Landscape Ecol (2021) 36:45–67
https://doi.org/10.1007/s10980-020-01099-1

RESEARCH ARTICLE

Landscape factors and how they influence whitefly pests in cassava fields across East Africa

Sarina Macfadyen · Wee Tek Tay · Andrew D. Hulthen · Cate Paull · Andrew Kalyebi · Frances Jacomb · Hazel Parry · Peter Sseruwagi · Zuberi Seguni · Christopher Abu Omongo · Donald Kachigamba · Michael Otim · Nancy Schellhorn · Andrew Polaszek · John Colvin · Paul De Barro

Received: 14 October 2019 / Accepted: 18 August 2020 / Published online: 15 October 2020

Abstract

Context African production landscapes are diverse, with multiple cassava cultivars grown in small patches amongst a diversity of other crops. Studies on how diverse smallholder landscapes impact herbivore pest outbreak risk have not been carried out in sub-Saharan Africa.

Objectives *Bemisia tabaci* is a cryptic pest species complex that cause damage to cassava through feeding and vectoring plant-virus diseases and are known to reach very high densities in certain contexts. However, the factors driving this phenomenon are unclear.

Methods *Bemisia* density data in cassava across a large number of sites representing a geographic gradient across Uganda, Tanzania and Malawi were collected. We tested whether in-field or landscape factors associated with land-use patterns underpinned *Bemisia* density variability and parasitism.

Results We found the *B. tabaci* SSA1 species dominated our study sites, although other species were also common in some cassava fields. Factors associated with the surrounding landscape were unimportant for explaining variability in adult density, but the in-field variables of cassava age and cultivar

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10980-020-01099-1) contains supplementary material, which is available to authorized users.

S. Macfadyen (✉) · W. T. Tay · F. Jacomb
CSIRO, GPO Box 1700, Canberra, ACT 2601, Australia
e-mail: sarina.macfadyen@csiro.au

A. D. Hulthen · C. Paull · H. Parry · N. Schellhorn · P. De Barro
CSIRO, Brisbane, QLD 4001, Australia

A. Kalyebi · C. A. Omongo · M. Otim
National Crops Resources Research Institute, Kampala, Uganda

A. Kalyebi · P. Sseruwagi · Z. Seguni
Mikocheni Agricultural Research Institute, Dar es Salaam, Tanzania

D. Kachigamba
Department of Agricultural Research Services, Bvumbwe Research Station, Limbe, Malawi

N. Schellhorn
RapidAIM Pty Ltd, Brisbane, QLD 4101, Australia

A. Polaszek
Department of Life Sciences, Natural History Museum, London, UK

J. Colvin
Natural Resources Institute, University of Greenwich, Chatham Maritime, Kent, UK

were very important. The density of nymphs and the parasitism of nymphs was heavily influenced by a diversity of landscape factors surrounding the field, including the size of focal cassava field, and area of cassava in the landscape. However, unlike the trend from many other studies on drivers of natural enemy populations, this pattern was not solely related to the amount of non-crop vegetation, or the diversity of crops grown in the landscape.
Conclusions Our findings provide management options to reduce whitefly abundance, including describing the characteristics of landscapes with high parasitism. The choice of cassava cultivar by the farmer is critical to reduce whitefly outbreak risk at the landscape-scale.

Keywords Crop diversity · *Bemisia tabaci* · *Bemisia afer* · Habitat manipulation · Parasitoids · Natural enemies · Pest density

Introduction

Small-holder production landscapes are inherently diverse both spatially and temporally. Large scale stochastic processes influence the environmental variability, and therefore the resources (such as water) available to farmers each season. This, together with planned household food consumption and external market demands influence the area committed to growing specific crops by smallholder farmers in East Africa. At a field scale, individual farmers make decisions about what crops they plant where, when they plant, apply inputs (fertilisers and pesticides) and harvest, and how they manage the non-productive parts of their land. Farmers' management decisions directly influence biotic processes, such as the amount of suitable host-plant material available at any time for invertebrate herbivores. The change in the amount of host-plant resources across a landscape can influence population density within a field unit and population persistence from one season to the next. For widely dispersing pest species, landscape-scale effects can undermine or completely ameliorate field-scale management actions (Gurr et al. 2018). An understanding of landscape-scale effects on pest species are needed for the development of area-wide management (AWM) approaches. AWM has been advocated for reducing the risk of pest outbreaks, increasing the activity of natural enemies, and reducing the risk of insecticide resistance (Schellhorn et al. 2015). However, AWM can sometimes be difficult to achieve because it requires coordination across different landholders.

There have been many studies examining the relationship between the composition, configuration, and diversity of land-use in agricultural landscapes and pest outbreaks. Empirical studies testing these relationships come predominantly from intensively farmed landscapes (often consisting of monoculture crops), in high-income countries (Chaplin-Kramer et al. 2011; Karp et al. 2018). Some studies show a clear link between increased amounts of non-crop habitats and increased natural enemy diversity at the landscape-level, but subsequent pest suppression is not always observed (Chaplin-Kramer et al. 2011; Veres et al. 2013; Karp et al. 2018; Duarte et al. 2018). It remains unclear whether such patterns also exist in highly diverse smallholder production landscapes (Zhou et al. 2014). Using household surveys in smallholder systems in Nigeria, Zhang et al. (2018) reported crop pest severity was less where the proportions of forest and unused land at the landscape scale was higher. However, other agronomic and socio-economic factors were also important for explaining reported pest severity. In contrast, increased landscape diversity, through management actions to increase the area of grasslands, has been shown to increase the risk of Lepidoptera stemborer damage to maize and sorghum, by providing a new source habitat for these pests (Midega et al. 2014). Given the low numbers of studies we can't yet say if the drivers associated with lower populations of pests in intensively farmed, homogenous landscapes will differ substantially from the drivers in more diverse smallholder landscapes.

Cassava is a common crop included in smallholder production landscapes in East Africa. Whilst it is grown over large areas of land, it is not managed as a monoculture crop, and has a diversity of cultivars, soil preparation techniques, and harvesting options. Cassava is a long-duration crop (remaining in the ground for periods between 8 and 18 months), with relatively low inputs in terms of fertilisers and pesticides, and therefore represents an ideal environment for population growth and activity of natural enemies. Furthermore, in some parts of East Africa, cassava is planted

in two seasonal windows per year, resulting in temporal overlap between fields at different stages of maturity. Farmers plant a diversity of different cassava cultivars with different traits, including tolerance or resistance to two major diseases; cassava mosaic disease (CMD) caused by a multiple Begomoviruses, and Cassava Brown Streak Disease (CBSD) caused by two RNA Ipomoviruses (Legg and Thresh 2000; Maruthi et al. 2002; Alicai et al. 2007). Both diseases are vectored by whitefly pests in the *Bemisia tabaci* complex, consisting of more than 36 different but morphologically similar species, while CBSD is potentially also vectored by other *Bemisia* species (Maruthi et al. 2002; 2005; Legg et al. 2011, 2014a, b; Ateka et al. 2017). However, these diseases can also be spread between fields in a region via the transfer of infected cuttings from one farmer to the next. The realization in recent year that *B. tabaci* is a pest species complex (Maruthi et al. 2004; Mugerwa et al. 2018; Elfekih et al. 2018; Vyskočilová et al. 2018; Mugerwa 2019; Kanakala and Ghanim 2019; Kunz et al. 2019) has required a re-assessment of the virus-vector relationships, and research on the phylogenetic position of species sampled from East Africa, using molecular approaches (Alicai et al. 2016; Boykin et al. 2018). The combination of disease impacts and damage caused by high *B. tabaci* populations in cassava production regions of East Africa is one of the key constraints on the productivity of smallholder farmers. Cassava breeding efforts have produced several new cultivars that are resistant or tolerant to either CMD or CBSD (Adriko et al. 2011; Katono et al. 2015). However, improved cultivars are usually planted alongside un-improved local landraces in a farming landscape as these cultivars have other traits that are desirable for farmers. Furthermore, a diversity of other crops are grown in small fields and gardens that may also act as host plants for *B. tabaci*. This mosaic of host plant resources may potentially influence the behaviour (Kalyebi et al. 2018) and landscape-wide population abundance of *B. tabaci* (Parry et al. 2020).

Until recently, research on and control of agricultural pests, such as species in the *B. tabaci* complex, was focussed at the individual field level with little reference to surrounding land-use (although see Kristensen et al. (2013) and Bianchi et al. (2015) for models of a whitefly parasitoids at multiple scales). These field-scale studies together with discrete behavioural experiments provide information about an insect's biology, by way of life-table analysis. Although laboratory and field-scale research has revealed valuable information about *B. tabaci,* they do not reflect other important aspects such as the spatial variation in the density of these pests in more complex field environments. Controlled studies have shown that *B. tabaci* populations respond differently to different cassava cultivars (Ariyo et al. 2005; Omongo et al. 2012; Katono et al. 2015; Kalyebi et al. 2018). Environmental conditions like temperature and rainfall also impact population dynamics, either directly by causing mortality of immature stages, or indirectly through altering populations of natural enemies (see Macfadyen et al. 2018 for a review). Therefore, we can observe very different levels of *B. tabaci* abundance across a landscape between regions and seasons (Kriticos et al. 2020). In this study, we aim to understand how the in-field and land-use factors in the surrounding landscape influence the observed variability in the density of this pest in cassava fields. Ultimately, we can use this knowledge to develop novel landscape management approaches that can be implemented by smallholder farmers and adopted at a relatively low cost. To achieve this, we collected data on *Bemisia* abundance from smallholder cassava farms, at a landscape-scale, at sites in three East African countries; Uganda, Tanzania, and Malawi. This large geographic gradient allowed us to examine *Bemisia* whitefly populations across a diversity of different cassava production contexts. Given that the species in the *B. tabaci* pest complex cannot be identified in the field using morphology alone, we developed a high-throughput sequencing method to identify the common species present in our focal fields (Tay et al. 2020). Using this data set we asked:

1. What are the common species of whitefly found on cassava and on nearby crops and non-crop host plants?
2. Which landscape or in-field factors influence the density of *Bemisia* adults and nymphs in cassava fields?
3. Which landscape or in-field factors influence the parasitism of *Bemisia* nymphs in cassava fields?

Methods

Cassava fields were selected in at least three different geographical regions in Uganda (Lira called "UG4", Kamuli "UG1" and "UG3", Kyotera "UG2" and "UG5") and Tanzania (Mwanza called "TZ1" and "TZ5", Dodoma "TZ4", and Dar es Salaam "TZ2" and "TZ3"), and two regions in Malawi (Northern called "ML1" and Central "ML2", Fig. 1). Kamuli, Kyotera, Dar es Salaam and Mwanza were sampled at two different periods (referred to as 12 different regions throughout for ease) (Fig. 1, Online Appendix 1, Table A1.1). These regions were chosen to represent different agro-ecological zones that grow significant amounts of cassava. The exception is TZ4, in the Dodoma region, where cassava production has only recently started. In total, we conducted three data collection trips across two years (trip 1: 1/8/2015—26/08/2015, trip 2: 5/4/2016—25/4/2016, trip 3: 29/10/2016—9/11/2016) but complete regions were sampled at each trip (i.e. all fields in a region). Each of these regions differ in their altitudes, temperature and rainfall profiles, and the agricultural systems used by farmers. For example, in Kamuli (average altitude of 1090 m above sea level, msl) and Kyotera (average altitude of 1210 msl) in Uganda, there are generally two planting windows for cassava per year, but in northern Malawi (average altitude of 547 msl), there is one (Online Appendix 1). Before sampling, regions were visited by extension officers and contact was made with local farmers to obtain their permission to sample in their fields.

In each of these 12 regions we searched for up to ten focal cassava fields to sample and each focal field was sampled once. Focal cassava fields selected needed to have: cassava of between three to seven months after planting; cassava as the dominant crop (although it could have an intercrop of another crop type if cassava was still dominant); cassava cultivars that could be identified and that were consistently planted across the field; at least 30 plants of the same cultivar to survey. Each focal field had to be at least 4 km (straight line distance) from the next nearest focal field. Prior to sampling farmers were interviewed to confirm information about the cultivars of cassava and other crops that they were growing and associated management activities, i.e. crop rotation and fungicide and pesticide use (Online Appendix 2). A field deemed suitable for sampling was assigned a focal field code and a full sampling protocol conducted. All data were recorded on predesigned and trialled electronic forms that were constructed using Open Data Kit (ODK) software (Hartung et al. 2010). ODK Software was run using Android tablets, and field collection identifiers related to unique barcodes which allowed all information and samples to be referenced back to the field and individual plant. Data was collected offline and uploaded to secure cloud servers after the sampling was completed.

There were 36 different cassava cultivars recorded from the focal fields, some of which were unique to a certain region (although they may have been dominant within the region). Each cassava cultivar that was surveyed was categorized into one of four groups; susceptible to both CMD and CBSD, tolerant to CMD (CBSD could be susceptible or resistant), tolerant to both CMD and CBSD, resistant to CMD (CBSD could be susceptible or resistant), based on the knowledge of scientists involved in the project. Four cultivars had an unknown disease rating and were categorized as susceptible. Four cultivars had an unknown disease rating and were categorized as susceptible.

Mapping the landscapes

Land-use information was captured within a minimum radius of 100 m from the centroid of the focal cassava field. Given that all the mapping was completed manually, this minimum distance was chosen as it captured a large amount of the land-use diversity surrounding the focal field, and at the same time could be completed in a reasonable time-period. Field boundaries and landscape features were digitised on Android tablets using offline maps and satellite image base layers, authored in ArcGIS Collector (ESRI 2015) and resulting spatial layers checked and cleaned in ArcGIS desktop (ESRI 2010). This information was confirmed by walking field boundaries and ground truthing the details which were then digitized and used to produce maps. Different land uses and features including roads, buildings and different crop and host plant, and non-host plant categories were then added. We focussed on the dominate crop and vegetation types that could be mapped with some accuracy and were present across enough fields to enable statistical analysis. Simple landscape metrics were generated,

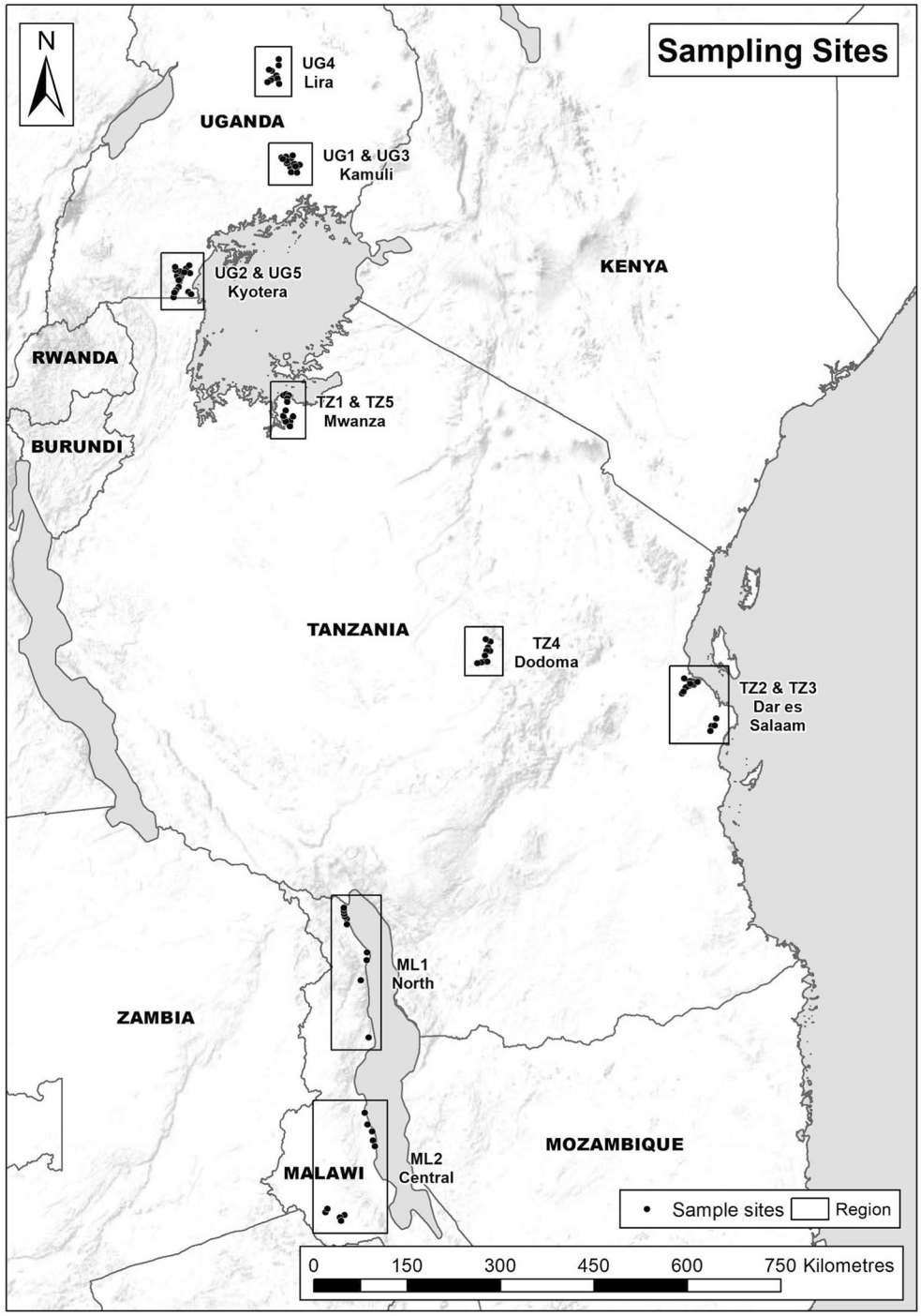

Fig. 1 Focal cassava field sampling sites across 12 regions. At each site, a focal cassava field was sampled and the landscape around the focal field mapped. See Online Appendix 1 for examples of focal field maps

such as percentage cover of land-use types around the focal field, size of the focal field, amount of cassava in the landscape, diversity and number of other crops, the amount of non-crop vegetation (i.e. grassland or wooded areas, example maps of three sites are provided in Online Appendix 1).

Counting adult and nymph *Bemisia* species, and collecting parasitoids

It is not possible to identify the different species within the *B. tabaci* pest complex morphologically in the field, so we collected samples of *Bemisia* species adults and nymphs for later molecular-based identification. Throughout the methods we refer to "whitefly" to encompass *Bemisia* whitefly species later clarified using molecular identification. To sample for whitefly adults and nymphs 30 cassava plants were randomly chosen and sampled in each focal field. The top five fully expanded leaves of each cassava plant were carefully turned over, (to minimise disturbing the adults) and the numbers of adult whitefly were counted and entered as a range. During analysis each range was given a number (0 = 0, 1–9 = 1, 10–50 = 2, 51–100 = 3, 101–200 = 4, > 200 = 5), hereafter ferred to as density categories. Similarly using a five point scoring (1–5 with 1 being no symptoms and five the most severe) each plant was scored for CBSD symptoms, CMD symptoms, cassava green mite (*Mononychellus tanajoa*) infestation symptoms, and sooty mould severity on the upper surfaces of the lower leaves (Sseruwagi et al. 2004). We attempted to use this data to determine whether the density of adult or nymph of *Bemisia* species influenced CMD and CBSD symptom expression in cassava fields. Previous studies have sought to link high abundance of whitefly with greater incidence or severity of cassava diseases at the field and pandemics at the regional scale (e.g., Maruthi et al. 2005; Rwegasira et al. 2011) usually with limited success. In our study we also found no clear relationship between CMD and CBSD symptoms and *B. tabaci* density in the focal fields (either for adults or nymphs). We show the full results in Online Appendix 4 but will not discuss them further in this study.

Further down each cassava plant stem, three leaves per plant were turned over and visually inspected for third to fourth instar nymphs. Cassava leaves are palmate and divided into lobes. If there were third-fourth instar nymphs anywhere on the central lobe of any of the three leaves a 10 cm diameter metal cutter was placed over the leaf and a leaf disc removed. Leaf discs (maximum three per plant) were then placed sequentially in a partitioned leaf disc holder, a modified clipboard (See images in Online Appendix 1). Each disc was labelled with a barcode, which related to the field identifiers. A digital image of each barcoded leaf disc was taken in the field. Images were examined later and the number of live third-fourth instar nymphs, and the number of empty pupal cases (where the adult had emerged) were recorded. We were unable to determine if a nymph had been parasitised from the photographs, so the nymph density count includes both parasitised and unparasitised live third to forth instar nymphs. The rearing data were used to estimate parasitism rate, once the adult parasitoids and whitefly had emerged.

For the adult whitefly data, the density categories were summed across the five leaves assessed per plant. We then calculated the mean from the 30 plants sampled in each focal field. We refer to this as adult density throughout. For nymph density, we used the count of the number of live nymphs per leaf disc. Before analysis, we examined the relationship between nymph count and actual leaf area included in the disc (as not all cassava lobes were wide enough to occupy the whole disc area). However, no obvious relationship was found, therefore we ignored leaf disc area, and used live nymph numbers. We summed the number found on each of the three discs, then took the mean across the 30 plants in the focal field.

Cassava discs cut from each of the focal field plants were combined (max three per plant) and placed into small black emergence containers (see Online Appendix 1 for images and details of the materials used to construct emergence containers). A barcode was stuck to the 5 ml clear screw-top vial on top of the emergence container. When the whitefly adults or parasitoids emerged, attracted by the light transmitted through the clear plastic vial they moved up into the larger 5 ml vial. The emergence containers with leaf discs were kept at room temperature out of direct sunlight and monitored while the whitefly adults and parasitoids emerged. After a minimum of 14 days the containers were checked for a final time. Ethanol was added to the vial to preserve the whitefly and parasitoids. Given the high numbers of parasitoids that emerged in the first sampling trips (over 80% parasitism in some fields in Uganda, and a maximum of 923 live rearings from one field), we had to reduce the number of leaf discs collected in the final trip. In two regions, only three focal fields (out of ten) were selected for rearing nymphs to obtain parasitism data (UG3 and UG5). To calculate the rate of parasitism per focal field, we used the numbers of adult whitefly and

parasitoids emerging from the emergence containers. We summed the data from all containers in each focal field. We divided the parasitoid adults by the sum of the live emergences (parasitoid adults plus whitefly adults) giving a proportion between zero to one. Eight fields with zero counts were removed before analysis: firstly, fields where no leaf discs were collected due to low/absent nymph numbers, and secondly fields with few leaf discs collected (due to low nymph density) and therefore no live emergences. Fields with leaf discs collected and live emergences recorded, but no parasitoids emerged were considered true zeros and retained in the data set.

The other crops beyond cassava and areas of non-crop vegetation were mapped and given a field code. A timed search method was used to quickly assess these fields for whitefly adults or nymphs. For 15 min plants were examined for the presence or absence of whitefly nymphs or adults. The search protocol was modified slightly in response to the different growth habits of crops. The presence and absence of whitefly adults/nymphs and the total number of plants searched were recorded. Timed searches of non-crop areas involved timed searches of any plants in the areas for whitefly adults and nymphs. When the search was completed, the name and/or description of the plant on which adults and nymphs were found was recorded. If the name of the plant could not be confidently confirmed a specimen was barcoded and pressed for identification later. If whitefly nymphs were found on the plant a sample was collected for genotyping. Genotype samples were barcoded and cross-referenced to the field code and preserved as above.

Molecular identification of whitefly nymphs on cassava

In addition to the density data, a sample of nymphs was collected for genotyping using a flat-based 5 ml plastic vial without a cap. Using the edge of the vial small $\sim$ 7 mm diameter leaf discs complete with nymphs were cut from the leaves. The leaf discs were placed into the vial (the same vial used to cut the leaf disc) and filled with $>$ 95% ethanol and barcoded and sealed with a screw-top lid. A maximum of three vials were collected per focal field. Adults collected using aspirators were also added to these vials. We targeted nymph specimens for genotyping because we could be certain of the host plant relationship as per Sseruwagi et al. (2006). In some cases, the field did not meet all our criteria to be considered a focal field (cassava too young or old, mixed varieties, too few plants). However, we still collected a sample of whitefly for genotyping and basic information on the characteristics of the cassava field (size of field, GPS location, etc.) ("genotype only" fields in Online Appendix 3).

DNA extraction and sequencing library preparation followed the methods described in Tay et al. (2020), based upon the Illumina 16S Metagenomic Sequencing Library Preparation (Part # 15044223 Rev. B) protocol. For the full molecular protocol see Tay et al. (2020). Briefly, nymphs were dislodged from leaves and visually sorted into parasitised and unparasitised groups under a dissecting microscope (by looking for evidence of developing hymenopteran larvae in the nymphal case). Samples of 20 or 40 unparasitised individuals per vial (representing populations from each field) were randomly selected for DNA extraction. Genomic DNA (gDNA) was extracted from each population using the Qiagen DNeasy Blood and Tissue Kit (cat. #69506), including the optional RNase A treatment. Each sample of gDNA was quantified using a Qubit Fluorometer and dsDNA HS Assay Kit and samples were standardised to 0.5 ng/µl in preparation for PCR amplification.

Two sets of primers (wfly-PCR-F1/R1 and wfly-PCR-F2/R2; Tay et al. 2020) were used to amplify the target mtCO1 region. PCR product was visualised on an agarose gel to determine success, before being cleaned and purified using Agencourt AMPure XP beads (cat. #A63882). Purified PCR product was quantified using Qubit and standardised to 0.5 ng/µl, before Index PCR to construct amplicon libraries. Indexed amplicons were then quantified and samples were combined in equal amounts, to create an F1/R1 pool and an F2/R2 pool. The pooled samples were run on a gel, the expected fragment size excised, and the amplicons cleaned and purified using the Zymoclean Gel DNA Recovery Kit (cat. # D4007). Quality and quantity of the amplicon libraries was ascertained using Qubit and Agilent Technologies TapeStation. Purified indexed amplicon pools were diluted to 4 nM and the sequencing run was performed with MiSeq Reagent Kit V3 (600 cycles).

High throughput amplicon sequence data from populations were analysed using Geneious 11.1.5 (Biomatters Ltd, Auckland, New Zealand) based on the workflow pipeline as described in Tay et al. (2020)

to determine the proportions of amplicons from each population that belonged to known and unknown *Bemisia* species. We ascertained *Bemisia* species identity based on characterised and NUMT-free/pseudogene-free partial mtCOI genes as described by Kunz et al. (2019). The species nucleotide boundary was set at 3% in the first instance (Vyskočilová et al. 2018; Kunz et al. 2019). Unmatched sequences were relaxed to 5% to allow for PCR-introduced nucleotide polymorphisms to be mapped to the reference sequences. We mapped the amplicon sequences to a total of 18 African *Bemisia* species, and included two putative novel species (Tay et al. 2020 GenBank accession numbers MN646951 and MN646952). Parasitoid, bacterial, fungal, and potential NUMTs/pseudogene sequences were assembled and identified by de novo assembly (for parameter see Tay et al. 2020).

Data analysis

All the data manipulation, graphing and statistical analysis were completed using R and associated packages (RStudio version 1.2.5033, R version 3.5.2, R core team 2018). In total, we had 101 focal fields in the three to seven months after planting age range with adult and nymph density data, and 79 fields with parasitism rate data that were included in the analysis. For all the response variables we checked for spatial autocorrelation between focal field sites by plotting spline correlograms (using spline.correlog in ncf package, Bjornstad 2019), however, we found no evidence of strong spatial autocorrelation. The explanatory variables were grouped into three categories: in-field factors, landscape factors relating to cropping components of the landscape, and landscape factors relating to the non-crop components of the landscape (Online Appendix 1). Before developing models, we tested for collinearity between pairwise combinations of the explanatory variables using Spearman correlation coefficient (magnitudes $\pm$ 0.5 were considered problematic, Online Appendix 1) and variance inflation factors (in the car package, Fox and Weisberg 2019, VIF $>$ 10 were considered problematic). We removed the landscape suitability score (LSS, description in Online Appendix 1), area of "other crops" in the landscape (a catch-all for rare and infrequently recorded crop types) as these were highly correlated with other variables. The area of beans was removed as this land-use type was not consistently recorded across regions. We standardized the continuous explanatory variables by subtracting the mean and dividing by the standard deviation.

We used a model selection approach to identify the most parsimonious models for each response variable (denisyt of adults, nymphs, and parasitism rates in the focal fields). Each model was constructed with and without groups of explanatory variables from each of the three categories and examined the change in AICc and the weight of each model ("model.sel" function in R package MuMIn, Barton 2018). To assist in this process, we simplified the months after planting variable, by making continuous in all models (rather than categorical). In all models the variable "region" was included as a random factor. For the adult and nymph models, a generalized linear mixed model with a negative binomial distribution was used. For the final model we calculated the pseudo-R-squared value using the "r.squaredGLMM" function. We include both the marginal R-squared (R^2m) which includes fixed effects and the conditional R-squared (R^2c) which includes fixed and random effects, both using the delta method. We assessed the model adequacy of the final model using residual plots, QQ plots, and checking for over dispersion and zero inflation (using functions in the package DHARMa, Hartig 2019). For the parasitism rate in the final model we plotted the predictions from the model with a few of the explanatory variables. We used "predictInterval" functions from the merTools package (Knowles and Frederick 2019) to produce a fitted line and the 95% confidence intervals.

For parasitism rate we used a binomial model (glmer function in the R package lme4, Bates et al. 2015) as the response variable was a proportion (0–1) and added a weight associated with the number of live emergences per field. Given that there were less fields (N = 79) in this data set we struggled to get a full model with all explanatory variables (i.e. the ratio of observations to explanatory variables was getting too small). This model failed to converge and was removed from the analysis.

Results

Our detailed mapping data quantified how each region differs in the mean area covered by crop and non-crop

vegetation around focal cassava fields (Online Appendix 1, Fig. A1.5). From the Lira sites in the north of Uganda (UG4) having 55% of the landscape covered by crops, and 31% by uncropped forest, grass or shrubland, through to sites in southern Malawi (ML2) having 24% crops and 66% forest, grass or shrubland (Online Appendix 1, Fig. A1.5). The amount and types of crops planted exhibited significant changes across regions, with the amount of cassava on average at 20% of the landscape but varying from 1 to 84%. Sites in Uganda and the Lake Zone of Tanzania (TZ1 and TZ5) had significant amounts of maize and sweet potato fields surrounding each focal cassava field (Online Appendix 1, Figs. A1.4 and A1.5). The field sizes in East Africa are generally very small, with the cassava fields chosen as focal fields being slightly larger than a usual cassava field (as we need at least 30 plants of a known cultivar). The cassava focal field sizes were largest in Uganda at 4053 m^2 ($\pm$ 1322), medium in Malawi at 2476 m^2 ($\pm$ 583), and the smallest in Tanzania at 1754 m^2 ($\pm$ 236). Change in the average number of crops grown in the landscape differed between regions, with Malawi central region having low diversity (ML2 mean = 1.2 crops, N = 5, s.e. = 0.2), and the Lira region in northern Uganda having the highest crop diversity (UG4 mean = 8.9, N = 10, s.e. = 0.6). Overall, insecticide use was very low across all the cassava regions surveyed, with about 15% of farmers reporting that they used insecticides on their crops (Online Appendix 2). There were only two farmers who confirmed insecticide use in cassava, but the pest being targeted was not always whitefly.

What are the common species of whitefly found on cassava and on nearby crops and non-crop host plants?

A large number of individual whiteflies were processed as part of the sequencing runs, with 149 samples processed in total (Table 1). Twelve of these samples consisted of adults only, and two samples had adults and nymphs combined (to increase the number of individuals per sample). The majority of the whitefly collected came from cassava host plants, but a few samples from other host plants also contained enough individuals to warrant inclusion (see details in methods). Overall most regions had on average less than one individual whitefly with unknown sequences per sample (exceptions being UG4, mean = 2.09, and TZ1 mean = 1.2 individuals). A complete list of the molecular identification results can be found in Online Appendix 3.

Of the 18 whitefly species used in our reference library, not all were detected in our samples. Notably the *B. tabaci* SSA3, *B. tabaci* MEAM1 and *B. tabaci* MEAM2 sensu Mugerwa et al. (2018) (*c.f* Tay et al. 2017) were not detected. Of the SSA species, *B. tabaci* SSA1 was the most commonly detected in the cassava fields we sampled across all regions (Fig. 2). *B. tabaci* SSA2 was present but only in a few Ugandan cassava fields (UG4 and UG1). On cassava, both *B. tabaci* Indian Ocean (IO) (TZ4) and *B. afer* (ML1, ML2) were relatively common at certain sites (Fig. 2). Some studies suggest that *B. tabaci* IO does not use cassava as a host plant (Misaka et al. 2020), however nymphs of this species were recorded on cassava and other plants throughout our study. *B. tabaci* SSA1 was also recorded on other crop and non-crop plants (e.g., wild groundnut, sweet potato), but often not in the frequency with which it was recorded on cassava (Fig. 3). Species such as *B. tabaci* MED and *B. tabaci* UGANDA1 were recorded only on non-cassava host plants (cowpea, sweet potato, and a rosella-like leafy vegetable). Sanger sequencing of individual specimens confirmed the sub-sampling of the HTS amplicon results.

Parasitoid DNA was consistently detected across our samples, demonstrating the utility of this method and associated primer pairs (Tay et al. 2020) for detecting species interactions. Despite deliberately selecting nymphs with no visual signs of parasitoid development, our detection of parasitoid partial mtCOI gene indicated that some nymphs were in the early stages of parasitism. Parasitoid DNA was absent in the samples that contained adults only, further confirming the robustness of this HTS method developed for this study. The proportion of parasitoid sequences seen in samples was highest in the Ugandan regions, and negligibly low in TZ4 and Malawi samples. Parasitoid sequences could be related to both the *Eretmocerus* and *Encarsia* genera that commonly parasitise the *B. tabaci* complex in East Africa (Polaszek et al. 1992; Otim et al. 2005).

Table 1 Summary of the molecular data collected on the identity of *Bemisia* whitefly species collected from cassava, other crops, and non-crop host plants

Country	Region	Number of samples processed	Number of fields or non-crop areas	Total number of whitefly sequenced (adults and nymphs)	Total number of reads obtained
Uganda	UG4	19	18	620	167,065
Uganda	UG1	16	10	233	287,717
Uganda	UG3	9	8	220	429,589
Uganda	UG2	10	10	400	406,131
Uganda	UG5	10	10	380	494,024
Tanzania	TZ1	11	8	124	187,086
Tanzania	TZ5	8	7	240	378,108
Tanzania	TZ4	12	11	245	558,093
Tanzania	TZ2	13	10	193	279,502
Tanzania	TZ3	17	16	440	321,660
Malawi	ML1	9	8	301	364,853
Malawi	ML2	15	10	383	740,583

A total of 149 samples analysed. The regions run from north UG4 Lira in Uganda, to south ML2 southern Malawi (see map in Fig. 1)

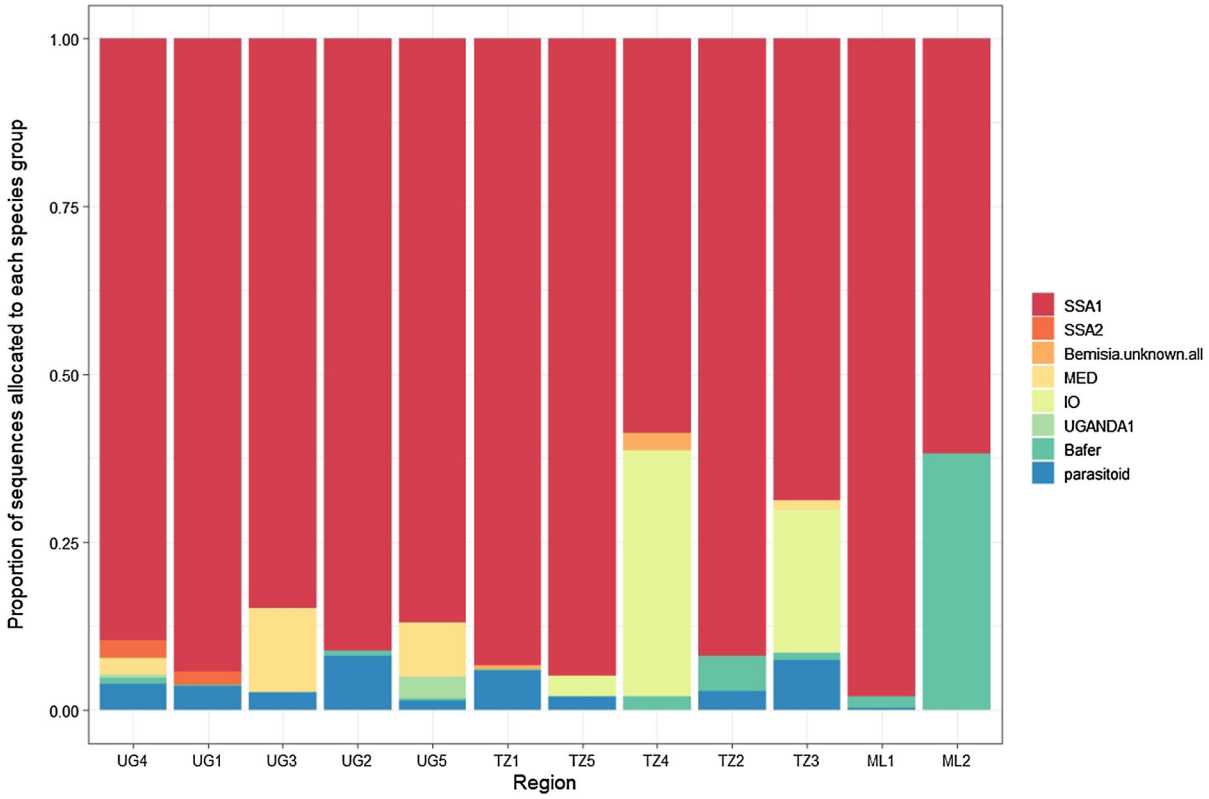

Fig. 2 The proportion of sequences allocated to *Bemisia* species from a reference library from samples collected from cassava, other crops and non-crop host plants across East Africa. In total 149 samples were assessed across these regions (Table 1)

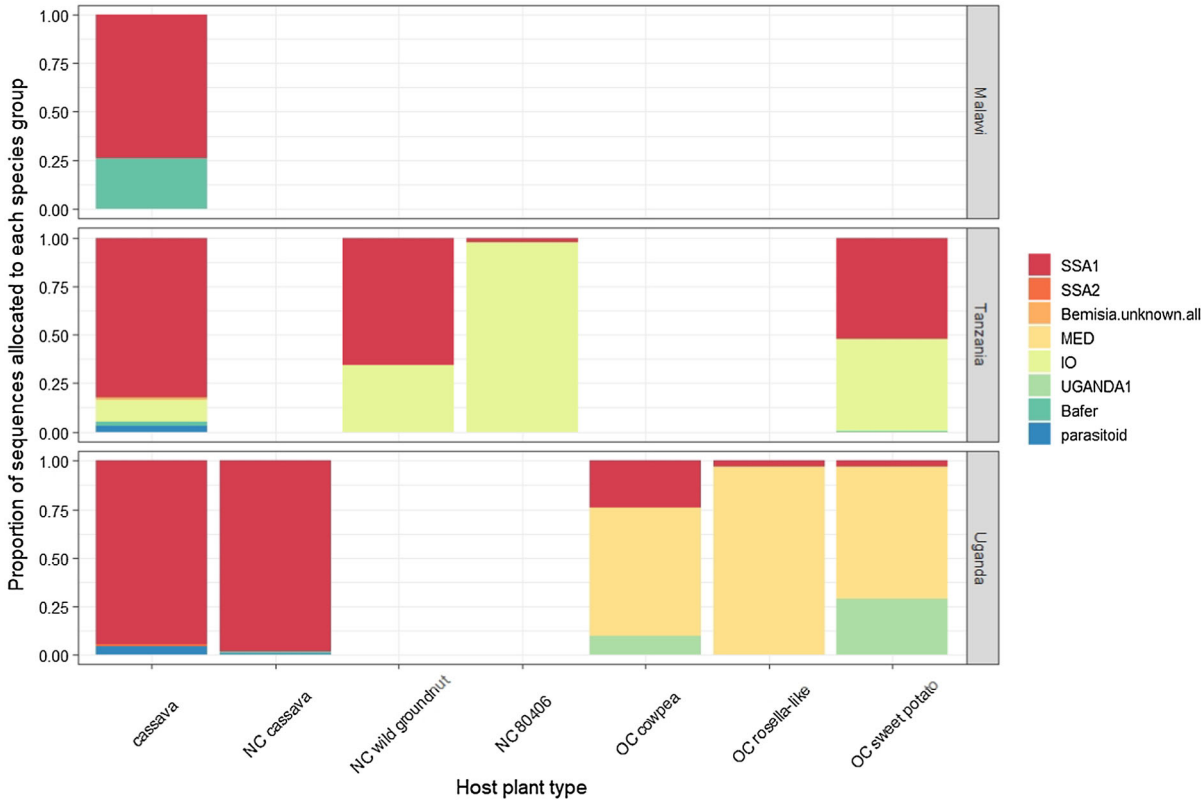

Fig. 3 The proportion of sequences allocated to *Bemisia* species from a reference library based on the host plants they were collected from. Most samples came from cassava host plants, either in cassava fields or in non-crop (NC) areas that contained cassava. Other crops (OC) were also processed if they had high numbers of whitefly nymphs. In total 149 samples were assessed (Table 1). NC80406 was a single cassava plant in a non-crop area adjacent to a sunflower field, that had a very high density of nymphs (*B. tabaci* IO)

Which landscape or in-field factors influence the density of *Bemisia* adult and nymphs in cassava fields?

The densities of adults and nymphs in each of the focal cassava fields differed across the regions (Fig. 4). The highest adult densities were observed in cassava fields in the Kyotera region of Uganda (UG2, n = 10, mean = 4.78, s.e. = 0.63), the Kamuli region of Uganda (UG1, n = 10, mean = 3.71, s.e. = 0.54), and the Coastal region near Dar es Salaam in Tanzania (TZ3, n = 8, mean = 4.71, s.e. = 0.64). In contrast, the highest nymph densities were observed in cassava fields in the Kyotera region of Uganda (UG2, n = 10, mean = 6.79, s.e. = 0.1.69 and UG5, n = 4, mean = 5.55, s.e. = 3.66). The final model for adult density included only in-field predictor variables, with months after planting and cultivar category showing the strongest patterns (Table 2, Fig. 5a, Online Appendix 1). The conditional R-squared value for the final model was 0.65. There were significantly lower numbers of adults at 5, 6, and 7 months after planting, compared to early in the season (3 and 4 months after planting). For the cultivar categories there were lower adult density on the susceptible cultivars compared to those cultivars that were tolerant to both diseases (Fig. 5a). Landscape factors outside the field appeared unimportant for predicting adult density (Table 2).

For nymph density, landscape factors outside the field were important (Table 2). The two best performing models contained firstly only in-field factors (months after planting, cultivar category, and intercrop, conditional R-squared of 0.57); and secondly in-field factors along with landscape factors associated with crops (size of focal cassava field, and area of cassava in the landscape, conditional R-squared 0.69).

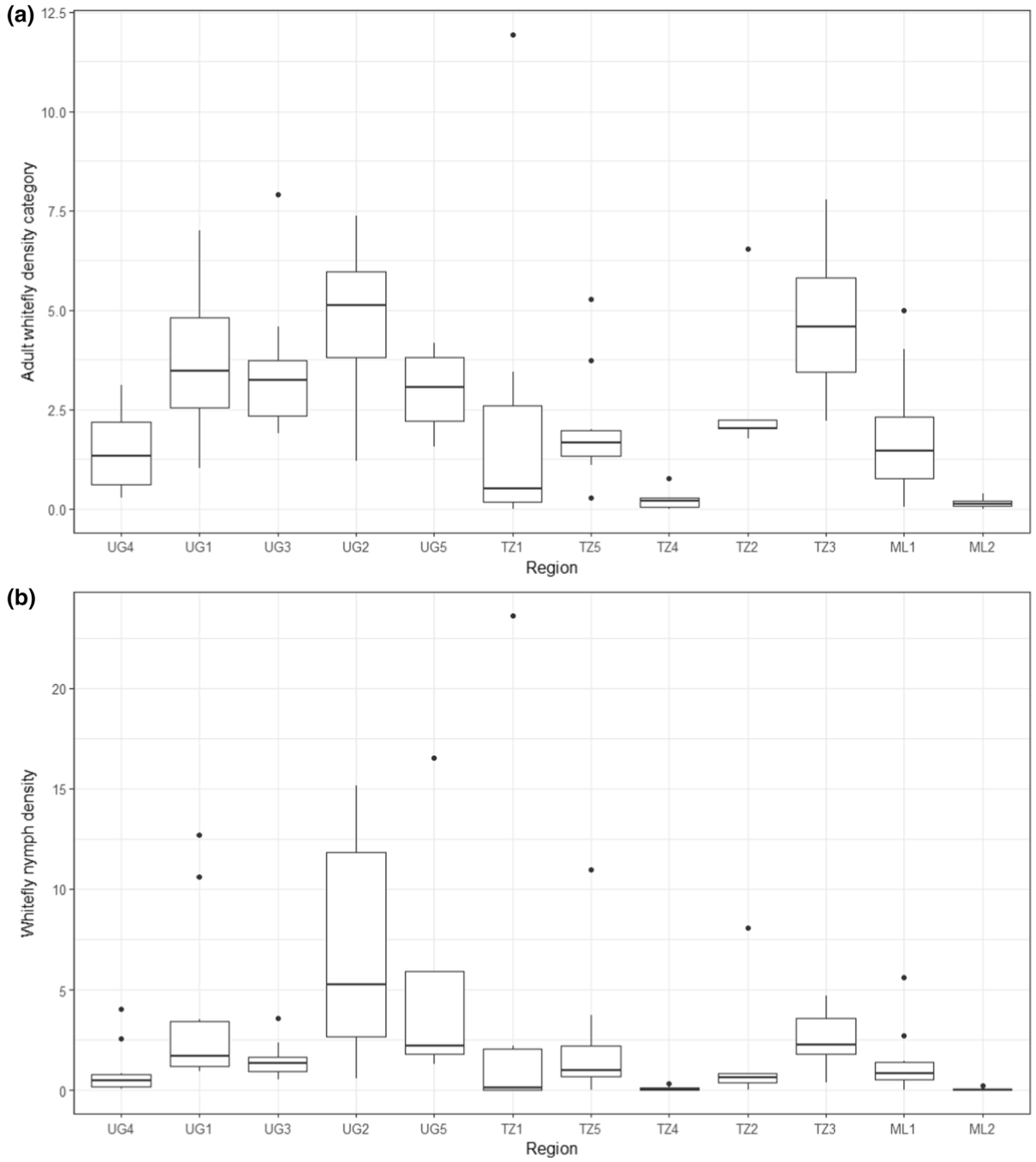

Fig. 4 Density of *Bemisia* adults (**a**) and nymphs (**b**) in cassava fields across 12 regions in East Africa (running from north UG4 Lira in Uganda, to south ML2 southern Malawi). The line in the boxplot shows the median values, the box boundaries the upper and lower quartiles, and the whiskers the highest and lowest values excluding outliers

Landscape factors associated with non-crop components of the landscape appeared unimportant for nymph density in the focal fields. In the final nymph model, the factors months after planting (Fig. 5b), cultivar category (Fig. 5b), cassava area in the landscape, and area of the focal field were all significant (Table 2, Online Appendix 1). As for adults, there were significantly lower numbers of nymphs at 5, 6,

Table 2 The importance of in-field and landscape factors for predicting *Bemisia* adult and nymph density and parasitism rate

ADULTS		a2#	a6	a5	a3	a7	a4	a1
In-field	months after planting*	***						
	var.cat	***						
	intercrop							
Landscape, crops	area							
	Cassava							
	no.crops							
	Sweet pot							
	Soybean							
	Pumpkin							
Landscape, non-crop	non-crop							
	Banana							
	Maize							
	AICc	**375.1**	380.2	385.2	398.6	405.0	396.8	391.6
	R^2m	0.164						
	R^2c	0.646						
NYMPHS		n2#	n6	n5#	n3	n7	n4	n1
In-field	months after planting*	***		**				
	var.cat	**		**				
	intercrop							
Landscape, crops	area			**				
	Cassava			**				
	no.crops							
	Sweet pot							
	Soybean							
	Pumpkin							
Landscape, non-crop	non-crop							
	Banana							
	Maize							
	AICc	**396.2**	397.8	**396.2**	401.9	408.5	409.8	401.5
	R^2m	0.176		0.297				
	R^2c	0.567		0.700				
PARASITOIDS		p2	p6	p5	p3	p7#	p4	p1
In-field	months after planting*							NA
	var.cat							
	intercrop							
Landscape, crops	area							
	Cassava					***		
	no.crops					***		
	Sweet pot					***		
	Soybean					***		
	Pumpkin					***		
Landscape, non-crop	non-crop					***		
	Banana					***		
	Maize					*		
	AICc	1126.7	1124.0	1044.8	1197.0	**1036.0**	1264.6	
	R^2m					0.117		
	R^2c					0.262		

Table 2 continued

The shaded boxes indicate the explanatory variables included in each model. For the adults and nymph data a negative binomial GLMM was used (N = 101 fields). For parasitism rate a binomial GLMM was used with proportion parasitized (0–1) as response and weights added as the live nymphs in each sample (N = 79 fields). The term "region" was included as a random effect in all models. The bold AICc show the model(s) with the lowest AICc values. Significance codes for final model are 0 '***', 0.001 '**', 0.01 '*', 0.05. (full model outputs in Online Appendix 1)

NA—model not assessed in this case

*Set as continuous in these models to help with model simplification

#Final model

and 7 months after planting, compared to early in the season (3 and 4 months after planting), and lower nymph density on the susceptible cultivars compared to cultivars that were tolerant to diseases (Fig. 5b). Interestingly, nymph density was higher in larger fields but was significantly lower in fields surrounded by a higher proportion of cassava (Online Appendix 1). Note that the area of the focal field and the amount of cassava in the landscape are related to each other although not strongly (Spearman correlation of 0.44, Online Appendix 1).

Which landscape or in-field factors influence the parasitism of *Bemisia* nymphs in cassava fields?

The parasitism rate of nymphs varied greatly between regions with the highest average of 83% parasitism in the Kamuli region of Uganda (UG1, n = 10, s.e. = 5.24), and 76% in the coastal region of Tanzania (TZ3, n = 8, s.e. = 5.64) (Fig. 6). There was a weak asymptotic relationship between nymph numbers in the focal field and parasitism rate (Online Appendix 1) suggesting that it is not only the density of nymphs that leads to high or low parasitism rate.

There are likely to be greater than 10 species of parasitoids that attack *B. tabaci* on cassava in East Africa (Guastella et al. 2015; Macfadyen et al. 2018). Given the numbers of parasitoids reared in this study we have identified and categorised them into three groups; *Encarsia* genera, *Eretmocerus* genera, and other parasitoids not easily grouped into these two genera. When we examined the live adult parasitoids, there was a greater proportion from the *Eretmocerus* genera in the Ugandan regions (Online Appendix 1), relative to *Encarsia* genera.

Landscape factors outside the field were more important in determining the variability in parasitism rate between the focal fields (Table 2) than in-field factors. Overall, in-field factors, such as cassava cultivar had no influence on the parasitism rate in the focal fields. The most parsimonious model contained all groups of crop and non-crop landscape factors, however, the model overall did not explain a lot of the variation in parasitism rate with a conditional R-squared of 0.26 (Table 2). The amount of cassava in the landscape was significant in the final model, but not the area of the focal field. The area of other crops (sweet potato, pumpkin, banana) displayed negative coefficients (Online Appendix 1), but this was sometimes a complex relationship with parasitism rate. For example, the highest parasitism rates were observed in fields with $< 20\%$ cassava in the landscape (with a peak at about 10%), and then decreased rates with greater amounts of cassava (Fig. 7). There was a decrease in parasitism rate with an increasing cover of non-crop land-use until 40–50% of the landscape, then no impact after that point (Fig. 7). The number of crops in the landscape (crop diversity) showed a positive coefficient in the final model, however again this pattern was complex. The parasitism rate increased from zero to four crops in the landscape, peaked at five crops, and decreased with higher diversity of crops in the landscape (Fig. 7).

Discussion

Cassava and other crop cultivation practices in East African smallholder production landscapes are incredibly diverse across space and seasonal conditions and we have a limited understanding of how pests and natural enemies respond to this diversity. In this study,

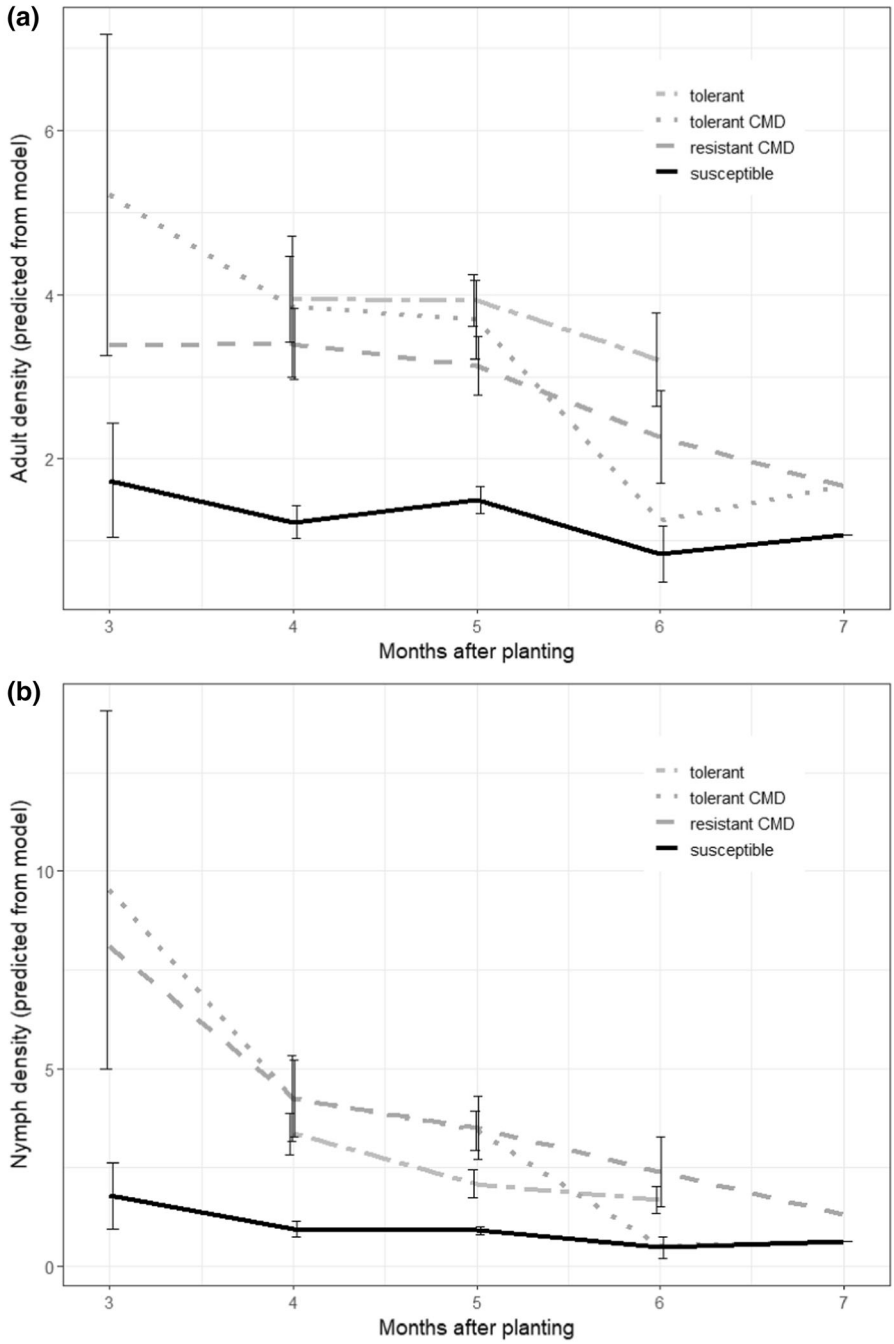

Fig. 5 Results from the final statistical model. *Bemisia* adults (**a**) and nymphs (final model 2, Online Appendix 1) **b** in cassava fields concerning months after planting and cassava cultivar category. The cassava categories represent cultivars that are tolerant to CMD and CBSD ("tolerant"), tolerant to CMD ("tolerant CMD"), resistant to CMD ("resistant CMD"), and landraces that are susceptible to both disease ("susceptible"). Full model outputs can be found in Online Appendix 1

we used a broad geographic survey approach to describe how these diverse production landscapes impact the variability in *Bemisia* density in cassava fields. We found large differences in the density of *Bemisia* adults and nymphs in focal fields across the geographic gradient. Within each region, we have

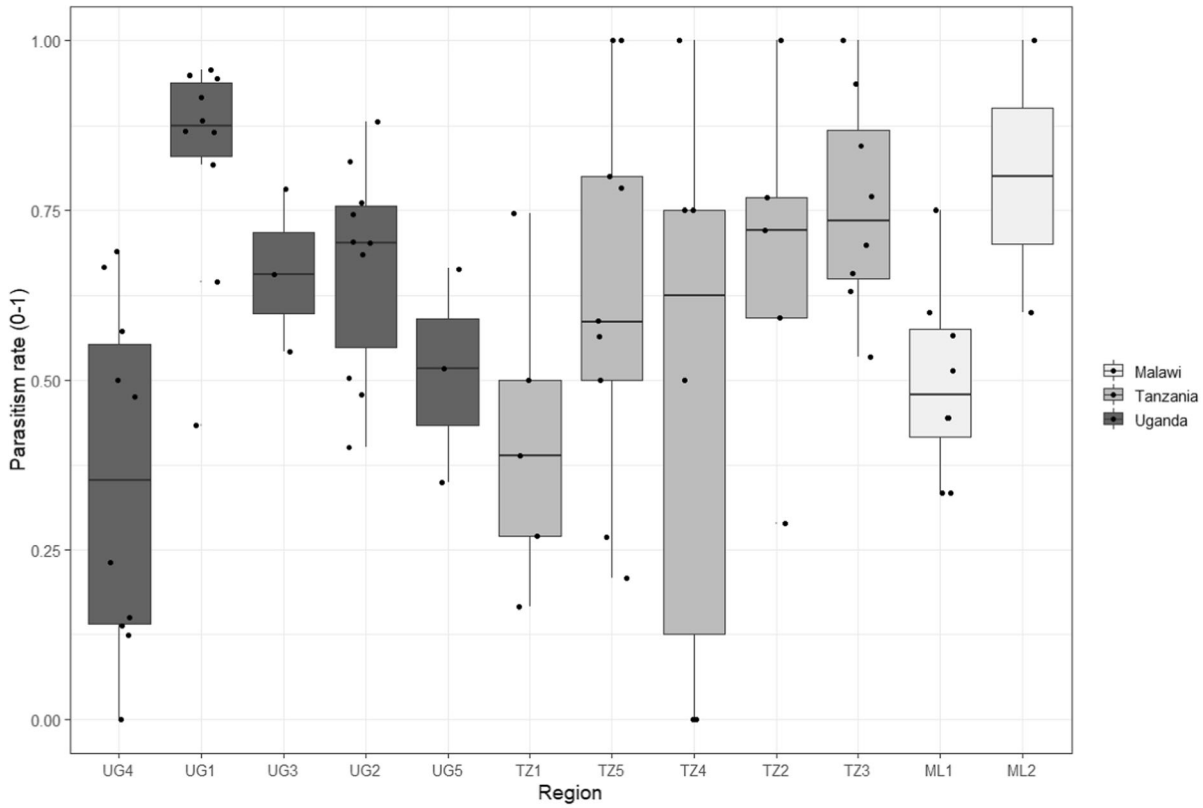

Fig. 6 Parasitism rate (as a proportion of parasitoids emerged overall live rearings) of *Bemisia* nymphs in cassava fields across 12 regions in East Africa (running from north UG4 Lira in Uganda, to south ML2 southern Malawi). Data set includes 79 fields

identified a range of in-field factors and landscape factors that contribute to the variability in *Bemisia* density observed between fields. For adults the in-field factors of cassava cultivar category and age of cassava (months after planting) were important, but for the nymph counts and parasitism factors outside the field in the broader landscape (e.g., area of cassava, size of the focal field) became important. These landscape factors will interact with regional factors (that were not examined in this study) to lead to high or low populations of *Bemisia* species in focal cassava fields. For example, we know that long-term climate and short-term weather patterns at the regional level can dampen or facilitate populations (Macfadyen et al. 2018). However, the in-field and landscape factors identified in this study have some potential for management by farmers, whereas climate patterns are generally outside of the control of individual smallholder farmers.

What are the common species of whitefly found on cassava and on nearby crops and non-crop host plants?

Given that the *B. tabaci* pest complex is still undergoing significant taxonomic and nomenclatorial changes (Boykin et al. 2018; Kunz et al. 2019) based on new molecular approaches for identifying species, reciprocal-crossing studies and consideration of implications from NUMTs/pseudogenes on *Bemisia* cryptic species status delimitation (Tay et al. 2017; Kunz et al. 2019), we spent a significant amount of time identifying samples of whitefly nymphs collected from cassava fields. This data set represents the most comprehensive molecular identification of samples from this region. We found that the *Bemisia* cyptic species community in cassava is dominated by the geographically widespread *B. tabaci* SSA1 species, however other *B. tabaci* and 'non-*tabaci*' species were also relatively common in some regions. This supports

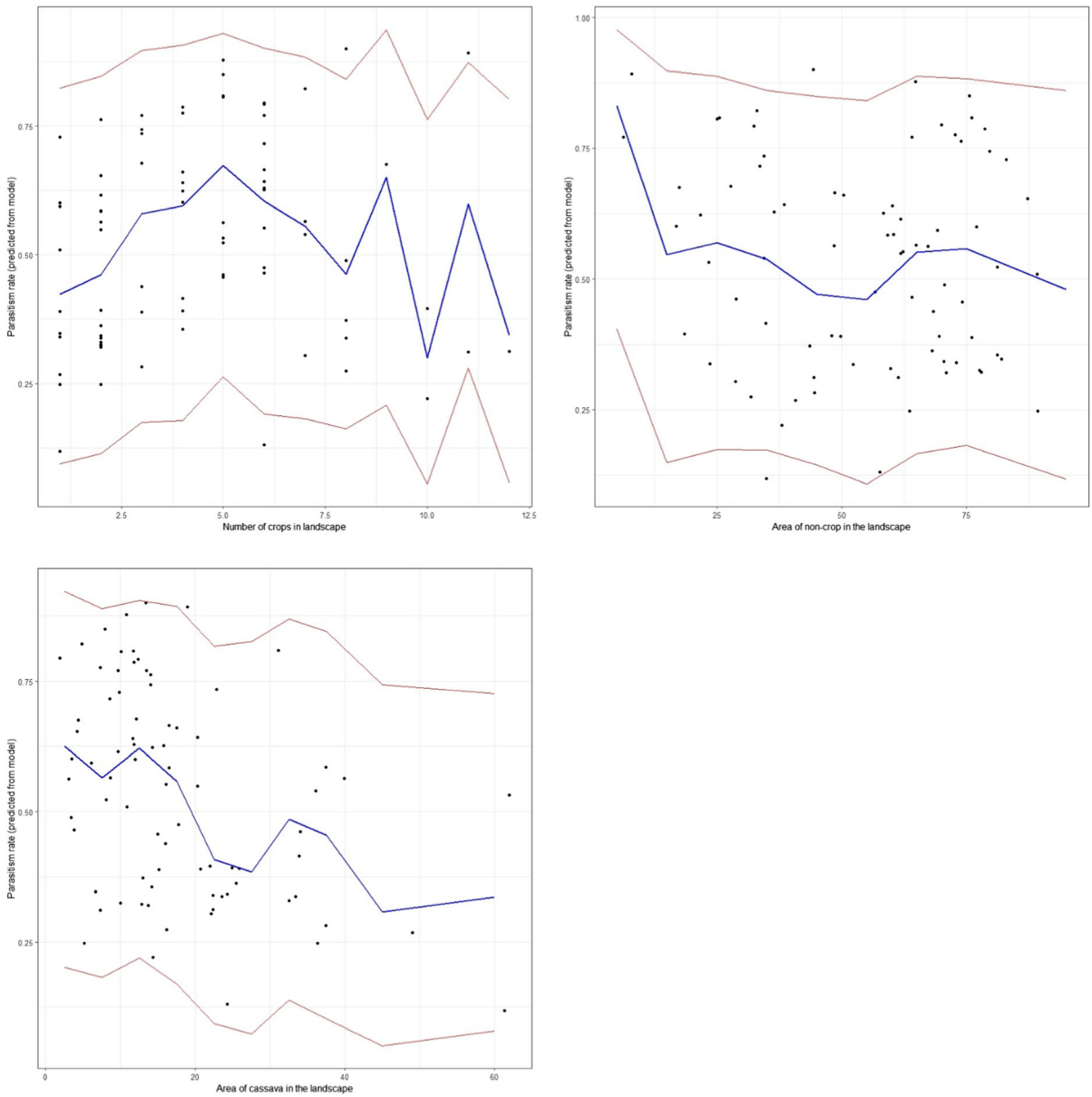

Fig. 7 Results from the final model. Parasitism rate (as a proportion of parasitoid adults emerged over all live rearings, between zero to one) of *Bemisia* nymphs in cassava fields in relation to the amount of cassava in the landscape, number of crops in the landscape and amount of non-crop in the landscape. Including region as a random effect. The blue line shows the fitted values and the two red lines the upper and lower 95% confidence interval

previous studies that also found *B. tabaci* SSA1 to be widespread on cassava in Tanzania (Tajebe et al. 2015), Kenya and Uganda (Mugerwa et al. 2012), although with many fewer samples. In this study, we have used the phylogenetic mtCOI grouping of SSA1 and consider it to be a single, putative species, based exclusively on the mtCOI classification. However, previous studies have identified mtCOI sub-groups and sub-clades within SSA1 using various methods, and in the future, there may be more biological diversity described within this putative species (Ally et al. 2019; Elfekin et al. 2019). Future molecular characterisation of the mtCOI barcoding gene region used for species identification and ascertainment of

genetic diversity should consider the impact of NUMTs/pseudogenes especially when utilising sub-optimal PCR markers (e.g., see Tay et al. 2017; Mugerwa et al. 2018; Elfekih et al. 2018; Kunz et al. 2019). In Malawi, *B. afer* was also present in our nymph samples and could represent an important vector for some of the cassava diseases (Maruthi et al. 2005). In Tanzania, the *B. tabaci* IO species was also relatively common in some cassava fields, on cassava plants. Misaka et al. (2020) did not record *B. tabaci* IO on cassava in a survey conducted in South Sudan, however, their study focussed on collecting adults and sequenced low numbers of individuals from cassava. In our study, *B. tabaci* SSA2 was found relatively frequently but in low density on other crop plants and non-crop plants in cassava production landscapes (Fig. 3). The *B. tabaci* SSA3 species was not detected in our study but its presence in cassava was previously reported further west in the Democratic Republic of Congo (Legg et al. 2014a, b).

We know that species in the *B. tabaci* complex differ in the range of plant species that they utilise as host plants (i.e. those that support successful growth and reproduction, Malka et al. 2018; Vyskočilová et al. 2018, 2019). Whilst adults were detected on a large diversity of plants in every production landscape, nymphs were recorded on a smaller selection of plant species and these were the focus of our molecular identifications. Sweet potato, given its frequent planting in many regions of East Africa, maybe a potential alternate host plant for *B. tabaci* SSA1 and its parasitoids. *B. tabaci* SSA1 was more detected on Tanzanian sweet potato samples than samples from Uganda (Fig. 3). However, more controlled host plant choice experiments are required to confirm this.

Which landscape or in-field factors influence the density of *Bemisia* adults and nymphs in cassava fields?

When examining factors that influence the density of adult *Bemisia* in cassava fields we found that in-field factors were the most important. In particular, the cassava cultivar sampled and the age of the cassava, with highest densities observed on improved cassava cultivars at 3 months after planting. It was surprising to see such dominance of cassava cultivars from such a geographically broad field-survey. Differences in whitefly infestation between cassava cultivars has been shown in replicated small plot trials in multiple locations in Nigeria (Ariyo et al. 2005) and Uganda (Omongo et al. 2012; Katono et al. 2015). However, these studies do not consider the alternative host plants present surrounding these trials, which may also be attractive and accessible to adult *Bemisia* species. In these plot trials the local landraces (unimproved cultivars that are assumed to be susceptible to diseases) also experienced lower adult *Bemisia* whitefly pressures.

For the variation in nymph density observed in the focal fields, factors in the surrounding landscape were important (along with cassava cultivar and months after planting). The area occupied by the focal field (positive relationship) and the amount of cassava surrounding the focal field (negative relationship) were both important factors in the final model. The number of nymphs on a cassava plant will a priori be related to both the numbers of adults, the number of eggs laid, and the presence of factors that induce mortality at the egg, first and second instar stage of development. These include processes such as competition, host plant defenses, environmental events, and predatory and parasitic natural enemies. For example, high rainfall events may wash away nymphs and deter adults from ovipositing on the upper leaves (Katono et al. 2019). However, these events should be unrelated to the landscape patterns observed in our study. There was a negative relationship between nymph density in the focal field and the area of cassava in the landscape. If mortality due to predatory natural enemies (e.g. ladybeetles and ants) is the main causative agent behind this pattern, then the natural enemy species may be gaining some benefit from increased amounts of cassava in the landscape. This theory is supported by the fact that non-crop factors appeared unimportant for nymph density in the focal fields. It may be that predatory natural enemies associated with the cropped components of the landscape can reach higher densities and therefore increase the mortality of whitefly nymphs in cassava.

Which landscape or in-field factors influence the parasitism of *Bemisia* nymphs in cassava fields?

Although unusual for intensive agricultural production systems, the high levels of parasitism observed in this survey were not unexpected for smallholder systems in

East Africa (especially given the low pesticide use, see Online Appendix 2). High parasitism rates on cassava (58–67%) have been recorded in trials in Uganda (Otim et al. 2005, 2008). Furthermore, the detection of parasitoid DNA in the *Bemisia* nymph samples also suggests that parasitoids are a common part of this community. The next step is to determine the impact these parasitoids have on pest population reduction at the field and landscape-level.

The parasitism rate of *Bemisia* in focal cassava fields was heavily influenced by factors occurring in the landscape surrounding the field. However, this pattern was not solely related to the amount of non-crop vegetation, or the diversity of crops grown in the landscape. Highest rates of parasitism were recorded in fields that were in landscapes with $< 20\%$ cassava (relatively low), < 40–50% non-crop vegetation (low-moderate) and had an intermediate level of crop diversity (~ 5 crops in the landscape) (Fig. 7). However, our final model did not explain much of the variation in parasitism rate observed, suggesting that other factors (that we did not measure here) may also be influencing parasitoid behaviour and utimatley their ability to cause mortatility to whitefly. It would be impossible to disentangle potential causative mechanisms without a manipulative field study. However, it is likely that the parasitoid species are using multiple whitefly hosts in the landscape, some of which may only be found on cassava, but others may have multiple host plants. For example, Guastella et al. (2015) recorded some parasitoid species attacking both *B. tabaci* and *B. afer* hosts in cassava growing regions of Tanzania. For a host-specific parasitoid, that parasitizes a herbivore that is also specific to certain host plants, landscape factors become less important. In southeast Asia, the cassava mealybug, *Phenacoccus manihoti*, colonized cassava fields earlier in high diversity landscapes, but overall abundance was the same in high and low diversity landscapes. Furthermore, there were no landscape effects seen for the parasitism of the mealybug by the host-specific parasitoid, *Apoanagyrus lopezi* (Le et al. 2018). In contrast, if the community of parasitoids attacking *Bemisia* species also attack other closely related herbivore hosts that use a range of host plants, then factors outside the field may influence their density and behaviour more.

Whilst we sampled a large number of focal cassava fields from different regions across East Africa, there are regions that may differ in their climate, agronomic practices and whitefly community which we have missed in this survey. Furthermore, there may be interactions between climate and landscape variables we have not included in this study. The static spatial landscape descriptors used in our study do not account for the dynamic nature of these smallholder production landscapes across the season. The 100 m radius we selected around centroid of the focal field was relatively small compared to other landscapes studies (e.g. Thies et al. 2003; Janković et al. 2016), however it was large enough to capture the land-use diversity surrounding the focal fields (see maps in Online Appendix 1). The limitation here is that we cannot rule out other drivers of whitefly population dynamics that may take place at larger spatial scales. For example, the spatial scale we used would capture the regular movements of between local land-use units as they search for oviposition sites (Kalyebi et al. 2018), but would not capture large-scale movements from one region to another as the cassava season comes to an end. Importantly, we did not classify the age structure of the surrounding cassava and the suitability of growth stage of other crop types and how these resources changed across time (see basic rotation information in Online Appendix 2). We know that *Bemisia* adults move around these production landscapes easily and maybe responding to a diversity of cues concerning oviposition choice (Kalyebi et al. 2018). In some regions, and on non-cassava host plants, there was a relatively high proportion of species outside of the *B. tabaci* complex recorded in our genotype samples. This may lead to some of the whitefly-parasitoid interactions recorded here to also involve *B. afer* and/or other *Bemisia* species. Currently, our high throughput sequencing approach cannot differentiate individual whitefly host species-parasitoid interactions (as we batch-processed 20–40 individuals in each population).

How can our findings help improve farmers' management of whitefly pests?

Given the low usage of pesticides in these production landscapes, control options for smallholder farmers are limited to cultural options and choice of cultivars. Avoiding some of the improved cultivars that appear to support higher population densities of *B. tabaci* species is very important. Fortunately, our study has

shown that cassava cultivar does not impact the top-down reduction in nymphs due to parasitoids (also seen by Otim et al. 2006). Conducting small plot trials to rank commonly used cassava cultivars in each region by their ability to support the cassava *Bemisia* species would provide a useful resource for farmers. There may be cultivars that are tolerant to disease but are relatively poor at supporting the *Bemisia* species found in each region. In the future, the development of cassava cultivars that are simultaneously resistant to *Bemisia* whitefly and diseases may be available to farmers. However, careful consideration relating to the deployment of these cultivars in complex smallholder production landscapes will nevertheless remain necessary (see modelling conducted by Parry et al. 2020). In the meantime, farmers can potentially reduce whitefly adult and nymph densities on cassava by facilitating a mosaic landscape that consists of smaller sized cassava fields, and by avoiding placing new cassava fields adjacent to fields with existing high adult whitefly populations (Kalyebi et al. 2018). Furthermore, in our study, the highest rates of parasitism were seen in fields that were in landscapes with relatively low amounts of cassava ($< 20\%$ cassava), low-moderate amounts of non-crop vegetation (< 40–50%), and an intermediate level of crop diversity ($\sim$ 5 crops in the landscape). There may be several ways that farmers may enhance landscapes with these characteristics. Finally, agronomists and extension workers supporting smallholder farmers need to be aware of the larger-scale and longer-term processes that operate on *Bemisia* species populations and also impact the risk of pest outbreak. For example, we know that some parts of Uganda have gone through an extended dry period which may make these regions more suitable for *B. tabaci* SSA1 growth and development (Kriticos et al. 2020), and in some cases, the double-cropping season may also support higher population abundances of *Bemisia* species at the landscape-level (Parry et al. 2020).

Acknowledgements This work was supported by the Natural Resources Institute, University of Greenwich from a grant provided by the Bill & Melinda Gates foundation (Grant Agreement OPP1058938). We would like to thank the broader cassava whitefly project team for their enthusiasm and support throughout this study. Thank you to the technical staff involved with the field work; Tom Omara, Patrick Ocitti, Wamani Sam and Kasifa Katono (National Crops Resources Research Institute, Uganda), Paschal Chacha, Thomas Mrosso, and Leonia Mlaki (Mikocheni Agricultural Research Institute, Tanzania) and Tiferra Mankwandha, Yolice Tembo and Adessa Kabanga (Department of Agricultural Research Services, Malawi).

References

Adriko J, Sserubombwe WS, Adipala E, Bua A, Thresh JM, Edema R (2011) Response of improved cassava varieties in Uganda to cassava mosaic disease (CMD) and their inherent resistance mechanisms. Afr J Agric Res 6:521–531

Alicai T, Omongo CA, Maruthi MN, Hillocks RJ, Baguma Y, Kawuki R, Bua A, Otim-Nape GW, Colvin J (2007) Re-emergence of cassava brown streak disease in Uganda. Plant Dis 91:24–29

Alicai T, Ndunguru J, Sseruwagi P, Tairo F, Okao-Okuja G, Nanvubya R, Kiiza L, Kubatko L, Kehoe MA, Boykin LM (2016) Cassava brown streak virus has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. Sci Rep 6:36164

Ally HM, Hamss HE, Simiand C, Maruthi MN, Colvin J, Omongo CA, Delatte H (2019) What has changed in the outbreaking populations of the severe crop pest whitefly species in cassava in two decades? Scientific reports 9(1):14796. https://doi.org/10.1038/s41598-019-50259-0

Ariyo OA, Dixon AGO, Atiri GI (2005) Whitefly *Bemisia tabaci* (Homoptera : Aleyrodidae) infestation on cassava genotypes grown at different ecozones in Nigeria. J Econ Entomol 98:611–617

Ateka E, Alicai T, Ndunguru J, Tairo F, Sseruwagi P, Kiarie S, Makori T, Kehoe MA, Boykin LM (2017) Unusual occurrence of a DAG motif in the Ipomovirus Cassava brown streak virus and implications for its vector transmission. PLoS ONE 12:e0187883

Barton K (2018). MuMIn: multi-model inference. R package version 1.42.1. https://CRAN.R-project.org/package=MuMIn

Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48

Bianchi FJJA, Walters BJ, ten Hove ALT, Cunningham SA, van der Werf W, Douma JC, Schellhorn NA (2015) Early-season crop colonization by parasitoids is associated with

native vegetation, but is spatially and temporally erratic. Agr Ecosyst Environ 207:10–16

Bjornstad ON (2019). ncf: Spatial Covariance Functions. R package version 1.2–8. https://CRAN.R-project.org/package=ncf

Boykin LM, Kinene T, Wainaina JM, Savill A, Seal S, Mugerwa H, Macfadyen S, Tay WT, De Barro P, Kubatko L, Alicai T (2018) Review and guide to a future naming system of African *Bemisia tabaci* species. Syst Entomol 43:427–433

Chaplin-Kramer R, O'Rourke ME, Blitzer EJ, Kremen C (2011) A meta-analysis of crop pest and natural enemy response to landscape complexity. Ecol Lett 14:922–932

Duarte GT, Santos PM, Cornelissen TG, Ribeiro MC, Paglia AP (2018) The effects of landscape patterns on ecosystem services: meta-analyses of landscape services. Landsc Ecol 33:1247–1257

Elfekih S, Etter P, Tay WT, Fumagalli M, Gordon K, Johnson E, De Barro P (2018) Genome-wide analyses of the *Bemisia tabaci* species complex reveal contrasting patterns of admixture and complex demographic histories. PLoS ONE 13:e0190555

Elfekih S, Tay WT, Polaszek A, Gordon KHJ, Kunz D, Macfadyen S, Walsh TK, Vyskočilová S, Colvin J, De Barro PJ (2019) On species delimitation, hybridization and population structure of cassava whitefly in Africa. BioRxiv. https://doi.org/10.1101/836072

Environmental Systems Research Institute, Inc. (2010) ArcGIS Desktop Version 10.2. Redlands, CA.

Environmental Systems Research Institute, Inc. (2015) Collector for ArcGIS Version 10.4. Redlands, CA.

Fox J, Weisberg S (2019). An {R} companion to applied regression, Third Edition. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Guastella D, Lulah H, Tajebe LS, Cavalieri V, Evans GA, Pedata PA, Rapisarda C, Legg JP (2015) Survey on whiteflies and their parasitoids in cassava mosaic pandemic areas of Tanzania using morphological and molecular techniques. Pest Manag Sci 71:383–394

Gurr GM, Reynolds OL, Johnson AC, Desneux N, Zalucki MP, Furlong MJ, Li Z, Akutse KS, Chen J, Gao X, You M (2018) Landscape ecology and expanding range of biocontrol agent taxa enhance prospects for diamondback moth management. A review. Agron Sustain Dev 38:23

Hartig F (2019). DHARMa: Residual diagnostics for hierarchical (Multi-Level/Mixed) regression models. R package version 0.2.4. https://CRAN.R-project.org/package=DHARMa

Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G (2010) Open data kit: tools to build information services for developing regions. Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development (2010): 1–12

Janković M, Plećaš M, Sandić D, Popović A, Petrović A, Petrović-Obradović O, Tomanović Ž, Gagić V (2016) Functional role of different habitat types at local and landscape scales for aphids and their natural enemies. J Pest Sci. https://doi.org/10.1007/s10340-016-0744-9

Knowles JE, Frederick F (2019) merTools: tools for analyzing mixed effect regression models. R package version 0.5.0. https://CRAN.R-project.org/package=merTools

Kalyebi A, Macfadyen S, Parry H, Tay WT, De Barro P, Colvin J (2018) African cassava whitefly, *Bemisia tabaci*, cassava colonization preferences and control implications. PLoS ONE 13:e0204862

Kanakala S, Ghanim M (2019) Global genetic diversity and geographical distribution of *Bemisia tabaci* and its bacterial endosymbionts. PLoS ONE 14:e0213946

Karp DS, Chaplin-Kramer R, Meehan TD, Martin EA, DeClerck F, Grab H, Gratton C, Hunt L, Larsen AE, Martínez-Salinas A, O'rourke ME (2018) Crop pests and predators exhibit inconsistent responses to surrounding landscape composition. Proc Natl Acad Sci USA 115:E7863–E7870

Katono K, Alicai T, Baguma Y, Edema R, Bua A, Omongo CA (2015) Influence of host plant resistance and disease pressure on spread of cassava brown streak disease in Uganda. Am J Exp Agric 7:284–293

Katono K, Macfadyen S, Omongo CA et al (2019) Influence of cassava morphological traits and environmental conditions on field population of *Bemisia tabaci* in Uganda. Submitted.

Kriticos DJ, Darnell RE, Yonow T et al (2020) Improving climate suitability for *Bemisia tabaci* in East Africa is correlated with increased prevalence of whiteflies and cassava diseases. Under review.

Kristensen NP, De Barro PJ, Schellhorn NA (2013) The initial dispersal and spread of an intentional invader at three spatial scales. PLoS ONE. https://doi.org/10.1371/journal.pone.0062407

Kunz D, Tay WT, Elfekih S, Gordon KHJ, DeBarro PJ (2019) Take out the rubbish—removing NUMTs and pseudogenes from the *Bemisia tabaci* cryptic species mtCOI database. bioRxiv. https://doi.org/10.1101/724765

Le TTN, Graziosi I, Cira TM, Gates MW, Parker L, Wyckhuys KAG (2018) Landscape context does not constrain biological control of *Phenacoccus manihoti* in intensified cassava systems of southern Vietnam. Biol Control 121:129–139

Legg JP, Thresh JM (2000) Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. Virus Res 71:135–149

Legg JP, Jeremiah SC, Obiero HM, Maruthi MN, Ndyetabula I, Okao-Okuja G, Bouwmeester H, Bigirimana S, Tata-Hangy W, Gashaka G, Mkamilo G (2011) Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. Virus Res 159:161–170

Legg JP, Shirima R, Tajebe LS, Guastella D, Boniface S, Jeremiah S, Nsami E, Chikoti P, Rapisarda C (2014a) Biology and management of *Bemisia* whitefly vectors of cassava virus pandemics in Africa. Pest Manag Sci 70:1446–1453

Legg JP, Sseruwagi P, Boniface S, Okao-Okuja G, Shirima R, Bigirimana S, Gashaka G, Herrmann HW, Jeremiah S, Obiero H, Ndyetabula I, Tata-Hangy W, Masembe C, Brown JK (2014b) Spatio-temporal patterns of genetic change amongst populations of cassava *Bemisia tabaci* whiteflies driving virus pandemics in East and Central Africa. Virus Res 186:61–75

Macfadyen S, Paull C, Boykin LM, De Barro P, Maruthi MN, Otim M, Kalyebi A, Vassão DG, Sseruwagi P, Tay WT, Delatte H (2018) Cassava whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in East African farming

landscapes: a review of the factors determining abundance. Bull Entomol Res. https://doi.org/10.1017/S0007485318000032

Malka O, Santos-Garcia D, Feldmesser E, Sharon E, Krause-Sakate R, Delatte H, van Brunschot S, Patel M, Visendi P, Mugerwa H, Seal S (2018) Species-complex diversification and host-plant associations in *Bemisia tabaci*: a plant-defence, detoxification perspective revealed by RNA-Seq analyses. Mol Ecol 27:4241–4256

Maruthi MN, Colvin J, Thwaites RM, Banks GK, Gibson G, Seal S (2004) Reproductive incompatibility and cytochrome oxidase I gene sequence variability amongst host-adapted and geographically separate *Bemisia tabaci* populations. Syst Entomol 29:560–568

Maruthi MN, Colvin J, Seal S, Gibson G, Cooper J (2002) Co-adaptation between cassava mosaic geminiviruses and their local vector populations. Virus Res 86:71–85

Maruthi MN, Hillocks RJ, Mtunda K, Raya MD, Muhanna M, Kiozia H, Rekha AR, Colvin J, Thresh JM (2005) Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). J Phytopathol 153:307–312

Misaka BC, Wosula EN, Marchelo-d'Ragga PW, Hvoslef-Eide T, Legg JP (2020) Genetic diversity of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) colonizing sweet potato and cassava in South Sudan. Insects 11:58

Midega CAO, Jonsson M, Khan ZR, Ekbom B (2014) Effects of landscape complexity and habitat management on stemborer colonization, parasitism and damage to maize. Agr Ecosyst Environ 188:289–293

Mugerwa H, Rey ME, Alicai T, Ateka E, Atuncha H, Ndunguru J, Sseruwagi P (2012) Genetic diversity and geographic distribution of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) genotypes associated with cassava in East Africa. Ecol Evol 2:2749–2762

Mugerwa H (2019). Molecular characterisation of cassava-colonising populations of the whitefly, *Bemisia tabaci*. A thesis submitted as partial fulfilment for the award of a PhD Degree of the University of Greenwich, United Kingdom, January 2019.

Mugerwa H, Seal S, Wang H-L, Patel MV, Kabaalu R, Omongo CA, Alicai T, Tairo F, Ndunguru J, Sseruwagi P, Colvin J (2018) African ancestry of New World *Bemisia tabaci*-whitefly species. Sci Rep 8:2734

Omongo CA, Kawuki R, Bellotti AC, Alicai T, Baguma Y, Maruthi MN, Bua A, Colvin J (2012) African cassava whitefly, *Bemisia tabaci*, resistance in African and South American cassava genotypes. J Integr Agric 11:327–336

Otim M, Legg DJ, Kyamanywa S, Polaszek A, Gerling D (2006) Population dynamics of *Bemisia tabaci* (Homoptera: Aleyrodidae) parasitoids on cassava mosaic disease-resistant and susceptible varieties. Biocontrol Sci Technol 16:205–214

Otim M, Legg J, Kyamanywa S, Polaszek A, Gerling D (2005) Occurrence and activity of *Bemisia tabaci* parasitoids on cassava in different agro-ecologies in Uganda. Biocontrol 50:87–95

Otim M, Kyalo G, Kyamanywa S, Asiimwe P, Legg JP, Guershon M, Gerling D (2008) Parasitism of *Bemisia tabaci* (Homoptera: Aleyrodidae) by *Eretmocerus mundus* (Hymenoptera: Aphelinidae) on cassava. Int J Trop Insect Sci 28:158

Parry HR, Kalyebi A, Bianchi F, Sseruwagi P, Colvin J, Schellhorn N, Macfadyen S (2020) Evaluation of cultural control and resistance-breeding strategies for suppression of whitefly infestation of cassava at the landscape scale: a simulation modeling approach. Pest Manag Sci 76(8):2699–2710

Polaszek A, Evans GE, Bennett FD (1992) *Encarsia* parasitoids of *Bemisia tabaci* (Hymenoptera: Aphelinidae, Homoptera: Aleyrodidae)—a preliminary guide to identification. Bull Entomol Res 82:375–392

R Core Team (2018). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rwegasira GM, Momanyi G, Rey MEC, Kahwa G, Legg JP (2011) Widespread occurrence and diversity of Cassava brown streak virus (Potyviridae: Ipomovirus) in Tanzania. Phytopathology 101:1159–1167

Schellhorn NA, Parry HR, Macfadyen S, Wang Y, Zalucki MP (2015) Connecting scales: achieving in-field pest control from areawide and landscape ecology studies. Insect Sci 22:35–51

Sseruwagi P, Sserubombwe WS, Legg JP, Ndunguru J, Thresh JM (2004) Methods of surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa: a review. Virus Res 100:129–142

Sseruwagi P, Maruthi MN, Colvin J, Rey MEC, Brown JK, Legg JP (2006) Colonization of non-cassava plant species by cassava whiteflies (*Bemisia tabaci*) in Uganda. Entomol Exp Appl 119:145–153

Tajebe LS, Boni SB, Guastella D, Cavalieri V, Lund OS, Rugumamu CP, Rapisarda C, Legg JP (2015) Abundance, diversity and geographic distribution of cassava mosaic disease pandemic-associated *Bemisia tabaci* in Tanzania. J Appl Entomol 139:627–637

Tay WT, Elfekih S, Court LN, Gordon KHJ, Delatte H, De Barro PJ (2017) The trouble with MEAM2: implications of pseudogenes on species delimitation in the globally invasive *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex. Genome Biol Evol 9:2732–2738

Tay WT, Court LN, Macfadyen S, Jacomb F, Vyskocilova S, Colvin J, De Barro PJ (2020) A high-throughput amplicon sequencing approach for population-wide species diversity and composition survey. https://doi.org/10.1101/2020.10.12.336545

Thies C, Steffan-Dewenter I, Tscharntke T (2003) Effects of landscape context on herbivory and parasitism at different spatial scales. Oikos 101:18–25

Veres A, Petit S, Conord C, Lavigne C (2013) Does landscape composition affect pest abundance and their control by natural enemies? A review. Agri Ecosyst Environ 166:110–117. https://doi.org/10.1016/j.agee.2011.05.027

Vyskočilová S, Tay WT, van Brunschot S, Seal S, Colvin J (2018) An integrative approach to discovering cryptic species within the *Bemisia tabaci* whitefly species complex. Sci Rep 8:10886

Vyskočilová S, Seal S, Colvin J (2019) Relative polyphagy of "Mediterranean" cryptic *Bemisia tabaci* whitefly species and global-pest status implications. J Pest Sci. https://doi.org/10.1007/s10340-019-01113-9

Zhang W, Kato E, Bianchi F, Bhandary P, Gort G, van der Werf W (2018) Farmers' perceptions of crop pest severity in

Nigeria are associated with landscape, agronomic and socio-economic factors. Agric Ecosyst Environ 259:159–167

Zhou K, Huang J, Deng X, van der Werf W, Zhang W, Lu Y, Wu K, Wu F (2014) Effects of land use and insecticides on natural enemies of aphids in cotton: first evidence from smallholder agriculture in the North China Plain. Agric Ecosyst Environ 183:176–184