BMC Genomics

RESEARCH Open Access

Mitochondrial genome of *Bactrocera* fruit flies (Tephritidae: Dacini): features, structure, and significance for diagnosis

Nathaly Lara Castellanos[1], Disna N. Gunawardana[1], Bede McCarthy[2], Puthigae Sathish[1], Sherly George[1] and Dongmei Li[1*]

Abstract

Background True fruit flies (Diptera: Tephritidae) are among the most destructive pests of fruit and vegetables worldwide and are on the top of quarantine pest lists. To respond effectively to a fruit fly invasion, we need to identify the species rapidly and reliably to understand its biological features and guide response decisions. Molecular techniques have been used to improve the diagnostic ability circumventing many difficulties of morphological identification. However, the commonly used Cytochrome Oxidase I (*COI*) gene lacks sufficient variation to distinguish species within *Bactrocera* species complexes. Here we conducted mitochondrial genome sequencing to identify additional genetic markers that could aid diagnosis of *Bactrocera* fruit fly species.

Results We assembled 82 complete mitochondrial genomes from 16 *Bactrocera* species, including 13 species for which no mitochondrial genome data were previously available, as well as one species each from *Dacus aneuvittatus*, *Dirioxa pornia* and *Zeugodacus gracilis*. Phylogenetic analysis of the Tephritidae family confirmed the monophyly of the *Bactrocera* genus but could not properly resolve species within species complexes. Comparative mitochondrial genome analysis revealed that intergenic spacer and NADH dehydrogenase genes, specifically *ND2* and *ND6*, harbour enough variations for new specific real-time PCR assays. Based on these findings, six TaqMan-based real-time PCR assays targeting *ND2*, *COI*, and *CO3* genes were successfully designed and assessed for their specificity and sensitivity in detecting *Bactrocera curvipennis*, a member of the *B. tryoni* complex. Of these, one real-time PCR assay targeting the ND2 gene proved to be the most specific and sensitive. It detects *B. curvipennis* specifically at the level of 1 copy/µL of target DNA.

Conclusions Mitochondrial sequence analysis and comparative studies indicate that mitochondrial genomes offer valuable genetic markers for accurate diagnosis of *Bactrocera* fruit flies. The successful development of the *B. curvipennis* real-time PCR assay highlights the importance of having additional genetic markers to advance the molecular diagnostics in economically important *Bactrocera* species.

Keywords Molecular identification, Species-specific real-time PCR, Intergenic spacer, DNA barcoding, Biosecurity, Phylogeny, *Bactrocera curvipennis*

*Correspondence:
Dongmei Li
Dongmei.Li@mpi.govt.nz

Full list of author information is available at the end of the article

Background

Biological invasions can lead to biodiversity loss, disruption of ecosystem services, reduced agricultural productivity [1–3]. The economic losses are estimated at $1,208 billion globally between 1980 and 2019, with a 702% increase in reported losses from 1980 to 1999 to 2000–2019 [1]. The unintentional transport of living organisms through global trade and international tourism increase the risk of biological invasions [4, 5]. True fruit flies (Diptera: Tephritidae) are on top of the quarantine pest list because they pose a major threat to horticultural crops and high potential of anthropogenic spread [2, 3]. Out of the 5,000 species within this family, only 250 highly polyphagous species are considered potential invasive species [2, 3, 6]. The most destructive fruit fly pests to global production of fresh fruits and vegetables belong to the genus *Bactrocera*, along with *Anastrepha*, *Ceratitis* and *Rhagoletis* [7, 8]. The economic impacts of these invasive fruit flies include substantial crop losses, increased cost due to pest management, eradication programs and biosecurity measures, and quarantine bans imposed by importing countries [9–11].

Accurate, rapid, and reliable identification of fruit flies to species level is essential for the adoption of effective quarantine actions to prevent the spread of invasive species. Inaccurate identification of specimens can lead to unnecessary quarantine measures or allow harmful species to establish in new regions [12]. Morphological differentiation between invasive and non-invasive species can be difficult especially for immature stages due to limited morphological characters and large intra-specific morphological variation [10, 13]. Various molecular techniques have been developed to overcome these challenges, including DNA barcoding, specific real-time PCR, and loop-mediated isothermal amplification (LAMP) [14–19]. Among them, DNA barcoding can discriminate most species [9, 20, 21], but it is time-consuming (taking up to four days) and depends on the quality of reference sequences [22–25]. In contrast, molecular identification using real-time PCR assays offer higher sensitivity, higher specificity and can considerably reduce the diagnostic time by eliminating post-PCR electrophoresis and amplicon sequencing [16, 17].

Molecular tools have significantly improved the diagnosis of fruit flies, but the selection of genetic markers for specific PCR assays or DNA barcoding purposes remains challenging. Mitochondrial genes, in particular the cytochrome oxidase I (*COI*), are the preferred genetic markers for molecular identifications due to their high mutation rates, rare gene recombination, maternal acquisition, and ease of sequencing using Sanger technique [26–28]. However, COI sequences do not allow distinction between closely related species within the same species complex, leaving at least 11.3% of Dacini fruit flies species non-identifiable [12, 24]. High similarity of COI sequences have been observed within invasive *Bactrocera* complexes [29–31], such as *B. dorsalis* (*B. dorsalis*, *B. carambolae*, *B. musae*, *B. borneoensis*, *B. incognita*, *B. cacuminata*, *B. verbascifoliae*, *B. caryecae*, *B. parafroggati*, *B. pallida*, *B. commensurata*, *B. occipitalis*, *B. kandiensis* and *B. raiensis* [30, 32]), *B. tryoni* (*B. tryoni*, *B. neohumeralis*, *B. erubescentis*, *B. mutabilis*, *B. curvipennis* and *B. ustulata* [31]), and *B. frauenfeldi* (*B. frauenfeldi*,, *B. trilineola*, *B. kirki*, *B. caledoniensis*, *B. psidii* and *B. parafrauenfeldi* [29]). This high sequence similarity makes species identification using real-time PCR challenging [14, 20, 31, 33, 34]. Complete mitochondrial genomes sequencing can reveal alternative genetic markers exhibiting greater interspecific variation, offering greater primer design flexibility and enable more accurate species-specific assays critical for biosecurity [11, 35]. Recent advances in high-throughput sequencing (HTS) technologies have accelerated the expansion of fruit fly mitochondrial genomes [36, 37], facilitating the discovery of more informative genetic markers that can enhance molecular diagnostics [11].

This study analysed 82 mitochondrial genomes from 19 species of *Bactrocera* and related genera to identify molecular markers that can be used to enhance the reliability of fruit fly diagnosis. Specifically, we aim to (i) address phylogenetic gaps in species resolution within species complexes, focusing on the *Bactrocera* genus, (ii) characterize mitogenomic features useful for species discrimination within this genus, and (iii) develop a real-time PCR assay for the rapid and accurate identification of *B. curvipennis.*

Results

Mitochondrial genome sequencing

Adult fruit fly specimens were morphologically identified, and their identities were confirmed through COI barcoding following standard protocols (Table S1, Supporting Information). Mitochondrial genomes from 82 specimens, representing diverse COI haplotypes, were sequenced using the Illumina NovaSeq 6000 platform, generating an average of 20.0 ± 6.8 million reads per sample, totalling 2.70 ± 1.55 Gbp (Table S2). The average sequencing depth was 1,193.44x, ranging from 230x to 7,435x (Table S2). Despite some DNA degradation due to the specimens collected dry from fruit-fly traps and long-time storage, the obtained data quality was high [36]. Complete circular mitochondrial genomes were assembled for 18 species across four genera within two tribes: 16 *Bactrocera* species, as well *Zeugodacus gracilis* from the Dacini tribe, and *Dirioxa pornia* from the Acantho-neurini tribe. A partial mitochondrial genome for *Dacus aneuvittatus* (Dacini) was recovered with only a missing section in the control region.

The arrangement of the 37 genes was identical to their order and transcription directions as in other Tephritid fruit flies [36–38] and follows the ancestral insect mitogenome arrangement. The majority strand (J-strand) encodes 23 genes, including 9 protein coding genes (PCGs)—*ATP6*, *ATP8*, *CYTB*, *COI*, *CO2*, *CO3*, *ND2*, *ND3* and *ND6*, and 14 tRNA genes. The minority strand (N-strand) encodes 14 genes, comprising 4 PCGs—*ND1*, *ND4*, *ND4L* and *ND5*, 8 tRNA genes and 2 rRNA genes—*12S* and *16S*.

Phylogenetic analysis

Maximum likelihood analyses (ML) using 13 PCGs and the entire mitochondrial genome from 479 Tephritidae individuals produced trees with nearly identical topologies featuring few unstable branches (Fig. 1, Figure S1–2, Supporting Information). The resulting phylogeny was well-supported for subfamily and tribe relationships, with most major nodes showing high posterior probability (> 0.95) (Figure S1). Two primary clades emerged, grouping Dacinae, Tripetinae, and Tephritinae, while Phytalmidae formed the basal group of Tephritidae. Although support for the monophyly of Dacinae was low, three monophyletic tribes, (Ceratidini + Gastrozonini) + Dacini, were strongly supported.

Within the *Dacini* tribe, *Dacus* and *Zeugodacus* identified as sister groups to *Bactrocera*, and the genera *Bactrocera* and *Dacus* were monophyletic. The genus *Zeugodacus* was recovered as paraphyletic based on the 13 PCGs reconstruction but monophyletic according the entire mitogenome (Figure S1). Within *Bactrocera*, the subgenera *Bactrocera*, *Daculus*, *Notodacus*, and *Tetradacus* showed strong monophyly, whereas *Afrodacus* was polyphyletic. None of the phylogenetic reconstructions were able to resolve certain species within the species complexes as monophyletic, including *B. dorsalis*, *B. carambolae*, *B. tryoni*, *B. neohumeralis* and *B. trilineola*, with supported nodes containing mixed samples from different species (Fig. 1). The *B. tryoni* complex was divided into two subclades, one of which included *B. curvipennis*, with *B. neohumeralis* and *B. tryoni* specimens mixed in both clades. Although not yet recognized as a species complex, *B. facialis* was polyphyletic, consisting of two clades, one of which nested within *B. passiflorae* clade, referred as *B. facialis* clade 2 in the tree. These discordances between molecular and morphological data across clades in the complex, as previously reported [10, 29, 32, 37], highlights the need for more stable morphological and molecular markers to reliably define these species.

Features of *Bactrocera* mitochondrial genomes: insights into species differentiation

The structural analysis of all available mitochondrial genomes of economically important species revealed that *Bactrocera* mitochondrial genomes remained stable across major clades and provided enough taxonomic resolution for species-level identification (Fig. 2). The nucleotide composition of fruit fly mitochondrial genomes showed a high AT bias, the AT contents varied from 66.55% in *B. tsuneomus* to 74.09% in the *B. frauelfeldi* complex. The length range of the complete mitochondrial genome was from 15,815.9 ± 2.0 bp in *B. oleae* to 15,977 bp in *B. latifrons* (Fig. 2A). The length variations of 13 PCGs, 22 tRNAs and 2 rRNAs were very minor, while the length variation among different specimens was mainly caused by the variation in noncoding regions especially the control region (size range from 923.8 ± 0.8 bp in *B. xanthodes* to 953 bp in *B. latifrons*; Fig. 2A). The three longest mitochondrial intergenic spacers (IGS) were between *trnQ-trnM* with average length of 129.2 ± 22.3 bp, *trnC-trnY* with 98.2 ± 15.2 bp and *trnR-trnN* with 21.6 ± 11.5 bp (Fig. 2B). These intergenic regions usually evolve faster than PCGs, indeed length variation between species was observed (Fig. 2A). Therefore, these regions should be explored as species-specific markers for the accurate identification of fruit flies. Only minor variations in stop codons were observed across some clades, with the vast majority remaining conserved (Table S3).

Evolutionary analysis of mitochondrial protein-coding genes

Complete mitochondrial genomes serve as valuable genetic markers, therefore, the availability and abundance of these genomes is essential for designing specific PCR assays to avoid cross-reaction with non-represented species [11]. Therefore, we analysed the evolutionary patterns of the mitochondrial PCGs in *Bactrocera* fruit flies, which are more readily accessible across taxa due to higher nucleotide conservation in coding regions compared to intergenic spaces. Our evolutionary patterns analysis of mitochondrial genomes evaluated key parameters such as nucleotide diversity (Pi), synonymous and nonsynonymous substitution rates (Ka/Ks ratio), barcoding gaps, and codon usage bias (Fig. 3, Figure S3-S4). The average Pi of the individual genes ranged from 0.056 for *ND4L* and 0.109 for *ND2*, the *COI* gene showed intermediary nucleotide diversity (0.079; Fig. 3A). Along with *ND2*, the genes *ND6*, *ND3* and *CYTB* had the highest nucleotide diversity, 0.108, 0.094 and 0.093, respectively. The Ka/Ks ratio is a widely used metric in comparative genomics to estimate selection pressure and the evolutionary rate [39]. For the 13 PCGs, *ND5* had the highest average Ka/Ks ratio across all available *Bactrocera*

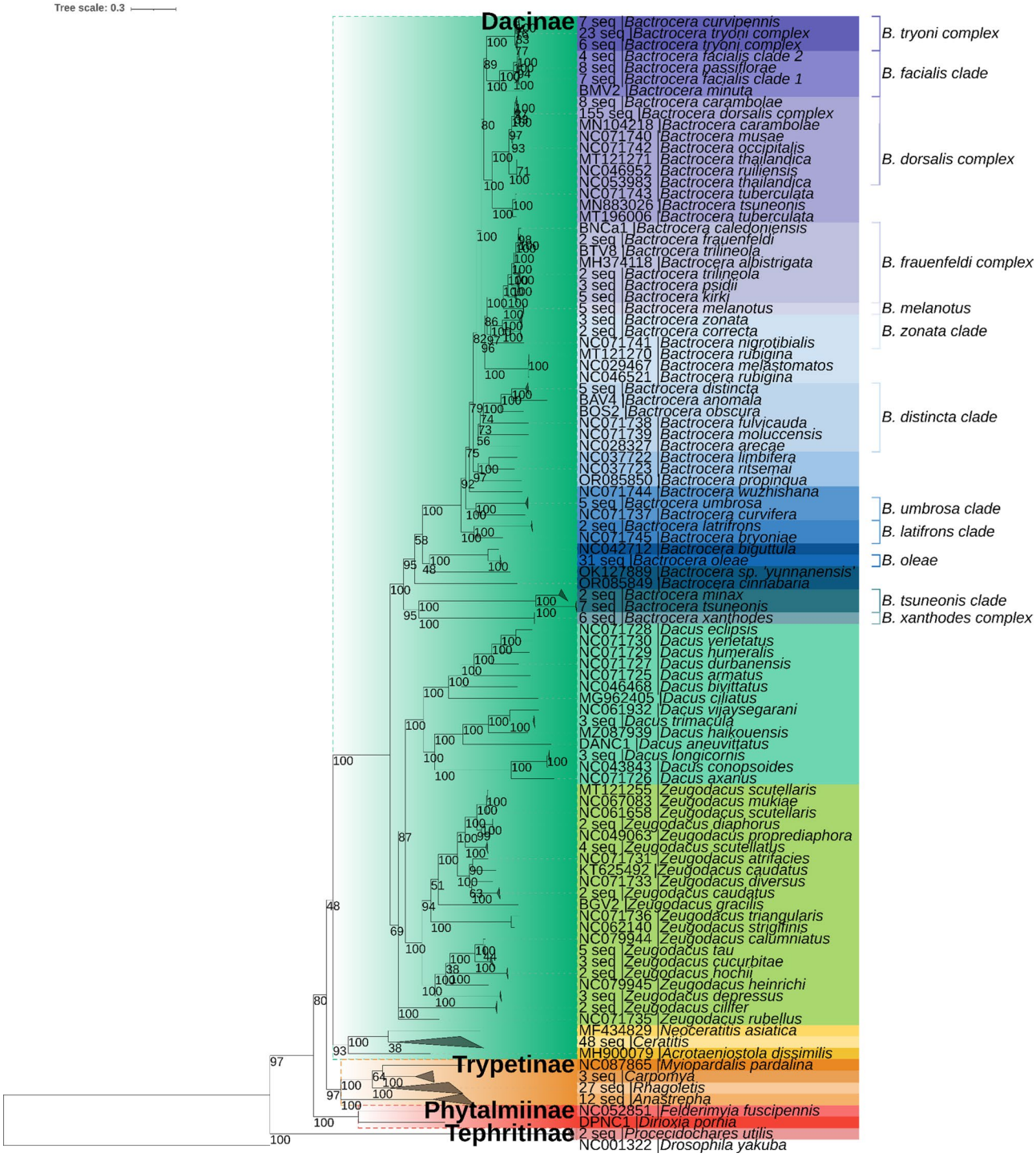

Fig. 1 Reconstructed phylogenetic tree of Tephritidae fruit flies inferred from the PCGs dataset with partitions using RAxML. Maximum likelihood (ML) bootstrap support values are shown at each node. Species and genera clades have been collapsed to improve visualization of the tree's topology. The *Bactrocera tryoni* complex includes *B. tryoni* and *B. neohumeralis*, while the *B. dorsalis* complex comprises primarily *B. dorsalis* and three *B. carambolae* specimens. The number of specimens used in each collapsed clade is specified. Boxes highlight subfamilies and brackets denote well-supported *Bactrocera* clades selected for further analyses

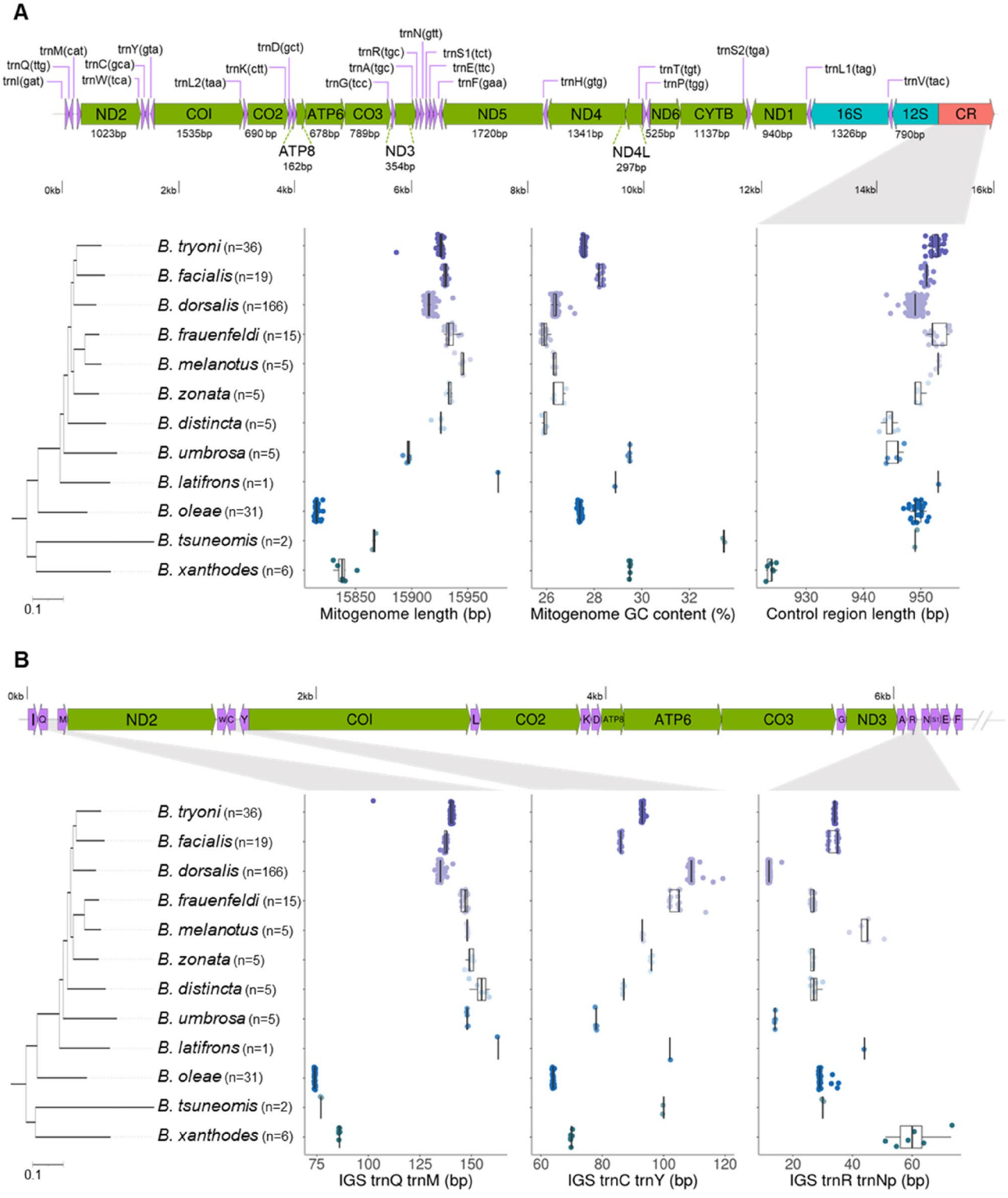

Fig. 2 Mitochondrial genome comparative analysis of *Bactrocera* fruit flies. **A** Schematic representation of mitochondrial gene arrangement. Protein-coding genes (PCGs), rRNAs, tRNAs, and the control region (CR) are shown in green, cyan, pink, and orange, respectively. Box plots of the genome length, GC content, and control region length are provided for selected *Bactrocera* clades. **B** Detailed schematic of major intergenic spaces (IGS) within *Bactrocera*, showing the boxplots of the lengths of the *trnQ-trnM*, *trnC-trnY*, and *trnR-trnN* IGS for selected clades. Each data point represents individual *Bactrocera* specimens, with point colours indicating different *Bactrocera* clades shown on Fig. 1. The number of specimens used for each clade is indicated in the pruned phylogenetic tree

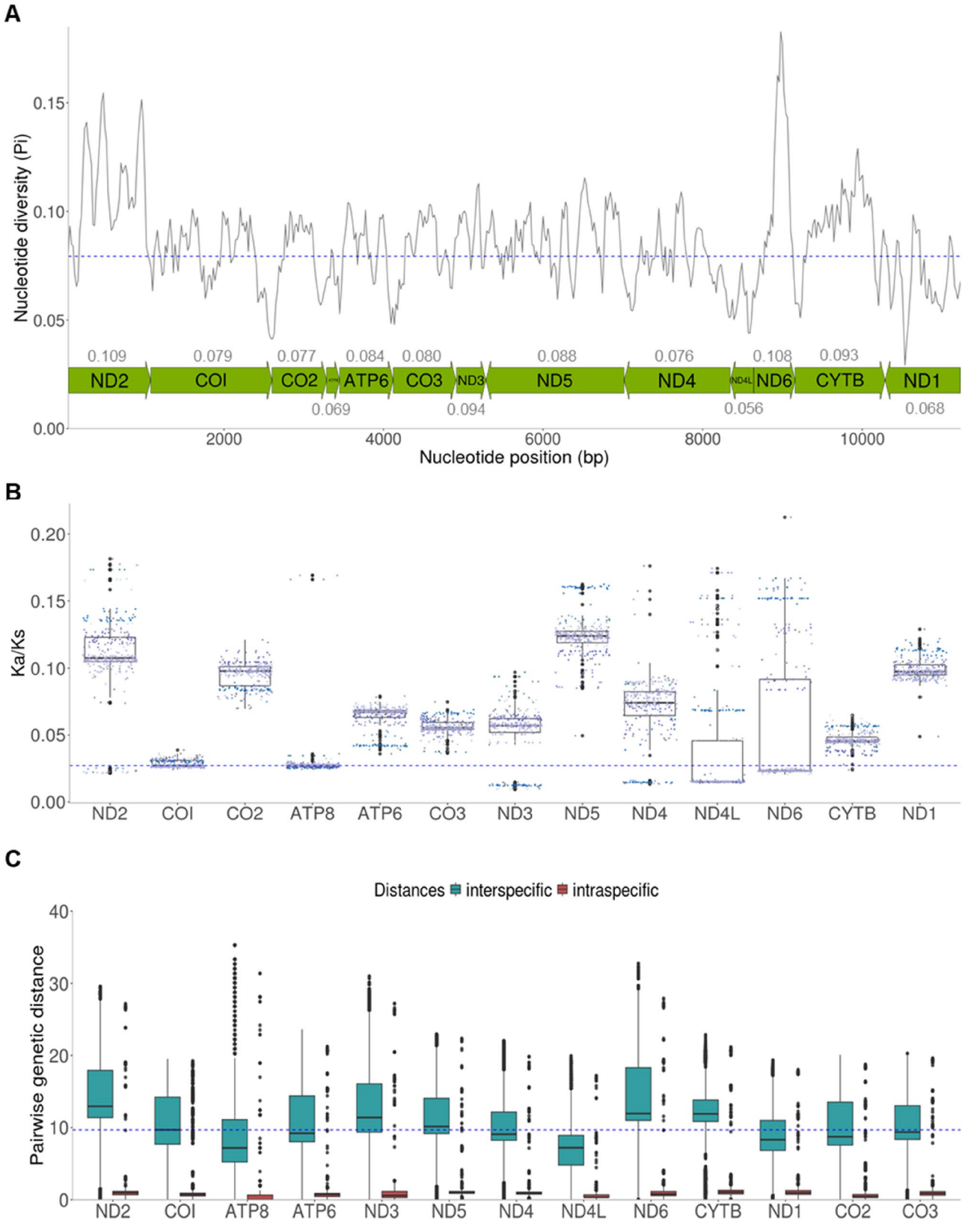

Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Evolutionary analysis of mitochondrial protein-coding genes (PCGs) in *Bactrocera* fruit flies. **A** Sliding window analysis of nucleotide diversity (Pi) across the PCGs, using a 100 bp window with a 20 bp step. Gene names and Pi values are indicated above or below the schematic representation of the genes. **B** Boxplots showing synonymous (Ks) and nonsynonymous (Ka) substitution rates for each PCG, with each data point represents an individual *Bactrocera* specimen. Point colours correspond to different *Bactrocera* clades shown on Fig. 1. **C** Boxplots of intra-specific and inter-specific pairwise genetic distances across PCGs to assess the barcoding gap. The median values for nucleotide diversity, Ka/Ks ratio, and interspecific genetic distances for the *COI* gene are indicated by the blue dotted line

mitogenomes, followed by *ND2* and *ND1*, whereas *COI* and *ATP8* had the lowest average values (Fig. 3B). Even though the *ATP8* gene showed a high rate of non-synonymous substitutions, its Ka/Ks ratio remained low due to the large number of synonymous substitutions (Figure S3). The average Ka/Ks values across all available *Bactrocera* mitogenomes varied from 0.029 (*COI*) to 0.124 (*ND5*) and consistently remained below 0.213 (below 1.0 the threshold for purifying selection), indicating that all PCGs have undergone purifying selection. Variation in the Ka/Ks ratio between different *Bactrocera* clades can be observed for genes *ND4L* and *ND6*, with the *B. dorsalis* complex showing the lowest Ka/Ks values. Intra- and inter-specific genetic distances based on the data of 13 PCGs were significantly overlapped regardless of the marker, providing evidence of the lack of adequate barcoding gaps for mitochondrial genes (Fig. 3C). For interspecific genetic distance, the highest values were observed for *ND2* (mean 13.93 ± 5.53%), followed by *ND6* (mean 13.85 ± 5.78%). Although the mean intra-specific genetic distance was very low for individual markers, maximum values overlap with the interspecific genetic distance (Fig. 3C). For instance, the commonly used COI barcode marker showed an interspecific genetic distance of 10.21 ± 4.18% with a maximum intra-specific genetic distance of 19.25%. Very weak codon bias was observed on the *Bactrocera* PCGs (Figure S4). In summary, although the *COI* gene has been extensively used for molecular identification of fruit flies, this gene exhibited lower evolutionary rates and moderate nucleotide diversity resulting in small differences among closely related species. Based in our results, the *ND2*, and *ND6* genes can provide better species resolution for distinguishing closely related species at the species level.

Performance comparison of mitochondrial genes for species-specific assays

Six Taqman-based real-time PCR assays targeting the *ND2*, *COI*, and *CO3* genes were designed (Table S4, Fig. 4A–B) and used for screening to evaluate potential cross-reactions and the amplification efficiency for known variants of *B. curvipennis* (Fig. 4C–H, Figure S5, S6). The initial screening of three ND2, one COI and two CO3 assays (Table S4) showed that the ND2 assay (ND2_1138F_BC, ND2_1192P_BC+ and ND2_1251R_BC) performed the best, with no cross-reactivity with non-target species, producing lower Cq values, and generating higher relative fluorescence (RFU > 1000) for all target species haplotypes (Fig. 4C, F). The COI assay showed cross-reactivity with other members of the *B. tryoni* complex, although lower RFU was observed for non-target species (Fig. 4D, F). Both CO3 assays exhibited cross-reactivity not only with the *B. tryoni* complex but also with additional *Bactrocera* species, showing non-specific amplification with higher Cq values (Fig. 4E, H and Figure S5C, F). The ND2 assay's higher specificity was further confirmed with additional primer combinations (Figure S5A, D and S5B, E).

Optimization and validation of the ND2 species-specific real-time PCR assay

Real-time PCR conditions for the ND2 assay (ND2_1138F_BC, ND2_1192P_BC + and ND2_1251R_BC) was optimized to increase sensitivity while still being specific. The optimized master mix composition and cycling conditions for the *B. curvipennis* duplex assay with 18S internal control targeting the 18S rRNA gene [17, 40, 41] for evaluation of the DNA quality are provided in Table S5 and S6. The annealing temperature was set at 60°C since higher temperatures resulted in increased Cq values and reduced florescence for certain *B. curvipennis* variants (Figure S7).

The linear dynamic range for the singleplex assay was assessed with serial dilutions of gBlock containing the ND2 primer and probe binding sites (Figure S8). The limit of detection (LOD) the singleplex assay was 1 copy/µL of synthetic DNA template with an average Cq of 36.86 for all replicates detected (Fig. 5). The amplification efficiency was 101.1% (Fig. 5), an thereby the analytical cut-off was stablished at 36 cycles based on the LOD [42]. The ND2 assay could also amplify 10 fg/µL of genomic DNA with Cq values around 35 cycles (Figure S8A, B) and the LOD was unaffected by the addition of the 18 S internal control (Figure S9C).

The diagnostic specificity the ND2 assay was validated with 31 fruit fly samples in the specificity tests and 21 additional samples in the blind panel (a complete list of the specimens tested is provided in Table S7). All the tested samples of *B. curvipennis* were successfully amplified the ND2 assay with a Cq lower than 25 cycles (Fig. 6). A single replicate of *Dirioxa pornia* DPNC2 resulted in a very weak amplification (Cq = 39.61), above the negative threshold of Cq > 36. The blind test produced consistent results and independently matched to the original

A

ND2	ND2_1138F_BC	Forward	CTACGACTATGCTTCGCG
	ND2_1192P_BC+	Probe	CAL610-ApdUpdUGApdUApdUAApdUpdUApdUCAApdCAApdCApdUpdUpdU-BHQ2plus
	ND2_1251R_BC	Reverse	GGAGATAAACGATAGGACC
COI	COI_1970F_BTC	Forward	TCACTTAGCTGGTATCTCATCG
	COI_2117P_BC	Probe	CAL610-TACCAGTTTTAGCGGGAGCTATTACTATAC-BHQ2
	COI_2176R_BTC	Reverse	AAGGATGTATTTAAGTTTCGAT
CO3	CO3_5171F_BTC	Forward	CTTTAACCCATTCCAAATCCCA
	CO3_5239P_BC	Probe	CAL610-AGCTCACCATAGTTTAATAGAAGGTAATCA-BHQ2
	CO3_5310R_BC	Reverse	CTAGAAGAACTGTAAAAAATAGACCC

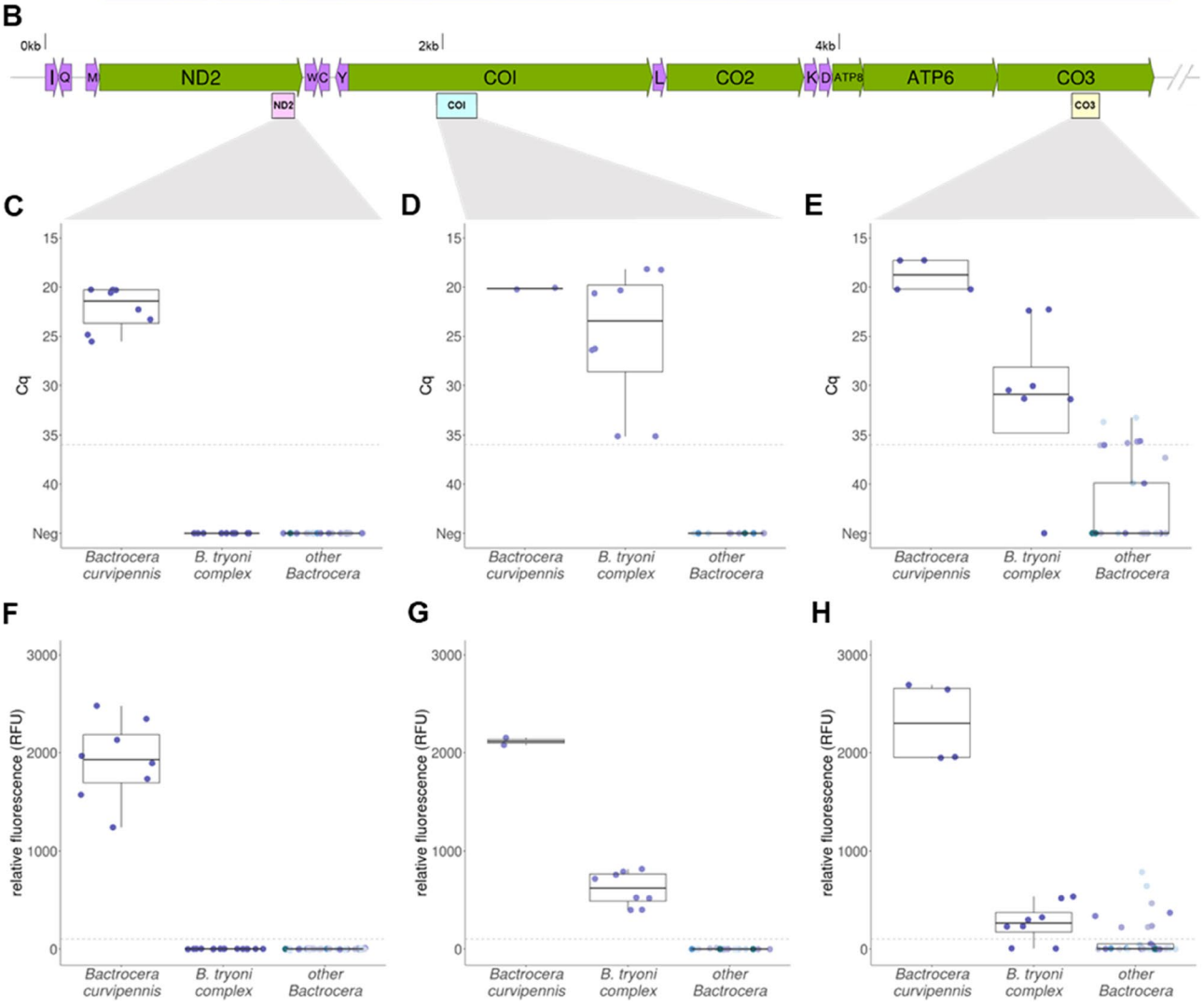

Fig. 4 Performance comparison of the specificity of real-time PCR assays designed for *Bactrocera curvipennis* detection. **A** Primer and probe sequences for each of the three assays targeting the *ND2*, *COI*, and *CO3* genes, represented by color-coded boxes. **B** Schematic overview of amplicon positions of each assay on the *B. curvipennis* mitochondrial genome (BCNC1). **C–E** Boxplot of Cq values for specific and non-specific amplifications across different samples, cut-off set at 36 cycles (grey dotted line). Neg: no amplification within 40 cycles. **F-H** Boxplot of relative fluorescence unit (RFU) for specific and non-specific amplifications, fluorescence threshold set at 100 RFU (grey dotted line). Primers (F: forward; R: reverse) and probes (P) sequences of each assay are shown in the same colour of boxes. Each data point represents samples from different *Bactrocera* clades shown on Fig. 1

identities of the samples provided. High assay reproducibility was demonstrated by low variation in %CV values for each of the samples tested within individual and between different runs (Table S8). Overall, these results demonstrated the ND2 assay robustness and reliability.

Discussion

This study generated 82 complete mitochondrial genomes from 19 fruit fly species, 16 of which were previously unavailable. Among these are highly polyphagous species with elevated risk of spreading to new locations, such as *B. curvipennis*, *B. facialis*, *B. kirki*, *B. melanotus*,

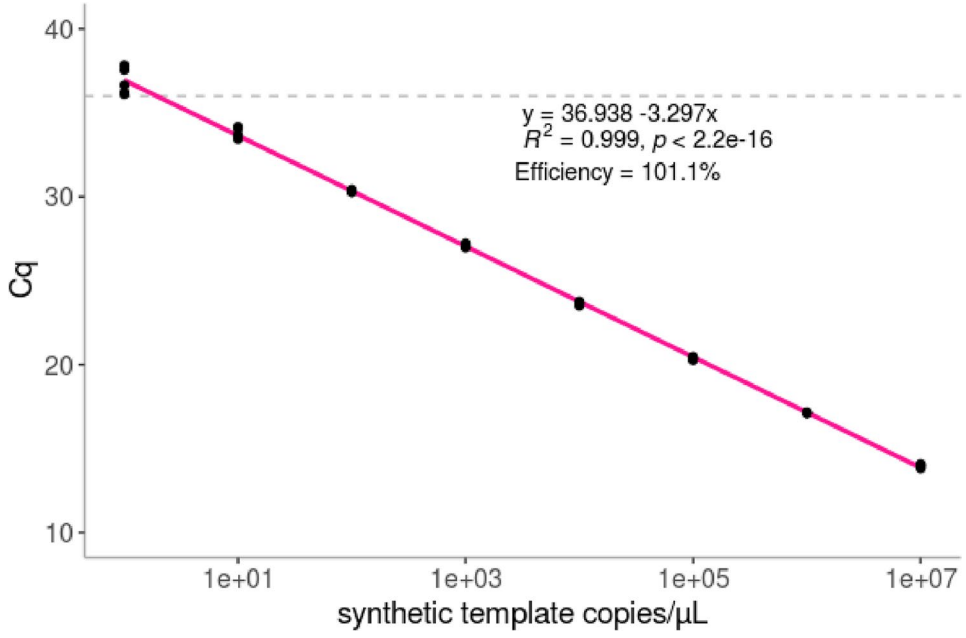

Fig. 5 Sensitivity of the ND2 real-time PCR assay for *Bactrocera curvipennis* detection. The gBlock synthetic template (Bac_TC_allin_control), containing concatenated primer and probe binding sites, was serially diluted, and tested with the singleplex real-time PCR assay. Linear regression of gBlock concentrations against Cq values generated a standard curve, plotting Cq values against the log copy number (range = 10^7 copies–1 copy) of the gBlock

B. passiflorae, *B. psidii*, *B. trilineola* and *B. xanthodes* [8]. Analysis of complete mitochondrial genomes provide useful molecular markers for both taxonomic and diagnostic studies [43]. Features like mitochondrial genome size, GC content, and control region length showed great variation across different *Bactrocera* clades, but their diagnostic utility is not widely applied due to the high cost and time required for whole genome sequencing. In contrast, the variability on the IGS lengths, specifically the *trnQ-trnM*, *trnC-trnY* and *trnR-trnN*, can be easily implemented as diagnostic markers to differentiate major clades as previously reported [43]. The findings of this study have also showed that the IGS length is conserved in *Bactrocera* species complexes analysed. However, the representation of *Bactrocera* species in the IGS database remains limited to a few species. Therefore, further investigation of more species is essential to more comprehensively assess IGS variability and its potential usage as diagnostic markers. To differentiate species within these complexes, genes with higher substitution rates, nucleotide diversity and interspecific genetic distance, such as *ND5*, *ND2* and *ND1*, can be better candidates for the development of species-specific diagnostic tools. This was exemplified by an increased specificity of the ND2 real-time PCR assay for the identification of *B. curvipennis* as compared with the assays for genes *COI* and *CO3* tested. Overall, these findings highlight that other mitochondrial genes and features offer higher diagnostic specificity than universal COI barcode, particularly for distinguishing species within complexes [12].

Our phylogenetic analyses of 479 mitochondrial genomes covering the major linages in the Tephritidae family provided a robust reconstruction, which was largely congruent with previous analyses and classifications schemes [10, 31, 36, 37, 44–46]. Monophyletic clades were observed for the genera of *Bactrocera*, *Dacus*, *Ceratitis*, *Rhagoletis*, *Carpomya* and *Anastrepha*, while *Zeugodacus* was resolved as paraphyletic with *Z. rubellus* and *Z. cilifer* positioned external to the *Zeugodacus* main clade (Fig. 1). The fact that mitochondrial genome data failed to distinguish closely related species within the well-known *B. dorsalis*, *B. tryoni*, and *B. frauenfeldi* species complexes, as well as *Ceratitis* FARQ complex (frugivorous flies - *C. fasciventris*, *C. anonae*, *C. rosa* and *C. quilicii*) [47], is not surprising. Notably, the inability to discriminate *B. facialis* and *B. passifloreae* using either the COI barcode or the entire mitochondrial genome is a novel finding, indicating that *B. facialis* and *B. passifloreae* are belonging to a species complex. Our results support expanding the *B. tryoni* complex to include the polyphagous fruit pest *B. curvipennis* from New Caledonia, consistent with previous molecular studies [31, 46]. Although the presence of *B. curvipennis* in Aneityum Island, a remote island in Vanuatu has been previously reported [48]subsequent evaluation by the European and Mediterranean Plant Protection Organization (EPPO) have considered it as absent [49]. This study sequenced *B. curvipennis* specimens from New Caledonia, obtained over different years and from diverse sources, thereby ensuring a sufficiently representative dataset to support

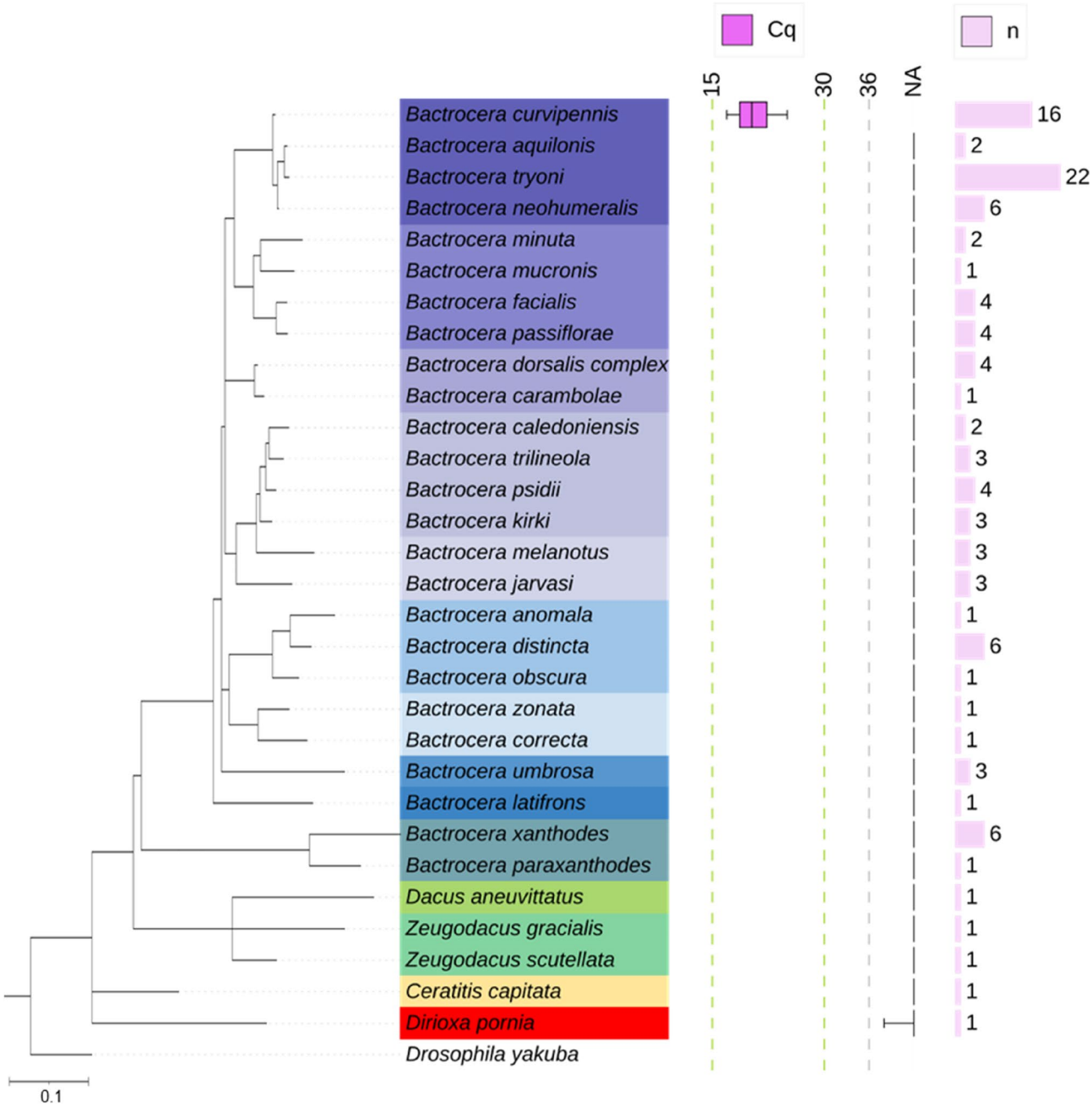

Fig. 6 Specificity of the ND2 real-time PCR assay for *Bactrocera curvipennis* detection. Reconstructed phylogenetic tree of the samples used in the development of the assay, inferred from the COI alignment using IQ-TREE. The boxplot of Cq values and bars representing the number of specimens displayed on the tree. Positive results with early amplification (Cq < 30) are indicated between the green dotted lines, while the grey dotted line represents the assay cut-off at 36 cycles

the conclusion drawn. Recent speciation events, hybridization, mitochondrial introgression, inter-specific gene flow likely contribute to the discrepancies between mitochondrial genetic diversity and morphology [10, 30, 32]. To address challenges posed by recent or ongoing speciation, employing nuclear genome wide data, such as RADseq (Restriction-site Associated DNA Sequencing) and HiMAP (Highly Multiplexed Amplicon-based Phylogenomics) can provide more accurate insights into species delimitation [29, 30, 32]. However, the continuous post-speciation gene flow [32] and the presence of additional, unrecognized species complexes among less studied fruit fly pests with more restricted distribution, highlights significant gaps in our knowledge which can hinder efforts to prevent new invasions or prioritize pest control measures [30]. Improving the accuracy of species

delimitation and resolving taxonomic discrepancies in fruit flies are crucial for strengthening biosecurity measures [30, 31].

Our findings demonstrate that mitochondrial genome features like the IGS and other rapid-evolving mitochondrial genes can be valuable to enhance species-detection resolution. In the battle against the spread of fruit fly pests, molecular methods such as DNA barcoding or species-specific assays (e.g., PCR, real-time PCR and LAMP) targeting the *COI* gene have been developed [12, 16, 17, 19, 50–52]. However, the COI-5P barcoding fragment that is widely used to develop these assays may not be the most informative for the molecular identification of *Bactrocera* fruit flies. Despite extensive BOLD records that have been generated for this fragment (up to 18,283 on 03/12/2024), its utility for diagnosis is compromised by (i) misidentifications in the reference databases, (ii) underrepresentation of non-pest species that do not respond to lures (only 272 of 750 species in *Bactrocera* genus), and (iii) considerable overlap of intra- and interspecific genetic divergences within species complexes [12, 22, 24]. For example, some assays targeting the *COI* gene are unable to distinguish *B. tryoni* from other closely related *B. tryoni* complex members [14, 50], *B. dorsalis* from other *B. dorsalis* complex species [14, 15], *B. latifrons* from *B. correcta* [53], *B. minax* from *B. tsuneomis* [52] and *C. capitata* from *C. cosyra* [54].

Remarkably, the whole length of IGS is relatively consistent within each *Bactrocera* clade and significantly varied among the other clades, suggesting that the IGS length can serve as a useful diagnostic tool for distinguishing clades. Mitochondrial IGS are usually limited in number and size, and represent the most variable part of insect mitochondrial genomes [55–58]. For instance, the *trnL2-CO2* IGS has been effective in characterizing genetic diversity within honey bee populations [59], while the *trnI-trmQ* and *trnQ-trnM* IGS gene region has been successfully used to differentiate the five most common fruit flies in South Africa (*C. capitata*, *C. cosyra*, *C. rosa*, *C. quilicii*, and *B. dorsalis*) [43]. Similarly, sequences covering the t*rnQ-trnM*, *trnC-trnY* and *trnR-trnN* IGS gene regions can be used for accurate differentiation of the main *Bactrocera* clades without downstream analysis and sequencing. However, the high AT content (> 80%) of IGS regions poses a challenge for developing species-specific assays, since AT-rich sequences reduce primer annealing temperatures, compromise the probe binding, and lower the amplification efficiency [60]. Additionally, the high variability of IGS regions increases the chances of mutations within primer and probe binding sites can lead to false negative results.

Mitochondrial PCGs are widely used in animal species-specific assays due to higher temporal stability, moderate AT content (~ 70% in *Bactrocera*), and large cellular copy number, which increases the likelihood of detection from the fragmented DNA frequently found in fruit-fly traps [61, 62]. Selecting the proper gene is critical for ensuring the analytical specificity and sensitivity required for border diagnostics. The nucleotide diversity, evolutionary rates and interspecific genetic distance indicates that the *ND5*, *ND2* and *ND1* genes are the least conserved genes among PCGs. While the *COI* gene was the most conserved gene as reported for other insects, including fruit flies [37], mosquitoes [63], beetles [58, 64, 65], parasitoid wasps [66], and soft scales [67]. Recent studies suggest that genes with rapid-evolving rates might be more suitable for the design of species-specific assays and capable of distinguishing closely related species [64]. A comparison of the analytical specificity of real-time PCR assays targeting the *COI*, *CO2* and *ND2* genes to distinguish *B. curvipennis* found that only the ND2 assay did not show cross-reactions with other *B. tryoni* complex members or other *Bactrocera* species. Subsequent test for sensitivity, diagnostic specificity, and reproducibility further validated the suitability of the ND2 assay. The *COI* representation of non-pest *Bactrocera* species is often lacking in the public databases, and the representation of other genetic markers that can be more suitable for the differentiation of closely related species is even less robust. Using genetic markers with limited representation in public databases can lead to cross-reaction with non-represented species. Therefore, the situations represent a trade-off dilemma between the representation in the databases and the diagnostic capability. Fortunately, advances in sequence technologies have facilitated the rapid expansion of mitochondrial genome databases [64], mitigating this limitation and enabling the broader implementation of alternative genetic makers for the diagnosis of invasive fruit flies.

The advent of high-throughput sequencing (HTS) offers a promising solution by enabling the cost-effective sequencing of entire mitochondrial genome, due to smaller size and high copy number of this organelle [61, 68–70]. HTS is not only more affordable, but also more technically robust as eliminates primer biases and non-standard amplification [68]. For instance, the cost of generating the entire mitochondrial genome using Illumina NovaSeq is currently less than US $100 for 10 million reads, which is more than sufficient to get the entire mitochondrial genome. Similar pricing and read depth are achievable using portable sequencing technologies, such as Oxford Nanopore Technologies (ONT), with the advantages of in situ sequencing, shorter timeframes and even covering highly AT rich sequences in the control region that are often missing by Illumina [71–75]. Despite these advancements, the accuracy of identifications at species-level based on the entire mitochondrial genome remains dependent on the availability and the

quality of the curated sequences in DNA repositories [74]. Therefore, expanding mitochondrial sequences databases, particularly for fruit flies and other quarantine pests, is essential for employing these technologies for species identification and biosecurity applications.

Conclusions

This study addresses significant gaps in sequence databases that have hindered the molecular identification of fruit flies, particularly in South Pacific Countries. By generating 82 complete mitochondrial genomes from 19 species across three genera of Tepritidae, this study provides critical genetic resources for diagnostic and taxonomic studies.

Our data also highlight the potential of alternative mitochondrial gene markers, such as *ND2*, *ND5* and IGS gene regions to improve the specificity and sensitivity of species-specific assays for closely related species. The evolutionary analysis of mitochondrial protein-coding genes (PCGs) showed that genes like *ND2* and *ND6* exhibited higher nucleotide diversity compared to *COI*. While *ND2* and *ND5* genes provide better species resolution, making them particularly useful for distinguishing closely related species at the species level. Additionally, the intergenic spacers, such as *trnQ-trnM* and *trnC-trnY*, displayed significant length variation between species, making them potential candidates for species-specific markers. Despite the success in screening potential genetic markers based on evolutionary rates, limited representation of many mitochondrial genes in public databases remains a challenge for their widespread use in fruit fly diagnostics.

The ND2 real-time PCR assay developed in this study demonstrates the practical application of these findings. The assay showed high sensitivity and specificity, with a limit of detection (LOD) of 1 copy/µL of the synthetic DNA, making it a reliable tool for the identification of *B. curvipennis*. When compared with the assays targeting *COI* and *CO3* genes, the ND2 assays demonstrated superior specificity in distinguishing *B. curvipennis* from other *B. tryoni* complex members without cross-reactions.

This research provide supports to improve surveillance and management of fruit flies, enhance the diagnostic accuracy and aid border security agencies in rapid decision-making to safeguard international trade; ultimately improving biosecurity response capability. The continuous expansion of genetic databases and innovations in sequencing technologies present further opportunities to refine molecular diagnosis and strengthen biosecurity responses. Additionally, these innovations can reduce the cost and time of fruit fly detection, benefiting the biosecurity agents and agricultural sectors that rely heavily on fruit production.

Methods

Specimen selection

Fruit fly species, representing five genera were obtained from interceptions in New Zealand and fruit fly traps from Australia, Cook Islands, Fiji, French Polynesia, New Caledonia, Tonga, and Vanuatu (Table S1 and S7). All adult specimens were identified using taxonomic keys and DNA barcoding (Supplementary Notes).

DNA extraction

For mitochondrial genome sequencing, total DNA from individual whole fruit fly specimens of 82 individuals of three genera (Table S2) was extracted using the DNeasy for Blood and Tissue kit (Qiagen, Germany) as per the manufacturer's instructions. The purified genomic DNA was quantified using the Qubit dsDNA Quantification Assay Kits (Invitrogen, Waltham, MA, USA) and normalized to 300 ng per sample.

For real-time PCR validation, additional DNA (Table S7) was extracted either a leg of adult, pieces of larva/pupa or an egg with DNeasy for Blood and Tissue kit (Qiagen, Germany) or PrepGEM universal kit (ZyGEM, New Zealand) according to manufacturers' instructions.

Mitochondrial genome sequencing, assembly, and annotation

Fragment length, Illumina library preparation and sequencing were performed by Novogene on the Agilent 5400 Fragment Analyzer System and Illumina NovaSeq 6000 platform paired-end 150 to attain at least 2.0–3.0 Gbp raw data per sample. The quality of sequences was evaluated using the FastQC software (Table S2). All ambiguous reads with an average quality value lower than Q20 and adaptor sequences were excluded from further analysis using Trimmomatic. De novo assembly was performed with MitoZ [76], MitoFlex [77] and GetOrganelle [78]. MitoFlex and MitoZ are modular pipelines specifically designed for mitochondrial genome assembly and annotation. MitoFlex and MitoZ represent a modified version of the assembler MEGAHIT to better assemble mitochondrial sequences, while GetOrganalle use Spades for *de novo* assembly of organelle genomes. For the MitoZ assembly module, it was specified that the reads stemmed from a specimen in the clade Arthropoda and all other default parameters were used. For the MitoFlex and GetOrganelle default parameter to assembly animal mitochondria were used. The final accessions were generated with consensus sequences from an alignment of the assembly methods that obtained circular sequence. After that, majority consensus sequences were generated using Samtools and coverage calculated using BEDtools [79]. Consensus sequences were then annotated using Geneious Prime 2021.1.1 (www.geneious.com/) and Mitos2 with the appropriate parameters:

Reference= "RefSeq 63 Metazoa" and Genetic Code = "5 Invertebrate" [80]. The two different annotation tools were compared and adjusted using the available fruit fly mitochondrial genomes. The genome sequences were submitted into GenBank under the accession numbers PV604183 - PV604264.

Phylogenetic analysis

The complete mitochondrial genome sequences of all 397 Tephritidae fruit fly specimens available in October 2024 were downloaded from NCBI (Table S9). Along with 82 from this study, 479 Tepthitidae mitochondrial genomes available representing 4 subfamilies, 8 tribes, 14 genera and 116 species were used for mitochondrial phylogenetic analyses, using *Drosophila yakuba* NC_001322.1 as the outgroup. Additionally, 13 complete PCGs were extracted from the 473 annotated mitochondrial genome sequences and concatenated into a single dataset using Geneious Prime 2021.1.1. (www.geneious.com). Multiple sequence alignment was performed using MAFFT version 7.505 with default parameters [81]. Possible incongruences in the alignment of protein-coding genes (PCGs) were verified using Geneious Prime. The alignments based on the entire mitochondrial genome and the concatenated PCGs were used for the phylogenetic analyses. Two different methodologies were used for the reconstruction of the phylogenetic relationships: (i) leaving the data unpartitioned with the same substitution model applied to all the sites in the alignment; and (ii) using a partition scheme based on genes and codon positions. For the first method, Maximum Likelihood (ML) tree was inferred for each of the three alignments by searching for the best substitution model using ModelFinder and calculating the bootstrap support with 10,000 replicates in IQ-TREE. PartitionFinder2 [82] was used to select the optimal partitioning schemes and substitution models for phylogenetic analyses using the "greedy" algorithm and branch lengths estimated as "unlinked" under the AICC model selection. The best partitioning scheme was used for phylogenetic reconstruction using 10,000 replicates in IQ-TREE. Phylogenetic analyses were also reconstructed using ML performed with RAxML 8.2.12 [83] using the GTRCAT model and a full ML bootstrap analysis with 100 replicates.

Mitochondrial genome comparative analysis

Comparative analysis of the *Bactrocera* mitochondrial genomes was performed using the concatenated PCGs dataset that includes the 325 annotated mitochondrial genome sequences distributed in 46 species. Nucleotide diversity (Pi) values of each PCG from 4SPE were determined using sliding window analyses of a window size of 100 bp and step size of 20 bp in DnaSP v6 [84]. The nonsynonymous substitution 184 rate (Ka), synonymous substitution rate (Ks), and Ka/Ks were calculated by the seqinr package [85]. The Ka/Ks ratio for each PCG, and each taxon was measured in comparison with the outgroup *Drosophila yakuba*. Genetic distances among fruit fly species were calculated using K80 model of evolution in the dist.dna function of ape for R. Barcoding gap analysis and the calculation of maximum intra-specific distances and the distance to nearest neighbour were performed employing the spider package in R [86]. The indices of codon usage and synonymous codon usage bias were measured using CodonW (https://codonw.sourceforge.net/).

In order to identify more effective genetic markers for the molecular identification of *Bactrocera* fruit flies, we analysed the structure of all available mitochondrial genomes, with a focus on economically important species [according to 8]. The selected clades included the complexes *B. dorsalis* (*B. dorsalis*, *B. carambolae*, *B. occipitalis*), *B. tryoni* (*B. tryoni*, *B. neohumeralis*, *B. curvipennis*), *B. frauenfeldi* (*B. frauenfeldi*, *B. albistigrata*, *B. kirki*, *B. trilineola*, *B. psidii*, *B. caledoniensis*), *B. zonata* (*B. zonata*, *B. correcta*), *B. facialis* (*B. facialis*, *B. passiflorae*), *B. latrifrons*, *B. oleae*, *B. melanotus*, *B. tsuneomis*, *B. xanthodes*, *B. distincta*, and *B. umbrosa*.

Development and validation of the qPCR assay

Species-specific primer design

The entire mitochondrial genome alignment was used to design real-time PCR assay for *B. curvipennis*. This alignment includes seven *B. curvipennis*, nine *B. neohumeralis* and twenty *B. tryoni* mitochondrial sequences. Additionally, all 591 available COI sequences from the *B. tryoni* complex defined by [31], namely *B. curvipennis*, *B. tryoni*, *B. neohumeralis*, *B. aquilonis*, *B. erubescentis*, *B. mutabilis* and *B. ustulata*, were also included in the aligment. Regions that showed differences specific for *B. curvipennis* were selected manually in the alignment and used for the design of specific probes and primers manually. The optimal condition for melting temperatures, GC content, dimer formation, and secondary structure formation were analysed through Oligo Analyzer Tool (IDT) and Geneious Prime 2021.1.1. The BHQplus stabilising technology was used to increase the melting temperature (Tm) of the probe, mismatch discrimination and binding to the AT rich regions in *ND2* gene [87].

Optimization of the assay conditions

All real-time PCR reactions were set up on a CFX96™ Touch Real-time platform (BioRad, Hercules, CA, USA). For each PCR assay, 2 µL of the DNA extract was added to a 20 µL final volume of reaction mix, containing 10 µL of the 2X mastermix, variable concentrations of primers and probes and DEPC water. Each reaction mix was performed in duplicate wells and all runs

included a positive control and a non-template control (NTC). Newly designed primers and probe for each assay (Table S4) were preliminary tested in different combinations using the target species and closely related species. Based on results from the preliminary specificity assessment, the best-performing primer-probe combination was selected for further optimisation. The selected primers and probes for each species was first tested using different annealing temperatures (58.1, 59.7, 61.6, 63.2 °C). Then, the assay was further optimised using mastermixes [PerfeCTa® qPCR ToughMix® (Quanta Bioscience, Beverly, MA, USA) and SsoAdvanced™ Universal Probes Supermix (BioRad, Hercules, CA, USA)], different primer concentrations (200–400 nM), probe gradients (100–300 nM), and with or without the addition of Bovine Serum Albumin (BSA, Sigma-Aldrich Co., Massachusetts, USA; 0.1–0.5 μg/μL). To coamplify the 18S ribosomal RNA (rRNA) gene as an internal control, the TaqMan™ 18S internal control (Applied Biosystems ™, CA, USA) primers and probe were incorporated with the *B. curvipennis* specific assay in duplex format.

Sensitivity and specificity of the species-specific real-time assay

A synthetic control (Integrated DNA Technologies gBlocks™ Gene Fragment) was used to determine the sensitivity of the qPCR assay. The synthetic control is a 283 bp template for different qPCR assay targets used in the detections of *B. curvipennis* and *B. tryoni* complex. The concatenated primers - probe sequence of the all-in-one synthetic control, is given in Figure S8. Connecting strings of Cs and Gs nucleotides were added before, between and after the primers/probes to increase CG content to meet the IDT manufacturer's requirement of minimum of 32% of GC content for a gBlock fragment. Extra bases at 3' and 5' ends were added to ensure primer binding. Sequences in between the primer and probe binding sites were not included for better differentiation of the COI or ND2 sequences of the target specimens from the gBlock positive control.

Analytical sensitivity of the assay was determined using the dilution series of the "all-in-one" template prepared in TE buffer ranging from 1 copy/μL to 1×10^7 copies/μL with each concentration in quintuplicate per reaction. Additionally, sensitivity was tested with ten-fold serial dilutions of genomic DNA from the target specimens (concentration ranging from 1 fg/μL to 10 ng/μL). Standard curve, PCR efficiency and R^2 were calculated automatically by the CFX manager software.

In addition to the specificity implied through the design process, the specificity of real-time PCR assay was tested using singleplex format against DNA extracted from target and the non-target species. In total, 31 fruit fly specimens were tested, including target species, closely related non-target species as well as species commonly intercepted at New Zealand's borders (Supplementary table S7). For the blind panel test, total of 21 specimens and 3 dilutions of the synthetic control were provided to the operators with no knowledge of the sample origin and identity (Supplementary table S7). The samples were tested with the ND2 real-time PCR assay in singleplex format. For each run, positive and non-template controls were included, and all the reactions were performed in triplicate wells.

Abbreviations

ML	Maximum likelihood
LAMP	Loop-mediated isothermal amplification
Gbp	Gigabase pair
PCG	Protein Coding Gene
ATP6	ATP synthase membrane subunit 6
ATP8	ATP synthase membrane subunit 8
CYTB	Cytochrome b
COI	Cytochrome c oxidase I
CO2	Cytochrome c oxidase 2
CO3	Cytochrome c oxidase 3
ND1	NADH dehydrogenase subunit 1
ND2	NADH dehydrogenase subunit 2
ND3	NADH dehydrogenase subunit 3
ND4	NADH dehydrogenase subunit 4
ND5	NADH dehydrogenase subunit 5
ND6	NADH dehydrogenase subunit 6
tRNA	transfer RNA
rRNA	Ribosomal RNA
CR	Control region
Pi	Average pairwise differences between all possible pairs of individuals
Ka/Ks ratio	The ratio of the number of nonsynonymous substitutions per non-synonymous site (Ka), in a given period of time, to the number of synonymous substitutions per synonymous site (Ks), in the same period
IGS	Intergenic Spacer region
HTS	High-Throughput Sequencing
ONT	Oxford Nanopore Technologies
HiMAP	Highly Multiplexed Amplicon-based Phylogenomics
RADseq	Restriction-site Associated DNA sequencing

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11872-8.

Supplementary Material 1.

Acknowledgements

We would like to thank Dr. Catia Delmiglio from Plant Health and Environment Laboratory (PHEL), MPI for the critically reviewing the manuscript. We would like to thank Zhidong Yu, Renee De Ruyter, Evan Brenton-Rule and Claire McDonald from the Operational Research Team, MPI for their support. We also acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities. Our special thanks go to the Entomology team members of the PHEL, MPI, specially for Yan Chen, for assisting in morphological identification of the specimens. Finally, we would like to thank the following researchers for providing the fruit fly specimens from Pacific countries: Piriaki Maaoo from the Ministry of Agriculture–Cook Islands; Siutoni Tupou from Ministry of Agriculture, Food and Forests–Tonga; Jainesh Ram from Biosecurity Authority Fiji; Faalelei Tunupopo from the Ministry of Agriculture and Fisheries–Samoa; Laura Hartamann from the Biosécurité de la Polynésie Française and Jeffline Tasale from Biosecurity Vanuatu.

Authors' contributions
DL, DG, BM and SG conceived the study. PS provided financial and project administration. DL and NLC designed the study. DG sourced all colony and trap-collected fruit flies and performed morphological identification. NLC conducted all laboratory experiments, primer design, HTS data analysis and mitochondrial genome assembly. DL supervised the genomic work, validation of the real-time PCR assay and data analysis, and co-wrote the manuscript. PS made a major contribution in revising the manuscript. All authors reviewed and approved the final manuscript.

Funding
The research was funded by the Operational Research programme of the Ministry for Primary Industries (MPI), New Zealand, grant number 406793.

Data availability
All the relevant data are included in the manuscript and additional data are in the supplementary information files.

Declarations

Ethics approval and consent to participate
No specific permits were required for all the sample collected for this study. No samples were collected in national parks and no endangered or threatened insects were included in this study, thus collection permits were not required for all the collections. All specimens imported into New Zealand were in accordance with the Import Health Standard, Sect. 22 of the Biosecurity Act 1993. Ethics approval was not required as insects are not classified as animals for the purposes of the Animal Welfare Act, 1999, New Zealand Legislation.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Ministry for Primary Industries, Plant Health & Environmental Laboratory, Diagnostic, Readiness and Surveillance (DRS), Biosecurity New Zealand, Ministry for Primary Industries, 231 Morrin Road, Auckland, New Zealand
[2]Ministry for Primary Industries, Plant Health & Environmental Laboratory, Diagnostic, Readiness and Surveillance (DRS), Biosecurity New Zealand, Ministry for Primary Industries, 14 Sir William Pickering Drive, Christchurch, New Zealand

Received: 16 April 2025 / Accepted: 2 July 2025
Published online: 29 July 2025

References

1. Turbelin AJ, Cuthbert RN, Essl F, Haubrock PJ, Ricciardi A, Courchamp F. Biological invasions are as costly as natural hazards. Perspect Ecol Conserv. 2023;21:143–50.
2. Papadopoulos NT, Meyer MD, Terblanche JS, Kriticos DJ. Fruit Flies: Challenges and Opportunities to Stem the Tide of Global Invasions. Ann Rev Entomol. 2024;69:355–73.
3. Trombik J, Ward SF, Norrbom AL, Liebhold AM. Global drivers of historical true fruit fly (Diptera: Tephritidae) invasions. J Pest Sci. 2023;96:345–57.
4. MacLachlan MJ, Liebhold AM, Yamanaka T, Springborn MR. Hidden patterns of insect establishment risk revealed from two centuries of alien species discoveries. Sci Adv. 2021;7:eabj1012.
5. Ahmed DA, Haubrock PJ, Cuthbert RN, Bang A, Soto I, Balzani P, et al. Recent advances in availability and synthesis of the economic costs of biological invasions. Bioscience. 2023;73:560–74.
6. EFSA P on PH, Bragard C, Dehnen-Schmutz K, Di Serio F, Gonthier P, Jacques M-A, et al. Pest categorisation of non-EU Tephritidae. EFSA J. 2020;18:e05931.
7. White IM, Elson-Harris MM. Fruit flies of economic significance: their identification and bionomics. 1992.
8. Vargas RI, Piñero JC, Leblanc L. An overview of pest species of *Bactrocera* fruit flies (Diptera: Tephritidae) and the integration of biopesticides with other biological approaches for their management with a focus on the Pacific region. Insects. 2015;6:297–318.
9. Krosch MN, Strutt F, Blacket MJ, Batovska J, Starkie M, Clarke AR, et al. Development of internal COI primers to improve and extend barcoding of fruit flies (Diptera: tephritidae: Dacini). Insect Sci. 2018;27:143–58.
10. San Jose M, Doorenweerd C, Leblanc L, Barr N, Geib S, Rubinoff D. Incongruence between molecules and morphology: A seven-gene phylogeny of Dacini fruit flies paves the way for reclassification (Diptera: Tephritidae). Mol Phylogenet Evol. 2018;121:139–49.
11. Andrews KJ, Bester R, Manrakhan A, Maree HJ. A multiplex PCR assay for the identification of fruit flies (Diptera: Tephritidae) of economic importance in South Africa. Sci Rep. 2022;12:13089.
12. Doorenweerd C, San Jose M, Leblanc L, Barr N, Geib S, Chung A, et al. Towards a better future for DNA barcoding: evaluating monophyly- and distance-based species identification using COI gene fragments of Dacini fruit flies. Mol Ecol Resour. 2024:24(6):e13987. https://doi.org/10.1111/1755-0998.13987.
13. Doorenweerd C, Leblanc L, Norrbom AL, Jose MS, Rubinoff D. A global checklist of the 932 fruit fly species in the tribe Dacini (Diptera, Tephritidae). Zookeys. 2018;730:19–56. https://doi.org/10.3897/zookeys.730.21786.
14. Dhami MK, Gunawardana DN, Voice D, Kumarasinghe L. A real-time PCR toolbox for accurate identification of invasive fruit fly species. J Appl Entomol. 2016;140:536–52.
15. Jiang F, Fu W, Clarke AR, Schutze MK, Susanto A, Zhu S, et al. A high-throughput detection method for invasive fruit fly (Diptera: Tephritidae) species based on microfluidic dynamic array. Mol Ecol Resour. 2016;16:1378–88.
16. Koohkanzade M, Zakiaghl M, Dhami MK, Fekrat L, Namaghi HS. Rapid identification of *Bactrocera zonata* (Dip.: Tephritidae) using TaqMan real-time PCR assay. PLoS ONE. 2018;13:e0205136.
17. Li D, Waite DW, Gunawardana DN, McCarthy B, Anderson D, Flynn A, et al. DNA barcoding and real-time PCR detection of *Bactrocera xanthodes* (Tephritidae: Diptera) complex. Bull Entomol Res. 2019;109:102–10.
18. Plant Health Australia. The Australian Handbook for the Identification of Fruit Flies. Version 3.1. 2018.
19. Starkie ML, Fowler EV, Zhu X, Agarwal A, Rako L, Schneider IC, et al. Loop-mediated isothermal amplification (LAMP) assays for detection of the new Guinea fruit fly *Bactrocera trivialis* (Drew) (Diptera: Tephritidae). Sci Rep. 2022;12:12602.
20. Blacket MJ, Semeraro L, Malipatil MB. Barcoding Queensland fruit flies (*Bactrocera tryoni*): impediments and improvements. Mol Ecol Resour. 2012;12:428–36.
21. Doorenweerd C, Jose MS, Leblanc L, Barr N, Geib S, Chung AYC et al. DNA barcodes and reliable molecular identifications in a diverse group of invasive pests: lessons from *Bactrocera* fruit flies on variation across the COI gene, introgression, and standardization. bioRxiv. 2020. https://doi.org/10.1101/2020.11.23.394510.
22. Drew Ra, Hancock I, Biogeography DL. Speciation and taxonomy within the genus *Bactrocera* Macquart with application to the *Bactrocera dorsalis* (Hendel) complex of fruit flies (Diptera: tephritidae: Dacinae). Zootaxa. 2022;5190:333–60.
23. Jiang F, Liang L, Li Z, Yu Y, Wang J, Wu Y, et al. A conserved motif within Cox 2 allows broad detection of economically important fruit flies (Diptera: Tephritidae). Sci Rep. 2018;8:2077.
24. Jiang F, Jin Q, Liang L, Zhang AB, Li ZH. Existence of species complex largely reduced barcoding success for invasive species of tephritidae: a case study in *Bactrocera* spp. Mol Ecol Resour. 2014;14:1114–28.
25. Smit J, Reijnen B, Stokvis F. Half of the European fruit fly species barcoded (Diptera, Tephritidae); a feasibility test for molecular identification. Zookeys. 2013;365:279–305. https://doi.org/10.3897/zookeys.365.5819.
26. Desalle R, Schierwater B, Hadrys H. MtDNA: the small workhorse of evolutionary studies. Front Biosci (Landmark Ed). 2017;22:873–87.
27. Kneubehl AR, Muñoz-Leal S, Filatov S, de Klerk DG, Pienaar R, Lohmeyer KH, et al. Amplification and sequencing of entire tick mitochondrial genomes for a phylogenomic analysis. Sci Rep. 2022;12:19310.
28. Nakahama N. Museum specimens: an overlooked and valuable material for conservation genetics. Ecol Res. 2021;36:13–23.
29. Doorenweerd C, San Jose M, Geib S, Dupuis J, Leblanc L, Barr N, et al. A phylogenomic approach to species delimitation in the Mango fruit fly (Bactrocera frauenfeldi) complex: A new synonym of an important pest species with variable morphotypes (Diptera: Tephritidae). Syst Entomol. 2023;48:10–22.
30. Doorenweerd C, San Jose M, Geib S, Barr N, Rubinoff D. Genomic data reveal new species and the limits of MtDNA barcode diagnostics to contain a

global pest species complex (Diptera: tephritidae: Dacinae). Syst Entomol. 2024;49:279–93.
31. Starkie ML, Cameron SL, Krosch MN, Phillips MJ, Royer JE, Schutze MK, et al. A comprehensive phylogeny helps clarify the evolutionary history of host breadth and lure response in the Australian Dacini fruit flies (Diptera: Tephritidae). Mol Phylogenet Evol. 2022;172:107481.
32. San Jose M, Doorenweerd C, Geib S, Barr N, Dupuis JR, Leblanc L, et al. Interspecific gene flow obscures phylogenetic relationships in an important insect pest species complex. Mol Phylogenet Evol. 2023;188:107892.
33. Armstrong KF, Ball SL. DNA barcodes for biosecurity: invasive species identification. Philos Trans R Soc Lond B Biol Sci. 2005;360:1813–23.
34. Smith PT, Kambhampati S, Armstrong KA. Phylogenetic relationships among *Bactrocera* species (Diptera: Tephritidae) inferred from mitochondrial DNA sequences. Mol Phylogenet Evol. 2003;26:8–17.
35. Dong Z, Wang Y, Li C, Li L, Men X. Mitochondrial DNA as a molecular marker in insect ecology: current status and future prospects. Ann Entomol Soc Am. 2021;114:470–6.
36. Zhang Y, De Meyer M, Virgilio M, Feng S, Badji K, Li Z. Phylogenomic resolution of the ceratitis farq complex (Diptera: Tephritidae). Mol Phylogenet Evol. 2021;161:107160.
37. Zhang Y, Li H, Feng S, Qin Y, De Meyer M, Virgilio M, et al. Mitochondrial phylogenomics reveals the evolutionary and biogeographical history of fruit flies (Diptera: Tephritidae). Entomol Generalis. 2022. https://doi.org/10.1127/entomologia/2022/1594.
38. Yong H-S, Chua K-O, Song S-L, Liew YJ-M, Eamsobhana P, Chan K-G. Complete mitochondrial genome of *Dacus Vijaysegarani* and phylogenetic relationships with congeners and other tephritid fruit flies (Insecta: Diptera). Mol Biol Rep. 2021;48:6047–56.
39. Li L-L, Xiao Y, Wang X, He Z-H, Lv Y-W, Hu X-S. The Ka /Ks and πa /πs ratios under different models of gametophytic and sporophytic selection. Genome Biol Evol. 2023;15:evad151.
40. Li D, Sooda A, Gunawardana DN, Thomas A, Chen Y, Kumarasinghe L. DNA barcodes for thrips species and development of multiplex real-time PCR assay for Frankliniella occidentalis pergande, Frankliniella panamensis hood, Thrips palmi Karny and Thrips tabaci Lindeman (Thysanoptera: Thripidae). New Z Entomol. 2023;46:16–34.
41. Li D, Fan Q-H, Waite DW, Gunawardana D, George S, Kumarasinghe L. Development and validation of a Real-Time PCR assay for rapid detection of Two-Spotted spider mite, *Tetranychus urticae* (Acari: Tetranychidae). PLoS ONE. 2015;10:e0131887.
42. Caraguel CGB, Stryhn H, Gagné N, Dohoo IR, Hammell KL. Selection of a cutoff value for Real-Time polymerase chain reaction results to fit a diagnostic purpose: analytical and epidemiologic approaches. J VET Diagn Invest. 2011;23:2–15.
43. Andrews KJ, Bester R, Manrakhan A, Maree HJ. Utilisation of a mitochondrial intergenic region for species differentiation of fruit flies (Diptera: Tephritidae) in South Africa. BMC Genomics. 2022;23:793.
44. Starkie ML, Cameron SL, Krosch MN, Sweet AD, Clarke AR. Biogeographic influences on the evolution and historical dispersal of the Australo-Pacific Dacini fruit flies (Tephritidae: Dacinae). Zoolog Scr. 2024;53:87–97.
45. Virgilio M, Jordaens K, Verwimp C, White IM, De Meyer M. Higher phylogeny of frugivorous flies (Diptera, tephritidae, Dacini): Localised partition conflicts and a novel generic classification. Mol Phylogenet Evol. 2015;85:171–9.
46. Krosch MN, Schutze MK, Armstrong KF, Graham GC, Yeates DK, Clarke AR. A molecular phylogeny for the tribe Dacini (Diptera: Tephritidae): Systematic and biogeographic implications. Mol Phylogenet Evol. 2012;64:513–23.
47. Drosopoulou E, Damaskou A, Markou A, Ekesi S, Khamis F, Manrakhan A, et al. The complete mitochondrial genomes of *Ceratitis rosa* and *Ceratitis quilicii*, members of the Ceratitis FAR species complex (Diptera: Tephritidae). Mitochondrial DNA Part B. 2021;6:1039–41.
48. Drew RAI, Drew RAI. The tropical fruit flies (Diptera: tephritidae: Dacinae) of the Australasian region and oceanic regions. Mem Qld Museum. 1989;26:1–521.
49. Leblanc L, Vueti ET, Drew RAI, Allwood AJ. Host Plant Records for Fruit Flies (Diptera: Tephritidae: Dacini) in the Pacific Islands. Proceed Hawaiian Entomological Soc. 2012;44:11–53.
50. Blacket MJ, Agarwal A, Zheng L, Cunningham JP, Britton D, Schneider I, et al. A LAMP assay for the detection of *Bactrocera tryoni* Queensland fruit fly (Diptera: Tephritidae). Sci Rep. 2020;10:9554.
51. Li D, Nair S, Anderson D, Doddala P, Gunawardana DN, George S. Real-time PCR assays for rapid detection of *Zeugodacus cucumis* and *Bactrocera jarvisi* (Diptera: Tephritidae) for quarantine application. J Appl Entomol. 2019;143:155–63.
52. Zheng L, Zhang Y, Yang W, Zeng Y, Jiang F, Qin Y, et al. New Species-Specific primers for molecular diagnosis of *Bactrocera minax* and *Bactrocera tsuneonis* (Diptera: Tephritidae) in China based on DNA barcodes. Insects. 2019;10:447.
53. Yu DJ, Zhang GM, Chen ZL, Zhang RJ, Yin WY. Rapid identification of *Bactrocera latifrons* (Dipt., Tephritidae) by real-time PCR using SYBR green chemistry. J Appl Entomol. 2004;128:670–6.
54. Dermauw W, Van Moerkercke Y, Ebrahimi N, Casteels H, Bonte J, Witters J. A loop-mediated isothermal amplification (LAMP) assay for rapid identification of *Ceratitis capitata* and related species. Curr Res Insect Sci. 2022;2:100029.
55. Andrianov B, Goryacheva I, Mugue N, Sorokina S, Gorelova T, Mitrofanov V. Comparative analysis of the mitochondrial genomes in Drosophila virilis species group (Diptera: Drosophilidae). Trends Evolutionary Biology. 2010;2:e4–4.
56. da Silva NM, de Souza Dias A, da Silva Valente VL, Valiati VH. Characterization of mitochondrial control region, two intergenic spacers and tRNAs of Zaprionus Indianus (Diptera: Drosophilidae). Genetica. 2009;137:325–32.
57. Xu M, Zhou S, Wan X. Phylogenetic implication of large intergenic spacers: insights from a mitogenomic comparison of Prosopocoilus stag beetles (Coleoptera: Lucanidae). Anim (Basel). 2022;12:1595.
58. Du C, Zhang L, Lu T, Ma J, Zeng C, Yue B, et al. Mitochondrial genomes of blister beetles (Coleoptera, Meloidae) and two large intergenic spacers in hycleus genera. BMC Genomics. 2017;18:698.
59. Chávez-Galarza J, López-Montañez R, Jiménez A, Ferro-Mauricio R, Oré J, Medina S, et al. Mitochondrial DNA variation in Peruvian honey bee (Apis mellifera L.) populations using the tRNAleu-cox2 intergenic region. Insects. 2021;12:641.
60. Bustin S, Huggett J. qPCR primer design revisited. Biomol Detect Quantification. 2017;14:19–28.
61. Elyasigorji Z, Izadpanah M, Hadi F, Zare M. Mitochondrial genes as strong molecular markers for species identification. Nucleus. 2023;66:81–93.
62. Fowler EV, Starkie ML, Blacket MJ, Mayer DG, Schutze MK. Effect of temperature and humidity on insect DNA integrity evaluated by real-time PCR. J Econ Entomol. 2024;117:1995–2002.
63. Hao Y-J, Zou Y-L, Ding Y-R, Xu W-Y, Yan Z-T, Li X-D, et al. Complete mitochondrial genomes of an.pheles stephensi and an. dirus and comparative evolutionary mitochondriomics of 50 mosquitoes. Sci Rep. 2017;7:7666.
64. Li C, Wang B, Ji Y, Huang L, Wang X, Zhao W, et al. Mitochondrial genome provides species-specific targets for the rapid detection of early invasive populations of Hylurgus ligniperda in China. BMC Genomics. 2024;25:90.
65. Shi F, Yu T, Xu Y, Zhang S, Niu Y, Ge S, et al. Comparative mitochondrial genomic analysis provides new insights into the evolution of the subfamily Lamiinae (Coleoptera: Cerambycidae). Int J Biol Macromol. 2023;225:634–47.
66. Huang L, Sun H-Q, Li C-J, Zhao W-X, Yao Y-X. The complete mitochondrial genome of *Platygaster robiniae* (Hymenoptera: Platygastridae): A novel tRNA secondary structure, gene rearrangements and phylogenetic implications. Int J Parasitology: Parasites Wildl. 2022;18:249–59.
67. Lu C, Huang X, Deng J. Mitochondrial genomes of soft scales (Hemiptera: Coccidae): features, structures and significance. BMC Genomics. 2023;24:37.
68. Antil S, Abraham JS, Sripoorna S, Maurya S, Dagar J, Makhija S, et al. DNA barcoding, an effective tool for species identification: a review. Mol Biol Rep. 2023;50:761–75.
69. Crampton-Platt A, Yu DW, Zhou X, Vogler AP. Mitochondrial metagenomics: letting the genes out of the bottle. GigaScience. 2016;5:s13742-016-0120-y.
70. Cameron SL. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny. Annual Rev Entomol. 2014;59:95–117.
71. De Vivo M, Lee H-H, Huang Y-S, Dreyer N, Fong C-L, de Mattos FMG, et al. Utilisation of Oxford nanopore sequencing to generate six complete gastropod mitochondrial genomes as part of a biodiversity curriculum. Sci Rep. 2022;12:9973.
72. Kipp EJ, Lindsey LL, Milstein MS, Blanco CM, Baker JP, Faulk C, et al. Nanopore adaptive sampling for targeted mitochondrial genome sequencing and bloodmeal identification in hematophagous insects. Parasites Vectors. 2023;16:68.
73. Prodhan MA, Widmer M, Kinene T, Kehoe M. Whole mitochondrial genomes reveal the relatedness of the browsing ant incursions in Australia. Sci Rep. 2023;13:10273.
74. Velasquez-Restrepo S, Orozco MC, Franco-Sierra ND, Martínez-Cerón JM, Díaz-Nieto JF. Identification of non-model mammal species using the minion DNA sequencer from Oxford nanopore. PeerJ. 2024;12:e17887.
75. Filipović I, Hereward JP, Rašić G, Devine GJ, Furlong MJ, Etebari K. The complete mitochondrial genome sequence of Oryctes rhinoceros

(Coleoptera: Scarabaeidae) based on long-read nanopore sequencing. PeerJ. 2021;9:e10552.
76. Meng G, Li Y, Yang C, Liu S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 2019;47:e63.
77. Li J-Y, Li W-X, Wang A-T, Zhang Y. MitoFlex: an efficient, high-performance toolkit for animal mitogenome assembly, annotation and visualization. Bioinformatics. 2021;37:3001–3.
78. Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, et al. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:241.
79. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2.
80. Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, et al. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res. 2019;47:10543–52.
81. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
82. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2017;34:772–3.
83. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35:4453–5.
84. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol Biol Evol. 2017;34:3299–302.
85. Charif D, Lobry JR. SeqinR 1.0–2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman HE, Vendruscolo M, editors. Structural approaches to sequence evolution: molecules, networks, populations. Berlin, Heidelberg: Springer; 2007. pp. 207–32.
86. Brown SDJ, Collins RA, Boyer S, Lefort M-C, Malumbres-Olarte J, Vink CJ, et al. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol Ecol Resour. 2012;12:562–5.
87. Sowers BA, Messenger E, Cook RM. Genotypomg With BHQplus fortified probes achiebe powerful discrimination. In: Biosearch Technologies. 2007.

Publisher's Note