



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA
MOLECULAR - PPGGEv

Edyane Moraes dos Santos

**Filogenômica das moscas-das-frutas do gênero *Anastrepha* do
complexo *fraterculus* (Wiedemann) (Diptera, Tephritidae)**

São Carlos-SP
2024

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA
MOLECULAR - PPGGEv

Edyane Moraes dos Santos

**Filogenômica das moscas-das-frutas do gênero *Anastrepha* do complexo
fraterculus (Wiedemann) (Diptera, Tephritidae)**

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

Orientadora: Profa. Dra. Patrícia Domingues de Freitas

Co-Orientador: Prof. Dr. Reinaldo Alves de Brito

EDYANE MORAES DOS SANTOS

Filogenômica das moscas-das-frutas do gênero *Anastrepha* do complexo *fraterculus* (Wiedemann) (Diptera, Tephritidae)

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

Aprovado em: 03/12/2024

BANCA EXAMINADORA

Profa. Dra. Patrícia Domingues de Freitas
UFSCAR

Prof. Dra. Maura Helena Manfrin
USP

Prof. Dr. Sergio Russo Matioli
USP

Prof. Dr. Marco Antônio Del Lama
UFSCAr

Prof. Dr. Roberto Ferreira Artoni
UFSCAr

Dedicatória

Dedico esse trabalho ao apoio incondicional da minha família, à minha mãe Franci Moraes e sobretudo às minhas irmãs Edyene Moraes e Edineia Reis pelo incentivo, amor e carinho.

AGRADECIMENTOS

Ao Prof. Dr. Reinaldo Brito por abrir as portas do seu laboratório quando tantas outras se fecharam, pela orientação e condução deste trabalho. À Profa. Dra. Patrícia Domingues de Freitas por conduzir parte da minha orientação e pela gentileza que sempre me tratou.

Aos amigos do Labev pelo companheirismo, mãos estendidas e acolhimento: Mariana Bisio, Beatriz Faustino, Bruno Gomes, Profa. Dra. Andrea Peripato e Natália Rosa. Em especial às minhas amigas Maria Isabel e Samira Chahad-Ehlers, pelos conselhos, conversas e risadas durante esse longo e difícil caminho. O final dessa jornada só foi possível pela com a ajuda de vocês.

Ao Dr. Carlos Congrains por toda ajuda nas análises e nos conhecimentos em bioinformática e à confiança em fornecer os dados para essa tese e ao Tiago Belintani pelo desprendimento e altruísmo em ajudar nos momentos de dúvidas e incertezas. À Dra. Silvia Nassif Del Lama pela amizade e conselhos.

E especial agradecimento à minha família, sobretudo às minhas irmãs Edinéia Reis e Edyene Moraes, aos meus sobrinhos Ana Luiza Reis, José Daniel Reis e Mariana Lima pelo amor incondicional e recíproco, pelo apoio nos momentos em que nem eu mesma acreditava mais em mim.

À secretária do Programa Ivanildes Menezes e à Profa. Dra. Fernanda Aníbal por todo apoio, ajuda e conselhos. Minha eterna gratidão!

À toda coordenação e aos professores do Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular e à Capes pela concessão da bolsa de pesquisa.

Epígrafe

“Não chores, meu filho;
Não chores, que a vida
É luta renhida: Viver é lutar.
A vida é combate,
Que os fracos abate,
Que os fortes, os bravos
Só pode exaltar”

Canção do Tamoio - Antônio Gonçalves Dias

RESUMO

O desafio de definir relações taxonômicas entre espécies crípticas ainda é uma questão em aberto e pode resultar em classificações tendenciosas e imprecisas. Isso ocorre porque essas espécies apresentam pouca variação fenotípica, mesmo diante de uma divergência evolutiva significativa. A análise filogenômica de espécies crípticas, com ênfase particular nas moscas-das-frutas, desempenha um papel crucial no âmbito taxonômico e evolutivo, além de ser uma ferramenta relevante no contexto do manejo de pragas. As moscas-das-frutas (Diptera: Tephritidae) do gênero *Anastrepha* são espécies de grande importância econômica em regiões tropicais e subtropicais, causando perdas extensivas à fruticultura devido a sua polifagia e fácil adaptação a diferentes hospedeiros. A família Tephritidae está entre as mais numerosas famílias de Diptera e inclui aproximadamente 4.000 espécies distribuídas em 500 gêneros, com cerca de 1.500 espécies de moscas-das-frutas associadas a prejuízos anuais. Devido às notáveis similaridades morfológicas entre as espécies em questão, a identificação destas encontra-se intrinsecamente vinculada ao uso de estudos morfométricos, que, até o momento, ainda se encontram em estágios iniciais de investigação para definir grupos taxonômicos e linhagens. A delimitação de espécies tem se mostrado um desafio complexo dentro deste grupo, resultando na formação de clados imprecisos e ramos curtos, sobretudo no caso da entidade específica de *A. fraterculus*, cujos morfotipos continuam a gerar incertezas filogenéticas. Ao realizar uma análise abrangente de um grande conjunto de genes ortólogos, utilizamos diversas estratégias para selecionar loci que permitem a construção de árvores filogenéticas com um conjunto menor de marcadores moleculares, que ainda permitem a identificação de espécies de *Anastrepha*. Ao analisar taxas evolutivas, índices de diversidade, hipóteses de classificação de linhagem incompleta, além de fenômenos de introgressão e fluxo gênico entre as espécies, buscamos esclarecer melhor os eventos históricos que tornaram a diferenciação de linhagens e espécies dentro do gênero *Anastrepha* um desafio, mesmo diante de um extenso conjunto de dados disponíveis. Neste contexto, adotamos uma abordagem que utiliza diversas ferramentas de bioinformática para identificar regiões no genoma que possam fornecer novos insights sobre as divergências observadas nas árvores filogenéticas de *Anastrepha*.

Palavras-chave: genômica, genes ortólogos, introgressão, bioinformática, filogenia

ABSTRACT

The challenge of defining taxonomic relationships among cryptic species is still an open question and can result in biased and imprecise classifications. This is because these species show little phenotypic variation despite significant evolutionary divergence. Phylogenomic analysis of cryptic species, with particular emphasis on fruit flies, plays a crucial role within the taxonomic and evolutionary scope, in addition to serving as a tool of significant relevance in the context of pest management. Fruit flies (Diptera: Tephritidae) of the genus *Anastrepha* are economically important species in tropical and subtropical regions, causing extensive fruit crops losses due to their polyphagy and easy adaptation to different hosts. The family Tephritidae is among the largest families of Diptera and includes approximately 4,000 species across 500 genera, with around 1,500 species of fruit flies associated to annual losses. Due to the remarkable morphological similarities between the species in question, their identification is intrinsically linked to the use of morphometric studies, which, to date, are still in the initial stages of investigation for defining taxonomic groups and lineages. Species delimitation has proven to be a complex challenge within this group, resulting in the formation of imprecise clades and short branches, especially in the case of the specific entity of *A. fraterculus*, whose morphotypes continue to pose phylogenetic uncertainties. Through a comprehensive analysis of a large set of orthologous genes, we used several strategies to seek efficient subsets of select loci that are still able of accurately infer phylogenetic trees and identify different lineages of *Anastrepha*. By analyzing evolutionary rates, diversity indices, hypotheses of incomplete lineage classification, as well as the phenomena of introgression and gene flow between species, we aim to better understand the historical events that have made the differentiation of lineages and species within the genus *Anastrepha* a challenging, despite the extensive set of available data. In this context, we adopted an approach that uses several bioinformatics tools to identify regions in the genome that can provide additional insights into the divergences observed in the phylogenetic trees of *Anastrepha*.

Keywords: genomics, orthologous genes, introgression, bioinformatics, phylogeny

LISTA DE FIGURAS

1. Introdução geral

Figura 2: Variações do acúleo para a complexo <i>fraterculus</i>	2
Figura 1: Representação esquemática do ciclo biológico das moscas-das-frutas do Complexo <i>fraterculus</i>	3

Capítulo 1: Comparative subsampling of phylogenomic informative loci to delimit species in *Anastrepha*

Figure 1: Spatial configuration of the sampling distribution of the 15 species of the genus <i>Anastrepha</i> , with samples distributed in North and South America.....	79
Figure 2: Multispecies coalescent species tree of <i>Anastrepha</i> species and five other Tephritidae species as outgroups based on 3,170 genes inferred in ASTRAL.....	86
Figure 3: Maximum likelihood inference from concatenated 3,168 genes of <i>Anastrepha</i> species with focus on the <i>fraterculus</i> group using <i>A. striata</i> and <i>A. bistrigata</i> as outgroups inferred in RAxML.	87
Figure 4: Summary of gene concordance and conflict based on PhyParts analysis against the species tree inferred on ASTRAL based on the <i>Anastrepha</i> dataset.....	88
Figure 5: Visualization of phylogenetic congruence between the topology produced by all loci.....	90
Figure 6: Chromomap distribution of loci across the <i>A. ludens</i> genome.....	93
Figure 7: Summary of concordance and conflict inferred on Astral and plotted with Phyparts and Phypartspiecharts.....	94
Figure 8: Summary of concordance and conflict inferred on Astral and plotted with Phyparts and Phypartspiecharts.....	95
Figure 9: Violin plots of values of A) dN/dS , C) Evolutionary rate, E) Saturation and G) Treeness for the <i>Anastrepha</i> dataset, and B) dN/dS , D) Evolutionary rate, F) Saturation and H) Treeness for the <i>fraterculus</i> dataset.....	98

LISTA DE TABELAS

Capítulo 2: Use of different strategies to subsample genomic data to investigate phylogenetic relationships in *Anastrepha*

Table 1 – Sampling information.77

Table 2 – Average values for different evolutionary and population parameters across the whole dataset and in each reduced subset.....97

SUMARIO

1. Introdução geral	1
1.1. Complexo <i>fraterculus</i> , Espécies crípticas, Ecologia e combate às Moscas-das-frutas.....	2
1.2 Filogenia do Complexo <i>fraterculus</i>	5
1.3 Influência da introgressão na evolução das <i>Anastrephas</i>	7
2. Objetivos.....	10
2.1 Objetivo geral.....	10
2.2 Objetivos específicos.....	10
Hipótese geral.....	10
Referências.....	11

1º Capítulo: Population Differentiation with Introgression

Resumo.....	27
Abstract	27
1. Introdução	28
2. Objetivos.....	30
2.1 Objetivo Geral.....	30
2.2 Objetivos específicos.....	30
3. Bases Genômicas da Diferenciação Populacional e Especiação	31
4. A Interação entre Fluxo Gênico e Deriva Genética na Estrutura Populacional.....	32
5. Dinâmicas de Estrutura Populacional com Contato Secundário	35
6. Padrões Genômicos de Introgressão.....	37
7. Interação entre Fluxo Gênico, Seleção e Estrutura Populacional	39
8. Efeitos Antropogênicos na Introgressão.....	42
9. Processos de Diferenciação e Fluxo Gênico em Ambientes Neotropicais.....	45
Conclusão	49
References	51

2º Capítulo: Use of different strategies to subsample genomic data to investigate phylogenetic relationships in *Anastrepha*

Resumo.....	72
Abstract.....	72
1. Introduction.....	74
2. Materials and methods.....	76
2.1 Data processing.....	76
2.2 Species Trees Inference and Concordance Analysis.....	79
2.3 Search for Phylogenetically Informative Genes.....	80
2.4 Incomplete Lineage Sorting (ILS) and Introgression.....	82
2.5 Phylogenetic and evolutionary parameters.....	82
2.6 Tests of tree compatibility.....	84
2.7 Gene Mapping.....	84
3. Results.....	84

3.1 Species tree inference and concordance analyses.....	85
3.2 Phylogenetic Inferences and the Search for Informative Genes.....	89
3.3 Association of phylogenetic parameters to reduced subsets.....	96
4. Discussion.....	98
References.....	104
Material Suplementar.....	113

1. Introdução geral

As moscas-das-frutas são espécies fitófagas que pertencem à ordem Diptera, subordem Brachycera, infraordem Muscomorpha, seção Schizophora, superfamília Tephritoidea e família Tephritidae (McAlpine, 1989). Os Tephritidae são representados por mais ou menos 4.700 espécies distribuídas em 500 gêneros, porém a classificação de algumas espécies dessa família ainda é considerada insatisfatória por problemas na taxonomia e identificação das espécies, devido a dificuldades práticas, pois a identificação é baseada exclusivamente na morfologia das fêmeas (Selivon, Perondini, Morgante; 2005; Han e Ro, 2009; Norrbom et al. 1999, 2014, 2004; Perre et al. 2016).

As principais pragas da fruticultura brasileira são caracterizadas por sua polifagia, o que lhes permite infestar uma ampla variedade de cultivos frutíferos de elevado valor comercial. Esta infestação resulta em consideráveis prejuízos para a economia global (White e Elson-Harris, 1992; Norrbom, 2004; Pereira-Rêgo *et al.* 2011; Bachmann *et al.*, 2019). O Brasil se configura como o terceiro maior produtor mundial de frutas do mundo (INCT, 2009; Brazilian Horti and Fruti Yearbook 2020, Abrafrutas 2020) e o parasitismo dessas espécies de dípteros frugívoros, especialmente dos Tephritídeos, causam grandes prejuízos econômicos anuais (Garcia e Norrbom, 2011; Morais *et al.* 2021).

Entre os Tephritídeos que causam os maiores prejuízos em frutíferas no Brasil e no mundo, com restrições quarentenárias, temos espécies dos gêneros *Anastrepha* (Schiner), *Ceratitis capitata* (Wiedemann) e *Bactrocera* (Drew & Hancock) (Zucchi, 2000; Uchoa, 2012; De Oliveira *et al.* 2019; Gallo-Franco e Toro-Perea, 2020; Troya et al. 2020). O gênero *Anastrepha* é endêmico das Américas com distribuição biogeográfica na região Neotropical com mais de 300 espécies identificadas (Aluja, 1994; Norrbom 2004a; Norrbom *et al.* 1999, 2012, 2018; Raga *et al.* 2004; Norrbom e Korytkowski 2009, 2011; Troya et al., 2020), com sete destas de grande importância econômica: *Anastrepha grandis* Macquart, 1846, *A. fraterculus* Wiedemann, 1830, *A. obliqua* Macquart, 1835, *A. pseudoparallela* Loew, 1873, *A. sororcula* Zucchi, 1979, *A. striata* (Schiner, 1868) e *A. zenildae* Zucchi, 1979 (Zucchi, 2000).

1.1. Complexo *fraterculus*, Espécies crípticas, Ecologia e combate às Moscas-das-frutas

O agrupamento taxonômico conhecido como *fraterculus* apresenta evolução recente e inclui cerca de 40 entidades taxonômicas (Norrbon e colaboradores, 2012; Fontoura et al., 2018; Rodriguez *et al.* 2018). São definidas como espécies crípticas que nada mais são do que táxons com grandes semelhanças morfológicas, porém, evidências indicam que estão em diferentes caminhos evolutivos, gerando controvérsias taxonômicas e evolutivas (Bickford *et al.* 2007; Struck *et al.* 2018). Dessa forma, o complexo *fraterculus* compõe um conjunto de espécies crípticas que são classificadas dentro de uma única espécie, embora caracteres morfológicos distintos discordem em sua identificação, levando a errôneas classificações, o que tem dificultado a filogenia desse grupo (Barr, 2005; Selivon, Perondini e Morgante *et al.* 2005a). Os estudos de espécies crípticas servem como ponte entre a multidisciplinaridade com suporte dos conhecimentos da taxonomia, padrões filogenéticos, estudos de ecossistemas e ajuda a entender as tendências de processos macro e microevolutivos das espécies (Pante, et al., 2015; Nygren, 2013; Bickford *et al.* 2007; Struck *et al.* 2018). Para as espécies de *A. fraterculus* ainda faltam recursos diagnósticos suficientes para diferenciá-los, mesmo com grande esforço mundial, advindo de diferentes fontes de conhecimentos (Prezotto *et al.* 2019).

Ainda assim, medidas morfométricas de asas, cabeça, tórax e acúleo da fêmea são caracteres morfológicos comumente usados na identificação das espécies intimamente relacionadas de *Anastrepha* (Hernández-Ortiz *et al.* 2004, 2015). Mesmo sendo caracteres que sofrem grande plasticidade, há forte variação sazonal nas populações devido a fatores bióticos e abióticos em escalas regionais e locais (Begon *et al.* 1996), essa plasticidade é observada com mais precisão na estrutura do acúleo demonstrando grande variação morfológica no complexo *fraterculus* (Lopes *et al.* 2013) (Figura 1).

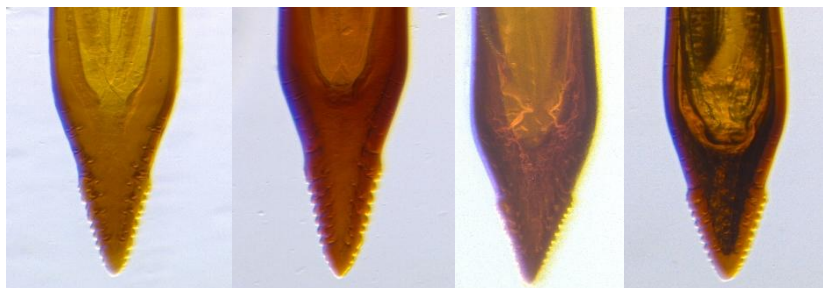


Figura 1: Variações do acúleo para a complexo *fraterculus* (Fonte: própria autoria)

Sobre a biologia do desenvolvimento, os indivíduos adultos de espécies do complexo *fraterculus* medem cerca de 6,5 mm a 7mm de comprimento, de coloração amarela e asas com faixas de cor marrom (Gallo *et al.* 2002). O processo de acasalamento ocorre mediante a prática de cortejo pelo macho à fêmea, ocorrendo comumente por meio do fenômeno conhecido como lekking. Esse padrão comportamental engloba uma série de exhibições visuais e acústicas, além da atração mediada por feromônios. O sucesso reprodutivo da espécie pode ser determinado pela compatibilidade ou incompatibilidade reprodutiva entre os indivíduos envolvidos (Segura *et al.* 2007; Gomez Cendra *et al.* 2011; Dias, 2016).

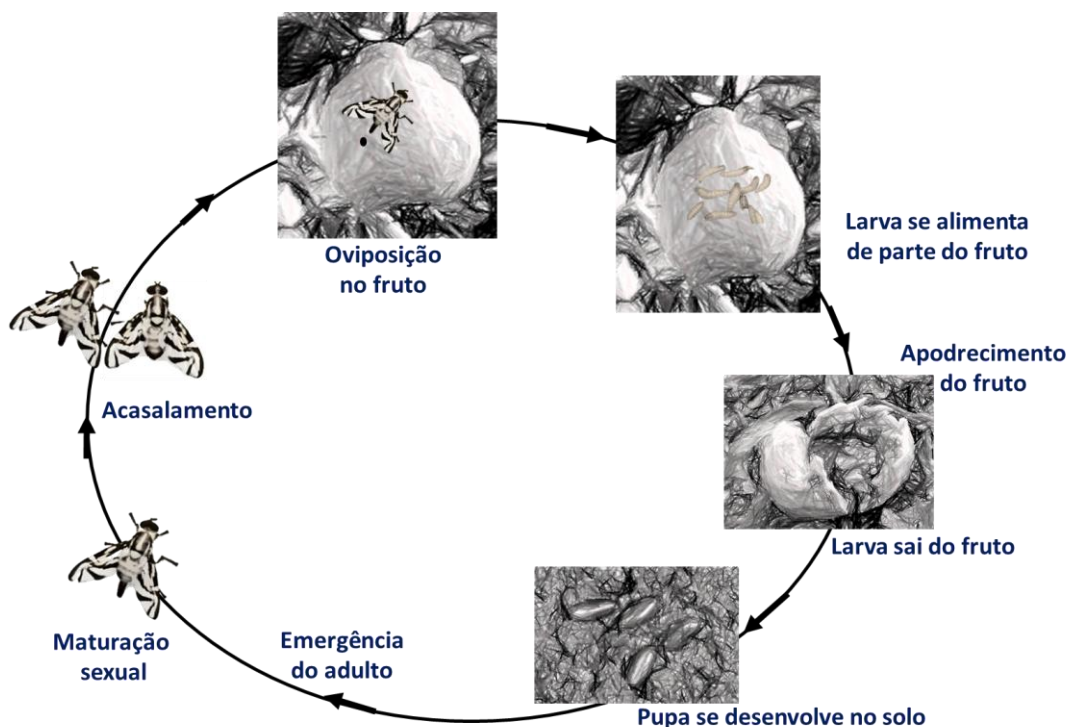


Figura 2: Representação esquemática do ciclo biológico das moscas-das-frutas do Complexo *fraterculus*.

Após o acasalamento, fêmeas escolhem frutos para oviposição, preferindo as verdes ou em processo de amadurecimento. As larvas em condições adequadas se desenvolvem na polpa do fruto, induzindo a podridão e deformação do fruto e sua decomposição prematura (Aguiar-Menezes *et al.* 2004). As larvas se desenvolvem no endocarpo do fruto e, após 6 a 11 dias, deixam o fruto para empupar no solo. Após 9 a 11 dias, ocorre a emergência dos adultos reiniciando o ciclo (Figura 2). A média de duração do ciclo em espécies do complexo dura entre 24 a 39 dias (Souza Filho *et al.* 2004). O número de hospedeiros varia de acordo com a espécie de *Anastrepha*, que abrange desde espécies que podem ser monófagas (preferência por uma única espécie vegetal), estenófagas

(plantas parasitadas são do mesmo gênero), oligófagas (parasitam vários gêneros da mesma família) e polífagas que se alimentam de espécies vegetais de famílias diversas (Zucchi, 2000). As moscas-das-frutas exploram aproximadamente 80 espécies vegetais de 15 famílias diferentes (Malavasi e Zucchi, 2000; Gómez-Cendra *et al.* 2016). Devido à sua capacidade de parasitar uma ampla diversidade de plantas, o controle de espécies-praga desse gênero representa um desafio significativo. Isso ocorre devido à limitada eficácia das técnicas tradicionais empregadas para seu manejo.

A forma mais conservadora e de maior difusão de combate a essas espécies-praga é o emprego principalmente de inseticidas, produtos fitossanitários de aplicação seletiva, associado a um atrativo alimentar em forma de isca tóxica. Esse controle químico apesar de efetivo, causa grandes consequências no âmbito ambiental e da saúde humana (Carvalho *et al.* 2000; Nava e Botton, 2010). O uso de agentes biológicos como espécies parasitóides *Pachycrepoideus viriendemmiae* (Hymenoptera: Pteromalidae), *Diachasmimorpha longicaudata* (Hymenoptera: Braconidae), *Doryctobracon brasiliensis* e *D. areolatus* (Hymenoptera: Braconidae), que predam as larvas das moscas-das-frutas, são usadas como controle biológico de *A. fraterculus* em plantações de *citrus* no Brasil (Zucchi e Canal, 1996; Paranhos *et al.* 2003; Poncio, 2015; Gonçalves, 2016; Paranhos *et al.* 2019). Outra técnica utilizada de controle foi adaptada da mosca varejeira *Cochliomyia hominivorax* (Coquerel), chamada técnica do inseto estéril (TIE). Trata-se de em um método genético usado em programas de manejo integrado a outras formas de combate a pragas (Knipling, 1959, 1985). Através desta, machos criados em larga escala são esterilizados por emissão de radiação gama. Uma vez liberados de forma inundativa, o macho ao copular com as fêmeas, não gera descendentes por serem estéreis, o que onera e diminui consideravelmente as populações das moscas-das-frutas em número de indivíduos (Montoya *et al.* 2010; Ramírez-Santos *et al.* 2021).

Por tudo isso, o abandono do emprego tradicional de controle de pragas e aplicações mais modernas e co-associadas de técnicas que causem menor prejuízo ambiental fazem parte do cerne da fruticultura moderna mundial (Corsato, 2004). Faz-se necessário, para um melhor direcionamento dos programas de manejo e controle, conhecer os processos de especiação, distribuição espacial e sua biologia evolutiva desse complexo de espécies, que fornecerá subsídios para um combate mais direcionado e efetivo (Oliveira, 2009).

1.2 Filogenia do Complexo *fraterculus*

A espécie nominal *Anastrepha fraterculus* é reconhecida como um complexo de espécies crípticas com oito morfotipos identificados distribuídos nas Américas Central e do Sul (Prezotto *et al.* 2019). Estudos relacionados a genética de populações (Steck, 1993; Alberti *et al.* 1999; Selivon, Perondini e Morgante, 2005; Selivon e Perondini, 2007; Sutton *et al.* 2015; Barr *et al.* 2017), morfometria (Norrbon *et al.* 1999; Selivon *et al.* 2005b; Perre *et al.* 2014; Hernández-Ortiz *et al.* 2004, 2012, 2015), citogenética (Selivon, Perondini e Morgante, 2005; Basso *et al.* 2003; Goday *et al.* 2006; Cáceres *et al.* 2009; Giardini *et al.* 2020), estudos moleculares (Morgante *et al.* 1980, Steck e Sheppard, 1993; Smith-Caldas *et al.* 2001; McPheron *et al.* 1999; Rocha e Selivon, 2004; Selivon *et al.* 2005; Barr e McPheron, 2005; Boykin *et al.* 2006; Manni *et al.* 2015; Isaza *et al.* 2017; Passos *et al.* 2018; Congrains *et al.* 2021, 2023) e outros estudos integrativos que tem sido conduzidos como parte de um esforço mais amplo para aprimorar a compreensão das interações entre as espécies intimamente relacionadas. Apesar da aplicação de uma abordagem multidisciplinar, subsiste a dificuldade em estabelecer a taxonomia do grupo com precisão.

Processos evolutivos como introgressão de genes que ocorrem entre espécies-irmãs, também influenciam diretamente na filogenia do grupo como encontrado em Diaz *et al.* (2018) entre as espécies *Anastrepha fraterculus*, *A. sororcula* e *A. obliqua*, sugerindo que as mesmas estão em processo incipiente de especiação. Mesmo trabalhos que utilizaram um grande conjunto de dados genômicos como Mengual *et al.* (2017) e Congrains (2021, 2023), ainda não foram suficientes para elucidar a filogenia do grupo, apontando regiões genômicas espécies-específicas que as diferenciam. Como uma forma de resolver essa questão taxonômica, abordagens multidisciplinares mais recentes, envolvendo dados morfológicos, fisiológicos, comportamentais e genéticos, mostram que a resolução do complexo de espécies pode determinar os limites entre as espécies, se baseando em diferentes fontes de conhecimento e de forma independente (Matthews e Matthews, 1978; De Queiroz, 2007; Schutze *et al.* 2014; Dias *et al.* 2016, Congrains 2021, 2023).

Atualmente, estudos filogenéticos e filogenômicos, com uma grande gama de dados genômicos, aliados a novas tecnologias de análises genéticas, permitem novas inferências de diferenciação entre grupos animais de táxons próximos (Johnson *et al.* 2018). Diversos estudos moleculares de diferentes partes do genoma tentaram elucidar a filogenia do grupo *fraterculus*, como estudos de Steck e Sheppard, 1993; McPheron *et al.* 1999; Smith-Caldas *et al.* 2001; Ludeña *et al.* 2010; Mengual *et al.* 2017, Congrains 2021, 2023). Porém, a filogenia para algumas espécies

do grupo *fraterculus* ainda apresenta filogenias mal resolvidas, sobretudo por se tratar de espécies de divergência recente e que passaram por rápidos processos de radiação, além da grande semelhança morfológica (Barr e McPheron, 2006, Condon *et al.* 2008, Clarke *et al.* 2005, Silva e Barr, 2008, Virgilio *et al.* 2009, 2013; Congrains *et al.* 2021).

Outra questão a ser ainda esclarecida é sobre a monofilia do grupo e o quanto representa uma linhagem evolutiva. Estudos morfométricos (Norrbon *et al.* 1999) e moleculares (Han e McPheron, 1997; McPheron *et al.* 1999); Han e Ro, 2009; Segura *et al.* 2006; Congrains *et al.* 2023) confirmam a monofilia, porém outros dados genéticos considerando diferentes marcadores (McPheron *et al.* 1999; Barr *et al.* 2005) rejeitam essa hipótese, assim como análises isoenzimáticas (Morgante *et al.* 1980) e de marcadores nucleares (Ruiz *et al.* 2007a; 2007b). A parafilia do grupo é apontada em estudos de McPheron *et al.* (1999), Norrbom *et al.* (1999b), Barr *et al.* (2005), Mengual *et al.* (2017). A relação entre espécies do gênero *Anastrepha* como *A. grandis*, *A. serpentina*, *A. striata* e *A. pickeli* mostram uma relação monofilética e parafilética entre as espécies *A. fraterculus*, *A. obliqua*, *A. distincta* e *A. bahiensis* (Velasco-Cuervo *et al.* 2019). Em *A. obliqua*, a monofilia foi bem caracterizada através de marcadores moleculares nucleares (Passos *et al.* 2018, Scally *et al.* 2016). A monofilia do grupo *fraterculus* foi comprovada por dados de transcriptoma por Congrains *et al.* (2021) e a parafilia entre espécies de seis morfotipos de *A. fraterculus* distribuídas nas Américas do sul e norte por Congrains *et al.* (2023). A definição problemática da taxonômica dos estudos citados pode ser influenciado pela amostragem ou caracteres insuficientes (Silva e Barr, 2006). Porém, em espécies de divergência recente e com rápida diferenciação, além de fatores evolutivos como fluxo gênico interespecífico e deriva genética que possibilitam filogenias imprecisas (Giarla e Esselstyn, 2015; Li *et al.* 2019).

Portanto, a investigação da biologia dessa praga, podendo abranger vários aspectos biológicos, com ênfase na investigação molecular, pode fornecer insights cruciais para o desenvolvimento de métodos eficazes de manejo e controle, considerando o seu significativo impacto econômico. Ademais, o entendimento da distribuição espacial e populacional das espécies, assim como de seus territórios, é fundamental para otimizar as estratégias de manejo (Park e Tollefson, 2006; Nicácio *et al.* 2019).

1.3 Influência da introgressão na evolução das *Anastrephas*

O processo de introgressão pode ser definido como um fenômeno da genética evolutiva no qual ocorre a transferência de genes a partir do cruzamento entre espécies diferentes e posterior retrocruzamento, tornando-se fonte importante de variação genética e do surgimento de novas espécies e fenótipos. Os alelos podem ser fixados uma vez que conferem alguma vantagem adaptativa, ou não são suficientemente deletérios (Anderson, 1949; Rieseberg e Carney, 1998; Twyford e Ennos, 2012; Arnold e Martin, 2009; Suarez-Gonzalez *et al.* 2018). O cerne de novos estudos sobre a hibridação concentra-se em investigar o processo do fluxo gênico nos processos adaptativos em espécies híbridas e em novas espécies (Arnold, 2006). Mallet (2005) afirma que a introgressão pode ser definido como um processo de fluxo gênico entre espécies onde há “invasão do genoma”, sobretudo quando há sobreposição de nichos e distribuição espacial facilitando assim, o fluxo gênico entre diferentes espécies.

Ao refletir sobre a natureza da origem da hibridação introgressiva, Willis *et al.* (2012) e Rhymer e Simberloff, (1996) relatam que esse evento pode acontecer normalmente na natureza, desde que esteja presente em habitats não perturbados, refletindo um dos importantes processos evolutivos. Processos contínuos de retrocruzamento podem mascarar os *loci* posteriormente introgressados, e mesmo as tecnologias mais avançadas de sequenciamento podem detectar falsos positivos (Huerta-Sánchez *et al.* 2014; Hibbins e Hahn, 2022). Os processos de intercruzamento podem aumentar ou diminuir a frequência de alelos adaptativos e não há garantia que esses alelos irão subsistir na espécie receptora de genes introgrididos (Martin e Jiggins 2017; Dagilis *et al.* 2021). Identificar esses loci que foram herdados não é uma tarefa muito fácil. Um grande conjunto de dados genômicos está sendo usado de genomas ou transcriptomas para estimar árvores gênicas em comparação à árvore de espécies (Fontaine *et al.* 2015; Edelman *et al.* 2019; Hibbins e Hahn, 2022).

A introdução de novos genes em espécies com importância econômica, sobretudo em espécies-praga, pode resultar em adaptações evolutivas, como a conquista de novos habitats, aumento no número de hospedeiros a serem explorados e criação de biótipos ou raças (Lewontin e Birch 1966; dos Santos *et al.* 2001; Arcella *et al.*, 2015). Mesmo que tenha sido reportado uma menor aptidão em híbridos, alguns que fogem dessa janela de aptidão podem se adaptar ao ambiente e dar origem a novas espécies e linhagens (Levin *et al.* 1996; Egan e Funk, 2009; Schwarz

et al. 2005; Abbott *et al.* 2013). Caso esse, que pode ser espelhado em espécies-pragas da família Tephritidae, espécies polífagas, cuja semelhança morfológica dificulta o processo de identificação. Mesmo que tenha sido reportado que algumas espécies apresentem diferentes horários de cópula e diferença na vocalização dos machos na atração das fêmeas, torna-se uma barreira de isolamento pré-zigótica (Malavasi *et al.* 1983; Morgante *et al.* 1993). Como algumas espécies da família Tephritidae possuem divergência recente sua filogenia ainda é de certo conturbada, essas discordâncias podem ter como uma das fontes a introgressão genética (Degnan e Rosenberg, 2009; Giarla e Esselstyn, 2015).

O gênero *Anastrepha*, sobretudo espécies do complexo *fraterculus*, possui um histórico de divergência recente. Por essa especificidade, Mallet (2005) afirma que espécies de evolução recente possuem barreiras de isolamento mais fracas que facilitam as trocas gênicas e processos introgressivos, processo este que pode levar à diminuição da divergência interespecífica entre as espécies. Árvores filogenéticas conflitantes podem ser indícios de sinais de introgressão (Baack e Rieseberg, 2007), e esses conflitos na filogenia de *Anastrepha* foram vastamente documentados para dados genômicos e mitocondriais (Congrains *et al.* 2021; 2023), uma vez que se trata de um complexo de espécies crípticas (Manni *et al.* 2015). Olhando evolutivamente para a ordem Diptera, além de modelos para estudos na genética comparativa, ecológicos e estudos de evolução, também foram encontradas evidências de introgressão em *Anopheles* (Choochote *et al.* 2014; Fontaine *et al.*, 2015) e presença de zonas híbridas entre drosofilídeos (Lachaise *et al.* 2000; Llopart *et al.* 2005). Especialmente em espécies-pragas, o processo de introgressão podem ser provocados por introduções de espécies de forma proposital ou acidental, aumentando o fitness adaptativo de pragas a um novo habitat e hospedeiros (Allendorf *et al.* 2001; 2010).

Esta tese está estruturada em dois capítulos. O primeiro, intitulado "Population Differentiation with Introgression", tem como objetivo discutir, de maneira acadêmica e teórica, como as espécies evoluem, mesmo diante de processos anteriormente considerados barreiras à evolução. Com os avanços tecnológicos, o aprimoramento das análises bioinformáticas e o desenvolvimento da genômica, foi possível analisar grandes volumes de dados, permitindo uma compreensão mais aprofundada de aspectos evolutivos que antes eram considerados raros na natureza. Nesse contexto, abordamos o caráter dinâmico e muitas vezes controverso do conceito de espécie, especialmente quando se tenta reconstruir a filogenia de certos grupos de espécies e

linhagens. Um exemplo emblemático são as moscas-das-frutas da família Tephritidae, particularmente o gênero *Anastrepha*. Para algumas espécies, como *A. fraterculus*, ainda não há uma definição clara de espécie devido à sua polifagia e à capacidade de adaptação a diferentes ambientes. Esses fatores contribuem para que *A. fraterculus* seja vista como uma ‘unidade biológica’ em contínuo processo evolutivo, com diferentes morfotipos que permanecem como uma incógnita filogenética.

Outro ponto central abordado neste capítulo é o processo de introgressão, analisado extensivamente a partir de dados empíricos, apresentados no segundo capítulo, intitulado "Comparative subsampling of phylogenomic informative loci to delimit species in *Anastrepha*". As análises realizadas demonstram que as incertezas taxonômicas dentro da família Tephritidae não se restringem a esse único processo evolutivo. A introgressão revelou-se um mecanismo fundamental não apenas para influenciar alterações fenotípicas e genotípicas, mas também para diferenciar espécies, mesmo em contexto de fluxo gênico. Isso é particularmente evidente em espécies com divergência recente, uma característica biológica reconhecida no gênero *Anastrepha*, que apresentam maior probabilidade de introgressão em seu genoma.

Com base no arcabouço teórico apresentado no primeiro capítulo, o segundo capítulo prossegue com análises empíricas utilizando um amplo conjunto de marcadores moleculares, permitindo uma investigação detalhada de outros parâmetros evolutivos, como seleção positiva e sinais filogenéticos. Apesar de termos identificado um sinal filogenético significativo capaz de diferenciar espécies, populações e até mesmo gêneros dentro da família Tephritidae, fatores como saturação de dados, classificação de linhagens (Incomplete Lineage Sorting-ILS) e introgressão emergem como processos capazes de mascarar a filogenia. Esses fenômenos são especialmente relevantes em espécies que estão evolutivamente em estágios de adaptação e desenvolvimento no ambiente.

Concluímos que os dois capítulos desta tese oferecem uma análise integrada e complementar: o primeiro discute teoricamente os processos que moldam a evolução das espécies, enquanto o segundo aplica essa base teórica à prática, explorando como esses processos podem separar ou unir filogeneticamente grupos de espécies. Essa dinâmica evolutiva depende tanto da natureza e das interações das espécies quanto da qualidade e abrangência dos dados utilizados, os quais são essenciais para compreender os mecanismos que moldam a evolução em diferentes contextos.

2. OBJETIVOS

2.1 Objetivo geral:

Explorar uma grande gama de mais de 3.000 genes potencialmente informativos para identificar genes e determinar o potencial filogenético afim de descrever adequadamente a filogenia das espécies pertencentes ao gênero *Anastrepha* e efetivamente diferenciar linhagens evolutivas de espécies de moscas-das-frutas da família Tephritidae em um conjunto menor de marcadores.

2.2 Objetivos específicos:

- Examinar os padrões de seleção positiva ao longo dos genes através do uso de abordagens baseadas nos valores de ω (dN/dS);
- Averiguar as pressões de seleção mais atuaram sobre as linhagens das espécies de *Anastrepha*;
- Explorar as árvores de genes e de espécies através de inferências filogenômicas e medir a resolução filogenética entre espécies de *Anastrepha*.
- Investigar através de diferentes parâmetros evolutivos o potencial filogenético do conjunto de genes órtologos
- Realizar através de diferentes metodologias em bioinformática, genes que possam ter potencial filogenético.
- Esclarecer a história evolutiva, para esse conjunto de loci, do gênero *Anastrepha* e dentro do complexo *A. fraterculus*.

Hipótese geral

A diversidade genética de mais de 3.000 genes potencialmente informativos permitirá a identificação de marcadores filogenéticos específicos e a construção de uma árvore filogenética mais robusta e precisa para as espécies do gênero *Anastrepha*, possibilitando a diferenciação eficaz das linhagens evolutivas das moscas-das-frutas da família Tephritidae.

REFERÊNCIAS

- ABBOTT, R.; ALBACH, D.; ANSELL, S.; ARNTZEN, J. W.; BAIRD, S. J. E.; BIERNE, N.; BOUGHMAN J. et al. Hybridization and speciation. **Journal of Evolutionary Biology**, 26: 229– 246, 2013.
- ANDERSON, E. **Hibridação Introgressiva**. Wiley: Nova York. 1949.
- ARCELLA, T.; HOOD, G. R.; POWELL, T. H.; SIM, S. B.; YEE, W. L.; SCHWARZ, D.;.. FEDER, J. L. Hybridization and the spread of the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae), in the northwestern United States. **Evolutionary applications**, 8(8), 834-846, 2015. <https://doi.org/10.1111/eva.12298>
- ALUJA, M. Bionomics and management of *Anastrepha*. **Annual Review of Entomology**, Stanford, v. 39, p. 155-178, Jan. 1994.
- ALBERTI, A.C.; CALCAGNO, G.; SAIDMAN, B.O.; VILARDI, J.C. Analysis of the genetic structure of a natural population of *Anastrepha fraterculus* (Diptera: Tephritidae). **Ann Entomol Soc Am.**, 92:731D736. 1999.
- ABRAFRUTAS. Associação Brasileira de Produtores Exportadores de Frutas e Derivados, 2020. Disponível em: <https://abrafrutas.org/2019/03/07/brasil-e-oterceiro-maior-produtor-de-frutas-do-mundo-diz-abrafrutas/>. Acesso em janeiro de 2021.
- ANUÁRIO BRASILEIRO DE HORTI & FRUTI. Anuário Brasileiro de Horti&Fruti, 2020. Disponível em: http://www.editoragazeta.com.br/sitewp/wp-content/uploads/2020/05/HORTIFRUTI_2020.pdf. Acesso em 12 janeiro de 2021.
- ALLENDORF, F.W.; LEARY, R.F.; SPRUELL, P.; WENBURG, J.K. The problem with hybrids: setting conservation guidelines. **Trends Ecol. Evolut.**, 16(11):613-622, 2001.
- ALLENDORF, F. W.; HOHENLOHE, P.A.; LUIKART, G. Genomics and the future of conservation genetics. **Nature**, 11(10):697–709, 2010.
- AGUIAR-MENEZES, E.L.; FERRARA, F.A.A.; MENEZES, E.B. Moscas-das-frutas. In: Cassino PCR, Rodrigues WC. [Eds.]. Citricultura Fluminense: principais pragas e seus inimigos naturais. Seropedica: EDUR. pp. 67 – 84. 2004.
- ARNOLD ML: **Evolution Through Genetic Exchange**. 2006, Oxford: Oxford University Press
- ARNOLD, M.L.; Martin, N.H.; Adaptation by introgression. **J Biol**, 8, 82, 2009. <https://doi.org/10.1186/jbiol176>
- BACHMANN, G. E. et al. Mate choice confers direct benefits to females of *Anastrepha fraterculus* (Diptera: Tephritidae). **PLoS One**, v. 14, n. 6, p. e0214698. 2019.
- BASSO, A.C.; SONVICO, A.; QUESADA-ALLUE, L. A.; MANSO, F. Karyotypic and molecular identification of laboratory stocks of the South American fruit fly *Anastrepha*

- fraterculus* (Wied.) (Diptera: Tephritidae). **Journal of Economic Entomology**, 96: 1237–1244. 2003. <https://doi.org/10.1093/jee/96.4.1237>.
- BARR, N. B.; CUI, L.; MCPHERON, B. A. Molecular systematics of nuclear period in genus *Anastrepha* (Tephritidae). **Annals of the Entomological Society of America**. 98(2): 173-180. 2005.
- BARR, N. B.; MCPHERON, B. A. Molecular phylogenetics of the genus *Ceratitis* (Diptera: Tephritidae). **Molecular phylogenetics and evolution**, 38(1), 216-230. 2006. <https://doi.org/10.1016/j.ympev.2005.10.013>.
- BARR, N.; RUIZ-ARCE, R.; OBREGÓN, O.; SHATTERS, R.; NORRBOM, A. L.; NOLAZCO, N.; THOMAS, D. Diagnostic characters within ITS2 DNA support molecular identification of *Anastrepha suspensa* (Diptera: Tephritidae). **Florida Entomologist**, 100(1), 182-185. 2017.
- BAACK, E. J.; RIESEBERG, L. H. A genomic view of introgression and hybrid speciation. **Current opinion in genetics & development**, 17(6), 513-518, 2007. <https://doi.org/10.1016/j.gde.2007.09.001>
- BEGON, M.; HARPER, J.L.; TOWNSEND, C.R. **Ecology: individuals, populations and communities**. 3rd edn. Blackwell Science, Oxford. 1996.
- BICKFORD, D.; LOHMAN, D. J.; SODHI, N. S.; NG, P. K.; MEIER, R.; WINKER, K.; ... & DAS, I. Cryptic species as a window on diversity and conservation. **Trends in ecology & evolution**, 22(3), 148-155. 2007.
- BOYKIN, L.M.; SHATTERS, R.G.; HALL, D.G.; BURNS, R.E.; FRANQUI, R. A. Analysis of host preference and geographical distribution of *Anastrepha suspensa* (Diptera: Tephritidae) using phylogenetic analyses of mitochondrial cytochrome oxidase I DNA sequence data. **Bull. Entomol. Res.**, 96, pp. 457-469. 2006. [10.1079/BER2006438](https://doi.org/10.1079/BER2006438).
- CARVALHO, R. DA S.; NASCIMENTO, A. S.; MATRANGOLO, W. J. R. Controle biológico. In: Malavasi, A.; Zucchi, R. A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil: conhecimento básico e aplicado**. Ribeirão Preto: Holos, 2000. p. 113- 117.
- CÁCERES, C.; SEGURA, D.F.; VERA, M.T.; WORNOPORN, V.; CLADERA, J.L.; TEAL, P. SAPOUNTZIS, P.; BOURTZIS, P.; ZACHAROPOULOU, A.; ROBINSON. A.S. Incipient speciation revealed in *Anastrepha fraterculus* (Diptera; Tephritidae) by studies on mating compatibility, sex pheromones, hybridization, and cytology. **Biological Journal of the Linnean Society**, 97: 152–165. 2009. [doi: 10.1111/j.1095-8312.2008.01193.x](https://doi.org/10.1111/j.1095-8312.2008.01193.x).

- CLARKE, A. R.; ARMSTRONG, K. F.; CARMICHAEL, A. E.; MILNE, J. R.; RAGHU, S.; RODERICK, G. K.; YEATES, D. K. Invasive phytophagous pests arising through a recent tropical evolutionary radiation: the *Bactrocera dorsalis* complex of fruit flies. **Annu. Rev. Entomol.**, 50, 293-319. 2005. [10.1146/annurev.ento.50.071803.130428](https://doi.org/10.1146/annurev.ento.50.071803.130428).
- CORSATO, C. D. A. **Moscas-das-frutas (Diptera: Tephritidae) em pomares de goiaba no Norte de Minas Gerais: biodiversidade, parasitóides e controle biológico**. 2004. 95 f. (Doutorado em Ciências) - Universidade de São Paulo, Piracicaba, SP. 2004.
- CONGRAINS, C.; ZUCCHI, R. A.; DE BRITO, R. A. Phylogenomic approach reveals strong signatures of introgression in the rapid diversification of neotropical true fruit flies (Anastrepha: Tephritidae). **Molecular Phylogenetics and Evolution**, 162, 107200. 2021. <https://doi.org/10.1016/j.ympev.2021.107200>.
- CONGRAINS, C.; DUPUIS, J. R.; RODRIGUEZ, E. J.; NORRBOM, A. L.; STECK, G.; SUTTON, B.; ... & GEIB, S. M. Phylogenomic analysis provides diagnostic tools for the identification of *Anastrepha fraterculus* (Diptera: Tephritidae) species complex. **Evolutionary Applications**, 16(9), 1598-1618. 2023. <https://doi.org/10.1111/eva.13589>
- CONDON, M. A.; SCHEFFER, S. J.; LEWIS, M. L.; SWENSEN, S. M. Hidden neotropical diversity: greater than the sum of its parts. **Science**, 320(5878), 928-931. 2008. DOI: [10.1126/science.1155832](https://doi.org/10.1126/science.1155832).
- CHOOCHOTE, W.; MIN, G. S.; INTAPAN, P. M.; TANTRAWATPAN, C.; SAEUNG, A.; LULITANOND, V. Evidence to support natural hybridization between *Anopheles sinensis* and *Anopheles kleini* (Diptera: Culicidae): possibly a significant mechanism for gene introgression in sympatric populations. **Parasites & vectors**, 7(1), 1-13, 2014.
- DAGILIS, A. J.; PEEDE, D.; COUGHLAN, J. M.; JOFRE, G. I.; D'AGOSTINO, E. R.; MAVENGERE, H.; ... MATUTE, D. R. 15 years of introgression studies: quantifying gene flow across Eukaryotes. **BioRxiv**, p. 2021.06. 15.448399, 2021.
- DEGNAN, J. H.; ROSENBERG, N. A. Gene tree discordance, phylogenetic inference and the multispecies coalescent. **Trends in ecology & evolution**, 24(6), 332-340, 2009. <https://doi.org/10.1016/j.tree.2009.01.009>.
- DE OLIVEIRA, ISAIAS et al. *Anastrepha* species (Diptera: Tephritidae): patterns of spatial distribution, abundance, and relationship with weather in three environments of midwestern Brazil. **The Florida Entomologist**, v. 102, n. 1, p. 113-120, 2019.

- DE QUEIROZ, K. Species concepts and species delimitation. **Syst Biol.**, 56: 879–86. pmid:18027281. 2007.
- DIAS, V. S.; SILVA, J. G.; LIMA, KATIA M.; PETITINGA, C. S. C. D.; HERNÁNDEZ-ORTIZ, V.; LAUMANN, R. A.; PARANHOS, B. J.; URAMOTO, K.; ZUCCHI, R. A.; JOACHIM-BRAVO, I. S. An integrative multidisciplinary approach to understanding cryptic divergence in Brazilian species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Biological Journal of the Linnean Society**, v. 117, n. 4, 1, p. 725–746. 2016.
- DIAZ, F.; LUÍS A.; LIMA, A.; NAKAMURA, A. M.; FERNANDES, F.; SOBRINHO JR, I.; DE BRITO, R. A. Evidence for introgression among three species of the *Anastrepha fraterculus* group, a radiating species complex of fruit flies. **Frontiers in Genetics**, 9, 359. 2018. <https://doi.org/10.3389/fgene.2018.003>.
- DOS SANTOS, P.; URAMOTO, K.; MATIOLI, R.S., Experimental Hybridization Among *Anastrepha* Species (Diptera: Tephritidae): Production and Morphological Characterization of F₁ Hybrids, **Annals of the Entomological Society of America**. Volume 94, Issue 5, 1 September 717–725, 2001. [https://doi.org/10.1603/0013-8746\(2001\)094\[0717:EHAASD\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0717:EHAASD]2.0.CO;2)
- EDELMAN, N. B.; FRANDSEN, P. B.; MIYAGI, M.; CLAVIJO, B.; DAVEY, J.; DIKOW, R. B.; ... MALLET, J. Genomic architecture and introgression shape a butterfly radiation. **Science**, 366(6465), 594-599, 2019.
- EGAN, S. P.; FUNK, D. J. Ecologically dependent postmating isolation between sympatric ‘host forms’ of *Neochlamisus bebbianae* leaf beetles. **Proceedings of the National Academy of Sciences**, 106: 19426–19431, 2009.
- FONTAINE, M. C.; PEASE, J. B.; STEELE, A.; WATERHOUSE, R. M.; NEAFSEY, D. E.; et al. Extensive introgression in a malaria vector species complex revealed by phylogenomics. **Science**. 347:1258524. 2015.
- GIARLA, T.C.; J.A. ESSELSTYN. The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. **Syst. Biol.**, 64 (2015), pp. 727-740, 2015. [10.1093/sysbio/syv029](https://doi.org/10.1093/sysbio/syv029)
- GIARDINI, M. C. et al. Geographic distribution of sex chromosome polymorphism in *Anastrepha fraterculus* sp. 1 from Argentina. **BMC Genetics**, v. 21, n. Suppl 2, p. 2020.

- GODAY, C.; SELIVON, D.; PERONDINI, A.L.P.; GRECIANO, P.G.; RUIZ, M.F. Cytological characterization of sex chromosomes and ribosomal DNA location in *Anastrepha* species (Diptera, Tephritidae). **Cytogenetic and Genome Research**, v. 114, p. 70–76, 2006.
- GALLO, D. et.al., **Entomologia Agrícola**. Piracicaba: FEALQ, 2002. 920p
- GALLO-FRANCO, J. J.; TORO-PEREA, N. Variations in the Bacterial Communities in *Anastrepha obliqua* (Diptera: Tephritidae) According to the Insect Life Stage and Host Plant. **Curr Microbiol.**, 77, 1283–1291. 2020. <https://doi.org/10.1007/s00284-020-01939-y>.
- GARCIA, F. R.; NORRBOM, A. L. Tephritoid flies (Diptera, Tephritoidea) and their plant hosts from the state of Santa Catarina in southern Brazil. **Florida Entomologist**, 151-157. 2011.
- GIARLA, T.C.; ESSELSTYN, J.A. The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. **Syst. Biol.**, 64, pp. 727-740. 2015. [10.1093/sysbio/syv029](https://doi.org/10.1093/sysbio/syv029).
- GONÇALVES, R.S. **Bioecologia e competição interespecífica de parasitoides (Hym.: Braconidae) de moscas-das-frutas (Diptera: Tephritidae)**. 2016. 184 f. Tese (Doutorado em Fitossanidade) - Universidade Federal de Pelotas, RS. 2016.
- GOMEZ CENDRA, P.; CALCAGNO, G.; BELLUSCIO, L.; VILARDI, J.C. Comportamento de corte masculino da mosca-das-frutas sul-americana, *Anastrepha fraterculus*, de uma linhagem laboratorial argentina. **Journal of Insect Science**, 11: 1–18. 2011. [doi: 10.1673/031.011.17501](https://doi.org/10.1673/031.011.17501).
- HAN, H.Y.; RO, K.E. Molecular phylogeny of the family Tephritidae (Insecta: Diptera): New insight from combined analysis of the mitochondrial 12S, 16S, and COII genes. **Mol Cells**, 27, 55–66. 2009. <https://doi.org/10.1007/s10059-009-0005-3>.
- HAN, H.-Y.; MCPHERON, B. A. Molecular phylogenetic study of Tephritidae (Insecta: Diptera) using partial sequences of mitochondrial 16S ribosomal DNA. **Molecular Phylogenetics and Evolution**, 7: 17–32. 1997. [doi: 10.1006/mpev.1996.0370](https://doi.org/10.1006/mpev.1996.0370).
- HAN, H.-Y.; K.E. RO. Molecular phylogeny of the family Tephritidae (Insecta: Diptera): new insight from combined analysis of the mitochondrial 12S, 16S, and COII genes. **Molecules and Cells**, 27: 55–66. 2009. [doi: 10.1007/s10059-009-0005-3](https://doi.org/10.1007/s10059-009-0005-3).
- HERNÁNDEZ-ORTIZ, V.; GÓMEZ-ANAYA, J.A.; SÁNCHEZ, A.; MCPHERON, B.A.; ALUJA, M. Morphometric analysis of Mexican and South American populations of

- the *Anastrepha fraterculus* complex (Diptera: Tephritidae), and recognition of a distinct Mexican morphotype. **Bulletin of Entomological Research**, 94: 487–499. 2004. doi: [10.1079/BER2004325](https://doi.org/10.1079/BER2004325).
- HERNÁNDEZ-ORTIZ, V.; BARTOLUCCI, A. F.; MORALES-VALLES, P.; FRÍAS, D.; SELIVON, D. Cryptic Species of the *Anastrepha fraterculus* Complex (Diptera: Tephritidae): A Multivariate Approach for the Recognition of South American Morphotypes. **Annals of the entomological society of américa**, v. 105, no. 2, 2012.
- HERNÁNDEZ-ORTIZ, V.; CANAL, N.A.; TIGRERO SALAS, J.O.; RUÍZ-HURTADO, F.M.; DZUL-CAUICH, J.F. Taxonomy and phenotypic relationships of the *Anastrepha fraterculus* complex in the Mesoamerican and Pacific Neotropical dominions (Diptera, Tephritidae). In: De Meyer, M. Clarke, A.R.; Vera, M.T.; Hendrichs, J. (Eds). Resolution of Cryptic Species Complexes of Tephritid Pests to Enhance SIT Application and Facilitate International Trade. **ZooKeys**, 540: 95–124. 2015. doi: [10.3897/zookeys.540.6027](https://doi.org/10.3897/zookeys.540.6027).
- HIBBINS, MARK S.; HAHN, MATTHEW W. Phylogenomic approaches to detecting and characterizing introgression. **Genetics**, v. 220, n. 2, p. iyab173, 2022.
- HUERTA-SÁNCHEZ, E.; ASAN, J. X.; BIANBA, Z.; PETER, B. M.; VINCKENBOSCH, N.; et al. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. **Nature**, 512:194–197, 2014.
- ISAZA, J. P.; ALZATE, J. F.; CANAL, N. A. Complete mitochondrial genome of the Andean morphotype of *Anastrepha fraterculus* (Wiedemann) (Diptera: Tephritidae). **Mitochondrial DNA Part B**, v. 2, n. 1, p. 210-211, 2017. doi: [10.1080/23802359.2017.1307706](https://doi.org/10.1080/23802359.2017.1307706).
- INCT (Intitutos Nacionais de Ciência e Tecnologia). Frutos tropicais: anual activity report. 2009. Instituto Nacional de Ciência e Tecnologia – **Frutos Tropicais**. http://www.frutostropicais.com.br/AREport_INCTFT_2009.pdf. Acesso em: março de 2022.
- JOHNSON, K. P.; DIETRICH, C. H.; FRIEDRICH, F.; BEUTEL, R. G.; WIPFLER, B.; PETERS, R. S.; ... & YOSHIZAWA, K. Phylogenomics and the evolution of hemipteroid insects. **Proceedings of the National Academy of Sciences**, 115(50), 12775-12780. 2018. <https://doi.org/10.1073/pnas.1815820115>.
- KNIPLING, E.F. Sterile-male method of population control. **Science**. 130:902–904. 1959. doi:[10.1126/science.130.3380.902](https://doi.org/10.1126/science.130.3380.902).

- KNIPLING, E.F. **Sterile insect technique as a screwworm control measure: The concept and its development.** In: Symposium on eradication of the screwworm from the United States and Mexico. O.H. Graham Editor. Misc. Publication. No. 62. Entomological Society of America. 1985.
- LACHAISE, D.; HARRY, M.; SOLIGNAC, M.; LEMEUNIER, F.; BENASSI, V.; CARIOU, M. L. Evolutionary novelties in islands: *Drosophila santomea* a new *melanogaster* sister species from São Tomé. **Proc. R. Soc. Lond. Ser. B**, 267:1487–1495, 2000
- LEVIN, D. A.; FRANCISCO-ORTEGA, J.; JANSEN, R. K. Hybridization and the extinction of rare plant species. **Conservation Biology**, 10: 10– 16, 1996.
- LEWONTIN, R. C.; BIRCH, L. C.. Hybridization as a source of variation for adaptation to new environments. **Evolution**. 315–336. 20. 1966.
- LLOPART, A.; LACHAISE, D.; COYNE, J. A. Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. **Genetics**, v. 171, n. 1, p. 197-210, 2005.
- LI, G.; FIGUEIRÓ, H.V.; EIZIRIK, E.; MURPHY, W.J. Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat Species. **Mol. Biol. Evol.**, 36 pp. 2111-2126. 2019. [10.1093/molbev/msz139](https://doi.org/10.1093/molbev/msz139).
- LOPES, G. N.; ARIAS, O. R.; CÔNSOLI, F. L.; ZUCCHI, R. A. The Identity of Specimens of the *Anastrepha fraterculus* Complex (Diptera, Tephritidae) with Atypical Aculeus Tip. **Neotropical entomology**, 42(6), 618-627. 2013.
- LUDEÑA, B.; BAYAS, R.; PINTAUD, J. C. Phylogenetic relationships of Andean-Ecuadorian populations of *Anastrepha fraterculus* (Wiedemann 1830) (Diptera: Tephritidae) inferred from COI and COII gene sequences. In: Annales de la Société entomologique de France, (Vol. 46, No. 3-4, pp. 344-350). **Taylor & Francis Group**. 2010. <https://doi.org/10.1080/00379271.2010.10697673>.
- MALAVASI, A.; R. A. ZUCCHI (eds.). **Moscas-das-frutas de importância econômica no Brasil: conhecimento básico e aplicado.** Ribeirão Preto, Holos Editora, 327 p. 2000.
- MALLET, J. Hybridization as an invasion of the genome. **Trends in ecology & evolution**, 20(5), 229-237, 2005. <https://doi.org/10.1016/j.tree.2005.02.010>
- MANNI, K.M.; LIMA, C.R.; GUGLIELMINO, S.B.; LANZAVECCHIA, M.; JURI, T.; VERA, J.; MALACRIDA, A. T. *et al.* Relevant genetic differentiation among Brazilian populations

- of *Anastrepha fraterculus* (Diptera, Tephritidae). **ZooKeys**, 540, pp. 157-173. 2015. [10.3897/zookeys.540.6713](https://doi.org/10.3897/zookeys.540.6713).
- MARTIN, S. H.; JIGGINS, C. D. Interpreting the genomic landscape of introgression. **Current Opinion in Genetics and Development**. 47:69-74, 2017.
- MALAVASI, A.; MORGANTE, J. S.; PROKOPY, R. J. Distribution and activities of *Anastrepha fraterculus* (Diptera: Tephritidae) flies on host and nonhost trees. **Ann. Entomol. Soc. Am**, 286–292. 76. 1983.
- MENGUAL, X.; KERR, P.; NORRBOM, A. L.; BARR, N. B.; LEWIS, M. L.; STAPELFELDT, A. M.; ... & GAIMARI, S. Phylogenetic relationships of the tribe Toxotrypanini (Diptera: Tephritidae) based on molecular characters. **Molecular Phylogenetics and Evolution**, 113, 84-112. 2017. <https://doi.org/10.1016/j.ympev.2017.05.011>.
- MATTHEWS, R.W.; MATTHEWS, J.R. **Insect behavior**. New York, NY: John Wiley & Sons. 1978.
- MORGANTE, J. S.; SELIVON, D.; SOLFERINI, V. N.; MATIOLI, S. R. Evolutionary patterns in specialist and generalist species of *Anastrepha*, pp. 15–20. In: M. Aluja and P. Liedo [eds.], **Fruit flies: biology and management**. Springer, New York, 1993.
- MORAIS, M. C.; RAKES, M.; PADILHA, A. C.; GRÜTZMACHER, A. D.; NAVA, D. E.; BERNARDI, O.; & BERNARDI, D. Susceptibility of Brazilian Populations of *Anastrepha fraterculus*, *Ceratitis capitata* (Diptera: Tephritidae), and *Drosophila suzukii* (Diptera: Drosophilidae) to Selected Insecticides. **Journal of Economic Entomology**, v. 114, Issue 3, p. 1291–1297. 2021. <https://doi.org/10.1093/jee/toab050>.
- MCALPINE, J.F. Phylogeny and classification of the Muscomorpha. In: Mcalpine, J.F. (Ed.) **Manual of nearctic diptera**. Ottawa: Biosystema Research Centre, p.1397-1518. 1989.
- MCPHERON, B.A.; HAN, H.Y.; SILVA, J.G.; NORRBOM, A.L. Phylogeny of genus *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based upon 16S rDNA mitochondrial sequences. In: Aluja M, Norrbom AL. (Eds) **Fruit flies (Tephritidae): phylogeny and evolution of behavior**. CRC Press, Boca Raton, Florida, 343–361. 1999.
- MONTOYA, P. et al. **MOSCAS DE LA FRUTA: FUNDAMENTOS Y PROCEDIMIENTOS PARA SU MANEJO**. S y G editores, Distrito Federal, México, 2010. 395p.

- MORGANTE, J.S.; MALAVASI, A.; BUSH, G.L. Biochemical systematics and evolutionary relationships of neotropical *Anastrepha*. **Annals of the Entomological Society of America**. 73: 622–630. 1980. doi: [10.1093/aesa/73.6.622](https://doi.org/10.1093/aesa/73.6.622).
- NAVA, D.E.; BOTTON, M. **Bioecologia e controle de *Anastrepha fraterculus* e *Ceratitis capitata* em pessegueiro**. Pelotas: Embrapa, 2010. 29p
- NICÁCIO, J.; OLIVEIRA, I. D.; UCHOA, M. A.; FACCENDA, O.; ABOT, A. R.; FERNANDES, M. G.; GARCIA, F. R. Spatial distribution and control levels of *Anastrepha* spp. (Diptera: Tephritidae) in guava orchards. **Anais da Academia Brasileira de Ciências**, 91. 2019. <https://doi.org/10.1590/0001-3765201920180428>.
- NORRBOM, A.L.; CARROLL, L.E.; THOMPSON, F.C.; WHITE, I.M.; FREIDBERG, A. Systematic database of names. In: THOMPSON, F.C. (Ed.). **Fruit Fly Expert Identification System and Systematic Information Database**. Myia, 9, 65–251 and *Diptera Data Dissemination Disk*, 1 [CD-ROM]. 1999.
- NORRBOM, A. L.; ZUCCHI, R. A.; ORTIZ, H. V. Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based on morphology, pp. 299–342. In: Aluja, M.; Norrbom, A. L. (eds.). **Fruit flies (Tephritidae): Phylogeny and evolution of behavior**. CRC Press, Boca Raton. 1999b.
- NORRBOM, A.L.; KORYTKOWSKI, C.A. New species of *Anastrepha* (Diptera: Tephritidae), with a key for the species of the *megacantha* clade. **Zootaxa**, 3478 (1), 510–552. 2012 <https://doi.org/10.11646/zootaxa.3478.1.43>.
- NORRBOM, A. L.; NEDER, L. E. New neotropical species of Trupanea Diptera: Tephritidae with unusual wing patterns. **Zootaxa**, 3821(4), 443–456. 2014. DOI: [10.11646/ZOOTAXA.3821.4.3](https://doi.org/10.11646/ZOOTAXA.3821.4.3)
- NORRBOM, A.L.; BARR, N.; KERR, P.H.; MENGUAL, X.; NOLAZCO ALVARADO, N.; RODRIGUEZ, E.J.; STECK, G.J.; SUTTON, B.D.; URAMOTO, K.; ZUCCHI, R.A. Synonymy of *Toxotrypana* Gerstaecker with *Anastrepha* Schiner (Diptera: Tephritidae). **Proceedings of the Entomological Society of Washington**, 120, 834–841. 2018 <https://doi.org/10.4289/0013-8797.120.4.834>.
- NORRBOM, A.L. **Host plant database for *Anastrepha* and *Toxotrypana* (Diptera: Tephritidae: Toxotrypanini)**. Diptera Data Dissemination Disk, 2. 2004. Available online: <https://www.ars.usda.gov/research/publications/publication/?seqNo115=108854>. (accessed on 17 March 2023).

- NORRBOM, A.L. **Updates to Biosystematic Database of World Diptera for Tephritidae through 1999**. Diptera Data Dissemination Disk, 2 [CD-ROM]. 2004.
- NORRBOM, A. L.; KORYTKOWSKI, C.A. A revision of the *Anastrepha robusta* species group (Diptera: Tephritidae). **Zootaxa**, 2182, 1–91. 2009.
- NORRBOM, A. L.; KORYTKOWSKI, C.A. New species of and taxonomic notes on *Anastrepha* (Diptera: Tephritidae). **Zootaxa**, 2740, 1–23. 2011.
- NORRBOM, A.L., KORYTKOWSKI, C.A., ZUCCHI, R.A., URAMOTO, K., VENABLE, G.L., MCCORMICK, J., DALLWITZ, M.J., **Anastrepha and Toxotrypana: descriptions, illustrations, and interactive keys**. Dallwitz M, Paine T, Zurcher E. Intkey for Windows. Version, 5. 2012.
- NYGREN, A. Cryptic Polychaete Diversity: A Review. **Zool. SCR.**, 43 pp. 172-183. 2013.
- OLIVEIRA, G. C. **Variações no gene yolk em moscas-das-frutas do grupo fraterculus**. 2009. 90 p. Dissertação (Mestrado em Ciências Biológicas) – Universidade Federal de São Carlos, 2009.
- PASSOS, J. F.; NASCIMENTO, D. B.; MENEZES, R. S.; ADAIME, R.; ARAUJO, E. L.; LIMA, K. M.; ... SILVA, J. G. Genetic structure and diversity in Brazilian populations of *Anastrepha obliqua* (Diptera: Tephritidae). **PLoS One**, 13(12), e0208997. 2018. <https://doi.org/10.1371/journal.pone.0208997>.
- PANTE, E.; PUILLANDRE, N.; VIRICEL, A.; ARNAUD-HAOND, S.; AURELLE, D.; CASTELIN, M.; ... SAMADI, S. Species are hypotheses: avoid connectivity assessments based on pillars of sand. **Molecular Ecology**, 24(3), 525-544. 2015.
- PARANHOS, B. A. J.; WALDER, J. M. M.; PAPADOPOULOS, N. T. A simple method to study parasitism and field biology of the parasitoid *Diachasmimorpha longicaudata* (Hymenoptera: Braconidae) on *Ceratitis capitata* (Diptera: Tephritidae). **Biocontrol Science and Technology**, 13: 631-639. 2003.
- PARANHOS, B.J. et al. Biological control of fruit flies in Brazil. **Pesq. Agropec. Bras.**, 54:e26037, 2019.
- PARK, Y.L.; TOLLEFSON, J.J. Spatial distributions of corn rootworm (Coleoptera: Chrysomelidae) eggs and larvae: implications for sampling. **J KANSAS ENTOMOL SOC.** 79: 129-135. 2006.

- PERRE, P.; FARIA, F. A.; JORGE, L. R.; ROCHA, A.; TORRES, R. D. S.; SOUZA-FILHO, M. F.; ... ZUCCHI, R. A. Toward an automated identification of *Anastrepha* fruit flies in the *fraterculus* group (Diptera, Tephritidae). **Neotropical entomology**, 45(5), 554-558. 2016.
- PEREIRA-RÊGO, D. R. G. et al. Morfometria de *Anastrepha fraterculus* (Wied) (Diptera: Tephritidae) relacionada a hospedeiros nativos, Myrtaceae. **Arquivos do Instituto Biológico**, v. 78, p. 37-43, 2020.
- PERRE, P.; JORGE, L.R.; LEWINSOHN, T.M.; ZUCCHI, R.A. Morphometric differentiation of fruit fly pest species of the *Anastrepha fraterculus* group (Diptera: Tephritidae). **Ann. Entomol. Soc. Am.**, 107 pp. 490-495. 2014. [10.1603/AN13122](https://doi.org/10.1603/AN13122).
- PREZOTTO, L.; PERONDINI, A.; HERNÁNDEZ-ORTIZ, V.; FRÍAS, D.; SELIVON, D. What Can Integrated Analysis of Morphological and Genetic Data Still Reveal about the *Anastrepha fraterculus* (Diptera: Tephritidae) Cryptic Species Complex? **Insects**, 10, 408. 2019.
- PONCIO, S. **Bioecologia e técnicas de criação de parasitoides (Hymenoptera) nativos de três espécies de *Anastrepha* no Brasil e no México**. 2015. 133p. Tese (Doutorado em Fitossanidade) - Universidade Federal de Pelotas, RS. 2015.
- RAGA, A.; PRESTES, D.; SOUZA FILHO, M.; SATO, M.; SILOTO, R.; GUIMARÃES, J.; ZUCCHI, R. Fruit fly (Diptera: Tephritoidea) infestation in citrus in the State of São Paulo, **Brazil. Neotrop. Entomol.** 33, 85–89. 2004.
- RAMÍREZ-SANTOS, E.; RENDON, P.; GOUVI, G.; ZACHAROPOULOU, A.; BOURTZIS, K.; CÁCERES, C.; BLOEM, K. A Novel Genetic Sexing Strain of *Anastrepha ludens* for Cost-Effective Sterile Insect Technique Applications: Improved Genetic Stability and Rearing Efficiency. **Insects**, v. 12, n. 6, 2021.
- RIESEBERG, L. H.; CARNEY, S. E. Revisão de Tansley no. 102 Hibridação de plantas. **Novo Phytol**, 140 : 599-624, 1998.
- ROCHA, L.S.; SELIVON, D. **Studies on highly repetitive DNA in cryptic species of the *Anastrepha fraterculus* complex**. In: Barnes, B.N. (Ed.) 6th International Symposium on Fruit Flies of Economic Importance, Stellenbosch, South Africa, 415–418. 2004.
- RUIZ, M. F.; EIRÍN-LOPEZ, J. M.; STEFANI, R. N.; PERONDINI, A. L. P.; SELIVON, D.; SÁNCHEZ, L. The gene doublesex of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. **Development Genes and Evolution**. 217: 725-731. 2007a.

- RUIZ, M. F.; MILANO, A.; SALVEMINI, M.; EIRÍN-LOPEZ, J. M.; PERONDINI, A. L. P.; SELIVON, D.; POLITO, C.; SACCONI, G.; SÁNCHEZ, L. The gene transformer of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. **PloS One**, 11: e1239. 2007b.
- RHYMER, J. M; SIMBERLOFF, D. Extinction by hybridization and introgression. **Annual Review of Ecology and Systematics**, 27: 83-109, 1996.
- SELIVON, D.; PERONDINI, A. L. P.; Morgante, J.S. A genetic-morphological characterization of two cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Annals of the Entomological Society of America**, v. 98, n. 3, p. 367–381, doi: [10.1603/0013-8746\(2005\)098\[0367:2005a](https://doi.org/10.1603/0013-8746(2005)098[0367:2005a).
- SELIVON, D.; PERONDINI, A. L. P.; ROCHA, L. S. Karyotype characterization of *Anastrepha* fruit flies (Diptera: Tephritidae). **Neotropical Entomology**. 34: 273–279, 2005b. doi: [10.1590/S1519-566X2005000200015](https://doi.org/10.1590/S1519-566X2005000200015).
- SELIVON, D.; PERONDINI, A. L. P. Especies crípticas del complejo *Anastrepha fraterculus* en Brasil. In: Hernández-Ortiz, V. (ed.). **Moscas de la fruta en Latinoamérica (Diptera: Tephritidae): diversidad, biología y manejo**. S y G editores, Distrito Federal, México. 101–118. 2007.
- SEGURA, M. D.; CALLEJAS, C.; FERNÁNDEZ, M. P.; OCHANDO, M. D. New contributions towards the understanding of the phylogenetic relationships among economically important fruit flies (Diptera: Tephritidae). **Bulletin of Entomological Research**. 96: 279-288. 2006.
- SEGURA, D.; PETIT-MARTY, N.; SCIURANO, R.; VERA, M.T.; CALCAGNO, G.; ALLINGHI, A.; GÓMEZ-CENDRA, P.; CLADERA, J.; VILARDI, J. Comportamento Lekking de *Anastrepha fraterculus* (Diptera: Tephritidae). **Entomologista da Flórida**. 90: 154-162. doi: 10.1653 / 0015-4040 (2007) 90 [154: LBOAFD] 2.0.CO; 2. 2007.
- SUAREZ-GONZALEZ, A.; LEXER, C.; CRONK, Q. C. Adaptive introgression: a plant perspective. **Biology letters**, 14(3), 20170688, 2018. <https://doi.org/10.1098/rsbl.2017.0688>
- SILVA, J. G.; BARR, N. B. Recent advances in molecular systematics of *Anastrepha* Schiner. fruit flies of economic importance: from basic to applied knowledge. In: International Symposium on fruit flies of economic importance, 7. Salvador. **Proceedings...** Salvador, 2006. p. 13-28. 2006.

- SILVA, J.G.; BARR, N.B. **Recent advances in molecular systematics of *Anastrepha* Schiffner**. In: Proceedings of 7th International Symposium on Fruit Flies of Economic Importance. 13-18 pp. 2008.
- SOUZA FILHO, M. F.; COSTA, V. A.; PAZINI, W. C. Manejo integrado de pragas na cultura da manga, p. 339-376. In: Rozane, D. E.; Darezzo, R.J.; Aguiar, R.L.; Aguilera, G.H.A.; Zambolim, L. (eds.), **Manga: Produção integrada, industrialização e comercialização**. Viçosa, UFV, 604p. 2004.
- SUTTON, B.D.; STECK, G.J.; NORRBOM, A. L.; RODRIGUEZ, E. J.; SRIVASTAVA, P.; NOLAZCO ALVARADO, N.; COLQUE, F.; ... PONCE, P. Nuclear ribosomal internal transcribed spacer 1 (ITS1) variation in the *Anastrepha fraterculus* cryptic species complex (Diptera, Tephritidae) of the Andean region. **ZooKeys**, 540:175–91. 2015.
- SMITH-CALDAS, M. R.; MCPHERON, B. A.; SILVA, J. G.; ZUCCHI, R. A. Phylogenetic relationships among species of the fraterculus group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. **Neotropical Entomology**, 30, 565-573. 2001.
- SCALLY, M.; INTO, F.; THOMAS, D. B.; RUIZ-ARCE, R.; BARR, N. B.; SCHUENZEL, E. L. Resolution of inter and intra-species relationships of the West Indian fruit fly *Anastrepha obliqua*. **Molecular Phylogenetics and Evolution**, 101, 286-293. 2016. <https://doi.org/10.1016/j.ympev.2016.04.020>.
- STRUCK, T. H.; FEDER, J. L.; BENDIKSBY, M.; BIRKELAND, S.; CERCA, J.; GUSAROV, V. I.; ... DIMITROV, D. Finding evolutionary processes hidden in cryptic species. **Trends in Ecology & Evolution**, 33(3), 153-163. <https://doi.org/10.1016/j.tree.2017.11.007>. 2018.
- STECK, G.J.; SHEPPARD, W.S. Mitochondrial DNA Variation in *Anastrepha fraterculus*. In: Aluja, M.; Liedo, P. (ed.). **Fruit flies: biology and management**. New York: Springer-Verlag, p. 9-14, 1993.
- SCHUTZE, M.K.; AKETARAWONG, N.; AMORNSAK, W.; ARMSTRONG, K. F.; AUGUSTINOS, A.A.; BARR, N.; et al. Synonymization of key pest species within the *Bactrocera dorsalis* species complex (Diptera: Tephritidae): taxonomic changes based on a review of 20 years of integrative morphological, molecular, cytogenetic, behavioural and chemoecological data. **Systematic Entomology**, 40(2), 456-471. 2014.

- SCHWARZ, D.; MATTA, B. M.; SHAKIR-BOTTERI, N. L.; MCPHERON, B. A. Host shift to an invasive plant triggers rapid animal speciation. **Nature**, 436: 546– 549, 2005.
- UCHÔA, M. **A Fruit flies (Diptera: Tephritoidea): biology, host plants, natural enemies, and the implications to their natural control**. Integrated Pest Management and Pest Control: Current and Future Tactics. InTech, Rijeka, Croatia, 271-300. 2012.
- TROYA, H.; NORRBOM, A.; PINEDA, J. Two new species of *Anastrepha* (Diptera: Tephritidae) from Ecuador. **Zootaxa**, 4820:366–372. 2020. <https://doi.org/10.11646/zootaxa.4820.2.9>.
- TWYFORD, A.; ENNOS, R. Hibridação de próxima geração e introgressão. **Hereditariedade**, 108, 179-189, 2012. <https://doi.org/10.1038/hdy.2011.68>.
- VELASCO-CUERVO, S. M.; AGUIRRE-RAMIREZ, E.; GALLO-FRANCO, J. J.; OBANDO, R. G.; CARREJO, N.; TORO-PEREA, N. Saving DNA from museum specimens: The success of DNA mini-barcodes in haplotype reconstruction in the genus *Anastrepha* (Diptera: Tephritidae). **Journal of advanced research**, 16, 123-134. 2019. <https://doi.org/10.1016/j.jare.2018.11.001>.
- VIRGILIO, M.; DE MEYER, M.; WHITE, I. M.; BACKELJAU, T. African *Dacus* (Diptera: Tephritidae: Molecular data and host plant associations do not corroborate morphology based classifications. **Molecular Phylogenetics and Evolution**, 51(3), 531-539. 2009. <https://doi.org/10.1016/j.ympev.2009.01.003>.
- VIRGILIO, M.; DELATTE, H.; QUILICI, S.; BACKELJAU, T.; DE MEYER, M. Cryptic diversity and gene flow among three African agricultural pests: *Ceratitis rosa*, *Ceratitis fasciventris* and *Ceratitis anonae* (Diptera, Tephritidae). **Molecular Ecology**, 22(9), 2526-2539. 2013. <https://doi.org/10.1111/mec.12278>.
- WILLIS, S. C.; MACRANDER, J.; FARIAS, I. P.; ORTÍ, G. Simultaneous delimitation of species and quantification of interspecific hybridization in Amazonian peacock cichlids (genus *Cichla*) using multi-locus data. **BMC Evol. Biol.**, 12:96, 2012.
- WHITE, I.M.; ELSON-HARRIS, M.M. **Fruit Flies of Economic Significance: Their Identification and Bionomics**. CAB International. 1992. 70p.
- ZUCCHI, R. A.; CANAL, D. N. A. **Braconídeos parasitóides de moscas-das-frutas na América do Sul**. In: Simpósio de controle biológico, 5., Foz de Iguaçu, Anais... Londrina: SEB, p. 89-92. 1996.

ZUCCHI, R. A. Taxonomia. In: Malavasi, A.; Zucchi, R.A. (Ed.). Moscas-das-frutas de importância econômica no Brasil: conhecimento Básico e aplicado. Ribeirão Preto: Holos, 2000. p.12-24.

1ª Capítulo

Population Differentiation with Introgression

Capítulo do livro “Conservation Genetics in the Neotropics” publicado na Springer International Publishing.

1º Capítulo: Diferenciação populacional com introgressão

Resumo

Espécies são usadas como indicadores de vulnerabilidade em nível populacional, em diferentes estudos, principalmente porque, com base em fundamentos teóricos, desempenham papel importante como parâmetros ecológicos e evolutivos e têm consequências práticas em vários campos, principalmente em estudos sobre biodiversidade. No entanto, há controvérsias sobre o que realmente forma uma determinada espécie, uma vez que essa ideia é frequentemente baseada em diferentes atributos biológicos e forças evolutivas. Assumindo que é assim, por causa da combinação de relevância e discordância, vários novos métodos têm proposto recentemente para ajudar a resolver esse dilema. Novos métodos e dados, embora promissores, ainda não atenuaram essas alegações. Na verdade, eles evidenciaram a alta complexidade do que forma uma determinada espécie na natureza, ao invés do que se pensava anteriormente. Essa descoberta trouxe uma nova perspectiva para os níveis de introgressão que ocorrem entre diferentes espécies, um fato que levou à conclusão de que esses processos parecem ser ainda mais significativos nos Neotrópicos. Assim, sob essa perspectiva, a seleção natural não é a única força que molda espécies e populações em seus ambientes, uma vez que está bem documentado que elas podem se diferenciar umas das outras, mesmo quando se trata da incidência de forças seletivas antagonicas, como fluxo gênico, deriva genética, recombinação e processos de introgressão, bem como que podem ser acessadas a partir da perspectiva conservacionista. No entanto, além de influenciar a filogenia de um número significativo de espécies, a introgressão pode favorecer variações fenotípicas e genotípicas que, por sua vez, podem ser responsáveis por adaptações, principalmente em populações forçadas a se ajustar a mudanças climáticas contínuas e perda de habitat. Com base em um arcabouço conceitual e teórico, o presente estudo comparou a mística da introgressão como força seletiva pungentemente negativa à maneira como ela pode gerenciar processos históricos e potencialmente habilitar variações genéticas importantes capazes de afetar processos ecológicos e evolutivos em diferentes espécies e em populações pequenas e grandes.

Palavras-chave: Evolução, Conservação, Seleção natural, Fluxo gênico, Neotropicos.

Abstract

Species are used as vulnerability indicators at population level, in different studies, mainly because, based on theoretical foundations, they play important part as ecological and evolutionary parameters and have practical consequences in several fields, mostly in studies about biodiversity. Yet, there are controversies about what really forms a given species, since this idea is often based on different biological attributes and evolutionary forces. Assumingly, it is so, because of the combination of relevance and disagreement several new methods have been recently proposing to help solving this dilemma. New methods and data, although promising, have not yet mitigated these claims. Actually, they have evidenced the high complexity of what forms a given species in nature, rather than it was previously thought. This finding has brought a new perspective to the introgression levels taking place among different species, a fact that led to the conclusion that these processes appear to be even more significant in the Neotropics. Thus, from this perspective, natural selection is not the only force shaping both species and populations in their environments, since it is well documented that they can differentiate themselves from each other, even when it comes to the incidence of antagonistic selective forces, such as gene flow, genetic drift, recombination and introgression processes, as well as that they can be accessed from the conservationist perspective. Nevertheless, besides influencing the phylogeny of a significant number of species, introgression can favor phenotypic and genotypic variations that, in their turn, can account for adaptations, mainly in populations forced to adjust to continuous climate changes and habitat loss. Based on a conceptual and theoretical framework, the current study compared the mystique of introgression as poignantly negative selective force to the way it can manage historical processes and potentially enable important genetic variations capable of affecting ecological and evolutionary processes in different species, and in small and large populations.

Keywords: Evolution, Conservation, Natural selection, Gene flow, Neotropics

1. Introdução

O livro mais famoso de Charles Darwin, "Sobre a Origem das Espécies", na verdade tem o título completo "Origem das Espécies por Meio da Seleção Natural, ou a Preservação das Raças Favorecidas na Luta pela Vida", que reflete como Darwin via o processo de formação das espécies, tendo a seleção natural como o principal impulsionador, mas não a única força envolvida com sua diferenciação. De acordo com Darwin, à medida que a seleção natural e a adaptação prosseguiam, elas lentamente afastariam as populações. Darwin então via a diferenciação como um continuum no qual a separação efetiva de diferentes "espécies" seria estabelecida mais ou menos arbitrariamente, de modo que as espécies não teriam realidade na natureza, ao contrário da visão fixista mantida na época. Apesar dessa continuidade, a separação entre populações promovida por sua adaptação a diferentes nichos, e a extinção de ancestrais e intermediários, seriam responsáveis pelas lacunas que encontramos entre as espécies (Darwin, 1859).

Mesmo na época de Darwin, a sugestão de que espécies seriam construções relativamente arbitrárias foi recebida com ceticismo por pesquisadores que acreditavam que qualquer historiador natural era capaz de ver a realidade das espécies ao investigá-las na natureza (Sharp, 1871). Com o passar do tempo, a visão de que espécies na natureza seriam reais foi filosoficamente reforçada por pesquisadores que consideravam as espécies não mais como tendo uma essência, como originalmente proposto por Aristóteles, mas inicialmente como uma classe e depois como um indivíduo (Hull, 1978), e a questão mudou para a compreensão das forças envolvidas em sua formação.

Embora a grande maioria dos pesquisadores hoje reconheça as espécies como unidades evolutivas reais, a compreensão do que determina essa realidade e quais forças evolutivas estariam envolvidas em sua manutenção continua sendo os principais pontos de discórdia. Alguns pesquisadores, como Mayr e Dobzhansky, sugeriram que a realidade das espécies seria mantida porque seus membros compartilhariam uma estrutura genética/ecológica comum, estabelecendo o que Mayr apelidou de "campos de recombinação" (Mayr, 1963). Essa interpretação expandiu a visão da síntese evolutiva que considerava principalmente os genes individualmente na especiação (no que Mayr depreciativamente se referiu como "genética beanbag") para reforçar a expressão da variação no contexto ecológico e fenotípico de um indivíduo. Essa ênfase no potencial de variação genética ser compartilhada e apropriadamente expressa em diferentes indivíduos em um contexto

ecológico fez Mayr enfatizar o isolamento reprodutivo como um aspecto essencial para o reconhecimento de espécies distintas, levando ao “Conceito Biológico de Espécie” (CBE).

Os fundamentos do CBE intuitivamente faziam sentido para vários pesquisadores, tanto que foi amplamente aceito e disseminado para se tornar, sem dúvida, a definição de espécie mais comum, mesmo fora da comunidade científica. No entanto, mesmo naquela época, os botânicos estavam encontrando vários casos em que híbridos eram formados entre diferentes "espécies" que ainda mantinham suas identidades e não conseguiam homogeneizar. Esses híbridos até levaram alguns a enfatizar o importante papel que a recombinação dessa variação interespecífica pode ter no processo adaptativo (Anderson e Stebbins Jr., 1954). O abismo entre botânicos e zoólogos culminou com Paul Ehrlich e Peter Raven (1969) propondo que a manutenção da identidade das espécies seria uma consequência de adaptações a nichos específicos, mais do que ao isolamento reprodutivo. As espécies permaneceriam isoladas porque a seleção natural poderia superar até mesmo níveis significativos de fluxo gênico e manter sua identidade. De acordo com esse "conceito ecológico de espécie", as espécies eram grupos de indivíduos que compartilhavam adaptações que lhes permitiriam ocupar um certo nicho ecológico, mesmo diante de algum fluxo gênico. Os limites das espécies seriam então estabelecidos pelo equilíbrio entre o fluxo genético e a seleção natural.

Apesar da prevalência do conceito de espécie biológica de Mayr, principalmente na década de 1970, outros pesquisadores importantes, mesmo no cerne da síntese evolutiva, já haviam sugerido que a existência de linhagens evolutivas seria instrumental para a caracterização de espécies (Simpson, 1951). A princípio, foi sugerido que espécies eram linhagens evolutivas que mantêm sua identidade de outras linhagens e retêm suas próprias tendências evolutivas e destino histórico. Tal visão não explicitava quais atributos permitiriam isso, muito menos qual seria seu “destino”, levando Joel Cracraft (1983) a simplificar a definição de espécies como sendo grupos de organismos que compartilham sinapomorfias, ou caracteres derivados, em seu conceito filogenético de espécies.

Desde sua primeira definição moderna, vários outros conceitos de espécie foram propostos, enfatizando diferentes aspectos biológicos, ou suas combinações. Apesar dessa complexidade, De Queiroz (2007), propôs que diferenças entre conceitos de espécie apenas levam de fato a diagnósticos diferentes e a uma incongruência em uma área cinzenta que abrange espécies que divergiram recentemente e que abrigam certos atributos. Segundo ele, nessa zona cinzenta, os processos que são considerados determinantes para cada um dos diferentes conceitos de espécie

podem acontecer em velocidades diferentes. Com o passar do tempo, porém, esses processos tendem a convergir se não houver forças evolutivas intensas atuando sobre as populações. Um aspecto relevante dessa visão é que se considerarmos a especiação de um ponto de vista macroevolutivo, de fato, não pareceria haver nenhuma inconsistência, porque depois de muito tempo decorrido, a maioria, se não todos, os conceitos de espécie convergiriam para as mesmas conclusões. No entanto, uma visão mais populacional do processo, facilitada hoje pelo aumento do número de marcadores genéticos que são cada vez mais capazes de diferenciar até mesmo grupos de indivíduos em uma população, pode revelar padrões na história demográfica de populações, como contrações e expansões do tamanho populacional. Devido a esses avanços, há maior capacidade de reconhecer a estrutura genética intraespecífica e diferentes níveis de introgressão ao longo de sua distribuição e história, o que aponta para a heterogeneidade filogenética entre indivíduos e entre seus genomas que teriam um impacto potencial em seu status de espécie, dependendo do conceito de espécie usado (Edwards *et al.* 2016; Smith *et al.* 2020; Hibbins e Hahn, 2021).

2. Objetivos

2.1 Objetivo Geral: Analisar o papel da introgressão como força seletiva em espécies, com foco em suas implicações ecológicas e evolutivas, especialmente no contexto dos Neotrópicos, e discutir os novos métodos e dados para entender a formação de espécies e os processos de adaptação.

2.2 Objetivos específicos:

- Avaliar os impactos da introgressão nos processos evolutivos e na filogenia das espécies, com ênfase em como ela pode favorecer variações genéticas importantes para a adaptação das populações.
- Investigar as controvérsias sobre a definição de espécies, explorando as diferentes forças biológicas e evolutivas que influenciam a formação de espécies, com foco na interação entre fluxo gênico, deriva genética e seleção natural.
- Examinar os novos métodos e dados que propõem soluções para os dilemas na definição de espécies, analisando como eles podem contribuir para uma compreensão mais precisa dos processos de introgressão.

3. Bases Genômicas da Diferenciação Populacional e Especiação

O que ficou claro, à medida que mais espécies e mais partes de seus genomas foram estudadas, é que o processo de especiação é muito mais complexo e heterogêneo do que se considerava anteriormente. Isso talvez justifique a infinidade de diferentes conceitos de espécies propostos, tanto que há alguns pesquisadores hoje que sugerem que devemos restabelecer a visão de Darwin que questiona a realidade das espécies na natureza (Mallet 2001; Mallet *et al.* 2007; Abbott *et al.* 2013). Essas descobertas estão mudando o paradigma de que o isolamento reprodutivo é inexorável para que a divergência ocorra, uma vez que estamos encontrando diferentes graus de isolamento reprodutivo, tanto ao longo da distribuição geográfica da espécie quanto de sua história (Kronforst *et al.* 2013; Kronforst e Papa, 2015; Zhang *et al.* 2016; Barrera-Guzmán *et al.* 2017). Essas descobertas enfraquecem a amplitude do conceito biológico de espécie (Mallet, 2001) ao sugerir que pode não haver grandes discontinuidades entre espécies, ao contrário do que Mayr e outros pesquisadores propuseram, pelo menos em seu início. Estudos usando dados genômicos que encontraram evidências de fluxo gênico durante a especiação também sugerem que alguns genes que evoluem sob seleção natural ou sexual, desempenham papéis importantes na diversificação de linhagem (Lamichhaney *et al.* 2015; Meier *et al.* 2017b). Assim, quando duas populações divergem na presença de fluxo gênico ou quando experimentam fluxo gênico secundário após sua diferenciação, é o equilíbrio entre fluxo gênico, deriva e seleção natural que determinará o resultado do processo adaptativo, particularmente no caso de seleção divergente, mas também com seleção de equilíbrio. A presença de fluxo gênico pode facilitar o processo adaptativo, mas também aumentar a probabilidade de que as populações possam se extinguir.

O surgimento de novas metodologias genômicas que aumentaram a compreensão do processo de diferenciação de espécies e demonstraram que o cruzamento é muito maior do que o previsto anteriormente, mesmo em animais (Mallet, 2007; Solís-Lemus *et al.* 2016, Jackson *et al.* 2017a; Schield *et al.* 2019; Congrains *et al.* 2021; Hirase e Yamasaki, 2021; Banker *et al.* 2022; Suvorov *et al.* 2022), levou a uma reavaliação do papel das forças evolutivas subjacentes no processo de formação de espécies. Com isso, o isolamento reprodutivo não parece ser uma propriedade que afeta necessariamente todo o genoma de uma só vez, podendo ocorrer em diferentes níveis, em diferentes partes do genoma ou em diferentes pontos da distribuição das espécies. Fica claro então que o isolamento reprodutivo não parece ser um atributo "tudo ou nada" no qual as espécies seriam completamente isoladas ou completamente compatíveis. Além disso, a sugestão de que o fluxo

gênico pode até envolver genes que são seletivamente importantes, como, por exemplo, genes que conferem adaptações a diferentes hospedeiros (Powell *et al.* 2014) ou a certos aspectos ambientais, levanta questões importantes sobre a formação de espécies que impactam alguns conceitos de espécie, como a própria natureza do isolamento reprodutivo, o papel da adaptação na especiação ou a própria existência de linhagens evolutivas bem definidas. Diferentes partes do genoma podem ter histórias diferentes, responder diferentemente a pressões de seleção semelhantes e ser diferencialmente permeáveis ao fluxo gênico (Zhang *et al.* 2016). Isso não só produziria diferentes relações filogenéticas inferidas de diferentes partes do genoma, mas também quando inferidas de indivíduos coletados de diferentes pontos da distribuição, o que pode levar a incongruências nas identificações de espécies dependendo do conceito considerado. Embora essa visão de que o isolamento reprodutivo ocorreria aos poucos já tivesse sido considerada mesmo na época de Dobzhansky (1937) e Muller (1942), razão pela qual os pesquisadores têm procurado genes de isolamento reprodutivo desde aquela época, a grande maioria dos pesquisadores considerou que a especiação, representada pelo isolamento no conceito biológico de espécie, não aconteceria apenas pela variação em um pequeno número de genes. Em vez disso, estes se expandiriam rapidamente para complexos de genes coadaptados em diferentes espécies e, eventualmente, seriam fixados, por exemplo, em inversões cromossômicas e polimorfismos balanceados, de acordo com Dobzhansky, de modo que, em geral, haveria pouca introgressão no genoma de diferentes espécies (Rieseberg e Burke, 2001).

4. A Interação entre Fluxo Gênico e Deriva Genética na Estrutura Populacional

Como mencionado anteriormente, um maior conhecimento da estrutura populacional e sua estrutura genética em diferentes grupos revelou uma presença mais disseminada de introgressões passadas e atuais influenciando a segregação da variação genética e produzindo padrões filogenéticos mais complexos. É possível que essa difusão do fluxo gênico e da introgressão não tenha sido detectada antes devido à dificuldade de identificar tais processos usando ferramentas genéticas disponíveis anteriormente, que eram baseadas na avaliação de mudanças de frequência de alelos em um ou alguns loci, ou mesmo um conjunto de vários microssatélites, para os quais as informações históricas são difíceis ou mesmo impossíveis de reconstruir (Sirén *et al.* 2011, Putman e Carbone, 2014).

Quando as populações são estruturadas, isto é, quando as espécies são divididas em populações que mostram diferenças nas frequências de alelos e em seu histórico de segregação, a distribuição da variação dentro e entre as populações é controlada por outro equilíbrio entre duas forças evolutivas, neste caso entre o fluxo gênico e a deriva genética. Assim como antes, se esse equilíbrio for controlado pelo fluxo gênico, a espécie tende a permanecer como uma única unidade evolutiva que compartilha a maior parte da variação genética e sua história entre diferentes indivíduos e partes de seus genomas. Por outro lado, se o equilíbrio for controlado pela deriva, as populações tendem a se tornar cada vez mais diferenciadas, mas também tendem a aumentar a endogamia dentro de cada uma dessas subpopulações, uma vez que têm tamanhos menores. Esse processo pode eventualmente levar à especiação, embora isso possa levar algum tempo ou especialmente em populações relativamente pequenas.

Se considerarmos populações que divergiram recentemente, a deriva genética teve pouco tempo para operar, então essas populações ainda segregariam polimorfismos antigos herdados da população ancestral. Com o tempo, a deriva e outras forças evolutivas, como a seleção, levariam esses polimorfismos à fixação, o que faria com que diferentes regiões do genoma abrigassem diferentes histórias evolutivas, refletindo o processo estocástico de segregação e fixação de polimorfismos. Como se considera que a maior parte da variação genética que segrega em populações é neutra (Kimura, 1991), esse processo é determinado principalmente pelo equilíbrio das taxas de deriva genética e mutação, que são ponderadas pelos tamanhos efetivos da população. Embora esse processo não tenha impacto na taxa de fixação de alelos neutros, independentemente do tamanho da população, a maneira como isso acontece varia em populações com tamanhos diferentes. Grandes populações segregam mais polimorfismos, enquanto populações pequenas têm maior chance de fixação e extinção. Dessa forma, populações maiores têm maior chance de reter polimorfismos ancestrais, quando comparadas a populações menores. Além disso, esse processo seria acelerado se houvesse seleção natural favorecendo a diversificação em diferentes populações, e seria reduzido se houvesse seleção de balanceamento. Esta última, de fato, pode complicar a identificação de regiões segregantes entre diferentes populações (Guerrero e Hahn, 2017), porque os polimorfismos ainda segregariam mesmo após a separação completa da linhagem.

Embora o processo de fixação de polimorfismos ancestrais por deriva seja estocástico, ele ainda é moldado pela história populacional subjacente, o que significa que inferências filogenéticas de diferentes regiões do genoma (chamadas de árvores genéticas) podem apresentar topologias

distintas umas das outras e da árvore de espécies subjacente. Esse processo seria exacerbado em populações recentemente divergentes, especialmente se elas tivessem grandes tamanhos efetivos, e seria menos importante em populações menores e com o passar do tempo (Carstens e Knowles, 2007). A segregação desses polimorfismos ancestrais pode levar a inferências filogenéticas potencialmente incongruentes derivadas de diferentes marcadores. Portanto, pode haver menos confiança em inferências filogenéticas que são baseadas em um pequeno número de marcadores, uma vez que elas podem não revelar adequadamente a história evolutiva global. No entanto, as informações filogenéticas obtidas de marcadores em todo o genoma podem ser usadas para inferir uma história evolutiva única e mais completa representada em uma árvore de espécies (Pamilo e Nei 1988). Para esse efeito, vários métodos foram propostos para inferir a árvore de espécies subjacente mais provável, mesmo na presença de múltiplas árvores genéticas que exibem topologias diversas (Drummond e Rambaut 2007; Larget *et al.* 2010; Drummond *et al.* 2012; Corl e Ellegren 2013; Jones, 2017).

Algumas abordagens que são genericamente chamadas de modelos de delimitação de espécies buscam inferir se árvores filogenéticas derivadas de diferentes partes do genoma permitem determinar se há estrutura genética em populações e restrição de processos de fluxo gênico (Sukumaran e Knowles, 2017). Isso é frequentemente tomado erroneamente como uma indicação de que essas unidades identificadas seriam espécies diferentes, embora essas estratégias de fato testem a existência de linhagens evolutivas que podem representar populações em evolução relativamente independente, ou talvez espécies distintas, usando o conceito de espécie filogenética. Essas estratégias estão se tornando cada vez mais poderosas e, embora tenham um problema potencial, já que a maioria delas assume que eventuais incongruências entre árvores gênicas e árvores de espécies seriam exclusivamente devido à estocasticidade da deriva gênica, algumas recentemente incorporaram o papel do fluxo gênico e da introgressão (Than *et al.* 2008; Jackson *et al.* 2017b; Solís-Lemus *et al.* 2017; Hibbins e Hahn, 2021). Os vários métodos que foram descritos, juntamente com a melhoria nas estratégias para produzir dados genômicos, mostram que diferentes táxons abrigam diferentes, mas relevantes, níveis de hibridização ao longo de sua história (Fontaine *et al.* 2015; Jones *et al.* 2018; Taylor e Manseau, 2020; Congrains *et al.* 2021; Pfennig, 2021) e há introgressão significativa entre diferentes espécies. Apenas para dar um exemplo, um grande estudo de 155 espécies de moscas-das-frutas *Drosophilid* as separou em nove grupos principais, oito dos quais mostraram evidências de introgressão (Suvorov *et al.* 2022). Agora parece que, ao contrário

do que se pensava anteriormente, filogenias sem evidências de fluxo gênico e introgressão são a exceção e não a regra (Edelman e Mallet, 2021).

5. Dinâmicas de Estrutura Populacional com Contato Secundário

Quando duas espécies trocam material genético, podemos antecipar três resultados diferentes: a troca de material genético entre espécies pode ser tão extensa que essas espécies se fundem em uma única; espécies podem experimentar introgressão envolvendo apenas partes do genoma, de alguns genes para uma gama mais ampla de genes; ou podem produzir híbridos que não levam à introgressão entre as espécies, seja porque são estéreis ou mesmo inviáveis (Allendorf *et al.* 2001), e isso pode causar um híbrido permanente ou zona de tensão onde há alguma sobreposição de distribuições de espécies. Todas essas possibilidades têm consequências importantes que serão consideradas aqui, mas primeiro investigaremos as condições que ajudam a definir qual desses resultados é mais provável quando espécies diferentes são encontradas em um ambiente.

Em geral, podemos definir introgressão como a persistência e fixação de alguns genes em uma população quando há uma troca de alelos entre dois grupos genéticos diferentes (Bohling, 2016). Embora doravante possamos nos referir a esses grupos como espécies, tal definição não se baseia em nenhum conceito específico de espécie, mas apenas reflete a indicação de que essas populações têm, em geral, diferentes histórias e variações genéticas. Em alguns casos, a introgressão promove variação fenotípica e genotípica, fornecendo ferramentas necessárias para a seleção natural atuar, favorecendo alelos que permitem que as populações se adaptem (Meier *et al.* 2017a). Nos referimos como introgressão adaptativa quando, seja por meio de fluxo gênico, hibridização, retrocruzamento, ligação e recombinação genética, a seleção fornece às espécies ou populações estruturadas maior adaptação e expansão de alcance. Como a principal limitação ao processo adaptativo é a disponibilidade de variabilidade genética, compartilhar uma parte do genoma de outra espécie pode permitir a recuperação da variação genética esgotada e oferecer respostas positivas às mudanças no ambiente. Isso pode permitir uma adaptação mais rápida a novos ambientes, permitindo a exploração de novos nichos, processos que estão se tornando muito importantes em ambientes em rápida mudança (Harrison e Larson, 2014; Schumer *et al.* 2018; Suarez-Gonzalez *et al.* 2018; Taylor e Larson, 2019; Jones *et al.* 2020; Pfennig, 2021), por exemplo, como resultado de impactos antrópicos e aquecimento global (Oziolor *et al.* 2019).

Ao passar pela peneira da seleção natural, a introgressão adaptativa pode ser, em alguns casos, uma importante fonte evolutiva de nova variação genética para espécies, especialmente se combinada com o impacto de interações pleiotrópicas e epistáticas com seu próprio background genético (Barton, 2001; Rieseberg *et al.* 2003; Tigano e Friesen, 2016; Suarez-Gonzalez *et al.* 2018). Isso ocorre principalmente porque é mais provável que os híbridos não sejam tão adaptados quanto suas espécies parentais a um nicho específico, o que tornaria a seleção mais provável de ser mais forte sobre eles (Mitchell *et al.* 2022). Considerando isso, a introgressão pode impulsionar a adaptação ecológica e ter um importante potencial de conservação, porque pode aumentar a resiliência evolutiva, mas não é provável que seja universal entre as espécies (Pfennig, 2021), a menos que seja para um número limitado de genes selecionados positivamente e aditivos para os quais a arquitetura genética permaneceria semelhante entre as espécies. Isso ocorreria porque esses processos provavelmente envolvem populações locais e interações locais, que podem exibir variação genética distinta de outras populações e mostrar adaptação ao ambiente local. Várias interações ecológicas, como coevolução e competição, são altamente influenciadas pela variação local, e esse também deve ser o caso da formação híbrida. Apenas muito raramente as espécies são panmíticas em sua distribuição, o que significa que a formação híbrida local ocorre em um subconjunto da variação genética total, o que pode levar à introgressão tendenciosa (Pfennig, 2021), causada por um aumento em cruzamentos heteroespecíficos localmente (Taylor *et al.* 2015; Abbott, 2017). Esse processo influencia os padrões de introgressão adaptativa não apenas no espaço, por exemplo, mais longe do centro da distribuição da espécie ou seu “ótimo ecológico”, o que pode fazer com que certos indivíduos que podem ter origens e características específicas, como comportamentos diferentes, sejam mais propensos à hibridização (Rosenthal, 2013; Taylor *et al.* 2015; Mitchell *et al.* 2022). Isso também levaria a uma potencial hibridização tendenciosa em seu genoma, particularmente se essas características forem controladas por poucos loci ou se estiverem evoluindo como supergenes. Um exemplo disso parece ser o caso de híbridos produzidos a partir de algumas espécies de *Helianthus* no Texas e Oklahoma que eram mais propensos a replicar combinações genômicas e algumas características adaptativas independentemente em sua distribuição (Mitchell *et al.* 2022). Nesse caso, há evidências de que partes dos genomas das espécies envolvidas evoluíram como supergenes com grandes unidades de seleção, limitando sua recombinação (Todesco *et al.* 2020).

6. Padrões Genômicos de Introgessão

Em geral, à medida que diferentes linhagens divergem, a incompatibilidade genética tende a aumentar e, conseqüentemente, as chances de hibridização e introgessão tendem a diminuir. Esta é uma das razões pelas quais Darwin acreditava que há descontinuidades na natureza. Essa relação varia muito entre diferentes grupos, não apenas por causa de atributos intrínsecos de cada grupo, mas também pela não aditividade do processo de diferenciação, que é supostamente influenciado pela evolução de interações epistáticas explicadas pelo modelo Bateson-Dobzhansky-Muller (Orr e Turelli, 2001). Neste modelo, substituições que podem ser neutras ou mesmo benéficas, no background genético de uma espécie, podem mostrar interações epistáticas negativas no contexto do background genético da outra espécie. O número de diferenças independentes entre espécies tende a aumentar à medida que divergem, seja por deriva, novas mutações ou impulsionadas pela seleção. Mesmo que o número dessas interações negativas, que na verdade tendem a aumentar quadraticamente, o que potencialmente pode levar com o tempo ao que foi chamado de efeito bola de neve (Matute *et al.* 2010; Presgraves, 2010). Espécies que divergiram mais recentemente neste modelo são mais propensas a compartilhar origens genéticas semelhantes, tornando mais provável que compartilhem os padrões de expressão de loci introgrididos, sejam eles neutros, selecionados positiva ou negativamente, o que aumenta a chance de que a introgessão possa ser mais disseminada pelo genoma. Essa introgessão pode aumentar a variação genética total pelo genoma e, embora alelos não dominantes fracamente deletérios possam persistir. Ela tenderia a envolver alelos neutros e fracamente vantajosos, porque os genes introgrididos já resistiram ao teste de seleção e sobrevivência em uma espécie diferente, ao contrário de novos alelos que surgiram por mutação (Templeton, 2021).

Os níveis de introgessão e a posição dos genes introgressados no genoma devem, portanto, ser influenciados pelo tempo desde a divergência, mas também pelas forças de seleção que atuam nessas regiões genômicas introgressadas. À medida que a divergência das espécies aumenta, portanto, torna-se menos provável que a variação no genoma flua livremente entre essas espécies, a menos que esses genes introgressados sejam adaptativos (Edmands, 2007), o que parece ser o caso até mesmo para algumas espécies icônicas, como os tentilhões darwinianos (Grant e Grant, 2021). Esse processo é difícil de generalizar, porque a arquitetura genética dos genes adaptativos varia muito, de ser determinada por um único locus a uma série de vários genes, em geral com pequenos efeitos, e essa adaptação depende também do ambiente ecológico e genético (epistasia)

onde essa variação é expressa, devido à própria natureza da segregação dessa variação. No entanto, à medida que a divergência aumenta, torna-se mais provável que apenas loci muito localizados no genoma, com efeito cada vez maior, estejam envolvidos nessa introgressão. Este é o caso da introgressão de genes de resistência a inseticidas da traça-praga *Helicoverpa armigera* em populações sul-americanas de *Helicoverpa zea* (Cordeiro *et al.* 2020; Valencia-Montoya *et al.* 2020), embora existam vários exemplos de alelos neutros ou ligeiramente deletérios que ainda introgridem através das fronteiras das espécies (Wen e Nakhleh, 2018; Cruzan *et al.* 2021; Nikolakis *et al.* 2022).

A existência de regiões no genoma que são resistentes à introgressão pode ser um dos impulsos evolutivos para a especiação que são mais comuns para espécies simpátricas, onde barreiras biológicas impedem o cruzamento de diferentes espécies que coexistem na mesma região geográfica. Dados recentes mostraram que, pelo menos por algum tempo, elas não parecem suficientes para limitar completamente a introgressão (Nosil *et al.* 2009; Wolf e Ellegren, 2017). Nesse caso, regiões do genoma com maiores taxas de recombinação têm maior tendência a serem permeáveis à introgressão, porque a seleção pode ser mais eficaz em alelos não ligados e pode se recombinar para longe de alelos deletérios (Roesti *et al.* 2012; Aeschbacher *et al.* 2017; Kim *et al.* 2018; Edelman e Mallet, 2021). Frequências e taxas de recombinação em regiões genômicas devem ser vistas com cautela, porque elas mudam entre diferentes espécies e cromossomos, mas, em geral, tendem a ser menores perto do centrômero e maiores nos telômeros, embora haja exceções devido à recombinação heterogênea em outras regiões do cromossomo. É possível que loci deletérios estejam ligados em regiões de baixa recombinação ou abriguem alelos polimórficos, vantajosos e adaptativos, e isso impacta a introgressão, bem como o isolamento reprodutivo (Kirkpatrick e Barton, 2006; Roesti *et al.* 2012; Charlesworth e Barton, 2018). Considera-se então que regiões com alta taxa de recombinação seriam mais permeáveis à introgressão (Schumer *et al.* 2018), enquanto regiões com menos recombinação mostrariam maior diferenciação entre espécies, um padrão que foi observado em diferentes sistemas naturais (Carneiro *et al.* 2014; Banker *et al.* 2022). Há muitos fatores que podem determinar o sucesso evolutivo de diferentes espécies, determinando seu sucesso ou extinção, e é possível que a introgressão seja um dos mais importantes por seus efeitos benéficos na diversidade e conservação de espécies.

7. Interação entre Fluxo Gênico, Seleção e Estrutura Populacional

Embora uma longa linha de pesquisa teórica e empírica tenha indicado que a viabilidade populacional é afetada positivamente pela variação genética (Lande e Shannon, 1996; Frankham e Ralls, 1998; Hohenlohe *et al.* 2021; Kardos *et al.* 2021), ainda é difícil sugerir uma ligação direta entre essas variáveis (Doyle *et al.* 2015), embora existam várias associações indiretas (Willi *et al.* 2022). Parte do problema decorre das diferenças inerentes entre diferentes espécies, mas algumas pessoas afirmam que um problema maior é que pesquisamos, em geral, a variação genética neutra, e não é isso que é usado pela seleção natural na adaptação (Teixeira e Huber, 2021). Por causa disso, existem várias abordagens que sugerem que devemos investigar e focar na variação genética funcional que impacta diretamente a aptidão, em vez de métodos padrão que visam preservar a variação genética em todo o genoma como um substituto para a primeira (Kyriazis *et al.* 2021). Embora o conhecimento da variação que impacta características relacionadas à aptidão para espécies ameaçadas de extinção claramente beneficiaria seu manejo, para várias espécies e cenários podemos não saber quais atributos e características seriam seletivamente importantes. Quando sabemos, em geral, esses atributos tendem a ser específicos da espécie e têm uma arquitetura genética complexa que torna sua identificação cara e demorada. Isso é tempo e recursos que, em geral, não podemos gastar. Além disso, hoje em dia, conseguimos pesquisar partes cada vez maiores do genoma a custos cada vez mais fáceis e baratos, o que facilita o uso padrão dos princípios da genética populacional e da evolução para trazer informações importantes não apenas para a conservação, mas também para lançar luz sobre partes do genoma envolvidas (Hoelzel *et al.* 2019; Mable 2019). Isso pode fornecer informações importantes de pedigree e parentesco que podem até ser usadas para fornecer informações demográficas que são inestimáveis para o manejo *in situ* e *ex situ* de populações ameaçadas de extinção (Galla *et al.* 2022).

Além disso, a investigação de partes do genoma que se introgridem entre diferentes espécies e seu potencial impacto na adaptação pode ser uma informação importante para ajudar a resolver esse enigma, porque ajudaria a identificar porções do genoma envolvidas com a adaptação a atributos ecológicos específicos, o que poderia ter enviesado a introgressão em seus genomas (Pfennig, 2021). Por exemplo, quando há introgressão entre diferentes espécies, podemos investigar se essa variação pode facilitar a adaptação das espécies receptoras a ambientes novos ou em mudança (Taylor *et al.* 2020; Blanco-Pastor, 2022), o que as faria expandir seus alcances e ocupar novos nichos e até mesmo evitar sua extinção local (Pfennig *et al.* 2016; Suarez-Gonzalez

et al. 2018; Oziolor *et al.* 2019; Hodel *et al.* 2022; Mitchell *et al.* 2022) . Essa estratégia tem sido usada, por exemplo, em grandes felinos, para revelar não apenas que há uma grande introgressão entre diferentes espécies, mas também quais genes introgridiram, levando à identificação de alguns genes de cor da pelagem e visão que podem ter sido importantes em sua diferenciação (Figueiro *et al.* 2017; Ramirez *et al.* 2022).

Como discutimos anteriormente, a probabilidade de introgressão e seu potencial impacto evolutivo são modulados pelo tamanho da população, tempo de divergência, bem como isolamento de outras populações. Em geral, grandes populações podem suportar mais prontamente o influxo de alelos deletérios ou levemente deletérios. Em grandes populações, indivíduos portadores dos alelos terão menor aptidão, o que não deve ser muito afetado pela epistasia, mais provavelmente fazendo com que essa carga genética seja mais prontamente eliminada, embora alelos deletérios recessivos possam se acumular, a menos que essas populações estejam passando por endogamia. Por outro lado, a introgressão de populações grandes para pequenas pode ser importante porque pode trazer nova variação genética para as últimas e neutralizar o impacto da deriva na depressão por endogamia e perda de variação genética, especialmente em populações muito pequenas ou isoladas.

Estratégias que promovem o fluxo gênico para recuperar a variação genética em populações pequenas ou isoladas têm sido comumente propostas, mas geralmente essa abordagem tem sido restrita a populações intimamente relacionadas, porque é sugerido que o risco de depressão por endogamia em alguns casos supera suas vantagens potenciais (Allendorf *et al.* 2001; Edmands 2007; Laikre, 2010). Alguns desses estudos, bem como programas de reprodução para espécies ameaçadas em cativeiro, comumente usam tecnologia de marcadores moleculares em marcadores genéticos neutros para avaliar o parentesco como um plano de manejo para aumentar a diversidade em uma população e reduzir o risco de extinção (Pacioni *et al.* 2018). Esses programas de manejo e recuperação exigem várias etapas de processamento e, entre elas, precisam ter genomas de referência das espécies de interesse, bem como de outras espécies filogeneticamente próximas, como o que foi feito com o pernilha-preta e outras aves Charadriiformes relacionadas (Galla *et al.* 2019). O problema é que, para várias espécies ameaçadas de extinção, pode não haver populações de origem alternativas para explorar, por isso foi sugerido que outras populações também deveriam ser consideradas (Jahner *et al.* 2019).

Alguns sugeriram recentemente que limitar o fluxo gênico apenas à população intraespecífica pode ser muito conservador (Bohling, 2016), especialmente porque vários casos foram encontrados na natureza ultimamente com sucesso notável (Schumer *et al.* 2018; Zecherle *et al.* 2021; Banker *et al.* 2022). Embora a introgressão possa produzir indivíduos com menos alelos deletérios do que o que estava originalmente presente nessas populações pequenas e potencialmente decrescentes, deve-se observar que isso deve ser cuidadosamente monitorado. Essa triagem é importante não apenas pelo risco de depressão por endogamia, mas também porque grandes populações podem ter acumulado um número maior de alelos deletérios recessivos ou parcialmente recessivos, que, devido ao seu tamanho, permanecem em frequências muito baixas. Se transferidos para populações menores, esses alelos letais e recessivos de pequeno efeito têm maior probabilidade de aumentar em frequência e serem expostos pela homozigose, levando à depressão por endogamia. Acredita-se que esse tenha sido o caso que envolveu a população de lobos de Isle Royale, que caiu após uma introdução que levou a vários descendentes sendo geneticamente relacionados e consanguíneos (Kyriazis *et al.* 2021). Por outro lado, novos alelos em um fundo genético diferente podem gerar uma nova combinação de alelos até então indisponíveis, o que poderia permitir que as populações cruzassem vales de aptidão, como aqueles apresentados em nichos diferentes. A transferência de genes de coloração de asas envolvidos com mimetismo entre diferentes espécies, fornecendo proteção contra predação em várias espécies de borboletas *Heliconius*, é um ótimo exemplo de como a nova variação genética pode permitir tais cruzamentos e promover a divergência ecológica (Zhang *et al.* 2016; Kozak *et al.* 2021; Meier *et al.* 2021).

Embora enfatizemos nesta revisão o impacto positivo da introgressão como um processo que pode facilitar a adaptação em alguns grupos de espécies, esse processo pode ter consequências negativas para a conservação da biodiversidade e levar à extinção de muitas populações e espécies, especialmente entre aquelas que são mais raras e menos frequentes localmente (Rhymer e Simberloff, 1996; Allendorf *et al.* 2001; Crispo *et al.* 2011). Nesse caso, populações que experimentam introgressão não adaptativa podem sofrer depressão por endogamia ou perda de habitat para populações híbridas (Allendorf *et al.* 2001). Mesmo traços antigos de introgressão, como observados em algumas espécies de lêmures (*Cheirogaleus sp. cf. medius*, *Cheirogaleus major*, *Cheirogaleus crossleyi* e *Cheirogaleus sibreei*) podem ter tido sua diversidade genética reduzida pela persistência de alelos introgrididos que influenciaram negativamente a adaptação a alguns ambientes (Williams *et al.* 2020). Além disso, a depressão por endogamia também pode ser

uma ameaça para grandes populações, mesmo com considerável diversidade genética, e usar populações com altos níveis de diversidade pode não ser eficaz na recuperação de pequenas populações (Robinson *et al.* 2019; Kyriazis *et al.* 2021), como indicado no caso dos lobos da Ilha Royale. Como o objetivo pode ser preservar fenótipos ou genótipos para fins evolutivos ou ecológicos, as decisões relativas à gestão ambiental envolvem instâncias políticas que podem ser complexas e demoradas (Coates *et al.* 2015; vonHoldt *et al.* 2018; Waples e Lindley, 2018). Talvez por esse motivo, a tomada de decisões sobre a conservação e preservação de ecossistemas deve levar em consideração dados que avaliem o status intrínseco de cada espécie visando preservar também seus ambientes. Uma possibilidade seria verificar o status das chamadas espécies “guarda-chuva”, que historicamente necessitam de grandes territórios de proteção e com elas, proteger, por extrapolação territorial, grande parte do habitat e outras espécies (Noss, 1990).

8. Efeitos Antropogênicos na Introgessão

Padrões históricos de fluxo gênico foram alterados por efeitos antropogênicos que trouxeram grandes mudanças ecológicas e ambientais (Thomas *et al.* 2004). Fatores como a introdução de novas espécies, dispersão não natural, fragmentação da paisagem e modificação do habitat contribuem para mudar a distribuição das espécies e levam à hibridização, introgessão e perda de variação intraespecífica (Rhymer e Simberloff, 1996). Além disso, espécies locais, evolutivamente bem estabelecidas, podem agora estar sujeitas à competição por nichos e recursos com espécies invasoras, ameaçando a biodiversidade (Seehausen *et al.* 2008; Crispo *et al.* 2011). Um dos exemplos clássicos é a truta cutthroat (*Oncorhynchus clarki lewisi*) que perdeu seu valor comercial após sofrer hibridização com a espécie introduzida truta arco-íris (*Oncorhynchus mykiss*), além da ameaça demográfica à população parental (Allendorf e Leary, 1988). Outro efeito pungente da introgessão é o que tem sido chamado de “reversão da especiação”, caracterizada pela perda da distinção fenotípica entre espécies que pode afetar os processos reprodutivos e o reconhecimento do parceiro sexual (Kearns *et al.* 2018), além de afetar as redes ecológicas e genéticas, levando a sérios declínios demográficos e extinção local (Vonlanthen *et al.* 2012; Kearns *et al.* 2018).

Um dos impactos antropogênicos mais relevantes, no entanto, tem sido as mudanças resultantes do aquecimento global. Os impactos causados pelas mudanças climáticas estão associados a diversas mudanças em diferentes táxons, por exemplo, em sua distribuição geográfica e densidade,

o que pode afetar processos ecológicos importantes em comunidades e ecossistemas (Franks e Hoffmann, 2012; Shaw e Etterson, 2012). Os impactos mais drásticos das mudanças climáticas podem levar à redução ou expansão do habitat das espécies e impactar os padrões de migração. Esses processos podem alterar a probabilidade de hibridização entre espécies, alterar o tempo de reprodução e, finalmente, afetar a endogamia e a introgressão de genes que podem não ser adaptativos (Vallejo-Marín e Hiscock, 2016; Quilodrán *et al.* 2020). Em ambientes em rápida mudança, a capacidade das espécies de se adaptar rapidamente se torna um fator limitante, mas pode ser reforçada quando a introgressão adaptativa está envolvida no processo, o que pode facilitar a dispersão de combinações de genes que apresentam maior plasticidade (Suarez-Gonzalez *et al.* 2018).

A capacidade de adaptação a novas condições ambientais e a velocidade com que esse processo ocorreu tornam mais provável que a seleção seja mais eficiente na variação genética existente, em vez de novas mutações (Orr e Unckless, 2014). Estudos sobre os efeitos das mudanças climáticas na introgressão ainda são incipientes, mas já sugerem que ela seja relevante. É o caso, por exemplo, da tolerância ao sal na espécie híbrida *Helianthus paradoxus* (Welch e Rieseberg 2002), da tolerância ao frio em árvores coníferas (*Picea sitchensis* e *Picea glauca*) (Hamilton *et al.* 2013), da formação de híbridos de borboletas entre *Lycaeides anna* e *Lycaeides melissa*, que parecem ser mais adaptáveis a novos ambientes (Nice *et al.* 2013), da tolerância térmica no copépode *Tigriopus californicus* (Pereira *et al.* 2014) e em espécies de corais do gênero *Acropora* (Mao *et al.* 2018). Em todos esses exemplos, os híbridos mostram picos de aptidão adaptativa, revelando que a recombinação interespecífica aumenta a diversidade genética para permanecer em condições extremas. Devido às mudanças climáticas cada vez mais aceleradas, a hibridização pode ser um fenômeno positivo em alguns casos, pois pode permitir a combinação e dispersão de novos genes e genótipos mais propícios à adaptação em condições extremas (Kremer *et al.* 2012).

Em contraste, a fragmentação populacional e as mudanças climáticas potencialmente ameaçam o fluxo gênico que ocorre entre populações locais (Aitken *et al.* 2008). Isso ocorre especialmente em populações pequenas, cujos pools genéticos não favorecem adaptações rápidas a mudanças ambientais, têm capacidade de migração limitada e/ou experimentaram o surgimento de barreiras geográficas que impedem a migração (Aitken *et al.* 2008; Wilczek *et al.* 2014). Durante expansões populacionais, a formação de novas zonas híbridas pode facilitar a introgressão (Buggs, 2007), que pode ser assimétrica para uma ampla gama de condições demográficas, incluindo quando espécies

invasoras têm uma densidade muito maior do que espécies locais ou quando há competição entre as duas espécies levando à extinção de espécies locais (Currat e Excoffier, 2004; Currat *et al.* 2008). Eventualmente, espera-se que a introgressão seja maior em espécies invasoras do que em espécies já estabelecidas, o que pode fazer com que tais genes aumentem em frequência em invasores e pode impactar a persistência de espécies naturais (Klopfstein *et al.* 2006; Quilodrán *et al.* 2020). Além disso, espécies cujas taxas de mudança genética são menores do que as taxas de mudança climática podem sofrer um declínio na aptidão ao longo do período evolutivo, levando a extinções locais (Atkins e Travis, 2010; Wang *et al.* 2010). Diante da mudança climática acelerada, os organismos precisam responder rapidamente por meio de processos de migração, mudanças nos parâmetros genéticos e por meio da plasticidade fenotípica que facilita o processo adaptativo (Hamilton e Miller 2016), bem como aumentar suas apêndices (a regra de Allen) (Ryding *et al.* 2021). Mudanças morfológicas como em corujas-do-mato (*Strix aluco*) (Karell *et al.* 2011), massa corporal e mudança na envergadura das asas em pássaros da Amazônia (Jirinec *et al.* 2021) e do Velho Mundo (Dubiner e Meiri, 2022) e seleção térmica em caracóis-pervinca-planos (*Littorina obtusata*) (Phifer-Rixey *et al.* 2008) são alguns exemplos de consequências de aumento da temperatura global. Populações cujas respostas não são rápidas o suficiente a essas mudanças climáticas correm maior risco de redução da aptidão e extinção (Aitken *et al.* 2008).

Programas de conservação têm usado resgate genético de espécies ameaçadas usando processos introgressivos, considerando cronogramas de adaptação e quanto da história evolutiva pode ser resgatada de subespécies extintas por meio de retrocruzamento (Hamilton e Miller, 2016). Um conceito ainda pouco conhecido de “desintrogressão” se aplica a isso, no qual por meio de retrocruzamento programado, linhagens perdidas e unidades genotípicas são recuperadas, uma estratégia usada para recuperar espécies extintas de tartarugas de Galápagos (*Chelonoidus spp.*) (Garrick *et al.* 2012; Edwards *et al.* 2013; Miller *et al.* 2017) e cavalo de Przewalski (*Equus przewalskii*) (Amador *et al.* 2011). Embora seja uma estratégia válida para estudos de conservação, especialmente em espécies ameaçadas para recuperar o pool genético original, estas podem sofrer com os efeitos potenciais de depressão por endogamia. Ao mesmo tempo, essa metodologia requer um alto nível de conhecimento sobre o genoma da espécie original e é difícil de ser usada em espécies selvagens e pequenas populações, tendo maior chance de sucesso quando a ancestralidade exógena é recente (Amador *et al.* 2013; Bech *et al.* 2022; vonHoldt *et al.* 2022).

9. Processos de Diferenciação e Fluxo Gênico em Ambientes Neotropicais

Os processos mencionados acima sugerem que espécies que divergiram recentemente têm maior probabilidade de introgreir, mostrando diferentes níveis de congruência filogenética entre diferentes características. Isso se deve à segregação de polimorfismos ancestrais ou porque a introgressão diferencial de genes pode levar a diferentes relações filogenéticas em regiões do genoma, especialmente porque alguns desses genes são impulsionados pela seleção e refletem o processo de uma rápida expansão de uma adaptação (Feder *et al.* 2005), mesmo entre fronteiras de espécies (Edelman *et al.* 2019; Nosil *et al.* 2021). Esse problema é especialmente complexo nos Neotrópicos, uma vez que várias espécies passaram por ciclos distintos de expansões e contrações de habitat ao longo de sua história e diferenciação (Avice *et al.* 1998; Knowles, 2001; de Brito *et al.* 2002) e podem ser separadas por fronteiras flutuantes, como rios sinuosos (Leite e Rogers, 2013; Naka e Brumfield, 2018). Como consequência, várias espécies passaram por períodos em que algumas populações ficaram temporariamente isoladas e interagiram localmente com outras espécies. Quando esses limites mudaram, seja por mudanças climáticas ou mudanças nas barreiras fluviais, as espécies com as quais elas interagem mudam, assim como suas condições ambientais. Em situações em que as interações ecológicas são intensas e importantes, como no caso de predação, parasitismo e competição, isso pode promover a diferenciação de linhagens específicas em diferentes partes da distribuição e promover introgressão diferencial em sua distribuição. Essas adaptações, de maneira semelhante ao que acontece com diferentes parceiros sexuais, reforçam soluções locais, por causa de suas complexas interações genéticas, o que cria o potencial para epistasia e processos não aditivos (Templeton, 2006).

Um dos exemplos para os quais as interações entre espécies podem ser importantes para sua diferenciação é o caso das verdadeiras espécies de moscas-das-frutas da família Tephritidae, uma vez que a maioria delas é fitófaga e põe seus ovos em frutas, atributos que criam interações ecológicas intensas (Christenson e Foote, 1960). A maior parte da diferenciação em Tephritidae é na verdade muito mais antiga e, conseqüentemente, mais complexa do que em outras espécies da família Tephritidae, o que parece ser o caso do gênero mais diverso da família, *Anastrepha*. Este gênero é amplamente distribuído na América do Sul e Central e abriga algumas das espécies de pragas mais importantes nas Américas (Zucchi, 2000, Norrbom *et al.* 2012). As espécies do gênero estão associadas a uma ampla gama de espécies de plantas, várias que mostram associações específicas de espécies, enquanto outras usam uma ampla gama de diferentes plantas hospedeiras.

Uma análise extensiva de *Anastrepha* usando vários genes nucleares revelou um padrão de divergência rápida e recente, com vários ramos curtos ainda não resolvidos, particularmente de espécies que divergiram recentemente (Mengual *et al.* 2017). Dados moleculares derivados de dados genômicos e transcriptômicos de diferentes espécies de *Anastrepha* com diferentes graus de parentesco filogenético revelaram grande incongruência filogenética em diferentes partes do genoma, embora, em geral, sua análise conjunta tenha fornecido grande suporte para diversas relações filogenéticas entre grupos de espécies e espécies alocadas a esses grupos (Congrains *et al.* 2021; Congrains *et al.* Subm.).

Talvez tão relevantes quanto as relações filogenéticas inferidas por vários estudos (Vaníčková *et al.* 2015; Dias *et al.* 2016; Mengual *et al.* 2017; Congrains *et al.* Subm.) são as descobertas que sugerem que houve um fluxo gênico considerável entre espécies de *Anastrepha* ao longo de sua diferenciação (Rezende *et al.* 2016; Díaz *et al.* 2018; Congrains *et al.* 2021). Esse fluxo gênico parece ter sido parte integrante de seu processo de diferenciação, uma vez que há evidências de que isso aconteceu mesmo entre espécies que divergiram há muito tempo, mas ainda mais para espécies intimamente relacionadas (Congrains *et al.* Subm.). Essa descoberta tem consequências importantes, dependendo de quais genes estão sendo trocados entre as espécies. Mesmo que os genes introgressivos não estejam envolvidos no processo adaptativo, eles podem pelo menos complicar o processo de identificação de marcadores específicos da espécie. Partes do genoma que não estão envolvidas na diferenciação podem ser compartilhadas, dependendo do grau de isolamento reprodutivo. Isso pode ajudar a explicar por que certos marcadores diferem entre populações e não servem como marcadores comuns para todas as populações da espécie, conforme observado (Thawornwattana *et al.* 2022, Congrains *et al.* Subm.).

Considerando que a seleção sexual e a adaptação do hospedeiro são importantes para a diferenciação de Tephritidae e esses processos podem estar envolvidos com mudanças comportamentais e genéticas, é possível que essas mudanças antecedam mudanças morfológicas observáveis e possam ajudar a explicar a existência de espécies crípticas em linhagens que divergiram recentemente, como pode ser o caso de espécies no complexo *fraterculus* (Dias *et al.* 2016). A existência de espécies crípticas pode dizer algo sobre o processo de diferenciação de espécies, pois pode indicar que caracteres que estão sendo importantes para sua diferenciação estão evoluindo em um ritmo diferente, talvez porque a taxa de evolução de caracteres morfológicos (ou traços morfológicos facilmente identificáveis) não seja a mesma que a taxa de evolução dos genes

anteriores (Struck *et al.* 2018), complicando assim a correta identificação das espécies, mesmo por especialistas.

Por sua vez, se o fluxo gênico abrange partes do genoma que estão envolvidas com o processo adaptativo, estas podem levar à produção de híbridos inviáveis ou estéreis, mas também podem levar à miscigenação em linhagens distintas de alelos com adaptações a um determinado contexto ambiental, como uma adaptação a um novo hospedeiro. Isso pode resultar em adaptações específicas presentes em um pequeno conjunto de indivíduos sendo compartilhadas com uma grande parte da população se isso trouxer vantagens adaptativas como, por exemplo, a capacidade de explorar uma nova planta hospedeira. Tal processo pode facilitar a geração e expansão de espécies polífagas e pode desempenhar um papel importante no processo adaptativo em um processo semelhante ao que aconteceu na introgressão do mimetismo mülleriano em *Heliconius* através das fronteiras de espécies (Martin *et al.* 2013; Moest *et al.* 2020), o que também, incidentalmente, tem consequências importantes para o controle de pragas, o que não é nosso foco aqui.

Pode parecer absurdo que o conhecimento de espécies de pragas, não exatamente aquelas que vêm à mente quando se pensa em conservação, possa contribuir para a compreensão da evolução e conservação de espécies nos Neotrópicos. Afirmamos, no entanto, que a combinação de divergência recente, isolamento reprodutivo potencialmente incompleto, adaptações a diferentes hospedeiros, seleção natural e sexual intensa e impactos antropogênicos tornam as espécies de *Anastrepha* um modelo apropriado para o estudo dos estágios incipientes da especiação e suas consequências na diferenciação e adaptação. Essa combinação de fatores pode, na verdade, ser muito comum para táxons neotropicais e pode ser uma das razões pelas quais uma abundância de espécies crípticas nos trópicos foi observada para vários grupos diferentes (Vieites *et al.* 2009; Funk *et al.* 2012; Gill *et al.* 2016; Huntley *et al.* 2019), então o estudo de *Anastrepha* pode ajudar a entender os padrões gerais de diversidade nos Neotrópicos.

Os Neotrópicos são caracterizados por grande diversidade de clima, solo, paisagem que cria uma ampla diversidade de ecossistemas (Mittermeier, 1986) e têm experimentado muitas oscilações climáticas desde o Pleistoceno (Vasconcellos e Colli, 2009). Essas mudanças climáticas do Pleistoceno foram até propostas por uma teoria controversa, chamada de 'refúgios florestais', que propôs que as contrações e expansões florestais produziram áreas historicamente preservadas que mantiveram alta diversidade, que favoreceram o isolamento reprodutivo entre espécies

presentes em diferentes comportamentos de refúgio entre espécies para gerar hibridização entre elas, refletindo na estrutura genética histórica das espécies (Haffer, 1969; Vanzolini e Williams 1981). Independentemente da existência real desses refúgios e áreas com alta diversidade (Bush 1994; Colinvaux, 1996), é provável que o aumento e diminuição da floresta durante as mudanças climáticas tenham afetado a distribuição das espécies e possam ter promovido o surgimento de zonas híbridas (Pulido-Santacruz *et al.* 2018; Pulido-Santacruz *et al.* 2020) ou permitido curtos períodos de introgressão entre espécies durante corredores florestais temporários na região Amazônica (Batalha-Filho *et al.* 2013). A região Amazônica, em particular, devido à sua extensão, limitou barreiras a vários ecossistemas diferentes que podem favorecer o fluxo gênico entre espécies (Dexter *et al.* 2017), como rios que promovem a diversificação de espécies de aves (Naka e Brumfield 2018), mas caracteristicamente têm cursos sinuosos que podem facilitar a introgressão. Além disso, a profusão de ambientes e ecossistemas produz 'mosaicos de paisagens' mais complexos que facilitam o surgimento de zonas híbridas (Rieseberg *et al.* 1999).

A paisagem complexa e a história dos Neotrópicos, com essas barreiras geográficas fracas ou semipermeáveis, podem promover o processo de hibridização e introgressão, especialmente se a nova informação genética fornecer habilidades para aumentar a distribuição geográfica por meio da adaptação a novos habitats, o que também pode criar oportunidades para especiação alopátrica e radiação adaptativa (Marques *et al.* 2019). O maior número de espécies mais intimamente relacionadas que são encontradas próximas umas das outras nos Neotrópicos pode aumentar essas fronteiras semipermeáveis que facilitam a introgressão, em um processo que pode ser o equivalente genético de fronteiras ecológicas flutuantes. Assim como em complexos de espécies, espécies crípticas também podem surgir de linhagens híbridas que passaram por processos de introgressão, e isso pode ser facilitado pela estrutura em escala de paisagem em biomas neotropicais que discutimos acima (Hebert *et al.* 2004; Elmer *et al.* 2007), como demonstrado, por exemplo, em espécies crípticas de orquídeas *Epidendrum* (Vega *et al.* 2013), introgressão e hibridização em sapos neotropicais (Sequeira *et al.* 2011; Vallinoto *et al.* 2017) e em espécies neotropicais de moscas-das-frutas (Congrains *et al.* 2021). Apesar de sua grande importância e vasta distribuição, ainda sabemos muito pouco sobre o impacto da introgressão e hibridização na diversidade e diferenciação neotropicais, e mesmo espécies potencialmente pragas podem trazer insights importantes para ajudar a entender sua evolução.

Conclusão

Programas de recuperação e conservação de espécies não devem considerar apenas medidas genéticas como a única fórmula para salvar espécies da extinção. Medidas de diversidade genética, por exemplo, não podem ser a única maneira de quantificar e qualificar se o status de uma população ou espécie pode ser considerado saudável para mantê-la em seu habitat. Deve ser apenas uma medida preditiva (Spielman *et al.* 2004; Teixeira e Huber, 2021). Não podemos desconsiderar que os impactos da introgressão podem ser benéficos para populações naturais, uma vez que desempenham um papel importante na evolução adaptativa, conforme exemplificado acima. Em alguns casos, ela fornece vantagens adaptativas para espécies que sofrem depressão por endogamia que podem ser revertidas, além de produzir novas espécies e linhagens diferencialmente adaptadas ao ambiente, especialmente porque fornece meios para uma adaptação ambiental mais rápida (Edelman e Mallet, 2021). O impacto da introgressão de espécies, bem como outros processos evolutivos, ainda não é muito bem explorado em espécies neotropicais, embora seja bem documentado em regiões temperadas. Em parte por isso, espécies recentemente divergentes, como as espécies do gênero *Anastrepha* aqui discutidas, podem ser modelos biológicos importantes para ajudar a elucidar histórias evolutivas complexas e fornecer modelos mais abrangentes para espécies neotropicais.

Um dos maiores desafios do reconhecimento do impacto histórico da introgressão em várias linhagens evolutivas diferentes, e a observação de que é muito mais comum do que se considerava anteriormente, é como isso impacta a definição do que constitui uma espécie (Mallet, 2008; Abbott *et al.* 2013; Shapiro *et al.* 2016). Começamos este capítulo descrevendo várias visões diferentes sobre conceitos de espécie e indicando a importância de “espécies” em diferentes campos do conhecimento, mas gostaríamos de terminar enfatizando que a introgressão tem grande consequência na conservação. Aspectos políticos importantes, e como isso foi escrito em leis e protocolos, usam definições de espécies que dificultam o reconhecimento ou limitam o valor dos híbridos, chegando até o ponto de remover o status de conservação protetora de alguns. Isso tem sido realmente promovido por vários biólogos conservacionistas ao longo do tempo, considerando que em vários casos, os híbridos são bicos sem saída evolutivos e colocariam em risco a sobrevivência das espécies originais (Rhymer e Simberloff, 1996; Levin, 2002), embora isso tenha sido demonstrado ser uma causa rara de extinção de espécies (Draper *et al.* 2021). Se considerarmos que parece que a maioria das espécies na natureza experimentou pelo menos algum nível de

introgressão ao longo de sua história, e isso é mais provável que aumente com as mudanças antropogênicas que estão acontecendo atualmente, talvez uma reconsideração dessas conclusões seja justificada à luz da ideia proposta por Darwin de que as espécies seriam construções um tanto arbitrárias que usamos para caracterizar e definir a biodiversidade. À medida que aumentamos nosso conhecimento das relações entre espécies e seu contexto histórico, esse conhecimento deve nos ajudar a estabelecer melhor critérios para preservar diferentes grupos de populações, indivíduos ou linhagens que garantiriam sua sobrevivência a longo prazo, bem como de seu ambiente.

Referências

- ABBOTT, R.; ALBACH, D.; ANSELL, S. *et al.* Hybridization and speciation. **J Evol Biol** 26: 229-246. 2013. DOI: [10.1111/j.1420-9101.2012.02599](https://doi.org/10.1111/j.1420-9101.2012.02599).
- ABBOTT, R. J. Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. **J Syst Evol**, 55: 238-258. 2017. <https://doi.org/10.1111/jse.12267>
- AESCHBACHER, S.; SELBY, J. P.; WILLIS, J. H. *et al.* Population-genomic inference of the strength and timing of selection against gene flow. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, p. 7061-7066, 2017. DOI: [10.1073/pnas.1616755114](https://doi.org/10.1073/pnas.1616755114).
- AITKEN, S. N.; YEAMAN, S.; HOLLIDAY, J. A. *et al.* Adaptation, migration or extirpation: climate change outcomes for tree populations. **Evolutionary Applications**, v. 1, p. 95-111, 2008. DOI: <https://doi.org/10.1111/j.1752-4571.2007.00013.x>.
- ALLENDORF, F. W.; LEARY, R. F. Conservation and distribution of genetic variation in a polytypic species, the cutthroat trout. **Conservation Biology**, v. 2, p. 170-184, 1988. DOI: <https://doi.org/10.1111/j.1523-1739.1988.tb00168.x>.
- ALLENDORF, F. W.; LEARY, R. F.; SPRUELL, P. *et al.* The problems with hybrids: setting conservation guidelines. **Trends in Ecology & Evolution**, v. 16, p. 613-622, 2001. DOI: [https://doi.org/10.1016/S0169-5347\(01\)02290-X](https://doi.org/10.1016/S0169-5347(01)02290-X).
- AMADOR, C.; FERNÁNDEZ, J.; MEUWISSEN, T. H. E. Advantages of using molecular coancestry in the removal of introgressed genetic material. **Genetics Selection Evolution**, v. 45, p. 13, 2013. DOI: [10.1186/1297-9686-45-13](https://doi.org/10.1186/1297-9686-45-13).
- AMADOR, C.; TORO, M. Á.; FERNÁNDEZ, J. Removing exogenous information using pedigree data. **Conservation Genetics**, v. 12, p. 1565-1573, 2011. DOI: [10.1007/s10592-011-0255-4](https://doi.org/10.1007/s10592-011-0255-4).
- ANDERSON, E.; STEBBINS Jr, G. L. Hybridization as an evolutionary stimulus. **Evolution**, v. 8, p. 378-388, 1954. DOI: <https://doi.org/10.1111/j.1558-5646.1954.tb01504.x>.
- ATKINS, K. E.; TRAVIS, J. M. J. Local adaptation and the evolution of species' ranges under climate change. **Journal of Theoretical Biology**, v. 266, p. 449-457, 2010. DOI: <https://doi.org/10.1016/j.jtbi.2010.07.014>.

- AVISE, J. C.; WALKER, D.; JOHNS, G. C. Speciation durations and Pleistocene effects on vertebrate phylogeography. **Proceedings of the Royal Society of London. Series B, Biological Sciences**, v. 265, p. 1707-1712, 1998.
- BANKER, S. E.; BONHOMME, F.; NACHMAN, M. W. Bidirectional introgression between *Mus musculus domesticus* and *Mus spretus*. **Genome Biology and Evolution**, v. 14, 2022. DOI: 10.1093/gbe/evab288.
- BARRERA-GUZMÁN, A. O.; ALEIXO, A.; SHAWKEY, M. D. et al. Hybrid speciation leads to novel male secondary sexual ornamentation of an Amazonian bird. **Proceedings of the National Academy of Sciences of the United States of America**, 2017. DOI: 10.1073/pnas.1717319115.
- BARTON, N. H. The role of hybridization in evolution. **Molecular Ecology**, v. 10, p. 551-568, 2001. DOI: <https://doi.org/10.1046/j.1365-294x.2001.01216.x>.
- BATALHA-FILHO, H.; FJELDSÅ, J.; FABRE, P.-H. et al. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. **Journal of Ornithology**, v. 154, p. 41-50, 2013. DOI: 10.1007/s10336-012-0866-7.
- BECH, N.; NIVELLE, D.; CARON, S. et al. Extent of introgressive hybridization in the Hermann's tortoise (*Testudo hermanni hermanni*) from the south of France. **European Journal of Wildlife Research**, v. 68, p. 37, 2022. DOI: 10.1007/s10344-022-01585-8.
- BLANCO-PASTOR, J. L. Alternative modes of introgression-mediated selection shaped crop adaptation to novel climates. **Genome Biology and Evolution**, v. 14, evac107, 2022. DOI: 10.1093/gbe/evac107.
- BOHLING, J. H. Strategies to address the conservation threats posed by hybridization and genetic introgression. **Biological Conservation**, v. 203, p. 321-327, 2016. DOI: <https://doi.org/10.1016/j.biocon.2016.10.011>.
- BUGGS, R. J. A. Empirical study of hybrid zone movement. **Heredity**, v. 99, p. 301-312, 2007. DOI: 10.1038/sj.hdy.6800997.
- BUSH, M. B. Amazonian speciation: a necessary complex model. **Journal of Biogeography**, v. 21, p. 5-17, 1994.
- CARNEIRO, M.; ALBERT, F. W.; AFONSO, S. et al. The genomic architecture of population divergence between subspecies of the European rabbit. **PLoS Genetics**, v. 10, e1003519, 2014. DOI: 10.1371/journal.pgen.1003519.

- CARSTENS, B. C.; KNOWLES, L. L. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. **Systematic Biology**, v. 56, p. 400-411, 2007. DOI: 10.1080/10635150701405560.
- CHARLESWORTH, B.; BARTON, N. H. The spread of an inversion with migration and selection. **Genetics**, v. 208, p. 377-382, 2018. DOI: 10.1534/genetics.117.300426.
- CHRISTENSON, L. D.; FOOTE, R. H. Biology of fruit flies. **Annual Review of Entomology**, v. 5, p. 171-192, 1960. DOI: 10.1146/annurev.en.05.010160.001131.
- COATES, D. J.; MCARTHUR, S. L.; BYRNE, M. Significant genetic diversity loss following pathogen driven population extinction in the rare endemic *Banksia brownii* (Proteaceae). **Biological Conservation**, v. 192, p. 353-360, 2015. DOI: <https://doi.org/10.1016/j.biocon.2015.10.013>.
- COLINVAUX, P. A. Quaternary environmental history and forest diversity in the Neotropics. In: JACKSON, J. B. C.; BUDD, A. F.; COATES, A. G. (Eds.). **Evolution and Environment in Tropical America**. Chicago: The University of Chicago Press, 1996. p. 359-405.
- CONGRAINS, C.; DUPUIS, J. R.; STECK, G. J. et al. Phylogenomic analysis provides diagnostic tools for identification of species complex. **Evolutionary Applications**, Submetido.
- CONGRAINS, C.; ZUCCHI, R. A.; DE BRITO, R. A. Phylogenomic approach reveals strong signatures of introgression in the rapid diversification of neotropical true fruit flies (*Anastrepha*: Tephritidae). **Molecular Phylogenetics and Evolution**, v. 162, 107200, 2021. DOI: <https://doi.org/10.1016/j.ympev.2021.107200>.
- CORDEIRO, E. M. G.; PANTOJA-GOMEZ, L. M.; DE PAIVA, J. B. et al. Hybridization and introgression between *Helicoverpa armigera* and *H. zea*: an adaptational bridge. **BMC Evolutionary Biology**, v. 20, p. 61, 2020. DOI: 10.1186/s12862-020-01621-8.
- CORL, A.; ELLEGREN, H. Sampling strategies for species trees: the effects on phylogenetic inference of the number of genes, number of individuals, and whether loci are mitochondrial, sex-linked, or autosomal. **Molecular Phylogenetics and Evolution**, v. 67, p. 358-366, 2013. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1055790313000596>.
- CRACRAFT, J. Species concepts and speciation analysis. In: JOHNSTON, R. F. (Ed.). **Current Ornithology**. New York: Springer, 1983. p. 159-188.

CRISPO, E.; MOORE, J. S.; LEE-YAW, J. A. et al. Broken barriers: human-induced changes to gene flow and introgression in animals: an examination of the ways in which humans increase genetic exchange among populations and species and the consequences for biodiversity. **BioEssays**, v. 33, p. 508-518, 2011. DOI: 10.1002/bies.201000154.

Aqui estão as referências formatadas conforme as normas ABNT 2024:

CRUZAN, M. B.; THOMPSON, P. G.; DIAZ, N. A. et al. Weak coupling among barrier loci and waves of neutral and adaptive introgression across an expanding hybrid zone. **Evolution**, v. 75, p. 3098-3114, 2021. DOI: <https://doi.org/10.1111/evo.14381>.

CURRAT, M.; EXCOFFIER, L. Modern humans did not admix with Neanderthals during their range expansion into Europe. **PLoS Biology**, v. 2, e421, 2004. DOI: 10.1371/journal.pbio.0020421.

CURRAT, M.; RUEDI, M.; PETIT, R. J. et al. The hidden side of invasions: massive introgression by local genes. **Evolution**, v. 62, p. 1908-1920, 2008. DOI: 10.1111/j.1558-5646.2008.00413.x.

DARWIN, C. *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*. London: John Murray, 1859.

DE BRITO, R. A.; MANFRIN, M. H.; SENE, F. M. Nested Cladistic Analysis of Brazilian Populations of *Drosophila serido*. **Molecular Phylogenetics and Evolution**, v. 22, p. 131-143, 2002.

Aqui estão as referências formatadas conforme as normas ABNT 2024:

DE QUEIROZ, K. Species concepts and species delimitation. **Systematic Biology**, v. 56, p. 879-886, 2007. DOI: 10.1080/10635150701701083.

DEXTER, K. G.; LAVIN, M.; TORKE, B. M. et al. Dispersal assembly of rain forest tree communities across the Amazon basin. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, p. 2645-2650, 2017. DOI: 10.1073/pnas.1613655114.

DIAS, V. S.; SILVA, J. G.; LIMA, K. M. et al. An integrative multidisciplinary approach to understanding cryptic divergence in Brazilian species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Biological Journal of the Linnean Society**, v. 117, p. 725-746, 2016. DOI: 10.1111/bij.12712.

- DÍAZ, F.; LUIZ, A. LIMA, A.; NAKAMURA, A. M. et al. Evidence for introgression among three species of the *Anastrepha fraterculus* group, a radiating species complex of fruit flies. **Frontiers in Genetics**, v. 9, 359, 2018. DOI: 10.3389/fgene.2018.00359.
- DOBZHANSKY, T. **Genetics and the origin of species**. New York: Columbia University, 1937.
- Aqui estão as referências formatadas conforme as normas ABNT 2024:
- DOYLE, J. M.; HACKING, C. C.; WILLOUGHBY, J. R. et al. Mammalian genetic diversity as a function of habitat, body size, trophic class, and conservation status. **Journal of Mammalogy**, v. 96, p. 564-572, 2015. DOI: 10.1093/jmammal/gyv061.
- DRAPER, D.; LAGUNA, E.; MARQUES, I. Demystifying negative connotations of hybridization for less biased conservation policies. **Frontiers in Ecology and Evolution**, v. 9, 637100, 2021. DOI: 10.3389/fevo.2021.637100.
- DRUMMOND, A. J.; RAMBAUT, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, v. 7, p. 214, 2007. DOI: 10.1186/1471-2148-7-214.
- DRUMMOND, A. J.; SUCHARD, M. A.; XIE, D. et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7. **Molecular Biology and Evolution**, v. 29, p. 1969-1973, 2012. Disponível em: <http://dx.doi.org/10.1093/molbev/mss075>.
- DUBINER, S.; MEIRI, S. Widespread recent changes in morphology of Old World birds, global warming the immediate suspect. **Global Ecology and Biogeography**, v. 31, p. 791-801, 2022. DOI: <https://doi.org/10.1111/geb.13474>.
- EDELMAN, N. B.; FRANDSEN, P. B.; MIYAGI, M. et al. Genomic architecture and introgression shape a butterfly radiation. **Science**, v. 366, p. 594+, 2019. DOI: 10.1126/science.aaw2090.
- EDELMAN, N. B.; MALLET, J. Prevalence and adaptive impact of introgression. **Annual Review of Genetics**, v. 55, p. 265-283, 2021. DOI: 10.1146/annurev-genet-021821-020805.
- EDMANDS, S. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. **Molecular Ecology**, v. 16, p. 463-475, 2007. DOI: 10.1111/j.1365-294X.2006.03148.x.
- EDWARDS, D. L.; EDGAR, B.; RYAN, C. G. et al. The genetic legacy of Lonesome George survives: Giant tortoises with Pinta Island ancestry identified in Galápagos. **Biological Conservation**, v. 157, p. 225-228, 2013. DOI: 10.1016/j.biocon.2012.10.014.

- EDWARDS, S. V.; XI, Z.; JANKE, A. et al. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. **Molecular Phylogenetics and Evolution**, v. 94, p. 447-462, 2016. DOI: 10.1016/j.ympev.2015.10.027.
- EHRLICH, P. R.; RAVEN, P. H. Differentiation of populations. **Science**, v. 165, p. 1228-1232, 1969. DOI: 10.1126/science.165.3899.1228.
- ELMER, K. R.; DÁVILA, J. A.; LOUGHEED, S. C. Cryptic diversity and deep divergence in an upper Amazonian leaf litter frog, *Eleutherodactylus ockendeni*. **BMC Evolutionary Biology**, v. 7, p. 247, 2007. DOI: 10.1186/1471-2148-7-247.
- FEDER, J. L.; XIE, X.; RULL, J. et al. Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, p. 6573-6580, 2005. Disponível em: http://www.pnas.org/cgi/content/abstract/102/suppl_1/6573. Acesso em: 12 dez. 2024.
- FIGUEIRO, H. V.; LI, G.; TRINDADE, F. J. et al. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. **Science Advances**, v. 3, e1700299, 2017. DOI: 10.1126/sciadv.1700299.
- FONTAINE, M. C.; PEASE, J. B.; STEELE, A. et al. Extensive introgression in a malaria vector species complex revealed by phylogenomics. **Science**, v. 347, 2015. Disponível em: <http://www.sciencemag.org/content/347/6217/1258524.abstract>. Acesso em: 12 dez. 2024.
- FRANKHAM, R.; RALLS, K. Inbreeding leads to extinction. **Nature**, v. 392, p. 441-442, 1998. DOI: 10.1038/33022.
- FRANKS, S. J.; HOFFMANN, A. A. Genetics of climate change adaptation. **Annual Review of Genetics**, v. 46, p. 185-208, 2012. DOI: 10.1146/annurev-genet-110711-155511.
- FUNK, W. C.; CAMINER, M.; RON, S. R. High levels of cryptic species diversity uncovered in Amazonian frogs. **Proceedings of the Royal Society B: Biological Sciences**, v. 279, p. 1806-1814, 2012. DOI: 10.1098/rspb.2011.1653.
- GALLA, S. J.; BROWN, L.; COUCH-LEWIS, Y. et al. The relevance of pedigrees in the conservation genomics era. **Molecular Ecology**, v. 31, p. 41-54, 2022. DOI: 10.1111/mec.16192.
- GALLA, S. J.; FORSDICK, N. J.; BROWN, L. et al. Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform

- conservation management. **Genes**, v. 10, n. 1, p. 9, 2019. Disponível em: <https://www.mdpi.com/2073-4425/10/1/9>. Acesso em: 12 dez. 2024.
- GARRICK, R. C.; BENAVIDES, E.; RUSSELLO, M. A. et al. Genetic rediscovery of an ‘extinct’ Galápagos giant tortoise species. **Current Biology**, v. 22, p. R10-R11, 2012. DOI: 10.1016/j.cub.2011.12.004.
- GILL, B. A.; KONDRATIEFF, B. C.; CASNER, K. L. et al. Cryptic species diversity reveals biogeographic support for the ‘mountain passes are higher in the tropics’ hypothesis. **Proceedings of the Royal Society B: Biological Sciences**, v. 283, p. 20160553, 2016. DOI: 10.1098/rspb.2016.0553.
- GRANT, P. R.; GRANT, B. R. Morphological ghosts of introgression in Darwin’s finch populations. **Proceedings of the National Academy of Sciences of the United States of America**, v. 118, e2107434118, 2021. DOI: 10.1073/pnas.2107434118.
- GUERRERO, R. F.; HAHN, M. W. Speciation as a sieve for ancestral polymorphism. **Molecular Ecology**, v. 26, p. 5362-5368, 2017. DOI: 10.1111/mec.14290.
- HAFFER, J. Speciation in Amazonian forest birds. **Science**, v. 165, p. 131-137, 1969.
- HAMILTON, J. A.; LEXER, C.; AITKEN, S. N. Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). **Molecular Ecology**, v. 22, p. 827-841, 2013. DOI: 10.1111/mec.12007.
- HAMILTON, J. A.; MILLER, J. M. Adaptive introgression as a resource for management and genetic conservation in a changing climate. **Conservation Biology**, v. 30, p. 33-41, 2016. DOI: 10.1111/cobi.12574.
- HARRISON, R. G.; LARSON, E. L. Hybridization, introgression, and the nature of species boundaries. **Journal of Heredity**, v. 105, p. 795-809, 2014. DOI: 10.1093/jhered/esu033.
- HEBERT, P. D. N.; PENTON, E. H.; BURNS, J. M. et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, p. 14812-14817, 2004. DOI: 10.1073/pnas.0406166101.
- HIBBINS, M. S.; HAHN, M. W. Phylogenomic approaches to detecting and characterizing introgression. **Genetics**, v. 220, 2021. DOI: 10.1093/genetics/iyab173.

- HIRASE, S.; YAMASAKI, Y. Y. Genomic evidence for speciation with gene flow in broadcast spawning marine invertebrates. **Molecular Biology and Evolution**, v. 38, p. 4683-4699, 2021. DOI: 10.1093/molbev/msab194.
- HODEL, R. G. J.; MASSATTI, R.; KNOWLES, L. L. Hybrid enrichment of adaptive variation revealed by genotype–environment associations in montane sedges. **Molecular Ecology**, v. 31, p. 3722-3737, 2022. DOI: 10.1111/mec.16502.
- HOELZEL, A. R.; BRUFORD, M. W.; FLEISCHER, R. C. Conservation of adaptive potential and functional diversity. **Conservation Genetics**, v. 20, p. 1-5, 2019. DOI: 10.1007/s10592-019-01151-x.
- HOHENLOHE, P. A.; FUNK, W. C.; RAJORA, O. P. Population genomics for wildlife conservation and management. **Molecular Ecology**, v. 30, p. 62-82, 2021. DOI: 10.1111/mec.15720.
- HULL, D. L. A matter of individuality. **Philosophy of Science**, v. 45, p. 335-360, 1978.
- HUNTLEY, J. W.; KEITH, K. D.; CASTELLANOS, A. A. et al. Underestimated and cryptic diversification patterns across Afro-tropical lowland forests. **Journal of Biogeography**, v. 46, p. 381-391, 2019. DOI: 10.1111/jbi.13505.
- JACKSON, N. D.; CARSTENS, B. C.; MORALES, A. E. et al. Species delimitation with gene flow. *n*, v. 66, p. 799-812, 2017. DOI: 10.1093/sysbio/syw117.
- JACKSON, N. D.; MORALES, A. E.; CARSTENS, B. C. et al. PHRAPL: Phylogeographic Inference Using Approximate Likelihoods. **Systematic Biology**, v. 66, p. 1045-1053, 2017. DOI: 10.1093/sysbio/syx001.
- JAHNER, J. P.; MATOCQ, M. D.; MALANEY, J. L. et al. The genetic legacy of 50 years of desert bighorn sheep translocations. **Evolutionary Applications**, v. 12, p. 198-213, 2019. DOI: 10.1111/eva.12708.
- JIRINEC, V.; BURNER, R. C.; AMARAL, B. R. et al. Morphological consequences of climate change for resident birds in intact Amazonian rainforest. **Science Advances**, v. 7, eabk1743, 2021. DOI: 10.1126/sciadv.abk1743.
- JONES, G. Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. **Journal of Mathematical Biology**, v. 74, p. 447-467, 2017. DOI: 10.1007/s00285-016-1034-0.

- JONES, M. R.; MILLS, L. S.; ALVES, P. C. et al. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. **Science**, v. 360, p. 1355-1358, 2018. DOI: 10.1126/science.aar5273.
- JONES, M. R.; MILLS, L. S.; JENSEN, J. D. et al. The origin and spread of locally adaptive seasonal camouflage in snowshoe hares. **The American Naturalist**, v. 196, p. 316-332, 2020. DOI: 10.1086/710022.
- KARDOS, M.; ARMSTRONG, E. E.; FITZPATRICK, S. W. et al. The crucial role of genome-wide genetic variation in conservation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 118, e2104642118, 2021. DOI: 10.1073/pnas.2104642118.
- KARELL, P.; AHOLA, K.; KARSTINEN, T. et al. Climate change drives microevolution in a wild bird. **Nature Communications**, v. 2, p. 208, 2011. DOI: 10.1038/ncomms1213.
- KEARNS, A. M.; RESTANI, M.; SZABO, I. et al. Genomic evidence of speciation reversal in ravens. **Nature Communications**, v. 9, p. 906, 2018. DOI: 10.1038/s41467-018-03294-w.
- KIM, B. Y.; HUBER, C. D.; LOHMUELLER, K. E. Deleterious variation shapes the genomic landscape of introgression. **PLoS Genetics**, v. 14, e1007741, 2018. DOI: 10.1371/jgen.1007741.
- KIMURA, M. The neutral theory of molecular evolution: A review of recent evidence. **Japanese Journal of Genetics**, v. 66, p. 367-386, 1991.
- KIRKPATRICK, M.; BARTON, N. Chromosome inversions, local adaptation and speciation. **Genetics**, v. 173, p. 419-434, 2006. DOI: 10.1534/genetics.105.047985.
- KLOPFSTEIN, S.; CURRAT, M.; EXCOFFIER, L. The fate of mutations surfing on the wave of a range expansion. **Molecular Biology and Evolution**, v. 23, p. 482-490, 2006. DOI: 10.1093/molbev/msj057.
- KNOWLES, L. L. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. **Molecular Ecology**, v. 10, p. 691, 2001.
- KOZAK, K. M.; JORON, M.; MCMILLAN, W. O. et al. Rampant genome-wide admixture across the *Heliconius* radiation. **Genome Biology and Evolution**, v. 13, 2021. DOI: 10.1093/gbe/evab099.

- KREMER, A.; RONCE, O.; ROBLEDO-ARNUNCIO, J. J. et al. Long-distance gene flow and adaptation of forest trees to rapid climate change. **Ecology Letters**, v. 15, p. 378-392, 2012. DOI: 10.1111/j.1461-0248.2012.01746.x.
- KRONFORST, M. R.; HANSEN, M. E.; CRAWFORD, N. G. et al. Hybridization reveals the evolving genomic architecture of speciation. **Cell Reports**, v. 5, p. 666-677, 2013. DOI: 10.1016/j.celrep.2013.09.042.
- KRONFORST, M. R.; PAPA, R. The functional basis of wing patterning in *Heliconius* butterflies: the molecules behind mimicry. **Genetics**, v. 200, 2015. DOI: 10.1534/genetics.114.172387.
- KYRIAZIS, C. C.; WAYNE, R. K.; LOHMUELLER, K. E. Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. **Evolution Letters**, v. 5, p. 33-47, 2021. DOI: 10.1002/evl3.209.
- LAIKRE, L. Genetic diversity is overlooked in international conservation policy implementation. **Conservation Genetics**, v. 11, p. 349-354, 2010. Disponível em: <http://www.springerlink.com/index/10.1007/s10592-009-0037-4>.
- LAMICHHANEY, S.; BERGLUND, J.; ALMÉN, M. S. et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. **Nature**, v. 518, p. 371, 2015. DOI: 10.1038/nature14181.
- LANDE, R.; SHANNON, S. The role of genetic variation in adaptation and population persistence in a changing environment. **Evolution**, v. 50, p. 434-437, 1996. DOI: 10.1111/j.1558-5646.1996.tb04504.x.
- LARGET, B.; KOTHA, S.; DEWEY, C. et al. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. **Bioinformatics**, v. 26, p. 2910-2911, 2010. DOI: 10.1093/bioinformatics/btq539.
- LEITE, R. N.; ROGERS, D. Revisiting Amazonian phylogeography: insights into diversification hypotheses and novel perspectives. In: **Organisms Diversity & Evolution**. Springer-Verlag, 2013. p. 1-26.
- LEVIN, D. A. Hybridization and Extinction. **American Scientist**, v. 90, p. 254, 2002. DOI: 10.1511/2002.3.254.
- MABLE, B. K. Conservation of adaptive potential and functional diversity: integrating old and new approaches. **Conservation Genetics**, v. 20, p. 89-100, 2019. DOI: 10.1007/s10592-018-1129-9.

- MALLET, J. The speciation revolution. **Journal of Evolutionary Biology**, v. 14, p. 887-888, 2001. Disponível em: <http://www.blackwell-synergy.com/doi/abs/10.1046/j.1420-9101.2001.00342>.
- MALLET, J. Hybrid speciation. **Nature**, v. 446, p. 279, 2007. DOI: 10.1038/nature05706.
- MALLET, J. Mayr's view of Darwin: was Darwin wrong about speciation? **Biological Journal of the Linnean Society**, v. 95, p. 3-16, 2008. DOI: 10.1111/j.1095-8312.2008.01089.x.
- MALLET, J.; BELTRAN, M.; NEUKIRCHEN, W. et al. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. **BMC Evolutionary Biology**, v. 7, 2007. DOI: 10.1186/1471-2148-7-28.
- MAO, Y.; ECONOMO, E. P.; SATOH, N. The Roles of Introgression and Climate Change in the Rise to Dominance of Acropora Corals. **Current Biology**, v. 28, p. 3373-3382.e3375, 2018. DOI: 10.1016/j.cub.2018.08.061.
- MARQUES, D. A.; MEIER, J. I.; SEEHAUSEN, O. A combinatorial view on speciation and adaptive radiation. **Trends in Ecology & Evolution**, v. 34, p. 531-544, 2019. DOI: 10.1016/j.tree.2019.02.008.
- MARTIN, S. H.; DASMAHAPATRA, K. K.; NADEAU, N. J. et al. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. **Genome Research**, v. 23, p. 1817-1828, 2013. DOI: 10.1101/gr.159426.113.
- MATUTE, D. R.; BUTLER, I. A.; TURISSINI, D. A. et al. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. **Science**, v. 329, p. 1518-1521, 2010. DOI: 10.1126/science.1193440.
- MAYR, E. **Animal species and evolution**. Cambridge, Mass: Harvard University Press, 1963.
- MEIER, J. I.; MARQUES, D. A.; MWAIKO, S. et al. Ancient hybridization fuels rapid cichlid fish adaptive radiations. **Nature Communications**, v. 8, p. 14363, 2017. DOI: 10.1038/ncomms14363.
- MEIER, J. I.; SALAZAR, P. A.; KUČKA, M. et al. Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. **Proceedings of the National Academy of Sciences of the United States of America**, v. 118, p. e2015005118, 2021. DOI: 10.1073/pnas.2015005118.

- MEIER, J. I.; SOUSA, V. C.; MARQUES, D. A. et al. Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. **Molecular Ecology**, v. 26, p. 123-141, 2017. DOI: 10.1111/mec.13838.
- MENGUAL, X.; KERR, P.; NORRBOM, A. L. et al. Phylogenetic relationships of the tribe Toxotrypanini (Diptera: Tephritidae) based on molecular characters. **Molecular Phylogenetics and Evolution**, v. 113, p. 84-112, 2017. DOI: 10.1016/j.ympev.2017.05.011.
- MILLER, J. M.; QUINZIN, M. C.; POULAKAKIS, N. et al. Identification of genetically important individuals of the rediscovered Floreana Galápagos giant tortoise (*Chelonoidis elephantopus*) provides founders for species restoration program. **Scientific Reports**, v. 7, p. 11471, 2017. DOI: 10.1038/s41598-017-11516-2.
- MITCHELL, N.; LUU, H.; OWENS, G. L. et al. Hybrid evolution repeats itself across environmental contexts in Texas sunflowers (*Helianthus*). **Evolution**, n/a, 2022. DOI: 10.1111/evo.14536.
- MITTERMEIER, R. A. Primate conservation priorities in the Neotropical region. In: BENIRSCHKE, K. (ed.). **Primates**. New York: Springer, 1986. p. 221-240.
- MOEST, M.; VAN BELLEGHEM, S. M.; JAMES, J. E. et al. Selective sweeps on novel and introgressed variation shape mimicry loci in a butterfly adaptive radiation. **PLoS Biology**, v. 18, 2020. DOI: 10.1371/journal.pbio.3000597.
- MULLER, H. J. Isolating mechanisms, evolution, and temperature. **Biology Symposium**, v. 6, p. 71-125, 1942.
- NAKA, L. N.; BRUMFIELD, R. T. The dual role of Amazonian rivers in the generation and maintenance of avian diversity. **Science Advances**, v. 4, eaar8575, 2018. DOI: 10.1126/sciadv.aar8575.
- NICE, C. C.; GOMPERT, Z.; FORDYCE, J. A. et al. Hybrid speciation and independent evolution in lineages of alpine butterflies. **Evolution**, v. 67, p. 1055-1068, 2013. DOI: 10.1111/evo.12019.
- NIKOLAKIS, Z. L.; SCHIELD, D. R.; WESTFALL, A. K. et al. Evidence that genomic incompatibilities and other multilocus processes impact hybrid fitness in a rattlesnake hybrid zone. **Evolution**, v. 76, p. 2513-2530, 2022. DOI: 10.1111/evo.14612.

- NORRBOM, A. L.; KORYTKOWSKI, C. A.; ZUCCHI, R. A. et al. **Anastrepha and Toxotrypana: descriptions, illustrations, and interactive keys**. Available from: <http://delta-intkey.com/>. Accessed: 24 Feb. 2018.
- NOSIL, P.; FEDER, J. L.; GOMPTER, Z. How many genetic changes create new species? **Science**, v. 371, p. 777-779, 2021. DOI: 10.1126/science.abf6671.
- NOSIL, P.; FUNK, D. J.; ORTIZ-BARRIENTOS, D. Divergent selection and heterogeneous genomic divergence. **Molecular Ecology**, v. 18, p. 375-402, 2009. DOI: 10.1111/j.1365-294X.2008.03946.x.
- NOSS, R. F. Indicators for monitoring biodiversity: A hierarchical approach. **Conservation Biology**, v. 4, p. 355-364, 1990. DOI: 10.1111/j.1523-1739.1990.tb00309.x.
- ORR, H. A.; TURELLI, M. The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. **Evolution**, v. 55, p. 1085-1094, 2001.
- ORR, H. A.; UNCKLESS, R. L. The population genetics of evolutionary rescue. **PLoS Genetics**, v. 10, e1004551, 2014. DOI: 10.1371/journal.pgen.1004551.
- OZILOR, E. M.; REID, N. M.; YAIR, S. et al. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. **Science**, v. 364, p. 455-457, 2019. DOI: 10.1126/science.aav4155.
- PACIONI, C.; RAFFERTY, C.; MORLEY, K. et al. Augmenting the conservation value of rehabilitated wildlife by integrating genetics and population modeling in the post-rehabilitation decision process. **Current Zoology**, v. 64, p. 593-601, 2018. DOI: 10.1093/cz/zox065.
- PAMILO, P.; NEI, M. Relationships between gene trees and species trees. **Molecular Biology and Evolution**, v. 5, p. 568-583, 1988. DOI: 10.1093/oxfordjournals.molbev.a040517.
- PEREIRA, R. J.; BARRETO, F. S.; BURTON, R. S. Ecological novelty by hybridization: experimental evidence for increased thermal tolerance by transgressive segregation in *Tigriopus californicus*. **Evolution**, v. 68, p. 204-215, 2014. DOI: 10.1111/evo.12254.
- PFENNIG, K. S. Biased Hybridization and Its Impact on Adaptive Introgression. **Trends in Ecology and Evolution**, v. 36, p. 488-497, 2021. DOI: 10.1016/j.tree.2021.02.010.
- PFENNIG, K. S.; KELLY, A. L.; PIERCE, A. A. Hybridization as a facilitator of species range expansion. **Proceedings of the Royal Society B: Biological Sciences**, v. 283, p. 20161329, 2016. DOI: 10.1098/rspb.2016.1329.

- PHIFER-RIXEY, M.; HECKMAN, M.; TRUSSELL, G. C. et al. Maintenance of clinal variation for shell colour phenotype in the flat periwinkle *Littorina obtusata*. **Journal of Evolutionary Biology**, v. 21, p. 966-978, 2008. DOI: 10.1111/j.1420-9101.2008.01549.x.
- POWELL, T. H. Q.; FORBES, A. A.; HOOD, G. R. et al. Ecological adaptation and reproductive isolation in sympatry: genetic and phenotypic evidence for native host races of *Rhagoletis pomonella*. **Molecular Ecology**, v. 23, p. 688-704, 2014. DOI: 10.1111/mec.12635.
- PRESGRAVES, D. C. Speciation Genetics: Search for the Missing Snowball. **Current Biology**, v. 20, p. R1073-R1074, 2010. Disponível em: [ScienceDirect](https://www.sciencedirect.com).
- PULIDO-SANTACRUZ, P.; ALEIXO, A.; WEIR, J. T. Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. **Proceedings of the Royal Society B: Biological Sciences**, v. 285, p. 20172081, 2018. DOI: 10.1098/rspb.2017.2081.
- PULIDO-SANTACRUZ, P.; ALEIXO, A.; WEIR, J. T. Genomic data reveal a protracted window of introgression during the diversification of a neotropical woodcreeper radiation. *n*, v. 74, p. 842-858, 2020. DOI: 10.1111/evo.13902.
- PUTMAN, A. I.; CARBONE, I. Challenges in analysis and interpretation of microsatellite data for population genetic studies. **Ecology and Evolution**, v. 4, p. 4399-4428, 2014. DOI: 10.1002/ece3.1305.
- QUILODRÁN, C. S.; MONTOYA-BURGOS, J. I.; CURRAT, M. Harmonizing hybridization dissonance in conservation. **Communications Biology**, v. 3, p. 391, 2020. DOI: 10.1038/s42003-020-1116-9.
- RAMIREZ, J. L.; LESCROART, J.; FIGUEIRÓ, H. V. et al. Genomic Signatures of Divergent Ecological Strategies in a Recent Radiation of Neotropical Wild Cats. **Molecular Biology and Evolution**, v. 39, msac117, 2022. DOI: 10.1093/molbev/msac117.
- REZENDE, V. B.; CONGRAINS, C.; LIMA, A. L. A. et al. Head transcriptomes of two closely related species of fruit flies of the *Anastrepha fraterculus* group reveals divergent genes in species with extensive gene flow. **G3**, v. 6, p. 3283-3295, 2016. DOI: 10.1534/g3.116.030486.
- RHYMER, J. M.; SIMBERLOFF, D. Extinction by hybridization and introgression. **Annual Review of Ecology and Systematics**, v. 27, p. 83-109, 1996.
- RIESEBERG, L. H.; BURKE, J. M. A genic view of species integration. **Journal of Evolutionary Biology**, v. 14, p. 883-886, 2001. DOI: 10.1046/j.1420-9101.2001.00339.x.

- RIESEBERG, L. H.; RAYMOND, O.; ROSENTHAL, D. M. et al. Major ecological transitions in wild sunflowers facilitated by hybridization. **Science**, v. 301, p. 1211-1216, 2003. DOI: 10.1126/science.1086949.
- RIESEBERG, L. H.; WHITTON, J.; GARDNER, K. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. **Genetics**, v. 152, p. 713-727, 1999. DOI: 10.1093/genetics/152.2.713.
- ROBINSON, J. A.; RÄIKKÖNEN, J.; VUCETICH, L. M. et al. Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. **Science Advances**, v. 5, p. 635-639, 2019. DOI: 10.1126/sciadv.aau0757.
- ROESTI, M.; HENDRY, A. P.; SALZBURGER, W. et al. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. **Molecular Ecology**, v. 21, p. 2852-2862, 2012. DOI: 10.1111/j.1365-294X.2012.05509.x.
- ROSENTHAL, G. G. Individual mating decisions and hybridization. **Journal of Evolutionary Biology**, v. 26, p. 252-255, 2013. DOI: 10.1111/jeb.12004.
- RYDING, S.; KLAAUSSEN, M.; TATTERSALL, G. J. et al. Shape-shifting: changing animal morphologies as a response to climatic warming. **Trends in Ecology & Evolution**, v. 36, p. 1036-1048, 2021. DOI: 10.1016/j.tree.2021.07.006.
- SCHILD, D. R.; PERRY, B. W.; ADAMS, R. H. et al. Allopatric divergence and secondary contact with gene flow: a recurring theme in rattlesnake speciation. **Biological Journal of the Linnean Society**, v. 128, p. 149-169, 2019. DOI: 10.1093/biolinnean/blz077.
- SCHUMER, M.; XU, C. L.; POWELL, D. L. et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. **Science**, v. 360, p. 656-659, 2018. DOI: 10.1126/science.aar3684.
- SEEHAUSEN, O. L. E.; TAKIMOTO, G.; ROY, D. et al. Speciation reversal and biodiversity dynamics with hybridization in changing environments. **Molecular Ecology**, v. 17, p. 30-44, 2008. DOI: 10.1111/j.1365-294X.2007.03529.x.
- SEQUEIRA, F.; SODRÉ, D.; FERRAND, N. et al. Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. **BMC Evolutionary Biology**, v. 11, p. 264, 2011. DOI: 10.1186/1471-2148-11-264.

- SHAPIRO, B. J.; LEDUCQ, J.-B.; MALLET, J. What is speciation? **PLoS Genetics**, v. 12, e1005860, 2016. DOI: 10.1371/journal.pgen.1005860.
- SHARP, D. The Reality of Species. **Nature**, v. 3, p. 426, 1871. DOI: 10.1038/003426g0.
- SHAW, R. G.; ETTERSON, J. R. Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. **New Phytologist**, v. 195, p. 752-765, 2012. DOI: 10.1111/j.1469-8137.2012.04230.x.
- SIMPSON, G. G. The Species Concept. **Evolution**, v. 5, p. 285-298, 1951. DOI: 10.2307/2405675.
- SIRÉN, J.; MARTTINEN, P.; CORANDER, J. Reconstructing population histories from single nucleotide polymorphism data. **Molecular Biology and Evolution**, v. 28, p. 673-683, 2011. DOI: 10.1093/molbev/msq236.
- SMITH, S. A.; WALKER-HALE, N.; WALKER, J. F. et al. Phylogenetic conflicts, combinability, and deep phylogenomics in plants. **Systematic Biology**, v. 69, p. 579-592, 2020. DOI: 10.1093/sysbio/syz078.
- SOLÍS-LEMUS, C.; BASTIDE, P.; ANÉ, C. PhyloNetworks: A package for phylogenetic networks. **Molecular Biology and Evolution**, v. 34, p. 3292-3298, 2017. DOI: 10.1093/molbev/msx235.
- SOLÍS-LEMUS, C.; YANG, M.; ANÉ, C. Inconsistency of species tree methods under gene flow. **Systematic Biology**, 2016. Disponível em: <http://sysbio.oxfordjournals.org/content/early/2016/06/02/sysbio.syw030.abstract>. Acesso em: 12 dez. 2024.
- SPIELMAN, D.; BROOK, B. W.; FRANKHAM, R. Most species are not driven to extinction before genetic factors impact them. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, p. 15261-15264, 2004. DOI: 10.1073/pnas.0403809101.
- STRUCK, T. H.; FEDER, J. L.; BENDIKSBY, M. et al. Finding evolutionary processes hidden in cryptic species. **Trends in Ecology & Evolution**, v. 33, p. 153-163, 2018. DOI: 10.1016/j.tree.2017.11.007.
- SUAREZ-GONZALEZ, A.; LEXER, C.; CRONK, Q. C. B. Adaptive introgression: a plant perspective. **Biology Letters**, v. 14, p. 20170688, 2018. DOI: 10.1098/rsbl.2017.0688.

- SUKUMARAN, J.; KNOWLES, L. L. Multispecies coalescent delimits structure, not species. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, p. 1607-1612, 2017. DOI: 10.1073/pnas.1607921114.
- SUVOROV, A.; KIM, B. Y.; WANG, J. et al. Widespread introgression across a phylogeny of 155 *Drosophila* genomes. **Current Biology**, v. 32, p. 111-123.e115, 2022. DOI: 10.1016/j.cub.2021.10.052.
- TAYLOR, R. S.; MANSEAU, M. The role of introgression and ecotypic parallelism in delineating intraspecific conservation units. **Molecular Ecology**, v. 29, p. 2793-2809, 2020. DOI: 10.1111/mec.15522.
- TAYLOR, S. A.; LARSON, E. L. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. **Nature Ecology & Evolution**, v. 3, p. 170-177, 2019. DOI: 10.1038/s41559-018-0777-y.
- TAYLOR, S. A.; LARSON, E. L.; HARRISON, R. G. Hybrid zones: windows on climate change. **Trends in Ecology & Evolution**, v. 30, p. 398-406, 2015. DOI: 10.1016/j.tree.2015.04.010.
- TEIXEIRA, J. C.; HUBER, C. D. The inflated significance of neutral genetic diversity in conservation genetics. **Proceedings of the National Academy of Sciences of the United States of America**, v. 118, e2015096118, 2021. DOI: 10.1073/pnas.2015096118.
- TEMPLETON, A. R. **Population genetics and microevolutionary theory**. Hoboken: John Wiley & Sons, Inc., 2006.
- TEMPLETON, A. R. **Population Genetics and Microevolutionary Theory**. Hoboken, New Jersey: John Wiley & Sons, Inc., 2021.
- THAN, C.; RUTHS, D.; NAKHLEH, L. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. **BMC Bioinformatics**, v. 9, p. 322, 2008. DOI: 10.1186/1471-2105-9-322.
- THAWORNWATTANA, Y.; SEIXAS, F. A.; YANG, Z. et al. Full-Likelihood Genomic Analysis Clarifies a Complex History of Species Divergence and Introgression: The Example of the erato-sara Group of *Heliconius* Butterflies. **Systematic Biology**, 2022. DOI: 10.1093/sysbio/syac009.
- THOMAS, C. D.; CAMERON, A.; GREEN, R. E. et al. Extinction risk from climate change. **Nature**, v. 427, p. 145-148, 2004. DOI: 10.1038/nature02121.

- TIGANO, A.; FRIESEN, V. L. Genomics of local adaptation with gene flow. **Molecular Ecology**, v. 25, p. 2144-2164, 2016. DOI: 10.1111/mec.13606.
- TODESCO, M.; OWENS, G. L.; BERCOVICH, N. et al. Massive haplotypes underlie ecotypic differentiation in sunflowers. **Nature**, v. 584, p. 602-607, 2020. DOI: 10.1038/s41586-020-2467-6.
- VALENCIA-MONTOYA, W. A.; ELFEKIH, S.; NORTH, H. L. et al. Adaptive introgression across semipermeable species boundaries between local *Helicoverpa zea* and invasive *Helicoverpa armigera* moths. **Molecular Biology and Evolution**, v. 37, p. 2568-2583, 2020. DOI: 10.1093/molbev/msaa108.
- VALLEJO-MARÍN, M.; HISCOCK, S. J. Hybridization and hybrid speciation under global change. **New Phytologist**, v. 211, p. 1170-1187, 2016. DOI: 10.1111/nph.14004.
- VALLINOTO, M.; CUNHA, D. B.; BESSA-SILVA, A. et al. Deep divergence and hybridization among sympatric Neotropical toads. **Zoological Journal of the Linnean Society**, v. 180, p. 647-660, 2017. DOI: 10.1093/zoolinnean/zw001.
- VANÍČKOVÁ, L.; HERNÁNDEZ-ORTÍZ, V.; BRAVO, I. S. J. et al. Current knowledge of the species complex *Anastrepha fraterculus* (Diptera, Tephritidae) in Brazil. **ZooKeys**, p. 211-237, 2015. DOI: 10.3897/zookeys.540.9791.
- VANZOLINI, P. E.; WILLIAMS, E. E. The vanishing refuge: A mechanism for ecogeographic speciation. **Papéis Avulsos de Zoologia (São Paulo)**, v. 34, p. 251-255, 1981.
- VASCONCELLOS, M. M.; COLLI, G. R. Factors affecting the population dynamics of two toads (Anura: Bufonidae) in a seasonal Neotropical savanna. **Copeia**, 2009, p. 266-276, 211. DOI: 10.1643/CE-07-099.
- VEGA, Y.; MARQUES, I.; CASTRO, S. et al. Outcomes of extensive hybridization and introgression in *Epidendrum* (Orchidaceae): Can we rely on species boundaries? **PLOS ONE**, v. 8, e80662, 2013. DOI: 10.1371/journal.pone.0080662.
- VIEITES, D. R.; WOLLENBERG, K. C.; ANDREONE, F. et al. Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, p. 8267-8272, 2009. DOI: 10.1073/pnas.0810821106.

- VONHOLDT, B. M.; BRZESKI, K. E.; AARDEMA, M. L. et al. Persistence and expansion of cryptic endangered red wolf genomic ancestry along the American Gulf coast. **Molecular Ecology**, v. 31, p. 5440-5454, 2022. DOI: 10.1111/mec.16200.
- VONHOLDT, B. M.; BRZESKI, K. E.; WILCOVE, D. S. et al. Redefining the role of admixture and genomics in species conservation. **Conservation Letters**, v. 11, e12371, 2018. DOI: 10.1111/conl.12371.
- VONLANTHEN, P.; BITTNER, D.; HUDSON, A. G. et al. Eutrophication causes speciation reversal in whitefish adaptive radiations. **Nature**, v. 482, p. 357-362, 2012. DOI: 10.1038/nature10824.
- WANG, T.; O'NEILL, G. A.; AITKEN, S. N. Integrating environmental and genetic effects to predict responses of tree populations to climate. **Ecological Applications**, v. 20, p. 153-163, 2010. DOI: 10.1890/08-2257.1.
- WAPLES, R. S.; LINDLEY, S. T. Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids. **Evolutionary Applications**, v. 11, p. 1518-1526, 2018. DOI: 10.1111/eva.12687.
- WELCH, M. E.; RIESEBERG, L. H. Habitat divergence between a homoploid hybrid sunflower species, *Helianthus paradoxus* (Asteraceae), and its progenitors. **American Journal of Botany**, v. 89, p. 472-478, 2002. DOI: 10.3732/ajb.89.3.472.
- WEN, D. Q.; NAKHLEH, L. Coestimating reticulate phylogenies and gene trees from multilocus sequence data. **Systematic Biology**, v. 67, p. 439-457, 2018. DOI: 10.1093/sysbio/syx085.
- WILCZEK, A. M.; COOPER, M. D.; KORVES, T. M. et al. Lagging adaptation to warming climate in *Arabidopsis thaliana*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, p. 7906-7913, 2014. DOI: 10.1073/pnas.1406314111.
- WILLI, Y.; KRISTENSEN, T. N.; SGRÒ, C. M. et al. Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. **Proceedings of the National Academy of Sciences of the United States of America**, v. 119, e2105076119, 2022. DOI: 10.1073/pnas.2105076119.
- WILLIAMS, R. C.; BLANCO, M. B.; POELSTRA, J. W. et al. Conservation genomic analysis reveals ancient introgression and declining levels of genetic diversity in Madagascar's hibernating dwarf lemurs. **Heredity**, v. 124, p. 236-251, 2020. DOI: 10.1038/s41437-019-0260-9.

- WOLF, J. B. W.; ELLEGREN, H. Making sense of genomic islands of differentiation in light of speciation. **Nature Reviews Genetics**, v. 18, p. 87-100, 2017. Review.
- ZECHERLE, L. J.; NICHOLS, H. J.; BAR-DAVID, S. et al. Subspecies hybridization as a potential conservation tool in species reintroductions. **Evolutionary Applications**, v. 14, p. 1216-1224, 2021. DOI: 10.1111/eva.13191.
- ZHANG, W.; DASMAHAPATRA, K. K.; MALLET, J. et al. Genome-wide introgression among distantly related *Heliconius* butterfly species. **Genome Biology**, v. 17, p. 25, 2016. DOI: 10.1186/s13059-016-0889-0.
- ZUCCHI, R. A. Espécies de *Anastrepha*, sinonímias, plantas hospedeiras e parasitóides. In: MALAVASI, A.; ZUCCHI, R. A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil: Conhecimento básico e aplicado**. Ribeirão Preto: Holos Editora, 2000. p. 41-48.

2° Capítulo

**Comparative subsampling of phylogenomic informative loci to delimit species
in *Anastrepha***

2º Capítulo

Comparative subsampling of phylogenomic informative loci to delimit species in *Anastrepha*

Resumo

As análises filogenômicas revolucionaram nossa compreensão das relações evolutivas, mas a incongruência entre os loci genômicos frequentemente complica a reconstrução de filogenias precisas. Aqui, reanalisamos dados genômicos para investigar as relações evolutivas entre moscas-frutas *Anastrepha*, incluindo sete grupos de espécies e múltiplas linhagens do complexo *Anastrepha fraterculus*. Nosso conjunto de dados abrangeu 3.170 grupos de genes ortólogos para 15 espécies de *Anastrepha* e cinco outros gêneros de tefritídeos. A inferência filogenética usando árvore de espécies e análises de concordância revelou consistentemente topologias bem suportadas para relações evolutivas mais profundas, enquanto divergências mais recentes, particularmente dentro do complexo *A. fraterculus*, exibiram altos níveis de incongruência filogenética. Fatores de concordância genética (gCF) e fatores de concordância de sítio (sCF) indicaram que apenas uma minoria de loci (<50%) suportou nos-chave, sugerindo a influência do fluxo gênico, classificação de linhagem incompleta ou outras forças evolutivas. Para abordar esses desafios, exploramos estratégias para selecionar subconjuntos reduzidos de loci filogeneticamente informativos. Entre as estratégias testadas, subconjuntos baseados em métricas de concordância e taxa evolutiva produziram topologias consistentes com análises de conjuntos de dados completos e níveis reduzidos de discordância. Notavelmente, esses subconjuntos mantiveram suporte robusto para relacionamentos mais profundos enquanto aumentavam a congruência em nós mais rasos. Investigações posteriores de parâmetros genômicos revelaram que genes nos conjuntos reduzidos exibiram taxas evolutivas mais baixas, valores mais altos de certeza de internó (IC) e, em alguns casos, sinais mais fortes de introgressão. Essas descobertas destacam o potencial de loci cuidadosamente selecionados para melhorar a resolução filogenética e mitigar os efeitos de sinais conflitantes. Nosso estudo fornece novos insights sobre a história evolutiva de *Anastrepha*, esclarece relacionamentos dentro do complexo *A. fraterculus* e oferece abordagens práticas para refinar análises filogenômicas em sistemas com histórias evolutivas complexas.

Palavras-chave: filogenômica, subamostragem, complexo *Anastrepha fraterculus*, praga de inseto, discriminação de espécies

Abstract

Phylogenomic analyses have revolutionized our understanding of evolutionary relationships, yet the incongruence among genomic loci often complicates the reconstruction of accurate phylogenies. Here, we reanalyzed genomic data to investigate the evolutionary relationships among *Anastrepha* fruit flies, including seven species groups and multiple lineages of the *Anastrepha fraterculus* complex. Our dataset encompassed 3,170 orthologous gene clusters for 15 *Anastrepha* species and five other tephritid genera. Phylogenetic inference using species tree and concordance analyses consistently revealed well-supported topologies for deeper evolutionary relationships, while more recent divergences, particularly within the *A. fraterculus* complex, exhibited high levels of phylogenetic incongruence. Gene concordance factors (gCF) and site concordance factors (sCF) indicated that only a minority of loci (<50%) supported key nodes, suggesting the influence of gene flow, incomplete lineage sorting, or other evolutionary forces. To address these challenges, we

explored strategies for selecting reduced subsets of phylogenetically informative loci. Among the strategies tested, subsets based on concordance and evolutionary rate metrics produced topologies consistent with full dataset analyses and reduced levels of discordance. Notably, these subsets maintained robust support for deeper relationships while increasing congruence at shallower nodes. Further investigation of genomic parameters revealed that genes in the reduced sets exhibited lower evolutionary rates, higher internode certainty (IC) values, and, in some cases, stronger signals of introgression. These findings highlight the potential for carefully selected loci to improve phylogenetic resolution and mitigate the effects of conflicting signals. Our study provides new insights into the evolutionary history of *Anastrepha*, clarifies relationships within the *A. fraterculus* complex, and offers practical approaches for refining phylogenomic analyses in systems with complex evolutionary histories.

Keywords: phylogenomics, subsampling, *Anastrepha fraterculus* complex, insect pest, species discrimination

1 INTRODUCTION

The genus *Anastrepha* encompasses over 300 identified species, several of which are pests of great economic importance in the Neotropical region (Norrbon *et al.* 2018). As with other Tephritidae, species identification in this genus is hampered by a general reliance on female morphology for identification (Norrbon *et al.* 2000). Others strategies have been tried, with varied results, particularly in more closely related species (Rodriguez *et al.* 2022; Dutra *et al.* 2011), but more emphasis has been given to the use of genetic data to investigate their phylogenetic relationships, with the use of different genetic regions, with variable success (Dias *et al.* 2016; Sobrinho and De Brito, 2012). Recent phylogenetic and phylogenomic studies using genomic and transcriptomic data have provided relevant phylogenetic relationships across different taxonomic levels, such as among species groups, but even among more closely related taxa, such as species in the *fraterculus* group, which harbors most of the pestiferous species in the genus (Congrains *et al.*, 2023; Congrains *et al.*, 2021). These data have been analyzed employing a combination of methods from concatenated supergenes to coalescence and congruence analysis of individual gene trees to explore genetic variability across multiple loci (Congrains *et al.*, 2023). The combination of different methods has revealed that a combination of evolutionary factors limits our ability to infer phylogenetic relationship among *Anastrepha*, such as introgression, recent and rapid divergence and high levels of ancestral polymorphism (Congrains *et al.*, 2023; Díaz *et al.* 2018).

The differentiation between ancestral polymorphism and introgression is not trivial (Hibbins and Hahn, 2021), but recent phylogenomic studies has indicated that recent divergence and introgression may contribute to complicate species identification and recognition in *Anastrepha* (Congrains *et al.*, 2023; Congrains *et al.*, 2021). As more genomic data has emerged, it has become clear that introgression has a much larger role in differentiation across several different taxa (Suvorov *et al.*, 2022; Thawornwattana *et al.* 2022). This may have happened because of historical and/or recurring events, and may even be involved with species adaptation to different selective pressures, a process that can increase pest adaptability to new environments, but influence proper phylogenetic inferences (Daubin *et al.* 2003). This may be one of the reasons why different genetic markers have produced discordant phylogenetic inferences in *Anastrepha*, particularly among more closely related species in the *fraterculus* group, and has limited our ability to tackle on the differentiation of several lineages in the *A. fraterculus* *s.l.*, which is considered to be a species complex (Selivon *et al.* 2022; Hernández-Ortiz *et al.* 2015). The ability to investigate a

greater number of specimens across their distribution may be essential to help investigating species boundaries and evolutionary forces involved in their differentiation, but the cost of whole genomic analyses may still be prohibitive.

The use of a reduced subset of markers for phylogenetic analyses, or phylogenomic subsampling, has emerged as a valuable approach (Chen *et al.* 2015; Mongiardino Koch, 2021) to offset the costs of producing and analyzing large datasets (Mongiardino Koch, 2021 Mongiardino Koch, 2021), particularly for population studies. The focus on a carefully selected subset of genes may enable a target on phylogenetically informative loci that help resolve contentious or unstable nodes while mitigating confounding factors such as saturation, alignment gaps (Fernández *et al.* 2016; Sharma *et al.* 2014) or even introgression and ancestral polymorphism (Steenwyk *et al.* 2023; Moran *et al.* 2021). Different criteria have been used in subsampling protocols aimed at identifying reliable loci, including information quantity (e.g., alignment length or data completeness) and quality (e.g., phylogenetic signal or resistance to systematic biases). These approaches vary widely, with some focusing on evolutionary rates, favoring loci with fast, slow, or intermediate rates, or using rate-homogeneous partitions to reduce noise (Cummins and McInerney, 2011; Li *et al.* 2021). Others employed the parameter referred to as phylogenetic informativeness (PI) to predict loci likely to resolve relationships accurately (Townsend *et al.* 2012; Alda *et al.* 2019) or Internode Certainty (IC), that indicates loci that properly infer correct phylogenetic relationships (Kobert *et al.* 2016). Different strategies have used these criteria individually, or combined, to achieve adequate subsampling, but the results have been conflicting (Mongiardino Koch, 2021; Fernández *et al.* 2016).

Previously, we have suggested that a set of ~100 loci that were phylogenetically informative which could be used for species identification in *Anastrepha* (Congrains *et al.* 2023). Here we test three different protocols to subsample a large array of genomic and transcriptomic data on different species and samples of *Anastrepha* aiming to identify a minimal set of loci that are both phylogenetically informative and robust against the effects of introgression and ancestral lineage sorting. The identification of a short subset of informative loci is critical for a broader analysis across large populations, but also to help design a cost-effective strategy for species identification in *Anastrepha*, paramount for the development of tools to monitor and control several important pest species that belong to this genus.

2. Materials and Methods

2.1 Data processing

We explored data produced from a pipeline that relies on the selection and alignment of orthologous genes through the integration of transcriptomic and genomic data in a phylogenetic framework. The pipeline we used is a slight modification of what has been described elsewhere (Congrains *et al.* 2023, 2021). Different than in previous analysis, we did not implement a filter based on taxon occupancy to prevent discarding potentially informative genes whose orthology inference has failed due to their high evolutionary rates. The dataset consisted of 3,170 orthologs from 36 *Anastrepha* specimens and five specimens from other genera of the Tephritidae family. The set of samples has a greater emphasis on species of the *fraterculus* group (Congrains *et al.* 2023) (Figure 1). We produced two different datasets with aligned sequences from different specimens to perform the analyses hierarchically, with one focusing on species of the *fraterculus* group (*fraterculus* dataset) and a more encompassing dataset that investigated patterns of evolution across the *Anastrepha* genus (*Anastrepha* dataset). The *fraterculus* dataset was comprised of 3,168 orthologs from 27 specimens of 8 species, that encompassed all samples from the *fraterculus* group, *A. psidivora*, which is considered *incertae sedis* and two outgroups, *A. striata* and *A. bistrigata*. It is important to mention that since different taxa in the *A. fraterculus* complex has not been formally separated into different species with proper names, we are using the term *A. fraterculus* s.l. to refer to an assemblage of lineages that forms a polyphyletic taxa, though we recognize several different lineages that have the support of different attributes as potentially valid separate taxa in the complex. When we consider all these taxa separately, we have a total of 13 taxa in the *fraterculus* dataset. The *Anastrepha* dataset consists of 3,170 orthologs from 36 specimens of 15 species (21 when considering lineages of *fraterculus* complex as separate taxa) of *Anastrepha* and expands the analyses also to other species groups and includes *Zeugodacus* (*Bactrocera*) *cucurbitae*, *Bactrocera dorsalis*, *Bactrocera oleae*, *Ceratitis capitata*, and *Rhagoletis zephyria* as outgroups (Table 1 and Figure 1).

Table 1. Sampling information.

Sample ^a	Species	Species group	City (State/Province)	Country	Coordinates lat/long ^b	Host
<i>A. fraterculus</i> ES	<i>Anastrepha fraterculus fraterculus</i>		Muniz Freire (ES)	Brazil	20°27'52.13"S / 41°24'54.88"W	<i>Plinia cauliflora</i>
<i>A. fraterculus</i> RJ	<i>Anastrepha fraterculus fraterculus</i>		Conceição do Jacareí (RJ)	Brazil	23° 1'54.32"S / 44° 9'54.14"W	Tropical- almond
<i>A. fraterculus</i> RS	<i>Anastrepha fraterculus fraterculus</i>		Dois Irmãos (RS)	Brazil	29°57'7"S / 51°11'33"W	Cattley guava
<i>A. fraterculus</i> RS2	<i>Anastrepha fraterculus fraterculus</i>		Vacaria (RS)	Brazil	28°30'44"S / 50°56'02"W	Trap
<i>A. fraterculus</i> SP1	<i>Anastrepha fraterculus fraterculus</i>		São Carlos (SP)	Brazil	22°1'49.84" S / 20°27'52.13" S	Savanna cherry
<i>A. fraterculus</i> AR	<i>Anastrepha fraterculus fraterculus</i>		Tucumán (TUC)	Argentina	27° 2'3.74"S / 65°19'45.33"W	Trap
<i>A. fraterculus</i> SC	<i>Anastrepha fraterculus fraterculus</i>		Itapema (SC)	Brazil	27°05'36"S / 48°37'08"W	Barbados cherry
<i>A. fraterculus</i> SP2	<i>Anastrepha fraterculus fraterculus</i>		Porto Ferreira (SP)	Brazil	21°50'59"S / 47°29'42"W	Rangpur
<i>A. fraterculus</i> BA	<i>Anastrepha fraterculus fraterculus</i>		Ubaitaba (BA)	Brazil	14°18'37.66"S / 39°19'18.08"W	Guava
<i>A. fraterculus</i> SP3	<i>Anastrepha fraterculus fraterculus</i>		Ilha Bela (SP)	Brazil	23°47'19.52"S / 45°21'42.02"W	Guava
<i>A. fraterculus</i> CUS	<i>Anastrepha fraterculus fraterculus</i>		Cusco	Peru	13°19'18.18" S / 71°57'24.9" W	<i>Prunus persica</i>
<i>A. fraterculus</i> LOR	<i>Anastrepha fraterculus fraterculus</i>		Loreto	Peru	03°15' 5.27" S / 72°54'28.44" W	Trap
<i>A. fraterculus</i> ANC	<i>Anastrepha fraterculus fraterculus</i>		Ancash	Peru	09°29'07.42" S / 78°17'58.52" W	Trap
<i>A. fraterculus</i> MDD	<i>Anastrepha fraterculus fraterculus</i>		Madre de Dios	Peru	12°33'28.33" S / 70°06'31.36" W	<i>Onychopetalum periquino</i>
<i>A. fraterculus</i> Col	<i>Anastrepha fraterculus fraterculus</i>		Rizalalda	Colômbia	05°11'18.85" N / 75°52'28.42" W	Trap
<i>A. fraterculus</i> Ecu	<i>Anastrepha fraterculus fraterculus</i>		Pichincha	Ecuador	00°12'02.70" S / 78°28'19.99" W	Chirimoya
<i>A. fraterculus</i> Mx ^c	<i>Anastrepha fraterculus fraterculus</i>		Ciudad de Mexico	Mexico	-	-

<i>A. turpiniae</i> MG	<i>Anastrepha turpiniae</i>	<i>fraterculus</i>	Três Marias (MG)	Brazil	18°12'13.28"S / 45°14'23.34"W	Guava
<i>A. turpiniae</i> SP	<i>Anastrepha turpiniae</i>	<i>fraterculus</i>	Araraquara (SP)	Brazil	21°48'55.81"S / 48°12'5.34"W	Guava
<i>A. distincta</i> SP	<i>Anastrepha distincta</i>	<i>fraterculus</i>	São Carlos (SP)	Brazil	21°57'33"S / 47°53'54"W	Inga
<i>A. obliqua</i> RJ	<i>Anastrepha obliqua</i>	<i>fraterculus</i>	Conceição do Jacareí (RJ)	Brazil	23° 1'54.32"S / 44° 9'54.14"W	Starfruit
<i>A. obliqua</i> GO	<i>Anastrepha obliqua</i>	<i>fraterculus</i>	Goiania (GO)	Brazil	16°41'58"S / 49°16'35"W	Jacote
<i>A. obliqua</i> PR1	<i>Anastrepha obliqua</i>	<i>fraterculus</i>	Capanema (PR)	Brazil	25°39'45.54"S / 53°48'28.74"W	Eugenia uvalha
<i>A. obliqua</i> PR2	<i>Anastrepha obliqua</i>	<i>fraterculus</i>	Marialva (PR)	Brazil	23°30'56.68"N / 51°49'34.11"W	Jocote
<i>A. obliqua</i> Col	<i>Anastrepha obliqua</i>	<i>fraterculus</i>	-	Colômbia	4°54'8.49"S / 74°41'20.46"W	Trap
<i>A. suspensa</i>	<i>Anastrepha suspensa</i>	<i>fraterculus</i>	-	United States*	41°33'57.11"N / 77°54'8.54"W	Trap
<i>A. ludens</i>	<i>Anastrepha ludens</i>	<i>fraterculus</i>	-	United States**	41° 7'37.30"N / 74°40'57.73"W	Trap
<i>A. psidivora</i>	<i>Anastrepha psidivora</i>	Unplaced	Cusco	Peru	12°53'42" S / 71°24'10" W	Trap
<i>A. bistrigata</i>	<i>Anastrepha bistrigata</i>	<i>striata</i>	São Carlos (SP)	Brazil	22° 1'49.84"S / 47°54'27.90"W	Guava
<i>A. striata</i>	<i>Anastrepha striata</i>	<i>striata</i>	Cusco	Peru	12°53'42.23" S / 71°24'09.98" W	Trap
<i>A. pseudoparallela</i>	<i>Anastrepha pseudoparallela</i>	<i>pseudoparallela</i>	Porto Ferreira (SP)	Brazil	21°50'59"S / 47°29'42"W	Passion fruit
<i>A. grandis</i>	<i>Anastrepha grandis</i>	<i>grandis</i>	Porto Ferreira (SP)	Brazil	21°50'59"S / 47°29'42"W	Pumpkin
<i>A. serpentina</i>	<i>Anastrepha serpentina</i>	<i>serpentina</i>	Araraquara (SP)	Brazil	21°48'55.81"S / 48°12'5.34"W	Pouteria campechiana
<i>A. hadracantha</i>	<i>Anastrepha hadracantha</i>	<i>mucronata</i>	Madre de Dios	Peru	12°33'11.23" S / 70°06'38.27" W	Trap
<i>A. leptozona</i>	<i>Anastrepha leptozona</i>	<i>leptozona</i>	Cusco	Peru	12°53'42" S / 71°24'10" W	Trap

A. curitis *Anastrepha curitis* *leptozona* Madre de Dios Peru 12°34'00.98" W / 70°06'05.72" S Trap

^aCountry abbreviations are: Arg: Argentina; Col: Colombia; Ecu: Ecuador. State/Department/Province, City/Locality Abbreviations correspond to these full names: SP: São Paulo, ES: Espírito Santo, RJ: Rio de Janeiro, RS: Rio Grande do Sul, SC: Santa Catarina, BA: Bahia, GO: Goiás, PR: Paraná, MG: Minas Gerais, TUC: Tucumán, CDMX: Ciudad de México, CUS: Cusco, LOR: Loreto, ANC: Ancash, and MDD: Madre de Dios

^bThese coordinates were obtained using Google Earth Pro v. 7.3.4.

^cPooled sample from GenBank with some information not available

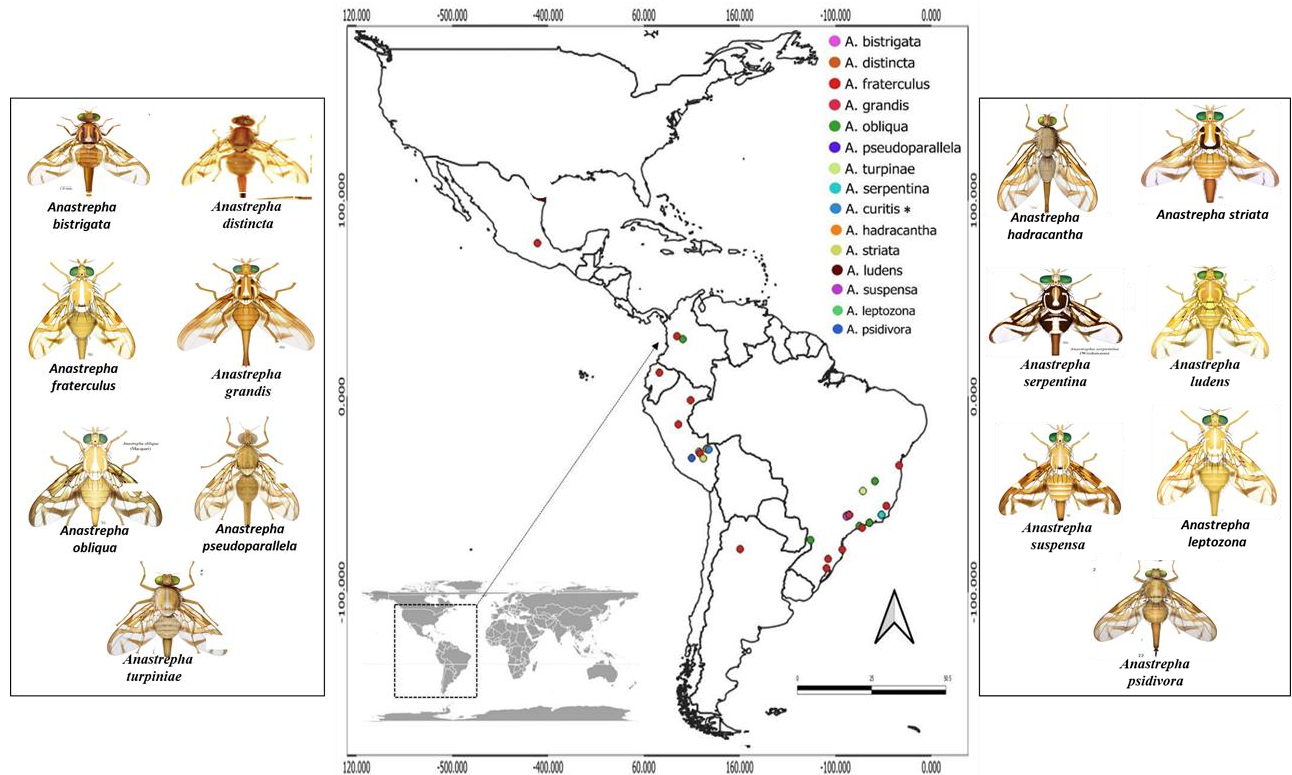


Figure 1: Spatial configuration of the sampling distribution of the 15 species of the genus *Anastrepha*, with samples distributed in North and South America.

2.2 Species Trees Inference and Concordance Analysis

We used Model-Test-NG (Darriba *et al.* 2019) to infer the best-fit nucleotide substitution model based on the Bayesian information criterion (BIC) for each alignment of orthologous genes, independently for each set, *fraterculus* and *Anastrepha*, which was used in RAXML-NG (Kozlov *et al.* 2019) to infer Maximum likelihood (ML) gene trees with 200 rapid bootstrap replicates. Species trees were inferred by concatenating gene alignments into a super-matrix using a concatenation method and also by combining gene trees through multispecies coalescent approaches for both datasets. Concatenation analysis involved incorporating the best-fit model for each gene and employing 200 bootstraps using IQ-TREE v. 2.1.2 (Minh *et al.* 2020). Multi-species

coalescent trees were estimated based on ML gene trees using default parameters in ASTRAL v. 5.7.7 (Zhang *et al.* 2018). Phylogenetic support was evaluated through the gene concordance factor, local posterior probabilities (PP), and quartet support. The gene concordance factor represents the proportion of gene trees with compatible topologies with a branch in the species tree (Minh and Hahn, 2020), calculated using IQ-TREE v. 2.1.2 (Minh *et al.* 2020). Local PP and quartet support were computed based on the quartet frequency in the gene trees (Sayyari and Mirarab, 2016), estimated using ASTRAL v. 5.7.7 (Zhang *et al.* 2018).

We investigated the proportion of genes within our comprehensive dataset that exhibit higher concordance with the species tree generated by the ASTRAL method using Phyparts (Smith *et al.* 2015), a bipartition-based method, for a concordance and conflict analysis with a bootstrap support threshold < 70%. Subsequently, we utilized the python script PhypartsPieCharts.py (<https://github.com/mossmatters/phyloscripts>) to visualize the results with pie charts that show the number of gene trees that were concordant, conflicting, or uninformative with respect to the species tree in each node. We used Gotree (Lemoine and Gascuel, 2021) throughout the workflow to remove zero length branches and to root trees.

2.3 Search for Phylogenetically Informative Genes

Previous studies have been successful at using molecular data to bring some phylogenetic resolution among different species of *Anastrepha*, especially those belonging to the *fraterculus* group, by using a large number of different loci. Here we investigated whether we could find similar resolution using a more reduced set of phylogenetically informative genes, and to that effect we tried three alternative methods. We chose strategies that would define the smallest number of loci that was still effective at identifying a set of taxonomically and topologically well-supported nodes, similar to what we described elsewhere (Congrains *et al.* 2023). The procedures here considered three sets of nodes that are informative at different hierarchical levels, that is, that separates genera, species groups in the genus *Anastrepha*, species, and lineages of the *A. fraterculus* complex (see results for details). We should mention that whenever a species group is represented by a single species, we allocated this node to the species group. On the other hand, lineages in the *fraterculus* species were considered as “species” when there was association to other traits that would support a stronger claim. This is akin to what we used elsewhere (Congrains *et al.* 2023), though we expanded the nodes previously considered to account for more divergent taxa.

The initial strategies used to select phylogenetically informative sets, which is exploratory in nature, relied on data derived from the IQ-TREE v. 2.1.2 software (Minh *et al.* 2020; Mo *et al.* 2022). From these analyses, we tested two approaches to identify sets of informative genes for elucidating relationships among *Anastrepha* species, and particularly species in the *fraterculus* group. For the first approach, we used a strategy that simply considered the number of parsimoniously informative sites per gene, as defined by (Nei and Kumar, 2000), and retained all genes with over 1,000 variable sites, derived from IQ-TREE2. We selected a set of 37 genes with over 1,000 variable sites that was referred to as Info_37. To make the other subsets comparable we also selected 37 genes for the following strategies. The second approach retained genes that had sCF values above 60% for phylogenetically informative nodes (the sum of all three informative hierarchical levels discussed above), based on the ML phylogenetic tree inferred by that gene, as well as an average gCF values above 50%. This set is referred to as sgC60_37. For the third approach, we used the SortaDate software (<https://github.com/FePhyFoFum/sortadate>) employing a combination of two criteria to choose a subset of genes, largest bipartition support and minimization of tip-to-root variation. SortaDate assesses similarities between gene trees and the species tree, focusing on tree length variation and retains only genes that meet certain optimality criteria for phylogenetic information (Smith *et al.* 2018). This set is referred to as Sortadate_37. We also explored subsets for each of these strategies with a more reduced number of loci, in this case 20 loci per subset (Info_20, sgC60_20, and Sortadate_20), which could make these analyses more amenable for further practical applications.

We evaluated the ability of the three strategies to produce a more reduced number of loci that were still informative to identify lineages at different phylogenetic levels in the *Anastrepha* phylogeny by contrasting the inferred species trees produced on ASTRAL using only the selected genes and the full datasets. This was investigated by looking at the Gene Concordance Factor (gCF) using the IQ-TREE2 to test the node support across the phylogeny, especially for well-defined groups, i.e., clades that are associated with taxonomically or morphologically recognized taxa (following Congrains, Dupuis, Steck, Rodriguez, Sutton, Branham, De Brito and Geib (Congrains *et al.* 2023) and represented on results). This was further analyzed by using PhyParts (Smith *et al.* 2015).

2.4 Incomplete Lineage Sorting (ILS) and Introgression

Given the history of recent divergence, incomplete lineage sorting (ILS) and introgression in Tephritidae, particularly in *Anastrepha* (Congrains *et al.* 2023; Díaz *et al.* 2018; Prezotto *et al.*, 2019), we conducted specific analyses to investigate the relationship between concordance factors and the bootstrap method generated by IQ-TREE. We used these concordance factors to test the assumptions of an Incomplete Lineage Sorting (ILS) model. We conducted pairwise analyses of all concordance factor statistics, generating correlation value assignment graphs.

We investigated for introgression using the Dsuite computational tool (available at <https://github.com/millanek/Dsuite>) to calculate Patterson's D statistics, known as ABBA-BABA metrics along the genome [49]. These parameters were adopted to infer signs of introgression between different lineages. Attention in this study focused on exploring alleles of common ancestry among combinations of three species groupings that are phylogenetically tenable. Due to the heterogeneous distribution of the locations of the genes across the genome, we also performed the f4-ratio analysis [50], which examines introgression signs delimited to specific genomic intervals. We used several sets of species (see Suppl. Table 1) to estimate a compounded general estimate of introgression per gene that would not depend on an individual species set, but rather, would indicate the general propensity of each marker to move across lineage (or “species”) boundaries. Benjamini-Hochberg (BH) correction to adjust p-values were considered to account for multiple tests and graphically represented in figures using Microsoft Excel and the R environment scripts.

For the purpose of investigating introgression, we treated different lineages of *A. fraterculus* as separate species when there were other traits to support that separation, following previous phylogenetic analyses that indicated several clades (Congrains, 2023, 2021; Hernández-Ortiz *et al.* 2015). Furthermore, some of these analyses were only performed for taxa for which we had more than one specimen sampled.

2.5 Phylogenetic and evolutionary parameters

Several parameters were estimated from aligned orthologs to indicate levels of polymorphism, phylogenetic signals, phylogenetic congruence, and levels of introgression using different procedures, which we used to contrast to each reduced subset and the full dataset. This analysis aimed to investigate whether these parameters are associated with different gene subsets. We estimated transition and transversion evolutionary rates calculated as the total length of the tree

divided by the number of terminals (Telford *et al.* 2014), and their levels of saturation, using the Saturation test (Philippe *et al.* 2011) on Phykit (Steenwyk *et al.* 2021). We also used this program to estimate evolutionary rates as one minus the sum of squared frequency of different bases at a given site and treeness (Phillips *et al.* 2003), which investigates how well evolutionary distances among specimens are better explained by trees, rather than networks, with higher values denoting a more significant signal-to-noise relationship.

We also used IQ-TREE2 (Minh *et al.* 2020; Mo *et al.* 2022) to produce a consensus phylogenetic tree (species tree) based on single gene Maximum-Likelihood trees. The analysis employed 1,000 bootstrap replicates to assess statistical robustness of the consensus tree. We employed a heuristic search process with 100 clusters for sequence grouping (--scf 100), activated verbose mode for the phylogenetic inference process with Coalescence Hidden Markov Model (CHMM) (--cf-verbose), and used the phylogenetic inference mode with the double-forest approach (--df-tree). These analyses were used to estimate gCF (gene concordance factor = gCF_N/gN%), sCF (site concordance factor with an average of 100 quartets = sCF_N/sN%), and sC (number of concordant sites with an average of more than 100 quartets). We estimated Internode Certainty per node per gene according to Kobert, Salichos, Rokas and Stamatakis (Kobert *et al.* 2016). Individual values were averaged over branches depending on the comparisons performed.

We used two approaches to investigate patterns of selection on individual gene regions – one that estimated the ratio of nonsynonymous to synonymous changes (ω) on clusters of orthologs, and another that explores patterns of selection across branches of phylogenetic trees. For the former, the number of synonymous (dS) and nonsynonymous (dN) changes and its ratio (dN/dS or ω) was estimated using the M0 model in CODEML implemented in the PAML software (Yang, 1997; Chen *et al.* 2009). A $\omega > 1$ is considered strong evidence of positive selection for amino acid substitutions, while $\omega \approx 0$ indicates purifying selection (Yang and Bielawski, 2000). This was implemented using a python script that uses Ete3 (Huerta-Cepas *et al.* 2016) and Biopython (Talevich *et al.* 2012) modules to calculate ω for each cluster of ortholog in parallel (available at <https://github.com/carloscongrains/pamlM0>). The analysis of patterns of selection on different branches was performed using the branch-site model Busted (Murrell *et al.* 2015), implemented in Datamonkey (Weaver *et al.* 2018). This approach identifies signatures of positive selection (amino acid changes) or purifying selection (amino acid maintenance) by the nonsynonymous to synonymous nucleotide substitution rate ratio (dN/dS) on tree branches (Murrell *et al.* 2015).

2.6 Tests of tree compatibility

The software IQ-TREE2 (Minh et al. 2020) was used to compare whether the topology produced by each subset is significantly different than the topology inferred by the *Anastrepha* dataset. This was done by comparing whether maximum likelihood scores of each subset is significantly different than the score of the full dataset using several different tests. We used Shimodaira-Hasegawa test (Shimodaira and Hasegawa, 1999), as well as the expected likelihood weights (Strimmer and Rambaut, 2002) and the approximately unbiased test (Shimodaira, 2002), performed on IQ-TREE2 (Minh *et al.* 2020).

A visual comparison of the topology produced by the full *Anastrepha* dataset and each subset topology was performed on R (R Core Team R, 2023), using the function `cophylo` of Phytools (Revell, 2024). This function creates co-phylogenetic plots to visualize incongruences between the full dataset and each subset topology.

2.7 Gene Mapping

We used Exonerate (Slater and Birney, 2005) to align one sequence for each ortholog cluster here studied against the genome of *A. ludens* (GenBank: GCA_028408465.1) (Congrains *et al.* 2024) using default parameters and the `est2genome` model to generate a physical map indicating the location of genes on chromosomes. To do so, we chose sequences from *A. ludens*, whenever available, otherwise we used sequences of *A. distincta*, which is closely related to *A. ludens*. The interactive mapping visualization across chromosomes was performed using ChromoMap v0.4.1, an R package that maps any chromosomal feature from known coordinates and illustrates the arrangement and physical or genetic distance between molecular markers or genes on the chromosomes of a genome (Pfeifer and Kapan, 2019). These analyses were performed on the full set of markers, as well as on the phylogenetically reduced sets of 20 and 37 markers.

3. Results

The *Anastrepha* dataset here investigated is derived from the reanalysis of whole genome sequencing, complete genome assemblies or transcriptomes from 36 specimens of 15 *Anastrepha* species which included seven species groups and seventeen samples from the *A. fraterculus* complex across South America and Mexico as well as five species of other tephritid genera. This dataset consisted of 3,170 clusters of orthologous genes that had an average of 1,432 bases per

cluster, providing a total of 4,551,036 bases and about 20.5% missing data for the ingroup. The *fraterculus* dataset was composed of a subset of the *Anastrepha* dataset, including 13 species of the *fraterculus* group, *A. psidivora*, *A. striata*, and *A. bistrigata*. This set was composed of 3,168 with an average of 1,545 bases per cluster, totaling 4,895,310 bases and about 21% of missing data. Phylogenetic inferences with the whole dataset produced basically the same topology produced elsewhere (Figures 1 and 2). We explored the lack of phylogenetic signal at different parts of the topology to investigate which evolutionary forces might be influencing this pattern.

3.1 Species tree inference and concordance analyses

The topologies of the species trees derived from the *Anastrepha* dataset, which encompasses all *Anastrepha* samples here studied, and the outgroups, as well as the *fraterculus* dataset, which restricted the sampling to samples from the *fraterculus* species group, *A. psidivora* and two outgroups, using different methodological approaches were highly congruent with previous analyses (Congrains *et al.* 2023) (Figures 2 and 3). Most nodes associated with separation of potentially valid different taxa (be it clades within the *A. fraterculus* complex, species, or genera) had high support (100% bootstrap for the concatenated approach and 1.0 PP) for the multispecies coalescent approach). Despite the high bootstrap and PP support, separate analyses of multispecies coalescent congruence across different genes showed high levels of phylogenetic incongruence across different genes, in general presenting values smaller than 50% of congruence for the gene concordance factors (gCF), site concordance factors (sCF) values, and quartet support. Because most of the incongruence was associated with the more closely related taxa and the results between the two datasets are very similar, we only show here the phytools tree based on the whole *Anastrepha* dataset (Figure 4).

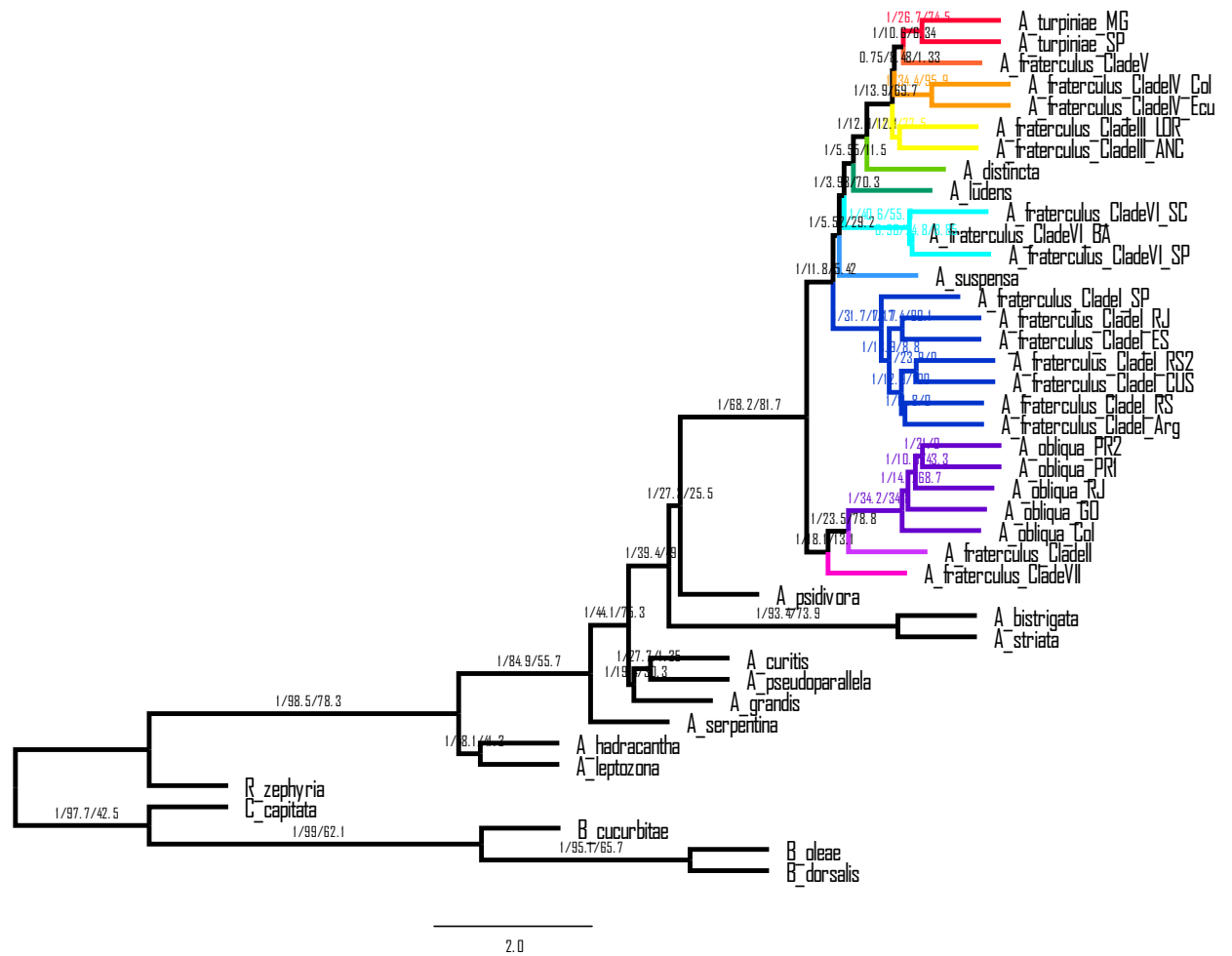


Figure 2. Multispecies coalescent species tree of *Anastrepha* species and five other Tephritidae species as outgroups based on 3,170 genes inferred in ASTRAL. Phylogenetic supports (bootstrap / gene concordance factor / quartet support) are shown on the nodes.

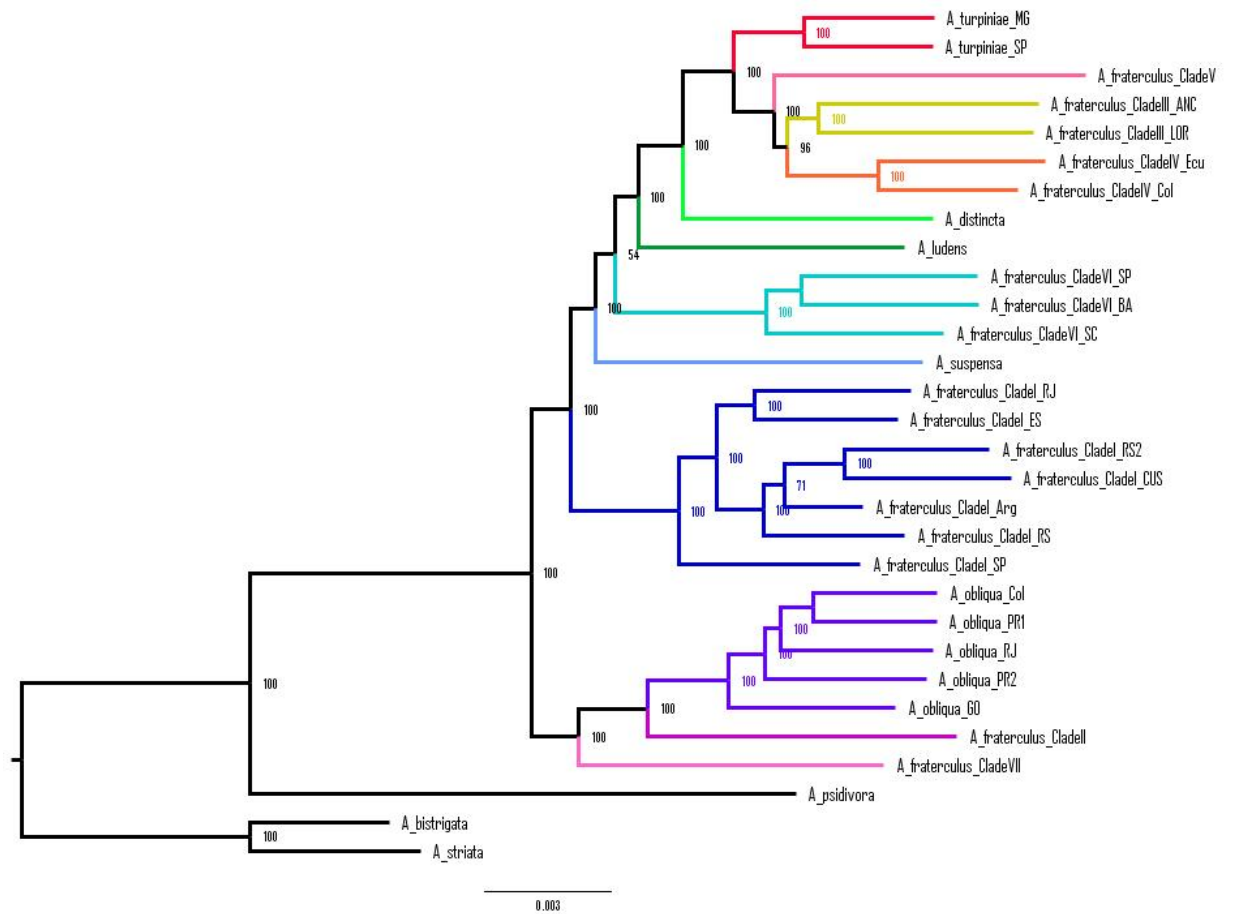


Figure 3. Maximum likelihood inference from concatenated 3,168 genes of *Anastrepha* species with focus on the *fraterculus* group using *A. striata* and *A. bistrigata* as outgroups inferred in RAxML. Bootstrap supports are shown close to the nodes

As expected, the deeper nodes in general had greater support and more concordance over different markers, but there were some shallower nodes with high support as well. Only a limited number of branches in the topology had congruence values over 50%, and these mostly were associated with deeper branches, which is particularly true for the nodes involving the *A. fraterculus* complex. Because the topologies produced from the full datasets do not vary from what previous analyses have shown, where a more detailed analysis of these relationships was presented (Congrains *et al.* 2023), we will only emphasize here aspects that are enriched by the congruence analyses. The main point of discordance between different datasets and methodologies entails the separation of clades of *A. fraterculus* in the West South America and *A. turpiniae*. The ML concatenated analyses of all *Anastrepha* data failed to separate *A. fraterculus* Clades III, from Peru, and IV, from Colombia and Ecuador, but placed *A. fraterculus* Clade V, also from Peru, more closely related to *A. turpiniae*. This was basically the same arrangement for the ASTRAL topologies with both datasets, though ASTRAL separated *A. fraterculus* Clades III, and IV. The ML analyses of more closely related species in the *fraterculus* dataset separates these different West South American *A. fraterculus* clades, but places *A. turpiniae* as a sister taxa.

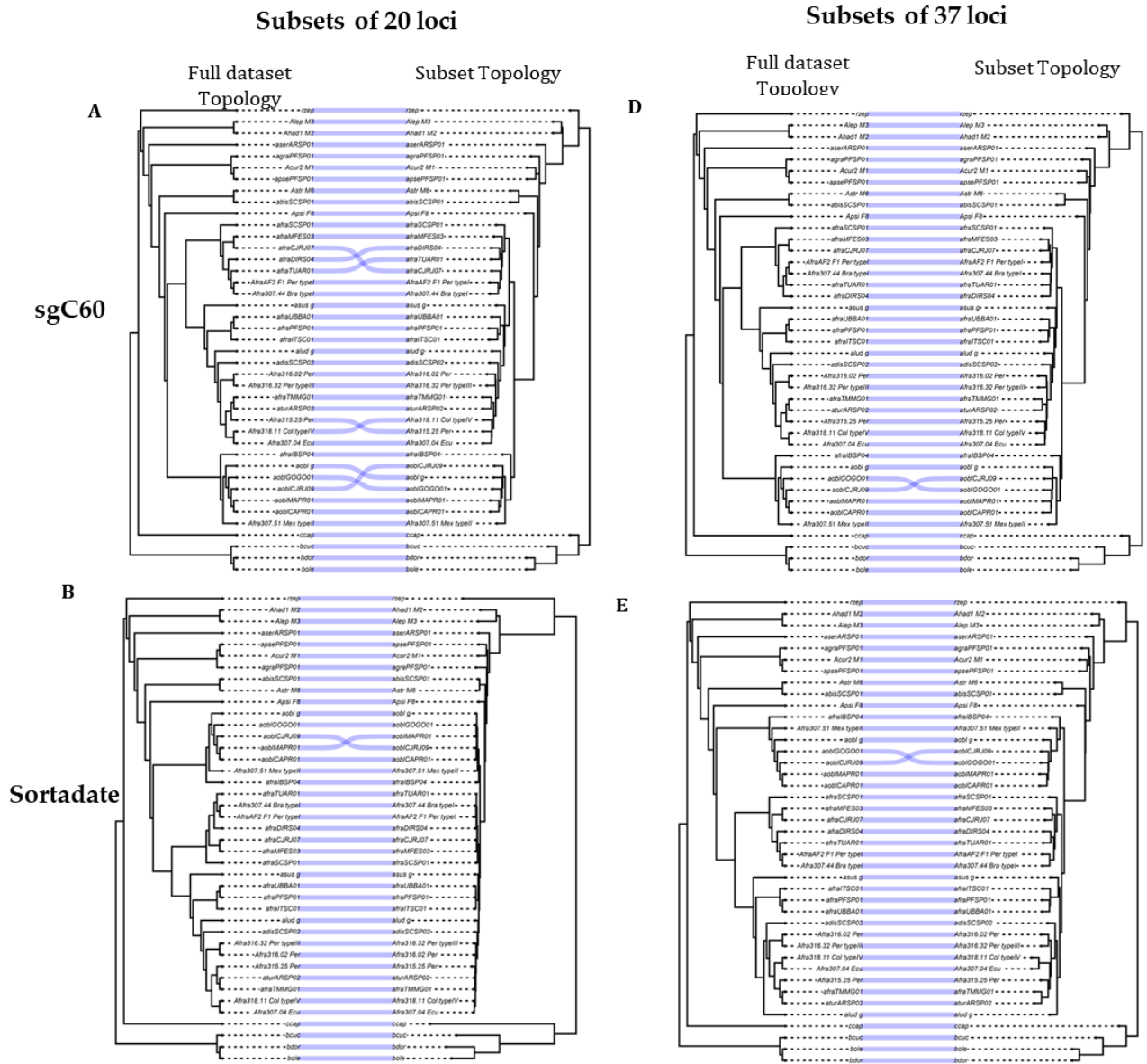
The lack of resolution among lineages in the West South American is reflected on the congruence analyses, that show over 60-70% discordance across different genes for most of these relationships and quartet supports between 2 and 20% (Figure 2). That is not too different from what is observed across the *fraterculus* species group, with gCF support values for several different species failing to exceed 35%. In fact, the *fraterculus* complex exhibited considerable variation in support, ranging from 11% to 70% for different lineages, some of which that could potentially represent different species (Dias *et al.* 2016; Congrains *et al.* 2023). The support for the *fraterculus* species group, on the other hand, is very robust (75% of the genes supporting this lineage) and showing *A. psidivora* as sister taxa to the *fraterculus* group. As we compare more distantly related taxa, the level of congruence across different genes tend to increase, being around 48% for the comparisons across species groups and reaching over 90% for the separation of genera, placing *Anastrepha* more related to *Rhagoletis* than to *Ceratitis*, *Bactrocera*, and *Zeugodacus*, a result corroborated by other phylogenetic analyses (He *et al.* 2024) as well as by the current taxonomy, which places the former two genera in Trypetinae, whereas the latter are in the Dacinae subfamily.

Despite high levels of incongruence across different loci on the genome, indicated by sCF and gCF values below 50% in several nodes, the consistency across datasets and methodologies suggests the existence of the lineages we observed, which also correlates well with genera, species groups and species (even in the *fraterculus* complex). We explored these results by investigating whether we could find a smaller assortment of genes that might be more phylogenetically informative and congruent than the average loci as measured by multiple parameters.

3.2 Phylogenetic Inferences and the Search for Informative Genes

The three different strategies chosen to find a reduced set of informative genes (See list of genes and some parameters on Suppl. Table 2) produced different topologies in general. Tests of tree compatibility shows indicates that all trees produced by subset strategies are significantly worse than the one produced by the full *Anastrepha* dataset using Shimodaira Hasegawa (SH), weighted SH, Expected Likelihood weight or unbiased test (data not shown). Despite these significant differences, these discrepancies mostly involve closely related samples, and this is more apparent when we chose sets of 20 genes rather than 37, which can be visualized on the cophylo analyses (Figure 5). Like before, these results were very similar when comparing the two different

main datasets, so we chose to retain information only for the *Anastrepha* dataset, considering that the markers chosen should provide information across different divergence spectra.



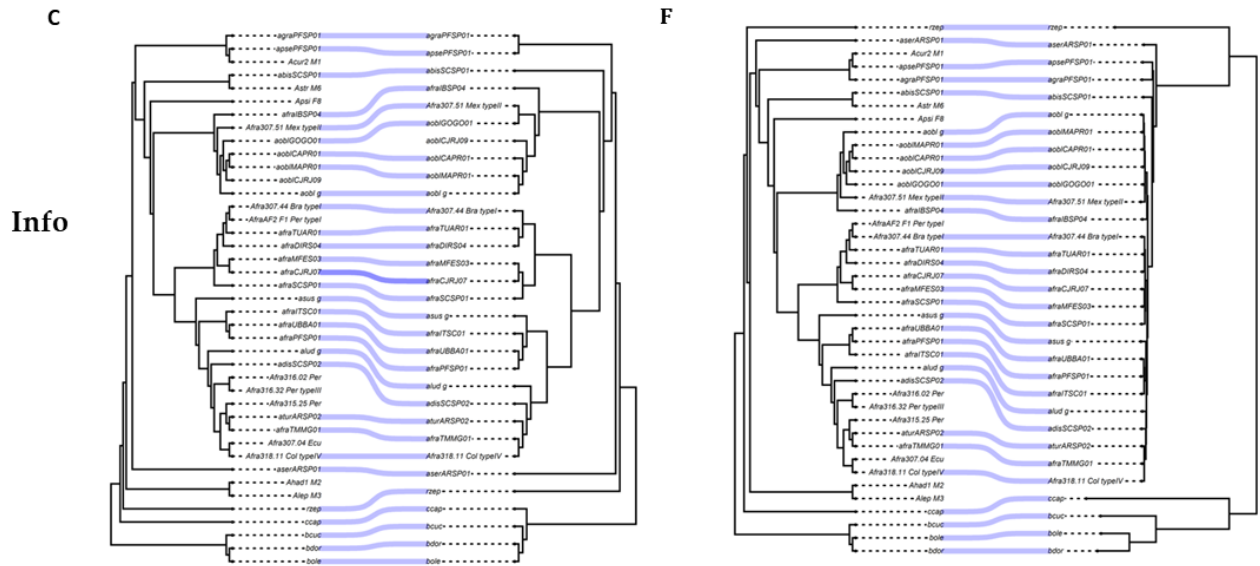


Figure 5: Visualization of phylogenetic congruence between the topology produced by all loci (on the left of each cophylo tree above) and each of the three gene subsets of 20 and 37 loci A) 20 highest $sC>60\%$, B) 20 highest genes inferred from Sortadate; C) 20 genes with highest number of informative characters D) 37 highest $sC>60\%$, E) 37 highest genes inferred from Sortadate; F) 37 genes with highest number of informative characters. Blue lines indicate matching taxa on contrasting trees.

The strategy of choosing the genes with highest number of informative characters (dubbed info sets) failed to provide phylogenetic information for all the specimens here studied in the larger datasets, whereas the other methods represented the whole array of taxa here investigated (Figure 4A-F shows ASTRAL trees for sets of 20 loci and 37 loci, whereas the ML inferences are shown on Suppl. Figures 1 and 2). These different strategies showed different efficiencies at recovering the original topology and congruence. Even though info_20 and info_37 subsets, based on number of informative loci, also produced topologies that were fully compatible with the overall tree, that encompassed a smaller number of taxa, and these were associated with higher congruence values for the topology of the remaining taxa (Figure 4C and Figure 4F). These datasets failed to capture enough data to include in the tree at least five species, *A. hadracantha*, *A. leptozona*, *A. curitis*, *A. striata* and *A. psidivora*, as well as a few specimens in the *fraterculus* species group. The other two strategies, Sortadate and sgC60 produced topologies that represented all samples here studied and better matched the phylogeny derived from the full datasets. More importantly, the latter strategies produced topologies with less phylogenetic discordance across loci.

In general, the percentage of genes supporting or not a specific topology, or nodes, varied, but were higher in the subsets of genes here chosen, when compared to the whole datasets. Around 40% of the loci were informative and compatible with the topology produced by all loci, a number

that decreased to 32% when only considering the variation among *Anastrepha* species. When we consider the sets of 20 loci, the sgC60_20 had over 53% genes informative for the whole assembly of species and 40% when only considering *Anastrepha*, but failed to resolve the relationships among West South America *A. fraterculus* clades (Clades III, IV and V) and *A. turpiniae*. The Sortadate_20 had values of 62% (all taxa) and 52% (only *Anastrepha*), respectively and the Info_20 had values of 67% and 61%, the highest among them, though it should be remembered that not all specimens are represented in this strategy. This pattern remains when we change to sets of 37, increasing on average by about 20% of genes being informative across all nodes with sC60_37 and Sortadate_37, and maintaining similar values for the informative set. We should mention that even though all sets are basically represented by different sets of loci (graphically represented on Figure 6), they consistently produced robust phylogenetic inference and compatible with the one produced by the complete dataset, with the only exception being the separation of *A. turpiniae* and relationship of that to other *A. fraterculus* from the West South America with 20 loci (Figures 7 and 8). Because the levels of support varied across different lineages and different representative sets, we explored other parameters to understand whether there were specific parameters associated with the sets of selected loci at different points in the hierarchy (Suppl. Figure 4 shows nodes that were used to represent separation of species, species groups and genera).

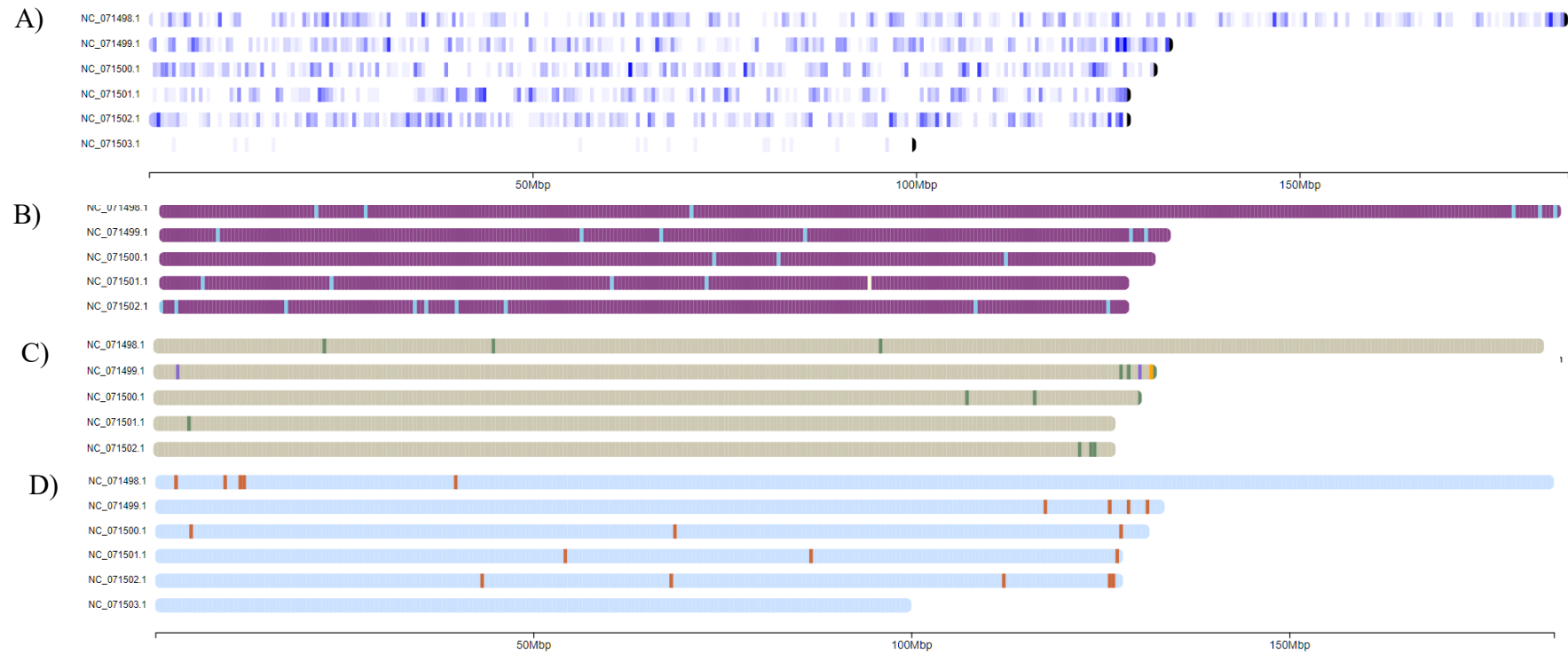


Figure 6: Chromomap distribution of loci across the *A. ludens* genome. A) Full set of 3170 genes, B) Distribution of Info_37 – 37 genes with more than 100 informative characters; C) Distribution of sgC60 – Set of genes with highest sC60%; D) Distribution of Sortadate_37 – Set of genes with longer and more robust internal branches inferred from Sortadate.

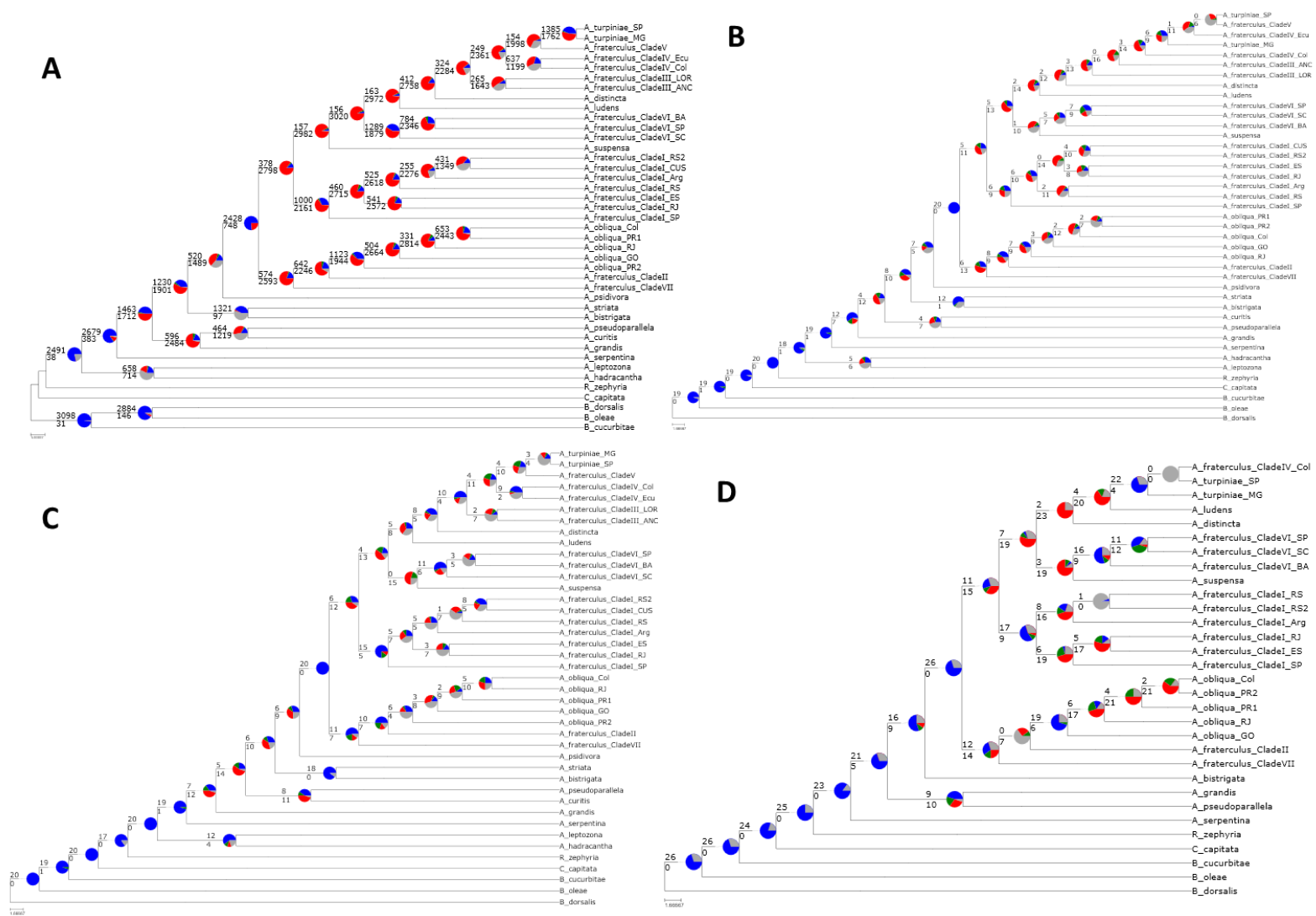


Figure 7: Summary of concordance and conflict inferred on Astral and plotted with Phyparts and Phypartspiecharts. A) all 3179 genes. B) 20 highest sC > 60%, C) 20 highest genes inferred from Sortadata; D) 20 genes with highest number of informative characters (ASTRAL). Piecharts indicate number of genes which support the node (blue), support the most frequent alternative (green), support any of the other alternatives (red), and failed to be informative for that node.

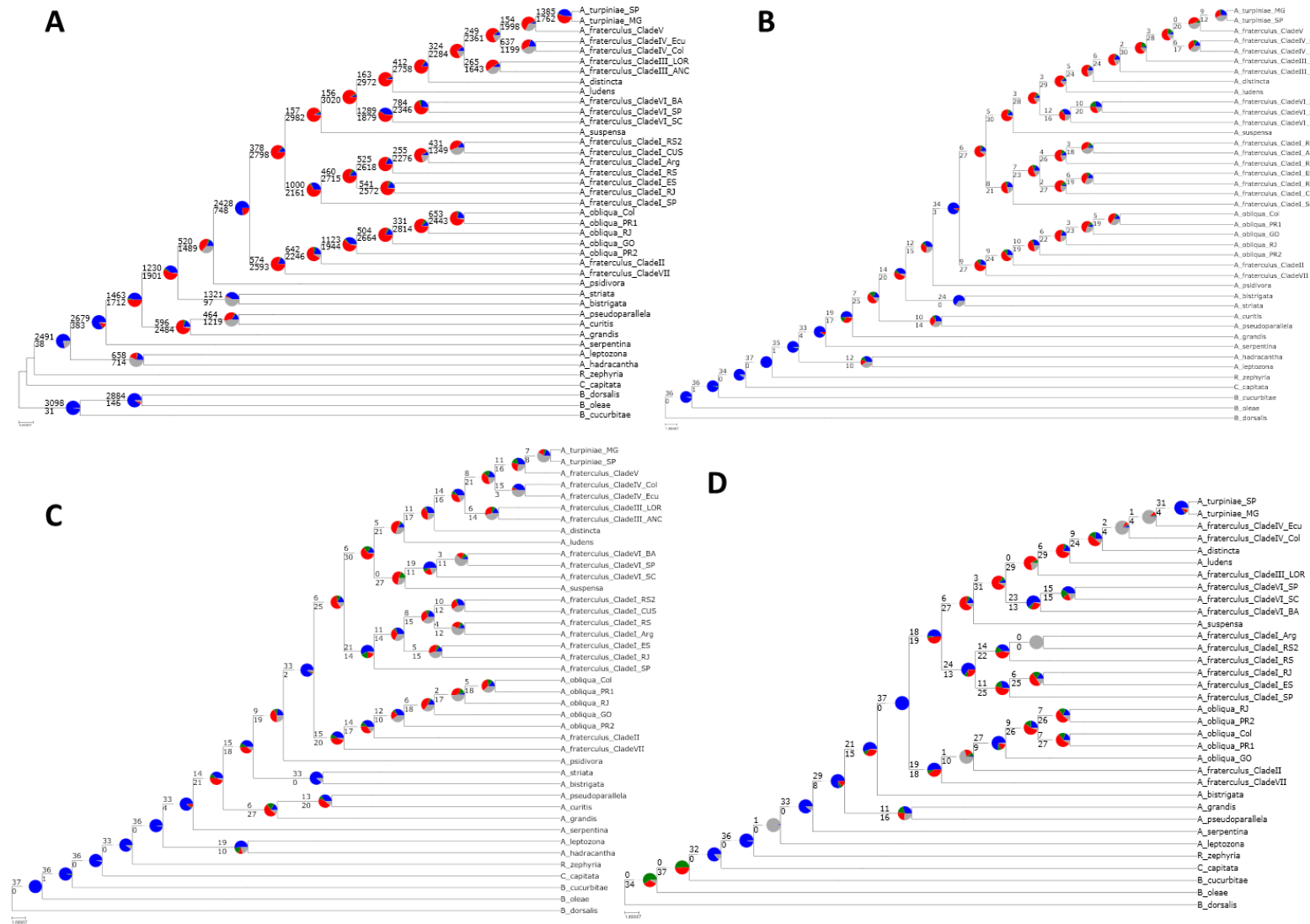


Figure 8 Summary of concordance and conflict inferred on Astral and plotted with Phyparts and Phypartspiecharts. A) all 3179 genes. B) 37 highest $sC > 60\%$, C) 37 highest genes inferred from Sortadata; D) 37 genes with highest number of informative characters (ASTRAL). Piecharts indicate number of genes which support the node (blue), support the most frequent alternative (green), support any of the other alternatives (red), and failed to be informative for that node.

3.3 Association of phylogenetic parameters to reduced subsets.

We investigated several phylogenetic and evolutionary parameters across different genes on the genome, summarized on Table 2 and visually represented on Figure 8. The focus of these analyses was more on their effect and association to the three subsets here described, rather than a thorough investigation of these parameters on the genome of *Anastrepha*. In general, though, these parameters indicate that the contigs here studied are evolving under purifying selection, indicated by average low values of omega (ω) over different regions of the genome for the two datasets, *Anastrepha* and *fraterculus* (Table 2, Figure 8A-B and Suppl. Figure 5), with very few exceptions with dN/dS values of at least 0.5, an indication they might be evolving under positive selection (Swanson *et al.* 2004). It is interesting that the *fraterculus* dataset, which compared more closely related taxa, had higher average values of ω than the *Anastrepha* dataset (0.10 and 0.06, respectively). These lower dN/dS values for more distantly related taxa could be due to saturation, and in fact, these values are higher on the *Anastrepha* dataset compared with the *fraterculus* dataset (Table 2, Figure 8C-D and Suppl. Figure 6), but they affected mostly the more divergent taxa and did not seem to impact much the phylogenetic signal. This is indicated by the average evolutionary rate values across the genome, which is higher when comparing more distant taxa in Tephritidae than among species and clades of *Anastrepha* (Table 2, Figure 8E-F and Suppl. Figure 7), indicating that the higher saturation values are not enough to quash the association between evolutionary distance and change. This is reflected also on average values of treeness (Table 2 and Figure 8G-H), that measures the degree to which the total length of the phylogenetic tree is retained in internal branches, which is higher in the *Anastrepha* dataset, compared to the *fraterculus* dataset. Despite the presence of evolutionary signal, there is a great deal of incongruence across different portions of the genome, reflected on the distribution of gCF and sCF values across the genome. We investigated that by contrasting the support of the main topology (sC) against the alternatives (sD1 and sD2) for each node and plotted the IC (Internode Certainty) index (Table 2). A comparison of these estimates with average f4-ratio per gene as a conservative estimate of introgression suggested that introgression decreases as you investigate more divergent taxa (Table 2).

Table 2 – Average values for different evolutionary and population parameters across the whole dataset and in each reduced subset.

	<i>Full dataset^a</i>	<i>sgC60_20^b</i>	<i>sgC60_37^b</i>	<i>Sortadate_20^c</i>	<i>Sortadate_37^c</i>	<i>Info_20^d</i>	<i>Info_37^d</i>
ω_{Busted} <i>fraterculus</i>	0.1039	0.1774**	0.1442	0.1267	0.1215	0.1436	0.1493
Saturation <i>fraterculus</i>	0.8161	0.8358	0.8244	0.8005	0.7737	0.8187	0.8697
Evolutionary rate <i>fraterculus</i>	0.0048	0.0035	0.0034*	0.0038	0.0035	0.0058	0.0055
Treeness <i>fraterculus</i>	0.2875	0.3777**	0.3703***	0.4216***	0.3997***	0.2336	0.2538
ω_{Busted} <i>Anastrepha</i>	0.289	0.296	0.3292	0.2506	0.2649	0.2413	0.1759
Saturation <i>Anastrepha</i>	0.0239	0.0215	0.0212	0.0225	0.0207	0.0333***	0.0324***
Evolutionary rate <i>Anastrepha</i>	0.3484	0.3672	0.3619	0.3593	0.3553	0.3582	0.363
Treeness <i>Anastrepha</i>	0.0048	0.0035	0.0034*	0.0038	0.0035	0.0058	0.0055
IC among all taxa	0.2063	0.2597	0.2745***	0.2527	0.2619***	0.1398**	0.1484***
IC among species	0.2279	0.3224*	0.3328***	0.3483*	0.312**	0.1524	0.1783
IC among species groups	0.1504	0.1916	0.1931	0.208	0.2213***	0.0152***	0.0178***
IC among genera	0.2149	0.2589	0.2492	0.2121	0.227	0.1435*	0.1427***
IC phylogenetically relevant nodes	0.3913	0.3586***	0.3603***	0.3663*	0.3698**	0.4232**	0.417***
IC non-phylogenetically relevant nodes	0.3323	0.2986	0.288	0.2883	0.2845	0.3686	0.3603
f4-ratio interpopulations	0.0252	0.0253	0.0253	0.0254	0.0253		
f4-ratio closely related species	0.0254	0.0288	0.0288	0.0305	0.0305*	0.0149***	0.0153***
f4-ratio among species of <i>fraterculus</i>	0.0268	0.0295	0.0293	0.0295	0.0294	0.0195***	0.0197***
f4 distantly related species	0.0118	0.0127	0.0127	0.0127	0.0128	0.0094***	0.0091***

^a Average values for the *Anastrepha* dataset for different parameters, unless indicated that values are from the *fraterculus* dataset. Significance values correspond to * p<0.05, ** p<0.01, *** p<0.001, after BH correction for multiple tests.

^b sgC60_20 and sgC60_37 refer to sgC60 sets of 20 and 37 genomic regions that show at least 60% average sCf values in phylogenetically relevant nodes (see text) and over 50% average gCf.

^c Sortadate_20 and Sortadate_37 correspond to subsets of 20 and 37 genomic regions, respectively, chosen by a process that focuses on evolutionary rate variation across the topology and bipartition values.

^d Info_20 and Info_37 correspond to subsets of genomic regions defined by loci with largest number of informative characters.

An analysis of the distribution of evolutionary and populational parameters across the three different subsets here chosen to subsample genomic information indicates some interesting results. In general, there are two important parameters that are reflected on the subsets here chosen, treeness, significantly higher on sgC60 and sortadate subsets, but not the info subset, and average IC, which is in general higher for the former two subsets as well, for many contrasts, significantly so. Interestingly, both info subsets had on average values of IC lower than the average gene on the genome (Table 2 and Figure 8, which also show the average values of random subsets of 20 and 37 loci for all attributes here investigated). For almost all cases, level of significance increases with the number of genes in the subset, that is, increases

when moving from 20 to 37 genes. Interestingly, there is also significant association with introgression estimate f4-ratio in info subsets, which is significantly lower than other genes on the genome (Table 2 and Figure 9), but this could be a consequence of the absence of several population samples, so much so that there is intrapopulation f4-ratio estimate from info.

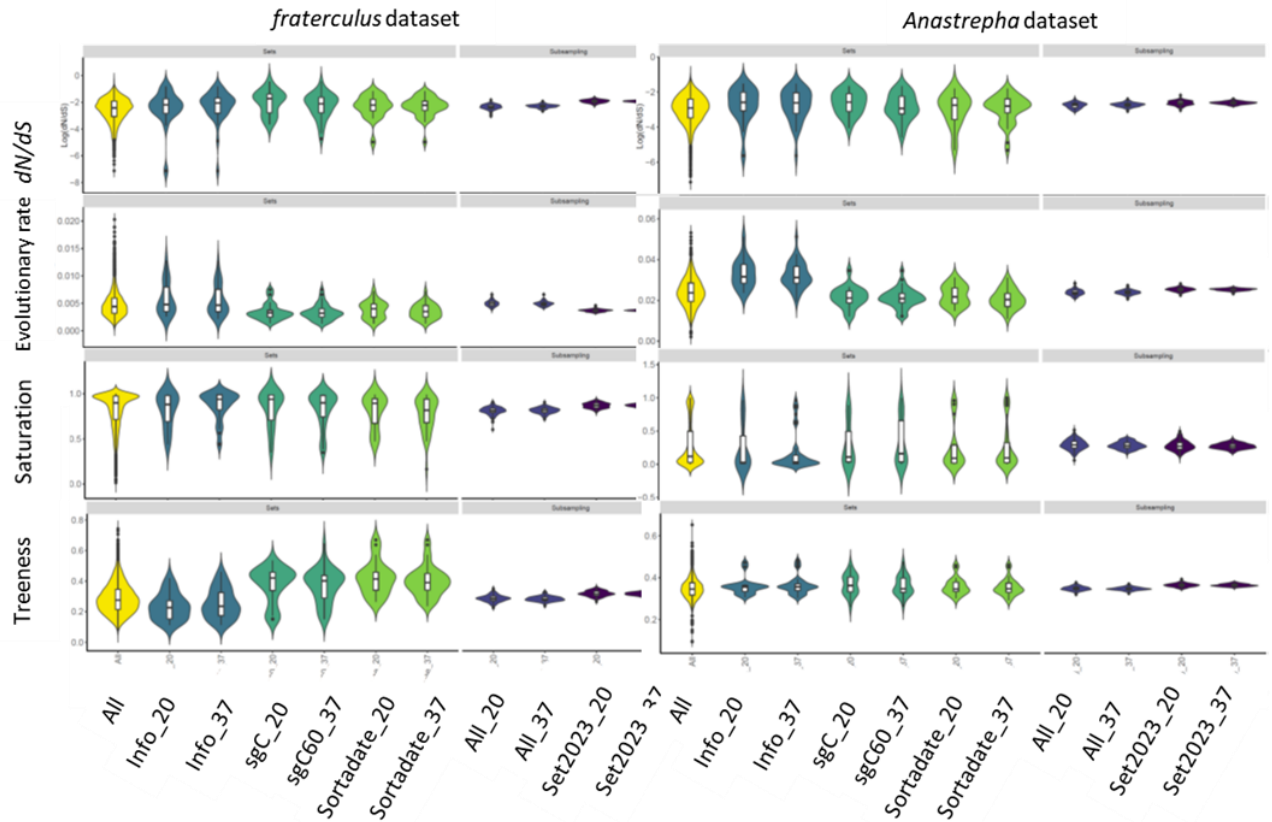


Figure 9: Violin plots of values of A) dN/dS , C) Evolutionary rate, E) Saturation and G) Treeness for the *Anastrepha* dataset, and B) dN/dS , D) Evolutionary rate, F) Saturation and H) Treeness for the *fraterculus* dataset. Sets are as follow: all (corresponding *Anastrepha* or *fraterculus* dataset), subsets sgC60_20 and sgC60_37, sortadate_20, sortadate_37, info_20 and info_37, as well as random sets of 20/37 loci from the whole dataset (all_20 and all_37 are resampling of the current datasets, whereas set2023_20 and set2023_37 are resampling of the original Congrains *et al.* 2023 dataset).

4. Discussion

Previously, we created a workflow to identify and align orthologs from different taxa that was very useful to infer phylogenetic relationships among different species of *Anastrepha* (Congrains *et al.* 2023). This is relevant when you consider that this genus is composed of a great number of closely related taxa, especially in the *fraterculus* group, whose divergence might even have been influenced by gene flow (Dias *et al.* 2016; Congrains *et al.* 2021; Selivon *et al.* 2022; Scally *et al.* 2016), which complicates inference of their phylogenetic relationships. This complexity limits the ability of a reduced set of markers to effectively identify different taxa, and affects practical and streamlined applications that would profit from the use of a

smaller set which would be cost- and time-effective. In that paper, some of us used the whole dataset, as well as a more reduced set of over 100 phylogenetically informative loci to identify lineages in the *fraterculus* group which in general agrees with the current taxonomy (Congrains *et al.* 2023). Here we explore three strategies to identify smaller subsets of markers that would remain useful not only for species identification, but also for phylogenetic inference among more closely related taxa and even lineages in the *A. fraterculus* complex. The search for minimal gene sets is central to comparative genomics (Koonin *et al.* 2005), though due to current database sizes, they are dependent on phylogenetic scoring across extensive gene sets which may introduce biases, as genes vary across time and space in evolutionary and populational parameters (Lafond *et al.* 2018). We tested three approaches to investigate phylogenetic relationships by leveraging different phylogenetic and population parameters that allowed us to identify subsets which were informative for the purpose of phylogenetic inference and lineage identification using either 20 or 37 groups of loci.

First, it is noteworthy that all three chosen strategies to find phylogenetically informative genomic regions for the *Anastrepha* phylogeny, be it considering 20 or 37 genomic regions, produced topologies that were significantly worse than the one produced by the full *Anastrepha* dataset using Shimodaira Hasegawa (SH), weighted SH, Expected Likelihood weight or approximately unbiased test (data not shown). Despite these significant differences, all three methods chosen to create a smaller subset of genomic regions produced topologies that were very similar to the one produced with the full dataset of 3,170 genomic regions (Figures 9A-C and 10A-C). These subsets correctly depicted basically all deeper relationships as well as most of the ones among more closely related taxa. Despite this overall congruence, there are a few differences among the different subsets. The subsets Info_20 and Info_37 failed to capture phylogenetic information for all samples, even though it produced the same topology for the samples for which information was available. It should be noted that these are most likely due to the fact that some genomic regions that show large numbers of variable characters were too big and had only portions of the gene recaptured for different samples. With that, the topology produced from Info_20 and Info_37 failed to place in the final topology at least five species, such as *A. hadracantha*, *A. leptozona*, *A. curitis*, *A. striata* and *A. psidivora*, as well as other specimens in the *fraterculus* species group. The sample reduction also affected introgression estimates, since there was not enough variation on info datasets to estimate intraspecific introgression. The topology of the specimens present in the tree, on the other hand, completely match the one obtained for the full dataset (Figure 10C). The other strategies produced significant phylogenetic information from all samples which was, for most taxa, similar to the

topology produced from the whole dataset, with minor differences between a few closely related taxa. The tree produced from the sC60_20 and the full dataset showed three main differences, one among different samples of *A. fraterculus* Clade I from Brazil and Argentina, a couple of *A. fraterculus* from West South America (samples from Peru, Ecuador and Colombia), and among different samples of *A. obliqua*, all pertaining to closely related samples. When we increased the number to 37 genomic regions there was only one intra-clade difference in *A. obliqua*, immaterial to the final topology.

As we mentioned before, we chose to test the reduced subsets of genomic regions using only data from the full *Anastrepha* dataset to investigate the effectiveness of these smaller subsets to infer proper phylogenetic relationships across different levels of the phylogenetic hierarchy. That is, we tested the effectiveness of these reduced datasets to infer relationships among more closely related, as well as among more distantly related taxa, based on the comparison with a genome-scale dataset. We should mention though that some of the few differences between the topology produced by the full *Anastrepha* dataset (Figure 2) and the *fraterculus* dataset (Figure 3) is exactly among a few closely related taxa or inside a recognized taxon. For example, the position of *A. turpiniae* along *A. fraterculus* Clade III, *A. fraterculus* Clade IV and *A. fraterculus* Clade V, as well as the position of *A. obliqua* from Colombia, which is basal in the full *Anastrepha* dataset, whereas it is in the middle of other Brazilian populations in the *fraterculus* group dataset. It is interesting that both subset methods, sgC60 and Sortadate, even with 37 genomic regions, produced topologies changing the position of *A. obliqua* from Colombia and Brazil, whereas the sgC60 with 20 genes also produced different topologies for the position of *A. turpiniae* and these West South American *A. fraterculus*. These incongruences can be clearly seen on the phyparts analyses (Figure 4), since the placing of *A. fraterculus* Clade V along *A. turpiniae* is supported only by about 5% of the genomic regions, whereas about 50% support alternative relationships.

Even though the subsets were mostly effective at recovering the phylogenetic relationships among taxa inferred from the full dataset, which also matched the topology previously obtained (Congrains *et al.* 2023). The different strategies used to identify each subset were based on different parameters, but when we investigate these parameters in each subset against the whole dataset, not always did we find significant differences between them. The Sortadate strategy, for instance, chooses regions that show the lowest evolutionary rate variation across different branches combined with the highest bipartition support. The latter was estimated considering the topology inferred by the whole *Anastrepha* dataset from ASTRAL or ML. Sortadate subsets (be it with 20 or 37 loci) had basically the average values for all

parameters here considered, from dN/dS , IC, evolutionary rates, treeness or introgression, just like what we found for the informative subsets. For the latter, we might not have expected it to be differentially picking any particular genomic regions, but since Sortadata uses bipartition support as one of its parameters, we might have expected it to show on average higher IC values than average. In fact, sC60% subsets were the only to show different average values when compared to other regions on the genome, for several parameters. In general, sC60% gene subset had higher average Internode Certainty (IC), which is not surprising, considering this parameter is based on sC values across different nodes, but also higher dN/dS rates estimated from more distantly related species (though not significant) and among more closely related taxa in the *fraterculus* dataset. Despite higher values of omega, in general sC60% subset had significantly lower evolutionary rates than other genes on the genome. Finally, sC60% also had marginally significantly higher f4-ratio estimates, indicating that in general these genes seem to show higher rates of introgression compared to other regions of the genome.

The combination of higher dN/dS values and, lower evolutionary rates found on the sgC60 and Sortadata subsets may seem counterintuitive. It should be mentioned that even these higher dN/dS than average still suggest a preponderance of purifying selection ($dN/dS < 0.2$), which is common across protein coding genes (Kimura, 1977; Dukler et al. 2022). It does suggest that there might be a disproportionate representation of non-synonymous changes in sgC60 subsets, compared to other changes. This could be a consequence of sexual selection driving rapid evolution, for instance in sex-related genes, but a cursory Gene Ontology analysis performed on these subsets failed to find significant enrichment (data now shown), though this could be due to the reduced number of genomic regions present. This pattern could also be due to adaptive evolution driving the differentiation of multiple clades, for instance, as adaptation to different hosts or environmental conditions. The direct impact of selection is very difficult to identify in closely related taxa because it is difficult to separate the polymorphisms still segregating at the population level from the ones being fixed by selection, a process that is highly impacted by stochastic and populational processes (Kryazhimskiy and Plotkin, 2008). Furthermore, in some instances, even a few evolutionary changes may have a significant adaptive impact, but that signal may be overwhelmed by the general purifying selection acting on the rest of the proteins (Bloom, 2017; Suzuki and Gojobori, 1999). Some studies have shown that few, or even a single non-synonymous change may have important adaptive impact, such as adaptation to different hosts (Lang *et al.* 2012; Dworkin and Jones, 2009), resistance to infestation (Shi *et al.* 2022; Schmelz *et al.* 2012) or changes in aposematic coloration (Rosser *et al.* 2024; Jay *et al.* 2018), that could be preferably moved across different lineages. Patterns

like that could help explain why you would see these patterns also associated with higher f_4 -ratio values, if these regions would be actively involved with the adaptation.

The general pattern of phylogenetic incongruence among different regions across the genome may be a consequence of differential responses to adaptation, as we indicate above, but most often is a consequence of other processes as well, such as ancestral lineage sorting, or the segregation of ancestral polymorphism caused by stochastic processes, and introgression, or the gene exchange among diverging lineages or species (Steenwyk *et al.* 2023; Mirarab *et al.* 2021; Edwards *et al.* 2016). Several studies have shown that a combination of these two processes may be the best combination to explain high values of phylogenetic incongruence across the genome, and it is not the purpose of this exercise to investigate this further here. It is the combination of these different processes, segregation of ancestral polymorphism, introgression and adaptive sieving, which complicates phylogenetic inferences in *Anastrepha* in general, and species in the *fraterculus* group in particular. A few studies have suggested that individual traits, such as morphology or behavior, or a few loci, in some cases even a single locus, such as COI or ITS, might be informative to study variation across *Anastrepha* (Prezotto *et al.* 2019; Mith-Caldas *et al.* 2001; Sutton *et al.* 2015). Even though these markers in general may be effective to investigate part of the variation across the genus, and even help differentiate several species, our studies have indicated that they alone fail to capture the whole scenario, exactly because of the complexity of evolutionary and population processes that have shaped these species histories (Congrains *et al.* 2023).

It is important to reinforce that the purpose of this exercise was to find subsets of genomic regions that were effective at separating different lineages that could correspond to different species/taxa, and this was exactly what was achieved. Considering that several of these taxa diverged recently in the presence of gene flow (Congrains *et al.* 2023; Prezotto *et al.* 2019; Scally *et al.* 2016), and several samples here studied belong to the same species, there should be no expectation that consistent topological congruence would be found across the genome, exactly because these are populational samples that are still experiencing all evolutionary forces, such as the balance between gene flow, drift and selection, among other forces that shape intra and interpopulational variation. The identification of genes that account for evolutionary processes such as introgression, gene flow, and hybridization is critical for clarifying phylogenetic relationships among closely related and cryptic species. Genome-wide scans for “barcode” loci are increasingly used to find informative loci. For *Anastrepha*, this task is challenging due to their adaptation to diverse sets of hosts and environments (Norrbon *et al.* 2018; Zucchi, 1988; Zucchi, 2000), which in other Tephritid species have proved complex to

study (Egan *et al.* 2015; Feder *et al.* 2014), as well as other complex populational processes, such as asymmetric gene flow observed in some species (Congrains *et al.* 2023; Díaz *et al.* 2018). Understanding how introgression influences speciation and hybridization aids in elucidating adaptive processes and the spatial and biological boundaries of species (Hedrick, 2013; Harrison and Larson, 2014) demand broader sampling not only of specimens from different localities and species, but also across the genome.

Finally, the analysis of sets of orthologs, followed by their reduction to smaller subsets of genes with phylogenetic potential, goes beyond an academic exercise. The genus *Anastrepha* includes many pest species that cause significant damage to fruit crops of great commercial value. Developing genetic markers for identifying these species offers a promising tool that can complement or even replace traditional morphometric methods. This is even more important, considering the preeminence of the use of morphometric data on adult females for proper identification of different taxa in the genus, even among more closely related species (Norrbon, *et al.* 2018; 2000). This approach can speed up and improve identification, allowing for more efficient and targeted control. Additionally, these strategies may enable the use of species-specific management strategies which can reduce the costs of insecticides and other expensive, environmentally harmful methods, promoting more sustainable solutions for producers.

The phylogenetic investigation of *Anastrepha* fruit flies here conducted, not only added substantial knowledge to the evolutionary patterns associated with their differentiation, but particularly provided tools that can be effectively employed in discriminating lineages of species whose determination still poses significant challenges. The subsets here produced provide a promising tool to further studies on this group, and help in the control of potential important pests, while underscoring the relevance of analyzing potentially relevant subsets of genes as potential phylogenetic markers for the study of evolutionary relationships among species of the genus *Anastrepha*.

References

- ABRAHAM, S.; CLADERA, J.; GOANE, L.; VERA, M. T. Factors affecting *Anastrepha fraterculus* female receptivity modulation by accessory gland products. **Journal of Insect Physiology**, v. 58, p. 1–6, 2012.
- ALDA, F. et al. Resolving deep nodes in an ancient radiation of neotropical fishes in the presence of conflicting signals from incomplete lineage sorting. **Systematic Biology**, v. 68, p. 573–593, 2019.
- ÁRNASON, Ú. et al. Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. **Science Advances**, v. 4, eap9873, 2018. DOI: 10.1126/sciadv.aap9873.
- BARR, N. B.; CUI, L. W.; MCPHERON, B. A. Molecular systematics of nuclear gene *period* in genus *Anastrepha* (Tephritidae). **Annals of the Entomological Society of America**, v. 98, p. 173–180, 2005.
- BARR, N. B. et al. Identifying *Anastrepha* (Diptera; Tephritidae) species using DNA barcodes. **Journal of Economic Entomology**, v. 111, p. 405–421, 2017. DOI: 10.1093/jee/tox300.
- BLOOM, J. D. Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. **Biology Direct**, v. 12, n. 1, 2017. DOI: 10.1186/s13062-016-0172-z.
- CHEN, G. K.; MARJORAM, P.; WALL, J. D. Fast and flexible simulation of DNA sequence data. **Genome Research**, v. 19, p. 136–142, 2009.
- CHEN, M. Y.; LIANG, D.; ZHANG, P. Selecting question-specific genes to reduce incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny. **Systematic Biology**, v. 64, p. 1104–1120, 2015.
- CONGRAINS, C. et al. Phylogenomic analysis provides diagnostic tools for identification of species complex. **Evolutionary Applications**, v. 16, p. 1598–1618, 2023. DOI: <https://doi.org/10.1111/eva.13589>.
- CONGRAINS, C. et al. Chromosome-scale genome of the polyphagous pest *Anastrepha ludens* (Diptera: Tephritidae) provides insights on sex chromosome evolution in *Anastrepha*. **G3: Genes|Genomes|Genetics**, 2024. DOI: 10.1093/g3journal/jkae239.
- CONGRAINS, C.; ZUCCHI, R. A.; DE BRITO, R. A. Phylogenomic approach reveals strong signatures of introgression in the rapid diversification of neotropical true fruit flies (*Anastrepha*: Tephritidae). **Molecular Phylogenetics and Evolution**, v. 162, 107200, 2021. DOI: <https://doi.org/10.1016/j.ympev.2021.107200>.

- CUMMINS, C. A.; MCINERNEY, J. O. A method for inferring the rate of evolution of homologous characters that can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases. **Systematic Biology**, v. 60, p. 833–844, 2011.
- DARRIBA, D. et al. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. **Molecular Biology and Evolution**, v. 37, p. 291–294, 2019. DOI: 10.1093/molbev/msz189.
- DAUBIN, V.; MORAN, N. A.; OCHMAN, H. Phylogenetics and the cohesion of bacterial genomes. **Science**, v. 301, p. 829–832, 2003. DOI: 10.1126/science.1086568.
- DIAS, V. S. et al. An integrative multidisciplinary approach to understanding cryptic divergence in Brazilian species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). **Biological Journal of the Linnean Society**, v. 117, p. 725–746, 2016. DOI: 10.1111/bij.12712.
- DÍAZ, F. et al. Evidence for introgression among three species of the *Anastrepha fraterculus* group, a radiating species complex of fruit flies. **Frontiers in Genetics**, v. 9, 359, 2018. DOI: 10.3389/fgene.2018.00359.
- DOYLE, V. P.; YOUNG, R. E.; NAYLOR, G. J. P.; BROWN, J. M. Can we identify genes with increased phylogenetic reliability? **Systematic Biology**, v. 64, p. 824–837, 2015.
- DUKLER, N. et al. Extreme purifying selection against point mutations in the human genome. **Nature Communications**, v. 13, 4312, 2022. DOI: 10.1038/s41467-022-31872-6.
- DUTRA, V. S. et al. Egg morphology of *Anastrepha* spp. (Diptera: Tephritidae) in the *fraterculus* group using scanning electron microscopy. In: **Proceedings of the Annual Entomological Society of America**, 1 jan. 2011, p. 16–24. 2011.
- Aqui estão as referências formatadas de acordo com as normas ABNT 2024:
- DWORKIN, I.; JONES, C. D. Genetic changes accompanying the evolution of host specialization in *Drosophila sechellia*. **Genetics**, v. 181, p. 721–736, 2009. DOI: 10.1534/genetics.108.093419.
- EDWARDS, S. V. Phylogenomic subsampling: a brief review. **Zoologica Scripta**, v. 45, p. 63–74, 2016.
- EDWARDS, S. V. et al. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. **Molecular Phylogenetics and Evolution**, v. 94, p. 447–462, 2016. DOI: <https://doi.org/10.1016/j.ympev.2015.10.027>.
- EGAN, S. P. et al. Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. **Ecology Letters**, v. 18, p. 817–825, 2015. DOI: 10.1111/ele.12460.

- FEDER, J. L. et al. Genome-Wide Congealing and Rapid Transitions across the Speciation Continuum during Speciation with Gene Flow. **Journal of Heredity**, v. 105, p. 810–820, 2014. DOI: 10.1093/jhered/esu038.
- FERNÁNDEZ, R.; EDGECOMBE, G. D.; GIRIBET, G. Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. **Systematic Biology**, v. 65, p. 871–889, 2016.
- HARRISON, R. G.; LARSON, E. L. Hybridization, introgression, and the nature of species boundaries. **Journal of Heredity**, v. 105, p. 795–809, 2014. DOI: 10.1093/jhered/esu033.
- HE, R. et al. Phylogenomic analysis and molecular identification of true fruit flies. **Frontiers in Genetics**, v. 15, 1414074, 2024. DOI: 10.3389/fgene.2024.1414074.
- HEDRICK, P. W. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. **Molecular Ecology**, v. 22, 2013.
- HERNÁNDEZ-ORTIZ, V. et al. Taxonomy and phenotypic relationships of the *Anastrepha fraterculus* complex in the Mesoamerican and Pacific Neotropical dominions (Diptera, Tephritidae). **ZooKeys**, v. 540, p. 95–124, 2015. DOI: 10.3897/zookeys.540.6027.
- HERNÁNDEZ-ORTIZ, V. et al. Cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae): A multivariate approach for the recognition of South American morphotypes. **Annals of the Entomological Society of America**, v. 105, p. 305–318, 2012. DOI: 10.1603/AN11123.
- HIBBINS, M. S.; HAHN, M. W. Phylogenomic approaches to detecting and characterizing introgression. **Genetics**, v. 220, 2021. DOI: 10.1093/genetics/iyab173.
- HUERTA-CEPAS, J.; SERRA, F.; BORK, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. **Molecular Biology and Evolution**, v. 33, p. 1635–1638, 2016. DOI: 10.1093/molbev/msw046.
- JAY, P. et al. Supergene evolution triggered by the introgression of a chromosomal inversion. **Current Biology**, v. 28, p. 1839–1845.e1833, 2018. DOI: 10.1016/j.cub.2018.04.072.
- KIMURA, M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. **Nature**, v. 267, p. 275–276, 1977. DOI: 10.1038/267275a0.
- KOBERT, K. et al. Computing the internode certainty and related measures from partial gene trees. **Molecular Biology and Evolution**, v. 33, p. 1606–1617, 2016. DOI: 10.1093/molbev/msw040.
- KOONIN, E. V. Orthologs, paralogs, and evolutionary genomics. **Annual Review of Genetics**, v. 39, p. 309–338, 2005. DOI: 10.1146/annurev.genet.39.073003.114725.

- KOZAK, K. M. et al. Rampant genome-wide admixture across the *Heliconius* radiation. **Genome Biology and Evolution**, v. 13, 2021. DOI: 10.1093/gbe/evab099.
- KOZLOV, A. M. et al. RAXML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. **Bioinformatics**, v. 35, p. 4453–4455, 2019. DOI: 10.1093/bioinformatics/btz305.
- KRIAZHIMSKIY, S.; PLOTKIN, J. B. The population genetics of dN/dS. *PLoS Genetics*, v. 4, e1000304–e1000304, 2008.
- LAFOND, M.; MEGHDARI MIARDAN, M.; SANKOFF, D. Accurate prediction of orthologs in the presence of divergence after duplication. **Bioinformatics**, v. 34, p. i366–i375, 2018. DOI: 10.1093/bioinformatics/bty242.
- LANG, M. et al. Mutations in the neverland gene turned *Drosophila pachea* into an obligate specialist species. **Science**, v. 337, p. 1658–1661, 2012. DOI: 10.1126/science.1224829.
- LEMOINE, F.; GASCUEL, O. Gotree/Goalign: toolkit and Go API to facilitate the development of phylogenetic workflows. **NAR Genomics and Bioinformatics**, v. 3, 2021. DOI: 10.1093/nargab/lqab075.
- LI, X. et al. Phylogenomics reveals accelerated late Cretaceous diversification of bee flies (Diptera: Bombyliidae). **Cladistics**, v. 37, p. 276–297, 2021.
- MALINSKY, M.; MATSCHINER, M.; SVARDAL, H. Dsuite - Fast D-statistics and related admixture evidence from VCF files. **Molecular Ecology Resources**, v. 21, p. 584–595, 2021. DOI: <https://doi.org/10.1111/1755-0998.13265>.
- MINH, B. Q.; HAHN, M. W.; LANFEAR, R. New methods to calculate concordance factors for phylogenomic datasets. **Molecular Biology and Evolution**, v. 37, p. 2727–2733, 2020. DOI: 10.1093/molbev/msaa106.
- MINH, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. **Molecular Biology and Evolution**, v. 37, p. 1530–1534, 2020. DOI: 10.1093/molbev/msaa015.
- MIRABAR, S.; NAKHLEH, L.; WARNOW, T. Multispecies coalescent: Theory and applications in phylogenetics. **Annual Review of Ecology, Evolution, and Systematics**, v. 52, p. 247–268, 2021. DOI: 10.1146/annurev-ecolsys-012121-095340.
- MO, Y. K.; LANFEAR, R.; HAHN, M. W.; MINH, B. Q. Updated site concordance factors minimize effects of homoplasy and taxon sampling. **Bioinformatics**, v. 39, 2022. DOI: 10.1093/bioinformatics/btac741.

- MOLLOY, E. K.; WARNOW, T. To include or not to include: The impact of gene filtering on species tree estimation methods. **Systematic Biology**, v. 67, p. 285–303, 2018. DOI: 10.1093/sysbio/syx077.
- MONGIARDINO KOCH, N. Phylogenomic subsampling and the search for phylogenetically reliable loci. **Molecular Biology and Evolution**, v. 38, p. 4025–4038, 2021. DOI: 10.1093/molbev/msab151.
- MORAN, B. M.; PAYNE, C.; LANGDON, Q.; POWELL, D. L.; BRANDVAIN, Y.; SCHUMER, M. The genomic consequences of hybridization. **eLife**, v. 10, e69016, 2021. DOI: 10.7554/eLife.69016.
- MURRELL, B. et al. Gene-wide identification of episodic selection. **Molecular Biology and Evolution**, v. 32, p. 1365–1371, 2015. DOI: 10.1093/molbev/msv035.
- NEI, M.; KUMAR, S. **Molecular Evolution and Phylogenetics**. Oxford University Press: New York, 2000.
- NOGUERALE, V.; ORTEGO, J. Genomic evidence of speciation by fusion in a recent radiation of grasshoppers. **Evolution**, v. 76, p. 2618–2633, 2022. DOI: <https://doi.org/10.1111/evo.14508>.
- NORRBOM, A. L.; KORYTKOWSKI, C. A.; ZUCCHI, R. A.; URAMOTO, K.; VENABLE, G. L.; MCCORMICK, J.; DALLWITZ, M. **Anastrepha and Toxotrypana: descriptions, illustrations, and interactive keys**. Available online: <http://delta-intkey.com/> (accessed on February, 24, 2018).
- NORRBOM, A. L.; ZUCCHI, R. A.; HERNÁNDEZ-ORTÍZ, V. Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based on morphology. In: ALUJA, M.; NORRBOM, A. L., Eds. **Fruit flies (Tephritidae): Phylogeny and evolution of behavior**. CRC Press: Boca Raton, 2000. p. 299–342.
- PATTERSON, N. et al. Ancient admixture in human history. **Genetics**, v. 192, p. 1065–1093, 2012. DOI: 10.1534/genetics.112.145037.
- PFEIFER, B.; KAPAN, D. D. Estimates of introgression as a function of pairwise distances. **BMC Bioinformatics**, v. 20, p. 207, 2019. DOI: 10.1186/s12859-019-2747-z.
- PHILIPPE, H. et al. Resolving difficult phylogenetic questions: Why more sequences are not enough. **PLoS Biology**, v. 9, e1000602, 2011. DOI: 10.1371/journal.pbio.1000602.
- PHILLIPS, M. J.; PENNY, D. The root of the mammalian tree inferred from whole mitochondrial genomes. **Molecular Phylogenetics and Evolution**, v. 28, p. 171–185, 2003. DOI: [https://doi.org/10.1016/S1055-7903\(03\)00057-5](https://doi.org/10.1016/S1055-7903(03)00057-5).

- PREZOTTO, L. F.; PERONDINI, A. L. P.; HERNÁNDEZ-ORTÍZ, V.; FRÍAS, D.; SELIVON, D. What can integrated analysis of morphological and genetic data still reveal about the *Anastrepha fraterculus* (Diptera: Tephritidae) cryptic species complex? **Insects**, v. 10, p. 408, 2019. DOI: 10.3390/insects10110408.
- R_CORE_TEAM. R: **A language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, Austria, 2023.
- REVELL, L. J. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). **PeerJ**, v. 12, e16505, 2024. DOI: 10.7717/peerj.16505.
- RODRIGUEZ, E. J. et al. Exceptional larval morphology of nine species of the *Anastrepha mucronota* species group (Diptera, Tephritidae). **ZooKeys**, v. 1127, 2022. DOI: 10.3897/zookeys.1127.84628.
- RODRIGUEZ, E. J. et al. Description of larvae of *Anastrepha amplidentata* and *Anastrepha durantae* with review of larval morphology of the *fraterculus* group (Diptera: Tephritidae). **Proceedings of the Entomological Society of Washington**, v. 123, p. 169–189, 2021.
- ROSSER, N. et al. Hybrid speciation driven by multilocus introgression of ecological traits. **Nature**, v. 628, p. 811–817, 2024. DOI: 10.1038/s41586-024-07263-w.
- SAYYARI, E.; MIRARAB, S. Fast coalescent-based computation of local branch support from quartet frequencies. **Molecular Biology and Evolution**, v. 33, p. 1654–1668, 2016. DOI: 10.1093/molbev/msw079.
- SCALLY, M. et al. Resolution of inter and intra-species relationships of the West Indian fruit fly *Anastrepha obliqua*. **Molecular Phylogenetics and Evolution**, v. 101, p. 286–293, 2016. DOI: <https://doi.org/10.1016/j.ympev.2016.04.020>.
- SCHMELZ, E. A. et al. An amino acid substitution inhibits specialist herbivore production of an antagonist effector and recovers insect-induced plant defenses. **Plant Physiology**, v. 160, p. 1468–1478, 2012. DOI: 10.1104/pp.112.201061.
- SELIVON, D. et al. Genetical, morphological, behavioral, and ecological traits support the existence of three Brazilian species of the *Anastrepha fraterculus* complex of cryptic species. **Frontiers in Ecology and Evolution**, v. 10, 2022. DOI: 10.3389/fevo.2022.836608.
- SHARMA, P. P. et al. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. **Molecular Biology and Evolution**, v. 31, p. 2963–2984, 2014.

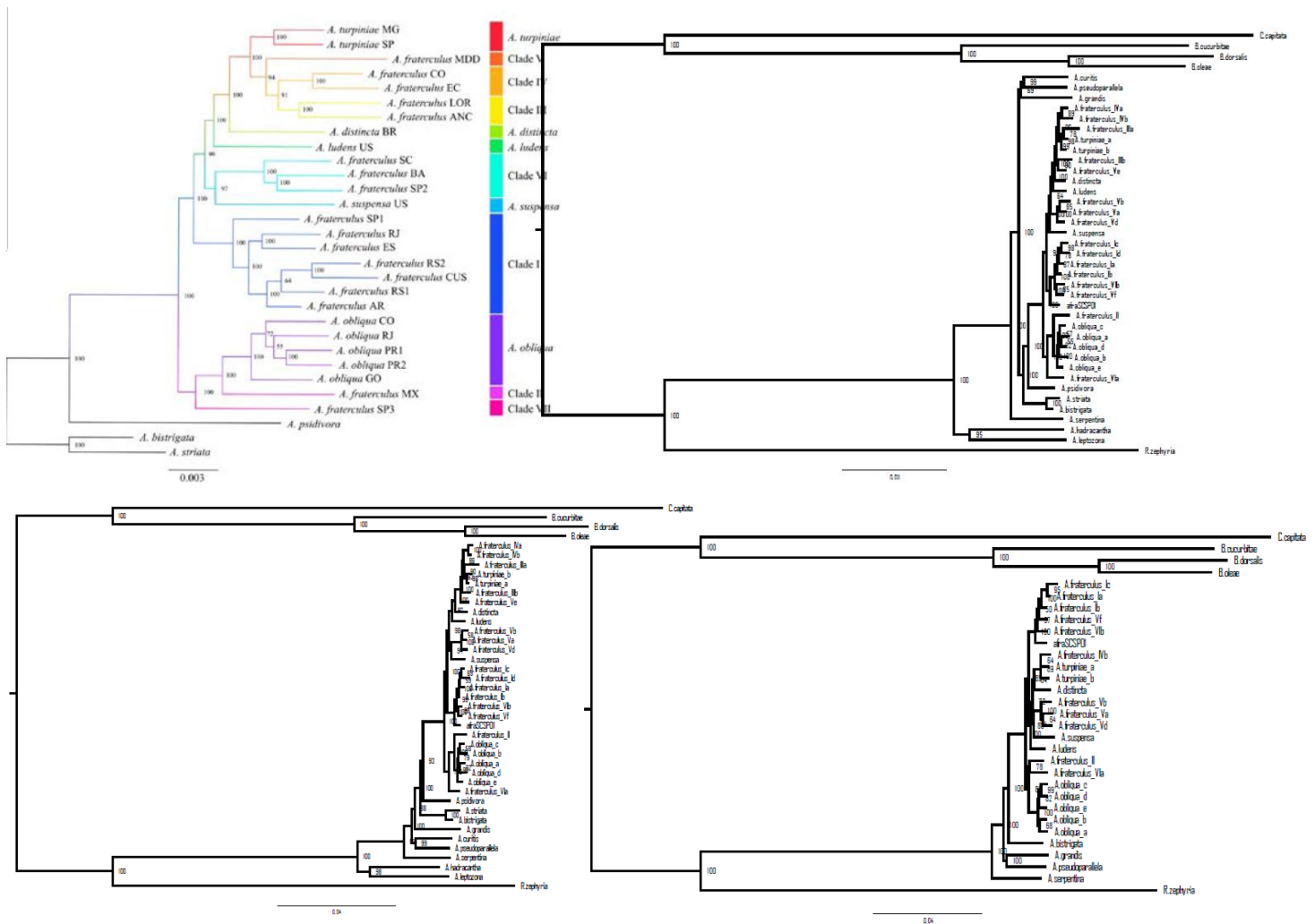
- SHI, Y. et al. Single amino acid variations drive functional divergence of cytochrome P450s in *Helicoverpa* species. **Insect Biochemistry and Molecular Biology**, v. 146, p. 103796, 2022. DOI: <https://doi.org/10.1016/j.ibmb.2022.103796>.
- SHIMODAIRA, H. An approximately unbiased test of phylogenetic tree selection. **Systematic Biology**, v. 51, p. 492–508, 2002. DOI: 10.1080/10635150290069913.
- SHIMODAIRA, H.; HASEGAWA, M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. **Molecular Biology and Evolution**, v. 16, p. 1114, 1999. DOI: 10.1093/oxfordjournals.molbev.a026201.
- SILVA, J. G.; BARR, N. B. Recent advances in molecular systematics of *Anastrepha* Schiner. In: SUGAYAMA, R.L.; ZUCCHI, R.A.; OVRUSKI, S.M.; SIVINSKI, J., eds. **Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance**, 10–15 Sept. 2006, Salvador, Brazil. Salvador-BA, Brazil: Color Press, 2008. p. 13–28.
- SLATER, G. S. C.; BIRNEY, E. Automated generation of heuristics for biological sequence comparison. **BMC Bioinformatics**, v. 6, p. 31, 2005. DOI: 10.1186/1471-2105-6-31.
- SMITH-CALDAS, M. R. B.; MCPHERON, B. A.; SILVA, J. G.; ZUCCHI, R. A. Phylogenetic relationship among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I gene. **Neotropical Entomology**, v. 30, p. 565–573, 2001.
- SMITH, S. A.; BROWN, J. W.; WALKER, J. F. So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. **PLoS ONE**, v. 13, e0197433, 2018. DOI: 10.1371/journal.pone.0197433.
- SMITH, S. A.; MOORE, M. J.; BROWN, J. W.; YANG, Y. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. **BMC Evolutionary Biology**, v. 15, p. 150, 2015. DOI: 10.1186/s12862-015-0423-0.
- SOBRINHO, I. S.; DE BRITO, R. A. Positive and purifying selection influence the evolution of *Doublesex* in the *Anastrepha fraterculus* species group. **PLoS ONE**, v. 7, e33446, 2012. DOI: 10.1371/journal.pone.0033446.
- STEENWYK, J. L. et al. PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. **Bioinformatics**, v. 37, p. 2325–2331, 2021. DOI: 10.1093/bioinformatics/btab096.
- STEENWYK, J. L. et al. Incongruence in the phylogenomics era. **Nature Reviews Genetics**, 2023. DOI: 10.1038/s41576-023-00620-x.

- STRIMMER, K.; RAMBAUT, A. Inferring confidence sets of possibly misspecified gene trees. **Proceedings of the Royal Society of London B**, v. 269, p. 137–142, 2002. DOI: 10.1098/rspb.2001.1862.
- SUTTON, B. D. et al. Nuclear ribosomal internal transcribed spacer 1 (ITS1) variation in the *Anastrepha fraterculus* cryptic species complex (Diptera, Tephritidae) of the Andean region. **ZooKeys**, p. 175–191, 2015. DOI: 10.3897/zookeys.540.6147.
- SUVOROV, A. et al. Widespread introgression across a phylogeny of 155 *Drosophila* genomes. **Current Biology**, v. 32, p. 111–123.e115, 2022. DOI: <https://doi.org/10.1016/j.cub.2021.10.052>.
- SUZUKI, Y.; GOJOBORI, T. A method for detecting positive selection at single amino acid sites. **Molecular Biology and Evolution**, v. 16, p. 1315–1328, 1999.
- SWANSON, W. J. et al. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. **Genetics**, v. 168, p. 1457–1465, 2004.
- TALEVICH, E. et al. Bio.Phylo: a unified toolkit for processing, analyzing and visualizing phylogenetic trees in Biopython. **BMC Bioinformatics**, v. 13, p. 209, 2012. DOI: 10.1186/1471-2105-13-209.
- TELFORD, M. J. et al. Phylogenomic analysis of echinoderm class relationships supports Asterozoa. **Proceedings of the Royal Society of London B**, v. 281, p. 20140479, 2014. DOI: 10.1098/rspb.2014.0479.
- THAWORNWATTANA, Y. et al. Full-Likelihood Genomic Analysis Clarifies a Complex History of Species Divergence and Introgression: The Example of the erato-sara Group of *Heliconius* Butterflies. **Systematic Biology**, 2022. DOI: 10.1093/sysbio/syac009.
- TOWNSEND, J. P.; SU, Z.; TEKLE, Y. I. Phylogenetic signal and noise: predicting the power of a data set to resolve phylogeny. **Systematic Biology**, v. 61, p. 835, 2012.
- WEAVER, S. et al. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. **Molecular Biology and Evolution**, v. 35, p. 773–777, 2018. DOI: 10.1093/molbev/msx335.
- YANG, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. **Bioinformatics**, v. 13, p. 555–556, 1997.
- YANG, Z.; BIELAWSKI, J. P. Statistical methods for detecting molecular adaptation. **Trends in Ecology & Evolution**, v. 15, p. 496–503, 2000.

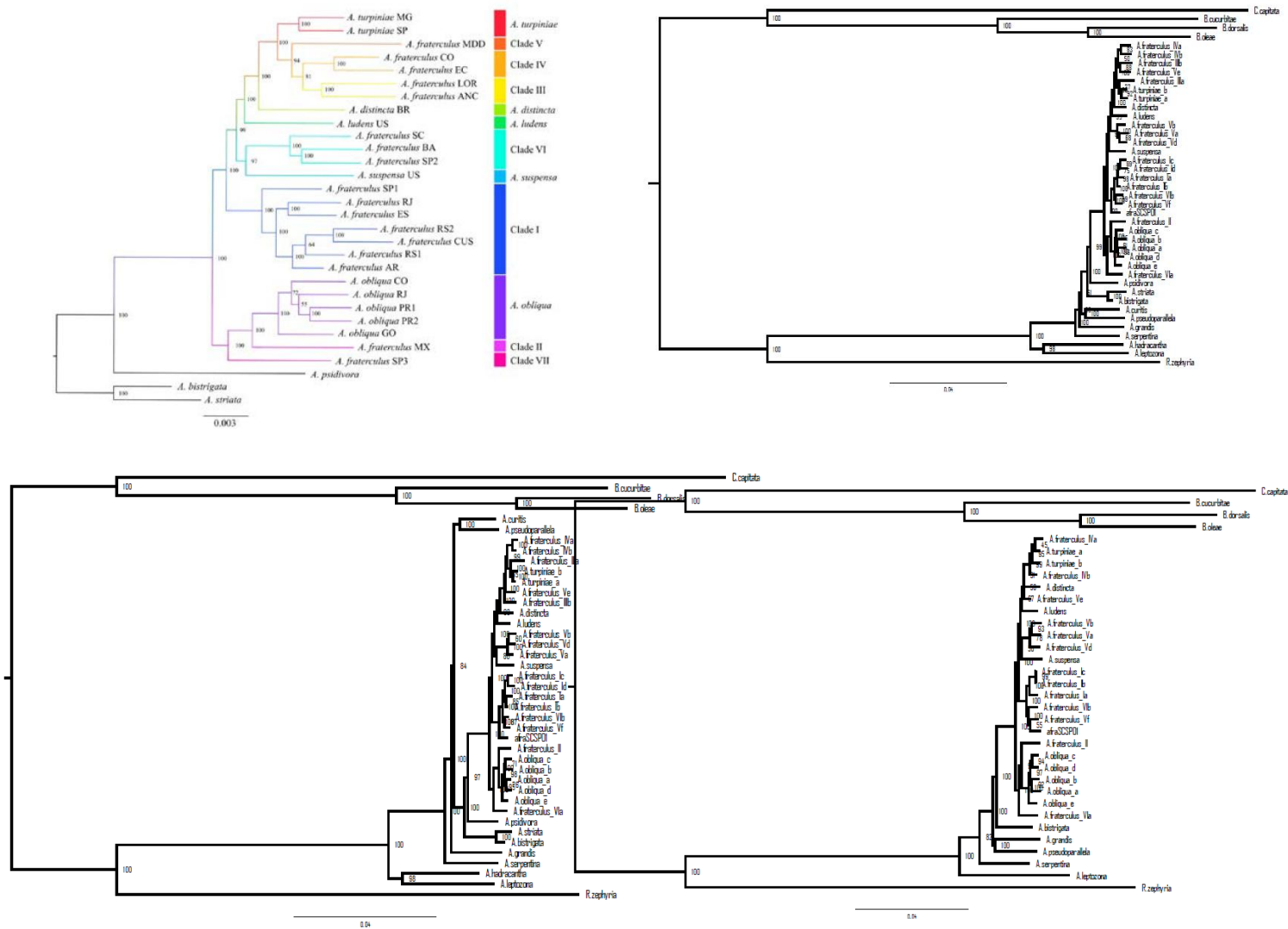
- ZHANG, C. et al. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. **BMC Bioinformatics**, v. 19, p. 153, 2018. DOI: 10.1186/s12859-018-2129-y.
- ZUCCHI, R. A. Espécies de *Anastrepha*, sinónímias, plantas hospedeiras e parasitóides. In: MALAVASI, A.; ZUCCHI, R. A. (Eds.). **Moscas-das-frutas de importância econômica no Brasil: Conhecimento básico e aplicado**. Ribeirão Preto: Holos Editora, 2000. p. 41–48.
- ZUCCHI, R. A. Moscas-das-frutas (Diptera: Tephritidae) no Brasil: Taxonomia, distribuição geográfica e hospedeiros. In: SOUZA, H. M. L. (Ed.). **Moscas-das-frutas no Brasil**. Campinas, SP: Anais Fundação Cargil, 1988. p. 1–10.

MATERIAL SUPPLEMENTAR

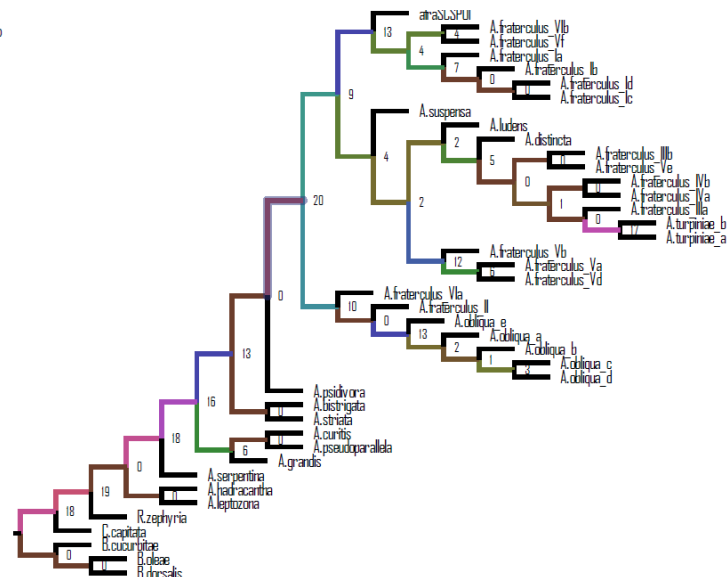
Supplementary Figures



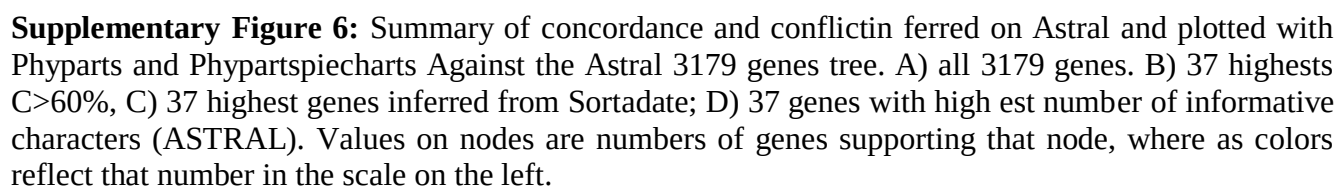
Supplementary Figure1: iqtree inference produced from diferente concatenated sets of genes. A) all3 179 genes. B) 20 highests C>60%, C) 20 highest genes inferred from Sortadate; D) 20 genes with highest number of informative characters (ASTRAL).

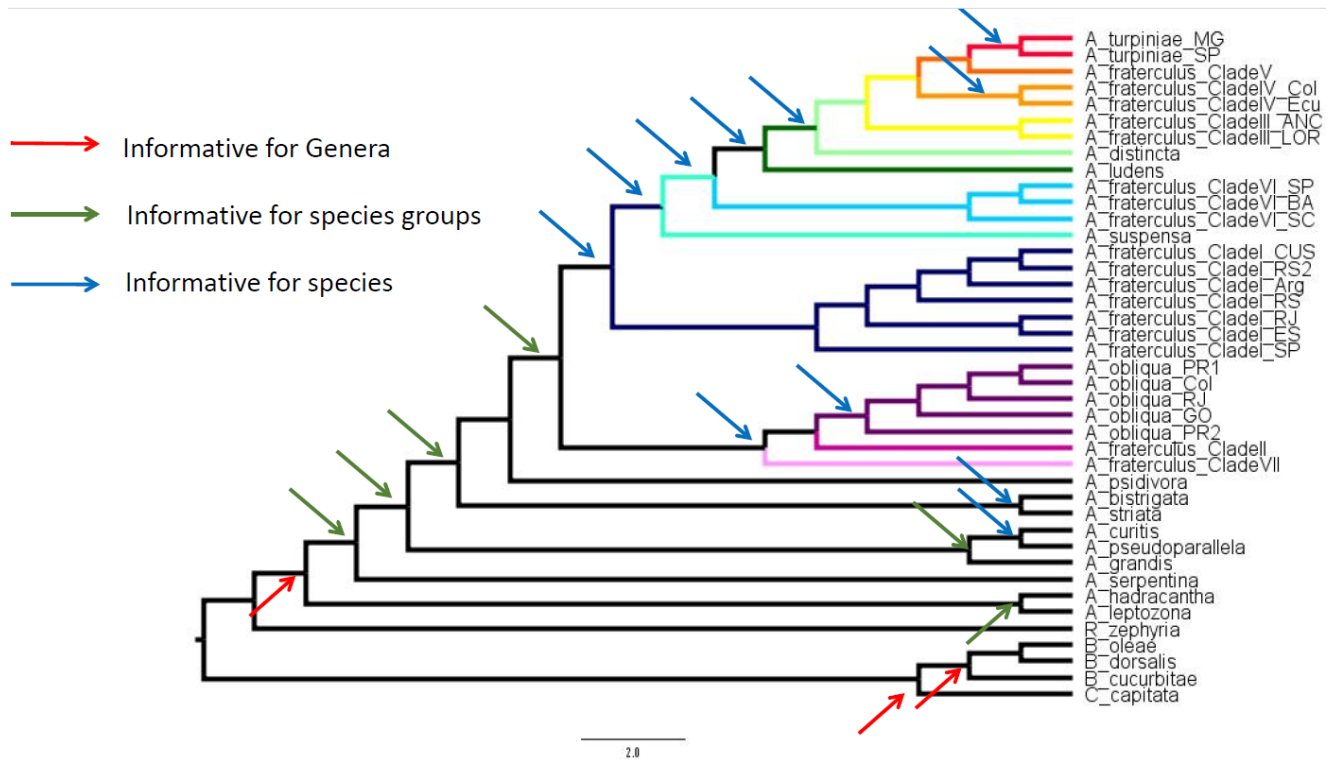


Supplementary Figure 2: iqtree inference produced from diferente concatenated sets of genes. A) all 3179 genes. B) 37 highests C>60%, C) 37 highest genes inferred from Sortadate; D) 37 genes with highest number of informative characters (ASTRAL).



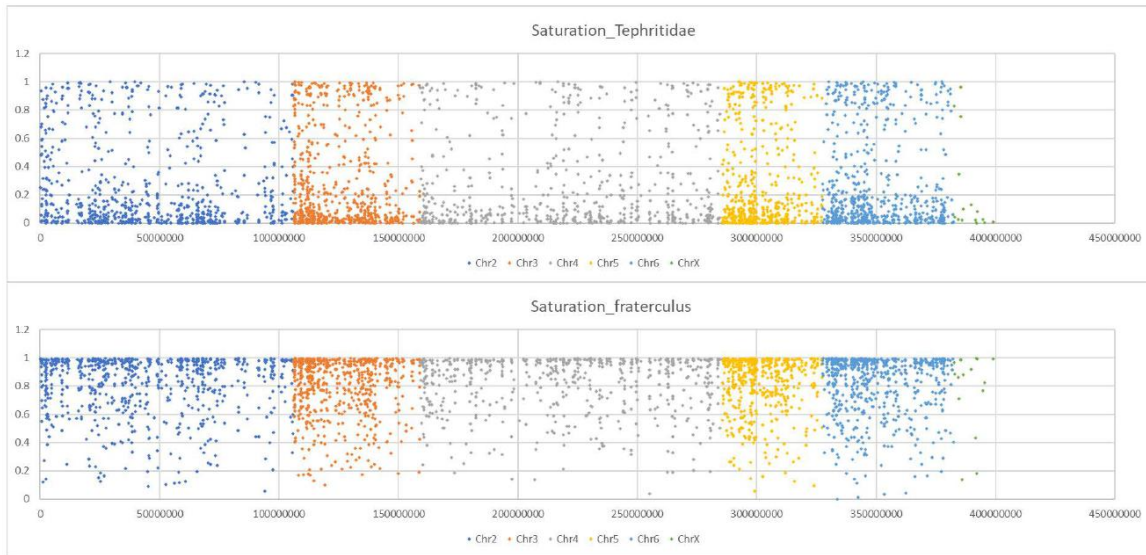
Supplementary Figure 5: Summary of concordance and conflict inferred on Astral and plotted with Phyparts and Phypartspiecharts Against the Astral 3179 genes tree. A) all 3179 genes. B) 20 highest C>60%, C) 20 highest genes inferred from Sortadata; D) 20 genes with highest number of informative characters (ASTRAL). Values on nodes are number sof genes supporting that node, where as colors reflect that number in the scale on the left.





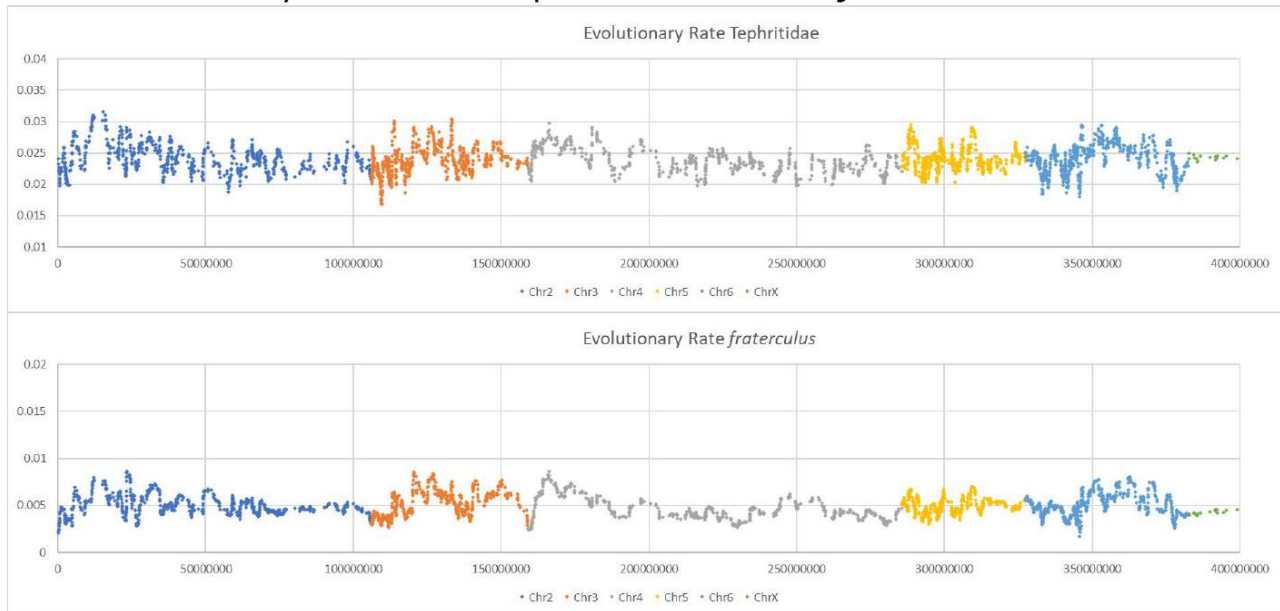
Suppl. Figure 4. Species tree inferred from the *Anastrepha* dataset indicating which clades were considered as surrogate indicators of phylogenetic information at different hierarchical levels. Red arrows indicate nodes that separate genera, Green arrows indicate nodes that separate species groups, whereas blue arrows indicate nodes that separate species.

Saturation in Tephritidae and *fraterculus* datasets



Suppl. Figure 6: Distribution of phylogenetic saturation inferred from the *fraterculus* and *Anastrepha* data sets across the genome, using *A.ludens* genome as reference.

Evolutionary Rate in Tephritidae and *fraterculus* datasets



Suppl. Figure 7: Distribution of evolutionary rates inferred from the *fraterculus* and *Anastrepha* datasets across the genome, using *A.ludens* genome as reference.

Suppl. Table 1

Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadat e 20	sortadat e 37	info 20	info 37
cluster1030_1	Chr3	686940	691471	FBgn0029863	Tocopherol Transfer Protein (Loc108363593)	X	X				
cluster1762_1	Chr4	1764966	1760656	FBgn0086254	Aldose Reductase-Like	X	X				
cluster1852_1	ChrX	9196598	9139777	FBgn0067102	Ceramide Glucosyltransferase	X	X				
cluster1873_1	Chr2	1829265	1830042	FBgn0015221	Soma Ferritin-Like	X	X				
cluster1967_1	Chr2	28650806	28644719	FBgn0020385	C-1-Tetrahydrofolate Cytoplasmic Isoform X2	X	X				
cluster2163_1	Chr3	667850	668331	FBgn0004828	Histone Partial	X	X				
cluster2407_1	Chr5	13346539	13349222	FBgn0032635	Predicted: Putative E3 Ubiquitin-Protein Ligase Ubr7 Isoform X1 [Rhagoletis Zephyria]	X	X				
cluster3068_1	Chr5	38488847	38489970	FBgn0032451	Magnesium Transporter Nipa2	X	X				
cluster3083_1	Chr3	26517786	26520203	FBgn0030151	Obg-Like Atpase 1	X	X				
cluster3234_1	Chr2	469774	468144	FBgn0038683	Mitochondrial Import Inner Membrane Translocase Subunit Tim44	X	X				
cluster3296_1	Chr5	11125756	11124348	FBgn0016698	Serine Threonine-Protein Phosphatase 2A Activator	X	X				
cluster3381_1	Chr3	6241812	6243287	FBgn0031060	Probable Trna N6-Adenosine Mitochondrial	X	X				
cluster343_1	Chr3	6646128	6725971	FBgn0264270	Sex-Lethal Homolog (Loc105208616)	X	X				

cluster3737_1	Chr3	375019	376548	FBgn0030468	Pre-Mrna-Splicing Factor 38B Isoform X1	X	X		X		
cluster3741_1	Chr3	2167022	2156818	FBgn0028360	Cell Division Cycle 7-Related Protein Kinase	X	X				
cluster3962_1	Chr5	32308435	32306676	FBgn0264672	Egf Domain-Specific O-Linked N-Acetylglucosamine Transferase	X	X				
cluster4000_1	Chr3	297131	298501	FBgn0030396	Ankyrin Repeat-Containing Protein Kinase A (Loc108366792)	X	X				
cluster4106_1	Chr2	5424748	5426279	FBgn0014366	Splicing Factor 3A Subunit 3	X	X				
cluster4936_1	Chr6	13819668	13820535	FBgn0035035	Probable Methyltransferase Btm2 Homolog	X	X				
cluster602_1	Chr3	44303016	44276974	FBgn0030052	Uncharacterized Loc118733309	X	X				
Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadat e 20	sortadat e 37	info 20	info 37
cluster2142_1	Chr4	83426328	83426973	FBgn0024814	Clathrin Light Chain		X				
cluster2173_1	Chr4	119392949	119393666	FBgn0029094	Histone Chaperone Asf1		X				
cluster2243_1	Chr2	37023822	37021321	FBgn0003882	Predicted: Low Quality Protein: Protein Tube-Like [Rhagoletis Zephyria]		X				
cluster2539_1	Chr2	4887960	4889548	FBgn0037279	Mannose-1-Phosphate Guanyltransferase Beta		X		X		
cluster2550_1	Chr2	2701898	2704018	FBgn0038763	N-Acetylglucosaminyl-Phosphatidylinositol De-N-Acetylase		X				

cluster2608_1	Chr2	39851067	39852127	FBgn0029980	Dehydrololichyl Diphosphate Synthase Complex Subunit Dhdds	X					
cluster2913_1	Chr4	66358863	66361286	FBgn0035981	Predicted: Protein Fam76A [Bactrocera Dorsalis]	X					
cluster2940_1	Chr4	114574709	114573202	FBgn0000261	Catalase	X					
cluster2971_1	Chr2	4347953	4347026	FBgn0039625	Beta- -N-Acetylgalactosaminyltransferase Bre-4	X					
cluster3007_1	Chr5	7564512	7563616	FBgn0002031	Protein L 37Cc	X					
cluster3095_1	Chr5	13784083	13787054	FBgn0001202	Protein Hook	X					
cluster3167_1	Chr2	2882849	2884759	FBgn0037718	Predicted: Dnaj Homolog Subfamily C Member 3 [Bactrocera Cucurbitae]	X					
cluster3179_1	Chr6	17560804	17559336	FBgn0010339	Gtp-Binding Protein 128Up	X					
cluster3242_1	Chr3	373433	371063	FBgn0003079	Raf Homolog Serine Threonine-Protein Kinase Phl	X					
cluster3367_1	Chr3	17792647	17798725	FBgn0024985	Rilp-Like Protein Homolog	X					
cluster3402_1	Chr2	22874931	22876490	FBgn0037329	Polymerase Delta-Interacting Protein 2	X			X		
cluster3414_1	Chr6	6866765	6868065	FBgn0033891	Nadh Dehydrogenase	X					
Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadate 20	sortadate 37	info 20	info 37
cluster1066_1	Chr2	39728752	39730487	FBgn0039186	Fha Domain-Containing Protein Fhaa (Loc118738518)			X	X		

cluster1369_1	Chr6	18679619	18676336	FBgn0014029	Septin-2	X	X
cluster1406_1	Chr4	125702690	125698731	FBgn0036504	Uncharacterized Protein Loc108377800	X	X
cluster1802_1	Chr6	4263840	4265670	FBgn0034703	Trna Pseudouridine(38 39) Synthase	X	X
cluster1812_1	Chr5	13228025	13226587	FBgn0031309	General Transcription Factor Iih Subunit 3	X	X
cluster2199_1	Chr4	4367495	4366185	FBgn0035471	Very-Long-Chain Enoyl- Reductase	X	X
cluster2395_1	Chr6	8019001	8018014	FBgn0039664	Predicted: Dtw Domain- Containing Protein 1 [Bactrocera Cucurbitae]	X	X
cluster2441_1	Chr2	9215716	9219831	FBgn0038312	Zinc Transporter Zip1 Isoform X2	X	X
cluster2507_1	Chr5	18669854	18639286	FBgn0027052	E3 Ubiquitin-Protein Ligase Chip	X	X
cluster3012_1	Chr6	38701889	38704370	FBgn0016687	Inorganic Pyrophosphatase	X	X
cluster3028_1	Chr4	66347257	66348963	FBgn0085484	Pyridoxal Kinase	X	X
cluster3084_1	Chr6	18859228	18860302	FBgn0034921	Mrna-Decapping Enzyme 1B	X	X
cluster3268_1	Chr3	15387275	15386459	FBgn0039904	Predicted: Host Cell Factor 2- Like [Bactrocera Oleae]	X	X
cluster3298_1	Chr2	11318196	11317112	FBgn0038772	Ring Finger Protein 113A	X	X
cluster3598_1	Chr5	27402172	27400782	FBgn0032016	Methyltransferase-Like Protein 14 Homolog	X	X
cluster3644_1	Chr3	4814006	4817601	FBgn0036262	Inositol Oxygenase	X	X

cluster3673_1	Chr2	11665328	11666276	FBgn0037342	Rna-Binding Protein 42			X	X		
cluster3963_1	Chr3	2061063	2062523	FBgn0029887	Luc7-Like Protein 3			X	X		
cluster4180_1	Chr3	6895424	6897671	FBgn0011710	Septin-1			X	X		
cluster5109_1	Chr2	2807890	2791460	FBgn0037715	Mitochondrial Pyruvate Carrier 2			X	X		
Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadate 20	sortadate 37	info 20	info 37
cluster1055_1	Chr6	43209577	43162106	FBgn0043070	Protein Ndr3 (Loc108377860) Predicted: Evolutionarily Conserved Signaling Intermediate In Toll Pathway, Mitochondrial [Bactrocera Cucurbitae]				X		
cluster2160_1	Chr2	2708200	2709456	FBgn0028436	Tetraspanin-33				X		
cluster2673_1	Chr2	9392418	9391571	FBgn0037848	28 Kda Heat- And Acid-Stable Phosphoprotein [Rhagoletis Zephyria]				X		
cluster2935_1	Chr3	6567121	6567688	FBgn0029715	B-Cell Receptor-Associated Protein 31				X		
cluster3389_1	Chr4	14805539	14807556	FBgn0035165	Ctd Nuclear Envelope Phosphatase 1 Homolog Isoform X1				X		
cluster3392_1	Chr3	23094827	23093661	FBgn0029067	Predicted: Atp-Dependent Rna Helicase Dbp4-Like [Bactrocera Latifrons]				X		
cluster3595_1	Chr4	12096522 1	12096292 2	NA	Decaprenyl-Diphosphate Synthase Subunit 2				X		
cluster3735_1	Chr4	48713688	48715106	FBgn0037044					X		

cluster4500_1	Chr3	5826511	5827998	FBgn0030177	Rna-Binding Protein Nob1				X		
cluster4574_1	Chr2	26676353	26674189	FBgn0037657	Parafibromin				X		
cluster4769_1	Chr6	18982750	18985376	FBgn0023477	Probable Transaldolase				X		
cluster5077_1	Chr6	49909558	49908268	FBgn0046692	Serine Threonine-Protein Kinase Ste20-Like				X		
cluster5207_1	Chr4	48682634	48683996	FBgn0011573	Hsp90 Co-Chaperone Cdc37				X		
cluster924_1	Chr5	6831549	6804216	FBgn0031948	Ethanolaminephosphotransferase 1 (Loc108358103)				X		
Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadate 20	sortadate 37	info 20	info 37
cluster1344_1	Chr5	4743751	4782829	FBgn0032208	E3 Ubiquitin-Protein Ligase Hectd1					X	X
cluster1676_1	ChrX	1864864	1706494	FBgn0039936	Gigyf Family Protein Cg11148 (Loc105220315)					X	X
cluster1794_1	Chr3	31911935	31901693	FBgn0250788	Spectrin Beta Chain Isoform X2					X	X
cluster1944_1	Chr2	27467487	27435604	FBgn0002431	E3 Ubiquitin-Protein Ligase Hyd Isoform X2					X	X
cluster2084_1	Chr6	34758508	34745753	FBgn0033688	Pre-Mrna-Processing-Splicing Factor 8					X	X
cluster2086_1	Chr2	6412209	6420639	FBgn0039120	Nuclear Pore Complex Protein Nup98-Nup96					X	X
cluster2822_1	Chr3	2043017	2052352	FBgn0087008	Supporter Of Activation Of Yellow Protein (Loc101461521)					X	X
cluster4288_1	Chr5	20275239	20269275	FBgn0016075	Collagen Alpha-1 Chain					X	X

cluster4553_1	Chr6	17880118	17872422	FBgn0013756	Nucleoprotein Tpr					X	X
cluster4599_1	-	-	-	-	-					X	X
cluster4643_1	Chr5	5770582	5763759	FBgn0011722	Predicted: Tiggrin [Rhagoletis Zephyria]					X	X
cluster4666_1	Chr6	573499	566947	FBgn0038968	Predicted: Uncharacterized Protein Loc108361800 Isoform X1 [Rhagoletis Zephyria]					X	X
cluster4686_1	Chr5	20360078	20365789	FBgn0000299	Collagen Alpha-1 Chain					X	X
cluster4697_1	Chr3	12963445	12935174	FBgn0083956	Atp-Binding Cassette Sub-Family A Member 13					X	X
cluster4746_1	Chr5	22886925	22882864	FBgn0086451	Predicted: Heat Repeat-Containing Protein 1 Homolog [Rhagoletis Zephyria]					X	X
cluster4789_1	Chr6	12228916	12218827	FBgn0015477	Dnaj Homolog Subfamily C Member 13 Isoform X1					X	X
cluster5165_1	Chr4	71080789	71033086	FBgn0036663	Glutamate Synthase					X	X
cluster5546_1	ChrX	9365285	9355203	FBgn0087002	Apolipophorins					X	X
cluster5592_1	Chr5	23832510	23840772	FBgn0283427	Fatty Acid Synthase					X	X
cluster5718_1	Chr2	2829452	2816307	FBgn0038651	Ectopic P Granules Protein 5 Homolog Isoform X2					X	X
Contig	Chrom	Start	Finish	Flybase	Gene	sgC6 0 20	sgC6 0 37	sortadate 20	sortadate 37	info 20	info 37
cluster1384_1	ChrX	4138851	4225428	FBgn0025741	Plexin-A2 Isoform X1						X
cluster2689_1	Chr6	51276937	51284050	FBgn0028741	1-Phosphatidylinositol 3-Phosphate 5-Kinase						X

cluster2728_1	Chr4	79422453	79437551	FBgn0267348	Laminin Subunit Gamma-1	X
cluster2797_1	Chr2	20714844	20719130	FBgn0037926	Elongator Complex Protein 1	X
cluster2869_1	Chr3	5268976	5277184	FBgn0267350	Phosphatidylinositol 4-Kinase Alpha	X
cluster2875_1	Chr6	16425526	16466844	FBgn0265434	Myosin Heavy Non-Muscle Isoform X2	X
cluster3372_1	Chr2	592646	578147	FBgn0005674	Bifunctional Glutamate Proline--Trna Ligase	X
cluster4626_1	Chr6	5746379	5738059	FBgn0259876	Condensin Complex Subunit 3	X
cluster4696_1	Chr2	70437481	70432040	FBgn0019938	Dna-Directed Rna Polymerase I Subunit Rpa1	X
cluster4725_1	Chr3	42585534	42571675	FBgn0040236	Predicted: Maestro Heat-Like Repeat-Containing Protein Family Member 1 [Bactrocera Cucurbitae]	X
cluster4737_1	Chr5	41642649	41635671	FBgn0267488	Cd109 Antigen	X
cluster5410_1	ChrX	13036251	13126035	FBgn0025740	Plexin-B	X
cluster5535_1	Chr6	49016633	49002886	FBgn0034618	Glycogen Debranching Enzyme	X
cluster5571_1	ChrX	9874652	10102721	FBgn0023213	Eukaryotic Translation Initiation Factor 4 Gamma 3 Isoform X1	X
cluster5632_1	Chr3	7959387	7975355	FBgn0027287	Predicted: Thyroid Receptor-Interacting Protein 11 Isoform X1 [Rhagoletis Zephyria]	X

cluster6202_1	ChrX	2926758	2932077	FBgn0051998	Uncharacterized Loc108363536 (Loc108363536)	X
cluster781_1	Chr4	109214095	109222382	FBgn0262714	Mucin-22-Like (Loc118753884)	X

Set	outgroup	test	Species I	Species II
-----	----------	------	-----------	------------

Suppl. Table 2 – List of set used to estimate introgression per loci. Species I and II are sister species, based on a species tree derived from Figure 1 tree, and test species are used to estimate deviations from ABBA-BABA expectations, considering different outgroups.

str_tur_fratIII	A_striata	A_distincta	A_turpiniae	A_fraterculus_III
	A_striata	A_ludens	A_turpiniae	A_fraterculus_III
	A_striata	A_fraterculus_VI	A_turpiniae	A_fraterculus_III
	A_striata	A_suspensa	A_turpiniae	A_fraterculus_III
	A_striata	A_fraterculus_I	A_turpiniae	A_fraterculus_III
	A_striata	A_obliqua	A_turpiniae	A_fraterculus_III
	A_striata	A_fraterculus_II	A_turpiniae	A_fraterculus_III
	A_striata	A_psidivora	A_turpiniae	A_fraterculus_III
lep_tur_fratIII	A_leptozona	A_distincta	A_turpiniae	A_fraterculus_III
	A_leptozona	A_ludens	A_turpiniae	A_fraterculus_III
	A_leptozona	A_fraterculus_VI	A_turpiniae	A_fraterculus_III
	A_leptozona	A_suspensa	A_turpiniae	A_fraterculus_III
	A_leptozona	A_fraterculus_I	A_turpiniae	A_fraterculus_III
	A_leptozona	A_obliqua	A_turpiniae	A_fraterculus_III
	A_leptozona	A_fraterculus_II	A_turpiniae	A_fraterculus_III
	A_leptozona	A_psidivora	A_turpiniae	A_fraterculus_III
	A_leptozona	A_striata	A_turpiniae	A_fraterculus_III
	A_leptozona	A_pseudoparallela	A_turpiniae	A_fraterculus_III
	A_leptozona	A_grandis	A_turpiniae	A_fraterculus_III
	A_leptozona	A_serpentina	A_turpiniae	A_fraterculus_III
out_tur_fratIII	outgroup	A_distincta	A_turpiniae	A_fraterculus_III
	outgroup	A_ludens	A_turpiniae	A_fraterculus_III
	outgroup	A_fraterculus_VI	A_turpiniae	A_fraterculus_III
	outgroup	A_suspensa	A_turpiniae	A_fraterculus_III
	outgroup	A_fraterculus_I	A_turpiniae	A_fraterculus_III
	outgroup	A_obliqua	A_turpiniae	A_fraterculus_III
	outgroup	A_fraterculus_II	A_turpiniae	A_fraterculus_III
	outgroup	A_psidivora	A_turpiniae	A_fraterculus_III
	outgroup	A_striata	A_turpiniae	A_fraterculus_III
	outgroup	A_pseudoparallela	A_turpiniae	A_fraterculus_III
	outgroup	A_grandis	A_turpiniae	A_fraterculus_III
	outgroup			

outgroup	A_serpentina	A_turpiniae	A_fraterculus_III
outgroup	A_leptozona	A_turpiniae	A_fraterculus_III
outgroup	A_hadracantha	A_turpiniae	A_fraterculus_III

Set	outgroup	test	Species I	Species II
str_fratIV_fratIII	A_striata	A_distincta	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_ludens	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_fraterculus_VI	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_suspensa	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_fraterculus_I	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_obliqua	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_fraterculus_II	A_fraterculus_IV	A_fraterculus_III
	A_striata	A_psidivora	A_fraterculus_IV	A_fraterculus_III
lep_fratIV_fratIII	A_leptozona	A_distincta	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_ludens	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_fraterculus_VI	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_suspensa	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_fraterculus_I	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_obliqua	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_fraterculus_II	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_psidivora	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_striata	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_pseudoparallela	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_grandis	A_fraterculus_IV	A_fraterculus_III
	A_leptozona	A_serpentina	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_distincta	A_fraterculus_IV	A_fraterculus_III
out_fratIV_fratIII	outgroup	A_ludens	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_fraterculus_VI	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_suspensa	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_fraterculus_I	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_fraterculus_II	A_fraterculus_IV	A_fraterculus_III

	outgroup	A_obliqua	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_fraterculus_II	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_psidivora	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_striata	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_pseudoparallela	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_grandis	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_serpentina	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_leptozona	A_fraterculus_IV	A_fraterculus_III
	outgroup	A_hadracantha	A_fraterculus_IV	A_fraterculus_III
str_fratVI_fratIII	A_striata	A_suspensa	A_fraterculus_VI	A_fraterculus_III
	A_striata	A_fraterculus_I	A_fraterculus_VI	A_fraterculus_III
	A_striata	A_obliqua	A_fraterculus_VI	A_fraterculus_III
	A_striata	A_fraterculus_II	A_fraterculus_VI	A_fraterculus_III
	A_striata	A_psidivora	A_fraterculus_VI	A_fraterculus_III

Set	outgroup	test	Species I	Species II
lep_fratVI_fratIII	A_leptozona	A_suspensa	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_fraterculus_I	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_obliqua	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_fraterculus_II	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_psidivora	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_striata	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_pseudoparallela	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_grandis	A_fraterculus_VI	A_fraterculus_III
	A_leptozona	A_serpentina	A_fraterculus_VI	A_fraterculus_III
out_fratVI_fratIII	outgroup	A_suspensa	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_fraterculus_I	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_obliqua	A_fraterculus_VI	A_fraterculus_III

	outgroup	A_fraterculus_II	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_psidivora	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_striata	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_pseudoparallela	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_grandis	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_serpentina	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_leptozona	A_fraterculus_VI	A_fraterculus_III
	outgroup	A_hadracantha	A_fraterculus_VI	A_fraterculus_III
out_fratIII_fratI	outgroup	A_obliqua	A_fraterculus_III	A_fraterculus_I
	outgroup	A_fraterculus_II	A_fraterculus_III	A_fraterculus_I
	outgroup	A_fraterculus_VII	A_fraterculus_III	A_fraterculus_I
	outgroup	A_psidivora	A_fraterculus_III	A_fraterculus_I
	outgroup	A_striata	A_fraterculus_III	A_fraterculus_I
	outgroup	A_pseudoparallela	A_fraterculus_III	A_fraterculus_I
	outgroup	A_grandis	A_fraterculus_III	A_fraterculus_I
	outgroup	A_serpentina	A_fraterculus_III	A_fraterculus_I
	outgroup	A_leptozona	A_fraterculus_III	A_fraterculus_I
	outgroup	A_hadracantha	A_fraterculus_III	A_fraterculus_I
out_fratIV_fratI	outgroup	A_obliqua	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_fraterculus_II	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_fraterculus_VII	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_psidivora	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_striata	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_pseudoparallela	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_grandis	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_serpentina	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_leptozona	A_fraterculus_IV	A_fraterculus_I
	outgroup	A_hadracantha	A_fraterculus_IV	A_fraterculus_I

Set	outgroup	test	Species I	Species II
-----	----------	------	-----------	------------

out_tur_fratIIa	outgroup	A_fraterculus_II	A_turpiniae	A_fraterculus_III
	outgroup	A_fraterculus_VII	A_turpiniae	A_fraterculus_III
out_fratII_fratI	outgroup	A_psidivora	A_fraterculus_II	A_fraterculus_I
	outgroup	A_striata	A_fraterculus_II	A_fraterculus_I
	outgroup	A_pseudoparallela	A_fraterculus_II	A_fraterculus_I
	outgroup	A_grandis	A_fraterculus_II	A_fraterculus_I
	outgroup	A_serpentina	A_fraterculus_II	A_fraterculus_I
	outgroup	A_leptozona	A_fraterculus_II	A_fraterculus_I
	outgroup	A_hadracantha	A_fraterculus_II	A_fraterculus_I
out_fratII_obli	outgroup	A_turpiniae	A_fraterculus_II	A_obliqua
	outgroup	A_fraterculus_III	A_fraterculus_II	A_obliqua
	outgroup	A_fraterculus_IV	A_fraterculus_II	A_obliqua
	outgroup	A_distincta	A_fraterculus_II	A_obliqua
	outgroup	A_ludens	A_fraterculus_II	A_obliqua
	outgroup	A_fraterculus_VI	A_fraterculus_II	A_obliqua
	outgroup	A_psidivora	A_fraterculus_II	A_obliqua
	outgroup	A_striata	A_fraterculus_II	A_obliqua
	outgroup	A_pseudoparallela	A_fraterculus_II	A_obliqua
	outgroup	A_grandis	A_fraterculus_II	A_obliqua
	outgroup	A_serpentina	A_fraterculus_II	A_obliqua
	outgroup	A_leptozona	A_fraterculus_II	A_obliqua
	outgroup	A_hadracantha	A_fraterculus_II	A_obliqua
out_fratVII_obli	outgroup	A_psidivora	A_fraterculus_VII	A_obliqua
	outgroup	A_striata	A_fraterculus_VII	A_obliqua
	outgroup	A_pseudoparallela	A_fraterculus_VII	A_obliqua
	outgroup	A_grandis	A_fraterculus_VII	A_obliqua
	outgroup	A_serpentina	A_fraterculus_VII	A_obliqua
	outgroup	A_leptozona	A_fraterculus_VII	A_obliqua
	outgroup	A_hadracantha	A_fraterculus_VII	A_obliqua
out_fratVI_fratI	outgroup	A_fraterculus_II	A_fraterculus_VI	A_fraterculus_I
	outgroup	A_fraterculus_VII	A_fraterculus_VI	A_fraterculus_I

outgroup	A_obliqua	A_fraterculus_VI	A_fraterculus_I
outgroup	A_psidivora	A_fraterculus_VI	A_fraterculus_I
outgroup	A_striata	A_fraterculus_VI	A_fraterculus_I
outgroup	A_pseudoparallela	A_fraterculus_VI	A_fraterculus_I
outgroup	A_grandis	A_fraterculus_VI	A_fraterculus_I
outgroup	A_serpentina	A_fraterculus_VI	A_fraterculus_I
outgroup	A_leptozona	A_fraterculus_VI	A_fraterculus_I
outgroup	A_hadracantha	A_fraterculus_VI	A_fraterculus_I
