

RESEARCH ARTICLE

Proteomic fingerprinting of Neotropical hard tick species (*Acari: Ixodidae*) using a self-curated mass spectra reference library

Rolando A. Gittens^{1,2}, Alejandro Almanza¹, Kelly L. Bennett^{1,3}, Luis C. Mejía^{1,3}, Javier E. Sanchez-Galan^{1,4}, Fernando Merchan⁵, Jonathan Kern^{5,6}, Matthew J. Miller^{7,8}, Helen J. Esser⁹, Robert Hwang¹⁰, May Dong¹⁰, Luis F. De León^{1,11}, Eric Álvarez¹², Jose R. Loaiza^{1,3,12*}

1 Centro de Biodiversidad y Descubrimiento de Drogas, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT AIP), Panama, Republic of Panama, **2** Centro de Neurociencias, INDICASAT AIP, Panama, Republic of Panama, **3** Smithsonian Tropical Research Institute, Panama, Republic of Panama, **4** Grupo de Investigación en Biotecnología, Bioinformática y Biología de Sistemas, Facultad de Ingeniería de Sistemas Computacionales, Universidad Tecnológica de Panamá, Panama, Republic of Panama, **5** Grupo de Investigación en Sistemas de Comunicaciones Digitales Avanzados, Facultad de Ingeniería Eléctrica, Universidad Tecnológica de Panamá, Panama, Republic of Panama, **6** ENSEIRB-MATMECA—Bordeaux INP, France, **7** Department of Anthropology, Pennsylvania State University, University Park, PA, United States of America, **8** University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, AK, United States of America, **9** Department of Environmental Sciences, Wageningen University, Wageningen, the Netherlands, **10** Department of Biology, Swarthmore College, Swarthmore, PA, United States of America, **11** Department of Biology, University of Massachusetts Boston, Boston, MA, United States of America, **12** Programa Centroamericano de Maestría en Entomología, Universidad de Panamá, Panama, Republic of Panama

* jloaiza@indicasat.org.pa



OPEN ACCESS

Citation: Gittens RA, Almanza A, Bennett KL, Mejía LC, Sanchez-Galan JE, Merchan F, et al. (2020) Proteomic fingerprinting of Neotropical hard tick species (*Acari: Ixodidae*) using a self-curated mass spectra reference library. *PLoS Negl Trop Dis* 14(10): e0008849. <https://doi.org/10.1371/journal.pntd.0008849>

Editor: Job E. Lopez, Baylor College of Medicine, UNITED STATES

Received: May 27, 2020

Accepted: October 2, 2020

Published: October 27, 2020

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: Financial support for this work was provided by SENACYT through the research grant GRID15-002 to JRL, LM, JSG, LFD and RAG. INDICASAT-AIP, UTP and STRI provided additional economic and logistic support. The SNI supports research activities by JRL (SNI 05-2016 & SNI 157-2017), JSG, LM, LFD, FM and RAG (SNI 91-

Abstract

Matrix-assisted laser desorption/ionization (MALDI) time-of-flight mass spectrometry is an analytical method that detects macromolecules that can be used for proteomic fingerprinting and taxonomic identification in arthropods. The conventional MALDI approach uses fresh laboratory-reared arthropod specimens to build a reference mass spectra library with high-quality standards required to achieve reliable identification. However, this may not be possible to accomplish in some arthropod groups that are difficult to rear under laboratory conditions, or for which only alcohol preserved samples are available. Here, we generated MALDI mass spectra of highly abundant proteins from the legs of 18 Neotropical species of adult field-collected hard ticks, several of which had not been analyzed by mass spectrometry before. We then used their mass spectra as fingerprints to identify each tick species by applying machine learning and pattern recognition algorithms that combined unsupervised and supervised clustering approaches. Both Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) classification algorithms were able to identify spectra from different tick species, with LDA achieving the best performance when applied to field-collected specimens that did have an existing entry in a reference library of arthropod protein spectra. These findings contribute to the growing literature that ascertains mass spectrometry as a rapid and effective method to complement other well-established techniques for

2015 & SNI 146-2017). RAG is also supported by SENACYT grants FID14-066, ITE15-016. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

taxonomic identification of disease vectors, which is the first step to predict and manage arthropod-borne pathogens.

Author summary

Hard ticks (Ixodidae) are external parasites that feed on the blood of almost every species of terrestrial vertebrate on earth, including humans. Due to a complete dependency on blood, both sexes and even immature stages, are capable of transmitting disease agents to their hosts, causing distress and sometimes death. Despite the public health significance of ixodid ticks, accurate species identification remains problematic. Vector species identification is core to developing effective vector control schemes. Herein, we provide the first report of MALDI identification of several species of field-collected Neotropical tick specimens preserved in ethanol for up to five years. Our methodology shows that identification does not depend on a commercial reference library of lab-reared samples, but with the help of machine learning it can rely on a self-curated reference library. In addition, our approach offers a complimentary tool with good accuracy and lower cost per sample than conventional and modern identification approaches such as morphology and molecular barcoding, provided that financial resources are available to invest in the acquisition of the MALDI equipment.

Introduction

Hard ticks (Ixodidae) are hematophagous ectoparasites that feed on almost every species of terrestrial vertebrate on earth, including *Homo sapiens sapiens* [1, 2]. Due to a complete dependency on blood as a food source, both sexes of adults and immature ticks are capable of transmitting disease pathogens to their hosts, causing significant morbidity and sometimes even death [3, 4]. Research on hard ticks has increased recently in the Neotropics, where a growing number of outbreaks of tick-borne related illnesses have been documented [5–8]. Despite these efforts, comprehensive studies about the ecology, behavior and control of hard ticks relevant to public health remain elusive in Central America due to the shortcomings of traditional taxonomic methods for species identification [9]. Taxonomic identification of Neotropical Ixodidae has traditionally relied on adult morphological characters [10]; however, morphological keys for immature stages (i.e., larvae and nymphs) are lacking and experts are often unable to reliably identify immature ticks to species [10, 11]. Moreover, morphological identification of ticks is unrealistic in epidemiological settings because assessing the role of ticks as disease vectors usually involves identifying hundreds of individuals for pathogen screening, an extremely time-consuming effort, which may be further impeded by the lack of qualified taxonomic specialists [12].

Matrix-assisted laser desorption/ionization (MALDI) time-of-flight mass spectrometry is an analytical technique that allows for sensitive and accurate detection of complex molecules such as proteins, peptides, lipids and nucleic acids [13–15]. The conventional MALDI approach has been used successfully for proteomic fingerprinting through pattern recognition for the identification of microorganisms such as pathogenic bacteria and fungi, which can be cultured in the laboratory and form discrete colonies with very consistent mass spectra that facilitates the development of reference libraries for identification of unknown samples [16, 17]. In fact, a commercial program offered by the manufacturers of the MALDI technology is

capable of determining statistical similarities between the spectra of unknown samples and a well-curated, proprietary reference library of bacteria and fungi to identify the species of the unknown specimen. This is analogous to the process of matching fingerprints, and offers a simplified comparison score that ranges from 0.0 to 3.0. Scores above or equal to 2.3 represent a confident match at the genus rank, and high probability at the species level, while values below 1.7 are considered as non-reliable identifications [16–18].

Although more challenging than identifying bacteria and fungi due to the size and heterogeneity of the specimen, MALDI has also been used to discriminate among species of invertebrates, including mosquitoes (Culicidae—*Anopheles*), fleas (Pulicidae—Ctenocephalide), biting midges (Ceratopogonidae—*Culicoides*), sandflies (Psychodidae—*Phlebotomus*, *Lutzomyia*) and ticks (Ixodidae—*Rhipicephalus*) [19–27]. A key finding from these studies is that protein spectra obtained from body sections or whole specimens were similar among individuals of the same morphological species but differed noticeably across different species. Therefore, MALDI protein spectra can be used as a tool to delimit species boundaries in arthropods that are vectors of pathogens. Nevertheless, fresh laboratory-reared specimens are routinely needed to build a reference library that meets the high-quality standards required for classification. This represents an important limitation for some arthropod groups, or assemblages, that are difficult to rear under laboratory conditions. In addition, epidemiological studies often rely on field-collected specimens preserved in ethanol for long-term storage in reference collections. To overcome these limitations, previous studies have opted for adjusting the comparison scores minimum-threshold limit for identification, lowering the manufacturer's recommended scores from 2.3 to 1.8 [22, 28] or even 1.3 [23, 29]. Hence, mass fingerprinting for the identification of field-collected specimens that do not exist in a reference spectra library (or for those from which reference spectra cannot be generated under ideal conditions) requires an alternative, objective approach [12]. Moreover, most existing applications of MALDI to identify arthropod disease vectors have focused on relatively species-poor vector assemblages from Europe. This technique has been tested less frequently in the new world tropics [20, 21, 23, 25, 28–37], where vector species richness is the greatest on Earth.

Here, we used MALDI as a scheme to identify Neotropical specimens of adult hard ticks derived from ethanol-preserved field collections. Specifically, we used machine learning and pattern recognition algorithms to classify protein spectra from the legs of field-collected specimens in order to identify a group of unknown samples with a self-curated reference library. MALDI is a promising tool for cataloging and quickly identifying large arthropod groups such as ticks [12]. Our results should contribute to the growing body of literature trying to address questions about feasibility, reliability and universality of the methodology for different environments and species that have not been evaluated before. Properly identifying disease vectors such as Ixodidae in highly diverse Neotropical countries, such as Panama, is a critical first step to predict and manage tick-borne zoonotic pathogens such as *Rickettsia* and arboviruses (i.e., arthropod-borne viruses).

Methods

Sample preparation

Ticks stored in ethanol for up to 5 years, and previously identified based on morphological characters, were taken from long-term storage in a -20°C freezer (S1 Table). A total of 103 specimens from the following species were included in this study: *Amblyomma mixtum* (*cajennense*), *Amblyomma calcaratum*, *Amblyomma dissimile*, *Amblyomma geayi*, *Amblyomma nodosum*, *Amblyomma oblongoguttatum*, *Amblyomma ovale*, *Amblyomma pecarium*, *Amblyomma sabanerae*, *Amblyomma varium*, *Amblyomma naponense*, *Amblyomma*

tapirellum, *Ixodes affinis*, *Ixodes boliviensis*, *Dermacentor nitens*, *Haemaphysalis juxtackochi*, *Rhipicephalus microplus* and *Rhipicephalus sanguineus*.

Samples were prepared following previously published protocols with minor modifications [22, 23]. Briefly, we removed either the left or the right anterior leg from each tick specimen using a scalpel. The leg was then put in a tube with 300 μ L ultrapure water followed by the addition of 900 μ L 100% ethanol. The tube was vortexed for 15 seconds and centrifuged using a Heraeus Biofuge Pico microcentrifuge (Thermo Fisher Scientific, Waltham, MA, USA) at 17,000 g for 2 minutes. After centrifugation, the supernatant was poured off from the sample tube, which was left to dry for 15 minutes. Subsequently, the leg was resuspended in 60 μ L 70% formic acid and 60 μ L 100% acetonitrile and homogenized in the microtube using a manual pestle. The sample was placed in a Branson 1510 ultra-sonicator (Branson, Danbury, CT, USA) for 60 minutes in ice water, and then vortexed for 15 seconds and centrifuged again at 17,000 g for 2 minutes.

For peptide detection with mass spectrometry, a saturated solution (10 mg/mL) of α -cyano-4-hydroxycinnamic acid (HCCA) matrix was prepared in 30:70 [v/v] acetonitrile: 0.1% trifluoroacetic acid (TFA) in water. An aliquot of 1 μ L from the sample supernatant was pre-mixed with an equal volume of HCCA matrix, and 1 μ L of the mix was quickly pipetted onto a polished steel MALDI plate in its respective target spot. All samples were placed and measured on three individual target spots with spectra from three technical replicates collected per spot. After letting the plate dry, it was inserted into the MALDI mass spectrometer to record the protein spectra from the tick's leg.

MALDI mass spectrometry parameters

We used an UltrafleXtreme spectrometer (Bruker Daltonics, Bremen, Germany) to generate the protein mass spectra of each specimen. The equipment has a MALDI source, a time-of-flight (TOF) mass analyzer, and a 2 kHz Smartbeam-II neodymium-doped yttrium aluminum garnet (Nd:YAG) solid-state laser ($\lambda = 355$ nm) that we used in positive polarization mode. All spectra were automatically acquired in the range of 2,000 to 20,000 m/z in linear mode for the detection of the most abundant protein ions. Each spectrum represented the accumulation of 5,100 shots with 300 shots taken at a time, and the acquisition was done in random-walk mode with a laser power in the range of 50% to 100% (global laser attenuation at 30%).

The software FlexAnalysis (Bruker) was used to pre-process and evaluate the mass spectra quality, based on the number of ion peaks and their intensity. Initially, all sample spectra were normalized by applying a general algorithm for baseline subtraction and smoothing provided by the software. Visual comparisons of the mass spectra from different tick species gave initial indications of dominant ion peaks that would suggest possible classification into discrete groups. Mass spectra that did not include at least one ion peak with an intensity of 1000 a.u. or more, were considered low quality and filtered out. All samples were placed and measured on three individual target spots, with three technical replicates of the mass spectra collected per spot.

Data analysis, clustering algorithms and statistics

The methodology has been described in detail previously by our group for the identification of adult mosquito legs [27], based on similar data analysis for face recognition [38, 39] and spectral classification using mass spectrometry [40, 41]. In brief, 239 mass spectra generated across 103 samples for all 18 species of morphologically identified Neotropical hard ticks were classified with a custom-made algorithm developed by our group using MATLAB (MathWorks, Natick, MA, USA). The algorithm is based on Principal Component Analysis (PCA) and

Linear Discriminant Analysis (LDA), which are linear transformation techniques from the field of Machine Learning that are commonly used for dimensionality reduction and classification. Dimensionality reduction can help decrease computational costs for classification, as well as avoid overfitting by minimizing the error in parameter estimation. Overfitting was also addressed by maximizing the number of specimens analyzed per species, while minimizing the number of technical replicas (i.e., only three spectra per specimen with good signal intensity were used for data analysis).

PCA is an “unsupervised” algorithm that generates vectors that correspond to the direction of maximal variance in the sample space. On the other hand, LDA is a “supervised” algorithm that considers class information to provide a basis that best discriminates the classes (i.e., tick species) [38]. For both PCA and LDA analyses, we calculated the Euclidean distance between the vector describing the test sample and the average vector describing each class to identify a test sample. The class with the minimum distance with respect to the test sample was assigned as the identified species for that test sample. The LDA was applied over the data set expressed in terms of the coefficients (i.e., principal components) obtained by the PCA. Thus, PCA reduced the dimensionality of the data, and the LDA provided the supervised classification.

The performance of the clustering algorithms was tested using Monte Carlo simulations over 1000 iterations per species to optimize training and cross-validation prediction success rates. For each iteration, the data elements in each class were split randomly in approximately, but not less than, 20% of the elements for testing and the rest of the elements for training, for each species. We used all the peaks in the spectra for the PCA analysis, and the first 150 principal components from the PCA stage that explained 99.9% of the total variance were then projected for the LDA algorithm, which also generated a 150-components data set. The number of components was chosen after a performance analysis, again using a Monte Carlo approach, that provided the best identification rates. Global and class positive identification rates were calculated to establish the classification capacity of the algorithm. The positive identification rate corresponds to the percent ratio between positive identifications performed by the algorithm and the real positive cases in the data.

For visualization purposes in the plots generated with our algorithm in MATLAB, species that were morphologically identified within the *Rhipicephalus* and *Ixodes* genera were separately compared against *Dermacentor* and *Haemaphysalis* for which there was only one species in each. All species that were morphologically identified within the *Amblyomma* genus were separately compared between themselves or against the *Ixodes* genera.

Results

Optical micrographs from 18 species of Neotropical hard ticks showed evident differences among species in terms of adult morphological features (Fig 1), which was well aligned with the expected unique mass spectra generated from each sample and taxon (Fig 2, S1 Fig, S2 Fig and S3 Fig). The global automatic acquisition rate was 77% for all species (Table 1), confirming that, overall, the mass spectra of field-collected and ethanol-preserved specimens allowed automatic acquisition of spectra. In fact, automatic acquisition of spectra results in faster and more objective data acquisition than performing spectra collection manually. However, the automatic spectra collection, coupled to the fact that species had different starting number of specimens, meant that the number of spectra per species for data analysis was not the same and, in some cases, did not meet the expected number of spectra per specimen (Table 1). Still, this was not an obstacle for our data analysis clustering algorithm. The percentage of automatic spectra acquisition with the MALDI ranged from 50% for *A. mixtum* (*cajennense*), *I. boliviensis* and *R. sanguineus* to 100% for several of the species, including *A. calcaratum*, *A. geayi*, *A. sabanerae*,

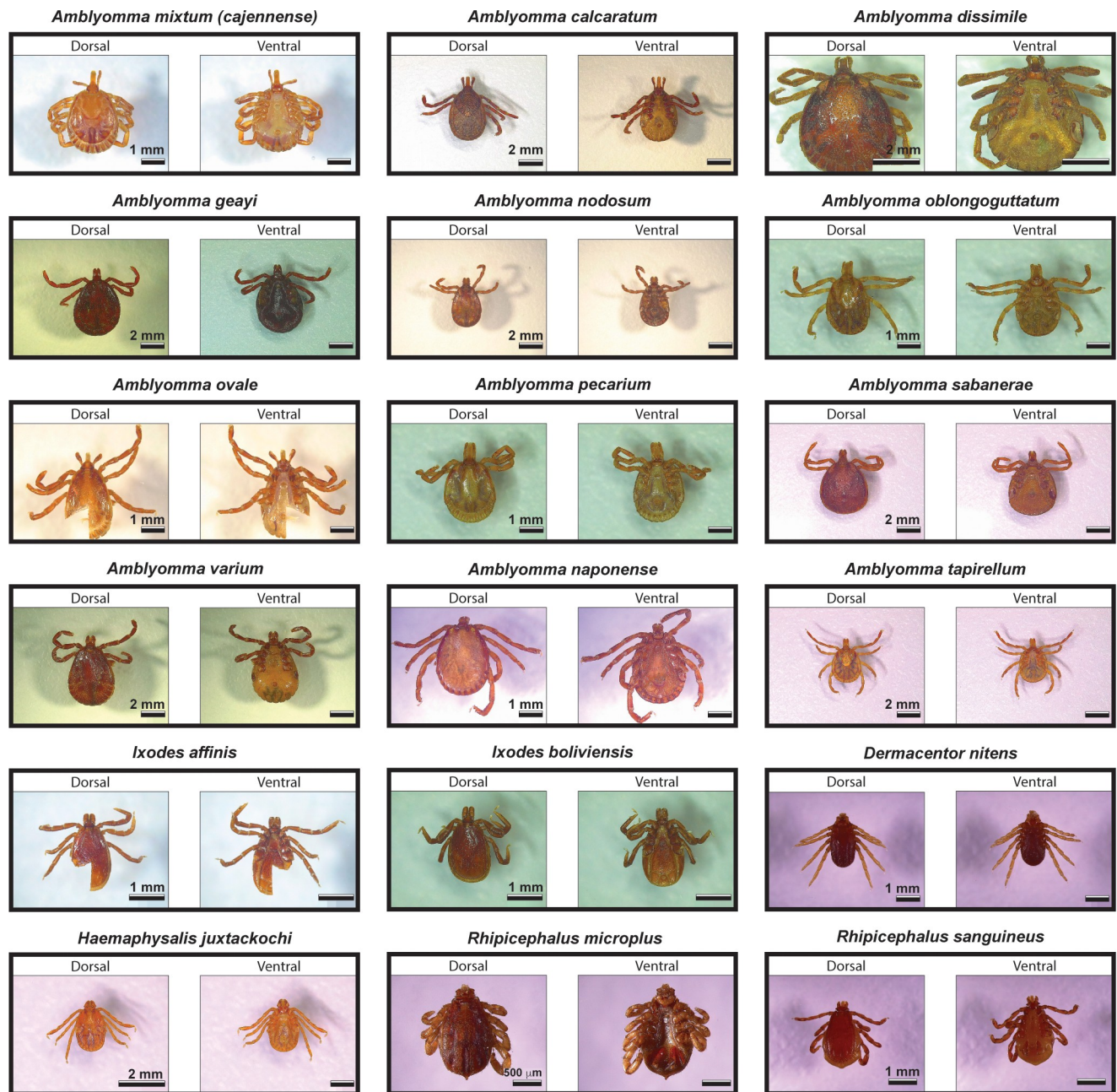


Fig 1. Optical micrographs of Neotropical hard ticks. The image shows the dorsal and ventral sides for all 18 species of hard ticks in the genera *Amblyomma*, *Dermacentor*, *Haemaphysalis*, *Ixodes*, and *Rhipicephalus* used to generate protein spectra with our MALDI mass spectrometry approach.

<https://doi.org/10.1371/journal.pntd.0008849.g001>

I. affinis, and *R. microplus*, covering a range from 6 to 56 spectra per species (Table 1). The time stored in ethanol or the location of sample origin did not seem to explain the variable percentages of automatic spectra collection (S1 Table). Spectra from freshly collected specimens stored dry at -20°C , used to establish the methodology, exhibited the best signals, with better-defined spectral peaks and higher signal-to-noise ratio.

In addition, the specimens within each species showed consistently similar protein profiles, regardless of their taxonomic genera, sex, collection date and/or sampling location (S1 Fig, S2 Fig, S3 Fig). Mean protein spectra for tick species differed visually among taxa and the

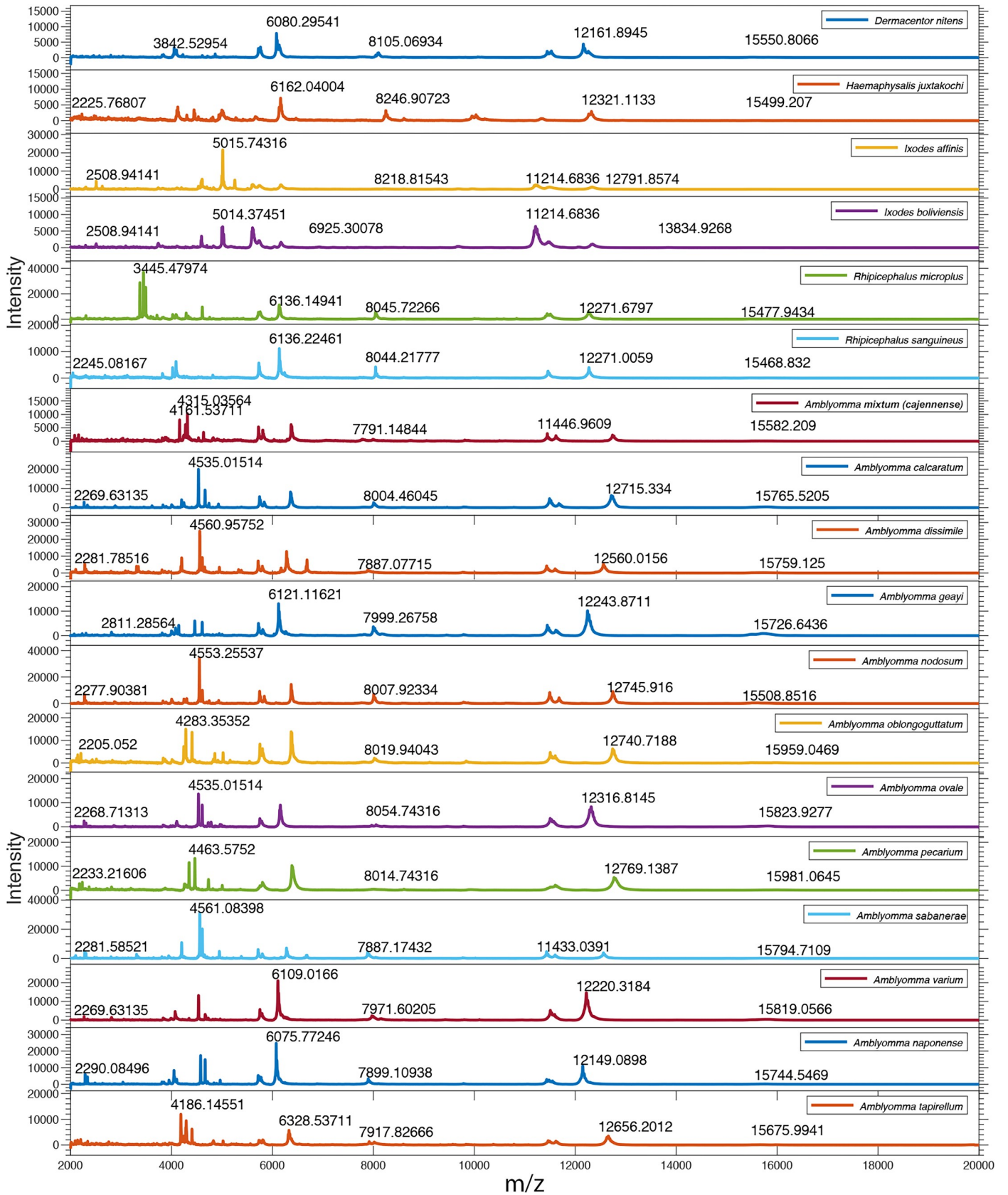


Fig 2. Baseline-corrected and smoothed spectra for 18 species of ticks in the genus *Amblyomma*, *Dermacentor*, *Haemaphysalis*, *Ixodes* and *Rhipicephalus*. Major ion peaks and their molecular weights are annotated in the range of 2,000 to 20,000 m/z for all species.

<https://doi.org/10.1371/journal.pntd.0008849.g002>

differences appeared to be related to their degree of phylogenetic relatedness (Fig 2). For example, species within the genera *Ixodes*, *Rhipicephalus*, and *Amblyomma* were more similar among themselves in terms of the ions peak number and mass over charge (m/z) position in their mass spectra than species from different genera. Nonetheless, some closely related species within the *Amblyomma* genus such as *A. mixtum* (*cajennense*), *A. varium*, and *A. tapirellum* also showed fairly distinct protein spectra (Fig 2), which motivated the application of clustering algorithms for their classification.

Distinct mass spectra profiles between morphologically identified ixodid species could be classified by an unsupervised PCA algorithm to identify specimens. The quantitative performance of the PCA algorithm was assessed per species (Table 2), and visually confirmed with the graphic clustering presented in 3D plots (Fig 3). The PCA global positive identification rate was 91.2%, with 14 out of 18 species having higher than 90% positive identification rate. The PCA graphs showed that most species separated in well-defined clusters, and the distance among clusters seemed to be related to the degree of phylogenetic relatedness as evidenced by the clear separation from the specimens of *Dermacentor* and *Rhipicephalus* with those from *Haemaphysalis* and *Ixodes* (Fig 3A and 3B), or just between the specimens of *Amblyomma* (Fig 3C). When comparing species within the genus *Amblyomma* against those from *Ixodes*, again the spectra from specimens of each species clustered together with limited overlap between groups and those from different genera were clearly separated (Fig 3D).

Table 1. Description of specimens subjected to analysis with the MALDI mass spectrometry procedure.

Species Name	# of specimens	Locality code	# of expected spectra	# of obtained spectra	MALDI automatic spectra acquisition rate (%)
<i>Amblyomma mixtum</i> (<i>cajennense</i>)	4	a	12	6	50%
<i>Amblyomma calcaratum</i>	5	a, b	15	15	100%
<i>Amblyomma dissimile</i>	4	c	12	9	75%
<i>Amblyomma geayi</i>	4	d	12	12	100%
<i>Amblyomma nodosum</i>	4	a	12	10	83%
<i>Amblyomma oblongoguttatum</i>	4	a, e	12	8	67%
<i>Amblyomma ovale</i>	4	e	12	11	92%
<i>Amblyomma pecarium</i>	4	e	12	11	92%
<i>Amblyomma sabanerae</i>	3	f	9	9	100%
<i>Amblyomma varium</i>	4	g	12	9	75%
<i>Amblyomma naponense</i>	5	f	15	9	60%
<i>Amblyomma tapirellum</i> *	26	e, g	78	56	72%
<i>Ixodes affinis</i>	4	e	12	12	100%
<i>Ixodes boliviensis</i>	4	e	12	6	50%
<i>Dermacentor nitens</i>	4	c	12	9	75%
<i>Haemaphysalis juxtackochi</i>	6	a, e	18	11	61%
<i>Rhipicephalus microplus</i>	10	c, d	30	30	100%
<i>Rhipicephalus sanguineus</i>	4	a	12	6	50%
Total	103	a-g	309	239	77%

(a) = Panama: West Panama, Las Pavas; (b) = Panama: Colon, Madden Road; (c) = Panama: Colon, Achiote; (d) = Panama: West Panama, Capira; (e) Panama: Colon, Barro Colorado Island; (f) Panama: Colon, Sierra Llorona Lodge; (g) Panama: Colon, Gamboa. (*) Indicates some specific specimens that upon collection were stored fresh in Silica Gel (For more metadata information about these samples see also S1 Table).

<https://doi.org/10.1371/journal.pntd.0008849.t001>

Table 2. Performance of PCA and LDA clustering algorithms.

Species Name	PCA Positive Identification Rate (%)	LDA Positive Identification Rate (%)	Spectra per Class	# Training Elements	# Test Elements
<i>Amblyomma mixtum</i> (<i>cajennense</i>)	100.0%	100.0%	6	4000	2000
<i>Amblyomma calcaratum</i>	100.0%	99.6%	15	12000	3000
<i>Amblyomma dissimile</i>	67.6%	67.6%	9	7000	2000
<i>Amblyomma geayi</i>	99.1%	99.6%	12	9000	3000
<i>Amblyomma nodosum</i>	100.0%	100.0%	10	8000	2000
<i>Amblyomma oblongoguttatum</i>	100.0%	100.0%	8	6000	2000
<i>Amblyomma ovale</i>	100.0%	100.0%	11	8000	3000
<i>Amblyomma pecarium</i>	99.8%	99.0%	11	8000	3000
<i>Amblyomma sabanerae</i>	69.3%	85.9%	9	7000	2000
<i>Amblyomma varium</i>	99.8%	100.0%	9	7000	2000
<i>Amblyomma naponense</i>	100.0%	100.0%	9	7000	2000
<i>Amblyomma tapirellum</i>	97.8%	97.8%	56	44000	12000
<i>Dermacentor nitens</i>	21.7%	45.6%	12	9000	3000
<i>Haemaphysalis juxtackochi</i>	90.9%	97.8%	6	4000	2000
<i>Ixodes affinis</i>	84.0%	89.5%	9	7000	2000
<i>Ixodes boliviensis</i>	96.8%	98.8%	11	8000	3000
<i>Rhipicephalus microplus</i>	93.1%	98.7%	30	24000	6000
<i>Rhipicephalus sanguineus</i>	100.0%	100.0%	6	4000	2000
Global	91.2%	94.2%	239	183000	56000

<https://doi.org/10.1371/journal.pntd.0008849.t002>

In addition, the LDA clustering analysis showed a global positive identification rate of 94.2% (Fig 4; Table 2), with 14 out of 18 species having higher than 97.8% positive identification rate. The range of positive identification rates went from 100% (best score possible) for *A. mixtum* (*cajennense*), *A. nodosum*, *A. oblongoguttatum*, *A. ovale*, *A. varium*, *A. naponense* and *R. sanguineus* to 45.6% for *D. nitens*. The 3D representation plots of the LDA clustering displayed that the separation between species was more pronounced than with PCA when comparing species from different genera, confirming the improved quantitative results of the performance of the LDA algorithm (Table 2).

Discussion

Our results show that MALDI mass spectra of highly abundant proteins in arthropod legs served as fingerprints to identify samples of 18 species of Neotropical hard ticks using machine learning and pattern recognition algorithms to create a self-curated reference library. We compared smoothed and baseline-corrected spectra generated from unknown field-collected tick samples against the mean spectra from a subset of the same field samples that had already been identified through traditional means. To systematize this process, we used PCA and LDA algorithms to classify mass spectra without prior establishment of a high-quality reference library, which typically requires laboratory-reared specimens that may not be possible to obtain for all species. Global positive identification rates of up to 94.2% were achieved with this methodology, offering a rapid, reliable and objective approach to identify hard tick species, which will likely improve as more specimens are evaluated and included in our database.

These outcomes agree with our previous work [27] in which we used a similar approach to classify field-collected samples of 11 morphologically-identified species of *Anopheles* mosquitoes. In that study, Neotropical *Anopheles* samples were stored dry in silica gel at -20°C ,

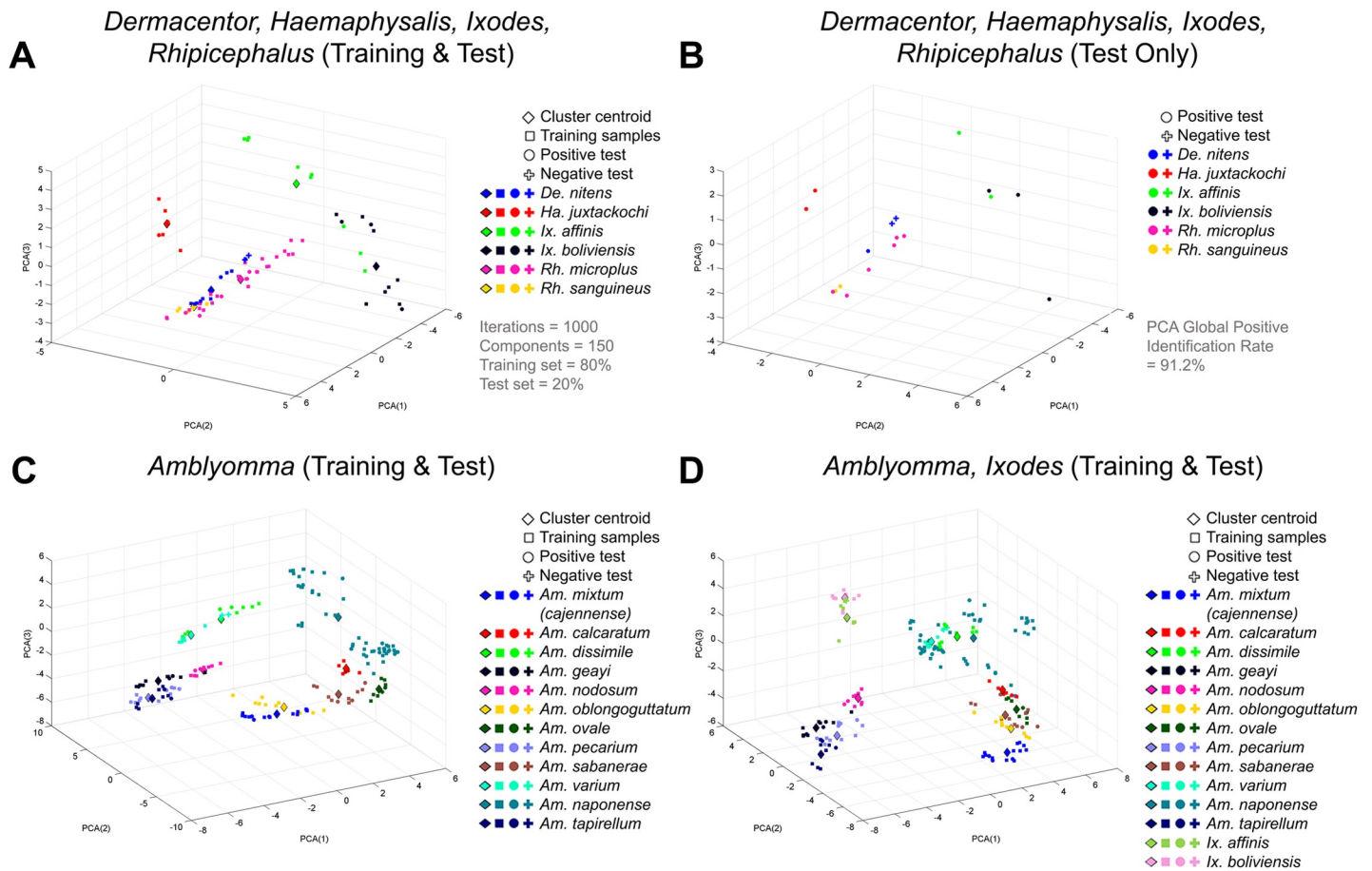


Fig 3. Principal component analysis (PCA) of individual species plotted against first, second and third principal components (PC). All species were classified using a Monte Carlo simulation with 1000 iterations, in which 80% of the samples were used as training set (□) and the remaining 20% as test set (• for positive identifications and + for negative ones). The cluster centroid of each species is also presented in the graph (◇). The plots show (A) the training and test sets for the species belonging to the *Dermacentor*, *Haemaphysalis*, *Ixodes* and *Rhipicephalus* genera, and (B) only the test sets for better visualization; as well as the training set and test set of (C) *Amblyomma* species alone or (D) *Amblyomma* in combination with *Ixodes* genera. The unsupervised PCA algorithm had a global positive identification rate of 91.2%. These 3D plots represent only one of the 1000 Monte Carlo iterations performed with the algorithm.

<https://doi.org/10.1371/journal.pntd.0008849.g003>

which seemed to avoid sample degradation and maintain spectral quality. This contrasts with the present study, where most of our specimens were stored in ethanol at -20°C for several years. Thus, our findings confirm that our novel analytical approach using MALDI and PCA/LDA clustering algorithms is robust for species classification regardless of the arthropod assemblage, sample storing conditions, and the lack of a high-quality reference library. In fact, the percentage of automatic spectra acquisition from the processed tick species was much higher (Table 1) than that obtained in our previous publication using mosquitoes, which ranged from 41.8% in *Anopheles albimanus* to 70.3% in *Anopheles triannulatus* [27].

Our results herein also show that both classification algorithms, PCA and LDA, were capable of clustering and recognizing spectra from up to 18 different tick species, including roughly 50% of ixodid taxa (e.g., both ecologically dominant and rare species) reported for Panama [27, 42]. LDA outcomes were more discriminant and robust than PCA overall, but PCA also classified species from different genera with over 91% accuracy and consistency. LDA was able to cluster each of the 18 species of ticks with validation and cross-validation scores above 94%, both between and within genera. As expected, the clustering algorithm was most accurate for distinctly related phylogenetic species (i.e., *Ixodes*, *Rhipicephalus* and *Haemaphysalis* genera),

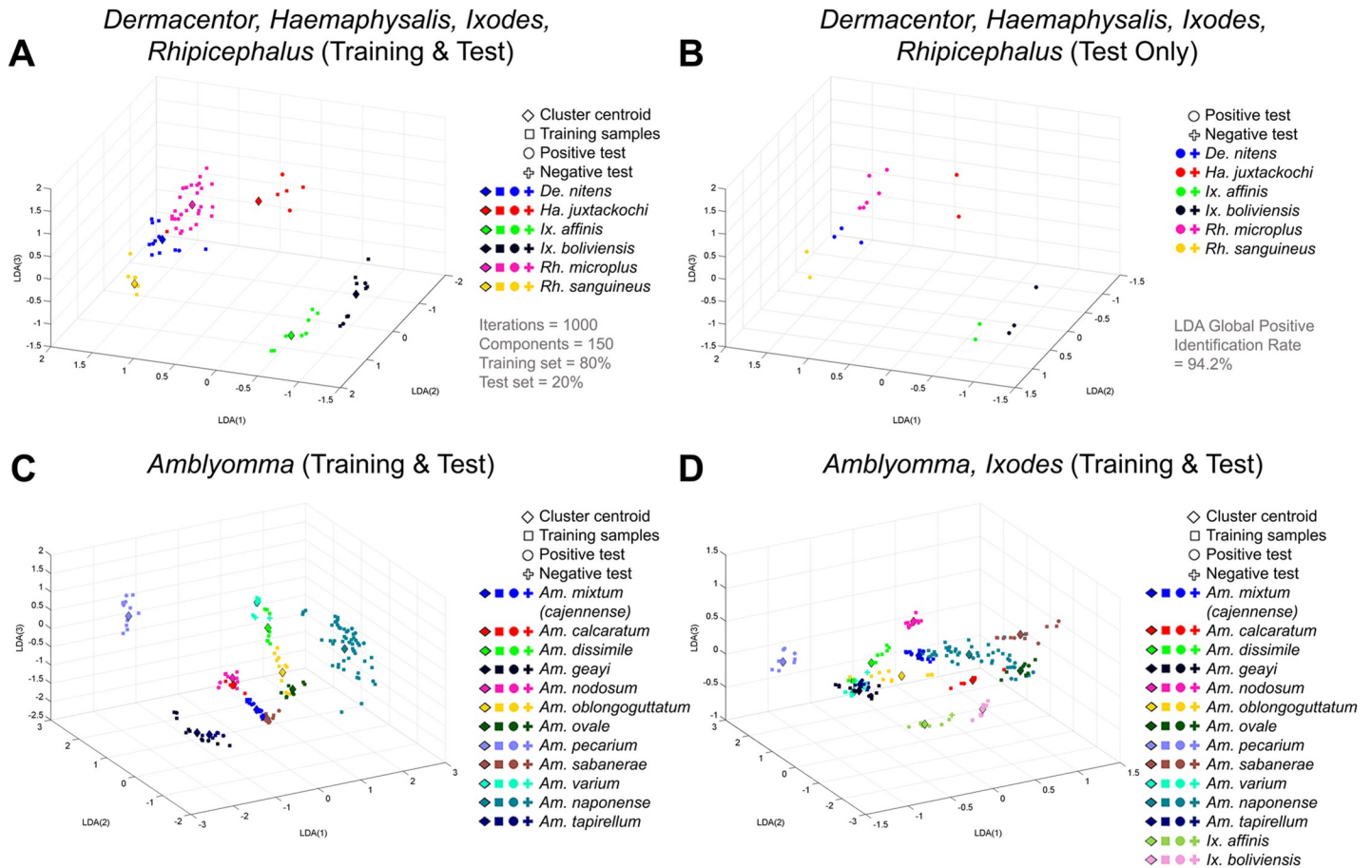


Fig 4. Linear Discriminant Analysis (LDA) applied to spectra from tick species of the genera *Amblyomma*, *Dermacentor*, *Haemaphysalis*, *Ixodes* and *Rhipicephalus*. The plots show (A) the training and test sets for species in the *Dermacentor*, *Haemaphysalis*, *Ixodes* and *Rhipicephalus* genera projected over the first three components of the LDA, as well as (B) only the test set for better visualization; and also the training and test sets for (C) the *Amblyomma* genus alone, as well as (D) the *Amblyomma* genus compared to the *Ixodes* genus. These 3D plots represent only one of the 1000 Monte Carlo iterations performed with the algorithm. The supervised LDA algorithm had a 94.2% global positive identification rate.

<https://doi.org/10.1371/journal.pntd.0008849.g004>

with higher than 97% success rate in most of these cases, than for closely related species (i.e., *Amblyomma* genus). However, *A. dissimile* and *D. nitens* depicted only moderate to low positive identification rates. Although this could be due to assemblage specific signals (i.e., high protein variability of conspecifics within these taxa), sample degradation and contamination, or technical errors such as spotting errors cannot be ruled out entirely. Future studies will have to corroborate the findings regarding these two species.

Although the number of samples analyzed for some ixodid species was relatively low, several of these taxa are considered cryptic species complexes [43] and have been implicated as vectors of human pathogens in Panama as well as more broadly, including *A. mixtum (cajennense)* and *D. nitens*, the likely vectors of *Rickettsia rickettsii*, known to cause Rocky Mountain spotted fever [44]. We also included samples of *A. tapirellum*, *A. oblongoguttatum* and *H. juxtakochi*, three species from which human pathogens have been previously isolated [45], such as: *Coxiella*-related bacteria, whose member *C. burnetii* can cause Q fever; *Ehrlichia*, which causes ehrlichiosis infection; and *Rickettsia*, which causes a variety of bacterial infections in humans and other animals. These results are important because our species identification platform can serve along with recently implemented metagenomic approaches as additional tools

for health ministries in Panama and other countries, to monitor, predict and manage tick-borne zoonotic pathogens [46].

Morphological taxonomic identification of ixodid ticks can be enhanced by molecular techniques such as DNA barcoding [8, 47], but this procedure is laborious, expensive and needs a highly trained lab technician. Studies show that typical DNA barcoding costs can range from \$2 to \$5 per sample, with difficult-to-extract samples increasing the cost two-fold or more [48, 49]; while costs associated to MALDI species identification have been calculated to be less than \$0.50 per sample, without considering the high equipment cost [50–52]. Furthermore, a comprehensive repository of DNA sequences (e.g., DNA barcodes) is needed in order to test species limits, yet only a handful of Neotropical tick species are represented in Genbank [53] or BOLD [54] repositories, which could limit identification to the most common taxa only. In addition, DNA barcoding occasionally fails to delimit species boundaries due to ambiguous evolutionary relationships among closely related tick species [47].

Modern methodologies of whole genome analysis of arthropod vectors using Illumina or Nanopore next generation sequencing platforms can be applied not only to delimit taxonomic boundaries among tick species, but also to examine vector evolution (i.e., positive selection and ecological diversification), demographic phenomena (i.e., expansion and bottlenecks) and molecular epidemiology (i.e., pathogen infection and genetic diversity). The cost of these modern technologies is decreasing rapidly, and they could quickly become a valid alternative for taxonomic studies in developing and middle-income countries of Central America, including Panama. Indeed, portable Nanopore MinION methodology can be performed at the site of interest, with a laptop computer by someone with very basic entomological knowledge, and at a very affordable price on a per-sample basis [55]. Nevertheless, the bioinformatic skills and cluster capacity to process whole genome sequences of tick samples might represent an impractical burden for some institutions in developing nations, which may not have the machinery or competency to analyze this kind of data. Moreover, using the Nanopore MinION next generation sequencing approach for the exclusive goal of achieving reliable taxonomic identification of tick species may represent an underutilized expenditure that might ultimately end up overkilling the budget of resource-limited institutions.

While MALDI mass spectrometry suffers from many of the shortcomings listed for other technologies, our approach can be used to identify both field-collected vectors and the pathogens they harbor in a short period of time, with a minimal amount of tissue and without the need of expert taxonomists. Our strategy to analyze protein spectra also overcomes the drawbacks of working without a reference library to classify unknown samples. We posit that MALDI mass spectra of highly abundant proteins from arthropod tissues is a powerful tool for species identification that can be easily adapted to other biological systems. However, we also believe that this technology will be best used as a complement to the traditional barcoding technique or modern next generation sequencing methodologies, to accurately confirm species boundaries across entire arthropod communities, while considering problematic vector taxonomy and the availability of local financial resources. Developing an additional tool for rapid and accurate arthropod species identification offers further flexibility to the fluctuating budgets of the research community in Central/South America.

The long-term goal of our analytical approach with MALDI is to develop a tool that can enhance currently available open-source, web-based platforms, such as MALDI UP [56], MicrobeMS [57], or Mass-Up [58]; or become a new *all-in-one* platform where users can upload mass spectra datasets of known specimens to increase the number of species covered (e.g., bacteria, fungi, insects) and directly test spectra from unknown specimens for identification with our clustering algorithms. This crowd-sourced approach could be more cost effective, given that it is not necessary to generate a reference library of well-curated samples.

Instead, field samples can be taxonomically assigned as they arrive to the laboratory using a correctly matched protein fingerprint, while unidentified samples can be identified with traditional methods and added as new entries into the growing self-curated reference database.

Conclusions

The present study used MALDI mass spectrometry as a tool to rapidly identify Neotropical specimens of adult hard ticks that had been preserved in ethanol for several years. Our algorithms were capable of identifying specimens from the 18 tick species evaluated, based on their protein spectra “fingerprint” with up to 94% cross-validation capability. This is the first report of the protein mass spectra from the leg for most of these Neotropical tick species. Large arthropod groups such as ticks are difficult to identify with currently available strategies from commercial vendors, forcing the user to lower the “quality” bar of a positive match to enhance the percentage of correct identification. Our MALDI/self-curated library approach, although still under development and serving as an auxiliary technique to traditional identification methods (and not necessarily replacing them), would reduce considerably the number of samples that would require morphological identification or DNA barcoding. This will reduce the time and cost needed to integrate these techniques in routine surveillance programs in Neotropical regions where tick diversity remains relatively uncharacterized.

Supporting information

S1 Fig. Baseline-corrected and smoothed spectra for tick specimens from the species *A. calcaratum*. Major ion peaks and their molecular weights are annotated in the range of 2,000 to 20,000 m/z for all specimens. The dataset shows consistently similar protein profiles, regardless of their sex, collection date and/or sampling location.
(TIF)

S2 Fig. Baseline-corrected and smoothed spectra for tick specimens from the species *R. microplus*. Major ion peaks and their molecular weights are annotated in the range of 2,000 to 20,000 m/z for all specimens. The dataset shows consistently similar protein profiles, regardless of their sex, collection date and/or sampling location.
(TIF)

S3 Fig. Baseline-corrected and smoothed spectra for tick specimens from the species *A. tapirellum*. Major ion peaks and their molecular weights are annotated in the range of 2,000 to 20,000 m/z for all specimens. The dataset shows consistently similar protein profiles, regardless of their sex, collection date and/or sampling location.
(TIF)

S1 Table. Metadata of specimens and species of hard tick (e.g., Ixodidae) collected in Panama.
(XLSX)

Acknowledgments

Special thanks to Philip Davis and Eric Rodríguez from the University of Panama, for assisting with tick collections during the study. We are grateful to Aishwarya Sunderrajan, from the Madras Institute of Technology, India, for her support editing some of the figures for this work. We want to thank the governmental personnel at the Panamanian Ministry of Environment (*MiAmbiente*) for supporting scientific collecting of ticks in Panama.

Author Contributions

Conceptualization: Rolando A. Gittens, Jose R. Loaiza.

Data curation: Rolando A. Gittens, Javier E. Sanchez-Galan, Fernando Merchan, Jose R. Loaiza.

Formal analysis: Rolando A. Gittens, Javier E. Sanchez-Galan, Fernando Merchan, Jose R. Loaiza.

Funding acquisition: Rolando A. Gittens, Luis C. Mejía, Javier E. Sanchez-Galan, Fernando Merchan, Matthew J. Miller, Luis F. De León, Jose R. Loaiza.

Investigation: Rolando A. Gittens, Alejandro Almanza, Kelly L. Bennett, Javier E. Sanchez-Galan, Fernando Merchan, Jonathan Kern, Helen J. Esser, Robert Hwang, May Dong, Eric Álvarez, Jose R. Loaiza.

Methodology: Rolando A. Gittens, Jose R. Loaiza.

Project administration: Rolando A. Gittens, Jose R. Loaiza.

Resources: Rolando A. Gittens, Luis C. Mejía, Javier E. Sanchez-Galan, Fernando Merchan, Matthew J. Miller, Luis F. De León, Jose R. Loaiza.

Software: Javier E. Sanchez-Galan, Fernando Merchan, Jonathan Kern.

Supervision: Rolando A. Gittens, Jose R. Loaiza.

Validation: Rolando A. Gittens, Javier E. Sanchez-Galan, Fernando Merchan, Jose R. Loaiza.

Visualization: Rolando A. Gittens, Alejandro Almanza, Javier E. Sanchez-Galan, Fernando Merchan, Jonathan Kern, Jose R. Loaiza.

Writing – original draft: Rolando A. Gittens, Javier E. Sanchez-Galan, Fernando Merchan, Jose R. Loaiza.

Writing – review & editing: Rolando A. Gittens, Alejandro Almanza, Kelly L. Bennett, Luis C. Mejía, Javier E. Sanchez-Galan, Fernando Merchan, Jonathan Kern, Matthew J. Miller, Helen J. Esser, Robert Hwang, May Dong, Luis F. De León, Eric Álvarez, Jose R. Loaiza.

References

1. Estrada-Pena A, Jongejan F. Ticks feeding on humans: a review of records on human-biting Ixodoidea with special reference to pathogen transmission. *Exp Appl Acarol.* 1999; 23(9):685–715. Epub 1999/12/03. <https://doi.org/10.1023/a:1006241108739> PMID: 10581710.
2. McCoy KD, Leger E, Dietrich M. Host specialization in ticks and transmission of tick-borne diseases: a review. *Front Cell Infect Microbiol.* 2013; 3:57. Epub 2013/10/11. <https://doi.org/10.3389/fcimb.2013.00057> PMID: 24109592; PubMed Central PMCID: PMC3790072.
3. Parola P, Labruna MB, Raoult D. Tick-borne rickettsioses in America: unanswered questions and emerging diseases. *Curr Infect Dis Rep.* 2009; 11(1):40–50. Epub 2008/12/20. <https://doi.org/10.1007/s11908-009-0007-5> PMID: 19094824.
4. Jongejan F, Uilenberg G. The global importance of ticks. *Parasitology.* 2004; 129 Suppl:S3–14. Epub 2005/06/09. <https://doi.org/10.1017/s0031182004005967> PMID: 15938502.
5. Estripeaut D, Aramburu MG, Saez-Llorens X, Thompson HA, Dasch GA, Paddock CD, et al. Rocky Mountain spotted fever, Panama. *Emerg Infect Dis.* 2007; 13(11):1763–5. Epub 2008/01/26. <https://doi.org/10.3201/eid1311.070931> PMID: 18217566; PubMed Central PMCID: PMC3375809.
6. Hidalgo M, Miranda J, Heredia D, Zambrano P, Vesga JF, Lizarazo D, et al. Outbreak of Rocky Mountain spotted fever in Cordoba, Colombia. *Mem Inst Oswaldo Cruz.* 2011; 106(1):117–8. Epub 2011/02/23. <https://doi.org/10.1590/s0074-02762011000100019> PMID: 21340366.

7. Tribaldos M, Zaldivar Y, Bermudez S, Samudio F, Mendoza Y, Martinez AA, et al. Rocky Mountain spotted fever in Panama: a cluster description. *J Infect Dev Ctries*. 2011; 5(10):737–41. Epub 2011/10/15. <https://doi.org/10.3855/jidc.2189> PMID: 21997944.
8. Arguello AP, Hun L, Rivera P, Taylor L. A fatal urban case of rocky mountain spotted fever presenting an eschar in San Jose, Costa Rica. *Am J Trop Med Hyg*. 2012; 87(2):345–8. Epub 2012/08/03. <https://doi.org/10.4269/ajtmh.2012.12-0153> PMID: 22855769; PubMed Central PMCID: PMC3414575.
9. Esser HJ, Herre EA, Bluthgen N, Loaiza JR, Bermudez SE, Jansen PA. Host specificity in a diverse Neotropical tick community: an assessment using quantitative network analysis and host phylogeny. *Parasit Vectors*. 2016; 9(1):372. Epub 2016/07/01. <https://doi.org/10.1186/s13071-016-1655-6> PMID: 27357506; PubMed Central PMCID: PMC4928246.
10. Guglielmone AA, Robbins RG, Apanaskevich DA, Petney TN, Estrada-Peña A, Horak IG. *The Hard Ticks of the World (Acari: Ixodida: Ixodidae)*. Dordrecht: Springer Netherlands; 2014.
11. Martins TF, Onofrio VC, Barros-Battesti DM, Labruna MB. Nymphs of the genus *Amblyomma* (Acari: Ixodidae) of Brazil: descriptions, redescriptions, and identification key. *Ticks Tick Borne Dis*. 2010; 1(2):75–99. Epub 2011/07/21. <https://doi.org/10.1016/j.ttbdis.2010.03.002> PMID: 21771514.
12. Murugaiyan J, Roesler U. MALDI-TOF MS Profiling-Advances in Species Identification of Pests, Parasites, and Vectors. *Front Cell Infect Microbiol*. 2017; 7:184. Epub 2017/05/31. <https://doi.org/10.3389/fcimb.2017.00184> PMID: 28555175; PubMed Central PMCID: PMC5430024.
13. Croxatto A, Prod'hom G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev*. 2012; 36(2):380–407. Epub 2011/11/19. <https://doi.org/10.1111/j.1574-6976.2011.00298.x> PMID: 22092265.
14. Singhal N, Kumar M, Virdi JS. MALDI-TOF MS in clinical parasitology: applications, constraints and prospects. *Parasitology*. 2016; 143(12):1491–500. Epub 2016/07/09. <https://doi.org/10.1017/S0031182016001189> PMID: 27387025.
15. Tan KE, Ellis BC, Lee R, Stamper PD, Zhang SX, Carroll KC. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: a bench-by-bench study for assessing the impact on time to identification and cost-effectiveness. *J Clin Microbiol*. 2012; 50(10):3301–8. Epub 2012/08/03. <https://doi.org/10.1128/JCM.01405-12> PMID: 22855510; PubMed Central PMCID: PMC3457442.
16. Ge MC, Kuo AJ, Liu KL, Wen YH, Chia JH, Chang PY, et al. Routine identification of microorganisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: Success rate, economic analysis, and clinical outcome. *J Microbiol Immunol Infect*. 2017; 50(5):662–8. Epub 2016/07/19. <https://doi.org/10.1016/j.jmii.2016.06.002> PMID: 27426930.
17. Lai YH, Wang YS. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry: Mechanistic Studies and Methods for Improving the Structural Identification of Carbohydrates. *Mass Spectrom (Tokyo)*. 2017; 6(Spec Iss 2):S0072. Epub 2017/09/30. <https://doi.org/10.5702/massspectrometry.S0072> PMID: 28959517; PubMed Central PMCID: PMC5610957.
18. Florio W, Tavanti A, Barnini S, Ghelardi E, Lupetti A. Recent Advances and Ongoing Challenges in the Diagnosis of Microbial Infections by MALDI-TOF Mass Spectrometry. *Front Microbiol*. 2018; 9:1097. Epub 2018/06/14. <https://doi.org/10.3389/fmicb.2018.01097> PMID: 29896172; PubMed Central PMCID: PMC5986882.
19. Suarez E, Nguyen HP, Ortiz IP, Lee KJ, Kim SB, Krzywinski J, et al. Matrix-assisted laser desorption/ionization-mass spectrometry of cuticular lipid profiles can differentiate sex, age, and mating status of *Anopheles gambiae* mosquitoes. *Anal Chim Acta*. 2011; 706(1):157–63. Epub 2011/10/15. <https://doi.org/10.1016/j.aca.2011.08.033> PMID: 21995923.
20. Lafri I, Almeras L, Bitam I, Caputo A, Yssouf A, Forestier CL, et al. Identification of Algerian Field-Caught Phlebotomine Sand Fly Vectors by MALDI-TOF MS. *PLoS Negl Trop Dis*. 2016; 10(1):e0004351. Epub 2016/01/16. <https://doi.org/10.1371/journal.pntd.0004351> PMID: 26771833; PubMed Central PMCID: PMC4714931.
21. Muller P, Pfluger V, Wittwer M, Ziegler D, Chandre F, Simard F, et al. Identification of cryptic *Anopheles* mosquito species by molecular protein profiling. *PLoS One*. 2013; 8(2):e57486. Epub 2013/03/08. <https://doi.org/10.1371/journal.pone.0057486> PMID: 23469000; PubMed Central PMCID: PMC3585343.
22. Yssouf A, Almeras L, Berenger JM, Laroche M, Raoult D, Parola P. Identification of tick species and disseminate pathogen using hemolymph by MALDI-TOF MS. *Ticks Tick Borne Dis*. 2015; 6(5):579–86. Epub 2015/06/09. <https://doi.org/10.1016/j.ttbdis.2015.04.013> PMID: 26051210.
23. Yssouf A, Flaudrops C, Drali R, Kernif T, Socolovschi C, Berenger JM, et al. Matrix-assisted laser desorption ionization-time of flight mass spectrometry for rapid identification of tick vectors. *J Clin Microbiol*. 2013; 51(2):522–8. Epub 2012/12/12. <https://doi.org/10.1128/JCM.02665-12> PMID: 23224087; PubMed Central PMCID: PMC3553915.

24. Yssouf A, Socolovschi C, Flaudrops C, Ndiath MO, Sougoufara S, Dehecq JS, et al. Matrix-assisted laser desorption ionization—time of flight mass spectrometry: an emerging tool for the rapid identification of mosquito vectors. *PLoS One*. 2013; 8(8):e72380. Epub 2013/08/27. <https://doi.org/10.1371/journal.pone.0072380> PMID: 23977292; PubMed Central PMCID: PMC3744494.
25. Yssouf A, Socolovschi C, Leulmi H, Kernif T, Bitam I, Audoly G, et al. Identification of flea species using MALDI-TOF/MS. *Comp Immunol Microbiol Infect Dis*. 2014; 37(3):153–7. Epub 2014/06/01. <https://doi.org/10.1016/j.cimid.2014.05.002> PMID: 24878069.
26. Bennett KL, Gomez Martinez C, Almanza A, Rovira JR, McMillan WO, Enriquez V, et al. High infestation of invasive *Aedes* mosquitoes in used tires along the local transport network of Panama. *Parasit Vectors*. 2019; 12(1):264. Epub 2019/05/28. <https://doi.org/10.1186/s13071-019-3522-8> PMID: 31133041; PubMed Central PMCID: PMC6537307.
27. Loaiza JR, Almanza A, Rojas JC, Mejia L, Cervantes ND, Sanchez-Galan JE, et al. Application of matrix-assisted laser desorption/ionization mass spectrometry to identify species of Neotropical Anophelids vectors of malaria. *Malaria journal*. 2019; 18(1):95. <https://doi.org/10.1186/s12936-019-2723-0> PMID: 30902057.
28. Dieme C, Yssouf A, Vega-Rua A, Berenger JM, Failloux AB, Raoult D, et al. Accurate identification of Culicidae at aquatic developmental stages by MALDI-TOF MS profiling. *Parasit Vectors*. 2014; 7:544. Epub 2014/12/03. <https://doi.org/10.1186/s13071-014-0544-0> PMID: 25442218; PubMed Central PMCID: PMC4273427.
29. Laroche M, Almeras L, Pecchi E, Bechah Y, Raoult D, Viola A, et al. MALDI-TOF MS as an innovative tool for detection of *Plasmodium* parasites in Anopheles mosquitoes. *Malar J*. 2017; 16(1):5. Epub 2017/01/05. <https://doi.org/10.1186/s12936-016-1657-z> PMID: 28049524; PubMed Central PMCID: PMC5209920.
30. Steinmann IC, Pfluger V, Schaffner F, Mathis A, Kaufmann C. Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for the identification of ceratopogonid and culicid larvae. *Parasitology*. 2013; 140(3):318–27. Epub 2012/11/23. <https://doi.org/10.1017/S0031182012001618> PMID: 23171762.
31. Dvorak V, Halada P, Hlavackova K, Dokianakis E, Antoniou M, Volf P. Identification of phlebotomine sand flies (Diptera: Psychodidae) by matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Parasit Vectors*. 2014; 7:21. Epub 2014/01/16. <https://doi.org/10.1186/1756-3305-7-21> PMID: 24423215; PubMed Central PMCID: PMC3896986.
32. Hoppenheit A, Murugaiyan J, Bauer B, Steuber S, Clausen PH, Roesler U. Identification of *Tsetse* (*Glossina* spp.) using matrix-assisted laser desorption/ionisation time of flight mass spectrometry. *PLoS Negl Trop Dis*. 2013; 7(7):e2305. Epub 2013/07/23. <https://doi.org/10.1371/journal.pntd.0002305> PMID: 23875040; PubMed Central PMCID: PMC3708848.
33. Karger A, Kampen H, Bettin B, Dautel H, Ziller M, Hoffmann B, et al. Species determination and characterization of developmental stages of ticks by whole-animal matrix-assisted laser desorption/ionization mass spectrometry. *Ticks Tick Borne Dis*. 2012; 3(2):78–89. Epub 2012/04/11. <https://doi.org/10.1016/j.ttbdis.2011.11.002> PMID: 22487425.
34. Kaufmann C, Ziegler D, Schaffner F, Carpenter S, Pfluger V, Mathis A. Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for characterization of *Culicoides nubeculosus* biting midges. *Med Vet Entomol*. 2011; 25(1):32–8. Epub 2010/12/02. <https://doi.org/10.1111/j.1365-2915.2010.00927.x> PMID: 21118284.
35. Rothen J, Githaka N, Kanduma EG, Olds C, Pfluger V, Mwaura S, et al. Matrix-assisted laser desorption/ionization time of flight mass spectrometry for comprehensive indexing of East African ixodid tick species. *Parasit Vectors*. 2016; 9:151. Epub 2016/03/17. <https://doi.org/10.1186/s13071-016-1424-6> PMID: 26979606; PubMed Central PMCID: PMC4792108.
36. Kaufmann C, Schaffner F, Ziegler D, Pfluger V, Mathis A. Identification of field-caught *Culicoides* biting midges using matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Parasitology*. 2012; 139(2):248–58. Epub 2011/10/20. <https://doi.org/10.1017/S0031182011001764> PMID: 22008297.
37. Diarra AZ, Almeras L, Laroche M, Berenger J-M, Koné AK, Bocoum Z, et al. Molecular and MALDI-TOF identification of ticks and tick-associated bacteria in Mali. *PLoS Negl Trop Dis*. 2017; 11(7):e0005762. <https://doi.org/10.1371/journal.pntd.0005762> PMID: 28742123
38. Belhumeur PN, Hespanha JP, Kriegman DJ. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell*. 1997; 19(7):711–20.
39. Turk M, Pentland A. Eigenfaces for recognition. *J Cogn Neurosci*. 1991; 3(1):71–86. Epub 1991/01/01. <https://doi.org/10.1162/jocn.1991.3.1.71> PMID: 23964806.

40. Ryan L, Hany F, Bruce D. Probabilistic Disease Classification of Expression-Dependent Proteomic Data from Mass Spectrometry of Human Serum. *J Comput Biol.* 2003; 10(6):925–46. <https://doi.org/10.1089/106652703322756159> PMID: 14980018.
41. Satten GA, Datta S, Moura H, Woolfitt AR, Carvalho MdG, Carlone GM, et al. Standardization and denoising algorithms for mass spectra to classify whole-organism bacterial specimens. *Bioinformatics.* 2004; 20(17):3128–36. <https://doi.org/10.1093/bioinformatics/bth372> PMID: 15217815
42. Fairchild GB, Kohla GM, Tipton VJ. The ticks of Panama (Acarina: Ixodoidea). *Ectoparasites of Panama.* Chicago: Field Museum of Natural History; 1966. p. 167–219.
43. Nava S, Beati L, Labruna MB, Cáceres AG, Mangold AJ, Guglielmo AA. Reassessment of the taxonomic status of *Amblyomma cajennense* (Fabricius, 1787) with the description of three new species, *Amblyomma tonelliae* n. sp., *Amblyomma interandinum* n. sp. and *Amblyomma patinoi* n. sp., and reinstatement of *Amblyomma mixtum* Koch, 1844, and *Amblyomma sculptum* Berlese, 1888 (Ixodida: Ixodidae). *Ticks Tick Borne Dis.* 2014; 5(3):252–76. <https://doi.org/10.1016/j.ttbdis.2013.11.004> PMID: 24556273
44. Lopes MG, May Junior J, Foster RJ, Harmsen BJ, Sanchez E, Martins TF, et al. Ticks and rickettsiae from wildlife in Belize, Central America. *Parasit Vectors.* 2016; 9:62. Epub 2016/02/03. <https://doi.org/10.1186/s13071-016-1348-1> PMID: 26831147; PubMed Central PMCID: PMC4736261.
45. Parola P, Raoult D. Ticks and tickborne bacterial diseases in humans: an emerging infectious threat. *Clin Infect Dis.* 2001; 32(6):897–928. Epub 2001/03/15. <https://doi.org/10.1086/319347> PMID: 11247714.
46. Bennett KL, Almanza A, McMillan WO, Saltonstall K, Vdovenko EL, Vinda JS, et al. Habitat disturbance and the organization of bacterial communities in Neotropical hematophagous arthropods. *PLoS One.* 2019; 14(9):e0222145. Epub 2019/09/07. <https://doi.org/10.1371/journal.pone.0222145> PMID: 31491005; PubMed Central PMCID: PMC6730880.
47. Miller MJ, Esser HJ, Loaiza JR, Herre EA, Aguilar C, Quintero D, et al. Molecular Ecological Insights into Neotropical Bird-Tick Interactions. *PLoS One.* 2016; 11(5):e0155989. Epub 2016/05/21. <https://doi.org/10.1371/journal.pone.0155989> PMID: 27203693; PubMed Central PMCID: PMC4874597.
48. Hebert PD, Gregory TR. The promise of DNA barcoding for taxonomy. *Syst Biol.* 2005; 54(5):852–9. Epub 2005/10/26. <https://doi.org/10.1080/10635150500354886> PMID: 16243770.
49. Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov EV. Is DNA barcoding actually cheaper and faster than traditional morphological methods: results from a survey of freshwater bioassessment efforts in the United States? *PLoS One.* 2014; 9(4):e95525. Epub 2014/04/24. <https://doi.org/10.1371/journal.pone.0095525> PMID: 24755838; PubMed Central PMCID: PMC3995707.
50. Dhiman N, Hall L, Wohlfiel SL, Buckwalter SP, Wengenack NL. Performance and cost analysis of matrix-assisted laser desorption ionization-time of flight mass spectrometry for routine identification of yeast. *J Clin Microbiol.* 2011; 49(4):1614–6. Epub 2011/01/29. <https://doi.org/10.1128/JCM.02381-10> PMID: 21270234; PubMed Central PMCID: PMC3122878.
51. El-Bouri K, Johnston S, Rees E, Thomas I, Bome-Mannathoko N, Jones C, et al. Comparison of bacterial identification by MALDI-TOF mass spectrometry and conventional diagnostic microbiology methods: agreement, speed and cost implications. *Br J Biomed Sci.* 2012; 69(2):47–55. Epub 2012/08/10. PMID: 22872927.
52. Lagace-Wiens PR, Adam HJ, Karlowsky JA, Nichol KA, Pang PF, Guenther J, et al. Identification of blood culture isolates directly from positive blood cultures by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry and a commercial extraction system: analysis of performance, cost, and turnaround time. *J Clin Microbiol.* 2012; 50(10):3324–8. Epub 2012/08/10. <https://doi.org/10.1128/JCM.01479-12> PMID: 22875888; PubMed Central PMCID: PMC3457416.
53. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2016; 44(D1):D67–D72. Epub 2015/11/20. <https://doi.org/10.1093/nar/gkv1276> PMID: 26590407.
54. Ratnasingham S, Hebert PDN. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol Ecol Notes.* 2007; 7(3):355–64. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790.
55. Srivathsan A, Baloglu B, Wang W, Tan WX, Bertrand D, Ng AHQ, et al. A MinION-based pipeline for fast and cost-effective DNA barcoding. *Mol Ecol Resour.* 2018. Epub 2018/04/20. <https://doi.org/10.1111/1755-0998.12890> PMID: 29673082.
56. Rau J, Eisenberg T, Sting R. MALDI-UP—An Internet Platform for the Exchange of MALDI-TOF Mass Spectra (User guide for <http://maldi-up.ua-bw.de/>). *eJournal.* 2016; 2016(01).
57. Lasch P, Wahab T, Weil S, Palyi B, Tomaso H, Zange S, et al. Identification of Highly Pathogenic Microorganisms by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry: Results of an Interlaboratory Ring Trial. *J Clin Microbiol.* 2015; 53(8):2632–40. Epub 2015/06/13. <https://doi.org/10.1128/JCM.00813-15> PMID: 26063856; PubMed Central PMCID: PMC4508426.

58. Lopez-Fernandez H, Santos HM, Capelo JL, Fdez-Riverola F, Glez-Pena D, Reboiro-Jato M. Mass-Up: an all-in-one open software application for MALDI-TOF mass spectrometry knowledge discovery. *BMC Bioinformatics*. 2015; 16:318. Epub 2015/10/07. <https://doi.org/10.1186/s12859-015-0752-4> PMID: [26437641](https://pubmed.ncbi.nlm.nih.gov/26437641/); PubMed Central PMCID: PMC4595311.