Novel sex-specific genes and diverse interspecific expression in the antennal transcriptomes of ithomiine butterflies

Francesco Cicconardi[1+], Billy J Morris[2], Jacopo Martelossi[3], David Ray[4], Stephen H Montgomery[1+]

[1]School of Biological Sciences, Bristol University, 24 Tyndall Ave, Bristol BS8 1TQ, United Kingdom

[2]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United Kingdom

[3]Department of Biological Geological and Environmental Science, University of Bologna, Via Selmi 3, 40126 Bologna, Italy

[4]Department of Biological Sciences, Texas Tech University, Lubbock TX 79409 USA

[+]Corresponding authors: Francesco Cicconardi, francicco@gmail.com; Stephen H. Montgomery, s.montgomery@bristol.ac.uk

Abstract. The olfactory sense is crucial for organisms, facilitating environmental recognition and inter-individual communication. Ithomiini butterflies exemplify this importance not only because they rely strongly on olfactory cues in both inter- and intra-sexual behaviours, but also because they show re-emergence of macro-glomerular complexes (MGCs). These specialized structures within the antennal lobe, lost in butterflies, but present in moths where they enable the integration of information from various types of pheromones, refining responses to specific cues. In this study we present high-quality genome assemblies for four Ithomiini species, investigating chemosensory evolution and associating expression profiles with neuroanatomical differences. We found that antennal transcriptomes across species exhibit profound divergence, indicating distinct species adaptations in environmental sensing. Noteworthy, sexual dimorphism is also characterised both in chemosensory genes and non-chemosensory genes, suggesting their relevance in behaviour. Lipid-related genes exhibit sexual dimorphism, potentially linked to pheromone production or host selection. The study broadens the understanding of antennal chemosensory adaptations in butterflies, highlighting the intricate interplay between genetic diversity, ecological specialization, and sensory perception with neuro-anatomical differences. Insights into chemosensory gene evolution, expression patterns, and potential functional implications enhance our knowledge of sensory adaptations in butterflies, laying the foundation for future investigations into the genetic drivers of behaviour, adaptation, and speciation in insects.

1 | Introduction

Insects constitute one of the planet's most successful and diverse eukaryotic classes, accounting for roughly fifty percent of all land-dwelling species (Mora et al. 2011), across an astonishingly broad spectrum of environments. Research into the anatomical, physiological, and behavioural facets of this diversity has demonstrated the evolutionary malleability and importance of insect sensory systems, including chemoreception (Missbach et al. 2014). A significant emphasis has been placed on the identification and functional characterization of olfactory receptors (Yan et al. 2020), as well as the neural circuits in which they are expressed and the odour driven

behaviours driven that they govern (Amin and Lin 2019). While olfactory reception and circuit evolution has been linked to some cases of ecological speciation (Olsson et al. 2006; Prieto-Godino et al. 2017; Auer et al. 2020; Auer et al. 2022), the ecological selection pressures shaping olfactory evolution often remains challenging to pinpoint, particularly given the diversity of receptor types and associated gene families.

There are two primary categories of insect olfactory receptors: odorant receptors (ORs) (Vosshall et al. 1999) and ionotropic receptors (IRs) (Benton et al. 2009). OR genes encode for 7-transmembrane (7-TM) proteins, which create a homodimer of a heterodimer odour-gated ion channel through a combination of ligand-specific ("tuning") receptor subunits (ORx) and a co-receptor, Orco (Mika and Benton 2021). OR proteins can identify odorants in lymph fluid, transforming chemical cues into neuroelectric signals, and transmitting them to the central nervous system, thereby influencing insect behaviour (Fleischer et al. 2018). IRs represent a markedly diverse subset of ionotropic glutamate receptors (iGluRs) (Benton et al. 2009). Predominantly, iGluRs bind the excitatory neurotransmitter glutamate and are instrumental in synaptic communication within the brain (Yelshanskaya et al. 2014). In contrast, IRs had been hypothesized to be the most ancient arthropod chemoreceptors, dating back to the Protostomia (Eyun et al. 2017). They have a primary and extensive presence in peripheral sensory systems, serving various functions including chemosensation, thermosensation, hygrosensation (Giesen and Garrity 2022), and potentially non-olfactory functions, such as mechanosensation (Bechstedt et al. 2012). Within IRs, "antennal" IRs are conserved throughout insects, often with numerous introns, and function in olfaction, thermosensation, and hygrosensation; while "divergent IRs" are expressed in peripheral and internal gustatory neurons, and contribute to taste and food assessment (Croset et al. 2010). Structurally, they are similar to iGluRs and a functional ion channel is formed by three-pass transmembrane of a homodimer of a heterodimer of IR subunits. In most cases, IR receptors consist of specific tuning receptors for different stimuli, alongside one or two broadly expressed co-receptors (Abuin et al. 2011). The canonical view is that the majority of olfactory sensory neurons (OSNs) generally express a pair of distinct ligand-selective ORs or IRs, in a one-receptor-to-one-neuron organization: a distinctive 'tuning' receptor designed to detect specific ligands or odorants. This has been recently challenged by data showing that how some neurons co-express multiple chemosensory receptors in *Aedes aegypti* (Herre et al. 2022; Task et al. 2022). Ligand-selective receptors are always associated with specific co-receptors (Orco for ORs, and either IR8a or IR25a for IRs), which do not recognise compounds but instead are needed to form heteromeric complexes with tuning ligand binding receptors (Schmidt and Benton 2020). Antennal lymph fluid also contains abundant secreted proteins and proteoglycans (Schmidt and Benton 2020), which influence the intrinsic physicochemical traits of the odours. Among these lymphatic proteins are the odorant-binding proteins (OBPs), which encode for small globular and soluble proteins. OBPs allow hydrophobic airborne odorants to dissolve into the lymph fluid, and bind odorant compounds with different degrees of affinity and specificity, shuttling them to the underlying receptors in the form of monomers and/or homodimers (Leal 2013; Larter et al. 2016).

The diversity and variable evolutionary rates of olfactory receptors suggest an intimate link to ecological variation and species selection regimes. Lepidopterans have frequently been utilized as models among various

insect species to explore the influence of ecological variability on the evolution of olfactory systems. Within them, Ithomiini butterflies offer an interesting system due to their chemical defences, reliance on chemical communication, and interspecific interactions. They are one of the most specious tribes of Neotropical butterflies with 393 species, and dominate butterfly communities in Neotropical forests (Beccaloni 1997a; Beccaloni 1997b). Their chemical defences are primarily derived from pyrrolizidine alkaloids (PA) from specific host plants, either obtained as larvae or through adult foraging and male-to-female provision through the spermatophore (Brown Jr 1984; Masters 1992; Massuda and Trigo 2009). Males generally have a stronger attraction towards these plants compared to females, but in some species, they also function as female attractants at short ranges and male repellents at long-range, especially in species where males establish territorial dominance defending resource patches (Pliske 1975a; Pliske 1975b). The complexity of these chemically driven behaviours is an indicator of a strong sexual dimorphism within adults in relation to PA sources, both chemical defences and pheromones, suggesting that this ecological context might have led to specific olfactory adaptations.

These adaptations likely lie in the antennae, antennal receptor cells and downstream olfactory processing areas. In general, however, diurnal butterflies lack the striking specialisations observed in moths. The antennal lobe of all lepidoptera is generally formed by ~60-70 morphological units, called glomeruli, which are each composed of axon terminals from antennal sensory neurons expressing the same olfactory receptor (Rospars 1983; Hansson and Stensmyr 2011; Carlsson et al. 2013). In many moths, the antennal lobe of males is characterized by macro-glomerular complexes (MGCs), specialized structures within the antennal lobe composed of interconnected and enlarged glomeruli that often respond to pheromones (Koontz and Schneider 1987; Sung et al. 2017; Williams et al. 2022). MGCs enable the integration of information from various types of pheromones and their associated odorants, facilitating precise and refined responses to specific cues. In butterflies, likely due to their increased reliance on visual cues, these structures are widely absent and were likely lost at the origin of the superfamily (Morris et al 2021). However, within the Ithomiini group, analogous structures to MGCs have re-emerged through convergent evolution (Montgomery and Ott 2015; Morris et al. 2021). In ithomiines the composition and size of these structures is highly variable across species, and in many cases exhibit a degree of sexual dimorphism, with certain MGC glomeruli being larger in males than in females, although in some species the enlarged glomeruli are shared between sexes (Morris et al. 2021). In contrast, in one genus, *Methona*, antennal glomeruli are of uniform size and lack enlarged or dimorphic glomeruli (Morris et al. 2021). This genus has a distinct mating strategy, which relies less on olfactory signalling than other ithomiini genera (Brown Jr 1984; Brown 1987; McClure et al. 2019). This strongly suggests a unique, derived and heightened reliance on olfactory reception in this tribe of butterflies, but with species-specific divergences in the presence/absence of traits, offering a case study to understand the evolution of new neurosensory traits at a molecular and ecological level.

Here, we aim to explore the genetic basis of these sexual dimorphism/adaptations, by sequencing the genomes of four Ithomiini species - *Mechanitis polymnia, Tithorea harmonia, Methona confusa* and *Greta morgana*

- representing deep divisions within the ithomiine phylogeny, and variable ecologies and neuromorphologies. With these resources we characterise and manually curate the antennal chemosensory receptors and odorant binding proteins, alongside data from 11 species from closely and more distantly related nymphalid butterflies. We use this data to test for ithomiine specific signatures of gene family evolution and selection, and divergent patterns of receptor expression using antennal transcriptomes between sexes and species in three Ithomiini butterflies. We find that in ithomiini butterflies: *i*) olfactory innovations did not involve massive antennal chemosensory gene (ACG) expansion events, or OR novelty; *ii*) *neuro-dimorphic species* have over- and sex-biased ACG expression, mirroring anatomical differences among sex and species, with the number of OR expression outliers directly reflecting the number of MGC glomeruli in each species; and *iii*) there is a strong transcriptomic diversity among species, possibly reflecting the different ecological niches in which these species are adapted.

2 | Materials and Methods

2.1 | DNA and RNA Extraction and Sequencing

Samples for *Mc. polymnia, T. harmonia, Me. confusa* and *G. morgane* were obtained from commercial pupae supplies (Stratford Butterfly Farm, UK; London Pupae Supplies, UK). These species represent three of the ten recognised subtribes within Ithomiini: Godyridina, Mechanitina, Tithoreina, (Brower et al. 2014), and variable microhabitat preferences (Elias et al. 2008; Hill 2010) and antennal lobe morphologies (Morris et al., 2021). High-quality, high-molecular-weight genomic DNA was extracted as in Cicconardi et al. (2023); 100 mg of tissue were dissected, snap frozen in liquid nitrogen and homogenized in 9.2 ml buffer G2 (Qiagen Midi Prep Kit) adding 19 µl of RNAseA, adding 0.2 µl of Protease K and incubated at ~50 °C for 2 hours. Samples were processed with a Qiagen Midi Prep Kit (Qiagen, Valencia, CA) following the manufacturer's instructions, and precipitated using 2 ml 70% EtOH and dissolved in water. To generate whole genome sequencing data the 10x Chromium Library Prep was adopted alongside Illumina sequencing using 150bp paired-end reads with NovaSeq FC S2, generating ~40Gbp per species, performed at the Institute of Applied Genomics (IGA), Udine, Italy.

For RNA extractions, pupae were allowed to eclose at 26°C and 80% humidity under a 12:12 day:night regime. Butterflies were then aged in these conditions for 4-6 days in 1.5m x 1.5m x 2m cages and fed on a 30% sugar solution. The cages included plants, *Cestrum nocturn* and *Solanum crispum*, from the Solanacea as natural stimuli, and cuttings of *Heliotropium* (a PA source for adults). Surviving butterflies were then flash frozen in liquid nitrogen and stored at -80°C. To capture sex-specific expression of antennal olfactory receptors, pairs of antennae were homogenised with RLT buffer by repeated aspiration with a 21-gauge needle, and RNA was extracted using a Qiagen RNeasy kit, following the manufacturer's instructions, including treatment with Qiagen RNase-free DNase to remove any remaining DNA. To achieve sufficient concentrations, three individuals were pooled in each sample according to their estimated RNA concentrations, with three-replicates per sex per species for *Mc. polymnia, T. harmonia, Me. confusa*. Sufficient samples of *G. morgane* were not available at the time of sampling. Polyadenylated Illumina RNAseq data (125 bp x 2) was carried out by University of Liverpool Centre for Genomic, for a total of nine samples.

2.2 | 10X Genomics linked-read Genome Assembly and Repeat Annotation

Sequenced Illumina paired-end reads from 10X Genomics libraries were input to the SUPERNOVA v2.1.1 assembler (10x Genomics, San Francisco, CA, USA) (Zheng et al. 2016) for *de novo* genome assembly. No trimming was needed as per the assembler documentation. The assembly pipeline follows Cicconardi et al. (2023). In brief the optimal amount of reads was adopted to maximise contiguity, duplication level and completeness, based on BUSCO (Benchmarking Universal Single-Copy Orthologs; v3.1.0, Insecta_odb9) statistics (Simão et al. 2015). Subsequently, assemblies were processed with PURGE HAPLOTIGS to remove haplocontigs and TIGMINT v1.1.2 (Jackman et al. 2018) was used to correct potential assembly errors. RNA-seq data was then used for scaffolding as implemented in P_RNA_SCAFFOLDER, followed by ARCS v1.1.0 (Yeo et al. 2018).

Transposable elements (TEs) *de novo* annotation can be highly inaccurate when analysing genomes from non-model species. This is primarily due to the frequent partial status of raw consensus sequences and the high number of unclassified repeats (Platt et al. 2016; Goubert et al. 2022; Martelossi et al. 2023; Sproul et al. 2023). To address this issue, we followed a standardized pipeline described Osmanski et al. (2023). Briefly, in the first step the pipeline employs REPEATMODELER v2.0.4 (Flynn et al. 2020) with REPEATSCOUT v1.0.6 (Price et al. 2005) to discover TEs and generate the initial repeat library. Next, we extended consensus sequences using a combination of the Ray lab's EXTRACT_ALIGN and Robert Hubley's DAVIDEXTENDEDCONSRAM.PL scripts, available at https://github.com/davidaray/bioinfo_tools. The extended consensus was further classified using REPEATCLASSIFIER from the REPEATMODELER package. TE-related proteins and structural features were collected with the TECURATE.SH script (https://github.com/davidaray/bioinfo_tools/blob/master/TEcurate.sh) implemented with DIAMOND v2.1.5.159 (Buchfink et al. 2021) on the REPEATPEPS.LIB libraries from REPEATMASKER repository, followed by TE+AID v.0-DEV (Goubert et al. 2022). The resulting libraries, one for each species, were then manually screened to link TE orders (*i.e.*: DNA, Rolling Circle, SINE, LINE, LTR) to unclassified TEs based on characteristic structural features of each order, following a similar approach to Ray et al. (2019). Finally, we adopted CD-HIT v4.8.1(Fu et al. 2012), with a sequence identity threshold of 0.80 with available TE libraries for Lepidoptera, to further extend the annotation to remained unknown TEs. All libraries were concatenated without removing redundancies, and REPEATMASKER v4.1.4 (Smit et al. 2013) was utilized to re-identify repetitive elements in all the genomes considered in this study.

2.3 | Bacterial Contamination & Assembly completeness assessment

After the genome assembly contaminants were removed using BLOBTOOLS v1.1.1 (Laetsch and Blaxter 2017) using BLASTN [-evalue 1e-25 -max_target_seqs 1] and the NCBI nucleotide collection (#seqs: 49,266,009, retrieved September 2018). Mitochondrial sequences were identified by blasting (BLASTN) and removed from the main assembly. A combination of BUSCO v3.1.0 (Benchmarking Universal Single-Copy Orthologs) (Simão et al. 2015) and the Lepidoptera set in ORTHODB v.10 (odb10) was implemented using default parameters [-m genome], and EXONERATE v2.46.2 (Slater and Birney 2005), to assess genome completeness and duplicated content.

2.4 | Species Phylogeny and Whole Genome Alignment

The complete single-copy orthologous genes (scOGs) identified with BUSCO were used to generate the species phylogeny using three Danainae species (*Danaus plexippus, D. melanippus* and *D. chrysippus*), five Heliconiinae (*Heliconius melpomene*, *H. erato demophoon*, *Eueides isabella*, *Dryas iulia*, *Speyeria mormoria*), the Nymphalinae *Melitaea cinxia,* and two recently available genomes of Ithomiinae (*Melinaea menophilus* and *Ml. marsaeus rileyi*) (Gauthier et al. 2023). From each locus the nucleotide sequence was following settings in Couto et al. (2023) and Cicconardi et al. (2017; 2020; 2022), described in more details below. Final alignments were concatenated to maximum likelihood (ML) phylogenetic tree estimated using IQ-TREE2 v2.1.3 COVID-edition (Minh et al. 2020), partitioning the supermatrix for each locus and codon position. IQ-TREE2 was run with the following settings: --runs 5 -m MFP with 5,000 ultrafast bootstrap replicates. As a complement to the ML tree, gene trees from scOGs were using IQ-TREE2 and used to generate a coalescent summary method species tree, as implemented in ASTRAL-III v5.6.3 (Zhang et al. 2018), in order to detect discordant topological signals due to incomplete lineage sorting (ILS).

The Bayesian algorithm of MCMCTree (Yang 2007) was performed adopting the approximate likelihood computation to estimate divergence times, estimating first branch lengths by ML, and then the gradient and Hessian matrix around these calculated in MCMCTree using the DNA supermatrix. Calibration nodes were constrained according to Cicconardi et al. (2023) using a uniform distribution. The analysis was run ten times each with 100k generations after a 10M generations as burn-in, logging every 200 generations. Convergence was checked using Tracer v 1.7.1 (Rambaut and Drummond 2007), verifying values form ESS higher than 200.

For the whole genome alignment, the ML phylogeny was used to guide the whole genome alignment using all the previously listed 11 Nymphalid soft-masked genomes, plus the new four new ithomiini genomes produced by this study. CACTUS (Paten et al. 2011; Armstrong et al. 2019; Armstrong et al. 2020) was run using genomes at chromosome level set as the reference.

2.4 | Genome and Chemosensory Gene Annotations.

Raw RNA-seq read data from each library were filtered using TRIMMOMATIC v0.39 (Bolger et al. 2014) (ILLUMINACLIP:$ILLUMINACLIP: 2:30:10; SLIDINGWINDOW: 5:10; MINLEN: 100), and pooled prior to performing the genome annotation. Multiple approaches were adopted (prediction of coding genes, *ab initio* and *de novo*) as implemented in the pipeline described in Cicconardi et al. (2023), which maximises the return from each approach to overcome their own limitations. Briefly, quality filtered reads were mapped using STAR v2.7.10A (Dobin et al. 2013), and the resulting BAM file used as training data for the BRAKER v2.1.5 pipeline (Brůna et al. 2021), which implements GENEMARK-ES SUITE v4.30 (Lomsadze et al. 2005) and AUGUSTUS v3.4.0 (Stanke et al. 2006). For the *de novo* transcriptome assemblies, TRINITY v2.10.0 (Iyer and Chinnaiyan 2011; Haas et al. 2013) was adopted to generate contigs that were subsequently aligned to the genome using MINIMAP2. Coordinates for the aligned contigs were used to extract nt sequences, and TRANSDECODER v5.5.0 (http://transdecoder.github.io/)

(minimum amino acid length > 50) was implemented to annotate coding regions, using homologs from the UniProt database (Bateman 2019) and Lepidoptera proteome (see below) found with deltaBLAST v.2.7.1+ (Boratyn et al. 2012); and PFam v33.1 domains (El-Gebali et al. 2019) with HMMscan v3.3.2 (Eddy 1998) (e < 1e−10). To generate the *ab initio* transcriptomes, BAM files were used as input for both StringTie v2.1.3b (Pertea et al. 2015) and Cufflinks v2.2.1 (Trapnell et al. 2010; Garber et al. 2011; Trapnell et al. 2012).

The different annotations (predicted, *de novo*, and two *ab initio*) were combined and use with STAR to re-map pooled reads, and Portcullis v1.1.2 (Mapleson et al. 2018) [--threshold 0.5] used to remove false positive spliced sites and to generate a splice-site database. Finally, all these elements (transcript and spice-site annotations) were combined together using Mikado v2.3.3 (Venturini et al. 2018) [--scoring insects.yaml -bt UniprotDB+Lepidoptera –mode split]. To annotate the genome of *Greta morgane,* only the prediction pipeline was adopted as RNA-seq data was not available, and only mRNAs and proteins obtained from the other ithomiini were used as training set for the BRAKER pipeline.

2.5 | Selection on Syntenic Single-Copy Ortholog Groups

The Comparative Annotation Toolkit (CAT) (Fiddes et al. 2018) was finally used leveraging the whole-genome alignment to produce an annotation set on every genome in that alignment. This extended the annotation of genomes by detecting loci not previously annotated, and by projecting the annotation from a reference it was used to identify syntenic orthologs; orthologs that can be identified using syntenic information; similar approached used by Jebbs et al. (2020). To this end, using a pipeline that parse CAT output (CATgff32OrthologyTable.py) we generated a list of single copy syntenic orthologs.

2.6 | Phylogenies of Chemosensory Gene Families and Orthology Assignment.

We inferred the evolutionary relationships of the odorant receptors (ORs), ionotropic receptors (IRs), and odorant binding proteins (OBPs) annotated from the 14 Nymphalid species using aa sequences. To do that we implemented a combination of manual and automatic procedures. We first collected protein sequences from curated datasets available in the literature (Vogt et al. 2015; Bastin-Héline et al. 2019; Yin et al. 2021), which were mapped onto the 14 genomes using Exonerate. From all Exonerate alignments CDS were extracted and conserved domains were identified, with CD-search (Marchler-Bauer et al. 2015) and PFAM v31 (El-Gebali et al. 2019) using HMMscan (Eddy 2011), combined with the TOPcons web-server (Tsirigos et al. 2015) to identify the presence of the peptide signal, and to predict the number of transmembrane helices (TMHs). At each locus, the best annotation was therefore automatically selected based on the optimal protein length, conserved domain length and score, presence of P-signal and best number of TMHs. Finally, we used the latest version of WebApollo, run using Docker, to manually check the annotations to validate the procedure and correct possible mistakes. Loci with no valid conserved domain hit were then excluded from subsequent analyses (e-value < $1x10^{-5}$). For each chemosensory gene family (CGF), amino-acid sequences were aligned using CLUSTALW v2.1 (settings: dnamatrix=IUB; gapopen=10; gapext=0.1; gapdist=10; iteration=TREE; numiter=1000; clustering=NJ),

and the phylogeny was inferred using maximum likelihood (ML) search as implemented in FastTree v2.1.11 SSE3, using Le-Gascuel 2008 model with pseudocounts and the slow exhaustive search algorithm to search for neighbour-joining. For the OR phylogeny, Orco was used as an outgroup, IR phylogeny we used the IGluR and for OBPs mid-point root. Gene orthology was subsequently assigned based on the phylogenetic tree or reference genes.

2.7 | RNA-seq Data Analyses.

Quality filtered reads were mapped to the corresponding reference genomes using STAR v2.7.10a [parameters: outSAMattributes NH HI AS NM MD; outFilterMultimapNmax 20; outFilterMismatchNmax 999; outFilterMismatchNoverLmax 0.04; alignIntronMin 20; alignIntronMax 1000000; alignMatesGapMax 500000, alignSJoverhangMin 8; alignSJDBoverhangMin 1; sjdbScore 1]. Expression abundance of each gene/isoform was calculated using RSEM (Li and Dewey 2011) and used as input for intraspecific and interspecific differential expression analysis using EBSeq (Leng et al. 2013) [rsem-run-ebseq], correcting for multiple tests [rsem-control-fdr] with a threshold for posterior probability of 0.95. We also checked for possible bias generating MA plots, PCA and dispersion estimates using normalized counts in DESeq2 implemented in R-project module (Love et al. 2014). For interspecies comparisons TPMs (transcripts per million) were adopted.

Enrichment of GOterms was performed using a combination of two different approaches, the hypergTest algorithm, implemented in the GOStats package (Falcon and Gentleman 2007) for R [annotation org.Dm.eg.db; conditional TRUE; testdirection over), and GOATOOLS (Klopfenstein et al. 2018) (*P* value cutoff 0.05); both using scOGs with a putative sign of diversifying positive selection (adjusted P value < 0.05), with all scOGs used as the background list. To reduce the false positive rate *conditional(p)* == TRUE (GOStats) was selected, a conditional algorithm that uses the structure of the GO graph to reduce subsequent tests (Alexa et al. 2006), only considering terms in common between GOStats and GOATOOLS (Klopfenstein et al. 2018) results.

2.8 | Selection Dynamics of Chemosensory Genes

For each OG of the three CGFs the nucleotide sequences were aligned, with a filtering procedure as implemented in Cicconardi et al. (2023). Briefly, nt sequences were quality filtered before the alignment with PREQUAL v1.02 (Whelan et al. 2018) [-pptype all] and after the alignment, performed with MACSE v2.03 (Ranwez et al. 2011), with HmmCleaner (Di Franco et al. 2019) and GBLOCKS v0.91b (Castresana 2000) under a "relaxed" condition. A ML gene tree was then generated as implemented in IQ-Tree2 v2.1.3 COVID-edition (Minh et al. 2020) [sampling GENESITE; m MFP]. To gain insights into the evolutionary history and the selective pressures on CGFs, we scan for shifts in selective regimes. To do that we used RELAX (Wertheim et al. 2014) to estimate the selection coefficients (*k*) of orthologous genes for all the chemosensory genes (ORs, IRs and OBPs) in the six ithomiini species. In brief, RELAX tests the hypothesis of evolutionary rate relaxation in selected branches of a phylogenetic tree compared with reference branches. A *k* value is computed to evaluate whether selective strength was relaxed ($k < 1$) or intensified ($k > 1$). We performed the test as implemented in the HYPHY

framework (Kosakovsky et al. 2020) to identify genes under intensified selection, and test whether different species experience intensification/relaxation for the same genes. To do so, each terminal branch leading to any of the six ithomiini species were tested using all the other internal and internal branches of the OG as the background. All p-values associated with k were subsequently adjusted for multiple comparisons using Bonferroni correction.

3 | Results

3.1 | New genomic resources for ithomiini butterflies

A total of 180 Gb of linked-reads on average per sample were generated, which resulted in an extremely high coverage per sample. The optimal coverage per sample was identified by resampling our sequencing data by 50M to 300M reads. This optimised contiguity and completeness for each sample (Table S1) resulting in a final assembly size close to the estimated genome size, computed by SUPERNOVA; with contiguity on average in the range of N50 ~ 4.6 Mb per sample (min: 1.2Mb for *G. morgane*, max: 9.4Mb for *T. harmonia*; Figure S1; Table S1); and very high completeness: ~97% of single-copy BUSCO genes, 1% missing, with 1% of duplicated genes (Table S1). In terms of gene content, with the exception of Danaini species, which harbour smaller number of genes (~11.3k genes on average), the rest of the Nymphalids have similar amounts of protein coding genes (~20k on average), with the exception of *G. morgane*, which included ~33k loci. We don't exclude possible gene expansions, but it this assembly has a lower contiguity (fragmented BUSCO: 1.6%) and possibly higher gene fission rate due to the annotation strategy lacking RNA-seq data (Figure S2).

As anticipated, genome size was influenced by the content of repeats (Figure S3), with retroelements (SINE + LTR + LINE) and DNA transposons particularly affecting genome size (randomForest Figure S4). Danainae are generally characterized by a lower proportion of repeats (16% on average) compared with the other species in which it resulted always greater than 21%. On average, compared with Heliconiini, Ithomiini have slightly lower repeat content, 26% vs 33%, although, correlating with their larger genome size (~500 Mb). Ithomiini have ~8Mb (~9% of the total TE content) of rolling circles (*Helitrons*), similar to other butterflies, but ~8 times less compared with Heliconiini (on average ~62Mb; ~48% of the total TE content), which seem to be particularly expanded in the latter. Within Ithomiini, *Melinaea spp.* show a particularly rich rolling circles, with ~15Mb, 6 times richer compared with other species, which seem to be expanded recently, as shown by the kimura distances (Table S1, Figure S5). These elements possibly contribute to the formation of promoters, exons, splice sites, polyadenylation sites, and microRNA-binding sites to conserved transcripts, impacting innovation (Thomas et al. 2014).

Across the 15 species selected, we identified 5,077 single copy BUSCO genes, which, once concatenated, resulted in an alignment of 4.1 Mb of which 1.7 Mb parsimony informative. The alignment was then used to build a species tree under maximum-likelihood framework and a gene tree reconciliation, to reconstruct a first approximation of the species tree. The two approaches returned identical topologies (Figure 1a, b). However, while the Ithomiini outgroup shows high coalescent units (CUs), the branching within Ithomiini

is characterized by very low CUs, possibly indicating a rapid diversification with a high amount of incomplete lineage sorting (ILS). Based on the phylogenetic tree and different calibration points (see Methods), we further calculated the divergence time and substitution rate. The analysis indicated that Danaiinae diverged between 41.7 and 70.5 million years ago (Ma) (95% CI; median 54.6 Ma), while the Ithomiini during the Oligocene, between 24.1 and 42.1 Ma (95% CI; median 32.3 Ma), both overlapping with the latest estimations for these groups (Kawahara et al. 2023). The topology of the phylogenetic tree shows some differences previously reported phylogeny of Ithomiini (Chazot et al. 2019), based on nine nuclear gene fragments, a mitochondrial fragment. Both studies overlap on the temporal framework of Ithomiini stem (CI: 24.1-42.1 Mya in this study; 22.75-30.99 Mya in Chazot et al. 2019), and in placing *Melinaea spp*. as the first diverging branch from the stem of the tribe. The difference is in the placing of the genus *Mechanitis*, which we recover as branching later in time.

3.2 | Diversity and Independent Duplications in Chemosensory Genes Across Nymphalids

To remove any bias caused by different annotation procedures we manually curated all chemosensory receptors in all 15 species of our dataset, which resulted in 65 ORs, 23 IRs, 16 IGluRs, and 35 OBPs on average per species, and a total of 2113 annotated genes (Figure 1-4; Table S2-S4). Not all loci show complete functional domains. Among all chemosensory gene families ORs show the largest turnover rate and diversification. Overall, looking at loci with complete domains, Ithomiini have on average 67 ORs per species, the same as for Heliconiinae, and in line with their estimated olfactory glomeruli number of related ithomiines (Montgomery & Ott, 2015; Morris et al 2021). The complete phylogenetic tree of ORs resulted in 49 orthologous groups (OGs). One of these OG, the OR4 shows a major expansion within Nymphalids, with a total of 107 genes across the 15 species. Of note, we found an Ithomiini-specific loss of one of the known pheromone receptor clades. In fact, while *Danaus* species have genes within OR2 and OR13; OR5 seems to be present only in Heliconiinae. This loss could be balanced by independent expansions within Ithomiini of the other "Novel" pheromone receptor clade ORs (OR30 and OR83). Specifically, for OR38, in-paralogs are present in the Heliconiini, *Dryas iulia* (3 copies), in *D. plexippus* (2 copies), and in the Ithomiini: *Ml. menophilus* (2 copies), *G. morgane* (4 copies), *Mt. confusa* (4 copies) and *T. harmonia* (6 copies). In contrast, for OR30, there is an expansion in the Heliconiini *Heliconius* species (2 copies) and an extra copy in *H. erato* (3 copies, in total), *D. plexippus* (3 copies), and in the Ithomiini *Mc. polymnia* (2 copies) and in the *Melinaea* species (2 copies). Independent and repeated duplications or OR51 are present in all Nymphalids. The OG duplicated three times at the base of Heliconiinae, followed other species-specific duplication events, including an independent duplication at the stem of Danainae. Similarly, OR46 was duplicated at the stem of Danainae, with a further duplication in *G. morgane* (3 copies located on the same scaffold). OR56, which seems to be associated to plant volatile compounds (Bastin-Héline et al. 2019), seems to be not only specific to Heliconiinae but expanded at least three times. Finally, here are two major expansions specific to Ithomiini, one within OR53, which has generally two copies in nymphalids but is further expanded in *Melinaea spp.* (3/4 copies), *Me. confusa* (4 copies), and *G. morgane* (3 copies), and one within the, which shows a duplication in Danainae, with several other independent expansions within Ithomiini, between four and six copies ((Figure 2; Table S5).

Compared with ORs, IGluRs and IRs show a much more conserved pattern. In total we identified 41 OGs, with 15 IGluR OGs and 26 IRs. Within IGluRs, we ascribe two ionotropic coreceptors (CoIR), CoIR8 and CoIR25, as they cluster within all the IGluRs. Within Ithomiini, almost all IGluRs are present in single copy, although for four receptors, six more short fragments were also identified in *G. morgane* and *M. confusa*. Among the other IRs the Lepidoptera-specific (LS-IRs) and divergent IRs, show the highest turnover. IR1b seems to present only in Ithomiini and was lost in the rest of Nymphalids, and duplicated in *Melinaea ssp*.. Within the divergent IRs, IR7d4 seems to be lost in all ithomiines, while IR7d2, IR7d4, and IR143 underwent duplications in different ithomiine species (Figure 3; Table S5).

Odorant binding proteins show a strong pattern of conservation, with very few exceptions within the antennal binging proteins (ABPs). These exceptions include, ABP8, which shows multiple duplications within *G. morgane*, with four copies, while, ABP1 shows numerous independent duplications in Heliconiini, *Melinea spp*, and Ithomiini. Specifically, a total of three copies for *Me. confusa* and *Mc. polymnia*, and five and six for *T. harmonia* and *G. morgane*, respectively. Notably, the OG of ABP6 shows huge expansion in all included Lepidoptera. After a first duplication in Heliconiini, the gene was duplicated several other times, with some losses, resulting in between 11 and 14 copies per species within Heliconiini. Within Ithomiini, ABP6 independently duplicated in different lineages, resulting in between three and five copies of the gene per species (Figure 4; Table S5).

3.3 | Interspecific and Intraspecific Diversity in Antennal Gene Expression

To understand sexual dimorphism at the level of antennal gene expression, we analysed transcription expression level in both sexes of three species, *Me. confusa*, *T. harmonia* and *Mc. polymnia*. A total of 686 million reads were obtained after sequencing all eighteen libraries. All libraries show good statistics in terms of GC distribution, quality of sequences and redundancy. On average 38M pair-reads were obtained for each sample, which resulted in a 33M uniquely mapped reads on average (87%). Although one sample (*Me. confusa* F1) shows a lower degree of percentage of uniquely mapped reads, the absolute values (22M) and PCA (Figure S6) shows no bias or possible artifacts and similar degree of variance compared with the other species (Figure S7-S8). Overall, females of *T. harmonia* and *Mc. polymnia* showed more variance compared with *Me. confusa*, where the trend is reversed. Considering a minimum threshold of 10 transcript per million (TPM) on average, the three species expressed similar numbers of genes (on average ~6.4k genes per species).

In the first step, we used positional orthology inference to identify single-copy OGs (see Methods) to investigate the expression profile across species. The PCA plot shows a strong pattern of interspecific divergence in expression patterns (Figure 1c), also evident from very long branches in the tree obtained clustering samples by gene expression (Figure S9). The topology shows *Me. confusa* and *T. harmonia* forming sister clades, as in the phylogenetic tree (Figure 1), showing that phylogenetic signal can be detected in gene expression regardless of the short branches at the base of Ithomiini. Overall male transcriptomes seem to be more similar compared to

females, which show higher interspecific heterogeneity. These findings suggest extensive functional divergence between sexes and species.

To characterize sex-specific genes in the antennal transcriptomes, we identified 499, 380, and 772 differentially expresses genes (DEG) between sexes in *Me. confusa*, *Mc. polymnia* and *T. harmonia,* respectively (Figure 5; Figure S10; Table S6; Bayes Factor > 0.95). The GO term enrichment analyses using DEGs only found significant terms in male-biased genes of *T. harmonia* (upregulated genes in males; adjusted p-value < 0.05). These 30 GO terms involve 362 of the 426 DEGs genes enriching almost exclusively biological processes involved in the biosynthesis of lipids, monocarboxylic acids, fatty acids, terpenoids, organonitrogens, their regulation, and response to external stimuli (Table S7). Notable loci within the 346 females-biased DEGs, include *doublesex* (*dsx*), which controls somatic sexual differentiation and courtship behaviour, mediates the development of sex-specific pheromone organs in butterflies (Prakash 2020); *juvenile hormone acid methyltransferase* (*jhamat*), which has effects on courtship behaviour in *Drosophila* (Wijesekera et al. 2016); *SLC22A*, a cation transporter implicated in the regulation of olfactory learning (Gai et al. 2016), and *Epidermal stripes and patches* (*Esp*), which encodes a protein involved in female remating receptivity (Findlay et al. 2014). Within *Mc. Polymnia*, there are 221 and 159 male-biased and female-biased genes in, respectively. In the male-biased genes we also find genes related to lipid processes such as *farjavit* (*frj*), a lysophospholipid acyltransferase, *Dpr-interacting protein γ*, and *Kinesin heavy chain 73* (*Khc-73*), and *bric a brac* (*bab*). Female-biased genes include *apolipophorin* (*apolpp*) and *Dpr-interacting protein ζ* (*DIP-κ*).

Finally, in contrast to the even balance of male and female biased genes In *Mc. polymnia* and *T. harmonia*, within *Me. confusa* we detect a very skewed proportion of male-biased and female-biased DEGs with 85 and 414, respectively. Male-biased DEGs involve seven members of the cuticular proteins (*Cpr*) highly expressed with very high fold-change (Log_2 FC > 4.4); but also *ebony* and *yellow-e* involved in dark pigmentation in butterflies (Article 2017); *TrpA1*, a thermotactic sensor; *defective proboscis extension response 18* (*dpr18*) and *pumilio* (*pum*) involved in synapse organization and long-term memory (Dubnau et al. 2003). In the female-biased DEGs we find *sarah* (*sra*), which is involved in the regulation of female receptivity, post-mating receptivity (Ejima et al. 2004); *murashka* (*mura*) implicated in long-term memory (Akalal et al. 2011); and *supernumerary limbs* (*slmb*), involved in the regulation of circadian clocks (Srikanta and Cermakian 2021).

Finally, to test for a deeper phylogenetic expression pattern between sexes, we assessed the amount of overlap in sex biased gene expression between species using single-copy orthologs. Echoing the results from the PCA, we find very little overlap between DEGs among the three species (18, 10 and 8 for *T. harmonia* vs. *Me. confusa, T. harmonia* vs. *Mc. polymnia* and *Me. confusa* vs. *Mc. polymnia*, respectively, Figure 5). Between *Me. confusa* and *T. harmonia* there are 18 DEGs in common, three of these have opposite biases, while 15 show the same bias (upregulated only in males or females), of these 13 are male-biased. Among these, are *germ cell-expressed bHLH-PAS* (*gce*) and *klumpfuss* (*klu*); while among the shared female-biased genes there is trehalose transporter, a mediator in the bidirectional transfer of trehalose and regulating trehalose levels in the haemolymph; and between *T. harmonia* and *Mc. polymnia* there are 10 DEGs in common, four with opposite

gene expressions, four unidirectionally expressed in females and two in males. Between *Me. confusa* and and *Mc. polymnia* there are the least DEGs in common, only eight, three and two unidirectionally expressed in females and two in males, respectively, and three discordant. There is only one gene shared in all comparisons: *biniou* (*bin*), which was always found as male-biased with a very high fold-change (Log_2 FC > 5).

Sex-biased Antennal Chemosensory Gene Expression Patterns

We detected high transcriptomic diversity for chemosensory genes highlighting deep transcriptomic differences that are likely the result of adaptations to recognize odorant volatiles during mate choice and host plant selection. With the exception of only 19 ORs, all the remaining 187 (91%) ORs are expressed across the three species, with higher a higher proportion of the OR gene family being expressed than for IRs + IGluRs or OBPs with 101 (85%) and 83 (86%), respectively. The OGs OR68, IR143, IR7d4, OBP15, and OBP32 are consistently not expressed in all the three species. Also, five duplicated ORs/IRs from *Me. confusa* ($OR42_{Loc44}$, $OR4_{Loc22}$, $IGluRIa_{Loc46}$, $IGluR_{Loc45}$, $IR75pA_{Loc47}$) three from *T. harmonia* ($OR24_{Loc40}$, $OR4_{Loc32}$, $OR247_{Loc30}$), and four *Mc. polymnia* ($OR24_{Loc11}$, $OR4_{Loc12}$, $OR4_{Loc13}$, $OR38_{Loc29}$) are not expressed and are therefore likely non-functional or expressed in the larval stage.

To explore sex-specific chemoreception we identified OR genes which display sex-biased DE in the analysis described above. Among all the 201 ORs tested ten show DE between sexes. Specifically, *Me. confusa* has one female-biased ($MpolOR44_{Loc34}$) and one male-biased ($MpolOR8_{Loc11}$) receptors. The male-biased gene belongs to the same OG that is also male-biased in another butterfly, *Heliconius cydno* (van Schooten et al. 2020), while the female-biased gene belongs to an OG where no functional information is available, therefore showing the association between this OG and sexual dimorphism for the first time. In *T. harmonia* we detected four DE ORs, all female-biased. Two of these genes belong to the same OG37 ($TharOR37_{Loc31}$, $TharOR37_{Loc33}$), and are located on the same scaffold, 35kbp apart on opposite strands, and are likely the result of an interspersed duplication that possibly also copied their associated regulatory elements. This OG is also lost from almost all species from our dataset and an intact functional domain is present only in three Ithomiini species (*T. harmonia, Ml. menophilus* and *Ml. marsaeus rileyi*). The other two female-biased genes ($TharOR46_{Loc08}$ and $TharOR51_{Loc57}$) belong to OGs that have been found to be sex-biased in *Heliconius cydno* (van Schooten et al. 2020); $TharOR46_{Loc08}$ with the same female-bias and $TharOR51_{Loc57}$ being male-biased in *Heliconius*. $TharOR46_{Loc08}$ was also found to be sex-biased in *Mc. polymni*a, although in the opposite direction ($MpolOR46_{Loc43}$, female-biased), together with the female-biased $MpolOR40_{Loc43}$. *Mc. polymni*a has two other DE ORs, both male-biased: $MpolOR23_{Loc31}$, also found to be male-biased in *Spodoptera litura* (Feng et al. 2015), and $MpolOR30_{Loc23}$, which belong a novel pheromone receptor clade (Walker III et al. 2019). Only one OR, OR46, is sex-biased across all species, but not in a consistent direction.

Among all the 86 IRs and IGluRs tested three IRs and three IGluRs display DE between sexes. In *Me. confusa* the co-receptor CoIR8 was detected as male-biased and the GluRIIb female-biased. In *T. harmonia* the "Lepidoptera specific" IR87 and the antennal IR31 are male-biased and female-biased, respectively. While the

IGluR1 is identified as male-biased. *Mc. polymnia*, instead shows a single female-biased receptor, the IGluRIa with a very high fold-change (FC_{log2} = 3.9). Finally, although OBPs constitute the higher fraction of antennal CSGs, none of the "canonical" pheromone binding proteins (PhBPs) show differential expression between the sexes, but two other transcripts do, both in, *T. harmonia*: the antennal binding protein 8 (ABP8) and X (ABPX), and both being female-biased. Notably, the functional domain of the ABPX in the Ithomiini, instead of showing the canonical conserved domain of other ABPs and GOBPs (PFAM:01395), as observed for this OG in Heliconiinae, seems to be more similar the functional domains of the PhBPs (smart00708). A shift that seems to occur also within the orthologs of ABP1 and OBP29, possibly hinting to a gain-of-function.

3.4 | Candidate ORs Linked to the Macro-Glomeruli Complex in Ithomiini Butterflies, and High Expression of Coreceptors

A relatively large proportion of total antennal gene expression is accounted for by chemosensory genes (CSGs; ORs, IRs and OBPs). On average more than 5% of expressed transcripts belong to these genes (*Me. confusa*: 5.9%; *T. harmonia*: 5.9%; *Mc. polymnia*: 5.2%). Females of *Me. confusa* and *T. harmonia* express more CSGs compared to males (~7% of fold-enrichment), while in *Mc. polymnia* males express more CSGs. Notably, OBPs constitute over the 96% of those transcripts, therefore the actual overall expression of CSGs over non-CSGs can be attributed to this gene family (Wilcoxon rank-sum test 'one-side' *P* value < $2.2x10^{-16}$). The shape of the distribution of gene expression differs between gene families, with ORs having the smallest variance and OBPs the largest. OBPs also have a bimodal distribution of expression (Hartigans' dip test D = 0.050155, *P* value < 0.005). Notably, the distributions of OR gene expression varies among species, with *Me. confusa* having lower expression of ORs than *Mc. polymnia* (medians: 1 TPM and 1.3 TPM; Wilcoxon rank-sum test 'one-side' *P* value = 0.017), and *Mc. polymnia* lower than *T. harmonia* (median: 2.9 TPM; Wilcoxon rank-sum test 'one-side' *P* value = $4.35x10^{-6}$). In contrast, IRs and OBPs do not show this significant interspecific variation.

To putatively attribute the expanded MGC glomeruli to ORs we also looked at the within CSGs expression in our species. In *Me. confusa* ORs seem to show a more restricted variation in the gene expression of OR with only Orco markedly more expressed, this is consistent with the monomorphic nature of their antennal lobe glomeruli, and the absence of macroglomeruli in this species. In the other two species, *Mc. polymnia* and *T. harmonia*, a longer tail of the distribution is present. In *Mc. Polymnia*, alongside the co-receptor Orco, there are two more dominant transcripts OR46 and OR51, which show sex-biased expression. In *T. harmonia*, excluding Orco, there are five dominant receptors, the two also found in *Mc. Polymnia* (OR45 and OR51), and three more: OR32, OR42, and OR38, which belong to the "novel" pheromone clade. Within IRs and IGluRs, the co-receptors CoIR8 and CoIR25a, and the antennal IR75q1 are highly expressed in all the species. Whilst, in the upper tail of the distributions of the other two species we find the co-receptors CoIR76; while in *T. harmonia* the most expressed IRs is the antennal IR75pB, which is three times more expressed that Orco (FC_{log2} > 1.64). OBPs are more most abundant component of the CSGs by far and within them, the ABPs and the PhBPs are the most dominant transcripts, specifically across all the species the most abundant OBPs are always the PhBP-C, PhBP-D,

and PhBP2; the ABPs: ABP2, ABP3, ABP8 and ABPX; and the GOBP5. While the bottom part of the distribution is always occupied almost exclusively by GOBPs with ABP6A, and PhBP-B in *T. harmonia*.

3.5 | Differential Selection Regimes Across Chemosensory Genes

To identify differential selection regimes across Ithomiini, we performed the RELAX test computing the *k* parameter on each annotated ORs, IRs and OBPs across all Ithomiini species. Comparisons of the *k* distributions also provide a proxy to understand which of the ACG family is the most and lesser dynamic within Ithomiini. The scan for relaxation/intensification shows very different patterns among chemosensory gene families. ORs are by far the most dynamic with a median *k* of 1 but with a standard variation of 13.5. In comparison IRs plus IGluRs and OBPs, with a median still around 1, but with a standard variation of 7.9 and 11.3 respectively. Within ORs we found eight loci to be under intensified selection and twice as many, 16 to be under relaxed selection, showing how strong the turnover of selection is acting on ORs. Five of the relaxed loci are at the bottom distribution of OR expressions, hinting to a possible low functionality. Among the eight loci under intensification, MpolOR51$_{Loc22}$ and TharOR38$_{Loc64}$ are among the most highly expressed loci, with OR38 belonging to the "novel" pheromone clade. Furthermore, the loci of OR38 and OR51 are also under intensification in *G. morgane* and *Ml. marsaeus rileyi*. In fact, there is a sign of intensification in two ORs of *Ml. marsaeus rileyi* (MmarOR51$_{Loc31}$ and MmarOR01$_{Loc51}$) and one in *Ml. menophilus* (MmenOR22$_{Loc03}$); two closely related species, suggesting that these loci might be related to mate or host plant differentiation. Within IRs and IGluRs, there are nine loci under relaxed selection, among them MpolIR40a$_{Loc01,}$ which is lowly expressed, and two that belong to the OG OR74, supporting a general trend toward the loss of these genes withing Nymphalids. Only three of these receptors are identified as evolving under intensified selection. The loci under intensification are not the IRs but the IGluRs, two in *G. morgane* (GotoIGluRIb$_{Loc01}$ and GotoIGluRIIc$_{Loc27}$) and one in *Mc. polymnia* (MpolNMDA$_{Loc38}$), genes that may play a role in synaptic plasticity, synaptogenesis, excitotoxicity, memory acquisition and learning. Finally, consistent with their general conservation, OBPs show the least variation in selection regime. The lack of selection turnover is an indication how important these genes are for the correct functionality of the whole chemosensory circuit.

4 | Discussion

The ability to detect odours plays a fundamental role in organisms, not only because it enables them to recognise environmental chemical signals, but also because it allows communication between individuals. This is particularly true in Ithomiini butterflies where the detection of specific compounds is linked to a strong inter- and intra-sexual behaviours. Therefore, chemosensory organs should harbour a complex pattern of gene expression not only strictly related to chemosensory genes, but a variety of non-chemosensory genes that have a great importance to support neuronal sensory function cells and to regulate the stimuli (Schmidt and Benton 2020; Scalzotto et al. 2022). It is therefore plausible that selection could act in modulating the expression of both

chemosensory and non-chemosensory genes generating distinct expression patterns within and between species.

In this study, we generate high-quality/highly contiguous reference genomes for four Ithomiini species, a tribe of diurnal butterflies with particular reliance on olfactory cues. The species chosen reflect phylogenetic and ecological differences across the tribe, and were then used to assess divergent patterns of chemosensory evolution, and to putatively associate expression profiles with distinct neuroanatomical differences between species. Specifically, based on patterns of neuromorphological variation between ithomiines and other butterflies, and within ithomiines we predicted: *i*) a molecular signature that parallels the expansion of specific glomeruli in the ithomiine antennal lobe (i.e. particularly high expression of a small number of ORs); *ii*) evidence of sexual dimorphism in some, but not all, or these genes; *iii*) interspecific differences, with *Methona* displaying reduced variance and dimorphism in chemosensory gene expression than other ithomiines, in line with the homogeneous structure of this species' antennal lobe, which is more typical of other butterflies (Morris et al., 2022).

In addressing these predictions, we show how the antennal transcriptome is profoundly divergent between related species, suggesting distinct adaptations of these species in sensing the environment (Figure 1c, 6). This diversity, to a lesser degree, has previously been found in other lepidopteran radiation, including *Heliconius* species, where the antennal transcriptomics shows the largest diversity of expression patterns compared with mouthparts and legs (Wu et al. 2022). Similarly, in the only other ithomiine study on the antennal transcriptomics, in two subspecies of *Ml. marsaeus*, Prunier et al. (2021) found twice as many DE transcripts in the antennae (1028 transcripts), compared with imaginal discs. In our study, this species uniqueness seems to be so profound that the number of DEGs between sexes overlap minimally across species. Within these DEGs in common there is *Gce*, which is a male-biased gene in *Drosophila* that bind to the juvenile hormone, which has been shown to determine sex-dimorphisms in the gut regulating intestinal stem cell proliferation (Drosophila et al. 2020); *Klu* a transcription factor involved in determination of the identities of neuroblast lineages in the central nervous system (Xiao et al. 2012); and *biniou* (*bin*), always male-biased across the three species, which encodes for a transcription factor with important regulator functions for the development of the visceral musculature of the midgut (Zaffran et al. 2001), and also expressed uniquely in *Drosophila* males (Brown et al. 2014). These present candidates for the exploration of conserved network gene expression in sex dimorphism in butterflies.

Looking at within species sexual dimorphism at the transcriptomic level for non-chemosensory genes, we also found some recurrent signal coming from genes related to lipid processes. In *T. harmoria*, we found 66 DEGs that almost exclusively enrich biological processes involved in the biosynthesis lipids and hormone metabolic process; while within *Mc. Polymnia*, there are male-biased genes, such as *farjavit* (*frj*), a lysophospholipid acyltransferase, related to synaptic transmission as *scramblase*, *Dpr-interacting protein γ*, and *Kinesin heavy chain 73* (*Khc-73*), which may be implicated in regulation of olfactory learning (Guven-ozkan et al. 2020), all involved in lipid-related pathways as is *bric a brac* (*bab*), which has been shown to control sex

pheromone choice by male in the moths of *Ostrinia nubilalis* (Unbehend et al. 2021). Similarly to the female-biased genes in which *apolipophorin* (*apolpp*), a lipid transporter (Ugrankar et al. 2019); and *Dpr-interacting protein ζ* (*DIP-κ*), involved in establishment of synaptic specificity at neuromuscular junction (Bornstein et al. 2021). Genetic studies in *D. melanogaster* have revealed that lipid transporters, such as ATP8B, have pivotal role in olfactory sensory neurons classes related to pheromone ORs (Liu et al. 2014; Soo et al. 2014). Although the role of lipid composition on OR signalling is unclear, these transporters are required to flip aminophospholipids (*e.g.*: phosphatidylserine) between membrane leaflets, potentially affecting cilia morphology (Garcia III et al. 2018).

The comprehensive investigation of ACGs across six representative species of the Ithomiini tribe has yielded a wealth of insights into the intricate interplay between genetic diversity, ecological specialization, and sensory adaptation. The multifaceted landscape of chemosensory gene evolution, expression, and potential functional significance offers a rich platform for understanding the genetic underpinnings of sensory perception and its role in driving behaviours and ecological interactions. In our study, the main ACGs were manually curated for the 15 species considered in our analyses. Compared to a previous study on *Ml. marsaeus* and *Ml. menophilus* ACGs (Prunier et al. 2021; Gauthier et al. 2023), we found a slightly different number of ACGs, probably due to a different criteria of characterisation of the gene. Specifically, we excluded all putative loci without a valid PFAM domain e-value score. For *H. melpomene* and *D. plexippus* we instead found a higher number. This let us observe a more accurate picture of chemosensory evolution at the gene family level; more precisely contextualise changes we registered within the Ithomiini tribe. Strong differences were observed within and among ACGs. In fact, compared with ORs, IGluRs, IRs and OBPs showed an unexpected much lower turnover rates with stronger signal of conservation. Among all IRs, the Lepidoptera-specific (LS-IRs) and Divergent IRs showed the highest turnover. We detected an Ithomiini-specific clade, the IR1b, possibly a vestigial clade of an old duplication that occurred within Danainae, similar to IR7d4 which was lost in all Ithomiini, but also several duplications in IR7d2, IR7d4, and IR143. Odorant binding proteins show high conservation, with very few exceptions within the antennal binging proteins (ABPs). These include, the expansions of ABP6a, a clade hugely expanded in Nymphalidae, ABP8 and ABP1. Odorant receptors, on the other hand, showed the highest turnover, with a number of gene losses and possible gain of function. In particular, these butterflies seem to have lost genes from the "Known" butterfly pheromone receptor clade, but at the same time expanding their repertoire within the "Novel" pheromone receptor clade (OR30 and OR83), associated with moths (Bastin-Héline et al. 2019); and two other Ithomiini-specific expansions within OR53 and OR42, orthologous to HarmOR40, which has affinity with terpenes (Guo et al. 2021), a class of compounds found in the androconia of the ithomiine butterfly *Ithomia salapia* (Mann et al. 2020).

From the point of view of ACG expression, we investigated two patterns: the differential ACG expression between sexes, and the relative expression within ACG expression. From a behavioural perspective, the presence of sex-specific receptors could provide insights into the genes that hold greater relevance for each sex. We can speculate that, while in males there is a strong selection towards sensing pheromones, mostly due to male-to-male competitive interactions; in females there is a strong selection for locating and choosing suitable host plant

for egg deposition, and for choosing males with high PA concentrations in their nuptial gifs. From the neuro-anatomical perspective, because each glomerulus in the antennal lobe is associated with one or two chemosensory receptors, and because MGCs, in moth typically involved in processing olfactory pheromone, we might expect to see skewed distribution of expressions in chemosensory receptors, with the one more associated with MGCs being the one more expressed, compared to others. Among the three species *Me. confusa*, a species which lacks MGCs (Morris et al. 2021), showed the least differentiation both in terms of sex-specific ACGs, with one male-biased and one female-biased OR, and the least skewed distribution of ACGs. In *Mc. polymnia*, a species with MGCs, we found four sex-biased ORs, these corresponds to two female-biased ORs: OR23 and OR40, the latter being orthologous to SlitOR40, being associated plant volatile receptor (Revadi et al. 2021); and two male-biased ORs: OR46 and OR30, the first being the second most expressed OR after Orco, and the second part of the "Novel" pheromone receptor clade, with possible implications for convergent evolution of reliance on long distance pheromone detection with moths. Finally, *T. harmoria*, which showed the largest skewed distribution pattern, with six ORs more highly expressed, and with four, one and two female-biased ORs, IRs and OBPs, respectively; and one and one male-biased IGluR and IR, respectively. While data on the antennal lobe morphology of *T. harmoria* is currently lacking, the presence of sexual dimorphism in hind-wing hair pencils, which are present only in males (Fox 1940), in common with other ithomiini that show enlarged glomeruli, would suggest that these outlying ORs will reflect similar neural adaptations in this species.

Finally, to gain insight into selection pressures we also explored differences at the nucleotide level, identifying in ORs the gene family with the most dynamic range of shifts, and the highest number of genes under intensification. Of these, two: OR51 in *Mc. polymnia* and OR38 in *T. harmoria*, are under intensification of selection and more highly expressed compared with other ORs, suggesting a fundamental role of these genes in these species. A diversification pattern in ORs that is also evident looking at the two closely related species: *Ml. marsaeus rileyi* and *Ml. menophilus*, only ~ 5 Mya with three ORs in intensification, with possible implications in assortative mating. Overall, the variation in gene counts, expression patterns and relaxation/intensification of selection indicate possible roles that ORs play in sensory perception and behavioural responses. In contrast, the more conserved patterns in both: IGluRs and OBPs suggest a more profound fundamental importance in maintaining core sensory functions, with a delicate balance between gene innovation and conservation that shapes the chemosensory landscape in butterflies.

5 | Conclusions

While in moths, the expansions of glomeruli in the antennal lobe, typically sexually dimorphic and responsive to pheromones, are quite common (Rospars and Hildebrand 1992; Christensen and Hildebrand 2002), in butterflies these anatomical modifications were lost and independently re-emerged in the hyper-diverse Ithomiini, the only tribe of butterflies currently known to possess this kind of olfactory specialization (Montgomery and Ott 2015; Morris et al. 2021). Here, by assembling the genomes from four representatives of the tribe and sequencing sex-specific antennal transcriptomic of four of them, we provide evidence that the convergent evolution of these

neuro-anatomical features does not involve olfactory innovations in Ithomiini, but on the other hand they adapted already available receptors, possibly achieving the same sensory function. We showed that *Me. confusa*, the lesser sexual dimorphic species, shows less gene expression differences compared with the other species taken into account, fitting the neuro-anatomy.

In summary, this study provides a comprehensive exploration of chemosensory gene evolution, expression patterns, and potential functional implications within the Ithomiini tribe. The identification of sexual dimorphism, the presence of macro-glomeruli complexes, and the detection of differential selection regimes enrich our understanding of sensory adaptations in butterflies. By elucidating the genetic foundations of chemosensory diversity and its ecological significance, this research contributes not only to the field of sensory ecology but also lays the groundwork for future investigations into the genetic drivers of behaviour, adaptation, and speciation in insects.

Author Contribution

S.H.M. conceived the study, with input from F.C. and B.J.M.; B.J.M. and F.C. extracted mRNA and gDNA; F.C., J.M. and D.R. curated transposable element curation; F.C. analysed genomic and transcriptomic data; F.C. wrote major parts of the manuscript; F.C., S.H.M., B.J.M., J.M. and D.R. contributed to the final version of the paper.

Acknowledgments

This work was supported by a NERC Independent Research Fellowship NE/N014936/1 to S.H.M.. F.C. and S.H.M. are grateful to the High-Performance Computing team at the Advanced Computing Research Center, University of Bristol for support.

Conflicts of Interest Statement

The authors declare no conflicts of interest.

Data Availability Statement

Genomic and transcriptomic raw reads together with genome assemblies have been made available via NCBI's Genbank BioProjects: PRJNA1023055; PRJNA1023057; PRJNA1023058; PRJNA1023059. Custom scripts and the updated TE library for Nymphalidae ALLFAMILIES_MOD-CDHIT.DEF.FA is available at https://github.com/francicco/-IthomiineChemosensoryProject.git.

ORCID

Francesco Cicconardi https://orcid.org/0000-0001-6509-6179
Billy J Morris https://orcid.org/0000-0003-3412-1794
Jacopo Martelossi https://orcid.org/0000-0003-4227-0669
David Ray https://orcid.org/0000-0002-3340-3987
Stephen H Montgomery https://orcid.org/0000-0002-5474-5695

References

Abuin L, Bargeton B, Ulbrich MH, Isacoff EY, Kellenberger S, Benton R. 2011. Functional Architecture of Olfactory Ionotropic Glutamate Receptors. *Neuron* 69:44–60.

Akalal DG, Yu D, Davis RL. 2011. The Long-Term Memory Trace Formed in the Drosophila ␣ / ␣ Mushroom Body Neurons Is Abolished in Long-Term Memory Mutants. 31:5643–5647.

Alexa A, Rahnenführer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22:1600–1607.

Amin H, Lin AC. 2019. Neuronal mechanisms underlying innate and learned olfactory processing in *Drosophila*. *Curr. Opin. Insect Sci.* [Internet] 36:9–17. Available from: https://doi.org/10.1016/j.cois.2019.06.003

Armstrong J, Hickey G, Diekhans M, Deran A, Fang Q, Xie D, Feng S, Stiller J, Genereux D, Johnson J, et al. 2019. Progressive alignment with Cactus: A multiple-genome aligner for the thousand-genome era. *bioRxiv*.

Armstrong J, Hickey G, Diekhans M, Fiddes IT, Novak AM, Deran A, Fang Q, Xie D, Feng S, Stiller J, et al. 2020. Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature* 587:246–251.

Article H. 2017. Genetic Basis of Melanin Pigmentation in Butter fl y Wings. 205:1537–1550.

Auer TO, Alvarez-Ocana R, Cruchet S, Benton R, Arguello JR. 2022. Copy number changes in co-expressed odorant receptor genes enable selection for sensory differences in drosophilid species. *Nat. Ecol. \& Evol.* 6:1343–1353.

Auer TO, Khallaf MA, Silbering AF, Zappia G, Ellis K, Álvarez-Ocaña R, Arguello JR, Hansson BS, Jefferis GSXE, Caron SJC, et al. 2020. Olfactory receptor and circuit evolution promote host specialization. *Nature* [Internet] 579:402–408. Available from: http://dx.doi.org/10.1038/s41586-020-2073-7

Bastin-Héline L, de Fouchier A, Cao S, Koutroumpa F, Caballero-Vidal G, Robakiewicz S, Monsempes C, François MC, Ribeyre T, Maria A, et al. 2019. A novel lineage of candidate pheromone receptors for sex communication in moths. *Elife* 8:1–17.

Bateman A. 2019. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.*

Beccaloni GW. 1997a. Vertical stratification of ithomiine butterfly (Nymphalidae : Ithomiinae) mimicry complexes : the relationship between adult flight height and larval host-plant height. *Biol. J. Linnaean Soc.* 62:313–341.

Beccaloni GW. 1997b. Ecology, natural history and behaviour of Ithomiine butterflies and their mimics in Ecuador (Lepidoptera: Nymphalidae: Ithomiinae). *Trop. Lepid. Res.* 8:103–124.

Bechstedt S, Senthilan PR, Piepenbrock D, Ovezmyradov G, Pauls S, Winkler M, Mo W, Howard J, Go MC. 2012. Drosophila Auditory Organ Genes and Genetic Hearing Defects. *Cell*.

Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. 2009. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* [Internet] 136:149–162. Available from: http://dx.doi.org/10.1016/j.cell.2008.12.001

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. *Biol. Direct* [Internet] 7:12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22510480%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3438057

Bornstein B, Meltzer H, Adler R, Alyagor I, Berkun V, Cummings G, Reh F, Keren-shaul H, David E, Riemensperger T, et al. 2021. Transneuronal Dpr 12 / DIP- d interactions facilitate compartmentalized dopaminergic innervation of Drosophila mushroom body axons. :1–21.

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al. 2014. Diversity and dynamics of the Drosophila transcriptome. *Nature* [Internet]:1–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24670639

Brown Jr KS. 1984. Chemical ecology of dehydropyrrolizidine alkaloids in adult Ithominae (Lepidoptera: Nymphalidae). *Rev. Bras. Biol.* 44:435–460.

Brown KS. 1987. Chemistry at the Solanaceae / Ithomiinae Interface Author (s): Keith S . Brown , Jr . Source : Annals of the Missouri Botanical Garden , 1987 , Vol . 74 , No . 2 (1987), pp . 359-397 Published by : Missouri Botanical Garden Press Stable URL : https://. *Missouri Bot.*

729 *Gard.* 74:359–397.
730 Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and
731 AUGUSTUS supported by a protein database. *NAR Genomics Bioinforma.* 3:1–11.
732 Buchfink B, Reuter K, Drost H. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* [Internet] 18. Available
733 from: http://dx.doi.org/10.1038/s41592-021-01101-x
734 Carlsson MA, Schäpers A, Nässel DR, Janz N. 2013. Organization of the olfactory system of nymphalidae butterflies. *Chem. Senses* 38:355–367.
735 Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540–552.
736 Chazot N, Elias M, Willmott KR, Lamas G, Freitas AVL, Piron F, Carlos P, James FA, Donna M, Silva L De. 2019. Renewed diversification following
737 Miocene landscape turnover in a Neotropical butterfly radiation. :1118–1132.
738 Christensen TA, Hildebrand JG. 2002. Pheromonal and host-odor processing in the insect antennal lobe : how different ? :393–399.
739 Cicconardi F, Marcatili P, Arthofer W, Schlick-Steiner BC, Steiner FM. 2017. Positive diversifying selection is a pervasive adaptive force
740 throughout the Drosophila radiation. *Mol. Phylogenet. Evol.* 112:230–243.
741 Cicconardi F, Milanetti E, Castro ECP De, Mazo-vargas A, Belleghem SM Van, Ruggieri AA, Rastas P, Hanly J, Evans E, Jiggins CD, et al. 2023.
742 Evolutionary dynamics of genome size and content during the adaptive radiation of Heliconiini butterflies. *Nat. Commun.* 14:5620.
743 Couto A, Young FJ, Atzeni D, Marty S, Flórez LM, Hebberecht L, Monllor M, Neal C, Cicconardi F, Mcmillan WO, et al. 2023. Rapid expansion
744 and visual specialisation of learning and memory centres in the brains of Heliconiini butterflies. *Nat. Commun.* 14.
745 Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson TJ, Benton R. 2010. Ancient protostome origin of chemosensory
746 ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* [Internet] 6:e1001064. Available from:
747 http://dx.plos.org/10.1371/journal.pgen.1001064
748 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq
749 aligner. *Bioinformatics* [Internet] 29:15–21. Available from:
750 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3530905&tool=pmcentrez&rendertype=abstract
751 Drosophila U, Millington JW, Rideout EJ. 2020. Sexual Dimorphism : Ecdysone Modulates Sex Differences in the Gut. *Curr. Biol.* [Internet]
752 30:R1327–R1330. Available from: https://doi.org/10.1016/j.cub.2020.08.088
753 Dubnau J, Chiang A-S, Grady L, Barditch J, Gossweiler S, McNeil J, Smith P, Buldoc F, Scott R, Certa U, et al. 2003. The staufen/pumilio pathway
754 is involved in Drosophila long-term memory. *Curr. Biol.* 13:286–296.
755 Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* [Internet] 14:755–763. Available from:
756 http://bioinformatics.oxfordjournals.org/content/14/9/755.short
757 Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput. Biol.* 7.
758 Ejima A, Tsuda M, Takeo S, Ishii K, Matsuo T, Aigaki T. 2004. Expression Level of *sarah*, a Homolog of DSCR1, Is Critical for Ovulation and
759 Female Courtship Behavior in Drosophila melanogaster. *Genetics* 2087:2077–2087.
760 El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. 2019. The Pfam protein
761 families database in 2019. *Nucleic Acids Res.*
762 Elias M, Gompert Z, Jiggins C, Willmott K. 2008. Mutualistic interactions drive ecological niche convergence in a diverse butterfly community.
763 *PLoS Biol.* 6.
764 Eyun S, Soh HY, Posavi M, Munro JB, Hughes DST, Murali SC, Qu J, Dugan S, Lee SL, Chao H, et al. 2017. Evolutionary History of Chemosensory-
765 Related Gene Families across the Arthropoda. 34:1838–1862.
766 Falcon S, Gentleman R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics* 23:257–258.
767 Feng B, Lin X, Zheng K, Qian K, Chang Y, Du Y. 2015. Transcriptome and expression profiling analysis link patterns of gene expression to
768 antennal responses in Spodoptera litura. ??? [Internet]:1–12. Available from: ???
769 Fiddes IT, Armstrong J, Diekhans M, Nachtweide S, Kronenberg ZN, Underwood JG, Gordon D, Earl D, Keane T, Eichler EE, et al. 2018.
770 Comparative Annotation Toolkit (CAT) - simultaneous clade and personal genome annotation. *Genome Res.* [Internet]:231118.
771 Available from: https://www.biorxiv.org/content/early/2017/12/11/231118
772 Findlay GD, Sitnik JL, Wang W, Aquadro CF, Clark NL, Wolfner MF. 2014. Evolutionary rate covariation identifies new members of a protein
773 network required for Drosophila melanogaster female post-mating responses. *PLoS Genet.* [Internet] 10:e1004108. Available from:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3894160&tool=pmcentrez&rendertype=abstract

Fleischer J, Pregitzer P, Breer H, Krieger J. 2018. Access to the odor world : olfactory receptors and their role for signal transduction in insects. *Cell. Mol. Life Sci.* 75:485–508.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. 117:9451–9457.

Fox MR. 1940. A Generic Review of the Ithomiinae (Lepidoptera, Nymphalidae). *Trans. Am. Entomol. Soc.* 66:161–207.

Di Franco A, Poujol R, Baurain D, Philippe H. 2019. Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. *BMC Evol. Biol.* 19:1–17.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. BIOINFORMATICS APPLICATIONS NOTE Sequence analysis CD-HIT : accelerated for clustering the next-generation sequencing data. 28:3150–3152.

Gai Y, Liu Z, Cervantes-sandoval I, Davis RL, Gai Y, Liu Z, Cervantes-sandoval I, Davis RL. 2016. Drosophila SLC22A Transporter Is a Memory Suppressor Gene that Influences Cholinergic Neurotransmission to the Mushroom Bodies Article Drosophila SLC22A Transporter Is a Memory Suppressor Gene that Influences Cholinergic Neurotransmission to the Mushroom . *Neuron* [Internet] 90:581–595. Available from: http://dx.doi.org/10.1016/j.neuron.2016.03.017

Garber M, Grabherr MG, Guttman M, Trapnell C. 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat. Methods* [Internet] 8:469–477. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21623353

Garcia III G, Raleigh DR, Reiter JF. 2018. How the Ciliary Membrane Is Organized Inside-Out to Communicate Outside-In. *Curr. Biol.* [Internet] 28:R421–R434. Available from: https://doi.org/10.1016/j.cub.2018.03.010

Gauthier J, Meier J, Legeai F, Mcclure M, Whibley A, Bretaudeau A, Boulain H, Parrinello H, Mugford ST, Durbin R, et al. 2023. First chromosome scale genomes of ithomiine butterflies (Nymphalidae : Ithomiini): Comparative models for mimicry genetic studies. *Mol. Ecol. Resour.*:872–885.

Gauthier J, Meier J, Legeai F, McClure M, Whibley A, Bretaudeau A, Boulain H, Parrinello H, Mugford ST, Durbin R, et al. 2023. First chromosome scale genomes of ithomiine butterflies (Nymphalidae: Ithomiini): Comparative models for mimicry genetic studies. *Mol. Ecol. Resour.*:1–14.

Giesen L Van, Garrity PA. 2022. More than meets the IR : the expanding roles of variant Ionotropic Glutamate Receptors in sensing odor , taste , temperature and moisture [version 1 ; peer review : 2 approved]. 6:1–7.

Goubert C, Craig RJ, Bilat AF, Peona V, Vogan AA, Protasio A V. 2022. A beginner ' s guide to manual curation of transposable elements. *Mob. DNA* [Internet]:1–19. Available from: https://doi.org/10.1186/s13100-021-00259-7

Guo M, Du L, Chen Q, Feng Y, Zhang J, Zhang X, Tian K, Cao S, Huang T, Jacquin-Joly E, et al. 2021. Odorant Receptors for Detecting Flowering Plant Cues Are Functionally Conserved across Moths and Butterflies. *Mol. Biol. Evol.* 38:1413–1427.

Guven-ozkan T, Busto GU, Jung J, Drago I, Davis RL. 2020. miR-92a Suppresses Mushroom Body-Dependent Memory Consolidation in Drosophila. 7:1–12.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* [Internet] 8:1494–1512. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3875132&tool=pmcentrez&rendertype=abstract

Hansson BS, Stensmyr MC. 2011. Evolution of insect olfaction. *Neuron* 72:698–711.

Herre M, Goldman O V, Lu T, Li H, Vosshall LB, Younger MA, Herre M, Goldman O V, Lu T, Caballero-vidal G, et al. 2022. Article Non-canonical odor coding in the mosquito ll Non-canonical odor coding in the mosquito. *Cell* [Internet] 185:3104-3123.e28. Available from: http://dx.doi.org/10.1016/j.cell.2022.07.024

Hill RI. 2010. Habitat segregation among mimetic ithomiine butterflies (Nymphalidae). :273–285.

Iyer MK, Chinnaiyan AM. 2011. RNA-Seq unleashed. *Nat. Biotechnol.* [Internet] 29:599–600. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21747384

Jackman SD, Coombe L, Chu J, Warren RL, Vandervalk BP, Yeo S, Xue Z, Mohamadi H, Bohlmann J, Jones SJM, et al. 2018. Tigmint: Correcting assembly errors using linked reads from large molecules. *BMC Bioinformatics* 19:1–10.

Jebb D, Huang Z, Pippel M, Hughes GM, Lavrichenko K, Devanna P, Winkler S, Jermiin LS, Skirmuntt EC, Katzourakis A, et al. 2020. Six

reference-quality genomes reveal evolution of bat adaptations. *Nature* [Internet] 583:578–584. Available from: http://dx.doi.org/10.1038/s41586-020-2486-3

Kawahara AY, Storer C, Carvalho APS, Plotkin DM, Condamine FL, Braga MP, Ellis EA, St Laurent RA, Li X, Barve V, et al. 2023. A global phylogeny of butterflies reveals their evolutionary history, ancestral hosts and biogeographic origins. *Nat. Ecol. Evol.*

Klopfenstein D V., Zhang L, Pedersen BS, Ramírez F, Vesztrocy AW, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, et al. 2018. GOATOOLS: A Python library for Gene Ontology analyses. *Sci. Rep.* 8:1–17.

Koontz MA, Schneider D. 1987. Sexual dimorphism in neuronal projections from the antennae of silk moths (Bombyx mori, Antheraea polyphemus) and the gypsy moth (Lymantria dispar). *Cell Tissue Res.* 249:39–50.

Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier D, Nekrutenko A, et al. 2020. HyPhy 2.5 - A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* 37:295–299.

Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. *F1000Research* [Internet] 6:1287. Available from: https://f1000research.com/articles/6-1287/v1

Larter NK, Sun JS, Carlson JR. 2016. Organization and function of Drosophila odorant binding proteins. *Elife* 5:1–22.

Leal WS. 2013. Odorant reception in insects: Roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* 58:373–391.

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendziorski C. 2013. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29:1035–1043.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.

Liu Y, Pearce MW, Honda T, Johnson TK, Charlu S, Sharma KR, Imad M, Burke RE, Zinsmaier KE, Ray A, et al. 2014. The Drosophila melanogaster Phospholipid Flippase dATP8B Is Required for Odorant Receptor Function. 10.

Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* 33:6494–6506.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:1–21.

Mann F, Szczerbowski D, Silva L De, Mcclure M, Elias M, Schulz S. 2020. 3-Acetoxy-fatty acid isoprenyl esters from androconia of the ithomiine butterfly Ithomia salapia. *J. Org. Chem.*:2776–2787.

Mapleson D, Venturini L, Kaithakottil G, Swarbreck D. 2018. Efficient and accurate detection of splice junctions from RNA-seq with Portcullis. *Gigascience* 7:1–11.

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, et al. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43:D222–D226.

Martelossi J, Nicolini F, Subacchi S, Pasquale D, Ghiselli F, Luchetti A. 2023. Multiple and diversified transposon lineages contribute to early and recent bivalve genome evolution. *BMC Biol.* [Internet] 21:1–23. Available from: https://doi.org/10.1186/s12915-023-01632-z

Massuda KAF, Trigo JOSÉR. 2009. Chemical defence of the warningly coloured caterpillars of Methona themisto (Lepidoptera : Nymphalidae : Ithomiinae). 5759:253–259.

Masters AR. 1992. Chemical defense in ithomiine butterflies (Nymphalidae: Ithomiinae). University of Florida

McClure M, Mahrouche L, Houssin C, Monllor M, Le Poul Y, Frérot B, Furtos A, Elias M. 2019. Does divergent selection predict the evolution of mate preference and reproductive isolation in the tropical butterfly genus Melinaea (Nymphalidae: Ithomiini)? *J. Anim. Ecol.* 88:940–952.

Mika K, Benton R. 2021. Olfactory Receptor Gene Regulation in Insects : Multiple Mechanisms for Singular Expression. 15:1–10.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37:1530–1534.

Missbach C, Dweck HKM, Vogel H, Vilcinskas A, Stensmyr MC, Hansson BS, Grosse-wilde E. 2014. Evolution of insect olfactory receptors. :1–22.

Montgomery SH, Ott SR. 2015. Brain Composition in Godyris zavaleta , a Diurnal Butterfly , Reflects an Increased Reliance on Olfactory Information. 891:869–891.

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B. 2011. How many species are there on Earth and in the ocean?Mace GM, editor. *PLoS Biol.* [Internet] 9:e1001127. Available from: http://dx.plos.org/10.1371/journal.pbio.1001127

Morris BJ, Couto A, Aydin A, Montgomery SH. 2021. Re-emergence and diversification of a specialized antennal lobe morphology in ithomiine butterflies ∗. :3191–3202.

Olsson SB, Linn CE, Michel A, Dambroski HR, Berlocher SH, Feder JL, Roelofs WL. 2006. Receptor expression and sympatric speciation: Unique olfactory receptor neuron responses in F1 hybrid Rhagoletis populations. *J. Exp. Biol.* 209:3729–3741.

Osmanski AB, Paulat NS, Korstian J, Grimshaw JR, Halsey M, Sullivan KAM, Moreno-Santillán DD, Crookshanks C, Roberts J, Garcia C, et al. 2023. Insights into mammalian TE diversity through the curation of 248 genome assemblies. *Science (80-.).* 380:eabn1430.

Paten B, Earl D, Nguyen N, Diekhans M, Zerbino D, Haussler D. 2011. Cactus: Algorithms for genome multiple sequence alignment. *Genome Res.* 21:1512–1528.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33:290–295.

Platt RN, Blanco-Berdugo L, Ray DA. 2016. Accurate transposable element annotation is vital when analyzing new genome assemblies. *Genome Biol. Evol.* 8:403–410.

Pliske TE. 1975a. Courtship behavior and use of chemical communication by males of certain species of ithomiine butterflies (Nymphalidae: Lepidoptera). *Ann. Entomol. Soc. Am.* 68:935–942.

Pliske TE. 1975b. Attraction of Lepidoptera to plants containing pyrrolizidine alkaloids. *Environ. Entomol.* 4:455–473.

Prakash A. 2020. Doublesex Mediates the Development of Sex-Specific Pheromone Organs in Bicyclus Butterflies via Multiple Mechanisms. 37:1694–1707.

Price AL, Jones NC, Pevzner PA. 2005. De novo identification of repeat families in large genomes. *Bioinformatics* 21 Suppl 1:i351-8.

Prieto-Godino LL, Rytz R, Cruchet S, Bargeton B, Abuin L, Silbering AF, Ruta V, Dal Peraro M, Benton R. 2017. Evolution of Acid-Sensing Olfactory Circuits in Drosophilids. *Neuron* 93:661-676.e6.

Prunier FP-, Persyn E, Legeai F, Mcclure M, Meslin C, Robin S, Alves- S, Mohammad A, Blugeon C, Montagné N, et al. 2021. Comparative transcriptome analysis at the onset of speciation in a mimetic butterfly — - The Ithomiini Melinaea marsaeus. :1704–1721.

Rambaut A, Drummond AJ. 2007. Tracer v14, Available from http://beast.bio.ed.ac.uk/Tracer. Available from: http://beast.bio.ed.ac.uk/Tracer

Ranwez V, Harispe S, Delsuc F, Douzery EJP. 2011. MACSE: Multiple alignment of coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6:e22594.

Ray DA, Grimshaw JR, Halsey MK, Korstian JM, Osmanski AB, Sullivan KAM, Wolf KA, Reddy H, Foley N, Stevens RD, et al. 2019. Simultaneous TE Analysis of 19 Heliconiine Butterflies Yields Novel Insights into Rapid TE-Based Genome Diversification and Multiple SINE Births and Deaths. *Genome Biol. Evol.* 11:2162–2177.

Revadi S V, Giannuzzi VA, Rossi V, Hunger GM, Conchou L, Rondoni G, Conti E, Anderson P, Walker WB, Jacquin-joly E, et al. 2021. Stage-specific expression of an odorant receptor underlies olfactory behavioral plasticity in Spodoptera littoralis larvae. :1–18.

Rospars JP. 1983. Invariance and sex-specific variations of the glomerular organization in the antennal lobes of a moth, Mamestra brassicae, and a butterfly, Pieris brassicae. *J. Comp. Neurol.* 220:80–96.

Rospars JP, Hildebrand JG. 1992. Cell & Tissue Anatomical identification of glomeruli in the antennal lobes of the male sphinx moth Manduca sexta. :205–227.

Scalzotto M, Ng R, Cruchet S, Saina M, Armida J, Su CY, Benton R. 2022. Pheromone sensing in Drosophila requires support cell - expressed Osiris 8. *BMC Biol.* [Internet]:1–18. Available from: https://doi.org/10.1186/s12915-022-01425-w

Schmidt HR, Benton R. 2020. Molecular mechanisms of olfactory detection in insects : beyond receptors.

van Schooten B, Meléndez-Rosa J, van Belleghem SM, Jiggins CD, Tan JD, Owen McMillan W, Papa R. 2020. Divergence of chemosensing during the early stages of speciation. *Proc. Natl. Acad. Sci. U. S. A.* 117:16438–16447.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212.

Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:1–11.

Smit A, Hubley R, Green P. 2013. RepeatMasker Open-4.0. 2013-2015 . *http://www.repeatmasker.org* [Internet]. Available from: http://repeatmasker.org

Soo T, Xia R, Zhang H, Jin X, Smith DP. 2014. Lipid flippase modulates olfactory receptor expression and odorant sensitivity in Drosophila.

111:7831–7836.

Sproul J, Hotaling S, Heckenhauer J, Powell A, Marshall D, Larracuente AM, Kelley J, Pauls SU, Frandsen PB. 2023. 600+ insect genomes reveal repetitive element dynamics and highlight biodiversity-scale repeat annotation challenges. *Genome Res.*:gr--277387.

Srikanta SB, Cermakian N. 2021. To Ub or not to Ub : Regulation of circadian clocks by ubiquitination and deubiquitination. :11–30.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: A b initio prediction of alternative transcripts. *Nucleic Acids Res.* 34:435–439.

Sung EJ, Ryuda M, Matsumoto H, Uryu O, Ochiai M, Cook ME, Yi NY, Wang H, Putney JW, Bird GS, et al. 2017. Cytokine signaling through Drosophila Mthl10 ties lifespan to environmental stress. *Proc. Natl. Acad. Sci. U. S. A.*

Task D, Brbic M, Schlegel P, Raji J, Jefferis GSXE, Li H, Menuz K, Potter CJ. 2022. Chemoreceptor co- - expression in Drosophila melanogaster olfactory neurons. :1–70.

Thomas J, Phillips CD, Baker RJ, Pritham EJ. 2014. Rolling-Circle Transposons Catalyze Genomic Innovation in a Mammalian Lineage. 6:2595–2610.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* [Internet] 7:562–578. Available from: http://www.nature.com/doifinder/10.1038/nprot.2012.016

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* [Internet] 28:511–515. Available from: http://www.nature.com/doifinder/10.1038/nbt.1621

Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* [Internet] 43:W401-407. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4489233&tool=pmcentrez&rendertype=abstract

Ugrankar R, Bowerman J, Hariri H, Jean S, Collins BM, Henne WM, Ugrankar R, Bowerman J, Hariri H, Chandra M, et al. 2019. Drosophila Snazarus Regulates a Lipid Droplet Population at Plasma Membrane-Droplet Contacts in Article Drosophila Snazarus Regulates a Lipid Droplet Population at Plasma Membrane-Droplet Contacts in Adipocytes. :557–572.

Unbehend M, Kozak GM, Koutroumpa F, Coates BS, Dekker T, Groot AT, Heckel DG, Dopman EB. 2021. European corn borer moths. *Nat. Commun.* [Internet]:1–11. Available from: http://dx.doi.org/10.1038/s41467-021-23026-x

Venturini L, Caim S, Kaithakottil GG, Mapleson DL, Swarbreck D. 2018. Leveraging multiple transcriptome assembly methods for improved gene structure annotation. *Gigascience* 7:1–15.

Vogt RG, Große-Wilde E, Zhou J-JJ. 2015. The Lepidoptera Odorant Binding Protein gene family: Gene gain and loss within the GOBP/PBP complex of moths and butterflies. *Insect Biochem. Mol. Biol.* [Internet] 62:142–153. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0965174815000521

Vosshall LB, Amrein H, Morozov PS, Rzhetsky A, Axel R. 1999. A Spatial Map of Olfactory Receptor Expression in the Drosophila Antenna. 96:725–736.

Walker III WB, Roy A, Anderson P, Schlyter F, Hansson BS, Larsson MC. 2019. Transcriptome Analysis of Gene Families Involved in Chemosensory Function in Spodoptera littoralis (Lepidoptera : Noctuidae). *BMC Genomics* 20:1–20.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2014. RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* [Internet] 32:1–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25540451

Whelan S, Irisarri I, Burki F. 2018. PREQUAL: Detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics* 34:3929–3930.

Wijesekera TP, Saurabh S, Dauwalder B. 2016. Juvenile Hormone Is Required in Adult Males for Drosophila Courtship. :1–11.

Williams AT, Verhulst EC, Haverkamp A. 2022. A unique sense of smell: development and evolution of a sexually dimorphic antennal lobe – a review. *Entomol. Exp. Appl.* 170:303–318.

Wu N, Evans E, Schooten B Van, Mele J, Ortiz Y, Soto-navarro SMP, Belleghem SM Van, Counterman BA, Papa R, Zhang W. 2022. Widespread Gene Expression Divergence in Butterfly Sensory Tissues Plays a Fundamental Role During Reproductive Isolation and Speciation. 39.

Xiao Q, Komori H, Lee C. 2012. klumpfuss distinguishes stem cells from progenitor cells during asymmetric neuroblast division. 2680:2670–

2680.

Yan H, Jafari S, Pask G, Zhou X, Reinberg D, Desplan C. 2020. Evolution , developmental expression and function of odorant receptors in insects.

Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.*

Yelshanskaya M V., Li M, Sobolevsky AI. 2014. Structure of an agonist-bound ionotropic glutamate receptor. *Science* [Internet] 345:1070–1074. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1256508

Yeo S, Coombe L, Warren RL, Chu J, Birol I. 2018. ARCS: Scaffolding genome drafts with linked reads. *Bioinformatics* 34:725–731.

Yin NN, Nuo SM, Xiao HY, Zhao YJ, Zhu JY, Liu NY. 2021. The ionotropic receptor gene family in Lepidoptera and Trichoptera: Annotation, evolutionary and functional perspectives. *Genomics* 113:601–612.

Zaffran S, Küchler A, Lee H, Frasch M. 2001. biniou (FoxF), a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in Drosophila. :2900–2915.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* [Internet] 19:15–30. Available from: http://dx.doi.org/10.1186/s12859-018-2129-y

Zheng GXY, Lau BT, Schnall-Levin M, Jarosz M, Bell JM, Hindson CM, Kyriazopoulou-Panagiotopoulou S, Masquelier DA, Merrill L, Terry JM, et al. 2016. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat. Biotechnol.* 34:303–311.

Figure 1. Phylogeny of Ithomiini Butterflies and Antennal Transcriptomic Diversity. a) *Dated phylogeny of Ithomiini butterflies in the context of other Nymphalids. The Ithomiini stem is dated between 24 and 42 Mya, overlapping with the Eocene–Oligocene transition (EOT), a period of global cooling, before the orogenesis of the Andes.* ***b****) Gene reconciliation tree topology (Astral III) with branch lengths corresponding to the coalescent units (CUs). Short branches are proxy of incomplete lineage sorting (ILS). Of note are the very short branches at the base of ithomiini species, which indicate rapid speciation, which can also be seen by the overlapping confidence intervals (CIs) in the date phylogeny (a).* c*) Principal component analysis (PCA) of the read counts of single-copy orthologous genes (scOGs), showing very distinct and diverse expression profiles of the antennal transcriptomics of the three Ithomiini butterflies studied. In colour female (blue) and male (orange) samples.*

Figure 2. Phylogeny of Odorant Receptor Gene Family. *The maximum likelihood phylogeny of ORs, the clear blue circle size indicates where the bootstrap is higher than 0.80 and are proportional to their values; 1) branch colours are associated to different species, while dashed line to lineages used as references, while sequences annotated in this study are highlighted in the first inner circle. More specifically okra and reddish colours highlight Danainae species, while green and blue members of the other Nymphalidae ; 2) different shades of green are correlated with the expression levels in the six samples, three males (M1-3) and females (F1-3); 3) Diagram showing the length of the annotated protein and the region that is occupied by the conserved domain (red). Black arrows indicate C-terminus and/or N-terminus that are missing from the protein; 4) fold-change (log_2 transformed) of the differentially expressed genes (DEGs), also indicated by the shape (hexagon for female-biased and star for male-biased). The protein structure in the centre depicts the general shape of a tetramer of a typical odorant receptor.*

Figure 3. Phylogeny of Ionotropic (Glutammate) Receptor Gene Family. *The maximum likelihood phylogeny of IRs and IGluRs, the clear blue circle size indicates where the bootstrap is higher than 0.80 and are proportional to their values; 1) branch colours are associated to different species, while dashed line to lineages used as references, while sequences annotated in this study are highlighted in the first inner circle. More specifically okra and reddish colours highlight Danainae species, while green and blue members of the other Nymphalidae 2) different shades of green are correlated with the expression levels in the six samples, three males (M1-3) and females (F1-3); 3) Diagram showing the length of the annotated protein and the region that is occupied by the conserved domains (shades of red). Black arrows indicate C-terminus and/or N-terminus that are missing from the protein; 4) fold-change (log_2 transformed) of the differentially expressed genes (DEGs), also indicated by the shape (hexagon for female-biased and star for male-biased). The protein structure in the centre depicts the general shape of a tetramer of a typical ionotropic receptor.*

Figure 4. Phylogeny of Odorant Binding Protein Gene Family. *The maximum likelihood phylogeny of OBPs, the clear blue circle sizes indicate where the bootstrap is higher than 0.80 and are proportional to their values; 1) branch colours are associated to different species, while dashed line to lineages used as references, while sequences annotated in this study are highlighted in the first inner circle. More specifically okra and reddish colours highlight Danainae species, while green and blue members of the other Nymphalidae 2) different shades of green are correlated with the expression levels in the six samples, three males (M1-3) and females (F1-3); 3) Diagram showing the length of the annotated protein and the region that is occupied by the conserved domains (shades of red). Black arrows indicate C-terminus and/or N-terminus that are missing from the protein; 4) fold-change (log_2 transformed) of the differentially expressed genes (DEGs), also indicated by the shape (hexagon for female-biased and star for male-biased). The protein structure in the centre depicts the general shape of a homo-dimer of a typical odorant binding protein.*

Figure 5. Antennal Differential Gene Expression in the Three Species of Ithomiini. *a) Volcano plots showing the relation between fold-change (FC) (log_2 transformed) and their Bayes Factors (BFs) of the antennal gene expression. In grey, genes with a FC below |1| and with a BF below the threshold of 0.95; in black with a FC > |1| but not significant; in blue and red genes found significantly differential expressed (DEGs; BF > 0.95), with a FC lower or higher than one, respectively. For each plot, female-biased and male-biased genes are on the left and right, respectively, for* Me. confusa*,* Mc. polymnia *and* T. harmonia*. d) Venn diagram showing the DEGs in common between species. The number in the main circles are the single-copy orthologous genes used for the comparisons. For each comparison, genes are grouped in boxes according to the concordance of expression (coloured boxes) or not (grey boxes).*

Figure 6. Chemosensory Gene Expression and the Distribution of Selective Shift. *For each chemosensory gene family, we show the distribution of their expression level (TPM) in each of the three genes (upper half) and the distribution of* k*, the relaxation/intensification index for each coding sequence in al the six Ithomiini species included in this study (bottom section). The horizontal dashed lines indicate the bottom threshold level (adjusted* P *value < 0.05), while the vertical one the shift between the relaxation (log_2(k) < 0 and intensification (log_2(k) > 0). The colours of the circles indicate the species, while their size is relative to the d*N*/d*S *(ω) for that particular locus.*