

RESEARCH

Open Access



Exploring near-infrared spectroscopy ability to predict the age and species of *Anopheles gambiae* sensu lato mosquitoes from different environmental conditions in Burkina Faso

Nicaise D. C. Djègbè^{1,2*}, Dari F. Da^{1*}, Bernard M. Somé¹, Lawata Inès G. Paré^{1,2}, Fatoumata Cissé^{1,2}, Jacques Kaboré², Thomas S. Churcher³ and Roch K. Dabiré¹

Abstract

Background Near infrared spectroscopy (NIRS) has shown ability in previous studies to predict age and species of laboratory-reared and wild mosquitoes with moderate to high accuracy. To validate the technique as a routine tool, it is necessary to assess NIRS accuracy on these variables under different environmental conditions susceptible to affect the mosquito cuticle and interfere with the machine accuracy. This study investigated the influence of environmental conditions on NIRS accuracy to determine the age and species of *Anopheles gambiae* sensu lato (*s.l.*).

Methods Environmental conditions of three important seasonal periods in Burkina Faso covering the onset, the peak and the end of the rainy season were mimicked in the laboratory using incubators. Emerged *An. gambiae* s.s. and *An. coluzzii* from laboratory colonies were reared in each period using temperature and relative humidity for predicting mosquito species by NIRS. Wild *An. gambiae* s.l. (n = 3788) were caught during the 3 different periods described above and analysed by NIRS to predict *Anopheles* species. Furthermore, first generation of wild *Anopheles* (n = 1014) was used to assess NIRS ability to classify mosquito age in each environmental condition. All data analysis were performed using a binomial logistic regression model.

Results NIRS discriminated between laboratory-reared *Anopheles* with 83% of accuracy independently of any environmental condition. Similar trend was found in wild-caught *Anopheles*. NIRS accuracies varied slightly in laboratory *Anopheles* (77–85%) and more strongly in their field counterparts (67–84%). In both cases, models developed from the season of interest were more accurate than models trained with insectary conditions or from a different period of the year, indicating temperature and humidity can impact NIRS accuracy. Models derived from laboratory-mosquitoes reared under fluctuating environmental conditions predicted field-derived mosquito species with a low accuracy (59%). Models trained on varying conditions reliably classified age into two categories (< 9 days or ≥ 9 days, 79–84% accuracy).

*Correspondence:

Nicaise D. C. Djègbè
dndjegbe@yahoo.fr

Dari F. Da
dafrenick@yahoo.fr

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



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusion NIRS was able to predict *An. gambiae s.l.* species and classified age into two categories under different environmental conditions with modest accuracy. Models trained using wild mosquitoes from one season could predict species in wild mosquitoes from a different season, though with slightly lower accuracy.

Keywords Near infrared spectroscopy, *Anopheles*, Temperature, Relative humidity, Environmental conditions, Age and species determination

Background

Malaria is the most widespread parasitic disease transmitted by *Anopheles* mosquitoes. Insecticide-treated nets (ITNs) and indoor residual spraying (IRS) have been the most important disease control interventions in endemic countries in Africa [1]. Vector control interventions aim to reduce *Anopheles* abundance or/and reduce their lifespan to limit *Plasmodium* transmission [2]. Entomological monitoring is needed to estimate the ability of the mosquito population to transmit the disease and evaluate control interventions [2]. These measurements would benefit from an understanding of the age and species of *Anopheles* vector populations. Current methods of estimating these entomological parameters are limited [3, 4]. For instance, molecular analysis methods to determine *Anopheles* species are technically laborious, require reagents, time-consuming and qualified workers, limiting the number of mosquitoes which can be processed. Also, age determination techniques based on observations of morphological changes in the reproductive system of female mosquitoes are laborious and require skilled workers, reducing their use on a large scale. To minimize these shortcomings, a variety of complementary or alternative methods are under development [3, 5–7]. NIRS, an innovative tool has the potential in predicting the entomological parameters of malaria such as age and species. This method has successfully identified laboratory and field *Anopheles* species [3, 8], age-grade [3, 7–10] and detected mosquito *Plasmodium* infection status [11, 12]. Compared to other methods, NIRS is fast, non-destructive, reagent free, waste free and requires less qualified personnel [4, 6, 11, 12]. While the technique appears promising, there are some challenges that must be addressed before NIRS can be used as a routine tool in the field [13]. For instance, the machine learning models need to be expanded using samples from a much wider diversity of mosquitoes, as models trained with mosquito data from one *Anopheles* population fail to reliably predict the characteristic of interest in other populations [10]. Models trained with spectra generated from laboratory-reared mosquitoes have been shown to predict field-mosquitoes age with relatively low accuracy [12]. Other studies have reported that the physiological status [14] and diet [15, 16] affect NIRS

ability to accurately predict mosquito age. NIRS relies on the property of hydrocarbons within the mosquito cuticle to quantitatively measure changes in organic compounds such as C–H, O–H, N–H, and C–O functional. It detects these changes through the rotation, bending and stretching of these functional groups following exposure to light of the near infrared spectrum. The lipid structure of the arthropod cuticle depends on their environmental conditions [17], leading to NIRS spectral variations.

Vector control programmes collect mosquitoes in different locations at different periods of the year. Each of these contexts is characterized by particular environmental conditions such as temperature, relative humidity, photoperiodicity and ecological factors which may be able to induce morphological, physiological and biochemical changes within the mosquitoes. Recent work has reported that *An. coluzzii* and *An. gambiae s.l.* grow faster and live shorter at the peak of the rainy season (August) than their counterparts at the onset of the dry season (December) [18]. Furthermore, mosquitoes of *An. gambiae* complex reared under conditions typical of the onset of the dry season (December–February) have a larger body size than those reared under conditions from the late dry season (April–May) and those of the peak of the rainy season (August) in the Sahel region [19, 20]. It is thought that *An. gambiae* reared under dry season conditions have a higher hydrocarbon content compared to rainy season specimens [20]. These same authors reported higher chain lengths in *n*-alkanes in response to the dry season. These findings suggest seasonality induced cuticular hydrocarbons reorganizations within mosquitoes which could potentially influence NIRS accuracy. This could be a challenge for predicting sample characteristics for a given condition on the basis of models derived from another environmental condition. Verifying the dependence of NIRS predictions to environmental conditions is an important step for implementing NIRS technique as a routine diagnostic tool. This study aims to explore NIRS ability to predict species and age-grade of *An. gambiae s.l.* mosquitoes under different temperature and humidity in Burkina Faso at three time periods (the onset, peak and end of the rainy season) and compare these to mosquitoes reared under insectary conditions.

Methods

Experimental design

This study evaluated NIRS accuracy to identify species and age-grade of *An. gambiae s.l.* population reared under different environmental conditions by mimicking the important malaria seasons observed in Burkina Faso. In this country, entomological variations (vector density, mosquito longevity) occur at three different periods of the year [21]. June period is the onset of the rainy season characterized by significant daily fluctuations in air temperature and a considerable drop in relative humidity. During this period the vectors return from aestivation and the malaria transmission repeats. The peak of the rainy season (August) is the second important period associated with low variations of temperature and relative humidity. This period is characterized by a high density of mosquitoes associated with a high transmission of *Plasmodium*. The third important period is the end of rainy season/onset of the dry season (December) during which important daily fluctuations in air temperature and a significant drop in relative humidity were observed, marks the mosquito aestivation period [22]. Wild *Anopheles* were sampled at different time points of the year (onset, peak and end of rainy season or onset of dry season) for age determination or species identification using the NIRS technique. *Anopheles coluzzii* and *An. gambiae s.s.*, two sibling species of the Gambiae complex, used in this study, significantly contribute to malaria transmission in Burkina Faso [23]. In the laboratory, these species were separately reared, while the field ones were identified using conventional PCR [26] after analysis with NIRS.

Environmental parameters simulation inside the incubator

In the IRSS laboratory, three incubators (Sanyo MLR 315H, Sanyo Electric Co., Osaka, Japan) were used to mimic cyclically environmental conditions of June, August and December for mosquito rearing [19, 22]. For

each period, average hourly temperature and relative humidity (RH) (Fig. 1) cycles were scheduled inside the climate chambers based on data recorded hourly in the village of Bama in previous work [19]. To mimic natural daily climate fluctuations, a 12-step cycle of temperature and RH was set up in the climate chambers. A photoperiod of light/dark 12:12 h cycle in all climate chambers was performed for all experiments due to the low variations in photoperiod in the tropics (about 1 h 15 min day length variation between the solstices in Bobo-Dioulasso, Burkina-Faso). In another incubator, insectary conditions were programmed (temperature $27^{\circ}\text{C} \pm 2$, relative humidity $70\% \pm 5$ with light/dark 12:12 h cycle) and acted as a control.

Mosquito rearing inside the incubator for NIRS analysis

Species identification of laboratory *An. gambiae s.s.* and *An. coluzzii*

In the first part of the experiment, the effect of environmental changes on NIRS accuracy using adult mosquitoes was tested. Laboratory-reared female *An. gambiae s.s.* and *An. coluzzii* were sourced from an outbred colony established in 2018. Since changes in mosquito cuticles happen during adulthood [20], larvae were developed under controlled conditions (temperature $27^{\circ}\text{C} \pm 2$, relative humidity $70\% \pm 5$ with light/dark 12:12 h) and then the resulting adult females were exposed under four environmental conditions. This approach could allow changes in the cuticular structure of mosquitoes exposed to different environmental conditions. Briefly, eggs laid by females of *An. gambiae s.s.* and *An. coluzzii* were separately hatched in trays containing water in the insectary. The resulting larvae were fed *ad libitum* with fish food (Tetramin® Baby) and pupae were pipetted into a plastic cup half-filled with spring water and then placed in cages covered with untreated mosquito netting for emergence. Emerged female mosquitoes were organized by groups in cages and then kept in incubators where environmental

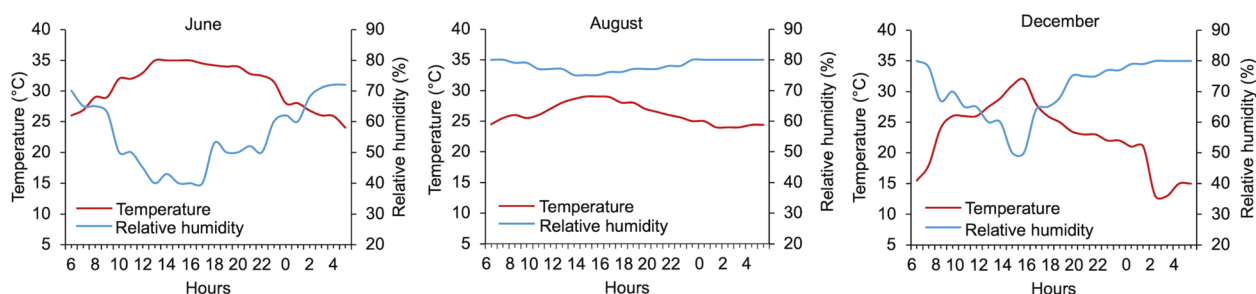


Fig. 1 Environmental parameters of temperature (red line) and relative humidity (blue line) variation of the different periods of the year. The hourly mean variation of temperature and relative humidity recorded during the months of June (onset of the rainy season, at the left), August (peak of the rainy season, at the middle) and December (onset of the dry season, at the right) in the village of Bama (western Burkina Faso). The values recorded, were reproduced in the incubators to simulate these environmental conditions

conditions were programmed as designed below (Fig. 2). Mosquitoes were fed *ad libitum* with 10% glucose solution that was changed daily to avoid mosquitoes contaminated by mould. For a given generation of adult's mosquitoes, approximately 60 females of *An. gambiae s.s.* and *An. coluzzii* in each experimental group were sampled at different time point (3, 6, 9, 12, 15 and 18 days post-emergence) for NIRS analysis to identify *Anopheles* species. To minimize any interference effect and increase sample size, the experiment was repeated across four consecutive mosquito generations, resulting in four replicates. To avoid any effect of the cage, the positions were changed over time in climate chambers. In addition, climate chambers were also changed between conditions to avoid any interference of the incubator in data analysis.

Species identification of field-derived *An. gambiae s.l.*

The influence of different environmental conditions on NIRS accuracy in predicting *Anopheles* species in the field was evaluated and compared with laboratory-reared sample predictions. Wild mosquitoes were collected in the villages surrounding Bobo-Dioulasso at the periods with similar conditions which were reproduced in the incubators: June, July–September, October–November corresponding respectively to the onset, the peak and the end of rainy season. For each period, mosquitoes were caught early in the morning in the living rooms dwellings using mouth aspirators [24]. *Anopheles gambiae s.l.* morphologically identified based on Giles and Coetzee keys [25], were transferred at IRSS insectary in standard conditions (temperature 27 ± 2 °C, $70 \pm 5\%$ relative humidity, 12:12 light/dark). To prevent interference from blood meals in the NIRS spectra, unfed *Anopheles* mosquitoes were immediately scanned using the machine. After each mosquito scanning, the carcass was analysed for species determination using conventional PCR protocol [26].

Wild-mosquito age determination

First generation (F_1) of wild female *An. gambiae s.l.* were used in this experiment. Blood-engorged *An. gambiae s.l.* were caught during rainy season early in the morning using a mouth aspirator [24] in the living rooms of human in the villages of Bama and Soumouso. As reported in previous studies [27, 28], almost 99% of Bama's mosquitoes are *An. coluzzii* while those from Soumouso include three major species (*An. gambiae s.s.*, *An. arabiensis* and *An. coluzzii*). Collected mosquitoes were kept live in cages according to the locality, fed with 10% glucose solution during couple days for eggs laying as described in previous work [18]. Mosquitoes were killed after oviposition, individually preserved in Eppendorf tubes and then analysed by conventional PCR for species identification [26]. Eggs collected from F_0 females of each locality, were organized by pools and then incubated in different conditions (insectary, August and December) to obtain F_1 generation mosquitoes. The adult female mosquitoes were kept in the same conditions, fed with 10% glucose sugar solution and, around 30 *Anopheles* were sampled every 3 days (from day 3 to 18 day or day 21 according to environmental conditions) for NIRS analysis to determine *Anopheles* age. Two replicates were performed for this experiment.

Mosquito scanning

Mosquitoes from each experiment were killed with chloroform vapor and their cephalo-thorax scanned using a LabSpec4 Standard-Res *i* (standard resolution, integrated light source) near-infrared spectrometer and a bifurcated reflectance probe mounted 2 mm from a spectral on white reference panel (ASD Inc.). RS3 spectral acquisition software (ASD Inc.) which averages spectra from 20 scans, was used to record absorbance at 2151 wavelengths from 350 to 2500 nm of the electromagnetic

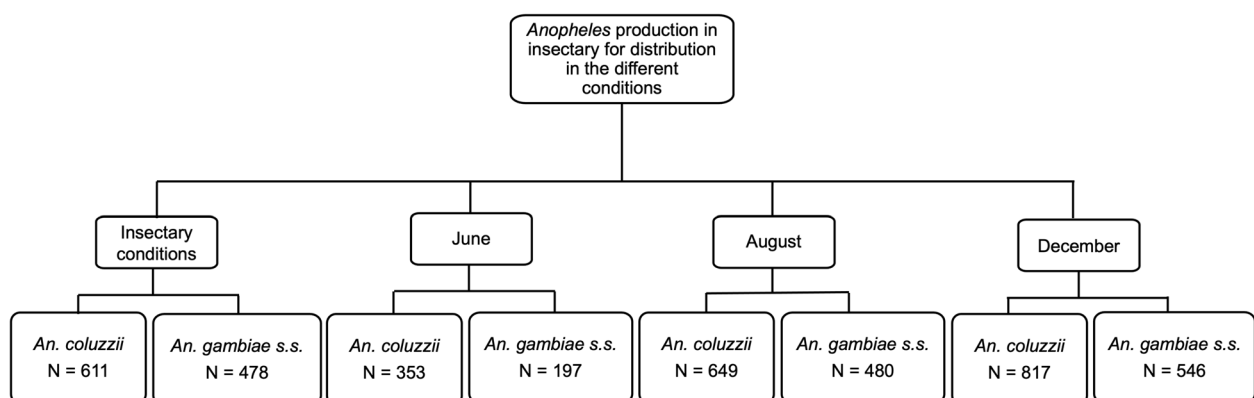


Fig. 2 Experimental design and sample size of *Anopheles* per environmental condition. N indicates the number of survival mosquitoes which were scanned by NIRS. These mosquitoes were used to explore mosquito species

spectrum. All specimens were scanned on the two sides focusing the light probe on the head and thorax region.

Data analysis

Data analyses were conducted using R software version 4.0.2. The wavelengths from 700 to 2350 nm were used to remove the excess noise arising from the sensitivity of the spectrometer often caused by too much light in the visible region. PLS (partial least squares) regression was used to analyse the relationship between spectra and entomological parameters (age and species) using packages downloaded from <https://github.com/pmesperanca/mlevcm> [29]. The age of each *Anopheles* was predicted using its spectrum. Additionally, a machine learning approach was used to develop binomial logistic regression model for performing *Anopheles* species identification and age classification (<9 days or ≥ 9 days) [12, 29]. In this analysis, two response classes $y=0$ and $y=1$ were attributed to *An. gambiae* s.s. and *An. coluzzii*, respectively. The same procedure was used when classifying ages into two groups: $y<9$ days for young mosquitoes and $y\geq 9$ days for old mosquitoes. All analyses were performed using simple models without spectra smoothing or penalized estimation of the coefficient function. For reducing the overfitting, data were split into three subsets: 60%, 20% and 20% for training, validation and testing (estimation of the generalized error) the model, respectively. Sixty percent (60%) of sample was used to fit the model to known (age or species) sample using different numbers of PLS components. The first twenty percent (20%) corresponded to the sample used to select an optimum number of PLS component for the model. The second twenty percent (20%) of sample was used to assess the final model against a blinded subset of data. The number of principal components was allowed to vary between 2 and 50. This process was repeated 50 times to reduce errors. The observations were balanced by random sampling of the total number of mosquitoes analysed in all models. Model accuracy was obtained through the value of the area under the receiver operating characteristic (ROC) curve (AUC). A value of AUC close to or equal 1 indicates a good accuracy. The accuracies from different models were compared using Fisher exact test. Since mosquitoes were reared in different conditions, a calibration model developed using samples from one condition was used to predict the species or age of mosquitoes from another condition.

Results

NIRS ability to identify *An. gambiae* s.l. species from different environmental conditions

A total of 4131 laboratory derived *Anopheles* of various ages and reared under different environmental conditions

were used to assess NIRS accuracy. In the field, 4343 mosquitoes were collected, with 12.8%, 41.5%, and 45.7% of *An. arabiensis*, *An. gambiae* s.s. and *An. coluzzii* respectively. Due to the low sample size of *An. arabiensis*, only spectra derived from *An. gambiae* s.s. and *An. coluzzii* were included in the data analysis. The preliminary analyses revealed distinct spectral profiles between the two groups of mosquitoes (laboratory-mosquitoes versus field mosquitoes) (Fig. 3).

The binomial logistic classification model trained on mosquitoes reared across all conditions in laboratory discriminated between *An. gambiae* s.s. and *An. coluzzii* with 83% of accuracy (*An. gambiae* s.s. = 84%; *An. coluzzii* = 82%; Fig. 4a). When a calibration model was generated from laboratory mosquitoes reared in one condition to predict those in other conditions, NIRS accuracies varied from 77 to 85% (Fig. 5a). In most situations, models trained with mosquitoes from the same condition were more accurate, i.e. models calibrated with mosquitoes reared under August conditions were the most reliable in predicting mosquito species reared under August conditions. The lowest accuracy was obtained when models derived from *Anopheles* reared under June conditions were used to predict *Anopheles* species from other rearing conditions. However, these models were also trained on the smallest number of mosquitoes. In contrast, the high accuracy was found in August conditions when the model derived from mosquitoes in these conditions were used to predict *Anopheles* species from another reared conditions (Fig. 5a). The overall accuracies of different comparisons using a predictive model from one condition to identify mosquito species across all conditions resulted no difference for distinguishing *An. gambiae* s.l. species ($p=1$). Altogether, models developed from insectary-reared mosquitoes and those from the peak of the rainy season were suitable for predicting species under different environmental conditions (Fig. 5a). Details of *Anopheles* species classification accuracy (*An. gambiae* s.s. and *An. coluzzii*) under each environmental condition were shown in Additional file 1: Table S1.

The analysis was extended to investigate whether these finding could be generalized to wild *Anopheles*. The model developed using wild mosquitoes across all periods classified *An. gambiae* s.s. and *An. coluzzii* with 84% accuracy (*An. gambiae* s.s. 86%; *An. coluzzii* 82%, Fig. 4b). This accuracy was not significantly different from that obtained with laboratory-reared *Anopheles* (84% versus 83%; $p=0.23$). Again, accuracies varied considerably between 67 and 84% when the model trained with samples collected in one period was used to predict samples from another period (Fig. 5b). Indeed, the model derived from mosquitoes at the end of the rainy season generated low accuracy (67%) in predicting onset-season

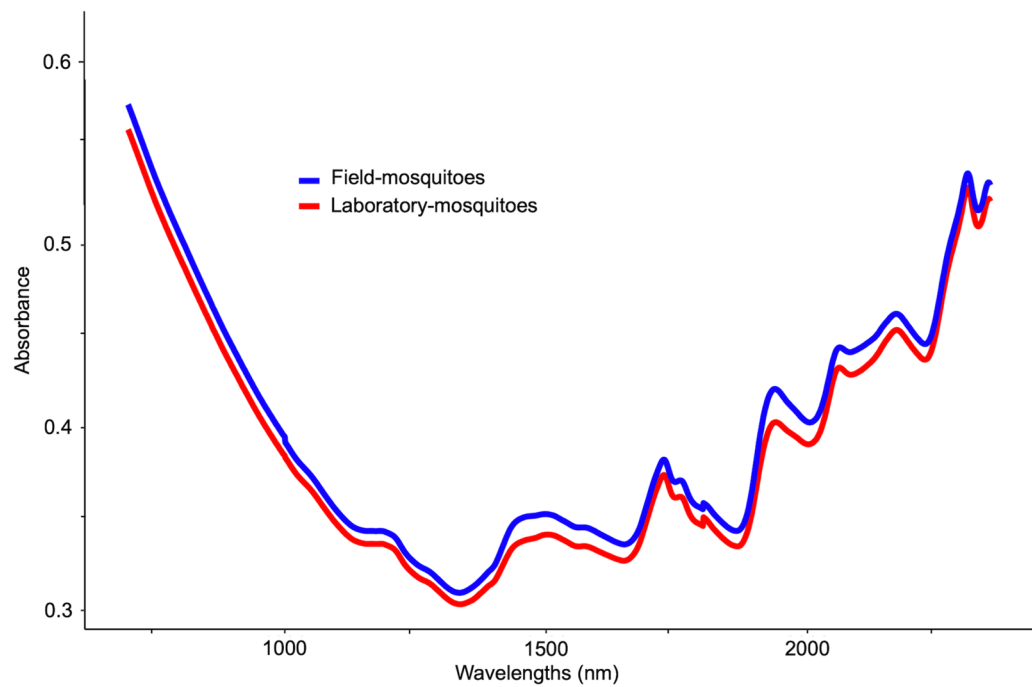


Fig. 3 Average spectra of laboratory-derived mosquitoes (red curve) and field-derived mosquitoes (blue curve); interval of wavelengths [700, 2350]

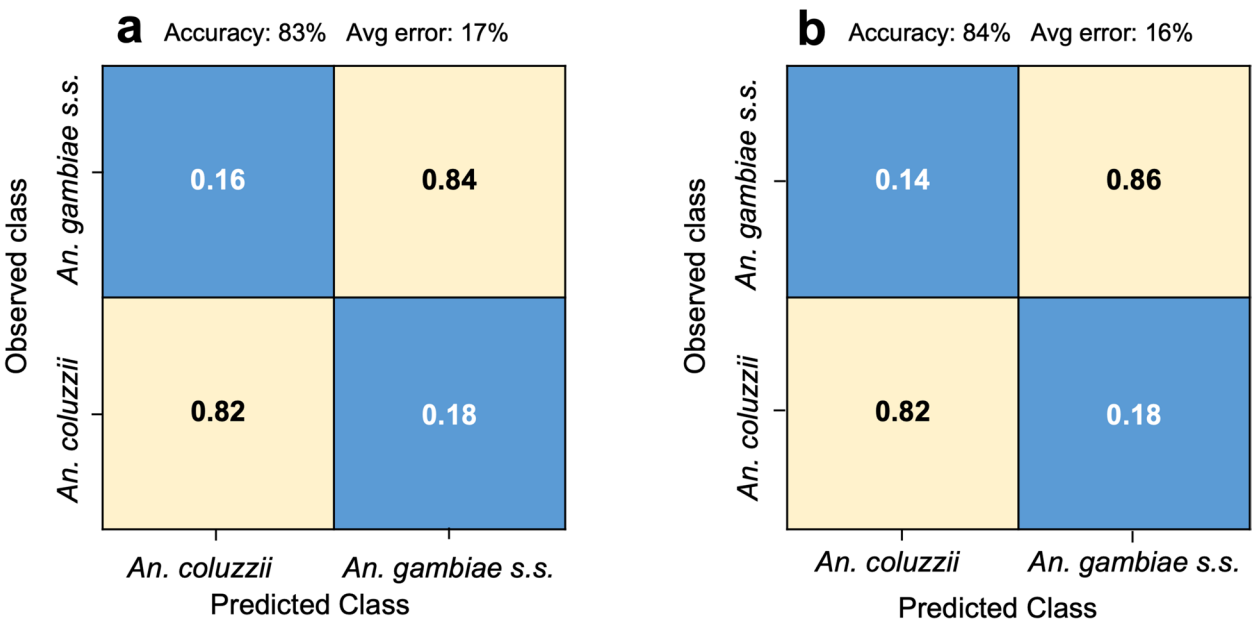


Fig. 4 NIRS accuracy to predict mosquito species reared under all conditions. **a** Laboratory derived *Anopheles*; model trained using mosquitoes from all environmental conditions and **b** Wild caught *Anopheles*; model trained with mosquitoes from all collection periods

mosquito species, though these models were trained on the smallest number of mosquitoes. The highest accuracy (84%) was found with the model trained on samples from the peak of the rainy season. Additional file 2: Table S2 provided detailed classification accuracy for *An. gambiae s.s.* and *An. coluzzii* under each collected period. As in laboratory *Anopheles*, the model trained from the peak of the rainy season specimens had the highest accuracy

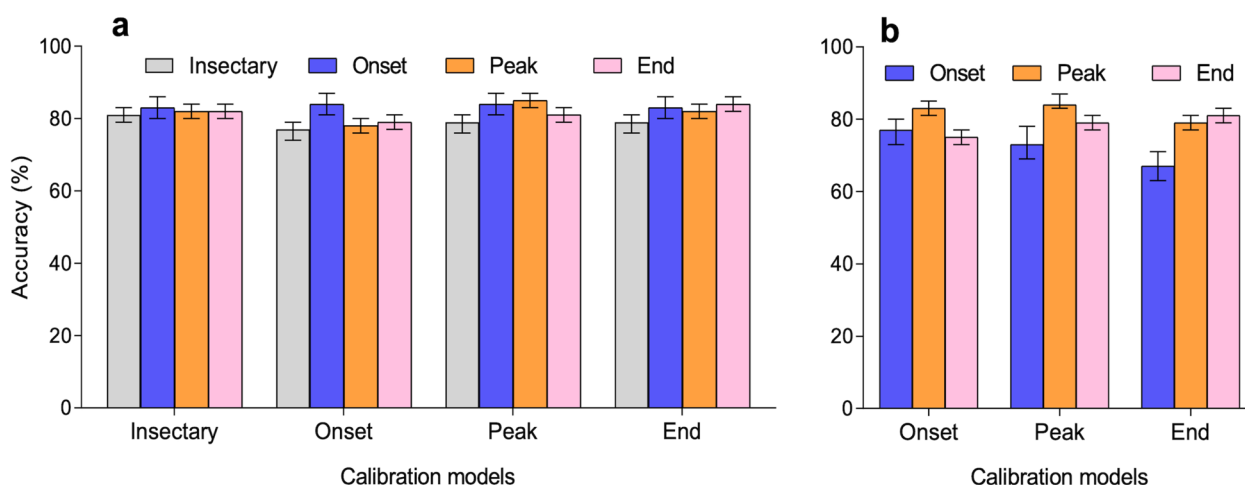


Fig. 5 NIRS accuracy to predict *An. gambiae* s.s. and *An. coluzzii* species using calibration models from each environmental condition. **a** Laboratory derived *Anopheles*; **b** Wild derived *Anopheles*. Predictive models were trained on sample derived from one environmental condition to predict those from another condition

across all populations (Fig. 5b). Finally, models derived from laboratory-mosquitoes reared under fluctuating environmental conditions predicted field-derived mosquito species with a low accuracy of 59% (*An. gambiae* s.s. 59%; *An. coluzzii* 58%).

NIRS accuracy to predict wild *An. gambiae* s.l. age under different environmental conditions

Around 2500 wild *Anopheles* were collected in the field, with 1000 females from Bama and 1500 from Soumouso. As expected, the three major species of *Anopheles* involved in malaria transmission (51.6% *An. gambiae* s.s., 38.2% *An. arabiensis* and 10.2% *An. coluzzii*) were found in Soumouso while Bama mosquitoes were all *An. coluzzii*. The first-generation female (F_1) of these mosquitoes was used for NIRS age analysis. Altogether, NIRS was unsuccessful determined mosquito chronological age in different environmental conditions (Fig. 6). The average difference between actual age and predicted age varied from 3.27 to 5.59 days. The lowest and highest average were found under end season conditions and insectary conditions, respectively.

Mosquito ability to transmit *Plasmodium* requires a minimum longevity covering the parasite extrinsic incubation period (EIP) estimated at 9 days [30]. Thus, the accuracy of the NIRS technique in classifying mosquito age as young (<9 days) or old (≥ 9 days) under different environmental conditions was evaluated. Regardless of mosquito group and environmental conditions, the binomial predictive model indicated that NIRS was able to classify mosquitoes as young (<9 days) and old (≥ 9 days), with 81% of accuracy for both ages' classifications. The similar accuracies were found under each environmental

condition with slight variations depending on the *Anopheles* group, ranging from 79 to 84% (Fig. 7). Additionally, although there is no significant difference, NIRS classified younger *Anopheles* (<9 days old) better than older ones (≥ 9 days) whatever environmental conditions and mosquito group (Table 1). These accuracies were obtained by training the model on samples of one condition and validating on a subsample of the same condition. However, NIRS was unable to distinguish between young and old mosquitoes when the model derived from mosquitoes reared in one environmental condition was used to predict the age of mosquitoes reared in another condition (<50% accuracy).

Discussion

NIRS can successfully discriminate between laboratory-reared and field-collected *An. gambiae* s.s. and *An. coluzzii* with varying degrees of accuracy depending on the environmental conditions. Overall, this work revealed that NIRS was able to distinguish both *An. gambiae* s.s. and *An. coluzzii* in the laboratory and in the field with greater than 80% of accuracy. Inter-environmental conditions accuracies of species identification vary slightly in laboratory mosquitoes but more in field specimens. Nevertheless, wild mosquitoes caught in one season can broadly predict species in mosquitoes caught months later, though with some loss of accuracy. The considerable variation in inter-period accuracies could be due to several potential factors: the diversity of larval habitats, nutriment (blood and sugar sources), gonotrophic/physiological status [12] and resistance status to the insecticides of the mosquitoes [29]. Since these factors weren't controlled in the field, the differences in

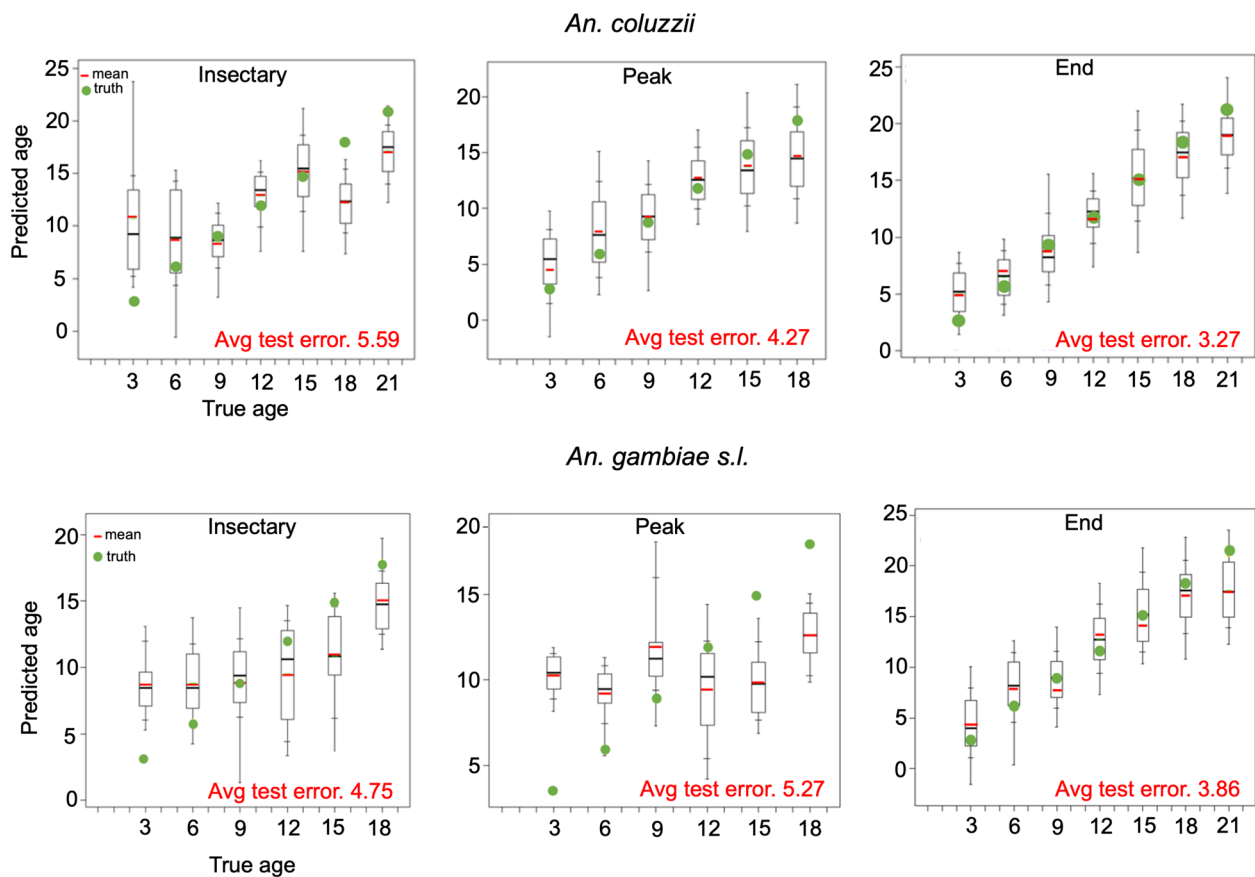


Fig. 6 NIRS accuracy to determine *Anopheles* chronological age under insectary, peak of the rainy season (August) and end of the rainy season/onset of the dry season (December) conditions programmed in incubators

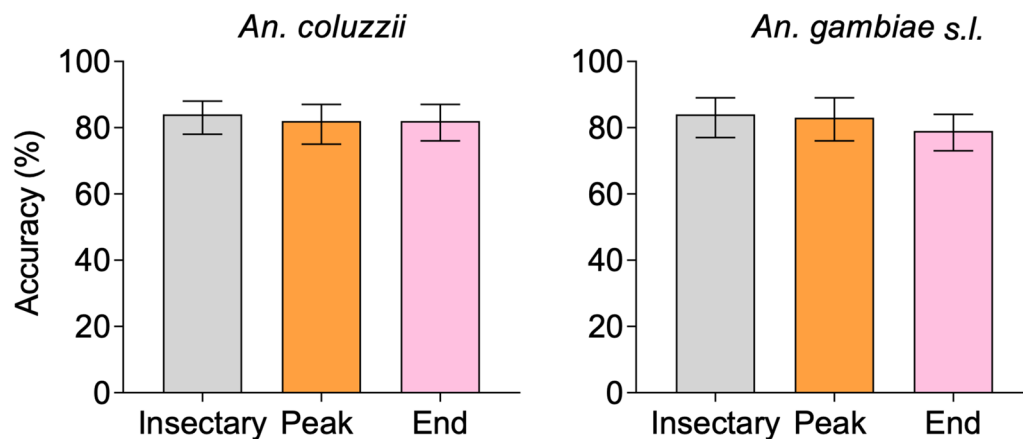


Fig. 7 NIRS accuracy to classify *Anopheles* (*An. coluzzii* and *An. gambiae s.l.*) in two age classes (< 9 days or ≥ 9 days) under insectary, peak of the rainy season (August) and end of the rainy season/onset of the dry season (December) conditions programmed in incubators

accuracy couldn't be solely attributed to the abiotic factors of various collection periods. Additionally, the spectra from laboratory mosquitoes differ from those of field mosquitoes in this study as reported by other authors

[7]. This study underlines once more the risk in extrapolating laboratory findings to the real-life environment, as previously suggested by other authors [31]. Results also revealed that the model derived from peak season

Table 1 NIRS accuracy to classify *Anopheles* age under different environmental conditions inside climate chambers

Within-sample accuracy (%)									
Periods	Number	<i>An. gambiae s.l.</i>				<i>An. coluzzii</i>			
		Accuracy	<9	≥9	p	Accuracy	<9	≥9	p
Insectary	342	84	87	80	0.35	84	84	83	1
Peak	288	83	86	79	0.32	82	82	81	0.83
End	333	79	81	77	0.83	81	83	79	0.79

samples had the highest accuracy for predicting samples from any other conditions, both in the laboratory and in the field. This period is characterized by a low fluctuation in temperature and RH, which could have less impact on absorbances through the mosquito cuticle compared to other periods. However, since each model was trained on a relatively small number of samples with varying sizes, these results should be validated using larger number of mosquitoes. This is particularly important given that previous studies have shown that increasing sample size can enhance the NIRS accuracy [10]. In addition, the preprocessing techniques and advance machine learning such as artificial neural networks and autoencoders could also improve the NIRS accuracy [32]. Furthermore, the study showed that the machine learning algorithm trained using laboratory-mosquitoes reared under different environmental conditions predicted field-derived mosquito species with a low accuracy. The mosquitoes were reared in laboratory under controlled conditions, free from factors like competition, nutrition, parasitism and several other factors that could create differences between laboratory specimens and those in the wild. This may mean that laboratory-reared mosquitoes do not reflect all the variations found in wild populations, so that predictive models derived from them are able to correctly predict the species of field specimens.

The results revealed that NIRS was able to classify both *An. coluzzii* and *An. gambiae s.l.* of age <9 days and ≥9 days with reasonable accuracies in each environmental condition. Although this technique is not useful to determine mosquito chronological age, it may provide a rapid means of determining an epidemiologically relevant measure of the proportion of the vector population that is enough old to transmit malaria. This is ecologically important because it may help in designing and monitoring vector prevention efforts and assess the risk of pathogen transmission in the human population. Overall, the analyses showed that the technique was slightly more accurate in predicting mosquitoes <9 days old compared to their counterparts ≥9 days old. This supports the assertion that major biochemical changes detected by NIRS occur

in mosquito cuticles from birth and stagnate between days 10 and 14 due to cuticle thickening [33]. Findings indicated that a calibration model derived from mosquitoes of specific environmental condition determined mosquito species of another condition with relatively good accuracy. However, classifying *Anopheles* age into two groups using a model trained on specimens from another environmental condition was challenging and resulted in low accuracy. This may have been driven in part by the low number of mosquitoes to train the models from some conditions or perhaps the diversity in spectra seen when temperature fluctuate. While statistical analyses require sufficient sample size to improve model accuracy [10], the number of mosquitoes using in this study substantially varied in the different seasons. Therefore, care should be taken interpreting the reliability of different models trained on data from different season. NIRS would be more useful if a calibration model derived from specimens in one environmental condition is able to predict age of *Anopheles* in another condition. This work suggests temperature and relative humidity could influence mosquito spectra and that future models should include some variability in these factors in the data used to train and validate the models. Similar results were reported by other authors who found that models derived from laboratory mosquitoes were unable to predict the age of field derived mosquitoes reared under ambient condition [7]. Similarly, models derived from one preserved mosquito were unsuitable for predicting the age of specimens stored in another preservative [34]. Based on these findings, it is suggested to estimate the age of mosquitoes from a given environmental condition using predictive models developed from mosquitoes reared under the same conditions. Ultimately, NIRS accuracy for *Anopheles* species identification and out-of-sample prediction should be improved by substantially increasing the number of mosquitoes used in the training and validating datasets. These mosquitoes should be reared in as diverse a set of environmental variables as possible to include the range likely in the wild mosquitoes being tested.

Conclusion

The findings revealed that NIRS is able to predict *An. gambiae s.l.* species under different environmental conditions with reasonable accuracy. In addition, models derived from *An. gambiae s.l.* in one environmental condition is able to predict *Anopheles* species in more other condition with relatively good accuracy. However, models trained using laboratory-mosquitoes reared under various environmental conditions predicted field-derived mosquito species with a low accuracy. NIRS provided a good classification of age under all environmental conditions for both *An. coluzzii* and *An. gambiae s.l.* when the age was classified into two groups age (<9 days or ≥9 days) using binary classification models. The models derived from *Anopheles* in one environmental condition failed to predict the age of mosquitoes in other conditions. Further studies are needed to investigate the accuracy of NIRS for determining *Anopheles* infection status or *Anopheles* resistance to insecticide status under different environmental conditions in order to develop appropriate calibration of the machine before field application.

Abbreviations

NIRS	Near infrared spectroscopy
ITNs	Insecticide-treated nets
IRS	Indoor residual spray
RH	Relative humidity
PCR	Polymerase chain reaction
PLS	Partial least squares regression
ROC	Receiver operating characteristic curve
AUC	Area under curve

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12936-025-05554-9>.

Additional file 1: Table S1. NIRS accuracy for predicting insectary *Anopheles* species reared under different environmental conditions inside climate chambers. The table has shown the NIRS accuracies of each *Anopheles* species in each environmental condition for insectary mosquitoes using different calibration models.

Additional file 2: Table S2. NIRS accuracy for predicting wild *Anopheles* species collected during different periods of rainy season. The table has described the NIRS accuracies of each *Anopheles* species collected during each period for wild mosquitoes using different calibration models.

Acknowledgements

The authors thank the residents of Bama, Klesso, Longo and Soumouso for their sincere cooperation during mosquito collection.

Author contributions

NDCC, DFD, and RKD conceived and designed the study. NDCC, DFD and BMS conducted the laboratory and field work. NDCC, LIGP and FC conducted the molecular analysis. NDCC, BMS and TSC were responsible for the data analysis. NDCC and DFD wrote the first draft of the manuscript. BMS, FC, JK, TSC and RKD reviewed the manuscript. All authors read and approved the final version of the manuscript.

Funding

The work was supported by the Centre d'Excellence Africain en Innovations Biotechnologiques pour l'Élimination des Maladies à transmission vectorielle (CEA-ITECH/MTV) of Nazi BONI University, Burkina Faso and UK Medical Research Council (MRC) Project Grant (MR/P01111X/1).

Availability of data and materials

The datasets are fully available without restriction from the corresponding authors by request. Further inquiries can be directed to the corresponding authors.

Declarations

Ethics approval and consent to participate

The project received approval from the local institutional ethics committee, reference number: A018-2017/CEIRES.

Consent for publication

"Not applicable" for that section.

Competing interests

The authors declare no competing interests.

Author details

¹Institut de Recherche en Sciences de La Santé, Direction Régionale, 399 Avenue de La Liberté, 01 BP 545, Bobo-Dioulasso 01, Burkina Faso. ²Université Nazi Boni, Bobo-Dioulasso, Burkina Faso. ³MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London W2 1PG, UK.

Received: 22 October 2024 Accepted: 28 August 2025

Published online: 30 September 2025

References

- Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature*. 2015;526:207–11.
- Carnevale P, Robert V. Les anophèles. Biologie, transmission du *Plasmodium* et lutte antivectorielle. IRD Edition. 2009.0.4000/books.irdeditions.10374.
- Sikulu M, Killeen GF, Hugo LE, Ryan PA, Dowell KM, Wirtz RA, et al. Near-infrared spectroscopy as a complementary age grading and species identification tool for African malaria vectors. *Parasit Vectors*. 2010;3:49.
- Johnson BJ, Hugo LE, Churcher TS, Ong OTW, Devine GJ. Mosquito age grading and vector-control programmes. *Trends Parasitol*. 2020;36:39–51.
- González Jiménez M, Babayan SA, Khazaeli P, Doyle M, Walton F, Reedy E, et al. Prediction of mosquito species and population age structure using mid-infrared spectroscopy and supervised machine learning. *Wellcome Open Res*. 2019;4:76.
- Mohammad N, Naudion P, Dia AK, Boëlle PY, Konaté A, Konaté L, et al. Predicting the age of field *Anopheles* mosquitoes using mass spectrometry and deep learning. *Sci Adv*. 2024;10:ead6990.
- Ong OTW, Kho EA, Esperança PM, Freebairn C, Dowell FE, Devine GJ, et al. Ability of near-infrared spectroscopy and chemometrics to predict the age of mosquitoes reared under different conditions. *Parasit Vectors*. 2020;13:160.
- Mayagaya VS, Michel K, Benedict MQ, Killeen GF, Wirtz RA, Ferguson HM, et al. Non-destructive determination of age and species of *Anopheles gambiae s.l.* using near-infrared spectroscopy. *Am J Trop Med Hyg*. 2009;81:622–30.
- Sikulu MT, Majambere S, Khatib BO, Ali AS, Hugo LE, Dowell FE. Using a near-infrared spectrometer to estimate the age of *Anopheles* mosquitoes exposed to pyrethroids. *PLoS ONE*. 2014;9:e90657.
- Lambert B, Sikulu-Lord MT, Mayagaya VS, Devine G, Dowell F, Churcher TS. Monitoring the age of mosquito populations using near-infrared spectroscopy. *Sci Rep*. 2018;8:5274.

11. Maia MF, Kapulu M, Muthui M, Wagah MG, Dowell FE, Baldini F, et al. Detection of malaria in insectary-reared *Anopheles gambiae* using near-infrared spectroscopy. *Epidemiology*. 2019;18:85.
12. Da DF, McCabe R, Somé BM, Esperança PM, Sala KA, Blight J, et al. Detection of *Plasmodium falciparum* in laboratory-reared and naturally infected wild mosquitoes using near-infrared spectroscopy. *Sci Rep*. 2021;11:10289.
13. Somé BM, Da DF, McCabe R, Djègbè NDC, Paré LIG, Wermé K, et al. Adapting field-mosquito collection techniques in a perspective of near-infrared spectroscopy implementation. *Parasit Vectors*. 2022;15:338.
14. Ntamatungiro AJ, Mayagaya VS, Rieben S, Moore SJ, Dowell FE, Maia MF. The influence of physiological status on age prediction of *Anopheles arabiensis* using near infra-red spectroscopy. *Parasit Vectors*. 2013;6:298.
15. Liebman K, Swamidoss I, Vizcaino L, Lenhart A, Dowell F, Wirtz R. The influence of diet on the use of near-infrared spectroscopy to determine the age of female *Aedes aegypti* mosquitoes. *Am J Trop Med Hyg*. 2015;92:1070–5.
16. Claudio-Piedras F, Recio-Tótoro B, Cime-Castillo J, Condé R, Maffei M, Lanz-Mendoza H. Dietary and *Plasmodium* challenge effects on the cuticular hydrocarbon profile of *Anopheles albimanus*. *Sci Rep*. 2021;11:11258.
17. Aw WC, Ballard JWO. The effects of temperature and diet on age grading and population age structure determination in *Drosophila*. *J Insect Physiol*. 2013;59:994–1000.
18. Djègbè NDC, Da DF, Somé BM, Paré LIG, Cissé F, Mamai W, et al. *Anopheles* aquatic development kinetic and adults' longevity through different seasons in laboratory and semi-field conditions in Burkina Faso. *Parasit Vectors*. 2024;17:181.
19. Mouline K, Mamai W, Agnew P, Tchonfienet M, Brengues C, Dabire R, et al. Physiology and development of the M and S molecular forms of *Anopheles gambiae* in Burkina Faso (West Africa). *Med Vet Entomol*. 2012;26:447–54.
20. Wagoner KM, Lehmann T, Huestis DL, Ehrmann BM, Cech NB, Wasserberg G. Identification of morphological and chemical markers of dry- and wet-season conditions in female *Anopheles gambiae* mosquitoes. *Parasit Vectors*. 2014;7:294.
21. Huestis DL, Lehmann T. Ecophysiology of *Anopheles gambiae* s.l.: persistence in the Sahel. *Infect Genet Evol*. 2014;28:648–61.
22. Hidalgo K, Mouline K, Mamai W, Foucreau N, Dabiré KR, Bouchereau A, et al. Novel insights into the metabolic and biochemical underpinnings assisting dry-season survival in female malaria mosquitoes of the *Anopheles gambiae* complex. *J Insect Physiol*. 2014;70:102–16.
23. Dabiré RK, Namountougou M, Diabaté A, Soma DD, Bado J, Toé HK, et al. Distribution and frequency of kdr mutations within *Anopheles gambiae* s.l. populations and first report of the Ace. 1G119S mutation in *Anopheles arabiensis* from Burkina Faso (West Africa). *PLoS ONE*. 2014;9:e101484.
24. Coluzzi M, Petrarca V, WHO Aspirator with paper cup for collecting mosquitos and other insects. Geneva, World Health Organization. 1973. <https://iris.who.int/handle/10665/65660>.
25. Gillies MT, Coetzee M. A supplement to the *anophelinae* of Africa South of the Sahara. *S Afr Inst Med Res*. 1987;55(Suppl):1–143.
26. Santolamazza F, Mancini E, Simard F, Qi Y, Tu Z, della Torre A. Insertion polymorphisms of SINE200 retrotransposons within speciation islands of *Anopheles gambiae* molecular forms. *Malar J*. 2008;7:163.
27. Gimonneau G, Pombi M, Dabiré RK, Diabaté A, Morand S, Simard F. Behavioural responses of *Anopheles gambiae* sensu stricto M and S molecular form larvae to an aquatic predator in Burkina Faso. *Parasit Vectors*. 2012;5:65.
28. Nignan C, Poda BS, Sawadogo SP, Maïga H, Dabiré KR, Gnankine O, et al. Local adaptation and colonization are potential factors affecting sexual competitiveness and mating choice in *Anopheles coluzzii* populations. *Sci Rep*. 2022;12:636.
29. Esperança PM, Da DF, Lambert B, Dabiré RK, Churcher TS. Functional data analysis techniques to improve the generalizability of near-infrared spectral data for monitoring mosquito populations. *bioRxiv*. 2020. (preprint).
30. Beier JC. Malaria parasite development in mosquitoes. *Annu Rev Entomol*. 1998;43:519–43.
31. Arcaz AC, Huestis DL, Dao A, Yaro AS, Diallo M, Andersen J, et al. Desiccation tolerance in *Anopheles coluzzii*: the effects of spiracle size and cuticular hydrocarbons. *J Exp Biol*. 2016;219:1675–88.
32. Milali Id MP, Sikulu-Lord MT, Kiware SS, Dowell Id FE, Corliss GF, Povinelliid RJ. Age grading *An. gambiae* and *An. arabiensis* using near infrared spectra and artificial neural networks. *PLoS ONE*. 2019;14:e0209451.
33. Caputo B, Dani FR, Horne GL, Petrarca V, Turillazzi S, Coluzzi M, et al. Identification and composition of cuticular hydrocarbons of the major Afrotropical malaria vector *Anopheles gambiae* s.s. (Diptera: Culicidae): Analysis of sexual dimorphism and age-related changes. *J Mass Spectrom*. 2005;40:1595–604.
34. Dowell FE, Noutcha AEM, Michel K. The effect of preservation methods on predicting mosquito age by near infrared spectroscopy. *Am J Trop Med Hyg*. 2011;85:1093–6.

Publisher's Note

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