



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



Beatriz Faustino Lima Mendonça

**EFEITO DE HOSPEDEIROS NA VARIABILIDADE DE CARACTERES
MORFOLÓGICOS E MARCADORES EPIGENÉTICOS EM MOSCAS-DAS-
FRUTAS *Anastrepha fraterculus* (DIPTERA; TEPHRITIDAE)**

São Carlos - SP

2023

BEATRIZ FAUSTINO LIMA MENDONÇA

Efeito de hospedeiros na variabilidade de caracteres morfológicos e marcadores epigenéticos em Moscas-das-frutas *Anastrepha fraterculus* (DIPTERA; TEPHRITIDAE)

Dissertação apresentada ao Programa de Pós-graduação em Genética Evolutiva e Biologia Molecular da Universidade Federal de São Carlos, *campus* São Carlos, como parte dos requisitos para a obtenção do título de Mestra em Genética Evolutiva e Biologia Molecular.

Orientador: Prof. Dr. Reinaldo Alves de Brito

São Carlos - SP

2023



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

Defesa de Dissertação de Mestrado da candidata Beatriz Faustino Lima Mendonça, realizada em 28/12/2023.

Comissão Julgadora:

Prof. Dr. Reinaldo Otavio Alvarenga Alves de Brito (UFSCar)

Prof. Dr. Francis de Moraes Franco Nunes (UFSCar)

Prof. Dr. Klaus Hartmann Hartfelder (USP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

*À toda minha família, por sempre
dedicarem suas vidas em prol
daqueles que amam.*

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus pelo dom da vida, e por sempre me mostrar que não importa a dificuldade enfrentada eu nunca estou só, Ele sempre está comigo!

A todos os professores que auxiliaram em minha formação acadêmica e como pesquisadora até hoje, em especial, agradeço ao Prof. Dr. Reinaldo A. De Brito pela orientação, acolhida e por ter aberto as portas do Laboratório de Genética de Populações e Evolução para o desenvolvimento da minha dissertação de mestrado, sou muito grata por todas as conversas, discussões, conselhos, e pela amizade construída. Foram com certeza dois anos de muita aprendizagem e crescimento pessoal e profissional, que levarei sempre comigo.

Aos meus pais, Jânio Cesar Vieira Lima e Maclene Faustino Lima, por sempre me incentivarem a estudar e buscar o melhor para minha carreira profissional, mesmo que para isso envolva ficar quase 3 mil quilômetros de distância e a saudade ser constante. Obrigada por entenderem e me apoiarem em tudo, os amos muito! A minha irmã Danielly, por me lembrar sempre o lado bom da vida e por todo companheirismo, e força que me deu até aqui. E a toda a minha família, em especial meus avós Pedro e Zilda por serem meus exemplos e por compartilharem o melhor de vocês.

Ao meu esposo, Alan Mendonça, por tudo que abriu mão para que eu viesse desenvolver este trabalho, por todo o seu amor, por ser o meu apoio nos momentos mais difíceis e por constantemente me lembrar do quanto eu sou capaz. Sem o seu incentivo eu não teria chegado até aqui. Você e a Kirinha são sempre motivos de maior alegria ao chegar em casa.

A todos os meus companheiros de laboratório pelo ambiente acolhedor, vocês fizeram com que estes dois anos fossem melhores, e aos amigos que o DGE e os horários de almoço me proporcionaram, Bruno, Mariana, Natalia, Maria Isabel, Samira, Bruna, Angelica, Lindomar, Victoria e Matheus, obrigada por estarem comigo em todos os momentos, pelas conversas e principalmente pela amizade. Vocês estarão marcados para sempre em minha vida.

A minha amiga e sempre orientadora Prof. Dr. Rubiani de Cássia Pagotto por sempre me incentivar na pesquisa. E a todos os meus amigos, em especial, Thandy por tudo que já fez e faz por mim, sua amizade é um grande presente, Taíza Vicentini e Joice Mara por todo o apoio, força e oração mesmo estando longe. Aos meus amigos Gabi e

Luis por se tornarem família e por todo o apoio que deram a mim e minha família em São Carlos. A família Muffato por todo o acolhimento, por todos os almoços, e também por se tornarem nossa família em São Carlos. Todos vocês fizeram com que eu me sentisse mais pertinho de casa e foram essenciais nessa jornada.

Gostaria também de agradecer ao Centro de Ciências Biológicas e da Saúde, Departamento de Genética e Evolução da Universidade Federal de São Carlos (UFSCar - campus São Carlos) e ao Programa de Genética Evolutiva e Biologia Molecular (PPGGEv).

Agradeço também as agências de fomento CAPES e CNPq, e pelo financiamento da pesquisa de mestrado recebida por meio do processo nº 2021/01364-2, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

RESUMO

No gênero *Anastrepha* (Diptera: Tephritidae) de moscas-das-frutas, o grupo *fraterculus* destaca-se por ter espécies altamente generalistas e pragas especializadas em diferentes frutos, sendo um dos mais importantes economicamente para a agricultura nacional e mundial. De modo geral, tais espécies são diferenciadas por meio de caracteres morfológicos, porém os limites dessa diferenciação são difíceis de serem determinados para alguns taxa, especialmente no grupo de espécies *fraterculus*, que apresenta inclusive espécies crípticas. Isso acontece em alguns casos pela separação recente entre espécies, mas também pela sobreposição das medidas dos caracteres. Esta variação fenotípica pode ocorrer, inclusive, pela atuação de fatores epigenéticos, uma vez que essas mudanças podem estar relacionadas com estímulos ambientais que desencadeiam respostas rápidas. Este trabalho teve por objetivo investigar o efeito de frutos hospedeiros em populações de *Anastrepha fraterculus*, visando identificar variações fenotípicas em diversos caracteres morfológicos presentes na cabeça, tórax, acúleo e asa, bem como investigar a presença de marcas epigenéticas, como metilação de DNA, entre espécimes que emergiram de diferentes frutos. Para este fim, foram utilizadas moscas derivadas de populações coletadas de cinco diferentes hospedeiros (cereja do rio grande, pitanga, uvaia, goiaba e guabiroba) no entorno da Universidade Federal de São Carlos – campus São Carlos. As análises morfométricas foram realizadas em 117 espécimes provenientes dos 5 frutos coletados, sendo analisados 18 caracteres morfológicos pelo método de estatística multivariada, como análise de componentes principais (PCA), discriminante (DFA) e distância de Mahalanobis (MD). As análises epigenéticas foram investigadas em um subconjunto dessas moscas, utilizando dados obtidos por meio de sequenciamento em um equipamento MinION da *Oxford Nanopore Technologies* de dois pools de cabeças por fruto, que permitem o estudo de metilação de DNA 5mC e 6mA. A chamada de metilação em sítios 6mA entre moscas que emergiram de diferentes frutos foi feita por meio do software mCaller e sua variação entre amostras (DMRs) investigada por uma análise não paramétrica. Por sua vez, a análise de regiões 5mC diferencialmente metiladas (DMRs) foi realizada por meio do pipeline pycoMeth, cujos resultados foram investigados para enriquecimento de categorias ontológicas de genes (GO). Os resultados de morfometria indicaram variabilidade fenotípica em caracteres morfológicos de moscas provenientes de cinco diferentes frutos suficiente para separar a grande maioria das amostras emergidas de frutos distintos. As análises epigenéticas permitiram registrar pela primeira vez a presença

de metilação de DNA em moscas-das-frutas *A. fraterculus* tanto em citosinas (5mC), como em adeninas (6mA). A Investigação sobre a distribuição dessas DMRs sugeriu genes importantes no processo de diferenciação dessa espécie a frutos distintos, bem como diversas classes enriquecidas de genes responsáveis por diferentes funções nos processos biológicos, inclusive na morfogênese de estruturas anatômicas e no desenvolvimento do sistema nervoso. Estes resultados fornecem uma base para futuras investigações sobre modificações epigenéticas e variabilidade fenotípica em moscas-das-frutas, destacando sua relevância na regulação de fenótipos, influência e escolha de hospedeiros e adaptação evolutiva. A compreensão desses padrões de adaptação é crucial não apenas para a taxonomia, mas também para elucidar os processos evolutivos, e as dinâmicas ecológicas, envolvidas nas interações entre as moscas e seus hospedeiros.

ABSTRACT

In the genus *Anastrepha* (Diptera: Tephritidae) of fruit flies, the *fraterculus* group stands out for having highly generalist species specialized in different fruits, being one of the most economically important for national and global agriculture. In general, such species are differentiated through morphological characters, however the limits of this differentiation are difficult to determine for some taxa, especially in the *fraterculus* species group, which presents inclusive cryptic species. This happens in some cases due to the recent separation between species, but also due to overlapping of character measurements. This phenotypic variation may even occur through the action of epigenetic factors, since these changes may be related to environmental stimuli that trigger rapid responses. This work aimed to investigate the effect of hosts fruits on populations of *Anastrepha fraterculus*, identifying phenotypic variations in several morphological characters present in the head, thorax, aculeus and wing, as well as investigating the presence of epigenetic marks, such as DNA methylation, between specimens that emerged from different fruits. For this purpose, flies derived from the ecosystem were used, collected from five different hosts (cherry do rio grande, pitanga, uvaia, guava and guabiroba) around the Federal University of São Carlos – São Carlos campus. Morphometric analyzes were carried out on 117 samples from the 5 fruits found, with 18 morphological characters being analyzed using the multivariate statistical method, such as principal component analysis (PCA), discriminant analysis (DFA) and Mahalanobis distance (MD). Epigenetic analyzes were investigated in a subset of these flies, using data obtained through sequencing on a MinION equipment from Oxford Nanopore Technologies of two pools of heads per fruit, which allows the study of 5mC and 6mA DNA methylation. Methylation calling at 6mA sites between flies that emerged from different fruits was done using the mCaller software and its variation between samples (DMRs) investigated by a non-parametric analysis. In turn, an analysis of differentially methylated 5mC regions (DMRs) was performed using the pycoMeth pipeline, the results of which were investigated for enrichment of gene ontological categories (GO). The morphometric results indicated phenotypic variability in morphological characters of flies from five different fruits enough to separate the vast majority of samples emerging from different fruits. Epigenetic analyzes allowed recording for the first time the presence of DNA methylation in fruit flies *A. fraterculus* at both cytosines (5mC) and adenines (6mA). The investigation into the distribution of these DMRs suggested important genes in the process of differentiating this species into

different fruits, as well as several enriched classes of genes responsible for different functions in biological processes, including the morphogenesis of anatomical structures and the development of the nervous system. These results provide a basis for future investigations into epigenetic modifications and phenotypic variability in fruit flies, highlighting their relevance in the regulation of phenotypes, host influence and choice, and evolutionary adaptation. Understanding these adaptation patterns is crucial not only for taxonomy, but also for elucidating the evolutionary processes, and ecological dynamics, involved in the interactions between flies and their hosts.

Sumário

Lista de Figuras	12
INTRODUÇÃO GERAL	14
Gênero <i>Anastrepha</i>	14
Plasticidade fenotípica	15
Epigenética	17
OBJETIVOS	21
CAPÍTULO 1 – Variação fenotípica de moscas <i>Anastrepha fraterculus</i> que emergiram de diferentes frutos	22
1.1 INTRODUÇÃO	22
1.2 MATERIAL E MÉTODOS	23
1.2.1 Coleta e Criação de Moscas	23
1.2.2 Medidas morfométricas	24
1.2.3 Análises estatísticas	25
1.3 RESULTADOS	26
1.4 DISCUSSÃO	29
CAPÍTULO 2 - Marcas epigenéticas de Moscas-das-Frutas <i>Anastrepha fraterculus</i>	34
2.1 INTRODUÇÃO	34
2.2 MATERIAL E MÉTODOS	36
2.2.1 Extração de DNA	36
2.2.2 Preparo de biblioteca e sequenciamento	38
2.2.3 Processamento de dados	39
2.2.4 Análises epigenéticas	40
2.2.4.1 Metilação 5mC	40
2.2.4.2 Metilação 6mA	42
2.3 RESULTADOS	43
2.3.1 Sequenciamento <i>Nanopore</i> para análises de metilação	43
2.3.2 Metilação de ilhas CpG no DNA e Regiões Diferencialmente Metiladas (DMRs)	45
2.3.3 Metilação 6mA	49
2.4 DISCUSSÃO	49
CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS BIBLIOGRÁFICAS	56
MATERIAL SUPLEMENTAR	67
Tabelas	67
Figuras	77

Lista de Figuras

Figura 1. Variáveis consideradas em medidas morfométricas para as 4 estruturas propostas.	25
Figura 2. PCA para os frutos cereja do rio grande, guabiroba, goiaba, pitanga e uvaia. A2 - Medida da abertura do oviduto; A3 - largura no início de seção serrilhada; A8 - comprimento da ponta do acúleo (A4+A5); A9 - razão do comprimento da região não serrilhada versus o comprimento da seção serrilhada (A4/A5); W1- comprimento da asa W7 - razão entre largura e comprimento da asa (W2/W1); TL – comprimento do tórax; TW – largura do tórax; HEL – comprimento da cabeça; HW- largura da cabeça.	27
Figura 3. Análise de função discriminante separando as amostras que emergiram de hospedeiros como Uvaia, pitanga, cereja do rio grande, goiaba e guabiroba, com elipses de 95% de confiança.	29
Figura 4. Resultado cumulativo do sequenciamento de bibliotecas Oxford Nanopore de DNA genômico em um MinION ao longo do tempo; (a) Biblioteca 1 (b) Biblioteca 2.	44
Figura 5. Contagens de categorias de metilação por amostra para intervalos significativos de CpG inferidos por pycoMeth de resultados de sequenciamento por Oxford Nanopore em um MinION.	46
Figura 6. Razão de log-verossimilhança de metilação por intervalo CpG significativo.	46
Figura 7. Distância do intervalo CpG ao gene TSS mais próximo.	47
Figura 8. Análise de gene ontology (GO) mostrando os 20 principais termos GO por ordem de fold enrichment,	47
Figura 9 Análise de gene ontology (GO) mostrando as relações entre os 20 principais termos GO enriquecidos. Quando duas vias (ou nós) são conectadas é devido compartilharem pelo menos 20% genes, linhas mais grossas significam maiores sobreposições de genes, enquanto nós mais escuros representam conjuntos de genes mais informativos significativamente e nós maiores representam conjuntos de genes maiores....	48
Figura S 1. Decaimento do número de poros por intervalo de tempo. (a) biblioteca 1; (b) biblioteca 2.	77
Figura S 2. Comprimento de leitura vs qualidade de leitura média usando pontos. (a) biblioteca 1 (b) biblioteca 2; e comprimento de leitura log ao longo do tempo. (c) biblioteca 1; (d) biblioteca 2.	77
Figura S 3. Número de reads gerados por canal. (a) biblioteca 1; (b) biblioteca 2.	78
Figura S 4. Distribuição da razão log-verossimilhança de metilação de CpG para intervalos significativos de CpG.	78
Figura S 5. Distribuição de intervalos CpG significativos por intervalos cromossômicos.	78
Figura S 6. (A) Distribuição ao longo do genoma para as 100 regiões mais significativas de (DMRs) para metilação 5mC de moscas que emergiram de diferentes frutos. (B) Distribuição ao longo do genoma de (DMRs) para metilação 6mA de moscas que emergiram de diferentes frutos.	79
Figura S 7. Termos GO relacionados através de uma árvore de cluster hierárquica.	79

LISTA DE TABELA

Tabela 1. Distância de Mahalanobis para os grupos estudados.....	29
Tabela 2. Identificação dos barcodes utilizados para cada amostra sequenciada.....	38
Tabela 3 Dados de qualidade do sequenciamento de DNA genômico por meio de bibliotecas Oxford Nanopore em um MinION.....	44
Tabela 4. Número de regiões com múltiplas bases 6mA metiladas. Número de loci metilados com cobertura maior ou igual a 5X para cada pool de amostra de diferentes frutos.....	49
Tabela S 2. Lambda de Wilk's para as amostras escolhidas para análise.....	67
Tabela S 3. Os 100 principais intervalos diferenciamete metilados significativos.	68
Tabela S 4. Termos e caminhos GO enriquecidos.....	72
Tabela S 5. Sítios com valores significantes de m6A estimados por mCaller e com probabilidade corrigida por FDR para eventos múltiplos.	74

INTRODUÇÃO GERAL

Gênero *Anastrepha*

A família Tephritidae (Diptera) pode ser encontrada em ambientes tropicais e subtropicais e é caracterizada por sua grande diversidade, contando com mais de 4000 espécies descritas, algumas das quais representam grandes ameaças à produção de frutos (BENELLI, 2014). Dentre os grupos pertencentes a esta família, o gênero *Anastrepha* destaca-se por ser o mais diverso nas Américas, com mais de 300 espécies descritas, e por compreender os insetos conhecidos popularmente como moscas-das-frutas (NORRBOM; KORYTKOWSKI, 2009; TROYA; NORRBOM; PINEDA, 2020).

O gênero *Anastrepha* tem como característica geralmente distintiva dos outros Tephritidae o ápice da veia M, fortemente curvado na direção anterior, além de possuírem asas com padrão C, S e V marrons ou amareladas (FOOTE, 1980; LAMPERT et al., 2020). Apesar de sua grande importância, uma vez que abriga algumas das mais relevantes espécies de pragas de frutas tropicais, muitas espécies permanecem não descritas, e suas relações com plantas hospedeiras são pouco conhecidas (NORRBOM; KORYTKOWSKI, 2011). Um melhor conhecimento das espécies deste grupo e suas respectivas preferências de hospedeiros e sazonalidades relacionadas à fenologia da planta hospedeira, é essencial para o entendimento da dinâmica populacional desses insetos (SOUZA-FILHO et al., 2009).

No Brasil são registradas cerca de 128 espécies de *Anastrepha* em todo o país (ZUCCHI E MORAES, 2023), algumas das quais de grande importância econômica por causarem problemas fitossanitários nas áreas de produção de frutos (CONTRERAS-MIRANDA et al., 2023), particularmente espécies do grupo *fraterculus* (RAGA et al., 2011). Esta notoriedade está relacionada a um pequeno número de espécies generalistas que as tornam pragas altamente polífagas que se utilizam de uma ampla variedade de frutos maduros e carnosos. Essa característica pode ter favorecido uma pré-aptidão para infestação de cultivados (ARAUJO; ZUCCHI, 2006; SILVA et al., 2010).

O conjunto de espécies conhecido como grupo *fraterculus* abrange, até o momento, 34 espécies taxonomicamente próximas (NORRBOM et al., 2015, 2021), algumas das quais de difícil distinção, como *Anastrepha turpiniae*, *Anastrepha sororcula* e a mosca das frutas sul-americana *Anastrepha fraterculus* (Wiedmann, 1830), *sensu lato*. Esta última sendo reconhecida como um complexo de espécies compreendendo oito

morfotipos, dos quais, três foram identificados no Brasil (HERNÁNDEZ-ORTIZ et al., 2015; SELIVON et al., 2022; VANÍČKOVÁ et al., 2015).

No entanto, as relações evolutivas e a taxonomia das espécies de *Anastrepha* não estão totalmente esclarecidas (CONGRAINS et al., 2023; DÍAZ et al., 2018; MENGUAL et al., 2017). Geralmente, parte da identificação das espécies deste grupo é baseada em atributos morfológicos complexos, com possível herança multifatorial, como padrões de asas, manchas no sub-escutelo e, principalmente, por meio de diferenças no ovopositor feminino (acúleo) (STONE, 1942; HERNÁNDEZ-ORTIZ; GUILLÉN-AGUILAR; LÓPEZ, 2010). Porém, os limites de diferenciação podem ser difíceis de serem delimitados, visto que, em alguns casos, ocorre a sobreposição nas mensurações destes caracteres taxonômicos, como na medida do ápice do ovipositor feminino, tornando a identificação das espécies do grupo uma tarefa desafiadora (ARAUJO; ZUCCHI, 2006; PERRE et al., 2014).

Além disso, é importante destacar que existe plasticidade destes marcadores morfológicos entre indivíduos de uma mesma espécie, e até mesmo dentro de uma população específica dessa espécie (ARAUJO; ZUCCHI, 2006; FOOTE, 1980; GUSTANI, 2018). Uma hipótese que pode estar contribuindo para as variações nos caracteres morfológicos taxonômicos reside na possível influência da planta hospedeira sobre o desenvolvimento dos indivíduos, uma vez que todo o desenvolvimento embrionário e larval ocorre dentro do fruto. A existência de tal processo indicaria plasticidade corroborando processos adaptativos que favoreceriam a capacidade das espécies em se ajustar ao ambiente em que se desenvolvem (NORRBOM et al., 2015).

Plasticidade fenotípica

A capacidade de um genótipo gerar fenótipos distintos quando exposto a diferentes condições ambientais ao longo de sua ontogenia é definida como plasticidade fenotípica. E estas variações podem ser irreversíveis, ou de respostas lentas, caracterizadas por ajustes graduais e progressivos nos traços fenotípicos ao longo de várias gerações ou fases do ciclo de vida (GOTTHARD; NYLIN; NYLIN, 1995; PIGLIUCCI, 2005; PIGLIUCCI; MURREN; SCHLICHTING, 2006). Portanto, essa variação é decorrente de respostas a diferentes pressões seletivas sobre processos bioquímicos e fisiológicos do desenvolvimento (RAGLAND et al., 2012). Essas respostas seguem um padrão que tende a se assemelhar em espécies mais proximamente relacionadas, possivelmente decorrente

de herança comum ou influência de fatores genéticos compartilhados. Elas também tendem a refletir a árvore filogenética devido a tendências evolutivas e à inércia filogenética. Isso significa que mantém certos padrões ao longo do tempo, mesmo diante de mudanças nas pressões seletivas (CHEVIN; LANDE, 2009).

Os mecanismos que impulsionam a evolução fenotípica podem ser facilitados em diferentes condições ambientais, promovendo diferentes resultados evolutivos (HUANG, 2015). A plasticidade permite que um organismo apresente diferentes fenótipos que podem ser adaptados a dois ou mais ambientes, sem a necessidade da evolução de polimorfismos genéticos (NIJHOUT, 2003). Esta diversidade promovida pela plasticidade ocorre devido às vias de desenvolvimento que moldam os fenótipos influenciados pelo ambiente consistirem em variados componentes genéticos que têm o potencial de responder às pressões de seleção, contribuindo para a adaptação da espécie ao ambiente em mudança (PFENNIG et al., 2010). Além disso, é importante ressaltar que uma mudança induzida pelo ambiente não afeta somente a morfologia, como também é frequentemente acompanhada de mudanças comportamentais e fisiológicas (BOURDEAU, 2009; CORNWALLIS; BIRKHEAD, 2008). Portanto, se esta alteração fenotípica devido à variação ambiental reduzir o valor adaptativo (*fitness*), provavelmente será eliminada pela evolução de mecanismos que de alguma forma reduzem a sensibilidade do fenótipo (mecanismos homeostáticos) (NIJHOUT, 2003).

A plasticidade fenotípica possui herdabilidade, tornando-a sujeita à evolução, essa característica pode ser condicionada tanto por loci ambientalmente responsivos (plásticos), quanto por loci não responsivos que interagem para criar uma dinâmica evolutiva complexa (SCHEINER, 2002; SCHEINER; BARFIELD; HOLT, 2020). Dessa forma, entender a direção da seleção e da mudança fenotípica em caracteres de populações naturais possibilita o reconhecimento de respostas diretas ou indiretas aos efeitos ambientais (ERCIT, 2016), além disso, a variação natural está intrinsecamente associada à força variável da seleção, o que proporciona uma compreensão de partes determinantes da diferenciação morfológica (HAYWARD et al., 2018).

Dessa forma, considerando que o ambiente em que o organismo se desenvolve tem um papel fundamental tanto na pressão de seleção (a influência do ambiente na sobrevivência e reprodução) quanto na expressão de variações fenotípicas (diferenças observáveis nas características do organismo) (HANDEL; EBERS; RAMAGOPALAN, 2010), o efeito ecológico potencial das plantas hospedeiras pode impactar no desenvolvimento dos insetos a elas relacionados. Em outras palavras, a interação entre o

organismo e o ambiente, especialmente no contexto das plantas hospedeiras, desempenha um papel crucial no modo como os insetos se desenvolvem e se adaptam ao seu entorno. Essa interação entre insetos fitófagos e seus hospedeiros os tornam ótimos modelos para se estudar o efeito ambiental na plasticidade fenotípica, uma vez que os hospedeiros são a única fonte de alimentos, particularmente de larvas de holometábolos. Tais efeitos, que podem ser tanto em características morfológicas, comportamentais e até genética (ALUJA et al., 2019), são altamente relevantes na interação inseto-planta de espécies do gênero *Anastrepha*, e em especial na espécie *Anastrepha fraterculus*, uma vez que esta é uma espécie altamente polífaga que se tornou uma importante praga agrícola por infestar uma grande diversidade de frutos e vegetais (NEL; BERTRAND; NEL, 2018; NORRBOM et al., 2015).

Epigenética

A relação entre as moscas das frutas e seus hospedeiros é profundamente interligada (ARAÚJO et al., 2022), uma vez que todo o desenvolvimento do embrião e larval ocorre dentro do fruto. Neste contexto, espera-se que o fruto exerça influência em diversas características destes insetos, podendo até estar sujeito a algum controle epigenético. As mudanças epigenéticas ocasionadas por estímulos ambientais podem induzir respostas rápidas que explicam a especiação. Por exemplo, larvas transferidas para uma nova planta podem, quando adultas, manter memórias dos odores e colonizar um novo hospedeiro, resultando em isolamento reprodutivo em relação à população de origem (LASSERRE, 2014).

Os mecanismos epigenéticos são fortemente vinculados ao meio ambiente, sendo que fatores como mudanças climáticas, estresse, alimentação e componentes químicos podem afetar diretamente o fenótipo (FEIL; FRAGA, 2012; MOGILICHERLA; ROY, 2023; OLIVARES-CASTRO et al., 2021). Estas respostas adaptativas são possibilitadas pela taxa de epimutações da metilação do DNA, que se demonstrou significativamente maior do que a taxa de mutações genéticas (SKINNER et al., 2014). Dessa forma, as populações têm a capacidade de responder ao meio ambiente por meio da variação epigenética antes que as mutações genéticas sejam acumuladas (KLIRONOMOS; BERG; COLLINS, 2013).

As mudanças epigenéticas são definidas como processos moleculares que ocorrem em torno do DNA e regulam a atividade do genoma independente da sequência gênica (SKINNER, 2011), ou seja, tais mudanças podem promover alterações no fenótipo sem afetar o genótipo. A maquinaria que regula as marcações epigenéticas é composta de processos como metilação do DNA e RNA, modificações de histonas, (MARGUERON; REINBERG, 2010; TOLLEFSBOL, 2017) e modificações estruturais da cromatina (HOLLIDAY, 1994), os quais são modulados por sinais ambientais.

Grande parte dos estudos relacionados à epigenética envolvem modificações de histonas e metilação do DNA. A primeira é caracterizada em sua maioria por alterações na porção amino-terminal causadas por acetilações, metilações, fosforilações e ubiquitinação, sendo as principais ocasionadas por acetilação e metilação da lisina (CORTINI et al., 2016). É proposto que o uso combinatório dessas modificações funcione como um sistema de marcação que podem afetar diretamente a estrutura da cromatina e modular a expressão gênica nessa região (QUINA; BUSCHBECK; DI CROCE, 2006). Este código é capaz de ser transmitido para a próxima geração de células como uma “memória” epigenética de programas transcricionais, também conhecido como o epigenoma (JENUWEIN; ALLIS, 2001). Existem dois mecanismos que atuam nestas alterações, um deles é a diminuição do contato pelo nucleossomo, podendo ser entre histona-histona ou histona-DNA, o que modifica tanto a estrutura como a estabilidade, e o outro que está relacionado a proteínas que modificam, de forma enzimática, a cromatina (KOUZARIDES, 2007).

A metilação do DNA é, até o presente, a mais estudada e consiste na adição covalente de um grupo metil a molécula de DNA. Em mamíferos, ocorre com mais frequência em dinucleotídeos CG, conhecidas como ilhas de CpG, estas são comumente encontradas em sequências regulatórias onde mudanças podem impactar a expressão de genes próximos (LEA et al., 2016). Já em insetos, especificamente em *Drosophila melanogaster*, a metilação ocorre principalmente em dinucleotídeos CpA e CpT (LYKO; RAMSAHOYE; JAENISCH, 2000). Esta importante reação de transferência inclui cinco proteínas de ligação metil – DNA (Mbd1, Mbd2, Mbd3, Mbd4 e MeCP2) e é catalisada por três grupos de enzimas DNA metiltransferases: DNMT1, DNMT2 e DNMT3 (ALBALAT, 2008; TOLLEFSBOL, 2017).

Dentre estas enzimas, o grupo DNMT1 se destaca pela importância na manutenção dos padrões de metilação primários, enquanto a DNMT3 é responsável pela metilação “de novo” do DNA em células germinativas e embriões iniciais (DAHLET et al., 2020;

OKANO et al., 1999; SUETAKE et al., 2004) em mamíferos. Já a DNMT2, apesar de suas funções não estarem completamente descritas, apresenta uma baixa atividade no DNA e, em comparação com outras DNMTs, parece ser mais conservadora (SCHAEFER; LYKO, 2010).

Entretanto, é relevante destacar que em *Drosophila* e diversos outros Diptera, inclusive provavelmente Tephritidae, como a espécie aqui estudada, apresentam apenas DNMT2 em seu genoma. Curiosamente, essa enzima não parece ser essencial para diversos processos de metilação, uma vez que moscas deficientes para DNMT2 mantiveram padrões de metilação, embora tenham tido redução da expressão gênica na transição materna-zigótica (TAKAYAMA et al., 2014). Tal resultado indica um papel provável de outras proteínas na determinação da metilação e, potencialmente, de outros mecanismos epigenéticos (DESHMUKH et al., 2020).

Os padrões de metilação do DNA em todo o genoma podem ser estabelecidos pela primeira vez durante o desenvolvimento e, uma vez definidos, são capazes de serem transferidos através das divisões celulares ao longo da vida do organismo, ou mudar em função do ciclo de vida (LEA et al., 2016; LOKK et al., 2014). Em estudos com *D. melanogaster*, por exemplo, foi possível observar uma mudança nos níveis genômicos de metilação ao longo de cada estágio do ciclo de vida, desde embriões a adultos, sendo encontrado em maiores frequências em estágios embrionários (DESHMUKH et al., 2018). É importante destacar ainda que a complexa arquitetura epigenética, envolvendo o cérebro e ovário, pode estar intrinsecamente relacionada com a regulação do comportamento alimentar destas moscas (ANREITER; KRAMER; SOKOLOWSKI, 2017).

Dessa forma, as condições ambientais podem afetar os padrões de metilação tanto durante o desenvolvimento (quando as alterações epigenéticas são consideradas sensíveis) como a posteriori, quando as metilações auxiliam a coordenar a resposta celular a novos estímulos ambientais (GUO et al., 2011). Portanto, evidências crescentes associam a variação morfológica a marcas epigenéticas, sugerindo que a metilação do DNA desempenha um papel significativo na mediação da resposta regulatória do gene às mudanças ambientais (LAUBACH et al., 2018; THORSON et al., 2017; VERHOEVEN; PREITE, 2014).

Apesar de ser uma área promissora para a compreensão de mecanismos adaptativos e filogenéticos (LASSERRE, 2014), existem poucos estudos que utilizam a maquinaria epigenética para entender a dinâmica da plasticidade de caracteres morfológicos de insetos. A variabilidade fenotípica encontrada em estudos anteriores de

Anastrepha fraterculus, bem como suas características comportamentais, tornam este complexo excelente para estudos que visam observar a atuação dos fatores epigenéticos em suas populações. Isso ocorre porque esta maquinaria está intimamente relacionada com o ambiente de desenvolvimento das espécies, podendo trazer respostas adaptativas e até mesmo contribuir para os processos de especiação. Em particular, a investigação de fatores epigenéticos que podem modular a infestação de diversos hospedeiros é importante não apenas por expandir essas informações para além de organismos modelo, mas principalmente por fornecer “insights” sobre regiões no genoma que podem ser alvo de investigações para o controle dessa importante praga no país.

Um dos maiores desafios enfrentados em estudos de epigenética está relacionado à identificação de marcas epigenéticas no sequenciamento do DNA, uma vez que a maioria das tecnologias não consegue distinguir diretamente entre bases metiladas e não metiladas (SCHATZ, 2017), especialmente em certos taxa nos quais níveis de metilação são reduzidos em fase adulta, como em *Drosophila* e potencialmente *Anastrepha* (LEWIS et al., 2020). Algumas metodologias como, PCR de LINE-1 seguida de pirosequenciamento e HPLC-UV (*High Performance Chromatography – Ultraviolet*), como MSCC (*Methyl-Sensitive Cut Counting*), técnicas com Microarray ou Bead Array e Sequenciamento por Bissulfito podem ser utilizadas para traçar perfis de metilação, tanto completo como regiões diferencialmente metiladas (KURDYUKOV; BULLOCK, 2016). Entre estes o tratamento com bissulfito de sódio vem sendo um padrão para estas análises, e consiste na conversão de citosinas não modificadas em uracila, enquanto as citosinas metiladas permanecem inalteradas; porém além de ser um processo mais trabalhoso, resulta em uma extensa fragmentação de DNA que limita sua amplificação (SIMPSON et al., 2017).

Com a intenção de minimizar estes problemas, novas tecnologias têm sido testadas a fim de aprimorar os resultados e otimizar as condições laboratoriais, dentre estas, destaca-se o dispositivo de sequenciamento MinION da *Oxford Nanopore Technologies*. Este é um dispositivo de molécula única baseado em nanoporos de alto rendimento, que pode sequenciar fragmentos de DNA longos e não amplificados (IP et al., 2015). O MinION registra variações dependentes da sequência na corrente iônica ao longo do tempo, à medida que o DNA atravessa nanoporos de proteína inseridos em uma membrana lipídica. Isso ocorre porque os sinais de corrente são sensíveis a modificações na base, como 5-metil citosina (5-mC) ou N6-metil adenina (6mA) (CHEN et al., 2023; SCHATZ, 2017).

Desse modo, assim como os padrões reproduzíveis podem ser interpretados como sequências curtas de DNA específicos, as bases metiladas podem gerar assinaturas reproduzíveis e interpretáveis (IP et al., 2015; RAND et al., 2017; SIMPSON et al., 2017) de uma forma direta, e sem demandar modificações como o bissulfito, que facilitam a identificação de regiões com marcas epigenéticas com muito mais sensibilidade, permitindo o estudo mesmo em tecidos com baixos níveis de metilação (SHAH; CAO; ELLISON, 2019) a um custo relativamente acessível.

OBJETIVOS

O objetivo geral deste trabalho foi investigar o efeito de hospedeiros em populações de *Anastrepha fraterculus*, visando identificar variações fenotípicas em diversos caracteres morfológicos, bem como investigar a presença de marcas epigenéticas entre populações que emergiram de diferentes frutos hospedeiros. Para isso, os objetivos específicos foram:

- Analisar variabilidade de caracteres morfológicos relacionados tanto a aspectos reprodutivos quanto de importância taxonômica em moscas-das-frutas *A. fraterculus* provenientes de frutos distintos.

- Encontrar padrões epigenéticos, como metilação 5mC e 6mA, pela primeira vez em *A. fraterculus* derivadas de frutos e populações distintas, utilizando dados obtidos por meio de sequenciamento com o equipamento MinION da *Oxford Nanopore Technologies*.

- Identificar regiões diferencialmente metiladas entre moscas *A. fraterculus* que emergiram de diferentes frutos, com o intuito de observar o efeito do hospedeiro sobre estas moscas.

- Identificar os conjuntos de genes presentes nas regiões diferencialmente metiladas, e realizar uma análise funcional para caracterizar a quais processos estão relacionados.

CAPÍTULO 1 – Variação fenotípica de moscas *Anastrepha fraterculus* que emergiram de diferentes frutos

1.1 INTRODUÇÃO

Espécies pertencentes ao grupo *fraterculus*, em geral, são identificadas através de marcadores morfológicos (ARAUJO; ZUCCHI, 2006; LAMPERT et al., 2020; PERRE et al., 2014; STONE, 1942) e análises morfométricas. No entanto, estudos recentes revelaram uma considerável variabilidade em vários desses marcadores, inclusive dentro de uma mesma população, e mesmo em características consideradas cruciais para a distinção entre espécies, como o ovipositor (ARAUJO; ZUCCHI, 2006; LAMPERT et al., 2020; PERRE et al., 2014). Uma das hipóteses levantadas para entender esta variação é a influência que a planta hospedeira pode ter sobre o desenvolvimento dos indivíduos.

A interação entre insetos fitófagos e seus hospedeiros tem sido amplamente explorada na busca por compreender a importância da condição de generalistas e especialistas, uma vez que, estas condições são fundamentais para compreender a distribuição, diferenciação, especialização e/ou especiação de novas espécies (VELASCO-CUERVO; GALINDO-GONZÁLEZ; TORO-PEREA, 2023). Esta associação é particularmente importante em moscas-das-frutas da família Tephritidae, como o grupo *fraterculus*, pois a relação inseto-planta é intensa e todo o desenvolvimento embrionário e larval ocorre dentro do hospedeiro. Assim pode-se esperar que o hospedeiro exerça uma influência significativa sobre essas moscas, considerando que cada planta apresenta características fenológicas, físicas e químicas específicas (MCPHERON; SMITH; BERLOCHER, 1988; VELASCO-CUERVO; GALINDO-GONZÁLEZ; TORO-PEREA, 2023).

Além disso, o fenótipo de um organismo frequentemente responde a condições ambientais diversas, sujeito a pressões de seleção distintas (GÓMEZ-CENDRA et al., 2016; NIJHOUT, 2003). Isso ocorre devido à plasticidade fenotípica, que capacita um organismo a expressar diferentes fenótipos adaptativos em dois ou mais ambientes, sem a necessidade da evolução de polimorfismos genéticos (PIGLIUCCI, 2005; WANG; ALTHOFF, 2019). Dessa forma, indivíduos do grupo *fraterculus*, em especial neste trabalho com *Anastrepha fraterculus*, constituem um excelente modelo para avaliar a influência do fruto hospedeiro na plasticidade fenotípica de caracteres morfológicos. Uma vez que, diferentes frutos hospedeiros podem contribuir para a variação morfológica e

comportamental nas populações selvagens (GÓMEZ-CENDRA et al., 2016) torna-se, fundamental compreender como as condições ambientais podem modular as características morfológicas desses organismos.

A caracterização das variações morfológicas entre diferentes frutos hospedeiros, pode-se valer por meio de técnicas e análises de morfometria. Esta técnica pode ser definida como uma fórmula matemática e estatística, para quantificar as diferenças entre as formas, transformando essas informações visuais em dados mensuráveis para identificar as espécies (MONTEIRO; REIS, 1999; SILVA, 2008). As análises morfométricas multivariadas têm se mostrado úteis para identificar diferenças específicas entre táxons mais relacionados e para a detecção de diferenças morfológicas em diversos organismos, incluindo moscas das frutas (HERNÁNDEZ-ORTIZ et al., 2004; KHAMIS et al., 2012). Além disso, essa abordagem pode ser empregada em pesquisas multidisciplinares, sendo comumente correlacionada a estudos genéticos (DIAS et al., 2016; ROHLF, 2002).

1.2 MATERIAL E MÉTODOS

1.2.1 Coleta e Criação de Moscas

As coletas de frutos foram realizadas no *campus* São Carlos da Universidade Federal de São Carlos (UFSCar), onde foram coletados 5 frutos: Uvaia (*Eugenia pyriformes*), pitanga (*E. uniflora*), cereja do rio grande (*E. involucrata*), goiaba (*P. guajava*) e guabiroba (*Campomanesia xanthocarpa*) em épocas diferentes, em função de seu aparecimento em estações distintas ao longo do ano. Em seguida, os frutos coletados foram mantidos em laboratório com temperatura e umidade controlada e fotoperíodo natural, alocados em caixas de acrílico com vermiculita, onde permaneceram por duas a três semanas, passando por uma etapa de peneiragem semanal, para a coleta de pupas. Estas pupas foram transferidas para gaiolas, contendo água e comida, nesse mesmo ambiente, até que adultos emergissem.

À medida que os adultos emergiram, as moscas foram separadas por sexo e mantidas em novas gaiolas separadas de acordo com cada população/fruto (neste trabalho foi considerado como população os grupos de indivíduos que emergiram do mesmo fruto), até que atingissem o amadurecimento sexual (cerca de 15 dias), a fim de garantir que não houvesse acasalamento. Dada a importância da identificação de espécies no gênero,

especialmente no grupo *fraterculus*, e considerando que o acúleo é uma das principais estruturas com valor taxonômico, as análises foram restritas exclusivamente às fêmeas. Estes indivíduos foram então sacrificados, por meio de congelamento a -20 °C, para posterior extração de DNA e análises morfométricas.

Desse modo, estruturas de asa (parte dorsal) e acúleo (parte ventral) foram retiradas e montadas em lâminas individuais sob uma lamela com bálsamo do Canadá, o que permitiu tanto análises morfométricas quanto para a identificação das espécies, seguindo a chave de identificação descrita por Zucchi (2000). Dos espécimes analisados, 117 foram devidamente identificados como *Anastrepha fraterculus* (Uvaia n= 22; Pitanga n= 30; Goiaba n= 42; Gabiroba n= 17; Cereja do cerrado n= 6), e foram utilizados para as análises.

1.2.2 Medidas morfométricas

Para as análises morfométricas consideramos quatro estruturas para as medidas lineares, como proposto por Hernández-Ortiz et al., (2012) e por Gómez-Cendra *et al.* (2016), sendo elas: asa direita (parte dorsal), acúleo (parte ventral), cabeça e tórax (Figura 1).

As imagens das estruturas aqui estudadas (Figura 1) foram captadas com auxílio do microscópio Estereoscópico Leica S APO (Leica Microsystems), e a estrutura da cabeça foi alocada em um suporte com régua visível para evitar a distorção durante a digitalização (Figura 1a). O estudo morfológico focou em um conjunto de caracteres relevantes já comumente utilizados para identificar morfotipos de espécimes de *A. fraterculus*, que estão bem estabelecidas nos trabalhos de Hernández-Ortiz *et al.* (2012) e Hernández-Ortiz *et al.* (2004), bem como, variáveis importantes para encontrar variabilidade e plasticidade fenotípica neste grupo entre hospedeiros distintos (GÓMEZ-CENDRA et al., 2016; GUSTANI, 2018). As medidas foram expressas como distâncias lineares entre dois pontos e razões entre duas variáveis, sendo estas realizadas no software Leica Las para as seguintes variáveis:

Variáveis da cabeça: HEL - comprimento do olho; HL – comprimento da cabeça; HW – largura da cabeça; FW - largura mínima da face

Variáveis do tórax: TL - comprimento do tórax; TW - largura do tórax.

Variáveis da asa: **W1**- comprimento da asa **W2** - largura da asa no ápice até o braço da banda V; **W4** - distância da extremidade proximal do braço proximal da banda V na margem posterior da asa; **W7** - razão entre largura e comprimento da asa ($W2/W1$).

Variáveis do acúleo: **A2** - Medida da abertura do oviduto; **A3** - largura no início de seção serrilhada; **A4** - comprimento da ponta basal da sessão não serrilhada; **A5** - comprimento da ponta apical da seção serrilhada; **A6** - comprimento do lado esquerdo da região serrilhada; **A8** - comprimento da ponta do acúleo ($A4+A5$); **A9** - razão do comprimento da região não serrilhada versus o comprimento da seção serrilhada ($A4/A5$); **A10** – razão do comprimento da ponta do acúleo versus abertura do oviduto ($A8/A2$).

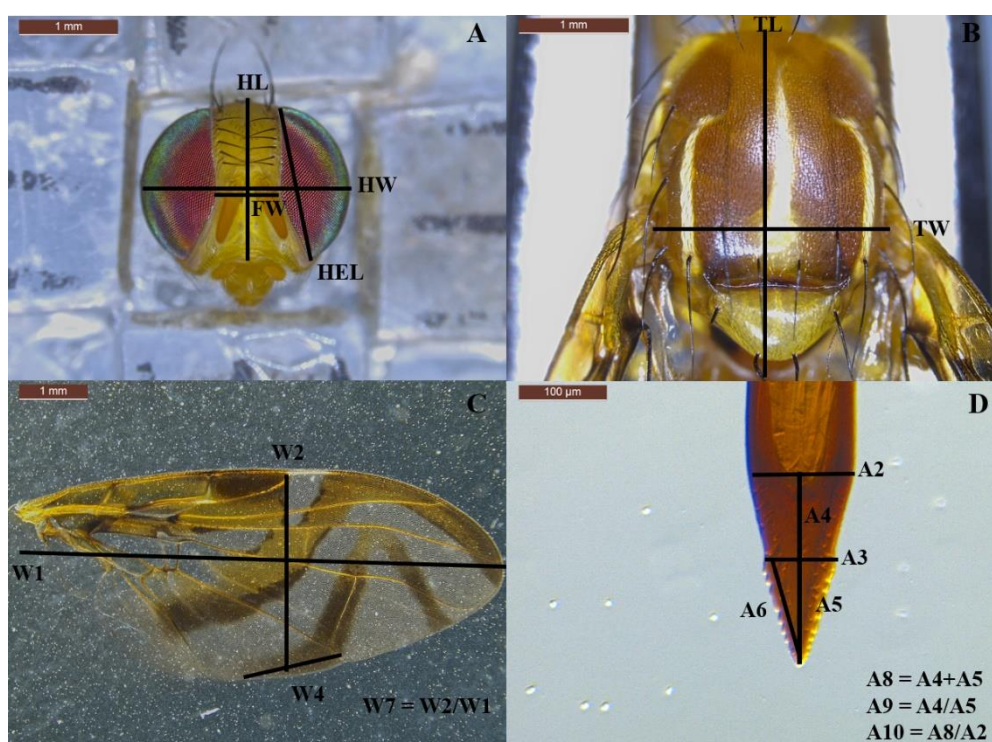


Figura 1. Variáveis consideradas em medidas morfométricas para as 4 estruturas propostas. (a) Cabeça: **HEL** - comprimento do olho; **HL** - comprimento da cabeça; **HW** - largura da cabeça; **FW** - largura mínima da face. (b) Tórax: **TL** - comprimento do tórax; **TW** - largura do tórax. (c) Asa: **W1**- comprimento da asa **W2**- largura da asa no ápice até o braço da banda V; **W4** - distância da extremidade proximal do braço proximal da banda V na margem posterior da asa; **W7** - razão entre largura e comprimento da asa ($W2/W1$). (d) Acúleo: **A2** - Medida da abertura do oviduto; **A3** - largura no início de seção serrilhada; **A4** - comprimento da ponta basal da sessão não serrilhada; **A5** - comprimento da ponta apical da seção serrilhada; **A6** - comprimento do lado esquerdo da região serrilhada; **A8** - comprimento da ponta do acúleo ($A4+A5$); **A9** - razão do comprimento da região não serrilhada versus o comprimento da seção serrilhada ($A4/A5$); **A10** - razão do comprimento da ponta do acúleo versus abertura do oviduto ($A8/A2$).

1.2.3 Análises estatísticas

Avaliamos a variabilidade fenotípica entre moscas *Anastrepha fraterculus* que emergiram de diferentes hospedeiros aqui coletados. Para isso, inicialmente foi aplicado o teste de normalidade de Ryan-Joiner no software *Minitab v.19* para todas as variáveis

morfológicas, a fim de averiguar se nossos dados apresentavam uma distribuição normal. Posteriormente realizamos uma análise de correlação de Person no software *Past v.4.12*. As variáveis que apresentaram uma correlação menor que 0,8, foram utilizadas na análise de componentes principais (PCA), uma vez que indicam uma relação mais fraca entre as variáveis. Isso indica que essas variáveis estariam menos linearmente relacionadas, o que pode ser desejável para uma análise de PCA para destacar diferentes aspectos da variação nos dados.

O PCA foi então aplicado, com a finalidade de observar quais caracteres explicavam melhor o nosso conjunto de dados. Durante a aplicação, variáveis que apresentaram uma mesma relação na explicação dos componentes principais foram removidas, optando por manter aquela de menor significância. Desse modo, realizamos uma análise de função discriminante (DFA), com tratamento de outliers para as variáveis que melhor explicaram o nosso conjunto de dados. O tratamento de outliers é realizado para mitigar o impacto de valores extremos que poderiam distorcer os resultados, assegurando uma análise mais robusta e confiável. A DFA classifica indivíduos desconhecidos com base nas características medidas, atribuindo uma probabilidade de sua classificação em um determinado grupo. Essa análise pressupõe que as amostras possuam uma distribuição normal para as características consideradas (MOORE, 2013).

Também aplicamos o teste Lambda de Wilks para avaliar a significância estatística, que varia de 0 a 1, das variáveis discriminantes e do poder discriminatório de cada modelo para cada variável. Essas medidas estão relacionadas ao critério da razão de verossimilhança dos dados, utilizado para agrupar os indivíduos de acordo com o seu hospedeiro. E por fim, foi medida a distância de Mahalanobis (MD), que é definida por (SCHOONHOVEN; LOON; DICKE, 2005) Mahalanobis (1936) como um cálculo baseado nas correlações entre variáveis das medidas morfológicas, com as quais distintos padrões podem ser identificados e analisados independente da escala das medições. Esse método leva em conta grupos de pares não ponderados com a média aritmética e é uma análise de agrupamento que estabelece as relações fenéticas entre as populações. Esta análise, expressa através de uma tabela, possibilitou a observação das amostras que apresentaram maiores diferenças entre os grupos (frutos).

1.3 RESULTADOS

A análise de correlação de Person indicou que as variáveis A4, A5, A6, A10 e W4 estavam altamente correlacionados com a variável A8 (comprimento do acúleo). Por esse motivo, nós as excluímos, uma vez que, estatisticamente elas explicavam basicamente a mesma informação. Além disso, a variável W2 apresentou correlação com W1 e foi removida para as análises subsequentes.

A análise de componentes principais (Figura 2) revelou que as variáveis estudadas explicaram 41% do eixo x (componente 1), sendo as principais contribuintes A3, A9, TL e W1. Para o eixo y (componente 2) foi explicado apenas 29%, sendo as variáveis HL e HW (medidas de cabeça) as que mais contribuíram. A partir da Figura 2, notamos que as variáveis HL e HW exercem uma influência positiva sobre as amostras de pitanga, assim como as variáveis A8 (comprimento da ponta do acúleo) e TL (comprimento do tórax) indicam uma forte explicação para as amostras de goiaba e guabiroba. Uvaia apresentou maior distribuição dentro dos dois eixos, porém a variável A3 (largura no início de seção serrilhada) esteve no eixo mais representativo para este grupo.

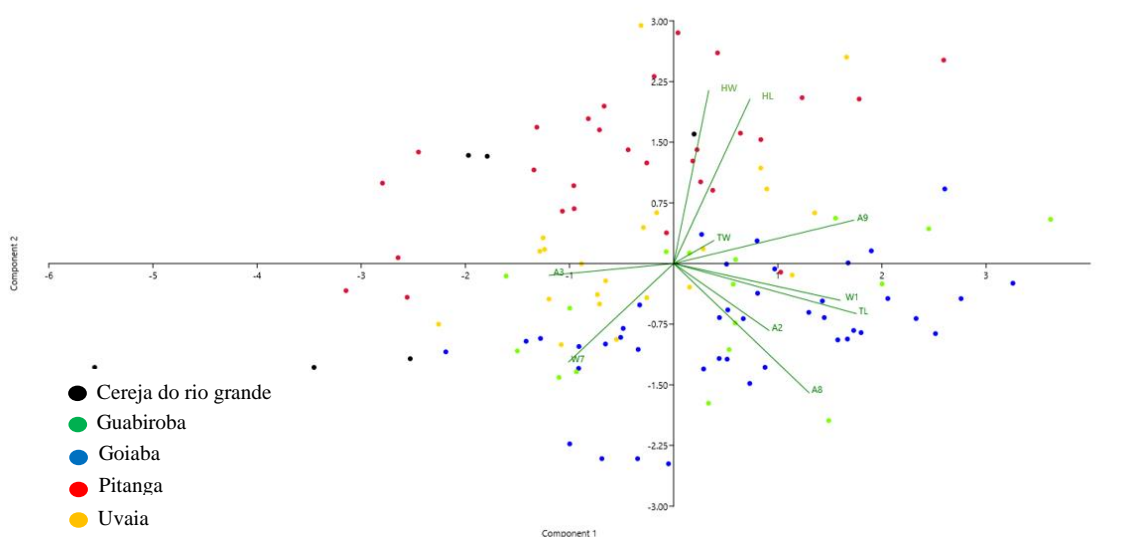


Figura 2. PCA para os frutos cereja do rio grande, guabiroba, goiaba, pitanga e uvaia. A2 - Medida da abertura do oviduto; A3 - largura no início de seção serrilhada; A8 - comprimento da ponta do acúleo (A4+A5); A9 - razão do comprimento da região não serrilhada versus o comprimento da seção serrilhada (A4/A5); W1- comprimento da asa W7 - razão entre largura e comprimento da asa (W2/W1); TL – comprimento do tórax; TW – largura do tórax; HEL – comprimento da cabeça; HW- largura da cabeça.

A análise de função discriminante (Figura 3) gerou quatro eixos ortogonais que podem melhor explicar os dados, sendo o eixo 1 e 2 os mais significativos (66,5% e 24,11%, respectivamente). Ao aplicar elipses de confiança de 95% para análises merísticas, observamos que três grupos foram separados diferencialmente. O primeiro grupo é composto majoritariamente por espécimes que emergiram de uvaia; o segundo

inclui espécimes de goiaba e guabiroba, enquanto o terceiro é composto por amostras de pitanga e cereja do rio grande.

A sobreposição de espécimes de goiaba e guabiroba indica uma similaridade entre os indivíduos considerando as variáveis analisadas. No entanto, mesmo apresentando valores estatisticamente semelhantes, elas não são idênticas. Uma das análises da função discriminante inclui a classificação (Tabela S1), que indica se uma amostra de um grupo realmente pertence a esse grupo. Esta análise revelou que 67% das amostras que emergiram de goiaba foram corretamente classificadas como goiaba, enquanto 32% foram consideradas mais próximas de guabiroba. Já entre os espécimes provenientes de guabiroba, 59% foram classificadas como pertencentes a este hospedeiro, enquanto 41% foram identificadas como pertencentes a goiaba.

As amostras provenientes de pitanga se encontram bem definidas em comparação com as dos outros grupos. Na análise de classificação, 82% dos espécimes foram corretamente identificados como pertencentes a pitanga, enquanto apenas 11% foram classificadas como cereja do rio grande e, os outros 7% pertencentes a uvaia ou guabiroba. A sobreposição com cereja de rio grande ocorre principalmente devido a 33% das amostras desse grupo serem classificadas como espécimes provenientes de pitanga, um resultado possivelmente influenciado pelo baixo número amostral para espécimes que emergiram deste fruto (n=6).

Além disso, os resultados do teste de Lambda de Wilks (Tabela S2) sugerem que as variáveis A3, A8, W1, HW, TL e TW foram as que melhor contribuíram para a separação dos grupos. Estas variáveis apresentaram um valor de lambda próximo a zero, indicando uma contribuição expressiva, e um p-valor dentro do limite de significância (p-value < 0,05).

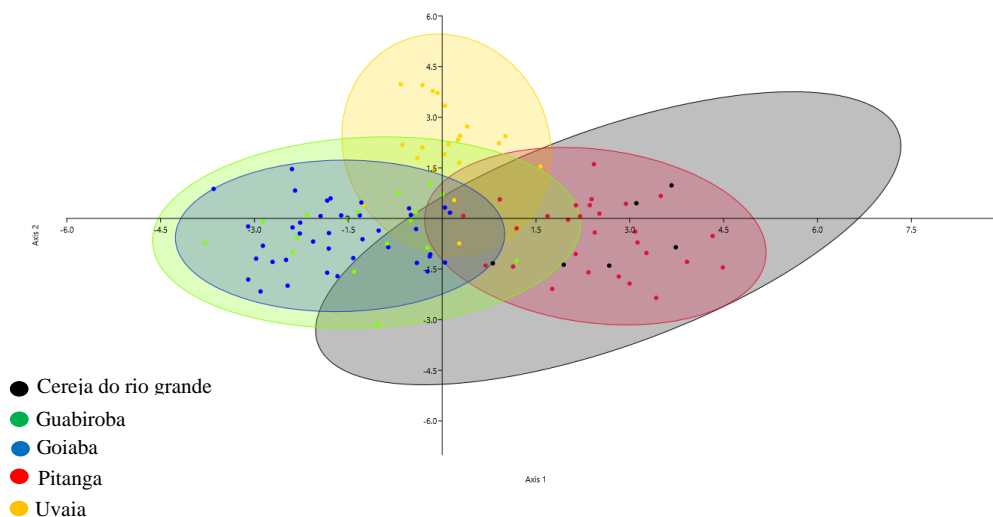


Figura 3. Análise de função discriminante separando as amostras que emergiram de hospedeiros como Uvaia, pitanga, cereja do rio grande, goiaba e guabiroba, com elipses de 95% de confiança.

A análise estatística da distância de Mahalanobis (Tabela 1) revelou as disparidades de similaridade entre os grupos, destacando as diferenças entre as amostras de diferentes frutas. Observou-se que as maiores disparidades estavam presentes nas moscas provenientes da Goiaba e Cereja do Rio Grande (23,243), Guabiroba e Cereja do Rio Grande (22,033), e Pitanga e Goiaba (17,077). Por outro lado, os indivíduos das frutas que apresentaram maior semelhança foram os que emergiram de Goiaba e Guabiroba (0,92384), evidenciando um valor muito mais próximo em comparação com outras moscas dos grupos (frutas) analisados.

Tabela 1. Distância de Mahalanobis para os grupos estudados.

	Cereja do Rio Grande	Uvaia	Pitanga	Goiaba	Guabiroba
Cereja do Rio Grande					
Uvaia	19,638				
Pitanga	8,0263	13,313			
Goiaba	23,243	10,245	17,077		
Guabiroba	22,033	8,8303	13,781	0,92384	

1.4 DISCUSSÃO

Insetos fitófagos representam modelos ótimos para explorar os mecanismos subjacentes à adaptação e divergência interespecífica. Isso ocorre devido à influência direta exercida pelos hospedeiros desses insetos nos estágios iniciais do ciclo de vida

(SCHOONHOVEN; LOON; DICKE, 2005). Esses hospedeiros não apenas desempenham um papel crucial como o fator ambiental mais próximo, mas também moldam o ambiente em que o organismo se desenvolve, exercendo influência tanto na pressão de seleção quanto na expressão de variações fenotípicas (HANDEL; EBERS; RAMAGOPALAN, 2010; NIJHOUT, 2003; SCHOONHOVEN; LOON; DICKE, 2005).

Um desafio significativo nos estudos de moscas-das-frutas, especialmente no grupo *fraterculus*, como é o caso desse estudo, é a diferenciação com base em caracteres morfológicos taxonômicos. Contudo, determinar os limites dessa diferenciação é uma tarefa complexa devido à considerável variação desses caracteres, tanto entre diferentes populações quanto dentro das mesmas (ARAUJO; ZUCCHI, 2006). Diante desse desafio, o presente estudo buscou fornecer uma compreensão mais aprofundada da complexidade na variação morfológica de *A. fraterculus*. Nossos resultados revelaram diferenças morfológicas entre moscas que emergiram de diferentes hospedeiros, cuja importância variou entre as diferentes variáveis mensuradas.

Cada análise desempenhou um papel específico, oferecendo insights distintos sobre os dados e proporcionando uma compreensão abrangente dos aspectos morfométricos das moscas-das-frutas estudadas. A Análise de Componentes Principais (PCA) reduziu a dimensionalidade dos dados, destacando algumas variáveis que influenciaram significativamente os eixos x e y e permitiram a melhor caracterização fenotípica da variabilidade intra-populacional entre frutos distintos. A Análise de Função Discriminante foi fundamental para entender as diferenças entre os grupos, evidenciando que as moscas provenientes de uvaia formaram um grupo distinto, enquanto goiaba e guabiroba apresentaram sobreposição, e pitanga e cereja do rio grande constituíram outro grupo identificável. É interessante ressaltar que tais similaridades espelham relações de parentesco entre essas espécies de hospedeiros de cereja do rio grande e pitanga, e mesmo entre essas e uvaia (GIARETTA et al., 2022; VASCONCELOS et al., 2017).

A validação da diferenciação entre grupos foi realizada através de elipses de confiança durante a análise discriminante, confirmando a distinção entre eles e destacando a notável similaridade entre as amostras de goiaba e guabiroba. A eficácia do modelo discriminante foi respaldada pela classificação das amostras, revelando altas taxas de acerto na identificação dos grupos, fortalecendo assim a confiança na capacidade do modelo em distinguir entre os diferentes hospedeiros. Finalmente, a análise da distância de Mahalanobis ofereceu uma perspectiva única sobre as disparidades de similaridade entre os grupos. Destacou-se que as moscas provenientes de goiaba e guabiroba eram mais

semelhantes entre si do que em relação a outros grupos analisados, evidenciando padrões distintos de morfometria e variação fenotípica. Tal similaridade contrasta com informações de relações de parentesco entre estas espécies, que sugerem que o gênero *Campomanesia* é mais aparentado de *Eugenia* do que de *Psidium* (VASCONCELOS et al., 2017). Por sua vez, os frutos desses hospedeiros apresentam características mais similares entre si do que com a de outros hospedeiros aqui estudados, o que pode ressaltar a associação entre plasticidade e características específicas dos frutos dos hospedeiros.

Caracteres como a largura no início de seção serrilhada (A3), o comprimento da ponta do acúleo (A8), comprimento da asa (W1), largura da cabeça (HW), comprimento do tórax (TL) e largura do tórax (TW), se mostraram cruciais na demarcação e diferenciação entre os grupos estudados. Essas variáveis não apenas desempenharam um papel fundamental, mas a significância estatística reforçou a relevância delas na caracterização das diferenças morfométricas. A relevância desses caracteres está por exemplo nos três primeiros, pois eles são comumente utilizados para a identificação das espécies (ARAUJO; ZUCCHI, 2006; HERNÁNDEZ-ORTIZ et al., 2004, 2012).

As variações encontradas em nossos estudos podem estar intrinsicamente relacionadas com a relação inseto-hospedeiro, uma vez que o ambiente em que o organismo se desenvolve tem um papel fundamental na pressão de seleção e na expressão de variações fenotípicas (HANDEL; EBERS; RAMAGOPALAN, 2010; NIJHOUT, 2003). A análise de várias características das moscas-das-frutas revelou que as propriedades dos hospedeiros, ou seja, os alimentos nos quais as larvas se alimentam e se desenvolvem, desempenham um papel crucial que podem estar contribuindo para a sua adaptação. A interação com diferentes hospedeiros gera variações morfológicas evidentes que podem dificultar, por exemplo, a distinção entre espécies intimamente relacionadas. Isso acontece porque, durante a rápida mudança genética, não há tempo suficiente para que características únicas sejam fixadas, tornando as variações morfológicas mais pronunciadas (DIAS et al., 2016; HERNÁNDEZ-ORTIZ et al., 2015).

É evidente que as moscas do gênero *Anastrepha* apresentam uma variabilidade intraespecífica complexa. Essa variação pode manifestar-se com uma estruturação filogenética específica em algumas características anatômicas, que apresentam significativo sinal filogenético para certos caracteres morfológicos em alguns níveis das relações de parentesco entre as moscas (GUSTANI, 2018). O sinal filogenético pode ser definido como o grau em que o parentesco filogenético entre os táxons está associado à sua similaridade fenotípica, desse modo, o sinal filogenético é alto quando espécies

intimamente relacionadas exibem valores de características semelhantes (KAMILAR; COOPER, 2013; KLINGENBERG; GIDASZEWSKI, 2010). É importante ressaltar que esse padrão filogenético não abrange todas os marcadores estudados, revelando uma complexidade na dinâmica evolutiva desses organismos. Notavelmente, certas características, como as medidas do acúleo por exemplo, apesar de demonstrarem sinal filogenético que confere utilidade no seu uso para identificação de certas espécies (GUSTANI, 2018), também apresenta evidência de plasticidade em moscas derivadas de frutos distintos (GUSTANI, 2018), e observado neste estudo. Esses resultados sugerem que a evolução morfológica em *Anastrepha* pode estar associada a uma resposta à seleção direcional, possivelmente em função da exploração de diferentes hospedeiros.

Foi possível observar em nossos resultados, diferenças morfométricas tanto entre populações de frutos diferentes, como dentro da mesma população nas quais 32% das moscas que emergiram de goiaba apresentaram semelhanças morfométricas com aquelas que emergiram de guabiroba, indicando considerável similaridade (0,92384) entre as características morfológicas desses dois grupos. Uma explicação possível para essa semelhança pode estar relacionada à época da coleta das amostras, uma vez que as frutas amadurecem em períodos diferentes. A goiaba amadurece de janeiro a março, enquanto a guabiroba amadurece de novembro a dezembro, de forma que estas podem representar a mesma população em pontos diferentes de seu ciclo anual.

A possibilidade de migração entre hospedeiros, adaptando-se às características específicas, foi sugerida em um estudo sobre *A. fraterculus* de uma mesma região da Argentina, que identificou diferenças morfológicas entre amostras de três hospedeiros (pêssego, goiaba e noz), mesmo compartilhando a mesma área geográfica (GÓMEZ-CENDRA et al., 2016). Entretanto amostras de goiaba e pêssego que foram coletadas em períodos diferentes devido a sazonalidade apresentaram maiores semelhanças morfométricas, assim como goiaba e guabiroba neste estudo. Os autores ressaltam que essas divergências morfométricas, especialmente entre moscas de diferentes hospedeiros, podem aumentar a probabilidade de acasalamentos entre indivíduos originados do mesmo hospedeiro, possivelmente servindo como ponto de partida para um processo de diferenciação simpátrica, cujos resultados finais ainda permanecem incertos.

A combinação de polimorfismos ancestrais, ou o sorteamento incompleto de linhagens (ILS, do inglês "*Incomplete Lineage Sorting*"), provocado pela separação recente entre espécies do grupo *fraterculus* com potencial hibridação ou introgressão pode tornar mais complexo o estudo de relações de parentesco entre espécies do grupo, e sua

adaptação a hospedeiros. Contudo, a existência de tais processos pode fazer com que haja a persistência de fluxo gênico mesmo entre espécies incipientes, como é o caso das borboletas *Heliconius*, que apresentam colorações aposemáticas distintas a predadores diferentes que sofrem introgressão em regiões de contato entre espécies (MARTIN et al., 2013). Tal processo pode sugerir que o genoma de espécies incipientes seja fluido e mais permeável em certas partes, permitindo adaptações locais e introgressão direcional em apenas certas partes do genoma, até mesmo entre espécies que divergiram há mais tempo. O mesmo pode estar ocorrendo aqui, se houver de fato adaptações a hospedeiros distintos, e pode trazer consequências importantes para o processo de dispersão e adaptação de moscas-das-frutas e seu controle.

Dessa forma, o processo de adaptação a diferentes hospedeiros pode envolver uma variedade de caracteres morfológicos (SEGURA et al., 2018). Este fenômeno, amplamente discutido na literatura, encontra respaldo em nossos dados preliminares. Há indícios substanciais de que fatores epigenéticos possam estar orquestrando esses efeitos, exercendo influência nos processos adaptativos e contribuindo para a especiação (LAUBACH et al., 2018; NORRBOM et al., 2015). Os padrões de compatibilidade, associados à plasticidade fenotípica, sugerem que mudanças comportamentais podem ser transmitidas à prole por meio de mecanismos epigenéticos, oferecendo uma forma alternativa à herança determinada pelo padrão mendeliano tradicional (LASSERRE, 2014).

CAPÍTULO 2 - Marcas epigenéticas de Moscas-das-Frutas *Anastrepha fraterculus*

2.1 INTRODUÇÃO

Modificações moleculares que podem afetar o fenótipo e regular a atividade do genoma sem alterações na sequência de nucleotídeos são também conhecidas como mecanismos epigenéticos (BIRD, 2007; SKINNER, 2011). Estas modificações podem ser transmitidas tanto mitoticamente, por meio de divisões celulares em células somáticas, quanto por meio da meiose, passando a informação entre gerações (FEIL; FRAGA, 2012; JENUWEIN; ALLIS, 2001) em diversos organismos distintos, incluindo em insetos (YAGOUND et al., 2020). Vários fatores afetam o processo de regulação de padrões epigenéticos, que podem ser alterados até mesmo em resposta a fatores ambientais, assim como por diversos processos da via epigenética (DE CARVALHO, 2023).

Diversos estudos focam na herança epigenética intergeracional, especialmente na maneira pela qual os indivíduos transmitem as informações epigenéticas para sua prole (GLASTAD; HUNT; GOODISMAN, 2019). O que torna a herança epigenética interessante é a possibilidade de que essas modificações epigenéticas sejam transmitidas para as gerações seguintes, impactando o fenótipo sem modificar a sequência de DNA. Isso implica que as experiências e influências ambientais de uma geração podem ter efeitos duradouros nas gerações futuras através dessas alterações epigenéticas hereditárias. Na verdade, estas informações podem ser estáveis ou não, tendo impactos significativos nos processos evolutivos. A possibilidade de um fenótipo, codificado por uma mudança epigenética desencadeada ou não pelo ambiente, ser transmitido e evoluir, resultando em adaptações independentes da sequência de DNA (PEREZ; LEHNER, 2019), é uma questão central nesse contexto. Alternativamente, essa transmissão estável pode ser atribuída ao controle genético sobre os efeitos epigenéticos que levam à mudança fenotípica (DE CARVALHO, 2023; STAJIC; JANSEN, 2021).

Existem três mecanismos principais que estão associados a processos epigenéticos e são candidatos a transportar informações epigenéticas entre gerações, sendo eles a metilação do DNA, modificações pós-traducionais de histonas e pequenos RNAs não codificantes (YAGOUND et al., 2020). Dentre estes mecanismos, a metilação de DNA consiste na adição covalente de grupos metil a moléculas de DNA e é até o momento a mais estudada. A presença desse mecanismo é particularmente evidente em vertebrados, cujo níveis de metilação de DNA são elevados em todo o genoma (KLUGHAMMER et

al., 2023). Essa característica difere significativamente dos organismos invertebrados, os quais, em geral, exibem níveis reduzidos de metilação, concentrando-se mais frequentemente em regiões codificadoras (AL ADHAMI et al., 2022; PROVATARIS et al., 2018; VILLAGRA; FRÍAS-LASSERRE, 2020). Essa observação sugere que a metilação desempenhe um papel no splicing alternativo de genes e na regulação da expressão (OLDROYD; YAGOUND, 2021; SUZUKI; BIRD, 2008).

Os estudos sobre metilação em insetos ganharam destaque, especialmente quando se observou que a redução da expressão de *dnmt3*, enzimas DNA metiltransferases responsáveis pela metilação "de novo" do DNA em células germinativas e embriões iniciais (DAHLET et al., 2020), em larvas de abelhas resultou em uma mudança do fenótipo de operária para rainha (GLASTAD; HUNT; GOODISMAN, 2019; KUCHARSKI et al., 2008). Esses resultados indicam que, mesmo em níveis considerados baixos de metilação, esse mecanismo epigenético pode desempenhar um papel crucial em respostas importantes em estudos de insetos (GLASTAD et al., 2011; GLASTAD; HUNT; GOODISMAN, 2019). Em *Drosophila melanogaster*, por exemplo, variações nos níveis genômicos de metilação em diferentes estágios do ciclo de vida podem desempenhar um papel significativo nos processos regulatórios durante o desenvolvimento (DESHMUKH et al., 2018, 2020).

A maioria dos estudos de metilação de DNA para eucariotos envolvem apenas dinucleotídeos CG (ilhas de CpG), em especial para metilação 5-metil citosina (5-mC) (SCHMITZ; LEWIS; GOLL, 2019). Entretanto, na última década foi demonstrado que metilação N6-metil adenina (6mA) para DNA desempenha funções cada vez mais importantes em eucariotos, como na regulação da transcrição gênica, replicação e reparo do DNA e atividades transponíveis (WANG et al., 2018; ZHANG et al., 2018), desempenhando papéis inclusive na tolerância ao estresse ambiental e no desenvolvimento embrionário, sendo este último observado até em *Drosophila* (LI et al., 2022; ZHANG et al., 2015). Estudos em *Drosophila* encontraram, mesmo que em baixos níveis, marcas de metilação 6mA e uma das explicações para estes valores em tecidos de *D. melanogaster* adultas é potencialmente decorrente da presença da proteína DMAD, que regula um processo de desmetilação durante o desenvolvimento embrionário (ZHANG et al., 2015), no entanto, estes mecanismos de regulação precisam ser ainda melhor investigados pois parecem incertos (BOULET et al., 2023).

Considerando que as condições ambientais podem afetar os padrões de metilação tanto durante o desenvolvimento, como também a posteriori, quando as metilações auxiliam na coordenação da resposta celular a novos estímulos ambientais (GUO et al., 2011), evidências crescentes associam a variação morfológica a marcas epigenéticas. Desse modo, a mosca-das-frutas *Anastrepha fraterculus*, que está presente em uma ampla gama de frutos hospedeiros, pode ser um excelente complexo de espécies para estudos que visam observar a atuação dos fatores epigenéticos em suas populações, em função da amplitude de hospedeiros que utiliza (ARAUJO; ZUCCHI, 2006) e da similaridade entre populações de distintas regiões. Todo o desenvolvimento larval e embrionário destes insetos holometábolos ocorre dentro do fruto hospedeiro, e o mesmo pode exercer uma influência significativa sobre estas moscas, tanto em aspectos comportamentais como morfológicos, por apresentarem características específicas (MCPHERON; SMITH; BERLOCHER, 1988; VELASCO-CUERVO; GALINDO-GONZÁLEZ; TORO-PEREA, 2023).

Diversos estudos indicam que os caracteres morfológicos que são em geral efetivos para a caracterização da diversidade do gênero e identificação da maioria das espécies (ARAUJO; ZUCCHI, 2006; MENGUAL et al., 2017; NORRBOM et al., 2015) são mais plásticos e não é tão efetivo para a separação de algumas espécies no grupo *fraterculus*, variando inclusive dentro de uma mesma população (ARAUJO; ZUCCHI, 2006; LAMPERT et al., 2020; PERRE et al., 2014). Além disso, evidências indicam que os frutos hospedeiros podem estar colaborando para essa variação (GÓMEZ-CENDRA et al., 2016; GUSTANI, 2018), como observado inclusive neste trabalho no qual mostramos que moscas *A. fraterculus* que emergiram de diferentes frutos apresentaram diferenças em seus caracteres morfológicos. Desse modo, esta variabilidade fenotípica em caracteres morfológicos pode estar mediada pela resposta a diferentes hospedeiros, inclusive por processos epigenéticos. Isto ocorre porque esta maquinaria está intimamente relacionada com o ambiente de desenvolvimento das espécies, podendo trazer respostas adaptativas, e até mesmo, contribuir para os processos de especiação.

2.2 MATERIAL E MÉTODOS

2.2.1 Extração de DNA

Um estudo piloto foi feito para testar três protocolos distintos de extração de DNA, em moscas provenientes de coleções do laboratório: extração por Trizol, tiocianato de guanidina e fenol-clorofórmio para avaliar a qualidade e comprimento do DNA, bem como seu rendimento total, fatores importantes para um sequenciamento baseado na plataforma Oxford Nanopore (dados não apresentados). Entre esses, o protocolo de extração por fenol-clorofórmio (KIM et al., 2021) foi o que apresentou os melhores resultados, sendo escolhido para dar continuidade ao trabalho. Tal protocolo foi utilizado de acordo com Kim (2021), com uma pequena modificação durante a etapa de lise celular, que foi feita sob nitrogênio líquido. Para tanto, adicionamos inicialmente 141 µl de tampão de lise aos pools de cabeça em um tubo de 1,5 ml, seguido de nitrogênio líquido. As amostras congeladas eram maceradas levemente com um pistilo, visando evitar quebras das longas moléculas de DNA. Em seguida, acrescentamos 142 µl da solução de lise para completar o volume sugerido no protocolo.

Durante o experimento piloto, realizamos a extração por fenol-clorofórmio separadamente para a cabeça da fêmea adulta, e para o restante do corpo, visando investigar se os padrões de metilação seriam diferentes na cabeça em relação ao resto do corpo. Os resultados do teste piloto indicaram maior presença de atividade epigenética no tecido da cabeça de *A. fraterculus* em relação ao resto do corpo, em concordância com o observado em moscas adultas de *Drosophila* (SHAH; CAO; ELLISON, 2019; TAKEMORI et al., 2007; ZHANG et al., 2015). Desse modo, concentramos nossos esforços na identificação de marcas epigenéticas apenas em tecidos da cabeça. Além disso, esta abordagem minimiza a possibilidade de detecção de metilação devido à potencial contaminação por microflora intestinal ou endossimbiontes, prevenindo falsos positivos no DNA genômico (DESHMUKH et al., 2018). Da mesma forma, tal experimento mostrou que a quantidade de DNA produzida por cabeça seria menor do que a requerida para o preparo de bibliotecas Nanopore, que seria de 1 µg de DNA genômico. Por esse motivo, e para tentar homogeneizar o padrão de metilação por amostra, utilizamos duas réplicas biológicas compostas de pools de 4 cabeças das moscas que emergiram de *E. pyrifera* (uvaia), *E. uniflora* (pitanga), *P. guajava* (goiaba) e *C. xanthocarpa* (guabiroba). Além disso, conduzimos duas réplicas de pools de 2 cabeças para *E. involucrata* (cereja do rio grande), em função de um menor tamanho amostral para esta.

2.2.2 Preparo de biblioteca e sequenciamento

O sequenciamento de DNA genômico das amostras provenientes de cinco hospedeiros diferentes (Tabela 2) foi feito utilizando o Kit *Native barcoding genomic DNA (with EXP – NBD104)* da *Oxford Nanopore Biotechnologies*, que permite que até 12 amostras sejam corridas em uma única *flow cell*. Este kit foi associado ao kit de sequenciamento de ligação - SQK-LSK109 (*Oxford Nanopore Biotechnologies*), que é altamente recomendado para análises epigenéticas. Utilizamos o protocolo recomendado pelo fabricante, que é resumido em seguida. O procedimento é dividido em várias etapas, sendo que a primeira consiste no reparo e na preparação final do DNA para a fixação dos adaptadores (*DNA repair end-prep*) e para isso foi utilizado o kit *NEBNext End Repair/dA-tailing* (New England Biolabs), que corrige o DNA para eventuais quebras nas fitas duplas e cria fragmentos de DNA “*blunt-end*” para a incorporação de barcodes que permitem a combinação de diversas amostras.

A segunda etapa do protocolo de sequenciamento consistiu em ligar os barcodes fornecidos no kit às extremidades do DNA (*Native barcode ligation*), e para isto foram utilizados os barcodes demonstrados na Tabela 2. A terceira etapa compreendeu a ligação dos adaptadores de sequenciamento às extremidades de DNA (*Adapter ligation and clean up*), que é essencial para o sequenciamento no dispositivo MinION e resulta da combinação do protocolo *Native barcoding* com o *ligation Kit*. Nesta etapa utilizamos o tampão de lavagem SFB (*Short Fragment Buffer*), a fim de reter fragmentos de todos os tamanhos. Em seguida iniciamos a preparação da lâmina de fluxo e as bibliotecas de DNA foram carregadas no sequenciador para dar início ao sequenciamento (*Priming and loading the SpotON flow cell*). Por fim foi realizada a aquisição dos dados brutos para eventual *basecalling* e processamento de dados.

Tabela 2. Identificação dos barcodes utilizados para cada amostra sequenciada.

Amostra	Fruto	Biblioteca	Barcode	Sequência
CeP1	Cereja	1	1	CACAAAGACACCGACAACCTTCTT
UvP1	Uvaia	1	4	TAGGGAAACACGATAGAATCCGAA
PiP1	Pitanga	1	8	ACGTAACCTGGTTTGTTCCTGAA
GoP1	Goiaba	1	9	AACCAAGACTCGCTGTGCCTAGTT
GbP1	Guabiroba	1	12	TCCGATTCTGCTTCTTTCTACCTG
CeP2	Cereja	2	2	ACAGACGACTACAAACGGAATCGA

UvP2	Uvaia	2	6	GACTACTTTCTGCCTTTGCGAGAA
PiP2	Pitanga	2	7	AAGGATTCATTCCCACGGTAACAC
GoP2	Goiaba	2	10	GAGAGGACAAAGGTTTCAACGCTT
GbP2	Guabiroba	2	11	TCCATTCCCTCCGATAGATGAAAC

As amostras são referentes a pools de cabeça para cada fruto hospedeiro utilizado. As duas primeiras letras da coluna de amostras são referentes as siglas dos frutos, enquanto P1 ou P2 é referente a biblioteca final do sequenciamento.

Ao longo do experimento, realizamos duas modificações no protocolo original. Utilizamos beads da Pronex em substituição às AMPure XP beads, em função do custo mais baixo. Além disso, as etapas de incubação para ressuspensão do DNA com *nuclease-free water* foram conduzidas por 30 minutos a 37 °C e no intuito de aumentar a eficiência da reação de ligação, estas foram feitas *overnight* a 4°C.

2.2.3 Processamento de dados

Os dados brutos do sequenciamento (arquivos fast5) foram capturados inicialmente pelo software MinKNOW, responsável por controlar o dispositivo de sequenciamento de nanoporos e coletar os dados brutos em tempo real. Esta etapa foi realizada em um servidor DELL PowerEdge T620 presente no laboratório, sendo essencial para garantir a adequada captura de dados durante o sequenciamento. Em seguida iniciamos o processo de *basecalling*, que consiste na tradução do sinal bruto do sequenciador em bases (ATCG). Para isto, utilizamos o software Guppy v.6.4.2 no módulo de alta acurácia (HAC) com a função “guppy_basecaller”.

Uma vez que a chamada de bases não foi realizada em tempo real com o software Guppy integrado ao MinKNOW, a segregação das *reads* por barcode e a subsequente remoção dos adaptadores e barcodes foi realizada separadamente por meio da função “guppy_barcode” e os comandos “--trim_adapters” e “--trim_barcodes”. Para executar este procedimento, foram necessários os arquivos em formato “fastq” gerados pelo “guppy_basecaller” na etapa anterior. Este processo gerou um arquivo de resumo do sequenciamento chamado *sequencing_summary.txt* o qual possibilitou a criação de gráficos que nos auxiliaram a observar os dados do sequenciamento (Figura 2) no software *Nanoplot v.1.40.0* (DE COSTER et al., 2018).

Softwares de análise e chamada de metilação requerem arquivos fast5 *single* como *input*. Como nossos arquivos estavam em formato *multi*, realizamos a conversão desses

arquivos. Para esse procedimento, utilizamos o comando `demux_fast5`, disponível no software `ont_fast5_api`, a fim de demultiplexar os arquivos brutos do sequenciamento (fast5) e separá-los pelos barcodes utilizados. Nesta análise, foi utilizado o arquivo “`barcoding_summary.txt`” resultante da demultiplexação dos arquivos “fastq” realizados pelo `guppy_barcode`. Em seguida, empregamos o comando “`multi_to_single_fast5`”, também do `ont_fast5_api`, para transformar os fast5 *multi* em fast5s (*single fast5*).

O passo seguinte foi realizar o *basecalling* das fast5s pertencentes a cada amostra utilizando o software Guppy v.6.4.2 com o comando “`guppy_basecaller`”. Esse processo não apenas traduz os sinais elétricos em bases, mas também filtra as *reads* de baixa qualidade. Devido à exigência de alto poder de computação e processamento, a análise foi realizada em uma ferramenta do Google online, também conhecida como *Google Colaboratory* (Google Colab), uma plataforma online que fornece acesso a recursos de computação como GPUs. Em seguida reunimos os arquivos “fastq” em um único arquivo usando “`cat /pass/*.fastq > fast5s_guppy.fastq`” para cada amostra. Além disso, os dados da cobertura do sequenciamento foram realizados pelo programa *Coverage_calculator* v.0.0.1.

2.2.4 Análises epigenéticas

2.2.4.1 Metilação 5mC

Após o processamento dos dados brutos, iniciamos as análises de metilação, visando identificar regiões diferencialmente metiladas (DMRs) entre moscas que emergiram de frutos diferentes. Inicialmente, a chamada de metilação 5mC foi realizada pelo software *Nanopolish* v.0.13.2, independentemente para cada amostra e réplica. Este processo teve início com a indexação dos arquivos de leitura (fastq) com os dados de sinal bruto do nanopore (fast5s) por meio do módulo “*nanopolish index*”. Posteriormente, foi realizado o alinhamento dos arquivos de chamada de base ao genoma de referência utilizando o software *Minimap2* v2.17, seguido pela indexação deste arquivo com o comando “`samtools index`”. Para essa etapa, utilizamos o genoma de *Anastrepha ludens* disponível no NCBI (GCF_028408465.1), uma vez que esta espécie é suficientemente próxima de *A. fraterculus* e a versão do genoma de *Anastrepha fraterculus* que dispomos não apresenta a mesma qualidade. O próximo passo foi realizar a chamada de metilação 5-methylcytosine em ilhas CpG, e a frequência com que cada posição foi calculada com

as funções “*nanopolish call-methylation*” e “*scripts/calculate_methylation_frequency.py*” respectivamente, do software *Nanopolish*.

Para a análise dos arquivos de chamada de metilação, aplicamos o *pipeline* do software *pycoMeth* v. 2.2.2, visto que este programa propõe algoritmos para análise de metilação diferencial eficientes para multiamostras e haplótipos (Snajder *et al.*, 2023). O *pipeline* é empregado para realizar análises adicionais a partir dos arquivos de saída gerados pelo “*nanopolish call-methylation*”. No entanto, para a versão mais recente do software, os dados de entrada precisaram ser convertidos para o formato Meth5 (.m5).

A conversão do formato da chamada de metilação foi realizada por meio do programa *MetH5Format* v.1.2.0 com a função “*meth5 create_m5*”. Os arquivos resultantes armazenam a razão log-verossimilhança de cada chamada de metilação, as coordenadas genômicas (início e fim), e o nome da leitura associada a cada chamada. Posteriormente, utilizamos o comando “*--input_m5_files --read_group_names --no_read_names --read_groups_key --output_file*” para juntar os arquivos .m5 correspondente a cada fruto, considerando as réplicas como uma só amostra.

A etapa subsequente envolveu a identificação dos cromossomos que continham chamadas de metilação. Para isto, empregamos o subcomando “*meth5 list_chunks*” também do programa *MetH5Format* v.1.2.0. Essa informação foi essencial para realizar uma análise específica para cada cromossomo que continha dados relevantes de metilação. Essa análise foi realizada por meio do subcomando “*Meth_seg*” do programa *pycoMeth* v.2.2.2. A etapa de segmentação refere-se a uma análise bayesiana, que visa identificar segmentos com base exclusivamente nas informações de metilação.

Os arquivos gerados para cada cromossomo, por meio da segmentação, foram concatenados e utilizados na etapa subsequente. Em seguida, o comando “*Meth_Comp*” do *pycoMeth* foi empregado para a análise de metilação diferencial, realizando um teste de segmentos para identificar regiões com metilação diferencial entre duas ou mais amostras. Nesta análise foi definido que a cobertura mínima seria de 10 leituras para que um intervalo fosse considerado; a razão mínima de log-verossimilhança absoluta de metilação é 2 com valores de p corrigidos para testes múltiplos por Benjamini/Hochberg (*fdr_bh*) e a hipótese testada foi a *llr_diff*, que realiza um teste não paramétrico para a mudança na razão log-verossimilhança de metilação.

Foi aplicado também o subcomando “*Comp_report*” para gerar um relatório para regiões diferencialmente metiladas significativas encontradas com o *pycoMeth*

Meth_comp, bem como relatórios individuais para as principais regiões, ou seja, a produção de informações mais detalhadas que se concentram em regiões específicas consideradas relevantes. Para isto foi considerado que a diferença mínima na taxa de metilação entre as amostras relatadas era 0,5 e a distância máxima de um intervalo do TSS (locais de início da transcrição) de um gene a ser relatado foi de 100 kb.

Por fim, as regiões diferencialmente metiladas mais significativas e com altos picos de metilação foram associadas a termos do *gene ontology* (GO) por meio do software *Blast2GO*, a fim de realizar uma análise de enriquecimento de conjuntos de genes. Desse modo, identificamos os termos GO enriquecidos para estes conjuntos de genes por meio da ferramenta *ShinyGO v.0.77*, com correção de testes múltiplos por FDR (False Discovery Rate) de 0,05.

2.2.4.2 Metilação 6mA

Para a identificação de metilação N6-metil adenina (6mA), utilizamos o *pipeline* do programa *mCaller v.1.0*. A primeira parte do processo envolveu a indexação dos arquivos de leitura (fastq) contendo os dados de sinal bruto do nanopore (fast5s) por meio do módulo "*nanopolish index*", como já feito na análise anterior. Em seguida, foi empregado o subcomando "*nanopolish eventalign*" do software *Nanopolish v.0.13.2* a fim de alinhar os eventos ou "*squiggles*" ao genoma de referência. Esse procedimento ajuda a descobrir diferenças na corrente que podem sugerir modificações de base.

Em seguida, foi executado o programa *mCaller v.1.0* para detectar modificações 6mA em *motifs* GATC. Esta função gerou um arquivo contendo previsões por leitura, incluindo informações como nome do cromossomo, nome da leitura, posição genômica, contexto da posição k-mer e probabilidade de metilação prevista para esta posição e leitura. Além disso, geramos um arquivo bed, por meio do subcomando do *mCaller* (*make_bed.py*), com previsões de metilação com cobertura mínima de 5 vezes e com uma fração mínima de observações de uma posição indicada como metilada, com uma probabilidade maior ou igual a 50% (*--mod_threshold 0,5*).

A relação entre frequências de metilação de m6A em sítios ao longo do genoma que apresentaram valores acima de 50% de amostras de frutos diferentes foram contrastados por um teste estatístico não-paramétrico Kruskal-Wallis. Para esse teste, duas análises distintas foram feitas. No primeiro caso, foi assumido que valores de

metilação abaixo de 50% foram considerados como sendo indicativos de ausência de metilação e o fruto hospedeiro foi usado como variável explanatória. Outro teste mais restrito contrastou apenas os sítios no genoma para os quais existia pelo menos uma amostra com valores acima de 50% de metilação. Para ambos os casos, valores de significância do teste não paramétrico de Kruskal-Wallis foi corrigido por testes múltiplos por FDR e significância de 0,05.

2.3 RESULTADOS

2.3.1 Sequenciamento *Nanopore* para análises de metilação

O procedimento de sequenciamento aplicado à primeira biblioteca (Biblioteca 1) foi conduzido ao longo de aproximadamente 50 horas, resultando em uma estimativa de 3.61 Gb e 934.85 mil reads. Este processo proporcionou uma cobertura genômica de aproximadamente 2x para cada conjunto de amostras, sendo que as primeiras 30 horas representaram cerca de 99% da produção total. Contudo, um declínio significativo no rendimento foi observado após 20 horas de sequenciamento embora aparentemente ainda havia disponibilidade imediata de poros para sequenciamento que não estavam sendo utilizados. Isso possivelmente foi causado pelo consumo da biblioteca na *flow cell*. Em resposta a essa situação, mais amostras da mesma biblioteca foram introduzidas no sistema com o intuito de otimizar o rendimento do sequenciamento. Contudo, no intervalo de 20 a 30 horas, foi observada uma queda abrupta de poros na *flow cell* (Figura S1a) e a partir de então a curva de produção de dados se estabilizou, quando optamos por encerrar o sequenciamento (Figura 4a).

Para a segunda biblioteca, o sequenciamento teve duração de aproximadamente 40 horas, resultando em aproximadamente 5.84 Gb e 1.48 M reads geradas. Neste caso, a cobertura foi de 2 a 3x do genoma completo para cada conjunto de amostras, sendo as primeiras 24 horas responsáveis pela maior parte da produção total do sequenciamento (Figura 4b). A perda abrupta dos poros (Figura S1b), assim como no sequenciamento da primeira biblioteca fez com que a curva de produção de dados se estabilizasse a partir de 30 horas de sequenciamento como podemos observar na Figura 4b. Em função desses parâmetros, buscamos estratégias adicionais para otimizar o desempenho do sequenciamento, visando garantir uma eficiência consistente ao longo de todo o

procedimento. Os dados do sequenciamento podem ser observados no material suplementar (Figura S2).

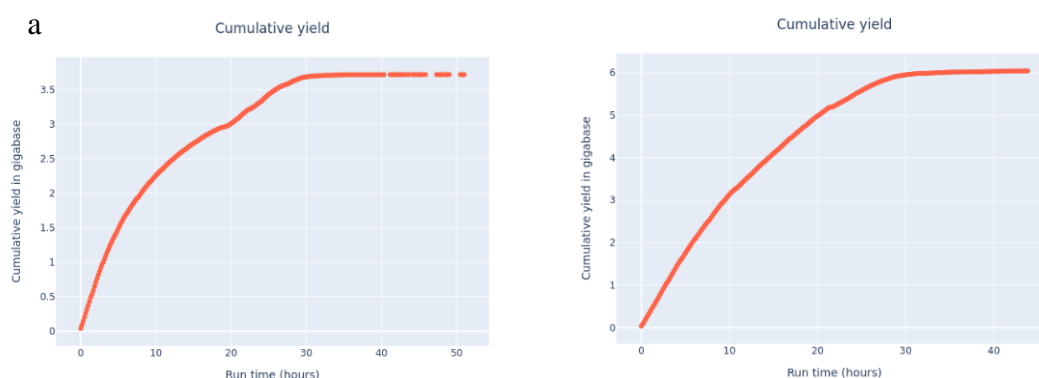


Figura 4. Resultado cumulativo do sequenciamento de bibliotecas Oxford Nanopore de DNA genômico em um MinION ao longo do tempo; (a) Biblioteca 1 (b) Biblioteca 2.

Ao longo do processo de sequenciamento, alguns poros classificados inicialmente como disponíveis podem apresentar um *status* inativo ou de indisponibilidade. Essa condição pode ocorrer devido a eventos como entupimento dos poros, ou por um processo natural de degradação durante o sequenciamento. Na Figura S3 a coloração amarela representa justamente estes poros inativos ao longo do sequenciamento. Vale ressaltar que uma *flow cell* nova apresenta 2048 poros distribuídos em 512 canais (4 para cada canal) (Oxford Nanopore Biotechnologies), e as utilizadas em nosso sequenciamento continham cerca de 1500 poros ativos para a primeira biblioteca e 1200 poros para a segunda. Além disso, uma análise da qualidade das *reads* geradas pelo sequenciamento também foi realizada pelo Nanoplot (Tabela 3), o que permitiu visualizar os índices de qualidade das leituras e o valor do N50, métrica definida pelo tamanho do *contig* mais curto para o qual *contigs* mais longos e de comprimento igual cobrem pelo menos 50% do genoma (ALHAKAMI; MIREBRAHIM; LONARDI, 2017).

Tabela 3 Dados de qualidade do sequenciamento de DNA genômico por meio de bibliotecas Oxford Nanopore em um MinION.

	Biblioteca 1	Biblioteca 2
Comprimento médio de leitura	1.664	4.075,4
Qualidade média de leitura	8,0	7,5
Comprimento mediano de leitura	1.664	1.348
Qualidade de leitura mediana	9,6	8,8

Número de leituras	934.850	1.483.794
Comprimento de leitura N50	9.489	11972
Bases totais	3.721.625.575	6.047.084.840
Número, porcentagem e megabases de leituras acima dos limites de qualidade		
>Q5	878931 (94.0%) 3552.9Mb	1370818 (92.4%) 5688.3Mb
>Q7	726710 (77.7%) 3007.3Mb	1057151 (71.2%) 4477.4Mb
>Q10	414018 (44.3%) 1711.1Mb	518708 (35.0%) 2196.0Mb
>Q12	133435 (14.3%) 555.1Mb	126646 (8.5%) 473.5Mb
>Q15	612 (0.1%) 0.8Mb	604 .0%) 0.5Mb

2.3.2 Metilação de ilhas CpG no DNA e Regiões Diferencialmente Metiladas (DMRs)

A chamada de metilação, executada no programa *Nanopolish* (SIMPSON et al., 2017), indicou mais de 1 milhão de regiões CpG por amostra. Tal software estima a razão logarítmica calculada (*log_likelihood_ratio*) para cada posição CG no genoma de referência. Quando o *log_likelihood_ratio* é positivo e acima de 2, há suporte estatístico para uma região com metilação, enquanto valores negativos abaixo de 2 indicam uma região CpG não metilada (SNAJDER et al., 2022; SIMPSON, 2019). Valores de *log likelihood ratio* entre -2 e 2 indicam chamada de metilação ambígua. Esses dados foram posteriormente utilizados na análise de regiões diferencialmente metiladas por meio do software *pycoMeth* (SNAJDER et al., 2022).

Uma análise de metilação diferencial entre as amostras que emergiram de diferentes frutos pelo software *pycoMeth* resultou em 60.733 regiões densas em CpG e 16.594 diferencialmente metiladas. Dentre estes intervalos, considerando *log-likelihood ratio* (LLR) maior que 1, 67 são considerados com altos níveis de metilação para os indivíduos de cereja do rio grande; 92 para pitanga; 151 para uvaia; 173 para goiaba e 175 para guabiroba, como observado graficamente na Figura 5.

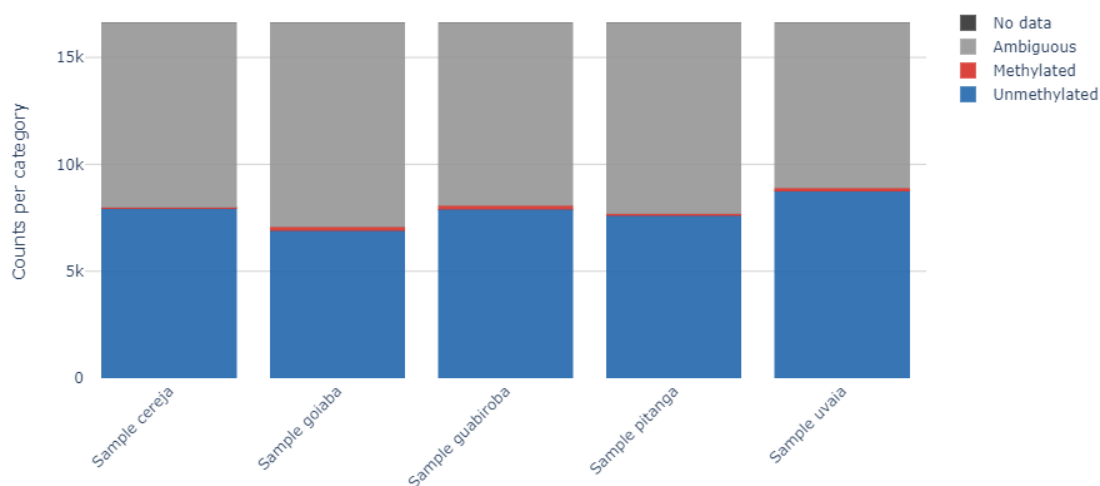


Figura 5. Contagens de categorias de metilação por amostra para intervalos significativos de CpG inferidos por *pycoMeth* de resultados de sequenciamento por Oxford Nanopore em um MinION.

Os resultados do *pycoMeth* mostram uma grande quantidade de intervalos categorizados como ambíguos, apresentando valores de *Log-Likelihood Ratio* (LLR) entre -1 a 1 que não podem ser claramente classificados como metilados ou não metilados, conforme indicado por Leger et al. (2022) (Figuras 5 e S4).

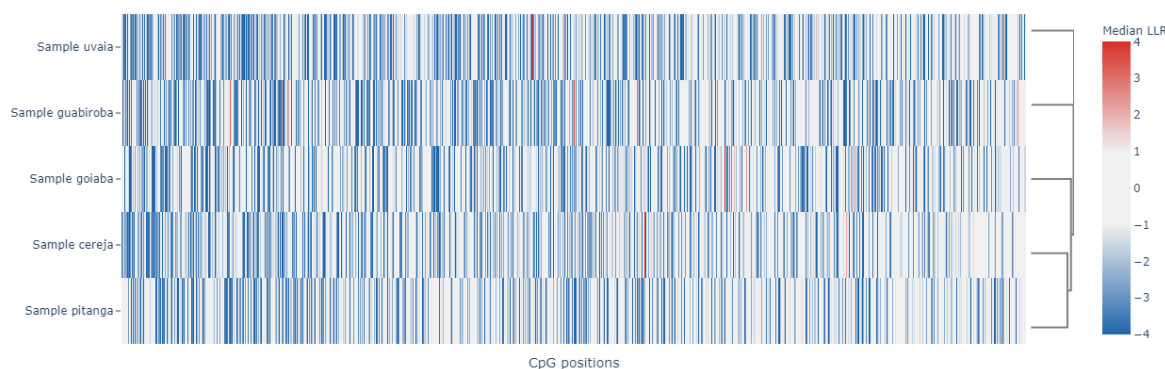


Figura 6. Razão de log-verossimilhança de metilação por intervalo CpG significativo.

Além disso, observamos um enriquecimento acentuado de intervalos com DMRs em ilhas de CpG significativas perto dos locais de início da transcrição como observado na Figura 7. A distribuição espacial desses intervalos CpG nos segmentos cromossômicos, bem como a lista com os 100 principais intervalos diferencialmente metilados significativos podem ser visualizados no material suplementar (Figura S5, Tabela S3). Além disso, uma distribuição ao longo do genoma dessas regiões pode ser observada na Figura S6(a).

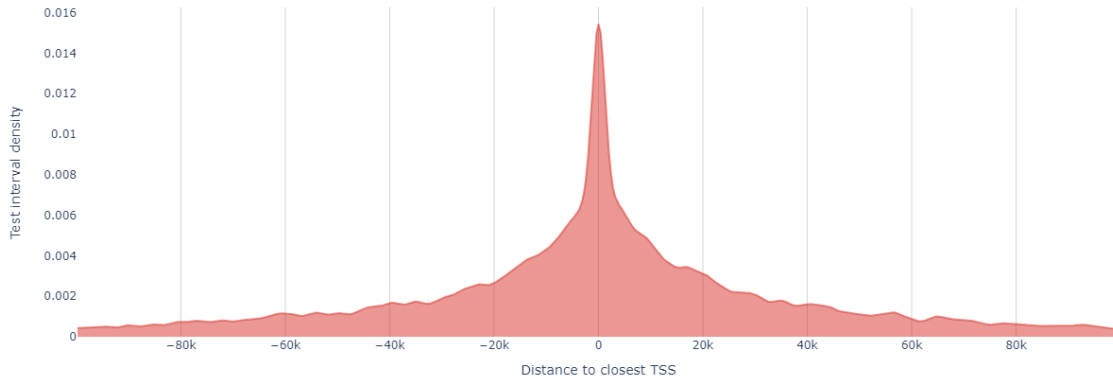


Figura 7. Distância do intervalo CpG ao gene TSS mais próximo.

Nosso conjunto de DMRs se sobrepõe a um total de 3500 genes, dos quais são potencialmente relevantes para o estudo da diferenciação de *A. fraterculus*, bem como sua adaptação a diferentes hospedeiros. Tal análise demanda estudos mais aprofundados de efeito e função. O conjunto de genes com DMR entre moscas emergidas de frutos distintos é muito grande e demanda cuidadosa exploração, mas uma visão global pode ser obtida por uma análise de enriquecimento de ontologia genética (GO). Neste conjunto de genes, identificamos um total de 560 termos significativamente enriquecidos com o valor de FDR $\leq 0,05$. Dentre eles, os 20 principais e mais significativos são pertencentes a ontologia de processos biológicos (Tabela S4) e podem ser observados na Figura 8, bem como as relações entre as vias enriquecidas encontradas (Figuras 9 e S7).

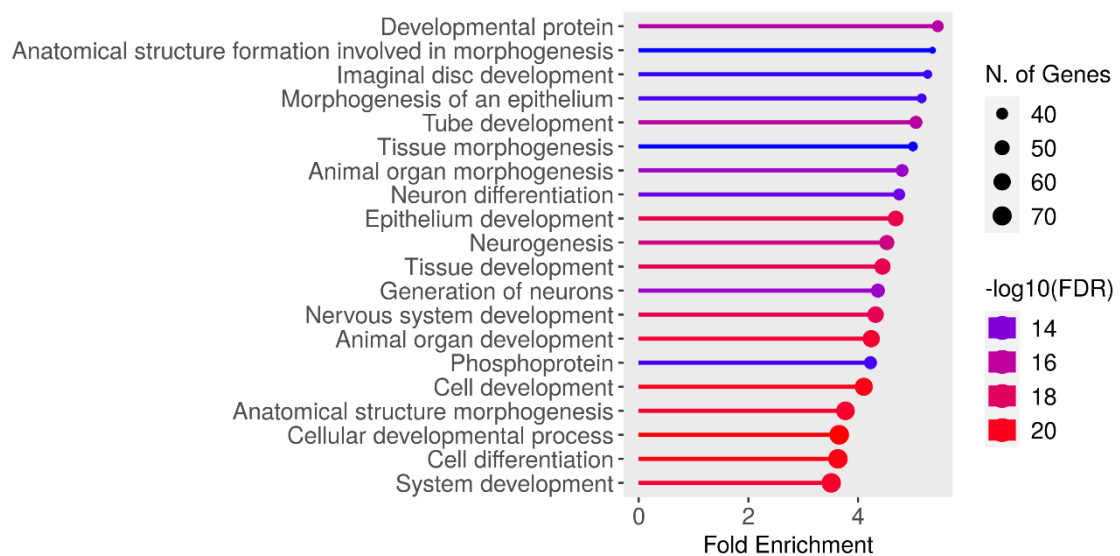


Figura 8. Análise de gene ontology (GO) mostrando os 20 principais termos GO por ordem de fold enrichment.

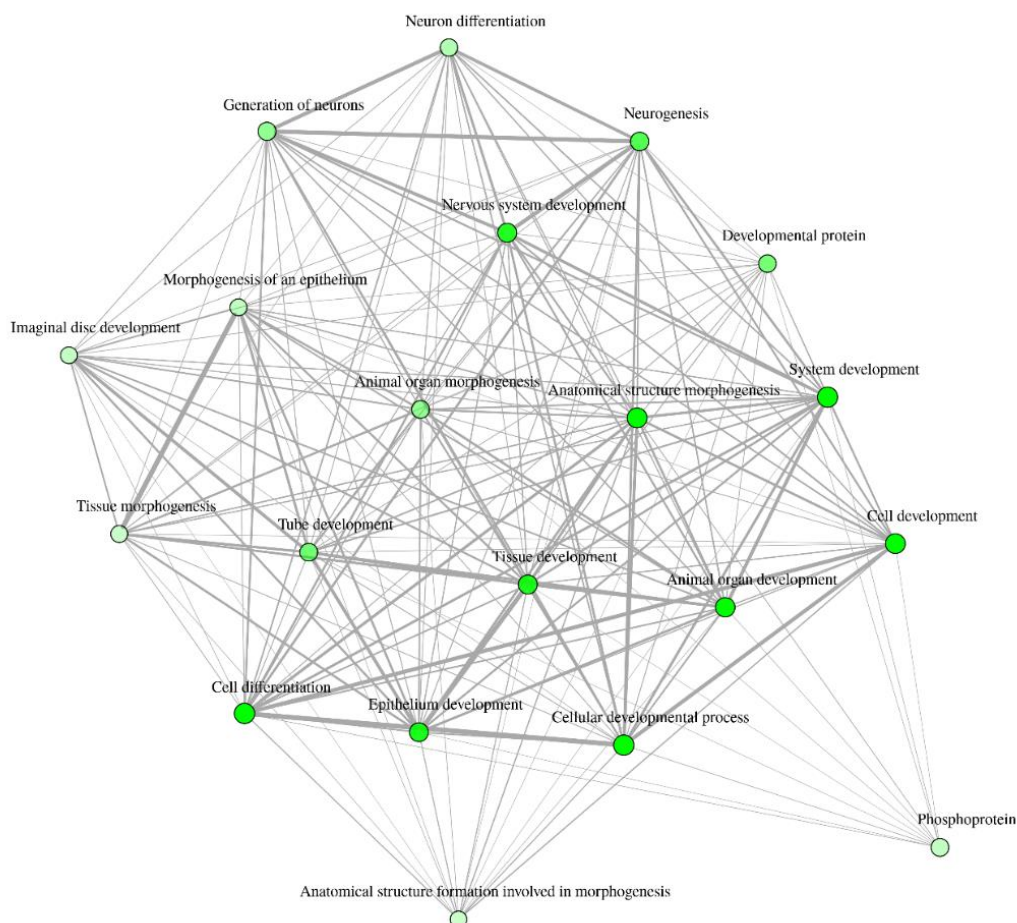


Figura 9 Análise de gene ontology (GO) mostrando as relações entre os 20 principais termos GO enriquecidos. Quando duas vias (ou nós) são conectadas é devido compartilharem pelo menos 20% genes, linhas mais grossas significam maiores sobreposições de genes, enquanto nós mais escuros representam conjuntos de genes mais informativos significativamente e nós maiores representam conjuntos de genes maiores.

As análises do *gene ontology* revelaram que a maioria dos genes estão relacionados com processos biológicos como, por exemplo, desenvolvimento e diferenciação celular, morfogênese da estrutura anatômica; desenvolvimento do epitélio e desenvolvimento do sistema nervoso (Tabela S4). A quantidade de genes envolvidos nos 20 principais termos enriquecidos variou entre 35 e 73 genes, sendo a via de processos de desenvolvimento celular com 73 e compartilha grande parte dos genes com o desenvolvimento de epitélio, diferenciação celular e desenvolvimento de tecidos (Figura 8). Além disso, a via do desenvolvimento do sistema nervoso apresenta cerca de 50 genes, e grande parte do mesmo se encontra com genes sobrepostos a enriquecimento para a neurogênese e geração de neurônios (Figura 9).

2.3.3 Metilação 6mA

A análise de metilação 6mA até o momento é exploratória com os resultados da chamada de metilação com o software mCaller para cada pool de amostras realizadas nas duas bibliotecas sequenciadas, como podemos observar na Tabela 4.

Tabela 4. Número de regiões com múltiplas bases 6mA metiladas. Número de loci metilados com cobertura maior ou igual a 5X para cada pool de amostra de diferentes frutos.

Amostras	Fruto	Biblioteca	#Regiões	≥ 5 reads
CeP1	Cereja	1	14.953	13.121
CeP2	Cereja	2	11.595	8.536
UvP1	Uvaia	1	2639	618
UvP2	Uvaia	2	12.327	8.150
PiP1	Pitanga	1	11.011	8.496
PiP2	Pitanga	2	6.141	2.971
GoP1	Goiaba	1	3.610	1.190
GoP2	Goiaba	2	5.661	2.504
GbP1	Guabiroba	1	2.616	442
GbP2	Guabiroba	2	8.568	5.118

As amostras são referentes a pools de cabeça para cada fruto hospedeiro utilizado. As duas primeiras letras da coluna de amostras são referentes às siglas dos frutos, enquanto que P1 ou P2 é referente a biblioteca final do sequenciamento.

Dentre esses loci, 23.758 apresentaram evidências de metilação nas amostras como um todo, enquanto, 4570 foram encontradas em pelo menos 50% das amostras e 802 estavam presentes em ao menos 80% das amostras. Além disso, a relação entre frequências de metilação de 6mA por meio do teste não paramétrico de Kruskal-Wallis indicou a presença de 136 regiões diferencialmente metiladas com valores significativos (FDR = 0,05) entre os cinco frutos analisados (Tabela Suplementar S5), a maioria das quais associadas a regiões não codificadoras de DNA. Uma distribuição ao longo do genoma dessas regiões pode ser observada na Figura S6 (b).

2.4 DISCUSSÃO

Nossos resultados mostram que a estratégia de sequenciamento por Oxford Nanopore Technologies em um MinION permitiu a identificação de sítios metilados 5mC e m6A no genoma de *Anastrepha fraterculus*, registrando pela primeira vez a presença de sítios diferencialmente metilados (DMRs) nessa espécie. Assim como descrito para *Drosophila* (DESHMUKH et al., 2018), modificações 5mC apresentaram frequências mais reduzidas ao longo do genoma. Os níveis de metilação em insetos são, em geral, mais baixos quando comparados aos de genomas de vertebrados (SCHMITZ; LEWIS; GOLL, 2019). No entanto, mesmo em níveis reduzidos, metilações em regiões GC tem se mostrado importantes para estes táxons que frequentemente exibem plasticidade fenotípica, podendo influenciar na regulação da expressão gênica e conservar os recursos genéticos ao gerenciar a conformação funcional do genoma (DESHMUKH et al., 2018; ZEE et al., 2016).

Pesquisas realizadas em *Drosophila* revelaram uma diferença significativa nos padrões de metilação do DNA ao longo dos diferentes estágios do ciclo de vida dessas moscas, sendo identificado uma maior incidência dessa modificação epigenética em estágios embrionários quando pode desempenhar um papel no desenvolvimento (DESHMUKH et al., 2018; ZHANG et al., 2015). Contudo, mesmo em baixos níveis, metilação de tecidos em adultos, predominantemente em tecidos da cabeça, podem desempenhar papel importante na regulação gênica como, por exemplo, na fotorrecepção (SHAH; CAO; ELLISON, 2019; TAKEMORI et al., 2007; YAO et al., 2018; ZHANG et al., 2015).

A comparação do padrão de metilação do sequenciamento de cabeças de moscas-das-frutas emergidas de cinco frutos hospedeiros distintos indicou a existência de regiões diferencialmente metiladas (DMRs), inferidas por meio do software pycoMeth para regiões 5mC utilizando testes não paramétricos, bem como uma análise de Kruskal-Wallis para regiões com metilação 6mA. Estas regiões podem ser definidas como regiões genômicas que possuem distintos padrões de metilação de DNA entre diferentes amostras (CHEN; LIN; FANN, 2016). Considerando que a distribuição da metilação de DNA no genoma segue padrões locais robustos, a metilação diferencial entre as regiões genômicas é considerada informativa tanto estatisticamente quanto biologicamente (MAGI et al., 2023).

As DMRs desempenham um papel importante em diversos mecanismos biológicos e podem estar envolvidas, inclusive, na coordenação da resposta celular a estímulos ambientais (AKBARI et al., 2022; GUO et al., 2011). Em insetos, a presença de metilação diferencial frequentemente está associada a organismos eussociais, envolvendo a plasticidade comportamental, o comportamento social e o forrageamento, como visto em estudos realizados com Hymenoptera (BEWICK et al., 2016; YAN et al., 2015). Em formigas, por exemplo, quando o gene promotor de crescimento (*egfr*) é metilado e expresso em larvas influencia no tamanho das operárias, e neste caso, é notório que a divisão de trabalho na colônia de formigas depende do tamanho das mesmas (RENARD; GUEYDAN; ARON, 2022).

Já em abelhas *Apis mellifera*, regiões diferencialmente metiladas indicaram que o desenvolvimento de uma larva em rainha ou operária envolve modificações epigenéticas (FORET et al., 2012). Além disso, estudos revelam que a complexa arquitetura epigenética envolvendo o cérebro de *Drosophila* pode estar intrinsecamente relacionada com a regulação do comportamento alimentar destas moscas (ANREITER; KRAMER; SOKOLOWSKI, 2017). Nesse contexto, a identificação dos conjuntos de genes que estão envolvidos nas DMRs significativas encontradas neste trabalho, bem como sua possível função, pode trazer importantes subsídios para a compreensão da biologia da mosca-das-frutas *A. fraterculus*, que é conhecida por ser uma importante praga agrícola associada a uma grande diversidade de hospedeiros.

Uma limitação do estudo aqui realizado foi o fato de termos utilizado pools de indivíduos para detecção de metilação. Infelizmente, não temos ainda software adequado para a análise adequada desses dados, uma vez que o pycoMeth assume que são estudados indivíduos e portanto a metilação é considerada como estando em três estados apenas: as duas cópias metiladas, as duas não metiladas, ou heterozigota. Por isso encontramos uma grande quantidade de resultados ambíguos para a análise de metilação, que sugere que a metilação está presente nas populações, mas não é completamente fixada. Devemos buscar estratégias adequadas para a correta análise desses dados, mas os resultados até aqui obtidos claramente indicam a presença de metilação presente na população, e distinta entre moscas emergidas de frutos distintos. Pensando em compreender as diversas funções biológicas que os conjuntos gênicos podem exercer, uma análise de enriquecimento foi empregada a uma lista de genes significativos, a fim de compreender melhor a biologia da espécie *A. fraterculus*, bem como, sua função.

A análise de enriquecimento de termos GO da rede de genes vinculados às regiões diferencialmente metiladas em sítios 5mC em moscas-das-frutas *A. fraterculus*, conforme apresentado neste estudo, proporcionou uma visão abrangente das múltiplas funções biológicas desempenhadas por esses genes. Destaca-se, entre essas funções, a morfogênese de estruturas anatômicas, evidenciada tanto pela presença de um considerável número de genes (67) relacionados a esse processo nas DMRs, quanto pela sua notável significância estatística (FDR 2.8E-20). Essa via desempenha um papel crucial na geração e organização das estruturas anatômicas, revelando insights valiosos sobre os processos subjacentes.

Este resultado é interessante pois vários estudos apontam que os caracteres morfológicos de moscas-das-frutas possuem variabilidade fenotípica, inclusive entre frutos hospedeiros diferentes em uma mesma região geográfica (ARAUJO; ZUCCHI, 2006; FOOTE, 1980; GÓMEZ-CENDRA et al., 2016). Esta variação morfológica pode estar relacionada a mecanismos de plasticidade fenotípica mediada por processos epigenéticos em resposta a diferentes hospedeiros. Ainda que seja uma hipótese audaciosa, a maquinária epigenética está frequentemente ligada a plasticidade tanto morfológica como comportamental e fisiológica (DUNCAN; CUNNINGHAM; DEARDEN, 2022; GLASTAD; HUNT; GOODISMAN, 2019; MANFREDINI; ARBETMAN; TOTH, 2019; WEN; UENO, 2021; WILLIAMS; SOKOLOWSKI, 2018), e não existem estudos que associem marcas epigenéticas a variação fenotípica de moscas-das-frutas.

A via enriquecida de morfogênese de estruturas anatômicas possui genes sobrepostos a diversas funções relacionadas ao sistema nervoso (Figuras 9 e S6). O desenvolvimento do sistema nervoso, neurogênese, geração de células no sistema nervoso e a diferenciação de neurônios foram outras vias importantes que estão associadas ao conjunto de genes encontrados neste trabalho. Esses achados fortalecem a hipótese proposta de que a metilação do DNA desempenha um papel crucial na regulação de funções cerebrais (XIE et al., 2023) em organismos em geral.

Interessantemente, em *Drosophila* genes metilados estão constantemente associados a um conjunto rigoroso de processos biológicos ligados à funções cerebrais, como na formação de memória (DESHMUKH et al., 2020). Além disso, neste mesmo táxon, mesmo em contextos em que a metilação ocorre fora do padrão 5mC, evidências sugerem funções específicas em processos como a fotorrecepção e o forrageamento dessas

moscas (ANREITER; KRAMER; SOKOLOWSKI, 2017; TAKEMORI et al., 2007), bem como funções em processos biológicos envolvidos com o neurônio para a metilação 6mA (SHAH; CAO; ELLISON, 2019; YAO et al., 2018; ZHANG et al., 2015). Tornando essencial estudos direcionados para este tecido na espécie *A. fraterculus*.

Nossos resultados para modificações epigenéticas em 6mA, embora de natureza exploratória até o momento, revelaram a presença de numerosos sítios metilados em amostras provenientes de diferentes frutos. A diferença no número de regiões encontradas, principalmente entre pools de amostras do mesmo fruto, deve-se não apenas à variabilidade intrínseca entre elas, mas também devido aos resultados do sequenciamento, que mostraram uma quantidade inferior de *reads* para as amostras da primeira biblioteca. Apesar dessa consideração, destaca-se a relevância dos resultados iniciais relacionados a essa modificação epigenética em *A. fraterculus*, uma vez que, cada vez mais há evidências indicando que esta modificação desempenha um papel significativo na regulação da expressão gênica em diversos processos biológicos a metilação 6mA (SHAH; CAO; ELLISON, 2019; YAO et al., 2018; ZHANG et al., 2015), por mais que os mecanismos que controlam esta marca epigenética ainda parecerem incertos (BOULET et al., 2023).

A análise de regiões com metilação 6mA permitiu encontrar através de um teste não paramétrico de Kruskal-Wallis mais de 100 regiões diferencialmente metiladas entre moscas que emergiram entre os 5 frutos analisados. Uma análise mais detalhada dessas regiões é importante, mas diversas dessas modificações acontecem em regiões para as quais também foram detectadas mudanças epigenéticas em 5mC. Interessantemente, a maioria das regiões investigadas para m6A está associada a regiões não codificadoras, e isso deve ser melhor explorado.

Considerando esta complexa maquinaria epigenética, evidências indicam que metilações do DNA medeiam as interações gene-ambiente, traduzindo sinais ambientais em mudanças na expressão gênica e alterando o fenótipo sem modificar a sequência de DNA (FEIL; FRAGA, 2012; RENARD; GUEYDAN; ARON, 2022). Acredita-se que respostas adaptativas preditivas possam ser determinadas por meio de marcas epigenéticas estabelecidas por sinais ambientais no início do desenvolvimento, levando alterações no metabolismo, morfologia e comportamento que podem ter efeitos na aptidão à medida que o organismo se desenvolve (DUNCAN; GLUCKMAN; DEARDEN, 2014; VARLEY et al., 2013). Desse modo, é natural que o hospedeiro possa exercer influência nos caracteres

de moscas-das-frutas, uma vez que todo o desenvolvimento de *A. fraterculus* ocorre dentro do fruto.

Além disso, a maquinaria epigenética que regula a expressão gênica pode ser transmitida tanto por meio de divisão celular como entre gerações. Esta última faz com que os fenótipos hereditários associados a este controle epigenético estejam sujeitos a seleção e influenciem na adaptação evolutiva (GLASTAD; HUNT; GOODISMAN, 2019; STAJIC; JANSEN, 2021). A transmissão intergeracional das marcas epigenéticas, apesar de muito discutida, já foi observada em insetos como abelhas, onde foi percebida uma herança estável de marcas epigenéticas entre gerações (YAGOUND et al., 2020) e inclusive em *Drosophila*, onde alterações epigenéticas induzidas por uma dieta paterna rica em açúcar foram transmitidas aos seus descendentes e afetou os fenótipos relacionados com a alimentação na próxima geração (ÖST et al., 2014). Esses exemplos destacam a capacidade das modificações epigenéticas em influenciar não apenas o desenvolvimento individual, mas também a evolução das espécies ao longo do tempo.

Estas informações são relevantes, pois indicam que em insetos, como as moscas-das-frutas, estas marcas epigenéticas não são completamente apagadas ou reprogramadas, ao contrário de alguns estudos de mamíferos (YAGOUND et al., 2020). Além disso, considerando a existência de vários registros que associam a variação epigenética a diversas características em insetos, como o desenvolvimento da asa, regulação de metamorfose e diapausa (DE CARVALHO, 2023), explorar os processos epigenéticos nas moscas-das-frutas se torna fundamental. Uma vez que estas moscas são consideradas altamente generalistas e apresentam uma grande variabilidade morfológica, como evidenciado inclusive neste trabalho, é importante destacar que os efeitos ocasionados pelos mecanismos epigenéticos podem afetar não apenas caracteres morfológicos, mas também processos fisiológicos que, em conjunto, podem mediar inclusive a escolha de hospedeiros, considerando a influência que os mesmos podem ter sobre este táxon. Portanto, a compreensão dos processos epigenéticos nas moscas-das-frutas é crucial para elucidar os mecanismos subjacentes à sua adaptabilidade e variabilidade fenotípica.

CONSIDERAÇÕES FINAIS

O presente estudo evidencia a complexidade na variação morfológica das moscas-das-frutas do grupo *fraterculus*, especialmente quando consideramos a influência dos diferentes hospedeiros. A análise morfométrica revelou padrões distintos entre os grupos, destacando a importância de variáveis específicas na diferenciação. A compreensão desses padrões de adaptação é crucial não apenas para a taxonomia, mas também para elucidar os processos evolutivos, e as dinâmicas ecológicas envolvidas nas interações entre as moscas e seus hospedeiros.

É importante destacar que os frutos coletados para esse estudo se encontravam dentro do mesmo espaço geográfico (Universidade Federal de São Carlos – *campus* São Carlos, SP) e ainda assim demonstraram diferenças morfológicas significativas entre espécimes que emergiram entre os 5 diferentes frutos. É possível que exista diferenciação entre estas populações, mas em se considerando a distância entre as populações, e a vagilidade desta espécie, isso parece-nos pouco provável. Isso sugere ser possível, até mesmo provável que os frutos hospedeiros estejam influenciando na diferenciação morfológica e fenotípica das moscas-das-frutas. Além disso, a frequente hipótese considerando a plasticidade fenotípica e fatores epigenéticos destaca a necessidade contínua de explorar abordagens integrativas para entender a evolução desses organismos e suas respostas às pressões ambientais.

Nesse contexto, o presente estudo oferece uma abordagem abrangente das modificações epigenéticas, com ênfase em metilação 5mC e 6mA no DNA, nas moscas-das-frutas *Anastrepha fraterculus*. Destaca-se a importância destes resultados, uma vez que identificam, pela primeira vez, todos os dados epigenéticos aqui abordados, bem como, regiões 5mC e m6A diferencialmente metiladas (DMRs) para a espécie *A. fraterculus* que emergiram de 5 diferentes frutos de uma mesma localidade.

Os resultados aqui obtidos apontam para diversos genes potencialmente relevantes para o estudo da diferenciação intrapopulacional em *Anastrepha fraterculus*. Tais genes podem também ter papel importante na diferenciação entre espécies também, o que torna seu estudo altamente relevante. Além disso, a análise dos termos Gene Ontology (GO) forneceu insights valiosos sobre as funções biológicas associadas aos conjuntos gênicos presentes nas DMRs encontradas, destacando a importância da morfogênese de estruturas anatômicas. Este achado assume um caráter intrigante, especialmente considerando a

variedade fenotípica em caracteres morfológicos observada em espécimes próximas de moscas-das-frutas do grupo *fraterculus* em diversos estudos. A variação morfológica entre diferentes frutos, como observado no presente estudo, sugere uma possível correlação entre modificações epigenéticas e respostas distintas a diversos hospedeiros. Adicionalmente, a presença de conjuntos de genes relacionados ao sistema nervoso pode indicar influência não apenas em características morfológicas, mas também em aspectos comportamentais e fisiológicos.

Concluindo, os resultados e as interpretações apresentados nesta discussão fornecem uma base sólida para futuras investigações sobre modificações epigenéticas em moscas-das-frutas, destacando sua relevância na regulação de fenótipos, escolha de hospedeiros e adaptação evolutiva e auxiliando na compreensão da biologia desta importante praga agrícola.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKBARI, V. et al. Genome-wide detection of imprinted differentially methylated regions using nanopore sequencing. **eLife**, v. 11, p. e77898, 5 jul. 2022.
- AL ADHAMI, H. et al. A comparative methylome analysis reveals conservation and divergence of DNA methylation patterns and functions in vertebrates. **BMC Biology**, v. 20, n. 1, p. 70, 23 mar. 2022.
- ALBALAT, R. Evolution of DNA-methylation machinery: DNA methyltransferases and methyl-DNA binding proteins in the amphioxus *Branchiostoma floridae*. **Development Genes and Evolution**, v. 218, n. 11–12, p. 691–701, dez. 2008.
- ALHAKAMI, H.; MIREBRAHIM, H.; LONARDI, S. A comparative evaluation of genome assembly reconciliation tools. **Genome Biology**, v. 18, n. 1, p. 93, dez. 2017.
- ALUJA, M. et al. Phenotypic plasticity in clutch size regulation among populations of a potential invasive fruit fly from environments that vary in host heterogeneity and isolation. **Bulletin of Entomological Research**, v. 109, n. 2, p. 169–177, abr. 2019.
- ANREITER, I.; KRAMER, J. M.; SOKOLOWSKI, M. B. Epigenetic mechanisms modulate differences in *Drosophila* foraging behavior. **Proceedings of the National Academy of Sciences**, v. 114, n. 47, p. 12518–12523, 21 nov. 2017.
- ARAÚJO, E. D. S. et al. Tritrophic relations and spatial distribution of fruit flies (Diptera: Tephritidae) in the Cerrado and Caatinga regions in Piauí, Brazil. **Revista Brasileira de Entomologia**, v. 66, n. 2, p. e20210117, 2022.

- ARAUJO, E. L.; ZUCCHI, R. A. Medidas do acúleo na caracterização de cinco espécies de *Anastrepha* do grupo *fraterculus* (Diptera: Tephritidae). **Neotropical Entomology**, v. 35, n. 3, p. 329–337, jun. 2006.
- BENELLI, G. Aggression in Tephritidae Flies: Where, When, Why? Future Directions for Research in Integrated Pest Management. **Insects**, v. 6, n. 1, p. 38–53, 30 dez. 2014.
- BEWICK, A. J. et al. Evolution of DNA Methylation across Insects. **Molecular Biology and Evolution**, p. msw264, 25 dez. 2016.
- BIRD, A. Perceptions of epigenetics. **Nature**, v. 447, n. 7143, p. 396–398, maio 2007.
- BOULET, M. et al. **Adenine methylation is very scarce in the drosophila genome and not erased by the Ten Eleven Translocation dioxygenase**. [s.l.] *elife*, 11 dez. 2023. Disponível em: <<https://elifesciences.org/reviewed-preprints/91655v2>>. Acesso em: 20 dez. 2023.
- BOURDEAU, P. E. Prioritized phenotypic responses to combined predators in a marine snail. **Ecology**, v. 90, n. 6, p. 1659–1669, jun. 2009.
- CHEN, D.-P.; LIN, Y.-C.; FANN, C. S. J. Methods for identifying differentially methylated regions for sequence- and array-based data: Table 1. **Briefings in Functional Genomics**, p. elw018, 20 jun. 2016.
- CHEN, P. et al. Portable nanopore-sequencing technology: Trends in development and applications. **Frontiers in Microbiology**, v. 14, p. 1043967, 1 fev. 2023.
- CHEVIN, L.-M.; LANDE, R. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? **Evolution**, v. 64, n. 4, p. 1143–1150, 10 out. 2009.
- CONGRAINS, C. et al. Phylogenomic analysis provides diagnostic tools for the identification of *Anastrepha fraterculus* (Diptera: Tephritidae) species complex. **Evolutionary Applications**, v. 16, n. 9, p. 1598–1618, set. 2023.
- CONTRERAS-MIRANDA, J. A. et al. Search hours for food attractant by *Anastrepha fraterculus* and *Ceratitis capitata* (Diptera: Tephritidae) adults in guava orchards. **Anais da Academia Brasileira de Ciências**, v. 95, n. 2, p. e20201880, 2023.
- CORNWALLIS, C. K.; BIRKHEAD, T. R. Plasticity in reproductive phenotypes reveals status-specific correlations between behavioral, morphological, and physiological sexual traits. **Evolution**, v. 62, n. 5, p. 1149–1161, maio 2008.
- CORTINI, R. et al. The physics of epigenetics. **Reviews of Modern Physics**, v. 88, n. 2, p. 025002, 26 abr. 2016.
- DAHLET, T. et al. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. **Nature Communications**, v. 11, n. 1, p. 3153, 19 jun. 2020.
- DE CARVALHO, C. F. Epigenetic effects of climate change on insects. **Current Opinion in Insect Science**, v. 57, p. 101029, jun. 2023.

- DE COSTER, W. et al. NanoPack: visualizing and processing long-read sequencing data. **Bioinformatics**, v. 34, n. 15, p. 2666–2669, 1 ago. 2018.
- DESHMUKH, S. et al. Levels of DNA cytosine methylation in the *Drosophila* genome. **PeerJ**, v. 6, p. e5119, 2 jul. 2018.
- DESHMUKH, S. et al. **Genome-wide comparison of DNA methylation between life cycle stages of *Drosophila melanogaster* using high-throughput sequencing techniques**. [s.l.] Genomics, 7 jan. 2020. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2020.01.06.895920>>. Acesso em: 17 dez. 2023.
- DIAS, V. S. et al. An integrative multidisciplinary approach to understanding cryptic divergence in Brazilian species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Biological Journal of the Linnean Society**, v. 117, n. 4, p. 725–746, abr. 2016.
- DÍAZ, F. et al. Evidence for Introgression Among Three Species of the *Anastrepha fraterculus* Group, a Radiating Species Complex of Fruit Flies. **Frontiers in Genetics**, v. 9, p. 359, 10 set. 2018.
- DUNCAN, E. J.; CUNNINGHAM, C. B.; DEARDEN, P. K. Phenotypic Plasticity: What Has DNA Methylation Got to Do with It? **Insects**, v. 13, n. 2, p. 110, 19 jan. 2022.
- DUNCAN, E. J.; GLUCKMAN, P. D.; DEARDEN, P. K. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? **Journal of Experimental Zoology Part B: Molecular and Developmental Evolution**, v. 322, n. 4, p. 208–220, jun. 2014.
- ERCIT, K. Temporal variation in selection on male and female traits in wild tree crickets. **Ecology and Evolution**, v. 6, n. 15, p. 5118–5128, ago. 2016.
- FEIL, R.; FRAGA, M. F. Epigenetics and the environment: emerging patterns and implications. **Nature Reviews Genetics**, v. 13, n. 2, p. 97–109, fev. 2012.
- FOOTE, R. H. Fruit Fly Genera South of the United States (Diptera: Tephritidae). 1980.
- FORET, S. et al. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. **Proceedings of the National Academy of Sciences**, v. 109, n. 13, p. 4968–4973, 27 mar. 2012.
- GIARETTA, A. et al. Phylogenetic Relationships Within the Hyper-Diverse Genus *Eugenia* (Myrtaceae: Myrteae) Based on Target Enrichment Sequencing. **Frontiers in Plant Science**, v. 12, p. 759460, 4 fev. 2022.
- GLASTAD, K. M. et al. DNA methylation in insects: on the brink of the epigenomic era. **Insect Molecular Biology**, v. 20, n. 5, p. 553–565, out. 2011.
- GLASTAD, K. M.; HUNT, B. G.; GOODISMAN, M. A. D. Epigenetics in Insects: Genome Regulation and the Generation of Phenotypic Diversity. **Annual Review of Entomology**, v. 64, n. 1, p. 185–203, 7 jan. 2019.

- GÓMEZ-CENDRA, P. V. et al. Morphometric Differentiation Among *Anastrepha fraterculus* (Diptera: Tephritidae) Exploiting Sympatric Alternate Hosts. **Environmental Entomology**, v. 45, n. 2, p. 508–517, abr. 2016.
- GOTTHARD, K.; NYLIN, S.; NYLIN, S. Adaptive Plasticity and Plasticity as an Adaptation: A Selective Review of Plasticity in Animal Morphology and Life History. **Oikos**, v. 74, n. 1, p. 3, out. 1995.
- GUO, J. U. et al. Neuronal activity modifies the DNA methylation landscape in the adult brain. **Nature Neuroscience**, v. 14, n. 10, p. 1345–1351, out. 2011.
- GUSTANI, E. C. **Evolução e integração de fenótipos complexos em insetos: Estudo de caso em *Teleopsis dalmanni* (Diptera: Diopsidae) e em espécies de *Anastrepha* do grupo *fraterculus* (Diptera: Tephritidae)**. São Carlos: Universidade Federal de São Carlos, 31 ago. 2018.
- HANDEL, A. E.; EBERS, G. C.; RAMAGOPALAN, S. V. Epigenetics: molecular mechanisms and implications for disease. **Trends in Molecular Medicine**, v. 16, n. 1, p. 7–16, jan. 2010.
- HAYWARD, A. D. et al. Evidence for Selection-by-Environment but Not Genotype-by-Environment Interactions for Fitness-Related Traits in a Wild Mammal Population. **Genetics**, v. 208, n. 1, p. 349–364, 1 jan. 2018.
- HERNÁNDEZ-ORTIZ, V. et al. Morphometric analysis of Mexican and South American populations of the *Anastrepha fraterculus* complex (Diptera: Tephritidae) and recognition of a distinct Mexican morphotype. **Bulletin of Entomological Research**, v. 94, n. 6, p. 487–499, dez. 2004.
- HERNÁNDEZ-ORTIZ, V. GUILLÉN-AGUILAR, J. LÓPEZ, L. Taxonomía e identificación de moscas de la fruta de importância económica em América. In: MONTOYA, P. TOLEDO, J. HERNÁNDEZ, E. (ed.). **Moscas de la fruta: Fundamentos y procedimientos para su manejo**. Df: S y G Editores, 2010. p. 49-79.
- HERNÁNDEZ-ORTIZ, V. et al. Cryptic Species of the *Anastrepha fraterculus* Complex (Diptera: Tephritidae): A Multivariate Approach for the Recognition of South American Morphotypes. **Annals of the Entomological Society of America**, v. 105, n. 2, p. 305–318, 1 mar. 2012.
- HERNÁNDEZ-ORTIZ, V. et al. Taxonomy and phenotypic relationships of the *Anastrepha fraterculus* complex in the Mesoamerican and Pacific Neotropical dominions (Diptera, Tephritidae). **ZooKeys**, v. 540, p. 95–124, 26 nov. 2015.
- HOLLIDAY, R. Epigenetics: An overview. **Developmental Genetics**, v. 15, n. 6, p. 453–457, jan. 1994.
- HUANG, J.-P. Revisiting rapid phenotypic evolution in sticklebacks: integrative thinking of standing genetic variation and phenotypic plasticity. **Frontiers in Ecology and Evolution**, v. 3, 8 maio 2015.
- IP, C. L. C. et al. MinION Analysis and Reference Consortium: Phase 1 data release and analysis. **F1000Research**, v. 4, p. 1075, 15 out. 2015.

- JENUWEIN, T.; ALLIS, C. D. Translating the Histone Code. **Science**, v. 293, n. 5532, p. 1074–1080, 10 ago. 2001.
- KAMILAR, J. M.; COOPER, N. Phylogenetic signal in primate behaviour, ecology and life history. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 368, n. 1618, p. rstb.2012.0341, 20120341, 19 maio 2013.
- KHAMIS, F. M. et al. Taxonomic Identity of the Invasive Fruit Fly Pest, *Bactrocera invadens*: Concordance in Morphometry and DNA Barcoding. **PLoS ONE**, v. 7, n. 9, p. e44862, 17 set. 2012.
- KIM, B. Y. et al. Highly contiguous assemblies of 101 drosophilid genomes. **eLife**, v. 10, p. e66405, 19 jul. 2021.
- KLINGENBERG, C. P.; GIDASZEWSKI, N. A. Testing and Quantifying Phylogenetic Signals and Homoplasy in Morphometric Data. **Systematic Biology**, v. 59, n. 3, p. 245–261, 1 maio 2010.
- KLIRONOMOS, F. D.; BERG, J.; COLLINS, S. How epigenetic mutations can affect genetic evolution: Model and mechanism. **BioEssays**, v. 35, n. 6, p. 571–578, jun. 2013.
- KLUGHAMMER, J. et al. Comparative analysis of genome-scale, base-resolution DNA methylation profiles across 580 animal species. **Nature Communications**, v. 14, n. 1, p. 232, 16 jan. 2023.
- KOUZARIDES, T. Chromatin Modifications and Their Function. **Cell**, v. 128, n. 4, p. 693–705, fev. 2007.
- KUCHARSKI, R. et al. Nutritional Control of Reproductive Status in Honeybees via DNA Methylation. **Science**, v. 319, n. 5871, p. 1827–1830, 28 mar. 2008.
- KURDYUKOV, S.; BULLOCK, M. DNA Methylation Analysis: Choosing the Right Method. **Biology**, v. 5, n. 1, p. 3, 6 jan. 2016.
- LAMPERT, S. et al. Distribution of *Anastrepha Schiner*, 1868 (Diptera, Tephritidae) in Brazil: new records from the state of Paraná; **Check List**, v. 16, n. 4, p. 799–804, 1 jul. 2020.
- LASSERRE, D. F. Modes of Sympatric Speciation in Phytophagous Insects of Diptera Tephritidae and the Importance of Epigenetics and Behavior Imprinting in Host Fidelity. **Photon**, v. 115, p. 179–194, 23 fev. 2014.
- LAUBACH, Z. M. et al. Epigenetics and the maintenance of developmental plasticity: extending the signalling theory framework. **Biological Reviews**, v. 93, n. 3, p. 1323–1338, ago. 2018.
- LEA, A. J. et al. Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*). **Molecular Ecology**, v. 25, n. 8, p. 1681–1696, abr. 2016.
- LEGER, A. et al. Genomic variations and epigenomic landscape of the Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel. **Genome Biology**, v. 23, n. 1, p. 58, dez. 2022.

- LEWIS, S. H. et al. -----Widespread conservation and lineage-specific diversification of genome-wide DNA methylation patterns across arthropods. **PLOS Genetics**, v. 16, n. 6, p. e1008864, 25 jun. 2020.
- LI, H. et al. DNA N6-Methyladenine Modification in Eukaryotic Genome. **Frontiers in Genetics**, v. 13, p. 914404, 24 jun. 2022.
- LOKK, K. et al. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. **Genome Biology**, v. 15, n. 4, p. 3248, abr. 2014.
- LYKO, F.; RAMSAHOYE, B. H.; JAENISCH, R. DNA methylation in *Drosophila melanogaster*. **Nature**, v. 408, n. 6812, p. 538–540, nov. 2000.
- MAHALANOBIS, P. C. On the generalised distância in statistics». **Proceedings of the National Institute of Sciences of India**. v. 2, n.1, p. 49–55,1936.
- MAGI, A. et al. High-resolution Nanopore methylome-maps reveal random hyper-methylation at CpG-poor regions as driver of chemoresistance in leukemias. **Communications Biology**, v. 6, n. 1, p. 382, 8 abr. 2023.
- MANFREDINI, F.; ARBETMAN, M.; TOTH, A. L. A Potential Role for Phenotypic Plasticity in Invasions and Declines of Social Insects. **Frontiers in Ecology and Evolution**, v. 7, p. 375, 15 out. 2019.
- MARGUERON, R.; REINBERG, D. Chromatin structure and the inheritance of epigenetic information. **Nature Reviews Genetics**, v. 11, n. 4, p. 285–296, abr. 2010.
- MARTIN, S. H. et al. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. **Genome Research**, v. 23, n. 11, p. 1817–1828, nov. 2013.
- MCPHERON, B. A.; SMITH, D. C.; BERLOCHER, S. H. Genetic differences between host races of *Rhagoletis pomonella*. **Nature**, v. 336, n. 6194, p. 64–66, nov. 1988.
- MENGUAL, X. et al. Phylogenetic relationships of the tribe Toxotrypanini (Diptera: Tephritidae) based on molecular characters. **Molecular Phylogenetics and Evolution**, v. 113, p. 84–112, ago. 2017.
- MOGILICHERLA, K.; ROY, A. Epigenetic regulations as drivers of insecticide resistance and resilience to climate change in arthropod pests. **Frontiers in Genetics**, v. 13, p. 1044980, 6 jan. 2023.
- MONTEIRO, L. R.; REIS, S. F. dos. Princípios de morfometria geométrica. *In: Princípios de morfometria geométrica*. p. 189-189, 1999.
- MOORE, M. K. Sex Estimation and Assessment. Em: **Research Methods in Human Skeletal Biology**. [s.l.] Elsevier, 2013. p. 91–116.
- NEL, P.; BERTRAND, S.; NEL, A. Diversification of insects since the Devonian: a new approach based on morphological disparity of mouthparts. **Scientific Reports**, v. 8, n. 1, p. 3516, 23 fev. 2018.

- NIJHOUT, H. F. Development and evolution of adaptive polyphenisms. **Evolution & Development**, v. 5, n. 1, p. 9–18, jan. 2003.
- NORRBOM, A. L. et al. New species and host plants of *Anastrepha* (Diptera: Tephritidae) primarily from Peru and Bolivia. **Zootaxa**, v. 4041, n. 1, p. 1, 16 nov. 2015.
- NORRBOM, A. L. et al. New species and host plants of *Anastrepha* (Diptera: Tephritidae) primarily from Suriname and Pará, Brazil. **Zootaxa**, v. 5044, n. 1, p. 1–74, 27 set. 2021.
- NORRBOM, A. L.; KORYTKOWSKI, C. A. A revision of the *Anastrepha robusta* species group (Diptera: Tephritidae). **Zootaxa**, v. 2182, n. 1, p. 1–91, 6 ago. 2009.
- NORRBOM, A. L.; KORYTKOWSKI, C. A. New species of and taxonomic notes on *Anastrepha* (Diptera: Tephritidae). **Zootaxa**, v. 2740, n. 1, p. 1, 14 jan. 2011.
- OKANO, M. et al. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. **Cell**, v. 99, n. 3, p. 247–257, out. 1999.
- OLDROYD, B. P.; YAGOUND, B. The role of epigenetics, particularly DNA methylation, in the evolution of caste in insect societies. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 376, n. 1826, p. 20200115, 7 jun. 2021.
- OLIVARES-CASTRO, G. et al. Insect Epigenetic Mechanisms Facing Anthropogenic-Derived Contamination, an Overview. **Insects**, v. 12, n. 9, p. 780, 31 ago. 2021.
- ÖST, A. et al. Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity. **Cell**, v. 159, n. 6, p. 1352–1364, dez. 2014.
- PEREZ, M. F.; LEHNER, B. Intergenerational and transgenerational epigenetic inheritance in animals. **Nature Cell Biology**, v. 21, n. 2, p. 143–151, fev. 2019.
- 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, 1 mar. 2014.
- PFENNIG, D. W. et al. Phenotypic plasticity's impacts on diversification and speciation. **Trends in Ecology & Evolution**, v. 25, n. 8, p. 459–467, ago. 2010.
- PIGLIUCCI, M. Evolution of phenotypic plasticity: where are we going now? **Trends in Ecology & Evolution**, v. 20, n. 9, p. 481–486, set. 2005.
- PIGLIUCCI, M.; MURREN, C. J.; SCHLICHTING, C. D. Phenotypic plasticity and evolution by genetic assimilation. **Journal of Experimental Biology**, v. 209, n. 12, p. 2362–2367, 15 jun. 2006.
- PROVATARIS, P. et al. Signatures of DNA Methylation across Insects Suggest Reduced DNA Methylation Levels in Holometabola. **Genome Biology and Evolution**, v. 10, n. 4, p. 1185–1197, 1 abr. 2018.

- QUINA, A. S.; BUSCHBECK, M.; DI CROCE, L. Chromatin structure and epigenetics. **Biochemical Pharmacology**, v. 72, n. 11, p. 1563–1569, nov. 2006.
- RAGA, A. et al. Host Ranges and Infestation Indices of Fruit Flies (Tephritidae) and Lance Flies (Lonchaeidae) in São Paulo State, Brazil. **Florida Entomologist**, v. 94, n. 4, p. 787–794, dez. 2011.
- RAGLAND, G. J. et al. Environmental interactions during host race formation: host fruit environment moderates a seasonal shift in phenology in host races of *Rhagoletis pomonella*. **Functional Ecology**, v. 26, n. 4, p. 921–931, ago. 2012.
- RAND, A. C. et al. Mapping DNA methylation with high-throughput nanopore sequencing. **Nature Methods**, v. 14, n. 4, p. 411–413, abr. 2017.
- RENARD, T.; GUEYDAN, C.; ARON, S. DNA methylation and expression of the *egfr* gene are associated with worker size in monomorphic ants. **Scientific Reports**, v. 12, n. 1, p. 21228, 8 dez. 2022.
- ROHLF, F. J. Geometric morphometrics and phylogeny. *In*: MACLEOD, N.; FOREY, P. (Eds.). Morphology, Shape and Phylogenetics. **Systematics Association Special Volume**, v. 64, p. 175-193, 2002.
- SCHAEFER, M.; LYKO, F. Solving the Dnmt2 enigma. **Chromosoma**, v. 119, n. 1, p. 35–40, fev. 2010.
- SCHATZ, M. C. Nanopore sequencing meets epigenetics. **Nature Methods**, v. 14, n. 4, p. 347–348, abr. 2017.
- SCHEINER, S. M. Selection experiments and the study of phenotypic plasticity. **Journal of Evolutionary Biology**, v. 15, n. 6, p. 889–898, 1 nov. 2002.
- SCHEINER, S. M.; BARFIELD, M.; HOLT, R. D. The genetics of phenotypic plasticity. XVII. Response to climate change. **Evolutionary Applications**, v. 13, n. 2, p. 388–399, fev. 2020.
- SCHMITZ, R. J.; LEWIS, Z. A.; GOLL, M. G. DNA Methylation: Shared and Divergent Features across Eukaryotes. **Trends in Genetics**, v. 35, n. 11, p. 818–827, nov. 2019.
- SCHOONHOVEN, L. M.; LOON, J. J. A. VAN; DICKE, M. **Insect-plant biology**. 2nd ed. Oxford ; New York: Oxford University Press, 2005.
- SEGURA, D. F. et al. Plant Chemicals and the Sexual Behavior of Male Tephritid Fruit Flies. **Annals of the Entomological Society of America**, v. 111, n. 5, p. 239–264, 14 set. 2018.
- SELIVON, D. et al. Genetical, Morphological, Behavioral, and Ecological Traits Support the Existence of Three Brazilian Species of the *Anastrepha fraterculus* Complex of Cryptic Species. **Frontiers in Ecology and Evolution**, v. 10, p. 836608, 27 abr. 2022.

- SHAH, K.; CAO, W.; ELLISON, C. E. Adenine Methylation in *Drosophila* Is Associated with the Tissue-Specific Expression of Developmental and Regulatory Genes. **G3 Genes|Genomes|Genetics**, v. 9, n. 6, p. 1893–1900, 1 jun. 2019.
- SILVA, T.G.A. **Caracterização morfológica e molecular de *Anastrepha bistrigata* Bezzi e *Anastrepha striata* Schiner (Diptera: Tephritidae)**. 2008. Dissertação (Mestrado em Ciências – Entomologia), Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2008.
- SILVA, J. G. et al. Diversity of *Anastrepha* spp. (Diptera: Tephritidae) and Associated Braconid Parasitoids From Native and Exotic Hosts in Southeastern Bahia, Brazil. **Environmental Entomology**, v. 39, n. 5, p. 1457–1465, 1 out. 2010.
- SIMPSON, J. T. et al. Detecting DNA cytosine methylation using nanopore sequencing. **Nature Methods**, v. 14, n. 4, p. 407–410, abr. 2017.
- SIMPSON LAB. **Nanopolish Documentation**. S.L: Oxford Nanopore, 2019.
- SKINNER, M. K. Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. **Epigenetics**, v. 6, n. 7, p. 838–842, jul. 2011.
- SKINNER, M. K. et al. Epigenetics and the Evolution of Darwin’s Finches. **Genome Biology and Evolution**, v. 6, n. 8, p. 1972–1989, ago. 2014.
- SNAJDER, R. et al. **pycoMeth: A toolbox for differential methylation testing from Nanopore methylation calls**. [s.l.] Bioinformatics, 19 fev. 2022. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2022.02.16.480699>>. Acesso em: 20 dez. 2023.
- SOUZA-FILHO, MF. et al. Diversity and seasonality of fruit flies (Diptera: Tephritidae and Lonchaeidae) and their parasitoids (Hymenoptera: Braconidae and Figitidae) in orchards of guava, loquat and peach. **Brazilian Journal of Biology**, v. 69, n. 1, p. 31–40, fev. 2009.
- STAJIC, D.; JANSEN, L. E. T. Empirical evidence for epigenetic inheritance driving evolutionary adaptation. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 376, n. 1826, p. 20200121, 7 jun. 2021.
- STONE, A. **The Fruitflies of the Genus *Anastrepha***. Washington: U.s. Department Of Agriculture, 1942. 112 p.
- SUETAKE, I. et al. DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction. **Journal of Biological Chemistry**, v. 279, n. 26, p. 27816–27823, jun. 2004.
- SUZUKI, M. M.; BIRD, A. DNA methylation landscapes: provocative insights from epigenomics. **Nature Reviews Genetics**, v. 9, n. 6, p. 465–476, jun. 2008.
- TAKAYAMA, S. et al. Genome methylation in *D. melanogaster* is found at specific short motifs and is independent of DNMT2 activity. **Genome Research**, v. 24, n. 5, p. 821–830, maio 2014.

- TAKEMORI, N. et al. Novel eye-specific calmodulin methylation characterized by protein mapping in *Drosophila melanogaster*. **PROTEOMICS**, v. 7, n. 15, p. 2651–2658, ago. 2007.
- THORSON, J. L. M. et al. Epigenetics and adaptive phenotypic variation between habitats in an asexual snail. **Scientific Reports**, v. 7, n. 1, p. 14139, 26 out. 2017.
- TOLLEFSBOL, T. O. An Overview of Epigenetics. Em: **Handbook of Epigenetics**. [s.l.] Elsevier, 2017. p. 1–6.
- TROYA, H.; NORRBOM, A. L.; PINEDA, J. E. B. Two new species of *Anastrepha* (Diptera: Tephritidae) from Ecuador. **Zootaxa**, v. 4820, n. 2, 28 jul. 2020.
- VANÍČKOVÁ, L. et al. Current knowledge of the species complex *Anastrepha fraterculus* (Diptera, Tephritidae) in Brazil. **ZooKeys**, v. 540, p. 211–237, 26 nov. 2015.
- VARLEY, K. E. et al. Dynamic DNA methylation across diverse human cell lines and tissues. **Genome Research**, v. 23, n. 3, p. 555–567, mar. 2013.
- VASCONCELOS, T. N. C. et al. Myrteae phylogeny, calibration, biogeography and diversification patterns: Increased understanding in the most species rich tribe of Myrtaceae. **Molecular Phylogenetics and Evolution**, v. 109, p. 113–137, abr. 2017.
- VELASCO-CUERVO, S. M.; GALINDO-GONZÁLEZ, L.; TORO-PEREA, N. An omics evolutionary perspective on phytophagous insect–host plant interactions in *Anastrepha obliqua*: a review. **Entomologia Experimentalis et Applicata**, v. 171, n. 1, p. 2–16, jan. 2023.
- VERHOEVEN, K. J. F.; PREITE, V. Epigenetic variation in asexually reproducing organisms: special section. **Evolution**, v. 68, n. 3, p. 644–655, mar. 2014.
- VILLAGRA, C.; FRÍAS-LASSERRE, D. Epigenetic Molecular Mechanisms in Insects. **Neotropical Entomology**, v. 49, n. 5, p. 615–642, out. 2020.
- WANG, S. P.; ALTHOFF, D. M. Phenotypic plasticity facilitates initial colonization of a novel environment. **Evolution**, v. 73, n. 2, p. 303–316, fev. 2019.
- WANG, X. et al. DNA methylation on N6-adenine in lepidopteran *Bombyx mori*. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, v. 1861, n. 9, p. 815–825, set. 2018.
- WEN, J.; UENO, T. Predator cue-induced plasticity of morphology and behavior in planthoppers facilitate the survival from predation. **Scientific Reports**, v. 11, n. 1, p. 16760, 18 ago. 2021.
- WILLIAMS, K. D.; SOKOLOWSKI, M. B. **Phenotypic plasticity**. [s.l.] Oxford University Press, 2018. v. 1
- XIE, J. et al. Dynamic Regulation of DNA Methylation and Brain Functions. **Biology**, v. 12, n. 2, p. 152, 18 jan. 2023.

YAGOUND, B. et al. Intergenerational transfer of DNA methylation marks in the honey bee. **Proceedings of the National Academy of Sciences**, v. 117, n. 51, p. 32519–32527, 22 dez. 2020.

YAN, H. et al. DNA Methylation in Social Insects: How Epigenetics Can Control Behavior and Longevity. **Annual Review of Entomology**, v. 60, n. 1, p. 435–452, 7 jan. 2015.

YAO, B. et al. Active N6-Methyladenine Demethylation by DMAD Regulates Gene Expression by Coordinating with Polycomb Protein in Neurons. **Molecular Cell**, v. 71, n. 5, p. 848- 857.e6, set. 2018.

ZEE, B. M. et al. Streamlined discovery of cross-linked chromatin complexes and associated histone modifications by mass spectrometry. **Proceedings of the National Academy of Sciences**, v. 113, n. 7, p. 1784–1789, 16 fev. 2016.

ZHANG, G. et al. N6-Methyladenine DNA Modification in Drosophila. **Cell**, v. 161, n. 4, p. 893–906, maio 2015.

ZHANG, Q. et al. N6-Methyladenine DNA Methylation in Japonica and Indica Rice Genomes and Its Association with Gene Expression, Plant Development, and Stress Responses. **Molecular Plant**, v. 11, n. 12, p. 1492–1508, dez. 2018.

ZUCCHI, R. A. Espécies de *Anastrepha*, sinonimias, plantas hospedeiras e parasitoides. In: MALAVASI, A. e ZUCCHI, R. A. (Ed.). **Moscas-das-frutas de importância econômica no Brasil: Conhecimento básico e aplicado**. Ribeirão Preto, Brazil: Holos, 2000. p.41-48.

ZUCCHI, R.A. & MORAES, R.C.B. 2008. **Fruit flies in Brazil** - *Anastrepha* species their host plants and parasitoids. Available in: www.lea.esalq.usp.br/anastrepha/, updated on November 10, 2020. Acesso em 12 de dez de 2020.

MATERIAL SUPLEMENTAR

Tabelas

	Cereja do cerrado	Uvaia	Pitanga	Goiaba	Gabiroba	Total
Cereja do cerrado	4	0	2	0	0	6
Uvaia	0	18	1	2	0	21
Pitanga	3	1	23	0	1	28
Goiaba	0	0	0	27	13	40
Gabiroba	0	2	1	4	10	17
Total	7	21	27	33	24	112

Tabela S 1. Lambda de Wilk's para as amostras escolhidas para análise.

	Wilks'	Partial	F-remove	p-value	Toler.	1-Toler.
A2	0,082	0,961359	0,99481	0,414	0,468174	0,531826
A3	0,089	0,881260	3,33478	0,013	0,469499	0,530501
A8	0,113	0,696614	10,77898	0,000	0,647446	0,352554
A9	0,082	0,954503	1,17972	0,324	0,662098	0,337902
W1	0,127	0,620826	15,11623	0,000	0,448073	0,551927
W7	0,084	0,940802	1,55733	0,192	0,809528	0,190472
HEL	0,096	0,819727	5,44297	0,001	0,341389	0,658611
HW	0,085	0,929679	1,87210	0,121	0,440157	0,559843
TL	0,110	0,716700	9,78329	0,000	0,306729	0,693271

TW	0,096	0,818094	5,50326	0,000	0,316743	0,683257
-----------	-------	----------	---------	-------	----------	----------

Tabela S 2. Os 100 principais intervalos diferenciamente metilados significativos.

rank	p-value	chromosome	start	end	Number of nearby TSS	closest tx id	distance to tss
#1	3,70E-62	NC_071500.1	6E+07	6E+07	11	xm_054099679.1	-13901
#2	1,93E-54	NC_071500.1	3E+07	3E+07	4	xm_054097723.1	11518
#3	3,92E-47	NC_071499.1	4E+06	4E+06	2	xm_054091699.1	255
#4	5,50E-45	NC_071503.1	7E+07	7E+07	15	xm_054113192.1	51385
#5	6,51E-43	NC_071498.1	1E+08	1E+08	16	xm_054090498.1	0
#6	3,15E-42	NC_071500.1	1E+08	1E+08	15	xm_054098744.1	-3881
#7	5,43E-42	NC_071500.1	6E+07	6E+07	6	xm_054101318.1	-33710
#8	1,38E-41	NC_071503.1	9E+07	9E+07	1	xm_054112944.1	-461
#9	5,41E-41	NC_071498.1	4E+07	4E+07	17	xm_054112456.1	-8370
#10	1,45E-40	NC_071503.1	3E+07	3E+07	1	xm_054111834.1	44234
#11	7,67E-40	NC_071498.1	1E+08	1E+08	13	xm_054089593.1	-3757
#12	1,12E-39	NC_071500.1	5E+07	5E+07	7	xm_054100913.1	-20157
#13	5,97E-39	NC_071503.1	3E+07	3E+07	1	xm_054111834.1	47270
#14	2,37E-37	NC_071498.1	1E+08	1E+08	16	xm_054089290.1	105
#15	2,63E-37	NC_071503.1	5E+07	5E+07	1	xm_054112613.1	-72639
#16	5,20E-37	NC_071500.1	9E+07	9E+07	0		
#17	2,16E-36	NC_071503.1	3E+07	3E+07	1	xm_054112461.1	24611
#18	2,74E-36	NC_071502.1	1E+08	1E+08	10	xm_054110566.1	-15434
#19	4,65E-36	NC_071498.1	2E+08	2E+08	14	xm_054091470.1	4918
#20	6,68E-36	NC_071500.1	1E+08	1E+08	9	xm_054098782.1	-22886
#21	8,76E-36	NC_071499.1	1E+08	1E+08	8	xm_054095295.1	0

#22	9,99E-36	NC_071498.1	2E+07	2E+07	58	xm_054103222.1	0
#23	1,27E-35	NC_071498.1	1E+08	1E+08	24	xm_054089375.1	0
#24	1,86E-35	NC_071499.1	1E+08	1E+08	22	xm_054095605.1	18644
#25	3,05E-35	NC_071499.1	4E+06	4E+06	6	trnawcca4	-11153
#26	3,89E-33	NC_071501.1	9E+07	9E+07	27	xm_054102442.1	750
#27	5,18E-33	NC_071501.1	4E+07	4E+07	14	xm_054102112.1	-4341
#28	7,91E-33	NC_071503.1	4E+07	4E+07	1	xm_054113159.1	-72846
#29	2,14E-32	NC_071500.1	3E+07	3E+07	23	xm_054099501.1	0
#30	2,43E-32	NC_071503.1	6E+07	6E+07	5	xm_054111935.1	-61
#31	3,45E-32	NC_071501.1	5E+07	5E+07	3	xm_054101448.1	-32473
#32	4,32E-32	NC_071501.1	1E+08	1E+08	1	xm_054102539.1	-9550
#33	6,11E-32	NC_071498.1	3E+07	3E+07	7	xm_054109621.1	2764
#34	6,74E-32	NC_071502.1	2E+07	2E+07	17	xm_054106268.1	15787
#35	7,79E-32	NC_071503.1	4E+07	4E+07	13	xm_054113152.1	3700
#36	1,38E-31	NC_071503.1	9E+07	9E+07	0		
#37	1,71E-31	NC_071502.1	3E+07	3E+07	16	xr_008455169.1	-35446
#38	2,24E-31	NC_071503.1	8E+07	8E+07	7	xm_054112863.1	561
#39	3,14E-31	NC_071502.1	5E+07	5E+07	9	xm_054108982.1	0
#40	3,15E-31	NC_071501.1	7E+07	7E+07	1	xr_008454790.1	-71932
#41	3,31E-31	NC_071498.1	3E+07	3E+07	22	xm_054092331.1	-312
#42	5,39E-31	NC_071502.1	1E+08	1E+08	49	xm_054110535.1	768
#43	9,20E-31	NC_071503.1	1E+07	1E+07	3	xm_054111821.1	51522
#44	1,08E-30	NC_071498.1	884504	892267	8	xm_054091530.1	-3320
#45	1,31E-30	NC_071501.1	6E+06	6E+06	5	xm_054102684.1	5375
#46	3,62E-30	NC_071500.1	9E+07	9E+07	5	xm_054100458.1	0

#47	3,98E-30	NC_071503.1	8E+07	8E+07	0		
#48	6,03E-30	NC_071500.1	1E+08	1E+08	12	xm_054098322.1	13817
#49	7,24E-30	NC_071498.1	4E+06	4E+06	4	xm_054091599.1	-987
#50	1,27E-29	NC_071502.1	1E+07	1E+07	7	xm_054111225.1	-22473
#51	1,50E-29	NC_071500.1	1E+08	1E+08	16	xr_008454299.1	-16038
#52	4,29E-29	NC_071503.1	1E+07	1E+07	4	xm_054112365.1	-9862
#53	4,66E-29	NC_071500.1	9E+07	9E+07	3	xm_054097635.1	3836
#54	1,74E-28	NC_071499.1	8E+07	8E+07	44	xm_054094395.1	6393
#55	2,57E-28	NC_071500.1	1E+08	1E+08	2	xm_054100642.1	-75470
#56	4,69E-28	NC_071502.1	9E+07	9E+07	20	xm_054109828.1	0
#57	5,68E-28	NC_071501.1	5E+07	5E+07	8	xm_054104453.1	22439
#58	7,66E-28	NC_071498.1	1E+08	1E+08	2	xm_054090351.1	0
#59	8,23E-28	NC_071503.1	1E+07	1E+07	4	xm_054112360.1	-43562
#60	1,06E-27	NC_071498.1	9E+07	9E+07	6	xm_054089202.1	7542
#61	1,89E-27	NC_071498.1	8E+07	8E+07	12	xm_054094486.1	-100
#62	2,99E-27	NC_071499.1	1E+08	1E+08	0		
#63	3,10E-27	NC_071498.1	2E+06	2E+06	11	xm_054099795.1	0
#64	3,16E-27	NC_071501.1	2E+06	2E+06	12	xm_054104689.1	-18815
#65	3,28E-27	NC_071501.1	8E+07	8E+07	13	xm_054102910.1	7239
#66	3,54E-27	NC_071498.1	1E+08	1E+08	9	xm_054090021.1	38610
#67	6,20E-27	NC_071503.1	5E+07	5E+07	4	xr_008455263.1	42542
#68	6,39E-27	NC_071503.1	2E+07	2E+07	4	xm_054112420.1	-26009
#69	8,29E-27	NC_071500.1	1E+08	1E+08	44	xm_054098465.1	0
#70	1,19E-26	NC_071499.1	3E+07	3E+07	4	xm_054092984.1	48790
#71	1,23E-26	NC_071502.1	3E+07	3E+07	16	xm_054107380.1	423

#72	3,26E-26	NC_071502.1	5E+07	5E+07	0		
#73	3,37E-26	NC_071503.1	7E+07	7E+07	1	xm_054111898.1	30129
#74	5,02E-26	NC_071502.1	1E+08	1E+08	2	xm_054108398.1	50956
#75	5,25E-26	NC_071499.1	9E+07	9E+07	5	xm_054094582.1	87634
#76	5,89E-26	NC_071498.1	2E+08	2E+08	10	xm_054098309.1	0
#77	6,21E-26	NC_071498.1	8E+07	8E+07	20	xm_054114066.1	0
#78	6,56E-26	NC_071499.1	9E+07	9E+07	0		
#79	6,56E-26	NC_071499.1	1E+08	1E+08	35	xm_054096115.1	8959
#80	7,99E-26	NC_071503.1	5E+07	5E+07	3	xm_054112183.1	49297
#81	8,53E-26	NC_071499.1	1E+08	1E+08	10	xm_054093387.1	-3147
#82	1,08E-25	NC_071503.1	7E+07	7E+07	11	xm_054112080.1	0
#83	1,13E-25	NC_071498.1	3E+07	3E+07	6	xm_054092576.1	-10675
#84	1,14E-25	NC_071503.1	2E+07	2E+07	0		
#85	1,24E-25	NC_071498.1	7E+07	7E+07	8	xm_054093880.1	16786
#86	2,14E-25	NC_071498.1	1E+08	1E+08	12	xm_054089956.1	0
#87	2,33E-25	NC_071500.1	3E+07	3E+07	4	xm_054099243.1	-61063
#88	2,52E-25	NC_071500.1	1E+08	1E+08	28	xm_054097758.1	-5276
#89	2,62E-25	NC_071502.1	4E+07	4E+07	26	xm_054111083.1	134
#90	3,08E-25	NC_071501.1	7E+06	7E+06	8	xm_054104690.1	-61803
#91	3,28E-25	NC_071499.1	8E+07	8E+07	0		
#92	3,99E-25	NC_071499.1	1E+08	1E+08	6	xm_054095494.1	67014
#93	6,88E-25	NC_071500.1	8E+07	8E+07	12	xm_054097202.1	-12801
#94	8,07E-25	NC_071499.1	1E+08	1E+08	0		
#95	9,83E-25	NC_071502.1	7E+07	7E+07	3	xm_054106459.1	0
#96	1,17E-24	NC_071498.1	2E+08	2E+08	20	xm_054091481.1	370

#97	1,19E-24	NC_071499.1	2E+07	2E+07	6	xm_054091585.1	0
#98	1,24E-24	NC_071503.1	2E+07	2E+07	1	xr_008455374.1	41695
#99	1,49E-24	NC_071500.1	1E+08	1E+08	0		
#100	2,10E-24	NC_071498.1	8E+07	8E+07	1	xm_054109117.1	-82903

Tabela S 3. Termos e caminhos GO enriquecidos.

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathways (click for details)
3.5E-21	73	1617	3.7	Cellular developmental process
6.4E-21	72	1605	3.6	Cell differentiation
6.9E-21	64	1263	4.1	Cell development
2.6E-20	72	1661	3.5	System development
2.8E-20	67	1439	3.8	Anatomical structure morphogenesis
3.0E-20	60	1146	4.2	Animal organ development
2.1E-19	53	916	4.7	Epithelium development
2.9E-19	55	1002	4.4	Tissue development
3.9E-19	56	1050	4.3	Nervous system development
1.5E-17	50	895	4.5	Neurogenesis
1.8E-16	43	689	5.1	Tube development
2.5E-16	40	594	5.5	Developmental protein
2.3E-15	46	854	4.4	Generation of neurons
2.6E-15	42	708	4.8	Animal organ morphogenesis
2.6E-14	40	682	4.8	Neuron differentiation
8.6E-14	36	565	5.2	Morphogenesis of an epithelium

8.8E-14	43	824	4.2	Phosphoprotein
1.1E-13	35	538	5.3	Imaginal disc development
1.8E-13	34	514	5.4	Anatomical structure formation involved in morphogenesis
1.8E-13	36	583	5	Tissue morphogenesis

Tabela S 4. Sítios com valores significantes de m6A estimados por mCaller e com probabilidade corrigida por FDR para eventos múltiplos.

Chrom	position	sequence	Kruskal-Wallis	chi-quadrado	FDR
NC_071498.1	1826080	TGTAGMTCTTA	0	0	4.20787E-06
NC_071498.1	138114037	TGAAGMTCCCG	0	0	4.20787E-06
NC_071501.1	51867746	AGCGGMTCTTC	0.109091	3.77737E-08	1.26236E-05
NC_071503.1	9058626	TAAAGMTCCTG	0.109091	3.77737E-08	1.26236E-05
NC_071503.1	47468364	GGCAGMTCTAT	0.109091	3.77737E-08	1.26236E-05
NC_071503.1	90956099	TGGCGMTCCTG	0.109091	3.77737E-08	1.26236E-05
NC_071502.1	53324820	AGAAGMTCCGC	0.218182	8.175E-07	2.94551E-05
NC_071498.1	15250277	TTACGMTCTGT	0.245455	1.37355E-06	3.36629E-05
NC_071500.1	122845920	ACAAGMTCTAT	0.245455	1.37355E-06	3.36629E-05
NC_071503.1	81388443	TTAAGMTCTAC	0.245455	1.37355E-06	3.36629E-05
NW_026530159.1	1674	GGAAGMTCTAC	0.245455	1.37355E-06	3.36629E-05
NC_071498.1	67772463	TCGAGMTCGAG	0.272727	2.18236E-06	5.04944E-05
NC_071499.1	119482162	ACAGGMTCGAA	0.272727	2.18236E-06	5.04944E-05
NC_071503.1	64176033	GAAAGMTCACA	0.272727	2.18236E-06	5.04944E-05
NC_071503.1	64176960	ACTAGMTCAAT	0.272727	2.18236E-06	5.04944E-05
NW_026530051.1	11363	GACAGMTCTTT	0.272727	2.18236E-06	5.04944E-05
NC_071499.1	31052659	GCCCCMTCGCA	0.3	3.31411E-06	7.15338E-05
NC_071499.1	31052684	GTGGGMTCGCT	0.3	3.31411E-06	7.15338E-05
NC_071500.1	96142479	GAAGGMTCTTC	0.327273	4.84832E-06	7.99495E-05
NC_071500.1	103790099	AGAAGMTCCAA	0.354545	6.87383E-06	8.41574E-05
NC_071498.1	28566808	ATATGMTCGCA	0.436364	1.69254E-05	8.83652E-05
NC_071498.1	146664775	GCAAGMTCTGT	0.436364	1.69254E-05	8.83652E-05
NC_071500.1	13633035	ACTAGMTCCAA	0.436364	1.69254E-05	8.83652E-05
NC_071501.1	99493615	ATTTGMTCTTA	0.436364	1.69254E-05	8.83652E-05
NC_071503.1	13595353	TTAAGMTCGGC	0.436364	1.69254E-05	8.83652E-05
NW_026530033.1	226125	GAAGGMTCATT	0.436364	1.69254E-05	8.83652E-05
NC_071498.1	60856339	AACAGMTCTTG	0.463636	2.1989E-05	0.000113612
NC_071499.1	105226062	CTCCGMTCTTT	0.463636	2.1989E-05	0.000113612
NC_071503.1	88685852	GGTCGMTCCAA	0.463636	2.1989E-05	0.000113612
NC_071499.1	79321631	TACAGMTCGTC	0.490909	2.81254E-05	0.000126236
NC_071500.1	25379156	CGCGGMTCTCT	0.490909	2.81254E-05	0.000126236
NC_071498.1	186745	AACAGMTCTTG	0.518182	3.54775E-05	0.000134652
NC_071498.1	14402702	TTGGGMTCTCA	0.518182	3.54775E-05	0.000134652
NC_071498.1	17086947	GCCAGMTCGAT	0.518182	3.54775E-05	0.000134652
NC_071498.1	45439806	TGCTGMTCAGA	0.518182	3.54775E-05	0.000134652
NC_071498.1	59201671	CTTCGMTCAGG	0.518182	3.54775E-05	0.000134652
NC_071498.1	138114869	GGCTGMTCGAC	0.518182	3.54775E-05	0.000134652
NC_071498.1	182317859	TTTAGMTCTTT	0.518182	3.54775E-05	0.000134652
NC_071499.1	3234446	TAAAGMTCTGA	0.518182	3.54775E-05	0.000134652
NC_071499.1	3234882	GCAAGMTCTGT	0.518182	3.54775E-05	0.000134652
NC_071499.1	35125813	AAAAGMTCATG	0.518182	3.54775E-05	0.000134652
NC_071499.1	61638097	TAGAGMTCATT	0.518182	3.54775E-05	0.000134652

NC_071499.1	124350859	TCAGGMTCCCTA	0.518182	3.54775E-05	0.000134652
NC_071500.1	11296252	AGCAGMTCTTC	0.518182	3.54775E-05	0.000134652
			Kruskal-	chi-	
Chrom	position	sequence	Wallis	quadrado	FDR
NC_071500.1	16756229	TGGGGMTCTGC	0.518182	3.54775E-05	0.000134652
NC_071500.1	25378311	GGCAGMTCCGGT	0.518182	3.54775E-05	0.000134652
NC_071500.1	64444362	TAAAGMTCTTA	0.518182	3.54775E-05	0.000134652
NC_071500.1	90806476	ATGAGMTCTCA	0.518182	3.54775E-05	0.000134652
NC_071500.1	127300239	AAACGMTCTGA	0.518182	3.54775E-05	0.000134652
NC_071501.1	49004891	AGAAGMTCCCC	0.518182	3.54775E-05	0.000134652
NC_071501.1	92382548	GGTAGMTCCAC	0.518182	3.54775E-05	0.000134652
NC_071502.1	325984	CGTAGMTCTCC	0.518182	3.54775E-05	0.000134652
NC_071502.1	37101796	CATAGMTCTCC	0.518182	3.54775E-05	0.000134652
NC_071502.1	97183577	GCCGGMTCGAA	0.518182	3.54775E-05	0.000134652
NC_071502.1	103111469	TGCAGMTCAAT	0.518182	3.54775E-05	0.000134652
NC_071502.1	103111595	CCCAGMTCTTG	0.518182	3.54775E-05	0.000134652
NC_071502.1	125970450	AATGGMTCAGA	0.518182	3.54775E-05	0.000134652
NC_071503.1	7269833	ATTAGMTCAAC	0.518182	3.54775E-05	0.000134652
NC_071503.1	10397106	AAGAGMTCCAA	0.518182	3.54775E-05	0.000134652
NC_071503.1	11839822	GAAGGMTCCGGA	0.518182	3.54775E-05	0.000134652
NC_071503.1	12550995	GCCGGMTCTCC	0.518182	3.54775E-05	0.000134652
NC_071503.1	14059767	TAGGGMTCCAC	0.518182	3.54775E-05	0.000134652
NC_071503.1	21967698	TCAAGMTCCCTG	0.518182	3.54775E-05	0.000134652
NC_071503.1	33907996	GTAAGMTCTTA	0.518182	3.54775E-05	0.000134652
NC_071503.1	53218726	CTAGGMTCCGGC	0.518182	3.54775E-05	0.000134652
NC_071503.1	56779321	TTCCGMTCCAT	0.518182	3.54775E-05	0.000134652
NC_071503.1	57884114	TGGTGMTCGAT	0.518182	3.54775E-05	0.000134652
NC_071503.1	90954864	ACGAGMTCAAT	0.518182	3.54775E-05	0.000134652
NC_071504.1	5819778	TTCTGMTCGAT	0.518182	3.54775E-05	0.000134652
NW_026530034.1	217200	CTGCGMTCCAC	0.518182	3.54775E-05	0.000134652
NC_071498.1	5792290	GCTAGMTCCGGA	0.545455	4.41966E-05	0.000298759
NC_071498.1	39889282	CTCGGMTCTCT	0.545455	4.41966E-05	0.000298759
NC_071498.1	69999707	TTCAGMTCATT	0.545455	4.41966E-05	0.000298759
NC_071498.1	153012600	TCTGGMTCTTA	0.545455	4.41966E-05	0.000298759
NC_071499.1	193410	TCTAGMTCTCG	0.545455	4.41966E-05	0.000298759
NC_071499.1	12361440	GATGGMTCCAT	0.545455	4.41966E-05	0.000298759
NC_071499.1	119189741	CTTCGMTCTCT	0.545455	4.41966E-05	0.000298759
NC_071499.1	119193235	GCCAGMTCGAT	0.545455	4.41966E-05	0.000298759
NC_071499.1	123561062	ACAGGMTCGAA	0.545455	4.41966E-05	0.000298759
NC_071499.1	125246471	CTTCGMTCGAC	0.545455	4.41966E-05	0.000298759
NC_071499.1	126762232	TTAGGMTCCCTG	0.545455	4.41966E-05	0.000298759
NC_071499.1	132107107	TATAGMTCTTG	0.545455	4.41966E-05	0.000298759
NC_071500.1	16388085	TCAGGMTCGCT	0.545455	4.41966E-05	0.000298759
NC_071500.1	18989862	TATCGMTCTGG	0.545455	4.41966E-05	0.000298759
NC_071500.1	40279546	TTAAGMTCCAG	0.545455	4.41966E-05	0.000298759
NC_071500.1	107514814	CTTCGMTCAGG	0.545455	4.41966E-05	0.000298759
NC_071500.1	127300049	GGCAGMTCAAA	0.545455	4.41966E-05	0.000298759
NC_071501.1	19804734	GCAAGMTCAGA	0.545455	4.41966E-05	0.000298759

NC_071501.1	53287732	TGGAGMTCCGG	0.545455	4.41966E-05	0.000298759
NC_071501.1	54345563	GACTGMTCCGGC	0.545455	4.41966E-05	0.000298759
			Kruskal-	chi-	
Chrom	position	sequence	Wallis	quadrado	FDR
NC_071501.1	92381888	GAATGMTCCCA	0.545455	4.41966E-05	0.000298759
NC_071502.1	60535164	AACCGMTCTTG	0.545455	4.41966E-05	0.000298759
NC_071502.1	103111400	GGCAGMTCCGGT	0.545455	4.41966E-05	0.000298759
NC_071502.1	109513794	CCTGGMTCCAGG	0.545455	4.41966E-05	0.000298759
NC_071502.1	119095163	TGATGMTCAAA	0.545455	4.41966E-05	0.000298759
NC_071503.1	2989636	GACGGMTCAAT	0.545455	4.41966E-05	0.000298759
NC_071503.1	8239720	AAAGGMTCATC	0.545455	4.41966E-05	0.000298759
NC_071503.1	47933366	TCTTGMTCCGGC	0.545455	4.41966E-05	0.000298759
NC_071503.1	57487391	TTTAGMTCCGGG	0.545455	4.41966E-05	0.000298759
NC_071503.1	64080366	TCAAGMTCAAC	0.545455	4.41966E-05	0.000298759
NC_071503.1	64140252	AAATGMTCAAA	0.545455	4.41966E-05	0.000298759
NC_071503.1	77779993	GTCAGMTCCGTC	0.545455	4.41966E-05	0.000298759
NC_071503.1	83754160	GTCAGMTCCGTC	0.545455	4.41966E-05	0.000298759
NC_071503.1	87869649	AGTAGMTCTTG	0.545455	4.41966E-05	0.000298759
NC_071503.1	94922282	AATTGMTCCAT	0.545455	4.41966E-05	0.000298759
NC_071504.1	4057596	CCTTGMTCCGTT	0.545455	4.41966E-05	0.000298759
NW_026530033.1	264599	CGGAGMTCTCC	0.545455	4.41966E-05	0.000298759
NW_026530051.1	78901	TTTAGMTCAAG	0.545455	4.41966E-05	0.000298759
NW_026530086.1	6017	CAATGMTCTTT	0.545455	4.41966E-05	0.000298759
NW_026530134.1	8596	CCAAGMTCTGT	0.545455	4.41966E-05	0.000298759
NC_071499.1	31052976	TCGGGMTCTTC	0.572727	5.4442E-05	0.000467073
NC_071498.1	40996205	TCAAGMTCCCTG	0.6	6.63811E-05	0.000471281
NC_071498.1	70556006	TTAAGMTCTAC	0.627273	8.0189E-05	0.000475489
NC_071499.1	117693999	TTAAGMTCTAC	0.627273	8.0189E-05	0.000475489
NW_026530033.1	112126	TTTAGMTCAAG	0.627273	8.0189E-05	0.000475489
NC_071503.1	2515881	TGAAGMTCCCTT	0.654545	9.60482E-05	0.000488113
NC_071503.1	88736358	TAAAGMTCCAA	0.654545	9.60482E-05	0.000488113
NW_026530041.1	63413	GACAGMTCTTT	0.654545	9.60482E-05	0.000488113
NC_071498.1	67772624	GCAAGMTCTTC	0.681818	0.000114148	0.000500736
NC_071499.1	113103800	CAAAGMTCCCTT	0.736364	0.000157865	0.000504944
NC_071499.1	113103811	AAAAGMTCCAGA	0.736364	0.000157865	0.000504944
NC_071503.1	48602223	CACAGMTCTTT	0.736364	0.000157865	0.000504944
NC_071503.1	48629723	AACAGMTCTTG	0.736364	0.000157865	0.000504944
NC_071499.1	124353390	ACCCGMTCTCC	0.763636	0.000183893	0.000521776
NC_071503.1	5866004	AAACGMTCCGT	0.763636	0.000183893	0.000521776
NC_071503.1	8927301	TCAAGMTCCCTG	0.763636	0.000183893	0.000521776
NW_026530157.1	4140	TCAGGMTCTCC	0.763636	0.000183893	0.000521776
NC_071499.1	30021876	GTAGGMTCCGCT	0.790909	0.000212988	0.000538607
NC_071499.1	31052144	AGTCGMTCCAGC	0.790909	0.000212988	0.000538607
NC_071503.1	47469460	GGAAGMTCTGC	0.790909	0.000212988	0.000538607
NC_071500.1	122845014	AACGGMTCCAGA	0.818182	0.00024537	0.000551231
NC_071501.1	65876400	ATCAGMTCCGT	0.818182	0.00024537	0.000551231
NC_071503.1	90077616	CGGTGMTCTTC	0.818182	0.00024537	0.000551231
NW_026530033.1	265022	CGGAGMTCCGCC	0.818182	0.00024537	0.000551231

NW_026530051.1	21103	CGGAGMTCGCC	0.818182	0.00024537	0.000551231
NW_026530051.1	79531	GACAGMTCTTT	0.818182	0.00024537	0.000551231

Figuras

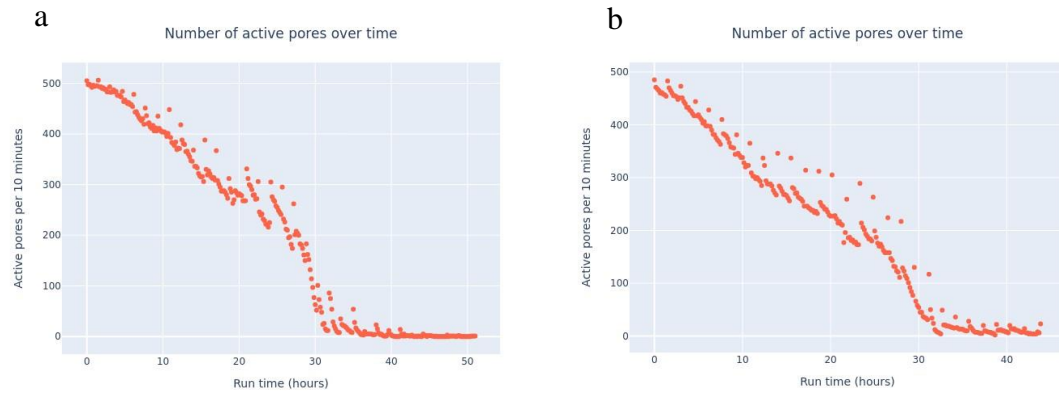


Figura S 1. Decaimento do número de poros por intervalo de tempo. (a) biblioteca 1; (b) biblioteca 2.

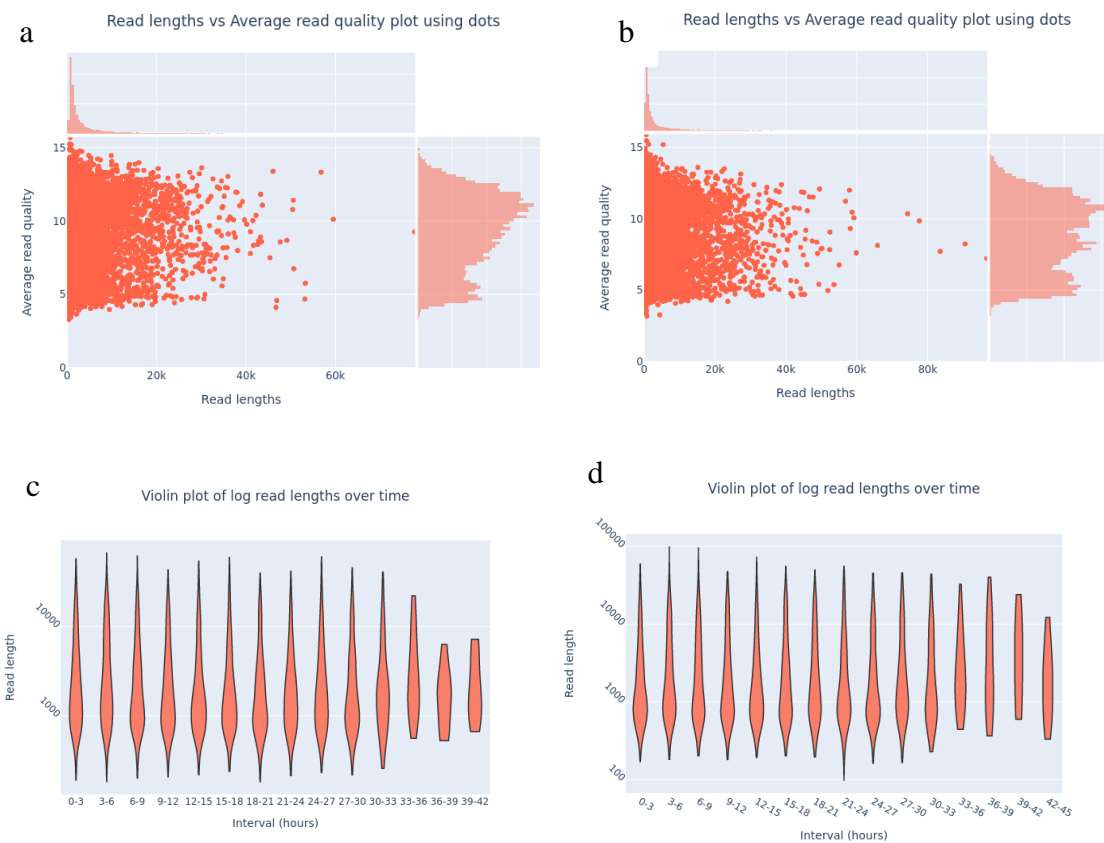


Figura S 2. Comprimento de leitura vs qualidade de leitura média usando pontos. (a) biblioteca 1 (b) biblioteca 2; e comprimento de leitura log ao longo do tempo. (c) biblioteca 1; (d) biblioteca 2.

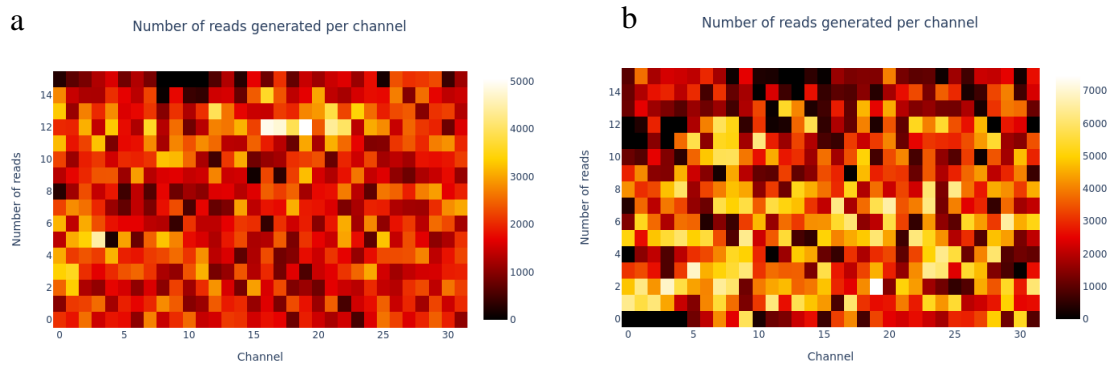


Figura S 3. Número de reads gerados por canal. (a) biblioteca 1; (b) biblioteca 2.

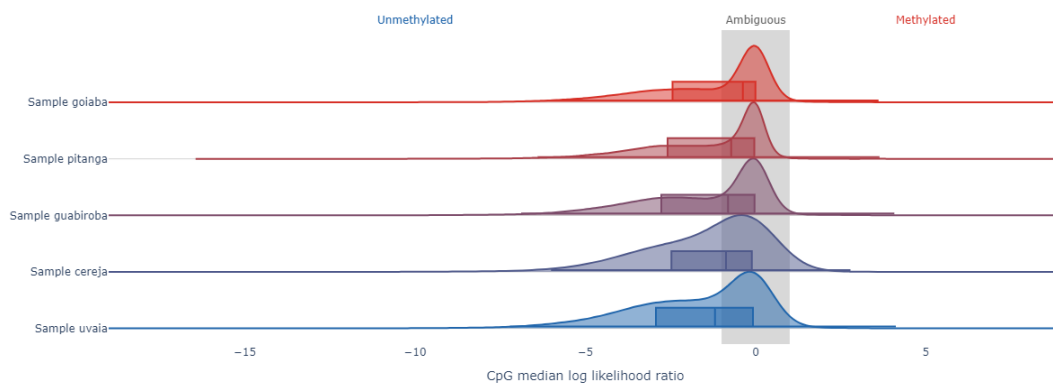


Figura S 4. Distribuição da razão log-verossimilhança de metilação de CpG para intervalos significativos de CpG.

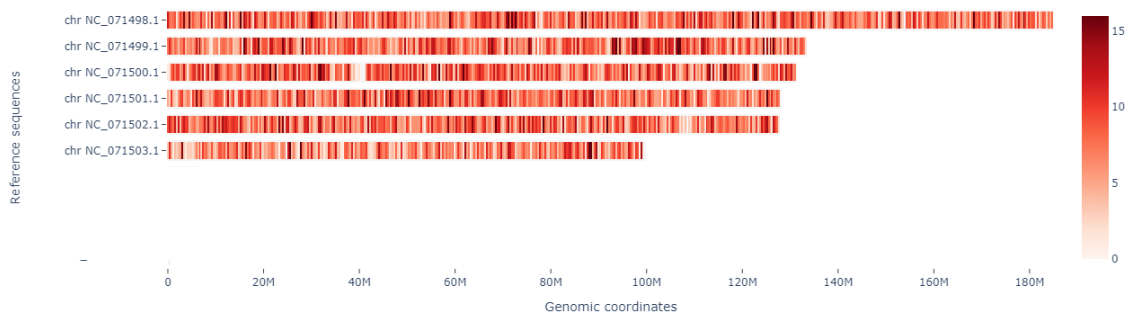


Figura S 5. Distribuição de intervalos CpG significativos por intervalos cromossômicos.

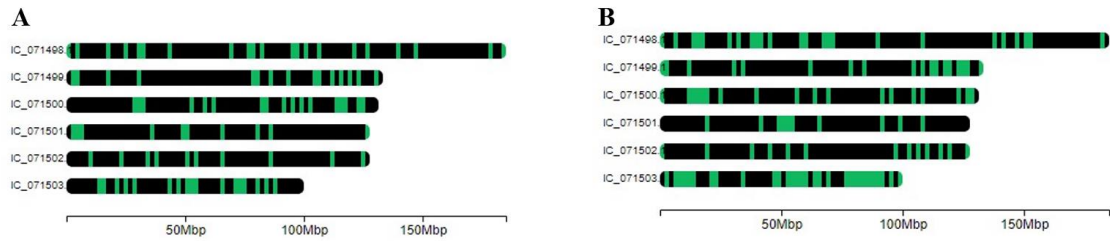


Figura S 6. (A) Distribuição ao longo do genoma para as 100 regiões mais significativas de (DMRs) para metilação 5mC de moscas que emergiram de diferentes frutos. (B) Distribuição ao longo do genoma de (DMRs) para metilação 6mA de moscas que emergiram de diferentes frutos.

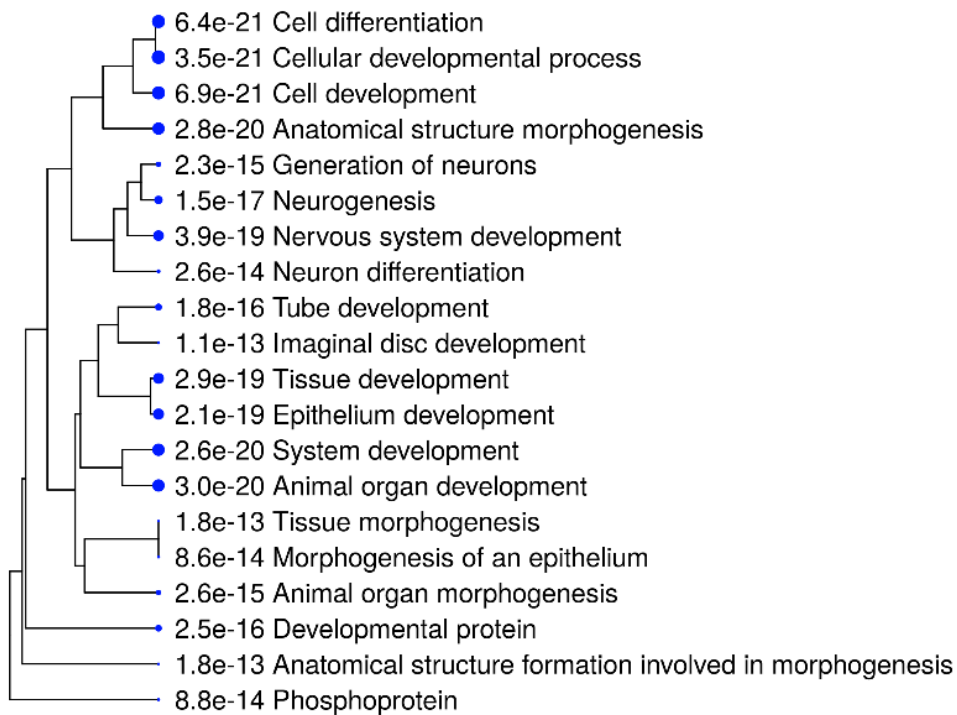


Figura S 7. Termos GO relacionados através de uma árvore de cluster hierárquica.