



UNIVERSIDADE DE SÃO PAULO – USP
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO – FMRP
DEPARTAMENTO DE GENÉTICA

**As contribuições da metilação diferencial no DNA para os processos sexo-
específico e mudanças de ambientes hospedeiros em *Drosophila gouveai* (cluster
Drosophila buzzatii)**

ADRIANO SILVA DOS SANTOS

RIBEIRÃO PRETO

2022

“Versão corrigida. A versão original está disponível tanto na Biblioteca da Unidade que abriga o Programa quanto na Biblioteca Digital de Teses e Dissertações da USP (BDTD)”

ADRIANO SILVA DOS SANTOS

As contribuições da metilação diferencial no DNA para os processos sexo-específico e mudanças de ambientes hospedeiros em *Drosophila gouveai* (cluster *Drosophila buzzatii*)

Tese apresentada ao Departamento de Genética, da Faculdade de Medicina de Ribeirão, como requisito para a obtenção do título de Doutor em Ciências.

Área de concentração: Genética.

Orientador (a): Profa Dra Maura Helena Manfrin.

Coorientador (a): Profa Dra Ester Silveira Ramos.

RIBEIRÃO PRETO

2022

Autorizo a reprodução total ou parcial deste trabalho para fins de estudo e didático, desde que citada a fonte.

FICHA CATALOGRÁFICA

Santos, Adriano Silva dos

As contribuições da metilação diferencial no DNA para os processos sexo-específico e mudanças de ambientes hospedeiros em *Drosophila gouveai* (cluster *Drosophila buzzatii*). Adriano Silva dos Santos; Orientadora: Maura Helena Manfrin. Coorientadora: Ester Silveira Ramos. Ribeirão Preto, 2022.

203p.: 21il.; 30 cm

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Genética.

Versão corrigida.

1. *Drosophila cactofflica*. 2. Metilação de DNA. 3. Sexo-específico. 4. Sítios CpGs. 5. Elemento transponível. 6. Cactos hospedeiros. 7. Ambiente ecológico.

“Esforça-te, e tem bom ânimo....porque o Senhor teu Deus é contigo, por onde quer que andares. (Josué 1:9)”

FOLHA DE APROVAÇÃO

Adriano Silva dos Santos

Título da tese: As contribuições da metilação diferencial no DNA para os processos sexo-específico e mudanças de ambientes hospedeiros em *Drosophila gouveai* (cluster *Drosophila buzzatii*).

Tese apresentada ao Departamento de Genética, da Faculdade de Medicina de Ribeirão, como requisito para a obtenção do título de Doutor em Ciências.

Área de concentração: Genética.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Dedico à minha família, com carinho.

AGRADECIMENTOS

À Deus por me ajudar, ensinar, sustentar, e possibilitar a realização deste projeto.

À Profa Dra Maura H. Manfrin pelo material de estudo, sugestões, e orientação.

À Profa Dra Ester S. Ramos pela oportunidade de aprender sobre epigenética, sugestões, e orientações.

Ao Prof. Fábio pela ajuda financeira, conversas sobre ciência, e biologia de Drosophilas.

Aos membros da Comissão de Pesquisa (FMRP/USP); profs Dr. Carlos Scrideli, Dr. José Alves-Filho, Dr. Eurico A. Neto, também ao Gilson e a Sandra, por ter oportunizado aprender um pouco mais sobre projetos, pesquisa, organização de estruturas físicas, carreira acadêmica, bem como os conhecimentos aprendidos nas reuniões, e o excelente ambiente.

Aos membros da Comissão da Pós-Graduação (FMRP/USP); Profs Dra Cristina Del-Ben, Dr. Luis Lambert, Dr. Luiz Navegantes, Dra Ana Rosa e Silva e a todos desta comissão, pela oportunidade de aprender mais sobre a organização da estrutura da pós-graduação, carreira profissional, e adquirir bons conhecimentos, e motivações na área profissional.

À técnica Ana Paula pelo excelente trabalho, e suporte na pós-graduação, junto à pós-graduação.

Aos membros da Comissão da Pós-Graduação em Genética (FMRP/USP); Profs Dr. Klaus Hartfelder, Dra Ester Ramos, Dr. Celso Mendes-Júnior, Dra Nilce Rossi, Dra Elza Tiemi, Dr. Aguinaldo Simões, Dr. Victo Evangelista, e Dra Aparecida Fontes, que me permitiram conhecer um pouco mais sobre organização da estrutura física da PGGen, carreira profissional, e administrativa.

Aos membros da Comissão de Direitos Humanos (FMRP/USP); Dra Maria Panuncio, Dra Carolina, Dra Luciane Loures, Márcia, Elke, por ter criado vivências acadêmicas com aprendizados sobre questões importantes, como bem estar, acolhimento, e suporte as pessoas na universidade.

Aos profs. Drs. Klaus Hartfelder, Francis Nunes e Dra Maíra Martins pelas críticas e sugestões ao trabalho, na banca de qualificação.

Aos colegas no laboratório de Genética Evolutiva: Dora B. Leal, Ana B. Zichinelli, Marcelo Moretto.

Ao técnico Paulo Epifânio pela amizade e suporte no laboratório de Genética Evolutiva.

Aos colegas no laboratório de Genética Bioquímica: Maria Luisa, Juliet Figueiredo, Igor Caetano, e Aline Brazorotto, pela amizade e ajuda com técnicas de biologia molecular.

À técnica Maria Canas e Cláudia Wiesel, no laboratório de Genética Bioquímica (FMRP/USP), por toda contribuição com técnicas de biologia molecular.

Aos colegas no laboratório de Epigenética e Reprodução (FMRP/USP): Jorge Marchesi, Carolina Verruma, pela ajuda com metodologias para estudos com epigenética.

À profa Dra Zilá Simões (FFCLRP/USP) pela disponibilidade de usar equipamentos no Laboratório de Biologia e Desenvolvimento de abelhas.

Ao prof. Me. Anderson Hander pelas sugestões acadêmicas, profissionalismos, e dicas com estruturas de linguagens, e revisões.

À profa Dra Karina Faria pela amizade, apoio, e colaborações científicas.

Aos colegas no Laboratório de Biologia e Desenvolvimento de abelhas: Franciene Rabiço, Fábio Barbin, Thiago Depintor, Luana Bataglia, Danielle Lucena, pelas ajudas e colaborações com metodologias e dicas de protocolos.

À Dra Vera Figueiredo, técnica no laboratório de Biologia e Desenvolvimento de abelhas, pelas ajudas com soluções e quantificações de amostras.

À Dra Agda pelo excelente trabalho junto ao Centro de Recursos Biológicos e Biologia Genômica (CreBio), e por sempre ajudar nos procedimentos de sequenciamento de DNA.

Ao Mendelson Mazucato, técnico no Laboratório de Genética Molecular e de Microorganismos (FMRP/USP), pela boa convivência, ajudas com protocolos, e bons conselhos.

À profa Dra Maria Fontes (FMRP/USP) pela ajuda com protocolos de Biologia Molecular.

Ao prof. Dr. Tiago Campos (FFCLRP/USP) pelas discussões sobre ciências e aprendizados acadêmicos.

À minha família, vó Maria, vô Joaquim (*in memoriam*), mãe Deusirene, pai Jorge, irmão André, primo Jhonatham, pelo suporte, carinho, e ajuda de sempre em diversos momentos da vida.

À minha namorada Isa e a sua família por me ajudarem no momento pós-COVID, e vivências familiares.

Aos meus colegas e irmãos amigos: Eduardo Bianchi, Karol, irmã Francis, Dani Mônico, Arthur, Gabriel Lavez, Samuel, Amanda, Davi, Daniel, Rafa, Lucas, Hugo, Victor, Diego, Luiz, Henrique, Solange, Gil, Luana, e a todos/as que contribuíram para uma boa convivência durante esses anos.

Aos membros do Departamento de Genética (FMRP/USP) por disponibilizar os seus recursos.

À Susie e ao Gustavo pelo suporte na secretária do programa de Genética-FMRP/USP.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A fundação de pesquisa FAPESP e as agências da CAPES e CNPq pelos auxílios financeiros aos estudos realizados no Laboratório de Genética Evolutiva.

A Universidade de São Paulo (USP) e a Faculdade de Medicina de Riberião Preto (FMRP) pelos auxílios na formação profissional na pós-graduação.

A todos que contribuíram direta ou indiretamente para a realização deste projeto. Obrigado!

Sumário

RESUMO.....	xx
ABSTRACT	xxi
1.Introdução	23
1.1 Considerações iniciais.....	23
1.2 Cluster <i>Drosophila buzzatii</i>	23
1.3 Biologia Reprodutiva e Desenvolvimento em <i>Drosophila</i>	28
1.4 Plasticidade de fenótipos em resposta à mudança de ambiente	32
1.5 Mecanismos epigenéticos em <i>Drosophila</i>	36
1.6 Modelo biológico <i>Drosophila gouveai</i> (cluster <i>D. buzzatii</i>).....	42
2.Hipótese de estudo.....	48
3.Objetivos.....	50
3.1 Objetivo Geral	50
3.2 Objetivos específicos	50
4.Material e Métodos	52
4.1 <i>Stock</i> de <i>Drosophila</i>	52
4.2 Enzimas	52
4.3 Tampões, Soluções, e Meios	52
4.4 Kits.....	54
4.5 Experimentos: ambientes heterogêneos.....	54
4.6 Extrações de DNA genômico	55
4.7 Quantificação de DNA genômico.....	56
4.8 Ensaio <i>Methylation Sensitive Amplification Polymorphism</i> (MSAP).....	56
4.8.1 Digestões enzimáticas	56
4.8.2 Ligação com adaptadores	57
4.8.3 Amplificações Pré-Seletivas	57
4.8.4 Amplificações Seletivas	57
4.9 Eletroforese	59
4.9.1 Gel de poliacrilamida	59
4.9.2 Gel de Agarose.....	59
4.10 Interpretações do padrão de metilação.....	60
4.11 Isolamento de regiões metiladas.....	61
4.12 Razões Relativas de Metilação (RRM).....	62
4.13 Análise de dados MSAP	63
4.14 Teste de análise de variância ANOVA	63
4.15 Sequenciamento de DNA.....	64
4.16 Análises de Bioinformática	64
4.17 Análises funcionais de Ontologia dos Genes (GO)	65
5.Resultados.....	67
5.1 Análises dos perfis de metilação	67
5.2 Perfis Moleculares do DNA genômico	68
5.3 Razões Relativas de Metilação.....	69
5.4 Perfis de Metilação – MSAP.....	72
5.5 Análises de Variações – ANOVA	77
5.7 Sequenciamento de DNA alvo de metilação.....	82
5.8 Regiões regulatórias moleculares	92
5.8.1 Regiões de sítio de início transcricional	92
5.8.2 Predição de regiões promotoras	92

5.8.3 Análises funcionais de Ontologia dos Genes (GO).....	94
6.Discussão	100
6.1 Padrões de Metilação no DNA.....	100
6.2 Perfis de metilação de DNA.....	101
6.3 Sequências DNA e mecanismos de regulação gênica	103
6.4 Metilação de DNA e Reprodução.....	106
6.5 Ecologia	110
7.Conclusões	112
8.Referências Bibliográficas	115
Apêndices e Anexos	126
Apêndice A. Verificação da integridade de moléculas de DNA genômico (gel agarose 0.8%) (A), e digestões enzimáticas (B) (gel poliacrilamida 10%), amostra de <i>D. gouveai</i>	127
Apêndice B. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AAC)/HM(TAG), com DNA genômico de pupas em meio padrão (Q), <i>C. hildmannianus</i> (C), e <i>P. machrisii</i> (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.	128
Apêndice C. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AGG)/HM(TGA), com DNA genômico de adultas e fêmeas que se desenvolveram em meio padrão (Q), <i>C. hildmannianus</i> (C), e <i>P. machrisii</i> (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.	129
Apêndice D. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(ACT)/HM(TTG), com DNA genômico de adultos e machos que se desenvolveram em meio padrão (Q), <i>C. hildmannianus</i> (C), e <i>P. machrisii</i> (P). Marcador de peso molecular 50pb (M). Controle negativo (C-) da PCR.	130
Apêndice E. Verificação de produtos de PCR (gel de agarose 1%), com re-amplificações amostras diferenciais de metilação de DNA, em tecidos reprodutivos de moscas em <i>D. gouveai</i>	131
Apêndice F. Matriz de dados binários MSAP com DNA genômico de pupas, em meio padrão (Q), <i>C. hidmannianus</i> (C), e <i>P. machrisis</i> (P).....	132
Apêndice G. Matriz de dados binários de MSAP com DNA genômico de moscas adultas e fêmeas, em meio padrão (Q), <i>C. hidmannianus</i> (C),e <i>P. machrisis</i> (P).	134
Apêndice H. Matriz de dados binários de MSAP com DNA genômico de moscas adultas e machos, em meio padrão (Q), <i>C. hidmannianus</i> (C),e <i>P. machrisis</i> (P).	136
Apêndice I. Predições de bioinformática para regiões promotoras em sequências de DNA.....	138
Apêndice J. Estrutura quaternária da proteína DNA metiltransferase 2 (Dnmt2) em <i>D. buzzatii</i> . Esta sequência foi obtido do seu genoma sequenciado e disponível em https://dbuz.uab.cat/welcome.php , Dnmt2 - scaffold27:847239-848215, e a proteína foi modelada usando o I-TASSER Server disponível em https://zhanglab.cmb.med.umich.edu/I-TASSER . A estrutura de Dnmt2 é composta por α -hélices (cor merge) e β -folhas (cor amarelo), e espaçadores (cores em branco e azul).....	139
Anexo A – Comitê de Ética	140
Anexo B – Manuscrito como primeiro autor, com os dados do projeto do doutorado, a ser submetido em revista científica.....	141
Anexo C – Outra produção Científica. Manuscrito como primeiro autor com colaboração acadêmica, mas externo ao projeto de doutorado, embora realizado durante o curso.	173

LISTA DE FIGURAS

- Figura 1.** Relações da morfologia do aparelho reprodutor masculino (edeago) nas espécies do cluster *D. buzzatii*, *D. koepferae* (isolinhagem B25D7, La Rioja, Argentina), *D. antonietae* (Serrana, São Paulo), *D. gouveai* (Analândia, São Paulo), *D. serido* (B53Q5, Milagres, Bahia), *D. seriema* (B95C3, Serra do Cipó, Minas Gerais), e *D. borborema* (D54B5, Grão Mogol, Minas Gerais). Adaptado de Tidon-Sklorz & Sene (2001).....26
- Figura 2.** Hipótese de relações filogenética com o marcador mitocondrial mtCOI no cluster *D. buzzatii*. Adaptado de Manfrin & Sene (2006).....26
- Figura 3.** Distribuição geográfica conhecida das espécies do cluster *Drosophila buzzatii* na América do Sul. As principais bacias hidrográficas, bem como a distribuição aproximada das principais áreas de florestas secas da América do Sul, estão apresentadas no mapa. FTSS: Florestas Tropicais Sazonalmente Secas. Fonte: Franco, 2009.....27
- Figura 4.** Padrão acústico do cortejo sonoro produzido por moscas machos (adultos) nas espécies do cluster *D. buzzatii*. Adaptada de Oliveira *et al.* (2013).....29
- Figura 5.** Ilustração do ambiente de desenvolvimento proposto por Waddington (1956). Esse modelo é conhecido como paisagem epigenética de Waddington. Célula “onipotente” do embrião em desenvolvimento que é representada como uma bola rolando por um “vale”. Essa bola pode encontrar vários pontos de ramificação. A escolha do ramo esquerdo ou direito depende, em certa medida, da forma local precisa da paisagem, e pode ter predisposições genéticas ou de influências de fatores ambientais. Modificado de Waddington (1956).....37
- Figura 6.** Ilustração com as hipóteses dos padrões de distribuição de m5C no embrião de *Drosophila*. Representação de duas regiões metiladas é indicada por meio de blocos em azul (A) e vermelho (B). No modelo 1, um subconjunto de núcleos é metilado na região A e outro na região B, em posições diferentes do embrião. No modelo 2, ambas as regiões são metiladas em um subconjunto de núcleos. No modelo 3, a metilação das duas regiões é distribuída, difusamente, por meios dos núcleos em todo o embrião. Esse modelo considera que todos os núcleos apresentam pouca metilação em ambas as regiões. Os três modelos apresentam padrões distintamente diferentes na distribuição da metilação, embora, ainda, não tenha sido testada a metilação em células específicas do embrião (como em todos os estudos até o momento), os padrões são indistinguíveis. Modificado de Boffelli *et al.* (2014).....42
- Figura 7.** Ilustrações das espécies de cactos *C. hildmaniannusi* (A) e *P. machrisis* (B).....45
- Figura 8.** Perfil de metilação obtido por meio da técnica MSAP com DNA de pupas, visualização em gel de poliacrilamida (10%), e com a indicação de quatro padrões analisados: ausência de metilação (+/+), metilação completa ou semi metilação de citosinas internas (-/+), e não informativa (-/-). M representa o marcador molecular (50pb - Promega®). As amostras estão representadas por Q (meio padrão de laboratório), C (meio de *Cereus hidmannianus*), e P (meio de *Pilosocereus machrisis*). Os ensaios enzimáticos estão representados por *HpaII* e *MspI*, que foram realizados em conjunto com *EcoRI*. C- representa o controle

negativo	da	reação	de
MSAP.....			67

Figura 9. Gráfico boxplot com razões relativas de metilação no DNA genômico no organismo (pupas) que se desenvolveram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*. Esses dados foram obtidos por meio de perfis moleculares gerados com os estados de metilação interna e externa de bases de citosinas representado (*HpaII-/MspI+*; *HpaII+/MspI-*), e ambos no total relativo de metilação, em cor vermelho.....70

Figura 10. Gráfico boxplot com razões relativas de metilação no DNA genômico no tecido reprodutivo (moscas adultas, fêmeas) que se desenvolveram e emergiram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*. Esses dados foram obtidos por meio de perfis moleculares gerados com os estados de metilação interna e externa de bases de citosinas representado (*HpaII-/MspI+*; *HpaII+/MspI-*), e ambos no total relativo de metilação, em cor rosa.....71

Figura 11. Gráfico boxplot com razões relativas de metilação no DNA genômico no tecido reprodutivo (moscas adultas, machos) que se desenvolveram e emergiram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*, com o segundo desenvolvimento consecutivo. Esses dados foram obtidos por meio de perfis moleculares gerados com os estados de metilação interna e externa de bases de citosinas representado (*HpaII-/MspI+*; *HpaII+/MspI-*), e ambos no total relativo de metilação, em cor azul.....71

Figura 12. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de pupas. Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*.....78

Figura 13. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de tecidos reprodutivos de fêmeas (TRF). Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*.....78

Figura 14. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de tecidos reprodutivos de machos (TRM). Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*.....79

Figura 15. Visualização em gel de poliacrilamida (10%) de perfil de metilação no DNA de amostras de tecidos reprodutivos em moscas fêmeas (TRF) e outras em machos (TRM), com corante de nitrato de prata. As amostras de DNA estão representadas por indivíduos criados no meio padrão de laboratório (Q), meio de cacto *C. hildmannianus* (C), e meio com *P. machrisii* (P). Os ensaios de MSAP foram realizados com as combinações E2-HM2 para TRF, e E2-HM1 para TRM, com uso de enzimas representadas por *EcoRI-HpaII* (H) e *EcoRI-MspI* (MI). Os produtos de PCR foram obtidos por meio da técnica MSAP, isolados os fragmentos de interesse (caixas em vermelho), e com o processo de reamplificação dessas regiões. Essas regiões são alvos de metilação externa de citosinas (bandas somente em *EcoRI-HpaII*), e metilação interna de citosinas (bandas somente em *EcoRI-MspI*), que seguiram para o processo de sequenciamento de DNA.

M representa o marcador molecular (50pb - Promega®). O controle negativo da reação de MSAP é indicado na lane (C-).....82

Figura 16. Alinhamento de sequências de nucleotídeos do gene *Piezo* em *Drosophila melanogaster*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA em moscas fêmeas criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).....88

Figura 17. Alinhamento de sequências de nucleotídeos do gene *eEF1alpha1* em *Drosophila melanogaster*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA em moscas fêmeas criadas *P. machrisii*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-HpaII*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).....88

Figura 18. Alinhamento de sequências de nucleotídeos que corresponde ao ncRNA LOC125776504 em *Bactrocera dorsalis*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Esta região foi amplificada em PCR, com amostras de DNA com moscas machos criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).....89

Figura 19. Alinhamento de sequências de nucleotídeos que corresponde ao elemento móvel *Galileo* em *Drosophila buzzatii*, com sequência completa (seq_t) e parcial (seq_P), e em *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA com moscas machos criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).....90

Figura 20. Dendograma obtido por meio do método Neighbor-Joining(NJ), com os dados das sequências de DNA alvo de metilação em *D. gouveai* (Query sequence_28961), com regiões gênicas disponíveis em *Drosophila buzzatii*. A região em *D. gouveai* corresponde ao TE *Galileo*. As sequências estão disponíveis na Tabela 11.....91

Figura 21. Estrutura dos dados com o método Neighbor-Joining(NJ), com o uso das sequências de DNA-alvos de metilação, em *Drosophila gouveai*. Essas amostras são provenientes de indivíduos que se desenvolveram em diferentes ambientes, com moscas fêmeas tais como meio padrão (GQO), cacto *C. hildmannianus* (GCO), e em *P. machrisii* (GCO). Uma parte das amostras representa indivíduos moscas machos que se desenvolveram em meio com cacto *C. hildmannianus* (GCT), e em *P. machrisii* (GCT). Os ensaios de MSAP foram realizados com enzimas representadas por *EcoRI-HpaII* (H) e *EcoRI-MspI* (M). Todas as sequências estão disponíveis na Tabela 11. Inferências estatísticas foram obtidas com 1000 *bootstrap*, com os valores indicados próximos aos cladogramas.....91

Figura 22. Network de proteínas com as interações entre eEF1A com a sinalização de transporte do núcleo citoplasma. A representação contém o Complexo do Poro Nuclear (NPC) com a cesta nuclear, anel do lúmen,

complexo Spoke, canal central, Simetria nups, complexo Nup358, e fibrilas citoplasmáticas. O complexo de transporte nuclear envolve as importinas (IPOA, IPOB), proteína adaptador (SPN1), exportinas (XPO, Ran, eEF1A, PHAX, CBC, NMD3). eEF1A é indicado pelo círculo vermelho. Complexo de junção-exon (EJC) é representado por núcleo interno EJC, escudo exterior ELC, e fatores de interação transiente. Complexo de transcrição-exportação (TREX), que envolve o subcomplexo THO.....98

LISTA DE TABELAS

- Tabela 1.** Relação de *primers* usados para as análises de MSAP, com adaptadores, pré-amplificações, e amplificações seletivas, com as suas respectivas combinações.....58
- Tabela 2.** Padrão de metilação de DNA genômico com as enzimas *HpaII* e *MspI*.....61
- Tabela 3.** Perfil molecular de DNA genômico em amostras de pupas, com a segunda geração consecutiva criada em meio padrão de laboratório (Padrão L), meio dos cactos *C. hidmannianus*, e outras com *P. machrisis*.....68
- Tabela 4.** Perfil molecular de DNA genômico em amostras de tecidos reprodutivos, com a segunda geração consecutiva de moscas fêmeas, que se desenvolveram no meio padrão de laboratório (Padrão L), nos meios de cactos *C. hidmannianus*, e e outras com *P. machrisis*.....68
- Tabela 5.** Perfil molecular de DNA genômico em amostras de tecidos reprodutivos, com a segunda geração consecutiva de moscas machos, que se desenvolveram no meio padrão de laboratório (Padrão L), nos meios de cactos *C. hidmannianus*, e outros com *P. machrisis*.....69
- Tabela 6.** Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de bases (pb), em pupas de *D. gouveai*, com desenvolvimento em meio padrão (Q), meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).....74
- Tabela 7.** Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de base (pb), em moscas fêmeas de *D. gouveai*, com desenvolvimento em meio padrão (Q), meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).....75
- Tabela 8.** Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de base (pb), em moscas machos de *D. gouveai*, com desenvolvimento em meio padrão (Q), nos meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).....76
- Tabela 9.** Teste de *One-way* ANOVA e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos (pupas) que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*).....79

Tabela 10. Teste de *One-way ANOVA* e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*), em moscas fêmeas.....80

Tabela 11. Teste de *One-way ANOVA* e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*), em moscas machos.....81

Tabela 12. Amostras obtidas as sequências de DNA alvos de metilação, em *Drosophila gouveai*, com dados moleculares sobre extensão em Pares de Bases nucleotídicas (pb), Porcentagem de Bases de Citosinas (PBC), análise de similaridade no BLASTn, Localização Genômica (LG), Similaridade e Número de acesso no Banco FlyBase ou GenBank.....85

Tabela 13. Análises de Bioinformática com buscas de regiões genômicas com similaridade (BLASTn), regiões transcricionais (BLASTx), predição de Sítios de Início de Transcrição, e Regiões Promotoras, nas sequências alvo de metilação no DNA de moscas adultas de *D. gouveai*.....87

Tabela 14. Análise do *Termo Gene Report* com as ferramentas do programa David v.6.8. Os genes são os mesmos alvos de metilação e sequenciados em *D. gouveai*. As informações dos genes são apresentadas por categoria, termo/função, Fase *Discovery Rate* (FDR), e *P-value*.....96

Tabela 15. Análise Descritiva dos genes por meio do programa DAVID v.6.8. Os genes são os mesmos alvos de metilação e sequenciados em *D. gouveai*. As informações funcionais dos genes estão apresentadas por categoria, termo/gene, similaridade (%), *Fase Discovery Rate* (FDR), e *P-value*.....97

LISTA DE APÊNDICES

- Apêndice A.** Verificação da integridade de moléculas de DNA genômico (gel agarose 0.8%) (A), e digestões enzimáticas (B) (gel poliacrilamida 10%), amostra de *D. gouveai*.....129
- Apêndice B.** Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AAC)/HM(TAG), com DNA genômico de pupas em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.....130
- Apêndice C.** Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AGG)/HM(TGA), com DNA genômico de adultas e fêmeas que se desenvolveram em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.....131
- Apêndice D.** Gel de poliacrilamida 10% com verificação de produto de MSAP, E(ACT)/HM(TTG), com DNA genômico de adultos e machos que se desenvolveram em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle negativo (C-) da PCR.....132
- Apêndice E.** Verificação de produtos de PCR (gel de agarose 1%), com re-amplificações amostras diferenciais de metilação de DNA, em tecidos reprodutivos de moscas em *D. gouveai*.....133
- Apêndice F.** Matriz de dados binários MSAP com DNA genômico de pupas, em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P).....134
- Apêndice G.** Matriz de dados binários de MSAP com DNA genômico de moscas adultas e fêmeas, em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P).....136
- Apêndice H.** Matriz de dados binários de MSAP com DNA genômico de moscas adultas e machos, em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P).....138
- Apêndice I.** Predições de bioinformática para regiões promotoras em sequências de DNA.....140
- Apêndice J.** Estrutura quaternária da proteína DNA metiltransferase 2 (Dnmt2) em *D. buzzatii*. Esta sequência foi obtido do seu genoma sequenciado e disponível em <https://dbuz.uab.cat/welcome.php>, Dnmt2 - scaffold27:847239-848215, e a proteína foi modelada usando o I-TASSER Server disponível em <https://zhanglab.ccmb.med.umich.edu/I-TASSER>. A estrutura de Dnmt2 é composta por α -hélices (cor merge) e β -folhas (cor amarelo), e espaçadores (cores em branco e azul).....139

LISTA DE ANEXOS

Anexo A – Comitê de Ética.....	142
Anexo B – Manuscrito como primeiro autor, com os dados do projeto do doutorado, a ser submetido em revista científica.....	143
Anexo C – Outra produção Científica. Manuscrito como primeiro autor com colaboração acadêmica, mas externo ao projeto de doutorado, embora realizado durante o curso.....	172

RESUMO

Santos, A.S. **As contribuições da metilação diferencial no DNA para os processos sexo-específico e mudanças de ambientes hospedeiros em *Drosophila gouveai* (cluster *Drosophila buzzatii*)**. 2022. Tese (Doutorado em Ciências Biológicas, na área de Genética) - Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, 2022. 203p.

Plasticidade fenotípica é um processo em que um genótipo pode resultar em múltiplos fenótipos quando organismos alteram de ambiente ecológico. Um mecanismo molecular interessante constitui as mudanças epigenéticas, como metilação do DNA, relacionadas às respostas fenotípicas rápidas quando os organismos alteram o uso de ambientes. *Drosophila gouveai* (cluster *Drosophila buzzatii*) se desenvolve, obrigatoriamente, em tecidos necróticos de cactos, principalmente em *P. machrisis*, na natureza. Populações de *D. gouveai* têm apresentado isolamento geográfico, com baixa na diversidade genética por eventos de endogamia, mas dados sugerem plasticidade fenotípica na viabilidade, tempo de desenvolvimento, e estruturas morfológicas quando os organismos se desenvolvem no cacto *C. hildmannianus* em relação aos tecidos *P. machrisii*. Nesse contexto, o presente estudo teve como objetivo analisar a variação nos estados de metilação do DNA, posteriormente, o sequenciamento das regiões genômicas, e a identificação das vias moleculares por meio de análises de bioinformática, com dados obtidos do DNA genômico obtidos de organismos que se desenvolvem em diferentes ambientes, como nos tecidos dos cactos *P. machrisis*, em *C. hildmannianus*, e ao meio, padrão de laboratório (controle), na espécie *D. gouveai*. As amostras biológicas originaram-se de cultivo de duas gerações consecutivas em cada tratamento, com coleta de pupas, e moscas adultas virgens (fêmeas, e outros machos). O DNA genômico foi digerido em duas reações independentes, uma com o uso da enzima *HpaII* (sensível à metilação), outra com o isoesquizomero *MspI* (não sensível à metilação), e ambas as reações com a enzima *EcoRI*, por meio da metodologia *Methylation Sensitive Amplification Polymorphism* (MSAP). O padrão molecular das restrições *EcoRI-HpaII/MspI* apresentou diferenças no conjunto total de bandas das amostras de pupas (324 regiões), tecidos reprodutivos de fêmeas (308 regiões), e tecidos reprodutivos masculino (201 regiões). Descobriram-se diferenças em *loci* sensíveis à metilação de DNA (MSL) em moscas fêmeas que foram criadas em meio padrão (MSL=82), em tecidos de cactos *C. hildmannianus* (MSL=65), e *P. machrisii* (MSL=52). Evidenciaram-se diferenças em *loci* metilados no DNA em moscas machos que se desenvolveram e emergiram no meio padrão (MSL=55), outras em tecidos de *C. hildmannianus* (MSL=54), e em *P. machrisii* (MSL=42). Com base no sequenciamento de DNA, identificaram-se sequências com localização em regiões do genoma, que abrangem íntron, éxons, RNA não codificador, regiões promotoras, codificadores gênicos, e elemento transponível. Essas regiões relacionam-se com processos biológicos, como apoptose celular, síntese de proteínas e comunicação núcleo-citoplasma. A metilação de DNA pode regular a atividade de transposição de elemento transponível, e contribuir para a estabilidade do genoma. Os dados corroboram a hipótese de que as mudanças no estado de metilação de citosinas no DNA estão relacionadas com vias moleculares de sinalização, o que contribui com os indivíduos que se desenvolveram em diferentes ambientes hospedeiros, como nos cactos *P. machrisis*, e em *C. hildmannianus*, na espécie cactofílica *D. gouveai*.

Palavras-Chaves: *Drosophila* cactofílica. Metilação de DNA. Sexo-específico. Sítios CpGs. Elemento transponível. Ambiente ecológico.

ABSTRACT

Santos, A.S. **Contribution of differential DNA methylation to sex-specific processes and changes in host environments in *Drosophila gouveai* (*Drosophila buzzatii* cluster).** 2022. Thesis (Doctorate in Biological Sciences, in the area of Genetics) - University of São Paulo, Ribeirão Preto Medical School, 2022. 203p.

Phenotypic plasticity is a process in which a genotype can result in multiple phenotypes when organisms change their ecological environment. An interesting molecular mechanism constitutes epigenetic changes, such as DNA methylation, related to rapid phenotypic responses when organisms change their use environments. *Drosophila gouveai* (cluster *Drosophila buzzatii*) develops, obligatorily, in necrotic tissues of cacti, mainly in *P. machrisis*, in nature. Populations of *D. gouveai* have shown geographic isolation, with low genetic diversity due to inbreeding events, but data suggest phenotypic plasticity in viability, developmental time, and morphological structures when organisms develop in *C. hildmannianus* cactus in relation to *P. machrisii* tissues. In this context, this study analyzed the variation in DNA methylation states, later, the sequencing of genomic regions, and the identification of molecular pathways through bioinformatics analysis, with data obtained from genomic DNA obtained from organisms that develop in different environments, such as in the tissues of *P. machrisis* cacti, in *C. hildmannianus*, and in the middle, laboratory standard (control), in the species *D. gouveai*. The biological samples originate from the cultivation of two consecutive generations in each treatment, with a collection of pupae and virgin adult flies (females and other males). The genomic DNA was digested in two independent reactions, one with the enzyme *HpaII* (sensitive to methylation), another with the isoschizomer *MspI* (not sensitive to methylation), and both reactions with the enzyme *EcoRI*, using the methodology Methylation-Sensitive Amplification Polymorphism (MSAP). The molecular pattern of *EcoRI-HpaII/MspI* constraints showed differences in the total set of bands from pupae samples (324 regions), female reproductive tissues (308 regions), and male reproductive tissues (201 regions). Differences were found at DNA methylation-sensitive loci (MSL) in female flies that were reared on standard medium (MSL=82), in tissues from *C. hildmannianus* (MSL=65), and *P. machrisii* cacti (MSL=52). Differences in methylated DNA *loci* were evidenced in male flies that developed and emerged on standard medium (MSL=55), others in *C. hildmannianus* tissues (MSL=54), and in *P. machrisii* (MSL=42). Based on DNA sequencing, sequences were identified with localization in regions of the genome, including intron, exons, non-coding RNA, promoter regions, gene coding, and transposable elements. These regions are related to biological processes, such as cellular apoptosis, protein synthesis and nucleus-cytoplasm communication. DNA methylation can regulate transposable element transposition activity and contribute to genome stability. The data support the hypothesis that changes in the methylation state of cytosines in DNA are related to molecular signaling pathways, which contributes to individuals who developed in different host environments, such as in *P. machrisis* cacti, and in *C. hildmannianus*, in the cactophilic species *D. gouveai*.

Keywords: Cactophilic *Drosophila*. DNA methylation. CpG sites. sex-specific. Transposable element. Ecological environment.

Introdução

1.Introdução

1.1 Considerações iniciais

1.2 Cluster *Drosophila buzzatii*

O “cluster” *Drosophila buzzatii* (grupo *Drosophila repleta*) é constituído por sete espécies proximamente relacionadas: *Drosophila buzzatii* (Patterson & Wheeler, 1942), *Drosophila koepferae* (Fontdevila & Wasserman, 1988), *Drosophila antonietae* (Tidon-Sklorz & Sene, 2001), *Drosophila serido* (Vilela & Sene, 1977), *Drosophila gouveai* (Tidon-Sklorz & Sene, 2001), *Drosophila seriema* (Tidon-Sklorz & Sene, 1995), e *Drosophila borborema* (Vilela & Sene, 1977)(Manfrin & Sene, 2006). Essas espécies de *Drosophila* apresentam o desenvolvimento larval com ocorrência obrigatoriamente em tecidos necróticos de diferentes cactos hospedeiros (Pereira *et al.* 1983; Manfrin & Sene, 2006), com identificações taxonômicas por meio da morfologia do aparelho reprodutor masculino (*edeago*)(Vilela & Sene, 1977; Tidon-Sklorz & Sene, 2001)(Figura 1).

As relações filogenéticas entre as espécies do cluster *D. buzzatii* foram inferidas com o marcador de inversões cromossômicas fixas (Ruiz & Wasserman, 1993), e marcadores moleculares, tais como: gene mitocondrial *COI* (Manfrin *et al.* 2001), genes nucleares α -esterase 5 (Santos *et al.* 2009) e *period* (Franco *et al.* 2010). Os resultados obtidos nesses trabalhos indicam, com o conjunto de marcadores, que o cluster *D. buzzatii* é um grupo monofilético (Manfrin & Sene, 2006). Uma das hipóteses da relação entre as espécies mais aceita é a obtida com o marcador mitocondrial *COI*, que indica a organização monofilética do grupo, com um clado mais basal para *D. buzzatii*, seguido por outro clado para *D. koepferae*, e um agrupamento com as espécies *D. antonitae*, *D. gouveai*, *D. seriema*, *D. borborema*, e *D. serido* (Figura 2)(Manfrin *et al.* 2001; Manfrin & Sene, 2006).

Informações com substituições no gene mt*COI* sugerem que o processo de divergência entre as espécies do cluster *D. buzzatii* ocorreu em diferentes tempos, com, aproximadamente, 6 a 12 milhões de anos na divergência da linhagem evolutiva de *D. buzzatii* das demais espécies desse cluster (Manfrin *et al.* 2001; Manfrin *et al.* 2006). Os autores indicaram, também, um processo mais recente de divergência nas demais espécies, com estimativa de 3-6 milhões de anos no processo de diversificação entre o clado de *D. antonietae* e o agrupamento com *D. serido*, *D. seriema*, *D. borborema*, e *D. gouveai* (Manfrin *et al.* 2001; Manfrin *et al.* 2006).

Outro aspecto importante refere-se à distribuição geográfica das espécies do cluster *D. buzzatii* que ocorrem em regiões da América do sul, em áreas abertas e sazonalmente secas (Figura 3), sendo considerada somente a espécie *D. buzzatii* como cosmopolita com ocorrência em diferentes locais do mundo, como Europa, África, e Oceania (Manfrin & Sene, 2006; Barrios-Leal *et al.* 2019). Uma hipótese sobre a diversificação das espécies do cluster *D. buzzatii* relaciona-se à associação ecológica dos organismos com o uso de diferentes sítios hospedeiros para o

desenvolvimento biológico, com implicação na colonização de novas áreas e na sua distribuição geográfica (Manfrin *et al.* 2001; Manfrin *et al.* 2006). De fato, o trabalho de O'Grady & Markow (2012) sugere que o processo de diversificação das espécies de *Drosophila* cactofílicas, no continente Americano, pode ter ocorrido por radiação adaptativa por meio da exploração de tecidos necróticos de diferentes espécies de Cactaceae.

A exploração de sítios necróticos de cactos tem sido relatada para as espécies do cluster *D. buzzatii* em diferentes cladódios, tais como: *Drosophila buzzatii* com organismos que se desenvolvem e emergem moscas adultas em tecidos de *Trichocereus terscheckii*, *Cereus* sp., *Pilosocereus* sp., *Opuntia sulphurea*, *O. quimilo*, *O. sulfúrea*, *O. vulagris*, e *O. monocanta* (Pereira *et al.* 1983; Soto *et al.* 2014); *Drosophila koepferae* com indivíduos que usam os substratos dos tecidos vegetais de *Trichocereus terscheckii*, *Neocardensia* sp., *Cereus validus*, e *Opuntia* sp. (Hasson *et al.* 1992); *Drosophila serido* usa como sítio criadores para o desenvolvimento biológico em cladódios de *Cereus hildmaniannus*, *C. fermambucensis*, *Pilosocereus arrabidae*, e *O. ficus-indica* (Pereira *et al.* 1983); *Drosophila antonietae* com ocorrência de indivíduos com uso de tecidos em decomposição dos cactos *C. hildmaniannus* e *O. monocanta* (Soto *et al.* 2007; Barrios-Leal *et al.* 2019); *Drosophila seriema* com indivíduos que usam os recursos da matéria vegetal em decomposição em cactos do gênero *Cereus* (Tidon-Sklorz & Sene, 1995). *Drosophila borborema* apresentam indivíduos que se desenvolvem em *Pilosocereus piauhyensis* e *Cephalocereus piauhyensis* (Vilela & Sene, 1977; Pereira *et al.* 1983). *Drosophila gouveai* possui organismos que se desenvolvem e emergem moscas adultas em tecidos necróticos de cactos *Pilosocereus machrisii*, *P. vilaboensis*, e *Opuntia ficus-indica* (Tidon & Sene, 2001; Manfrin & Sene, 2006; Moraes *et al.* 2009).

Os tecidos vegetais de cactos contêm diferenças em compostos químicos, tais como: açúcar, água, ácidos graxos, esteroides, triterpenos, e alcaloides (Kircher *et al.* 1967; Fogleman & Danielson, 2001; Stintzing & Carle, 2005). Compostos alcaloides são indicados com potencial citotóxico em processos metabólicos e no ciclo celular (viabilidade celular) dos organismos (Kircher *et al.* 1967; Dutra *et al.* 2018). Com base em marcadores de múltiplos estados e biológicos em espécies de *Drosophila* cactofílica, tem sido indicada uma hipótese que o cacto palmar *Opuntia* sp. apresenta menor complexidade química do que as espécies de cactos colunares, o que implica diferenças na história de vida dos organismos (Oliveira *et al.* 2012).

A heterogeneidade em tecidos de cactos tem sido relatada com a grande diversidade de compostos alcaloides, em perfis químicos de tecidos vegetais por espécie de cacto (Fogleman & Danielson, 2001; Panis *et al.* 2016). Por exemplo, *Opuntia sulphurea* contém os alcaloides mescalina e 4-hydroxyproline; *Trichocereus terscheckii* apresenta, em seus tecidos, os compostos alcaloides de trichocereine e mescaline, N,N-dimethylmescaline (Panis *et al.* 2016); *Lophocereus*

schotti contém, em seus tecidos, os alcaloides pilocereine e lophocereine (Kircher *et al.* 1967); *Pilosocereus maxonii* apresenta, em sua composição química, alcaloides N-metil-3,4-dimetoxifenetilamina, 3,4-dimetoxifenetilamina, tiramina e N-metiltiramina (Pummangura *et al.* 1977); *Echinocereus merkeri*, os compostos químicos de N,N-dimetil-3,4-dimetoxifenetilamina, N-metil-3,4-dimetoxifenetilamina, hordenina, e salsoline (Aguirell *et al.* 1969); *Cereus jamacaru* contém os compostos alcaloides de tiramina e N-metiltiramina (Davet, 2005).

Grupos químicos complexos em tecidos de plantas são indicados como uma barreira na troca de hospedeiros por diferentes espécies de insetos fitófagos (Becerra, 1997). O uso de diferentes tecidos vegetais demanda um ajuste no sistema fisiológico e químico no organismo de inseto durante o processo adaptativo aos sítios hospedeiros (Panis *et al.* 2016). De fato, indicam-se os compostos alcaloides com possível potencial tóxico para desenvolvimento e viabilidade dos indivíduos de espécies de *Drosophila* cactofílicas (Kircher *et al.* 1967; Soto *et al.* 2014).

Outro aspecto que chama atenção refere-se ao mecanismo biológico que as espécies de *Drosophila* utilizam para reconhecer os sítios criadores em cactos necróticos, que envolve a localização dos cladódios, aceitação/reconhecimento, e uso dos seus substratos químicos (Markow, 2019). Esse mecanismo pode envolver diferentes sentidos na percepção do meio interno e externo, em *Drosophila*, tais como: visão e olfato (localização dos hospedeiros); visão, olfato, tato, e mecanorreceptores (aceitação e reconhecimento); e sistemas fisiológicos no metabolismo e desintoxicação de compostos (uso dos substratos químicos e desenvolvimento larval) (Markow, 2019). Em conjunto, esses fatores influenciam a percepção e adaptação dos organismos ao uso de diferentes recursos ecológicos, com efeito no processo de reprodução e viabilidade dos indivíduos (Hasson *et al.* 2019; Markow, 2019).

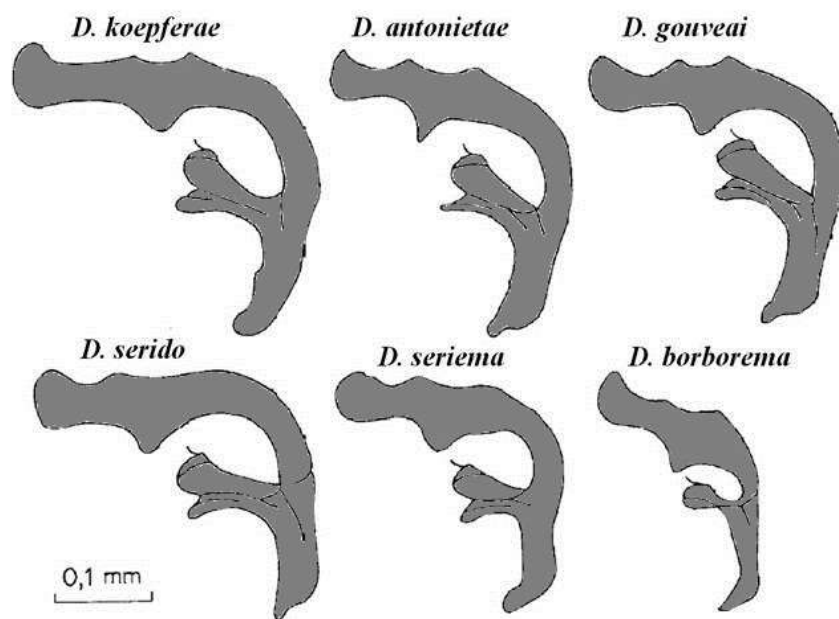


Figura 1. Relações da morfologia do aparelho reprodutor masculino (edeago) nas espécies do cluster *D. buzzatii*, *D. koepferae* (isolinagem B25D7, La Rioja, Argentina), *D. antonietae* (Serrana, São Paulo), *D. gouveai* (Analândia, São Paulo), *D. serido* (B53Q5, Milagres, Bahia), *D. seriema* (B95C3, Serra do Cipó, Minas Gerais), e *D. borborema* (D54B5, Grão Mogol, Minas Gerais). Adaptado de Tidon-Sklorz & Sene (2001).

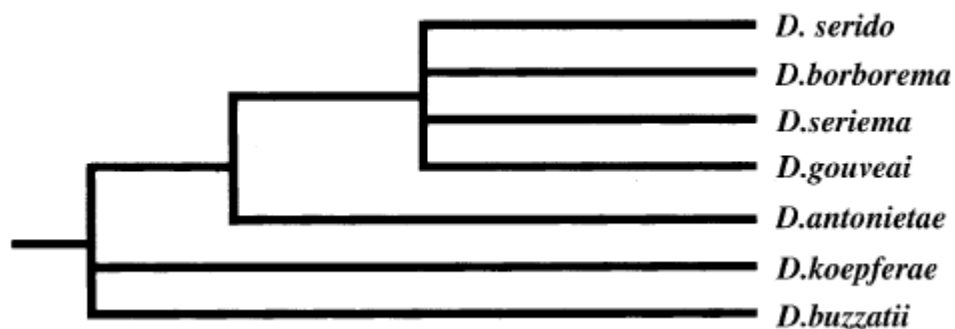


Figura 2. Hipótese de relações filogenética com o marcador mitocondrial mtCOI no cluster *D. buzzatii*. Adaptado de Manfrin & Sene (2006).

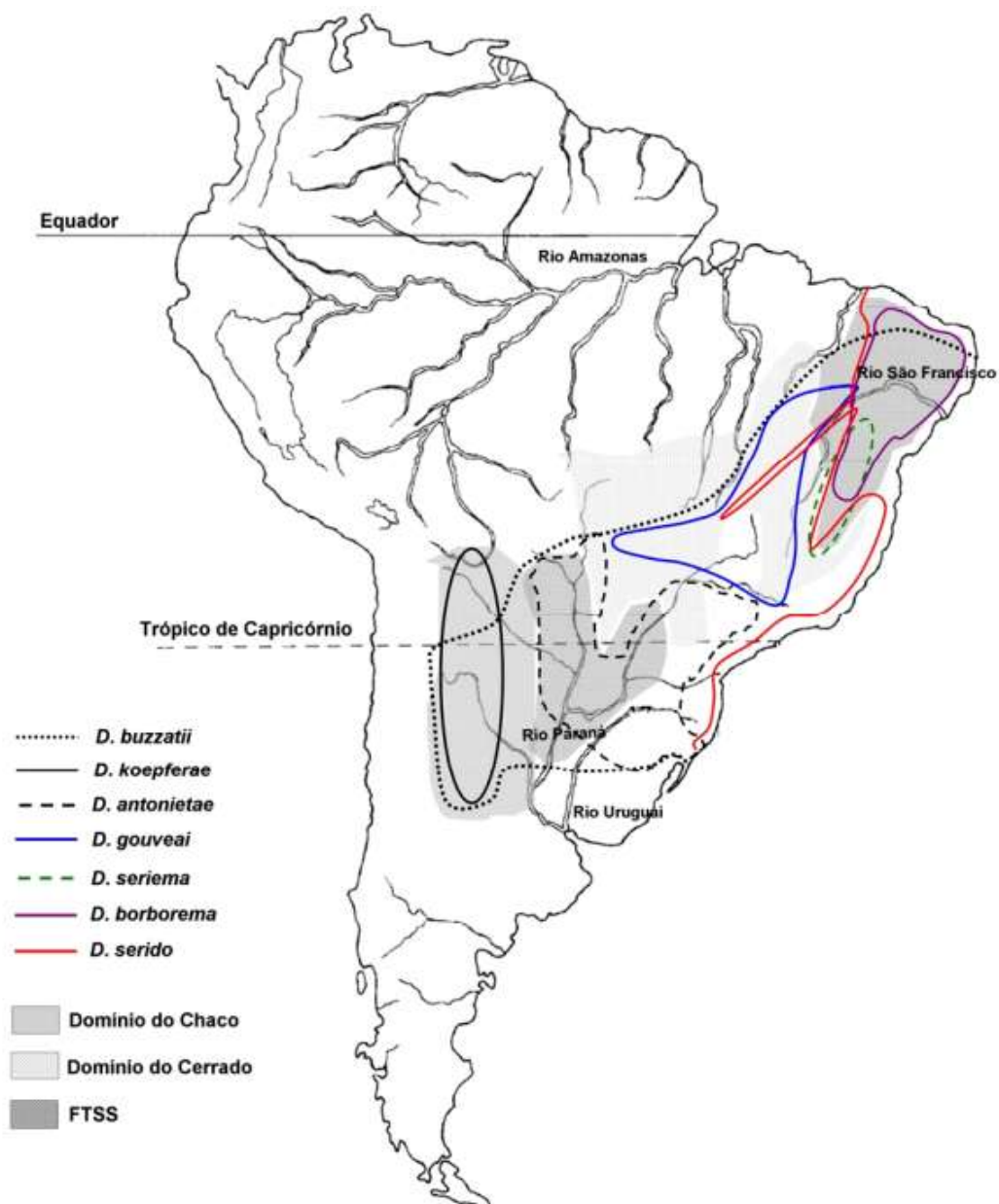


Figura 3. Distribuição geográfica conhecida das espécies do cluster *Drosophila buzzatii* na América do Sul. As principais bacias hidrográficas, bem como a distribuição aproximada das principais áreas de florestas secas da América do Sul, estão apresentadas no mapa. FTSS: Florestas Tropicais Sazonalmente Secas. Fonte: Franco, 2009.

1.3 Biologia Reprodutiva e Desenvolvimento em *Drosophila*

O mecanismo usado no processo reprodutivo entre as espécies de *Drosophila* apresenta mudanças em características morfológicas, fisiológicas, e comportamentais (Markow, 1996; Markow & O'Grady, 2008), assim como no uso de diferentes recursos ecológicos. Estudos encontraram diferentes graus de isolamento reprodutivo pré-copulatório, produção de gametas, e pós-zigótico em espécies do cluster *D. buzzatii*, e grupos de *Drosophila* (Markow, 1996; Maddi-Ravazzi *et al.* 1997; Oliveira *et al.* 2013).

O isolamento pré-copulatório envolve uma característica importante que é a corte sonoro (Oliveira *et al.* 2013). Esse processo ocorre quando os indivíduos machos e adultos realizam o batimento coordenado de suas asas durante o cortejo das moscas fêmeas (Ewing, 1983). O cortejo e batimento das asas podem ser caracterizados por parâmetros acústicos, tais como: duração de rajada (BD), frequência de portadora (CF), comprimento de pulso (PL), número de pulso (PN), intervalo entre rajadas (IBI), e intervalo entre pulsos (IPI) (Oliveira *et al.* 2013). Adicionalmente, o padrão acústico apresentam diferenças na corte sonora entre as espécies cactofílicas do cluster *D. buzzatii* (Machado *et al.* 2002; Oliveira *et al.* 2013) (Figura 4).

Dentre as características sonoras, o Intervalo Entre Pulsos (IPI) apresenta variação rítmica de forma espécie-específica no cluster *D. buzzatii* (Oliveira *et al.* 2013). Outro componente importante refere-se à duração média do corte sonoro que apresenta diferença interespecífica nas espécies deste cluster (Machado *et al.* 2002).

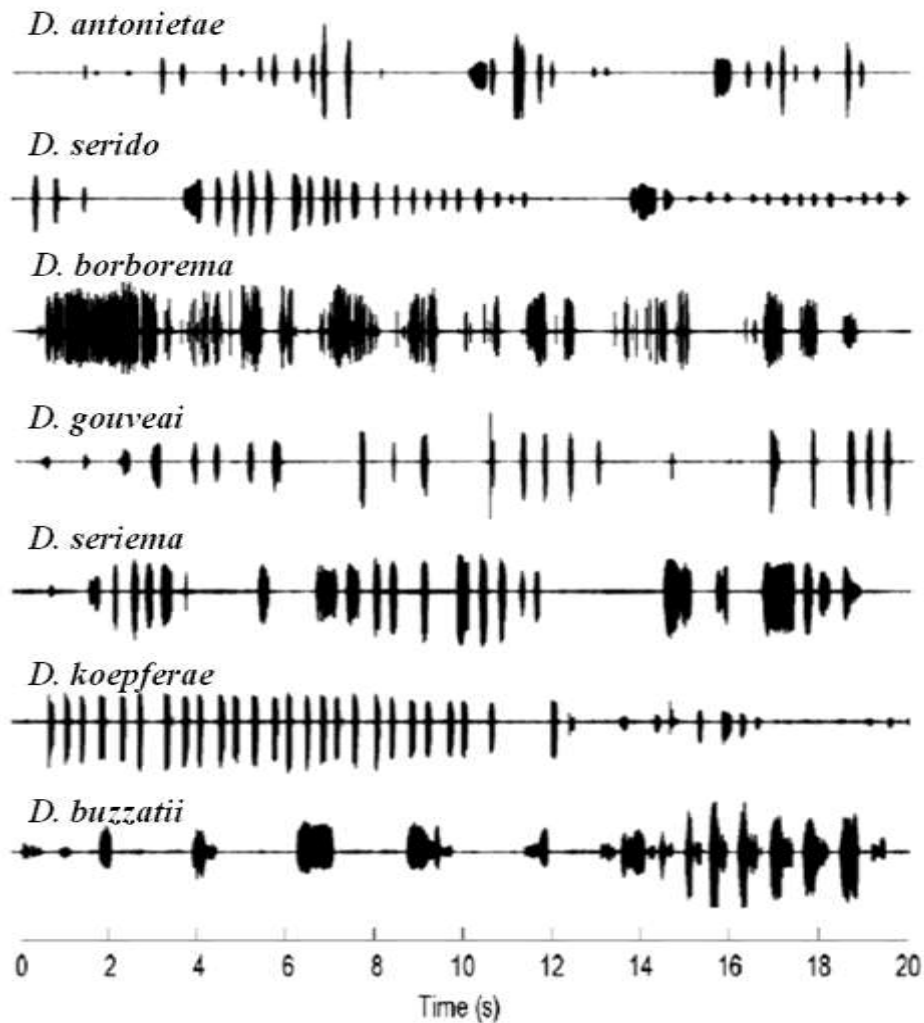


Figura 4. Padrão acústico do cortejo sonoro produzido por moscas machos (adultos) nas espécies do cluster *D. buzzatii*. Adaptada de Oliveira *et al.* (2013).

O padrão sonoro é percebido por meio das antenas, nas aristas (estruturas emplumadas das antenas), com localização no segundo segmento da antena que aumenta as vibrações e são sentidos pelo órgão de Johnston (Tyler, 2000).

Uma característica importante refere-se à capacidade de entregar e receber os sinais de corte apropriados, o que depende da espécie de *Drosophila* (Markow & O'Grady, 2005). Após a percepção dos sinais de corte, a mosca fêmea apresenta o comportamento de namoro que pode incluir sinais de rejeição ou aceitação do macho, com o uso de componentes olfativos, auditivas ou visuais durante a cópula (Markow & O'Grady, 2005). Em caso de rejeição, a mosca fêmea pode apresentar o comportamento de se movimentar com as patas traseiras, com afastamento por meio das patas do meio, que inclui movimentos das asas para a produção do som de rejeição, e o afastamento, rapidamente, do macho (Tyler, 2000). Em contraste, no comportamento de aceitação do macho, a mosca fêmea aceita a aproximação do macho, que realiza o comportamento de tocar as patas na cabeça da fêmea, e produz o som de cortejo com as asas (Tyler, 2000). Após aceitação do

macho, a mosca estende a probólide para tocar a genitália da fêmea, com posterior acasalamento, e transferência de espermatozoides por meio do *edeago* para o receptor seminal e espermoteca na fêmea (Tyler, 2000).

Há variações no sistema de reprodução que podem ser observadas nas gônadas reprodutivas em moscas fêmeas, com mudanças no tamanho de ovos (0.44 mm para 1.09 mm) e no número de ovariolos (n=9 a 101) entre as diferentes espécies de *Drosophila* (Markow, 1996). Essa produção de gametas e a estratégia reprodutiva podem variar entre as espécies de *Drosophila* com relação a vários fatores, tais como: tamanho da gônada, idade de reprodução, longevidade, e nutrição (Markow, 1996).

Em contraste aos dados de origem em fêmeas, tem sido chamada atenção pelas diferenças no sistema reprodutivo que envolve o material reprodutivo masculino (Markow, 1996). Esses compostos no material masculino, em relação ao processo de ejaculação, compreendem espermatozoides e o líquido seminal, que é produzido nas glândulas acessórias nos indivíduos machos (Knowles & Markow, 2001). De fato, o material ejaculado contém diferentes substâncias, que envolvem carboidratos, lipídios, e proteínas das glândulas acessórias (APCs) (Markow, 1996).

Adicionalmente, dados têm documentado variações em características no conteúdo e composição espermática, tais como: tamanho, forma, e número dos espermatozoides entre as espécies de *Drosophila* (Markow, 1996). Esse fenótipo de tamanho dos espermatozoides apresenta alteração espécie-específica em *Drosophila*, tais como em *Drosophila simulans* (1.14 mm), *D. buzzatii* (1.59 a 167 mm), e *Drosophila bifurca* (58.29 mm) (Markow, 1996; Bundgaard & Baker, 2000). Uma interpretação dos dados morfológico espermáticos em *D. buzzatii* sugerem estruturas heteromórficas nessa espécie (Markow, 1996). Outro elemento importante refere-se à oscilação do número de espermatozoides entre as espécies de *Drosophila*, como *Drosophila pachea* (n=44) e *D. pseudoobscura* (25.000)(Markow, 1996).

Com base no conjunto de dados reprodutivos em *Drosophila*, descobriram-se diferentes estratégias, tais como: múltiplos acasalamentos em uma única manhã, em *Drosophila hydei* e *Drosophila nigrospiracula*; único cruzamento em toda a sua história de vida, em *Drosophila subobscura* (Markow, 2015).

Logo após o processo de fecundação, com a transferência de espermatozoides para o aparelho reprodutor feminino, ocorre a formação da massa de reação no órgão reprodutor feminino com duração de 1-16 horas, nas espécies de *Drosophila* (Knowles & Markow, 2001). Esse tempo de duração da massa de reação pode apresentar diferenças interespecíficas em *Drosophila*, e inclui o processo de ovoposição (Knowles & Markow, 2001).

Em contraste, o sistema produtivo pode incluir a Competição de Espermatozoides (SC) de machos por fêmeas (Markow, 1996). Evidências descobriram que as moscas fêmeas podem realizar

o processo de reacasalamento, com substituição de esperma, e múltipla paternidade da prole, nas espécies cactofílicas *D. buzzatii*, *D. koepferae* e *D. antonietae* (Bundgaard & Baker, 2000; Hurtado *et al.* 2013; Goya *et al.* 2019). Essa competição de esperma e o alto custo reprodutivo são sugeridos para o organismo de machos, no acasalamento com as fêmeas, em *D. buzzatii* (Hurtado *et al.* 2013). Desse modo, os diferentes mecanismos reprodutivos contribuem com a dinâmica de suas populações na natureza, que incluem os gastos energéticos e metabólicos para a formação da próxima geração.

Adicionalmente, dados morfológicos e celulares permitiram caracterizar diferentes fases no desenvolvimento biológico *D. melanogaster*, que incluem as seguintes etapas, conforme Tyler (2000):

Fase de Ovo: a estrutura compreende envelope vitelino, córion, conjunto de células (com conteúdo de DNA, RNA, e proteínas), micropilo, e o filamento respiratório. A morfologia da estrutura é oval, e com cor esbranquiçada.

Fase de larva de 1º instar: organismo com aproximadamente 22 a 24 horas, após a fecundação, e com a eclosão do ovo. A morfologia da estrutura da larva apresenta aspecto filamentosos, pequena, e esbranquiçada.

Fase de larva de 2º instar: o organismo da larva de 1º instar, com aproximadamente 25 horas, que se alimentou do substrato, e segue para a fase de 2º instar. Tem-se o aumento da estrutura corporal da larva de 2º instar em relação à de 1º instar.

Fase de larva de 3º instar: o organismo da larva de 2º instar, com, aproximadamente, 24 horas, que se alimentou do substrato, e ocorre a diferenciação em larva de 3º instar. Tem-se o acréscimo em tecido adiposo, estrutura corporal da larva de 3º instar em relação à de 2º instar, e com cor de aspecto amarelado.

Fase de Pupa: o organismo da larva de 3º instar, com aproximadamente 30 horas, e com comportamento de movimento da larva do substrato líquido em direção à parte física e seca, e formação de pupa. As estruturas da larva passam por metamorfose (*imago*), com diferenciação celular, reestruturação de tecidos, mas com a preservação das estruturas do sistema nervoso, túbulo de Malpighi, e das gônadas. O aspecto da pupa torna-se, progressivamente, de cor amarelo escuro, em diferentes tons, conforme o organismo se desenvolve.

Fase de adulto: o organismo contém um conjunto de células mitóticas indiferenciadas (discos imaginais e histoblastos) dentro da larva que se diferencia na maioria das estruturas da mosca adulta. As células do disco imaginal são diferenciadas em estruturas de asas, pernas, olhos, partes da boca, e ducto genital. As células do histoblastos são diferenciadas nos tecidos do abdômen e órgãos internos da mosca adulta. O processo de ovo para mosca adulta pode ocorrer com, aproximadamente, 216 horas. Por fim, a maturação sexual da mosca macho e mosca fêmea podem apresentar diferenças em tempo (horas) espécie-específica.

Estudos evidenciam isolamento reprodutivo pós-zigótico em experimentos com cruzamentos de diferentes linhagens evolutivas no cluster *D. buzzatii* (Madi-Ravazzi *et al.* 1997; Machado *et al.*

2006). Esses autores indicam que há diferentes níveis de isolamento reprodutivo nas espécies cactofílicas do cluster *D. buzzatii*, mas sem o relato de caso de isolamento reprodutivo completo (Madi-Ravazzi *et al.* 1997; Machado *et al.* 2006). Os resultados indicam um isolamento pré-zigótico assimétrico no cruzamento entre *D. gouveai* (fêmeas, *stock* B50) e *D. buzzatii* (machos, *stock* R2) (Machado *et al.* 2006).

Os fenótipos de baixa taxa de fecundidade foram descobertos nos cruzamentos intraespecíficos em linhagens de *D. seriema*, *D. antonietae*, *D. gouveai* (Machado *et al.* 2006). Esse aspecto reprodutivo foi encontrado, também, nos cruzamentos interespecíficos entre *D. gouveai* e *D. buzzatii* (Machado *et al.* 2006). Adicionalmente, dados citogenéticos indicam que erros na sinapse entre cromossomo politênico X e microcromossomo podem relacionar-se ao fenótipo de infertilidade nas espécies do cluster (Madi-Ravazzi *et al.* 1997; Machado *et al.* 2006). Isso indica que alterações nas interações entre regiões cromossômicas podem ter importância no processo reprodutivo, e na infertilidade em espécies do cluster *D. buzzatii*.

Outro aspecto interessante refere-se ao fato de que em relação às moscas fêmeas, após a realização do segundo acasalamento, a fertilidade do primeiro acasalamento pode ser reduzida em 71.4% para *D. buzzatii* e 33.3% para *D. koepferae* (Hurtado *et al.* 2013; Goya *et al.* 2019). Isto sugere que ocorrem diferenças na estratégia reprodutiva das espécies cactofílicas dentro do cluster *D. buzzatii*.

Interessantemente, os componentes da nutrição (carboidratos e glicose) das moscas parentais (P) podem alterar o fenótipo da sua prole, com o efeito transgeracional (F2), em *D. melanogaster* (Emborski & Mikheyev, 2019). Todavia, ainda não estão claros os mecanismos moleculares para a regulação entre genótipo-fenótipo, quando os organismos se desenvolvem com o uso de diferentes recursos nutricionais e ecológicos, nas espécies de *Drosophila*.

1.4 Plasticidade de fenótipos em resposta à mudança de ambiente

O conceito de fenótipo compreende características que são observáveis ou mensuráveis de um organismo, tais como: processos fisiológicos, desenvolvimento, morfologia, e comportamentos (Whitlock, 2014). Mudanças em fenótipos pode ser resultado de características herdadas por meio de variação genética na prole com origem parental (Whitlock, 2014). A diversidade de fenótipo é um elemento que pode contribuir para o processo de adaptação das populações às mudanças de fatores ambientais, que inclui fatores tais como: temperatura, nutrição, e estresse (Whitlock, 2014). De fato, dados na literatura discutem a diversificação de fenótipos com diversos eventos, como a plasticidade de fenotípica.

Plasticidade fenotípica é conhecida como a capacidade de um único genótipo de expressar vários fenótipos em resposta à sua criação em diferentes ambientes (Whitman & Agrawal, 2009;

Fox *et al.* 2019). Essas mudanças nos fenótipos podem compreender diferentes níveis, como mudanças morfológicas, fisiológicas, e de organização comportamental/social (Whitman & Agrawal, 2009; Fox *et al.* 2019). O mecanismo de plasticidade fenotípica envolve mudanças em rede regulatórias de genes, hormônios, no desenvolvimento de tecidos e nos organismos que podem melhorar o ajuste dos organismos aos diferentes ambientes (Ketterson *et al.* 2014). Uma questão interessante é o custo e limite da plasticidade fenotípica dos organismos na adaptação das populações aos fatores externos do ambiente (Stearns, 2014).

Dados promissores têm sido obtidos com o modelo *Drosophila*-Yeast-Cactus, que compreende as espécies cactofílicas do cluster *D. buzzatii*, com evidências de plasticidade fenotípica em organismos que usaram tecidos de diferentes espécies de cactos hospedeiros (Soto *et al.* 2007; Hasson *et al.* 2019). Estes organismos se desenvolvem em diferentes condições ecológicas, que envolvem a composição química dos cactos, composição da comunidade de bactérias e fungos nos tecidos vegetais em decomposição na natureza (Kircher *et al.* 1967; Manfrin & Sene, 2006; Matzkin *et al.* 2014). Traços fenotípicos com importância no processo de adaptação de populações têm sido descobertos entre as espécies de *Drosophila* quando os organismos se desenvolveram nos substratos químicos, na matéria em decomposição, em diferentes espécies de cactos (Kircher *et al.* 1967; Matzkin *et al.* 2014).

Descobertas demonstram comportamento de ovoposição com preferência para os tecidos dos cactos *Opuntia sulphurea* (palmar) e *Trichocereus terscheckii* (colunar), em *Drosophila buzzatii* e *Drosophila koepferae*, respectivamente (Hasson *et al.* 2019). Isto foi sugerido com base na diferença nos números de ovos, no processo de ovoposição, encontrado em tecidos vegetais dos cactos *O. sulphurea* (66%) e *T. terscheckii* (7%), em *D. buzzatii* (Hasson *et al.* 2019). Adicionalmente, descobriu-se um acréscimo de 10% no número de ovariolos, em moscas fêmeas, em *D. buzzatii* quando se desenvolveram em cladódios de *O. sulphurea* (Hasson *et al.* 2019). Em sua espécie irmã *D. koepferae*, relatou-se maior amostragem de indivíduos em tecidos de *T. terscheckii* do que em *O. sulphurea* (Hasson *et al.* 2019). Uma explicação pode ser a complexidade química dos ambientes hospedeiros, com dados que sugerem diferenças na composição de alcaloides: tais como mescalina e tricocereína (N,N dimetilmescalina), em *T. terscheckii*; e derivados de prolina semelhantes a 4-hidroxi prolina, em *O. sulphurea* (Panis *et al.* 2016). De fato, os dados indicam que as moscas fêmeas podem apresentar um comportamento de preferência na escolha dos sítios hospedeiros no processo de ovoposição, com a seleção dos melhores recursos para o desenvolvimento da prole.

Tem sido reconhecido que a regulação materna pode ocorrer na composição de produto gênico, mRNA, e proteínas que são depositadas em seus ovos, no início do processo de desenvolvimento, que são influenciadas por mudanças no ambiente nutricional, em *D. melanogaster*

(Crofton *et al.* 2018). Genes relacionados ao processo traducional, transporte, e macromoléculas estão envolvidos com nutricional materna que representa um custo energético no processo de início do desenvolvimento do embrião (Crofton *et al.* 2018). Assim, moscas (maternas) apresentam plasticidade na deposição de mRNA em resposta ao ambiente nutricional, que contribui para o desenvolvimento do embrião no ambiente (Crofton *et al.* 2018).

Fenótipos de viabilidade dos organismos podem ser alterados com a mudança de cactos hospedeiros (Soto *et al.* 2007). Dados com cultivo de larvas, em meios com diferentes cactos hospedeiros, indicam que as espécies irmãs *D. buzzatii* e *D. koepferae* apresentam maior viabilidade de larvas em ambiente com tecidos de *Opuntia* do que em *Tichocereus* (Vrdoljak *et al.* 2019). De acordo com esses autores, em experimento com acréscimo na proporção de alcaloides em meio de cultura, as larvas apresentaram redução da viabilidade quando se desenvolvem neste ambiente de cultura, em *D. buzzatii* e *D. koepferae* (Vrdoljak *et al.* 2019). Neste estudo descobriu-se que dieta pobre em proteínas resulta na diminuição da viabilidade dos indivíduos somente em *D. buzzatii* (Vrdoljak *et al.* 2019). Esses resultados sugerem que os organismos apresentam diferentes graus de ajustes aos compostos em ambientes de diferentes cactos hospedeiros.

Plasticidade fenotípica tem sido relatada no desempenho e sobrevivência dos organismos aos cladódios de cactos (Soto *et al.* 2007). Os autores encontraram um bom desempenho dos organismos, nas espécies cactofílicas *Drosophila antonietae* e *Drosophila gouveai*, no desenvolvimento em tecidos de cactos *Pilosocereus machrisis* (Soto *et al.* 2007). Em contrastes, a sobrevivência diferiu, dramaticamente, entre *D. antonietae* e *D. gouveai* em tecidos de *Cereus hilmannianus* (Soto *et al.* 2007). Nesse contexto, reduziu-se a viabilidade de *D. antonietae* no cacto *P. machrisis*, e *D. gouveai* apresentou baixa sobrevivência em *C. hilmannianus* (Soto *et al.* 2007). Essas características podem ser importantes para o *fitness* e sucesso reprodutivos de organismos na interação com ambientes ecológicos.

Aliada a essas características fenotípicas, mudanças têm sido documentadas na velocidade de atividade metabólica das larvas para a formação de pupas em populações de espécies de *Drosophila* cactofílicas (Coleman *et al.* 2018). Dados indicam que há diferenças genéticas, distribuídas no genoma que pode implicar a atividade das larvas e formação de pupas, embora ainda não esteja clara a contribuição de *loci* no comportamento de pupas em espécies de *Drosophila* cactofílicas (Coleman *et al.* 2018). Esses autores sugerem que a estrutura dos sítios criadores, tais como a forma, tamanho e volume dos tecidos de cactos, são fatores que podem contribuir para a atividade das larvas nas espécies de *Drosophila* cactofílicas (Coleman *et al.* 2018).

O tempo de desenvolvimento de larvas de 1º instar a adulto, em espécies de *Drosophila* cactofílicas, pode ser alterado com o uso de da matéria em decomposição de tecidos de cactos (Hasson *et al.* 2019). Os resultados de experimentos indicam rápido tempo de desenvolvimento de

D. buzzatii em *O. sulphurea* e *D. koepferae* em *T. terscheckii* (Hasson *et al.* 2019). Em suas espécies irmãs, *D. antonietae* e *D. gouveai* apresentam desenvolvimento mais rápido em *P. machrisis* do que em *C. hildmanianus* (Soto *et al.* 2007). Isto sugere que há diferentes níveis de organização em mecanismos genéticos e fisiológicos para a exploração de diferentes recursos (ex. habilidade de obter nutrientes e eliminar toxinas), nessas espécies de *Drosophila* cactofílica (Hasson *et al.* 2019). Desse modo, o tempo de desenvolvimento pode contribuir para o processo para emergir moscas adultas, nova geração, e colaborar para a dispersão e busca de outros nichos ecológicos.

Evidências têm encontrado plasticidade fenotípica na morfologia da asa de moscas quando se desenvolvem em diferentes cactos hospedeiros (Manfrin *et al.* 2006; Soto *et al.* 2008; Hasson *et al.* 2019). Dentre as estruturas das asas, houve maior tamanho das asas em moscas adultas quando os organismos se desenvolveram em tecidos de *O. sulphurea* do que em *T. terscheckii*, nos dados com *D. buzzatii* e *D. koepferae* (Hasson *et al.* 2019). Nas espécies cactofílicas irmãs, moscas adultas apresentaram maior tamanho da asa quando os organismos se desenvolveram no cacto *C. hilmannianus* do que em *P. machrisis*, em *D. antonietae* e *D. gouveai* (Soto *et al.* 2008). Essas descobertas em relação às estruturas das asas sugerem uma implicação no custo de energia no processo de voo e na dispersão dos organismos na natureza.

Outro aspecto central refere-se ao tempo de resistência à fome após emergir moscas adultas, nas espécies de *Drosophila*, que podem ser influenciadas pelo cacto hospedeiro (Hasson *et al.* 2019). Esses autores apresentam dados com evidências que as moscas adultas podem apresentar maior tempo de resistência à fome quando se desenvolveram em tecidos de *O. sulphurea* do que em *T. terscheckii*, em *D. buzzatii* (100 horas; 60 horas) e *D. koepferae* (175 horas; 140hs) respectivamente (Hasson *et al.* 2019). Isto pode contribuir para a busca de novos sítios criadores em cactos hospedeiros, nas moscas cactofílicas.

O conjunto de evidências indicam diferenças em características com plasticidade fenotípica que compreendem o processo de ovoposição, reprodução, viabilidade, comportamento larval para pupação, tempo de desenvolvimento, sobrevivências, resistência à fome e estrutura das asas em indivíduos nas espécies de *Drosophila* com o uso de tecidos necróticos em diferentes cactos (Soto *et al.* 2007; Matzkin *et al.* 2014; Hasson *et al.* 2019). Entretanto, ainda não estão elucidados os mecanismos moleculares que regulam as informações de um genótipo para múltiplos fenótipos.

Uma hipótese proposta refere-se ao fato de que mudanças epigenéticas, tais como a metilação no epigenoma, podem contribuir com a plasticidade de fenótipos em insetos, com respostas células específicas ou tecido-específico (Maleska, 2016), e com a colonização de diferentes ambientes ecológicos (Schrey *et al.* 2016). Ressalta-se que os dados na literatura, ainda,

são escassos sobre as mudanças químicas no DNA quando os organismos se desenvolvem em diferentes ambientes ecológicos, e suas implicações para a plasticidade de fenótipos.

1.5 Mecanismos epigenéticos em *Drosophila*

Em 1956, o biólogo de desenvolvimento, Conrad Waddington, obteve resultados com a proposta de novos conceitos relacionados à capacidade de influências ambientais induzirem a formação de novos fenótipos nos organismos e a sua manutenção nas gerações consecutivas (Waddington, 1956). Obtiveram-se esses resultados com os embriões de *D. melanogaster* que resultam em diferentes fenótipos em características de tórax e asas, quando os organismos se desenvolveram com mudanças de estímulos, tais como: temperatura do ambiente, e compostos químicos (Waddington, 2012).

Uma hipótese elaborada por Waddington (1956) refere-se ao fato de que o processo de desenvolvimento biológico ocorre com a “canalização”, e, frente a distúrbios e influência externa estressante, pode ocorrer uma perturbação das tendências no desenvolvimento normal em condições naturais. Esse processo é seguido por etapas que são constituídas por um conjunto de “decisões” que pode ser representadas como “vales” e “bifurcações” em um ambiente de paisagem de desenvolvimento (Figura 5) (Waddington, 1956). Nesse modelo, o termo epigenótipo foi sugerido para explicar variações entre genótipo-fenótipo, em que o genótipo pode expressar diferentes fenótipos no desenvolvimento de *Drosophila* (Waddington, 2012). Adicionalmente, os dados sugerem que ocorre um conjunto de fatores no processo de desenvolvimento, tais como influências do ambiente, ambiente de desenvolvimento, network funcional, e redes gênicas que contribuem com a organização do sistema biológico (Waddington, 2012).

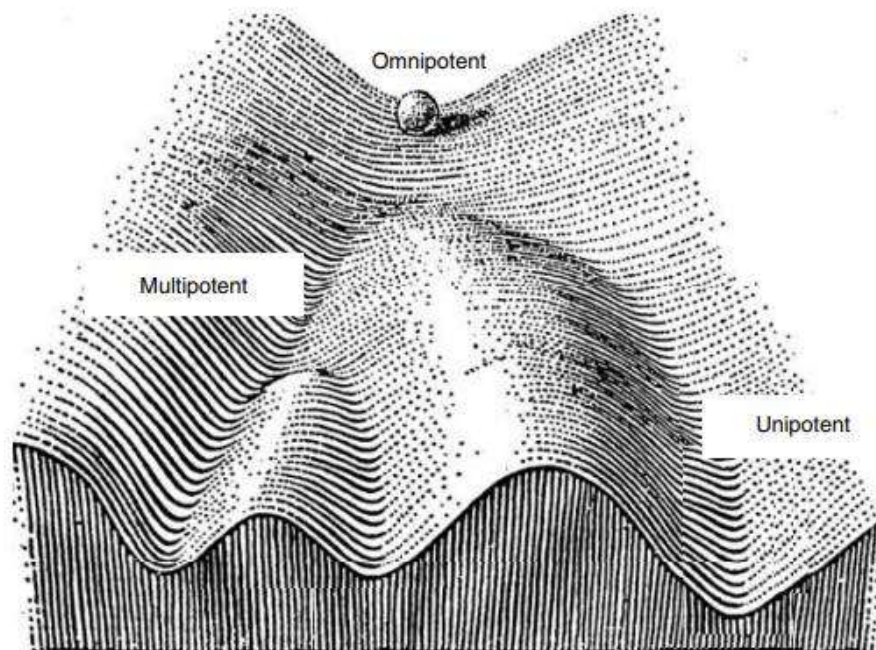


Figura 5. Ilustração do ambiente de desenvolvimento proposto por Waddington (1956). Esse modelo é conhecido como paisagem epigenética de Waddington. Célula “onipotente” do embrião em desenvolvimento que é representada como uma bola rolando por um “vale”. Essa bola pode encontrar vários pontos de ramificação. A escolha do ramo esquerdo ou direito depende, em certa medida, da forma local precisa da paisagem, e pode ter predisposições genéticas ou de influências de fatores ambientais. Modificado de Waddington (1956).

O termo epigenética é reconhecido por mudanças na regulação gênica hereditária mitoticamente ou meiótica sem alterações na sequência de DNA (Fitz-James & Cavalli, 2022). Nesse sentido, a herança epigenética envolve a transmissão de informações que não são codificadas nas sequências de DNA entre uma célula parental para a célula filha ou de uma geração para outra geração de organismos (Fitz-James & Cavalli, 2022).

De modo geral, diferentes tipos de processos epigenéticos têm sido identificados, documentados, e compreendem fenômenos, tais como: 1) metilação ocorre por meio da substituição de um átomo de hidrogênio (H) por um grupo metil (CH_3) na posição cinco do carbono na base de citosina no DNA genômico, que é catalisada por metiltransferases (Fitz-James & Cavalli, 2022); 2) fosforilação ocorre com a adição de um grupo fosfato (PO_4) em um aminoácido de uma cadeia proteica, que é um dos mecanismos na regulação de proteínas (Fitz-James & Cavalli, 2022). O processo de fosforilação é realizado pelas enzimas quinases, e a desfosforilação por meio das enzimas fosfatases (Fitz-James & Cavalli, 2022); 3) ubiquitinação ocorrem por meio de conjunto de proteínas (E1, E2, E3) que adicionam marcas químicas em moléculas de proteína em células, com o consumo de ATP neste processo (Glastad *et al.* 2019; Fitz-James & Cavalli, 2022). Moléculas marcadas por ubiquitina são reconhecidas por um proteossoma, em que ocorre a degradação (Glastad *et al.* 2019; Fitz-James & Cavalli, 2022); 4) acetilação é regulada por duas família de

enzimas, sendo as histonas-acetiltransferases responsáveis pela acetilação, e as histonas-desacetilases com função de desacetiladoras (Glastad *et al.* 2019). Esse processo é reconhecido pela remodelação do nucleossoma cromatina, em que a acetilação é associada com abertura da cromatina e ativação transcricional, enquanto a desacetilação está associada à repressão da transcrição gênica por meio da condensação da cromatina (Glastad *et al.* 2019); 5) classe de RNAs não são traduzidos em proteínas, mas o produto de RNA apresenta função específica, como os não-codificadores (ncRNAs) com regulação de processos celulares (Glastad *et al.* 2019). Esse grupo de ncRNAs abrangem *PIWI-interacting RNAs* (piRNA), *microRNAs* (miRNAs), *small interfering RNAs* (siRNAs), e *long noncoding RNAs* (lncRNAs) (Glastad *et al.* 2019); 6) interação entre cromatinas, em que segmentos de DNA distantes da região promotora de um gene em milhares de pares de bases; entretanto, tem a capacidade de regulação da atividade transcricional desse gene (Ling & Hoffman, 2007). Esse processo ocorre quando os cromossomos se descondensam com o preenchimento da estrutura do núcleo e formação de territórios cromossômicos distintos que se intercalam com territórios cromossômicos vizinhos, e ainda mais distantes (Ling & Hoffman, 2007). Associações de interações de longo alcance intracromossômica e intercromossômica têm sido documentadas, e proteínas de interação com o DNA têm sido indicadas para participar dessas interações (Ling & Hoffman, 2007). Os genes que são expressos monoalelicamente, como genes impressos, têm sido relacionados com essas interações de longo alcance (Ling & Hoffman, 2007; Fitz-James & Cavalli, 2022).

Uma característica do fenômeno epigenético é evidenciada quando um genótipo pode apresentar fenótipos alternativos, baseado no estado epigenético de um ou mais *locus* distribuídos no genoma (Fitz-James & Cavalli, 2022). Isso tem sido observado em sistemas epigenéticos clássicos que foram encontrados por meio de uma aparente instabilidade genética ou desvios das proporções de heranças mendelianas (Richards, 2006).

Atualmente, estudos enfocam o epigenoma, definido como uma abordagem geral do genoma para analisar modificações epigenéticas, que compreendem a meta central de definir as características nas sequências de DNA que participam dos processos epigenéticos (Fitz-James & Cavalli, 2022). Dentre as modificações epigenética, a metilação de DNA tem sido amplamente investigada no genoma dos eucariotos, com a presença de “ilhas” CpG (bases de citosina seguida por guanina), que constituem alvos de modificações covalentes no DNA (Fitz-James & Cavalli, 2022).

Inicialmente, houve uma dificuldade nas investigações do conteúdo de metilação no genoma de *Drosophila*, considerada livre ou com baixos níveis não detectáveis de metilação (Krauss & Reuter, 2011). Isto pode estar relacionado à falta de metodologias disponíveis com sensibilidade para detectar bases metiladas nos genomas nas espécies do gênero *Drosophila* (Krauss & Reuter,

2011). Embora as conclusões tenham sido inferidas para o gênero *Drosophila*, com a investigação de DNA genômico somente de *D. melanogaster* (Krauss & Reuter, 2011). Nas últimas décadas, com novas abordagens de técnicas de biologia molecular, estudos têm sido realizados com *Drosophila*, e o panorama começou a ser modificado (Krauss & Reuter, 2011).

No início do século XX, houve a confirmação da presença de modificações químicas na molécula do DNA genômico em *D. melanogaster* (Achwal *et al.* 1984; Lyko *et al.* 1999; Gowher *et al.* 2000). Essa modificação química foi encontrada em 5-metil-citosinas (5mC) no DNA por meio da técnica de cromatografia líquida de alta pressão (HPLC), em *D. melanogaster* (Achwal *et al.* 1984; Lyko *et al.* 1999). Resultados indicam que a presença de 5mC ocorre com diferentes proporções no DNA de indivíduos ao longo do desenvolvimento biológico em *Drosophila*, com, aproximadamente, 1% de metilação no DNA genômico (Lyko *et al.* 1999).

O mecanismo molecular de metilação no DNA genômico, em *Drosophila*, contém uma única enzima DNA metiltransferases reconhecida como “Dnmt2 only”, devido ao fato de não conter outras DNA metiltransferases canônicas (DNMT1 e DNMT3) (Kraus & Reuter, 2011). Tem sido indicado que a adição de grupo metil em bases de citosinas no DNA é um subproduto da Dnmt2, em *Drosophila*, com funções em tRNAs e DNA (Kraus & Reuter, 2011). Essa proteína Dnmt2 é conservada em diferentes linhagens evolutivas de *Drosophila*, mas com diferenças seletivas positivas nas enzimas que podem alterar a sua função molecular e epigenética (Vieira *et al.* 2018). Em contraste, a família de proteínas Ten-Eleven Translocation (TETs) é responsável pela oxidação e remoção de grupos metil na molécula de DNA, na dinâmica de metilação do genoma (Kraus & Reuter, 2011).

Recentemente, estudos têm investigado a presença de 5mC no DNA genômico, em *D. melanogaster*, com o emprego de diferentes metodologias, dentre elas destacam-se: *Affinity Chromatography* (AC); *Chromatin Immunoprecipitation with 5mC Antibodies* (ChIP), *Capillary Electrophoresis* (CE), *High Pressure Liquid Chromatography* (HPLC), *Slot Blotting* (SB), *Bisulfite Sequencing* (BS), e *Methylation-Sensitive Restriction Endonucleases* (MSRE) (Kraus & Reuter, 2011).

Dentre os diversos métodos, a abordagem com enzimas de restrição tem sido amplamente usada nas análises de metilação em genoma de eucariotos, como a técnica *Methylation-Sensitive Amplified Polymorphism* (MSAP), desenvolvida com uma adaptação de *Amplified fragment length polymorphism* (AFLP) (Yaish *et al.* 2013; Zhou *et al.* 2013; Bulut *et al.* 2020). Essa metodologia emprega um conjunto de enzimas de restrição endonucleases (RE), reconhecidas como *MspI* e *HpaII*, que fazem a varredura por todo o genoma, em regiões com diferenças no estado metilação no DNA (Yaish *et al.* 2013). A metodologia MSAP baseia-se na amplificação seletiva de DNA e comparação entre os perfis de metilação obtidos por *EcoRI-MspI* e *EcoRI-HpaII*, com a distinção

entre sequências metiladas e não metiladas (Yaish *et al.* 2013). Esses dois conjuntos de enzimas RE, *MspI* e *HpaII*, clivam o alvo de sequência CCGG, porém *HpaII* apresenta sensibilidade para dinucleotídeo central CpG é metilado (Yaish *et al.* 2013).

Em paralelo, Deshmukh *et al.* (2018) relatou que a estrutura e organização do genoma apresentam variações na composição de bases CpGs entre as espécies de *Drosophila*, que foi estimado em *Drosophila willistoni* (CG=35%), *Drosophila virilis* (CG=37%), *Drosophila mojavensis* (CG=37%), *Drosophila melanogaster* (CG=42%), e *Drosophila pseudoobscura* (CG=44%).

Aliada às variações na composição CpGs nos genomas nas espécies de *Drosophila*, descobriu-se uma diversidade na metilação no DNA (5mC) no haplogenoma, com a identificação de 22.910 resíduos 5mC no DNA de *Drosophila willistoni*, 21.769 resíduos 5mC no DNA em *Drosophila virilis*, 18.327 resíduos 5mC no DNA de *Drosophila mojavensis*, 300 resíduos 5mC no DNA de *Drosophila melanogaster*, e 16.675 resíduos 5mC no DNA de *Drosophila pseudoobscura* (Deshmukh *et al.* 2018). Esses dados sugerem uma complexidade nos alvos de metilação em sequências de DNA que podem alterar em posições ao longo do genoma entre as espécies de *Drosophila*.

O trabalho de Kraus & Reuter (2011) relatou que diferentes regiões gênicas são alvo de metilação que abrangem sítios promotores, éxons, enhancers, intergênicas, repetições longas terminais (LTR), retrotransposons e elementos genéticos móveis com papel na estabilidade no genoma, em *D. melanogaster*. Adicionalmente, as regiões diferencialmente metiladas estão relacionado a sexo-específico, compensação de dosagem em espécie-específica, com maior quantidade de regiões para fêmeas do que em machos de *D. willistoni* (Garcia *et al.* 2007). Os autores indicaram que a metilação do DNA pode ter diversas funções dentro do gênero *Drosophila*, que pode ser relacionada à sua história evolutiva das espécies e sua ecologia (Garcia *et al.* 2007).

As mudanças nos estados epigenéticos podem contribuir para diversidade de fenótipos que é reconhecido como “epimutações”, que compreendem evidências de alterações epigenéticas hereditárias sem modificações na sequência de DNA, e resultam em diferenças na regulação da expressão gênica e/ou em fenótipos observáveis (Fitz-James & Cavalli, 2022). As epimutações podem contribuir com a sobrevivência em curto prazo de organismos em uma mudança rápida de ambiente (Fitz-James & Cavalli, 2022). Uma questão biológica importante refere-se a quais são as contribuições relativas da epigenética e genética para a formação de novos fenótipos em relação à adaptação dos organismos (Fitz-James & Cavalli, 2022). Nesse contexto, descobertas de Lyko *et al.* (1999) indicam que alterações no padrão de metilação no DNA genômico podem prejudicar a viabilidade de moscas transgênicas (expressão *Dnmt1* e *Dnmt3a*), o que pode implicar funções no desenvolvimento de *Drosophila*.

Outro aspecto central é efeito de origem parental com modificações dos padrões epigenéticos distribuídas nos cromossomos X e Y, em *D. melanogaster* (Lemos *et al.* 2014). Resultados obtidos de cruzamentos de moscas com diferenças na transmissão dos cromossomos X e Y, materno_(M) (XY_MX_P) x (XY_PX_M) paternos (P) sugerem que modificação na transmissão do cromossomo Y entre linhagens germinativas masculinas e femininas apresentam diferenças na expressão gênica em todo o genoma (Lemos *et al.* 2014). Genes diferencialmente expressos contribuem, especificamente, com os tecidos de testículos e células do intestino médio (Lemos *et al.* 2014). Isto sugere que as modificações na informação presente na cromatina com efeito na origem parental contribuem para a regulação de diferentes genes na prole F1 (Lemos *et al.* 2014).

Encontraram-se dados obtidos por Zhang *et al.* (2020), silenciamento de genes por meio do efeito de posição de variegação (PEV), em *D. melanogaster*. Fenótipo na cor dos olhos compostos em *Drosophila*, cores intermediárias entre branco e vermelho em diferentes regiões em uma mesma mosca constituem o efeito de PEV (Zhang *et al.* 2020). Esse fenótipo também apresenta respostas aos micronutrientes de alimentação de larvas de *D. melanogaster*, como os componentes tóxicos de cádmio (Zhang *et al.* 2020).

Compostos nutricionais ingeridos por moscas maternas, com alto teor de açúcar, resultam em mudanças de fenótipos na sua prole, que sugerem reprogramação de via metabólica de glicose, em *D. melanogaster* (Buescher *et al.* 2013). Relataram-se os componentes dos recursos alimentares de origem parental com efeito de transmissão de fenótipos de forma transgeracional em *Drosophila* (Matzkin *et al.* 2013). Uma hipótese levantada refere-se ao fato de que podem ocorrer diferenças na distribuição da metilação do DNA nos ovos, como: modelo 1) com diferentes subconjuntos em diferentes regiões (Boffelli *et al.* 2014); modelo2) metilação concentrada em um subconjunto de núcleos embrionários (Boffelli *et al.* 2014); modelo 3) todos os núcleos têm alguma metilação (Boffelli *et al.* 2014). Entretanto, os estudos, ainda, não conseguiram distinguir essas possibilidades referentes às dificuldades em técnicas moleculares (Boffelli *et al.* 2014) (Figura 6).

Os dados na literatura indicam diversidade em conteúdo CpG e metilação nos genomas de *Drosophila*, que podem implicar funções biológicas na plasticidade de fenótipos, e, possivelmente, ter uma relação com a história de vida e ecologia das espécies de *Drosophila* (Garcia *et al.* 2007; Buescher *et al.* 2013; Lemos *et al.* 2014).

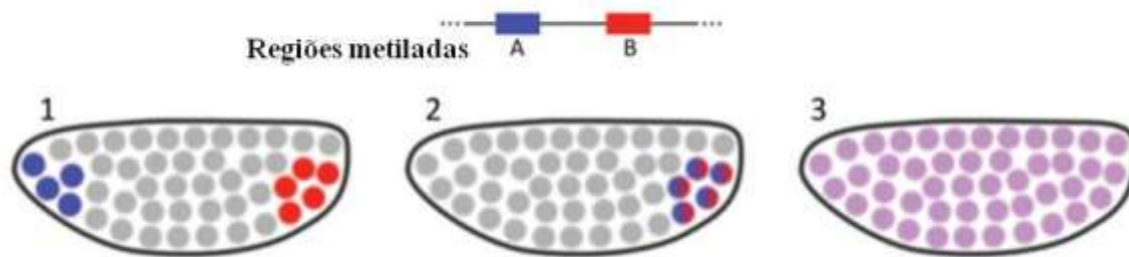


Figura 6. Ilustração com as hipóteses dos padrões de distribuição de m5C no embrião de *Drosophila*. Representação de duas regiões metiladas é indicada por meio de blocos em azul (A) e vermelho (B). No modelo 1, um subconjunto de núcleos é metilado na região A e outro na região B, em posições diferentes do embrião. No modelo 2, ambas as regiões são metiladas em um subconjunto de núcleos. No modelo 3, a metilação das duas regiões é distribuída, difusamente, por meios dos núcleos em todo o embrião. Esse modelo considera que todos os núcleos apresentam pouca metilação em ambas as regiões. Os três modelos apresentam padrões distintamente diferentes na distribuição da metilação, embora, ainda, não tenha sido testada a metilação em células específicas do embrião (como em todos os estudos até o momento), os padrões são indistinguíveis. Modificado de Boffelli *et al.* (2014).

1.6 Modelo biológico *Drosophila gouveai* (cluster *D. buzzatii*)

Dentre as espécies do cluster *Drosophila buzzatii* (grupo *Drosophila repleta*), *Drosophila gouveai* Tidon-Sklorz & Sene (2001) apresenta características diagnósticas que compreendem: morfótipo do *edeago* tipo B (Silva & Sene, 1991), inversão cromossômica $2e^8$ (Ruiz & Wasserman, 1993; Manfrin *et al.* 2006), e placa metafásica do tipo I (Baimai *et al.* 1993). Esta espécie apresenta desenvolvimento biológico obrigatoriamente em tecidos necróticos de cactos na natureza (Pereira *et al.* 1983).

A distribuição geográfica de populações de *D. gouveai*, na natureza, pode ser representada por duas rotas distintas, sendo estas: (I) porção norte e nordeste do Brasil, que compreendem áreas das localidades de Petrolina (Pernambuco), Xique-Xique, Santo Inácio, Barreiras, Ibotirama, Caririnha (Bahia); (II) porção do centro e sudeste do Brasil, nas localidades de Cristalina (Goiás), Furnas (Minas Gerais), Miranda, Bela Vista (Mato Grosso do Sul), Altinópolis, e Analândia (São Paulo) (Pereira *et al.* 1983; Tidon-Sklorz & Sene, 2001). A vegetação nestas regiões é distinta, que compreende a formação de Florestas Tropicais Sazonalmente Secas nas áreas do norte e nordeste, e formações rochosas, campos rupestres e granitos nas áreas do centro e sudeste, no Brasil (Tidon-Sklorz & Sene, 2001).

Evidências indicam que os indivíduos das populações de *D. gouveai* ocorrem em áreas com grande variedade de espécies de cactos dos gêneros *Pilosocereus* e *Cereus*, na porção norte e nordeste do Brasil (Pereira *et al.* 1983). Em contraste, os indivíduos de populações de *D. gouveai* estão associados com o cacto *Pilosocereus machrisis*, na porção sudeste do país (Tidon-Sklorz & Sene, 2001; Moraes & Sene, 2002; Manfrin *et al.* 2006).

Informações obtidas por Tidon-Sklorz & Sene (2001) indicam que *D. gouveai* apresenta como nicho de reprodução, alimentação, e desenvolvimento biológico, principalmente, em espécies de cactos colunares. As populações de *D. gouveai* ocorrem principalmente em sítios criadores do cacto colunar *Pilosocereus machrisii* Dawson, 1960 em Altinópolis (São Paulo), e pode eventualmente usar os tecidos do cacto palmar *Opuntia ficus-indica* Miller, 1788 em Bela Vista, Mato Grosso do Sul (Pereira *et al.* 1983).

A organização e estrutura geográfica populacional de *D. gouveai*, em uma das suas populações, pode ser influenciada por associação de organismos com a disponibilidade do seu cacto hospedeiro, e formação de grupos endogâmicos (Moraes *et al.* 2002). Os autores relataram a presença do cacto hospedeiro *P. machrisii*, nesta localidade, que apresenta estrutura colunar e seu tecido em decomposição pode permanecer somente por alguns dias (Moraes *et al.* 2002; Manfrin *et al.* 2006). Isto sugere que os tecidos em decomposição poderiam servir como substratos para o desenvolvimento de uma geração desta mosca (Moraes *et al.* 2002). De fato, uma tendência tem sido indicada que os indivíduos de *Drosophila* se dispersam apenas para encontrar um tecido de cacto em decomposição (Moraes *et al.* 2002; Manfrin *et al.* 2006). Esta população de *D. gouveai* está teoricamente isolada em uma área de colina de arenito, e não ocorre a troca de material genético por meio de fluxo gênico com outras populações (Moraes *et al.* 2002; Manfrin *et al.* 2006). Portanto, a organização de estrutura endogênica populacional pode ser um dos eventos que reduz a variação genética nesta população de *D. gouveai* (Moraes *et al.* 2002; Manfrin *et al.* 2006).

Dados genéticos obtidos com microssatélites, com repetições dinucleotídeos e trinucleotídeos, permitiram agrupamento em duas populações de *D. gouveai*, sendo estas: (I) indivíduos da população de em Analândia (SP); (II) outro grupo com os indivíduos das populações nas localidades de Altinópolis (SP), Cristalina (GO), Furnas (MG), e Ibotirama (BA) (Moraes & Sene, 2007). Os resultados deste estudo indicaram que as populações estão no processo de isolamento e redução do fluxo gênico, com perda da diversidade genética e endogamia, e também encontraram alelos privados, como por exemplo, na localidade de Serrinha (Ibotirama, SP), em *D. gouveai* (Moraes & Sene, 2007).

Adicionalmente, variações em haplótipos no gene mitocondrial *Citocromo Oxidase subunidade I* (mtCOI) sugerem dois agrupamentos entre as populações de *D. gouveai*, sendo estas: (I) um agrupamento com espécimes das populações nas localidades de Altinópolis (SP), Cristalina

(GO), e Furnas (MG); (II) um segundo agrupamento com espécimes na população localizada em Analândia (SP) (Manfrin *et al.* 2001).

Em contraste, diferenças morfológicas com associação entre altitude e a estrutura das asas (tamanho e forma), em machos e fêmeas, que permitem agrupar os espécimes de populações de *D. gouveai* entre as regiões geográficas: (grupo I) Ibotirama; (grupo II) localidades de Cristalina (GO), Furnas (MG), e Altinópolis (SP) (Moraes & Sene, 2007).

Evidências indicam que os sítios de criação, em tecidos necróticos de cactos, podem influenciar a plasticidade de fenótipos em *D. gouveai* (Soto *et al.* 2007, 2008). De acordo com Soto *et al.* (2008), descobriu-se a formação de asas grandes em adultos que se desenvolveram em *C. hildmaniannus* em comparação com os que foram criados em *P. machrisis*. O fenótipo de viabilidade de larvas para adultos é maior em tecidos de *P. machrisis* do que em *C. hildmaniannus* (Soto *et al.* 2007). Adicionalmente, os organismos dos indivíduos apresentam respostas com diferentes tempos de desenvolvimento ao uso de compostos químicos presentes nos tecidos de cactos, com relato de, aproximadamente, 280 horas em *P. machrisis* e 295 horas em *C. hildmaniannus*, em *D. gouveai* (Soto *et al.* 2007). Outro fenótipo importante é o parêntese reprodutor masculino (*edeago*), que apresenta rápida mudança em resposta aos compostos do sítio de desenvolvimento larval, com maior tamanho em moscas adultas que emergiram em tecidos do cacto *P. machrisis* em relação a *C. hildmaniannus* (Soto *et al.* 2010) (Figura 7).

A plasticidade fenotípica nos traços de viabilidade, tempo de desenvolvimento, morfologia do aparelho reprodutor masculino, e estrutura da asa resultam em diferenças no *fitness* dos indivíduos de *D. gouveai*, no uso dos cactos *P. machrisis* do que em *C. hildmaniannus*, que pode contribuir no processo de adaptação ecológica (Soto *et al.* 2007, 2008, 2010). De acordo com os dados expostos anteriormente, o tempo de desenvolvimento pode ser um resultado que envolve o metabolismo dos compostos dos tecidos de cactos, com maior eficiência no cacto *P. machrisis* do que *C. hildmaniannus* (Soto *et al.* 2007). Diferenças na viabilidade podem resultar na sobrevivência diferenciada dos organismos na exploração de recursos em tecidos de cactos hospedeiros (Soto *et al.* 2008). Ademais, o aparelho reprodutor masculino pode contribuir para o mecanismo de reprodução desses organismos em relação ao acasalamento entre machos e fêmeas (Soto *et al.* 2010). Por fim, o tamanho das asas pode contribuir para o processo de corte sexual e dispersão geográfica das moscas em *D. gouveai* (Soto *et al.* 2010).

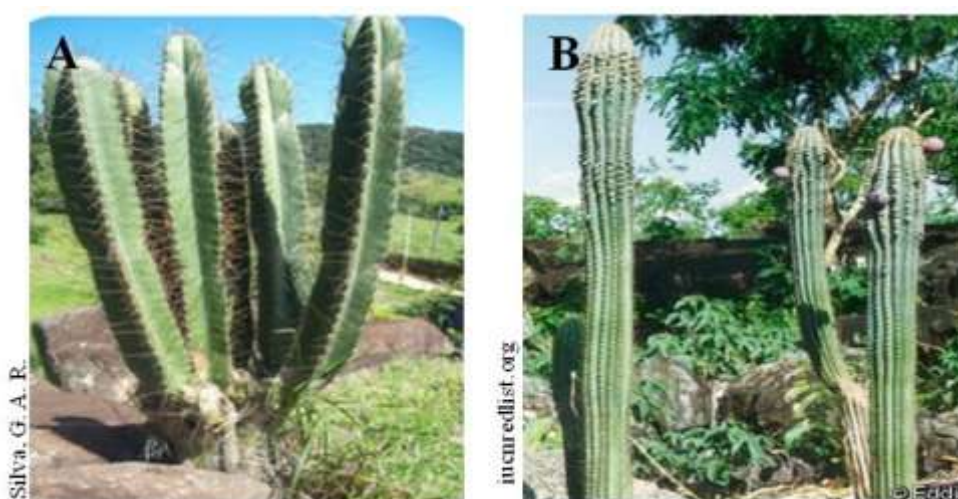


Figura 7. Ilustrações das espécies de cactos *C. hildmaniannusi* (A) e *P. machrisis* (B).

Resultados obtidos por Machado *et al.* (2006) indicam diferenças nos fenótipos de fecundidade e reprodução de espécime de *D. gouveai*, nas localidades de Ibotirama (BA) em relação a Bela Vista (MS). Esses autores encontraram isolamento reprodutivo interespecífico, em cruzamentos experimentais, com os indivíduos de *D. gouveai* (Ibotirama, BA) e *D. buzzatii* (regiões da Bahia), consideradas espécies relacionadas no cluster *D. buzzatii* (Machado *et al.* 2006).

O conjunto de dados genéticos e morfológicos indica que há graus de diferenciação populacional, que podem estar relacionados a isolamento demográfico, diferenças no processo de reprodução, e respostas aos nichos ecológicos nos cactos *P. machrisis* e *C. hildmaniannus* (Moraes & Sene, 2007). Tem sido observada uma perda na diversidade genética em populações de *D. gouveai*, que pode estar relacionada a processos de endogamia, embora tenham sido relatadas plasticidades em fenótipos em relação aos compostos dos cactos *P. machrisis* e *C. hildmaniannus* (Moraes & Sene, 2007). A distribuição na localidade de Ibotirama, no estado Bahia, chama atenção pelo fato de que os espécimes da população de *D. gouveai* estão em contato com uma diversidade de cactos, como *Pilosocereus* e *Cereus* (Pereira *et al.* 1983; DeBrito *et al.* 2002), e também apresentam isolamento reprodutivo com outra espécie irmã no cluster *D. buzzatii* (Machado *et al.* 2006).

Um desafio intrigante é investigar mudanças epigenéticas, metilação de DNA, que podem contribuir para a regulação de genótipo, vias moleculares, que resultam em múltiplos fenótipos em *D. gouveai* quando os organismos usam diferentes espécies de cactos hospedeiros, como *P. machrisis* e *C. hildmaniannus*. Evidências sugerem que os compostos de alcaloides podem influenciar no padrão de metilação do DNA (Bulut *et al.* 2020). É importante ressaltar que, ainda, não há dados sobre variações de *loci* metilados no DNA genômico em pupas e moscas adultas, em que ocorre o processo de diferenciação de tecidos, e efeito de origem parental que podem contribuir

para a transmissão de informações epigenéticas para a prole quando ocorre a mudança de ambientes ecológicos, em *Drosophila*.

Hipótese

2.Hipótese de estudo

Mudanças no estado de metilação de citosinas no DNA podem contribuir para a regulação diferencial de regiões genômicas e em vias de sinalização, quando indivíduos se desenvolvem em diferentes ambientes hospedeiros, como *Cereus hildmmanianus* e *Pilosocereus machrisis*, na espécie cactofílica *D. gouveai*. Isto poderia contribuir para a regulação de genótipo e plasticidade de fenótipos durante as mudanças de nichos ecológicos na espécie cactofílica *D. gouveai*.

Objetivos

3.Objetivos

3.1 Objetivo Geral

O objetivo geral do estudo consiste em analisar a variação nos estados de metilação do DNA e identificar suas regiões genômicas, com as vias moleculares, nos indivíduos quando se desenvolvem em diferentes ambientes ecológicos, nos cactos *C. hildmmanianus*, *P. machrisis*, e meio padrão de laboratório (controle), em *Drosophila gouveai*.

3.2 Objetivos específicos

- Analisar as variações em *loci* sensíveis à metilação de citosinas no DNA genômico em indivíduos (pupas, tecidos reprodutivos em moscas fêmeas e outros em machos) que se desenvolveram em diferentes nichos ecológicos, em *D. gouveai*.
- Isolar e sequenciar regiões alvos de metilação no DNA em indivíduos adultos (ovários e testículos).
- Determinar a localização genômica das sequências alvos de metilação e suas funções biológicas em indivíduos adultos, com uso de diferentes ambientes.
- Compreender as funções das vias moleculares e celulares que envolvem o sistema dos organismos que usaram diferentes recursos nutricionais, em *D. gouveai*.

Material e Métodos

4. Material e Métodos

4.1 Stock de *Drosophila*

As amostras biológicas são provenientes de indivíduos representantes de *Drosophila gouveai* (cluster *Drosophila buzzatii*, grupo *Drosophila repleta*) (stock J79). Foram realizadas coletas prévias de moscas fêmeas naturalmente inseminadas na localidade de Ibotirama (Estado da Bahia, Brasil)(12.16°S, 43.04°W). O stock foi criado com um conjunto de fêmeas da mesma espécie e localidade, chamado massal, e mantido no Laboratório de Genética e Evolução, da Faculdade de Filosofia, Ciências, e Letras de Ribeirão Preto.

4.2 Enzimas

Proteinase K	Boehinger Mannheim
RNase	Thermo Scientific
<i>Taq</i> DNA polimerase	Promega
<i>Eco</i> RI	Thermo Scientific
<i>Hpa</i> II	Thermo Scientific
<i>Msp</i> I	Thermo Scientific
T4 DNA Ligase	Thermo Scientific

4.3 Tampões, Soluções, e Meios

Tampão de Proteinase K	0,375M NaCl; 5mM EDTA
Tampão TBE 10X 0,9 M pH 8,0 (1000ml)	0,89M Tris-Base; 0,89M ácido bórico (PM = 61,83); 20 mM EDTA
Tampão de corrida TBE 1X 0,9 M (1000ml)	100 mL de Tampão TBE 10X; 900 mL de H ₂ O destilada
Tampão de amostra	900 µL de bromofenol; 900 µL de xilenocianol; 900 µL de TBE 10X; 4,5 mL de Ficol 30%; 1,8 mL de EDTA 0,5 M pH 8,0; 3,6 g de sacarose
Tampão 1X Tango Buffer	33 mM Tris-acetato (pH 7.9 at 37°C), 10 mM acetate de magnésio, 66 mM acetate de potásio, 0.1 mg/mL BSA

Tampão 10X T4 DNA Ligase	400 mM Tris-HCl, 100 mM MgCl ₂ , 100 mM DTT, 5 mM ATP
Tampão Enzimático	Tris-HCl 20 mM (pH 7,5), KCl 50 mM, DTT 1 mM, EDTA 0,1 mM e glicerol a 50% (v/v)
Solução de acrilamida/bisacrilamida (29:1)	29 g de acrilamida (PM= 71,08); 1 g de bisacrilamida (N,N'-metileno-bis-acrilamida; PM= 154,17), 100 mL de água destilada
Solução de EDTA 0,5 M	186 g de EDTA (PM = 372,24); 1 L de água destilada
Solução hipersaturada de persulfato de potássio	650 mg de persulfato de potássio (PM = 270,32); 6,5 mL de água destilada
Solução fixadora	167 mL de álcool etílico; 7 mL de ácido acético glacial e 826 mL de H ₂ O
Solução reveladora	22,5g de NaOH (PM= 40,00); 1L de água destilada
Solução de nitrato de prata	100 g de nitrato de prata; 1 L de água
Solução de RNase	10mg/ml H ₂ O
Solução PEG 50%	50% (w/v) de polietilenoglicol 4000
Mix para gel não desnaturante 10%	6,66 mL de solução acrilamida/bisacrilamida; 1,4 mL de glicerol; 9,7 mL de água destilada; 2,0 mL de TBE 10%

TEMED	(N,N,N',N'-tetrametiletilenodiamina)
SYBR Safe	10,000X, in DMSO

4.4 Kits

Go Taq DNA Polimerase	Promega
GFX PCR DNA and Band Purification	GE Healthcare
ABI PRISM BigDye™ Terminator	Thermo Scientific
Cycle Sequencing Ready Reaction	

4.5 Experimentos: ambientes heterogêneos

Os indivíduos de *D. gouveai* (stock J79) foram criadas em condições idênticas por 24 gerações, com cruzamentos entre parentais e prole, em 35 ml de meio padrão de laboratório. Esse meio de cultura padrão é composto por farinha de milho, mel, *Sacharomices cerevise*, Agar, e Nipagin (4-hidroxibenzoato de metila). Os cruzamentos entre as moscas foram realizados para fornecer condições de endogamia, com aumento do parentesco entre os organismos, e diminuição da variação genética entre os indivíduos.

Em paralelo, amostras de tecidos vegetais foram coletadas, separadamente, para os cactos *Cereus hildmmanianus* (região de Sertãozinho-SP) e *Pilosocereus machrisis* (região de Altinópolis-SP), na natureza, e preservados em freezer -20°C no Laboratório de Genética e Evolução. Os meios de cactos foram preparados com modificações do protocolo de Fanara *et al.* (1999), com a seguinte composição: 500g de cada tecido separadamente por espécie de cacto, previamente descongelado e homogeneizado em liquidificador, 5g de fermento biológico (*Sacharomices cerevise*), 5g de Agar, 20 ml de Nipagin, completa com 800 ml H₂O para o volume de 1l de meio de cultura.

Experimentos foram conduzidos com cultivos em três ambientes, que compreenderam o meio padrão de laboratório (controle), meio com tecidos de *C. hildmmanianus*, e outro meio com *P. machrisis*. Esses três meios oferecem a oportunidade de avaliar as respostas fisiológicas e moleculares dos organismos de indivíduos de *D. gouveai* em mudanças de cactos hospedeiros, que ocorrem nas áreas de distribuição geográfica da espécie, e em um ambiente com meio padrão de laboratório. Isto inclui as respostas de um genótipo parental com mudanças de fenótipo da prole em diferentes ambientes ecológicos, em *D. gouveai*. Dados prévios de Soto *et al.* (2007, 2008)

relataram mudanças em traços relacionados ao *fitness* em *D. gouveai*, quando os indivíduos se desenvolvem em *C. hildmmanianus*, e *P. machrisis*.

Inicialmente, conjuntos de 50 casais de moscas adultas (50 machos e 50 fêmeas) e maduras sexualmente, com três réplicas, em vidros com 35 ml de meio padrão de laboratório. Essas moscas adultas foram obtidas, previamente, com cruzamentos endogâmicos de *D. gouveai* em laboratório. Em cada vidro com os casais de moscas, com meio padrão, foram observados por 24 horas para verificar a postura de ovos, e, posteriormente, com 24 horas para a presença de larvas de 1° instar. Todos os experimentos foram realizados com condições de $24 \pm 1^\circ\text{C}$ e 12hs/12hs (luz/escuro), em fotoperíodo.

Na segunda etapa, foram coletadas 30 larvas de 1° instar, em meio padrão, e transferidas para vidros com os seguintes meios de cultura: meio padrão de laboratório, meio com tecidos de *C. hildmmanianus*, e outro meio com *P. machrisis*. Os experimentos foram realizados com três réplicas biológicas, com as mesmas condições mencionadas anteriormente em fotoperíodo. Os organismos se desenvolveram de larvas de 1° instar a moscas adultas em cada meio de cultura.

Na terceira etapa, as moscas adultas foram transferidas para novos meios de cultura com a mesma composição, que os organismos usaram anteriormente para o processo de desenvolvimento. As moscas adultas amadureceram sexualmente, processo seguido por cruzamento, e formação de uma nova geração. Esse experimento foi conduzido para obter amostras de indivíduos com a segunda exposição ao mesmo ambiente ecológico, e no controle.

As amostras biológicas foram identificadas e coletadas nas fases de pupas (± 215 hs) e moscas adultas (± 269 hs), com indivíduos criados nos meios de padrão de laboratório, *C. hildmmanianus*, e *P. machrisis*. As características fenotípicas dessas amostras foram as seguintes: pupa em fase estacionária, e coloração amarela; moscas adultas com 24 hs após emergirem da pupa, virgens (fêmeas, machos), e com desenvolvimento de asas, cabeça e tórax. Em amostras de moscas adultas, foram realizados os procedimentos de dissecação dos tecidos reprodutivos de ovários, e outros em testículos, com o auxílio de lupas estereomicroscópicas. Todas as amostras biológicas foram coletadas com nitrogênio líquido, e armazenadas em freezer -80°C .

4.6 Extrações de DNA genômico

As moléculas de DNA genômico foram isoladas por meio do método de extração e precipitação com NaCl (Olerup & Zetterquist, 1992). Os procedimentos foram realizados com *pool* de indivíduos ($n=5$ a 10), nas fases de pupas, e tecidos reprodutivos de fêmea e outro de machos. Essas amostras foram homogeneizadas em 80 μL de Tampão de Proteinase K [0,375M NaCl; 5mM EDTA], 10 μL de 20% SDS, 8,0 μL Proteinase K (20 ng/mL), e 283,3 10 μL de água Milli-Q,

RNase (10mg/ml), e incubadas a 55°C por 16hs. Em seguida, foram adicionados 120 µL de 5M NaCl, com agitação manual por 15 segundos, e centrifugação a 13.000 rpm por 8 minutos. O sobrenadante foi transferido para outro microtubo, adicionado 1,0 mL de etanol a -20°C, homogeneizado e centrifugado a 13.000 rpm por 10 minutos a 4°C. Por fim, o sobrenadante foi descartado e, posteriormente, o *pellet* foi desidratado. O DNA foi hidratado com 50 µL de água Milli-Q.

4.7 Quantificação de DNA genômico

As amostras de DNA genômico foram quantificadas por meio do espectrofotômetro Nanodrop 2000c (Thermo Scientific, Vilnius, Lituânia), com avaliação das amostras nos comprimentos das ondas 260 e 280 nm. Razões com valores de 1,8 a 2,0 foram considerados ideais. Posteriormente, a integridade das moléculas de DNA foi avaliada por meio de gel de agarose 0,8% (Apêndice A) em eletroforese com Tampão TBE 1X, e corridas com 80mA por 40 minutos. Por fim, a visualização das moléculas de DNA em luz UV, no equipamento transiluminador.

4.8 Ensaios *Methylation Sensitive Amplification Polymorphism* (MSAP)

4.8.1 Digestões enzimáticas

As moléculas de DNA genômico, em cada amostra, foram digeridas em duas reações independentes com as enzimas *EcoRI* e *HpaII*, e outra com *EcoRI* e *MspI* (Thermo Scientific). A enzima *EcoRI* com clivagem nos sítios 5'-GAATTC-3'. Já as enzimas *HpaII* ou *MspI* com clivagem nos sítios 5'-CCGG-3', reconhecidas como isoesquizômeros. Em contraste, ambas as enzimas apresentam diferenças na sensibilidade de metilação de citosinas no DNA, sendo reconhecido que *HpaII* é sensível à metilação interna de bases de citosinas, e *MspI* é sensível a metilação externa de citosinas.

Na primeira reação enzimática, 300ng de DNA de cada amostra foram pipetados em dois microtubos independentes, e acrescentado 10 X Tampão buffer Tango (33 mM Tris-acetato (pH 7,9), 10 mM acetato de magnésio, 66 mM acetato de potássio, 0,1 mg/mL BSA) e 5U *EcoRI* na mistura da reação enzimática, que compreenderam os componentes: 13,9 µL H₂O Milli-Q; 2 µL 10X Tampão Buffer Tango; 0,6 µL Enzima *EcoRI*; 3,5 µL DNA genômico, com volume final de 20µL. As amostras foram mantidas a 37°C por, aproximadamente, 16 horas, e inativação a 80°C por 20 minutos, em termociclador.

Na segunda reação enzimática, foi adicionado 5U de *HpaII* para uma amostra e em paralelo 5U *MspI* para a outra, e mantidas a 37°C por aproximadamente 16 horas, e inativadas a 80°C por 20 minutos, em termociclador. De forma que uma amostra de DNA foi tratada com *HpaII* e outra com *MspI*, em duas reações independentes (Apêndice A). Todas as amostras biológicas foram

submetidas às reações com *EcoRI/HpaII* e outra com *EcoRI/MspI*. Posteriormente, os fragmentos gerados foram ligados com sequências de adaptadores específicos complementares às extremidades clivadas pelas duas enzimas.

4.8.2 Ligação com adaptadores

Foram pipetados 5 µL do produto de digestão em microtubos, e acrescentados 2,0 µL de 10X Tampão T4 DNA ligase, 5 Weiss U da enzima T4 DNA ligase (Thermo Scientific), 2 pmol do adaptador para região de *EcoRI* (5'-CTCGTAGACTGCGTACC-3'/5'-AATTGGTACGCAGTCTAC-3') e 20 pmol do adaptador para região de *HpaII/MspI* (5'-GACGATGAGTCTAGAA-3'/5'-CGTTCTAGACTCATC-3'), 2,0 µL tampão PEG (50% W/V), e 8,5 µL de H₂O Milli-Q. Os adaptadores *EcoRI* foram previamente pareados por meio de gradual decréscimo de temperatura de 56° a 10°C por 46 minutos, e o mesmo foi realizado com os adaptadores *HpaII/MspI*, antes da reação de ligação. As reações de ligação dos adaptadores com as moléculas clivadas foram realizadas com incubação a 16°C por 14 horas, e inativação a 65°C por 30 minutos. Os produtos de ligação foram diluídos 1:10 com H₂O Milli-Q.

4.8.3 Amplificações Pré-Seletivas

Reações de pré-amplificações foram realizadas com 3,0 µL do produto de ligação (1:10), 2,0 µL Tampão 10X Buffer, 1,6 µL MgCl⁺² (25 mM), 0,5 µL dNTPs (2.5 mM), 1U *Taq* DNA polimerase (Promega Corporation[®]), com os primers 2pmol E0 (5'-GACTGCGTACCAATTCA-3') e 2pmol HM0 (5'-GATGAGTCTAGAACGGT-3'). As reações com *Polymerase Chain Reaction* (PCR) foram realizadas com o ciclo de programação: 94°C por 1 minuto (desnaturação), 25 ciclos com 94°C por 30 segundos (desnaturação), 65°C por 30 segundos (pareamento) e 72°C por 1 minuto (extensão), e extensão final de 72°C por 10 minutos.

4.8.4 Amplificações Seletivas

Reações de amplificações seletivas foram realizadas com 3,0 µL do produto de pré-amplificação, 2,0 µL Tampão 10X Buffer, 1,6 µL MgCl⁺² (25 mM), 0,5 µL dNTPs (2.5 mM), 1U *Taq* DNA polimerase, com os *primers* 2 pmol dos E1-E6 e 2 pmol HM1-HM6, e completa o volume de 20 µL com H₂O Milli-Q (Tabela 1). As reações de PCR foram realizadas com os ciclos de programação: 94°C por 1 minuto, 13 ciclos por 94°C por 30 segundos, 65°C por 30 segundos (redução 0,7°C por ciclo), 72°C por 1 minuto, e 23 ciclos com 94°C por 30 segundos, 56°C por 30 segundos, 72°C por 1 minuto, e extensão final com 72°C por 10 minutos.

A síntese dos *primers* (oligonucleotídeos) para as reações de ligação de adaptadores, pré-amplificação, e amplificações seletivas foram obtidas de acordo com o descrito por Zhou *et al.* (2013). Foram avaliados 12 combinações de *primers* seletivos no presente estudo (Tabela 1).

Tabela 1. Relação de *primers* usados para as análises de MSAP, com adaptadores, pré-amplificações, e amplificações seletivas, com as suas respectivas combinações.

	EcoRI		MspI/HpaII	
	Adaptador E1	5'-CTCGTAGACTGCGTACC-3'	Adaptador HM1	5'-GACGATGAGTCTAGAA-3'
	Adaptador E2	5'-AATTGGTACGCAGTCTAC-3'	Adaptador HM2	5'-CGTCTAGACTCATC-3'
	PreAmp E0	5'-GACTGCGTACCAATTCA-3'	PreAmp HM0	5'-GATGAGTCTAGAACGGT-3'
Amplificação Seletiva	(1)E1	5'-GACTGCGTACCAATTCAAC-3'	HM1	5'-GATGAGTCTAGAACGGTTC-3'
	(2)E2	5'-GACTGCGTACCAATTCACC-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(3)E3	5'-GACTGCGTACCAATTCACG-3'	HM3	5'-GATGAGTCTAGAACGGTAG-3'
	(4)E4	5'-GACTGCGTACCAATTCACT-3'	HM4	5'-GATGAGTCTAGAACGGTTG-3'
	(5)E5	5'-GACTGCGTACCAATTCAGC-3'	HM5	5'-GATGAGTCTAGAACGGTCC-3'
	(6)E6	5'-GACTGCGTACCAATTCAGG-3'	HM6	5'-GATGAGTCTAGAACGGTGA-3'
	(7)E1	5'-GACTGCGTACCAATTCAAC-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(8)E1	5'-GACTGCGTACCAATTCAAC-3'	HM3	5'-GATGAGTCTAGAACGGTAG-3'
	(9)E2	5'-GACTGCGTACCAATTCACC-3'	HM1	5'-GATGAGTCTAGAACGGTTC-3'
	(10)E3	5'-GACTGCGTACCAATTCACG-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(11)E3	5'-GACTGCGTACCAATTCACG-3'	HM4	5'-GATGAGTCTAGAACGGTTG-3'
	(12)E4	5'-GACTGCGTACCAATTCACT-3'	HM5	5'-GATGAGTCTAGAACGGTCC-3'

4.9 Eletroforese

4.9.1 Gel de poliacrilamida

Os produtos de ampliações seletivas de MSAP foram visualizados em géis de poliacrilamida 10%, não desnaturante. Os géis foram preparados com os seguintes componentes: 6,66 mL de solução acrilamida + bisacrilamida (29:1); 1,4 mL de glicerol; 9,7 mL de água destilada; 2,0 mL de TBE 1X; 15,0 µL TEMED (N,N,N',N'-tetrametiletilenodiamina); 500 µL Solução hipersaturada de persulfato de potássio (100mg/mL).

Com o suporte de duas placas de vidros (16 cm x 14 cm), fixadas à base com grampos e os pentes do gel, os componentes dos géis de poliacrilamida foram adicionados ao recipiente. Posteriormente, houve o processo de polimerização por cerca de 20 minutos. Na próxima etapa, foram retirados os pentes, e a placa colocada em cuba de eletroforese vertical contendo o tampão TBE 1X (o mesmo usado no preparado dos géis) nas porções superior (anódica) e inferior (catódica). Em cada poço dos géis, aplicou-se 5µL do produto de amplificação seletiva e 4µL de tampão de carregamento (*loading buffer* – TBE 1X; Xilenocianol 0,25%; Azul de Bromofenol 0,25%; EDTA 0,1M pH8,0; Ficoll 15%). O marcador molecular usado foi de 50pb (Promega Corporation®). A cuba foi conectada a uma fonte de alta voltagem e realizada a separação dos fragmentos por 2 horas, com 200V e 400mA.

A coloração dos géis foi realizada com modificação do protocolo de Sanguinetti *et al.* (1994). Os géis foram colocados em recipientes contendo 100 ml de solução fixadora (167 ml de álcool etílico, 7 ml de ácido acético glacial, e 826 ml de H₂O destilada), com adição de 4,0 ml de solução nitrato de prata (0,1 g/ml) por gel, e agitação por cerca de 15 minutos. Em seguida, a solução fixadora foi descartada e os géis lavados em água deionizada a 60°C por 15 segundos. Por fim, adicionou-se 100 mL de solução reveladora (2,25 g de NaOH, 1L de H₂O destilada), e 2 ml de formaldeído por cada gel, com agitação dos géis até aparecer bandas nítidas. A solução reveladora foi descartada e bloqueia a reação com a adição de 100 ml de solução fixadora. Os dados dos géis foram foto-documentados, e isoladas as regiões com informação sobre o estado de metilação de DNA.

4.9.2 Gel de Agarose

Foram preparados e usados géis de agarose em duas concentrações, tais como gel de agarose 0,8% para avaliar a integridade das moléculas de DNA genômico, e outro de gel de agarose 1% para verificar às reamplificações dos produtos de PCR, fragmentos isolados em gel de

poliacrilamida. Os compostos dos géis de agarose 08%: 0,8g de agarose e 100ml de Tampão 1X TBE. Já nos géis de agarose 1%, foi usado 0,7g de Agarose e 70ml de Tampão 1X TBE. Esses compostos dos géis seguiram o processo de homogeneização em uma temperatura de aproximadamente de 100°C, com leve agitação manual. Em seguida, aplicou-se 10 µL e 7µL do corante SYBR safe na solução do gel, na concentração de 0,8% e 1% respectivamente, e os componentes foram colocados no molde de uma placa com pentes. O processo de polimerização ocorreu por cerca de 30 minutos. Posteriormente, foram retirados os pentes e colocado o gel em cuba horizontal de eletroforese, contendo Tampão TBE 1X, e as corridas ocorreram com 80mA por 40 minutos. A visualização dos fragmentos foi realizada por meio de exposição à luz UV no equipamento transiluminador.

4.10 Interpretações do padrão de metilação

Os padrões de metilação foram avaliados quanto à presença e ausência de bandas por meio de comparações lado a lado dos perfis moleculares obtidos nos ensaios *EcoRI/HpaII* e *EcoRI/MspI*, com ampliações seletivas de PCR (Tabela 2). Foram selecionados os fragmentos de DNA com extensão molecular de 100pb a 1000pb para as análises estatísticas de MSAP, e determinação do tamanho molecular de cada *locus* no gel com base no marcador de peso molecular 50pb (Promega Corporation®). As bandas foram convertidas em matriz binária, em que a presença de banda (+) corresponde a “1” e a sua ausência por (-) “0”.

Informações dos padrões de metilação foram avaliadas em quatro estados (*EcoRI-HpaII/MspI*): ausência de metilação (1/1), metilação completa ou interna de bases de citosinas (0/1), hemi-metilação ou externa de citosinas (1/0), não informativo ou ausência de sítio de restrição (0/0) (Tabela 2).

Tabela 2. Padrão de metilação de DNA genômico com as enzimas *HpaII* e *MspI*

Tipos	Sítios alvos		Sensibilidade de enzima		Padrão de Bandas	
			<i>HpaII</i>	<i>MspI</i>	<i>EcoRI-HpaII</i>	<i>EcoRI-MspI</i>
I	CCGG	5mCCGG	Sim	Sim	+	+
	GGCC	GGCC				
II	C5mCCGG		Sim	Não	+	-
	GGCC					
III	5mCCGG		Não	Sim	-	+
	GGCC5m					
IV	5mCCGG		Não	Não	-	-
	GGCC5m					
V	5mCTGG	CNNG	Não	Não	-	-
	GACC	GNNC				

“+” representa presença de bandas em gel de eletroforese; “-” ausência de bandas em gel de eletrofores; 5mC representa sítio de metilação em bases de citosinas. N indica uma substituição por uma das quatro bases nucleotídicas (A,T,C,G). Modificado de Zhou *et al.* (2013).

4.11 Isolamento de regiões metiladas

Bandas com informações do estado de metilação de DNA foram isoladas dos géis de poliacrilamida 10%, com o auxílio de espátula, e colocados os fragmentos em tubos de eppendorf, com etiquetas das amostras (tecidos reprodutivos) e tratamentos (meio de cultura).

Os parâmetros para a seleção dos fragmentos foram os seguintes: regiões com extensão molecular de 200 a 700pb; bandas nítidas e com estado de metilação interna de citosinas (*EcoRI-MspI*); bandas nítidas e com estado de metilação externa de citosinas (*EcoRI-HpaII*); o padrão das bandas consiste de amplificação em somente um dos ensaios (1/0; 0/1); as regiões foram obtidas

para os tecidos reprodutivos de fêmeas e outras de machos, em moscas criadas em meio padrão, *C. hildmmanianus*, e *P. machrisis*.

Os fragmentos de DNA (isolados dos géis) foram incubados com 100 µL de tampão de eluição (TE) e aquecidos a 95°C por 20 minutos (Sanguinetti *et al.* 1994), seguido por centrifugação de 5.000 rpm por 2 minutos. Essas amostras foram enriquecidas com reamplificações em PCR, com os componentes: 7,0 µL de amostra de DNA, 2,0 µL Tampão 10X Buffer, 1,6 µL MgCl⁺² (25 mM), 0,5 µL dNTPs (2.5 mM), 1U taq DNA polimerase, com os mesmos *primers* da amplificação seletiva, e completou o volume de 20 µL com H₂O Milli-Q. O ciclo da programação da PCR consistiu em: 94°C por 1 minuto, 23 ciclos com 94°C por 30 segundos, 56°C por 30 segundos, 72°C por 1 minuto, e extensão final com 72°C por 10 minutos.

Os produtos de PCR foram verificados com gel de agarose 1% em eletroforese. Por fim, as amostras de produto de PCR foram purificadas com os reagentes do *kit Gel Band and DNA Purification* (GE Healthcare, Little Chalfont, Buckinghamshire, Reino Unido), de acordo com as instruções do fabricante. As amostras foram preservadas em freezer -20°C.

4.12 Razões Relativas de Metilação (RRM)

Foram avaliadas as variações nos estados de metilação no DNA, em pupas, e tecidos reprodutivos de fêmeas e outro de machos, com os indivíduos criados nos ambientes de meio padrão, *C. hildmmanianus* e *P. machrisis*. As razões foram determinadas para metilação interna (RRM-I) e externa de citosinas (RRM-E), e metilação total (RRM-T), com o uso da fórmula:

RRM-I

$$\text{Razão } RRM-I = \frac{\text{N.º bandas com metilação interna de citosinas}}{\text{N.º total de bandas}}$$

em que N.º bandas com metilação interna de citosinas corresponde ao número total de bandas com os marcadores seletivos – MSAP para *EcoRI/MspI*, e N.º total de bandas corresponde a todas as regiões obtidas com os mesmo marcadores seletivos nos ensaios com *EcoRI-HpaII/MspI*.

RRM-E

$$\text{Razão } RRM-E = \frac{\text{N.º bandas hemi-metilados de citosinas}}{\text{N.º total de bandas}}$$

em que N.º bandas com metilação externa de bases de citosinas corresponde ao número total de bandas para *EcoRI/ HpaII*, ausentes em *EcoRI-MspI*, e com o N.º total de bandas que corresponde a todas as bandas obtidas com *EcoRI-HpaII/MspI*.

RRM-T

$$\text{Razão RRM-T} = \text{razão RRM-I} + \text{razão RRM-E}$$

em que *RRM-I* corresponde ao valor da razão da metilação interna de bases de citosinas, com a soma de *RRM-E* que significa o valor da razão da metilação externa de citosinas.

4.13 Análise de dados MSAP

Os *loci* foram classificados como suscetíveis à metilação (*Methylation -susceptible loci* - MSL) e não metilados (*Non-methylated loci* - NML), com o uso do pacote MSAP (Pérez-Figueroa, 2013), no programa R 2.14.2. Dados de MSL compreendem dois padrões (-/+, +/-) e NML engloba o padrão de ausência de metilação (-/-, +/+). Variação no estado de metilação de DNA foi estimada no programa MSAP, e inferência sobre o índice médio de Diversidade Shannon para MSL e NML. As variações no estado de metilação foram consideradas significativas estatisticamente com $p < 0.05$.

4.14 Teste de análise de variância ANOVA

As variações no padrão de metilação no MSAP foram estimadas por meio do teste de Análise de Variância (ANOVA), com o uso das variáveis no estado de metilação interna de citosinas, e metilação externas de citosinas de DNA, nas amostras de pupas, e tecidos reprodutivos de fêmeas e outra de machos. Essas comparações foram realizadas com indivíduos criados em meio padrão, *C. hildmmanianus*, e *P. machrisis*. O teste de múltiplas comparações de Tukey foi realizado para verificar quais comparações foram significativas. Os resultados das variações foram considerados com significância estatística com $P < 0,05$.

$$D_{ms} = q \cdot \sqrt{QMR}$$

N.º de repetições

em que q corresponde ao valor da amplitude da probabilidade, QMR indica o valor soma de quadrados, $N.º$ repetições representa o número de réplicas de tratamento.

O coeficiente de variação (CV) (Eberhard, 1998) foi inferido sobre os dados do padrão de bandas da MSAP obtidos nos indivíduos que usaram os recursos de meio padrão, os cactos *C. hildmmanianus* e *P. machrisis*, com o uso da fórmula estatística:

$$CV(\%) = \frac{\text{Desvio padrão} \times 100}{\text{Média}}$$

Os valores de CV foram classificados em baixa dispersão ou dados homogêneos (CV < 15%), dispersão média (CV= 15 a 30%), e alta dispersão ou dados heterogêneos (CV> 30%).

4.15 Sequenciamento de DNA

Os produtos de PCR, com as regiões alvos de metilação, que foram purificados seguiram para o processo de sequenciamento de DNA. As reações de sequenciamento de DNA foram preparadas de acordo com as instruções do *kit BigDye Terminator Cycle Sequencing Ready Reaction* (Perkin-Elmer, Foster City, CA, EUA). Para cada reação de sequenciamento, foram usados os seguintes componentes: 6,5 µL de DNA purificado; 1,0 µL de Big Dye; 1,5 µL de Tampão Big Dye [5,0X]; 1,0 µL *primer* [10 picomol/ µL]. O volume final da reação de sequenciamento foi de 10 µL. Isso foi realizado quando a amostra continha valores próximos a 50ng de DNA. Os “*primers*” usados no sequenciamento direto foram os mesmos da amplificação seletiva.

Os ciclos de programação de PCR para sequenciamento consistiram em: 96°C por 2 segundos, 25 ciclos de 96°C por 10 segundos, 56°C por 5 segundos e 60°C por 4 minutos. Na precipitação dos produtos de PCR, aplicou-se 80 µL de isopropanol 65% e incubados por 15 minutos, em temperatura ambiente e com ausência de luminosidade, e seguido por centrifugação a 4000 rpm por 45 minutos. Os sobrenadantes foram removidos e realizou-se a aplicação de 200 µL de etanol 70%, com temperatura de -20°C, com duas repetições. O sequenciamento do DNA foi realizado em um sequenciador ABI Prism 377 (Applied Biosystems, Foster City, CA, EUA).

4.16 Análises de Bioinformática

A qualidade das sequências foi avaliada usando a análise de Phred (Ewing & Green, 1998), com o auxílio do programa Codoncode Aligner¹, os valores de Phred>15 foram considerados satisfatórios para as análises das bases nucleotídicas nas sequências de DNA. As sequências foram comparadas, corrigidas e editadas com o auxílio dos programas Chromas Lite 2.0. Os alinhamentos das sequências foram realizadas por meio do ClustalW Multiple Alignment 1.8 (Thompson *et al.* 1994), implementado no programa BioEdit 7.1.9 (Hall, 1999), e obtidas as sequências *consensos*.

Foram realizadas buscas de similaridades por meio de comparações entre as sequências de consultas e sequências disponíveis nos bancos de dados genômicos, com o auxílio do programa BLASTn, disponível em <https://www.ncbi.nlm.nih.gov/>, e FlyBase (disponível em <https://www.flybase.org/>), NCBI (disponível em: <http://ncbi.nlm.nih.gov/>), e FlyBase (disponível

¹ Disponível em <https://www.codoncode.com/aligner/>

em <http://flybase.org/>). O genoma de *Drosophila melanogaster* (BDGP6.32) foi usado como sequências de referências no banco do Ensembl. O alinhamento local entre a sequência de consulta e as sequências disponíveis resulta em valores de pontuações, que são associados a um valor de *expected-value* (*e-value*). Os *scores* do *e-value* indicam os números de alinhamentos locais que seriam esperados com pontuações iguais ou maiores que o encontrado em buscas aleatórias nos bancos de dados genômicos.

A organização da estrutura dos dados das sequências de DNA, com origem em moscas fêmeas e machos em diferentes ambientes, foi analisada por meio de árvore filogenética, com o modelo de vizinhos próximos – Neighbour-Joining. Essa organização das sequências foi recuperada com 1000 repetições de *bootstrep*, com uma árvore *consenso*.

Identificações dos genes foram obtidas no banco do FlyBase. Os códigos de cada região alvo de metilação em *D. gouveai* foram obtidos com base nos ortólogos em *D. melanogaster*.

As identificações das regiões de exons e íntrons nas sequências de DNA foram obtidas por meio de buscas em bancos de dados de transcrição e suas respectivas proteínas no programa BLASTx². Em seguida, as sequências de nucleotídeos correspondentes ao local de sítio de início da transcrição (TSS) foram determinadas com o auxílio da ferramenta ExPasy, disponível em <https://web.expasy.org/translate/>. Por fim, as regiões promotoras foram preditas com o auxílio do programa *Neural Network Promoter Prediction* (NNPP, versão 2.2), disponível em http://www.fruitfly.org/seq_tools/promoter.html (Reese, 2001). O *score* mínimo para estimar as regiões promotoras foi *score* > 0,8.

4.17 Análises funcionais de Ontologia dos Genes (GO)

Todas as regiões alvos de metilação no DNA, em *Drosophila gouveai* com indivíduos criados em diferentes ambientes, foram identificadas as posições genômicas destas regiões e os genes. O conjunto de genes foi avaliado com análises funcionais, com o auxílio do banco de dados do DAVID V.6.8 (*Database for Annotation, Visualization and Integrated Discovery*) (Dennis *et al.* 2003; Huang *et al.* 2007). Este método foi usado para identificar os processos biológicos e as suas relações fisiológicas (componentes celulares e funções moleculares), com a integração de dados de diversos bancos com uso funcional. O banco de dados do FlyBase, com *Drosophila melanogaster*, foi usado como *background* nas análises funcionais. Foram usadas todas as categorias dos bancos de dados interno de *Gene Ontology* (GO), tais como: a integração das informações biológicas de genes/proteínas com suas categorias de anotações funcionais; enriquecimento e relações entre genes e termos GO; classificação funcional dos genes; agrupamento com anotação funcional; proteína-

² Disponível em <http://ncbi.nlm.nih.gov/BLAST>

proteína interação. Por fim, a visualização da dinâmica de genes e vias metabólicas que foram mais relevantes com o uso do KEGG. Os dados de cada categoria foram considerados estatisticamente significativos com $p \leq 0,05$, e correções de $FDR \leq 0,05$.

Resultados

5. Resultados

5.1 Análises dos perfis de metilação

Analisaram-se quatro padrões de metilação nos perfis moleculares obtidos por meio da técnica *Methylation Sensitive Amplification Polymorphism* (MSAP) (Figura 8). Os dados dos perfis foram usados para elaborar as matrizes binárias, sendo a presença de marcas (+) consideradas “1” e ausência de marcas (-) como “0” (Apêndices B, C, e D).

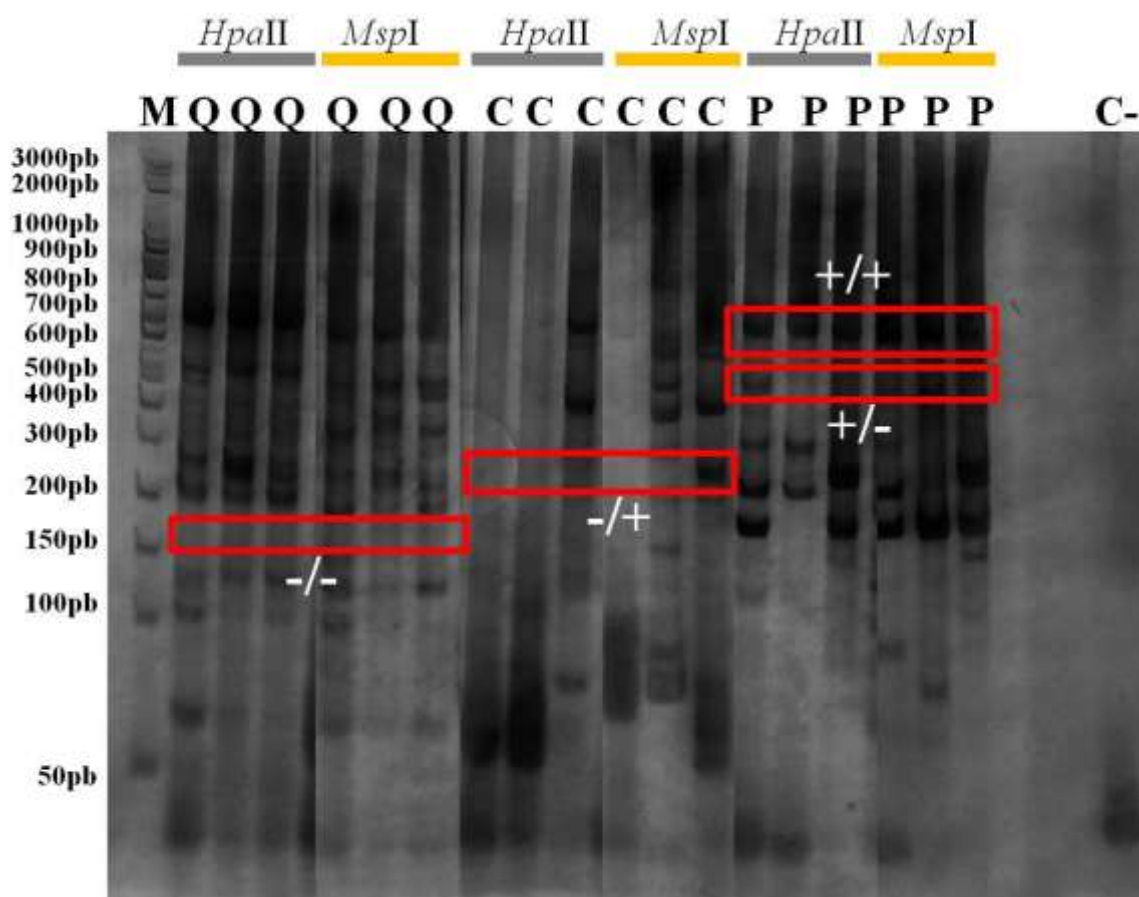


Figura 8. Perfil de metilação obtido por meio da técnica MSAP com DNA de pupas, visualização em gel de poliacrilamida (10%), com a indicação de quatro padrões analisados: ausência de metilação (+/+), metilação completa ou semi metilação de citosinas internas (-/+), e não informativa (-/-). M representa o marcador molecular (50pb - Promega®). As amostras estão representadas por Q (meio padrão de laboratório), C (meio de *Cereus hidmannianus*), e P (meio de *Pilosocereus machrisis*). Os ensaios enzimáticos estão representados por *HpaII* e *MspI*, que foram realizados em conjunto com *EcoRI*. C- representa o controle negativo da reação de MSAP.

5.2 Perfis Moleculares do DNA genômico

Avaliaram-se 12 combinações de *primers* seletivos nos ensaios de MSAP, que compreenderam variações na escala molecular entre 100pb a 1000pb, nas amostras de DNA genômico em *D. gouveai*. Dados dos padrões moleculares nos ensaios de clivagem, *EcoRI-HpaII/MspI*, e ampliações seletivas evidenciaram conjuntos totais de regiões de fragmentos de DNA em populações de células em pupas (n=324 bandas) (Tabela 3), em tecidos reprodutivos feminino (n=308 bandas) (Tabela 4), e outras em tecidos reprodutivos masculinos (201 bandas) (Tabela 5), quando os organismos foram criados em meio padrão de laboratório (quirera), e nos cactos *C. hidmannianus*, e *P. machrisis* (Apêndices F, G, e H). Todos os dados genômicos em tecidos reprodutivos, em moscas fêmeas e outras em machos, foram obtidos com indivíduos virgens.

Tabela 3. Perfil molecular de DNA genômico em amostras de pupas, com a segunda geração consecutiva criada em meio padrão de laboratório (Padrão L), meio dos cactos *C. hidmannianus*, e e outras com *P. machrisis*

	Padrão L	<i>C. hidmannianus</i>	<i>P. machrisis</i>
Metilação interna 5mC	25	39	43
Metilação externa 5mC	15	20	26
Regiões não metiladas	65	35	56
Total	105 <i>loci</i>	94 <i>loci</i>	125 <i>loci</i>

Tabela 4. Perfil molecular de DNA genômico em amostras de tecidos reprodutivos, com a segunda geração consecutiva de moscas fêmeas, que se desenvolveram no meio padrão de laboratório (Padrão L), nos meios de cactos *C. hidmannianus*, e e outras com *P. machrisis*.

	Padrão L	<i>C. hidmannianus</i>	<i>P. machrisis</i>
Metilação interna 5mC	47	38	31
Metilação externa 5mC	36	29	21
Regiões não metiladas	64	28	15
Total	146 <i>loci</i>	95 <i>loci</i>	67 <i>loci</i>

Tabela 5. Perfil molecular de DNA genômico em amostras de tecidos reprodutivos, com a segunda geração consecutiva de moscas machos, que se desenvolveram no meio padrão de laboratório (Padrão L), nos meios de cactos *C. hidmannianus*, e outros com *P. machrisis*.

	Padrão L	<i>C. hidmannianus</i>	<i>P. machrisis</i>
Metilação interna 5mC	29	42	21
Metilação externa 5mC	27	12	22
Regiões não metiladas	29	13	06
Total	85 loci	67 loci	49 loci

5.3 Razões Relativas de Metilação

Os padrões de metilação obtidos nos perfis moleculares compreendem os estados de metilação completa ou semi-metilação de citosinas internas e a semi-metilação externa de citosinas, que são reconhecidas por metilação total relativa. Esses dados de metilação representam parte da porção do epigenoma que está metilada, em populações de células, que é independente das citosinas na sequência 5'-CCGG-3', devido ao fato de englobar dados de metilação interna e externa desse alvo.

Variações nas razões relativas de metilação foram encontradas no DNA genômico de indivíduos de pupas e órgãos reprodutivos de moscas adultas, quando os organismos se desenvolveram no meio padrão de laboratório, e em tecidos dos cactos *C. hildmannianus*, e *P. machrisii*, em *D. gouveai* (Figuras 9, 10, e 11). O DNA genômico contém alvos de metilação total relativa com maior representação em indivíduos que usaram os meios de *C. hildmannianus* (razão=0.62) do que no padrão de laboratório (razão 0.38), na fase de pupa. O estado de metilação interna de bases de citosinas é diferentes entre os indivíduos que usaram *C. hildmannianus* (razão=0.41) e *P. machrisii* (razão=0.34), embora diferentes do meio padrão de laboratório (razão=0.24). Os resultados de metilação externa ou hemi-metilação de citosina foram semelhantes entre os indivíduos que usaram os ambientes de *C. hildmannianus* (razão=0.21) e *P. machrisii* (razão=0.21), porém com maior proporção do que no meio padrão (razão = 0.14).

Os estados de metilação no DNA genômico apresentaram mudanças epigenéticas com maior representatividade em indivíduos que se desenvolveram em *P. machrisii* e *C. hildmannianus* do que em meio padrão, nos dados de tecidos reprodutivos de moscas adultas fêmeas e também de machos,

em *D. gouveai*. Informações sobre o estado de metilação interna de citosinas indicam maior quantidade de bases metiladas no DNA quando os organismos se desenvolveram e emergiram do tecido de *P. machrisii* (razão=0.46) em relação a *C. hildmannianus* (razão=0.39) e no meio padrão (razão=0.32). Por outro lado, os dados do estado de metilação externa de citosinas foram semelhantes entre as moscas fêmeas que usaram tecidos de *P. machrisii* (razão=0.31) e *C. hildmannianus* (razão=0.32), embora distinto quando usaram o recurso do meio padrão (razão=0.24).

Mudanças nos estados de metilação foram mais representativas no DNA em organismos que usaram os recursos de *P. machrisii* (razão=0.88) do que no meio padrão (razão=0.66), nas moscas adultas, com origem masculina, em *D. gouveai*. O estado de metilação internas de citosinas foi mais representativo em indivíduos que emergiram de tecidos de *C. hildmannianus* (razão=0.63) do que *P. machrisii* (razão=0.43), embora com diferenças em relação ao meio padrão (razão=0.34), em moscas machos (Figura 10). Em contraste, o estado de metilação externa de bases citosinas contém diferenças quando os organismos se desenvolveram e emergiram de tecidos de *P. machrisii* (razão=0.45), meio padrão de laboratório (razão=0.32), e *C. hildmannianus* (razão=0.18).

Os resultados em conjuntos indicam que a dinâmica nas razões de metilação relativa de bases de citosinas no DNA gênomico foram mais perceptíveis em organismos que se desenvolveram em tecidos de cactos do que em meio padrão, nas linhagens feminina e masculina do que em pupas (Figuras 9, 10, e 11), em *D. gouveai*.

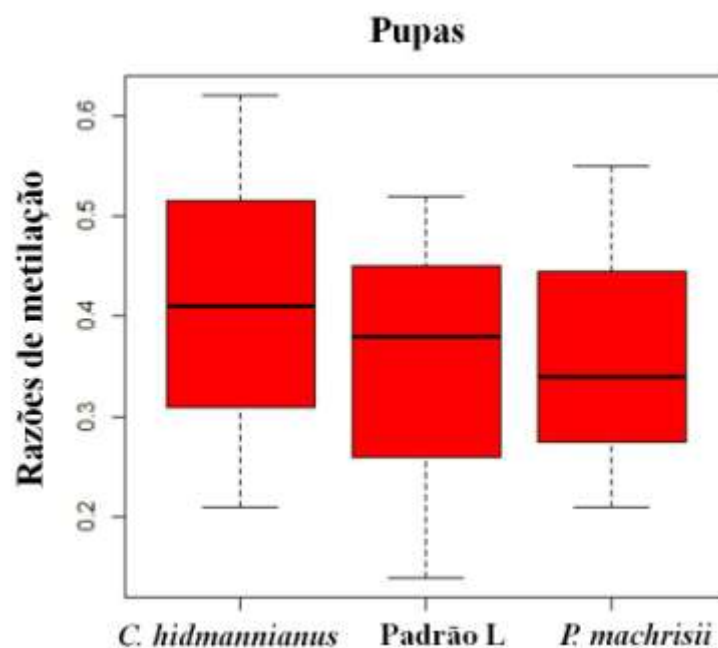


Figura 9. Gráfico boxplot com razões relativas de metilação no DNA genômico no organismo (pupas) que se desenvolveram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*. Esses dados foram obtidos por meio de perfis moleculares gerados com os

estados de metilação interna e externa de bases de citosinas representado ($HpaII-/MspI+$; $HpaII+/MspI-$), e ambos no total relativo de metilação, em cor vermelha.

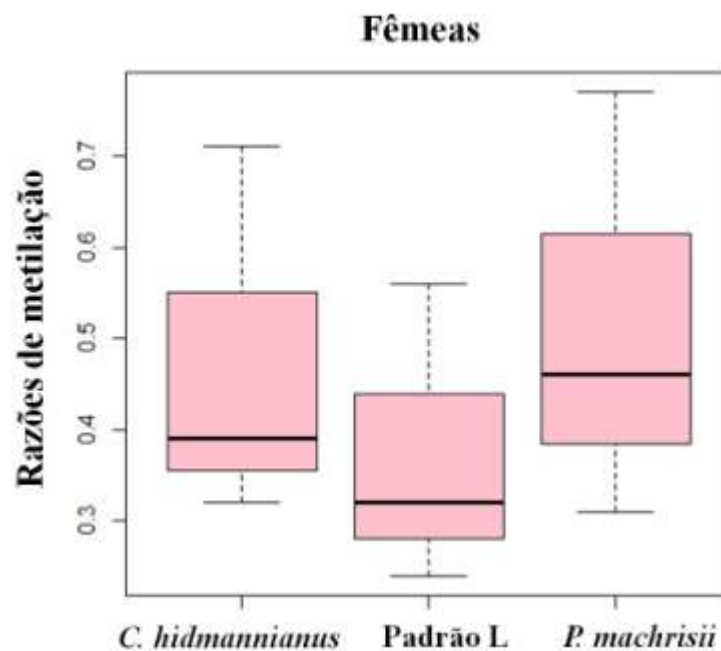


Figura 10. Gráfico boxplot com razões relativas de metilação no DNA genômico no tecido reprodutivo (moscas adultas, fêmeas) que se desenvolveram e emergiram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*. Esses dados foram obtidos por meio de perfis moleculares gerados com os estados de metilação interna e externa de bases de citosinas representado ($HpaII-/MspI+$; $HpaII+/MspI-$), e ambos no total relativo de metilação, em cor rosa.

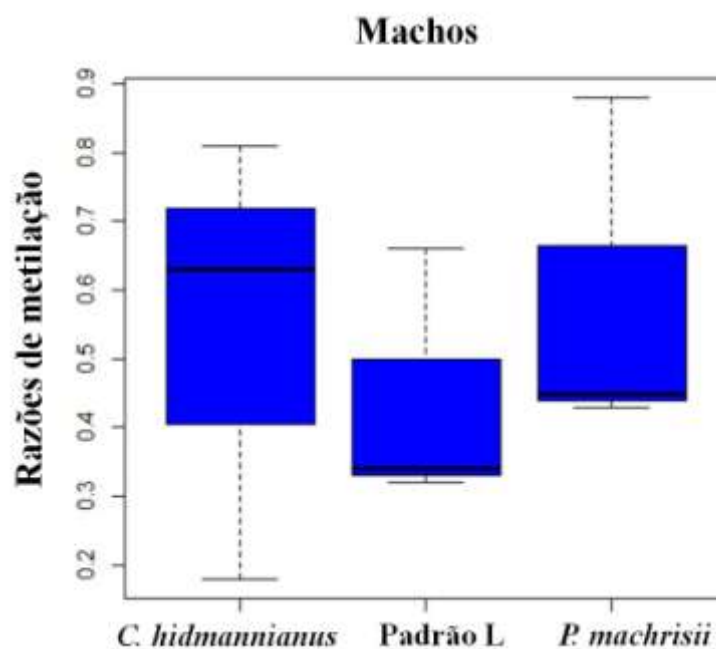


Figura 11. Gráfico boxplot com razões relativas de metilação no DNA genômico no tecido reprodutivo (moscas adultas, machos) que se desenvolveram e emergiram nos ambientes com meio padrão de laboratório (Padrão L), tecidos de cactos *C. hildmannianus*, e *P. machrisii*, com o segundo desenvolvimento

consecutivo. Esses dados foram obtidos por meio de perfis moleculares gerados com os estados de metilação interna e externa de bases de citosinas representado (*HpaII-/MspI+*; *HpaII+/MspI-*), e ambos no total relativo de metilação, em cor azul.

5.4 Perfis de Metilação – MSAP

Os perfis de metilação compreenderam presença de bandas *EcoRI/HpaII* e ausência de bandas em *EcoRI/MspI*, e ausência de bandas *EcoRI/HpaII* e presença de bandas em *EcoRI/MspI*. Esses dados das bandas correspondem à metilação interna e externa de bases de citosinas, com múltiplos *loci* sensíveis à metilação (MSL) na varredura de todo o DNA genômico.

Identificaram-se mudanças em múltiplos *loci* sensíveis à metilação no DNA dos organismos que usaram no processo de desenvolvimento o meio padrão de laboratório (MSL=40 *loci*), meios de cactos *C. hildmannianus* (MSL=59 *loci*), e *P. machrisii* (MSL=69 *loci*), na fase de pupa. Essas regiões moleculares apresentaram variações mais representativas nas escalas moleculares entre 150 a 600pb. Nas análises dos perfis, foram encontrados 16 *loci* sensíveis metilação genômica, nos amplicons de PCR com a combinação E(ACC)/HM(TTC), quando os organismos usaram os tecidos vegetais de *P. machrisii* (Tabela 6). Já com as combinações E(AAC)/HM(TAC) e E(ACC)/HM(TTC), foram amplicados 12 *loci* sensíveis metilação, cada combinação de *primers*, no DNA dos indivíduos que se desenvolveram em meio de *C. hildmannianus*. Em contraste, foram encontrados 9 *loci* (MSL) com a combinação E(ACG)/HM(TAC) para os organismos que se desenvolveram no meio padrão.

A metilação interna e externa de bases citosinas apresenta diferenças moleculares, embora não apresenta significado estatístico com $P < 0.05$ em relação ao ambiente de desenvolvimento, no meio padrão de laboratório ($P=0.05$ e $P=0.08$), e nos meios de cactos *C. hildmannianus* ($P=0.06$ e $P=0.12$), e *P. machrisii* ($P=0.08$ e $P=0.12$), em amostras de pupas, com análise no programa MSAP.

Variações nos perfis de metilação foram encontradas no DNA, em moscas fêmeas, que se desenvolveram no ambiente com meio padrão (MSL=82), meio com tecidos de *C. hildmannianus* (MSL=65), e *P. machrisii* (MSL=52) (Tabela 7). Essas regiões foram mais abundantes na escala molecular entre 150 a 800pb. Diferenças foram encontradas com a combinação E(ACC)/HM(TTC) no DNA nessas moscas fêmeas que se desenvolveram e emergiram no ambiente com o meio padrão (MSL=18), tecidos vegetais dos cactos *C. hildmannianus* (MSL=12), e *P. machrisii* (MSL=6). Por outro lado, dois *loci* com informação do estado de metilação foram encontrados com E(ACG)/HM(TAC).

Os resultados das análises dos perfis de metilação interna e externa de bases de citosinas não apresentaram diferenças estatísticas com relação ao recurso do meio padrão de laboratório ($P=0.09$ e $P=0.14$), nos cactos *C. hildmannianus* ($P=0.07$ e $P=0.08$), e *P. machrisii* ($P=0.06$ e $P=0.07$), considerando significativas $P < 0.05$.

A metilação no DNA genômico apresenta maior quantidade de múltiplos *loci* MSL com as combinações E(ACT)/HM(TTG) e E(AGG)/HM(TGA), e com baixa representação de regiões sensíveis à metilação com o alvo E(AGC)/HM(TCC), em moscas machos (Tabela 8). O conjunto de dados genômicos indica maior representação de *loci* sensíveis à metilação na escala molecular entre 200 a 700pb, com relação às ampliações próximas de 1000pb em PCR. Essas regiões do genoma apresentaram diferenças em quantidade de alvos de metilação, com a combinação E(ACT)/HM(TTG), em organismos que se desenvolveram no ambiente com meio padrão (MSL=15). Variações foram encontradas em número de regiões alvos de metilação no DNA, com a combinação E(AGG)/HM(TGA), em indivíduos que usaram os compostos de tecidos de *C. hildmannianus* (MSL=11), e *P. machrisii* (MSL=7). Em contraste, encontramos três e seis *loci* alvos de metilação com a combinação E(AGC)/HM(TCC), em machos adultos.

Dentre os *loci* MSL, informações sobre o estado de metilação externa de citosinas foram significativas em moscas machos que se desenvolveram nos tecidos de *C. hildmannianus* ($P>0.03$) em relação ao cacto *P. machrisii* ($P=0.06$) e no meio padrão de laboratório ($P=0.11$), no programa MSAP. Já o estado de metilação interna de citosinas não foi significativo no DNA de indivíduos em usaram o meio padrão ($P=0.11$), *C. hildmannianus* ($P=0.14$), e *P. machrisii* ($P=0.07$).

Comparações entre os dados MSL e NSL sugerem que os componentes genéticos e epigenéticos são importantes quando os organismos usam os recursos de diferentes ambientes nutricionais. Tendo em vista que encontramos significado estatístico ($p<0.05$) nas variações de *loci* não sensíveis à metilação NML (bandas em ambos os ensaios *EcoRI-HpaII/MspI*; +/+; ausência de bandas ambas ensaios *EcoRI-HpaII/MspI*; -/-) em todas as fases, nos diferentes ambientes. Isto pode estar relacionado à diferença de regiões alvos de metilação no DNA genômico, quando os organismos se desenvolveram em diferentes ambientes, na espécie cactofílica *D. gouveai*.

Tabela 6. Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de bases (pb), em pupas de *D. gouveai*, com desenvolvimento em meio padrão (Q), meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).

Ambientes	Combinação de <i>primers</i> seletivos	Nº <i>Loci</i>	MSL	NML	Extensão dos fragmentos
Q	E(AAC)/HM(TTC)	12	6	6	150-800
C		17	7	10	
P		22	16	6	
Q	E(ACT)/HM(TTG)	22	5	17	100-700
C		17	8	9	
P		13	1	12	
Q	E(AGC)/HM(TCC)	17	5	12	100-700
C		7	5	2	
P		19	14	5	
Q	E(AGG)/HM(TGA)	12	6	6	100-500
C		4	2	2	
P		7	3	4	
Q	E(AAC)/HM(TAC)	7	7	0	100-800
C		12	12	0	
P		13	13	0	
Q	E(AAC)/HM(TAG)	10	0	10	150-700
C		11	7	4	
P		12	4	8	
Q	E(ACC)/HM(TTC)	2	2	0	150-900
C		18	12	6	
P		19	15	4	
Q	E(ACG)/HM(TAC)	23	9	14	100-750
C		8	6	2	
P		20	3	17	
Número de <i>loci</i> (Desvio padrão)	Q	105 (±7,21)	40 (±2,82)	65 (±6,24)	
	C	94 (±5,23)	59 (±3,37)	35 (±3,62)	
	P	125 (±5,12)	69 (±6,39)	56 (±5,01)	
Diversidade de Shannon's	MSL	I=0,48 (SD:0.12)			
	NML	I=0,37 (SD:0.09)			
Teste de Wilcoxon		W=71	P=0,046		

Tabela 7. Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de base (pb), em moscas fêmeas de *D. gouveai*, com desenvolvimento em meio padrão (Q), meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).

Ambientes	Combinação de <i>primers</i> seletivos	Nº <i>Loci</i>	MSL	NML	Extensão dos fragmentos
Q	E(AAC)/HM(TTC)	19	12	7	100-1000
C		5	5	0	
P		2	2	0	
Q	E(ACC)/HM(TAC)	15	4	11	100-1000
C		8	4	2	
P		6	6	0	
Q	E(ACG)/HM(TAG)	10	2	8	100-700
C		11	7	4	
P		3	3	0	
Q	E(ACT)/HM(TTG)	17	15	2	100-900
C		8	4	4	
P		7	7	0	
Q	E(AGG)/HM(TGA)	20	3	17	100-1000
C		14	9	5	
P		7	7	0	
Q	E(AAC)/HM(TAC)	17	13	4	100-600
C		4	1	3	
P		4	4	0	
Q	E(AAC)/HM(TAG)	8	4	4	100-800
C		5	3	2	
P		0	0	0	
Q	E(ACC)/HM(TTC)	22	18	4	100-1000
C		16	12	4	
P		13	6	7	
Q	E(ACG)/HM(TAC)	4	2	2	100-800
C		0	0	0	
P		4	2	2	
Q	E(ACG)/HM(TTG)	7	7	0	100-1000
C		11	11	0	
P		15	11	4	
Q	E(ACT)/HM(TCC)	7	2	5	100-1000
C		13	9	4	
P		6	4	2	
Número de <i>loci</i> (Desvio Padrão - DS)	Q	146 (±6,23)	82 (±5,93)	64 (±4,81)	
	C	95 (±4,86)	65 (±3,98)	28 (±1,86)	
	P	67 (±4,48)	52 (±3,06)	15 (±2,29)	
Diversidade de Shannon's	MSL	I=0,43 (DS: 0.11)			
	NML	I=0,23 (DS: 0.07)			
Teste de Wilcoxon		W=137	P=0,0001		

Tabela 8. Número de *loci* que foram detectados com a combinação de *primers* seletivos de MSAP, com número total (N.º *loci*) e *loci* sensíveis à metilação (MSL), não sensível a metilação (NML), e extensão dos fragmentos em pares de base (pb), em moscas machos de *D. gouveai*, com desenvolvimento em meio padrão (Q), nos meios de cactos *C. hildmannianus* (C), e *P. machrisii* (P).

Ambientes	Combinação de primers seletivos	Nº Loci	MSL	NML	Extensão dos fragmentos
Q	E(AAC)/HM(TTC)	4	1	2	150-900
C		9	7	2	
P		5	3	2	
Q	E(ACC)/HM(TAC)	7	3	4	100-1000
C		2	2	0	
P		6	6	0	
Q	E(ACT)/HM(TTG)	17	15	2	100-1000
C		5	5	0	
P		8	6	2	
Q	E(AGC)/HM(TCC)	5	3	2	100-900
C		8	6	2	
P		3	3	0	
Q	E(AGG)/HM(TGA)	19	11	8	100-900
C		15	11	4	
P		9	7	2	
Q	E(AAC)/HM(TAC)	5	3	2	100-900
C		12	10	2	
P		4	4	0	
Q	E(ACC)/HM(TTC)	13	9	4	100-900
C		5	5	0	
P		7	6	7	
Q	E(ACG)/HM(TTG)	15	10	5	100-700
C		11	8	3	
P		7	7	0	
Número de <i>loci</i> (Desvio padrão - DS)	Q	85 (±5,03)	55 (±5,03)	29 (±2,13)	
	C	67 (±4,27)	54 (±2,91)	13 (±1,50)	
	P	49 (±2,03)	42 (±1,67)	13 (±2,39)	
Diversidade de Shannon's	MSL	I=0,43 (DS:0.14)			
	NML	I=0,29 (DS:0.06)			
Teste de Wilcoxon		W=20	P=0.106		

5.5 Análises de Variações – ANOVA

Os dados de metilação no DNA apresentam maior variação média em amostras obtidas em meios de cacto (*C. hildmannianus*; e *P. machrisii*) do que no meio padrão, em amostras de pupa (Figura 11). Todavia essas alterações nos perfis de metilação de DNA não apresentaram diferenças estatísticas no teste *One-way* ANOVA ($F=1.1453$; $P\text{-valor}=0.3444$) (Tabela 9).

Adicionalmente, uma variação média de regiões metiladas no DNA, em moscas fêmeas, foi observada com maior evidência nos organismos que se desenvolveram no meio padrão do que em meios de cactos (*C. hildmannianus*; e *P. machrisii*) (Figura 13). Essa variação no perfil de metilação apresentou significância estatística no teste *One-way* ANOVA ($F=2.6618$; $P\text{-valor}=0.1025$) (Tabela 10).

Em contraste, observa-se uma semelhança na variação de número de regiões metiladas no DNA entre os organismos que se desenvolveram no meio padrão e no cacto *C. hildmannianus* do que em *P. machrisii* (Figura 14), em moscas machos. Entretanto, as mudanças nos perfis de metilação de DNA não apresentaram suporte estatístico no teste *One-way* ANOVA ($F=0.2685$; $P\text{-valor}=0.7681$) (Tabela 11).

Encontrou-se uma variação significativa entre os dados de metilação genômica dos organismos que se desenvolveram no meio padrão de laboratório e no cacto *P. machrisii* ($P\text{-valor ajustado}=0.098$) (Tabela 10), em moscas fêmeas. Todavia, as amostras de pupas e adultos machos apresentaram baixo suporte nas diferenças de metilação entre os organismos que usaram o meio padrão e os tecidos de cactos (*C. hildmannianus* e *P. machrisii*) (Tabelas 9 e 11).

Adicionalmente, o Coeficiente de Variação (CV) indicou dados com dispersão média no estado de metilação de DNA nos organismos de indivíduos que foram criadas em meio padrão (28.33%), no meio de cacto *C. hildmannianus* (23.13%), e outro em *P. machrisii* (20.66%), em pupas de *D. gouveai*. Em contraste, o valor de CV indicou dispersão média no estado de metilação de DNA nos organismos que se desenvolveram no meio de cacto *P. machrisii* (22.22%), e com dispersão homogênea dos dados em meio padrão (11.28%) e no cacto *C. hildmannianus* (12.11%), em moscas fêmeas de *D. gouveai*.

Dados sobre o valor de CV para amostras de moscas machos indicaram padrão com alta dispersão dos dados de metilação no DNA em organismos que se desenvolveram no meio de cacto *P. machrisii* (36.28%), com média dispersão de dados em meio com tecidos de *C. hildmannianus* (29.67%), e com tendência homogênea de dados para os indivíduos em meio padrão (12.26%), em moscas machos de *D. gouveai*.

Em conjunto, os dados indicam diferenças sutis quando os organismos se desenvolveram em diferentes ambientes, meio padrão de laboratório e em tecidos de cactos (Figuras 12, 13, e 14), que podem envolver mudanças em regiões genômicas específicas.

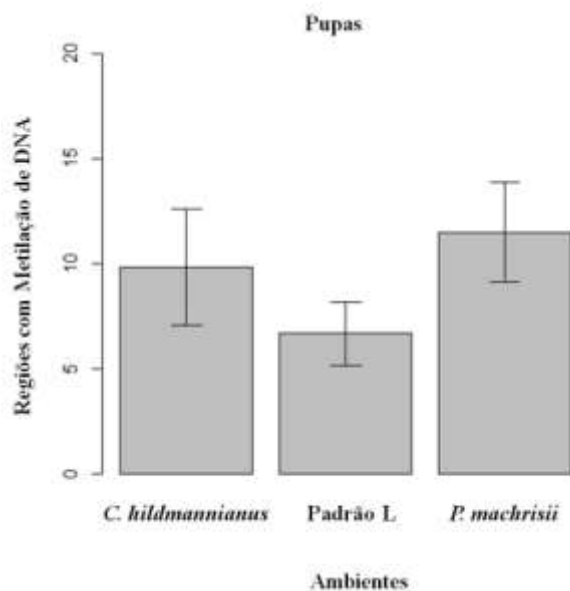


Figura12. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de pupas. Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*.

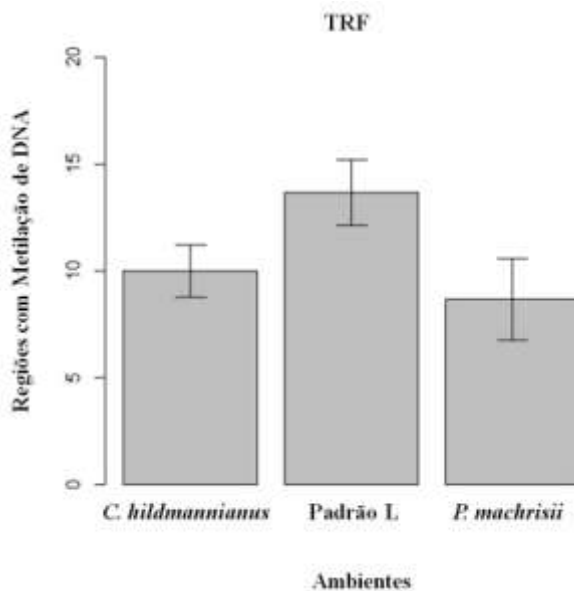


Figura 13. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de tecidos reprodutivos de fêmeas (TRF). Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*

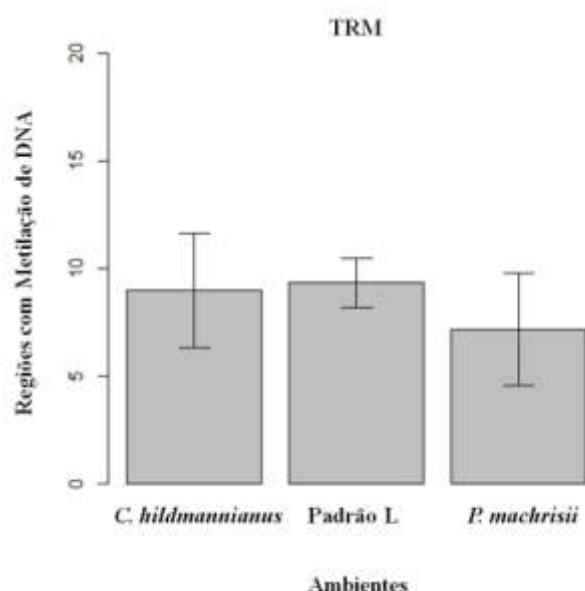


Figura 14. Representação em gráfico de barras com a variação média, e erro padrão, nos dados de metilação relativa (interna e externa de citosinas) em amostras de tecidos reprodutivos de machos (TRM). Essas amostras foram obtidas em meio padrão de laboratório (Padrão L), meio do cacto *C. hildmannianus*, e outro em *P. machrisii*.

Tabela 9. Teste de *One-way ANOVA* e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos (pupas) que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*).

One-way ANOVA					
Fontes de variações	d.f.	MS	SQ	F	P-valor
Dentro do grupo	2	72.33	36.167	1.1453	0.3444
Resíduos	15	473.67	31.578	1.145	0.3444
Total	17	546,000			

Post Hoc Tukey	d.m.	lwr	upr	P-valor ajustados
Padrão-Cereus	3.166667	11.593817	5.260484	0.6025233
Pilosocerus-Cereus	1.666667	6.760484	10.093817	0.8658760
Pilosocerus-Padrão	4.833333	3.593817	13.260484	0.3234175

Meios de cultura – Padrão (meio padrão de quirera), Cereus (tecidos do cacto *C. hildmannianus*), Pilosocerus (*P. machrisii*). Graus de liberdade (d.f.). Quadrados Médios (MQ). Soma dos Quadrados (SQ). Diferenças médias (d.m.). Valores de significâncias estatística representados por $P=0$ ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘ΔΔ’, 0.1 ‘Δ’.

Tabela 10. Teste de *One-way ANOVA* e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*), em moscas fêmeas.

<i>One-way ANOVA</i>					
Fontes de variações	d.f.	MS	SQ	F	P-valor
Dentro do grupo	2	80.444	40.222	2.6618	0.1025 ^Δ
Resíduos	15	226.667	15.111		0.1635 ^Δ
Total	17	307,111			

<i>Post Hoc Tukey</i>	d.m.	lwr	upr	P-valor ajustados
Padrão-Cereus	3.666667	2.162925	9.4962585	0.2625469
Pilosocerus-Cereus	1.333333	7.162925	4.4962585	0.8253483
Pilosocerus-Padrão	5.000000	10.829592	0.8295919	0.0986350 ^Δ

Meios de cultura – Padrão (meio padrão de quirera), *Cereus* (tecidos do cacto *C. hildmannianus*), *Pilosocerus* (*P. machrisii*). Graus de liberdade (d.f.). Quadrados Médios (MQ). Soma dos Quadrados (SQ). Diferenças médias (d.m.). Valores de significâncias estatística representados por $P=0$ ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘ΔΔ’, 0.1 ‘Δ’.

Tabela 11. Teste de *One-way ANOVA* e *Post Hoc Tukey* com as bandas com metilação interna e externa, obtidas na MSAP, em indivíduos que se desenvolveram em diferentes ambientes (meio Padrão, *C. hildmannianus*, e *P. machrisii*), em moscas machos.

One-way ANOVA					
Fontes de variações	d.f.	MS	SQ	F	P-valor
Dentro do grupo	2	16.33	8.1667	0.2685	0.7681
Resíduos	15	456.17	30.4111		0.7681
Total	17	472,500			

Post Hoc Tukey	d.m.	lwr	upr	P-valor ajustados
Padrão-Cereus	0.3333333	7.936678	8.603344	0.9939777
Pilosocerus-Cereus	1.8333333	10.103344	6.436678	0.8348754
Pilosocerus-Padrão	2.1666667	10.436678	6.103344	0.7781727

Meios de cultura – Padrão (meio padrão de quirera), *Cereus* (tecidos do cacto *C. hildmannianus*), *Pilosocerus* (*P. machrisii*). Graus de liberdade (d.f.). Quadrados Médios (MQ). Soma dos Quadrados (SQ). Diferenças médias (d.m.). Valores de significâncias estatística representados por $P=0$ ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘ Δ ’, 0.1 ‘ Δ ’.

5.6 Isolamento das regiões-alvos de metilação

Foram isoladas 15 bandas em amostras de DNA, em tecidos reprodutivos de fêmeas e também em machos, com diferenças no estado de metilação nos ensaios MSAP (Figura 15). Essas regiões apresentam os seguintes parâmetros: estado externa de citosinas (presença de bandas somente em *EcoRI-HpaII*); metilação interna de citosinas (presença de bandas somente em *EcoRI-MspI*); ampliações com padrão de bandas nítidas; extensão molecular entre 200pb a 800pb; mudanças de metilação no DNA em tecidos reprodutivos; diferenças nos estados de metilação nas amostras de DNA com relação aos ambientes (meio padrão de laboratório, tecidos de cactos *C. hildmannianus*, e *P. machrisii*).

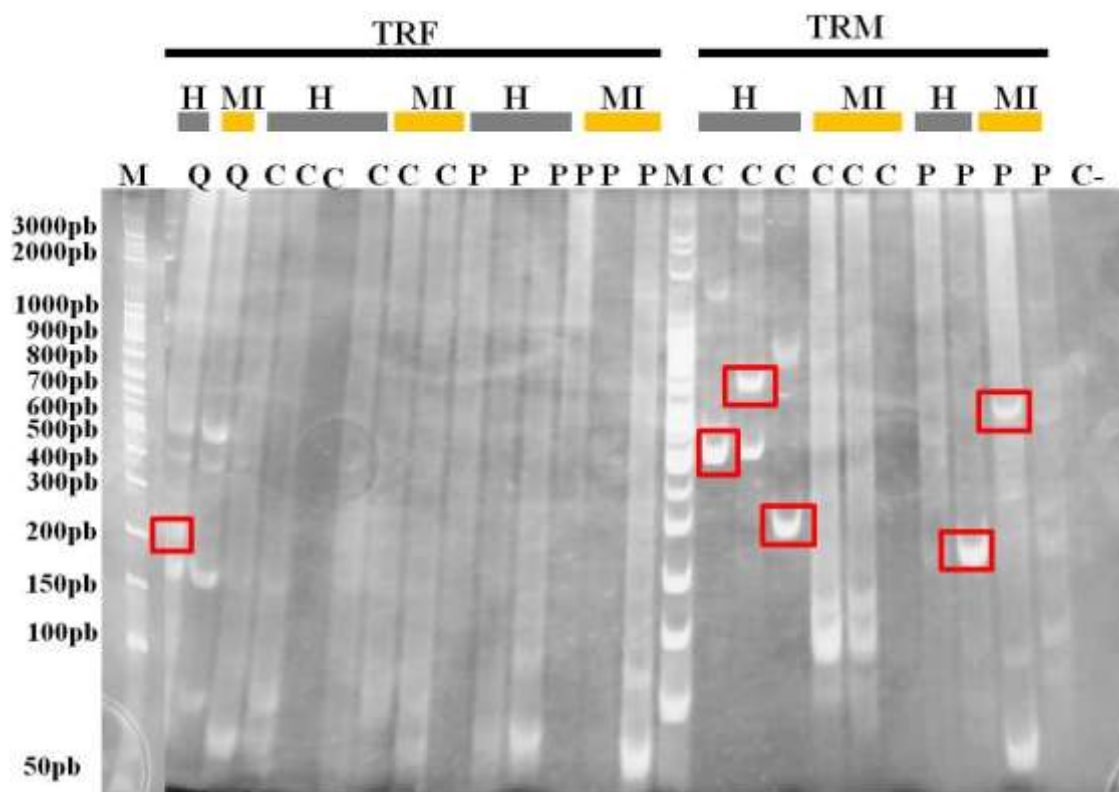


Figura 15. Visualização em gel de poliacrilamida (10%) de perfil de metilação no DNA de amostras de tecidos reprodutivos em moscas fêmeas (TRF) e outras em machos (TRM), com corante de nitrato de prata. As amostras de DNA estão representadas por indivíduos criados no meio padrão de laboratório (Q), meio de cacto *C. hidmannianus* (C), e meio com *P. machrisis* (P). Os ensaios de MSAP foram realizados com as combinações E2-HM2 para TRF, e E2-HM1 para TRM, com uso de enzimas representadas por *EcoRI-HpaII* (H) e *EcoRI-MspI* (MI). Os produtos de PCR foram obtidos por meio da técnica MSAP, isolados os fragmentos de interesse (caixas em vermelho), e com o processo de reamplificação dessas regiões. Essas regiões são alvos de metilação externa de citosinas (bandas somente em *EcoRI-HpaII*), e metilação interna de citosinas (bandas somente em *EcoRI-MspI*), que seguiram para o processo de sequenciamento de DNA. M representa o marcador molecular (50pb - Promega®). O controle negativo da reação de MSAP é indicado na lane (C-).

5.7 Sequenciamento de DNA alvo de metilação

De modo geral, 15 bandas com variação no estado de metilação de citosinas, que compreendem DNA genômico em amostras de moscas fêmeas e outras em machos criados em diferentes ambientes, foram realizados os procedimentos de reamplificação (Apêndice E), purificação, e sequenciamento de DNA. Obtiveram-se 30 sequências de DNA para as regiões alvos de metilação.

O processo de triagem foi realizado com análise de qualidade das sequências de DNA com os parâmetros de Phred>15, consideradas satisfatórias para seguir nas análises moleculares, e os

padrões de picos em chromatograma. Sequências com baixas resoluções nucleotídicas foram retiradas das análises.

Obtiveram-se sete sequências de DNA *consensos* com informações de metilação interna e externa em amostras de tecidos reprodutivos de fêmeas e outras em machos, quando os indivíduos se desenvolveram em ambientes com meio padrão de laboratório, em cacto *C. hildmannianus*, e *P. machrisii* (Tabela 12). Sítios com múltiplos dinucleotídios CpG e CpC foram encontradas nas sequências de DNA.

Uma sequência de DNA foi obtida em fêmeas criadas em meio padrão, com ensaio *EcoRI-MspI*, mas sem similaridade com outras regiões genômicas por meio da análise da ferramenta do BLASTx. Isto pode estar relacionada à dificuldade de buscas com fragmentos menores do que 200pb, nos bancos de dados do GenBank, FlyBase, e no Ensembl.

Uma sequência de DNA para alvo de metilação obtidas em machos adultos, criados em meio padrão, não apresentou qualidade satisfatória, e foi retirada das análises.

Identificaram-se similaridades das sequências alvos de metilação no DNA genômico de *D. gouveai*, moscas fêmeas, com os genes *Piezo*, e *Factor elongation alpha 1 (eEF1alpha1)* com regiões genômicas, principalmente, nas espécies *Drosophila melanogaster* e *Drosophila buzzatii* (Tabela 12).

Gene *Piezo* apresenta extensão de 242pb, com similaridade de 20pb com *D. melanogaster*, que compreende região de éxon, em amostra de moscas fêmeas, criadas com meio de cacto *C. hildmannianus* (Figura 16).

Identificou-se similaridade entre as sequências de buscas em *D. gouveai*, alvo de metilação em moscas adultas e machos, que compreenderam regiões do cromossomo 3L, LOC125776504, elemento transponível (TE) *Galileo*, e gene *eEF1alpha1*, com as sequencias genômicas em *Bactrocera dorsalis*, *D. melanogaster*, e *D. buzzatii*.

As sequências de DNA com as regiões do cromossomo 3L em *D. gouveai*, com a busca de 127pb da sequência de consulta, sendo 19pb de similaridade com *D. melanogaster*. Essa amostra de *D. gouveai* foi obtida de moscas machos criados no cacto *C. hildmannianus*. Não foi possível determinar a região de éxons, íntrons ou genes ao longo do cromossomo. Isto pode estar relacionado à pequena extensão do fragmento, com 127pb.

Uma sequência de DNA com extensão de 315pb apresenta similaridade com 60pb no genoma de *Bactrocera dorsalis*, correspondendo a região do LOC125776504 (RNA não codificador). Essa região é alvo de metilação DNA em machos criados em cacto *C. hildmannianus*, em *D. gouveai* (Figura 17).

O elemento transponível *Galileo*, com amplificação de 307pb, é uma sequência alvo de metilação de DNA em *D. gouveai*, quando os indivíduos foram criados em meio com o cacto *C.*

hildmannianus. Essa região de TE *Galileo* apresentam similaridade com 49pb deste elemento no genoma de *D. buzzatii* (Figura 18). Resultados indicam que esta região metilada pode estar localizada próxima de regiões de quebras cromossômicas (Figura 21).

Interessantemente, o gene *eEF1alpha1*, com amplificação de 218pb, é um alvo de metilação de DNA em *D. gouveai*. Essa região do *eEF1alpha1* apresenta 64pb de similaridade com o mesmo gene no genoma de *D. melanogaster*, e presente também em *D. buzzatii* (Figura 19). Essa sequência envolve uma região de éxon em moscas adultas de origem masculina que emergiram em tecidos de *P. machrisii*, em *D. gouveai*.

O conjunto de sequências foi usado na inferência por análise de distâncias genética, com o método Neighbor-Joining (NJ). Uma hipótese para a organização dos dados alvos de metilação é estruturado com um agrupamento entre as sequências, embora com poucas similaridades entre as sequências alvo de metilação quando os organismos usaram diferentes recursos no desenvolvimento biológico (Figura 20).

As duas regiões com similaridade com gene *eEF1alpha1*, em amostras de moscas fêmeas e machos criadas em meio do cacto *P. machrisii*, correspondem às suas regiões transcricionais, na análise do programa BLASTx (Tabela 13). Esses dados apresentam similaridade (>55.0%) com sequências disponíveis em *Marpesia eleuthea* (borboleta; GenBank: ABV53950.1; *E-value*: 0.005) e *Tenebrio molitor* (besouro; GenBank: KAH0819316.1; *E-value*: 2e-06).

Tabela 12. Amostras obtidas as sequências de DNA alvos de metilação, em *Drosophila gouveai*, com dados moleculares sobre extensão em Pares de Bases nucleotídicas (pb), Porcentagem de Bases de Citosinas (PBC), análise de similaridade no BLASTn, Localização Genômica (LG), Similaridade e Número de acesso no Banco FlyBase ou GenBank

Amostras	Sequências	Escala molecular (pb)	PBC(%)	BLASTn	LG	E-value	Similaridade (%)	Acession number
GQOM <i>MspI</i>	ACGCAGTCAATtATGACTCTGTCACCGTACAGTGAATTGGTAC ACACTCAAACAGTAAGTTTACAGAA	68pb	22,1%	No found	No found	No found	No found	No found -
GCOM <i>MspI</i>	TCCACAGAAACCTGCTGTTGTCTaCATGgTACGCGgAAGCAA CTGACCAAGCCTCCAGGCAGCCCCGCCCGCTGCTGAGCACA CACATATACACCATGGCAGCCGCTAgATGCCCCACCCTCCAC TGGCCTGGGCTGTGCGCAGTCCACCGCGAGTGCAGTG ACTCAAAACAGGGCCAGAGCAGCGGCaAGACCGCAcAGCGC GAGGGGAGAACAGCACACCACTGGGTTcTGCTCC	242 pb	36,4%	Gene <i>Piezo</i>	éxon	0.028	100%	FBgn0264953
GPOH <i>HpaII</i>	TTGTGATGGAACCGTTCTAGACTCATCGGTCTGTATTCTGC GAACCGTTCTAGACTCATAGTCGGGGCTCTTTAGACAGATAC CTTTGTTTAGCATACTACGCACAAGCGGCTGTTTTGCCCTAT ATGAAAACCTGACGCGCTAGAGATGACGCCAATATACGCCG TCGGAGACGATGATAACCACCTGATAAAACCTGGTATGTGC CCCCTGCTCGACCGCGAATGTAGTGACTCAAAATAGGGCC AGAGTGTCGACCAAACCTTTTTTATTTTAAGGGAAAGTGACTA ACCTTCTCGAACTTCTCGATGGTTCGCTGGTCGATACCACCC CACTGGTAGATCAAGTGACCGGTCTGT	360 pb	25%	Gene <i>eEF1alpha</i> 1	íntron	1.14e-16	92,2%	FBgn0284245
GCTM <i>MspI</i>	AGCAACAAATTGATGTTAAAAGAATATTATTTTTACCTTACC TTAAAATCACCTTAAACCTTTCACAGTTGGAACCGTTCTAGA CTCATCCATGTGCAAAAACGTCTAAGCAACAAATGGATGTT AA	127pb	19,7%	<u>3L</u>	2457431 - 2457432 9	0.054	100%	No found
GCTM-1 <i>MspI</i>	TGAACGTATTAGACTGGGCACTGGGCAAGCTGCTATTCGAA GCATTATAGCAAGCAACATACGCACACAGCAGACGGACCTC ACTATAATCGACTCAGCTCATCTTCTGATTAGCTATAAATA TACTCCTTGACTGCGTACCAATTCAACTAAGAAGAACTTTAA	315pb	24,1%	LOC12577 6504	ncRNA	1e-04	80.56%	XR_00742176 8.1

	GAATCCTTAAAGTTCATTGAAAACGTATTAGACTGGGCACTG GGCAAGCTGATTTCGAAGCATTATAGCAAGCAACATACGCAC ACAGCAGACGGCCTCACTATAATCGACTCAGCTCATCATTCT GATTAGCTATAAATATACTCCTGC							
GCTM-2 <i>MspI</i>	TGACTGCGTACCAATTCAACTCTTCGGTGCTGTGTACATATC ATACATTATGCATACATAACAAATGCTCGGCTGCAAAGCAG CACACAACAATTTAAATTTCTTCTAGCCTCGAGTCCATTGGT TCCCTTTTGAAGGAGCTGGAGTCCCCAAGGCTCAGTGCTGG GGCATTATGTACGAAAGCGTACTGCGCTTCGAACTACTAGC AGGCGCGCACACCATTGACTTTCCTGATGATGTAGGTGTCGT CTCCATAGGCAGAAGTTCAAAGATATTGAGGATATGTGCAG CGCAAATATCAATCGC	307pb	24,1%	Transposon <i>Galileo</i>	TE	9e-06	84.5%	AY756168.1
GPTH <i>HpaII</i>	CCACCAATATTCATGGTGAATTGGTACGCAGTCAGAGCTTAT CTGCTCATCATcAAACCCAgACAAAAACCGGGGGGGGGTTTT TCCAAAGAaAACATGCTGACATCGCTTCACAGACCGGTCCT TGATCTACCAGTGCGGTGGTATCGACAAGCGAACCATCGAG AAGTTCGAGAAGGTTAGTCACTTTCCTTCCATCGCGCGTCC TTTGTCAT	218 pb	26,6%	Gene <i>eEF1alpha</i> 1	éxon	1.2259 4e-21	87.5%	FBgn0000557

Nota. As amostras de DNA genômico foram obtidas de indivíduos cultivados em meio padrão de laboratório (GQOM), e nos cactos de *C. hildmannianus* (GCOH), e *P. machrisii* (GPOH), em tecidos de ovários de moscas fêmeas. Outras amostras de DNA foram obtidas para indivíduos criados nos meios de cactos com *C. hildmannianus* (GCTM), e *P. machrisii* (GPTH), em tecidos de testículos, em machos e adultos. Os ensaios enzimáticos compreenderam o uso de *EcoRI-HpaII* (*HpaII*-H) e *EcoRI-MspI* (*MspI*-M). TE- Elemento transposível/móvel. UTR – Região Não Traduzida (não codificante).

Tabela 13. Análises de Bioinformática com buscas de regiões genômicas com similaridade (BLASTn), regiões transcricionais (BLASTx), predição de Sítios de Início de Transcrição, e Regiões Promotoras, nas sequências alvo de metilação no DNA de moscas adultas de *D. gouveai*.

Amostras	BLASTn	BLASTx	Sítio de Início de Transcrição	Regiões Promotoras
GQOM	Não encontrado	Não encontrado	SIM	NÃO
GCOM	Gene <i>Piezo</i>	Não encontrado	SIM	SIM (<i>score</i> 0.93)
GPOH	Gene <i>eEF1alpha1</i>	eEF1alpha1	SIM	SIM (<i>score</i> 0.91)
GCTM	3L	Não encontrado	SIM	NÃO
GCTM-1	LOC125776504	Não encontrado	NÃO	SIM (<i>score</i> 0.86)
GCTM-2	Transposon <i>Galileo</i>	Não encontrado	SIM	NÃO
GPTH	Gene <i>eEF1alpha1</i>	eEF1alpha1	SIM	SIM (<i>score</i> 0.95)

Nota. As amostras de DNA genômico são provenientes de indivíduos (tecidos reprodutivos de fêmeas) que se desenvolveram em meio padrão de laboratório (GQOM), no meio do cacto *C. hildmannianus* (GCOH), e em *P. machrisii* (GPOH). As amostras de DNA foram analisadas indivíduos (tecidos reprodutivos de machos) criados no meio do cacto *C. hildmannianus* (GCTM), e em *P. machrisii* (GPTH). Os ensaios enzimáticos compreenderam o uso de *EcoRI-HpaII* (*HpaII*-H) e *EcoRI-MspI* (*MspI*-M).

```

      14010      14020      14030      14040      14050      14060      14070      14080
Dmelanogaster_Piezo  CGTAAGATTGCGACACTTCAATTACGAATTAATTAATTAAGTGAAGTATTGCGAGACTGCGGGCATTAACTGCAACAA
Dbuzzatii_piezo     .....G.T.....T...G...
Query sequence      -----

      14110      14120      14130      14140      14150      14160      14170      14180
Dmelanogaster_Piezo  TCGAGAACCAAATGCGGCCATGTGTGCCGCTGATCGTGAAGAC-AGCGTTTGTGCTAATGTTTTG-GGTGACATCAGACA
Dbuzzatii_piezo     .T....T..T....T..C..C.....C..A..-T..C..C....G....C.-.....G..G..T..
Query sequence      -----TC.ACAG.A..CT..TG....CTAC...G.AC.C..AAG..A.TGAC..

      14210      14220      14230      14240      14250      14260      14270      14280
Dmelanogaster_Piezo  GATCGCCGACGGGACAGCACTCTGGCGGACTTCATTGCCCCACTGCAGATTACTGTGGGATCGGCTGOCCTCAAGCTACCTC
Dbuzzatii_piezo     ....G..T..T..T....CT...T..TA.T..A..G...T.....A.....C..A..G.....C.....
Query sequence      .....CC...T.CT.-----A.ACATA.T..AC...G..A--G.C..TAGA.G.CCCACCC..CA....CTGG..C-TG..

      14310      14320      14330      14340      14350      14360      14370      14380
Dmelanogaster_Piezo  CCTCAAAGTTCCTAAAGAAGGCCGGCATGTGATCAAAAATCTATTGGTCCGTCTGTGGATCTGGCTACTCGTCTGGTTA
Dbuzzatii_piezo     .....T..G...T..C.A....C..C.....G....C....C...
Query sequence      .....GTG..GT.ACT.A...C.G...A.A.CA.C.GC--.G.C.GCACA.C...AGG.GA..A.A.CAC..CAC.--...C

      14410      14420      14430      14440      14450      14460      14470      14480
Dmelanogaster_Piezo  TGGCGAQAACATGACCGOCTTCCGCATCTGCTACATGGCCCTGTTTCTATTCTTCTGCTAGTCTTTCAATCTTCGTCCAA
Dbuzzatii_piezo     G.....G.T.....T....T....T....TT....C.....C.....
Query sequence      -----

      14510      14520      14530      14540      14550      14560      14570      14580
Dmelanogaster_Piezo  TACGGCTTCTGGCTGTTTCTGATCTTCTATGGCATGTCCACTTATATGATCTACACATATCAATTCGACAAAGTTCGAC
Dbuzzatii_piezo     .....G.....T.....A...
Query sequence      -----

```

Figura 16. Alinhamento de seqüências de nucleotídeos do gene *Piezo* em *Drosophila melanogaster*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA em moscas fêmeas criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).

```

      1210      1220      1230      1240      1250      1260      1270
B_dorsalis-ncRNA-LOC125776504 AATTGTTTCAGATCGAATTAAGCTATAGCATATAGCTOCCATACAAACTGAACAGATAGTTACTACCAGAAATG
D_buzzatii_scaffold28      ..GC..C.--AT.....GC.T.....CAGC.A.----.TGC..AC.G...CG.ACCTC..T.T..TC.
Query_sequence            ..GC..C.--AT.....GC.T.....AGC.A.----.GC..AC.G...CG.ACCTC..T.T..TC.

      1310      1320      1330      1340      1350      1360      1370
B_dorsalis-ncRNA-LOC125776504 AACCGAAGTTAACGTTTTTCTGTTTAAATGTAACCTTAAACTATGGTACAAAATATAAAATATCACTGC
D_buzzatii_scaffold28      ..ATAT.T.C----C..G-----
Query_sequence            ..ATAT.C.C----C..GACTGC..ACC.AT.C...A.G.AG.ACT.TA.G..TCC.T...GT...T..A

      1410      1420      1430      1440      1450      1460      1470
B_dorsalis-ncRNA-LOC125776504 CAAAAATTCAAAATGTGGCTTTCCTCTTCTACTTCTGGAACTTCTCCATTAACCTCCACCAATTTTTTC
D_buzzatii_scaffold28      GCTG.T.CGAAOC...A.A..AAG.AA.A.A.GCACA.A.C.G..GG.CT..C..TAAT.GA.TCAOC.CA
Query_sequence            -----

```

Figura 17. Alinhamento de seqüências de nucleotídeos que corresponde ao ncRNA LOC125776504 em *Bactrocera dorsalis*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Esta região foi amplificada em PCR, com amostras de DNA com moscas machos criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).


```

          310      320      330      340      350      360      370
D_buzzatii - TE Galileo seq_t AAAAGAAAACCTCATCTGTCAATGTATCCCAATGAAAATGATAATCCCAATGCTTTGTTAGCTTTGGA
D_buzzatii - TE Galileo seq_p C.GT.T.G.TG.TG.GA.--.C.C.CTGT..TG.TTC.GGTGG..T.T.TG..TG.AA.GATAG.GCA.T
Query Sequence -----

          410      420      430      440      450      460      470
D_buzzatii - TE Galileo seq_t ACCATACAACACATAGACTGGACAACCTGGAACAAATTAATGTGCACACTTATTTTTGACATTAGAACCCC
D_buzzatii - TE Galileo seq_p .TAGACTTTTTG..T.TT.AA.T.G.CATGTGTT..TGACA...AT.C.TA.AGG..CGCCCTT.CAAGG
Query Sequence -----

          510      520      530      540      550      560      570
D_buzzatii - TE Galileo seq_t TTTTCGGGTCGGGAGAGAAAAATTGCTGTGACAAAATCATTCCGCTTCGAGCCCTCGACGGTGATGGTCG
D_buzzatii - TE Galileo seq_p AA.GA.A.A.TATTTTC.TG.CC...AATATC.TC..AA.G.ATAT.G.ATTACAGG..T.A.T..TACAA
Query Sequence -----

          610      620      630      640      650      660      670
D_buzzatii - TE Galileo seq_t TCTTTCCCTTGTCTCTCTTTCTCTTTCTGCCCATGCTTTTTACTCTCTGCTC-TCTTCTCACAAAAATATA
D_buzzatii - TE Galileo seq_p AT..GGAA.GA.AACAA.G-ATC.CT.CTTTGT...CACG.CAAAAT.AAA.CAAGATC..TT.G..A--
Query Sequence -----

          710      720      730      740      750      760      770
D_buzzatii - TE Galileo seq_t ACAAAAATTGAAATTTCTTCTCGACGGATCCCATTTGTTCCCTTCTGAGCGAAGGATATACCGAATTCCTT
D_buzzatii - TE Galileo seq_p G..C..A.ACG.ACC.AA.CATTTTT..ATTT.A.TCCATAA.ATCC.AAATGCTT.C..TGA..G.T..
Query Sequence -----

          810      820      830      840      850      860      870
D_buzzatii - TE Galileo seq_t TGTTAAGCOTTTTCATGATTATTTGTGAAAGGATGAAAGTGTGTGATGTGCATACAGAAATTTTTTC--
D_buzzatii - TE Galileo seq_p .A.--.TAGA.G.A...CAA..CTAGCTAATATAAT..AA.A.AGC.G.TCC..CCTT.C.G.A..TCA
Query Sequence -----

          910      920      930      940      950      960      970
D_buzzatii - TE Galileo seq_t TTCC-TGCTATATGTGTAGGCATTTATGTGCTGTGTGCAATATGCTTGTGCTGCTTATGACACATTAGAAC
D_buzzatii - TE Galileo seq_p A.GAAGAA.CT.CA.ATCCA.AT...GAAATAGAA.TGT..T..GAT.CAACAAAT...AACAT.C.A..T
Query Sequence -----

         1010      1020      1030      1040      1050      1060      1070
D_buzzatii - TE Galileo seq_t AGCGATGTGAGCTTGACATACACACACATTTACACTTCTTTCGCTATATGTGTAGGCATTTATGTGCTGT
D_buzzatii - TE Galileo seq_p T...GGA.--T..CCTGC...AG.TGTTGCAA.G.AC.GA.T.TG.AG.T...CCA..C.-----TT.AA
Query Sequence -----

```

Figura 18. Alinhamento de seqüências de nucleotídeos que corresponde ao elemento móvel *Galileo* em *Drosophila buzzatii*, com seqüência completa (seq_t) e parcial (seq_P), e em *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA com moscas machos criadas em *C. hildmannianus*, na espécie *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-MspI*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).

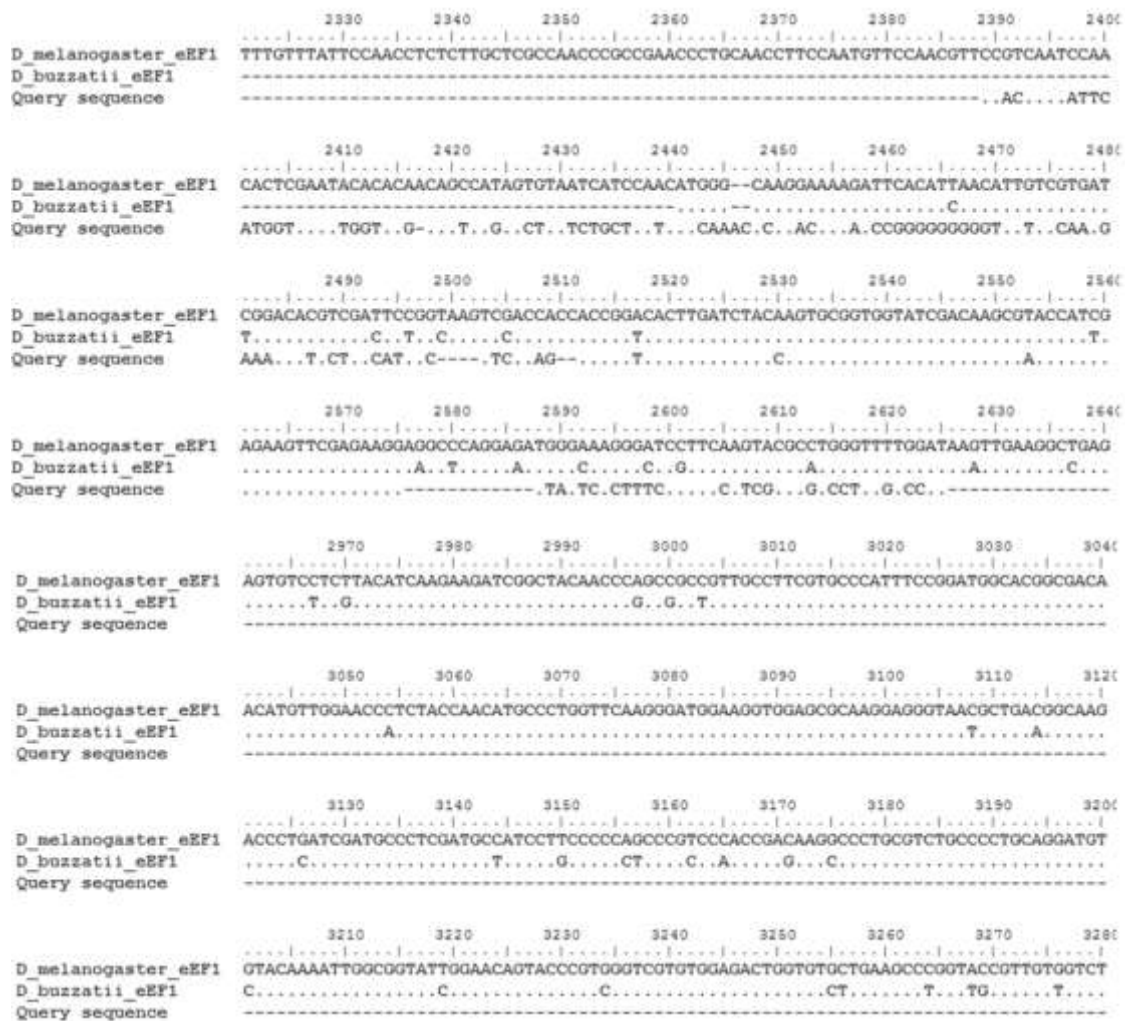


Figura 19. Alinhamento de seqüências de nucleotídeos do gene *eEF1*alpha1 em *Drosophila melanogaster*, *Drosophila buzzatii*, e *Drosophila gouveai* (Query sequence). Essa região foi amplificada em PCR, com amostras de DNA em moscas machos que foram criadas com o cacto *P. machrisii*, em *D. gouveai*. O ensaio de MSAP realizado com *EcoRI-HpaII*. Os símbolos indicam regiões com similaridade (.) e substituição/deleção de bases (-).

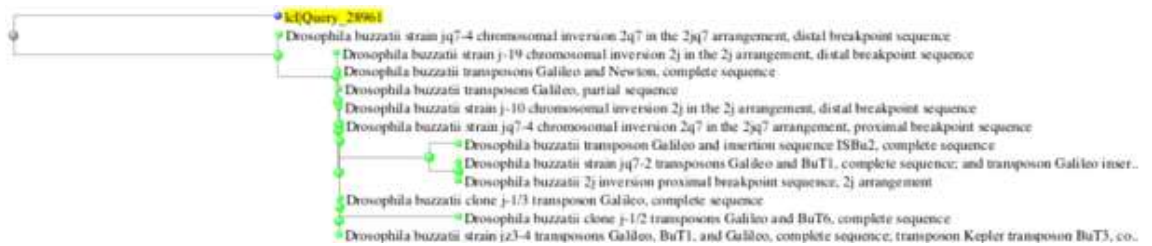


Figura 20. Dendrograma obtido por meio do método Neighbor-Joining(NJ), com os dados das seqüências de DNA alvo de metilação em *D. gouveai* (Query sequence_28961), com regiões gênicas disponíveis em *Drosophila buzzatii*. A região em *D. gouveai* corresponde ao TE *Galileo*. As seqüências estão disponíveis na Tabela 12

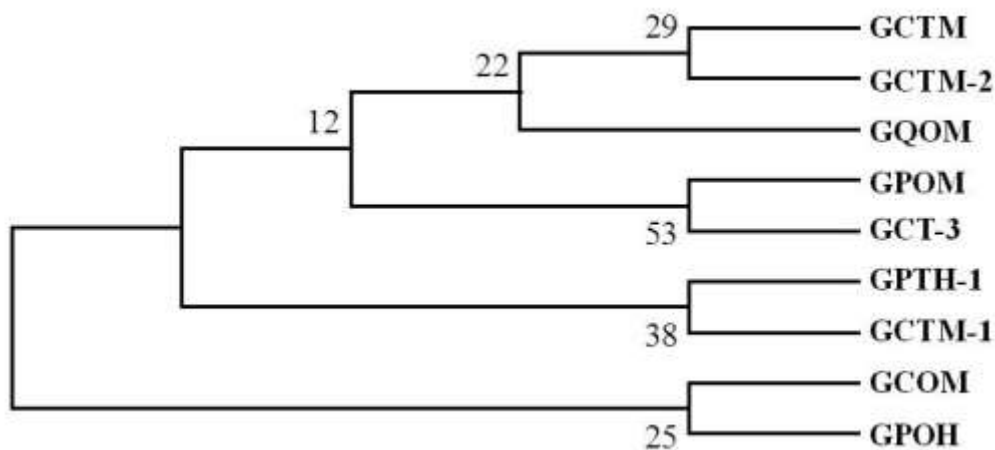


Figura 21. Estrutura dos dados com o método Neighbor-Joining(NJ), com o uso das seqüências de DNA-alvos de metilação, em *Drosophila gouveai*. Essas amostras são provenientes de indivíduos que se desenvolveram em diferentes ambientes, com moscas fêmeas tais como meio padrão (GQO), cacto *C. hildmannianus* (GCO), e em *P. machrisii* (GCO). Uma parte das amostras representa indivíduos moscas machos que se desenvolveram em meio com cacto *C. hildmannianus* (GCT), e em *P. machrisii* (GCT). Os ensaios de MSAP foram realizados com enzimas representadas por *EcoRI-HpaII* (H) e *EcoRI-MspI* (M). Todas as seqüências estão disponíveis na Tabela 12. Inferências estatísticas foram obtidas com 1000 *bootstrap*, com os valores indicados próximos aos cladros.

5.8 Regiões regulatórias moleculares

5.8.1 Regiões de sítio de início transcricional

De modo geral, identificamos sete regiões preditas com funções de sítios de início de transcrição nas sequências de DNA, em *D. gouveai* (Tabela 13). Um destes sítios de transcrição foi predito na sequência de DNA em moscas fêmeas criadas em meio padrão, embora sem a determinação de genes próximos. Esse fragmento de DNA pode ser uma região com função de transcrição para grupos de genes, que, devido ao tamanho da sequência nucleotídeos, pode ter dificultado para determinar genes vizinhos. Em contraste, duas regiões preditas com funções transcricionais foram relacionadas com os genes *Piezo* e *eEF1alpha1*. Essas amostras são de moscas fêmeas que foram criadas em meios de cactos, com *C. hildmannianus* e outras em *P. machrisii*. Observou-se que essas regiões com sítios de início de transcrição ocorrem no DNA de moscas fêmeas quando se desenvolvem em diferentes ambientes.

Descobriram-se quatro sítios preditos sítios de início de transcrição, que compreendem as regiões do cromossomo 3L, transposon *Galileo*, e o gene *eEF1alpha1*, em machos adultos (Tabela 13). Os dados indicam que três desses sítios de transcrição ocorrem no DNA em indivíduos que se desenvolveram e emergiram no meio de cacto *C. hildmannianus* (Tabela 13). Um sequência predita para sítio de transcrição envolve o gene *eEF1alpha1* em indivíduos adultos que se desenvolveram no meio de *P. machrisii*.

Por fim, observou-se que as sequências alvos de metilação podem estar relacionadas com sítios de início de transcrição. Isto pode colaborar com o processo de regulação gênica nos organismos de indivíduos quando se desenvolvem em diferentes ambientes ecológicos, em *D. gouveai*.

5.8.2 Predição de regiões promotoras

Com relação às regiões regulatórias, foram preditas cinco regiões com funções promotoras nas sequências de DNA, em *D. gouveai* (Tabela 13)(Apêndice I). Uma dessas regiões promotoras está localizada na sequência do gene *Piezo*, posição inicial em 159 e fim em 209 da sequência de DNA, em organismos que se desenvolveram em meio com cacto *C. hildmannianus*, e emergiram moscas fêmeas. Outra região promotora foi encontrada com relação a sequência do gene *eEF1alpha1*, com moscas criadas em tecidos de *P. machrisii*. Entretanto, foram preditas três possíveis regiões promotoras localizadas nesta sequência, sendo estas: 1) início na posição 114 e fim em

164 (*score* 0.85); 2) início em 146 e fim em 196 (*score* 0.87); 3) início em 228 e fim em 278 (*score* 0.91)(Anexo G). De fato, essa sequência do gene *eEF1alpha1* apresentou três possibilidades de regiões promotoras, mas utilizou-se o parâmetro estatístico para determinar essa posição como promotora deste gene, posição com início em 228 e fim em 278.

Os dados das sequências indicam duas regiões com predição para a função de promotor, que estão localizadas no LOC125776504 (ncRNA) e o gene *eEF1alpha1*, em adultos machos de *D. gouveai* (Anexo G). A região promotora no LOC125776504 foi predita com posição inicial em 109 e fim em 159. Essas duas regiões regulatórias foram preditas no DNA de organismos que se desenvolveram no meio de cacto *C. hildmannianus*, e emergiram moscas adultas e machos. Por fim, foi predita uma região promotora na sequência de *eEF1alpha1*, quando os organismos usaram o meio com tecidos do cacto *P. machrisii*, e emergiram moscas adultas e machos.

Por último, observou-se que as sequências de DNA-alvo de metilação estão relacionadas com regiões promotoras que podem apresentar função da regulação da expressão de genes, em indivíduos fêmeas e machos que usaram diferentes recursos ecológicos, em *D. gouveai*.

5.8.3: Análises funcionais de Ontologia dos Genes (GO)

Identificações foram realizadas entre os dados obtidos com *D. gouveai* e os disponíveis para *Drosophila melanogaster*, com os códigos obtidos no Flybase para os genes (IDs): *Piezo* (FBgn0264953); *eEF1alpha1* (FBgn0284245); TE *Galileo* (FBte0001170); *eEF1alpha1* (FBgn0000557).

Todos os genes foram usados nas análises de Ontologia de genes (GO). Esses resultados foram obtidos nas análises integradas sobre os termos de processos biológicos, componentes celulares, e funções moleculares para os genes alvos de metilação em *D. gouveai*, com o auxílio do programa David (GO e KEGG) (Tabela 14 e 15; Figura 23).

Tabela 14. Análise do *Termo Gene Report* com as ferramentas do programa David v.6.8. Os genes são os mesmos alvos de metilação e sequenciados em *D. gouveai*. As informações dos genes são apresentadas por categoria, termo/função, Fase *Discovery Rate* (FDR), e *P-value*.

Categoria	Termo	Fold Change	FDR	P-value
UP_KW_DOMAIN <i>Piezo</i>	Transmembrana	2.9E0	2.3E-1	1.2E-1
UP_SEQ_FEATURE <i>Piezo</i>	TRANSMEM:Helical	2.4E0	1.0E0	2.6E-1
GOTERM_CC_DIRECT <i>Piezo</i>	Integral componente de membrana	2.3E0	7.0E-1	2.8E-1
UP_KW_CELULAR_COMPONENT <i>Piezo</i>	Membrana	1.1E0	1.0E0	7.7E-1

*Score de Enriquecimento: 0.55.

Tabela 15. Análise descritiva dos genes por meio do programa DAVID v.6.8. Os genes são os mesmos alvos de metilação e sequenciados em *D. gouveai*. As informações funcionais dos genes estão apresentadas por categoria, termo/gene, similaridade (%), *Fase Discovery Rate* (FDR), e *P-value*.

Categoria	Termo	%	FDR	P-value
INTERPRO	Translation elongation factor EF1A, eukaryotic/archaeal	40.0	7.3E-3	7.3E-4
UP_SEQ_FEATURE	REGION:G4	40.0	1.8E-2	1.7E-3
UP_SEQ_FEATURE	REGION:G3	40.0	1.7E-2	1.8E-3
UP_SEQ_FEATURE	REGION:G5	40.0	1.8E-2	1.7E-3
UP_SEQ_FEATURE	REGION:G2	40.0	1.8E-2	1.7E-3
UP_SEQ_FEATURE	REGION:G1	40.0	1.8E-2	1.7E-3
GOTERM_CC_DIRECT	Eukaryotic translation elongation factor 1 complex	40.0	9.1E-3	1.8E-3
INTERPRO	Translation elongation factor EFTu/EF1A,C-terminal	40.0	1.1E-2	2.2E-3
GOTERM_CC_DIRECT	Translation elongation	40.0	2.3E-2	2.6E-3
UP_SEQ_FEATURE	DOMAIN:tr-type G	40.0	2.4E-2	2.7E-3
INTERPRO	Translational elongation factor EF1A/initiation factor IFgama, C-terminal	40.0	1.1E-2	3.3E-3
GOTERM_MF_DIRECT	Translational elongation factor activity	40.0	2.0E-2	3.9E-3
INTERPRO	Translation elongation factor EFTu/EF1A, domain 2	40.0	1.1E-2	4.4E-3
UP_SEQ_FEATURE	DOMAIN:Tr-type G	40.0	3.9E-2	5.0E-3
UP_KW _MOLECULAR _FUNCTION	Elongation factor	40.0	1.1E-2	5.3E-3
INTERPRO	Elongation factor, GTP-binding domain	40.0	1.3E-2	6.6E-3
INTERPRO	Translation elongation/initiation factor/Ribossomal, beta-barrel	40.0	1.5E-2	9.1E-3
UP_SEQ_FEATURE	NP_BIND:GTP	40.0	1.6E-1	2.3E-2
KEGG_PATHWAY	Nucleocytoplasmic transport	40.0	3.0E-2	3.0E-2
GOTERM_MF_DIRECT	GTPase activity	40.0	6.1E-2	3.1E-2
GOTERM_MF_DIRECT	GTP-binding	40.0	6.1E-2	3.7E-2
UP_KW_LIGAND	GTP-binding	40.0	9.3E-2	4.6E-2
UP_KW_BIOLOGICAL_PROCESS	Protein biosynthesis	40.0	2.4E-1	6.0E-2
GOTERM_BP_DIRECT	Translation	40.0	4.3E-1	9.5E-2

Gene *Piezo*

Com respeito ao gene *Piezo*, dados moleculares indicam que esse gene possui funções relacionadas aos canais mecanossensíveis, que envolve comportamento e resposta a estímulo, no processo biológico de identificação e avaliação da dureza do substrato por moscas fêmeas durante o processo de ovoposição, em *D. melanogaster*. Os dados sugerem que o gene *Piezo* apresenta marcas de metilação no DNA em moscas fêmeas que se desenvolveram no meio de cacto *C. hildmannianus*, em *D. gouveai*. Isto pode indicar que uma informação química na molécula do DNA pode estar relacionada ao comportamento de reconhecimento de recursos ecológicos, no processo de ovoposição.

Gene *eEF1alpha*

Informações sobre o gene *eEF1alpha* indicam funções previstas para as atividades de GTPase e fator de alongação da maquinaria de tradução (complexo ribossômico). Prevista para compor o complexo de alongação de tradução no sistema eucariótico. Ademais, esse gene participa no processo de determinação da vida adulta, que compreende a expressão gênica em diversas estruturas, o sistema circulatório, cabeça, e célula glial lateral (Gene ID 36271). Dados têm discutido o papel da diferença de expressão de *eEF1alpha* na longevidade de organismos em *D. melanogaster*, embora ainda não esteja clara a regulação do gene e os diferentes fatores ambientais relativos à característica de longevidade (Shikama & Brack, 1994).

Os dados apresentados neste trabalho indicam que o gene *eEF1alpha* é alvo de metilação em sequências de DNA, quando moscas (machos e fêmeas) se desenvolveram nos tecidos do cacto *P. machrisii*. Dados das funções biológicas e processos moleculares para *eEF1alpha* sugerem a função de GTPase, no processo de tradução, e comunicação núcleo-citoplasma (Tabelas 14 e 15).

Por fim, as mudanças no estado de metilação no gene *eEF1alpha* podem contribuir para a regulação de expressão gênica, com implicações no processo traducional e comunicação núcleo-citoplasma, quando os organismos se desenvolvem em diferentes cactos hospedeiros, na espécie cactofílica *D. gouveai*.

Elemento Transponível *Galileo*

Obtiveram-se dados na literatura para o TE *Galileo* (GenBank – acesso AY756168.1) que sugerem funções na reorganização e estabilidade do genoma, com localização em inversões no cromossomo 2 e próximo a pontos alvos de quebras cromossômicas, em *Drosophila buzzatii* (Delprat *et al.* 2009) (Figura 20). Adicionalmente, descobrimos que a sequência de DNA do TE *Galileo* apresenta mudanças no estado de metilação quando os organismos que se desenvolveram em tecidos do cacto *C. hildmannianus*, e emergiram machos deste ambiente. Isto sugere que a metilação em partes do elemento TE *Galileo* pode contribuir para a regulação da sua replicação e estabilidade dos organismos quando exploram diferentes cactos, embora tenha se observado a metilação desse TE somente nos tecidos reprodutivos de machos.

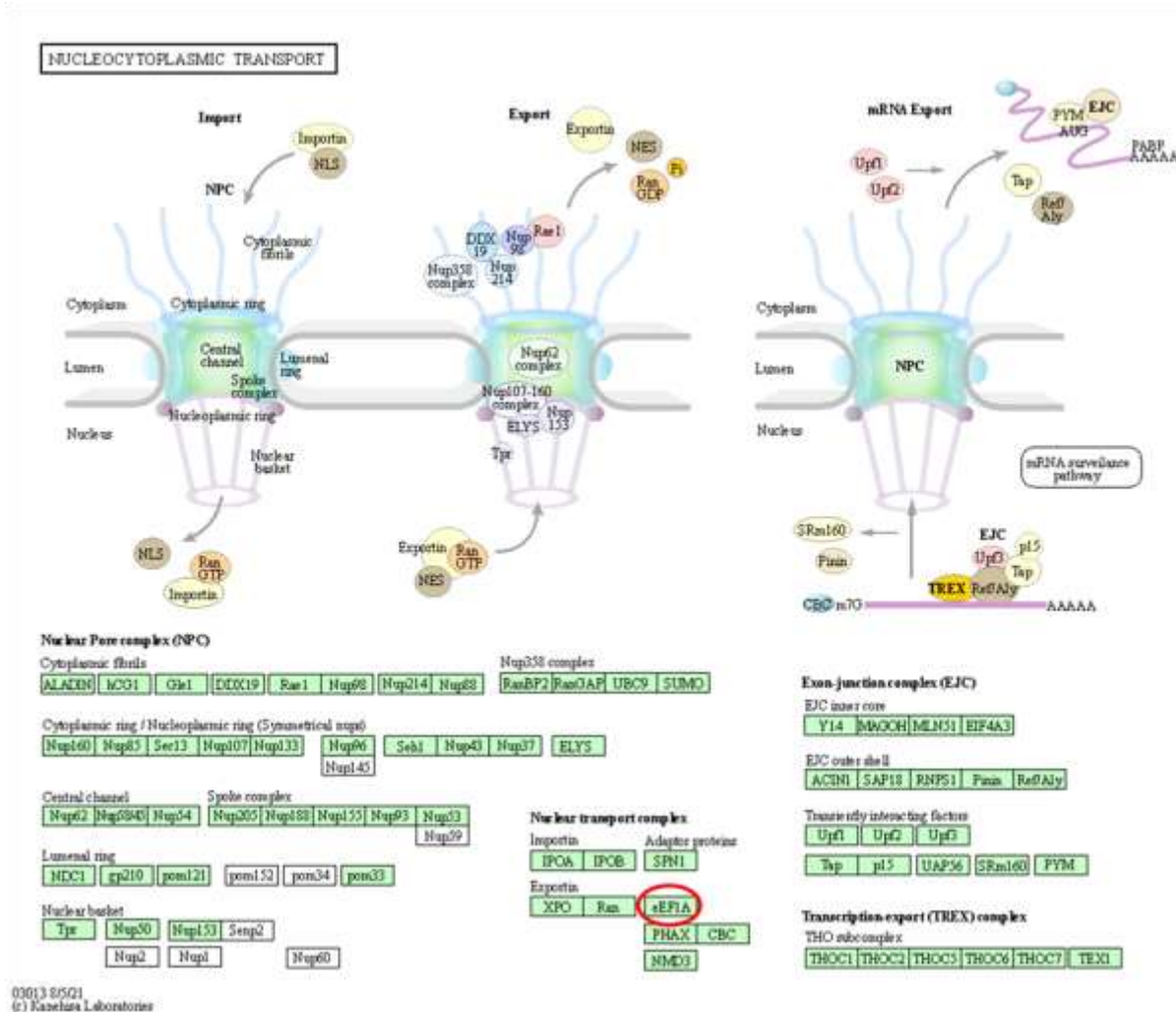


Figura 22. Network de proteínas com as interações entre eEF1A com a sinalização de transporte do núcleo-citoplasma. A representação contém o Complexo do Poro Nuclear (NPC) com a cesta nuclear, anel do lúmen, complexo Spoke, canal central, Simetria nups, complexo Nup358, e fibrilas citoplasmáticas. O complexo de transporte nuclear envolve as importinas (IPOA, IPOB), proteína adaptador (SPN1), exportinas (XPO, Ran, eEF1A, PHAX, CBC, NMD3). eEF1A é indicado pelo círculo vermelho. Complexo de junção-exon (EJC) é representado por núcleo interno EJC, escudo exterior ELC, e fatores de interação transiente. Complexo de transcrição-exportação (TREX), que envolve o subcomplexo THO.

Discussão

6. Discussão

6.1 Padrões de Metilação no DNA

Os resultados apresentados neste trabalho demonstram que o fenômeno de metilação diferencial no DNA apresenta mudanças em regiões genômicas e vias moleculares quando os organismos se desenvolvem em diferentes recursos ecológicos, na espécie cactofílica *D. gouveai*. Os perfis de metilação do DNA, por meio MSAP, combinado com abordagem de sequenciamento direto permitiu compreender as regiões metiladas e as vias moleculares que contribuem na mudança de sítios hospedeiros, e com abordagem de algumas questões como o efeito de origem parental.

Os dados de metilação de DNA estão de acordo com trabalhos anteriores, em que foi evidenciada a presença de metilação no DNA genômico em espécies do gênero *Drosophila*, com regiões com aproximadamente 1% do seu genoma (Lyko *et al.* 2000; Deshmukh *et al.* 2018).

Regiões diferencialmente metiladas são documentadas com preferência em sítios CpT, CpA, seguidos por CpC, e CpG (Lyko *et al.* 2000). Os dados indicaram sítios metilados nos alvos CpC e CpG, que coincidem com os resultados descobertos nos genomas de *D. melanogaster* (Deobagkar *et al.* 2004; Panikar *et al.* 2015), e *D. willistoni* (Garcia *et al.* 2007), com o uso de enzimas isoesquizomeros *MspI* (5'C↓CGG3') e *HpaII* (5'C↓CGG3'). Isso sugere que há regiões que podem apresentar motivo de DNA metilados e com funções conversadas entre as espécies de *Drosophila*.

Analisou-se o padrão de metilação diferencial no DNA genômico, e os resultados são consistentes entre os grupos de pupas em *D. gouveai*, e dados obtidos com *D. melanogaster* (Deshmukh *et al.* 2018). Em relação às mudanças de ambiente de desenvolvimento, testadas por meio padrão de laboratório, e nos cactos *C. hildmannianus*, e outros em *P. machrisii*, os dados indicam pouca influência no padrão de metilação no DNA em pupas na espécie cactofílica em *D. gouveai*. Uma explicação apresentada na literatura é a hipótese que o estado de metilação genômico em pupa de *Drosophila* apresenta função em processos ontogenéticos (Boffelli *et al.* 2014; Lyko *et al.* 2000), com diferenças em populações de células e tecidos-epececífiocs, e heterogeinidade epigenética em múltiplos sítios CpGs. Isto sugere que o processo ontogenético, com a complexidade na formação de tecidos e órgãos, que pode ser um fator que dificulta a análise da metilação diferencial no DNA genômico dos organismos nessa fase com relação ao ambiente, em *D. gouveai*.

Descobriu-se um contraste no nível de metilação relativa no DNA genômico nos tecidos reprodutivos de fêmeas e machos, em *D. gouveai*. Diferença no padrão de metilação no DNA foi relatada entre moscas fêmeas e machos, com as enzimas endonucleases *HaeIII* (5'GG↓CC3') e os isoesquizômeros *MspI* (5'C↓CGG3') e *HpaII* (5'C↓CGG3'), em *Drosophila willistoni* (Garcia *et al.* 2007), que corroboram os dados com diferença em relação ao sexo. Em contraste, a metilação no DNA de forma sexo-específico não foi documentada em *Drosophila tropicalis*. Isso pode indicar mecanismos epigenéticos diferentes entre as espécies de *Drosophila*. Nesse contexto, a metilação no DNA genômico pode apresentar função sexo-específico em *D. gouveai*, com semelhanças com *D. willistoni*.

É importante destacar que diferenças no estado de metilação no DNA têm sido documentadas entre fêmeas e machos, que envolvem diferentes processos biológicos em eucariotos, tais como: inativação do cromossomo X, *imprinting* genômico, microambiente nutricional (ácido fólico, vitamina B), atividade das enzimas DNA metiltransferases (DNMTs) e demetilases (TETs), processos de desenvolvimento, e fertilidade reprodutiva (Jones *et al.* 2012; Schmitz *et al.* 2019; Glastad *et al.* 2019). Os dados sugerem que machos e fêmeas apresentam diferenças no estado de metilação do DNA, embora não estejam claros os mecanismos moleculares em *D. gouveai*.

Os organismos com genomas semelhantes, em *D. gouveai*, podem apresentar variação de metilação no DNA com respeito a processos biológicos como sexo-específico, e respostas nas alterações de tecidos de cactos no desenvolvimento biológico de *D. gouveai* (Figura 9 e 10). Uma explicação pode estar relacionada à variação em composição química de tecidos vegetais, como os grupos de alcaloides, que são indicados com potencial tóxico que pode resultar em diferenças fenotípicas. Uma proposta interessante é que os organismos apresentam alvos de DNA genômico em *Drosophila* que contribuem para o processo de adaptação dos organismos, e a plasticidade de fenótipos.

6.2 Perfis de metilação de DNA

Um aspecto geral foi observado no DNA genômico com diferenças nos padrões moleculares em amostras de tecidos reprodutivos de moscas fêmeas (n=308 *loci*) e outras em machos (n=201 *loci*), em *D. gouvei*. Entretanto, observou-se uma tendência com maior intensidade de regiões metiladas na escala molecular entre 150pb a 800pb, em amostras de DNA em tecidos reprodutivos de fêmeas, e entre 200 pb a 700pb, em amostras de tecidos reprodutivos de machos, em *D. gouvei*. Prévios estudos obtiveram

informações semelhantes com escala molecular de 100pb a 600pb e 300pb a 1000pb com regiões mais intensamente metiladas no DNA genômico em insetos, como *Sogatella furcifera* (Zhou *et al.* 2013) e *D. melanogaster* (Panikar *et al.* 2015). Isto indica variações de metilação em regiões moleculares no genoma de *D. gouveai*, que corroboram dados de metilação com prévios estudos com outros insetos.

As razões de metilação de DNA indicam mudanças com relação ao ambiente (sítios de desenvolvimento) para moscas fêmeas e machos, em *D. gouveai*. Regiões genômicas e genes diferencialmente metilados estão relacionados à diferenciação de castas e sexo em cupins, *Zootermopsis nevadensis* (Glastad *et al.* 2016), e à formação de rainhas e soldados em formigas, *Camponotus floridanus* (Chittka *et al.* 2012). Esses autores sugerem maior evidência de marcas de metilação no DNA de organismos de fêmeas do que em machos. De fato, esses dados estão de acordo com o que se encontrou em relação à metilação do DNA em *D. gouveai*.

Diferenças foram encontradas em múltiplos *loci* sensíveis de metilação (MSL) de DNA e respostas dos organismos (populações de células) com relação ao ambiente, em *D. gouveai*, tais como: meio padrão (fêmeas - MSL=82; machos - MSL=55), tecidos de cacto *C. hildmannianus* (fêmeas - MSL=65; machos - MSL=54), e em *P. machrisii* (fêmeas - MSL=52; machos - MSL=42). Essa descoberta é interessante em virtude de indicar diferenças de metilação em relação ao sexo e respostas a mudanças que podem estar relacionados aos sítios hospedeiros de desenvolvimento biológico. Isto sugere mudanças na estrutura do epigenoma, com foco na metilação de DNA, quando os indivíduos se desenvolveram e/ou exploraram diferentes recursos ecológicos, na espécie cactofílica *D. gouveai*.

Descobertas de Lian *et al.* (2018) indicam que a restrição de alimentos e o uso de meio padrão de laboratório, no processo de desenvolvimento de larva para adultos, não resultam em diferenças marcantes na metilação de DNA genômico em *D. melanogaster*. Os resultados obtidos indicam discreta variação no meio de cultura padrão de laboratório em relação aos tecidos de cactos, na espécie cactofílica *D. gouveai*.

Há variações na metilação de DNA, no contexto de alvos moleculares, quando os organismos de moscas fêmeas e machos se desenvolveram e emergiram em diferentes recursos ecológicos. Descobriu-se maior quantidade de regiões de metilação, no DNA em tecidos reprodutivo de fêmeas, com a combinação E(ACC)/HM(TTC), quando os indivíduos se desenvolveram em meio padrão (MSL=18 *loci*), tecidos vegetais do cacto *C. hildmannianus* (MSL=12 *loci*), e em *P. machrisii* (MSL=6 *loci*) (Tabela 7). Em contraste, observou-se maior quantidade de alvos de metilação no

genoma com a combinação E(ACT)/HM(TTG) quando os organismos se desenvolveram no meio padrão (MSL=15 *loci*), e com a combinação E(AGG)/HM(TGA) nos indivíduos que usaram os tecidos de *C. hildmannianus* (MSL=11 *loci*), e no cacto *P. machrisii* (MSL=7 *loci*), em tecidos reprodutivos de machos de *D. gouveai*. Esses conjuntos de dados reforçam a existência de variação de metilação de DNA sexo-específico e nas mudanças de ambiente de desenvolvimento em *D. gouveai*.

Dados obtidos por Drewell *et al.* (2014) indicam diferenças no número de sítios CpGs metilados no DNA genômico em amostras de ovários e espermas, que podem implicar o processo de desenvolvimento de zangões e rainhas, na abelha *Apis mellifera*. Esse fenômeno é explicado por efeito de origem parental e de linhagem que pode resultar em regiões diferencialmente metiladas no DNA e alterar a expressão de genes, em *Apis mellifera* (Wu *et al.* 2020). Evidências indicam que regiões com metilação no DNA genômico pode ser transmitidas entre os parentais para a sua prole (F1), em espécie de vespas do gênero *Nasonia* (Wang *et al.* 2016), que sugerem uma herança intergeracional. Os dados deste trabalho podem sugerir uma herança intergeracional quando os organismos usam diferentes sítios hospedeiros no desenvolvimento *D. gouveai*. No entanto, são necessários mais dados para confirmar a herança intergeracional quando os organismos usam diferentes nichos ecológicos.

Um processo interessante refere-se ao fato de que a metilação de alvos CCGG afeta a conformação da estrutura helicoidal do DNA e respostas ao estresse de torção na molécula (Hörberg & Reymer, 2018). As sequências com os alvos de bases de Ctosinas não Metilados (C) e Metiladas em (C^m) no DNA, CCCG/CC^mGC e CC^mGA/CTGA, apresentam uma bimodalidade de deslocamento CpG, com torção e assimetria na rigidez no DNA metilado que podem bloquear a sequência alvo (Hörberg & Reymer, 2018). Com base nessas previsões de moléculas de DNA, a variação de metilação e não metilação de citosinas nas sequências de DNA pode significar uma mudança na conformação do DNA e apresentar implicações na estrutura do epigenoma em organismos, em *D. gouveai*.

6.3 Sequências DNA e mecanismos de regulação gênica

Os dados sugerem que as sequências alvos de metilação de DNA estão localizadas mais intensamente em cromossomos autossômicos, 2R, 2L, e 3L, no genoma de *D. gouveai*. Embora dados de genoma amplo tenham indicado metilação em regiões de cromossomos autossômicos e sexuais em *D. melanogaster* (Boffelli *et al.*

2014). Dessa forma, os resultados não indicaram uma relação clara com a compensação de dose relacionada ao sexo, como sugerido em mamíferos (Jones *et al.* 2012). Isto reforça que o padrão de metilação no DNA abrange diferentes regiões na estrutura do genoma, o que pode estar relacionado com fatores do ambiente, em *D. gouveai*.

As regiões diferencialmente metiladas no DNA apresentam similaridades com genes e locais no genoma quando os organismos se desenvolvem em diferentes condições: moscas fêmeas em tecidos do cacto *C. hildmannianus* (gene *Piezo*) e no cacto *P. machrisii* (gene *Factor elongation alpha 1*); moscas machos em tecidos de *C. hildmannianus* (regiões do cromossomo 3L, LOC125776504, elemento transponível *Galileo*), e no cacto *P. machrisii* (gene *Factor elongation alpha 1*).

Adicionalmente, as sequências de DNA diferencialmente metiladas contribuem para novas percepções sobre os mecanismos de regulação gênica e elementos no genoma, que envolvem regiões localizadas em éxons e íntrons de genes, regiões promotoras, sítios de início de transcrição, não codificadores de RNAs (ncRNAs), Regulador Não Traducional (UTR), e elemento transponível, nos organismos durante a exploração de nichos ecológicos em *Drosophila* cactofílica.

Um aspecto funcional indica que genes diferencialmente metilados no DNA participam de processos que envolvem mecanorreceptores na atividade de ovoposição (gene *Piezo*), e no processo de síntese transcricional e comunicação núcleo-citoplasma (gene *Factor elongation alpha 1*), em moscas fêmeas de *D. gouveai*. Em contraste, outros processos biológicos estão participando da viabilidade celular e estabilidade do genoma (ncRNAs, TE *Galileo*), e, adicionalmente, no processo de síntese transcricional e comunicação núcleo-citoplasma (gene *Factor elongation alpha 1*), em moscas machos de *D. gouveai*.

Cabe chamar atenção para propriedades e funções da metilação de bases de citosinas no DNA de genoma de eucariotos, tais como: (I) os grupos metil apresentam propriedades hidrofóbicas com tendência ao empacotamento e torção das fitas de DNA; (II) as bases de citosinas metiladas interagem com grupos complexos de proteínas de ligação aos sítios CpG (MeCP1, MeCP2) que tornam a cromatina mais condensada; (III) a compactação de DNA em regiões promotoras ricas em CpGs pode diminuir a interação com alguns fatores de transcrição ligados ao grupo metil em genoma de mamíferos; (IV) o fenômeno da metilação de DNA possui função de repressão na maquinaria de RNA polimerase na transcrição ou diminuição da afinidade de fatores de transcrição aos sítios de DNA, que pode resultar na diminuição do produto

transcricional (mRNA) ou silenciamento gênico (Nan *et al.* 1998; Bird & Wolffe, 1999; Lyko *et al.* 1999; Schmitz *et al.* 2019).

Os dados indicam que ocorre metilação diferencial em regiões promotoras preditas para os genes *Piezo*, em moscas fêmeas criadas no cacto *C. hildmannianus*, e no gene *eEF1alpha1* em fêmeas criadas em *P. machrisii*, em *D. gouveai*. Essa espécie de *Drosophila* tem sido encontrada em sítios criadores na natureza com relação ao cacto *P. machrisii*, que é usado preferencialmente, em relação a outros cactos colunar e palmar (Tidon-Sklorz & Sene, 2001). Considerando-se os dados de promotores em eucariotos mencionados anteriormente, a adição de grupo metil em regiões reguladoras pode diminuir o produto transcricional de *Piezo* quando fêmeas são criadas em cactos não usualmente na natureza (*C. hildmannianus*), e também diminuir a síntese de mRNA ou comunicação núcleo-citoplasma em moscas fêmeas que se desenvolveram no cacto usado na natureza (*P. machrisii*).

De fato, metilação diferencial de DNA foram encontradas com respeito a regiões com funções de síntese, tais como DNA ribossômico (rDNA) em moscas fêmeas de *D. willistoni* e *D. melanogaster* (Garcia *et al.* 2007; Panikar *et al.* 2015). Isto sugere que fêmeas podem apresentar menor capacidade de reconhecimento ao tecido do cacto *C. hildmannianus*, e, possivelmente, reduzir a síntese de proteínas ou transporte de mRNA na comunicação núcleo-citoplasma, quando os organismos se desenvolvem no cacto *P. machrisii*.

Dados na literatura têm indicado que modificações químicas no DNA podem resultar em fenótipo no comportamento de forrageamento (gene *For*), no cortejo sexual (Kramer *et al.* 2011; Anreiter *et al.* 2017), e diferenciação de oócitos em fêmeas (Lovino, 2014), em *D. melanogaster*. Os dados chamam atenção para o mecanismo de reconhecimento do substrato de ovoposição em *D. gouveai*.

Adicionalmente, descobriram-se regiões promotoras preditas em regiões metiladas de DNA próximas de ncRNAs, em machos criados em *C. hildmannianus*, e no promotor de *eEF1alpha1* em machos criadas no cacto *P. machrisii*. Nesse contexto, a metilação de citosinas em regiões promotoras pode resultar na diminuição de ncRNAs, no organismo de moscas machos que se desenvolveram no cacto não usualmente na natureza, *C. hildmannianus*.

Resultados de Soto *et al.* (2007) indicam que os compostos do cacto *C. hildmannianus* podem diminuir a viabilidade entre larvas para adultos de *D. gouveai*, quando os organismos usam este recurso, em relação ao cacto *P. machrisii*. Os dados indicam que a metilação, diferencialmente no DNA, pode implicar aspectos de

viabilidade dos organismos quando alteram o uso de recursos ecológicos, em machos de *D. gouveai*.

Em contraste, a metilação de sítios CpGs no DNA com posição em região do corpo gênico, conhecido como *gene body methylation* (GBM), éxons, íntrons e intergênicas, tem sido relatada sem efeito ou com baixa função na regulação da expressão gênica, em *Zootermopsis nevadensis* e *Apis mellifera* (Glastad *et al.* 2016; Maleszka, 2016). Nesse sentido, os de sequências de DNA, com alvo de metilação de bases de citosinas, em regiões internas do corpo gênico podem não resultar em diferenças drásticas na expressão de genes, em *D. gouveai*.

Encontraram-se, também, sequências de DNA metiladas com alvo em elemento transponível *Galileo*, em machos que emergiram no meio do *cacto C. hildmannianus*. Essa sequência fica localizada próxima a outros elementos transponíveis (Newton, Kepler, BuT1, BuT3, e BuT6) e região de breakpoint (2j)(Figura 21). A metilação de DNA tem sido relacionada com a regulação de elementos transponíveis e retrotransposons em mamíferos, e em *D. melanogaster* (Bird & Wolffe, 1999; Schmitz *et al.* 2019). Evidências de Deprat *et al.* (2009) sugerem que três inversões polimórficas no genoma de *D. buzzatii* podem estar relacionadas com a mesma família de transposons *Galileo*, que é indicado devido à estrutura incomum dos membros e atividade transposicional recente. Com base nesses dados, levanta-se uma hipótese de que o elemento transponível *Galileo* pode estar sob o processo de regulação por metilação diferencial no DNA, quando os organismos usam recursos ecológicos potencialmente tóxicos, em *D. gouveai*. Dessa forma, a metilação de DNA pode contribuir com a estabilidade do genoma por meio da repressão no processo de replicação de DNA desse elemento, em machos de *D. gouveai*.

6.4 Metilação de DNA e Reprodução

Os dados deste trabalho indicam que podem ocorrer alterações no padrão de metilação do DNA genômico de *D. gouveai*, e envolver os mecanismos de reprogramação, bem como repostas fisiológicas ao ambiente nutricional, considerando-se que a dinâmica tem sido relatada na desmetilação de pronúcleo masculino após 4 horas de fertilização de oócitos *in vitro*, em camundongos (Santos *et al.* 2002). Perda passiva de metilação ocorre de forma gradual até o estágio de mórulas, com o mecanismo de metilação de *novo*, e formação da massa celular interna (ICM) (Santos *et al.* 2002). Os autores sugerem que a reprogramação da metilação envolve as proteínas

DNMTs e TETs, com funções no *imprinting*, controle da expressão gênica, e no estabelecimento da totipotência nuclear, em mamíferos (Santos *et al.* 2002). Entretanto, ainda não elucidaram-se os mecanismos de reprogramação epigenética em *Drosophila*.

Bourc'his *et al.* (2001) descobriram uma incompleta reprogramação do padrão de metilação de DNA em clone de embriões bovinos. O processo de desmetilação do pró-núcleo paterno tem sido relatado em camundongos e humanos, na fase de mórula (Dean *et al.* 2001). Aberrações no processo de metilação de DNA têm sido relatadas entre as etapas de mórula e blastocistos, com não metilação de DNA em regiões de eucromatina nos cromossomos, e heterocromatina centromérica mais metilada do que em embriões normal em camundongos (Bourc'his *et al.* 2001; Dean *et al.* 2001). Esses autores indicam que eventos com alteração anormal dependem do tempo de desenvolvimento e reprogramações epigenética podem contribuir com altas incidências de anomalias fisiológicas ou após o nascimento do embrião.

Um sistema clássico é a síndrome de Silver-Russell (SSR; OMIM 180860), em humanos, que foi independentemente descrita por Silver *et al.* (1953) e Russell (1954). Essa síndrome é caracterizada por retardo no crescimento intrauterino que persiste no período pós-natal, baixa estatura, face triangular, clinodactilia e/ou encurtamento do quinto dedo de mãos, macrocefalia aparente e a assimetria corporal, alterações de orelhas, filtro nasolabial bem demarcado, hipotonia e hipotrofia muscular, atraso motor e neuropsicomotor (Monk *et al.* 2002).

Adicionalmente, alterações estruturais em anel no cromossomo 15 ou deleções 15q em pacientes com características de SSR, com características de heterogeneidade genética, sem relação clara com herança ligada ao sexo (Monk *et al.* 2002; DeLaval *et al.* 2006). Dados moleculares indicam que o gene *IGF1R* (*receptor de fator semelhante à insulina I*), com localização no braço curto do cromossomo 15, compõe a família de fatores de crescimento semelhante à insulina (*IGF*) (DeLaval *et al.* 2006). Essa família gênica inclui *IGF1* e *IGF2*, com função na regulação do crescimento fetal. Experimento com camundongos indicam que o gene *IGF2*, cuja expressão do alelo é o paterno e não expressa o materno, compreende o mecanismo de *imprinting* genômico, que regula o crescimento fetal e pós-natal, em camundongos (DeLaval *et al.* 2006).

Evidências indicam que o alelo não metilado materno ICR1 (regiões de controle de *imprinting*) é ligado a várias cópias da proteína CTCF (fatores de ligação-CCCTC), que cria um limite de cromatina, que é um inibidor de *IGF2* e acentuador da expressão *H19*, em camundongos. Esse processo molecular resulta na não expressão de *IGF2* materno (Verona *et al.* 2003). De fato, as diferenças na metilação do DNA nas regiões

de ICR garantem que o alelo paterno do gene *IGF2* seja expresso e o *H19* (ncRNAs) com origem materna, no processo de desenvolvimento (Moss & Wallrath, 2007).

O mecanismo fisiológico em *Drosophila* contém órgão-chave com funções metabólicas conservadas com vertebrados, que envolvem análogo de insulina, fatores de crescimento semelhante a insulina e alvo de sinalização da rapamicina (TOR), com regulação de açúcares circulantes, armazenamento e metabolização de energia (Baker & Thummel, 2007; Schlegel & Stainier, 2007). Moscas fêmeas podem transmitir fenótipos com alteração da massa corporal para a prole na geração F2, com relação a características de obesidade, que envolve mecanismos epigenéticos e reprogramação do metabolismo quando o organismo se desenvolve em diferentes condições do ambiente (glicose, glicogênio, trealose), em *D. melanogaster* (Buescher *et al.* 2013). Os dados sugerem que os alvos de metilação de DNA podem participar de processos de reconhecimento de substrato de ovoposição e fisiológicos (apoptose celular, síntese de proteína e comunicação núcleo-citoplasma) quando os organismos se desenvolvem os recursos do cacto *C. hildmannianus*, e *P. machrisii*, em *D. gouveai*.

Um aspecto central é que micronutrientes e/ou compostos orgânicos presentes em tecidos vegetais podem alterar o desenvolvimento, com diferenças no padrão de metilação no DNA genômico dos organismos. Evidências obtidas por Hara *et al.* (1994) demonstraram que compostos de alcalóides em cultura de células contribuem com uma notável aumento na atividade de metiltransferases. Adicionalmente, a concentração de alcaloides pode estar relacionada a mudanças no padrão de metilação do DNA e regulação da expressão de genes, com participação de processos metabólicos em plantas (Bernoville *et al.* 2020). A concentração de alcaloide pode ser um fator que contribui para a alteração de fenótipos nas espécies cactofílicas *Drosophila buzzatii* e *D. koepferae*, irmãs de *D. gouveai*, quando se desenvolvem em diferentes cactos colunares hospedeiros (Soto *et al.* 2014). De fato, uma hipótese é que o conteúdo de alcaloide nos tecidos vegetais nas diferentes espécies de cactos podem alterar os mecanismos de metilação do DNA (enzima Dnmt2 e outras enzimas ainda não elucidadas), e contribuir para a regulação de fenótipos em espécies de *Drosophila* cactofílica.

Adicionalmente, Mahajan *et al.* (2019) relataram que a deficiência de folato e vitamina B12 nos tecidos maternos (F1) altera o padrão de metilação global do DNA, com o aumento da deficiência de folato na placenta (F1) e sob todas as condições nos tecidos fetais (F2), em camundongos. Os compostos de folato e vitamina B12 são cofatores importantes para a atividade das DNMTs, com relação importante na metilação de DNA em camundongos, que podem alterar fenótipos em ambas as

gerações F1 e F2 (Mahajan *et al.* 2019). Esses autores indicam que a proporção dietética de folato e B12 resultou em alterações na expressão de transportadores, miRNAs e metilação do DNA genômico em associação com DNMTs. Dados indicam que vitaminas B1 e C apresentam efeitos relacionados com mudanças no estado de metilação de DNA, em bactérias de *Mycobacterium bovis* (Song *et al.* 2020). Nesse contexto, os compostos de folato e as vitaminas B e C parecem ser cofatores importantes para as funções das enzimas DNMTs, que são encontradas em diferentes eucariotos. Nós obtivemos previamente a proteína de Dnmt2 na espécie *D. buzzatii* ancestral comum neste cluster (Apêndice J), embora ainda não esteja elucidada a evolução deste mecanismo epigenético nestas espécies cactofílicas.

Descobertas relataram a presença da concentração de 1,5 mg de vitamina C por 100 g⁻¹, em peso seco do fruto, no substrato de *C. hildmannianus* (Pereira *et al.* 2012), e 4,0 mg L⁻¹ de vitamina B5 presentes em células de tecidos vegetais, em *C. peruvianus* (Oliveira *et al.* 1995). Outras evidências sugerem o conteúdo de 7 a 22 mg de vitamina C por 100 g⁻¹, e compostos de vitamina B1 (0,14 mgL⁻¹), B2 (0,60 mgL⁻¹), e B3 (0,46 mgL⁻¹), em cladódios de cacto *Opuntia* sp. (Kamble *et al.* 2017). Os componentes químicos do cacto *Pilosocereus pachycladus* compreendem as substâncias de álcool alifático, esteroides, e fenólicas, mas ainda não está clara a concentração das vitaminas nos cladódio (Brito-Filho, 2015). No contexto químico, diferenças na composição de vitaminas B, e C podem ser um dos fatores que contribuem para as mudanças nos estados de metilação do DNA genômico em *Drosophila*, com o uso de diferentes micronutrientes presentes nos recursos ecológicos.

Uma perspectiva para futuros estudos seria a investigação de herança transgeracional (parental, e prole F2), com amostras de tecidos ovários e outras com testículos em adultos de *Drosophila gouveai*, quando as moscas se desenvolvem em diferentes meios de cultura. Os estudos poderiam avaliar a diferença na constituição de compostos químicos na metilação de DNA, que envolvem meios suplementados com alcaloides (I); vitamina B (II); vitamina C (III); vitamina B + vitamina C (IV); tecidos de cacto *C. hildmannianus* (V); tecidos de cacto *P. machrisii* (VI). O uso de tecnologias como *Chromatin Immunoprecipitation by Sequencing* (Chip-Seq) e *Whole-Genome Bisulfite Sequencing* (WGBS) pode contribuir para elucidar mudanças químicas na cromatina e/ou na metilação do genoma (Mehrmohamadi *et al.* 2021), quando os organismos se desenvolvem em diferentes ambientes/nichos ecológicos, em *Drosophila*.

6.5 Ecologia

Os dados indicam uma tendência para maior variação da metilação do DNA entre o ambiente oferecido por tecidos do cacto *C. hildmannianus* do que em *P. machrisii*. Coletas de amostras biológicas sugerem que o cacto colunar *P. machrisii* é o sítio criador principal das larvas de *D. gouveai*, e com menor frequência outros cactos hospedeiros (Tidon-Sklorz & Sene, 2001). Com base nesses dados, uma explicação seria o aumento da variação de metilação de DNA quando os organismos exploram outros recursos ecológicos, que não incluem o cacto hospedeiro *P. machrisii*.

Dados indicam que formigas vermelhas *Solenopsis invicta* (Santschi, 1916) apresentam alterações nos padrões de metilação do DNA, com relação aos organismos que são invasoras e as que são nativas à região geográfica (Schrey *et al.* 2016). Uma possibilidade levantada pelos autores é que as alterações na metilação do DNA podem estar relacionadas a rápidas mudanças no ambiente, em que ocorrem respostas fisiológicas aos recursos ecológicos.

Os animais, no processo de invasão aos novos ambientes ecológicos, em casos com baixa diversidade genética, têm apresentado plasticidade de fenótipos em organismos (Carneiro & Lyko, 2020). Um evento relatado é a variabilidade de metilação do DNA intrapopulacional no pardal invasor (*Passer domesticus*), com respostas rápidas às alterações em condições ambientais (Carneiro & Lyko, 2020). Isto pode implicar o ajuste do *fitness* dos organismos, no processo de adaptação, em diferentes ambientes.

Dados de Moraes & Sene (2007) indicam baixa diversidade genética em indivíduos das populações de *D. gouveai* que ocorrem em Altinópolis (São Paulo; Brasil), com evento de endogamia e isolamento populacional, com outras regiões geográficas (Minas Gerais, Goiás, e Bahia; Brasil). Em contraste, a plasticidade de fenótipo de viabilidade e tempo de desenvolvimento, larvas para adultos, foram rápidas com relação à mudança dos tecidos vegetais de *P. machrisis* para o cacto *C. hildmaniannus* (Soto *et al.* 2007). Os dados indicam que ocorre metilação diferencial do DNA quando os organismos se desenvolvem em tecidos vegetais de *P. machrisis* do que em *C. hildmaniannus*, e no meio padrão (controle), em pupas e adultos de *D. gouveai*. Essas mudanças de metilação de DNA pode significar um ajuste fisiológico no metabolismo de *D. gouveai* quando os organismos se desenvolvem em diferentes recursos ecológicos.

Conclusão

7. Conclusões

A hipótese formulada foi testada e os objetivos propostos foram todos cumpridos. Aceitou-se a hipótese de que a metilação diferencial de bases de citosinas no DNA genômico está relacionada a mudanças de ambientes hospedeiros, como *Cereus hildmmanianus* e *Pilosocereus machrisis*, no desenvolvimento biológico na espécie cactofílica *D. gouveai*.

Os dados obtidos no presente estudo permitem como principais conclusões:

- A metilação no DNA no organismo de pupas podem estar relacionadas com processos de desenvolvimento ontogenético, com múltiplos processos de diferenciação de tecidos e órgãos, que geram ruídos nas informações de metilação do genoma, sem suporte estatístico de variações de metilação com relação ao ambiente, em *D. gouveai*.

- Variações de metilação de bases de citosinas permitem sugerir processos de diferenciação sexo-específico, com maior evidência de metilação relacionado aos tecidos reprodutivos de moscas fêmeas do que machos.

- Mudanças no estado de metilação no DNA ocorrem quando os organismos se desenvolvem em compostos de meio padrão, em tecidos de cactos *C. hildmmanianus*, e no cacto *P. machrisis*. Os dados sugerem que ocorre um sutil aumento da metilação no DNA de indivíduos que usaram o meio de cacto *C. hildmmanianus*, que não ocorre usualmente na natureza, do que em *P. machrisis*, que é sugerido como sítio de desenvolvimento principal, na natureza, para *D. gouveai*.

- Identificaram-se regiões alvos de metilação no genoma, que compreendem sequências, tais como os elementos estruturais: éxons e íntrons de genes, regiões promotoras, sítios de início de transcrição, não codificadores de RNAs (ncRNAs), Regulador Não Traducional (UTR), e elemento transponível, quando os organismos usam diferentes nichos ecológicos.

- A metilação de DNA nas sequências do gene *Piezo* pode estar relacionado com a mudança no processo de reconhecimento de ovoposição, em moscas fêmeas, e no desenvolvimento biológico, nos cactos *C. hildmmanianus*, e no cacto *P. machrisis*, em *D. gouveai*.

- Variações na metilação de bases de citosinas nas sequências do DNA estão relacionadas com processos biológicos de apoptose celular, ncRNA, e estabilidade do genoma, no desenvolvimento biológico em moscas machos, com uso de recursos de tecidos dos cactos *C. hildmmanianus*, e no cacto *P. machrisis*, em *D. gouveai*.

- Sequências alvos de metilação no DNA podem estar relacionadas com vias de síntese de proteínas, e comunicação núcleo-citoplasma, quando os organismos usam os tecidos vegetais do cacto *P. machrisis*, em *D. gouveai*.

- Microcompostos presentes em tecidos de cactos, como alcaloides e vitaminas (B, e C) podem contribuir para diferenças nos estados de metilação. Embora novos estudos precisem elucidar esse processo nas interações ecológicas em *Drosophila*.

Referências Bibliográficas

8.Referências Bibliográficas

- Achwal, C. W.; Ganguly, P.; Chandra, H. S. (1984) Estimation of the amount of 5-methylcytosine in *Drosophila melanogaster* DNA by amplified ELISA and photoacoustic spectroscopy. **EMBO J.**, v.3, n.2, 263-266.
- Aguirell, S.; Lundström, J.; Masoud, A. (1969) Cactaceae alkaloids VII:alkaloids of *Echinocereus merkeri*. **J. Pharm. Sci.**, v.58, n.11, 1413-1414.
- Anreiter, I.; Kramer, J. M.; Sokolowski, M. B. (2017) Epigenetic mechanisms modulate differences in *Drosophila* foraging behavior. **Proc. Natl. Acad. Sci. U S A.**, v.114, n.47, 12518-12523.
- Baker, J. L.; Olsen, L. W.; Sørensen, T. I. (2007) Childhood body-mass index and the risk of coronary heart disease in adulthood. **N Engl J Med.**, v.357, n.23, 2329-2337.
- Baimai, V.; Sene, F. M.; Pereira, M. A. Q. R. (1983) Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* group. **Genetica**, v.67, 81-92.
- Barrios-Leal, D.; Rocha, J. N.; Manfrin, M. H. (2019) Genetics and distribution modeling: the demographic history of the cactophilic *Drosophila buzzatii* species cluster in open áreas of South America. **J. Hered.**, v. 110, 22-33.
- Becerra, J. X. (1997) Insects on plants: macroevolutionary chemical trends in host use. **Science**, 276, 253-256.
- Bernonville, T. D.; Maury, S.; Delaunay, A.; Daviaud, C.; Chaparro, C.; Tost, J.; O'Connor, S. E.; Courdavault, V. (2020) Developmental Methylation of the Medicinal Plant *Catharanthus roseus* Unravels the Tissue-Specific Control of the Monoterpene Indole Alkaloid Pathway by DNA Methylation. **Int J Mol Sci.**, v.21, n.17, 6028.
- Bird, A. P.; Wolffe, A. P. (1999) Methylation-induced repression--belts, braces, and chromatin. **Cell.**, v.99, n.5, 451-4.
- Boffelli, D.; Takayama, S.; Martin, D. I. (2014) Now you see it: genome methylation makes a comeback in *Drosophila*. **Bioessays.**, v.36, n.12, 1138-1144.
- Bourhis, D.; Bourhis, D. L.; Patin, D.; Niveleau, A.; Comizzoli, P.; Renard, J. P.; Viegas-Péquignot, E. (2001) Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. **Curr Biol.**, v.11, n.19, 1542-1546.
- Brito-Filho, Severino Gonçalves De. (2015) **Estudo fitoquímico e atividade antimicrobiana de *Pilosocereus pachycladus* F. RITTER (CACTACEAE)**. Tese. Departamento de Produtos Naturais e Sintéticos Bioativos, Universidade Federal da Paraíba. 150p.

- Buescher, J. L.; Musselman, L. P.; Wilson, C. A.; Lang, T.; Keleher, M.; Baranski, T. J.; Duncan, J. G. (2013) Evidence for transgenerational metabolic programming in *Drosophila*. **Dis. Model Mech**, v.6, n.5, 1123–1132.
- Bulut, B.; Aydinli, Z.; Türktas-Erken, M. (2020) MSAP analysis reveals diverse epigenetic statuses in opium poppy varieties with different benzyisoquinoline alkaloid content. **Turk. J. Biol.**, v.44, 103-109.
- Bundgaard, J.; Barker, J. S. F. (2000) Remating, sperm transfer, and sperm displacement in the cactophilic species *Drosophila buzzatii* Patterson & Wheeler (Diptera: Drosophilidae). **Biol. J. Linn. Soc.**, v. 71, 145–164.
- Carneiro, V. C.; Lyko, F. (2020) Rapid Epigenetic Adaptation in Animals and Its Role in Invasiveness. **Integr Comp Biol.**, v.60, n.2, 267-274.
- Chittka, A.; Wurm, Y.; Chittka, L. (2012) Epigenetics: the making of ant castes. **Curr Biol.**, v.22, n.19, 835-838.
- Coleman, J. M.; Benowitz, K. M.; Jost, A. G.; Matzkin, L. M. (2018) Behavioral evolution accompanying host shifts in cactophilic *Drosophila* larvae. **Ecol. and Evol.**, v. 8, 6921-6931.
- Crofton, A. E.; Cartwright, E. L.; Feitzinger, A. A.; Lott, S. E. (2018) Effect of Larval Nutrition on Maternal mRNA Contribution to the *Drosophila* Egg. **G3**, v.8, n.6, 1933-1941.
- Davet, Aline. (2005) **Estudo fitoquímico e biológico do cacto – *Cereus jamacaru* De Candolle, Cactaceae.** Dissertação. Departamento de Ciências Farmacêuticas, Universidade Federal do Paraná. 100p.
- Dean, W.; Santos, F.; Stojkovic, M.; Zakhartchenko, V.; Walter, J.; Wolf, E.; Reik, W. (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. **Proc Natl Acad Sci U S A.**, v.98, n.24, 13734-1378.
- DeBrito, R. A.; Manfrin, M. H.; Sene, F. M. (2002) Nested cladistic analysis of Brazilian populations of *Drosophila serido*. **Mol Phylogenet Evol.**, v. 22, n.1, 131-43.
- DeLaval, K.; Wagschal, A.; Feil, R. (2006) Epigenetic deregulation of imprinting in congenital diseases of aberrant growth. **Bioessays.**, v.28, n.5, 453-459.
- Deobagkar, D.; Deshpande, A.; Chatterjee, S.; Kelkar, A. (2004) CpC methylation is present in *Drosophila melanogaster* and undergoes changes during its life cycle. **Drosoph Inf Serv**, v.87, n.7, 1-4.
- Delprat, A.; Negre, B.; Puig, M.; Ruiz, A. (2009) The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. **PLoS One.**, v.4, n.11, e7883.

- Dennis, G. Jr.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. **Genome Biol.**, v.4, n.5, 1-11.
- Deshmukh, S.; Ponnaluri, V. C.; Dai, N.; Pradhan, S.; Deobagkar, D. (2018) Levels of DNA cytosine methylation in the *Drosophila* genome. **PeerJ.**, v.6, e5119.
- Drewell, R. A.; Bush, E. C.; Remnant, E. J.; Wong, G. T.; Beeler, S. M.; Stringham, J. L.; Lim, J.; Oldroyd, B. P. (2014) The dynamic DNA methylation cycle from egg to sperm in the honey bee *Apis mellifera*. **Development.**, v.141, n.13, 2702-2711.
- Dutra, J. C. V.; Ferreira, J. M.; Pereira, P. R. C.; Oliveira, J. B. H., Gervásio, S. V.; Xavier, M. B.; Mota, M. M.; Luz, A. C. D.; Pretti, I. R.; França, H. S.; Jamal, C. M.; Batitucci, M. D. C. P. (2018) *Cereus jamacaru* D.C. Hydroalcoholic Extract Promotes Anti-Cytotoxic and Antitumor Activity. **Pharmaceuticals**, v.11, n.4, 130-148.
- Emborski, C.; Mikheyev, A. S. (2019) Ancestral diet transgenerationally influences offspring in a parent-of-origin and sex-specific manner. **Philos Trans R Soc Lond B Biol Sci.**, v.374, n.1768, 20180181.
- EWING, A. W. (1983) Functional aspects of *Drosophila* courtship. **Biology Review**, v.58, 275-292.
- Ewing, B.; Green, P. (1998) Base-Calling of automated sequencer traces using *Phred*. II. Error probabilities. **Genome Res.**, n.8, 186-194.
- Fanara, J. J.; Fontdevila, A.; Hasson, E. (1999) Oviposition preference, viability, developmental time and body size in the cactophilic sibling species *Drosophila koepferae* and *D. buzzatii* in association to their natural hosts. **Evol. Ecol.**, v.13, 173-190.
- Fitz-James, M. H.; Cavalli, G. (2022) Molecular mechanisms of transgenerational epigenetic inheritance. **Nat Rev Genet.**, v. 23, n.6, 325-341.
- Fogleman, J. C.; Danielson, P. B. 2001. Chemical Interactions in the Cactus-Microorganism-*Drosophila* model system of the Sonoran Desert. **Amer. Zool.**, v.41, 877-889.
- Fontdevila, A.; Pla, C.; Hasson, E.; Wasserman, M.; Sanchez, A.; Naveira, H.; Ruiz, A. (1988) *Drosophila koepferae*: a New Member of the *Drosophila serido* (Diptera: Drosophilidae) Superspecies Taxon, **Ann. Entomol. Soc. Am.**, v. 81, 380–385.
- Fox, R. J.; Donelson, J. M.; Schunter, C.; Ravasi T.; Gaitán-Espitia, J. D. (2019) Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. **Phil. Trans. R. Soc. B**, n. 374, 20180174.
- Fulneček, J.; Kovařík, A. (2014) How to interpret methylation sensitive amplified polymorphism (MSAP) profiles? **BMC Genet.**, v.15, n.2, 1-9.

- Franco, F. F.; Silva-Bernardi, E.; Sene, F. M.; Hasson, E.; Manfrin, M. H. (2010) Intra- and interspecific divergence in the nuclear sequences of the clock gene period in species of the *Drosophila buzzatii* cluster. **J. Zool. Syst. Evol. Res.**, v. 48, 322-331.
- Franco, Fernando de Faria (2009) **História evolutiva do “cluster” *Drosophila buzzatii* (grupo *D. repleta*): eventos históricos e diversificação de espécies no Brasil**. Tese. Departamento de Genética, Universidade de São Paulo. 197p.
- Garcia, R. N.; D'Ávila, M. F.; Robe, L. J.; Loreto, E. L. Da S.; Panzera, Y.; Heredia, F. O. De; Valente, V. L. Da S. (2007) First evidence of methylation in the genome of *Drosophila willistoni*. **Genetica**, v.131, 91-105.
- Glastad, K. M.; Hunt, B. G.; Goodisman, M. A. D. (2019) Epigenetics in Insects: Genome Regulation and the Generation of Phenotypic Diversity. **Annu Rev Entomol.**, v.64, 185-203.
- Glastad, K.; Gokhale, K.; Liebig, J.; Goodisman, M. A. D. (2016) The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. **Sci Rep**, v.6, 37110.
- Gowher, H.; Leismann, O.; Jeltsch, A. (2000) DNA of *Drosophila melanogaster* contains 5-methylcytosine. **EMBO J.**, v.19, n.24, 6918-6923.
- Goya, L. F.; Hurtado, J.; Hasson, E. (2019) Female remating rate and pattern of sperm use suggest intense sperm competition in *Drosophila antonietae* (Diptera: Drosophilidae). **Evol Ecol**, v.33, 637–650.
- Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symp. Ser.**, v.41, 95-98.
- Hara, M.; Kitamura, T.; Fukui, H.; Tabata, M. (1993) Induction of berberine biosynthesis by cytokinins in *Thalictrum minus* cell suspension cultures. **Plant Cell Rep.**, v.12, n.2, 70-73.
- Hasson, E.; DePanis, D.; Hurtado, J.; Mensch, J. (2019) Host Plant Adaptation in Cactophilic Species of the *Drosophila buzzatii* Cluster: Fitness and Transcriptomics. **J. Hered.**, v.110, n.1, 46-57.
- Hasson, E.; Naveira, H.; Fontdevila, A. (1992) The breeding sites of Argentinian cactophilic species of the *Drosophila mulleri* complex (subgenus *Drosophila-repleta* group). **Rev. Chil. Hist. Nat.**, v.65, 319-326.
- Hörberg, J.; Reymer, A. (2018) A sequence environment modulates the impact of methylation on the torsional rigidity of DNA. **Chem. Commun.**, v.54, n.84, 11885-11888.
- Huang, D. W.; Sherman, B. T.; Tan, Q.; Kir, J.; Liu, D.; Bryant, D.; Guo, Y.; Stephens, R.; Baseler, M. W.; Clifford, H.; Lempicki, R. A. (2007) The DAVID Gene Functional

- Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. **Nucleic Acids Res.**, v.35,169-175.
- Hurtado, J.; Iglesias, P. P.; Lipko, P.; Hasson, E. (2013) Multiple paternity and sperm competition in the sibling species *Drosophila buzzatii* and *Drosophila koepferae*. **Mol Ecol.**, v.22, n.19, 5016-5026.
- Jones, P. A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. **Nat Rev Genet.**, v.13, n.7, 484-492.
- Kamble, S. M.; Debaje, P.; Ranveer, R. C.; Sahoo, A. K. (2017). Nutritional Importance of Cactus: A Review. **Biosci.**, v. 10, n.37, 7668-7677.
- Ketterson, E. D.; Atwell, J. W.; McGlothlin, J. W. (2014) **Evolution of Hormones and Behaviors**. In: Losos, J. B. The Princeton guide to Evolution. Princeton University, New Jersey, 639-646.
- Kircher, H. W., Heed, W. B., Russell, J. S., Grove, J. (1967) Senita Cactus Alkaloids: Their Significance to sonorant desert *Drosophila* ecology. **J. Insect Physiol.**, v.13, 1869-1874.
- Knowles, L. L.; Markow, T. A. (2001) Sexually antagonistic coevolution of a postmating-prezygotic reproductive character in desert *Drosophila*. **Proc Natl Acad Sci U S A**, v.98, n.15, 8692-8696.
- Kramer, J. M.; Kochinke, K.; Oortveld, M. A. W.; Marks, H.; Kramer, D.; Jong, E. K. De; Asztalos, Z.; Westwood, J. T.; Stunnenberg, H. G.; Sokolowski, M. B.; Keleman, K.; Zhou, H.; Bokhoven, H. V.; Schenck, A. (2011) Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. **Plos Biology**, v.9, 1-15.
- Krauss, V.; Reuter, G. (2011) DNA methylation in *Drosophila*--a critical evaluation. **Prog Mol Biol Transl Sci.**, v.101, 177-191.
- Lemos, B.; Branco, A. T.; Jiang, P. P.; Hartl, D. L.; Meiklejohn, C. D. (2014) Genome-wide gene expression effects of sex chromosome imprinting in *Drosophila*. **G3**, v.4, n.1, 1-10.
- Lian, T.; Gaur, U.; Wu, Q. I.; Tu, J.; Sun, B.; Yang, D.; Fan, X.; Mao, X.; Yang, M. (2018) DNA methylation is not involved in dietary restriction induced lifespan extension in adult *Drosophila*. **Genet. Res.**, v.100, 1-6.
- Ling, J. Q.; Hoffman, A. R. (2007) Epigenetics of long-range chromatin interactions. **Pediatr Res.**, v.61, n.5, 11R-16R.
- Lovino, N. (2014) *Drosophila* epigenome reorganization during oocyte differentiation and early embryogenesis. **Brief Funct Genomics**, v.13, n.3, 246-253.

- Lyko, F.; Ramsahoye, B. H.; Jaenisch, R. (2000) DNA methylation in *Drosophila melanogaster*. **Nature**, v.408, n.6812, 538-540.
- Lyko, F.; Ramsahoye, B. H.; Kashevsky, H.; Tudor, M.; Mastrangelo, M. A.; Orr-Weaver, T. L.; Jaenisch, R. (1999) Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in *Drosophila*. **Nat Genet.**, v.23, n.3, 363-366.
- Machado, L. P. B.; Castro, J. P.; Madi-Ravazzi, L. (2002) Evaluation of the courtship and of the hybrid male sterility among *Drosophila buzzatii* cluster species (Diptera, Drosophilidae). **Braz. J. Biol.**, v.62, n.4, 601-608.
- Machado, L. P.; Madi-Ravazzi, L.; Tadi, W. J. (2006) Reproductive relationships and degree of synapsis in the polytene chromosomes of the *Drosophila buzzatii* species cluster. **Braz. J. Biol.**, v.66, n.4, 279-293.
- Madi-Ravazzi, L.; Bicudo, H. E.; Manzato, J. A. (1997) Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. **Cytobios**, v.89, 21-30.
- Mahajan, A.; Sapehia, D.; Thakur, S.; Mohanraj, P. S.; Bagga, R.; Kaur, J. (2019) Effect of imbalance in folate and vitamin B12 in maternal/parental diet on global methylation and regulatory miRNAs. **Sci Rep.**, v.9, n.1, 17602.
- Maleszka, R. (2016) Epigenetic code and insect behavioural plasticity. **Curr Opin Insect Sci.**, v.15, 45-52.
- Manfrin, M. H.; DeBrito, R. O. A.; Sene, F. M. (2001) Systematics and Evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) Cluster Using mtDNA. **Ann. Entomol. Soc. Am.**, v.94, 333-346.
- Manfrin, M. H.; Sene, F. M. (2006) Cactophilic *Drosophila* in South America: A model for evolutionary studies. **Genetica**, v.126, 57-75.
- Markow, T. A. (1996) Evolution of *Drosophila* mating systems. **Evol. Biol.**, v.29, 73-81.
- Markow, T. A. (2015) The secret lives of *Drosophila* flies. **Elife**, v.4, e06793.
- Markow, T. A. (2019) Host use and host shifts in *Drosophila*. **Curr Opin Insect Sci.**, v.31, 139-145.
- Markow, T. A.; O'Grady, P. (2008), Reproductive ecology of *Drosophila*. **Funct. Ecol.**, v.22, 747-759.
- Markow, T. A.; O'Grady, P. M. (2005) Evolutionary genetics of reproductive behavior in *Drosophila*: connecting the dots. **Annu Rev Genet.**, v.39, 263-291.
- Matzkin, L. M., Johnson, S., Paight, C., Markow, T. A. (2013) Preadult parental diet affects offspring development and metabolism in *Drosophila melanogaster*. **PLoS One**, v.8, n.3, e59530.

- Matzkin, L.M. (2014) Ecological Genomics of host shifts in *Drosophila mojavensis*. **Adv. Exp. Med. Biol.**, v.781, 233-247.
- Mehrmohamadi, M.; Sepehri, M. H.; Nazer, N.; Norouzi, M. R. (2021) A Comparative Overview of Epigenomic Profiling Methods. **Front Cell Dev Biol.**, v.9, 714687.
- Monk, D.; Bentley, L.; Hitchins, M.; Myler, R. A.; Clayton-Smith, J.; Ismail, S.; Price, S. M.; Preece, M. A.; Stanier, P.; Moore, G. E. (2002) Chromosome 7p disruptions in Silver Russell syndrome: delineating an imprinted candidate gene region. **Hum Genet.**, v.111, 376-87.
- Moraes, E. M.; Sene, F. M. (2002) Breeding structure of an isolated cactophilic *Drosophila* population on a sandstone table hill. **J. Zool. Syst. Evol. Research.**, v.40, 123-128.
- Moraes, E. M.; Sene, F. M. (2007), Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. **J. Zool. Syst. Evol. Res.**, v.45, 336-344.
- Moraes, E. M.; Yotoko, K. S. C.; Manfrin, M. H.; Solferini, V. N.; Sene, F. M. (2009), Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. **J. Biogeogr.**, v.36, 2136-2147.
- Moss, T. J.; Wallrath, L. L. (2007) Connections between epigenetic gene silencing and human disease. **Mutat Res.**, v.618, 163-174.
- Nan, X.; Ng, H. H.; Johnson, C. A.; Laherty, C. D.; Turner, B. M.; Eisenman, R. N.; Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. **Nature.**, v.28, n.393, 386-389.
- O`Grady, P. M.; Markow, T. A. (2012) **Rapid morphological, behavioral, and ecological evolution in *Drosophila*: comparisons between the endemic Hawaiian *Drosophila* and the cactophilic replete species group.** In: Singh, Rama S.; Xu, Jianping, Kulathinal, Rob J. *Rapidly Evolving genes and genetic systems*. United Kingdom: Oxford, 176-186.
- Olerup, O.; Zetterquist, H. (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. **Tissue Antigens.**, v.39, n.5, 225-235.
- Oliveira, C. C.; Manfrin, M. H.; Sene, F. M.; Etges, W. J. (2013) **Evolution of males courtship songs in the *Drosophila buzzatii* species cluster.** In: Michalak, P. *Speciation: Natural Processes, Genetics and Biodiversity*. Nova Biomedical, New York, 137-164.

- Oliveira, D. C. S. G.; Almeida, F. C.; O'Graddy, P.; Armella, M. A.; Desalle, R.; Etges, W. J. (2012) Monophyly, divergence times, and evolution of host plant use inferred from a revised phylogeny of the *Drosophila repleta* species group. **Mol. Phylogenet. Evol.**, v.64, n.3, 533-544.
- Oliveira, S. A.; Machado, M. F. P. S.; Prioli, A. J.; Mangolin, C. A. (1995) In vitro propagation of *Cereus peruvianus* mill. (cactaceae). **In Vitro Cell Dev Biol – Plant**, v.31, 47–50.
- Panikar, C. S.; Rajpathak, S. N.; Abhyankar, V.; Deshmukh, S.; Deobagkar, D. D. (2015) Presence of DNA methyltransferase activity and CpC methylation in *Drosophila melanogaster*. **Mol Biol Rep.**, v.42, n.12, 1615-1621.
- Panis, D. N. De; Padró, J.; Furió-Tarí, P.; Tarazona, S.; Carmona, P. S. M.; Soto, I. M.; Dopazo, H.; Conesa, A.; Hasson, E. (2016) Transcriptome modulation during host shift is driven by secondary metabolites in desert *Drosophila*. **Mol. Ecol.**, v.25, 4534-4550.
- Patterson, J. T.; Wheeler, M. R. (1942). Description of new species of the subgenera *Hirtodrosophila* and *Drosophila*. **Univ. Tex. Publs.** 4213: 67–109.
- Pereira, M. A. Q. R.; Vilela, C. R.; Sene, F. M. (1983) Notes on breeding and feeding sites of some species of the *repleta* group of the genus *Drosophila* (Diptera, Drosophilidae). **Cien. Cult.**, v.35, n.1, 1313-1319.
- Pereira, M. C.; Jablonski, A.; Rios, A.; Vizzotto, M.; Flôres, S. H. (2012) **Perfil de carotenoides e teor de vitamina C em frutos Brasileiros**. In: Simpósio internacional de plantas medicinais e nutracêuticos. Embrapa, Conferência do Instituto Nacional de Ciências e Tecnologia de frutos tropicais, Universidade Federal do Sergipe, ALICE., v.3, 1.
- Pérez-Figueroa, A. (2013) msap: a tool for the statistical analysis of methylation-sensitive amplified polymorphism data. **Mol Ecol Resour.**, v.13, n.3, 522-527.
- Pummangura, S.; Nichols, D. E.; McLaughlin, J. L. (1977) Cactus alkaloids XXXIII:β-Phenethylamines from the Guatemalan cactus *Pilosocereus maxonii*. **J. Pharm. Sci.**, v.66, n.10, 1485-1487.
- Reese, M. G. (2001) Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. **Comput Chem**, v. 26, n.1, 51-6.
- Richards, E. J. (2006) Inherited epigenetic variation – revisiting soft inheritance. **Nat. Rev. Genet.**, v.7, n.5, 395-401
- Ruiz, A.; Wasserman, M. (1993) Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. **Heredity**, v.70, n.1, 582-596.
- Russell, A. (1954) A syndrome of intra-uterine dwarfism recognizable at birth with cranio-facial dysostosis, disproportionately short arms, and other anomalies (5 examples). **Proc R Soc Med.**, v.47, n.12, 1040-1044.

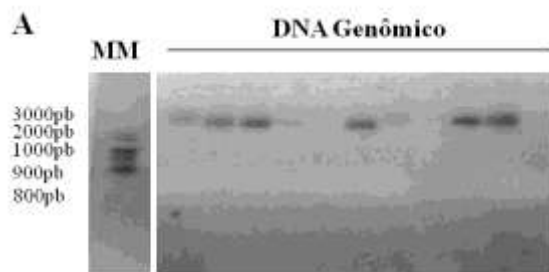
- Sanguinetti, C. J.; Dias, N. E.; Simpson, A. J. (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. **Biotechniques.**, v.17, n.5, 914-921.
- Santos, F.; Hendrich, B.; Reik, W.; Dean, W. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. **Dev Biol.**, v.241, n.1, 172-182.
- Santos, M. H.; Hasson, E.; Sene, F. M.; Manfrin, M. H. (2009) Esterasa-5 gene sequences from the *Drosophila buzzatii* cluster. **Dros. Inf. Serv.**, v.92, n.1,44-56.
- Shikama, N.; Ackermann, R.; Brack, C. (1994) Protein synthesis elongation factor *EF-1 alpha* expression and longevity in *Drosophila melanogaster*. **Proc Natl Acad Sci U S A**, v.91, n.10, 4199-203.
- Schlegel, A.; Stainier, D. Y. (2007) Lessons from "lower" organisms: what worms, flies, and zebrafish can teach us about human energy metabolism. **PLoS Genet.**, v.3, n.11, e199.
- Schmitz, R. J.; Lewis, Z. A.; Goll, M. G. (2019) DNA Methylation: Shared and Divergent Features across Eukaryotes. **Trends Genet.**, v.35, n.11, 818-827.
- Schrey, A. W.; Robbins, T. R.; Lee, J.; Dukes, D. W. Jr.; Ragsdale, A. K.; Thawley, C. J.; Langkilde, T. (2016) Epigenetic response to environmental change: DNA methylation varies with invasion status. **Environ Epigenet.**, v.2, n.2, 1-5.
- Silva, A. F. G.; Sene, F. M. (1991) Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). **Rev. Bras. Entomol.**, n.55, 455-468.
- Silver, H. K.; Kiyasu, W.; George, J.; Deamer, W. C. (1953) Syndrome of congenital hemihypertrophy, shortness of stature, and elevated urinary gonadotropins. **Pediatrics.**, v.12, n.4,368-376.
- Song, N.; Zhu, Y.; Cui, Y.; Lv, M.; Tang, Y.; Cui, Z.; Dang, G.; Zheng, H.; Liu, S. (2020) Vitamin B and Vitamin C Affect DNA Methylation and Amino Acid Metabolism in *Mycobacterium bovis* BCG. **Front Microbiol.**, v.11, 812.
- Soto, I. M.; Carreira, V. P.; Corio, C.; Padró, J.; Soto, E. M.; Hasson, E. (2014) Differences in tolerance to host cactus alkaloids in *Drosophila koepferae* and *D. buzzatii*. **Plos One**, v.09, 1-9.
- Soto, I. M.; Hasson, E. R.; Manfrin, M. H. (2008) Wing morphology is related to host plants in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera, Drosophilidae). **Biol. J. Linn. Soc.**, n.95, 655-665.
- Soto, I. M.; Manfrin, M. H.; Sene, F. M.; Hasson, E. (2007) Viability and developmental time in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera: Drosophilidae) are dependent on the cactus host. **Ann. Entom. Soc. Am.**, v.100, n.04, 490-496.
- Soto, I. M.; Soto, E. M.; Corio, C.; Carreira, V. P.; Manfrin, M.; Hasson, E. (2010) Male genital and wing morphology in the cactophilic sibling species *Drosophila gouveai* and

- Drosophila antonietae* and their hybrids reared in different host plants. **Environ. Entomol.**, v.39, n.03, 865-873.
- Stearns, S. C. (2014) **Natural selection, adaptation, and fitness: overview**. In: Losos, J. B. The Princeton guide to Evolution. Princeton University, New Jersey, 208-214.
- Stintzing, F. C.; Carle, R. (2005) Cactus stems (*Opuntia* spp.): A review on their chemistry, technology, and uses. **Mol. Nutr. Food Res.**, n.49, 175-194.
- Thompson, J. D.; Higgins, D. G.; Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res.**, v.22, 4673-4680.
- Tidon-Sklorz, R.; Sene, F. M. (1995) *Drosophila seriema*: a new member of the *Drosophila serido* (Diptera, Drosophilidae) superspecies taxon. **Ann. Entomol. Soc. Am.**, v.88, n.2, 139-142.
- Tidon-Sklorz, R.; Sene, F. M. (2001) Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). **Iheringia (Zoologia)**, n.90, 141-146.
- Tyler, M. S. (2000) **Developmental Biology: a guide for experimental study**. 2^a ed. Elsevier Science: Sunderland. 199.
- Verona, R. I.; Mann, M. R.; Bartolomei, M. S. (2003) Genomic imprinting: intricacies of epigenetic regulation in clusters. **Annu Rev Cell Dev Biol.**; v.19, 237-259.
- Vieira, G. C.; D'Ávila, M. F.; Zanini, R.; Deprá, M.; Valente, V. L. S. (2018) Evolution of DNMT2 in drosophilids: Evidence for positive and purifying selection and insights into new protein (pathways) interactions. **Genet Mol Biol.**, v.41, 215-234.
- Vilela, C. R.; Sene, F. M. (1977) Two new tropical species of the repleta group of the genus *Drosophila* (Diptera, Drosophilidae). **Pap. Avulsos Zool.**, v.30, n.20, 295-299.
- Vrdoljak, J.; Padró, J.; DePanis, D.; Soto, I. M.; Carreira, V. P. (2019) Protein-alkaloid interaction in larval diet affects fitness in cactophilic *Drosophila* (Diptera: Drosophilidae). **Biol. J. Linn. Soc.**,v.127, 44-55.
- Waddington, C. H. (1956) Genetic Assimilation of the Bithorax Phenotype. **Evolution**, v.10, n.1, 1–13.
- Waddington, C. H. (2012) The Epigenotype, **Int. J. Epidemiol.**, v. 41, 10–13.
- Wang, X.; Werren, J. H.; Clark, A. G. (2016) Allele-Specific Transcriptome and Methylome Analysis Reveals Stable Inheritance and Cis-Regulation of DNA Methylation in *Nasonia*. **PLoS Biol.**, v.14, n.7, e1002500.
- Whitlock, M. C. (2014) **Evolutionary Processes**. In: Losos, J. B. The Princeton guide to Evolution. Princeton University, New Jersey, 320-321.

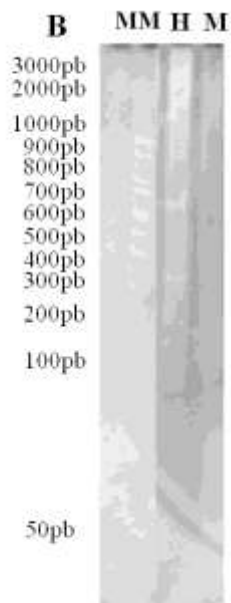
- Whitman, D.W.; Agrawal, A. A. (2009) **What is phenotypic plasticity and why is it important?** In: Whitman, D. W.; Ananthakrishnan, T. N. Phenotypic Plasticity of Insects: Mechanisms and Consequences. Science Publishers: Enfield, 1–63.
- Wu, X.; Galbraith, D. A.; Chatterjee, P.; Jeong, H.; Grozinger, C. M.; Yi, S. V. (2020) Lineage and Parent-of-Origin Effects in DNA Methylation of Honey Bees (*Apis mellifera*) Revealed by Reciprocal Crosses and Whole-Genome Bisulfite Sequencing. **Genome Biol Evol.**, V.12, n.8, 1482-1492.
- Yaish M. W.; Peng M.; Rothstein S. J. (2014) Global DNA Methylation Analysis Using Methyl-Sensitive Amplification Polymorphism (MSAP). In: Sanchez-Serrano J., Salinas J. (eds) **Arabidopsis Protocols**. Methods in Molecular Biology (Methods and Protocols), vol 1062. Humana Press, Totowa, NJ. 720.
- Zhang, Y.; Wolosker, M. B.; Zhao, Y.; Ren, H.; Lemos, B. (2020) Exposure to microplastics cause gut damage, locomotor dysfunction, epigenetic silencing, and aggravate cadmium (Cd) toxicity in *Drosophila*. **Sci. Total Environ.**, n.744, 140979.
- Zhou, X.; Chen, J.; Liang, M. S.; Wang, F. (2013) Differential DNA methylation between two wing phenotypes adults of *Sogatella fucifera*. **Genesis**, v.51, n.12, 819-826.

Apêndices e Anexos

Apêndice A. Verificação da integridade de moléculas de DNA genômico (gel agarose 0.8%) (A), e digestões enzimáticas (B) (gel poliacrilamida 10%), amostra de *D. gouveai*.

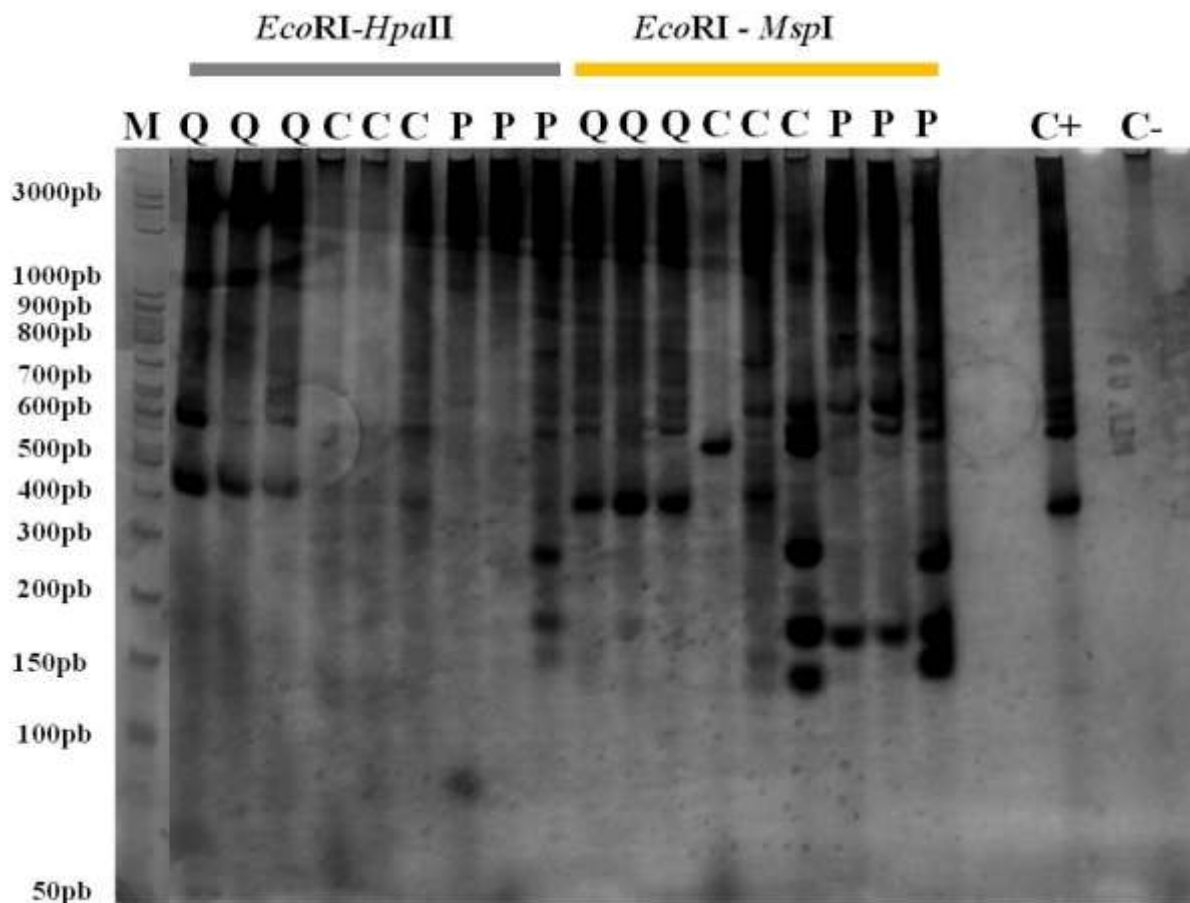


Integridade de moléculas de DNA genômico.
MM: Marcador de peso molecular.

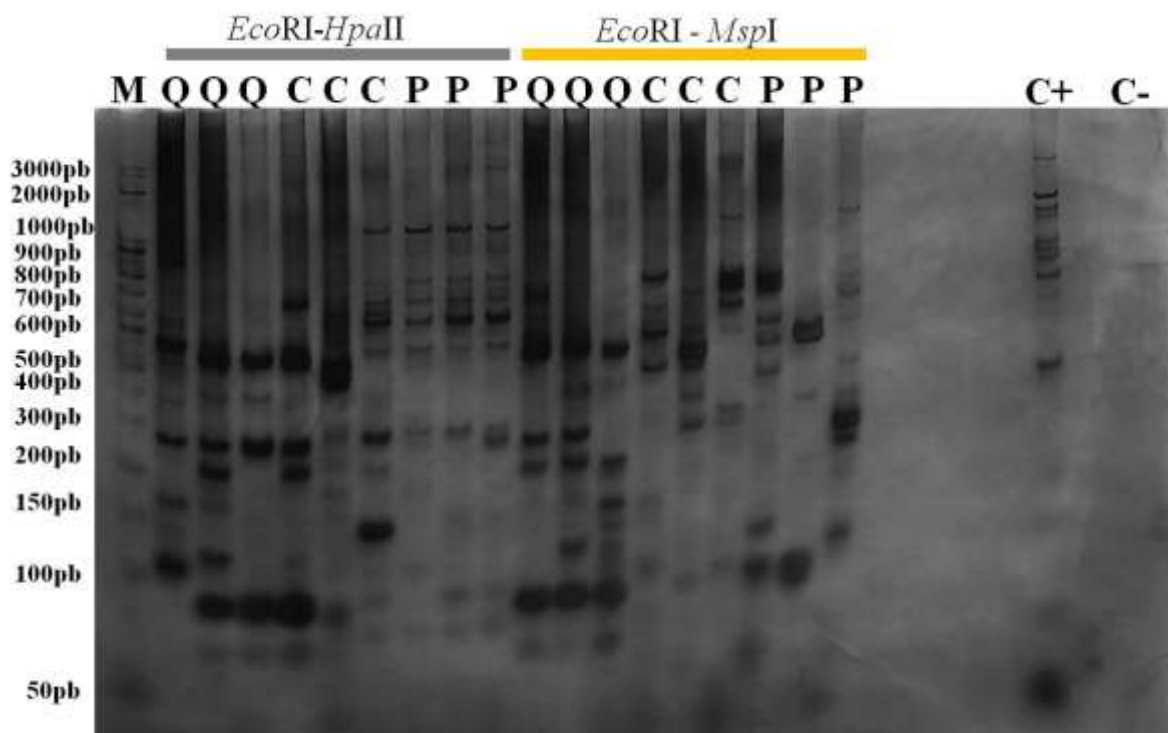


Digestões de DNA genômico com *Hpa*II (H) ou *Msp*I (M). MM Marcador de peso molecular.

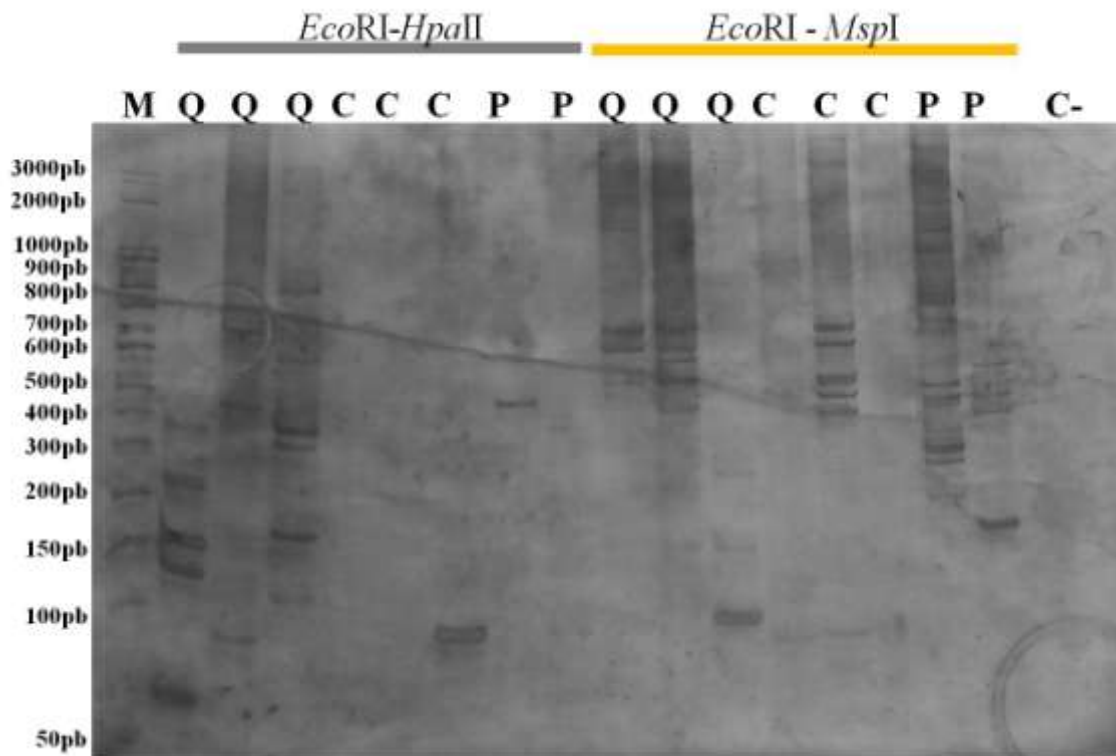
Apêndice B. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AAC)/HM(TAG), com DNA genômico de pupas em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.



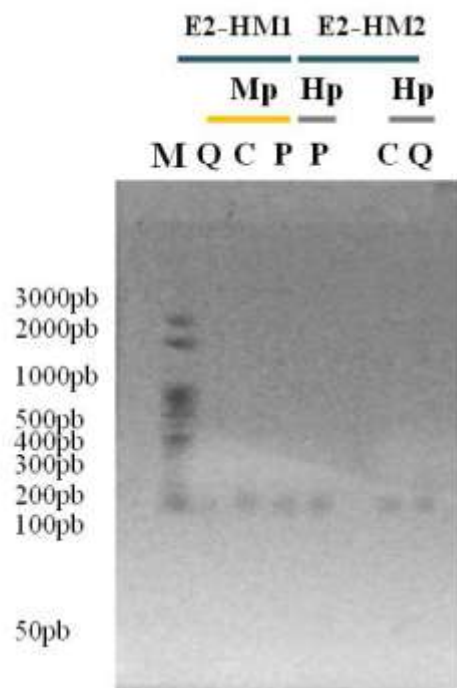
Apêndice C. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(AGG)/HM(TGA), com DNA genômico de adultas e fêmeas que se desenvolveram em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle positivo (C+), e controle negativo (C-) da PCR.



Apêndice D. Gel de poliacrilamida 10% com verificação de produto de MSAP, E(ACT)/HM(TTG), com DNA genômico de adultos e machos que se desenvolveram em meio padrão (Q), *C. hildmannianus* (C), e *P. machrisii* (P). Marcador de peso molecular 50pb (M). Controle negativo (C-) da PCR.



Apêndice E. Verificação de produtos de PCR (gel de agarose 1%), com re-amplificações amostras diferenciais de metilação de DNA, em tecidos reprodutivos de moscas em *D. gouveai*.



Regiões alvos de metilação, isoladas, e Re-amplificadas em PCR, com os primers E2-HM1 e E2-HM2, com MSAP. Amostras de tecidos reprodutivos de moscas fêmeas, obtidas em meio padrão de laboratório (Q), tecidos de *Cereus hidmannianus* (C), e outras em *Pilosocereus machrisis* (P). Ensaio enzimáticos com *MspI* (Mp), e *HpaII* (Hp). M: Marcador de peso molecular.

TC,b11,MSP,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0
TC,b12,HPA,0
TC,b12,MSP,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0
TC,b13,HPA,0
TC,b13,MSP,0
TC,b14,HPA,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0
TC,b14,MSP,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0
TC,b15,HPA,0
TC,b15,MSP,0
TC,b16,HPA,0
TC,b16,MSP,0,1,1,1
TC,b17,HPA,0
TC,b17,MSP,0
TP,c1,HPA,1,0
TP,c1,MSP,0
TP,c2,HPA,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0
TP,c2,MSP,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
TP,c3,HPA,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
TP,c3,MSP,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0
TP,c4,HPA,1,0,1,0,0
TP,c4,MSP,0,0,0,1,0,0,1,1,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0
TP,c5,HPA,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0
TP,c5,MSP,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0
TP,c6,HPA,0,1,1,0
TP,c6,MSP,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0
TP,c7,HPA,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0
TP,c7,MSP,0
TP,c8,HPA,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0
TP,c8,MSP,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0
TP,c9,HPA,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0
TP,c9,MSP,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,1,0,0
TP,c10,HPA,0,1,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,1,0
TP,c10,MSP,0,1,0
TP,c11,HPA,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0
TP,c11,MSP,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0
TP,c12,HPA,0
TP,c12,MSP,0
TP,c13,HPA,0
TP,c13,MSP,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0
TP,c14,HPA,0,1,0
TP,c14,MSP,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,1,0
TP,c15,HPA,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0
TP,c15,MSP,0
TP,c16,HPA,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0
TP,c16,MSP,0

Apêndice I. Predições de bioinformática para regiões promotoras em sequências de DNA.

Sequencia do gene *Piezo*

Promoter predictions for 1 eukaryotic sequence with score cutoff 0.80 (transcription start shown in larger font):

Promoter predictions for seq0 :

Start	End	Score	Promoter Sequence
159	209	0.93	AGTGCAGTGACTCAAACAGGGCCAGAGCAGCGGC a AGAC C GCAcAGCGC

Sequência do gene *eEF1alpha*

Promoter predictions for 1 eukaryotic sequence with score cutoff 0.80 (transcription start shown in larger font):

Promoter predictions for seq0 :

Start	End	Score	Promoter Sequence
114	164	0.85	TGTTTTGCCCTATATGAAAACCTGACGCGCTAGAGATGAC G CCAATATAC
146	196	0.87	GAGATGACGCCAATATACGCCGTCGGAGACGATGATAACC A CCTGATAAA
228	278	0.91	TGTAGTGACTCAAATAGGGCCAGAGGTGTCGACCAAAC T TTTTATTTTA

Sequência do ncRNA - LOC125776504

Promoter predictions for 1 eukaryotic sequence with score cutoff 0.80 (transcription start shown in larger font):

Promoter predictions for seq0 :

Start	End	Score	Promoter Sequence
109	159	0.86	TGATTAGCTATAAATATACTCCTTGACTGCGTACCAATTC A ACTAAGBAG

Sequência do gene *eEF1alpha*

Promoter predictions for 1 eukaryotic sequence with score cutoff 0.80 (transcription start shown in larger font):

Promoter predictions for seq0 :

Start	End	Score	Promoter Sequence
54	184	0.95	cAAA E CCAgCAAAAACCGGGGGGGGTTTTTCCAAGAb A ACATGETGA

Apêndice J. Estrutura quaternária da proteína DNA metiltransferase 2 (Dnmt2) em *D. buzzatii*. Esta sequência foi obtido do seu genoma sequenciado e disponível em <https://dbuz.uab.cat/welcome.php>, Dnmt2 - scaffold27:847239-848215, e a proteína foi modelada usando o I-TASSER Server disponível em <https://zhanglab.ccmb.med.umich.edu/I-TASSER>. A estrutura de Dnmt2 é composta por α -hélices (cor merge) e β -folhas (cor amarelo), e espaçadores (cores em branco e azul).





DECLARAÇÃO

O Projeto de Pesquisa intitulado “*Variabilidade epigenética: reprogramação e evolução durante o desenvolvimento de Drosophila buzzatii e Drosophila gouveai em diferentes cactos hospedeiros*”, sob a responsabilidade da **Professora Maura Helena Manfrin e da Professora Ester Silveira Ramos**, tem por objetivo contribuir com o conhecimento sobre os mecanismos de reprogramação epigenética do DNA genômico e as suas consequências na regulação das expressões gênicas e seus efeitos fenotípicos.

Por não satisfazer à condição de estudo experimental envolvendo animais pertencentes ao *filo Chordata* e *subfilo Vertebrado*, conforme diretriz do CONCEA - Lei nº 11.794 de 08/10/2008, declaramos que **NÃO COMPETE** à Comissão de Ética no Uso de Animais da Faculdade de Medicina de Ribeirão Preto – USP emitir parecer acerca dos aspectos éticos relacionados a projetos de pesquisa dessa natureza.

Ribeirão Preto, 04 de março de 2022


Prof. Dr. Luiz Carlos Carvalho Navegantes
Coordenador da CEUA-FMRP-USP

Anexo B – Manuscrito como primeiro autor, com os dados do projeto do doutorado, a ser submetido em revista científica.

Requisito regimental da Pós-graduação da Faculdade de Medicina de Ribeirão Preto-USP (CPG-FMRP-USP), que determina o encaminhamento junto com a Tese de ao menos um manuscrito derivado dos dados do trabalho. Desta forma, é apresentada a seguir a versão de um manuscrito que etende a exigência citada.

Differential DNA methylation contributes to sex-specific processes and changing ecological environments in the cactophilic species, *Drosophila gouveai* (cluster *Drosophila buzzatii*)

Adriano S. Santos¹, Ester S. Ramos¹, Maura H. Manfrin^{1,2}

¹Postgraduate, Department of Genetics, Faculty of Medicine of Ribeirão Preto, FMRP-USP, Ribeirão Preto, São Paulo, Brazil.

²Department of Biology, Faculty of Philosophy, Sciences and Letters of Ribeirão Preto, FFCLRP-USP, Ribeirão Preto, São Paulo, Brazil.

Autores de correspondência: adrianosantos@usp.br; esramos@usp.br; maura.manfrin@usp.br.

Abstract

Larval organisms in *Drosophila gouveai* develop obligatorily in cactus tissues, with evidence of phenotypic plasticity in viability, developmental time, and morphological structures when individuals change breeding sites in *P. machrisii* and *C. hildmannianus* cactus. One molecular mechanism discovered is that DNA methylation marks may contribute to genotype-phenotype regulation, although the processes in phenotypic plasticity have not been elucidated. This study analyzed the variation in DNA methylation states (CCGG-sites), with the sequencing of genomic regions, and the identification of the molecular pathways involved in the changes of creator sites, in the cactus *P. machrisis*, in *C. hildmmanianus*, and standard laboratory medium (control), in *D. gouveai*. We discovered differences in molecular patterns in pupae (324 regions), female flies (308 regions), and male flies (201 regions) in the Methylation-Sensitive Amplification Polymorphism (MSAP) assays. Variations in the ratios of methylated loci in genomic DNA were found in adult organisms that emerged in different environments: tissues of *P. machrisii* cactus (females=0.77%; males=0.88%), in *C. hildmannianus* (females= 0.71%; males=0.81%), and in the standard medium (females=0.56%; males=0.66%). Additionally, methylation target regions (CpC or CpG) involve íntrons sites, exons, non-coding RNA, promoter regions, and transposable elements in bioinformatics analyses with DNA sequences. Finally, these data are related to different biological processes, such as protein synthesis, and nucleus-cytoplasm communication. We conclude that changes in cytosine methylation states in DNA are related to molecular pathways that may contribute to changes in the ecological environments of *D. gouveai*.

Keywords: Cactophilic *Drosophila*. Phenotypic plasticity. DNA methylation. CpG-sites. sex-specific. Transposable element. Host cacti.

1.Introduction

The concept phenotypic plasticity is recognized as the ability of a single genotype to express multiple phenotypes in response to its upbringing in different environments (Whitman & Agrawal, 2009; Fox *et al.* 2019). These authors indicate that phenotypes can comprise different levels, such as morphological, physiological, and behavioral/social organization changes (Whitman & Agrawal, 2009; Fox *et al.* 2019), which can influence populations' adaptation processes (Stearns, 2014).

The *Drosophila*-cactus model comprises the species *Drosophila gouveai* Tidon-Sklorz & Sene (2001) (cluster *Drosophila buzzatii*, *Drosophila replete* group) and is characterized by edeago type B morphotype (Silva & Sene, 1991), chromosome inversion 2e8 (Ruiz & Wasserman, 1993; Manfrin *et al.* 2006), and type I metaphase plaque (Baimai *et al.* 1993). This species develops obligatorily in necrotic cactus tissues and reports that the cladodes of *Pilosocereus machrisis* are major created sites in nature (Pereira *et al.* 1983; Tidon-Sklorz & Sene, 2001). The distribution of *D. gouveai* populations has been reported in areas with the great diversity of cactus species of the genera *Pilosocereus* and *Cereus*, in northern and northeastern Brazil (Pereira, Vilela & Sene, 1983). For example, in the locality of Ibotirama (in Bahia state, Brazil), *D. gouveai* populations are in contact with different cacti, such as *P. machrisis* and *C. hildmaniannus* (Pereira *et al.* 1983; DeBrito *et al.* 2002). In contrast, individuals from populations of *D. gouveai* are associated with the cactus *Pilosocereus machrisis*, in the southeastern portion of the country, for example, in the locality of Altinópolis (state of São Paulo, Brazil) (Tidon-Sklorz & Sene, 2001; Moraes & Sene, 2002; Manfrin *et al.* 2006).

Interestingly, DNA microsatellite *loci* indicate three clusters in *D. gouveai* populations, concerning geographic regions and low genetic diversity in populations, which has suggested barriers to gene flow with inbreeding (Moraes *et al.* 2002). Additionally, data about reproductive isolation was found concerning another sister species in the *D. buzzatii* cluster (Machado *et al.* 2006).

Phenotypic plasticity was discovered in traits when larvae developed at different breeding sites in tissues of cactus species in *D. gouveai* (Soto *et al.* 2007, 2008, 2010). One set of phenotypes shows rapid responses to changes in ecological environments, such as development times (DT) of approximately 280 hours in *P. machrisis* and 295 hours in *C. hildmaniannus* (Soto *et al.* 2007); larval to adult viability is higher in *P. machrisis* tissues than in *C. hildmaniannus* (Soto *et al.* 2007); the structure of large wings in adults that developed in *C. hildmaniannus* compared to those that grew in *P.*

machrisis (Soto *et al.* 2008); the morphological structure of the male reproductive pair (edeago) is more prominent in adults that emerged in *P. machrisis* cactus tissues than in *C. hildmaniannus* (Soto *et al.* 2010). The life history of this species includes cycles with differences in development in cactus tissues, covers traits and aspects of survival, reproduction, morphology, and dispersal through the flight of these organisms in different ecological environments, and contributes to their fitness (Soto *et al.* 2007, 2008, 2010).

Information indicates that the plasticity of phenotypes, as discovered in insects (ants and bees), can be affected by changes in epigenetic states, such as DNA methylation. Still, the mechanisms in genotype-phenotype regulation in interaction with the environment are not yet elucidated. This epigenetic term is recognized by changes in mitotically or meiotically inherited gene regulation without changes in DNA sequence (Fitz-James & Cavalli, 2022). These authors indicate that DNA methylation, with the addition of a methyl group at carbon 5 of cytosine bases (CpG; cytosine bases followed by guanine) in DNA, without changes in nucleotide bases, in the genomes of several eukaryotes (Fitz-James & Cavalli, 2022).

The dynamics of genomic DNA methylation have been discovered concerning the development and sex-specific differentiation in *Drosophila melanogaster* and *Drosophila willistoni* (Lyko *et al.* 2000; Garcia *et al.* 2007), which involve regions of introns, exons, promoter regions, and retrotransposons (Kraus & Reuters, 2011). Promising DNA methylation results were obtained with enzymatic assays, with differences in the sensitivity of methylation of cytosine bases (*HpaII*, *MspI*) (CpC and CpG sites) in the genome of *Drosophila* species (Lyko *et al.* 2000; Deobagkar *et al.* 2004; Garcia *et al.* 2007; Panikar *et al.* 2015).

Methylation in genomic DNA involves a single DNA methyltransferase recognized as "Dnmt2 only", with the function of adding a methyl group to cytosines, but without the presence of other canonical DNA methyltransferases (DNMT1 and DNMT3), *Drosophila* (Kraus & Reuter, 2011). In contrast, the oxidation and removal of methyl groups in the DNA molecules are carried out through the Ten-Eleven Translocation (TETs) family of proteins in genome demethylation (Kraus & Reuter, 2011).

One proposed idea is that DNA methylation may contribute to the plasticity of phenotypes in insects, cell-or tissue-specific (Maleska, 2016), and colonizing different ecological environments (Schrey *et al.* 2016). This work analyzes the variation in methylation states in the genome and in its DNA sequences, as well as the identification

of functions in molecular pathways when organisms develop in different ecological environments (breeding sites), such as *P. machrisis* cacti, in *C. hildmmanianus*, and laboratory standard medium (control), in *D. gouveai*.

2. Material and methods

2.1 Drosophila stock

The biological samples come from individuals representing *Drosophila gouveai* (*Drosophila buzzatii* cluster, *Drosophila replete* group) (stock J79), with previous collections of naturally inseminated female flies in the locality of Ibotirama (State of Bahia, Brazil)(12.16°S, 43.04°W).

2.2 Experiments: heterogeneous environment

Flies were reared under identical conditions for 24 generations, with crosses between parents and offspring, in 35 ml of standard laboratory medium (corn flour/honey/*Sacharomices cerevise*/Agar/Nipagin; methyl 4-hydroxybenzoate), in *D. gouveai*. These crosses between flies result in increased relatedness between organisms and decreased genetic variation between individuals. We also collected pieces of *Pilosocereus machrisis* cacti, and others of *Cereus hildmmanianus*, in the wild. The laboratory stored these plant tissues in a freezer at -20°C.

Once maintained the *D. gouveai* stocks in the laboratory, then maintained fifty pairs of sexually mature adult flies (50 males and 50 females) in bottles with 35 ml of standard laboratory medium in three replications. Each bottle with flies was inspected after 24 hours, analyzing the presence of eggs or first-instar larvae. All experiments were performed under the same conditions with 24± 1°C and 12hs/12hs (light/dark) in the photoperiod equipment. We collected a set of 30 first-instar larvae in standard medium. We transferred them to bottles with different environments (breeding sites), which comprised standard laboratory medium, others in medium with *P. machrisis* tissues, and others in *C. hildmanmanianus*, with three replicates.

The preparation of cactus media was performed according to Fanara *et al.* (1999), with modifications. The cactus media consisted of 500g of cactus tissue, separately for each species, added 5g of biological yeast (*Sacharomyces cerevise*), 5g of Agar, 20ml of Nipagin and 800ml of H₂O to complete the volume of 1 l, and 35 ml of cactus medium was added to each bottle in the experiment.

Finally, the adult flies that emerged in each culture medium were transferred to other bottles with the same experimental conditions. We collected biological samples in

the pupal (± 215 hs) and adult fly (± 269 hs) stages in the second generation of organisms reared in standard laboratory medium, in *P. machrisis* tissues, and *C. hildmmanianus*. Pupae were collected, and reproductive tissues (females and others in males) were dissected with liquid nitrogen and stored in a -80°C freezer.

2.3 Isolation of genomic DNA

Genomic DNA was isolated from pupae and adult fly samples (female and other male reproductive tissues), with a *pool* $n=5$ to 10, according to modifications in the protocol of extraction and precipitation of DNA with NaCl (Olerup & Zetterquist, 1992). Genomic DNA was precipitated with ethanol 100% in the presence of NaCl 5M. The quantification of the genomic DNA samples was performed using the Nanodrop 2000c spectrophotometer (Thermo Scientific, Vilnius, Lithuania), and the integrity through the mobility of the molecules in 0.8% agarose gel.

2.4 MSAP assay

Methylation-sensitive amplification polymorphism (MSAP) analyzes were performed according to the protocol described by Zhou *et al.* (2013). Briefly, 300 ng genomic DNA for each sample, in two independent reactions, was digested with enzymes *EcoRI* (5U) ($\text{G}\downarrow\text{AATTC}$ -sites) and with *HpaII* (5U) or *MspI* (5U) ($\text{C}\downarrow\text{CGG}$ -sites) (Thermo Scientific); in a total volume of 20 μL at 37°C for 16 hrs, and inactivated at 80°C for 20 min. The digested samples *EcoRI/HpaII* and *EcoRI/MspI* (5 μl) were ligated to 2 pmol of *EcoRI* adapter and 20 pmol of *MspI/HpaII* adapter with 5U T4 DNA ligase (Thermo Scientific) at 16°C for 14 hours and inactivated at 65°C for 30 min. Pre-selective amplification was performed using 3 μL of binding products (1:10), 2pmol E0 and 2pmol HM0, in a final volume of 20 μL containing 2 μL of PCR buffer, 1.6 μL of MgCl_2 (25 mM), 0.5 μL dNTPs (2.5 mM) and 1 U Taq polymerase (Promega Corporation). PCR reactions were performed using the following program: 94°C for 1 min, 25 cycles at 94°C for 30 sec, 65°C for 30 sec and 72°C for 1 min, and a final extension at 72°C for 10 min. The sequences of *primers* for adapters, pre-amplification, and selective amplifications were obtained as described by Zhou *et al.* (2013) (see Table 1).

Selective amplification was performed using a pre-amplification product, combinations of six *EcoRI* primers (E1-E6) and six *HpaII/MspI* primers (HM1 - HM6), as listed in Table 1. The selective PCR reactions were performed in volumes of 20 μL containing 3 μL of the pre-amplified products, 2pmol of *EcoRI* selective primer (E1 to

E6), and 2pmol of *HpaII/MspI* selective primer (HM1 to HM6), which had two additional selective nucleotide bases compared to pre-amplified primers. PCR reaction cycle: 94°C for 1 min, 13 cycles at 94°C for 30 sec, 65°C for 30 sec (reducing 0.7° C per cycle), and 72°C for 1 min; and another 23 cycles of PCR amplification after the program: 94°C for 30sec, 56°C for 30 sec, 72°C for 1 min, and a final extension at 72°C for 10 min. The selective PCR products (amplicons) were analyzed and isolated on a 10% polyacrylamide gel (PAGE). Electrophoresis was performed at a constant power of 200 V and 400 mA for 1.5 h. The gels were stained with silver nitrate (Bassam *et al.* 1991) and digitalized.

2.5 Analysis of methylation profiles

The DNA genomic, molecular pattern was obtained with the isoschizomers *HpaII* and *MspI*, with differences in the sensitivity of the enzymes in the methylation of cytosine bases. Enzyme *HpaII* is sensitive to internal methylation of cytosine bases, and *MspI* is sensitive to external methylation of cytosines. The amplicons obtained with MSAP assays with clear bands were analyzed with molecular extension from 100bp to 1000bp. The bands were converted into a binary matrix, where the presence of a (+) band corresponds to “1” and its absence to (-) “0”.

The interpretation of DNA methylation data through the *EcoRI-HpaII* and *EcoRI-MspI* assays, and methylation profiles through MSAP, were classified into the following categories: (Type I) bands in the *HpaII* (+) and *MspI* (+) lanes, absence of methylation, or external to cytosine bases; (Type II) bands present in *HpaII* (+) lanes, but absent in *MspI* (-), were considered hemimethylated, with external methylation of cytosines; (Type III) bands visible in *MspI* (+), but absent in *HpaII* (-), fully methylated internal cytosines, with both strands of DNA; (Type IV and Type V) absent in both *HpaII* (-) and *MspI* (-) bands, without DNA methylation information (Table 1. SuplMaterial), according to Zhou *et al.* (2013) and Fulnecek & Kovarik (2014). Changes in DNA methylation states were evaluated in *Drosophila* samples raised in different environments (standard medium, *P. machrisii* cacti, and another in *C. hildmannianus*) using MSAP data with the ratios: (I) number of bands with external methylation of cytosines/total number of bands (external methylation ratio); (II) total number of bands with internal methylation of cytosines/ number of total bands (internal methylation ratio); (III) external methylation ratio + internal methylation ratio (Relative Total Methylation Ratio).

The binary matrices were evaluated using the ‘*msap*’, package R (Pérez-Figueroa 2013). Data were analyzed as methylation susceptible *loci* (MSL) and unmethylated *loci* (NML), in DNA samples (different developmental stages) obtained in the host environment, and Shannon Diversity Index was estimated. Finally, was conducted Analysis of Variance (*One-way* ANOVA) and Tukey test to evaluate the significance of the variation in methylation states concerning the environment, using the package in the R program. The scores were statistically significant, with $P=0.001$, $P=0.01$, $P=0.05$, $P=0.10$.

2.6 Isolation of methylated regions, DNA sequencing, and bioinformatics analysis

MSAP banding patterns, *HpaII*+/*MspI*- or *HpaII*-/*MspI*+, which correspond to regions with internal or external methylation of cytosines in the DNA, were randomly selected and excised directly from the 10% polyacrylamide gel. In the next step, the fragments were re-amplified, recovered on 1% agarose gel, and purified with reagents from the Gel Band and DNA Purification kit (GE Healthcare, Little Chalfont, Buckinghamshire, UK), according to the manufacturer's instructions.

DNA sequencing reaction was prepared according instructions to the BigDye Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer, Foster City, CA, USA). The DNA sequencing was performed in an ABI Prism 377 sequencer (Applied Biosystems, Foster City, CA, USA) with the same primers selective PCR. The quality of the sequences was assessed using the Phred (Ewing & Green, 1998), and ClustalW Multiple Alignment 1.8 (Thompson *et al.* 1994) in BioEdit 7.1.9 (Hall, 1999).

Exons and introns were determined by combining the transcript data and their respective proteins with the help of the BLASTn and BLASTx tools (<http://ncbi.nlm.nih.gov/BLAST>). The nucleotide sequences corresponding to the transcription start site (TSS) were determined using the translate Expasy tool (<https://web.expasy.org/translate/>). We predicted promoter regions with the Neural Network Promoter Prediction tool (http://www.fruitfly.org/seq_tools/promoter.html).

We evaluated the gene set with functional analyses using the DAVID V.6.8 database (Database for Annotation, Visualization, and Integrated Discovery) (Dennis *et al.* 2003; Huang *et al.* 2007). This method was used to identify biological processes and their physiological relationships (cellular components and molecular functions), with the integration of data from different databases to perform functional enrichment using

the Gene Ontology (GO) tool and the visualization of the dynamics of genes and metabolic pathways that were most relevant with the use of KEGG.

3. Results

3.1 DNA methylation patterns

The genomic DNA samples, with pupae and others from reproductive tissues of adult flies, were doubly digested with the isoschizomers *HpaII* (sensitive to methylation) and another with *MspI* (not sensitive to methylation) (Table 1), and both with *EcoRI*. We tested twelve selective primers combinations (E1-E6 and HM1-HM6). We obtained data on unmethylated CCGG sites (bands common to both enzymatic digestions), internal methylation of cytosines at C^mCGG sites (bands present in *MspI* digestions but not in *HpaII*), and external methylation of cytosines in ^mCCGG sites (bands present in *HpaII*, but not in *MspI* digestions).

A dataset comprising 833 DNA regions was analyzed for molecular patterns in MSAP assays, in pupae (n=324 bands), female reproductive tissues (n=308 bands), and male reproductive tissues (201 bands). *Loci* sensitive to cytosine methylation (MSL) in DNA contain variations between biological phases concerning developmental sites. First, methylation differences were observed in organisms reared in a standard laboratory medium (pupas=40 *loci*; female flies=82 *loci*; male flies=55 *loci*). Second, DNA methylation data indicate changes in individuals raised with *P. machrisii* tissues (pupas=69 *loci*; females=52 *loci*; males=42 *loci*). Finally, we observed differences in DNA methylation states in organisms that developed in *C. hildmannianus* tissues (pupas=59 *loci*; females=65 *loci*; males=54 *loci*) (Tables 2, and 3).

The *loci* non-sensitive to methylation (NML) of cytosines in DNA were obtained between the developmental stages and the environments, in standard medium (pupa=65 *loci*; female=64 *loci*; male=29 *loci*), tissues of *P. machrisii* (pupa=56 *loci*; female=15 *loci*; male=13 *loci*) and *C. hildmannianus* (pupa =35 *loci*; female=28 *loci*; male=13 *loci*).

Shannon's diversity index presents different values in biological samples and *loci* (MSAP): pupae (MSL I=0.48; NML I=0.37); female reproductive tissues (MSL I=0.43; NML I=0.23); male reproductive tissues (MSL I=0.43; NML I=0.29).

Internal methylation profiles of cytosines in DNA (C^mCGG sites) indicate differences in samples in pupae and adult flies concerning the environment (breeding sites). We obtained changes in methylation profiles that correspond to the following

data, in *P. machrisii* cactus medium (ratio in pupa =0.34; ratio in female =0.46; ratio in male =0.43) and *C. hildmannianus* (ratio in pupa =0.41; ratio in females =0.39; ratio in males=0.63) and laboratory standard medium (ratio in pupa =0.24; ratio in females =0.32; ratio in males =0.34) (Figures 1, 2, and 3).

Profiles with external methylation data of cytosines in DNA (^mCCGG sites) suggest differences between developmental stages concerning host environments (Figures 1, 2, and 3). These variations methylation are observed throughout the results: cactus *P. machrisii* (ratio in pupa =0.21; ratio in females =0.31; ratio in males =0.45), and *C. hildmannianus* (ratio in pupa =0.21; ratio in female =0.32; ratio in male =0.18), and laboratory standard medium (ratio in pupa =0.14; ratio in females =0.24; ratio in males =0.32).

The relative total methylation changes indicated more significant variation between the sexes and more significant relation to host cacti tissues than in the standard laboratory medium (control) (Figures 1, 2, and 3). Genomic DNA methylation shows alterations in the reproductive tissue concerning the environment, with the data we obtained in *P. machrisii* cactus tissues (ratio in females =0.77; ratio in males =0.88), in *C. hildmannianus* (ratio in females =0.71; ratio in males =0.81), and in the standard medium (ratio in females =0.56; ratio in males =0.66).

The combinations (ACC)/HM(TTC) and E(AAC)/HM(TAC) in the genomic DNA show changes in methylation-sensitive loci more prominent with of pupae and reproductive tissues of females in different environments. In contrast, the MSL *loci* were more representative of the combinations E(ACT)/HM(TTG) and E(AGG)/HM(TGA) in the DNA of reproductive tissues of adult males who were raised in different environments (Tables 1, and 2).

Among the number of MSL *loci*, information on the external methylation status of cytokines was significant in male flies that were reared on *C. hildmannianus* tissues ($P=0.03$) concerning the *P. machrisii* cactus ($P=0.06$) and in the standard laboratory medium ($P=0.11$), in the “msap” program.

DNA methylation in female flies shows a difference in the average pattern of regions between the organisms that developed in the standard medium concerning cactus tissues (*C. hildmannianus*; and *P. machrisii*), with significance in the One-way ANOVA test ($F=2.6618$; $P\text{-value}=0.1025$). These variations were also significant with in relation to the methylated loci and in the interactions with the breeding sites (standard medium and *P. machrisii*), with an adjusted $P\text{-value}=0.09$, in the Tukey test. Statistical scores were considered significant with $P=0.1$.

3.2 DNA sequencing and bioinformatics analyzes

A set of 15 bands was selected with differential cytosines methylation, with a molecular range from 200 bp to 1000 bp. Among these samples, nine satisfactory DNA sequences with Phred values >15 for cytosine methylation targets in reproductive tissues of adult flies, including females and males, which were developed in standard laboratory medium, medium with tissues of the *C. hildmannianus* cactus, and another in *P. machrisii*.

Searches for similarities between five DNA sequences, with differential methylation of cytosines, were found with other gene regions in the NCBI, FlyBase, and Genome Ensembl, with Blastn tool: *piezo* (242pb; E-value:0.028; FBgn0264953), *eEF1alpha1* (360pb; E -value:1.14e-16; FBgn0284245), ncRNA (315pb; E-value:1e-04; XR_007421768.1), transposon *Galileo* (307pb; E-value:9e-06; AY756168.1) (Figure 4), *eEF1alpha1*(218pb; E-value:1.22594e-21; FBgn0000557) (Table 4).

These data were obtained on female flies bred on *C. hildmannianus* (*piezo*) and on *P. machrisii* (*eEF1alpha1*), and on other male flies bred on *C. hildmannianus* (ncRNA, Transposon *Galileo*) and on *P. machrisii* (*eEF1alpha1*). Unfortunately, the other three DNA sequences could not accurately determine the gene regions, which may be related to other factors, such as the low molecular length of the query sequences (<200bp).

We analyzed gene transcriptional regions, with the Blastx tool, which was retrieved only for the *eEF1alpha1* genes (butterfly; GenBank: ABV53950.1; E-value:0.005) (beetle; GenBank: KAH0819316.1 ; E-value: 2e-06). In the next step, we identified four transcriptions start sites using the translate ExPasy tool in the *Piezo*, Transposon *Galileo*, and *eEF1alpha1* sequences. Finally, we identified the predicted five promoter regions in the gene sequences, which correspond to *Piezo*, *eEF1alpha1*, ncRNA, and another in *eEF1alpha1*, with the Neural Network Promoter Prediction tool (Table 5).

We analyzed the gene sequences using the Ontology gene tools (GO and KEGG) in the David program. This method searches for terms of biological processes, cellular components, and molecular functions, as well as visualizing interaction data.

Our results indicate that the *Piezo* gene has functions related to mechanosensitive channels, which involve behavior and response to stimuli in the biological process of identification and evaluation of substrate hardness by female flies during the oviposition process.

A second biological process, with the *eEF1alpha1* gene (Figures 5 and 6), indicates a role in GTPase and elongation factor activities of the translation machinery (ribosomal complex), which may contribute to the process in the adult life cycle (Gene ID 36271) (Figure 4). Finally, TE *Galileo* indicates a function related to genome organization and stability, in inversions in chromosome 2 and close to chromosomal break target points (GenBank accession: AY756168.1), in *Drosophila buzzatii* (Delprat *et al.* 2009) (Fig1 SuplMaterial).

4. Discussion

The results presented in this work demonstrate that the phenomenon of differential methylation in DNA presents changes in genomic regions and molecular pathways when organisms develop in different ecological resources in the cactophilic species *D. gouveai*. DNA methylation profiles, using MSAP, combined with a direct sequencing approach, allowed us to understand the methylated regions and the molecular pathways that contribute to the change of host sites and to address issues such as parental origin's effect.

The DNA methylation data agree with previous works, in which the presence of methylation in the genomic DNA was evidenced in species of the genus *Drosophila*, with regions with approximately 1% of its genome (Lyko *et al.* 2000; Deshmukh *et al.* 2018).

Differently methylated regions are documented preferentially at CpT, CpA, followed by CpC, and CpG sites (Lyko *et al.* 2000). The data indicated methylated sites in the CpC and CpG targets, which coincide with the results found in the genomes of *D. melanogaster* (Deobagkar *et al.* 2004; Panikar *et al.* 2015), and *D. willistoni* (Garcia *et al.* 2007), with the use of isoschizomer enzymes *MspI* (5'C↓CGG3') and *HpaII* (5'C↓CGG3'). This suggests regions may present a methylated DNA motif with discussed functions among *Drosophila* species.

Data obtained by Drewell *et al.* (2014) indicate differences in the number of methylated CpG sites in genomic DNA in samples of ovaries and sperm, which may imply the development process of drones and queens in the bee in *Apis mellifera*. This phenomenon is explained by the parental and lineage effect that can result in differentially methylated regions in DNA and alter gene expression in *Apis mellifera* (Wu *et al.* 2020). Evidence indicates that regions with methylation in genomic DNA can be transmitted between parents to their offspring (F1) in wasp species of the *Nasonia*

genus (Wang *et al.* 2016), which suggests an intergenerational inheritance. Data from this work may suggest an intergenerational inheritance when organisms use different host sites in *D. gouveai* development. However, more data are needed to confirm intergenerational inheritance when organisms use different ecological niches.

An exciting process is methylation of CCGG targets affects DNA helical structure conformation and responses to torsion stress in the molecule (Hörberg & Reymer, 2018). The sequences with the base targets of unmethylated (C) and methylated in (Cm) cytosines bases in DNA, CCGG/CCmGC and CCmGA/CTGA, show a bimodality of CpG displacement, with twist and asymmetry in rigidity in the methylated DNA that can block the target sequence (Hörberg & Reymer, 2018). Based on these predictions of DNA molecules, the variation of methylation and non-methylation of cytosines in DNA sequences may signify a change in DNA conformation and present implications for the structure of the epigenome in organisms in *D. gouveai*.

Two functional aspects indicate that differentially methylated DNA genes participate in processes that involve mechanoreceptors in oviposition activity (*Piezo* gene), and in the transcriptional synthesis and nucleus-cytoplasm communication (*Factor elongation alpha 1* gene), in female flies of *D. gouveai*. In contrast, other biological processes participate in cellular viability and genome stability (ncRNAs, TE *Galileo*) and, additionally, in transcriptional synthesis and nucleus-cytoplasmic communication (*Factor elongation alpha 1* gene), in male flies of *D. gouveai*.

Differential DNA methylation was found concerning regions with synthesis functions, such as ribosomal DNA (rDNA) in female flies of *D. willistoni* and *D. melanogaster* (Garcia *et al.* 2007; Panikar *et al.* 2015). This suggests that females may have a lower ability to recognize *C. hildmannianus* cactus tissue and possibly reduce protein synthesis or mRNA transport in nucleus-cytoplasmic communication when organisms develop on the *P. machrisii* cactus.

Data in the literature have indicated that chemical changes in DNA can result in phenotypes in foraging behavior (*For* gene), sexual courtship (Kramer *et al.* 2011; Anreiter *et al.* 2017), and oocyte differentiation into females (Lovino, 2014), in *D. melanogaster*. The data call attention to the mechanism of oviposition substrate recognition in *D. gouveai*.

Soto *et al.* (2007) indicate that compounds from the cactus *C. hildmannianus* can decrease the viability between larvae to adults of *D. gouveai*, when the organisms use this resource concerning the cactus *P. machrisii*. The data indicate that different DNA

methylation may imply aspects of organism viability when altering the use of ecological resources in males of *D. gouveai*.

Our data indicate that differential methylation occurs in promoter regions predicted for *Piezo* genes in female flies reared on *C. hildmannianus* cactus and in the *eEF1alpha1* gene in females reared on *P. machrisii* in *D. gouveai*. This species of *Drosophila* have been found in natural breeding sites with the cactus *P. machrisii*, which is used principally concerning other cacti (Tidon-Sklorz & Sene, 2001). Considering the promoter data in eukaryotes mentioned above, the addition of a methyl group in regulatory regions can decrease the *Piezo* transcriptional product when females are reared in cacti unusually in nature (*C. hildmannianus*), and also decrease mRNA synthesis or nucleus-cytoplasm communication in female flies that developed on the cactus used in nature (*P. machrisii*).

The physiological mechanism in *Drosophila* contains key organs with metabolic functions conserved with vertebrates, which involve insulin analog, insulin-like growth factors, and target of rapamycin (TOR) signaling, with the regulation of circulating sugars, storage and metabolism of energy (Baker & Thummel, 2007; Schlegel & Stainier, 2007). Female flies can transmit phenotypes with altered body mass to offspring in the F2 generation concerning obesity characteristics, which involves epigenetic mechanisms and metabolism reprogramming when the organism develops under different environmental conditions (glucose, glycogen, trehalose) in *D. melanogaster* (Buescher *et al.* 2013). The data suggest that DNA methylation targets may participate in oviposition and physiological substrate recognition processes (protein synthesis and nucleus-cytoplasm communication) when organisms develop *C. hildmannianus*, and *P. machrisii*, in *D. gouveai*.

A central aspect is that micronutrients and organic compounds present in plant tissues can alter development, with differences in the methylation pattern in the genomic DNA of organisms. Hara *et al.* (1994) demonstrated that alkaloid compounds in cell culture contribute to a remarkable increase in methyltransferase activity. Additionally, the concentration of alkaloids may be related to changes in the pattern of DNA methylation and regulation of gene expression, with the participation of metabolic processes in plants (Bernoville *et al.* 2020). The alkaloid concentration concerning the alteration of phenotypes in the cactophilic species *Drosophila buzzatii* and *D. koepferae*, sisters of *D. gouveai*, when they develop in different columnar cacti hosts (Soto *et al.* 2014). One hypothesis is that the alkaloid content in plant tissues in different cactus species can alter DNA methylation mechanisms (Dnmt2 enzyme and other

enzymes not yet elucidated) and contribute to the regulation of phenotypes in cactophilic *Drosophila* species.

Data from Moraes & Sene (2007) indicate low genetic diversity in individuals of *D. gouveai* populations that occur in Altinópolis (São Paulo; Brazil), with inbreeding and population isolation, from other geographic regions (Minas Gerais, Goiás, and Bahia in Brazil). In contrast, the plasticity of viability phenotype and developmental time, larvae to adults, were fast for changing plant tissues from *P. machrisis* to the cactus *C. hildmaniannus* (Soto *et al.* 2007). Our data indicate that differential DNA methylation occurs when organisms develop in plant tissues of *P. machrisis* than in *C. hildmaniannus*, and in standard (control) medium, in pupae and adults of *D. gouveai*. These DNA methylation changes could mean a physiological adjustment in *D. gouveai*, when organisms develop in different ecological resources.

5. Conclusions

Our data suggest that methylation variations of cytosine bases participate in sex-specific differentiation processes, with more significant evidence of methylation related to reproductive tissues of female flies than in males. The second aspect is the role of DNA methylation in oviposition processes, and genome stability when organisms alter the use of tissue resources in the cacti *P. machrisis*, and in *C. hildmaniannus*, in the cactophilic species *D. gouveai*. Environmental micro compounds (alkaloids and vitamins B or C) may contribute to *Drosophila*'s differences in DNA methylation states. However, this molecular mechanism in the ecological interactions of these organisms has not yet been elucidated.

Acknowledgment

The authors were funded by the Universidade de São Paulo (USP) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (CAPES 88882.378679/2019-1 to A.S.S.). Financial resources from the National Research Center (CNPq; Prof. Fábio M. Sene).

6. Bibliographic references

- Anreiter, I.; Kramer, J. M.; Sokolowski, M. B. (2017) Epigenetic mechanisms modulate differences in *Drosophila* foraging behavior. *Proc. Natl. Acad. Sci. U S A.*, v.114, n.47, 12518-12523.
- Baker, J. L.; Olsen, L. W.; Sørensen, T. I. (2007) Childhood body-mass index and the risk of coronary heart disease in adulthood. *N Engl J Med.*, v.357, n.23, 2329-2337.
- Baimai, V.; Sene, F. M.; Pereira, M. A. Q. R. (1983) Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* group. *Genetica*, v.67, 81-92.
- Bernonville, T. D.; Maury, S.; Delaunay, A.; Daviaud, C.; Chaparro, C.; Tost, J.; O'Connor, S. E.; Courdavault, V. (2020) Developmental Methylation of the Medicinal Plant *Catharanthus roseus* Unravels the Tissue-Specific Control of the Monoterpene Indole Alkaloid Pathway by DNA Methylation. *Int J Mol Sci.*, v.21, n.17, 6028.
- Buescher, J. L.; Musselman, L. P.; Wilson, C. A.; Lang, T.; Keleher, M.; Baranski, T. J.; Duncan, J. G. (2013) Evidence for transgenerational metabolic programming in *Drosophila*. *Dis. Model Mech*, v.6, n.5, 1123–1132.
- DeBrito, R. A.; Manfrin, M. H.; Sene, F. M. (2002) Nested cladistic analysis of Brazilian populations of *Drosophila serido*. *Mol Phylogenet Evol.*, v. 22, n.1, 131-43.
- Deobagkar, D.; Deshpande, A.; Chatterjee, S.; Kelkar, A. (2004) CpC methylation is present in *Drosophila melanogaster* and undergoes changes during its life cycle. *Drosoph Inf Serv*, v.87, n.7, 1-4.
- Dennis, G. Jr.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.*, v.4, n.5, 1-11.
- Deshmukh, S.; Ponnaluri, V. C.; Dai, N.; Pradhan, S.; Deobagkar, D. (2018) Levels of DNA cytosine methylation in the *Drosophila* genome. *PeerJ.*, v.6, e5119.
- Drewell, R. A.; Bush, E. C.; Remnant, E. J.; Wong, G. T.; Beeler, S. M.; Stringham, J. L.; Lim, J.; Oldroyd, B. P. (2014) The dynamic DNA methylation cycle from egg to sperm in the honey bee *Apis mellifera*. *Development.*, v.141, n.13, 2702-2711.
- Ewing, B.; Green, P. (1998) Base-Calling of automated sequencer traces using *Phred*. II. Error probabilities. *Genome Res.*, n.8, 186-194.
- Fanara, J. J.; Fontdevila, A.; Hasson, E. (1999) Oviposition preference, viability, developmental time and body size in the cactophilic sibling species *Drosophila koepferae* and *D. buzzatii* in association to their natural hosts. *Evol. Ecol.*, v.13, 173-190.

- Fitz-James, M. H.; Cavalli, G. (2022) Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet.*, v. 23, n.6, 325-341.
- Fox, R. J.; Donelson, J. M.; Schunter, C.; Ravasi T.; Gaitán-Espitia, J. D. (2019) Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Phil. Trans. R. Soc. B*, n. 374, 20180174.
- Fulneček, J.; Kovařík, A. (2014) How to interpret methylation sensitive amplified polymorphism (MSAP) profiles? *BMC Genet.*, v.15, n.2, 1-9.
- Garcia, R. N.; D'Ávila, M. F.; Robe, L. J.; Loreto, E. L. Da S.; Panzera, Y.; Heredia, F. O. De; Valente, V. L. Da S. (2007) First evidence of methylation in the genome of *Drosophila willistoni*. *Genetica*, v.131, 91-105.
- Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, v.41, 95-98.
- Hara, M.; Kitamura, T.; Fukui, H.; Tabata, M. (1993) Induction of berberine biosynthesis by cytokinins in *Thalictrum minus* cell suspension cultures. *Plant Cell Rep.*, v.12, n.2, 70-73.
- Hörberg, J.; Reymer, A. (2018) A sequence environment modulates the impact of methylation on the torsional rigidity of DNA. *Chem. Commun.*, v.54, n.84, 11885-11888.
- Huang, D. W.; Sherman, B. T.; Tan, Q.; Kir, J.; Liu, D.; Bryant, D.; Guo, Y.; Stephens, R.; Baseler, M. W.; Clifford, H.; Lempicki, R. A. (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Nucleic Acids Res.*, v.35,169-175.
- Kramer, J. M.; Kochinke, K.; Oortveld, M. A. W.; Marks, H.; Kramer, D.; Jong, E. K. De; Asztalos, Z.; Westwood, J. T.; Stunnenberg, H. G.; Sokolowski, M. B.; Keleman, K.; Zhou, H.; Bokhoven, H. V.; Schenck, A. (2011) Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *Plos Biology*, v.9, 1-15.
- Krauss, V.; Reuter, G. (2011) DNA methylation in *Drosophila*--a critical evaluation. *Prog Mol Biol Transl Sci.*, v.101, 177-191.
- Lovino, N. (2014) *Drosophila* epigenome reorganization during oocyte differentiation and early embryogenesis. *Brief. funct. genomics*, v.13, n.3, 246-253.
- Lyko, F.; Ramsahoye, B. H.; Jaenisch, R. (2000) DNA methylation in *Drosophila melanogaster*. *Nature*,v.408, n.6812, 538-540.
- Machado, L. P.; Madi-Ravazzi, L.; Tadi, W. J. (2006) Reproductive relationships and degree of synapsis in the polytene chromosomes of the *Drosophila buzzatii* species cluster. *Braz. J. Biol.*, v.66, n.4, 279-293.

- Maleszka, R. (2016) Epigenetic code and insect behavioural plasticity. *Curr Opin Insect Sci.*, v.15, 45-52.
- Manfrin, M. H.; Sene, F. M. (2006) Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genetica*, v.126, 57-75.
- Moraes, E. M.; Sene, F. M. (2007), Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. *J. Zool. Syst. Evol. Res.*, v.45, 336-344.
- Olerup, O.; Zetterquist, H. (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. *Tissue Antigens.*, v.39, n.5, 225-235.
- Panikar, C. S.; Rajpathak, S. N.; Abhyankar, V.; Deshmukh, S.; Deobagkar, D. D. (2015) Presence of DNA methyltransferase activity and CpC methylation in *Drosophila melanogaster*. *Mol Biol Rep.*, v.42, n.12, 1615-1621.
- Pereira, M. A. Q. R.; Vilela, C. R.; Sene, F. M. (1983) Notes on breeding and feeding sites of some species of the *repleta* group of the genus *Drosophila* (Diptera, Drosophilidae). *Cien. Cult.*, v.35, n.1, 1313-1319.
- Pérez-Figueroa, A. (2013) msap: a tool for the statistical analysis of methylation-sensitive amplified polymorphism data. *Mol Ecol Resour.*, v.13, n.3, 522-527.
- Ruiz, A.; Wasserman, M. (1993) Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity*, v.70, n.1, 582-596.
- Schlegel, A.; Stainier, D. Y. (2007) Lessons from "lower" organisms: what worms, flies, and zebrafish can teach us about human energy metabolism. *PLoS Genet.*, v.3, n.11, e199.
- Schrey, A. W.; Robbins, T. R.; Lee, J.; Dukes, D. W. Jr.; Ragsdale, A. K.; Thawley, C. J.; Langkilde, T. (2016) Epigenetic response to environmental change: DNA methylation varies with invasion status. *Environ Epigenet.*, v.2, n.2, 1-5.
- Silva, A. F. G.; Sene, F. M. (1991) Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). *Rev. Bras. Entomol.*, n.55, 455-468.
- Soto, I. M.; Carreira, V. P.; Corio, C.; Padró, J.; Soto, E. M.; Hasson, E. (2014) Differences in tolerance to host cactus alkaloids in *Drosophila koepferae* and *D. buzzatii*. *Plos One*, v.09, 1-9.
- Soto, I. M.; Hasson, E. R.; Manfrin, M. H. (2008) Wing morphology is related to host plants in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera, Drosophilidae). *Biol. J. Linn. Soc.*, n.95, 655-665.

- Soto, I. M.; Manfrin, M. H.; Sene, F. M.; Hasson, E. (2007) Viability and developmental time in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera: Drosophilidae) are dependent on the cactus host. *Ann. Entom. Soc. Am.*, v.100, n.04, 490-496.
- Soto, I. M.; Soto, E. M.; Corio, C.; Carreira, V. P.; Manfrin, M.; Hasson, E. (2010) Male genital and wing morphology in the cactophilic sibling species *Drosophila gouveai* and *Drosophila antonietae* and their hybrids reared in different host plants. *Environ. Entomol.*, v.39, n.03, 865-873.
- Stearns, S. C. (2014) **Natural selection, adaptation, and fitness: overview.** In: Losos, J. B. The Princeton guide to Evolution. Princeton University, New Jersey, 208-214.
- Thompson, J. D.; Higgins, D. G.; Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, v.22, 4673-4680.
- Tidon-Sklorz, R.; Sene, F. M. (2001) Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia (Zoologia)*, n.90, 141-146.
- Wang, X.; Werren, J. H.; Clark, A. G. (2016) Allele-Specific Transcriptome and Methylome Analysis Reveals Stable Inheritance and Cis-Regulation of DNA Methylation in *Nasonia*. *PLoS Biol.*, v.14, n.7, e1002500.
- Whitman, D.W.; Agrawal, A. A. (2009) **What is phenotypic plasticity and why is it important?** In: Whitman, D. W.; Ananthakrishnan, T. N. Phenotypic Plasticity of Insects: Mechanisms and Consequences. Science Publishers: Enfield, 1–63.
- Wu, X.; Galbraith, D. A.; Chatterjee, P.; Jeong, H.; Grozinger, C. M.; Yi, S. V. (2020) Lineage and Parent-of-Origin Effects in DNA Methylation of Honey Bees (*Apis mellifera*) Revealed by Reciprocal Crosses and Whole-Genome Bisulfite Sequencing. *Genome Biol Evol.*, V.12, n.8, 1482-1492.
- Zhou, X.; Chen, J.; Liang, M. S.; Wang, F. (2013) Differential DNA methylation between two wing phenotypes adults of *Sogatella fucifera*. *Genesis*, v.51, n.12, 819-826.

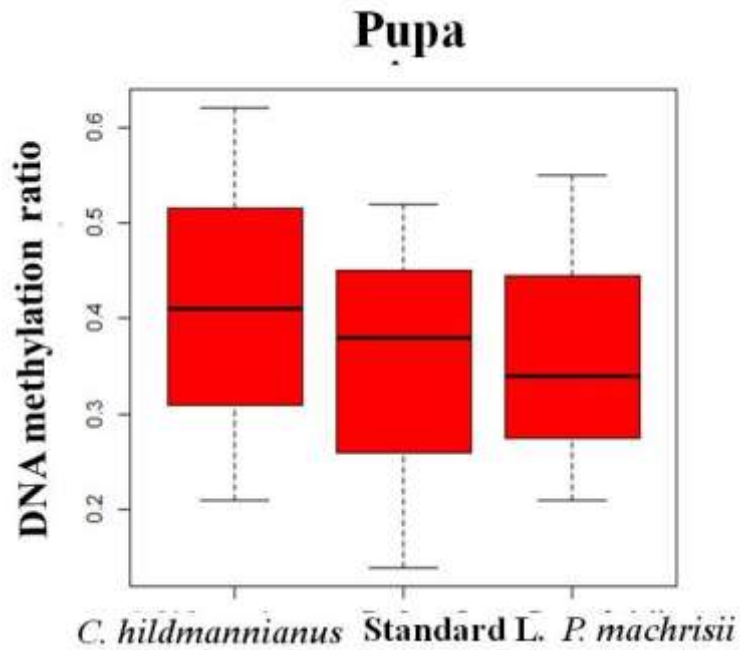


Figure 1. Boxplot plot with relative ratios of methylation in genomic DNA in the organism (pupae) that developed in environments with standard laboratory medium (Standard L), tissues from *C. hildmannianus*, and *P. machrisii* cacti. These data were obtained from molecular profiles generated with the internal and external methylation states of cytosine bases represented (*HpaII*-/*MspI*+; *HpaII*+/*MspI*-), and both in the relative total of methylation, in red.

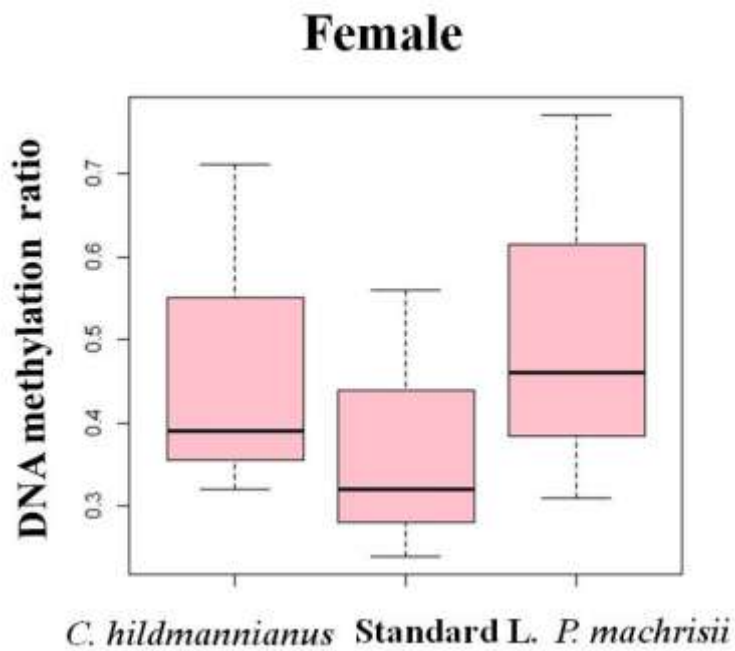


Figure 2. Boxplot plot with relative ratios of methylation in genomic DNA in reproductive tissue (adult, female flies) that developed and emerged in environments with standard laboratory

medium (Standard L), tissues from *C. hildmannianus*, and *P. machrisii* cacti. These data were obtained through molecular profiles generated with the internal and external methylation states of cytosine bases represented (*HpaII*-/*MspI*+; *HpaII*+/*MspI*-), and both in the relative total of methylation, in pink.

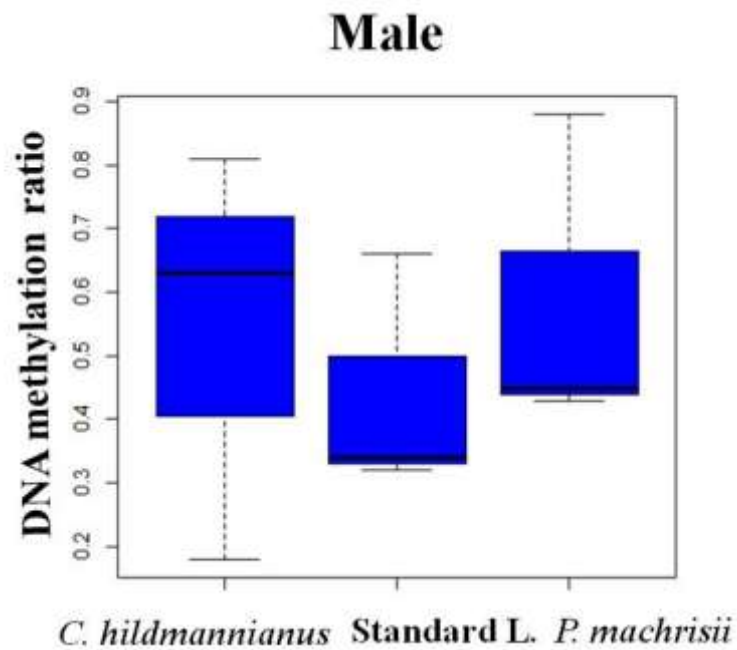


Figure 3. Boxplot plot with relative ratios of methylation in genomic DNA in reproductive tissue (adult, male flies) that developed and emerged in environments with standard laboratory medium (Standard L), tissues from *C. hildmannianus*, and *P. machrisii* cacti, with the second consecutive development. These data were obtained through molecular profiles generated with the internal and external methylation states of cytosine bases represented (*HpaII*-/*MspI*+; *HpaII*+/*MspI*-), and both in the relative total of methylation, in blue color.

```

          310       320       330       340       350       360       370
D_buzzatii - TE Galileo seq_t A A A A G A A A C C T C A T C T G T C A A T G T A T C C C A A T G C A A A T G A A A T G T A T C C C A A T G C T T T T T T A G C T T T G G A
D_buzzatii - TE Galileo seq_p C . O T . T . G . T G . T G . G A . - - . C . C . C T O T T . . T G . T T C . G O T G G . . T . T . T G . . T G . A A . G A T A G . O C A . T
Query Sequence -----

          410       420       430       440       450       460       470
D_buzzatii - TE Galileo seq_t A C C A T A C A A C A C A T A G A C T G G A C A C T G G A A C A A A T T A A T G T G C A C A C T T A T T T T T T G A C A T T A G A A C C C C
D_buzzatii - TE Galileo seq_p . T A G A C T T T T T G . . T . T T . A A . T . G . C A T G T G T T . . T G A C A . . . A T . C . T A . A G G . . C G C C T T . C A A G G
Query Sequence -----

          510       520       530       540       550       560       570
D_buzzatii - TE Galileo seq_t T T T T C G G T C G G G A G A G A A A A A T T G C T G A T G A C A A A A T C A T T C C G C T C G A G C G C T C G A C G G T G A T G G T C G
D_buzzatii - TE Galileo seq_p A A . G A . A . A . T A T T T T C . T G . C C . . . A A T A T C . T C . . A A . G . A T A T . G . A T T A C A G G . . T . A . T . . T A C A A
Query Sequence -----

          610       620       630       640       650       660       670
D_buzzatii - TE Galileo seq_t T C T T T C C C T T G C T C T C T T T C T T T C T G C C C A T G C T T T T A C T C T C T G C T C - T C T T C T C A C A A A A A T A T A
D_buzzatii - TE Galileo seq_p A T . . G G A A . G A . A A C A A . G - A T C . C T . C T T T O T . . . C A C G . C A A A A T . A A A . C A A G A T C . . T T . G . . A - - .
Query Sequence -----

          710       720       730       740       750       760       770
D_buzzatii - TE Galileo seq_t A C A A A A A T T G A A A T T T C T T C T C G A C C G A T C C C A T T G O T T C C C T T C T G A C C G A A G A G T A T A C C G A A T T C T T
D_buzzatii - TE Galileo seq_p G . . C . . A . A C G . A C C . A A . C A T T T T T . . A T T T . A . T C C A T A A . A T C C . A A A T G C T T . C . . T G A . . G . T . .
Query Sequence -----

          810       820       830       840       850       860       870
D_buzzatii - TE Galileo seq_t T G T T A G C G T T T T C A T G A T T A T T T G T G A A A G G G T A T G A A A G T G T T G A T G T G C A T A C A G A A T A T T T T T T C - -
D_buzzatii - TE Galileo seq_p . A . . - - . T A G A . G . A . . . . C A A . . C T A G C T A A T A T A A T . . A A . A . A G C . G . T C C . . C C T T . C . G . A . . T C A
Query Sequence -----

          910       920       930       940       950       960       970
D_buzzatii - TE Galileo seq_t T T C C - T C G T A T A T G T G T A G G C A T T A T G T G C C T G T G T G C A A T A T G C T T G C T G C C T G T A G C A C A T T A G A A C
D_buzzatii - TE Galileo seq_p A . G A G A A . C T . C A . A T C C A . A T . . . G A A T A G A A . T G T . . T . . G A T . C A A C A A T . . . . A A C A T . C . A . . T
Query Sequence -----

         1010      1020      1030      1040      1050      1060      1070
D_buzzatii - TE Galileo seq_t A G C G A T G T G A O C T T G C A T A C A C A C A C A C A T T T A C A C T T C T T C G C T A T A T G T G A G C A T T T A T G T C C T G T
D_buzzatii - TE Galileo seq_p T . . . G G A . - . T . . C C T G C . . . . A G . T G T T G C A A . G . A C . G A . T . T G . A G . T . . . . C C A . . C . - - - - - T T . A A
Query Sequence -----

```

Figure 4. Alignment of nucleotide sequences corresponding to the *Galileo* mobile element in *Drosophila buzzatii*, with complete (seq_t) and partial (seq_P) sequence, and in *Drosophila gouveai* (Query sequence). This region was amplified by PCR, with DNA samples from male flies raised in *C. hildmannianus*, in the species *D. gouveai*. The MSAP assay was performed with *EcoRI-MspI*. The symbols indicate regions with similarity (.) and substitution/deletion of bases (-).

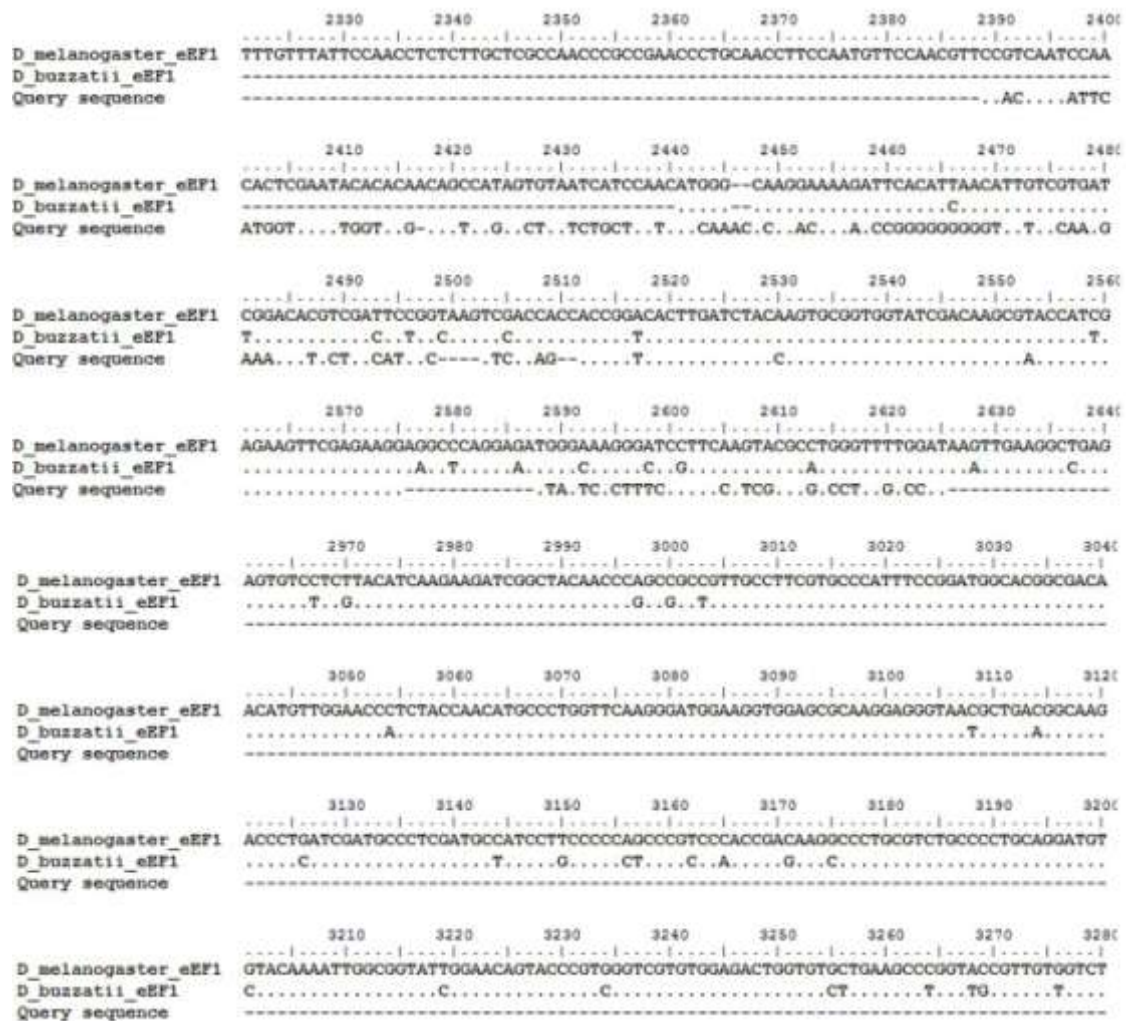


Figure 5. Nucleotide sequence alignment of the *eEF1alpha1* gene in *Drosophila melanogaster*, *Drosophila buzzatii*, and *Drosophila gouveai* (Query sequence). This region was amplified in PCR, with DNA samples in male flies that were reared with the cactus *P. machrisii*, in *D. gouveai*. MSAP assay performed with *EcoRI-HpaII*. Symbols indicate regions with similarity (.) and base substitution/deletion (-).

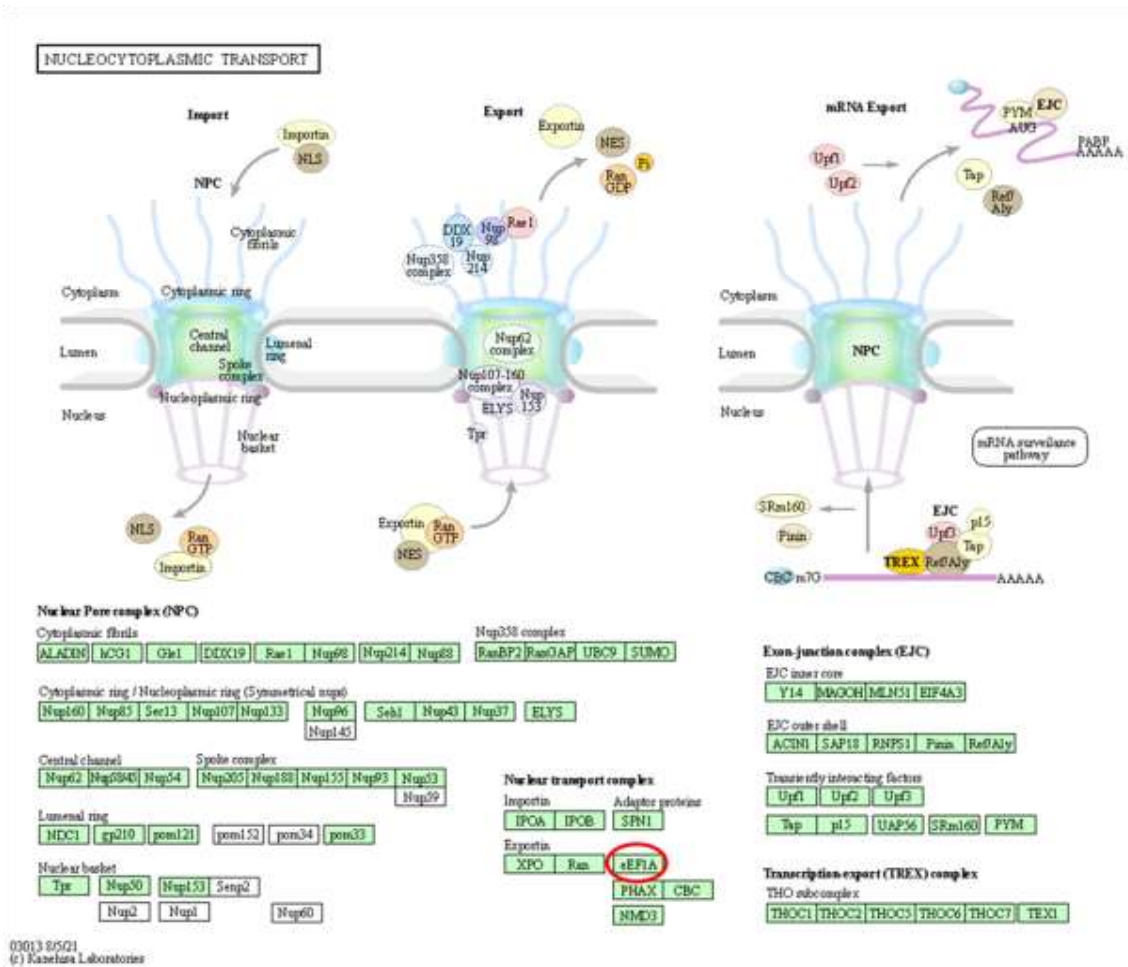


Figure 6. Protein network with interactions between eEF1A and nucleus-cytoplasmic transport signaling. The representation contains the Nuclear Pore Complex (NPC) with the nuclear basket, lumen ring, Spoke complex, central channel, Symmetry nups, Nup358 complex, and cytoplasmic fibrils. The nuclear transport complex involves importins (IPOA, IPOB), adapter protein (SPN1), exportins with XPO, Ran, eEF1A, PHAX, CBC, NMD3. eEF1A is indicated by the red circle. Exon-junction complex (EJC) is represented by EJC inner core, EJC outer shell, and transient interaction factors. Transcription-export complex (TREX), which involves the THO subcomplex.

Table 1. List of primers used for MSAP analyses, with adapters, pre-amplifications, and selective amplifications, with their respective combinations.

		EcoRI	MspI/HpaII	
	Adaptador E1	5'-CTCGTAGACTGCGTACC-3'	Adaptador HM1	5'-GACGATGAGTCTAGAA-3'
	Adaptador E2	5'-AATTGGTACGCAGTCTAC-3'	Adaptador HM2	5'-CGTTCTAGACTCATC-3'
	PreAmp E0	5'-GACTGCGTACCAATTCA-3'	PreAmp HM0	5'-GATGAGTCTAGAACGGT-3'
Amplificação Seletiva	(1)E1	5'-GACTGCGTACCAATTCAAC-3'	HM1	5'-GATGAGTCTAGAACGGTTC-3'
	(2)E2	5'-GACTGCGTACCAATTCACC-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(3)E3	5'-GACTGCGTACCAATTCACG-3'	HM3	5'-GATGAGTCTAGAACGGTAG-3'
	(4)E4	5'-GACTGCGTACCAATTCAC-3'	HM4	5'-GATGAGTCTAGAACGGTTG-3'
	(5)E5	5'-GACTGCGTACCAATTCAGC-3'	HM5	5'-GATGAGTCTAGAACGGTCC-3'
	(6)E6	5'-GACTGCGTACCAATTCAGG-3'	HM6	5'-GATGAGTCTAGAACGGTGA-3'
	(7)E1	5'-GACTGCGTACCAATTCAAC-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(8)E1	5'-GACTGCGTACCAATTCAAC-3'	HM3	5'-GATGAGTCTAGAACGGTAG-3'
	(9)E2	5'-GACTGCGTACCAATTCACC-3'	HM1	5'-GATGAGTCTAGAACGGTTC-3'
	(10)E3	5'-GACTGCGTACCAATTCACG-3'	HM2	5'-GATGAGTCTAGAACGGTAC-3'
	(11)E3	5'-GACTGCGTACCAATTCACG-3'	HM4	5'-GATGAGTCTAGAACGGTTG-3'
	(12)E4	5'-GACTGCGTACCAATTCAC-3'	HM5	5'-GATGAGTCTAGAACGGTCC-3'

Table 2. Number of *loci* that were detected with the combination of selective primers MSAP, with total number (No. *loci*) and methylation sensitive (MSL), non-methylation sensitive *loci* (NML), and paired fragment extension of base (bp), in female flies of *D. gouveai*, with development in standard medium (Q), cactus *C. hildmannianus* (C), and *P. machrisii* medium (P).

Environments	Combination of selective primers	N° <i>Loci</i>	MSL	NML	Extension of the fragments
Q	E(AAC)/HM(TTC)	19	12	7	100-1000
C		5	5	0	
P		2	2	0	
Q	E(ACC)/HM(TAC)	15	4	11	100-1000
C		8	4	2	
P		6	6	0	
Q	E(ACG)/HM(TAG)	10	2	8	100-700
C		11	7	4	
P		3	3	0	
Q	E(ACT)/HM(TTG)	17	15	2	100-900
C		8	4	4	
P		7	7	0	
Q	E(AGG)/HM(TGA)	20	3	17	100-1000
C		14	9	5	
P		7	7	0	
Q	E(AAC)/HM(TAC)	17	13	4	100-600
C		4	1	3	
P		4	4	0	
Q	E(AAC)/HM(TAG)	8	4	4	100-800
C		5	3	2	
P		0	0	0	
Q	E(ACC)/HM(TTC)	22	18	4	100-1000
C		16	12	4	
P		13	6	7	
Q	E(ACG)/HM(TAC)	4	2	2	100-800
C		0	0	0	
P		4	2	2	
Q	E(ACG)/HM(TTG)	7	7	0	100-1000
C		11	11	0	
P		15	11	4	
Q	E(ACT)/HM(TCC)	7	2	5	100-1000
C		13	9	4	
P		6	4	2	
<i>Loci</i> Number (Standard Deviation - DS)	Q	146 (±6,23)	82 (±5,93)	64 (±4,81)	
	C	95 (±4,86)	65 (±3,98)	28 (±1,86)	
	P	67 (±4,48)	52 (±3,06)	15 (±2,29)	
Shannon's Diversity	MSL	I=0,43 (DS: 0.11)			
	NML	I=0,23 (DS: 0.07)			
Wilcoxon test		W=137	P=0,0001		

Table 3. Number of *loci* that were detected with the combination of selective *primers* MSAP, with total number (No. *loci*) and methylation sensitive (MSL), non-methylation sensitive *loci* (NML), and paired fragment extension of base (bp), in male flies of *D. gouveai*, with development in standard medium (Q), in the medium of cactus *C. hildmannianus* (C), and *P. machrisii* (P).

Environments	Combination of selective primers	N° <i>Loci</i>	MSL	NML	Extension of the fragments
Q	E(AAC)/HM(TTC)	4	1	2	150-900
C		9	7	2	
P		5	3	2	
Q	E(ACC)/HM(TAC)	7	3	4	100-1000
C		2	2	0	
P		6	6	0	
Q	E(ACT)/HM(TTG)	17	15	2	100-1000
C		5	5	0	
P		8	6	2	
Q	E(AGC)/HM(TCC)	5	3	2	100-900
C		8	6	2	
P		3	3	0	
Q	E(AGG)/HM(TGA)	19	11	8	100-900
C		15	11	4	
P		9	7	2	
Q	E(AAC)/HM(TAC)	5	3	2	100-900
C		12	10	2	
P		4	4	0	
Q	E(ACC)/HM(TTC)	13	9	4	100-900
C		5	5	0	
P		7	6	7	
Q	E(ACG)/HM(TTG)	15	10	5	100-700
C		11	8	3	
P		7	7	0	
<i>Loci</i> Number	Q	85 (±5,03)	55 (±5,03)	29 (±2,13)	
(Standard	C	67 (±4,27)	54 (±2,91)	13 (±1,50)	
Deviation - DS)	P	49 (±2,03)	42 (±1,67)	13 (±2,39)	
Shannon's	MSL	I=0,43			
Diversity		(DS:0.14)			
	NML	I=0,29			
		(DS:0.06)			
Wilcoxon test		W=20	P=0.106		

Table 4. Samples obtained from methylation target DNA sequences in *Drosophila gouveai*, with molecular data on length in Nucleotide Base Pairs (bp), percentage of cytosine bases (PBC), BLASTn similarity analysis, Genomic Localization (L.G.), Similarity and Access Number in FlyBase or GenBank.

Samples	Sequences	Molecular scale (bp)	PBC(%)	BLASTn	L.G.	E-value	Similaridade (%)	Acession number
GQOM <i>MspI</i>	ACGCagTCAATtATGACTCTGTCACCGTACAGTGAATTGGTAC ACACTCAAACAGTAAGTTTACAGAA	68pb	22,1%	No found	No found	No found	No found	No found -
GCOM <i>MspI</i>	TCCACAGAAACCTGCTGTTGTCTaCATGgTACGCGgAAGCAA CTGACCAAGCCTCCAGGCAGCCCCGCCGCTGCTGAGCACA CACATATACACCATGGCAGCCGCTAgATGCCCCACCCTCCAC TGGCCTGGGCCTGTGCGCAGTCTCCACCGCGAGTGCAGTG ACTCAAAACAGGGCCAGAGCAGCGGCaAGACCGCAcAGCGC GAGGGGAGAACAGCACACCACTGGGTTcTGCTCC	242 pb	36,4%	Gene <i>Piezo</i>	éxon	0.028	100%	FBgn0264953
GPOH <i>HpaII</i>	TTGTGATGGAACCGTTCTAGACTCATCGGTCTGTATTCCCTGC GAACCGTTCTAGACTCATAGTCGGGGCTCTTTAGACAGATAC CTTTGTTTAGCATACTACGCACAAGCGGCTGTTTTGCCCTAT ATGAAAACCTGACGCGCTAGAGATGACGCCAATATACGCCG TCGGAGACGATGATAACCACCTGATAAAACCTGGTATGTGC CCCCTGCTCGACCGCAATGTAGTGACTCAAAATAGGGCC AGAGTGTCGACCAAACCTTTTTTATTTTAAGGGAAAGTGACTA ACCTTCTCGAACTTCTCGATGGTTTCGCTGGTCGATACCACCC CACTGGTAGATCAAGTGACCGGTCTGT	360 pb	25%	Gene <i>eEF1alpha</i> 1	íntron	1.14e-16	92,2%	FBgn0284245
GCTM <i>MspI</i>	AGCAACAAATTGATGTTAAAAGAATATTATTTTTACCTTACC TTAAAATCACCTTAAACCTTTCACAGTTGGAACCGTTCTAGA CTCATCCATGTGCGAAAAACGTCTAAGCAACAAATGGATGTT AA	127pb	19,7%	<u>3L</u>	2457431 - 2457432 9	0.054	100%	No found
GCTM-1 <i>MspI</i>	TGAACGTATTAGACTGGGCACTGGGCAAGCTGCTATTCGAA GCATTATAGCAAGCAACATACGCACACAGCAGACGGACCTC ACTATAATCGACTCAGCTCATCATTCTGATTAGCTATAAATA TACTCCTTGACTGCGTACCAATTCAACTAAGAAGAACTTTAA	315pb	24,1%	LOC12577 6504	ncRNA	1e-04	80.56%	XR_00742176 8.1

	GAATCCTTAAAGTTCATTGAAAACGTATTAGACTGGGCACTG GGCAAGCTGATTTCGAAGCATTATAGCAAGCAACATACGCAC ACAGCAGACGGCCTCACTATAATCGACTCAGCTCATCATTCT GATTAGCTATAAATATACTCCTGC							
GCTM-2 <i>MspI</i>	TGACTGCGTACCAATTCAACTCTTCGGTGCTGTGTACATATC ATACATTATGCATACATAACAAATGCTCGGCTGCAAAGCAG CACACAACAATTTAAATTTCTTCTAGCCTCGAGTCCATTGGT TCCCTTTTGAAGGAGCTGGAGTCCCCAAGGCTCAGTGCTGG GGCATTATGTACGAAAGCGTACTGCGCTTCGAACTACTAGC AGGCGCGCACACCATTGACTTTCCTGATGATGTAGGTGTCGT CTCCATAGGCAGAAGTTCAAAGATATTGAGGATATGTGCAG CGCAAATATCAATCGC	307pb	24,1%	Transposon <i>Galileo</i>	TE	9e-06	84.5%	AY756168.1
GPTH <i>HpaII</i>	CCACCAATATTCATGGTGAATTGGTACGCAGTCAGAGCTTAT CTGCTCATCATcAAACCCAgACAAAAACCGGGGGGGGTTTT TCCAAAGAAaACATGCTGACATCGCTTACAGACCGGTCCT TGATCTACCAGTGCGGTGGTATCGACAAGCGAACCATCGAG AAGTTCGAGAAGGTTAGTCACTTTCCTTCCATCGCGCGTCC TTTGTCCAT	218 pb	26,6%	Gene <i>eEF1alpha</i> 1	éxon	1,2259 4e-21	87.5%	FBgn0000557

Note. Genomic DNA samples were obtained from individuals cultured in standard laboratory medium (GQOM), and in *C. hildmannianus* (GCOH), and *P. machrisii* (GPOH) cacti, in tissues from the ovaries of female flies. Other DNA samples were obtained for individuals reared in cactus media with *C. hildmannianus* (GCTM), and *P. machrisii* (GPTH), in testis tissue, in males and adults. The enzymatic assays comprised the use of *EcoRI-HpaII* (*HpaII*-H) and *EcoRI-MspI* (*MspI*-M). TE- Transposable/mobile element. UTR – Untranslated Region (non-coding).

Table 5. Bioinformatics analysis with searches for genomic regions with similarity (BLASTn), transcriptional regions (BLASTx), prediction of Transcription Start Sites, and Promoter Regions, in the methylation target sequences in the DNA of *D. gouveai* adult flies.

Samples	BLASTn	BLASTx	Transcription Start Sites	Promoter Regions
GQOM	No found	No found	YES	No found
GCOM	Gene <i>Piezo</i>	No found	YES	YES (score 0.93)
GPOH	Gene <i>eEF1alpha1</i>	eEF1alpha1	YES	YES (score 0.91)
GCTM	3L	No found	YES	No found
GCTM-1	LOC125776504	No found	No found	YES (score 0.86)
GCTM-2	Transposon <i>Galileo</i>	No found	YES	No found
GPTH	Gene <i>eEF1alpha1</i>	eEF1alpha1	YES	YES (score 0.95)

Note. The genomic DNA samples came from individuals (reproductive tissues from females) that were developed in standard laboratory medium (GQOM), in *C. hildmannianus* cactus medium (GCOH), and in *P. machrisii* (GPOH). DNA samples were analyzed in individuals (male reproductive tissues) reared in *C. hildmannianus* (GCTM) cactus medium, and in *P. machrisii* (GPTH). The enzymatic assays comprised the use of *EcoRI-HpaII* (HpaII-H) and *EcoRI-MspI* (MspI-M).

Figure1. SuplMaterial - Dendrogram obtained using the Neighbor-Joining(NJ) method, with data from DNA sequences targeting methylation in *D. gouveai* (Query sequence_28961), with genomic regions available in *Drosophila buzzatii*). The region in *D. gouveai* corresponds to TE Galileo.

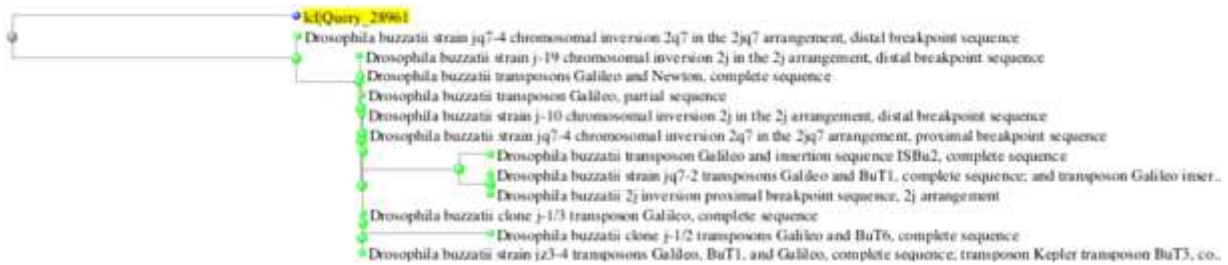


Table1. SuplMaterial - Genomic DNA methylation pattern with *HpaII* and *MspI* enzymes.

Types	Target sites		Enzyme sensitivity		Band Pattern	
			<i>HpaII</i>	<i>MspI</i>	<i>EcoRI-HpaII</i>	<i>EcoRI-MspI</i>
I	CCGG GGCC	5mCCGG GGCC	YES	YES	+	+
II	C5mCGG GGCC		YES	NO	+	-
III	5mCCGG GGCC5m		NO	YES	-	+
IV	5mCCGG GGCC5m		NO	NO	-	-
V	5mCTGG GACC	CNNG GNNC	NO	NO	-	-

“+” represents the presence of bands in the electrophoresis gel; “-” absence of bands in electrophore gel; 5mC represents the methylation site in cytosine bases. N indicates a substitution by one of the four nucleotide bases (A,T,C,G). Modified from Zhou *et al.* (2013)

Anexo C – **Outra produção Científica.** Manuscrito como primeiro autor com colaboração acadêmica, mas externo ao projeto de doutorado, embora realizado durante o curso.

Positive selection and biochemical changes contribute to adaptive evolution of GstD1 protein in cactophilic *Drosophila* species

Adriano Silva dos Santos^{1*}; Silvana Giuliatti¹; Dora Yovana Barrios-Leal¹; Maura Helena Manfrin^{1,2}

¹Postgraduate, Department of Genetics, Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil.

²Department of Biology, Faculty of Philosophy, Sciences and Letters of Ribeirão Preto, University of São Paulo, Brazil.

*Corresponding author.

Adriano Silva dos Santos. Postgraduate, Department of Genetics, Faculty of Medicine of Ribeirão Preto, University of São Paulo, Av. Bandeirantes, 3900, Ribeirão Preto, Brazil.

E-mail addresses: adrianosantos@usp.br (A. S. Santos), silvana@fmrp.usp.br (S. Giuliatti), genyovana@gmail.com (D. Y. Barrios-Leal), maura.manfrin@usp.br (M. H. Manfrin).

Abstract

Genetic changes in codons can result in different properties in proteins; contribute to the diversification of phenotypes and interactions of organisms with the environment. The *Drosophila buzzatii* cluster (*Drosophila repleta* group) is composed of seven species, which obligatorily use cactus necrotic tissues. These plant tissues contain a group of alkaloids involved in the defense against predatory insects and potentially toxic. In *Drosophila*, the *GstD1* gene is recognized for its role in metabolic detoxification and is overexpressed when larvae use the resources of different host cactus species. Here, we use the species of the *D. buzzatii* cluster to analyze the evolutionary process of the *GstD1* sequence gene and the physicochemical properties of the GstD1 protein, including its interaction with the alkaloid substrate. We analyzed the phylogenetic and selection signals in DNA sequences, and used bioinformatics tools for modeling 3D structures of protein of GstD1. GstD1 proteins were docked with mescaline, and the hydrophobic properties were analyzed. The phylogenetic signal is incongruence between the *GstD1* and *mtCOI* genes for the species of the *D. buzzatii* cluster. Were observed four positive selection signals that result in changes in the amino acid residues located within or near the active sites. GstD1 proteins show differences in a slight hydrophobic areas and energy values in interactions with mescaline. Our data indicate that the GstD1 proteins have adjustments that can favor the process of detoxification in host cacti.

Keywords: GstD1; Cactophilic *Drosophila*; molecular evolution; docking; detoxification.

1.0 Introduction

Genetic changes in codons can result in different properties in proteins, with implications for physiological mechanisms and diversification of phenotypes (Low *et al.* 2007; Gloss *et al.* 2014; Saisawang *et al.* 2014). The Glutathione S-Transferase (GST) superfamily is interesting for evolutionary studies due to its diversification by processes that include gene duplication and loss, selective codon signals, amino acid substitution, and differential regulation (Low *et al.* 2007; Saisawang *et al.* 2012). The GSTs enzymes have roles in physiological processes, protection against oxidative stress, hormonal biosynthesis, and detoxification (Saisawang *et al.* 2012). One of the GSTs enzymes, Glutathione S-Transferase D1 (GstD1), has received attention for its metabolic detoxification function, such as in pesticides (dichloro-diphenyl-trichloroethane) (Low *et al.* 2007), cytotoxic compounds (4-hydroxynonenal) (Saisawang *et al.* 2012), and allelochemical substances from plants involved in the use of the host by herbivorous insects (Gloss *et al.* 2014). However, knowledge is unclear about how molecular variations in GstD1 can contribute to insect-plant interactions and changes in ecological environments.

The *Drosophila/cactus* model is excellent for ecological research, with species that use the necrotic tissues of cacti for breeding obligatorily, and most belong to the *Drosophila repleta* group (O'Grady and Markow, 2012; ^aMarkow, 2019). Cactophilic *Drosophila* species diversified on the American continent is hypothetical adaptive radiation through exploration of necrotic tissues of different species of Cactaceae (O'Grady and Markow, 2012). These species are defined as generalists when using different species or specialists when associated with one cactus species (Pfeiler and Markow, 2011; O'Grady and Markow, 2012; Etges and Oliveira 2014).

In the *Drosophila repleta* group, the *Drosophila buzzatii* cluster is composed of seven cactophilic species: *Drosophila buzzatii*, *Drosophila koepferae*, *Drosophila antonietae*, *Drosophila serido*, *Drosophila gouveai*, *Drosophila seriema*, and *Drosophila borborema* (Manfrin and Sene, 2006). These species occur in open and seasonally dry areas in South America (Manfrin and Sene, 2006; Barrios-Leal *et al.* 2019). Information based on mitochondrial *COI* gene (Manfrin *et al.* 2001), nuclear genes *α-esterase 5* (Santos *et al.* 2009), and *period* (Franco *et al.* 2010) support the monophyly of the *D. buzzatii* cluster. These authors indicated that the internal ramifications have some inconsistencies, but in general, the phylogenetic hypothesis suggested a sequence of branching processes with *D. buzzatii* as a first branch, then *D. koepferae* in a second branch following by *D. antonietae*, *D. serido*, and a more derived clade with *D. gouveai*, *D. seriema*, and *D. borborema* as sister taxa (Manfrin *et al.* 2001; Santos *et al.* 2009; Franco *et al.* 2010).

Considering the phylogenetic hypothesis proposed by Oliveira *et al.* (2012), based on multiple molecular markers and mapping the use of the host, the *Opuntia* cactus has been suggested as the ancestral host, and columnar cacti is a condition derived from existing breeders within the *D.*

repleta group, as also in *D. buzzatii* cluster (Manfrin and Sene, 2006; Oliveira *et al.* 2012; Soto *et al.* 2014). For *D. buzzatii* and *D. koepferae*, the diversification can have occurred in the history of specialization in columnar cacti and tolerance to alkaloids (Soto *et al.* 2014). We already know from experimental research that breeding in alternative non-natural cacti hosts alter the morphological characters, developmental time, viability, most of them representing fitness reduction, in *D. gouveai*, *D. antonietae*, *D. koepferae*, and *D. buzzatii* (Soto *et al.* 2007, 2008, 2014).

The exploration of this ecological resource demanded an adaptive adjustment process in morphological phenotypes, various physiologic systems, and a detoxification process associated with the secondary compounds of the cacti, for example, the toxic alkaloids (Kircher *et al.* 1967; Matzkin, 2014; Panis *et al.* 2016). The alkaloids have been recognized as a chemical barrier and potentially toxic to the survival and development of *Drosophila* species (Kircher *et al.* 1967; Soto *et al.* 2014). The alkaloids mescaline and trichocerein, for example, are related to changes in survival, developmental time, and wing morphology in *D. buzzatii* (Padró *et al.* 2018; Soto *et al.* 2014). For mescaline, columnar cactus *Trichocereus terscheckii* and palmar *Opuntia sulphurea* show constitutive differences in concentration in their tissues, with a more significant fraction in columnar cactus (Panis *et al.* 2016). In experiments, *D. buzzatii* showed differences in the levels of expression of genes related to detoxification in larvae individuals that developed in their secondary host *T. terscheckii* cacti compared to its primary host *O. sulphurea*, with *GstD1* overexpression in host plant change (Matzkin, 2014; Panis *et al.* 2016).

In closely related species, *D. arizonae* and *D. mojavensis*, were discovered seven non-synonymous substitutions, with two fixed differences (Leu-7-Gln and His-39-Gln) in the pocket of the active site of the *GstD1* enzyme, when flies used different cactus species (Matzkin, 2008). These amino acid changes provide evidence of an adaptive process for the *GstD1* protein potentially associated with different cacti hosts (Matzkin, 2008). However, the interaction process between the *GstD1* protein and alkaloids when organisms use other ecological resources is not yet elucidated.

Considering the information about *GstD1* and the adaptive process related to secondary compounds of cacti hosts, we used the closely related species of the *Drosophila buzzatii* cluster to investigate the following objectives: (I) understand the evolutionary process associated with the *GstD1* gene considering the phylogenetic relationship of the flies; (II) evaluated the signals of selection in gene and protein sequences; (III) analyze the molecular interactions of *GstD1* proteins with the compound mescaline. The mescaline alkaloid has characterized its chemical structure; unfortunately, other alkaloids are unavailable for biological studies. Moreover, the alkaloid mescaline was the target candidate of the study due to its low concentration in tissues of palmar cactus compared to columnar, and also related to changes in survival, development time, wing morphology, and differences in gene expression in *D. buzzatii* (Soto *et al.* 2014, 2016; Padró *et al.*

2018). We tested the hypothesis that selection signals in *GstD1* gene sequences result in changes in physical-chemical properties and the interaction of proteins with mescaline.

2.0 Material and methods

2.1 Biological samples and sequences

We analyzed samples from six species of the *Drosophila buzzatii* cluster: *D. koepferae*, *D. antonietae*, *D. serido*, *D. gouveai*, *D. borborema*, and *D. seriema* (Table S1). DNA extraction from each individual was performed using *DNeasy*® *tissue* kit (QIAGEN Inc., Hilden, Germany). For *D. buzzatii* the information was extracted from the genome (scaffold 305: 1573-2228) available in <http://dbuz.uab.cat/welcome.php>.

Two pairs of primers for amplification and sequencing of the *GstD1* gene were designed based on *Drosophila GstD1* sequences available in GenBank (EU079442.1 and EU079471.1) for amplification and sequencing of the *GstD1* gene: 1 – F 5' – MATGGYTGACTTCTAYTAYC – 3' and R 5' – TTCTCGTACCACTTGTTACGTTNKCR – 3'; 2 – F 5' – CCCGAGTTCGTGAAGATCAAYCCNCA – 3' and R 5' – GGCTYAWTCGAAGTAYTTCTTG – 3'. The amplifications were performed in 20 µL reaction with 50-75 ng of DNA, 2.0 mM MgCl₂+ buffer, 2.0 mM dNTPs, 0.4 µM of each primer and 0.5 U *Taq* polymerase. The PCR conditions were one step of 94°C for 1 min 30 sec., 35 cycles of 94°C for 30 sec., 46°C for 30 sec. and 72°C for 2 min, with a final extension of 72°C for 10 min.

The DNA sequencing reaction was prepared according to the BigDye Terminator Cycle Sequencing Ready Reaction kit instruction (Perkin-Elmer, Foster City, CA, USA). The DNA sequencing was performed in an ABI Prism 377 sequencer (Applied Biosystems, Foster City, CA, USA). The quality of the sequences was assessed using the Phred (Ewing and Green, 1998). Alignment and sequence editing were performed in ClustalW Multiple Alignment 1.8 (Thompson *et al.* 1994) and BioEdit 7.1.9 (Hall, 1999), respectively. Heterozygous sites (double peaks in chromatograms) were inferred using the Chromas program. New sequencings was performed for these samples, and when confirmed the double peak, the heterozygous sites with amino acid changes were conserved, and sites without amino acid substitution were standardized with the IUPAC code.

2.2 Phylogenetic analyses

To compare phylogenetic signal and codon selection, we included sequences of mtCOI from 11 species previously published, comprising the species analyzed for *GstD1* (Kokudai *et al.* 2011; Franco and Manfrin, 2013; Franco *et al.* 2015) (Table S1). As outgroups, we use the cactophilic species *D. mojavensis*, *D. arizonae* and *D. navojoa*, and *D. virilis* as non-cactophilic. Considering

the mtCOI phylogenetic hypothesis established for the *D. buzzatii* cluster (Manfrin *et al.* 2006), we used two independent markers (mtCOI and *GstD1*) to compare the phylogenetic relationships between the species in the cluster.

The phylogenetic hypothesis using *GstD1* and *COI* sequences was estimated by Bayesian Inference (BI) tree in MrBayes v.3.1.2 (Huelsenbeck, 2001). The General Time Reversible (GTR+G) model of nucleotide substitution was determined as the best fit model in jModeltest v.2.1.2 (Posada, 1998) for both genes. A majority-rule consensus tree was constructed with runs of 10 million generations; each tree was sampled every 1,000 generations and with 25% burn-in. The convergence and stability of the parameters were verified using Tracer v.1.6 (effective sample size, ESS \geq 200) (Rambaut *et al.* 2018), and a consensus tree was visualized using FigTree v.1.3.1 (Available at <http://tree.bio.ed.ac.uk/software/figtree>).

We calculated the congruence level between phylogenetic information from *GstD1* and mtCOI performing distance matrices (Table S2), using the p-distance in Mega v.6.0 (Tamura *et al.* 2013). The Kendall's coefficient (W) of concordance was calculated with the function CADM.global with 999 permutations and 95% significance in the R-package “ape”. The results of W generate values 0 (complete incongruence) to 1 (complete congruence) (Campbell *et al.* 2011).

2.3 Selection signal analysis

To evaluate the codon selections and amino acids properties, the analysis was performed Maximum Likelihood (ML) topology using *GstD1* and mtCOI sequences in MEGA v.6.0 (Tamura *et al.* 2013). The selection signals were analyzed using the ratio (ω) between the rate of non-synonymous substitutions (dN) by synonymous substitutions (dS) ($\omega = dN/dS$) (Doron-Faigenboim *et al.* 2005), using the Selecton Server (Available at <http://selecton.tau.ac.il/>). The models analyzed were M8 (beta, $\omega \geq 1$, proportions $p1 = 1 - p0$) (Yang *et al.* 2000), M8a (beta, $\omega = 1$) (Swanson *et al.* 2003) and M7 (beta) (Yang *et al.* 2000). The values of $\omega < 1$, $\omega = 1$, and $\omega > 1$ represent signs of purifying, neutral and positive selection, respectively. Statistical significance was assessed using the X^2 test, assuming 2 degrees of freedom and $p < 0.05$.

The selective influence assessment of 31 physical-chemical properties of amino acid residues of *GstD1*, considering the ML phylogenetic tree, was evaluated according to McClellan and McCracken (2001) in the TreeSAAP v.3.2 (Woolley *et al.* 2003). This analysis uses a gradient of categories to classify changes in amino acid residues in property between conservative (C1 to C5) and destabilizing (C6 to C8), and the analysis of the values of z-scores indicates the direction of the selection. We focus on the analysis of positive selection, which is indicated by positive z-score values and can represent greater changes in the physical-chemical properties of proteins with

structural and functional consequences. Statistical significance was assessed using $P < 0.05$, $P < 0.01$ and $P < 0.001$.

2.4 Structure modelling analyses

The complete coding sequences (CDs) of *GstD1* were translated into amino acid residues using the EMBOSS Transeq (Available at https://www.ebi.ac.uk/Tools/st/emboss_transeq/) with a standard codon table. Posteriorly, the three-dimensional structure of *GstD1* protein was modeled by homology in Modeller 9v.7 (Sali *et al.* 1995). Initially, we used a three-dimensional structure from *Drosophila melanogaster* (access to the Protein Data Bank (PDB): 3MAK) as the template to obtain for *D. mojavensis* 3D structure. Then, this model was used as a template for the species of the *D. buzzatii* cluster.

The models obtained were analyzed using the Ramachandran Plot Analysis (Available at <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) (Lovell *et al.* 2003). We investigate changes in structural elements between proteins, alignments of the tertiary structure of the protein were performed between samples of species from the cluster *D. buzzatii* using Chimera v.1.14 (Pettersen *et al.* 2004). The *D. buzzatii* protein structure was used as a reference. We analyze the residues that constitute the active sites of GStD1 using metaPocket v.2.0 (Huang, 2009) (Available at <https://projects.biotec.tu-dresden.de/metapocket/>). To study the interaction of proteins with the substrate, we performed the molecular docking of the G-site, in the GStD1 protein, with the alkaloid mescaline (PubChem access CID: 4076; 3,4,5-Trimethoxyphenethylamine) that was suggested as potentially toxic for *Drosophila* (Panis *et al.* 2016), using the PyRx program (Available at <http://pyrx.sourceforge.io/>) (Wolf, 2009). Results with more negative values of interaction affinity energy (Kcal/mol) indicate a better position of the ligand at the binding site (Pagadala *et al.* 2017). For the docking analyses, we used the X-ray resolved structure of the crystallized GStD1 protein (PDB access: 3MAK), protein docking data with DDT (Low *et al.* 2007), to select the G-site structure for the molecular interactions, with independent runs (Pagadala *et al.* 2017). We performed the first interaction of GStD1 with mescaline. We consider the mescaline alkaloid as a target of *Drosophila/cactus* interaction due to the chemical characterization of this alkaloid in cacti, and effects related to gene expression and phenotypes (Soto *et al.* 2014; Panis *et al.* 2016), but other compounds can also interact with GStD1. Finally, changes in the ligand and protein interactions and the hydrophobic forces were manually analyzed and visualized using the PyMOL v.2.3 (DeLano, 2018).

3.0 Results

The 630 base pairs of *GstD1* coding region, single exon, were obtained for species of *D. buzzatii* cluster and the Phred program indicated a high quality for all sequences (values between 20 and 50). We found four heterozygous sites (positions 204, 345, 351, and 354; C>T), only one site changes the amino acid (position 351 - codon 117) in *D. seriema*.

3.1 Phylogenetic analysis

The BI trees based on *GstD1* and other mt*COI* sequences recovered the monophyly of the *D. buzzatii* cluster (Figure 1; Bayesian Posterior Probabilities, BPP = 1). For the *GstD1* BI tree, a polytomy defined the intra-cluster relationship with a branch for *D. buzzatii*, another one for *D. koepferae*, and two well-supported clades: one encompassing *D. antonietae* and *D. serido* as sister species (BPP = 1.0), and other with *D. gouveai*, *D. borborema* and *D. seriema* (BPP = 0.94), with the last two sister taxa (BPP = 1.0) (Figure1).

The Kendall coefficient ($W = 0.66$, Global test value = 0.67, $X^2 = 26.40$, $P = 0.11$) was not significant, indicating that the phylogenetic information of *GstD1* and mt*COI* is not congruent. We reject the hypothesis of similarity between phylogenetic trees.

3.2 Selection test

The Selecton test for the variation in mt*COI* DNA sequences suggested purifying selection for this molecule (Table S3). For the *GstD1* DNA sequences, the Selecton test indicated six codons under positive selection (positions 2, 117, 133, 171, 191, and 209), with changes in amino acid residues (see Table 1). Positive selection was found for the heterozygous site at codon 117 (changed Leu to Val) in *D. seriema*. The substitution in codon 117 results in an amino acid change within the G-site. The substitution at codons 171, 191, and 209 resulted in amino acid changes within the H-site in *D. gouveai*. The substitution in codon 2 is close to the active sites in all species (Figure 2)

The *GstD1* proteins were identified with ten physico-chemical changes and thirteen sites with positive selection signals (Table 2). Destabilizing selection signals were identified in the amino acid residue Val-2-Ala in *D. buzzatii*, Asp-3-Ala in *D. gouveai*, Val-19-Ala in *D. serido* and *D. antonietae*, and Ala/Pro-178-Glu for all species. For *D. buzzatii*, DNA and protein selection tests indicated a purifying selection for the second position with the substitution Val-2-Ala. Positive selection for changes in codons 2, 171, and 191 indicate changes directly related to the physico-chemical properties of the *GstD1* protein (Figure S1).

We found interspecific variations in amino acid residues at the G and H sites in the GstD1 proteins. At site G, *D. seriema* has a substitution of His-39-Glu, which means a change from positive to polar charge and a nonsynonymous change from the third base of codon 39, from CAU to CAA. *Drosophila serido* and *D. koepferae* show a substitution of Leu-117-Val, within the G-site, that is, an alteration of non-polar charged amino acids to an aliphatic is the result of a change in the first position of the CUN codon to GUN, where N represents any basis.

3.3 Molecular modeling and docking

The primary protein of GstD1 for the species of *D. buzzatii* cluster is composed of 209 amino acid residues, with a secondary structure comprising nine α -helices and four β -sheets (Figure 3). The N-terminal domain comprises three α -helices and four β -sheets, and the C-terminal domain comprises six α -helices, in the 3D GstD1 protein. For all cluster species, the number of amino acid residues in site G changed from 40-aa, as in *D. serido*, to 51-aa in *D. buzzatii*. *Drosophila serido*, *D. antonietae*, and *D. buzzatii* showed the highest number of amino acid residue changes in the GstD1 protein (Figure 2). The species in the cluster have positions of conserved residues among proteins, but with some differences in the arrangement of amino acids at the G-site (Figure 2). For the H-site in *D. buzzatii* cluster, the number of amino acid residues changed from 7-aa, as *D. serido*, to 22-aa in *D. buzzatii*.

In the structural alignments of GstD1 protein between the species of *D. buzzatii* cluster, was found a structural variation of 0.18 to 0.26 Å (Figure S2), with the high structural similarity between proteins. The C-terminal region is more variable in amino acid residues than the N-terminal region (Figure 2). Residue 209 of the protein tail has amino acid residue substitutions (Figure 1), such as Asp residue, negative charge, in *D. buzzatii*; amino acid Glu, negative charge, in *D. koepferae*, *D. gouveai*, and *D. seriema*; Lys residue, positive charge, in *D. antonietae*, *D. borborema*, and *D. serido*.

Molecular docking suggests differences in the interaction energy in the complex between GstD1 proteins and the mescaline ligand. *Drosophila seriema*, *D. gouveai*, and *D. koepferae* have low interaction energy values, with values between -5.9 Kcal/mol and -5.6 Kcal/mol (Figure 4). The greater distances of mescaline with residue 10 (central) of GstD1 correspond to the lower interaction energy. For *D. antonietae* and *D. borborema* the energy values calculated were -5.3 Kcal/mol in molecular interactions. *Drosophila serido* and *D. buzzatii* showed the highest energy values of interactions calculated between -5.1 and -4.8 Kcal/mol, respectively. In these last two species, the mescaline ligand was positioned at the top of the pocket of the GstD1 proteins.

Molecular surface changes near the G-site of the GstD1 protein, and properties of amino acids, indicate an adjustment of hydrophobic properties for *D. seriema*, *D. gouveai*, and *D. koepferae*, with a slight gain in portions with less hydrophobic properties close to the G-site (Figure 5). For *D. buzzatii*, GstD1 presents a slight decrease in an area with less hydrophobic property, close to the active site (Figure 5).

4.0 Discussion

The biological diversification of the *Drosophila* species that use cacti as host involved adaptive processes to deal with the detoxification of the secondary compounds of the cacti species (Markow, 2019; Soto *et al.* 2012, 2014). It is already known that in this association, flies that developed in different cacti that are non-natural hosts have different gene expression related to the physiological and detoxification system, such as some GstD1, a member of the Gst gene family (Low *et al.* 2007; O`Grady and Markow, 2012; Oliveira *et al.* 2012; Saisawang *et al.* 2012, 2014; Panis *et al.* 2016), so this gene is recognized for its association with resource exploitation of different host cacti and host changes in cactophilic *Drosophila* (Low *et al.* 2007; Matzkin 2008; López-Olmos *et al.* 2017). The results of our work indicated that in the cactophilic species of *Drosophila buzzatii* cluster, the *GstD1* gene has a different evolutionary dynamic compared to *mtCOI* gene, with signs of positive selection in DNA and protein sequence suggesting differences in metabolic processes between species, as evidenced in the docking simulations with mescaline alkaloid with differences in interaction energy.

4.1 Phylogenetic incongruence

Our results showed incongruence in the phylogenetic signal between *GstD1* and *mtCOI*. This incongruence between the mitochondrial DNA and nuclear DNA is expected due to differences in parental inheritance and effective size (mitochondrial with effective population size four times smaller than nuclear) (Funk and Omland, 2003; Toews and Brelsford, 2012), and several studies have shown this mitonuclear disagreement (Gompert *et al.* 2008; Franco *et al.* 2015, Ré *et al.* 2010; Després, 2019). Furthermore, differences have been found in the selective pressure of the mitochondrial genome compared to the nuclear genome (Toews and Brelsford, 2012; DeBiasse *et al.* 2014), corroborating our data. In mtDNA, strong selective pressures act to maintain mitochondrial function by removing deleterious mutations, with the action of purifying (negative) selection (Shtolz and Mishmar, 2019). These authors indicate that selection can also act on mitochondria to allow the adaptation of cells and organisms to new environments and physiological conditions, with the action of positive selection. For *GstD1* and *mtCOI*, we found signs of purifying and positive selection in the *GstD1* sequences (nuclear DNA) evolution, while *mtCOI* (mtDNA)

showed purifying selection. In *Drosophila*, the *GstD1* gene contributes to the detoxification process and adaptive response to toxic components in different environments/cacti (Low *et al.* 2007; Matzkin, 2008; López-Olmos *et al.* 2017). Our data indicate selection signals in *GstD1* that may be associated with adaptive processes in *Drosophila* with the environment.

4.2 Selection and Protein structure and function

Selective changes in *GstD1* codons may represent noise in the phylogenetic analysis in the *D. buzzatii* cluster species compared to phylogeny with mtCOI (Figure 1; see Table 1), which can be explained by selective differences in the genes (Table S3; Table 1). Our results suggest that selection at codons 2, 171, and 191 may be related to changes in physical and chemical properties of the composition, helical contact area, β structure trends, and solvent-accessible reduction rate, an adjustment in the GstD1 protein. Therefore, positive selection result in amino acid changes and may contribute to functional or structural adjustments in GstD1.

Interestingly, the substitution in the codon 171-Lys found in the *D. buzzatii* cluster (see table 1) can contribute to the interaction with toxic compounds. The 171-Lys residue was associated with a better docking binding between GstD1 and the DDT compound in *D. melanogaster* (Low *et al.* 2007). In resonance data, it has been reported that site 171 does not appear to significantly change the peak intensity ratio with/or without DDT binding, but its immediate neighbors (residues 170 and 172), which change or alter the intensity ratio (Low *et al.* 2010). In *D. mojavensis*, data suggest that the Gly-171-Lys substitution did not evolve in response to DDT selection; one hypothesis is that this substitution may affect the kinetics of DDTase activity or other activities (Matzkin *et al.* 2006). These authors suggest that GstD1 is evolving adaptively in *D. mojavensis*, with data indicating 1.4 times more abundance has been reported in the transcription of the *GstD1* gene in flies raised in organ tube cacti compared to flies raised in a medium containing agria. Thus, the amino acid 171-lys can contribute to better kinetics of the enzymatic activity of GstD1 in the interaction with toxic chemical compounds in the environment.

For *D. buzzatii*, the purifying selection signal at codon 2 in the GstD1 protein results in the alanine residue, which means a loss in the tendency in the conformation of the β -sheets structure and reduced accessibility to the solvent, according to biochemical data (Betts and Russell, 2003). This species presents a subtle loss of hydrophobic regions in the protein compared to *D. seriema*, *D. gouveai*, and *D. koepferae* (Figure 4). Changes in hydrophobic portions can cause effects such as intra and intermolecular interactions, with changes in the interaction of the protein with soluble compounds (Eisenhaber, 1996, 1999). We hypothesized that change in the hydrophobic portions of GstD1 result in less accessibility and stability in the interaction with soluble components in *D. buzzatii*.

In the GstD1 enzyme, the active G-site is an essential region for activity and specificity due to interaction with potentially toxic substrates (Low *et al.* 2007; Matzkin, 2008). On G-site, the positive selection sign in codon 117 is related to the change from Leu-117-Val residue in *D. serido* and *D. koepferae* (see Table 1). The valine residue has a C β branch linked to two non-hydrogen substituents and represents a gain in volume and chemical interaction with a group of lipids within the G site (Betts and Russell, 2003). In *D. melanogaster*, conformational changes in residues 117-Val and 118-Phe have been reported to be important to the DDT binding (Low *et al.* 2010). *Drosophila hexastigama* presents a positive selection signal at codon 116-Thr, which has been suggested to have evolved in the bonding of substrates in the adaptative process in different chemical environments (López-Olmos *et al.* 2017). These authors reported a conformational change in the intermediate space 116-117-118 of the amino acids in the structure of GstD1 proteins of *D. hexastigma* and *D. melanogaster*. Position 117 is essential in the interactions of the GstD1 protein with chemical substrates. These factors suggest adaptive changes in *D. serido* and *D. koepferae* that can improve the activity of GstD1 in metabolic processes. Although we do not have information on population variation, the occurrence of this amino acid in these two species and that do not occur in closely related species, suggest independent evolutionary processes.

We also found the His-39-Gln substitution at the G site in *D. seriema*, although no selection signal was found in the DNA and protein. In populations of *D. arizonae* and *D. mojavensis*, his-39-Gln fixed differences were found, which in the location in the active site with change in polarity, that can result in significant changes in the activity and specificity of the protein during the adaptative process to the cacti of the Baja/California and Sonora/Mexico regions (Matzkin, 2008), corroborating our data. The substitution of histidine (positive polar charge) for glutamine (polar) results in changes in molecular weight and the average volume of residues, with histidine (155Da; 166.3Å) and glutamine (146Da; 148.0Å) (Richards, 1977). In addition, the isoelectric point is different between histidine (7.4) and glutamine (5.6), at an almost neutral physiological pH. Thus, our interpretation of the data suggests that changes in polar charge can result in improvements in G-site activity and specificity with cactus substrates in *D. seriema*.

4.3 Molecular docking

Selective changes in codons and amino acids may represent adjustments in the functions of the GstD1 protein in *D. seriema*, *D. gouveai*, *D. koepferae*, *D. serido*, and *D. buzzatii*. Our molecular docking data showed lower interaction energy between the G-site and mescaline in *D. seriema*, *D. gouveai*, and *D. koepferae*, suggesting a better interaction, and a higher value for *D. buzzatii* and *D. serido*. For *D. antonietae* and *D. borborema*, the interaction energy of GstD1 with mescaline showed intermediate values. Selective changes in *D. serido* and *D. koepferae* did not contribute to the interaction of GstD1 protein with mescaline. The mescaline is a common alkaloid found in different cacti genus such as *Trichocereus pachanoi*, *Opuntia cylindrica* (Turner and Heyman, 1960; Cruse, 1973), *Gymnocalycium*, *Turbincarpus* (Starha 1996, 1999), *Pereskia* (Turra *et al.* 2007), *Lophophora williamsii* (Shulgin, 1979), *Trichocereus terscheckii* and *Opuntia sulphurea* (Shulgin, 1979; Panis *et al.* 2016). In this context, we propose that the species of cactophilic *Drosophila* may have different performances in the metabolism of alkaloids concerning hosts cacti.

Information on breeding and feeding niches for the species of the *D. buzzatii* cluster is scarce. Literature data suggest that *D. buzzatii* preferably uses *Opuntia* sp. with better fitness than *T. terscheckii*, while *D. koepferae* preferably uses *T. terscheckii* with better fitness than *Opuntia* sp. (Soto *et al.* 2014; Panis *et al.* 2016; Hasson *et al.* 2019). In addition, in experiments in which the species grew in media with different mescaline concentrations, *D. koepferae* had better adaptive characters values than *D. buzzatii* (Panis *et al.* 2016). Thus, this difference in the observed interaction energy in docking simulations can be an adaptive change that explains the better performance of *D. koepferae*, which preferably uses *T. terscheckii*, which has a concentration of 0.4 mg of the mescaline alkaloid. In contrast, *D. buzzatii* uses preferential *O. sulphurea*, which contains 0.05 mg of this alkaloid, and may not be adjusted to metabolize a large concentration of mescaline. Probably, the GstD1 protein shows differences in the interaction between organisms with ecological niches.

Unfortunately, specific information about the presence of mescaline in other cacti is not available, such as *Cereus*, *Cephalocereus*, *Pilosocereus*, and others. In general, in our previous knowledge of field collections, we know that *D. gouveai* can use tissues from cactus *Pilosocereus machrisii*, and *Opuntia ficus-indica* (Tidon and Sene, 2001; Manfrin and Sene, 2006). *Drosophila koepferae* uses the tissues of *T. terscheckii*, *Cereus validus*, *Opuntia* sp. (Hasson *et al.* 1992). *Drosophila antonietae* use the resources of *C. hildmaniannus* (Soto *et al.* 2007). *Drosophila serido* use the resources of the columnar cacti *Cephalocereus piauhyensis* and *C. hildmaniannus*, and *Opuntia ficus-indica* palma cactus (Pereira *et al.* 1983). *Drosophila buzzatii* use *T. terscheckii*, *C.*

hildmaniannus, and *O. sulphurea* (Pereira *et al.* 1983; Soto *et al.* 2014). *Opuntia* sp. has higher amounts of sugar and free lipid compounds (cholesterol, sitosterol, and fatty acids) (Stintzing and Carle 2005) than columnar cacti. Columnar cacti are indicated with high chemical complexity due to the presence of toxic substances, such as alkaloids, atypical fatty acids, and triterpenes (Kircher 1967; Fogleman and Danielson, 2001; Stintzing and Carle 2005). *Drosophila buzzatii* showed high energetic value in the interaction of GstD1 with mescaline, which may be related to a lower amount of this alkaloid in the primary host *Opuntia*. In this regard, information from the molecular docking of proteins with ligands suggests that negative interaction values (kcal/mol) can improve the biological activity of proteins (Tripathi *et al.* 2013). Our results suggest that the GstD1 protein can have lesser activity or efficiency in detoxifying the alkaloid mescaline in *D. buzzatii*.

Drosophila buzzatii is recognized as a generalist for using tissue resources from more than three cactus species (Manfrin *et al.* 2006). We expected to find a better fit and interaction of GstD1 protein with alkaloids in *D. buzzatii*, which was different from our data. We also recognize that other proteins can contribute to the detoxification mechanism, such as the P450 and ADH gene families. In fact, it is recognized that genes (e.g., P450 and ADH) show differential expression when larvae use the resources of different species of palmar and columnar cacti, in *D. buzzatii* (Panis *et al.* 2016). Finally, another possibility is the presence of *GstD1* alleles with a better adjustment in the interaction of the protein with alkaloid compounds. Future studies can be carried out with population analyzes of *D. buzzatii* in different species of cacti, with an investigation of *GstD1* alleles and their protein-alkaloid interaction.

5.0 Conclusions

The complexity of chemical groups in plants can be a barrier in exchanging hosts for phytophagous insects (Becerra, 1997), which demands an adjustment in the insect detoxification system during the adaptive process. The detoxification system involves a set of proteins, such as P450s, UGTs, ADH, Esterases, and Gsts (Matzkin, 2014; Panis *et al.* 2016). The information obtained in our study suggests that the evolutionary process in GstD1 comprises the following factors: (I) the variation in the *GstD1* gene sequence allows recovering phylogenetic relationships between species of the *D. buzzatii* cluster, but with changes in branches that may be related the selection signals in this gene; (II) we found positive selection signals in codons and changes in GstD1 protein properties, with a role in structure and interaction, in different species of the cluster; (III) GstD1 proteins showed differences in the energy of interaction between GstD1 and mescaline, and with favorable interactions for *D. seriema*, *D. gouveai* and *D. koepferae*. We accept the tested hypothesis, with results that indicate selective and amino acid changes (e.g., codons 2, 117, 171, and 191) that are related to structure and interaction (e.g., α -helices, β - sheets, and solvent

accessibility) of the protein GstD1 with the substrate. In conclusion, selective and polarity changes on the active sites of the GstD1 protein can contribute to the metabolic system in *Drosophila* to the use of different hosts.

6.0 Data Availability Statement

Data are available on Genbank: <https://www.ncbi.nlm.nih.gov/genbank/>.

Author Contributions

ASS, SG, MHM designed the study; ASS collected the data; ASS, SG, DYBL and MHM analyzed the data, and wrote the manuscript.

Acknowledgment

The authors were funded by the Universidade de São Paulo (USP) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (CAPES 88882.378679/2019-1 to A.S.S.; PNPd/CAPES, 88882.317598/2019-01 to D.Y. B.L.), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (447697/2014-5 to M.H.M.) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (03/05031-0; 2011/51652-2 to M.H.). The team of the Laboratory of Evolutionary Genetics, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto-USP-SP, for their assistance in the part of molecular biology, statistical analysis and scientific discussions.

7.0 References

- Barrios-Leal, D.Y.; Neves-da-Rocha, J.; Manfrin, M.H., 2019. Demographic history of the cactophilic *Drosophila buzzatii* species cluster in open areas of South America. *J. Hered.*, 10, 22-33. <https://doi.org/10.1093/jhered/esy042>
- Becerra, J.X., 1997. Insects on plants: macroevolutionary chemical trends in host use. *Science*, 276, 253-256. <https://doi.org/10.1126/science.276.5310.218>
- Betts, M.J., Russell, R.B., 2003. Amino acid properties and consequences of substitutions. In: Barners, M.R.; Gray, I.C. (eds) *Bioinformatics for Geneticists*. Wiley, New York, 289-316. <https://doi.org/10.1002/0470867302.ch14>
- Campbell, V., Legendre, P., Lapointe, F.J., 2011. The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis. *BMC Evol. Biol.*, 11, 1-15. <https://doi.org/10.1186/1471-2148-11-64>
- Cruse, R.R., 1973. Desert Plant Chemurgy: a current review. *Econ. Bot.*, 27, 210-230. <https://doi.org/10.1007/BF02872991>
- DeBiasse, M.B., Nelson, B.J., Hellberg, M.E., 2014. Evaluating summary statistics used to test for incomplete lineage sorting: mito-nuclear discordance in the reef sponge *Callispongia vaginalis*. *Mol. Ecol.*, 23, 225-238. <https://doi.org/10.1111/mec.12584>
- DeLano, W., 2018. The PyMOL Molecular Graphics System. Available in: <http://www.pymol.org>. (accessed 20 January 2018).

- Després, L., 2019. One, two or more species? Mitonuclear discordance and species delimitation. *Mol. Ecol.*, 28, 3845-3847. <https://doi.org/10.1111/mec.15211>
- Doron-Faigenboim, A., Stern, A., Mayrose, I., Bacharach, E., Pupko, T., 2005. Selection: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics*, 21, 2101-2103. <https://doi.org/10.1093/bioinformatics/bti259>
- Eisenhaber, F., 1996. Hydrophobic regions on protein surfaces. Derivation of the salvation energy from their area distribution in crystallographic protein structures. *Protein Sci.*, 5, 1676-1686. <https://doi.org/10.1002/pro.5560050821>
- Eisenhaber, F., 1999. Hydrophobic regions on protein surfaces. *Perspect. Drug Discov Des.*, 17, 27-42. <https://doi.org/10.1023/A:1008766422140>
- Etges, W.J., Oliveira, C.C., 2014. Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. X. Age-specific dynamics of adult epicuticular hydrocarbon expression in response to different host plant. *Ecol. Evol.*, 4, 2033-2045. <https://doi.org/10.1002/ece3.108>
- Ewing, B., Green, P., 1998. Base-Calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res.*, 8, 186-194. <https://doi.org/10.1101/gr.8.3.186>
- Franco, F.F., Silva-Bernardi, E., Sene, F.M., Hasson, E., Manfrin, M.H., 2010. Intra- and interspecific divergence in the nuclear sequences of the clock gene period in species of the *Drosophila buzzatii* cluster. *J. Zool. Syst. Evol. Res.*, 48, 322-331. <https://doi.org/10.1111/j.1439-0469.2010.00564.x>
- Franco, F.F., Lavagnini, T.C., Sene, F.M., Manfrin, M.H., 2015. Mito-nuclear discordance with evidence of shared ancestral polymorphism and selection in cactophilic species of *Drosophila*. *Biol. J. Linn. Soc.*, 116, 197-210. <https://doi.org/10.1111/bij.12554>
- Franco, F.F., Manfrin, M.H., 2013. Recent demographic history of cactophilic *Drosophila* species can be related to quaternary palaeoclimatic changes in South America. *J. Biogeogr.*, 40, 142-154. <https://doi.org/10.1111/j.1365-2699.2012.02777.x>
- Fogleman, J.C., Danielson, P.B., 2001. Chemical Interactions in the Cactus-Microorganism-*Drosophila* model system of the Sonoran Desert. *Amer. Zool.*, 41, 877-889. <https://doi.org/10.1093/icb/41.4.877>
- Funk, D.J., Omland, K.E., 2003. Species-level paraphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.*, 34, 397-423. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132421>
- Gloss, A.D., Vassão, D.G., Hailey, A.L., Nelson-Dittrich, A.C., Schramm, K., Reichelt, M., Rast, T.J., Weichsel, A., Cravens, M.G., Gershenson, J., Montfort, W.R., Whiteman, N.K., 2014. Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the drosophilidae. *Mol. Biol. Evol.*, 31, 2441-2456. <https://doi.org/10.1093/molbev/msu201>
- Gompert, Z., Forister, M.L., Fordyce, J.A., Nice, C.C., 2008. Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. *Mol. Ecol.*, 17, 5231-5244. <https://doi.org/10.1111/j.1365-294X.2008.03988.x>
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, 41, 95-98. <https://doi.org/10.14601/Phytopathol.Mediterr-14998u1.29>
- Hasson, E., Panis, D.N., Hurtado, J., Mensch, J., 2019. Host plant adaptation in cactophilic species of the *Drosophila buzatii* cluster: fitness and transcriptomics. *J. Hered.*, 110, 46-57. <https://doi.org/10.1093/jhered/esy043>
- Hasson, E., Naveira, H., Fontdevila, A., 1992. The breeding sites of Argentinian cactophilic species of the *Drosophila mulleri* complex (subgenus *Drosophila-repleta* group). *Rev Chil Hist Nat.*, 65, 319-326.
- Huang, B., 2009. metaPocket: a meta approach to improve protein ligand binding site prediction. *Omics*, 13, 325-330. <https://doi.org/10.1089/omi.2009.0045>
- Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17, 754-755. <https://doi.org/10.1093/bioinformatics/17.8.754>

- Kircher, H.W., Heed, W.B., Russell, J.S., Grove, J., 1967. Senita Cactus Alkaloids: Their Significance to sonoran desert *Drosophila* ecology. *J. Insect Physiol.*, 13, 1869-1874. [https://doi.org/10.1016/0022-1910\(67\)90023-6](https://doi.org/10.1016/0022-1910(67)90023-6)
- Kokudai, C.B.S., Sene, F.M., Manfrin, M.H., 2011. Sympatry and Asymmetric Introgression between the Cactophilic Species *Drosophila serido* and *Drosophila antonietae* (Diptera: Drosophilidae). *Ann. Entomol. Soc. Am.*, 104, 434-442. <https://doi.org/10.1603/AN10096>
- López-Olmos, K., Markow, T.A., Machado, C.A., 2017. Evolution of GSTD1 in cactophilic *Drosophila*. *J. Mol. Evol.*, 84, 285-294. <https://doi.org/10.1007/s00239-017-9798-4>
- Lovell, S.C., Davis, I.W., Arendall III, W.B., Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by C α geometry: phi,psi and C-beta deviation. *Proteins*, 50, 437-450. . <https://doi.org/10.1002/prot.10286>
- Low, W.Y., Feil, S.C., Ng, H.L., Gorman, M.A., Morton, C.J., Pyke, J., McConville, M.J., Bieri, M., Mok, Y.F., Robin, C., Gooley, P.R., Parker, M.W., Batterham, P., 2010. Recognition and detoxification of the insecticide DDT by *Drosophila melanogaster* Glutathione S-Transferase D1. *J. Mol. Biol.*, 399, 358-366. <https://doi.org/10.1016/j.jmb.2010.04.020>
- Low, W.Y., Ng, H.L., Morton, C.J., Parker, M.W., Batterham, P., Robin, C., 2007. Molecular Evolution of Glutathione S-Transferase in the genus *Drosophila*. *Genetics*, 177, 363-1375. <https://doi.org/10.1534/genetics.107.075838>
- Manfrin, M.H., DeBrito, R.A., Sene, F.M., 2001. Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Ann. Entomol. Soc. Am.*, 94, 333-346. [https://doi.org/10.1603/0013-8746\(2001\)094\[0333:SAEOTD\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0333:SAEOTD]2.0.CO;2)
- Manfrin, M.H., Sene, F.M., 2006. Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genetica*, 126, 57-75. <https://doi.org/10.1007/s10709-005-1432-5>
- ^aMarkow, T.A., 2019. Ecological and evolutionary genomics: the cactophilic *Drosophila* model system. *J. Hered.*, 110, 1-3. <https://doi.org/10.1093/jhered/esy062>
- ^bMarkow, T.A., 2019. Host use and host shifts in *Drosophila*. *Curr. Opin. Insect. Sci.*, 31, 139-145. <https://doi.org/10.1016/j.cois.2019.01.006>
- Matzkin, L.M., Watts, T.D., Bitler, B.G., Machado, C.A., Markow, T.A., 2006. Functional genomics of cactus host shifts in *Drosophila mojavensis*. *Mol. Ecol.*, 15, 4635-4643. <https://doi.org/10.1111/j.1365-294X.2006.03102.x>
- Matzkin, L.M., 2008. The molecular basis of host adaptation in cactophilic *Drosophila*: molecular evolution of a Glutathione S-Transferase gene (*GstD1*) in *Drosophila mojavensis*. *Genetics*, 178, 1073-1083. <https://doi.org/10.1534/genetics.107.083287>
- Matzkin, L.M., 2014. Ecological Genomics of host shifts in *Drosophila mojavensis*. *Adv. Exp. Med. Biol.*, 781, 233-247. https://doi.org/10.1007/978-94-007-7347-9_12
- McClellan, D.A., McCracken, K.G., 2001. Estimating the influence of selection on the variable amino acid sites of the cytochrome *b* protein functional domains. *Mol. Biol. Evol.*, 18, 917-925. <https://doi.org/10.1093/oxfordjournals.molbev.a003892>
- O'Grady, P.M., Markow, T.A., 2012. Rapid morphological, behavioral, and ecological evolution in *Drosophila*: comparisons between the endemic Hawaiian *Drosophila* and the cactophilic replete species group. In: Singh, Rama S.; Xu, Jianping, Kulathinal, Rob J. Rapidly Evolving genes and genetic systems. United Kingdom: Oxford, 176-186. <https://doi.org/10.1093/acprof:oso/9780199642274.003.0018>
- Oliveira, D.C.S.G., Raychoudhury, R., Lavrov, D., Werren, J.H., 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic Wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol. Biol. Evol.*, 25, 2167-2180. <https://doi.org/10.1093/molbev/msn159>
- Oliveira, D.C.S.G., Almeida, F.C., O'Grady, P.M., Armella, M.A., DeSalle, R., Etges, W.J., 2012. Monophyly, divergence times, and evolution of host plant use inferred from a revised phylogeny of the *Drosophila replete* species group. *Mol. Phylogenetics Evol.*, 64, 533-544. <https://doi.org/10.1016/j.ympev.2012.05.012>

- Padró, J., Panis, D.N.D., Vrdoljak, J., Carmona, P.M., Colines, B., Hasson, E., Soto, I.M., 2018. Experimental evolution of alkaloid tolerance in sibling *Drosophila* species with different degrees of specialization. *Evol. Biol.*, 45, 170-181. <https://doi.org/10.1007/s11692-017-9441-8>
- Pagadala, N., Syed, K., Tuszynski, J., 2017. Software for molecular docking: a review. *Biophys Rev.*, 9, 91-102. <https://doi.org/10.1007/s12551-016-0247-1>
- Panis, D.N., Padró, J., Furió-Tarí, P., Tarazona, S., Carmona, P.S.M., Soto, I.M., Dopazo, H., Conesa, A., Hasson, E., 2016. Transcriptome modulation during host shift is driven by secondary metabolites in desert *Drosophila*. *Mol. Ecol.*, 25, 4534-4550. <https://doi.org/10.1111/mec.13785>
- Pereira, M.A.Q.R., Vilela, C.R., Sene, F.M., 1983. Notes on breeding and feeding sites of some species of the *repleta* group of the genus *Drosophila* (Diptera, Drosophilidae). *Cien. Cult.*, 35, 1313-1319.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.*, 25, 1605-1612. <https://doi.org/10.1002/jcc.20084>
- Pfeiler, E., Markow, T.A., 2011. Phylogeography of the cactophilic *Drosophila* and other Arthropods associated with cactus necrose in the Sonoran desert. *Insects*, 2, 218-231. <https://doi.org/10.3390/insects2020218>
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics*, 14, 817-818. <https://doi.org/10.1093/bioinformatics/14.9.817>
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.*, 1-5. <https://doi.org/10.1093/sysbio/syy032>
- Ré, F.C.D., Loreto, E.L.S., Robe, L.J., 2010. Gene and species trees reveal mitochondrial and nuclear discordance in the *Drosophila cardini* group (Diptera: Drosophilidae). *Invertebr. Biol.*, 129, 353-367. <https://doi.org/10.1111/j.1744-7410.2010.00207.x>
- Richards, F.M., 1977. Areas, volumes, packing, and protein structure. *Ann. Rev. Biophys. Bioeng.*, 6, 151-176. <https://doi.org/10.1146/annurev.bb.06.060177.001055>
- Saisawang, C., Wongsantichon, J., Ketterman, A.J., 2012. A preliminary characterization of the cytosolic glutathione transferase proteome from *Drosophila melanogaster*. *Biochem. J.*, 442, 181-190. <https://doi.org/10.1042/BJ20111747>
- Saisawang, C., Ketterman, A.J., 2014. Micro-Plasticity of genomes as illustrated by the evolution of Glutathione Transferases in 12 *Drosophila* species. *PloS ONE*, 9, e109518. <https://doi.org/10.1371/journal.pone.0109518>
- Sali, A., Potterton, L., Yuan, F., Vlijmen, H.V., Karplus, M., 1995. Evaluation of comparative protein modeling by MODELLER. *Proteins*, 23, 318-326. <https://doi.org/10.1002/prot.340230306>
- Santos, M.H., Hasson, E., Sene, F.M., Manfrin, M.H., 2009. *Esterase-5* gene sequences from the *Drosophila buzzatii* cluster. *Dros. Inf. Serv.*, 92, 44-56.
- Soto, E.M., Goenaga, J., Hurtado, J.P., Hasson, E., 2012. Oviposition and performance in natural hosts in cactophilic *Drosophila*. *Evol. Ecol.*, 26, 975-990. <https://doi.org/10.1007/s10682-011-9531-5>
- Soto, I.M., Carreira, V.P., Corio, C., Padró, J., Soto, E.M., Hasson, E., 2014. Differences in tolerance to host cactus alkaloids in *Drosophila koepferae* and *D. buzzatii*. *PloS ONE*, 9, e88370. <https://doi.org/10.1371/journal.pone.0088370>
- Soto, I.M., Hasson, E.R., Manfrin, M.H., 2008. Wing morphology is related to host plants in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera, Drosophilidae). *Biol. J. Linn. Soc.*, 95, 655-665. <https://doi.org/10.1111/j.1095-8312.2008.00980.x>
- Soto, I.M., Manfrin, M.H., Sene, F.M., Hasson, E., 2007. Viability and developmental time in cactophilic *Drosophila gouveai* and *Drosophila antonietae* (Diptera: Drosophilidae) are

- dependent on the cactus host. *Ann. Entomol. Soc. Am.*, 100, 490-496. [https://doi.org/10.1603/0013-8746\(2007\)100\[490:VADTIC\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2007)100[490:VADTIC]2.0.CO;2)
- Shulgin, A.T., 1979. Chemistry of Phenethylamines related to mescaline. *J. Psychedelic Drugs*, 11, 41-52. <https://doi.org/10.1080/02791072.1979.10472091>
- Shtolz, N., Mishmar, D., 2019. The Mitochondrial Genome—on Selective Constraints and Signatures at the Organism, Cell, and Single Mitochondrion Levels. *Front. Ecol. Evol.*, 7, 342, 1-9. <https://doi.org/10.3389/fevo.2019.00342>
- Starha, R., 1996. Alkaloids from the cactus genus *Gymnocalycium* (Cactaceae). *Biochem. Syst. Ecol.*, 4, 85-86. [https://doi.org/10.1016/0305-1978\(95\)00097-6](https://doi.org/10.1016/0305-1978(95)00097-6)
- Starha, R., Chybidziurova, A., Lacny, Z., 1999. Alkaloids of genus *Turbinicarpus* (Cactaceae). *Biochem. Syst. Ecol.*, 27, 839-841. [https://doi.org/10.1016/S0305-1978\(99\)00019-8](https://doi.org/10.1016/S0305-1978(99)00019-8)
- Stintzing, F.C., Carle, R., 2005. Cactus stems (*Opuntia* spp.): a review on their chemistry, technology, and uses. *Mol. Nutr. Food Res.*, 49, 175-194. <https://doi.org/10.1002/mnfr.200400071>
- Swanson, W.J., Nielsen, R., Yang, Q., 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol. Biol. Evol.*, 20, 18-20. <https://doi.org/10.1093/oxfordjournals.molbev.a004233>
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.*, 30, 2725-2729. <https://doi.org/10.1093/molbev/mst197>
- Tidon-Sklorz, R., Sene, F.M., 2001. Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia, Sér. Zool., Porto Alegre*, 90, 141-146. <https://doi.org/10.1590/S0073-47212001000100014>
- Toews, D.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.*, 21, 3907-3930. <https://doi.org/10.1111/j.1365-294X.2012.05664.x>
- Turner, W.J., Heyman, J.J., 1960. Presence of Mescaline in *Opuntia cylindrica*. *J. Org. Chem.*, 25, 2250-2251. <https://doi.org/10.1021/jo01082a623>
- Turra, A.F., Marçal, F.J.B., Baretta, I.P., Takemura, O.S., Laverde-Júnior, A., 2007. Avaliação das propriedades antioxidantes e susceptibilidade antimicrobiana de *Pereskia grandifolia haworth* (Cactaceae). *Health Sciences Archives - UNIPAR*, 11, 9-14.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22, 4673-80. <https://doi.org/10.1093/nar/22.22.4673>
- Tripathi, S.K., Muttineni, R., Singh, S.K., 2013. Extra precision docking, free energy calculation and molecular dynamics simulation studies of CDK2 inhibitors. *J. Theor. Biol.*, 334, 87-100. <https://doi.org/10.1016/j.jtbi.2013.05.014>
- Wolf, L.K., 2009. PyRx. *Chem. Eng. News*, 87.
- Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A., McClellan, D.A., 2003. TreeSAAP: Selection on Amino Acid Properties using phylogenetic trees. *Bioinformatics*, 19, 671-672. <https://doi.org/10.1093/bioinformatics/btg043>
- Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M.K., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155, 431-449. <https://doi.org/10.1093/genetics/155.1.431>

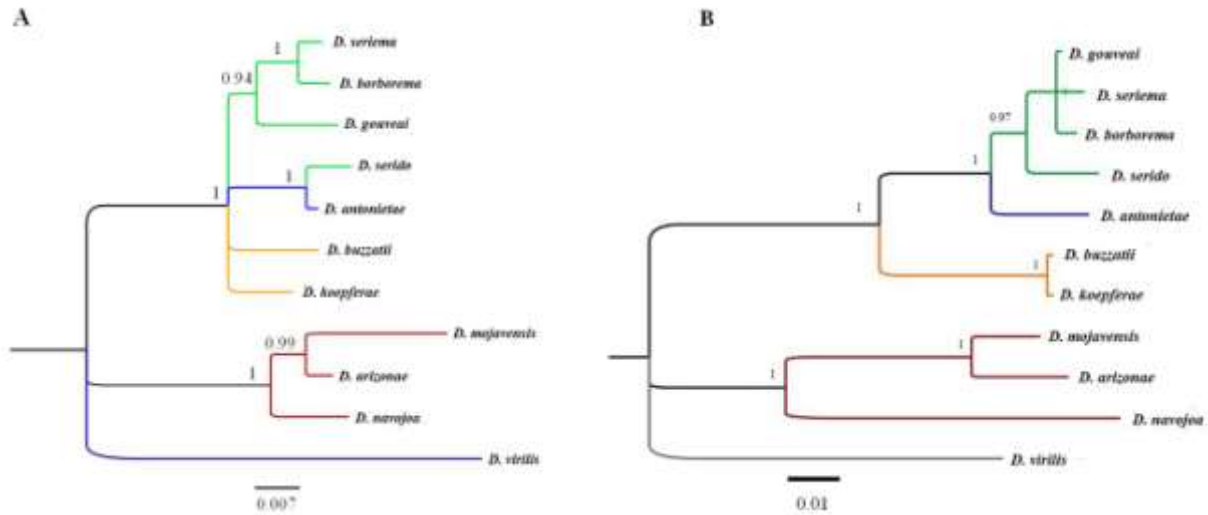


Figure 1: (A) Bayesian phylogenetic hypothesis for the *Drosophila* species based on *GstD1* DNA sequences. The coloured branches are exclusive to *D. koepferae* and *D. buzzatii* (orange), *D. antonietae* (blue), *D. serido*, *D. gouveai*, *D. borborema* and *D. seriema* (green). For the external group, *D. navojoa*, *D. arizonae* and *D. mojavenis* (red), and *D. virilis* (blue or grey). *Drosophila virilis* was used as an external group, which uses other foods, not cactus specialists. The internal nodes of the tree were organized with a posterior probability (PP 0,8), located next to each branch. The scale bar represents genetic distances calculated. (B) Bayesian phylogenetic hypothesis with mtCOI DNA sequences for the *Drosophila* species.

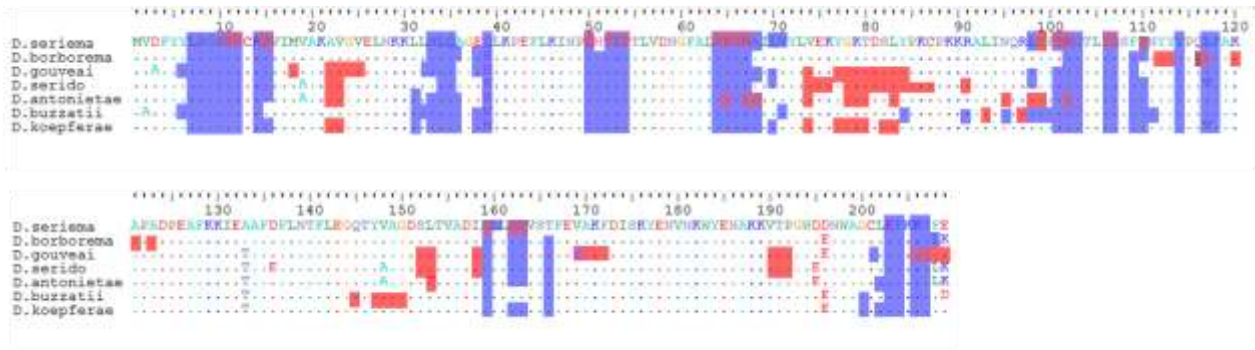


Figure 2: Multiple alignments of the primary structures of the GstD1 proteins of the *D. buzzatii* cluster species. The N-terminal region comprises amino acid residues 1-79, and the C-terminal comprises residues 86-208. The residues that make up the active sites G, in blue color, and H, in red color, were estimated. The magenta sites represent the residues with possible double functions at the active sites.

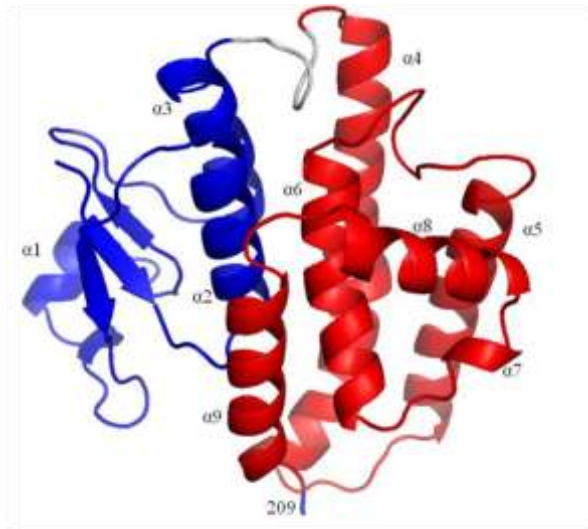


Figure 3: In silico prediction of the GstD1 protein, with the visualization of α -helices and β -sheets structures, with the representation of the N-terminal (blue color) and C-terminal (red color) regions. Residue 209 (highlighted in blue) corresponds to the tail of the protein.

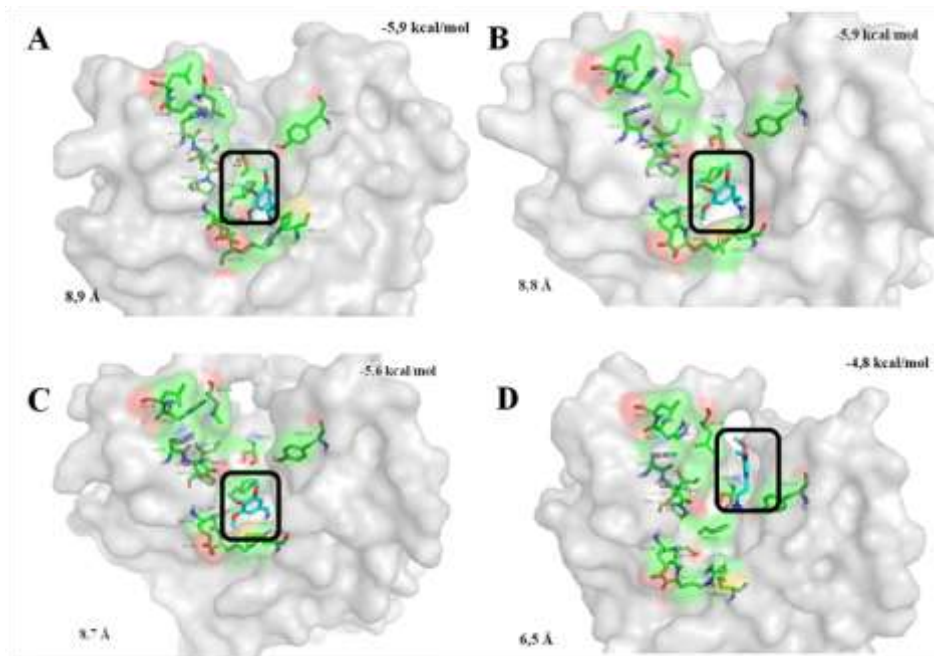


Figure 4: Molecular docking of the GstD1 protein with the mescaline ligand for *D. seriema* (A), *D. gouveai* (B), *D. koepferae* (C) and *D. buzzatii* (D). The protein is presented in gray, the G-site with its amino acid residues in green, red and blue, and the ligand in blue and red. The values of interaction energy (top, right side) and distance between residue 10 and the ligand (bottom, left side) are shown.

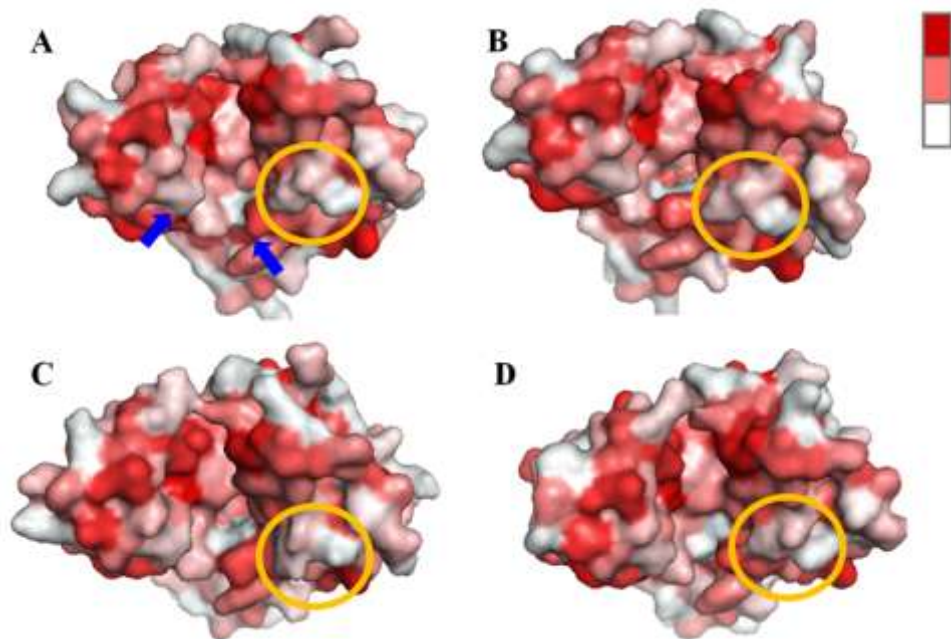


Figure 5: In silico model of the global surface of the GstD1 protein, with representation of hydrophobic regions in red and regions with less hydrophobic properties in pink and white. In visualization the proteins of *D. seriema* (A), *D. koepferae* (B), *D. gouveai* (C) and *D. buzzatii* (D), with indication of a subtle change in hydrophobic portions (yellow circle). The blue arrows indicate the basal region of the G site. The scale of hydrophobic properties represented by regions in red and regions with less hydrophobic properties in pink and white (upper part, right side).

Table 1: Results of positive selection in codons (Selecton) of *GstDI* gene sequences, which is presented in *Drosophila* species, codon position, ω values, amino acid residues and confidence interval. Val = valine; Thr = threonine; Lys = lysine.

Species	Codon position	ω values	Amino acid residues	Active site	Confidence interval
<i>D. gouveai</i> , <i>D. koepferae</i> , <i>D. seriema</i> , <i>D. serido</i> , <i>D. borborema</i> and <i>D. antonietae</i> .	2	1.6	Val	-	0.13, 4.9
<i>D. serido</i> and <i>D. koepferae</i>	117	4.8	Val	Site G*	0.46, 4.9
<i>D. gouveai</i> , <i>D. serido</i> , <i>D. antonietae</i> , <i>D. buzzatii</i> and <i>D. koepferae</i>	133	2.5	Thr	-	0.13, 4.9
All species in the cluster	171	3.0	Lys	Site H*	0.46, 4.9
All species in the cluster	191	2.8	Thr	Site H*	0.13, 4.9
<i>D. borborema</i> , <i>D. serido</i> and <i>D. antonietae</i>	209	4.6	Lys	Site H*	0.46, 4.9

Note. * Selective change in active site for some species of *Drosophila*.

Table 2: Protein level results (TreeSAAP) to physicochemical properties, categories, amino acid sites and Z-scores for positive selection of GstD1 protein.

Property	Category	amino acid residue	Z-scores
Total non-bonded energy	8	151 a 158, 187 e 191	27,483***
Composition	6	171	5,046***
Turn tendencies	6	178	2,638**
Helical contact area	6	151 a 156, 164 e 171	13,996***
Alpha-helical tendencies	6	165 a 168	11,888***
Beta-structure tendencies	6	2, 3, 19 and 137	1,878*
Isoelectric point	6	160, 164, 166 and 178	-1,677*
Solvent accessible reduction ratio	6	2, 3 and 139	-1,904*
Short and medium range non bonded energy	6	152 a 168	10,97***
Chromatographic index	6	172 to 175	-2,218*

Note. Statistical significance is represented by $p < 0,05^*$, $p < 0,01^{**}$ and $p < 0,001^{***}$. Only amino acids with selection signal were show in the table.

Figure S1: Scheme with codon selection signals and changes in physicochemical properties of GstD1 proteins.

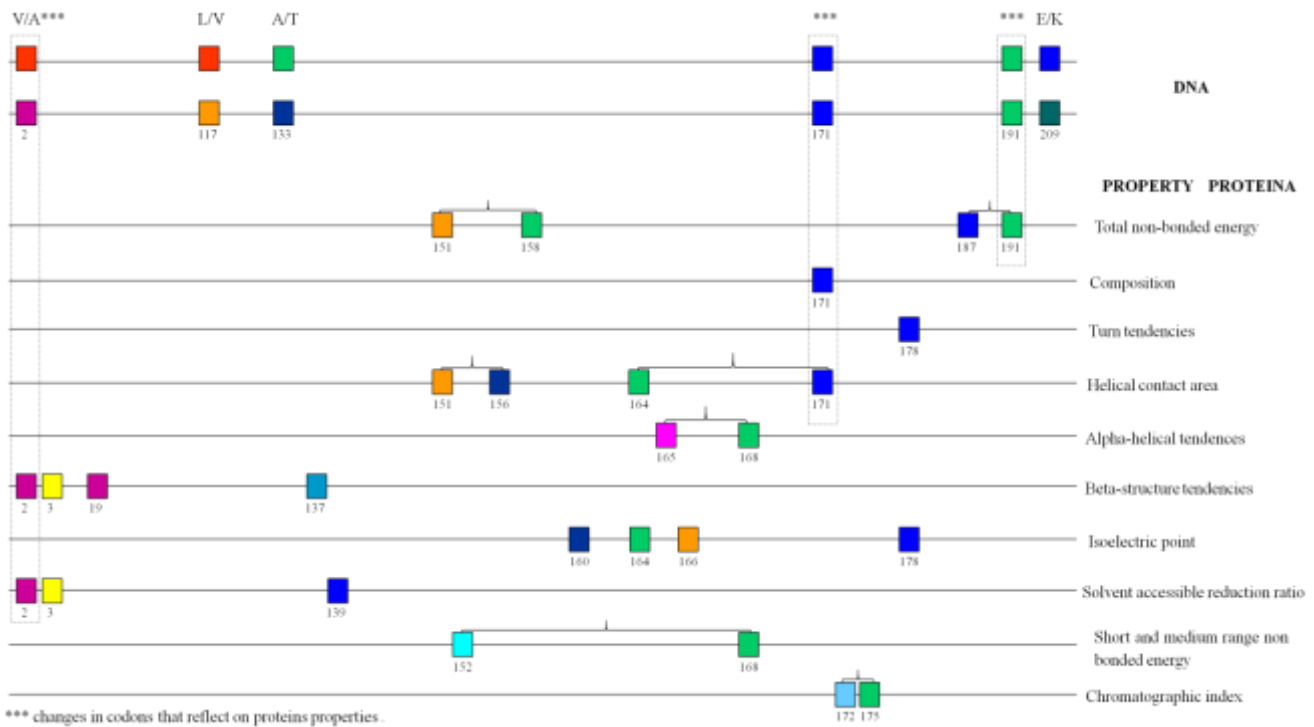


Figure S2: Multiple alignments, in the program Chimera v.1.14, between the seven tertiary proteins GstD1 of the species of the *D. buzzatii* cluster. Structural change measures between *D. buzzatii* (template) with *D. koepferae* (0.203 Å), *D. gouveai* (0.237 Å), *D. seriema* (0.189 Å), *D. serido* (0.226 Å), *D. borborema* (0.262 Å), and *D. antonietae* (0.235 Å). The colors yellow, violet and blue correspond to the grouping of *D. koepferae*, *D. gouveai*, *D. seriema*, *D. serido*, *D. borborema*, and *D. antonietae*, and *D. buzzatii* in gray. Structural changes indicated by black arrows.

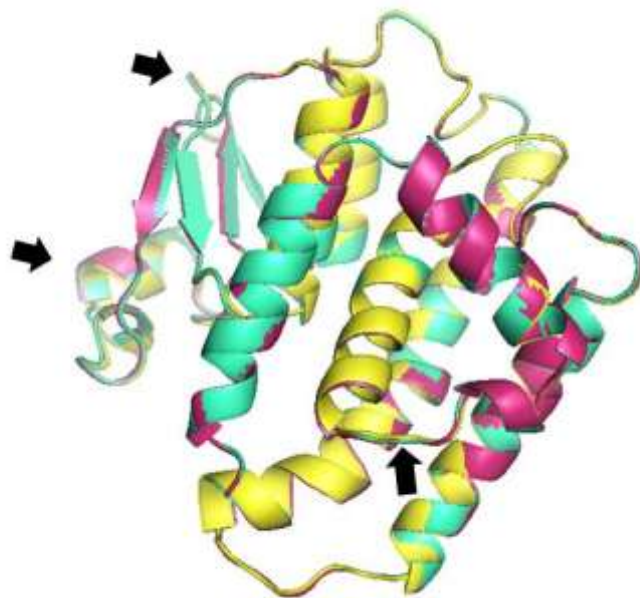


Table S1: Biological samples for species, location and number of individuals (N) for the study of *Glutathione S-Transferase D1 (GstD1)* and *Cytochrome Oxidase I (COI)* genes, and Genbank Access Number (Accessions).

Species and location	N <i>GstD1</i>	<i>GstD1</i> Accessions	N <i>COI</i>	<i>COI</i> Accessions
<i>Drosophila buzzatii</i>	1	MW805280	1	KF632603.1
<i>Drosophila koepferae</i> Famantina, Argentina (B26) (SUYDB60)	1	MW805281	1	KP404433.1
<i>Drosophila serido</i> Bom Jesus, Rio Grande do Norte, Brazil (R38) Mucugê, Minas Gerais, Brazil (N45)	1	MW805284	1	JN124588.1
<i>Drosophila seriema</i> Santana do Riacho, Minas Gerais, Brazil (N58) Cachoeira Ferro Doido, Bahia, Brazil (N42)	1	MW805283	1	JN124793.1
<i>Drosophila antonietae</i> Sertãozinho, São Paulo, Brazil (R71) N61C5	1	MW805286	1	HM371179.1
<i>Drosophila gouveai</i> Petrolina, Pernambuco, Brazil (N36) N36	1	MW805282	1	JN124690.1
<i>Drosophila borborema</i> Casa Nova, Bahia, Brazil (R41) N47	1	MW805285	1	JN124749.1
<i>Drosophila mojavensis</i> MJBC95 enmu97	1	EU079442.1	1	DQ383728.1
<i>Drosophila arizonae</i> RVSD17 1271_14	1	EU079471.1	1	MF443102.1
<i>Drosophila navojoa</i> NAV12	1	EU079446.1		
NAV.A828b			1	DQ383666.1
<i>Drosophila virilis</i> Vi-L160	1	XM_002054301.2	1	MH423345.1

Table S2: Genetic distance matrices for the *GstDI* (top) and *COI* (bottom) genes among the eleven species of *Drosophila*, calculated in the Mega v.6.0. The species are represented by D. moj (*D. mojavenensis*), D. ari (*D. arizonae*), D. nav (*D. navojoa*), D. buz (*D. buzzatii*), D. koe (*D. koepferae*), D. gou (*D. gouveai*), D. ser (*D. seriema*), D. ser (*D. serido*), D. bor (*D. borborema*), D. ant (*D. antonietae*), D. vir (*D. virilis*).

	D. moj	D.ari	D. nav	D.buz	D.koe	D.gou	D.ser	D.serid	D.bor	D.ant	D.vir
D. moj	0	0.0275	0.0372	0.0841	0.0889	0.0906	0.0809	0.0922	0.0841	0.0922	0.1067
D.ari	0.0275	0	0.0194	0.0631	0.0711	0.0711	0.0663	0.0711	0.0663	0.0695	0.0889
D. nav	0.0372	0.0194	0	0.0695	0.0744	0.0760	0.0711	0.0744	0.0711	0.0744	0.0906
D.buz	0.0841	0.0631	0.0695	0	0.0210	0.0275	0.0226	0.0323	0.0226	0.0258	0.0825
D.koe	0.0889	0.0711	0.0744	0.0210	0	0.0275	0.0194	0.0275	0.0194	0.0258	0.0841
D.gou	0.0906	0.0711	0.0760	0.0275	0.0275	0	0.0210	0.0291	0.0210	0.0242	0.0922
D.ser	0.0809	0.0663	0.0711	0.0226	0.0194	0.0210	0	0.0258	0.0064	0.0210	0.0825
D.serid	0.0922	0.0711	0.0744	0.0323	0.0275	0.0291	0.0258	0	0.0210	0.0080	0.0906
D.bor	0.0841	0.0663	0.0711	0.0226	0.0194	0.0210	0.0064	0.0210	0	0.0161	0.0825
D.ant	0.0922	0.0695	0.0744	0.0258	0.0258	0.0242	0.0210	0.0080	0.0161	0	0.0889
D.vir	0.1067	0.0889	0.0906	0.0825	0.0841	0.0922	0.0825	0.0906	0.0825	0.0889	0

	D. moj	D.ari	D. nav	D.buz	D.koe	D.gou	D.ser	D.serid	D.bor	D.ant	D.vir
D. moj	0	0.0362	0.1068	0.1376	0.1376	0.1394	0.1449	0.1358	0.1394	0.1231	0.1268
D.ari	0.0362	0	0.1105	0.1431	0.1431	0.1413	0.1467	0.1449	0.1413	0.1358	0.1304
D. nav	0.1068	0.1105	0	0.1394	0.1394	0.1286	0.1340	0.1376	0.1304	0.1304	0.1376
D.buz	0.1376	0.1431	0.1394	0	0.0000	0.0652	0.0706	0.0688	0.0652	0.0706	0.1358
D.koe	0.1376	0.1431	0.1394	0.0000	0	0.0652	0.0706	0.0688	0.0652	0.0706	0.1358
D.gou	0.1394	0.1413	0.1286	0.0652	0.0652	0	0.0054	0.0199	0.0036	0.0271	0.1322
D.ser	0.1449	0.1467	0.1340	0.0706	0.0706	0.0054	0	0.0253	0.0090	0.0326	0.1376
D.serid	0.1358	0.1449	0.1376	0.0688	0.0688	0.0199	0.0253	0	0.0235	0.0326	0.1358
D.bor	0.1394	0.1413	0.1304	0.0652	0.0652	0.0036	0.0090	0.0235	0	0.0307	0.1358
D.ant	0.1231	0.1358	0.1304	0.0706	0.0706	0.0271	0.0326	0.0326	0.0307	0	0.1340
D.vir	0.1268	0.1304	0.1376	0.1358	0.1358	0.1322	0.1376	0.1358	0.1358	0.1340	0

Table S3: Values of the dN / dS (ω) ratio for each codon in the mtCOI gene fragment analyzed using the Selecton program. AA = amino acid, IC = ω Confidence Interval.

Position	AA	ω	IC	Position	AA	ω	IC
1	P	0.0084	6.5e-10,0.057	47	P	0.0085	6.5e-10,0.057
2	G	0.0097	6.5e-10,0.057	48	D	0.0091	6.5e-10,0.057
3	A	0.0064	6.5e-10,0.057	49	M	0.0067	6.5e-10,0.057
4	L	0.0099	6.5e-10,0.057	50	A	0.0063	6.5e-10,0.057
5	I	0.0056	6.5e-10,0.057	51	F	0.0150	6.5e-10,0.057
6	G	0.0100	6.5e-10,0.057	52	P	0.0085	6.5e-10,0.057
7	D	0.0130	0.0024,0.27	53	R	0.0130	6.5e-10,0.057
8	D	0.0120	6.5e-10,0.057	54	M	0.0054	6.5e-10,0.057
9	Q	0.0180	6.5e-10,0.27	55	N	0.0100	6.5e-10,0.057
10	I	0.0056	6.5e-10,0.057	56	N	0.0097	6.5e-10,0.057
11	Y	0.0140	6.5e-10,0.057	57	M	0.0054	6.5e-10,0.057
12	N	0.0110	6.5e-10,0.057	58	S	0.0100	6.5e-10,0.057
13	V	0.0063	6.5e-10,0.057	59	F	0.0068	6.5e-10,0.057
14	I	0.0060	6.5e-10,0.057	60	W	0.0140	6.5e-10,0.057
15	V	0.0069	6.5e-10,0.057	61	L	0.0077	6.5e-10,0.057
16	T	0.0064	6.5e-10,0.057	62	L	0.0100	6.5e-10,0.057
17	A	0.0070	6.5e-10,0.057	63	P	0.0150	6.5e-10,0.27
18	H	0.0240	6.5e-10,0.27	64	P	0.0095	6.5e-10,0.057
19	A	0.0064	6.5e-10,0.057	65	A	0.0064	6.5e-10,0.057
20	F	0.0068	6.5e-10,0.057	66	L	0.0094	6.5e-10,0.057
21	V	0.0690	6.5e-10,0.27	67	T	0.0730	6.5e-10,0.27
22	M	0.0059	6.5e-10,0.057	68	L	0.0089	6.5e-10,0.057
23	I	0.0056	6.5e-10,0.057	69	L	0.0120	6.5e-10,0.057
24	F	0.0068	6.5e-10,0.057	70	L	0.0094	6.5e-10,0.057
25	F	0.0068	6.5e-10,0.057	71	V	0.0063	6.5e-10,0.057
26	M	0.0054	6.5e-10,0.057	72	S	0.0240	6.5e-10,0.27
27	V	0.0070	6.5e-10,0.057	73	S	0.0100	6.5e-10,0.057
28	M	0.0054	6.5e-10,0.057	74	M	0.0058	6.5e-10,0.057
29	P	0.0091	6.5e-10,0.057	75	V	0.0094	6.5e-10,0.057
30	I	0.0056	6.5e-10,0.057	76	E	0.0094	6.5e-10,0.057
31	M	0.0059	6.5e-10,0.057	77	N	0.0180	6.5e-10,0.27
32	I	0.0056	6.5e-10,0.057	78	G	0.0096	6.5e-10,0.057
33	G	0.0110	6.5e-10,0.057	79	A	0.0064	6.5e-10,0.057
34	G	0.0096	6.5e-10,0.057	80	G	0.0096	6.5e-10,0.057
35	F	0.0068	6.5e-10,0.057	81	T	0.0064	6.5e-10,0.057
36	G	0.0096	6.5e-10,0.057	82	G	0.0400	6.5e-10,0.27
37	N	0.0097	6.5e-10,0.057	83	W	0.0140	6.5e-10,0.057
38	W	0.0140	6.5e-10,0.057	84	T	0.0066	6.5e-10,0.057
39	L	0.0110	6.5e-10,0.057	85	V	0.0079	6.5e-10,0.057
40	V	0.0064	6.5e-10,0.057	86	Y	0.0320	6.5e-10,0.27
41	P	0.0085	6.5e-10,0.057	87	P	0.0120	6.5e-10,0.057
42	L	0.0097	6.5e-10,0.057	88	P	0.0190	6.5e-10,0.27
43	M	0.0610	6.5e-10,0.057	89	L	0.0120	6.5e-10,0.057
44	L	0.0110	6.5e-10,0.057	90	S	0.0079	6.5e-10,0.057
45	G	0.0130	6.5e-10,0.057	91	S	0.1100	6.5e-10,0.27
46	A	0.0080	6.5e-10,0.057	92	G	0.0098	6.5e-10,0.057

Continuation							
Position	AA	ω	IC	Position	AA	ω	IC
93	I	0.0060	6.5e-10,0.057	140	L	0.0094	6.5e-10,0.057
94	A	0.0095	6.5e-10,0.057	141	F	0.0068	6.5e-10,0.057
95	H	0.0140	6.5e-10,0.057	142	V	0.0088	6.5e-10,0.057
96	G	0.0097	6.5e-10,0.057	143	W	0.0140	6.5e-10,0.057
97	G	0.0097	6.5e-10,0.057	144	S	0.0088	6.5e-10,0.057
98	A	0.0081	6.5e-10,0.057	145	V	0.0065	6.5e-10,0.057
99	S	0.0079	6.5e-10,0.057	146	V	0.0063	6.5e-10,0.057
100	V	0.0063	6.5e-10,0.057	147	I	0.0056	6.5e-10,0.057
101	D	0.0091	6.5e-10,0.057	148	T	0.0065	6.5e-10,0.057
102	L	0.0094	6.5e-10,0.057	149	A	0.0064	6.5e-10,0.057
103	A	0.0064	6.5e-10,0.057	150	L	0.0096	6.5e-10,0.057
104	I	0.0056	6.5e-10,0.057	151	L	0.0110	6.5e-10,0.057
105	F	0.0110	6.5e-10,0.057	152	L	0.0087	6.5e-10,0.057
106	S	0.0078	6.5e-10,0.057	153	L	0.0081	6.5e-10,0.057
107	L	0.0100	6.5e-10,0.057	154	L	0.0096	6.5e-10,0.057
108	H	0.0150	6.5e-10,0.27	155	S	0.0079	6.5e-10,0.057
109	L	0.0095	6.5e-10,0.057	156	L	0.0096	6.5e-10,0.057
110	A	0.0099	6.5e-10,0.057	157	P	0.0085	6.5e-10,0.057
111	G	0.0096	6.5e-10,0.057	158	V	0.0084	6.5e-10,0.057
112	I	0.0056	6.5e-10,0.057	159	L	0.0094	6.5e-10,0.057
113	S	0.0078	6.5e-10,0.057	160	A	0.0110	6.5e-10,0.057
114	S	0.0079	6.5e-10,0.057	161	G	0.0150	6.5e-10,0.27
115	I	0.0056	6.5e-10,0.057	162	A	0.0064	6.5e-10,0.057
116	L	0.0094	6.5e-10,0.057	163	I	0.0060	6.5e-10,0.057
117	G	0.0110	6.5e-10,0.057	164	T	0.0065	6.5e-10,0.057
118	A	0.0079	6.5e-10,0.057	165	M	0.0054	6.5e-10,0.057
119	V	0.0063	6.5e-10,0.057	166	L	0.0089	6.5e-10,0.057
120	N	0.0097	6.5e-10,0.057	167	L	0.0094	6.5e-10,0.057
121	F	0.0068	6.5e-10,0.057	168	T	0.0066	6.5e-10,0.057
122	I	0.0072	6.5e-10,0.057	169	D	0.0190	6.5e-10,0.27
123	T	0.0064	6.5e-10,0.057	170	R	0.0130	6.5e-10,0.057
124	T	0.0064	6.5e-10,0.057	171	N	0.0097	6.5e-10,0.057
125	V	0.0069	6.5e-10,0.057	172	L	0.0096	6.5e-10,0.057
126	I	0.0063	6.5e-10,0.057	173	N	0.0097	6.5e-10,0.057
127	N	0.0100	6.5e-10,0.057	174	T	0.0072	6.5e-10,0.057
128	M	0.0054	6.5e-10,0.057	175	S	0.0079	6.5e-10,0.057
129	R	0.0130	6.5e-10,0.057	176	F	0.0073	6.5e-10,0.057
130	S	0.0079	6.5e-10,0.057	177	F	0.0100	6.5e-10,0.057
131	T	0.0710	6.5e-10,0.27	178	D	0.0230	6.5e-10,0.27
132	G	0.0096	6.5e-10,0.057	179	P	0.0190	6.5e-10,0.27
133	I	0.0056	6.5e-10,0.057	180	A	0.0064	6.5e-10,0.057
134	T	0.0066	6.5e-10,0.057	181	G	0.0110	6.5e-10,0.057
135	L	0.0093	6.5e-10,0.057	182	G	0.0100	6.5e-10,0.057
136	D	0.0210	6.5e-10,0.27	183	G	0.0150	6.5e-10,0.27
137	R	0.0140	6.5e-10,0.057	184	D	0.0200	6.5e-10,0.27
138	M	0.0055	6.5e-10,0.057				
139	P	0.0085	6.5e-10,0.057				