



HAL
open science

Contrasting genomic and phenotypic outcomes of hybridization between pairs of mimetic butterfly taxa across a suture zone

Jérémy Gauthier, Donna Lisa De-silva, Zachariah Gompert, Annabel Whibley, Céline Houssin, Yann Le Poul, Melanie McClure, Claire Lemaitre, Fabrice Legeai, James Mallet, et al.

► To cite this version:

Jérémy Gauthier, Donna Lisa De-silva, Zachariah Gompert, Annabel Whibley, Céline Houssin, et al.. Contrasting genomic and phenotypic outcomes of hybridization between pairs of mimetic butterfly taxa across a suture zone. *Molecular Ecology*, 2020, 29 (7), pp.1328-1343. 10.1111/mec.15403 . hal-02800429

HAL Id: hal-02800429

<https://hal.science/hal-02800429>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Contrasting genomic and phenotypic outcomes of**
2 **hybridization between pairs of mimetic butterfly taxa**
3 **across a suture zone**
4

5 Jérémy Gauthier^{1,2}, Donna Lisa de Silva³, Zachariah Gompert⁴, Annabel Whibley⁵, Céline
6 Houssin³, Yann Le Poul^{3,6}, Melanie McClure³, Claire Lemaitre¹, Fabrice Legeai¹, James
7 Mallet⁷ and Marianne Elias³

8
9 ¹ Univ Rennes, Inria, CNRS, IRISA, F-35000 Rennes, France

10 ² Geneva Natural History Museum, 1208 Geneva, Switzerland

11 ³ Institut de Systématique, Évolution, Biodiversité, CNRS, MNHN, EPHE, Sorbonne
12 Université, Université des Antilles, Paris, France

13 ⁴ Department of Biology, Utah State University, Logan, UT 84322-5305, USA

14 ⁵ School of Biological Sciences, University of Auckland, Auckland, New Zealand

15 ⁶ Ludwig-Maximilians Universität München, Fakultät für Biologie, Biozentrum,
16 Grosshaderner Strasse 2, 82152 Planegg-Martinsried, Germany

17 ⁷ Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA
18 02138, USA

19
20 Corresponding Author:

21 Jérémy Gauthier

22 Geneva Natural History Museum, 1208 Geneva, Switzerland

23 mr.gauthier.jeremy@gmail.com
24
25
26

27 **Abstract.**

28 Hybrid zones, whereby divergent lineages come into contact and eventually hybridize, can
29 provide insights on the mechanisms involved in population differentiation and reproductive
30 isolation, and ultimately speciation. Suture zones offer the opportunity to compare these
31 processes across multiple species. In this paper we use reduced-complexity genomic data to
32 compare the genetic and phenotypic structure and hybridization patterns of two mimetic
33 butterfly species, *Ithomia salapia* and *Oleria onega* (Nymphalidae: Ithomiini), each
34 consisting of a pair of lineages differentiated for their wing colour pattern and that come into
35 contact in the Andean foothills of Peru. Despite similarities in their life history, we highlight
36 major differences, both at the genomic and phenotypic level, between the two species. These
37 differences include the presence of hybrids, variations in wing phenotype, and genomic
38 patterns of introgression and differentiation. In *I. salapia*, the two lineages appear to hybridize
39 only rarely, whereas in *O. onega* the hybrids are not only more common, but also genetically
40 and phenotypically more variable. We also detected loci statistically associated with wing
41 colour pattern variation, but in both species these loci were not over-represented among the
42 candidate barrier loci, suggesting that traits other than wing colour pattern may be important
43 for reproductive isolation. Our results contrast with the genomic patterns observed between
44 hybridizing lineages in the mimetic *Heliconius* butterflies, and call for a broader investigation
45 into the genomics of speciation in Ithomiini - the largest radiation of mimetic butterflies.

46

47 **Keywords.** Hybridization, Ithomiini, introgression, differentiation, admixture mapping,
48 mimicry

49

50 **Introduction**

51 Speciation is the ultimate process responsible for the considerable biological diversity
52 observed on Earth. Hybrid zones, whereby divergent lineages come into contact and
53 potentially hybridize, can provide insights into the mechanisms involved in population
54 differentiation and reproductive isolation, and ultimately speciation (Barton & Hewitt, 1989;
55 Ravinet et al., 2017; Safran & Nosil, 2012). When hybrid zones span an environmental
56 transition, populations across the hybrid zone diverge not only because of genetic drift but
57 also due to local adaptation to different environments. Over time, drift and selection can lead
58 to the emergence of barriers to gene flow that increase reproductive isolation, resulting in
59 heterogeneous patterns of differentiation and introgression across the genome (Barton &
60 Bengtsson, 1986; Ravinet et al., 2017; Safran & Nosil, 2012). Genomic regions with low rates
61 of introgression are more likely to be associated with divergent selection and reproductive
62 isolation (Gompert & Buerkle, 2009; *Heliconius* Genome Consortium, 2012; Mallet, 2005;
63 Mallet & Barton, 1989; Jay et al., 2018). Assessing the genetic structure of hybrid zones can
64 therefore shed light on the evolutionary processes at play during the early stages of speciation
65 by revealing the number and distribution of loci presenting deviant patterns of differentiation
66 and introgression compared to the genome-wide average (Bierne, Welch, Loire, Bonhomme,
67 & David, 2011). Additionally, while genetic mapping of adaptive traits has classically relied
68 on controlled crosses, which cannot be performed in many organisms, hybrid zones enable the
69 application of admixture mapping approaches that take advantage of natural mixing and
70 recombination to investigate the genetic basis underlying adaptive phenotypic variation
71 (Buerkle & Lexer, 2008; Gompert & Buerkle, 2013; Pallares, Harr, Turner, & Tautz, 2014).

72

73 Many studies have focused on hybrid zones to unravel the processes generating local
74 adaptation (e.g. Jones et al., 2012; Larson, Andrés, Bogdanowicz, & Harrison, 2013; Soria-
75 Carrasco et al., 2014) and reproductive isolation (e.g. Christe et al., 2016; Teeter et al., 2008)

76 and have revealed different patterns (Barton & Hewitt, 1985, 1989; Gompert & Buerkle,
77 2009; Kronforst et al., 2013; Rieseberg, Whitton, & Gardner, 1999; Teeter et al., 2010; Via &
78 Hawthorne, 2002). These differences may stem from differences in the organisms studied, but
79 may also stem from differences in the environmental conditions faced by populations of these
80 organisms, such that comparative interpretations are rather limited. To investigate the
81 repeatability of genetic and phenotypic differentiation, reproductive isolation and
82 introgression patterns across incipient species diverging under similar environmental
83 conditions are needed. Suture zones are areas where multiple recently diverged pairs of taxa
84 come into contact and hybridize (Remington, 1968), and typically span a sharp environmental
85 gradient or a dispersal barrier (Dasmahapatra, Lamas, Simpson, & Mallet, 2010; Endler,
86 1977; Moritz et al., 2009). With replicated pairs of divergent and hybridizing lineages, suture
87 zones offer an exceptional opportunity to compare levels and patterns of hybridization and
88 reproductive isolation in relation to genomic and phenotypic divergence in a common
89 environmental setting (Moritz et al., 2009; Nosil, Funk, & Ortiz-Barrientos, 2009; Rissler &
90 Smith, 2010).

91 Müllerian mimicry in butterflies, in which multiple defended species locally converge
92 on warning wing colour patterns and form mimicry 'rings' (Bates, 1862; Muller, 1879),
93 provides an excellent system to unravel the mechanisms underlying adaptation and speciation.
94 Two large neotropical mimetic butterfly tribes, Heliconiini and Ithomiini (Nymphalidae, 77
95 and 393 species, respectively) are particularly well suited. Heliconiine and ithomiine species
96 typically comprise multiple geographical subspecies that differ in colour patterns (Brown,
97 Sheppard, & Turner, 1974). Because warning colour patterns in mimetic butterflies are under
98 strong positive frequency-dependent selection locally (Kapan, 2001; Mallet & Barton, 1989),
99 divergent mimetic subspecies are often separated by narrow hybrid zones maintained by
100 migration-selection balance (Mallet & Barton, 1989). Mimicry generates postzygotic

101 reproductive isolation via higher predation on intermediate-patterned, non-mimetic hybrids
102 (Merrill et al., 2012) and prezygotic reproductive isolation if there is also assortative mating
103 for colour pattern among subspecies (Jiggins, Naisbit, Coe, & Mallet, 2001; McClure et al.,
104 2019; Merrill et al., 2011; 2012). Mimicry is therefore a strong ecological driver of speciation,
105 and is believed to have triggered the diversification of large radiations of heliconiine and
106 ithomiine butterflies (Jiggins et al., 2001; Kozak et al., 2015). Studies of genetic
107 differentiation and the basis of colour pattern variation in mimetic butterflies have almost
108 exclusively focused on heliconiine butterflies (particularly the genus *Heliconius*), where a few
109 major-effect genes (dubbed the mimicry ‘toolkit’, (Joron et al., 2006)) have been found to
110 control wing pattern variation (Martin et al., 2012; Mazo-Vargas et al., 2017; Nadeau et al.,
111 2016; Reed et al., 2011; Westerman et al., 2018) and to be highly differentiated across hybrid
112 zones, while the rest of the genome seems highly permeable (Nadeau et al., 2014). By
113 contrast, because of practical limitations (difficulties in maintaining captive populations and
114 making controlled crosses), much less is known about population genetics of Ithomiini
115 species (but see Dasmahapatra, Lamas, Simpson, & Mallet, 2010; McClure & Elias, 2016;
116 McClure et al., 2019), let alone of the genetics of wing pattern variation. Yet, Ithomiini
117 numerically dominate forest communities of day-flying Lepidoptera and are believed to be
118 some of the main drivers of both Müllerian and Batesian mimicry (whereby palatable species
119 mimic unpalatable aposematic species) among Lepidoptera in the Neotropics (Bates, 1862;
120 Beccaloni, 1997; Muller, 1879). Because of the ecological importance of Ithomiini, shedding
121 light on how populations are structured across hybrid zones and elucidating the genetic basis
122 of wing pattern variation would significantly advance our understanding of adaptation and
123 speciation in mimetic butterflies.

124

125 The foothills of the Andes in the region of Tarapoto are transitional between lowland
126 rainforest and mid-elevation mountain forest. This area is a major suture zone for a range of
127 organisms (Roberts et al., 2006; Smith et al., 2014; Weir, 2006) including heliconiine and
128 ithomiine mimetic butterflies (Dasmahapatra, Lamas, Simpson, & Mallet, 2010; Nadeau et
129 al., 2014; Whinnett et al., 2005) that harbour divergent wing colour patterns across the suture
130 zone. Here we take advantage of the Tarapoto suture zone to assess the patterns of genomic
131 and phenotypic divergence and infer the extent of reproductive isolation across the hybrid
132 zone for two widespread ithomiine species, *Ithomia salapia* and *Oleria onega*, that each
133 harbour divergent wing colour patterns across the Tarapoto suture zone. The two species have
134 broadly similar life histories (e.g., forest habitats, mimicry, Solanaceae hostplants) and
135 although they are somewhat differently distributed throughout the Neotropics, in the area of
136 Tarapoto their populations have similar distributions. In this area, *I. salapia* comprises two
137 subspecies: *I. salapia aquinia*, on the Amazonian side, which has a transparent yellow colour
138 pattern surrounded by an orange and black line with small white dots; and *I. salapia derasa*,
139 on the Andean side above about 500m alt., has a similar yellow pattern but surrounded by a
140 thick black line with large white dots (Figure 1). In addition to being found on the Amazonian
141 side of the Escalera mountains, *I. salapia aquinia* is also found on the other side of the
142 mountains, in the lowlands of the Río Mayo valley near Tarapoto.

143 On the Andean side (including in the lowlands of the Río Mayo) *O. onega* ssp. nov. 2
144 has translucent white wings with black patterning that splits the apical part of the forewing
145 into two white ‘windows’, while in the Amazonian subspecies *O. onega janarilla* the
146 forewing black patterning nearly splits the apical part of the forewing into four windows
147 (Figure 1).

148 Thus, whereas *I. salapia* has a strictly altitude-based distribution of populations, *O.*
149 *onega* has populations that differ geographically East-West, independently of altitude. For

150 each species, the two subspecies belong to distinct mimicry rings, which they numerically
151 dominate. Therefore, *I. salapia* and *O. onega* likely play a major role in divergence and
152 maintenance of their respective mimicry rings. Individuals with intermediate patterns are
153 sometimes found in the contact zone, suggesting that occasional interlineage hybrids are
154 produced.

155

156 In this paper we carry out population genomic analyses and quantify wing colour
157 pattern variation in samples from Andean, Amazonian and intermediate populations of *I.*
158 *salapia* and *O. onega* to address the following questions: (1) what are the overall patterns of
159 differentiation, admixture and introgression between subspecies for each taxon pair? (2) To
160 what extent does introgression vary across the genome, and are genetic regions associated
161 with colour pattern among those that exhibit higher differentiation and reduced introgression?
162 (3) How repeatable are the patterns of differentiation, introgression and genotype-phenotype
163 association across the two taxon pairs?

164

165 **Material and Methods**

166 *Sampling*

167 Sampling was performed in five study sites in the region of Tarapoto in Peru (Figure 1a;
168 details in Supplementary Table 1: gives each site, GPS coordinates and the number of each
169 sex of *I. salapia* and *O. onega* sampled, before and after filtering). One hundred and twelve
170 (112) *I. salapia* specimens were sampled from five localities, and 149 *O. onega* specimens
171 were sampled in five localities.

172

173 *Genotyping by sequencing (GBS)*

174 DNA was extracted from ¼ of thorax of each individual using the Qiagen DNeasy blood and
175 tissue kit, following the manufacturer protocol. We generated reduced genomic complexity
176 libraries for each specimen using a GBS (genotyping by sequencing) approach (Gompert et
177 al., 2012; Parchman et al., 2012). Briefly, genomic DNA was digested with the restriction
178 endonucleases EcoRI and MseI and resulting fragments were ligated to double-stranded
179 adaptor oligonucleotides. These adaptors consisted of the Illumina sequencing priming sites
180 followed barcodes that allow for the identification of sequences for each individual. These
181 barcodes allowed us to multiplex all individuals into one library. Sequencing of the library
182 was completed by the National Center for Genome Research (Santa Fe, NM, USA) on an
183 Illumina HiSeq platform; 100 base single-end sequencing reads were generated.

184

185 *SNP calling*

186 First, sequencing primers were removed, sequences were demultiplexed and associated with
187 each individual based on internal barcode sequences. SNP-calling was performed on samples
188 from each species separately using `DiscoSnp-RAD`, a *de novo* reference-free and assembly-
189 free method (Gauthier et al., 2017; Uricaru et al., 2015). SNPs are identified from particular
190 arrangements in the De Bruijn graph built using a k-mer size of 31 and a minimal coverage of
191 2 for each allele (Gauthier et al., 2017; Uricaru et al., 2015). Individuals with more than 90%
192 missing genotypes were excluded (Supplementary Table 1) resulting in a final dataset
193 consisting of 105 samples for *I. salapia* and 142 for *O. onega*. SNPs scored in at least 80% of
194 the samples (i.e. sites with < 20% missing data) and with a minor allele frequency above 0.01
195 were retained using `vcftools` (Danecek et al., 2011) resulting in a dataset of 17,779 SNPs
196 for *I. salapia* and 15,894 SNPs for *O. onega*.

197

198 *Population structure analyses*

199 Population genetic structure was investigated using a subset of SNPs, where only one SNP
200 per GBS locus was considered so as to minimize the effects of linkage disequilibrium that
201 would occur within loci, as often recommended (Falush, Stephens, & Pritchard, 2003). A total
202 of 8,219 SNP for *I. salapia* and 5,133 SNP for *O. onega* were retained. To investigate genetic
203 structure, we used principal component analysis implemented for genetic data in the
204 adegenet R package (Jombart & Ahmed, 2011). We used Bayesian admixture analysis
205 implemented in Structure (Pritchard, Stephens, & Donnelly, 2000) to estimate admixture
206 proportions, that is, the proportion of each individual's genome inherited from each of K
207 hypothetical source populations. We ran analyses with K from 1 to 6 with 3 independent
208 Markov chains each, using 200,000 steps and including 10,000 burn-in steps. We checked the
209 results obtained in each run to verify convergence of the chains to a stable posterior
210 distribution. The most likely number of clusters was identified using Evanno's method
211 (Evanno, Regnaut, & Goudet, 2005) implemented in Structure Harvester (Earl &
212 vonHoldt, 2012).

213

214 *Genome-wide introgression and estimates of differentiation*

215 To investigate introgression among each population pair and to search for loci potentially
216 associated with reproductive isolation, we used a genomic cline approach using bgc
217 (Gompert & Buerkle, 2012). Loci acting as barriers to gene flow and linked regions should
218 exhibit reduced introgression into the foreign genomic background. Locus-specific
219 introgression is characterized by the probability ϕ of being inherited from a given parental
220 population (here, the Amazonian lineage; the probability of being inherited from the Andean
221 lineage is therefore $1 - \phi$). These probabilities are compared to the genome-wide average

222 probability, which corresponds to the hybrid index. Introgression patterns can be summarized
223 by two locus-specific genomic cline parameters: α , the genomic center parameter, and β , the
224 genomic cline rate parameter. The genomic cline center parameter α specifies an increase
225 (positive values of α) or decrease (negative values of α) in the probability of ancestry of the
226 focal population (here, the Amazonian lineage). Positive or negative α values denote an
227 asymmetry in the direction of introgression with hybrids having increased or decreased
228 ancestry from one or the other ancestral lineage, respectively. The genomic cline rate
229 parameter β specifies the cline steepness, with an increase (positive values of β) or decrease
230 (negative values of β) in the rate of transition from low to high probability of ancestry
231 (Gompert & Buerkle, 2011). Positive or negative β values are associated with a high or low
232 level of gene flow, respectively. We estimated the posterior probability distribution of hybrid
233 indices and cline parameters with `bgc`. MCMC of 50,000 steps including 10,000 burn-in
234 steps for *I. salapia* samples and 100,000 with 30,000 burn-in steps for *O. onega* samples were
235 used to reach mixing, and convergence was verified graphically by plotting log-likelihood
236 distributions. For a given SNP, outlier introgression from the genome-wide average was
237 identified as credible when the 95% credible intervals of the cline parameters α and β
238 excluded zero. These SNPs deviating from global pattern should reflect unusual patterns of
239 evolution acting on these loci. We used the admixture model implemented in `entropy`
240 (version 1.2) to estimate admixture proportions and intertaxon ancestry (Gompert et al. 2014).
241 This model explicitly estimates the proportion of each individuals' genome where the two
242 allele copies are derived from different source populations (i.e., the proportion of the genome
243 with intertaxon ancestry). `entropy` also incorporates uncertainty in genotypes due to limited
244 sequence coverage and sequencing errors. We fit the model using Markov chain Monte Carlo
245 (MCMC). We ran the MCMC algorithm three times with 15,000 iterations following a 5,000

246 iteration burnin, and with a thinning interval of 5. We assumed the number of source
247 populations (K) was two.

248

249 Genome-wide weighted and per-SNP genetic differentiation F_{ST} were estimated using
250 Weir and Cockerham's method (Weir & Cockerham, 1984) implemented in `vcftools`
251 (Danecek et al., 2011). To do so, only samples from parental lineages, that is, samples from
252 the initial localities distant from the hybrid zone, were kept. To identify outlier SNPs with
253 elevated genetic differentiation, the threshold was fixed to the 95th percentile of F_{ST}
254 distribution obtained by random sampling with replacement of 100,000 values.

255

256 *Wing pattern analyses*

257 Photographs of dorsal and ventral sides of detached wings of 90 and 94 of the
258 genotyped specimens of *I. salapia* and *O. onega*, respectively, were taken with a Nikon D90
259 digital camera and a 105 mm lens on a white background with a piece of millimeter paper for
260 scale. For each specimen, dorsal and ventral patterns of fore- and hindwings were quantified
261 using Colour Pattern Modeling (CPM, (Le Poul et al., 2014) as follows: wings were first
262 extracted from their background, resulting in eight images per specimen (2 wings [forewing
263 and hindwing] x 2 lateral sides [left/right] x 2 vertical sides [ventral/dorsal]); for each image,
264 wing pattern was described by semi-automatically categorizing wings into a finite number of
265 colours (yellow, black, orange and white for *I. salapia*; white, black and orange for *O. onega*).
266 Damaged wings were discarded, and when left and right wings were available only one
267 randomly chosen side was used in subsequent analyses. Homologous wings were then aligned
268 according to both shape and pattern (Le Poul et al., 2014), and a binary principal components
269 analysis based on one-hot encoding of colours (i. e., where each colour is encoded by a string

270 of bits among which only one takes the value 1) was performed on the colour of homologous
271 pixels shared by all wings. Principal component (PC) scores were used as a quantitative
272 measure of colour pattern in subsequent analyses.

273

274 *Admixture association mapping*

275 To identify SNPs associated with variation in wing patterns, we performed association
276 mapping using the Genome-wide Efficient Mixed Model Association tool (GEMMA) (Zhou &
277 Stephens, 2012, 2014). We used the multivariate linear mixed model (mvLMM) to test
278 marker associations with multiple phenotypes and to estimate genetic correlations among
279 complex phenotypes. To do so, we retained all wing pattern PCs that explained more than 1%
280 of the variation in each species as variables. This included 14 variables accounting for 57.1%
281 of the variation in wing pattern for *I. salapia* and 18 variables explaining 54.9% of the
282 variation for *O. onega*. Both the variation linked to sex and the confounding effect of
283 population structure (Freedman et al., 2004; Price et al., 2006) were integrated into the models
284 by implementing a relatedness matrix between individuals generated using GEMMA (Zhou &
285 Stephens, 2012) and the first population structure PC (obtained using *adegenet* R package
286 (Jombart & Ahmed, 2011)) as a covariate. Analyses were carried out using the option `-lmm 1`
287 to perform a Wald test evaluating the probability of the null hypothesis that the marker effect
288 sizes for all phenotypes were zero. For the identification of SNPs significantly associated with
289 wing pattern, the threshold was fixed to a p-value adjusted using (i) a classical Bonferroni
290 correction, which divides the significance threshold by the number of multiple comparisons,
291 that is, the number of molecular markers multiplied by the number of variables, and (ii) using
292 the false discovery rate (FDR) method (Benjamini & Hochberg 1995). We considered SNPs
293 identified by both of these correction methods as significantly associated with wing pattern.

294 To test whether outlier SNPs by the three approaches, i.e. differentiation genome-scan,
295 introgression pattern and admixture mapping, are distributed randomly or if there is
296 enrichment in shared outlier SNPs, we used two methods: a Pearson's Chi-squared test to
297 compare shared outlier to a random distribution and a bootstrap of 1,000 random samplings
298 among SNPs to estimate confidence interval of such shared SNPs.

299

300 *Similarities with other Lepidoptera genomes*

301 Loci containing SNPs identified as outliers in the genetic differentiation or the
302 differential introgression approaches were screened by BLAST against all annotated butterfly
303 reference genomes in LepBase v4 (Challis, Kumar, Dasmahapatra, Jiggins, & Blaxter, 2016)
304 to investigate the gene content of homologous genomic regions. This was performed using the
305 BLASTn tool available on the LepBase platform (Priyam et al., 2019). Best hits, their location
306 in reference genomes and genes were then investigated manually.

307

308 **Results**

309 Our reduced complexity genotype by sequencing approach coupled with Illumina sequencing
310 produced 77.8 million reads distributed relatively evenly between samples with a mean of
311 340,940 reads per individual (sd: 157,991) for *I. salapia* samples and a mean of 270,602 reads
312 per individual (sd: 103,734) for *O. onega*. From this sequencing data, SNP calling and
313 filtering steps resulted in final datasets of 17,779 SNPs from 6,972 loci for *I. salapia* samples
314 and 15,894 SNPs from 4,524 loci for *O. onega* samples.

315

316 *Population genetic structure*

317 For *I. salapia* we considered two parental lineages, an Amazonian lineage corresponding to
318 subspecies *aquinia* sampled from two sites, Km-26 Yurimaguas-Tarapoto (1) and San Miguel
319 de Achinamiza (2), and an Andean lineage corresponding to subspecies *derasa* and sampled
320 from one site Puente Aguas Verdes (5). Between these sites a zone that is geographically and
321 altitudinally intermediate and where those lineages are in contact was sampled at two sites
322 (sites Km-42 Tarapoto-Yurimaguas (i3) and La Florida (i4), hybrid zone) (Figure 1a). The
323 genetic structure, identified by multivariate analyses on genetic data (Figure 2a) as well as
324 Structure (Figure 2b), highlight that the samples likely segregate in 3, best clustering, or 2
325 groups (Evanno's method: higher ΔK of 5037.42 for $K=3$ and 718.12 for $K=2$). These results
326 shows that the two parental lineages are distinct and show a low level of admixture. They
327 have a weighted genome-wide differentiation F_{ST} of 0.177 between the two. In the
328 intermediate populations, we did not find putative F1 hybrids (i.e. individuals with hybrid
329 index close to 0.5 and high intertaxon ancestry). Rather, four individuals (subpopulation 3.1,
330 Figure 2b) were genetically similar to individuals from the parental Amazonian *aquinia*
331 lineage, and the 28 other samples had equal levels of admixture with the majority of their
332 genetic content associated with the Andean *derasa* population (Figure 2b). Hybrid indices
333 estimated for these individuals is also indicative of this (Figure 2c).

334 The distribution of *O. onega* hybrids in the study area is somewhat similar to that observed in
335 *I. salapia*. The genetic structure, identified by multivariate analyses on genetic data (Figure
336 2a) as well as Structure (Figure 2b), highlight the samples likely segregate in 2, best
337 clustering, or 3 groups (Evanno's method: highest ΔK of 9887.07 for $K=2$ and 7.76 for $K=3$).
338 This species also consists of two parental lineages corresponding to an Amazonian
339 subspecies, *O. onega janarilla*, collected from two sites, Km-26 Yurimaguas-Tarapoto (1)
340 and San Miguel de Achinamiza (2) and an Andean subspecies, *O. onega* ssp. nov. 2
341 (Gallusser, 2002; Dasmahapatra, Lamas, Simpson, & Mallet, 2010), sampled at Puente

342 Serranoyacu (5). In addition to these sites, a geographically and altitudinally intermediate
343 zone was also sampled (sites Shapaja-Chazuta (o3) and a site spanning Quebrada Yanayacu to
344 Laguna del Mundo Perdido (o4), hybrid zone) (Figure 1b). The two parental lineages are
345 genetically divergent with a weighted genome-wide differentiation F_{ST} of 0.372, more than
346 twice as high as between *I. salapia* parental lineages. The samples from the hybrid zone
347 comprise a mix of individuals with low and intermediate levels of admixture. Most
348 individuals have a low level of admixture (39 out of 46 had $q < 0.2$), and are genetically
349 closest to the Amazonian parental lineage. The stronger link with the Amazonian lineage
350 suggests a directionality in hybridization different from that of the *I. salapia* hybrid zone.
351 Three individuals show hybrid indices that suggest almost equal contributions from each
352 parental lineage (Figure 2a,b,c). One of these has a high level of intertaxon ancestry, which
353 suggests it is an F1 hybrid. The two other individuals have lower heterozygosity, which is
354 consistent with recent backcrossing. In conclusion, both species show some evidence of gene
355 flow and introgression, but both also exhibit strongly bimodal phenotypes (sensu Jiggins &
356 Mallet, 2000) in the region of the hybrid zone, suggesting strong reproductive isolation in
357 both species. However, *O. onega* displays somewhat more evidence of ongoing hybridization
358 and gene flow than *I. salapia*.

359

360 *Genomic patterns of introgression and differentiation*

361 Introgression varied across the genome for each taxon pair, with distinct patterns of
362 introgression for the two studied species, *I. salapia* and *O. onega*, as demonstrated by the
363 distributions of the genomic cline center (α) and rate parameters (β) (Figure 3). In each
364 species the center parameter (α) is highly variable, with point estimates (posterior median)
365 ranging from -3.067 to 2.864 for *I. salapia* and from -5.785 to 5.907 for *O. onega*. With
366 respect to α , many loci show introgression patterns that differ credibly from the genome-wide

367 average. Loci with positive center parameter α are more likely to be inherited from the
368 Amazonian parental lineage than the rest of the genome. Conversely, loci with negative α are
369 more likely to be inherited from the Andean lineage. For *I. salapia*, 2125 SNPs (11.95 % of
370 all SNPs) have a center parameter α that differ from the genome-wide average, consistent
371 with different levels of introgression than the average of the rest of the genome. Among the
372 SNPs with credible evidence of differential introgression, most of them (1637 out of 2125, i.
373 e., 77%) have excess ancestry from the Amazonian populations whereas the genome-wide
374 average is more closely associated with the Andean population (Figure 2b,c). In *O. onega*,
375 where intermediate populations are genetically closer to the Amazonian lineage, 2146 SNPs
376 (13.50 % of all SNPs) have a center parameter α that significantly deviates from the genome-
377 wide average. The distribution is reversed compared to *I. salapia*, however, with most of
378 those SNPs (1406 out of 2146, i. e., 66%) having a lower probability of being inherited from
379 the Amazonian parental lineage. This is expected, as there is more statistical power to detect
380 differential introgression of alleles of the less common (i.e., minor) ancestry type.

381 Regarding the genomic cline rate parameter (β), the profiles for the two species are markedly
382 different. In *I. salapia* species, β hardly shows any variation, ranging from -0.681 to 0.738,
383 whereas the variation observed in *O. onega* is higher by tenfold, ranging from -7.710 to 8.440.
384 Such a large variation in *O. onega* is probably due to the heterogeneity of the hybridization
385 profiles observed in the hybrid zone (Figure 2b). While in *I. salapia* no SNPs are different
386 from the genome-wide expectation, in *O. onega*, 1274 SNPs have a genomic cline rate (β)
387 credibly different from the genome-wide pattern, with 447 and 827 SNPs having higher and
388 lower values, respectively. These 447 SNPs with steeper introgression patterns than the
389 genome-wide average are characteristic of SNPs putatively associated with barrier loci (i.e.,
390 in LD with barrier loci).

391 This difference observed between *I. salapia* and *O. onega* can be explained, at least in part, by
392 the fact that there is hardly any variation among hybrids in *I. salapia* in our sampling. Perhaps
393 an even larger sample than at present would show additional variation. In any case, the
394 absence of variation in the *I. salapia* hybrids is a result in itself but limits the capacity to
395 detect variation in patterns of introgression.

396 Genetic differentiation (F_{ST}) between parental lineages was heterogeneous across the genome,
397 ranging from ~0 to 1.000 with a weighted mean value of 0.177 for *I. salapia* and from ~0 to
398 1.000 with a weighted mean value of 0.372 for *O. onega*. The genome scan also identifies
399 outlier SNPs that deviate from the genome-wide distribution and have unusually high levels
400 of differentiation. In *I. salapia*, the 95th percentile threshold corresponds to an F_{ST} value of
401 0.415, and was exceeded by 890 SNPs. In *O. onega*, the 95th percentile threshold corresponds
402 to an F_{ST} value of 0.686, and was exceeded by 795 SNPs.

403

404 *Phenotypic variation*

405 For both species, the first principal component on colour pattern separates Andean from
406 Amazonian lineages (Fig. 4). However, the two species differ in where along this axis
407 specimens from intermediate populations fall. For *I. salapia*, most specimens in the hybrid
408 zone cluster with the Andean *I. salapia derasa*, while only four (with a predominantly
409 Amazonian genetic background) cluster with Amazonian *I. salapia aquinia* (Figure 4). No
410 individual in this sample has a markedly intermediate color pattern along this first principal
411 component. By contrast, most *O. onega* from the hybrid zone have an intermediate position
412 along the first principal component. For both species, the second axis highlights variation
413 associated with sex and as a result males and females are segregated along this axis. While
414 sexual dimorphism is moderate and of the same magnitude in both lineages of *O. onega*, it is
415 more pronounced in the Andean lineage *I. salapia derasa* and virtually absent in the

416 Amazonian lineage *I. salapia aquinia*. Sex was therefore included as a factor in our
417 phenotype-genotype analyses, along with genetic structure.

418

419 *Phenotype-genotype relationship and association mapping*

420 Despite the low variability among hybrid individuals from the hybrid zone, especially for *I.*
421 *salapia*, we attempted to associate genetic variation in specific SNPs to wing pattern
422 variation. All principal components (PCs) explaining at least 1% of total phenotypic variation
423 were included in association mapping (i.e. 14 variables for *I. salapia*, jointly explaining
424 57.1% of the variance, and 18 for *O. onega*, jointly explaining 54.9% of the variance).
425 Association mapping using GEMMA revealed several PCs for which a large phenotypic
426 variation is explained, specifically more than 80% phenotypic variation explained (pve) for
427 7/14 PCs for *I. salapia* and for 4/18 PCs for *O. onega*. When combining all PCs that were
428 retained (i.e., all those that explained at least 1% of the phenotypic variation), 59.3% of the
429 phenotypic variation is explained for *I. salapia* and 26.3% for *O. onega* (Supplementary
430 Figure 2). The PCs explaining most genetic variation are also those harbouring the largest
431 proportions of wing pattern variation. Specifically, 84.3% of the wing pattern variation is
432 explained for *I. salapia* and 65.3% for *O. onega*. The multivariate linear mixed model of
433 GEMMA performs tests to evaluate the probability that SNPs are associated with phenotypic
434 variation and outputs the corresponding p-value resulting from a Wald test. Retaining
435 significant SNPs concurrently in Bonferroni p-value correction and FDR approach, 88 SNPs
436 (0.49% of all SNPs) were significantly associated with wing patterns in *I. salapia* and 109
437 SNPs (0.69% of all SNPs) in *O. onega* (Figure 5).

438 We then focused on the differentiation and introgression patterns of SNPs associated with
439 wing pattern. A very small number of SNPs, 17 for *I. salapia* and 4 for *O. onega*, combine

440 strong association with wing pattern and high levels of differentiation, but in *I. salapia* the
441 SNP with the strongest association with wing pattern is also an outlier in genome scan for
442 differentiation (Figure 5).

443 With all of these approaches combined, we observe both differences and similarities between
444 *I. salapia* and *O. onega* in the number of SNPs similarly identified as outliers by multiple
445 approaches. The proportion of SNPs combining both high differentiation levels (F_{ST}) and
446 differential introgression (α), and which are characteristic of loci potentially involved in
447 adaptation in hybrid samples, are of the same order in *I. salapia* and in *O. onega* (1.81%,
448 321/17,779, for *I. salapia*, and 1.59%, 253/15,894, for *O. onega*, Figure 6), and higher than
449 expected at random (Pearson's Chi-squared test, p-value = 1.32e-114 and 95% CI = [103, 135]
450 for *I. salapia* and Pearson's Chi-squared test, p-value = 2.99e-54 and 95% CI = [92, 123] for
451 *O. onega*). SNPs previously identified as potentially involved in adaptation and adaptive
452 introgression, combining high differentiation levels (F_{ST}) and differential introgression (α), do
453 not have a specific enrichment in SNPs significantly associated with wing pattern. In *I.*
454 *salapia* and *O. onega* the numbers of SNPs that fit this description are low, respectively only
455 two and one SNPs, and do not differ from a random distribution (Pearson's Chi-squared test,
456 p-value = 0.74 and 95% CI = [0, 4] for *I. salapia* and Pearson's Chi-squared test, p-value =
457 0.57 and 95% CI = [0, 4] for *O. onega*). The main difference between the two species consists
458 in the SNPs with differential positive genomic cline rate values (β) and potentially involved in
459 reproductive isolation. None of the SNPs in *I. salapia* have positive β while a non-trivial
460 proportion do in *O. onega*, i.e. 447 SNPs or 2.81%. Among these SNPs only a small fraction
461 is also significantly associated with wing pattern variation (5) and not enriched compared to a
462 random distribution (Pearson's Chi-squared test, p-value = 0.26 and 95% CI = [1, 6]).
463 Moreover, among the SNPs involved in both, high differentiation level and positive β (71)

464 and potentially involved in reproductive isolation, none of them is associated with wing
465 pattern variations (Figure 6).

466

467 *Similarities with other Lepidoptera genomes*

468 We used BLASTn to investigate the potential functional roles of the loci carrying the SNPs
469 highlighted by the approaches listed above. We identified homologous regions in the *D.*
470 *plexippus* genome for loci containing SNPs significantly associated with wing patterns (12/67
471 loci for *I. salapia* and 21/99 for *O. onega*; Supplementary Table 3). None of the genes known
472 to control colour pattern variation in Lepidoptera and identified in the *D. plexippus* genome,
473 i.e. *optix*, *cortex*, *WntA*, *ebony* and *aristaless*, were identified.

474 On the other side the BLASTn of loci with SNPs potentially involved in adaptation and
475 adaptive introgression or reproductive isolation highlighted candidate genomic region and
476 gene in the genome of *D. plexippus* (37/277 loci for *I. salapia* and 66/218 for *O. onega*
477 potentially involved in adaptation and adaptive introgression and 73/389 loci potentially
478 involved in reproductive isolation for *O. onega*; Supplementary Table 2). We here report the
479 list of scaffold containing these loci of interest which are potential candidates for genes
480 involved in local adaptation and reproductive isolation, and on which further functional
481 analyses could be performed to investigate underlying biological functions (Supplementary
482 Table 2).

483

484 **Discussion**

485 The comparison of genome-wide patterns of genetic differentiation, introgression and
486 genotype-phenotype associations in two species, *I. salapia* and *O. onega*, that face similar
487 environmental transitions revealed some surprisingly large phenotypic and genomic

488 differences. Below, we discuss potential reasons for the differences observed in light of
489 biological and ecological information.

490

491 *Genomic and phenotypic differentiation patterns across the Tarapoto suture zone: similarities*
492 *and differences*

493 Both *I. salapia* and *O. onega* are distributed across an important environmental gradient in the
494 region of Tarapoto in Peru, and both species consist of an Andean and Amazonian lineage.

495 Our analysis of wing pattern variation confirms that Amazonian and Andean lineages of both
496 species are phenotypically different (as is seen by the human eye) and also reveals a subtle
497 sexual dimorphism not readily discernible.

498 Phenotypic differentiation between populations of each species is associated with strong
499 overall genomic differentiation, especially in *O. onega*. These findings are consistent with
500 those obtained by Dasmahapatra, Lamas, Simpson, & Mallet (2010) using four loci, which
501 also revealed inter-lineage differentiation for these taxa, with the strongest genetic
502 differentiation occurring in *O. onega*.

503 However, the genomic and phenotypic population structure of hybrid populations differ
504 between *I. salapia* and *O. onega*. Firstly, while all but four of the *I. salapia* individuals
505 sampled in the hybrid zone are genetically closer to the Andean population, most individuals
506 in the *O. onega* hybrid populations we sampled are genetically closer to Amazonian
507 populations. Secondly, one individual in *O. onega* is likely a F1, and two other individuals are
508 recent backcrosses, while no such genetically intermediate individuals were found in our
509 samples of *I. salapia*. Thirdly, the phenotypic structure of hybrid populations mirrors the
510 genomic patterns. Along the first PC individuals in intermediate populations of *I. salapia* are
511 phenotypically closest to the Andean parental lineage (*derasa*), to which they are also closest

512 genetically. In our sample, intermediate color patterns are not observed in these populations,
513 nor in parental populations. By contrast, individuals in intermediate populations of *O. onega*
514 have intermediate phenotypes between the two lineages, with a tendency to be closer to the
515 Amazonian lineage (*janarilla*), to which they are also closest genetically.

516 Overall, the patterns detected suggest past gene flow in both species (most individuals have a
517 similar, low hybrid index), with potentially more recent (but rare) gene flow in *O. onega* -
518 although we cannot rule out the fact that we may have missed recent hybrids in *I. salapia*.
519 Genomic differentiation across hybrid zones in Müllerian mimetic butterflies have mostly
520 been documented in the genus *Heliconius*. While *Heliconius* sub-specific lineages sometimes
521 exhibit high genome-wide differentiation across hybrid zones (Martin, Davey, Salazar, &
522 Jiggins, 2019; Van Belleghem et al., 2018), this appears not to be the case in the Tarapoto
523 suture zone. In this region, Nadeau *et al.* (2014) found that in phenotypically differentiated
524 lineages of *H. erato* and *H. melpomene* only loci around pattern gene loci showed genetic
525 differentiation, while the rest of the genome was highly permeable to gene flow, with F_{ST}
526 values ranging from 0.0112 to 0.0280 (see also Martin et al., 2013). This stands in stark
527 contrast to the strong overall differentiation we revealed in ithomiine butterflies from the
528 Tarapoto suture zone ($F_{ST} = 0.177$ for *I. salapia* and $F_{ST} = 0.372$ for *O. onega*).

529 While intermediate populations of *O. onega* show a high extent of genetic heterogeneity, in *I.*
530 *salapia* all but four individuals from intermediate populations are remarkably similar in their
531 genetic composition. This suggests that intermediate populations of *I. salapia* are hardly
532 exchanging genes with Andean and Amazonian populations, and may be in the process of
533 forming a distinct taxon.

534

535 *Genomic patterns of introgression*

536 Variation in introgression patterns across the genome can help pinpoint loci involved in
537 adaptation and reproductive isolation (Gompert & Buerkle, 2011; Gompert et al., 2012;
538 Gompert, Mandeville, & Buerkle, 2017). In particular, highly divergent SNPs with deviant
539 genomic cline center parameters (α) or positive genomic cline rate parameter (β) (i. e.,
540 exhibiting a steep cline) should be more common in regions of the genome involved in local
541 adaptation or reproductive isolation.

542 Here, intermediate populations of both *I. salapia* and *O. onega* present SNPs with outlier
543 values in their genomic cline center parameters (α), meaning that these SNPs have an ancestry
544 different from that of the average of the genome. The SNPs exhibiting deviant α should be
545 enriched for genomic regions involved in adaptation or reproductive isolation. Such SNPs
546 (77.0%) are shifted towards Amazonian ancestry in *I. salapia*, whereas the majority of SNPs
547 with deviant α (65.5%) are shifted towards Andean ancestry in *O. onega*. While this may
548 indicate introgression from the parental lineage that is least represented in the genomic
549 background of intermediate populations, in our case such asymmetry may also result from a
550 lower power to detect introgression from the dominant parental background. . Whether some
551 of those SNPs result from adaptive introgression, as has been revealed in *Heliconius*
552 butterflies (*Heliconius* Genome Consortium, 2012; Jay et al., 2018), warrants further study.

553 Patterns of the parameter cline rate β markedly differ between *I. salapia* and *O. onega*. While
554 in *I. salapia* no SNPs show outlier cline steepness, in *O. onega* many SNPs show narrower or
555 wider clines compared to the genome average. Overall, in *O. onega*, the distribution of
556 genomic cline parameters is wider and more heterogeneous than in *I. salapia*, suggesting less
557 constraints in the hybridization process. Such heterogeneity in *O. onega* allows the
558 identification of loci with specific introgression levels. Highly divergent genomic regions that
559 have low levels of introgression are likely associated with reproductive isolation (Gompert &
560 Buerkle, 2011; Gompert et al., 2012). Low levels of introgression can be the result of several

561 evolutionary processes involving both extrinsic mechanisms, such as divergent selection and
562 environment-dependent selection against hybrids, and intrinsic mechanisms such as an
563 environment-independent reduced hybrid fitness caused by Bateson-Dobzhansky-Muller
564 incompatibilities (Gompert & Buerkle, 2011; Gompert et al., 2012). Correlation of genetic
565 patterns with other evidence (e. g., candidate traits) may shed light on the mechanisms of
566 speciation and reproductive isolation (Ravinet et al., 2017).

567

568 *Genetic bases of colour pattern variation*

569 Our admixture mapping analysis revealed SNPs associated with color pattern in *I. salapia* (88
570 SNPs, representing 0.49% of all SNPs) and *O. onega* (109 SNPs, representing less than
571 0.69% of all SNPs).

572 In nymphalid butterflies, wing pattern variation can be explained by combinations of
573 conserved pattern elements (Martin & Reed, 2014) and tends to be controlled by small
574 numbers of loci (Van Belleghem et al., 2017; Zhang et al., 2017). Previous studies, including
575 studies on mimetic *Heliconius* (Joron et al., 2006; Martin et al., 2012; Nadeau, 2016; Reed et
576 al. 2011; Westerman et al., 2018) and *Papilio* (Timmermans et al., 2014) identified a list of
577 candidate genes such as *WntA*, *optix*, *cortex*, *ebony* and *aristaless*. Some of these have been
578 functionally characterized (Martin & Reed, 2014; Nadeau, 2016; Nadeau et al., 2016). A
579 recent study on *D. plexippus*, the most closely related species to Ithomiini for which a
580 reference genome is available, highlighted the role of *WntA* in vein shape (Mazo-Vargas et
581 al., 2017).

582 None of our candidate loci correspond to genes known to be involved in wing colour pattern
583 in other butterflies. This is likely due to the relatively low-resolution genotype-by-sequencing

584 approach adopted here, such that we may have missed gene regions that were not covered by
585 our loci.

586 Moreover, only a small fraction of loci with SNPs associated with wing pattern (23.9% for *I.*
587 *salapia* and 21.2% for *O. onega*) map to an orthologous region in the *D. plexippus* genome.
588 This deficit is related to the relatively large divergence time between our focal species and *D.*
589 *plexippus* (ca. 42 million years ago, Chazot et al., 2019), which limits our ability to find
590 orthologous regions and more specifically to find regions involved in non-coding regulatory
591 loci. We may therefore have missed loci that contain known genes involved in wing pattern
592 development.

593 Finally, the extremely low level of hybridization observed in *I. salapia* reduces the statistical
594 power of admixture mapping and hampers detection of genomic regions associated with wing
595 pattern variation. The function of most regions identified in our analyses are unknown and
596 represent a starting point for further analyses of these regions, as those regions may contain
597 novel genes in these pathways.

598

599 *Colour pattern and reproductive isolation*

600 Wing colour pattern is known to cause pre- and post-zygotic reproductive isolation in
601 Müllerian mimetic butterflies (e. g., Chamberlain, Hill, Kapan, Gilbert, & Kronforst, 2009;
602 Jiggins, Naisbit, Coe, & Mallet, 2001; Mallet & Barton, 1989; Merrill et al., 2012; Merrill et
603 al., 2011; Naisbit, Jiggins, Linares, Salazar, & Mallet, 2002), including Ithomiini (McClure et
604 al., 2019).

605 In our admixture mapping analysis, we found that only two and one of the significantly
606 differentiated introgression outliers were also associated with wing pattern variation in *I.*
607 *salapia* and *O. onega*, respectively. These figures do not differ from random expectations.

608 These results suggest that wing colour pattern may be moderately involved in reproductive
609 isolation in both species, but since our genomic data do not cover the entire genome, we
610 cannot rule out the fact that we may have missed some important loci involved in wing
611 pattern coloration and with deviant genomic clines.

612 In mimetic butterflies, hybrid individuals with intermediate colour pattern may suffer more
613 predation because they are not recognized as unpalatable (e. g., Merrill et al., 2012), which
614 may in turn select for assortative mating for wing colour pattern through reinforcement (e. g.,
615 Kronforst, Young, & Gilbert, 2007), resulting in reproductive isolation between
616 phenotypically differentiated lineages. Whether individuals with intermediate phenotype
617 suffer increased predation has never been tested in *I. salapia* and *O. onega*, but predation
618 experiments on *Heliconius* species carried out in the same region demonstrated the ability of
619 predators to discriminate fine phenotypic differences (Arias et al., 2016; Chouteau, Arias, &
620 Joron, 2016). Assortative mating seems likely in *I. salapia* and *O. onega* (MM and ME, pers.
621 obs.), and has been documented by genetic and phenotypic characterization of the reared
622 offspring of females collected in hybrid populations of *O. onega* (De Silva, 2010: chapter 5).

623 There are fewer phenotypically intermediate individuals in *I. salapia* than in *O. onega*. This
624 difference might be explained by the mimicry rings to which the two species belong. While
625 the mimicry rings of *I. salapia* lineages are readily discriminated and show little variation
626 within each mimicry ring, the forms *O. onega* belongs to are more variable with overlapping
627 phenotypes (ME pers. obs.; Supplementary Figure 1). Because of the greater variation and
628 overlap of the two *O. onega* mimicry rings in Tarapoto, selection against hybrids with
629 intermediate phenotypes may be reduced compared with that in *I. salapia*, thereby allowing
630 the persistence of greater levels of gene flow between lineages. Whether the absence of
631 intermediate phenotypes in *I. salapia* is due to high mortality of hybrids through predation,

632 strong assortative mating, hybrid incompatibilities or all of these is currently unknown and
633 deserves further examination.

634

635 *Other putative adaptive traits*

636 In both species, only a small number of SNPs potentially involved in adaptation or
637 reproductive isolation (i. e., highly differentiated SNPs that show deviant α or significantly
638 positive β) are also associated with wing pattern. This suggests that other traits may play a
639 role in reproductive isolation. Because interactions with local host plants at the larval stage
640 and the ability to fully exploit them often impact fitness in phytophagous insects (Simon et
641 al., 2015), larval hostplant shifts are believed to be an important driver of reproductive
642 isolation. However, in *O. onega* the two lineages utilize the same larval hostplants, *Solanum*
643 *mite* and related *Solanum* sect. *Pteroidea* species (de-Silva, Vásquez, & Mallet, 2011;
644 Gallusser, 2002). Similarly, *I. salapia derasa* larvae commonly feed on *Witheringia*
645 *solanacea* (Beccaloni, 1997), a plant also used by *I. salapia aquinia* (JM, MM and ME,
646 unpublished observations). Shifts in hostplant are therefore unlikely to explain divergence
647 between lineages in either of these species, as is the case in another ithomiine genus,
648 *Melinaea*, present in the same region (McClure & Elias, 2016).

649 The two lineages of *O. onega* have divergent egg-laying behavior: females of the Amazonian
650 population (*janarilla*) lay eggs on the hostplants, while females of the Andean population
651 (ssp. nov. 2) tend to lay eggs off the host plant (Gallusser, 2002). Eggs are typically laid up to
652 0.5 m away from the nearest host plant individual, on twigs, leaf litter or live non-host plant,
653 which reduces egg predation (de-Silva, Vásquez, & Mallet, 2011). Differences in egg-laying
654 behavior have been shown to cause reduced hybrid fitness in butterflies (McBride & Singer,
655 2010). This could be the case here, too, if hybrid females lay eggs off the plant, and if first
656 instar larvae are incapable of locating their host plant.

657 Other putative adaptive traits include adaptations to distinct habitats (higher elevations and
658 cooler temperatures for Andean lineages) and potentially microhabitats where co-mimics are
659 most abundant (e. g., Elias, Gompert, Jiggins, & Willmott, 2008).

660 Finally, as many butterfly species, Ithomiini probably rely on sexual pheromones during mate
661 choice (Schulz et al., 2004), and differences in sexual pheromones may incur discrimination
662 between lineages. Notably, putative male pheromones have been shown to differ between the
663 two lineages of *O. onega* (Stamm, Mann, McClure, Elias, & Schulz, 2019).

664 The role of these traits in reproductive isolation remains to be further explored using both
665 experimental and genomic approaches.

666

667 **Acknowledgements.** We thank the Peruvian authorities and Dr Gerardo Lamas (Museo de
668 Historia Natural, Universidad Mayor de San Marcos) for research permits (096-2004-
669 INRENA-IFFS-DCB, 021C/C-2005-INRENA-IANP and 236-2012-AG-DGFFS-DGEFFS).

670 We also thank Armando Silva-Vásquez and Fraser Simpson for their precious help in the
671 field. Molecular work was carried out at the Service de Systématique Moléculaire du Muséum
672 National d'Histoire Naturelle (UMS 2700 - OMSI). The support and resources from the
673 Center for High Performance Computing at the University of Utah are gratefully
674 acknowledged. We thank three anonymous reviewers for their useful comments that led us to
675 improve our manuscript.

676

677 **Data accessibility.** Sequence reads are archived at the NCBI SRA in the BioProject
678 PRJNA575968. Scripts describing the whole analytic process have been uploaded to GitHub
679 (https://github.com/JeremyLGauthier/Scripts_Gauthier_et.al_2019_ME).

680

681 **Author contributions :** ME and ZG designed the study. LdS, JM , MM and ME performed
682 sampling. LdS, AW, ZG and ME performed labwork. JG analysed the molecular data, with
683 contributions from ZG, AW, CL and FL. ME, JG, CH and YLP analysed phenotypic data. All
684 authors took part in discussions concerning the analyses and result interpretations. JG wrote
685 the paper, with contributions from all authors.

686

687 **Funding.** This research was funded by a CNRS ATIP grant, two ANR grants (SPECREP
688 ANR-14-CE02-0011 and CLEARWING ANR-16-CE02-0012) and a Human Frontier Science
689 Program (RGP0014/2016) grant awarded to ME. LdS was a postdoc on the ATIP grant and
690 JG and MM were postdocs on the ANR SPECREP grant.

691

692

693

694 **References:**

695

696 Arias, M., le Poul, Y., Chouteau, M., Boisseau, R., Rosser, N., Théry, M., & Llaurens, V.
697 (2016). Crossing fitness valleys: empirical estimation of a fitness landscape associated
698 with polymorphic mimicry. *Proceedings. Biological Sciences / The Royal Society*,
699 283(1829).

700 Barton, N. H., & Bengtsson, B. O. (1986). The barrier to genetic exchange between
701 hybridising populations. *Heredity*, 57(3), 357.

702 Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology
703 and Systematics*, 16(1), 113–148.

704 Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*,
705 341(6242), 497–503.

706 Bates, H. W. (1862). XXXII. Contributions to an insect fauna of the Amazon valley.
707 Lepidoptera: Heliconidae. *Transactions of the Linnean Society of London*, 23(3), 495–
708 566.

709 Beccaloni, G. W. (1997). Ecology, natural history and behaviour of Ithomiine butterflies and
710 their mimics in Ecuador (Lepidoptera: Nymphalidae: Ithomiinae). *Tropical Lepidoptera
711 Research*, 8(2), 103–124.

712 Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and
713 powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*,
714 57, 289–300.

715 Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis:
716 why genome scans may fail to map local adaptation genes. *Molecular Ecology*, 20(10),
717 2044–2072.

718 Brown K. S., Sheppard Philip Macdonald, & Turner John Richard George. (1974).
719 Quaternary refugia in tropical America: evidence from race formation in *Heliconius*
720 butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 187(1088), 369–
721 378.

722 Buerkle, C. A., & Lexer, C. (2008). Admixture as the basis for genetic mapping. *Trends in
723 Ecology & Evolution*, 23(12), 686–694.

724 Challis, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D., & Blaxter, M. (2016). Lepbase:
725 the Lepidopteran genome database (p. 056994). doi: 10.1101/056994

726 Chamberlain, N. L., Hill, R. I., Kapan, D. D., Gilbert, L. E., & Kronforst, M. R. (2009).
727 Polymorphic butterfly reveals the missing link in ecological speciation. *Science*,
728 326(5954), 847–850.

729 Chazot, N., Wahlberg, N., Freitas, A. V. L., Mitter, C., Labandeira, C., Sohn, J.-C., ...
730 Heikkilä, M. (2019). Priors and Posteriors in Bayesian Timing of Divergence Analyses:
731 The Age of Butterflies Revisited. *Systematic Biology*, 68(5), 797–813.

732 Chouteau, M., Arias, M., & Joron, M. (2016). Warning signals are under positive
733 frequency-dependent selection in nature. *Proceedings of the National Academy of
734 Sciences of the United States of America*, 113(8), 2164–2169.

735 Christe, C., Stölting, K. N., Bresadola, L., Fussi, B., Heinze, B., Wegmann, D., & Lexer, C.
736 (2016). Selection against recombinant hybrids maintains reproductive isolation in

- 737 hybridizing *Populus* species despite F1 fertility and recurrent gene flow. *Molecular*
738 *Ecology*, 25(11), 2482–2498.
- 739 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... 1000
740 Genomes Project Analysis Group. (2011). The variant call format and VCFtools.
741 *Bioinformatics*, 27(15), 2156–2158.
- 742 Dasmahapatra, K. K., Lamas, G., Simpson, F., & Mallet, J. (2010). The anatomy of a “suture
743 zone” in Amazonian butterflies: a coalescent-based test for vicariant geographic
744 divergence and speciation. *Molecular Ecology*, 19(19), 4283–4301.
- 745 de-Silva, D. L. (2010). Ecology and Evolution in Neotropical Butterflies of the Subtribe
746 Oleriina (Lepidoptera: Nymphalidae: Danainae: Ithomiini). PhD thesis, University of
747 London, 275pp.
- 748 de-Silva, D. L., Vásquez, A. S., & Mallet, J. (2011). Selection for enemy-free space: eggs
749 placed away from the host plant increase survival of a neotropical ithomiine butterfly.
750 *Ecological Entomology*, 36(6), 667–672.
- 751 Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program
752 for visualizing STRUCTURE output and implementing the Evanno method.
753 *Conservation Genetics Resources*, 4(2), 359–361.
- 754 Elias, M., Gompert, Z., Jiggins, C., & Willmott, K. (2008). Mutualistic interactions drive
755 ecological niche convergence in a diverse butterfly community. *PLoS Biology*, 6(12),
756 2642–2649.
- 757 Endler, J. A. (1977). Geographic variation, speciation, and clines. *Monographs in Population*
758 *Biology*, 10, 1–246.
- 759 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals
760 using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611–
761 2620.
- 762 Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using
763 multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4),
764 1567–1587.
- 765 Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N.,
766 ... Altshuler, D. (2004). Assessing the impact of population stratification on genetic
767 association studies. *Nature Genetics*, 36(4), 388–393.
- 768 Gallusser, S. A. (2002). Biology, behaviour and taxonomy of two *Oleria onega* subspecies
769 (Ithomiinae, Nymphalidae, Lepidoptera) in north-eastern, Peru (Université de
770 Neuchâtel). Retrieved from <http://doc.rero.ch/record/2627>
- 771 Gauthier, J., Mouden, C., Suchan, T., Alvarez, N., Arrigo, N., Riou, C., ... Peterlongo, P.
772 (2017). DiscoSnp-RAD: de novo detection of small variants for population genomics (p.
773 216747). doi: 10.1101/216747
- 774 Gompert, Z., & Buerkle, C. A. (2009). A powerful regression-based method for admixture
775 mapping of isolation across the genome of hybrids. *Molecular Ecology*, 18(6), 1207–
776 1224.
- 777 Gompert, Z., & Buerkle, C. A. (2011). Bayesian estimation of genomic clines. *Molecular*
778 *Ecology*, 20(10), 2111–2127.
- 779 Gompert, Z., & Buerkle, C. A. (2012). bgc: Software for Bayesian estimation of genomic
780 clines. *Molecular Ecology Resources*, 12(6), 1168–1176.

- 781 Gompert, Z., & Buerkle, C. A. (2013). Analyses of genetic ancestry enable key insights for
782 molecular ecology. *Molecular Ecology*, 22(21), 5278–5294.
- 783 Gompert, Z., Lucas, L. K., Nice, C. C., Fordyce, J. A., Forister, M. L., & Buerkle, C. A.
784 (2012). Genomic regions with a history of divergent selection affect fitness of hybrids
785 between two butterfly species. *Evolution*, 66(7), 2167–2181.
- 786 Gompert Z., Lucas L. K., Buerkle C. A., Forister M. L., Fordyce J. A., Nice C. C. (2014).
787 Admixture and the organization of genetic diversity in a butterfly species complex
788 revealed through common and rare genetic variants. *Molecular Ecology*; 23(18):4555–
789 4573.
- 790 Gompert, Z., Mandeville, E. G., & Buerkle, C. A. (2017). Analysis of Population Genomic
791 Data from Hybrid Zones. *Annual Review of Ecology, Evolution, and Systematics*, 48(1),
792 207–229.
- 793 *Heliconius* Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of
794 mimicry adaptations among species. *Nature*, 487(7405), 94–98.
- 795 Jay, P., Whibley, A., Frézal, L., Rodríguez de Cara, M. Á., Nowell, R. W., Mallet, J., ...
796 Joron, M. (2018). Supergene Evolution Triggered by the Introgression of a
797 Chromosomal Inversion. *Current Biology: CB*, 28(11), 1839–1845.e3.
- 798 Jiggins, C. D., Naisbit, R. E., Coe, R. L., & Mallet, J. (2001). Reproductive isolation caused
799 by colour pattern mimicry. *Nature*, 411(6835), 302–305.
- 800 Jiggins, C.D., & Mallet, J. (2000). Bimodal hybrid zones and speciation. *Trends in Ecology*
801 *and Evolution* 15:250-255.
- 802 Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide
803 SNP data. *Bioinformatics* , 27(21), 3070–3071.
- 804 Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., ... Kingsley,
805 D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks.
806 *Nature*, 484(7392), 55–61.
- 807 Joron, M., Papa, R., Beltrán, M., Chamberlain, N., Mavárez, J., Baxter, S., ... Jiggins, C. D.
808 (2006). A conserved supergene locus controls colour pattern diversity in *Heliconius*
809 butterflies. *PLoS Biology*, 4(10), e303.
- 810 Kapan, D. D. (2001). Three-butterfly system provides a field test of müllerian mimicry.
811 *Nature*, 409(6818), 338–340.
- 812 Kozak, K. M., Wahlberg, N., Neild, A. F. E., Dasmahapatra, K. K., Mallet, J., & Jiggins, C.
813 D. (2015). Multilocus species trees show the recent adaptive radiation of the mimetic
814 *Heliconius* butterflies. *Systematic Biology*, 64(3), 505–524.
- 815 Kronforst, M. R., Hansen, M. E. B., Crawford, N. G., Gallant, J. R., Zhang, W., Kulathinal, R.
816 J., ... Mullen, S. P. (2013). Hybridization reveals the evolving genomic architecture of
817 speciation. *Cell Reports*, 5(3), 666–677.
- 818 Kronforst, M. R., Young, L. G., & Gilbert, L. E. (2007). Reinforcement of mate preference
819 among hybridizing *Heliconius* butterflies. *Journal of Evolutionary Biology*, 20(1), 278–
820 285.
- 821 Larson, E. L., Andrés, J. A., Bogdanowicz, S. M., & Harrison, R. G. (2013). Differential
822 introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution*, 67(12),
823 3653–3661.

- 824 Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V., & Joron, M. (2014).
825 Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature*
826 *Communications*, 5, 5644.
- 827 Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology &*
828 *Evolution*, 20(5), 229–237.
- 829 Mallet, J., & Barton, N. H. (1989). Strong Natural Selection in a Warning-Color Hybrid Zone.
830 *Evolution*, 43(2), 421–431.
- 831 Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., ...
832 Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in
833 *Heliconius* butterflies. *Genome Research* 23:1817-1828.
- 834 Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate
835 variation shapes barriers to introgression across butterfly genomes. *PLoS Biology*, 17(2),
836 e2006288.
- 837 Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., ... Reed, R. D.
838 (2012). Diversification of complex butterfly wing patterns by repeated regulatory
839 evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences of the*
840 *United States of America*, 109(31), 12632–12637.
- 841 Martin, A., & Reed, R. D. (2014). Wnt signaling underlies evolution and development of the
842 butterfly wing pattern symmetry systems. *Developmental Biology*, 395(2), 367–378.
- 843 Mazo-Vargas, A., Concha, C., Livraghi, L., Massardo, D., Wallbank, R. W. R., Zhang, L., ...
844 Martin, A. (2017). Macroevolutionary shifts of WntA function potentiate butterfly wing-
845 pattern diversity. *Proceedings of the National Academy of Sciences of the United States*
846 *of America*, 114(40), 10701–10706.
- 847 McBride, C. S., & Singer, M. C. (2010). Field studies reveal strong postmating isolation
848 between ecologically divergent butterfly populations. *PLoS Biology*, 8(10), e1000529.
- 849 McClure, M., & Elias, M. (2016). Unravelling the role of host plant expansion in the
850 diversification of a Neotropical butterfly genus. *BMC Evolutionary Biology*, 16(1), 128.
- 851 McClure, M., Mahrouche, L., Houssin, C., Monllor, M., Le Poul, Y., Frérot, B., ... Elias, M.
852 (2019). Does divergent selection predict the evolution of mate preference and
853 reproductive isolation in the tropical butterfly genus *Melinaea* (Nymphalidae:
854 Ithomiini)? *The Journal of Animal Ecology*. doi: 10.1111/1365-2656.12975
- 855 Merrill R. M., Wallbank R. W. R., Bull V., Salazar P. C. A., Mallet J., Stevens M., & Jiggins
856 C. D. (2012). Disruptive ecological selection on a mating cue. *Proceedings of the Royal*
857 *Society B: Biological Sciences*, 279(1749), 4907–4913.
- 858 Merrill, R. M., Gompert, Z., Dembeck, L. M., Kronforst, M. R., McMillan, W. O., & Jiggins,
859 C. D. (2011). Mate preference across the speciation continuum in a clade of mimetic
860 butterflies. *Evolution; International Journal of Organic Evolution*, 65(5), 1489–1500.
- 861 Moritz, C., Hoskin, C. J., MacKenzie, J. B., Phillips, B. L., Tonione, M., Silva, N., ...
862 Graham, C. H. (2009). Identification and dynamics of a cryptic suture zone in tropical
863 rainforest. *Proceedings of the Royal Society B: Biological Sciences*, 276(1660), 1235–
864 1244.
- 865 Muller, F. (1879). *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies.
866 *Proceedings of the Entomological Society of London*.
- 867 Nadeau, N. J. (2016). Genes controlling mimetic colour pattern variation in butterflies.

- 868 *Current Opinion in Insect Science*, 17, 24–31.
- 869 Nadeau, N. J., Pardo-Diaz, C., Whibley, A., Supple, M. A., Saenko, S. V., Wallbank, R. W.
870 R., ... Jiggins, C. D. (2016). The gene *cortex* controls mimicry and crypsis in butterflies
871 and moths. *Nature*, 534(7605), 106–110.
- 872 Nadeau, N. J., Ruiz, M., Salazar, P., Counterman, B., Medina, J. A., Ortiz-Zuazaga, H., ...
873 Papa, R. (2014). Population genomics of parallel hybrid zones in the mimetic butterflies,
874 *H. melpomene* and *H. erato*. *Genome Research*, 24(8), 1316–1333.
- 875 Naisbit, R. E., Jiggins, C. D., Linares, M., Salazar, C., & Mallet, J. (2002). Hybrid sterility,
876 Haldane's rule and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics*, 161(4),
877 1517–1526.
- 878 Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous
879 genomic divergence. *Molecular Ecology*, 18(3), 375–402.
- 880 Pallares, L. F., Harr, B., Turner, L. M., & Tautz, D. (2014). Use of a natural hybrid zone for
881 genomewide association mapping of craniofacial traits in the house mouse. *Molecular*
882 *Ecology*, 23(23), 5756–5770.
- 883 Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A.
884 (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine.
885 *Molecular Ecology*, 21(12), 2991–3005.
- 886 Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D.
887 (2006). Principal components analysis corrects for stratification in genome-wide
888 association studies. *Nature Genetics*, 38(8), 904–909.
- 889 Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using
890 multilocus genotype data. *Genetics*, 155(2), 945–959.
- 891 Priyam, A., Woodcroft, B. J., Rai, V., Munagala, A., Moghul, I., Ter, F., ... Wurm, Y. (2019).
892 Sequenceserver: a modern graphical user interface for custom BLAST databases,
893 *Molecular Biology and Evolution*, msz185.
- 894 Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., ... Westram, A.
895 M. (2017). Interpreting the genomic landscape of speciation: a road map for finding
896 barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450–1477.
- 897 Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., ...
898 McMillan, W. O. (2011). *optix* drives the repeated convergent evolution of butterfly
899 wing pattern mimicry. *Science*, 333(6046), 1137–1141.
- 900 Remington, C. L. (1968). Suture-zones of hybrid interaction between recently joined biotas.
901 *Evolutionary Biology* (pp. 321–428).
- 902 Rieseberg, L. H., Whitton, J., & Gardner, K. (1999). Hybrid zones and the genetic
903 architecture of a barrier to gene flow between two sunflower species. *Genetics*, 152(2),
904 713–727.
- 905 Rissler, L. J., & Smith, W. H. (2010). Mapping amphibian contact zones and
906 phylogeographical break hotspots across the United States. *Molecular Ecology*, 19(24),
907 5404–5416.
- 908 Roberts, J. L., Brown, J. L., May, R. von, Arizabal, W., Schulte, R., & Summers, K. (2006).
909 Genetic divergence and speciation in lowland and montane peruvian poison frogs.
910 *Molecular Phylogenetics and Evolution*, 41(1), 149–164.

- 911 Safran, R. J., & Nosil, P. (2012). Speciation: The origin of new species. *Nature Education*
912 *Knowledge*, 3(10), 17.
- 913 Schulz, S., Beccaloni, G., Brown, K. S., Jr., Boppre, M., Freitas, A. V. L., Ockenfels, P., &
914 Trigo, J. R. (2004). Semiochemicals derived from pyrrolizidine alkaloids in male
915 ithomiine butterflies (Lepidoptera: Nymphalidae: Ithomiinae). *Biochemical Systematics*
916 *and Ecology*, 32.
- 917 Simon, J.-C., d'Alençon, E., Guy, E., Jacquin-Joly, E., Jaquiéry, J., Nouhaud, P., ... Streiff,
918 R. (2015). Genomics of adaptation to host-plants in herbivorous insects. *Briefings in*
919 *Functional Genomics*, 14(6), 413–423.
- 920 Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D.,
921 ... Brumfield, R. T. (2014). The drivers of tropical speciation. *Nature*, 515(7527), 406–
922 409.
- 923 Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J.
924 S., ... Nosil, P. (2014). Stick insect genomes reveal natural selection's role in parallel
925 speciation. *Science*, 344(6185), 738–742.
- 926 Stamm, P., Mann, F., McClure, M., Elias, M., & Schulz, S. (2019). Chemistry of the
927 Androconial Secretion of the Ithomiine Butterfly *Oleria onega*. *Journal of Chemical*
928 *Ecology*, 45(9), 768–778.
- 929 Teeter, K. C., Payseur, B. A., Harris, L. W., Bakewell, M. A., Thibodeau, L. M., O'Brien, J.
930 E., ... Tucker, P. K. (2008). Genome-wide patterns of gene flow across a house mouse
931 hybrid zone. *Genome Research*, 18(1), 67–76.
- 932 Teeter, K. C., Thibodeau, L. M., Gompert, Z., Buerkle, C. A., Nachman, M. W., & Tucker, P.
933 K. (2010). The variable genomic architecture of isolation between hybridizing species of
934 house mice. *Evolution; International Journal of Organic Evolution*, 64(2), 472–485.
- 935 Timmermans Martijn J. T. N., Baxter Simon W., Clark Rebecca, Heckel David G., Vogel
936 Heiko, Collins Steve, ... Vogler Alfred P. (2014). Comparative genomics of the
937 mimicry switch in *Papilio dardanus*. *Proceedings of the Royal Society B: Biological*
938 *Sciences*, 281(1787), 20140465.
- 939 Uricaru, R., Rizk, G., Lacroix, V., Quillery, E., Plantard, O., Chikhi, R., ... Peterlongo, P.
940 (2015). Reference-free detection of isolated SNPs. *Nucleic Acids Research*, 43(2), e11.
- 941 Van Belleghem, S. M., Baquero, M., Papa, R., Salazar, C., McMillan, W. O., Counterman, B.
942 A., ... Martin, S. H. (2018). Patterns of Z chromosome divergence among *Heliconius*
943 species highlight the importance of historical demography. *Molecular Ecology*, 27(19),
944 3852–3872.
- 945 Van Belleghem, S. M., Rastas, P., Papanicolaou, A., Martin, S. H., Arias, C. F., Supple, M.
946 A., ... Papa, R. (2017). Complex modular architecture around a simple toolkit of wing
947 pattern genes. *Nature Ecology & Evolution*, 1(3), 52.
- 948 Via, S., & Hawthorne, D. J. (2002). The genetic architecture of ecological specialization:
949 correlated gene effects on host use and habitat choice in pea aphids. *The American*
950 *Naturalist*, 159 Suppl 3, S76–S88.
- 951 Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-Statistics for the analysis of population
952 structure. *Evolution*, 38(6), 1358–1370.
- 953 Weir, J. T. (2006). Divergent timing and patterns of species accumulation in lowland and
954 highland neotropical birds. *Evolution*, 60(4), 842–855.

- 955 Westerman, E. L., VanKuren, N. W., Massardo, D., Tenger-Trolander, A., Zhang, W., Hill, R.
956 I., ... Kronforst, M. R. (2018). Aristaless controls butterfly wing color variation used in
957 mimicry and mate choice. *Current Biology: CB*, 28(21), 3469–3474.e4.
- 958 Whinnett, A., Zimmermann, M., Willmott, K. R., Herrera, N., Mallarino, R., Simpson, F., ...
959 Mallet, J. (2005). Strikingly variable divergence times inferred across an Amazonian
960 butterfly “suture zone.” *Proceedings of the Royal Society B: Biological Sciences*,
961 272(1580), 2525–2533.
- 962 Zhang, L., Martin, A., Perry, M. W., van der Burg, K. R. L., Matsuoka, Y., Monteiro, A., &
963 Reed, R. D. (2017). Genetic basis of melanin pigmentation in butterfly wings. *Genetics*,
964 205(4), 1537–1550.
- 965 Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for
966 association studies. *Nature Genetics*, 44(7), 821–824.
- 967 Zhou, X., & Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for
968 genome-wide association studies. *Nature Methods*, 11(4), 407–409.
- 969

970 **Figure captions :**

971

972 **Figure 1.** The study organisms and sites studied in N.E. Peru. a. Photos of representative
973 specimens from each population of the two studied species (dorsal side shown against a dark
974 background to highlight transparency and ventral side shown against a white background to
975 highlight colour pattern). b. Sampling sites for *I. salapia* populations (top) with Amazonian
976 sites in red (1) Km-26 Yurimaguas-Tarapoto and (2) San Miguel de Achinamiza, the Andean
977 sites in blue (5) Puente Aguas Verdes and sites within the hybrid zone in purple (i3) Km-42
978 Tarapoto-Yurimaguas and (i4) La Florida. For *O. onega* populations (bottom), Amazonian
979 sites are in green (1) Km-26 Yurimaguas-Tarapoto and (2) San Miguel de Achinamiza,
980 Andean sites are in yellow (5) Puente Serranoyacu and the sites in the hybrid zone are in
981 apple green (o3) Shapaja-Chazuta (o4) from Quebrada Yanayacu to Laguna del Mundo
982 Perdido. Color codes are conserved for all other figures. c. Photos of putative hybrid
983 specimens with intermediate color patterns (ventral side). Photo credits: Céline Houssin

984

985 **Figure 2.** Population structure of pure and hybrid populations of *Ithomia salapia* (top) and
986 *Oleria onega* (bottom). a. Principal component Analysis (PC1: horizontal axis, PC2: vertical
987 axis), the percentage of total inertia explained by each axis is indicated in parentheses and the
988 histograms in the top corners represent the inertia percentages of the first principal
989 components. b. Structure plot for $K = 2$ and $K = 3$. The number of individuals that were used
990 is indicated for each site. c. Plot of the hybrid index of each sample from the hybrid
991 populations. The points represent the mean hybrid index value estimated from the posterior
992 distribution and black lines indicate 95% credible intervals. d. Plot of intertaxon ancestry and
993 hybrid index. Population color codes are the same as those in Figure 1b.

994

995 **Figure 3.** Scatterplots representing the relationships between the genomic cline center
996 parameter (α), representing SNP ancestry; the genomic cline rate parameter (β), representing
997 the steepness of the cline; and the differentiation level, F_{ST} , estimated for each SNP. Plots for
998 *I. salapia* and *O. onega* are on the left and right, respectively. Each data point is colored in grey,
999 and darkness increases with point density (i. e., darker areas contain more points). Blue lines
1000 frame sets of SNPs for which the genomic cline center parameters (α) significantly deviates
1001 from the genome-wide pattern. Green lines frame sets of SNPs for which the genomic cline
1002 rate parameters (β) significantly deviates from the genome-wide pattern. Note that for *I.*
1003 *salapia*, no SNPs have genomic cline rate parameters (β) that deviate from the genome-wide
1004 pattern. SNPs on the right hand side of the orange lines harbour a significantly higher
1005 differentiation (high F_{ST}) than the genome average.

1006

1007 **Figure 4.** Phenotypic position of 90 *I. salapia* (top) and 94 *O. onega* (bottom) in the wing
1008 color space consisting of the two main principal components from the colour pattern
1009 modeling approach. Color indicates sample populations as in Figure 1b. Females and males
1010 are depicted by circles and triangles, respectively. Representative images of the average
1011 phenotypes for population and sex are shown on each side of the figure.

1012

1013 **Figure 5.** The relationship between the significance of association with color pattern
1014 (represented as $-\log_{10}(p\text{-Wald})$) and F_{ST} for *I. salapia* (left) and *O. onega* (right). Yellow
1015 points indicate SNPs significantly associated with wing patterns (after both Bonferroni and

1016 FDR corrections). Orange points highlight SNPs with high F_{ST} values and red points highlight
1017 SNPs with significant association both to wing pattern and to high F_{ST} .

1018

1019 **Figure 6.** Venn diagram combining number of SNPs identified as supported by each approach
1020 (differentiation, introgression and admixture mapping) and shared between them. Note that
1021 for introgression patterns, no SNP showed deviant genomic cline rate parameters (β) in *I.*
1022 *salapia*. This parameter is therefore not represented in the diagram.

1023

1024

1025 **Supplementary material :**

1026

1027 **Supplementary Figure 1.** Mimicry ring example for each studied species and lineages, in
1028 black frameworks, including various other butterfly species.

1029

1030 **Supplementary Figure 2.** Barplots with error bars of the Phenotypic Variation Explained
1031 (PVE) by genetic for each variable (PC) explaining more than 1% of the wing pattern
1032 variation.

1033

1034 **Supplementary Table 1.** Sampling information including species, population, sex, location,
1035 region, GPS positions, sampling date. For each sample, the number of reads sequenced and
1036 SNPs called has been given.

1037

1038 **Supplementary Table 2.** BLAST results of locus with outlier SNPs identified as potentially
1039 involved in local adaptation, adaptive introgression and reproductive isolation, i.e. differential
1040 genomic cline center (α), high differentiation level (F_{ST}) and differential positive genomic
1041 cline rate (β).

1042

1043 **Supplementary Table 3.** BLAST results of loci with SNPs significantly associated with wing
1044 pattern variation.

1045

1046

a.

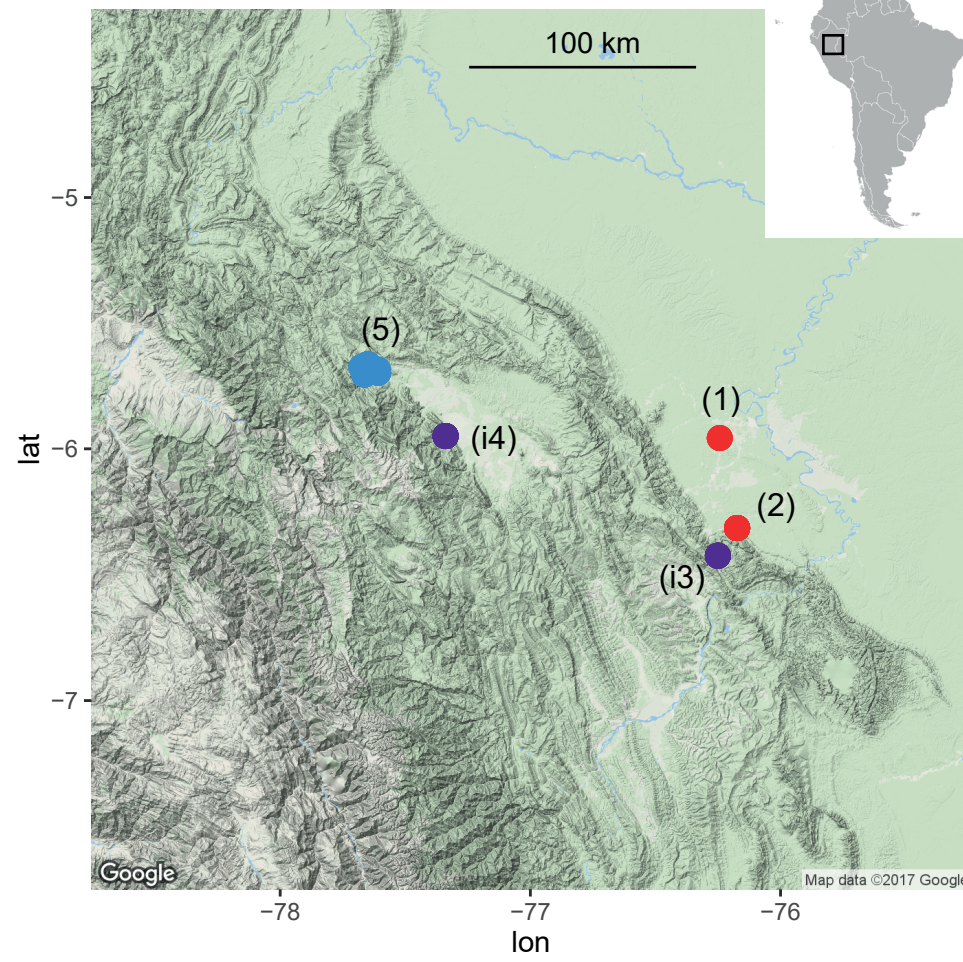
Ithomia salapia aquinia (Amazon)



Ithomia salapia derasa (Andes)



b.



c.

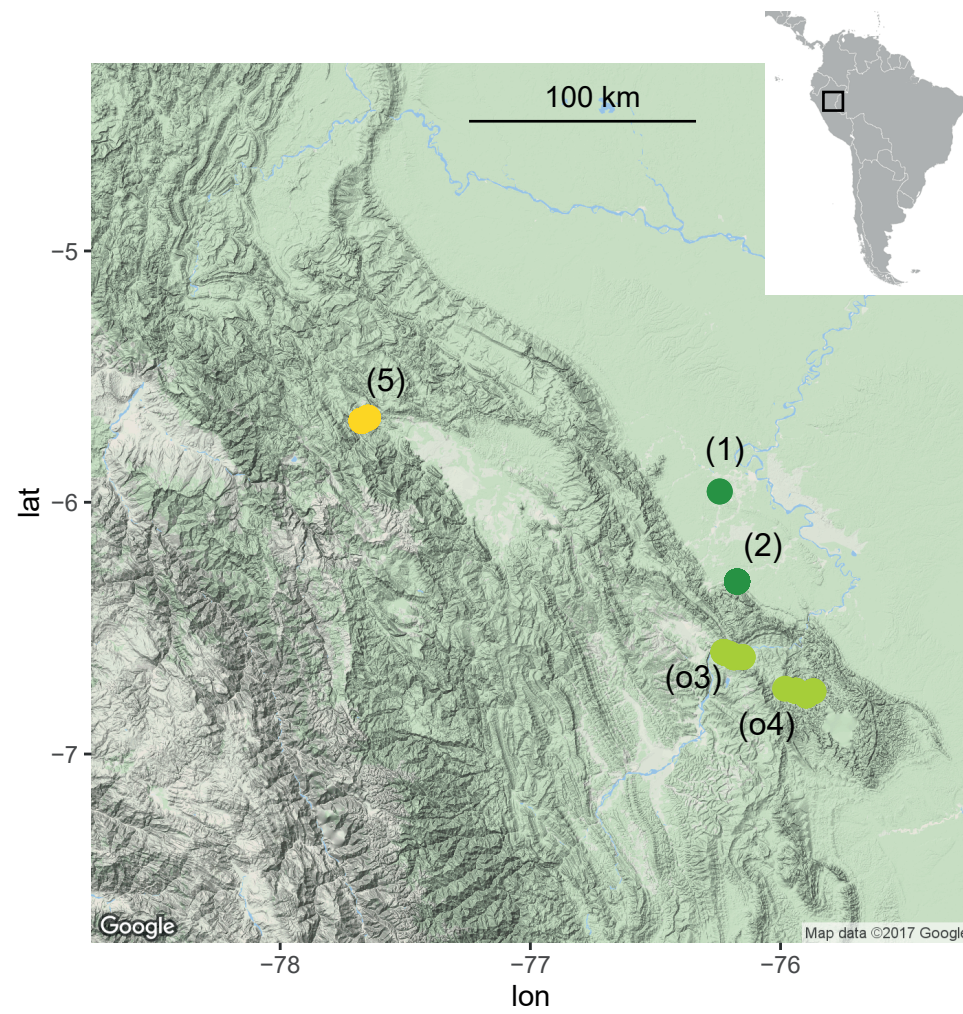
Ithomia salapia hybrid



Oleria onega janarilla (Amazon)



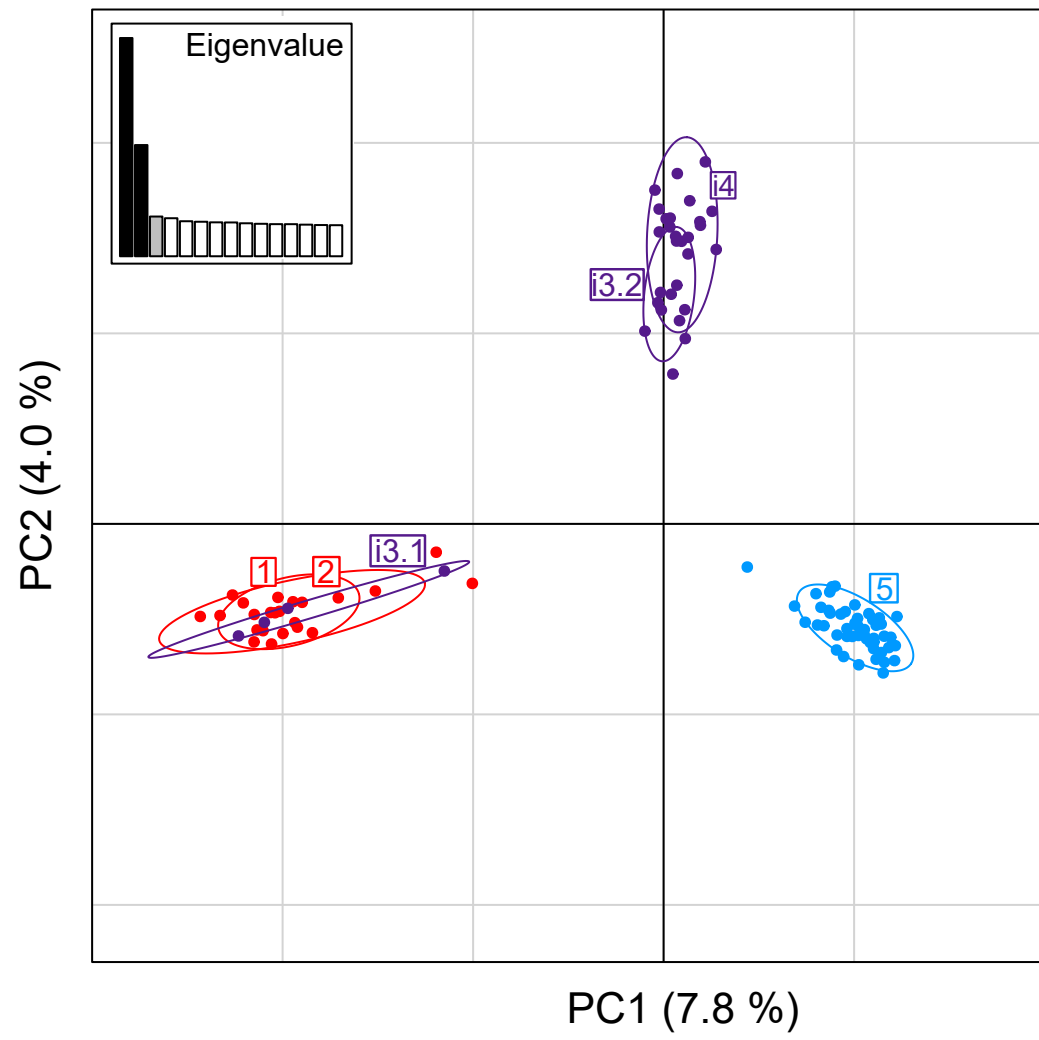
Oleria onega ssp nov 2 (Andes)



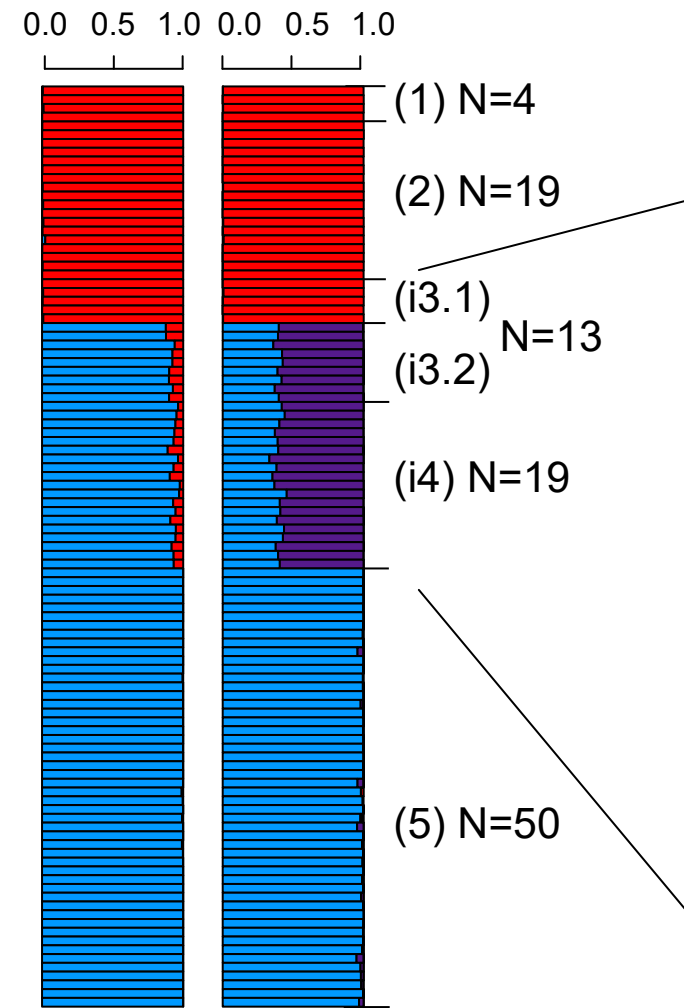
Oleria onega hybrid



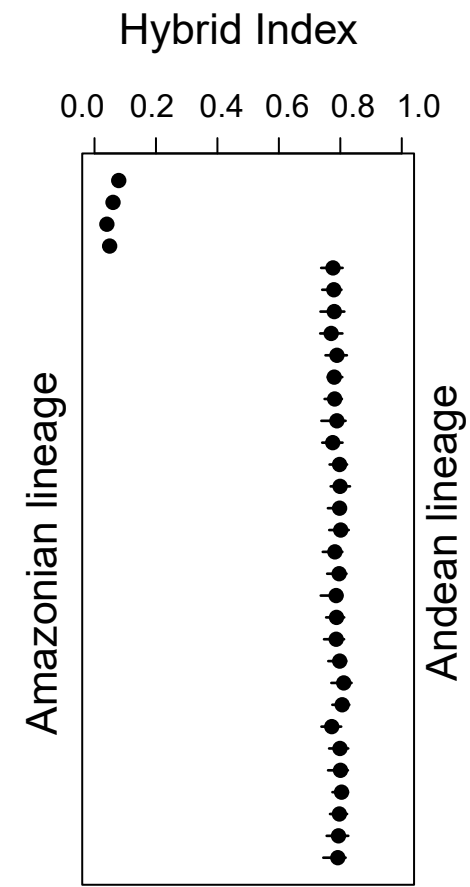
a. *Ithomia salapia*



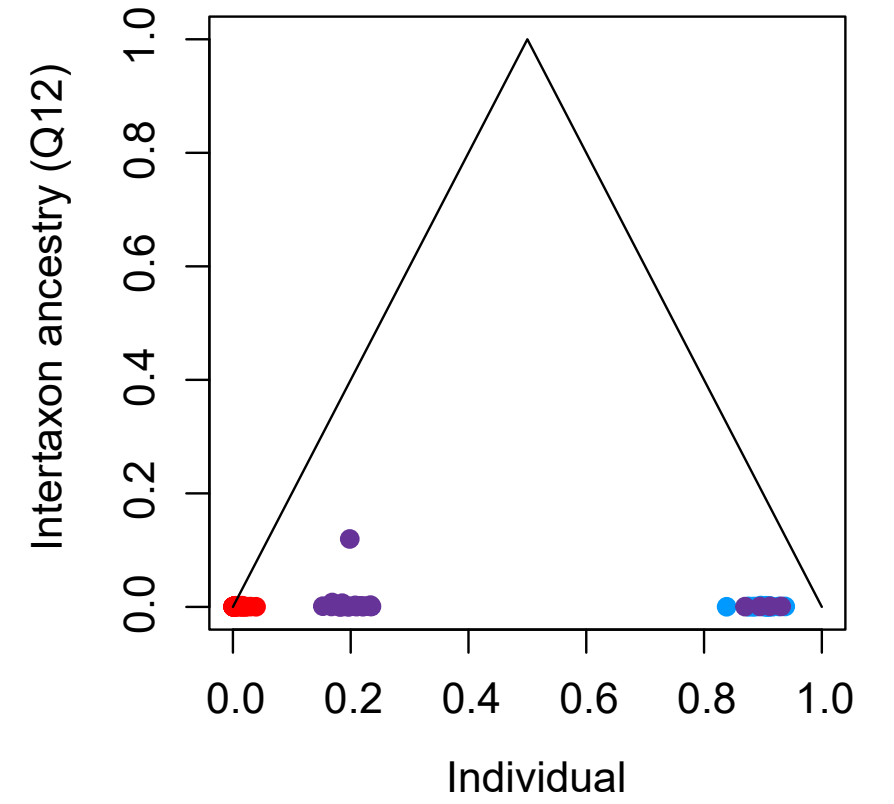
b.



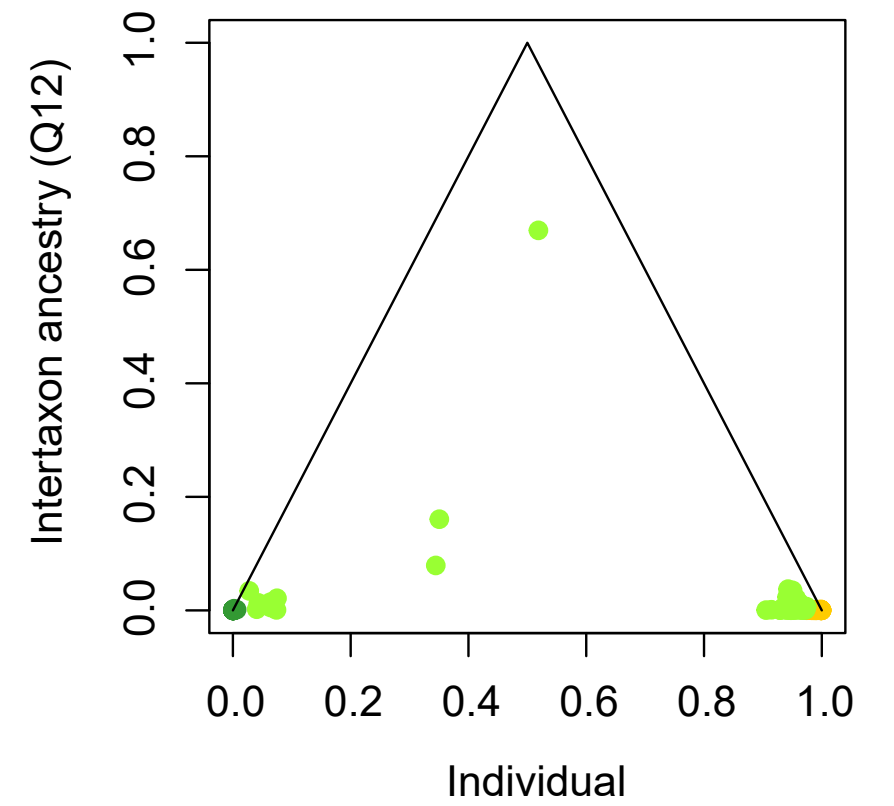
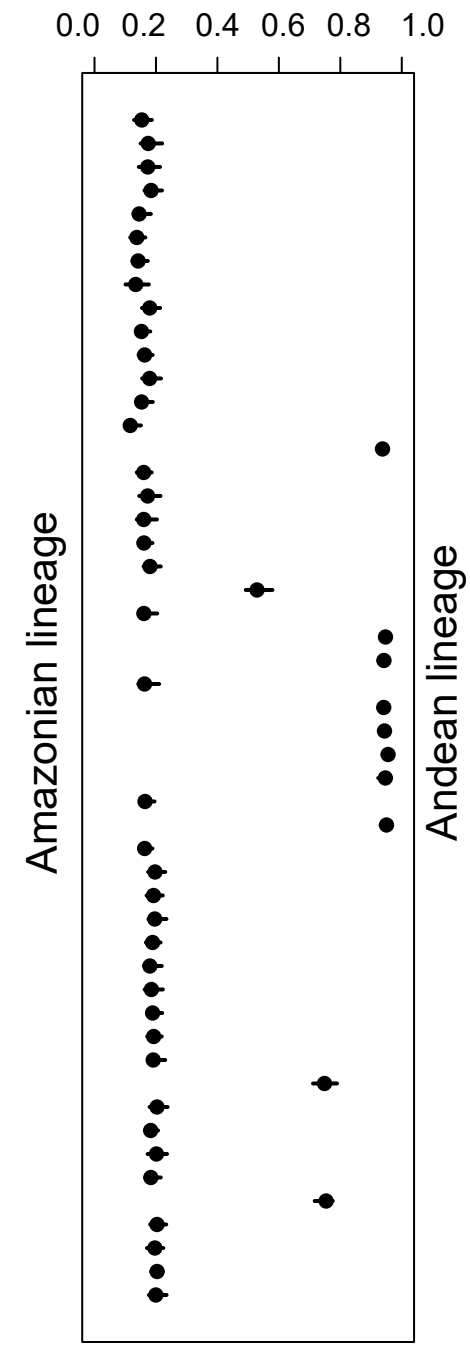
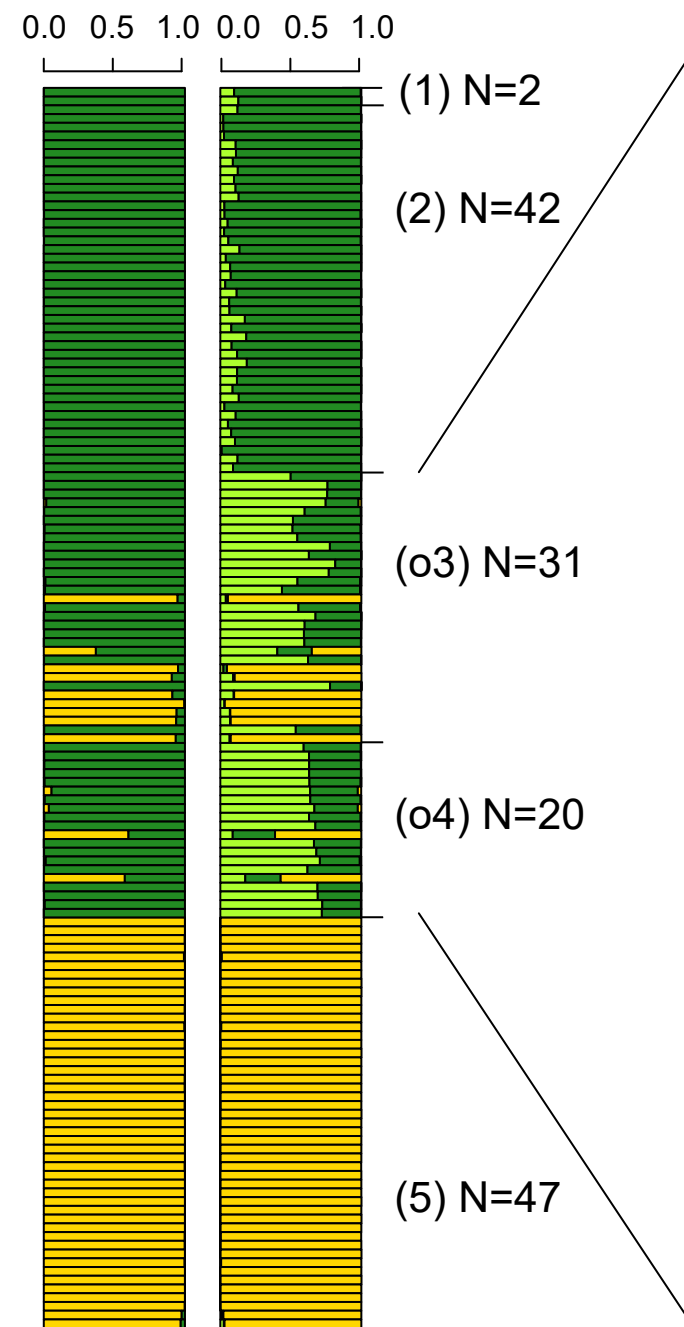
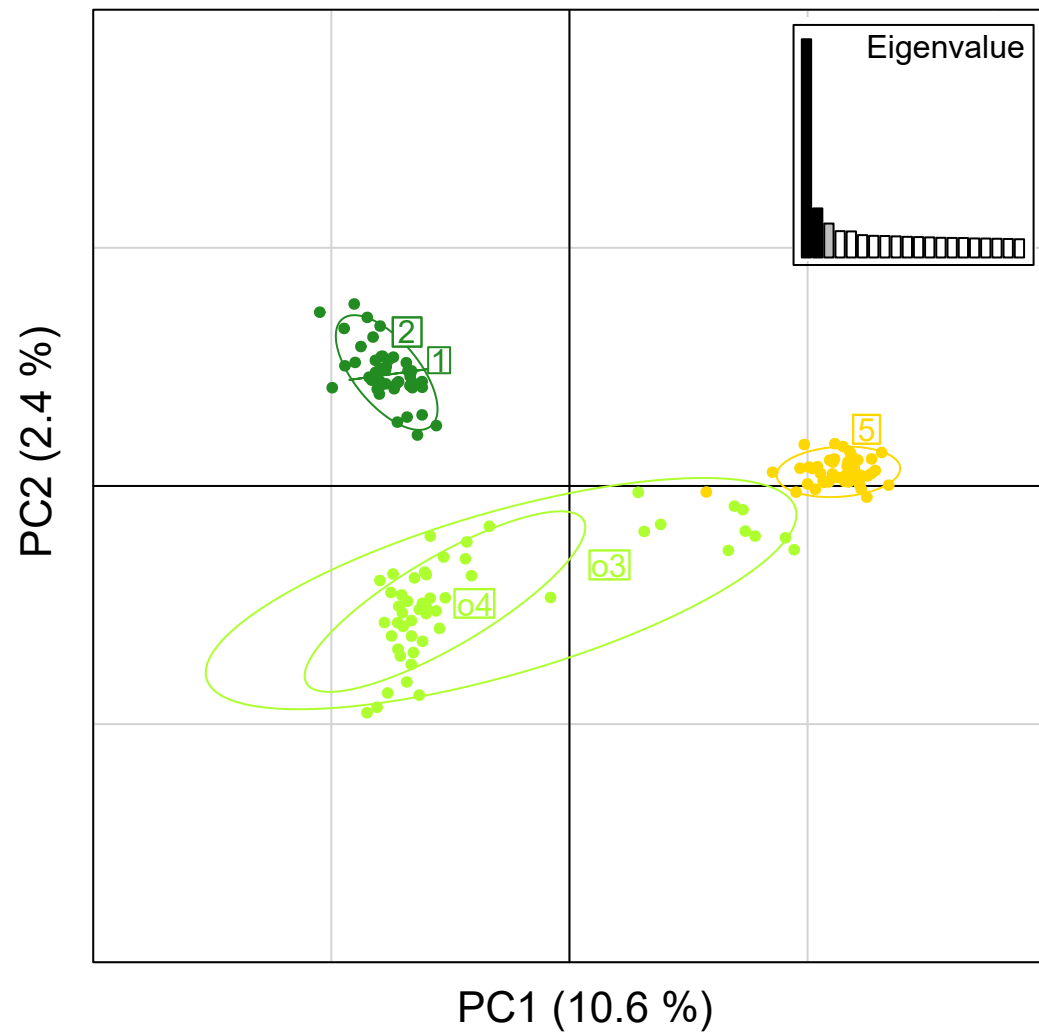
c.



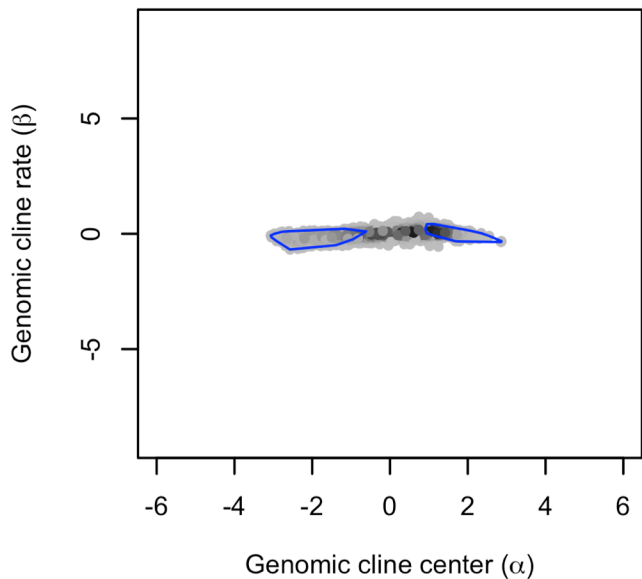
d.



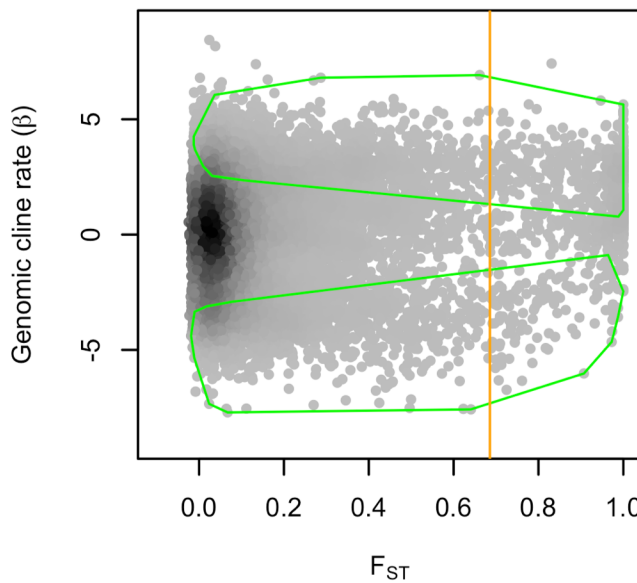
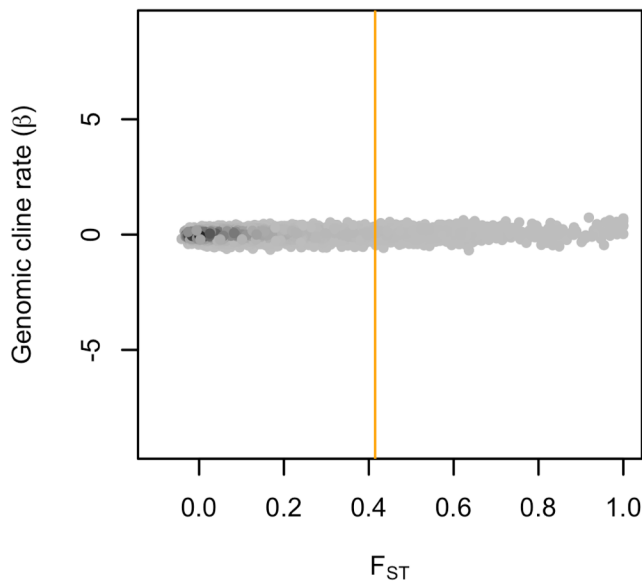
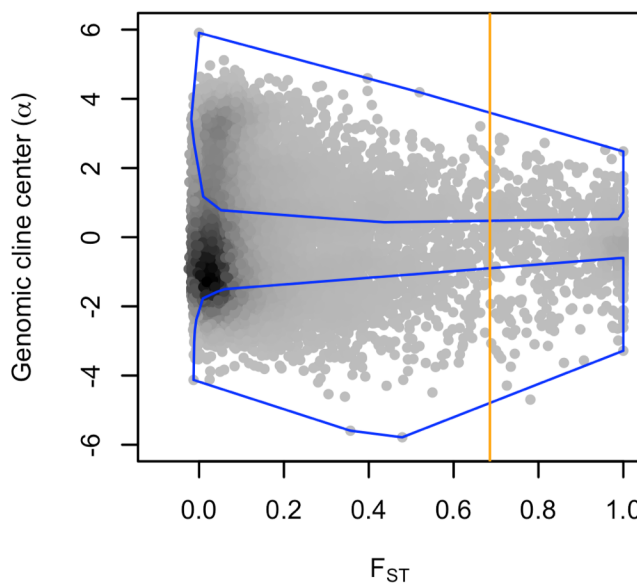
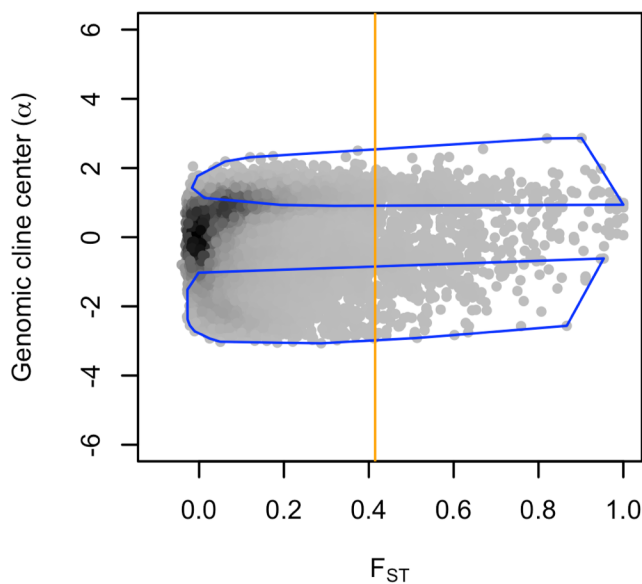
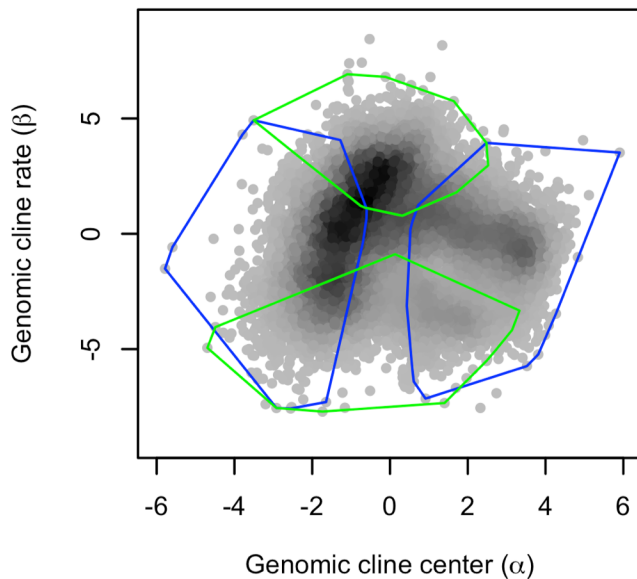
Oleria onega



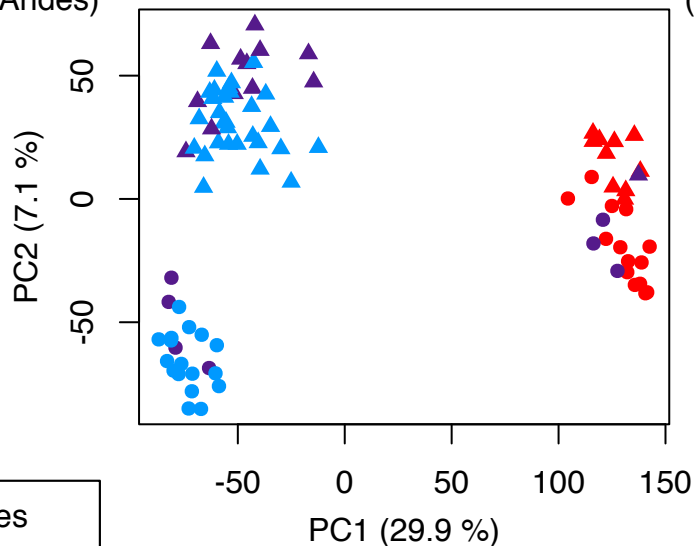
Ithomia salapia



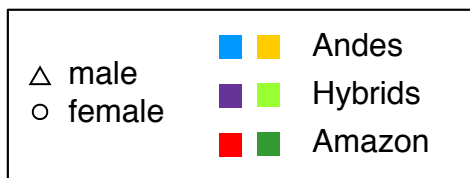
Oleria onega



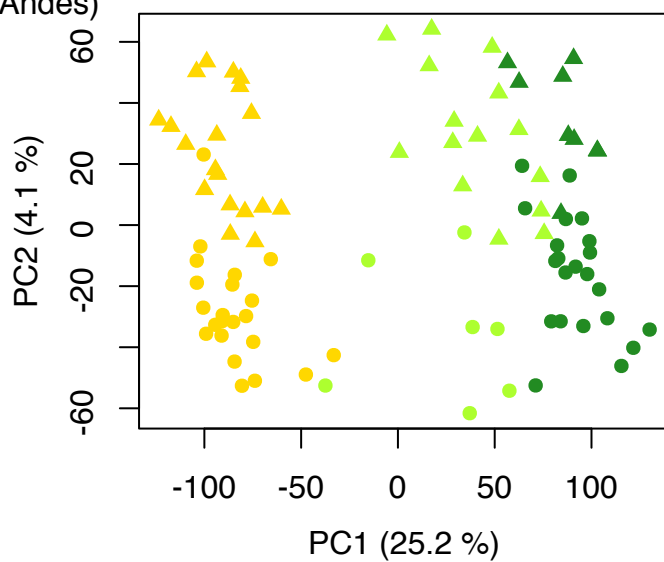
Ithomia salapia derasa
(Andes)



Ithomia salapia aquinia
(Amazon)



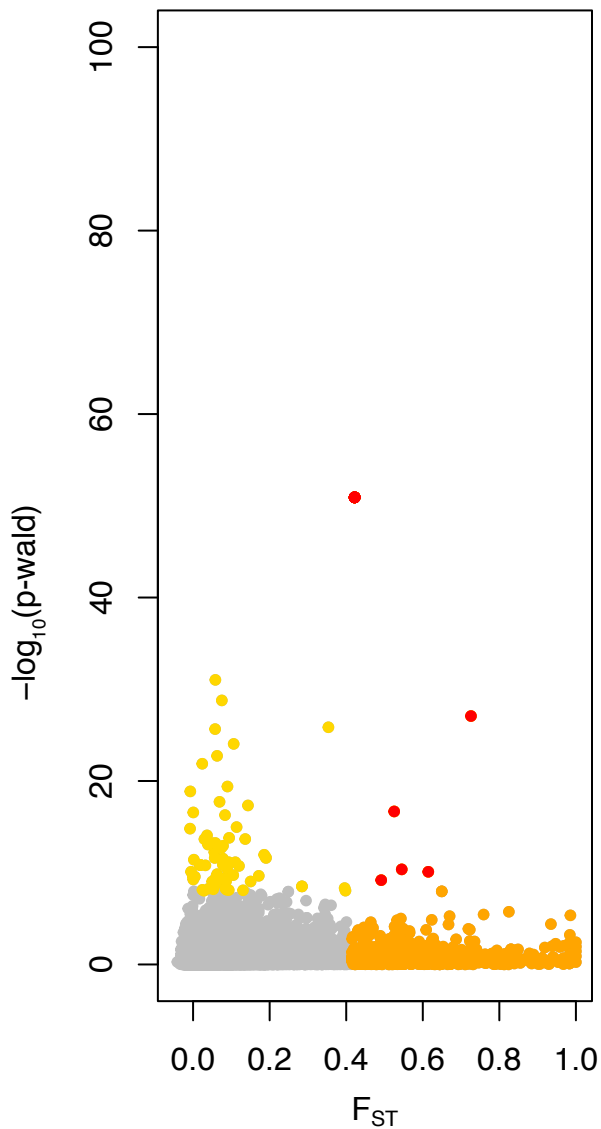
Oleria onega ssp nov 2
(Andes)



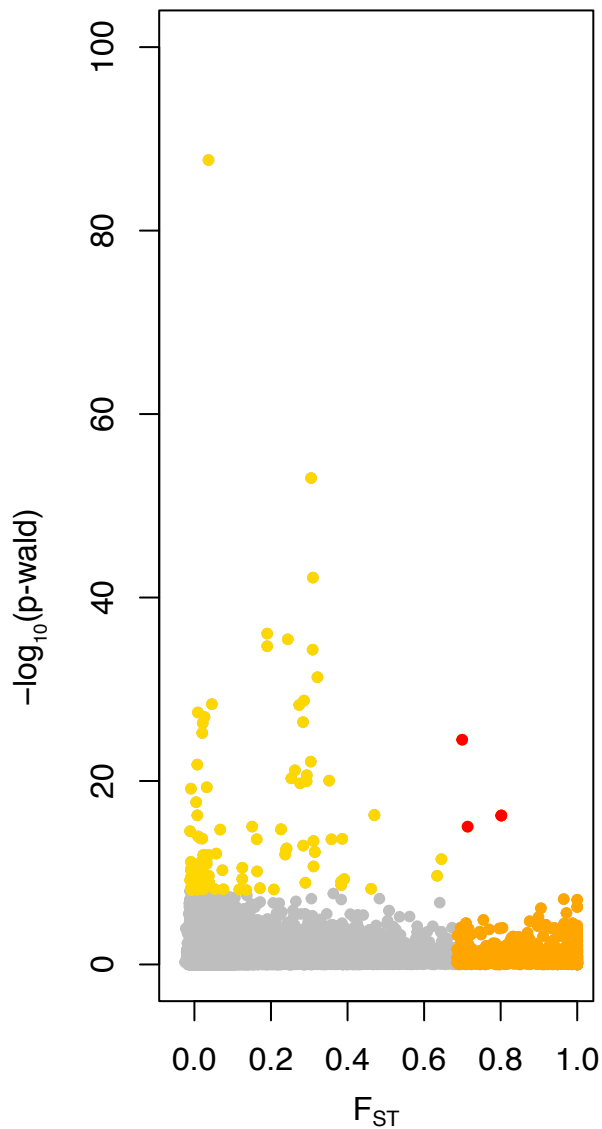
Oleria onega janarilla
(Amazon)



Ithomia salapia

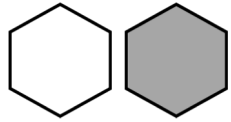


Oleria onega

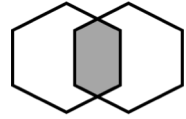


Ithomia salapia

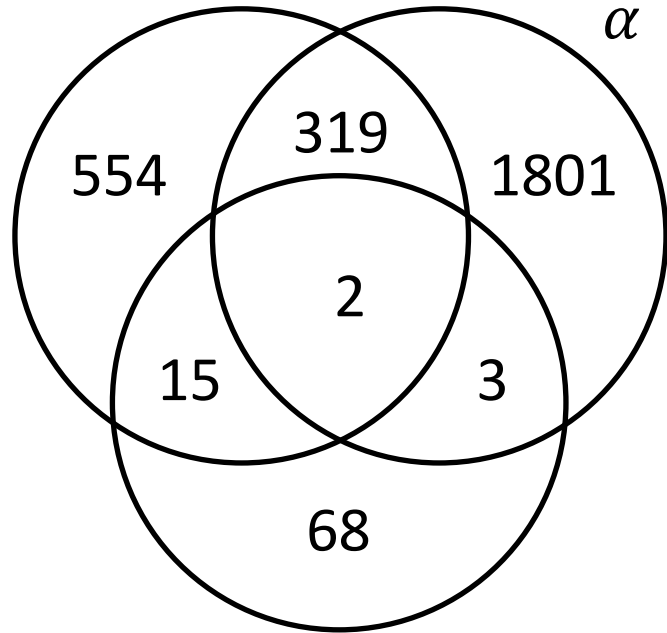
Differentiation
genome-scan



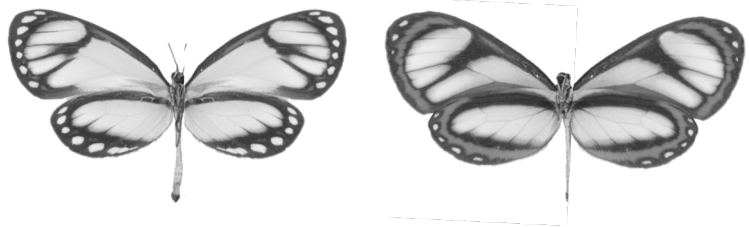
Introgression
pattern



α

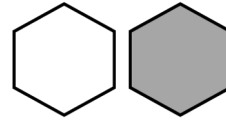


Admixture mapping

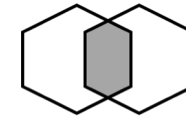


Oleria onega

Differentiation
genome-scan

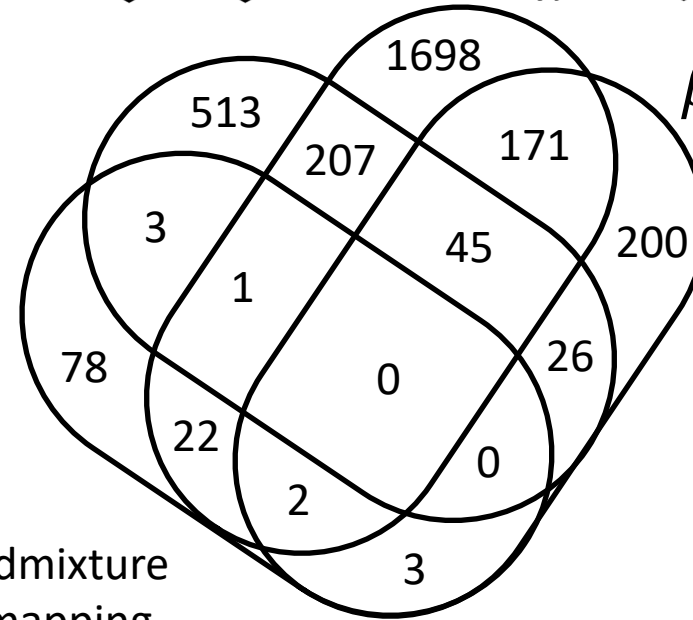


Introgression
pattern



α

$\beta+$



Admixture
mapping



Contrasting genomic and phenotypic outcomes of hybridization between pairs of mimetic butterfly taxa across a suture zone

SUPPLEMENTARY MATERIAL

Supplementary Figure 1. Mimicry ring example for each studied species and lineages, in black frameworks, including various other butterfly species.

Supplementary Figure 2. Barplots with error bars of the Phenotypic Variation Explained (PVE) by genetic for each variable (PC) explaining more than 1% of the wing pattern variation.

Supplementary Table 1. Sampling information including species, population, sex, location, region, GPS positions, sampling date. For each sample, the number of reads sequenced and SNPs called has been given.

Supplementary Table 2. BLAST results of locus with outlier SNPs identified as potentially involved in local adaptation, adaptive introgression and reproductive isolation, i.e. differential genomic cline center (α), high differentiation level (F_{ST}) and differential positive genomic cline rate (β).

Supplementary Table 3. BLAST results of loci with SNPs significantly associated with wing pattern variation.

Locus_ID	BLASTn Best-Hit								Danaus plexippus cds	Protein homology
	Danaus plexippus scaffold	start	stop	%_id	alignment length	e-value	Score			
<i>Ithomia salapia</i>										
Adaptive introgression loci (diverging genomic cline center (α) + high differentiation level (F _{ST}))										
locus_20794	DPSCF300001	2255382	2255463	85.37	82	4,00E-19	95.1	DPOGS207061-TA		
locus_49453	DPSCF300003	1757941	1758023	80.72	83	3,00E-14	78.8	DPOGS203415-TA		
locus_19916	DPSCF300006	780102	780018	83.53	85	4,00E-18	91.5			
locus_20816	DPSCF300008	93324	93266	87.30	63	9,00E-14	77.0			
locus_14403	DPSCF300008	1238880	1238796	96.47	85	1,00E-32	140			
locus_28434	DPSCF300009	1937088	1937006	79.52	83	1,00E-12	73.4	DPOGS208993-TA		
locus_28611	DPSCF300010	1978358	1978428	81.69	71	1,00E-11	69.8			
locus_32677	DPSCF300012	1198547	1198467	81.71	82	1,00E-11	69.8			
locus_91	DPSCF300014	846573	846515	89.83	59	8,00E-15	80.6			
locus_17576	DPSCF300015	681524	681610	86.36	88	3,00E-20	98.7			
locus_24148	DPSCF300017	331577	331503	85.33	75	2,00E-16	86.0			
locus_1838	DPSCF300017	764025	763943	89.16	83	2,00E-23	109			
locus_13643	DPSCF300028	1192556	1192640	77.65	85	5,00E-11	68.0	DPOGS206136-TA		
locus_7322	DPSCF300064	159016	158934	83.13	83	5,00E-17	87.8	DPOGS208522-TA		
locus_22050	DPSCF300073	640050	639972	86.08	79	1,00E-18	93.3			
locus_15377	DPSCF300081	548500	548415	84.88	86	1,00E-19	96.9	DPOGS205859-TA	Down syndrome critical region protein 3 homolog	
locus_18829	DPSCF300091	307474	307391	85.71	84	3,00E-20	98.7			
locus_16344	DPSCF300114	46930	46851	81.25	80	1,00E-13	77.0	DPOGS213191-TA		
locus_15693	DPSCF300121	264102	264039	84.38	64	4,00E-12	71.6			
locus_12790	DPSCF300127	454398	454316	83.13	83	5,00E-17	87.8	DPOGS209486-TA		
locus_11228	DPSCF300127	454413	454480	83.82	68	1,00E-12	73.4	DPOGS209486-TA		
locus_11261	DPSCF300131	603666	603750	80.00	85	1,00E-13	77.0	DPOGS202495-TA		
locus_15366	DPSCF300153	187359	187435	81.82	77	1,00E-13	77.0	DPOGS214916-TA		
locus_14723	DPSCF300154	292667	292598	85.71	70	3,00E-14	78.8			
locus_874	DPSCF300176	896628	896547	84.15	82	2,00E-17	89.7			
locus_11740	DPSCF300177	83862	83792	85.92	71	6,00E-16	84.2	DPOGS207516-TA		
locus_21620	DPSCF300200	112178	112250	82.19	73	1,00E-12	73.4	DPOGS204071-TA	Lactase-phlorizin hydrolase	
locus_5119	DPSCF300211	597	508	76.67	90	3,00E-13	75.2			
locus_11544	DPSCF300211	279599	279549	90.20	51	1,00E-11	69.8			
locus_21834	DPSCF300219	461394	461479	88.37	86	5,00E-24	111	DPOGS213665-TA		
locus_2581	DPSCF300232	508892	508818	82.67	75	1,00E-13	77.0			
locus_21052	DPSCF300268	80273	80342	84.29	70	1,00E-13	77.0			
locus_36372	DPSCF300307	138643	138561	86.75	83	3,00E-20	98.7			
locus_24886	DPSCF300324	76126	76041	86.05	86	2,00E-21	102	DPOGS202701-TA		
locus_63	DPSCF300338	140335	140254	89.02	82	6,00E-23	107	DPOGS207656-TA		
locus_29426	DPSCF300391	9101	9025	87.01	77	4,00E-19	95.1	DPOGS201653-TA		
locus_47081	DPSCF300391	179887	179804	80.95	84	8,00E-15	80.6	DPOGS203107-TA		
<i>Oleria onega</i>										
Adaptive introgression loci (diverging genomic cline center (α) + high differentiation level (F _{ST}))										

locus_39916	DPSCF300001	1760764	1760847	83.33	84	2,00E-17	89.7	DPOGS206900-TA	
locus_6600	DPSCF300001	2445655	2445571	77.65	85	5,00E-11	68.0	DPOGS207070-TA	Bleomycin hydrolase
locus_18737	DPSCF300001	2898435	2898352	84.52	84	1,00E-18	93.3	DPOGS206854-TA	
locus_291	DPSCF300001	4348908	4348825	79.07	86	1,00E-11	69.8		
locus_24405	DPSCF300004	1289625	1289544	84.15	82	2,00E-17	89.7	DPOGS211019-TA	
locus_18889	DPSCF300004	1527412	1527329	88.10	84	6,00E-23	107		
locus_11213	DPSCF300006	94744	94826	81.93	83	2,00E-15	82.4	DPOGS201471-TA	
locus_30280	DPSCF300006	804531	804587	87.72	57	4,00E-12	71.6	DPOGS201447-TA	
locus_6020	DPSCF300009	179886	179969	90.48	84	1,00E-25	116	DPOGS208940-TA	Contactin
locus_63224	DPSCF300009	765102	765185	89.29	84	5,00E-24	111	DPOGS208957-TA	
locus_41437	DPSCF300009	1494818	1494900	95.18	83	1,00E-30	132		
locus_5544	DPSCF300014	973146	973230	87.06	85	7,00E-22	104		
locus_53694	DPSCF300019	1083985	1084069	78.82	85	1,00E-12	73.4	DPOGS212316-TA	
locus_6837	DPSCF300024	24576	24661	83.72	86	4,00E-18	91.5		
locus_939	DPSCF300027	157690	157608	89.16	83	2,00E-23	109		
locus_36129	DPSCF300038	1105994	1106051	91.38	58	2,00E-15	82.4		
locus_28726	DPSCF300041	1563919	1563853	92.54	67	3,00E-20	98.7		
locus_7639	DPSCF300042	1148866	1148782	85.88	85	8,00E-21	100	DPOGS207841-TA	
locus_20657	DPSCF300048	132467	132383	96.47	85	1,00E-31	136		
locus_5259	DPSCF300049	818458	818521	87.50	64	3,00E-14	78.8		
locus_39337	DPSCF300050	197880	197799	87.80	82	7,00E-22	104	DPOGS214621-TA	
locus_5676	DPSCF300051	459989	460071	85.54	83	1,00E-19	96.9	DPOGS207483-TA	
locus_40167	DPSCF300051	482393	482477	88.24	85	2,00E-23	109	DPOGS207445-TA	
locus_51320	DPSCF300052	511744	511813	90.00	70	4,00E-18	91.5		
locus_9796	DPSCF300053	1109584	1109645	83.87	62	5,00E-11	68.0		
locus_30776	DPSCF300056	274712	274639	81.08	74	4,00E-12	71.6	DPOGS205507-TA	
locus_4041	DPSCF300057	52461	52377	77.65	85	5,00E-11	68.0		
locus_11635	DPSCF300058	44474	44542	86.96	69	5,00E-16	84.2	DPOGS208156-TA	
locus_16883	DPSCF300064	259762	259678	84.71	85	4,00E-19	95.1	DPOGS208524-TA	
locus_14881	DPSCF300064	1707322	1707239	89.29	84	5,00E-24	111	DPOGS208460-TA	Cadherin-related tumor suppressor
locus_40108	DPSCF300066	454834	454751	84.52	84	4,00E-18	91.5		
locus_14418	DPSCF300073	565519	565435	92.94	85	6,00E-29	127		
locus_16119	DPSCF300082	1222334	1222416	86.75	83	8,00E-21	100	DPOGS206300-TA	
locus_36677	DPSCF300089	376974	376890	84.71	85	4,00E-19	95.1	DPOGS205916-TA	Serine/threonine-protein kinase N3
locus_35425	DPSCF300094	144496	144424	82.89	76	1,00E-13	77.0		
locus_25295	DPSCF300095	151289	151371	91.76	85	9,00E-27	120		
locus_17924	DPSCF300095	220263	220183	81.48	81	3,00E-14	78.8		
locus_19646	DPSCF300098	588403	588320	94.05	84	2,00E-29	129		
locus_31990	DPSCF300122	617164	617245	82.93	82	2,00E-16	86.0	DPOGS214469-TA	Potassium voltage-gated channel subfamily D member 2
locus_30027	DPSCF300125	447801	447882	84.34	83	5,00E-17	87.8		
locus_23486	DPSCF300128	685283	685351	82.61	69	4,00E-12	71.6	DPOGS200181-TA	Chromodomain-helicase-DNA-binding protein 4
locus_8696	DPSCF300131	292633	292709	79.22	77	5,00E-11	68.0	DPOGS202524-TA	Dynein heavy chain, cytoplasmic
locus_17973	DPSCF300133	205697	205773	80.52	77	4,00E-12	71.6		
locus_9927	DPSCF300145	423602	423519	90.48	84	1,00E-25	116	DPOGS206960-TA	
locus_17455	DPSCF300154	359812	359895	78.57	84	4,00E-12	71.6	DPOGS208109-TA	
locus_38972	DPSCF300162	224765	224849	84.71	85	4,00E-19	95.1	DPOGS202151-TA	
locus_40071	DPSCF300168	21644	21726	81.93	83	2,00E-15	82.4	DPOGS210608-TA	
locus_3709	DPSCF300168	103826	103878	88.68	53	1,00E-11	69.8		
locus_21454	DPSCF300170	429546	429467	86.25	80	4,00E-19	95.1		

locus_14350	DPSCF300171	455112	455200	88.76	89	4,00E-25	114		
locus_38803	DPSCF300182	11272	11350	87.34	79	3,00E-20	98.7	DPOGS216113-TA	
locus_13278	DPSCF300184	311089	310999	81.32	91	2,00E-17	89.7	DPOGS204141-TA	
locus_21125	DPSCF300188	254025	253943	80.72	83	3,00E-14	78.8	DPOGS207347-TA	
locus_27398	DPSCF300206	88373	88457	88.24	85	2,00E-23	109		
locus_14969	DPSCF300210	32672	32755	83.33	84	2,00E-17	89.7	DPOGS203241-TA	E3 ubiquitin-protein ligase Smurf1
locus_12981	DPSCF300212	876257	876175	84.34	83	4,00E-18	91.5	DPOGS213742-TA	
locus_38548	DPSCF300219	579537	579452	84.88	86	1,00E-19	96.9	DPOGS213705-TA	
locus_23097	DPSCF300231	390393	390313	92.59	81	9,00E-27	120	DPOGS204749-TA	
locus_38495	DPSCF300242	161615	161690	82.89	76	3,00E-14	78.8	DPOGS203439-TA	
locus_16210	DPSCF300258	210889	210825	83.08	65	5,00E-11	68.0		
locus_20793	DPSCF300270	226701	226777	81.82	77	1,00E-13	77.0	DPOGS208007-TA	
locus_6434	DPSCF300271	119860	119810	94.12	51	3,00E-14	78.8		
locus_42203	DPSCF300296	235411	235349	87.30	63	3,00E-13	75.2		
locus_3798	DPSCF300325	4824	4749	84.21	76	6,00E-16	84.2	DPOGS211639-TA	
locus_52767	DPSCF300382	158605	158685	90.12	81	5,00E-24	111	DPOGS206254-TA	
locus_21333	DPSCF300547	1509	1588	92.50	80	4,00E-25	114		

Oleria onega

Reproductive isolation loci (high diverging positive genomic cline rate (β))

locus_56496	DPSCF300002	448758	448841	83.33	84	2,00E-17	89.7	DPOGS204463-TA	
locus_23233	DPSCF300002	1313553	1313619	85.07	67	1,00E-13	77.0		
locus_14217	DPSCF300009	1484490	1484408	97.59	83	3,00E-32	138		
locus_41437	DPSCF300009	1494818	1494900	95.18	83	1,00E-30	132		
locus_6715	DPSCF300010	503894	503822	83.56	73	3,00E-14	78.8		
locus_13325	DPSCF300010	1174123	1174177	96.36	55	3,00E-18	91.5		
locus_7285	DPSCF300010	2856044	2856110	97.01	67	1,00E-24	113		
locus_16456	DPSCF300011	910922	910840	81.93	83	2,00E-15	82.4	DPOGS211875-TA	
locus_2045	DPSCF300012	17595	17663	88.57	70	6,00E-16	84.2		
locus_5796	DPSCF300013	724317	724236	85.37	82	4,00E-19	95.1	DPOGS210690-TA	
locus_5544	DPSCF300014	973146	973230	88.24	85	2,00E-23	109		
locus_13402	DPSCF300014	2159403	2159488	95.35	86	3,00E-32	138		
locus_5347	DPSCF300018	791599	791676	83.33	78	2,00E-15	82.4	DPOGS202792-TA	
locus_635	DPSCF300019	386290	386224	91.04	67	4,00E-19	95.1		
locus_6133	DPSCF300021	679108	679192	94.12	85	5,00E-30	131		
locus_17850	DPSCF300021	1255675	1255754	82.50	80	2,00E-15	82.4		
locus_3400	DPSCF300022	1406128	1406086	100.00	43	3,00E-14	78.8		
locus_2334	DPSCF300025	217171	217087	89.41	85	1,00E-24	113	DPOGS210344-TA	Low-density lipoprotein receptor-related protein 2
locus_53	DPSCF300028	1715782	1715719	87.50	64	8,00E-15	80.6		
locus_7065	DPSCF300030	1210754	1210819	89.39	66	4,00E-16	84.2		
locus_56182	DPSCF300032	800216	800170	95.83	48	3,00E-13	75.2		
locus_9208	DPSCF300033	37338	37255	83.33	84	1,00E-17	89.7	DPOGS213564-TA	Protein disulfide-isomerase
locus_39358	DPSCF300034	225066	225150	85.88	85	8,00E-21	100	DPOGS204184-TA	
locus_36129	DPSCF300038	1105994	1106051	91.38	58	2,00E-15	82.4		
locus_9061	DPSCF300041	1435324	1435406	87.95	83	2,00E-22	105	DPOGS215762-TA	
locus_18124	DPSCF300041	1632655	1632739	84.71	85	4,00E-19	95.1	DPOGS215771-TA	Uncharacterized helicase C17H9.02
locus_7639	DPSCF300042	1148866	1148782	85.88	85	8,00E-21	100	DPOGS207841-TA	
locus_15855	DPSCF300044	922991	923073	83.53	85	2,00E-15	82.4		

locus_9976	DPSCF300047	27674	27596	83.54	79	7,00E-16	84.2	DPOGS215265-TA	
locus_13421	DPSCF300048	729948	730032	87.06	85	7,00E-22	104	DPOGS206628-TA	
locus_16663	DPSCF300048	860848	860929	92.68	82	3,00E-27	122		
locus_11355	DPSCF300052	339442	339358	80.00	85	1,00E-13	77.0	DPOGS208609-TA	
locus_51320	DPSCF300052	511744	511813	90.00	70	4,00E-18	91.5		
locus_16883	DPSCF300064	259762	259678	84.71	85	4,00E-19	95.1	DPOGS208524-TA	
locus_1104	DPSCF300066	72184	72264	82.72	81	8,00E-15	80.6		
locus_14418	DPSCF300073	565519	565435	92.94	85	6,00E-29	127		
locus_26323	DPSCF300074	116559	116475	84.71	85	4,00E-19	95.1	DPOGS205078-TA	
locus_3251	DPSCF300082	579057	579142	83.72	86	1,00E-18	93.3	DPOGS206318-TA	DNA-directed RNA polymerase I subunit rpa1
locus_6613	DPSCF300092	467872	467788	89.41	85	2,00E-23	109		
locus_17924	DPSCF300095	220263	220183	81.48	81	3,00E-14	78.8		
locus_17571	DPSCF300099	149029	149094	93.94	66	2,00E-21	102		
locus_6969	DPSCF300112	257821	257733	87.64	89	1,00E-24	113		
locus_42925	DPSCF300117	317371	317306	83.58	67	5,00E-11	68.0		
locus_80	DPSCF300119	307017	306933	98.82	85	2,00E-35	149		
locus_8829	DPSCF300128	685276	685191	77.91	86	1,00E-11	69.8	DPOGS200181-TA	Chromodomain-helicase-DNA-binding protein 4
locus_23486	DPSCF300128	685283	685351	82.61	69	4,00E-12	71.6	DPOGS200181-TA	Chromodomain-helicase-DNA-binding protein 4
locus_10447	DPSCF300129	295733	295816	91.67	84	9,00E-27	120	DPOGS215550-TA	
locus_8696	DPSCF300131	292633	292709	79.22	77	5,00E-11	68.0	DPOGS202524-TA	Dynein heavy chain, cytoplasmic
locus_24658	DPSCF300153	302796	302868	80.82	73	1,00E-11	69.8	DPOGS214919-TA	Structural maintenance of chromosomes protein 4 (Fragment)
locus_50550	DPSCF300160	254056	254135	81.25	80	3,00E-14	78.8		
locus_4637	DPSCF300160	502195	502271	89.61	77	7,00E-22	104		
locus_1279	DPSCF300160	533929	534011	96.39	83	1,00E-31	136		
locus_25495	DPSCF300170	555014	554930	92.94	85	6,00E-29	127	DPOGS204700-TA	Inhibitor of growth protein 4
locus_760	DPSCF300171	509326	509246	91.36	81	4,00E-25	114		
locus_15055	DPSCF300172	330603	330522	81.71	82	8,00E-15	80.6	DPOGS205104-TA	ATP-binding cassette sub-family G member 4
locus_44871	DPSCF300176	877906	877987	81.71	82	8,00E-15	80.6	DPOGS201276-TA	
locus_13278	DPSCF300184	311089	310999	81.32	91	2,00E-17	89.7	DPOGS204141-TA	
locus_10628	DPSCF300189	257	334	82.05	78	3,00E-14	78.8	DPOGS212552-TA	
locus_45809	DPSCF300192	328851	328931	90.12	81	5,00E-24	111		
locus_50	DPSCF300196	630167	630237	92.96	71	2,00E-22	105	DPOGS207656-TA	
locus_36202	DPSCF300224	20899	20974	89.47	76	2,00E-21	102		
locus_3993	DPSCF300235	508167	508086	84.34	83	5,00E-17	87.8		
locus_4021	DPSCF300250	293523	293607	95.29	85	1,00E-31	136		
locus_19317	DPSCF300258	53527	53451	80.52	77	4,00E-12	71.6	DPOGS212418-TA	
locus_16210	DPSCF300258	210889	210825	83.08	65	5,00E-11	68.0		
locus_11201	DPSCF300283	220103	220179	84.42	77	2,00E-16	86.0	DPOGS210186-TA	
locus_21486	DPSCF300298	103299	103230	82.19	73	4,00E-12	71.6	DPOGS215444-TA	
locus_13053	DPSCF300300	50254	50175	91.25	80	1,00E-24	113		
locus_14218	DPSCF300304	184614	184531	79.76	84	3,00E-13	75.2	DPOGS210561-TA	Cyclin-dependent kinase 2
locus_9530	DPSCF300407	236218	236141	84.62	78	6,00E-16	84.2	DPOGS205394-TA	Fructose-1,6-bisphosphatase isozyme 2
locus_878	DPSCF300424	30626	30707	89.02	82	6,00E-23	107		
locus_8293	DPSCF300446	77824	77901	92.31	78	5,00E-24	111		
locus_17	DPSCF300575	2185	2124	90.32	62	2,00E-16	86.0		

Locus_ID	BLASTn Best-Hit							Danaus plexippus cds	Protein homology
	Danaus plexippus scaffold	start	stop	%_id	alignment length	e-value	Score		
<i>Ithomia salapia</i>									
locus_575	DPSCF300006	929124	929055	85.71	70	2E-15	82.4		
locus_10434	DPSCF300037	729941	730017	81.82	77	1E-12	73.4		
locus_6641	DPSCF300038	1144841	1144764	86.42	81	4E-19	95.1	DPOGS212179-TA	Serine/threonine kinase
locus_5230	DPSCF300046	93689	93772	96.43	84	3E-32	138		
locus_11053	DPSCF300054	645838	645902	89.23	65	3E-14	78.8		
locus_3841	DPSCF300071	431640	431701	88.71	62	8E-15	80.6		
locus_15377	DPSCF300081	548500	548415	84.88	86	1E-19	96.9	DPOGS205859-TA	Down syndrome critical region protein 3 homolog
locus_10724	DPSCF300164	45882	45969	85.23	88	3E-20	98.7		
locus_16002	DPSCF300171	38821	38903	80.72	83	3E-14	78.8	DPOGS214065-TA	Carnitine O-palmitoyltransferase 2
locus_2581	DPSCF300232	508892	508818	82.67	75	1E-13	77.0		
locus_9663	DPSCF300395	68400	68317	89.29	84	5E-24	111		
locus_4352	DPSCF300598	15776	15694	80.72	83	3E-14	78.8	DPOGS200524-TA	
<i>Oleria onega</i>									
locus_6600	DPSCF300001	2445655	2445571	80.00	85	1,00E-13	77.0	DPOGS207070-TA	Bleomycin hydrolase
locus_54320	DPSCF300001	2487997	2488051	89.09	55	1,00E-12	73.4	DPOGS207073-TA	
locus_2138	DPSCF300001	5999940	5999893	91.67	48	1,00E-11	69.8		
locus_5821	DPSCF300002	991352	991435	78.57	84	4,00E-12	71.6	DPOGS204406-TA	
locus_18347	DPSCF300012	275034	274963	86.11	72	2,00E-16	86.0		
locus_50	DPSCF300013	1476318	1476235	90.48	84	1,00E-25	116	DPOGS207656-TA	
locus_2696	DPSCF300016	1060706	1060623	83.33	84	2,00E-17	89.7	DPOGS213145-TA	Thioredoxin-like protein 1
locus_291	DPSCF300046	701709	701774	92.42	66	1,00E-19	96.9	DPOGS204311-TA	
locus_4636	DPSCF300051	713048	712964	87.06	85	7,00E-22	104	DPOGS207490-TA	Teneurin-a
locus_302	DPSCF300063	548496	548419	81.18	85	3,00E-14	78.8		
locus_14881	DPSCF300064	1707322	1707239	91.67	84	9,00E-27	120	DPOGS208460-TA	Cadherin-related tumor suppressor
locus_2602	DPSCF300073	701149	701226	88.46	78	1,00E-19	96.9		
locus_8241	DPSCF300086	693323	693404	86.59	82	3,00E-20	98.7		
locus_24045	DPSCF300115	366338	366397	95.00	60	3,00E-19	95.1		
locus_6523	DPSCF300120	68784	68700	83.53	85	4,00E-18	91.5	DPOGS215317-TA	Innexin inx1
locus_12311	DPSCF300162	175614	175527	81.82	88	2,00E-16	86.0		
locus_6481	DPSCF300196	262147	262068	82.50	80	2,00E-15	82.4		
locus_1830	DPSCF300209	256021	255940	86.59	82	3,00E-19	95.1		
locus_47502	DPSCF300221	281710	281775	82.86	70	5,00E-11	68.0		
locus_878	DPSCF300424	30618	30696	92.50	80	5,00E-24	111		
locus_14178	DPSCF300431	40916	41000	77.65	85	5,00E-11	68.0		



*Aeria eurimedia
negricola*



*Ithomia salapia
derasa*
(Andes)



*Napeogenes pharo
lamia*



*Scada reckia
junina*



*Scada reckia
quotidiana*



*Napeogenes pharo
ssp nov*



*Scada reckia
ethica*



*Scada zibia
batesi*



*Episcada sulphurea
ssp nov*



*Ithomia salapia
aquinia*
(Amazon)



Hypoleria xenophis



*Pteronymia primula
primula*



*Mcclungia cymo
subtilis*



*Napeogenes inachia
patentia*



*Napeogenes pharo
pharo*



*Napeogenes inachia
pozziana*



*Napeogenes juanjuiensis
juanjuiensis*



*Pseudoleria aelia
pachiteae*



*Oleria onega
ssp nov 2*
(Andes)



*Hypoleria virginia
oriana*



*Hyoscada zarepha
flexibilis*



*Napeogenes sylphis
rindgei*



*Hypoleria virginia
orianula*



*Hyoscada illinissa
ssp nov 4*



*Oleria onega
janarilla*
(Amazon)



*Oleria onega
lerida*



*Oleria agarista
agarista*



*Oleria astrea
tigilla*



*Oleria gunilla
lotta*



*Oleria gunilla
serdolis*

