AFLP Linkage Map of Hybridizing Swallowtail Butterflies, Papilio glaucus and Papilio canadensis

CLAYTON B. WINTER and ADAM H. PORTER

From the Graduate Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, MA 01003 (Winter and Porter); and the Department of Plant, Soil & Insect Sciences, Hatch Laboratory Building, 140 Holdsworth Way, University of Massachusetts, Amherst, MA 01003 (Porter).

Address correspondence to Adam H. Porter at the address above, or e-mail: aporter@ent.umass.edu.

High-density linkage maps provide powerful tools for studying the genetic basis of ecologically relevant adaptations and the genomic scope of introgression. We backcrossed an F1 hybrid male Papilio glaucus/Papilio canadensis tiger swallowtail butterfly to a pure P. glaucus female and constructed amplified fragment length polymorphism linkage maps from the progeny. The paternal map contains 309 markers distributed among 29 linkage groups, with a corrected map distance of 1167 cM (logarithm of the odds [LOD] = 4.0). The average linkage group contained 10.65 ± 4.85 markers separated by 32.7 ± 3.8 cM, with statistically significant clustering. The paternal hybrid map had 18.65% more markers than the maternal P. glaucus map, which provides a rough estimate of the extent of genetic differentiation between the species. The maternal map contains 253 markers among 28 linkage groups, without the X and Y chromosomes. Segregation distortion from expected Mendelian ratios was observed for 94/1096 scored loci (8.6%, P < 0.05). The X chromosome map includes 7 markers spanning 29.3 cM (LOD = 3.0). These naturally hybridizing, female heterogametic species are used to study important questions in the maintenance of species boundaries, sex chromosome introgression, sex-limited mimicry, and host plant use. The map will facilitate research into the physiological, ecological, and evolutionary genetics of these phenomena.

Key words: AFLP, hybrid zone, Lepidoptera, Papilionidae

The central motivation of ecological genetics is to understand microevolutionary processes and consequences in their natural context, as they vary temporally, geographically, and across the genome. Linkage maps provide great power to dissect these phenomena (Stinchcombe and Hoekstra 2008). This is perhaps especially so in systems that show natural hybridization where contemporary gene flow and selection are prone to overcome historical patterns of differentiation in areas of contact, and classical hybrid zone analysis (Barton and Gale 1993) can quantify the strengths of selection and gene flow as they vary both across traits and across the genome.

The eastern tiger swallowtail butterfly, Papilio glaucus, hybridizes with its close relative, Papilio canadensis, in a hybrid zone that extends from Wisconsin to eastern New York. These butterflies show a suite of interesting trait differences, and the ecology, physiology, and classical genetics of many of them have been well studied (Rockey et al. 1987; Hagen and Scriber 1989; Scriber et al. 1995; Andolfatto et al. 2003; Putnam et al. 2007). A great many fall on the sex chromosomes (Sperling 1994). It is perhaps the best-studied system documenting the evolutionary ecology of sex chromosome traits—especially interesting because the Lepidoptera have XY (also known as WZ) female and XX (ZZ) male sex determination—providing an important counterexample in testing models of sex chromosome evolution under sexual selection (Charlesworth et al. 1987; Mank et al. 2007) and the consequences for speciation (Qvarnström and Bailey 2008). Papilio glaucus, in particular, is something of a model species for the study of female-limited mimicry, a trait inherited on the Y chromosome with an X-linked modifier. Further, the filtering of the diverse, species-specific, ecologically relevant traits as they introgress across the hybrid zone provides unusual power to dissect the selection pressures on them (Barton and Gale 1993; Porter 2009); introgression seems quite likely in at least some parts of the genome (Stump et al. 2003; Putnam et al. 2007). We are therefore poised to make a great deal of progress on the ecological genetics of the evolutionary pressures on the sex chromosomes, with the help of a high-density linkage map of the Papilio genome. Here, we describe the backbone of this map based on amplified fragment length polymorphism (AFLP) markers (Vos et al. 1995), using a backcross family to emphasize the bases of the ecological and taxonomic differences between P. glaucus and P. canadensis.
Materials and Methods

We used the AFLP markers (Vos et al. 1995) for linkage mapping. Restriction enzymes cut genomic DNA into fragments of varying length; a known adapter sequence is ligated to the ends of the fragments, then subject to polymerase chain reaction (PCR) with primers that recognize the cut ends and an additional 3 base pairs. This permits 64 primer combinations per pair of restriction enzymes and roughly 30–50 scorable markers each. AFLP markers have been used to generate linkage maps in a variety of taxa (e.g., Liu et al. 2003; Capo-chichi et al. 2004; De Vos et al. 2007; Ukrainetz et al. 2008), including several other Lepidopterans (Tan et al. 2001; Wang and Porter 2004; Jiggins et al. 2005; Tobler et al. 2005).

Insect Rearing

In 2004, we raised 3 generations of tiger swallowtail butterflies. The first generation consisted of the offspring of 35 wild-caught female P. glaucus from Lawrenceville, VA, kept in a population cage with potted cherry (Prunus spp.) and clippings of tulip tree (Liriodendron tulipifera). The resulting larvae were reared on cherry, using potted trees and clippings. Pupae comprising this P1 generation were split into 2 groups: one developed normally in an incubator at 25 °C (14:10-h light:dark cycle), and the second was kept at 10 °C to slow their development until the F1 hybrid generation was available. Eclosing butterflies were marked, segregated by sex, fed to satiation with dilute honey every 2–3 days, and kept at 20 °C to inhibit flight while promoting gonad development.

Females from the 25 °C group were tethered outdoors following Deering and Scriber (2002) in Leyden, MA, where they mated with wild P. canadensis males. The males were subsequently frozen, and the females were allowed to oviposit in separate cages on tulip tree clippings. Neonate F1 larvae were kept segregated by family and transferred to potted cherry to complete development.

An adult male F1 hybrid from a yellow-morph P. glaucus mother was mated twice to a yellow-morph, P1 female that had been reared at 10 °C. This female was caged with a potted cherry, and the resulting 110 backcross larvae (designated 04-033) were raised to pupation on potted cherry. Of these, 50 (mostly pupae, with several emerged adults) were sexed based on external morphology and used for genotyping. We genotyped pupae because female backcrosses segregate for an X-linked, obligate pupal diapause allele (Rockey et al. 1987) from their paternal P. canadensis grandfather, leading to reduced female overwintering survival and subsequent sex-ratio bias that confounds linkage mapping. An additional 7 pupae and one prepupa had indeterminate external morphology and were sexed based on markers segregating in the 50 individuals above; this yielded 58 backcross offspring and the 2 parents for map construction.

Genotyping

DNA was extracted using the phenol–chloroform method, and AFLP markers amplified from EcoRI and Msel restriction enzymes following Wang and Porter (2004). PCR products for 64 primer combinations were multiplexed and visualized on an ABI-377 sequencer by Laragen, Inc. (Los Angeles, CA). We analyzed the resulting gels using Genescan 3.1 (Applied Biosystems, Foster City, CA). The sorting program BinThere (N. Garnhart, http://hcgs.unh.edu/protocol/aflp/AFLPbinthere.html) was used to produce spreadsheet data; that process produces occasional errors when band peaks fall close to bin boundaries, so we manually confirmed the binning against the original gels. Occasional bands that appeared in our negative control (without butterfly tissue) were removed from analysis. For each band, every individual was scored 1 (present), 0 (absent), or ? (unknown). Only unambiguous bands (with few unknowns) were included in the final analysis.

We named markers following Wang and Porter (2004) as illustrated here for the marker ACCAG217. The first 3 letters (ACC) represent the nucleotides appended to the EcoRI primer, the second 3 letters (CAG) represent those for the Msel primer (together, a primer combination), and the number (217) represents the size of the fragment in base pairs.

Linkage Mapping

AFLP markers are typically dominant, resulting in 4 independent linkage maps. Using JoinMap 3.0 (Van Ooijen and Voorrips 2001), we generated a paternal map from markers present in the father, absent in the mother, and present in half of the offspring; this map includes the paternal X chromosome. A maternal map was produced from markers present in the mother, absent in the father, and present in half of the offspring. A third map, where the marker is present in both parents and absent in 25% of the offspring, has lower statistical power and is not reported here. A fourth map is generated from markers that are present in the father, the mother, all the male offspring, and half of the female offspring. This maps independent markers on the paternal X chromosome and has half of the statistical power of the autosomal maps. We also report X- and Y-linked markers (identified by sex-biased inheritance patterns) that did not fit into any of these maps. We used Kosambi centimorgan units for the maps.

Results and Discussion

From 64 primer combinations, 1096 unambiguously scored markers, of which 829 (76%) were polymorphic, comprised the data set. We excluded markers deviating from the expected 1:1 segregation ratio (independent χ² tests, α ≤ 0.05), a conservative threshold given the large number of markers.

The F1 paternal map (Figure 1) contained 309 markers, yielding 29 linkage groups (LOD = 4.0) averaging 10.65 ± 4.85 loci per group. At least one of these groups represents the X chromosome, but it cannot yet be distinguished from the autosomes based on the segregation pattern in this
family. The maternal map (Supplementary Material), which excludes the X chromosome, contained 253 markers and yielded 28 linkage groups (LOD = 4.0), each with at least 4 markers. This suggests a haploid number of $N = 29$ chromosomes, one fewer than the typical chromosome count in the Papilionidae (Robinson 1971); this number may be revised once the independent maps are assembled.

The map distances on the F1 paternal map (Figure 1) total 1038.5 cM including the X chromosome, with a mean of $35.8 \pm 9.8$ cM (median 30 cM) per linkage group. The

Figure 1. The F1 paternal map, derived from 309 AFLP markers distributed into 29 linkage groups (LOD = 4.0) including the X chromosome. Right side: marker designation; left side: distance in cM from the first marker in each linkage group; LG: linkage group. An additional 2 fragments, AAGCAT332 and AACCTA105, were unplaced. These groups do not correspond with linkage groups identified through maternal inheritance (Supplementary Material).
mean interval distance between adjacent markers was 3.71 ± 4.73 cM (median 1.8 cM). The markers at the ends of each linkage group are unlikely to be precisely at the ends of the chromosomes, so we add $29 \times 3.71$ cM to the total, yielding an uncorrected map size estimate of 1253.6 cM. Female *P. glaucus*, like most Lepidoptera (Robinson 1971; but see Wang and Porter 2004), show negligible recombination, so that linkage-group relationships in the maternal map are readily determined, although not the order of the genes. To correct the F₁ paternal map's length, we note that any spurious recombination in the maternal map provides an estimate of mapping error. Such error may be due to
a combination of differing sample sizes among markers (when unknowns are dropped), scoring errors on the gels (including undetected binning errors and occasional mis-scored bands), and occasional AFLP “collisions” where the same primer combination amplifies independent loci of the same fragment size (Gort et al. 2006). We found the maternal map to be 86.5 cM (median distance between markers: 0.003 cM; mean: 0.38 ± 0.82). This suggests a corrected F1 paternal map length estimate of 1167 cM, with a corrected mean of 32.7 cM per linkage group. Although this correction is intended to reduce estimation bias, we expect the size estimate to be revised as more markers are integrated into the map.

A map of the X chromosome (Figure 2) was generated from 9 markers present in the father (heterozygous), the mother (hemizygous), all the male offspring, and half of the female offspring. Female segregation ratios deviated from the expected 1:1 ($\chi^2$ tests, $\alpha \leq 0.05$) in 2 of the 9 markers, which we excluded. The 7 remaining markers mapped to a single linkage group (LOD = 3.0) spanning 29.3 cM. Besides this map, we also identified (Table 1) an additional 8 markers inherited on the X chromosome and 3 markers

---

**Figure 1.** Continued
showing Y-linked inheritance (present in all females and no males).

The *Papilio* linkage map at 1167 cM is much shorter than the maps of other butterflies. The *Heliconius erato* map spans 1430 cM (Kapan et al. 2006) and 1616 cM for the *Heliconius melpomene* map (Jiggins et al. 2005), each with 21 linkage groups, as well as a hybrid *Colias* butterfly map (Wang and Porter 2004), which holds 452 loci spanning 2541.7 cM.

**Species Differentiation**

We expect the F1 paternal map to show more polymorphism than the *P. glaucus* maternal map because the F1 father is heterozygous for diagnostic differences between the taxa in addition to providing a sample of the standing polymorphism within *P. glaucus*. The ratio provides an estimate of the proportion of the genome that differs between these taxa. Lepidopteran chromosomes are all about the same size (Robinson 1971), so estimating that the X chromosome carries the average of 10 markers/chromosome yields 301 paternal, autosomal markers.

**Clustering of Markers**

To test for a nonrandom distribution of markers within linkage groups, the frequency distribution of distances between consecutive markers was compared against an exponential distribution with the same mean, which represents randomly distributed distances among markers. The 280 distances (309 markers among 29 linkage groups) were sorted into bin sizes of 2 cM and frequencies tabulated. Significant clustering was found (*P* = 0.007; Figure 3). Varying the bin size affects resolution of the analysis and therefore its statistical significance, with clustering being less apparent using 1 cM bins (*P* = 0.043) or 3 cM bins (*P* = 0.033). With smaller, more numerous bin sizes, statistical power drops and fewer samples appear in bins holding larger map distances. With fewer, larger bins, meaningful positive and negative deviations in the lower size classes are pooled, masking the clustering pattern. Significant clustering of loci was also observed for the *Colias* map (Wang and Porter 2004).

**Segregation Distortion**

In independent *χ²* tests, 94 (~8.6%) of the 1096 markers showed segregation distortion from expected Mendelian ratios (*P* < 0.05) and were not included in the maps. Because we tested each marker individually for segregation distortion, this proportion is conservative from the perspective of map quality but is likely to overestimate the true number of distorted markers. Using a sequential Bonferroni correction for multiple tests (Rice 1989), we found that the proportion of loci showing unambiguous segregation distortion is on the order of 1.5%. Segregation distortion includes the biologically “real” effects of potentially incompatible loci and meiotic drivers, but with AFLP markers, it also includes spurious collisions (Gort...
et al. 2006) of separate loci with the same size and primer combination.

Prospects

This linkage map provides a valuable tool to the community of researchers studying the genetics, development, physiology, and evolutionary ecology of Papilio. On one hand, we find roughly 16% of the markers in the Papilio map represent species-level differences, implying a significant contribution from the autosomes in the maintenance of species differences. Indeed, differences in larval patterning and adult wing pattern elements between P. glaucus and P. canadensis show automosomal inheritance (Scriber et al. 2003). On the other, a “large X-effect” on the inheritance of species-diagnostic differences is seen in many Lepidoptera and in Papilio, in particular, implying a significant role of sex chromosomes and perhaps sexual selection in their speciation history (Sperling 1994; Prowell 1998). This map has already been useful in demonstrating different divergence times among X chromosome regions (Putnam et al. 2007), and the methods are being applied to Y linked and automosomal portions of the genome.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

National Science Foundation (DEB 9981608 and DEB 0129234).

Acknowledgments

This study forms the basis for MS thesis of C.B.W. in the Organismic and Evolutionary Biology Graduate Program. We especially thank B. Wang for his invaluable advice during data collection and analysis. J. Ahern, G. Boettner, C. Hoeffer, S. Lerman, A. Paradis, J. Seavey, M. Sei, and B. Taft also provided important advice and insights. N. Johnson, B. Normark, M. Riley, and 2 anonymous referees contributed valuable criticism of the manuscript.

References


Received January 22, 2009; Revised July 09, 2009; Accepted July 10, 2009

Corresponding Editor: Rob DeSalle