Applied Spectroscopy

www.s-a-s.org

Detection and Identification of *Wolbachia pipientis* Strains in Mosquito Eggs Using Attenuated Total Reflection Fourier Transform Infrared (ATR FT-IR) Spectroscopy

Dale Christensen[1,2], Aazam Khoshmanesh[1],
David Perez-Guaita[1,3], Inaki Iturbe-Ormaetxe[4], Scott O'Neill[5] (iD),
and Bayden R. Wood[1] (iD)

Abstract

The global fight against mosquito-borne viral diseases has in recent years been bolstered by the introduction of the endosymbiotic bacteria *Wolbachia* to vector populations, which in host mosquitoes suppresses the transmissibility of several viruses. Researchers engaged on this front of the battle often need to know the *Wolbachia* infection status of individual mosquitoes, as the intervention progresses and the mosquitoes become established in the target population. Previously, we successfully applied attenuated total reflection Fourier transform infrared spectroscopy to the detection of *Wolbachia* in adult *Aedes aegypti* mosquitoes; here we apply the same principles to *Aedes* eggs, with sensitivity and selectivity > 0.95. Further, we successfully distinguish between infections in eggs of the *w*Mel and *w*MelPop strains of *Wolbachia pipientis*, with a classification error of 3%. The disruption of host lipid profile by *Wolbachia* is found to be a key driver in spectral differences between these sample classes.

Introduction

An estimated 3.9 billion people across 128 countries are at risk of contracting dengue fever, a potentially fatal viral disease transmitted primarily by the *Aedes aegypti* mosquito.[1] One study estimates a global annual infection rate of 390 million, of which approximately 25% manifest with clinical symptoms.[2] According to the World Health Organization, reported infections have increased precipitously in recent decades, and the number of countries where dengue is endemic now exceeds 100.[3]

Wolbachia pipientis are Gram-negative intracellular endosymbiotic bacteria first described by Hertig and Wolbach in 1924 in *Culex pipiens* mosquitoes.[4] Although they are not naturally present in *A. aegypti* populations, when transinfected into *A. aegypti*[5,6] they have been shown to suppress the transmission of the dengue virus,[7–10] along with other viral diseases such as Zika and chikungunya.[11,12] Their use

as a biocontrol agent is magnified by the phenomenon of cytoplasmic incompatibility, which renders the offspring of *Wolbachia*-infected males and uninfected females unviable by disrupting their embryonic development.[13] By this

[1]Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, Australia
[2]ANSTO Australian Synchrotron, Clayton, Australia
[3]Department of Analytical Chemistry, University of Valencia, Burjassot, Spain
[4]Institute of Vector-Borne Diseases, Monash University, Clayton, Australia
[5]World Mosquito Program, Ho Chi Minh City, Vietnam

Corresponding authors:
Aazam Khoshmanesh, School of Chemistry, Monash University, Wellington Rd., Clayton, AU-VIC Victoria 3800, Australia.
Email: aazam.khoshmanesh@monash.edu
Bayden R. Wood, School of Chemistry, Monash University, Wellington Rd., Clayton, AU-VIC Victoria 3800, Australia.
Email: bayden.wood@monash.edu

mechanism a relatively small proportion of infected individuals can increase to dominate a local population of mosquitoes within a few generations.

In the course of both laboratory research and field roll-out, scientists in the World Mosquito Program (WMP, worldmosquitoprogram.org, formerly the Eliminate Dengue Program) frequently need to know whether Wolbachia is present in individual mosquitoes in order to determine the establishment of these mosquitoes in the local A. aegypti population. This is presently determined using quantitative polymerase chain reaction (qPCR) assays, which apply DNA primers for targeting Wolbachia-specific genes. These tests are accurate, robust, and sensitive, and are rightly regarded as a "gold standard" for the detection of Wolbachia in mosquitoes. However, they have a large per-sample cost due to expensive polymerase chain reaction (PCR) instruments and reagents; results are also not timely for efficient field implementation, as the bulk and complexity of the instrument mean captured mosquitoes must be shipped from field collection sites to a central lab environment. These factors have driven the search for a faster, cheaper method for detecting Wolbachia in mosquitoes.

One promising candidate is attenuated total reflection Fourier transform infrared (ATR FT-IR) spectroscopy. ATR FT-IR instruments can be significantly cheaper and more portable than qPCR instruments, and no expensive reagents are required in the trivial sample preparation stage. The infrared technique has also been applied to a range of insects (including mosquitoes) in the past with some success.[14]

Jiménez et al.[15] used ATR FT-IR spectra in conjunction with machine learning to predict mosquito age, and to classify each as belonging to either Anopheles gambiae or An. arabiensis. While the accuracy of species classification was only 83%, the age distributions predicted via the spectroscopic model were "statistically indistinguishable from true modelled distributions".[15] Sroute et al.[16] used FT-IR to classify mosquito samples as A. aegypti, A. albopictus, A. japonicus, or A. triseriatus, with an overall accuracy of 94%. Mwanga et al.[17] used ATR FT-IR spectra of bloodfed female mosquitoes to determine the vertebrate from which the blood meal had come (goat, cow, chicken, or human), with an accuracy of 98%.

Attenuated total reflections FT-IR has also been successfully deployed to detect Wolbachia in adult mosquitoes. Khoshmanesh et al.[18] successfully applied ATR FT-IR to distinguish male mosquitoes from females, young (two day) from old (10 day), and Wolbachia-infected from uninfected. However, the sensitivity and specificity of these models decreased markedly when the sample population changed from lab-reared to wild mosquitoes. This is most likely due to greater variability in diet, age, and condition in wild populations compared to lab-reared colonies. Since mosquito eggs are an inherently more homogeneous matrix, we

analyzed them using similar methods, hypothesizing that a greater spectral distinction between infected and uninfected eggs may be evident. Researchers at the WMP have established the possibility of detecting Wolbachia in single eggs through reverse-transcription PCR,[*] but for economic and logistical reasons this is not currently deployed as part of their field strategy.

Other studies have also applied near-infrared spectroscopy (NIRS) to detect W. pipientis in mosquitoes and report the ability to distinguish between bacterial strains.[19] While these approaches have had some success, the near-infrared region comprises overlapping combination and overtone bands. Because of this, NIRS tends to give less insight into the biomolecular factors driving the spectral separation between classes.[20] By attempting to replicate this spectral distinction using a mid-infrared ATR FT-IR instrument, we hope to add this insight on top of a robust classification model.

Materials and Methods

Sample Origin and Preparation

All egg samples were descendants of mosquitoes captured from WMP field sites in Queensland, Australia. These mosquitoes were reared as part of the established colony at Monash University, as described previously by McMeniman et al.,[21] with some minor differences. Briefly, adult mosquitoes were maintained in cages inside a controlled room temperature at 26 °C with 65–75% relative humidity and 12 h:12 h light–dark cycle. They were allowed access to 10% sucrose ad libitum as well as to human blood. Eggs were collected on filter paper strips placed around water containers. The colony contains numerous sub-populations from which we selected four for analysis: WT (wildtype mosquitoes, uninfected with Wolbachia), wMel (mosquitoes transinfected with the wMel strain of Wolbachia), wMelPop (mosquitoes transinfected with the wMelPop strain of Wolbachia), and wMel.tet (wMel mosquitoes treated with the antibiotic tetracycline to eliminate Wolbachia infection from population). These classes are summarized in Table I. Both wMel and wMelPop strains have been evaluated for use in WMP's long-term field deployment strategy; while both suppress viral transmissibility within the host,

Table I. Summary of sample classes.

Class	n	Infected	Uninfected
WT	72		✓
wMel	52	✓	
wMel.tet	52		✓
wMelPop	57	✓	
Total	233	109	124

*w*MelPop also introduces a range of detrimental fitness effects, which impede its establishment and longevity within wild populations.[22] As multiple eggs (approximately 8–12) were required to cover the attenuated total reflection (ATR) crystal and achieve sufficient signal, the number of "samples" from each class corresponds to the number of spectra, not individual eggs; each spectrum measured eggs from one class only.

Data Collection

Once laid on damp filter paper, the eggs were lightly pressed and stored between paper towel and absorbent kitchen cloth to remove excess moisture and preclude hatching before analysis.

A Bruker model EQUINOX 55 (Bruker Optic, Germany) FT-IR spectrometer fitted with an N_2-cooled mercury–cadmium–tellurium (MCT) detector and a Golden Gate diamond ATR accessory (Specac Ltd, UK) was used for spectral acquisition of the mosquito eggs. Eggs were deposited on to the ATR crystal until it was fully covered; the anvil was then applied to ensure adequate contact between the eggs and the crystal. Spectra were recorded with a spectral resolution of $6\,cm^{-1}$ from 32 co-added interferograms, using a Happ–Genzel apodization function in the Fourier transform. The order in which samples were analyzed was randomized so as to rule out systematic errors introduced by changes in the lab environment. The Bruker system was controlled using an IBM-compatible PC running OPUS v.6.0 software.

Data Analysis

The collected data were analyzed using PLS_Toolbox (Eigenvector Research, USA), which operates in a Matlab environment (The MathWorks Inc., USA).

Several combinations of spectral pre-processing techniques were applied to the data, and a number of restricted spectral windows were considered for further analysis. Both preliminary observations and model performance confirmed the efficacy of a simple approach: the application of a Savitzky–Golay second derivative[23] (third-order polynomial, 11-point smoothing window), standard normal deviate, and mean center. These processes were applied to a spectral window of $1800–900\,cm^{-1}$.

A k-nearest neighbor (KNN) classification, principal component analysis (PCA), and partial least squares discriminant analysis (PLS-DA) were performed on the resultant data set. PCA is a widely used chemometric method,[24] implemented here as an unsupervised method of examining spectral variation within the whole population, with the spectral significance of the PC loadings explored in some depth. KNN, an unsophisticated algorithm, assigns a class to each sample in the independent test set by the "vote" of the k closest calibration set samples. It is shown here first to demonstrate the relative ease with which the different sample classes can be identified. However, KNN provides no explanation as to what spectral features are driving this classification. Finally, two PLS-DA models were constructed, harnessing the explanatory power of a supervised method to further explore the spectral differences between classes. When helpful, these PLS-DA models are class-orthogonalized, a technique which helps to isolate the variables most responsible for discriminating between sample classes.[25]

For the supervised methods (KNN, PLS-DA), approximately 70% of the samples from each class ($n = 161$) were randomly assigned to a model calibration data set. This data set was employed for the creation of models, using leave-one-out cross-validation to optimize the number of latent variables (LV) in the PLS-DA. The generalization error of the models was evaluated through independent validation using the remaining 30% ($n = 72$) of samples. The performance of each of the classification models was then compared using a series of confusion matrices.

Results and Discussion

Spectral Quality and Band Assignment

An early decision was made to focus our analysis on the region between 1800 and $900\,cm^{-1}$, the information-rich lipid region ($3100–2800\,cm^{-1}$) being largely overwhelmed by the O–H stretch from residual water in the samples. Although the water band and $1640\,cm^{-1}$ also causes some interference, the underlying biological bands still proved to be useful in our analyses. Figure 1 shows the average second-derivative spectra in this range, broken down by sample class. Table II presents our assignment of the key peaks present in these spectra.

The second-derivative spectra show clearly the average difference between classes in the ester carbonyl band ($1744\,cm^{-1}$), with the two infected classes having stronger absorbance at this energy. More subtle differences are seen in the amide I and amide II profiles, indicating the possible usefulness of protein content to distinguish between sample classes. The methyl deformation mode at $1378\,cm^{-1}$ and the carboxylate stretch at $1400\,cm^{-1}$ are present in all sample classes, but their strength in oppositely correlated. This may suggest an inverse relationship between lipids and proteins or a disruption of the host lipid profile by the *Wolbachia* endosymbiont, depending on whether the carboxylate signal is mostly attributable to fatty acids or to amino acid residue in proteins. The final difference visible to the eye from these spectra arises in the DNA bands at 1080 and $964\,cm^{-1}$, where the tetracycline-treated mosquitoes seem to have weaker absorbance on average.
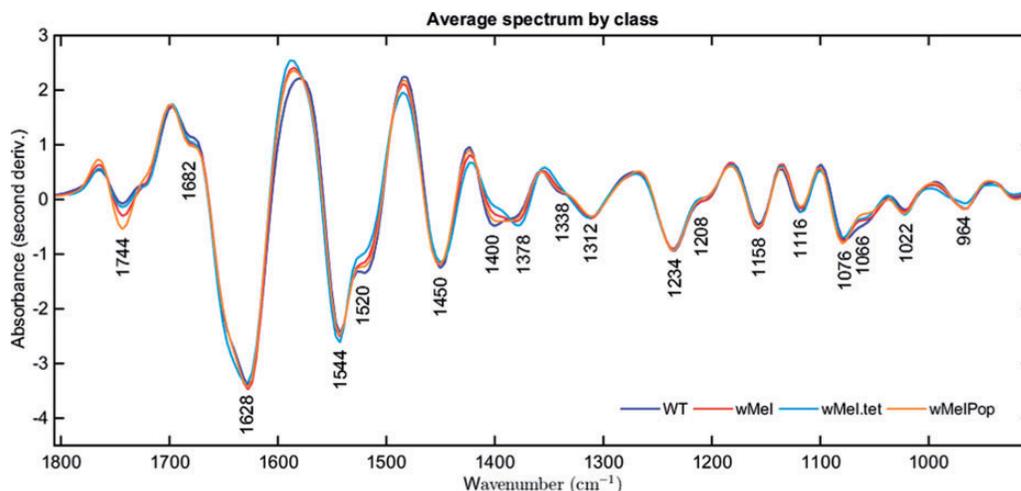
Figure 1. Average second-derivative spectra broken down by sample class, with key peaks labeled. Note in particular the differences between class at the ester carbonyl (1744 cm^{-1}), methyl deformation (1400 cm^{-1}), carboxylate stretch (1378 cm^{-1}), and DNA bands (1076, 964 cm^{-1}).

Table II. Assignment of vibrational peaks in ATR FT-IR spectra.

Wavenumber (cm^{-1})	Vibrational mode[a]	Assignment
964	$\nu_s PO_4^{2-}$	DNA (phosphomonoester)
1038, 1066	$\nu C-O$	Glycogen
1080	$\nu_s PO_2^-$	DNA (phosphodiester)
1158, 1116	$\nu C-O$	RNA (ribose)
1234	$\nu_{as} PO_2^-$	DNA (phosphodiester)
1312	$\delta NH_2 \uparrow\uparrow \nu C-N$	Proteins (amide III)
1378	δCH_3	Lipids
1400	$\nu_s COO^-$	Lipids (fatty acids), Proteins (amino acid residues)
1450	δCH_2	Lipids (acyl chains), Proteins (backbone)
1544	$\delta NH_2 \uparrow\downarrow \nu C-N$	Proteins (amide II)
1628	$\nu C=O \uparrow\downarrow (\nu C-N, \delta NH_2)$	Proteins (amide I)
1744	$\nu CO_2 R$	Lipids (triglyceride backbone)

$\uparrow\uparrow$ In-phase combination.
$\uparrow\downarrow$ Out-of-phase combination.
[a]Complex modes given with some minor contributors listed in parentheses.[26]

K-Nearest Neighbor Classification

The calibration set of 161 samples was used to train the model. Figure 2 shows the results of this model applied to the 72-sample independent test set with $k = 3$. The predictive power of this model is excellent, with an overall classification error of just 4.2%. However, as noted above, it provides no insight to the biochemical phenomena that produce such a distinction between sample classes; for that, more sophisticated analyses are required.

Principal Component Analysis

The calibration and test samples from above were combined into a single data set for PCA. The scores and loadings for the first three principal components (PCs) (accounting for a cumulative 85% of total sample variance) are shown in Fig. 3.

The spectral variation described by PC-1 incorporates 46% of the total sample variance, and indicates a positive correlation between lipids (1744 cm^{-1}, ester carbonyl) and
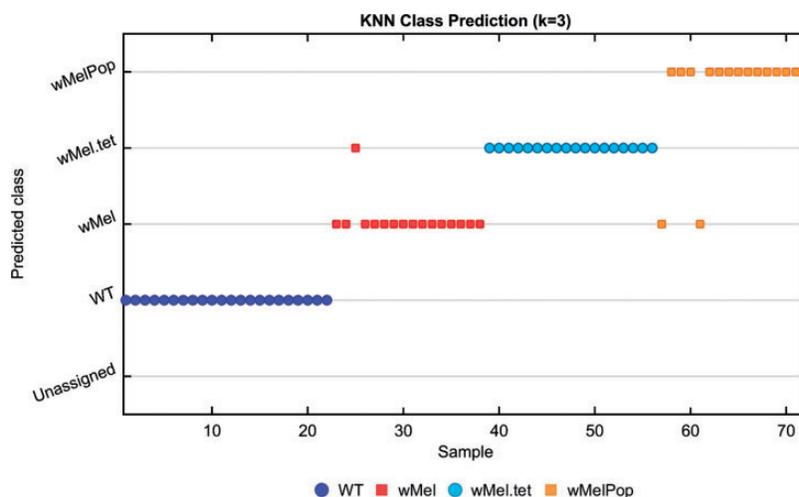
Figure 2. KNN prediction of class membership for the independent test set ($n = 72$). Prediction accuracy was greatest with $k = 3$. No samples were left unassigned by the model, and only three were misclassified.

nucleic acids (1220, 1084 cm^{-1}, DNA phosphodiester), and an inverse correlation between these and protein (1612 cm^{-1}, amide I). However, PC-1 seems not to be meaningfully correlated with *Wolbachia* infection status (Figs. 3a and 3c). It is possible that the considerable residual moisture in the samples is also being picked up in this PC, indicating varying degrees of dehydration uncorrelated with *Wolbachia* infection.

On the other hand, PC-2 and PC-3 describe promising distinction between sample classes (Figs. 3b, 3d, and 3e). We observed in Fig. 3b that the two non-infected sample classes (WT and *w*Mel.tet) had formed promising clusters away from the other classes. A third cluster had also formed, comprised of both infected sample classes (*w*Mel and *w*MelPop). The bulk of the separation between classes occurred along PC-2, with significant amide I and II bands in the loadings suggesting that the protein composition of the samples played the most decisive role. We might expect simple groupings of infected and non-infected samples along this axis, but interestingly, the cluster of infected samples sat at the midpoint, with WT and *w*Mel.tet occupying respectively the low and high extremes. It is not immediately obvious why the two uninfected sample classes are so different in this respect, but the initial presence of *Wolbachia* followed by treatment with a broad spectrum antibiotic clearly has a compounding effect.

The dominance of the ester carbonyl peak at 1744 cm^{-1} in the PC-3 loadings suggests a population-wide variation in lipid content, and the now-familiar inverse correlation with the amide I band at 1650 cm^{-1} is the only other significant datum in this PC. However, the diagnostic significance of variation along this axis is only marginal, and only in conjunction with the information in PC-2; there is little class

separation along this PC, despite its accounting for nearly 12% of variation in the sample population.

Principal component analysis was also applied to (a) lipid and protein (3100–2800, 1700–1500 cm^{-1}), and (b) lipid and nucleic acid (3100–2800, 1300–900 cm^{-1}) spectral windows. These did not provide superior performance to the analysis based on the entire fingerprint region and so are not shown here.

Partial Least Squares Discriminant Analysis

While our PCA showed modest separation between infected and non-infected sample classes, there was no ability to distinguish between *Wolbachia* strains, nor were we able to establish exactly what distinguished the two uninfected classes from the infected. For these questions we turned to a supervised method, PLS-DA.

The first PLS-DA model was designed to predict only the presence or absence of *Wolbachia*. The infected sample classes *w*Mel and *w*MelPop were assigned a pseudo-variable $Y = 0$, while WT and *w*Mel.tet were assigned $Y = 1$.

Figures 4a and 4b demonstrate the accuracy of this classification model, with a classification error of just 4.2% and an area under the receiver operator characteristic curve (AUROC) of 0.9983. Figure 4c is not the regression vector, but the component of LV1 when the PLS-DA model is class-orthogonalized (OPLS-DA). This orthogonalization allows us to clearly see the spectral variations associated with the differences between infected and uninfected samples in the first LV.

Our second PLS-DA model was designed to distinguish between *w*Mel-infected and *w*MelPop-infected mosquito eggs. For this analysis, *w*MelPop was assigned a
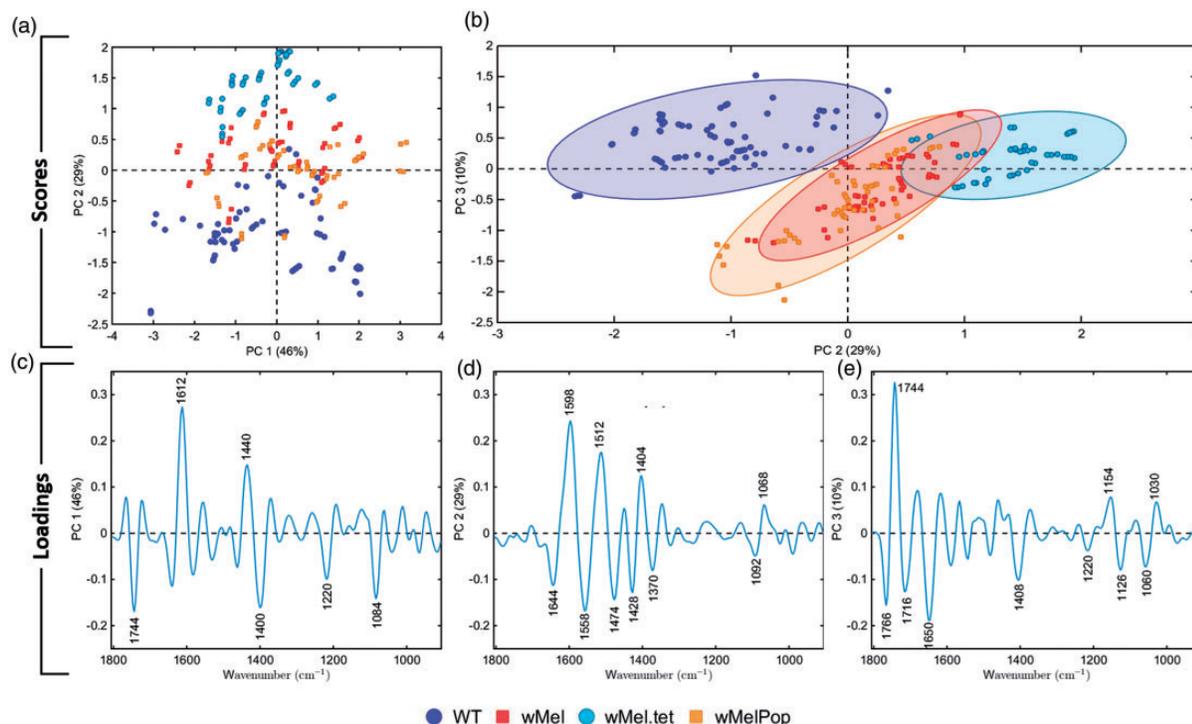
Figure 3. PCA applied to entire data set. (a) Scores plot, showing little class separation along PC1, but a definite sorting along PC2. (b) PC2 versus PC3 scores plot, which shows the best separation between classes in the model. Ellipses are 95% confidence limits for each class. (c–e) Loadings for PC1-3. Of particular note is the inverse relationship between the ester carbonyl and protein bands shown in PCs 1 and 3. This relationship is a recurring theme in our other analyses.

pseudo-variable $Y = 1$ and wMel was assigned $Y = 0$, while WT and wMel.tet samples were excluded. After the model was constructed, it was tested on the subset of the aforementioned test set which belongs to these classes ($N = 33$), with the results shown in Figs. 4d and 4e. As in the previous model, the third figure panel describes not the regression vector, but LV1 for a class-orthogonalized model. Once again, the first LV of the OPLS-DA shows that the class discrimination is driven largely by variance in the lipid profile. The wMelPop-infected eggs show an exaggerated version of the same effect that distinguished infected from uninfected eggs, most obvious in the enlarged ester carbonyl peak ($1744 \, \mathrm{cm}^{-1}$).

The results of both PCA and PLS-DA models indicate that the lipid and protein profiles of mosquito eggs are significantly altered by the presence of *Wolbachia*. This is consistent both with our previous ATR FT-IR work in adult mosquitoes,[18] and with current microbiological and entomological literature.[27] In particular, the ester carbonyl's strong correlation with uninfected eggs, and the proclivity of *Wolbachia*-infected samples to have a relatively high protein content, are consistent with our findings in Khoshmanesh et al.[18] These characteristics are exaggerated in wMelPop-infected eggs relative to wMel, providing a spectral basis for distinguishing the two classes with excellent accuracy (97%).

Comparing Classification Model Performances

Previously, Sikulu-Lord et al.[19] and Khoshmanesh et al.[18] have reported diagnostic models for detection of *Wolbachia* based on infrared spectroscopy. Using NIRS, Sikulu-Lord et al.[19] produced binary classifiers for WT/ wMel, WT/wMelPop, and wMel/wMelPop, with each of these broken into two separate models for male and female mosquitoes. The classification error for these models applied to independent test sets ranged from 2 to 21%. Khoshmanesh et al. used mid-infrared spectra of female only mosquitoes were able to distinguish between infected and uninfected mosquitoes with errors of just 3–5%. These studies form a benchmark for evaluating the success of classification models from mosquito eggs.

Figure 5 shows a series of confusion matrices describing our classification models.

The KNN model is not directly comparable to these previous studies but did produce an excellent error rate of just 3%. Our PLS-DA model for detecting the presence of *Wolbachia* had an overall prediction accuracy of 96%, equal or superior to the classification models of Sikulu-Lord et al.[19] and Khoshmanesh et al.[18] Sikulu-Lord's classifiers for distinguishing wMel and wMelPop had an overall accuracy of 88%, which was surpassed by our PLS-DA model at 97%.
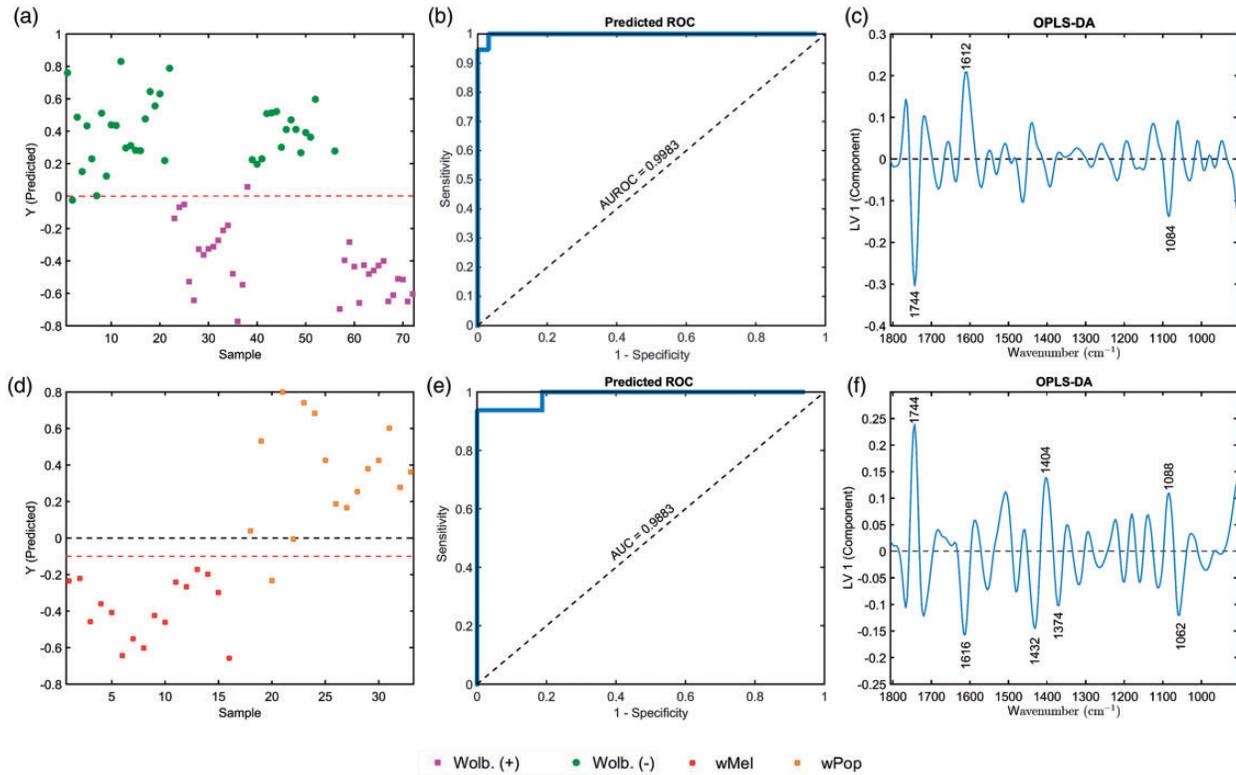
Figure 4. Two PLS-DA models, the first (a–c) for distinguishing *Wolbachia* infection status, and the second (d–f) for distinguishing between the *w*Mel and *w*MelPop strains of *Wolbachia*. (a, d) The classification results for each of the tested samples. (b, e) The model's receiver operator characteristic curve (ROC). An area under this curve (AUROC) of close to 1.0 demonstrates the great success of the classification models. Please note that for the purposes of (e), *w*MelPop was considered the "positive" class, with sensitivity and specificity to be interpreted accordingly. (c, f) LV1 for each of these models when class-orthogonalized (OPLS-DA), a method for drawing out the variables most involved in separating the sample classes. Both models demonstrate the importance of the inverse relationship between the ester carbonyl peak (1744 cm^{-1}) and the amide I peak (1612 cm^{-1}).
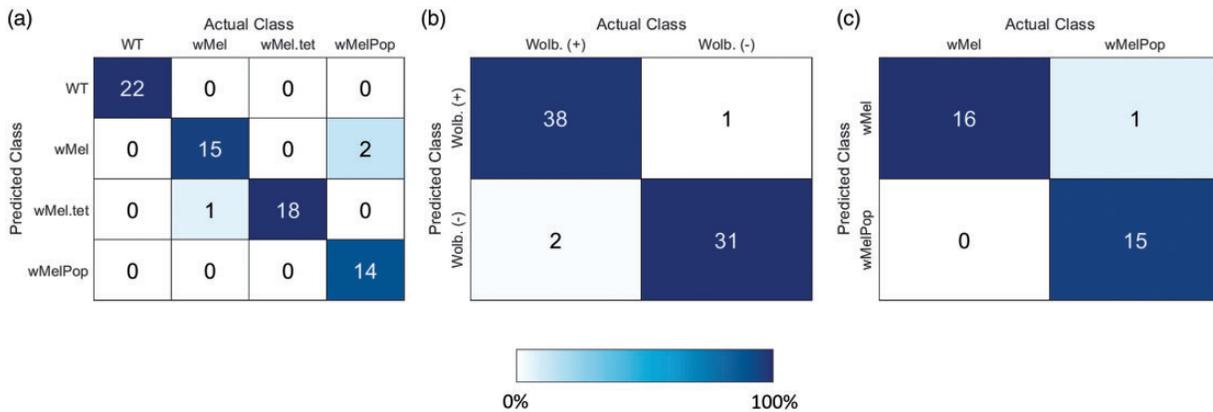

Figure 5. Confusion matrices for our three classification models: (a) KNN four-class model, (b) PLS-DA *Wolbachia* infection status model, and (c) PLS-DA *w*Mel versus *w*MelPop model. The color bar represents the percentage of samples in each actual class that are predicted in each class. The strong dark blue down the diagonal of each matrix indicates the excellent accuracy of these models applied to the test sets.

Implications for Further Research

The success of PLS-DA at predicting not only infection status but also strain of *Wolbachia* confirms our hypothesis that mosquito eggs are an excellent target for biospectroscopic analysis. However, it is appropriate to take time to situate these results in the context of the broader WMP effort, in order that their significance not be under- or overstated.

The models above are based on eggs collected in the laboratory, not in the wild, and as such they may suffer from the same fragility seen by Khoshmanesh et al. when analyzing adult mosquitoes.[18] However, the samples' relative homogeneity and time of collection, very early in the life cycle, leads us to hypothesize that eggs would be less susceptible to these interfering effects. Now that we have established the viability of using eggs for the detection of *Wolbachia*, further research must be undertaken to apply these techniques to samples collected in the field. Given the dominance of water signals on PC-1, any field studies should consider a dehydration step in sample preparation, so as not to introduce confounding variation into the data set.

As mentioned above, each of our sample spectra was derived from several eggs, in order to establish sufficient spectral signal. There is no practical need to test individual eggs, as the detection of *Wolbachia* in any egg is sufficient to establish its presence in the mother. However, any given cluster of mosquito eggs may be the progeny of several individual females, which may in turn differ in *Wolbachia* infection status. These challenges likely rule out surveying populations with the same level of granularity as current testing (performed with adult mosquitoes), but are unlikely to preclude monitoring the proportion of *Wolbachia* establishment on a wider scale.

In this study, we have been able to assume that every sample is an *A. aegypti* egg. In order to be a viable screening method for samples collected in the field, we must also be able to sift out eggs of other mosquito species accidentally collected. Herein lies the inherent drawback of using eggs; adults are much easier to ID species. However, the gains in sample homogeneity discussed above are not insignificant, and so a model that includes an initial screening step for species is worth pursuing. The success of previous studies in identifying adult mosquito species by ATR FT-IR suggests optimism is warranted in meeting this challenge.[15,16]

Conclusion

The use of ATR FT-IR spectroscopy on mosquito eggs produced clean, replicable spectra with biologically significant absorbance peaks clearly resolved in the second derivative. Analysis of egg spectra by PCA and PLS-DA revealed the significance of the ester carbonyl, amide I, and phosphodiester bands in detecting the presence of *Wolbachia*. Uninfected samples are strongly correlated with the ester carbonyl band at $1744\,\text{cm}^{-1}$, and inversely correlated with the amide I peak at $1630\,\text{cm}^{-1}$, a finding consistent with previous ATR FT-IR studies on adult mosquitoes.[18]

A PLS-DA model was able to predict the *Wolbachia* infection status of an independent test set with a sensitivity of 0.95 and specificity of 0.97. A second PLS-DA model was able to further distinguish between *w*Mel-infected and *w*MelPop-infected eggs, with a classification error of just 3%. KNN modeling predicted the class of the same independent test set as WT, *w*Mel, *w*Mel.tet, or *w*MelPop, with an overall accuracy of 96%.

The scale and logistical framework of the WMP currently precludes using eggs to measure the spread of *Wolbachia* through a mosquito population, in part because of the relative ease of identifying the species of adult mosquitoes. The established detection method (PCR) becomes infeasible both economically and temporally for eggs, because the presence of non-target species in the collection area leads to an increase in total samples processed along with a species pre-screening test. However, we have demonstrated the efficacy of the ATR FT-IR/PLS-DA approach as a viable option for inclusion in future fieldwork program designs, by the WMP or other organizations. In addition, the molecular features we have identified as key to discriminating between *w*Mel and *w*MelPop will aid the development of strain-specific diagnostic tools.

ORCID iDs

Scott O'Neill ⓘ https://orcid.org/0000-0002-4131-3615
Bayden R. Wood ⓘ https://orcid.org/0000-0003-3581-447X

Note

* Unpublished data, Institute for Vector-Borne Diseases, communicated April 2021.

References

1. O.J. Brady, P.W. Gething, S. Bhatt, J.P. Messina, et al. "Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus". PLoS Negl. Trop. Dis. 2012. 6(8): e1760.
2. S. Bhatt, P.W. Gething, O.J. Brady, J.P. Messina, et al. "The Global Distribution and Burden of Dengue". Nature. 2013. 496(7446): 504–507.
3. World Health Organization (WHO). "Dengue and Severe Dengue". 2020. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue [accessed Apr 15 2021].
4. M. Hertig, S.B. Wolbach. "Studies on Rickettsia-Like Micro-Organisms in Insects". J. Med. Res. 1924. 44(3): 329.

5. D.A. Joubert, T. Walker, L.B. Carrington, J.T. De Bruyne, et al. "Establishment of a *Wolbachia* Superinfection in *Aedes aegypti* Mosquitoes as a Potential Approach for Future Resistance Management". PLoS Pathog. 2016. 12(2): e1005434.

6. J.K. Axford, P.A. Ross, H.L. Yeap, A.G. Callahan, et al. "Fitness of wAlbB *Wolbachia* infection in *Aedes aegypti*: Parameter Estimates in an Outcrossed Background and Potential for Population Invasion". 2016. 94(3): 507–516.

7. A.A. Hoffmann, B. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, et al. "Successful Establishment of *Wolbachia* in *Aedes* Populations to Suppress Dengue Transmission". Nature. 2011. 476(7361): 454–457.

8. T. Walker, P. Johnson, L. Moreira, I. Iturbe-Ormaetxe, et al. "The wMel *Wolbachia* Strain Blocks Dengue and Invades Caged *Aedes aegypti* Populations". Nature. 2011. 476(7361): 450–453.

9. P. Ryan, A. Turley, G. Wilson, T. Hurst, et al. "Establishment of wMel *Wolbachia* in *Aedes aegypti* Mosquitoes and Reduction of Local Dengue Transmission in Cairns and Surrounding Locations in Northern Queensland, Australia". Gates Open Res. 2019. 3: 1547.

10. C. Indriani, W. Tantowijoyo, E. Rancès, B. Andari, et al. "Reduced Dengue Incidence Following Deployments of *Wolbachia*-Infected *Aedes aegypti* in Yogyakarta, Indonesia: A Quasi-Experimental Trial Using Controlled Interrupted Time Series Analysis". Gates Open Res. 2020. 4: 50.

11. L.A. Moreira, I. Iturbe-Ormaetxe, J.A. Jeffery, G. Lu, et al. "A *Wolbachia* Symbiont in *Aedes aegypti* Limits Infection with Dengue, Chikungunya, and Plasmodium". Cell. 2009. 139(7): 1268–1278.

12. H.L.C. Dutra, M.N. Rocha, F.B.S. Dias, S.B. Mansur, et al. "*Wolbachia* Blocks Currently Circulating Zika Virus Isolates in Brazilian *Aedes aegypti* Mosquitoes". Cell Host Microbe. 2016. 19(6): 771–774.

13. S. Sinkins, H. Braig, S.L. O'Neill. "*Wolbachia* Superinfections and the Expression of Cytoplasmic Incompatibility". Proc. R. Soc. London, Ser. B. 1995. 261(1362): 325–330.

14. J.B. Johnson, M. Naiker. "Mid-Infrared Spectroscopy for Entomological Purposes: A Review". J Asia-Pac. Ent. 2020. 23(3): 613–621.

15. M.G. Jiménez, S.A. Babayan, P. Khazaeli, M. Doyle, et al. "Prediction of Mosquito Species and Population Age Structure Using Mid-Infrared Spectroscopy and Supervised Machine Learning". Wellcome Open Res. 2019. 4: 76.

16. L. Sroute, B.D. Byrd, S.W. Huffman. "Classification of Mosquitoes with Infrared Spectroscopy and Partial Least Squares-Discriminant Analysis". Appl. Spec. 2020. 74(8): 900–912.

17. E.P. Mwanga, S.A. Mapua, D.J. Siria, H.S. Ngowo, et al. "Using Mid-Infrared Spectroscopy and Supervised Machine-Learning to Identify Vertebrate Blood Meals in the Malaria Vector, *Anopheles arabiensis*". Malar. J. 2019. 18(1): 187.

18. A. Khoshmanesh, D. Christensen, D. Perez-Guaita, I. Iturbe-Ormaetxe, et al. "Screening of *Wolbachia* Endosymbiont Infection in *Aedes aegypti* Mosquitoes Using Attenuated Total Reflection Mid-Infrared Spectroscopy". Anal. Chem. 2017. 89(10): 5285–5293.

19. M.T. Sikulu-Lord, M.F. Maia, M.P. Milali, M. Henry, et al. "Rapid and Non-Destructive Detection and Identification of Two Strains of *Wolbachia* in *Aedes aegypti* by Near-Infrared Spectroscopy". PLoS Negl. Trop. Res. 2016. 10(6): e0004759.

20. J.A. Adegoke, K. Kochan, P. Heraud, B.R. Wood. "A Near-Infrared "Matchbox Size" Spectrometer to Detect and Quantify Malaria Parasitemia". Anal Chem. 2021. 93(13): 5451–5458.

21. C.J. McMeniman, R.V. Lane, B.N. Cass, A.W. Fong, et al. "Stable Introduction of a Life-Shortening *Wolbachia* Infection into the Mosquito *Aedes aegypti*". Science. 2009. 323(5910): 141–144.

22. T. Nguyen, H. Le Nguyen, Y. Thu, N. Sinh, et al. "Field Evaluation of the Establishment Potential of wMelPop *Wolbachia* in Australia and Vietnam for Dengue Control". Parasit Vectors. 2015. 8(8): 563.

23. A. Savitzky, M.J. Golay. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". Anal. Chem. 1964. 36(8): 1627–1639.

24. C.B. Cordella. "PCA: The Basic Building Block of Chemometrics". In: I.S. Krull, editor. Analytical Chemistry. London: IntechOpen, 2012. Pp. 1–35.

25. J. Trygg, S. Wold. "Orthogonal Projections to Latent Structures (O-PLS)". J. Chemomet. 2002. 16(3): 119–128.

26. A. Barth. "Infrared Spectroscopy of Proteins". Biochim. Biophys. Acta, Bioenerg. 2007. 1767(9): 1073–1101.

27. J.C. Molloy, U. Sommer, M.R. Viant, S.P. Sinkins. "*Wolbachia* Modulates Lipid Metabolism in *Aedes albopictus* Mosquito Cells". Appl. Env. Microbiol. 2016. 82(10): 3109.