

ISSN 1175-5326 (print edition)

 ZOOTAXA

 ISSN 1175-5334 (online edition)



# A newly recognized species in the *Anopheles* Hyrcanus Group and molecular identification of related species from the Republic of South Korea (Diptera: Culicidae)

CONG LI<sup>1</sup>, JOHN S. LEE<sup>2</sup>, JENNIFER L. GROEBNER<sup>2</sup>, HUENG-CHUL KIM<sup>3</sup>, TERRY A. KLEIN<sup>4</sup>, MONICA L. O'GUINN<sup>2</sup> & RICHARD C. WILKERSON<sup>1</sup>

<sup>1</sup> Department of Entomology, Walter Reed Army Institute of Research, 503 Robert Grant Ave., Silver Spring Maryland, 20910-7500, USA; e-mail: wilkersonr@msc.si.edu

<sup>2</sup> U.S Army Medical Research Institute of Infectious Diseases, 1425 Porter St., Frederick Maryland, 21702

<sup>3</sup> 5<sup>th</sup> Medical Detachment, 18<sup>th</sup> Medical Command, U.S. Army, APO AP 96205-5247

<sup>4</sup> 18<sup>th</sup> MEDCOM, Unit 15281, APO AP 96205-0054

## Abstract

We report here a previously unrecognized mosquito species from the Republic of South Korea that is closely related to *Anopheles* (*Anopheles*) *sinensis*. We also present an rDNA ITS2-based method for identification of this and the other morphologically similar *Anopheles* from the country. The PCR assay is a multiplex of seven primers in a single reaction that unambiguously identifies all included species. The other species include *An. sinensis*, *An. lesteri*, *An. pullus*, and a previously reported unknown *Anopheles*. Based on the ITS2 sequence, the two unnamed species have genetic distances from *An. sinensis* of 9.1% (unknown 1) and 10.7% (unknown 2), and are 10.6% different from each other. Adult females of *An. sinensis* and unknown 2 are morphologically indistinguishable, while unknown 1, *An. lesteri* and *An. pullus* can usually be separated from each other and from *An. sinensis* using available keys and descriptions. This PCR identification tool offers vector biologists and malaria epidemiologists a means to identify the potential vectors of malaria parasites in South Korea.

Key words: malaria, Anopheles, PCR identification, Hyrcanus Group, South Korea

# Introduction

Recent malaria transmission in the Republic of South Korea has resulted in an increased effort to incriminate the vector or vectors of this disease (reviewed in Wilkerson et al. 2003). Reported in the above study was the existence of an unknown species ("unknown

 $\overline{\mathbf{939}}$ 

1") related to *Anopheles (Anopheles) sinensis* Wiedemann in South Korea, and a description of the rDNA Internal Transcribed Spacer 2 (ITS2) sequence for those two species plus two other members of the Hyrcanus Group (Harbach, 2004) in South Korea that could be confused with *An. sinensis*; *An. pullus* Yamada and *An. lesteri* Baisas & Hu. Wilkerson et al. (2003) also reconfirmed that *An. pullus* (the senior synonym) and *An. yatsushiroensis* Xu & Feng were genetically identical in Korea and showed that *An. anthropophagus* Xu & Feng is a junior synonym of *An. lesteri*. Continued collection of specimens, rearing of progeny broods, and sequencing of representative specimens resulted in the recognition of another unknown species ("unknown 2") related to *An. sinensis*. The two "unknown" species have been formally described and named (Rueda, 2005). Rueda (2005) provides a discussion of distinguishing characters for the species which usually will serve to separate all but *An. sinensis* and unknown 2. Below, we provide the ITS2 sequence for *An.* unknown 2 and give a PCR method for identification of five of the Hyrcanus Group species found in South Korea. *An. sineroides* Yamada was not included since it is easily distinguishable morphologically.

# **Materials and Methods**

**Source of specimens.** PCR identifications were carried out on the same specimens reported in Wilkerson et al. (2003). An additional 50 specimens identified morphologically as belonging to the Hyrcanus Group (Tanaka et al. 1979, Lee 1998) were also tested. These include examples from three progeny broods of the previously unrecognized *An*. "unknown number 2" (Table 1). Morphological and DNA vouchers of the above are deposited in the Smithsonian Institution, National Museum of Natural History. A synoptic collection is kept by one of us (Hueng-Chul Kim) in his laboratory.

Morphological identification. Adult female specimens included in this study were first identified morphologically using either Tanaka et al. (1979) or Lee (1998).

**DNA isolation and sequencing.** DNA was isolated from individual adult mosquitoes by phenol-chloroform extraction as described in Wilkerson et al. (1993). The rDNA ITS2 was amplified using conserved sequence found in the 5.8S subunit, ITS2 forward (5'-TGTGAACTGCAGGACACATGAA–3') and in the 28S subunit, ITS2 reverse (5'-ATGCTTAAATTTAGGGGGGTAGTC–3') (Cornel et al. 1996). PCR products were directly sequenced using Big Dye 3.0 (Applied Biosystems Inc. –ABI) with an ABI 3100 sequencer (ABI). The sequence was then edited and analyzed using Sequencher (v4.2, AB). Sequence of *An. sinensis, An. lesteri, An. pullus,* and *An.* unknown number 1 are those of Wilkerson et al. (2003). GenBank accession numbers for the above are in Wilkerson et al. (2003). The accession number for the *An.* unknown 2 reported here is AY753740. Alignment was carried out with Clustal W (v. 1.82), with final alignments done manually. The boundaries of the rDNA ITS2 were estimated following Cornel et al. (1996). Sequences shown in Fig. 1 are of the ITS2 only. Complete amplicons have an additional

92 bases from the 5.8S subunit and 43 bases from the 28S subunit (total ITS2 and flanking sequence given in Fig. 1).

**TABLE 1.** Summary of collection localities for species belonging to the *Anopheles* Hyrcanus Group reported here for the first time.

| Province<br>(Locality)  | Coordinates                     | Date              | Collector  | Habi-<br>tat | Country prefix,<br>Collection No.,<br>(progeny brood<br>no.)                    | Species<br>identified by<br>PCR and/or<br>sequence |
|-------------------------|---------------------------------|-------------------|------------|--------------|---|--|
| Korea                   |                                 |                   |            |              |   |  |
| Ogeum-ri, Paju-<br>Paju | 37° 40' 00" N<br>126° 54' 00" E |                   | W.J. Lee   | cow<br>shed  | KS2-3(1,5)  | sinensis   |
|                         |                                 |                   |            |              | KS2-3(2-4,6,7)  | pullus   |
| Jeonju                  | 35° 49' 19" N<br>127° 8' 56" E  | July 26,<br>2002  | W.J. Lee   | cow<br>shed  | KS2-4(1-5)  | sinensis   |
| Camp Casey              | 37° 54' 15" N<br>127° 05' 23" E |                   | T.A. Klein | cow<br>shed  | KS2-5(1)  | lesteri  |
| Cheongpyeong            | 36° 40'N<br>128° 22'E           | August 2,<br>2002 | HC. Kim    | cow<br>shed  | KS2-6(1,3,5)  | sinensis   |
|                         |                                 |                   |            |              | KS2-6(2)  | Unknown 2  |
| Kunsan                  | 35° 58' 43" N<br>126° 47' 41" E |                   | T.A. Klein | cow<br>shed  | KS2-7(1-6)  | sinensis   |
| Majeong-ri, Paju        | 37° 50'N<br>126° 49'E           | June 5,<br>2004   | W.J. Lee   | cow<br>shed  | KS4-2(1), KS4-<br>2(17)   | Unknown 2  |
|                         |                                 |                   |            |              | KS4-<br>2(3,7,9,11,16,19-<br>21,24-27,31,34-<br>35,37-38,42-<br>43,45-48,50,52) | sinensis   |

**Primer design and multiplex PCR mixture** (Table 2). Species-specific reverse primers were designed based on species-specific sequence in the ITS2. When these primers are combined with universal ITS2 forward primer, they produce PCR products of different lengths for each species. This makes it possible to identify the five very similar species of the Hyrcanus Group in South Korea in a single reaction. The multiplex reaction also contains the universal ITS2 reverse primer so that if the DNA of a non-target species is tested, or no species-specific primer is included, a band representing the entire ITS2 will result (Fig. 2, lane 11). This serves as an internal control for the presence of amplifiable DNA.

**Distance analysis.** A Neighbor Joining (NJ) analysis as implemented in PAUP\* v.4 (Swofford 1998) was carried out using uncorrected "p" setting, with gaps treated as missing data. A distance matrix is presented in Table 3.

| (33)       unknown1<br>lester1<br>pullus       AATTAGAARTCAAACGCCCACTAGCATGGCCGCCCCCGCCGCCGCCCGCC   | ZOOTAXA | sinensis | AATTAGAAGTGGAAACGTGGACTTACGCAGTGATTGGTGCTGGTCACCACGTCAC-G  |      |
|---|---------|----------|--|------|
| Initiativity       Initiativity <t< td=""><td>(939)</td><td></td><td></td><td></td></t<>  | (939)   |          |  |      |
| pullus       ATTTGGETGTATACACAATAACTACCCGTGATGGTCCCACTGCACTG  |         |          |  |      |
| <pre>sinensis<br/>unknown1<br/>grCGTGAANAARGATGATGATGATGAGGCCCCCCCCCCCCCTTCATT-AAAACGTGTT 116<br/>unknown2<br/>GrCGTGAANAARGACTAAGTGAGGCACCCCCCCTCCCCTTCATT-AAAACGTGTT 116<br/>Unknown2<br/>GrCGTGCAANAAACGCCTAAGAAGTGGCGCACCCCCCTCCCATTC-AAAACGTGTT 117<br/>pullus<br/>GrCGTGCATAAAAACGCCTAAGAAGCTGCCCACCCCCCTCCATTC-AAAACGTGTT 117<br/>pullus<br/>GrCGTGCATAAAAACGCCCTAAGAAGCCAAGTAGAAGCGGCTGTGTCCCCGGGG 168<br/>unknown1<br/>TGAAAAGCCCCTAAGAAGCCAAGAAGCAAGTAGAAGGGCTGTGTCCCCGGG 164<br/>lesteri<br/>TGAAAAGCCCCTAAGAAGCAAGTAGAAGGGCTGTGTCCCCGGGG 164<br/>lesteri<br/>TGAAAAGCCCCTAAGAA-CAAGCAAGTAGAAGGGCTGTGTCCCCGGGG 164<br/>lesteri<br/>TGAAAAGCCCCTAAGAA</pre>   |         |          |  |      |
| unknown1<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unknown2<br>unkno |         | pullus   |  | 58   |
| unknown2<br>Lesteri<br>Justoren 2<br>Justoren 2   |         | sinensis | GTCGTGAATAATGATGTAAGATGGGGTCTCGTCGACCCGCTTGCATTT-AAAACGTGTGT   | 115  |
| Lesteri<br>pullus GTCGTGCATANGCCTAGAGG-GGGTCCGCGGCCGGCTGCATTGCATCCGGAGCTTGTG 11<br>pull GTCGGCATANGAGGCCTAGGA-GGGGCGCGCGCGCGCGCGCGCGCGCGCGGCTGGCTGGCT   |         | unknown1 | GTCGTGAAAAATGATGTAAGATGTTGTCTCGTCGACTCGCTTGCATTT-AAAACGTGTGT   | 116  |
| <pre>pullus GTCGTGCATAATGATCGAGACGCGCCCCCGCCCCCTATCATTC-AAGAACGCGC 116 pul  sinensis GTTGGAAAAGCCCTAAGAACGCGACACATAGAAGAACCCCCGGCGCGGCGCGGCGGGCG</pre>  |         | unknown2 |  |      |
| <pre>pul pul pul pul pul pul pul pul pul pul</pre>  |         |          |  |      |
| sinensis<br>unknownl<br>TGATAAGCAGCCTAAGCAAGCTAGAAAC  |         | pullus   | Sensitive and an antiparticipation and antiparticipation antiparticipati | 116  |
| unknown1 TCATAACCCCTAGGA-CGACAACAGATAGAAGGGCTGUTTCCCCCGG 163<br>unknown2 TGATAACCCCTAGGACGACAAGTAGAAACGGCTGUTTCCCCGG 163<br>unknown2 TGATAACCCCCGAGAGCAGACAAGTAGATAACAGCAGCAGAGTGUTTCCCCCGG 164<br>lesteri T  |         |          |  | 1.60 |
| unknown2       TGAAAAACCCCTAAGCA-CCACACACTAGTAGTAACGCCCTTCTTCCCCCGG 164         lesteri       TGTAAGCCCCTAGGACAAACTAGTAACTAGCACCCACACACTCCCCGG 159         sinensis       ACGGCGAGGAAGTATATTGAGCAGGCGCTCCTT-TGCTAGGTATGGAACAGG 227         unknown1       ACGGCGAGGAAGTATATTGAGCAGGCGCTCCTT-TGCCCTGCTAGGAACAGG 222         unknown2       ACGGCGAGGAAGTATATTGAGCAGGCGCTCTT-TGCCCTGCTAGGAACAGG 221         unknown2       ACGGCGAGGAAGTATATTCTGAGCAGGCGCCTCTT-TGCCCTGCTGTAATGGACGAGG 221         pullus       ACGGCGAGGAAGTATATTCTGAGCAGGCGCCCTTTCTGCTGCTGTA   |         |          |  |      |
| lesteri       T=GAAGGCCCCGAGACAAGAGCAGTAGTAGACAGCAGCAGCGAGTGTTTCCCCCGGA         pullus       T=GTTAGCCATGATATCAAGAGCAGTAGTAGGCGTGTCCTGGGG         sinensis       ACGGCGGAGGAAGTATATTGAGCAGGCGCGTCTT-TGCCTGCTGGGAAGAGG         unknown1       ACGGCGGAGGAAGTATATTGAGCAGGCGCGTCTT-TGCCTGCTGTGGAACAGG         unknown2       ACGGCGGAGGAAGTATCTGAGCAGGCGCGTCTT-TGCCTGCTGTGGAAGAGGCGC         pullus       ACGGCGGAGAATACCCAGGCGGCGCTCTT-TGCCTGCTGTGGAACAGG         pullus       ACGGCGGAGAATACCCATGCAGGCGCCCTCTGTGGCTGGA-AGGCAGCGG         sinensis       TGTCTTCCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTT   |         |          |  |      |
| pullus       TGTAAGCCATGATATCAAACTAGTAGTAGCGTGTTCCTGGG 159         sinensis       ACGGCGAGGAAGTATATTGAGCAGGCGGTCCTT-TGCCAGGTAGGAACAGG 221         unknown1       ACGGCGAGGAAGTATATTGAGCAGGCGGCTCTT-TGCCAGGTAGGAACAGG 221         unknown2       ACGGCGAGGAAGTATTCTGAGCAGGCGCCCTT-TGCCAGGTGGTAGGAACAGG 221         pullus       ACGGCGAGGAAGTATTCTGAGCAGGCGCCCTT-TGCCAGGTGTGTAATGGACGAG 221         pullus       ACGGCGAGGAATACCCATGAGGGCGCCTTGTGTGTGTGT-TGC-GTGGA-CAGG 211         unk2   |         |          |  |      |
| <pre>sinensis ACGCCGAGGAACTATATTGACAGCCGCGCTCCTT TGCTATGTCTAGGTATGGACAGG 227 unknown1 ACGCCGGAGGAACTATTCTCAGCAGCAGCGCTCTTTTGCTCAGCAGGTATGGACAGG 222 unknown1 ACGCGGGAGGAACTATTCTCTGAGCAGGCGCCTCTTTTTCTCTGCGGACACGACGG 221 lesteri TTGCCGGAGAATACCCCATGCAGGCCGCCCTCTGTGTGTGCT-ATGGACGGCGCCGCGGGGGGGAGGACGACAGG 211 unk2 sinensis TCTCTCCTCTTCTATTTTATTTTTTTTTAAATTGAGGTAAGGCACGAACAGCTTCGA 266 unknown1 TCTCTCCCCTTCTCTTTTAA-TTTTTTTTAAATTGAGGTAAGGCACGAACAGCTTCGA 279 unknown2 TCTCTCCTCTTCTTTTAA-TTTTTTTAAAATTGAGGTAAGGACCAACAGCTTCG 279 lesteri TCTCTCCTCTTCTTTTATTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTCG 279 unknown2 TCTCTCCTCTTCTTTTATTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTCG 279 lesteri TCTCTCCTCTTCTTTTTTATTTAAAATTGAGGTAAGATTTCCAACGGTTCG 279 unknown2 TCTGTTCCTCTTTTTTTATTTTAAAATTGAGGTAAGATTTCCAACGGTTCG 279 lesteri TCTGTCTCCTCTTTTTTTATTTTAAAATTGAGGTAAGATTTCCAACGGTTCG 279 lesteri TCTGTCTCCTCTTTTTTTATTTTAAAATTGAGGTAAGATTTCCAACGGTTCG 279 lesteri TCGAGATAGCATG-TACGCAAATAATCATTGTTTGGAACCCCTGAACAACGGGAACGT 344 unknown1 TCGAGACGCACGCATG-TACGCCAATTAATCATTGTTGGAACCCCCTGAACACGGAACGC 336 lesteri TCGAGACAGTAGCACGCCCAACGAATTATCATTGTTGGAACCCCCTGAACACGGAACACG 337 pullus TCGAGACAGTGGTCGGACGCAACGAATGGATCGTTGGAACCCCCTGAACACGGAACACG 336 lesteri TCGAGACAGTGGTCGGACGCAACGAATTGGAACACCACGAACACGAACGGAACACT 323 TCGGGACAGTGGTCGGACGCAACGAATTGGAACTATTGAACCAGACACGCAACGAACG</pre>  |         |          |  |      |
| unknown1 ACGGCGGAGGA <u>AGTATTCTGAGCAGGCGCTC</u> CTT-TGCCCTGTGTGGTAT-GAACAGG 222<br>unknown2 ACGGCGGAGGA <u>AGTATTCTGAGCAGGAGGCCC</u> CTGTGCTGTGTGTAATGGAGCAGG 221<br>lesteri<br>pullus ACGGAGGAAGTATTCTAGGCAGGCGCCCTGGCGTGTGTGTGT-AC-ATGGAGCAGG 211<br>unk2<br>unk2<br>unk2<br>unk2<br>unknown1 TGTGTTCCTCTTCTATTTTAA-TTTTTTTAAAATTGAGGTAAGGCACGAACAGGTTCGA 279<br>unknown2 TGTGTTCCTCTTCTATTTTAA-TTTTTTTTAAAATTGAGGTAAGGACCAA-ACGTTTCGA 279<br>pullus TGTGTTCCTCTTCTATTTTATTTTTTTTAAAATTGAGGTAAGGAACGAA-ACGTTTCGA 279<br>pullus TGTGTTCCTCTTCTATTTTATTTTTTTTAAAATTGAGGTAAGGTA   |         | pullus   |  | 155  |
| unknown2       ACGGCGGGGGAGAGAGTATTCTGAGCAGGGGCGCCCTGACGGTGTGTAAATGGACCAGG 221         lesteri       TTGGCGGAGGAGTTCTGGCGGGGCGCCTGTGGTGTGG  |         | sinensis | ACGGCGGAGGAAGTATATTGAGCAGGCGCGTCCTT-TGCTATGTGTAGGTATGGAACAGG   | 227  |
| lesteri       TTGGCGGAAG       TTGGCGGAGGCGCCCCTCACCCTCTGTGCTAG-ATCGACCAGG 224         nullus       AcGABGGAATACCCATGCAGGCGCGTCGTTGTGTGTGTGTGT_AGATCGACCAGG 211         unk2       unk2         sinensis       TGTCTCCTCTTCTATTTTAAATTGAGGTAAGGCACCAA-CGGTTCGA 287         unknown1       TGTCTTCCTCTTCTATTTTAAATTGAAGGTAAGGCACCAA-ACGTTTGGA 279         unknown2       TGTCTTCCTCTCTTTTTAAATTTTTTTAAAATTGAGGTAAGGCACCAA-ACGTTTGCA 279         pullus       TGTCTTCCTCTCTTCTATTTTAA-TTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTGA 279         pullus       TGTCTTCCTCTCTTCTATTTTTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTA- 265         sinensis       TGGCGACAGCAGCTGCGCAATAATCATTGTATGGAACCCCTGAACAACGGAACACT 344         unknown1       TCGAGACGGCATG-TACGCCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT 336         unknown2       TTGAGACAGTAG-GACGCAACAGAAGAGGAACCTGC-GAAGGTGAACCACGGAACACT 323         pullus       TCGAGACAGTAG-GACGCAACCAAGAGGAACTTGCAGAA-TGGATGCAACGAACACACAAAGGGAACACT 323         sinensis       TATGGCACTAGGAGAACACTACCCAGA-TTTGTTGT-AGCGGGCCCACACAACAACAAAGGAACACT         unknown2       AAAGGCACCAAGAGAACACTACCCAGA-TTTGTTGGATAG-GGGACCAGAGAATCAA 388         lesteri       C   |         | unknown1 | ACGGCGGAGGA <u>AGTATTCTGAGCAGGCGCGTC</u> CTT-TGCCGTGTGTAGGTATGGAACAGG  | 222  |
| pullus       ACGGAGGAATACCCATGCAGGCGCGTTCTGTTCTGT   |         | unknown2 | ACGGCGGAGGA <u>AGTATTCTGAGCAGGAGCGTC</u> CTT-TGCTGTGTGTAAATGGAGCAGG  | 221  |
| unk2         sinensis       TGTCTTCCTCTCTCTATTTTAAATTTTTTTAAAATTGAGGTAAGGCACCAACGTTCGA 287         unknown1       TGTCTTCCTCTCTCTATTTTAA-TTTTTTTAAAATTGAGGTAAGGCACCAA-ACGTTTCGA 279         unknown2       TGTCTTCCTCTCTCTTTTTTAA-TTTTTTTAAAATTGAGGTAAGATTTCCAAAGTTTCCA-279         pullus       TGTCTTCCTCTCTCTTTTTTTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTTCCA-279         pullus       TGTCTTCCTCTCTCTTTTTTTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTCT-279         pullus       TGTCTTCCTCTCTCTTTTTTTTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTCT-279         pullus       TGTCTTCCTCTCTCTTTTTATTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTCT-279         pullus       TGTCTTCCTCTCTTTTATTTTTTTAAAATTGAGGTAAGATTTCCA-ACGTTTCT-279         pullus       TGTGTTCCTCTCTCTTTTATTTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTCCA-265         lesteri       TGTGAGACGTAGC-TACGCAAATAATCATTGATGGAACCCCTGAACAAGGGAACACT 344         unknown1       TGGAGACGTGG-TACGCAAAGAGATTATCCATGTATGGAACCCCTGAACAAGGGAACACT 323         sinensis       TGTGGCACTAGGGACACTACCCCAGA-TTTGTTGTA-GACGCACCAACGAACAACAGGAACACT 323         unknown1       TATGGCACTAGGGAACACTACCCCAGA-TTTGTGTA-GACGGACCC   |         | lesteri  | TTGGCGGAAGTTCTAGGCAGGCGCGCCCTGACGCTGTGTGTAGATGGAGCAGG  | 224  |
| sinensis<br>uknown1<br>tGTCTTCCTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTT  |         | pullus   | ACGGAGGAAATACCCATGCAGGCGCGTTCTGTTGCTGTGT-TAGGTGGA-CAGG   | 211  |
| unknown1 TGTCTTCCTCTTTTTTTTTTTTTTTTTTTTTAAATTGAGGTAAGGACCAA-ACGTTTCGA 279<br>unknown2 TGTCTTCCTCTTCTATTTTAA-TTTTTTTAAAATTGAGGTAAGATTTCCAACGTTTCG- 279<br>pullus TGTCTTCCTCTTCTATTTAA-TTTTTTTAAAATTGAGGTAAGATTTCCAACGTTTA- 265<br>lesteri<br>TCGAGATAGCATG-TACGCAAATAATCATTGTTGTAGGA-ACCCCTGAACAACGGAACACT 344<br>unknown1 TCGACAGGCATG-TACGCTAATAATCATTGTATGGAACCCCTGAACAACGGAACACT 336<br>lesteri<br>TCGAGATAGTGGATG-GACGCAAATAATCATTGTAGGAACCCCTGAACAACGGAACACT 337<br>pullus TCGAGACAGTGG-TGCGAAGACTGGAACGC-GAACAACGGAACACT 337<br>pullus TCGAGATAGTGGATGGATGGGCTGCAAGAGATTGGAA-ACCCCTGAACAACGGAACACT 337<br>pullus TCGAGATAGTGGATTGGAATGATGTTGTAGAC-ACCCCGGACAACGGAACACT 323<br>  |         |          | unk2   |      |
| unknown2       TGTCTTCCTGTTCTATTTTAA-TTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTTCG-279         lesteri       TGTCTTCCTCTCTATTTTTATTTTAAAATTGAGGTAAGATTTCCAAAGGTTTCG-279         pullus       TGTCTTCCTCTTTTATATTTTAAAATTGAGGTAAGATTACCA-ACGTTTCT-279         sinensis       TCGAGATAGCATG-TACGCAAATAATCATTGTAGGA-ACACCCTGAACAACGGAACACT         unknown1       TCGACACGCACTG-TACGCTAATAATCATTGTATGGAACCCCTGAACAACGGAACACT         unknown2       TTGAGACAGTATG-GACGCAAATAATCATTGATGGA-ACCCCTGAACAACGGAACACT         pullus       TCGACACGGAATGGG-TGCGAGAGAGCTGGAATCG-CAGATGAACAGGGAACACT         sinensis       TATGGCACTAGGAATGGGTCCAAGAGATGGAACG-GGGCCCGACGAACAATAAT         unknown1       TATGGCACTAGGGAACACTACCCAGA-TTTGTT <u>TTTTTTTTTTTTTTTTTTTTTTTTTTTTT</u>   |         | sinensis | TGTCTTCCTCTTCTATTTTAATTTTTTTTTTTAAAATTGAGGTAAGGCACCAACACGTTTCGA  | 287  |
| lesteri       TGTCTTCCTCATCTATTTTATTTTAAAAATTGAGGTAAGATTTCCA-ACGTTTC-       279         pullus       TGTCTTCCTCTCTATTTATTATTTAAAATTGAGGTAAGATTACCA-ACGTTTA-       265         les ◆       1       1         sinensis       TCGAGATAGCATG-TACGCAAATAATCATTGTATGGAACCCTGAACAACGGAACACT       336         unknown1       TCGAGACGGCATG-TACGCCTAATAATCATTGTATGGAACCCCTGAACAACGGAACACG       336         lesteri       TCGAGACAGTAG-GACGCACAACAACAGCGAACGGAACACG       337         pullus       TATGGCACTGGA-TGGACGCTGCAACGAACGGAATCG-GAAGCAAACGGAACACG       337         pullus       TATGGCACTAGAGAACACTACCCAGA-TTTGTATGGAACCCCTGAACAACGGAACACG       339         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATAG-CGGGTCCTAACGATACA       388         lesteri       CTATTGCCACACACAGAACCTACCCAGA-TTTGTGCATAG-CGGGTCCGAAGATACA       388         lesteri       CTATTGCAACACACACACACCACGCAGATCCCCGACACCACGCGAGGCCCCTGC       382         unknown2       AAACGCACCA-TAAACACTACCCAGATTCTCACAGAGGGAGAGGCCACTCG-GGGCCACTGCG-GGGGCA       388         lesteri       CTATTATCACTA-AAGGTCAATATACACTCACGAGGGAGGCCACTCG-GGGGCA       382         unknown1       -GCAAACAACA-AAGGTCAAAATATCATTACACTCACGAGGGAGGCCACTCG-GGGCCA       382         unknown1       -GCAAACAACA-AAGGTCA-AAGGTCAATATCACTCACGAGGTGAGGCCACTCG-GTGGTCA       440         unknown1       <  |         | unknown1 | TGTCTTCCTCTTCTATTTTATTTTTTTAAAATTGAGGTAAGGCACCAA-ACGTTTCGA   | 279  |
| pullus       TGTCTTCCTTTTATTTATTTATTTAAATTGAGGTAAGATTACCA-ACGTTTA-<br>les       265         sinensis       TCGAGATAGCATG-TACGCAAATAATCATTGTAGGAACCCCTGAACAACGGAACACT       344         unknown1       TCGAGACGGCATG-TACGCCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT       336         unknown2       TTGAGACAGCTATG-GACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT       337         pullus       TCGAGACAGTCGTGGGACGCACGAATTGGAACGGAACACGGAACACT       337         pullus       TATGGCACTAGAGAACACTACCCAGA-TTTGTATGTAGC_GGGCCCGACAACAAATAT       402         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGTATAG-CGGGTCGGACGATACA       389         unknown2       AATAGCACAAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA       388         lesteri       CTTATTGCAAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA       388         lesteri       CTTATTGCAAACACTACCCAGATTGTGTGACGAGGCGACGCCACTCG-GAGGTCA       388         lesteri       CTATTACCAAACACAACCAACCAGATTGTGTGACGAGGCGACTCGC-AGCATACA       388         lesteri       C   |         | unknown2 | TGTCTTCCTCTTCTATTTTAA-TTTTTTAAAATTGAGGTAAGATTTCCAAAGGTTTCG-  | 279  |
| sinensis       TCGAGATAGCATG-TACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT       344         unknown1       TCGAGACGGCATG-TACGCTAATAATCATTGTATGGAACCCCTGAACAACGGAACACT       336         unknown2       TTGAGACAGTGGGATG-GACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT       336         lesteri       TCGAGATAGTGGGATGGGCTGCAAGAGACTGAGAACTGGA-ACCCGGAACAACGGAACACT       337         pullus       TCGAGACAGTGGTGGGAACGAACTAGCGAACTATGAAACCGAACAACGGAACACT       323         sinensis       TATGGCACTAGGGAACACTACCCAGA-TTTGTTAGC-GGGGCTCGACAACAAACGGAACACT       323         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-GGGTCCTAACGATACA       389         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGCATAG-CGGATGGACGATACA       389         unknown1       TATGGCACAACAACTACCCAGA-TTTGTGTGATAG-CGGATGGACGATACA       389         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTGATAG-CGGATGGACGATACA       389         unknown1       TATGGCACAAACAA-AAGGTCAATATCACTGATAG-CGGATGGAGATACA       380         unknown1   |         | lesteri  |  |      |
| sinensis TCGAGATAGCATG-TACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACT 344<br>unknown1 TCGAGACAGCAG-TACGCTAATAATCATTGTATGGAACCCCTGAACAACGGAACACT 336<br>unknown2 TTGAGACAGTATG-GACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACG 336<br>lesteri TCGAGACAGTAG-GACGCACGAAGAGCTGGAATCGGAAGTTGAACAACGGAACACT 337<br>pullus TCGAGACAGTGGTGGGACGCACGAAGGACTGGAATCGGAAGTTGAACAACGGAACACT 323<br>sinensis TATGGCACTAGAGAACACTACCCAGA-TTTGTT <u>ATGTTAG-CGGGCTGGACAACGAACAACGGAAACT 323</u><br>unknown1 TATGGCACTAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATCGGACGATACA 389<br>unknown2 AATAGCACAAGGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388<br>lesteri CTATTAGCAAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 386<br>pullus AAACGCACCA-TAAACACTACCCAGAATCGTGGC-ATAA-CGACTGGAAGATGCA 386<br>sinensis -ACAGCAAACA-AAGGTCAAACAATCACCCAGTGTTGAGAGAGCGGACGAGCACTCG-AGGATGCA 386<br>unknown1 -GCAAACAA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown1 -GCAAACAA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCT-AGAATCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown1 - GCAAACAA-AAGGTCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown1 - GCAAACAA-AAGGTCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown1 - GCAAACAACA-AAGGTCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 442<br>sinensis C 459 594<br>unknown1 T 441 576<br>lesteri A 438 573<br>pullus A 443 578  |         | pullus   |  | 265  |
| unknown1       TCGAGACGGCATG-TACGCTAATAATCATTATATGGAACCCCTGAACAACGGAACACT 336         unknown2       TTGAGACAGTATG-GACGCAAATAATCATTATTATGGAACCCCTGAACAACGGAACACG 336         lesteri       TCGAGATAGTGGATGGGATGGCAAGGACTGGAATCGGAAGTGAACAACGGAACACT 323         pullus       TCGAGACAGTGGTGGGATGGCAAGGAATCGAATGGAACCA-ACGGACAACT 323         sinensis       TATGGCACTAGGGAACACTACCCAGA-TTTGTATGTAGC-CGGGCTGGACAACAATAAT 402         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGCTCGACAACAATAAT 402         unknown1       TATGGCACAGAGAACACTACCCAGA-TTTGTGCATGG-CGGCTCGACAACAATAA         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGAGTGCGACGATACA 389         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 389         unknown2       AATAGCACAAGAACACTACCCAGA-TTTGTGTATAG-CGAGTGAGACGATACA 386         pullus       AAAGCCACAAGAAACAACTAACCCAGACGTGGTGTGTGAAGAGAGCGAGC   |         |          |  |      |
| unknown2       TTGAGACAGTATG-GACGCAAATAATCATTGTATGGAACCCCTGAACAACGGAACACG 336         lesteri       TCGAGACAGTGGTGGGACGCAGAGAGCTGGAATCGGAAGTTGAACAACGGAACACT 323         rulus       TCGAGACAGTGGTGGGACGCAACGAATTGGAACTATTGAAACCAGACAACGGAACACT 323         sinensis       TATGGCACTAGAGAACACTACCCAGA-TTTGTATGGA-CGGGCCCTAACGAACAACAACAACAACAACAACAACAACAACAACAAC   |         |          |  |      |
| lesteri       TCGAGATAGTGGAATGGGCTGCAAGAGACTGGAATCGGAAGTTGAACAACGGAACACT 337         rulus       TCGAGACAGTGGTGGGACGCAACGAATTGGAACTATTGAAACCAGACAACGGAACACT 323         sinensis       TATGGCACTAGAGAACACTACCCAGA-TTTGTT <u>ATGTTAG-CGGGCTGGACAAC</u> AATAAT 402         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGTCCTAACGATACA 389         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388         lesteri       CTATTAGCAAACACTACCCAGA-TTTGTGGTATAG-CGGATGGACGATACA 388         lesteri       AAACGCACCA-TAAACACTACCCAGA-TTTGTGGA-AGAGCGACGACGACGCACTGC 382         sinensis       -ACAGCAAACA-AAGGTCAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458         unknown1       -GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AGTTCTACCCG-AGAATCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AGTCTCTACCTG-AGAATCAATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AAAGCGTGGCATACAATTCATCAATTATCACTTACGAGTGAGGCCCCTG-GTGGTCA 440         lesteri       AGTCTCTACCTG-AGAATCAATTATCACTTACGAGTGAGGCCCCTG-GTGGTCA 442         sinensis       G 459 594         unknown1       T 441 576         unknown2       A 443 578  |         |          |  |      |
| pullus       TCGAGACAGTGGTGGGACGCAACGAATTGGAACTATTGAAACCAGACAAGGGAACACT 323         sinensis       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGGCTGGACAACAATAAT 402         unknown1       TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGCTCCTAACGATACA 389         unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGCGACGATACA 388         lesteri       CTATTAGCAAACACTACCCAGAATCCTGCGATAA-CGACTGGACGATACA 386         pullus       AAACGCACCA-TAAACACTACCCAGATTGTGAGAGGAGCGGACGAGCATCGCAAGCCTGG 382         sinensis       -ACAGCAAACA-AAGGTCCAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458         unknown1       -GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACA-AAGGTCAATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         unknown1       -ACAGCGCAGCTGCAACAATTCATCAATTATCACTTACGAGTGAGGCCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACAA-AAGGTCATTATCACTTACGAGTGAGGCCCACTCG-GTGGTCA 442         sinensis       G 459       594         unknown1       T 441       576         unknown2       A 443       578   |         |          |  |      |
| sinensis TATGGCACTAGAGAACACTACCCAGA-TTTGTT <u>ATGTTAG</u> - <u>CGGGCTGGACAA</u> CAATAAT 402<br>unknown1 TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGTCCTAACGATACA 389<br>unknown2 AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388<br>lesteri CTATTAGCAAACACTACCCAGA-TTTGTGTATAG-CGAGTGGAGATACA 386<br>pullus AAACGCACCA-TAAACACTACCCAGATCGTGCATAA-CGACTGCGAGATGCA 386<br>sinensis -ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAGGCCACTGCG-GTGGTCA 458<br>unknown1 -GCAAACAA-AAGGTCAATTATCACTCCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown2 AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440<br>sinensis G 459 594<br>unknown1 T 441 576<br>unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         |          |  |      |
| unknown1 TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGTCCTAACGATACA 389<br>unknown2 AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388<br>lesteri CTATTAGCAAACACTACCCAGAATCCGTGCATAA-CGACTGGAAGATGCA 386<br>pullus AAACGCACCA-TAAACACTACCCAGATGTTGAGAGAGCGGACGAGCATCGCAAGCCTGCG 382<br>sin<br>sinensis -ACAGCAAACA-AAGGTCAAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458<br>unknown1 -GCAAACAA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown2 AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCTG-AGAATCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>sinensis G 459 594<br>unknown1 T 441 576<br>unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         | pullus   |  | 325  |
| unknown1 TATGGCACTAGAGAACACTACCCAGA-TTTGTGCATGG-CGGTCCTAACGATACA 389<br>unknown2 AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388<br>lesteri CTATTAGCAAACACTACCCAGAATCCGTGCATAA-CGACTGGAAGATGCA 386<br>pullus AAACGCACCA-TAAACACTACCCAGATCCTGAGAGAGCGGACGAGCATCGCAAGCCTGCG 382<br>sin<br>sinensis -ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAAGGCCACTCG-GTGGTCA 458<br>unknown1 -GCAAACAA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown2 AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCTG-AGAATCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>sinensis G 459 594<br>unknown1 T 441 576<br>unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         | sinensis | TATGGCACTAGAGAACACTACCCAGA-TTTGTTATGTTAG-CGGGCTGGACAACAATAAT   | 402  |
| unknown2       AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA 388         lesteri       CTATTAGCAAACACTACCCAGAATCCGTGCATAA-CGACTGGAAGATGCA 386         pullus       AAACGCACCA-TAAACACTACCCAGGATCCGTGGC-ATAA-CGACTGC-GAGGATCCG 382         sinensis       -ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAAGGCCACTCG-GTGGTCA 458         unknown1       -GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         unknown2       AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AGTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AGTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440         lesteri       AAAGCGTGGCATACAATCAATCAATCAATATCACTTACAAACATAGTGGGACCCGTACGGTCA 442         sinensis       G 459 594         unknown1       T 441 576         lesteri       A 438 573         pullus       A 443 578  |         |          |  |      |
| pullusAAACGCACCA-TAAACACTACCCAGTTGTTGAGAGAGCGGACGAGCATCGCAAGCCTGCG 382sinensis-ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458unknown1-GCAAACAA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440unknown2AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTTCTACCTG-AGAATCATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTCTACCTG-AGAATCATTATCACTT-ACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTCTACCTG-AGAATCATTATCACTTACAAACATAGTGGGACCCGTACGGTCA 442sinensisG 459 594unknown1T 441 576lesteriA 438 573pullusA 443 578  |         | unknown2 | AATAGCACAAGAGAACACTACCCAGA-TTTGTGTATAG-CGGATGGACGATACA   | 388  |
| sinensis -ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458<br>unknown1 -GCAAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>unknown2 AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440<br>lesteri AGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 437<br>pullus AAAGCGTGGCATACAATTCATCAATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 442<br>sinensis G 459 594<br>unknown1 T 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         | lesteri  | CTATTAGCAAACACTACCCAGAATCCGTGCATAA-CGACTGGAAGATGCA   | 386  |
| sinensis-ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA 458unknown1-GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440unknown2AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 437pullusAAAGCGTGGCATACAATTCATCAATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 442sinensisG 459 594unknown1T 441 576lesteriA 438 573pullusA 443 578   |         | pullus   |  | 382  |
| unknown1-GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA440unknown2AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA440lesteriAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA437pullusAAAGCGTGGCATACAATTCATCAATTATCACTTACGAGTGGGACCCGTACGGTCA442sinensisG 459594unknown1T 441576lesteriA 438573pullusA 443578  |         |          | sin  |      |
| unknown2AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA 440lesteriAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 437pullusAAAGCGTGGCATACAATTCATCAATTATCACTTACAAACATAGTGGGACCCGTACGGTCA 442sinensisG 459 594unknown1T 441 576unknown2A 441 576lesteriA 438 573pullusA 443 578  |         | sinensis | -ACAGCAAACA-AAGGTCAAACAATTATCACTCC-AAGAGTGAGGCCACTCG-GTGGTCA   | 458  |
| lesteri<br>pullusAGTTCTACCTG-AGAATCATTATCACTTACGAGTGAGGCCACTCG-GTGGTCA 437sinensisG 459 594unknown1T 441 576unknown2A 441 576lesteriA 438 573pullusA 443 578  |         | unknown1 | -GCAAACAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA   | 440  |
| pullusAAAGCGTGGCATACAATTCATCAATTATCACTTACAAACATAGTGGGACCCGTACGGTCA 442sinensisG 459 594unknown1T 441 576unknown2A 441 576lesteriA 438 573pullusA 443 578  |         | unknown2 | AGTTATCAACA-AAGGTCAATTATCACTCACGAGTGAGGCCACTCG-GTGGTCA   | 440  |
| sinensis G 459 594<br>unknown1 T 441 576<br>unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         | lesteri  | 그는 것은 것 것 같아요. 그는 것 같아요. 그는 것 같아요. 이 같이 있는 것 같아요. 이 것 않아요. 이 것 같아요. 이 집 ? 이 것 같아요. 이 집 ? 이 것 같아요. 이 집 ? 이 |      |
| unknown1 T 441 576<br>unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         | pullus   | AAAGCGTGGCATACAATTCATCAATTATCACTTACAAACATAGTGGGACCCGTACGGTCA   | 442  |
| unknown2 A 441 576<br>lesteri A 438 573<br>pullus A 443 578   |         |          |  |      |
| lesteri A 438 573<br>pullus A 443 578   |         |          |  |      |
| pullus A 443 578  |         |          |  |      |
|   |         |          |  |      |
|   |         | -        |  |      |

**FIGURE 1.** Ribosomal DNA ITS2 sequence for five *Anopheles* Hyrcanus Group species from the Republic of South Korea. Bases in the gray areas are common to all species. Species-specific primers and direction of amplification are indicated by arrows:  $\sin = An$ . *sinensis*,  $\operatorname{unk1} = An$ . unknown 1,  $\operatorname{pull} = An$ . *pullus*,  $\operatorname{unk2} = An$ . unknown 2,  $\operatorname{les} = An$ . *lesteri*. Two numbers are given at the end of the figure: 1) ITS2 length, and 2) total length of amplified fragment, which includes 135 bases from the flanking regions.

4

**TABLE 2.** Species-specific primers used to identify five species in Hyrcanus Group from South Korea. Reactions were a multiplex of these five primers and the universal ITS2F and ITS2R primers. Diagnostic band size(s) are on the right. \*The *An. sinensis* primer was first reported by Ma et al. (1998).

| ZOOTAXA |
|---------|
| 939     |

| Species       | Species-specific primers (name: sequence) | Diagnostic bands |  |  |
|---------------|---|------------------|--|--|
| An. sinensis  | sin: 5'- GTTGTCCAGCCCGCTAACAT -3'*        | 488bp            |  |  |
| An. lesteri   | les: 5'- AGCCCATTCCACTATCTCGAAG -3'       | 391bp            |  |  |
| An. pullus    | pul: 5'- TTGATATCATGGCTTAACACCG -3'       | 227bp            |  |  |
| An. unknown 1 | unk1: 5'- CATTTTTCACGACTGCGACGG -3'       | 287bp 162bp      |  |  |
| An. unknown 2 | unk2: 5' – GACGCTCCTGCTCAGAATACT – 3'     | 288bp            |  |  |

**TABLE 3.** Neighbor Joining distance matrix, using uncorrected "p" with gaps treated as missing data.

|   |           | 1       | 2       | 3       | 4       | 5 |
|---|-----------|---------|---------|---------|---------|---|
| 1 | sinensis  | -       |         |         |         |   |
| 2 | unknown 1 | 0.09106 | -       |         |         |   |
| 3 | unknown 2 | 0.10698 | 0.10560 | -       |         |   |
| 4 | lesteri   | 0.26778 | 0.27309 | 0.23632 | -       |   |
| 5 | pullus    | 0.33152 | 0.33957 | 0.31407 | 0.29896 | - |

**Mosquito PCR identification.** Adult mosquitoes used here were preserved either in 100% ethanol or in some cases originated from dry pin-pointed museum specimens. DNA was extracted by boiling a single mosquito leg in 20  $\mu$ l of TE for 10 minutes and then the DNA was spun at high speed (10,000g) for 10 minutes. One  $\mu$ l of template supernatant was added to a 25  $\mu$ l PCR reaction containing 1X buffer (10mM Tris-HCl, pH8.3, 50mM KCl); 2.5mM MgCl2; 200 $\mu$ M dNTP; 6pmol of each ITS2 forward, reverse and unknown 1 (unk1) primers, and 4pmol each of *An. sinensis* (sin), *An. lesteri* (les), *An. pullus* (pul), and unknown 2 (unk2) primers; and 0.5 U Taq polymerase (ABI). The primer sequences are listed in Table 2 and their locations in the ITS2 are shown in Fig. 1. PCR reactions were performed using the following parameters: 95°C for 2 min, 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min, final extension at 72°C for 7 min. PCR products were visualized on a 2% agarose gel containing ethidium bromide.

# **Results and Discussion**

ZOOTAXA

939

We present here discovery of a second unknown *Anopheles* Hyrcanus Group species from South Korea that is closely related to *An. sinensis* and is also related to a previously discovered unknown *Anopheles* from the same region. We also provide an rDNA ITS2-based method for identification of all five morphologically similar species belonging to the Hyrcanus Group in South Korea.

We raised three progeny broods of the *An*. unknown 2 and sequenced the ITS2 from at least one individual of each. The ITS2 sequences were identical in all individuals sequenced, showing no indels or muations. The ITS2 of the *An*. unknown 2 is 441 bases long with a total amplicon, including flanking sequence, of 576, the same as *An*. unknown 1. It has a GC content of 45%, similar to *An*. unknown 1 (47%), and the other related species: *An. sinensis* 45%, *An. lesteri* 47%, *An. pullus* 47% (Wilkerson et al. 2003). Note that in the figure caption in Wilkerson et al. (2003) numbers 5 and 7 are reversed. The table itself and the description in the text are correct.

A pairwise comparison of the five species treated here (Table 3) shows that the two unknown species and *An. sinensis* are much more similar to each other, about a 10% difference between them, than to *An. lesteri* and *An. pullus*, which are approximately 24–34% different from each other and the *sinensis*/unknowns. These differences are sufficient for design of species-specific primers.



**FIGURE 2.** Results of amplification of rDNA ITS2 of *An. sinensis* (lane, progeny brood number): (1) KS8(67), (2) KS8(94); *An. lesteri*: (3) KS8(59), (4) KS8(88); *An. pullus*: (5) KS8(76), (6) KS8(86); *An.* unknown 1: (7) KS8(12), (8) KS7(27); *An.* unknown 2: (9) KS2-6(2) and (10) KS4-2(1); (11) positive control, complete ITS2 amplicon of KS8(67), *An. sinensis*; (12) negative control, no template. M: DNA ladder consisting of lambda DNA digested by *Hind* III, and phiX174 DNA digested with *Hae* III (Sigma, St. Louis, MO).

As stated in the introduction, *An. sinensis* and unknown 2 are not separable morphologically, and key characters to separate *An. sinensis* from unknown 1, *An. lesteri* and *An. pullus* are not completely reliable (see Rueda (2005) for a discussion of available characters). Use of a multiplex PCR assay with five species-specific primers and the universal ITS2F and ITS2R primers resulted in unambiguous species identifications of the five species. The addition of the ITS2R primer to the reaction mix served as a positive control for detection of *Anopheles* DNA that is not from one of the five species-specific band was observed with and without the complete ITS2 amplicon (ITS2 plus flanking regions) due to direct primer competition that depends on the starting concentration of the mosquito DNA. Examples of this can be seen in lanes 7–10 of Fig. 2.

We successfully tested the multiplex PCR reaction mix described above against at least one individual from progeny broods of 96 *An. sinensis*, 29 *An. pullus*, 10 *An. lesteri*, 3 *An.* unknown 1, and 3 *An.* unknown 2, all from South Korea. Molecular-based diagnostic methods such as the one reported here, will enable entomologists to correctly identify and correlate potential vectors with the diseases they carry.

### Acknowledgements

The material presented here reflects the views of the authors and should not be construed to represent those of the Department of the Army or the Department of Defense. This research was performed under a Memorandum of Understanding between the Walter Reed Army Institute of Research and the Smithsonian Institution, with institutional support provided by both organizations. Financial support was provided in part by NIH AI RO154139 to Jan Conn.

#### References

- Cornel, A.J., Porter, C.H. & Collins, F.H. (1996) Polymerase chain reaction species diagnostic assay for Anopheles quadrimaculatus cryptic species (Diptera: Culicidae) based on ribosomal DNA ITS2 sequences. Journal of Medical Entomology, 33, 109–116.
- Harbach, R.E. (2004) The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. *Bulletin of Entomological Research*, 94, 537–553.
- Lee, K.W. (1998) A revision of the illustrated taxonomic keys to genera and species of female mosquitoes of Korea (Diptera, Culicidae). 5<sup>th</sup> Medical Detachment, 168 Medical Battalion, 18<sup>th</sup> Medical Command, U.S. Army, 40 pp. (available at http://wrbu.si.edu)
- Ma, Y., Qu, F., Xu, J. & Zheng, Z. (1998) Sequence differences of rDNA-ITS2 and species-diagnostic PCR assay of Anopheles sinensis and Anopheles anthropophagus from China. Journal of the Medical College People's Liberation Army, 13, 123–128.
- Rueda, L.M. (2005) Two new species of *Anopheles (Anopheles)* Hyrcanus Group (Diptera: Culicidae) from the Republic of South Korea. *Zootaxa*, 941, 1–26.

Swofford, D.L. (1998) PAUP\*. Phylogenetic analysis using parsimony (\*and other methods), Ver-

ANOPHELES SINENSIS

ZOOTAXA

939



sion 4. Sinauer Associates, Sunderland, Massachusetts.

- Tanaka, K., Mizusawa, K. & Saugsted, E.S. (1979) A revision of the adult and larval mosquitoes of Japan (including the Ryukyu Archipelago and Ogasawara Islands) and Korea (Diptera: Culicidae). Contributions of the American Entomological Institute, 16, ii–vii, pp. 1–987.
- Wilkerson, R.C., Parsons, T.J., Albright, T.A., Klein, T.A. & Braun, M.J. (1993) Random amplified polymorphic DNA (RAPD) markers readily distinguish cryptic mosquito species (Diptera: Culicidae: Anopheles). Insect Molecular Biology, 1, 205–211.
- Wilkerson, R.C., Li, C., Rueda, L.M., Kim, H-C, Klein, T.A. Song, G-H & Strickman, D. (2003) Molecular confirmation of *Anopheles (Anopheles) lesteri* from the Republic of South Korea and its genetic identity with *An. (Ano.) anthropophagus* from China (Diptera: Culicidae). *Zootaxa*, 378, 1–14.