

RESEARCH

Open Access



MALDI-TOF mass spectrometry combined with machine learning algorithms to identify protein profiles related to malaria infection in human sera from Côte d'Ivoire

Fateneba Kone^{1,2,3†}, Lucie Conrad^{1†}, Jean T. Coulibaly^{2,3,4,5}, Kigbafori D. Silué^{2,3}, Sören L. Becker^{1,4,5,6}, Brama Kone^{3,7} and Issa Sy^{1*}

Abstract

Background In sub-Saharan Africa, *Plasmodium falciparum* is the most prevalent species of malaria parasites. In endemic areas, malaria is mainly diagnosed using microscopy or rapid diagnostic tests (RDTs), which have limited sensitivity, and microscopic expertise is waning in non-endemic regions. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) is nowadays the standard method in routine microbiology laboratories for bacteria and fungi identification in high-income countries, but is rarely used for parasite detection. This study aims to employ MALDI-TOF MS for identifying malaria by distinguishing *P. falciparum*-positive from *P. falciparum*-negative sera.

Methods Sera were obtained from 282 blood samples collected from non-febrile, asymptomatic people aged 5 to 58 years in southern Côte d'Ivoire. Infectious status and parasitaemia were determined by both RDTs and microscopy, followed by a categorization into two groups (*P. falciparum*-positive and *P. falciparum*-negative samples). MALDI-TOF MS analysis was carried out by generating protein spectra profiles from 131 *Plasmodium*-positive and 94 *Plasmodium*-negative sera as the training set. Machine learning (ML) algorithms were employed for distinguishing *P. falciparum*-positive from *P. falciparum*-negative samples. Subsequently, a subset of 57 sera (42 *P. falciparum*-positive and 15 *P. falciparum*-negative) was used as the validation set to evaluate the best two of the four models trained.

Results MALDI-TOF MS was able to generate good-quality spectra from both *P. falciparum*-positive and *P. falciparum*-negative serum samples. High similarities between the protein spectra profiles did not allow for distinguishing the two groups using principal component analysis (PCA). When four supervised ML algorithms were tested by ten-fold cross-validation, *P. falciparum*-positive sera were discriminated against *P. falciparum*-negative sera with a global accuracy ranging from 73.28% to 81.30%, while sensitivity ranged from 70.23% to 83.97%. The independent test performed with a subset of 57 serum samples showed accuracies of 85.96% and 89.47%, and sensitivities of 90.48% and 92.86%, respectively, for LightGBM and RF.

Conclusion MALDI-TOF MS combined with ML might be applied for detection of protein profiles related to *P. falciparum* malaria infection in human serum samples. Additional research is warranted for further optimization such as specific biomarkers detection or using other ML models.

[†]Fateneba Kone and Lucie Conrad have contributed equally to this article.

*Correspondence:

Issa Sy
issa.sy@uks.eu

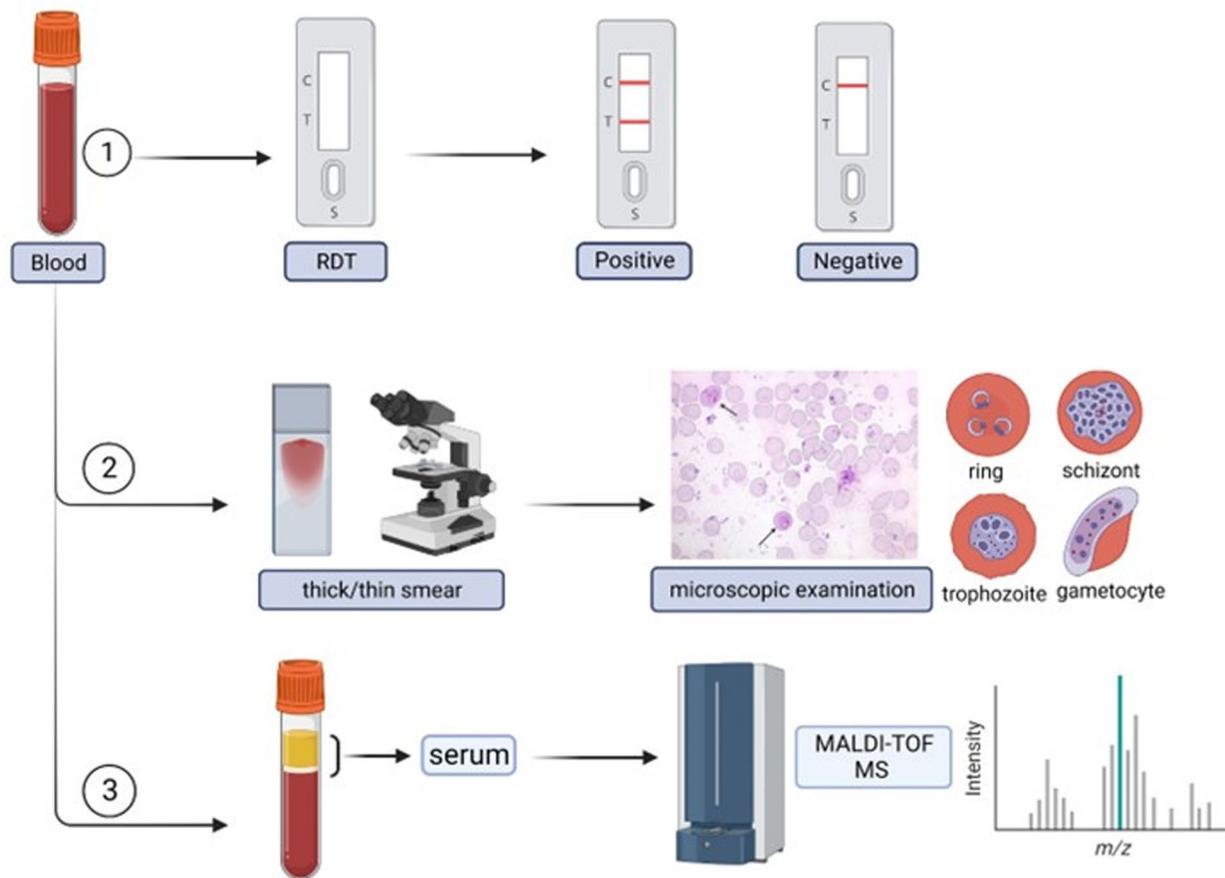
Full list of author information is available at the end of the article



Keywords Malaria identification, *Plasmodium falciparum*, Serum, Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry, Machine learning (ML)

Graphical Abstract

MALDI-TOF MS and Malaria diagnosis



Background

Malaria is a protozoan disease caused by parasites of the genus *Plasmodium*, which are transmitted by *Anopheles* mosquitoes [1]. Species differentiation is of great importance for clinical decision-making and selection of the adequate treatment regimen [2]. It is the most important parasitic disease causing around 249 million cases in 85 malaria-endemic countries and 608,000 related deaths in 2022 [3]. The African region, especially sub-Saharan Africa, is the most affected

region with 233 million cases, representing 94% of global cases and an estimated 580,000 annual deaths in 2022, most of which are attributable to *Plasmodium falciparum* [3, 4]. In Cote d’Ivoire, it is estimated that about 7,343,000 million malaria cases and 1555 deaths occur each year [3]. Infections and malaria-related deaths affect mostly pre-school-aged children (pre-SAC) (i.e., children under 5 years old), who are the most vulnerable populations (76% of global deaths) [3].

Malaria is mainly diagnosed by morphologic analysis of stained blood smears (thin and thick smears) and species differentiation is based on distinct morphological features, such as Schüffner's dots, Maurer's clefts, ring forms, shape of gametocytes [5]. Rapid diagnostic tests (RDTs) are also commonly employed, but their sensitivity varies and is lower for species other than *P. falciparum* [6]. Alternative methods such as nucleic acid amplification tests (NAATs) [e.g., real-time PCR, nested PCR, multiplex PCR, loop-mediated isothermal amplification (LAMP), or nucleic acid sequence-based amplification (NASBA)] based on the detection of 18S rRNA or other specific gene targets [e.g., telomere-associated repetitive element 2 (TARE-2), var acidic terminal sequence (var-ATS), mitochondrial DNA, or cytochrome oxidase III (COX-III)] can also be employed [7, 8]. The limitations of currently available techniques (e.g., the limited sensitivity of RDTs, challenging and sometimes unclear microscopy, and the scarcity of specific PCR kits in resource-limited countries, as well as the waning parasitological expertise in routine microbiology laboratories in non-endemic regions) are encouraging the development of new alternative approaches [8].

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) is a rapid, accurate, and cost-effective technique based on the analysis of protein spectra profiles. Indeed, it is nowadays used as the standard diagnostic method for the identification of bacteria, mycobacteria, and fungi in clinical laboratories in high-income countries [9–12]. The applicability of this technique has been investigated as a new tool for the identification of parasites, particularly for malaria diagnosis through haemozoin detection [13, 14]. More recently, a proof-of-concept study has investigated the utilization of MALDI-TOF MS as a new technique for diagnosing malaria by detecting *P. falciparum* isolated from infected red blood cells (iRBCs) using a culture-based approach [15]. The new trend of introducing artificial intelligence (AI), machine learning (ML), and deep learning (DL) into microbiology laboratories has demonstrated its applicability for improved diagnostic accuracy and automating the microscopic analysis of blood smears to detect malaria [16–18]. Similarly, the use of such approaches combining ML algorithms and MALDI-TOF MS could extend the applicability of MALDI-TOF MS to the direct analysis of more complex matrices such as blood, or serum for malaria diagnosis. Hence, a detailed assessment of MALDI protein spectra profiles of infected serum samples could be of a great interest in detecting *P. falciparum* malaria. Against this background, this study aimed to employ MALDI-TOF MS coupled with ML algorithms for the identification of *P. falciparum* malaria infection in human sera by assessing its identification

accuracy for differentiating human serum of non-*P. falciparum* infected participants from serum infected with *P. falciparum*.

Methods

Ethics statement

All the procedures involving humans were conducted in strict accordance with the Institutional local Guidelines and approved by the National Ethical Committee for Life Science and Health of the Ministry of Health and Public Hygiene of Côte d'Ivoire under the number 056-21/MSHP/CNESVS-km. Consents were obtained from all adult participants as well as from parents or legal guardians of participants aged below 18 years. Clinical malaria cases were provided with medication in accordance with the country guidelines.

Origin of samples

Only asymptomatic participants were included. Individuals with an acute or uncontrolled illness (for example severe anaemia defined as haemoglobin < 8.0 d/dL, or persistent fever) assessed by a clinician, or who had received anthelmintics in the last four weeks, were excluded. Venous (for serum) and finger prick blood (for RDT) samples were obtained from consenting participants and/or with the consent of the parents/tutors aged between five and 58 years old in five malaria-endemic areas (Bekpou, Taboutou, Allepila, Ouellezue, and Odogueie) in the southern region of Côte d'Ivoire between November and December 2021 (Fig. 1). Fingerprick blood was used to perform the RDT and blood smear on-site and the venous blood samples collected were transported to the local laboratory for serum isolation.

Antigenic and microscopic identification

RDTs (CareStart pLDH(pan), AccessBio, Somerset, NJ, USA, catalog number RMNM-02571) and blood smear techniques (thin and thick smears) were performed using finger prick blood samples following the manufacturer's instructions to determine the infectious status, the parasite density, and the causative malaria infection. Parasite density was calculated using only thick blood smear microscopy readings. Individual parasitaemia was determined by assuming a standard white blood cells (WBCs) count of 8000/μL of blood [19]. Both asexual and gametocyte stages of the parasite were counted, in parallel with WBCs, until reaching 200 WBCs, or 500 WBCs if less than 10 parasites were detected [20]. A slide was considered negative if no asexual or gametocyte stage was found after counting 500 WBCs. Serum samples were subsequently prepared with the venous blood by centrifugation (5634×g for 10 min) and frozen in cryotubes before being transferred to the Institute of

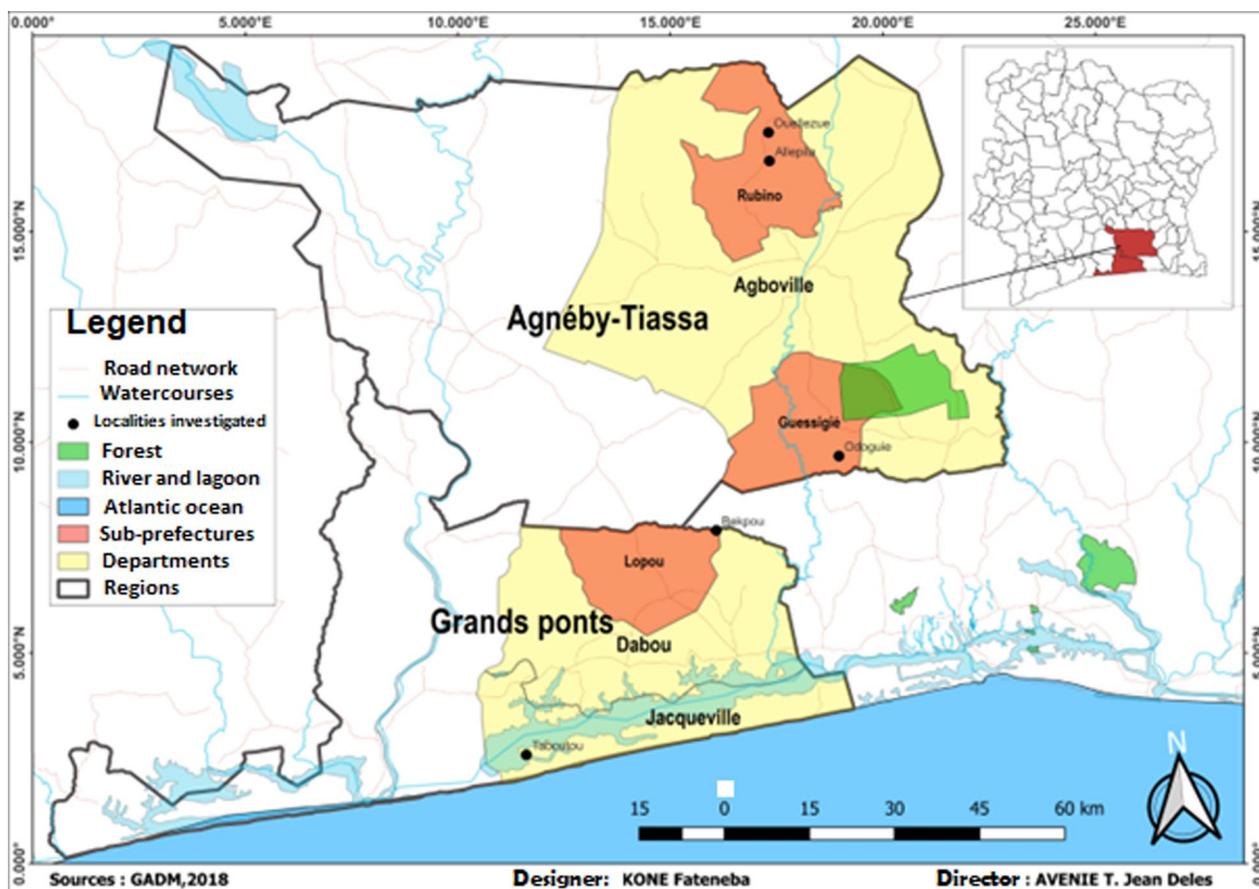


Fig. 1 Map showing the study area: localization of the five malaria-endemic villages (Ouellezue, Allepila, Odoguie, Bekpou, and Taboutou) in Southern Côte d'Ivoire

Medical Microbiology and Hygiene in Germany for further MALDI-TOF MS analysis. Upon receipt in Homburg (Germany), all samples were kept at -20°C pending further examination.

MALDI-TOF MS analysis

Sample preparation and protein extraction

A total of 282 serum samples selected accordingly to the positive RDTs and microscopy results (i.e., only *P. falciparum*, positive for both techniques), were defrosted at room temperature under a biosafety cabinet and measured by MALDI-TOF MS between April and Mai 2022. A protein extraction protocol using formic acid and acetonitrile described elsewhere was adapted and employed in this study [21, 22]. In brief, 50 μL of serum was placed in a tube containing a pinch of Zirconium beads and mixed by vortex. After centrifugation at $18,312\times g$ for 2 min, the supernatant was discarded and 300 μL of HPLC-grade water and 900 μL Absolut ethanol were added to the pellet ($\sim 10\ \mu\text{L}$ of serum + beads) and mixed, followed by another centrifugation step at $18,312\times g$ for

2 min. After discarding the supernatant, 10 μL of 70% (v/v) formic acid and 10 μL acetonitrile were added and mixed by vortex. This procedure was applied in triplicates (i.e., three different days) to each serum sample to ensure reproducibility.

MALDI-target plate preparation and measurements

The protein extracts obtained from the previous section were centrifuged at $18,312\times g$ for 2 min. For each replicate, one μL of the clear supernatant was spotted four different spots onto the MALDI-TOF MS target plate (Bruker Daltonics, Bremen, Germany) then allowed to dry completely before overlaying 1 μL of α -cyano-4-hydroxycinnamic acid (CHCA) matrix solution (Bruker Daltonics, Bremen, Germany), composed of saturated CHCA 50% (v/v) of acetonitrile, 2.5% (v/v) of trifluoroacetic acid and 47.5% (v/v) of LC-MS grade water. Each of the four spots was then measured two times to obtain 8 spectra per protein extraction (i.e., 24 spectra per sample). Of note, not all samples in one category were acquired separately from the other category.

Simultaneous acquisition has also occurred on the same plate. A total of 173 *P. falciparum*-positive, and 109 *P. falciparum*-negative sera were analysed using the FlexControl® software version 3.4 (Bruker Daltonics, Bremen, Germany) for spectra acquisition. Bacterial test standard (BTS) (Bruker Daltonics, Bremen, Germany), which is an extract of *Escherichia coli* that is spiked with two high molecular weight proteins, was used to calibrate the mass spectrometer. Dried MALDI plates were then placed into the Microflex LT Mass Spectrometer (Bruker Daltonics) for measurements.

MALDI-TOF MS parameters

Measurements were performed using the AutoXecute algorithm implemented in the FlexControl® software version 3.4. For each spot, a total of 240 laser shots (40 shots each, in six random positions) were carried out automatically to generate protein mass profiles in linear positive ion mode with a laser frequency of 60 Hz, a high voltage of 20 kV, a pulsed ion extraction of 180 ns, a total of at least 80 shots allowed before automated termination, an evaluation mass range of ± 200 Da, and a minimum intensity threshold of 600 a.u. Mass charge ratios range (m/z) were measured between 2 and 20 kDa.

Spectra treatment, classification and comparison analysis

Pre-processing parameters A total of 6768 (282 samples \times 3 biological replicates \times 8 technical replicates) raw spectra were generated. The dataset was randomly split into training (i.e., 80% of the data, including 131 *P. falciparum*-positive, and 94 *P. falciparum*-negative sera) and test (i.e., 20% of the data, including 42 *P. falciparum*-positive, and 15 *P. falciparum*-negative sera) sets using Python (version 3.13.1) and exported into the online Clover MS Data Analysis® software (<https://platform.clovermsdataanalysis.com>, Clover Bioanalytical Software, Granada, Spain) for further investigation. Default parameters reported elsewhere were used during preprocessing [23, 24]. Variance stabilization and Savitzky–Golay filter (window length 11; polynomial order 3) were applied for smoothing spectra, the baseline was removed using the top-hat filter method (factor 0.02). Replicated spectra were aligned using the following parameters: allowed shift, medium; constant tolerance, 0.2 Da; linear tolerance, 2000 ppm in order to create one average spectrum per sample, which will be used for classification and comparison analysis.

Classification: training using tenfold cross-validation of four ML algorithms A peak matrix was generated from pre-processed spectra and used for comparison analysis. Spectra were normalized using Total Ion Current (TIC) normalization, followed by a “threshold method” (factor

0.01), where peaks with an intensity below 1% of the maximum intensity were not included; a constant tolerance of 0.5 Da; and a linear tolerance of 500 ppm [23, 24].

The spectra were classified into two categories (positive and negative) according to the infectious status and subjected to different classification algorithms in order to distinguish *P. falciparum*-positive sera from *P. falciparum*-negative sera. Supervised (i.e., a guided learning system with training data associated with the pre-defined labels) ML algorithms were used to assess the classification. Four widely used supervised ML algorithms for MALDI-TOF mass spectra analysis [linear super vector machine (SVM) with PCA technique applied using 87 components (95.16% variance), partial least squares-discriminant analysis (PLS-DA) with PCA technique applied using 9 components (43.95% variance), random forest (RF), and light gradient boosting model (GBM)] were applied [25, 26]. These algorithms were trained using k-fold cross-validation ($k=10$) for internal validation. Oversampling was also applied using SMOTE algorithm. A confusion matrix (generating values such as accuracy, specificity, sensitivity, F1-score, positive prediction value or precision, and negative prediction value), as well as the area under receiver operating characteristic (AUROC) curve, were used as performance metrics of the supervised ML algorithms to evaluate their discrimination power, with higher values indicating greater discrimination [26].

Independent test of the two best models The two best models of the four tested during the training phase were evaluated by testing MALDI-spectra generated from a subset of 57 serum samples (42 positives and 15 negatives). The spectra were pre-processed using the same parameters as the training phase.

Results

Prevalence and malaria confirmation

A total of 486 individuals aged between five and 58 years old (mean = 11.2 ± 5.1) participated in this study, including 270 males, 214 females, and 2 people for whom gender information was not reported. RDT and microscopy analyses showed *P. falciparum* infection prevalences of 69.3% (337/486) and 65.6% (319/486) respectively. Most of the positive samples were *P. falciparum* (95.6%; 305/319) with parasite densities ranging between 48 and $34'320$ p/ μ L. *Plasmodium malariae* and mixed *P. malariae/falciparum* were detected in 6/319 (1.9%) and 8/319 (2.5%) positive samples, respectively (Fig. 2; Additional file).

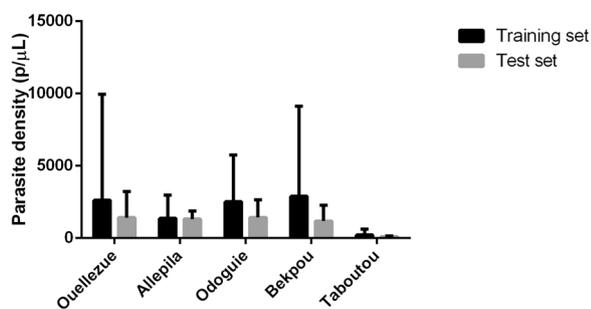


Fig. 2 Average parasite density (p/μL) determined by blood smear in the five malaria-endemic villages (Ouellezue, Allepila, Odogouie, Bekpou, and Taboutou) in Southern Côte d'Ivoire

Spectra analysis and cross-validation performance

Good quality spectra with high intensities were generated for 173 *P. falciparum*-positive (131 training, and 42 test) and 109 *P. falciparum*-negative (94 training, and 15 test) sera. Spectra profiles of both groups are very similar with many common peaks (Fig. 3).

When supervised ML algorithms were employed, tenfold cross-validation results for positive sera showed accuracies of 81.68% (107/131) for LightGBM, 83.97% (110/12) for RF, 74.05% (97/131) for PLS-DA, and 70.23% (92/131) for SVM. For negative sera, accuracies were 80.92% (106/131), 77.86% (102/131), 75.57% (99/131), and 76.34% (100/131), respectively for LightGBM, RF, PLS-DA, and SVM. Global accuracies were

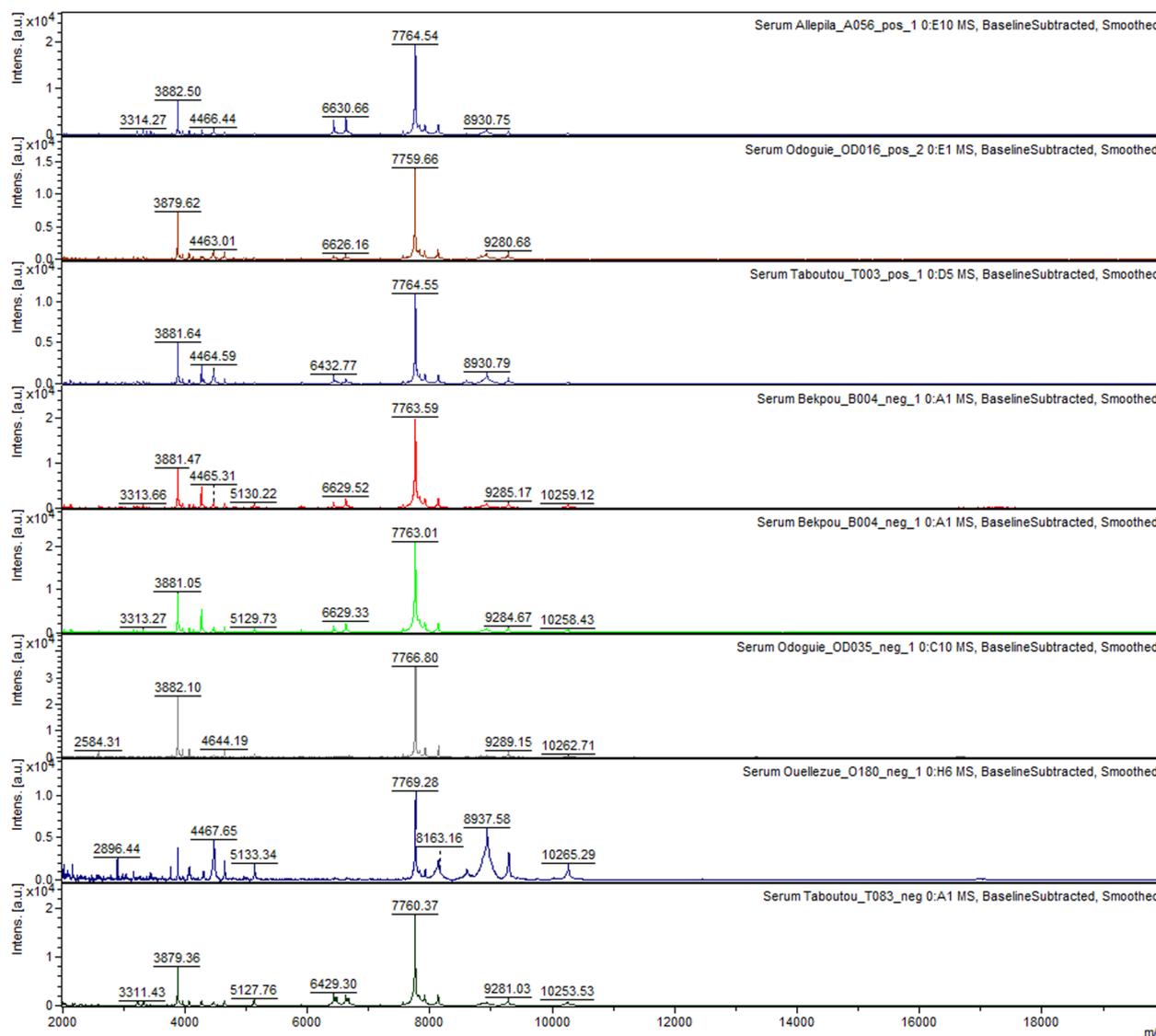


Fig. 3 Representative protein spectra profiles of *P. falciparum*-positive and *P. falciparum*-negative sera

81.3%, 80.92%, 74.81%, and 73.28%, respectively for LightGBM, RF, PLS-DA, and SVM. Other metrics such as sensitivity, specificity, and F1-score (harmonic mean of precision and sensitivity) showed higher values for LightGBM (sensitivity: 81.68%; specificity: 80.92%, F1-score: 81.37%) and RF (sensitivity: 83.97; specificity: 77.86%, F1-score: 81.48%) than PLS-DA (sensitivity: 74.05%; specificity: 75.57%, F1-score: 74.62%) and SVM (sensitivity: 70.23%; specificity: 76.34%, F1-score: 72.44) (Table 1). The receiver operating curves (ROC),

indicating the statistical significance had area under the curve (AUC) values of 0.90, 0.89, 0.83, and 0.76 for LightGBM, RF, PLS-DA, and SVM, respectively (Fig. 4). Distance plots visualization of the different algorithms confirmed the performance metrics as the separation of the clusters is clearer with RF than with PLS-DA and SVM (Fig. 5).

Table 1 Tenfold cross-validation results showing discrimination between *P. falciparum*-positive sera from *P. falciparum*-negative sera with *P. falciparum*-positive sera considered as positive category: scores (in %) obtained with four different classifying algorithms (SVM, PLS-DA, RF, and LightGBM)

| | Accuracy % | Sensitivity % | Specificity % | Error rate % | PPV % | NPV % | F1-score % |
|----------|------------|---------------|---------------|--------------|-------|-------|------------|
| SVM | 73.28 | 70.23 | 76.34 | 26.72 | 74.8 | 71.94 | 72.44 |
| PLS-DA | 74.81 | 74.05 | 75.57 | 25.19 | 75.19 | 74.44 | 74.62 |
| RF | 80.92 | 83.97 | 77.86 | 19.08 | 79.14 | 82.93 | 81.48 |
| LightGBM | 81.3 | 81.68 | 80.92 | 18.7 | 81.06 | 81.54 | 81.37 |

PPV: positive predictive values; NPV: negative predictive values

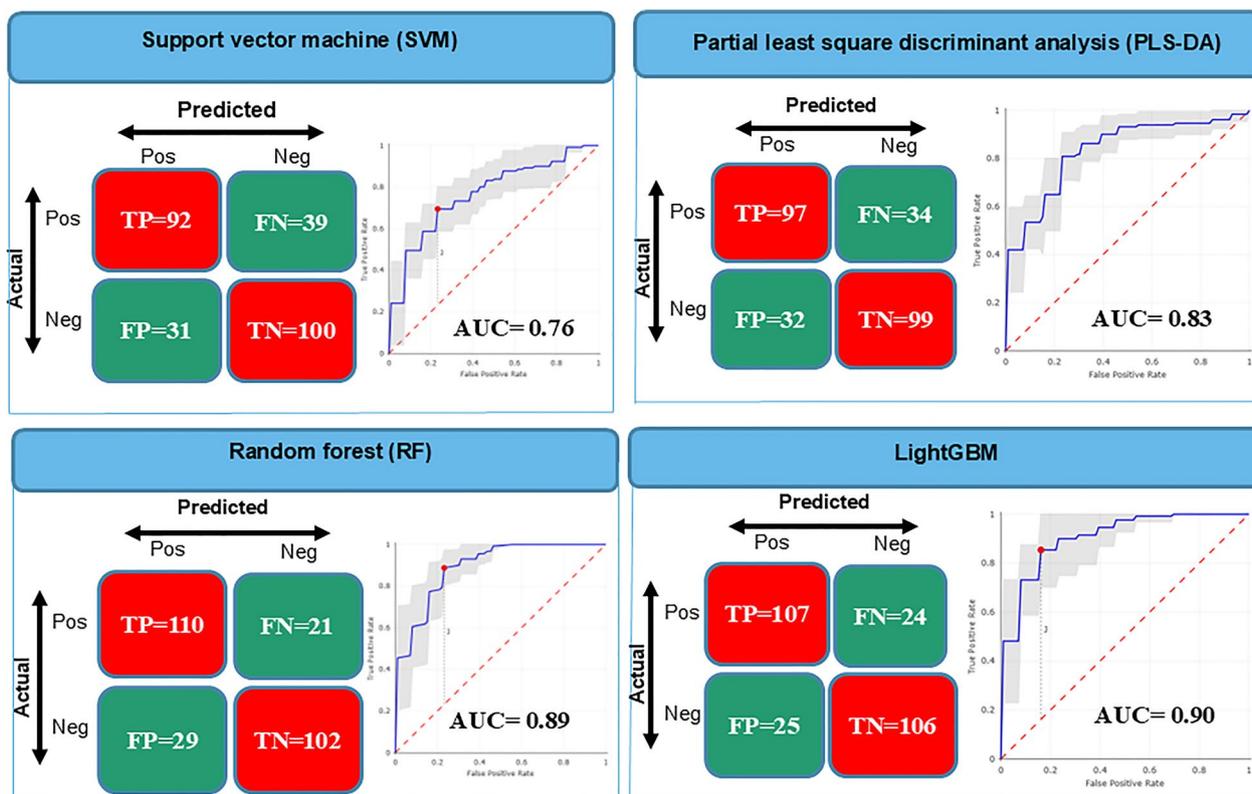


Fig. 4 Confusion matrix and ROC curves showing the classification results of *P. falciparum*-positive and *P. falciparum*-negative sera using MALDI-TOF MS spectra and supervised ML algorithms. Pos: positive; Neg: negative; TP: true positive; TN: true negative; FP: false positive; FN: false negative; ROC: receiver operating characteristic; AUC: area under the curve

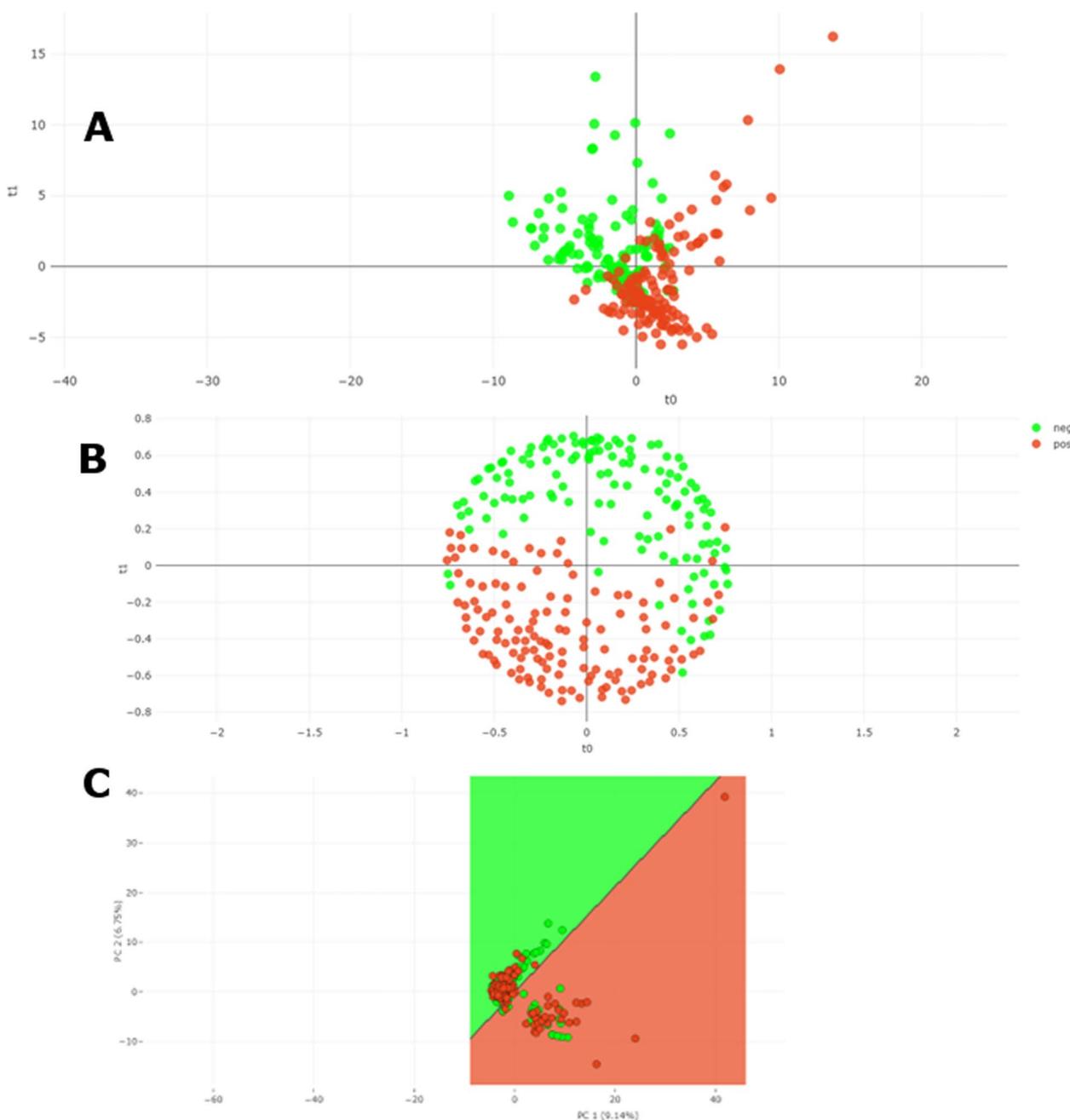


Fig. 5 Distance plots of tenfold cross-validation results using supervised ML algorithms in Clover MS Data Analysis® software, based on peak matrix generated with a threshold of 1% and TIC normalization: distance plots of positive and negative spectra. **A** Two-dimensional view of partial least squares-discriminant analysis (PLS-DA); **B** Two-dimensional view of Random Forest (RF); **C** Two-dimensional view of linear support vector machine (SVM). Pos: positive (in red); Neg: negative (in green)

Independent test of the ML-based models

The best two ML-based models (Light GBM and RF) of the training phase tested with a subset of 57 sera samples (42 positives and 15 negatives) showed accuracies of 90.48% (38/42), and 73.33% (11/15) respectively for positive and negative, for LightGBM algorithm. For the

RF algorithm, accuracy rates were 92.86% (39/42), and 80% (12/15) for positive and negative sera, respectively (Fig. 6). Globally, these two models have accuracy rates of 85.96%, and 89.47% respectively for LightGBM and RF. While specificity performances were 73.33% for LightGBM, and 80% for RF (Table 2).

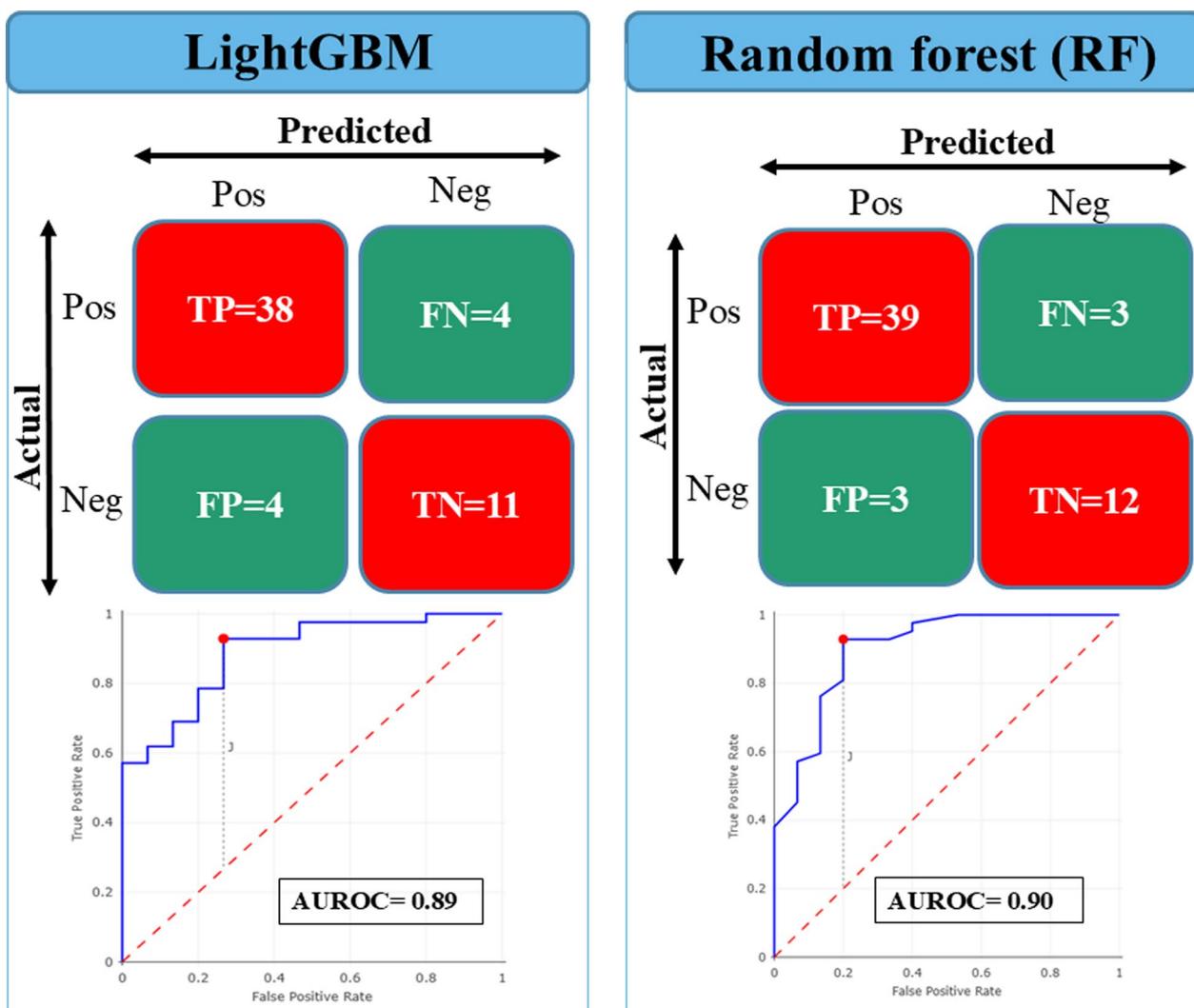


Fig. 6 Confusion matrix and ROC curves showing the independent test results of *P. falciparum*-positive and *P. falciparum*-negative sera using MALDI-TOF MS spectra and supervised ML algorithms (LightGBM, and RF). Pos: positive; Neg: negative; TP: true positive; TN: true negative; FP: false positive; FN: false negative; AUC: area under the curve

Table 2 Independent test results of two ML algorithms (LightGBM, and RF) showing discrimination performances between *P. falciparum*-positive sera from *P. falciparum*-negative sera with *P. falciparum*-positive sera considered as positive category

| | Accuracy % | Sensitivity % | Specificity % | Error rate % | PPV % | NPV % | F1-score % |
|----------|------------|---------------|---------------|--------------|-------|-------|------------|
| LightGBM | 85.96 | 90.48 | 73.33 | 14.04 | 90.48 | 73.33 | 90.48 |
| RF | 89.47 | 92.86 | 80 | 10.53 | 92.86 | 80 | 92.86 |

LightGBM: light gradient boosting model; RF: random forest; PPV: positive predictive values; NPV: negative predictive values

Discussion

MALDI-TOF MS is nowadays a widely used technique in clinical routine microbiology laboratories in high-income countries [27]. Recently, new applications (e.g., parasite

identification) have also been investigated, but have yet to be applied in clinical samples [15, 28]. Combined with ML, MALDI-TOF MS enables a more in-depth study (e.g., comparative analysis of infected vs uninfected specimens), especially for unpurified or complex biological

matrices such as blood or serum samples, represents a step forward in the quest for new, clinically applicable diagnostic methods.

In this study, MALDI-TOF MS associated with ML algorithms allowed distinguishing *P. falciparum*-positive from *P. falciparum*-negative samples using protein spectra profiles generated from serum samples. Despite high similarities of the spectra profiles of both groups, trained classification models using supervised ML algorithms such as LightGBM and RF showed significant abilities to distinguish *P. falciparum*-positive from *P. falciparum*-negative sera with a global accuracy of 81.3%, and 80.9% for LightGBM and RF, respectively. In addition, the ability to correctly classify new independent sera reaches accuracies of 85.96%, and 89.47% for LightGBM and RF, respectively.

In the context of exploring alternative MALDI-based methods to improve malaria diagnosis, many other studies, particularly in entomology, have attempted to identify the different species of the parasite vector (i.e., *Anopheles* mosquitoes), the origin of their blood meal, and if they carry the malaria parasite or not [29–36].

Other studies have attempted malaria screening through the detection of haemozoin using laser desorption time-of-flight (LD-TOF) MS [13, 14, 37]. More recently, Stauning and colleagues have reported the detection of malaria using MALDI-TOF MS by targeting directly the *Plasmodium* parasite extracted from human blood [15]. The authors also reported about 30 *P. falciparum*-specific peaks, and none of them could be found in this present study due to the differences in the methodological approach. Indeed, Stauning and colleagues analyzed *P. falciparum* directly isolated from RBCs of human blood, while this study focused on *P. falciparum*-positive sera. As regards the data available to date, this is the first study to use MALDI-TOF MS combined with ML algorithms for identifying *P. falciparum*-infected human sera from endemic regions.

The present study is limited by the sample pre-processing and the type of material used (human serum), which will be more likely to display host and immune-related proteins, making it difficult to detect *Plasmodium*-specific peaks. Also, the negative group consisted of participants whose serum was negative for *P. falciparum*, but who were not tested for other infections (e.g., bloodstream infection), which could affect specificity if such infections occurred. Other limitations are the restricted origin of the samples (all sera were isolated from the same endemic country, Côte d'Ivoire), the limited size of the dataset, and the lack of species diversity leading to limited specificity results since the prediction models were trained with a positive category (i.e., positive sera) comprising only one type of species (i.e., *P. falciparum*).

Nevertheless, *P. falciparum* is globally the most prevalent species in sub-Saharan Africa, but is also reported elsewhere outside Africa, having these ML-based models developed with such a dataset could be helpful in quickly identifying most malaria cases [3, 38]. However, to correctly deploy appropriate public health strategies based on reliable epidemiological data, and limit the severity and the spread of the parasite by ensuring effective treatment, it is crucial to precisely identify the parasite species [39, 40]. Hence, further investigations are needed (i.e., training more robust models with a dataset including all types of species from various regions in endemic countries) to optimize the method for better performance and enable *Plasmodium* species differentiation. Furthermore, employing alternative approaches such as deep learning (e.g., Siamese Neural Network (SNN), or Triplet Neural Networks (TNN)) combined with high-resolution proteomics would enable intensity variation analysis to detect potential specific biomarkers for malaria [41].

From a clinical diagnostic perspective, more effort should be dedicated to the direct application of MALDI-TOF MS on biological samples from humans (e.g., human blood and/or serum). Indeed, this could be achieved as has already been demonstrated for other diseases such as COVID-19, where plasma proteome fingerprint predicted high (hospitalized) and low-risk (outpatients) cases with 92% accuracy; or serum analysis of COVID-patients in comparison with control cases achieved 99% accuracy [42, 43]. Likewise, MALDI/ML analysis of serum protein fingerprints showed significant differences between liver cancer patients and healthy controls [44].

MALDI-TOF mass spectrometry offers many advantages, including the ability to carry out rapid, and reliable analyses. However, even if sample analysis is inexpensive, the acquisition price remains high (~one hundred thousand euros) compared to microscopy, which is the gold standard for malaria diagnosis.

Conclusion

The results of this study provide evidence of the general applicability of the MALDI-TOF MS/ML combination for the identification of *P. falciparum*-positive human serum stemming from endemic areas in Côte d'Ivoire. Further investigations are needed to optimize this method for better performance in terms of accuracy, sensitivity, and specificity, but also to enable species differentiation and quantification.

Abbreviations

| | |
|--------------|--|
| MALDI-TOF MS | Matrix-Assisted laser desorption/ionization-time of flight mass spectrometry |
| CHCA | α -Cyano-4-hydroxycinnamic acid |
| BTS | Bacterial test standard |
| ML | Machine learning |
| PCA | Principal component analysis |

| | |
|-----------|--|
| SVM | Support vector machine |
| PLS-DA | Partial least squares-discriminant analysis |
| RF | Random forest |
| Light GBM | Light gradient boosting model |
| ROC | Receiver operating characteristic |
| AUROC | Area under receiver operating characteristic |

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12936-025-05362-1>.

Supplementary Material 1. Epidemiological data and *Plasmodium* prevalence using RDTs and microscopy.

Acknowledgements

We acknowledge the community workers as well as authorities in all villages of the health districts of Agboville, Dabou, and Jacqueville in southern Côte d'Ivoire where the study was conducted. We would like to thank Sophie E. Müller from the Institute of Medical Microbiology and Hygiene (IMMH) for her help in using Python software for a random splitting of the data.

Author contributions

JTC, KDS, SLB, BK, and IS designed the study. FK, JTC, KDS, and BK provided the samples and performed the RDT and microscopic identification. FK, LC, and IS performed the MALDI-TOF MS analysis. IS wrote the original draft. BK and SLB acquired the funding and supervised the work. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work has been funded by the Fonds pour la Science, la Technologie et l'Innovation (FONSTI) by the Ministry of Higher Education and Scientific Research of Côte d'Ivoire under the number *P10 MALDI-TOF*.

Availability of data and materials

All data supporting the findings of this study are available within the paper and its Supplementary Information.

Declarations

Ethics approval and consent to participate

Approved by the National Ethical Committee for Natural Science and Health ("Comité National d'Éthique des Sciences de la Vie et de la Santé, CNESVS") of the Ministry of Health, Public Hygiene and Universal Health Coverage ("Ministère de la Santé, de l'Hygiène Publique et de la Couverture Maladie Universelle") under the number *056-21/MSHP/CNESVS-km*.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Institute of Medical Microbiology and Hygiene, Saarland University, Homburg, Germany. ²UFR Biosciences, Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire. ³Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Côte d'Ivoire. ⁴Swiss Tropical and Public Health Institute, Allschwil, Switzerland. ⁵University of Basel, Basel, Switzerland. ⁶Helmholtz Institute for Pharmaceutical Research Saarland, Saarbrücken, Germany. ⁷Université Pélégro Gon Coulibaly, Korhogo, Côte d'Ivoire.

Received: 28 September 2024 Accepted: 1 April 2025

Published online: 18 April 2025

References

- Garcia LS. Malaria. *Clin Lab Med*. 2010;30:93–129.
- Milner DA. Malaria pathogenesis. *Cold Spring Harb Perspect Med*. 2018;8:a025569.
- WHO. World Malaria Report 2023. Geneva, World Health Organization. <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023>. Accessed 27 June 2024.
- White NJ, Pukrittayakamee S, Hien TT, Faiz MA, Mokuolu OA, Dondorp AM. Malaria. *Lancet*. 2014;383:723–35.
- CDC. DPDx: Laboratory identification of parasites of public health concern. <https://www.cdc.gov/dpdx/malaria/index.html>. Accessed 16 Dec 2024.
- CDC. Malaria diagnostic tests. <https://www.cdc.gov/malaria/hcp/diagnosis-testing/malaria-diagnostic-tests.html>. Accessed 31 July 2024.
- Zheng Z, Cheng S. Advances in molecular diagnosis of malaria. *Adv Clin Chem*. 2017;80:155–92.
- Calderaro A, Piccolo G, Chezzi C. The laboratory diagnosis of malaria: a focus on the diagnostic assays in non-endemic areas. *Int J Mol Sci*. 2024;25:695.
- Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P-E, Rolain JM, et al. Ongoing Revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Infect Dis*. 2009;49:543–51.
- Angeletti S. Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. *J Microbiol Methods*. 2017;138:20–9.
- De Carolis E, Vella A, Vaccaro L, Torelli R, Spanu T, Fiori B, et al. Application of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *J Infect Dev Ctries*. 2014;8:1081–8.
- Wieser A, Schneider L, Jung J, Schubert S. MALDI-TOF MS in Microbiological diagnostics-identification of microorganisms and beyond (Mini Review). *Appl Microbiol Biotechnol*. 2012;93:965–74.
- Scholl PF, Kongkasuriyachai D, Demirev PA, Feldman AB, Lin JS, Sullivan DJ, et al. Rapid detection of malaria infection in vivo by laser desorption mass spectrometry. *Am J Trop Med Hyg*. 2004;71:546–51.
- Christner M, Frickmann H, Klupp E, Rohde H, Kono M, Tannich E, et al. Insufficient sensitivity of laser desorption-time of flight mass spectrometry-based detection of hemozoin for malaria screening. *J Microbiol Methods*. 2019;160:104–6.
- Stauning MA, Jensen CS, Staalsøe T, Kurtzhals JAL. Detection and quantification of *Plasmodium falciparum* in human blood by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: a proof of concept study. *Malar J*. 2023;22:285.
- Marletta S, L'Imperio V, Eccher A, Antonini P, Santonicco N, Girolami I, et al. Artificial intelligence-based tools applied to pathological diagnosis of microbiological diseases. *Pathol Res Pract*. 2023;243:154362.
- Ikerionwu C, Ugwuishiwu C, Okpala I, James I, Okoronkwo M, Nnadi C, et al. Application of machine and deep learning algorithms in optical microscopic detection of *Plasmodium*: a malaria diagnostic tool for the future. *Photodiagn Photodyn Ther*. 2022;40:103198.
- Silka W, Wiczorek M, Silka J, Woźniak M. Malaria detection using advanced deep learning architecture. *Sensors (Basel)*. 2023;23:1501.
- Moody A. Rapid diagnostic tests for malaria parasites. *Clin Microbiol Rev*. 2002;15:66–78.
- Bassa FK, Ouattara M, Silué KD, Adiassan LG, Baikoro N, Koné S, et al. Epidemiology of malaria in the Taabo Health and Demographic Surveillance System, South-Central Côte d'Ivoire. *Malar J*. 2016;15:9.
- Abdrabou AMM, Sy I, Bischoff M, Arroyo MJ, Becker SL, Mellmann A, et al. Discrimination between hypervirulent and non-hypervirulent ribotypes of *Clostridioides difficile* by MALDI-TOF mass spectrometry and machine learning. *Eur J Clin Microbiol Infect Dis*. 2023;42:1373–81.
- Feucherolles M, Nennig M, Becker SL, Martiny D, Losch S, Penny C, et al. Combination of MALDI-TOF mass spectrometry and machine learning for rapid antimicrobial resistance screening: the case of *Campylobacter* spp. *Front Microbiol*. 2021;12:804484.
- Candela A, Arroyo MJ, Sánchez-Molleda Á, Méndez G, Quiroga L, Ruiz A, et al. Rapid and reproducible MALDI-TOF-based method for the detection of vancomycin-resistant *Enterococcus faecium* using classifying algorithms. *Diagnostics (Basel)*. 2022;12:328.
- Ebersbach JC, Sato MO, de Araújo MP, Sato M, Becker SL, Sy I. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

- for differential identification of adult *Schistosoma* worms. *Parasit Vectors*. 2023;16:20.
25. Weis CV, Jutzeler CR, Borgwardt K. Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. *Clin Microbiol Infect*. 2020;26:1310–7.
 26. Blázquez-Sánchez M, Guerrero-López A, Candela A, Belenguer-Llorens A, Moreno JM, Sevilla-Salcedo C, et al. Rapid, automatic typing of *Clostridioides difficile* ribotypes using MALDI-TOF MS. *bioRxiv* (preprint). 2024.
 27. Seng P, Rolain J-M, Fournier PE, La Scola B, Drancourt M, Raoult D. MALDI-TOF-mass spectrometry applications in clinical microbiology. *Future Microbiol*. 2010;5:1733–54.
 28. Calderaro A, Chezzi C. MALDI-TOF MS: a reliable tool in the real life of the clinical microbiology laboratory. *Microorganisms*. 2024;12:322.
 29. Karisa J, Ominde K, Tuwei M, Bartilol B, Ondieki Z, Musani H, et al. Utility of MALDI-TOF MS for determination of species identity and blood meal sources of primary malaria vectors on the Kenyan Coast. *Wellcome Open Res*. 2023;8:151.
 30. Nabet C, Kone AK, Dia AK, Sylla M, Gautier M, Yattara M, et al. New assessment of *Anopheles* vector species identification using MALDI-TOF MS. *Malar J*. 2021;20:33.
 31. Briolant S, Costa MM, Nguyen C, Dusfour I, Pommier de Santi V, Girod R, et al. Identification of French Guiana Anopheline mosquitoes by MALDI-TOF MS profiling using protein signatures from two body parts. *PLoS ONE*. 2020;15: e0234098.
 32. Chaumeau V, Piarroux M, Kulabkeeree T, Sawasdichai S, Inta A, Waththana-worawit W, et al. Identification of Southeast Asian *Anopheles* mosquito species using MALDI-TOF mass spectrometry. *PLoS ONE*. 2024;19: e0305167.
 33. Niare S, Berenger J-M, Dieme C, Doumbo O, Raoult D, Parola P, et al. Identification of blood meal sources in the main African malaria mosquito vector by MALDI-TOF MS. *Malar J*. 2016;15:87.
 34. Fall FK, Diarra AZ, Bouganali C, Sokhna C, Parola P. Using MALDI-TOF MS to identify mosquitoes from Senegal and the origin of their blood meals. *Insects*. 2023;14:785.
 35. Tandina F, Niare S, Almeras L, Davoust B, Doumbo OK, Raoult D, et al. Identification of mixed and successive blood meals of mosquitoes using MALDI-TOF MS protein profiling. *Parasitology*. 2020;147:329–39.
 36. Laroche M, Almeras L, Pecchi E, Bechah Y, Raoult D, Viola A, Parola P. MALDI-TOF MS as an innovative tool for detection of *Plasmodium* parasites in *Anopheles* mosquitoes. *Malar J*. 2017;16:5.
 37. Nyunt M, Pisciotta J, Feldman AB, Thuma P, Scholl PF, Demirev PA, et al. Detection of *Plasmodium falciparum* in pregnancy by laser desorption mass spectrometry. *Am J Trop Med Hyg*. 2005;73:485–90.
 38. Calderaro A, Piccolo G, Montecchini S, Buttrini M, Rossi S, Dell'Anna ML, et al. High prevalence of malaria in a non-endemic setting: comparison of diagnostic tools and patient outcome during a four-year survey (2013–2017). *Malar J*. 2018;17:63.
 39. Daily JP, Minuti A, Khan N. Diagnosis, Treatment, and prevention of malaria in the US: a review. *JAMA*. 2022;328:460–71.
 40. Lover AA, Baird JK, Gosling R, Price RN. Malaria elimination: time to target all species. *Am J Trop Med Hyg*. 2018;99:17–23.
 41. Merchan F, Contreras K, Gittens RA, Loaiza JR, Sanchez-Galan JE. Deep metric learning for the classification of MALDI-TOF spectral signatures from multiple species of neotropical disease vectors. *Artif Intell Life Sci*. 2023;3:100071.
 42. Lazari LC, Ghilardi FDR, Rosa-Fernandes L, Assis DM, Nicolau JC, Santiago VF, et al. Prognostic accuracy of MALDI-TOF mass spectrometric analysis of plasma in COVID-19. *Life Sci Alliance*. 2021;4: e202000946.
 43. Yan L, Yi J, Huang C, Zhang J, Fu S, Li Z, et al. Rapid detection of COVID-19 using MALDI-TOF-based serum peptidome profiling. *Anal Chem*. 2021;93:4782–7.
 44. Park H-G, Jang K-S, Park H-M, Song W-S, Jeong Y-Y, Ahn D-H, et al. MALDI-TOF MS-based total serum protein fingerprinting for liver cancer diagnosis. *Analyst*. 2019;144:2231–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.