

# A framework for estimating the effects of sequential reproductive barriers: Implementation using Bayesian models with field data from cryptic species

Jean Peccoud,<sup>1,2,\*</sup>  David R. J. Pleydell,<sup>1,3,\*</sup>  and Nicolas Sauvion<sup>1,4</sup> 

<sup>1</sup>BGPI, Univ Montpellier, INRA, CIRAD, Montpellier SupAgro, Montpellier, France

<sup>2</sup>Current Address: Laboratoire Écologie et Biologie des Interactions, UMR 7267 CNRS, Université de Poitiers, 86000 Poitiers, France

<sup>3</sup>Current Address: UMR Animal, Santé, Territoires, Risques et Écosystèmes, INRA, CIRAD, Montpellier SupAgro, Université de Montpellier, Montpellier, France

<sup>4</sup>E-mail: nicolas.sauvion@inra.fr

Received July 20, 2013

Accepted August 27, 2018

Determining how reproductive barriers modulate gene flow between populations represents a major step toward understanding the factors shaping the course of speciation. Although many indices quantifying reproductive isolation (RI) have been proposed, they do not permit the quantification of cross-direction-specific RI under varying species frequencies and over arbitrary sequences of barriers. Furthermore, techniques quantifying associated uncertainties are lacking, and statistical methods unrelated to biological process are still preferred for obtaining confidence intervals and *P*-values. To address these shortcomings, we provide new RI indices that model changes in gene flow for both directions of hybridization, and we implement them in a Bayesian model. We use this model to quantify RI between two species of the psyllid *Cacopsylla pruni* based on field genotypic data for mating individuals, inseminated spermatophores and progeny. The results showed that preinsemination isolation was strong, mildly asymmetric, and indistinguishably different between study sites despite large differences in species frequencies; that postinsemination isolation strongly affected the more common hybrid type; and that cumulative isolation was close to complete. In the light of these results, we discuss how these developments can strengthen comparative RI studies.

**KEY WORDS:** Asymmetrical isolation, Bayesian modeling, finite mixture models, gene flow, hybridization, reproductive isolation.

Speciation involves the build-up of reproductive isolation (RI) at several key parts of the populations' life cycles, which are referred to as reproductive barriers. Understanding how these barriers act in conjunction to reduce gene flow and permit the divergence of populations into species has been an important goal of speciation research (Coyne and Orr 2004; Sobel et al. 2010; Butlin et al. 2012). As a result, the last 15 years have seen a burgeoning of methods to estimate the strength of reproductive barriers from field data on natural populations (Ramsey et al. 2003; Malausa

et al. 2005; Martin and Willis 2007; Matsubayashi and Katakura 2009; Sanchez-Guillen et al. 2012; Sobel and Streisfeld 2015; Pombi et al. 2017). Field estimates indeed provide the most pertinent results to help identify local factors affecting the course of speciation (Nosil et al. 2009; Via 2009; Sobel et al. 2010; Butlin et al. 2012).

This objective requires that RI estimates represent evolutionary relevant quantities, mainly potential gene flow, while correcting for differences in species frequencies that do not reflect

\*These authors contributed equally to this work.

phenotypic variations—effects we collectively refer to as “contingency.” In the progression toward this goal, many indices to quantify RI have been developed [reviewed in Sobel and Chen (2014)]. Key developments include formulas to quantify cumulative RI over sequential reproductive barriers (Ramsey et al. 2003); corrections for unequal species frequencies and allochory (Martin and Willis 2007); and the integration of these developments into RI indices that maintain a linear relation to the probability of gene flow—a desirable property when comparing populations and species (Sobel and Chen 2014).

Despite this growing diversity and sophistication of RI indices, and of the studies using them, two deficiencies of current methods remain apparent. First, although RI is commonly asymmetrical (e.g., Lowry et al. 2008; Matsubayashi and Katakura 2009; Sanchez-Guillen et al. 2012; Brys et al. 2014; Kaufmann et al. 2017; Martin et al. 2017), we lack indices that can estimate directional (cross-type-specific) RI over an arbitrary combination of reproductive barriers, while controlling for contingency. Second, speciation research would benefit from more studies reporting and discussing uncertainty in RI (e.g., Merrill et al. 2011; Lackey and Boughman 2017). The RI literature is dominated by the discussion of point estimates for which there exists a lack of associated uncertainty measures. Thus, it remains difficult to demonstrate whether apparent differences in RI (observed between different barriers, subpopulations or species) reflect real phenotypic differences or merely sampling error. Accordingly, practitioners seeking a richer statistical analysis—involving confidence intervals or significance tests for example—have been constrained to adopt less biologically motivated indices such as those provided by generalized linear models (Takami et al. 2007; Polacik and Reichard 2011; Falk et al. 2012; Peccoud et al. 2014; Kostyun and Moyle 2017).

Why a more complete statistical framework for RI estimation has not emerged may partly stem from the fact that the calculation of RI is frequently complexified by the need to correct for contingency and the effects of reproductive barriers not under scrutiny, or to combine the effects of several barriers (Sobel and Chen 2014). Accounting for asymmetry in RI would further complexify existing formulas and pose a substantial challenge regarding the construction of confidence intervals and significance tests for these indices.

We suggest that these issues can be resolved by focusing attention on estimating potential gene flow—rather than RI per se—induced by both within- and between-species crosses. Focusing on potential gene flow facilitates statistical estimation, from field data, of contingency-independent RI indices (in both cross directions) at any reproductive barrier or over any arbitrary sequence of barriers. Moreover, this approach naturally lends itself to Bayesian uncertainty analysis. In other branches of ecology and evolution, Bayesian techniques have long been popular for

numerous reasons, including: they provide a natural paradigm to account for multiple sources of uncertainty; they facilitate the incorporation of prior knowledge; they are applicable to a wide variety of models; and inference based on posterior distributions of model parameters is easy and intuitive (Gelman et al. 1995; Clark 2005; Cressie et al. 2009; Beaumont 2010; Hoban et al. 2012; Gompert et al. 2017). We illustrate these benefits with a Bayesian model designed to quantify the weight of evidence for spatial heterogeneity in RI using genetic data from natural populations of the psyllid *Cacopsylla pruni*.

## Methods

### MODELING SEQUENTIAL REPRODUCTIVE BARRIERS

Consider two species *A* and *B* interacting at a reproductive barrier. Let  $G_{XY}$  denote the probability that an individual sampled from the next generation comes from an  $X \times Y$  cross (with  $X, Y \in \{A, B\}$  and the maternal species always noted first) in the absence of further isolation after the barrier. Thus,  $G_{XY}$  is the potential gene flow (which we may simply call “gene flow” afterwards) induced by  $X \times Y$  crosses. Let  $\mathbf{G} = \{G_{AA}, G_{BB}, G_{AB}, G_{BA}\}$  be the set of all such proportions, which sum to one.

Estimating RI as the decrease of interspecific gene flow (Sobel and Chen 2014) requires a measure of gene flow that solely depends on phenotypic differences between species, not on their local relative frequencies, which may be affected by contingency. We call these contingency-independent gene flow rates “barrier porosities” to convey that they solely depend on phenotypic differences expressed at the barrier. We denote barrier porosities as  $\beta = \{\beta_{AA}, \beta_{BB}, \beta_{AB}, \beta_{BA}\}$ , these sum to one and each  $\beta_{XY}$  element equals 1/4 in the absence of RI.

The ratio of porosity, over its null expectation when  $RI = 0$ , indicates the strength of RI at a barrier (Sobel and Chen 2014). Thus, a bidirectional RI index that considers both hybrid cross-types is obtained by taking porosity to hybridization in both directions,  $\beta_{AB} + \beta_{BA}$ , divided by its expected value in the absence of RI (1/2):

$$RI = 1 - 2(\beta_{AB} + \beta_{BA}). \quad (1)$$

Similarly, RI affecting just one hybrid cross-type ( $X \neq Y$ ) is:

$$RI_{XY} = 1 - 4\beta_{XY}. \quad (2)$$

These indices vary linearly with gene flow, take value zero when porosities are 1/4, and take value one when porosity to hybridization is zero.

Directional RI indices allow between cross-type differences (asymmetry) in RI to be quantified as:

$$\Delta = RI_{AB} - RI_{BA}.$$

Given the simplicity of these developments, the main modeling task is to establish the relationships between gene flow  $G$  and barrier porosities  $\beta$ . To this aim, we introduce “null gene flow”  $E_0[G_{XY}]$  to denote potential gene flow in the absence of RI at the studied barrier.  $E_0[G_{XY}]$  can be visualized as the flow of genes going through the previous barrier and arriving at the focal barrier. Table 1 provides examples of  $E_0[G_{XY}]$  and  $G_{XY}$  for different sources of RI.

Postzygotic isolation is modeled as follows. Potential gene flow  $G_{XY}$  is the relative frequency of  $XY$  genotypes in the progeny, after the studied barrier. It is proportional to their frequency before the barrier,  $E_0[G_{XY}]$ , multiplied by their probability to survive (or pass) through the barrier (defined as  $S_{XY}$ ):

$$G_{XY} \propto E_0[G_{XY}] S_{XY}. \quad (3)$$

Under equal species frequencies and random mating,  $G_{XY} = \beta_{XY}$  by definition and  $E_0[G_{XY}] = 1/4$ . Assuming that survival  $S_{XY}$  is constant, and is therefore independent of species frequencies, the above equation translates to:

$$\beta_{XY} \propto S_{XY}$$

where  $1/4$  was dropped as a constant in the proportionality relationship.

Substituting  $S_{XY}$  in equation (3) yields:

$$G_{XY} \propto E_0[G_{XY}] \beta_{XY}. \quad (4)$$

Postzygotic barrier porosities are therefore proportional to the ratio of potential over null gene flow—a metric that enables RI quantification when null expectations (elements of  $E_0[G]$ ) are unequal [see appendix D of Sobel and Chen (2014)].

For total gene flow to equal one, equation (4) requires normalization:

$$G_{XY} = \frac{E_0[G_{XY}] \beta_{XY}}{\sum_{X \in \{A, B\}} \sum_{Y \in \{A, B\}} E_0[G_{XY}] \beta_{XY}}. \quad (5)$$

Equation (5) satisfies that when all null gene flows  $E_0[G_{XY}]$  equal  $1/4$  (equal species frequencies), the porosity  $\beta_{XY}$  equals gene flow  $G_{XY}$ . Conversely, in the absence of RI ( $\beta_{XY}$  is  $1/4$  for all combinations), gene flows equal null gene flows.

Given estimates of gene flow before and after a barrier, the porosities can be recovered by rearranging and normalizing equation (4) so that elements of  $\beta$  sum to 1, hence:

$$\beta_{XY} = \frac{G_{XY}/E_0[G_{XY}]}{\sum_{X \in \{A, B\}} \sum_{Y \in \{A, B\}} (G_{XY}/E_0[G_{XY}])}. \quad (6)$$

In contrast to postzygotic barriers that increase progeny mortality, prezygotic barriers do not usually incur a fitness cost to parents. Hence, prezygotic isolation must be modeled in such a

**Table 1.** Formulations for null (in the absence of RI at a studied barrier) gene flow due to females of species  $X$  and males of species  $Y$ , ( $E_0[G_{XY}]$ ), and for potential gene flow following modification by the barrier ( $G_{XY}$ ).

Source of RI	$E_0[G_{XY}]$	$G_{XY}$
Ecological niche difference	$f_X \times m_Y^\dagger$	$p.^\ddagger$ of encounters involving females of species $X$ and males of species $Y$ (a)
Allochrony	(a) <sup>§</sup>	$p.$ of encounters between sexually active mates involving females of species $X$ and males of species $Y$ (b)
Conspecific mate preference	(b)	$p.$ of mating pairs involving females of species $X$ and males of species $Y$ (c)
Mechanical incompatibility	(c)	$f_X \times$ (average proportion of species $Y$ sperm per female of species $X$ ) (d)
Gametic incompatibility	(d)	$p.$ of zygotes from females of species $X$ and males of species $Y$ (e)
Early hybrid mortality	(e)	$p.$ of progeny from females of species $X$ and males of species $Y$ (f)
Late hybrid mortality	(f)	$p.$ of older progeny from females of species $X$ and males of species $Y$

Barriers are shown in their order of appearance in the life cycle. <sup>†</sup> $f_X$ : proportion of species  $X$  in females;  $m_Y$ : proportion of species  $Y$  in males. <sup>‡</sup> $p.$ : proportion. <sup>§</sup>Letters in parentheses correspond to formulations proposed for  $G_{XY}$  in the rightmost column. Formulations may be amended depending on the specifics of the biological model. For instance individuals may be interpreted as gametes (e.g., pollen for males), and proportion of encounters can be estimated from range or habitat overlap.

way that it does not directly affect fitness. To do so, we express the proportion of  $XY$  zygotes (that we expect if no isolation exists after the studied barrier) among those having a mother from species  $X$ :

$$\frac{G_{XY}}{G_{XA} + G_{XB}} = \frac{E_0[G_{XY}]\beta_{XY}}{E_0[G_{XA}]\beta_{XA} + E_0[G_{XB}]\beta_{XB}}. \quad (7)$$

To derive  $G_{XY}$ , the relative contribution of species  $X$  females to the next generation ( $G_{XA} + G_{XB}$ ) must be specified. If, at the focal barrier, female reproductive success does not vary between the species, then  $G_{XA} + G_{XB}$  is the frequency of species  $X$  in females, which we call  $f_X$ , and equation (7) becomes:

$$G_{XY} = f_X \frac{E_0[G_{XY}]\beta_{XY}}{E_0[G_{XA}]\beta_{XA} + E_0[G_{XB}]\beta_{XB}}. \quad (8)$$

Barrier porosities can be obtained from gene flows by rearranging equation (7) into (proof not shown):

$$\frac{\beta_{XY}}{\beta_{XA} + \beta_{XB}} = \frac{G_{XY}/E_0[G_{XY}]}{G_{XA}/E_0[G_{XA}] + G_{XB}/E_0[G_{XB}]}, \quad (9)$$

and specifying  $\beta_{XA} + \beta_{XB}$  appropriately. If female reproductive success can be assumed equal between species, then  $\beta_{XA} + \beta_{XB} = 1/2$ , so:

$$\beta_{XY} = \frac{1}{2} \frac{G_{XY}/E_0[G_{XY}]}{G_{XA}/E_0[G_{XA}] + G_{XB}/E_0[G_{XB}]}. \quad (10)$$

If the assumption of equal reproductive success between females is not warranted, alternative formulations for equations (8) and (10) may be desirable. Such developments should be tailored to the specifics of the biological system and are beyond the scope of the current work.

Note that equation (7) and its by-products assume that the fraction of  $XY$  zygotes contributed by females of species  $X$  is proportional to its null-expected value in the absence of RI,  $E_0[G_{XY}]$ . This implies that the probability of hybridization per interspecific encounter, represented by the ratio  $G_{XY}/E_0[G_{XY}]$ , does not depend on species frequencies, and reflects the barrier porosity  $\beta_{XY}$ .

Once obtained, barrier porosities can be used to model a sequence of  $b$  barriers with porosities  $\beta^1 \dots \beta^b$ . The product of these porosities is proportional to the probability that genes from two parents flow through all these barriers to eventually produce an offspring. The combined porosity of these barriers to  $X \times Y$  gene flow is thus given by:

$$\beta_{XY}^{1:b} = \frac{\prod_{i=1}^b \beta_{XY}^i}{\sum_{X \in \{A, B\}} \sum_{Y \in \{A, B\}} \prod_{i=1}^b \beta_{XY}^i}, \quad (11)$$

whose denominator ensures that the combined porosities of all four  $XY$  combinations sum to one.

Equations (5) and (8) permit barrier porosities  $\beta_{XY}$ , hence RI, to be estimated via statistical techniques that confront modeled gene flows  $G_{XY}$  with data collected at different points of

the reproductive cycle. We will demonstrate this approach with a Bayesian implementation. Alternatively, a simpler approach would use equations (6) or (10) to obtain point estimates of barrier porosities by specifying  $G_{XY}$  according to observations (examples given in Table 1). RI is then derived from porosities with equations (1) or (2).

## STUDY MODEL

Our model system, *Cacopsylla pruni* Scopoli (Sternorrhyncha: Psyllidae), includes two unnamed cryptic species that are strongly genetically divergent but have yet to show ecological or morphological differences (Sauvion et al. 2007; Peccoud et al. 2013). These species co-occur at several sites in Southern France (Sauvion et al. 2007) on shrubs of genus *Prunus*, on which the insects feed, reproduce, and die in spring (Fig. 1). Progeny reach adulthood after approximately 2 months, migrate shortly after to conifers for overwintering and return to *Prunus* in early spring to mate.

Rearing *C. pruni* in controlled conditions has proven extremely difficult (Jarausch and Jarausch 2016). However, the nonoverlapping generations and co-occurrence of the *C. pruni* species at several sites make them good candidates for field-based estimates of RI within their life-cycle (Fig. 1). To this aim, we genotyped mating adults, inseminated spermatophores, and progeny as species *A* or *B* or as hybrids.

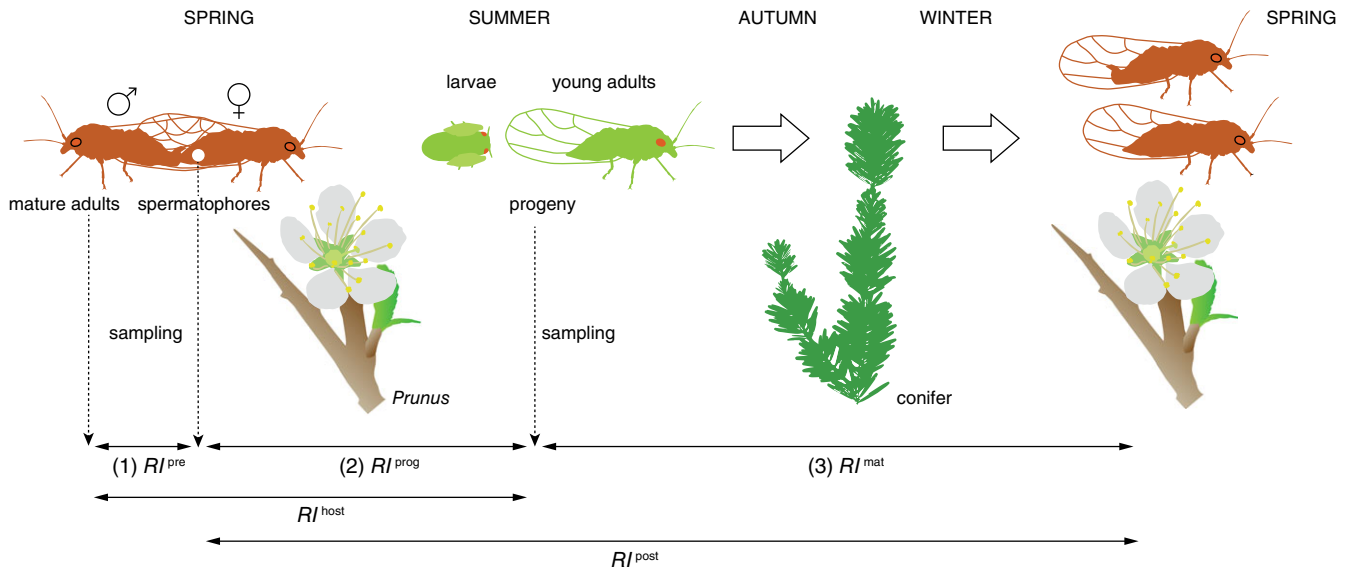
## SAMPLE COLLECTION AND SPECIES ASSIGNMENT

Psyllids were collected in spring 2010 on *Prunus* in southern France at three sites: near Tautavel (42°47'38N, 2°41'56E), Grabels (43°39'35N, 3°49'12E), and Bompas (42°43'43N, 2°56'31E). We also used collections obtained in spring 2008 near Torrealles (42°44'29N, 2°59'6E). Each sampling site consisted of a small number of bushes or hedges of *Prunus* and covered a few dozen meters at most.

Mature adults were sampled at all sites. Progeny (larvae and young adults of the subsequent generation) were sampled at Tautavel and Grabels. Psyllids that fell from beaten branches onto flat nets were either stored in ethanol and/or frozen. Mating pairs caught on nets were stored in separate tubes.

Purification of DNA followed Peccoud et al. (2013). Abdomens of mature females were retained and softened in 70% ethanol for spermatophore extraction. Spermatophores were identified as glossy white pellets in spermatheca under a stereo microscope and transferred separately to DNA purification wells. To minimize between-species DNA contamination risk, each batch of dissections, DNA purifications and amplifications of spermatophore DNA was performed on females of the same species.

Each DNA sample was assigned to species *A* or *B* using a single diagnostic PCR of the Internal Transcribed Spacer 2 (*ITS2*) gene, which yields an amplicon of a specific size for a



**Figure 1.** Life cycle of *Cacopsylla pruni* and sampling used to estimate reproductive isolation (RI) between its cryptic species at various barriers, or combinations of barriers. Barriers are shown as horizontal arrows and their effects are estimated with RI indices defined in the main text.

given species (Peccoud et al. 2013). Individuals showing two bands—putative hybrids—were reprocessed through DNA extraction (reusing their carcasses after washing in water) and PCR to minimize the risk of DNA contamination being interpreted as hybridization. To identify the maternal species of each confirmed hybrid, we Sanger-sequenced a mitochondrial region encompassing the *COI* gene (sequences are available under Genbank under accession numbers MH577765 to MH577790). Supporting text I details purification, primers, the genotyping of hybrids and possible sources of error.

#### MODELING REPRODUCTIVE ISOLATION IN *CACOPSYLLA PRUNI*

The genotype data of spermatophores, progeny and mature adults allowed estimation of RI arising: (1) between colonization of *Prunus* by mature adults and insemination, (2) between insemination and the sampling of progeny on *Prunus*, (3) following the sampling of progeny, overwintering on conifers and return of mature adults on *Prunus*. Indices related to (1), (2), and (3) are given superscripts “pre” (for *pre-insemination*), “prog” (*progeny*), and “mat” (*mature adults*), respectively (Fig. 1).

Gene flow due to inseminations at sampling site  $i$ ,  $G_{iXY}^{\text{pre}}$ , reflects the proportion of spermatophores of species  $Y$  per female of species  $X$  (Table 1). All dissected females were inseminated, thus there was no evidence that female mating success differed between species. Accordingly, we used equation (8) to model  $G_{iXY}^{\text{pre}}$ . At this stage, gene flow before the barrier reflects species proportions among mating adults of each sex, thus:

$$E_0[G_{iXY}^{\text{pre}}] = f_{iX}m_{iY},$$

where  $f_{iX}$  and  $m_{iY}$  are the proportions of species  $X$  and species  $Y$  among mature females and males of site  $i$ , respectively. Thus, equation (8) becomes:

$$G_{iXY}^{\text{pre}} = \frac{f_{iX}m_{iY}\beta_{iXY}}{m_{iA}\beta_{iXA} + m_{iB}\beta_{iXB}}.$$

At a subsequent barrier, null gene flow is gene flow through the previous barrier. Thus, from equation (5), gene flows at the two postzygotic barriers are:

$$G_{iXY}^{\text{prog}} = \frac{G_{iXY}^{\text{pre}}\beta_{iXY}^{\text{prog}}}{\sum_{X \in \{A,B\}} \sum_{Y \in \{A,B\}} (\beta_{iXY}^{\text{prog}} G_{iXY}^{\text{pre}})},$$

$$G_{iXY}^{\text{mat}} = \frac{G_{iXY}^{\text{prog}}\beta_{iXY}^{\text{mat}}}{\sum_{X \in \{A,B\}} \sum_{Y \in \{A,B\}} (\beta_{iXY}^{\text{mat}} G_{iXY}^{\text{prog}})}.$$

$G_{iXY}^{\text{prog}}$  reflects the proportion of  $XY$  genotypes among progeny, and  $G_{iXY}^{\text{mat}}$  reflects the proportion of  $XY$  genotypes among returning adults. With these specifications, porosities  $\beta^{\text{pre}}$ ,  $\beta^{\text{prog}}$ , and  $\beta^{\text{mat}}$  can be estimated from genotype data using Bayesian modelling (below).

Equation (11) was used to compute the combined porosities of successive barriers (Fig. 1). Thus, we quantified  $\beta^{\text{host}}$  and  $RI^{\text{host}}$ , representing RI on spring host-plants [barriers (1)+(2)];  $\beta^{\text{post}}$  and  $RI^{\text{post}}$ , representing global postinsemination RI [barriers (2)+(3)]; and  $\beta^{\text{total}}$  and  $RI^{\text{total}}$  for barriers (1)+(2)+(3). The “absolute contribution” of barriers (Ramsey et al. 2003; Sobel and Chen 2014) was quantified as the difference in cumulative RI either side of each barrier (Ramsey et al. 2003).



## SPATIAL HETEROGENEITY IN PREINSEMINATION ISOLATION

To determine the degree of spatial heterogeneity in preinsemination RI,  $\beta^{\text{pre}}$ , we incorporated a finite mixture model (FMM) (McLachlan and Peel 2000) as a parsimonious model of hidden spatial structure (Pleydell and Chretien 2010). This FMM allocates each study site to one of  $k \in \{1 \dots n\}$  “site groups,” where each site in a group shares identical porosities  $\beta^{\text{pre}}$  and  $n$  is the number of study sites. This introduces vectors  $\mathbf{z}$ , which allocates sites to groups,  $\mathbf{w}$ , which weights the importance of groups and  $\mathbf{k}$ , an indicator vector that activates/disactivates groups (see supporting text II).

## BAYESIAN INFERENCE

Bayesian analysis of RI in *C. pruni* required making inference from the posterior distribution:

$$f(\beta^{\text{pre}}, \beta^{\text{prog}}, \beta^{\text{mat}}, \mathbf{f}, \mathbf{m}, \mathbf{k}, \mathbf{z}, \mathbf{w} | \mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y}, p^{\text{Obs}}, p^{\text{Mis}}),$$

with new terms defined below. Uninformative priors were adopted for all parameters.

Likelihood functions for species frequencies among sexes  $\mathbf{m}$  and  $\mathbf{f}$  were obtained assuming the number of individuals of each species among sampled males  $\mathbf{u}$  and females  $\mathbf{v}$  follow multinomial distributions with probabilities  $\mathbf{m}$  and  $\mathbf{f}$ , respectively.

At barrier (1), the likelihood was evaluated using genotype data for inseminated females  $\mathbf{x}$  and spermatophores  $\mathbf{y}$ . The numbers of species *A* and *B* spermatophores extracted from a female of species *X* were assumed to follow a multinomial distribution with probabilities  $G_{XA}/(G_{XA} + G_{XB})$  and  $G_{XB}/(G_{XA} + G_{XB})$ , respectively. This assumes independent inseminations—indeed males of the related species *C. pyricola* inseminate one spermatophore per female (Burts and Fisher 1967; Krysan 1990).

At barrier (2), counts of the four genotypes among progeny,  $p^{\text{Obs}}$ , were assumed to follow a multinomial distribution with probabilities  $G^{\text{prog}}$ . The likelihood also accounted for two hybrids (from Tautavel) of unknown maternal ancestry,  $p^{\text{Mis}}$ , that failed to amplify at the mitochondrial region (see supporting text I).

The likelihood at barrier (3) was derived similarly to that of barrier (2), from genotype data  $\mathbf{u}$  and  $\mathbf{v}$ , reflecting possible between-year differences in genotype frequency. Further model details, and a glossary defining all variables, are provided in supporting text II.

The posterior distribution was sampled using Markov chain Monte Carlo (MCMC) (Gelman et al. 1995). Site-group activation indicators,  $\mathbf{k}$ , were sampled using Reversible Jump MCMC (Green 1995). The model and MCMC algorithm were written and executed in NIMBLE 6.10 (de Valpine et al. 2017) within R 3.4.1 (R Development Core Team 2017). One hundred MCMC chains of 600,000 iterations were run, the first 100,000 iterations were removed as burn-in and samples were saved each 50 iterations. Concatenated output ( $10^6$  samples in total) was analyzed using R package CODA (Plummer et al. 2006). Source code and data are available at [https://bitbucket.org/DRJP/reproductive\\_isolation\\_mcmc/](https://bitbucket.org/DRJP/reproductive_isolation_mcmc/).

## Results

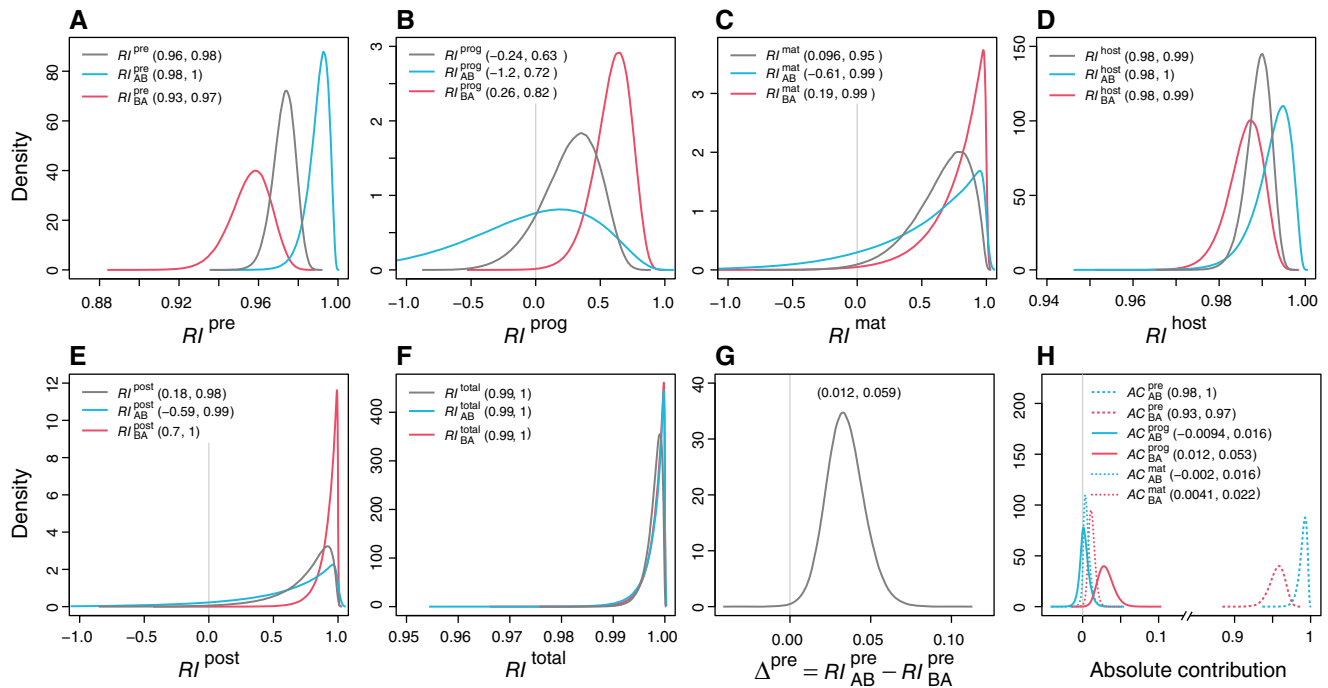
Table 2 summarizes the genotypic data and shows large differences in species frequencies across sites.

We did not model premating isolation as only 46 mating pairs were caught on sampling nets, all at Tautavel. Thirty-five involved individuals of species *A*, and 11 involved individuals of species *B*. No heterospecific pairs were found. Species proportions in mating pairs were indistinguishable from those in mature adults ( $\chi^2 = 0.045$ ,  $P = 0.84$ , 1 d.f.) but differed significantly from those

**Table 2.** Genotype data (assignment to *Cacopsylla pruni* species *A* or *B*) from mature adults, spermatophores, and progeny at the four sampled sites.

	Tautavel		Grabels		Bompas		Torreilles	
	<i>A</i>	<i>B</i>	<i>A</i>	<i>B</i>	<i>A</i>	<i>B</i>	<i>A</i>	<i>B</i>
Females	307	87	127	41	30	31	60	146
Males	211	57	43	18	27	23	26	64
Intra-inseminated females	194	76	110	38	30	30	53	141
Cross-inseminated females	1	11	1	3	0	1	0	5
Spermatophores in <i>A</i> females	826	1	249	1	80	0	83	0
Spermatophores in <i>B</i> females	14	202	3	65	1	70	5	212
Progeny from <i>A</i> mothers	1935	3	382	0				
Progeny from <i>B</i> mothers	10	424	2	142				

For logistic reasons, not all genotyped females were dissected. “Cross-inseminated” refers to females carrying at least one allospecific spermatophore. For the progeny column headings, *A* or *B*, indicate the paternal species.



**Figure 2.** Posterior probability distributions of reproductive isolation (RI) between *Cacopsylla pruni* species measured at three reproductive barriers (panels A, B, C) and their combinations (panels D, E, F); (G) asymmetry in preinsemination RI; (H) absolute contributions of reproductive barriers to overall RI. Ninety-five percent credibility intervals are shown in parentheses. See Figure 1 for a representation of the different forms of RI measured in *C. pruni*.

expected under random mating ( $\chi^2 = 40.7$ ,  $P < 0.001$ , 1 d.f.). The sampling time of 41 mating pairs sampled over the course of a single day showed little difference between species (Mann–Whitney = 244,  $P > 0.19$ ), providing no evidence for differences in timing of mating activities.

Most spermatophores (1812 of the 1990 extracted) were successfully genotyped (missing data is discussed in supporting text I). Interspecific inseminations were detected at all sites (Table 1) and involved 1.38% of genotyped spermatophores. This indicates strong but incomplete preinsemination isolation ( $RI^{pre}$ , Fig. 2A). The proportion of MCMC samples in which  $RI^{pre}$  differed between sites was  $\sim 0.001$ , providing only negligible evidence for between-site variation. In terms of asymmetry,  $RI^{pre}$  was stronger in  $A \times B$  crosses than in the opposite direction,  $\Delta^{pre}$  being positive (Fig. 2G). Other barriers showed little evidence for asymmetry, as posterior distributions of directional RI indices for reciprocal crosses largely overlapped ( $\Delta$  not shown).

Results support positive postinsemination isolation against  $B \times A$  hybrids of the progeny ( $RI_{BA}^{prog}$ , Fig. 2B), meaning these hybrids were less frequent than expected from cross-inseminations. The absence of hybrid genotypes in mature adults (Table 2) rendered  $RI^{mat}$  positive for  $B \times A$  hybrids (Fig. 2C), indicating mortality between emigration from spring hosts and return to these hosts the subsequent year. For the opposite cross direction, uncertainty was large, due to the strong isolation against

$A \times B$  insemination and subsequent low expected frequency of  $A \times B$  hybrids.

The combinations of these successive reproductive barriers led to strong  $RI^{host}$ ,  $RI^{post}$  and essentially complete overall RI (Fig. 2D–F). Preinsemination barriers contributed by far the most to overall RI, as shown by the high absolute contribution  $AC^{pre}$  (Fig. 2H).

## Discussion

### BENEFITS AND ASSUMPTIONS OF THE APPROACH

This work introduces the notion of barrier porosities, which represent contingency-independent probabilities of gene flow, to facilitate RI estimation. Our formulations extend the current RI quantification framework (Sobel and Chen 2014) in several ways.

First, they standardize the construction of RI indices for any type of barrier by modelling null gene flows  $E_0[G]$  and potential gene flows  $G$  (Table 1) at each barrier.

In addition, by explicitly considering all four cross-types (within and between species), this approach leads naturally to the construction of directional RI indices ( $RI_{XY}$ , eq. (2)). These indices share the properties of Sobel and Chen’s (2014) bidirectional  $RI$  and satisfy the growing interest in measuring asymmetry in  $RI$  (Lowry et al. 2008; Matsubayashi and Katakura 2009; Sanchez-Guillen et al. 2012; Yukilevich 2012; Brys et al. 2014). A notable

difference between  $RI_{XY}$  (eq. (2)) and the bidirectional RI indices of Sobel and Chen (2014) and our equation (1) is the lower bound of  $RI_{XY}$  at  $-3$  rather than  $-1$ , which makes it asymmetric around 0. However, a value of  $-3$  accurately informs that directional gene flow is 300% higher than expected under random mating (1/4). It therefore seems sensible to prioritize linearity with gene flow (barrier porosity) of the RI index over symmetry around 0.

Our formulation also simplifies the quantification of cumulative effects of sequential and potentially asymmetrical barriers on RI—it is sufficient to compute a normalized product of sequential porosities estimated separately (eq. (11)). This reduces the difficulty of formulating RI over sequential barriers, where exigent checking for correctness in respect to a particular combination of barriers is typically required (Sobel and Chen 2014). Because barrier porosities are probabilities, and are designed to be contingency-independent, they can be combined (multiplied) for any sequence of barriers studied by any method ranging from field surveys to laboratory experiments.

Finally, because our porosity-centered specification permits comparison of modeled and observed gene flow, it readily accommodates Bayesian inference and hence credibility intervals for RI-related indices. The potential of Bayesian modeling is demonstrated here with a finite mixture model designed to detect spatial heterogeneity in RI. These developments can help identify local factors conditioning RI. In particular, were RI to vary according to species frequencies, one may question the two main assumptions underlying frequency-independent RI indices: (i) hybrid survival rates are unaffected by genotype frequencies in the progeny (implied by eq. (4)), neglecting possible effects of competition on hybrids, and (ii) the risk of hybridization per interspecific encounter is stable (eq. (7)). While we could not evaluate the former assumption due to uneven sampling of progeny and the scarcity of hybrids, the latter is discussed in the next section.

## INTENSITY AND CONTRIBUTION OF REPRODUCTIVE BARRIERS

The Bayesian model used to analyze genotypic data from *C. pruni* populations demonstrated high, asymmetrical preinsemination isolation ( $RI^{\text{pre}}$ ) with little evidence for between-site variation, and positive postinsemination isolation ( $RI^{\text{prog}}$  and  $RI^{\text{mat}}$ ) against  $B \times A$  crosses (Fig. 2). The combination of these barriers results in essentially complete RI in both directions.

Preinsemination isolation is dominated by premating isolation, given the absence of heterospecific mating pairs among the 46 collected. Conspecific mate preference could be mediated by olfaction (Soroker et al. 2004; Wenninger et al. 2008; Guedot et al. 2009) and/or acoustic signals (Percy et al. 2006; Tishechkin 2007; Wenninger et al. 2009), both of which contribute to species recognition and mate attraction in other psyllid species. Mechanical isolation (Sota and Kubota 1998; Holwell et al. 2010) ap-

pears unlikely, as variation in male genitalia morphology could not be detected by optical and electron microscopy (N. Sauvion, unpubl. data). The same can be said for temporal isolation [reviewed in Taylor and Friesen (2017)], as the timing of mating did not significantly differ between species according to mating pairs caught within the course of a day. At larger temporal scales, synchrony between reproductive cycles is supported by the similar species proportions across larval stages at Tautavel ( $\chi^2 = 2.0556$ ,  $P > 0.35$ , 2 d.f.).

We detected that preinsemination RI significantly differs according to the direction of crosses (Fig. 2A, G), suggesting that *B* females and/or *A* males are on average less discriminatory than their allospecific counterparts in respect to species recognition. Asymmetric prezygotic isolation is frequently observed in mate preference assays (Jaenike et al. 2006; Rafferty and Boughman 2006; Takami et al. 2007; Dopman et al. 2010; Raychoudhury et al. 2010; Merrill et al. 2011; Veen et al. 2011; Sanchez-Guillen et al. 2012), but is only rarely measured in the field (Bournez et al. 2015). In comparison to laboratory studies, the asymmetry we observed involves much higher levels of RI (Fig. 2A). This suggests that asymmetry in prezygotic isolation can persist late in the speciation process, as does prezygotic RI in general (e.g., Coyne and Orr 1997; Mallet et al. 2007; Merrill et al. 2011), and/or that premating RI can be higher in the field than in the laboratory (Jennings and Etges 2010).

Interestingly, we found no convincing evidence that preinsemination RI varied between the four sampling sites, despite large differences in relative species frequencies (Table 1). Hence, the assumption of a stable hybridization risk per interspecific encounter (implied by eq. (7)) is not called into question. Keeping in mind that our ability to challenge this assumption is limited by the number of sampling sites, our observations inform us on the bases of incomplete prezygotic RI in *C. pruni*. A stable risk of mating per interspecific encounter may indicate a certain degree of conspecific mate preference that is both relatively insensitive to site-specific factors, and similar among individuals of the same species and sex (e.g., Merrill et al. 2011). The hypothesis of between-individual variation in mate preference, potentially due to polymorphism at underlying loci, would less parsimoniously explain incomplete RI. Indeed, it would not explain why cross-inseminations appear more frequent at species *A*-rich sites, unless the less discriminatory individuals essentially occurred among males of species *A*. This hypothesis thus also requires that mate choice be mostly exercised by males. Mate-choice experiments would help to evaluate these hypotheses.

The predominant contribution of prezygotic barriers to overall RI (Fig. 2H) naturally follows from their early occurrence in the species life-cycle and has been reported in various sympatric species (Ramsey et al. 2003; Malausa et al. 2005; Kay and Husband 2006; Lowry et al. 2008; Matsubayashi and Katakura



2009; Sanchez-Guillen et al. 2012). This predominance does not indicate that postinsemination barriers contributed little to the divergence of *A* and *B* species. These barriers may have reinforced premating isolation [see Servedio and Noor (2003); Coyne and Orr (2004) for a review on reinforcement] and have certainly permitted genetic divergence between *C. pruni* species (Sauvion et al. 2007; Peccoud et al. 2013) in the face of cross-insemination. Some barriers ( $Rf^{prog}$ , Fig. 1) operate between insemination and progeny growth, at least for *B* × *A* crosses (Fig. 2B), and others ( $Rf^{mat}$ , Fig. 2C) affect survival of grown hybrids up to their return on *Prunus* shrubs. In terms of causes,  $Rf^{prog}$  possibly reflects low sperm efficacy in allospecific females (e.g., Matute 2010) and/or reduced hybrid survival up to sampling.

Although the scarcity of hybridization in *C. pruni* limits the precision of certain estimates, our case study illustrates how the proposed framework provides estimates of reproductive barriers at an arbitrary number of sampling points through the species life cycle. Future models could incorporate refinements such as independent development and/or survival rates for each sex and developmental stage, more sophisticated models of spatio-temporal variation, or other sources of prior information.

#### AUTHOR CONTRIBUTIONS

J.P. and N.S. initiated the study and obtained biological data. J.P. and D.R.J.P. developed the porosity-based approach. D.R.J.P. conceived the Bayesian implementation and code. J.P., D.R.J.P., and N.S. wrote the manuscript.

#### ACKNOWLEDGMENTS

We thank René Rieu for his advice on spermatophore extraction, and Gaël Thébaud, Gérard Labonne, and François Bonnot for helpful comments on the study and drafts of this article. We thank Josiane Peyre and Patrick Limon for their help in genotyping. This work utilized computing resources from INRA's MIGALE cluster (<http://migale.jouy.inra.fr>) and benefitted from the assistance of Eric Montaudon and Véronique Martin. Part of this work was funded by INRA grant SDIPS (Speciation and molecular Diagnosis of Insect Pest Species complexes).

#### DATA ARCHIVING

Mitochondrial sequence data are deposited at Genbank under accession numbers MH577765 to MH577790, source code and raw data are available at [https://bitbucket.org/DRJP/reproductive\\_isolation\\_mcmc/](https://bitbucket.org/DRJP/reproductive_isolation_mcmc/).

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### LITERATURE CITED

Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* 41:379–406.  
 Bournez, L., N. Cangi, R. Lancelot, D. R. J. Pleydell, F. Stachurski, J. Bouyer, D. Martinez, T. Lefrançois, L. Neves, and J. Pradel. 2015. Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique. *Parasites Vectors* 8:504.

Brys, R., A. Vanden Broeck, J. Mergeay, and H. Jacquemyn. 2014. The contribution of mating system variation to reproductive isolation in two closely related *Centaureum* species (Gentianaceae) with a generalized flower morphology. *Evolution* 68:1281–1293.  
 Burts, E. C., and W. R. Fisher. 1967. Mating behavior, egg production and egg fertility in the pear psylla. *J. Econ. Entomol.* 60:1297–1300.  
 Butlin, R., A. Debelle, C. Kerth, R. R. Snook, L. W. Beukeboom, R. F. C. Cajas, W. Diao, M. E. Maan, S. Paolucci, F. J. Weissing, et al. 2012. What do we need to know about speciation? *Trends Ecol. Evol.* 27:27–39.  
 Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecol. Lett.* 8:2–14.  
 Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. *Evolution* 51:295–303.  
 ———. 2004. *Speciation*. Sinauer, Sunderland, USA.  
 Cressie, N., C. A. Calder, J. S. Clark, J. M. V. Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecol. Appl.* 19:553–570.  
 de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2017. Programming with models: writing statistical algorithms for general model structures with NIMBLE. *J. Comput. Graphical Stat.* 26:403–413.  
 Dopman, E. B., P. S. Robbins, and A. Seaman. 2010. Components of reproductive isolation between north american pheromone strains of the European corn borer. *Evolution* 64:881–902.  
 Falk, J. J., C. E. Parent, D. Agashe, and D. I. Bolnick. 2012. Drift and selection entwined: asymmetric reproductive isolation in an experimental niche shift. *Evol. Ecol. Res.* 14:403–423.  
 Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian data analysis*. Chapman and Hall, London.  
 Gompert, Z., E. G. Mandeville, and C. A. Buerkle. 2017. Analysis of population genomic data from hybrid zones. *Annu. Rev. Ecol. Evol. Syst.* 48:207–229.  
 Green, P. J. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82:711–732.  
 Guedot, C., J. G. Millar, D. R. Horton, and P. J. Landolt. 2009. Identification of a sex attractant pheromone for male winterform pear psylla, *Cacopsylla pyricola*. *J. Chem. Ecol.* 35:1437–1447.  
 Hoban, S., G. Bertorelle, and O. E. Gaggiotti. 2012. Computer simulations: tools for population and evolutionary genetics. *Nat. Rev. Genet.* 13:110–122.  
 Holwell, G. I., C. Winnick, T. Tregenza, and M. E. Herberstein. 2010. Genital shape correlates with sperm transfer success in the praying mantis *Ciulfina klassi* (Insecta: Mantodea). *Behav. Ecol. Sociobiol.* 64:617–625.  
 Jaenike, J., K. A. Dyer, C. Cornish, and M. S. Minhas. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* 4:1852–1862.  
 Jarausch, W., and B. Jarausch. 2016. A permanent rearing system for *Cacopsylla pruni*, the vector of 'Candidatus Phytoplasma prunorum'. *Entomol. Exp. Appl.* 159:112–116.  
 Jennings, J. H., and W. J. Etges. 2010. Species hybrids in the laboratory but not in nature: a reanalysis of premating isolation between *Drosophila arizonae* and *D. mojavensis*. *Evolution* 64:587–598.  
 Kaufmann, J., T. L. Lenz, M. Kalbe, M. Milinski, and C. Eizaguirre. 2017. A field reciprocal transplant experiment reveals asymmetric costs of migration between lake and river ecotypes of three-spined sticklebacks (*Gasterosteus aculeatus*). *J. Evol. Biol.* 30:938–950.  
 Kay, K. M., and B. Husband. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. *Evolution* 60:538–552.  
 Kostyun, J. L., and L. C. Moyle. 2017. Multiple strong postmating and intrinsic postzygotic reproductive barriers isolate florally diverse species of *Jaltomata* (Solanaceae). *Evolution* 71:1556–1571.

- Krysan, J. L. 1990. Laboratory study of mating-behavior as related to diapause in overwintering *Cacopsylla pyricola* (Homoptera, Psyllidae). *Environ. Entomol.* 19:551–557.
- Lackey, A. C. R., and J. W. Boughman. 2017. Evolution of reproductive isolation in stickleback fish. *Evolution* 71:357–372.
- Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos. Trans. R Soc. Lond. Ser. B Biol. Sci.* 363:3009–3021.
- Malausa, T., M. T. Bethenod, A. Bontemps, D. Bourguet, J. M. Cornuet, and S. Ponsard. 2005. Assortative mating in sympatric host races of the European corn borer. *Science* 308:258–260.
- Mallet, J., M. Beltran, W. Neukirchen, and M. Linares. 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evol. Biol.* 7:1–16.
- Martin, H., P. Touzet, M. Dufay, C. Gode, E. Schmitt, E. Lahiani, L. F. Delph, and F. Van Rossum. 2017. Lineages of *Silene nutans* developed rapid, strong, asymmetric postzygotic reproductive isolation in allopatry. *Evolution* 71:1519–1531.
- Martin, N. H., and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61:68–82.
- Matsubayashi, K. W., and H. Katakura. 2009. Contribution of multiple isolating barriers to reproductive isolation between a pair of phytophagous ladybird beetles. *Evolution* 63:2563–2580.
- Matute, D. R. 2010. Reinforcement of Gametic Isolation in *Drosophila*. *PLoS Biol.* 8:e1000341.
- McLachlan, G., and D. Peel. 2000. *Finite mixture models*. John Wiley & Sons, Inc, Hoboken, USA.
- Merrill, R. M., Z. Gompert, L. M. Dembeck, M. R. Kronforst, W. O. McMillan, and C. D. Jiggins. 2011. Mate preference across the speciation continuum in a clade of mimetic butterflies. *Evolution* 65:1489–1500.
- Nosil, P., L. J. Harmon, and O. Seehausen. 2009. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* 24:145–156.
- Peccoud, J., M. de la Huerta, J. Bonhomme, C. Laurence, Y. Outreman, C. M. Smadja, and J. C. Simon. 2014. Widespread host-dependent hybrid unfitness in the pea aphid species complex. *Evolution* 68:2983–2995.
- Peccoud, J., G. Labonne, and N. Sauvion. 2013. Molecular test to assign individuals within the *Cacopsylla pruni* complex. *Plos One* 8:e72454.
- Percy, D. M., G. S. Taylor, and M. Kennedy. 2006. Psyllid communication: acoustic diversity, mate recognition and phylogenetic signal. *Invertebr. Syst.* 20:431–445.
- Pleydell, D. R. J., and S. Chretien. 2010. Mixtures of GAMs for habitat suitability analysis with overdispersed presence/absence data. *Comput. Stat. Data Anal.* 54:1405–1418.
- Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6: 7–11.
- Polacik, M., and M. Reichard. 2011. Asymmetric reproductive isolation between two sympatric annual killifish with extremely short lifespans. *Plos One* 6:e22684.
- Pombi, M., P. Kengne, G. Gimonneau, B. Tene-Fossog, D. Ayala, C. Kamdem, F. Santolamazza, W. M. Guelbeogo, N. Sagnon, V. Petrarca, et al. 2017. Dissecting functional components of reproductive isolation among closely related sympatric species of the *Anopheles gambiae* complex. *Evol. Appl.* 10:1102–1120.
- R Development Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Rafferty, N. E., and J. W. Boughman. 2006. Olfactory mate recognition in a sympatric species pair of three-spined sticklebacks. *Behav. Ecol.* 17:965–970.
- Ramsey, J., H. D. Bradshaw, and D. W. Schemske. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* 57:1520–1534.
- Raychoudhury, R., C. A. Desjardins, J. Buellesbach, D. W. Loehlin, B. K. Grillenberger, L. Beukeboom, T. Schmitt, and J. H. Werren. 2010. Behavioral and genetic characteristics of a new species of *Nasonia*. *Heredity* 104:278–288.
- Sanchez-Guillen, R. A., M. Wullenreuther, and A. Cordero Rivera. 2012. Strong asymmetry in the relative strengths of prezygotic and postzygotic barriers between two damselfly sister species. *Evolution* 66:690–707.
- Sauvion, N., O. Lachenaud, G. Genson, J. Y. Rasplus, and G. Labonne. 2007. Are there several biotypes of *Cacopsylla pruni*? *Bull. Insectol.* 60:185–186.
- Servedio, M. R., and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* 34:339–364.
- Sobel, J. M., and G. F. Chen. 2014. Unification of methods for estimating the strength of reproductive isolation. *Evolution* 68:1511–1522.
- Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. The biology of speciation. *Evolution* 64:295–315.
- Sobel, J. M., and M. A. Streisfeld. 2015. Strong premating reproductive isolation drives incipient speciation in *Mimulus aurantiacus*. *Evolution* 69:447–461.
- Soroker, V., S. Talebaev, A. R. Harari, and S. D. Wesley. 2004. The role of chemical cues in host and mate location in the pear psylla *Cacopsylla bidens* (Homoptera: Psyllidae). *J. Insect Behav.* 17:613–626.
- Sota, T., and K. Kubota. 1998. Genital lock-and-key as a selective agent against hybridization. *Evolution* 52:1507–1513.
- Takami, Y., N. Nagata, M. Sasabe, and T. Sota. 2007. Asymmetry in reproductive isolation and its effect on directional mitochondrial introgression in the parapatric ground beetles *Carabus yamato* and *C. albrechti*. *Popul. Ecol.* 49:337–346.
- Taylor, R. S., and V. L. Friesen. 2017. The role of allochrony in speciation. *Mol. Ecol.* 26:3330–3342.
- Tishechkin, D. Y. 2007. New data on vibrational communication in psyllids from the families aphalaridae and triozidae (Homoptera, Psyllinea). *Zoologicheskii Zhurnal* 86:547–553.
- Veen, T., J. Faulks, R. Rodriguez-Munoz, and T. Tregenza. 2011. Premating reproductive barriers between hybridising cricket species differing in their degree of polyandry. *Plos One* 6:e19531.
- Via, S. 2009. Natural selection in action during speciation. *Proc. Natl. Acad. Sci. USA* 106:9939–9946.
- Wenninger, E. J., D. G. Hall, and R. W. Mankin. 2009. Vibrational communication between the sexes in *Diaphorina citri* (Hemiptera: Psyllidae). *Ann. Entomol. Soc. Am.* 102:547–555.
- Wenninger, E. J., L. L. Stelinski, and D. G. Hall. 2008. Behavioral evidence for a female-produced sex attractant in *Diaphorina citri*. *Entomol. Exp. Appl.* 128:450–459.
- Yukilevich, R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in *Drosophila*. *Evolution* 66:1430–1446.

Associate Editor: S. Flaxman

Handling Editor: Mohamed A. F. Noor

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.