Heredity (2015) 114, 593–600

www.nature.com/hdy

ORIGINAL ARTICLE

A recombination suppressor contributes to ecological speciation in *OSTRINIA* moths

CB Wadsworth, X Li and EB Dopman

Despite unparalleled access to species' genomes in our post-genomic age, we often lack adequate biological explanations for a major hallmark of the speciation process—genetic divergence. In the presence of gene flow, chromosomal rearrangements such as inversions are thought to promote divergence and facilitate speciation by suppressing recombination. Using a combination of genetic crosses, phenotyping of a trait underlying ecological isolation, and population genetic analysis of wild populations, we set out to determine whether evidence supports a role for recombination suppressors during speciation between the Z and E strains of European corn borer moth (*Ostrinia nubilalis*). Our results are consistent with the presence of an inversion that has contributed to accumulation of ecologically adaptive alleles and genetic differentiation across roughly 20% of the *Ostrinia* sex chromosome (~4 Mb). Patterns in *Ostrinia* suggest that chromosomal divergence may involve two separate phases—one driving its transient origin through local adaptation and one determining its stable persistence through differential introgression. As the evolutionary rate of rearrangements in lepidopteran genomes appears to be one of the fastest among eukaryotes, structural mutations may have had a disproportionate role during adaptive divergence and speciation in *Ostrinia* and in other moths and butterflies.
Heredity (2015) **114,** 593–600; doi:10.1038/hdy.2014.128; published online 28 January 2015

INTRODUCTION

A key feature of the speciation process is the evolving structure of genetic differentiation between populations that culminates in genetic divergence across much of the genomes of daughter taxa. Indeed, genomic patterns of evolutionary divergence across the speciation continuum promise to reveal not only the functional nature of adaptive evolution and reproductive isolation (Storz, 2005) but also the processes underlying speciation (Pinho and Hey, 2010). Consequently, a major research theme has emerged that aims to characterize genome-wide patterns of differentiation between diverging populations, races and species. Pockets of divergence or 'genomic islands of speciation,' as well as larger 'continents,' have been uncovered and could reflect the effect of genes contributing to reproductive isolation ('barrier loci') (Dopman *et al.*, 2005; Turner *et al.*, 2005; Feder, Egan *et al.*, 2012). However, connecting these fundamental genetic signatures with their biological cause remains elusive (Noor and Bennett, 2009; Dopman, 2011; Nachman and Payseur, 2012; Cruickshank and Hahn, 2014).

Chromosomal rearrangements such as inversions are thought to facilitate speciation and promote the origin of evolutionary divergence by suppressing recombination. When speciation occurs in the face of gene flow, suppressed recombination in a heterokaryotypic hybrid offspring may cause the entire inverted segment to behave as a single linked unit (Noor *et al.*, 2001; Rieseberg, 2001). Consequently, only a small number of barrier loci might be required to protect a large chromosomal band, possibly several megabases in length, from gene flow. Because of restricted gene flow, barrier and locally adaptive alleles are thought to accumulate within inversions leading to increasing divergence until speciation is complete (Navarro and Barton, 2003). Finally, adaptive alleles that are captured by inversions may rapidly drive their fixation, creating long-range hitch-hiking across the nonrecombining rearranged interval (Kirkpatrick and Barton, 2006; Kirkpatrick, 2010; Feder *et al.*, 2011). In the absence of structural mutations, reduced recombination in collinear regions could also lead to evolutionary divergence; however, the footprint of genetic differentiation and opportunity for the accumulation of barrier loci might be limited (Butlin, 2005; West and Via, 2008; Feder *et al.*, 2012). Thus, changes in recombination rate because of structural mutations might be the key to speciation with gene flow, and they may therefore help to explain many instances of genomic islands of speciation.

In this study, we evaluate the role of suppressed recombination for the evolution of a modestly sized genomic island of speciation between a pair of moth lineages. The 'Z' and 'E' strains of European corn borer (ECB, *Ostrinia nubilalis*) are incipient species that colonized North America ~100 ya without an appreciable bottleneck from allopatric locations in Italy (mixture of E and Z strains) and Hungary (exclusively Z-strain; Smith, 1920; Caffrey, 1927; Klun and Cooperators, 1975; Dopman, 2011). Strains are recently diverged (~100 000 ya; Malausa *et al.*, 2007) and are textbook examples of speciation (Coyne and Orr, 2004), in which one species is splitting into two through the evolution of many forms of reproductive isolation (Dopman *et al.*, 2010). The sex (Z) chromosome harbors several known genetic factors for adaptation and/or barrier traits, including those underlying behavioral isolation (Roelofs *et al.*, 1987; Dopman *et al.*, 2004) and temporal isolation (Glover *et al.*, 1992; Dopman *et al.*, 2005). Current and historical gene exchange between

Department of Biology, Tufts University, Medford, MA, USA
Correspondence: Dr EB Dopman, Department of Biology, Tufts University, 200 Boston Avenue, Suite 4739, Medford 02155, MA, USA.
E-mail: erik.dopman@tufts.edu
Received 25 March 2014; revised 25 November 2014; accepted 26 November 2014; published online 28 January 2015

ECB strains is likely, as indicated by hybridization in Europe and North America (Dopman *et al.*, 2010; Coates *et al.*, 2013), geographic variation in strength of reproductive isolation (ranging from 0.91 to 0.99; Dopman *et al.*, 2010) and molecular evidence for an isolation-with-migration model of divergence ($2\,Nm \sim 10$; Malausa *et al.*, 2007). Whereas most loci show extensive shared polymorphism (for example, $F_{ST} < 0.05$), a possible island of speciation along an ~1-cM swath of the Z chromosome consists of four genes (*Tpi* and three olfactory receptors, *ORs*) and reveals ECB strains as nearly reciprocally monophyletic (for example, $F_{ST} > 0.7$; Dopman *et al.*, 2005; Dopman, 2011; Lassance *et al.*, 2011). Evolutionary divergence can be explained by several tightly linked barrier loci on the sex chromosome or by a regional selective sweep at these or other loci. However, such interpretations must be viewed as incomplete without proper tests of suppressed recombination and its repercussions for the genomic zone of influence for adaptation and/or barrier loci. Indeed, a possible signature of an inversion between Z and E strains was detected on the Z chromosome (Dopman *et al.*, 2004), in which the lowest estimated recombination rate across the entire 1697-cM genetic linkage map was found in the vicinity of the *Tpi*/*ORs* loci (Figure 1).

If inversions are important during speciation and the evolution of genetic differentiation, at least three testable predictions have been made (Faria and Navarro, 2010). First, patterns of gene flow should be higher within collinear versus rearranged regions (Rieseberg, 2001; Feder *et al.*, 2005; Kulathinal *et al.*, 2009; McGaugh and Noor, 2012). Second, traits involved in reproductive isolation should map to regions located within chromosomal rearrangements (Noor *et al.*, 2001; Feder *et al.*, 2005; Lowry and Willis, 2010). Finally, genotyping hybrid offspring in the laboratory ought to reveal evidence for suppressed recombination (Feder *et al.*, 2003; Kulathinal *et al.*, 2009). By using a combination of genetic crosses, phenotyping of a trait contributing to reproductive isolation, and population genetic analysis, we set out to test these predictions for the ECB sex chromosome.

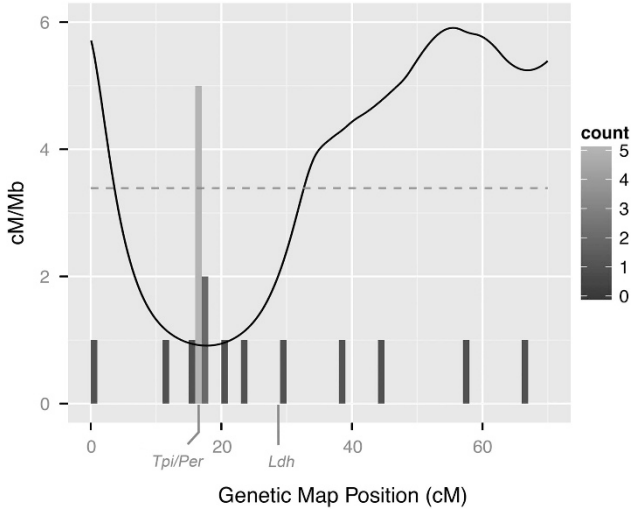

Figure 1 Local recombination rates (cM/Mb, black line) and marker densities (count) on the ECB sex (Z) chromosome. The dashed red line indicates the genome-wide average. Genetically differentiated loci *Tpi* and *ORs* ($F_{ST} > 0.7$) are located in the low-recombination region (~17 cM). Adapted from Dopman *et al.* (2004).

MATERIALS AND METHODS

Recombination suppression

To estimate variation in recombination frequencies surrounding the genetically differentiated region of the ECB sex chromosome, we generated three-point test crosses within and between Z and E strains. Crosses used insects collected as caterpillars, pupae and adults from the New York State, USA ($n \sim 500$ males and $n \sim 500$ females each from East Aurora, Geneva and Bouckville). Laboratory populations were maintained by mass rearing ~100 males and ~100 females each generation. The female sex-pheromone blend is strain-specific. Z-strain females produce a 3:97 ratio of E/Z11-14:OAc, whereas E-strain females use a 99:1 ratio. Strain identities of all parental stocks were confirmed by genotyping *pgFAR*, the autosomal locus that determines the female sex-pheromone blend (Lassance *et al.*, 2010; Supplementary Table S1).

Recombination is restricted to males in many Lepidoptera (Dopman *et al.*, 2004). Therefore, in backcross pedigrees only F1 males are informative. Here, between-strain backcrosses (referred to as 'Z × E') involved a Z grandmother, an E grandfather, an F1 male and a recurrent Z-strain mother. Within-strain crosses (referred to as 'Z × Z') had a similar design but were exclusively from the Z-strain.

The ECB diverged from *Bombyx mori* (the silk moth) ~100 Mya (Pringle *et al.*, 2007); however, the species pair shows a high degree of conservation in gene order like many other Lepidoptera (Pringle *et al.*, 2007; d'Alencon *et al.*, 2010; Kroemer *et al.*, 2011). Therefore, we relied on the *B. mori* genome to identify evenly spaced sex-linked mapping loci for the ECB, whose genome is incomplete. Markers consisted of *Tpi* at 9 Mb, *Per* at 13 Mb and *Ldh* at 17.4 Mb (International Silkworm Genome Consortium, 2008). We adopted methods developed elsewhere for *Tpi* and *Ldh* (Dopman *et al.*, 2005), whereas initial *Per* sequence for the ECB was obtained using degenerate PCR.

DNA from mapping family individuals was extracted from adult legs following using a DNeasy Blood and Tissue kit (QIAGEN, Valencia, CA, USA). Genotyping assays for offspring were developed from segregating polymorphisms and consisted of diagnostic PCR products that amplified polymorphic indels and/or restriction enzyme cut sites (*Eco*R1, *Alu*1 and *Bsr*1 for *Tpi*, *Per* and *Ldh*, respectively; Supplementary Table S1). Only female offsprings were genotyped and, because female Lepidoptera are the heterogametic sex and possess only one Z chromosome, one allele per locus was detected. Genotyping results were confirmed by Sanger sequencing ($n \sim 100$ sequenced alleles per marker). One Z × E cross had already been genotyped for *Tpi* and *Ldh* in a prior study (Dopman *et al.*, 2005).

Colocalization of reproductive isolation

As prior results (Figure 1) and preliminary data suggested evidence for reduced recombination near *Tpi* in F1 hybrids, we tested for colocalization of traits involved in reproductive isolation. A subset of Z × E backcross offspring was mapped for a major sex-linked factor (named *Pdd*) controlling diapause emergence time (Glover *et al.*, 1992; Dopman *et al.*, 2005), defined as the time to pupation for over-wintering caterpillars under environmental conditions conducive to breaking diapause. Differences at *Pdd* contribute to temporal isolation between Z and E strains in North America by conferring an ~30-day shift in emergence and thus adult-mating flights (Wadsworth *et al.*, 2013). Shifts in adult flights between univoltine Z and bivoltine E populations eliminate as much as 85% gene flow in nature (Dopman *et al.*, 2010).

We measured diapause timing variation by inducing and then breaking diapause in recombinant ECB caterpillars. Following earlier studies (Glover *et al.*, 1992; Dopman *et al.*, 2005), diapause was induced by a 12:12 light:dark photoperiod and then broken 35 days later with a 16:8 light:dark cycle. Diapause emergence time was noted every 2 days. As females are hemizygous and *Pdd* is inherited as a Mendelian locus, backcross females show patterns of emergence consistent with either their grandfather (here, E-strain) or grandmother (here, Z-strain; Glover *et al.*, 1992; Dopman *et al.*, 2005). Hence, recombinant backcross females show E-like earlier emergence times and pupate in less than 22 days (range = 8–22 days, mean = 15.5 days, s.e. ± 0.55), or they show Z-like later emergence times and pupate in more than 18 days (range = 18–80, mean = 43.94 days, s.e. ± 1.64; Glover *et al.*, 1992).

Higher introgression in collinear regions

We evaluated patterns of genetic variation among field-caught female ECB: 18 insects were E-strain, 18 were Z-strain and at least one insect was from an outgroup species, the Asian corn borer (ACB, *O. furnacalis*). ECB caterpillars, pupae or adults were collected over a large geographic range, and thus genetic similarities should be viewed as indicating a combination of ongoing or recent gene exchange. Samples included insects from USA ($n=7$ E-strain and $n=5$ Z-strain from New York; $n=7$ E-strain from North Carolina and $n=4$ Z-strain from North Carolina; $n=4$ Z-strain from Iowa), Italy ($n=4$ E-strain and $n=3$ Z-strain) and Hungary ($n=2$ Z-strain). Moths were classified into strain by diagnostic gas chromatographic profiles of female sex-pheromone blend (that is, 99:1 versus 3:97 ratio of E/Z11-14:OAc; for example, Dopman *et al.*, 2004) and/or by genotyping the *pgFAR* locus.

Molecular markers were developed from genes that were evenly distributed across the *B. mori* sex chromosome. Using ECB transcripts that were developed from a separate study, reciprocal BLASTs were performed to obtain ECB–*B. mori* gene pairs with predicted locations on the ECB sex chromosome (*B. mori* nucleotide CDS, BLASTn, *e*-value$<$1e$-$40). ECB transcripts were then searched with tBLASTx against the matching *B. mori* genomic sequence to identify predicted introns. Primer pairs for 23 loci (Supplementary Table S2) that had interlocus intervals ranging from 0.1 to 2.5 Mb (mean$=$0.96 Mb, s.e.$\pm$0.16) were designed to amplify ECB introns using PrimerBlast (NCBI). Sequences from field-caught insects were aligned and edited by eye using Genious Pro 5.5 (Biomatters Inc., San Francisco, CA, USA). Markers were named based on *Bombyx* gene identifiers.

Measures of genetic variation (π, θ_w), genetic differentiation (d_a, d_{xy}, F_{ST} (Hudson *et al.*, 1992), S_{nn} (Hudson, 2000)) and the allele-frequency spectrum (Tajima's *D* (Tajima, 1989b), Fay and Wu's *H* (Fay and Wu, 2000)) were calculated using DnaSP (Rozas *et al.*, 2003). The genealogical sorting index (*gsi*; Cummings *et al.*, 2008) was calculated to quantify the degree of exclusive ancestry of Z- and E-strain moths. *gsi* ranges from 0 (no exclusivity) to 1 (monophyletic), and was calculated in R (Team RDC, 2013) on the Tufts high-performance computing research cluster using the genealogicalSorting library (www.genealogicalsorting.org). Genealogies were constructed using neighbor-joining with a Tamura-Nei 93 model of evolution to calculate distances. Statistical significance of *gsi* was assessed through permutation tests ($n=10\,000$) on each of 1000 bootstrap replicates for a total of 10 000 000 replicated genealogies. Indel polymorphisms were considered a fifth base when calculating genetic differentiation and genetic exclusivity. All figures were constructed using the ggplot2 library (Wickham, 2009) in R.

RESULTS

Recombination suppression

We used three-point crosses within- and between-strain to test for evidence of suppressed recombination associated with a high-frequency chromosomal rearrangement between Z and E strains of ECB. No significant differences in recombination frequency were observed among different within-strain backcross families (Z$\times$Z, $n=2$ families) or among between-strain backcross families (Z$\times$E, $n=7$ families) and therefore families were combined for each backcross type. Within strain, we observed six recombinants between the *Tpi* and *Per* pair (7%, $n=86$ offsprings), six between the *Per* and *Ldh* pair (7%, $n=86$ offsprings) and twelve between *Tpi* and *Ldh* (14%, $n=86$ offsprings; Figure 2a). Hence, in agreement with that found in *B. mori*, the predicted gene order within the Z-strain of ECB is *Tpi—Per—Ldh* (Figure 2b).

Between-strain crosses revealed heterogeneity in recombination frequency compared with that seen within-strain (Figure 2a). ZxE crosses yielded slightly elevated recombination frequencies for the *Per–Ldh* gene pair (seven recombinants, 9.1%, $n=77$ offsprings; Fisher's exact test, *P*-value$=$0.7739) and slightly reduced recombination between *Tpi* and *Ldh* (seven recombinants, 9.1%, $n=77$ offsprings; Fisher's exact test, *P*-value$=$0.4643). However, a statistically significant suppression in recombination was observed between *Tpi* and *Per*.

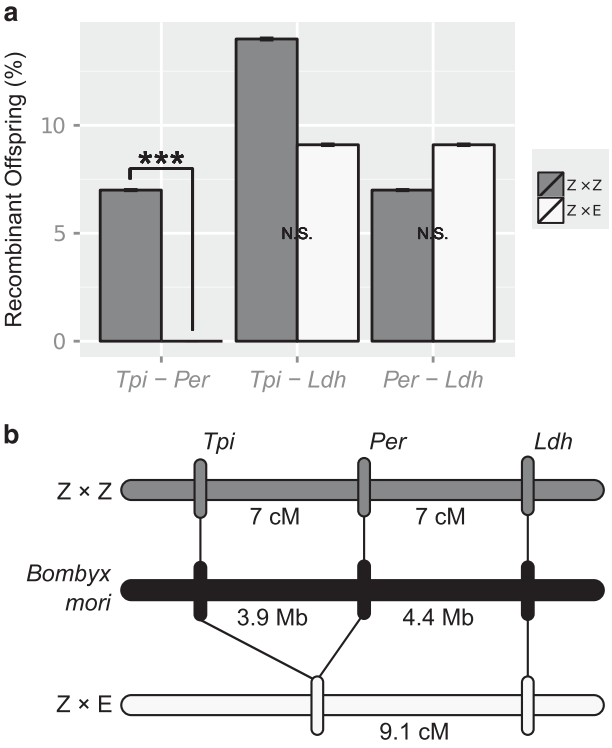

Figure 2 Suppressed recombination on the sex chromosome in ECB hybrids. (**a**) Percent recombinant offspring between *Tpi*, *Per* and *Ldh* in within-strain (Z$\times$Z) and between-strain (ZxE) crosses. (**b**) Inferred linear order of markers in *Bombyx mori* and the two ECB crosses. ***P-value$<$0.001.

Zero recombinants were found in over 600 offsprings ($n=666$ offsprings; Fisher's exact test, *P*-value$=$1.91e$-$06).

Colocalization of reproductive isolation

We phenotyped a subset of backcross offspring to determine the linkage relationship between the nonrecombining *Tpi–Per* gene pair and *Pdd*. Diapause emergence timing of 333 hemizygous females from backcross families ranged from 2 to 100 days; yet, a clear bimodal distribution emerged (Figure 3). Eight offspring (2.4%) had emergence times that fell within the period of overlap of parental Z and E strains (day 18–22) and their *Pdd* genotype was considered indeterminate. The remaining 325 offspring had emergence times consistent with pure-strain insects. Of these, one female with E-strain markers at the *Tpi–Per* gene pair displayed Z-strain like diapause and emerged later, and one female with Z-strain markers displayed E-strain like diapause and emerged earlier. Thus, two females (0.62%) were apparent recombinants between the *Tpi–Per* gene pair and the *Pdd* locus. In the other 323 offspring (99.38%), *Pdd* was perfectly associated with their genotype at the *Tpi–Per* gene pair.

Higher introgression in collinear regions

We characterized genetic variation for 23 loci to test for differences in patterns of genetic variation between loci with predicted locations inside (referred to as 'rearranged') and outside (referred to as 'collinear') of the nonrecombining *Tpi–Per* chromosomal region. A total of 779 sequences were obtained, resulting in $\sim$9 kb of aligned data and 224 segregating sites. Seven loci fell within the *Tpi–Per* interval and 16 were outside (8 upstream and 8 downstream).

We tested for variation in patterns of genetic differentiation using four different indices. Rearranged loci exhibited statistically

significantly elevated evolutionary divergence for both F_{ST} (median rearranged = 0.14990, median collinear = 0.06418, $W = 88$, P-value = 0.01764) and S_{nn} (median rearranged = 0.6569, median collinear = 0.5742, $W = 90$, P-value = 0.01258; Figure 4). Similarly, the number of net substitutions per site between strains (d_a, median rearranged = 0.09%, median collinear = 0.029%, $W = 74$, P-value = 0.121) and the average number of substitutions per site between strains (d_{xy}, median rearranged = 0.76%, median collinear = 0.66%, $W = 54$, P-value = 0.5647) were both higher for rearranged loci (Figure 4), although these differences were not statistically significant.

We also used *gsi* to explore gene tree topology measures of genetic differentiation (Table 1). Loci with significant values of *gsi* ranged from 0.01 to 0.68 with an average of 0.22. E-strain ECB had seven loci with significant genealogical exclusivity, whereas the Z-strain had 15. There were no significant differences in the magnitude of *gsi* for rearranged and collinear loci ($t = -1.3504$, degrees of freedom = 19.019, P-value = 0.1927); however, all rearranged loci possessed significant genealogical exclusivity in either one or both strains, in contrast to ~50% of collinear loci (Fisher's exact test, P-value = 0.0574).

To evaluate patterns of nucleotide polymorphism, we compared π and θ_w between groups. Average genetic variation for rearranged loci was approximately one-half of that for collinear loci (Table 2), although only statistically significant (or trending) for E-strain moths (mean rearranged $\pi = 0.386\%$, mean collinear $\pi = 0.812\%$, $t = -1.7158$, degrees of freedom = 16.839, P-value = 0.05227; mean $\theta_w = 0.347\%$, mean collinear $\theta_w = 0.916\%$, $t = -2.3966$, degrees of freedom = 15.098, P-value = 0.01496). For both strains Tajima's D showed a slight skew toward an excess of intermediate-frequency alleles compared with low-frequency alleles in rearranged loci (that is, more positive values; Table 2), whereas Fay and Wu's H revealed a relative excess of high-frequency derived alleles compared with intermediate-frequency alleles (that is, more negative values; Table 2). However, neither estimate was statistically significant.

DISCUSSION

Although our post-genomic age has produced billions of base pairs of DNA sequence data for many organisms across the speciation continuum, we often lack adequate biological explanations for the evolution of fundamental genomic features such as genetic divergence. Inversions are thought to have far-reaching impacts on evolutionary divergence during speciation, especially when divergence occurs in the face of gene flow. However, demonstrating the importance of inversions to the speciation process requires evidence for (i) recombination suppression in hybrid offspring, (ii) localization of barrier or adaptation loci in rearranged regions and (iii) enhanced sequence divergence (Faria and Navarro, 2010). In this study, we combined genetic crosses (Figure 2), phenotyping of traits underlying reproductive isolation (Figure 3) and observational studies of DNA polymorphism in wild populations (Figure 4, Table 1) to show that Z and E strains of ECB meet all three criteria. Although we cannot physically confirm the presence of an inversion yet through cytology, these results are consistent with the presence of an inversion that has contributed to accumulation of ecologically adaptive alleles and genetic

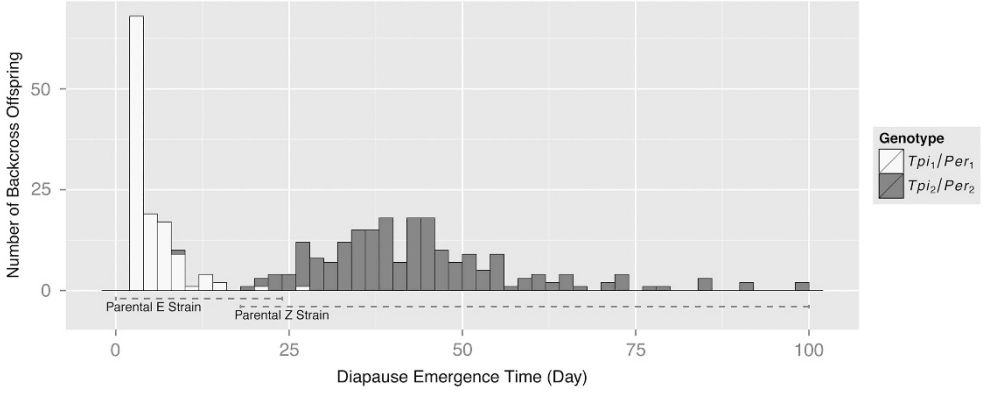

Figure 3 Diapause emergence time maps to a putative sex-linked inversion. Bimodal distribution of diapause emergence time for female backcross offspring genotyped for *Tpi* and *Per*. Note that Tpi_E/Per_E are E-strain alleles and Tpi_Z/Per_Z are Z-strain alleles. Dashed lines indicate parental distribution values.

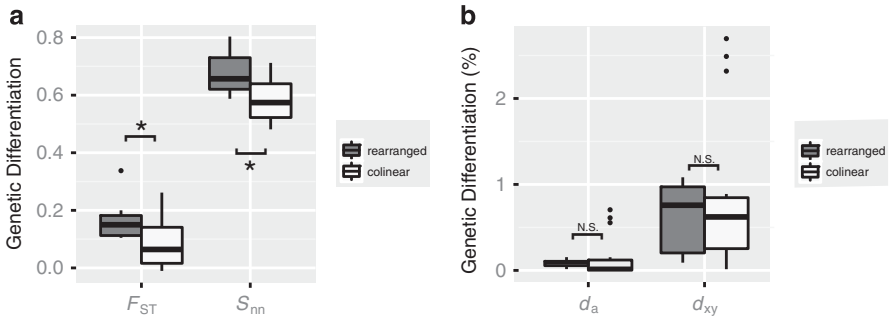

Figure 4 Elevated evolutionary divergence between ECB strains. (**a**) F_{ST} and S_{nn}. (**b**) d_a and d_{xy}. *P-value < 0.05.

Table 1 Genetic exclusivity between ECB strains for colinear and rearranged loci

Location	*Locus*	gsi_E	P-*value*[a]	gs_{iZ}	P-*value*[a]
Rearranged	*Tpi*	0.14	N.S.	0.26	***
Rearranged	*Per*	0.11	N.S.	0.21	*
Rearranged	680	0.28	**	0.35	***
Rearranged	524	0.18	N.S.	0.35	***
Rearranged	710	0.17	N.S.	0.26	**
Rearranged	1368	0.19	N.S.	0.29	**
Rearranged	683	0.18	*	0.12	N.S
Colinear	2071	0.23	N.S.	0.23	N.S
Colinear	2076	0.13	N.S.	0.17	N.S
Colinear	2104	0.1	N.S.	0.12	N.S
Colinear	2005	0.17	N.S.	0.28	*
Colinear	622	0.14	N.S.	0.2	N.S
Colinear	643	0.26	***	0.23	***
Colinear	647	0.33	**	0.32	**
Colinear	563	0.02	N.S.	0.5	***
Colinear	9691	0.68	***	0.52	***
Colinear	13318	0.01	N.S.	0.23	**
Colinear	13333	0.2	N.S.	0.22	N.S
Colinear	8010	0.28	***	0.33	***
Colinear	12292	0.08	N.S.	0.04	N.S
Colinear	13832	0.19	**	0.28	***
Colinear	587	0.12	N.S.	0.29	**
Colinear	2781	0.07	N.S	0.12	N.S

***P-value<0.001; **P-value<0.01; *P-value<0.05. N.S. not significant.
[a]Significance determined by permutation ($n=10\,000$) of bootstrap replicated ($n=1000$).

Table 2 Patterns of genetic variation among rearranged and colinear loci

	Loci	*Length*[a]	S[b]	π[c](%)	θ[d](%)	D[e]	H[f]
E Strain							
Rearranged	7	3541	47	0.386	0.347	0.214	−0.229
Colinear	16	5172	107	0.812	0.916	−0.492	−0.178
P-value				0.052	0.015	0.888	0.473
Z Strain							
Rearranged	7	3549	60	0.486	0.514	−0.165	−3.266
Colinear	16	5172	106	0.812	0.849	−0.379	−0.883
P-value				0.127	0.102	0.646	0.131

[a]Base pairs. [b]Segregating sites. [c]Heterzygosity from average number of differences per site.
[d]Heterzygosity from number of polymorphic sites. [e]Tajima, 1989a, b. [f]Fay and Wu, 2000.

differentiation across roughly 20% of the ECB sex chromosome (~4 Mb).

Inversions and reproductive isolation

A major appeal of chromosomal rearrangement models of speciation is their ability to help explain apparent examples of speciation with gene flow (Pinho and Hey, 2010). Indeed, a serious theoretical weakness of nonallopatric speciation is that recombination randomizes associations between genes conferring local adaptation (hybrid unfitness) and assortative mating (Felsenstein, 1981; Gavrilets, 2003; Butlin, 2005). Although the specifics of chromosomal rearrangement models of speciation differ (Noor *et al.*, 2001; Rieseberg, 2001; Navarro and Barton, 2003; Kirkpatrick and Barton, 2006; Faria and Navarro, 2010), a common feature is the removal of this antagonism and a buildup of adaptation and barrier loci within rearrangements during a nonallopatric phase(s) that then allows for reinforcement and the completion of speciation. Thus, a critical prediction for speciation with gene flow is that rearrangements such as inversions should be present and contain multiple prezygotic and/or postzygotic barrier loci with major effects on reproductive isolation. Determining the degree to which these expectations are borne out in nature is challenging. First, we have limited data on quantitative measures of reproductive isolation, defined as $RI = 1 - (\text{between species fitness}) \div (\text{within species fitness})$ and ranging from 0 (no effect) to 1 (complete *RI*; Ramsey *et al.*, 2003; Coyne and Orr, 2004). Second, the genomic positions of the underlying barrier loci is often unknown. Monkeyflowers are a rare exception, in which loci of large reproductive isolation effect that contribute to floral traits, flowering time and hybrid male sterility map to chromosomal rearrangements in the *Mimulus* genome (Ramsey *et al.*, 2003; Lowry and Willis, 2010; Fishman *et al.*, 2013).

The putative sex-linked inversion we document in *Ostrinia* appears consistent with the large-effect, multicomponent prediction of chromosomal rearrangement models of speciation. First, multiple 'speciation phenotypes' (Shaw and Mullen, 2011) in ECB have known sex-linked factors, including male pheromone response and diapause timing (McLeod, 1978; Dopman *et al.*, 2004; 2005; Wanner *et al.*, 2010; Ikten *et al.*, 2011). Second, a recent study indicated that large isolation effect sizes are common between ECB strains ($RI \geqslant 0.5$ for 4 of 7 forms of isolation; Dopman *et al.*, 2010). Given this prior work and the large size of the putative inversion (7 cM or ~4 Mb; Figure 2b), there would appear to be good reason to expect diverse and potent barrier loci in an inversion on the ECB sex chromosome. Indeed, we found that only two of 325 offspring were recombinants (0.62%) between the nonrecombining region (*Tpi–Per* loci, Figure 2b) and the factor(s) determining diapause emergence time (*Pdd*; Figure 3). Thus, a major cause of temporal isolation (mean $RI = 0.65$) is likely inside or near the chromosomal rearrangement. Similarly, the sex-linked *ORs* are known to underlie strain-specific male antennal response to Z versus E-strain female pheromones (that is, 3:97 versus 99:1 ratio of E/Z11-14:OAc) and they are tightly linked to *Tpi* and thus the sex-linked rearrangement (Wanner *et al.*, 2010; Lassance *et al.*, 2011). Hence, a rearrangement may have promoted coupling between multiple genes conferring ecological (*Pdd*) and behavioral (*ORs*) divergence, thereby facilitating speciation. Clearly, there is a need for formal experimentation to determine whether chromosomal rearrangements in *Ostrinia* and in other systems harbor multiple loci of large isolation effect that contribute to varied components of reproductive isolation.

Evolution of genetic divergence

Results from several systems including *Ostrinia* and *Drosophila* (McGaugh and Noor, 2012) imply high genetic divergence within inversions and minimal divergence outside of them, and a reciprocal relationship for polymorphism. Such patterns may be common in the early stages of speciation because of the coupling of selection and recombination. Globally beneficial mutations and adaptive barrier alleles that arise within fixed inversions are expected to get trapped and cannot easily migrate between lineages (Navarro and Barton, 2003). Further, according to the Kirkpatrick–Barton model (Kirkpatrick and Barton, 2006), selective establishment and fixation of inversions occurs because locally adaptive alleles at loci within inversions (for example, underlying ecological and behavioral divergence) cannot recombine with chromosomes from other populations and thus they are favored because they do not suffer the 'Achilles' heel' of being located on the same chromosome as immigrant disadvantageous allele(s). Hence, genetic hitch-hiking associated with either

selective fixation of inversions or their internal mutations might lead to inversion-specific reductions in polymorphism within one or both lineages, while also creating enhanced differentiation for relative divergence measures that are sensitive to within lineage variation (for example, $F_{ST;}$ Charlesworth, 1998).

An additional role of suppressed recombination for evolutionary divergence involves repeated bouts of gene exchange and disruptive selection. Specifically, barrier loci within inversions under the Navarro–Barton model (Navarro and Barton, 2003) might produce a molecular signal of differentiation that is spread across the inversion or at least near its breakpoints (Rieseberg, 2001; Noor *et al.*, 2001; Strasburg *et al.*, 2009; McGaugh and Noor, 2012) because of selective purging of maladapted (and nonrecombining) regions encompassed by inversions upon their introduction into sister taxa. Such a process might be analogous to 'differential introgression,' in which gene flow is restricted at individual barrier loci but is high elsewhere (Barton and Bengtsson, 1986; Wu, 2001; West and Via, 2008; Harrison, 2012). Over the long term, differential introgression is expected to produce specific signatures. For example, increases in absolute genetic divergence (for example, d_{xy}) as new mutations arise and fix independently in daughter taxa (Charlesworth *et al.*, 1997; Charlesworth, 1998; Nachman and Payseur, 2012; Cruickshank and Hahn, 2014).

Despite limitations of our current sampling scheme (18 samples per strain from both allopatric and sympatric localities), results from *Ostrinia* suggest that selection and suppressed recombination were important for evolutionary divergence. Compared with 'collinear' loci, patterns of genetic variation at 'rearranged' loci are consistent with a selective sweep(s), including 50% lower genetic variation (primarily in E-strain moths; Table 2) and 2–3× greater differentiation for measures that are sensitive to selection at linked sites (F_{ST} and $d_{a;}$ Figure 4). Support for differential introgression is less obvious. The putative inversion shows modest increases in absolute divergence (1.15× greater values of d_{xy}) and also genealogical exclusivity (1.14× greater values of S_{nn}). Moreover, although patterns in the allele-frequency spectrum (D and H) are consistent with rare migration within the inversion and higher migration elsewhere (Table 2; Tajima, 1989a; Simonsen *et al.*, 1995; Przeworski, 2002; Fay and Wu, 2005), these changes are relatively minor. Lack of genetic evidence for differential introgression runs counter to observations in nature. An appreciable number of hybrid offsprings occur at sampled sympatric localities in New York (5–15%; Dopman *et al.*, 2010; Coates *et al.*, 2013) and yet differentiation between Z and E strains at *Tpi* has been maintained at these sites for more than 40 generations (Glover *et al.*, 1991; Dopman *et al.*, 2005; Dopman, 2011). Disruptive selection following introgression of the nonrecombining region that includes *Tpi* might be important in maintaining divergence at New York localities, where sweeps and incomplete lineage sorting alone cannot easily account for patterns of nucleotide variation across loci. Equivocal genetic support for differential introgression should not be surprising because statistical power is limited during early stages of speciation (for example, for $d_{xy;}$ Cruickshank and Hahn, 2014). Hence, lineages that are in the early stages of divergence like ECB strains may be too closely related to show strong statistical signs of differential gene flow even at genome regions where it occurs.

More extensive investigation of *Ostrinia* moths and other systems might validate the scenario we propose in which evolutionary divergence during speciation occurs by two complementary but distinct mechanisms—one driving the transient origin of differentiation through local adaptation and selection on linked sites, and one determining its stable persistence through reproductive isolation and/or differential introgression (see also Dopman, 2011; Nachman and Payseur, 2012; Cruickshank and Hahn, 2014). If an inversion is present as a low-frequency polymorphism in allopatry and greatly differentiates lineages only in geographic regions where they co-occur (for example, as in the 'mixed-geographic model' (Feder *et al.*, 2011)), this might suggest a direct role of recombination suppression in promoting evolutionary divergence by local adaptation and selective sweeps (for example, to eliminate the creation of less fit hybrids or to keep high fitness loci together). In contrast, the presence of an older, high-frequency inversion that differentiates lineages regardless of geography could indicate an indirect role of recombination suppression in facilitating accumulation of additional genetic changes once inversions establish or fix. That is, the rearrangement could protect chromosomal regions from gene flow and allow for additional genetic changes to accrue (for example, through local sweeps of internal beneficial mutations or differential introgression of the region encompassed by the rearrangement). Broader comparisons between recently sympatric Z and E strains of *Ostrinia* living in North America and their allopatric, European ancestors may help distinguish between direct and indirect roles of recombination suppression during divergence.

CONCLUSIONS

Our studies of recombination, reproductive isolation and genetic variation in the wild imply that structural mutations have facilitated speciation with gene flow between Z and E strains of ECB. Our results add to the limited number of studies that provide empirical support for this notion, including apple maggot flies (Feder *et al.*, 2003; 2005; Michel *et al.*, 2010; Powell *et al.*, 2013), fruit flies (Noor *et al.*, 2001; Kulathinal *et al.*, 2009), stickleback fishes (Kitano *et al.*, 2009), sunflowers (Rieseberg, 2001; Strasburg *et al.*, 2009) and monkey-flowers (Lowry and Willis, 2010; Fishman *et al.*, 2013). Our finding that structural mutations contribute to speciation may be the first of many for Lepidoptera. The rate of rearrangement in lepidopteran genomes appears to be one of the fastest among eukaryotes at ~2 breakages/Mb/My, or ~3× faster than nematodes and more than an order of magnitude faster than flies, mammals and plants (Ranz *et al.*, 2001; Coghlan and Wolfe, 2002; d'Alencon *et al.*, 2010). As lepidopterans (and nematodes) have diffusely organized centromeres (holocentric), these high evolutionary rates may stem from an increased likelihood of reintegration of double-strand break fragments (d'Alencon *et al.*, 2010). Hence, tens or even hundreds of fixed rearrangements could characterize moth or butterfly species pairs that diverged just several hundred thousand years ago, implying that structural mutations may have had a disproportionate role over the evolutionary history of the second largest order of insects, and may commonly promote adaptive divergence and speciation in moths and butterflies.

DATA ARCHIVING

Data available from the Dryad Digital Repository: http://doi.org/10.5061/dryad.2vg05

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Charles Linn for insect stocks, Jean-Marc Lassance for initial *Per* primers and Genevieve Kozak for comments on an early version of the manuscript. The study was supported by the USDA (EBD: 2010-65106-20610), NSF (EBD: DEB-1257251; CBW: 2011116050) and Tufts University (XL:

Summer Scholar Program). We thank Jeff Feder, Aurora Ruiz-Herrera and three anonymous referees for suggestions on this final version.

Barton N, Bengtsson BO (1986). The barrier to genetic exchange between hybridising populations. *Heredity* **57**: 357–376.
Butlin RK (2005). Recombination and speciation. *Mol Ecol* **14**: 2621–2635.
Caffrey D (1927). *A Progress Report on the Investigations of the European Corn Borer*. Governmental Printing Office: Washington, DC.
Charlesworth B (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Mol Biol Evol* **15**: 538–543.
Charlesworth B, Nordborg M, Charlesworth D (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet Res* **70**: 155–174.
Coates BS, Johnson H, Kim K-S, Hellmich RL, Abel CA, Mason C *et al.* (2013). Frequency of hybridization between *Ostrinia nubilalis* E-and Z-pheromone races in regions of sympatry within the United States. *Ecol Evol* **3**: 2459–2470.
Coghlan A, Wolfe KH (2002). Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Res* **12**: 857–867.
Coyne JA, Orr HA (2004). *Speciation*1st edn.Sinauer Associates Inc: Sunderland, MA.
Cruickshank TE, Hahn MW (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol* **23**: 3133–3175.
Cummings MP, Neel MC, Shaw KL (2008). A genealogical approach to quantifying lineage divergence. *Evolution* **62**: 2411–2422.
d'Alencon E, Sezutsu H, Legeai F, Permal E, Bernard-Samain S, Gimenez S *et al.* (2010). Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. *Proc Natl Acad Sci USA* **107**: 7680–7685.
Dopman EB (2011). Genetic hitchhiking associated with life history divergence and colonization of North America in the European corn borer moth. *Genetica* **139**: 565–573.
Dopman EB, Bogdanowicz SM, Harrison RG (2004). Genetic mapping of sexual isolation between E and Z pheromone strains of the european corn Borer (*Ostrinia nubilalis*). *Genetics* **167**: 301–309.
Dopman EB, Pérez L, Bogdanowicz SM, Harrison RG (2005). Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proc Natl Acad Sci USA* **102**: 14706–14711.
Dopman EB, Robbins PS, Seaman A (2010). Components of reproductive isolation between North American pheromone strains of the European Corn Borer. *Evolution* **64**: 881–902.
Faria R, Navarro A (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol Evol* **25**: 660–669.
Fay J, Wu C-I (2000). Hitchhiking under positive Darwinian selection. *Genetics* **155**: 1405–1413.
Fay JC, Wu C-I (2005). Detecting hitchhiking from patterns of DNA polymorphism. In Nurminsky D (ed) *Selective Sweep*. Kluwer Academic/Plenum Publishers: New York, NY.
Feder J, Roethele J, Filchak K, Niedbalski J (2003). Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* **163**: 939–953.
Feder JL, Egan SP, Nosil P (2012). The genomics of speciation-with-gene-flow. *Trends Genet* **28**: 342–350.
Feder JL, Gejji R, Powell THQ, Nosil P (2011). Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* **65**: 2157–2170.
Feder JL, Gejji R, Yeaman S, Nosil P (2012). Establishment of new mutations under divergence and genome hitchhiking. *Philos T R Soc B* **367**: 461–474.
Feder JL, Xie X, Rull J, Velez S, Forbes A, Leung B *et al.* (2005). Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *P Natl Acad Sci USA* **102** (Suppl 1): 6573–6580.
Felsenstein J (1981). Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**: 124–138.
Fishman L, Stathos A, Beardsley PM, Williams CF, Hill JP (2013). Chromosomal rearrangements and the genetics of reproductive barriers in *Mimulus* (monkey flowers). *Evolution* **67**: 2547–2560.
Gavrilets S (2003). Perspective: models of speciation: what have we learned in 40 years? *Evolution* **57**: 2197–2215.
Glover TJ, Knodel JJ, Robbins PS, Eckenrode CJ, Roelofs WL (1991). Gene flow among three races of European corn borers (Lepidoptera: Pyralidae) in New York State. *Environ Entomol* **20**: 1356–1362.
Glover TJ, Robbins PS, Eckenrode CJ, Roelofs WL (1992). Genetic control of voltinism characteristics in European corn borer races assessed with a marker gene. *Arch Insect Biochem* **20**: 107–117.
Harrison RG (2012). The language of speciation. *Evolution* **66**: 3643–3657.
Hudson RR (2000). A new statistic for detecting genetic differentiation. *Genetics* **155**: 2011–2014.
Hudson RR, Slatkin M, Maddison WP (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**: 583–589.
Ikten C, Skoda SR, Hunt TE, Molina-Ochoa J, Foster JE (2011). Genetic variation and inheritance of diapause induction in two distinct voltine ecotypes of *Ostrinia nubilalis* (Lepidoptera: Crambidae). *Ann Entomol Soc Am* **104**: 567–575.
International Silkworm Genome Consortium (2008). The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem Molec* **38**: 1036–1045.
Kirkpatrick M (2010). How and why chromosome inversions evolve. *PLoS Biol* **8**: e1000501.
Kirkpatrick M, Barton N (2006). Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419–434.
Kitano J, Ross JA, Mori S, Kume M, Jones FC, Chan YF *et al.* (2009). A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**: 1079–1083.
Klun JA, Cooperators (1975). Insect sex pheromones: Intraspecific pheromonal variability of *Ostrinia nubilalis* in North America and Europe. *Environ Entomol* **4**: 891–894.
Kroemer JA, Coates BS, Nusawardani T, Rider SD, Fraser LM, Hellmich RL (2011). A rearrangement of the Z chromosome topology influences the sex-linked gene display in the European corn borer. *Ostrinia nubilalis*. *Mol Genet Genomics* **286**: 37–56.
Kulathinal RJ, Stevison LS, Noor MAF (2009). The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet* **5**: e1000550.
Lassance J-M, Bogdanowicz SM, Wanner KW, Löfstedt C, Harrison RG (2011). Gene genealogies reveal differentiation at sex pheromone olfactory receptor loci in pheromone strains of the European Corn Borer, *Ostrinia nubilalis*. *Evolution* **65**: 1583–1593.
Lassance J-M, Groot AT, Liénard MA, Antony B, Borgwardt C, Andersson F *et al.* (2010). Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* **466**: 486–489.
Lowry DB, Willis JH (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol* **8**: e1000500.
Malausa T, Leniaud L, Martin J-F, Audiot P, Bourguet D, Ponsard S *et al.* (2007). Molecular differentiation at nuclear loci in French host races of the European corn borer (*Ostrinia nubilalis*). *Genetics* **176**: 2343–2355.
McGaugh SE, Noor MAF (2012). Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Philos T R Soc B* **367**: 422–429.
McLeod D (1978). Genetics of diapause induction and termination in the European corn borer, *Ostrinia nubilalis* (Lepidoptera: Pyralidae), in southwestern Ontario. *Can Entomol* **110**: 1351–1353.
Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL (2010). Widespread genomic divergence during sympatric speciation. *Proc Natl Acad Sci USA* **107**: 9724–9729.
Nachman MW, Payseur BA (2012). Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos T R Soc B* **367**: 409–421.
Navarro A, Barton N (2003). Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* **57**: 447–459.
Noor MA, Grams KL, Bertucci LA, Reiland J (2001). Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* **98**: 12084–12088.
Noor MAF, Bennett SM (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* **103**: 439–444.
Pinho C, Hey J (2010). Divergence with gene flow: models and data. *Annu Rev Ecol Evol Syst* **41**: 215–230.
Powell THQ, Hood GR, Murphy MO, Heilveil JS, Berlocher SH, Nosil P *et al.* (2013). Genetic divergence along the speciation continuum: the transition from host race to species in rhagoletis (Diptera: Tephritidae). *Evolution* **67**: 2561–2576.
Pringle EG, Baxter SW, Webster CL, Papanicolaou A, Lee SF, Jiggins CD (2007). Synteny and chromosome evolution in the lepidoptera: evidence from mapping in *Heliconius melpomene*. *Genetics* **177**: 417–426.
Przeworski M (2002). The signature of positive selection at randomly chosen loci. *Genetics* **160**: 1179–1189.
Ramsey J, Bradshaw H Jr, Schemske D (2003). Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* **57**: 1520–1534.
Ranz JM, Casals F, Ruiz A (2001). How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res* **11**: 230–239.
Rieseberg LH (2001). Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**: 351–358.
Roelofs W, Glover T, Tang XH, Sreng I, Robbins P, Eckenrode C *et al.* (1987). Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proc Natl Acad Sci USA* **84**: 7585–7589.
Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
Shaw KL, Mullen SP (2011). Genes versus phenotypes in the study of speciation. *Genetica* **139**: 649–661.
Simonsen KL, Churchill GA, Aquadro CF (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**: 413–429.
Smith HE (1920). Broom Corn, the probable host in which *Pyrausta nubilalis*, Hubn., reached America. *J Econ Entomol* **13**: 425–430.
Storz JF (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* **14**: 671–688.
Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH (2009). Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Mol Biol Evol* **26**: 1341–1355.
Tajima F (1989a). DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-subpopulation model. *Genetics* **123**: 229–240.

Tajima F (1989b). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Team RDC (2013). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.

Turner TL, Hahn MW, Nuzhdin SV (2005). Genomic Islands of Speciation in *Anopheles gambiae*. *PLoS Biol* **3**: e285.

Wadsworth CB, Woods WA, Hahn DA, Dopman EB (2013). One phase of the dormancy developmental pathway is critical for the evolution of insect seasonality. *J Evol Biol* **26**: 2359–2368.

Wanner KW, Nichols AS, Allen JE, Bunger PL, Garczynski SF, Linn CE *et al.* (2010). Sex pheromone receptor specificity in the European corn borer moth, *Ostrinia nubilalis*. *PLoS ONE* **5**: e8685.

West J, Via S (2008). The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol Ecol* **17**: 4334–4345.

Wickham H (2009). *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York Incorporated.

Wu C-I (2001). The genic view of the process of speciation. *J Evol Biol* **14**: 851–865.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)