

Development of a PCR-Based Molecular Sex-Determination Assay for the Endangered Dung Beetle *Gymnopleurus mopsus*

Dong Eon Kim

National Institute of Ecology

Eunji Jeong

Bioleap Incorporated

Jongdae Park

National Institute of Ecology

Kuem Hee Jang

National Institute of Ecology

Young Jun Jung

`june5763@nie.re.kr`

National Institute of Ecology

Method Article

Keywords: *Gymnopleurus mopsus*, molecular sexing, transformer-2(tra2), sex-associated marker, conservation genetics, endangered species

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11286765/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Reliable sex identification is important for the conservation breeding and restoration of endangered species, particularly when morphological sex determination is difficult or unavailable. *Gymnopleurus mopsus* is an endangered dung beetle regarded as regionally extinct in South Korea, yet no molecular method for sex identification has been reported for this species. In this study, comparative analysis of whole-genome sequencing (WGS) data identified a sex-associated sequence difference within the transformer-2 (*tra2*) locus. A conventional PCR assay targeting this region (*tra2*-region3) produced a consistent sex-associated amplification pattern in all specimens examined: males yielded a single amplicon of approximately 1,168 bp, whereas females exhibited a characteristic double-band profile of approximately 1,012 and 1,168 bp. These amplification profiles were fully concordant with phenotypically determined sex within the examined specimens. TA cloning followed by bidirectional Sanger sequencing confirmed the identity of the PCR amplicons and revealed a 156-bp sequence difference associated with the additional ~ 1,012-bp amplicon observed in females, consistent with the sequence variation identified from the WGS data. Although the functional relationship between this sequence variation and the sex-determination pathway remains unresolved, the observed amplification pattern provides a readily distinguishable molecular marker for sex identification in *G. mopsus*. This study provides the first PCR-based approach for molecular sex identification in a dung beetle species. The *tra2*-region3 marker may serve as a useful tool for sex identification in *G. mopsus* and for supporting conservation breeding and restoration programs.

Introduction

The dung beetle (*Gymnopleurus mopsus* Pallas, 1781; Coleoptera: Scarabaeidae) is an ecologically important species that contributes to dung removal, soil aeration, nutrient cycling, and secondary seed dispersal, thereby maintaining ecosystem functioning and grassland health (Nichols et al. 2008). Dung beetles are generally sensitive to habitat modification, changes in livestock management, and exposure to veterinary pharmaceuticals. In South Korea, *G. mopsus* underwent a severe historical decline and is currently regarded as regionally extinct. Consequently, *G. mopsus* is currently designated as a Class II Endangered Wildlife species under the Wildlife Protection and Management Act and has become a priority species for national restoration programs (MOE 2018; NIBR 2023). Current conservation programs for *G. mopsus* focus on ex situ breeding, genetic resource management, and the long-term reintroduction of individuals into suitable habitats (MOE 2018). The success of these conservation breeding programs depends not only on maintaining genetic diversity but also on achieving balanced sex ratios and accurately monitoring population structure throughout the breeding process (Simmons and Ridsdill-Smith 2011). However, reliable sex identification remains challenging in immature individuals and in adults with damaged or incompletely preserved diagnostic structures (Lim et al. 2015). Morphological sex identification in *G. mopsus* is possible in adults based on sexual differences in the protibial spur, which differs in curvature and apparent length between males and females (Lim et al. 2015). However, this approach is primarily applicable to adult specimens with intact diagnostic structures, potentially limiting reliable sex determination in immature individuals and specimens with

damaged appendages. These limitations highlight the need for molecular approaches that enable reliable sex determination when morphological identification is difficult or unavailable. In insects, somatic sex determination is regulated through a highly conserved genetic cascade involving transformer (*tra*), transformer-2 (*tra2*), doublesex (*dsx*), intersex (*ix*), and sex-lethal (*sxl*) (Schütt and Nöthiger 2000; Gempe and Beye 2011). Within this pathway, the Tra/Tra2 protein complex regulates sex-specific alternative splicing of *dsx* pre-mRNA, generating male- and female-specific Dsx isoforms that direct somatic sexual differentiation (Sánchez 2008; Verhulst et al. 2010). Although the overall architecture of this regulatory pathway is generally conserved across Scarabaeidae (Mitchell et al. 2024), considerable sequence divergence within intronic regions and splice junctions can complicate the development of universal PCR-based diagnostic markers (Minezawa et al. 2020; Shukla and Palli 2021). Consequently, molecular markers developed for one species cannot be assumed to be directly applicable to closely related taxa without empirical validation. Among these genes, *dsx* has been the most extensively investigated target because of its characteristic sex-specific alternative splicing patterns (Geuverink and Beukeboom 2014). However, variation in transcript isoforms, gene structure, and amplification specificity may reduce its utility as a diagnostic marker in some insect taxa. Genes involved in the insect sex-determination pathway represent biologically relevant candidate regions for screening sex-associated genomic variation. However, their functional involvement in sex determination does not necessarily imply the presence of sex-specific differences at the genomic DNA level, and candidate loci therefore require empirical evaluation in the target species. To our knowledge, no PCR-based molecular sex-determination marker has yet been reported for a dung beetle species. Based on this region, we developed and empirically evaluated a PCR-based molecular sex-determination assay for *G. mopsus*. This study provides a genome-informed molecular resource with potential applications in conservation breeding and restoration management of *G. mopsus*.

Materials and methods

Sample Collection and Genomic DNA Extraction

A total of ten adult *G. mopsus* specimens were used in this study, comprising eight wild-origin individuals (four males, M1–M4; four females, F1–F4) and two whole-genome sequencing (WGS) reference specimens (one male positive control [MPC] and one female positive control [FPC]) provided by the Research Center for Endangered Species, National Institute of Ecology (NIE), Seocheon, South Korea. Thoracic muscle tissue was mechanically homogenized using a TissueLyser II (QIAGEN, Hilden, Germany). Genomic DNA (gDNA) was extracted using the ZymoBIOMICS DNA Miniprep Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's protocol. DNA concentration and purity (A260/A280) were measured using a QIAxpert spectrophotometer (QIAGEN, Hilden, Germany). Purified gDNA was stored at – 20°C until subsequent PCR analysis.

Whole-genome sequencing, genome assembly, and candidate gene selection

Whole-genome sequencing (WGS) was performed using genomic DNA from one female (*G_mopsus_F*) and one male (*G_mopsus_M*) individual. Sequencing libraries were prepared using the Hieff DNA library kit (YEASEN, Shanghai, China), and library quality was assessed using a TapeStation HS D5000 system. Paired-end sequencing generated 151-bp reads, and the raw sequencing data were preprocessed using fastp v1.0.1 to remove adapter sequences and low-quality reads and bases. Genome size and sequence characteristics were estimated from 21-mer frequency distributions generated using KMC v3.2.4 and analyzed with GenomeScope 2.0. De novo genome assembly was independently performed for the female and male samples using MaSuRCA v4.1.4. Assembly consistency was evaluated by mapping the quality-filtered reads back to the corresponding draft genome assemblies using the BWA-MEM algorithm, and genome completeness was assessed using BUSCO v6.0.0 with the insecta_odb12 dataset.

Reference sequences of insect sex-determination genes were retrieved from the NCBI GenBank database, including *doublesex* (*dsx*; XM_071197859.1 and XM_023052271.2), *transformer-2* (*tra2*; LC385013.1), *intersex* (*ix*; XM_023056442.2), and *sex-lethal* (*sxl*; LC385009.1). BLASTn searches were performed against the female and male draft genome assemblies to identify genomic scaffolds corresponding to the candidate genes. Scaffolds were selected based on BLAST matches with a cumulative alignment length of ≥ 200 bp, and the corresponding female and male scaffold sequences were compared using dot-plot analyses to identify sequence conservation and potential sex-associated sequence or structural differences. Candidate regions showing potentially informative male–female sequence differences and suitable primer-binding sites were selected for subsequent PCR evaluation.

Primer design and candidate locus evaluation

Based on sequence differences identified between the male and female genomic scaffolds, female-specific (Fdsx-F/R) and male-specific (Mdsx-F/R) primer sets targeting the *dsx* locus were initially designed. Primer performance was evaluated using the validation specimens under two PCR conditions: annealing at 58°C for 20 s for 38 cycles (Condition 1) and annealing at 60°C for 20 s for 35 cycles (Condition 2). Following the unsuccessful evaluation of the selected *dsx*-based primer sets, three candidate regions within the *tra2* gene (*tra2*-region1, *tra2*-region2, and *tra2*-region3) were selected based on the WGS-derived sequence comparisons and evaluated for molecular sex determination. The primer pair targeting *tra2*-region3 consisted of *tra2*-F3 (5'-CTCAATTCTCAAAAATGTTGTTTCG-3') and *tra2*-R3 (5'-AATGCCGCTCCTGTTTGTTC-3').

Conventional PCR and electrophoresis

Conventional PCR and electrophoresis

PCR amplification was performed in 20- μ L reaction mixtures containing 10 μ L of Tomic™-cPCR premix (Bioto, Daejeon, Korea), 1 μ L each of forward and reverse primers (10 μ M), 2 μ L of genomic DNA (approximately 5–10 ng), and nuclease-free water to a final volume of 20 μ L. Amplification was carried

out using a Veriti 96-Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) under the following conditions: initial denaturation at 95°C for 3 min; 35 cycles of denaturation at 95°C for 30 s, annealing at 58°C for 20 s, and extension at 72°C for 45 s; followed by a final extension at 72°C for 5 min. PCR products were separated on 1.5% agarose gels prepared in 1× TAE buffer by electrophoresis at 180 V for 20 min. Gels were stained with ethidium bromide and visualized under ultraviolet illumination.

TA cloning and Sanger sequencing validation

To characterize the nucleotide sequences of the diagnostic amplicons and verify that the observed band-size polymorphism matched the genomic predictions derived from the WGS data, PCR products from representative male and female specimens were subjected to TA cloning using the TOPcloner™ TA Kit (Enzymomics, Daejeon, Korea) according to the manufacturer's instructions. Fresh PCR products were ligated into the pTOP TA V2 vector and transformed into competent *Escherichia coli* DH5α cells (Enzymomics). Transformants were plated onto LB agar containing ampicillin, X-gal, and IPTG for blue-white selection. Colony PCR was conducted to verify insert sizes, and multiple independent white colonies representing the ~ 1,168-bp and ~ 1,012-bp fragments were selected. Plasmid DNA was isolated and sequenced by Macrogen Inc. (Seoul, Korea) via bidirectional Sanger sequencing using vector-specific primers (M13F and M13R). The resulting sequence chromatograms were manually inspected for quality using SnapGene Viewer (Dotmatics, Boston, MA, USA), and consensus sequences were aligned against the male and female *G. mopsus* draft WGS scaffolds using ClustalW implemented in MEGA6.

Results and Discussion

Identification and evaluation of candidate sex-determination loci

Local BLASTn analysis using Scarabaeidae reference sequences for the candidate sex-determination genes (*dsx*, *tra2*, *ix*, and *sxl*) identified homologous regions in the *G. mopsus* draft genome assemblies, with individual BLAST alignments showing sequence identities ranging from approximately 72.9% to 94.9%. These values represent the similarity between the Scarabaeidae reference gene sequences and the corresponding *G. mopsus* genomic regions, rather than the similarity between female and male *G. mopsus* genomes. Subsequent direct comparison of the corresponding female and male scaffolds revealed substantially higher sequence conservation. In particular, the *dsx*- and *ix*-associated scaffolds showed approximately 96.58–99.82% sequence identity between the sexes, whereas the *tra2*-associated scaffolds were generally collinear but contained several discontinuous alignment regions, suggesting local sequence differences, insertion/deletion events, or assembly-associated variation.

Based on the sequence differences identified within the *dsx* locus, female-specific (*Fdsx*) and male-specific (*Mdsx*) primer pairs were initially designed and experimentally evaluated. Under the initial PCR

condition (58°C annealing, 38 cycles), the resulting amplification profiles showed poor concordance with phenotypic sex (Fig. 1A). Female-specific primers produced inconsistent amplification among female specimens and generated non-specific products in male specimens, whereas male-specific primers failed to produce reproducible sex-specific amplification patterns. Increasing the annealing temperature to 60°C and reducing the number of amplification cycles to 35 did not eliminate the non-specific amplification or produce a consistent sex-specific pattern (Fig. 1B).

These results indicated that the selected *dsx* regions and primer sets were unsuitable for reliable sex-associated PCR discrimination under the conditions evaluated. Although *dsx* is a key component of the insect somatic sex-determination pathway, sequence conservation alone does not necessarily ensure its suitability as a PCR-based diagnostic marker. Local sequence variation, genomic organization, or primer-binding characteristics may affect amplification specificity and reproducibility. Therefore, experimental evaluation of multiple candidate loci was necessary to identify a more reliable molecular marker.

Screening and validation of the *tra2*-region3 marker

Following the unsuccessful validation of the *dsx*-based markers, three candidate regions within the transformer-2 (*tra2*) locus (*tra2*-region1, *tra2*-region2, and *tra2*-region3) were evaluated using representative male and female samples (M2, F2, MPC, FPC, and NTC) under the optimized PCR condition of 58°C annealing for 35 cycles (Fig. 2). The three candidate regions exhibited different amplification profiles. *Tra2*-region1 showed apparent differences in band size between male and female samples; however, additional background amplification limited its diagnostic reliability. *Tra2*-region2 did not consistently produce the expected amplification product. In contrast, *tra2*-region3 generated a distinct and reproducible sex-associated amplification pattern and was therefore selected for further empirical evaluation. The *tra2*-region3 marker was subsequently evaluated using male specimens (M1–M4), female specimens (F1–F4), positive controls (MPC and FPC), and a non-template control (NTC) (Fig. 3). All male specimens produced a single amplicon of approximately 1,168 bp, whereas all female specimens exhibited a characteristic double-band pattern consisting of approximately 1,012-bp and 1,168-bp fragments. Within the limited sample set examined, the observed PCR profiles were fully concordant with phenotypically determined sex. Thus, within the specimens available for this study, the *tra2*-region3-targeting assay produced a consistent amplification pattern associated with phenotypic sex.

Interestingly, the experimentally observed female amplification pattern was more complex than initially expected from the WGS-based sequence comparison. Whereas the genomic comparison suggested localized differences between the female and male *tra2*-associated scaffolds, PCR amplification showed that females produced both the approximately 1,168-bp fragment observed in males and an additional approximately 1,012-bp fragment. This pattern suggested the presence of sex-associated sequence variation within or near the *tra2*-region3 locus, which was subsequently examined by TA cloning and Sanger sequencing.

Sequence validation and characterization of the *tra2*-region3 amplicons

To elucidate the molecular basis of the sex-differentiated banding profile and determine whether the experimental amplicons corresponded to the genomic predictions from the WGS data, the cloned PCR products representing both sexes were sequenced (Supplementary Data 1). Sequence alignment demonstrated that the ~ 1,168-bp amplicons obtained from males shared high sequence identity (99.7–100%) with the male reference locus (MPC_WGS), with minor length variation (1,168–1,171 bp) attributable to small homopolymeric adenine tract variations. In females, the common ~ 1,168-bp amplicon showed complete sequence conservation with the male sequence, whereas the female-specific 1,012-bp amplicon harbored a distinct 156-bp deletion that precisely matched the sequence predicted from the female reference genome (FPC_WGS) (Supplementary Data 1). This sequence-level concordance indicates that the observed double-band pattern in females reflects authentic sex-associated genomic variation within the *tra2-region3* locus rather than non-specific PCR artifacts. Although the precise chromosomal organization and cytogenetic mechanism underlying this dimorphism remain unresolved, the consistent presence of the additional ~ 1,012-bp amplicon in all females examined provides a readily distinguishable sex-associated PCR pattern. The sequence data further indicate that this pattern is associated with a 156-bp sequence difference within the targeted region. However, whether this sequence variation is directly involved in sex determination or sexual differentiation in *G. mopsus* remains to be elucidated through further genomic and functional studies.

Conclusion

This study identified a reproducible sex-associated PCR pattern in the endangered dung beetle *G. mopsus* and developed a molecular sexing approach based on the *tra2-region3*-targeting assay. To our knowledge, this represents the first PCR-based molecular sexing approach reported for a dung beetle species. Although the *doublesex* (*dsx*) locus exhibited high genomic sequence similarity between males and females, experimental evaluation demonstrated inconsistent amplification and non-specific cross-reactivity, indicating that it is unsuitable for reliable molecular sex determination in this species. In contrast, the *tra2-region3*-targeting assay produced a consistent sex-associated amplification pattern in all specimens examined that showed complete agreement with the phenotypic sex of all specimens examined, distinguishing males by a single amplicon (~ 1,168 bp) and females by a characteristic double-band profile (~ 1,012 bp and ~ 1,168 bp). Furthermore, TA cloning and bidirectional Sanger sequencing confirmed concordance between the experimental amplicons and the draft female and male WGS scaffolds, confirming that the observed band-size polymorphism corresponds to a 156-bp sequence difference associated with the female-specific amplification pattern. Although broader validation using additional individuals or geographically distinct populations would further strengthen the general applicability of the assay, the present study provides a sequence-validated proof-of-concept molecular resource with potential applications in conservation management, captive breeding, and restoration research for *G. mopsus*.

Declarations

Experimental procedures involving *G. mopsus* were conducted under a permit issued by the Ministry of Climate, Energy and Environment (MCEE) of the Republic of Korea (Permit No. EG202508ECP0001). Only specimens that were already deceased were used for molecular analyses, and no live individuals were sacrificed specifically for this study.

Competing interests:

The authors declare that they have no conflict of interest.

Funding

This study was supported by the National Institute of Ecology (Grant No. NIE-B-2026-48) through its core research project program, funded by the Ministry of Environment, Republic of Korea.

Author Contribution

Dong Eon Kim, Eunji Jeong, Kuem Hee Jang and Young Jun Jung conceived, designed, and supervised the study. Dong Eon Kim, Jongdae Park, and Eunji Jeong performed DNA extractions, PCR experiments, and gel electrophoresis analysis. Eunji Jeong and Young Jun Jung analyzed the molecular data and drafted the manuscript. All authors reviewed and approved the final manuscript.

Acknowledgements

We express our gratitude to the technical staff at the Endangered Species Research Center, National Institute of Ecology (NIE), for providing animal specimens and assisting with laboratory logistics.

Data availability

Sequence datasets and alignments are available from the corresponding authors upon reasonable request.

References

1. Gempe T, Beye M (2011) Function and evolution of sex determination mechanisms in insects. *Cell Mol Life Sci* 68:1205–1222
2. Geuverink E, Beukeboom LW (2014) Phylogenetic distribution and evolution of the *doublesex* gene in insects. *Insect Biochem Mol Biol* 55:1–12
3. Lim CS, Baek HM, Park SH, Purevdorj J, Hong MH, Han KD, Bae YJ (2015) Morphological characteristics of *Gymnopleurus mopsus* (Coleoptera: Scarabaeidae) with a new sex discrimination

- method. Entomol Res Bull 31(2):126–129
4. Minezawa F, Unagami M, Su ZH (2020) Molecular characterization and sex-specific expression of *doublesex* in stag beetles (Coleoptera: Lucanidae). Gene 730:144315
 5. Ministry of Environment (2018) Comprehensive plan for the conservation of endangered wildlife (2018–2027) [in Korean]. Ministry of Environment, Republic of Korea
 6. Mitchell LC, Moczek AP, Nadolski EM (2024) A conserved somatic sex determination cascade instructs trait-specific sexual dimorphism in horned dung beetles. Evol Dev 26:e70004
 7. Moczek AP, Rose DJ (2009) Differential expression of *doublesex* recruits ancient leg-patterning pathways to generate novel beetle weaponry. BMC Biol 7:21
 8. National Institute of Biological Resources (NIBR) (2023) Red Data Book of Republic of Korea Insects III. National Institute of Biological Resources: Incheon, Republic of Korea
 9. Nichols E, Spector S, Louzada J, Larsen T, Amezcuita S, Favila ME (2008) Ecological functions and ecosystem services provided by dung beetles. Biol Conserv 141:1461–1474
 10. Sánchez L (2008) Sex-determining genes in insects: regulation and evolution. Int J Dev Biol 52:837–855
 11. Schütt C, Nöthiger R (2000) Structure, function and evolution of sex-determining systems in Diptera insects. Development 127:667–677
 12. Shukla JN, Palli SR (2021) Sex determination mechanisms in Coleoptera: progress and prospects. Insect Mol Biol 30:245–256
 13. Simmons LW, Ridsdill-Smith TJ (2011) Ecology and Evolution of Dung Beetles. Wiley-Blackwell, Oxford
 14. Verhulst EC, van de Zande L, Beukeboom LW (2010) Insect sex determination: it all evolves around transformer. Curr Opin Genet Dev 20(4):376–383

Figures

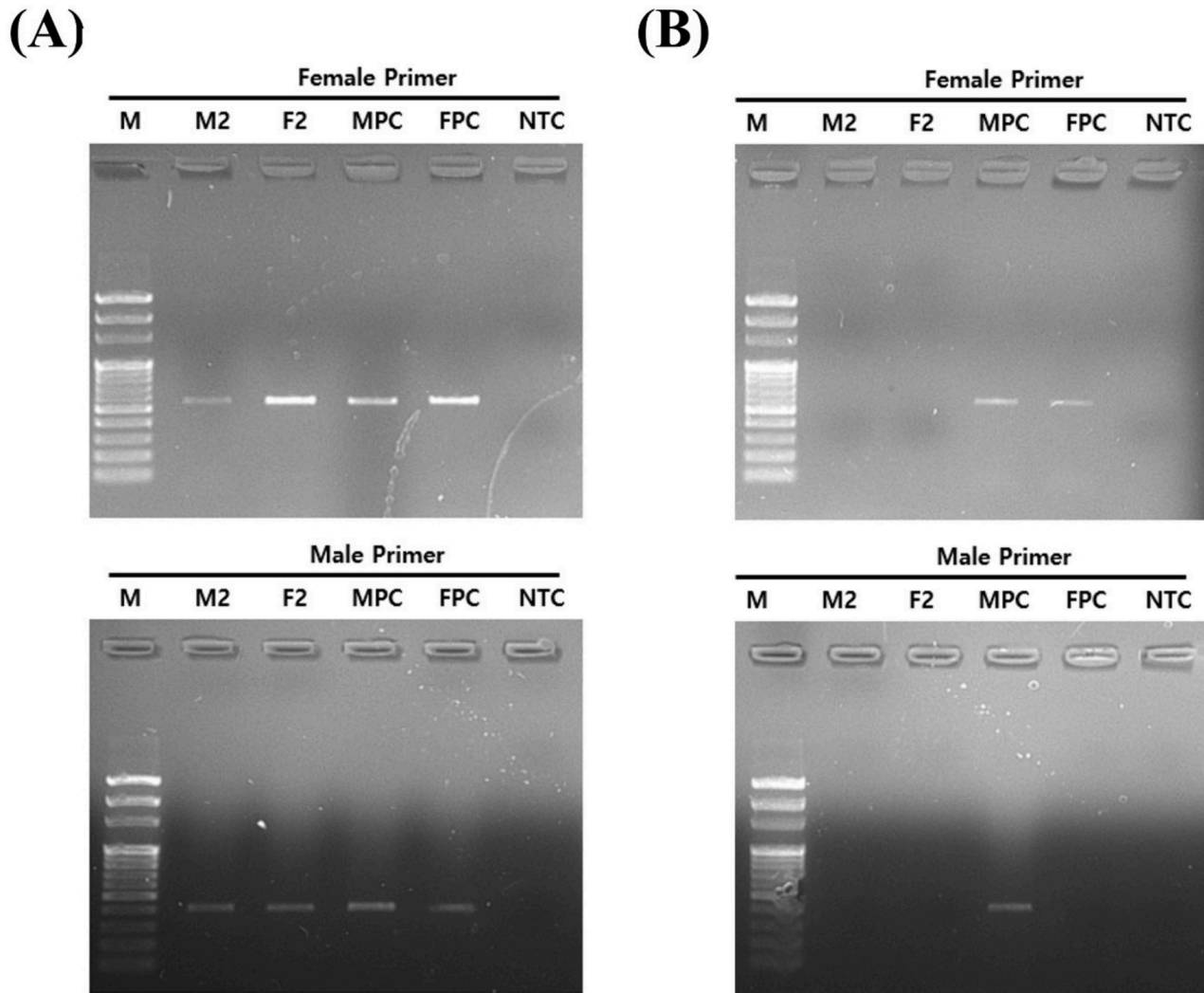
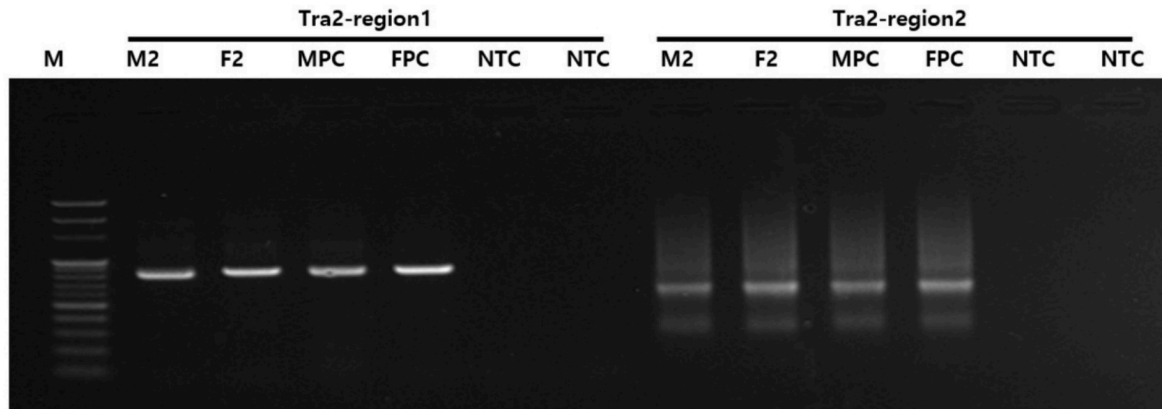


Figure 1

Empirical evaluation and marker invalidation of *doublesex* (*dsx*) gene primers for sex determination in the dung beetle (*G. mopsus*). PCR amplification results using female- and male-specific *dsx* primer sets across 4 samples (males: M2, MPC; females: F2, FPC) and NTC. Both the (A) primary condition (58 °C, 38 cycles; top) and (B) secondary condition (60 °C, 35 cycles; bottom) exhibited non-sex-concordant band patterns and non-specific cross-reactivity, indicating that the evaluated *dsx*-based primer sets did not produce a reproducible sex-associated amplification pattern under the conditions tested (Legend: M, 100 bp ladder; M2, independent male gDNA; F2, independent female gDNA; MPC, male positive control; FPC, female positive control); NTC, Non-template control).

(A)



(B)

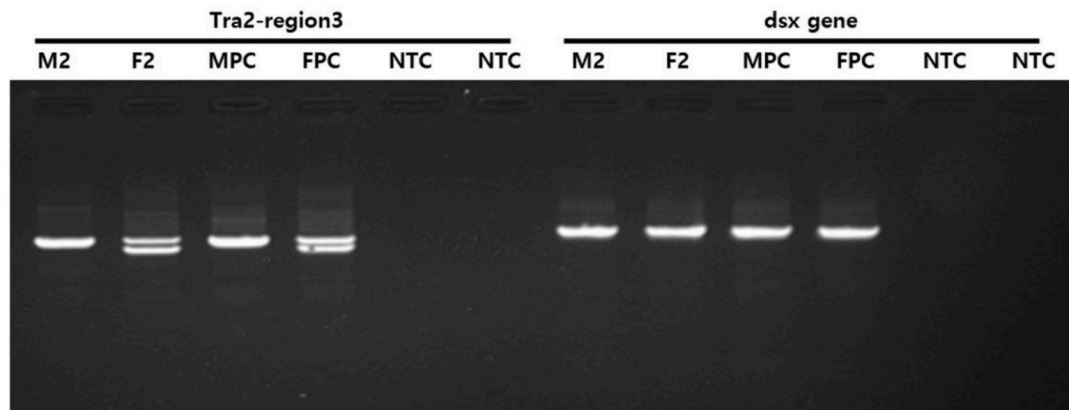


Figure 2

Screening and selection of candidate loci within the *transformer-2* (*tra2*) gene of the dung beetle (*G. mopsus*). Comparative evaluation of three target candidate regions (*tra2*-region1, *tra2*-region2, and *tra2*-region3) in representative male and female samples (MPC, FPC, M2, F2, NTC) under optimized conditions (58 °C, 35 cycles). (A) While *tra2*-region1 and region2 failed to display clear sex-specific diagnostic patterns, (B) *tra2*-region3 produced a distinct sex-associated amplification pattern, clearly distinguishing a single band in males (~1,168 bp) from a double band in females (~1,012 bp and ~1,168 bp). (Legend: M, 100 bp ladder; MPC, male positive control; FPC, female positive control; M2, male sample 2; F2, female sample 2; NTC, non-template control)

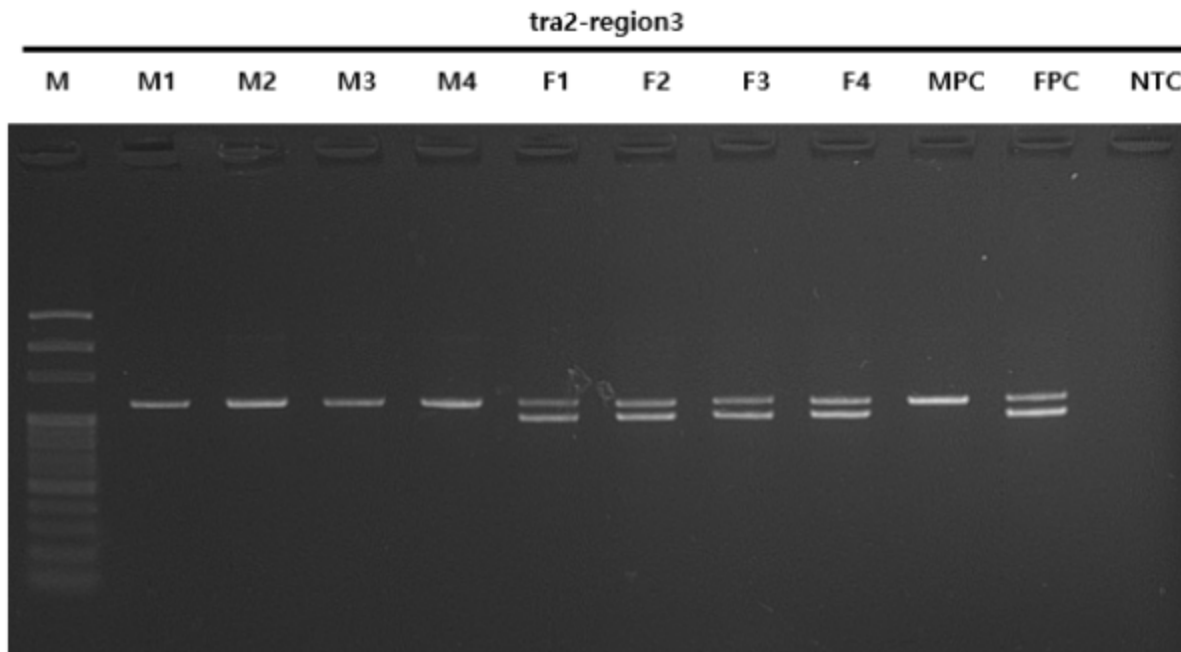


Figure 3

Experimental validation of the *tra2-region3* molecular marker for sex determination in *Gymnopleurus mopsus*.

PCR amplification profiles of validation specimens (M1–M4, F1–F4, MPC, and FPC) along with a non-template control (NTC) using the selected *tra2-region3* primer set under optimized conditions (58 °C annealing, 35 cycles). The assay demonstrates complete concordance with phenotypic sex: male individuals (M1–M4 and MPC) consistently exhibit a single specific diagnostic band of approximately 1,168 bp, whereas female individuals (F1–F4 and FPC) display a clear and reproducible double-band pattern consisting of approximately 1,012 bp and 1,168 bp fragments. No amplification is observed in the negative control (NTC). (Legend: M, 100 bp DNA ladder; M1–M4, male validation samples 1–4; F1–F4, female validation samples 1–4; MPC, male positive control; FPC, female positive control; NTC, non-template control)