DR. RAFAEL AUGUSTO HOMEM (Orcid ID : 0000-0001-9649-1825)
DR. CHRISTOPHER M JONES (Orcid ID : 0000-0002-6504-6224)

Article type : Original Article

Title page

Odorant binding proteins promote flight activity in the migratory insect, *Helicoverpa armigera*

Short title

Odorant binding proteins and insect flight

Shang Wang[a,b], Melissa Minter[b,c], Rafael A. Homem[b], Louise V. Michaelson[d], Herbert Venthur[e,f], Ka S. Lim[b], Amy Withers[g], Jinghui Xi[a], Christopher M. Jones[b,h,1], Jing-Jiang Zhou[a,b,1]

[a]College of Plant Sciences, Jilin University, Changchun, Jilin Province 130062, P.R. China
[b]Biointeractions and Crop Protection, Rothamsted Research, Harpenden, Herts, AL5 2JQ, UK
[c]Department of Biology, University of York, Heslington Way, York, UK
[d]Plant Sciences, Rothamsted Research, Harpenden, Herts; AL5 2JQ, UK
[e]Laboratorio de Química Ecológica, Departamento de Ciencias Químicas y Recursos Naturales, Universidad de La Frontera, Casilla 14-D, Temuco, Chile
[f]Centro de Investigación Biotecnológica Aplicada al Medio Ambiente, CIBAMA, Universidad de La Frontera, Temuco, Chile
[g]Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[h]Vector Biology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK

[1]Corresponding authors:

Christopher M. Jones, Vector Biology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK, +447792965202, chris.jones@lstmed.ac.uk; ORCID 0000-0002-6504-6224

Jing-Jiang Zhou, College of Plant Sciences, Jilin University, Changchun, Jilin Province 130062, P.R. China, jjzhouchina@163.com; ORCID 0000-0002-8203-601X

Keywords: Insect migration, Helicoverpa, odorant binding proteins

Abstract

Migratory insects are capable of actively sustaining powered flight for several hours. This extraordinary phenomenon requires a highly efficient transport system to cope with the energetic demands placed on the flight muscles. Here, we provide evidence that the role of the hydrophobic ligand binding of odorant binding proteins (OBPs) extends beyond their typical function in the olfactory system to support insect flight activity via lipid-interactions. Transcriptomic and candidate gene analyses show that two phylogenetically clustered OBPs (OBP3/OBP6) are consistently over-expressed in adult moths of the migrant Old-World bollworm, *Helicoverpa armigera*, displaying sustained flight performance in flight activity bioassays. Tissue-specific over-expression of *OBP6* was observed in the antennae, head and thorax in long-fliers of *H. armigera.* Transgenic *Drosophila* flies over-expressing a *H. armigera* transcript of OBP6 (HarmOBP6) in the flight muscle attained higher flight speeds on a modified tethered flight system. Quantification of lipid molecules using mass-spectrometry showed a depletion of triacylglyerol and phospholipids in flown moths. Protein homology models built from the crystal structure of a fatty acid carrier protein identified the binding site of OBP3 and OBP6 for hydrophobic ligand binding with both proteins exhibiting a stronger average binding affinity with TAG and phospholipids compared with other groups of ligands. We propose that HarmOBP3 and HarmOBP6 contribute to the flight capacity of a globally invasive and highly migratory noctuid moth, and in doing so, extend the function of this group of proteins beyond their typical role as chemosensory proteins in insects.

Introduction

Insect flight is one of the most energetically demanding processes in the animal kingdom. Long-distance insect migrants can actively sustain periods of flight for several hours. To achieve these remarkable feats of endurance, migratory insects have evolved a suite of morphological, sensory and physiologically traits that form part of an inherited 'migratory syndrome' (Dingle, 2014; Liedvogel, Akesson, & Bensch, 2011; Roff & Fairbairn, 2007). Comparative genomics and quantitative trait analyses reveal considerable genetic variation for single migratory traits but the associated molecular genetic mechanisms and biochemical pathways remain poorly understood.

The vital role of chemical cues in host location and oviposition (Bruce & Pickett, 2011; Hansson & Stensmyr, 2011; Mescher & De Moraes, 2015) means that the involvement of a sophisticated olfactory system in migration and flight ability is an attractive proposition (Getahun et al., 2016; McCormick et al., 2017). For example, the odorant receptor family (OR), central to the olfactory system of pterygotes, emerged prior to the evolution of winged flight in insects as an adaptation to terrestrial life (Brand et al., 2018). New evidence suggests that OR-based signal transduction in *Drosophila* is necessary for efficient odour localization in flight (Getahun et al., 2016). Our recent transcriptomic work (RNAseq) in the Old World bollworm, *Helicoverpa armigera*, show that specific odorant binding proteins (OBPs), OBP3 and OBP6, are highly and consistently over-expressed in moths displaying sustained flight activity (Jones et al., 2015). This suggests that OBPs have a direct or indirect role in supporting insect flight and their function extends beyond their part in host-seeking and mating behaviour.

Insect OBPs are small, water soluble extracellular transporter proteins (13–16 kDa) (Lartigue et al., 2002; Tegoni, Campanacci, & Cambillau, 2004; Zhou, 2010), and possess extreme diversity between species with as little as 8% amino acid conservation (Pelosi, Zhou, Ban, & Calvello, 2006; Zhou, He, Pickett, & Field, 2008). They are generally thought to contribute to the sensitivity of the olfactory system by participating in the binding, solubilization and transportation of hydrophobic ligands through the sensillum lymph of the antennae (Grosse-Wilde, Svatos, & Krieger, 2006; Leal, 2013; Tsuchihara et al., 2005) and protecting odours from enzymatic degradation (Chertemps et al., 2012; Gomez-Diaz, Reina, Cambillau, & Benton, 2013). Some OBPs, however, are found in non-chemosensory tissues and may participate in other physiological processes (Graham et al., 2001; Guo et al., 2011; Missbach, Vogel, Hansson, & Grosse-Wilde, 2015; Pelosi et al., 2006). The *Drosophila* OBP28a is not required for odorant transport and signal

transduction, implying a different function altogether (Larter, Sun, & Carlson, 2016). The homologues of OBP6 and OBP3 in *H. armigera* are highly expressed in non-olfaction tissues in other noctuid moths, *Agrotis ipsilon* and *Helicoverpa assulta* (Gu et al., 2014; Li et al., 2015). In arthropods, OBPs are found exclusively in insects (Pelosi, Iovinella, Felicioli, & Dani, 2014) and comparative genomics suggests that the evolution of this protein family provided a mechanism to mediate the transport of hydrophobic chemical signals present in a terrestrial environment (Vieira & Rozas, 2011).

The Noctuidae family of moths possess some of the most important and polyphagous agricultural insect pests globally. A key characteristic that makes them such devastating pests is their ability to spread hundreds of kilometres in response to deteriorating local conditions. This exacerbates their potential to invade new territories as observed with the current fall armyworm (*Spodoptera frugiperda*) which has spread eastwards into the Asian continent and the rapid expansion of *H. armigera* in the Americas following its recent incursion (Fitt, 1989; Jones, Parry, Tay, Reynolds, & Chapman, 2019). Adult moths from both species can climb to high-altitudes and sustain active flight for several hours (Chapman et al., 2010). This requires an enormous amount of fuel consumption, metabolism and intra-cellular transport to the flight muscles. Given the well-established hydrophobic binding capacity of OBPs and their over-expression in *H. armigera*, it is possible that this group of proteins act as lipid transport carriers in *H. armigera* – the main flight fuel of migratory insects (Van der Horst & Ryan, 2012).

In the present study, we use a combination of behavioural, molecular, transgenic and protein modelling approaches to (i) determine the tissue-specificity of two OBPs consistently expressed in *H. armigera* adult moths demonstrating sustained flight activity, (ii) show that the transgenic overexpression of one of these OBPs leads to enhanced flight performance in *Drosophila*, (iii) identify the primary lipids depleted in *H. armigera* following flight and (iv) the key residues responsible for lipid binding. Overall, our findings provide evidence that a subset of OBPs are responsible for binding key lipids commonly used by insect migrants and this relationship promotes insect flight.

Materials and Methods

H. armigera strains

The adult *H. armigera* used in this study originated from a long-term laboratory strain, *Bayer* (courtesy of the Max Planck Institute, Jena, Germany) and a wild-caught population from Spain (courtesy of the University of Valencia). The moths used in the RNA-seq were from a colony established from northern Greece. Insects were reared under a constant light regime of L:D 14:10 at 26 ± 1ºC in the insectaries of Rothamsted Research and the flight mill trials were conducted under the same conditions. Larvae were reared individually in 37-ml clear plastic pots containing a chickpea artificial diet and allowed to pupate before transfer to clean pots filled with vermiculite. Adult emergence was checked daily and any emerged individuals were set aside for flight mill trials or for rearing onto the next generation.

Flight propensity of H. armigera measured by tethered flight mill

A series of flight mill experiments were conducted to determine the effects of over-expression of candidate genes associated with migration or flight in *H. armigera* displaying contrasting flight abilities. A detailed description of the flight mill system is explained elsewhere (Jones et al., 2015; Minter et al., 2018). Insects from *Bayer* and *Spain* strains were reared through at least one generation in the insectary prior to flight mill trials and each strain flown in independent experiments. Adult moths assigned to flight mill trials were placed in 4-10ºC to facilitate the attachment of ~60 mg pins to the thorax with adhesive glue. Each moth was provided with 10% honey water solution *ad libum* prior to flight. Moths were attached randomly to one of 16 flight mills via a pin and allowed to rest on a paper platform until the first flight was initiated by the insect. All flights took place between 1900 and 0900 under a 10-hour dark cycle from 2000 to 0600. At approximately 0900 the next morning, individuals were taken off the mills and placed into individual pots for weighing. Any dead, unhealthy (e.g. broken wings or damage through improper handling) or escaped individuals were recorded and excluded from further analyses. All individuals were snap-frozen or placed in RNAlater within two hours and stored at -80^{o}C for downstream molecular analysis.

The aggregated response variables were calculated for all individuals. We considered four response variables as being important discriminants of 'strong' and 'weak' fliers based on previous experiments; Total Distance Flown (m), Average Speed Flown (m/s), Maximum Speed attained (m/s) and number of bouts. Seven individuals from each strain were selected for RT-qPCR of candidate genes from the ends of the flight activity distribution based on total distance and number of flight bouts.

Tissue-specific candidate gene expression profiling in H. armigera flown on the flight mills

Initially, we determined the expression of 20 candidate genes from the head and thorax of 28 individual moths flown on the mills. The head (including the antennae) and thorax were removed using dissection instruments and placed in separate Eppendorf tubes with lysis buffer. The samples were homogenised using pellet pestles (Sigma-Aldrich). RNA was extracted using ISOLATE II RNA Mini Kit (Bioline) and RNA was eluted in RNase-free water. cDNA was synthesised from 230 ng total RNA using SuperScript® IV Reverse Transcriptase (Invitrogen) and Oligo$(dT)_{20}$ (Invitrogen).

Twenty candidate genes were screened for gene expression levels. qPCR primers were screened over a five-fold serial dilution of a cDNA sample ($1/10^{th}$ to $1/6000^{th}$) and the primer efficiency calculated. qPCR reactions were completed on the RotorGene 6000 (Qiagen) with conditions of 95°C for 2 minutes, followed by 40 cycles of 95°C 10 seconds, 57°C 15 seconds and 72°C 20 seconds, followed by a melt curve analysis. Each reaction contained 10 µl of SYBR Green JumpStart Taq ready mix (Sigma-Aldrich), 300nM of each primer and 5 µl of cDNA $1/50^{th}$ dilution. The control genes *β-actin* and *elongation factor 1-α* were used for normalization (Wang, Dong, Desneux, & Niu, 2013; Yan et al., 2013) and all reactions were run in duplicate. Ct values were adjusted for primer pair efficiency. Expression levels were compared using a two-sided t-test on the dCt values ($p < 0.05$) and are presented as $\log_{10}$ fold-change using ddCt (Schmittgen & Livak, 2008). The RNA-seq was performed on moths flown and not-flown (N = 4 per group) according to previously described methods (Jones et al., 2015). All genes were considered significantly expressed at a False Discovery Rate of $p < 0.1$.

Following the identification of strong OBP expression profiles from the twenty candidate genes we determined the tissue-specific expression of OBP6 in the antennae, heads, thoraces, abdomens, legs and wings of *H. armigera* flown on the flight mills. Tissues were dissected from 18 adults and promptly immersed in liquid nitrogen and stored at -80ºC. RNA was extracted using RNA-Solv reagent (Omega) following the manufacturer's protocol. Total RNA was quantified and checked for purity and integrity using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) and gel electrophoresis. HiScript® II Q RT SuperMix for qPCR with gDNA wiper (R223-01, Vazyme, Nanjing China) was used for cDNA synthesis.

For tissue-specific expression profiling, RT-qPCR primer pairs were designed using Primer 5 software (Untergasser et al., 2012) and the same control genes used as above. mRNA levels were

measured by RT-qPCR using the ChamQTM SYBR® qPCR Master Mix (Vazyme, Nanjing, China). Each amplification reaction contained 1 µl of synthesized cDNA, 10 µl of 2×ChamQTM SYBR qPCR Master Mix, 0.4 µl of 10µM of forward primer, 0.4 µl of reverse primer, and 8.2 µl water in a 20 µl reaction mix. Reaction were performed on an ABI 7500 Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA) under the following conditions: 30 sec denaturation at 95 °C and 40 cycles of 95 °C for 10 sec, 60 °C for 30 sec followed by a melt curve for specificity analysis. Fold-changes were calculated from the mean of three biological replicates with the ddCt method and using the abdomen as the calibrator. The relative expression levels were compared using the dCt values ($p < 0.05$) as described above.

Quantitative triacylglycerol analysis

Total lipids were extracted from moth tissue ground in liquid nitrogen (Usher et al., 2017). The molecular species of triacylglycerols were analysed by electrospray ionisation triple quadrupole mass spectrometry (ESI-MS) (using API 4000 QTRAP; Applied Biosystems). Triacylglycerols (TAGs) are defined by the presence of one acyl fragment and the mass/charge of the ion formed from the intact lipid (neutral loss profiling) (Krank, Murphy, Barkley, Duchoslav, & McAnoy, 2007). This allowed identification of one triacylglycerol acyl species and the total acyl carbons and total number of acyl double bonds in the other two chains. The procedure does not allow identification of the other two fatty acids individually nor the positions (sn-1, sn-2, or sn-3) that individual acyl chains occupy on the glycerol. Triacylglycerol was quantified after background subtraction, smoothing, integration, isotope deconvolution and comparison of sample peaks with those of the internal standard (using LipidView™, AB-Sciex, Framingham, MA, USA). The profiling samples were prepared by combining 10 µL of the total lipid extract with 990 µL of isopropanol/methanol/50 mM ammonium acetate/dichloromethane (4:3:2:1). Samples were infused at 15 µL min^{-1} with an autosampler (LC mini PAL, CTC Analytics, Switzerland). The scan speed was 100 µ s^{-1}. The collision energy, with nitrogen in the collision cell set to +25 V; declustering potential +100 V; entrance potential 14 V; and exit potential +14 V. Sixty continuum scans are averaged in the multiple channel analyser mode. For product ion analysis, the first quadrupole mass spectrometer (Q1) was set to select the triacylglycerol mass and Q3 for the detection of fragments fragmented by collision induced dissociation. The mass spectral responses of various triacylglycerol species are variable, owing to differential ionisation of individual molecular triacylglycerol species. For all analyses gas pressure was set on 'low', and the mass

analysers were adjusted to a resolution of 0.7 μ full width height. The source temperature was set to 100 °C; the interface heater deployed, +5.5 kV applied to the electrospray capillary; the curtain gas set at 20 (arbitrary units; and the two ion source gases set at 45 (arbitrary units). The data are normalised to the internal standard Tri15:0 (Sigma Aldrich) and further normalized to the weight of the initial sample.

Quantitative Phospholipid Analysis

Quantitative analyses to measure, phospholipids (PL) (phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylglycerol (PG), and phosphatidylserine (PS) was carried out using electrospray ionization tandem triple-quadrupole mass spectrometry (API 4000 QTRAP; Applied Biosystems; ESI-MS/MS). The lipid extracts are diluted and resuspended in CHCl3/MeOH/300 mM ammonium acetate in water, 300/665/35. Internal standards obtained and quantified as previously described (Devaiah et al., 2006). Samples were directly infused at 15 μL/min with an autosampler (HTS-xt PAL, CTC-PAL Analytics AG, Switzerland). Data acquisition and acyl group identification of the polar lipids was performed with modifications (Ruiz-Lopez, Haslam, Napier, & Sayanova, 2014). The internal standards are supplied by Avanti (Alabaster, AL, USA), incorporated as, 0.085 nmol of di24:1-PC, 0.08 nmol of di14:0-PE, 0.08 nmol of di18:0-PI, and 0.032 nmol of di18:0-PS and 0.08 nmol of di14:0-PG.

The molecular species of polar lipids are defined on the basis of the presence of a head-group fragment and the mass/charge of the intact lipid ion formed by ESI. However, tandem ESI-MS/MS precursor and product ion scanning, based on head group fragment, do not determine the individual fatty acyl species. Instead, polar lipids are identified at the level of class, total acyl carbons and total number of acyl carbon–carbon double bonds.

The data was processed using the program Lipid View Software (AB-Sciex, Framingham, MA, USA) where isotope corrections are applied. The peak area of each lipid was normalized to the internal standard and further normalized to the weight of the initial sample. A parametric two-sided t-test was used to compare lipid content between flown and not flown moths (N = 4-5 moths per group).

Phylogenetic analysis of H. armigera OBPs

N-terminal signal peptides of OBPs were predicted by Signal IP 4.0 (http://www.cbs.dtu.dk/services/SignalP/). Alignment of amino acid sequences (without signal

peptides) was performed by MAFFT (https://www.ebi.ac.uk/Tools/msa/mafft/). The phylogenetic trees of OBPs were constructed using MEGA6 software by the maximum-likelihood method at bootstrap 1000 with the p-distance model (Gong, Zhang, Zhao, Xia, & Xiang, 2009).

Development of a novel flight mill for Drosophila melanogaster and other small insects

We designed a new set of flight mills to accommodate smaller insects to examine the flight ability of wild type and transgenic *Drosophila* flies (Fig. S1). These flight mills are similar in structural design to the mills used in the *H. armigera* experiments, comprising of a flight arm and rotational disc to allow flies to move around an axis by means of a very low friction magnetic bearing (Fig. S1).

As part of the study we developed a robust standard operating procedure for tethering *Drosophila*. Briefly, an individual fly was lightly anesthetized with CO_2 and transferred to a custom-made paper groove which had been made to allow accurate positioning of an anesthetized fly (Fig. S1). The paper groove was placed on the platform with CO_2 passing through the groove bottom. When the flies were under CO_2 anaesthesia, the tip of a small flight mill arm (a 5-cm-long, 0.2-mm-diameter copper wire) was tethered onto the dorsal side of the anesthetized fly's thorax with Super Glue under a stereomicroscope (Olympus SZ40). Individual flies and the small flight mill arms were gently handled with either a small brush or jeweler's vacuum tweezers. Once the glue was dry and hard, the tethered flies were moved to the experimental chamber, fed with sucrose solution from a small piece of filter paper and allowed to rest in the recording chambers to adapt to the experimental environment overnight prior to data collection. At 10:00 the following day, filter papers were removed from the recording chambers and data collection was started using the same software as the larger mills. Experiments were run till approximately 14:30 to ensure each mill had run for at least 3 hours. Any flies which looked damaged, unhealthy or had escaped from the flight arm were disregarded from further analyses.

Generation of transgenic Drosophila expressing HarmOBP6

All *Drosophila* strains were maintained on standard food (Bloomington formulation) at 24^{o}C and 65% RH on a 12/12-hour light/dark cycle. Proteinase K treatment and phenol/chloroform extraction were used to isolate genomic DNA (gDNA) from adult *D. melanogaster* flies for use in PCR.

HarmOBP6 (B5X24_HaOG200803 with the addition of a stop codon) was codon optimised for expression in *D. melanogaster* and synthetized by GeneArt™ (ThermoFisher Scientific). The codon-optimised sequence was transferred from the sub-cloning plasmid pMA (GeneArt™) to the attB-carrying plasmid pUAST (pUASTattB_EF362409) using *EcoRI* and *XhoI* restriction sites. The pUAST-*Harm-OBP6* construct was microinjected into syncytial blastoderm embryos of a an integration strain (y w M(eGFP, vas-int, dmRFP)ZH-2A; Pattp40) (Dundas et al., 2006) that carries an *attP* docking site on the second chromosome (*attP40*) and the *phiC31* integrase gene under the control of a germline-specific (*vasa*) promoter on the X chromosome. This strain was sourced from the Fly Facility-University of Cambridge. The GAL4 strains (w[1118]; Pw[+mW.hs]=GawBDJ757) was sourced from Bloomington Drosophila Stock Centre (BDSC-8184). Microinjections were performed in house using an inverted microscope (eclipse TieU Nikon, Japan) equipped with a 10x/0.25 lens, 10x/22 eyepiece and fluorescence illumination. Injection mixtures comprised of 0.5x phosphate buffer (pH 6.8, 0.05 mM sodium phosphate, 2.5 mM KCl), 300 ng/µl of the pUAST-*Harm-OBP6* construct and 200 mg L^{-1} fluorescein sodium salt and delivered by a FemtoJet express micro injector (Eppendorf, Hamburg, Germany) controlled by a motorised micromanipulator TransferMan NK2 (Eppendorf, Hamburg, Germany). Injection needles were prepared following (Miller, Holtzman, & Kaufman, 2002).

Microinjection survivors were back-crossed and the F1 progeny was screened for the presence of the *white* marker gene (orange eye phenotype). Positive flies were inter-crossed to generate homozygous flies (red eyes) which were selected to establish the final strain. Control flies carrying an empty pUAST plasmid (*UAS-empty* strain) were generated following the same protocols described above.

Tethered flight of transgenic Drosophila and statistical analysis of flight response variables

Three flight mill experiments were performed to compare the flight ability of transgenic *Drosophila* flies carrying *HarmOBP6* (GAL4>UAS-OBP6 line) with control flies (GAL4>UAS line). In addition, we were also interested in how flight activity changes with the age of the fly. A total of eight flies were flown simultaneously per run with each trial consisting of a mixture of *HarmOBP* and control flies.

1. Experiment 1: flies generated from crosses between male *UAS-OBP6* (*UAS-empty* for controls) and female *muscle-GAL4* strains. *GAL4>UAS-OBP6* virgin females (N = 28)

were flown on the mills alongside *GAL4>UAS* virgin female control flies (N = 28). The age of the flies in this experiment ranged from 24h to 144h after emergence.

2. Experiment 2: flies were generated from crosses between female *UAS-OBP6* (*UAS-empty* for controls) and male *muscle-GAL4*. Both *GAL4>UAS-OBP6* (N = 23) and *GAL4>UAS* controls (N = 21) female flies were mated prior to the flight mill trials. Flies were either 2d, 6d or 15 days old after emergence.
3. Experiment 3: *GAL4>UAS-OBP6* (N = 43) and *GAL4>UAS* (N = 34) control flies were generated as Experiment 2 but without mating. The age of the flies ranged from 7d to 26d old after emergence.

After preliminary trials, we determined that a one-hour cut-off period was sufficient to measure flight performance with difference between the average speeds attained between 1 hour and 3 hours of flight (example from Experiment 3 in Fig. S2). We were primarily interested in the two response variables, the average speed flown (AVGSP) (m/s) and maximum speed attained (MAXSP) (m/s). We hypothesise that the average or maximum speed of flight is a much more useful metric to distinguish the flight activity of flies such as *Drosophila* since they are not capable of sustaining hours of flight like larger insects (e.g. Lepidoptera). The distribution of each flight parameter was assessed using the *fitdist* package (Delignette-Muller & Dutang, 2015) using QQ plots and goodness of fit statistics. AVGSP and MAXSP were both normally distributed (Fig. S2).

Data was fitted using Generalised Linear Mixed Models (GLMMs) using the package lme4 package in R (Bates, Machler, Bolker, & Walker, 2015). To model AVGSP and MAXSP as a function of the covariates we used a Gaussian linear mixed-effects model. Fixed covariates are *strain* (transgenic or control) and *age* (categorical). An interaction term *strain x age* was included. To incorporate differences between the flight mills on which the individual was flown we include *mill* as a random effect. Best fit GLMMs were created using a backward stepwise approach from the maximally complex model which included the interaction. Explanatory variables were retained in the best fit model according to significance ($p < 0.05$) in Likelihood Ratio Tests (LRT). Model assumptions were verified using residual-fitted plots. Predictions of response variables from each model were made using Least Square Means (LSMs) in the package *lsmeans* (Lenth, 2016) and differences between groups assessed using Tukey *post-hoc* tests.

Homology structure modelling of H. armigera OBPs

The amino acid sequences of HarmOBP6 and HarmOBP3 were used as a target while the template was the crystal structure of the blowfly *Phormia regina* OBP56a (PregOBP56a) (PDB code: 5DIC). The pheromone binding protein 1 from the silkmoth *Bonbyx mori* BmorPBP1 (1DQE) was used as the template for HarmPBP1 structure modelling. Five hundred models of each OBP were obtained using MODELLER9.14 (http://salilab.org/modeller) and the best initial model was selected according to the lowest discrete optimized protein energy (DOPE) score provided by the software. The stereochemistry of the best model was assessed using the theoretical validation package ProCheck (Laskowski, Macarthur, Moss, & Thornton, 1993), and the overall structure was visualized using PyMOL software (http://www.pymol.org). Further refinement steps were carried out with NAMD v2.9 (parallel molecular dynamics code for biomolecular system simulation) installed in the high-performance computer (HPC) Lautaro Linux cluster at Centro de Modelación y Computación Científica (CMCC) from Universidad de La Frontera. CHARMM36 force field was used for all the simulations. The selected protein model was solvated with TIP3P water model in a cubic box with a minimum distance of 10 Å between the protein and the edge of the box. Neutralization of the protein-water system was performed by adding Na^+ or Cl^- randomly placed in the box. Likewise, the system was simulated under periodic boundary conditions with a cutoff radius of 12 Å for non-bonded interactions and a time step of 2 fs. Alpha-carbons (Cα) of secondary structures were fixed with a constant force of 1 kcal/mol/Å. A first energy minimization of 10 000 steps was performed followed by heating through short simulations of 1 ps at 50, 100, 150, 200, 250 and 300 K. Long simulations were kept at 300 K and 1 bar pressure in the NTP (referred to a constant number of particles, temperature and pressure) during 50 ns. Root-mean-square deviation (RMSD) trajectory tool was used to calculate the RMSD with reference to the starting structure (Fig. S3). Therefore, when the plotted RMSD showed small fluctuations (~1-1.5 Å), coordinates were analyzed by ProCheck every 100 frames to obtain the best structure (lowest energy). Finally, putative binding site and its volume were calculated by CASTp server (http://sts-fw.bioengr.uic.edu/castp/calculation.php) (Dundas et al., 2006).

Molecular docking

The refined structures of HarmOBP6 and HarmOBP3 were used as the target for molecular docking with AutoDock Vina (Trott & Olson, 2010). Likewise, a refined 3D structure of the pheromone binding protein HarmPBP1 was used as the reference template for the molecular

docking tasks based on its reported function in binding sex pheromones (Dong et al., 2017; Ye et al., 2017). Energy minimization and optimization for the ligands used in this study were performed using MM2 minimization methods in the Chem3D 16.0 Software (Perkin Elmer). For HarmOBP6, polar hydrogens were added using the interface AutoDock Tools, as well as torsional bonds for ligands. A grid box with 26×26×26 points and a default space of 1 Å was prepared via AutoGrid following the predicted binding site by CASTp server. For every docking run, an exhaustiveness of 500 was considered and the best binding modes were selected according to the lowest free binding energy (kcal mol^{-1}). The triacylglycerols and phospholipids were energy minimized following the same protocol for fatty acids and semiochemicals. Considering that AutoDock Vina allows a maximum of 32 rotatable bonds, these compounds and their binding to HarmOBP6 and HarmOBP3 were submitted to DINC server (http://dinc.kavrakilab.org/) (Antunes et al., 2017; Dhanik, McMurray, & Kavraki, 2013). This server was used to dock the lipids into HarmOBP6 and HarmOBP3 binding site following the above grid box parameters and with all rotatable bonds active. DINC allows docking for large molecules based on the AutoDock algorithm and fragmentation processes, for which fragments that show best binding are incrementally expanded by adding atoms of the ligand to it in each of several rounds. Thus, both fully flexible and bound conformations of lipid molecules were extracted and docked again into HarmOBP6, HarmOBP3 and HarmPBP1 (control) using Autodock Vina.

Results and Discussion

Two H. armigera-specific odorant binding proteins are overexpressed in the thorax of moths displaying prolonged flight activity

Adult moths from two colonised strains of *H. armigera* (*Bayer* and *Spain*) were flown overnight on a computerised tethered flight mill system that experimentally quantifies the flight performance of individual insects in the absence of external stimuli (Minter et al., 2018). Previous flight mill studies with noctuid moths show an inverse relationship between the total distance flown and the number of individual flight bursts to discriminate those insects engaging in prolonged or more appetitive behaviour (Jones et al., 2015). We used this relationship to assign individual moths into two distinct flight activity groups, 'short-distance' (SD) or 'long-distance' (LD), for downstream gene expression analyses (Fig. *1A*).

We undertook a candidate gene approach to determine the differential expression of twenty genes in the two strains of *H. armigera* flown on the flight mills. As a baseline control, and to validate

some of our previously detected candidate genes from whole transcriptome studies (Jones et al., 2015), an RNA-seq of moths flown and not flown on the flight mills was performed. Eight of our 20 candidate genes were significantly up-regulated in the flown group with OBP6 showing the highest and most consistent level of up-regulation (Fig. S4). Each gene has a reported role in insect migration or sustained flight activity including those involved in circadian and photoreceptor processes (Reppert, Guerra, & Merlin, 2016), lipid metabolism (Arrese & Soulages, 2010), odorant binding proteins (Jones et al., 2015), flight muscle structure (Zhan et al., 2014) and the metabolism of proline and phenylalaine/tryptophan (Arrese & Soulages, 2010; Rio, Attardo, & Weiss, 2016) (Fig. *1B*).

In the *Bayer* strain four genes were significantly over-expressed in the thorax of the LD moths, all of which were up-regulated exclusively in the thorax and not the head (Fig. *1B*, Table S1). These four genes encode two OBPs (*OBP3* and *OBP6*), the *protein henna-like isoform X3* and a *fatty-acid synthase-like* gene. In the thorax of individuals from the strain *Spain*, three genes were significantly over-expressed in the LD group; *myofilin*, *OBP3* and *protein henna-like isoform X3* ; three genes were significantly over expressed in the SD group, *collagen alpha subunit-1(IV)*, *cry-1* and *phospholipase A2-like*. Two genes were differentially expressed in the head of *H. armigera* individuals from the *Spain* strain (although the magnitude of this expression was small) (Fig. *1B*).

Following the detection of OBP over-expression both in this study and from transcriptome profile analysis (RNA-seq (Fig. *S4*); (Jones et al., 2015)), we showed that the relative expression levels of *OBP3* and *OBP6* in individual *H. armigera* displayed a significant positive correlation with flight performance which was strongest in the thorax (*HarmOBP3*: head: $R = 0.49$, $p = 0.006$, thorax: $R = 0.81$, $p < 0.001$; *HarmOBP6*: head: $R = 0.31$, $p = 0.18$, thorax: $R = 0.65$, $p = 0.002$) (Fig. *1C*). Furthermore, we quantified the expression of *HarmOBP6* in the antenna, head, thorax, abdomen, leg and wing of SD and LD moths. *HarmOBP6* was significantly over-expressed in the antennae ($p = 0.016$) , thoraces ($p = 0.009$) and wings ($p = 0.05$) and this expression was significantly upregulated in LD moths compared with those in the SD group (Fig. *1D*, Table S2).

The simple phenotypic comparisons of SD and LD insects presented here provide a measurement of flight performance in terms of the raw physiological capacity to fly. We recognise that a full spectrum of flight behaviours exists and that these are controlled by intricate internal and external processes. For example, the migratory flight behaviour of the Monarch butterfly (*Danaus plexippus*) is controlled in response to environmental changes (temperature, photoperiod) via

internal genetic, and possibly epigenetic, cascades (Merlin & Liedvogel, 2019). Here we focus on raw flight capacity and use our expression profiling to speculate on the functional role of OBPs in insect flight.

Phylogenetic analysis of the odorant binding proteins implicated in H. armigera flight

An alignment of the protein sequences of HarmOBP3 (Acc No: AEB54582), and HarmOBP6 (Acc No: AEB54587) is provided in Fig. *S5*. Based on the sequence alignment, HarmOBP6 belongs to the classic OBP subgroup, which contains typical characteristic sequence features of six conserved cysteine residues and the classic insect OBP motif: C_1-X_{15-39}-C_2-X_3-C_3-X_{21-44}-C_4-X_{7-12}-C_5-X_8-C_6 (Fig. *S5*.) (Zhou et al., 2008). The phylogenetic analysis shows that HarmOBP3 and HarmOBP6 are clustered into the same branch with a 100% bootstrapping support between OBP3 and OBP6, indicating that they share a high homologous amino acid sequence similarity and most likely a similar function (Fig. *S6*). There is an 81% amino acid identity between OBP3 and OBP6. HarmOBP6 is also closely clustered with other *Helicoverpa* OBPs such as *Helicoverpa assulta*, HassOBP6 (Acc No: AEX07270) and *Heliothis virescens*, HvirOBP0136 (Acc No: ACX53819) (Fig. *S6*).

Transgenic Drosophila expressing OBP6 in the flight muscle attain higher speeds on a novel flight mill system

To functionally validate the role of OBPs in flight activity we generated a transgenic *D. melanogaster* strain that over-expresses *HarmOBP6* in muscle cells and assessed the performance of these flies on a newly designed flight mill system for small dipterans (Fig. *S1*). We chose OBP6 based on its magnitude of expression in a Chinese strain of *H. armigera* previously reported as well as the flown/not flown comparison (Fig. *S4*) but postulate that the high conservation between the protein sequences of OBP6 and OBP3 (see phylogenetic analysis above) would lead to similar results had we chosen OBP3. Transgenic strains were generated using the φC31 integration system (Bischof, Maeda, Hediger, Karch, & Basler, 2007). Genomic integration of *HarmOBP6* in generated transgenic flies, hereafter referred to as *UAS-OBP6* strain, was confirmed by PCR and sequencing (Fig. *2A*). The *GAL4/UAS* expression system (Brand & Perrimon, 1993) was employed to induce the expression of *HarmOBP6* in muscle cells by using a muscle-specific GAL4 driver strain (Seroude, Brummel, Kapahi, & Benzer, 2002) (referred to as *muscle-GAL4* strain). The over-expression of *HarmOBP6* in transgenic flies generated from the cross between the *UAS-OBP6* and *muscle-GAL4* strains (GAL4>UAS-OBP6 flies), was confirmed by RT-PCR and RT-

qPCR (Fig. *2A*). The expression of *HarmOBP6* increased by more than 15 times in *GAL4>UAS-OBP6* flies when compared to parental *muscle-GAL4* and *UAS-OBP6* (Fig. *2A*).

We performed a series of flight mill experiments with three separate experimental trials. First, we compared the one-hour flight activity of *Gal4>UAS-OBP6* transgenic and *Gal4>UAS* control flies. These flies are genetically identical where the only difference was the absence of *HarmOBP6* in the controls. The average and maximum speeds (m/s) attained during the one-hour of flight activity were analysed using Generalised Linear Mixed Models (GLMMs) as a function of the covariates strain and age (Table 1). There was no difference in the average or maximum speed between *Gal4>UAS-OBP6* and *Gal4>UAS* control flies when *Gal4>UAS-OBP6* originated from crosses using *UAS-OBP6* as the male parent (Experiment 1, Table 1). In this experiment there was evidence for increased speeds in older (over 48h old) *Gal4>UAS-OBP6* flies.

By contrast, *Gal4>UAS-OBP6* flies originating from the reciprocal cross (*UAS-OBP6* as the female parent) flew consistently faster and attained higher maximum speeds than control flies (Fig. *2B-2E*, Table 1) and this pattern was observed in both mated and virgin F1 flies (Experiment 2 and 3). There was an effect of age in both experiments, flies from the older age groups (those flies emerging after one week) flew faster than the younger cohort. The discrepancy in the F1 flight activity results between *UAS-OBP6* male and female parental lines could be due to maternal effects as observed in laboratory crosses of 'short' and 'long' flight phenotypes from other moth species (Gu & Danthanarayana, 1992).

Quantification of TAG and phospholipids in flown H. armigera

We hypothesise that OBPs function as a fuel carrier for the supply of lipids to the flight muscles during prolonged flight in *H. armigera*. To determine candidate lipid molecular species for binding with OBPs we compared the total lipid content of age-matched moths flown on the flight mills with those reared to adults and not forced to undergo flight. Six lipid classes were assayed using electrospray ionization tandem triple-quadrupole mass spectrometry (ESI-MS) including triacylglycerol (TAG), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylinositol (PI), phosphatidylglycerol (PG) and phosphatidylcholines (PC).

Unsurprisingly total TAG levels were a) the most abundant class of the lipids analysed and b) underwent the most pronounced decline in flown moths (1.8-fold reduction from 739.6 to 410.5 $nmolsg^{-1}fw$, $p = 0.006$) (Fig. *3A*). Sustained flight activity in insects is primarily powered by the

mobilisation of TAG in the insect fat body into diacylglycerol (DAG) which is then shuttled in the haemolymph to the flight muscle (Van der Horst & Ryan, 2012). The role of phospholipid metabolism in insects is far less understood although in this experiment we show a consistent depletion in each phospholipid class following flight (Fig. *3B*, see Fig. S7 for individual lipid species and Table S3 for test statistics). In mammals the relative abundance of the two most common phospholipids, PC and PE (also the two most abundant classes in *H. armigera* moths measured by ESI-MS, Fig. 3*B*), regulates the size and dynamics of lipid droplets and energy metabolism (van der Veen et al., 2017). Phospholipids are critical to membrane structure and function; the fatty acyl components of the phospholipids can provide another potential energy source. When cells are subject to starvation, levels of phospholipid classes decrease (Steinhauser et al., 2018). Lipid droplets are storage organelles at the centre of lipid and energy homeostasis. They have a unique architecture consisting of a hydrophobic core of neutral lipids which is dominated by TAG, enclosed by a phospholipid monolayer that is decorated by a specific set of proteins (Olzmann & Carvalho, 2019). Utilizing the reserves of TAG from lipid droplets for energy will release phospholipids which can also be metabolised.

The protein structure and lipid binding site of H. armigera OBPs

We used the three-dimension structure of an OBP from the blowfly, *Phormia regina*, (PregOBP56a) (Ishida, Ishibashi, & Leal, 2013), as a template for HarmOBP6 and HarmOBP3, and the pheromone binding protein from the silkmoth *Bombyx mori* BmorPBP1 (1DQE) as a template for HarmPBP1, to build structural models and predict the binding efficacy to a range of fatty acids. We used HarmPBP1 as a positive control for semiochemical binding in relation to its observed role in female sex pheromone response (Ye et al., 2017). As expected, the structures of both HarmOBP6 and HarmOBP3 resemble typical insect OBPs composing of six α-helices connected by loops and three disulfide bridges that contribute to overall structural stability (Fig. *4A* & Fig. *S8*). Binding site prediction indicates the OBP has a pocket of 772.8 $Å^3$ volume and 917.1 $Å^2$ area for OBP6 and 777.9 $Å^3$ volume and 642.4 $Å^2$ area for OBP3 with a "Tunnel" conformation suitable for lipid binding (Fig. *4B* & Fig. *S8*).

To quantify the strength of molecular interactions between the over-expressed OBPs and potential substrates, molecular docking was conducted to determine binding energies with a range of fatty acids and olfactory odorants (semiochemicals) (Table 2). A total of 33 compounds were selected to dock with the predicted HarmOBP6, HarmOBP3 and HarmPBP1 protein structures, including 9

fatty acids, 15 semiochemicals (identified from the Pherobase database https://www.pherobase.com), L-proline (amino acid) (Rio et al., 2016), D-trelahose (sugar) and a selection of DAG/TAG/phospholipid species analysed by ESI-MS (Table 2).

The *H. armigera* OBPs possessed the lowest overall binding energies with TAG and phospholipids (Table 2). Apart from phosphatidylethanolamine (PE), HarmOBP3 had a greater binding affinity to each long-chain lipid than OBP6, with the mean docking values for modelled HarmOBP6 and HarmOBP3 -16.30±0.80 kcal mol^{-1} and -18.20±1.84 kcal mol^{-1} respectively. The lowest values were observed for HarmOBP3:TAG (52:2) and HarmOBP:phosphatidylinositol (PI 36:3) (Table 2). The predicted binding model for PI in the pocket of OBP3 and OBP6 is shown in Fig. *4C* with optimal predictions for TAG and other phospholipids in Fig. *S8*. In contrast, the binding predictions between HarmPBP1 and lipid molecules were highly inconsistent (Table 2). As expected from its putative role in sex pheromone transportation (Ye et al., 2017), HarmPBP1 bound semiochemicals and fatty acids with greater negative values compared to the OBPs (Table 2). There was little difference in semiochemical or fatty acid docking values between OBP6 and OBP3. Overall, these molecular docking patterns support the hypothesis that the *H. armigera* OBPs investigated in this study have a binding affinity for long-chain fatty acids which is either supplementary to their role in olfaction or represents an entirely new physiological function. There is now compelling evidence that OBPs perform physiological functions beyond olfaction (Pelosi, Iovinella, Zhu, Wang, & Dani, 2018). The sensilla of *Drosophila* maintain a robust response to a wide range of odours even when all abundantly expressed antennal OBP genes are deleted demonstrating that many OBPs are not essential to the olfactory response (Xiao et al., 2019). Humidity-detection (hygrosensation) relies on a single OBP (OBP59a) within *Drosophila* antenna (Sun et al., 2018) contravening the typical model that OBPs exclusively transport hydrophobic odorants to receptors. The diverse array of non-olfactory roles for OBPs in Dipteran include bacterial-induced haematopoiesis in tsetse flies (Benoit et al., 2017), the transportation of sex-pheromones in *Helicoverpa sp.* (Sun et al., 2012) and eggshell formation in the mosquito *Aedes aegypti* (Marinotti et al., 2014). The degree of redundancy in OBP function and the circumstances under which dual or split roles are performed is currently unknown but tissue-specific functional genomics will undoubtedly begin to uncover the broader range of OBP operate.

Summary

It is recently become accepted that the versatility of OBPs is greater than previously thought and this group of proteins represent a highly adaptive set of hydrophobic carriers performing multiple physiological functions beyond their classical role in chemoreception (Pelosi, Iovinella, Zhu, Wang, & Dani, 2018). Our findings on two *H. armigera* OBPs are consistent with this view and we propose an additional physiological role in regulating insect flight in a migratory Lepidopteran organism. The affinity of OBPs for long-chain fatty acids (Ishida et al., 2013) lends support to the hypothesis that OBPs act as carriers of hydrophobic free fatty acids produced from upstream lipid metabolism as part of the flight fuel pathway. The OBP homology structure models and binding affinities for a range of substrates described here support this. The precise mechanism(s) of how over-expressed OBPs contribute to flight performance at the biochemical and cellular level needs further study. Coping with the extreme energy demands of sustained migratory flight in insects is just one of several traits that make up the heritable 'migratory syndrome' (Roff & Fairbairn, 2007). Investigating the pathways and mechanisms that support such a fascinating feat of endurance is an excellent means to understand animal migration at the genetic level.

Acknowledgements

We thank the three anonymous Reviewers for their critical comments and suggestions for improvement of the initial draft. The Max Planck Institute in Jena, Germany and the University of Valencia provided vital *H. armigera* insect material. This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) as part of a Future Leader Fellowship (grant number: BB/N012011/1) (to CMJ) and by Jilin University, China as part of a Tang Aoqing Chair Professorship (No. TAQ(JZ)-2017[7]-201811) to JJZ. LVM is supported by the BBSRC Institute Strategic Programme Tailoring Plant Metabolism (BBS/E/C/000I0420).

References

Antunes DA*, et al.* (2017) DINC 2.0: A New Protein-Peptide Docking Webserver Using an Incremental Approach. *Cancer Research* 77(21):E55-E57.

Arrese EL & Soulages JL (2010) Insect Fat Body: Energy, Metabolism, and Regulation. *Annual Review of Entomology,* Annual Review of Entomology, Vol 55, pp 207-225.

Bates D, Machler M, Bolker BM, & Walker SC (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1):1-48.

Bischof J, Maeda RK, Hediger M, Karch F, & Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phi C31 integrases. *Proceedings of the National Academy of Sciences of the United States of America* 104(9):3312-3317.

Benoit *et al*. (2017) Symbiont-induced odorant binding proteins mediate insect host hematopoiesis. eLife 6:e19535.

Brand AH & Perrimon N (1993) Targeted gene-expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118(2):401-415.

Brand P, *et al.* (2016) The origin of the odorant receptor gene family in insects. *Elife* 7:e38340

Bruce TJA & Pickett JA (2011) Perception of plant volatile blends by herbivorous insects - Finding the right mix. *Phytochemistry* 72(13):1605-1611.

Chapman JW*, et al.* (2010) Flight Orientation Behaviors Promote Optimal Migration Trajectories in High-Flying Insects. *Science* 327(5966):682-685.

Chertemps T*, et al.* (2012) A carboxylesterase, Esterase-6, modulates sensory physiological and behavioral response dynamics to pheromone in Drosophila. *Bmc Biology* 10.

Delignette-Muller ML & Dutang C (2015) fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* 64(4):1-34.

Devaiah SP*, et al.* (2006) Quantitative profiling of polar glycerolipid species from organs of wild-type Arabidopsis and a phospholipase D alpha 1 knockout mutant. *Phytochemistry* 67(17):1907-1924.

Dhanik A, McMurray JS, & Kavraki LE (2013) DINC: A new AutoDock-based protocol for docking large ligands. *Bmc Structural Biology* 13.

Dingle H (2014) *Migration: The Biology of Life on the Move* (Oxford University Press) 2nd Ed.

Dundas J*, et al.* (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research* 34:W116-W118.

Fitt GP (1989) The ecology of Heliothis species in relation to agroecosystems. *Annual Review of Entomology* 34:17-52.

Getahun MN*, et al.* (2016) Intracellular regulation of the insect chemoreceptor complex impacts odour localization in flying insects. *Journal of Experimental Biology* 219(21):3428-3438.

Gomez-Diaz C, Reina JH, Cambillau C, & Benton R (2013) Ligands for Pheromone-Sensing Neurons Are Not Conformationally Activated Odorant Binding Proteins. *Plos Biology* 11(4).

Gong DP, Zhang HJ, Zhao P, Xia QY, & Xiang ZH (2009) The Odorant Binding Protein Gene Family from the Genome of Silkworm, Bombyx mori. *Bmc Genomics* 10.

Graham LA*, et al.* (2001) Characterization and cloning of a Tenebrio molitor hemolymph protein with sequence similarity to insect odorant-binding proteins. *Insect Biochemistry and Molecular Biology* 31(6-7):691-702.

Grosse-Wilde E, Svatos A, & Krieger J (2006) A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor in vitro. *Chemical Senses* 31(6):547-555.

Gu HN & Danthanarayana W (1992) Quantitative genetic analysis of dispersal in *Epiphyas postvittana*. 1. Genetic variation in flight capacity. *Heredity* 68:53-60.

Gu S-H*, et al.* (2014) Molecular Characterization and Differential Expression of Olfactory Genes in the Antennae of the Black Cutworm Moth Agrotis ipsilon. *Plos One* 9(8).

Guo W*, et al.* (2011) CSP and Takeout Genes Modulate the Switch between Attraction and Repulsion during Behavioral Phase Change in the Migratory Locust. *Plos Genetics* 7(2).

Hansson BS & Stensmyr MC (2011) Evolution of Insect Olfaction. *Neuron* 72(5):698-711.

Ishida Y, Ishibashi J, & Leal WS (2013) Fatty Acid Solubilizer from the Oral Disk of the Blowfly. *Plos One* 8(1).

Jones CM*, et al.* (2015) Genomewide transcriptional signatures of migratory flight activity in a globally invasive insect pest. *Molecular Ecology* 24(19):4901-4911.

Jones CM, Parry H, Tay WT, Reynolds DR, & Chapman JW (2019) Movement Ecology of Pest Helicoverpa: Implications for Ongoing Spread. *Annual Review of Entomology, Vol 64,* Annual Review of Entomology, ed Douglas AE), Vol 64, pp 277-295.

Jones, CM, et al. (2020), Odorant binding proteins promote insect flight in Helicoverpa armigera, Dryad, Dataset, https://doi.org/10.5061/dryad.dr7sqv9w4

Krank J, Murphy RC, Barkley RM, Duchoslav E, & McAnoy A (2007) Qualitative analysis and quantitative assessment of changes in neutral glycerol lipid molecular species within cells. *Lipidomics and Bioactive Lipids: Mass-Spectrometry-Based Lipid Analysis,* Methods in Enzymology, ed Brown HA), Vol 432, pp 1-20.

Larter NK, Sun JS, & Carlson JR (2016) Organization and function of Drosophila odorant binding proteins. *Elife* 5.

Lartigue A*, et al.* (2002) X-ray structure and ligand binding study of a moth chemosensory protein. *Journal of Biological Chemistry* 277(35):32094-32098.

Laskowski RA, Macarthur MW, Moss DS, & Thornton JM (1993) PROCHECK – a programme to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26:283-291.

Leal WS (2013) Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annual Review of Entomology, Vol 58,* Annual Review of Entomology, ed Berenbaum MR), Vol 58, pp 373-391.

Lenth RV (2016) Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* 69(1):1-33.

Li ZQ*, et al.* (2015) Transcriptome comparison of the sex pheromone glands from two sibling Helicoverpa species with opposite sex pheromone components. *Scientific Reports* 5.

Li *et al*., (2008) Multiple functions of an odorant-binding protein in the mosquito Aedes aegypti. *Biochemical and Biophysical Research Communications* 372:464-468.

Liedvogel M, Akesson S, & Bensch S (2011) The genetics of migration on the move. *Trends in Ecology & Evolution* 26(11):561-569.

Marinotti *et al*., (2014). Integrated proteomic and transcriptomic analysis of the *Aedes aegypti* eggshell. *BMC Developmental Biology* 14:15.

McCormick AC*, et al.* (2017) Comparing the Expression of Olfaction-Related Genes in Gypsy Moth (Lymantria dispar) Adult Females and Larvae from One Flightless and Two Flight-Capable Populations. *Frontiers in Ecology and Evolution* 5.

Merlin C & Liedvogel (2019) The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond. *Journal of Experimental Biology* 222: jeb191890

Mescher MC & De Moraes CM (2015) Role of plant sensory perception in plant-animal interactions. *Journal of Experimental Botany* 66(2):425-433.

Miller DFB, Holtzman SL, & Kaufman TC (2002) Customized microinjection glass capillary needles for P-element transformations in Drosophila melanogaster. *Biotechniques* 33(2):366-+.

Minter M*, et al.* (2018) The tethered flight technique as a tool for studying life-history strategies associated with migration in insects. *Ecological Entomology* 43(4):397-411.

Missbach C, Vogel H, Hansson BS, & Grosse-Wilde E (2015) Identification of Odorant Binding Proteins and Chemosensory Proteins in Antennal Transcriptomes of the Jumping Bristletail Lepismachilis y-signata and the Firebrat Thermobia domestica: Evidence for an Independent OBP-OR Origin. *Chemical Senses* 40(9):615-626.

Olzmann JA & Carvalho P (2019) Dynamics and functions of lipid droplets. *Nature Reviews Molecular Cell Biology* 20(3):137-155.

Pelosi P, Zhou JJ, Ban LP, & Calvello M (2006) Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences* 63(14):1658-1676.

Pelosi P, Iovinella I, Felicioli A, & Dani FR (2014) Soluble proteins of chemical communication: an overview across arthropods. *Frontiers in Physiology* 5.

Pelosi P, Iovinella I, Zhu J, Wang GR, & Dani FR (2018) Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biological Reviews* 93(1):184-200.

Rio RVM, Attardo GM, & Weiss BL (2016) Grandeur Alliances: Symbiont Metabolic Integration and Oblicate Arthopod Hematophagy. *Trends in Parasitology* 32(9):739-749.

Reppert SM, Guerra PA, & Merlin C (2016) Neurobiology of Monarch Butterfly Migration. *Annual Review of Entomology, Vol 61,* Annual Review of Entomology, ed Berenbaum MR), Vol 61, pp 25-42.

Roff DA & Fairbairn DJ (2007) The evolution and genetics of migration in insects. *Bioscience* 57(2):155-164.

Ruiz-Lopez N, Haslam RP, Napier JA, & Sayanova O (2014) Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. *Plant Journal* 77(2):198-208.

Schmittgen TD & Livak KJ (2008) Analyzing real-time PCR data by the comparative C-T method. *Nature Protocols* 3(6):1101-1108.

Seroude L, Brummel T, Kapahi P, & Benzer S (2002) Spatio-temporal analysis of gene expression during aging in Drosophila melanogaster. *Aging Cell* 1(1):47-56.

Steinhauser ML*, et al.* (2018) The circulating metabolome of human starvation. *Jci Insight* 3(16).

Sun *et al*. (2012) Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. *PLoS One* 7(1):e30040.

Sun *et al.* (2018) Humidity response depends on the small soluble protein Obp59a in Drosophila. eLife 7:e39249.

Tegoni M, Campanacci V, & Cambillau C (2004) Structural aspects of sexual attraction and chemical communlication in insects. *Trends in Biochemical Sciences* 29(5):257-264.

Trott O & Olson AJ (2010) Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry* 31(2):455-461.

Tsuchihara K*, et al.* (2005) An odorant-binding protein facilitates odorant transfer from air to hydrophilic surroundings in the blowfly. *Chemical Senses* 30(7):559-564.

Untergasser A*, et al.* (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Research* 40(15).

Usher S, *et al.* (2017) Tailoring seed oil composition in the real world: optimising omega-3 long chain polyunsaturated fatty acid accumulation in transgenic Camelina sativa. *Scientific Reports* 7.

Van der Horst DJ & Ryan RO (2012) *Lipid Transport* pp 317-345.

Van der Veen JN, *et al.* (2017) The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1859(9, Part B):1558-1572.

Vieira FG & Rozas J (2011) Comparative Genomics of the Odorant-Binding and Chemosensory Protein Gene Families across the Arthropoda: Origin and Evolutionary History of the Chemosensory System. *Genome Biology and Evolution* 3:476-490.

Wang ZJ, Dong YC, Desneux N, & Niu CY (2013) RNAi Silencing of the HaHMG-CoA Reductase Gene Inhibits Oviposition in the Helicoverpa armigera Cotton Bollworm. *Plos One* 8(7).

Xiao *et al*. (2019) Robust olfactory responses in the absence of odorant binding proteins. eLife 8:e51040.

Yan S, *et al.* (2013) Molecular Cloning, Characterization, and mRNA Expression of Two Cryptochrome Genes in Helicoverpa armigera (Lepidoptera: Noctuidae). *Journal of Economic Entomology* 106(1):450-462.

Zhan S, *et al.* (2014) The genetics of monarch butterfly migration and warning colouration. *Nature* 514(7522):317-321.

Zhou JJ, He XL, Pickett JA, & Field LM (2008) Identification of odorant-binding proteins of the yellow fever mosquito Aedes aegypti: genome annotation and comparative analyses. *Insect Molecular Biology* 17(2):147-163.

Zhou JJ (2010) Odorant-binding proteins in insects. *Vitamins and Hormones: Pheromones,* Vitamins and Hormones, ed Litwack G), Vol 83, pp 241-272.

Author Contributions

C.M.J. and J.J.Z conceived and designed the study. S.W., M.M., R.A.H., L.V.M. and H.V. performed the experimental work. K.S.L. designed and provided the tethered flight mill equipment. C.M.J, J.J.Z., S.W., A.W., M.M., L.V.M. and H.V. conducted data analysis. J.X.

provided advice, academic and financial support for S.W. C.M.J. and J.J.Z. wrote the initial draft of the manuscript. All authors edited and made comments on the final draft.

Data Accessibility

The candidate gene qPCR, OBP6 tissue-specific expression and lipid quantification ESI-MS data have been archived in Dryad (doi:10.5061/dryad.dr7sqv9w4). The R script used to analyse the flight mill data is available on request from C.M. Jones.

Table 1. Estimated regression parameters, standard errors, *t* values for GLMM models for *Drosophila* flight mill experiments.

Experiment	Response	Regression parameter	Estimate	Std. error	*t* value
#1 F0 = ♂*UAS-OBP6 x* ♀*muscle-GAL4*	*avgsp*	Intercept	0.316	0.021	14.70
F1 = Virgin ♀		StrainOBP6	-0.019	0.026	-0.72
		AgeOver48h	0.048	0.026	1.84
	maxsp	Intercept	0.481	0.033	12.35
		StrainOBP6	-0.017	0.035	-0.49
		AgeOver48h	0.067	0.037	1.84
#2 F0 = ♀*UAS-OBP6 x* ♂*muscle-GAL4*	*avgsp*	Intercept	0.226	0.015	14.76
F1 = Mated ♀		StrainOBP6	0.015	0.016	0.93
		Age6D	0.097	0.019	5.11
		Age15D	0.054	0.019	2.81
	maxsp	Intercept	0.326	0.024	13.68
		StrainOBP6	0.010	0.025	0.40
		Age6D	0.179	0.030	6.05
		Age15D	0.135	0.030	4.50
#3F0 = ♀*UAS-OBP6 x* ♂*muscle-GAL4*	*avgsp*	Intercept	0.251	0.020	12.50
F1 = Virgin ♀		StrainOBP6	0.019	0.015	1.34
		Age2weeks	0.039	0.021	1.90
		Age4weeks	0.046	0.021	2.22

	maxsp	Intercept	0.373	0.034	11.03
		StrainOBP6	0.046	0.024	1.95
		Age2weeks	0.088	0.034	2.62
		Age4weeks	0.071	0.033	2.14

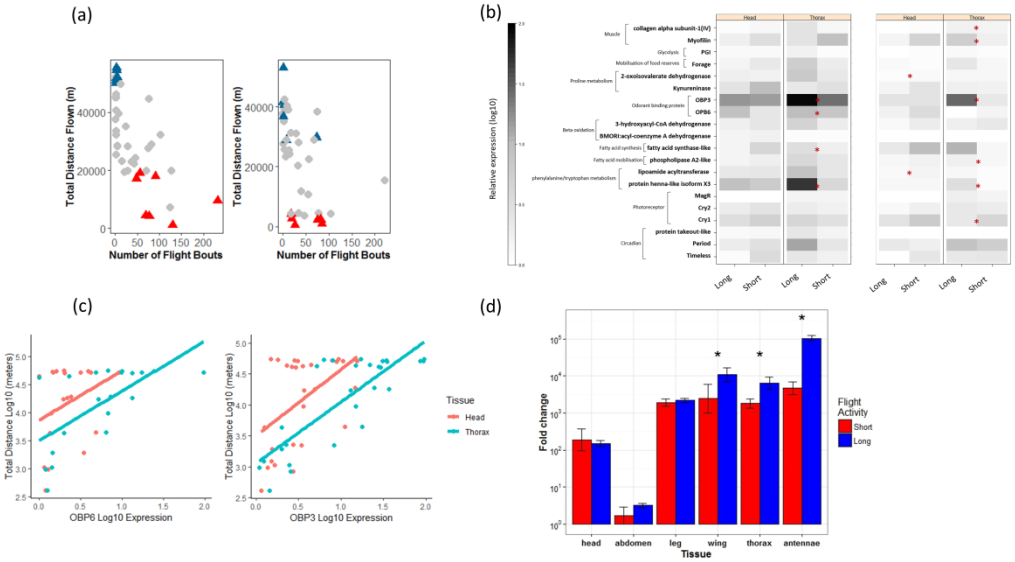
Note: avgsp is the average speed and maxsp is the maximum speed flown on the flight mill

Table 2. Molecular docking between HarmOBP6 and HarmOBP3 with fatty acids, semiochemicals, TAG and phospholipids**.**

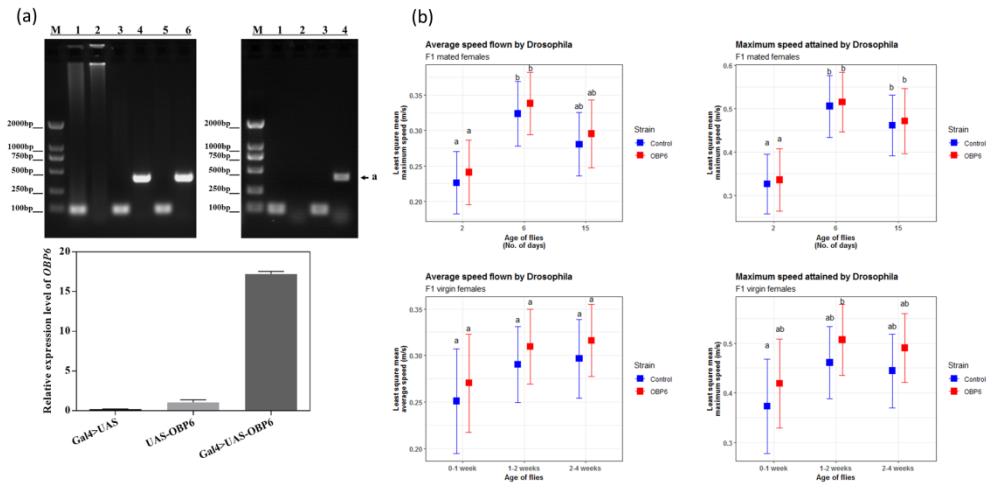
Ligands	Ligands	Binding energy (kcal mol^{-1})		
		OBP6	OBP3	PBP1*
Fatty acids	1,2-diacylglycerol	-5.1	-4.6	-4.9
	α-linolenic acid	-5.8	-5.7	-7.8
	cis-vaccenic acid	-5.6	-5.3	-6.9
	D-trehalose	-5.2	-5.5	-5.5
	γ-linolenic acid	-6.2	-5.7	-7.5
	Linoleic acid	-5.4	-5.3	-7.3
	L-proline	-4.5	-4.4	-4.5
	Oleic acid	-5.4	-5.5	-7.0
	Palmitic acid	-5.3	-5.1	-6.5
	Palmitoleic acid	-5.5	-5.9	-6.9
	Stearic acid	-5.6	-5.5	-6.7
	trans-vaccenic acid	-5.5	-6.0	-7.1
Semiochemicals	2-phenylacetaldehyde	-5.1	-4.6	-5.9
	2-phenylethanol	-4.9	-4.7	-5.8
	benzaldehyde	-4.8	-4.6	-5.6
	heptanal	-4.2	-3.9	-4.6
	hexadecanal	-5.0	-4.9	-6.5
	hexadecanol	-5.0	-4.8	-6.3
	nonanal	-4.7	-3.9	-5.4
	phenylmethanol	-4.7	-4.8	-5.4
	salicylaldehyde	-5.0	-5.0	-5.2
	tetradecanal	-5.0	-4.5	-6.0
	(*Z*)-7-hexadecenal	-5.2	-5.0	-6.7
	(*Z*)-9-hexadecenal	-5.5	-5.0	-6.6
	(*Z*)-9-tetradecenal	-5.2	-4.8	-6.4
	(*Z*)-11-hexadecenal	-5.4	-4.9	-6.7
	(*Z*)-11-hexadecenol	-5.3	-5.0	-6.4
TAG/phospholipid	triacylglycerol – TAG (52:3)	-13.0	-21.0	-7.5

	triacylglycerol – TAG (52:2)	-18.0	-22.0	50.8
	phosphatidylethanolamine – PE 36:4	-17.5	-3.8	5.1
	phosphatidylserine – PS (36:2)	-18.0	-19.4	19.1
	phosphatidylinositol – PI (36:3)	-19.9	-21.4	26.0
	phosphatidylglycerol – PG (34:3)	-17.6	-18.7	21.3
	phosphatidylcholines – PC (36:5)(1)	-13.7	-18.2	-0.7
	phosphatidylcholines – PC (36:5)(2)	-14.0	-19.3	-14.8
	phosphatidylcholines – PC (36:5)(3)	-15.3	-19.7	-1.5

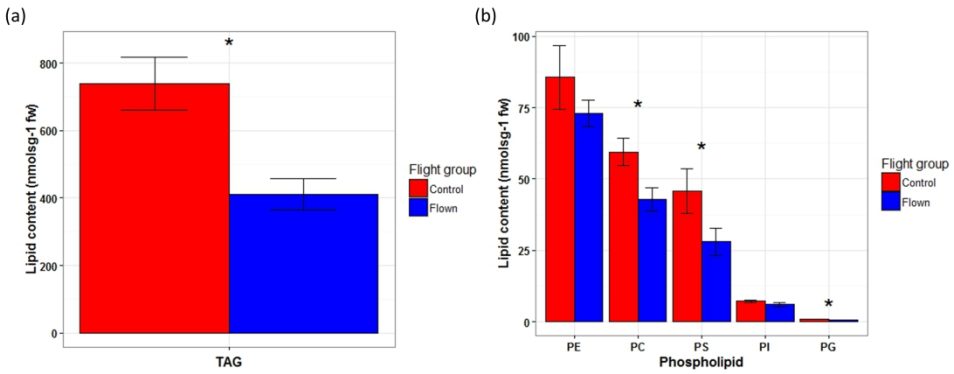
Numbers in parentheses indicate isomers for phosphatidylcholines as carbon atoms:unsaturations. (1) represents 18:3/18:2; (2) represents 18:2/18:3 and (3) represents 16:0/20:5. *Corresponds to HarmPBP1 used as reference target with a reported function in binding sex pheromones. "-" indicates no binding predicted by molecular docking based on the settings used.



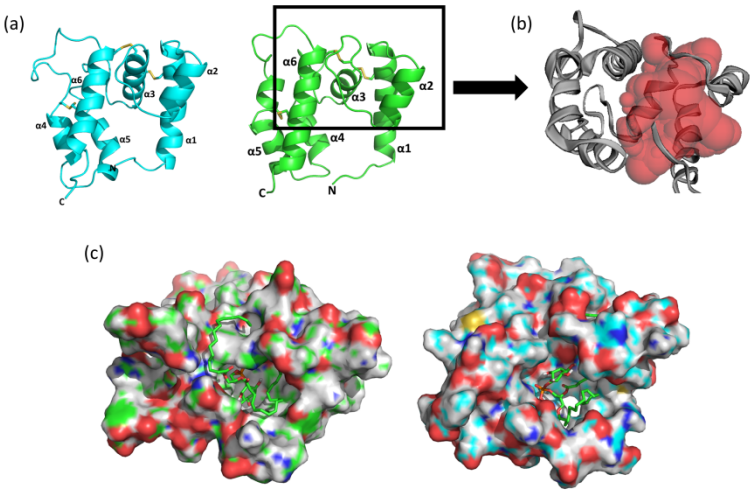
mec_15556_f1.tif



mec_15556_f2.tif



mec_15556_f3.tif



mec_15556_f4.tif