The New Internal Transcribed Spacer 2 Diagnostic Tool Clarifies the Taxonomic Position and Geographic Distribution of the North American Malaria Vector Anopheles Punctipennis

James M. Hodge
Virginia Polytechnic Institute and State University

Andrey A. Yurchenko
Institute of General Genetics and Cytology

Dmitriy A. Karagodin
Institute of Cytology and Genetics

Reem A. Masri
Virginia Polytechnic Institute and State University

Ryan C. Smith
Iowa State University

Mikhail I. Gordeev
Moscow Region State University: Moskovskij gosudarstvennyj oblastnoj universitet

Maria V Sharakhova (msharakh@vt.edu)
Department of Entomology and the Fralin Life Sciences Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
https://orcid.org/0000-0002-5790-3548

Research

Keywords: Anopheles punctipennis, mosquito, molecular diagnostics, internal transcribed spacer, restriction fragment length polymorphism

DOI: https://doi.org/10.21203/rs.3.rs-83154/v1

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

**Background:** The malaria mosquito *Anopheles punctipennis* Say, a widely distributed species in North America, is capable of transmitting human malaria and is actively involved in the transmission of the ungulate malaria parasite *Plasmodium odocoilei*. However, molecular diagnostic tools based on Internal Transcribed Spacer 2 (ITS2) of ribosomal DNA are lacking for *An. punctipennis*. *An. punctipennis* is a former member of the *Anopheles maculipennis* complex but its systematic position remains unclear.

**Results:** In this study, we collected and sequenced ITS2 from 276 *An. punctipennis* specimens collected from the eastern and midwestern United States. Our analysis demonstrated a consistent sequence length and showed no indications of intragenomic variation among the samples based on ITS2, suggesting that *An. punctipennis* represents a single species in the studied geographic locations. In addition to the ITS2 sequence-based molecular tool, we developed a simple and robust Restriction Length Polymorphism approach for species identification. *An. punctipennis* was found in urban, rural, and forest settings, suggesting its potential broad role in pathogen transmission. Phylogeny based on ITS2 sequence comparison demonstrated the close relationship of this species with other members of the Maculipennis group.

**Conclusions:** This study developed molecular tools based on ITS2 sequences for the malaria vector *An. punctipennis* and clarified the phylogenetic position of the species within the Maculipennis group.

**Background**

Malaria is one of the most dangerous infectious diseases transmitted by mosquitoes [1]. Although the majority of malaria transmission occurs in Africa, in 2016 the World Health Organization (WHO) reported a 20% annual increase of malaria cases in 9 countries in North and South America [2]. Malaria was introduced to North America in the 17th century during the slave trade and, because of the presence of native competent malaria vectors, it rapidly spread through the entire continent [3]. In the United States (US) numerous outbreaks of malaria occurred from as far north as Massachusetts to subtropical southern states in the 18th and 19th centuries. Starting in the 1930s, intensive efforts to control malaria were undertaken and, in 1954, malaria was declared eliminated from the US [3]. Nevertheless, each year, the Center for Disease Control and Prevention reports between 1,000 and 1,500 cases of malaria in the US [4], with 63 outbreaks of locally transmitted malaria occurring since 1957 [5]. Climate change, human migration, political instability, and the presence of competent malaria vectors increase the risk of malaria and other tropical diseases being imported and transmitted in regions where they were eradicated [5].

Three species are considered dominant malaria vectors in North America – *An. freeboui*, *An. pseudopunctipennis*, and *An. quadrimaculatus* [6]. In addition to these species, *An. punctipennis*, Say, the most common and widely distributed anopheline mosquito species in North America [7], was an important contributor to malaria transmission in the US in the past [8]. Similar to *An. quadrimaculatus*, this species is capable of transmitting two major human malaria parasites, *Plasmodium vivax* and *P. falciparum* [9]. Moreover, *An. punctipennis* can also transmit *P. odocoilei*, which infects wild hooved animals, such as deer. A recent study conducted in 17 states across the US estimated a 25% *P. odocoilei* infection rate in white-tailed deer [10]. Malaria in wildlife populations may also have an impact on biodiversity, as infection reduces the deer population by killing fawns [11]. This can also present problems for humans due to the possibility of zoonotic transmission to livestock ungulates vectored from wild ungulates [11]. The ability of *An. punctipennis* to transmit arboviruses such as Cache Valley [12] and eastern equine encephalomyelitis [13] viruses, as well as a nematode *Dirofilaria immitis* [14] has also been reported.

Continuous study of mosquitoes in the field requires rapid and robust identification of the species. The availability of PCR-based molecular technologies enable the use of ribosomal DNA (rDNA) for species diagnostics [15]. rDNA genes are organized in arrays of multiple copies in eukaryotic genomes and undergo “concerted evolution,” which leads to unification of gene and intergenic sequences [16]. Unlike Intragenic Spacer and Intergenic Transcribed Spacer 1, Intergenic Transcribed Spacer 2 (ITS2) is a reliable tool for species diagnostics in the *Anopheles* genus [15]. ITS2 sequences have been utilized for species diagnostics in the *An. gambiae* complex [17], the *An. funestus* group [18], and the *An. bancroftii* group [19]. One nucleotide difference in ITS2 led to the discovery of the M and S forms of *An. gambiae* [17], which have now been elevated to species status as *An. gambiae* (former S form) and *An. coluzzii* (former M form) [20]. Three new Eurasian species – *An. artemiev* [21], *An. persiensis* [22], and *An. daciae* [23] — were described based on differences in their ITS2 sequences. ITS2 sequences have been successfully used for the identification of North American mosquitoes, including a newly identified species, *An. hemisi* [24], and four new species in the Quadrimaculatus complex [25]. More recently ITS2-based diagnostic tools were developed for species from the *An. crucians* complex [26]. Despite the wide distribution in North America and its obvious role in pathogen transmission, a molecular diagnostic tool using ITS2 from *An. punctipennis* has not been developed yet.

In the past, *An. punctipennis* was a member of the former *An. maculipennis* complex, which included 7 Eurasian and 6 North American species [27]. Although considered distant by the classical morphological approach, *An. punctipennis* was placed within the *An. maculipennis* complex based on interspecies hybridization with other members of the complex and the structure of its polytene chromosomes [28]. R. Harbach, in 2004, proposed a new classification for the genus *Anopheles* that was largely based on morphological characteristics and ITS2 sequence analyses [29]. According to this classification, the *An. maculipennis* complex was transformed to the Maculipennis group, which was subdivided into two North American subgroups, Freebomi and Quadrimaculatus, and one Eurasian Maculipennis subgroup. *An. punctipennis* was excluded from the Maculipennis group and placed into a separate Punctipennis group that included *An. punctipennis*, *An. perplexens*, and the *An. crucians* complex.

In this study, we collected malaria mosquitoes from 10 sites in eastern (Virginia and Florida) and midwestern states (Iowa and Minnesota) and developed 2 assays based on the sequencing of ITS2 of the rDNA and Restriction Fragment Length Polymorphism (RFLP) of ITS2 for molecular diagnostics of *An. punctipennis*. In addition, ITS2 sequences were used to clarify the phylogenetic relationships of *An. punctipennis*. 

**Results**
Internal Transcribed Spacer 2 sequencing

The primary goal of this study was to develop a molecular tool based on ITS2 for the identification of *An. punctipennis* that has been proven to be the most efficient source for species identification for the *Anopheles* genus [15]. To accomplish this goal, mosquitoes were collected from multiple locations in several eastern and midwestern states in the US (Table 1, Fig. 1). A total of 494 *Anopheles* samples, which included 276 *An. punctipennis* samples, were successfully sequenced. Species were identified by comparing the obtained sequences with sequences available in Gene Bank using NCBI BLAST [30]. Because ITS2 sequences are currently unavailable for *An. punctipennis*, this species was identified using additional sequencing of the mitochondrial cytochrome c oxidase subunit I (COI) gene; samples were identified as *An. punctipennis* if they had ≥ 99% identity with *An. punctipennis* COI sequences from Gene Bank [30]. This allowed us to match the ITS2 sequence with the COI sequence from the same individual. In addition, samples from Iowa were analyzed morphologically as *An. punctipennis* at the adult stage [31].

The ITS2 sequences for *An. punctipennis* were unknown prior to this research. The sequence samples were then analyzed in Sequencer (Gene Codes Corporation, Ann Arbor MI) to verify their quality. This allowed searching for ambiguous nucleotides as well as identifying the primers. Once analyzed, the ITS2 sequences obtained from forward and reverse primers from the same individual mosquito were assembled into combined sequences. Primer sequences from the 5.8S 3’ and the 28S 5’ ends were trimmed from the sequence. ITS2 sequences of sampled individuals of *An. punctipennis* from various locations showed no intraspecific differences among the samples tested. All of the samples had a consistent sequence length and there was no indication of interindividual variation among them based on ITS2.

Restriction Fragment Length Polymorphism approach

In addition to ITS2 sequencing, we developed a Restriction Fragment Length Polymorphism (RFLP) approach for the rapid identification of *An. punctipennis* among the other species that can be simultaneously present in the same field collections. The sizes of PCR products of ITS2 for the species that were obtained in the mosquito collections were the following: *An. bradleyi* 391 bp, *An. crucians* A 644 bp, *An. crucians* C 387 bp, *An. crucians* D 476 bp, *An. crucians* E, 378 bp, *An. punctipennis* 451 bp, and *An. quadrimaculatus* 470 bp. Thus, based on PCR product length, it is hard to distinguish *An. punctipennis* from *An. quadrimaculatus* and *An. crucians* D. Based on the restriction enzyme map of ITS2 sequences, we chose a restriction enzyme, Nael, that cleaves the PCR product of *An. punctipennis*, resulting in the appearance of two fragments of 168 bp and 283 bp. There are no sites for this restriction enzyme in the ITS2 sequences of any of the other species that were found together in all mosquito collections (Table 1) and other species included in this study (Table 2). We tested this approach on 50 samples from the various sample locations. We tested 10 samples from 3 locations in Virginia, 9 samples each from the Defuniak Springs location in Florida and the Des Moines location in Iowa, and 3 samples from the Minnetrista location in Minnesota. The overall analysis of the 50 samples of *An. punctipennis* taken from the local collections and evaluated by the RFLP-ITS2 assay indicated that 100% of the samples displayed a double band. We then setup up and ran the assay with all the collected Anopheline species as depicted in Fig. 2. As shown in Fig. 2, *An. punctipennis* is clearly the only species collected that displays this double band result, indicating that this method is reliable for the identification of *An. punctipennis*.

Anopheles species distribution in the eastern and midwestern US

In addition to the development of a tool for molecular diagnosis of *An. punctipennis*, the species composition was analyzed in all locations from the eastern and midwestern United States (Table 1, Fig. 1). In the Midwest, *An. punctipennis* was found in Minnetrista, Minnesota where ~8% of the specimens were represented by *An. punctipennis* and nearly 92% were *An. quadrimaculatus*. This location is indicated on the map as KW (Fig. 1A). Additional samples from the Midwest were collected in Iowa near Des Moines (IO). In contrast, this collection included 94% *An. punctipennis* and 6% *An. quadrimaculatus*.

Another collection set was made from mosquitoes collected in the Appalachian Mountains in Virginia, with elevations ranging from 600 m to 1.2 km (Fig. 1B). Collections in Virginia were made in 3 different environments: town, village, and forest. The town collection was made at the Duck Pond (DP) on the Virginia Tech campus, which is in the center of the small college town of Blacksburg, Virginia. In this setting, 94% of the mosquitoes were *An. punctipennis* with a high population density. The village collection was made in a rural setting in nearby Pandalas Pond (PP) in Montgomery County, Virginia. This location contained farms and plenty of wildlife. The proportion of *An. punctipennis* in this location reached 65% of the samples collected. The rest was represented by 22% *An. quadrimaculatus*, 3.7% *An. crucians* C, and 5.6% each of *An. crucians* D and *An. crucians* E. The third collection was made in a forest location at the high elevation of 1.2 km in Giles County, Virginia near Mountain Lake (MM). In this location, 100% of the mosquitoes were *An. punctipennis*, again with a much higher population density than in rural location.

The next set of collections were made in Florida along the coast. Mosquitoes were also collected in Western Florida near Panama City Beach and in Eastern Florida near Orlando in urban, rural, and forest locations. *An. punctipennis* was found in only one urban location in Defuniak Springs, Florida in Bruce Creek (BC). The collection consisted of approximately 73% *An. punctipennis*, 23% *An. crucians* A, 2 % *An. crucians* C, and 2% *An. crucians* E. In the coastal areas, the prevalent mosquitoes were from the *An. crucians* complex. In the mineral springs in Panacea, Florida (PF) the sample collection included 92% *An. bradleyi* and 8% *An. quadrimaculatus*. *An. bradleyi* was previously considered a brackish water mosquito [31], which is consistent with the current observation. In the more urban setting along the coast, Carrabelle, Florida, approximately 39% *An. crucians* C and approximately 61% *An. crucians* D were found. Finally, in the forest settings of Apalachicolona National Forest in Apalachicolona, Florida (SR), a remote setting with few humans, 100% of the collection was *An. crucians* D.

Phylogenetic analysis based on Internal Transcribed Spacer 2

Phylogenetic analyses based on ITS2 sequences using the maximum-likelihood approach [32] allowed us to construct a reliable tree of the major phylogenetic subgroups of the studied species (Fig 3). The analysis included 12 species: *An. punctipennis*; 4 North American species from the Maculipennis group (*An. freeborni, An. hermsi, An. occidentalis, and An. quadrimaculatus*); 5 Eurasian species (*An. beklemishevi, An. martinius, An. messeae, An. persiensis, and An. sacharovi*), and an outgroup Asian species, *An. sinensis*. Overall, the tree was separated into two large groups corresponding to species from the Maculipennis
shown in Table 1. The collections from Iowa were provided by colleagues from Iowa State University. Global positions of the mosquito collection sites are

A collection of approximately 500 Anopheles mosquitoes were sampled from multiple locations across four states during the summers of 2017 and 2018 as

Mosquito collection samples suggesting that material from the eastern and midwestern US. Sequence analysis demonstrated no variations in length or nucleotide polymorphism within the studied

North America. Two versions of the tool, based on either sequence analysis or a cost-effective RFLP approach, were optimized based on field collected

This study reports the development of a new molecular diagnostic tool for malaria mosquitoes [15], had never been developed for An. punctipennis prior to this study. Accurate identification of mosquitoes is important for proper surveillance, control, and diagnostic measures and is pivotal for genomic research, identification of vector competence, and identification of insecticide resistant genes associated with the species.

In this study, ITS2 fragments of rDNA were sequenced from 494 field collected mosquito samples in the eastern and midwestern US. Among them, 276 samples were identified as An. punctipennis based on morphological characteristics and COI sequences available at NCBI [30]. All of the An. punctipennis samples had a consistent ITS2 sequence length and there was no indication of interindividual variations. Thus, our study developed a molecular diagnostic for An. punctipennis based on ITS2 sequence and determined that An. punctipennis represents a single species in the eastern and midwestern states in the US. However, samples from California and other Western states were not included in this study, which requires further attention in the future.

In addition, the RFLP approach based on ITS2 sequences was optimized for 7 species of malaria mosquitoes that were available through our mosquito collections (Table 1). The presence of a unique site in An. punctipennis ITS2 sequences for the restriction enzyme NaeI allowed to cleave the PCR product of the ITS2 into two fragments that could be easily identified by gel-electrophoresis. This method represents a simple and robust technique for identifying An. punctipennis among other local anopheline species. The approach also reduces the time and cost of molecular identification and is a robust method for species identification in field collected mosquitoes.

In this study, Anopheles mosquito collections were generated from samples collected at 11 locations within 4 states (Florida, Iowa, Minnesota, and Virginia) from urban, rural, and forest settings. An. punctipennis was found in 6 locations; only the coastal collection sites in Florida had no An. punctipennis samples (Fig. 1). Thus, our study showed that An. punctipennis is a very abundant species of Anopheles in the eastern and midwestern US. Despite the fact that An. punctipennis was considered as a woodland mosquito [31], this species was detected in all of the settings that were examined (i.e., urban, rural, or forest), suggesting a possible broad role of this mosquito in transmission of pathogens. Interestingly, our study determined the presence of An. quadrivittatus at high altitude in the Appalachian Mountains (Fig. 1), where the presence of this species had not previously been determined [31].

Finally, our study sheds some light on the phylogenetic relationships of An. punctipennis and its systematic position within other species from the genus Anopheles. An. punctipennis was a former member of the An. maculipennis complex [27] but later was excluded from the complex based on morphological characteristics. According to R. Harbach [29], An. punctipennis belongs to a separate Punctipennis group that includes An. punctipennis, An. perplexens, and the An. crucians complex. Because of the absence of ITS2 sequences, An. punctipennis was not included in the phylogenetic analysis that was conducted for other species from the Maculipennis group [33-35]. The only phylogeny that included An. punctipennis was based on sequencing the D2 region of its rDNA [36]. This analysis clearly separated the An. quadrimaculatus clade from the An. freeborni—An. occidentalis—An. earlei—An. punctipennis clade and suggested An. walkeri as the most basal species for the entire group. However, the branches within these two clades remained poorly resolved. In our study, phylogenetic analysis was performed for 12 species, including 10 species from the Maculipennis group, two species from the An. crucians complex and an outgroup species, An. sinensis. The maximum-likelihood tree clearly separated the Maculipennis species from the An. crucians complex species and placed An. punctipennis as a sister taxon to An. quadrimaculatus. Thus, the phylogenetic analysis based on ITS2 sequences allowed us to reliably identify that An. punctipennis represents a species within the Maculipennis group and is a close relative to An. quadrimaculatus. Further analysis based on multiple genes or the entire genome will provide more details about the phylogenetic relationships of An. punctipennis and clarify its systematic position within the Maculipennis group of malaria mosquitoes.

Conclusions

This study reports the development of a new molecular diagnostic tool for An. punctipennis, one of the most widely spread species of malaria mosquitoes in North America. Two versions of the tool, based on either sequence analysis or a cost-effective RFLP approach, were optimized based on field collected material from the eastern and midwestern US. Sequence analysis demonstrated no variations in length or nucleotide polymorphism within the studied samples suggesting that An. punctipennis represents a single species in the area of study. The phylogenetic analysis based on ITS2 sequences placed An. punctipennis within the Maculipennis group species and separate from species within the An. crucians complex.

Methods

Mosquito collection

A collection of approximately 500 Anopheles mosquitoes were sampled from multiple locations across four states during the summers of 2017 and 2018 as shown in Table 1. The collections from Iowa were provided by colleagues from Iowa State University. Global positions of the mosquito collection sites are
shown on a map using the program ArcGIS Pro [37] in Fig 1. Mosquitoes were collected from larval breeding sites, typically small ponds or bodies of water. These breeding sites were located near aquatic plants known as the filamentous algae. More specifically the plants among these breeding sites were identified as Cladophora sp., Spirogyra sp., and Pithophora sp. The pan dipping method was used to collect samples of Anopheles in their natural habitat. A pan was dipped under the water so that the surface water spilled into the pan. Once in the pan, a disposable pipette was implemented to draw the larvae from the water, which was then placed into a collection container. We used a 50 ml centrifuge tube for temporary storage while in the field. Once collected, the larvae were fixed in 70% ethanol. The mosquito larvae were individually labeled for identification and placed in 1.5 ml tubes and stored at -20°C.

DNA extraction

DNA extractions were conducted from mosquito larvae, except at the location in Iowa, where DNA were extracted from adults. DNA extractions were attempted for ~500 specimens. Each individual carcass was placed in a sterile 1.5 ml tube to prevent contamination. Extraction of genomic DNA from each specimen was conducted using the standard protocol for Qiagen DNeasy Blood and Tissue Kits (Qiagen, Germantown, MD, USA) with some modification. DNA elution was performed with 50 ml of molecular grade water (Research Products International, Mt. Prospect, IL, USA). Purified DNA was checked using a Nanodrop spectrometer (Thermo Fisher Scientific, Haverhill, MA, USA) to determine quantity and quality and then stored at -20°C.

DNA amplification

For rDNA ITS2 amplification, we used the direct primer design by J. Proft [38] 5'-ATCACCTGGCTCTCGTGGATCG-3' (Tm = 71.8°C) and the reverse primer design by Y. Novikov 5'-ATGCTTAATTAGGGGTA-3' (Tm = 56.8°C) [39]. For the PCR reaction, we used Hot Start ImmoMix™ (Bioline, Taunton, MA, USA). PCR reaction mixtures were conducted using 3 ml of genomic DNA, 15 ml 2×ImmoMix, 1 ml of both forward and reverse primers at 10 µM concentration, and 10 ml of molecular grade water. PCR amplification was conducted using a thermal cycler (Eppendorf, Hauppauge, NY, USA) with the following parameters: 95°C for 10 minutes, followed by 30 cycles of 94°C for 15 seconds, 55°C for 30 seconds, and 72°C for 30 seconds, and the final step at 72°C for 5 minutes.

The mitochondrial cytochrome oxidase gene (COI) gene was amplified using the following primers that were designed using Primer 3 software: Forward 5'-TGAGACCTTTATTTAGGGGTA-3' (Tm = 64.0°C) and Reverse 5'-ATAGTAGAAATGGGGGCCGG-3' (Tm = 63.7°C). PCR amplification was conducted with the following parameters for COI: 95°C for 10 minutes, following by 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds, and a final step at 72°C for 7 minutes. PCR products were analyzed by electrophoresis in 1% agarose gel. GelRed® Nucleic Acid Gel Stain (Biotium, Fremont, CA, USA) was used to visualize the bands. For DNA sequencing, amplicons were purified with a Wizard™PCR Clean Up kit (Promega, Fitchburg, WI, USA). The product was stored at -20°C until processed for sequencing.

DNA sequencing, analysis, and submission

10 µl of PCR products were mixed with 3 ml forward or reverse primers at 3.2 mM concentration. The specimens were then sent to sequencing using the Sanger Platform at the BioComplexity Institute at Virginia Polytechnic Institute and State University.

A total of 494 samples were successfully sequenced. The raw sequences obtained from local collections were uploaded to Sequencher 5.3 software (Gene Codes Corp. Ann Arbor, MI) for further analysis. Sequences were analyzed for quality and the forward and reverse sequences were assembled to obtain the complete sequence of ITS2. In preparation for submission to NCBI ambiguous nucleotides and primer sequences were trimmed off the sequence [30]. A total of 113 sequences, which were sequenced from both ends, were combined using Sequencher 5.3 software and the sequence can be found in NCBI [30] with accession numbers MN460678-MN460726 and MN460728-MN460791.

DNA restriction

To obtain DNA fragments with different lengths for An. punctipennis, 4 ml of the ITS2 PCR product, 2 ml of the 10× CutSmart® Buffer, 5 units (0.5 µl) of NaeI restriction enzyme (New England BioLabs, INC., Ipswich, MA, USA), and 18 ml of ddH₂O were added to a sterile PCR tube. The tubes were then placed in a thermocycler at 37°C for 60 minutes. PCR products were analyzed by electrophoresis in 3% agarose gel. GelRed® Nucleic Acid Gel Stain was used to visualize the bands (Fig. 2). Lane 1 in all gels was loaded with a 1 kb DNA ladder (Millipore Sigma, Sigma Aldrich USA).

Phylogenetic analysis

We downloaded ITS2 sequences representing 12 Anopheline species. The sequences were aligned using the webPRANK server [40] and then the model of nucleotide substitutions was estimated using jModelTest 2 software [41] and Bayesian information criterion. The maximum-likelihood tree was constructed using MEGA X [32] software with the HKY-G model and tested using 500 bootstrap-replications. All sites with gaps or missing data were excluded from the alignment during tree construction. Branches with bootstrap support less than 50% were collapsed.

Abbreviations

DNA: deoxyribonucleic acid; ITS2: internal transcribe spacer 2; PCR: polymerase chain reaction; RFLP: restriction fragment length polymorphism; COI: cytochrome c oxidase subunit I; rDNA: ribosomal DNA; NCBI: National Center for Biotechnology Information; BLAST: Basic Logical Alignment Search Tool; US: United States.

Declarations

Ethics approval and consent to participate
Availability of data and material

ITS2 sequencing data will be available via NCBI upon the publication of the manuscript.

Competing interests

Authors declare no competing interests.

Consent of publication

This manuscript does not contain any individualized data.

Funding

The research was supported by US Department of Agriculture, National Institute of Food and Agriculture Project No. VA-160058 and Russian State Budget Project #0324-2019-0042.

Authors’ Contributions

MVS conceived and designed the experiments. JMH, MVS, MIG, and RCS collected mosquitoes. JMH, AAY, and RAM performed the experiments. JMH, AAY, DAK, and MVS conducted data analysis and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The study was performed using US Department of Agriculture, National Institute of Food and Agriculture Project No. VA-160058 and the Russian State Budget Project #0324-2019-0042. The funding agencies had no role in the design of the study or in the collection, analysis, and interpretation of data, and in the writing of the manuscript. We thank Janet Webster for proofreading the text.

Author details

1Department of Entomology and the Fralin Life Sciences Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA; 2Laboratory of Evolutionary Genomics of Insects, the Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia; 3Kurchatov Genomics Center, the Federal Research Center Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Science, Novosibirsk, Russia; 4Department of Entomology, Iowa State University, Ames, IA, USA; 5Department of General Biology and Ecology, Moscow Regional State University, Moscow, Russia.

References

1. Tolle MA: Mosquito-borne diseases. Curr Probl Pediatr Adolesc Health Care 2009, 39:97-140.
2. WHO: World Malaria Report 2017 2017:1-196.
3. Moreno-Madrinan MJ, Turell M: History of Mosquitoborne Diseases in the United States and Implications for New Pathogens. Emerg Infect Dis 2018, 24:821-826.
4. Robert LL, Santos-Ciminera PD, Andre RG, Schultz GW, Lawyer PG, Nigro J, Masuoka P, Wirtz RA, Neely J, Gaines D, et al: Plasmodium-infected Anopheles mosquitoes collected in Virginia and Maryland following local transmission of Plasmodium vivax malaria in Loudoun County, Virginia. J Am Mosq Control Assoc 2005, 21:187-193.
5. Rossati A, Bargiacchi O, Kroumova V, Zaramella M, Caputo A, Garavelli PL: Climate, environment and transmission of malaria. Infez Med 2016, 24:93-104.
6. Sinka ME, Rubio-Palis Y, Manguin S, Patil AP, Temperley WH, Gething PW, Van Boeckel T, Kabaria CW, Harbach RE, Hay SI: The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic precis. Parasit Vectors 2010, 3:72.
7. Baker RH, Kitzmiller JB: Salivary Gland Chromosomes of Anopheles Punctipennis. J Hered 1964, 55:9-17.
8. Jensen T, Dritz DA, Fritz GN, Washino RK, Reeves WC: Lake Vera revisited: parity and survival rates of Anopheles punctipennis at the site of a malaria outbreak in the Sierra Nevada foothills of California. Am J Trop Med Hyg 1998, 59:591-594.
9. Strickman D, Gaffigan T, Wirtz RA, Benedict MQ, Rafferty CS, Barwick RS, Williams HA: Mosquito collections following local transmission of Plasmodium falciparum malaria in Westmoreland County, Virginia. J Am Mosq Control Assoc 2000, 16:219-222.
10. Martensens ES, McNerney N, Brightman H, Ferebee K, Walsh T, McShea WJ, Forrester TD, Ware L, Joyner PH, Perkins SL, et al: Hidden in plain sight: Cryptic and endemic malaria parasites in North American white-tailed deer (Odocoileus virginianus). Sci Adv 2016, 2:e1501486.
11. Guggisberg AM, Sayler KA, Wisely SM, Odom John AR: Natural History of Plasmodium odocoilei Malaria Infection in Farmed White-Tailed Deer. mSphere 2018, 3.
12. Anderson JF, Armstrong PM, Misencik MJ, Bransfield AB, Andreadis TG, Molaei G: Seasonal Distribution, Blood-Feeding Habits, and Viruses of Mosquitoes in an Open-Faced Quarry in Connecticut, 2010 and 2011. J Am Mosq Control Assoc 2018, 34:1-10.
13. Shepard JJ, Andreadis TG, Thomas MC, Molaei G: Host associations of mosquitoes at eastern equine encephalitis virus foci in Connecticut, USA. Parasit Vectors 2016, 9:474.
Table 1. Mosquito collections from the eastern and midwestern United States.
| State | County | Site location | Latitude  | Longitude | Number | An. punctipennis | An. quadrimaculatus s.s. | An. bradley A | An. crucians A | An. crucius C |
|-------|--------|---------------|-----------|-----------|--------|-----------------|--------------------------|--------------|---------------|---------------|
| MN    | Hennepin | Minnetrista   | 44.95972  | -93.709693| 38     | 3               | 35                       | 0            | 0             | 0             |
| IA    | Polk   | Des Moines    | 41.53825  | -93.57757 | 50     | 47              | 3                       | 0            | 0             | 0             |
| FL    | Wakulla| Panacea       | 30.0347   | -84.3894  | 12     | 0               | 1                       | 11           | 0             | 0             |
|      | Walton | DeFuniak Springs | 30.6887   | -86.0945  | 60     | 46              | 0                       | 0            | 12            | 1             |
|      |        | Greenway Trail| 30.3708   | -86.1807  | 49     | 0               | 0                       | 0            | 0             | 14            |
| IA    | Franklin| Carrabelle    | 29.9151   | -84.524   | 46     | 0               | 0                       | 0            | 0             | 18            |
|      |        | Appalachiania | 29.973512 | -84.61206 | 12     | 0               | 0                       | 0            | 0             | 0             |
| FL    | Osceola| Melbourne     | 28.136    | -80.897   | 25     | 0               | 18                      | 0            | 0             | 4             |
| VA    | Montgomery| Pandapas Pond| 37.282164 | -80.465947| 102    | 83              | 12                      | 0            | 2             | 2             |
|      |        | Duck Pond     | 37.22592  | -80.427208| 50     | 47              | 3                       | 0            | 0             | 0             |
| VA    | Giles | Mountain Lake | 37.356167 | -80.536283| 50     | 50              | 0                       | 0            | 0             | 0             |
|       |        |               |           |           |        |                 |                          |              |               |               |
| **Total** |       |               |           |           | 494    | 276             | 72                      | 11           | 14            | 39            |

Table 2. Accession numbers and corresponding species names of the ITS2 sequences used for the phylogenetic analysis.

| Accession number | Species                      |
|------------------|------------------------------|
| MG028992.1       | An. sinensis                 |
| MN460678-N460726, MN460728-MN460791 | An. punctipennis |
| AY593958         | An. beklemishivi             |
| M64484           | An. freeborni                |
| M64483           | An. hermsi                   |
| M64482           | An. occidentalis             |
| AF504204         | An. messae                   |
| FN665794.1       | An. persiensis               |
| AJ849885.1       | An. martinius                |
| EU346655.1       | An. sacharovi                |
| DQ662410.1       | An. sacharovi                |
| AQU32550         | An. quadrimaculatus          |
| AQU32503         | An. smaragdus (An. quadrimaculatus B) |
| AQU32504         | An. diluvialis (An. quadrimaculatus C1) |
| AQU32505         | An. inundatus (An. quadrimaculatus C2) |
| AQU32506         | An. maverlius (An. quadrimaculatus D) |
| AY245553         | An. cricians A               |
| AY386966         | An. cricians E               |
| AY386967         | An. bradley                  |
| AY386965         | An. cricians C               |
| AY386964         | An. cricians D               |
Species composition in mosquito collections in eastern and midwestern United States. Species distribution [31] is shown on the map by different shades (A). The map was developed using the program ArcGis Pro [36]. Species abundance is shown by different colors in pie charts. Details of the site locations and species compositions in Virginia and Florida are shown in panels B and C, respectively. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
Figure 2

Restriction fragment length polymorphism approach for detecting *An. punctipennis*. Results of gel electrophoresis of the ITS2 PCR product digested with restriction enzyme *NaeI* is shown for the following species: 1) *An. crucians A*; 2) *An. crucians C*; 3) *An. crucians D*; 4) *An. crucians E*; 5) *An. bradley*; 6) *An. quadrimaculatus*; and 7) *An. punctipennis*. Column M represents a 1 kb DNA ladder. Only the *An. punctipennis* ITS2 PCR product was cleaved by the enzyme resulting in the appearance of two bands.
Figure 3

Maximum-likelihood tree based on the ITS2 sequences. Bootstrap values are shown for the branches with support higher than 50%. Branches with lower values were collapsed. The phylogenetic analysis placed An. punctipennis within the Maculipennis group species.