

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Insecta Mundi

Center for Systematic Entomology, Gainesville,
Florida

10-14-2022

Genomics reveals a new genus and species from a single female specimen (Lepidoptera: HesperIIDae: HesperIinae: HesperIini: Moncina)

Ernst Brockmann

Jing Zhang

Qian Cong

Nick V. Grishin

Follow this and additional works at: <https://digitalcommons.unl.edu/insectamundi>



Part of the [Ecology and Evolutionary Biology Commons](#), and the [Entomology Commons](#)

This Article is brought to you for free and open access by the Center for Systematic Entomology, Gainesville, Florida at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Insecta Mundi by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

A journal of world insect systematics

INSECTA MUNDI

0957

Genomics reveals a new genus and species
from a single female specimen
(Lepidoptera: HesperIIDae: HesperIinae: HesperIini: Moncina)

Ernst Brockmann

Laubacher Str. 4, 35423 Lich, Hessen, Germany

Jing Zhang

Eugene McDermott Center for Human Growth and Development and Department of Biophysics
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-8816 USA

Qian Cong

Eugene McDermott Center for Human Growth and Development and Department of Biophysics
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-8816 USA

Nick V. Grishin

Howard Hughes Medical Institute and Departments of Biophysics and Biochemistry
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-9050 USA

Date of issue: October 14, 2022

Center for Systematic Entomology, Inc., Gainesville, FL

Brockmann E, Zhang J, Cong Q, Grishin NV. 2022. Genomics reveals a new genus and species from a single female specimen (Lepidoptera: Hesperiiidae: Hesperiiinae: Hesperiiini: Moncina). *Insecta Mundi* 0957: 1–8.

Published on October 14, 2022 by
Center for Systematic Entomology, Inc.
P.O. Box 141874
Gainesville, FL 32614-1874 USA
<http://centerforsystematicentomology.org/>

INSECTA MUNDI is a journal primarily devoted to insect systematics, but articles can be published on any non-marine arthropod. Topics considered for publication include systematics, taxonomy, nomenclature, checklists, faunal works, and natural history. *Insecta Mundi* will not consider works in the applied sciences (i.e. medical entomology, pest control research, etc.), and no longer publishes book reviews or editorials. *Insecta Mundi* publishes original research or discoveries in an inexpensive and timely manner, distributing them free via open access on the internet on the date of publication.

Insecta Mundi is referenced or abstracted by several sources, including the Zoological Record and CAB Abstracts. *Insecta Mundi* is published irregularly throughout the year, with completed manuscripts assigned an individual number. Manuscripts must be peer reviewed prior to submission, after which they are reviewed by the editorial board to ensure quality. One author of each submitted manuscript must be a current member of the Center for Systematic Entomology.

Guidelines and requirements for the preparation of manuscripts are available on the *Insecta Mundi* website at <http://centerforsystematicentomology.org/insectamundi/>

Chief Editor: David Plotkin, insectamundi@gmail.com

Assistant Editor: Paul E. Skelley, insectamundi@gmail.com

Layout Editor: Robert G. Forsyth

Editorial Board: Davide Dal Pos, Oliver Keller, M. J. Paulsen

Founding Editors: Ross H. Arnett, Jr., J. H. Frank, Virendra Gupta, John B. Heppner, Lionel A. Stange, Michael C. Thomas, Robert E. Woodruff

Review Editors: Listed on the *Insecta Mundi* webpage

Printed copies (ISSN 0749-6737) annually deposited in libraries

Florida Department of Agriculture and Consumer Services, Gainesville, FL, USA
The Natural History Museum, London, UK
National Museum of Natural History, Smithsonian Institution, Washington, DC, USA
Zoological Institute of Russian Academy of Sciences, Saint-Petersburg, Russia

Electronic copies (Online ISSN 1942-1354) in PDF format

Archived digitally by Portico
Florida Virtual Campus: <http://purl.fcla.edu/fcla/insectamundi>
University of Nebraska-Lincoln, Digital Commons: <http://digitalcommons.unl.edu/insectamundi/>
Goethe-Universität, Frankfurt am Main: <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:hebis:30:3-135240>

Copyright held by the author(s). This is an open access article distributed under the terms of the Creative Commons, Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited. <http://creativecommons.org/licenses/by-nc/3.0/>

Genomics reveals a new genus and species
from a single female specimen
(Lepidoptera: HesperIIDae: HesperIIDae: HesperIIDae: Moncina)

Ernst Brockmann

Laubacher Str. 4, 35423 Lich, Hessen, Germany
brockmann-ernst@arcor.de

Jing Zhang

Eugene McDermott Center for Human Growth and Development and Department of Biophysics
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-8816 USA
jingzhang.first@gmail.com

Qian Cong

Eugene McDermott Center for Human Growth and Development and Department of Biophysics
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-8816 USA
congqian1986@gmail.com

Nick V. Grishin

Howard Hughes Medical Institute and Departments of Biophysics and Biochemistry
University of Texas Southwestern Medical Center
5323 Harry Hines Blvd., Dallas, TX, 75390-9050 USA
grishin@chop.swmed.edu

Abstract. New taxa in HesperIIDae (Lepidoptera: Papilionoidea) are traditionally proposed after inspection of male genitalia, which largely form the basis for HesperIIDae taxonomy. However, with genomic DNA sequencing, even a single female specimen can be placed in a phylogenetic context of existing classification and taxonomically assigned with confidence. Genomic sequencing of an unusually patterned HesperIIDae female from San Martin, Peru, characterized by pearly spots outlining an inverted heart pattern on the rust-colored ventral hindwing, reveals that it represents an undescribed genus and species named here as *Gemmia buechei* Brockmann and Grishin, **new genus** and **new species**.

Key words. Skipper butterflies, South America, biodiversity, female genitalia, museomics.

ZooBank registration. <https://zoobank.org/2FA538FA-7D65-4097-9BBA-71CD1B2795E5>

Introduction

Neotropical regions are rich in undescribed species. While many of them may be cryptic and evade recognition by visual inspection (Hebert et al. 2004), others are distinctive and can be recognized as new at the first glance (Turland et al. 2012). However, it may be a challenge to place such distinctive species within a taxonomic hierarchy, in particular when only a single specimen is known. HesperIIDae taxonomy relies heavily on the analysis of male genitalia, and descriptions of genus-group taxa traditionally report the structure of male genitalia. However, genera are defined as monophyletic groups of species, and a confident phylogeny that includes all close relatives is a reliable way to define them (Cong et al. 2019; Li et al. 2019; Zhang et al. 2019, 2020, 2022). Here, we illustrate this approach and propose a new HesperIIDae genus based on a single distinctively patterned female specimen. We believe that bringing a new species and the new genus to the attention of researchers has advantages over waiting to find more of its specimens, in particular males. The genomics-based approach that we use puts the phylogenetic placement of this new taxon on a strong footing.

Materials and Methods

Specimens used in this study were from the following collections: American Museum of Natural History, New York, NY, USA (AMNH); Los Angeles County Museum of Natural History, Los Angeles, CA, USA (LACM); Museum für Naturkunde, Berlin, Germany (MFNB); National Museum of Natural History, Smithsonian Institution, Washington, DC, USA (USNM); and the private research collection of Ernst Brockmann (EBC). Standard entomological techniques were used for dissection (Robbins 1991), i.e., the distal part of the abdomen was broken off, soaked for 45 minutes in 10% KOH at 60°C, dissected, and subsequently stored in a small glycerol-filled vial on the pin under the specimen. Genitalia and wing venation terminology follows Steinhauser (1981). Length measurements are in metric units and were made from photographs of specimens taken next to a scale and magnified on a computer screen. Photographs of specimens were taken with a Panasonic Lumix DMC-FZ50 with Macro converter Panasonic DMW-LC55E, following Häuser et al. (2005); dissected genitalia were photographed in glycerol with the Nikon D200 camera without a lens and through a microscope at 2x and 4.5x magnification. Images were assembled and edited in Photoshop CS5. Genitalia photographs were taken in several focus slices and stacked in Photoshop to increase depth of field. Genomic DNA was extracted from a leg, and libraries were prepared and sequenced as previously reported (Li et al. 2019). Phylogenetic tree construction was performed following Zhang et al. (2022). In brief, the maximum-likelihood tree was constructed from protein-coding regions in the Z chromosome using IQ-TREE v1.6.12 under a GTR+GAMMA model (Nguyen et al. 2015). To estimate the confidence of each node, we generated 100 replicates of 10,000 codons randomly sampled from the total set of codons in the Z chromosome and constructed maximum-likelihood trees for each replicate. The support values of each node were summarized from these replicate trees using the sumtrees routine in the DendroPy package (Sukumaran and Holder 2010). The method to find diagnostic DNA characters is described in Cong et al. (2019) and we used the *Lerema (Lerema) accius* (J. E. Smith, 1797) genome assembly as a reference (Cong et al. 2015).

Results

An unusually patterned female specimen of Hesperidae was noticed among the material given to the first author by Michael Büche (Fig. 1a, b). In wing pattern, it did not match any known species and was believed to be undescribed. To place this probably new species among Hesperidae, we obtained its whole genome and constructed genomic phylogenetic trees of representative species. According to our phylogeny, the species is confidently placed in the *Phlebodes* Hübner, [1819] subgroup of related genera (marked by blue branches in Fig. 1d) within subtribe Moncina A. Warren, 2008. The new species is sister to a clade of *Phlebodes* and *Dubia* Grishin, 2019, not closely allied to any existing genera, and therefore belongs to a new genus, described here, together with this species.

***Gemmia buechei* Brockmann and Grishin, new genus and new species**

<https://zoobank.org/89315B0A-D839-453A-B9A7-81439D67C0F7>

<https://zoobank.org/66B847C1-C51D-450B-939B-0938D7AE7A5B>

Diagnosis of the new genus. Differs from close relatives, such as *Phlebodes* and *Dubia*, by ventral hindwing pale spots in cells M_1 - M_2 and M_2 - M_3 being offset towards wing base, not aligned with other spots in postdiscal row; in female genitalia, ductus bursae with bursa copulatrix longer, reaching thorax, sterigma longer, and more expanded in its anterior portion, with rounder sides. Due to the lack of males and monotypic composition of the genus, it is best diagnosed by a combination of the following DNA characters: lac704.5.28:T96C, lac7147.4.1:C46A, lac884.3.3:A590T, lac221.9.2:A233A (not G), lac886.16.8:C464C (not A), lac49.46.10:G109G (not A), lac357.30.3:G1355G (not A), lac1895.5.35:T37T (not A), lac1416.2.1:A449A (not G), lac580.72.1:T75T (not C), lac3441.2.1:A436A (not G), and lac886.16.8:T1215T (not C). See Appendix 1 for abbreviations and sequences.

Description of the new genus. We hypothesize that the following characters may be shared among species of the genus. Medium-sized, forewing length around 20 mm. Body brown, abdominal sternites with a median dark line, paler on either side of this line, especially in the distal half where they are overscaled with cream-colored

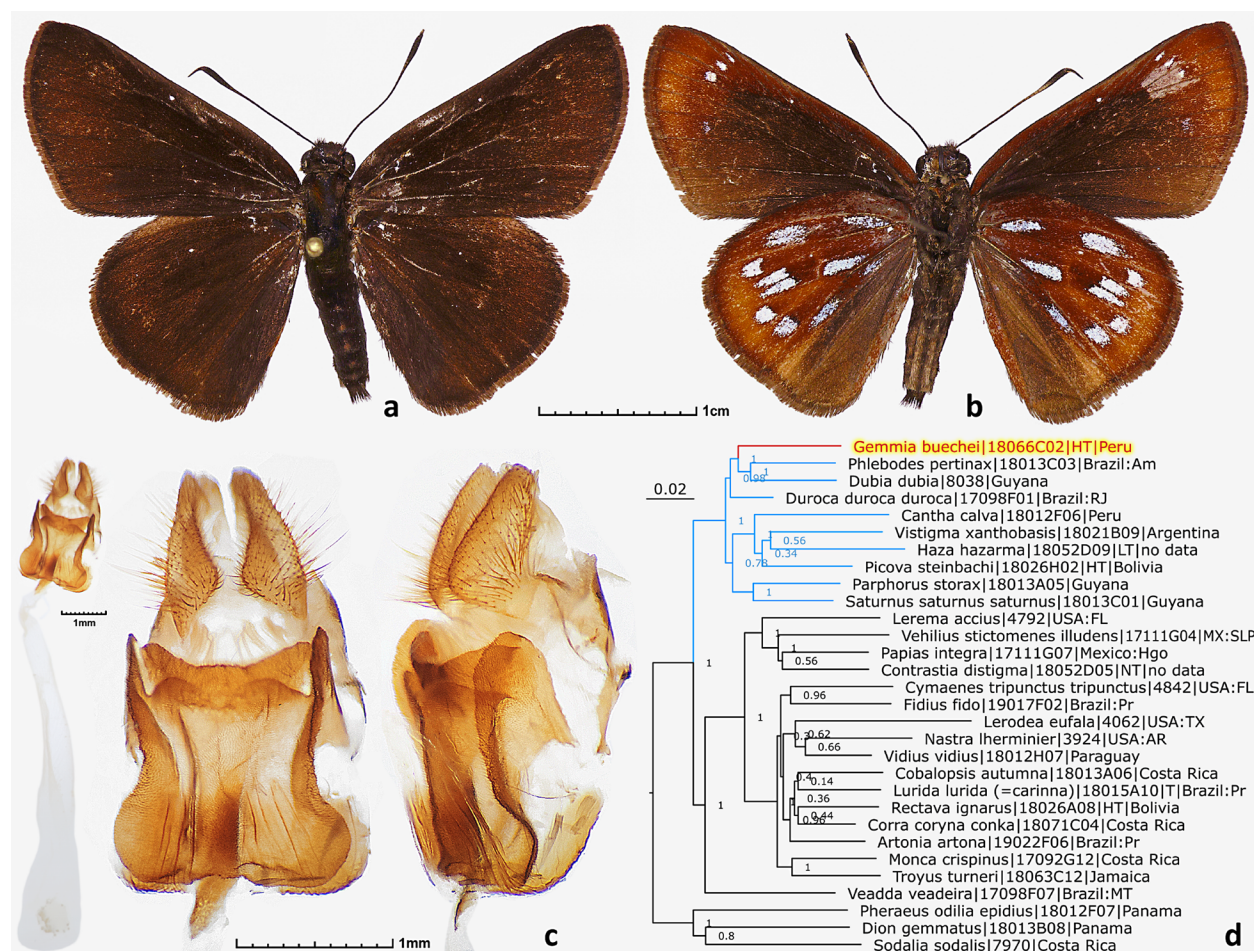


Figure 1. The holotype and phylogenetic placement of *Gemmia buechei* Brockmann and Grishin, gen. n., sp. n. **a)** Dorsal and **b)** Ventral views of the holotype, female, NVG-18066C02, from Peru: San Martin, additional data are in the text. **c)** Genitalia in ventral view, enlarged sterigma in ventral and right ventrolateral views: left, middle and right, respectively. **d)** Phylogenetic tree constructed from protein-coding regions of the Z chromosome: statistical support values (bootstrap fractions) are shown by the nodes, numbers after species names refer to DNA samples. Scale bars are shown by the images. For the tree, the unit is the estimated number of base pair substitutions per position (i.e., the scale bar corresponds to 2 substitutions per 100 positions). We sequenced a syntype of *Megistias carinna* Schaus, 1902, a name currently considered a junior subjective synonym of *Lurida lurida* (Herrich-Schäffer, 1869), thus the notation *Lurida lurida* (=carinna).

scales. Antennae about half of forewing length, dark-brown, ventrally pale near the club, nudum brown. Wings brown, could be with pale spots, ventral side brighter colored, with rusty, orange, and yellow scales and white or pearly spots. In female genitalia (Fig. 1c), ductus bursae with bursa copulatrix very long, spanning the entire abdomen up to thorax, no signa, ductus gradually thickening towards bursa; ductus bursae ventrally sclerotized by sterigma, ductus seminalis connecting near sterigma, around the sclerotized area; sterigma expanded, nearly rectangular with concave sides in ventral view and on each side with a lateral ridge rounded caudad and protruding past lamella antevaginalis; lamella postvaginalis notched in the middle, on the sides connected through the expanded ridge to mostly convex, straight in the middle lamella antevaginalis.

Type species. *Gemmia buechei* Brockmann and Grishin, **new species**.

Species included in the genus. Only the type species.

Parent taxon for the genus. Subtribe Moncina A. Warren, 2008.

Diagnosis of the new species. This mostly brown species is identified by nine or ten pearly spots on the ventral hindwing outlining an inverted heart pattern on a rusty-colored background.

Description of the new species. Female: (n=1, Fig. 1a,b) right forewing length (wing base to apex) 19 mm in holotype. Palpi missing in the holotype. Fringes brown. Dorsally, wings dark chocolate brown, forewing with traces of pale subapical spots, costal area of hindwing paler. Ventrally, rust-brown to chestnut colored, with darker, nearly black-brown discal forewing area and paler, yellowish submarginal areas of wing and planer brown areas towards inner margins of both wings (except hindwing cell 3A rust-brown); forewing with 3 subapical pearly spots in a straight line and thickening from costa towards outer margin from a dash to nearly round, and a faint spot (patch of pearly scales) in the middle of cell M_3 -CuA₁; hindwing with ten large, mostly elongated pearly spots arranged to outline an inverted heart pattern: triplet of spots towards base, weakest (nearly missing on left hindwing) in cell CuA₂-1A+2A, and heptad in postdiscal area arranged to form “3” on left wing, pattern alternatively described as two spot (one near base and one in outer half) in each of cells Sc+R₁-RS and CuA₂-1A+2A, and one spot in each of cells: discal (in distal, posterior area), RS-M₁, M₁-M₂, M₂-M₃, M₃-CuA₁, and CuA₁-CuA₂. For genitalia description (Fig. 1c) see description of the genus above, because we hypothesize that they would be mostly similar for all species in this genus. Male: unknown or unrecognized.

Barcode. The COI barcode sequence of the holotype (sample NVG-18066C02, GenBank accession ON256154, 658 base pairs) is:

```
AAC TTTATATTTTATTTTGG AATTGAGCAGGAATATTAGGAACATCCTTAAGTTTATTAATTCGTACAGAATTAGGAAATCCAGG
TTCTTTAATTGGCGATGATCAAATTTATAATACTATCGTAACAGCTCATGCATTTATTATAATTTTTTTTATAGTAATACCTATTA
TAATTGGGGGATTGGTAATTGATTAGTTCCATTAATATTAGGAGCCCTGATATAGCTTTCCCCCGAATAAATAACATAAGATTT
TGAATATTACCCCATCCTTAATATTATTAATCTCAAGAAGAATTGTAGAAAATGGTGCAGGAACAGGATGAAGTGTATATCCCCC
TCTTTCCTCTAATATTGCTCATCAAGGAGCATCTGTTGACTTAGCAATTTTTTCCTTACATTTAGCTGGAATTTCTTCTATTTTAG
GAGCTATTAAC TTTTATTACTACAATTATCAATATACGAATTAGAAATTTAGCATTCGATCAAATACCTTTATTGTCTGATCAGTAG
GTATTACTGCATTATTATTACTTTTATCTTTACCTGTATTAGCTGGAGCTATTACTATATTATTAAGTATCGAAATTTAAATACA
TCATTTTTTGATCTGCTGGAGGAGGAGATCCTATTTTATATCAACATTTATTT
```

Types. Holotype ♀ (Fig. 1), presently in the research collection of Ernst Brockmann (Lich, Germany), to be deposited in the collection of the Museo de Historia Natural, Universidad Mayor de San Marcos, Lima, Peru, bears the following three rectangular labels: white printed and hand-printed [Peru IV.2012 | Dept. San Martin | 1500-1800 m | ex coll. Michael Büche], white printed [DNA sample ID: | NVG-18066C02 | c/o Nick V. Grishin], red printed [HOLOTYPE ♀ | *Gemmia buechei* | Brockmann & Grishin].

Type locality. Peru: San Martin, elevation 1500–1800 m.

Etymology. The genus name *Gemmia* is a feminine noun in the nominative singular, given for the gem-like pearly spots on the ventral hindwing of the type species. The specific epithet *buechei* is named to honor its collector, Michael Büche, who generously gave many interesting specimens to the first author. The species epithet is a noun in the genitive case.

Acknowledgments

We acknowledge Jinhui Shen, Leina Song, and Ming Tang for excellent technical assistance and the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources. We are grateful to David Grimaldi and Courtney Richenbacher (AMNH: American Museum of Natural History, New York, NY, USA), Weiping Xie (LACM: Los Angeles County Museum of Natural History, Los Angeles, CA, USA), Wolfram Mey and Viola Richter (MFNB: Museum für Naturkunde, Berlin, Germany), and Robert K. Robbins, John M. Burns, and Brian Harris (USNM: National Museum of Natural History, Smithsonian Institution, Washington, DC, USA) for facilitating access to the collections under their care and stimulating discussions. We thank Crispin S. Guppy and Bernard Hermier for critical reviews of the manuscript and helpful suggestions. This study was supported in part by the grants (to N.V.G.) from the National Institutes of Health GM127390 and the Welch Foundation I-1505.

Literature Cited

- Cong Q, Borek D, Otwinowski Z, Grishin NV. 2015. Skipper genome sheds light on unique phenotypic traits and phylogeny. *BMC Genomics* 16: 639.
- Cong Q, Zhang J, Shen J, Grishin NV. 2019. Fifty new genera of Hesperidae (Lepidoptera). *Insecta Mundi* 0731: 1–56.
- Häuser CL, Steiner A, Holstein J, Scoble MJ (eds.). 2005. Digital imaging of biological type specimens. A manual of best practice results from a study of the European Network for Biodiversity Information. European Network for Biodiversity Information; Stuttgart. viii + 309 p.
- Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14812–14817.
- Li W, Cong Q, Shen J, Zhang J, Hallwachs W, Janzen DH, Grishin NV. 2019. Genomes of skipper butterflies reveal extensive convergence of wing patterns. *Proceedings of the National Academy of Sciences of the United States of America* 116(13): 6232–6237.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1): 268–274.
- Robbins RK. 1991. Evolution, comparative morphology, and identification of the Eumaeine butterfly genus *Rekoa* Kaye (Lycaenidae: Theclinae). *Smithsonian Contributions to Zoology* 498: 1–64.
- Steinhauser SR. 1981. Revision of the *proteus* group of the genus *Urbanus* Hubner (Lepidoptera: Hesperidae). *Bulletin of the Allyn Museum* 62: 1–41.
- Sukumaran J, Holder MT. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26(12): 1569–1571.
- Turland VA, Warren AD, Delan L. 2012. A new genus and species of Moncini from Jamaica, West Indies (Lepidoptera, Hesperidae, Hesperinae). *Tropical Lepidoptera Research* 22(2): 66–73.
- Zhang J, Cong Q, Shen J, Brockmann E, Grishin NV. 2019. Genomes reveal drastic and recurrent phenotypic divergence in firetip skipper butterflies (Hesperidae: Pyrrhopyginae). *Proceedings of the Royal Society B: Biological Sciences* 286(1903): 20190609.
- Zhang J, Cong Q, Shen J, Grishin NV. 2022. Taxonomic changes suggested by the genomic analysis of Hesperidae (Lepidoptera). *Insecta Mundi* 921: 1–135.
- Zhang J, Cong Q, Shen J, Opler PA, Grishin NV. 2020. Genomic evidence suggests further changes of butterfly names. *The Taxonomic Report of the International Lepidoptera Survey* 8(7): 1–40.

Received April 26, 2022; accepted September 24, 2022.

Review editor David Plotkin.

Appendix 1

DNA sequences of *Lerema accius* (J. E. Smith, 1797) exons with diagnostic characters of *Gemmia* Brockmann and Grishin, **new genus**. Sequences of *Lerema accius* nuclear exons (NCBI genomic assembly accession GCA_001278395.1) (Cong et al. 2015) with diagnostic characters for *Gemmia* are provided below. The position used as a character state is shown in bold italic font. Base pair in this position is the one present in the *L. accius* reference genome, and may not correspond to the actual base pair in *Gemmia*. The characters are given as abbreviations in the lines starting from the > symbol, e.g., lac704.5.28:T96C means position 96 in exon 28 of gene 5 from scaffold 704 of the *L. accius* reference genome is C, changed from T in the ancestor; (not G) means that this position is not G in *Gemmia*. If the function of a corresponding protein is known or can be predicted, it is stated after the | symbol.

```
>lac704.5.28:T96C | Predicted: Down syndrome cell adhesion molecule-like protein Dscam2
TGTTACTTCGAATTTTCTTTCTGTTTATTTAACCGTGGTTCTTTTACGTACCGATCTTTTATTCCAGTCCCCCAAACATTCTGCC
ATTTTCGTTTGGCGAAAAACCAGCAAATGTTGGCGAATATTTACAAGCATCGTGCACCGTTAATCTCGGTGATCTTCCGCTTACGA
TCGCTTGGACGTTCAATGGGCACCTTGATATCACAACGAAACAGTGACTATTCTATGAGCAACTCTAAAAGAAGTAGCCTCCTTATA
ATCGAATCGGTTCGACGCTAAACACGCTGGTTCCTATACTTGCGTCGGTGAAAATCGCGCGGGAAAATCGACATACTCCTCGGATTT
AGTT
>lac7147.4.1:C46A
ATGACAACGAAAGTAAAGAAAATCGTTGGAAATGATGCATTCCAACGAATTAACCTCCTATATCAG
>lac884.3.3:A590T | uncharacterized protein LOC101737811 isoform X1
ACAATCCAACGAATGTCTTCCACGAACGTTGTTATTAGGATATGGTGTACGTCCTCACTTCACCGTGTAAGGATGGCGAGGACAA
GTGTGATAGGGTCCAAGACGAAATTATTCAGACTTGGGGAGTGTATTTTAGCGGTTTGAGGTACATTGGGGCCAACCTCGCGTTAA
ACTTCACGTCAGATTGCTTCAAGAGCAACACGCTGGTTTTCCAAATGCACGGCGGATACTTCTGCTCGTACAAAAGCCTTAAATT
AGACATTATACCACCAGAAAATGAAATAGAACATGGATCTCTGTCTTCGGACTCTTCCGGAAAAACTCACGTGTCAAAAATAACA
TTAGAGTCTCGCGTTTTTGCAAAACTACAGACATTTCAAAGAAGGAAGGTCTATATCACCTAGTAAAATTTCAAGGAAGTCGGATA
AGGAGTATTCTAAAAGCCATAGTCCAGTGGAGAAAAGGCTTTAAACAGAAGTTCCTCAAGTTTAAAGCTTCTGCGAGTTTGAACAT
GGCATTGAATCATAGGGAAGGCGTAAGGCAACGAGACGATAGTAATGAGCGAAAGAAATTGACCCGAATCACATATACGAGCAT
TCACCCGACAATTTAGACGTTAAAGAATTCACGACCAGTGATGAAAACGTGTGATATATCAAAAATCGAGTCACAAGAGGAATTATC
CGCCAGCAAAAATAGACTTAAGTGAAACCGCAGAGAGTTACGATTTGAGAAACGAATCTGATGGAATGCCTAAATATCGGTACTTGG
CGTTGAGTTTTCTTAAATCGGAAATCCGGCCAAGTTACAACGCGTCCAGCTGCAGAGCTTGCAATTTGCTTCAATACTCAATAAAG
AAGAAGCAAGAAGCAGTCCAAGAGATTAAAGAAAGAATCTATCAGAAGTCTGCTTTAGCGAATGAGTCTTGGAAGTTTACCGGAGAA
TGCTTTCGAAATCAACACAAGTCCAGGAGTAGATCTCAGAGGAATCTTTTCTCGCCGGAAGAGAATGAAAGGTTTCAGAAGGAAC
GGTCGTCGTCTCAGAGTGATTGAGCAAGAATGGAAGAATCAAAGATGACAAAACACCCATTTCGAATGGTCCAAGACTCACTTTA
AGTTGAACGATCTGTGTCTTAC
>lac221.9.2:A233A (not G) | protein rhomboid
ATGGAGCAATTAACGAAGAAGATGGATCGCCGGTGGCGGTGCATCAGGGGAGAAGAATCCTCCCGGACGTGGTGCCGCCGCCCG
GGCCTGGGTGGAGCACGAGGAGTCGGAGCGACTGACCGATAGCAGGACGAACAGTCCCGGGGAGAACAGCGGCTGCTCAGCCCCG
ACGTCCACAAACAGACTCGTCTCCAAGCTGGACTCACTTTAGGAGTCAAGGATATAAAGAAAACATACGAAATGTCTCAAATTGCG
AAAAAAGTGAGACCTACAGTGAAATCGAAGTTCCAGAAAAGGAAGGAGAAACTTATCGAAATTCTAAAGCCGCCTTACTTCATCAT
CACCATGATAATAATTATA
>lac886.16.8:C464C (not A) | uncharacterized protein LOC101745102
GTTGATTCTTTTACCGTGAACCAAAAACCTAAGTGGTAAATTCGTGCTGCATGAAAGTATACGACGTGTCTCAACATAGATCAACTCA
TACAAGCCAACATTCTGAACCAATGAGTACATCCTCTACTGAAAACCAAACCTTCTTCAGACAAAAGCCACCCAAAGTCATTTAATT
TGGTAAACTATTTATGAAACAGAAAAGTATTAGTAATGACGGAATTCATGGTATGGATCAATTAGATAGATCGGAATGTTGGCC
TTCTAGTTACAGGTGGTGAGAGTGGTGAATCAATGGGAGAGCAAAAACCTGGGGACTCAAAATCTATATCAACAGATAGACCCCAA
GCCGAATTACCATTAGATGCTGATTATAATAGCTCGGCTGTCTATGATGAAGTCCGATCAGTTAATCCATTTAATAGAGTATACG
ACGAAGTTATTGAAGAAGAGACAATCTTGAAACGACCAAATTAAGCGGCAGCGAAACGAATTTATATGCTGTTGTTAACAAACC
GCACATTAAGAGCCAAATATAAATATAAATAACAGCCCATGTAAGTCTAGGAGTAGGAAATCTTGTTCTTCAACATCATCGGCT
ACAAGTGTTAGCATTTCAAGTTGTTCTGAATCAGATGGCACACAAATTACAAGAATGAACAAAGTGCTGTTAAAGAGTCTGTGA
TCATAAATCTACATCTACACACTTAGAGGCTGATACGTTAGATAAAAGTATGCAGACGTCAAATATCCAAACATCAACAATTTCTC
ATGAACGCGAAGTCTTAAATAGTGGAGCCGTCATTCTTAGAAAAGTTAAAGAGGGCGACTGTGAGAAACCCGTTTTTGTACTC
TATCCTAATTATACACTGCCAGACATAAGTTTCTTAAATGGTAGGCCAAATATTTACTTGAATCCGATAAAAGTAAACATTTCCCC
TAAACCTAATGAAAGTAAAGAAGACACGTATGCAAGTTAAAGGGAAAAGGCCTTTTTTCATGCAATGATGTAGAAATTTTAAAGAAAA
AAGGTTTAGGCCATATCAAAGACTGGGACTCTCTTAACCTCTTGTTGCCTTTGGAGTGCAAGCAAATGTTGTGAGAATTACCAGA
GCTAACCCAAATAAAAGAAAAAGAGGGTGATCTAAATGCAATGACAAATACTGTAACCTCATCACCTTCGTGTAAGTCTAAGAAT
AGACCAATGAGTTGTGATTGCAATAATTTGATTGGAATAACACGGCTGTGTCTTCGAGCTCCAGTACAGCGACACAACCTTCATC
GGGATATCGAGGATCCTCTACAATGTTGACGGACTCTCTGCACAAAATAGCCCCGCCCTGCTGGAATTTCAATCCTCTTTTGG
```

TTTATCGTTATGATAGTGCTACAAGCTCCGAGGCGAGTTGCGTCAATAATGACGGACTACGAATTATTCCTACCATACCGAAGCGA
TCCTTATCGTTGGCCGATCAAACCTCGCATGGCCAAACAAGGCGAGTTGGCACCTCCGAGACCTCCATTACCAAAGAGTATCTTACG
TAAATCCATGGACAAAACCTCGCAAGTCCACTAGTCATACGAAACGTTACAGCATGTTTGAAATGGATGAATTTATTCAAGATCCAA
TTGTGTGCGCTACGGCTGCCGTAGAGCATAAACTAAACGAAGGTCTTTACAAGAACCTTATTATTACAAAATCAAATATGGA
TTATAGAAAAGAACAATGATCTAGCTGCTAAACGACTATCTCAGCAATTCCTAGATGCCGCCGAGAAAGACGCTGATTACAATGAA
TATTATCAAGATGAAGGTGTGGGAACAGAAAGCAGCCTTGAATCAGGAAAATCAAATGAAATAAAATATCATAGACCACACACACC
TCCATTACCGAAACCAAGAATAAAAAATGGAGTACATTGATTTTCCACCTCCGGGAGCACTTATTAGTAGCGCTGATTTACAAC
AACTAGAAGAATTCTTGAAACAAAGTGGATTCAACTGCCAAAATATGGATGAATGGGATCAAAATCAAGTA
>lac49.46.10:G109G (not A) | importin subunit alpha-7
GGTCTCGACAAAATAGAATTCCTGCAGTCGCATGAGAACCTTGATATATATCAGAAATCATTTGAAATAATTGAGAATTACTTTGG
ATCGGAAGGAGAGGATAATAGA**G**TAGCGCCCGAAGTGCTTCCACAACCTGATCAGTACCAATTTAACACCGACCAATCAGTGCCTA
TGGGTGGTTACCAAGTTC
>lac357.30.3:G1355G (not A) | A-kinase anchor protein 10, mitochondrial isoform X1
CAACCGAACATATTTGAAAAGAAATCTAAATTGTCTCTAACGTTACATGACATACTTGTGACAAAAAGTCTATTAAGTATTTTAT
CTTATTTATGGAACAATTTGGTAAGCAATTGTTAATAAGATGTTGGATGGAAGTACAAATTTTAAGATGCAAGTTTTATTTAATA
GTAAGTTAAATGAAGATGATAAGCGTCTACAGTTAAGTCGTACCAAAAGCTTTGATGTTTCAGTGCCTACTGATACCGAAAAGAA
AGATGAAGATTCAATGTCACAAAGGTTCAACTGGAAATCCAGTACAAGTAATCTGAAGCAAACCAAGTCTGAGTGCTCAACCAGT
TCATGTGGGAACCTTAGCAGAGCATATTAGACTTGTAGAAAACCTTACACAAGATCAGAAATGTTTAATGTAACAGAAGAAGCCA
TTAATCTGTTCAAGAAGTATGTAGCATTAGAAGCTCCCAATAAATTAATATTACCGATGAGCTGAGGAAGTCAGTGATAAGTAA
TATATGTCAACCAGAAGGTCAAATAAAAGAGACTGCTTTGAACCGGTGCAACAGGCATTATTTTACCACAAATGGAACCCACTTT
ATGGAGTTTCGAAATAGTCCCTTATATGTAAAAGCACAAAATAGATATTTTAACTATGGGCTCAATTGAACCTCCAGATATATTGTA
TAATGAAAATATATTATCTATTTTACAGAATACTTAGAGCAAGAATCTTGAGTATCTTATTAGAATTTTAAATAGCCATAATGC
ATTTTAGGGACAATATTACTGATAATACAGACTCCAGTCCAGAACAGAAGCAAGGCGATGCTATTATCTTGTATGATAAATATTTT
TCGCTTCAAGCAACCAATCCTGTGGGATTTCTACTGAAATAAGGTTGAAAGTGGAGAGTGATATTTGTGCAAGGCGGGCCTCT
CCCTACAAGCTTTTGACATGCCATACAAAGATTATATTTTAATACATTATCAAAGTATGTTAAACATTTTGAATCAGAGTTATATT
ATTCTTACCTATCTGAAATGGTGCATTCCGTGGACAGATTATGGTCTTCCAAATTCAAATCACACTCAGATTGTAGTAGTGAATT
TAGTATAAGTACTCAGAATACTTTACTAGCAATGGGAGATCCTCTCTTTAAAAAGAAACGGAACCTATTCTGTACCTGATATGTCA
ATAGATTCAAACAGCTTTATAATGCAGATGCATTATGGCAAAGAAAAAACATGACGGGCTAAGTCTAGGAAGAATAAATAGTTT
GGGTAGATTTCGAGTCAAATTTGAGCCTGATCCAGATAAAAAAGATAGGTGAGTCTTAAAAAAATGGTTA**G**CAAACCTTGTGCCAA
CAAATACATCCAAAGTTGAAGAAGAGATGGCATGGCAGATAGCACATATGATTGTTAAAGATGTAAGTACTGACTTAACCATGGCACCT
CCTGATAATGAAAGTGATGATGTA
>lac1895.5.35:T37T (not A) | Down syndrome cell adhesion molecule-like protein Dscam2
AACTATCCGTCGTTCAACGATCACATTCATCCGCCG**T**CGCGATTCTCCGACGACACGGAGTGCGAGGGGTCACCA
TCGGCGGAAACGCGAGCGGTACGCGGGCCACGCGGCTCACTTGGCCCACTCGGCCCACTCCGAGCCTAAAGTG
ACTCGGAAGCCTTCCAAATCAACGTC
>lac1416.2.1:A449A (not G) | mRNA cap guanine-N7 methyltransferase
ATGAGTTGTTTAGAAAGTAATTACGAGGCAAGTGGTAACGAAAAAGCCTTAGACCTCGATAGGGATAAAGTCAATTTGTTAAAAAG
GAAACATTCTGACTGTGATACTGGTGAAGGTGATGCTTCTGTATCAGAAGTTCCAAGTAAGGTGGAGCGCAGCGATGAACACTCAA
GCGTCGTCGCTGCACACTATAATCACCTCGAAGAAAAGGGTTTAAAGAAAGATATAAATCACCTATATTTTATCTTAGAACTT
CAACAACCTGGGTAAAGTGTGCTTATACAGGATCACATAGATAAGGTGAGTGAAAAAGACTATGGAAAACCGTTGAGAGTCCCTA
GACATATGTTGCGGGAAAGGAGGCGATTTAAGTAAGTGGCAAAAGCAGTAGGAATGAAGCATGTTGTATTGCTGATATCGCTG
AAGTGTCCGTACAGCAATGTG**A**AGCACGTTACATGGACTTACGTCGACGATGTGGCCGGCTGTTACAGGCTGAATTCATAGCGGC
TGATTGCACTAAGGATACTCTTAGAGATAAGTATCAAGACCCATCAATCAATTTTGACTTTGTGAGCTGCCAGTTTGGATTGCAT
TACAGTTTTGAAAGCTATCTCAAGCTAGAAGGATGTTGACTAACATATCTGAGTGTTTGAAACCAATGGCTTTTTCTTTGGAAC
AATACCTAATGCTGATGAAATTGTTTCAAGAGCTAAAAATCACCAGACAGAAGTTTTGGAAATAGAGTATTTAATATTAAGCTTT
TGTTTGATCCAGCAGATGGATATCCACTGTTTGGTGCAAAATATGATTTTACCTTGAAGGTGTAGTCGATTGCCCTGAGTTTTTA
GTGGATTTTAAAGTACTTGTTAAGCTCGCTGCAGATATGGATTAGAAGTATTTATAAAGCAGGATTTGCTGATTTTTTTAAGGA
CTATTACAGAGAAATATGAAAAGTTATTACATAGAATAATGTGTTTTGAAAGTTATCCTGCTCCACAGGGAAACATTTGATAGCAG
AGGACTCTGAATATGAACATGCTAAGCAGAGATGGGAGGAGATAAGTAATAAGACTGAACATCAACGCTTAGGTACTATGAGCAAG
AGTGAATGGGAAGTTGCA
>lac580.72.1:T75T (not C) | glycylpeptide N-tetradecanoyltransferase 2
ATGGAAGATAATAAGGCAAACCAACCCGGCGGATATCAAATACACATGAAGATAAACAGCCAAAGAGGAAAA**T**AAAAACAAGAA
AAAGCGAAGTGGCGGGGACGGGGACATTGAGGGTTCGGGGTCAGACCTAAAAACGGACATTGTAACCATTCTACTGGTGATATAC
CACAAATCTATCTTTAAAGACCTTAAATGGCAATGGAAGTACTAAATTTACAACAAAAGCCTGCTAAACAACTGAGGAAGC
TTTGCAATAAATCTTATCAATTTTGGTCCACTCAACCCGTACCAACAATGGATGAGAAAAATAGTTTCGAATGAACCTATTGAACCA
CCTAAATCACCCGAAGAAATACGATCTGAACCTTATACATTGCCAGATGGGTTTCAATGGGATACTTTGAATCTTAATGAGCCTT
TAGTATTAAAGGAACCTTATTCTTTGCTTAATGAAAATTATGTGGAGGATGATGATTGCATGTTTCGTTTTGACTATCAAACCTGA
TTTTTTGAAGTGGGCGTTGCAGCCTCCTGGATGGCAAATGGAGTGGCACTGTGGAGTGAGAGTAGTAAATCTGGTAGGTTAGTT
GGATTTATATCAGCAATTCAGCTGACTTAAGAATATATGACCATACACAGACTGTCTAGAGATTAACCTTTTATGTGTTTCATAA

AAAGTTACGTGCTAAAAAGAGTTGCTCCAGTTCTTATTAGGGAAATAACACGCAGAGTTAATTTAACTGGCATTTTCCAAGGTGTAT
 ATACTGCTGGTATTGTACTACCCAAACCCATTGCAACTTGTAGGTATTGGCATAGGTCACCTCAATCCTAAAAAGTTAATTGACATC
 AAATTTAGCCACCTATCCAGGAATATGACTATGCAGAGAACCCTTAAATTATTTAACTACCGGATTACCTCGTACTCCTGGTTA
 CCGTAAAAATGGAACTCAAGATTGTGAAAAAGTTGTAAAGTTACTCAATGAGTACTTGAAAAAGTTTGACTTGGCTCCTATTTTTT
 CTGAGGAAGATTTCAAACATTGGTTTATACCACAAATAGGTATAATTGATAGCTATGTTGTTGAATCATCTGATGGCTCCATCAC
 TGATTTTGTAGTTACTACACTCTACCTTCTACCGTTGTTTATCATCCAGTGCATAAGACACTGAAAGCAGCATATTCTATCTAC
 AATGTATCTACCAAAACACCTTGGACTGAACTCATGCTTGATGCATTGATATCTGCTAAAAATTCTGGTTTGTATGTTTTTAATGC
 ATTGGATTTGATGGATAATAAAGAATTTTTGGAGCCTTTAAAATTTGGTATTGGTGAT
 >lac3441.2.1:A436A (not G) | cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 2
 ATGTTTCATAAACCCAGATAGTTATTTCCGACAGAAATATTATAAAGAAGATTTTCAGGACGAGCTGAACGACTTATTCAACAAAA
 GTTCCGATTTTCGGTTTAGCAGTGAATGGAAGTTGCCGGAGCGCAGTTGTGGTTTACAGCAGCGCCATGGAAGGTGAAGACGTTGG
 ATTCGTTGAAATCGCGACTAAATTTTTAACAAAAGTCAGCTTAATGACTTTAGTATTGAAGAGTGGAGCAGTCACACGCGGCGCCGG
 AACCTGCGGGCGAGGTGGGCTGGAAGCTGCGCTGCCTCGTGAACCCCGAGTTCTTGACACAAGCCTGGACCAAAATCTATGAATG
 CGCTTGACATACAATGTTGTGCCACTGGAAGCTATTGCAGATAAGAAAATGGTATCTTTGCACCTCTGTGAGGCTCCTGGAGCCT
 TTATTACATCCTTAAACCACTTCCTAAAACCTACGTCATCATGACTTACAG
 >lac886.16.8:T1215T (not C) | uncharacterized protein LOC101745102
 GTTGATTCTTTTACCGTGAACCAAAAACTAAGTGGTAATTCGTCGTGCATGAAAGTATACGACGTGTCTCAACATAGATCAACTCA
 TACAAGCCAACATTCTGAACCAATGAGTACATCCTCTACTGAAAACCAAACTTCTTCAGACAAAAGCCACCCAAAGTCATTTAATT
 TGGTAAAACTATTTATGAAACAGAAAAGTATTAGTAATGACGGAATTCATGGTATGGATCAATTAGATAGATCGGAATGTTGGCC
 TTCTAGTTACAGGTGGTGAGAGTGGTGAATCAATGGGAGAGCAAAAACCTGGGGACTCAAAATCTATATCAACAGATAGACCCCAA
 GCCGAATTACCATTAGATGCTGATTATAATAGCTCGGCTGTCTATGATGAAGTCCGATCAGTTAATCCATTTAATAGAGTATACG
 ACGAAGTTATTGAAGAAGAAGACAATCTTGAAACGACCAAAATTAAGCGGCAGCGAAACGAATTTATATGCTGTTGTTAAACAAAC
 GCACATTTAAAGAGCCAATATAAATATAAATAACAGCCCATGTAAACTAGGAGTAGGAAATCTTGTCTTCACAATCATCGGCT
 ACAAGTGTAGCATTTCAAGTTGTTCTGAATCAGATGGCACACAAATTACAAGAATGAACAAAGTGCTGTTAAAGAGTCTGTGA
 TCATAAATCTACATCTACACACTTAGAGGCTGATACGTTAGATAAAAAGTATGCAGACGTCAAATATCCAAACATCAACAATTTCTC
 ATGAACGCGAAGTCTTAAAATAGTGGAGCCGTCATTCTTAGAAAAGTTAAAAGAGGGCGACTGTGAGAAACCCGTTTTTGTACTC
 TATCCTAATTATACACTGCCAGACATAAGTTTCTTAAATGGTAGGCCAAATATTTACTTGAATCCGATAAAAGTAAACATTTCCCC
 TAAACCTAATGAAAGTAAAGAAGACACGTATGCAAGTTAAAGGGAAAAGGCCCTTTTTTCATGCAATGATGTAGAAATTTTAAAGAAAA
 AAGGTTTAGGCCATATCAAAGACTGGGACTCTCTTAACTTCTTGTGTCCTTTGGAGTGCAAGCAAATGTTGTGAGAATTACCAGA
 GCTAACCCAAATAAAAGAAAAAGAGGGTGTATCTAAATGCAATGACAAATACTGTAACCTCATCACCTTCGTGTAAGTCTAAGAAT
 AGACCAATGAGTTGTGAATTGCAATAATTTGATTGGAATAACACGGCTGTGTCTTCGAGCTCCAGTACAGCGACACAACCTTCATC
 GGGATATCGAGGATCCTCTACAATGTTGACGGACTCTTCTGCACAAAATAGCCCCGCCCTGCTGGAAATTTCAATCCTCTTTTGTG
 TTTATCGTTATGATAGTGCTACAAGCTCCGAGGCGAGTTGCGTCAATAATGACGGACTACGAATTATTCCTACCATACCGAAGCGA
 TCCTTATCGTTGGCCGATCAAACCTCGCATGGCCAAACAAGGCGAGTTGGCACCTCCGAGACCTCCATTACCAAAGAGTATCTTACG
 TAAATCCATGGACAAAACCTCGCAAGTCCACTAGTCATACGAAACGTTACAGCATGTTTGAAATGGATGAATTTATTCAAGATCCAA
 TTGTGTGCGCTACGGCTGCCGTAGAGCATAAACTAAACGAAGGTCTTTACAAGAACCCTTATTATTTACAAAATCAAAATATGGA
 TTATAGAAAGAACAATGATCTAGCTGCTAAACGACTATCTCAGCAATTCCTAGATGCCGCCGAGAAAGACGCTGATTACAATGAA
 TATTATCAAGATGAAGGTGTGGGAACAGAAAGCAGCCTTGAATCAGGAAAAATCAAATGAAATAAAATATCATAGACCACACACACC
 TCCATTACCGAAACCAAGAACTAAAAAAATGGAGTACATTGATTTTCCACCTCCGGGAGCACTTATTAGTAGCGCTGATTTACAAC
 AACTAGAAGAATTTCTGAAACAAAGTGATTCAACTGCCAAAATATGGATGAATGGGATCAAAATCAAGTA