with the lid partially closed, filled with moistened vermiculite, and maintained under laboratory conditions.

Part 1: Pre-mating isolation (species assortative mating) assay

To evaluate whether males and females of each species exerted any mating biases, discriminating according to population identity, we conducted mating trials with two populations representing *P. americana* (A) and *P. pennsylvanica* (P), respectively. On 12 August 2013, we collected adult males and females from sites A and P. Because females in this assay were likely already inseminated (i.e. mated before collection from the wild), the experimental design has the advantage of using individuals with realistic mating histories (i.e. previously mated and potentially choosy). For example, in many species of insects, virgin females show weaker degrees of choosiness and mate preference than their mated counterparts (see Judge *et al.*, 2010 and references within, but also see Gershman *et al.*, 2014). Bugs were weighed prior to holding in individual cages. The following morning, male–female pairs were formed by assigning bugs to one of four possible treatments, resulting in a full-factorial design of con- and heterospecific matings. These treatments were comprised of a female from population A + a male from population P (hereafter AP): $N = 23$, P female + A male (PA): $N = 21$, A female + A male (AA): $N = 21$, P female + P male (PP): $N = 23$. Mating arenas consisted of an inverted transparent plastic cup placed over the blank side of an index card (7.3 cm × 12.3 cm), in the same environmental conditions as previously described. Females were added to the arena first at 10.00 hours, followed by the males of appropriate treatment approximately 30 min later. Arenas were checked every 5 min after that for a total of 180 min, noting whether/when pairs successfully copulated. The assay essentially employed a 'no-choice' design which has been shown to be valuable for assaying mating preferences in several systems (Shackleton *et al.*, 2005; Dougherty & Shuker, 2014). Importantly, this method seems most appropriate for the mating system of these bugs than a scenario where one female is presented with multiple males because, under natural conditions, it is unlikely that females are receiving stimuli from multiple males simultaneously. Mating is initiated by males, which upon encountering a female, effectively exclude other males via mate guarding or 'coupling' (Punzalan *et al.*, 2008b). From the coupled position, males engage in ritualised behaviours, apparently courting females using various tactile and stridulatory signals, with females able to exercise choice over whether or not to copulate (Punzalan *et al.*, 2008b).

To evaluate possible differences between conspecific (AA or PP) and heterospecific (AP or PA) pairs in the frequency of mating, we first used a χ^2 test to evaluate an association between treatments and successful copulation. We performed a second analysis: a logistic mixed regression of mating success (0 or 1) as the response variable, and male population identity and female population identity (A or P) as categorical predictors, their interaction, and female residual weight (in mg, after cube root transformation and regression on pronotum width in mm) as a covariate. The main effects of male and female population identity were meant to account for possible population differences in male persistence and female receptivity. The interaction term was interpreted as the strength of the assortative mating; positive values of this term reflect a greater likelihood of successful mating for conspecific pairs while negative values reflect a higher likelihood of matings among heterospecific pairs. The inclusion of the covariate was motivated by previous studies of conspecific matings in several species of ambush bugs (Dodson & Marshall, 1984; Punzalan *et al.*, 2008b) that showed a propensity to mate is related to female reproductive condition (i.e. relative weight is correlated with egg maturation).

Part 2: Post-mating prezygotic isolation (reproductive interference) assay

To assess whether heterospecific matings resulted in any direct fitness consequences for fecundity and fertility, we maintained the pairs in Part 1 for 31 days under laboratory conditions; eggs were collected and substrate replaced approximately every 2 days. After overwintering, eggs were returned to the environmental chambers and monitored for hatching for 14 days. Again, since females were probably already mated before the experiment, any treatment differences in fecundity and fertility would potentially reflect processes mediated by postcopulatory mating biases (e.g. species differences in sperm competition and cryptic choice). Although we we could not determine the proportion of eggs fertilised via heterospecific males (vs. eggs fertilised by stored sperm from previous matings in the wild) in the clutches generated in the AP and PA treatments, our goal was simply to determine whether heterospecific matings might affect the average number of eggs produced and hatched. That is, any reductions in components of female fitness compared to the conspecific pairings should be indicative of negative impacts of heterospecific mating. Differences among the four treatments (i.e. AA, AP, PA, PP) with respect to success/complete failure to produce eggs was first evaluated using a chi-square test. Among those pairs that succeeded, we evaluated treatment effects on the total number of eggs laid (log-transformed) per pair, and the proportion hatched (arcsine square root), using separate two-way ANOVAs (i.e. with the male population, female population and their interaction as predictors).

Part 3: Hybridisation in laboratory crosses

To experimentally assess whether A and P populations were indeed capable of producing viable hybrids, we collected 206 fourth and fifth instar nymphs from the A site, between 1 and 10 July 2013 and these were maintained in the laboratory until adulthood. From these, we obtained 62 virgin females that were used for subsequent controlled crosses of A females with A or P males (both wild collected 29–30 July 2013; crosses began on 2 August 2013). Unfortunately, juveniles were not found

at the P site, preventing the reciprocal crosses. However, our design allowed for subsequent comparisons between conspecific and heterospecific crosses (i.e. Parts 4 and 5). Each A female was maintained with either an A or P male ($N = 31$ for each treatment) in individual cages. Cages were checked every 2 days for eggs and, if necessary, males were replaced with another male from the appropriate population. Females were allowed to lay eggs for 28 days before eggs were collected and overwintered (as described previously). The following spring, eggs were brought into laboratory conditions and the total number of successfully hatched offspring was recorded; although we made a note of whether a female laid at least one egg, we did not count the number of eggs laid per female. Nymphs were reared to adulthood in individual cages and then euthanised and preserved for later phenotypic measures. Treatment (AA and AP) differences in probability of producing at least one viable offspring were performed using a chi-square test. Of those females that succeeded, treatment differences in the average number of offspring produced was analysed using the Mann-Whitney *U*-test of log-transformed offspring numbers.

Part 4: Post-zygotic isolation (hybrid viability)

One straightforward mechanism for the maintenance of distinct parapatric species is via reduced hybrid fitness (i.e. post-zygotic isolation). We inferred hybrid juvenile viability in two ways. First, we estimated the viability of known F1 hybrids, using the offspring obtained from crosses in Part 3. Nymphs hatched from AA ($N = 222$) and AP ($N = 490$) crosses were held separately by treatment in 900 -ml glass jars (approximately 50 randomly assigned nymphs per jar) and provisioned daily with approximately 200 live *Drosophila melanogaster*. Beginning 9 June 2014 (when many nymphs had grown to third or fourth instar), nymphs were moved to population-specific colony cages and every 6 days, cages were also provisioned with approximately 30 calliphorid pupae. After 21 days, the bugs were moved to individual cages and reared to adulthood under common garden conditions. The proportion of individuals that survived until adulthood was compared to the expected distribution (given their proportional representation; i.e. bugs hatched from each treatment) using a chi-square test.

Drawbacks of the aforementioned assay are that it does not capture postzygotic selection on genetic variation produced from crosses of both (reciprocal) directions and, in particular, it cannot detect possible hybrid breakdown manifested in further generations (e.g. F2 hybrids, backcrosses). Limitations imposed by biological and logistic considerations (i.e. univoltine life cycles with obligate winter diapause and very low reproductive success of laboratory-reared bugs) prohibited assays on such crosses in the laboratory. Instead, our approach was to compare the proportional juvenile viability of putative hybrids (i.e. M population) to those from populations A and P, when reared under common garden laboratory conditions. In 2013, we collected eggs from 40 to 50 wild-caught females of population A, P, and M, respectively and maintained them in separate cages (according to population) under laboratory conditions. The following spring, 1- to–3-day-old first instar nymphs were placed in population-specific ($N = 521$, 419 and 500, for A, P, and M, respectively) colony cages. Nymphs were monitored every 2 days, and upon adulthood, bugs were moved to individual cages. At age 14 days, adults were euthanised, mounted on entomological pins and then photographed for later measurement of phenotypic characters (i.e. Part 5). It is important to note that because M is suspected to be comprised of mixed or hybrid individuals, our assay of juvenile viability in this population potentially reflects the combined viability of both 'pure' and hybrid offspring (i.e. depending on maternal and paternal identities). Nevertheless, any differences among populations here have a straightforward interpretation regarding average juvenile survivorship, with important implications for intrinsic rates of population increase. Population differences in the counts of individuals successfully emerging into adulthood (i.e. juvenile viability) was analysed using a chi-square contingency test with expected counts calculated according to the proportional representation of nymphs from each population at the beginning of the assay.

Part 5: Phenotypic evidence of hybridisation in wild populations

To determine whether morphological differences among populations reflected genetically-based differences, we reared bugs in a common environment (i.e. removed the additive environmental contributions to phenotypic variance) and then measured a suite of traits from progeny reared in Part 4. We expected that individuals from M would, on average, exhibit phenotypes intermediate to those seen in A and P populations. We also compared these to data from known hybrids generated in the laboratory, i.e. the offspring of AP crosses in Part 3. We focused on four traits that have been previously identified as phenotypically divergent between *P. americana* and *P. pennsylvanica*; these were pronotum width (PN), two measures of cuticular melanism on the pronotum: mean dorsal darkness (MD) and mean lateral darkness (ML) as well as an antennal ratio (AR). The latter trait is a composite metric, indicating the length of the terminal (i.e. fourth) antennal segment divided by the sum of the two proximal (i.e. second and third) antennal segments (Figure S2). According to the identification keys provided by Evans (1931) and Kormilev (1962), *P. pennsylvanica* is distinguished as having relatively long terminal antennal segments (AR $\sim$ 1.3) than *P. americana* (AR $\sim$ 1.0). Although those previous studies did not specify female characters as diagnostic, a study of wild-collected individuals from a broad geographical range indicated female ML in *P. pennsylvanica* to be strikingly darker than in *P. americana* (Punzalan & Rowe, 2015). Bugs were pinned, photographed from the dorsal and lateral aspect, under standardised lighting conditions (protocol described in Punzalan *et al.*, 2008b) and antennal segments measured using an ocular micrometer and stereoscopic microscope. Dark colouration, or melanisation, was quantified from digital photographs as mean pixel value (darkness on a greyscale) of a sampled region on the dorsal (MD) or lateral (ML) surface of the thorax, using Scion® Image (methods described in Punzalan & Rowe, 2015 and the Figure S3). Pronotum width in mm was also obtained from digital photographs.

We first evaluated our expectation of population and sex-differences in multivariate phenotypes using a MANOVA that included the four traits as response variables and with sex, population (with four levels corresponding to A, P, M and AP hybrids) and their interaction as independent variables. Subsequently, we focused on the two traits (ML and AR) considered to be pertinent for distinguishing between species. We performed analyses for each sex and trait separately, using one-way ANOVAS with trait as the dependent variable and with the population as the predictor. Post-hoc tests (Tukey's HSD) were used to assess our expectation of significant phenotypic differences between 'pure' (A and P) populations but with 'mixed' or hybrid (M and AP) bugs exhibiting intermediate phenotypes. For phenotypic data, five univariate outliers (from separate analyses by sex and for each trait) were identified; three were from the M population with antennal deformities and two were from the P population with anomolous values for PN and MD suspected to be data entry errors. These individuals were excluded from all analyses. Phenotype data was otherwise normally distributed and did not require transformation. All statistical analyses were performed using R (http://www.R-project.org) or Systat ®. V.10.

Results

Part 1: No evidence of species-assortative mating

Overall, the number of pairs successfully copulating was comparable between conspecific (27 of 42, or 64%) and heterospecific pairs (26 of 46, or 57%) ($\chi^2 = 0.552$, d.f. = 1, $P = 0.457$). The analysis of the mixed model including female weight as a covariate showed that female reproductive condition predicted the likelihood of successful copulation in the laboratory assay (Table 1). With respect to the primary goal of the assay, however, we found no evidence of assortative mating (preferences) between populations; pairs were equally likely to mate irrespective of the population of origin or male–female pair combinations (Table 1).

Part 2: No cost of heterospecific matings on fecundity and fertility

Treatments (i.e. hetero- and conspecific pairs) did not differ in their frequency of producing at least one clutch of eggs (approximately 56 of 88, or 64% success overall); failure to produce eggs (often because of death of one or both individuals) was marginally significantly different among the four treatments ($X^2 = 6.904$, $P = 0.075$, d.f. = 3). Every laid clutch yielded at least one successful hatch. Among those pairs that did produce eggs, treatments did not differ in the quantity of eggs laid over the duration of the assay (Fig. 1). Similarly, the proportion of hatched eggs did not differ among treatments, though there was a marginally nonsignificant main effect of male origin (Table 2, Fig. 1), suggesting higher fertility of males from population A. Collectively, we did not find evidence of negative fertility or fecundity consequences associated with heterospecific matings.

Table 1. Predictors (and standard error of the estimates, SE) of successful copulation in conspecific and heterospecific pairs of *Phymata americana* and *P. pennsylvanica*, in laboratory trials. *Origin* refers to the population/species identity of an individual.

Model term	*Estimate*	SE	*t*	*P*
Constant	0.861	0.478	1.80	0.072
Female origin	−0.448	0.654	−0.69	0.493
Male origin	−0.633	0.665	−0.95	0.341
Female × male origin	0.532	0.935	0.57	0.569
Residual female weight	2.763	1.006	2.75	0.006

Estimates are from a mixed-model logistic regression. Full Model Log-Likelihood = −59.143, $X^2 = 9.998$, d.f. = 4, $P = 0.040$.

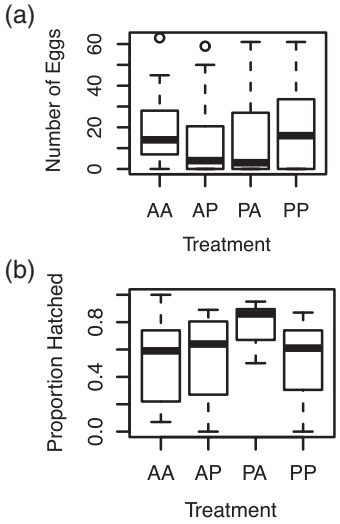

Fig. 1. Fecundity (a) and fertility (b) of conspecific- and heterospecific-mated pairs in Part 2. AA, AP, PA and PP, indicate treatments where the species identities (A or P) of females and males are denoted by the first and second letter, respectively. Fecundity refers to the number of eggs laid and fertility refers to the proportion of eggs hatched.

Part 3: Formation of viable hybrid offspring

Laboratory crosses confirmed that viable F1 hybrids are possible, at least in the one direction tested in this study (i.e. A females mated with P males). The two crosses did not differ in the proportion of females that laid at least one egg (AA: 25 of 31, AP: 26 of 31, Fisher's Exact test two-tailed $P > 0.999$). Among those crosses that produced progeny (i.e. eggs were hatched by 42 of 62 or 68% of females), the heterospecific crosses produced significantly more progeny than conspecific crosses ($U = 130.00$, $N_{AA} = 19$, $N_{AP} = 22$, $P = 0.039$, Fig. 2).

Part 4: No evidence of hybrid disadvantage but possible hybrid breakdown

Survival of juveniles from hybrid (A female × P male) crosses did not differ significantly from the those of comparable pure crosses (proportional survival: AP = 62 of 490 or 13%, AA = 24 of 222 or 11%; Fisher's Exact test, two-tailed $P = 0.536$). This

Table 2. Direct effects of conspecific and heterospecific matings on two components of fitness in *Phymata americana* and *P. pennsylvanica*, in laboratory trials.

Response variable, predictor	*SS*	d.f.	*MS*	*F*	*P*
Fecundity					
Female origin	1.383	1	1.383	2.03	0.160
Male origin	0.294	1	0.294	0.43	0.514
Female × male origin	0.224	1	0.224	0.33	0.568
Error	35.354	52	0.680		
	Full model multiple $R^2 = 0.057$				
Fertility					
Female origin	0.275	1	0.275	2.50	0.120
Male origin	0.376	1	0.376	3.42	0.070
Female × male origin	0.256	1	0.256	2.32	0.134
Error	5.72	52	0.110		
	Full model multiple $R^2 = 0.121$				

Origin refers to the population/species identity of an individual. Effects were evaluated using separate two-way ANOVA. Fecundity represents the number of eggs laid and fertility represents the proportion of eggs hatched.

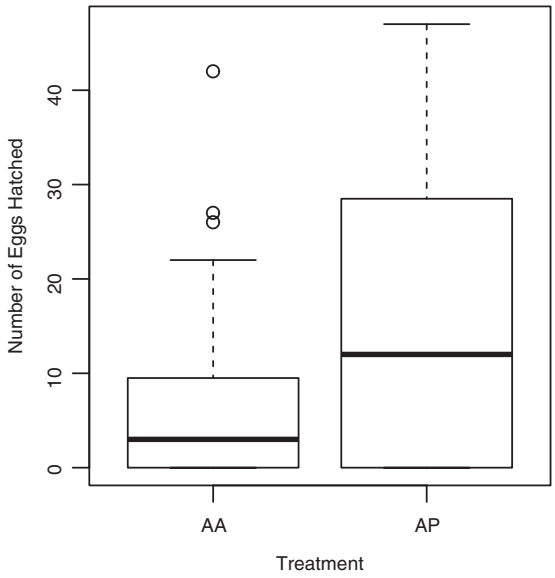

Fig. 2. Hatching success of hybrid (AP) versus pure (AA) crosses in Part 3. AP, indicates female *Phymata americana* mated to a male *P. pennsylvanica,* whereas AA refers to crosses between male and female *P. americana*.

indicates, at least in the unidirectional cross (and under laboratory conditions), imperceptible selection against F1 hybrids under laboratory conditions. However, comparisons of proportional juvenile survival (to adulthood) among A, P and M populations revealed a possible source of post-zygotic isolation; hybrid (M) populations yielded half as many (8%) progeny surviving to adulthood than A (16%) and P (16%) raised in the same conditions ($X^2 = 25.17$, $P = 0.0001$, d.f. = 2).

Part 5: Divergence between wild populations and hybrid phenotypes

We detected significant genetically based phenotypic divergence among populations (population: Wilks' Lambda = 1591, $F_{12,526.8} = 44.04$, $P < 0.0001$), as well as sexual dimorphism (sex: Wilks' Lambda = 0.0895, $F_{4,199} = 506.36$, $P < 0.0001$), but also a population-by-sex interaction (Wilks' Lambda = 0.6955, $F_{12,526.8} = 6.46$, $P < 0.0001$). Consistent with our *a priori* assumption that P and A sites corresponded to *P. pennsylvanica* and *P. americana*, respectively, we observed significant differences between populations (P > A) in both male antennal morphology (AR: $F_{3,79} = 108.30$, $P < 0.0001$, Fig. 3) and female lateral colouration (ML: $F_{3,123} = 85.47$, $P < 0.0001$, Fig. 3). Furthermore, the intermediate phenotypes (and generally greater phenotypic variance) seen in the 'mixed' population (M) and their similarity in mean phenotypes generated by the AP crosses is consistent with the former population composed of hybrids (Figs 3 and 4, Figure S4).

Discussion

There is growing recognition that the ecological and genetic factors that maintain reproductive barriers between closely related species is also pertinent to the maintenance of geographic ranges (Goldberg & Lande, 2006). Closely related species have the potential to mate heterospecifically and hybridise, which can have various outcomes including range expansion and species collapse (reviewed in Abbott *et al.*, 2013). At the same time, the potential for hybridisation is thought to be key to the evolution of mechanisms of prezygotic isolation and, thus, the process of speciation. Here we used morphological, behavioural and fitness (component) data to assess species delimitation and hybridisation between a pair of closely related species in the bug genus *Phymata.* We found evidence of genetically-based morphological differences between sampled populations of *P. americana* and *P. pennsylvanica* consistent with their previous treatment as distinct 'species'. However, we did not detect any evidence of species-assortative mating in the laboratory. Although low power to detect assortative mating is conceivable, we should point out that our assay was capable of detecting an effect of female weight on (positively covarying with) mating probability. This result is in accord with previous studies in the field and laboratory (Dodson & Marshall, 1984; Punzalan *et al.*, 2008b). We believe this reflects female receptivity that coincides

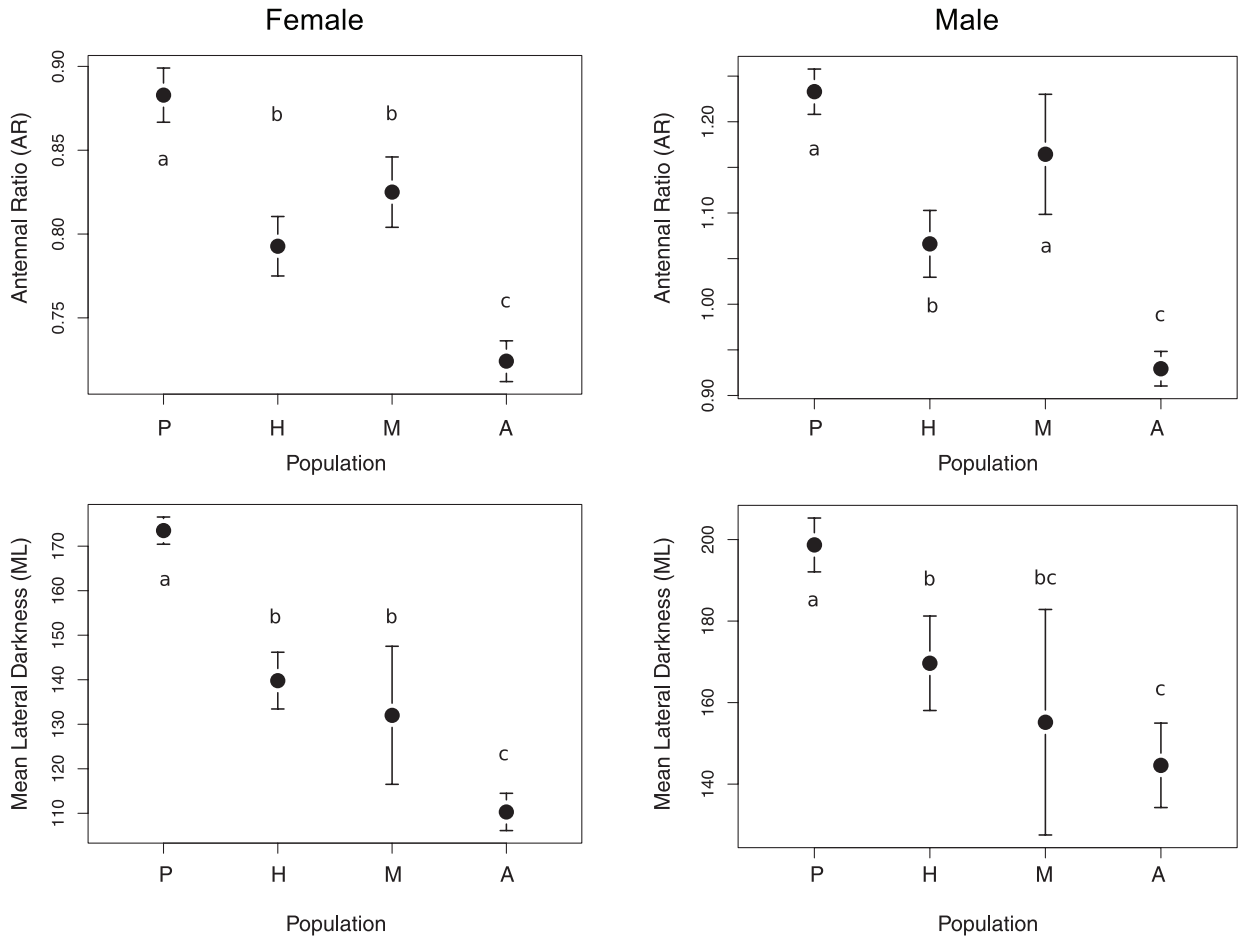

Fig. 3. Phenotypic means and 95% CIs for female (left panels) and male (right panels) progeny of *Phymata pennsylvanica* (P) and *P. americana* (A), putative mixed or hybrid (M) populations, and from laboratory hybrid crosses (H), with respect to antennal ratio (AR), and mean lateral melanism (ML, measured in units of mean pixel value). Lower case letters indicate significantly different groups using Tukey's HSD ($\alpha = 0.05$).

with the extent of maturation of ova available for fertilisation, although it is equally consistent with male preference for fecund females. This observation is also consistent with the possibility that female mating decisions are influenced by trade-offs between feeding and mating (Rowe *et al.*, 1996; Ortigosa & Rowe, 2002), assuming relatively heavy females may be those who had fed more recently. Similar previous assays have also revealed strong effects of male courtship vigor in predicting the outcome of mating interactions (Punzalan *et al.*, 2008b). The present study shows that, by comparison, any potential mating biases toward mating with conspecifics are, at best, weak.

Even when precopulatory mating biases are absent, one possible contributor to reproductive isolation is direct costs associated with heterospecific mating. Although some studies have reported direct negative consequences of heterospecific matings for female components of fitness (reviews in Groening & Hochkirk, 2008; Svensson, 2013), we did not observe any treatment differences in either fecundity or fertility in the studied populations.

We show here that, ultimately, the two species are capable of generating viable F1 hybrid offspring, at least in the unidirectional crosses we were able to perform. Curiously, we found that the relative success (i.e. number of viable offspring) was higher for heterospecific crosses than for matings between members of the same species. It is unclear whether this reflects intrinsic population differences in male vigour (i.e. higher in P males) with respect to courtship and crytptic female choice or is simply because of possible differences in the average age/condition of wild-caught males from either collection site. As we could not perform the fully reciprocal crossing design, we could not evaluate the degree to which hybridisation might be asymmetric–a scenario that appears to be taxonomically widespread (Turelli & Moyle, 2007). That is, we cannot rule out that crosses in the reverse direction (i.e. A males with P females) might result in inviable zygotes. Nonetheless, the production of viable offspring indicates at least some potential for gene flow between the two species. AP crosses resulted in F1 progeny of both sexes, with a slightly female-biased tertiary (i.e. at maturity) sex ratio (61% of 53 adults). Although it is unknown to what degree this reflects any sex-biases at fertilisation and/or sex differences in juvenile viability, this observation is consistent with previous studies reporting female-biased collections of juveniles

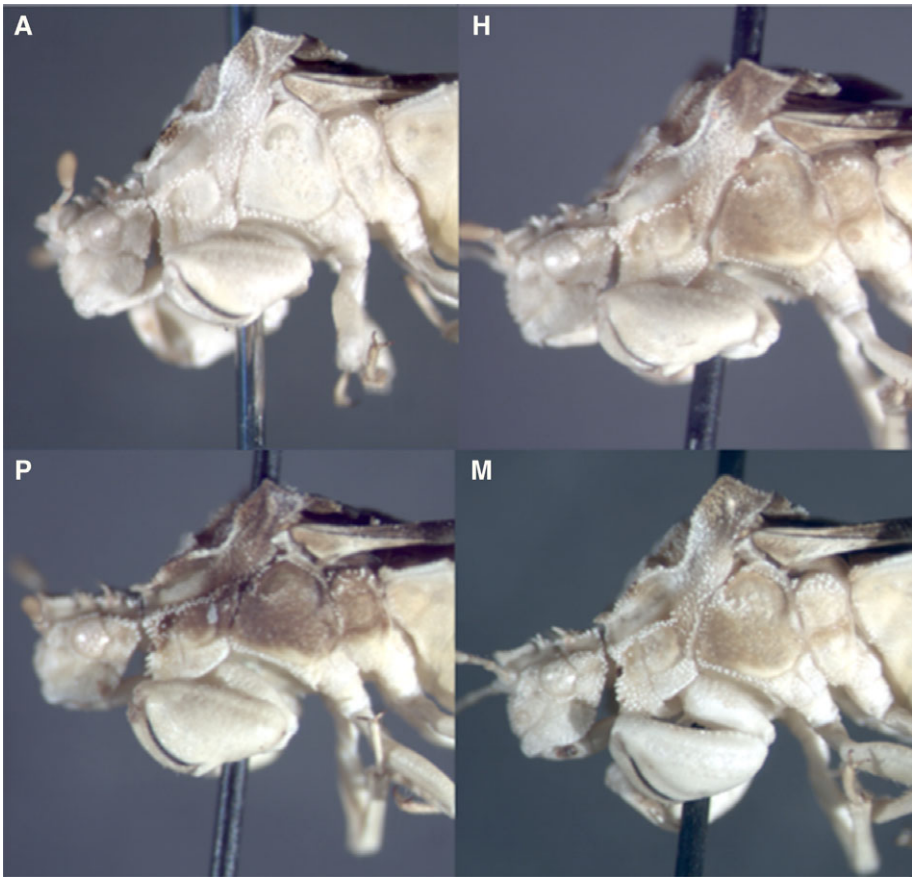

Fig. 4. Representative photographs of female lateral melanism (ML) in laboratory-reared bugs from the three populations (A, P and M), as well as from F1 hybrid (H) crosses. [Colour figure can be viewed at wileyonlinelibrary.com].

in the wild (i.e. 63% in Punzalan *et al.*, 2008a) and comparable to the percentage of female progeny yielded from the populations represented in the common garden study (i.e. population A: 57%, population P: 59%, M: 68%). Clearly, these data do not indicate the complete absence of one sex, as is sometimes seen for hybrid crosses (i.e. Haldane's rule).

We found some circumstantial evidence that the progeny of putative hybrid populations (M) had lower juvenile viability than progeny of parental species. Although the analyses of viability data did not account for differences in larval rearing density within colony cages, the rank order of initial densities (A > M > P) does not correspond to proportional juvenile survival, suggesting that our results are not simply the result of density-dependent effects (e.g. competition, larval cannibalism). Also, a recent observation (D. Punzalan, unpublished) based on juvenile survivorship among individually reared nymphs from the same three populations recovered a similar qualitative result: the M population had the lowest proportion of survivors. We interpret these as tentative results, although the severe reduction in hybrid fitness in later generations appears to be common, arising from a multitude of mechanisms including intrinsic genetic incompatibilities and ecologically-mediated natural selection (reviewed in Burton *et al.*, 2013). We also showed that F1 hybrid offspring exhibit intermediate (to the parental species) phenotypes similar to those found in natural populations thought to be composed of hybrids and, consistent with the expectation (see Barton & Gale, 1993), increased phenotypic variance.

We acknowledge the limitations of the present study, which represents only one pair of 'parental species' populations, in addition to a probable hybrid population. Unfortunately, the true distribution of hybrid populations, including the number of 'contact zones' is not presently known. Ideally, similar pairs of populations or appropriate sampling transects would be identified, allowing for replication of experiments similar to those presented here. Although we only formally assayed one putative hybrid population, populations bearing similar, intermediate phenotypes appear to be common in localities that are also geographically situated between A and P sites (Figure S1 and D. Punzalan, unpublished). This probably reflects frequent hybridisation and a contact zone of substantial size, not unlike the situation reported in many other insects (e.g. Harrison & Rand, 1989; Spence, 1990). Such introgression between populations is a likely contributor to the frequent confusion and misidentification associated with this pair of *Phymata* species. Consequently, the appropriate systematic position and nomenclature for these, i.e. whether 'good' species versus subspecific rank, may need to be revisited, particularly with appropriate molecular genetic tools. Development of such genetic resources could also

shed light on general aspects of the importance of introgression; 'natural hybridisation' has been touted as particularly useful for fine-scale mapping of the genes involved in reproductive isolation (see Harrison & Larson, 2014).

As a side note, lab-reared male *P. americana* were noticeably paler in lateral colouration (ML) than their *P. pennsylvanica* counterparts (Figs 1 and 2) but also much paler than usually seen in wild caught *P. americana*, which are usually even darker than *P. pennsylvanica* (see Punzalan & Rowe, 2015). Color differences between wild caught and laboratory-reared insects have been reported in other taxa (e.g. Pegram *et al.*, 2013) and expression of melanism in *P. americana* is known to be strongly condition dependent (Punzalan *et al.*, 2008a). However, our results suggest weaker condition dependence of melanic coloration in *P. pennsylvanica* and, thus, the divergence between these two species in the underlying determinants of colour development.

Thus, our results are consistent with natural hybridisation between *P. americana* and *P. pennsylvanica* and can be partly attributed to a lack of mate discrimination; both species readily engaged in heterospecific matings, which did not result in direct costs with respect to reductions in fecundity or fertility. Although F1 hybrid offspring from lab crosses did not suffer lower viability, our data indicates the possibility of hybrid breakdown, manifested after subsequent generations of hybridisation. Whether the proportion of hybrids that emerge as adults suffer from further reductions in fitness from late-life components (e.g. mating success) has not been formally explored but some data from MM pairings, benchmarked against hetero- and conspecific pairings (i.e. Part 2), suggests comparable mean fecundity and fertility (D. Punzalan, unpublished).

The ecological and historical causes of morphological divergence between *P. americana* and *P. pennsylvanica* are not well understood, nor are the factors that might determine their respective geographic ranges. In fact, we consider the maintenance of the observed phenotypic differences between A and P populations striking, given their close geographical proximity coupled with what appears to be a lack of reproductive isolation. In theory, this could be accomplished with steep ecological gradients (Kirkpatrick & Barton, 1997) or pronounced resource specialisation (Price & Kirkpatrick, 2009), yet the two species appear to share similar habits: generalist predators with no obvious specificity regarding host plant/hunting patch. Note that in the present study, the hatching success of *P. pennsylvanica* did not appear to suffer despite diapaused eggs being subjected to the natural overwintering conditions typical of *P. americana*. This suggests that climate (e.g. temperature, day length) during the egg stage cannot, alone, account for the northern limit of *P. pennsylvanica*. Similarly, it is unclear how long hybrid populations persist, though a balance between dispersal/gene flow and selective forces (e.g. local adaptation and hybrid breakdown) may allow the hybrid zone to be relatively stable over time (Barton & Hewitt, 1985). We acknowledge that in laboratory studies, it is difficult to capture all stimuli or environmental conditions that occur in a natural setting and that these could, conceivably, influence the outcome of mating interactions and juvenile development/viability. However, the present data collectively suggests weak partial post-zygotic isolation between *P. americana* and *P. pennsylvanica* but no evidence of pre-zygotic mechanisms to reinforce viability selection against hybrids.

A potential alternative explanation of the observed data is that the populations sampled here merely represent subspecific variation along some geographical cline. The apparent lack of reproductive isolating mechanisms observed in the present study suggests that the traditional designation of *P. americana* and *P. pennsylvanica* as separate species may be unwarranted. Consistent with this, preliminary analyses also indicate low degrees of molecular genetic differentiation (C. Weirauch, pers. comm.). More 'mixed' populations would need to be sampled to fully address the possibility of clinal variation, though we should point out that analyses of intraspecific color pattern variation (i.e. within *P. americana* and *P. pennsylvanica*, respectively, using specimens from a broad geographic range) did not show a significant latitudinal cline (see S5 and S6 in Punzalan & Rowe, 2015). Nonetheless, the present study makes it clear that there is genetically-based morphological variation among the sampled populations, and more studies are needed to elucidate the factors that maintain the characteristic phenotypes and geographic distributions, despite what appears to be considerable potential for introgression.

Acknowledgements

We are grateful to H. Rodd for the use of laboratory space, to B. Cole for access to materials, and to L. Yun for fly cultures. Thanks to C. Weirauch, P. J. Malé, and two anonymous reviewers for comments on the manuscript, and to museum staff for access to their respective collections. We are also grateful to *rare* Charitable Research Reserve and the Niagara Region Conservation Authority for access to collection sites. D. P. was supported by a Natural Sciences and Engineering Research Council grant to L. R.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/een.12380

Figure S1. Partial map of localities for specimens of *Phymata americana* (red) and *pennsylvanica* (blue) in S. Ontario and nearby U.S. states. Localities composed of suspected hybrids are indicated in purple. Inset shows sites representative of *americana* (population A: red triangle), *pennsylvanica* (P: blue triangle) and a putative mixed/hybrid zone (M: purple triangle), from which we collected live individuals for assays in the present study. Locality data was obtained from specimens housed in a number of museums including: American Museum of Natural History, Canadian National Museum for Insects and Arachnids, Carnegie Museum of Natural History, National (Smithsonian) Museum of Natural History, Royal Ontario Museum, University of California Riverside Entomology Museum, University of Michigan Museum of Zoology and the University of Guelph Insect Collection. Additional localities and specimens were

obtained from field collections by DP. Putative hybrids refer to individuals of morphologically ambiguous identity using established keys and determined/adjudged by DP.

Figure S2. Species differences in the antennal ratio (AR). Depicted are the right antenna of laboratory reared bugs, representative of *Phymata americana* (A) and *pennsylvanica* (P). AR is calculated as the length of the terminal antennal segment divided by the sum of the lengths of the two proximal antennal segments.

Figure S3. Measurement of the melanic colour pattern. Specimens were placed on a custom-made stage with an adjustable (position) greyscale and length standard: four printed squares (hypotenuse 4.03 mm) of black, 50%, gray 25% grey and white. These corresponded to average adjusted darkness values of 261, 143, 92 and 60, respectively. Bugs were photographed using a Nikon Coolpix™ 995 digital camera with a LED ring light (Nikon SL-1). Images were analysed using Scion® Image (on Microsoft © XP) to obtain measures of two colour pattern traits: MD, the mean darkness of a circular patch on the dorsal surface of the prothorax, between the left or right posterior lobe and longitudinal ridge (panel A), and ML, the lateral surface mesothorax (panel B). Darkness refers to the 'value' (average number of black pixels) over a pre-determined location on the integument (indicated by circles). For each photograph, each image was recalibrated according to the known values from the greyscale standards.

Figure S4. Treatment differences in pronotum width and the mean dorsal melanism. Depicted are phenotypic means and 95% CIs for progeny of *P. pennsylvanica* (P) and *P. americana* (A), putative mixed or hybrid (M) populations and from laboratory hybrid crosses (H). Lower case letters indicate significantly different groups using Tukey's HSD ($\alpha = 0.05$). Pronotum width (PN) is measured in mm and mean dorsal darkness (MD) is measured in units of the mean pixel value.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N. *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Arnold, M.L. (1997) *Natural Hybridization and Evolution*. *Oxford Series in Ecology and Evolution*. Oxford University Press, Oxford, U.K.

Balduf, W.V. (1939) Food habits of *Phymata pennsylvanica americana* Melin (Hemip.). *Canadian Entomologist*, **71**, 66–74.

Bargielowski, I.E., Lounibos, L.P. & Carrasquilla, M.C. (2013) Evolution of resistance to satyrization through reproductive character displacement in populations of invasive dengue vectors. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 2888–2892.

Barton, N.H. & Gale, K.S. (1993) *Genetic Analysis of Hybrid Zones* (ed. by R. G. Harrison), pp. 13–45. Oxford University Press, New York, New York.

Barton, N.H. & Hewitt, G.M. (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Bridle, J.R. & Vines, T.H. (2006) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, **22**, 140–147.

Burton, R.S., Pereira, R.J. & Barreto, F.S. (2013) Cytonuclear genomic interactions and hybrid breakdown. *Annual Review of Ecology and Systematics*, **44**, 281–302.

Dobzhansky, T. (1936) Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, **21**, 113–135.

Dobzhansky, T. (1937) *Genetics and the Origin of Species*. Columbia University Press, New York, New York.

Dodson, G. & Marshall, L.D. (1984) Mating patterns in an ambush bug *Phymata fasciata*. *American Midland Naturalist*, **112**, 50–57.

Dougherty, L.R. & Shuker, D.M. (2014) Precopulatory sexual selection in the seed bug *Lygaeus equestris*: a comparison of choice and no-choice paradigms. *Animal Behaviour*, **89**, 207–214.

Evans, J.H. (1931) A preliminary revision of the ambush bugs of North America, (Hemiptera, Phymatidae). *Annals of the Entomological Society America*, **24**, 711–738.

Gershman, S.N., Delcourt, M. & Rundle, H.D. (2014) Female preferences for male cuticular hydrocarbons in *Drosophila serrata* are not affected by female age or mating status. *Journal of Evolutionary Biology*, **27**, 1279–1286.

Goldberg, E.E. & Lande, R. (2006) Ecological and reproductive character displacement on an environmental gradient. *Evolution*, **60**, 1344–1357.

Groening, J. & Hochkirk, A. (2008) Reproductive interference between animal species. *Quarterly Review of Biology*, **88**, 257–282.

Harrison, R.G. & Larson, E.L. (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.

Harrison, R.G. & Rand, D.M. (1989) Mosaic hybrid zones and the nature of species boundaries. *Speciation and its Consequences* (ed. by D. Otte and J. A. Endler), pp. 111–133. Sinauer, Sunderland, Massachusetts.

Judge, K.A., Tran, K.-C. & Gwynne, D.T. (2010) The relative effects of mating status and age on the mating behavior of female field crickets. *Canadian Journal of Zoology*, **88**, 219–223.

Kirkpatrick, M. & Barton, N.H. (1997) Evolution of a species' range. *American Naturalist*, **150**, 1–23.

Kormilev, N.A. (1962) Revision of the Phymatinae. *Philippine Journal of Science*, **89**, 287–486.

Lima, S.L. & Dill, L.M. (1990) Behavioural decisions made under the risk of predation: a review and prospectus. *Canadian Journal of Zoology*, **68**, 619–640.

Maw, H.E.L., Footit, R.G., Hamilton, K.G.A. & Scudder, G.G.E. (2000) *Checklist of the Hemiptera of Canada and Alaska*. NRC Research Press, Ottawa, Canada.

Muller, H.J. (1942) Isolating mechanisms, evolution and temperature. *Temperature, Evolution, Development*. *Biological Symposia: A Series of Volumes Devoted to Current Symposia in the Field of Biology*, Vol. **6** (ed. by T. Dobzhansky), pp. 71–125. Jaques Cattell Press, Lancaster, Pennsylvania.

Ortigosa, A. & Rowe, L. (2002) The effect of hunger on mating behaviour and sexual selection for male body size in *Gerris buenoi*. *Animal Behaviour*, **64**, 369–375.

Pegram, K.V., Nahm, A.C. & Rutowski, R.L. (2013) Warning color changes in response to food deprivation in the Pipevine Swallowtail butterfly (*Battus philenor*). *Journal of Insect Science*, **13**, 1–16.

Price, T.D. & Kirkpatrick, M. (2009) Evolutionarily stable range limits set by interspecific competition. *Proceedings of the Royal Society of London B*, **276**, 1429–1434.

Punzalan, D. & Rowe, L. (2015) Evolution of sexual dimorphism in phenotypic covariance structure in *Phymata*. *Evolution*, **69**, 1597–1609.

Punzalan, D., Cooray, M., Rodd, F.H. & Rowe, L. (2008a) Condition dependence of sexually dimorphic colouration and longevity in the ambush bug *Phymata americana*. *Journal of Evolutionary Biology*, **21**, 1297–1306.

Punzalan, D., Rodd, F.H. & Rowe, L. (2008b) Sexual selection on sexually dimorphic traits in the ambush bug *Phymata americana*. *Behavioral Ecology*, **19**, 860–870.

Ribeiro, J.M. & Spielman, A. (1986) The satyr effect: a model predicting parapatry and species extinction. *American Naturalist*, **128**, 513–528.

Rowe, L., Krupa, J.J. & Sih, A. (1996) An experimental test of condition-dependent mating behavior and habitat choice by water striders in the wild. *Behavioural Ecology*, **7**, 474–479.

Seehausen, O. (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, **19**, 198–207.

Servedio, M.R. (2004) The what and the why of research on reinforcement. *PLoS Biology*, **2**, e420.

Servedio, M.R. & Noor, M.A.F. (2003) The role of reinforcement in speciation: theory and data. *Annual Review in Ecology Systematics*, **34**, 339–364.

Shackleton, M.A., Jennions, M.D. & Hunt, J. (2005) Fighting success and attractiveness as predictors of male mating succes in the black field cricket, *Teleogryllus commodus*: the effectiveness of no-choice tests. *Behavioural Ecology and Sociobiology*, **58**, 1–8.

Spence, J.R. (1990) Introgressive hybridization in Heteroptera: the example of *Limnoporus* Stal (Gerridae) species in western Canada. *Canadian Journal of Zoology*, **68**, 1770–1782.

Svensson, E.I. (2013) Beyond hybridization: diversity of interactions with heterospecifics, direct fitness consequences and the effects on mate preferences. *Journal of Evolutionary Biology*, **26**, 270–273.

Swanson, D.R. (2013) A review of the ambush bugs (Heteroptera: Reduviidae: Phymatinae) of Michigan: identification and additional considerations for two common eastern species. *Great Lakes Entomologist*, **46**, 154–164.

Ting, J.J., Woodruff, G.C., Leung, G., Shin, N.R., Cutter, A.D. & Haag, E.S. (2014) Intense sperm-mediated sexual conflict promotes reproductive isolation in *Caenorhabditis* nematodes. *PLoS Biology*, **12**, e1001915.

Turelli, M. & Moyle, L.C. (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*, **176**, 1059–1088.

Violle, C., Nemergut, D.R., Pu, Z. & Jiang, L. (2011) Phylogenetic limiting similarity and competitive exclusion. *Ecology Letters*, **14**, 782–787.

Yong, T.H. (2005) Prey capture by a generalist predator on flowering and nonflowering ambush sites: are inflorescences higher quality hunting sites? *Environmental Entomology*, **34**, 969–976.

Accepted 4 November 2016

Associate Editor: Rosa Menendez

