Title Page

Title:

Rapid evaporative ionisation mass spectrometry (REIMS): A potential and rapid tool for the identification of insecticide resistance in mosquito larvae.

Authors:

Jasmine Morgan[1], J. Enrique Salcedo-Sora[2*], Iris Wagner[3], Robert J Beynon[3], Omar Triana-Chavez[4] and Clare Strode[1*]

Author affiliations:

[1]Department of Biology, Edge Hill University, Ormskirk, Lancashire, L39 4QP, UK

[2]GeneMill, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Crown Street, Liverpool, L69 7TU, UK

[3]Centre for Proteome Research, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK.

[4] Instituto de Biología, Facultad de Ciencias Exactas y Naturales (FCEN), University of Antioquia, Medellín, Colombia

Corresponding authors: *

Clare Strode, BIO208, Biosciences building, Department of Biology, Edge Hill University, Ormskirk, Lancashire, L39 4QP, 01695 657 373, strodecl@edgehill.ac.uk

J. Enrique Salcedo Sora, +44 (0)151 794 6455, J.Salcedo-Sora@liverpool.ac.uk

Abstract

Insecticide resistance is a significant challenge facing the successful control of mosquito vectors globally. Bioassays are currently the only method for phenotyping resistance. They require large numbers of mosquitoes for testing, the availability of a susceptible comparator strain and often insectary facilities. This study aimed to trial the novel use of rapid evaporative ionisation mass spectrometry (REIMS) for the identification of insecticide resistance in mosquitoes. No sample preparation is required for REIMS and analysis can be rapidly conducted within hours. Temephos resistant *Aedes aegypti* (Linnaeus) larvae from Cúcuta, Colombia and temephos susceptible larvae from two origins (Bello, Colombia, and the lab reference strain New Orleans) were analysed using REIMS. We tested the ability of REIMS to differentiate three relevant variants: population source, lab versus field origin and response to insecticide. The classification of these data was undertaken using linear discriminant analysis (LDA) and random forest. Classification models built using REIMS data were able to differentiate between *Ae. aegypti* larvae from different populations with 82% (± 0.01) accuracy, between mosquitoes of field and lab origin with 89% (± 0.01) accuracy and between susceptible and resistant larvae with 85% (± 0.01) accuracy. LDA classifiers had higher efficiency than random forest with this data set. The high accuracy observed here identifies REIMS as a potential new tool for rapid identification of resistance in mosquitoes. We argue that REIMS and similar modern phenotyping alternatives should complement existing insecticide resistance management tools.

Key words:

Insecticide resistance, rapid evaporative ionisation mass spectrometry, REIMS, *Aedes aegypti*, larvae, Colombia

Introduction

Insecticide resistance is one of the most significant challenges posed to mosquito control programmes. The control of mosquito vectors, including *Aedes aegypti* (Linnaeus) the principal vector for the dengue, Zika and chikungunya viruses, relies heavily on the use of insecticides to reduce disease burden. There are only four insecticides classes which are licensed for use in public health: organophosphates, organochlorines, pyrethroids and carbamates. Resistance has now been reported in *Ae. aegypti* to all four of these chemical classes (Ranson et al. 2010, Vontas et al. 2012, Moyes et al. 2017). Insecticide resistance in *Ae. aegypti* is also spread worldwide with reports in South America (Guedes et al. 2020), North America (Marcombe et al. 2014), Asia (Amelia-Yap et al. 2018), Europe (Seixas et al. 2017), Africa (Weetman et al. 2018), and Oceania (Demok et al. 2019). This trend is compromising effective vector control (Viana-Medeiros et al. 2007, Bisset et al. 2011, Marcombe et al. 2011).

Insecticide resistance management (IRM) which aims to prevent, slow, or reverse the emergence of resistance is therefore crucial for sustainable vector control. The first step in IRM is to monitor local populations for the development of insecticide resistance whilst establishing its impact on effective vector control (Dusfour et al. 2019). Current methods for resistance monitoring include bioassays, biochemical assays, and molecular testing. Biochemical assays and molecular testing are used to identify the specific mechanisms responsible for insecticide resistance, allowing for appropriate IRM strategies to be implemented (Hemingway et al. 2013). However, insecticide bioassays (e.g. WHO tube and CDC bottle assays) are the only current method for identifying (phenotyping) resistance in mosquitoes. Bioassays have low sensitivity, lengthy completion times (24 hours) and often only detect high levels of resistance which maybe too late for alternative measures to be deployed (Dusfour et al. 2019). Other limitations include the requirement of large numbers of

70 individual mosquitoes, and the availability of a comparable susceptible strain (World Health
71 Organization (WHO) 2016). Alternative phenotyping methods that can surpass those
72 limitations are necessary.

73 Rapid evaporative ionisation mass spectrometry (REIMS) is a relatively new technology
74 which provides a rapid method of mass spectrometry without the need for any sample
75 preparation. Samples are burned by diathermy and the resultant aerosols are collected,
76 ionized, and analysed by mass spectrometry (Schäfer et al. 2009, Balog et al. 2010, 2013,
77 2015). The spectra, collected in negative ion mode, largely reflect the lipid composition of
78 the sample, and is collected over a wide range of m/z values. The spectra are then are
79 discretised by binning, creating a data matrix that is further processed by dimension reduction
80 and classification (Balog et al. 2010). The potential applications of REIMS are vast with its
81 previous successful applications including distinguishing cancerous tissue from healthy tissue
82 (Alexander et al. 2017, St John et al. 2017, Phelps et al. 2018), authentication of food
83 products (Balog et al. 2016, Black et al. 2017, Verplanken et al. 2017, Guitton et al. 2018,
84 Rigano et al. 2019), microbial species identification (Strittmatter et al. 2013, 2014),
85 monitoring of bacterial growth and recombinant protein expression (Sarsby et al. 2021), and
86 the identification of rodent species and sex from faecal matter (Davidson et al. 2019). REIMS
87 has also been shown to be a highly effective method for species and sex determination in
88 *Drosophila* adults and larvae (Wagner et al. 2020).

89 Here we present a proof-of-concept for the novel use of REIMS as a rapid tool for the
90 identification of insecticide resistance in *Ae. aegypti* larvae. We analysed three *Ae. aegypti*
91 populations, previously profiled for susceptibility to the larvicide temephos (Morgan et al.
92 2021): a resistant population originating from field collected mosquitoes from Cúcuta
93 (Colombia) and two susceptible populations, one field originating population from Bello
94 (Colombia) and a susceptible laboratory reference strain, New Orleans. The results

demonstrate the potential of REIMS for phenotyping insecticide resistant mosquitoes with relevant discriminatory power and faster and less labour-intensive methods which may be used to complement existing IRM strategies.

Materials and methods

Mosquito samples and rearing

Aedes aegypti larvae from three populations previously tested for susceptibility to temephos (Morgan et al. 2021) were used in this study. Two field populations were used, one temephos resistant (field resistant (FR) and one susceptible (field susceptible (FS)), the susceptible *Ae. aegypti* laboratory strain New Orleans (lab susceptible (LS)) was also used (Fig.1). *Ae. aegypti* were reared to fourth instar larvae following a standard rearing protocol and under standard conditions within Edge Hill University Vector Research Group insectaries. Standard conditions were 27^{o}C and 70% relative humidity with an 11-hour day/night cycle with 60-minute dawn/dusk simulation periods, using a lighting system of 4× Osram Dulux 26W 840 lights. Eggs were submerged in a hatching broth of 350 ml dH_2O, 0.125 g nutrient broth (Sigma-Aldrich, Dorset, UK) and 0.025 g brewer's yeast (Holland & Barrett, Ormskirk, UK) for 48 hours (Zheng et al. 2015). Once hatched, larvae were reared at a density of 0.5 larva/ml in dH_2O and fed ground fish food (AQUARIAN® advanced nutrition) at increasing quantities per day (day 3 = 0.08 mg/larva, day 4 = 0.16 mg/larva, day 5 = 0.31 mg/larva, day 6 = 0 mg/larva) (Carvalho et al. 2014). For each experimental group (FR, FS, LS) four biological replicates were conducted, using eggs from different females each submerged on different days. Seven days after egg submission larvae were removed and stored at -20^{o}C until REIMS analysis. The storage period ranged from 32-36 weeks (Table 1). The number of larvae analysed per biological replicate ranged from 8-15 with a total of 42-51 larvae per experimental group (Table 1).

Rapid evaporative ionisation mass spectrometry analysis

Rapid evaporative ionisation mass spectrometry analysis was conducted following the detailed methods outlined by Wagner *et al.* (2020). Larvae were burned using a monopolar electrosurgical pencil (Erbe Medical UK Ltd, Leeds); the electric current was provided to the pencil by a VIO 50 C electrosurgical generator, a black conductive rubber mat acted as the counter electrode to enable the flow of electricity through the sample. The entire biomass of each larva was burned, and the aerosols produced were aspirated through tubing attached to the pencil into the REIMS source using a nitrogen powered venturi valve. Leucine enkephalin (Waters, UK) in propan-2-ol (CHROMASOLV, Honeywell Riedel-de-Haën) was used as a lock mass solution and continuously introduced via a whistle in the venturi tube at a flow rate of 30 µl min−1. REIMS was conducted using a Synapt G2Si instrument ion mobility equipped quadrupole time of flight mass spectrometer (Waters, UK). A heated impactor (Kanthal metal coil at 900°C) within the REIMS source was used to decluster the ionized particles. Mass spectra were acquired in negative ion mode at a rate of 1 scan per second over a mass/charge range of *m/z* 50–1200. All larvae were analysed in a single day in a random order created by a random number generator within Microsoft excel.

Data analysis

The raw data files were imported into the Offline Model Builder software (OMB-1.1.28; Waters Research Centre, Hungary). Each data file/sample contains the burn event of only one larva, therefore the option to create one spectrum per sample was selected. The background was subtracted, and the spectra corrected using the lock mass (leucine enkephalin, m/z 554.26). The normalised intensities were then binned into 0.1 m/z wide groups. The binned mass spectra data were then imported into R (version 3.6.3) (R Core Team 2020) for further analysis.

Dimension reduction was carried out by principal components analysis (PCA) using the R package factoextra (version 1.0.7) (Kassambara and Mundt 2020). Different numbers of principal components were then extracted (10,20,40,60,80,100) and used for classification of samples into categories: population, population type and resistance status. Classification was conducted using two different model types; linear discriminant analysis (LDA) and random forest (RF), with the data randomly split into 70% training data and 30% test data. Each model was built using variable numbers of principal components (PCs) extracted using PCA and the most accurate model selected and used for analysis. LDA models with varying numbers of PCs were built using the R package MASS (version 7.3.53) (W. N. Venables and B. D. Ripley 2002), model validation was conducted by plotting receiver operating characteristic curves (ROC) and selecting the model with the highest area under ROC curve (AUC) (Supp Fig.S2-11). Random forest models were validated using the R package caret (version 6.0.88) (Kuhn 2021) to select the model with optimum PCs, number of variables available for splitting at each tree node (mtry) and tree number. The random forest models with the highest overall accuracy following building in caret were selected for use in the analysis with models built using the R package randomForest (version 4.6.14) (Liaw and Wiener 2002). Random under sampling in the caret package was used to balance classes prior to RF analysis as this showed increase in overall model performance. Class imbalance did not affect performance of LDA models, as no difference in classification accuracy was observed between the different groups within the models, therefore no over or under sampling was required. LDA and RF models with parameters as selected by model validation were each ran 20 times using a different random split of test (30%) and training (70%) data. The model statistics: percentage accuracy, standard error of means (SEM) and range, were then averaged across all 20 replicates. LDA following PCA was also used to visualise the separation of

samples, plots were created using the R packages ggplot2 (version 3.3.2) (Wickham 2016) and ggpubr (version 0.4.0) (Kassambara 2020).

The experimental design is outlined in Fig.1. A code for analysing REIMS data using LDA and random forest classification models which can be applied to other similar datasets is available in Supplementary File 1. All raw data files are available in the MetaboLights database under the accession number MTBLS4129. The data matrix, created in OMB and used for subsequent analysis in R is available in Supplementary table 1.

Results

Population source

Visualisation of the data, following PCA-LDA analysis showed a clear discrimination between *Ae. aegypti* larvae from different geographical origins (Fig.2A). All three populations; field susceptible, field resistant and lab susceptible were separated in linear discriminant one whilst the field resistant population separated from the two susceptible populations in linear discriminant 2, thus demonstrating that LD1 is representative of population and LD2 of resistance to insecticide. A PCA-LDA conducted on the data with randomly assigned classifications showed no separation (Supp Fig.S1) demonstrating that the observed separation of classifications is due to variations between populations and not due to chance. The LDA model built using the REIMS data was able to correctly classify 82% (± 0.01) of *Ae. aegypti* larvae into the correct population (Fig.2B). The lab susceptible population had the highest accuracy (90% ± 2.0) and had the largest sample number whilst the population with the lowest sample number, field resistant, had the lowest accuracy (77% ± 2.2). When classification was conducted using a random forest model accuracy was lower, but the model was still able to correctly assign 76% of individual *Ae. aegypti* larvae to the correct population (Fig.2C).

Population type (lab and field)

A clear separation is observed when *Ae. aegypti* larvae from field origin are compared to larvae from a standard laboratory reference strain using PCA-LDA (Fig.3A & B). The classification models had high accuracy with 89% (± 0.01) of individual larvae classified to the correct population type with the PCA-LDA model (Fig.3C) and 83% (± 0.01) correctly classified by random forest (Fig.3D). Larvae from field origin had higher classification accuracy (86% ± 1.8) than those of lab origin (80% ± 2.4) when the RF model was used. When the LDA model was used the accuracy was similar for both groups (Field = 90% ± 0.8, Lab = 89 ± 2.0).

Insecticide sensitivity profile

Analysis of the REIMS data was also conducted to investigate the potential for determination between insecticide resistant and susceptible *Ae. aegypti* larvae (Fig.4). PCA-LDA classification models show 85% (± 0.01) accuracy in assigning larvae to the correct resistance status, with 75% (± 2.8) of temephos resistant larvae being correctly assigned (Fig.4c). The classification accuracy was higher for susceptible individuals (89% ± 1.1), this is likely due to the larger sample size of susceptible individuals available for training the model (Fig.4C). Whilst the random forest classification model was less accurate it still had a correct classification rate of 78% (± 0.02) correctly classifying 73% (± 3.3) of resistant individuals and 79% of susceptible individuals (Fig.4D).

A similar classification accuracy is achieved when field resistant larvae are compared only to susceptible larvae from a laboratory strain (Fig.5) as when field resistance larvae are compared to susceptible larvae from field origin (Fig.6). When only a field susceptible comparator strain is used the classification accuracy was 88% (± 0.01) using LDA (Fig.6C) and 84% (± 0.02) using RF (Fig.6D). When only a lab susceptible comparator strain is used the classification accuracy was similar with accuracies of 87% with LDA (Fig.5C) and 82%

with RF (Fig.5D). The similarity in classification accuracy observed here demonstrates that a field equivalent susceptible strain may not be necessary for identification of insecticide resistance in field *Ae. aegypti* larvae using this method, which is beneficial with the decreasing availability of field relevant susceptible populations.

Discussion

Early detection of resistance in mosquito populations is key to effective IRM and in reducing its effect on transmission of disease (Dusfour et al. 2019). The current principal methods for monitoring resistance are bioassays, biochemical assays, and molecular testing. Biochemical assays and molecular testing can be used to identify resistance in mosquitoes and are also important for the identification of mechanisms conferring resistance which can be useful when deciding on the most effective control method and in the development of novel control strategies (Brogdon 1989, World Health Organization (WHO) 1998, Corbel and N'Guessan 2013, Hemingway et al. 2013, Faucon et al. 2017, Dusfour et al. 2019). Current understanding of resistance has been developed through molecular and biochemical studies which have identified common resistance mechanisms including target site insensitivity and metabolic detoxification (Hemingway et al. 2004). Identification of these resistance mechanisms has been vital to increasing understanding of resistance.

Biochemical and molecular assays are important for increasing understanding of resistance mechanisms however there is an operational need for scalable rapid identification tools which are less labour intensive thereby yielding faster results which therefore have the potential to have more direct impact on decision making in the field. Insecticide bioassays are currently the only method for phenotyping resistance in mosquitoes (World Health Organization 2013, World Health Organization (WHO) 2016). They are limited to detecting high levels of resistance only which is often too late for alternative control methods to be deployed and high level of variation between experiments is often observed (Owusu et al. 2017). Bioassays also

require large numbers of mosquitoes, the availability of a comparable susceptible strain and insectary facilities (World Health Organization 2013, World Health Organization (WHO) 2016).

This study presents proof of concept for the use of rapid evaporative ionisation mass spectrometry (REIMS) as a faster tool for monitoring of insecticide resistance which has the potential to directly inform vector control decision making. The data obtained by REIMS analysis was able to categorise resistance with 85% (± 0.01) accuracy. This method also benefits from requiring no sample preparation, and rapid data acquisition. For this study relatively small sample numbers were used, but high accuracy was still obtained. Accuracy of classification models has potential to increase as the size of the training data set is increased, therefore subsequent testing with higher sample numbers may yield an even greater accuracy, however higher variability of samples (diet, ages, environmental factors etc.) would need to be included in order to produce a robust model capable of dealing with fully wild samples (Dobbin et al. 2008, Figueroa et al. 2012, Hanberry et al. 2012, Beleites et al. 2013, Luan et al. 2020). The tool was also able to differentiate between different mosquito populations with 82% (± 0.01) accuracy, suggesting other applications for the tool aside from resistance monitoring.

We also compared two different classification model types, linear discriminant analysis (LDA) and random forest (RF) both of which are commonly applied to classification of samples using REIMS data (Cameron et al. 2016, St John et al. 2017, Davidson et al. 2019, Gredell et al. 2019, Wagner et al. 2020, Sarsby et al. 2021). LDA is often the classification method of choice for spectrometry-based phenotyping, including REIMS (Bonetti 2018, D'Hue et al. 2018, Gredell et al. 2019, Kenar et al. 2019, Liu et al. 2021, Wang et al. 2021). The results of this study showed that LDA classification models were able to achieve comparable accuracy to the more complex random forest models and in the case of our data

266 performed better. Use of a simpler but equally accurate model is important in enabling the
267 data analysis to be accessible to a variety of personnel working within vector control. The
268 PCA-LDA method has previously been shown to be effective at classifying groups which
269 show large differences in biochemical profile, however for groups with more subtle
270 differences machine learning methods may have higher accuracy than LDA (Gromski et al.
271 2015, Gredell et al. 2019). The higher accuracy of the LDA model used in this study
272 comparatively to the RF model suggests that the differences in molecular profile between the
273 groups studied; geographical origin, population type and resistance status may be distinct.
274 This provides further promise for the use of REIMS in insecticide resistance monitoring as
275 larger differences in lipid signatures are easier to detect than subtle differences. The use of
276 multiple classification models to accurately classify REIMS data has previously been shown
277 to be important due to the high complexity of REIMS data. Dimension reduction, as
278 conducted in this study, has also been shown to be a critical step in REIMS data analysis
279 (Gredell et al. 2019).

280 Whilst the REIMS method is a fast and effective method it does have some disadvantages
281 when compared with alternative methods. The technique is destructive, meaning that the
282 sample cannot be used for further analysis. However, application of the technique to adult
283 mosquitoes provides the opportunity for partial dissection (e.g. leg removal) prior to REIMS
284 which will allow for further genetic or biochemical testing. The mass spectroscopy
285 equipment involved in REIMS is estimated to cost around $500,000 USD (Logrono 2020),
286 whilst costs of the initial set up of REIMS facilities are high, once equipment is available the
287 cost per sample is low due to rapid sampling turnover. Costs are also saved elsewhere without
288 the need for high staffing costs and insectary facilities. The speed at which samples can be
289 analysed allows for high sample turnover which therefore reduces cost, 100 mosquito larvae
290 could be analysed, and an answer generated in as little as 2-3 hours. In other applications

291 including cancer diagnostic REIMS has been identified to be a more cost-effective method
292 than other molecular techniques with costs around £1.60 per sample (Paraskevaidi et al.
293 2020). The REIMS method identifies differences in the lipid/metabolite profile of samples
294 however specific molecule detection is not the objective of this method, which is designed
295 instead to detect unique patterns in mass spectrum that enable classification (Wagner et al.
296 2020). Whilst we propose the use of REIMS as a potential rapid resistance identification tool
297 with direct operational impact the technique is not intended to be used for identification of
298 the mechanisms conferring the detected resistance.

299 Near-infrared spectroscopy (NIRS) is another rapid technique that has been utilised for
300 examining invertebrates which is non-destructive and cost-effective (Johnson 2020). The
301 high sensitivity spectrometers required for NIRS analysis cost an estimated $45,000 -
302 $60,000 USD (Ferguson et al. 2009, Fernandes et al. 2018, Maia et al. 2019). The technique
303 has been used successfully to differentiate mosquito species and age (Ferguson et al. 2009,
304 Sikulu et al. 2010, 2011, Dowell et al. 2015, González Jiménez et al. 2019) and can also
305 identify mosquitoes which are infected with arboviruses, *Plasmodium* and *Wolbachia*
306 (Sikulu-Lord et al. 2016, Fernandes et al. 2018, Maia et al. 2019). The ability of NIRS to
307 estimate age of mosquitoes has also been applied to the detection of insecticide resistance
308 (Sikulu et al. 2014, Lambert et al. 2018), as insecticide resistance has been shown to decrease
309 with age (Lines and Nassor 1991, Rajatileka et al. 2011, Jones et al. 2012). However there
310 has been no studies which investigate the use of NIRS to directly measure insecticide
311 resistance. The accuracy of NIRS for mosquito species determination is reported to be 78 –
312 90% (Ferguson et al. 2009, Sikulu et al. 2010, 2011, González Jiménez et al. 2019), lower
313 than the 91 – 100% REIMS accuracy for species differentiation in *Drosophila* (Wagner et al.
314 2020). As NIRS has not been used to directly monitor insecticide resistance, comparisons
315 between REIMS and NIRS accuracy for this purpose cannot be made.

316 This study focussed on identifying resistance to temephos however resistance to one
317 insecticide rarely occurs in isolation. *Ae. aegypti* from both Cúcuta and Bello have previously
318 been reported to have resistance to the pyrethroid permethrin and Cúcuta also to lambda-
319 cyhalothrin (Granada et al. 2021). Whilst the current study provides proof of concept for the
320 potential use of REIMS in identifying resistance, further study is needed to establish whether
321 the tool can be used to differentiate between resistance to different insecticides, an
322 application which could be beneficial to vector control programmes. Knock down resistance
323 (*kdr)*, mutations in the sodium channel gene frequently associated with pyrethroid resistance,
324 has also been reported in *Ae. aegypti* from Bello and Cúcuta. The varying frequencies of *kdr*
325 alleles demonstrates that these populations are not genetically homogenous (Granada et al.
326 2021). Whilst gaining an understanding of the genetic basis of resistance is important (e.g. in
327 tracking resistance and development of new interventions) it has little direct impact on the
328 rapid decision making needed in the field (Vontas and Mavridis 2019). This study aims to
329 provide a method which fulfils the need for more rapid resistance phenotyping tools to
330 contribute to existing strategies without delving into the mechanisms contributing to this
331 however there is also a further potential application of REIMS in investigating the genetic
332 basis of resistance.

333 To reduce the confounding effects of phenotypic differences between populations unrelated
334 to resistance, this study used two different susceptible populations of *Ae. aegypti,* one of field
335 origin and a lab strain. Whilst this experimental design does reduce these confounding
336 effects, as shown when comparing gene expression (Morgan et al. 2021), it cannot mitigate
337 them completely and therefore other phenotypic differences between populations may be
338 contributing to the high REIMS accuracy. This cannot be fully avoided when using field
339 collected populations of mosquitoes.

Further testing is required to establish sensitivity of REIMS to more granular levels of resistance, resistance in other medically important mosquito species, resistance to a variety of insecticides as well as resistance in adult mosquitoes. Determining whether the preservation method of mosquito samples (e.g., desiccation, storage temperatures, fixation) affects results also has implications for field application. The results presented here identified REIMS as a promising alternative tool for the identification of insecticide resistance in mosquitoes. REIMS and similar modern phenotyping methods should be standardised and incorporated into existing insecticide resistance management strategies.

Supplementary Material

Supplementary File 1: R Code for analysing REIMS data. R coding for analysing REIMS data matrices, following data binning in OMB, using LDA and random forest classification models.

Supplementary Table 1: The REIMS data matrices. REIMS data following binning in OMB. Data organised by population type, population, and resistance status. Mass spectra displayed in 0.1 *m/z* wide bins from 50 – 1200 *m/z.*

Supplementary Figures S1 – S11: Supplementary figures and figure legends. Separation of data with random group assignment (Fig S1). LDA and RF validation plots (Fig S2-S11).

Acknowledgments

The authors thank The University of Liverpool for the support to JES-S, IW and RJB, Edge Hill University for support to JM and CS and University of Antioquia for support to OT-C. The REIMS instrumentation was supported by a grant from the Biological and Biotechnological Sciences Research Council (BB/L014793/1) to RJB.

Author Contributions

JM: data curation; formal analysis; investigation; methodology; resources; visualization; writing – original draft, writing – review & editing. JES-S: conceptualization; supervision; writing – review and editing. IW: methodology, investigation, writing – review and editing. RJB: methodology; resources; writing – review and editing. OT-C: resources; writing – review and editing. CS: conceptualization; supervision; writing – review and editing.

References Cited

Alexander, J., L. Gildea, J. Balog, A. Speller, J. McKenzie, L. Muirhead, A. Scott, C. Kontovounisios, S. Rasheed, J. Teare, J. Hoare, K. Veselkov, R. Goldin, P. Tekkis, A. Darzi, J. Nicholson, J. Kinross, and Z. Takats. **2017**. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. Surg. Endosc. 31: 1361–1370.

Amelia-Yap, Z. H., C. D. Chen, M. Sofian-Azirun, and V. L. Low. **2018**. Pyrethroid resistance in the dengue vector Aedes aegypti in Southeast Asia: Present situation and prospects for management. Parasites and Vectors. 11: 332.

Balog, J., S. Kumar, J. Alexander, O. Golf, J. Huang, T. Wiggins, N. Abbassi-Ghadi, A. Enyedi, S. Kacska, J. Kinross, G. B. Hanna, J. K. Nicholson, and Z. Takats. **2015**. In Vivo Endoscopic Tissue Identification by Rapid Evaporative Ionization Mass Spectrometry (REIMS). Angew. Chemie. 127: 11211–11214.

Balog, J., D. Perenyi, C. Guallar-Hoyas, A. Egri, S. D. Pringle, S. Stead, O. P. Chevallier, C. T. Elliott, and Z. Takats. **2016**. Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. J. Agric. Food Chem. 64: 4793–4800.

Balog, J., L. Sasi-Szabó, J. Kinross, M. R. Lewis, L. J. Muirhead, K. Veselkov, R. Mirnezami, B. Dezső, L. Damjanovich, A. Darzi, J. K. Nicholson, and Z. Takáts. **2013**. Intraoperative Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry. Sci. Transl. Med. 5: 194ra93-194ra93.

Balog, J., T. Szaniszlo, K.-C. Schaefer, J. Denes, A. Lopata, L. Godorhazy, D. Szalay, L. Balogh, L. Sasi-Szabo, M. Toth, and Z. Takats. **2010**. Identification of Biological Tissues by Rapid Evaporative Ionization Mass Spectrometry. Anal. Chem. 82: 7343–7350.

Beleites, C., U. Neugebauer, T. Bocklitz, C. Krafft, and J. Popp. **2013**. Sample size planning for classification models. Anal. Chim. Acta. 760: 25–33.

Bisset, J. A., M. M. Rodríguez, Y. Ricardo, H. Ranson, O. Pérez, M. Moya, and A. Vázquez. **2011**. Temephos resistance and esterase activity in the mosquito Aedes aegypti in Havana, Cuba increased dramatically between 2006 and 2008. Med. Vet. Entomol. 25: 233–239.

Black, C., O. P. Chevallier, S. A. Haughey, J. Balog, S. Stead, S. D. Pringle, M. V. Riina, F. Martucci, P. L. Acutis, M. Morris, D. S. Nikolopoulos, Z. Takats, and C. T. Elliott. **2017**. A real time metabolomic profiling approach to detecting fish fraud using rapid evaporative ionisation mass spectrometry. Metabolomics. 13: 153.

Bonetti, J. 2018. Mass spectral differentiation of positional isomers using multivariate statistics. Forensic Chem. 9: 50–61.

Brogdon, W. G. 1989. Biochemical resistance detection: An alternative to bioassay. Parasitol. Today. 5: 56–60.

Cameron, S. J. S., F. Bolt, A. Perdones-Montero, T. Rickards, K. Hardiman, A. Abdolrasouli, A. Burke, Z. Bodai, T. Karancsi, D. Simon, R. Schaffer, M. Rebec, J. Balog, and Z. Takáts. **2016**. Rapid Evaporative Ionisation Mass Spectrometry (REIMS) Provides Accurate Direct from Culture Species Identification within the Genus Candida. Sci. Rep. 6: 36788.

Carvalho, D. O., D. Nimmo, N. Naish, A. R. McKemey, P. Gray, A. B. B. Wilke, M. T. Marrelli, J. F. Virginio, L. Alphey, and M. L. Capurro. **2014**. Mass production of genetically modified Aedes aegypti for field releases in Brazil. J. Vis. Exp. e3579.

Corbel, V., and R. N'Guessan. **2013**. Distribution, Mechanisms, Impact and Management of Insecticide Resistance in Malaria Vectors: A Pragmatic Review. *In* Anopheles Mosquitoes - New Insights into Malar. Vectors. InTech.

D'Hue, C., M. Moore, D.-J. Summerlin, A. Jarmusch, C. Alfaro, A. Mantravadi, A. Bewley, D. Gregory Farwell, and R. G. Cooks. **2018**. Feasibility of desorption electrospray ionization mass spectrometry for diagnosis of oral tongue squamous cell carcinoma. Rapid Commun. Mass Spectrom. 32: 133–141.

Davidson, N. B., N. I. Koch, J. Sarsby, E. Jones, J. L. Hurst, and R. J. Beynon. **2019**. Rapid identification of species, sex and maturity by mass spectrometric analysis of animal faeces. BMC Biol. 17: 66.

Demok, S., N. Endersby-Harshman, R. Vinit, L. Timinao, L. J. Robinson, M. Susapu, L. Makita, M. Laman, A. Hoffmann, and S. Karl. **2019**. Insecticide resistance status of Aedes aegypti and Aedes albopictus mosquitoes in Papua New Guinea. Parasit. Vectors. 12: 333.

Dobbin, K. K., Y. Zhao, and R. M. Simon. **2008**. How Large a Training Set is Needed to Develop a Classifier for Microarray Data? Clin. Cancer Res. 14: 108–114.

Dowell, F., A. Lenhart, L. Vizcaino, I. Swamidoss, K. Liebman, and R. Wirtz. **2015**. 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. 92: 1070–1075.

Dusfour, I., J. Vontas, J. P. David, D. Weetman, D. M. Fonseca, V. Corbel, K. Raghavendra, M. B. Coulibaly, A. J. Martins, S. Kasai, and F. Chandre. **2019**. Management of insecticide resistance in the major Aedes vectors of arboviruses: Advances and challenges. PLoS Negl. Trop. Dis.

Faucon, F., T. Gaude, I. Dusfour, V. Navratil, V. Corbel, W. Juntarajumnong, R. Girod, R. Poupardin, F. Boyer, S. Reynaud, and J. P. David. **2017**. In the hunt for genomic markers of metabolic resistance to pyrethroids in the mosquito Aedes aegypti: An integrated next-generation sequencing approach. PLoS Negl. Trop. Dis. 11:

443 e0005526.

444 **Ferguson, H. M., G. F. Killeen, K. Michel, R. A. Wirtz, M. Q. Benedict, F. E. Dowell,**
445 **and V. S. Mayagaya**. **2009**. Non-destructive Determination of Age and Species of
446 Anopheles gambiae s.l. Using Near-infrared Spectroscopy. Am. J. Trop. Med. Hyg. 81:
447 622–630.

448 **Fernandes, J. N., L. M. B. dos Santos, T. Chouin-Carneiro, M. G. Pavan, G. A. Garcia,**
449 **M. R. David, J. C. Beier, F. E. Dowell, R. Maciel-de-Freitas, and M. T. Sikulu-**
450 **Lord**. **2018**. Rapid, noninvasive detection of Zika virus in *Aedes aegypti* mosquitoes by
451 near-infrared spectroscopy. Sci. Adv. 4: eaat0496.

452 **Figueroa, R. L., Q. Zeng-Treitler, S. Kandula, and L. H. Ngo**. **2012**. Predicting sample
453 size required for classification performance. BMC Med. Inform. Decis. Mak. 12: 8.

454 **González Jiménez, M., S. A. Babayan, P. Khazaeli, M. Doyle, F. Walton, E. Reedy, T.**
455 **Glew, M. Viana, L. Ranford-Cartwright, A. Niang, D. J. Siria, F. O. Okumu, A.**
456 **Diabaté, H. M. Ferguson, F. Baldini, and K. Wynne**. **2019**. Prediction of mosquito
457 species and population age structure using mid-infrared spectroscopy and supervised
458 machine learning. Wellcome Open Res. 4: 76.

459 **Granada, Y., A. M. Mejía-Jaramillo, S. Zuluaga, and O. Triana-Chávez**. **2021**.
460 Molecular surveillance of resistance to pyrethroids insecticides in Colombian Aedes
461 aegypti populations. PLoS Negl. Trop. Dis. 15: e0010001.

462 **Gredell, D. A., A. R. Schroeder, K. E. Belk, C. D. Broeckling, A. L. Heuberger, S.-Y.**
463 **Kim, D. A. King, S. D. Shackelford, J. L. Sharp, T. L. Wheeler, D. R. Woerner, and**
464 **J. E. Prenni**. **2019**. Comparison of Machine Learning Algorithms for Predictive
465 Modeling of Beef Attributes Using Rapid Evaporative Ionization Mass Spectrometry
466 (REIMS) Data. Sci. Rep. 9: 5721.

467 **Gromski, P. S., H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner, and R.**
468 **Goodacre**. **2015**. A tutorial review: Metabolomics and partial least squares-discriminant
469 analysis – a marriage of convenience or a shotgun wedding. Anal. Chim. Acta. 879: 10–
470 23.

471 **Guedes, R. N. C., K. Beins, D. Navarro Costa, G. E. Coelho, and H. S. da S. Bezerra**.
472 **2020**. Patterns of insecticide resistance in *Aedes aegypti*□: meta□analyses of surveys in
473 Latin America and the Caribbean. Pest Manag. Sci. 76: 2144–2157.

474 **Guitton, Y., G. Dervilly-Pinel, R. Jandova, S. Stead, Z. Takats, and B. Le Bizec**. **2018**.
475 Rapid evaporative ionisation mass spectrometry and chemometrics for high-throughput
476 screening of growth promoters in meat producing animals. Food Addit. Contam. Part A.
477 35: 900–910.

478 **Hanberry, B. B., H. S. He, and D. C. Dey**. **2012**. Sample sizes and model comparison
479 metrics for species distribution models. Ecol. Modell. 227: 29–33.

480 **Hemingway, J., N. J. Hawkes, L. McCarroll, and H. Ranson**. **2004**. The molecular basis
481 of insecticide resistance in mosquitoes. Insect Biochem. Mol. Biol. 34: 653–665.

482 **Hemingway, J., J. Vontas, R. Poupardin, J. Raman, J. Lines, C. Schwabe, A. Matias,**
483 **and I. Kleinschmidt**. **2013**. Country-level operational implementation of the Global
484 Plan for Insecticide Resistance Management. Proc. Natl. Acad. Sci. U. S. A. 110: 9397–
485 9402.

Johnson, J. 2020. Near□infrared spectroscopy (NIRS) for taxonomic entomology: A brief review. J. Appl. Entomol. 144: 241–250.

Jones, C. M., A. Sanou, W. M. Guelbeogo, N. Sagnon, P. C. Johnson, and H. Ranson. **2012**. Aging partially restores the efficacy of malaria vector control in insecticide-resistant populations of Anopheles gambiae s.l. from Burkina Faso. Malar. J. 11: 24.

Kassambara, A. 2020. ggpubr: "ggplot2" Based Publication Ready Plots.

Kassambara, A., and F. Mundt. **2020**. factoextra: Extract and Visualize the Results of Multivariate Data Analyses Title.

Kenar, A., B. Çiçek, F. N. Arslan, G. Akin, Ş. N. K. Elmas, and I. Yilmaz. **2019**. Electron Impact–Mass Spectrometry Fingerprinting and Chemometrics for Rapid Assessment of Authenticity of Edible Oils Based on Fatty Acid Profiling. Food Anal. Methods. 12: 1369–1381.

Kuhn, M. 2021. caret: Classification and Regression Training.

Lambert, B., M. T. Sikulu-Lord, V. S. Mayagaya, G. Devine, F. Dowell, and T. S. Churcher. **2018**. Monitoring the Age of Mosquito Populations Using Near-Infrared Spectroscopy. Sci. Rep. 8: 5274.

Liaw, A., and M. Wiener. **2002**. Classification and Regression by randomForest. R News. 2: 18–22.

Lines, J. D., and N. S. Nassor. **1991**. DDT resistance in Anopheles gambiae declines with mosquito age. Med. Vet. Entomol. 5: 261–265.

Liu, H., Q. Meng, X. Zhao, Y. Ye, and H. Tong. **2021**. Inductively coupled plasma mass spectrometry (ICP-MS) and inductively coupled plasma optical emission spectrometer (ICP-OES)-based discrimination for the authentication of tea. Food Control. 123: 107735.

Logrono, J. B. 2020. Thesis Assessment Of Methods To Screen For Carotenoids IN Yellow-Fleshed Potato Germplasm.

Luan, J., C. Zhang, B. Xu, Y. Xue, and Y. Ren. **2020**. The predictive performances of random forest models with limited sample size and different species traits. Fish. Res. 227: 105534.

Maia, M. F., M. Kapulu, M. Muthui, M. G. Wagah, H. M. Ferguson, F. E. Dowell, F. Baldini, and L. Ranford-Cartwright. **2019**. Detection of Plasmodium falciparum infected Anopheles gambiae using near-infrared spectroscopy. Malar. J. 18: 85.

Marcombe, S., F. Darriet, M. Tolosa, P. Agnew, S. Duchon, M. Etienne, M. M. Y. Tcha, F. Chandre, V. Corbel, and A. Yébakima. **2011**. Pyrethroid resistance reduces the efficacy of space sprays for Dengue control on the island of Martinique (Caribbean). PLoS Negl. Trop. Dis. 5: e1202.

Marcombe, S., A. Farajollahi, S. P. Healy, G. G. Clark, and D. M. Fonseca. **2014**. Insecticide resistance status of United States populations of Aedes albopictus and mechanisms involved. PLoS One. 9: e101992.

Morgan, J., J. E. Salcedo-Sora, O. Triana-Chavez, and C. Strode. **2021**. Expansive and Diverse Phenotypic Landscape of Field *Aedes aegypti* (Diptera: Culicidae) Larvae with

Differential Susceptibility to Temephos: Beyond Metabolic Detoxification. J. Med. Entomol.

Moyes, C. L., J. Vontas, A. J. Martins, L. C. Ng, S. Y. Koou, I. Dusfour, K. Raghavendra, J. Pinto, V. Corbel, J. P. David, and D. Weetman. **2017**. Contemporary status of insecticide resistance in the major Aedes vectors of arboviruses infecting humans. PLoS Negl. Trop. Dis. 11: e0005625.

Owusu, H. F., N. Chitnis, and P. Müller. **2017**. Insecticide susceptibility of Anopheles mosquitoes changes in response to variations in the larval environment. Sci. Rep. 7: 3667.

Paraskevaidi, M., S. J. S. Cameron, E. Whelan, S. Bowden, M. Tzafetas, A. Mitra, A. Semertzidou, A. Athanasiou, P. R. Bennett, D. A. MacIntyre, Z. Takats, and M. Kyrgiou. **2020**. Laser-assisted rapid evaporative ionisation mass spectrometry (LA-REIMS) as a metabolomics platform in cervical cancer screening. EBioMedicine. 60: 103017.

Phelps, D. L., J. Balog, L. F. Gildea, Z. Bodai, A. Savage, M. A. El-Bahrawy, A. V. Speller, F. Rosini, H. Kudo, J. S. McKenzie, R. Brown, Z. Takáts, and S. Ghaem-Maghami. **2018**. The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS). Br. J. Cancer. 118: 1349–1358.

R Core Team. **2020**. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Viennna, Austria.

Rajatileka, S., J. Burhani, and H. Ranson. **2011**. Mosquito age and susceptibility to insecticides. Trans. R. Soc. Trop. Med. Hyg. 105: 247–253.

Ranson, H., J. Burhani, N. Lumjuan, and W. C. Black IV. **2010**. Insecticide resistance in dengue vectors. TropIKA. 1.

Rigano, F., S. Stead, D. Mangraviti, R. Jandova, D. Petit, N. Marino, and L. Mondello. **2019**. Use of an “Intelligent Knife” (iknife), Based on the Rapid Evaporative Ionization Mass Spectrometry Technology, for Authenticity Assessment of Pistachio Samples. Food Anal. Methods. 12: 558–568.

Sarsby, J., L. McLean, V. M. Harman, and R. J. Beynon. **2021**. Monitoring recombinant protein expression in bacteria by rapid evaporative ionisation mass spectrometry. Rapid Commun. Mass Spectrom. 35: e8670.

Schäfer, K.-C., J. Dénes, K. Albrecht, T. Szaniszló, J. Balog, R. Skoumal, M. Katona, M. Tóth, L. Balogh, and Z. Takáts. **2009**. In Vivo, In Situ Tissue Analysis Using Rapid Evaporative Ionization Mass Spectrometry. Angew. Chemie Int. Ed. 48: 8240–8242.

Seixas, G., L. Grigoraki, D. Weetman, J. L. Vicente, A. C. Silva, J. Pinto, J. Vontas, and C. A. Sousa. **2017**. Insecticide resistance is mediated by multiple mechanisms in recently introduced Aedes aegypti from Madeira Island (Portugal). PLoS Negl. Trop. Dis. 11: e0005799.

Sikulu-Lord, M. T., M. P. Milali, M. Henry, R. A. Wirtz, L. E. Hugo, F. E. Dowell, and G. J. Devine. **2016**. Near-Infrared Spectroscopy, a Rapid Method for Predicting the Age of Male and Female Wild-Type and Wolbachia Infected Aedes aegypti. PLoS Negl. Trop. Dis. 10: e0005040.

Sikulu, M., K. M. Dowell, L. E. Hugo, R. A. Wirtz, K. Michel, K. H. Peiris, S. Moore, G. F. Killeen, and F. E. Dowell. **2011**. Evaluating RNAlater® as a preservative for using near-infrared spectroscopy to predict Anopheles gambiae age and species. Malar. J. 10: 186.

Sikulu, M., G. F. Killeen, L. E. Hugo, P. A. Ryan, K. M. Dowell, R. A. Wirtz, S. J. Moore, and F. E. Dowell. **2010**. Near-infrared spectroscopy as a complementary age grading and species identification tool for African malaria vectors. Parasit. Vectors. 3: 49.

Sikulu, M. T., S. Majambere, B. O. Khatib, A. S. Ali, L. E. Hugo, and F. E. Dowell. **2014**. Using a Near-Infrared Spectrometer to Estimate the Age of Anopheles Mosquitoes Exposed to Pyrethroids. PLoS One. 9: e90657.

St John, E. R., J. Balog, J. S. McKenzie, M. Rossi, A. Covington, L. Muirhead, Z. Bodai, F. Rosini, A. V. M. Speller, S. Shousha, R. Ramakrishnan, A. Darzi, Z. Takats, and D. R. Leff. **2017**. Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: towards an intelligent knife for breast cancer surgery. Breast Cancer Res. 19: 59.

Strittmatter, N., E. A. Jones, K. A. Veselkov, M. Rebec, J. G. Bundy, and Z. Takats. **2013**. Analysis of intact bacteria using rapid evaporative ionisation mass spectrometry. Chem. Commun. 49: 6188.

Strittmatter, N., M. Rebec, E. A. Jones, O. Golf, A. Abdolrasouli, J. Balog, V. Behrends, K. A. Veselkov, and Z. Takats. **2014**. Characterization and Identification of Clinically Relevant Microorganisms Using Rapid Evaporative Ionization Mass Spectrometry. Anal. Chem. 86: 6555–6562.

Verplanken, K., S. Stead, R. Jandova, C. Van Poucke, J. Claereboudt, J. Vanden Bussche, S. De Saeger, Z. Takats, J. Wauters, and L. Vanhaecke. **2017**. Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: The case of boar taint. Talanta. 169: 30–36.

Viana-Medeiros, P. F., A. J. Martins, I. A. Braga, J. B. P. Lima, D. Valle, and I. R. Montella. **2007**. Insecticide Resistance Mechanisms of Brazilian Aedes aegypti Populations from 2001 to 2004. Am. J. Trop. Med. Hyg. 77: 467–477.

Vontas, J., E. Kioulos, N. Pavlidi, E. Morou, A. Della Torre, and H. Ranson. **2012**. Insecticide resistance in the major dengue vectors Aedes albopictus and Aedes aegypti.

Vontas, J., and K. Mavridis. **2019**. Vector population monitoring tools for insecticide resistance management: Myth or fact? Pestic. Biochem. Physiol. 161: 54–60.

W. N. Venables and B. D. Ripley. **2002**. Modern Applied Statistics with S.

Wagner, I., N. I. Koch, J. Sarsby, N. White, T. A. R. Price, S. Jones, J. L. Hurst, and R. J. Beynon. **2020**. The application of rapid evaporative ionization mass spectrometry in the analysis of *Drosophila* species—a potential new tool in entomology. Open Biol. 10: 200196.

Wang, Y., T. He, J. Wang, L. Wang, X. Ren, S. He, X. Liu, Y. Dong, J. Ma, R. Song, J. Wei, A. Yu, Q. Fan, X. Wang, and G. She. **2021**. High performance liquid chromatography fingerprint and headspace gas chromatography-mass spectrometry combined with chemometrics for the species authentication of Curcumae Rhizoma. J.

Pharm. Biomed. Anal. 202: 114144.

Weetman, D., B. Kamgang, A. Badolo, C. Moyes, F. Shearer, M. Coulibaly, J. Pinto, L. Lambrechts, and P. McCall. **2018**. Aedes Mosquitoes and Aedes-Borne Arboviruses in Africa: Current and Future Threats. Int. J. Environ. Res. Public Health. 15: 220.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis.

World Health Organization. **2013**. Test procedures for insecticide resistance monitoring in malaria vector mosquitoes. World Heal. Organ. Tech. Rep. Ser. 22.

World Health Organization (WHO). **1998**. Techniques to detect insecticide resistance mechanisms (field and laboratory manual).

World Health Organization (WHO). **2016**. Monitoring and Managing Insecticide Resistance in Aedes mosquito Populations, Who.

Zheng, M.-L., D.-J. Zhang, D. D. Damiens, H. Yamada, and J. R. L. Gilles. **2015**. Standard operating procedures for standardized mass rearing of the dengue and chikungunya vectors Aedes aegypti and Aedes albopictus (Diptera: Culicidae) - I - egg quantification. Parasit. Vectors. 8: 1–7.

Fig. legends

Fig. 1: Block diagram of the experimental approach. This study utilised insecticide resistant and susceptible larvae of the mosquito *Ae. aegypti*. The resistant larvae originated from Cúcuta, Colombia and the susceptible larvae had dual origin, field samples from Bello, Colombia (Field Susceptible) and the New Orleans lab strain (Lab Susceptible). Individual larvae from each experimental group were analysed using REIMS to acquire individual mass spectra for each sample. The data acquired through REIMS was background and lock mass corrected and binned into 0.1 m/z groups. Dimension reduction was conducted using PCA before LDA and random forest classification model building and testing.

Fig. 2: REIMS discrimination of *Ae. aegypti* samples by population. Combined PCA-LDA separation of the three *Ae. aegypti* populations using REIMS mass spectra (A). Dimension reduction was conducted using principal components analysis (PCA), 40 principal components were selected for linear discriminant analysis (LDA). The number of PCs was determined by selecting the model with the lowest area under the ROC curve (AUC) (Supp

643 Fig.S2). Separation is shown in both linear discriminant one and linear discriminant 2. All
644 populations separated in linear discriminant 1 whilst field resistant separated from the two
645 susceptible populations in LD2. Classification of samples into population using PCA-LDA
646 (B) and random forest models (C), showing percentage of samples classified to each group,
647 standard error of the mean (SEM) and the percentage range across all replicates. Models were
648 built and tested 20 times each with a different set of training (70%) and test (30%) data.
649 Accuracy percentages, SEM and range were averaged across all 20 replicates. The PCA-LDA
650 classification model had a higher accuracy (82% ± 0.01) than the random forest model (76%
651 ± 0.02), correctly assigning 82% of individuals to their respective population. Random forest
652 models were built using 20 PCs to obtain the highest accuracy of models tested (Supp
653 Fig.S3).

654 **Fig. 3: REIMS discrimination of *Ae. aegypti* by population type (lab and field).**
655 Combined PCA-LDA separation of lab and field *Ae. aegypti* populations using REIMS mass
656 spectra (A & B). Dimension reduction was conducted using principal components analysis
657 (PCA), 40 principal components were selected for linear discriminant analysis (LDA). The
658 number of PCs was determined by selecting the model with the lowest area under the ROC
659 curve (AUC) (Supp Fig.S4). Classification of samples into resistance status using PCA-LDA
660 (C) and random forest models (D), showing percentage of samples classified to each group,
661 standard error of the mean (SEM) and the percentage range across all replicates. Models were
662 built and tested 20 times each with a different set of training (70%) and test (30%) data.
663 Accuracy percentages, SEM and range were averaged across all 20 replicates. The LDA-PCA
664 classification model had a higher accuracy (89% ± 0.01) than the random forest model (83%
665 ± 0.02), correctly assigning 89% of individuals to their respective resistance status. Random
666 forest models were built using 20 PCs to obtain the highest accuracy of models tested (Supp
667 Fig.S5).

668 **Fig. 4: REIMS discrimination of resistant and susceptible *Ae. aegypti*.** Combined PCA-
669 LDA separation of resistant and susceptible *Ae. aegypti* populations using REIMS mass
670 spectra (A & B). Dimension reduction was conducted using principal components analysis
671 (PCA), 40 principal components were selected for linear discriminant analysis (LDA). The
672 number of PCs was determined by selecting the model with the lowest area under the ROC
673 curve (AUC) (Supp Fig.S6). Classification of samples into resistance status using PCA-LDA
674 (C) and random forest models (D), showing percentage of samples classified to each group,
675 standard error of the mean (SEM) and the percentage range across all replicates. Models were
676 built and tested 20 times each with a different set of training (70%) and test (30%) data.
677 Accuracy percentages, SEM and range were averaged across all 20 replicates. The LDA-PCA
678 classification model had a higher accuracy (85% ± 0.01) than the random forest model (78%
679 ± 0.02), correctly assigning 85% of individuals to their respective resistance status. Random
680 forest models were built using 20 PCs to obtain the highest accuracy of models tested (Supp
681 Fig.S7).

682 **Fig. 5: REIMS discrimination of field resistant and lab susceptible *Ae. aegypti* larvae.**
683 Combined PCA-LDA separation of the resistant and lab susceptible populations using
684 REIMS mass spectra (A & B). Dimension reduction was conducted using principal
685 components analysis (PCA), 20 principal components were selected for linear discriminant
686 analysis (LDA). The number of PCs was determined by selecting the model with the lowest
687 area under the ROC curve (AUC) (Supp Fig.S8). Classification of samples into population
688 using PCA-LDA (C) and random forest models (D), showing percentage of samples
689 classified to each group, standard error of the mean (SEM) and the percentage range across
690 all replicates. Models were built and tested 20 times each with a different set of training
691 (70%) and test (30%) data. Accuracy percentages, SEM and range were averaged across all
692 20 replicates. The LDA-PCA classification model had a higher accuracy (87% ± 0.02) than

693 the random forest model (82% ± 0.02), correctly assigning 87% of individuals to their
694 respective resistance status. Random forest models were built using 10 PCs to obtain the
695 highest accuracy of models tested (Supp Fig.S9).

696 **Fig. 6: REIMS discrimination of field resistant and field susceptible *Ae. aegypti* larvae.**
697 Combined PCA-LDA separation of the resistant and field susceptible populations using
698 REIMS mass spectra (A & B). Dimension reduction was conducted using principal
699 components analysis (PCA), 20 principal components were selected for linear discriminant
700 analysis (LDA). The number of PCs was determined by selecting the model with the lowest
701 area under the ROC curve (AUC) (Supp Fig.S10). Classification of samples into population
702 using PCA-LDA (C) and random forest models (D), showing percentage of samples
703 classified to each group, standard error of the mean (SEM) and the percentage range across
704 all replicates. Models were built and tested 20 times each with a different set of training
705 (70%) and test (30%) data. Accuracy percentages, SEM and range were averaged across all
706 20 replicates. The LDA-PCA classification model had a higher accuracy (88% ± 0.01) than
707 the random forest model (84% ± 0.02), correctly assigning 88% of individuals to their
708 respective resistance status. Random forest models were built using 20 PCs to obtain the
709 highest accuracy of models tested (Supp Fig.S11).

Tables

Table 1: Summary data of the *Ae. aegypti* samples analysed via REIMS. Time larvae stored at -20^{o}C in weeks for each replicate and the number of larvae analysed in each replicate and the total number for each experimental group (*n*).

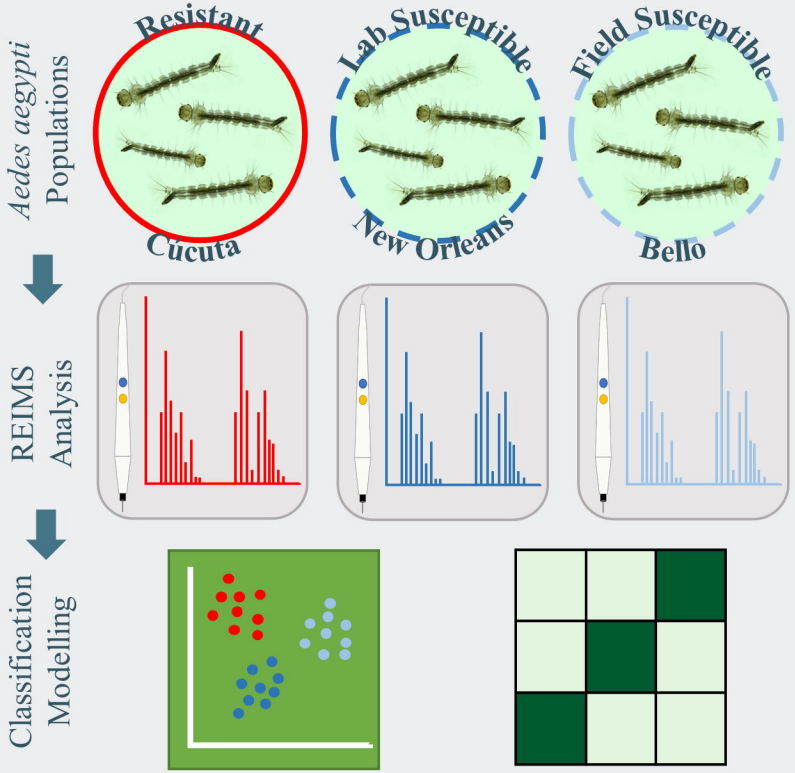
Population	Replicate	Storage Weeks	*n*
Lab Susceptible	1	36	8
	2	36	15
	3	36	13
	4	32	15
	Total	32-36	51
Field Susceptible	1	32	12
	2	34	13
	3	33	13
	4	32	13
	Total	32-34	51
Field Resistant	1	36	9
	2	32	14
	3	36	10
	4	36	9
	Total	32-36	42

Rapid evaporative ionisation mass spectrometry (REIMS): A potential rapid tool for the identification of insecticide resistance in mosquito larvae

Background and rationale

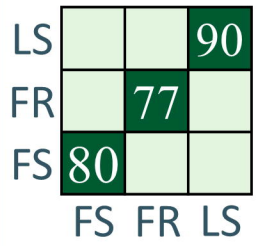
- *Aedes aegypti* is the principle vector for dengue, Zika and chikungunya.
- Control of *Ae. aegypti* and other vector mosquitoes is threatened by widespread insecticide resistance.
- Operational need for rapid tools for identifying and monitoring insecticide resistance.

Experimental Design



Results

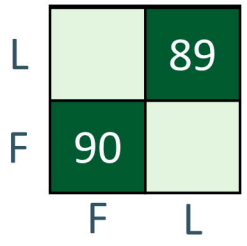
Population



Field Susceptible = FS
Field Resistant = FR
Lab Susceptible = LS

Overall model accuracy = 82%

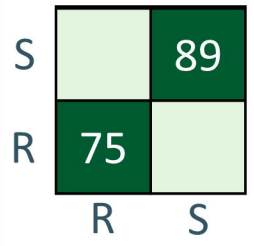
Population Type



Field = F
Lab = L

Overall model accuracy = 89%

Resistance Status

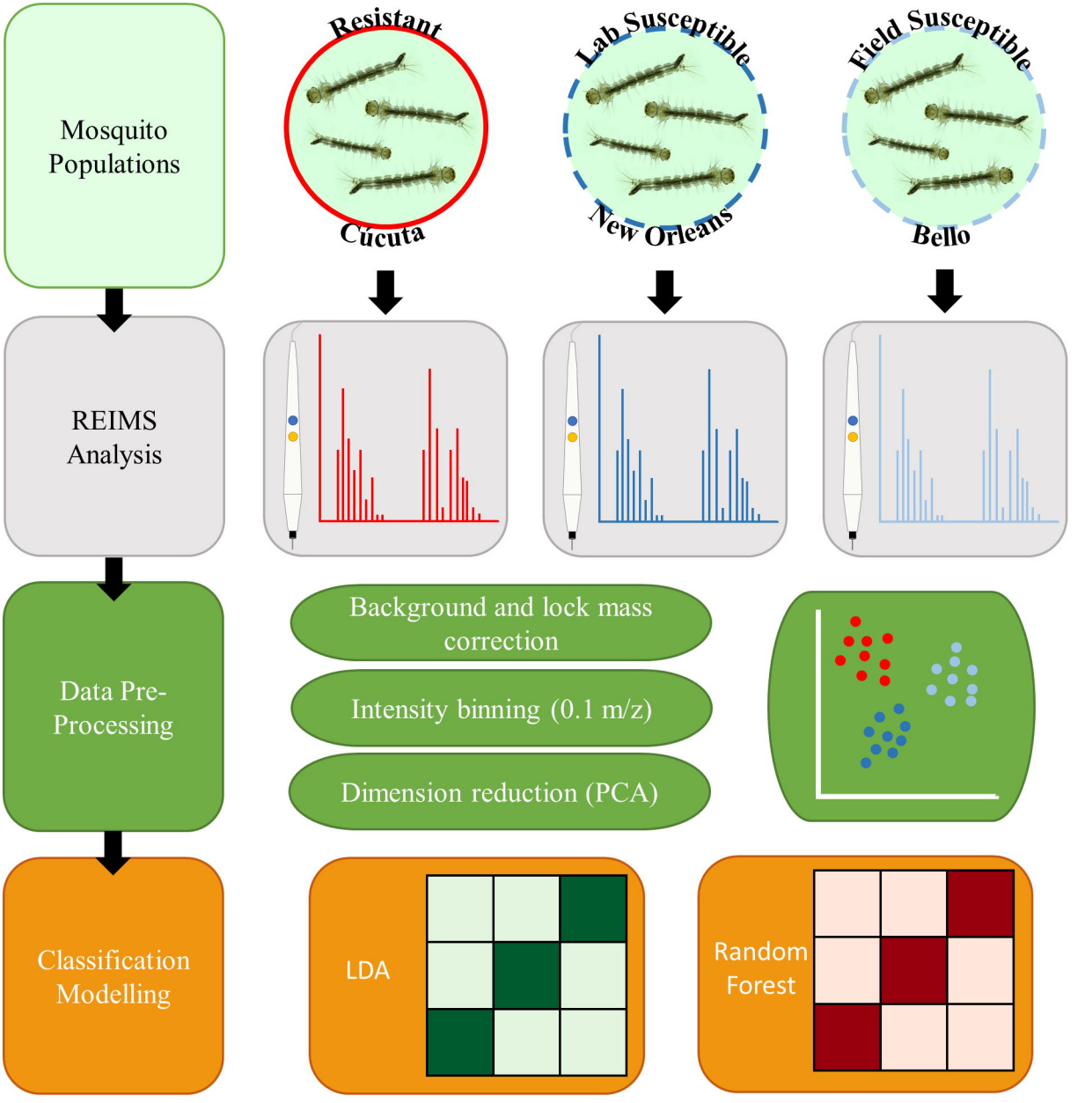


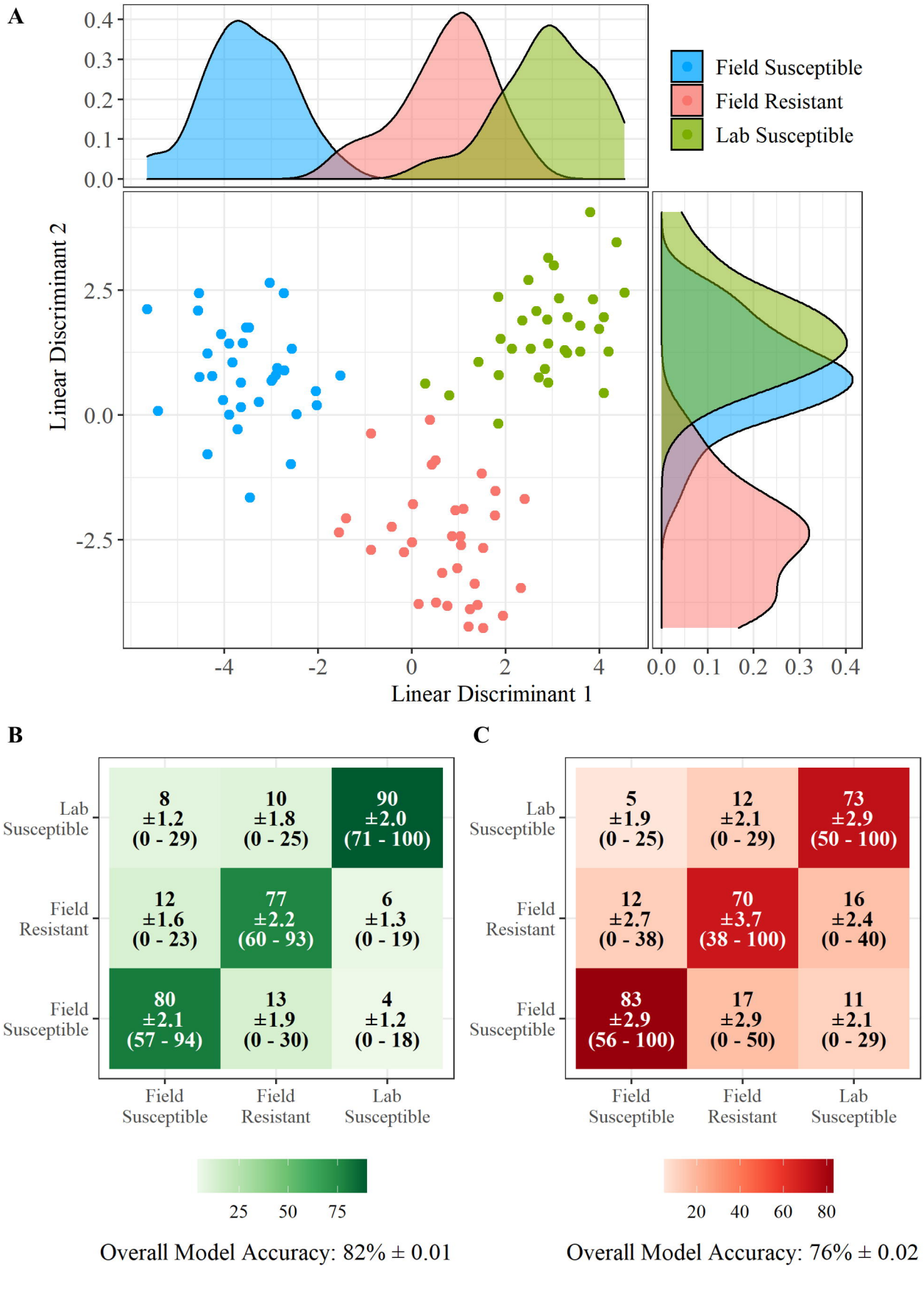
Resistant = R
Susceptible = S

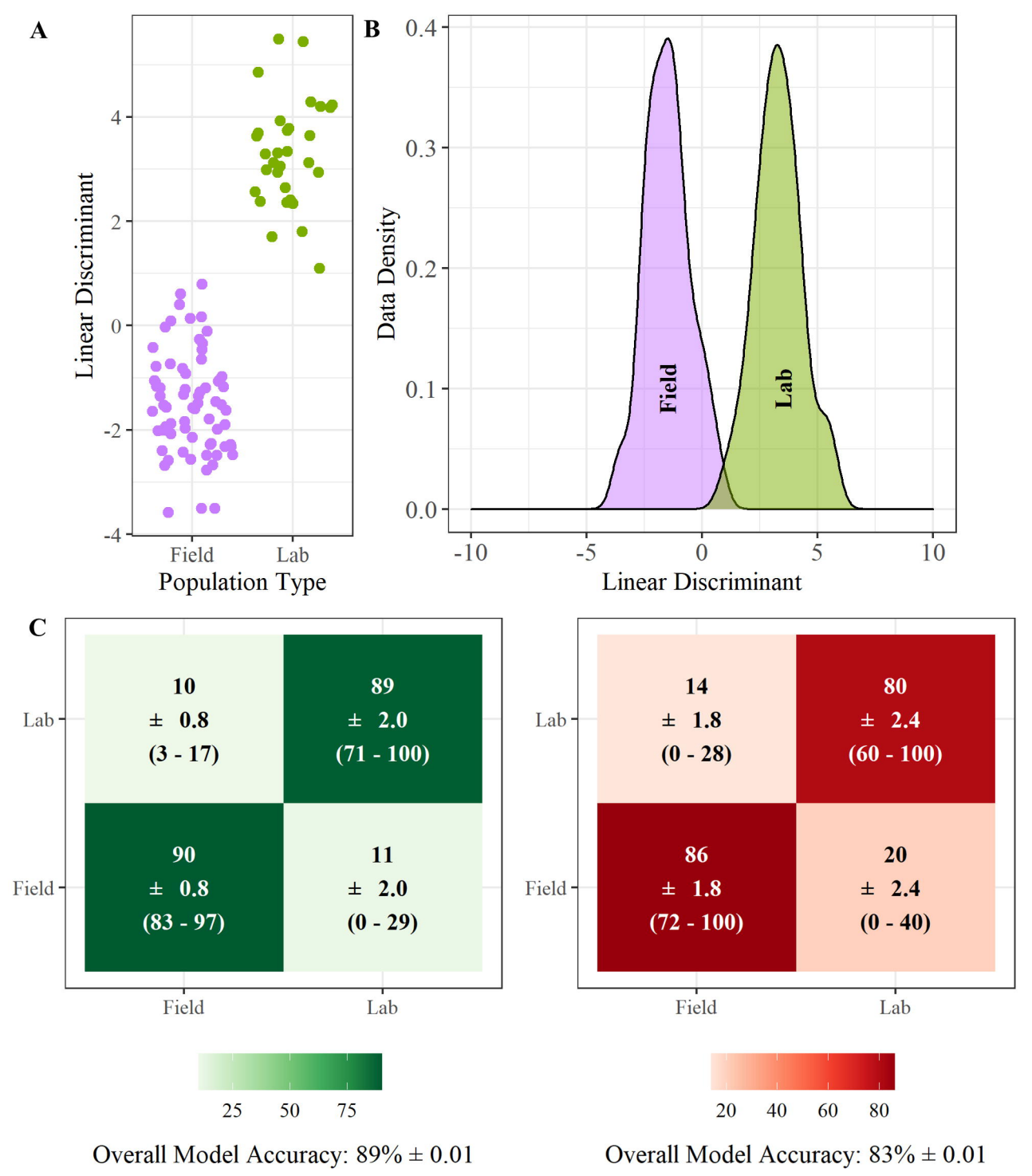
Overall model accuracy = 85%

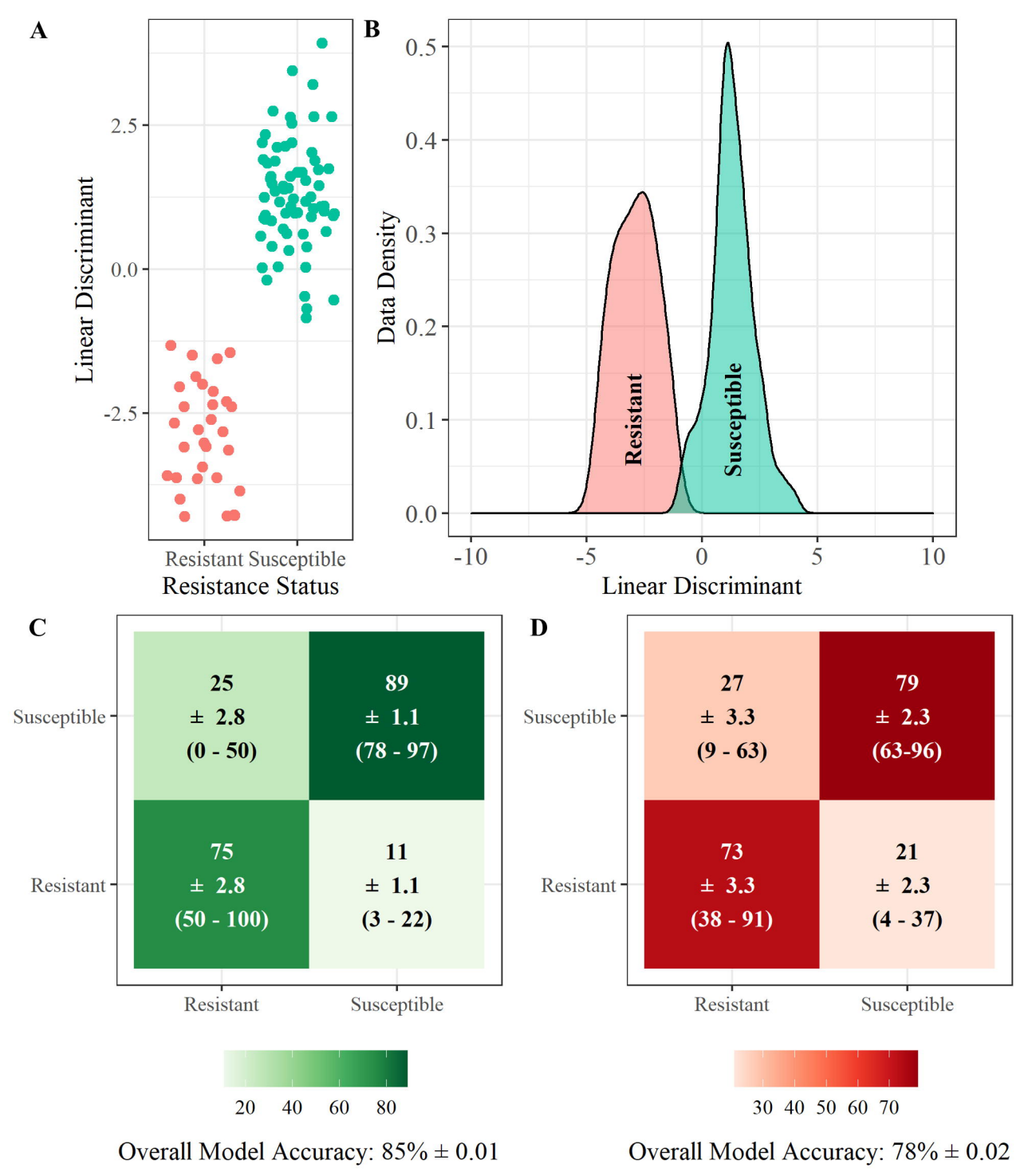
Conclusions

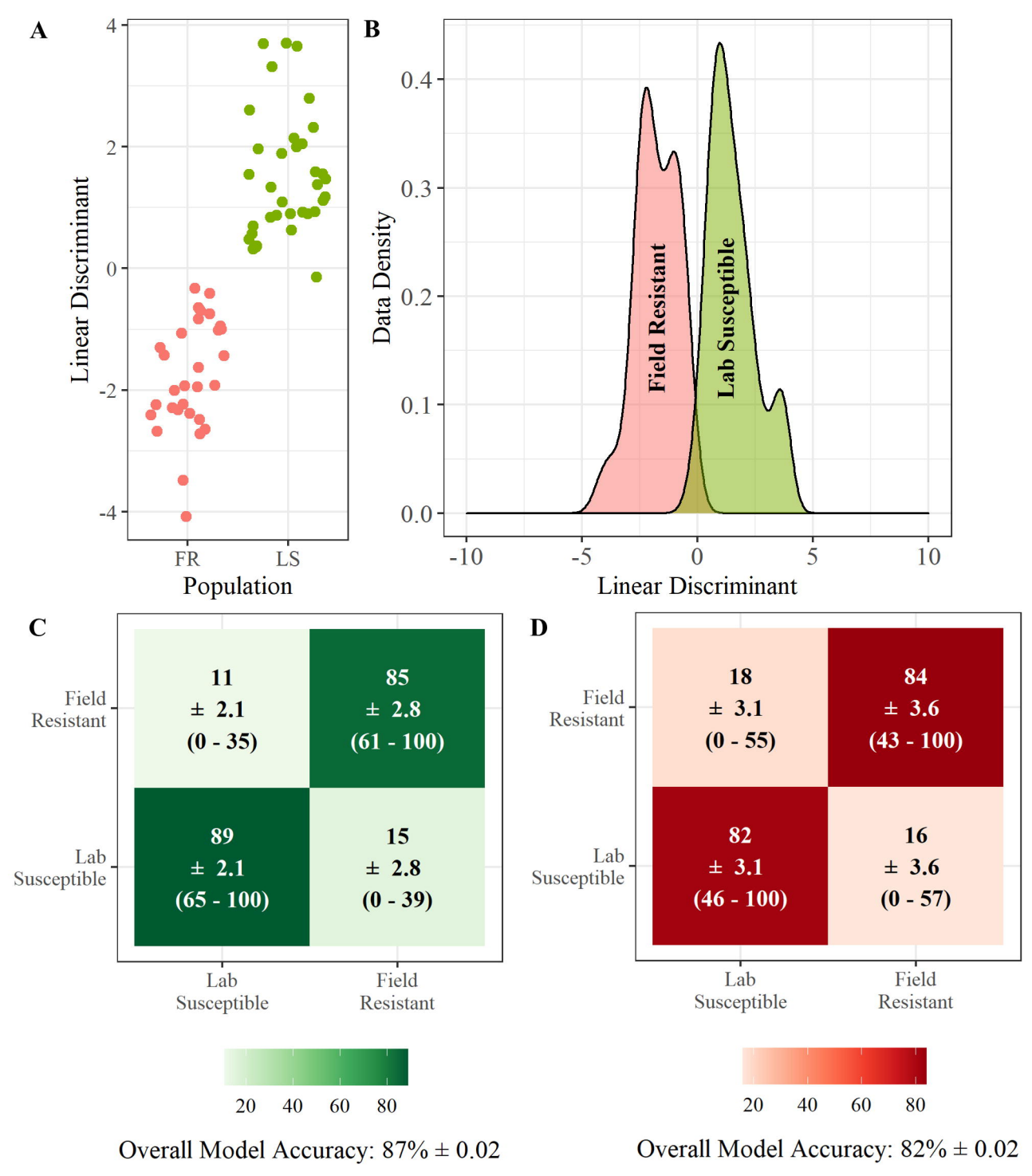
- REIMS identified as a potential novel tool for rapid identification of insecticide resistance with 85% accuracy.
- High accuracy in geographical origin (82%) and population type (89%) is also reported suggesting other potential applications of REIMS in vector surveillance.

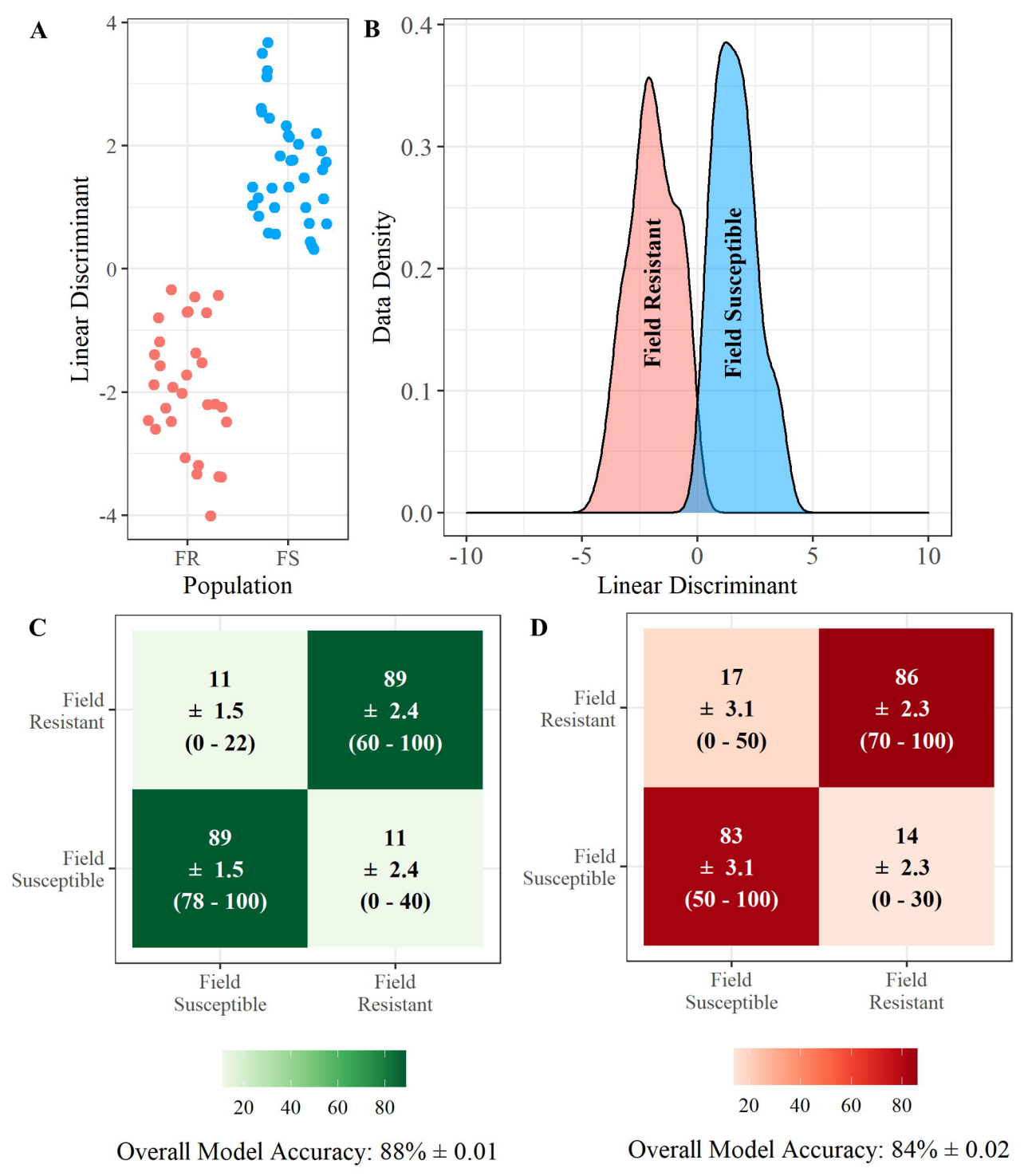

Mosquito Populations
Resistant
Cúcuta
Lab Susceptible
New Orleans
Field Susceptible
Bello
REIMS Analysis
Data Pre-Processing
Background and lock mass correction
Intensity binning (0.1 m/z)
Dimension reduction (PCA)
Classification Modelling
LDA
Random Forest


A
Field Susceptible
Field Resistant
Lab Susceptible
Linear Discriminant 2
Linear Discriminant 1
B
Lab Susceptible
Field Resistant
Field Susceptible
8 ±1.2 (0 - 29)
10 ±1.8 (0 - 25)
90 ±2.0 (71 - 100)
12 ±1.6 (0 - 23)
77 ±2.2 (60 - 93)
6 ±1.3 (0 - 19)
80 ±2.1 (57 - 94)
13 ±1.9 (0 - 30)
4 ±1.2 (0 - 18)
Field Susceptible
Field Resistant
Lab Susceptible
25 50 75
Overall Model Accuracy: 82% ± 0.01
C
Lab Susceptible
Field Resistant
Field Susceptible
5 ±1.9 (0 - 25)
12 ±2.1 (0 - 29)
73 ±2.9 (50 - 100)
12 ±2.7 (0 - 38)
70 ±3.7 (38 - 100)
16 ±2.4 (0 - 40)
83 ±2.9 (56 - 100)
17 ±2.9 (0 - 50)
11 ±2.1 (0 - 29)
Field Susceptible
Field Resistant
Lab Susceptible
20 40 60 80
Overall Model Accuracy: 76% ± 0.02


A
Linear Discriminant
4
2
0
-2
-4
Field
Lab
Population Type
B
Data Density
0.4
0.3
0.2
0.1
0.0
Field
Lab
-10
-5
0
5
10
Linear Discriminant
C
Lab
Field
10
± 0.8
(3 - 17)
89
± 2.0
(71 - 100)
90
± 0.8
(83 - 97)
11
± 2.0
(0 - 29)
Field
Lab
25
50
75
Overall Model Accuracy: 89% ± 0.01
Lab
Field
14
± 1.8
(0 - 28)
80
± 2.4
(60 - 100)
86
± 1.8
(72 - 100)
20
± 2.4
(0 - 40)
Field
Lab
20
40
60
80
Overall Model Accuracy: 83% ± 0.01


A
Linear Discriminant
2.5
0.0
-2.5
Resistant Susceptible
Resistance Status
B
Data Density
0.5
0.4
0.3
0.2
0.1
0.0
Resistant
Susceptible
-10
-5
0
5
10
Linear Discriminant
C
Susceptible
Resistant
25
± 2.8
(0 - 50)
89
± 1.1
(78 - 97)
75
± 2.8
(50 - 100)
11
± 1.1
(3 - 22)
Resistant
Susceptible
20 40 60 80
Overall Model Accuracy: 85% ± 0.01
D
Susceptible
Resistant
27
± 3.3
(9 - 63)
79
± 2.3
(63-96)
73
± 3.3
(38 - 91)
21
± 2.3
(4 - 37)
Resistant
Susceptible
30 40 50 60 70
Overall Model Accuracy: 78% ± 0.02


A
Linear Discriminant
4
2
0
-2
-4
FR
LS
Population
B
Data Density
0.4
0.3
0.2
0.1
0.0
Field Resistant
Lab Susceptible
-10
-5
0
5
10
Linear Discriminant
C
Field Resistant
Lab Susceptible
11 ± 2.1 (0 - 35)
85 ± 2.8 (61 - 100)
89 ± 2.1 (65 - 100)
15 ± 2.8 (0 - 39)
Lab Susceptible
Field Resistant
20 40 60 80
Overall Model Accuracy: 87% ± 0.02
D
Field Resistant
Lab Susceptible
18 ± 3.1 (0 - 55)
84 ± 3.6 (43 - 100)
82 ± 3.1 (46 - 100)
16 ± 3.6 (0 - 57)
Lab Susceptible
Field Resistant
20 40 60 80
Overall Model Accuracy: 82% ± 0.02


A
Linear Discriminant
4
2
0
-2
-4
FR
FS
Population
B
Data Density
0.4
0.3
0.2
0.1
0.0
Field Resistant
Field Susceptible
-10
-5
0
5
10
Linear Discriminant
C
Field Resistant
Field Susceptible
11
± 1.5
(0 - 22)
89
± 2.4
(60 - 100)
89
± 1.5
(78 - 100)
11
± 2.4
(0 - 40)
Field Susceptible
Field Resistant
20
40
60
80
Overall Model Accuracy: 88% ± 0.01
D
Field Resistant
Field Susceptible
17
± 3.1
(0 - 50)
86
± 2.3
(70 - 100)
83
± 3.1
(50 - 100)
14
± 2.3
(0 - 30)
Field Susceptible
Field Resistant
20
40
60
80
Overall Model Accuracy: 84% ± 0.02