Neural divergence and hybrid disruption between ecologically isolated *Heliconius* butterflies

Stephen H Montgomery*, Matteo Rossi, W. Owen McMillan, Richard M Merrill

*s.montgomery@bristol.ac.uk

SUPPLEMENTARY INFORMATION APPENDIX, CONTENTS:

1. Neural divergence and speciation (supplementary introduction)

Table S1: Examples of neural divergence during ecological divergence between populations/species with incomplete reproductive isolation

Species	Environmental context	Neural divergence	Evidence for: Heritable divergence	Evidence for: Selection	Evidence for: Hybrid effects	Evidence for: Behavioural relevance	Key references
A) Vertebrates							
Astyanax mexicanus	Micro-habitat divergence (surface/cave systems)	Brain morphology	Confirmed using common garden experiments	Not tested?	Evidence of intermediate traits inferred from [1]	Not tested, but likely linked to differences in use of sensory modalities.	[1–3]
Coregonus cupleaformis	Micro-habitat divergence (limnetic/benthic)	Neural gene expression	Confirmed using common garden experiments	Overlap between eQTLs and peaks in genome-wide patterns of divergence	Misexpression of neural genes in hybrids	Not tested, but likely linked to differences in use of sensory modalities.	[4,5]
Cichlidae	Micro-habitat divergence (adaptive radiation)	Brain morphology	Absent among congeneric sp., implied by [6]	Not tested?	Not tested?	Links to brain morphology not tested but see [7]	[8,9]
Gasterosteus aculeatus	Micro-habitat divergence (benthic/generalist)	Brain morphology	Not tested?	Not tested?	Not tested?	Not tested?	[10]
Gasterosteus aculeatus	Micro-habitat divergence (limnetic/benthic)	Brain morphology	Not tested?	Not tested?	Not tested?	Not tested?	[11]
Poecilia reticulata	Micro-habitat divergence (up/down stream and low/high predation)	Brain morphology	Not tested?	Not tested?	Not tested?	Not tested?	[12]
Poecilia mexicana	Micro-habitat divergence (surface/darkness/toxic hydrogen sulphide streams)	Brain morphology	Rejected using common harden experiments	Not tested?	Not tested?	Not tested?	[13]
Pungitius pungitius	Micro-habitat divergence (river/lake/sea)	Brain morphology	Not tested?	Not tested?	Not tested?	Not tested?	[12]

B) Invertebrates

Culicidae	Host shifts	Olfactory circuits	Inferred from lab experiments	Not tested?	Not tested?	Host odour detection	[15]
Drosophila melanogaster/sechellia	Host plant shifts	Olfactory circuits/antennal lobe morphology	Inferred from lab experiments	Not tested?	Not tested?	Host plant odour detection	[16,17]
Drosophila pseudoobscura/subobscura	Micro-habitat divergence hypothesised, but there is limited ecological data and contemporary populations occur on separate continents	Brain morphology	Implied by using lab populations experiments	Not tested?	Not tested?	Evidence for divergent mating behaviours, but no formal tests for an association with divergent brain morphology	[18]
Heliconius erato cyrbia/himera	Micro-habitat divergence	Brain morphology	Confirmed using common garden experiments	Not tested?	Not tested?	Not tested?	[19]
Heliconius melpomene/cydno	Micro-habitat divergence	Brain morphology, neural gene expression	Confirmed using common garden experiments	Evidence from PST/FST analyses and reduced gene flow around differentially expressed loci	Intermediate brain morphology and misexpression of neural genes in hybrids	Not tested, but likely linked to habitat dependent light regimes	Current study
Rhagoletis pomonella	Host plant shifts	Olfactory circuits	Inferred from lab experiments	Evidence of fitness trade offs [20]	Not tested?	Host plant odour detection	[21,22]

43

2. Materials and methods

2 i. Sampling of wild individuals

Although they began to diverge ~2 million years ago [23], *melpomene* and *cydno* have a long history of persistent gene flow [23,24], and in that sense speciation is considered to be incomplete. Distributed across Central and South America, the species boundary is maintained by ecological divergence and disruptive selection against hybrids [25–28], which now occur at low frequencies [29]. Although deep intra-clade splits are similar in age [23], populations of *melpomene* are currently ascribed to 'races' (defined by colour pattern variation), while several *cydno* lineages (*timareta, pachinus*) have been promoted to species level [30]. The geographic distribution of the *melpomene* and *cydno* clades allowed us to sample a series of populations, accounting for both geographic divergence on small and continental scales, as well as ecological divergence. All individuals were collected using hand nets and kept alive in glassine envelopes until brain tissue could be fixed within a few hours of collection. Sampling of wild individuals (Table S2) was focused on four countries:

Panama

In Panama, *H. c. chioneus* is found in closed forest habitats whereas *H. m. rosina*, occurs in secondary forest [25,31]. We sampled 10 individuals of each species along Pipeline road, Gamboa (elevation 60 m), which transects open to closed forest, and the nearby Soberanía National Park. Samples were collected under permits SEX/A-3-12, SE/A-7-13 and SE/AP-14-18.

Peru

H. timareta is a member of the *cydno* clade restricted to mid-elevation forest on the eastern Andes. In Peru, *H. t. thelxinoe* is in mosaic sympatry with *H. m. amaryllis*, with which it shares a co-mimetic wing pattern. Like low-elevation *H. cydno*, *H. timareta* is specialised for closed forests [26,32], suggesting micro-habitat partitioning from *H. melpomene* is maintained across the *cydno* clade. Because they are isolated by ecology but not by mimicry ring, this pair provides a 'control' for neuroanatomical divergence associated with visual mate cues. 10 individuals of each species were samples in the Escalera region near Tarapoto, Departamento de San Martín (elevation 300-1295 m). Samples were collected under permits 0289-2014-MINAGRI-DGFFS/DGEFFS, 020-014/GRSM/PEHCBM/DMA/ACR-CE, 040–2015/GRSM/PEHCBM/DMA/ACR-CE, granted to Dr Neil Rosser.

Costa Rica

H. c. galanthus and *H. pachinus* are parapatric species, within the cydno clade, that are restricted to opposite coastal drainages in Costa Rica. In spite of evidence for ongoing gene flow, few hybrids have been collected suggesting that strong reproductive isolation maintains the species barrier [33,34]. Except for differences in colour pattern, there are no known ecological differences between the two taxa suggesting they are ecologically equivalent and the product of allopatric speciation across the Central Valley [33–35]. This provides a 'control' speciation event where we do not expect neuroanatomical divergence between species. 10 *H. c. galanthus* were sampled at La Selva Biological Station (elevation 30-130 m) and Orosí (elevation ~1300 m). 10 *H. pachinus* were sampled at Las Cruces Biological Station (elevation <20 m) and Le Leona eco-lodge on the edge of Corcovado National Park (elevation ~1000 m). A small number of *H. m. rosina* were also collected from these locations. Samples were collected under permit SINAC-SE-GASP-PI-R-2015.

French Guiana

At the eastern extreme of its geographic distribution, *H. melpomene* is allopatric with *cydno*. *H. m. melpomene* shares its general ecology with its western relatives, with some exceptions. *H. m. melpomene* is more oligophagus in its larval food plants [36] and there is some suggestion that it uses the forest interior to a greater extent [37], although data supporting this observation is lacking. As we sampled 10 *H. m. melpomene* from forest edge habitats in the Arrondissement of Cayenne (elevation 0-150m), we consider them to have been exposed to similar micro-habitats as *melpomene* in Peru and Panama. We therefore use this population to construct a test of character displacement in brain morphology between sympatric *melpomene/cydno* species. At the time of sampling no permits were required to sample outside National Parks in French Guiana.

2 ii. Insectary reared animals

To determine whether any variation we observed was due to environmentally-induced plasticity, we performed common garden experiments focusing on the Panamanian species pair, *H. c. chioneus* and *H. m. rosina*. Insectary-reared individuals were obtained from wild-caught females. Adults were kept under standard conditions in outbred stock cages (c. 1 x 2 x 2 m) of mixed sex and equal densities at the Smithsonian Tropical Research Institute's Gamboa insectaries. These cages are maintained on the edge of the butterfly's native habitat, and light conditions do not substantially deviate from the forest edge environment. Stock cages contained a minimum of 10 females. Because *H. m. rosina* are monophagus, larvae were reared on the species' preferred host plant (*Passiflora menispermifolia* and *P. triloba* respectively). To assess whether hybrid individuals show intermediate or disrupted

phenotypes we produced multiple *H. c. chioneus* x *H. m. rosina* crosses in both directions, during two distinct field seasons (2013, 2019). Samples from the 2013 crosses were used for neuroanatomical measurements, while samples from both sets were used to collect gene expression data. We focus on F1 individuals because they are the represent a large portion of hybrids found in natural *Heliconius* hybrids [29] and must survive to produce back crosses with parental species. F1 larvae were reared on *P. triloba.* For both pure species and hybrid crosses, eggs were collected from the host plants on a daily basis over an ~8-week period, and isolated until hatching. Individual larvae were then raised on new growth shoots in outdoor larval cages. After eclosion, adults were aged for 2-3 weeks for the neuroanatomical samples, and 9-15 days for gene expression samples (Table S2, S3). Both sexes are sexually and behaviourally mature at ~8 days [38].

2 iii. Neuroanatomy protocols

Brains were fixed *in situ* using a $ZnCl_2$-formaldehyde solution, following Ott [39]. Further methodological details and anatomical descriptions of the *Heliconius* brain are available in Montgomery et al. [40]. Briefly, brain structure was revealed using immunofluorescence staining against a vesicle-associated protein at presynaptic sites, synapsin (anti-SYNORF1; obtained from the Developmental Studies Hybridoma Bank, University of Iowa, Department of Biological Sciences, Iowa City, IA 52242, USA; RRID: AB_2315424) and Cy2-conjugated affinity-purified polyclonal goat anti-mouse IgG (H+L) antibody (Jackson ImmunoResearch Laboratories, West Grove, PA), obtained from Stratech Scientific Ltd., Newmarket, Suffolk, UK (Jackson ImmunoResearch Cat No. 115-225-146, RRID: AB_2307343). All imaging was performed on a confocal laser-scanning microscope (Leica TCS SP5 or SP8, Leica Microsystem, Mannheim, Germany) using a 10x dry objective with a numerical aperture of 0.4 (Leica Material No. 11506511), a mechanical *z*-step of 2µm and an *x-y* resolution of 512 x 512 pixels. The *z*-dimension was scaled by 1.52 to correct the artefactual shortening [40]. We assigned image regions to brain components, or neuropils, using the Amira 5.5 (Thermo Fisher Scientific) *labelfield* module and defining outlines based on the brightness of the synapsin immunofluorescence. We reconstructed total central brain volume (CBR), six paired neuropils in the optic lobes (OL), six paired and one unpaired neuropils in the central brain (CBR) in all wild individuals, using the *measure statistics* module to estimate component volumes. In insectary samples the POTu, a small posteriorly located neuropil, was inconsistently stained and was not measured, and in hybrids only neuropils with evidence of divergence between *melpomene* and *cydno* were segmented. The total volume of segmented structures in the CBR was subtracted from total CBR volume to obtain a measure of the remaining, unsegmented CBR (rCBR), which is used as an allometric control throughout. Note that in insectary samples rCBR does not include POTU, and among comparisons including

hybrids rCBR is simply CBR minus AOTU volume. Due to the lack of volumetric asymmetry in *Heliconius* neuropils [40] we measured the volume of paired neuropils from one hemisphere, chosen at random unless one hemisphere was damaged, and multiplied the measured volume by two. All volumes were $\log_{10}$-transformed before data analysis.

2 iv. Statistical analyses of neuropil volumes

We identified non-allometric differences between brain component sizes using nested linear models, analysed in the lme4 R package [41]. Linear models included each brain component as the dependent variable, rCBR and taxonomic/experimental grouping as an independent variable, with sex and country (where relevant) included as random factors. The likelihoods of nested models were compared using a Chi^2-test. Correction for multiple testing was performed using a sequential Bonferroni procedure [42]. For neuropils showing a significant clade/species effect, we subsequently explored the scaling parameters responsible for group differences using SMATR v.3.4-3 [43]. Using the standard allometric scaling relationship: $\log y = \beta \log x + \alpha$, where y is the brain component of interest and x is rCBR, we performed tests for significant shifts in the allometric slope (β) between taxa, followed by two further tests which assume a common slope: 1) for differences in α that suggest discrete 'grade-shifts' in the relationship between two variables, 2) for major axis-shifts along a common slope. Deviation from a shared scaling relationship, by slope or elevation, can indicate an adaptive change in the functional relationship between two brain structures [44].

In addition to our allometrically controlled regressions, we performed two further analyses to explore the role of selection in neuroanatomical divergence. First, using data from wild-caught samples, we performed a Mantel test between pairwise differences in neuropil volumes and two estimates of F_{ST} from Arias et al. [30], based on AFLPs and mtDNA. *H. m. amaryllis* was not included in Arias et al. we therefore use *H. m. malleti* as a surrogate, as these two colour pattern races are geographically and phylogenetically close [45]. Pairwise differences in neuropil volumes were taken as $\log_{10}(|Population_A - Population_B|+1)$. Partial Mantel tests, controlling for pairwise differences in rCBR volumes, were performed using ECODIST [46] with Pearson correlations and 1000 permutations.

Second, with insectary reared samples, we calculated P_{ST} using the PSTAT package [47] initially using a c/h^2 ratio of 1, where c is the proportion of the total variance presumed to be due to additive genetic effects across populations, and h^2 is the trait heritability. Quantitative genetic parameters for invertebrate neuroanatomy are sorely lacking in the literature, with only a small number of heritability estimates (0.123-0.376) for linear dimensions of *Drosophila* mushroom body size [48]. We therefore also varied the c/h^2 ratio assuming c equals 0.25, 0.50, 0.75, 1.00, and h^2 equals 0.25, 0.50, 0.75, resulting in ratios of (0.33, 0.67, 1.00, 1.33, 1.50, 2.00, 3.00, 4.00) to test how sensitive the P_{ST} estimates are to these assumptions. P_{ST}

calculations were performed on raw, $\log_{10}$-transformed neuropil volumes, and on residual volumes after regressing neuropil volumes against rCBR using the res() function. To test if an individual neuropil's P_{ST} was significantly higher than expected by neutral divergence, we calculated a p-value as the proportion of the F_{ST} distribution [24] (see below) that was above each P_{ST} value, where a P_{ST} value above the 95th percentile of the F_{ST} distribution is taken as evidence of selection.

Finally, we identified intermediate traits in hybrids we also performed Principal Component Analysis and ANOVAs among parental and hybrid individuals, with post-hoc Tukey tests to compare group means, using base R packages [49] (R Core Team, 2013).

2 v. RNA extraction and sequencing

All samples used for our comparative transcriptomics were reared in common garden conditions (see above). Brains were dissected out of the head capsule in cold (4 ºC) 0.01M PBS solution and include the CBR, OL and ommatidia. Full details of library preparation and sequencing are available in Rossi et al. [50], but briefly, in 2014, mRNA was extracted from whole brains of age-matched *H. c. chioneus* (n =11), *H. m. rosina* (n = 12), and F1 hybrids in 2014 (n = 4), using TRIzol Reagent (Thermo Fisher, Waltham, MA, USA) and a RNeasy kit (Qiagen, Valencia, CA, USA). Samples were treated with DNase I (Ambion, Darmstadt, Germany), and Illumina libraries were prepared and sequenced with 100bp paired-end reads at Edinburgh Genomics (Edinburgh, UK). In 2019, mRNA was extracted from whole brains of age-matched *H. c. chioneus* (n = 5)*, H. m. rosina* (n = 5), and F1 hybrids (n = 12), using TRIzol Reagent and a PureLink RNA Mini Kit, with PureLink DNase digestion on column (Thermo Fisher, Waltham, MA, USA). Illumina 150bp paired-end RNA-seq libraries were prepared and sequenced at Novogene (Hong Kong, China).

After trimming adaptor and low-quality bases from raw reads using TrimGalore v.0.4.4 (www.bioinformatics.babraham.ac.uk/projects), Illumina reads were mapped to the *H. melpomene* 2 genome [51]/*H. melpomene* 2.5 annotation [52] using STAR v.2.4.2a in 2-pass mode [53]. We kept only reads that mapped in 'proper pairs', using Samtools [54]. The number of reads mapping to each gene was estimated with HTseq v. 0.9.1 (model = union) [55].

2 vi. Statistical analyses of gene expression data

Differential gene expression analyses between groups were conducted in DESeq2 [56], including sex and sequencing batch as random factors. We considered only those genes showing a 2-fold change in expression level, and at adjusted (false discovery rate 5%) p-values < 0.05, to be differentially expressed due to the potential for tissue composition to drive significant, but low-fold change differences in expression [57].

We conducted a Principal Component Analysis on rlog-transformed gene count data (as implemented in DESeq2) to inspect clustering of expression profiles among groups (species, or hybrids) for all genes, and for differentially expressed genes only.

We performed ANOVAs on normalized gene expression counts of species and hybrids, with post-hoc Tukey tests, and categorized gene expression levels in hybrids as follows:

1. "*melpomene*-like": where F1 vs. *cydno* p<0.05 *and* F1 vs *melpomene* p>0.05
2. "*cydno*-like": where F1 vs. *cydno* p>0.05 *and* F1 vs *melpomene* p<0.05
3. "intermediate": where F1 vs. *cydno* p>0.05 *and* F1 vs *melpomene* p>0.05, **or** *cydno* p<0.05 *and* F1 vs *melpomene* p<0.05, with F1s having an intermediate mean expression level between parental species
4. "transgressive": where *cydno* p<0.05 *and* F1 vs *melpomene* p<0.05, with F1 having higher or lower mean expression compared to both parental species.

We estimated phenotypic differentiation in gene expression (P_{ST}) from normalized gene counts in *H. m. rosina* and *H. c. chioneus*, following Uebbing et al. [58]. For the main analysis we set heritability to be 0.5, but examined the effects of varying heritability in the supplementary results. The distribution of genome-wide genetic differentiation (F_{ST}) between *H. m. rosina* and *H. c. chioneus* were retrieved from Martin et al. [24]. To test if an individual gene's P_{ST} was significantly higher than expected by neutral divergence, we calculated p-values as the proportion of the P_{ST} distribution that was above each F_{ST} value. (P_{ST} values above the 95th percentile of the F_{ST} distribution, were considered as showing evidence of selection). We subsequently explored how the frequency of this index of selection varied between gene expression categories (differentially expressed or not, and hybrid categories as defined above) by comparing the proportion of genes showing $P_{ST} > F_{ST}$ q(95%), in the various gene categories).

To further test whether genes highlighted by these analyses contribute to divergence between *melpomene* and *cydno*, we asked whether genes with $P_{ST} > F_{ST}$ q(95%), or with intermediate expression in hybrids, were more likely to occur in regions of the genome with low levels of gene flow. We retrieved estimated admixture proportions (f_d) between *H. m. rosina* and *H. c. chioneus,* and population recombination rates (*rho*) from Martin et al. [59]. We then investigated the relationship between f_d (estimated in 100kb non-overlapping windows) and P_{ST}, accounting for variation in recombination rate, as a way to study whether selection acts against introgression of foreign alleles. In this analysis, we fitted the following generalized linear mixed models (glmm): $f_d \sim rho + P_{ST} + (1|chromosome)$, assuming a Gaussian distribution. We also explored whether genes with intermediate expression in F1s

showed higher levels of P_{ST}, and lower levels of f_d, compared to other genes, with glmm models: *intermediate_Y_or_N ~ X + (1|chromosome), where X =* f_d or P_{ST}, assuming a binomial distribution. Inclusion of chromosome as a random factor provides partial correction for the effects of physical linkage between sites, however, we acknowledge this analysis may still be prone to inflated effect sizes due to non-independence of genomic regions. To test for differences in levels of P_{ST} and f_d values among all gene categories we conducted a Kruskal-Wallis test with post-hoc Dunn test (with Bonferroni correction).

Finally, to infer possible overrepresentation of specific molecular functions among gene categories, we first used InterProScan v.5 [60] to retrieve gene ontology (GO) terms associated with every gene annotated in the Hmel2.5 genome. We then conducted a gene set enrichment analysis (Fisher's exact test, $p < 0.01$) with the TopGO package in R [61], using the "elim" algorithm, which corrects for non-independence among GO terms. We identified GO term enrichment tests for differentially expressed genes, genes showing intermediate expression in F1 hybrids, and genes with $P_{ST} > F_{ST}$ q(95%), relative to all other genes, using Fisher's exact tests at $\alpha = 0.05$.

3. Supplementary results

3 i. Intra-clade variation in brain morphology

In cydno, no neuropil shows significant intra-clade divergence in scaling with an allometric control (Table S4B, Table S5B), whereas several do in melpomene. Below, we illustrate this difference with two neuropils, the medulla (A,C) and AOTU (B,D). Post-hoc tests in SMATR (following Table S5B) identify no significant shifts in the *cydno* clade for either the medulla (figure S1A) or AOTU (Figure S1B), consistent with our lme4 analyses (Table S4B). In the *melpomene* clade, *H. m. rosina* has a divergent slope from *H. m. amaryllis* for the medulla (X^2 = 4.096, p = 0.043), but this difference is not significant compared to *H. m. melpomene* (X^2 = 2.585, p=0.108) (Figure S1C). However, *H. m. melpomene* does have major-axis shift in medulla size relative to *H. m. amaryllis* (Wald-statistic = 21.235, p<0.001) and *H. m. rosina* (X^2 = 7.386, p = 0.007) consistent with coordinated expansion of the medulla and rCBR (Figure S1C). For the AOTU (Figure S1D), *H. m. rosina* has a divergent scaling relationship with rCBR compared to *H. m. amaryllis* (X^2 = 8.382, p = 0.004) and, to a lesser extent, *H. m. melpomene* (X^2 = 3.352, p = 0.067), whereas and *H. m. amaryllis* and *H. m melpomene* have grade shift on the *y*-axis (test=6.908, p=0.009) indicating non-allometric variation in AOTU size between these population.

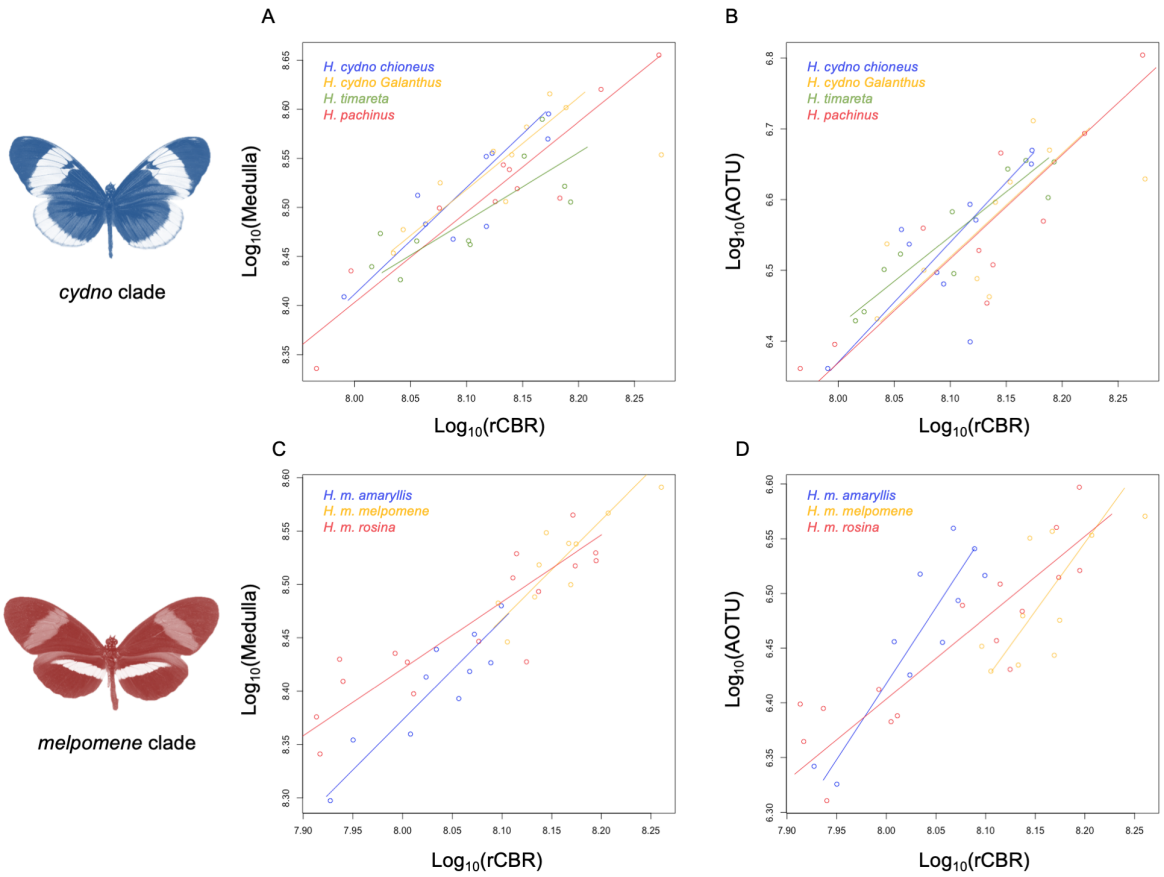

Figure S1. Example of intra-clade variation in neuropil scaling in the *cydno* (A,B) and *melpomene* (C,D) clades for the medulla (A,C) and AOTU (B,D). See also Table S4B and Table S5B.

3 ii Tests of neutrality on neuropil volumes

Pairwise volumetric differences in the visual neuropils, or total optic lobe volume, are not generally associated with F_{ST} across populations (Table S8A). Although power is likely limited by the number of populations, the only neuropils that suggest a relationship between trait divergence and F_{ST} are the lobula (LOB), antennal lobe (AL) and both mushroom body components (MBCA, MBLOPE), with only LOB showing an association at $p<0.05$ for F_{ST} estimates from both mtDNA and AFLP data. None of these four neuropils show evidence of non-allometric shifts in relative size between the *melpomene* and *cydno* clades. In contrast, all neuropils that show this pattern of divergence lack associations with F_{ST}. Under the association phenotypic drift is linearly associated with neutral genetic divergence this provides evidence for a role of selection in driving divergence in brain composition across the *cydno-melpomene* clade.

P_{ST} estimates based on comparisons between *H. m. rosina* and *H. c. chioneus* reared in common-garden conditions are consistent with this interpretation (Table S8B). With the exception of components of the central complex (PB and CB) all P_{ST} estimates for raw volumes are above the 95th percentile of the F_{ST} distribution, most likely reflecting divergence in total brain size. However, after accounting for allometric variation through regressions against central brain volume (rCBR), significant $P_{ST}>F_{ST}$ effects are only detected for total optic lobe size (OL), lamina (LAM), medulla (ME), lobula (LOB), ventral lobe of the lobula (vLOB), and the anterior optic tubercule (AOTU). All are robust to correcting for multiple tests. Varying the c/h^2 ratio suggests these results are widely robust to assumptions about the proportion of genetic variance and heritability (Table S8C). All significant $P_{ST}>F_{ST}$ results are recovered except under low a c/h^2 ratio, where the proportion of total variance accounted for by additive genetic variance is low, and heritability is high. A scenario we suspect is unlikely. Even under this scenario, LAM, vLOB and AOTU show significant $P_{ST}>F_{ST}$ results before correcting for multiple tests. Under high c/h^2 ratios above 1, the lobula plate (LOP) also has significant $P_{ST}>F_{ST}$.

Taken together at least LAM, ME, vLOB, aME and AOTU show greater degrees of phenotypic divergence between Panamanian *H. m. rosina* and *H. c. chioneus*, *and* an absence of an association with neutral divergence across the *cydno-melpomene* clade, which is highly suggestive of adaptive evolution, which ultimately affects overall OL size. Our data also suggests that the LOP and LOB have been under divergent selection between *melpomene* and *cydno*, but potentially with less consistency across the clade.

3 iii. Hybrid disruption

In addition to having intermediate volumes relative to rCBR volume, hybrids also show some evidence of intermediate scaling between pairs of neuropils. Scaling analyses between pairs of visual neuropils in SMATR identify several comparisons with significant deviation in scaling across parental species and F1 hybrids, affecting either the slope or elevation of the scaling relationship (Table S7Di, Dii). Many of these cases reflect pairs of neuropils with direct connections in other insects, including the LOB and vLOB [62], ME and AOTU [63], LOB and AOTU [63], LAM and aME [64], or where there are likely indirect functional connections, e.g. LAM and LOP, which are connected via projections to the MED [65], or the aME and AOTU which both process polarised light [64,66,67].

Post-hoc analyses of tests with $p<0.10$ suggest that F1 hybrids show potentially intermediate scaling relative to *cydno* and *melpomene*. For example, the elevation constant for scaling between the LAM and aME in F1 hybrids is intermediate between *H. cydno* and *H. melpomene*, although in both cases it is marginally non-significant (*H. cydno* wald = 3.665, p = 0.056; *H. melpomene* wald = 3.758, p = 0.053; Figure S1A). In other cases, scaling in F1 hybrids is significantly different from one parental species, but not both (Table S7D iii; Figure S1B,C). However, different pairs of neuropils show different parental similarities, with some scaling like *H. melpomene* (e.g. aME~AOTU), while other scale like *H. cydno* (e.g. aME~VLOB, LAM~LOB). We suggest that this provides a second potential avenue for hybrid disruption if information flow between multiple neuropils are unbalanced.

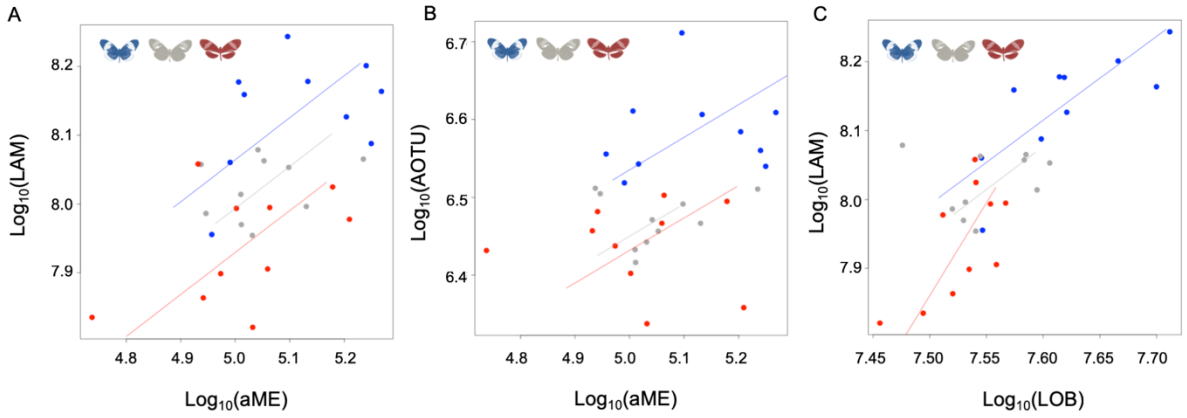

Figure S2. Examples of scaling between pairs of neuropils: A) LAM~aME (intermediate), B) AOTU~aME (*melpomene*-like); C) LAM~LOB (*cydno*-like).

3 iv. Divergence in gene expression

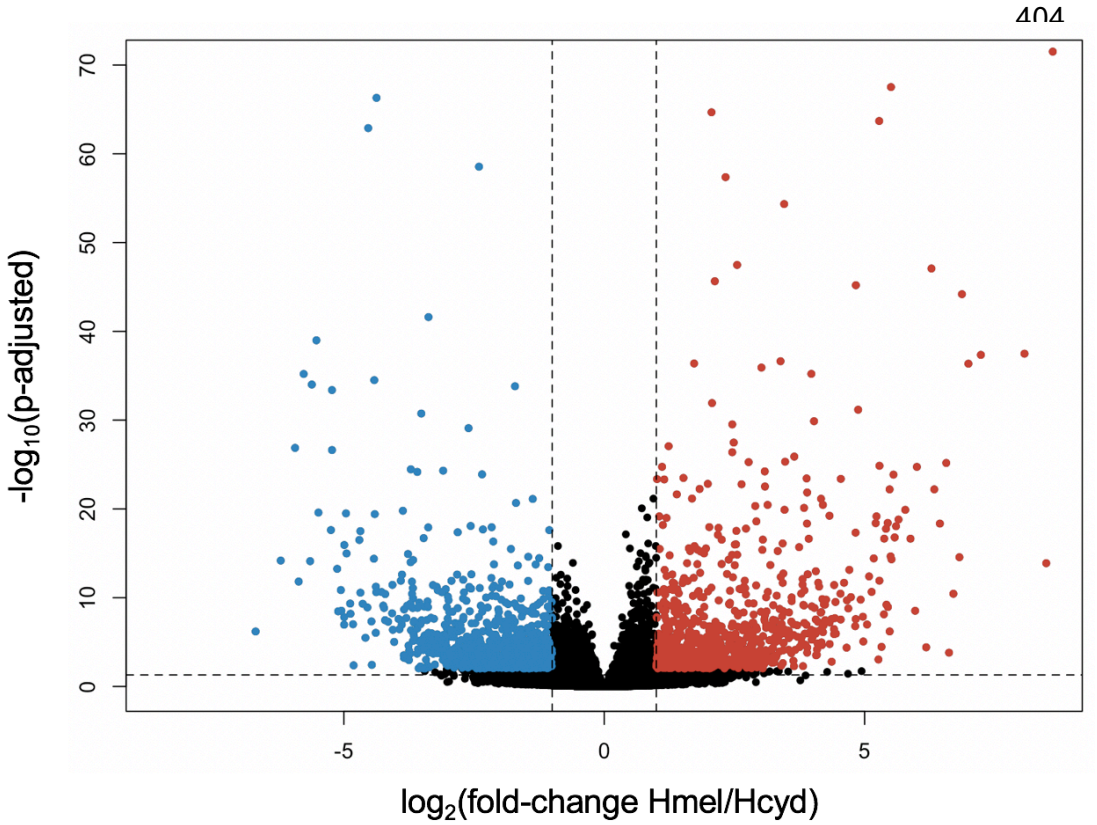

Figure S3. Volcano plot for neural gene expression comparisons between *Heliconius melpomene* and *H. cydno*. Vertical dotted lines indicate the thresholds of a 2-fold change in expression (at x values of -1 and 1), the horizontal dotted line indicates significance (p-adjusted<0.05) in the test for differential expression (as conducted in DESeq2). Differentially expressed genes are colored in blue if up-regulated in *cydno*, and red if up-regulated in *melpomene*. Note that 3 outliers (with very low associated p-values) were removed for clarity.

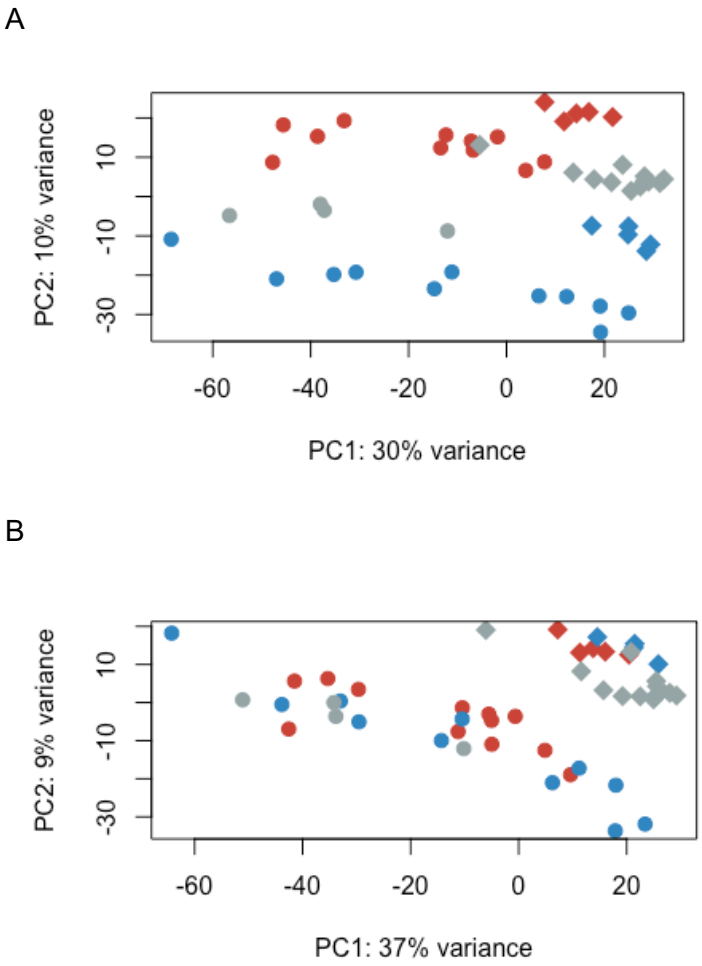

Figure S4. Principal component analyses of expression level profiles of (**A**) all genes and of (**B**) genes that were not detected as differentially expressed (see also, Figure 4). *H. cydno* samples are colored in blue, F1 hybrids in gray, *H. melpomene* in red. Hybrids show reduced intermediary in gene expression level when taking all genes into account, and are not intermediate when non-differentially expressed genes are analysed. However, a trend for dominance of the *melpomene* alleles is also evident across all genes (of all genes: 8.7% are *melpomene-like*, 7.5% *cydno*-like, 3.9% statistically intermediate, 1.3% transgressive, 78.3% show no difference between species), Sequencing batch is denoted by the dot shape: circular (batch 2014) and rhomboid (batch 2019).

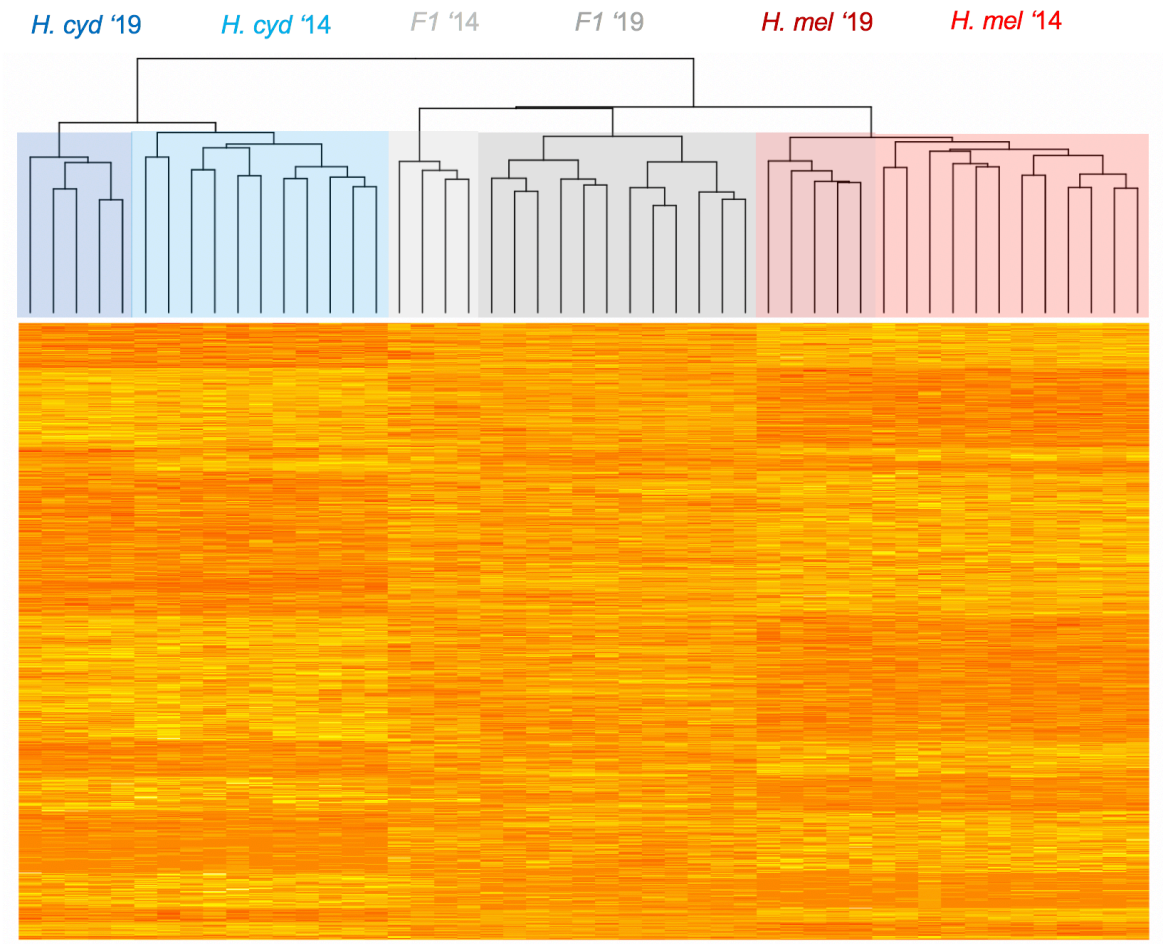

Figure S5. Dendogram of neural expression profiles for *H. cydno*, F1 hybrids and *H. melpomene,* for genes detected to be differentially expressed between *cydno* and *melpomene* (different sequencing batches, 2014 and 2019, are highlighted with different shades of blue *for H. cydno,* gray for F1s, and red for *H. melpomene*).

3 v. Classification of gene expression in F1 hybrids

Our classification of gene expression patterns in hybrids relative to both parental species provides insights into the disruptive nature of hybridisation on expression profiles. Figure S6 provides illustrative examples of the pattern of variance characteristic of each gene category. Among differentially expressed genes, 589 were *melpomene*-like, 344 were *cydno*-like, 12 were 'transgressive', and 701 were intermediate between parental distributions. Considering all genes regardless of their differential expression between species, 1686 were classed as *melpomene*-like, 1449 were *cydno*-like, 259 were 'transgressive', and 746 were intermediate

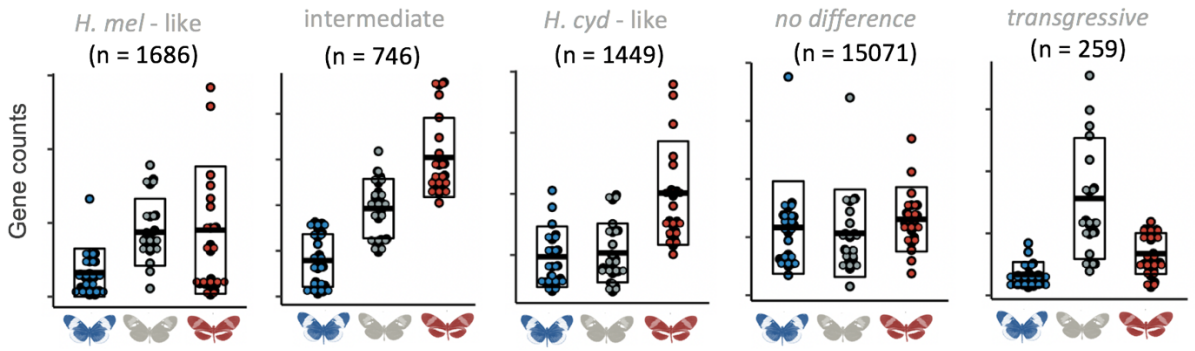

Figure S6. Example of expression profiles for genes assigned to the different categories mentioned (e.g. intermediate in F1s). "n =" denotes the number of genes classified in each category. y-axis indicates the (rlog) normalized gene count. Dots correspond to individual samples, and are colored in blue for *H. cydno* samples, in gray for F1 hybrids, and in red for *H. melpomene.* Horizontal black bars indicate mean, with boxplots delineating + and - one standard deviation, of normalized gene counts.

3 vi. Tests of neutrality on gene expression

PST estimates vary significantly between gene categories (Kruskal-Wallis test with post-hoc Dunn test with Bonferroni correction, p<0.001 in pairwise comparisons except between transgressive and "no differentially expressed" genes), with 11% (n=432/3881) of genes showing intermediate or species-like expression in hybrids having PST estimates within the 95th percentile of the FST distribution, compared to 0.02% (n=3/15330) for non-differentially expressed genes. The proportion of genes with significant PST estimates is highest within genes with intermediate expression in hybrids 23% (n=169/746), followed by *melpomene* and *cydno*-like genes (9% (n=159/1686) and 7% (n=104/1449) respectively), and lowest within transgressive genes (0% (n=0/259)). To test how these results varied under different assumed heritabilities, we recalculated PST under five h^2 values (Figure S5). As expected, the results are consistent across this range, with increased support for selection at lower h^2 values.

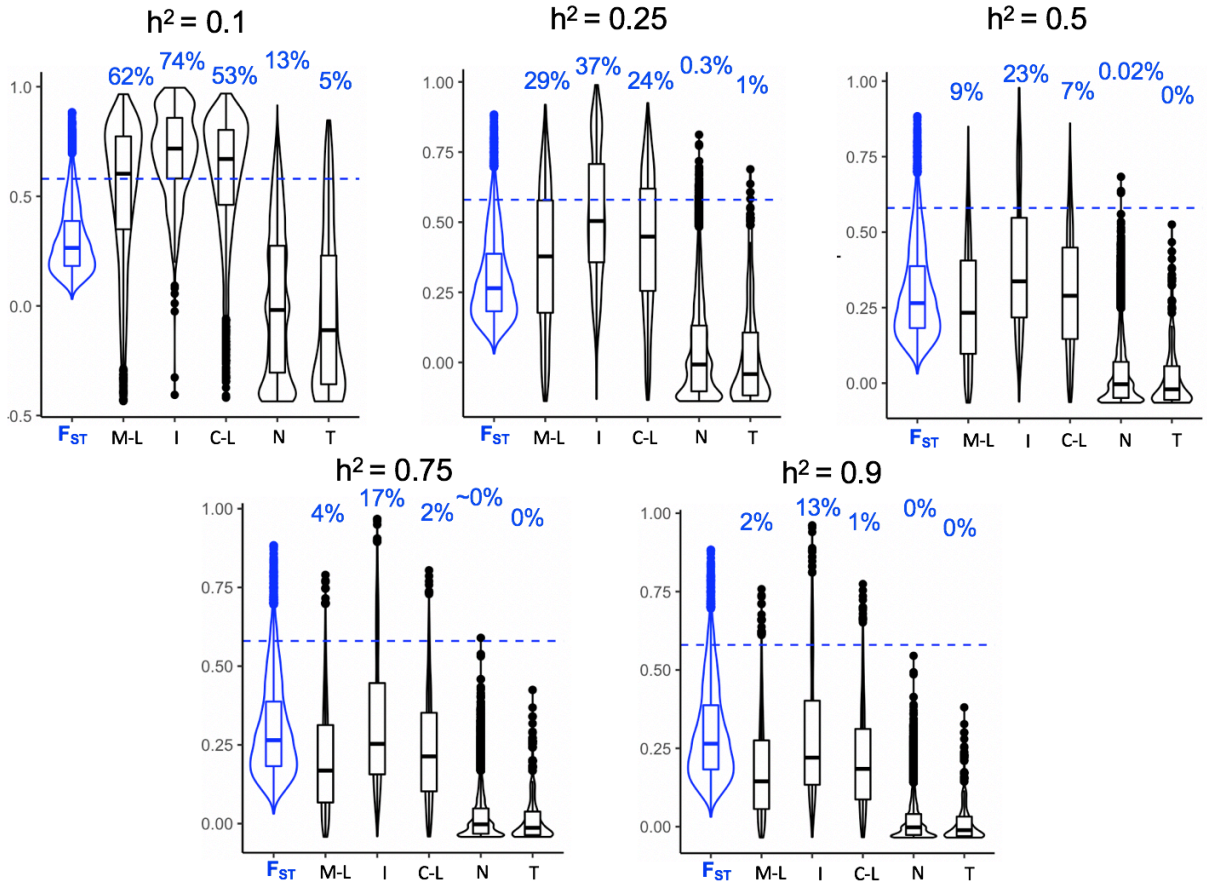

Figure S7. Median, interquartile range and distributions of FST and PST values (for different gene categories), where PST was estimated with varying levels of heritability (h^2), indicated on top of each panel. Percentages (in blue) indicate the percentages of genes with PST value higher than the 95% quantile of FST (indicated by a horizontal dotted blue line), in each category.

3 vii. Variation in P_{ST} and f_d among gene categories

P_{ST}

	H. c-like	intermed.	H. mel-like	no diff.
intermed.	P<0.0001	-	-	-
H. mel-like	P<0.001	P<0.0001	-	-
no diff.	P<0.0001	P<0.0001	P<0.0001	-
transgres.	P<0.0001	P<0.0001	P<0.0001	P=0.094

f_d

	H. c-like	intermed.	H. mel-like	no diff.
intermed.	P<0.001	-	-	-
H. mel-like	P=1.000	P<0.010	-	-
no diff.	P<0.0001	P<0.0001	P<0.0001	-
transgres.	P=0.646	P<0.001	P=0.185	P=1.000

Figure S8. Full pairwise comparisons from a Kruskal-Wallis test with post-hoc Dunn test, with Bonferroni correction (see Figure 5).

4. Supplementary references

1. Moran, D., Softley, R., and Warrant, E.J. (2015). The energetic cost of vision and the evolution of eyeless Mexican cavefish. Sci. Adv. *1*, e1500363.

2. James B. Jaggard, Evan Lloyd, Anders Yuiska, Adam Patch, Yaouen Fily, Johanna E. Kowalko, Lior Appelbaum, Erik R Duboue, and Alex C. Keene (2020). Cavefish brain atlases reveal functional and anatomical convergence across independently evolved populations. Sci. Adv. *6*, eaba3126.

3. Rétaux, S., Alié, A., Blin, M., Devos, L., Elipot, Y., and Hinaux, H. (2016). Neural Development and Evolution in Astyanax mexicanus. In Biology and Evolution of the Mexican Cavefish (Elsevier), pp. 227–244. Available at: https://linkinghub.elsevier.com/retrieve/pii/B9780128021484000128 [Accessed October 3, 2019].

4. Filteau, M., Pavey, S.A., St-Cyr, J., and Bernatchez, L. (2013). Gene Coexpression Networks Reveal Key Drivers of Phenotypic Divergence in Lake Whitefish. Mol. Biol. Evol. *30*, 1384–1396.

5. Renaut, S., and Bernatchez, L. (2011). Transcriptome-wide signature of hybrid breakdown associated with intrinsic reproductive isolation in lake whitefish species pairs (Coregonus spp. Salmonidae). Heredity *106*, 1003–1011.

6. Sylvester, J.B., Rich, C.A., Loh, Y.-H.E., van Staaden, M.J., Fraser, G.J., and Streelman, J.T. (2010). Brain diversity evolves via differences in patterning. Proc. Natl. Acad. Sci. *107*, 9718–9723.

7. Maan, M.E., Seehausen, O., and Groothuis, T.G.G. (2017). Differential Survival between Visual Environments Supports a Role of Divergent Sensory Drive in Cichlid Fish Speciation. Am. Nat. *189*, 78–85.

8. Maan, M.E., Hofker, K.D., Alphen, J.J.M. Van, and Seehausen, O. (2006). Sensory Drive in Cichlid Speciation. *167*.

9. Huber, R., van Staaden, M.J., S, K. Les, and Liem, K.F. (1997). Microhabitat use, trophic patterns, and the evolution of brain structure in African Cichlids. Brain. Behav. Evol. *50*, 167–182.

10. Park, P.J., and Bell, M. a. (2010). Variation of telencephalon morphology of the threespine stickleback (Gasterosteus aculeatus) in relation to inferred ecology. J. Evol. Biol. *23*, 1261–1277.

11. Keagy, J., Braithwaite, V.A., and Boughman, J.W. (2018). Brain differences in ecologically differentiated sticklebacks. Curr. Zool. *64*, 243–250.

12. Kotrschal, A., Deacon, A.E., Magurran, A.E., and Kolm, N. (2017). Predation pressure shapes brain anatomy in the wild. Evol. Ecol. *31*, 619–633.

13. Eifert, C., Farnworth, M., Schulz-Mirbach, T., Riesch, R., Bierbach, D., Klaus, S., Wurster, A., Tobler, M., Streit, B., Indy, J.R., *et al.* (2015). Brain size variation in extremophile fish: local adaptation versus phenotypic plasticity: Brain size variation in extremophile fish. J. Zool. *295*, 143–153.

14. Gonda, a., Herczeg, G., and MerilÄ, J. (2009). Adaptive brain size divergence in nine-spined sticklebacks (Pungitius pungitius)? J. Evol. Biol. *22*, 1721–1726.

15. Mcbride, C.S., Baier, F., Omondi, A.B., Spitzer, S.A., Lutomiah, J., Sang, R., Ignell, R., and Vosshall, L.B. (2014). Evolution of mosquito preference for humans linked to an odorant receptor. Nature *515*, 222–227.

16. Dekker, T., Ibba, I., Siju, K.P., Stensmyr, M.C., and Hansson, B.S. (2006). Olfactory shifts parallel superspecialism for toxic fruit in Drosophila melanogaster sibling, D. sechellia. Curr. Biol. CB *16*, 101–9.

17. Prieto-godino, L.L., Rytz, R., Cruchet, S., Ruta, V., Peraro, M.D., Benton, R., Abuin, L., and Silbering, A.F. (2017). Evolution of Acid-Sensing Olfactory Circuits in Article Evolution of Acid-Sensing Olfactory Circuits in Drosophilids. 661–676.

18. Keesey, I.W., Grabe, V., Knaden, M., and Hansson, B.S. (2020). Divergent sensory investment mirrors potential speciation via niche partitioning across Drosophila. eLife *9*, e57008.

19. Montgomery, S.H., and Merrill, R.M. (2016). Divergence in brain composition during the early stages of ecological specialization in Heliconius butterflies. J. Evol. Biol. *30*, 571–582.

20. Filchak, K.E., Roethele, J.B., and Feder, J.L. (2000). Natural selection and sympatric divergence in the apple maggot Rhagoletis pomonella. *407*, 4.

21. Linn, C., Feder, J.L., Nojima, S., Dambroski, H.R., Berlocher, S.H., and Roelofs, W. (2003). Fruit odor discrimination and sympatric host race formation in Rhagoletis. 10–13.

22. Tait, C., Batra, S., Ramaswamy, S.S., Feder, J.L., and Olsson, S.B. (2016). Sensory specificity and speciation: a potential neuronal pathway for host fruit odour discrimination in *Rhagoletis pomonella*. Proc. R. Soc. B Biol. Sci. *283*, 20162101.

23. Kozak, K.M., Wahlberg, N., Neild, A.F.E., Dasmahapatra, K.K., Mallet, J., and Jiggins, C.D. (2015). Multilocus Species Trees Show the Recent Adaptive Radiation of the Mimetic Heliconius Butterflies. Syst. Biol. *64*, 505–524.

24. Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., and Jiggins, C.D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Res. *23*, 1817–1828.

25. Estrada, C., and Jiggins, C.D. (2002). Patterns of pollen feeding and habitat preference among Heliconius species. Ecol. Entomol. *27*, 448–456.

26. Merot, C., Salazar, C., Merrill, R.M., Jiggins, C.D., and Joron, M. (2017). What shapes the continuum of reproductive isolation? Lessons from Heliconius butterflies. Proc. R. Soc. B Biol. Sci. *284*, 10.

27. Merrill, R.M., Wallbank, R.W.R., Bull, V., Salazar, P.C.A., Mallet, J., Stevens, M., and Jiggins, C.D. (2012). Disruptive ecological selection on a mating cue. Proc. R. Soc. B Biol. Sci. *279*, 4907–4913.

28. Naisbit, R.E., Jiggins, C.D., and Mallet, J. (2001). Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. Proc. R. Soc. Lond. B Biol. Sci. *268*, 1849–1854.

29. Mallet, J., Beltrán, M., Neukirchen, W., and Linares, M. (2007). Natural hybridization in heliconiine butterflies: the species boundary as a continuum. BMC Evol. Biol. *7*, 28.

30. Arias, C.F., Salazar, C., Rosales, C., Kronforst, M.R., Linares, M., Bermingham, E., and McMillan, W.O. (2014). Phylogeography of *Heliconius cydno* and its closest relatives: disentangling their origin and diversification. Mol. Ecol. *23*, 4137–4152.

31. Seymoure, B.M. (2016). Heliconius In A New Light: The Effects of Light Environments on Mimetic Coloration, Behavior, and Visual Systems.

32. Mérot, C., Mavárez, J., Evin, A., Dasmahapatra, K.K., Mallet, J., Lamas, G., and Joron, M. (2013). Genetic differentiation without mimicry shift in a pair of hybridizing *Heliconius* species (Lepidoptera: Nymphalidae): Mimicry and Hybridization in Butterflies. Biol. J. Linn. Soc. *109*, 830–847.

33. Kronforst, M.R., Young, L.G., Blume, L.M., and Gilbert, L.E. (2006). MULTILOCUS ANALYSES OF ADMIXTURE AND INTROGRESSION AMONG HYBRIDIZING HELICONIUS BUTTERFLIES. Evolution *60*, 1254–1268.

34. Kronforst, M.R., Young, L.G., Kapan, D.D., McNeely, C., O'Neill, R.J., and Gilbert, L.E. (2006). Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. Proc. Natl. Acad. Sci. *103*, 6575–6580.

35. James Mallet, W. Owen McMillan, Chris D. Jiggins (1998). Mimicry and warning color at the boundary between races and species. In Endless Forms: Species and Speciation (Oxford: Oxford University Press).

36. Benson, W.W., Brown, K.S., and Gilbert, L.E. (1975). COEVOLUTION OF PLANTS AND HERBIVORES: PASSION FLOWER BUTTERFLIES. Evolution *29*, 659–680.

37. Naisbit, R.E. (2001). Ecological divergence and speciation in Heliconius cydno and H. melpomene.

38. James Mallet (1980). A laboratory study of gregarious roosting in the butterfly Heliconius melpomene.

39. Ott, S.R. (2008). Confocal microscopy in large insect brains: Zinc–formaldehyde fixation improves synapsin immunostaining and preservation of morphology in whole-mounts. J. Neurosci. Methods *172*, 220–230.

40. Montgomery, S.H., Merrill, R.M., and Ott, S.R. (2016). Brain composition in *Heliconius* butterflies, posteclosion growth and experience-dependent neuropil plasticity: Anatomy and plasticity of *heliconius* brains. J. Comp. Neurol. *524*, 1747–1769.

41. Douglas Bates, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green, John Fox Package "lme4." Available at: https://github.com/lme4/lme4/.

42. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. *57*, 289–300.

43. Warton, D.I., Duursma, R.A., Falster, D.S., and Taskinen, S. (2012). smatr 3- an R package for estimation and inference about allometric lines: The smatr 3 - an R package. Methods Ecol. Evol. *3*, 257–259.

44. Montgomery, S.H. (2013). The Human Frontal Lobes: Not Relatively Large but Still Disproportionately Important? A Commentary on Barton and Venditti. Brain. Behav. Evol. *82*, 147–149.

45. The Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature *487*, 94–98.

46. Goslee, S.C., and Urban, D.L. (2007). The **ecodist** Package for Dissimilarity-based Analysis of Ecological Data. J. Stat. Softw. *22*. Available at: http://www.jstatsoft.org/v22/i07/ [Accessed May 30, 2020].

47. Silva, S., Blondeau,Da, and Silva, A., Da (2018). Pstat: An R Package to Assess Population Differentiation in Phenotypic Traits. R J. *10*, 447.

48. Zwarts, L., Vanden Broeck, L., Cappuyns, E., Ayroles, J.F., Magwire, M.M., Vulsteke, V., Clements, J., Mackay, T.F.C., and Callaerts, P. (2015). The genetic basis of natural variation in mushroom body size in Drosophila melanogaster. Nat. Commun. *6*, 10115.

49. R Core Team R: a language and environement for statistical computing. 16.

50. Rossi, M., Hausmann, A.E., Thurman, T.J., Montgomery, S.H., Papa, R., Jiggins, C.D., McMillan, W.O., and Merrill, R.M. (2020). Visual mate preference evolution during butterfly speciation is linked to neural processing genes (Evolutionary Biology) Available at: http://biorxiv.org/lookup/doi/10.1101/2020.03.22.002121 [Accessed May 30, 2020].

51. Davey, J.W., Chouteau, M., Barker, S.L., Maroja, L., Baxter, S.W., Simpson, F., Merrill, R.M., Joron, M., Mallet, J., Dasmahapatra, K.K., *et al.* (2015). Major Improvements to the Heliconius melpomene Genome Assembly Used to Confirm 10 Chromosome Fusion Events in 6 Million Years of Butterfly Evolution. G3 *6*, 695–708.

52. Pinharanda, A., Rousselle, M., Martin, S.H., Hanly, J.J., Davey, J.W., Kumar, S., Galtier, N., and Jiggins, C.D. (2019). Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*. J. Evol. Biol. *32*, 194–204.

53. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

55. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

56. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

57. Montgomery, S.H., and Mank, J.E. (2016). Inferring regulatory change from gene expression: the confounding effects of tissue scaling. Mol. Ecol. *25*, 5114–5128.

58. Uebbing, S., Künstner, A., Mäkinen, H., Backström, N., Bolivar, P., Burri, R., Dutoit, L., Mugal, C.F., Nater, A., Aken, B., *et al.* (2016). Divergence in gene expression within and between two closely related flycatcher species. Mol. Ecol. *25*, 2015–2028.

754 59. Martin, S.H., Davey, J.W., Salazar, C., and Jiggins, C.D. (2019). Recombination rate
755 variation shapes barriers to introgression across butterfly genomes. PLOS Biol. *17*,
756 e2006288.

757 60. Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H.,
758 Maslen, J., Mitchell, A., Nuka, G., *et al.* (2014). InterProScan 5: genome-scale protein
759 function classification. Bioinformatics *30*, 1236–1240.

760 61. Alexa, A., and Rahnenfuhrer, J. Gene set enrichment analysis with topGO. Available at:
761 http://www.mpi-sb.mpg.de/~alexa.

762 62. Kinoshita, M., Shimohigasshi, M., Tominaga, Y., Arikawa, K., and Homberg, U. (2015).
763 Topographically distinct visual and olfactory inputs to the mushroom body in the
764 Swallowtail butterfly, *Papilio xuthus*: Visual inputs in butterfly mushroom body. J. Comp.
765 Neurol. *523*, 162–182.

766 63. Homberg, U., Hofer, S., Pfeiffer, K., and Gebhardt, S. (2003). Organization and neural
767 connections of the anterior optic tubercle in the brain of the locust,Schistocerca gregaria.
768 J. Comp. Neurol. *462*, 415–430.

769 64. Uwe Homberg, and Stefan Würden Movement-sensitive, polarization-sensitive, and
770 light-sensitive neurons of the medulla and accessory medulla of the locust, Schistocerca
771 gregaria. J. Comp. Neurol. *386*, 329–346.

772 65. Borst, A. (2009). Drosophila's View on Insect Vision. Curr. Biol. *19*, R36–R47.

773 66. Pfeiffer, K., Kinoshita, M., and Homberg, U. (2005). Polarization-Sensitive and Light-
774 Sensitive Neurons in Two Parallel Pathways Passing Through the Anterior Optic
775 Tubercle in the Locust Brain. J. Neurophysiol. *94*, 3903–3915.

776 67. Mappes, M., and Homberg, U. (2007). Surgical lesion of the anterior optic tract abolishes
777 polarotaxis in tethered flying locusts, Schistocerca gregaria. J. Comp. Physiol. A *193*,
778 43–50.

779 68. Dell'Aglio, D.D., Troscianko, J., McMillan, W.O., Stevens, M., and Jiggins, C.D. (2018).
780 The appearance of mimetic *Heliconius* butterflies to predators and conspecifics.
781 Evolution *72*, 2156–2166.

782 69. Jiggins, C.D., Naisbit, R.E., Coe, R.L., and Mallet, J. (2001). Reproductive isolation
783 caused by colour pattern mimicry. Nature *411*, 302–305.

784 70. Merrill, R.M., Naisbit, R.E., Mallet, J., and Jiggins, C.D. (2013). Ecological and genetic
785 factors influencing the transition between host-use strategies in sympatric *Heliconius*
786 butterflies. J. Evol. Biol. *26*, 1959–1967.

787