UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

Emeline Boni Campanini

Evolução molecular e padrões de expressão de genes da família das proteínas ligantes a odores (OBPs) em duas espécies de moscas-das-frutas do grupo *Anastrepha fraterculus*

São Carlos – SP

Abril/2016

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

Emeline Boni Campanini

Evolução molecular e padrões de expressão de genes da família das proteínas ligantes a odores (OBPs) em duas espécies de moscas-das-frutas do grupo *Anastrepha fraterculus*

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutora em Ciências com área de concentração em Genética Evolutiva.

Orientador: Prof. Dr. Reinaldo Alves de Brito

São Carlos – SP

Abril/2016

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

Campanini, Emeline Boni
C186e Evolução molecular e padrões de expressão de genes da família das proteínas ligantes a odores (OBPs) em duas espécies de moscas-das-frutas do grupo Anastrepha fraterculus / Emeline Boni Campanini. -- São Carlos : UFSCar, 2016.
104 p.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2016.

1. Anastrepha. 2. Subfamílias das OBPs. 3. Seleção positiva. 4. Expressão diferencial. 5. PCR em tempo real. I. Título.

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado da candidata Emeline Boni Campanini, realizada em 18/04/2016:

Prof. Dr. Reinaldo Otavio Alvarenga Alves de Brito
UFSCar

Prof. Dr. Caio Cesar de Melo Freire
UFSCar

Prof. Dr. Diogo Meyer
USP

Profa. Dra. Tatiana Teixeira Torres
USP

Prof. Dr. Sergio Russo Matioli
USP

Não é por acidente que vemos verde por quase toda a parte. Não é por acidente que nos achamos empoleirados num minúsculo ramo no meio de uma frondosa e florescente árvore da vida; não é por acidente que estamos cercados por milhões de outras espécies que comem, crescem, decompõem-se, nadam, andam, voam, cavam, espreitam, perseguem, fogem, suplantam em velocidade ou esperteza. Sem plantas verdes que nos superem numericamente no mínimo dez vezes, não haveria energia para nos mover. Sem as corridas armamentistas em eterna escalada entre predadores e presas, parasitas e hospedeiros, sem a "guerra da natureza" de Darwin, sem sua "fome e morte", não existiriam sistemas nervosos capazes de ver e muito menos de apreciar e compreender. Estamos cercados por infindáveis formas belíssimas e fascinantes, e não é por acidente, e sim uma consequência direta da evolução pela seleção natural não aleatória – única na vida, **o maior espetáculo da Terra**.

(Richard Dawkins)

Dedico esse trabalho aos meus pais Rubens e Rosemeire, e aos meus irmãos Andrei e Eike, por todo o amor, apoio, compreensão e pela fé incondicional que sempre depositaram em meus sonhos...

AGRADECIMENTOS

Ao professor Reinaldo Alves de Brito pela orientação, por toda a ajuda e confiança em mim depositada. Obrigada por todos os ensinamentos e por me inspirar com seu encantamento pelo estudo da Evolução.

Aos queridos amigos do Laboratório de Genética de Populações: Aline, André, Carlos, Cris, Diana, Felipe, Fernando, Ider, Isa (minha terapeuta querida!), Janaína, Lívia (minha BBF!), Manu, Nancy, Paty, Samira e Victor (Nat tb!) pela amizade, por toda a ajuda no desenvolvimento desse trabalho e por tudo o que vivemos juntos. Obrigada também à professora Andréa Peripato e a nossa técnica Isabel Tscherne.

Ao professor Federico Hoffmann pela oportunidade de realizar o doutorado sanduíche sob sua supervisão e por todo o aprendizado obtido durante os sete inesquecíveis meses que passei na Mississippi State University. A Mike Vandewege por toda ajuda e todos os ensinamentos, de programação em Phyton à futebol americano. Aos amigos Dafne, Eli, Fred e Tâmara por todo o apoio e por serem a minha família em Starkville.

A Ricardo Nascimento pela companhia em todos aqueles finais de semana trabalhando na tese, por todo o apoio na etapa final desse trabalho e por estar sempre presente.

A minha família: avó, tios e primos que, mesmo sem entender nada de Evolução ou do meu trabalho, entendem a minha paixão pelo assunto e estão sempre na torcida.

À Universidade Federal de São Carlos e ao programa de Pós-graduação em Genética Evolutiva e Biologia Molecular, em especial a Ivanildes Menezes por todo auxílio.

Aos meus amigos de São Carlos, cujos nomes não vou citar, porque felizmente são muitos, aos amigos desde a infância e aos amigos dos treinos (Trupe Uaah e Studio Paula Uemura).

Aos membros da banca examinadora, professores Caio Freire, Diogo Meyer, Sérgio Matioli e Tatiana Torres pelas valiosas sugestões que contribuíram para o aperfeiçoamento desse trabalho.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de doutorado e auxílio financeiro (2012/17160-8).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de doutorado sanduíche no exterior (PDSE - 99999.004252/2014-04).

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

RESUMO

As proteínas ligantes a odores (OBPs – *odorant-binding proteins*) são de grande importância para a sobrevivência e reprodução, pois participam do passo inicial da cascata de transdução dos sinais olfatórios, solubilizando e transportando os sinais químicos (odores e feromônios) até os receptores olfativos. A análise comparativa dos genes OBPs entre espécies próximas pode ajudar na compreensão de como o repertório desses genes é mantido sob seleção natural, além de fornecer informações acerca de como as diferenças observadas podem afetar as respostas olfatórias e, consequentemente, levar à diferenciação dessas espécies. Estudamos genes OBP em duas espécies-irmãs *Anastrepha fraterculus* e *Anastrepha obliqua*, as quais têm preferência por diferentes frutos hospedeiros, usando dados de transcriptomas e de PCR quantitativa. Identificamos 24 sequências OBP para *A. fraterculus* e 25 para *A. obliqua*, que corresponderam a 21 genes OBP de *Drosophila melanogaster*. Análises filogenéticas separaram as OBPs de *Anastrepha* em quatro ramos, que representam quatro subfamílias dessa família gênica: classic, minus-C, plus-C e dimer. Evidências de seleção positiva foram observadas nos genes da subfamília classic *OBP56h-1*, *OBP56h-2* e *OBP57c*, e para o gene da subfamília plus-C *OBP50a*, e pelo menos um evento de duplicação gênica que precede a especiação dessas duas espécies. Quatro sítios selecionados positivamente resultavam em mudanças radicais nas propriedades dos aminoácidos. Inferências utilizando a estrutura terciária predita para essas OBPs revelaram que pelo menos um desses sítios faz parte da cavidade ligante ao odor de *OBP50a*, sendo que uma mudança nessa região pode alterar a especificidade de uma OBP. Análises de expressão por PCR quantitativa em diferentes estágios reprodutivos das moscas mostraram que todos os nove genes testados possuíam expressão gênica significativamente diferente entre *A. fraterculus* e *A. obliqua* para mais de um perfil reprodutivo, sendo que *OBP56a*, *OBP56d*, *OBP57c* e os dois parálogos *OBP56h* foram os que mais apresentaram diferenças entre as duas espécies. Todos os resultados gerados pelo presente trabalho indicam que pelo menos sete genes OBP podem estar envolvidos na diferenciação entre *A. fraterculus* e *A. obliqua* e, potencialmente, na diferenciação do grupo *fraterculus*.

Palavras-chave: *Anastrepha;* subfamílias das OBPs; seleção positiva; expressão diferencial; PCR em tempo real.

ABSTRACT

Odorant-binding proteins (OBPs) are of great importance for survival and reproduction since they participate in initial steps of the olfactory signal transduction cascade, solubilizing and transporting chemical signals to the olfactory receptors. A comparative analysis of OBPs between closely related species may help explain how these genes evolve and are maintained under natural selection and how differences in these proteins can affect olfactory responses, and consequently lead to species differentiation. We studied OBP genes in the closely related species *Anastrepha fraterculus* and *Anastrepha obliqua*, which, albeit generalists, have different host preferences, using transcriptomes and real time quantitative PCR data. We identified 24 different OBP sequences from *Anastrepha fraterculus* and 25 from *A. obliqua*, which correspond to 21 *Drosophila melanogaster* OBP genes. Phylogenetic analysis separated *Anastrepha* OBPs sequences in four branches that represent four subfamilies: classic, minus-C, plus-C and dimer. We found evidence of positive selection in three classic subfamily genes *OBP56h-1*, *OBP56h-2* e *OBP57c* and in the plus-C subfamily gene *OBP50a*, and at least one duplication event that preceded the speciation of these two species. Four positively selected sites putatively resulted in radical changes in amino acid properties. Inferences on tertiary structures of putative proteins from these genes revealed that at least one positively selected change involves the binding cavity (the odorant binding region) in the plus-C *OBP50a*, which is important because changes in the binding cavity could change OBPs specificity. Differential gene expression analysis at different reproductive stages showed that all nine OBP genes tested were significantly differentially expressed between *A. fraterculus* and *A. obliqua* at several reproductive profiles, but *OBP56a*, *OBP56d*, *OBP57c* and both *OBP56h* paralogs showed the highest differences in expression levels. The results generated in this study indicated that at least seven OBP genes may be involved in the *A. fraterculus* e *A. obliqua* differentiation, and in the *fraterculus* group differentiation as well.

Keywords: *Anastrepha;* OBP subfamilies; positive selection; differential expression; real time PCR.

Capítulo I

Introdução Geral e Objetivos

Capítulo I – Introdução Geral e Objetivos

1.1. Moscas-das-frutas do grupo *fraterculus*

Mosca-das-frutas é a denominação recebida por todas as espécies de moscas frugívoras pertencentes à família Tephritidae. As espécies dessa família destacam-se por sua grande importância econômica, uma vez que suas larvas causam danos expressivos às culturas de frutas carnosas (DUARTE; MALAVASI, 2000). A família Tephritidae possui cerca de 500 gêneros, sendo *Anastrepha* o maior deles, com 197 espécies atualmente reconhecidas (SEGURA et al., 2013). A maioria das espécies de *Anastrepha* distribui-se pela região neotropical do continente americano, estando estabelecidas no sul dos Estados Unidos, em toda a América Central e América do Sul, com exceção no Chile onde ocorrem apenas de forma esporádica ao norte do deserto do Atacama (MALAVASI; ZUCHHI; SUGAYAMA, 2000). No Brasil, são encontradas 112 espécies desse gênero, distribuídas em 13 grupos distintos (URAMOTO; WALDER; ZUCCHI, 2004; ZUCCHI, 2008). As mais relevantes espécies-praga da fruticultura nacional pertencem ao grupo denominado *fraterculus*, composto por 29 espécies que apresentam grande similaridade morfológica, tornando difícil a correta identificação e diferenciação entre elas (MALAVASI; ZUCHHI; SUGAYAMA, 2000).

As espécies do grupo *fraterculus* são holometábolas, isto é, completam o seu desenvolvimento através de quatro estágios: ovo, larva, pupa e fase adulta (Figura 1) (SALLES, 1995), e são denominadas multivoltinas, pois apresentam mais de uma geração anual (SELIVON, 2000). Uma vez tendo ocorrido a oviposição em um fruto hospedeiro, são necessárias entre 54 a 72 horas para as larvas eclodirem. As larvas se locomovem à medida que se alimentam e formam galerias dentro dos frutos, o que favorece a infecção secundária de fungos e bactérias, acelerando ainda mais o apodrecimento e tornando-os inviáveis para o consumo (CRUZ et al., 2000). Possuem fototropismo negativo, evitando

luz direta e se enterrando rapidamente no solo após cerca de oito dias da emergência. O tempo de desenvolvimento pupal é de aproximadamente 16 dias, e o tempo de vida dos adultos é em média, de 196 dias para os machos e 172 dias para as fêmeas (SALLES, 2000). Embora haja alguma especialização, as espécies do grupo *fraterculus* são consideradas polífagas (ZUCCHI, 2000a). Na natureza, os adultos se alimentam basicamente de exsudado de plantas, sucos de frutos, seiva, pólen e fezes de pássaros, sendo esses dois últimos importantes fontes proteicas necessárias para o desenvolvimento das glândulas sexuais em ambos os sexos (ZUCOLOTO, 2000) e, no caso das fêmeas, de maior importância uma vez que a quantidade de proteínas interfere diretamente nas taxas de oviposição (OVIEDO et al., 2011).

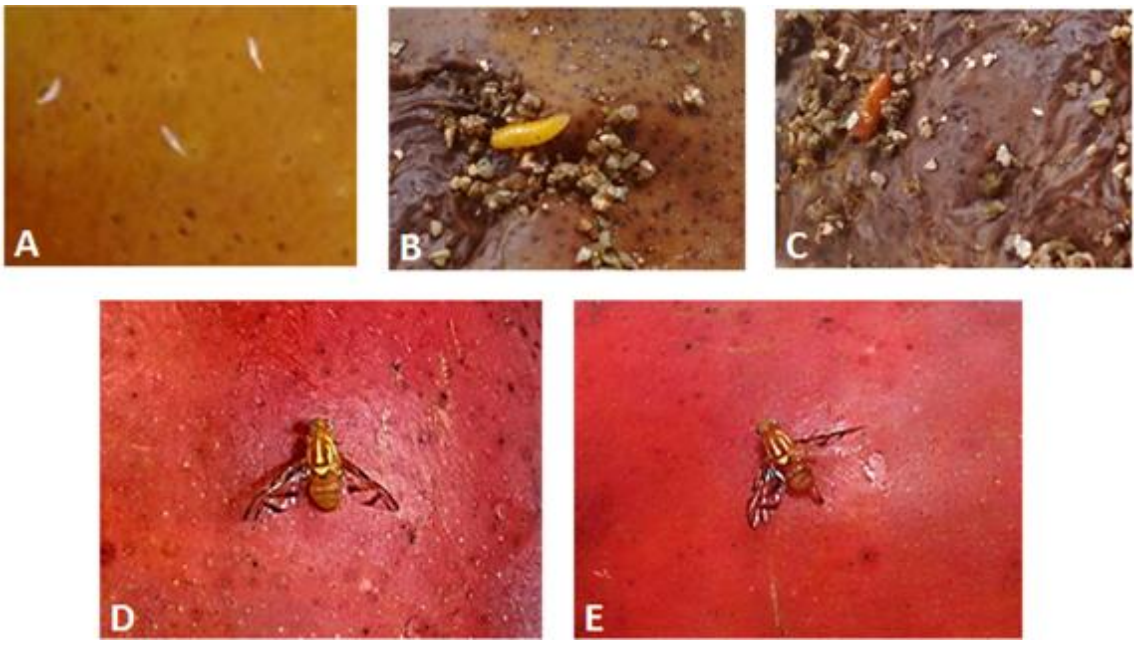


Figura 1. Fases da vida das moscas-das-frutas do gênero *Anastrepha*, mostradas a partir da espécie *Anastrepha obliqua*. A) ovos; B) larva; C) pupa; D) macho adulto; E) fêmea adulta.

Os adultos apresentam claro dimorfismo sexual, evidenciado pela presença de um protuberante ovipositor nas fêmeas (Figura 1D,1E) (SIVINSKI; DODSON, 1992). Apresentam formação de "leks" durante o período reprodutivo, o que consiste na formação de semicírculos ao redor da fêmea por no mínimo três machos (MORGANTE;

MALAVASI; PROKOPY, 1983), com liberação de feromônios e demonstrações de cortejo que podem envolver "displays" de asas, movimentos rápidos de cabeça ("head-rocking") e emissões acústicas (ALUJA et al., 2000; KASPI; YUVAL, 1999). Essas agregações ocorrem nas folhagens das árvores e não são relacionadas com recursos para a oviposição das fêmeas. Geralmente, as fêmeas interagem com vários machos antes de escolher um parceiro (MORGANTE; MALAVASI; PROKOPY, 1983). Ao contrário da maioria dos insetos, nas moscas-das-frutas são os machos que produzem os feromônios sexuais para a atração do sexo oposto, fato que pode ser de grande valor para o controle dessas espécies-praga nos pomares, por meio da atração e captura das fêmeas (VILELA; KOVALESKI, 2000).

A cópula dura em média de 40 a 60 minutos, e sua duração está positivamente relacionada com o sucesso reprodutivo, uma vez que a transferência de esperma é ineficiente em cópulas curtas (SEGURA et al., 2007). Após a cópula, as fêmeas buscam os sítios de oviposição. Essa é uma etapa crítica do ciclo de vida, uma vez que o fruto hospedeiro influencia o desenvolvimento dos estágios imaturos e também a sobrevivência e a fecundidade dos adultos por meio de ambos: da variação na qualidade do substrato oferecido e da susceptibilidade do hospedeiro a conter parasitoides (SUGAYAMA; MALAVASI, 2000), principalmente os da família Braconidae, considerados os mais importantes inimigos naturais (ZUCCHI, 2000b). Durante a oviposição, as fêmeas depositam no fruto hospedeiro o feromônio HMP ("Host marking pheromone") para sinalizar para outras fêmeas que aquele fruto já foi infestado, o que não impede, no entanto, que várias fêmeas utilizem o mesmo fruto hospedeiro (SUGAYAMA; MALAVASI, 2000). Os adultos vivem e se reproduzem próximos ao local de nascimento, uma vez que movimentos de dispersão ocorrem apenas por curtas distâncias, na busca de

parceiros e sítios de oviposição, sendo interrompidos quando o recurso é localizado (ALUJA, 1993).

As espécies do grupo *fraterculus* apresentam grande similaridade morfológica, o que torna difícil a diferenciação entre elas. A correta identificação entre as espécies desse grupo é importante não apenas do ponto de vista biológico mas também econômico, já que é uma etapa crítica para o controle das moscas nos pomares (MONTOYA et al., 2000). Alguns marcadores já foram propostos para diferenciar as espécies do grupo, como por exemplo a morfologia do ápice do ovipositor, presença ou ausência de manchas no subescutelo e o padrão de coloração e morfologia das asas. Esses caracteres morfológicos, quando avaliados em conjunto, permitem uma correta identificação na maioria dos casos (PERRE et al., 2014; SANTOS; URAMOTO; MATIOLI, 2001). No entanto, as medidas do ovipositor, que é o principal marcador morfológico, apresentam sobreposição interespecífica, já tendo sido observada variação ao longo da distribuição geográfica e até mesmo em exemplares obtidos de um mesmo fruto hospedeiro (ARAUJO; ZUCCHI, 2006). Devido à alta plasticidade fenotípica e consequentes limitações dos marcadores morfológicos, marcadores moleculares mitocondriais e nucleares têm sido buscados (FERNANDES, 2010; FRITZ; SCHABLE, 2004; LIMA, 2011; SELIVON; PERONDINI; MORGANTE, 2005; SMITH-CALDAS et al., 2001. Porém, apesar dos vários esforços, ainda não foram encontrados marcadores moleculares capazes de diferenciar com exatidão essas espécies.

1.2. O processo de especiação do grupo *fraterculus* investigado a partir de duas espécies-irmãs, *Anastrepha fraterculus* e *Anastrepha obliqua*

O grupo *fraterculus* compartilha uma história evolutiva bastante recente, o que é evidenciado não só pela alta similaridade morfológica observada entre as espécies, mas também pela obtenção de híbridos interespecíficos viáveis em laboratório (HENNING; MATIOLI, 2006; SANTOS; URAMOTO; MATIOLI, 2001; SELIVON; PERONDINI;

MORGANTE, 2005). Tais fatos são indicativos de que a completa separação das espécies ainda não ocorreu (WU, 2001), ou em outras palavras, que o processo de especiação nesse grupo ainda está ocorrendo. Para ocorrer especiação, é necessário que uma população acumule e mantenha variação genética suficiente e se divida em duas ou mais subpopulações. Após a subdivisão, os modelos mais estabelecidos de especiação consideram que é necessário que o fluxo gênico seja eliminado, ou pelo menos reduzido significativamente para que não ocorra a formação de híbridos ou tipos intermediários, garantindo assim o isolamento reprodutivo entre elas (GAVRILETS, 2014).

Barreiras contra o fluxo gênico podem ser físicas, que isolam completamente as subpopulações (modelo denominado especiação alopátrica), mas também a especiação pode ser resultado de seleção natural, seleção sexual, deriva genética e do balanço entre essas forças atuando em populações que não estão separadas por uma barreira física, mas nas quais o fluxo gênico é restrito ou pelas distâncias geográficas entre subpopulações adjacentes (modelo de especiação parapátrica) ou por divergência de nicho ecológico (modelo de especiação simpátrica) (GAVRILETS, 2014; RIDLEY, 2006), embora esta última seja considerada mais rara de ocorrer (RITCHIE, 2007). Após certo tempo de isolamento reprodutivo, barreiras pré- e pós-zigóticas se estabelecerão e limitarão a capacidade destas populações de trocar material genético mesmo que a barreira reprodutiva inicial deixe de existir, e só então o processo de especiação estaria completo (GAVRILETS, 2014).

Tem sido sugerido que as espécies do grupo *fraterculus* divergiram há pouco tempo, de forma que podem ainda não ter acumulado diferenças genéticas significativas além daquelas oriundas do processo de especiação, o que permite a obtenção de híbridos em condições de laboratório, embora na natureza as espécies simpátricas evitem o acasalamento interespecífico (LIMA-MENDONÇA et al., 2014). Isso sugere que

barreiras reprodutivas pré-zigóticas estão ocorrendo, o que nos leva a questionamentos acerca de quais características comportamentais e fisiológicas estão envolvidas nesse processo. A investigação dos genes responsáveis pela expressão dessas características pode nos permitir identificar os genes e atributos potencialmente envolvidos na especiação do grupo.

No presente estudo, duas espécies-irmãs, *Anastrepha fraterculus* (Wiedemann 1830) e *Anastrepha obliqua* (Macquart 1835), foram utilizadas como modelos para a compreensão do processo de especiação ocorrido no grupo *fraterculus*. Essas espécies foram escolhidas por estarem mais amplamente distribuídas no país (MALAVASI; ZUCHI; SUGAYAMA, 2000), apresentando grande sobreposição em sua distribuição (Figura 2) e por apresentarem diferenças em importantes características como horário de cópula e escolha do fruto hospedeiro. *A. obliqua* apresenta dois picos no horário de cópula (padrão bimodal), pela manhã entre 6 e 8 horas, e pela tarde, entre 16 e 18 horas, enquanto *A. fraterculus* apresenta apenas um pico (padrão unimodal) entre 4 e 8 horas da manhã (CHAHAD-EHLERS et al., comunicação pessoal).

Embora sejam consideradas oligotróficas, com *A. fraterculus* desenvolvendo-se em até 110 hospedeiros e *A. obliqua* em até 48 hospedeiros distintos (ZUCCHI, 2008, última atualização em 10/11/15), essas espécies mostram preferências particulares. Um teste de túnel de vento revelou que *A. fraterculus* era mais atraída por extratos de plantas da família Myrtacea (como goiaba e pitanga), e que nesses frutos a taxa de oviposição foi maior do que em frutos de plantas de outras famílias (GREGORIO; SANT'ANA; REDAELLI, 2010). Por outro lado, *A. obliqua* era mais atraída por extratos de plantas da família Anacardiaceae (como manga e seriguela) (CRUZ-LÓPEZ et al., 2006; LÓPEZ-GUILLÉN et al., 2011). Além disso, durante o comportamento de "leks" os machos emitem componentes voláteis específicos à espécie para atrair as fêmeas, que são capazes

de reconhecer aqueles provenientes de indivíduos de sua própria espécie e assim evitar o acasalamento interespecífico (LIMA-MENDONÇA et al., 2014; LÓPEZ-GUILLÉN et al., 2011).

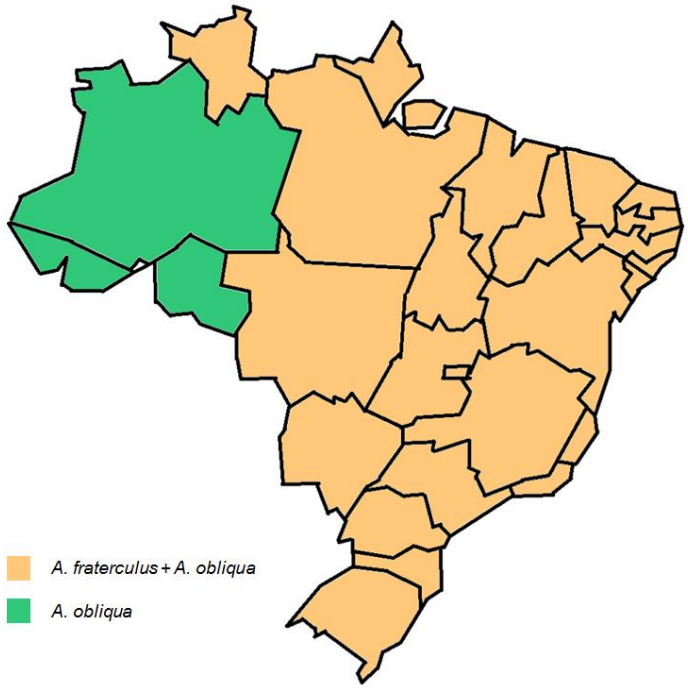


Figura 2. Distribuição das espécies *A. fraterculus* e *A. obliqua* no Brasil. Fonte: www.lea.esalq.usp.br/anastrepha.

As diferenças ecológicas já descritas entre as duas espécies sugerem que genes envolvidos com a manutenção do ciclo circadiano, com a busca por parceiros e com a busca por hospedeiros, podem estar envolvidos na diferenciação entre elas e possivelmente do grupo *fraterculus* como um todo. Genes quimiossensoriais, aqueles que codificam proteínas responsáveis pela percepção de sinais químicos do ambiente, geralmente exercem uma influência ainda maior nos processos de especiação (BENTON, 2006; SÁNCHEZ-GRACIA; ROZAS, 2008). Dessa forma, dentre os genes e fenótipos que poderiam estar envolvidos no processo de especiação do grupo *fraterculus*, esse trabalho investigou em *A. fraterculus* e em *A. obliqua* uma família gênica que codifica proteínas que atuam no sistema olfatório, envolvidas tanto com a recepção de odores, importante pela diferença na escolha de fruto hospedeiro que essas espécies apresentam,

quanto a recepção de ferômonios, que permitem aos indivíduos identificar o parceiro espécie-específico e cujas diferenças podem contribuir com o isolamento reprodutivo, como já foi descrito em espécies de *Drosophila* (COYNE; CRITTENDEN; MAH, 1994).

1.3. O sistema olfatório das moscas-das-frutas

O olfato é responsável pela transformação da informação sobre a estrutura de um odor em uma sucessão de representações neurais, por fim transformadas em uma dada percepção (SU; MENUZ; CARLSON, 2009). É utilizado pelos animais para extrair informações de substâncias químicas voláteis existentes no ambiente, permitindo dessa forma que localizem fontes de alimentos, parceiros reprodutivos e evitem predadores. Os insetos, em geral, exibem um rico repertório de comportamentos olfato-dirigidos, com grande sensibilidade e poder discriminatório mesmo sob o controle de um sistema nervoso relativamente simples (BENTON, 2006).

Em espécies do gênero *Drosophila*, o sistema olfatório é restrito ao terceiro segmento antenal (com cerca de 1200 neurônios olfatórios; ORNs – "olfactory receptor neurons") e às papilas maxilares (com cerca de 120 ORNs). Na superfície dessas estruturas receptoras localizam-se as sensilas, cerdas quimiossensoriais porosas que envolvem os ORNs e onde ocorre o contato primário com os sinais químicos voláteis do ambiente (SHANBHAG; MÜLLER; STEINBRECHT, 1999; STOCKER, 1994; SU; MENUZ; CARLSON, 2009). O terceiro segmento antenal é coberto por três diferentes tipos de sensilas – basicônicas, tricóides e coelocônicas, enquanto as papilas maxilares, estruturas mais simples, apresentam apenas as sensilas basicônicas (VOSSHALL; STOCKER, 2007). Os três tipos diferem em tamanho, morfologia e função. Sensilas basicônicas e coelocônicas respondem aos odores em geral e tricóides aos feromônios, sendo essa divisão de funções aparentemente conservada entre os insetos já estudados (DE BRUYNE; BAKER, 2008).

Cada sensila pode conter de um a quatro ORN, e cada ORN expressa duas proteínas transmembrana que atuam em conjunto para a detecção do odor (SU; MENUZ; CARLSON, 2009; VOSSHALL; STOCKER, 2007). Uma dessas proteínas é efetivamente o receptor de odor (OR – "odorant receptor") no qual os sinais químicos irão se ligar e é o que diferencia os ORNs nas sensilas (BUCK; AXEL, 1991). Em *Drosophila melanogaster*, por exemplo, já foram identificados 62 ORs distintos (ROBERTSON; WARR; CARLSON, 2003). Já a outra proteína, denominada OR83b ou Orco, atua como um co-receptor comum a todos os ORNs, sendo altamente conservada entre os insetos e essencial para a inserção dos ORs na membrana dendrítica (BENTON, 2006; LARSSON et al., 2004; VOSSHALL et al., 1999). O complexo formado pelo OR específico mais OR83b forma um canal iônico dependente de ligantes que resulta na despolarização do ORN o qual está inserido, o que por sua vez ativa as vias de codificação da informação olfatória (WICHER et al., 2008).

ORNs que expressam o mesmo OR convergem seus axônios para um mesmo glomérulo dentre os 50 glomérulos presentes em cada um dos dois lobos antenais de *Drosophila*, que é o órgão dos insetos equivalente ao bulbo olfatório dos mamíferos (BENTON, 2006; LAISSUE et al., 1999). Esses são denominados neurônios sensoriais primários, e a informação olfatória é passada por sinapse aos neurônios sensoriais secundários, ou neurônios de projeção, presentes em torno dos lobos antenais. Os neurônios de projeção fazem então sinapse com os neurônios sensoriais terciários, presentes nos chifres laterais e corpos de cogumelo, que são as estruturas finais da recepção de odores em insetos (LIANG; LUO, 2010), sendo os corpos de cogumelo associados à manutenção das memórias olfativas (HEISENBERG, 2003). Nessas regiões cerebrais, a informação é integrada com informações de outros sentidos (como visão e

paladar) e com experiências passadas para moldar a percepção olfatória e instruir o comportamento a partir dela (SU; MENUZ; CARLSON, 2009).

Para que ocorra a ativação dos ORNs, os sinais químicos voláteis do ambiente, a maioria de caráter hidrofóbico, devem primeiramente ser absorvidos pelos poros das sensilas e passar através de sua linfa aquosa para atingir os ORs (Figura 3) (FAN et al., 2011; SÁNCHEZ-GRACIA; VIEIRA; ROZAS, 2009). Por essa razão, a solubilização e o transporte desses compostos são os passos iniciais da recepção de odores, e as proteínas ligantes a odores (OBPs – "Odorant-binding proteins") são as responsáveis por ambas as funções (PELOSI, 1994; SÁNCHEZ-GRACIA; ROZAS, 2008; VOGT; RIDDIFORD, 1981), além de protegerem esses compostos contra modificações enzimáticas (WANG; HASAN; PIKIELNY, 1999).

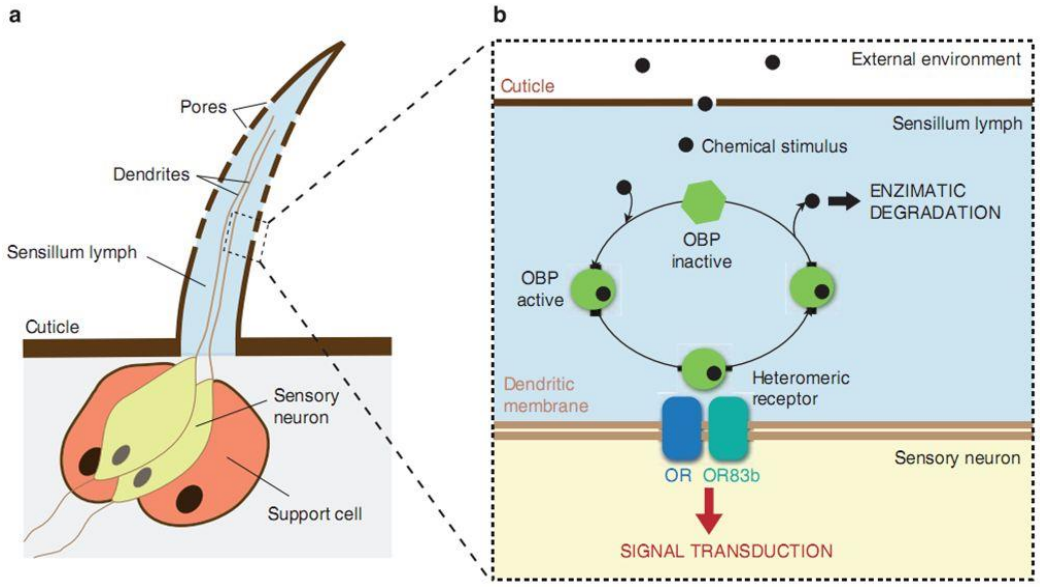


Figura 3. Reconhecimento e início da transdução dos sinais olfatórios em insetos. a) estrutura geral de uma sensila; b) solubilização e transporte dos sinais químicos através da linfa aquosa da sensila. Fonte: SÁNCHEZ-GRACIA e cols. 2009.

As OBPs são pequenas proteínas globulares solúveis (10-30 kDa) sintetizadas por células acessórias aos ORNs e secretadas em grandes quantidades (> 10mM) na linfa aquosa da sensila (PELOSI, 1994), já tendo sido descritas de 41 a 61 OBPs distintas em

espécies de *Drosophila* (GALINDO; SMITH, 2001; HEKMAT-SCAFE et al., 2002; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007). Cada OBP possui diferentes especificidades aos odores e está presente em porções distintas das sensilas (PELOSI; MAIDA, 1995), o que sugere que participam da detecção de odor por restrição do espectro de odores disponíveis em receptores adjacentes, atuando de maneira combinatorial com um certo número de ORs para aumentar o poder discriminatório da percepção de odor (HEKMAT-SCAFE et al., 2002; STEINBRECHT, 1996).

A interação entre OBPs e ORs pode se dar de diferentes formas (LEAL, 2013). Estudos já demonstraram tanto que o complexo OBP/ligante se liga diretamente aos ORs (LAUGHLIN et al., 2008; WANG et al., 2010; XU et al., 2005), como também que as OBPs apenas liberam os odores aos ORS, não participando diretamente da ativação do ORN (MAIDA; ZIEGELBERGER; KAISSLING, 2003; POPHOF, 2004; WOJTASEK; LEAL, 1999). No primeiro caso, a ligação do odor a OBP induz mudanças conformacionais na estrutura terciária, que seria reconhecida por um ORs em específico. Dessa forma, o estímulo do OR se dá não apenas pela molécula de odor, mas pelo complexo OBP/ligante. No segundo caso, essas mudanças conformacionais são pH-dependentes, sendo que o complexo OBP/ligante se mantém unido no pH 6,5 da sensila, mas não no pH 4,9 da superfície da membrana dendrítica, de forma que o odor é assim liberado pela OBP ao OR.

Além de estarem envolvidas na transdução dos sinais olfatórios, as OBPs também participam da inativação dos sinais ao se ligar a eles mas não os transportarem, tornando-os indisponíveis e protegendo os ORs contra saturação em casos de estímulos intensos. Essa função é de grande importância uma vez que a degradação enzimática é muito lenta (cerca de alguns minutos), o que atrasaria em muito a interrupção da recepção do estímulo (KAISSLING, 2001). Outras proteínas também têm sido investigadas quanto à

participação na recepção de odores, como as proteínas sensoriais transmembrana (SNMP – "sensory neuron membrane protein") e as proteínas quimiossensoriais (CSP – "chemosensory proteins"), mas ainda não há um consenso sobre suas funções e se atuam na recepção do olfato ou do paladar (BENTON, 2009**;** CAMPANACCI et al., 2003; PELOSI, 2005; SÁNCHEZ-GRACIA; VIEIRA; ROZAS, 2009).

Diferente dos animais vertebrados que contam com a visão altamente desenvolvida, os insetos dependem bastante do olfato para sua sobrevivência e reprodução (FAN et al., 2011; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007), uma vez que informações olfatórias controlam várias interações sociais e sexuais tanto entre indivíduos da mesma espécie como também entre indivíduos de espécies diferentes. No caso do grupo *fraterculus*, destacamos a importância da detecção dos feromônios liberados pelos machos dentro dos "leks", o que é importante tanto em relação aos machos, para que ocorra a agregação característica do comportamento de corte, quanto em relação às fêmeas, para que consigam identificar parceiros espécie-específicos. Da mesma forma, também destacamos a identificação pelas fêmeas dos frutos hospedeiros para a oviposição. Além de nosso interesse por seu possível papel na especiação do grupo, o sistema olfatório dessas espécies também tem sido estudado com o intuito de aperfeiçoar as estratégias de monitoramento e controle dessas moscas nos pomares, principalmente por meio de sistemas de armadilhas contendo atraentes alimentares ou sexuais (SALLES, 1995).

1.4. Proteínas ligantes a odores (OBPs) e a família gênica que as codificam

O sistema olfatório dos insetos faz uso tanto de proteínas receptoras transmembrana quanto de proteínas ligantes solúveis para ativar os neurônios sensoriais primários e consequentemente ativar a cascata de transdução da percepção de odores. Esse trabalho investigou os genes que codificam as OBPs, por essas proteínas participarem da primeira etapa da recepção olfatória e exercerem influência direta no que

chega até os ORs para ser decodificado. O repertório de genes que codificam para OBPs em insetos constitui uma família multigênica, o que as tornam ainda mais interessantes para o estudo da especiação do grupo *fraterculus*, já que as famílias gênicas, devido à sua rápida evolução, podem desempenhar um importante papel na construção de diferenças morfológicas, fisiológicas e metabólicas entre as espécies (HAHN, 2007).

As sequências de aminoácidos das OBPs são caracterizadas por um pequeno tamanho (de 135 a 220 aminoácidos em moscas-das-frutas), por apresentarem um peptídeo sinal na extremidade N-terminal e serem extremamente divergentes (HEKMAT-SCAFE et al., 2002; SÁNCHEZ-GRACIA; VIEIRA; ROZAS, 2009). Dentro do gênero *Drosophila*, as OBPs que formam o repertório de uma mesma espécie compartilham, em média, apenas 20% de identidade em suas sequências de aminoácidos (HEKMAT-SCAFE et al., 2002), e o mesmo padrão de alta divergência é observado quando se compara ortólogos entre espécies diferentes de insetos (VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007; ZHOU; FIELD; HE, 2010). A grande diversidade encontrada entre essas proteínas está associada a capacidade de distinção entre vários diferentes ligantes, e deve ser resultado de evolução adaptativa aos mais variados odores (FAN et al., 2011; HEKMAT-SCAFE et al., 2002).

As OBPs são constituídas de seis alfa-hélices unidas por "loops" de tamanhos variados, sendo a estrutura terciária compacta devido à presença de três pontes dissulfeto formadas entre seis resíduos de cisteínas (C1↔C3; C2↔C5; C4↔C6) (GALINDO; SMITH, 2001; VOGT; RIDDIFORD, 1981). Essas cisteínas são encontradas em posições altamente conservadas em todas as OBPs de todos os insetos estudados, o que se tornou um padrão característico pelo qual essas proteínas podem ser reconhecidas (CALVELLO et al., 2003; FAN et al., 2011; SCALONI et al., 1999; TEGONI; CAMPANACCI; CAMBILLAU, 2004). Os resíduos de cisteínas, associados com informações funcionais

e relações filogenéticas, também têm sido usados para classificar essas proteínas em subfamílias. Em *Drosophila*, as OBPs são divididas em quatro subfamílias: classic (OBPs com seis resíduos de cisteínas), plus-C (mais de seis cisteínas), minus-C (alguns membros apresentam menos de seis cisteínas) e dimer (OBPs com dois domínios completos de seis cisteínas) (HEKMAT-SCAFE et al., 2002; VIEIRA et al., 2012; ZHOU et al., 2004).

A conservação na posição das cisteínas das OBPs indica que as pontes dissulfeto são essenciais para a estrutura terciária ser funcional (LEAL, 2013; SCALONI et al., 1999; VIEIRA; ROZAS, 2011; VOGT; RIDDIFORD, 1981). Também já foi demonstrado que elas ajudam a definir a cavidade ligante hidrofóbica existente nessas proteínas, que é onde os odores se ligam (LAGARDE et al., 2011; LEITE et al., 2009; TSITSANOU et al., 2013). Um alto grau de conservação da estrutura terciária é notado entre os insetos, mas diferentes OBPs mostram grande versatilidade e habilidade para acomodar ligantes de diferentes formas e estruturas químicas (TEGONI; CAMPANACCI; CAMBILLAU, 2004), uma vez que diferentes aminoácidos em uma mesma posição contribuem para mudanças tanto na estrutura quanto na posição da cavidade ligante das OBPs (ARYA et al., 2010; VENTHUR et al., 2014).

A família multigênica que codifica as OBPs é composta por um número variável de membros, já tendo sido descritos desde 11 genes na lagarta-da-beterraba *Spodoptera exigua* (ZHU et al., 2013) até mais de 100 no mosquito *Aedes aegypti* (MANOHARAN et al., 2013). *D. melanogaster* apresenta 51 genes e um pseudogene, que apresentam entre zero e três íntrons, a maioria dos quais localizados em posições conservadas entre os diferentes genes (HEKMAT-SCAFE et al., 2002). Além da baixa similaridade das sequências gênicas dos membros, o que sugere que ocorre uma rápida evolução dessas sequências (FORÊT; MALESZKA, 2006; SÁNCHEZ-GRACIA; ROZAS, 2008), uma outra característica dessa família gênica nos insetos é a localização em “clusters” gênicos

no genoma, o que sugere que esses genes surgiram por duplicação em tandem (HEKMAT-SCAFE et al., 2002; PELOSI, 2005).

Em 12 espécies de *Drosophila* analisadas (espécies com cinco pares de cromossomos, quatro autossomos e um sexual), os genomas mostraram um preservado arranjo de *OBPs*, nas quais 37 dos 51 genes formavam "clusters" de quatro a nove genes cada. Um "cluster" localizava-se no cromossomo sexual, e dois dos "clusters" continham também um gene que codifica um OR. Essa conservação observada ao longo de 80 milhões de anos de evolução entre essas espécies sugere a ação de seleção natural prevenindo a quebra desses "clusters" (VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007). Além disso, essas espécies também compartilham elementos regulatórios em comum para a expressão das OBPs (GONG et al., 2007; MATSUO et al., 2007).

A família gênica das OBPs é bastante antiga, com seu ancestral comum mais recente (MRCA – "most recent common ancestor") tendo surgido, no mínimo, desde a origem dos insetos. A maioria dos genes da família tem um único sítio de "splice", localizado logo depois da sequência sinal, o que deve ser reflexo de um ancestral comum a esses genes (GALINDO; SMITH, 2001; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007). O principal mecanismo de evolução dessa família gênica é o modelo de "birth-and-death", no qual novos membros são formados por duplicação gênica, divergindo progressivamente em sequência e função, podendo também ser perdidos por uma deleção ou por processos de pseudogenização (NEI; ROONEY, 2005; SÁNCHEZ-GRACIA; VIEIRA; ROZAS, 2009). Em espécies de *Drosophila*, as taxas de surgimento ("birth") de novos genes *OBP* são muito mais altas que às do resto do genoma, refletindo a evolução altamente dinâmica dessa família gênica (HAHN, 2007; VIEIRA; ROZAS, 2011). Mudanças adaptativas promovidas potencialmente por mudanças ambientais também são frequentemente observadas na evolução das famílias quimiossensoriais nos

insetos, envolvendo principalmente características reprodutivas, ecológicas e comportamentais (WILLET, 2000). Devido à alta diversidade molecular encontrada nos genes OBP dos insetos, bem como à sua rápida taxa evolutiva, essa família gênica é considerada um alvo muito promissor para o estudo dos processos evolutivos envolvidos em eventos de especiação recente (SIEGLAFF et al., 2009; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007).

1.5. Objetivos

O objetivo desse trabalho foi avaliar o padrão de evolução molecular da família gênica das OBPs em *A. fraterculus* e em *A. obliqua*, bem como determinar o padrão de expressão desses genes durante as diferentes fases da vida adulta de machos e fêmeas, a fim de determinar o quanto da diversidade encontrada poderia estar envolvido na diferenciação dessas espécies.

Para isso, os seguintes objetivos específicos foram buscados:

1) Identificar genes da família das OBPs em bibliotecas de cDNA de tecidos cefálico e reprodutivo de machos e fêmeas de *A. fraterculus* e de *A. obliqua*;

2) Analisar a dinâmica das forças evolutivas envolvidas com os padrões de evolução molecular da família gênica das OBPs;

3) Averiguar se os genes OBPs estão sendo diferencialmente expressos entre *A. fraterculus* e *A. obliqua*, por análises "in silico" a partir dos transcriptomas, e por PCR em tempo real;

4) Inferir sobre o papel das OBPs na diferenciação de *A. fraterculus* e *A. obliqua* e do grupo *fraterculus*.

1.6. Referências Bibliográficas

ALUJA, M. The study of movements in Tephritidae fruit flies: review of concepts and recent advances. In: ALUJA, M.; LIEDO, P. (Ed.). **Fruit Flies: biology and management**. New York: Springer-Verlag, 1993. p. 105–113.

ALUJA, M. et al. Behavior of flies in the genus *Anastrepha* (Trypetinae: Toxotripanini). In: ALUJA, M.; NORRBOM, A.L. **Fruit flies (Tephritidae): phylogeny and evolution of behavior**. CRC Press, Washington, DC, USA, 2000. p. 375–408.

ARAUJO, E.; ZUCCHI, R. Medidas do acúleo na caracterização de cinco espécies de *Anastrepha* do grupo *fraterculus* (Diptera: Tephritidae). **Systematics, Morphology and Phisiology**, v. 35, n. 3, p. 329–337, 2006.

ARYA, G. H. et al. Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*. **Genetics**, v. 186, n. 4, p. 1475–1485, 2010.

BENTON, R. On the ORigin of smell: odorant receptors in insects. **Cellular and Molecular Life Sciences**, v. 63, n. 14, p. 1579–1585, 2006.

BENTON, R. Molecular basis of odor detection in insects. **Annals of the New York Academy of Sciences**, v. 1170, n. 1, p. 478–481, 2009.

BUCK, L.; AXEL, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. **Cell**, v. 65, n. 1, p. 175–187, 1991.

CALVELLO, M. et al. Soluble proteins of chemical communication in the social wasp Polistes dominulus. **Cellular and molecular life sciences: CMLS**, v. 60, n. 9, p. 1933–1943, set. 2003.

CAMPANACCI, V. et al. Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 9, p. 5069–5074, 2003.

COYNE, J. A.; CRITTENDEN, A. P.; MAH, K. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. **Science (New York, N.Y.)**, v. 265, n. 5177, p. 1461–1464, 1994.

CRUZ, I. et al. Morfologia do aparelho reprodutor e biologia do desenvolvimento. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 55–66.

CRUZ-LÓPEZ, L. et al. A new potential attractant for *Anastrepha obliqua* from *Spondias mombin* fruits. **Journal of Chemical Ecology**, v. 32, n. 2, p. 351–365, 2006.

DE BRUYNE, M.; BAKER, T. C. Odor detection in insects: volatile codes. **Journal of Chemical Ecology**, v. 34, n. 7, p. 882–897, 2008.

DUARTE, A.; MALAVASI, A. Tratamentos quarentenários. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 187–192.

FAN, J. et al. An overview of odorant-binding protein functions in insect peripheral olfactory reception. **Genetics and molecular research: GMR**, v. 10, n. 4, p. 3056–3069, 2011.

FERNANDES, F. **Análise multilocus de parâmetros populacionais, evolução molecular e diferenciação em espécies de moscas-dasfrutas do grupo fraterculus (Diptera, Tephritidae).** Dissertação de Mestrado, Universidade de São Carlos (UFSCar), 2010.

FORÊT, S.; MALESZKA, R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). **Genome research**, v. 16, n. 11, p. 1404–1413, 2006.

FRITZ, A. H.; SCHABLE, N. Microsatellite loci from the Caribbean Fruit Fly, *Anastrepha suspensa* (Diptera: Tephritidae). **Molecular Ecology Notes**, v. 4, n. 3, p. 443–445, 2004.

GALINDO, K.; SMITH, D. P. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. **Genetics**, v. 159, n. 3, p. 1059–1072, 2001.

GAVRILETS, S. Models of Speciation: Where Are We Now? **Journal of Heredity**, v. 105, n. S1, p. 743–755, 2014.

GONG, D.-P. et al. Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx mori*. **Insect Biochemistry and Molecular Biology**, v. 37, n. 3, p. 266–277, 2007.

GREGORIO, P. L. F.; SANT'ANA, J.; REDAELLI, L. R. Percepção química e visual de *Anastrepha fraterculus* (Diptera, Tephritidae) em laboratório. **Iheringia. Série Zoologia**, v. 100, n. 2, 2010.

HAHN, M. W. Bias in phylogenetic tree reconciliation methods: implications for vertebrate genome evolution. **Genome Biology**, v. 8, n. 7, p. R141, 2007.

HEISENBERG, M. Mushroom body memoir: from maps to models. **Nature Reviews Neuroscience**, v. 4, n. 4, p. 266–275, 2003.

HEKMAT-SCAFE, D. S. et al. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. **Genome research**, v. 12, n. 9, p. 1357–1369, 2002.

HENNING, F.; MATIOLI, S. R. Mating time of the West Indian fruit fly *Anastrepha obliqua* (Macquart) (Diptera: Tephritidae) under laboratory conditions. **Neotropical Entomology**, v. 35, n. 1, p. 145–148, 2006.

KAISSLING, K.-E. Olfactory Perireceptor and Receptor Events in Moths: A Kinetic Model. **Chemical Senses**, v. 26, n. 2, p. 125–150, 2001.

KASPI, R.; YUVAL, B. Lek site selection by male Mediterranean Fruit Flies. **Journal of Insect Behavior**, v. 12, n. 2, p. 267–276, 1999.

LAGARDE, A. et al. The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. **Journal of molecular biology**, v. 414, n. 3, p. 401–412, 2011.

LAISSUE, P. P. et al. Three-dimensional reconstruction of the antennal lobe in *Drosophila melanogaster*. **The Journal of Comparative Neurology**, v. 405, n. 4, p. 543–552, 1999.

LARSSON, M. C. et al. Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. **Neuron**, v. 43, n. 5, p. 703–714, 2004.

LAUGHLIN, J. D. et al. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. **Cell**, v. 133, n. 7, p. 1255–1265, 2008.

LEAL, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. **Annual Review of Entomology**, v. 58, p. 373–391, 2013.

LEITE, N. R. et al. Structure of an Odorant-binding protein from the mosquito *Aedes aegypti* suggests a binding pocket covered by a pH-sensitive "Lid". **PLoS ONE**, v. 4, n. 11, p. e8006, 2009.

LIANG, L.; LUO, L. The olfactory circuit of the fruit fly *Drosophila melanogaster*. **Science China. Life Sciences**, v. 53, n. 4, p. 472–484, 2010.

LIMA. **Evolução Molecular, análise multilocus e diferenciação entre espécies do grupo *fraterculus* (Diptera: Tephritidae).** Dissertação de mestrado (Genética e Evolução). Universidade Federal de São Carlos, 2011.

LIMA-MENDONÇA, A. et al. Semiochemicals of fruit flies of the genus *Anastrepha*. **Química Nova**, v. 37, n. 2, 2014.

LÓPEZ-GUILLÉN, G. et al. Olfactory responses of *Anastrepha obliqua* (Diptera: Tephritidae) to volatiles emitted by calling males. **Florida Entomologist**, v. 94, n. 4, p. 874–881, 2011.

MAIDA, R.; ZIEGELBERGER, G.; KAISSLING, K.-E. Ligand binding to six recombinant pheromone-binding proteins of *Antheraea polyphemus* and *Antheraea pernyi*. **Journal of Comparative Physiology. B, Biochemical, Systemic, and Environmental Physiology**, v. 173, n. 7, p. 565–573, 2003.

MALAVASI, A.; ZUCHHI, R.; SUGAYAMA, R. Biogeografia. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 93–98.

MANOHARAN, M. et al. Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. **Genome Biology and Evolution**, v. 5, n. 1, p. 163–180, 2013.

MATSUO, T. et al. Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. **PLoS Biology**, v. 5, n. 5, p. e118, 2007.

MONTOYA, P. et al. Biological control of *Anastrepha* spp. (Diptera: Tephritidae) in mango rorchards through augmentative releases of *Diachasmimorpha longicaudata* (Ashmead) (Hymenoptera: Braconidae). **Biological Control**, v. 18, n. 3, p. 216–224, 2000.

NEI, M.; ROONEY, A. P. Concerted and Birth-and-death evolution of multigene families. **Annual Review of Genetics**, v. 39, n. 1, p. 121–152, 2005.

OVIEDO, A. et al. Management of protein intake in the fruit fly *Anastrepha fraterculus*. **Journal of Insect Physiology**, v. 57, n. 12, p. 1622–1630, 2011.

PELOSI, P. Odorant-binding proteins. **Critical Reviews in Biochemistry and Molecular Biology**, v. 29, n. 3, p. 199–228, 1994.

PELOSI, P. Diversity of Odorant-binding proteins and chemosensory proteins in insects. **Chemical Senses**, v. 30, n. Supplement 1, p. i291–i292, 2005.

PELOSI, P.; MAIDA, R. Odorant-binding proteins in insects. **Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology**, v. 111, n. 3, p. 503–514, 1995.

PERRE, P. et al. Morphometric differentiation of fruit fly pest species of the *Anastrepha fraterculus* group (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 107, n. 2, p. 490–495, 2014.

POPHOF, B. Pheromone-binding proteins contribute to the activation of olfactory receptor neurons in the silkmoths *Antheraea polyphemus* and *Bombyx mori*. **Chemical Senses**, v. 29, n. 2, p. 117–125, 2004.

RIDLEY, M. Especiação. In: RIDLEY, M., **Evolução**. Porto Alegre: Artmed, 2006. p. 408–441.

RITCHIE, M. G. Sexual selection and speciation. **Annual Review of Ecology, Evolution, and Systematics**, v. 38, n. 1, p. 79–102, 2007.

ROBERTSON, H. M.; WARR, C. G.; CARLSON, J. R. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100 Suppl 2, p. 14537–14542, 2003.

SALLES, L. **Bioecologia e controle da mosca-das-frutas sul-americana**. Pelotas: EMBRAPA – CPACT, 1995.

SALLES, L. Biologia e ciclo de vida de *Anastrepha fraterculus* (Wied.). In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 81–86.

SÁNCHEZ-GRACIA, A.; ROZAS, J. Divergent evolution and molecular adaptation in the *Drosophila* odorant-binding protein family: inferences from sequence variation at the OS-E and OS-F genes. **BMC evolutionary biology**, v. 8, p. 323, 2008.

SÁNCHEZ-GRACIA, A.; VIEIRA, F. G.; ROZAS, J. Molecular evolution of the major chemosensory gene families in insects. **Heredity**, v. 103, n. 3, p. 208–216, 2009.

SANTOS, P.; URAMOTO, K.; MATIOLI, S. R. Experimental hybridization among *Anastrepha* species (Diptera: Tephritidae): production and morphological characterization of hybrids. **Annals of the Entomological Society of America**, v. 94, n. 5, p. 717–725, 2001.

SCALONI, A. et al. Structural analysis and disulfide-bridge pairing of two Odorant-binding proteins from *Bombyx mori*. **Biochemical and Biophysical Research Communications**, v. 266, n. 2, p. 386–391, 1999.

SEGURA, D. F. et al. Lekking behavior of *Anastrepha fraterculus* (Diptera: Tephritidae). **The Florida Entomologist**, v. 90, n. 1, p. 154–162, 2007.

SEGURA, D. F. et al. Methoprene treatment reduces the pre-copulatory period in *Anastrepha fraterculus* (Diptera: Tephritidae) sterile males: Methoprene accelerates sexual onset in *A. fraterculus*. **Journal of Applied Entomology**, v. 137, p. 19–29, 2013.

SELIVON, D. Relações com as plantas hospedeiras. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 87–91.

SELIVON, D.; PERONDINI, A. L. P.; MORGANTE, J. S. A Genetic–morphological characterization of two cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 98, n. 3, p. 367–381, 2005.

SHANBHAG, S.; MÜLLER, B.; STEINBRECHT, R. Atlas of olfactory organs of *Drosophila melanogaster*. **International Journal of Insect Morphology and Embryology**, v. 28, n. 4, p. 377–397, 1999.

SIEGLAFF, D. H. et al. Comparative genomics allows the discovery of cis-regulatory elements in mosquitoes. **Proceedings of the National Academy of Sciences**, v. 106, n. 9, p. 3053–3058, 2009.

SIVINSKI, J.; DODSON, G. Sexual dimorphism in *Anastrepha suspensa* (Loew) and other tephritid fruit flies (Diptera: Tephritidae): Possible roles of developmental rate, fecundity, and dispersal. **Journal of Insect Behavior**, v. 5, n. 4, p. 491–506, 1992.

SMITH-CALDAS, M. R. B. et al. Phylogenetic relationships among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. **Neotropical Entomology**, v. 30, n. 4, p. 565–573, 2001.

STEINBRECHT, R. A. Are Odorant-binding proteins involved in odorant discrimination? **Chemical Senses**, v. 21, n. 6, p. 719–727, 1996.

STOCKER, R. F. The organization of the chemosensory system in *Drosophila melanogaster*: a review. **Cell and Tissue Research**, v. 275, n. 1, p. 3–26, 1994.

SU, C.-Y.; MENUZ, K.; CARLSON, J. R. Olfactory perception: Receptors, cells, and circuits. **Cell**, v. 139, n. 1, p. 45–59, 2009.

SUGAYAMA, R.; MALAVASI, A. Ecologia comportamental. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 103–108.

TEGONI, M.; CAMPANACCI, V.; CAMBILLAU, C. Structural aspects of sexual attraction and chemical communication in insects. **Trends in Biochemical Sciences**, v. 29, n. 5, p. 257–264, 2004.

TSITSANOU, K. E. et al. Crystal and solution studies of the "Plus-C" odorant-binding protein 48 from *Anopheles gambiae*: control of binding specificity through three-dimensional domain swapping. **The Journal of biological chemistry**, v. 288, n. 46, p. 33427–33438, 2013.

URAMOTO, K.; WALDER, J. M. M.; ZUCCHI, R. A. Biodiversidade de moscas-das-frutas do gênero *Anastrepha* (Diptera, Tephritidae) no campus da ESALQ-USP, Piracicaba, São Paulo. **Revista Brasileira de Entomologia**, v. 48, n. 3, 2004.

VENTHUR, H. et al. Ligand binding and homology modelling of insect odorant-binding proteins. **Physiological Entomology**, v. 39, n. 3, p. 183–198, 2014.

VIEIRA, F. G. et al. Unique features of Odorant-binding proteins of the parasitoid wasp *Nasonia vitripennis* revealed by genome annotation and comparative analyses. **PLoS ONE**, v. 7, n. 8, p. e43034, 2012.

VIEIRA, F. G.; ROZAS, J. 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**, v. 3, p. 476–490, 2011.

VIEIRA, F. G.; SÁNCHEZ-GRACIA, A.; ROZAS, J. Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. **Genome Biology**, v. 8, n. 11, p. R235, 2007.

VILELA, E.; KOVALESKI, A. Feromônios. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 99–102.

VOGT, R. G.; RIDDIFORD, L. M. Pheromone binding and inactivation by moth antennae. **Nature**, v. 293, n. 5828, p. 161–163, 1981.

VOSSHALL, L. B. et al. A spatial map of olfactory receptor expression in the *Drosophila* antenna. **Cell**, v. 96, n. 5, p. 725–736, 1999.

VOSSHALL, L. B.; STOCKER, R. F. Molecular architecture of smell and taste in *Drosophila*. **Annual Review of Neuroscience**, v. 30, p. 505–533, 2007.

WANG, P. et al. Natural variation in odorant recognition among Odorant-binding proteins in *Drosophila melanogaster*. **Genetics**, v. 184, n. 3, p. 759–767, 2010.

WANG, Q.; HASAN, G.; PIKIELNY, C. W. Preferential expression of biotransformation enzymes in the olfactory organs of *Drosophila melanogaster*, the antennae. **Journal of Biological Chemistry**, v. 274, n. 15, p. 10309–10315, 1999.

WICHER, D. et al. *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. **Nature**, v. 452, n. 7190, p. 1007–1011, 2008.

WILLET, C. S. Evidence for directional selection acting on Pheromone-binding Proteins in the genus *Choristoneura*. **Molecular Biology and Evolution**, v. 17, n. 4, p. 553–562, 2000.

WOJTASEK, H.; LEAL, W. S. Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. **The Journal of Biological Chemistry**, v. 274, n. 43, p. 30950–30956, 1999.

WU, C.-I. The genic view of the process of speciation: Genic view of the process of speciation. **Journal of Evolutionary Biology**, v. 14, n. 6, p. 851–865, 2001.

XU, P. et al. *Drosophila* OBP LUSH is required for aof pheromone-sensitive neurons. **Neuron**, v. 45, n. 2, p. 193–200, 2005.

ZHOU, J.-J. et al. Revisiting the odorant-binding protein LUSH of *Drosophila melanogaster*: evidence for odour recognition and discrimination. **FEBS letters**, v. 558, n. 1-3, p. 23–26, 2004.

ZHOU, J.-J.; FIELD, L. M.; HE, X. L. Insect Odorant-binding proteins: Do they offer an alternative pest control strategy? **Outlooks on Pest Management**, b. v. 21, n. 1, p. 31–34, 2010.

ZHU, J.-Y. et al. Identification and tissue distribution of odorant binding protein genes in the beet armyworm, *Spodoptera exigua*. **Journal of Insect Physiology**, v. 59, n. 7, p. 722–728, 2013.

ZUCCHI, R. A. Taxonomia. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000a. p. 10–24.

ZUCCHI, R. A. Lista das espécies de *Anastrepha*, sinonímias, plantas hospedeiras e parasitoides. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000b. p. 41–48.

ZUCCHI, R. A. **Fruit flies in Brazil - *Anastrepha* species their host plants and parasitoids**, 2008. Disponível em: <www.lea.esalq.usp.br/anastrepha>. Acesso em: 16 nov. 2015.

ZUCOLOTO, F. Alimentação e nutrição de moscas-das-frutas. In: MALAVASI, A.; ZUCCHI, R.A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil**. Holos Editora, Ribeirão Preto, 2000. p. 67–80.

Capítulo II

Identificação de membros da família gênica das OBPs em transcriptomas de *Anastrepha fraterculus* e *Anastrepha obliqua* e descoberta de genes sob seleção positiva

Capítulo II – Identificação de membros da família gênica das OBPs em transcriptomas de *Anastrepha fraterculus* e *Anastrepha obliqua* e descoberta de genes sob seleção positiva

---

Considerações Iniciais

Uma vez que o estudo das diferenças moleculares entre espécies requer a identificação de genes e/ou alelos que estão fixados entre diferentes linhagens, investigamos o número e as diferenças nos genes OBPs entre *A. fraterculus* e *A. obliqua*, bem como suas relações filogenéticas com parálogos de *D. melanogaster* e de outros insetos pertencentes à ordem Tephritidae. No entanto, como as diferenças moleculares observadas entre linhagens podem simplesmente ter surgido aleatoriamente, não estando portanto envolvidas com novas mudanças ou adaptações dirigidas por seleção, faz-se necessário investigar se há evidência de seleção positiva agindo sobre eles. Dessa forma, a primeira parte desse trabalho consistiu na análise dos membros da família gênica das OBPs e na compreensão dos processos que têm moldado a evolução desses genes nas duas espécies de *Anastrepha* estudadas. Tais dados serão apresentados a seguir em forma de manuscrito científico[1] que já foi submetido à revista "BCM Evolutionary Biology".

[1] **CAMPANINI, EB**; DE BRITO, RA. Molecular evolution of Odorant-binding proteins gene family in two closely related *Anastrepha* fruit flies. BMC Evol Biol, submitted.

Molecular evolution of Odorant-binding proteins gene family in two closely related *Anastrepha* fruit flies

2.1. Abstract

Background: Odorant-binding proteins (OBPs) are of great importance for survival and reproduction since they participate in initial steps of the olfactory signal transduction cascade, solubilizing and transporting chemical signals to the olfactory receptors. A comparative analysis of OBPs between closely related species may help explain how these genes evolve and are maintained under natural selection and how differences in these proteins can affect olfactory responses. We studied OBP genes in the closely related species *Anastrepha fraterculus* and *A. obliqua*, which have different host preferences, using data from RNA-seq cDNA libraries of head and reproductive tissues from male and female adults, aiming to understand the speciation process occurred between them.

Results: We identified 23 different OBP sequences from *Anastrepha fraterculus* and 24 from *A. obliqua*, which correspond to 20 *Drosophila melanogaster* OBP genes. Phylogenetic analysis separated *Anastrepha* OBPs sequences in four branches that represent four subfamilies: classic, minus-C, plus-C and dimer. Both species showed five plus-C members, which is the biggest number found in tephritids until now. We found evidence of positive selection in four genes and at least one duplication event that preceded the speciation of these two species. Inferences on tertiary structures of putative proteins from these genes revealed that at least one positively selected change involves the binding cavity (the odorant binding region) in the plus-C *OBP50a*.

Conclusions: *A. fraterculus* and *A. obliqua* have a bigger OBP repertoire than the others tephritids studied, though the total number of *Anastrepha* OBPs may be larger, since we studied only a limited number of tissues. The contrast of these closely related species reveals that there are several amino acid changes between the homologous genes, which might be related to their host preferences. The plus-C OBP that has one amino acid under positive selection located in the binding cavity may be under a selection pressure to recognize and bind a new odorant. The other positively selected sites found may be involved in important structural and functional changes, especially that ones in which site-specific changes would radically change amino acid properties.

Keywords: Sister species; recent speciation event; phylogenetic analyses; OBP subfamilies; OBP putative tertiary structure.

2.2. Background

The study of the genetics of species differences generally requires the identification of fixed genes between lineages. Most of these genes are simply "ordinary" traits that diverged between lineages with no direct role on isolation (ORR; IRVING, 2001) while some could be involved with new changes or adaptations driven by selection. Up until recently, studies on speciation would then have to investigate these genes by interspecific crosses or other means to assess reproductive isolation (NOOR; FEDER, 2006) but new technical advances have enabled their identification by looking at selection at the gene level. Genes involved in reproduction and mate choice tend to evolve rapidly and display signatures of positive selection (RICE, 1992; SWANSON; VACQUIER, 2002; VANDEWEGE et al., 2013), as well as genes facilitating chemoreception (FORÊT; MALESZKA, 2006; GARDINER et al., 2008; KOSIOL et al., 2008; SÁNCHEZ-GRACIA; ROZAS, 2008), because they are needed to interpret information about the environment, such as the presence of food or predators. Olfactory information, specifically, also controls social and sexual interactions between individuals of the same species, wherein the detection of odors and pheromones essential for survival and reproduction (BENTON, 2006). In insects, the solubilization and transport of chemical signals through the aqueous lymph of insect's sensilla to the olfactory receptors is the initial step of the transduction cascade of olfactory signals, and is mediated by the Odorant-binding proteins – OBPs (VOGT; RIDDIFORD, 1981).

OBPs are small soluble globular proteins (10-30 kDa), with six highly conserved cysteine residues (referred to as C1 to C6) in characteristic positions of all known insect OBPs. These cysteines form three disulphide bonds that stabilize their tertiary structure (PELOSI; MAIDA, 1995) and help define a hydrophobic binding cavity (LAGARDE et al., 2011; LEITE et al., 2009; TSITSANOU et al., 2013). Different OBPs have specific affinities to odors and are present in distinctive portions of the sensilla, suggesting that

they participates in odor detection by restricting the available spectrum of odors into adjacent receptors (PELOSI; MAIDA, 1995). OBPs' conserved cysteine residues, associated with functional information and phylogenetic relationships, have been used to classify OBPs into subfamilies: classic OBPs (six cysteines), plus-C OBPs (more than six cysteines) and minus-C OBPs (some members with less than six cysteines) (HEKMAT-SCAFE et al., 2002; VIEIRA et al., 2012; ZHOU et al., 2004). Additionally, OBP transcripts encoding two complete OBP domains were identified in *Drosophila melanogaster*, and classified as a subfamily named dimer (VIEIRA et al., 2012) It is expected that more generalist species would have a larger OBP repertoire to recognize various plant chemicals, in contrast to more specialist species, in which the power of natural selection to maintain the whole OBP repertoire may have been weaker (OHTA et al., 2014).

OBP genes are part of a gene family with low sequence similarity among its members, but high conservation at the structural level (TEGONI; CAMPANACCI; CAMBILLAU, 2004). OBP genes are located in large gene clusters, suggesting that they arose by tandem duplication (GALINDO; SMITH, 2001) and evolve mainly via the birth-and-death model (NEI; ROONEY, 2005), in which newly duplicated members progressively diverge in sequence and function, or may be lost to deletion or pseudogenization. The high disparity of gene sequences among OBPs implies a rapid rate of evolution in this gene family, suggesting that these genes might have evolved under the influence of positive selection (FORÊT; MALESZKA, 2006; SÁNCHEZ-GRACIA; AGUADÉ; ROZAS, 2003; SÁNCHEZ-GRACIA; ROZAS, 2008), whereby even the conserved cysteine residues may be lost (VIEIRA et al., 2012).

Here, we studied OBP genes in two closely related species of Tephritidae, *Anastrepha fraterculus* and *A. obliqua*, which are important fruit pests in South America.

These species have diverged recently and exhibit a limited number of morphologic (PERRE et al., 2014) and genetic distinguishing characters (SELIVON; PERONDINI; MORGANTE, 2005). Though *A. fraterculus* has been associated with a wide number of hosts, it prefers several Myrtaceae fruits (PEREIRA-RÊGO et al., 2011), being considered one of the main economic pests in South America. *A. obliqua*, on the other hand, though an important pest species as well, has been associated to a smaller number of hosts, several of those Anacardiaceae (CAMARGO; ODELL; JIRÓN, 1996). Because OBPs are important targets for natural and sexual selection, their role in host and mate choice has previously been established in several species (LAUGHLIN et al., 2008; LI et al., 2013; MATSUO et al., 2007). The investigation of OBP evolution in these closely related species may provide clues about this group's differentiation and host preference. We identified *A. fraterculus* and *A. obliqua* OBP members using RNA-seq data from reproductive and cephalic tissues of several different reproductive stages in male and female adults: before and after copulation and females after oviposition, and analyzed the patterns of molecular evolution between these two species. Several OBP genes were evolving under positive selection and we speculate how these amino acid changes would affect protein structure and their consequences for adaptation.

2.3. Methods

2.3.1. Transcriptome libraries, assembly and annotation

We used populations of *A. fraterculus* from Southeastern (22º 01' 03" S, 47 º 53' 27" W) and of *A. obliqua* from Midwest (16º 41' 58" S, 49 º 16' 35" W) regions of Brazil, that have been maintained in a controlled environment room at 25ºC ± 5ºC (60–90% humidity), and natural photoperiod. Transcriptome libraries were made for each species separately, using two reproductive stages for both genders (virgin and post-mating), and a third one for females (post-oviposition). All profiles were made with biological replication, totaling 10 reproductive profiles per species. For each profile, we extracted

the total RNA from head and reproductive tissues, totaling 20 libraries for each species. We used a pool of 10 individuals per library, amounting to two hundred individuals in total.

More details on sample preparation, molecular procedures, assembly and annotation of the RNA-seq data are described elsewhere[2]. Briefly, total RNA was extracted using the TRizol/chloroform protocol (CHOMCZYNSKI; MACKEY, 1995). RNA-seq libraries were constructed from four µg of total RNA using the TruSeq Stranded Total RNA Sample Prep Kit (Illumina) protocol, according to the manufacturer's instructions. Pools of 12 libraries were run on an Illumina HiSeq2000 on a lane with runs of 2 x 100 bp paired-end reads. All reads were trimmed for quality and length with SeqyClean (https://github.com/ibest/seqyclean), keeping only reads with a minimum sequence length of 50 pb, a minimum of 0.01 for the parameter 'max-avg-error' and 0.05 for 'max-error-at-ends', and an average Phred quality score ≥ 20. Processed reads were assembled in two independent transcriptomes, one for each species, using the Trinity short read assembler (release 2013-02-25) (GRABHERR et al., 2011), using default parameters.

We identified *A. fraterculus* and *A. obliqua* OBP sequences in the assemblies by BLASTx searches (ALTSCHUL et al., 1990), using the Gene Ontology and *D. melanogaster* database as a reference. Only the first match was considered, and contigs with an e-value threshold of 10^{-5} that corresponded to OBP genes were retained and their open reading frames inferred using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf). We used PrediSi (HILLER et al., 2004) to identify signal peptide sequences, since this is an important attribute of OBP proteins. OBPs of *A. fraterculus* and *A. obliqua* were

[2] REZENDE, VB; CASTILLO, CC.; LIMA, ALA; **CAMPANINI, EB**; NAKAMURA, AM; OLIVEIRA, JL; CHAHAD-EHLERS, S; SOBRINHO-JUNIOR, I; DE BRITO, RA. Candidate genes involved in differentiation between two non-model species of fruit flies (*Anastrepha*: Tephritidae) screened from head transcriptomes. G3, submitted.

designated as Afra and Aobl respectively, followed by the name of the respective OBP, based on their similarity with *D. melanogaster* OBPs (Table 1). Sequences associated with the same OBP were differentiated with a numerical postscript.

2.3.2. Alignments and phylogenetic analysis

We initially performed a multiple alignment of the inferred OBP amino acid sequences for each species with MAFFT (KATOH; STANDLEY, 2013) using default settings. Average amino acid identities were obtained from the Percent Identity Matrix in the MAFFT alignment results. The multiple alignment of nucleotide sequences was manually adjusted in Bioedit v.7.0.5.3 (HALL, 1999), using the amino acid alignment as a guide. Three different methods, MaxChi (SMITH, 1992), GENECONV (SAWYER, 1989) and RDP (MARTIN; RYBICKI, 2000), were implemented in the RDP3 program (MARTIN et al., 2010) to investigate for recombination events.

Because OBPs are very divergent, we inferred the phylogenetic relationships in two steps. First, we aligned all *A. fraterculus* and *A. obliqua* OBPs here identified with OBP sequences of *Ceratitis capitata* and of *D. melanogaster* (obtained from GenBank - Additional file 1), using MAFFT as previously described. These species were chosen because *D. melanogaster* has the best curated genome and *C. capitata* is the closest species to *Anastrepha* with OBP sequences available. We used jModelTest (POSADA, 2008) to estimate the best-fitting nucleotide model of substitution, that was used to infer Maximum likelihood phylogenetic relationships among OBPs with PhyML ver.3.0 (GUINDON et al., 2010). This first phylogenetic tree was reconstructed in order to corroborate the annotation and, consequently, the division in subfamilies.

Based on the confirmation of the subfamilies' division in our first phylogenetic tree, we performed a new alignment combining all *A. fraterculus* and *A. obliqua* OBPs and including OBP sequences from the tephritids *C. capitata* (SICILIANO et al., 2014), *Bactrocera dorsalis* (ZHENG et al., 2013), *Rhagoletis pomonella* (SCHWARZ et al.,

2009), and *Rhagoletis suavis* (RAMSDELL et al., 2010). This alignment was performed with MAFFT separating and aligning OBP sequences by subfamily, in such a way that the alignments by subfamilies were combined in one general alignment using the Profile Alignment Mode in ClustalX2 (LARKIN et al., 2007), which produced a better alignment and a second phylogenetic tree, which was inferred as previously described. MAFFT also provided us with an identity matrix, which we used to estimate the average amino acid identity by subfamily. Due to the high divergence in the N-terminal region, for the phylogenetic and evolutionary analysis we removed the region before the first cysteine residue in all sequences prior to the alignments, as it was done in other studies (GONG et al., 2009; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007; ZHOU et al., 2008; ZHOU et al., 2010).

2.3.3. Evolutionary analysis

Since OBPs from different subfamilies show great divergence, we chose to perform the evolutionary analysis by subfamilies. We used separate alignments and phylogenetic trees for each one of the four OBP subfamilies, using sequences of *A. fraterculus*, *A. obliqua*, *C. capitata* and *D. melanogaster* (Additional file 2). Before the phylogenetic reconstruction, we tested the alignments for sequence saturation using DAMBE 5 (XIA, 2013). We investigated patterns of molecular evolution and positive selection with the strict branch-site test comparing models A vs. A-null (ZHANG, 2005), using the software CODEML, implemented in PAML v.4 (YANG, 2007). The comparison between models was tested using likelihood-ratio tests (LRTs) for hierarchical models (ANISIMOVA; BIELAWSKY; YANG, 2001), and we used a Bonferroni correction for the number of branches tested. A significantly higher likelihood for the alternative model than that of the null model would indicate evidence of positive selection. We also estimated pairwise Ka, Ks rates and its ratio Ka/Ks, using the KaKs_Calculator (ZHANG et al., 2006), for the putative orthologous genes using the MS

model (POSADA, 2003). Genes showing Ka/Ks rates higher than 0.5 were considered as potentially evolving under positive selection (SWANSON et al., 2004).

The Bayes Empirical Bayes (BEB) method implemented in PAML was used to estimate the posterior probability that a given site is evolving under positive selection. Furthermore, we analyzed whether amino acid replacements in positively selected sites alter physicochemical properties in the proteins using Conant-Stadler amino acid property set (CONANT; WAGNER; STADLER, 2007) with the PRIME method available in the Datamonkey web server (DELPORT et al., 2010). To make inferences about the position of the BEB positively selected sites on the tertiary structure of the proteins, we predicted the tertiary structure of *Anastrepha* OBPs under positive selection with PHYRE2 (KELLEY et al., 2015). The program used *Ae. aegypti* OBP1 (PDB: 3K1E, chain A - LEITE et al., 2009) as reference to infer the tertiary structure of *Anastrepha OBP56h-1* and *OBP56h-2*, and *A. gambiae* OBP4 (PDB: 3Q8I, chain A - Davrazou et al., unpublished data) and *A. gambiae* OBP48 (PDB: 4KYN, chain B - TSITSANOU et al., 2013) as reference to *OBP57c* and *OBP50a*, respectively. Since gene sequences in both species were similar and because *AfraOBP50a-1* and *AfraOBP50a-2* were incomplete, we used *A. obliqua* sequences to make the predictions.

2.4. Results and Discussion

2.4.1. OBP genes in *Anastrepha fraterculus* and *Anastrepha obliqua*

We identified a similar number of sequences associated with OBPs in the two species: 23 different sequences in *A. fraterculus* and 24 in *A. obliqua*, which corresponded to 20 different *D. melanogaster* OBP genes (Table 1). In comparison, free-living insect OBP families described to date range from 11 genes in the beet armyworm *Spodoptera exigua* (ZHU et al., 2013) to more than a hundred in some species of mosquitoes (MANOHARAN et al., 2013), although the parasitic body louse *Pediculus humanus* (which has a reduced genome), has only five OBPs (KIRKNESS et al., 2010). The length

of complete open reading frames (Table 1) matched what has been described for several arthropod species (VIEIRA; ROZAS, 2011).

Table 1. Odorant-binding proteins identified in *A. fraterculus* and *A. obliqua* transcriptomes.

A. fraterculus OBPs	ORF length (aa)	Signal Peptide (aa)	Accession Number	*A. obliqua* OBPs	ORF length (aa)	Signal Peptide (aa)	Accession Number	Subfamily
AfraOBP8a	156	1-20	KU317957	AoblOBP8a	156	1-20	KU317933	minus-C
AfraOBP19a	149	1-26	KU317958	AoblOBP19a	149	1-26	KU317934	classic
AfraOBP19b [NT]	130	NA	KU317976	AoblOBP19b	157	1-24	KU317935	classic
AfraOBP19c	232	no	KU317959	AoblOBP19c	229	no	KU317936	classic
AfraOBP19d	157	1-22	KU317960	AoblOBP19d	157	1-22	KU317937	classic
AfraOBP47b	194	1-22	KU317961	AoblOBP47b	194	1-22	KU317938	plus-C
AfraOBP49a	215	1-20	KU317962	AoblOBP49a-1	215	1-20	KU317939	plus-C
				AoblOBP49a-2	178	1-19	KU317940	plus-C
AfraOBP50a-1 [NT]	139	NA	KU317977	AoblOBP50a	259	1-18	KU317941	plus-C
AfraOBP50a-2 [NT]	139	NA	KU317978					plus-C
AfraOBP50e	231	1-20	KU317963	AoblOBP50e	231	1-20	KU317942	plus-C
AfraOBP56d-1 [NT]	105	NA	KU317979	AoblOBP56d-1	138	1-18	KU317943	classic
AfraOBP56d-2	138	1-18	KU317964	AoblOBP56d-2	138	1-18	KU317944	classic
AfraOBP56h-1	136	1-19	KU317965	AoblOBP56h-1	136	1-19	KU317945	classic
AfraOBP56h-2	125	1-19	KU317966	AoblOBP56h-2	125	1-19	KU317946	classic
AfraOBP57c	178	1-50	KU317967	AoblOBP57c	178	1-50	KU317947	classic
AfraOBP59a	309	1-26	KU317968	AoblOBP59a	309	1-26	KU317948	classic
AfraOBP83cd	241	1-19	KU317969	AoblOBP83cd	241	1-19	KU317949	dimer
AfraOBP83ef	282	1-31	KU317970	AoblOBP83ef	282	1-31	KU317950	dimer
AfraOBP83g	143	1-17	KU317971	AoblOBP83g	143	1-17	KU317951	minus-C
AfraOBP99a	143	1-17	KU317972	AoblOBP99a	143	1-17	KU317952	minus-C
AfraOBP99b	153	1-17	KU317973	AoblOBP99b	153	1-17	KU317953	minus-C
AfraOBP99c	151	1-17	KU317974	AoblOBP99c	151	1-17	KU317954	minus-C
AfraOBP99d	152	1-20	KU317975	AoblOBP99d-1	151	1-20	KU317955	minus-C
				AoblOBP99d-2	151	1-20	KU317956	minus-C

[NT] N-terminus missing; NA not applied;

The longest OBPs are *AfraOBP59a* and *AoblOBP59a*. These sequences are unusual, because they have a specific region containing 134 amino acids that is situated between the first (C1) and the second (C2) cysteine, whereas others in the same subfamily have ~30 amino acids in this region. Even though an unusual length for *OBP59a* had already been demonstrated in other arthropods (VIEIRA; ROZAS, 2011), in *Anastrepha* these OBPs are 106 amino acids longer than in *D. melanogaster*. Signal peptides, one of characteristic hallmarks of the OBP gene family, were predicted at the hydrophobic N-

terminal for the almost all OBPs, except for *AfraOBP19c* and *AoblOBP19c*. Though not all OBPs described have signal peptides (ZHOU et al., 2010), we believe that the PrediSi program failed to find a match in the SwissProt database for both *Anastrepha OBP19c* signal peptide due to sequence divergence, not because these OBPs lack a signal peptide.

A comparison of putatively homologous OBPs in *A. fraterculus* and *A. obliqua* reveals great amino acid similarity, so much so that homologs of *OBP19a* and *OBP99a* had identical amino acid sequences in both species. In spite of that, the majority of orthologous OBPs differ by at least a few amino acids, similar to what was described for related species of aphids (QIAO et al., 2009; ZHOU et al., 2010). However, even a single amino acid change may impact the overall 3D structure and/or the binding affinities of OBPs (AHMED et al., 2014; WANG et al., 2013). For instance, a polymorphism in *Obp57e* was shown to be responsible for differences in host plant preferences between *D. sechellia* and *D. melanogaster* (MATSUO et al., 2007), and some polymorphisms were associated with natural variation in olfactory behavior in response to benzaldehyde in *Obp99a*, *Obp99c*, and *Obp99d* in *D. melanogaster* (WANG et al., 2007). Likewise, the few changes observed between *A. fraterculus* and *A. obliqua* OBPs may result in significant differences in olfactory responses, which is yet to be determined.

We did an alignment with all *A. fraterculus* OBP sequences, and another one with all *A. obliqua* OBP sequences. Two independent recombination tests were performed using these alignments, because recombination interferes with phylogenetic inferences and may generate higher rates of false positives in positive selection tests (ANISIMOVA; BIELAWSKY; YANG, 2001). Both tests performed failed to find significant evidence for recombination. These alignments revealed that even though putatively orthologous OBP copies in different species were very similar, there was great divergence among OBP genes, so much so that only the cysteine residues were conserved across all OBPs

of the same species. Sequence divergence was higher than what has been described for other insects. The average amino acid identity among all OBPs was 16.65% in *A. fraterculus* (ranging from 8.1 to 97.1%) and 16.22% in *A. obliqua* (varying from 8.0 to 98.7%), compared with 20.4% of identity between all OBPs in *D. melanogaster* (HEKMAT-SCAFE et al., 2002), and 20% between all OBPs found in *Solenopsis invicta* (GOTZEK et al., 2011).

The higher divergence among OBPs in *A. fraterculus* and *A. obliqua*, when compared with other insects, may suggest a greater selective pressure for OBP differentiation, or relaxed selection, which also led us to investigate for patterns of evolution in these genes**.** On the other hand, the high amino acid similarity found between the homologs in *A. fraterculus* and *A. obliqua* probably indicates that the recent time of divergence between these species may not have been enough to drastically alter their sequences, which could be a reflection of retention of similar physiological functions across species or simply of the recent time since their divergence. Since these species have diverged recently and have accumulated only a few changes, the current methodologies to investigate for positive selection between these species are not very efficient, but we may use these data in contrast to other Tephritidae and Drosophilidae to investigate for patterns of selection in OBPs as a whole. We chose to investigate *Anastrepha* OBPs because they may be involved with the group's differentiation since they may be directly involved with host preference and mate choice. Therefore, patterns of selection in these genes may help us understand evolution and speciation not only in the genus *Anastrepha*, but also across this important family, which may make it useful in pest control programs.

2.4.2. Phylogenetic relationships among OBP genes

A ML phylogenetic tree that included all OBP sequences derived from *Anastrepha*, *C. capitata* and *D. melanogaster* (Figure 1) indicates four clades that correspond to four distinct protein subfamilies based on *D. melanogaster*'s classification (HEKMAT-SCAFE et al., 2002), hence the importance of including this species in the analysis. The four clades were highly supported (SHlike aLRT branch supports ranging from 0.855 and 0.966) and were composed of OBPs from all species here investigated, indicating that these subfamilies diverged before their common ancestor, which has already been described for others dipterans (OHTA et al., 2014; SICILIANO et al., 2014; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007). The support for each separate OBP subfamily in *Anastrepha* was also confirmed by the conservation of cysteine patterns (Table 2).

The average amino acid identity was higher in the dimer subfamily for both species, but the number of members in each subfamily may have inflated these values, since larger families show lower average identity (Table 2). Classic OBPs have the expected standard pattern of six conserved cysteine residues, whereas dimer OBPs have twelve cysteine residues, a pattern formed by the junction of two consecutive minus-C OBP domains (SÁNCHEZ-GRACIA; ROZAS, 2008). The dimer subfamily was identified in *A. fraterculus*, *A. obliqua* and *D. melanogaster*, though it has not been described for *C. capitata* yet. As in *D. melanogaster*, *A. fraterculus* and *A. obliqua* minus-C OBPs were divided in two lineages, one in which the members retain six conserved cysteine residues (a clade that includes *OBP8a*, *OBP99c* and *OBP99d*), and another one (that includes *OBP83g*, *OBP99a* and *OBP99b*) in which its members show only four conserved cysteine residues (Figure 1). The reduced number of residues in some members is caused by the loss of the conserved pair of cysteines C2 and C5, which forms a disulphide bond (FAN et al., 2011), which happened after the minus subfamily diverged

from the classic subfamily and might have a functional relevance, perhaps generating a more flexible structure (HEKMAT-SCAFE et al., 2002; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007).

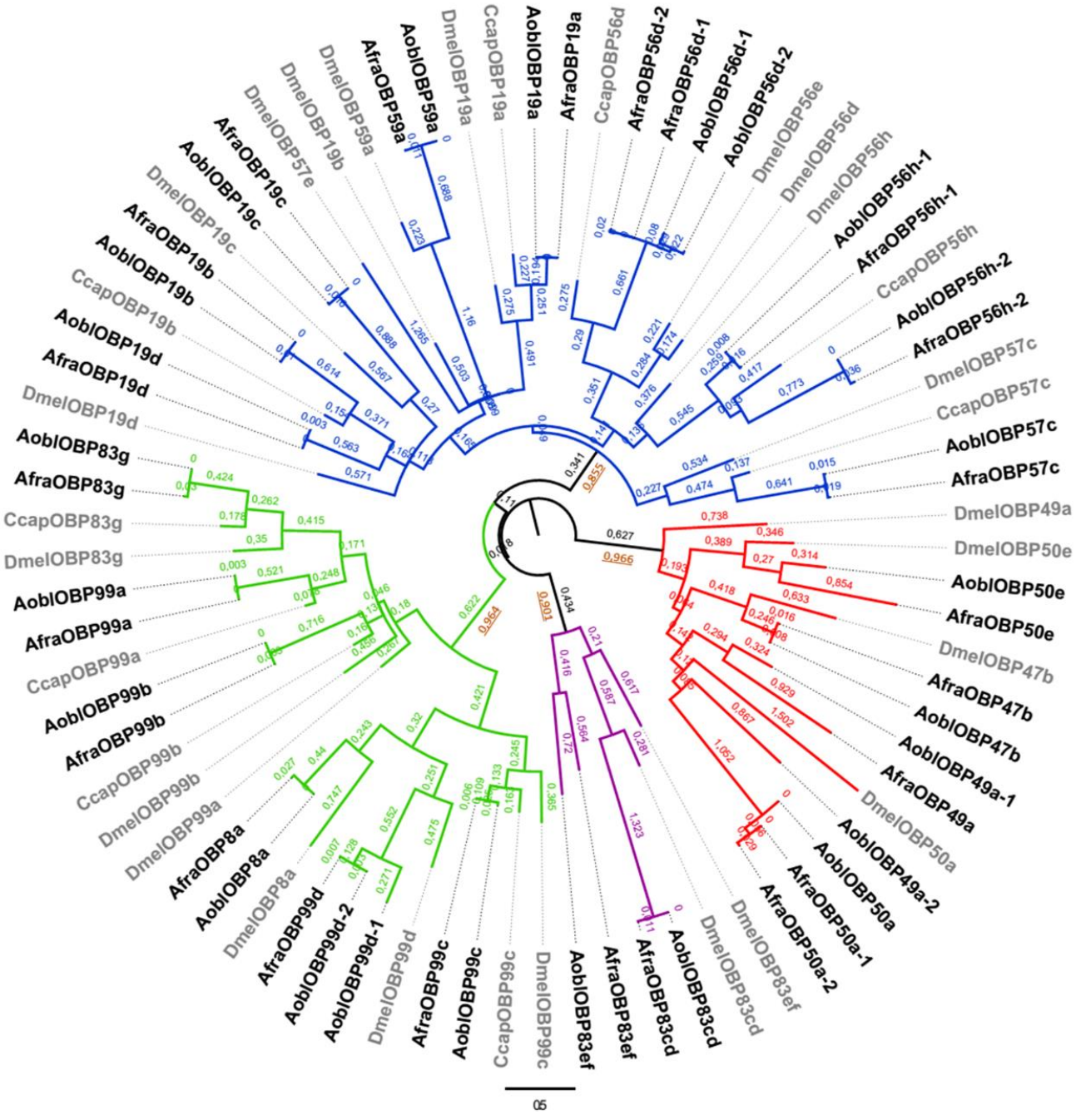


Figure 1. Mid-point rooted ML phylogenetic relationships of the OBPs. The branches are color coded for each subfamily: classic (blue), minus-C (green), plus-C (red) and dimer (purple). Branch lengths are estimated by amino acid substitutions per site. Brown and underlined numbers represent the SHlike aLRT branch supports for the four subfamilies.

Table 2. Attributes of *Anastrepha* Odorant-binding proteins subfamilies.

	Subfamily	Number of members	ORF length (aa)	Protein core region	Amino acid identity (%)
A. fraterculus	classic	10	125-309	C1 - X_{24-33} - C2 - X_3 - C3 - X_{35-47} - C4 - X_{8-18} - C5 - X_8 - C6	28.20
	minus-C	6	143-156	C1 - X_{26} - C2 - X_3 - C3 - X_{39-40} - C4 - X_{10} - C5 - X_8 - C6 and C1 - X_{23-30} - C3 - X_{38} - C4 - X_{18-19} - C6	39.14
	plus-C	5	194-231	C1a - X_{11-13} - C1b - C1c - C1 - X_{21-48} - C2 - X_3 - C3 - X_{43} - C4 - X_{20-33} - C5 - X_9 - C6 - X_8 - C6a - X_{10} - C6b - X_9 - C6c	45.14
	dimer	2	241-282	C1 - X_{28-29} - C2 - X_3 - C3 - X_{31} - C4 - X_{10-11} - C5 - X_8 - C6 - X_{17-28} - C1' - X_{24-25} - C2' - X_3 - C3' - X_{35} - C4' - X_{9-12} - C5' - X_8 - C6'	65.75
A. obliqua	classic	10	125-309	C1 - X_{24-37} - C2 - X_3 - C3 - X_{35-42} - C4 - X_{8-12} - C5 - X_8 - C6	29.54
	minus-C	7	143-156	C1 - X_{26} - C2 - X_3 - C3 - X_{39-40} - C4 - X_{10} - C5 - X_8 - C6 and C1 - X_{23-30} - C3 - X_{38} - C4 - X_{18} - C6	38.10
	plus-C	5	178-259	C1a - X_{11-13} - C1b - C1c - C1 - X_{21-48} - C2 - X_3 - C3 - X_{43} - C4 - X_{20-33} - C5 - X_9 - C6 - X_8 - C6a - X_{10} - C6b - X_9 - C6c	40.25
	dimer	2	241-282	C1 - X_{28-29} - C2 - X_3 - C3 - X_{31} - C4 - X_{10-11} - C5 - X_8 - C6 - X_{17-28} - C1' - X_{24-25} - C2' - X_3 - C3' - X_{35} - C4' - X_{9-12} - C5' - X_8 - C6'	65.14

Both protein core regions are reported for the minus-C subfamily, with four and six conserved cysteine residues.

A. fraterculus and *A. obliqua* plus-C OBPs showed three additional conserved cysteine residues before C1 (referred to as C1a, C1b and C1c) and three others after C6 (C6a, C6b and C6c), making for a total of 12 conserved cysteine residues. Three sequences associated with *OBP50a* did not have the cysteine C6c. Similar to what was described for *Drosophila*, we found a conserved hydrophobic proline after the cysteine C6a in *Anastrepha*'s plus-C members, as well as nine residues between C5 and C6, instead of eight in the other subfamilies (VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007). Recently, Siciliano et al. (SICILIANO et al., 2014) reported a plus-C OBP in *C. capitata*, related to *DmelOBP49a*, but in their phylogenetic tree, this OBP was grouped with classic OBPs. OBPs in the plus-C subfamily, such as the ones described here, generally show a

substantial increase in length size compared with other OBPs, mostly because of the extended C-termini (FORÊT; MALESZKA, 2006; ZHOU et al., 2004). The plus-C subfamily was also described in the pea aphid *Acyrthosiphon pisum* (ZHOU et al., 2010), and the high support in the tree indicates that their origin probably precedes the divergence of Diptera.

We also performed a phylogenetic inference that was restricted to tephritid OBPs, including *C. capitata*, *B. dorsalis*, *R. pomonella* and *R. suavis* (Figure 2), which would enable us to have a better understanding of the patterns of evolution of OBPs in these important pest species. As in the previous phylogenetic inference (Figure 1), we observe a basal division in four subfamilies with one exception. *Anastrepha OBP59a*, a classic OBP, in this inference was grouped with the dimer subfamily members, probably due to its extended length, which could have led to long-branch attraction. We also observed that *RpomOBP83cd*, a dimer member, grouped with minus-C subfamily, and some *B. dorsalis* OBPs were grouped in different subfamilies than was previously described (ZHENG et al., 2013). In our phylogenetic inference of tephritid OBPs, *CcapOBP49a* was grouped with plus-C subfamily members (SHlike aLRT branch support 0.992 and branch length 1.046), not with the classic subfamily, as previously described (SICILIANO et al., 2014). In general, each *Anastrepha* OBP clustered with their putative homolog from other tephritid species, with high support for the majority of branches.

Even though the overall number of OBP genes seems to be equivalent among tephritids, we point out that species of *Anastrepha* have a larger OBPs repertoire than the other tephritids studied: 23 and 24 OBP genes, respectively. In comparison, 17 different OBPs were described for *C. capitata*, the closest species of *A. fraterculus* and *A. obliqua* studied, 13 of which corresponded to the classic OBP subfamily, three to the minus-C subfamily and one to the plus-C subfamily (SICILIANO et al., 2014). We also found a

larger number of plus-C OBPs when compared to other tephritids. We detected five apiece for *A. fraterculus* and *A. obliqua*, whereas two were described in *R. pomonella* and only one in *R. suavis*, *B. dorsalis* and *C. capitata*. The lack of genomic sequences for *Anastrepha* makes the analyses here not complete, since we may only rely on levels of divergence and phylogenetic expectations that could be verified solely upon investigation of the evolutionary patterns found in the species' genome.

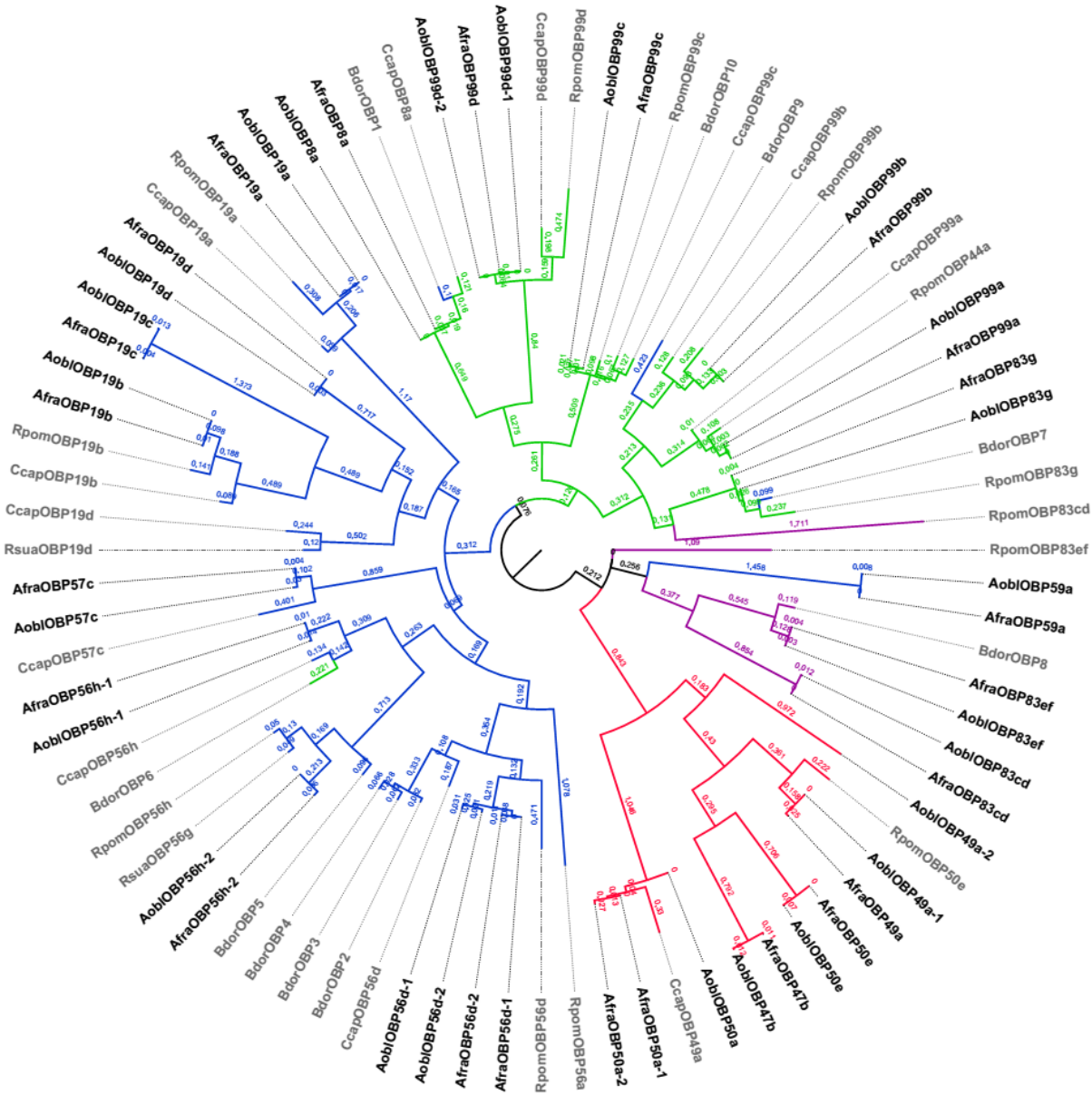


Figure 2. Mid-point rooted ML phylogenetic relationships of the tephritids OBPs. The branches are color coded for each subfamily: classic (blue), minus-C (green), plus-C (red) and dimer (purple). Branch lengths are estimated by amino acid substitutions per site.

We observed some instances in which more than one sequence in the *A. fraterculus* and *A. obliqua* OBP repertoire was associated with the same *D. melanogaster* OBP. In *A. fraterculus*, two different sequences were associated with the genes *OBP50a*, *OBP56d* and *OBP56h*, and in *A. obliqua*, with the genes *OBP49a*, *OBP56d*, *OBP56h* and *OBP99d*. These findings could be due to gene duplication events, though we cannot rule out intraspecific variation, because we used a pool of flies from a single population to make the cDNA libraries in each species. Analyzing the values of pairwise identity between the two copies of the same OBP of a species and comparing copies of the same OBP between species, different *Anastrepha* sequences associated with the same OBP seem to be due to intraspecific variation, with exception of two cases: *OBP56h* and *OBP49a*.

We believe that the copies homologous to *OBP56h* are consequence of a duplication event that preceded the divergence of *A. fraterculus* and *A. obliqua*. *AfraOBP56h-1* and *AoblOBP56h-1* have a pairwise identity of 93.6%, and *AfraOBP56h-2* and *AoblOBP56h-2* an identity of 96.3%, whereas the pairwise identity between *OBP56h* sequences of the same species is 29.3% in *A. fraterculus* and 33.3% in *A. obliqua*, clearly indicating them to be paralogous. Moreover, the phylogenetic analysis of plus-C members reveals that *AoblOBP49a-2* groups with the others *OBP49a*, including *DmelOBP49a* (Additional file 2). However, *AoblOBP49a-2* is 37 amino acids shorter than *AoblOBP49a-1* and *AfraOBP49a* and quite different. While the pairwise amino acid identity between *AoblOBP49a-1* and *AfraOBP49a* is 95%, it is only 31% between the two copies of *A. obliqua* (*AoblOBP49a-1* and *AoblOBP49a-2*), suggesting as well that *AoblOBP49a-2* is not an ortholog of *DmelOBP49a*, such as *AoblOBP49a-1* and *AfraOBP49a* but, rather, a paralog. The fact that we failed to find this copy in *A. fraterculus* may indicate that it diverged after the separation of the species, which would

be extremely interesting, but considering the high levels of divergence found among these paralogs, it is more likely that they diverged before the separation of *Anastrepha* species, and did not show significant levels of expression in *A. fraterculus*, which is also relevant.

2.4.3. Analysis of positive selection in OBPs

Similar to all insect OBPs studied so far (GOTZEK et al., 2011), *Anastrepha* OBPs share low sequence similarity, rendering evolutionary analyses difficult. Therefore, we did not use the alignment that included all subfamilies because we identified that there was saturation at synonymous positions in more divergent comparisons (data not shown), which might lead to a higher number of false positives in the branch-site test of positive selection (FLETCHER; YANG, 2010). We performed independent evolutionary analyses on each of the four OBP subfamily identified (Table 2; Additional file 2), which did not show evidence of sequence saturation. This, unfortunately, prevented us from investigating evidence of positive selection amongst the four subfamilies, but allowed us to investigate for positive selection within any of such subfamily.

We detected evidence of positive selection, using the strict branch-site test, for sequences associated to two classic OBP genes: *OBP56h* (*AfraOBP56h-1* and *AoblOBP56h-1* separated from *AfraOBP56h-2* and *AoblOBP56h-2* on different branches) and *OBP57c* (*AfraOBP57c* and *AoblOBP57c*). The significant results in the strict branch-site test for *Anastrepha OBP56h* suggest that after the likely duplication event in this gene, indicated by the pairwise identity analysis, positive selection acted to differentiate the two copies, which could have been to discrete odor specificities. Moreover, we detected positive selection on one plus-C gene: *OBP50a* (*AfraOBP50a-1*, *AfraOBP50a-2* and *AoblOBP50a*) (Table 3). The *OBP50a* copies from *A. fraterculus* are more closely related to one another than to the corresponding copy from *A. obliqua*, and have over 97% similarity, which could be an indication of intraspecific polymorphism in *A. fraterculus*. However, the evidence of positive selection acting in these genes may indicate a recent

duplication in that species, which could be best investigated by an analysis on the genome. Contrary to other insects such as *Apis mellifera* and *B. mori* (FORÊT; MALESZKA, 2006; GONG et al., 2009), we failed to find evidence of positive selection on the minus-C subfamily.

Table 3. Positive selection detected in *Anastrepha* OBP genes.

Gene	LRT[1]	Positively selected sites[2]
AfraOBP56h-1 *AoblOBP56h-1*	10,444**	85
AfraOBP56h-2 *AoblOBP56h-2*	7,555**	40, 67, 68, 145, 152
AfraOBP57c *AoblOBP57c*	8,403**	5, 34, 61, 65, 76, 94, 146, 150
AfraOBP50a-1 *AfraOBP50a-2* *AoblOBP50a*	8,416*	4, 56, 59, 146, 152, 156, 164, 177, 181

[1]Likelihood Ratio Test results; [2]These numbers reflect the site positions in the overall alignment, not necessarily the position in each specific OBP. Amino acid positions involving changes in the physicochemical properties influenced by positive selection are underlined. **p<0.00625; *p<0.0125.

A Bayes Empirical Bayes (BEB) method identified several amino acid sites under positive selection (Figure 3; Table 3), some of which were in regions under strong purifying selection (Figure 3b; 3d). For classic OBP sequences associated with *OBP56h* and *OBP57c*, the positively selected amino acids were located in the protein core region (between cysteines C1 and C6), except for the amino acid in position 152, the position right after C6. For plus-C *OBP50a*, we found the largest number of positively selected sites: nine sites, four of them located in the protein core region. The amino acid in position 181 is the cysteine C6c (the twelfth cysteine residue) in all plus-C sequences except in those associated with *OBP50a*, where a cysteine was replaced by asparagine. Though in general cysteines are conserved in OBPs, even highly conserved cysteine residues may be lost (VIEIRA et al., 2012).

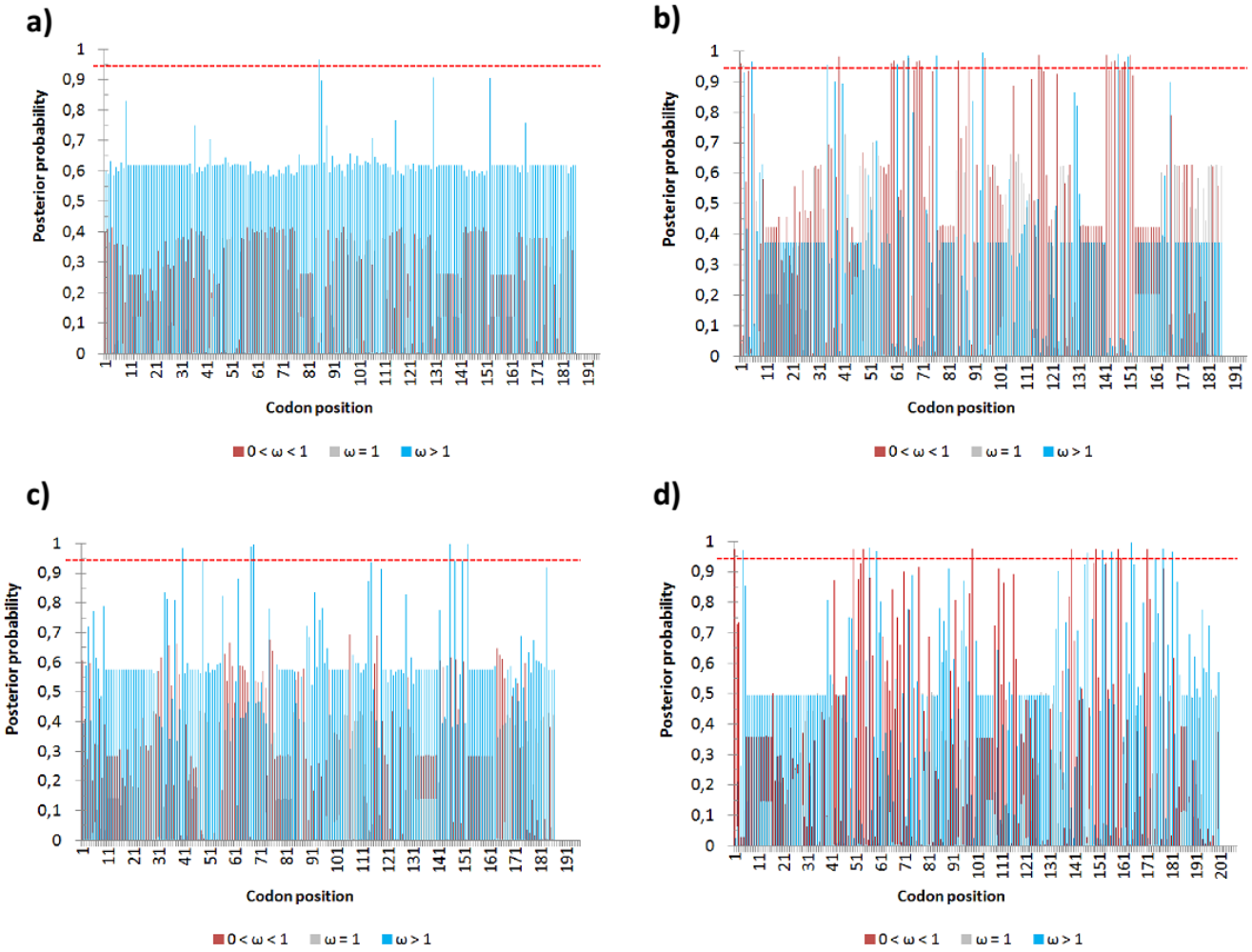


Figure 3. Bayes Empirical Bayes showing posterior probabilities of sites under purifying (0<ω<1), neutral (ω=1) and positive selection (ω>1). Amino acid in codon position '1' is cysteine C1. Dashed red lines indicate the posterior probability threshold of 0.95. a) results for *AoblOBP56h-1* and *AfraOBP56h-1*; b) results for *AoblOBP56h-2* and *AfraOBP56h-2*; c) results for *AoblOBP57c* and *AfraOBP57c*; d) results for *AoblOBP50a*, *AfraOBP50a-1* and *AfraOBP50a-2*.

We investigated whether amino acid changes in OBPs detected by the BEB analysis were associated with site-specific changes related to five amino acid properties: chemical composition, polarity, volume, isoeletric point and hydropathy ($p<0.05$), because not all non-synonymous substitutions are alike. If non-synonymous substitutions result in similar properties, the maintenance of the structure and/or chemical function of the region bearing such changes would be considered as conserved. On the other hand,

radical non-synonymous changes drastically change important physicochemical attributes. When these changes involve sites under positive selection, it is more likely that they would be those that promote structural and functional changes in a protein. Among the five non-synonymous changes in the duplicated copy of *Anastrepha OBP56h-2* we found two amino acid positions (68 and 145) that are associated with radical changes in the physicochemical properties, one associated to changes in isoelectric point, and another to changes in chemical composition. The high number of amino acid substitutions, particularly leading to radical changes, may be an indication of the effect of positive selection driving this recently duplicated copy to a different odor affinity.

Anastrepha OBP57c and *Anastrepha OBP50a* also have one amino acid position each that leads to radical changes in the physicochemical properties. A change in the amino acid position 150 of *Anastrepha OBP57c* alters its chemical composition, whereas a change in amino acid 59 of *Anastrepha OBP50a* affects its volume. Changes in chemical composition and volume of OBPs are important because they may be related to conformational changes in these globular proteins. Altering the structure of α-helices may modify their flexibility and, therefore, modify the interacting motifs of the protein (SÁNCHEZ-GRACIA; ROZAS, 2008). Likewise, substitutions that affect the isoeletric point are important since they may be involved with changes in the proteins' solubility, which, in the aqueous lymph of sensillas, may be related with a greater efficiency in transport activity.

An analysis of synonymous and non-synonymous substitution rates between putatively homologous copies in *A. fraterculus* and *A. obliqua* reveals several copies with high values for Ks (Table 4) indicating that even though the number of changes between homologous copies may not be high, there is a high proportion of non-synonymous changes. We have identified two OBPs with Ka/Ks above one and seven others that had

Ka/Ks over 0.5, whereas the median Ka/Ks ratio in putatively orthologous OBPs is 0.44. Though we should consider that the Ka/Ks ratio may not be reliable when mutation numbers are small, e.g., for within populations estimates (KRYAZHIMSKIY; PLOTKIN, 2008) or even for recently diverged lineages, which seems to be the case here, a comparison of these data with other genes in the same populations may provide an interesting parameter to consider when evaluating neutral evolution.

Table 4. Estimates for Ka, Ks and Ka/Ks ratio between *A. fraterculus* and *A. obliqua* putatively orthologous OBP genes.

OBP genes compared	Ka	Ks	Pairwise Ka/Ks	P-value (Fisher)
AfraOBP8a - AoblOBP8a	0.01302	0.04044	0.32192	0.02418
AfraOBP19a - AoblOBP19a	0.00001	0.01018	0.00100	0.08109
AfraOBP19b - AoblOBP19b	0.01250	0.02512	0.49751	0.34228
AfraOBP19c - AoblOBP19c	0.02404	0.01055	2.27963	0.35887
AfraOBP19d - AoblOBP19d	0.00275	0.00005	50	0.36788
AfraOBP47b - AoblOBP47b	0.01987	0.05524	0.35971	0.03121
AfraOBP49a - AoblOBP49a-1	0.01922	0.04152	0.46284	0.04973
AfraOBP49a - AoblOBP49a-2	0.75439	1.89710	0.39765	0
AfraOBP50a-1 - AoblOBP50a	0.03385	0.12629	0.26803	0.00095
AfraOBP50a-2 - AoblOBP50a	0.04541	0.22418	0.20258	0
AfraOBP50e - AoblOBP50e	0.00448	0.03749	0.11959	0.00282
AfraOBP56d-1 - AoblOBP56d-1	0.06552	0.15216	0.43063	0.0365
AfraOBP56d-2 - AoblOBP56d-2	0.14590	0.20809	0.70117	0.13362
AfraOBP56d-1 - AoblOBP56d-2	0.06276	0.12104	0.51850	0.12677
AfraOBP56d-2 - AoblOBP56d-1	0.14717	0.21836	0.67400	0.10769
AfraOBP56h-1 - AoblOBP56h-1	0.01884	0.05216	0.36124	0.07521
AfraOBP56h-2 - AoblOBP56h-2	0.03487	0.04064	0.85809	0.72667
AfraOBP56h-1 - AoblOBP56h-2	0.70700	1.99231	0.35522	0
AfraOBP56h-2 - AoblOBP56h-1	0.72197	1.84218	0.39191	0
AfraOBP57c - AoblOBP57c	0.02325	0.03906	0.59518	0.36258
AfraOBP59a - AoblOBP59a	0.00409	0.02980	0.13711	0.00594
AfraOBP83cd - AoblOBP83cd	0.00951	0.02519	0.37748	0.16455
AfraOBP83ef - AoblOBP83ef	0.00202	0.01555	0.13002	0.04322
AfraOBP83g - AoblOBP83g	0.02066	0.06384	0.32355	0.08097
AfraOBP99a - AoblOBP99a	0.00002	0.02374	0.00100	0.01413
AfraOBP99b - AoblOBP99b	0.00826	0.01162	0.71080	0.33160
AfraOBP99c - AoblOBP99c	0.02607	0.05666	0.46011	0.06971
AfraOBP99d - AoblOBP99d-1	0.01599	0.02393	0.66817	0.40804
AfraOBP99d - AoblOBP99d-2	0.01000	0.02179	0.45910	0.12015
Median	**0.01902**	**0.03975**	**0.44486**	

A similar analysis performed on 164 contigs that were identified as highly differentiated between *A. fraterculus* and *A. obliqua* shows median Ka = 0.0026, Ks = 0.0282 and Ka/Ks= 0.09 and eleven contigs with a Ka/Ks above 0.5, one of them an OBP. A contrast between these two data sets reveals that OBPs as a whole evolve faster and have proportionately more non-synonymous substitutions between these two species than other genes in the genome (t-test = 3.74, $p<0.001$) and suggests that several OBPs may be evolving under positive selection, though we cannot rule out that they are evolving under relaxed selection.

2.4.4. Positively selected sites in OBP 3-D structures

Positively selected sites experience a faster rate of amino acid replacement that is a consequence of mutations being fixed at a higher rate than expected by chance. Therefore, knowing the location of these amino acid sites in the protein's tertiary structure may help understand how these changes would affect their function. Structurally, OBPs share a common fold with α-helices connected by loops and interlinked by disulphide bonds, but despite their structural homology, they are predicted to bear binding cavities of different shapes (TEGONI; CAMPANACCI; CAMBILLAU, 2004). Positive selection in OBPs may promote functional divergence on the binding specificities, which is important because such structural diversity may enable OBPs to recognize and bind to a wider range of organic molecules and naturally occurring odorants (SÁNCHEZ-GRACIA; ROZAS, 2008).

The predicted 3D structures were used to locate the amino acid positions under positive selection according to PAML analysis (Figure 4). The tertiary structure of *A. gambiae OBP4* (*AgamOBP4*) (Davrazou et al., unpublished data) was selected by Phyre2 as a reference for *AoblOBP57c* and *AfraOBP57c*. *AgamOBP4* is a protein with 124 amino acid residues, in which three disulphide bridges link α1 to α3, the start of α3 to α6, and the middle of α6 to α5. In *Anastrepha* (Figure 4c), the positively selected sites were found

in α2, α3, α4 and α6, and loops between α1 and α2, and between α3 and α4. Notably the positively selected amino acids located in α3 and especially in α6 may be structurally very important because they are located near disulphide bonds, specially the last one that causes changes in the chemical composition.

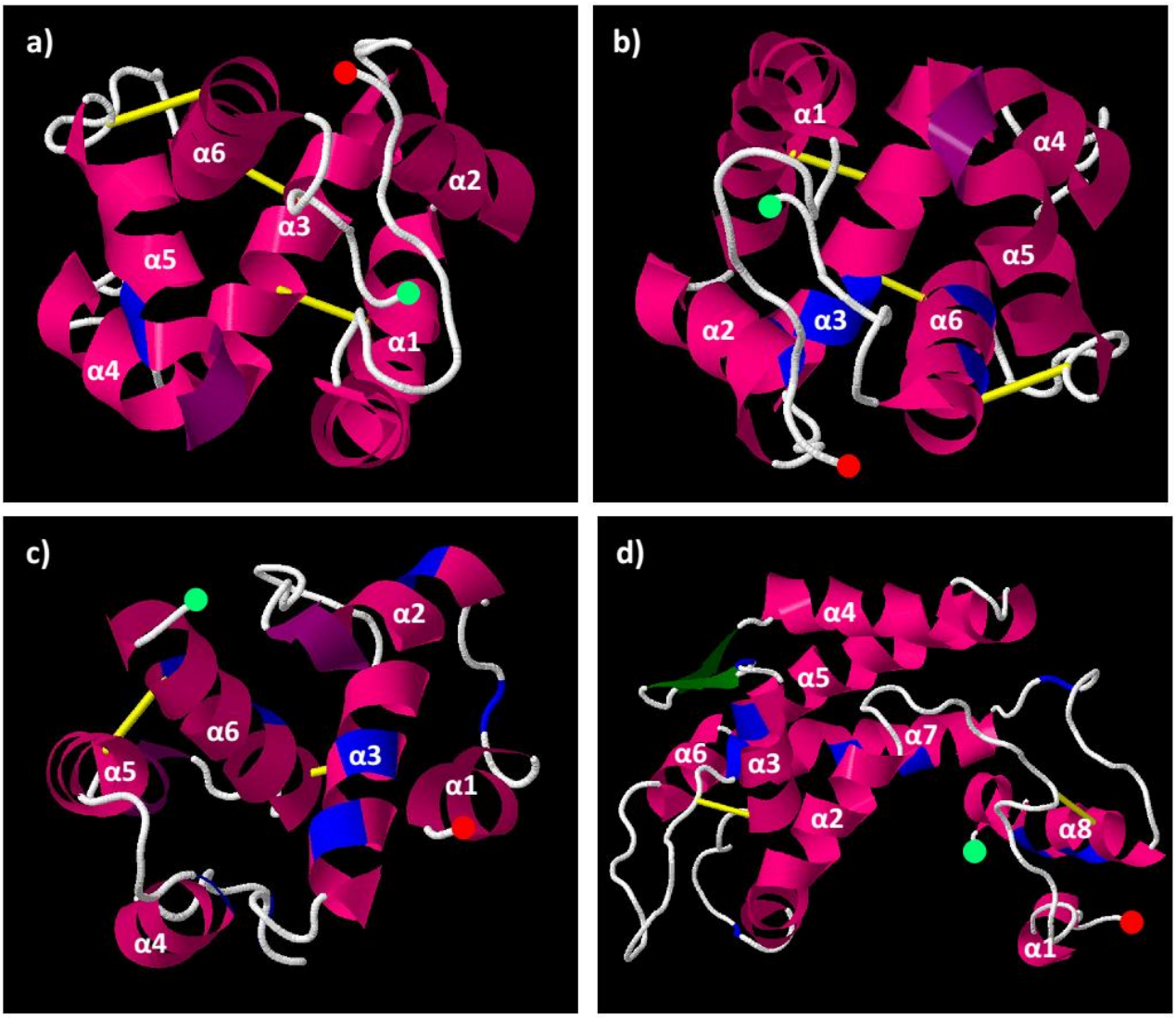


Figure 4. Cartoon representation of hypothetical positions of the positively selected sites in *Anastrepha* OBPs, based in their predicted 3D structures. α-helices are shown in pink, β-sheets in green, loops in white, disulphide bonds in yellow and positively selected sites in blue. N- and C-terminus residues are showed with red and light green circles, respectively. a) *AoblOBP56h-1* and *AfraOBP56h-1* representation; b) *AoblOBP56h-2* and *AfraOBP56h-2* representation; c) *AoblOBP57c* and *AfraOBP57c* representation; d) *AoblOBP50a*, *AfraOBP50a-1* and *AfraOBP50a-2* representation.

The tertiary structure of *Ae. aegypti OBP1*, selected by Phyre2 as reference for *Anastrepha OBP56h* due to their similarity, is a protein with 125 amino acid residues, comprised of six α-helices connected by loops between helices and knitted together by three disulphide bridges between α1 and α3, α3 and the top of α6, and α6 and the top of α5. Sixteen hydrophobic residues in helices α1, α3, α4, α5 and in loops between helices α3 and α4, and between α5 and α6, form the binding cavity. Binding cavities in the OBPs homodimers form channels, a long and continuous hydrophobic tunnel, where the odorant binds (LEITE et al., 2009). We did not find any site under positive selection in *Anastrepha* that correspond to sites that form the binding cavity in *AaegOBP1*. The site under positive selection in *AfraOBP56h-1* and *AoblOBP56h-1* is located in α5 (Figure 4a). The sites under positive selection in *AfraOBP56h-2* and *AoblOBP56h-2* (Figure 4b) are located in α2, α3 e α6, in which the latter two helices bear sites that lead to changes in the physicochemical properties, according to analysis in PRIME.

A. gambiae plus-C *OBP48*, selected by Phyre2 as reference for *Anastrepha OBP50a*, has 172 residues and eight α-helices that are stabilized by six disulphide bridges (TSITSANOU et al., 2013). *AgamOBP48* is structurally divided into three domains: the "core", the "NC-term" and the "cap". The central "core" domain shares a common architecture with classic OBPs and consists of four bridgedα-helices (α1, α3, α5 and α7), bearing six conserved cysteines that form three disulphide bridges connecting α1 to α3, α3 to the start of α7, and the middle of α7 to α5. This domain also contains two non-bridged α-helices α2 and α4, and a loop of eight amino acids (50s loop). Sites under positive selection in *Anastrepha* OBP50a sequences are found in the "core" region, in α3 and α7, in the loop between α2 and α3, and in the loop between α3 and β-sheet1 (Figure 4d). A seventh cysteine in *AgamOBP48* is located at the top of α3, but it is not involved in a disulphide bond, since not all cysteine residues are necessarily involved in disulphide

bonds (TSITSANOU et al., 2013). For instance, *A. fraterculus* and *A. obliqua* plus-C *OBP50a* lost the twelfth conserved cysteine residue in a positively selected change that may have relevant structural consequences, but this loss does not seem to affect their function.

The "NC-term" domain comprises the N-terminal residues 1-25, the C-terminal 150s loop (148-155) and the following C-terminal α8 helix (156-172), forming three disulphide bonds. The two subunits orient themselves so that the "NC-term" domain of one monomer inserts into the center of the neighboring monomer, to form a compact homodimer. The N-terminal component of this domain is characterized by two additional cysteines located in adjacent positions before C1 (TSITSANOU et al., 2013), pattern also observed in all *Anastrepha* plus-C OBPs. *Anastrepha OBP50a* has sites under positive selection in this domain, two ones in α8 and another one in the loop (Figure 4d). Finally, the "Cap" domain, in which we failed to find sites under positive selection in *Anastrepha*, encompasses helix α6 and the loop near amino acid 120, and this interaction between the helix and the loop is characterized by multiple hydrogen-bonding that apparently serves to stabilize protein structure (TSITSANOU et al., 2013).

AgamOBP48 shows a single binding site, although the authors argue that given the symmetry of the dimer, ligand binding to two equivalent binding sites, named "NC-term" pockets, cannot be excluded (TSITSANOU et al., 2013). We focused only on the single binding site to compare with *Anastrepha OBP50a*. The binding cavity is formed by 23 amino acid residues in the helices α5, α6 and α7, and in the loops between helices α2 and α3, and between α7 and α8. We found one amino acid under positive selection in the position exactly correspondent to the amino acid that forms the binding cavity in α7 of *AgamOBP48*. The other amino acid in α7, as well as the amino acids in the loop

between α2 and α3, and in the loop between α7 and α8 did not correspond to the same position to the binding cavity, but they are juxtaposed to the amino acids that form it.

Because we are only using computational inferences, we cannot be certain about the 3D structures inferred for *Anastrepha OBP56h*, *OBP57c* and *OBP50a*, nor if the *Anastrepha* OBP binding cavity is formed in the same site positions that in the 3D structures of *Ae. aegypti* and *A. gambiae* OBPs used as reference. However, the OBPs' high conservation at the structural level, observed in all insects studied, provide us some basis for this extrapolation (GALINDO; SMITH, 2001; TEGONI; CAMPANACCI; CAMBILLAU, 2004; ZHOU et al., 2009). The exception to this structure conservation is a unique feature of mosquito OBPs, their C-terminal loop covers the binding cavity, forming a ''lid'' for the release of ligands (WOGULIS et al., 2006), therefore we expect that the C-terminal in *Anastrepha* may be quite different, potentially altering our 3D inference for the C-terminal part. Only *Anastrepha OBP50a* showed one amino acid under positive selection that possibly correspond to amino acids that form the binding cavity for odors. However, even if no replacement occurs in amino acids that are directly involved with the binding function, radical non-synonymous substitutions placed both in the α-helix or even in the loops are important because they might alter the size and shape of the binding cavity. For instance, by modifying the position of the disulphide bonds, as demonstrated elsewhere (LARTIGUE, 2003; TEGONI; CAMPANACCI; CAMBILLAU, 2004).

Although the interaction mechanisms between OBPs and the OBP/ligand complex with ORs are still not well understood, studies showed that after binding the ligand, some OBPs are induced to a conformational change. These pH-dependent conformational changes were associated with changes in binding affinity, and it was reported to be common to some OBPs in distinct insects (LEITE et al., 2009; MAO et al., 2010;

SANDLER et al., 2000; WOGULIS et al., 2006; XU; LEAL, 2008; XU et al., 2010). The OBP/ligand complex releases the ligand after they reach specific odorant receptors (FAN et al., 2011). Therefore, amino acid changes that lead to conformational changes, such as those found in *A. fraterculus* and *A. obliqua* OBPs may also interfere with the interactions between OBPs and ORs, even when they are not directly involved with the binding sites.

2.4.5. OBPs' molecular evolution in *Anastrepha fraterculus* and *Anastrepha obliqua* may reflect specific adaptation

Despite the similar number of OBPs and their sequences, there is significant difference between OBPs of *A. obliqua* and *A. fraterculus* that may reflect specific adaptation. A wind tunnel test revealed that *A. fraterculus* antennae were responsive to Myrtaceae extracts, which also affected its oviposition rate (GREGORIO; SANT'ANA; REDAELLI, 2010). On the other hand, adults of *A. obliqua* were attracted to Anacardiaceae ripe fruits (CRUZ-LÓPEZ et al., 2006; LÓPEZ-GUILLÉN et al., 2011), so much so that at least nine volatile compounds from the Anacardiaceae *Spondias mombin* elicited antennal response from both sexes of this species (CRUZ-LÓPEZ et al., 2006). Therefore, despite being considered generalists (MORGANTE et al., 1993), both species have their preferences and show some host specificity for oviposition and feeding.

Species of *Anastrepha* in general show lekking behavior (ALUJA et al., 2000), which has been described for *A. obliqua* (ALUJA et al., 2000; ALUJA et al., 1989), as well as *A. fraterculus* (SEGURA et al., 2007). In these leks, a number of males compete for space and display to have access to females. Several factors have been associated with success in the leks, chief among them their pheromones (SEGURA et al., 2007). A comparison revealed that even though *A. fraterculus* and *A. obliqua* shared common chemical compounds on their pheromones, they also showed several different compounds emitted by calling males (LÓPEZ-GUILLÉN et al., 2011). These ecological and reproductive differences may have been the driving force behind the rapid rates of

evolution we identified amongst their OBP sequences, and suggest that the evolution of OBP genes may have had a significant impact in the evolution of species differences in this group.

2.5. Conclusions

In this study, we used transcriptome data to identify over 23 different OBP genes in *A. fraterculus* and *A. obliqua*, which is the largest and the most diverse number of OBPs yet reported for a Tephritidae. We found great similarity in amino acid and DNA sequences among orthologous OBPs in *A. fraterculus* and *A. obliqua*, which may be a reflection of their recent divergence or evolutionary conservatism. However, OBPs from *A. fraterculus* and *A. obliqua* showed a faster rate of evolution when comparing to other genes among these species, a higher Ka/Ks ratio and evidence of positive selection on at least four Anastrepha OBP genes: *OBP56h-1, OBP56h-2, OBP57c* and *OBP50a*. We also found four positively selected sites in which site-specific changes would radically change amino acid properties, and likely promote structural and functional changes. One amino acid under positive selection in *OBP50a* is located in the binding cavity according the putative 3D-structure inference, which is important because such change may promote functional divergence of the binding specificities, and enable this protein to recognize and bind a new odorant. The other changes that are not directly involved with the binding function may also be important because they may alter the size and shape of the binding cavity or the solubility of the whole molecule. Considering that, as was shown in other insects, few amino acid changes in OBPs may result in significant differences in olfactory responses, our results stress out the importance of OBPs for the evolution and divergence of *A. fraterculus* and *A. obliqua*.

2.6. References

AHMED, T. et al. Three amino acid residues bind corn odorants to McinOBP1 in the polyembryonic endoparasitoid of *Macrocentrus cingulum* Brischke. **PLoS ONE**, v. 9, n. 4, p. e93501, 2014.

ALTSCHUL, S. et al. Basic local alignment search tool. **Journal of Molecular Evolution**, v. 215, p. 403–410, 1990.

ALUJA, M. et al. Behaviour of *Anastrepha ludens*, *A. obliqua* and *A. serpentina* (Diptera:Tephritidae) on a wild mango tree (*Mangifera indica*) harbouring three McPhail traps. **International Journal of Tropical Insect Science**, v. 10, n. 03, p. 309–318, 1989.

ALUJA, M. et al. Behavior of flies in the genus *Anastrepha* (Trypetinae: Toxotripanini). In: ALUJA, M.; NORRBOM, A.L. **Fruit flies (Tephritidae): phylogeny and evolution of behavior**. CRC Press, Washington, DC, USA, 2000. p. 375–408.

ANISIMOVA, M.; BIELAWSKY, J.; YANG, Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. **Molecular Biology and Evolution**, v. 18, p. 1585–1592, 2001.

BENTON, R. On the ORigin of smell: odorant receptors in insects. **Cellular and Molecular Life Sciences**, v. 63, n. 14, p. 1579–1585, 2006.

CAMARGO, C. A.; ODELL, E.; JIRÓN, L. F. Interspecific interactions and host preference of *Anastrepha obliqua* and *Ceratitis capitata* (Diptera: Tephritidae), two pests of mango in Central America. **Florida Entomologist**, v. 79, n. 2, p. 266–268, 1996.

CHOMCZYNSKI, P.; MACKEY, K. Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. **BioTechniques**, v. 19, n. 6, p. 942–945, 1995.

CONANT, G. C.; WAGNER, G. P.; STADLER, P. F. Modeling amino acid substitution patterns in orthologous and paralogous genes. **Molecular Phylogenetics and Evolution**, v. 42, n. 2, p. 298–307, 2007.

CRUZ-LÓPEZ, L. et al. A new potential attractant for *Anastrepha obliqua* from *Spondias mombin* fruits. **Journal of Chemical Ecology**, v. 32, n. 2, p. 351–365, 2006.

DELPORT, W. et al. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. **Bioinformatics**, v. 26, n. 19, p. 2455–2457, 2010.

FAN, J. et al. An overview of odorant-binding protein functions in insect peripheral olfactory reception. **Genetics and molecular research: GMR**, v. 10, n. 4, p. 3056–3069, 2011.

FLETCHER, W.; YANG, Z. The effect of insertions, deletions, and alignment errors on the Branch-site test of positive selection. **Molecular Biology and Evolution**, v. 27, n. 10, p. 2257–2267, 2010.

FORÊT, S.; MALESZKA, R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). **Genome research**, v. 16, n. 11, p. 1404–1413, 2006.

GALINDO, K.; SMITH, D. P. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. **Genetics**, v. 159, n. 3, p. 1059–1072, 2001.

GARDINER, A. et al. *Drosophila* chemoreceptor gene evolution: selection, specialization and genome size. **Molecular Ecology**, v. 17, n. 7, p. 1648–1657, 2008.

GONG, D.-P. et al. The odorant binding protein gene family from the genome of Silkworm, *Bombyx mori*. **BMC Genomics**, v. 10, n. 1, p. 332, 2009.

GOTZEK, D. et al. Odorant binding proteins of the Red imported fire ant, *Solenopsis invicta*: An example of the problems facing the analysis of widely divergent proteins. **PLoS ONE**, v. 6, n. 1, p. e16289, 2011.

GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v. 29, n. 7, p. 644–652, 2011.

GREGORIO, P. L. F.; SANT'ANA, J.; REDAELLI, L. R. Percepção química e visual de *Anastrepha fraterculus* (Diptera, Tephritidae) em laboratório. **Iheringia. Série Zoologia**, v. 100, n. 2, 2010.

GUINDON, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. **Systematic biology**, v. 59, n. 3, p. 307–321, 2010.

HALL, T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95–98, 1999.

HEKMAT-SCAFE, D. S. et al. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. **Genome research**, v. 12, n. 9, p. 1357–1369, 2002.

HILLER, K. et al. PrediSi: prediction of signal peptides and their cleavage positions. **Nucleic acids research**, v. 32, n. Web Server issue, p. W375–379, 2004.

KATOH, K.; STANDLEY, D. M. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. **Molecular Biology and Evolution**, v. 30, n. 4, p. 772–780, 2013.

KELLEY, L. A. et al. The Phyre2 web portal for protein modeling, prediction and analysis. **Nature Protocols**, v. 10, n. 6, p. 845–858, 2015.

KIRKNESS, E. F. et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. **Proceedings of the National Academy of Sciences**, v. 107, n. 27, p. 12168–12173, 2010.

KOSIOL, C. et al. Patterns of positive selection in six mammalian genomes. **PLoS Genetics**, v. 4, n. 8, p. e1000144, 2008.

KRYAZHIMSKIY, S.; PLOTKIN, J. B. The population genetics of dN/dS. **PLoS Genetics**, v. 4, n. 12, p. e1000304, 2008.

LAGARDE, A. et al. The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. **Journal of molecular biology**, v. 414, n. 3, p. 401–412, 2011.

LARKIN, M. A. et al. Clustal W and Clustal X version 2.0. **Bioinformatics (Oxford, England)**, v. 23, n. 21, p. 2947–2948, 2007.

LARTIGUE, A. The crystal structure of a cockroach Pheromone-binding protein suggests a new ligand binding and release mechanism. **Journal of Biological Chemistry**, v. 278, n. 32, p. 30213–30218, 2003.

LAUGHLIN, J. D. et al. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. **Cell**, v. 133, n. 7, p. 1255–1265, 2008.

LEITE, N. R. et al. Structure of an Odorant-binding protein from the mosquito *Aedes aegypti* suggests a binding pocket covered by a pH-Sensitive "Lid". **PLoS ONE**, v. 4, n. 11, p. e8006, 2009.

LI, H. et al. Identification of differential expression genes associated with host selection and adaptation between two sibling insect species by transcriptional profile analysis. **BMC Genomics**, v. 14, n. 1, p. 582, 2013.

LÓPEZ-GUILLÉN, G. et al. Olfactory responses of *Anastrepha obliqua* (Diptera: Tephritidae) to volatiles emitted by calling males. **Florida Entomologist**, v. 94, n. 4, p. 874–881, 2011.

MANOHARAN, M. et al. Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. **Genome Biology and Evolution**, v. 5, n. 1, p. 163–180, 2013.

MAO, Y. et al. Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. **Proceedings of the National Academy of Sciences**, v. 107, n. 44, p. 19102–19107, 2010.

MARTIN, D. P. et al. RDP3: a flexible and fast computer program for analyzing recombination. **Bioinformatics**, v. 26, n. 19, p. 2462–2463, 2010.

MARTIN, D.; RYBICKI, E. RDP: detection of recombination amongst aligned sequences. **Bioinformatics**, v. 16, n. 6, p. 562–563, 2000.

MATSUO, T. et al. Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. **PLoS Biology**, v. 5, n. 5, p. e118, 2007.

MORGANTE, J. et al. Evolutionary patterns in specialist and generalist species of *Anastrepha*. In: ALUJA, M.; LIEDO, P. **Fruit flies Biology and Management**. New York, NY: SpringerV erlag, 1993. p. 15–20.

NEI, M.; ROONEY, A. P. Concerted and Birth-and-death evolution of multigene families. **Annual Review of Genetics**, v. 39, n. 1, p. 121–152, 2005.

NOOR, M. A. F.; FEDER, J. L. Speciation genetics: evolving approaches. **Nature Reviews Genetics**, v. 7, n. 11, p. 851–861, 2006.

OHTA, S. et al. Identification of odorant-binding protein genes expressed in the antennae and the legs of the onion fly, *Delia antiqua* (Diptera: Anthomyiidae). **Applied Entomology and Zoology**, v. 49, n. 1, p. 89–95, 2014.

ORR, H. A.; IRVING, S. Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota-USA hybridization. **Genetics**, v. 158, n. 3, p. 1089–1100, 2001.

PELOSI, P.; MAIDA, R. Odorant-binding proteins in insects. **Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology**, v. 111, n. 3, p. 503–514, 1995.

PEREIRA-RÊGO, D. R. G. et al. Morfometria de *Anastrepha fraterculus* (Wied) (Diptera: Tephritidae) relacionada a hospedeiros nativos, Myrtaceae. **Arquivos do Instituto Biológico**, v. 78, p. 3743, 2011.

PERRE, P. et al. Morphometric Differentiation of Fruit Fly Pest Species of the *Anastrepha fraterculus* GGroup (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 107, n. 2, p. 490–495, 2014.

POSADA, D. Using MODELTEST and PAUP* to select a model of nucleotide substitution. **Current Protocols in Bioinformatics,** Chapter 6, p. unit 6.5, 2003.

POSADA, D. jModelTest: phylogenetic model averaging. **Molecular biology and evolution**, v. 25, n. 7, p. 1253–1256, 2008.

QIAO, H. et al. Discrimination of alarm pheromone (E)-β-farnesene by aphid odorant-binding proteins. **Insect Biochemistry and Molecular Biology**, v. 39, n. 5-6, p. 414–419, 2009.

RAMSDELL, K. M. M. et al. Expressed sequence tags from cephalic chemosensory organs of the northern walnut husk fly, *Rhagoletis suavis*, including a putative canonical odorant receptor. **Journal of insect science (Online)**, v. 10, p. 51, 2010.

RICE, W. R. Sexually antagonistic genes: experimental evidence. **Science**, v. 256, n. 5062, p. 1436–1439, 1992.

SÁNCHEZ-GRACIA, A.; AGUADÉ, M.; ROZAS, J. Patterns of nucleotide polymorphism and divergence in the odorant-binding protein genes OS-E and OS-F: analysis in the *melanogaster* species subgroup of *Drosophila*. **Genetics**, v. 165, n. 3, p. 1279–1288, 2003.

SÁNCHEZ-GRACIA, A.; ROZAS, J. Divergent evolution and molecular adaptation in the *Drosophila* odorant-binding protein family: inferences from sequence variation at the OS-E and OS-F genes. **BMC evolutionary biology**, v. 8, p. 323, 2008.

SANDLER, B. H. et al. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. **Chemistry & biology**, v. 7, n. 2, p. 143–151, 2000.

SAWYER, S. Statistical tests for detecting gene conversion. **Molecular biology and evolution**, v. 6, n. 5, p. 526–538, 1989.

SCHWARZ, D. et al. Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella*. **BMC Genomics**, v. 10, n. 1, p. 633, 2009.

SEGURA, D. F. et al. Lekking behavior of *Anastrepha fraterculus* (Diptera: Tephritidae). **The Florida Entomologist**, v. 90, n. 1, p. 154–162, 2007.

SELIVON, D.; PERONDINI, A. L. P.; MORGANTE, J. S. A genetic–morphological characterization of two cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 98, n. 3, p. 367–381, 2005.

SICILIANO, P. et al. Sniffing out chemosensory genes from the Mediterranean fruit fly, *Ceratitis capitata*. **PLoS ONE**, v. 9, n. 1, p. e85523, 2014.

SMITH, J. Analyzing the mosaic structure of genes. **Journal of Molecular Evolution**, v. 34, p. 126–129, 1992.

SWANSON, W. J.; VACQUIER, V. D. The rapid evolution of reproductive proteins. **Nature Reviews. Genetics**, v. 3, n. 2, p. 137–144, 2002.

SWANSON, W. J. et al. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. **Genetics**, v. 168, n. 3, p. 1457–1465, 2004.

TEGONI, M.; CAMPANACCI, V.; CAMBILLAU, C. Structural aspects of sexual attraction and chemical communication in insects. **Trends in Biochemical Sciences**, v. 29, n. 5, p. 257–264, 2004.

TSITSANOU, K. E. et al. Crystal and solution studies of the "Plus-C" odorant-binding protein 48 from Anopheles gambiae: control of binding specificity through three-dimensional domain swapping. **The Journal of biological chemistry**, v. 288, n. 46, p. 33427–33438, 2013.

VANDEWEGE, M. W. et al. Evolution of the ABPA subunit of androgen-binding Protein expressed in the submaxillary glands in new and old world Rodent taxa. **Journal of Molecular Evolution**, v. 76, n. 5, p. 324–331, 2013.

VIEIRA, F. G. et al. Unique features of Odorant-binding proteins of the parasitoid wasp *Nasonia vitripennis* revealed by genome annotation and comparative analyses. **PLoS ONE**, v. 7, n. 8, p. e43034, 2012.

VIEIRA, F. G.; ROZAS, J. 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**, v. 3, p. 476–490, 2011.

VIEIRA, F. G.; SÁNCHEZ-GRACIA, A.; ROZAS, J. Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. **Genome Biology**, v. 8, n. 11, p. R235, 2007.

VOGT, R. G.; RIDDIFORD, L. M. Pheromone binding and inactivation by moth antennae. **Nature**, v. 293, n. 5828, p. 161–163, 1981.

WANG, P. et al. Association of polymorphisms in Odorant-binding protein genes with variation in olfactory response to benzaldehyde in *Drosophila*. **Genetics**, v. 177, n. 3, p. 1655–1665, 2007.

WANG, S.-Y. et al. Specific involvement of two amino acid residues in cis-nerolidol binding to odorant-binding protein 5 AlinOBP5 in the alfalfa plant bug, *Adelphocoris lineolatus* (Goeze). **Insect Molecular Biology**, v. 22, n. 2, p. 172–182, 2013.

WOGULIS, M. et al. The crystal structure of an odorant binding protein from *Anopheles gambiae*: Evidence for a common ligand release mechanism. **Biochemical and Biophysical Research Communications**, v. 339, n. 1, p. 157–164, 2006.

XIA, X. DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. **Molecular Biology and Evolution**, v. 30, n. 7, p. 1720–1728, 2013.

XU, W.; LEAL, W. S. Molecular switches for pheromone release from a moth pheromone-binding protein. **Biochemical and Biophysical Research Communications**, v. 372, n. 4, p. 559–564, 2008.

XU, X. et al. NMR structure of navel orangeworm moth pheromone-binding protein (AtraPBP1): implications for pH-sensitive pheromone detection. **Biochemistry**, v. 49, n. 7, p. 1469–1476, 2010.

YANG, Z. PAML 4: Phylogenetic analysis by maximum likelihood. **Molecular Biology and Evolution**, v. 24, n. 8, p. 1586–1591, 2007.

ZHANG, J. Evaluation of an improved Branch-site likelihood method for detecting positive selection at the molecular level. **Molecular Biology and Evolution**, v. 22, n. 12, p. 2472–2479, 2005.

ZHANG, Z. et al. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. **Genomics, Proteomics & Bioinformatics**, v. 4, n. 4, p. 259–263, 2006.

ZHENG, W. et al. Identification and expression profile analysis of Odorant binding bproteins in the Oriental fruit fly *Bactrocera dorsalis*. **International Journal of Molecular Sciences**, v. 14, n. 7, p. 14936–14949, 2013.

ZHOU, J.-J. et al. Revisiting the odorant-binding protein LUSH of *Drosophila melanogaster*: evidence for odour recognition and discrimination. **FEBS letters**, v. 558, n. 1-3, p. 23–26, 2004.

ZHOU, J.-J. et al. Identification of odorant-binding proteins of the yellow fever mosquito *Aedes aegypti*: genome annotation and comparative analyses. **Insect molecular biology**, v. 17, n. 2, p. 147–163, 2008.

ZHOU, J.-J. et al. Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrthosiphon pisum*. **Insect molecular biology**, a. v. 19 Suppl 2, p. 113–122, 2010.

ZHOU, S. et al. Plasticity of the chemoreceptor repertoire in *Drosophila melanogaster*. **PLoS Genetics**, v. 5, n. 10, p. e1000681, 2009.

ZHU, J.-Y. et al. Identification and tissue distribution of odorant binding protein genes in the beet armyworm, *Spodoptera exigua*. **Journal of Insect Physiology**, v. 59, n. 7, p. 722–728, 2013.

Additional files

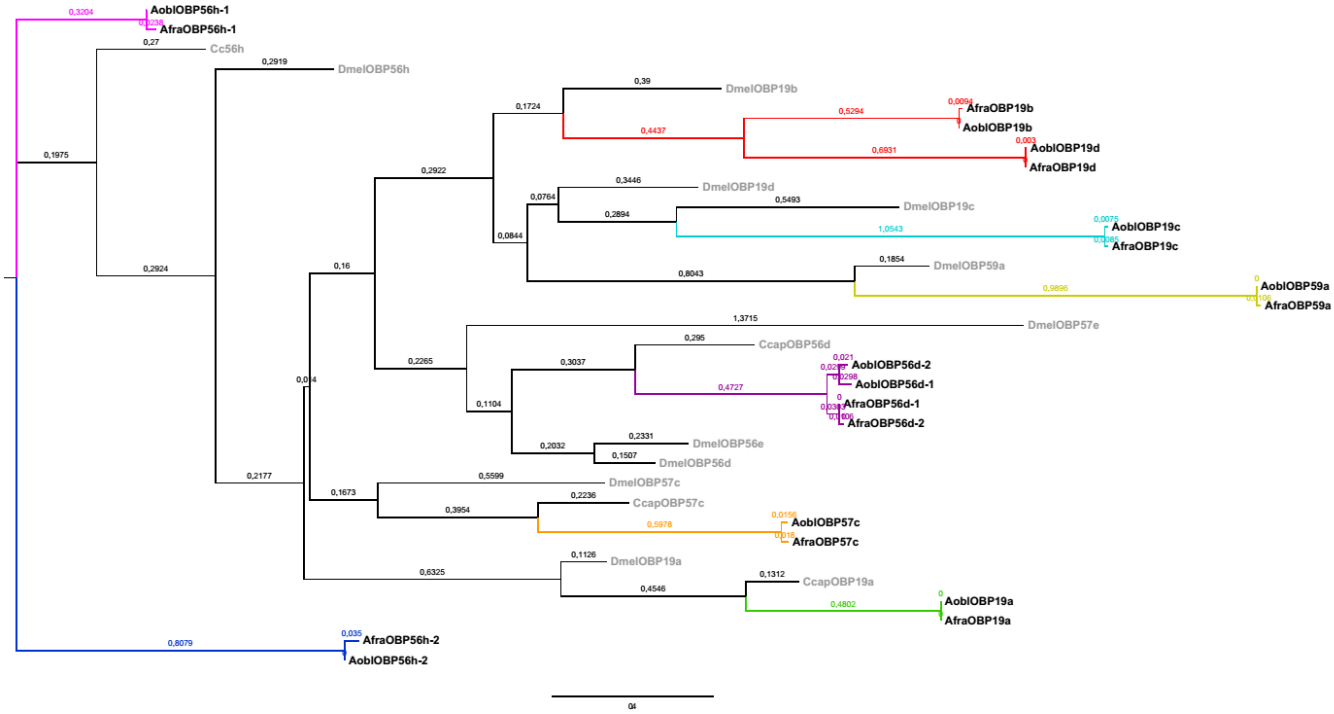
Additional file 1. Odorant-binding proteins genes from *Ceratitis capitata*, *Drosophila melanogaster*, *Bactrocera dorsalis*, *Rhagoletis suavis* and *Rhagoletis pomonella* used in the phylogenetic analyses.

Species	Gene	Subfamily	Accession Number
C. capitata	*CcapOBP8a*	minus-C	XM_004521128.1
C. capitata	*CcapOBP19a*	classic	XM_004524969.1
C. capitata	*CcapOBP19b*	minus-C	XM_004524970.1
C. capitata	*CcapOBP19d*	classic	XM_004524978.1
C. capitata	*CcapOBP44a*	minus-C	XM_004535885.1
C. capitata	*CcapOBP49a*	plus-C	XM_004522926.1
C. capitata	*CcapOBP56d*	classic	XM_004517746.1
C. capitata	*CcapOBP56h*	classic	XM_004517747.1
C. capitata	*CcapOBP57c*	classic	XM_004522799.1
C. capitata	*CcapOBP99b*	minus-C	XM_004521126.1
C. capitata	*CcapOBP99c*	minus-C	XM_004521129.1
C. capitata	*CcapOBP99d*	minus-C	XM_004521127.1
D. melanogaster	*DmelOBP8a*	minus-C	NM_167184.1
D. melanogaster	*DmelOBP19a*	classic	NM_167700.2
D. melanogaster	*DmelOBP19b*	classic	NM_134547.2
D. melanogaster	*DmelOBP19c*	classic	NM_134548.2
D. melanogaster	*DmelOBP19d*	classic	NM_078697.2
D. melanogaster	*DmelOBP47b*	plus-C	NM_136825.2
D. melanogaster	*DmelOBP49a*	plus-C	NM_136968.2
D. melanogaster	*DmelOBP50a*	plus-C	NM_166050.2
D. melanogaster	*DmelOBP50e*	plus-C	NM_137115.2
D. melanogaster	*DmelOBP56d*	classic	NM_137600.2
D. melanogaster	*DmelOBP56e*	classic	NM_137601.3
D. melanogaster	*DmelOBP56h*	classic	NM_137604.2
D. melanogaster	*DmelOBP57c*	classic	NM_137637.2
D. melanogaster	*DmelOBP57e*	classic	NM_137644.1
D. melanogaster	*DmelOBP59a*	classic	NM_176249.2
D. melanogaster	*DmelOBP83cd*	dimer	NM_141355.2
D. melanogaster	*DmelOBP83ef*	dimer	NM_169125.2
D. melanogaster	*DmelOBP83g*	minus-C	NM_169126.1
D. melanogaster	*DmelOBP99a*	minus-C	NM_143450.3
D. melanogaster	*DmelOBP99b*	minus-C	NM_143456.3
D. melanogaster	*DmelOBP99c*	minus-C	NM_143454.2
D. melanogaster	*DmelOBP99d*	minus-C	NM_143455.1
B.dorsalis	*BdorOBP1*	classic	KC559112.1
B.dorsalis	*BdorOBP2*	classic	KC559113.1
B.dorsalis	*BdorOBP3*	classic	KC559114.1
B.dorsalis	*BdorOBP4*	classic	KC559115.1
B.dorsalis	*BdorOBP5*	classic	KC559116.1
B.dorsalis	*BdorOBP6*	minus-C	KC559117.1

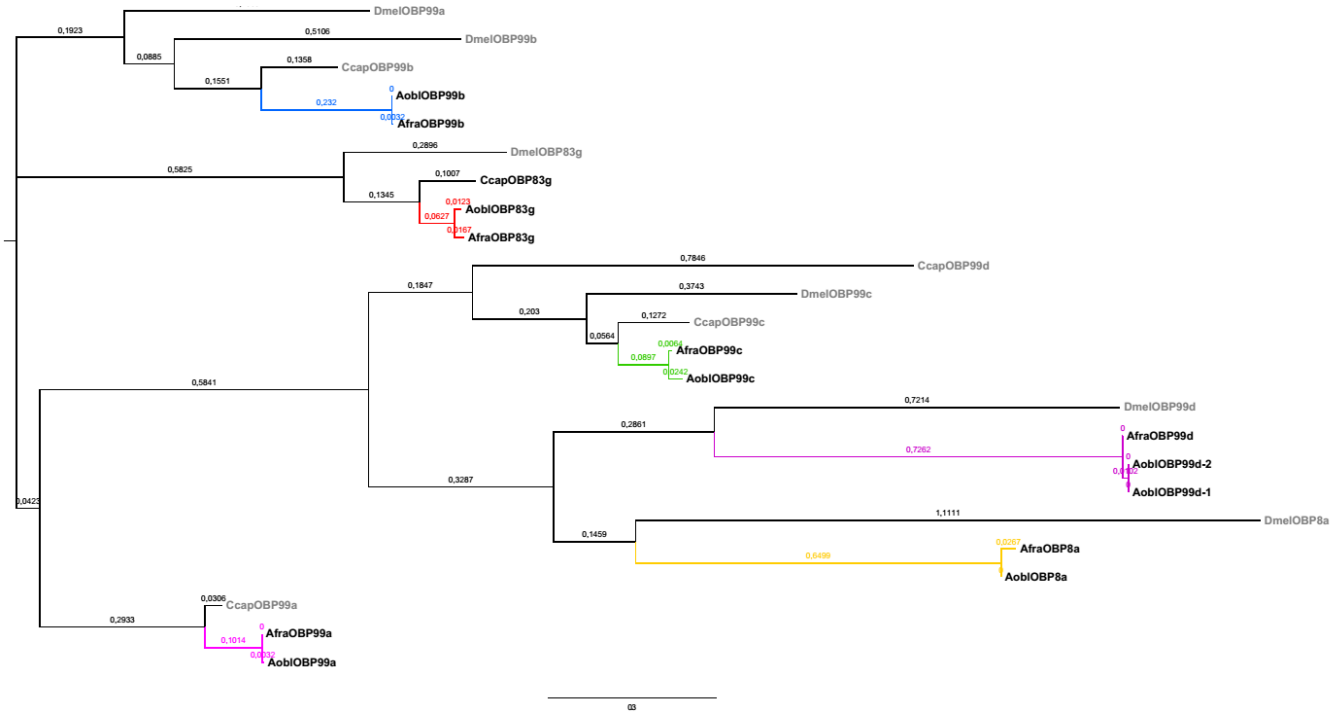
B.dorsalis	*BdorOBP7*	classic	KC559118.1
B.dorsalis	*BdorOBP8*	dimer	KC559119.1
B.dorsalis	*BdorOBP9*	classic	KC559120.1
B.dorsalis	*BdorOBP10*	minus-C	KC559121.1
R. suavis	*RsuaOBP19d*	classic	EX453827.1
R. suavis	*RsuaOBP56g*	classic	EX453825.1
R. pomonella	*RpomOBP19a*	classic	EZ126705.1
R. pomonella	*RpomOBP19b*	classic	Z138033.1
R. pomonella	*RpomOBP44a*	minus-C	EZ137697.1
R. pomonella	*RpomOBP50e*	plus-C	EZ131620.1
R. pomonella	*RpomOBP56a*	classic	EZ139735.1
R. pomonella	*RpomOBP56d*	classic	EZ116239.1
R. pomonella	*RpomOBP56h*	classic	EZ138985.1
R. pomonella	*RpomOBP83cd*	dimer	EZ136344.1
R. pomonella	*RpomOBP83ef*	dimer	EZ137089.1
R. pomonella	*RpomOBP83g*	minus-C	EZ136242.1
R. pomonella	*RpomOBP99b*	minus-C	EZ135703.1
R. pomonella	*RpomOBP99c*	minus-C	EZ138892.1
R. pomonella	*RpomOBP99d*	minus-C	EZ119053.2

Additional file 2. Non-rooted phylogenetic trees by subfamily used in the PAML evolutionary analysis. Each colored branch represent a different branch-site test analysis. Branch lengths are estimated by amino acid substitutions per site.

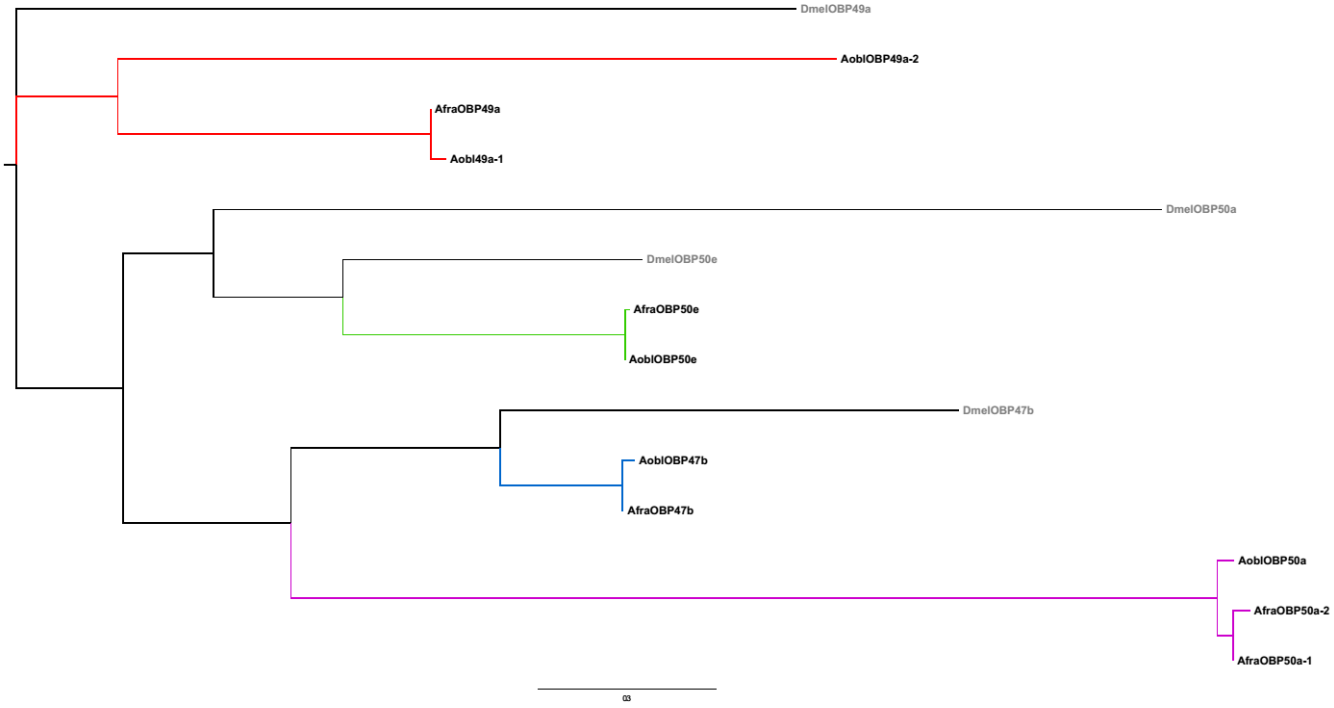
Classic subfamily



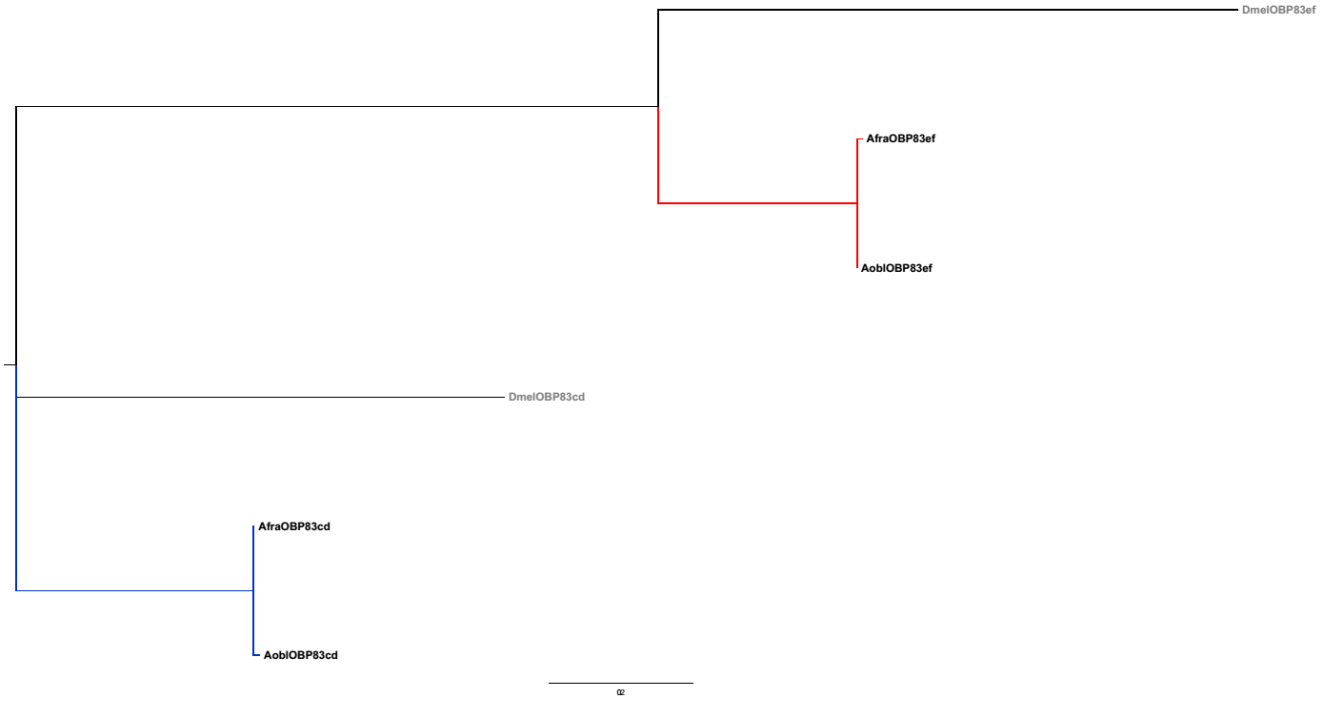
Minus-C subfamily



Plus-C subfamily



Dimer subfamily



Capítulo III

Análise de padrão de expressão gênica em genes OBP de *Anastrepha fraterculus* e *Anastrepha obliqua*

Capítulo III – Análise de padrão de expressão gênica em genes OBP de *Anastrepha fraterculus* e *Anastrepha obliqua*

Considerações Iniciais

Nossa investigação a partir dos transcriptomas de cabeça e tecido reprodutivo de *A. fraterculus* e *A. obliqua* nos permitiu identificar pela primeira vez genes OBPs nessas espécies, bem como averiguar os padrões moleculares de divergência dos membros dessa família gênica. Embora tenhamos encontrado importantes diferenças moleculares nos genes OBP entre essas duas espécies, muitas vezes as diferenças observadas entre as espécies não se dão devido às diferenças entre suas sequências gênicas, mas sim por diferenças no padrão de expressão dos genes. Dessa forma, um trabalho sobre a análise das diferenças entre as OBPs de *A. fraterculus* e *A. obliqua* não ficaria completo se não investigássemos também a ocorrência de expressão diferencial dos genes OBP nessas espécies.

Para essa análise, buscamos nos transcriptomas de tecidos cefálicos genes OBP diferencialmente expressos entre as duas espécies, considerando tanto machos quanto fêmeas. Estes genes foram então avaliados por PCR quantitativa (qPCR) em diferentes perfis reprodutivos de *A. fraterculus* e *A. obliqua*. Também consideramos para as análises de expressão por qPCR os genes identificados sob seleção positiva na etapa anterior desse trabalho. Optamos por realizar as análises de expressão utilizando apenas tecido cefálico, não incluindo o tecido reprodutivo como anteriormente, uma vez que é na cabeça que estão localizados os órgãos olfatórios das moscas. Tais dados serão apresentados a seguir em forma de artigo científico a ser submetido à revista “Scientific Reports”.

Differential expression of nine Odorant-binding protein genes in *Anastrepha fraterculus* and *Anastrepha obliqua* fruit flies

3.1. Abstract

The economically important South American *Anastrepha* fruit flies of the *fraterculus* group exhibit limited number of distinguishing characters and have insipient reproductive isolation among some of its members, which suggests that their differentiation may have been recent. We studied the differential expression of Odorant-Binding Proteins (OBPs) gene family, which participate in initial steps of the olfactory signal transduction cascade, in two closely related species of the *fraterculus* group, *Anastrepha fraterculus* and *A. obliqua*. These species show important ecological and reproductive differences related with olfactory perception, for instance different host preferences and ability to recognize species-specific pheromones and thus avoid interspecific matings. Therefore, we expected differences in the expression levels of OBP genes putatively involved with their divergence. We selected nine OBP genes, six that were differentially expressed in cDNA libraries derived from adult heads between *A. fraterculus* and *A. obliqua* (*OBP56a*, *OBP56d*, *OBP56h-2*, *OBP83cd*, *OBP99c* and *OBP99d*), and three OBP genes that had been previously reported to be under positive selection (*OBP50a*, *OBP56h-1*, *OBP57c*), and investigated their patterns of expression by qPCR in male and female head tissues from these species at several reproductive stages. According to the expression patterns obtained at different reproductive stages in each species, we hypothesized that some OBP genes may be potentially involved with the location of food resources (*OBP56a* in *A. obliqua* males and females, *OBP56d* and *OBP99c* in *A. fraterculus* males and *OBP83cd* for males of both species), while others seem to be involved in pheromone perception (*OBP56h-2*, *OBP99d* and *OBP56d*, in both species, whereas other OBPs showed particular patterns of expression, changing according to sex and/or species considered). We also observed that expression levels of OBPs changed significantly after mating for males and females of both species. In the interspecific comparisons, we found that all nine OBP genes here analyzed showed significant levels of differential expression between *A. fraterculus* and *A. obliqua*, especially *OBP56a*, *OBP56d*, *OBP57c* and the two copies of *OBP56h*, suggesting that they may hold important roles in their olfactory perception differences, and consequently, potentially involved in their differentiation.

Keywords: Real time PCR; Tephritidae; *fraterculus* group; Gene expression; Genetics of speciation.

3.2. Introduction

Fruit flies from the *Anastrepha fraterculus* group (Diptera:Tephritidae) have diverged recently, exhibiting limited number of morphologic and genetic distinguishing characters (PERRE et al., 2014; SELIVON; PERONDINI; MORGANTE, 2005). Furthermore, in laboratory conditions, interspecific viable hybrids can be obtained for some of the species in the group (SANTOS; URAMOTO; MATIOLI, 2001; SELIVON; PERONDINI; MORGANTE, 2005), which suggests reproductive isolation is still insipient and speciation may be recent among species in this group. Even though knowledge on the speciation in species on this group is still scant, a lot has been learned from investigating other tephritids, such as medflies. In that case, genes involved with host race formation and olfactory and gustatory reception have been shown to have an important role in their differentiation (SCHWARZ et al., 2009). These genes may impact speciation because olfactory responses control social and sexual interactions between individuals of the same species, through the detection of odors and pheromones essential for survival and reproduction (BENTON, 2006; VIEIRA; SÁNCHEZ-GRACIA; ROZAS, 2007).

The initial steps of the transduction cascade of olfactory signals in insects, mediated by the odorant-binding proteins (OBPs), are the solubilization and transport of chemical signals through the aqueous lymph of sensillas to reach the olfactory receptors (SÁNCHEZ-GRACIA; ROZAS, 2008; VOGT; RIDDIFORD, 1981). Different OBPs have specific affinities to odorants and their high molecular divergence, as well as the distinct expression patterns reported to OBP genes in insects, suggest that these proteins act as a filter, selecting the odorants to trigger olfactory responses (FAN et al., 2011). Therefore, it is a combination of the number and types of OBPs expressed in a species, and at which developmental stage and tissues these genes are expressed, that may

determine the specificity and sensitivity to the odorants (MCGRAW et al., 2004; PELOSI; MAIDA, 1995; STEINBRECHT, 1996; VOGT, 1999).

Because OBPs are important targets for natural and sexual selection, their role in host and mate choice has previously been established in several species (MATSUO et al., 2007; LAUGHLIN et al., 2008). Here, we studied differential expression of OBP genes in two closely related species of the *fraterculus* group, *Anastrepha fraterculus* and *A. obliqua*, which are important fruit pests in South America. These species show important ecological and reproductive differences related with olfactory perception. Though *A. fraterculus* has been associated with a wide number of hosts, it prefers several Myrtaceae fruits (PEREIRA-RÊGO et al., 2011), being considered one of the main economic pests in South America. *A. obliqua*, on the other hand, though an important pest species as well, has been associated to a smaller number of hosts, several of those Anacardiaceae (CAMARGO; ODELL; JIRÓN, 1996). Olfactory reception also plays an important role in courtship behavior of these species, during lek aggregations, which is a common feature for both species. In that case, females may be able to recognize favorable male pheromones, or even species-specific pheromones and thus avoid interspecific matings (BŘÍZOVÁ et al., 2013; LÓPEZ-GUILLÉN et al., 2011; SEGURA et al., 2007).

The differential detection of host odors and reproductive partners triggers several behaviors associated with distinct combinations of odorants, may subject genes in the OBP family to strong selection pressures (VOGT et al., 2002), which can gradually lead to differentiation both to the OBP genes repertory and their expression levels. A previous study analyzed selection pressures in the OBP genes repertory of *A. fraterculus* and *A. obliqua*, identifying several genes under positive selection (CAMPANINI; DE BRITO, unpublished data), thus in the present study we investigated the patterns of expression by qPCR of nine OBP genes in male and female heads of *A. fraterculus* and *A. obliqua*. We

analyzed how these genes are expressed at different reproductive stages in each species and whether OBP genes are differentially expressed between species, which could indicate genes involved in their differentiation (GIBSON, 2002).

3.3. Methods

3.3.1. *In silico* gene expression analysis

The *in silico* OBP gene expression analysis between *A. fraterculus* and *A. obliqua* was conducted using cDNA libraries derived from adult heads, where the flies' olfactory organs (antennae and maxillary palps) are located. The data included five profiles per species at different reproductive stages: virgin and post-mating for both genders and post-oviposition for females, done in duplicates, which consisted of ten pooled individuals from a population, totalizing 20 individuals sampled by reproductive stage. We named *A. fraterculus* (Af) libraries as Af_VF = virgin female (collected 10 days after pupae eclosion); Af_PF = post-mating female (10 days old females collected 15-20 h after the first successful copula); Af_PO = post-oviposition female (post-mating females collected immediately after a successful oviposition); Af_VM = virgin male (collected 10 days after pupae eclosion); Af_PM = post-mating male (10 days old males collected 15-20 h after the first successful copula). We used a similar nomenclature for *A. obliqua* (Ao) libraries. More details about the cDNA libraries' construction and analysis are described elsewhere (REZENDE et al., unpublished data).

Differential expression analyses were conducted separately for each reproductive stage from a methodology that uses statistical tools for transcript quantification, using the scripts included in Trinity package (GRABHERR et al., 2011). We used as reference a set of 23 sequences previously identified as *A. obliqua* OBP transcripts (CAMPANINI; DE BRITO, unpublished data; NAKAMURA et al., 2016) (Additional File 1). We used sequences from a single species because orthologs are very similar between *A. obliqua* and *A. fraterculus* (CAMPANINI; DE BRITO, unpublished data). Initially, we used the

Bowtie 2 software (LANGMEAD; SALZBERG, 2012) to align the reads of each library back to the reference. The second step was to estimate transcript abundance by RSEM (LI; DEWEY, 2011), using default options. RSEM uses an iterative process to fractionally assign reads to each transcript based on the probabilities of the reads being derived from each transcript, taking into account positional biases created by RNA-seq library generating protocols (HAAS et al., 2013). Finally, differential transcript expression was quantified using EdgeR (ROBINSON; MCCARTHY; SMYTH, 2010). This program performs an additional TMM (trimmed mean of *M*-values) scaling normalization that aims to account for differences in RNA across all comparisons.

3.3.2. Selected OBP genes and qPCR primers design

We chose a set of OBP genes that was found to be under differential expression in the *in silico* analysis, as well as OBP genes that were under positive selection previously identified (CAMPANINI; DE BRITO, unpublished data) to perform quantitative PCR (qPCR) analyses. Nine OBP genes were considered (Additional File 2) and two primer pairs were designed for each one, using the software Primer 3 version 4.0.0 (available at http://bioinfo.ut.ee/primer3/). We aligned *Anastrepha* OBP sequences with their putative *D. melanogaster*'s orthologs, to identify putative intron positions. Whenever possible, primer pairs were designed across introns to control for potential amplification of contaminated DNA. We designed OBP primers with a size between 18-22 bp, annealing temperature between 55-60°C, and ~50% of GC-content, that would amplify products of 90-210 bp, avoiding dimers and hairpins. We used regions with no substitutions between *A. fraterculus* and *A. obliqua* sequences, which allowed us to use the same primers for both species. We also investigated if the selected primer region had a high coverage in our transcriptomes, with the Tablet program (MILNE et al., 2013).

Primer concentration and primer efficiency tests were conducted for the OBP primer pairs. We tested the concentrations of 0.3, 0.5 and 0.6 µM in a final reaction

volume of 10 µl, and used dissociation curve analysis to determine primer-specific amplification. Once the best concentrations were determined, primer efficiency was tested using eight cDNA concentrations from a serial dilution starting at 5 ng/µl. For both tests, we used a pool consisting of five heads of *A. fraterculus* females, with three technical replicates. Efficiency (E) between 95-105% and standard curve correlation coefficient (R^2) higher than 0.95 were our thresholds to considerer a primer pair as efficient (BUSTIN et al., 2009). When both primer pairs were efficient, we chose the one with the closest E value to 100% and R^2 value to 1. Primer pairs selected for qPCR analyses and their results for efficiency tests are shown in the Additional File 2.

3.3.3. Profiles analyzed by qPCR

Our experimental design to the qPCR analysis consisted of sixteen profiles equally divided between females and males, sampled with three biological replicates, being each replicate composed of a pool of five heads, totalizing 15 different individuals sampled per profile. The profiles analyzed represented three reproductive stages of the adult life, and were named IF = immature females (collected 24 hours after pupae eclosion); VF = virgins females (collected 10 days after pupae eclosion); PF = post-mating females (10 days old females after the first sucessful mating); IM = immature male (collected 24 hours after pupae eclosion); VM = virgin male (collected 10 days after pupae eclosion); and PM = post-mating males (10 days old males after the first sucessful mating).

For the post-mating samples, virgin 10 days old males and females, maintained in different cages since the pupae eclosion and at the same controlled environment room (25°C, 60–90% humidity and natural photoperiod), were put together to mate in a proportion of 1:1. When mating started, the couple was gently removed to another cage. Only matings with duration longer than 40 minutes were considered, to avoid incomplete matings. By doing so, we ensure that we capture the very first successful mating for both sexes. After the male had dismounted, females and males were separated and collected at

six different times: 3, 6, 9, 12, 24 and 48 hours post-mating (PF3, PF6, PF9, PF12, PF24 and PF48 profiles for females; PM3, PM6, PM9, PM12, PM24 and PM48 for males). Post-mating individuals were maintained under the same controlled environment until the dissection, to avoid external variations in gene expression levels.

3.3.4. Isolation of total RNA and cDNA synthesis

Total RNA was extracted using the TRIzol/chloroform protocol (CHOMCZYNSKI; MACKEY, 1995). RNA quality was visualized in agarose gel electrophoresis for integrity and absorbance was measured in NanoVue™ Plus Spectrophotometer (GE Helthcare). Samples were quantified with Qubit® 2.0 Fluorometer, using the RNA BR assay kit (Invitrogen™). Before transcription, total RNA was treated with DNase I Amplification Grade (Invitrogen™), according to the manufacturer's protocol, to remove residual genomic DNA. One µg of treated RNA was converted into cDNA using iScript™ cDNA Synthesis Kit (Bio-Rad) and cDNA samples were diluted to 5 ng/µl to the qPCR assays.

3.3.5. qPCR assays

We performed the qPCRs in a CFX96 Touch™ Real-Time PCR Detection Systems (Bio-Rad), using SsoFast™ EvaGreen® Supermix (Biorad). Reactions were made with a final primer concentration of 0.6 µl in a final volume of 10 µl, which was the best primer concentration in all primer concentrations tests. Cycling parameters for all primers were 30 seconds at 95°C, followed by 40 cycles of dissociation at 95 °C for 10 seconds and annealing and extension at 60 °C for 1 min. A fluorescence reading was made at the end of each extension step. For melt curve analysis we used a protocol with temperatures that varied from 65 °C to 95 °C with increments of 0.5 °C for 5 seconds and continuous fluorescent measurements. No template controls (NTC) and inter-run calibrators were included in all qPCR plates. To check reproducibility, three technical replicates were carried out for each sample.

Relative quantification were calculated for each OBP gene in the Bio-Rad CFX Manager™ software (Bio-Rad), using the comparative $2^{\Delta\Delta Ct}$ method (LIVAK; SCHMITTGEN, 2001) in contrast with three *Anastrepha* reference genes: *rpS17* (Ribosomal proteinS17), *rpL18* (Ribosomal proteinL18) and *ef1a* (Elongation factor-1α) (NAKAMURA et al., 2016). In these analyses, we contrasted both, the within species differential expression of OBP genes among different reproductive stages, and the between species differential expression of each OBP gene. Data were statistically analyzed by ANOVA and Tukey's test using Prism 5.01 software (GraphPad Software, San Diego, CA, USA).

3.4. Results and Discussion

3.4.1. Differentially expressed OBP genes according to *in silico* analysis

Differential expression analysis from RNA-seq data is possible because different transcripts have different coverage in the transcriptome, reflecting their different expression levels (HAAS et al., 2013). We compared differential expression in OBP genes between *A. fraterculus* and *A. obliqua* from RNA-seq data aiming at selecting genes to be further investigated by qPCR analysis. In our *in silico* analysis, we found six OBP genes significantly differentially expressed (FDR <0.001), four of them in females: *OBP56a* and *OBP56d* between virgins; *OBP56a* and *OBP99d* between post-mated; and *OBP56d* and *OBP56h-2* between post-oviposition (Figure 1; Table 1). For males, we found the same four found in females plus *OBP99c* and *OBP83cd*. Between virgins *OBP99d*, *OBP56d*, *OBP56h-2* and *OBP56a* were differentially expressed, and between post-mated *OBP56a*, *OBP56d*, *OBP56h-2*, *OBP83cd* and *OBP99c* (Figure 1; Table 1).

The majority of OBPs genes was up-regulated in *A. obliqua* profiles, the only exception being *OBP56a*, which is up-regulated in *A. fraterculus*, but, interestingly at a level of differential expression much higher than the other significant gene comparisons here identified. Despite the use of two library replicates and a deep sequencing coverage

for these libraries, which in general generates more informational reads and increases the statistical power to detect differentially expressed genes (TARAZONA et al., 2011), we noticed a high standard deviation in some cases, mainly Af_VM for *OBP56a* and Ao_PF for *OBP99d*, which represents fpkm values considerably different between the two libraries replicates. However, these differences do not invalidate the results obtained.

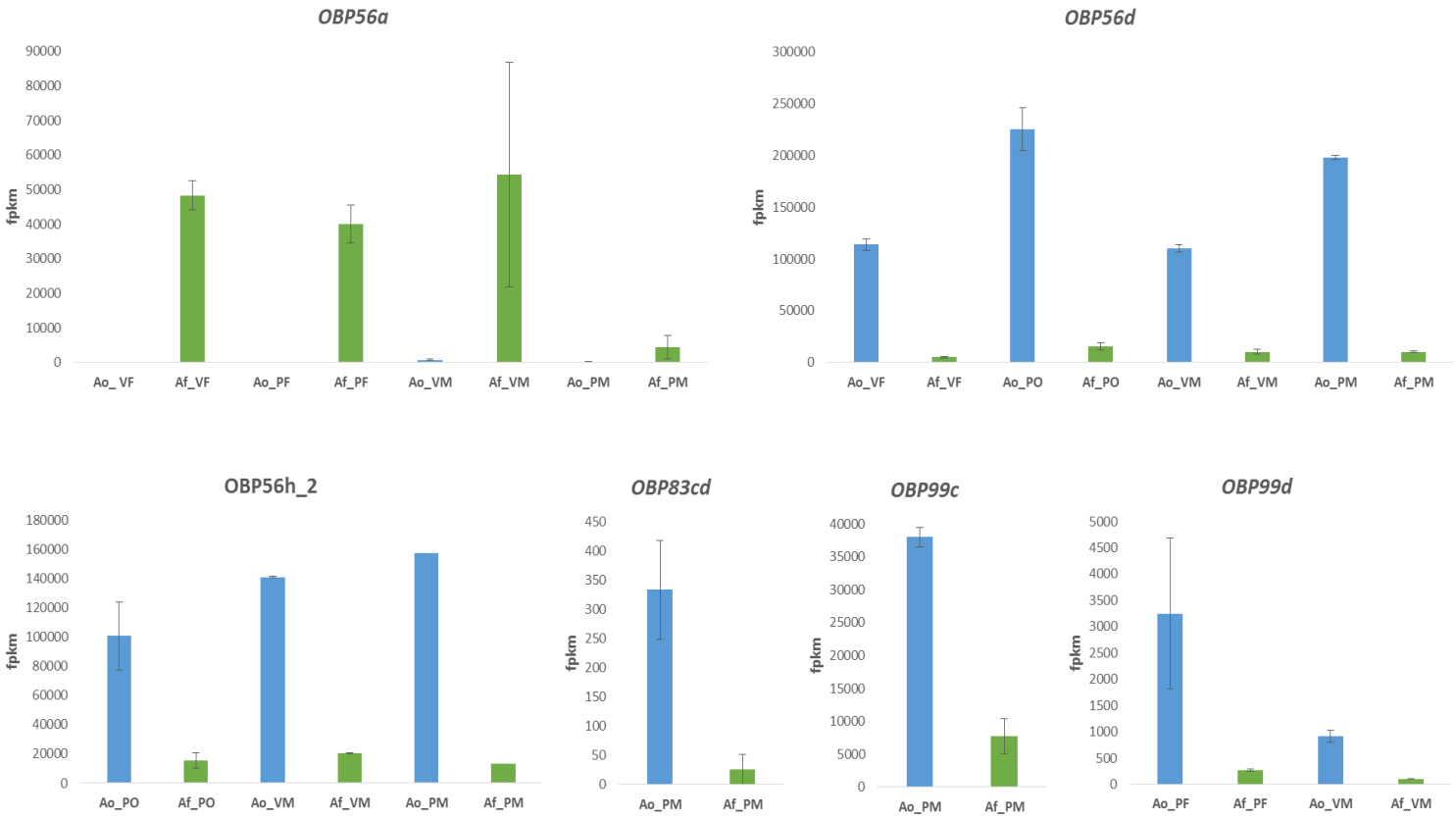


Figure 1. Differentially expressed OBP genes according to *in silico* analysis. Af = *A. fraterculus* (profiles are shown in green); Ao = *A. obliqua* (profiles are shown in blue); VF = virgin female; VM = virgin male; PF = post-mating female; PM = post-mating male; PO = post-oviposition female. RNA-seq expression was measured by Fragments Per Kilobase of transcript per Million mapped reads (fpkm).

Table 1. Significance and X-fold values of the *in silico* differential expression analysis between *A. fraterculus* and *A. obliqua* OBP genes.

Gene	Differentially expressed profile	FDR	Up-regulated species	X-fold values
OBP56a	VF	1,03E-131	*A. fraterculus*	447
	PF	2,23E-06	*A. fraterculus*	149
	VM	5,14E-04	*A. fraterculus*	65
	PM	5,62E-04	*A. fraterculus*	16
OBP56d	VF	2,49E-74	*A. obliqua*	23
	PO	1,44E-05	*A. obliqua*	6
	VM	1,03E-25	*A. obliqua*	9
	PM	3,73E-41	*A. obliqua*	19
OBP56h-2	PO	1,44E-05	*A. obliqua*	6
	VM	6,99E-23	*A. obliqua*	6
	PM	7,65E-25	*A. obliqua*	12
OBP83cd	PM	6,90E-05	*A. obliqua*	9
OBP99c	PM	6,90E-05	*A. obliqua*	5
OBP99d	PF	0,000732	*A. obliqua*	12
	VM	2,97E-06	*A. obliqua*	6

VF = virgin female; VM = virgin male; PF = post-mating female; PM = post-mating male; PO = post-oviposition female. FDR = false discovery rate.

Studies in *Drosophila* indicate that the detection of differing host odors by sibling species may be reflected in the expression level of the olfactory genes involved (DEKKER et al., 2006; KOPP et al., 2008; MCBRIDE, 2007). Therefore, we expected to find some OBP genes to show differential expression between *A. fraterculus* and *A. obliqua* because they have different host preferences and also because individuals seem to be able to recognize pheromones of their own species (CAMARGO; ODELL; JIRÓN, 1996; LÓPEZ-GUILLÉN et al., 2011; SEGURA et al., 2007; PEREIRA-RÊGO et al., 2011), which are two important ecological and reproductive attributes that may be involved in the species' differentiation. The six OBP genes revealed by our *in silico* analysis, as well as three other *Anastrepha* OBP genes that were found to be under

positive selection, *OBP50a*, *OBP56h-1* and *OBP57c* (CAMPANINI; DE BRITO, unpublished data), were thus considered for the qPCR analysis (Additional File 2).

3.4.2. Intraspecific differential expression analysis

Quantitative real-time PCR (qPCR) was carried out to determine the patterns of gene expression of the nine selected OBP genes. Initially, we wanted to investigate how these genes were expressed in each species. Several studies in insects have reported that the more relevant differences would occur between reproductive stages, instead of different tissues or life stages (AYROLES et al., 2009; DALTON et al., 2010; GONG et al., 2009; LI et al., 2013; MCGRAW et al., 2004; FORÊT; MALESZKA, 2006; PELOSI et al., 2006; ZHU et al., 2013), thus we focused our qPCR analysis in comparisons between immature, mature virgin and post-mating individuals in both males and females. Comparisons in *D. melanogaster* between virgin and mated females showed that the peak of differentially expressed genes was at 1-3 h post-mating in the whole adult (MCGRAW et al., 2004), 6 h post-mating in reproductive tissues (MACK et al., 2006) and 72 h post-mating in head and brain tissues, with the majority of genes (including *OBP99b*) showing lower expression at this later time (DALTON et al., 2010). Since there was no previous information available for *Anastrepha* species, we considered these results in *Drosophila* to sampled post-mating males and females at 3, 6, 9, 12, 24 and 48 h post-copula.

We investigated two distinct questions: 1) whether immature and mature virgin individuals had significant differential expression and 2) whether expression significantly changes after mating (Figures 2-5). We conducted and reported our species-specific analysis separately for males and females because a study in *D. melanogaster* showed sex-dependent differences in levels of expression of chemosensory genes, which indicates that they experience, interact with, and adapt to their chemical environments differently (ZHOU et al., 2009). Only significant comparisons in the Tukey's Multiple Comparison Test were considered as differentially expressed profiles ($p<0,05$).

A. fraterculus showed significant differential expression between immature and virgin females for *OBP56d*, *OBP56h-2*, *OBP99d* and *OBP50a*, all of them up-regulated in virgins (Figure 2). In males, *OBP56d*, *OBP56h-2*, *OBP57c*, *OBP99c*, *OBP99d* and *OBP83cd* were significantly differentially expressed between immature and virgins, being *OBP56d*, *OBP99c* and *OBP83cd* more expressed in immature males, and the other ones more expressed in virgins (Figure 3). Comparing the expression between virgin and post-mating profiles, we found that *OBP56a* was up-regulated 3 h post-mating in females, but all the other genes were down-regulated post-mating. In males, four genes *OBP56d*, *OBP56h-2*, *OBP57c* and *OBP99d* were up-regulated, and *OBP56h-1* was down-regulated in post-mating profiles. The other genes did not show significant differences in their expression patterns between virgin and post-mating profiles.

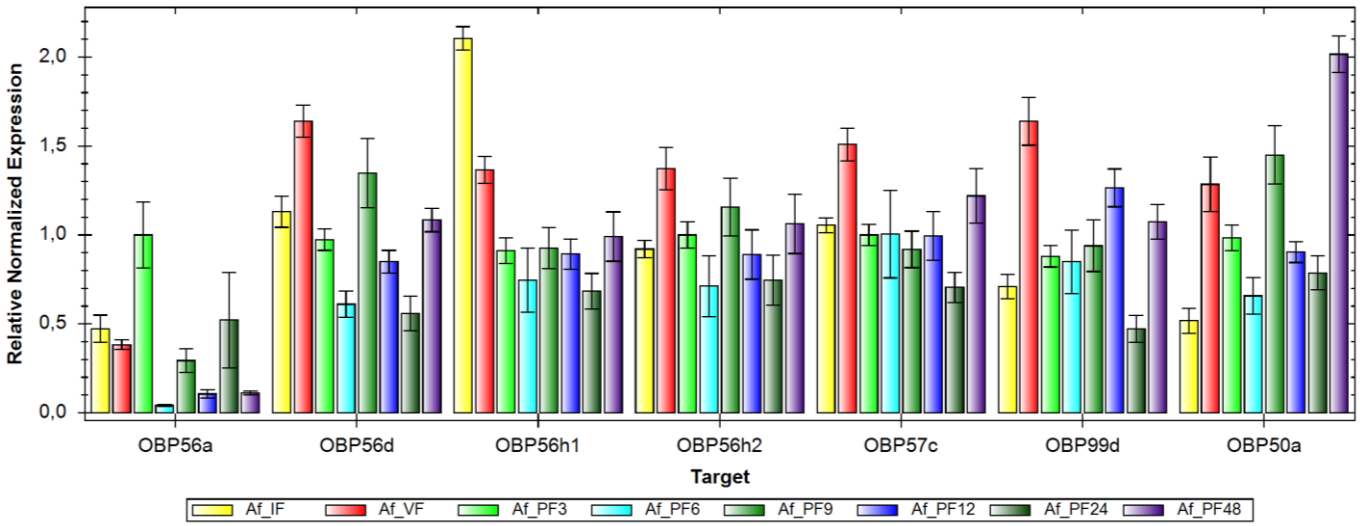


Figure 2. Relative normalized expression of OBP genes in *A. fraterculus* female reproductive stages. Af_IF = imature female; Af_VF = virgin female; Af_PF3 = 3 h post-mating female; Af_PF6 = 6 h post-mating female; Af_PF9 = 9 h post-mating female; Af_PF12 = 12 h post-mating female; Af_PF24 = 24 h post-mating female; Af_PF48 = 48 h post-mating female.

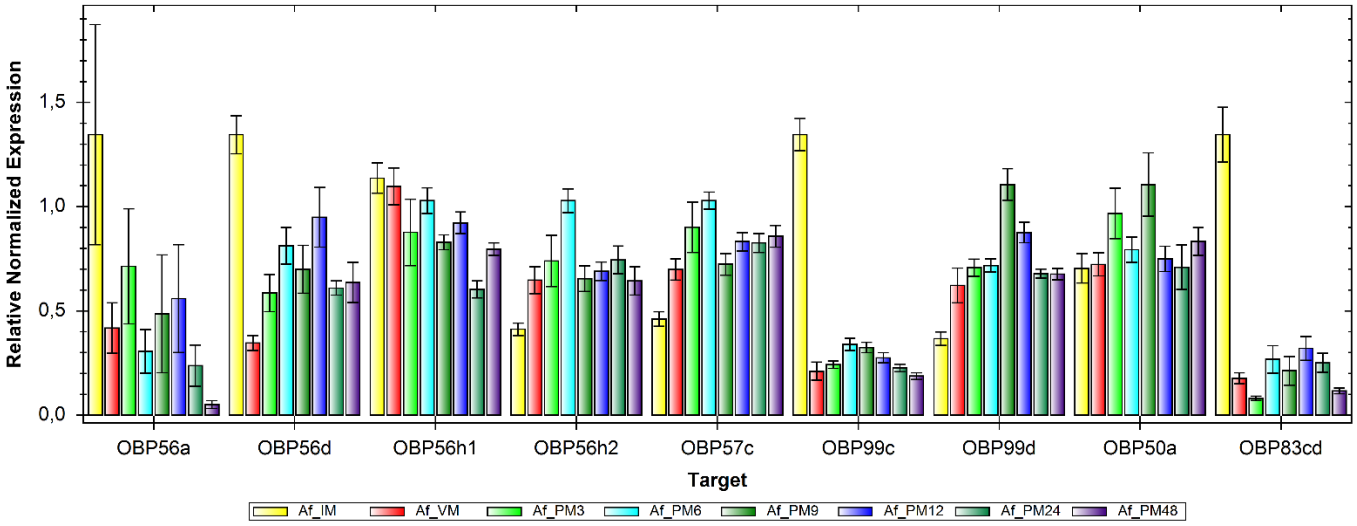


Figure 3. Relative normalized expression of OBP genes in *A. fraterculus* male reproductive stages. Af_IM = imature male; Af_VM = virgin male; Af_PM3 = 3 h post-mating male; Af_PM6 = 6 h post-mating male; Af_PM9 = 9 h post-mating male; Af_PM12 = 12 h post-mating male; Af_PM24 = 24 h post-mating male; Af_PM48 = 48 h post-mating male.

A. obliqua showed significant differential expression between immature and virgin females for *OBP56a*, *OBP56d*, *OBP56h-2*, *OBP57c* and *OBP99d*, in which only *OBP56a* was more expressed in immature females (Figure 4). All OBP genes tested were differentially expressed between immature and virgin males. *OBP56a* and *OBP83cd* were more expressed in immature males, and the other genes were more expressed in virgins (Figure 5). In females, all OBPs were up-regulated post-mating, except *OBP56a*, that did not show significant differences between virgin and post-mating profiles. Although the expression levels of *OBP99d* decreased significantly at 6 h post-mating in females, we noticed an increase at 24 h post-mating. Between virgins and post-mating males, *OBP56h-1*, *OBP57c*, *OBP99c* and *OBP99d* were down-regulated post-mating, at 3 or 6 h post-mating. Finally, *OBP56h-2* and *OBP83cd* showed peculiar patterns of expression in *A. obliqua* males, with peaks of up- and down-regulations in the profiles analyzed.

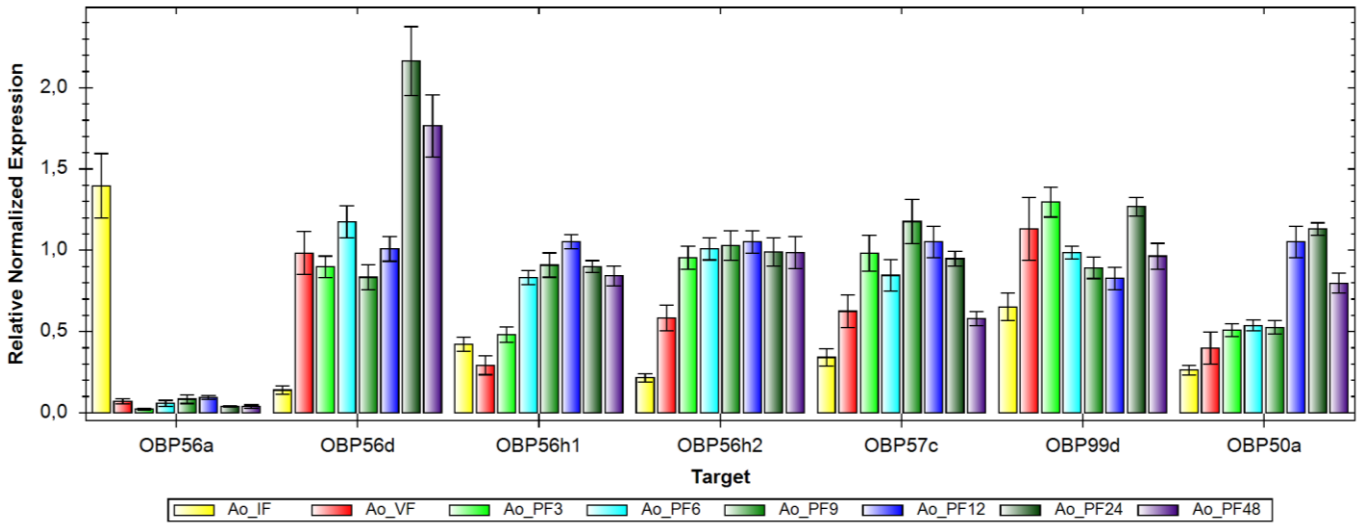


Figure 4. Relative normalized expression of OBP genes in *A. obliqua* female reproductive stages. Ao_IF = immature female; Ao_VF = virgin female; Ao_PF3 = 3 h post-mating female; Ao_PF6 = 6 h post-mating female; Ao_PF9 = 9 h post-mating female; Ao_PF12 = 12 h post-mating female; Ao_PF24 = 24 h post-mating female; Ao_PF48 = 48 h post-mating female.

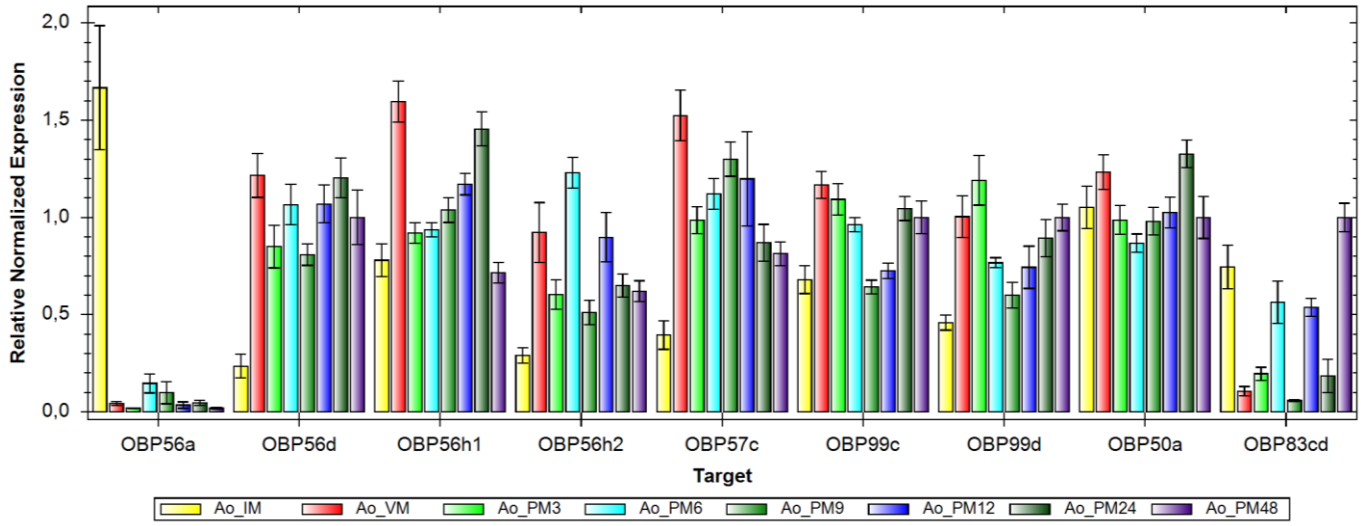


Figure 5. Relative normalized expression of OBP genes in *A. obliqua* male reproductive stages. Ao_IM = immature male; Ao_VM = virgin male; Ao_PM3 = 3 h post-mating male; Ao_PM6 = 6 h post-mating male; Ao_PFM9 = 9 h post-mating male; Ao_PM12 = 12 h post-mating male; Ao_PM24 = 24 h post-mating male; Ao_PM48 = 48 h post-mating male.

Before maturation, immature females tend to search for protein sources and plant hosts for foraging and eventual oviposition, shifting from a post-eclosion developmental phase during which they reach reproductive maturity to one in which they prepare for maximum reproduction (MACK et al., 2006). They need to get carbohydrates and proteins for their own development and for that of their eggs. Likewise, immature males spend a large proportion of their time feeding to accumulate energy upon on which their reproductive success will depend (DREW; YUVAL, 2000). Based on these needs, we hypothesize that genes up-regulated in immature flies are probably involved with the location of food resources during foraging activity, instead of location of mate partners. In *A. obliqua*, we noticed that *OBP56a* was more expressed in immature males and females, showing an opposite expression pattern to other OBPs studied in this species. In corroboration of this supposition, expression patterns of *OBP56a* did not show changes in post-mating profiles of this species. On the other hand, *A. fraterculus OBP56a* did not show significant differential expression between immature and mature virgin flies (although the high standard deviation in immature males could be affecting our inference). Therefore, the *OBP56a* might not have a similar function in *A. fraterculus*. Studies in *Drosophila* showed that this OBP has been associated with mating and copulation latency (MACKAY et al., 2005; MCGRAW et al., 2004). Another gene potentially involved with the location of food resources may be *OBP83cd*, at least in males. *OBP56d* and *OBP99c* were also up-regulated in immature males, but only in *A. fraterculus*.

Sexual maturation and mating in insects are generally accompanied by several physiological and behavioral changes, many of them related to the need to locate a mate and subsequently, in the case of females, to switch from mate searching to oviposition behavior (GOMULSKI et al., 2012). In the tephritid *Ceratitis capitata*, studies revealed

that sexually mature virgin females are attracted to the odor of male-produced pheromone, while mated females are more interested in finding suitable oviposition sites and are therefore more attracted to the odor of host fruits (JANG et al., 1994; JANG, 1995; LANDOLT; HEATH; CHAMBERS, 1992). Considering these findings, we supposed that OBP genes up-regulated in mature virgins, in comparison with immature flies, may be involved in pheromone perception. *A. fraterculus* and *A. obliqua* share three genes up-regulated in mature virgins, *OBP56h-2*, *OBP99d* and *OBP56d*, although the latter only in *A. fraterculus* females, not in males. Other OBP genes previously mentioned as significantly up-regulated in virgin flies showed particular patterns of expression, changing according to sex and/or species considered. Differences in OBP gene expression between males and females or between species are important because they could result in different olfactory perception. For instance, our results suggest that *A. fraterculus* and *A. obliqua* immature males may respond differently to the volatiles carried by OBP56d and OBP99c, as well as virgin males may respond differently to the volatiles carried by OBP56d.

Our comparisons between mature virgins and post-mating profiles revealed that the expression levels of OBPs changed significantly after mating. In *A. fraterculus* females, all OBPs were down-regulated post-mating, except *OBP56a*, up-regulated 3 h post-mating. In *A. obliqua* females we found the opposite, all OBPs were up-regulated post-mating, which is the same pattern observed in *D. melanogaster* females (ZHOU et al., 2009), except *OBP56a*, that did not show significant differences between virgin and post-mating profiles. Males of both species showed *OBP56h-1* down-regulated post-mating, and *A. obliqua* also showed three more down-regulated genes, *OBP57c*, *OBP99c* and *OBP99d*, this latter up-regulated in *A. fraterculus* post-mating males, as well as *OBP56h-2* and *OBP83cd*.

Sex-specific differences found in OBPs gene expression were expected between virgins and post-mating profiles, since each sex has different behavioral and physiological attributes to maximize their reproductive potentials. After mating, females become less attracted to males (probably due a lower expression of pheromones stimulated by seminal fluid proteins transferred from the males), tending to become unreceptive to further mates and look for oviposition sites (AVILA et al., 2011; DALTON et al., 2010; FERVEUR, 1997; GOMULSKI et al., 2012). On the other hand, mature males are actively involved in lek formation, in which the recognition of male pheromones is extremely important for male aggregation and for female attraction (LANDOLT; HEATH; CHAMBERS, 1992; LEAL et al., 2005). After mating, leks are no longer needed and pheromone production tend to decrease (LIGHT et al., 1999). A trancriptome analysis in *C. capitata* showed that sexual maturation induces profound expression changes in females and modest variations in the male. By contrast, post-mating changes were modest in females, and consistent both during maturation and post-mating in males (GOMULSKI et al., 2012), a pattern that was not repeated for OBP genes in *Anastrepha*.

Some of OBP genes described here, *OBP56a*, *OBP56d*, *OBP56h*, *OBP57c*, *OBP99c* and *OBP99d* were also found differentially expressed at different reproductive stages of *D. melanogaster* (ZHOU et al., 2009). Similarly, in the tephritid *Bactrocera dorsalis*, OBPs differentially expressed are potentially orthologous, or at least phylogenetically related with differentially expressed *Anastrepha* OBPs. For instance, *BdorOBP2*, *BdorOBP3*, and *BdorOBP4* were grouped in a same branch with *Anastrepha OBP56d, BdorOBP5* with *OBP56h-2* and *BdorOBP10* with *OBP99c* (CAMPANINI; DE BRITO, unpublished data). Although the patterns of expression are not similar, the recurrence of the same OBP genes differentially expressed in different species may indicate their importance in the reproductive process of these fruit flies.

In *D. melanogaster*, *OBP56a*, *OBP56d* and *OBP56h* genes belong to the same gene cluster, while *OBP99c* and *OBP99d* belong to a different one (HEKMAT-SCAFE et al., 2002). As shown in *D. melanogaster*, the expression patterns of OBP genes differ among members of the same gene cluster in both *A. fraterculus* and *A. obliqua*, which corroborates the hypothesis that OBP genes in the same gene cluster can be regulated independently. This independent regulation is supposed to be a necessary requirement for subfunctionalization or neofunctionalization during evolution, when daughter genes of duplication events either allow refinement and/or expansion in perception of the chemical environment, or yet the acquisition of specialized chemosensory functions (ZHOU et al., 2009).

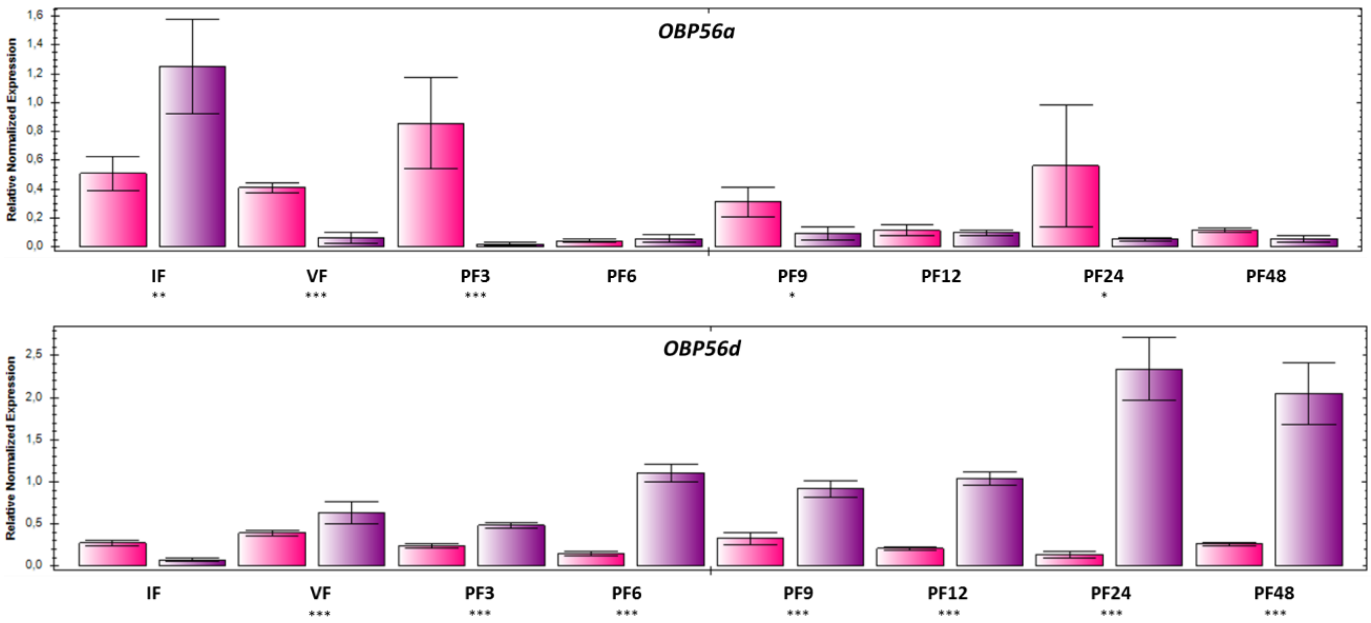
3.4.3. OBPs differential expression between species

Our intraspecific gene expression analysis indicated differences of expression patterns between *A. fraterculus* and *A. obliqua* throughout the reproductive stages. Since differences in levels of expression of chemosensory genes suggest that species experience, interact with, and adapt to their chemical environments differently (VOSSHALL, 2001; ZHOU et al., 2009), we focused on comparisons between these two species for each OBP gene studied. Once again, we performed female-female and male-male comparisons because OBPs has shown sex-biased gene expression in several insects studied so far (ANHOLT et al., 2003; FENG et al., 2015; YANG et al., 2016; ZHANG et al., 2015; ZHANG; PANG; ZHANG, 2015; ZHOU et al., 2009).

Comparing the significantly differential gene expression between females of both species (Figure 6, IF and VF profiles), we noticed that in the reproductive stages of immature and virgin females, most OBPs were significantly more expressed in *A. fraterculus*. The only exception was *OBP56a* in immature flies (profile not analyzed *in silico*), and *OBP56d* in virgins, both more expressed in *A. obliqua*, which is the same pattern previously observed in the *in silico* analysis. In males, however, we did not

observe the same pattern for these reproductive stages (Figure 7, IM and VM profiles), since some OBPs (*OBP56a*, *OBP56h-2* and *OBP57c*) were more expressed in *A. fraterculus*, other (*OBP56h-1*, *OBP99d*, *OBP50a* and *OBP83cd*) in *A. obliqua*, and yet two of them (*OBP56d* and *OBP99c*) showed an antagonistic pattern, more expressed in a species in immature males but in the other species in virgins.

The interspecific comparison among post-mating profiles revealed that *OBP56d*, *OBP57c*, and *OBP50a* showed the highest levels of differentiation between *A. fraterculus* and *A. obliqua* female expression, and *OBP56d*, *OBP56h-1* and *OBP57c* showed the highest levels of differentiation between *A. fraterculus* and *A. obliqua* male expression. We highlight the recurrence of *OBP56d* and *OBP57c* in both male and female comparisons, and the fact that *OBP56h-1* was the gene with the lowest differentiation between these species in post-mating females. *OBP57c* gene expression was shown to increase in females of *D. melanogaster* exposed to odors of other females, which could suggest a function related to female pheromone reception (ZHOU et al., 2009). To the best of our knowledge, no putative functions were associated to *OBP50a*, *OBP56d* and *OBP56h*.



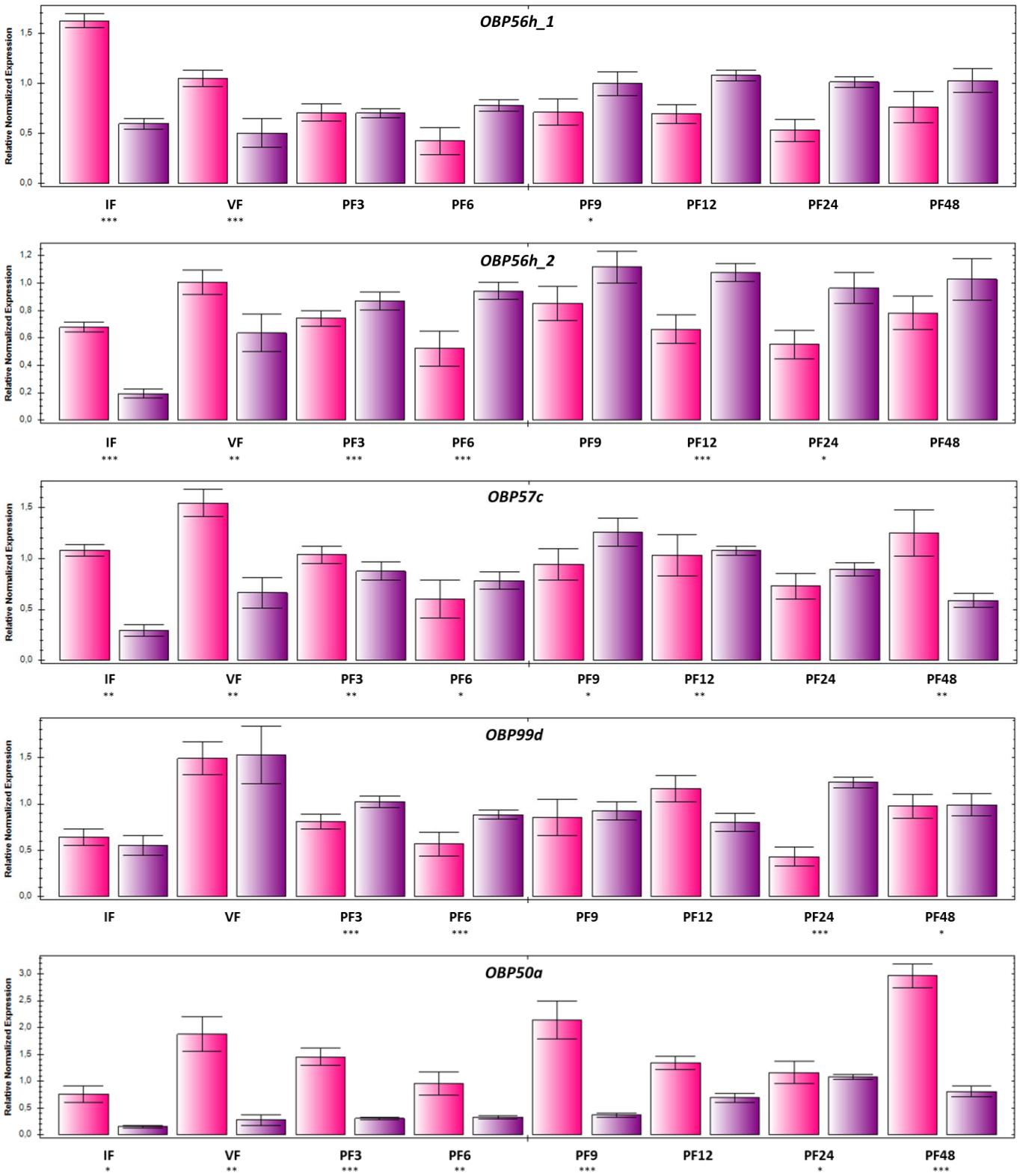
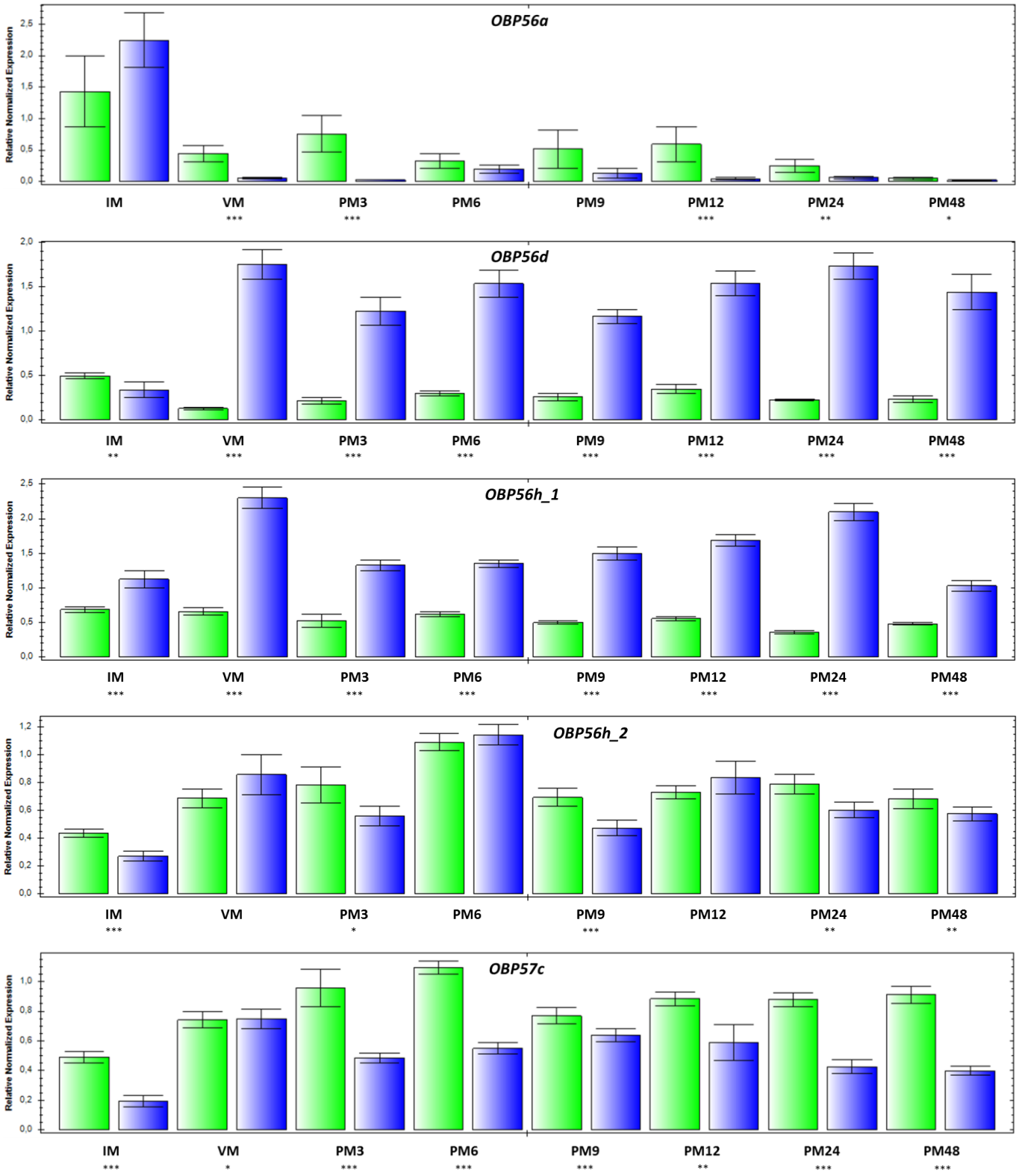


Figure 6. Relative normalized expression of OBP genes in *A. fraterculus* (pink bars) and *A. obliqua* (purple bars) female reproductive stages. IF = immature female; VF = virgin female; PF3 = 3 h post-mating female; PF6 = 6 h post-mating female; PF9 = 9 h post-mating female; PF12 = 12 h post-mating female; PF24 = 24 h post-mating female; PF48

= 48 h post-mating female. Unpaired t-test**:** * p-value > 0.05; ** p-value > 0.01; *** p-value > 0.001.



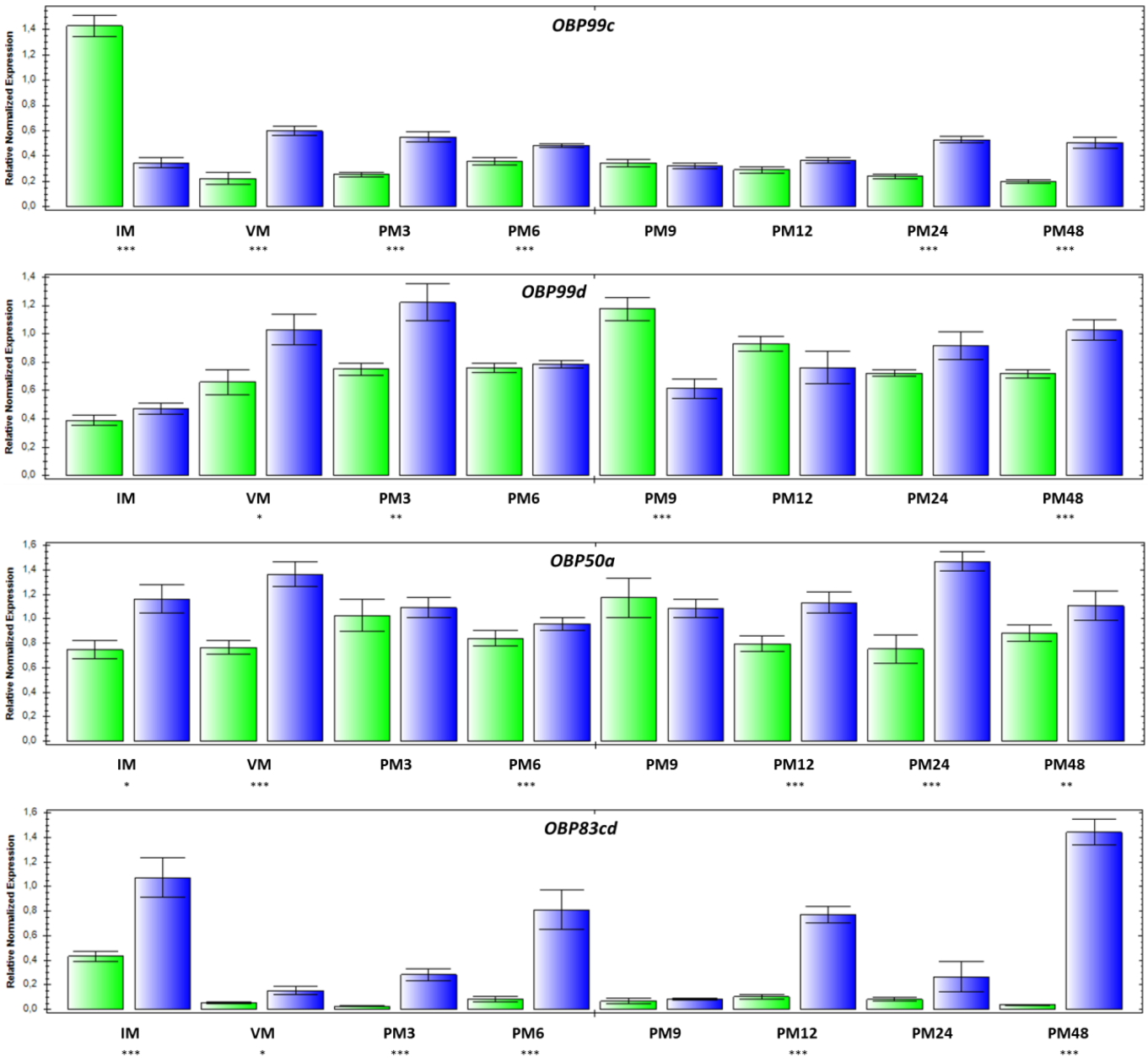


Figure 7. Relative normalized expression of OBP genes in *A. fraterculus* (green bars) and *A. obliqua* (blue bars) male reproductive stages. IF = immature male; VF = virgin male; PF3 = 3 h post-mating male; PF6 = 6 h post-mating male; PF9 = 9 h post-mating male; PF12 = 12 h post-mating male; PF24 = 24 h post-mating male; PF48 = 48 h post-mating male. Unpaired t-test**: *** p-value > 0.05; ** p-value > 0.01; *** p-value > 0.001.

Although our qPCR results cannot exactly corroborate the *in silico* results, because the profiles analyzed are not the same, the comparisons of each OBP expression between *A. fraterculus* and *A. obliqua* were concordant in both analyses. In summary, all nine OBP

genes investigated by qPCR showed high levels of differential expression between *A. fraterculus* and *A. obliqua*, but *OBP56a*, *OBP56d*, *OBP57c* and the two copies of *OBP56h* may be the main OBPs involved in the divergence of these species. *OBP56d* has two copies described for both species that may represent intraspecific variation (CAMPANINI; DE BRITO, unpublished data), thus we considered these two copies as a unique gene to conduct the *in silico* gene expression analysis. The intraspecific variability previously reported in this gene, associated with the significant differential expression patterns found here, makes it a great candidate to be involved in *Anastrepha* species differentiation. On the other hand, the two copies of *OBP56h* in *Anastrepha* may have arisen via gene duplication that preceded the divergence of *A. fraterculus* and *A. obliqua* (CAMPANINI; DE BRITO, unpublished data). The presence of two paralogous copies of a gene could have favored functional diversification, and promoted different expression patterns (ARYA et al., 2010). We clearly observed this differentiation in the interspecific contrast between males, in which *OBP56h-1* and *OBP56h-2* showed very distinct expression patterns, which reinforce that these genes are paralogs.

Differential gene expression in OBP genes is an important factor to be investigated in the study of species divergence, since changes in the amount of specific OBPs available in the sensillas could induce different olfactory responses, which is important in the adaptation of different species to different ecological niches (YANG et al., 2016). In *D. melanogaster*, for instance, a male-specific pheromone, cVA, is involved with courtship inhibition when detected in high concentrations (KURTOVIC; WIDMER; DICKSON, 2007) and with male-male aggression when detected in low concentrations (WANG; ANDERSON, 2010). Considering that OBPs are involved in species-specific behaviors related to host and mate location, differences in OBP gene expression between *A. fraterculus* and *A. obliqua* could explain their two main ecological differences, the

preference for different host fruits and the avoidance of mate partners that are not from their own species. Although we cannot assume that the OBPs here analyzed are in fact responsible for such ecological differences, we found significant differences in their expression between *A. fraterculus* and *A. obliqua*, indicating that these genes are differentially regulated in these species, which may result in different olfactory responses. The nine OBP genes here described under differential selection may be used in further investigations in other *fraterculus* group species to increase the knowledge about the group' speciation process, as well as be considered targets for species-specific control of these pest fruits by RNAi technology.

3.5. Conclusions

The present study identified nine OBP genes differentially expressed between *A. fraterculus* and *A. obliqua*, four of them previously described to be under positive selection in these species. Intraspecific expression analyses at different reproductive stages showed four OBP genes, *OBP56a, OBP56d, OBP99c* and *OBP83cd*, up-regulated in immature flies, and three genes, *OBP56h-2*, *OBP99d* and *OBP56d*, up-regulated in virgin flies in these species. Considering when these genes are expressed, it is possible that the former may be more involved in the search for food, whereas the latter may be more associated with mating. Furthermore, we found that OBP expression levels changed significantly after mating in all sexes for both species. Interspecific comparisons showed high levels of differential expression for OBP genes between *A. fraterculus* and *A. obliqua*, especially for *OBP56a*, *OBP56d*, *OBP57c* and the two copies of *OBP56h*, which suggests that these genes may have an important role in the evolution of differences between these species.

3.6. References

ALTSCHUL, S. et al. Basic Local Alignment Search Tool. **Journal of Molecular Evolution**, v. 215, p. 403–410, 1990.

ANHOLT, R. R. H. et al. The genetic architecture of odor-guided behavior in *Drosophila*: epistasis and the transcriptome. **Nature Genetics**, v. 35, n. 2, p. 180–184, 2003.

ARYA, G. H. et al. Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*. **Genetics**, v. 186, n. 4, p. 1475–1485, 2010.

AVILA, F. W. et al. Insect seminal fluid proteins: identification and function. **Annual Review of Entomology**, v. 56, p. 21–40, 2011.

AYROLES, J. F. et al. Systems genetics of complex traits in *Drosophila melanogaster*. **Nature Genetics**, v. 41, n. 3, p. 299–307, 2009.

BENTON, R. On the ORigin of smell: odorant receptors in insects. **Cellular and Molecular Life Sciences**, v. 63, n. 14, p. 1579–1585, 2006.

BŘÍZOVÁ, R. et al. Pheromone analyses of the *Anastrepha fraterculus* (Diptera: Tephritidae) cryptic species complex. **Florida Entomologist**, v. 96, n. 3, p. 1107–1115, 2013.

BUSTIN, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. **Clinical Chemistry**, v. 55, n. 4, p. 611–622, 2009.

CAMARGO, C. A.; ODELL, E.; JIRÓN, L. F. Interspecific interactions and host preference of *Anastrepha obliqua* and *Ceratitis capitata* (Diptera: Tephritidae), two pests of mango in Central America. **Florida Entomologist**, v. 79, n. 2, p. 266–268, 1996.

CHOMCZYNSKI, P.; MACKEY, K. Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. **BioTechniques**, v. 19, n. 6, p. 942–945, 1995.

DALTON, J. E. et al. Dynamic, mating-induced gene expression changes in female head and brain tissues of *Drosophila melanogaster*. **BMC Genomics**, v. 11, n. 1, p. 541, 2010.

DEKKER, T. et al. Olfactory shifts parallel superspecialism for toxic fruit in *Drosophila melanogaster* sibling, *D. sechellia*. **Current biology: CB**, v. 16, n. 1, p. 101–109, 2006.

DOS SANTOS, P.; URAMOTO, K.; MATIOLI, S. R. Experimental hybridization among *Anastrepha* species (Diptera:Tephritidae): Production and morphological characterization of hybrids. **Annals of the Entomological Society of America**, v. 94, n. 5, p. 717–725, 2001.

DREW, R.; YUVAL, B. The evolution of fruit fly feeding behaviour. In: **Fruit flies (Tephritidae): phylogeny and evolution of behavior**. Boca RatonFlorida, , USA: CRC Press: Aluja M, Norrbom AL, eds., 2000. p. 731–749.

FAN, J. et al. An overview of odorant-binding protein functions in insect peripheral olfactory reception. **Genetics and molecular research: GMR**, v. 10, n. 4, p. 3056–3069, 2011.

FENG, B. et al. Transcriptome and expression profiling analysis link patterns of gene expression to antennal responses in Spodoptera litura. **BMC Genomics**, v. 16, n. 1, 2015.

FERVEUR, J.-F. The pheromonal role of cuticular hydrocarbons in *Drosophila melanogaster*. **BioEssays**, v. 19, n. 4, p. 353–358, 1997.

FORÊT, S.; MALESZKA, R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). **Genome research**, v. 16, n. 11, p. 1404–1413, 2006.

GIBSON, G. Microarrays in ecology and evolution: a preview. **Molecular Ecology**, v. 11, n. 1, p. 17–24, 2002.

GOMULSKI, L. M. et al. Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitis capitata*. **PLoS ONE**, v. 7, n. 1, p. e30857, 2012.

GONG, D.-P. et al. The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*. **BMC Genomics**, v. 10, n. 1, p. 332, 2009.

GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v. 29, n. 7, p. 644–652, 2011.

HAAS, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. **Nature Protocols**, v. 8, n. 8, p. 1494–1512, 2013.

HEKMAT-SCAFE, D. S. et al. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. **Genome research**, v. 12, n. 9, p. 1357–1369, 2002.

JANG, E. B. et al. Attraction of female Mediterranean fruit flies to the five major components of male-produced pheromone in a laboratory flight tunnel. **Journal of Chemical Ecology**, v. 20, n. 1, p. 9–20, 1994.

JANG, E. B. Effects of mating and accessory gland injections on olfactory-mediated behavior in the female Mediterranean fruit fly, *Ceratitis capitata*. **Journal of Insect Physiology**, v. 41, n. 8, p. 705–710, ago. 1995.

KOPP, A. et al. Evolution of gene expression in the *Drosophila* olfactory system. **Molecular Biology and Evolution**, v. 25, n. 6, p. 1081–1092, 2008.

KURTOVIC, A.; WIDMER, A.; DICKSON, B. J. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. **Nature**, v. 446, n. 7135, p. 542–546, 2007.

LANDOLT, P. J.; HEATH, R. R.; CHAMBERS, D. L. Oriented flight responses of female Mediterranean fruit flies to calling males, odor of calling males, and a synthetic

pheromone blend. **Entomologia Experimentalis et Applicata**, v. 65, n. 3, p. 259–266, 1992.

LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357–359, 2012.

LAUGHLIN, J. D. et al. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. **Cell**, v. 133, n. 7, p. 1255–1265, 2008.

LEAL, W. S. et al. Kinetics and molecular properties of pheromone binding and release. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 15, p. 5386–5391, 2005.

LI, B.; DEWEY, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 1, p. 323, 2011.

LIGHT, D. M. et al. Minor and intermediate components enhance attraction of female Mediterranean fruit flies to natural male odor pheromone and its synthetic major components. **Journal of Chemical Ecology**, 1999.

LI, H. et al. Identification of differential expression genes associated with host selection and adaptation between two sibling insect species by transcriptional profile analysis. **BMC Genomics**, v. 14, n. 1, p. 582, 2013.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. **Methods (San Diego, Calif.)**, v. 25, n. 4, p. 402–408, 2001.

LÓPEZ-GUILLÉN, G. et al. Olfactory responses of *Anastrepha obliqua* (Diptera: Tephritidae) to volatiles emitted by calling males. **Florida Entomologist**, v. 94, n. 4, p. 874–881, 2011.

MACKAY, T. F. C. et al. Genetics and genomics of *Drosophila* mating behavior. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102 Suppl 1, p. 6622–6629, 2005.

MACK, P. D. et al. Mating-responsive genes in reproductive tissues of female *Drosophila melanogaster*. **Proceedings of the National Academy of Sciences**, v. 103, n. 27, p. 10358–10363, 2006.

MATSUO, T. et al. Odorant-binding proteins *OBP57d* and *OBP57e* affect taste perception and host-plant preference in *Drosophila sechellia*. **PLoS Biology**, v. 5, n. 5, p. e118, 2007.

MCBRIDE, C. S. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 12, p. 4996–5001, 2007.

MCGRAW, L. A. et al. Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*. **Current Biology**, v. 14, n. 16, p. 1509–1514, 2004.

MILNE, I. et al. Using Tablet for visual exploration of second-generation sequencing data. **Briefings in Bioinformatics**, v. 14, n. 2, p. 193–202, 2013.

NAKAMURA, A. M. et al. Reference genes for accessing differential expression among developmental stages and analysis of differential expression of OBP genes in *Anastrepha obliqua*. **Scientific Reports**, v. 6, p. 17480, 2016.

PELOSI, P. et al. Soluble proteins in insect chemical communication. **Cellular and molecular life sciences: CMLS**, v. 63, n. 14, p. 1658–1676, 2006.

PELOSI, P.; MAIDA, R. Odorant-binding proteins in insects. **Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology**, v. 111, n. 3, p. 503–514, 1995.

PEREIRA-RÊGO, D. R. G. et al. Morfometria de *Anastrepha fraterculus* (Wied) (Diptera: Tephritidae) relacionada a hospedeiros nativos, Myrtaceae. **Arquivos do Instituto Biológico**, v. 78, p. 3743, 2011.

PERRE, P. et al. Morphometric differentiation of fruit fly pest species of the *Anastrepha fraterculus* group (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 107, n. 2, p. 490–495, 2014.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 2010.

SÁNCHEZ-GRACIA, A.; ROZAS, J. Divergent evolution and molecular adaptation in the *Drosophila* odorant-binding protein family: inferences from sequence variation at the OS-E and OS-F genes. **BMC Evolutionary Biology**, v. 8, n. 1, p. 323, 2008.

SCHWARZ, D. et al. Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella*. **BMC Genomics**, v. 10, n. 1, p. 633, 2009.

SCOTT, C. P.; WILLIAMS, D. A.; CRAWFORD, D. L. The effect of genetic and environmental variation on metabolic gene expression. **Molecular Ecology**, v. 18, n. 13, p. 2832–2843, 2009.

SEGURA, D. F. et al. Lekking Behavior of *Anastrepha fraterculus* (Diptera: Tephritidae). **The Florida Entomologist**, v. 90, n. 1, p. 154–162, 2007.

SELIVON, D.; PERONDINI, A. L. P.; MORGANTE, J. S. A genetic–morphological characterization of two cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 98, n. 3, p. 367–381, 2005.

STEINBRECHT, R. A. Are Odorant-binding proteins involved in odorant discrimination? **Chemical Senses**, v. 21, n. 6, p. 719–727, 1996.

TARAZONA, S. et al. Differential expression in RNA-seq: a matter of depth. **Genome Research**, v. 21, n. 12, p. 2213–2223, 2011.

VIEIRA, F. G.; SÁNCHEZ-GRACIA, A.; ROZAS, J. Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. **Genome Biology**, v. 8, n. 11, p. R235, 2007.

VOGT, R. G. Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). **Chemical Senses**, v. 24, n. 5, p. 481–495, 1999.

VOGT, R. G. et al. A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). **The Journal of Experimental Biology**, v. 205, n. Pt 6, p. 719–744, 2002.

VOGT, R. G.; RIDDIFORD, L. M. Pheromone binding and inactivation by moth antennae. **Nature**, v. 293, n. 5828, p. 161–163, 1981.

VOSSHALL, L. B. How the brain sees smells. **Developmental Cell**, v. 1, n. 5, p. 588–590, 2001.

WANG, L.; ANDERSON, D. J. Identification of an aggression-promoting pheromone and its receptor neurons in *Drosophila*. **Nature**, v. 463, n. 7278, p. 227–231, 2010.

WHITEHEAD, A.; CRAWFORD, D. L. Variation within and among species in gene expression: raw material for evolution. **Molecular Ecology**, v. 15, n. 5, p. 1197–1211, 2006.

YANG, B. et al. Sexually biased expression of odorant-binding proteins and chemosensory proteins in Asian corn borer *Ostrinia furnacalis* (Lepidoptera: Crambidae). **Applied Entomology and Zoology**, 2016.

ZHANG, S.; PANG, B.; ZHANG, L. Novel odorant-binding proteins and their expression patterns in grasshopper, *Oedaleus asiaticus*. **Biochemical and Biophysical Research Communications**, v. 460, n. 2, p. 274–280, 2015.

ZHANG, T. et al. Male- and female-biased gene expression of olfactory-related genes in the antennae of asian corn borer, *Ostrinia furnacalis* (Guenée) (Lepidoptera: Crambidae). **PLOS ONE**, v. 10, n. 6, p. e0128550, 2015.

ZHOU, S. et al. Plasticity of the chemoreceptor repertoire in *Drosophila melanogaster*. **PLoS Genetics**, v. 5, n. 10, p. e1000681, 2009.

ZHU, J.-Y. et al. Identification and tissue distribution of odorant binding protein genes in the beet armyworm, *Spodoptera exigua*. **Journal of Insect Physiology**, v. 59, n. 7, p. 722–728, 2013.

Additional Files

Additional File 1. *A. obliqua* OBP sequences used as reference to the *in silico* gene expression analyses.

Gene	Accession number
AoblOBP8a	KU317933
AoblOBP19a	KU317934
AoblOBP19b	KU317935
AoblOBP19c	KU317936
AoblOBP19d	KU317937
AoblOBP47b	KU317938
AoblOBP49a-1	KU317939
AoblOBP49a-2	KU317940
AoblOBP50a	KU317941
AoblOBP50e	KU317942
AoblOBP56a	KX018812
AoblOBP56d-1	KU317943
AoblOBP56h-1	KU317945
AoblOBP56h-2	KU317946
AoblOBP57c	KU317947
AoblOBP59a	KU317948
AoblOBP83cd	KU317949
AoblOBP83ef	KU317950
AoblOBP83g	KU317951
AoblOBP99a	KU317952
AoblOBP99b	KU317953
AoblOBP99c	KU317954
AoblOBP99d-1	KU317955

Additional File 2. Primer pairs selected for qPCR analyses and results for efficiency tests.

Gene	Positive selection*	*In silico* differential expression	Sequences accession number	Primers	Sequences	E (%)	R^2
OBP50a	yes	no	Af = KU317977 Ao = KU317941	Foward Reverse	CATAGCATCGACGAAAACGG CCGCCTTCTCTAATTGCTCG	97.5	0,997
OBP56a	no	females e males	Af = KP939314 Ao = KX018812	Foward Reverse	GAGGCCAAATTCGATGTGTT ACACTTTTCCCGATCGTTTG	101.6	0.990
OBP56d	no	females e males	Af = KU317979 Ao = KU317943	Foward Reverse	AAGTCGTTAGGTCCCTTGG TAGCACTTGTACCTCTCGAG	98.1	0.998
OBP56h-1	yes	no	Af = KU317965 Ao = KU317945	Foward Reverse	TGTCACCGTTCTAATGTGTCA ATTATCCTTGGCATCGCTGG	98.5	0.991
OBP56h-2	yes	females e males	Af = KU317966 Ao = KU317946	Foward Reverse	CAATGCGCCATGGTCAAAC CGCTAGCCAACTTTGCATC	102.6	0.995
OBP57c	yes	no	Af = KU317967 Ao = KU317947	Foward Reverse	GGCTTGTCTAGAGCAGCACA TTCTCGTCCATGAAGCCCAG	99.1	0.991
OBP83cd	no	males	Af = KU317969 Ao = KU317949	Foward Reverse	GTGCGTTACAAGGAGTGGTC TCGCAATTATCCGCCAATGG	101	0.973
OBP99c	no	males	Af = KU317974 Ao = KU317954	Foward Reverse	TGAAGCTGAGGTATTGGCCAT AGACACTTGTGACCGCGATA	99.4	0.998
OBP99d	no	females e males	Af = KU317975 Ao = KU317954	Foward Reverse	TACCCGGACGAGGAGATAGT TCGCTCAATATCGGCACCAA	99	0.997

*according to CAMPANINI & DE BRITO, unpublished data. Af = *A. fraterculus*; Ao = *A. obliqua*; E = efficiency; R^2 = standard curve correlation coefficient.

CONSIDERAÇÕES FINAIS

Nossa investigação acerca da evolução molecular da família gênica das OBPs em *A. fraterculus* e *A. obliqua* mostrou um número similar de membros entre essas espécies, e também que os ortólogos entre elas apresentam grande similaridade em suas sequências. Contudo, mesmo que as diferenças observadas não sejam muitas, algumas delas são potencialmente drásticas e estes dados em geral mostram que essa família gênica tem evoluído mais rápido que o restante dos genes dessas espécies. Em função disso, encontramos quatro membros que estão evoluindo sob seleção positiva entre estas espécies, o que significa que existe uma pressão de seleção que favorece a manutenção de substituições de aminoácidos. Encontramos também algumas destas mudanças em sítios que alteram propriedades físico-químicas ou ainda que alteram a cavidade ligante das OBPs e são ainda mais importantes porque podem estar promovendo divergência funcional entre elas. Investigamos o padrão de expressão destes genes, bem como de outros genes que identificamos *in silico* como apresentando expressão diferencial entre estas espécies, em diferentes fases do desenvolvimento e da reprodução destas moscas e detectamos que a expressão gênica diferencial ocorre para pelo menos nove genes dessa família gênica entre *A. fraterculus* e *A. obliqua*. Observamos também que alguns destes genes são mais expressos em indivíduos imaturos enquanto outros em indivíduos maduros sexualmente, o que pode indicar diferentes papéis dessas OBPs na sobrevivência e reprodução dessas espécies. Concluímos que o presente trabalho apresenta um conjunto de OBPs que estão potencialmente envolvidas na diferenciação entre *A. fraterculus* e *A. obliqua*, que poderão agora ser investigadas em outras espécies do grupo *fraterculus* para a compreensão do processo de especiação ocorrido nesse grupo.