Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork

Blood host preferences and competitive inter-species dynamics within an African malaria vector species complex inferred from signs of animal activity around aquatic larval habitats distributed across a gradient of fully domesticated to fully pristine ecosystems in southern Tanzania.

Thesis presented by

Katrina Walsh

for the degree of

Research Master

University College Cork

School of Biological, Earth and Environmental Sciences

Head of School/Department: Prof. Astrid Wingler

Supervisor(s): Prof. Gerard Killeen, Dr. Fidelma Butler, Dr. Emmanuel Kaindoa

2023

Table of Contents

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

.

Acknowledgements

Firstly, I would like to sincerely thank Prof. Gerard Killeen, the AXA Research Fund and UCC, for providing me with such an extraordinary and unforgettable experience. I am very grateful for the consistent mentorship and support from Prof. Killeen over the last two years, all the way from the initial preparations until the final stages of thesis writing, and thankful to him for sharing his vast expertise and valuable advice. I would also like to thank Dr. Fidelma Butler for contributing her knowledge and for her words of support and encouragement. My sincerest thanks to Dr. Emmanuel Kaindoa, for sharing his knowledge and for welcoming me as a resident guest at the Environmental Health and Ecological Sciences Department at the Ifakara Health Institute during my time in Tanzania.

This study would also not have been possible without my colleagues and other project team members, who all contributed significantly towards this study and my own personal experience. Firstly, I would like to express my appreciation for Deogratius Kavishe, for leading the overall project team in the field, for providing the PCR data, and for his tireless efforts with respect to managing the endless list of logistical procedures that made the field work all possible. A very warm and special thanks to Lily Duggan and Lucia Tarimo, for not only collecting the wildlife data used in this study, but also for never being short on words of encouragement, for bringing lots of laughter and for becoming such dear friends.

I would also like to sincerely thank all the ILUMA village game scouts (VGS) who accompanied the investigators in the field for keeping us safe, for carrying out routine camp activities, and for their participation with respect to data collection and processing. In particular, I would like to acknowledge the late Octavian Mrope for his admirable leadership of the VGS team while in the field, and extend my thanks to Azizi and Semka, for their enthusiasm and dedication towards completing the additional round of larval occupancy surveys and larvae collection. I would also like to thank the entire VGS field team for their

patience when I spoke with them in Kiswahili, and for making the time in the bush so memorable. I am also grateful for all the invaluable assistance from the VGS and insectary technicians at the Msakamba camp, and to all the drivers and Nyerere National Park rangers for their time.

Thanks to my parents, Laurence and Martina, for always being so supportive, interested and assuring throughout my education. And to Peter, for his patience, commitment and for keeping me sane. And of course, to all my other family and friends who showed their support, understanding and listened to my stories. Finally, I would like to thank the many people of Ifakara, for welcoming me into their neighbourhood and community with so much warmth and kindness. Asanteni sana.

Abstract

The effectiveness of current first-choice vector control interventions on the semi-zoophagic vector *Anopheles arabiensis* are limited, among other things, by insecticide resistance and their ability to acquire non-human bloodmeals. Across much of southern Tanzania, in areas where humans and cattle are readily available, it is the predominant member sibling species of the *Anopheles gambiae* complex. However, little is known about its population dynamics or blood utilization behaviours in pristine natural ecosystems, where these known preferred hosts are scarce or completely absent.

This study investigated larval habitat occupancy and species composition of the *An. gambiae* complex, together with the availability of various potential mammalian blood host species, across a gradient of fully domesticated to fully pristine ecosystems in southern Tanzania. Potential aquatic habitats were surveyed at 40 locations encompassing permanent human settlements, a community-owned Wildlife Management Area with varying degrees of human activity, and Nyerere National Park (NNP), which had very little. To investigate the effects of potential host availability on the species composition of the complex, all direct observations, tracks and signs of humans, livestock and wild animals observed around the surveyed larval habitats were recorded in parallel. The resulting data were analysed by logistic regression using generalized linear mixed models.

Odds of habitat occupancy by *An. gambiae* complex larvae decreased by 62% ($P<0.0001$) across the full range of natural ecosystem integrities observed, from fully domesticated to completely pristine areas, suggesting that the availability of suitable blood hosts had a modest effect on overall habitat utilization by this taxon. However, while only *An. arabiensis*, a key vector of residual malaria transmission, was found in fully domesticated ecosystems, its non-vector sibling species *An. quadriannulatus* also occurred in conserved areas and dominated the most pristine natural ecosystems. The proportion of *An. arabiensis* versus *An. quadriannulatus*

was positively associated with the number of times humans and/or cattle (P=0.0007) were detected at a location and negatively associated with distance inside NNP and away from human settlements (P<0.0001). Nevertheless, *An. arabiensis* was found even in absolute pristine environments that were >40km away from any signs of human or livestock, suggesting this species can survive on blood from one or more wild animal species. High proportions of *An. quadriannulatus* inside NNP were positively associated with the number of times impala were detected (P<0.0001), suggesting they may be a preferred blood source for this non-vector, giving it a competitive advantage over *An. arabiensis* where this antelope is abundant. Despite being detected less frequently, bushpig were also positively associated with high proportions of *An. quadriannulatus*, suggesting they provide a second preferred blood source, particularly in the miombo woodlands of the buffer zone where impala were scarce.

Overall, it seems that the availability of preferred hosts influences the competitive balance between these sibling species and that refugia populations of *An. arabiensis* can persist in wild areas, thus presenting both challenges and opportunities for control interventions: While such refuge populations deep inside pristine conservation area may confound attempts to repeat historical successes with eliminating this species from areas outside its natural range, it may also constitute a reservoir of diverse, unselected genomes with original wild-type insecticide susceptibility traits that could repopulate areas around human settlements that had otherwise been considered lost from the natural gene pool.

Chapter 1

1. Introduction

1.1 Introduction to malaria transmission and the Anopheles gambiae complex.

Malaria in humans is a vector-borne disease by which *Plasmodium* parasites, predominantly *Plasmodium falciparum* and *Plasmodium vivax*, are transmitted from one person to another through the biting activity of *Anopheles* mosquitoes (WHO, 2023). While there are many biological factors that predict the intensity of malaria transmission by these vectors, including the ability to support sporogony development, the rate of bites on an infectious host, and vector longevity and survival rates (Macdonald, 1957; Garrett-Jones, 1964; Onori and Grab, 1980; Cohuet *et al.*, 2010), transmission ultimately relies on the acquisition of human bloodmeals. Therefore, although many *Anopheles* mosquitoes have the physiological vectorial competence to support the full sporogonic development process of human malaria parasites (Service and Townson, 2002), all the way from ingested gametocytes in the midgut through to infectious sporozoites in the salivary glands (Beier, 1998), actual transmission of the disease is driven by a few key species that innately prefer feeding upon humans (Kiszewski *et al.*, 2004; Sinka *et al.*, 2010; Sinka *et al.*, 2012). The fact that most of the global malaria burden is centred in sub-Saharan Africa can be readily explained by the coincidence of tropical climatic conditions with the presence of highly human-specialised mosquitoes, the most important of which are *An. funestus*, *An. gambiae sensu stricto (s.s.)*, *An. coluzzii* and *An. arabiensis* (Kiszewski *et al.*, 2004). Immense progress has been made to control malaria, averting 1.6 billion cases and over 10 million deaths in the last two decades (WHO, 2022), which have been largely attributed to the use of insecticide treated nets (ITNs) (Bhatt *et al.*, 2015). Despite this, the behavioural variation of important malaria vectors represents a biological limitation to further progress (Ferguson *et al.*, 2010; Durnez and Coosemans, 2013; Killeen, 2014) while insecticide

resistance threatens to undermine and unravel progress achieved to date (Hemingway *et al.*, 2016; Ranson and Lissenden, 2016; WHO, 2022).

An. gambiae s.s. is historically the most notorious vector of malaria across East Africa and is the nominate species of the *An. gambiae* complex, often referred to as *An. gambiae sensu lato (s.l.)*, which is comprised of at least eight other sibling species, most of which are morphologically indistinguishable (Coetzee *et al.*, 2013; Barrón *et al.*, 2019). This species complex, which is ubiquitous across sub-Saharan Africa includes three of the four most important vectors across the continent, two of which are found in east Africa (Coetzee, Craig and Le Sueur, 2000; Kiszewski *et al.*, 2004; Sinka *et al.*, 2012). *An. arabiensis* was first identified as a separate species during genetic studies in the 1960's (Davidson and Jackson, 1962) and historically shares overlapping geographic distributions with *Anopheles gambiae s.s.* (Gillies and Coetzee, 1987; Coetzee, Craig and Le Sueur, 2000), except that the former are better adapted to arid landscapes compared to the latter, which instead prefers wetter, more humid conditions (Lindsay, Parson and Thomas, 1998). While *An. coluzzii*, formerly known as the *An. gambiae* molecular M form, is also a hugely important vector for reasons that remain to be understood, it is confined to the central and western regions of Africa (Coetzee *et al.*, 2013). Excluding two saltwater species, *An. merus* and *An. melas*, that are respectively found along the east and west coasts of Africa (White, 1974; Gillies and Coetzee, 1987), as well as *An. bwambae* that is limited to geothermal springs in Uganda (White, 1973; Gillies and Coetzee, 1987), all other members of the *An. gambiae* complex undergo their aquatic lifecycle in freshwater habitats (Holstein, 1954; Gillies and De Meillon, 1968; Gillies and Coetzee, 1987).

1.2 Aquatic habitat utilisation of the Anopheles gambiae complex.

Anopheles gambiae complex larvae exhibit an extremely brief aquatic lifecycle, in which the combined processes of oviposition, larval development and pupal development into emerging adults are all completed within 5 to 14 days (Holstein, 1954; Gillies and De Meillon, 1968). While the stereotypical larval habitat of this species has been described as open, sunlit pools with only partial shade (Holstein, 1954; Gillies and De Meillon, 1968; Gimnig *et al.*, 2001), in reality these species are largely opportunistic and can be found in almost any water body with reasonably clean water (Holstein, 1954), and therefore have a highly ubiquitous distribution. These habitats are often anthropogenically formed, for example drainage ditches, borrow pits, tyre tracks, rice fields, and human or livestock prints (White, Magayuka and Boreham, 1972; Gillies and De Meillon, 1968). The complex can also be found in naturally-occurring waterbodies, including swamps or springs, pools created by retreating streambeds in the dry season, and other naturally occurring rain-fed depressions or puddles in the wet season (Gillies and De Meillon, 1968). Furthermore, even small highly ephemeral waterbodies can provide highly productive breeding sites, capable of facilitating impressive levels of malaria transmission (Fillinger *et al.*, 2004). The ubiquitous nature of *An. gambiae* complex larvae across such spatial, environmental, and temporal heterogeneities therefore create an unpredictably distributed target for larval source management vector control interventions (Killeen and Reed, 2018). Although, larval source management is advised as a supplemental control measure in some favourable cases, notably where aquatic habitats are relatively "few, fixed and findable" (Fillinger and Lindsay, 2011), the most successful and universally applicable first-line interventions that are still being used to date (Lengeler, 2004; Pluess *et al.*, 2010; Bhatt *et al.*, 2015; WHO, 2022) are those that depend on adult mosquito behaviour (Pates and Curtis, 2005; Durnez and Coosemans, 2013; Killeen, 2014).

1.3 The influence of host choice on malaria transmission and control.

Emerging, blood-seeking *An. gambiae* complex adults exhibit species-specific host choice behaviours that affect their vectorial capacity and the effectiveness of vector control interventions targeting those free flying adults. For example, *An. gambiae s.s.* exhibit a very strong, almost singular preference for human blood, correspondingly associated with one of the highest known human blood indices (HBI, defined as the proportion of blood meals a vector population obtains from humans), and so mediate exceptionally high rates of transmission (Garrett-Jones, 1964; White, 1974; Garrett-Jones, Boreham and Pant, 1980; Gillies and Coetzee, 1987; Costantini *et al.*, 1999; Kiszewski *et al.*, 2004). Hence, the highly anthropophagic nature of *An. gambiae s.s*. is a determining factor that makes this mosquito one of the most historically efficient vectors of life-threatening malaria (White, 1974; Kiszewski *et al.*, 2004; Besansky, Hill and Costantini, 2004; Killeen *et al.*, 2017a). In contrast, the highly zoophagic member, *An. quadriannulatus* almost exclusively feeds on animals other than humans and so are considered to be of little medical interest and negligible importance with respect to malaria transmission (White, 1974; Coluzzi *et al.*, 1979; Coluzzi, 1984; Gillies and Coetzee, 1987), despite its physiological susceptibility to the *Plasmodium* parasite (Takken *et al.*, 1999).

Despite the huge influence of vector preferences for human blood preference upon malaria transmission intensity, such anthropophagic traits among mosquitoes also leave them vulnerable to insecticidal control interventions that protect sleeping humans. Indoor residual spraying (IRS) and ITNs, mostly in the form of long-lasting insecticide treated nets (LLINs), are still almost universally deployed across Africa, and indeed many other parts of the tropics, as first-line interventions against vector mosquitoes (WHO, 2017; WHO, 2021; WHO, 2022). In contrast to conventional ITNs, whereby re-treatment of the net at least once a year is required for the insecticidal properties to be effective (WHO, 2002) and has therefore been difficult to

implement (Guillet *et al.*, 2001), LLINs are manufactured to have a longer lifespan of 3-5 years and survive multiple washes, so are therefore recommended (Guillet *et al.*, 2001; WHO, 2017).

These interventions have been highly effective at reducing malaria transmission (Lengeler, 2004; Pluess *et al.*, 2010; Bhatt *et al.*, 2015) because they exploit nocturnal blood-seeking and indoor resting behaviours (Killeen, 2014) that are typically associated with anthropophagy in general (Elliott, 1972) and characteristic to *An. gambiae s.s.* (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987). For example, the nocturnal biting times of *An. gambiae s.s.* coincide with the usual sleeping hours of most humans (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987), so contact with ITNs, ideally LLINs, kill or deter anthropophagic and endophagic mosquitoes when they are actively seeking a human bloodmeal indoors, while also physically protecting sleeping humans from infection during peak biting hours (Pates and Curtis, 2005). Furthermore, high coverage of LLINs can also benefit those that are unprotected by supressing overall density and survival of anthropophagic mosquito populations like *An. gambiae s.s.* (Hawley *et al.*, 2003; Killeen *et al.*, 2007b). Other synanthropic behaviours associated with anthropophagy are targeted with lethal doses of insecticide delivered by IRS, which targets the endophilic habits of adult mosquitoes that cause them to rest indoors on sprayed surfaces after acquiring a bloodmeal (Pates and Curtis, 2005). As a result of widespread uptake of LLINs and IRS targeting these behaviours, populations of *An. gambiae s.s.* have declined across much of their former natural range (Sinka *et al.*, 2016) and have even been effectively eliminated from large parts of east Africa, as documented in depth in some parts of Kenya (Bayoh *et al.*, 2010; Mwangangi *et al.*, 2013) and southeastern Tanzania (Russell *et al.*, 2010; Killeen *et al.*, 2013; Lwetoijera *et al.*, 2014; Kaindoa *et al.*, 2017).

However, local elimination of *An. gambiae s.s.* across large tracts of east Africa coincided with increased relative abundances of its morphologically indistinguishable sibling species *An. arabiensis*, resulting in a decisive species shift within the complex over a remarkably brief

period of a few short years (Russell *et al.*, 2010; Bayoh *et al.*, 2010; Mwangangi *et al.*, 2013; Killeen *et al.*, 2013; Sougoufara *et al.*, 2016). This shift in species composition was facilitated by the limitations of LLINs and IRS, as they have far less impact on less exclusively human-specialized vector mosquitoes like *An. arabiensis,* which exhibit behavioural phenotypic plasticity that allow them to minimize exposure to the active ingredients of LLINs and IRS (Muirhead-Thomson, 1960; Elliott, 1972; Pates and Curtis, 2005; Govella and Ferguson, 2012; Govella, Chaki and Killeen, 2013) and thus mediate residual malaria transmission (Durnez and Coosemans, 2013; Killeen, 2014; WHO, 2016).

In comparison to *An. gambiae s.s.*, that feed almost exclusively on human blood, *An. arabiensis* have long been known to also feed readily upon cattle (White, 1974; Gillies and Coetzee, 1987). Furthermore, as they do not appear to acquire bloodmeals from any other animal in domesticated landscapes (Killeen 2001), *An. arabiensis* seemingly demonstrate strong specialization upon these two hosts and can flexibly feed on either or both of them. Moreover, blood meal choices made by *An. arabiensis* are highly dependent upon the relative availabilities of cattle and humans, even across very fine scales of a few metres (White, Magayuka and Boreham, 1972; Killeen *et al.*, 2001). Host choice behaviour of *An. arabiensis* can therefore be explained by mechanisms of true phenotypic plasticity in the strict sense, rather than merely genetic variation, because the huge variability in blood meal choices observed by (White, Magayuka and Boreham, 1972) and modelled by (Killeen *et al.*, 2001), occurs at distances too short to be underpinned by distinct subpopulations of genotypes exhibiting different host choice preferences. Furthermore, although genetic components do appear to influence host species choice (Main *et al.*, 2016), and associated behaviours like biting times (Govella *et al.*, 2023), these heritable components of such behavioural variation appear to be relatively minor (Main *et al.*, 2016; Govella *et al.*, 2023).

Although feeding on cattle directly lowers vectorial capacity and transmission intensity through a lower HBI (Macdonald, 1957; Garrett-Jones, 1964), even *An. arabiensis* populations exhibiting a high degree of zoophagic activity will usually acquire human blood with sufficient frequency to mediate levels of transmission that are often considered moderate by the standards of Africa, but nevertheless, more than sufficient to sustain endemic populations of the malaria parasite (Beier, Killeen and Githure, 1999; Smith *et al.*, 2005; Killeen *et al.*, 2017a). These host choice behaviours makes it particularly difficult to control *An. arabiensis*, not only because feeding on cattle enable them to directly evade the insecticide pressures of LLINs and IRS interventions (Durnez and Coosemans, 2013; Killeen, 2014) that target the stereotypical behaviours associated with anthropophagic anophelines (Elliott, 1972; Gillies, 1988; Killeen, 2014), but also because preferences for feeding upon non-human hosts also appear to be associated in evolutionary terms with several other plastic behavioural phenotypes that further allow this vector of residual transmission to evade exposure to these insecticide delivery formats (Elliott, 1972; Killeen, 2014; Govella, Chaki and Killeen, 2013; Killeen and Chitnis, 2014).

Mosquitoes that feed upon animals other than humans typically exhibit exophagic and crepuscular feeding behaviours, with most biting activity occurring outdoors and peaking at dawn and/or dusk (Elliott, 1972), whereas highly anthropophagic mosquitoes usually feed indoors at night, where and when humans have historically been most vulnerable (Elliott, 1972; Gillies, 1988; Killeen, 2014). *An. arabiensis* are naturally associated with long foraging windows with peak biting activity at dawn when people tend to wake up and get out of bed (Dukeen and Omer, 1986; Braack *et al.*, 1994; Githeko *et al.*, 1996), and can also feed earlier in the evening during hours when people are usually still active (Dukeen and Omer, 1986; Russell *et al.*, 2011; Yohannes and Boelee, 2012). Crepuscular biting activity therefore occurs outdoors beyond the immediate reach of indoor-based interventions, representing a major

fundamental limitation to the efficacy of LLINs and IRS and hence, clearly contribute to residual malaria transmission (Durnez and Coosemans, 2013; Killeen, 2014; Sherrard-Smith *et al.*, 2019).

Outdoor resting, or exophily, is another characteristic trait of zoophagic mosquitoes and is often associated with outdoor biting (Elliott, 1972; White, 1974; Gillies and Coetzee, 1987). While fully zoophagic mosquitoes that are also almost completely exophilic, such as *An. pharoensis*, *An. tenebrosus* and *An. zeimanni*, rarely enter houses (Kampango *et al.*, 2023), they are also correspondingly inefficient secondary or incidental vectors if at all (Gillies and Coetzee, 1987). In contrast, like most of the other main malaria vectors of Africa, *An. arabiensis* is both exophagic and endophagic, typically feeding just as readily outdoors as indoors (Huho *et al.*, 2013; Killeen, 2014), so the majority of human exposure to this more efficient vector can occur outdoors in settings where it begins feeding early in the evening and continues feeding into the late morning (Govella, Okumu and Killeen, 2010), as described in the previous paragraph. The limitations of indoor intervention with respect to exophilic behaviours have long been recognised (Gillies, 1956; Muirhead-Thomson, 1960; Kouznetsov, 1977; Molineaux *et al.*, 1979), as these resting adults can not always be effectively reached by insecticides even if they have fed indoors (Pates and Curtis, 2005). For example, host-seeking *An. arabiensis* adults that enter houses or cattle sheds, tend to exit early, and so can avoid lethal doses of insecticide where interventions have been applied (Kitau *et al.*, 2012; Okumu *et al.*, 2013b; Okumu *et al.*, 2013a). Furthermore, adults that enter houses at night when people are protected under LLINs, can rapidly exit, and continue foraging for an unprotected host, even after one or more unsuccessful encounters (Killeen *et al.*, 2016).

Because all these different but closely associated evasive behaviours are all exhibited by *An. arabiensis*, and interact synergistically to render this species very resilient to attack with LLINs and IRS, it remains an abundant vector of residual malaria transmission across much of Africa

today (Russell *et al.*, 2010; Bayoh *et al.*, 2010; Ferguson *et al.*, 2010; Mwangangi *et al.*, 2013). Supplementary new control measures that either exploit these behavioural interactions, or target different behaviours and/or life stages, will be required if elimination of this species is to be achieved (Killeen *et al.*, 2013; Killeen *et al.*, 2017c).

1.4 Portfolio effect concept and investigating the ecology of potential refugia populations.

The portfolio concept originates from the field of economics but has also been adapted to ecological applications, to help identify the complex biological and ecological relationships and mechanisms in populations or ecosystems, that make them resilient to environmental or anthropogenic perturbations (Schindler, Armstrong and Reed, 2015). Furthermore, the portfolio concept can also be applied to help identify the mechanisms of vector resilience that create major challenges for traditional control intervention, effectively stabilizing mosquito populations and the residual malaria transmission they mediate (Killeen and Reed, 2018). Specifically, fine-scale variations in aquatic habitat abundance and availabilities of preferred hosts that occur across heterogenous landscapes interact with the flexible host-feeding behaviours of *An. arabiensis* described above to create such portfolio effects (Killeen and Reed, 2018). While these heterogeneities can occur at such fine scales as a few meters, they may also across larger geographic scales of many kilometres.

The host preferences and associated behaviours exhibited by *An. arabiensis* is indicative of an ecological niche adapted to domesticated land-use, particularly in arid areas (White, 1974; Coluzzi *et al.*, 1979; Lindsay, Parson and Thomas, 1998) where pastoralism is often the predominant livelihood (African Union, 2013). However, examples of *An. arabiensis* populations occurring in uninhabited areas, outside what is normally thought of as their typical niche in domesticated arid landscapes, have been identified inside African conservation areas. For example, adult *An. arabiensis* mosquitoes were caught during a study inside Mana Pools National Park, Zimbabwe (Prior and Torr, 2002). However, these experiments were completed

at a permanent research camp with a small resident population of cattle, which may therefore represent a refuge of environmental conditions like that of a small village in an otherwise pristine landscape. Although, a population of *An. arabiensis* has been documented at the Malahlapanga springs inside Kruger National Park in South Africa (Braack *et al.*, 1994; Munhenga *et al.*, 2011; Munhenga *et al.*, 2014), and it has been suggested that this malaria vector may be surviving on blood meals from wild animals rather than their known preferred human and cattle hosts (Braack *et al.*, 1994), this survey site was <20 km from the park boundary and adjacent villages. It therefore remains unclear whether *An. arabiensis* can exploit alternative hosts in wild areas, potentially creating an additional population-stabilizing portfolio effect by sustaining wild refugia populations far away from vector control interventions and insecticide pressure.

As vectors expressing phenotypic plasticity in their blood feeding behaviours underpin the persistent residual malaria transmission (Durnez and Coosemans, 2013; Killeen, 2014; WHO, 2016) that has contributed to the stalled progress towards malaria elimination, fully characterising the host choice ecology of this critically important semi-zoophagic vector may provide useful insights into the future potential and limitations of malaria vector control. Furthermore, physiological resistance to pyrethroids used in LLINs and IRS has clearly undermined progress to date and threatens a catastrophic resurgence of morbidity and mortality, so new solutions are urgently required (Hemingway *et al.*, 2016; Ranson and Lissenden, 2016). Therefore, this study contributes to a broader overall project which primarily aims to establish whether wild refugia populations of *An. arabiensis* exist inside conserved wilderness areas that are free from selective insecticide pressures and may therefore still exhibit original wild-type pyrethroid susceptibility traits that may prove to be an invaluable source population for managing insecticide resistance in surrounding villages.

Overall, the goal of this study at the outset was to shed light on the ecological behaviours that may create population-stabilizing portfolio effects which facilitate the robust survival of such wild populations and discuss their implications for ongoing and future vector control. More specifically, the *a priori* intention of this study, although some unforeseen new insights were fortuitously encountered along the way, was to investigate aquatic larval habitat occupancy of *An. arabiensis* across heterogenous environments with very different distributions of available mammalians hosts, ranging from fully domesticated settlements to absolute pristine ecosystems, to determine whether this key vector of residual malaria transmission could indeed survive in such wild refugia, presumably by feeding upon one or more species of wild mammal.

Chapter 2

2. Methods

This study addresses two objectives of a broader overall project, the primary goal of which was to collect live mosquitoes among potential refugia populations in wild areas to test for insecticide susceptibility (Kavishe *et al.*, Unpublished). Simultaneously, complementary objectives of this larger overall project, (Killeen, 2023) involved extensive land-use and wildlife surveys based around the same sampling sites (Duggan *et. al*., Unpublished). All three complimentary data sets and the investigators responsible for collecting them were mutually co-dependent with respect to addressing their individually assigned objectives and the overall goal of the project as a whole. Therefore, the integrated study design and the collectively managed logistical procedures that were required to successfully implement the data collection for all these distinct objectives, across such a large study area presenting such challenging environmental conditions, are emphasized.

2.1 Study area.

This study was conducted in the Kilombero valley in the Morogoro region of southern Tanzania, which is recognised as East Africa's largest wetland at low altitude (Mombo *et al.*, 2011; RIS, 2002). The valley is part of the Rufiji River Catchment Basin and is marked by the Udzungwa mountains in the north and the Mahenge mountains in the south (RIS, 2002). The wetlands of the Kilombero are a unique and biodiverse ecosystem, within 7,950 square kilometres have been designated as a Ramsar site since 2002 (RIS, 2002) (Figure 2.1.). As well as being of high ecological and hydrological importance, the wetlands and surrounding area are of major economic significance, supporting a variety of agricultural, forestry and fishing livelihoods (Mombo *et al.*, 2011).

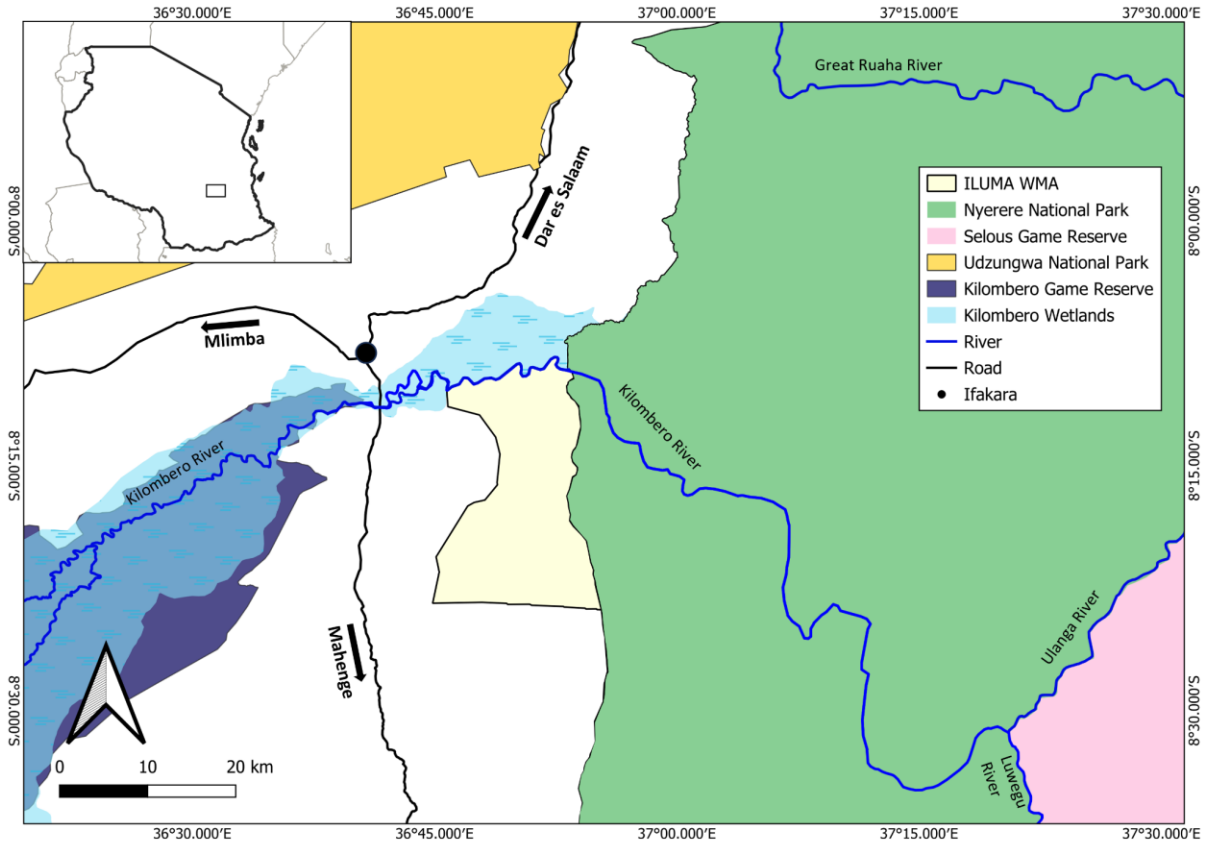

Figure 2.1. Map showing the Kilombero wetlands and the study area of the ILUMA Wildlife Management Area and Nyerere National Park.

The region has what are referred to as a short and a long rainy season, from approximately November to February and from March to April, respectively. The short rains are typically intermittent, sporadic and unreliable, whereas the long rains are usually more predictable and consistent. The dry season generally lasts from June to December, during which time the area experiences little to no rainfall. The malaria transmission systems and vector populations in the Kilombero valley have been exceptionally well-characterised (Takken *et al.*, 1998; Freyvogel and Kihaule, 1968; Smith *et al.*, 1993; Smith *et al.*, 1995; Charlwood *et al.*, 1995b; Charlwood, Vij and Billingsley, 2000; Charlwood, 1997; Charlwood *et al.*, 1997; Charlwood *et al.*, 1995a; Drakeley *et al.*, 2000; Russell *et al.*, 2011; Killeen *et al.*, 2007a; Killeen *et al.*, 2006; Lwetoijera *et al.*, 2014), and malaria vector research continues to progress with the Ifakara Health Institute being based in Ifakara town, with recent studies from (Katusi *et al.*, 2022; Kaindoa *et al.*, 2017; Mapua *et al.*, 2022; Okumu and Finda, 2021; Finda *et al.*, 2019; Nambunga *et al.*, 2020), to

name a few. Today, the two main vectors include *An. funestus* Giles and *An. arabiensis* Patton (Lwetoijera *et al.*, 2014; Maia *et al.*, 2016; Kaindoa *et al.*, 2017; Mapua *et al.*, 2022).

Located south-east of Ifakara town in the Kilombero and Ulanga districts, lies more than 500km^2 of community lands under the stewardships of 15 villages which is collectively managed through the Ifakara-Lupiro-Mangula Wildlife Management Area (ILUMA WMA) (Figure 2.1). It was formally established as a wildlife management area in 2015 for the purpose of practicing sustainable land use and promoting wildlife conservation in the buffer zone between a rigorously conserved state-protected area (formerly the Selous Game Reserve, now Nyerere National Park) and village communities, while also creating opportunities for local communities to generate income and advance rural development (Ministry of Natural Resources and Tourism, 2022). Miombo woodland is the predominant natural land cover type across most of the ILUMA WMA, although a sizeable trait of dense groundwater forest stretches along the south-bank of the Kilombero River. On the north-bank, floodplain grassland extends to the rainforest covered Udzungwa mountains. These very distinct ecosystems support sizeable populations of wild herbivores and carnivores, including iconic species like elephant, African buffalo, lion, leopard and African wild dogs.

The western boundary of ILUMA WMA borders lands designated for agriculture and grazing extending as far as the main road connecting Ifakara and Mahenge (Figure 2.1). Due to its proximity to domesticated land and human activity, many areas within ILUMA WMA close to the western and southern boundaries are vulnerable to unauthorised exploitation by humans, and large sections of these areas have even been cleared for agriculture, grazing and settlement (Duggan *et al.*, Unpublished). Other types of unauthorised human disturbance occurring across the designated conservation area, include timber harvesting, charcoal burning, fishing, and game meat hunting. However, as one moves further into ILUMA WMA and away from human settlement, natural ecosystems become less exposed to human disturbance and are increasingly

inhabited by abundant populations of wild mammals that offer mosquito populations diverse alternatives to humans and livestock as potential blood sources.

The eastern boundary of ILUMA WMA is marked by a road running in a north-south direction along its border with Nyerere National Park (NNP) (Figure 2.1). Established in 2019, NNP is Tanzania's largest and newest national park. NNP was formerly a part of the Selous Game Reserve and covers an area of more than 30,000 square kilometres of pristine wilderness (TANAPA, 2019). Presently, most tourist activities and development occur in the northern section of the park, along the Rufiji River between Ikwiriri and Kisaki, that have been previously used for photographic tourism rather than big game hunting. At the time of the study, however, the western area of NNP that borders ILUMA WMA had not yet been extensively developed, so it experiences minimal human activity. The miombo woodlands of ILUMA WMA, extend eastward beyond the border road and well into the national park (Figure 2.1), providing an extensive habitat for the same species that are found in ILUMA WMA. Additionally, open grasslands, acacia scrub and undisturbed floodplains in NNP favour species that are not regularly seen in ILUMA WMA, such as impala, waterbuck and kudu.

This large study area was fundamental for addressing the objectives of this specific study and the broader goals of the overall project, as it represents a geographical gradient of natural ecosystem integrity ranging from fully domesticated land uses in the west to completely pristine natural ecosystems to the east. Agricultural land and permanent villages to the west of ILUMA WMA offer typical larval habitats for mosquitoes of the *An. gambiae* complex, including early-stage rice paddies, ridge and furrow agriculture, livestock hoofprints, human-made tracks and artificial wells and drains (Holstein, 1954; Gillies and De Meillon, 1968; White, Magayuka and Boreham, 1972). In the wild areas of eastern ILUMA WMA and in NNP, waterholes, prints of various wild animals and seasonal pools were considered as water bodies that could provide potential breeding sites. Additionally, receding groundwater levels during

the dry season cause streams to stop flowing leaving small, stagnant isolated pools of groundwater within streambeds that can represent highly productive larval habitats (Gillies and De Meillon, 1968; Shililu *et al.*, 2007).

This comprehensive range of de facto land use practices and ecosystem characteristics were also considered to be indicative of the potential host species available to host-seeking *An. arabiensis* adults as potential sources of blood. Since cattle and humans are the known preferred blood source of *An. arabiensis* (White, 1974; Gillies and Coetzee, 1987), the abundance of this vector in and around villages may be readily rationalised because these two host species are plentiful in settlements and surrounding domesticated land. At the other end of the gradient, it has been suggested that wild game animals, particularly bovids, may act as suitable alternative hosts for wild *An. arabiensis* populations in wilderness areas (Braack *et al.*, 1994; Charlwood, Vij and Billingsley, 2000; Prior and Torr, 2002; Munhenga *et al.*, 2011), such as the pristine areas of ILUMA WMA and NNP. As WMAs act as buffer zones between national parks and permanently inhabited areas, wildlife, humans and cattle could all represent possible blood sources for *An. arabiensis* inside the ILUMA WMA. Furthermore, ILUMA WMA encompasses varying degrees of human disturbance and so the availabilities of cattle, human and wildlife species as potential blood sources was expected to vary considerably across fine geographical scales. Given the known blood-feeding behavioural plasticity amongst *An. arabiensis* populations (Section 1.3) and the anticipated variability in host species availability provided by the study area, it was therefore expected that these heterogeneities could be interpreted as a portfolio effect that stabilises vector populations (Killeen and Reed, 2018).

2.2 Sampling sites, study design and logistical field procedures.

This study was carried out using a rolling cross-sectional design with four rounds of surveys encompassing a total of 40 defined locations across NNP, ILUMA WMA and the villages immediately to the west of it, over the course of two years. Most sampling was carried out

during the wet season as rainfall influences mosquito abundance (Gillies and De Meillon, 1968; Charlwood *et al.*, 1995a). Larval occupancy surveys together with complimentary land use and wildlife surveys, were completed sequentially at the 40 defined locations, referred to herein as *camps,* that encompassed a range of land uses and ecosystem integrity states with a wide range of mammalian species abundance and diversity.

Because many of the camps were inaccessible to vehicles during the wet season, the whole data collection process was completed on foot except for the camps located up to 131 km inside NNP. These NNP camps were accessed via the Kilombero River using a motorboat in the wet season, and by car via established roads during the dry season. Additionally, transect surveys of land use and the activities of humans, livestock and wildlife were completed along the routes walked between camps within a survey circuit, and so these transits on what were referred to as *moving days*, needed to be completed on foot. Because of these challenges, the development of novel, logistically feasible procedures for collecting robust data and large numbers of live mosquito specimens, while also ensuring the safety of team members in a remote and challenging environment, were fundamental for the entire study framework.

A fenced camp with essential basic infrastructure like thatch roofs, tables, chairs, large tents, a kitchen and solar-powered electricity supply was established as the hub for all scientific and logistical processes in the field and named *Msakamba* (camp 1 in figure 2.2). The central location of Msakamba (figure 2.4) made it possible to reach any camp in ILUMA WMA within a day's walk, which ensured that live adult and larvae samples could usually be returned in good condition within 48 hours of collection. It also ensured that food supplies, as well as recharged batteries and power packs for mobile phones and field equipment could be regularly delivered to the mobile field team moving from camp to camp every two days. Village game scouts (VGS) recruited from stakeholder communities that are responsible for patrolling and protecting the ILUMA WMA, and for escorting any visitors to the conservation area, were also

engaged as essential team members for this study. Msakamba was occupied and maintained on a permanent basis by a team of VGS and technicians, who were responsible for maintaining the field insectary, rearing field-caught mosquitoes and carrying out experimental assessments on their insecticide resistance phenotypes.

Figure 2.2. Images of the central camp, Msakamba, including A; the surrounding fence, B; the kitchen and food storage area, C; tents for sleeping, D; sorting of wild-caught mosquitoes, E; a field insectary tent for housing mosquito adults and larvae, F; rearing containers of adult mosquitoes (mesh cages) and larvae (water basins) at the field insectary.

For the original protocol, a total of 28 camps were selected (Figure 2.4) after scouting potential locations in ILUMA WMA and considering the suggestions of the VGS based on their vast personal knowledge of the area. The broad geographic distribution of the camps was planned to encompass as wide of a range of ecosystem states as possible, by including all parts of ILUMA WMA and the neighbouring domesticated land to the west (Figures 2.4 to 2.7). However, the exact position of a camp location was ultimately determined by the requirement of perennial surface water that *Anopheles* larvae could be collected from. Crucially, the availability of water for cooking, drinking (after filtering as seen in Figure 2.3) and bathing was essential for sustaining the field team in the absence of regular vehicle support.

Figure 2.3. Filtering water from a nearby stream using the MSR Autoflow™ XL Gravity Filter 10L .

In most cases, the selection of camp locations with accessible surface water also ensured that firewood and at least some shade were available *in situ*. Accessibility by foot during the wet season was also a critical factor to consider, to ensure that each camp could be safely reached, and the transport of live mosquitoes could be completed even during periods of heavy rain and flooding. The presence of one or more glades or valleys with numerous perennial waterbodies, like waterholes, ponds and streambeds near the camp was also a requirement of the methodology used for collecting mosquito larvae and adults (Sections 2.3 and 2.4).

Once the selection process was complete, each camp was given a unique number and a name, and were then divided into quarterly *circuits* that roughly corresponded to the southeast, southwest, northeast, and northwest of the ILUMA WMA (Figure 2.4). The circuits were designed to manage the physical challenges of the study, and so camps were grouped into a circuit based on their geographic proximity to one another and the practicality of visiting them all in a circular route that started and ended centrally at the Msakamba camp (Figure 2.4). Each circuit had a planned route that minimised walking distances between camps which ensured that the team had enough time to rest in the afternoon before proceeding with data collection in the evening. Camps were visited consecutively for two nights each and involved intensive daily mosquito collection procedures (Section 2.3). Therefore, the circuit design for sequentially visiting camps in a rolling cross-sectional survey was crucial for sustaining optimal long-term data collection by limiting investigator fatigue and allowing them to take breaks of a few days in between periods of continuous field work that lasted about two weeks for a single circuit.

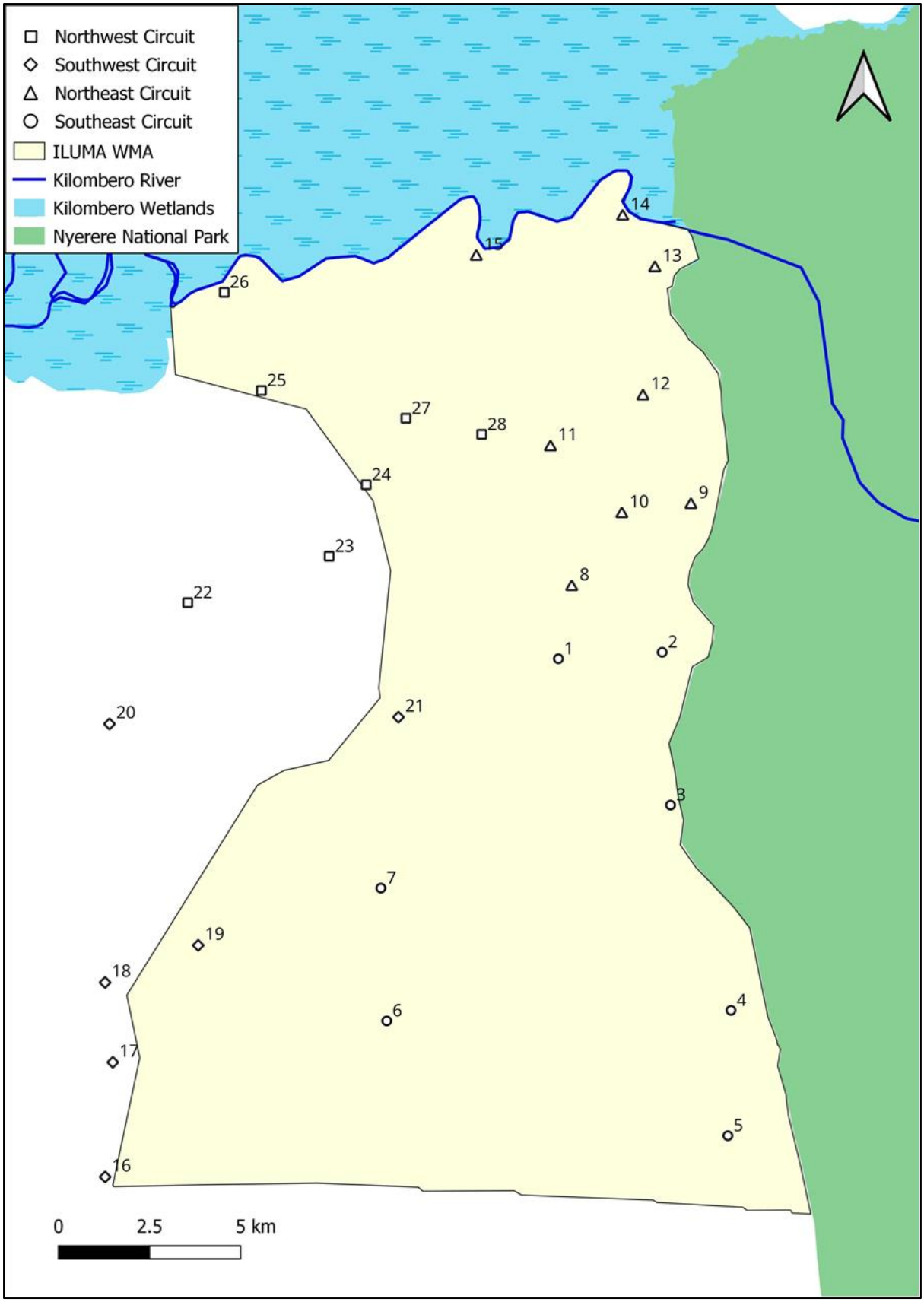

Figure. 2.4. The location and number of camps, and their respective geographic circuits. Camp number one is the central camp, Msakamba.

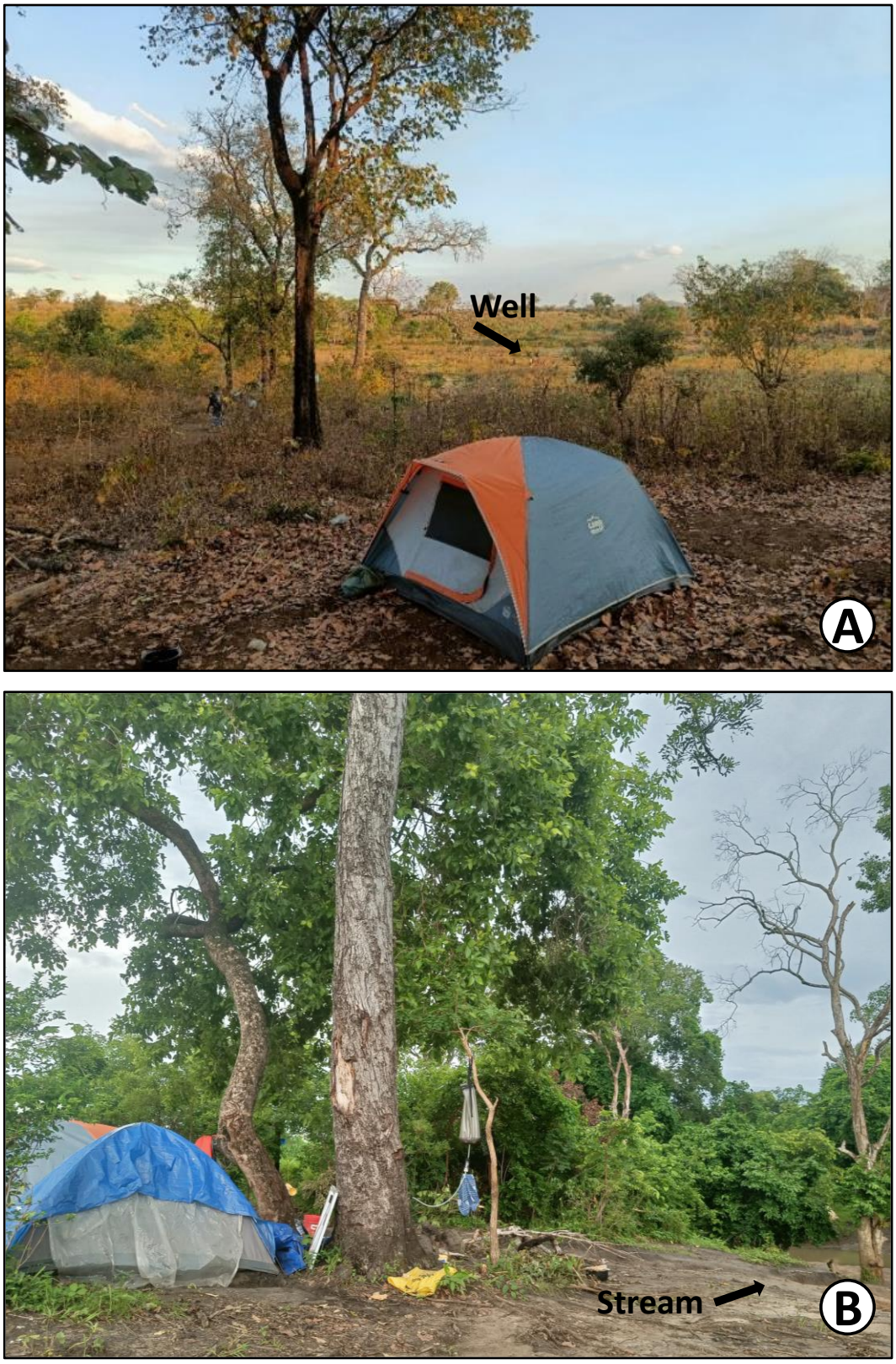

Figure 2.5. Camp number 18, *Mavimba Pori*, an example of miombo woodland that had been converted for agricultural land near the western border of ILUMA. The camp was set up close to a well during the dry season (A) and was located next to a flowing stream during the wet season (B), to fetch water for cooking and filtering drinking water.

Fig 2.6. Camp number 11, *Kisima cha Seba*, located inside ILUMA WMA (A), and camp number 3, *Bwawa la Nyati* located inside ILUMA WMA and close to the NNP border. Both camps are examples of intact, mature miombo woodland.

Figure 2.7. Examples of camp locations at the legal fishing camps inside the groundwater forest of ILUMA WMA on the south bank of the Kilombero, that practiced small-scale sustainable fishing. **A;** The biggest fishing camp inside ILUMA, *Mikeregembe*, where camp number 14 was located. **B;** Camp number 26, based next to the smaller fishing camp of *Funga*.

The original protocol planned for a total of three rounds to be completed from January to March, March to May, and August to October 2022, representing the short rainy season, long rainy season, and the dry season, respectively. It was planned that each round of surveys in the longitudinal rolling cross-sectional study design visited all 28 camps detailed in figure 2.4 and table 2.1, but in each round a few camps were omitted for pragmatic reasons such as a lack of surface water during the dry season or inaccessibility due to severe flooding during intense rains. Also, in the first round, the north-west circuit was omitted because transport, handling and rearing procedures for the collected mosquito specimens failed, and so it was decided to adjust these procedures and start afresh with round two.

However, no camp in ILUMA WMA appeared to lack signs of human disturbance, and only a few remained relatively pristine, so four new mobile camps inside NNP were added (numbers 29-32, Figure 2.8) to the study design at the end of round three in November 2022, forming a new circuit that was surveyed with vehicle support for logistical and safety reasons. These camps were located inside the boundary of NNP, immediately to the east of ILUMA WMA to capture absolute pristine environments and were accessed by vehicle via the NNP ranger post at Boma Ulanga for which that circuit was named (Figure 2.9). These were then repeated at the start of the fourth and final round of data collection which was completed from February to July 2023, representing the whole wet season and the beginning of the dry season for that calendar year. Considering the initial unexpected sibling species identification and insecticide phenotype results obtained from the first four NNP camps, it was decided to extend the sampling frame deeper into the park and adjust the field protocol to collect and immediately preserve additional collections of larvae that would not be used to rear adults in the Msakamba insectary.

Fig. 2.8. The view from camp number 31, *Zanzibar*, a fully pristine camp inside NNP away from any signs of humans or livestock.

To obtain samples as far away from human beings and as deep into pristine ecosystems as possible, another, eight camps inside NNP were added to the end of the fourth and final round. An additional five camps (*Kilombero circuit)* that were accessed by boat via the Kilombero river were visited (Figure 2.9), followed by camps 38-40 that were accessed by vehicle along main park roads via the Msolwa ranger post and were therefore named accordingly as the *Molswa circuit* (Figure 2.9). Camp number 38 marks the furthest possible aerial distance from any permanent human habitation, but it was later noted that this location was 10km from a small tourist camp (Figure 2.9). The other two camps from the Msolwa circuit (39, 40) were in a north-east direction from ILUMA WMA, and were respectively located 21.9km and 41.5km from the nearest point of the NNP border. However, the latter was close to a human settlement built within the NNP, namely the construction site for the Julius Nyerere hydroelectric power plant, formerly known as Stiegler's Gorge (Figure 2.9). The number and name of all camp locations and the number of times they were surveyed is presented in table 2.1.

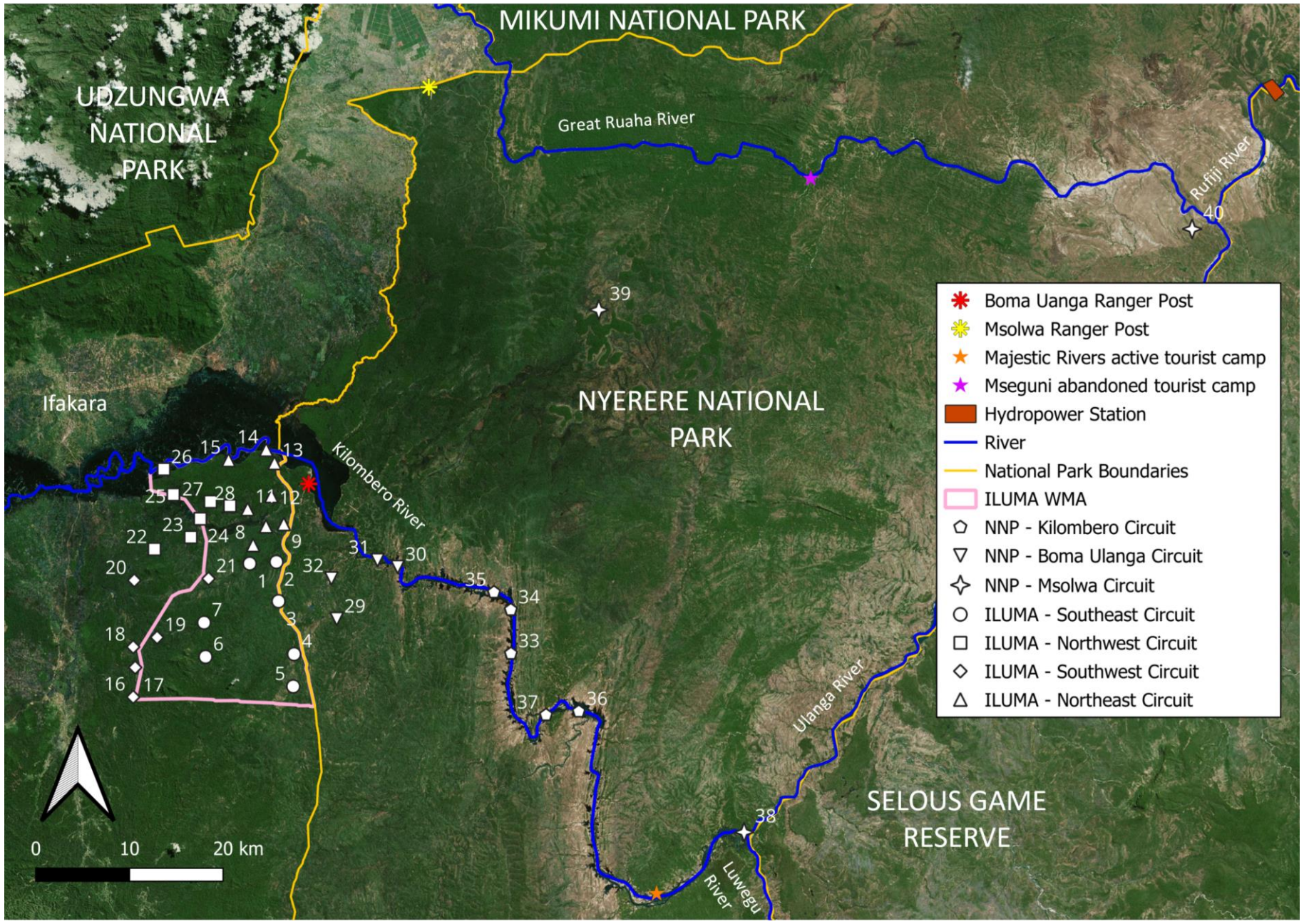

Figure 2.9. Aerial image showing the camps from the original protocol and the additional camps located deep inside the pristine environment of NNP.

Table 2.1. Each camp number and name with the corresponding circuit, location and the number of rounds they were surveyed in. The historical landcover for each camp is also given which was assigned as described in section 2.8.

Number.	*Camp name*	*Circuit*	*Location*	*Landcover*	*Round*
1	Msakamba	Southeast	Inside ILUMA	Miombo woodland	1,2,3,4
2	Bwawa la Msiba wa Deo	Southeast	Inside ILUMA	Miombo woodland	2,3,4
3	Bwawa la Nyati	Southeast	Inside ILUMA	Miombo woodland	2,3,4
4	Bwawa la Nandete	Southeast	Inside ILUMA	Miombo woodland	2,3,4
5	Korongo la Bundu	Southeast	Inside ILUMA	Miombo woodland	2,3,4
6	Bwawa la Namamba	Southeast	Inside ILUMA	Miombo woodland	2,3,4
7	Bwawa la Chakacheni	Southeast	Inside ILUMA	Miombo woodland	2,3,4
8	Bwawa la Njuju	Northeast	Inside ILUMA	Miombo woodland	1,2,3,4
9	Bwawa la Chamvi	Northeast	Inside ILUMA	Miombo woodland	1,2,3,4
10	Bwawa la Miembeni	Northeast	Inside ILUMA	Miombo woodland	2,3,4
11	Kisima cha Seba	Northeast	Inside ILUMA	Miombo woodland	1,2,4
12	Bwawa la Maya	Northeast	Inside ILUMA	Miombo woodland	1,2,3,4
13	Bwawa la Mrope	Northeast	Inside ILUMA	Groundwater forest	1,2,4
14	Mikeregembe	Northeast	Inside ILUMA	Groundwater forest	1,2,3,4
15	Mdalangwila	Northeast	Inside ILUMA	Groundwater forest	1,2,3,4
16	Bwawa la Mnyuamachi	Southwest	Outside ILUMA	Miombo woodland	1,2,3,4
17	Tuliza Moyo	Southwest	Outside ILUMA	Miombo woodland	1,2,3,4
18	Mavimba Pori	Southwest	Outside ILUMA	Miombo woodland	1,2,3,4
19	Bwawa la Selesusi	Southwest	Inside ILUMA	Miombo woodland	1,2,3,4
20	Makingi	Southwest	Outside ILUMA	Miombo woodland	1,2,3,4
21	Bwawa la Mpunga	Southwest	Inside ILUMA	Miombo woodland	1,2,3,4
22	Kisaki	Northwest	Outside ILUMA	Miombo woodland	2,3,4
23	Uwanja wa Ndege	Northwest	Outside ILUMA	Miombo woodland	2,3,4
24	Bwawa la Mkwajuni	Northwest	Inside ILUMA	Miombo woodland	2,3,4
25	Bwawa la Mamba Luhogi	Northwest	Inside ILUMA	Miombo woodland	2,3,4
26	Funga	Northwest	Inside ILUMA	Groundwater forest	2,3,4
27	Bwawa la Mlenda	Northwest	Inside ILUMA	Miombo woodland	2,3,4
28	Bwawa la Semka	Northwest	Inside ILUMA	Groundwater forest	2,3,4
29	Kambi ya Simba	Boma Ulanga	Inside NNP	Miombo woodland	3,4
30	Bwawa la Kiboko Zanzibar	Boma Ulanga	Inside NNP	Acacia savanna	3,4
31	Zanzibar	Boma Ulanga	Inside NNP	Miombo woodland	3,4
32	Bwawa la Moto	Boma Ulanga	Inside NNP	Miombo woodland	3,4
33	Kambi ya Mamba	Kilombero	Inside NNP	Acacia savanna	4
34	Kambi ya Machuma	Kilombero	Inside NNP	Acacia savanna	4
35	Serengeti Ndogo	Kilombero	Inside NNP	Acacia savanna	4
36	Kambi ya Makutano	Kilombero	Inside NNP	Acacia savanna	4
37	Kambi ya Mawe	Kilombero	Inside NNP	Acacia savanna	4
38	Shughuli Kubwa	Msolwa	Inside NNP	Miombo woodland	4
39	Bwawa la Chatu	Msolwa	Inside NNP	Miombo woodland	4
40	Bwawa la Umeme	Msolwa	Inside NNP	Acacia savanna	4

2.3 Day-to-day field protocol and sampling methodology of trapping live adult mosquitoes.

The team of three investigators responsible for collecting data and mosquito specimens were accompanied by eight VGS, whom in addition to providing food, water, shelter and security, were also integral to the implementation of the field protocol. Walking from one camp to the next was only possible under the guidance of at least two armed VGS members, whom also assisted with tracking for the surveys of humans, livestock and wild animals. Field procedures conducted inside NNP were also completed under the oversight and guidance of armed park rangers. Other VGS members assisted with carrying equipment and supplies, as well as the routine logistical and scientific activities such as setting up camp, cooking, filtering water for drinking, and collecting and maintaining mosquitoes before they were transported back to the central camp insectary at Msakamba. The field protocol was driven by the primary objective for the broader study that was to find insecticide susceptible mosquitoes, so all field activities, including the larval surveys, and the land use and mammalian activity surveys reported herein were centred around the daily collection of live adult mosquitoes.

After arriving at a new camp location, a shaded area next to a water source was scouted to identify a suitable place for setting up the tents, cooking, resting and processing mosquitoes. The batteries for four Centres of Disease and Control (CDC) light traps that were used for trapping live mosquitoes were then charged as soon as possible using portable solar panels. On overcast days, especially during the rainy season, charged batteries were received from the Msakamba central camp every two days by the team responsible for taking the collected mosquito samples. A site within a valley or open natural glade, at a minimum distance of 50m from the camp, was identified and a five-panel 25m long interception screen made from standard mosquito netting was set up as shown in figure 2.10. After hiking to and setting up camp, the afternoon was used to rest before starting the collection of live adult mosquitoes from dusk through to just before midnight.

Figure 2.10. The set up of the interception screen trap for mosquitoes (Burkot *et al*., 2013) in a typical open natural glade surrounded by dense woodland.

Live mosquitoes were trapped using two methods: an interception screen (Burkot *et al.*, 2013) (Figure 2.10) and CDC light traps (manufactured by John W. Hock Company, product number 512). Collections at the interception screen were carried out five times after dusk for every hour between 19:30 and 23:30, and once before dawn at 05:30, specifically to target the known peak outdoor biting activity of adult *An. arabiensis* in the Kilombero valley (Russell *et al.*, 2011; Maia *et al.*, 2016). Both sides of the screen were scanned in succession, using a torch to carefully inspect each panel in an up-and-down motion for resting mosquitoes. Once identified, the mosquito was suctioned into a collection cup using a custom-made, standard *Prokopack*, which is a handheld aspirator powered with a 12V GWESPECS battery (Vazquez-Prokopec *et al.*, 2009). Separate collection cups were used for every hour and were immediately labelled with the date and time of sampling. A small ball of cotton wool was soaked into a 10% w/v glucose solution and placed on top of the collection cup as a food source before storing the cup in a metal frame within a self-cooling, self-humidifying ventilated backpack designed specifically for the purpose (Kavishe *et al.*, Unpublished). The backpack was suspended from a tree on the end of a length of string to keep the samples off the ground and to avoid predation by ants, and at a minimum of 20m from the fire to avoid mortality caused by smoke. Additionally, four CDC light traps ran overnight between 19:00 and 07:00, each in a different

location to capture any potential variability in host-seeking behaviour or other dispersal patterns. A trap was set up in one of four distinct ways: One suspended in an open valley or natural glade, one next to a stream, one in the camp site, and one next to an occupied tent. Again, the traps inside the camp and next to the tent were placed as far away from the campfire as possible.

The following morning at 07:00, *Anopheles* mosquitoes from the interception screen and the CDC light traps were transferred into clean, labelled paper cups using a mouth aspirator and supplied with fresh glucose solution soaked in cotton wool (Figure 2.11).

Figure 2.11. The VGS transferring *Anopheles* adults from the collection cups into labelled paper cups using a mouth aspirator.

An experimental design form template (ED1) from a standardised mosquito collection informatics platform (Kiware *et al.*, 2016) was completed to include variables such as *date, camp number, collection method, habitat type, start time, finish time, experiment day, volunteer initials* and a column for any further comments. Samples in the paper cups were labelled with the date and time, and coded with the ED1 serial number and form row, as well as the variable habitat type, for sample processing and tracing in the field insectary at Msakamba. The cups with the *Anopheles* mosquitoes were then placed back in the specially designed backpack, which was soaked in water and also had two soaked towels placed inside it to keep the

temperature and humidity optimal for mosquito survival (Kavishe *et al*., Unpublished) (Figure 2.12).

Figure 2.12. Prepared samples of *Anopheles* mosquitoes from the interception screen and CDC light traps in the customized metal frame and backpack (Kavishe *et al.*, Unpublished), before being transported back to the field insectary at Msakamba.

The backpack was then resuspended from a tree branch that provided sufficient shade and was regularly soaked with water throughout the day to ensure continued optimal environmental conditions for the mosquitoes. After the mosquitoes were prepared on the first morning of a new camp location, larval surveys and associated surveys of land use and humans, livestock, wildlife and land use were completed between 08:00 and 12:00 noon, and the afternoon was then used to rest and enter the field data into Microsoft Excel® on a ruggedized laptop (DELL Latitude 5420 Rugged) before beginning the next mosquito collection starting at 19:30 after the CDC light traps were switched on. On the second morning, the field team would break camp and hike with all the equipment to the following location after two VGS arriving from the Msakamba central camp delivered supplies and collected the backpack of mosquito samples to be returned to the field insectary for sorting, counting, and rearing. After arriving at the new camp location, the routine was repeated as described above.

2.4 Larval survey methodology.

Larval surveys are typically undertaken as a routine surveillance method for guiding and monitoring the effectiveness of larval source management operations, or investigating the environmental and biological factors underpinning larval habitat receptivity and productivity (O'Malley, 1989). In this study, larval occupancy surveys were carried out as a methodological approach to illustrate the distribution of *An. gambiae* complex populations as indicated by how well aquatic sites were colonised across long distances that encompass different land uses. Occupancy surveys are also advantageous as they identify the source of populations as opposed to catching adult mosquitoes where their life history and flight distances are unknown.

Quantitative larval surveys were carried out within a 2km radius of each camp by characterising aquatic habitats, followed by dipping (Service, 1993) and identifying larvae for a maximum of four hours. The upper limit of 2km from the camp location was decided upon to prevent geographic overlapping between neighbouring camp locations and to minimise spatial autocorrelation effects. Surveys were intensive and conducted in a challenging climate, where habitats were often in open, unshaded areas, so a time limit of four hours was decided upon to mitigate against investigator fatigue and ensure optimal data collection for the full duration of an entire survey circuit. This upper limit of survey duration was based on a prior experience of the overall project and an initial pilot of the field procedures, both of which indicated that longer surveys would prove unsustainable, with potential deleterious consequences for both data quality and the surveyor's wellbeing, especially given the extreme heat in the afternoon. Starting at the camp, surveys were initiated at the nearest waterbodies known to the VGS, and any other potential aquatic habitat that were seen along the way were also surveyed. Although this purposive sampling was not a randomised approach, it proved practically feasible and maximised the number of habitats sampled around each camp in an efficient manner over the limited time frame of one morning. If the 2km radius was reached before the time limit was up,

a different route back to the camp was taken and any potential habitats that were identified on the way were surveyed. All the surveys in the fourth round, which accounted for all the data collected in 2023, was completed by a trained VGS rather than the investigator herself.

2.4.1 Identifying and characterising potential habitats.

Larval occupancy surveys were completed using an adapted version of a standardised form previously used for larval habitat surveys in Dar es Salaam, Tanzania (Fillinger *et al.*, 2008). The form contained details on the camp location including *camp number*, *name*, *GPS coordinates*, and a general description of the surrounding environment which were filled out at the camp before beginning the survey (Appendix 1). Each row on the form was filled out for one potential habitat only. After identifying a potential habitat and prior to dipping, the Global Positioning System (GPS) coordinates were taken using a handheld Garmin® Etrex 10 device. Habitat types were identified by assigning a numeric code from 1 to 12 that corresponded to twelve broad categories of water bodies (Appendix 1). Some examples of the various habitat types identified during surveys can be seen in figure 2.13. Further information including a more detailed description of the potential habitat and the presence or absence of *short vegetation*, *tall vegetation* and *floating vegetation* were recorded by ticking the appropriate column in the row (Appendix 1). Visual estimates for *water depth* and *perimeter* were taken and assigned to their corresponding categories indicated on the form (Appendix 1). Once the potential habitat was characterised, dips were taken to look for the presence and absence of larvae.

Figure 2.13. Examples of common larval habitat types that were found during the study, A; human-dug well, B; rice paddy, C; pool along the bottom of a streambed, D; waterhole, E; swampy area, F; animal print (hippo).

2.4.2 Dipping and field identification.

A sample of water, informally referred to in the field as a *dip*, was taken by briefly submerging a standard white 350ml dipper just below the water surface into the habitat at 45-degree angle so that the suction created by water displacement draws both water and larvae into the dipper cup (Figure 2.14), even where obstructing vegetation would frustrate sampling with conventional limnological sampling tools like sweep nets. The dipper was then removed immediately, to prevent captured larvae from escaping.

Fig. 2.14. Example of the standard dipping technique to sample larvae (Service, 1993) (A) and inspecting the dipper for anopheline and culicine larvae (B).

All dips were taken along the waterbody perimeter and were methodically numbered and spaced according to estimates of the habitat perimeter and criteria defined as follows. If a potential habitat was <2m in perimeter (Appendix 1), one or two dips were taken depending on size and feasibility. For human, livestock and wild animal prints or any other habitat that was too small to be effectively dipped, a standard 28ml turkey baster was used to suction water and larvae from the habitat into the dipping cup. These habitats were classified as having one dip. For medium sized waterbodies categorised between 2m and 20m, as well as two larger waterbody perimeter categories greater than 20m, and 200m, respectively (Appendix 1), the spacing between each dip was estimated by the number of paces taken by the individual carrying out the survey to ensure dips were taken at regular intervals around the full perimeter.

Aquatic habitats that were categorised as being approximately 2m to 20m in perimeter, were dipped for a total maximum of five times. Therefore, if the habitat was about 10m in perimeter or less, dips were taken once every two paces walked around the perimeter. If the habitat was estimated to be between 10m and 20m, dips were taken once every four paces. For larger waterbodies between 20m and 200m, and those that were greater than 200m, a maximum of 10 dips were taken once every 20 paces and every 50 paces, respectively. These criteria were developed to maximise sampling efficiency, consistency, and productivity after a pilot survey. Even if *An. gambiae* complex larvae were identified immediately, dips were continually taken around the full perimeter or until the maximum number of dips was reached, because as many of these larvae as possible were collected, kept alive, and sent back to Msakamba to be reared for insecticide susceptibility testing.

Dips were inspected for the presence or absence of early and/or late instar anopheline and culicine larvae, as well as the presence or absence of pupal stages (Figure 2.14). Anopheline and culicine larvae were identified in the dipper based on clear characteristic and morphological differences that do not require magnification; anopheline larvae lie parallel to the water surface and do not have a siphon, whereas culicines hang from the water surface at an angle and have a clear, elongated siphon (Rozendaal, 1997). Further identification of pupal stages was not completed as differentiating the genera was not possible under field conditions described herein. Any anopheline individuals were further inspected using a 10x magnified hand lens to distinguish *An. gambiae* complex larvae from other anopheline species, based on a characteristic white collar immediately behind the head (Holstein, 1954) (Figure 2.15). Once completed, the presence or absence of *An. gambiae* complex larvae and larvae from other taxa were recorded in the form (Appendix 1). All individuals identified as *An. gambiae* complex larvae that were collected, were kept alive and sent to Msakamba for rearing and insecticide susceptibility testing as described in the following subsection.

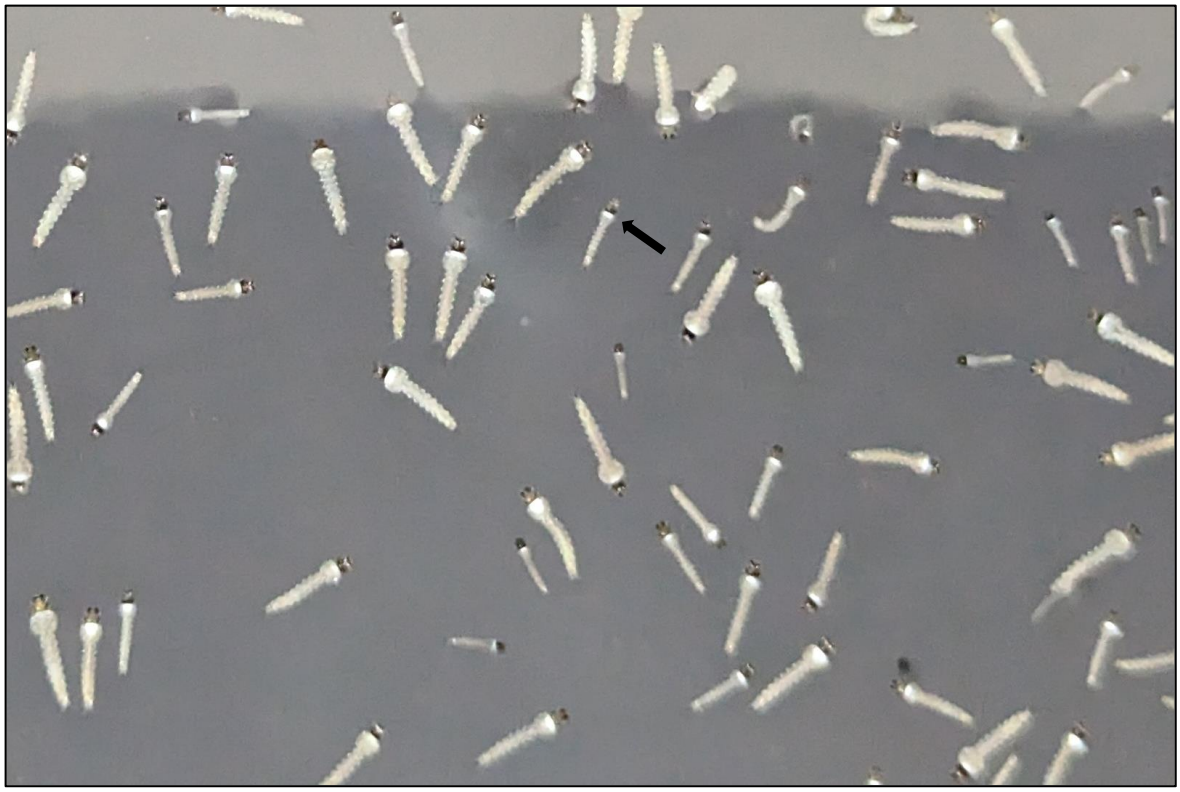

Figure 2.15. An image of *An. gambiae* complex larvae under the microscope as identified by a white collar immediately behind the head, which is a distinguishing morphologic trait of this species complex (Holstein, 1954).

2.4.3 Collection of larvae from aquatic habitats.

For each dip containing anopheline larvae identified as *An. gambiae* complex, individuals were taken from the dipper using a clean, disposable, 3ml plastic pipette and placed inside a 500ml capacity water bottle, filled with approximately 350ml of water that was freshly filtered *in situ* at the camp prior to starting the survey. The purpose of filling the bottle approximately three quarters full was to avoid sloshing that could cause mortality or reduce fitness while providing sufficient oxygen. Approximately 1mm wide air holes were punctured in the bottle caps to further supply fresh oxygen. To make the process of transporting larvae back to Msakamba logistically manageable, larvae from several habitats that were similar in appearance and close together were pooled into one bottle, reducing the number of bottles that had to be carried on foot. Habitats that appeared to have exceptionally high numbers of larvae, were collected as a separate *batch* sample so that one bottle contained larvae from one highly productive habitat

only. Each sample was labelled with the form serial number and form row (used as a primary key for sample tracing) and a unique sample label code (secondary key), along with the sample type (1; an individual larva, 2; pooled sample, i.e., larvae from multiple habitats, 3; batch i.e., from the same habitat). Once the survey was complete and the team returned to the camp, the water in each bottle was exchanged for newly filtered water and larvae were fed using finely ground TetraMin® fish food. Replacing water from their natural habitat with filtered drinking water was intended to minimise mortality from natural ammonia accumulation (Akpodiete and Tripet, 2021a; Akpodiete and Tripet, 2021b). Bottles were placed in the shade to avoid overheating from the sun until the following morning, when each sample was provided with fresh water and fish food again and transported back to Msakamba for sample processing in either the customised backpack with the adult mosquitoes as described earlier (Figure 2.12), or suspended on the end of string handles to prevent splashing which was used later during the study (Figure 2.16).

Figure 2.16. Hanging bottle design used later in the study to transport mosquitoes by foot across long distances (Kavishe *et al.*, Unpublished). This design prevented splashing and improved survival rates compared to using the backpack designed for transporting adults.

All field-identified *An. gambiae* complex larvae from January to November 2022 were reared to adults (known as the F0 generation), which were then fed and propagated through one or two further generations (F1 and F2, respectively), before they were used for insecticide susceptibility testing. Adult *An. gambiae* complex mosquitoes that were caught in the wild, but reared in captivity could only be identified by polymerase-chain reaction (PCR) analysis (Scott, Brogdon and Collins, 1993) in the Ifakara Health Institute laboratory in Ifakara, after these F0 adults had either already yielded eggs or died. These were used for preliminary analysis only (Section 3.4.1) and were not included in the final analyses (Sections 3.4.3 and 3.4.4) for several reasons. Following the observation that the *An. gambiae* population at many camps consisted of a mixture of *An. arabiensis* and *An. quadriannulatus*, it was decided that PCR tests (Scott, Brogdon and Collins, 1993) on the F0 adult mosquitoes cannot give robust representations of sibling species composition. This is because the specimens were handled multiple times during the collection, transportation and rearing phases that could easily lead to survival bias towards the fittest of those two competing species under such shared, artificial environmental conditions. Another limitation was that all anopheline species that were present in an aquatic habitat during the surveys were collected in the same sample and were retained and transported back to the Msakamba insectary in the same bottles, irrespective of whether they had been identified as *An. gambiae* complex or not. Once the samples were returned to the field insectary for the rearing phase, samples of anopheline larvae were then visually re-identified in the field insectary in Msakamba using the same methodological approach and anything not identified as *An. gambiae* complex was removed from the sample during the rearing phase. This procedural flow may have exacerbated potential survival biases arising from overcrowding and predation (Koenraadt and Takken, 2003; Koenraadt *et al.*, 2004).

It was therefore decided to add a supplementary procedure for collecting more larvae and preserving them *in situ* to address these questions about how species composition varied with

location, ecosystem integrity and potential host availability. The purpose of standardising and limiting the number of dips in the primary larval collection protocol was to strike a compromise between the amount of time spent at each waterbody and the number of different habitats that could be surveyed for assessing how occupancy varied across the ecological gradient of the study site. As it was essential to maintain the existing survey protocol without changing it, so that data collection to address the original questions about occupancy and insecticide resistance phenotypes were consistent and comparable, it was decided to supplement those collections of live larvae with more purposive and intensive collections from the same aquatic habitats to obtain larger, more carefully disaggregated samples and preserve them *in situ* for robust PCR analysis (Scott, Brogdon and Collins, 1993).

When a habitat was positively identified for the presence of *An. gambiae* complex, the first individual conducting the occupancy survey continued to collected larvae as per the original protocol that involved the live transportation to the field insectary as described above using the hanging bottle design (Figure 2.16). However, during this final round, separate samples of larvae were collected in parallel by a second trained VGS following immediately behind the first surveyor. This second surveyor collected larvae from aquatic habitats that were positively identified during the survey using the same dipping technique but far more exhaustively. Field identified *An. gambiae* complex only were immediately preserved into a 50ml Falcon® tube filled with ethanol using a clean, disposable plastic pipette. Directly preserving larvae right at the aquatic habitat at which they were collected from, eliminated any potential bias or human error that could distort the sibling species composition during the transport and rearing processes and therefore, ensured robust formal species identification. This approach also increased the sample sizes to be used for subsequent analyses as the only focus of the second VGS was to identify and preserve as many *An. gambiae* complex larvae as possible from each aquatic habitat found to contain them. Also, given the specific interest in sibling species

composition, and likelihood of strong covariance of individual dip samples with habitats, combined with substantial variance among habitats, it was decided to keep separate, specific samples for each individual habitat, so that anticipated within habitat covariance and between-habitat variance could be accounted for in the statistical analysis as nested random effects.

Unlike the previous surveys, over rounds one to three in 2022, wherein field-identified *An. gambiae* complex larvae from multiple habitats were often pooled together, the larvae from different habitats were all stored as separate batch samples so that each tube contained preserved larvae from a singular aquatic site. A target of ten batch samples per camp was set to enable robust quantification and statistical evaluation of population composition heterogeneity around each camp location. Each habitat-specific batch sample was labelled with the camp number, the serial number and form-row number on the larval surveillance form and given a *breeding site identification number (BSID)* from 1 to 10, so that each batch sample could be traceable to a particular aquatic habitat from a particular camp. These batch samples of preserved field identified *An. gambiae* complex larvae were returned to Msakamba and were transported to the laboratory in Ifakara for PCR analysis (Scott, Brogdon and Collins, 1993).

These more carefully separated and clearly distinguished samples were also used to assess how heterogeneities in sibling species composition within the *An. gambiae* complex may vary with the availability of different mammalian species as potential blood sources for the adult mosquitoes that oviposited in those habitats, just as demonstrated for mixed populations of *An. arabiensis* and *An. gambiae s.s.* in fully domesticated environments elsewhere (Charlwood and Edoh, 1996; Minakawa, Seda and Yan, 2002). To perform these associations during data analysis, radial surveys of humans, livestock, wildlife activity and land use were conducted as follows around each camp in parallel with the larval surveys (Duggan *et al.*, Unpublished).

2.5 Surveys of mammalian activity for assessments of host species association.

Each larval occupancy survey was accompanied by a radial survey of human, livestock and wildlife activity and land use, that was completed by one or two other investigators using a specifically developed key (Duggan *et al*., Unpublished) (Appendix 2). The surveys recorded all possible evidence of human activity and natural resource land use, as well as various forms of evidence for the presence and abundance of humans, livestock and 45 different species of wild animals, and estimates for the overall land use along the survey route (Duggan *et al.*, Unpublished). Observation categories of humans and wildlife included direct sightings, tracks, spoor, and other signs such as sounds, indicators of species-specific behaviours and dwellings (Appendix 2). These observations were further classed as *active* when the track, spoor or other sign was produced within the last 24 hours, *recent* when the track, spoor or other sign was produced within 2 to 5 days, and *old* when they were estimated to be 6 days to 6 months old (Appendix 2). When an entity or activity was observed it was immediately recorded as described in Duggan *et al*., (Unpublished). These radial surveys of mammalian host availability and land use were completed in parallel with the larval occupancy surveys by following the same route as the investigator(s) implementing the larval survey, and recording detections of humans, livestock and wild animals around the perimeter of the habitat that was being dipped (Figure 2.17).

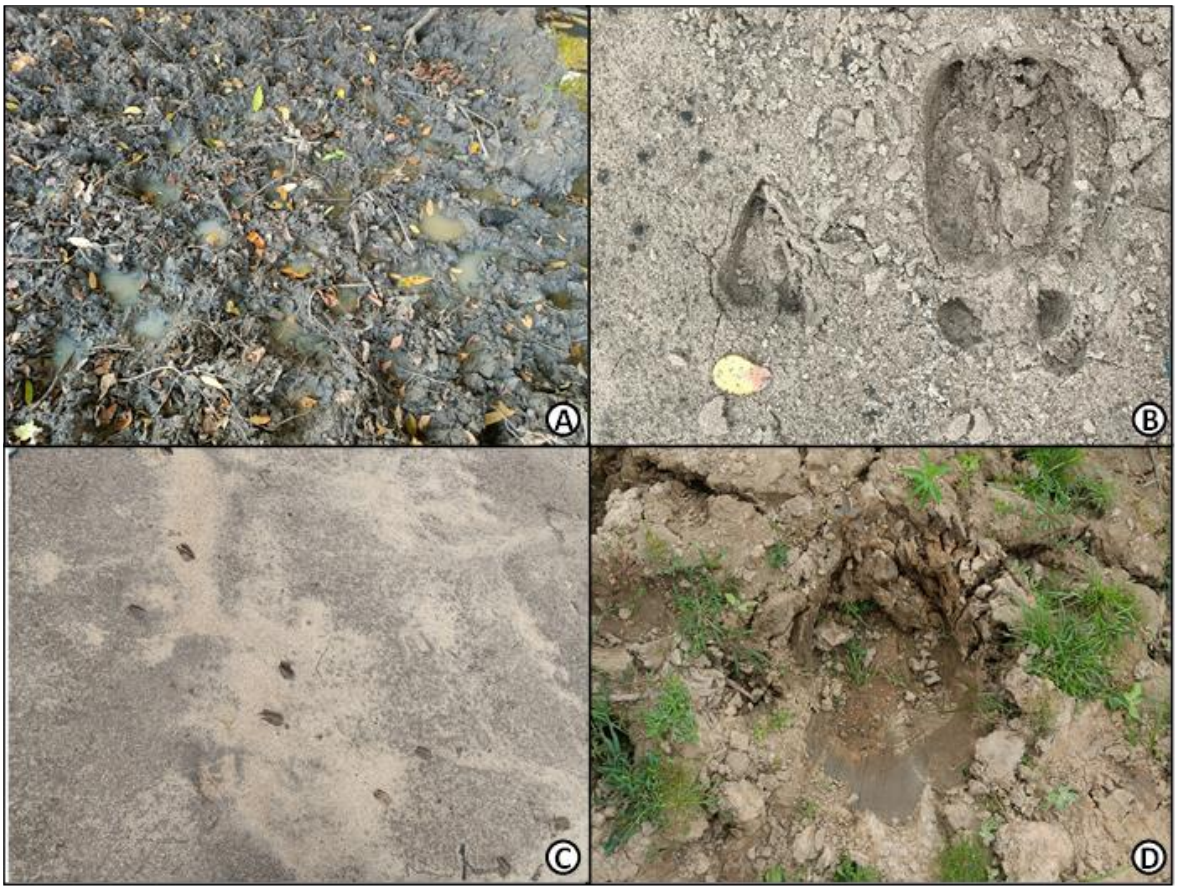

Figure 2.17. Examples of animal prints that were recorded as detections during the radial surveys of human, livestock, wildlife activity and land use. A; collection of cattle hoofprints surrounding a pool, B; hartebeest and a buffalo print in sand, C; common duiker prints in sand, D; a hippo print on the riverbank.

The former surveys also continued recording signs of potential blood host activity in between potential larval habitats to capture any additional detections that would be representative of the environment around the camp at which the larval survey was being conducted. These radial surveys of potential blood hosts produced a large dataset of detections with most relating to tracks and spoor rather than direct sightings or other signs. As only the PCR-identified larval samples that were collected and preserved *in situ* under the revised protocol during the fourth round were used for the final analyses of sibling species composition of the *An. gambiae* complex (presented in sections 3.4.3 and 3.4.4), only the fourth round of complimentary radial surveys, completed at exactly the same times and places, were used to test for associations between the species composition and the availability of various potential hosts. Hence, any

detection from the radial surveys that complemented the larval survey data from rounds 1,2 and 3, respectively, were omitted from analysis for this particular purpose. Therefore, any larva from an aquatic habitat could be spatially and temporally associated with the frequency and types of human, livestock or wildlife detections that were recorded around the aquatic habitats that the larvae were collected from. Recorded detections from every survey over all four rounds were, however, used to calculate a formal and objective *natural ecosystem integrity index* that was used to evaluate the validity of a more subjective, but also more powerful and less laborious indicator of ecosystem integrity that was formulated by the investigators, which proved to be an invaluable predictor of larval occupancy (Section 3.3) and *An. gambiae* complex sibling species composition (Section 3.4.3), and is described as follows.

2.6 Quantifying natural ecosystem integrity.

2.6.1 The Subjective Natural Ecosystem Integrity Index.

Following the initial year of data collection and prior to any analyses, a *subjective natural ecosystem integrity index* (SNEII) was devised, as a simple and intuitive means of assessing the ecological intactness of a location relative to its pristine, natural state and the level of degradation, if any, that had occurred there. This index was created by assigning each camp a score from 0% to 100%. An SNEII score of 0% represents a fully domesticated landscape, where human settlement has long been established and the ancillary activities that are associated with permanent settlements and unrestricted land use, such as land clearing, agriculture and extractive practices like charcoal burning are ubiquitous. At the other extreme end of the scale, a score of 100% represents a fully pristine, natural ecosystem, with no evidence of human activity of any kind. Individual scores were carefully assigned based on a consensus among the three investigators of recollected impressions of the landcover and land use patterns. Field notes that were recorded and any photographs that were taken during the study were also considered in discussions to agree on consensus scores. An initial SNEII for

camp numbers 1 to 32 was formulated in December 2022, after data collection had been completed for that calendar year. As 8 extra camps were added for the final round of data collection in 2023, the SNEII was subsequently edited to accommodate these additional 8 camps further inside NNP that were visited once. The scores were initially drafted and then finalised in Microsoft Excel® after further detailed conversations reached a conclusive consensus.

The primary and most important criterion considered by the three investigators when assigning a SNEII score was the perceived intensity of land degradation caused by the following: (1) the conversion of land for agriculture, (2) deforestation for the purpose of charcoal production, timber harvesting or human settlement (3) livestock herding. The intensity of degradation of the land surrounding a 2km radius of the camp was initially determined by estimating the proportion of remaining intact ecosystems using personal observations that were made while present in the camp and during the larval habitat surveys (Section 2.4) and radial surveys of human, livestock, wildlife activity and land use (Section 2.5). The estimated proportion of intact land was then used as a reference starting point for assigning the score based on other considerations. For example, a camp with an estimated 10% of intact natural landcover was initially given an approximate score of 10 before that was considered for adjustment based on other factors as described below. Camps that evidently belonged on the extreme ends of the scale were assigned first before those that were less clear and required a more in-depth discussion. Most camps that required extra discussion were areas where intermittent patches of deforestation, small-scale agriculture and human settlement occurred in between expanses of intact woodland. These were all located inside the ILUMA WMA, in areas that were either regenerating from previous encroachment, or beginning to experience an influx of human activities.

The perceived presence of wildlife and humans as indicators of ecosystem integrity were considered to be of secondary importance when assigning scores, but were nevertheless useful deliberations following the agreement of a consensus score with respect to the intensity of land degradation. Personal observations clearly indicated that species diversity generally increased in more intact natural ecosystems, but that some wild species were found more often than others in areas of degraded land. These criteria were useful when deciphering the scores of multiple camps within ILUMA WMA that had similar proportions of intact woodland and land degradation. It was also useful for areas of regenerating woodland that had been previously encroached upon and degraded by humans. Regenerated woodland with observed signs of high species diversity, influenced the investigators to assign a higher SNEII score, whereas woodlands in the early stages of regeneration that lacked obvious signs of wildlife diversity were instead adjusted downward. However, using wildlife as an indicator of ecosystem integrity often proved troublesome because it also required careful consideration of the fact that wild animal communities change naturally with land cover type, season, and even short-term weather conditions. The ability to distinguish land cover types was also a notable factor when assigning a value for this index as it was important to not mistake natural grasslands, scrublands, and valleys for deforestation.

From personal observations, it seemed to the investigators that the presence of humans was negatively associated with ecosystem integrity. However, it was not a consistently reliable criterion as some human activities including the practice of legal, sustainable fishing inside the WMA, had little perceptible impact on the surrounding natural environment. Although meat and fish poaching were not considered to have any direct effect on ecosystem integrity by the investigators, any evidence of these illegal activities prompted the investigators to assign a slightly lower final score depending on their perceived intensity. After the initial discussion was completed and the initial SNEII scores were drafted, these values were discussed again to

ensure that the order and distribution of the final SNEII scores were assigned as well as possible regarding the above criteria. The overall outcomes of this subjective ecosystem scoring processes are detailed in table 2.2 and narrated as follows.

All final scores of less than or equal to 10% were given to camps that were based at permanent human settlements outside ILUMA WMA (Table 2.2) and surrounded by domesticated land which was primarily used for rice farming and other tillage agriculture. Camps that were given SNEII scores of 15%, 25%, 30%, 35%, 45%, 48%, respectively, were all located inside the ILUMA boundary (Table 2.2) but were heavily encroached and included large areas of degraded land for agriculture and human settlement. A decrease in wildlife diversity was also seen around these camps based on recollected visual observations. Only two camps were assigned values between 50% and 70%, both of which were located inside ILUMA WMA (Table 2.2). *Bwawa la Mpunga* (camp number 21), was assigned a score of 55% and was based at a waterhole and surrounding swamp that had been exploited for rice farming. Despite this, the surrounding woodland was not as heavily degraded as other locations on the western boundary of ILUMA WMA. Similarly, *Bwawa la Miembeni* (camp number 10), was located in a large valley that was also exploited for rice farming. However, the surrounding woodland was fully intact and therefore received a higher score compared to *Bwawa la Mpunga* (Table 2.2). Although the location of the main camp Msakamba was surrounded by signs of land degradation and deforestation, it was assigned a score of 70% as progressive regeneration of the woodland and increased presence of wildlife was visually observed during the study period (Table 2.2).

All other camps with higher scores between 70% and 90% were in the groundwater forest in the north of ILUMA WMA or in the miombo woodlands that were located closer to the NNP border. Three small, legal fishing camps in the north of ILUMA WMA, on the southern bank of the Kilombero River (camp numbers 14,15 and 26), as seen in figure 2.7, were assigned high

SNEII scores of 75%, 72%, and 77%, respectively, despite being human settlements with permanent structures. This is because small-scale sustainable fishing and very limited extractive activities are permitted in these fishing camps, but are carefully monitored. Also, the vigilance of the residents, who often actively support the conservation activities of the ILUMA WMA, acts as a deterrent for unauthorised human activity in the areas surrounding their settlement (Duggan *et al.*, Unpublished). Correspondingly, the forests and wetlands that surround these authorised human settlements remain fully intact (Figure 2.7). Two camps inside NNP were given a score of 95% as they had still the remnants of old tourist camps. Six camps in NNP had the highest possible score of 100%, in which habitats were absolute pristine and did not have a single observed sign of poaching activity.

The lowest score inside NNP was assigned to one camp located 42km inside the park boundary. A score of 70% was given to camp number 40, *Bwawa la Umeme* because the area around this camp was recently cleared in relation to the Julius Nyerere Hydropower Station located on the Rufiji River (Figure 2.9). Despite this evidence of human intervention, large proportions of the area around the camp were in good condition and personal observations indicated that wildlife was abundant and diverse, albeit less so than in any other NNP camp.

Table 2.2. The values and corresponding ranks for the subjective natural ecosystem integrity index (SNEII) based on a consensus of recollected impressions and the objective natural ecosystem integrity index (ONEII) based on principal component analysis (PCA) of natural resource use, direct sightings and signs of tracks, spoor and other signs for wildlife, livestock and humans, together with every recorded estimate of the proportion of land used for rice farming and other tillage crops (Section 2.6.2).

			Subjective Ecosystem Integrity Index		Objective Ecosystem Integrity Index	
Camp	Camp Name	Location	Value [a]	Rank [b]	Value [c]	Rank [d]
1	Msakamba	Inside ILUMA	70%	16.5	-0.301	17
2	Bwawa la Msiba wa Deo	Inside ILUMA	84%	24.0	-0.361	28
3	Bwawa la Nyati	Inside ILUMA	85%	25.5	-0.323	22
4	Bwawa la Nandete	Inside ILUMA	35%	10.0	-0.283	14
5	Korongo la Bundu	Inside ILUMA	30%	9.0	-0.279	13
6	Bwawa la Namamba	Inside ILUMA	40%	11.0	-0.266	12
7	Bwawa la Chakacheni	Inside ILUMA	15%	7.0	-0.255	10
8	Bwawa la Njuju	Inside ILUMA	81%	21.0	-0.325	24
9	Bwawa la Chamvi	Inside ILUMA	86%	27.0	-0.353	26
10	Bwawa la Miembeni	Inside ILUMA	68%	15.0	-0.397	30
11	Kisima cha Seba	Inside ILUMA	82%	22.0	-0.293	16
12	Bwawa la Maya	Inside ILUMA	87%	28.0	-0.359	27
13	Bwawa la Mrope	Inside ILUMA	90%	29.0	-0.325	23
14	Mikeregembe	Inside ILUMA	75%	19.0	-0.311	20
15	Mdalangwila	Inside ILUMA	72%	18.0	-0.310	19
16	Bwawa la Mnyuamachi	Outside ILUMA	6%	5.0	1.185	5
17	Tuliza Moyo	Outside ILUMA	4%	3.0	1.219	4
18	Mavimba Pori	Outside ILUMA	10%	6.0	0.158	7
19	Bwawa la Selesusi	Inside ILUMA	25%	8.0	-0.253	9
20	Makingi	Outside ILUMA	0%	1.0	1.286	3
21	Bwawa la Mpunga	Inside ILUMA	55%	14.0	-0.263	11
22	Kisaki	Outside ILUMA	5%	4.0	4.722	1
23	Uwanja wa Ndege	Outside ILUMA	3%	2.0	2.750	2
24	Bwawa la Mkwajuni	Inside ILUMA	45%	12.0	-0.138	8
25	Bwawa la Mamba Luhogi	Inside ILUMA	48%	13.0	0.391	6
26	Funga	Inside ILUMA	77%	20.0	-0.287	15
27	Bwawa la Mlenda	Inside ILUMA	83%	23.0	-0.306	18
28	Bwawa la Semka	Inside ILUMA	85%	25.5	-0.382	29
29	Kambi ya Simba	Inside NNP	97%	32.0	-0.478	36
30	Bwawa la Kiboko Zanzibar	Inside NNP	100%	37.5	-0.520	40
31	Zanzibar	Inside NNP	95%	30.5	-0.458	35
32	Bwawa la Moto	Inside NNP	100%	37.5	-0.513	39
33	Kambi ya Mamba	Inside NNP	98%	33.0	-0.415	32
34	Kambi ya Machuma	Inside NNP	100%	37.5	-0.500	37
35	Serengeti Ndogo	Inside NNP	95%	30.5	-0.451	34
36	Kambi ya Makutano	Inside NNP	100%	37.5	-0.446	33
37	Kambi ya Mawe	Inside NNP	100%	37.5	-0.504	38
38	Shughuli Kubwa	Inside NNP	100%	37.5	-0.339	25

39	Bwawa la Chatu	Inside NNP	99%	34.0	-0.399	31
40	Bwawa la Umeme	Inside NNP	70%	16.50	-0.315	21

[a] The subjective natural ecosystem integrity score.
[b] The rank of each camp based on the subjective ecosystem integrity scores, where the lowest rank represents fully degraded, domesticated land and the highest rank represents an absolute pristine ecosystem.
[c] PC1 values that were derived from a PCA accounting for all recorded detections of humans, livestock and wildlife, and land use activities that were standardised using z-scores.
[d] The rank of each camp based on standardised PC1 values, where the lowest rank represents the most degraded camp and the highest rank represents the most pristine camp.

2.6.2 Validation of the Subjective Natural Ecosystem Integrity Index.

To assess the validity of this intuitive approach to quantifying ecosystem integrity, this subjective index was tested for correlation with an alternative, *objective natural ecosystem integrity index* (ONEII), which used a far more laborious approach based on statistical syntheses of the data from the radial surveys of human, livestock and wildlife activity and land use (Section 2.5). A principal component analysis (PCA) of every recorded detection of natural resource use, direct sightings and signs of tracks, spoor and other signs for wildlife, livestock and humans, together with every recorded estimate of the proportion of land used for rice farming and other tillage crops (Appendix 2) was conducted using the *prcomp* function in open software R version 4.3.1. Given that most of the variance was attributed to the first principal component (PC1), estimates of this PC1 for each camp were derived and then normalised by z-scores as described in Duggan *et al.* (Unpublished). As the PCA was based on recorded detections collected within a 2km radius of the camp and each camp was located 4-15km apart, the PC1 scores for each camp location was already assumed to be reasonably spatially independent. The direction of the scaled PC1 values based on all the recorded detections across all the completed radial surveys were interpreted intuitively with the most positive and most negative values for total detections clearly being associated with the most fully domesticated and fully pristine ecosystems, respectively (Table 2.2). The PC1 based indicators of ecosystem integrity were then ranked so that the lowest ranked camp represented a fully domesticated

ecosystem with a permanent human settlement, while the highest rank camp represented a pristine ecosystem with an intact natural wildlife community in the complete absence of humans or livestock (Table 2.2).

A Spearman's rank correlation test using the *cor.test* function in R version 4.3.1, demonstrated a strong positive correlation between the ranked SNEII scores and the ranked standardised PC1 indicators of ecosystem integrity (Figure 2.18). Like the SNEII, the PCA-based ONEII also ranked the camps outside ILUMA WMA as the most degraded. However, for the ONEII, one encroached camp inside ILUMA WMA, *Bwawa la Mamba Luhogi* (camp number 25), was ranked lower than a camp outside the protection of ILUMA WMA (*Mavimba Pori,* camp number 18) (Table 2.2). This highlights the severity of encroachment inside ILUMA WMA and its impact on natural ecosystems. Rice farming and other tillage farming were all evident in low scoring camps outside ILUMA WMA and in the encroached areas inside ILUMA WMA, and was the predominant land use in the three lowest scoring camps, suggesting that agriculture is a key determinant of poor ecosystem integrity. The ONEII ranked the top ten highest ranked camps inside NNP.

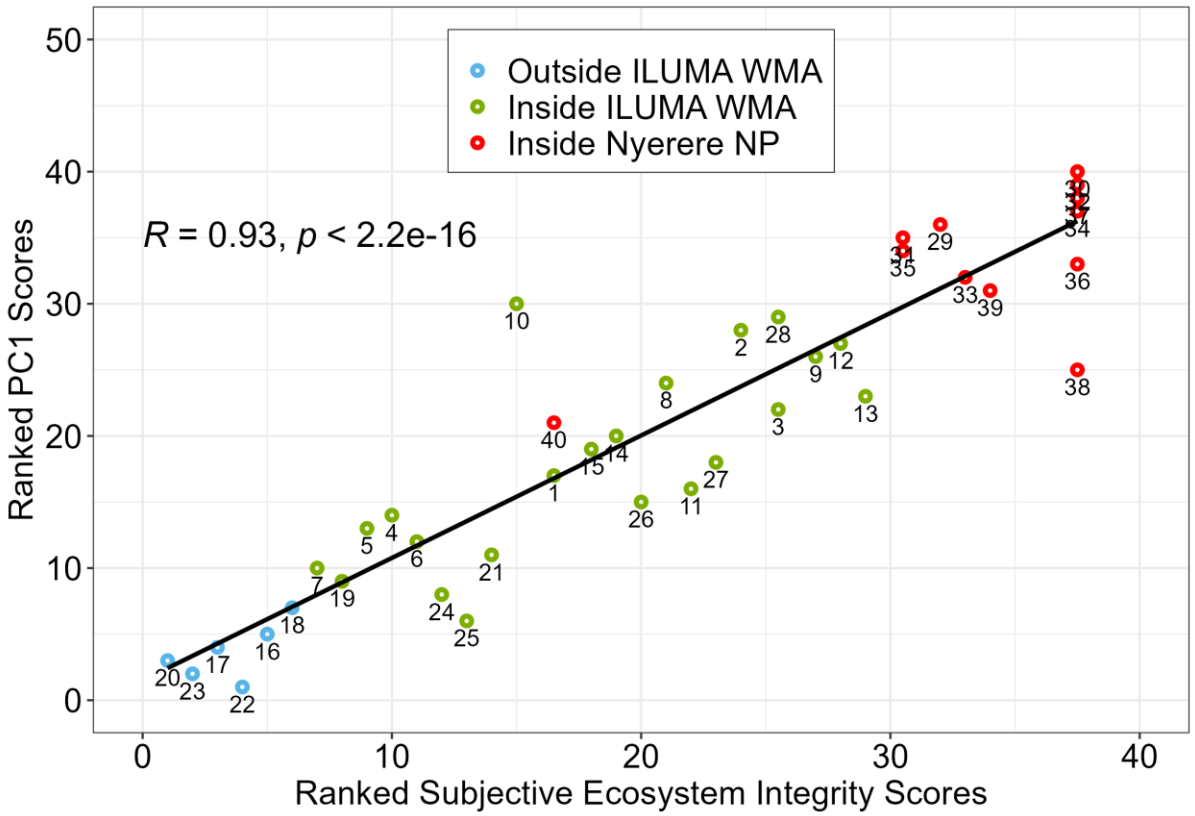

Figure 2.18. The ranked standardised PC1 values of the ONEII for each camp number plotted against the ranked SNEII scores for each camp. Note that while this graph is almost identical to that in Duggan *et al.* (Unpublished), it includes 8 extra camps, one of which (number 40), is a useful exception with unusually low ecosystem integrity for a location so far inside NNP which helps to prove the general rule of close correlation.

Apart from two moderate outliers, *Shuguli Kubwa* (camp number 36) and *Bwawa la Miembeni* (camp number 10), that were respectively ranked lower and higher compared to SNEII ranks, the high correlation displayed in figure 2.18 indicates that this novel methodology of formulating an ecosystem integrity index based on repeated recollective impressions of the study sites was a highly effective approach, and even has advantages compared to the use of the ONEII. The more convenient and intuitive SNEII is measured on a scale from 0% to 100%, thus providing a simplified interpretation of the ecological state at each study site where larvae surveys were conducted. Furthermore, multivariate regression analysis demonstrated that the ONEII was far less sensitive than the SNEII to the observed occurrence and intensity of various human activities (Duggan *et. al.*, Unpublished). Throughout this study, the SNEII was therefore

consistently used as the sole synthetic indicator of natural ecosystem integrity in the subsequent analyses of larval occupancy (Section 2.9.3) and the association between environmental parameters and *An. gambiae* complex sibling species composition (Section 2.9.4)

2.7. Creating additional spatial and geographic variables; 'Distance' and 'Historical Landcover'.

Before performing the statistical analyses on the occupancy data and sibling species composition data, two additional variables were added to the dataset to take further spatial and environmental factors of each camp location into account. First, a continuous variable *distance* was created to account for any potential variance attributable to the natural dispersal of mosquitoes (Service, 1997), from suitable environments to unsuitable ones, which might directly impact occupancy rates or the species composition of populations (further discussed in section 2.9.4). The values for each camp were calculated by measuring the number of kilometres (km) from the camp coordinates to the nearest point on the NNP park boundary line (Figure 2.9), using the geographic information system (GIS) open-source software, QGIS Version 3.30.2. Camps that were located outside the NNP boundary were assigned a negative value for the distance in km, while camps that were located inside the NNP boundary were assigned a positive value. Camp number 40, *Bwawa la Umeme,* which already proved to be an exception of the SNEII because of the observations of deforestation and human activity inside the national park (Section 2.6), also required consideration with respect to the distance variable. Despite being one of the furthest camps, located approximately 40km from the nearest point on the NNP boundary (Figure 2.9), it was also located just 16km from the Julius Nyerere Hydropower plant (Figure 2.9), where there is also a large permanent human settlement resident to those working within the NNP, which may provide a more suitable environment for mosquito populations. Therefore, this camp was again considered to be an exception, and so to improve model fits during statistical analyses, the value in km for this camp only was adjusted

to the distance from this known permanent human settlement inside the park rather than the distance to the nearest NNP boundary.

The categorical variable named *historical landcover* was also created to account for any potential environmental variability associated with different natural landcover types that may influence occupancy and species composition (further discussed in section 2.9.4). There were three distinguished historical landcover types found across the study area; evergreen groundwater forest, moist miombo woodland savanna, and dry acacia savanna. The predominant landcover type was assigned to each camp and was based on investigator recollections, the historical knowledge of VGS regarding fully domesticated areas, and satellite imagery available through Google Earth®. The landcover type was easily distinguished using satellite imagery as the density of vegetation for each of these landcover types vary significantly. The perennial groundwater forest in the north of ILUMA WMA was much denser with a predominantly closed canopy cover that was dark green in colour compared to the deciduous miombo woodlands that had much more broken canopy cover. Camps located in dry acacia savanna were identified as having acacia trees and scrub with only patchy canopy cover, which can easily be distinguished by an expansive landscape of sparse vegetation and light-coloured soil. Camps that were located in the villages outside the western ILUMA WMA border were all assigned miombo woodland, as this was known to be their historical predominant landcover type according to the VGS who would have been resident there.

Once each camp was assigned a SNEII score, a value for the distance from the camp to the nearest NNP boundary or human settlement inside NNP, and the historical landcover type, the datasets were prepared for final analysis as follows.

2.8. Data entry, cleaning and preparation.

To address the aims and objectives of this study, three complete datasets including the larval occupancy data (Section 2.4.1 and 2.4.2, Appendix 1), the PCR results from collected field-identified *An. gambiae* complex larvae (Section 2.4.3), and the corresponding dataset from the radial surveys involving humans, livestock, wildlife activity and land use (Section 2.5, Appendix 2), were all required for the full and final statistical analysis. Recorded observations in the larval survey forms (Appendix 1) were entered into Microsoft Excel® directly after each survey was completed and cleaned using R Version 4.3.1 at the end of rounds 3 and 4 respectively, largely by examining frequency tables broken down by date for implausible values. As the data were directly entered in a horizontal format with observations for each habitat represented by a single line of data with multiple variables, no further restructuring was required for the larval occupancy analysis described in section 2.9.3. These data were linked using Microsoft Access® by camp name and number to a table with the relevant spatial and geographical variables for each camp location, namely SNEII (Section 2.6), distance and landcover (Section 2.7).

To validate the field identification methodology (Section 2.9.2) and to calculate the proportional species composition of field-identified *An. gambiae* complex larvae (Sections 2.9.5 and 2.9.6), the PCR amplification results for each camp from the fourth round of surveys were aggregated by breeding site identification number (BSID) to give the total number of *An. arabiensis*, the total number of *An. quadriannulatus*, and the total number of unamplified larvae for each batch sample of field preserved larvae from specific individual aquatic habitats. The total number of larvae per habitat-specific batch sample at a given camp was calculated by summing all amplified and unamplified larvae for each BSID, while the proportion confirmed to be members of the *An. gambiae* complex was calculated by summing the total number of *An. arabiensis* and the total number of *An. quadriannulatus* identified and dividing by the total

number of larvae tested. This dataset was then linked by camp name and number to SNEII, distance and historical landcover. To plot the results of the larval validation presented in section 3.2, and the effect of spatial and geographic parameters on proportional sibling species composition presented in section 3.4.3, the data set was aggregated by camp number so that each row contained the sum of *An. arabiensis*, the sum of *An. quadriannulatus*, the number of unamplified larvae, the total number of larvae per batch sample and the total number identified as *An. gambiae* complex.

Data from the 40 radial surveys of human, livestock, wildlife activity and land use across each camp location from the fourth round of surveys in 2023 were cleaned, restructured and linked to the existing radial survey data, using open software R Version 4.3.1, Microsoft Excel®, and Microsoft Access®, exactly as described as *Duggan et al.* (Unpublished), so that each of the many specific types of detections (e.g., old hartebeest spoor, direct sighting of a charcoal burner, recent cattle tracks, etc.) at each camp was represented by a separate variable. However, the complete dataset that entailed of recorded detections from 116 radial surveys across the whole study period from January 2022 to July 2023 was used for the PCA of the ONEII (2.6.2). A subset was also created of all 40 radial surveys from round four only and new variables for the total observations of each species was derived by summing every observation count and class for each species. Total detections for each species were normalised using z-score means: $z = \frac{x-\mu}{\sigma}$, where z is the total number of detections of any kind for a given species at each camp, x is the number of detections for a species at each camp, μ is the mean number of detections for the same species across all camps, and σ is the standard deviation of the mean number of detections for that species across all camps. This normalised indicator of the number of detections for each species at each camp was then linked with the PCR results that had been aggregated by BSID and already merged with the spatial and geographic parameters.

2.9 Data analysis.

2.9.1 Distribution of biotic, abiotic, spatial, temporal attributes of sampled waterbodies.

The frequency of each habitat attribute was calculated using the *table* function in R version 3.4.1, and the proportions of habitat observations with a given attribute value were calculated in Microsoft Excel® as n/N, where n was the sum of all habitats with the attribute in question and N was the sum of every habitat that was surveyed across all four rounds.

2.9.2 Validation of field identification methodology for larvae from the An. gambiae complex.

Anopheles larvae encountered during the larval occupancy surveys were positively identified in the field as members of the *An. gambiae* complex if a white collar immediately behind the head was visually observed. Therefore, to validate the field methodology used and to enable face-value interpretation of the occupancy rates recorded in the field and analysed by a generalised linear mixed model (GLMM) as described in the following section 2.9.3, positive PCR amplification rates from field identified larvae were fitted to a GLMM using the *glmm* package in R.

Habitat-specific batch samples of collected larvae from the fourth round either amplified as members of the *An. gambiae* complex or did not amplify at all, hence the proportion of field identified larvae that positively amplified as *An. gambiae* complex was treated as a binomial outcome and weighted by the total number of mosquitoes from that were tested by PCR (Scott, Brogdon and Collins, 1993). The BSID nested within camp number was treated as a random effect to account for any covariance within habitats and camp locations. Distance, SNEII, and historical landcover type were all assessed as univariate fixed effects to account for any spatial or environmental effects on the amplification rates of *An. gambiae* complex that might be associated with them, but none proved significant. Consequently, a GLMM with an intercept only and no other fixed effects was used to calculate an overall mean amplification rate for *An. gambiae* complex (G) with 95% confidence intervals (CIs) across all camps, using the

following formulae, $G = \frac{e^{\beta_0}}{1+e^{\beta_0}}$ with 95% CIs = $\frac{e^{\beta_0 \pm 1.96\,(\sigma_0)}}{1+e^{\beta_0+\sigma_0}}$, respectively, where β_0 is the intercept and σ_0 is the standard error of the mean for the intercept. To illustrate how consistently high these results were regardless of where samples were obtained from, rates of successful amplification of *An. gambiae* complex at each camp were displayed as a scatter plot with the points distributed across a horizontal axis that represents the distance of the camp from the nearest NNP boundary or the nearest human settlement inside the park (km) on a square root scale. The predicted means and confidence intervals for this plot, were calculated based on a second GLMM that included distance as a fixed effect as well as the intercept. Using the outputs from this GLMM the predicted mean (G) and a combined standard error of the means (σ) for each camp were calculated in Microsoft Excel® as $G = \frac{e^{\beta_0+\beta_1 X_{1,i}}}{1+e^{\beta_0+\beta_1 X_{1,i}}}$, and σ $=\sqrt{{\sigma_0}^2 + {\sigma_1 x_{1,i,}}^2}$, respectively, where β_0 is the intercept, β_1 is the effect size attributed to distance, $x_{1,i}$ represents the values of distance at camp i, and σ_0 and σ_1 are the standard errors of the means for β_0 and β_1, respectively. The 95% CI for the predicted means at each camp were then calculated as $G = \frac{e^{\beta_0+\beta_1 X_{1,i} \pm 1.96*\sigma}}{1+e^{\beta_0+\beta_1 X_{1,i} \pm 1.96*\sigma}}$. The predicted means and confidence intervals were then plotted using the *geom_line* and *geom_ribbon* functions, respectively.

2.9.3 Occupancy of aquatic habitats by An. gambiae complex larvae.

The *mgcv* package and *gamm* function were used to fit GLMMs examining the dependence of the proportion of aquatic habitats occupied by *An. gambiae* complex larvae upon various biotic and abiotic factors, using a systematic forward-step selection process. Camp number was added as a random effect to account for covariance of occupancy status across potential larval habitats within the area around individual surveyed camps, and random variance that may occur between camps and is not explained by any of the fixed effects examined and was consistent across survey rounds. Temporal autocorrelation was accounted for by including the number of

weeks since the study commenced as a first order continuous autoregression term nested within the camp number random effect. This also helped mitigate the tendency of the rolling cross sectional study design to spuriously exaggerate the significance of various fixed effects examined because of the temporal and environmental covariance between surveys within a circuit which were completed in sequence at similar times and places. As *An. gambiae* complex larvae were recorded as either present or absent in a habitat, this binary dependent variable was fitted to a binomial distribution with a logit link function.

Univariate analysis was initially completed with all the temporal, environmental, abiotic, and biotic parameters included in separate models as the sole independent variable with a fixed effect. These parameters included all the attributes that were recorded in the larvae survey form (Appendix 1), SNEII, distance to the nearest NNP boundary or settlement inside the park and historical landcover type. The effect of SNEII, the number of dips and distance to the NNP boundary or settlement inside the park were assessed as continuous fixed effects, whereas all other variables including, individual implementing the survey, season, landcover, habitat type, perimeter, water depth and vegetation type were treated as categorical factors. The reference group for a categorical variable was assigned to the level that was most frequently recorded, except for the covariate, round, in which round 4 was treated as the reference group because it represented all surveys carried out by a different individual who clearly detected larvae with greater sensitivity than the investigator who preceded him over the first three rounds.

For categorical factors with many different values, including habitat type and habitat perimeter (Appendix 1), several recorded categories were combined into fewer, larger categories based on having similar crude mean occupancy rates, as estimated by the univariate analysis, and based on their similar physical characteristics and appearances to the investigator, thus reducing the degrees of freedom and increasing statistical power. For the habitat type variable, waterholes and rice paddies were pooled together as a single category as they had statistically

indistinguishable occupancy rates and were often similar in appearance. Streams were combined with springs and swampy areas based on similar criteria. All other habitat types that were not significantly different or were too scarce for which there were too few observations to assess as a separate category, were pooled into the selected reference group of pools, puddles, tracks and depressions, and categorised as *other habitat types.* The new reference group included all the other possible anthropogenic habitat types; artificial ditches or drains, human-dug wells, artificial containers, ridge and furrow agriculture, and other agriculture. These man-made habitats also had similar attributes to the habitats as the original reference group. The category flooded valleys was redundant due to small sample size (n=7) and therefore, was also added to the new reference category. This reduced the number of habitat types by from 12 categories (Appendix 1) to now three categories for GLMM analysis. The habitat perimeter variable was also recategorized by pooling the two larger groups (*21-200m* and *>200m*) together to form the category *>20m,* which reduced the original four categories (Appendix 1) to three. The Akaike information criterion (AIC) score was used to assess the goodness of fit of each model and the *anova* function was used to test for statistically significant differences between the alternative model fits.

The final multivariate model was then assembled by adding independent variables in order of their significance in the univariate analysis results, with variables being removed again based on the principle of parsimony and lack of evidence for improvements in the goodness of fit. For example, if a variable was statistically significant in the univariate analysis but lost significance or didn't improve the goodness of fit when added to a model, it was dropped from the model and excluded from further consideration. The final model only included the attributes that remained statistically significant at the end of the model building process.

The results of the GLMMs were expressed as odds ratios (OR) calculated as e^{β_i} , where β_i is the fixed effect in question. For categorical variables, the OR of β_i was compared against the

reference group for that variable. The OR for distance was calculated as $e^{10\beta_i}$, to represent the change in odds of occupancy for every ten kilometres further towards or beyond the NNP boundary. The OR of SNEII was expressed as $e^{100\beta_i}$ to represent the change in occupancy when moving from a fully domesticated ecosystem to a pristine ecosystem. The 95% confidence intervals for all these ORs were calculated as $e^{\beta_i \pm 1.96*\sigma}$, where β_i is the fixed effect in question and σ is the standard error of the mean of the given fixed effect.

Variations in the proportion of habitats occupied by *An. gambiae* complex larvae in relation to variation in SNEII, were represented graphically as a scatterplot using the *ggplot2* and *scales* packages in R. The data were aggregated by date, so that each data point on the scatterplot would represent the proportion of occupied habitats for one whole larval survey on a given day. The predicted means and confidence intervals for the linear trends observed were calculated based on a GLMM identical to the final best-fitting multivariate model except that all fixed effects other than SNEII were treated as random effects. Using the outputs from this GLMM, the predicted means and confidence intervals for plotting were then calculated based on the same formulae used for plotting in the above section 2.9.2, where β_0 is the intercept, β_1 is the effect size attributed to SNEII, and $x_{1,i}$ represents the values of SNEII at camp i, and σ_0 and σ_1 are the standard errors of the means for β_0 and β_1, respectively.

2.9.4 Conceptual analytical framework for assessing An. gambiae sibling species composition.

The PCR results obtained from adult F_0 mosquitoes that were reared from wild-caught F_0 larvae demonstrated that in addition to *An. arabiensis*, wild populations near or inside NNP also included *An. quadriannulatus*. While performing initial GLMMs to assess the potential effects that may influence the relative abundance of each sibling species at each camp, it was evident that the environmental and spatial variables included (SNEII, distance and historical landcover) covaried with the normalised indicators for the number of detections for each species, collected during the radial surveys. Although SNEII, distance and landcover were important to include in the occupancy analysis, to ensure that as many environmental drivers of larval occupancy as possible were accounted for, their potential for both direct causal effects upon *An. gambiae* complex species composition and indirect causal effects via direct effects upon potential host availability required careful consideration. This is because these fundamental ecological parameters also directly mediate the availability and distribution of the various mammals that may act as potential blood sources for mosquitoes. The conceptual analytical framework developed to guide this analysis (Figure 2.19), therefore illustrates how these three geographic variables may be reasonably expected to, directly and indirectly, influence proportional species composition of *An. gambiae* complex mosquitoes across the environmentally and ecologically diverse study area, serving as the foundational reasoning for how the GLMM analyses were conducted.

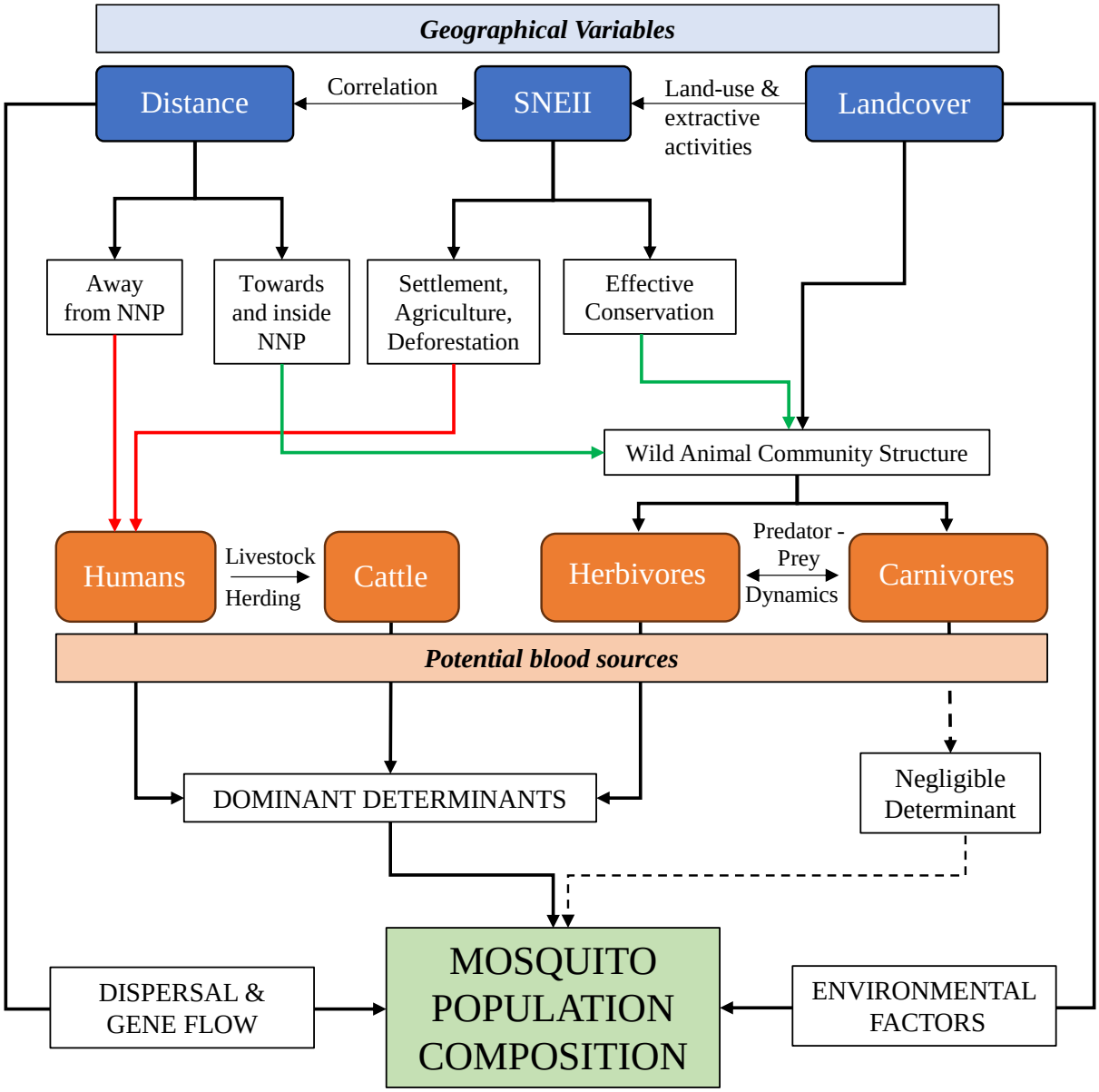

Figure 2.19. A flow diagram illustrating the direct effects of distance, SNEII and landcover on mosquito population composition, and the indirect effects that are mediated through the potential availability and distribution of blood sources.

Firstly, historical landcover and distance to the nearest NNP boundary or settlement inside the park (Section 2.7) may have substantive direct abiotic effects on mosquito population composition (Figure 2.19). The purpose of adding landcover as an additional variable was to account for any potential variance attributable to the environmental factors, that are associated with each of the strikingly different landcover types across the study area, and may well

influence occupancy and species composition. For example, it was observed that evergreen groundwater forest, miombo woodland and acacia savannah, were all associated with different soil types and different levels of canopy cover. These natural differences between landcover types may therefore impact environmental factors such as the level of shade or predominant surface water characteristics, all of which may influence the availability, attractiveness and productivity of aquatic habitats (Holstein, 1954; Gillies and De Meillon, 1968; Gimnig *et al.*, 2001). Furthermore, climatic differences between the landcover types, in terms of temperature and rainfall, for example, may also affect species composition as one particular species may be more tolerant to certain conditions compared to the other (White, 1974; Gillies and Coetzee, 1987).

Also, mosquito dispersal is influenced by the geographic distribution of natural resources, including blood hosts and nectar, resting sites, mates, and aquatic habitats (Service, 1997), so the distribution and gene flow of populations are directly affected. Hence, the addition of distance (Section 2.7) to the list of potential independent variables considered for inclusion in the GLMMs, to account for the direct effects of dispersal (Service, 1997) away from the preferred human and cattle hosts of *An. arabiensis* upon its relative abundance with respect to *An. quadriannulatus*. Additionally, the distance variable may also address any variance in population composition that may be attributable to long-distance windborne dispersal (Service, 1997) of *An. quadriannulatus* away from its optimal ecological niche.

However, although location, ecosystem integrity and landcover may all have direct effects on mosquito population composition (Figure 2.19), these spatial, geographic and environmental variables also fundamentally influence how animal communities are structured, which in turn drives the abundance and distribution of the numerous different mammalian species that may act as potential blood sources for different mosquitoes with different host preferences (Hess, Haves and Tempelis, 1968; Kay, Boreham and Edman, 1979; Lyimo and Ferguson, 2009)

(Figure 2.19). With respect to landcover, some species would be more frequently associated with a certain landcover type as it provides a preferred habitat given the availability of natural resources that are required by that species. For example, it was visually observed during the study, and consistent with the literature (Estes, 2012), that impala was almost exclusively associated with dry acacia savanna, while buffalo were predominantly associated with moist miombo woodlands and red duiker were closely associated with dense groundwater forest.

Historical land cover can also influence the SNEII, as certain types of predominant vegetation are associated with different human activities that have distinctive effects on the intensity of land degradation, which was the primary criterion used by the investigators to assign the SNEII scores. For example, miombo woodland provides more suitable environmental conditions for agriculture than acacia savanna, and is therefore more likely to become badly deforested. Also, miombo trees were also regularly observed to be exploited for charcoal burning during the study, especially in the dry season. On the other hand, the groundwater forest provides ample opportunity for selective harvesting of hardwood trees for high value timber and were, therefore, often deforested less drastically than miombo woodlands, and were far less vulnerable than dry acacia savannah where trees of any kind are scarce. All these intricacies, in turn, impact the densities and distributions of humans, cattle, and different wild animals, thus creating a highly heterogeneous landscape with respect to the preferred blood resources each mosquito species depends on.

Furthermore, the distance into or away from the protected area of NNP has obvious direct effects on SNEII and both are, in turn, direct determinants of mammalian community composition, creating a complex web of causality and correlation (Figure 2.19). Indeed, this study area was specifically chosen due to the clear geographic gradient which runs from west to east, representing a complete transition from fully domesticated land use, to a buffer zone where humans, livestock and wildlife are all present, to wild pristine areas devoid of any

humans or livestock. As the most important and primary criterion for the assigned SNEII values, land degradation due to human activity has negative consequences for wildlife communities and is also associated with the presence of humans and livestock (Duggan *et al.*, Unpublished). Therefore, the furthest distances away from NNP and the lowest SNEII scores occurred at fully domesticated camps, where humans and livestock were clearly the most abundant host species available and so may influence the species composition to favour mosquitoes that prefer these hosts as blood sources. Similarly, as one moves eastwards away from these domesticated landscapes into ILUMA WMA and then NNP, where the natural ecosystems are increasingly intact, animal communities clearly become more abundant and diverse, providing a variety of alternative blood sources that mosquitoes could potentially exploit, thus affecting mosquito species composition.

It was therefore considered that the numbers of detections of humans, livestock and wildlife would be closely correlated with all three of these geographic variables, thus confounding conventional multivariate GLMM analyses of all these variables together, by precluding meaningful model identification and interpretation because both levels of the causal chain illustrated in figure 2.19 would be included at the same time. In view of this, a separate model was developed that used only these broadly geographic factors to address the fundamental environmental and spatial effects on *An. gambiae* complex species composition across the study area (Section 2.9.5), before then investigating how variations in the mosquito population composition were associated with the availability of individual potential host species (Section 2.9.6). If the multivariate host species models of humans, livestock and wild animals had also included the effects of distance, SNEII or landcover, it was considered likely that these covariates would have absorbed much of the variance that would otherwise be attributed to the preferred host species of mosquitoes. This could then readily result in the direct causal effect of a given host species being underestimated or even not identified.

Assuming that wild populations of *An. gambiae* complex survive in the absence of humans and livestock, and must presumably do so by acquiring blood meals from wild herbivores, the predator-prey interactions of wild animals in the study area were also considered when building the models of sibling species composition as a function of the availability of individual host species. Large carnivores commonly recorded in this study area included lion, hyena and leopard (Appendix 2), and it was assumed that their abundance was associated with that of herbivores for a number of reasons. The population ecology of these carnivores, including their distribution and frequency with which they visit particular locations, are influenced by the availability and distribution of prey (Orsdol, Hanby and Bygott, 1985; Spong, 2002; Hopcraft, Sinclair and Packer, 2005; Loveridge *et al.*, 2009; De Boer *et al.*, 2010; Valeix *et al.*, 2009). However, these interactions are highly dynamic, and so the presence of carnivores can also deter herbivores that subsequently avoid certain ambush sites and locations where predators are perceived to be present (Riginos and Grace, 2008; Valeix *et al.*, 2009; Anderson *et al.*, 2010). Because of this complex web of causality and covariance (Figure 2.19), and also because carnivores constitute a negligible fraction of overall mammalian blood mass when herbivores are abundant (Hatton *et al.*, 2015) (Figure 2.19), the numbers of detections of these large carnivores were also omitted from the multivariate host association model, even though they were all clearly associated with *An. gambiae* complex sibling species composition in univariate regression analysis (Section 3.4.3). Descriptions of how the GLMMs were fitted, interpreted, and displayed for *An. gambiae* complex sibling species composition based on (1) distance, SNEII, and landcover, and (2) the number of detections for each animal species, are described in the following sections 2.9.5 and 2.9.6, respectively.

2.9.5 Effects of distance, ecosystem integrity and landcover on An. gambiae complex species composition.

Following revision of the field procedures to ensure samples of larvae caught at all camp locations were preserved *in situ* (Section 2.4.3), GLMMs were also fitted with logit link function to a binomial distribution, treating the dependent variable as the proportion of *An. arabiensis* rather than *An. quadriannulatus*, weighted according to the total number of amplified *An. gambiae* complex specimens identified. The BSID nested within camp number was treated as a random effect to account for any covariance within habitats and camp locations. Univariate analysis of distance, SNEII and historical landcover type were initially completed and the best model fit was derived by the same algorithm based on the principles of parsimony and goodness of fit as described in section 2.9.3.

The proportion of *An. arabiensis* at the mean camp according to the fixed effects in the best fit multivariate model was calculated as $\frac{e^{\beta_0}}{1+e^{\beta_0}}$, where β_0 is the intercept of the model. The ORs and 95% CIs for the univariate and multivariate analyses were expressed for distance and landcover as described in section 2.9.3. The proportion of *An. arabiensis* rather than *An. quadriannulatus* were aggregated by camp to generate a scatter plot where the vertical axis represents the proportional species composition, and the horizontal axis represents the distance variable. The predicted means (A) and standard error of the means (σ), with 95% CIs for the proportion of the PCR-identified members of the *An. gambiae* complex that were *An. arabiensis*, were plotted using an adjusted version of the formulae described in section 2.9.2. to accommodate for more than one fixed effect in the final best fit multivariate model. The adjusted formula used was $A = \frac{e^{\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}+\cdots\beta_n X_{n,i}}}{1+e^{\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}+\cdots\beta_n X_{n,i}}}$, and $\sigma = \sqrt{{\sigma_0}^2 + \sigma_1 {x_{1,i,}}^2 + \sigma_2 {x_{2,i,}}^2 + \cdots \sigma_n {x_{n,i,}}^2}$, respectively, where β_0 is the intercept, $\beta_{1,}\ \beta_2 \ldots\ldots \beta_n$, are the effect sizes attributed to the fixed effects included in the best fit multivariate model, and

$x_{1,i}, x_{2,i}, \ldots x_{n,i}$ represents the values of each respective fixed effect at camp i, and σ_0, $\sigma_1 \ldots \sigma_n$ are the standard errors of the means for β_0, $\beta_0 \ldots \beta_n$, respectively. The 95% CI for the predicted means at each camp were then calculated as $G = \frac{e^{\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}+\ldots..\beta_n X_{n,i}\pm 1.96*\sigma}}{1+e^{\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}+\ldots..\beta_n X_{n,i}\pm 1.96*\sigma}}$. The predicted means and confidence intervals were then plotted using the *geom_line* and *geom_ribbon* functions, respectively.

2.9.6 Associations between An. gambiae complex species composition and surveyed indicators of potential host availability.

A second set of GLMMs of PCR-identified specimens that were identified as *An. arabiensis* rather than *An. quadriannulatus,* were fitted to investigate the effects that humans, cattle, and various wild mammals that might act as potential blood sources, might have on population composition of this mosquito species complex. BSID nested within camp number were again treated as random effects and the spatial and geographic parameters from the best fit GLMM as described in section 2.9.5 were omitted based on the logical conceptual framework explained in section 2.9.4. A Pearson's correlation test was completed to assess the covariance between detected activity levels of humans and cattle, motivating the calculation of a new variable by summing the total numbers of detections of humans and the number of total detections of cattle herds. The new variable was then normalised as z-score means exactly as it had been done for the total number of detections for a given mammalian species at each camp as described in section 2.8. Univariate outputs of the total detections of humans, the total detections of cattle, and the combined new variable were compared in terms of goodness of fit, and the latter was used for multivariate analysis because it yielded the lowest AIC scores. The normalised total detections for all other species were individually assessed as the sole fixed effects in univariate models, and the same forward-step selection process as described in section 2.9.3 was completed to identify the best-fit multivariate model. Univariate and multivariate results of the

species of interest were calculated exactly as described in section 2.9.3. Note, however, that they need to be interpreted differently because the OR reflect the change in the odds of a specimen being *An. arabiensis* rather than *An. quadriannulatus,* per increase of one standard deviation above the mean number of detections of the animal species in question. The predicted means and confidence intervals for the relative abundance of *An. arabiensis* at each camp that were based on the normalised total detections of the animal species from the best model-fit, were calculated using the same formulae as described in section 2.9.5 and were graphically displayed on the same horizontal and vertical axes.

Chapter 3

3. Results

3.1 Biotic, abiotic, spatial, temporal attributes of sampled waterbodies.

A total of 8,266 dips across 1,944 potential larval habitats were completed over the sampling period from January 2022 to July 2023. The relative frequencies of each attribute for the total number of habitats are summarised in Table 3.1. Larvae from the genus *Anopheles* were present in a total of 1,058 habitats. Out of these positive habitats for *Anopheles* larvae, 72.3% were also identified as having *An. gambiae* complex present, demonstrating an overall occupancy rate of 39.4% for this species complex (Table 3.1).

Culex larvae were present in 763 habitats indicating lower overall occupancy (39.2%) than *An.* larvae (Table 3.1). The combined presence of both early and late instar larvae was the most frequent observation for both *Anopheles* and *Culex* mosquitoes (32.3%, 23.6%), followed by late stage only (15.1%, 10.0%), and early stage only (6.9%, 5.6%), respectively (Table 3.1). Pupal stages were identified in only 122 habitats (6.3%).

Table 3.1 Frequencies and proportions of the recorded biological, physical, spatial, and temporal attributes of all the habitats surveyed. Total habitat observations are equal to the sum of the number of habitats that possessed the given attribute at the time it was surveyed, totalled across all four rounds from January 2022 to July 2023.

Attribute	Total Habitat Observations	Proportion of Total Habitats (%)
Anopheles larvae		
Absent	886	45.6
Present	1058	54.4
Early instar only	134	6.9
Late instar only	293	15.1
Early and late instar	631	32.3
An. gambiae complex present	765	39.4
Culex larvae		
Absent	1181	60.8
Present	763	39.2
Early instar	110	5.6
Late instar	194	10.0
Early and late instar	459	23.6

Pupal stage		
Absent	1822	93.7
Present	122	6.3
Habitat Type and Description		
Pools, puddles tracks & depressions	1343	69.1
Springs and swampy areas	125	6.4
Flowing streams	97	5.0
Flooded valleys	7	0.4
Waterholes	173	8.9
Artificial drains or ditches	4	0.2
Human dug wells	105	5.4
Artificial containers	2	0.1
Rice paddies	69	3.5
Ridge and furrow agriculture	11	0.6
Other agriculture	5	0.3
Other	3	0.1
Habitat Perimeter		
<2m	270	13.9
2m-20m	1092	56.2
21-200m	402	20.7
>200m	180	9.2
Number of Dips		
1	174	8.9
2	182	9.4
3	395	20.3
4	337	17.3
5	585	30.1
6	50	2.6
7	52	2.7
8	34	1.8
9	16	0.8
10	119	6.1
Water Depth		
<0.1	1256	64.6
0.1 – 0.5	650	33.4
>0.5	38	2.0
Short Vegetation		
Absent	1301	66.9
Present	643	33.1
Tall Vegetation		
Absent	1580	81.3
Present	356	18.7

Floating Vegetation		
Absent	1780	91.6
Present	164	8.4
Location		
Inside ILUMA WMA	1072	55.2
Inside NNP	294	15.1
Outside ILUMA WMA	577	29.7
Historical Landcover		
Miombo woodland	1671	86.0
Groundwater forest	123	6.3
Acacia savanna	150	7.7
Season		
Wet	1365	70.2
Dry	579	29.8
Round		
1	234	12.0
2	591	30.4
3	559	28.8
4	560	28.8

Pools, puddles, tracks, and depressions were the most frequent habitat type accounting for more than two thirds of all habitats surveyed, followed distantly by waterholes, and springs and swampy areas, which only accounted for 6% of all habitats (Table 3.1). Human-dug wells and rice paddies were the most frequently surveyed man-made habitats, representing approximately 5% and 4% of all habitats surveyed, respectively (Table 3.1). Other artificial or man-made habitats combined only accounted for just over 1% of all surveyed habitats (Table 3.1). The frequency distribution of habitat type is also reflected by distribution patterns of the habitat's other physical attributes. More than half of all habitats had a perimeter between 2m and 20m (Table 3.1), which naturally included many from the pools, puddles, tracks, and depressions category. Whereas aquatic habitats exceeding 20m in perimeter, accounted for less than one third of all those surveyed (Table 3.1), and would have been mostly associated with less frequently sampled waterbodies such as rice paddies, springs and swampy areas, waterholes,

and flooded valleys (Table 3.1). As the number of dips taken was dependent on habitat size (Section 2.4.2), and more than 85% of the habitats were dipped five times or less, all habitats with a perimeter less than 20m fitted into this modest size category, as did many that were greater than 20m but were not fully accessible. Almost two-thirds of habitats were shallower than 10cm (Table 3.1) and were absent in most habitats especially the pools, puddles, tracks, and depressions category.

About 70% of habitats were surveyed during the wet season (Table 3.1) when surface water and mosquitoes were most abundant. The first round of the study had the lowest proportion of habitats sampled (Table 3.1) because one circuit was omitted, so only 20 out of a possible 28 camps were visited. Furthermore, data from the first six camps surveyed were omitted for analysis due to missing data in the *number of dips* column, which was only introduced to the data collection form (Appendix 1) two weeks into the study. Although the number of camps slightly differed per round, the number of habitats surveyed during rounds 2, 3 and 4 remained roughly consistent (Table 3.1). The habitats surveyed in round two were conducted in all the original 28 camps located inside ILUMA WMA and outside it in the villages to the west (Table 3.1). By the time the third round was conducted in August 2022, water sources in four camps in the north-east circuit had fully dried out and therefore were not surveyed. Round three also included four new camps inside NNP, collectively known as the NNP-Boma Ulanga circuit (Table 2.1). For round 4, potential larval habitats were surveyed across a total of 40 camps, including the initial 28 camps, the additional four camps that were added at the end of round 3, and a further eight camps located deeper into NNP that were visited during two separate circuits, namely the NNP-Kilombero River and NNP-Msolwa circuits (Table 2.1). The number of habitats surveyed inside the national park accounted for only 15% of the total over the 4 rounds (Table 3.1), as they were added at the last round in the final stage of the study. A total of 22 out of 40 camps were located inside ILUMA WMA and all of them were surveyed at

least twice (Table 2.1), so they correspondingly accounted for more than half of all aquatic habitats inside miombo woodland. Only 6% of habitats were surveyed in the groundwater forest (Table 3.1), because this landcover type was only found in the north of ILUMA WMA, where the availability of surface water was mainly limited to waterholes and small puddles. Dry acacia savanna landcover was exclusively found in a small number of camps well inside NNP that were surveyed only once or twice (Table 2.1, Table 3.1).

3.2 Validation of field identification methodology for larvae from the An. gambiae complex.

The DNA of 2,468 field-identified *An. gambiae* complex larvae taken from 286 aquatic habitats across 39 camps were tested for species identity by PCR (Scott, Brogdon and Collins, 1993). The number of aquatic habitats at each camp that larvae were collected from ranged from 1 to 10, with an overall average of 7 habitats per camp. The *a priori* target of 10 aquatic habitats per camp was not reached at some camps when there was a limited availability of occupied habitats that contained what appeared to be *An. gambiae* complex larvae. Camp number 29, *Kambi ya Simba*, was surveyed for occupancy but there were no available batch samples to be identified to species level by PCR.

The overall mean amplification rate of 96.6% [94.0%, 96.8%] (Figure 3.1), was very high, and an amplification of 5% failure rate would be normally considered excellent, even for insectary reared positive controls. Neither distance (OR = 1.09, P = 0.585), nor ecosystem integrity (OR = 0.99, P = 0.922), nor historical landcover type (OR = 1.01, P = 0.794) had any effect on amplification rates, indicating that any larva collected under these varying geographic, and environmental conditions can be confidently identified in the field as *An. gambiae* complex if a single white stripe is evident immediately behind the head. It also validates the results fitted by the GLMM in section 3.3 which predicts *An. gambiae* complex occupancy rates across ecosystem integrity and specific habitat attributes. Above all, it concludes that the field

identification approach proved to be very reliable, with the exception of one outlier in Funga (camp 26) (Figure 3.1).

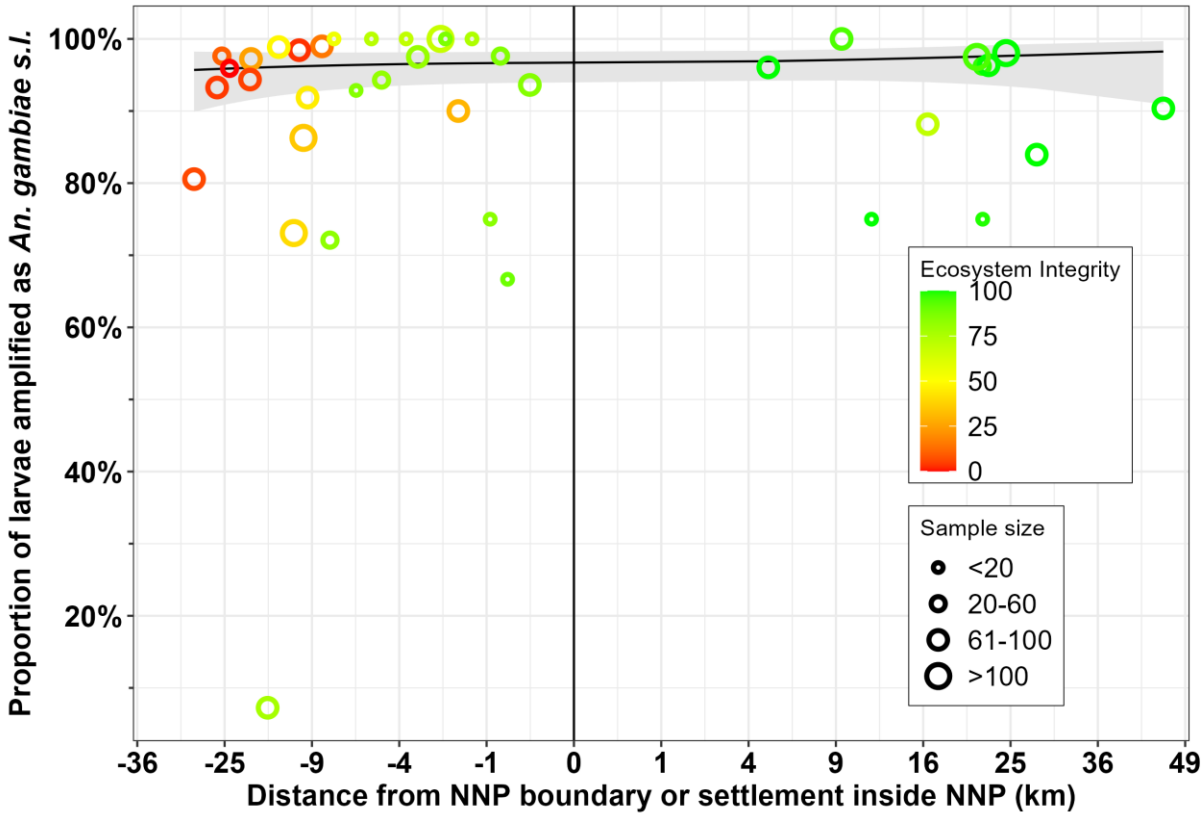

Figure 3.1. A graphical illustration of how the proportion of field-identified larvae that successfully amplified as *An. gambiae* complex did not vary with SNEII or distance from the nearest point to the NNP boundary or the nearest human settlement within the park. Each point represents the proportion of field-identified larvae confirmed to be *An. gambiae* complex by PCR (Scott, Brogdon and Collins, 1993) across all batch samples collected at each camp. The colour and size of each point represent the SNEII and the total sample size for each camp, respectively. Predicted means and confidence intervals were calculated from a GLMM fitted to the proportion of amplified specimens, using a binomial distribution and logit link for this binary dependent variable in which the non-significant effect of distance was treated as the only fixed effect with breeding site identification number nested within camps treated as random effects.

3.3 Occupancy of aquatic habitats by An. gambiae complex larvae.

Taking a forward-step selection process to building the model disclosed any confounding effects of habitat attributes on occupancy by *An. gambiae* complex larvae. The effect of season on occupancy was no longer significant once round and habitat type were accounted for (Table 3.2), probably because each round was designed to represent a period in the wet or dry season and therefore captures most of the variance with season. The covariate habitat type may also have had confounding effects as occupancy was higher in waterholes and rice paddies (Table 3.2), which were mostly seasonal, and were therefore more common after rainfall during the wet season. Floating vegetation and tall vegetation also proved to have no effect on the occupancy of *An. gambiae* complex once other covariates were accounted for (Table 3.2). Habitat type may account for any variance associated with these covariates as these vegetation types were often closely associated with specific types of waterbodies. For example, tall vegetation would be characteristic of late-stage rice paddies and floating vegetation was commonly observed in waterholes.

Table 3.2. Generalised additive mixed model univariate and multivariate outputs for the proportion of habitats occupied by *An. gambiae* complex larvae based on spatial, temporal, abiotic and biotic habitat attributes fitted to a binomial distribution and a logit link function. Variation between camps across rounds was accounted for as a random effect. First order temporal autoregression was accounted for by nesting the number of weeks since the first survey in camp number. Statistically significant effects are highlighted in bold.

	Univariate			**Multivariate**		
Fixed Effects	OR [95% CI]	t	P	OR [95% CI]	t	P
Intercept[a]	NA	NA	NA	1.35 [0.83, 2.21]	1.21	0.2268
Survey Implementor						
Individual 1	1.00	NA	NA	NE	NE	NE
Individual 2	**3.91 [2.91, 5.24]**	**9.09**	**<<0.0001**	NE	NE	NE
Temporal and spatial characteristics						
Round number						
1	**0.36 [0.23, 0.56]**	**-4.50**	**<<0.0001**	**0.27 [0.17, 0.43]**	**-5.60**	**<<0.0001**[d]
2	**0.29 [0.21, 0.41]**	**-7.19**	**<<0.0001**	**0.21 [0.15, 0.31]**	**-8.41**	**<<0.0001**[d]
3	**0.19 [0.13, 0.27]**	**-9.20**	**<<0.0001**	**0.18 [0.12, 0.26]**	**-9.04**	**<<0.0001**[d]
4	1.00	NA	NA	1.00	NA	NA
Season						
Wet	1.00	NA	NA	1.00	NA	NA
Dry	**0.41 [0.30-0.55]**	**-5.93**	**<<0.0001**	0.37 [0.09-1.44]	-1.44	0.1513[e]
Land cover						
Miombo woodland	1.00	NA	NA	1.00	NA	NA
Groundwater forest	0.82 [0.43, 1.56]	-0.60	0.5467	0.76 [0.37, 1.60]	-0.71	0.4746[e]
Acacia savannah	**1.93 [1.08, 3.44]**	**2.23**	**0.0262**	0.89 [0.45, 1.76]	-0.33	0.7400[e]
Distance	1.01[0.99, 1.03][b]	1.59	0.1120	0.97 [0.96, 0.99][b]	-1.83	0.0678[e]
Location						
Inside ILUMA	1.00	NA	NA	1.00	NA	NA
Outside ILUMA	1.09 [0.68, 1.77]	0.36	0.7165	0.60 [0.32, 1.13]	-1.59	0.1132[e]
Inside Nyerere NP	1.34 [0.83, 2.18]	1.20	0.2313	0.96 [0.54, 1.73]	-0.12	0.9042[e]
SNEII	0.91 [0.91, 0.92][c]	-0.33	0.7410	**0.38 [0.37, 0.38][c]**	**-3.80**	**<0.0001**[d]
Abiotic habitat characteristics						
Habitat type						
All other habitats	1.00	NA	NA	1.00	NA	NA
Waterholes & rice paddies	**2.58 [1.95, 3.42]**	**6.61**	**<<0.0001**	**1.78 [1.22, 2.58]**	**3.01**	**0.0028**[d]
Springs, streams & swamps	**1.72 [1.30, 2.26]**	**3.85**	**<0.0001**	1.02 [0.71, 1.46]	0.08	0.9332[e]
Perimeter						
<2m	1.10 [0.84, 1.45]	0.69	0.4900	**1.61 [1.15, 2.26]**	**2.74**	**0.0058**[d]
2m-20m	1.00	NA	NA	1.0	NA	NA
>20m	**1.84 [1.51, 2.24]**	**4.80**	**<<0.0001**	1.19 [0.90, 1.58]	1.29	0.2194[e]
Number of dips	**1.14 [1.09, 1.18]**	**6.03**	**<<0.0001**	**1.15 [1.08, 1.22]**	**4.34**	**<<0.0001**[d]

Water depth						
<0.1	1.00	NA	NA	1.00	NA	NA
0.1-0.5	0.98 [0.80, 1.18]	-0.25	0.8029	0.82 [0.65, 1.03]	-1.71	0.0871[e]
>0.5	0.54 [0.28, 1.05]	-1.81	0.0711	0.59 [0.28,1.26]	-1.36	0.1732[e]
Biotic habitat characteristics						
Short vegetation						
Absent	1.00	NA	NA	1.00	NA	NA
Present	**1.68 [1.37, 2.05]**	**5.02**	**<<0.0001**	**1.30 [1.03, 1.63]**	**2.22**	**0.0261[d]**
Floating vegetation						
Absent	1.00	NA	NA	1.00	NA	NA
Present	**1.62 [1.17, 2.225]**	**2.89**	**0.004**	0.95 [0.64, 1.43]	-0.23	0.8187[e]
Tall vegetation						
Absent	1.00	NA	NA	1.00	NA	NA
Present	1.28 [1.01, 1.64]	2.02	0.0433	0.76 [0.56, 1.02]	-1.86	0.0632[e]

Random Effects	σ
Camp number	0.351
First order continuous autoregression	Phi
Weeks since start of study/Camp number	0.232

OR; Odds ratio.
95% CI; The 95% confidence interval
[a] Larval occupancy under reference conditions.
NA; Not applicable because several different intercepts were estimated for more than one fitted univariate model, or not applicable to the reference group.
NE; Not estimated due to confounding effects.
[b] Odds ratio for every kilometre increase from the NNP boundary.
[c] Odds ratio between fully domesticated and fully pristine habitats
[d] As estimated from the final best fit GLMM where the effect was included based on a significant value <0.05.
[f] As estimated from the point at which the effect was no longer significant (>0.05) in the multivariate model fit and was therefore no excluded from the final model.
σ; standard deviation

Ecosystem integrity, survey round, habitat type, habitat perimeter, short vegetation and the number of dips were all significant predictors of *An. gambiae* complex larval occupancy in aquatic habitats (Table 3.2). Ecosystem integrity was the third most important predictor of larval occupancy, indicating a significant decrease in occupancy rates between aquatic habitats in fully domesticated ecosystems and aquatic habitats in fully pristine ecosystems (Table 3.2), suggesting that larval occupancy was somewhat lower for aquatic habitats in pristine areas further away from human settlements and domesticated land use (Figure 3.2). Despite this, *An. gambiae* complex larvae, were nevertheless present with mean occupancy rates exceeding

25%, even in absolute pristine environments far away from any signs of humans or livestock (Figure 3.2). This indicates that adult mosquitoes from the *An. gambiae* complex are ovipositing year-round in aquatic habitats even >40km away from the nearest resident human or livestock hosts.

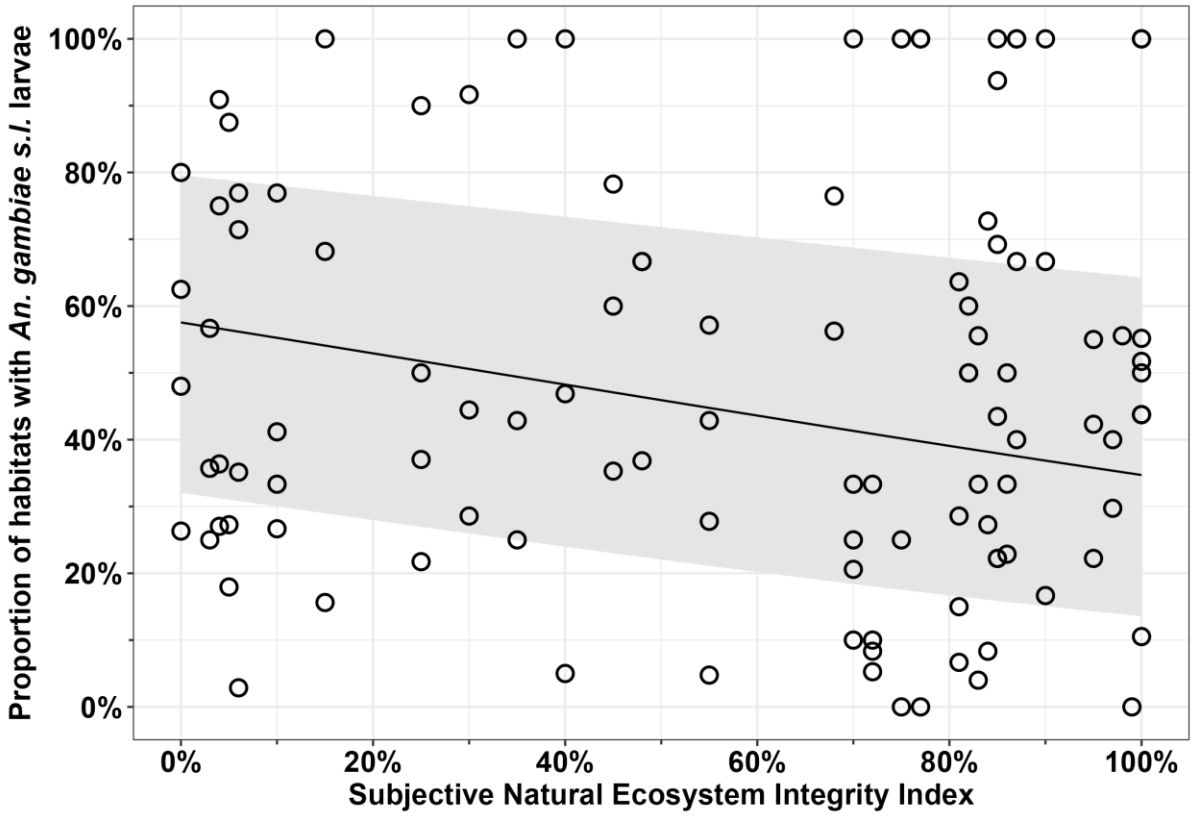

Figure 3.2. The proportion of habitats occupied by *An. gambiae* complex larvae for each survey conducted over four rounds plotted as a function of ecosystem integrity. The estimated mean trend with 95% confidence intervals reflects the predicted values for each surveyed camp location based on a fitted GLMM identical to the multivariate model detailed in table 3.2, except that all fixed effects other than the subjective natural ecosystem integrity index (SNEII) were instead treated as random effects, so that the predictions reflect the absolute means rather than the means for the reference values of each variable.

The odds of occupancy by *An. gambiae* complex larvae were 73% (P <<0.0001), 79% (P <0.0001) and 82% (P <<0.0001) lower, respectively for each of the consecutive rounds completed in 2022 when compared to all the habitats surveyed in round 4 (Table 3.2), which was completed the following year in 2023. This significant difference may probably be attributable to the two individuals that implemented the surveys, with the new person who took

over for round 4 exhibiting an exceptionally meticulous approach. As round and surveyors were obviously closely associated, the latter variable was dropped from the multivariate model because round captured all the variance attributed to the former plus any additional variance that occurred in between replicates.

Abiotic and biotic predictors of *An. gambiae* complex occupancy included habitat type, the number of dips taken, habitat perimeter and the presence of short vegetation (Table 3.2). Although the springs, swamps and streams category, was no longer significant in the multivariate analysis, habitats that were classified as waterholes or rice paddies had higher occupancy rates compared to the reference category of other habitats (Table 3.2). The apparent effect of habitat perimeter differed between univariate and multivariate analysis (Table 3.2), probably because it was highly correlated with the number of dips made. The number of dips was the second most important predictor of habitat occupancy and the effect size remained consistence across univariate and multivariate analysis. In the final model fit, each additional dip increased the odds of a habitat being occupied with *An. gambiae* complex larvae by 15% (Table 3.2). Therefore, the effect size estimated for larger habitats that had perimeters greater than 20m, was reduced and was no longer statistically significant (Table 3.2), probably because it was confounded by the higher number of dips taken from them. The model also revealed that the odds of occupancy was higher by almost two thirds in the smallest habitats (<2m in perimeter), which were dipped or sampled with a turkey baster only once or twice, compared to those of intermediate size (perimeters between 2m and 20m) (Table 3.2). The effect of short vegetation was less pronounced in the final model than in the univariate analysis (Table 3.2).

3.4 Sibling species composition of field-identified An. gambiae complex larvae.

3.4.1 Preliminary data exploration using F1 adults raised from larvae.

PCR results from the *An. gambiae* complex F1 adults that were raised from collected larvae during surveys in the first three survey rounds in 2022, demonstrated that these *An. gambiae* complex populations comprised not only of *An. arabiensis*, but also *An. quadriannulatus*, proving that these sibling species can co-exist in the same environment. However, the proportions accounted for by these two species differed between camps as *An. arabiensis* only were found in fully domesticated areas and highly encroached camps but *An. quadriannulatus* were found in camps with SNEII scores of 50% or higher (Figure 3.3). Furthermore, initial data exploration clearly indicated *An. quadriannulatus* were found predominantly in pristine areas which were well conserved and were located further away from human settlements (Figure 3.3), suggesting that this non-vector utilizes wild animal hosts for blood feeding. Preliminary graphical data on potential associations with the abundance of particular host species indicated that the relative abundance of *An. arabiensis* declined as SNEII increased and the numbers of cattle herds detected became fewer (Figure 3.3). Higher proportions of *An. arabiensis* larvae in the locations where cattle were present suggested competitive interactions between the two sibling species that may be potentially driven by the availability of this known preferred host for this species (Gillies and Coetzee, 1987; Charlwood and Edoh, 1996; Minakawa, Seda and Yan, 2002), giving them a competitive advantage over *An. quadriannulatus*. However, further investigation was required to test this hypothesis, and in particular to address concerns about larval rearing processes at the Msakamba insectary distorting sample composition (Section 2.4.3), and to find out whether *An. quadriannulatus* completely replaces *An. arabiensis* deeper inside NNP, even further away from humans and livestock. Hence the protocol for collecting field identified mosquitoes was revised so that additional samples of larvae were preserved *in*

situ (Section 2.4.3) and nine extra camps much further inside NNP were added for the final survey round in 2023 (Section 2.2).

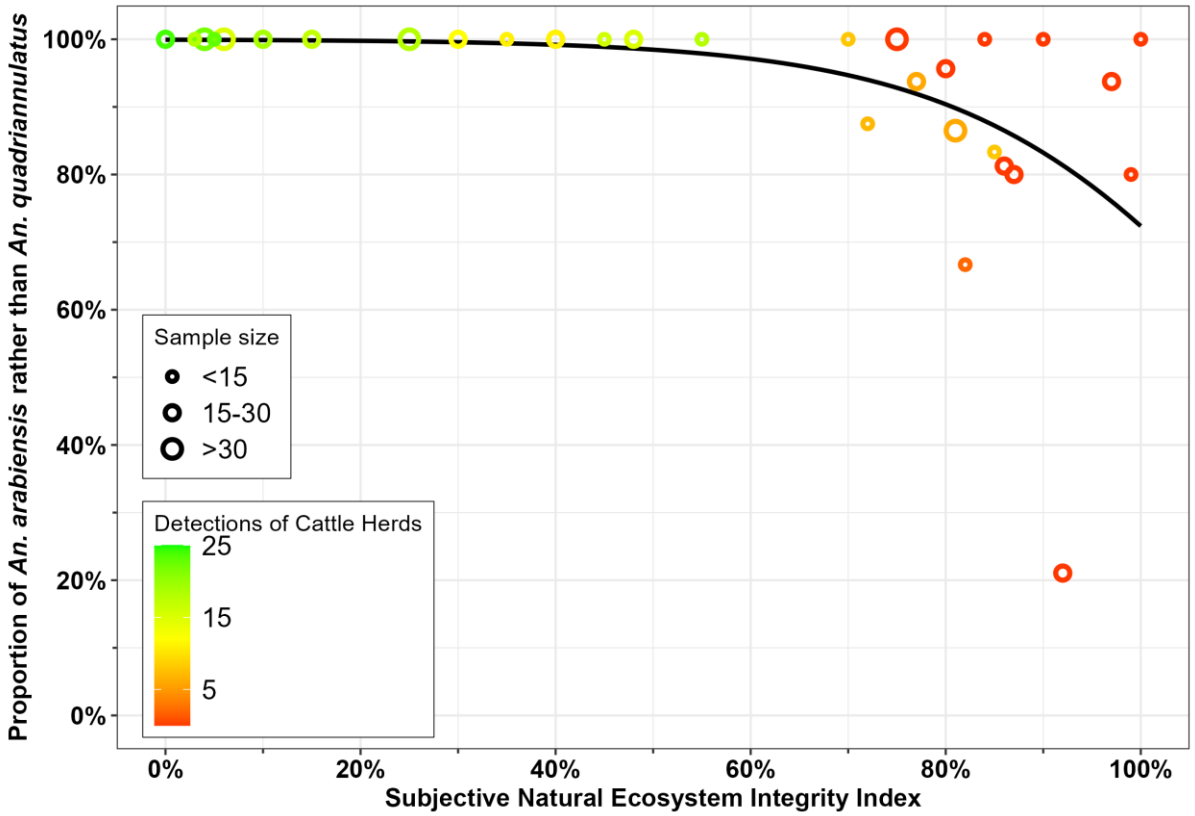

Figure 3.3 A scatterplot graph of the proportion of *An. gambiae* complex F1 adults that were identified as *An. arabiensis* rather than *An. quadriannulatus* by PCR amplification (Scott, Brogdon and Collins, 1993) for each camp plotted against the SNEII, with sample size and the frequency of which cattle herds were detected represented by symbol size and colour, respectively. The graph was generated with *ggplot* in R and the trendline was fitted with the *glm* option using the *geom_smooth* command, specifying a binomial distribution and logit link function for the dependent variable.

3.4.2 Evidence for a truly competitive relationship between Anopheles arabiensis and Anopheles quadriannulatus.

Although variable species composition within the *An. gambiae* complex could be explained by independently varying densities of the two sibling species found within the study area, this does not appear to be the case: As illustrated in Figure 3.4, the absolute numbers of *An. arabiensis* identified at each habitat when surveyed have a weak, but nevertheless negative correlation ($\rho = -0.31$, $p = <0.0001$) with the absolute numbers of *An. quadriannulatus* found. In simple terms, wherever there were more *An. arabiensis*, there tended to be less *An. quadriannulatus*, and *vice versa*, confirming that these two species compete with each other in the strict sense (Box 1) to some extent.

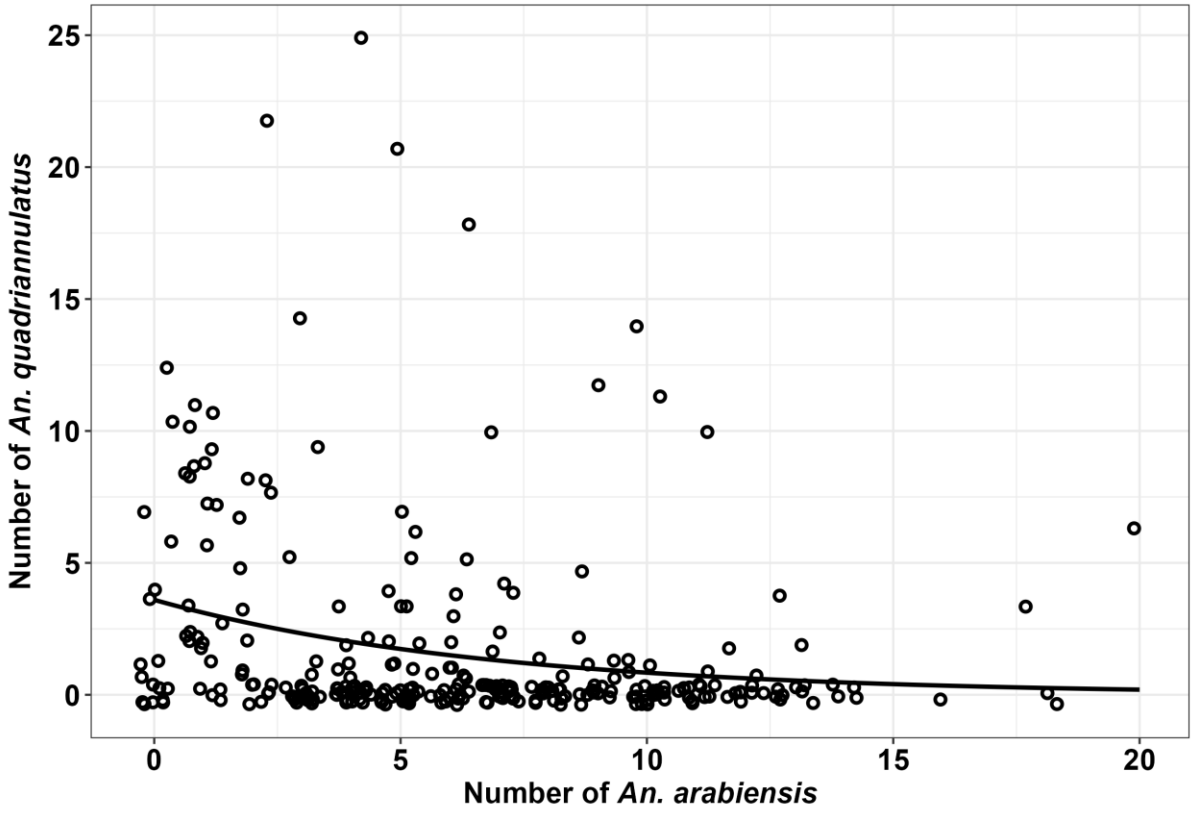

Figure 3.4. The number of *An. arabiensis* larvae per habitat against the number of *An. quadriannulatus* larvae per habitat, as determined by the full fourth round of surveys.

Box 1. Defining competition, competitive displacement and competitive co-existence in strict ecological terms.

Competition occurs when two species require the same limiting resources (Tilman, 1982; Grover, 1997). The superior competitor reduces the availability of resources, mainly through exploitative or interference mechanisms (Schoener, 1983; Aschehoug et al., 2016), thus weakening the survival and reproductive success of the latter competitor. Therefore, in a competitive relationship between two species, the relative abundance of one species is affected by the abundance of its competitor (Volterra, 1931; Lotka, 1932), just as illustrated for *An. arabiensis* and *An. quadriannulatus* in figure 3.4. The most extreme outcome of competition is known as *competitive displacement* (DeBach, 1966), when one species exhibits such a clear competitive advantage that its dominance leads to the elimination of the other. Hence, this concept is based on the principle of competitive exclusion (Hardin, 1960), wherein it is assumed, sometimes too simplistically, that two species can only co-exist if they compete for different limiting resources (Gause, 1934; Hardin, 1960).

However, complete competitive displacement is rarely observed between naturally co-occurring species that seemingly require the same resources. A famous example is that of the 'paradox of the plankton', where Hutchinson (1961) raises the question as to why there is such diversity of co-existing plankton, despite apparently requiring the same limiting resources. Indeed, although such *competitive co-existence* sounds counter-intuitive, it can be explained by stochastic dynamics driven by external environmental fluctuations (Hutchinson, 1961; Richerson, Armstrong and Goldman, 1970), as well as intrinsic chaos (Scheffer et al., 2003; Roy and Chattopadhyay, 2007) and meso-scale spatial heterogeneity (Bracco, LaCasce and Provenzale, 2000), so that ecosystem equilibrium is never reached and no single species can dominate (Hutchinson, 1961).

Similarly, *Anopheles* populations are often found in sympatry (Gillies and Coetzee, 1987), with aquatic stages sharing the same habitats (Gillies and Coetzee, 1987; Charlwood and Edoh, 1996; Minakawa et al., 1999; Gimnig et al., 2001), where larval stages need to compete for limiting resources such as nutrients and space (Koenraadt et al., 2004; Service, 1993; Dye, 1984; Barbosa, Peters and Greenough, 1972; Carpenter, 1983). While *Anopheles* larval habitats certainly exhibit the kind of stochastic dynamics considered to enable competitive co-existence of two or more species (Hutchinson, 1961; Chesson, 1986; Chesson, 2000), it is also worth noting that the competitive balance between species within larval populations is clearly influenced by the local abundance (Charlwood and Edoh, 1996; Minakawa, Seda and Yan, 2002) and safe accessibility (Gillies and Smith, 1960; Smith, 1962; Bayoh et al., 2010; Mutuku et al., 2011; Kawada et al., 2012) of suitable blood hosts for the relevant adult populations to feed upon. The influence of such factors, particularly the safe accessibility of blood meals, on the competitive dynamics and population composition of co-existing mosquito species are discussed further in section 4.4 and are emphasized by (Lounibos, 2007). Indeed, Lounibos (2007) concludes that the dominant competitor reduces the population abundance of the weaker competitor mainly through larval resource or interference competition mechanisms, but also that complete displacement (DeBach, 1966) between competing mosquito species does not tend to occur. Empirical field observations of sympatric populations of more than one mosquito species in most locales, even though they clearly share habitats and resources therein, therefore, also deviates from the classical but simplistic principle of competitive exclusion (Gause, 1934; Hardin, 1960), and instead resembles competitive co-existence like that observed between plankton species (Hutchinson, 1961).

3.4.3 Effects of distance, ecosystem integrity and landcover on An. gambiae complex species composition.

Using the field-identified larvae from round four that were collected under the revised protocol and preserved immediately *in situ*, distance from the NNP boundary or nearest human settlement inside the park, SNEII, and historical landcover were all found to be determinants of the sibling species composition of the *An. gambiae* complex by univariate analysis (Table 3.3). However, historical landcover type was excluded from the final multivariate model (Table 3.3), possibly because it correlated with the distance variable. For example, acacia savanna was exclusively found inside NNP and was the dominant landcover type in the most isolated parts the park along the Kilombero river. Despite the visually obvious correlation between distance inside the park and ecosystem integrity (Figure 3.5), the SNEII still accounts for additional variance that is not already captured by distance. The best fit GLMM of these spatial and environmental covariates indicate that larval habitats located closer to human settlement and at locations with low ecosystem integrity scores, have higher proportions of *An. arabiensis* rather than *An. quadriannulatus* (Table 3.3, Figure 3.5). As larval habitats are located further away from the boundary and in areas of improved ecosystem integrity, the inverse effect occurs so that the odds of *An. gambiae* complex larvae amplifying as *An. arabiensis* decline by 63% for every 10km further past the boundary and by 5% for every percentage point increase in the SNEII (Table 3.3, Figure 3.5), suggesting that *An. quadriannulatus* may have a competitive advantage in pristine ecosystems far away from human settlements and their associated cattle herds.

Table 3.3. Univariate and multivariate outputs of generalized linear mixed models (GLMMs) for the effects of spatial and geographic attributes on the proportion of field identified *An. gambiae* complex larvae that were identified as *An. arabiensis* rather than *An. quadriannulatus* by PCR (Scott, Brogdon and Collins, 1993). All model outputs were fitted to a binomial distribution with logit link function. A random effect that nested breeding site identification number within camp was included to account for variation between and covariance within larval populations. Statistically significant effects are highlighted in bold.

	Univariate			**Multivariate**		
Fixed Effects	OR [95% CI]	z	P	Proportion[a] or OR[b] [95% CI]	z	P
Intercept	NA	NA	NA	0.99 [0.99, 0.99][a]	6.22	<<0.0001
Spatial and geographic attributes						
Distance	**0.20 [0.19, 0.21][c]**	**-8.49**	**<<0.0001**	**0.37 [0.35, 0.39][c]**	**-4.25**	**<0.0001[e]**
SNEII	**0.91 [0.88, 0.94][d]**	**-6.55**	**<<0.0001**	**0.95 [0.92, 0.98][d]**	**-3.71**	**0.0002[e]**
Land cover						
Miombo Woodland	1.00	NA	NA	1.00	NA	NA
Groundwater Forest	0.86 [0.03, 22.59]	-0.08	0.9332	1.03 [0.17, 6.24][b]	0.03	0.9773[f]
Acacia Savanna	**0.01 [<0.00, 0.11]**	**-3.75**	**0.0001**	1.67 [0.52, 5.35][b]	0.87	0.3858[f]
Random Effects				σ		
Breeding site identification number/Camp number				0.833		
Camp number				0.802		

[a] Proportion identified as *An. arabiensis* under reference conditions for all fixed effects.
[b] OR; Odds ratio.
95% CI; The 95% confidence interval estimated for the proportion or OR.
NA; Not applicable because several different intercepts were estimated for more than one fitted univariate model, or not applicable to the reference group.
[c] Odds ratio for every 10km further towards the NNP boundary or past it and deeper into the park
[d] Odds ratio for each percentage point increase along the subjective natural ecosystem integrity index (SNEII).
[e] As estimated from the final best fit GLMM where the effect was included based on a significant value <0.05 and contributed to a lower AIC score.
[f] As estimated from the point at which the effect was no longer significant (>0.05) in the multivariate model fit and increased the AIC score, and was therefore excluded from the final model.
σ standard deviation

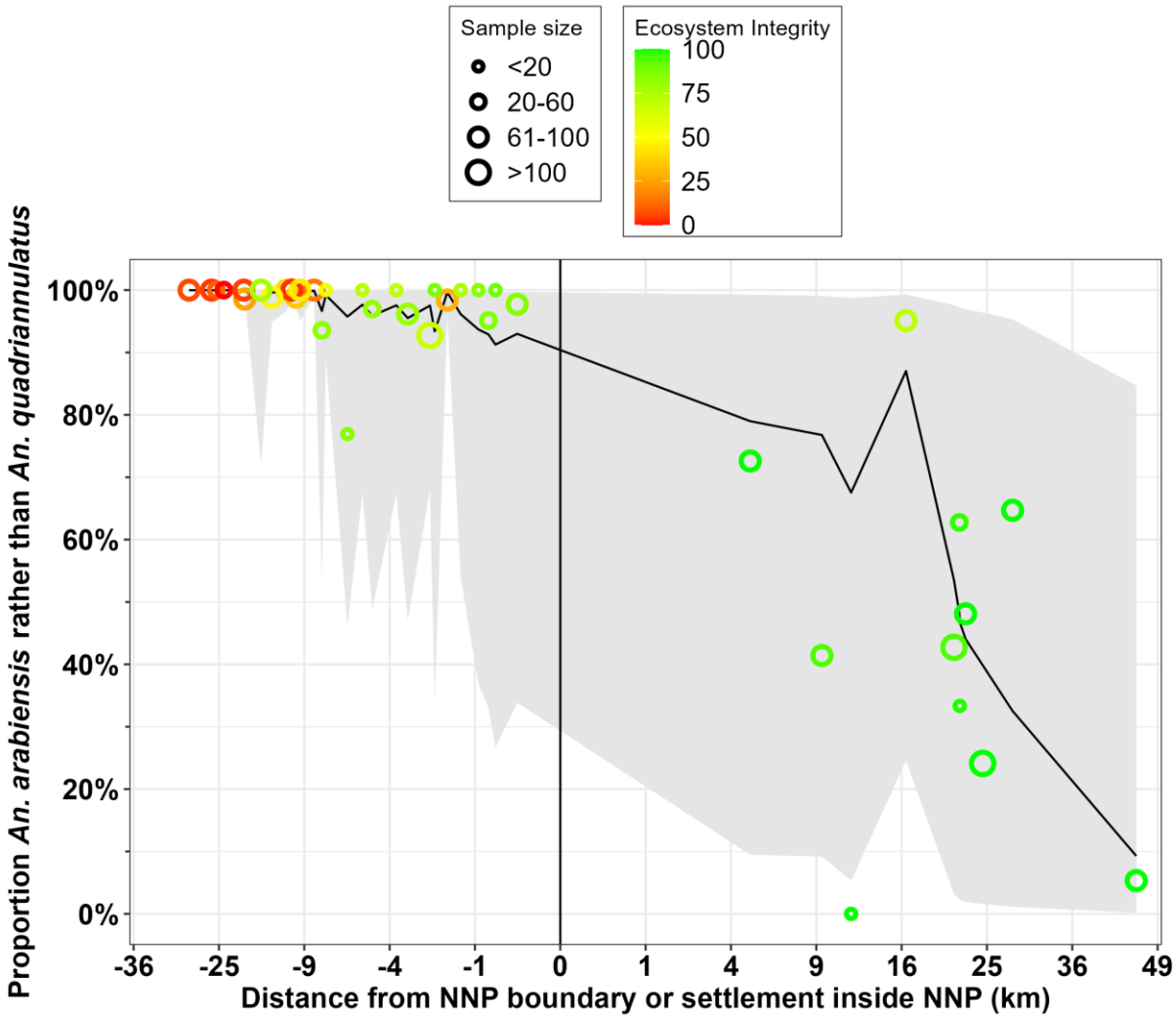

Figure 3.5. The proportion of field-identified *An. gambiae* complex collected and preserved *in situ* in the fourth round with the revised protocol as described in section 2.4.3 that were identified as *An. arabiensis* rather than *An. quadriannulatus* by PCR (Scott, Brogdon and Collins, 1993). The predicted mean values and their 95% confidence intervals for each surveyed camp location are plotted as a black line and grey ribbon, and were calculated based on the estimated intercept and the fixed effects of distance and subjective natural ecosystem integrity index (SNEII) from the final multivariate model detailed in Table 3.3.

The contents of figure 3.5 are represented cartographically in figure 3.6, illustrating clearly how *An. arabiensis* dominate fully domesticated habitats with villages and all throughout ILUMA WMA, even though very low proportions (<7%) of *An. quadriannulatus* occurred in its most pristine areas, generally along the NNP boundary.

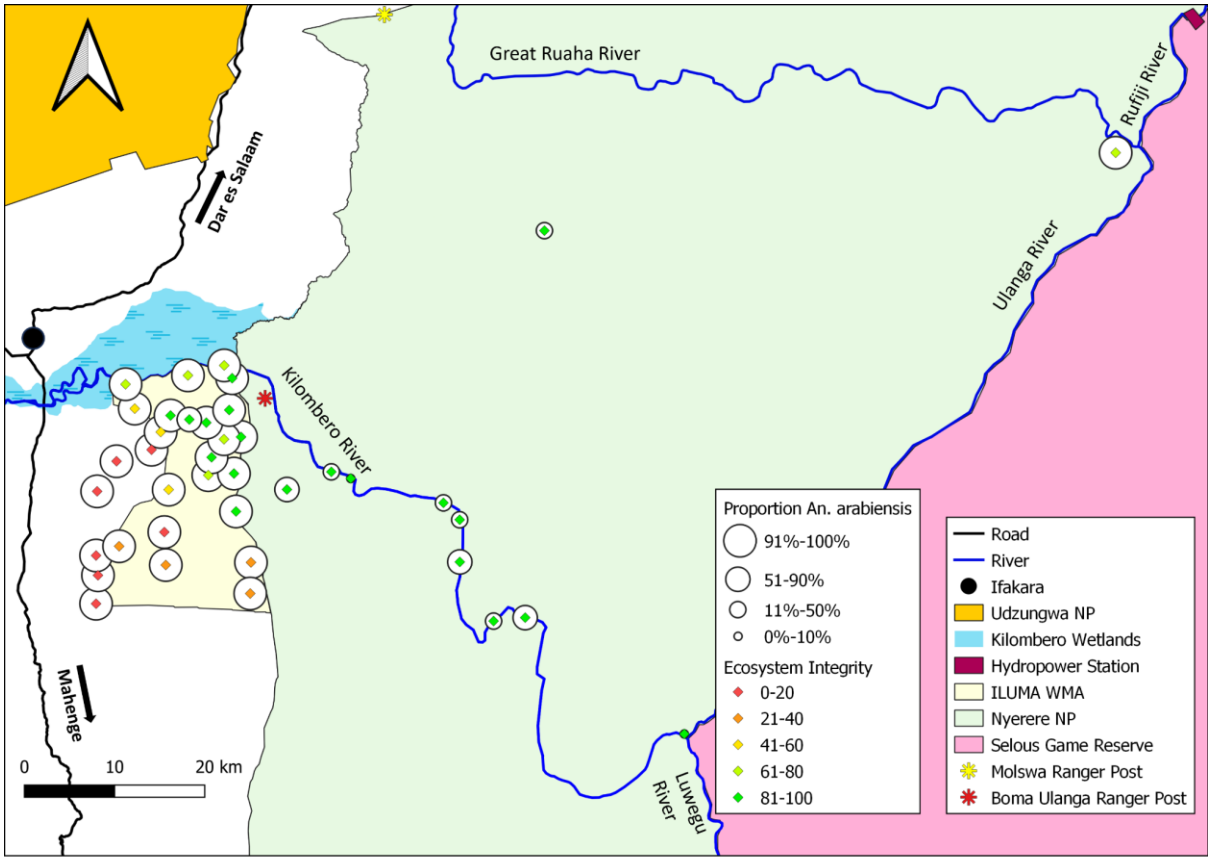

Figure 3.6. Map of ILUMA WMA and NNP illustrating how the proportion of *An. arabiensis* versus *An. quadriannulatus* varies geographically with respect to distance inside the park and away from human settlements, as well as the subjective natural ecosystem integrity (SNEII) at each camp location.

A conspicuous change in species composition was observed at locations inside NNP, all of which were at least 5km from the nearest boundary of NNP itself or that of an adjoining part of the Selous Game Reserve or Mikumi National Park (Figures 3.5 and 3.6). The SNEII reaches its maximum value at most camps inside NNP (Table 2.2 in section 2.6.1) and most of these had much higher proportions of *An. quadriannulatus* than any other camp in ILUMA and of the neighbouring villages to the west (Figures 3.5 and 3.6). Nevertheless, although the proportions of *An. arabiensis* generally declined with increasing distance inside NNP and away from substantial permanent human settlements (Figures 3.5 and 3.6), that geographic trend was remarkably erratic with lots of variability and with wide confidence intervals (Figure 3.5), indicating considerable uncertainty regarding predictions of the fitted model from table 3.3.

A clear outlier deep within NNP, *Bwawa la Umeme* (Figure 2.9, camp number 40), is a useful exception that proves the rule because it had the highest proportions of *An. arabiensis* (95%) and the lowest SNEII score inside the national park (Figure 3.5 and 3.6). This camp was located approximately 16km from the hydropower station and a large permanent human settlement within the national park, and had also been exposed to intensive human activity such as deforestation over the previous two years, suggesting that this population may be able to readily access human blood meals.

In contrast, *Shughuli Kubwa* (Figure 2.9; Camp 38), was located 47km from the nearest boundary of NNP other than its border with the adjoining Selous Game Reserve, representing the furthest surveyed camp away from major human settlements. At this exceptionally pristine location, only 5 of 76 (6%) of identified specimens were *An. arabiensis* (Figures 3.5 and 3.6), indicating a clear competitive advantage of *An. quadriannulatus* in this wild area. Nevertheless, it is notable that even in this extraordinarily pristine location, *An. arabiensis* persisted in small numbers. Indeed, camp number 30, *Bwawa la Kiboko Zanzibar* (Figure 2.9) which had a very small sample size (N=3), was the only camp where there were no *An. arabiensis* (Figures 3.5 and 3.6), so it seems that this species was ubiquitous across all surveyed pristine parts of the park.

While these data in themselves suggest that *An. arabiensis* may be able to exploit at least some wild animal species as sources of blood, and the NNP rangers could not recall ever having seen any signs of poaching or other unauthorized human activities as far downstream as *Shuguli Kubwa*, it is noteworthy that several small but permanently occupied ranger posts and tourist camps were scattered along the banks of the Kilombero and Ulanga Rivers, including one only 11km upstream from *Shuguli Kubwa* (Figure 2.9, Camp 38). While these minor outposts were all very small (Between 2 and 30 occupants at any given time), they nevertheless represented perennial sources of human blood that are within the regular flight ranges of *Anopheles*

mosquitoes, relative to the sampling frame depicted in figure 2.9. It was therefore not possible to rule out the possibility that the persistence of *An. arabiensis* in even the most pristine parts of the park, arises from the year-round availability of human hosts at fixed locations, albeit in a remarkably sparse and scattered population. As explained in the discussion, however, a subsequent complementary study confirmed the presence of *An. arabiensis* in the even more remote location of Mseguni (Figure 2.9), with no such small outposts of human existence nearby (Kavishe *et. al.*, Unpublished). It therefore seems likely that they can survive in complete wilderness by availing of blood from species other than humans and cattle.

3.4.4 Associations between An. gambiae complex species composition and surveyed indicators of potential host availability.

Note, however, that the predicted means calculated from the GLMM detailed in table 3.3 have wide confidence intervals (Figure 3.5), suggesting that the notable variability of *An. gambiae* complex species composition within the park may arise from heterogeneities in more specific factors other than distance and ecosystem integrity, such as blood resource availability and distribution. The following analysis was therefore undertaken to determine whether the competitive balance between *An. arabiensis* and *An. quadriannulatus* could be attributed to surveyable indicators of the activities of all the various mammalian species present in the study area, including humans and cattle.

Cattle herds and humans were initially examined as two distinct variables, but convergence failures during the multivariate analysis indicated that these variables were highly correlated, as confirmed in Figure 3.7. Fortunately, both humans and cattle are known to be strongly preferred hosts of *An. arabiensis* (White, 1974; Gillies and Coetzee, 1987; Killeen *et al.*, 2001), so a new variable was calculated for the combined total number of people and cattle herds

detected. This variable yielded a lower AIC value and was therefore treated as a fixed effect variable in the following univariate and multivariate analyses of the influence of various, diverse potential host species on the sibling species composition of the *An. gambiae* complex.

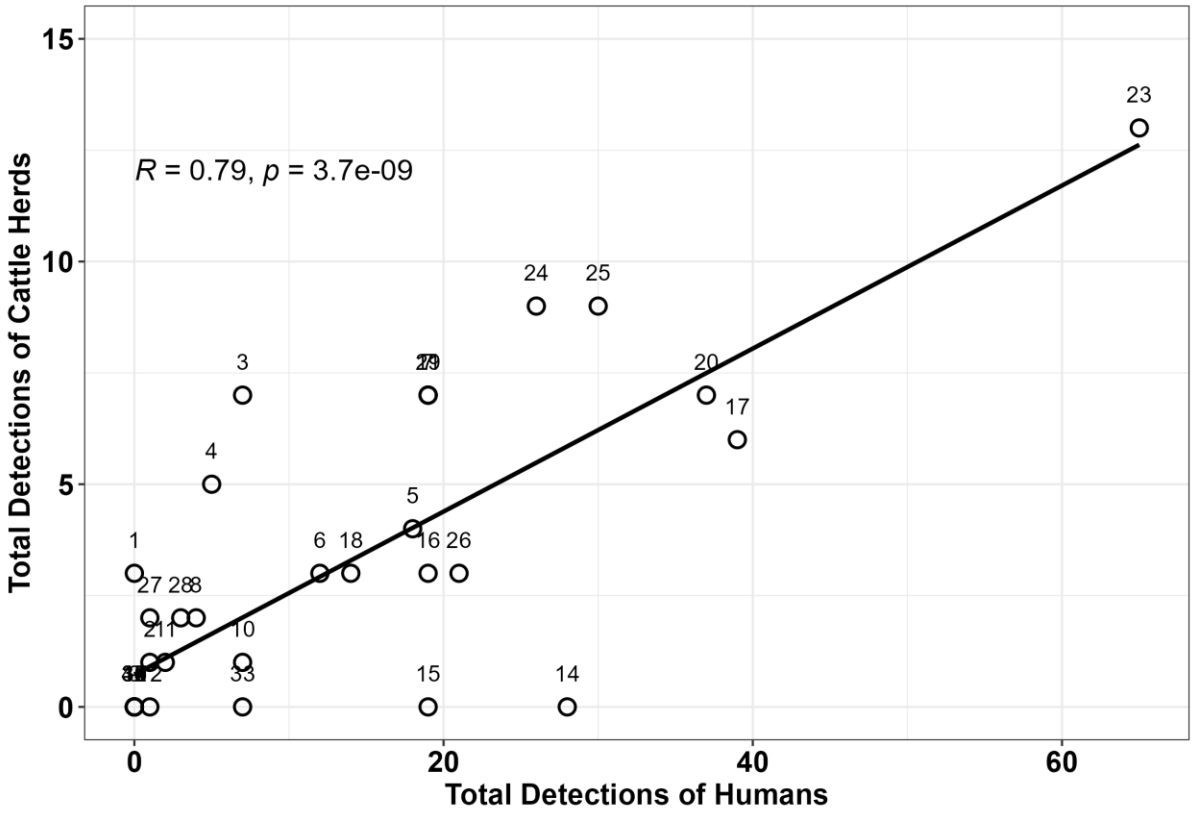

Figure 3.7. The detection frequency of humans plotted against the detection frequency of cattle herds at each camp number demonstrating a strong linear correlation, as tested using a Pearson's linear correlation test.

The best fit model from multivariate analysis indicates that cattle and humans, impala and bushpig are all mammalian species that appear to influence the proportions of *An. arabiensis* and *An. quadriannulatus* (Table 3.4). Hippos and waterbuck appeared to be significant covariates in the univariate analysis but were later excluded from the multivariate models once impala was accounted for (Table 3.4), indicating covariance between these three species, all of which were detected most frequently around waterbodies inside NNP. Similarly, common duiker was also excluded from the multivariate model once bushpig had been added (Table 3.4), again suggesting covariance between these two species, probably arising from their

common preference for moister woodlands and forests (Estes, 2012). Slender mongoose, baboons and warthog were also dropped from the multivariate analysis (Table 3.4) for similar statistical reasons, suggesting that their geographical distribution covaried with that of the species retained within the model. From all the wild animal species given in the key (Appendix 2), only red duiker, wildebeest, Sharpe's grysbok, klipspringer, African wild dog, side-striped jackal, banded mongoose, white-tailed mongoose, honey badger, African clawless otter and porcupine were not detected during radial surveys in round 4 and are therefore not represented in table 3.4. Lions, hyena, and leopard were all considered to be closely associated with wild herbivore populations, albeit in complex ways, and to contribute negligibly to overall mammalian biomass (Section 2.9.4), so they were not considered for inclusion in the multivariate models even though they were associated with *An. gambiae* complex population composition in the univariate analysis (Table 3.4).

Table 3.4. Univariate and multivariate outputs of generalised linear mixed models (GLMMs) of the effects of recorded detections of individual animal species on the proportion of *An. arabiensis* rather than *An. quadriannulatus*. All models were fitted to a binomial distribution with logit link function to the dependent variable. A random effect comprised of breeding site identification number nested within camp number was included to account for variation between and covariance within larval populations at individual waterbodies. Statistically significant effects are highlighted in bold.

	Univariate			**Multivariate**		
Fixed Effects	OR [95% CI]	t	P	Proportion[a] or OR[b] [95% CI]	t	P
Intercept	NA	NA	NA	0.98 [0.97, 0.99][a]	8.82	<<0.0001
Humans and Cattle Herds	**804 [65, 9910]**	**5.22**	**<<0.0001**	**45.5 [5.0, 416.6][b]**	**3.38**	**0.0007[c]**
Herbivores						
Hippo	**0.10 [0.04, 0.22]**	**-5.84**	**<<0.0001**	0.47 [0.21, 1.08][b]	-1.79	0.0742[d]
Impala	**0.11 [0.05, 0.23]**	**-5.62**	**<<0.0001**	**0.29 [0.15, 0.45][b]**	**-4.69**	**<0.0001[c]**
Waterbuck	**0.20 [0.08, 0.53]**	**-3.22**	**0.0013**	1.04 [0.58, 1.88][b]	0.15	0.8798[d]
Warthog	**0.29 [0.10, 0.79]**	**-2.41**	**0.0160**	1.59 [0.94, 2.69][b]	-1.73	0.0834[d]
Common Duiker	**5.31 [1.26, 22.89]**	**2.28**	**0.0225**	2.20 [0.88, 5.49][b]	1.69	0.0915[d]
Bushpig	**0.30 [0.10, 0.86]**	**-2.23**	**0.0258**	**0.51 [0.29, 0.86][b]**	**-2.39**	**0.0167[c]**
Eland	0.39 [0.14, 1.10]	-1.77	0.0763	NE	NE	NE
Buffalo	0.41 [0.15, 1.14]	-1.71	0.0883	NE	NE	NE
Elephant	0.44 [0.15, 1.22]	-1.57	0.1160	NE	NE	NE
Suni	1.97 [0.62, 6.34]	1.14	0.2530	NE	NE	NE
Hartebeest	0.54 [0.17, 1.75]	-1.02	0.3060	NE	NE	NE
Zebra	1.34 [0.50, 3.58]	0.59	0.5560	NE	NE	NE
Bushbuck	0.73 [0.23, 2.25]	-0.56	0.5790	NE	NE	NE
Sable	0.75 [0.25, 2.27]	-0.51	0.6080	NE	NE	NE
Dik-dik	0.78 [0.22, 2.73]	-0.39	0.6990	NE	NE	NE
Reedbuck	0.99 [0.33, 2.98]	-0.02	0.9810	NE	NE	NE
Large Carnivores						
Hyena	**0.17 [0.06, 0.45]**	**-3.58**	**0.0003**	NE	NE	NE
Leopard	**0.29 [0.11, 0.76]**	**-0.49**	**0.0127**	NE	NE	NE
Lion	0.52 [0.18, 1.53]	-1.18	0.2350	NE	NE	NE
Small Carnivores						
Slender mongoose	**0.20 [0.07, 0.61]**	**-2.84**	**0.0045**	0.96 [0.47, 1.94][b]	-0.12	0.9026[d]
Water mongoose	2.13 [0.84, 5.41	1.59	0.1110	NE	NE	NE
African wildcat	0.47 [0.15, 1.42]	-1.35	0.1780	NE	NE	NE
Genet	0.62 [0.17, 2.28]	-0.73	0.4680	NE	NE	NE
Civet	0.75 [0.28, 2.02]	-0.57	0.5720	NE	NE	NE
Primates and Prosimians						
Baboon	**0.24 [0.09, 0.66]**	**-2.78**	**0.0055**	0.75 [0.45, 1.25][b]	-1.10	0.2731[d]
Vervet monkey	1.84 [0.41, 8.34]	0.79	0.4290	NE	NE	NE
Greater galago	0.75 [0.33, 1.73]	-0.65	0.8310	NE	NE	NE
Blue monkey	1.13 [0.36, 3.58]	0.21	0.8320	NE	NE	NE
Macroscelidea						
Sengi	0.87 [0.29. 2.58]	-0.25	0.8020	NE	NE	NE
Tubulidentata						

Aardvark	0.60 [0.20, 1.78]	-0.81	0.4180	NE	NE	NE

Random Effects	Σ
Breeding site identification number/Camp number	0.868
Camp number	1.144

[a] Proportion identified as *An. arabiensis* based on the estimated intercept of the model.
[b] OR; Odds ratio per standard deviation increase from the mean.
95% CI; The 95% confidence interval estimated for the proportion or OR.
NA; Not applicable because several different intercepts estimated for more than one fitted univariate model, or not applicable to the reference group.
NE; not estimated for variables that were not-significant in the univariate analysis or large carnivores that were significant in the univariate analyses
[c] As estimated from the final best fit GLMM where the effect was included based on a significant value <0.05 and contributed to a lower AIC score.
[d] As estimated from the point at which the effect was no longer significant (>0.05) in the multivariate model fit and increased the AIC score and was therefore excluded from the final model.
σ standard deviation

Humans and cattle had the greatest effect on the relative abundances of *An. arabiensis* and *An. quadriannulatus* collected from larval habitats (Table 3.4). Consistent with the initial descriptive observations (Figure 3.3), such a high odds ratio indicates that where cattle or humans are present, populations of *An. gambiae* complex are overwhelmingly dominated by *An. arabiensis*. When their preferred blood sources become scarce or completely absent, however, proportions of *An. arabiensis* decline, seemingly losing its competitive advantage. Detections of humans or cattle were most frequently recorded in fully domesticated habitats, where the proportion of *An. arabiensis* was consistently greater than 91% (Figure 3.8). The highest number humans and cattle detections were in the villages outside the ILUMA western boundary, and both of these species were less frequently detected at larval habitat sites inside ILUMA. Nevertheless, these potential hosts were present across most of the WMA, including places that were a mere kilometre from the NNP boundary. Only two camps inside ILUMA WMA, *Bwawa la Chamvi* (camp number 9), and *Bwawa la Mrope* (camp number 13), were devoid of these hosts (Figure 3.8), however, given the extent of human and livestock encroachment at the surrounding camps, the high proportion of *An. arabiensis* found in these locations were likely within flying distance of the nearest human or cow. Only one camp inside NNP (*Kambi ya Mamba*, camp number 33) (Figure 2.9), exhibited any signs of the presence of

people (a few tracks of one or two fish poachers) and it is notable that it had higher proportions of *An. arabiensis* than most other camps inside NNP (Figure 3.8). This indicates that even a small number of humans may be enough to give *An. arabiensis* an advantage in a competitive relationship that they appear to dominate when allowed to. No detections of humans or cattle were found at *Bwawa la Umeme* (camp number 40), which was located 16km away from a large permanent human settlement (Figure 2.9), suggesting that the high proportions of *An. arabiensis* at this location (Figure 3.8) may be dispersing over long distances to acquire human blood.

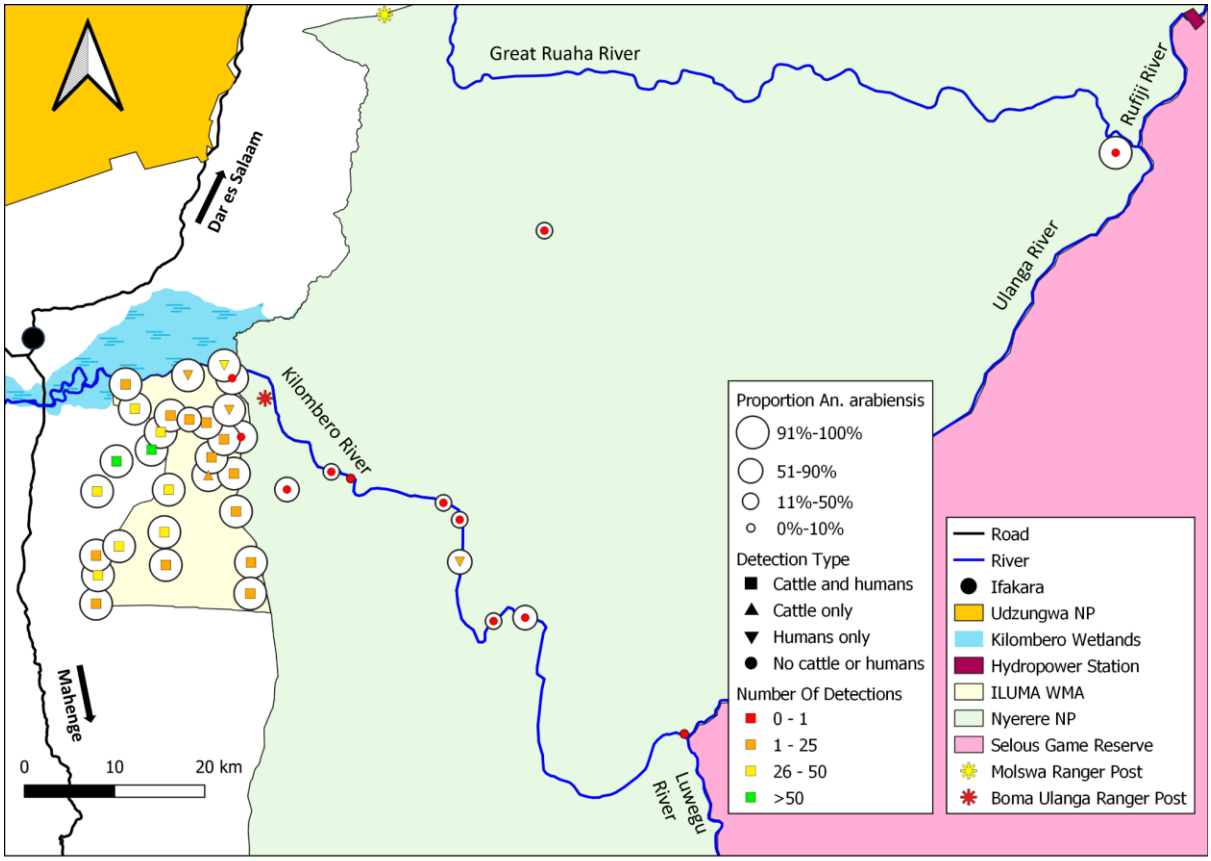

Figure 3.8. Map of ILUMA WMA and NNP illustrating how the proportion of *An. arabiensis* versus *An. quadriannulatus* varies geographically with respect to the number of cattle and/or human detections at each camp.

Interestingly, the relative abundance of *An. arabiensis* was strongly and negatively associated with detections of impala (Table 3.4). Impala were the predominant and most abundant antelope recorded inside NNP, particularly at camps along the Kilombero river where An. *quadriannulatus* were, on average, just as common as or more abundant than *An. arabiensis* (Figures 3.9 and 3.11), suggesting that this common antelope of dry savanna mosaic habitats (Estes, 2012), may well be a preferred blood source for *An. quadriannulatus*.

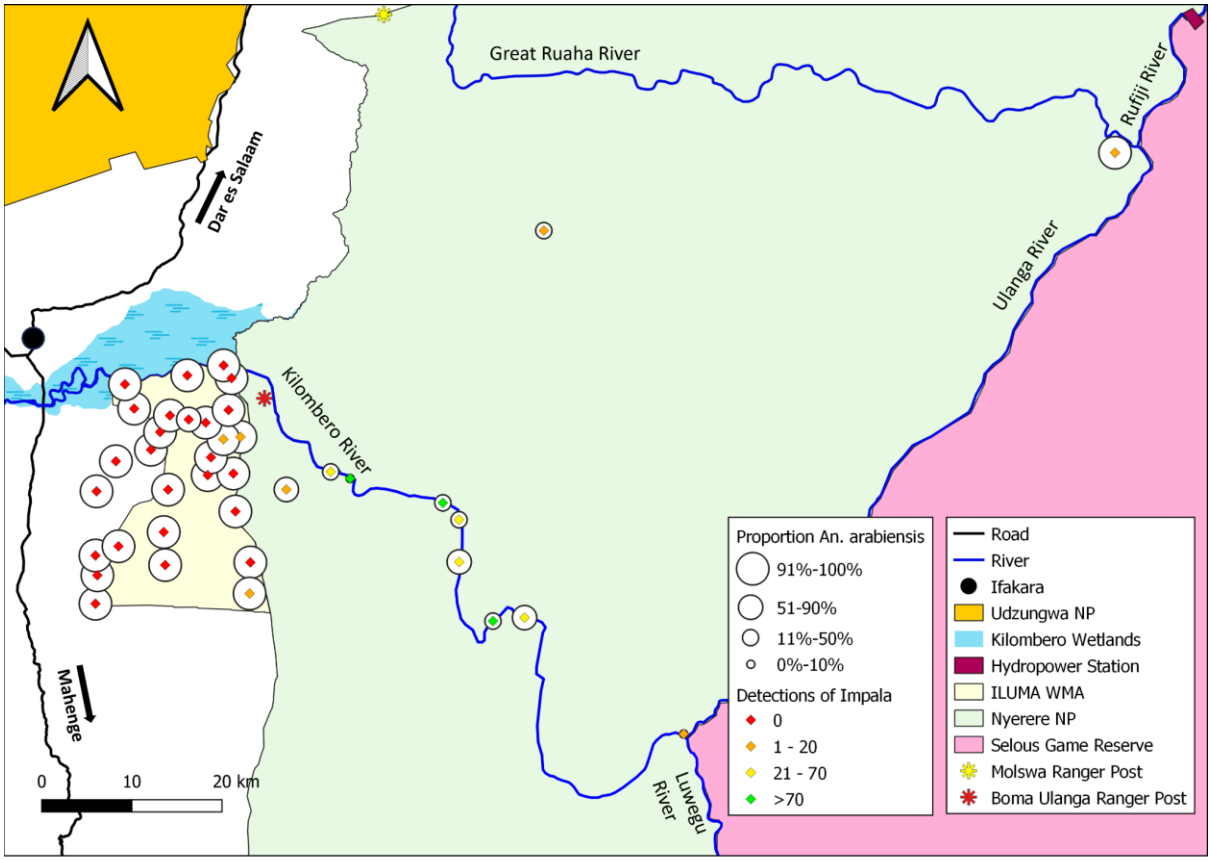

Figure 3.9. Map of ILUMA WMA and NNP illustrating how the proportion of *An. arabiensis* versus *An. quadriannulatus* varies geographically with respect to the number of impala detections at each camp.

The abundance of bushpig also seemed to be negatively associated with the relative abundance of *An. arabiensis* (Table 3.3), despite being detected far less frequently than either impala, humans or cattle, suggesting that bushpigs may provide a second preferred blood source for *An. quadriannulatus*. Bushpigs were most commonly detected in the groundwater forests of the north-eastern parts of ILUMA WMA and in the encroached areas at the interfaces between

humans and wildlife (Figure 3.10), indeed with the three highest number of detections occurring at *Mskamba* (camp number 1), *Bwawa la Semka* (camp number 28) and *Bwawa la Nandete* (camp number 4) (Figures 2.9 and 3.10), all of which have moderate to high SNEII scores (Figure 3.6), and were experiencing or recovering from encroachment at the time. Although only small proportions (<10%) of *An. quadriannulatus* were observed at these locations (Figure 3.10 and 3.11), they were still present despite these areas also exhibiting plenty of evidence for human and/or cattle activity (Figures 3.8 and 3.11). This suggests that bushpigs may sustain small populations of *An. quadriannulatus* that might not otherwise persist in the fringe areas where humans, livestock and wildlife co-occur.

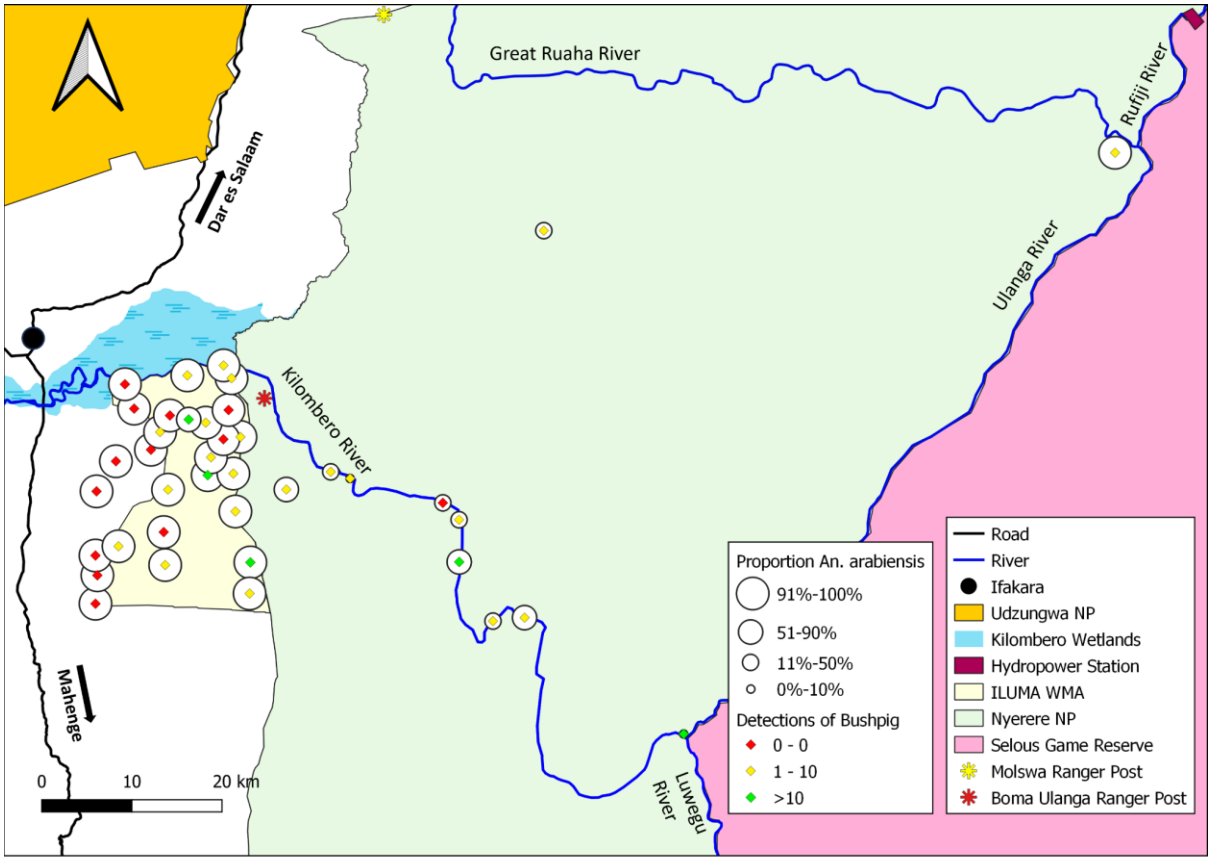

Figure 3.10. Map of ILUMA WMA and NNP illustrating how the proportion of *An. arabiensis* versus *An. quadriannulatus* varies geographically with respect to the number of bushpig detections at each camp.

The estimated effect size for combined detections of cattle and humans on the relative abundance of *An. arabiensis* is much greater than the effect sizes estimated for impala or bushpig (Table 3.4), suggesting that *An. arabiensis* have a very strong competitive advantage over *An. quadriannulatus* when their preferred blood sources (cattle and humans) are present even at low densities, only allowing *An. quadriannulatus* to compete with it in wild areas far away from human settlement and livestock.

Taken at face value in mathematical terms, such a considerable effect size for humans and cattle (Table 3.4) would suggest that *An. arabiensis* might be expected to disappear in the complete absence of their two known preferred hosts. Instead, however, they persist at relative abundances that are generally readily detectable and usually exceed 20% deep inside NNP (Figure 3.11). Even at the furthest possible distance from any human settlement (Figure 3.11), five *An. arabiensis* larvae were identified and each were collected from a different aquatic habitat, suggesting that it is unlikely for these multiple specimens to have hatched from eggs laid by the same mother that had fed on a person or cow, and then flown almost 50km before oviposition. Therefore, it may well be likely that these specimens are from one or more self-sustaining wild populations that are feeding on one or more alternative wild hosts that have not been captured by the best fit GLMM in Table 3.4.

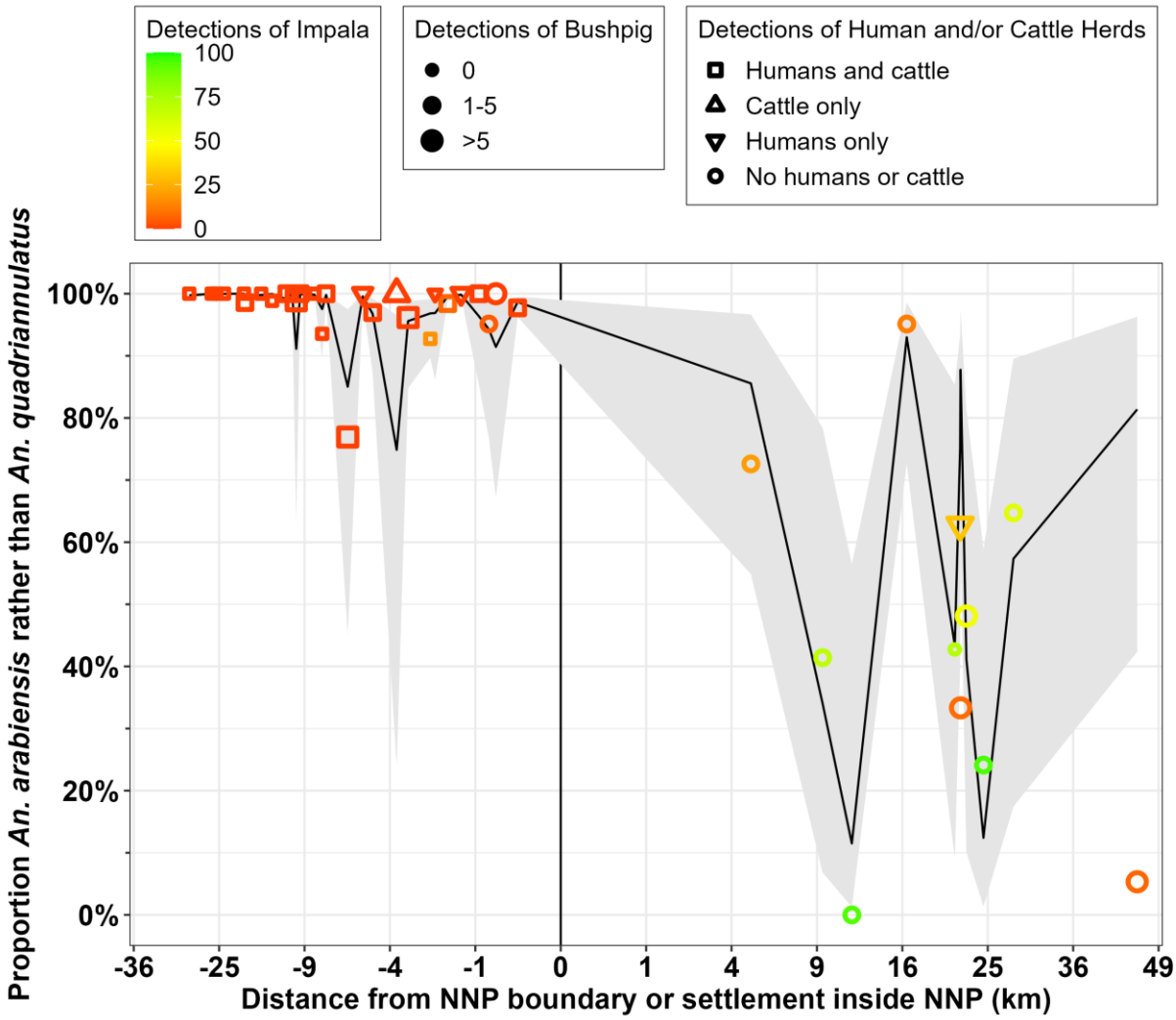

Figure 3.11. The proportion of field-identified *An. gambiae* complex collected and preserved *in situ* in the fourth round with the revised protocol as described in section 2.4.3 that were identified as *An. arabiensis* rather than *An. quadriannulatus* by PCR (Scott, Brogdon and Collins, 1993). The predicted mean values and their 95% confidence intervals for each surveyed camp location are plotted as a black line and grey ribbon, and were calculated based on the total number of detections for cattle and humans, impala and bushpig and the generalized linear mixed model (GLMM) presented in Table 3.4.

Reassuringly, the fluctuations of relative abundance between these two mosquito species across a gradient of fully domesticated to pristine natural ecosystems, is far better explained by data from surveys of mammalian activity (Figure 3.11), than by broader geographic parameters, namely distance and the SNEII (Figure 3.5). The calculated predicted means from the final host species model that captures most of the considerable variability between camps within NNP and yields for narrower confidence intervals (Figure 3.11) than those in figure 3.5. The one

major outlier is also informative: the model prediction also failed to capture the low proportions of *An. arabiensis* and high proportions of *An. quadriannulatus* inside NNP recorded at camp number 38, Shughuli Kubwa (Figure 2.9), where there appeared to be few impala and bushpig (Figure 3.11), and so *An. arabiensis* should predominate this location. This suggests that populations of *An. quadriannulatus* are feeding on other wild animals and that impala and bushpig are not the only hosts available to *An. quadriannulatus* in pristine ecosystems.

Chapter 4

4. Discussion

The original aims of the study were to determine occupancy rates of aquatic habitats with *An. gambiae* complex larvae across a gradient of natural ecosystem integrity, to demonstrate that *An. arabiensis* occupy aquatic habitats in pristine areas and to try to associate these populations with the availability of specific potential blood host species. These findings emphasized the ecology of *An. arabiensis* mosquitoes in understudied natural environments including a buffer zone where humans and livestock occur alongside wildlife, and in remote areas of untouched wilderness. Although no specific mammal species could be identified as a potential alternative blood source to cattle or humans that *An. arabiensis* are known to prefer, the results do clearly demonstrate a relationship between *An. arabiensis* and a second sibling species from the same complex *An. quadriannulatus*, which it co-exists within well conserved natural ecosystems. Notably, while *An. arabiensis* is a vector of residual malaria transmission (Russell *et al.*, 2010; Bayoh *et al.*, 2010; Durnez and Coosemans, 2013; Lwetoijera *et al.*, 2014), *An. quadriannulatus* is generally considered to be a non-vector (White, 1974; Coluzzi *et al.*, 1979; Gillies and Coetzee, 1987), and this study inferred impala and bushpig as likely preferred blood sources for this relatively harmless species, allowing it to partially displace its more dangerous sibling in areas where these mammals occur.

4.1 Influence of the surveyor and habitat attributes on larval occupancy.

The remarkably large differences in reported occupancy rates between the investigator and the trained VGS member may have been influenced by several factors including subtle variations in sampling methodology and technique. Mosquito larvae are extremely sensitive to predation threats and so if disturbed, they will rapidly swim downwards from the water surface or sidewards along it. Simple technique adjustments, like avoiding casting shadows across the

habitat before sampling it, and improving the angle and motion of dipper deployment (O'Malley, 1989; Service, 1993) may have increased the odds of *An. gambiae* complex larvae capture. More specifically, it was notable that the VGS member also often meticulously investigated the habitat perimeter prior to dipping, thus increasing detection sensitivity by dipping at points where larvae were already observed. While this could of course introduce biases of its own, including the additional year of surveys completed by the VGS volunteer improved the statistical power of the model fit, indicating that the patterns exhibited were consistent across these two years with two different surveyors. It should be noted however, that such large differences between different individual surveyors are quite normal (Chaki *et al.*, 2009; Chaki *et al.*, 2011) as increases in larval habitat detection among *community-owned resource persons* has previously been clearly documented in an operational municipal larval source management programme in Dar es Salaam, Tanzania. Interestingly, volunteers that were familiar with the area and were recruited through local community leaders were better at detecting larval habitats than those who were recruited by programme leaders and administrative staff (Chaki *et al.*, 2011).

Pools, puddles, tracks, and depressions were the most frequently surveyed habitat category encompassing a wide range of different waterbodies with similar appearances. Although this broad habitat category was available year-round, it was observed that specific waterbodies that were available within this category were influenced both seasonally and geographically. For example, during the wet season, this category included any surface-water pools of accumulated rainwater and any human, animal, vehicle tracks or other indentations formed in sodden soil. During the dry season, surface-water under this category was limited to stagnant isolated pools formed by retreating streambeds, which were observed to be the predominant habitat type during the driest months when waterholes, swamps, flooded valleys, and rice paddies dried up.

The habitat attributes found to favour *An. gambiae* complex larvae in this study were consistent with those reported by classic literature, which includes small, open sunlit habitats (Holstein, 1954; Gillies and De Meillon, 1968). Although habitat occupancy was higher in habitats with vegetation, it was observed to be often sparsely distributed. Rice paddies positive for *An. gambiae* complex in this study were typically in the early stages of cultivation when the crop height was low, an observation that was also consistent with those reported in Gillies and De Mellion, 1968). Furthermore, as per the protocol and to avoid trampling the crop, dips were taken along the margins of the field that would have been more exposed to the sun and provided optimal conditions for developing larvae (Holstein, 1954; Gillies and De Meillon, 1968).

The results showed that rice paddies, waterholes and small habitats (<2m), were all characteristic attributes of higher *An. gambiae* complex occupancy rates compared to any other habitat type (Table 3.2). As there was no difference in occupancy rates between fully domesticated settlements outside ILUMA WMA and inside NNP (Table 3.2) and no rice paddies exist in NNP because the area is under strict conservation, this indicates that waterholes and the surrounding prints of wild animals may well be predominant breeding sites for wild populations of *An. gambiae* complex mosquitoes. Therefore, to conduct further research on the aquatic stages or, when attempting to catch adult mosquitoes in wild areas far from human activity, it is recommended to set up camp near waterholes rather than alongside other habitats such as flowing streambeds, which proved to be occupied less (Table 3.2). However, although not detected in the model, it was also observed by both investigators that when water levels were low during the dry season, small pools found on the surfaces of exposed rocky outcrops in riverbeds were often occupied with *An. gambiae* larvae, particularly along the Kilombero River deep inside NNP.

The overall occupancy rate estimated from this study was 54% for all *Anopheles* larvae and 39% for *An. gambiae* complex larvae, which was reasonably consistent with some other

reported occupancy rates, in rural towns or settlements over the last two decades, which ranged from about 35% to 70% (Gimnig *et al.*, 2001; Fillinger *et al.*, 2004; Carlson, Byrd and Omlin, 2004; Fillinger and Lindsay, 2006; Shililu *et al.*, 2007; Mwangangi *et al.*, 2007; Kenea, Balkew and Gebre-Michael, 2011; Kweka *et al.*, 2012; Gimnig *et al.*, 2020; Epopa *et al.*, 2020). For example, Fillinger *et al.* (2004) identified 67% of surveyed aquatic habitats as being occupied with *An. gambiae* complex larvae, while a second study in Kenya also demonstrated a similar occupancy rate of 51% for all *Anopheles* larvae (Fillinger and Lindsay, 2006). More recently, a study by Epopa *et al.* (2020) demonstrated anopheline occupancy rates of 46% which was conducted in an area of low human density between two villages. Indeed, because the presence of anopheline larvae in aquatic habitats is highly ubiquitous (Holstein, 1954), and the larval ecology varies between anopheline species (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987), even within the *An. gambiae* complex (Diabaté *et al.*, 2005; Simard *et al.*, 2009; Gimonneau *et al.*, 2012; Diabaté *et al.*, 2008), one can't totally compare or over interpret differences in the reported occupancy rates of anopheline larvae between these studies. Additionally, as larval surveys are ordinarily conducted in areas of human settlement, the distribution of sampled breeding sites is biased towards man-made habitats that represent productive breeding sites, so natural habitats far away from people may not be fully represented.

4.*2 Influence of spatial and geographic factors on larval occupancy.*

Although the overall occupancy rates were relatively consistent with past and present reports, this study demonstrated for the first time, a modest but steady decrease in occupancy rates over a gradient of ecosystem integrity (Figure 3.2) that extended almost 80km inside a newly designated and as of now, undeveloped national park with negligible human intrusion. The unanticipated extension of the sampling frame into NNP at a later stage of the study, meant that larval surveys at these camps were only replicated once or twice compared to the camp

locations inside ILUMA WMA and in the villages along its western boundary, that were part of the original protocol and almost all surveyed three or four times (Table 2.1). Correspondingly, less occupancy data was available for these additional camps that generally represented the furthest distances from human settlement and had the highest SNEII scores overall. However, the statistical power of the final fitted model in table 3.2, as indicated by the low p-values and narrow confidence intervals, suggests that although the number of replicates was somewhat lower for the most pristine camps, the number of habitats surveyed at each camp were sufficient to give a robust model fit.

The SNEII is a variable that captures the overall general condition of the ecological state of an area and therefore reflects a myriad of other more specific factors that could act as hidden drivers of lower occupancy rates in pristine locations. In particular, it is closely associated with the distance from human settlement, with the lowest SNEII scores assigned to those in fully domesticated settlements and higher SNEII scores generally assigned much further away from permanent human settlements, with the highest being in NNP. In other settings, aquatic habitats that are located closer to human settlements have been found to have higher occupancy and larval density than those located further away (Charlwood and Edoh, 1996; Minakawa, Seda and Yan, 2002; Epopa *et al.*, 2020), indicating that proximity to known preferred blood sources influences the rates at which *An. gambiae* complex use suitable habitats. However, these studies were conducted across scales of hundreds of metres, while the study reported herein assesses occupancy trends spanning up to a total of 70km of distances inside or outside of NNP. Reduced occupancy observed in areas furthest away from humans and their cattle is plausible for *An. arabiensis* given these are their only two known preferred blood sources (White, 1974; Gillies and Coetzee, 1987; Killeen *et al.*, 2001). However, although occupancy of *An. gambiae* complex decline overall, aquatic habitats in camps with higher SNEII scores were also

occupied by a second species of the same complex, *An. quadriannulatus*, with increasing abundance relative to *An. arabiensis*.

4.3 Distribution of An. quadriannulatus.

As *An. arabiensis* was thought to be the only persisting member of the *An. gambiae* complex in the Kilombero valley since the elimination of *An. gambiae s.s.* by LLINs circa 2010 (Russell *et al.*, 2010; Killeen *et al.*, 2013), it was assumed from the onset that it would be the sole sibling species present, even in the most pristine corners of ILUMA WMA, because of its ability to exhibit zoophagic tendencies. Instead, the data revealed the presence of the much more zoophagic and relatively harmless species, *An. quadriannulatus* (White, 1974; Coluzzi *et al.*, 1979; Coluzzi, 1984; Gillies and Coetzee, 1987) which was completely unanticipated given that this species had only once been identified over the long history of mosquito research in the Kilombero valley (Maia *et al.*, 2016). However, the sympatric populations of the two sibling species reported herein are consistent with the historical literature across Africa, which indicated that *An. quadriannulatus* occur in the same geographic regions as *An. arabiensis* (White, 1974). However, the distribution of the former is much patchier than the latter (White, 1974; Coluzzi *et al.*, 1979).

Originally identified as a separate species by Davidson (1964), populations of *An. quadriannulatus* were initially reported in areas of southeast Africa including Zimbabwe, Zambia, Eswatini and South Africa (Davidson, 1964; White, 1974; Gillies and Coetzee, 1987). Soon after, populations exhibiting more endophilic tendencies were discovered in Ethiopa (White, 1974), which have since been recognised as a separate species, *An. quadriannulatus* species B (Hunt, Coetzee and Fettene, 1998), and has since been renamed as *An. amharcus* (Coetzee *et al.*, 2013). The patchy distribution of these two allopatric species are therefore thought to be the relics of a zoophagic, exophilic, ancestral species that was found throughout the ancient rainforests of Africa (Ayala and Coluzzi, 2005). In Tanzania, records of *An.*

quadriannulatus populations are lacking almost entirely. Documented originally on the island of Zanzibar (Odetoyinbo and Davidson, 1968), this species has reportedly been observed there again since, but no further published data of *An. quadriannulatus* on the island is available (Coetzee *et al.*, 2013). On the mainland, records are limited to an apparent report from Mbozi in southwestern Tanzania, which is referred to in Kweka *et al.* (2020), and a very recent study that revealed the presence of an *An. quadriannulatus* population at Lake Manyara, an arid region of the Rift Valley in northern Tanzania (Kweka *et al.*, 2020).

In general, studies of *An. gambiae* complex are conducted to understand the ecology of the most important malaria vectors and are therefore most frequently carried out in inhabited areas, where human-biting mosquitoes of epidemiological importance exist and interact with abundant human populations. Although behaviourally plastic malaria vectors like *An. arabiensis* feature strongly, highly zoophagic species like *An. quadriannulatus* may be underrepresented. As demonstrated by this study, finding self-sustaining populations of *An. quadriannulatus* may require sampling of conserved wilderness areas at least 10km from the nearest human settlement. Hence, this species may have a wider distribution than previously appreciated because of their apparent inability to compete with *An. arabiensis* wherever humans and cattle abound, which finds them in remote, pristine areas. Furthermore, the observations perhaps shed light on the apparent scarcity of *An. quadriannulatus* across Africa over recent years when compared with historical records. When much of the classical literature on African anopheles was written half a century ago (White, 1974; Gillies and Coetzee, 1987), rural settlements across the continent, were generally speaking, fewer, smaller and more often surrounded by natural ecosystems with healthy wildlife populations. Today, the opposite is true: Increasingly pristine wilderness is found only inside protected areas that are surrounded by growing populations of humans and livestock. It is therefore plausible that the already relict range of *An. quadriannulatus* has contracted even more over that time.

4.4 Competitive co-existence of An. arabiensis and An. quadriannulatus populations

The presence of *An. quadriannulatus* in varying proportions also gave insights into the potential drivers of their ability to co-exist with *An. arabiensis* in pristine wilderness, despite their apparently competitive relationship (See Box 1 for discussion of the general principles of competitive co-existence and a brief overview of how it relates to the larvae of sympatric mosquito populations). According to Gillies and Coetzee (1987), the range of aquatic habitats used by all the various freshwater species of the *An. gambiae* complex appear to be very similar, without any obvious differences between sibling species. While some distinctions in habitat use have been documented more recently between other members of the same complex, namely *An. gambiae s.s.* and *An. coluzzii*, these are nevertheless relatively subtle (Diabaté *et al.*, 2005; Simard *et al.*, 2009). Interestingly, the initial discovery of *An. quadriannulatus* specifies wild animal footprints and shallow pools as the habitats they were typically found in (Paterson, Paterson and Van Eeden, 1963), which is remarkably consistent with the observed characteristics of occupied habitat types found in the most pristine parts of the study area described herein.

The co-occurrence of both *An. arabiensis* and *An. quadriannulatus* in the same habitats is not surprising, as sympatric species of the *An. gambiae* complex are often found sharing the same habitat (Charlwood and Edoh, 1996; Minakawa *et al.*, 1999; Gimnig *et al.*, 2001; Edillo *et al.*, 2002). The relative abundance of one species compared to another in shared habitats depends on their ability to compete with each other under local conditions, which in turn depends upon a variety of factors, such as the presence of predators, food availability and larval densities, that affect the development or survival of various life stages (Schneider, Takken and McCall, 2000; Diabaté *et al.*, 2008; Koenraadt and Takken, 2003). Furthermore, environmental variability can also influence larval competitive interactions, which may lead to composition

shifts over time and space (Schneider, Takken and McCall, 2000; Kirby and Lindsay, 2009; Munhenga *et al.*, 2014).

Although the habitats surveyed during this study may appear similar, *An. arabiensis* and *An. quadriannulatus* may have subtly distinctive environmental preferences within an aquatic habitat, determined by factors like water quality and temperature that may change considerably over temporal scales. For example, a study in Kruger National Park, South Africa yielded evidence of a seasonal species composition shift, from *An. arabiensis* predominating in the wet season to a sudden proliferation of *An. quadriannulatus* in the dry season, and suggested that the latter could withstand higher saline concentrations found in stagnant dry-season water bodies (Munhenga *et al.*, 2014). Furthermore this view was supported by the surprizing appearance of *An. merus*, a saltwater species from the same complex normally associated with coastal environments, at the same time (Munhenga *et al.*, 2014).

Additionally, while both *An. arabiensis* and *An. quadriannulatus* are both clearly adapted to arid conditions (White, 1974; Gillies and Coetzee, 1987; Moffett, Shackelford and Sarkar, 2007), the latter is considered to exhibit greater tolerance of cooler, more temperate conditions (White, 1974). This suggest that *An. quadriannulatus* may have temporal competitive advantages when weather conditions in NNP generally become cooler at the end of the rainy season and start of the dry season around May and June. Previous studies have shown that water temperature has considerable effects on mixed populations of *An. arabiensis* and *An. gambiae s.s.* larvae, with the forming exhibiting higher survival rates in high water temperatures >30°C (Kirby and Lindsay, 2009) but is outcompeted by the latter at lower temperatures of 30°C or less (Schneider, Takken and McCall, 2000; Kirby and Lindsay, 2009). This suggests that higher water temperatures and the hotter, drier environmental conditions of the acacia savannah inside NNP may provide *An. arabiensis* with a competitive edge in these pristine ecosystems.

Although interspecific larval competition may be a contributing factor to the fluctuating species composition observed in this study, it is also important to consider its full life cycle as each life stage is interdependent. For example, competitive interactions at larval stages influence adult population density and fitness (Gillies and Smith, 1960; Ho, Ewert and Chew, 1989; Juliano, 1998; Gimnig *et al.*, 2002; Muriu *et al.*, 2013; Barreaux *et al.*, 2018). Similarly, larvae cannot occupy habitats without a female first successfully acquiring a blood meal and then finding a suitable aquatic habitat for oviposition somewhere reasonably nearby. The ability of adult females to acquire such limiting resources should also thereby influence larvae population abundance and distribution, so it is perhaps unsurprising that the competitive balance between *An. arabiensis* and *An. quadriannulatus* larvae (Section 3.4.2), and between other pairs of species within the same species complex (Gillies and Smith, 1960; Smith, 1962; Charlwood and Edoh, 1996; Minakawa, Seda and Yan, 2002; Bayoh *et al.*, 2010; Mutuku *et al.*, 2011) are so clearly influenced by the availability of preferred blood hosts nearby, and by vector control measures that protect some of those hosts against attack.

The close association of *An. quadriannulatus* with two wild animal species (Table 3.4, Figure 3.11) suggests that the availability of preferred blood resources may either be a limiting or enabling factor for individual species that strongly influences mosquito population composition. And so, although *An. arabiensis* displace *An. quadriannulatus* wherever they can access their preferred hosts (cattle and humans), *An. quadriannulatus* does not completely displace *An. arabiensis* where these hosts are absent and impala and bushpig are present, but merely supresses their abundance. The influence of host availability on species composition ratios have previously been documented within the *An. gambiae s.l.* complex, between *An. arabiensis* and the highly human-specialised species *An. gambiae s.s.* More specifically, Charlwood and Edoh (1996) found that the relative abundance of *An. arabiensis* larvae was higher in aquatic habitats that were close to cattle, whereas Minakawa *et al.*, (2002)

demonstrated that higher densities *An. gambiae s.s.* were found in aquatic habitats close to houses. In view of this, it is suggested that the availability of blood meals influences adult population dynamics which, in turn, influences species ratio outcomes in larval habitats.

As discussed in detail in section 1.3, deliberate interference with mosquito access to human blood sources through LLIN and IRS interventions caused a species shift in the *An. gambiae* complex, so that the nominate species was competitively reduced (Lounibos, 2007) and eliminated across much of its range (Sinka *et al.*, 2016), notably across large tracts of east Africa, by *An. arabiensis* (Russell *et al.*, 2010; Bayoh *et al.*, 2010; Mwangangi *et al.*, 2013; Killeen *et al.*, 2013; Sougoufara *et al.*, 2016). Such distinct competitive reductions in response to vector intervention has also been previously documented for other malaria vectors. In Kenya, *An. funestus* were historically replaced by zoophagic and exophilic species from the same species group, namely *An. rivulorum* and *An. parensis*, following IRS that presumably affected the competitive balance in larval habitats (Gillies and Smith, 1960; Smith, 1962; Gillies and Furlong, 1964; Kawada *et al.*, 2012). Furthermore, one of the most important vectors of malaria in South and Central America, namely *An. darlingi*, was formerly eliminated in Guyana (formerly British Guiana), in response to an IRS event that subsequently created a species shift towards the more zoophilic vector, *An. aquasalis* (Giglioli, 1951).

These classic examples of shifting species composition in response to intervention tools were therefore directly facilitated by selectively protecting humans from attack. Once this preferred host of the originally dominant mosquito species was no longer readily accessible, more zoophilic species were released from competition (Qureshi and Connolly, 2021), thus increasing their density and abundance. Here, we document similar species composition shifts, but across a spatial rather than a temporal scale, which is clearly underpinned by geographic variations in the natural availability and distribution of different blood resources, two of which (humans and cattle) are closely associated with insecticide coverage.

4.5 Host availability and inferred host preference.

Blood source utilisation of *An. arabiensis* has been long-established, however the blood feeding patterns of *An. quadriannulatus* is less understood. The association between *An. arabiensis* with detections of cattle or human blood sources (Table 3.4) is consistent with the known host preference behaviours of this malaria vector (White, 1974; Gillies and De Meillon, 1968; Gillies and Coetzee, 1987; Killeen *et al.*, 2001). The remarkably high intercept and the effect of humans and cattle displayed in table 3.4, clearly captures the specialised feeding behaviours demonstrated by *An. arabiensis*, for which no evidence of any other significant host could be detected with fitted process explicit models (Killeen *et al.*, 2001). The consistency of these quantitatively dominant effects of humans and cattle (Table 3.4) with the known host preferences of *An. arabiensis* is reassuring with respect to the reliability of this methodological approach used.

As *An. quadriannulatus* are generally thought to be of little medical importance (White, 1974; Coluzzi *et al.*, 1979; Gillies and Coetzee, 1987), their blood feeding behaviours have rarely been studied and remain poorly understood. Where this vector occurs in domesticated areas, it predominantly feeds on cattle (White, 1974) but will also feed on humans in some cases (Pates *et al.*, 2001; Torr *et al.*, 2008; Seyoum *et al.*, 2012). According to the classic literature, it is thought that *An. quadriannulatus* primarily survive on wild animals (Gillies and Coetzee, 1987; White, 1974). However, conclusive bloodmeal analysis for this species is almost lacking entirely, presumably because they are of such little medical interest. A study by Prior and Torr (2002) inside Mana Pools National Park, Zimbabwe, identified mixed blood-meals in a few specimens that included DNA from cattle and other unidentified hosts that were presumed to be wild ungulates. However, it seems that a definite link between *An. quadriannulatus* and a specific wild host species has not yet been published. While the *a priori* intention of the host association model was to investigate the blood sources of the vector *An. arabiensis*, it also

provided insight into the feeding behaviours of this highly zoophagic and seemingly non-vector in some of the most pristine and remote parts of the largest national park in East Africa.

As detailed in table 3.4, impala were identified to be the most plausible, common blood source for these *An. quadriannulatus* populations which may be attributed to several factors associated with the ecology of this antelope. Impala were the most frequently detected antelope in NNP where high populations of *An. quadriannulatus* persisted. Although a medium-sized antelope, breeding herds of impala can comprise of 120 individuals (Estes, 2012). This presents a large quantity of mammalian biomass, which is an important factor determining the host choice of blood-seeking adult mosquitoes (Lehane, 2005; Takken and Verhulst, 2013) by providing larger surface areas for biting (Port, Boreham and Bryan, 1980; Smith *et al.*, 2006), but also greater olfactory cues including higher levels of carbon dioxide which is a key kairomone for foraging mosquitoes (Gillies and Wilkes, 1972; Takken and Knols, 1999). Furthermore, impala have a flexible foraging ecology that encourages small home ranges and a sedentary existence.

Impala are predominantly grazers in the wet season when grasses are plentiful but will readily switch to browsing in the dry season, and so have a year-round abundant and reliable food supply that can support large herds (Estes, 2012). Therefore, they do not need to migrate long distances for resources and as a result, have much smaller home ranges (often less than $1km^2$) than larger more specialised antelopes (Estes, 2012). Furthermore, when the grasslands are dry, browsing impalas reside close to water sources where they need to drink daily and depend on the trees, shrubs and bushes that grow around waterholes and streambeds (Estes, 2012). Impala are also least active after dusk and spend most of this time lying down (Estes, 2012), which coincides with the peak crepuscular biting activity associated with zoophagy in general (Elliott, 1972) and that of *An. quadriannulatus* as recorded in southern Zambia, where they also feed frequently on humans (Seyoum et al., 2012). Considering their ecology, impala may represent especially reliable and abundant blood resources that are concentrated around waterbodies

where emergent blood-seeking mosquitoes may breed, and therefore an astute choice for *An. quadriannulatus* to have evolved to specialise on.

Host-specialised blood utilisation behaviours around anopheline mosquitoes is very common occurring in 82% of all studied anopheline species (Lyimo and Ferguson, 2009; Takken and Verhulst, 2013). Such specialisation to feed on only one or two vertebrate species is thought to have evolved under the selection of several factors, notably the rates at which various potential host species are encountered (Lyimo and Ferguson, 2009). For example, it has been shown that the speciation of the highly specialised *An. gambiae s.s.* arose with the arrival of Bantu agriculturists in the forest belt of central Africa (Ayala and Coluzzi, 2005). The lethal effects of trypanosomiasis transmitted by the forests' tsetse flies were detrimental for cattle which subsequently made humans the largest vertebrate host available to the ancestral *Anopheles* species, and so the number of encounters between humans and mosquitoes increased (Ayala and Coluzzi, 2005). Therefore, the unusually high historical abundance of impala across all the mixed acacia savannah habitat of eastern and southern Africa (Estes, 2012), may well have underpinned the evolution of *An. quadriannulatus* to specialise on this antelope that remain widespread across protected areas today.

However, it is also important to consider the limitations of associative observational studies when inferring preferred blood resources. As other mammalian species were not as frequently detected inside NNP, the models detailed in table 3.4 had less statistical power for evaluating the influence of these other potential blood host species. And so, a lack of other attributable influences may not necessarily imply that these numerous alternative mammals are not utilised as blood meals. It may well be the case that *An. quadriannulatus* simply feed flexibly and opportunistically upon whatever the most abundant hosts available happen to be, with impala being the far the most abundant in this case.

That being said, the statistical identification of bushpig as a second potential preferred blood host provides encouraging evidence of the validity and power of the approach taken, because even the small number of detections of this species when compared to many other wild animals, proved sufficient to give plausible evidence that *An. quadriannulatus* may well utilise them as blood sources. Like impala, bushpigs are also social animals, and live in groups of up to 15 individuals (Estes, 2012). Their preferred habitats include woodland and forest that provide dense vegetation (Estes, 2012) which is consistent with their observed distribution in this study, with smaller pockets of *An. quadrinannulatus* found around such locations within ILUMA WMA rather than the open acacia savanna of the nearby national park. Bushpig forage at night, and often thrive in swampy areas at the interface between humans and wild animals (Kukielka *et al.*, 2016). Like impala, they are also highly water dependent with a sedentary existence, and permanently stay within remarkably small home ranges with ready access to surface water (Estes, 2012), and may therefore also offer quite reliable, readily accessible feeding opportunities for blood-seeking mosquitoes.

Of course, it is reasonable to question the reliability with which the activities of humans, livestock and such an impressive diversity of wild mammals were surveyed, especially given the dense vegetation cover across much of the study area, and the cryptic or fluctuating habits of many of these mammalian species. However, practical experience during the study and detailed analysis of the animal activity survey data (Duggan *et al.*, Unpublished), both confirm that indirect observations of tracks, spoor and other signs were quite sensitive and reliable indicators, revealing the presence of almost every species envisaged at the outset, including 15 that were never directly sighted over the course of the two years. Although the activities of humans, livestock and wild animals were detected in several different ways, footprints and hoofprints were by far the most frequent indicator found around waterbodies. Indeed, the wet soil around the perimeter of many water bodies, particularly during the wet season, facilitated

clear track indentations that could be confidently identified by the investigators and VGS carrying out the radial surveys. Although the age of the tracks was recorded during the survey, these distinctions were not considered for the analysis because the most recent prints were always more likely to be recorded, as they usually covered the tracks of older detected prints in such wet areas around waterbodies.

While the survey methodology proved useful for inferring host species, its greatest limitation of this study overall is that it had an observational design and can therefore only demonstrate indirect associations rather than direct causality. Unfortunately, although over 4,000 live adult mosquitoes were caught over the two years of this study, not a single blood-fed individual was caught, despite the use of an interception screen specifically chosen to capture at least some blood-fed mosquitoes (Burkot *et al.,* 2013). Therefore, a direct link between wild mosquito populations and the host species that they were feeding on could not be demonstrated conclusively. Nevertheless, this novel methodology of associating the species composition through larval surveys with indicators of host availability has proven to be a simple, effective, and surprisingly powerful approach that yielded clear evidence for four different preferred hosts of the *An. gambiae* complex across the study area, two for *An. arabiensis* and two for *An. quadriannulatus.*

Having said all that, the existence of *An. arabiensis* deep inside NNP could not be explained by any of the variables retained in the final multivariate model (Table 3.4), so it was not possible to address the original objective of the study, which was to associate wild populations of *An. arabiensis* with one or more alternative hosts other than humans and cattle. Prior to data collection, it was hypothesised that perhaps African buffalo might provide a plausible blood source for wild type mosquitoes because of their close phylogenetic relationship with cattle (Janecek *et al.,* 1996), combined with the simple fact that they are large-sized mammals that often form large herds (Estes, 2012), thus providing a considerable biomass for host-seeking

mosquitoes to exploit. These bovids were also suggested as a potential blood source for a wild population of *An. arabiensis* identified in Kruger National Park, South Africa, based on the observation that African buffalo frequented aquatic habitats that were occupied by larvae (Braack *et al.*, 1994). Furthermore, bloodmeal analysis from a study in Uganda identified DNA of African buffalo from a blood-fed *An. gambiae* complex mosquito but unfortunately the specimen was not identified to species level (Crabtree *et al.*, 2013). As the host association model in table 3.4 did not provide any evidence that the abundance of African buffalo or any wild herbivore favoured *An. arabiensis*, questions remain regarding the ecological niche of this vector species outside of landscapes that are at least partially domesticated, with humans and/or cattle available to support this mosquito species.

4.6 Plausibility of refugia An. arabiensis populations in pristine wilderness areas.

Given that *An. arabiensis* appears so well adapted to dry environments (Lindsay, Parson and Thomas, 1998), their presence deep inside NNP might possibly be explained by dispersal from surrounding domesticated lands. Dispersal refers to the movements of mosquitoes over distances as short as a few meters or as long as several hundred kilometres as they routinely forage for hosts and oviposition sites or ride the wind to new areas (Service, 1997). Host-seeking adults from the *Anopheles gambiae* complex usually have modest flight range of a few hundred metres to a few kilometres (Gillies, 1961; Gillies and Coetzee, 1987; Kaufmann and Briegel, 2004) but they also sometimes disperse much further by exploiting high altitude winds (Huestis *et al.*, 2019; Florio *et al.*, 2020; Atieli *et al.*, 2023). Indeed, extremely long-distance windborne movements of gravid females across hundreds of kilometres of the Sahel, apparently driven by the seasonal availability of suitable aquatic habitats, has recently been documented for *An. coluzzii*, another member of the *An. gambiae* complex, (Huestis *et al.*, 2019). Furthermore, it has also been demonstrated that members of the *An. gambiae* complex, including *An. arabiensis*, can survive, oviposit and blood feed again after being exposed to

high altitude flight for up to 11 hours (Sanogo *et al.*, 2021). Indeed, as no wild alternative hosts for *An. arabiensis* could be identified for this vector species inside pristine wilderness areas (Table 3.4), it could be argued that these are *sink* populations (Pulliam, 1988; Dias, 1996) that may not be independently viable without sustained influx of already blood-fed adults from surrounding *source* populations (Pulliam, 1988; Dias, 1996) in domesticated villages.

However, it is difficult to envisage how such long-range dispersal alone could sustain such readily detected densities of *An. arabiensis* so deep inside NNP for several reasons. Firstly, the reciprocal of this spatial distribution pattern for *An. arabiensis* was not observed: Despite the presence of abundant *An. quadriannulatus* inside NNP, no specimens of this species were found in any of the nearby villages or even across most of the ILUMA conservation area. Notably, almost all (9/11) sampled location within the buffer zone represented by the WMA where *An. quadriannulatus* larvae were found also had detectable numbers of impala, bushpig or both. It therefore seems likely that local population viability, for *An. quadriannulatus* at least, is determined primarily by local condition, particularly the availability of preferred hosts, at quite fine spatial scales.

Secondly, each individual *An. arabiensis* larva caught at Shughuli Kubwa (camp number 40), located 47km inside NNP from the park boundary (Figure 2.9), were distributed across multiple breeding habitats, suggesting that they originated from more than one female. Also, several unfed, apparently host-seeking adult female *An. arabiensis* were also captured at the same remote location. Furthermore, since the study was complete, another 52 *An. arabiensis* larvae were collected from a new location called *Mseguni* inside NNP (Figure 2.9), which is located even further away from humans and cattle than Shughuli Kubwa. This strongly suggests the existence of a well-established population that acquire bloodmeals locally, rather than an otherwise non-viable *sink* population sustained by external *source* populations outside the park.

However, the most convincing evidence for a viable, self-sustaining refuge population of *An. arabiensis* deep inside NNP is the observation that some mosquito lineages reared from wild caught larvae and adults at these locations exhibit full susceptibility to pyrethroids, probably indicative of ancestral wild-type genomes (Kavishe *et al.*, Unpublished), persisting far away from the selective pressures of insecticides that are widely used in domesticated settlements. Having said that, resistance traits were also derived from wild-caught specimens well inside NNP (Kavishe *et al.*, Unpublished), confirming that long-distance dispersal nevertheless mediates a considerable amount of gene flow into the park. However, the way in which the prevalence of these resistance traits dissipates rapidly as one gets deeper into the park suggests that they also confer appreciable fitness costs in pristine habitats where they are redundant (Kavishe *et al.*, Unpublished).

4.7 Implications of wild An. arabiensis populations in ILUMA WMA and NNP.

It was clear that unauthorised human and livestock encroachment was extensive throughout most of the ILUMA WMA and had a substantial negative impact on the availability and distribution of wild animals (Duggan *et al*, Unpublished). In particular, illegal cattle herding represents an influx of preferred blood sources for *An. arabiensis* and all major forms of encroachment reduce the availability of wild mammalian blood sources, thus displacing the non-vector *An. quadriannulatus* in favour of the malaria vector it competes with. Through improved management and conservation efforts of community-based conservation areas, to effectively control unauthorised, illegal human activities and facilitate the return of wild animals to the WMA, a natural reduction of *An. arabiensis* in favour of *An. quadriannulatus* may well follow. Therefore, natural competitive suppression of this important malaria vector would represent a previously undocumented form of natural capital that arises from supporting and protecting conservation areas.

While it may be possible to suppress *An. arabiensis* populations in conservation areas to some extent, this vector was not completely displaced, even in the wildest areas with clear evidence of abundant preferred wild herbivore species for *An. quadriannulatus* being readily available. This highlights how the ecological plasticity of *An. arabiensis* extends deep into wild areas, where they can even compete for resources against this much more zoophagic member of the same complex. Furthermore, the apparent existence of such refugia populations of *An. arabiensis* in these remote pristine ecosystems, confirms the initial hypothesis that alternative blood resources create an additional population-stabilising portfolio effect (Killeen and Reed, 2018) for this species, and further emphasizing the major vector control challenges posed by its flexible feeding behaviours.

Following the initial observation that the relative abundance of *An. arabiensis* decreased while that of *An. quadriannulatus* increased as the surveys reached further inside NNP, it was hoped that these competitive interactions in response to changing host availabilities might ultimately enable elimination (Killeen *et al.*, 2013) through competitive displacement (DeBach, 1966), or competitively reduce the vector population (Lounibos, 2007) by the latter relatively harmless sibling species, as historically observed for other vectors (Giglioli, 1951; Gillies and Smith, 1960; Gillies and Furlong, 1964; Smith, 1962). However, since *An. arabiensis* are clearly and robustly the dominant competitive species wherever any humans or cattle whatsoever are found, and given that a refuge population of this species even exists deep inside NNP where there are none, the prospects for eliminating this key vector of residual malaria transmission appear remote. Not only are wild refugia populations feeding upon wild animals largely out of the reach of current first-choice LLIN and IRS interventions that protect humans where and when they sleep, usually in reasonably well-established houses and shelters, but they may also limit the population suppression effects (Killeen *et al.*, 2017a) of new complementary approaches like spatial repellents and veterinary endectocides (Killeen *et al.*, 2017c) that

respectively extend insecticidal coverage to humans and livestock outdoors (Killeen *et al.*, 2017b). Furthermore, in parts of neighbouring Zambia where sympatric populations of *An. quadriannulatus* and *An. arabiensis* occur in domesticated settings, it seems that the former is much more vulnerable to insecticide attack than the latter, as evidenced by an abrupt species composition shift towards the latter vector species immediately after effective scale up of IRS (Chinula *et al.*, 2018). It might therefore be difficult to target adult *An. arabiensis* with insecticide-based interventions without also affecting *An. quadriannulatus* populations.

On the other hand, two notable historical examples of successful elimination of this species from Brazil (Soper and Wilson, 1943; Killeen *et al.*, 2002; Killeen, 2003) (Retrospectively confirmed as *An. arabiensis* by molecular techniques (Parmakelis *et al.*, 2008)), and Egypt (Shousha, 1948) (By far the most likely culprit based on ecological niche mapping (Moffett, Shackelford and Sarkar, 2007; Sinka *et al.*, 2016)), primarily through comprehensive larvicide application, suggest it can nevertheless be tackled decisively at source as larvae, before they can express the evasive behaviours that make the adults so difficult to control. In both cases, however, *An. arabiensis* was an invasive species that had established itself within a limited, albeit quite extensive new geographic range, within which it proved possible to contain, suppress and eventually eliminate. It is very difficult to envisage such "scorched earth" larval source management campaigns being feasible or affordable across the vast natural range of this species in Africa, especially if that includes refuge populations in the many large tracts of largely uninhabited wilderness scattered across the continent.

On a more positive note, although wild refugia populations deep inside NNP pose clear challenges for vector elimination, they also provide a potential opportunity for reversing the deleterious effects of pyrethroid resistance on the progress of malaria control to date (Hemingway *et al.*, 2016; Ranson and Lissenden, 2016; Churcher *et al.*, 2016; Hancock *et al.*, 2020; Moyes *et al.*, 2020; Mosha *et al.*, 2022; Accrombessi *et al.*, 2023). Self-sustaining

populations that do not come into contact with humans or livestock, and are therefore free from the selective pressures of insecticides, may represent invaluable reservoirs of relatively unmodified mosquito genomes, complete with original wild-type insecticide susceptibility traits that would otherwise have been lost from the natural gene pool. Indeed, two lineages of *An. arabiensis* exhibit fully pyrethroid susceptibility have been isolated from *An. arabiensis* specimens caught inside NNP and even at one of the most pristine locations in the ILUMA WMA (Kavishe *et al.*, Unpublished). The existence of such large reservoirs of wild-type genomes that have never been bottlenecked by insecticide pressure, could enable insecticide resistance management strategies that exploit astute insecticide combinations to select back these traits (White *et al.*, 2014) in populations living in the surrounding villages, thereby restoring the effectiveness of current (WHO, 2017) and future (Govella *et al.*, 2015; Masalu *et al.*, 2017; Ogoma *et al.*, 2017; Mmbando *et al.*, 2018; Swai *et al.*, 2023) interventions that remain dependent on this exceptionally useful insecticide class.

Conclusion

Despite a moderate decline in larval occupancy as ecosystems become more pristine further away from domesticated areas, *An. gambiae* complex larvae nevertheless ubiquitously colonise a considerable fraction of available aquatic habitats. Along the same gradient, unexpected variability in species composition within *An. gambiae* complex across these occupied habitats was readily explained by apparently competitive interactions between *An. arabiensis* and *An. quadriannulatus,* a second sibling species found in the study area, which seems primarily driven by the relative availabilities of the different large mammals each prefers as sources of blood. The strong host preference of the notorious vector of residual malaria, *An. arabiensis*, for humans or cattle appears to give them a strong competitive advantage over *An. quadriannulatus* wherever these two host species occur, even at remarkably low densities inside quite pristine conservation areas.

Aquatic habitats occupied by *An. arabiensis* deep inside pristine areas, even in the complete absence of their preferred human and cattle hosts, indicate the presence of self-sustaining refugia populations that rely on bloodmeals acquired from one or more wild animal species that have yet to be unidentified. This observation casts further doubt upon the feasibility of eliminating *An. arabiensis* from large tracts of its natural range in Africa. On the other hand, however, these refugia populations may represent naturally-occurring wild reservoirs of persisting insecticide susceptibility traits that might enable more ambitious insecticide resistance management strategies than might otherwise be possible.

References

Accrombessi, M., Cook, J., Dangbenon, E., Yovogan, B., Akpovi, H., Sovi, A., Adoha, C., Assongba, L., Sidick, A. and Akinro, B. (2023) 'Efficacy of pyriproxyfen-pyrethroid long-lasting insecticidal nets (LLINs) and chlorfenapyr-pyrethroid LLINs compared with pyrethroid-only LLINs for malaria control in Benin: a cluster-randomised, superiority trial', *The Lancet*, 401(10375), pp. 435-446.

African Union 2013. Policy framework for pastoralism in Africa: Securing, protecting and improving the lives, livelihoods and rights of pastoralist communities.

Akpodiete, N. O. and Tripet, F. (2021a) 'Improvement of water quality for mass anopheline rearing: evaluation of the impact of ammonia-capturing zeolite on larval development and adult phenotypic quality', *Parasites & Vectors*, 14(1), pp. 268.

Akpodiete, N. O. and Tripet, F. (2021b) 'Laboratory and microcosm experiments reveal contrasted adaptive responses to ammonia and water mineralisation in aquatic stages of the sibling species Anopheles gambiae (sensu stricto) and Anopheles coluzzii', *Parasites & Vectors*, 14, pp. 1-19.

Anderson, T. M., Hopcraft, J. G. C., Eby, S., Ritchie, M., Grace, J. B. and Olff, H. (2010) 'Landscape-scale analyses suggest both nutrient and antipredator advantages to Serengeti herbivore hotspots', *Ecology*, 91(5), pp. 1519-1529.

Aschehoug, E. T., Brooker, R., Atwater, D. Z., Maron, J. L. and Callaway, R. M. (2016) 'The mechanisms and consequences of interspecific competition among plants', *Annual Review of Ecology, Evolution, and Systematics*, 47, pp. 263-281.

Atieli, H. E., Zhou, G., Zhong, D., Wang, X., Lee, M.-c., Yaro, A. S., Diallo, M., Githure, J., Kazura, J. and Lehmann, T. (2023) 'Wind-assisted high-altitude dispersal of mosquitoes and other insects in East Africa', *Journal of Medical Entomology*, pp. tjad033.

Ayala, F. J. and Coluzzi, M. (2005) 'Chromosome speciation: humans, Drosophila, and mosquitoes', *Proceedings of the National Academy of Sciences*, 102(suppl_1), pp. 6535-6542.

Barbosa, P., Peters, T. M. and Greenough, N. (1972) 'Overcrowding of mosquito populations: responses of larval Aedes aegypti to stress', *Environmental Entomology*, 1(1), pp. 89-93.

Barreaux, A. M., Stone, C. M., Barreaux, P. and Koella, J. C. (2018) 'The relationship between size and longevity of the malaria vector Anopheles gambiae (ss) depends on the larval environment', *Parasites & vectors*, 11, pp. 1-9.

Barrón, M. G., Paupy, C., Rahola, N., Akone-Ella, O., Ngangue, M. F., Wilson-Bahun, T. A., Pombi, M., Kengne, P., Costantini, C. and Simard, F. (2019) 'A new species in the major malaria vector complex sheds light on reticulated species evolution', *Scientific Reports,* 9(1), pp. 14753.

Bayoh, M. N., Mathias, D. K., Odiere, M. R., Mutuku, F. M., Kamau, L., Gimnig, J. E., Vulule, J. M., Hawley, W. A., Hamel, M. J. and Walker, E. D. (2010) 'Anopheles gambiae: historical population decline associated with regional distribution of insecticide-treated bed nets in western Nyanza Province, Kenya', *Malaria journal,* 9(1), pp. 1-12.

Beier, J. C. (1998) 'Malaria parasite development in mosquitoes', *Annual review of entomology,* 43(1), pp. 519-543.

Beier, J. C., Killeen, G. F. and Githure, J. I. (1999) 'entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa', *The American journal of tropical medicine and hygiene,* 61(1), pp. 109-113.

Besansky, N. J., Hill, C. A. and Costantini, C. (2004) 'No accounting for taste: host preference in malaria vectors', *Trends in parasitology,* 20(6), pp. 249-251.

Bhatt, S., Weiss, D., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C., Henry, A. and Eckhoff, P. (2015) 'The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015', *Nature,* 526(7572), pp. 207-211.

Braack, L., Coetzee, M., Hunt, R., Biggs, H., Cornel, A. and Gericke, A. (1994) 'Biting pattern and host-seeking behavior of Anopheles arabiensis (Diptera: Culicidae) in northeastern South Africa', *Journal of medical entomology,* 31(3), pp. 333-339.

Bracco, A., LaCasce, J. and Provenzale, A. (2000) 'Velocity probability density functions for oceanic floats', *Journal of physical oceanography,* 30(3), pp. 461-474.

Burkot, T. R., Russell, T. L., Reimer, L. J., Bugoro, H., Beebe, N. W., Cooper, R. D., Sukawati, S., Collins, F. H. and Lobo, N. F. (2013) 'Barrier screens: a method to sample blood-fed and host-seeking exophilic mosquitoes', *Malaria journal,* 12(1), pp. 1-9.

Carlson, J. C., Byrd, B. D. and Omlin, F. X. (2004) 'Field assessments in western Kenya link malaria vectors to environmentally disturbed habitats during the dry season', *BMC Public Health,* 4, pp. 1-7.

Carpenter, S. R. (1983) 'Resource limitation of larval treehole mosquitoes subsisting on beech detritus', *Ecology,* 64(2), pp. 219-223.

Chaki, P. P., Dongus, S., Fillinger, U., Kelly, A. and Killeen, G. F. (2011) 'Community-owned resource persons for malaria vector control: enabling factors and challenges in an operational programme in Dar es Salaam, United Republic of Tanzania', *Human resources for health,* 9(1), pp. 1-11.

Chaki, P. P., Govella, N. J., Shoo, B., Hemed, A., Tanner, M., Fillinger, U. and Killeen, G. F. (2009) 'Achieving high coverage of larval-stage mosquito surveillance: challenges for a community-based mosquito control programme in urban Dar es Salaam, Tanzania', *Malaria Journal,* 8(1), pp. 1-12.

Charlwood, J. (1997) 'Vectorial capacity, species diversity and population cycles of anopheline mosquitoes (Diptera: Culicidae) from indoor light-trap collections in a house in southeastern Tanzania', *African entomology,* 5(1), pp. 93-101.

Charlwood, J. and Edoh, D. (1996) 'Polymerase chain reaction used to describe larval habitat use by Anopheles gambiae complex (Diptera: Culicidae) in the environs of Ifakara, Tanzania', *Journal of Medical Entomology,* 33(2), pp. 202-204.

Charlwood, J., Kihonda, J., Sama, S., Billingsley, P., Hadji, H., Verhave, J., Lyimo, E., Luttikhuizen, P. and Smith, T. (1995a) 'The rise and fall of Anopheles arabiensis (Diptera: Culicidae) in a Tanzanian village', *Bulletin of Entomological Research,* 85(1), pp. 37-44.

Charlwood, J., Smith, T., Billingsley, P., Takken, W., Lyimo, E. and Meuwissen, J. (1997) 'Survival and infection probabilities of anthropophagic anophelines from an area of high prevalence of Plasmodium falciparum in humans', *Bulletin of Entomological Research,* 87(5), pp. 445-453.

Charlwood, J., Smith, T., Kihonda, J., Heiz, B., Billingsley, P. and Takken, W. (1995b) 'Density independent feeding success of malaria vectors (Diptera: Culicidae) in Tanzania', *Bulletin of Entomological Research,* 85(1), pp. 29-35.

Charlwood, J., Vij, R. and Billingsley, P. (2000) 'Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of east Africa', *The American journal of tropical medicine and hygiene,* 62(6), pp. 726-732.

Chesson, P. (2000) 'General theory of competitive coexistence in spatially-varying environments', *Theoretical population biology,* 58(3), pp. 211-237.

Chesson, P. L. (1986) 'Environmental variation and the coexistence of species', *Community ecology,* 240, pp. 54.

Chinula, D., Hamainza, B., Chizema, E., Kavishe, D. R., Sikaala, C. H. and Killeen, G. F. (2018) 'Proportional decline of Anopheles quadriannulatus and increased contribution of An. arabiensis to the An. gambiae complex following introduction of indoor residual spraying with pirimiphos-methyl: an observational, retrospective secondary analysis of pre-existing data from south-east Zambia', *Parasites & vectors,* 11(1), pp. 1-7.

Churcher, T. S., Lissenden, N., Griffin, J. T., Worrall, E. and Ranson, H. (2016) 'The impact of pyrethroid resistance on the efficacy and effectiveness of bednets for malaria control in Africa', *Elife,* 5, pp. e16090.

Coetzee, M., Craig, M. and Le Sueur, D. (2000) 'Distribution of African malaria mosquitoes belonging to the Anopheles gambiae complex', *Parasitology today,* 16(2), pp. 74-77.

Coetzee, M., Hunt, R. H., Wilkerson, R., Della Torre, A., Coulibaly, M. B. and Besansky, N. J. (2013) 'Anopheles coluzzii and Anopheles amharicus, new members of the Anopheles gambiae complex', *Zootaxa,* 3619(3), pp. 246-274.

Cohuet, A., Harris, C., Robert, V. and Fontenille, D. (2010) 'Evolutionary forces on Anopheles: what makes a malaria vector?', *Trends in parasitology,* 26(3), pp. 130-136.

Coluzzi, M. (1984) 'Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control', *Bulletin of the World Health Organization,* 62(Suppl), pp. 107.

Coluzzi, M., Sabatini, A., Petrarca, V. and Di Deco, M. A. (1979) 'Chromosomal differentiation and adaptation to human environments in the Anopheles gambiae complex', *Transactions of the Royal Society of tropical Medicine and Hygiene,* 73(5), pp. 483-497.

Costantini, C., Sagnon, N., Torre, A. and Coluzzi, M. (1999) 'Mosquito behavioural aspects of vector-human interactions in the Anopheles gambiae complex', *Parassitologia,* 41(1/3), pp. 209-220.

Crabtree, M. B., Kading, R. C., Mutebi, J.-P., Lutwama, J. J. and Miller, B. R. (2013) 'Identification of host blood from engorged mosquitoes collected in western Uganda using cytochrome oxidase I gene sequences', *Journal of wildlife diseases,* 49(3), pp. 611-626.

Davidson, G. (1964) 'Anopheles gambiae, a complex of species', *Bulletin of the World Health Organization,* 31(5), pp. 625.

Davidson, G. and Jackson, C. E. (1962) 'Incipient speciation in Anopheles gambiae Giles', *Bulletin of the World Health Organization,* 27(2), pp. 303.

De Boer, W. F., Vis, M. J., De Knegt, H. J., Rowles, C., Kohi, E. M., Van Langevelde, F., Peel, M., Pretorius, Y., Skidmore, A. K. and Slotow, R. (2010) 'Spatial distribution of lion kills determined by the water dependency of prey species', *Journal of Mammalogy,* 91(5), pp. 1280-1286.

DeBach, P. (1966) 'The competitive displacement and coexistence principles', *Annual review of entomology,* 11(1), pp. 183-212.

Diabaté, A., Dabire, R. K., Kim, E. H., Dalton, R., Millogo, N., Baldet, T., Simard, F., Gimnig, J. E., Hawley, W. A. and Lehmann, T. (2005) 'Larval development of the molecular forms of Anopheles gambiae (Diptera: Culicidae) in different habitats: a transplantation experiment', *Journal of Medical Entomology,* 42(4), pp. 548-553.

Diabaté, A., Dabiré, R. K., Heidenberger, K., Crawford, J., Lamp, W. O., Culler, L. E. and Lehmann, T. (2008) 'Evidence for divergent selection between the molecular forms of Anopheles gambiae: role of predation', *BMC evolutionary biology,* 8, pp. 1-11.

Dias, P. C. (1996) 'Sources and sinks in population biology', *Trends in ecology & evolution,* 11(8), pp. 326-330.

Drakeley, C., Akim, N., Sauerwein, R., Greenwood, B. and Targett, G. (2000) 'Estimates of the infectious reservoir of Plasmodium falciparum malaria in The Gambia and in Tanzania', *Transactions of the Royal Society of tropical Medicine and Hygiene,* 94(5), pp. 472-476.

Dukeen, M. Y. and Omer, S. (1986) 'Ecology of the malaria vector Anopheles arabiensis Patton (Diptera: Culicidae) by the Nile in northern Sudan', *Bulletin of entomological research,* 76(3), pp. 451-467.

Durnez, L. and Coosemans, M. (2013) 'Residual transmission of malaria: an old issue for new approaches', *Anopheles mosquitoes: new insights into malaria vectors/Manguin, Sylvie,* pp. 671-704.

Dye, C. (1984) 'Competition amongst larval Aedes aegypti: the role of interference'.

Edillo, F. E., Touré, Y. T., Lanzaro, G. C., Dolo, G. and Taylor, C. E. (2002) 'Spatial and habitat distribution of Anopheles gambiae and Anopheles arabiensis (Diptera: Culicidae) in Banambani village, Mali', *Journal of Medical Entomology,* 39(1), pp. 70-77.

Elliott, R. (1972) 'The influence of vector behavior on malaria transmission', *The American Journal of Tropical Medicine and Hygiene,* 21(5_Suppl), pp. 755-763.

Epopa, P. S., Millogo, A. A., Collins, C. M., North, A. R., Benedict, M. Q., Tripet, F., O'Loughlin, S., Dabiré, R. K., Ouédraogo, G. A. and Diabaté, A. (2020) 'Anopheles gambiae (sl) is found where few are looking: assessing mosquito diversity and density outside inhabited areas using diverse sampling methods', *Parasites & Vectors,* 13(1), pp. 1-11.

Estes, R. D. (2012) *The behavior guide to African mammals: including hoofed mammals, carnivores, primates.* Univ of California Press.

Ferguson, H. M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M. S., Gimnig, J. E., Fish, D. and Killeen, G. F. (2010) 'Ecology: a prerequisite for malaria elimination and eradication', *PLoS medicine,* 7(8), pp. e1000303.

Fillinger, U., Kannady, K., William, G., Vanek, M. J., Dongus, S., Nyika, D., Geissbühler, Y., Chaki, P. P., Govella, N. J. and Mathenge, E. M. (2008) 'A tool box for operational mosquito larval control: preliminary results and early lessons from the Urban Malaria Control Programme in Dar es Salaam, Tanzania', *Malaria journal,* 7, pp. 1-25.

Fillinger, U. and Lindsay, S. W. (2006) 'Suppression of exposure to malaria vectors by an order of magnitude using microbial larvicides in rural Kenya', *Tropical Medicine & International Health,* 11(11), pp. 1629-1642.

Fillinger, U. and Lindsay, S. W. (2011) 'Larval source management for malaria control in Africa: myths and reality', *Malaria journal,* 10(1), pp. 1-10.

Fillinger, U., Sonye, G., Killeen, G. F., Knols, B. G. and Becker, N. (2004) 'The practical importance of permanent and semipermanent habitats for controlling aquatic stages of Anopheles gambiae sensu lato mosquitoes: operational observations from a rural town in western Kenya', *Tropical medicine & international health,* 9(12), pp. 1274-1289.

Finda, M. F., Moshi, I. R., Monroe, A., Limwagu, A. J., Nyoni, A. P., Swai, J. K., Ngowo, H. S., Minja, E. G., Toe, L. P. and Kaindoa, E. W. (2019) 'Linking human behaviours and malaria vector biting risk in south-eastern Tanzania', *PloS one,* 14(6), pp. e0217414.

Florio, J., Verú, L. M., Dao, A., Yaro, A. S., Diallo, M., Sanogo, Z. L., Samaké, D., Huestis, D. L., Yossi, O. and Talamas, E. (2020) 'Diversity, dynamics, direction, and magnitude of high-altitude migrating insects in the Sahel', *Scientific Reports,* 10(1), pp. 20523.

Freyvogel, T. A. and Kihaule, P. M. (1968) 'Report on a limited anopheline survey at Ifakara, south-eastern Tanzania', *Acta Tropica,* pp. 1-27.

Garrett-Jones, C. (1964) 'The human blood index of malaria vectors in relation to epidemiological assessment', *Bulletin of the World Health Organization,* 30(2), pp. 241.

Garrett-Jones, C., Boreham, P. and Pant, C. (1980) 'Feeding habits of anophelines (Diptera: Culicidae) in 1971–78, with reference to the human blood index: a review', *Bulletin of entomological research,* 70(2), pp. 165-185.

Gause, G. F. (1934) *The struggle for Existence.* Baltimore: Williams and Wilkins.

Giglioli, G. (1951) 'Nation-wide malaria eradication projects in the Americas. III. Eradication of Anopheles darlingi from the inhabited areas of British Guiana by DDT residual spraying', *Journal. National Malaria Society (US),* 10(2), pp. 142-161.

Gillies, M. (1988) 'Anopheline mosquitoes. Vector behaviour and bionomics', *Malaria: Principle and practices of malariology,* 1, pp. 453-485.

Gillies, M. and Furlong, M. (1964) 'An investigation into the behaviour of Anopheles parensis Gillies at Malindi on the Kenya coast', *Bulletin of Entomological Research,* 55(1), pp. 1-16.

Gillies, M. and Wilkes, T. (1972) 'The range of attraction of animal baits and carbon dioxide for mosquitoes. Studies in a freshwater area of West Africa', *Bulletin of Entomological Research,* 61(3), pp. 389-404.

Gillies, M. T. (1956) 'The problem of exophily in Anopheles gambiae', *Bulletin of the World Health Organization,* 15(3-5), pp. 437.

Gillies, M. T. (1961) 'Studies on the dispersion and survival of Anopheles gambiae Giles in East Africa, by means of marking and release experiments', *Bulletin of Entomological Research,* 52(1), pp. 99-127.

Gillies, M. T. and Coetzee, M. (1987) 'A supplement to the Anophelinae of Africa South of the Sahara', *Publ S Afr Inst Med Res,* 55, pp. 1-143.

Gillies, M. T. and De Meillon, B. (1968) 'The Anophelinae of Africa south of the Sahara (Ethiopian zoogeographical region)', *The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region).*

Gillies, M. T. and Smith, A. (1960) 'The effect of a residual house-spraying campaign in East Africa on species balance in the Anopheles funestus group. The replacement of A. funestus Giles by A. rivulorum Leeson', *Bulletin of Entomological Research,* 51(2), pp. 243-252.

Gimnig, J. E., Ombok, M., Bayoh, N., Mathias, D., Ochomo, E., Jany, W. and Walker, E. D. (2020) 'Efficacy of extended release formulations of Natular™(spinosad) against larvae and adults of Anopheles mosquitoes in western Kenya', *Malaria Journal,* 19, pp. 1-13.

Gimnig, J. E., Ombok, M., Kamau, L. and Hawley, W. A. (2001) 'Characteristics of larval anopheline (Diptera: Culicidae) habitats in Western Kenya', *Journal of medical entomology,* 38(2), pp. 282-288.

Gimnig, J. E., Ombok, M., Otieno, S., Kaufman, M. G., Vulule, J. M. and Walker, E. D. (2002) 'Density-dependent development of Anopheles gambiae (Diptera: Culicidae) larvae in artificial habitats', *Journal of Medical Entomology,* 39(1), pp. 162-172.

Gimonneau, G., Pombi, M., Choisy, M., Morand, S., Dabiré, R. K. and Simard, F. (2012) 'Larval habitat segregation between the molecular forms of the mosquito Anopheles gambiae in a rice field area of Burkina Faso, West Africa', *Medical and veterinary entomology,* 26(1), pp. 9-17.

Githeko, A. K., Adungo, N. I., Karanja, D. M., Hawley, W. A., Vulule, J. M., Seroney, I. K., Ofulla, A. V., Atieli, F. K., Ondijo, S. O. and Genga, I. O. (1996) 'Some Observations on the Biting Behavior ofAnopheles gambiae ss, Anopheles arabiensis, andAnopheles funestusand Their Implications for Malaria Control', *Experimental parasitology,* 82(3), pp. 306-315.

Govella, N. J., Chaki, P. P. and Killeen, G. F. (2013) 'Entomological surveillance of behavioural resilience and resistance in residual malaria vector populations', *Malaria Journal,* 12(1), pp. 1-9.

Govella, N. J. and Ferguson, H. (2012) 'Why use of interventions targeting outdoor biting mosquitoes will be necessary to achieve malaria elimination', *Frontiers in physiology,* 3, pp. 199.

Govella, N. J., Johnson, P. C., Killeen, G. F. and Ferguson, H. M. (2023) 'Heritability of biting time behaviours in the major African malaria vector Anopheles arabiensis', *Malaria Journal,* 22(1), pp. 238.

Govella, N. J., Ogoma, S. B., Paliga, J., Chaki, P. P. and Killeen, G. (2015) 'Impregnating hessian strips with the volatile pyrethroid transfluthrin prevents outdoor exposure to vectors of malaria and lymphatic filariasis in urban Dar es Salaam, Tanzania', *Parasites & Vectors,* 8(1), pp. 1-9.

Govella, N. J., Okumu, F. O. and Killeen, G. F. (2010) 'Insecticide-treated nets can reduce malaria transmission by mosquitoes which feed outdoors', *The American journal of tropical medicine and hygiene,* 82(3), pp. 415.

Grover, J. P. (1997) *Resource competition.* Springer Science & Business Media.

Guillet, P., Alnwick, D., Cham, M. K., Neira, M., Zaim, M., Heymann, D. and Mukelabai, K. (2001) 'Long-lasting treated mosquito nets: a breakthrough in malaria prevention', *Bulletin of the World Health Organization,* 79.

Hancock, P. A., Hendriks, C. J., Tangena, J.-A., Gibson, H., Hemingway, J., Coleman, M., Gething, P. W., Cameron, E., Bhatt, S. and Moyes, C. L. (2020) 'Mapping trends in insecticide resistance phenotypes in African malaria vectors', *PLoS biology,* 18(6), pp. e3000633.

Hardin, G. (1960) 'The competitive exclusion principle: an idea that took a century to be born has implications in ecology, economics, and genetics', *science,* 131(3409), pp. 1292-1297.

Hatton, I. A., McCann, K. S., Fryxell, J. M., Davies, T. J., Smerlak, M., Sinclair, A. R. and Loreau, M. (2015) 'The predator-prey power law: Biomass scaling across terrestrial and aquatic biomes', *Science,* 349(6252), pp. aac6284.

Hawley, W. A., Phillips-Howard, P. A., ter Kuile, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., Nahlen, B. L., Gimnig, J. E., Kariuki, S. K. and Kolczak, M. S. (2003) 'Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya', *The American journal of tropical medicine and hygiene,* 68(4_suppl), pp. 121-127.

Hemingway, J., Ranson, H., Magill, A., Kolaczinski, J., Fornadel, C., Gimnig, J., Coetzee, M., Simard, F., Roch, D. K. and Hinzoumbe, C. K. (2016) 'Averting a malaria disaster: will insecticide resistance derail malaria control?', *The Lancet,* 387(10029), pp. 1785-1788.

Hess, A., Haves, R. and Tempelis, C. (1968) 'The use of the forage ratio technique in mosquito host preference studies', *Mosquito News,* 28(3).

Ho, B. C., Ewert, A. and Chew, L.-M. (1989) 'Interspecific competition among Aedes aegypti, Ae. albopictus, and Ae. triseriatus (Diptera: Culicidae): larval development in mixed cultures', *Journal of medical entomology,* 26(6), pp. 615-623.

Holstein, M. H. (1954) *Biology of anopheles gambiae: research in French west Africa.* World Health Organization.

Hopcraft, J. G. C., Sinclair, A. and Packer, C. (2005) 'Planning for success: Serengeti lions seek prey accessibility rather than abundance', *Journal of Animal Ecology,* 74(3), pp. 559-566.

Huestis, D. L., Dao, A., Diallo, M., Sanogo, Z. L., Samake, D., Yaro, A. S., Ousman, Y., Linton, Y.-M., Krishna, A. and Veru, L. (2019) 'Windborne long-distance migration of malaria mosquitoes in the Sahel', *Nature,* 574(7778), pp. 404-408.

Huho, B., Briët, O., Seyoum, A., Sikaala, C., Bayoh, N., Gimnig, J., Okumu, F., Diallo, D., Abdulla, S. and Smith, T. (2013) 'Consistently high estimates for the proportion of human exposure to malaria vector populations occurring indoors in rural Africa', *International journal of epidemiology*, 42(1), pp. 235-247.

Hunt, R. H., Coetzee, M. and Fettene, M. (1998) 'The Anopheles gambiae complex: a new species from Ethiopia', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 92(2), pp. 231-235.

Hutchinson, G. E. (1961) 'The paradox of the plankton', *The American Naturalist*, 95(882), pp. 137-145.

Janecek, L. L., Honeycutt, R. L., Adkins, R. M. and Davis, S. K. (1996) 'Mitochondrial gene sequences and the molecular systematics of the artiodactyl subfamily Bovinae', *Molecular phylogenetics and evolution*, 6(1), pp. 107-119.

Juliano, S. A. (1998) 'Species introduction and replacement among mosquitoes: interspecific resource competition or apparent competition?', *Ecology*, 79(1), pp. 255-268.

Kaindoa, E. W., Matowo, N. S., Ngowo, H. S., Mkandawile, G., Mmbando, A., Finda, M. and Okumu, F. O. (2017) 'Interventions that effectively target Anopheles funestus mosquitoes could significantly improve control of persistent malaria transmission in south–eastern Tanzania', *PloS one*, 12(5), pp. e0177807.

Kampango, A., Pinto, J., Abílio, A. P., Machoe, E., Matusse, J. and McCall, P. J. (2023) 'Characterisation of human exposure to nocturnal biting by malaria and arbovirus vectors in a rural community in Chókwè district, southern Mozambique', *Wellcome Open Research*, 8(193), pp. 193.

Katusi, G. C., Hermy, M. R., Makayula, S. M., Ignell, R., Govella, N. J., Hill, S. R. and Mnyone, L. L. (2022) 'Seasonal variation in abundance and blood meal sources of primary and secondary malaria vectors within Kilombero Valley, Southern Tanzania', *Parasites & Vectors*, 15(1), pp. 1-14.

Kaufmann, C. and Briegel, H. (2004) 'Flight performance of the malaria vectors Anopheles gambiae and Anopheles atroparvus', *Journal of vector ecology*, 29, pp. 140-153.

Kawada, H., Dida, G. O., Sonye, G., Njenga, S. M., Mwandawiro, C. and Minakawa, N. (2012) 'Reconsideration of Anopheles rivulorum as a vector of Plasmodium falciparum in western Kenya: some evidence from biting time, blood preference, sporozoite positive rate, and pyrethroid resistance', *Parasites & vectors*, 5, pp. 1-8.

Kay, B., Boreham, P. and Edman, J. (1979) 'Application of the" Feeding index" concept to studies of mosquito host-feeding patterns', *Mosquito News,* 39(1), pp. 68-72.

Kenea, O., Balkew, M. and Gebre-Michael, T. (2011) 'Environmental factors associated with larval habitats of anopheline mosquitoes (Diptera: Culicidae) in irrigation and major drainage areas in the middle course of the Rift Valley, central Ethiopia', *J Vector Borne Dis,* 48(2), pp. 85-92.

Killeen, G., Tami, A., Kihonda, J., Okumu, F., Kotas, M., Grundmann, H., Kasigudi, N., Ngonyani, H., Mayagaya, V. and Nathan, R. (2007a) 'Cost-sharing strategies combining targeted public subsidies with private-sector delivery achieve high bednet coverage and reduced malaria transmission in Kilombero Valley, southern Tanzania', *BMC Infectious Diseases,* 7(1), pp. 1-20.

Killeen, G. F. (2003) 'Following in Soper's footsteps: northeast Brazil 63 years after eradication of Anopheles gambiae', *The Lancet Infectious Diseases,* 3(10), pp. 663-666.

Killeen, G. F. (2014) 'Characterizing, controlling and eliminating residual malaria transmission', *Malaria journal,* 13, pp. 1-22.

Killeen, G. F. (2023) *Population-stabilizing portfolio effects of fine-scale environmental variations in natural resource availability to malaria vector mosquitoes; Characterization and implications for malaria vector control strategies*: University College Cork (Accessed: September 2023).

Killeen, G. F. and Chitnis, N. (2014) 'Potential causes and consequences of behavioural resilience and resistance in malaria vector populations: a mathematical modelling analysis', *Malaria Journal,* 13(1), pp. 1-16.

Killeen, G. F., Fillinger, U., Kiche, I., Gouagna, L. C. and Knols, B. G. (2002) 'Eradication of Anopheles gambiae from Brazil: lessons for malaria control in Africa?', *The Lancet infectious diseases,* 2(10), pp. 618-627.

Killeen, G. F., Govella, N. J., Lwetoijera, D. W. and Okumu, F. O. (2016) 'Most outdoor malaria transmission by behaviourally-resistant Anopheles arabiensis is mediated by mosquitoes that have previously been inside houses', *Malaria journal,* 15, pp. 1-10.

Killeen, G. F., Kihonda, J., Lyimo, E., Oketch, F. R., Kotas, M. E., Mathenge, E., Schellenberg, J. A., Lengeler, C., Smith, T. A. and Drakeley, C. J. (2006) 'Quantifying behavioural interactions between humans and mosquitoes: evaluating the protective efficacy of insecticidal nets against malaria transmission in rural Tanzania', *BMC infectious diseases,* 6(1), pp. 1-10.

Killeen, G. F., Kiware, S. S., Okumu, F. O., Sinka, M. E., Moyes, C. L., Massey, N. C., Gething, P. W., Marshall, J. M., Chaccour, C. J. and Tusting, L. S. (2017a) 'Going beyond personal protection against mosquito bites to eliminate malaria transmission: population suppression of malaria vectors that exploit both human and animal blood', *BMJ Global Health,* 2(2), pp. e000198.

Killeen, G. F., Marshall, J. M., Kiware, S. S., South, A. B., Tusting, L. S., Chaki, P. P. and Govella, N. J. (2017b) 'Measuring, manipulating and exploiting behaviours of adult mosquitoes to optimise malaria vector control impact', *BMJ Global Health,* 2(2), pp. e000212.

Killeen, G. F., McKenzie, F. E., Foy, B. D., Bøgh, C. and Beier, J. C. (2001) 'The availability of potential hosts as a determinant of feeding behaviours and malaria transmission by African mosquito populations', *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 95(5), pp. 469-476.

Killeen, G. F. and Reed, T. E. (2018) 'The portfolio effect cushions mosquito populations and malaria transmission against vector control interventions', *Malaria Journal,* 17, pp. 1-7.

Killeen, G. F., Seyoum, A., Sikaala, C., Zomboko, A. S., Gimnig, J. E., Govella, N. J. and White, M. T. (2013) 'Eliminating malaria vectors', *Parasites & vectors,* 6(1), pp. 1-10.

Killeen, G. F., Smith, T. A., Ferguson, H. M., Mshinda, H., Abdulla, S., Lengeler, C. and Kachur, S. P. (2007b) 'Preventing childhood malaria in Africa by protecting adults from mosquitoes with insecticide-treated nets', *PLoS medicine,* 4(7), pp. e229.

Killeen, G. F., Tatarsky, A., Diabate, A., Chaccour, C. J., Marshall, J. M., Okumu, F. O., Brunner, S., Newby, G., Williams, Y. A. and Malone, D. (2017c) 'Developing an expanded vector control toolbox for malaria elimination', *BMJ Global Health,* 2(2), pp. e000211.

Kirby, M. J. and Lindsay, S. W. (2009) 'Effect of temperature and inter-specific competition on the development and survival of Anopheles gambiae sensu stricto and An. arabiensis larvae', *Acta tropica,* 109(2), pp. 118-123.

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E. and Sachs, J. (2004) 'A global index representing the stability of malaria transmission', *The American journal of tropical medicine and hygiene,* 70(5), pp. 486-498.

Kitau, J., Oxborough, R. M., Tungu, P. K., Matowo, J., Malima, R. C., Magesa, S. M., Bruce, J., Mosha, F. W. and Rowland, M. W. (2012) 'Species shifts in the Anopheles gambiae complex: do LLINs successfully control Anopheles arabiensis?', *PloS one,* 7(3), pp. e31481.

Kiware, S. S., Russell, T. L., Mtema, Z. J., Malishee, A. D., Chaki, P., Lwetoijera, D., Chanda, J., Chinula, D., Majambere, S. and Gimnig, J. E. (2016) 'A generic schema and data collection forms applicable to diverse entomological studies of mosquitoes', *Source code for biology and medicine,* 11, pp. 1-11.

Koenraadt, C., Majambere, S., Hemerik, L. and Takken, W. (2004) 'The effects of food and space on the occurrence of cannibalism and predation among larvae of Anopheles gambiae sl', *Entomologia Experimentalis et Applicata,* 112(2), pp. 125-134.

Koenraadt, C. and Takken, W. (2003) 'Cannibalism and predation among larvae of the Anopheles gambiae complex', *Medical and veterinary entomology,* 17(1), pp. 61-66.

Kouznetsov, R. (1977) 'Malaria control by application of indoor spraying of residual insecticides in tropical Africa and its impact on community health', *Tropical doctor,* 7(2), pp. 81-91.

Kukielka, E. A., Jori, F., Martínez-López, B., Chenais, E., Masembe, C., Chavernac, D. and Ståhl, K. (2016) 'Wild and domestic pig interactions at the wildlife–livestock interface of Murchison Falls national Park, Uganda, and the potential association with African swine fever outbreaks', *Frontiers in veterinary science,* 3, pp. 31.

Kweka, E. J., Mazigo, H. D., Lyaruu, L. J., Mausa, E. A., Venter, N., Mahande, A. M. and Coetzee, M. (2020) 'Anopheline mosquito species composition, kdr mutation frequency, and parasite infectivity status in northern Tanzania', *Journal of Medical Entomology,* 57(3), pp. 933-938.

Kweka, E. J., Zhou, G., Munga, S., Lee, M.-C., Atieli, H. E., Nyindo, M., Githeko, A. K. and Yan, G. (2012) 'Anopheline larval habitats seasonality and species distribution: a prerequisite for effective targeted larval habitats control programmes', *PloS one,* 7(12), pp. e52084.

Lehane, M. J. (2005) *The biology of blood-sucking in insects.* Cambridge University Press.

Lengeler, C. (2004) 'Insecticide-treated bed nets and curtains for preventing malaria', *Cochrane database of systematic reviews,* (2).

Lindsay, S., Parson, L. and Thomas, C. (1998) 'Mapping the range and relative abundance of the two principal African malaria vectors, Anopheles gambiae sensu stricto and An. arabiensis, using climate data', *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 265(1399), pp. 847-854.

Lotka, A. J. (1932) 'The growth of mixed populations: two species competing for a commonfood supply:' *Journal of the Washington Academy of Sciences,* 22, pp. 461–469.

Lounibos, L. (2007) 'Competitive displacement and reduction', *Journal of the American Mosquito Control Association,* 23(2 Suppl), pp. 276.

Loveridge, A. J., Valeix, M., Davidson, Z., Murindagomo, F., Fritz, H. and Macdonald, D. W. (2009) 'Changes in home range size of African lions in relation to pride size and prey biomass in a semi-arid savanna', *Ecography,* 32(6), pp. 953-962.

Lwetoijera, D. W., Harris, C., Kiware, S. S., Dongus, S., Devine, G. J., McCall, P. J. and Majambere, S. (2014) 'Increasing role of Anopheles funestus and Anopheles arabiensis in malaria transmission in the Kilombero Valley, Tanzania', *Malaria journal,* 13(1), pp. 1-10.

Lyimo, I. N. and Ferguson, H. M. (2009) 'Ecological and evolutionary determinants of host species choice in mosquito vectors', *Trends in parasitology,* 25(4), pp. 189-196.

Macdonald, G. (1957) 'The epidemiology and control of malaria', *The Epidemiology and Control of Malaria.*

Maia, M. F., Kreppel, K., Mbeyela, E., Roman, D., Mayagaya, V., Lobo, N. F., Ross, A. and Moore, S. J. (2016) 'A crossover study to evaluate the diversion of malaria vectors in a community with incomplete coverage of spatial repellents in the Kilombero Valley, Tanzania', *Parasites & vectors,* 9(1), pp. 1-13.

Main, B. J., Lee, Y., Ferguson, H. M., Kreppel, K. S., Kihonda, A., Govella, N. J., Collier, T. C., Cornel, A. J., Eskin, E. and Kang, E. Y. (2016) 'The genetic basis of host preference and resting behavior in the major African malaria vector, Anopheles arabiensis', *PLoS Genetics,* 12(9), pp. e1006303.

Mapua, S. A., Hape, E. E., Kihonda, J., Bwanary, H., Kifungo, K., Kilalangongono, M., Kaindoa, E. W., Ngowo, H. S. and Okumu, F. O. (2022) 'Persistently high proportions of Plasmodium-infected Anopheles funestus mosquitoes in two villages in the Kilombero valley, South-Eastern Tanzania', *Parasite Epidemiology and Control,* 18, pp. e00264.

Masalu, J. P., Finda, M., Okumu, F. O., Minja, E. G., Mmbando, A. S., Sikulu-Lord, M. T. and Ogoma, S. B. (2017) 'Efficacy and user acceptability of transfluthrin-treated sisal and hessian decorations for protecting against mosquito bites in outdoor bars', *Parasites & vectors,* 10(1), pp. 1-8.

Minakawa, N., Mutero, C. M., Githure, J. I., Beier, J. C. and Yan, G. (1999) 'Spatial distribution and habitat characterization of anopheline mosquito larvae in Western Kenya', *The American journal of tropical medicine and hygiene,* 61(6), pp. 1010-1016.

Minakawa, N., Seda, P. and Yan, G. (2002) 'Influence of host and larval habitat distribution on the abundance of African malaria vectors in western Kenya', *The American journal of tropical medicine and hygiene,* 67(1), pp. 32-38.

Mmbando, A. S., Ngowo, H., Limwagu, A., Kilalangongono, M., Kifungo, K. and Okumu, F. O. (2018) 'Eave ribbons treated with the spatial repellent, transfluthrin, can effectively protect against indoor-biting and outdoor-biting malaria mosquitoes', *Malaria journal,* 17, pp. 1-14.

Ministry of Natural Resources and Tourism (2022) 'Birth of ILUMA Wildlife Management Area', *Kilombero and Lower Rufiji Wetlands Ecosystem Project Newsletter*.

Moffett, A., Shackelford, N. and Sarkar, S. (2007) 'Malaria in Africa: vector species' niche models and relative risk maps', *PloS one,* 2(9), pp. e824.

Molineaux, L., Shidrawi, G., Clarke, J., Boulzaguet, J. and Ashkar, T. (1979) 'Assessment of insecticidal impact on the malaria mosquito's vectorial capacity, from data on the man-biting rate and age-composition', *Bulletin of the World Health Organization,* 57(2), pp. 265.

Mombo, F., Speelman, S., Huylenbroeck, G. V., Hella, J. and Pantaleo, M. (2011) 'Ratification of the Ramsar convention and sustainable wetlands management: Situation analysis of the Kilombero Valley wetlands in Tanzania'.

Mosha, J. F., Kulkarni, M. A., Lukole, E., Matowo, N. S., Pitt, C., Messenger, L. A., Mallya, E., Jumanne, M., Aziz, T. and Kaaya, R. (2022) 'Effectiveness and cost-effectiveness against malaria of three types of dual-active-ingredient long-lasting insecticidal nets (LLINs) compared with pyrethroid-only LLINs in Tanzania: a four-arm, cluster-randomised trial', *The Lancet,* 399(10331), pp. 1227-1241.

Moyes, C. L., Athinya, D. K., Seethaler, T., Battle, K. E., Sinka, M., Hadi, M. P., Hemingway, J., Coleman, M. and Hancock, P. A. (2020) 'Evaluating insecticide resistance across African districts to aid malaria control decisions', *Proceedings of the National Academy of Sciences,* 117(36), pp. 22042-22050.

Muirhead-Thomson, R. (1960) 'The significance of irritability, behaviouristic avoidance and allied phenomena in malaria eradication', *Bulletin of the World Health Organization,* 22(6), pp. 721.

Munhenga, G., Brooke, B. D., Chirwa, T. F., Hunt, R. H., Coetzee, M., Govender, D. and Koekemoer, L. L. (2011) 'Evaluating the potential of the sterile insect technique for malaria control: relative fitness and mating compatibility between laboratory colonized and a wild population of Anopheles arabiensis from the Kruger National Park, South Africa', *Parasites & Vectors,* 4(1), pp. 1-11.

Munhenga, G., Brooke, B. D., Spillings, B., Essop, L., Hunt, R. H., Midzi, S., Govender, D., Braack, L. and Koekemoer, L. L. (2014) 'Field study site selection, species abundance and monthly distribution of anopheline mosquitoes in the northern Kruger National Park, South Africa', *Malaria journal,* 13, pp. 1-11.

Muriu, S. M., Coulson, T., Mbogo, C. M. and Godfray, H. C. J. (2013) 'Larval density dependence in Anopheles gambiae ss, the major African vector of malaria', *The Journal of animal ecology,* 82(1), pp. 166.

Mutuku, F. M., King, C. H., Mungai, P., Mbogo, C., Mwangangi, J., Muchiri, E. M., Walker, E. D. and Kitron, U. (2011) 'Impact of insecticide-treated bed nets on malaria transmission indices on the south coast of Kenya', *Malaria journal,* 10(1), pp. 1-14.

Mwangangi, J. M., Mbogo, C. M., Muturi, E. J., Nzovu, J. G., Githure, J. I., Yan, G., Minakawa, N., Novak, R. and Beier, J. C. (2007) 'Spatial distribution and habitat characterisation of Anopheles larvae along the Kenyan coast', *Journal of Vector Borne Diseases,* 44(1), pp. 44.

Mwangangi, J. M., Mbogo, C. M., Orindi, B. O., Muturi, E. J., Midega, J. T., Nzovu, J., Gatakaa, H., Githure, J., Borgemeister, C. and Keating, J. (2013) 'Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years', *Malaria journal,* 12, pp. 1-9.

Nambunga, I. H., Ngowo, H. S., Mapua, S. A., Hape, E. E., Msugupakulya, B. J., Msaky, D. S., Mhumbira, N. T., Mchwembo, K. R., Tamayamali, G. Z. and Mlembe, S. V. (2020) 'Aquatic habitats of the malaria vector Anopheles funestus in rural south-eastern Tanzania', *Malaria journal,* 19, pp. 1-11.

Odetoyinbo, J. and Davidson, G. (1968) 'The Anopheles gambiae complex and its role in the malaria transmission in the islands of Zanzibar and Pemba, United Republic of Tanzania', *Unpublished World Health Organization Document,* WHO/Mal/68.660.

Ogoma, S. B., Mmando, A. S., Swai, J. K., Horstmann, S., Malone, D. and Killeen, G. F. (2017) 'A low technology emanator treated with the volatile pyrethroid transfluthrin confers long term protection against outdoor biting vectors of lymphatic filariasis, arboviruses and malaria', *PLoS neglected tropical diseases,* 11(4), pp. e0005455.

Okumu, F. and Finda, M. (2021) 'Key characteristics of residual malaria transmission in two districts in south-eastern Tanzania—implications for improved control', *The Journal of Infectious Diseases,* 223(Supplement_2), pp. S143-S154.

Okumu, F. O., Kiware, S. S., Moore, S. J. and Killeen, G. F. (2013a) 'Mathematical evaluation of community level impact of combining bed nets and indoor residual spraying upon malaria

transmission in areas where the main vectors are Anopheles arabiensis mosquitoes', *Parasites & vectors,* 6(1), pp. 1-13.

Okumu, F. O., Mbeyela, E., Lingamba, G., Moore, J., Ntamatungiro, A. J., Kavishe, D. R., Kenward, M. G., Turner, E., Lorenz, L. M. and Moore, S. J. (2013b) 'Comparative field evaluation of combinations of long-lasting insecticide treated nets and indoor residual spraying, relative to either method alone, for malaria prevention in an area where the main vector is Anopheles arabiensis', *Parasites & vectors,* 6, pp. 1-20.

Onori, E. and Grab, B. (1980) 'Indicators for the forecasting of malaria epidemics', *Bulletin of the World Health Organization,* 58(1), pp. 91.

Orsdol, K. v., Hanby, J. P. and Bygott, J. (1985) 'Ecological correlates of lion social organization (Panthers, leo)', *Journal of zoology*, 206(1), pp. 97-112.

O'Malley, C. 'Guidelines for larval surveillance'. *Proceedings of the Seventy-Sixth Annual Meeting of the New Jersey Mosquito Control Association, Inc*, 44-55.

Parmakelis, A., Russello, M. A., Caccone, A., Marcondes, C. B., Costa, J., Forattini, O. P., Mureb Sallum, M. A., Wilkerson, R. C. and Powell, J. R. (2008) 'Historical analysis of a near disaster: Anopheles gambiae in Brazil', *American Journal of Tropical Medicine and Hygiene,* 78(1), pp. 176.

Paterson, H., Paterson, J. and Van Eeden, G. 'A New Member of the Anopheles gambiae Complex. A Preliminary Report'. *Medical Proceedings*: Johannesburg, 414-18.

Pates, H. and Curtis, C. (2005) 'Mosquito behavior and vector control', *Annu. Rev. Entomol.,* 50, pp. 53-70.

Pates, H., Takken, W., Curtis, C., Huisman, P., Akinpelu, O. and Gill, G. (2001) 'Unexpected anthropophagic behaviour in Anopheles quadriannulatus', *Medical and Veterinary Entomology*, 15(3), pp. 293-298.

Pluess, B., Tanser, F. C., Lengeler, C. and Sharp, B. L. (2010) 'Indoor residual spraying for preventing malaria', *Cochrane database of systematic reviews*, (4).

Port, G., Boreham, P. and Bryan, J. H. (1980) 'The relationship of host size to feeding by mosquitoes of the Anopheles gambiae Giles complex (Diptera: Culicidae)', *Bulletin of Entomological Research,* 70(1), pp. 133-144.

Prior, A. and Torr, S. (2002) 'Host selection by Anopheles arabiensis and An. quadriannulatus feeding on cattle in Zimbabwe', *Medical and Veterinary Entomology*, 16(2), pp. 207-213.

Pulliam, H. R. (1988) 'Sources, sinks, and population regulation', *The American Naturalist,* 132(5), pp. 652-661.

Qureshi, A. and Connolly, J. B. (2021) 'A systematic review assessing the potential for release of vector species from competition following insecticide-based population suppression of Anopheles species in Africa', *Parasites & Vectors,* 14, pp. 1-14.

Ranson, H. and Lissenden, N. (2016) 'Insecticide resistance in African Anopheles mosquitoes: a worsening situation that needs urgent action to maintain malaria control', *Trends in parasitology,* 32(3), pp. 187-196.

Richerson, P., Armstrong, R. and Goldman, C. R. (1970) 'Contemporaneous disequilibrium, a new hypothesis to explain the „Paradox of the Plankton"', *Proceedings of the National Academy of Sciences,* 67(4), pp. 1710-1714.

Riginos, C. and Grace, J. B. (2008) 'Savanna tree density, herbivores, and the herbaceous community: Bottom-up vs. top-down effects', *Ecology,* 89(8), pp. 2228-2238.

RIS 2002. Information Sheet on RAMSAR Wetlands. TZ1173RIS. Available at: https://rsis.ramsar.org/RISapp/files/RISrep/TZ1173RIS.pdf

Roy, S. and Chattopadhyay, J. (2007) 'Towards a resolution of 'the paradox of the plankton': A brief overview of the proposed mechanisms', *Ecological complexity,* 4(1-2), pp. 26-33.

Rozendaal, J. A. (1997) *Vector control: methods for use by individuals and communities.* World Health Organization.

Russell, T. L., Govella, N. J., Azizi, S., Drakeley, C. J., Kachur, S. P. and Killeen, G. F. (2011) 'Increased proportions of outdoor feeding among residual malaria vector populations following increased use of insecticide-treated nets in rural Tanzania', *Malaria journal,* 10(1), pp. 1-10.

Russell, T. L., Lwetoijera, D. W., Maliti, D., Chipwaza, B., Kihonda, J., Charlwood, J. D., Smith, T. A., Lengeler, C., Mwanyangala, M. A. and Nathan, R. (2010) 'Impact of promoting longer-lasting insecticide treatment of bed nets upon malaria transmission in a rural Tanzanian setting with pre-existing high coverage of untreated nets', *Malaria journal,* 9(1), pp. 1-14.

Sanogo, Z. L., Yaro, A. S., Dao, A., Diallo, M., Yossi, O., Samaké, D., Krajacich, B. J., Faiman, R. and Lehmann, T. (2021) 'The effects of high-altitude windborne migration on survival, oviposition, and blood-feeding of the African malaria mosquito, Anopheles gambiae sl (Diptera: Culicidae)', *Journal of Medical Entomology,* 58(1), pp. 343-349.

Scheffer, M., Rinaldi, S., Huisman, J. and Weissing, F. J. (2003) 'Why plankton communities have no equilibrium: solutions to the paradox', *Hydrobiologia,* 491, pp. 9-18.

Schindler, D. E., Armstrong, J. B. and Reed, T. E. (2015) 'The portfolio concept in ecology and evolution', *Frontiers in Ecology and the Environment,* 13(5), pp. 257-263.

Schneider, P., Takken, W. and McCall, P. (2000) 'Interspecific competition between sibling species larvae of Anopheles arabiensis and An. gambiae', *Medical and Veterinary Entomology,* 14(2), pp. 165-170.

Schoener, T. W. (1983) 'Field experiments on interspecific competition', *The american naturalist,* 122(2), pp. 240-285.

Scott, J. A., Brogdon, W. G. and Collins, F. H. (1993) 'Identification of single specimens of the Anopheles gambiae complex by the polymerase chain reaction', *The American journal of tropical medicine and hygiene,* 49(4), pp. 520-529.

Service, M. (1993) *Mosquito ecology: field sampling methods.* Second edn.: Springer.

Service, M. (1997) 'Mosquito (Diptera: Culicidae) dispersal—the long and short of it', *Journal of medical entomology,* 34(6), pp. 579-588.

Service, M. W. and Townson, H. (2002) 'The anopheles vector', *Essential malariology*: CRC Press, pp. 59-84.

Seyoum, A., Sikaala, C. H., Chanda, J., Chinula, D., Ntamatungiro, A. J., Hawela, M., Miller, J. M., Russell, T. L., Briët, O. J. and Killeen, G. F. (2012) 'Human exposure to anopheline mosquitoes occurs primarily indoors, even for users of insecticide-treated nets in Luangwa Valley, South-east Zambia', *Parasites & vectors,* 5, pp. 1-10.

Sherrard-Smith, E., Skarp, J. E., Beale, A. D., Fornadel, C., Norris, L. C., Moore, S. J., Mihreteab, S., Charlwood, J. D., Bhatt, S. and Winskill, P. (2019) 'Mosquito feeding behavior and how it influences residual malaria transmission across Africa', *Proceedings of the National Academy of Sciences,* 116(30), pp. 15086-15095.

Shililu, J., Mbogo, C., Ghebremeskel, T., Githure, J. and Novak, R. (2007) 'Mosquito Larval Habitats in Semiarid Ecosystem in Eritrea: Impact of Larval Habitat Management on Anopheles arabiensis Population'.

Shousha, A. T. (1948) 'Species-eradication: The Eradication of Anopheles gambioe from Upper Egypt, 1942-1945', *Bulletin of the World Health Organization,* 1(2), pp. 309.

Simard, F., Ayala, D., Kamdem, G. C., Pombi, M., Etouna, J., Ose, K., Fotsing, J.-M., Fontenille, D., Besansky, N. J. and Costantini, C. (2009) 'Ecological niche partitioning between Anopheles gambiae molecular forms in Cameroon: the ecological side of speciation', *BMC ecology*, 9, pp. 1-24.

Sinka, M. E., Bangs, M. J., Manguin, S., Coetzee, M., Mbogo, C. M., Hemingway, J., Patil, A. P., Temperley, W. H., Gething, P. W. and Kabaria, C. W. (2010) 'The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis', *Parasites & vectors*, 3, pp. 1-34.

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., Mbogo, C. M., Hemingway, J., Patil, A. P. and Temperley, W. H. (2012) 'A global map of dominant malaria vectors', *Parasites & vectors*, 5(1), pp. 1-11.

Sinka, M. E., Golding, N., Massey, N. C., Wiebe, A., Huang, Z., Hay, S. I. and Moyes, C. L. (2016) 'Modelling the relative abundance of the primary African vectors of malaria before and after the implementation of indoor, insecticide-based vector control', *Malaria journal*, 15(1), pp. 1-10.

Smith, A. (1962) 'Malaria in the Taveta Area of Kenya and Tanganyika. Part III. Entomological Findings Three Years after the Spraying Period', *East African medical journal*, 39(9), pp. 553-64.

Smith, D., Dushoff, J., Snow, R. and Hay, S. (2005) 'The entomological inoculation rate and Plasmodium falciparum infection in African children', *Nature*, 438(7067), pp. 492-495.

Smith, T., Charlwood, J., Kihonda, J., Mwankusye, S., Billingsley, P., Meuwissen, J., Lyimo, E., Takken, W., Teuscher, T. and Tanner, M. (1993) 'Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission', *Acta tropica*, 54(1), pp. 55-72.

Smith, T., Charlwood, J., Takken, W., Tanner, M. and Spiegelhalter, D. (1995) 'Mapping the densities of malaria vectors within a single village', *Acta tropica*, 59(1), pp. 1-18.

Smith, T., Maire, N., Dietz, K., Killeen, G. F., Vounatsou, P., Molineaux, L. and Tanner, M. (2006) 'Relationship between the entomologic inoculation rate and the force of infection for Plasmodium falciparum malaria', *The American journal of tropical medicine and hygiene*, 75(2_suppl), pp. 11-18.

Soper, F. L. and Wilson, D. B. (1943) *Anopheles gambiae in Brazil, 1930 to 1940*. Rockefeller Foundation.

Sougoufara, S., Harry, M., Doucouré, S., Sembène, P. and Sokhna, C. (2016) 'Shift in species composition in the Anopheles gambiae complex after implementation of long-lasting insecticidal nets in Dielmo, Senegal', *Medical and veterinary entomology*, 30(3), pp. 365-368.

Spong, G. (2002) 'Space use in lions, Panthera leo, in the Selous Game Reserve: social and ecological factors', *Behavioral Ecology and Sociobiology*, 52, pp. 303-307.

Swai, J. K., Soto, A. C., Ntabaliba, W. S., Kibondo, U. A., Ngonyani, H. A., Mseka, A. P., Ortiz, A., Chura, M. R., Mascari, T. M. and Moore, S. J. (2023) 'Efficacy of the spatial repellent product Mosquito Shield™ against wild pyrethroid-resistant Anopheles arabiensis in south-eastern Tanzania', *Malaria Journal*, 22(1), pp. 249.

Takken, W., Charlwood, J., Billingsley, P. and Gort, G. (1998) 'Dispersal and survival of Anopheles funestus and A. gambiae sl (Diptera: Culicidae) during the rainy season in southeast Tanzania', *Bulletin of entomological research*, 88(5), pp. 561-566.

Takken, W., Eling, W., Hooghof, J., Dekker, T., Hunt, R. and Coetzee, M. (1999) 'Susceptibility of Anopheles quadriannulatus theobald (Diptera: Culicidae) to Plasmodium falciparum', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 93(6), pp. 578-580.

Takken, W. and Knols, B. G. (1999) 'Odor-mediated behavior of Afrotropical malaria mosquitoes', *Annual review of entomology*, 44(1), pp. 131-157.

Takken, W. and Verhulst, N. O. (2013) 'Host preferences of blood-feeding mosquitoes', *Annual review of entomology*, 58, pp. 433-453.

TANAPA (2019) *Tanzania National Parks*: Tanzania National Parks Authority. Available at: https://www.tanzaniaparks.go.tz/national_parks/nyerere-national-park (Accessed: July 2023).

Tilman, D. (1982) *Resource competition and community structure.* Princeton university press.

Torr, S., Della Torre, A., Calzetta, M., Costantini, C. and Vale, G. (2008) 'Towards a fuller understanding of mosquito behaviour: use of electrocuting grids to compare the odour-orientated responses of Anopheles arabiensis and An. quadriannulatus in the field', *Medical and veterinary entomology*, 22(2), pp. 93-108.

Valeix, M., Fritz, H., Loveridge, A. J., Davidson, Z., Hunt, J. E., Murindagomo, F. and Macdonald, D. W. (2009) 'Does the risk of encountering lions influence African herbivore behaviour at waterholes?', *Behavioral Ecology and Sociobiology*, 63, pp. 1483-1494.

Vazquez-Prokopec, G. M., Galvin, W. A., Kelly, R. and Kitron, U. (2009) 'A new, cost-effective, battery-powered aspirator for adult mosquito collections', *Journal of medical entomology*, 46(6), pp. 1256-1259.

Volterra, V. (1931) 'Variations and fluctuations of the number of individuals in animal species living together', *Animal ecology*, pp. 412-433.

White, G. (1973) 'Comparative studies on sibling species of the Anopheles gambiae Giles complex (Dipt., Culicidae). III. The distribution, ecology, behaviour and vectorial importance of species D in Bwamba County, Uganda, with an analysis of biological, ecological, morphological and cytogenetical relationships of Ugandan species D', *Bulletin of Entomological Research*, 63(1), pp. 65-97.

White, G. (1974) 'Anopheles gambiae complex and disease transmission in Africa', *Transactions of the Royal Society of Tropical Medicine and hygiene*, 68(4), pp. 278-298.

White, G., Magayuka, S. A. and Boreham, P. (1972) 'Comparative studies on sibling species of the Anopheles gambiae Giles complex (Dipt., Culicidae): bionomics and vectorial activity of species A and species B at Segera, Tanzania', *Bulletin of Entomological Research*, 62(2), pp. 295-317.

White, M. T., Lwetoijera, D., Marshall, J., Caron-Lormier, G., Bohan, D. A., Denholm, I. and Devine, G. J. (2014) 'Negative cross resistance mediated by co-treated bed nets: a potential means of restoring pyrethroid-susceptibility to malaria vectors', *PloS one*, 9(5), pp. e95640.

WHO 2002. Instructions for treatment and use of insecticide-treated mosquito nets: use insecticide-treated mosquito nets to sleep in peace and protect your health.

WHO 2016. WHO malaria terminology: 2021 Update. WHO.

WHO 2017. Achieving and maintaining universal coverage with long-lasting insecticidal nets for malaria control. Global Malaria Programme.

WHO 2021. Global technical strategy for malaria 2016–2030: 2021 Update. WHO.

WHO 2022. World malaria report. Geneva: World Health Organisiation.

WHO (2023) *Malaria*. Available at: https://www.who.int/news-room/fact-sheets/detail/malaria (Accessed: September 2023).

Yohannes, M. and Boelee, E. (2012) 'Early biting rhythm in the afro-tropical vector of malaria, Anopheles arabiensis, and challenges for its control in Ethiopia', *Medical and veterinary entomology,* 26(1), pp. 103-105.

Appendix

Survey Round (RND): _____ **ED1-SEN:**

Form Type (FT): EDSO1; Larval habitat survey for presence of aquatic mosquitoes and environmental attributes **Date (DT):** ____/____/____ **Form Serial Number (SEN):** _____

Survey or Experiment Identifier (SEI): **Cluster/Camp (CR):** _____ **Camp Name or Description (CND):** _____

GPS Identifier (GID): **Camp GPS Way Point Number (CGWPN):** _____ **Camp GPS Coordinates:** Southing (CGCDS); _____ Easting (CGCDE) _____

Camp Comments (CC): _____

Habitat type (HT) codes:

1: Small puddles, tracks and depressions
2: Swampy areas, seepages and springs
3: Flowing stream bed
4: Flooded valley
5: Pond or water hole
6: Artificial drain or ditch
7: Wells and other human-made holes
8: Water storage & other artificial containers
9: Rice paddy
10: Ridge and furrow agriculture (Matuta)
11: Other agriculture
12: Others (describe)

Form Row (FR)	Habitat GPS Way Point Number (HGWPN)	Habitat Type (HT)	Habitat Description (HD)	Habitat perimeter (HPER)				Number of Dips (ND)	Emerging or Floating Vegetation (EFV)				Water Depth (WD)			Aquatic Stage (AS)									Sample Label Code (SLC)	Sample Type (ST) (Ind=1, Batch=2, Pool=3)	Habitat Comments (HCM)
																Anopheles				Culex			Pupae				
				< 2m(HPER=1)	< 20m (HPER=2)	20-200m (HPER=3)	> 200m (HPER=4)		Absent (EFVA)	Short (EFVS)	Tall (EFVT)	Floating (EFVF)	<0.1m (WD=1)	0.1-0.5m (WD=2)	>0.5m (WD=3)	Absent (ASAA)	Early (ASAE)	Late (ASAL)	Gambiae present (ASAGP)	Absent (ASCA)	Early (ASCE)	Late (ASCL)	Absent (ASPA)	Present (ASPP)			
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1																											
2																											
3																											
4																											
5																											
6																											
7																											
8																											
9																											
10																											
11																											
12																											
13																											
14																											
15																											
16																											
17																											
18																											
19																											
20																											

Appendix 1. The data entry form used to characterise an aquatic habitat and identify the presence or absence of aquatic stage mosquitoes.

Appendix 2. A uniquely developed key developed by Duggan *et al*. (Unpublished) specifically designed for recording various types of detections for humans, human activity, livestock and wildlife in ILUMA WMA and the surrounding villages.

Entity or Activity Observed (EAO)	**Observation Category (OC)**	**Additional Observation Attributes (AOA)**	**Observation Count or Class (OCC)**
1X "As you go" instances of specific human land and/or natural resources usages.			
11: Livestock Herding			**1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
12: Charcoal Burning			**1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
13: Timber Harvesting			**1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
14: Fishing			**1**: Active, **2**: Recent, **3**: Old, **9**: 0Not determined or recorded
15: Hunting			**1**: Active, **2**: Recent, 3: Old, **9**: Not determined or recorded
16: Human Settlement		**1,2,3…10** structures [n], 11-20 structures **[20],** 21-50 structures **[50],** >50 structures **[100]**, not determined or recorded **[9]**	**1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
17: Water Body			**1**: Bwawa. **2**: (Ma)dimbwi outside korongo **3:** (Ma)dimbwi inside korongo **4**: Flowing korongo **5**: Bonde
18: Human dug well			**1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
19: Meat Poaching			**1:** Active**, 2:** Recent, **3:** Old**, 9:** Not determined or recorded
2X: "As you go" sightings, tracks and signs of humans and animals			
21: Humans	1: Direct sighting, 2: Tracks, 3: Spoor, **4**: Other signs (Explain in notes)	**1**: ILUMA personnel, **2**: NNP Rangers, **3**: Hunting company staff, **4**: Other authorized personnel from ILUMA and other authorized institutions, **5**: Community legal presence inside conservation area, **6**: Community legal presence outside the conservation area, **7**: Charcoal transport, **8**: Timber harvesting, **9**: Timber transport, **10**: Fish harvesting, **11**: Fish transport, **12**: Community illegal presence inside the conservation area, **13:** Charcoal burner.	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
22: Livestock and work animals	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other	**1**: Cattle, **2**: Goats, **3**: Sheep, **4**: Rooster, **5**: Chicken, **6**: Dog	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded

23: Wild Herbivores	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other signs (Explain in notes)	1X: Reduncini (**11**: Reedbuck, **12**: Waterbuck, **13**: Puku), 2X: Spiral Horned Antelopes (**21**: Bushbuk, **22**: Eland, **23**: Greater Kudu) 3X: Alcephalines (**31**: Hartebeest, **32**: Wildebeest), 4X: Hippotragini (**41**: Sable Antelope), 5X: Suidae (**51**: Warthog, **52**: Bushpig), 6X Plantigrade non-ungilates (**61**: Hippo, **62**: Elephant), 7X: Bovinae (**71**: African Buffalo), 8X: Equidae (**81**: Plains Zebra), 9X: Neotragini (**91**: Suni), 10X: Cephalophini (**101**: Common Duiker, **102**: Red Duiker), 11X: Aepycerotini (**111**: Common Impala), 12X: Madoquini (**121**: Dikdik), 13X: Raphicerini (**131**: Sharpe's Grysbok), 14X: Oreotragini (**141**: Klipspringer)	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
24:Wild Carnivores	1: Direct sighting, 2: Tracks, 3: Spoor, 4: Other signs (Explain in notes)	1X: Pantherinae (**11**: Lion, **12**: Leopard), 2X: Felidae (**21**: African wild cat), 3X: Viverridae (**31**: African Civet, **32**: Cape Genet), 4X: Hyaenidae (**41**: Spotted Hyena), 5X: Canidae (**51**: African wild dog, **52**: Side - Striped Jackal), 6X: Herpestidae (**61**: Slender Mongoose, **62**: Banded Mongoose **63:** Water Mongoose, **64:** White Tailed Mongoose), 7X: Mustelids (**71**: Honey Badger), 8X Lutrinae (**81**: African Clawless Otter)	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
25:Wild Primates and Prosimians	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other	**1**: Lesser Galago, **2**: Greater Galago, **3**: Baboon, **4**: Vervet, **5**: Blue Monkey	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
26: Rodents	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other	1: Hystricidae (**11**: Porcupine, **12**: Brush tailed porcupine)	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
27: Macroscelidea	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other	1: Rhynchoninae (**11:** Sengei)	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
28: Tubulidentata	**1**: Direct sighting, **2**: Tracks, **3**: Spoor, **4**: Other	1: Orycteropodidae (**11:** Aardvark)	Distance (m) from track for direct observations, or age for tracks and signs: **1**: Active, **2**: Recent, **3**: Old, **9**: Not determined or recorded
3X: Summary impressions of land cover attributes averaged over entire segment			
31: Rice farming			**1**: 0%, **2**: 1 - 25%, **3**: 26 - 50%, **4**: 51 - 75%, **5**: 76 - 100%
32: Other Tillage Crops			**1**: 0%, **2**: 1 - 25%, **3**: 26 - 50%, **4**: 51 - 75%, **5**: 76 - 100%
33: Grass and forb height			**1:** <1ft, **2:** 1 - 3ft, **3:** 3 -6ft, **4:** 6ft+
34: Obstruction of visibility by Shrubs, Bushes and Trees			**1**: 0%, **2**: 1 - 25%, **3**: 26 - 50%, **4**: 51 - 75%, **5**: 76 - 100%
35: Ground Type			**1**: Hard & Dry, **2:** Sand, **3:** Damp Soil, **4**: Mud, **5:** Water, **6:** Rock**, 7:** Gravel & Pebbles