Next-generation sequencing, isolation and characterization of 14 microsatellite loci of *Canthon cyanellus* (Coleoptera: Scarabaeidae)

Luis Rodrigo Arce-Valdés1 · Rosa Ana Sánchez-Guillén1 · Janet Nolasco-Soto1 · Mario E. Favila2

Received: 28 April 2021 / Accepted: 15 September 2021 / Published online: 13 October 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

**Abstract**

**Background** We used Illumina paired-end sequencing to isolate and characterize microsatellites of *Canthon cyanellus*, a Neotropical roller dung beetle, encompassing several lineages within its distribution range.

**Methods and results** We examined *C. cyanellus* specimens collected at eight different localities in Mexico (two or three specimens per locality). We initially performed amplification tests with 16 loci, but two of which were unsuccessful. The 14 remaining microsatellites were polymorphic, with 2–16 alleles each. The expected and observed heterozygosity ranged from 0.11 to 0.76 and from 0.20 to 0.78, respectively.

**Conclusions** These microsatellites will help to assess structure at the population and lineage levels, identify zones of potential hybridization between lineages, and draw a more precise geographic delimitation of *C. cyanellus* lineages.

**Keywords** *Canthon cyanellus* · Microsatellites · Isolation · Mexico

**Introduction**

*Canthon cyanellus* is a Neotropical roller dung beetle (Scarabaeidae: Scarabaeinae) that feeds on dung and carrion but only reproduces on the latter [1, 2]. This diurnal species inhabits forests, edges, and hedgerows [3]. Similar to other dung beetle species, *C. cyanellus* provides ecological services such as nutrient cycling, carrion burial, seed dispersal, and fly control [4]. The cuticular color of *C. cyanellus* varies widely across its distribution from northern Brazil to Texas [5, 6]. The coloration polymorphism ranges from a predominantly reddish-brown in northern South America and Central America to entirely green or even blue in tropical regions of Mexico and Texas. This variation in cuticular color enabled Halffter [7] to establish three subspecies. *C. cyanellus cyanellus* LeConte, 1859 is distributed in Mexico along the Pacific and Gulf of Mexico slopes to Texas; it is mostly green, but green- or blue-colored specimens can be found in the northern plains of the Gulf of Mexico [7]. *C. cyanellus sallei* is distributed from Guatemala to Peru [8]; its dorsal surface is predominantly reddish-brown with greenish shades, elytra that are reddish-brown with greenish–black markings, and a light colored pronotal disc, usually with reddish or yellowish–brown hues. *C. cyanellus violetae* is endemic to Chiapas in southern Mexico, with pronotum and pygidium that are typically a reddish–brown [7]. However, the taxonomy of this group has been one of the most complex of the *Canthon* genus. The coloration polymorphism and the absence of diagnosable and reliable morphological characters have created uncertainty about its genealogical limits, which has led to the conflicting recognition and delimitation of the species or subspecies that comprise it.

The reproductive behavior, chemical communication, and ecology of *C. cyanellus* have been extensively studied [1–4, 9–14, 14–21]. Ortíz-Domínguez and collaborators [19] carried out experimental interpopulation crosses of *C. cyanellus* from the Gulf of Mexico coastal plains. These authors analyzed the composition of cuticular hydrocarbons used in sexual recognition and chemical communication [22]. Male–female interactions between individuals from populations geographically separated by more than 600 km were mainly aggressive. They fought for a food ball instead of rolling it cooperatively, as occurs between male–female pairs from the same population. Crosses between males and females from distant populations showed low fecundity and...
fertility, suggesting that C. cyanellus is undergoing an incipient reproductive isolation process.

Phylogeographic studies based on one nuclear (ITS2) and two mitochondrial (COI and 16S) loci resolved the C. cyanellus genealogy into seven clades from tropical regions of Mexico and one clade from Colombia [23, 24]. Five of these clades were well supported, while the remaining clades were not. The geographic distribution of haplotypes suggests limited historical gene flow and recent dispersal events between some clades [23, 24]. Nolasco-Soto and collaborators [24] did not align with the recovered clades, which therefore could not be separated into different species. Studies by Nolasco-Soto and collaborators [23, 24] indicate that the phylogeographic history of C. cyanellus in Mexico is complex.

In this context, we used Illumina sequencing to isolate and characterize microsatellite loci of the C. cyanellus complex. Microsatellite markers (short DNA fragments, 100–500 bp) are polymorphic, codominant markers that consist of tandem repeats of simple sequence motifs with high mutation rates (10⁻⁶–10⁻² mutations per site per generation) [26]. Characteristics of microsatellite markers can contribute to a better understanding of the population structure of this species and enable the quantification of gene flow between clades/populations. The geographical delimitation of the populations and lineages of C. cyanellus will allow for the study of ecological and geographical factors that may influence their genetic differentiation and diversification in the Neotropics.

Methods

Sample collection and DNA extraction

Specimens of C. cyanellus were collected from eight localities in the Pacific and the Gulf of Mexico slopes, the Sierra Madre Oriental, and the southeast and southwest of the state of Chiapas (see Table 1 and Supplementary Material 2). The geographic coordinates and elevation (masl) were recorded at each location. Twenty to thirty baited pitfall traps were placed at each location and left for 48 h; traps were filled with damp soil and fish or squid as bait. All the beetles collected were transported alive to the laboratory, where they were subsequently frozen at −70 °C. Both hind legs of each specimen were excised and ground to extract DNA following the DNeasy Blood & Tissue kit (QIAGEN) protocol.

Next-generation sequencing, De Novo contig assembly, and selection of microsatellites

Three genomic DNA samples of C. cyanellus (sampled from Tuxpan, Los Tuxtlas and El Vergel localities) were used to construct the Illumina library following the Illumina TruSeq DNA PCR-Free kit sequencing library protocol (Illumina Inc. USA). The library was sequenced using a 250-bp paired-end run on a MiSeq platform (Illumina Inc. USA). The quality of raw reads was checked with the software FastQC v0.11.3 [27] and filtered with cutadapt v1.8.3 [28]. Quality filtering consisted of clipping adapters, removing sequences shorter than 20 nt, and trimming the 3’ end to a threshold of −q = 20. The software SPAdes 3.10.1 [29] was used for de novo assembly, and the QDD 3.1.2 pipeline [30] was used for the detection, characterization, and primer design of the variable number tandem repeat (VNTR) loci; both programs were run with their default parameters.

The following criteria (based on Meglécz et al. [30]) were used to select primer pairs for testing from the set of primer design sequences detected in the “Tuxpan” and “Los Tuxtlas” samples: (i) select only one primer pair per locus, (ii) the sequence should contain only pure microsatellites in the target region with the motifs AAT or AAG; (iii) the primer alignment score of the amplified sequence should be 2 or 3; (iii) markers should be selected from different ranges (at least 20 bp between each other) for PCR product length to facilitate multiplexing; and (iv) right and left primer lengths should be ≤ 60 bp.

| State/locality (code) | n | Coordinates (latitude, longitude) | Elevation (m a.s.l.) |
|-----------------------|---|---------------------------------|---------------------|
| Chiapas, El Vergel (Ver)* | 2 | 14.7029, −92.2672 | 22 |
| Chiapas, Palenque (Nbet) | 2 | 17.2815, −91.6466 | 140 |
| Guerrero, Ixtapa (Ixt) | 2 | 17.6583, −101.5752 | 120 |
| Jalisco, Chamela (Cha) | 2 | 19.4997, −105.0229 | 90 |
| Tamaulipas, Gómez-Farías (GF) | 2 | 23.0480, −97.4661 | 379 |
| Oaxaca, Huautulco (Hua) | 2 | 15.7800, −96.0900 | 30 |
| Veracruz, Tuxpan (Tp)* | 3 | 20.9544, −97.4661 | 52 |
| Veracruz, Los Tuxtlas (Tx)* | 3 | 18.5833, −95.0667 | 120 |

n sample size

*One sample from each of these localities was used for whole-genome sequencing for microsatellite loci discovery.
Validation and genotyping of microsatellites

We performed PCR to confirm the amplification and polymorphism of the loci selected. PCR was carried out with the universal fluorescent labeling method [31] on 18 C. cyanellus samples from eight populations (Table 1). PCR amplifications were made in a total reaction volume of 15 µL using the GoTaq® Flexi DNA Polymerase PCR kit (Promega Corporation, USA). Each reaction contained 1X Green GoTaq® Flexi Buffer, 1.5 mM MgCl2, 200 µM dNTPs, 0.05 u/µL GoTaq® Polymerase, 0.5 µM reverse primer, 0.25 µM modified forward primer including the M13 sequence tail at the 5’ end, 0.25 µM labeled M13 forward primer with either FAM, HEX, or NED fluorescent probes, 1 µL of genomic DNA (15–25 ng), and ddH2O to 15 µL. PCRs were standardized to the following thermal cycling conditions: 94 °C for 3 min, followed by 30 cycles of 94 °C for 1 min, 50 °C for 1 min, and 70 °C for 2 min, and a final extension of 72 °C for 3 min. Automated fragment analysis of PCR products was conducted using gel capillary electrophoresis by Macrogen (Macrogen Inc., South Korea). Allele sizes were determined with GeneMarker 2.6.3 software (SoftGenetics, USA), and binning was performed with the Autobin [32] Excel® macro.

Polymorphism analyses

The total number of alleles (A), effective number of alleles (Ae), observed heterozygosity (Ho), and expected heterozygosity (He), inbreeding index (FIS), and number of private alleles per locality were estimated using the macro GenAlEx 6.503 [33]. Finally, we calculated the polymorphism information content (PIC) and the probability of identity (PID) for each locus using the software PowerMarker 3.25 [34] and Gimlet 1.3.3 [35], respectively.

Results

Genome sequencing and characterization of microsatellite loci

Between six and eight million paired-end reads were generated per sample (Table 2). Most reads had the expected sequence length (250 bp) and were of sufficiently high quality to pass the filtering; only ca. 1% of the reads of each sample were filtered out.

Between 90,000 and 115,000 contigs were successfully assembled per sample using the accepted reads. The average contig length ranged from 2000 to 3000 bp with an N50 of approximately 7000 bp.

A total of approximately 9000 VNTR loci were found per sample using the contigs; approximately 8000 of them were identified as perfect or pure microsatellites. As per the QDD3 glossary, we defined pure microsatellites as loci composed of a single, 2–6 bp motif, with no interruption, and at least five repetitions [30]. Some 135,000 primer pairs were generated per sample from the VNTR loci found and characterized.

Overall, the VNTRs of three C. cyanellus samples from each locality had a very similar genomic characterization (Fig. 1). Almost 90% of all the VNTR loci found met the definition of perfect microsatellites; approximately 9% were categorized as compound microsatellites, and the rest were categorized as minisatellite loci (Fig. 1a). Compound microsatellites were defined as microsatellite loci followed

---

**Table 2** Contig assembly and microsatellite loci discovery in three samples of C. cyanellus genomic DNA

| Sample locality | El Vergel (Ver) | Los Tuxtlas (Tx) | Tuxpan (Tp) |
|-----------------|----------------|-----------------|-------------|
| Paired raw reads | 6,872,992       | 8,252,350       | 6,059,490   |
| Sequence length  | 35–251 bp       | 35–251 bp       | 35–251 bp   |
| Paired reads after filtering | 8,282,801 (99.4%) | 8,207,156 (99.5%) | 5,984,230 (98.8%) |
| Assembled contigs | 89,037          | 115,230         | 113,636     |
| N50             | 8,607 bp        | 6,037 bp        | 6,434 bp    |
| N95             | 480 bp          | 441 bp          | 445 bp      |
| Average contig length | 2,922.3 bp | 2,231.5 bp | 2,263.8 bp |
| Minimum contig length | 128 bp         | 128 bp          | 128 bp      |
| Maximum contig length | 87,324 bp      | 66,814 bp       | 76,419 bp   |
| Total number of primer pairs | 137,585        | 138,176         | 132,930     |
| Total VNTR loci  | 9,794           | 9,705           | 9,333       |
| Perfect microsatellites | 8,718 (89.0%)  | 8,591 (88.5%)   | 8,315 (89.1%) |
| Final primer testing loci | 0              | 8               | 8           |

Perfect microsatellites are defined as VNTR loci composed of a single, 2–6 bp long motif with no interruptions and at least five repetitions.

VNTR Variable number tandem repeat
by 3–4 tandem repetitions of a 2–6 bp motif separated by a distance equal to or shorter than the longest of two tandem repetitions. Minisatellite loci were defined as two or more perfect or compound microsatellite loci in the same target region [30].

The frequency of the VNTR classes was inversely related to motif size (Fig. 1b). Dinucleotide VNTRs accounted for approximately 65% of all the loci, followed by trinucleotides with approximately 30% and tetranucleotides with 5%.

The average number of reiteration units seen in VNTR loci was close to the minimum threshold (five) employed by the QDD3 pipeline. The mean number of repetitions was 8.9 for dinucleotides, 5.5 for trinucleotides, 5.3 for tetranucleotides, 5.3 for pentanucleotides, and 8.4 for hexanucleotides; however, a higher number of repetitions was found in loci with shorter motifs (Fig. 1c; Dinucleotides: median = 5, Q3 = 9, Max = 66; Trinucleotides: median = 5, Q3 = 6, Max = 25; Tetranucleotides: median = 5, Q3 = 5, Max = 16).
We also found differences in the frequency of the various motifs in the VNTR loci. AT was the most common motif in dinucleotide VNTRs, and CG was the rarest (Fig. 1d); AAT and AGC were the most and least common motifs, respectively, in trinucleotide loci (Fig. 1e); AAAT was the most common motif in tetranucleotide loci, while ACAG and AACC were the rarest, with only one locus each.

**Validation of microsatellite loci**

Fourteen of the 16 primer pairs synthesized were successfully amplified and yielded single PCR products of the expected size; PCR amplification was not achieved in one locus, and reading issues occurred in another locus (Supplementary Material 3). The amplification success rate of the 14 markers was 95.2%, with only 12 missing data after amplifying 18 C. cyanellus samples. Most of these loci amplified fragments with lengths differing between them in three nucleotides. This result was expected, as primers were designed only on trinucleotide-perfect microsatellites; this facilitated the assignment of raw fragment lengths—often referred to as binning—into allele classes [36]. However, Ccy09 and Ccy12 showed a mostly continuous distribution of raw fragment sizes with differences of less than three nucleotides. Based on this, we decided to bin their alleles into the mononucleotide allele class (Supplementary Material 4).

**Polymorphism analyses**

The number of alleles in polymorphic loci ranged from 2 to 16, and the observed heterozygosity ranged from 0.11 to 0.82 (Table 3). Ten loci showed moderate polymorphism (4–6 alleles), four were marginally polymorphic (2–3 alleles), and only two were highly polymorphic (9 and 16 alleles). The effective number of alleles ($A_E$) was approximately half the number of observed alleles.

Eight loci showed high (>0.20) absolute $F_{IS}$ values, indicating large differences between their expected ($H_E$) and observed ($H_O$) heterozygosity. These loci displayed positive $F_{IS}$ values, indicating a deficit of observed heterozygote samples vs. expected under HW equilibrium. Polymorphism information content (PIC) ranged from 0.190 to 0.889, with a mean of 0.494.

The moderate genetic diversity observed in these markers conferred a relatively high probability of identity ($P_{ID}$). This parameter measures the probability of two unrelated individuals having the same genotype [37]. $P_{ID}$ values ranged from 60.1% (Ccy11) to 0.1% (Ccy12). However, when $P_{ID}$ was computed over the full set of 14 markers using the product rule, it reached a value of $1.299 \times 10^{-12}$ in unrelated individuals and $1.563 \times 10^{-4}$ in siblings.

Since we examined only a few (2–3) samples per locality, we measured the ability of these 14 markers to measure genetic structure by identifying the number of private alleles of each locus. Eleven of the 14 loci possessed at least one private allele at one locality (Fig. 2). Ten of these 11 loci had few (1–3) private alleles, and locus Tuxt06 had eight private alleles across six different localities.

**Discussion**

**Characterization of microsatellite loci**

Perfect microsatellite loci were the most common type of VNTR in C. cyanellus genomes, which has been similarly reported for the coconut rhinoceros beetle Oryctes rhinoceros (Linnaeus, 1758) [38] and the red flour beetle Tribolium castaneum (Herbst, 1797) [39]. This is in contrast to a recent comparison of 29 beetle genomes that showed imperfect microsatellites were the most common short-sequence repeat type [40]. Comparisons between genomic characterizations of microsatellite loci must be taken with caution due to a current lack of standardization of detection criteria and VNTR classes.

The most frequent VNTR class has been reported to be a highly heterogeneous feature of beetle genomes, particularly in Scarabaeidae species [40]. However, after excluding mononucleotide microsatellites (as we did in our analyses), dinucleotides and trinucleotides seem to be the most frequent microsatellite motif sizes [40, 41]. Our results are consistent with this observation, as we found a trend of decreasing VNTR frequency with an increase in motif size. Our results are also consistent with the general pattern of short sequence repeats to favor poly(A) and poly(T) motifs over poly(G) and poly(C) in eukaryotic genomes [40, 42, 43].

**Non-tandem repeat polymorphisms in Ccy09 and Ccy12**

The raw allele reads of Ccy09 and Ccy12 were not consistent with the expected three-nucleotide difference in size; as a result, we binned these loci as mononucleotide microsatellites. These markers might show evidence of non-tandem repeat polymorphisms, such as SNPs or indel mutations, and thus would not be perfect microsatellites. Complex mutational processes that produce variations in allele length at microsatellite loci have been reported in several eukaryote groups, including yeasts [44, 45], fish [46], and insects [39].

 Imperfect microsatellites have pros and cons when used as molecular markers for DNA fingerprinting and population genetics compared to perfect microsatellites [47]. Additional alleles provide more information that is useful for paternity testing and mapping studies and can achieve a
Table 3  Characterization of 14 polymorphic microsatellite markers of *C. cyanellus*

| Locus | GenBank  | Forward Primer | Reverse Primer | Motif      | n  | Size    | A    | A<sub>E</sub> | H<sub>O</sub> | H<sub>E</sub> | F<sub>IS</sub> | PIC    | PID    |
|-------|----------|----------------|----------------|------------|----|---------|------|-------------|-------------|-------------|-----------|--------|--------|
| Ccy01 | MW455319 | GGGTATCTGTGGGACCTAAT | TGGATTAAATAACTGTACCAGCA | AAT       | 17 | 212–227 | 6    | 4.48        | 0.76        | 0.78        | 0.016     | 0.743  | 0.056  |
| Ccy02 | MW455320 | CGATTGCGGTGGTCTTACAGAGA | AAAGCAACAACCTGTCTCAT | AAT       | 18 | 96–108  | 5    | 2.10        | 0.39        | 0.52        | 0.257     | 0.494  | 0.194  |
| Ccy03 | MW455321 | CTTCTGAAGAGAACACAACCGA | ACTAAACTACATTATCTTGGCA | AAT       | 17 | 177–183 | 3    | 1.35        | 0.18        | 0.26        | 0.315     | 0.236  | 0.520  |
| Ccy04 | MW455322 | AGTCGCGAAGAACCTGACGAGA | TCTTACTGAATACGTGTATTGTAACA | AAT       | 18 | 120–129 | 4    | 1.42        | 0.22        | 0.29        | 0.246     | 0.280  | 0.451  |
| Ccy05 | MW455323 | ACAATATTGTGTCGTGCA | TCGAAATGTTCTCTGAGAAATTCTTGG | AAT       | 17 | 304–310 | 4    | 2.42        | 0.53        | 0.59        | 0.097     | 0.510  | 0.216  |
| Ccy06 | MW455324 | TCTTGGTGTAAAAGTCTTACA | TGAATCGATAGCAGGAATAAGAGC | AAT       | 17 | 247–260 | 3    | 1.53        | 0.41        | 0.34        | -0.196    | 0.313  | 0.405  |
| Ccy07 | MW455325 | TGGCTAGTGTGATCTGAGGACA | GCTACATTCTCTGCCTTGACC | AAT       | 17 | 195–207 | 4    | 1.74        | 0.41        | 0.42        | 0.029     | 0.385  | 0.312  |
| Ccy08 | MW455326 | TCGAAATTTGATGCAACCCCGGA | TCGCTCGATATCACCAGGA | AAT       | 16 | 145–151 | 4    | 2.40        | 0.19        | 0.58        | 0.679     | 0.496  | 0.235  |
| Ccy09*| MW455327 | ATCTTGATTTGTTATTTGTAACCTT | TCTTGAGGTGAGATAAGATCTGCA | AAT       | 16 | 204–214 | 9    | 6.24        | 0.38        | 0.84        | 0.553     | 0.822  | 0.019  |
| Ccy10 | MW455328 | TGAATGTAATACACTGACCAGCA | GGTATCGTGTGGGCACCTAT | AAT       | 16 | 214–226 | 5    | 3.91        | 0.69        | 0.74        | 0.076     | 0.703  | 0.076  |
| Ccy11 | MW455329 | GCGAACAATCTAATCTTATCGCTT | ACTAAATGCTCTGTGATCGCTGCT | AAT       | 18 | 188–197 | 3    | 1.25        | 0.11        | 0.20        | 0.450     | 0.190  | 0.601  |
| Ccy12*| MW455330 | ATCCAGGCAAACCGAATCTCTT | ACACACAACATACAGATCTCTCGA | AAT       | 17 | 243–267 | 16   | 12.04       | 0.82        | 0.92        | 0.102     | 0.889  | 0.001  |
| Ccy13 | MW455331 | ACAGAAGTCTATGAAAGGACA | ACTTACTCCGAATTCTTATCTT | AAT       | 18 | 176–179 | 2    | 1.60        | 0.28        | 0.38        | 0.259     | 0.305  | 0.423  |
| Ccy14 | MW455332 | AAGATCCCTAATGCTAGTGGCA | CGGACCAGAACTAACCTCAC | AAG       | 18 | 142–151 | 4    | 2.54        | 0.33        | 0.61        | 0.450     | 0.551  | 0.168  |
| Average|          |               |                 |             | 5.14 | 3.22    | 0.41 | 0.53        | 0.238       | 0.494       | 0.263     |

*n* sample size, *A* number of alleles, *A<sub>E</sub>* effective number of alleles, *H<sub>O</sub>* observed heterozygosity, *H<sub>E</sub>* expected heterozygosity, *F<sub>IS</sub>* inbreeding coefficient, *PIC* polymorphism information content, *PID* probability of identity

*Loci that showed alleles with size differences shorter than three nucleotides*
higher resolution of population structure [47, 48]. Ccy09 and
Ccy12 showed higher genetic diversity and PIC values than
the other loci. However, the full advantages of imperfect
microsatellites can only be achieved by using DNA sequenc-
ing. Moreover, mutational models of imperfect microsatel-
lites are more complex than those of perfect microsatellites
[47].

Polymorphism analyses

Overall, our markers displayed moderate levels of genetic
diversity with an effective number of alleles that suggests
there are few dominant alleles per locus. Additionally, a
deficit of heterozygote samples was found compared to
what would be expected under HW equilibrium. This was
expected due to the small sample size used to evaluate the
primers and the use of specimens from multiple localities to
sample as many alleles as possible. Since a potential hidden
genetic structure might be causing the heterozygote deficit—
often referred to as the Wahlund effect [49] —, we decided
not to test for departure from HW equilibrium.

More accurate estimates of diversity for each marker
could be obtained by increasing the genotyped sample size.
However, with this small sample size, we estimated a $P_{ID}$
of $1.299 \times 10^{-12}$ for unrelated individuals and $1.563 \times 10^{-4}$
for siblings over the full set of 14 markers. Based on these
values, the probability of sampling two unrelated individuals
with the same genotypes across the 14 markers is 1 in ca.
770 billion and 1 in 6000 for sibling beetles. This highlights
the highly informative power of the full set of 14 microsat-
ellite markers. Finally, the robustness of these markers to
measure the genetic structure in *C. cyanellus* populations
and lineages should be evaluated using larger number of
samples from each locality.

Conclusion

The set of 14 *C. cyanellus* microsatellites that were isolated
and characterized provides a valuable genetic tool to assess
structure at the population or lineage levels and analyze
hybridization processes. These microsatellites will help to
better understand the geographic distribution of the lineages
in this species complex. Further studies including additional
populations from Central and South America would contrib-
ute to the establishment of a sound taxonomic delimitation
for these lineages.
Supplementary Information  The online version contains supplementary material available at https://doi.org/10.1007/s11033-021-06761-8.

Acknowledgements  Sequencing was performed at the Genomics Core Facility of Servicio Central de Soporte a la Investigación Experimental (SCSIIE), Universidad de Valencia. This study was funded by a research grant (CB 282922) from Consejo Nacional de Ciencia y Tecnología (CONACyT) awarded to Rosa Ana Sánchez Guillén. María Elena Sánchez-Salazar edited the English manuscript.

Author contributions  All the authors contributed to the conception and design of this study. Samples were collected by JN-S. Wet lab techniques were performed by JN-S and LRA-V. Data analyses were carried out by RAS-G and LRA-V. The manuscript was written and edited equally by all the authors. All the authors read and approved the final version of the manuscript.

Funding  This study was funded by a research grant from Consejo Nacional de Ciencia y Tecnología (CONACyT) awarded to Rosa Ana Sánchez Guillén (CB 282922).

Data availability  The database of microsatellite genotypes created and analyzed in this study is included in the Supplementary Material 1.xlsx file that accompanies the published article. Additional data supporting the findings of this study are available, upon request, from the corresponding author.

Code availability  Not applicable.

Declarations

Conflict of interest  The authors declare that they have no conflicts of interest that are relevant to the content of this article.

Ethical approval  Not applicable.

Consent for publication  Neither the article nor portions of it have been previously published elsewhere. The manuscript is not under consideration for publication in any other journal and will not be submitted elsewhere until the Molecular Biology Reports editorial process is completed. All authors consent to the publication of the manuscript in Molecular Biology Reports should the article be accepted.

References
1. Favila ME, Díaz A (1996) Canthon cyanellus cyanellus LeConte (Coleoptera: Scarabaeidae) makes a nest in the field with several brood balls. Coleopt Bull 50:52–60
2. Favila ME (2001) Historia de vida y comportamiento de un escarabajo necrófago: Canthon cyanellus cyanellus LeConte (coleoptera: Scarabaeinae). Folia Entomológica Mex 40:245–278
3. Arellano L, León-Cortés JL, Ovaskainen O (2008) Patterns of abundance and movement in relation to landscape structure: a study of a common scarab (Canthon cyanellus cyanellus) in Southern Mexico. Landsc Ecol 23:69–78. https://doi.org/10.1007/s10980-007-9165-8
4. Portela Salomão R, González-Tokman D, Dátillo W et al (2018) Landscape structure and composition define the body condition of dung beetles (Coleoptera: Scarabaeinae) in a fragmented tropical rainforest. Ecol Indic 88:144–151. https://doi.org/10.1016/j.ecolind.2018.01.033
5. Robinson M (1948) A review of the species of Canthon inhabiting the United States (Scarabaeidae: Coleoptera). Trans Am Entomol Soc 1890–74:83–100
6. Solís-Blanco A, Kohlmann B (2002) El género Canthon (Coleoptera: Scarabaeidae) en Costa Rica. G Ital Entomol 10:1–68
7. Halfter-Salas G (1961) Monografía de las especies norteamericanas del género Canthon Hoffm. (Coleoptera: Scarabaeidae). Univ Zulia 20:225–320
8. Blackwelder RE (1944) Checklist of the coleopterous insects of Mexico, Central America, the West Indies, and South America. US Government Printing Office, Washington
9. Bellés X, Favila M (1983) Protection chimique du nid chez Canthon cyanellus cyanellus Le Conte [Col. Scarabaeidae]. Bull Société Entomol Fr 88:602–607
10. Cortez V, Favila ME, Verdú JR, Ortiz AJ (2012) Behavioral and antennal electrophysiological responses of a predator ant to the pygidial gland secretions of two species of neotropical dung roller beetles. Chemoecology 22:29–38
11. Cortez V, Verdú JR, Ortiz AJ et al (2015) Chemical diversity and potential biological functions of the pygidial gland secretions in two species of Neotropical dung roller beetles. Chemoecology 25:201–213. https://doi.org/10.1007/s00049-015-0189-2
12. Favila M, Chamorro-Florescano I (2008) Male reproductive status affects contest outcome during nidification in Canthon cyanellus cyanellus LeConte (Coleoptera: scarabaeidae). Behaviour 145:1811–1821. https://doi.org/10.1163/156853908X445637
13. Favila M, Chamorro-Florescano I (2009) The reproductive status of both sexes affects the frequency of mating and the reproductive success of males in the ball roller beetle Canthon cyanellus cyanellus (coleoptera: Scarabaeidae). Behaviour 146:1499–1512. https://doi.org/10.1163/156853909X445560
14. Chamorro-Florescano IA, Favila ME, Macías-Ordóñez R (2011) Ownership, size and reproductive status affect the outcome of food ball contests in a dung roller beetle: when do enemies share? Evol Ecol 25:277–289
15. Chamorro-Florescano IA, Favila ME (2016) Male success in intrasexual contests extends to the level of sperm competition in a species of dung roller beetle. Ethology 122:53–60. https://doi.org/10.1111/eth.12443
16. Chamorro-Florescano IA, Favila ME, Macías-Ordóñez R (2017) Contests over reproductive resources in female roller beetles: outcome predictors and sharing as an option. PLoS ONE 12:e0182931. https://doi.org/10.1371/journal.pone.0182931
17. Favila ME, Ortiz-Domínguez M, Chamorro-Florescano I, Cortez-Gallardo V (2012) Comunicación química y comportamiento reproductor de los escarabajos rodadores del estiércol (Scarabaeinae: Scarabaeini): aspectos ecológicos y evolutivos, y sus posibles aplicaciones [pp. 141–164]. Temas Sel En Ecol Quim Insectos JC Rojas EA Malo Ed El Col Front Sur Tapachula Mex
18. Favila ME, Nolasco J, Florescano IC (2005) Sperm competition and evidence of sperm fertilization patterns in the carrion ball-roller beetle Canthon cyanellus cyanellus LeConte (Scarabaeidae: Scarabaeinae). Behav Ecol Sociobiol 59:38. https://doi.org/10.1007/s00265-005-0006-y
19. Ortiz-Domínguez M, Favila ME, Mendoza-López MR et al (2006) Epicuticular compounds and sexual recognition in the ball-roller scarab, Canthon cyanellus cyanellus, Entomol Exp Appl 119:23–27. https://doi.org/10.1111/j.1570-7458.2006.00388.x
20. Ortiz-Domínguez M, Favila ME, Mendoza-López MR (2006) Mate recognition differences among allopatric populations of the scarab Canthon cyanellus cyanellus (Coleoptera: Scarabaeidae). Ann Entomol Soc Am 99:1248–1256. https://doi.org/10.1603/0013-8746(2006)99[1248:MRDAAP]2.0.CO;2
21. Ortiz-Domínguez M, Favia-Castillo ME, González D, Zuñiga D (2010) Genetic differences in populations of the ball roller scarab Canthon cyanellus cyanellus (coleoptera: Scarabaeidae): a preliminary analysis. Elytron 24:99–106

22. Chung H, Carroll SB (2015) Wax, sex and the origin of species: dual roles of insect cuticular hydrocarbons in adaptation and mating. BioEssays 37:822–830. https://doi.org/10.1002/bies.201500014

23. Nolasco-Soto J, González-Astorga J, Espinosa Monteros A et al (2017) Phylogeographic structure of Canthon cyanellus (coleoptera: Scarabaeidae), a Neotropical dung beetle in the Mexican transition Zone: insights on its origin and the impacts of pleistocene climatic fluctuations on population dynamics. Mol Phylogenet Evol 109:180–190. https://doi.org/10.1016/j.ympev.2017.01.004

24. Nolasco-Soto J, Favia ME, De Los E, Monteros A et al (2020) Variations in genetic structure and male genitalia suggest recent lineage diversification in the neotropical dung beetle complex Canthon cyanellus (Scarabaeidae: Scarabaeinae). Biol J Linn Soc 131:505–520. https://doi.org/10.1093/biolinnean/blaa131

25. Ratnasingham S, Hebert PDN (2013) A DNA-based registry for individual identification data. Mol Ecol Notes 2:377–379. https://doi.org/10.1111/j.1471-8286.2002.00238.x

26. Amos W, Hoffman JL, Frodsham A et al (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10–14. https://doi.org/10.1111/j.1471-8286.2006.01560.x

27. Li CC (1976) First course in population genetics. Boxwood, Pac Grove CA

28. Manjeri G, Muhamad R, Faridah QZ, Tan SG (2014) Development of single locus DNA microsatellite markers in Oryctes rhinoceros (Linnaeus) using S’ anchored RAMs-PCR method. J Genet 93:92–96. https://doi.org/10.1007/s12041-012-0189-8

29. Behura SK, Severson DW (2015) Motif mismatches in microsatellites: insights from genome-wide investigation among 20 insect species. DNA Res 22:29–38. https://doi.org/10.1093/dnares/dsu036

30. Song X, Yang T, Yan X et al (2020) Comparison of microsatellite distribution patterns in twenty-nine beetle genomes. Gene 757:144919. https://doi.org/10.1016/j.gene.2020.144919

31. Santana QC, Coetzee MPA, Steenkamp ET et al (2009) Microsatellite discovery by deep sequencing of enriched genomic libraries. Biotechniques 46:217–223. https://doi.org/10.2144/000113085

32. Katti MV, Ranjekar PK, Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol 18:1161–1167. https://doi.org/10.1093/oxfordjournals.molbev.a003903

33. Töth G, Gaspári Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10:967–981. https://doi.org/10.1101/gr.7.9.967

34. Eder MLR, Rosa AL (2020) Non-tandem repeat polymorphisms at microsatellite loci in wine yeast species. Mol Genet Genomics 295:685–693. https://doi.org/10.1007/s00438-020-01652-2

35. Sampaio P, Correia A, Gusmão L et al (2007) Sequence analysis reveals complex mutational processes for allele length variation at two polymorphic microsatellite loci in Candida albicans. Communicating current research and educational topics and trends in applied microbiology. Formatex, Badajoz, pp 926–935

36. Angers B, Bernatchez L (1997) Complex evolution of a salmonid microsatellite locus and its consequences in inferring allometric divergence from size information. Mol Biol Evol 14:230–238. https://doi.org/10.1093/oxfordjournals.molbev.a025759

37. Schlötterer C, Zangerl B (1999) The use of imperfect microsatellites for DNA fingerprinting and population genetics. In: Epplen JT, Labuohn T (eds) DNA profiling and DNA fingerprinting. Birkhäuser, Basel, pp 153–165

38. Viard F, Franck P, Dubois M-P et al (1998) Variation of microsatellite size homoplasy across electromorphs, loci, and populations in three invertebrate species. J Mol Evol 47:42–51

39. Von Wahlund S (1928) Zusammensetzung von population und korrelationsers cheimungen von Standpunkt der veierbungs lehreaus betachlet. Hereditas 11:65–106

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.