UFRGS

Evolução do DNA mitocondrial em *Bombus* (Hymenoptera, Apidae): de *barcodes* a genomas completos

Leonardo Tresoldi Gonçalves

Junho de 2024

Capa: *Bombus wurflenii*. Foto tirada em 24 de agosto de 2023, em Evolène, Suíça, por Leonardo Tresoldi Gonçalves.

Universidade Federal do Rio Grande do Sul
Instituto de Biociências
Departamento de Genética
Programa de Pós-Graduação em Genética e Biologia Molecular

Evolução do DNA mitocondrial em *Bombus* (Hymenoptera, Apidae): de *barcodes* a genomas completos

Leonardo Tresoldi Gonçalves

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do título de **Doutor em Genética e Biologia Molecular**.

Orientadora: Prof.ª Dr.ª Maríndia Deprá
Coorientadora: Dr.ª Elaine Aparecida Françoso

Porto Alegre
Junho de 2024

Instituições e Fontes Financiadoras

Este trabalho foi desenvolvido no Laboratório de *Drosophila* do Departamento de Genética da Universidade Federal do Rio Grande do Sul (UFRGS) e no Laboratório do Prof. Dr. Mark J. F. Brown da Royal Holloway, University of London (RHUL, Egham, Surrey, Reino Unido). O trabalho foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), e pelo Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM) da UFRGS. A bolsa de doutorado foi concedida pelo CNPq, e a bolsa de doutorado no exterior pelo programa CAPES-PRINT no âmbito da UFRGS. As taxas estudantis da RHUL foram parcialmente financiadas pelo Centre for Ecology, Evolution and Behaviour, Department of Biological Sciences, RHUL. As participações em eventos científicos foram financiadas pelo CNPq, pela CAPES, pela European Society of Evolutionary Biology (ESEB) e pela Society for the Study of Evolution (SSE).

"– [...] Porque a história da evolução é de que a vida escapa a todas as barreiras. A vida se liberta. A vida se expande para novos territórios. Dolorosamente, talvez até perigosamente. Mas a vida dá um jeito. – Malcolm balançou a cabeça. – Eu não pretendia ser filosófico, mas aí está."

Michael Crichton,
Jurassic Park

Agradecimentos

Preciso começar agradecendo às minhas orientadoras, pois sem elas esta tese e esta trajetória não existiriam – ao menos não do jeito que existem agora! Agradeço à Maríndia, que me recebeu de portas abertas no laboratório, compartilhou sua experiência e conhecimento, e sempre valorizou e estimulou minha independência, projetos e ideias. À Elaine, que me ensinou muito sobre o fascinante mundo das *Bombus*, intensificou minha obsessão por DNA mitocondrial, e me ajudou a construir muitas das melhores memórias que guardo do meu doutorado-sanduíche. Muito obrigado por contribuírem na minha formação enquanto pesquisador e enquanto indivíduo.

Agradeço a todos que agregaram para a minha jornada no PPGBM, um programa de pós-graduação de excelência. Aos professores e colegas, pelas trocas de conhecimento sempre frutíferas, e à equipe da secretaria (Elmo e Gabriel), pelo suporte impecável. Concluo este doutorado com muito orgulho de ter feito parte do PPGBM.

Aos amigos do Laboratório de *Drosophila*: Anelise, Henrique, Natasha, Pedro, Thays e Vitor. Foi maravilhoso compartilhar momentos com vocês dentro e fora da UFRGS. Berenice, Danielle, Helena e Pedro, obrigado pela disponibilidade e por facilitarem nossa vida no laboratório. À Prof.ª Vera, agradeço por tornar o ambiente tão agradável e pelas histórias sempre interessantes.

Aos amigos (membros e agregados) do Laboratório de Evolução Molecular, que se tornou uma segunda casa: Aléxia, Alice, Analu, Giovanna, João Pedro, Lucas, Luana e Sebastián. À Prof.ª Loreta, pela acolhida, ensinamentos, colaborações e por gentilmente aceitar participar do meu exame de qualificação.

Ao Filipe, por sua amizade, apoio e colaboração ao longo desta jornada. Nossa paixão pela ciência e nossas conversas estimulantes criaram uma parceria que foi essencial para atravessar os tempos de pandemia. Obrigado também pela sugestão do título do Capítulo 1 que integra essa tese!

A todos que enriqueceram minha experiência do doutorado-sanduíche no Reino Unido. À Annie, pela hospitalidade, confiança e gentileza. À família da Elaine, Alexandre (He-Man) e Nina, que tornaram minha estadia no Reino Unido ainda mais divertida e memorável. À *Motherland*, especialmente à Erum, pelos momentos de descontração. Ao Jonathan e ao Matheus, pelos *happy hours* e aventuras em Londres e Oxford. Aos colegas

e professores da Royal Holloway, que agregaram muito durante o doutorado-sanduíche e me acompanharam em muitos almoços, pubs e horas do chá: Anees, Cecylia, Elisa, Elli, Lachlan, Leo, Lucy, Malin, Marianne, Mark, Matti, Morgan e Tom.

Ao Lars, uma das melhores surpresas do doutorado-sanduíche. Obrigado por dividir tantos momentos bons comigo, trazer leveza durante este período, me apresentar a novas possibilidades e me ensinar coisas que eu nem mesmo sabia que precisava aprender.

À Luana (Luti), agora minha amiga intercontinental. Agradeço pelas palavras de apoio que me ergueram em momentos desafiadores, pelas várias horas em chamadas de vídeo e por ser um exemplo de foco e dedicação. Essa inspiração foi fundamental para que eu me mantivesse firme na jornada de desenvolvimento desta tese.

A todos meus demais amigos, que me deram suporte nos melhores e piores momentos dos últimos quatro anos. Sou sortudo porque vocês são muitos, então não vou ousar citar nomes, com medo de esquecer alguém! A importância de vocês para a conclusão dessa etapa é imensa.

A todos que participaram do projeto de ciência cidadã desenvolvido durante esta tese. É recompensador ver pessoas fora da bolha acadêmica interessadas em compreender e preservar nossa biodiversidade. Foi um prazer aprender e partilhar meu conhecimento com vocês. Um agradecimento especial à Luciana, Flávia e Wellington, por viabilizarem o acesso à *Bombus bellicosus*.

Ao Pedro, meu companheiro na vida e na ciência. Obrigado pelo apoio incondicional, pela paciência, por todo o carinho e amor, e por se fazer presente mesmo a milhares de quilômetros de distância. À família do Pedro, pela hospitalidade e bons momentos de sempre, mas especialmente pela acolhida durante a pandemia. Muito obrigado por trazerem tranquilidade e leveza em tempos tão difíceis.

À minha família, em especial aos meus pais, Mara e Alessandro, e ao meu irmão, Bruno. Obrigado por todo o suporte e incentivo. Sei que pode ser abstrato entender o que faço, mas agradeço por reconhecerem a importância desta etapa para mim e por vibrarem com as minhas conquistas.

Sumário

Resumo

Himenópteros (abelhas, formigas e vespas) apresentam altas taxas de substituição em seu DNA mitocondrial (mtDNA), tornando-os modelos valiosos para o estudo da evolução do mtDNA. Entre eles, as mamangavas (abelhas do gênero *Bombus*) se destacam como polinizadores importantes em ecossistemas naturais e agrícolas, além de apresentarem grande diversidade ecológica. Esta tese visa compreender os processos evolutivos do mtDNA em *Bombus* sob diferentes perspectivas. No capítulo 1, demonstramos que tanto a região canônica quanto uma região interna (*mini-barcode*) do gene mitocondrial *cox1* apresentam desempenho similar na identificação por *DNA barcoding* de *Bombus* e outras abelhas da família Apidae. Também fornecemos um panorama dos *barcodes* de referência disponíveis para a família no *Barcode of Life Data System* (BOLD). No capítulo 2, montamos 40 novos genomas mitocondriais de *Bombus* para reconstruir as relações filogenéticas do gênero. Nossos achados revelaram discordância entre topologias obtidas por diferentes métodos, além de divergências em relação a hipóteses anteriores baseadas em marcadores nucleares. Também demonstramos que a ordem dos genes de tRNA no cromossomo mitocondrial varia consideravelmente entre espécies, podendo ser utilizada como sinapomorfia para os subgêneros. Por fim, evidenciamos um relaxamento da seleção purificadora nos genes mitocondriais de mamangavas parasitas sociais do subgênero *Psithyrus*. No capítulo 3, usamos *Bombus* como modelo para explorar a forte correlação entre taxas evolutivas de genes mitocondriais e genes nucleares que atuam na mitocôndria, corroborando a teoria da coevolução mitonuclear mesmo entre espécies proximamente relacionadas. Além disso, reportamos que espécies de ambientes mais frios apresentam taxas de evolução do mtDNA mais aceleradas. Por fim, no capítulo 4, apresentamos os genomas mitocondrial e nuclear de *B. bellicosus*, uma espécie ameaçada endêmica do sul da América do Sul. Estes dados fornecem subsídios para a conservação da espécie e para estudos futuros sobre sua biologia e evolução. Esta tese amplia significativamente o conhecimento sobre a evolução do mtDNA em *Bombus*, destacando sua relevância para taxonomia, filogenia, evolução molecular, adaptação e conservação. Os resultados contribuem para uma melhor compreensão dos processos micro e macroevolutivos que moldaram a diversidade deste importante grupo de polinizadores.

Abstract

Hymenopterans (bees, ants, and wasps) present high substitution rates in their mitochondrial DNA (mtDNA), making them invaluable models for studying mtDNA evolution. Among them, bumblebees (*Bombus*) stand out as crucial pollinators in natural and agricultural ecosystems, besides presenting striking ecological diversity. This thesis aims to understand the evolutionary processes of mtDNA in *Bombus* from different perspectives. In chapter 1, we demonstrated that both the canonical region and an internal region (mini-barcode) of mitochondrial gene *cox1* perform similarly in DNA barcoding of bumblebees and other apid bees. We also provided an overview of the reference barcodes available for the family at the Barcode of Life Data System (BOLD). In chapter 2, we assembled 40 new mitochondrial genomes to reconstruct the phylogenetic relationships of *Bombus*. Our findings revealed discordance among topologies obtained with different methods, besides disagreements with previous hypotheses based on nuclear markers. We also demonstrated that the order of tRNA genes on the mitochondrial chromosome varies considerably among species, and can be used as a synapomorphic character for subgenera. Finally, we revealed a relaxation on the purifying selection acting on mitochondrial genes of socially parasitic bumblebees (subgenus *Psithyrus*). In chapter 3, we used bumblebees as a model to explore the strong correlation between evolutionary rates of mitochondrial genes and nuclear genes that act on mitochondria, supporting the theory of mitonuclear coevolution among closely related species. Moreover, we reported that species from colder environments experience faster mtDNA evolutionary rates. Finally, in chapter 4, we presented the mitochondrial and nuclear genomes of the bellicose bumblebee (*B. bellicosus*), a threatened species endemic from southern South America. These data provide a resource for the conservation of the species and for future studies on its biology and evolution. This thesis significantly expands knowledge about the evolution of mtDNA in *Bombus*, highlighting its relevance for taxonomy, phylogeny, molecular evolution, adaptation, and conservation. The results contribute to a better understanding of the micro- and macro-evolutionary processes that shaped the diversity of this important group of pollinators.

Introdução

DNA mitocondrial, um marco em estudos evolutivos

O DNA mitocondrial (mtDNA) tem sido amplamente utilizado em estudos de ecologia molecular e filogeografia há mais de quatro décadas (Ballard e Whitlock 2004). Suas características únicas tornam essa molécula uma ferramenta crucial na biologia evolutiva. Por ser um genoma haploide, geralmente não-recombinante e com herança uniparental, o tamanho efetivo populacional (N_e) do mtDNA é considerado 1/4 do N_e do DNA nuclear. Isso é vantajoso para estudos populacionais, pois o N_e reduzido acelera a fixação de alelos, e a herança uniparental e a ausência de recombinação viabilizam o rastreamento de informações genealógicas e de história demográfica (Avise et al. 1984; DeSalle et al. 2017). Outro fator importante é que as taxas de substituição do mtDNA são, em média, nove a 25 vezes mais rápidas do que as do DNA nuclear (Lynch 2006). Isso provavelmente se deve à alta frequência de erros de replicação no genoma mitocondrial (Larsson 2010; Melvin e Ballard 2017; Szczepanowska e Trifunovic 2017) e à frequente exposição do DNA e da maquinaria mitocondrial a espécies reativas de oxigênio (Barja 1998; Lane 2011; Anderson et al. 2020). Coletivamente, estas características permitem um rápido acúmulo de mutações no mtDNA, mesmo entre linhagens proximamente relacionadas, o que é muito atrativo do ponto de vista da filogenética e da genética de populações. Além disso, a ordem dos genes no cromossomo mitocondrial pode ser acessada de modo relativamente simples, o que permite o uso dessas informações em um contexto filogenético (Richardson et al. 2013; Françoso et al. 2020; Sun et al. 2021b).

Um dos adventos mais emblemáticos envolvendo o mtDNA nos últimos 20 anos foi a formalização da técnica de *DNA barcoding* (Hebert et al. 2003; DeSalle e Goldstein 2019). O "código de barras de DNA" baseia-se no uso de loci que apresentem alta variação entre espécies, mas pouca variação entre indivíduos da mesma espécie. Isso permite que as sequências dos loci em questão sejam comparadas com um banco de dados com sequências previamente identificadas, possibilitando a identificação rápida e confiável de espécimes por meio de análises de distância genética. Os loci usados como *DNA barcodes* variam para cada grupo taxonômico, mas para animais o mais utilizado é um fragmento de cerca de 650 pb da região 5' do gene mitocondrial citocromo *c* oxidase subunidade I (*cox1* ou

COI; Ratnasingham e Hebert 2007; DeSalle e Goldstein 2019). Os *DNA barcodes* gerados são depositados no BOLD (*Barcode of Life Data System*), banco de dados gerido pelo *Consortium for the Barcode of Life*, que atualmente conta com mais de 17 milhões de *barcodes* (Ratnasingham e Hebert 2007).

Apesar de revolucionário, o mtDNA não é uma solução universal para estudos evolutivos. Suas complexidades e limitações estão cada vez mais evidentes. Mesmo que as altas taxas de substituição e a herança uniparental sejam comuns, a genética mitocondrial varia significativamente entre linhagens eucarióticas. As taxas de substituição mitocondrial apresentam ampla variação (Oliveira et al. 2008; Nabholz et al. 2009; Richardson et al. 2013), e há várias exceções em que os genomas mitocondriais são herdados de maneira biparental ou realizam recombinação (Barr et al. 2005; Breton e Stewart 2015). Além disso, a presença de heteroplasmia – múltiplos genomas mitocondriais por indivíduo – viola a natureza clonal e haploide do mtDNA em muitas linhagens (Rand 2001; Iannello et al. 2019; Ricardo et al. 2020). Por fim, evidências crescentes mostram que a evolução do mtDNA é influenciada pela sua relação com o genoma nuclear, como discutido a seguir.

Interação mitonuclear: o DNA mitocondrial além da mitocôndria

A origem endossimbiótica da mitocôndria destaca-se como um dos eventos evolutivos mais marcantes no surgimento da vida complexa na Terra (Sagan 1967). Além de sua função principal na produção de energia por meio da respiração celular, a mitocôndria desempenha diversos outros papéis vitais em eucariotos, incluindo o metabolismo de lipídios, a síntese de aminoácidos, a regulação da homeostase de íons e a ativação da morte celular programada (Nunnari e Suomalainen 2012). O estabelecimento dessa endossimbiose foi marcado pela perda ou transferência de diversos genes mitocondriais para o genoma da célula hospedeira, o que deu origem ao genoma nuclear (Hill 2015). Nos animais, o genoma mitocondrial geralmente consiste em um único cromossomo circular de DNA de fita dupla, contendo cerca de 15 kbp. Este tamanho é irrisório mesmo quando comparado ao menor dos genomas nucleares animais conhecido pela ciência (o do nematoide *Pratylenchus coffeae*, com 20 Mbp; Burke et al. 2015).

Nos animais bilaterais, o genoma mitocondrial é altamente compactado, tipicamente contendo 37 genes (Figura 1). Destes, 24 genes codificam a própria

maquinaria de tradução da mitocôndria (22 tRNAs e dois rRNAs). Os 13 genes restantes codificam subunidades do sistema de transporte de elétrons responsável pela fosforilação oxidativa (OXPHOS), que processa carboidratos e lipídios para gerar dióxido de carbono, água e ATP. No entanto, a vasta maioria das subunidades necessárias para a OXPHOS é codificada pelo genoma nuclear. Dos cinco complexos envolvidos na OXPHOS, quatro são compostos por subunidades codificadas tanto pelo genoma nuclear quanto pelo genoma mitocondrial (Figura 2). Consequentemente, estes genomas devem evoluir em concerto para manter a OXPHOS e outras funções mitocondriais — um processo denominado coevolução mitonuclear (Rand et al. 2004).

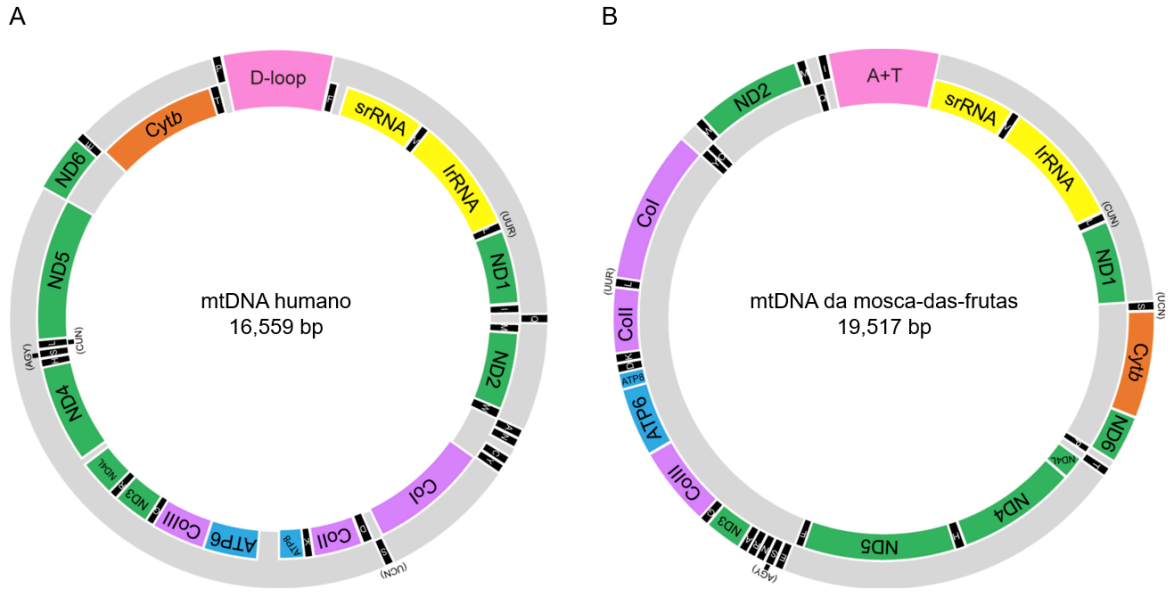

Figura 1. Organização do genoma mitocondrial de humanos e de moscas-das-frutas (*Drosophila melanogaster*). Mesmo ao comparar linhagens evolutivamente distantes de animais, a estrutura e conteúdo do genoma mitocondrial são conservados. **A.** O genoma mitocondrial humano codifica 13 proteínas, 12 tRNAs e dois rRNAs. ND, NADH desidrogenases, subunidades do Complexo I mitocondrial; Cyt*b*, citocromo *b*, subunidade do Complexo III; Co, citocromo *c* oxidases, subunidades do Complexo IV; ATP, ATP sintases, subunidades do Complexo V; ssRNA, subunidade 12S de rRNA; lsRNA, subunidade 16S de rRNA. Os 22 tRNAs estão representados pela abreviação IUPAC dos aminoácidos que eles transportam. Uma região não-codificante de aproximadamente 1 kbp, chamada de D-loop, inclui sequências promotoras da transcrição e uma das origens da replicação. **B.** O genoma mitocondrial da mosca-das-frutas codifica os mesmos genes que o genoma mitocondrial humano, mas a ordem dos genes no cromossomo mitocondrial é um pouco diferente. O genoma mitocondrial da mosca-das-frutas é cerca de 3 kbp maior

que seu equivalente humano, principalmente devido ao maior tamanho da região reguladora rica em A+T. Adaptado de Chen et al. (2019).

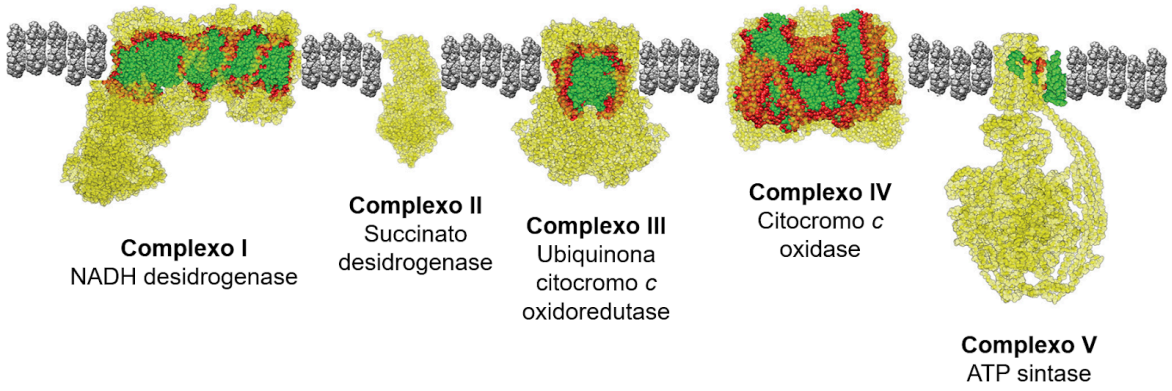

Figura 2. Complexos do sistema de transporte de elétrons, a principal maquinaria celular envolvida na síntese de ATP em eucariotos, que ficam inseridos na membrana interna mitocondrial. Os resíduos codificados pelo genoma mitocondrial estão em verde, resíduos codificados pelo genoma nuclear estão representados em amarelo, e as interações físicas entre resíduos codificados por diferentes genomas estão realçadas em vermelho. A maior parte das subunidades nesse sistema é codificada pelo núcleo (cerca de 75 genes), enquanto apenas 13 subunidades são codificadas pelo genoma mitocondrial. Notavelmente, o Complexo II é inteiramente codificado pelo genoma nuclear. Adaptado de Hill et al. (2019).

Evidências crescentes indicam que essa interdependência entre genomas nuclear e mitocondrial influencia a evolução e a ecologia dos eucariotos (Hill 2015), embora geralmente seja subestimada em estudos envolvendo populações naturais. Por exemplo, vários estudos sugerem que a interação mitonuclear é um mecanismo subjacente ao início da divergência populacional e à formação de barreiras reprodutivas, que são a base para eventos de especiação (Ellison e Burton 2006; Ellison e Burton 2010; Trier et al. 2014; Bar-Yaacov et al. 2015; Camus et al. 2017; Burton 2022). Além disso, a disrupção dos genomas nuclear e mitocondrial coadaptados resulta em incompatibilidades mitonucleares, comprometendo o funcionamento do sistema de transporte de elétrons e a montagem dos complexos de proteínas, além de aumentar anormalmente a produção de espécies reativas de oxigênio (Wolff et al. 2014). As interações mitonucleares também podem desempenhar um papel importante na resposta a doenças (Holmbeck et al. 2015; Andrews et al. 2020), na adaptação de nicho climático (Bernardo et al. 2019; Havird et al. 2020; Wang et al. 2021) e até mesmo na seleção sexual (Hill e Johnson 2013; Hill 2018). Portanto, a

interação entre estes genomas age de forma a criar pressões seletivas intrínsecas que favoreçam genótipos mitonucleares coadaptados que mantenham o funcionamento mitocondrial (Burton 2022).

Uma consequência da coevolução mitonuclear é a correlação de taxas evolutivas (ERC) mitonucleares, ou seja, a tendência de taxas evolutivas de genes mitocondriais e genes nucleares que interagem com produtos da mitocôndria variarem em conjunto, apesar da diferença entre as taxas de substituição do DNA nuclear e do mtDNA (de Juan et al. 2013; Piccinini et al. 2021) (Figura 3). Essas assinaturas de coevolução mitonuclear já foram detectadas em diversas linhagens, como bivalves (Piccinini et al. 2021), mamíferos (Weaver et al. 2022) e insetos holometábolos, incluindo himenópteros (Yan et al. 2019). No entanto, esses estudos geralmente se concentram em linhagens filogeneticamente amplas, que divergiram há muito tempo. Uma questão em aberto é se a ERC pode ser detectada, por exemplo, em nível taxonômico de espécie (Weaver et al. 2022).

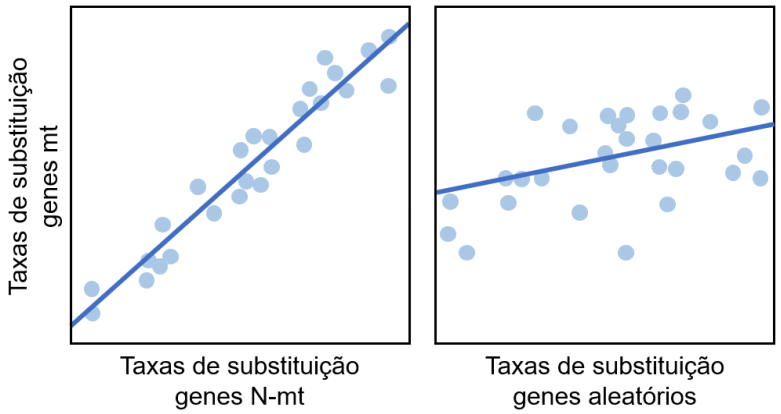

Figura 3. Em um cenário de coevolução mitonuclear, há uma correlação entre as taxas de substituição de genes mitocondriais (mt) e genes codificados no núcleo que agem na mitocôndria (N-mt). Apesar das diferenças inerentes entre taxas de substituição dos genomas mitocondrial e nuclear, essa correlação se mantém de modo a preservar o funcionamento dos complexos de proteínas da mitocôndria, que são compostos por subunidades mt e N-mt (veja também a Figura 2). Em contrapartida, essa correlação não se mantém ao compararmos genes mt e ortólogos aleatórios do núcleo que não estejam envolvidos com processos mitocondriais.

Himenópteros, um modelo para a evolução do DNA mitocondrial e das interações mitonucleares

Hymenoptera é uma das ordens de insetos mais ricas em espécies, incluindo mais de 150 mil espécies descritas e uma estimativa de 1 milhão de espécies (Sharkey 2007; Aguiar et al. 2013). Hymenoptera consiste em espécies diversas, incluindo moscas-serra e vespas da madeira[1] ("Symphyta"), vespas parasitoides ("Parasitica"), e vespas com ferrão (Aculeata, grupo que abrange formigas e abelhas). Esta ordem é também uma das mais biologicamente diversas, compreendendo espécies fitófagas, micófagas, onívoras e predadoras. Apresentam um amplo espectro entre vida solitária e eusocialidade, e uma riqueza impressionante de estratégias de parasitismo (Whitfield 2003). Outra característica distintiva dos himenópteros é o sistema de determinação do sexo haplodiploide, onde as fêmeas são diploides e os machos são haploides (Heimpel e Boer 2008).

Nos himenópteros, as taxas de substituição do genoma mitocondrial são excepcionalmente mais altas do que as do genoma nuclear (Oliveira et al. 2008; Kaltenpoth et al. 2012; Yan et al. 2019). Por exemplo, as taxas de substituição sinônima nos genes mitocondriais de vespas *Nasonia* são cerca de 40 vezes maiores do que nos genes nucleares (Oliveira et al. 2008), uma das taxas mais aceleradas conhecidas. Em abelhas sem ferrão *Tetragonula*, o mtDNA evolui cerca de 26 vezes mais rápido do que o DNA nuclear (Hereward et al. 2020). No entanto, os mecanismos por trás desta diferença marcante entre taxas de substituição ainda são especulativos (Baer et al. 2007). Linhagens com mtDNA de evolução rápida devem possuir genes nucleares que interagem com a mitocôndria evoluindo igualmente rápido, aumentando a probabilidade de evolução compensatória e de incompatibilidade mitonuclear (Havird e Sloan 2016). Além disso, como os machos são haploides, há apenas um alelo nuclear para cada loco interagindo com os genes mitocondriais. Machos resultantes de cruzamentos híbridos entre espécies de *Nasonia*, por exemplo, mostram uma disfunção no sistema de transporte de elétrons devido a incompatibilidades mitonucleares (Ellison et al. 2008).

Outras peculiaridades dos genomas mitocondriais de Hymenoptera ressaltam seu potencial como modelo para estudar a evolução mitocondrial e a coadaptação mitonuclear. Diferente de outras ordens de insetos, os himenópteros exibem um excesso de rearranjos

[1] Tradução livre, do inglês “sawflies” e “wood wasps”. Os representantes de Symphyta não possuem nome popular em português.

em seus genomas mitocondriais, que frequentemente variam entre espécies próximas (Dowton e Austin 1999; Dowton et al. 2003; Zheng et al. 2018; Françoso et al. 2020; Françoso et al. 2023). Há também registros de heteroplasmia em diversas linhagens (Magnacca e Brown 2010; Meza-Lázaro et al. 2018; Ricardo et al. 2020), bem como duplicação total do genoma mitocondrial em abelhas *Tetragonula* (Françoso et al. 2023).

Também parece haver uma influência do hábito de vida parasita na evolução do mtDNA em Hymenoptera. Linhagens de himenópteros parasitas sociais ou parasitoides apresentam taxa de evolução do mtDNA acelerada quando comparadas a espécies não-parasitas próximas (Xiao et al. 2011; Zhu et al. 2018; Schrader et al. 2021). Ainda não está claro se essa aceleração nas taxas evolutivas é resultado do menor N_e dessas espécies, como ocorre em outros parasitas (Castro et al. 2002; Oliveira et al. 2008; Jakovlić et al. 2021), ou se há uma razão funcional subjacente a essas alterações. Em himenópteros, é muito comum que parasita e hospedeiro sejam espécies evolutivamente próximas, uma generalização conhecida como "Regra de Emery" (Emery 1909). Portanto, o grupo também oferece um excelente contexto filogenético para investigar a influência do mtDNA no parasitismo social.

Bombus, as mamangavas-de-chão

Depois das abelhas melíferas (*Apis*), as mamangavas[2] (*Bombus*) se destacam como um dos grupos de abelhas mais estudados pela ciência. Elas desempenham um papel crucial como polinizadoras de plantas silvestres e cultivadas, o que lhes confere grande importância ecológica e econômica. *Bombus* agrupa cerca de 250 espécies, classificadas em 15 subgêneros (Williams et al. 2022). Em sua maioria, os subgêneros podem ser divididos em dois grandes clados, *"short-faced"* e *"long-faced"*, cujos nomes se referem à morfologia da cabeça e ao comprimento da probóscide (Cameron et al. 2007). A maior riqueza de espécies ocorre nas regiões Paleártica e Oriental, com aproximadamente 175

[2] De origem tupi, "mamangava" significa "abelha de grande porte". No Brasil, esse termo é popularmente usado para designar abelhas dos gêneros *Bombus* e *Xylocopa*. Para evitar confusões, alguns autores denominam espécies de *Xylocopa* como "mamangavas-de-toco" ou "mamangavas-de-pau-podre", e espécies de *Bombus* como "mamangavas-de-chão" (Garófalo 2005). Essas denominações se referem ao substrato em que as abelhas de cada um dos gêneros constroem seus ninhos. No entanto, nunca ouvi esses termos sendo usados coloquialmente, e todas as pessoas com quem tive contato durante o projeto de ciência cidadã (Capítulo 4 desta tese) e durante as demais etapas deste doutorado usavam simplesmente o termo "mamangava" para se referir às espécies de *Bombus*. Portanto, é o termo que adotei ao redigir os trechos em português desta tese.

espécies, enquanto as regiões Neártica e Neotropical apresentam menor diversidade, com cerca de 60 e 25 espécies, respectivamente (Williams et al. 2008; Hines 2008). No Brasil, ocorrem apenas oito espécies, todas pertencentes ao subgênero *Thoracobombus*: *B. applanatus*, *B. bahiensis*, *B. bellicosus*, *B. brasiliensis*, *B. brevivillus*, *B. morio*, *B. pauloensis* e *B. transversalis* (Williams 1998; Santos Júnior et al. 2015; Françoso et al. 2016).

As mamangavas são um importante modelo para estudos da ecologia e do comportamento social (Goulson 2010; Amsalem et al. 2015). Possuem um ciclo de vida complexo, geralmente anual, marcado por eventos de hibernação, fases solitárias e fases eusociais (Figura 4) — embora algumas espécies brasileiras apresentem ciclo de vida bianual e não hibernem (Moure e Sakagami 1962; Zucchi e Kerr 1974; Camillo e Garófalo 1989). Notavelmente, um grupo de espécies de mamangavas abandonaram completamente o modo de vida eusocial e adaptaram-se a um modo de vida parasita social, onde fêmeas solitárias usurpam o ninho de mamangavas sociais e exploram as operárias hospedeiras para cuidar da sua prole (Figura 4). O parasitismo social obrigatório é característico de um dos subgêneros de *Bombus*, *Psithyrus*, embora tenha surgido independentemente em pelo menos outras duas espécies de mamangavas (Lhomme e Hines 2019).

Apesar de ocuparem uma ampla variedade de ecossistemas, as mamangavas são mais comumente encontradas em altitudes elevadas (> 1000 m) ao redor do mundo, caracterizadas por clima frio e baixa disponibilidade de oxigênio (Cameron et al. 2007; Hines 2008). Por outro lado, muitas espécies brasileiras adaptaram-se a terras baixas e quentes (Cameron e Williams 2003), que contrastam com as condições ancestrais do gênero (Hines 2008). Em animais ectotérmicos, como as mamangavas e outros insetos, mudanças na temperatura externa afetam as reações bioquímicas da mitocôndria (Simčič et al. 2014). Já a concentração de oxigênio disponível influencia a eficiência na síntese de ATP e na produção de radicais livres (Fuhrmann e Brüne 2017). Portanto, fatores como temperatura e oxigênio moldam a evolução do mtDNA e são cruciais para o sucesso da coevolução mitonuclear (Hill 2015). O contraste ecológico existente entre espécies proximamente relacionadas, como é o caso de *Bombus*, somado à rápida evolução do mtDNA e seu potencial adaptativo, tornam estas abelhas modelos excelentes para o entendimento de interações mitonucleares em uma abordagem filogenética.

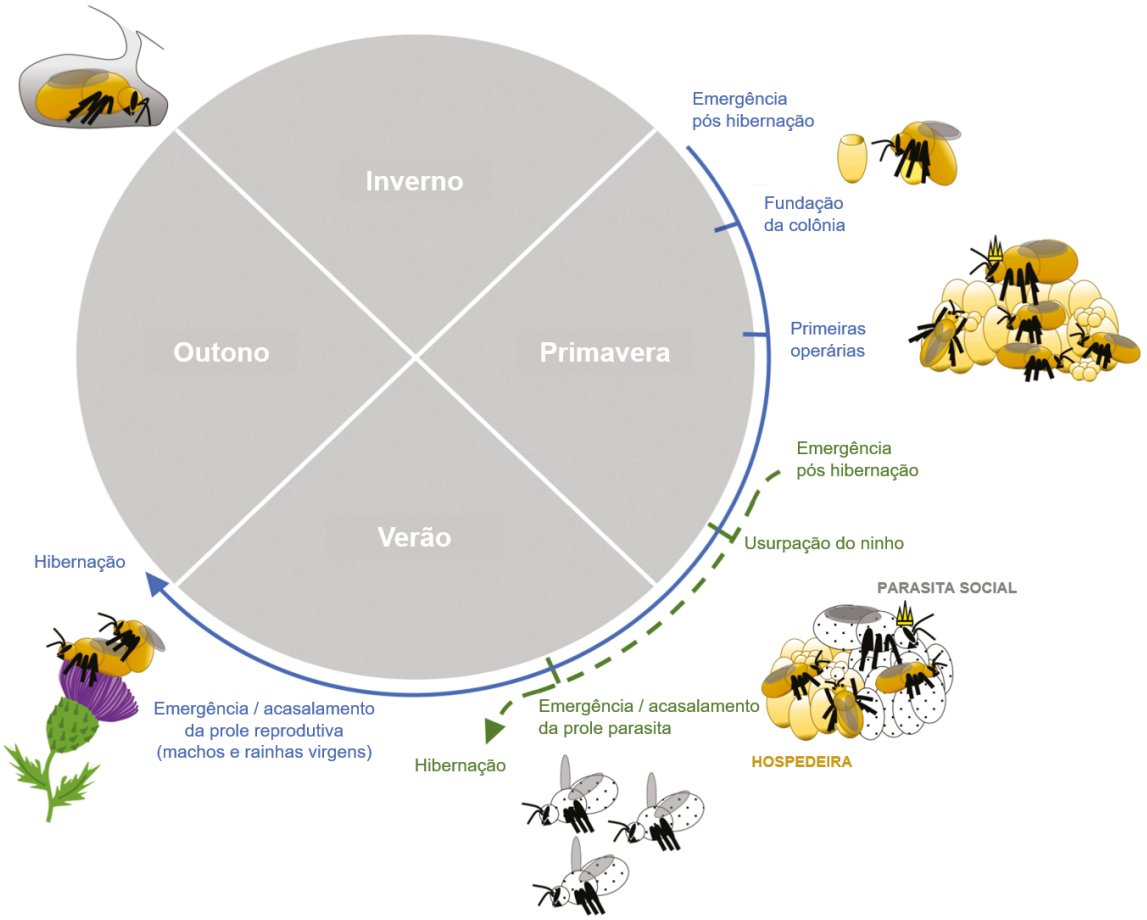

Figura 4. Ciclo de vida geral das mamangavas sociais (linha azul contínua) e das mamangavas parasitas sociais (linha verde tracejada). As rainhas de mamangavas sociais passam por uma fase solitária, saindo da hibernação no início da primavera para iniciar seus ninhos e produzir as primeiras operárias da colônia. Quando estas operárias eclodem na metade da primavera, elas assumem as funções de cuidar da prole e buscar alimento, para manter a rainha e as gerações seguintes de operárias. No final do verão, as colônias passam a produzir machos e fêmeas reprodutivas (rainhas virgens). A prole reprodutiva sai do ninho e acasala, e as rainhas recém-fecundadas hibernam a partir do início do outono. As fêmeas de mamangavas parasitas sociais emergem mais tarde na primavera e usurpam os ninhos hospedeiros geralmente quando a primeira leva de operárias está sendo produzida. Seu ciclo de vida é mais curto, resumindo-se à produção de machos e fêmeas reprodutivas no início do verão. Como acontece com suas hospedeiras, a prole das mamangavas parasitas abandona o ninho, acasala e hiberna. Adaptado de Lhomme e Hines (2019).

Diversas espécies de *Bombus* ao redor do mundo estão em declínio devido a uma série de fatores, incluindo mudanças climáticas, uso de agrotóxicos, parasitas e vírus emergentes, e perda de habitat (Cameron et al. 2011; Rasmont e Iserbyt 2012; Goulson

2015; Sirois-Delisle e Kerr 2018; Suzuki-Ohno et al. 2020). No Brasil, três espécies de mamangavas (*B. bellicosus*, *B. brevivillus* e *B. brasiliensis*) correm risco de extinção (Martins e Melo 2010; Martins et al. 2015; Krechemer e Marchioro 2020), mas ainda não figuram na Lista Vermelha de Espécies Ameaçadas por falta de dados suficientes. Uma das espécies-foco desta tese, *B. bellicosus*, já é considerada extinta em parte de sua área de ocorrência devido ao impacto do aquecimento global em suas populações (Martins e Melo 2010; Martins et al. 2015). Estudos recentes demonstram que haplótipos mitocondriais estão intimamente ligados à capacidade de adaptação climática e à resistência a doenças em mamangavas (Manlik et al. 2023). Compreender o papel do mtDNA nesse contexto, portanto, pode ser fundamental para a conservação dessas espécies.

Objetivos

Objetivo geral

Investigar o papel do mtDNA na adaptação e diversificação de abelhas do gênero *Bombus*, explorando suas aplicações na taxonomia, sistemática filogenética, evolução molecular e conservação da biodiversidade.

Objetivos específicos

A. Comparar a performance de duas regiões do gene mitocondrial *cox1* na identificação de abelhas da família Apidae: a região canônica usada em estudos de *DNA barcoding* e um *mini-barcode* que abrange um trecho interno à região canônica;

B. Fornecer um panorama geral dos *DNA barcodes* de referência disponíveis para a família Apidae no BOLD;

C. Avaliar a utilidade de genomas mitocondriais completos, bem como dos rearranjos estruturais no cromossomo mitocondrial, no entendimento da história evolutiva de *Bombus*;

D. Explorar as mudanças no regime de seleção dos genomas mitocondriais de *Bombus*, com foco nas espécies parasitas sociais obrigatórias do subgênero *Psithyrus*;

E. Investigar os padrões de coevolução entre os genomas nuclear e mitocondrial de *Bombus* e a possível relação com variáveis ambientais;

F. Descrever os genomas nuclear e mitocondrial de *Bombus bellicosus*, uma espécie endêmica do sul da América do Sul potencialmente ameaçada de extinção.

Capítulo 1

Shorter, better, faster, stronger? Comparing the identification performance of full-length and mini-DNA barcodes for apid bees (Hymenoptera: Apidae)

Leonardo Tresoldi Gonçalves, Elaine Françoso, Maríndia Deprá

Artigo publicado em 2022 no periódico
Apidologie
https://doi.org/10.1007/s13592-022-00958-x

Resumo

Abelhas da família Apidae são polinizadores essenciais nos ecossistemas, ocorrendo ao redor do mundo e compreendendo cerca de 5900 espécies. Embora sejam identificadas principalmente por meio da morfologia, o método de *DNA barcoding* vem sendo explorado como uma ferramenta suplementar na taxonomia de abelhas. Regiões menores de *DNA barcodes*, conhecidas como *mini-barcodes*, também foram implementadas com sucesso na identificação de abelhas corbiculadas. No entanto, o desempenho dos *mini-barcodes* foi testado apenas em um escopo taxonômico restrito. Neste estudo, analisamos todas as 18167 sequências do gene *cox1* da família Apidae disponíveis no *Barcode of Life Data System* para fornecer uma visão geral dos dados disponíveis, buscar por *barcoding gaps* a nível de gênero, testar se *barcodes* canônicos e *mini-barcodes* têm desempenho semelhante na identificação de espécimes e sinalizar táxons de abelhas que possam se beneficiar de estudos que implementem *DNA barcodes*. Nosso conjunto de dados incluiu cinco subfamílias, 25 tribos, 71 gêneros e 1012 espécies, sendo a maioria pertencente a tribos de abelhas corbiculadas. A maioria dos gêneros analisados mostrou um bom desempenho nas análises de *barcoding gap*. Além disso, *barcodes* canônicos e *mini-barcodes* exibiram uma probabilidade de identificação correta semelhante, demonstrando que ambos os tipos de marcador são equivalentes na identificação de abelhas. Por fim, discutimos alguns exemplos para mostrar como *barcodes* canônicos e *mini-barcodes* podem ajudar a resolver inconsistências taxonômicas e fomentar estudos futuros com abelhas.

Apidologie (2022) 53:55

https://doi.org/10.1007/s13592-022-00958-x

Original article

Shorter, better, faster, stronger? Comparing the identification performance of full-length and mini-DNA barcodes for apid bees (Hymenoptera: Apidae)

Leonardo Tresoldi Gonçalves[1,2], Elaine Françoso[3], Maríndia Deprá[1,2]

[1] Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[2] Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[3] Department of Biological Sciences, Royal Holloway University of London, Egham TW20 0EX, UK

Received 11 October 2021 – Revised 2 June 2022 – Accepted 27 July 2022

Abstract – Apid bees are essential ecosystem pollinators, occurring worldwide and comprising over 5900 species. Although they are identified mainly using morphology, DNA barcoding has been explored since its proposal as a supplementary tool in bee taxonomy. Smaller regions of barcode markers—mini-barcodes—were also successfully employed in corbiculate bee identification, but the performance of mini-barcodes was only tested in a narrow taxonomic scope. Here, we scrutinized all 18167 apid bee *cox1* sequences from the Barcode of Life Data System to provide an overview of the available data, search for barcoding gaps at genus level, test if full-length and mini-barcode regions perform similarly in specimen identification, and flag bee taxa that may benefit from studies implementing DNA barcodes. Our dataset encompassed five subfamilies, 25 tribes, 71 genera, and 1012 species, although it was biased towards corbiculate tribes. Most of the surveyed genera showed good performance in the barcoding gap analyses. Moreover, full-length and mini-barcodes displayed a similar probability of correct identification, demonstrating that both marker types are equivalent in bee identification. Finally, we discuss some examples to show how full-length and mini-barcodes can help solve taxonomic inconsistencies and foment future studies of apid bees.

Apoidea / Anthophila / Barcode of Life Data System / COI / *cox1* / integrative taxonomy

1. INTRODUCTION

The correct identification of species underpins most biological studies. In the last decades, molecular tools have been applied in organism identification, diversity surveys, and species delimitation (Roe et al. 2017). For animals, the 5′ region of the mitochondrial gene cytochrome *c* oxidase subunit I (*cox1*) was formalized as a DNA barcode, allowing a quick, efficient, and reliable tool for molecular identification (Hebert et al. 2003). DNA barcoding relies on comparing genetic distances of intraspecific and interspecific specimens. Generated sequences can be deposited in reference databases for future comparisons with novel data (Ratnasingham and Hebert 2007). Besides its applications in specimen identification, *cox1* barcodes may unveil cryptic diversity, shed light on species boundaries, and aid in phylogenetic and phylogeographic studies (DeSalle and Goldstein 2019).

Bees are known for their fundamental pollination role in ecosystems and their commercial and

Corresponding author: L. T. Gonçalves, tresoldigoncalves@gmail.com
Manuscript editor: Marina Meixner

Published online: 09 September 2022

scientific importance (Michener 2007). Among the bees, the Apidae form a clade comprising over 5900 valid species within 34 tribes (Table I; Bossert et al. 2019). Bee identification is based primarily on morphological characters and morphometric measurements (e.g., Michener 2007; Bustamante et al. 2021; Boustani et al. 2021; Schaller and Roig-Alsina 2021; Nogueira et al. 2022). However, morphology alone may be misleading since some bee species are challenging to distinguish, forming cryptic species complexes (e.g., some bumblebee species; Williams et al. 2012; Martinet et al. 2019). Specimens of certain life stages or castes may also lack informative characters, making it impossible to identify eggs, larvae, and most pupae to species (Michener 2007). Therefore, DNA barcoding has been employed as a supplementary tool to shed light on bee taxonomy (Schmidt et al. 2015; González-Vaquero et al. 2016; Packer and Ruz 2017).

The standard DNA barcode proposed by Hebert et al. (2003) corresponds to the so-called Folmer region, a 648 bp fragment at the 5′ end of the mitochondrial gene *cox1* amplified by the primers designed by Folmer et al. (1994). This region was initially chosen because it is informative and relatively easy to amplify, besides being sufficiently conserved within species yet variable between species (Hebert et al. 2003). Smaller regions of barcode markers—mini-barcodes—were developed for accurate identification in samples with degraded DNA (Hajibabaei et al. 2006). These markers can also be handy and cost-effective in high-throughput sequencing projects (Yeo et al. 2020). For bees, mini-barcodes may be applied to environmental samples, archived specimens (Françoso and Arias 2013), and commercial products such as honey (Schnell et al. 2010). A mini-barcode based on a 175 bp region of *cox1* was proposed for specimen identification of corbiculate bees (Figure 1) (Françoso and Arias 2013). However, the performance of this marker remains insufficiently tested: previous studies have only focused on a narrow taxonomic scope (e.g., Françoso and Arias 2013; Blasco-Lavilla et al. 2019), were limited to a regional bee fauna (e.g., Magnacca and Brown 2012; Sheffield et al. 2017), or evaluated mini-barcodes in a broader sense but did not include apid bees (e.g., Meusnier et al. 2008; Yeo et al. 2020).

In this context, we tested in silico if full-length barcodes and mini-barcodes perform similarly in specimen identification and species discovery of apid bees. We datamined all Apidae *cox1* sequences deposited in the Barcode of Life Data Systems (BOLD), aiming to (1) verify the existence of barcoding gaps in full-length and mini-barcodes at the generic level in all available apid genera, (2) test if full-length and mini-barcode regions have similar success rates of specimen identification, and (3) flag bee taxa that may benefit from integrative studies implementing DNA barcodes. Furthermore, we provide here an overview of the information available in BOLD concerning barcode sequences of apid bees.

2. MATERIALS AND METHODS

2.1. Data retrieval and filtering

We retrieved all BOLD sequences labeled as "Apidae" on November 27, 2020, generating separate FASTA files for each genus. To ensure robust analyses, we followed several filtering steps described by Bianchi and Gonçalves (2021a). Briefly, we only maintained sequences belonging to the 5′ region of *cox1* (labeled in BOLD as "COI-5P") and removed those entries without species-level identification (e.g., *Bombus* sp.). Then, we conducted preliminary alignments on MAFFT 7.0 (Katoh et al. 2019), seeking sequences with nonsense mutations, insertions, and deletions. These entries were removed from the datasets since we assumed they resulted from low-quality sequencing, erroneous amplification (e.g., nuclear mitochondrial DNA segments), or lab contamination. After the filtering steps, a final alignment round was conducted on MAFFT with default parameters.

AliView (Larsson 2014) was used to inspect the alignments and trim the sequences to the barcode region amplified by the primer pair BarbeeF (Françoso and Arias 2013) and MtD9 (Simon

Table I Taxonomic coverage of this study, sorted by subfamily and tribe following the revised generic classification of Bossert et al. (2019). In parentheses, the number of valid genera (sensu Bossert et al. 2019) and species (a rough estimative according to the Integrated Taxonomic Information System online database; http://www.itis.gov) for each tribe. Sampled sequence count is also provided, and estimated species coverage is given in percent (%)

Subfamily	Tribe	Genera	Species	Sequences	Species coverage (%)
Anthophorinae	Anthophorini	3 (7)	69 (794)	375	8.69
Apinae	Apini	1 (1)	8 (8)	2063	100.00
	Bombini	* 1 (1)	179 (280)	5566	63.93
	Centridini	2 (2)	18 (271)	49	6.64
	Euglossini	5 (5)	143 (248)	1615	57.66
	Meliponini	19 (51)	93 (518)	1712	17.95
Eucerinae	Ancylaini	0 (2)	0 (16)	0	0.00
	Emphorini	3 (10)	20 (120)	81	16.67
	Eucerini	6 (27)	131 (801)	549	16.35
	Exomalopsini	2 (5)	7 (156)	10	4.49
	Tapinotaspidini	5 (8)	23 (146)	36	15.75
Nomadinae	Ammobatini	1 (7)	2 (117)	3	1.71
	Ammobatoidini	2 (5)	6 (32)	10	18.75
	Biastini	1 (3)	2 (12)	3	16.67
	Brachynomadini	0 (5)	0 (26)	0	0.00
	Coelioxoidini	0 (1)	0 (4)	0	0.00
	Caenoprosopidini	0 (2)	0 (2)	0	0.00
	Epeolini	2 (8)	70 (309)	267	22.65
	Ericrocidini	3 (9)	9 (44)	14	20.45
	Hexepeolini	0 (1)	0 (1)	0	0.00
	Isepeolini	2 (2)	6 (21)	8	28.57
	Melectini	2 (9)	14 (206)	64	6.80
	Neolarrini	1 (1)	2 (16)	3	12.50
	Nomadini	1 (1)	95 (701)	738	13.55
	Osirini	1 (5)	2 (52)	7	3.85
	Protepeolini	1 (1)	2 (5)	3	40.00
	Rhathymini	0 (2)	0 (19)	0	0.00
	Townsendiellini	0 (1)	0 (3)	0	0.00
Xylocopinae	Allodapini	4 (16)	25 (248)	129	10.08
	Ceratinini	1 (1)	43 (339)	911	12.68
	Ctenoplectrini	1 (1)	5 (20)	11	25.00
	Manueliini	0 (1)	0 (3)	0	0.00
	Tetrapediini	0 (1)	0 (25)	0	0.00
	Xylocopini	1 (1)	38 (400)	351	9.50
		71 (203)	1012 (5963)	14,578	16.97

**Bombus* subgenera were also treated separately; our sample encompasses the 15 recognized Bombus subgenera sensu Williams et al. (2008)

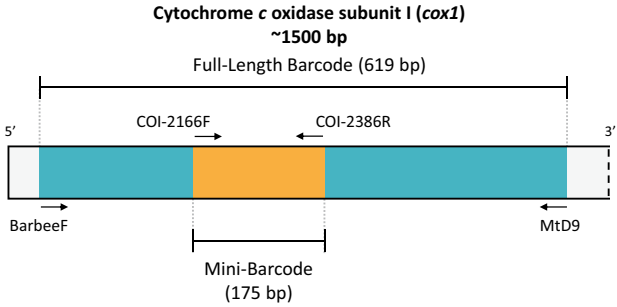

Figure 1. Schematic representation of the 5′ terminal of the mitochondrial gene cytochrome *c* oxidase subunit I (*cox1*). Primer pair BarbeeF and MtD9 amplify the full-length barcode in bees (619 bp); primer pair COI-2166F and COI-2386R amplify the mini-barcode (175 bp). Adapted from Françoso and Arias (2013)

et al. 1994), the first one specially developed to generate a 619 bp fragment of the Folmer region in bees (Figure 1). Sequences shorter than 400 bp were removed, and the remaining sequences composed our primary dataset (Full-Length Barcode Dataset). A secondary dataset (Mini-Barcode Dataset) was compiled, trimming the sequences from the Full-Length Barcode Dataset to a 175 bp mini-barcode region, which is amplified by the primer pair COI-2166F and COI-2386R (mini-barcode II; Françoso and Arias 2013) (Figure 1). As a final filtering step, we double-checked scientific names, correcting misspellings and non-valid names (i.e., synonyms) according to the Integrated Taxonomic Information System (www.itis.gov) and recent literature.

To guarantee intra- and interspecific comparisons, the analyses described below comprise only genera featuring at least two species, with at least one of the species represented by two or more sequences. Sequences identified as subspecies were treated at the species level. Although our analyses focused on the generic level, we also analyzed *Bombus* subgenera given the diversity of species and the high number of sequences recovered for this genus (see "Results"). We followed the simplified subgeneric classification of Williams et al. (2008). Our results are presented using the revised generic classification of Apidae from Bossert et al. (2019), and we include *Lanthanomelissa* as a valid genus of Tapinotaspidini (Ribeiro et al. 2021).

2.2. Data analysis

We generated separate FASTA files for each genus represented in the datasets for the barcoding gap analyses. The R package Spider (Brown et al. 2012) was used to estimate pairwise uncorrected p-distances for all sampled sequences within each genus. We opted to use uncorrected p-distances because they yield better or similar results in distance-based analyses when compared to other models of nucleotide substitution (e.g., K2P; see Collins et al. 2012; Srivathsan and Meier 2012). Intra- and interspecific distances of each genus were visualized in a boxplot. Boxplots are handy tools for data visualization: the line that divides the box into two parts represents the median of the data; box ends show the upper (Q3) and lower (Q1) quartiles; whiskers extend to Q3 + 1.5xIQR and Q1-1.5xIQR; dots show outlier values (McGill et al. 1978).

Based on the boxplots obtained for each genus, we followed Badotti et al. (2017) to sort *cox1* efficacy into three categories: *good*, *intermediate*, and *poor*. Efficacy was considered *good* when whiskers displayed a gap between intra- and interspecific comparisons, *intermediate* whenever the whiskers of intra- and interspecific comparisons overlapped, and *poor* when the boxes overlapped. Moreover, we used the function localMinima() implemented in Spider to set a threshold value for the scrutinized genera that could serve as a reference in future DNA barcoding studies. This function optimizes a

putative threshold value based on a gap in the density plot of genetic distances, disregarding sequence labels (Brown et al. 2012). Additionally, we assessed the number of informative characters of full-length and mini-barcodes for each genus using the function pis() of the R package ips (Heibl et al. 2019). Last, a Pearson Correlation (function cor.test() in base R) was used to examine the relationship between the number of informative characters of each marker.

We compared identification success between full-length and mini-barcodes by calculating the Probability of Correct Identification (PCI). We specifically adopted here the PCI metrics classified by Erickson et al. (2008) as "discrete species assignment", which considers the maximum intraspecific distance and the minimum interspecific distance (nearest-neighbor distance) for each species. Identification of species was considered successful if the maximum intraspecific distance of a species was less than its minimum interspecific distance. Then, we calculated the PCI for each genus as the proportion of species successfully identified. If the PCI of full-length and mini-barcodes differed for a given genus, the observed proportions were converted to a 2×2 contingency table and compared with a Fisher's exact test using the function fisher.test() implemented in base R (R Core Team 2021).

3. RESULTS

Our raw dataset consisted of 18167 *cox1* sequences. After the filtering steps and maintaining only genera with viable intra- and interspecific comparisons, 14578 sequences remained (Table I). We used these sequences to compile the full-length barcode and the mini-barcode datasets. A total of 393 sequences (2.69% of the dataset) had their labels changed due to misspelled or invalid species names. These changes are appended as supplementary material with all sampled species and sequences (Online Resources 1 and 2).

Regarding species coverage, our sample encompasses five subfamilies, 25 tribes, 71 genera, and 1012 species (around 17% of valid apid species; Table I). Sequence coverage by species ranged from 1 to 1162, with 77.86% of sampled species represented by less than ten sequences. Concerning sequence abundance, our final datasets present a strong bias towards Apinae (75.49%), followed by Xylocopinae (9.62%), Nomadinae (7.68%), Eucerinae (4.64%), and Anthophorinae (2.57%) (Table I). Corbiculate tribes (Apini, Bombini, Meliponini, and Euglossini) contributed with most of the sequences for the final datasets, although absolute species richness was higher for Bombini, Euglossini, and Eucerini, respectively (Table I). Our datasets, however, lack sequence data for eight apid tribes.

For most genera, *cox1* efficacy was considered *good* for both full-length and mini-barcodes (77.02 and 71.62%, respectively). Some genera, however, displayed *intermediate* (17.57 and 25.67%) and *poor* (5.41 and 2.70%) performances. Figure 2 illustrates the barcoding gap classifications we adopted in this study. The PCI of the two markers sometimes differed among genera, with full-length barcodes performing better for some taxa—*Anthophora*, *B.* (*Bombus*), *B.* (*Megabombus*), *Euglossa*, and *Epeolus*—but worst for others, *Caenonomada*, *Diadasia*, and *Triepeolus*. Intra- and interspecific distances fluctuated considerably among genera and between markers, affecting threshold values inferred by function localMinima(). The average threshold for the full-length barcodes was 2.88%, whereas mini-barcodes displayed an average of 3.31%. Marker performance and threshold values of each genus are presented in Table II.

The number of informative sites varied among genera, ranging from 0 to 112 for mini-barcodes and 0 to 366 for full-length barcodes (Figure 3A). The relationship between the number of informative sites of the markers was positive and strong ($r = 0.968$, $p < 0.001$). However, the informativeness was heavily affected by the sample size. For instance, *Bombus* presented the highest number of informative sites for both markers and was also the genus with the highest number of sequences. In contrast, genera represented by less than five sequences (such as *Erichocis*, *Leiopodus*, *Melectoides*, and *Neolarra*) often lacked informative sites.

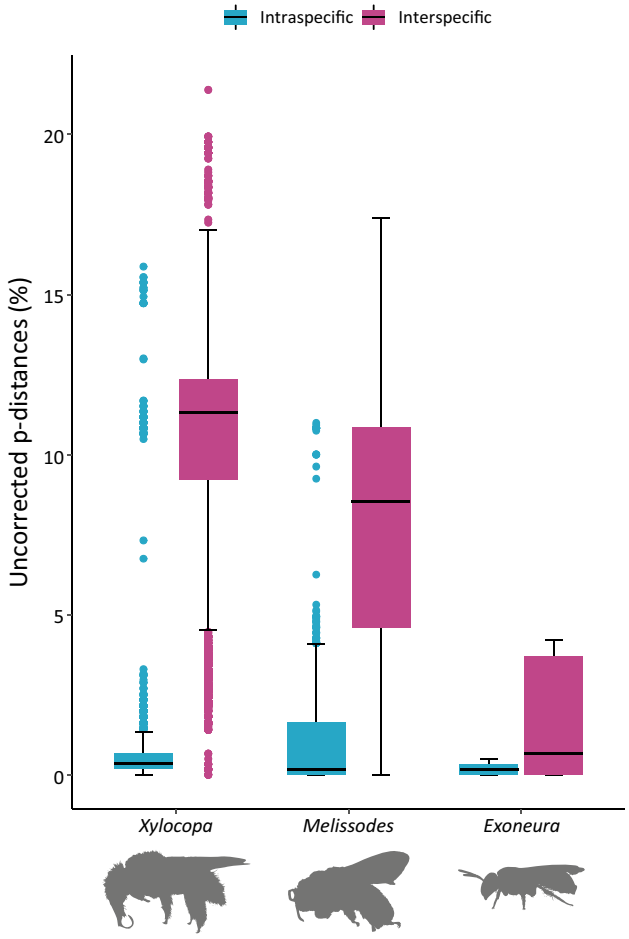

Figure 2. Examples of the barcoding gap performance classifications implemented in this study. The boxplots refer to comparisons done with the full-length dataset. *Xylocopa* displayed a *good* performance since intra- and interspecific boxes displayed a clear gap; the performance of *Melissodes* was classified as *intermediate* since the whiskers of the intra- and interspecific comparisons overlapped; intra- and interspecific boxes of *Exoneura* overlapped, implying a *poor* performance

Overall, full-length barcodes presented equal or higher PCI than mini-barcodes, whereas mini-barcodes of *Bombus (Psithyrus)* displayed a higher PCI than full-length barcodes (Table II). However, none of the differences were statistically significant (Table II). In general, PCI rates were higher for genera that displayed *good* and *intermediate* performances (Figure 3B). Full-length barcodes exhibited, on average, higher PCI rates for genera with *poor* performance. In short, our findings show that full-length barcodes and mini-barcodes have equivalent performance in bee identification.

4. DISCUSSION

Since its formal proposal in 2003, DNA barcoding has achieved the status of a revolutionary and reliable tool to support taxonomic studies (DeSalle and Goldstein 2019). Shorter DNA barcode sequences—mini-barcodes—have been conveniently employed in the taxonomic

Table II Barcoding gap, probability of correct identification (PCI), and local minima results for full-length (full) and mini-barcode (mini) datasets. When PCI differed between markers, the *p*-value of the Fisher's exact test is also shown

	Barcoding gap		PCI (%)			Local minima (%)[a]	
	Full	Mini	Full	Mini	*p*	Full	Mini
Anthophorinae							
Anthophorini							
Amegilla	Intermediate	Intermediate	71.43	71.43		2.15	2.80
Anthophora	Good	Intermediate	71.43	66.67	0.649	1.68	1.64
Habropoda	Good	Good	100.00	100.00		4.89	4.94
Apinae							
Apini							
Apis	Good	Good	62.50	62.50		0.76	0.56
Bombini							
Bombus	Good	Good	63.64	62.09	0.603	1.15	0.27
B. (*Alpigenobombus*)	Good	Good	66.67	66.67		4.48	6.26
B. (*Alpinobombus*)	Intermediate	Intermediate	42.86	42.86		1.19	1.30
B. (*Bombias*)	Good	Good	66.67	66.67		2.43	3.36
B. (*Bombus*)	Good	Intermediate	73.33	62.50	0.846	1.67	0.42
B. (*Cullumanobombus*)	Good	Good	75.00	75.00		3.91	4.46
B. (*Megabombus*)	Good	Intermediate	93.75	93.75		2.38	2.35
B. (*Melanobombus*)	Good	Good	100.00	100.00		1.01	0.29
B. (*Mendacibombus*)	Good	Good	100.00	100.00		5.82	5.43
B. (*Orientalibombus*)	Good	Good	100.00	100.00		0.52	0.27
B. (*Psithyrus*)	Good	Good	66.67	75.00	0.782	1.96	2.32
B. (*Pyrobombus*)	Good	Good	58.82	50.00	0.947	1.69	1.00
B. (*Subterraneobombus*)	Intermediate	Intermediate	80.00	80.00		0.88	0.84
B. (*Thoracobombus*)	Intermediate	Intermediate	64.29	64.29		1.86	1.64
Centridini							
Centris	Good	Good	100.00	100.00		2.38	2.83
Epicharis	Good	Good	100.00	100.00		6.04	6.93
Euglossini							
Eufriesea	Intermediate	Intermediate	35.29	29.41	0.991	0.68	0.82
Euglossa	Good	Intermediate	19.35	14.52	1.000	0.82	0.86
Eulaema	Good	Good	33.33	25.00	0.988	0.08	0.30
Exaerete	Intermediate	Intermediate	25.00	25.00		2.82	2.56
Ctenoplectrini							
Ctenoplectra	Good	Good	50.00	50.00		5.73	5.96
Emphorini							
Diadasia	Intermediate	Good	77.78	77.78		1.93	5.71
Melitoma	Good	Good	100.00	100.00		5.52	6.37
Ptilothrix	Good	Good	100.00	100.00		4.32	6.77

Table II (continued)

	Barcoding gap		PCI (%)			Local minima (%)[a]	
	Full	Mini	Full	Mini	*p*	Full	Mini
Ericrocidini							
Ericrocis	Good	Good	100.00	100.00		4.94	5.76
Hopliphora	Poor	Poor	0.00	0.00		NA	NA
Mesoplia	Good	Good	100.00	100.00		4.42	5.96
Eucerini							
Alloscirtetica	Good	Good	100.00	100.00		4.93	6.76
Eucera	Good	Good	84.21	73.68	0.810	1.59	1.42
Florilegus	Good	Good	100.00	100.00		1.57	NA
Melissodes	Intermediate	Intermediate	50.00	44.44	0.990	1.12	1.08
Svastra	Good	Good	100.00	100.00		4.77	5.39
Thygater	Intermediate	Intermediate	77.78	66.67	0.822	0.69	1.71
Exomalopsini							
Anthophorula	Good	Good	100.00	100.00		8.76	NA
Exomalopsis	Good	Good	100.00	100.00		3.56	8.59
Isepeolini							
Isepeolus	Good	Good	100.00	100.00		2.79	3.12
Melectoides	Good	Good	100.00	100.00		4.98	6.69
Melectini							
Melecta	Good	Good	100.00	100.00		4.16	5.76
Thyreus	Good	Good	100.00	100.00		3.44	4.15
Meliponini							
Cephalotrigona	Good	Good	100.00	100.00		2.89	3.99
Liotrigona	Good	Good	100.00	50.00	0.879	1.13	1.35
Melipona	Intermediate	Intermediate	23.53	11.76	0.996	0.58	5.01
Partamona	Good	Good	40.00	0.00	1.000	1.40	2.08
Plebeia	Intermediate	Intermediate	0.00	0.00		1.68	0.38
Scaptotrigona	Intermediate	Intermediate	50.00	16.67	0.984	0.47	0.30
Scaura	Good	Good	50.00	50.00		2.87	5.17
Tetragona	Good	Good	100.00	100.00		3.16	4.20
Tetragonisca	Intermediate	Intermediate	0.00	0.00		0.11	0.28
Trigona	Good	Good	28.57	28.57		3.20	3.29
Osirini							
Epeoloides	Good	Good	100.00	100.00		4.35	5.40
Protepeolini							
Leiopodus	Good	Good	100.00	100.00		4.70	4.40
Tapinotaspidini							
Arhysoceble	Good	Good	100.00	100.00		3.56	4.70
Caenonomada	Poor	Intermediate	66.67	33.33	0.929	3.46	3.65
Chalepogenus	Good	Good	100.00	100.00		3.03	4.57
Lanthanomelissa	Good	Good	100.00	100.00		1.45	1.53
Paratetrapedia	Good	Good	100.00	100.00		2.28	3.26

Table II (continued)

	Barcoding gap		PCI (%)			Local minima (%)[a]	
	Full	Mini	Full	Mini	*p*	Full	Mini
Nomadinae							
Ammobatini							
Oreopasites	Good	Good	100.00	100.00		3.99	5.12
Ammobatoidini							
Ammobatoides	Good	Good	100.00	100.00		5.51	6.05
Holcopasites	Good	Good	100.00	100.00		4.31	6.63
Biastini							
Biastes	Good	Good	100.00	100.00		3.85	3.61
Epeolini							
Epeolus	Good	Intermediate	80.00	75.00	0.802	1.87	2.54
Triepeolus	Poor	Intermediate	50.00	50.00		1.31	1.53
Neolarrini							
Neolarra	Good	Good	100.00	100.00		4.75	6.85
Nomadini							
Nomada	Good	Good	62.20	52.44	0.983	0.70	0.29
Xylocopinae							
Allodapini							
Braunsapis	Good	Good	0.00	0.00		3.32	4.28
Exoneura	Poor	Poor	0.00	0.00		2.40	0.39
Exoneurella	Good	Good	100.00	100.00		2.87	2.90
Macrogalea	Good	Good	66.67	66.67		10.20	2.63
Ceratinini							
Ceratina	Good	Good	68.75	68.75		0.71	0.96
Xylocopini							
Xylocopa	Good	Good	78.26	69.57	0.852	1.64	2.12
						2.88	3.31

[a]NA: when local minima failed to return a threshold for the given dataset

identification of ancient or degraded DNA (Meusnier et al. 2008; Cardeñosa et al. 2017; Erickson et al. 2017). A specific mini-barcode was developed for corbiculate bees (Françoso and Arias 2013) and has been applied to identify century-old museum specimens (Françoso and Arias 2013) and to assess bee biodiversity (e.g., Brettell et al. 2020; Nakamura et al. 2020). Traditional Sanger sequencing of full-length barcodes can be expensive when dealing with thousands of samples or ancient material (Yeo et al. 2020, 2021). In the genomics era, mini-barcodes with good taxonomic resolution can be employed in large-scale biodiversity studies using next-generation sequencing pipelines, costing equal or less than a morphology-based diagnostic system (Stein et al. 2014; Roe et al. 2017; Xing et al. 2021).

Here we retrieved all Apidae *cox1* sequences from BOLD to compare barcoding gap performance and identification success of full-length barcodes and mini-barcodes. We detected barcoding gaps for most of the genera and both markers, suggesting that these markers perform equally well in species

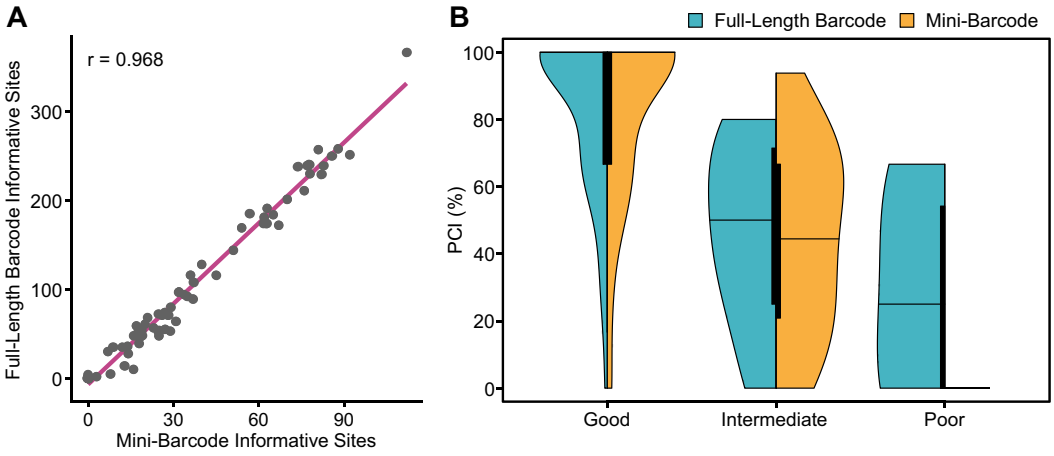

Fig. 3 **A** Relationship between the number of informative sites of full-length and mini-barcodes. Each point refers to a genus of the dataset. The solid line represents the regression line between variables. **B** Violin plot comparing identification success between full-length and mini-barcodes, sorted by barcoding gap categories. Lines show the medians; box limits indicate the 25th and 75th percentiles; colored areas extend 1.5 times the interquartile range from the 25th and 75th percentiles

discovery for apid bees—that is, when one of the barcode regions was considered *good*, the other region displayed the same performance. Furthermore, our results indicate that both markers also perform similarly concerning identification success, showing that mini-barcodes are a reliable supplementary tool to specimen identification and species delimitation considering a broader taxonomic scope. Here we also provide barcode threshold values for 78 bee genera that can aid future taxonomic projects (Table II), which can help determine clusters for further assessment with additional methods.

4.1. BOLD overview

Several DNA barcoding projects have been conducted for pollinators in the past years, especially in North America and Europe (e.g., Sheffield et al. 2009; Magnacca and Brown 2012; Packer and Ruz 2017). The available data on BOLD is overall biased towards corbiculate Apidae, often recognized as the commercially most important group of bees (Martins et al. 2014). Sequences belonging to *Apis* (Apini) and *Bombus* (Bombini) represented more than half of the dataset. These genera are widely studied due to their economic importance in crop pollination and honey production. Because of this commercial relevance, such taxa tend to be well characterized in reference databases (Virgilio et al. 2010).

After the filtration steps, eight apid tribes were absent in our dataset, reflecting the paucity of data for some taxa (Table I). These tribes comprise around 100 solitary and cleptoparasitic species. Even though most known bee species are solitary (Michener 2007), genetic and molecular studies usually focus on social species (Neumann and Seidelmann 2006), leading to an underrepresentation of solitary bee species in public barcode databases. Identifying solitary bees can be a nightmare even for experienced taxonomists because these species are usually small and exhibit cryptic morphology (Magnacca and Brown 2012). DNA barcoding can be a valuable tool to aid in specimen identification and species delimitation in this scenario. Since solitary bees also play a significant role in pollination services, DNA barcoding efforts focused on these species may also be fundamental to studying macroecological phenomena such as the worldwide decline in pollinators and animal-pollinated plant species (Pornon et al. 2017; Vamosi et al. 2017).

We detected 393 sequences (2.7% of the dataset) with invalid or misspelled species names during the filtration process. Although this number is somewhat concerning, previous surveys of BOLD sequences found a much higher proportion of invalid or misspelled names for other taxa (e.g., around 12% for true bugs; Bianchi and Gonçalves 2021a). A reliable reference database is fundamental to identifying specimens using DNA barcoding (DeSalle and Goldstein 2019), and incorrect taxonomy will inevitably hinder the effectiveness of this tool. Since identification errors are inherent to any public DNA repository (Meiklejohn et al. 2019; Bianchi and Gonçalves 2021b), data from these sources must be used with caution.

4.2. Barcoding gaps and taxonomic inconsistencies

As shown by the boxplots and the local minima analyses, barcoding gap values varied widely among the scrutinized genera, which could be explained by the different coalescence times of each lineage (Fujita et al. 2012). From a single-locus point of view, a recurrent debate is that the evolutionary story of a gene (like *cox1*) does not necessarily depict the evolution of the species (Knowles 2009). Furthermore, evolutionary events such as introgression, incomplete lineage sorting, heteroplasmy, and hybridization may further hinder single-locus approaches like traditional DNA barcoding (Moritz and Cicero 2004; Magnacca and Brown 2010). Inconsistencies that may arise from these biological factors may be mitigated, for instance, by multi-marker barcoding approaches (e.g., Cruaud et al. 2017). However, a myriad of operational biases—that is, non-biological factors—may be much more relevant to undermine DNA barcoding effectiveness, including (but not limited to) inaccurate reference taxonomy, misidentifications, spelling errors, contamination, and low-quality sequences (Mutanen et al. 2016). Since a gap between intra- and interspecific distances does not necessarily imply correct identification in DNA barcoding studies (see Collins and Cruickshank 2012), we separately evaluated identification success by calculating the PCI.

Most genera showed high PCI and a *good* barcoding gap performance for both full-length and mini-barcodes. This pattern is consistent with previous studies using similar metrics to evaluate barcode efficiencies for other taxonomic groups such as fungi (Badotti et al. 2017), nematodes (Gonçalves et al. 2021), and true bugs (Bianchi and Gonçalves 2021a). However, genera with *intermediate* or *poor* performance—which overall exhibited lower PCI—require special attention as these results may hint at operational biases and taxonomic inconsistencies. Although our objective here was not to discuss taxonomic details about the sampled taxa, we bring below some examples.

Melissodes (Eucerini) is a diverse genus of solitary bees (ca. 129 species) whose taxonomic determination at the species level is challenging (Wright et al. 2020). Moreover, identification keys for these species often refer to color and chaetotaxy (e.g., LaBerge 1956a, b), morphological characters that are not always available and are prone to deterioration in archived specimens. Thus, sequences generated for this genus are susceptible to misidentification in public databases due to problems in reference taxonomy, which may explain the PCI around 50%. Moreover, we found an *intermediate* performance in the barcoding gap analyses for *Melissodes* (Figure 2), and several of the sequences of this genus were labeled with invalid names (Table S2). These results may flag a relatively high number of misidentifications in BOLD and show that the taxonomic identity of sequences from this genus should be verified and analyzed for unknown cryptic diversity.

Exoneura (Allodapini), although common in temperate parts of Australia, is characterized by its intricate and volatile taxonomy (Michener 2007; Yong et al. 2020). Accordingly, *Exoneura* was one of the few genera that displayed a *poor* barcoding gap performance in our analysis, with a PCI of 0%. Although these results are partly explained by the low sample size (see Meyer and Paulay 2005), they emphasize the need for a taxonomic revision and deeper investigations concerning specimen identification

and species delimitation of this genus. To our knowledge, there is no recent taxonomic review for *Exoneura*, which could be one of the reasons behind the poor performance. Therefore, we believe that genera with *poor* barcoding gap and low PCI values, such as *Exoneura*, should be prioritized in future taxonomic works. As shown in multiple studies, DNA barcodes can be valuable resources in integrative approaches to unravel cryptic species complexes (e.g., Sheffield et al. 2017; Milam et al. 2020; Williams 2021).

4.3. Comparing barcodes: the longer, the better?

Although mini-barcodes represent roughly 25% of the canonical barcode region, we found that they still maintain significant information, explaining why the performance and identification success of full-length and mini-barcodes are very similar. Overall, both markers showed a high number of informative sites, and as new sequences become available for each genus, this number will inevitably increase. These results emphasize the remarks of Françoso and Arias (2013) about this mini-barcode region, reaffirming its effectiveness for specimen identification. Moreover, because of the greater relative genetic divergence of mini-barcodes, we found that their threshold values tend to be higher.

Since our work focused on the in silico evaluation of the barcodes, studies implementing other characters (such as morphology) must validate our results. However, our findings reinforce the utility of mini-barcodes in bee taxonomy, especially in scenarios where the amplification of full-length barcodes is compromised due to DNA degradation (e.g., archived specimens and environmental DNA). Moreover, genetic methods of specimen identification and species delimitation are prominent tools for discovering putative cryptic species; multiple evidence approaches are then necessary to corroborate these findings.

Studies related to bee identification are crucial to their conservation and to maintain ecosystem services (Zayed 2009; Lozier and Zayed 2017). DNA barcodes can be valuable allies to unveil cryptic diversity, flag taxonomic inconsistencies, and improve species discovery (DeSalle and Goldstein 2019; Bianchi and Gonçalves 2021a). By scrutinizing *cox1* sequences retrieved from BOLD, we showed that both full-length and mini-barcodes could be successfully employed in bee identification. We believe the reported results and analyses may help researchers identify species groups needing taxonomic revision as the first step in an integrative taxonomy workflow. Our findings may aid future research concerning apid bee diversity, taxonomy, and systematics.

SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at https://doi.org/10.1007/s13592-022-00958-x.

ACKNOWLEDGEMENTS

We thank Dr. Filipe Michels Bianchi for the feedback on an earlier draft of the manuscript and two anonymous reviewers for their valuable comments. LTG is supported by a doctoral fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

AUTHOR CONTRIBUTION

LTG and MD conceptualized the research; LTG conducted data analyses and wrote the manuscript draft; LTG, EF, and MD reviewed and edited the manuscript; MD supervised the research. All authors read and approved the final manuscript.

AVAILABILITY OF DATA AND MATERIAL

All sequences analyzed during this study are publicly available at the Barcode of Life Data System (https://www.boldsystems.org/). Accession numbers for these sequences are available as Online Resources.

CODE AVAILABILITY

Not applicable.

DECLARATIONS

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

Competing interests The authors declare no competing interests.

REFERENCES

Badotti F, de Oliveira FS, Garcia CF et al (2017) Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). BMC Microbiol 17:42. https://doi.org/10.1186/s12866-017-0958-x

Bianchi FM, Gonçalves LT (2021a) Borrowing the Pentatomomorpha tome from the DNA barcode library: scanning the overall performance of *cox1* as a tool. J Zool Syst Evol Res 59:992–1012. https://doi.org/10.1111/jzs.12476

Bianchi FM, Gonçalves LT (2021b) Getting science priorities straight: how to increase the reliability of specimen identification? Biol Lett 17:rsbl.2020.0874, 20200874. https://doi.org/10.1098/rsbl.2020.0874

Blasco-Lavilla N, Ornosa C, Michez D, De la Rúa P (2019) Contrasting patterns of genetic and morphological diversity in the bumblebee *Bombus lucorum* (Hymenoptera: Apidae: *Bombus*) along a European gradient. J Insect Conserv 23:933–943. https://doi.org/10.1007/s10841-019-00178-2

Bossert S, Murray EA, Almeida EAB et al (2019) Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Mol Phylogenet Evol 130:121–131. https://doi.org/10.1016/j.ympev.2018.10.012

Boustani M, Rasmont P, Dathe HH et al (2021) The bees of Lebanon (Hymenoptera: Apoidea: Anthophila). Zootaxa 4976:1–146. https://doi.org/10.11646/zootaxa.4976.1.1

Brettell LE, Riegler M, O'Brien C, Cook JM (2020) Occurrence of honey bee-associated pathogens in *Varroa*-free pollinator communities. J Invertebr Pathol 171:107344. https://doi.org/10.1016/j.jip.2020.107344

Brown SDJ, Collins RA, Boyer S et al (2012) Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol Ecol Resour 12:562–565. https://doi.org/10.1111/j.1755-0998.2011.03108.x

Bustamante T, Fuchs S, Grünewald B, Ellis JD (2021) A geometric morphometric method and web application for identifying honey bee species (*Apis* spp.) using only forewings. Apidologie 52:697–706. https://doi.org/10.1007/s13592-021-00857-7

Cardeñosa D, Fields A, Abercrombie D et al (2017) A multiplex PCR mini-barcode assay to identify processed shark products in the global trade. PLoS ONE 12:e0185368. https://doi.org/10.1371/journal.pone.0185368

Collins RA, Boykin LM, Cruickshank RH, Armstrong KF (2012) Barcoding's next top model: an evaluation of nucleotide substitution models for specimen identification. Methods Ecol Evol 3:457–465. https://doi.org/10.1111/j.2041-210X.2011.00176.x

Collins RA, Cruickshank RH (2012) The seven deadly sins of DNA barcoding. Mol Ecol Resour 13:969–975. https://doi.org/10.1111/1755-0998.12046

Cruaud P, Rasplus J-Y, Rodriguez LJ, Cruaud A (2017) High-throughput sequencing of multiple amplicons for barcoding and integrative taxonomy. Sci Rep 7:41948. https://doi.org/10.1038/srep41948

DeSalle R, Goldstein P (2019) Review and interpretation of trends in DNA barcoding. Front Ecol Evol 7:302. https://doi.org/10.3389/fevo.2019.00302

Erickson DL, Reed E, Ramachandran P et al (2017) Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. AoB PLANTS 9:plx015. https://doi.org/10.1093/aobpla/plx015

Erickson DL, Spouge J, Resch A et al (2008) DNA barcoding in land plants: developing standards to quantify and maximize success. Taxon 57:1304–1316. https://doi.org/10.1002/tax.574020

Folmer O, Black M, Hoeh W et al (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Françoso E, Arias MC (2013) Cytochrome c oxidase I primers for corbiculate bees: DNA barcode and mini-barcode. Mol Ecol Resour 13:844–850. https://doi.org/10.1111/1755-0998.12135

Fujita MK, Leaché AD, Burbrink FT et al (2012) Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol Evol 27:480–488. https://doi.org/10.1016/j.tree.2012.04.012

Gonçalves LT, Bianchi FM, Deprá M, Calegaro-Marques C (2021) Barcoding a can of worms: testing *cox1* performance as a DNA barcode of Nematoda. Genome 64:705–717. https://doi.org/10.1139/gen-2020-0140

González-Vaquero RA, Roig-Alsina A, Packer L (2016) DNA barcoding as a useful tool in the systematic study of wild bees of the tribe Augochlorini (Hymenoptera: Halictidae). Genome 59:889–898. https://doi.org/10.1139/gen-2016-0006

Hajibabaei M, Smith MA, Janzen DH et al (2006) A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes 6:959–964. https://doi.org/10.1111/j.1471-8286.2006.01470.x

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc B Biol Sci 270:313–321. https://doi.org/10.1098/rspb.2002.2218

Heibl C, Cusimano N, Krah F-S (2019) Interfaces to Phylogenetic Software in R. https://CRAN.R-project.org/package=ips

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization.

Brief Bioinform 20:1160–1166. https://doi.org/10.1093/bib/bbx108

Knowles LL (2009) Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. Syst Biol 58:463–467. https://doi.org/10.1093/sysbio/syp061

LaBerge WE (1956a) A revision of the bees of the genus *Melissodes* in North and Central America. Part I (Hymenoptera, Apidae). Univ Kans Sci Bull 37:911–1194

LaBerge WE (1956b) A revision of the bees of the genus *Melissodes* in North and Central America. Part II (Hymenoptera, Apidae). Univ Kans Sci Bull 38:533–578

Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278. https://doi.org/10.1093/bioinformatics/btu531

Lozier JD, Zayed A (2017) Bee conservation in the age of genomics. Conserv Genet 18:713–729. https://doi.org/10.1007/s10592-016-0893-7

Magnacca KN, Brown MJ (2010) Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus (Nesoprosopis)* bees (Hymenoptera: Colletidae). BMC Evol Biol 10:174. https://doi.org/10.1186/1471-2148-10-174

Magnacca KN, Brown MJF (2012) DNA barcoding a regional fauna: Irish solitary bees. Mol Ecol Resour 12:990–998. https://doi.org/10.1111/1755-0998.12001

Martinet B, Lecocq T, Brasero N et al (2019) Integrative taxonomy of an arctic bumblebee species complex highlights a new cryptic species (Apidae: *Bombus*). Zool J Linn Soc 187:599–621. https://doi.org/10.1093/zoolinnean/zlz041

Martins AC, Melo GAR, Renner SS (2014) The corbiculate bees arose from New World oil-collecting bees: implications for the origin of pollen baskets. Mol Phylogenet Evol 80:88–94. https://doi.org/10.1016/j.ympev.2014.07.003

McGill R, Tukey JW, Larsen WA (1978) Variations of box plots. Am Stat 32:12. https://doi.org/10.2307/2683468

Meiklejohn KA, Damaso N, Robertson JM (2019) Assessment of BOLD and GenBank – their accuracy and reliability for the identification of biological materials. PLoS ONE 14:e0217084. https://doi.org/10.1371/journal.pone.0217084

Meusnier I, Singer GA, Landry J-F et al (2008) A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9:214. https://doi.org/10.1186/1471-2164-9-214

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 3:e422. https://doi.org/10.1371/journal.pbio.0030422

Michener CD (2007) The bees of the world, 2nd edn. Johns Hopkins University Press, Baltimore

Milam J, Johnson DE, Andersen JC et al (2020) Validating morphometrics with DNA barcoding to reliably separate three cryptic species of *Bombus* Cresson (Hymenoptera: Apidae). Insects 11:669. https://doi.org/10.3390/insects11100669

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. PLoS Biol 2:e354. https://doi.org/10.1371/journal.pbio.0020354

Mutanen M, Kivelä SM, Vos RA et al (2016) Species-level para- and polyphyly in DNA barcode gene trees: strong operational bias in European Lepidoptera. Syst Biol 65:1024–1040. https://doi.org/10.1093/sysbio/syw044

Nakamura S, Yamamoto S, Sawamura N et al (2020) Pollination effectiveness of European honeybee, *Apis mellifera* (Hymenoptera: Apidae), in an Oriental persimmon, *Diospyros kaki* (Ericales: Ebenaceae), orchard. Appl Entomol Zool 55:405–412. https://doi.org/10.1007/s13355-020-00696-5

Neumann K, Seidelmann K (2006) Microsatellites for the inference of population structures in the Red Mason bee *Osmia rufa* (Hymenoptera, Megachilidae). Apidologie 37:75–83. https://doi.org/10.1051/apido:2005060

Nogueira DS, De Oliveira FF, De Oliveira ML (2022) Revision of the *Tetragona clavipes* (Fabricius, 1804) species-group (Hymenoptera: Apidae: Meliponini). Zootaxa 5119:1–64. https://doi.org/10.11646/zootaxa.5119.1.1

Packer L, Ruz L (2017) DNA barcoding the bees (Hymenoptera: Apoidea) of Chile: species discovery in a reasonably well known bee fauna with the description of a new species of *Lonchopria* (Colletidae). Genome 60:414–430. https://doi.org/10.1139/gen-2016-0071

Pornon A, Andalo C, Burrus M, Escaravage N (2017) DNA metabarcoding data unveils invisible pollination networks. Sci Rep 7:16828. https://doi.org/10.1038/s41598-017-16785-5

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Ratnasingham S, Hebert PDN (2007) BOLD: the Barcode of Life Data System (http://www.barcodinglife.org). Mol Ecol Notes 7:355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Ribeiro TMA, Martins AC, Silva DP, Aguiar AJC (2021) Systematics of the oil bee genus *Lanthanomelissa* (Apidae: Tapinotaspidini) and its implications for the biogeography of South American grasslands. J Zool Syst Evol Res 59:1013–1027. https://doi.org/10.1111/jzs.12472

Roe A, Dupuis J, Sperling F (2017) Molecular dimensions of insect taxonomy in the Genomics Era. In: Foottit RG, Adler PH (eds) Insect Biodiversity. John Wiley & Sons, Ltd, Chichester, UK, pp 547–573

Schaller A, Roig-Alsina A (2021) A revision of the bee genus *Ancyloscelis* (Hymenoptera: Apidae) in

Argentina. Zootaxa 4980:521–540. https://doi.org/10.11646/zootaxa.4980.3.4

Schmidt S, Schmid-Egger C, Morinière J et al (2015) DNA barcoding largely supports 250 years of classical taxonomy: identifications for Central European bees (Hymenoptera, Apoidea partim). Mol Ecol Resour 15:985–1000. https://doi.org/10.1111/1755-0998.12363

Schnell IB, Fraser M, Willerslev E, Gilbert MTP (2010) Characterisation of insect and plant origins using DNA extracted from small volumes of bee honey. Arthropod-Plant Interact 4:107–116. https://doi.org/10.1007/s11829-010-9089-0

Sheffield CS, Hebert PDN, Kevan PG, Packer L (2009) DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Mol Ecol Resour 9:196–207. https://doi.org/10.1111/j.1755-0998.2009.02645.x

Sheffield CS, Heron J, Gibbs J et al (2017) Contribution of DNA barcoding to the study of the bees (Hymenoptera: Apoidea) of Canada: progress to date. Can Entomol 149:736–754. https://doi.org/10.4039/tce.2017.49

Simon C, Frati F, Beckenbach A et al (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87:651–701. https://doi.org/10.1093/aesa/87.6.651

Srivathsan A, Meier R (2012) On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics 28:190–194. https://doi.org/10.1111/j.1096-0031.2011.00370.x

Stein ED, Martinez MC, Stiles S et al (2014) Is DNA barcoding actually cheaper and faster than traditional morphological methods? Results from a survey of freshwater bioassessment efforts in the United States. PLoS ONE 9:e95525. https://doi.org/10.1371/journal.pone.0095525

Vamosi JC, Gong Y-B, Adamowicz SJ, Packer L (2017) Forecasting pollination declines through DNA barcoding: the potential contributions of macroecological and macroevolutionary scales of inquiry. New Phytol 214:11–18. https://doi.org/10.1111/nph.14356

Virgilio M, Backeljau T, Nevado B, Meyer MD (2010) Comparative performances of DNA barcoding across insect orders. BMC Bioinformatics 11:. https://doi.org/10.1186/1471-2105-11-206

Williams PH (2021) Not just cryptic, but a barcode bush: PTP re-analysis of global data for the bumblebee subgenus *Bombus s. str.* supports additional species (Apidae, genus *Bombus*). J Nat Hist 55:271–282. https://doi.org/10.1080/00222933.2021.1900444

Williams PH, Brown MJF, Carolan JC et al (2012) Unveiling cryptic species of the bumblebee subgenus *Bombus s. str.* worldwide with COI barcodes (Hymenoptera: Apidae). Syst Biodivers 10:21–56. https://doi.org/10.1080/14772000.2012.664574

Williams PH, Cameron SA, Hines HM et al (2008) A simplified subgeneric classification of the bumblebees (genus *Bombus*). Apidologie 39:46–74. https://doi.org/10.1051/apido:2007052

Wright KW, Miller KB, Song H (2020) A molecular phylogeny of the long-horned bees in the genus *Melissodes* Latreille (Hymenoptera: Apidae: Eucerinae). Insect Syst Evol 1–16. https://doi.org/10.1163/1876312X-bja10015

Xing R-R, Hu R-R, Wang N et al (2021) Authentication of sea cucumber products using NGS-based DNA mini-barcoding. Food Control 129:108199. https://doi.org/10.1016/j.foodcont.2021.108199

Yeo D, Srivathsan A, Meier R (2020) Longer is not always better: optimizing barcode length for large-scale species discovery and identification. Syst Biol 69:999–1015. https://doi.org/10.1093/sysbio/syaa014

Yeo H, Yeoh TX, Ding H et al (2021) Morphology and mini-barcodes: the inclusion of larval sampling and NGS-based barcoding improves robustness of ecological analyses of mosquito communities. J Appl Ecol 1365–2664:13966. https://doi.org/10.1111/1365-2664.13966

Yong DL, Barton PS, Okada S et al (2020) Conserving focal insect groups in woodland remnants: the role of landscape context and habitat structure on cross-taxonomic congruence. Ecol Indic 115:106391. https://doi.org/10.1016/j.ecolind.2020.106391

Zayed A (2009) Bee genetics and conservation. Apidologie 40:237–262. https://doi.org/10.1051/apido/2009026

Capítulo 2

Mitochondrial phylogenomics of bumblebees, *Bombus* (Hymenoptera: Apidae): a tale of structural variation, shifts in selection constraints, and tree discordance

Leonardo Tresoldi Gonçalves, Elaine Françoso, Maríndia Deprá

Artigo publicado em 2023 no periódico
Zoological Journal of the Linnean Society
https://doi.org/10.1093/zoolinnean/zlad178

Resumo

O DNA mitocondrial (mtDNA) de mamangavas (*Bombus*) tem sido amplamente utilizado para estudos filogenéticos, porém sua evolução ainda é pouco explorada. Neste trabalho, reportamos uma análise abrangente de 60 mitogenomas de *Bombus*, incluindo 40 inéditos, com o objetivo de investigar a estrutura, composição, e informatividade do mtDNA de *Bombus* em um contexto filogenético. Nosso conjunto de dados confirmou a monofilia de *Bombus* e seus subgêneros, apesar de encontrarmos uma considerável discordância entre árvores em nós mais profundos, dependendo dos métodos de inferência ou da composição da matriz de dados. Quanto à estrutura dos mitogenomas, nossos resultados indicaram que genes de tRNA frequentemente sofrem rearranjos, e rearranjos únicos sugerem uma ancestralidade compartilhada entre subgêneros de *Bombus*, destacando seu potencial na classificação dos subgêneros. Além disso, nossos resultados questionam a hipótese de que mtDNA com evolução mais rápida apresenta uma taxa de rearranjos mais elevada. Por fim, investigamos mudanças no regime de seleção de genes do mtDNA de espécies parasitas sociais obrigatórias do subgênero *Psithyrus*, e constatamos que seu mtDNA evoluiu sob seleção relaxada. Nossos achados demonstram a utilidade do mtDNA para fornecer novas perspectivas sobre as relações filogenéticas, evolução e diversificação de características do genoma de *Bombus*. Também ressaltamos o potencial da mitogenômica comparada para revelar aspectos previamente desconhecidos da evolução de *Bombus*, abrindo novas oportunidades para pesquisas futuras nesta área.

Zoological Journal of the Linnean Society, 2023, XX, 1–15
https://doi.org/10.1093/zoolinnean/zlad178
Advance access publication 24 November 2023
Original Article

OXFORD

Original Article

Mitochondrial phylogenomics of bumblebees, *Bombus* (Hymenoptera: Apidae): a tale of structural variation, shifts in selection constraints, and tree discordance

Leonardo Tresoldi Gonçalves[1,2,*], Elaine Françoso[3] and Maríndia Deprá[1,2]

[1]Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[2]Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[3]Centre for Ecology, Evolution and Behaviour, Department of Biological Sciences, School of Life Sciences and the Environment, Royal Holloway University of London, Egham, TW20 0EX, UK

[*]Corresponding author. Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. E-mail: tresoldigoncalves@gmail.com

ABSTRACT

The mitochondrial DNA (mtDNA) of bumblebees (*Bombus*) has been widely used for phylogenetic studies, but its evolution is still underexplored. Here we report a comprehensive analysis of 60 bumblebee mitogenomes, including 40 newly assembled ones, to investigate bumblebee mtDNA structure, composition, and informativeness under a phylogenetic framework. Our mtDNA dataset supports the monophyly of *Bombus* and its subgenera, although we found a high degree of tree discordance in deeper nodes when using different inference methods or matrix composition. Concerning mitogenome structure, our results show that tRNA genes were often rearranged, with unique rearrangements indicating shared ancestry across bumblebee subgenera, illustrating their potential for subgeneric classification. Our results also challenge the notion that faster evolving mtDNA exhibits higher gene rearrangement rates. Finally, we explicitly assessed shifts in selection constraints of mtDNA genes in obligate social parasites of subgenus *Psithyrus* and found that their mtDNA evolved under relaxed selective constraints. Our findings show the utility of mtDNA in providing insights into bumblebee phylogenetics, evolution, and genome trait diversification. We also highlight the potential for comparative mitogenomics to uncover previously unknown aspects of bumblebee evolution, offering exciting opportunities for future research in this field.

Keywords: Apidae; evolutionary rate; Hymenoptera; mitochondrial genome; social parasitism

INTRODUCTION

Mitochondria play a pivotal role in providing energy for the cellular functions of eukaryotes. These organelles have their own genome, the mitogenome, which encodes proteins involved in ATP synthesis through oxidative phosphorylation. Insect mitogenomes are compact circular molecules that typically comprise 37 genes, including 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, and two ribosomal RNA (rRNA) genes (Wolstenholme 1992). Mitochondrial genes are widely used as markers in population genetics, phylogenetics, and taxonomic studies, particularly for insects (Cameron 2014). Apart from nucleotide polymorphisms, changes in mitochondrial gene order can also provide insights into phylogenetic relationships and genome evolution. Mitogenomes are conserved concerning gene order and content because they encode crucial proteins for eukaryotic life (Brown *et al.* 1979, Wolstenholme 1992). Thus, mitogenome rearrangements are considered rare genomic changes, which are often homoplasy-free and reflect shared ancestry (Rokas and Holland 2000). While rearrangements are typically characterized for major lineages, few studies have explored mitogenome evolution and rearrangements at lower taxonomic scopes, such as genus or subgenus.

Hymenopterans are exciting systems for studying mitochondrial DNA (mtDNA) evolution due to the high substitution and rearrangement rates observed in their mitogenomes (Dowton and Austin 1999, Dowton *et al.* 2003, Oliveira *et al.* 2008, Zheng *et al.* 2018, Françoso *et al.* 2023), particularly in tRNA gene clusters (Fig. 1) (Dowton and Austin 1999, Dowton *et al.* 2003, 2009, Françoso *et al.* 2020). Unlike other invertebrate groups, hymenopteran mitogenomes display rearrangements at lower

Received 8 June 2023; revised 6 September 2023; accepted 26 October 2023

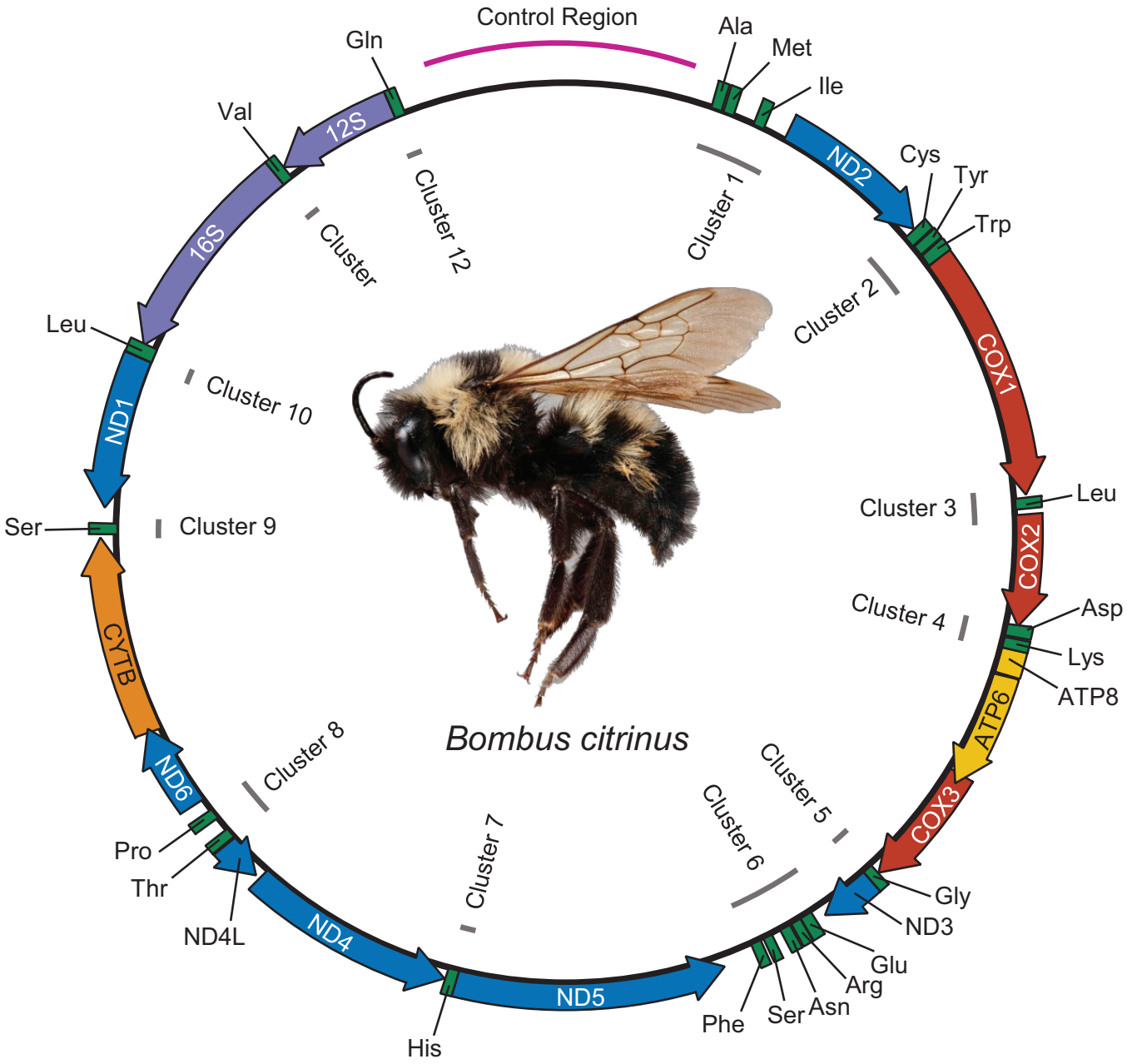

Figure 1. Typical bumblebee mitochondrial genome, using *Bombus (Psithyrus) citrinus* as an example. The arrowheads indicate the direction of protein-coding gene (PCG) transcription. Transfer RNA (tRNA) genes are represented by the three-letter IUPAC-IUB abbreviation for their corresponding amino acid. Gene sizes are proportional to their nucleotide length. Inner lines (in grey) refer to tRNA clusters, defined according to PCG junctions. Specimen image courtesy of Margarita Miklasevskaja at PCYU with funding from NSERC-CANPOLIN.

taxonomic levels, and synapomorphic rearrangements can be identified for species groups. In the case of bees, rearrangements in tRNA genes have been proposed as putative synapomorphies at the family or genus level (Françoso *et al.* 2020). For instance, a shuffle between $tRNA^{Lys}$ and $tRNA^{Asp}$ is present in all bee families when compared to a wasp outgroup; a translocation of $tRNA^{Ser1}$ and $tRNA^{Glu}$ from tRNA cluster 6 to cluster 1 was found in all analysed *Apis* species (details of taxon authors are in Table 1); and a translocation of the $tRNA^{Lys}$ from tRNA cluster 4 to cluster 1 has been identified in the stingless bee genus *Melipona* (Françoso *et al.* 2020).

Bumblebees (Hymenoptera: Apidae: *Bombus*) are major pollinators of wild flora and crops, comprising around ~280 species divided into 15 subgenera (Williams *et al.* 2022). The current phylogenetic hypothesis for *Bombus* is primarily based on five genes (four nuclear genes and one mitochondrial gene) and encompasses most known bumblebee species (Cameron *et al.* 2007, Santos Júnior *et al.* 2022). Recently, a genus-wide phylogenomic study utilizing nuclear data from 17 species supported previous topologies and refined subgeneric relationships (Sun *et al.* 2021). The majority of bumblebee species can be assigned to one of two major clades: a 'short-faced' clade (SF) and a 'long-faced' clade (LF), which broadly relate to differences in head morphology and tongue length (Cameron *et al.* 2007). While mitogenomes have been described for some bumblebee species, the evolution of mitogenomes in *Bombus* and its

Table 1. References and GenBank accession numbers for the mitochondrial genomes used in the analyses. Mitochondrial genomes that were assembled in this study are marked with an asterisk.

Species	Size (bp)	Dataset Reference	Acc. number	*
***Bombus (Alpigenobombus)* Skorikov, 1914**				
B. breviceps Smith, 1852	16 743	Zhao *et al.* (2017b)	MF478986	
B. kashmirensis Friese, 1909	16 793	Zhao *et al.* (2019)	MH998261	
***Bombus (Alpinobombus)* Skorikov, 1914**				
B. balteatus Dahlbom, 1832	18 099	Christmas *et al.* (2022)	BK063618	*
B. polaris Curtis, 1835	18 658	Sun *et al.* (2021)	BK063644	*
***Bombus (Bombias)* Robertson, 1903**				
B. confusus Schenck, 1859	16 832	Sun *et al.* (2021)	BK063651	*
B. nevadensis Cresson, 1874	18 238	Bossert *et al.* (2019)	BK063630	*
***Bombus (Bombus)* Latreille, 1802**				
B. cryptarum florilegus Panfilov, 1956	15 763	Takahashi *et al.* (2018b)	AP018158	
B. hypocrita sapporoensis Cockerell, 1911	15 468	Takahashi *et al.* (2016)	AP017370	
B. ignitus Smith, 1869	16 434	Cha *et al.* (2007)	DQ870926	
B. lantschouensis Vogt, 1908	16 153	Zhao *et al.* (2021)	BK063627	*
B. longipennis Friese, 1918	17 711	Zhou *et al.* (2021a)	MW741884	
B. lucorum (Linnaeus, 1761)	18 990	Lin *et al.* (2019a)	BK063625	*
B. terrestris (Linnaeus, 1758)	24 708	Crowley *et al.* (2023c)	OU342939	
B. terrestris canariensis Pérez, 1895	17 300	Ruiz *et al.* (2021)	MW959771	
B. terrestris lusitanicus Krüger, 1956	17 049	Cejas *et al.* (2020)	MK570128	
B. terrestris terrestris (Linnaeus, 1758)	17 232	Cejas *et al.* (2020)	MK570129	
B. terricola Kirby, 1837	20 452	Kent *et al.* (2018)	BK063649	*
***Bombus (Cullumanobombus)* Vogt, 1911**				
B. cullumanus (Kirby, 1802)	16 792	Sun *et al.* (2021)	BK063614	*
B. griseocollis (DeGeer, 1773)	17 587	Grab *et al.* (2019)	BK063647	*
***Bombus (Kallobombus)* Dalla Torre, 1880**				
B. soroeensis (Fabricius, 1777)	16 177	Sun *et al.* (2021)	BK063633	*
***Bombus (Megabombus)* Dalla Torre, 1880**				
B. consobrinus Dahlbom, 1832	17 966	Zhao *et al.* (2017a)	MF995069	
B. hortorum (Linnaeus, 1761)	21 620	Crowley *et al.* (2021)	BK063648	*
B. supremus Morawitz, 1887	19 280	Zhao *et al.* (2021)	BK063638	*
B. trifasciatus Smith, 1852	18 681	Lin *et al.* (2019a)	BK063641	*
B. ussurensis Radoszkowski, 1877	15 807	Yoon *et al.* (2020)	BK063640	*
***Bombus (Melanobombus)* Dalla Torre, 1880**				
B. ladakhensis Richards, 1928	15 877	Zhao *et al.* (2021)	BK063629	*
B. lapidarius (Linnaeus, 1758)	17 817	Tang *et al.* (2015)	KT164641	
B. pyrosoma Morawitz, 1890	18 897	Zhao *et al.* (2019)	MH998260	
B. sichelii Radoszkowski, 1860	17 165	Lin *et al.* (2019a)	BK063635	*
***Bombus (Mendacibombus)* Skorikov, 1914**				
B. convexus Wang, 1879	19 996	Lin *et al.* (2019a)	BK063615	*
B. superbus (Tkalcu, 1968)	16 855	Sun *et al.* (2021)	BK063632	*
B. waltoni Cockerell, 1910	19 349	Lin *et al.* (2019b)	MK252702	
***Bombus (Orientalibombus)* Richards, 1929**				
B. haemorrhoidalis Smith, 1852	16 595	Sun *et al.* (2021)	BK063620	*
***Bombus (Psithyrus)* Lepeletier, 1832**				
B. bohemicus Seidl, 1837	20 582	Lin *et al.* (2019a)	BK063613	*
B. campestris (Panzer, 1801)	24 740	Crowley *et al.* (2023e)	HG995151	
B. citrinus (Smith, 1854)	17 692	Bossert *et al.* (2019)	BK063616	*
B. skorikovi (Popov, 1927)	24 179	Sun *et al.* (2021)	BK063634	*
***Bombus (Pyrobombus)* Dalla Torre, 1880**				
B. bifarius Cresson, 1878	20 026	Heraghty *et al.* (2020)	BK063617	*
B. hypnorum (Linnaeus, 1758)	15 614	Crowley *et al.* (2023d)	OU427032	

Table 1. Continued

Species	Size (bp)	Dataset Reference	Acc. number	*
B. impatiens Cresson, 1863	17 161	Sadd *et al.* (2015)	BK063623	*
B. lepidus Skorikov, 1912	19 530	Zhao *et al.* (2021)	BK063626	*
B. melanopygus Nylander, 1848	18 141	Tian *et al.* (2019)	BK063624	*
B. perplexus Cresson, 1863	17 226	Grab *et al.* (2019)	BK063646	*
B. picipes Richards, 1934	18 017	Sun *et al.* (2021)	BK063636	*
B. pratorum (Linnaeus, 1761)	21 229	Crowley *et al.* (2023b)	BK063650	*
B. sylvicola Kirby, 1837	20 535	Christmas *et al.* (2021)	BK063612	*
B. vancouverensis nearcticus Handlirsch, 1888	20 554	Heraghty *et al.* (2020)	BK063643	*
B. vancouverensis vancouverensis Cresson, 1878	17 062	Ghisbain *et al.* (2020)	BK063642	*
B. vosnesenskii Radoszkowski, 1862	18 652	Heraghty *et al.* (2020)	BK063639	*
***Bombus (Sibiricobombus)* Vogt, 1911**				
B. asiaticus Morawitz, 1875	19 752	Zhao *et al.* (2019)	MH998259	
B. sibiricus (Fabricius, 1781)	20 048	Zhao *et al.* (2019)	MH998258	
***Bombus (Subterraneobombus)* Vogt, 1911**				
B. difficillimus Skorikov, 1912	16 810	Sun *et al.* (2021)	BK063619	*
B. melanurus Lepeletier, 1835	17 173	Zhao *et al.* (2021)	BK063631	*
B. personatus Smith, 1879	15 892	Zhao *et al.* (2021)	BK063645	*
***Bombus (Thoracobombus)* Dalla Torre, 1880**				
B. fervidus (Fabricius, 1798)	16 440	Grab *et al.* (2019)	BK063621	*
B. filchnerae Vogt, 1908	16 804	Zhou *et al.* (2021b)	MW741886	
B. impetuosus Smith, 1871	16 973	Lin *et al.* (2019a)	BK063622	*
B. laesus Morawitz, 1875	15 712	Lin *et al.* (2019a)	BK063628	*
B. opulentus Smith, 1861	18 218	Sun *et al.* (2021)	BK063637	*
B. pascuorum (Scopoli, 1763)	21 904	Crowley *et al.* (2023a)	HG995285	
***Apis* Linnaeus, 1758**				
A. cerana Fabricius, 1793	15 895	Tan *et al.* (2011)	NC_014295	
A. dorsata Fabricius, 1793	15 892	Chhakchhuak *et al.* (2016)	NC_037709	
A. florea Fabricius, 1787	17 694	Wang *et al.* (2013)	NC_021401	
A. mellifera sahariensis Baldensperger, 1932	16 569	Haddad *et al.* (2017)	NC_035883	
A. nigrocincta Smith, 1860	15 855	Takahashi *et al.* (2018a)	NC_038114	
A. nuluensis Tingek, Koeniger and Koeniger, 1996	15 843	Eimanifar *et al.* (2017)	NC_036235	
***Melipona* Illiger, 1806**				
M. bicolor Lepeletier, 1836	15 001	Silvestre *et al.* (2008)	AF466146	
M. fasciculata Smith, 1854	14 753	Unpublished	MH680930	
M. scutellaris Latreille, 1811	14 862	Pereira *et al.* (2016)	NC_026198	

subgenera remains unexplored under a robust phylogenetic framework. Specifically, the phylogenetic utility of mitochondrial PCGs has not been thoroughly examined, and the potential phylogenetic informativeness of mitochondrial rearrangements has not been assessed.

Bumblebees exhibit a fascinating range of ecological diversity. Most bumblebee species form social colonies with dominance hierarchies and reproductive division of labour (Free 1955). However, cuckoo bumblebees (a group of 26 species) are obligate parasites of other bumblebee species, exploiting the social structure and food resources of their hosts to rear their own brood (Lhomme and Hines 2019). Cuckoo bumblebees are classified into their own subgenus, *Psithyrus*, which is sister to the non-parasitic subgenus *Thoracobombus*. Previous studies have suggested that elevated substitution and rearrangement rates in mtDNA may be associated with parasitic life histories in Hymenoptera, particularly in wasps (Xiao *et al.* 2011, Zhu *et al.* 2018). Furthermore, such significant shifts in life history often result in changes in the selection forces that act on both nuclear and mitochondrial genomes, primarily due to reduced effective population sizes (N_e) of these species. For instance, positive selection has been observed in the mtDNA of parasitoid wasps (Oliveira *et al.* 2008), while socially parasitic ants show consistent signs of relaxing purifying selection in their nuclear genomes (Schrader *et al.* 2021). Given the contrasting life histories of *Psithyrus* and *Thoracobombus*, these sister-subgenera provide a system to investigate if obligate social parasitism also shaped the selection regime of bumblebee mitogenomes.

In this study, we conducted a comprehensive analysis of bumblebee mtDNA evolution using a dataset of 60 mitogenomes, including 40 newly assembled sequences. We evaluated the structure and composition of the mitogenomes and assessed their informativeness under a phylogenomic framework. Furthermore, we explored shifts in selection constraints in *Bombus* mitogenomes, focusing on the obligate social parasites of subgenus *Psithyrus*. Our study aimed to test the following hypotheses: (i) bumblebee mitogenomes resolve previous uncertainties while largely aligning with prior phylogenetic hypotheses, such as the monophyly of LF and SF groups; (ii) mitochondrial gene order and content are highly conserved, and gene rearrangement events are rare; (iii) unique gene rearrangements can serve as synapomorphies across bumblebee subgenera; (iv) faster-evolving mitogenomes exhibit higher rates of gene rearrangement; (v) the mitochondrial genes of *Psithyrus* exhibit positive or relaxed purifying selection. Our results provide valuable insights into bumblebee phylogenetics, evolution, and the diversification of genome traits, highlighting the utility of mitogenomes as a valuable resource in these research areas.

MATERIALS AND METHODS

Data retrieval, assembly, and annotation

We retrieved 40 publicly available Illumina paired-end datasets from the NCBI Sequence Read Archive (SRA) database, covering 39 *Bombus* species (Supporting Information, Table S1). Most of these libraries were originally prepared for whole-genome sequencing or to obtain ultra-conserved element loci (Supporting Information, Table S1). Datasets were converted to FASTQ using fastq-dump of the SRA TOOLKIT v.2.11.0 (https://trace.ncbi.nlm.nih.gov/Traces/sra/). Read quality was assessed using FastQC v.0.11.9 (Andrews 2010), and sequence adapters were trimmed with TRIMMOMATIC 0.39 (Bolger *et al.* 2014) with default parameters for Illumina data processing. The resulting reads were utilized as input for MitoFinder v.1.4 (Allio *et al.* 2020), a specialized pipeline for mitochondrial genome assembly and annotation. We employed the RefSeq mitogenomes of *B. waltoni* (NC_045283), *B. hypocrita sapporensis* (NC_011923), *B. terrestris lusitanicus* (NC_045178), and *B. terrestris terrestris* (NC_045179) as reference sequences. To ensure the quality of the assemblies and downstream analyses, we manually inspected each generated contig for all species and retained only the longest contig that contained all PCGs and rRNA genes. We also double-checked for evidence of pseudogenization or nuclear mitochondrial DNA (numt) contamination, such as premature stop codons. Annotations provided by MitoFinder were cross-checked with ARWEN (Laslett and Canbäck 2008) and MITOS2 WebServer (Bernt *et al.* 2013b).

Comparative analyses

The base composition of the mitogenomes and the pairwise *p*-distances for each gene were assessed using the base.freq() and dist.dna() functions of R package ape v.5.6-2 (Paradis and Schliep 2019), respectively. Strand asymmetry was calculated using the formulas AT skew = (A – T)/(A + T) and GC skew = (G – C)/(G + C) (Perna and Kocher 1995). Sequence divergence heterogeneity was assessed with AliGROOVE v.1.08 (Kück *et al.* 2014) with the default sliding window size.

Phylogenetic inference

Phylogenetic analyses were based on the 40 assembled mitogenomes and 20 bumblebee mitogenomes available on GenBank (Table 1), covering all 15 recognized *Bombus* subgenera (Williams *et al.* 2008). We used as outgroups the mitogenomes of six honey bee species (*Apis*) and three stingless bee species of the genus *Melipona* (Table 1). The GenBank mitogenomes were re-annotated following the procedures described above. The 13 mitochondrial PCGs and the two rRNA genes were extracted and processed separately. Stop codons were removed from the PCGs before subsequent analyses. The PCG sequences were aligned using the codon-aware program MACSE v.2.03 (Ranwez *et al.* 2018), which preserves the reading frame and prohibits indels within codons. The rRNA gene sequences were aligned using MAFFT v.7 (Katoh *et al.* 2019) with the Q-INS-i iterative refinement algorithm, which accounts for secondary RNA structure. The resulting alignments were concatenated, and ambiguously aligned fragments were removed with GBlocks v.0.91b (Talavera and Castresana 2007) using the default settings.

Three data matrices were prepared for phylogenetic analyses: PCG12RNA (first and second codon positions of the PCGs and the two rRNA genes), PCG123RNA (all codon positions combined), and a dataset of translated amino acids (AA). PartitionFinder v.2.1.1 (Lanfear *et al.* 2016) was employed to determine the optimal model for the partitioned alignments using a greedy search algorithm and Bayesian information criterion (BIC). Phylogenetic analyses were performed using the partitioned alignments and two different algorithms: Bayesian inference (BI) and maximum likelihood (ML). We performed the BI in MrBayes v.3.2.7a (Ronquist *et al.* 2012) through the CIPRES Science Gateway (Miller *et al.* 2010) with two simultaneous runs of 50 million generations, sampling trees every 5000 generations and a burn-in fraction of 0.25. We confirmed the convergence of BI runs using TRACER v.1.7.1 (Rambaut *et al.* 2018). ML trees were constructed using RAxML v.8.2.12 (Stamatakis 2014) through an ML + rapid bootstrap (BS) algorithm with 1000 replicates. Due to computational limitations, the AA dataset was exclusively analysed using the ML approach. We used the PCG123 + RNA BI tree for subsequent analyses because of the higher support values and similarity with the nuclear genome-wide phylogeny (Sun *et al.* 2021).

Gene order analysis

To investigate gene rearrangements, we registered the order of PCGs, tRNA, and rRNA genes and mapped the rearrangements on to the obtained PCG123 + RNA BI tree to visualize shared gene orders among species. Additionally, we employed qMGR (Zhang *et al.* 2020) to calculate the rearrangement score for each mitogenome, with the ancestral pancrustacean gene order serving as a reference (Lavrov *et al.* 2004). In brief, qMGR quantifies the extent of rearrangement in mitogenomes by measuring accumulated neighbour changes for each rearranged gene (Zhang *et al.* 2020). To examine whether species with fast-evolving mitogenomes exhibit higher rates of gene rearrangement, we compared the rearrangement scores of each

species with their respective root-to-tip distances. We extracted the root-to-tip distances for each species using Newick Utilities (Junier and Zdobnov 2010), serving as proxies for mitogenome evolutionary rates (Bernt *et al.* 2013a). Root-to-tip distances and rearrangement scores were standardized for comparative purposes using the scale() function available in base R (R Core Team 2021).

Selection tests

Because of the unique social parasitism of cuckoo bumblebees (subgenus *Psithyrus*), we tested whether a proportion of sites (i.e. codons) in *Psithyrus* (test branch) underwent positive selection compared with non-parasitic lineages (background branches) using the branch-site model in CODEML (Yang and Nielsen 2002, Zhang *et al.* 2005), implemented in PAML (Yang 2007). This method assesses the selective forces in the dataset through the ω value, the ratio of nonsynonymous (dN) to synonymous (dS) substitution rates, assuming that ω varies among sites in the alignments throughout the test branches of a phylogeny. We compared the alternative model (model A, which allows a subset of sites to have $\omega > 1$ in the test branch) with the null model (which applies a restriction to $\omega \leq 1$ to detect positively selected sites). Statistical significance between models was assessed using a likelihood ratio test (LRT). When the LRT was significant, we used the Bayes empirical Bayes (BEB) approach to calculate the posterior probability (PP) that individual codon sites are putatively under positive selection. We corrected significant *P*-values using a false discovery rate analysis (FDR) (Benjamini and Yekutieli 2001) implemented in base R as the function p.adjust() (R Core Team 2021), and q-values represent corrected *P*-values. We also independently ran these positive selection tests setting the non-parasitic *Thoracobombus* as the foreground to assess the extent of selection in other bumblebee lineages.

To distinguish between positive selection and relaxed purifying selection, we utilized the RELAX branch method (Wertheim *et al.* 2015) implemented in HyPhy (Pond *et al.* 2005) through the Datamonkey Adaptive Evolution Server (https://www.datamonkey.org/). RELAX compares the ω values between the background phylogeny and the lineages of interest, testing for relaxed or intensified selection using the selection intensity parameter k, where $k > 1$ indicates intensified/positive selection and $k < 1$ indicates relatively relaxed selection constraints in the test branches (Wertheim *et al.* 2015). Then, RELAX conducts an LRT to compare the alternative and null models. To confirm that selection relaxation is restricted to *Psithyrus*, we again independently assessed shifts in selection constraints setting *Thoracobombus* as a test branch in RELAX.

RESULTS

General features of novel bumblebee mitogenomes

A total of 40 bumblebee mitogenomes were assembled and annotated, representing 13 subgenera of *Bombus* (Table 1). All newly assembled mitogenomes were obtained as single contigs, with sizes ranging from 15 712 bp in *B. laesus* to 24 179 bp in *B. skorikovi*, and an average size of 18 200 bp (Table 1). These mitogenomes contained the standard set of 13 protein-coding genes (PCGs), two rRNA genes, and a control region (Fig. 1). However, eight mitogenomes lacked up to three tRNAs in cluster 1, probably due to challenges in assembling the repetitive regions flanking the control region.

Consistent with previously reported bumblebee mitogenomes, the novel sequences exhibited an AT-biased composition, with AT content ranging from 79.83% (*B. convexus*) to 89.07% (*B. pratorum*). AT-skew values were predominantly positive, indicating an excess of A over T, although 12 out of the 40 newly assembled mitogenomes showed negative AT values (Supporting Information, Table S2). GC-skew values were uniformly negative, reflecting the typical pattern observed in insect mitogenomes (Wei *et al.* 2010) (Table S2). Among the PCGs, *ATP8*, *NAD2*, and *NAD6* exhibited the highest levels of divergence, while *COI*, *COII*, and *Cytb* showed the highest conservation across the mitogenomes (Supporting Information, Fig. S1). AliGROOVE results indicated an overall absence of significant compositional heterogeneity within our dataset (Supporting Information, Fig. S2).

Phylogenetic tree

Bombus was recovered as a monophyletic group with robust support in all analyses (PP = 1.0; BS = 100; Fig. 2; Supporting Information, Figs S3–S8). Similarly, all subgenera of *Bombus* were resolved as monophyletic with strong support (PP ≥ 0.9; BS ≥ 80), except for *Cullumanobombus* in the ML trees (Supporting Information, Figs S4–S6) and the PCG12 + RNA BI tree (Supporting Information, Fig. S8). Subgenera were mainly subdivided into the two large clades, LF and SF, across all trees, confirming our initial hypothesis and the results of previous studies (Fig. 2; Supporting Information, Figs S3–S8). *Kallobombus* was consistently placed as the sister-group of the SF + LF clades (Fig. 2; Supporting Information, Figs S3–S8). Within the SF clade, the 'montane grassland' (MG) subgroup (*sensu* Williams *et al.* 2022), which includes *Alpigenobombus*, *Melanobombus*, *Sibiricobombus*, and *Cullumanobombus*, was identified as monophyletic. Likewise, the 'lowland grassland' (LG) subgroup (*sensu* Williams *et al.* 2022), comprising species from subgenus *Thoracobombus*, was also recovered as monophyletic (Fig. 2). Shallow-level relationships (i.e. among congeneric species) were consistently well-supported in all phylogenetic trees, regardless of the data matrix and reconstruction methodology used (Fig. 2; Supporting Information, Figs S3–S8).

The main discrepancies observed in the phylogenetic trees were restricted to poorly supported nodes (Fig. 3). Most trees exhibited an unresolved deeper node, placing *Mendacibombus* and *Bombias* as a polytomy sister to all bumblebee subgenera (Supporting Information, Figs S3–S8), except for the PCG123 + RNA BI tree, which supported *Mendacibombus* as the sister-group to all bumblebee subgenera (Fig. 2). In addition, the recovered trees displayed conflicting and unresolved relationships among LF subgenera, except for the placement of *Psithyrus* and *Thoracobombus* as sister-groups (Supporting Information, Figs S3–S8). Concerning the SF clade, ML trees showed unresolved subgeneric relationships (Supporting Information, Figs S3–S6), while the two BI trees recovered conflicting topologies. The PCG123 + RNA BI tree placed *Melanobombus* external to a trichotomy formed by *Sibiricobombus*, *Cullumanobombus*, and *Alpigenobombus* (Fig. 2; Supporting Information, Figs S3, S7).

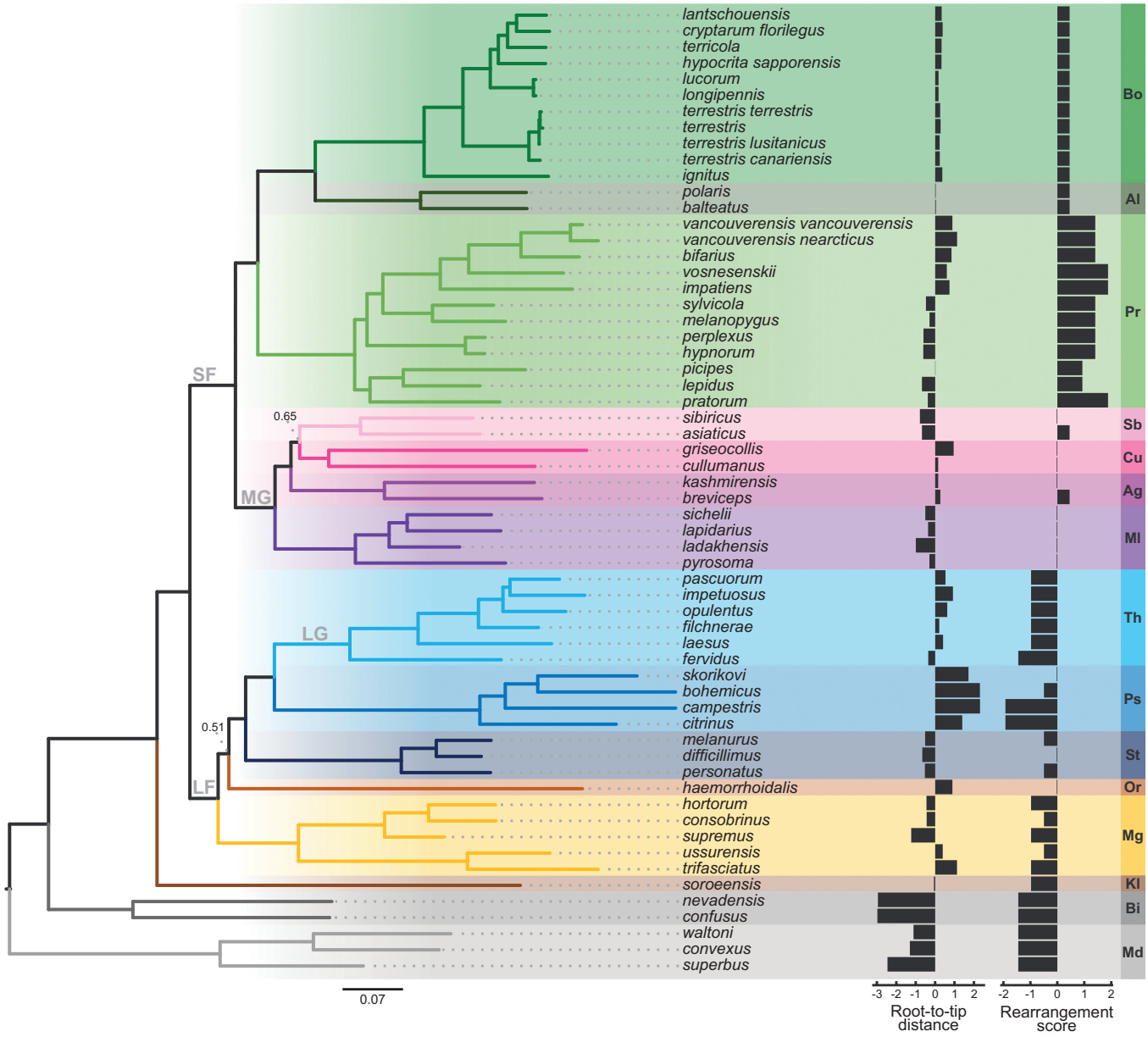

Figure 2. Phylogenetic tree of *Bombus* built based on the PCG123 + RNA dataset using Bayesian analysis. Only low node support values (PP < 0.90) are shown. The scale bar indicates estimated substitutions per site. Groups of bumblebee subgenera are labelled: LF, the 'long-faced' group; SF, the 'short-faced' group; MG, the 'montane grassland' group; LG, the 'lowland grassland' group (*sensu* Williams *et al.* 2022). Outgroups were pruned from the phylogenetic tree to improve visualization. The standardized root-to-tip distance and rearrangement score for each species are shown on the right, and the X axis represents the number of standard deviations from the mean. The subgenus that each bumblebee species belongs to is colour-coded: Bo, *Bombus*; Al, *Alpinobombus*; Pr, *Pyrobombus*; Sb, *Sibiricobombus*; Cu, *Cullumanobombus*; Ag, *Alpigenobombus*; Ml, *Melanobombus*; Th, *Thoracobombus*; Ps, *Psithyrus*; St, *Subterraneobombus*; Or, *Orientalibombus*; Mg, *Megabombus*; Kl, *Kallobombus*; Bi, *Bombias*; Md, *Mendacibombus*.

However, the PCG12 + RNA BI tree fully resolved subgeneric relationships, positioning *Melanobombus* as external to the other three subgenera and revealing a paraphyletic *Cullumanobombus* (Supporting Information, Figs S3, S8).

Gene order and rearrangements

When comparing the structure of mitogenomes, we found that PCG and rRNA gene order and orientation are conserved among bumblebee species, matching the proposed ancestral pancrustacean gene order (Boore 1999). However, we identified several tRNA gene rearrangements, including translocations, inversions, adjacent shuffling, and tandem duplications. Interestingly, the extent of rearrangement was not correlated with branch lengths (Fig. 2), indicating that species with rapidly evolving mitogenomes do not necessarily exhibit higher rearrangement rates (Supporting Information, Table S3). Species from the clade comprising subgenera *Pyrobombus*, *Alpinobombus*, and *Bombus s.s.* displayed the highest rearrangement scores, while *Psithyrus* exhibited the longest branch lengths (Fig. 2).

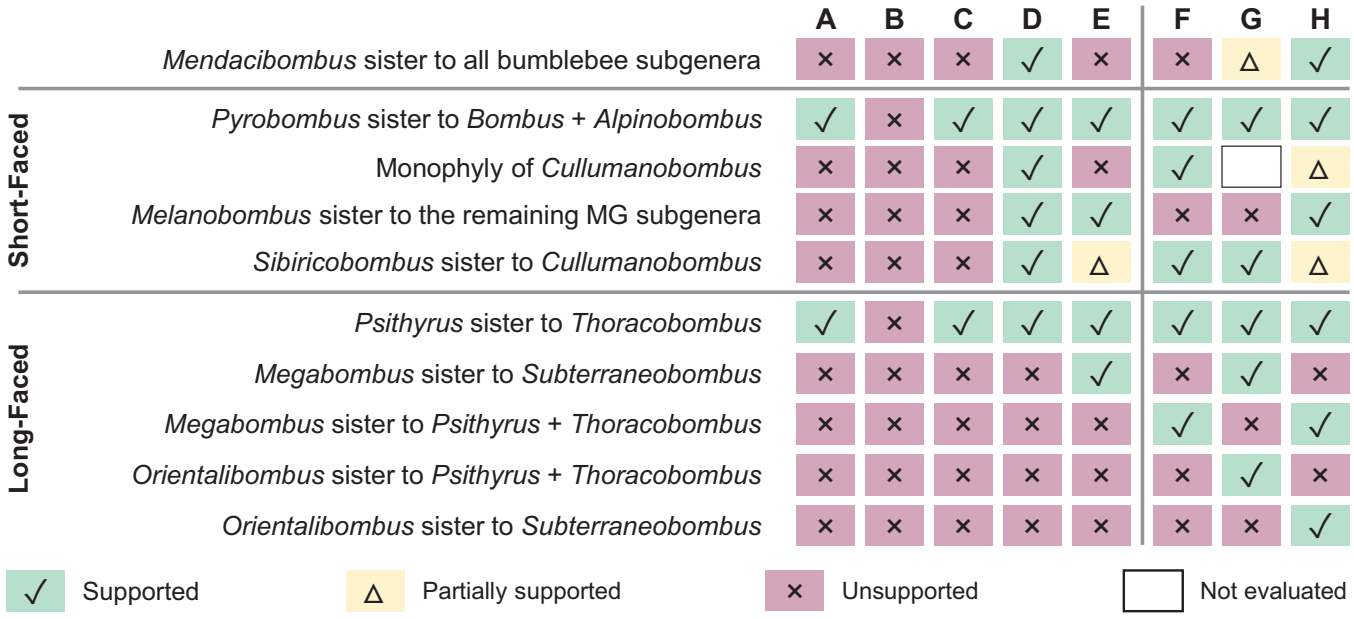

Figure 3. Major points of tree conflict in phylogenetic relationships of *Bombus* subgenera. Rows correspond to phylogenetic hypotheses, and columns correspond to the results from different datasets and methods (A–E, this study; F–H, other studies). A, maximum likelihood, PCG123 + RNA dataset; B, maximum likelihood, PCG12 + RNA dataset; C, maximum likelihood, AA dataset; D, Bayesian inference, PCG123 + RNA dataset; E, Bayesian inference, PCG12 + RNA dataset; F, topology from Cameron *et al.* (2007) using one mitochondrial (mt) and four nuclear (nu) genes; G, topology from Sun *et al.* (2021) using 2918 nu genes; H, topology from Santos Júnior *et al.* (2022) adding two mt genes to the matrix of Cameron *et al.* (2007). Nodes with PP ≥ 0.90 or BS ≥ 80 were considered supported, and partially supported nodes mainly refer to the paraphyly of *Cullumanobombus* in some of the recovered trees.

We classified the rearrangements into major and minor events based on the magnitude of changes and the number of tRNA genes affected. Major rearrangements involved significant changes in the order of tRNA genes (as in tRNA clusters 1 and 6; Fig. 1), whereas minor rearrangements entailed relatively small or punctual changes (as in tRNA clusters 2, 8, 11, and 12; Fig. 1).

In cluster 1, we found four distinct tRNA gene orders: (i) $tRNA^{Ala}$-$tRNA^{Ile}$-$tRNA^{Met}$, (ii) $tRNA^{Ala}$-$tRNA^{Met}$-$tRNA^{Ile}$, (iii) $tRNA^{Met}$-$tRNA^{Ile}$-$tRNA^{Ala}$, and (iv) $tRNA^{Met}$-$tRNA^{Ala}$-$tRNA^{Ile}$. These tRNA genes were typically located on the light strand, but we also detected inversion events for all three tRNAs (Supporting Information, Table S4). Rearrangement events within cluster 1 were homoplasic, occurring independently multiple times during the evolutionary history of bumblebees. Additionally, the copy number of $tRNA^{Met}$ varied among species, ranging from one (most species) to four (*Bombus haemorrhoidalis* and *Bombus skorikovi*) (Supporting Information, Table S4). Notably, the clade comprising *Alpinobombus* and *Bombus s.s.* shared a $tRNA^{Met}$ duplication event (Fig. 4), with the additional copy being lost in *Bombus hypocrita sapporensis* and *Bombus ignitus* (Supporting Information, Table S4). Furthermore, the available mitogenome sequences of *Bombus s.s.* in GenBank lack the annotation of the duplicated $tRNA^{Met}$.

Within cluster 6, we found three different tRNA orders: (i) $tRNA^{Arg}$-$tRNA^{Asn}$-$tRNA^{Glu}$-$tRNA^{Ser1}$-$tRNA^{Phe}$ in *Mendacibombus*, *Bombias*, and the SF clade; (ii) $tRNA^{Asn}$-$tRNA^{Arg}$-$tRNA^{Glu}$-$tRNA^{Ser1}$-$tRNA^{Phe}$ in *Kallobombus* and the LF clade (excluding *Psithyrus*); and (iii) $tRNA^{Glu}$-$tRNA^{Arg}$-$tRNA^{Asn}$-$tRNA^{Ser1}$-$tRNA^{Phe}$ exclusive to *Psithyrus* (Fig. 4). We also detected unique arrangements in *Bombus bohemicus* ($tRNA^{Glu}$-$tRNA^{Asn}$-$tRNA^{Arg}$-$tRNA^{Ser1}$-$tRNA^{Phe}$) and *Bombus fervidus* ($tRNA^{Arg}$-$tRNA^{Ser1}$-$tRNA^{Asn}$-$tRNA^{Glu}$-$tRNA^{Phe}$). In *Bombus consobrinus*, $tRNA^{Arg}$ translocated from cluster 6 to cluster 12 (Supporting Information, Table S4).

In Cluster 2, tRNA order was largely conserved across species: $tRNA^{Cys}$-$tRNA^{Tyr}$-$tRNA^{Trp}$. However, we observed a shuffling event between $tRNA^{Tyr}$ and $tRNA^{Cys}$ in *Bombus consobrinus* and *B. skorikovi*, and duplication of $tRNA^{Tyr}$ in *Bombus picipes*. Within cluster 8, the ancestral condition of bumblebee mitogenomes was $tRNA^{Thr}$-$tRNA^{Pro}$, and the shuffling between these tRNA genes was synapomorphic for the SF clade (Fig. 4). Furthermore, both subspecies of *Bombus vancouverensis* exhibited a duplication of $tRNA^{Pro}$. Lastly, $tRNA^{Val}$ translocated from cluster 11 to cluster 12 in all *Pyrobombus* species (Fig. 4).

Selection tests in *Psithyrus* and *Thoracobombus*

We employed PAML and HyPhy to investigate signs of positive or relaxed purifying selection in mitochondrial PCGs of cuckoo bumblebees. By conducting a branch-site test in PAML, we identified a significant signature of positive selection in *NAD2* and *NAD6* of *Psithyrus* (q = 0.004), with several individual sites being detected under positive selection (Supporting Information, Table S5). Utilizing RELAX, we detected evidence of selection relaxation in *COI*, *COII*, *Cytb*, *NAD5*, and *NAD6* of *Psithyrus* (Supporting Information, Table S5).

When examining signs of selection in *Thoracobombus*, the non-parasitic sister-subgenus of *Psithyrus*, we observed

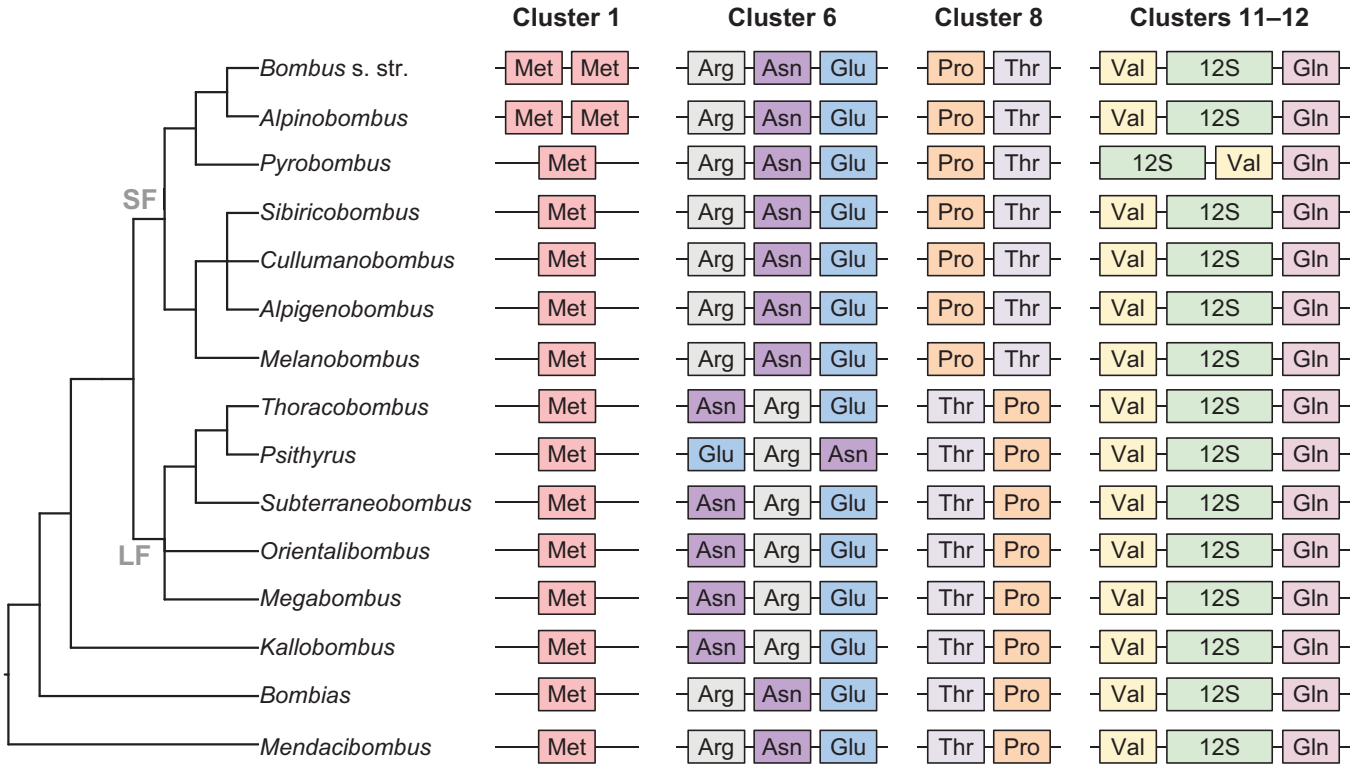

Figure 4. Putative structural synapomorphies of bumblebee mitogenomes plotted against the PCG123 + RNA BI tree. Groups of bumblebee subgenera are labelled: LF, the 'long-faced' group; SF, the 'short-faced' group. Transfer RNA (tRNA) genes are represented by the three-letter IUPAC-IUB abbreviation for their corresponding amino acid. In cluster 1, we emphasize the marked duplication event of $tRNA^{Met}$ shared by *Bombus s.s.* and *Alpinobombus*; however, this cluster comprises other tRNA genes that rearranged multiple times (see Results). In clusters 11–12, the 12S ribosomal RNA gene is also depicted.

positive selection in *NAD2* (q = 0.004) using PAML (Supporting Information, Table S5). In contrast, RELAX revealed no significant changes in the selection constraints acting upon *Thoracobombus* mitogenomes, except for an intensified purifying selection in *COIII* (P = .001) (Supporting Information, Table S5). These findings demonstrate that positive selection in *NAD2* is shared between the two subgenera, while the widespread relaxation of selection in mitochondrial genes is a characteristic feature of cuckoo bumblebees.

DISCUSSION

Phylogenetic relationships

We utilized mitochondrial phylogenomics to infer the evolutionary relationships of bumblebees, a diverse and ecologically important group of bees. Our analyses yielded well-supported phylogenetic trees that confirmed the monophyly of *Bombus* and its subgenera, supporting the currently accepted taxonomic classification (Williams *et al.* 2008). The resulting trees corroborated the presence of two major clades, the LF and SF clades, as documented in previous studies (Cameron *et al.* 2007), and the monophyly of the MG clade (Williams *et al.* 2022). Despite the limited taxon sampling for evaluating shallow nodes, our dataset provided robust support for these relationships. However, our findings revealed some discrepancies with previous studies using nuclear genes, highlighting instances of mitonuclear discordance in the phylogenetic placement of certain subgenera.

Overall, our results were consistent with the most recent phylogeny of *Bombus* subgenera inferred from 2918 nuclear loci (Sun *et al.* 2021). However, we observed two major topological differences. First, in the nuclear data phylogeny recovered by Sun *et al.* (2021), *Megabombus* and *Subterraneobombus* were positioned as sister-subgenera within the LF clade (Supporting Information, Fig. S3). In our mtDNA-based analyses, we only recovered this relationship in the PCG12 + RNA BI tree (Supporting Information, Fig. S8), while the remaining trees exhibited unresolved topologies or placed *Subterraneobombus* as sister to *Thoracobombus* + *Psithyrus* (Fig. 3; Supporting Information, Figs S4–S7). Second, Sun *et al.* (2021) identified discordant relationships within the MG clade, particularly concerning the position of *Melanobombus*. Their ASTRAL tree mirrored our findings, placing *Melanobombus* as an external group to the remaining MG subgenera (Supporting Information, Fig. S2). However, when a concatenated matrix was employed for phylogenetic reconstruction, Sun *et al.* (2021) recovered a different relationship, with (*Cullumanobombus*, *Sibiricobombus*) as sister to (*Melanobombus*, *Alpigenobombus*). None of our phylogenetic trees supported this relationship (Fig. 3; Supporting Information, Fig. S3).

Furthermore, our phylogenetic trees can be compared to the topologies obtained by Cameron *et al.* (2007), utilizing one mitochondrial and four nuclear genes, and Santos Júnior *et al.* (2022), which expanded upon the matrix of Cameron *et al.* (2007) by including two additional mitochondrial genes

(Fig. 3; Supporting Information, Fig. S3). Our phylogenetic trees exhibited slight differences compared to theirs, primarily regarding the position of *Subterraneobombus*, *Megabombus*, and *Orientalibombus*. While Cameron *et al.* (2007) and Santos Júnior *et al.* (2022) recovered *Megabombus* as the external group to *Psithyrus* and *Thoracobombus*, we found *Subterraneobombus* occupying that position in the PCG 123 + RNA BI tree (Fig. 3; Supporting Information, Fig. S7). Moreover, the placement of *Orientalibombus* and *Megabombus* within the LF clade was mostly inconclusive in our phylogenies (Fig. 3), except for the PCG12 + RNA BI tree, which showed *Subterraneobombus* and *Megabombus* as sister-subgenera (Fig. 3; Supporting Information, Fig. S8). In contrast, Santos Júnior *et al.* (2022) reported *Subterraneobombus* and *Orientalibombus* as sister-subgenera (Fig. 3; Supporting Information, Fig. S3). Interestingly, these topologies conflict with the findings of Sun *et al.* (2021) based on nuclear loci (Fig. 3; Supporting Information, Fig. S3).

In our dataset, we included two species belonging to the subgenus *Cullumanobombus*, which were recovered in all trees as part of the MG clade. However, the precise phylogenetic placement of *Cullumanobombus* within this clade varied. We identified three distinct scenarios for *Cullumanobombus*, influenced by the data matrix and methodology employed: well-supported monophyly (Supporting Information, Fig. S7), paraphyly due to low statistical support (Supporting Information, Figs S4–S6), or paraphyly with high support (Supporting Information, Fig. S8). In the latter, *B. griseocollis* was recovered external to *Sibiricobombus*, and *B. cullumanus* was placed external to *Alpigenobombus* (Supporting Information, Fig. S8). Notably, Santos Júnior *et al.* (2022) also encountered instances of paraphyletic *Cullumanobombus* in their analyses due to low statistical support, while Sun *et al.* (2021) did not assess the monophyly of this subgenus. Considering our restricted sampling and the relatively weak signal provided by mtDNA for understanding the phylogenetic relationships of the MG clade, a comprehensive investigation using an expanded dataset is necessary to test the monophyly of *Cullumanobombus*.

The taxonomy and systematics of *Bombus* have long puzzled biologists (Moure and Sakagami 1962), and our findings highlight the challenges in inferring the phylogenetic relationships of this group. The mitonuclear discordance and the sensibility to the chosen method emphasize the complexity of unravelling evolutionary relationships in *Bombus*, underscoring the importance of considering multiple data sources in bumblebee phylogenetics. Our results demonstrate that mitogenomes contribute to resolving shallower phylogenetic relationships within *Bombus* but have limited power in disentangling the deeper nodes of the bumblebee tree of life.

Gene rearrangements

We investigated the gene rearrangements in bumblebee mitogenomes and found that PCG and rRNA gene order remained conserved, while rearrangements in tRNA genes were frequent. Interestingly, we observed recurrent rearrangements in tRNA clusters 1 and 6, which aligns with previous studies conducted on other hymenopteran species (Dowton and Austin 1999, Oliveira *et al.* 2008, Mao *et al.* 2015, Françoso *et al.* 2020). Furthermore, unique rearrangements were observed in certain subgenera or subgeneric groups (Fig. 4), offering valuable insights into the evolutionary history of bumblebees and providing additional characters for subgeneric taxonomy.

Insect mitogenome rearrangements are commonly explained by the tandem duplication and random loss (TDRL) model, involving the duplication of a contiguous gene set followed by the random loss of one copy of each duplicated gene (Boore 2000). Here we identified mitogenomes with tandem duplicated tRNA genes, which could indicate ongoing TDRL events. However, some rearrangements are inconsistent with the TDRL model, such as the long-range translocations or inversion of tRNA genes. It has been proposed that intramitochondrial recombination may drive this type of rearrangement in invertebrate mitogenomes, notably in hymenopterans (Dowton and Austin 1999, Mao *et al.* 2014, 2015). Furthermore, rearrangements by slipped-strand mispairing are prone to occur in the replication origin regions (Levinson and Gutman 1987, Macey *et al.* 1998), which explains the higher frequency of rearrangements in clusters 1 and 6 that coincide with the origin of replication of the heavy and light mtDNA strands, respectively (Brown *et al.* 2005, Duarte *et al.* 2008). These tRNA clusters have been identified as regions of high rearrangement frequency in other bees, such as the tribe Meliponini (Silvestre *et al.* 2002, 2008, Wang *et al.* 2021) and genus *Tetrapedia* (Françoso *et al.* 2020). Since mitogenomic rearrangements within the same genus are generally uncommon in insects (Cameron 2014), the multiple and complex events of tRNA gene rearrangements we detected suggest a certain degree of structural plasticity in bumblebee mitogenomes.

Although tRNA gene rearrangements were common, certain positions remained conserved across all species, particularly in clusters 3–5 and 7–10. One notable exception was a single shuffling event between $tRNA^{Thr}$ and $tRNA^{Pro}$ in cluster 8, which was synapomorphic among SF bumblebees (Fig. 4). The conserved position of $tRNA^{Pro}$ in bee mitogenomes is attributed to the change in transcriptional polarity at these sites and the role these genes may play in mRNA maturation (Dowton *et al.* 2009, Françoso *et al.* 2020). This rare and conserved shuffling event might reflect the functional significance and constraints associated with the specific position of these tRNA genes, aligning with previous hypotheses (Dowton *et al.* 2003).

We did not find a direct association between the degree of rearrangement and evolutionary rates (Fig. 2), contrary to previous studies that have suggested such a relationship (Shao *et al.* 2003, Hassanin 2006, Xu *et al.* 2006, Bernt *et al.* 2013a, Zou *et al.* 2022). Although the high degree of rearrangements in bumblebees can be attributed to the inherent high substitution rates in hymenopteran mitogenomes, it remains uncertain if this relationship holds at lower taxonomic levels. For instance, our phylogeny exhibited accelerated substitution rates in cuckoo bumblebees (*Psithyrus*), as often observed in parasitic taxa (as discussed below). However, the most rearranged mitogenomes belonged to subgenus *Pyrobombus*, specifically *B. impatiens*, *B. pratorum*, and *B. vosnesenskii*. Further investigations are necessary to determine whether the high degree of rearrangement in these species is adaptive.

Besides their phylogenetic utility, mtDNA rearrangements hold crucial implications for speciation. Their capacity to disrupt

gene flow between populations may facilitate the formation of new species (Burton and Barreto 2012, Hill *et al.* 2019). This disruption stems from the coordinated functioning of mitochondrial and nuclear genes, essential for electron transport chain complexes and cellular respiration. Any disturbance to this coadaptation may create barriers to gene flow once coadapted genotypes from one population become incompatible with those of another, further contributing to speciation (Burton and Barreto 2012, Hill 2016). After speciation, these rearrangements may persist within the mtDNA of the resulting lineages, effectively acting as molecular fossils of their evolutionary history (Richardson *et al.* 2013). This explains why many of the rearrangements we observed are shared among related species, tracing back to their common ancestors. Supporting the findings for other taxa (Tan *et al.* 2019), it is likely that the causes of mitogenome rearrangements in bumblebees are multifactorial and lineage-specific, warranting additional research to unravel the underlying mechanisms.

Relaxed selection constraints in *Psithyrus*

Our findings revealed evidence of relaxed purifying selection in several mitochondrial genes of *Psithyrus* bumblebees. Purifying selection typically removes deleterious mutations and is essential for mitochondrial PCGs due to their fundamental role in ATP production (Stewart *et al.* 2008, Palozzi *et al.* 2018). The relaxed purifying selection suggests that these genes in *Psithyrus* bumblebees undergo more changes or variations than expected. This pattern could be attributed to reduced N_e and decreased functional constraints on these genes in social parasites.

Subgenus *Psithyrus* exhibited longer branch lengths in our phylogenetic trees, illustrating the accelerated substitution rates within this clade. Parasitic lineages often show accelerated mtDNA evolution compared to their non-parasitic counterparts, possibly due to reduced N_e (Castro *et al.* 2002, Jakovlić *et al.* 2021 , Oliveira *et al.* 2008). The nearly neutral theory predicts that a small N_e leads to the accumulation of slightly deleterious mutations through genetic drift, relaxing purifying selection (Ohta 1972). Since cuckoo bumblebees occupy a higher level in food webs and are quite rare in nature, their N_e is expected to be lower than their host species (Suhonen *et al.* 2015, Lhomme and Hines 2019). Thus, the reduced N_e may contribute to the relaxation of purifying selection on mitochondrial genes in *Psithyrus* bumblebees.

The relaxed purifying selection observed in these genes may also stem from decreased functional constraints. Social parasitism has independently evolved multiple times in Hymenoptera (Michener 2007) and is often associated with degenerative processes characterized by the loss of behavioural, physiological, and morphological traits (Schrader *et al.* 2021). Cuckoo bumblebees lack the pollen-collecting apparatus on their hind legs, cannot produce a worker caste, and have limited wax production for nest-building (Lhomme and Hines 2019). These losses extend to the molecular level, as *Psithyrus* bumblebees lost 11 odorant receptor genes (Sun *et al.* 2021). As obligate parasites with limited dispersal capabilities and complete reliance on their host workers for thermoregulation and foraging, cuckoo bumblebees probably experience distinct selective pressures on their mitochondrial genes compared to non-parasitic bumblebees, owing to their reduced metabolic needs. Furthermore, the selective constraints on mitochondrial genes directly impact insect mobility (Mitterboeck *et al.* 2017, Chang *et al.* 2020), and a relaxation of purifying selection may explain the slower and less energetic flight observed in cuckoo bumblebees compared to their non-parasitic counterparts (Lhomme and Hines 2019, Fisogni *et al.* 2021). Since obligate social parasitism has also arisen in other non-*Psithyrus* bumblebee species—in subgenus *Alpinobombus* with *Bombus natvigi* and *Bombus hyperboreus*, and in subgenus *Thoracobombus* with *Bombus inexpectatus* (Lhomme and Hines 2019)—an intriguing avenue for future research would be to investigate whether the relaxed purifying selection is a convergent phenomenon in the mitogenomes of these species or if it is unique to subgenus *Psithyrus*.

CONCLUSIONS

Here we provide new insights into the mtDNA evolution of bumblebees. Our findings support the monophyly of *Bombus* and its subgenera, while revealing discrepancies with nuclear DNA topologies at certain deep nodes. Moreover, we demonstrate the prevalence of mitochondrial tRNA rearrangements, which hold potential as informative markers for subgeneric classification. We found no association between rearrangement and evolutionary rates, challenging the prevailing notion that faster-evolving mitogenomes exhibit higher gene rearrangement rates. These results highlight the need for a more nuanced understanding of the factors influencing mitogenome evolution in hymenopterans. Finally, the observed relaxed selection constraints on mitochondrial genes in *Psithyrus* bumblebees provide valuable insights into the mitochondrial biology and evolutionary history of these parasitic species. Our study highlights the potential of comparative mitogenomics in uncovering previously unexplored aspects of bumblebee evolution and paves the way for exciting avenues of future research in this field.

SUPPLEMENTARY DATA

Supplementary data are available at *Zoological Journal of the Linnean Society* online.

ACKNOWLEDGEMENTS

We thank Pedro Henrique Pezzi for his valuable feedback on this manuscript. This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY

Resulting alignments, phylogenies, and scripts are available as Supporting Information. The newly assembled mitogenomes are available on GenBank under the accession numbers listed in Table 1.

REFERENCES

Allio R, Schomaker-Bastos A, Romiguier J *et al.* MitoFinder: efficient automated large-scale extraction of mitogenomic data in

target enrichment phylogenomics. *Molecular Ecology Resources* 2020;**20**:892–905. https://doi.org/10.1111/1755-0998.13160

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 2001;**29**:1165–88. https://doi.org/10.1214/aos/1013699998

Bernt M, Bleidorn C, Braband A *et al.* A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. *Molecular Phylogenetics and Evolution* 2013a;**69**:352–64. https://doi.org/10.1016/j.ympev.2013.05.002

Bernt M, Donath A, Jühling F *et al.* MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* 2013b;**69**:313–9. https://doi.org/10.1016/j.ympev.2012.08.023

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20. https://doi.org/10.1093/bioinformatics/btu170

Boore JL. Animal mitochondrial genomes. *Nucleic Acids Research* 1999;**27**:1767–80. https://doi.org/10.1093/nar/27.8.1767

Boore JL. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In: Sankoff D, Nadeau JH (eds), *Comparative Genomics*. Dordrecht: Springer Netherlands, 2000, 133–147. https://doi.org/10.1007/978-94-011-4309-7_13

Bossert S, Murray EA, Almeida EAB *et al.* Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. *Molecular Phylogenetics and Evolution* 2019;**130**:121–31. https://doi.org/10.1016/j.ympev.2018.10.012

Brown WM, George M, Wilson AC. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences* 1979;**76**:1967–71. https://doi.org/10.1073/pnas.76.4.1967

Brown TA, Cecconi C, Tkachuk AN *et al.* Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes & Development* 2005;**19**:2466–76. https://doi.org/10.1101/gad.1352105

Burton RS, Barreto FS. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? *Molecular Ecology* 2012;**21**:4942–57. https://doi.org/10.1111/mec.12006

Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology* 2014;**59**:95–117. https://doi.org/10.1146/annurev-ento-011613-162007

Cameron SA, Hines HM, Williams PH. A comprehensive phylogeny of the bumble bees (*Bombus*). *Biological Journal of the Linnean Society* 2007;**91**:161–88. https://doi.org/10.1111/j.1095-8312.2007.00784.x

Castro LR, Austin AD, Dowton M. Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera. *Molecular Biology and Evolution* 2002;**19**(7):1100–1113. https://doi.org/10.1093/oxfordjournals.molbev.a004168

Cejas D, López-López A, Muñoz I *et al.* Unveiling introgression in bumblebee (*Bombus terrestris*) populations through mitogenome-based markers. *Animal Genetics* 2020;**51**:70–7. https://doi.org/10.1111/age.12874

Cha SY, Yoon HJ, Lee EM *et al.* The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). *Gene* 2007;**392**:206–20. https://doi.org/10.1016/j.gene.2006.12.031

Chang H, Qiu Z, Yuan H *et al.* Evolutionary rates of and selective constraints on the mitochondrial genomes of Orthoptera insects with different wing types. *Molecular Phylogenetics and Evolution* 2020;**145**:106734. https://doi.org/10.1016/j.ympev.2020.106734

Chhakchhuak L, De Mandal S, Gurusubramanian G *et al.* The near complete mitochondrial genome of the Giant honey bee, *Apis dorsata* (Hymenoptera: Apidae: Apinae) and its phylogenetic status. *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis* 2016;**27**:3483–4. https://doi.org/10.3109/19401736.2015.1066359

Christmas MJ, Jones JC, Olsson A *et al.* Genetic barriers to historical gene flow between cryptic species of alpine bumblebees revealed by comparative population genomics. *Molecular Biology and Evolution* 2021;**38**:3126–43. https://doi.org/10.1093/molbev/msab086

Christmas MJ, Jones JC, Olsson A *et al.* A genomic and morphometric analysis of alpine bumblebees: ongoing reductions in tongue length but no clear genetic component. *Molecular Ecology* 2022;**31**:1111–27. https://doi.org/10.1111/mec.16291

Crowley L; University of Oxford and Wytham Woods Genome Acquisition Lab. The genome sequence of the garden bumblebee, *Bombus hortorum* (Linnaeus, 1761). *Wellcome Open Research* 2021;**6**:270. https://doi.org/10.12688/wellcomeopenres.17187.1

Crowley LM, Sivell O, Sivell D, University of Oxford and Wytham Woods; Genome Acquisition Lab. The genome sequence of the Common Carder Bee, *Bombus pascuorum* (Scopoli, 1763). *Wellcome Open Research* 2023a;**8**:142. https://doi.org/10.12688/wellcomeopenres.19251.1

Crowley LM, Sivell O, Sivell D, *et al.* The genome sequence of the Early Bumblebee, *Bombus pratorum* (Linnaeus, 1761). *Wellcome Open Research* 2023b;**8**:143. https://doi.org/10.12688/wellcomeopenres.19250.1

Crowley LM, Sivell O, Sivell D, *et al.* The genome sequence of the Buff-tailed Bumblebee, *Bombus terrestris* (Linnaeus, 1758). *Wellcome Open Research* 2023c;**8**:161. https://doi.org/10.12688/wellcomeopenres.19248.1

Crowley L, Sivell O, University of Oxford and Wytham Woods Genome Acquisition Lab et al. The genome sequence of the Tree Bumblebee, *Bombus hypnorum* (Linnaeus, 1758). *Wellcome Open Research* 2023d;**8**:21. https://doi.org/10.12688/wellcomeopenres.18778.1

Crowley LM; University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective et al. The genome sequence of the Field Cuckoo-bee, *Bombus campestris* (Panzer, 1801). *Wellcome Open Research* 2023e;**8**:77. https://doi.org/10.12688/wellcomeopenres.18985.1

Dowton M, Austin AD. Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in the Hymenoptera. *Molecular Biology and Evolution* 1999;**16**:298–309. https://doi.org/10.1093/oxfordjournals.molbev.a026111

Dowton M, Castro LR, Campbell SL *et al.* Frequent mitochondrial gene rearrangements at the hymenopteran nad3-nad5 junction. *Journal of Molecular Evolution* 2003;**56**:517–26. https://doi.org/10.1007/s00239-002-2420-3

Dowton M, Cameron SL, Dowavic JI *et al.* Characterization of 67 mitochondrial tRNA gene rearrangements in the hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. *Molecular Biology and Evolution* 2009;**26**:1607–17. https://doi.org/10.1093/molbev/msp072

Duarte GT, Junqueira ACM, Azeredo-Espin AML. The mitochondrial control region of blowflies (Diptera: Calliphoridae): a hot spot for mitochondrial genome rearrangements. *Journal of Medical Entomology* 2008;**45**:667–676. https://doi.org/10.1093/jmedent/45.4.667

Eimanifar A, Kimball RT, Braun EL *et al.* The complete mitochondrial genome of *Apis nuluensis* Tingek, an Asian honey bee (Insecta: Hymenoptera: Apidae). *Mitochondrial DNA. Part B, Resources* 2017;**2**:552–3. https://doi.org/10.1080/23802359.2017.1365655

Fisogni A, Bogo G, Massol F *et al.* Cuckoo male bumblebees perform slower and longer flower visits than free-living male and worker bumblebees. *Belgian Journal of Zoology* 2021;**151**:193–203. https://doi.org/10.26496/bjz.2021.93

Françoso E, de Souza Araujo N, Ricardo PC *et al.* Evolutionary perspectives on bee mtDNA from mito-OMICS analyses of a solitary species. *Apidologie* 2020;**51**:531–44. https://doi.org/10.1007/s13592-020-00740-x

Françoso E, Zuntini AR, Ricardo PC *et al.* Rapid evolution, rearrangements and whole mitogenome duplication in the Australian stingless bees *Tetragonula* (Hymenoptera: Apidae): a steppingstone towards understanding mitochondrial function and evolution. *International*

Journal of Biological Macromolecules 2023;**242**:124568. https://doi.org/10.1016/j.ijbiomac.2023.124568
Free JB. The division of labour within bumblebee colonies. *Insectes Sociaux* 1955;**2**:195–212. https://doi.org/10.1007/bf02224381
Ghisbain G, Lozier JD, Rahman SR *et al.* Substantial genetic divergence and lack of recent gene flow support cryptic speciation in a colour polymorphic bumble bee (*Bombus bifarius*) species complex. *Systematic Entomology* 2020;**45**:635–52. https://doi.org/10.1111/syen.12419
Grab H, Branstetter MG, Amon N *et al.* Agriculturally dominated landscapes reduce bee phylogenetic diversity and pollination services. *Science* 2019;**363**:282–4. https://doi.org/10.1126/science.aat6016
Haddad N, Adjlane N, Loucif-Ayad W *et al.* Mitochondrial genome of the North African Sahara Honeybee, *Apis mellifera sahariensis* (Hymenoptera: Apidae). *Mitochondrial DNA Part B* 2017;**2**:548–9. https://doi.org/10.1080/23802359.2017.1365647
Hassanin A. Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. *Molecular Phylogenetics and Evolution* 2006;**38**:100–16. https://doi.org/10.1016/j.ympev.2005.09.012
Heraghty SD, Sutton JM, Pimsler ML *et al. De novo* genome assemblies for three North American bumble bee species: *Bombus bifarius, Bombus vancouverensis,* and *Bombus vosnesenskii*. *G3* 2020;**10**:2585–92. https://doi.org/10.1534/g3.120.401437
Hill GE. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecology and Evolution* 2016;**6**:5831–42. https://doi.org/10.1002/ece3.2338
Hill GE, Havird JC, Sloan DB *et al.* Assessing the fitness consequences of mitonuclear interactions in natural populations. *Biological Reviews of the Cambridge Philosophical Society* 2019;**94**:1089–104. https://doi.org/10.1111/brv.12493
Jakovlić I, Zou H, Chen J-H, *et al.* Slow crabs - fast genomes: Locomotory capacity predicts skew magnitude in crustacean mitogenomes. *Molecular Ecology* 2021;**30**(21):5488–5502. https://doi.org/10.1111/mec.16138
Junier T, Zdobnov EM. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 2010;**26**:1669–70. https://doi.org/10.1093/bioinformatics/btq243
Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* 2019;**20**:1160–6. https://doi.org/10.1093/bib/bbx108
Kent CF, Dey A, Patel H *et al.* Conservation genomics of the declining North American bumblebee *Bombus terricola* reveals inbreeding and selection on immune genes. *Frontiers in Genetics* 2018;**9**:316. https://doi.org/10.3389/fgene.2018.00316
Kück P, Meid SA, Groß C *et al.* AliGROOVE—visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinformatics* 2014;**15**:294. https://doi.org/10.1186/1471-2105-15-294
Lanfear R, Frandsen PB, Wright AM *et al.* PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 2016;34:772–3. https://doi.org/10.1093/molbev/msw260
Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics (Oxford, England)* 2008;**24**:172–5. https://doi.org/10.1093/bioinformatics/btm573
Lavrov DV, Brown WM, Boore JL. Phylogenetic position of the Pentastomida and (pan)crustacean relationships. *Proceedings Biological Sciences* 2004;**271**:537–44. https://doi.org/10.1098/rspb.2003.2631
Levinson G, Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* 1987;**4**:203–21. https://doi.org/10.1093/oxfordjournals.molbev.a040442
Lhomme P, Hines HM. Ecology and evolution of cuckoo bumble bees. *Annals of the Entomological Society of America* 2019;**112**:122–40. https://doi.org/10.1093/aesa/say031
Lin G, Huang Z, Wang L *et al.* Evolutionary rates of bumblebee genomes are faster at lower elevations (Y Kim, Ed). *Molecular Biology and Evolution* 2019a;**36**:1215–9. https://doi.org/10.1093/molbev/msz057
Lin G, Jiang K, Su T *et al.* Complete mitochondrial genome of *Bombus waltoni* (Hymenoptera: Apidae). *Mitochondrial DNA Part B* 2019b;**4**:1484–5. https://doi.org/10.1080/23802359.2019.1601528
Macey JR, Schulte JA, Larson A *et al.* Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. *Molecular Biology and Evolution* 1998;**15**:71–5. https://doi.org/10.1093/oxfordjournals.molbev.a025849
Mao M, Austin AD, Johnson NF *et al.* Coexistence of minicircular and a highly rearranged mtDNA molecule suggests that recombination shapes mitochondrial genome organization. *Molecular Biology and Evolution* 2014;**31**:636–44. https://doi.org/10.1093/molbev/mst255
Mao M, Gibson T, Dowton M. Higher-level phylogeny of the Hymenoptera inferred from mitochondrial genomes. *Molecular Phylogenetics and Evolution* 2015;**84**:34–43. https://doi.org/10.1016/j.ympev.2014.12.009
Michener CD. *The Bees of the World.* Baltimore: Johns Hopkins University Press, 2007.
Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES science gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE).* New Orleans, LA, USA: IEEE, 2010, 1–8.
Mitterboeck TF, Liu S, Adamowicz SJ *et al.* Positive and relaxed selection associated with flight evolution and loss in insect transcriptomes. *GigaScience* 2017;**6**:1–14. https://doi.org/10.1093/gigascience/gix073
Moure JS, Sakagami SF. As mamangabas sociais do Brasil (*Bombus* Latreille) (Hymenoptera, Apoidea). *Studia Entomologica* 1962;**5**:65–194.
Ohta T. Population size and rate of evolution. *Journal of Molecular Evolution* 1972;**1**:305–14. https://doi.org/10.1007/BF01653959
Oliveira DCSG, Raychoudhury R, Lavrov DV *et al.* Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Molecular Biology and Evolution* 2008;**25**:2167–80. https://doi.org/10.1093/molbev/msn159
Palozzi JM, Jeedigunta SP, Hurd TR. Mitochondrial DNA purifying selection in mammals and invertebrates. *Journal of Molecular Biology* 2018;**430**:4834–48. https://doi.org/10.1016/j.jmb.2018.10.019
Paradis E, Schliep K. ape 50: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019;**35**:526–8. https://doi.org/10.1093/bioinformatics/bty633
Pereira U de P, Bonetti AM, Goulart LR *et al.* Complete mitochondrial genome sequence of *Melipona scutellaris*, a Brazilian stingless bee. *Mitochondrial DNA Part A* 2016;**27**:3387–8. https://doi.org/10.3109/19401736.2015.1018233
Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution* 1995;**41**:353–8. https://doi.org/10.1007/BF00186547
Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics (Oxford, England)* 2005;**21**:676–9. https://doi.org/10.1093/bioinformatics/bti079
R Core Team. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing, 2021.
Rambaut A, Drummond AJ, Xie D *et al.* Posterior summarization in bayesian phylogenetics using tracer 1.7. *Systematic Biology* 2018;**67**:901–4. https://doi.org/10.1093/sysbio/syy032
Ranwez V, Douzery EJP, Cambon C *et al.* MACSE v.2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Molecular Biology and Evolution* 2018;**35**:2582–4. https://doi.org/10.1093/molbev/msy159
Richardson AO, Rice DW, Young GJ *et al.* The 'fossilized' mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biology* 2013;**11**:29. https://doi.org/10.1186/1741-7007-11-29

Rokas A, Holland PWH. Rare genomic changes as a tool for phylogenetics. *Trends in Ecology & Evolution* 2000;**15**:454–9. https://doi.org/10.1016/s0169-5347(00)01967-4

Ronquist F, Teslenko M, Mark P van der *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 2012;**61**:539–42. https://doi.org/10.1093/sysbio/sys029

Ruiz C, Cejas D, Muñoz I *et al.* Characterizing the mitogenome of the endemic bumblebee subspecies from the Canary Islands for conservation purposes. *Sociobiology* 2021;**68**:e5910. https://doi.org/10.13102/sociobiology.v68i3.5910

Sadd BM, Barribeau SM, Bloch G *et al.* The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biology* 2015;**16**:76. https://doi.org/10.1186/s13059-015-0623-3

Santos Júnior JE, Williams PH, Dias CAR *et al.* Biogeography and diversification of bumblebees (Hymenoptera: Apidae), with emphasis on neotropical species. *Diversity* 2022;**14**:238. https://doi.org/10.3390/d14040238

Schrader L, Pan H, Bollazzi M *et al.* Relaxed selection underlies genome erosion in socially parasitic ant species. *Nature Communications* 2021;**12**:2918. https://doi.org/10.1038/s41467-021-23178-w

Shao R, Dowton M, Murrell A *et al.* Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. *Molecular Biology and Evolution* 2003;**20**:1612–9. https://doi.org/10.1093/molbev/msg176

Silvestre D, Francisco F de O, Weinlich R *et al.* A scientific note on mtDNA gene order rearrangements among highly eusocial bees (Hymenoptera, Apidae). *Apidologie* 2002;**33**:355–6. https://doi.org/10.1051/apido:2002021

Silvestre D, Dowton M, Arias MC. The mitochondrial genome of the stingless bee *Melipona bicolor* (Hymenoptera, Apidae, Meliponini): sequence, gene organization and a unique tRNA translocation event conserved across the tribe Meliponini. *Genetics and Molecular Biology* 2008;**31**:451–60. https://doi.org/10.1590/s1415-47572008000300010

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–3. https://doi.org/10.1093/bioinformatics/btu033

Stewart JB, Freyer C, Elson JL *et al.* Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nature Reviews Genetics* 2008;**9**:657–62. https://doi.org/10.1038/nrg2396

Suhonen J, Rannikko J, Sorvari J. The rarity of host species affects the co-extinction risk in socially parasitic bumblebee *Bombus* (*Psithyrus*) species. *Annales Zoologici Fennici* 2015;**52**:236–42. https://doi.org/10.5735/086.052.0402

Sun C, Huang J, Wang Y *et al.* Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. *Molecular Biology and Evolution* 2021;**38**:486–501. https://doi.org/10.1093/molbev/msaa240

Takahashi J-I, Nishimoto M, Wakamiya T *et al.* Complete mitochondrial genome of the Japanese bumblebee, *Bombus hypocrita sapporensis* (Insecta: Hymenoptera: Apidae). *Mitochondrial DNA. Part B, Resources* 2016;**1**:224–5. https://doi.org/10.1080/23802359.2016.1155423

Takahashi J, Hadisoesilo S, Okuyama H *et al.* Analysis of the complete mitochondrial genome of *Apis nigrocincta* (Insecta: Hymenoptera: Apidae) on Sangihe Island in Indonesia. *Conservation Genetics Resources* 2018a;**10**:755–60. https://doi.org/10.1007/s12686-017-0922-y

Takahashi J, Sasaki T, Nishimoto M *et al.* Characterization of the complete sequence analysis of mitochondrial DNA of Japanese rare bumblebee species *Bombus cryptarum florilegus*. *Conservation Genetics Resources* 2018b;**10**:387–91. https://doi.org/10.1007/s12686-017-0832-z

Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 2007;**56**:564–77. https://doi.org/10.1080/10635150701472164

Tan HW, Liu GH, Dong X *et al.* The complete mitochondrial genome of the asiatic cavity-nesting honeybee *Apis cerana* (Hymenoptera: Apidae). *PLoS One* 2011;**6**:e23008. https://doi.org/10.1371/journal.pone.0023008

Tan MH, Gan HM, Lee YP *et al.* Comparative mitogenomics of the Decapoda reveals evolutionary heterogeneity in architecture and composition. *Scientific Reports* 2019;**9**:10756. https://doi.org/10.1038/s41598-019-47145-0

Tang M, Hardman CJ, Ji Y *et al.* High-throughput monitoring of wild bee diversity and abundance via mitogenomics. *Methods in Ecology and Evolution* 2015;**6**:1034–43. https://doi.org/10.1111/2041-210X.12416

Tian L, Rahman SR, Ezray BD *et al.* A homeotic shift late in development drives mimetic color variation in a bumble bee. *Proceedings of the National Academy of Sciences of the United States of America* 2019;**116**:11857–65. https://doi.org/10.1073/pnas.1900365116

Wang AR, Kim MJ, Park JS *et al.* Complete mitochondrial genome of the dwarf honeybee, *Apis florea* (Hymenoptera: Apidae). *Mitochondrial DNA* 2013;**24**:208–10. https://doi.org/10.3109/19401736.2012.744986

Wang CY, Zhao M, Wang SJ *et al.* The complete mitochondrial genome of *Lepidotrigona flavibasis* (Hymenoptera: Meliponini) and high gene rearrangement in *Lepidotrigona* mitogenomes. *Journal of Insect Science* 2021;**21**:10. https://doi.org/10.1093/jisesa/ieab038

Wei SJ, Shi M, Chen XX *et al.* New views on strand asymmetry in insect mitochondrial genomes. *PLoS One* 2010;**5**:1–10. https://doi.org/10.1371/journal.pone.0012708

Wertheim JO, Murrell B, Smith MD *et al.* RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular Biology and Evolution* 2015;**32**:820–32. https://doi.org/10.1093/molbev/msu400

Williams PH, Cameron SA, Hines HM *et al.* A simplified subgeneric classification of the bumblebees (genus *Bombus*). *Apidologie* 2008;**39**:46–74. https://doi.org/10.1051/apido:2007052

Williams PH, Françoso E, Martinet B *et al.* When did bumblebees reach South America? Unexpectedly old montane species may be explained by Mexican stopover (Hymenoptera: Apidae). *Systematics and Biodiversity* 2022;**20**:1–24. https://doi.org/10.1080/14772000.2022.2092229

Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology* 1992;**141**:173–216. https://doi.org/10.1016/s0074-7696(08)62066-5

Xiao JH, Jia JG, Murphy RW *et al.* Rapid evolution of the mitochondrial genome in Chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. *PLoS One* 2011;**6**:e26645. https://doi.org/10.1371/journal.pone.0026645

Xu W, Jameson D, Tang B *et al.* The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. *Journal of Molecular Evolution* 2006;**63**:375–92. https://doi.org/10.1007/s00239-005-0246-5

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 2007;**24**:1586–91. https://doi.org/10.1093/molbev/msm088

Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* 2002;**19**:908–17. https://doi.org/10.1093/oxfordjournals.molbev.a004148

Yoon KA, Kim K, Kim WJ *et al.* Characterization of venom components and their phylogenetic properties in some Aculeate bumblebees and wasps. *Toxins* 2020;**12**:47. https://doi.org/10.3390/toxins12010047

Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution* 2005;**22**:2472–9. https://doi.org/10.1093/molbev/msi237

Zhang J, Kan X, Miao G *et al.* qMGR: a new approach for quantifying mitochondrial genome rearrangement. *Mitochondrion* 2020;**52**:20–3. https://doi.org/10.1016/j.mito.2020.02.004

Zhao X, Huang J, Sun C *et al.* Complete mitochondrial genome of *Bombus consobrinus* (Hymenoptera: Apidae). *Mitochondrial DNA. Part B, Resources* 2017a;**2**:770–2. https://doi.org/10.1080/23802359.2017.1390422

Zhao X, Wu Z, Huang J *et al.* Complete mitochondrial genome of *Bombus breviceps* (Hymenoptera: Apidae). *Mitochondrial DNA. Part B,*

Resources 2017b;**2**:604–6. https://doi.org/10.1080/23802359.2017.1372710

Zhao F, Yan J, Jiang K *et al.* Nearly complete mitochondrial genomes of four bumblebee species (Hymenoptera: Apidae: *Bombus*). *Mitochondrial DNA Part B* 2019;**4**:183–4. https://doi.org/10.1080/23802359.2018.1545535

Zhao F, Morandin C, Jiang K *et al.* Molecular evolution of bumble bee vitellogenin and vitellogenin-like genes. *Ecology and Evolution* 2021;**11**:8983–92. https://doi.org/10.1002/ece3.7736

Zheng BY, Cao LJ, Tang P *et al.* Gene arrangement and sequence of mitochondrial genomes yield insights into the phylogeny and evolution of bees and sphecid wasps (Hymenoptera: Apoidea). *Molecular Phylogenetics and Evolution* 2018;**124**:1–9. https://doi.org/10.1016/j.ympev.2018.02.028

Zhou F, Song F, Yao L *et al.* Next-generation sequencing of the mitochondrial genome of *Bombus longipennis* Friese, 1918 (Hymenoptera: Apidae). *Mitochondrial DNA. Part B, Resources* 2021a;**6**:3049–51. https://doi.org/10.1080/23802359.2021.1976691

Zhou F, Zhang G, Yao L *et al.* The complete mitochondrial genome of *Bombus filchnerae* (Hymenoptera: Apidae) and phylogenetic analysis. *Mitochondrial DNA. Part B, Resources* 2021b;**6**:2727–9. https://doi.org/10.1080/23802359.2021.1966329

Zhu J, Tang P, Zheng BY *et al.* The first two mitochondrial genomes of the family Aphelinidae with novel gene orders and phylogenetic implications. *International Journal of Biological Macromolecules* 2018;**118**:386–96. https://doi.org/10.1016/j.ijbiomac.2018.06.087

Zou H, Lei HP, Chen R *et al.* Evolutionary rates of mitochondrial sequences and gene orders in Spirurina (Nematoda) are episodic but synchronised. *Water Biology and Security* 2022;**1**:100033. https://doi.org/10.1016/j.watbs.2022.100033

Capítulo 3

Patterns of mitonuclear coevolution in bumblebee genomes

Leonardo Tresoldi Gonçalves, Maríndia Deprá, Elaine Françoso

Manuscrito a ser submetido ao periódico
Molecular Biology and Evolution

Resumo

As mitocôndrias desempenham um papel central na respiração celular, necessitando de uma coevolução íntima com o genoma nuclear para funcionarem adequadamente. Esta coevolução, conhecida como coevolução mitonuclear, ainda não é totalmente compreendida em uma escala evolutiva a nível de espécie. Neste estudo, investigamos a coevolução mitonuclear em mamangavas (*Bombus*), um grupo ecologicamente diverso de polinizadores com rápida evolução mitocondrial (mt). Utilizando dados genômicos de um conjunto abrangente de 55 espécies de *Bombus*, quantificamos a correlação das taxas evolutivas (ERC, *evolutionary rate correlation*) entre genes mt e genes nucleares que interagem com a mitocôndria (N-mt). Encontramos uma ERC forte entre os genes mt e N-mt, mas não entre os genes mt e outros genes nucleares aleatórios, o que confirma a coevolução mitonuclear. Além disso, observamos que a intensidade da ERC parece ser consistente entre diferentes linhagens de *Bombus*, em contraste com observações anteriores em outros táxons. Por fim, detectamos que espécies de mamangavas de ambientes mais frios apresentaram taxas evolutivas mt aumentadas, ao passo que as taxas N-mt não parecem ser afetadas pela temperatura. Discutimos nossos achados considerando as dinâmicas da coevolução mitonuclear em *Bombus* e seu possível papel na modelagem da adaptação a diferentes nichos ecológicos.

Capítulo 4

The genome of the bellicose bumblebee (*Bombus bellicosus*): a threatened pollinator in a changing South American landscape

Leonardo Tresoldi Gonçalves, Pedro Henrique Pezzi, Flávia Regina Girardi Montagner, Wellington Vasconcelos de Souza, Elaine Françoso, Maríndia Deprá

Manuscrito a ser submetido ao periódico
Insect Conservation and Diversity

Resumo

Apresentamos o primeiro genoma de uma espécie neotropical de mamangava, *Bombus bellicosus*. Esta é uma espécie endêmica do sul da América do Sul que está se tornando localmente extinta devido à perda de habitat e às mudanças climáticas. Durante o distanciamento social da COVID-19 no Brasil, lançamos um projeto de ciência cidadã nas redes sociais para encontrar populações de *B. bellicosus*, e coletamos um espécime para sequenciamento do genoma. A análise do genoma revelou baixa diversidade genética em *B. bellicosus*, em comparação com uma espécie próxima amplamente distribuída (*B. pascuorum*). No entanto, a ausência de corridas em homozigosidade indicou uma falta de endocruzamento recente, o que oferece uma perspectiva cautelosamente otimista para a conservação desta espécie. Análises de história demográfica sugeriram que populações de *B. bellicosus* expandiram em períodos glaciais do passado, em contraste com as mamangavas paleárticas que sofreram um declínio marcante durante as glaciações. Nossos achados fornecem informações valiosas para a conservação dessa espécie e para estudos futuros sobre sua biologia e conservação, especialmente em um cenário de mudanças ambientais rápidas.

Considerações Finais

Nossa compreensão do DNA mitocondrial (mtDNA) está passando por uma transformação nas últimas décadas. Por muito tempo, o mtDNA foi relegado ao papel de um relógio molecular que acumulava mutações em um ritmo constante ao longo das gerações, servindo como ferramenta para entender a história genealógica dos organismos (Galtier et al. 2009). No entanto, surgem cada vez mais evidências de que as mutações no mtDNA não são apenas eventos estocásticos, mas sim frutos de processos de seleção e adaptação dos organismos. Estas mutações afetam a bioquímica e a performance mitocondrial, desempenhando um papel crucial em eventos macroevolutivos como a especiação (Ellison e Burton 2006; Ballard e Pichaud 2014; Hill 2015; Holmbeck et al. 2015; Burton 2022). Nesta tese, exploramos as nuances do mtDNA em um grupo de polinizadores de extrema importância ecológica e econômica: as mamangavas, abelhas do gênero *Bombus*. Caracterizadas por uma evolução acelerada do mtDNA, estas abelhas são um sistema fascinante para investigar o papel dessa molécula na adaptação e na diversificação.

Desde o início dos anos 2000, o *DNA barcoding* se destaca como uma ferramenta promissora no estudo da biodiversidade, utilizando um fragmento de mtDNA para identificar animais (Hebert et al. 2003). Apesar de controvérsias (Moritz and Cicero 2004; DeSalle et al. 2005), a técnica segue agregando valor em diversas áreas da ciência, mesmo com a recente popularização e diminuição de custos do sequenciamento de genomas completos (DeSalle et al. 2017). No entanto, muitas espécies ainda carecem de recursos genéticos e não possuem sequer um único *DNA barcode* sequenciado. Isso fica ainda mais evidente se considerarmos a biodiversidade do Sul Global (Deplazes-Zemp et al. 2018). No **Capítulo 1**, exploramos os vieses amostrais nos *barcodes* de referência para abelhas da família Apidae no BOLD, principal banco de dados para sequências desse tipo. Também encontramos que tanto a região canônica de *barcode* do gene *cox1* (Hebert et al. 2003) quanto o *mini-barcode* proposto por Françoso e Arias (2013) apresentam desempenho similar na identificação de espécimes. A utilização desses marcadores será crucial para o levantamento contínuo de recursos genéticos da biodiversidade Neotropical, especialmente diante da acelerada perda de biodiversidade que tem marcado o Antropoceno. Esperar que os estudos com genomas completos se tornem viáveis pode ser tarde demais para a

conservação da biodiversidade. Além disso, os *mini-barcodes* se mostram promissores para identificação de espécimes antigos em coleções e outras amostras com DNA degradado (Françoso e Arias 2013), como o DNA ambiental.

O mtDNA também se destaca há décadas como ferramenta na sistemática filogenética, especialmente no âmbito da zoologia. No **Capítulo 2**, empregamos genomas mitocondriais completos para reconstruir a filogenia de *Bombus*. Notavelmente, detectamos conflitos entre topologias dependendo do método ou da matriz de dados empregados. Essa discrepância entre topologias, observada inclusive em estudos com genomas nucleares completos (Sun et al. 2021a), demonstra a complexidade em inferir as relações filogenéticas de *Bombus*. Também exploramos a ordem dos genes no cromossomo mitocondrial para entender relações evolutivas entre os subgêneros de *Bombus*. Ao contrário da maioria dos insetos, onde a ordem gênica do mtDNA se manteve conservada por milhões de anos (Cameron 2014), o mtDNA de *Bombus* se revelou dinâmico. Especificamente, a ordem dos genes de tRNA apresentou grande variabilidade entre as espécies, e muitos dos rearranjos detectados são sinapomorfias para grupos de subgêneros.

Muito mais do que uma ferramenta para identificar seres vivos ou para construir filogenias, o mtDNA também nos permite levantar novas hipóteses sobre processos evolutivos e ecológicos. Por meio de análises de evolução molecular nos **Capítulos 2 e 3**, demonstramos que as alterações observadas no mtDNA, outrora consideradas neutras ou resultantes de simples processos de deriva genética (Ballard e Whitlock 2004), podem estar relacionadas à adaptação. No **Capítulo 2**, evidenciamos um relaxamento da seleção purificadora em diversos genes mitocondriais de mamangavas do subgênero *Psithyrus*, parasitas sociais obrigatórias de outras espécies de *Bombus*. Um próximo passo crucial reside na integração destes resultados com experimentos em fisiologia comparada, a fim de elucidar se este relaxamento é produto da diminuição do N_e nestas espécies ou se está associado a um relaxamento da seleção sobre o funcionamento mitocondrial, resultando em uma atividade mitocondrial menos eficiente. Adicionalmente, o sistema parasita-hospedeiro em questão pode ser explorado para investigar se espécies onde o parasitismo social obrigatório surgiu de forma independente apresentam convergência nestes padrões evolutivos (a exemplo de *B. inexpectatus* e *B. hyperboreus*; Lhomme e Hines 2019).

Além disso, no **Capítulo 3**, demonstramos pela primeira vez uma coevolução estrita entre proteínas da OXPHOS codificadas no núcleo e na mitocôndria em *Bombus*. Embora esta coevolução já tenha sido identificada em outros grupos de animais (Piccinini et al. 2021; Weaver et al. 2022), inclusive em himenópteros (Yan et al. 2019), não estava claro se ela persistia ao analisar espécies próximas e recentes, ou sua relação com a diversificação. Nossos resultados confirmam a assinatura de coevolução mitonuclear em *Bombus* e também sugerem que espécies que habitam regiões mais frias apresentam taxas evolutivas aceleradas no mtDNA. Esses achados abrem portas para diversas pesquisas futuras. Em gradientes altitudinais ou latitudinais, por exemplo, será que os padrões observados se confirmam a nível populacional? Há valor adaptativo na força da interação mitonuclear ou na aceleração das taxas evolutivas entre diferentes linhagens proximamente relacionadas? Estudos integrativos, que considerem os resultados aqui apresentados dentro de um contexto filogenético, aprofundarão o conhecimento sobre essas questões.

Por último, no **Capítulo 4**, sequenciamos o primeiro genoma completo de uma espécie neotropical de mamangava: a enigmática e ameaçada *B. bellicosus*. Por meio de um projeto de ciência cidadã nas redes sociais, realizado durante a pandemia de COVID-2019, tivemos a oportunidade de encontrar um espécime desta abelha. Simultaneamente, pudemos realizar um trabalho de educação ambiental, promovendo a conservação de polinizadores entre o público engajado nesta iniciativa. Os resultados do sequenciamento genômico revelaram que, pelo menos para o indivíduo coletado, não havia indícios de cruzamentos consanguíneos recentes. Ademais, encontramos uma considerável diversidade genética, se considerarmos a proporção de loci heterozigotos no genoma do indivíduo amostrado, o que representa certa esperança sob o ponto de vista da genética da conservação. Embora existam registros de extinções locais (Martins e Melo 2010), ainda faltam dados abrangentes sobre o status de conservação de *B. bellicosus* em toda a sua área de ocorrência. Acreditamos que o genoma gerado neste estudo abrirá portas para pesquisas em genética de populações dessa espécie, que conquistou o apreço da comunidade durante nossa campanha de ciência cidadã. Dados como esse são cruciais para embasar decisões de conservação e pressionar o poder público a promover ações efetivas para a proteção dos polinizadores. Enxergamos nestes resultados uma bandeira para a preservação dessa espécie e de todo o rico ecossistema que ela representa.

Esta tese traz uma visão centrada no mtDNA sobre processos evolutivos em abelhas, mais especificamente em *Bombus*. Os resultados demonstram o potencial desta molécula como ferramenta para investigar diversos aspectos da biologia evolutiva, desde a adaptação individual até a especiação em larga escala. Esperamos que os métodos e as descobertas aqui apresentadas sirvam de base para futuros estudos, contribuindo para o avanço do conhecimento científico e a conservação da biodiversidade.

Referências

(Introdução e Considerações Finais)

Aguiar AP, Deans AR, Engel MS, Forshage M, Huber JT, Jennings JT, Johnson NF, Lelej AS, Longino JT, Lohrmann V et al. (2013) Order Hymenoptera. Zootaxa. doi: 10.11646/zootaxa.3703.1.12

Amsalem E, Grozinger CM, Padilla M and Hefetz A (2015) Chapter Two - The physiological and genomic bases of bumble bee social behaviour. In: Zayed A and Kent CF (eds) Advances in Insect Physiology. Academic Press, pp 37–93

Anderson AP, Luo X, Russell W and Yin YW (2020) Oxidative damage diminishes mitochondrial DNA polymerase replication fidelity. Nucleic Acids Res 48:817–829.

Andrews SJ, Fulton-Howard B, Patterson C, McFall GP, Gross A, Michaelis EK, Goate A, Swerdlow RH and Pa J (2020) Mitonuclear interactions influence Alzheimer's disease risk. Neurobiol Aging 87:138.e7-138.e14.

Avise JC, Neigel JE and Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. J Mol Evol 20:99–105.

Baer CF, Miyamoto MM and Denver DR (2007) Mutation rate variation in multicellular eukaryotes: causes and consequences. Nat Rev Genet 8:619–631.

Ballard JWO and Pichaud N (2014) Mitochondrial DNA: more than an evolutionary bystander. Funct Ecol 28:218–231.

Ballard JWO and Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13:729–744.

Barja G (1998) Mitochondrial free radical production and aging in mammals and birds. Ann N Y Acad Sci 854:224–238.

Barr CM, Neiman M and Taylor DR (2005) Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol 168:39–50.

Bar-Yaacov D, Hadjivasiliou Z, Levin L, Barshad G, Zarivach R, Bouskila A and Mishmar D (2015) Mitochondrial involvement in vertebrate speciation? The case of mito-nuclear genetic divergence in chameleons. Genome Biol Evol 7:3322–3336.

Bernardo PH, Sánchez-Ramírez S, Sánchez-Pacheco SJ, Álvarez-Castañeda ST, Aguilera-Miller EF, Mendez-de la Cruz FR and Murphy RW (2019) Extreme mito-nuclear discordance in a peninsular lizard: the role of drift, selection, and climate. Heredity 123:359–370.

Breton S and Stewart DT (2015) Atypical mitochondrial inheritance patterns in eukaryotes. Genome 58:423–431.

Burke M, Scholl EH, Bird DM, Schaff JE, Colman SD, Crowell R, Diener S, Gordon O, Graham S, Wang X et al. (2015) The plant parasite *Pratylenchus coffeae* carries a minimal nematode genome. Nematology 17:621–637.

Burton RS (2022) The role of mitonuclear incompatibilities in allopatric speciation. Cell Mol Life Sci 79:103.

Cameron SA, Hines HM and Williams PH (2007) A comprehensive phylogeny of the bumble bees (*Bombus*). Biol J Linn Soc 91:161–188.

Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF and Griswold TL (2011) Patterns of widespread decline in North American bumble bees. Proc Natl Acad Sci 108:662–667.

Cameron SA and Williams PH (2003) Phylogeny of bumble bees in the New World

subgenus *Fervidobombus* (Hymenoptera: Apidae): Congruence of molecular and morphological data. Mol Phylogenet Evol 28:552–563.
Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. Annu Rev Entomol 59:95–117.
Camillo E and Garófalo CA (1989) Analysis of the niche of two sympatric species of Bombus (Hymenoptera, Apidae) in southeastern Brazil. J Trop Ecol 5:81–92.
Camus MF, Wolff JN, Sgrò CM and Dowling DK (2017) Experimental support that natural selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian *Drosophila melanogaster*. Mol Biol Evol 34:2600–2612.
Castro LR, Austin AD and Dowton M (2002) Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera. Mol Biol Evol 19:1100–1113.
Chen Z, Zhang F and Xu H (2019) Human mitochondrial DNA diseases and *Drosophila* models. J Genet Genomics 46:201–212.
de Juan D, Pazos F and Valencia A (2013) Emerging methods in protein co-evolution. Nat Rev Genet 14:249–261.
Deplazes-Zemp A, Abiven S, Schaber P, Schaepman M, Schaepman-Strub G, Schmid B, Shimizu KK and Altermatt F (2018) The Nagoya Protocol could backfire on the Global South. Nat Ecol Evol 2:917–919.
DeSalle R, Egan MG and Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos Trans R Soc B Biol Sci 360:1905–1916.
DeSalle R and Goldstein P (2019) Review and interpretation of trends in DNA barcoding. Front Ecol Evol 7:302.
DeSalle R, Schierwater B and Hadrys H (2017) MtDNA: The small workhorse of evolutionary studies. Front Biosci Landmark Ed 22:873–887.
Dowton M and Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement “hot spot” in the Hymenoptera. Mol Biol Evol 16:298–309.
Dowton M, Castro LR, Campbell SL, Bargon SD and Austin AD (2003) Frequent mitochondrial gene rearrangements at the hymenopteran nad3-nad5 junction. J Mol Evol 56:517–526.
Ellison CK and Burton RS (2010) Cytonuclear conflict in interpopulation hybrids: the role of RNA polymerase in mtDNA transcription and replication. J Evol Biol 23:528–538.
Ellison CK and Burton RS (2006) Disruption of mitochondrial function in interpopulation hybrids of *Trigopus californicus*. Evolution 60:1382–1391.
Ellison CK, Niehuis O and Gadau J (2008) Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. J Evol Biol 21:1844–1851.
Emery C (1909) Uber den Ursprung der dulotischen, parasitischen und myrekophilen Ameisen. Biol Cent 29:352–362.
Françoso E and Arias MC (2013) Cytochrome c oxidase I primers for corbiculate bees: DNA barcode and mini-barcode. Mol Ecol Resour 13:844–850.
Françoso E, de Oliveira FF and Arias MC (2016) An integrative approach identifies a new species of bumblebee (Hymenoptera: Apidae: Bombini) from northeastern Brazil. Apidologie 47:171–185.
Françoso E, de Souza Araujo N, Ricardo PC, Santos PKF, Zuntini AR and Arias MC (2020) Evolutionary perspectives on bee mtDNA from mito-OMICS analyses of a solitary species. Apidologie 51:531–544.
Françoso E, Zuntini AR, Ricardo PC, Santos PKF, De Souza Araujo N, Silva JPN, Gonçalves LT, Brito R, Gloag R, Taylor BA et al. (2023) Rapid evolution,

rearrangements and whole mitogenome duplication in the Australian stingless bees *Tetragonula* (Hymenoptera: Apidae): A steppingstone towards understanding mitochondrial function and evolution. Int J Biol Macromol 124568.
Fuhrmann DC and Brüne B (2017) Mitochondrial composition and function under the control of hypoxia. Redox Biol 12:208–215.
Galtier N, Nabholz B, Glémin S and Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol Ecol 18:4541–4550.
Garófalo CA (2005) *Bombus*: as mamangavas de chão e sua importância como agentes polinizadores. http://www.apacame.org.br/mensagemdoce/80/polinizacao2.htm. Accessed 20 May 2024
Goulson D (2010) Bumblebees: behaviour, ecology, and conservation, 2nd ed. Oxford University Press, Oxford ; New York
Goulson D (2015) Neonicotinoids impact bumblebee colony fitness in the field; a reanalysis of the UK's Food & Environment Research Agency 2012 experiment. PeerJ 3:e854.
Havird JC, Shah AA and Chicco AJ (2020) Powerhouses in the cold: mitochondrial function during thermal acclimation in montane mayflies. Philos Trans R Soc B Biol Sci 375:20190181.
Havird JC and Sloan DB (2016) The roles of mutation, selection, and expression in determining relative rates of evolution in mitochondrial versus nuclear genomes. Mol Biol Evol 33:3042–3053.
Hebert PDN, Cywinska A, Ball SL and DeWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc B Biol Sci 270:313–321.
Heimpel GE and Boer JG de (2008) Sex determination in the Hymenoptera. Annu Rev Entomol 53:209–230.
Hereward JP, Smith TJ, Gloag R, Brookes DR and Walter GH (2020) Tests of hybridisation in *Tetragonula* stingless bees using multiple genetic markers. doi: 10.1101/2020.03.08.982546
Hill GE (2015) Mitonuclear ecology. Mol Biol Evol 32:1917–1927.
Hill GE (2018) Mitonuclear mate choice: a missing component of sexual selection theory? BioEssays 40:1700191.
Hill GE, Havird JC, Sloan DB, Burton RS, Greening C and Dowling DK (2019) Assessing the fitness consequences of mitonuclear interactions in natural populations. Biol Rev 94:1089–1104.
Hill GE and Johnson JD (2013) The mitonuclear compatibility hypothesis of sexual selection. Proc R Soc B Biol Sci 280:20131314.
Hines HM (2008) Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst Biol 57:58–75.
Holmbeck MA, Donner JR, Villa-Cuesta E and Rand DM (2015) A *Drosophila* model for mito-nuclear diseases generated by an incompatible tRNA-tRNA synthetase interaction. Dis Model Mech 8:843–854.
Iannello M, Puccio G, Piccinini G, Passamonti M and Ghiselli F (2019) The dynamics of mito-nuclear coevolution: A perspective from bivalve species with two different mechanisms of mitochondrial inheritance. J Zool Syst Evol Res 57:534–547.
Jakovlić I, Zou H, Chen J, Lei H, Wang G, Liu J and Zhang D (2021) Slow crabs - fast genomes: Locomotory capacity predicts skew magnitude in crustacean mitogenomes. Mol Ecol 30:5488–5502.
Kaltenpoth M, Showers Corneli P, Dunn DM, Weiss RB, Strohm E and Seger J (2012)

Accelerated evolution of mitochondrial but not nuclear genomes of Hymenoptera: new evidence from crabronid wasps. PLoS ONE 7:e32826.
Krechemer F da S and Marchioro CA (2020) Past, present and future distributions of bumblebees in South America: Identifying priority species and areas for conservation. J Appl Ecol 57:1829–1839.
Lane N (2011) Mitonuclear match: optimizing fitness and fertility over generations drives ageing within generations. BioEssays News Rev Mol Cell Dev Biol 33:860–869.
Larsson N-G (2010) Somatic mitochondrial DNA mutations in mammalian aging. Annu Rev Biochem 79:683–706.
Lhomme P and Hines HM (2019) Ecology and evolution of cuckoo bumble bees. Ann Entomol Soc Am 112:122–140.
Lynch M (2006) The origins of eukaryotic gene structure. Mol Biol Evol 23:450–468.
Magnacca KN and Brown MJ (2010) Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus (Nesoprosopis)* bees (Hymenoptera: Colletidae). BMC Evol Biol 10:174.
Manlik O, Mundra S, Schmid-Hempel R and Schmid-Hempel P (2023) Impact of climate change on parasite infection of an important pollinator depends on host genotypes. Glob Change Biol 29:69–80.
Martins AC and Melo GAR (2010) Has the bumblebee *Bombus bellicosus* gone extinct in the northern portion of its distribution range in Brazil? J Insect Conserv 14:207–210.
Martins AC, Silva DP, De Marco P and Melo GAR (2015) Species conservation under future climate change: the case of *Bombus bellicosus*, a potentially threatened South American bumblebee species. J Insect Conserv 19:33–43.
Melvin RG and Ballard JWO (2017) Cellular and population level processes influence the rate, accumulation and observed frequency of inherited and somatic mtDNA mutations. Mutagenesis 32:323–334.
Meza-Lázaro RN, Poteaux C, Bayona-Vásquez NJ, Branstetter MG and Zaldívar-Riverón A (2018) Extensive mitochondrial heteroplasmy in the neotropical ants of the *Ectatomma ruidum* complex (Formicidae: Ectatomminae). Mitochondrial DNA Part A 29:1203–1214.
Moritz C and Cicero C (2004) DNA Barcoding: Promise and Pitfalls. PLoS Biol 2:e354.
Moure JS and Sakagami SF (1962) As Mamangabas Sociais do Brasil (*Bombus* Latreille) (Hymenoptera, Apoidea). Stud Entomol 5:65–194.
Nabholz B, Glémin S and Galtier N (2009) The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol Biol 9:54.
Nunnari J and Suomalainen A (2012) Mitochondria: in sickness and in health. Cell 148:1145–1159.
Oliveira DCSG, Raychoudhury R, Lavrov DV and Werren JH (2008) Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Mol Biol Evol 25:2167–80.
Piccinini G, Iannello M, Puccio G, Plazzi F, Havird JC and Ghiselli F (2021) Mitonuclear coevolution, but not nuclear compensation, drives evolution of OXPHOS complexes in bivalves. Mol Biol Evol 38:2597–2614.
Rand DM (2001) The units of selection on mitochondrial DNA. Annu Rev Ecol Syst 32:415–448.
Rand DM, Haney RA and Fry AJ (2004) Cytonuclear coevolution: the genomics of

cooperation. Trends Ecol Evol 19:645–653.
Rasmont P and Iserbyt S (2012) The Bumblebees Scarcity Syndrome: Are heat waves leading to local extinctions of bumblebees (Hymenoptera: Apidae: *Bombus*)? Ann Société Entomol Fr NS 48:275–280.
Ratnasingham S and Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Mol Ecol Notes 7:355–364.
Ricardo PC, Françoso E and Arias MC (2020) Mitochondrial DNA intra-individual variation in a bumblebee species: A challenge for evolutionary studies and molecular identification. Mitochondrion 53:243–254.
Richardson AO, Rice DW, Young GJ, Alverson AJ and Palmer JD (2013) The "fossilized" mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11:29.
Sagan L (1967) On the origin of mitosing cells. J Theor Biol 14:225-IN6.
Santos Júnior JE, Santos FR and Silveira FA (2015) Hitting an unintended target: phylogeography of *Bombus brasiliensis* Lepeletier, 1836 and the first new Brazilian bumblebee species in a century (Hymenoptera: Apidae). PLoS ONE 10:e0125847.
Schrader L, Pan H, Bollazzi M, Schiøtt M, Larabee FJ, Bi X, Deng Y, Zhang G, Boomsma JJ and Rabeling C (2021) Relaxed selection underlies genome erosion in socially parasitic ant species. Nat Commun 12:2918.
Sharkey MJ (2007) Phylogeny and classification of Hymenoptera. Zootaxa. doi: 10.11646/zootaxa.1668.1.25
Simčič T, Pajk F, Jaklič M, Brancelj A and Vrezec A (2014) The thermal tolerance of crayfish could be estimated from respiratory electron transport system activity. J Therm Biol 41:21–30.
Sirois-Delisle C and Kerr JT (2018) Climate change-driven range losses among bumblebee species are poised to accelerate. Sci Rep 8:14464.
Sun C, Huang J, Wang Y, Zhao X, Su L, Thomas GWC, Zhao M, Zhang X, Jungreis I, Kellis M et al. (2021a) Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. Mol Biol Evol 38:486–501.
Sun Y, Daffe G, Zhang Y, Pons J, Qiu J-W and Kupriyanova EK (2021b) Another blow to the conserved gene order in Annelida: evidence from mitochondrial genomes of the calcareous tubeworm genus *Hydroides*. Mol Phylogenet Evol 160:107124.
Suzuki-Ohno Y, Yokoyama J, Nakashizuka T and Kawata M (2020) Estimating possible bumblebee range shifts in response to climate and land cover changes. Sci Rep 10:19622.
Szczepanowska K and Trifunovic A (2017) Origins of mtDNA mutations in ageing. Essays Biochem 61:325–337.
Trier CN, Hermansen JS, Sætre G-P and Bailey RI (2014) Evidence for mito-nuclear and sex-linked reproductive barriers between the hybrid Italian Sparrow and its parent species. PLoS Genet 10:e1004075.
Wang S, Ore MJ, Mikkelsen EK, Lee-Yaw J, Toews DPL, Rohwer S and Irwin D (2021) Signatures of mitonuclear coevolution in a warbler species complex. Nat Commun 12:4279.
Weaver RJ, Rabinowitz S, Thueson K and Havird JC (2022) Genomic signatures of mitonuclear coevolution in mammals. Mol Biol Evol 39:msac233.
Whitfield J (2003) Phylogenetic Insights into the Evolution of Parasitism in Hymenoptera. Advances in Parasitology. Elsevier, pp 69–100

Williams PH (1998) An annotated checklist of bumble bees with an analysis of patterns of description (Hymenoptera: Apidae, Bombini). Bull Nat Hist Mus Entomol 67:79–152.

Williams PH, Cameron SA, Hines HM, Cederberg B and Rasmont P (2008) A simplified subgeneric classification of the bumblebees (genus *Bombus*). Apidologie 39:46–74.

Williams PH, Françoso E, Martinet B, Orr MC, Ren Z, Júnior JS, Thanoosing C and Vandame R (2022) When did bumblebees reach South America? Unexpectedly old montane species may be explained by Mexican stopover (Hymenoptera: Apidae). Syst Biodivers 20:1–24.

Wolff JN, Ladoukakis ED, Enríquez JA and Dowling DK (2014) Mitonuclear interactions: Evolutionary consequences over multiple biological scales. Philos Trans R Soc B Biol Sci. doi: 10.1098/rstb.2013.0443

Xiao J-H, Jia J-G, Murphy RW and Huang D-W (2011) Rapid evolution of the mitochondrial genome in Chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. PloS One 6:e26645.

Yan Z, Ye G and Werren JH (2019) Evolutionary rate correlation between mitochondrial-encoded and mitochondria-associated nuclear-encoded proteins in insects. Mol Biol Evol 36:1022–1036.

Zheng B-Y, Cao L-J, Tang P, van Achterberg K, Hoffmann AA, Chen H-Y, Chen X-X and Wei S-J (2018) Gene arrangement and sequence of mitochondrial genomes yield insights into the phylogeny and evolution of bees and sphecid wasps (Hymenoptera: Apoidea). Mol Phylogenet Evol 124:1–9.

Zhu J, Tang P, Zheng B-Y, Wu Q, Wei S and Chen X (2018) The first two mitochondrial genomes of the family Aphelinidae with novel gene orders and phylogenetic implications. Int J Biol Macromol 118:386–396.

Zucchi R and Kerr WE (1974) Aspectos bionômicos de *Exomalopsis aureopilosa* e *Bombus atratus* incluindo considerações sobre a evolução do comportamento social (Hymenoptera, ANpoidea). PhD Thesis, Universidade de São Paulo