Community and Network Ecology of

Biting Diptera-Host Interactions

Ben Bellekom

St Cross College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

April 2023

Declaration of authorship

I declare that the work presented in this thesis is my own, conducted under the supervision of Professor Owen T. Lewis (University of Oxford), and Dr Talya D. Hackett (University of Oxford). I lead-authored all four data chapters in this thesis and have acknowledged the contributions of all co-authors in the author contributions section at the start and end of each chapter. I designed and ran the studies, analysed the data, and wrote the manuscripts, whilst co-authors assisted with data collection, supervision, and review of the manuscripts.

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford. No part of this thesis has already been accepted, or is concurrently being submitted, for any degree, diploma, certificate, or other qualification in this University or elsewhere.

Ben Bellekom

Student no. 1127404

April 2023

Abstract

Biting flies (Diptera) are responsible for the transmission of a wide range of pathogens that infect humans and animals. Biting Diptera are highly taxonomically and ecologically diverse and display a range of blood-feeding behaviours. Most species are generalist and have the potential to take blood meals from hosts of different species through opportunistic feeding. However, some Diptera species also display a consistent host preference at the class or species level. These patterns of host use are not thoroughly documented and have important implications for the transmission of disease within and between host species, and pathways for the emergence and spread of current and novel diseases. Interactions between Diptera and their hosts can now be reliably identified using DNA barcoding and metabarcoding of Diptera blood meals. However, most biting Diptera research has focused on individual species and their interactions with specific hosts. In this thesis I investigate the community and network ecology of biting Diptera – host interactions. First, using a database of Diptera-host interactions compiled from the literature, I highlight the potential insights that can emerge from using ecological interaction networks to examine biting Diptera community composition, interaction structure, and associated disease transmission risks, and how these can change in response to perturbations. Using the same dataset, I then applied a network approach to investigate the effects of macroecological and local anthropogenic factors on biting Diptera-host network structure. I found that the evenness of biting Diptera-host interactions was higher in Village/Urban habitats, but evenness or the overall specialism of the biting Diptera-host networks did not vary on a global scale with latitude. Interaction networks should be populated with the most complete data possible. In compiling my interaction database, I found a gap in the current understanding of how the degradation of host DNA in blood meals as a result of digestion and storage conditions reduces blood meal identification success. I investigated the effect of digestion time and storage condition on the identification success of biting midge blood meals. I found that increased blood meal digestion time, but not the storage condition, reduced the success of host identification. Finally, using biting Diptera collected in Ghana, I examined empirically the effect of anthropogenic landscape modification on biting Diptera community composition and structure. Biting Diptera community composition differed little across an anthropogenic landscape modification gradient, indicating a shared community. In addition,

network structure differed across different levels of habitat modification. I then identified interactions between Diptera competent of vectoring a range of pathogens and susceptible hosts, which may have the potential to act as disease transmission events. The findings reported in this thesis highlight the potential application of ecological networks to research on biting Diptera and provide new insights into the impact of anthropogenic landscape modification on biting Diptera-host community and interaction structure.

Acknowledgements

First, I would like to thank my supervisors Owen Lewis and Talya Hackett, whose help and guidance made this thesis possible. Thank you for helping me get through the hard pandemic years and providing me with endless advice and encouragement. I have truly loved having you both as supervisors and I wouldn't be the scientist I am today without you. I would also like to thank Talya for her help and guidance during my field season and for supplying me with a range of dogs to pet during meetings.

I would like to thank Abby Bailey and Dave Hemprich-Bennett for their company, technical advice, and support during the entirety of my DPhil. My long laboratory sessions would not have been the same without having Abby to chat with, and I'm truly grateful to have had you as a constant source of friendship and encouragement. Thank you also to everyone in the CERO lab group, the weekly meetings, pub trips, and pub quizzes have been a highlight of my time here. This thesis would not have been possible without the dedication and assistance from the Target Malaria team in both the UK and Ghana, thank you for all your hard work. I'd like to thank Fred Aboagye-Antwi for his assistance and guidance throughout this project. The Ghanian field team, Ezekiel, Bernard, and Naana played a pivotal part in the collection of my data, and I am grateful for the huge amount of work you put in. I'd like to thank Samuel Anie-Amoah for his enthusiasm and assistance in collecting field samples. Without you, all the work would not have been possible. I'd also like to thank all the collaborators that contributed to the formation of this thesis.

A massive thank you to all the friends that I have made across the DTP and Wolfson boat club. In particular, my board game and climbing friends, Marcus & Sarah, Ian, Ben, Nacho, David, Mathew, Saran, and Roberta. Your friendship has meant the world to me and kept me sane over the last few years. Thank you to everyone in the Wolfson boat club; training and racing with you all has been an unforgettable experience. Thank you to the pub wolves, James, Adam, Jorn, Mary, the Andrews, Dennis, and Wojciech for your constant source of friendship, advice, and entertainment. I will thoroughly miss our pub trips, and you have all made my DPhil a truly enjoyable experience. Thank you also to my friends Tom, Joe, and Rob for all your encouragement and the laughs we have had together.

Finally, I want to say a huge thank you to my family, Antony, Linda, and Robbie, and partner Christina. You have all been the constant that has got me through a difficult but enjoyable four and half years. None of this would have been possible without the love, support, help and advice you have given me. It has been quite a wild ride getting here, and I can’t thank you enough for always believing in me, even when I didn’t.

For Grandy

Jacobus (John, Zack) Bellekom

14/1/1926 - 13/1/2017

Your unwavering belief in me has led me to heights I would never have imagined.

I wish you were here to see it.

Good man John.

Table of Contents

Chapter 1: General introduction

Views from Abutia Amagame

1.1 The burden of disease on human populations

Vector-borne diseases represent a significant threat to public health, accounting for 17% of the global human disease burden (Chua et al., 2023), and are significant factors influencing population dynamics and welfare of many vertebrates including both wildlife and domestic livestock. Hematophagous (blood feeding) arthropods, such as ticks, fleas, and insects, transmit a wide spectrum of pathogens of epidemiological and epizootic concern, including viruses, bacteria, helminths, and protozoans (Harrington et al., 2020; Santiago et al., 2017). Biting flies (Diptera) (Figure 1.) are competent vectors for many of these pathogens and have the potential to spread them between populations of susceptible hosts. Mitigation of vector-borne disease burden requires a comprehensive understanding of the socioeconomic impacts, methods of transmission, and community ecology of the vectors themselves. In this thesis I explore the ecological interaction networks linking blood feeding Diptera to their vertebrate hosts and consider their epidemiological relevance and implications for future vector control initiatives. As key background to the thesis, in this General Introduction I first review the global importance of vectors, I then describe vector-borne disease transmission cycles and the effect of landscape and climatic change on transmission potential. I then identify existing vector control measures and future control initiatives. Further, I introduce ecological Interaction networks as a means of examining biting Diptera-host interactions and identify key metrics relevant to biting Diptera network analysis. Additionally, I highlight previous and current molecular methods for the detection of biting Diptera-host interactions. I conclude the Introduction by setting out key knowledge gaps and questions that the thesis addresses.

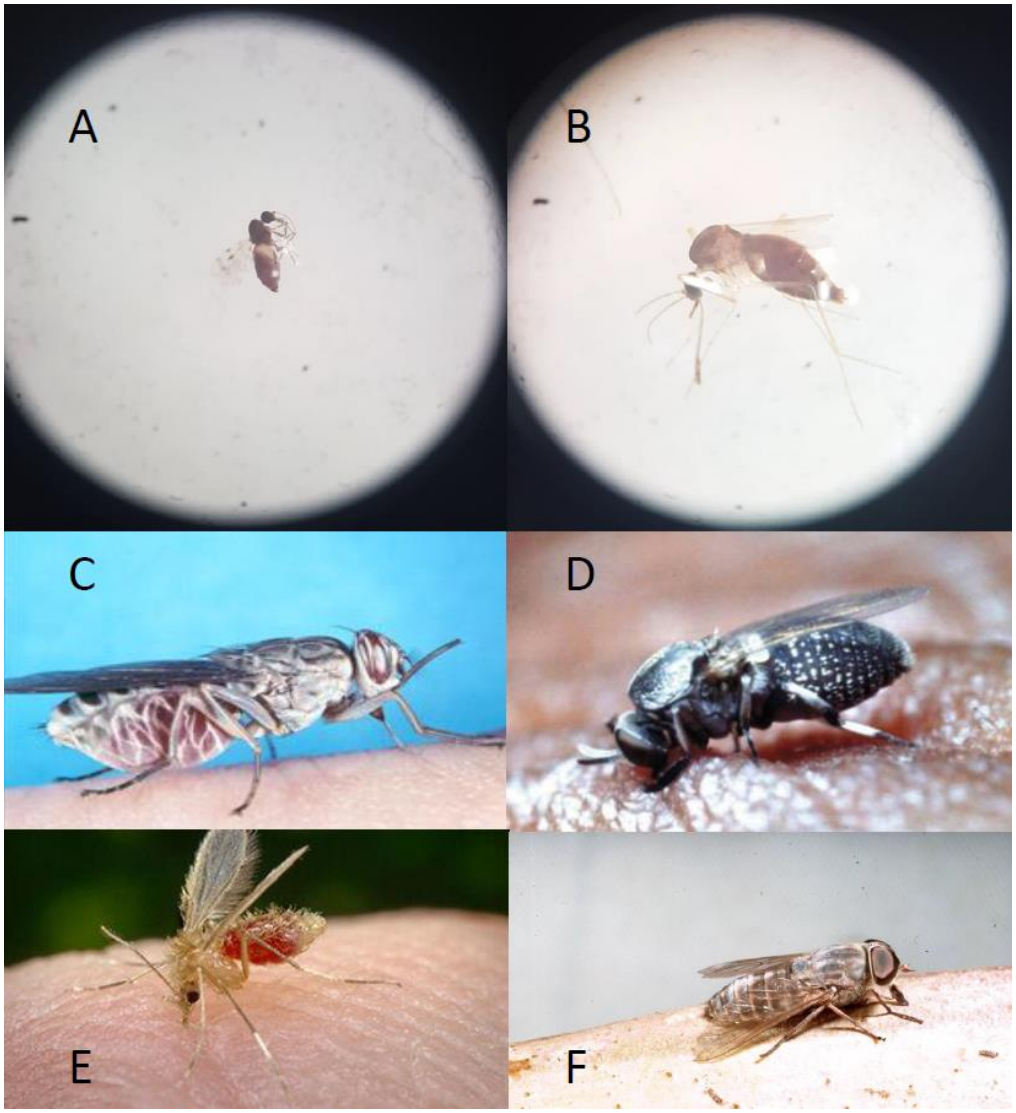


Figure 1. Biting Diptera families of particular epidemiological interest: Biting midge (Ceratopogonidae) (A), Mosquito (Culicidae) (B), Tsetse fly (Glossinidae) (C), Black fly (Simuliidae) (D), Sandfly (Psychodidae) (E), and Tabanid (Tabanidae) (F). Photo source: CDC (C), TDR Image Library (D), ECDC (E), and Frank E French (F).

Despite recent reductions in the prevalence of many vector-borne infectious human diseases as a result of control initiatives and improved medical care and sanitation services (Bhatt et al., 2015; Tizifa et al., 2018), there are still more than 1 billion infection cases and over 700,000 deaths annually (Campbell-Lendrum et al., 2015; WHO, 2020). Indeed, whilst malaria has been the target of widespread elimination efforts, infection has persisted and even increased in certain locations and there were a reported 241 million malaria cases with

an estimated 627,000 malaria linked deaths in 2020, an increase of 69,000 deaths from the previous year (Dhiman & Sarkar, 2017; Dhiman, 2019; Feng et al., 2022). Further, just under half of the global human population is currently living at risk of contracting dengue fever, which accounts for over 40,000 deaths each year. Vector-borne diseases continue to affect the most vulnerable members of society, with children under five accounting for 67% of malaria deaths (UNICEF, 2017). Many vector-borne diseases cause chronic suffering, life-long morbidity and disability. Indeed, one sixth of global disability occurs as a result of vector-borne disease (WHO, 2014). For example, there are approximately 4.1 million people living with leishmaniasis, and 65 million people living with lymphatic filariasis, both of which may cause severe disabilities (Rocklöv & Dubrow, 2020). Vector-borne diseases can also impact a population's reproductive health. The emergence of the Zika virus, a mosquito vectored disease, in French Polynesia, the Pacific islands, and Brazil posed a significant public health threat due to the link between Zika fever and severe congenital malformations in infants (Chouin-Carneiro et al., 2020; Yun & Lee, 2017).

1.2 The economic impacts of human disease

Vectored diseases of humans are economically significant, and their control is directly linked to poverty reduction and economic development. The direct economic burden of malaria alone has been estimated at over $12 billion in Africa, though the lost economic growth is likely to be many times higher (CDC, 2021; Gallup & Sachs, 2001). Indeed, the presence of malaria has been linked with slowed economic development, with countries affected by malaria growing up to 1% less per capita than those without malaria between 1965 and 1990 (Breman et al., 2001) and average income levels a third of those in countries without malaria. The estimated global cost of dengue illness was $8.9 billion in 2013 (Hung et al., 2020). The direct treatment expenses for vectored diseases, such as African trypanosomiasis, may be up to 40% of a rural community's household income and approaching double the minimum level of income in the case of malaria treatment (World Health Organization, 2015).

Multifaceted control and prevention measures are required to combat infection, which include surveillance, implementation of control measures, and medical treatment including

vaccination programmes. However, such programmes can be highly cost-effective; for example, prevention of just 15 cases of West Nile neuroinvasive disease would have been sufficient to offset the cost of emergency aerial spraying following increased West Nile virus (WNV) activity during a 2005 outbreak in California (Barber et al., 2010). Further, prevention measures such as insecticide-treated nets are highly cost-effective and can limit malaria transmission risk for as little as $1.30 per person per year (Yukich et al., 2008), though the efficacy of treated nets and spraying is becoming increasingly hampered by emerging insecticide resistance (Ranson & Lissenden, 2016). The costs associated with the surveillance of epidemiologically important biting Diptera are greatly outweighed by the cost required to respond to the introduction of vector-borne disease (Vazquez-Prokopec et al., 2010).

1.3 The economic impacts of livestock diseases

Vector-borne disease outbreaks may also cause substantial economic losses due to animal deaths and the restrictions on the international movement of ruminant livestock. Some countries, such as Tanzania, rely on the taxation of livestock exports as a main source of government revenue, and livestock rearing is a major form of employment (Peyre et al., 2015). Thus, export bans, due to a vector-borne disease outbreaks, leads to decreasing livestock prices and a subsequent impact on farmers livelihoods. Many diseases, such as Rift Valley Fever (RVF) and Bluetongue disease, are maintained in an enzootic cycle among the vector species and wild and domesticated ruminant species, with a number acting as disease reservoirs (De La Torre et al., 2021; Gould et al., 2017). RVF is a zoonotic (transmissible between vertebrate animals and humans) disease that is vectored by a wide range of mosquito species, primarily of the *Aedes* and *Culex* genera (Biteye et al., 2019), and causes periodic epizootics and epidemics primarily within sub-Saharan Africa (Nielsen et al., 2020). The disease affects domestic ruminants, though occasional spill over to other domestic and wild animals and humans can occur. Economic losses from previous RVF outbreaks are estimated to be as high as $470 million, though exact quantification is difficult and the true losses may be higher (Peyre et al., 2015). Bluetongue virus (BTV) is vectored by *Culicoides* midges and is notably non-zoonotic (Maclachlan et al., 2015). The economic losses attributed to enzootic and epizootic BTV infections are estimated at $3 billion (Alkhamis et al., 2020), All ruminant species are susceptible to BTV, though is it most commonly reported

in sheep. BTV is found globally in temperate and tropic regions with distribution dependent on the occurrence of distinct ecosystems of different midge vectors and the presence of susceptible hosts (Alkhamis et al., 2020). However, recent shifts in the global distribution of BTV have been recorded, with encroachments into novel regions of northern Europe (Rushton & Lyons, 2015).

1.4 The health care and economic burden of vector-borne disease

Over 80% of the global population live in areas at risk of infection by vector-borne diseases and nearly half live in regions at risk from two or more (Golding et al., 2015). However, the majority of the public health and economic burden affects low-income countries in parts of sub-Saharan Africa, Americas, and South Asia, the inhabitants of which may be at risk from 5 or more major vector-borne diseases (Golding et al., 2015). Vector-borne diseases disproportionately affect the poorest populations due to the prevalence of disease in tropical climates, and the limited levels of socioeconomic development and healthcare availability (Campbell-Lendrum et al., 2015; J. Kweka et al., 2019). Indeed the per capita mortality rate from vector-borne disease in developing regions is 300 times greater than those in developed nations (Campbell-Lendrum et al., 2015). The burden of 'major' disease infections, such as malaria, has been strongly negatively linked to the national gross domestic product per person, with 76% of reported malaria cases occurring in just 17 African countries (Benelli & Beier, 2017). Neglected tropical diseases (NTD) are those diseases that have been eradicated in most developed societies but persist in underdeveloped regions (Álvarez-Hernández et al., 2020), such as leishmaniasis and Chikungunya (Oviedo-Pastrana et al., 2017). The burden of NTD also typically falls upon developing African and American countries, and the Middle East (Karimkhani et al., 2016). Neglected tropical diseases affect nearly 2 billion of the world's poorest and marginalized people and account for approximately 200,000 deaths per year. Moreover, NTDs account for 22-25 million disability-adjusted life years (DALY), equivalent to approximately 40% of the malaria linked DALYs (Álvarez-Hernández et al., 2020; Fitzpatrick et al., 2017).

The association between poverty and vector-borne disease in low-income developing nations may lead to self-reinforcing poverty traps for the poorest and most under-privileged

members of the population (Franklinos et al., 2019), in which a significant amount of economic capital is required to escape poverty, and the lack of capital is itself a direct cause of poverty in the future, perpetuating the cycle. Rural populations in developing nations are particularly at risk of these, as they typically rely on subsistence farming, have poor access to health care facilities, and there is a high risk of disease transmission (Bonds et al., 2010; Franklinos et al., 2019). Moreover, the underlying production systems in these regions are directly linked to disease transmission risk. For example, domestic ruminants are a feeding source for vectors of a range of diseases including Rift Valley fever (Biteye et al., 2019). Indeed, efforts to break the poverty trap cycle, via increased development of rural areas and shifts to more profitable produce, may result in unintended increase in transmission risk. For example, economic development in regions of British Guiana led to an increase in rice farming and subsequent displacement of cattle. This resulted in a shift in host usage by the malarial vector *Anopheles aquasalis* to the more abundant humans leading to a localised outbreak of *P. vivax* malaria (Giglioli, 1963). Consequently, vector-borne disease and poverty are intrinsically linked in coupled ecological-economic systems and the widespread presence and transmission of vector-borne disease in developing nations may limit economic development (Franklinos et al., 2019; Ngonghala et al., 2017).

1.5 Vector-borne disease transmission pathways

Pathogen-arthropod transmission systems primarily develop due the female Diptera's need to obtain a blood meal for oogenesis, which facilitates disease infection and transmission to the host (Koenraadt et al., 2021). Vector-borne disease may be maintained in sylvatic (enzootic), rural, and urban cycles (Go et al., 2014). The sylvatic transmission cycle is comprised of transmission of a pathogen by a competent vector from an infected wild reservoir to a naïve wild animal (Valentine et al., 2019). Over 60% of human infectious diseases and 75% of emerging infectious diseases are zoonotic and can be maintained in a sylvatic cycle (Weaver & Barrett, 2004; White & Razgour, 2020). Sylvatic cycles are primarily constrained to sylvatic habitats, though cycles also occur in rural and urban habitats, for example West Nile virus may be maintained in susceptible avian reservoir hosts and competent vectors in these habitats (Paz, 2015; Weaver & Barrett, 2004). Stochastic spillover of pathogens primarily constrained to sylvatic cycles into humans can occur via tangential

transmission when amplification levels are sufficiently high, or when a human encroaches into sylvatic habitats (Hoyos et al., 2021). Whilst humans are often dead-end hosts and are likely not to develop sufficient viremia to transmit the pathogen back to a competent vector, infection can still lead to clinical disease (Weaver & Barrett, 2004). Transmission could occur via an opportunistic vector species acting as a bridge vector between the sylvatic and urban habitat (Miot et al., 2020; Vasilakis et al., 2011). Indeed, the variation in land use from rural or sylvatic to urban can create a wide range of ecotones on a micro and macro spatial scale within which zoonotic agents often circulate, infecting wild and domestic animals, and humans (Hassell et al., 2017). Consequently, the risk of spillover transmission is likely to be significantly higher at the sylvatic-urban interface (Valentine et al., 2019).

Within rural habitats, zoonoses, such as Japanese encephalitis virus and Venezuelan equine encephalitis virus, and their enzootic or bridge vectors, may exploit the abundance of livestock and their proximity to humans to form a rural epizootic transmission cycle. Within these cycles, livestock act as bridge hosts, and, via single-step or two-step transmission, in which a pathogen is transmitted from a reservoir to an intermediate host before transmission to humans (Althouse et al., 2018). This epidemiologically links wildlife with humans (Hassell et al., 2017), increasing the likelihood of spillover and subsequent epidemics (Nguyen-Tien et al., 2019; Paz, 2015; Weaver & Barrett, 2004).

Urban transmission cycles maintain diseases between competent vectors, often anthropophilic Diptera, and domestic animal and human hosts. Urban transmission is particularly prevalent within densely populated, low income tropical regions, in which frequent transmission results in the propagation of epidemics and pandemics and causes substantial morbidity and mortality (Gubler, 2011; Miot et al., 2020). However, transmission of vector-borne pathogens, via urban cycles, can also occur at low levels that are sufficient only to maintain the pathogen in the population. Sylvatic arboviruses may spill over and establish evolutionarily and ecologically independent urban cycles when a sylvatically infected human enters an urban environment and interacts with a competent vector (Grubaugh et al., 2019; Weaver & Reisen, 2010). For example, dengue virus is maintained within a sylvatic cycle of nonhuman primates and enzootic mosquito vectors. However, several hundred years ago, multiple phylogenetically distinct serotypes that do not require sylvatic vectors and hosts became endemic in urban cycles that possess sufficient human

population density to facilitate horizontal transmission (Weaver & Reisen, 2010). Moreover, as established urban serotypes required little to no modification to infect urban vectors and hosts, they can infect wild animals and spillback into sylvatic maintenance cycles, limiting the ability to control or eradicate an arbovirus (Figueiredo, 2019). The elucidation of interactions comprising pathogen-arthropod transmission systems, across all cycles classifications, may aid in the identification of reservoir and amplification hosts, highlight important vector species and identify spillover and spillback transmission events.

1.5.1 The effect of landscape and climatic changes on Diptera and disease transmission

Greater global connectivity, mobility, and the combined effects of rapid urban expansion and climate and environmental change are having a significant impact on the incidence and geographical distribution of existing and emerging vector-borne diseases and their associated Diptera vectors (Franklinos et al., 2019; Kolimenakis et al., 2021; Pybus et al., 2015). Increased mobility through business travel, labour movements, migration, and tourism may disseminate pathogens into non-endemic localities (Pybus et al., 2015). For example, localised cases of dengue fever in Europe are highly correlated with the number of travellers arriving from endemic regions, and occur when the infection is passed to a competent *Aedes albopictus* mosquito resulting in a secondary case of infection when it next feeds on a human (Rocklöv & Dubrow, 2020).

Over the last 50 years, air travel passengers have increased by 9% per annum and shipping traffic has increased by more than 27% since 1993 (Tatem et al., 2006). Increasing global connectivity and mobility has resulted in the establishment of routes between previously unconnected regions. Of particular importance is the increased connectivity between tropical regions in which disease burden is higher and transmission may occur year round (Pybus et al., 2015). Yellow fever has been exported, via air travel, from Africa to Asia, an area in which approximately 2 billion immunologically naïve people co-occur with the competent mosquito vector, *Ae. aegypti*, resulting in a major transmission risk (Näslund et al., 2021). Modern methods of transportation can facilitate the escape and establishment of Diptera vectors at a greater rate and range than seen before; with significant ramifications for disease transmission as biological invasion of disease vectors result in the expansion or

re-emergence of vector-borne diseases (Gratz, 1999). For example, whilst malaria is endemic in Brazil, the introduction of *An. gambiae*, a species with greater vectoral capacity than the local vectors, resulted in malaria epidemics that resulted in over 16,000 deaths (Parmakelis et al., 2008). Moreover, increased air travel has been implicated in the establishment of exotic mosquitoes on multiple Pacific islands, and container ships are known to have introduced numerous alien Diptera species, such as *Ae. albopictus,* into naïve regions in both the New and Old World (Effler et al., 2005; Pybus et al., 2015; Tatem et al., 2006).

Urbanisation is a key driver in land-use change, particularly in developing nations, in which approximately 90% of population growth occurs within cities. Indeed, it is estimated that 68% of the global population will live in an urban environment by 2050 (UN., 2019). As human, domestic, and peridomestic animal density is related to disease transmission potential and the emergence of novel disease, increasing urbanisation is likely to have significant epidemiological consequences (Hassell et al., 2017; Kolimenakis et al., 2021). For example, in China, areas of high human population density formed hubs of environmental suitability for dengue vectors, resulting in 90% of dengue cases being concentrated in urban villages (Yue et al., 2018).

Increasing urbanisation results in the increased abundance of Diptera vectors that have suitable environmental tolerances to benefit from anthropogenic environmental change and are able to thrive in urbanised systems (Townroe & Callaghan, 2014). Consequently, landscape modification has allowed vector species, such as *Ae. albopictus*, to greatly increase their global range with a subsequent increase in mosquito-borne disease transmission (Patz et al., 2004). Moreover, synanthropic species are abundant in urban habitats and may act as reservoirs for zoonotic pathogens. Therefore, human activity that results in an increasing abundance and contact with synanthropic species, and their associated Diptera vectors, are likely to increase the potential of disease spillover to humans or domesticated animals (Kolimenakis et al., 2021).

Urban expansion in developing nations is often unplanned and is typically associated with inadequate housing and minimal basic services, such as water and waste management (Gibb, Franklinos, et al., 2020; Wilson et al., 2020). Rapid expansion and inadequate piped water supplies may result in water being stored in a variety of containers. Disposal of waste close to human habitation, deteriorating infrastructure, and stored water provides breeding

habitat for a range of Dipteran vectors. This can reduce reliance upon seasonal precipitation cycles and result in increased interactions between disease vectors and humans (Gratz, 1999; Mattah et al., 2017).

Anthropogenic landscape modification, such as deforestation, for urban expansion and agricultural development has grown rapidly within the last century; in parallel with this growth is the growth of human-wildlife interactions via shared vectors, and subsequent exposure to novel pathogens (Patz et al., 2004). Deforestation has a broad range of effects, including the modification of predator and host densities (Chaves et al., 2020), and potentially impacting the community structure and abundance of Diptera vectors and altering their host-seeking behaviour (Ferraguti et al., 2016; Mayi et al., 2019; Orta-Pineda et al., 2021). Variation in host diversity and abundance due to landscape modification may have a significant impact on zoonotic pathogen prevalence and subsequent spillover risk through a variety of mechanisms, including the dilution effect, ecological and competitive release of vectors and hosts, and displacement of reservoir hosts (Keesing et al., 2010). Moreover, low diversity and hyper abundance of livestock host species in agricultural settings can impact vector community interaction structure and subsequent disease transmission potential if they act as intermediate or amplifier hosts (Franklinos et al., 2019; Jones et al., 2013). Spatial overlap between hosts and overlap in vector ranges are key factors in the transmission of vector-borne pathogens. Anthropogenically induced habitat fragmentation shifts existing landscapes into islands, with increased isolation, reduced area, and a greater ratio of habitat edges which act as an interface between humans, wildlife, and their associated vectors (Barbier, 2021). Consequently, increased human encroachment and modification of sylvatic habitat creates new opportunities for zoonotic pathogen exposure and enzootic disease to spill over into human populations (Gibb, Redding, et al., 2020).

Human-driven climate change can have significant implications for the transmission dynamics, geographic spread, emergence and re-emergence of vector-borne diseases through direct and indirect pathways (Rocklöv & Dubrow, 2020). Direct effects include the effect of climate variability on Diptera vectors, pathogens, and human and non-human hosts, whilst indirect effects include habitat alterations that may impact fitness of vectors or hosts (Rocklöv & Dubrow, 2020). Appropriate attribution of the direct effect of climate change on disease emergence and transmission dynamics may be challenging due to the regional

variability in climatic conditions, the confounding effects of concurrent stressors such as habitat alteration, and the fluctuations in population level susceptibility and dynamics of human behaviour (Baker et al., 2021). Long term climatic and epidemiological data, and robust statistical modelling are required to examine the extent of the association between climate change and disease emergence.

Life cycle traits of epidemiologically significant biting Diptera are affected by abiotic conditions, such as temperature and rainfall patterns (Fukui et al., 2022; Hamlet et al., 2018). As biting Diptera are ectotherms they display a unimodal response to temperature, though the thermal niches of different species may not overlap (Cunze et al., 2016). Increasing temperatures are expected to increase biting Diptera abundance, survival, feeding activity, and distribution. Moreover, vector-borne disease development rates and transmission cycles are climate sensitive, and climatic variations may result in changes to seasonal cycles and the frequency and scale of outbreaks (Baylis, 2017). Increasing temperatures in some regions can exceed a vector's thermal optima and limits, resulting in lower transmission risk; however, adaptation or invasion of thermally suitable competent vectors could minimise this decrease in transmission potential (Couper et al., 2021; Lim et al., 2021).

As the geographic distributions of many biting Diptera are limited by the abiotic conditions of their surroundings, increasing global temperatures and its associated effects will increase the suitability of novel habitat at higher altitudes and latitudes for invasion (Iwamura et al., 2020). Consequently, many vector-borne diseases are also expected to expand their range under climate change scenarios. For example, malaria transmission is predicted to increase due to a shift in spatial range to higher, newly suitable, latitudes and subsequent exposure to naïve populations, resulting in an additional 75 million people at risk of endemic malaria exposure (Ryan et al., 2020).

1.6 Existing and future control measures

Vector control initiatives are the principal method of controlling vector-borne disease transmission and the only means of population protection for a range of diseases, such as Zika virus and West Nile disease (Wilson et al., 2020). Whilst the first vector-borne disease transmission pathway to humans was identified by Sir Ronald Ross in 1897, who

demonstrated that *Anopheles* mosquitoes transmit malaria parasites (Ross, 1897), vector control had taken place prior to this due to the awareness of the connection between disease, swamps, and vectors. Indeed, notable concerted control efforts go back to at least the 1800s, though transmission mitigation methods, for example the use of bed nets or curtains, are recorded to have been used since at least 550 B.C (Wilson et al., 2020). Vector control methods primarily aim to reduce pathogen transmission through the reduction of human contact with the vector and may be broadly classified into chemical or non-chemical approaches (Chanda et al., 2013; Shroff et al., 2020; Wilson et al., 2020).

Chemical control methods can target vectors at the immature larval stage, using chemical larvicides which offer a low cost and effective control method, and are generally short-lived and have a low toxicity toward mammals and fish species (Shroff et al., 2020). Chemical methods may also target adult vectors using a range of approaches, such as long-lasting insecticide-treated bed nets (LLINs), indoor residual spraying (IRS) of insecticides, and topical repellents (Koenraadt et al., 2021). Whilst chemical control measures play a vital role in current prevention initiatives, the limited uptake of IRS, misuse of LLINs, and rising resistance to insecticides (Chanda et al., 2013; Koenraadt et al., 2021; Shroff et al., 2020) may limit the efficacy of future chemical-based control methods.

Non-chemical approaches utilise environmental and biological control methods. Larval source management is the elimination of the larval habitat and breeding sites, though it relies upon a comprehensive understanding of the interactions between the vector and their environment (McCann et al., 2017). Moreover, the efficacy of this approach is limited in natural habitats such as swamps and in agriculturally developed land containing irrigation ditches (Patz et al., 2004; Tusting et al., 2013). Whilst the introduction of biocontrols, such as microbial larvicides and larvivorous fish (MartíNez-Ibarra et al., 2002) have been shown to be efficacious, they are heavily context dependent and many biting Diptera larval habitats cannot support their presence (Benelli et al., 2016). Improvement of housing, through the addition of window screens and reduction of entry points, has been linked to a reduction in the number of indoor dwelling vectors and subsequent transmission cases (McCann et al., 2017). However, achieving a high community coverage may be challenging, particularly in developing nations, which are most at risk of disease transmission, without the inclusion of

behaviour change communication strategies (McCann et al., 2017; Van Den Berg et al., 2007).

To improve the efficacy of control measures, adopting an integrated Vector Management (VM) approach is recommended, which avoids the limitations of previous control initiatives which are broadly inflexible to change in circumstance and rely on a single form of control, such as IRS, overlooking the importance of complementary interventions (Benelli & Beier, 2017). However, like traditional approaches, IVM should be guided by the ecology of their biting Diptera targets and an understanding of how environmental factors affect their community structure, distribution and abundances (Beier et al., 2008; Chanda et al., 2013; Golding et al., 2015).

Notable developments of biological control methods targeting mosquito populations include Sterile Insect Technique (SIT), *Wolbachia* endosymbiotic bacteria, and genetically modified self-sustaining transgenic mosquitos (Benelli et al., 2016). SIT is a top down approach that involves raising and sterilizing a large number of males of the target species, which are then released into the wild to mate with wild females and produce no progeny (Alphey et al., 2010; Benelli et al., 2016). A top-down SIT control approach relies on a centralised production of sterilized males and release over a broad geographical area to reduce the target population. In comparison, a bottom-up control approach is decentralised and occurs on a smaller scale, often at the community level, to control localised populations (Alphey et al., 2010; Benelli et al., 2016). The SIT approach has proven successful against a range of agricultural pests; however, the SIT male mosquitoes have a lower fitness compared to the wild type, reducing the efficacy of the control method. The maternally inherited *Wolbachia* bacterium modifies mosquito life history, reducing their susceptibility to infection, shortening their life span, and consequently limiting their vectoral capacity. The release of *Ae. aegypti* mosquitoes infected with *Wolbachia* to combat Dengue transmission may prove to be a promising control strategy (Iturbe-Ormaetxe et al., 2011; Wilson et al., 2020). Self-sustaining transgenic mosquitoes are produced via gene drive, a method of genetic modification to spread traits through biasing inheritance (Kyrou et al., 2018). This technique may be used to suppress populations, through sterilising or reducing fertility of females while leaving males unaffected to continue to act as carriers, or by biasing the sex ratio towards males (Pollegioni et al., 2020). However, as mosquitoes from self-sustaining

strategies continue to propagate once released, they cannot be retrieved if subsequent unintended ecological consequences occur. Consequently, it is imperative to elucidate the community ecology and interactions of the target species, prior to release.

1.7 Ecological Interaction networks as a means of examining biting Diptera-host interactions

As hematophagous Diptera are embedded within complex ecological communities in which they may act as predator, prey, and competitor, their interactions can have an important role in the population dynamics of co-occurring species (Ferguson et al., 2010). Changes to biting Diptera community structure, for example as a result of environmental modification or a targeted control initiative, could trigger an ecological cascade that competitively releases and enhances the transmission potential of another vector species in the community (Gillies & Furlong, 1964; Gillies & Smith, 1960). However, there is currently a limited understanding of the degree to which environmental modification and control initiatives impact biting Diptera-host interactions, resource utilisation, and community structure (Ferguson et al., 2010). Indeed, as Diptera community assemblages are rarely studied, the full ecological and epidemiological implications of their interactions are currently unknown. Ultimately, the successful mitigation of vector-borne disease burden, elucidation of the effect of anthropogenic landscape modification and climate, and the examination of transmission systems, cannot occur without a comprehensive understanding of the ecology, community structure, abundances, and interactions of the wider biting Diptera community.

Ecological network modelling offers a holistic approach to describe the interdependencies between sets of interacting species (such as biting Diptera and their hosts), the underlying structure of the communities, and predict the dynamics of the component species (Evans et al., 2013; Fath et al., 2007; Proulx et al., 2005). At its core, ecological interaction networks are visual and mathematical representations of a set of discrete objects (nodes), often comprised of individuals that make up a species population (Ings et al., 2009), though they may represent individuals (Godfrey, 2013; Melián et al., 2011), and their relationships, represented by a series of edges (often described as links) (Blüthgen, 2010). Networks in ecological research are often depicted as two interacting trophic levels (bipartite networks)

with nodes and links that are proportional to a species abundance, and interaction frequency respectively (Dormann et al., 2009).

Interactions can be classified as unweighted, in which the edges are represented by a binary value (1 or 0 representing interaction presence or absence) or weighted, in which the frequency or abundance of an interaction represents its strength (weight) (Landi et al., 2018; Miranda et al., 2019). In unweighted networks, as interaction strength is not represented, all interactions are considered ecologically equivalent, whilst weighted interactions may take into account a variety of factors, such as resource selection (i.e. host preference) and abundances (Blüthgen, 2010; Miranda et al., 2019). Consequently, weighted networks can be considered more biologically 'realistic' as the intensity of interactions is incorporated (Moreira et al., 2018).

The relations between discrete objects can be classed as symmetric (undirected), in which species affect each other equally, or asymmetric (directed), in which a species may affect another with differing amounts of reciprocation (Landi et al., 2018; Poulin, 2010). Edges visualise a wide range of ecological relationships and can be broadly separated into trophic antagonistic interactions, including predator-prey and primary consumer-basal resource feeding and parasitism (Ings et al., 2009); mutualistic interactions, including pollination and seed dispersal (Bendel et al., 2019); competitive interactions, interference for a common resource (Buss & Jackson, 1979); and commensalism, a species gaining a positive effect without an adverse effect for the host (Mougi, 2016). Ecological networks can also be particularly relevant in an epidemiological context as edges may represent pathways of transmission between individuals, thus allowing for the identification of the component nodes of disease transmission cycles, such as reservoir species and bridge vectors. (Godfrey, 2013). For example, a similar approach has previously been applied in contact modelling to examine human-human interactions that may represent the spread of Covid-19 (Maheshwari & Albert, 2020).

1.7.1 Network metrics

The analysis of an ecological network is inherently hierarchical and ranges from examination of the entire network down to individual nodes and edges (Lau et al., 2017). To examine the structural properties of a network, a large number of quantitative (weighted) and qualitative (unweighted) metrics (sometimes referred to as indices) have been developed, that are used to characterise the system at each level: whole network, sub-network, and individual (Blüthgen et al., 2008; Dormann et al., 2009; Evans et al., 2016). Network level metrics integrate information from the entire set of interacting nodes and edges, such metrics include H2' network specialisation (Blüthgen et al., 2006), connectance (Jordano, 1987), and nestedness (Patterson & Atmar, 1986). Sub-network level analysis focuses on the identification of specific subsets of node and their associated edges, such as modules and motifs; a subgraph that represents the patterns of interactions between a small number of species (Simmons et al., 2019). Node level metrics, such as d' specialisation (Blüthgen et al., 2006), degree (Dáttilo & V. Rico-Gray, 2018), and between and closeness centrality (Martín González et al., 2010), summarise similar organisational features as network level metrics, and may quantify the differences in the relative importance of a node or edge (Niquil et al., 2012). The degree of resolution that can be examined is highly dependent on the characteristics and level of detail provided to the given model (Lau et al., 2017).

1.7.2 Metrics of interest for biting Diptera network analysis

Examples of network metrics that are of particular ecological and epidemiological relevance include H2' network specialisation, centrality, and interaction evenness.

H2' (Figure 2.) is a network-level metric that quantifies the degree of specialisation of a network. H2' is mathematically based on the deviation of a species realised number of interactions and those expected for each species' marginal interaction totals (the sum of expected interactions with other nodes in the network) (Dormann et al., 2009; Morris et al., 2014). In an ecological context, H2' can be used for comparisons across different interaction networks, for example, the variation in network specialisation between differing levels of habitat modification. The resulting H2' values range between 0 (no specialisation) and 1

(perfect specialisation, for the given interaction totals) (Dormann et al., 2009), and is robust to variations in network dimensions and sampling intensity (Blüthgen et al., 2006).

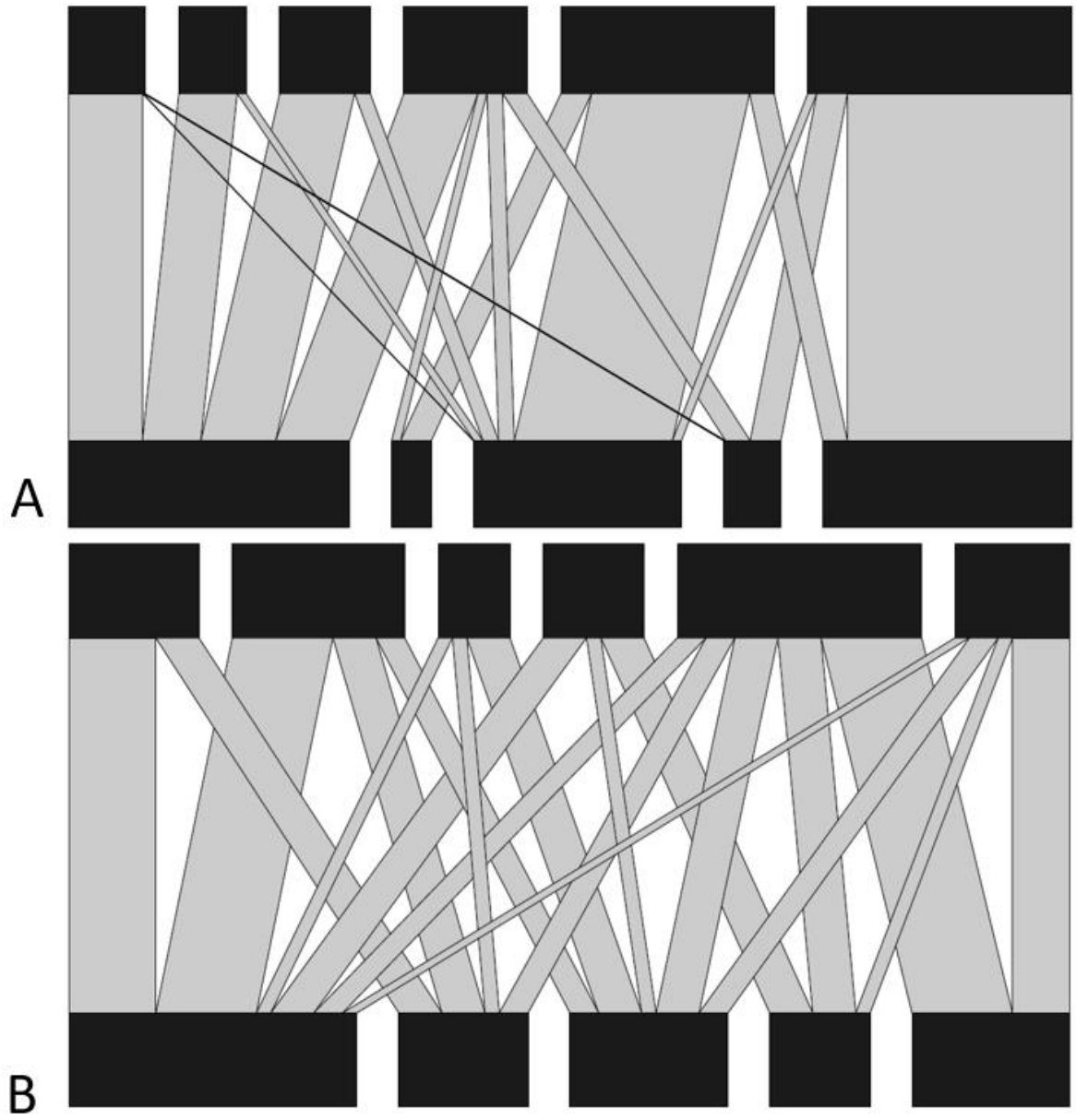


Figure 2. Example bipartite networks with a high value of H2' network specialisation (A) (H2'=0.68) and a low H2' network specialisation (B) (H2'=0.29)

The dynamics and stability of a complex network is highly dependent on its topological structure; however, not all nodes are equally important (Martín González et al., 2010). Weighted centrality metrics, degree, closeness, and betweenness quantify the relative topological importance of a node, with a greater importance resulting in a faster collapse of network structure when they are selectively removed in comparison to random removal (Freeman, 1978). Degree centrality calculates the number of edges the target node has in

the network. Closeness centrality describes the proximity of a node to all other nodes in the network and is quantified as the average shortest distance from a target node to every other node in the network (Swain et al., 2022). Betweenness centrality describes the importance of a node as a connector between parts of the network, and is quantified as the number of shortest paths between two nodes that pass through the target node (Freeman, 1978; Martín González et al., 2010). High closeness values indicate that a node may rapidly affect, or be affected by, other nodes in the network, whilst high betweenness values indicate the node is important for the cohesiveness of the network (Freeman, 1978; Swain et al., 2022). Epidemiological uses of centrality metrics include the examination of a vector or host's centrality to elucidate its relative importance for the maintenance and propagation of a disease through the system (Hoyos et al., 2021).

Interaction evenness (Figure 3.) is a network level metric based on Shannon diversity that describes the homogeneity in the distribution of the interactions frequencies across all nodes in a network, with greater values reflecting a more uniform spread of interactions among nodes (Kaiser-Bunbury & Blüthgen, 2015). Interaction evenness can indicate the degree to which a community is dominated by a small set of species and their associated links, with epidemiological relevance if the dominant species are involved in the maintenance and transmission of vectored disease.

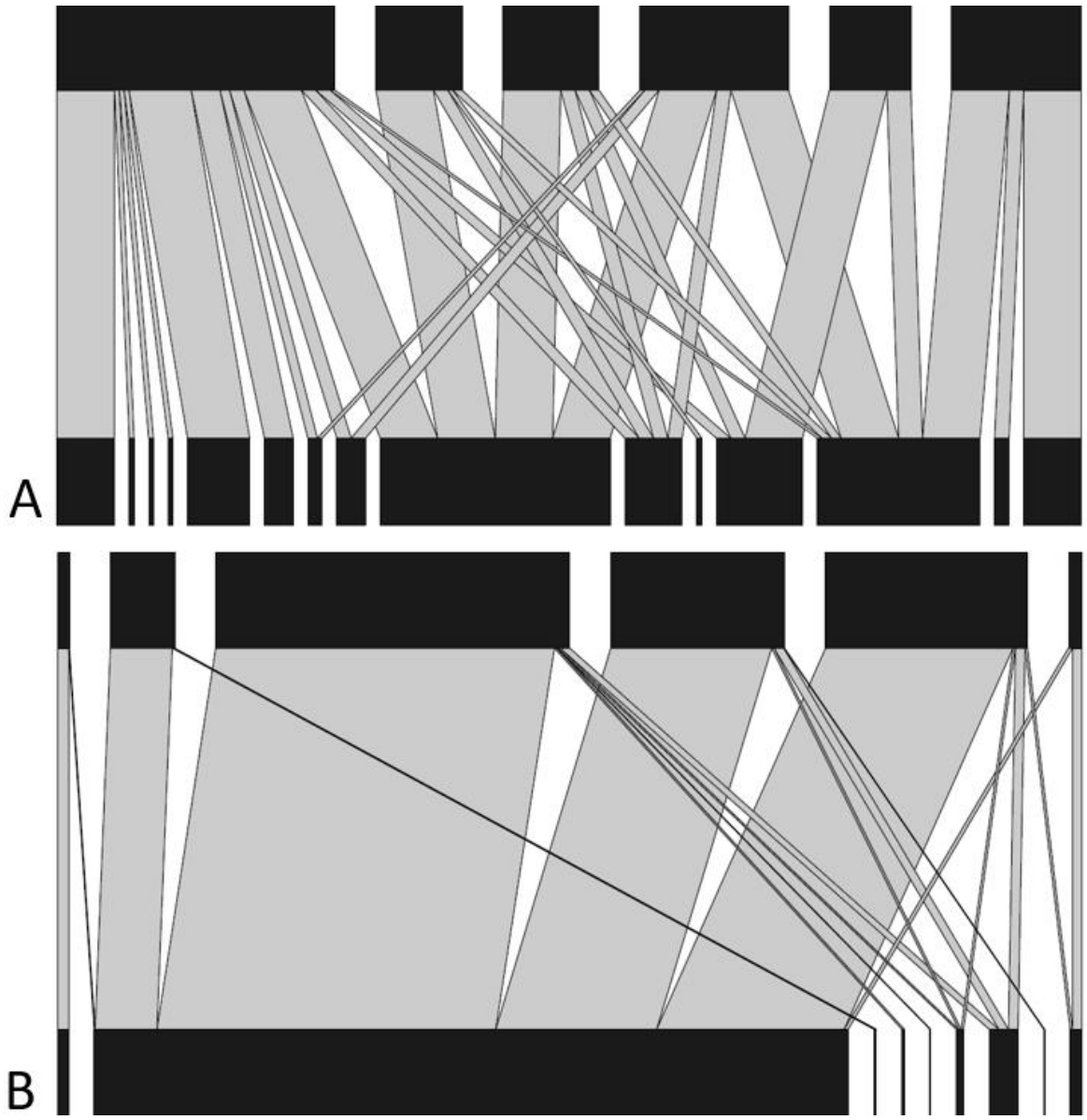


Figure 3. Example bipartite network with a high interaction evenness (A) (IE=0.7) and a low interaction evenness (B) (IE=0.39)

The selection of appropriate metrics should be question and context driven, with a clear rationale for selection. The calculation of a broad range of metrics, without the appreciation of their mathematical and ecological underpinning, may result in the calculation of metrics that report similar outputs, resulting in redundant analysis and the potential for multicollinearity between selected metrics (Morrison et al., 2022). Further, sampling completeness should be considered during metric selection, to maximise metric robustness to variations in network size and minimise the effects of sampling bias (Fründ et al., 2016).

1.8 Molecular detection of biting Diptera-host interactions

1.8.1 Previous methods of identifying interactions

To draw accurate conclusions from ecological interaction networks, a network should be populated with component interaction data that are as complete and reliable as possible (Cuff et al., 2022; Pompanon et al., 2012). In the past it has been difficult to characterise the interaction structure of complex communities with high species diversity, due to the logistical, sampling, and technical constraints (Evans et al., 2016). Interaction networks were constructed using methods that primarily relied upon field observations and morphological identification by taxonomic specialists, for example, antagonistic interactions generated through the direct observation of feeding events or the morphological analysis of gut (Amundsen & Sánchez-Hernández, 2019; Reñones et al., 2002) or faecal contents (Caryl et al., 2012; Sousa et al., 2019). However, these approaches are labour-intensive and the reliability of interactions may be limited as they can be difficult to observe, due to the individual species' ecology, subject to potential observer bias, and require considerable taxonomic expertise and the presence of sufficiently undigested prey items (Evans et al., 2016; Sheppard & Harwood, 2005; Sousa et al., 2019).

Indeed, the limitations of these methods may be particularly prevalent in the analysis of biting Diptera interactions, as the origin of a blood meal cannot be identified through visual examination and Diptera-host interactions are difficult to observe, due to the limited size of biting Diptera and the brief feeding period (Pompanon et al., 2012). Further, reliance on biting Diptera interaction data inferred through sampling approaches such as human landing catches (HLC) and animal landing catches will result in uneven and inherently biased networks. Molecular detection of species interactions has been increasingly utilised in the past decades, with particular emphasis on the identification of mixed samples, such as gut contents (Clare et al., 2019) . Moreover, the integration of molecular based methods can overcome the issues raised by morphological methods (Cuff et al., 2022), with continued advances in molecular methods providing increasingly reliable and complex interaction data.

Previous attempts to capture the dietary breadth required for the examination of network dynamics used a wide range of analytical methods, including, protein electrophoretic (Walrant & Loreau, 1995; Walter & O'Neill, 1986), stable isotope (Gratton et al., 2008;

Ostrom et al., 1996), and serological based methods (Deagle et al., 2005; Harper et al., 2006). Protein electrophoretic approaches enable the identification of gut contents in steno and monophagous predators, though are broadly unsuitable for generalist predators, in which multiple prey species can be simultaneously present in the gut, as it results in uninterpretable overlapped protein bands (Pompanon et al., 2012). Isotopic signatures (characteristic stable isotopes present in a predator's tissue derived from consumed food) provide information on the long-term diets of predators, rather than identifications from a recent meal. However, they are often not consistent or strong enough to produce a high resolution insight into trophic dynamics and are subject to a large amount of variability (Lecomte et al., 2011). Serological methods, including enzyme-linked immunosorbent assays (ELISAs) (Jansen et al., 2009) and precipitin tests (Logue et al., 2016) require significant time investment to produce host serological libraries and an *a priori* assumption of blood meal composition, often resulting in a library of limited taxonomic resolution. To overcome the limitations of morphological and early analytical approaches, and following the dramatic decrease in sequencing costs, the detection of species interactions is now broadly conducted using molecular DNA based detection approaches.

1.8.2 Current methods of identifying interactions

A commonly employed DNA based technique, DNA barcoding, utilises general or targeted primers, and Sanger sequencing for the identification of a single gene per sample (Borland & Kading, 2021; Clare et al., 2019). The advent of High-throughput sequencing (HTS), commonly referred to as Next Generation Sequencing (NGS), offers the opportunity to generate millions of sequences per sample and the parallel processing and identification of a wide number of DNA amplicons at the same time (Pompanon et al., 2012). HTS, when used in tandem with DNA barcoding and group-specific or general primers (Piñol et al., 2014), has rapidly become the standard for accurate identification of mixed samples, including bulk samples and environmental DNA (eDNA) (Zinger et al., 2019) in a process known as metabarcoding (Batovska et al., 2018; Cuff et al., 2022; Pompanon et al., 2012). Metabarcoding has allowed for previously unobtainable resolution of interactions, in particular where prey items were visually unidentifiable (Pompanon et al., 2012). Metabarcoding has been used in a range of contexts, including examining the diet of

vertebrates, such as bats (Hemprich-Bennett, Kemp, et al., 2021), pigs (Robeson et al., 2018), and birds (Cabodevilla et al., 2021), and invertebrates, such as phlebotomine sand flies (Kocher et al., 2017) and mosquitoes (Batovska et al., 2018). Moreover, metabarcoding may be used to identify bloodmeals in sanguivores, including leeches (Drinkwater et al., 2019, 2020; Schnell et al., 2012) and a range of biting Diptera (Gaithuma et al., 2020; Hopken et al., 2021; Reeves et al., 2018).

Regardless of the starting material (bulk samples, eDNA, blood meals etc) all metabarcoding approaches utilise broadly similar experimental steps. These steps are comprised of sampling and field or laboratory preservation of material, DNA extraction, Polymerase Chain Reaction (PCR) amplification of a selected taxonomically informative genomic region (e.g. CO1 or 16S loci), sequencing of amplicons using high throughput sequencing, and generated sequence analysis using a variety of bioinformatics methods and pipelines (Zinger et al., 2019). Molecular identification of pathogen presence in biting Diptera blood meals typically follows similar experimental steps, with the extraction of RNA and the addition of a reverse transcription step to convert arboviral RNA into cDNA prior to PCR. Pathogen detection typically involves removal of the head and thorax for quantification of viral titre, though blood meals may also be used for screened for pathogens (Chouin-Carneiro Id et al., 2020). Despite the apparent methodological uniformity and simplicity, each experimental step can introduce unintended artefacts and potential sources of bias (Lamb et al., 2019; Taberlet et al., 2018). Sources of bias include variation in collection and storage methods, which may enrich the sample with specific taxa depending on how samples are collected. For example, variation in efficacy and attractiveness of biting Diptera trap type (Busula et al., 2015; Hesson et al., 2015), and the use of live-animal and human landing sampling which inherently produces an abundance of known blood meals. Sample preservation, the degree to which the sample DNA is degraded, and extraction methods can also introduce their own source of biases (Reeves et al., 2016). Variation in PCR amplification efficiency is a well-known source of bias that is particularly prevalent with the increasing use of HTS techniques. PCR primer bias can result in the preferential amplification of certain taxa over others (Deagle et al., 2014), and produce false negatives and positives, through the presence of PCR inhibitors and PCR errors (replication errors and chimeric fragments) respectively (Zinger et al., 2019). Tag and index jumping introduce non-negligible amounts of contamination and

may disproportionately impact samples with low DNA concentration (Esling et al., 2015). Moreover, HTS instruments produce their own error rates making it difficult to distinguish between real biological nucleotide differences and sequencing artefacts (Schirmer et al., 2015).

1.8.3 Bioinformatic approaches

As it is frequently not possible to obtain taxonomic information from all generated sequences, the analysis of metabarcoding sequence data has predominantly been conducted using molecular operational taxonomic units (OTU), in which reads are assigned to a cluster based on an arbitrary predefined clustering dissimilarly threshold, typically 97% (Clare et al., 2016; Gordon et al., 2019; Westcott & Schloss, 2015). A variety of algorithms and software packages have developed to delimit OTUs (Liu et al., 2021), including MOTHUR (Schloss et al., 2009), Qiime (Uclust), and the automatic allocation system employed by the Barcode of Life Database (BOLD) to assign sequences to Barcode Index Numbers (BINs) (Boers et al., 2019; Ratnasingham & Hebert, 2007). Clustering has the benefit of reducing the computational requirements for downstream analysis and limiting the potential impact of sequencing errors, as erroneous sequences are merged with correct sequences (Chiarello et al., 2022). The predominant distinction between methods of defining molecular OTUs is between closed-reference and *de novo* OTUs. In closed-reference OTUs sequencing reads are assigned to an OTU if they are similar to an associated reference library sequence, and if the same reference library is maintained, multiple sets of independently processed data can be compared (Callahan et al., 2017). In contrast, *de novo* OTUs are grouped as a function of the pairwise sequence similarities; , consequently, *de novo* clusters are features of a particular data set and may cannot be compared across datasets (Callahan et al., 2017).

Variation in the sequence divergence threshold assigned during molecular OTU delimitation may significantly alter diversity and network architecture. Large thresholds risk failing to exclude molecular artefacts and incorrectly grouping distinct taxa into a single OTU, whilst smaller thresholds risk artificially increasing the number of OTUs through errors (Clare et al., 2016; Hemprich-Bennett, Oliveira, et al., 2021). Moreover, the divergence threshold may significantly impact network topology when molecular OTU data is incorporated into ecological interaction networks. Indeed, minor variations in the threshold and subsequent

node resolution have been shown to significantly alter network topography and a wide range of key network metrics, highlighting the need for caution when utilising molecular data derived from sequence divergence clustering (Hemprich-Bennett, Oliveira, et al., 2021).

To limit issues associated with OTU clustering there has been a recent methodological shift towards denoising methods, which resolve Amplicon Sequence Variants (ASVs) from amplicon data without imposing an arbitrary predefined clustering dissimilarly threshold (Chiarello et al., 2022). ASV methods, using packages such as DADA2 (Callahan et al., 2016), utilise a unique error model that is generated based on the quality of the sequencing run to differentiate between the true biological variations and those generated through sequencing errors, whilst also accounting for sequence abundance and similarities (Callahan et al., 2016; Joos et al., 2020; Nearing et al., 2018). The resultant sequences are then collapsed and defined as individual ASVs (Chiarello et al., 2022). Unlike OTUs, ASVs may be resolved down to the level of single nucleotide differences, which can greatly increase the taxonomic resolution of the results (Chiarello et al., 2022; Joos et al., 2020). Moreover, as ASVs represent the DNA sequence of the organism, ASVs generated independently, though targeting the same gene loci, from different samples or studies can be compared (Callahan et al., 2017). Whilst the development of ASVs has become increasingly utilised, the inference of ASVs does not solve all problems inherent with sequence analysis. ASV analysis using datasets with insufficient size or quality may result in real biological variations being classified as errors, there is no guarantee of an ecological coherence between genomes with the same ASV (Callahan et al., 2017), and a greater sensitivity for detecting species present can sometimes come at the expense of specificity (Chiarello et al., 2022).

The challenges associated with the sources of experimental bias and differences in bioinformatical approaches are particularly relevant when integrating metabarcoding data into interaction networks (Cuff et al., 2022; Quintero et al., 2022). Whilst the issues associated with metabarcoding do not limit its utility in network analysis, careful consideration of each stage of the experimental approach, including assessment of the benefits and weaknesses of the range of bioinformatical approaches, is required to produce reliable and accurate interaction data with the greatest possible resolution and allow robust conclusions to be drawn.

1.9 Overall aims

In this thesis, I investigate how biting Diptera-host community interactions are structured, and how these interactions vary at a range of spatial scales and under different environmental contexts. Most previous biting Diptera research focuses on individual Diptera species or small sets of interacting species, and broadly overlooks the communities in which they are embedded. My thesis aims to fill current knowledge gaps by establishing the range of potential applications of ecological network approaches and molecular blood-meal metabarcoding in biting Diptera research, demonstrating the effectiveness of this approach using interaction data drawn from the literature as well as new empirical data collected at two field sites in Ghana.

My thesis has four primary aims:

1. To highlight the applicability of ecological network analysis in the context of biting Diptera-host interactions derived from molecular analysis of host DNA contained in Diptera blood meals.
2. To quantify the effects of macroecological and anthropogenic factors on the interaction structure of biting Diptera-host networks.
3. To inform practical approaches to blood meal analysis by examining the effect of digestion time and storage conditions on the PCR amplification success of host DNA contained in biting midge (*Culicoides*) blood meals.
4. To examine and describe the differences in the community structure and composition of networks of biting Diptera-host interactions across differing levels of anthropogenic landscape modification.

1.10 Thesis structure

This body of work is presented as an integrated thesis comprising a General Introduction (Chapter 1) followed by four data chapters (listed below) which are formatted as independent research articles and include an Abstract, Introduction, Methods, Results, and Discussion. Each chapter is preceded by a status of publication statement and, where applicable, an Author contribution statement. The thesis concludes with Chapter 6, which contains a general discussion of the thesis and focuses on discussion of key patterns and findings that have emerged from the previous chapters, alongside thesis limitations and recommendations for future research. In addition, research papers that were completed external to this thesis, but are derived from data generated during its creation, are included as appendices (Appendix 1.), alongside protocols for research that was planned and partially completed during work on the thesis (Appendix 2.).

Data chapters:

- In Chapter 2, I use interaction data extracted from the literature to discuss the value of a network perspective for understanding biting Diptera-host interactions. I identify applications of network approaches in vector research, including highlighting pathways of disease transmission, reservoirs of infection, and emerging and previously unrecognised vectors. Finally, I identify a range of outstanding questions in the field.
- In Chapter 3, I explore the effect of macroecological and local anthropogenic factors on the properties of networks linking biting Diptera and their vertebrate hosts using a dataset of biting-host interactions extracted from the literature.
- In Chapter 4, I use a laboratory experiment to investigate the effect of digestion time and the impact of two storage conditions (-20°C freezer and ambient room temperature) on the PCR amplification success of host DNA contained in two species of biting midges.
- In Chapter 5, I use empirical data obtained from sampling at two field sites in rural Ghana to investigate the effect of differing levels of anthropogenic landscape modification on biting Diptera community composition and Diptera-host interaction structure.

1.11 Covid-19 impact statement

The output of this thesis was impacted by the COVID-19 pandemic as access to the Ghanian field sites was heavily interrupted during the period of 2020-2022 because of lockdown and continuing travel restrictions. I conducted field sampling in 2019 as part of a pilot study, prior to a larger sampling effort planned for the summer of 2020. As the field sites were no longer accessible during this period, the work presented in Chapter 4 is based on samples collected during the pilot season and the continued sampling by members of the Target Malaria research consortium, which took place until February 2020. Increased sampling at the field sites would likely have resulted in a greater network resolution and confidence in my village-based networks. Further, due to COVID-19 related closures, and then the temporary repurposing of the sequencing facility to COVID-19 testing, the sequencing of all biting Diptera and identification of blood meal host origin was severely delayed, and sequence data were not received until September 2022.

In addition, I designed a chapter that aimed to examine the variation in biting Diptera community composition and interaction structure with increasing distance to humans. Due to COVID-19 I was unable to conduct the field sampling for this project, and consequently managed a Ghanian based field team, which included the cooperation of Ghanian park rangers, during the 2022 field season. As this sampling effort was designed to take place in 2020, the significant delays in sampling have resulted in the exclusion of the work from the main body of this thesis. However, all samples required to conduct this research have been collected and are currently in freezer storage in Oxford; the project will be completed external to this thesis. The field sampling protocols, which I designed, are included in Appendix 2.

1.12 References

Alkhamis, M. A., Aguilar-Vega, C., Fountain-Jones, N. M., Lin, K., Perez, A. M., & Sánchez-Vizcaíno, J. M. (2020). Global emergence and evolutionary dynamics of bluetongue virus. *Scientific Reports 2020 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-78673-9

Alphey, L., Benedict, M., Bellini, R., Clark, G. G., Dame, D. A., Service, M. W., & Dobson, S. L. (2010). Sterile-Insect Methods for Control of Mosquito-Borne Diseases: An Analysis. *Vector Borne and Zoonotic Diseases*, *10*(3), 295. https://doi.org/10.1089/VBZ.2009.0014

Althouse, B. M., Guerbois, M., Cummings, D. A. T., Diop, O. M., Faye, O., Faye, A., Diallo, D., Sadio, B. D., Sow, A., Faye, O., Sall, A. A., Diallo, M., Benefit, B., Simons, E., Watts, D. M., Weaver, S. C., & Hanley, K. A. (2018). Role of monkeys in the sylvatic cycle of chikungunya virus in Senegal. *Nature Communications 2018 9:1*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-03332-7

Álvarez-Hernández, D. A., Rivero-Zambrano, L., Martínez-Juárez, L. A., & García-Rodríguez-Arana, R. (2020). Overcoming the global burden of neglected tropical diseases. *Therapeutic Advances in Infectious Disease*, *7*. https://doi.org/10.1177/2049936120966449

Amundsen, P. A., & Sánchez-Hernández, J. (2019). Feeding studies take guts – critical review and recommendations of methods for stomach contents analysis in fish. *Journal of Fish Biology*, *95*(6), 1364–1373. https://doi.org/10.1111/JFB.14151

Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L. F., Wesolowski, A., & Metcalf, C. J. E. (2021). Infectious disease in an era of global change. *Nature Reviews Microbiology 2021 20:4*, *20*(4), 193–205. https://doi.org/10.1038/s41579-021-00639-z

Barber, L. M., Schleier, J. J., & Peterson, R. K. D. (2010). Economic Cost Analysis of West Nile Virus Outbreak, Sacramento County, California, USA, 2005. *Emerging Infectious Diseases*, *16*(3), 480. https://doi.org/10.3201/EID1603.090667

Barbier, E. B. (2021). Habitat loss and the risk of disease outbreak. *Journal of Environmental Economics and Management*, *108*, 102451. https://doi.org/10.1016/J.JEEM.2021.102451

Batovska, J., Lynch, S. E., Cogan, N. O. I., Brown, K., Darbro, J. M., Kho, E. A., & Blacket, M. J. (2018). Effective mosquito and arbovirus surveillance using metabarcoding. *Molecular Ecology Resources*, *18*(1), 32–40. https://doi.org/10.1111/1755-0998.12682

Baylis, M. (2017). Potential impact of climate change on emerging vector-borne and other infections in the UK. In *Environmental Health: A Global Access Science Source* (Vol. 16). BioMed Central Ltd. https://doi.org/10.1186/s12940-017-0326-1

Beier, J. C., Keating, J., Githure, J. I., MacDonald, M. B., Impoinvil, D. E., & Novak, R. J. (2008). Integrated vector management for malaria control. *Malaria Journal*, *7*(Suppl 1), S4. https://doi.org/10.1186/1475-2875-7-S1-S4

Bendel, C. R., Kral-O'Brien, K. C., Hovick, T. J., Limb, R. F., & Harmon, J. P. (2019). Plant–pollinator networks in grassland working landscapes reveal seasonal shifts in network structure and composition. *Ecosphere*, *10*(1), e02569. https://doi.org/10.1002/ecs2.2569

Benelli, G., & Beier, J. C. (2017). Current vector control challenges in the fight against malaria. *Acta Tropica*, *174*, 91–96. https://doi.org/10.1016/J.ACTATROPICA.2017.06.028

Benelli, G., Jeffries, C. L., & Walker, T. (2016). Biological Control of Mosquito Vectors: Past, Present, and Future. *Insects*, *7*(4). https://doi.org/10.3390/INSECTS7040052

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K. E., Moyes, C. L., Henry, A., Eckhoff, P. A., Wenger, E. A., Briët, O., Penny, M. A., Smith, T. A., Bennett, A., Yukich, J., Eisele, T. P., Griffin, J. T., Fergus, C. A., … Gething, P. W. (2015). The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature 2015 526:7572*, *526*(7572), 207–211. https://doi.org/10.1038/nature15535

Biteye, B., Fall, A. G., Seck, M. T., Ciss, M., Diop, M., & Gimonneau, G. (2019). Host-feeding patterns of Aedes (Aedimorphus) vexans arabiensis, a Rift Valley Fever virus vector in the Ferlo pastoral ecosystem of Senegal. *PLOS ONE*, *14*(10), e0215194. https://doi.org/10.1371/JOURNAL.PONE.0215194

Blüthgen, N. (2010). Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic and Applied Ecology*, *11*(3), 185–195. https://doi.org/10.1016/J.BAAE.2010.01.001

Blüthgen, N., Fründ, J., Vázquez, D. P., & Menzel, F. (2008). WHAT DO INTERACTION NETWORK METRICS TELL US ABOUT SPECIALIZATION AND BIOLOGICAL TRAITS. *Ecology*, *89*(12), 3387–3399. https://doi.org/10.1890/07-2121.1

Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecology*, *6*(1), 9. https://doi.org/10.1186/1472-6785-6-9

Boers, S. A., Jansen, R., & Hays, J. P. (2019). Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. In *European Journal of Clinical Microbiology and Infectious Diseases* (Vol. 38, Issue 6, pp. 1059–1070). Springer Verlag. https://doi.org/10.1007/s10096-019-03520-3

Bonds, M. H., Keenan, D. C., Rohani, P., & Sachs, J. D. (2010). Poverty trap formed by the ecology of infectious diseases. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1685), 1185–1192. https://doi.org/10.1098/RSPB.2009.1778

Borland, E. M., & Kading, R. C. (2021). Modernizing the Toolkit for Arthropod Bloodmeal Identification. *Insects*, *12*(1), 37. https://doi.org/10.3390/insects12010037

Breman, J., Egan, A., & Keusch, G. (2001). The intolerable burden of malaria: a new look at the numbers. *The American Journal of Tropical Medicine and Hygiene*, *64*(1_suppl), iv–iv. https://doi.org/10.4269/AJTMH.2001.64.IV

Buss, L. W., & Jackson, J. B. C. (1979). Competitive Networks: Nontransitive Competitive

Relationships in Cryptic Coral Reef Environments. In *The American Naturalist* (Vol. 113, pp. 223–234). The University of Chicago PressThe American Society of Naturalists. https://doi.org/10.2307/2460201

Busula, A. O., Takken, W., Loy, D. E., Hahn, B. H., Mukabana, W. R., & Verhulst, N. O. (2015). Mosquito host preferences affect their response to synthetic and natural odour blends. *Malaria Journal*, *14*(1), 133. https://doi.org/10.1186/s12936-015-0635-1

Cabodevilla, X., Mougeot, F., Bota, G., Mañosa, S., Cuscó, F., Martínez-García, J., Arroyo, B., & Madeira, M. J. (2021). Metabarcoding insights into the diet and trophic diversity of six declining farmland birds. *Scientific Reports 2021 11:1*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-00519-9

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal 2017 11:12*, *11*(12), 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581. https://doi.org/10.1038/NMETH.3869

Campbell-Lendrum, D., Manga, L., Bagayoko, M., & Sommerfeld, J. (2015). Climate change and vector-borne diseases: what are the implications for public health research and policy? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1665), 1–8. https://doi.org/10.1098/RSTB.2013.0552

Caryl, F. M., Raynor, R., Quine, C. P., & Park, K. J. (2012). The seasonal diet of British pine marten determined from genetically identified scats. *Journal of Zoology*, *288*(4), 252–259. https://doi.org/10.1111/J.1469-7998.2012.00951.X

CDC. (2021). *CDC - Malaria - Malaria Worldwide - Impact of Malaria*. https://www.cdc.gov/malaria/malaria_worldwide/impact.html

Chanda, E., Govere, J. M., Macdonald, M. B., Lako, R. L., Haque, U., Baba, S. P., & Mnzava, A. (2013). Integrated vector management: A critical strategy for combating vector-borne diseases in South Sudan. *Malaria Journal*, *12*(1), 1–9. https://doi.org/10.1186/1475-2875-12-369/TABLES/3

Chaves, L. S. M., Fry, J., Malik, A., Geschke, A., Sallum, M. A. M., & Lenzen, M. (2020). Global consumption and international trade in deforestation-associated commodities could influence malaria risk. *Nature Communications 2020 11:1*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-14954-1

Chiarello, M., McCauley, M., Villéger, S., & Jackson, C. R. (2022). Ranking the biases: The choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. *PLOS ONE*, *17*(2), e0264443. https://doi.org/10.1371/JOURNAL.PONE.0264443

Chouin-Carneiro Id, T., David, M. R., De Bruycker Nogueira, F., Barreto Dos Santosid, F., & Lourenç O-De-Oliveira, R. (2020). *Zika virus transmission by Brazilian Aedes aegypti and Aedes albopictus is virus dose and temperature-dependent*. https://doi.org/10.1371/journal.pntd.0008527

Chua, P. L. C., Seposo, X. T., & Hashizume, M. (2023). Heat exposure and the transmission of infectious diseases. *Heat Exposure and Human Health in the Context of Climate Change*, 189–221. https://doi.org/10.1016/B978-0-12-819080-7.00003-3

Clare, E. L., Chain, F. J. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data1. *Genome*, *59*(11), 981–990. https://doi.org/10.1139/GEN-2015-0184/SUPPL_FILE/GEN-2015-0184SUPPLF_ALL_THREELANES_WITH_ALL_MIDS_STILL_PRESENT.REVERSEPRIMER.TXT

Clare, E. L., Fazekas, A. J., Ivanova, N. V., Floyd, R. M., Hebert, P. D. N., Adams, A. M., Nagel, J., Girton, R., Newmaster, S. G., & Fenton, M. B. (2019). Approaches to integrating genetic data into ecological networks. *Molecular Ecology*, *28*(2), 503–519. https://doi.org/10.1111/MEC.14941

Couper, L. I., Farner, J. E., Caldwell, J. M., Childs, M. L., Harris, M. J., Kirk, D. G., Nova, N., Shocket, M., Skinner, E. B., Uricchio, L. H., Exposito-Alonso, M., & Mordecai, E. A. (2021). How will mosquitoes adapt to climate warming? *ELife*, *10*. https://doi.org/10.7554/ELIFE.69630

Cuff, J. P., Windsor, F. M., Tercel, M. P. T. G., Kitson, J. J. N., & Evans, D. M. (2022). Overcoming the pitfalls of merging dietary metabarcoding into ecological networks. *Methods in Ecology and Evolution*, *13*(3), 545–559. https://doi.org/10.1111/2041-210X.13796

Cunze, S., Kochmann, J., Koch, L. K., & Klimpel, S. (2016). Aedes albopictus and Its Environmental Limits in Europe. *PLOS ONE*, *11*(9), e0162116. https://doi.org/10.1371/JOURNAL.PONE.0162116

Dáttilo, W., & V. Rico-Gray. (2018). *Ecological Networks in the Tropics*. https://doi.org/10.1007/978-3-319-68228-0

De La Torre, E., Moreira, N., Saegerman, C., De Clercq, K., Salinas, M., Maldonado, A., Jarrín, D., Vaca, M. S., Pachacama, S., Espinoza, J., Delgado, H., Barrera, M., & Byrne, A. (2021). Bluetongue Virus Infections in Cattle Herds of Manabí Province of Ecuador. *Pathogens 2021, Vol. 10, Page 1445*, *10*(11), 1445. https://doi.org/10.3390/PATHOGENS10111445

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, *10*(9). https://doi.org/10.1098/RSBL.2014.0562

Deagle, B. E., Tollit, D. J., Jarman, S. N., Hindell, M. A., Trites, A. W., & Gales, N. J. (2005). Molecular scatology as a tool to study diet: Analysis of prey DNA in scats from captive Steller sea lions. *Molecular Ecology*, *14*(6), 1831–1842. https://doi.org/10.1111/J.1365-294X.2005.02531.X

Dhiman, R. C., & Sarkar, S. (2017). El Niño Southern Oscillation as an early warning tool for malaria outbreaks in India. *Malaria Journal*, *16*(1), 1–7. https://doi.org/10.1186/S12936-017-1779-Y/FIGURES/4

Dhiman, S. (2019). Are malaria elimination efforts on right track? An analysis of gains achieved and challenges ahead. *Infectious Diseases of Poverty 2019 8:1*, *8*(1), 1–19.

https://doi.org/10.1186/S40249-019-0524-X

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal*, *2*(1), 7–24. https://doi.org/10.2174/1874213000902010007

Drinkwater, R., Jucker, T., Potter, J. H. T., Swinfield, T., Coomes, D. A., Slade, E. M., Gilbert, M. T. P., Lewis, O. T., Bernard, H., Struebig, M. J., Clare, E. L., & Rossiter, S. J. (2020). Leech blood-meal invertebrate-derived DNA reveals differences in Bornean mammal diversity across habitats. *Molecular Ecology*, *00*, 1–14. https://doi.org/10.1111/mec.15724

Drinkwater, R., Schnell, I. B., Bohmann, K., Bernard, H., Veron, G., Clare, E., Gilbert, M. T. P., & Rossiter, S. J. (2019). Using metabarcoding to compare the suitability of two blood-feeding leech species for sampling mammalian diversity in North Borneo. *Molecular Ecology Resources*, *19*(1), 105–117. https://doi.org/10.1111/1755-0998.12943

Effler, P. V., Pang, L., Kitsutani, P., Vorndam, V., Nakata, M., Ayers, T., Elm, J., Tom, T., Reiter, P., Rigau-Perez, J. G., Hayes, J. M., Mills, K., Napier, M., Clark, G. G., & Gubler, D. J. (2005). Dengue fever, Hawaii, 2001-2002. *Emerging Infectious Diseases*, *11*(5), 742–749. https://doi.org/10.3201/EID1105.041063

Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, *43*(5), 2513–2524. https://doi.org/10.1093/NAR/GKV107

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., & Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology*, *30*(12), 1904–1916. https://doi.org/10.1111/1365-2435.12659

Evans, D. M., Pocock, M. J. O., & Memmott, J. (2013). The robustness of a network of ecological networks to habitat loss. *Ecology Letters*, *16*(7), 844–852. https://doi.org/10.1111/ELE.12117

Fath, B. D., Scharler, U. M., Ulanowicz, R. E., & Hannon, B. (2007). Ecological network analysis: network construction. *Ecological Modelling*, *208*(1), 49–55. https://doi.org/10.1016/J.ECOLMODEL.2007.04.029

Feng, X., Feng, J., Zhang, L., Tu, H., & Xia, Z. (2022). Vector control in China, from malaria endemic to elimination and challenges ahead. *Infectious Diseases of Poverty 2022 11:1*, *11*(1), 1–11. https://doi.org/10.1186/S40249-022-00971-3

Ferguson, H. M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M. S., Gimnig, J. E., Fish, D., & Killeen, G. F. (2010). Ecology: A Prerequisite for Malaria Elimination and Eradication. *PLoS Medicine*, *7*(8), e1000303. https://doi.org/10.1371/journal.pmed.1000303

Ferraguti, M., Martínez-De La Puente, J., Roiz, D., Ruiz, S., Soriguer, R., & Figuerola, J. (2016). Effects of landscape anthropization on mosquito community composition and abundance. *Scientific Reports*, *6*. https://doi.org/10.1038/srep29002

Figueiredo, L. T. M. (2019). Human Urban Arboviruses Can Infect Wild Animals and Jump to

Sylvatic Maintenance Cycles in South America. *Frontiers in Cellular and Infection Microbiology*, *9*(JUL), 259. https://doi.org/10.3389/FCIMB.2019.00259

Fitzpatrick, C., Nwankwo, U., Lenk, E., de Vlas, S. J., & Bundy, D. A. P. (2017). An Investment Case for Ending Neglected Tropical Diseases. *Disease Control Priorities, Third Edition (Volume 6): Major Infectious Diseases*, 411–431. https://doi.org/10.1596/978-1-4648-0524-0_CH17

Franklinos, L. H. V., Jones, K. E., Redding, D. W., & Abubakar, I. (2019). The effect of global change on mosquito-borne disease. In *The Lancet Infectious Diseases* (Vol. 19, Issue 9, pp. e302–e312). Lancet Publishing Group. https://doi.org/10.1016/S1473-3099(19)30161-6

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Fründ, J., McCann, K. S., & Williams, N. M. (2016). Sampling bias is a challenge for quantifying specialization and network structure: lessons from a quantitative niche model. *Oikos*, *125*(4), 502–513. https://doi.org/10.1111/oik.02256

Fukui, S., Kuwano, Y., Ueno, K., Atsumi, K., & Ohta Id, S. (2022). Modeling the effect of rainfall changes to predict population dynamics of the Asian tiger mosquito Aedes albopictus under future climate conditions. *PLOS ONE*, *17*(5), e0268211. https://doi.org/10.1371/JOURNAL.PONE.0268211

Gaithuma, A., Yamagishi, J., Hayashida, K., Kawai, N., Namangala, B., & Sugimoto, C. (2020). Blood meal sources and bacterial microbiome diversity in wild-caught tsetse flies. *Scientific Reports 2020 10:1*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-61817-2

Gallup, J. L., & Sachs, J. D. (2001). The economic burden of malaria. *The American Journal of Tropical Medicine and Hygiene*, *64*(1-2 Suppl), 85–96. https://doi.org/10.4269/AJTMH.2001.64.85

Gibb, R., Franklinos, L. H. V, Redding, D. W., & Jones, K. E. (2020). Ecosystem perspectives are needed to manage zoonotic risks in a changing climate. *BMJ*, *371*, m3389. https://doi.org/10.1136/bmj.m3389

Gibb, R., Redding, D. W., Chin, K. Q., Donnelly, C. A., Blackburn, T. M., Newbold, T., & Jones, K. E. (2020). Zoonotic host diversity increases in human-dominated ecosystems. *Nature 2020 584:7821*, *584*(7821), 398–402. https://doi.org/10.1038/s41586-020-2562-8

Giglioli, G. (1963). ECOLOGICAL CHANGE AS A FACTOR IN RENEWED MALARIA TRANSMISSION IN AN ERADICATED AREA. A LOCALIZED OUTBREAK OF A. AQUASALIS-TRANSMITTED MALARIA ON THE DEMERARA RIVER ESTUARY, BRITISH GUIANA, IN THE FIFTEENTH YEAR OF A. DARLINGI AND MALARIA ERADICATION. *Bulletin of the World Health Organization*, *29*(2), 131–145. http://www.ncbi.nlm.nih.gov/pubmed/14056265

Gillies, M. T., & Furlong, M. (1964). *COMMONWEALTH INSTITUTE OF ENTOMOLOGY BULLETIN OF ENTOMOLOGICAL RESEARCH 0045/-AN INVESTIGATION INTO THE BEHAVIOUE OF ANOPHELES-'/ " PARENSIS GILLIES AT MALXNDI ON THE KENYA COAST East African Institute of Malaria and V'actor-borne Diseases*.

Gillies, M. T., & 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*(02), 243. https://doi.org/10.1017/S0007485300057953

Go, Y. Y., Balasuriya, U. B. R., & Lee, C. (2014). Zoonotic encephalitides caused by arboviruses: transmission and epidemiology of alphaviruses and flaviviruses. *Clinical and Experimental Vaccine Research*, *3*(1), 58. https://doi.org/10.7774/CEVR.2014.3.1.58

Godfrey, S. S. (2013). Networks and the ecology of parasite transmission: A framework for wildlife parasitology. In *International Journal for Parasitology: Parasites and Wildlife* (Vol. 2, Issue 1, pp. 235–245). Elsevier. https://doi.org/10.1016/j.ijppaw.2013.09.001

Golding, N., Wilson, A. L., Moyes, C. L., Cano, J., Pigott, D. M., Velayudhan, R., Brooker, S. J., Smith, D. L., Hay, S. I., & Lindsay, S. W. (2015). Integrating vector control across diseases. *BMC Medicine*, *13*(1), 1–6. https://doi.org/10.1186/S12916-015-0491-4/FIGURES/2

Gordon, R., Ivens, S., Ammerman, L. K., Fenton, M. B., Littlefair, J. E., Ratcliffe, J. M., & Clare, E. L. (2019). Molecular diet analysis finds an insectivorous desert bat community dominated by resource sharing despite diverse echolocation and foraging strategies. *Ecology and Evolution*, *9*(6), 3117. https://doi.org/10.1002/ECE3.4896

Gould, E., Pettersson, J., Higgs, S., Charrel, R., & de Lamballerie, X. (2017). Emerging arboviruses: Why today? *One Health*, *4*, 1–13. https://doi.org/10.1016/J.ONEHLT.2017.06.001

Gratton, C., Donaldson, J., & Zanden, M. J. Vander. (2008). Ecosystem linkages between lakes and the surrounding terrestrial landscape in northeast Iceland. *Ecosystems*, *11*(5), 764–774. https://doi.org/10.1007/S10021-008-9158-8/TABLES/2

Gratz, N. G. (1999). Emerging and resurging vector-borne diseases. *Annual Review of Entomology*, *44*(1), 51–75. https://doi.org/10.1146/annurev.ento.44.1.51

Grubaugh, N. D., Ladner, J. T., Lemey, P., Pybus, O. G., Rambaut, A., Holmes, E. C., & Andersen, K. G. (2019). Tracking virus outbreaks in the twenty-first century. *Nature Microbiology*, *4*(1), 10–19. https://doi.org/10.1038/s41564-018-0296-2

Gubler, D. J. (2011). Dengue, Urbanization and Globalization: The Unholy Trinity of the 21st Century. *Tropical Medicine and Health*, *39*(4 Suppl), 3. https://doi.org/10.2149/TMH.2011-S05

Hamlet, A., Jean, K., Perea, W., Yactayo, S., Biey, J., Van Kerkhove, M., Ferguson, N., & Garske, T. (2018). The seasonal influence of climate and environment on yellow fever transmission across Africa. *PLOS Neglected Tropical Diseases*, *12*(3), e0006284. https://doi.org/10.1371/journal.pntd.0006284

Harper, G. L., Sheppard, S. K., Harwood, J. D., Read, D. S., Glen, D. M., Bruford, M. W., & Symondson, W. O. C. (2006). Evaluation of temperature gradient gel electrophoresis for the analysis of prey DNA within the guts of invertebrate predators. *Bulletin of Entomological Research*, *96*(3), 295–304. https://doi.org/10.1079/BER2006426

Harrington, L. C., Foy, B. D., & Bangs, M. J. (2020). Considerations for Human Blood-Feeding

and Arthropod Exposure in Vector Biology Research: An Essential Tool for Investigations and Disease Control. *Vector Borne and Zoonotic Diseases*, *20*(11), 807. https://doi.org/10.1089/VBZ.2020.2620

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and Disease Emergence: Dynamics at the Wildlife–Livestock–Human Interface. In *Trends in Ecology and Evolution* (Vol. 32, Issue 1, pp. 55–67). Elsevier Ltd. https://doi.org/10.1016/j.tree.2016.09.012

Hemprich-Bennett, D. R., Kemp, V. A., Blackman, J., Struebig, M. J., Lewis, O. T., Rossiter, S. J., & Clare, E. L. (2021). Altered structure of bat–prey interaction networks in logged tropical forests revealed by metabarcoding. *Molecular Ecology*, *30*(22), 5844–5857. https://doi.org/10.1111/MEC.16153

Hemprich-Bennett, D. R., Oliveira, H. F. M., Le Comber, S. C., Rossiter, S. J., & Clare, E. L. (2021). Assessing the impact of taxon resolution on network structure. *Ecology*, *102*(3). https://doi.org/10.1002/ECY.3256

Hesson, J. C., Ignell, R., Hill, S. R., Östman, Ö., & Lundström, J. O. (2015). Trapping biases of *Culex torrentium* and *Culex pipiens* revealed by comparison of captures in CDC traps, ovitraps, and gravid traps. *Journal of Vector Ecology*, *40*(1), 158–163. https://doi.org/10.1111/jvec.12145

Hopken, M. W., Reyes-Torres, L. J., Scavo, N., Piaggio, A. J., Abdo, Z., Taylor, D., Pierce, J., & Yee, D. A. (2021). Temporal and Spatial Blood Feeding Patterns of Urban Mosquitoes in the San Juan Metropolitan Area, Puerto Rico. *Insects*, *12*(2), 1–21. https://doi.org/10.3390/INSECTS12020129

Hoyos, J., Carrasquilla, C., León, C., Montgomery, J. M., Salyer, S. J., Komar, N., & González, C. (2021). Host selection pattern and flavivirus screening of mosquitoes in a disturbed Colombian rainforest. *Scientific Reports |*, *11*(1), 18656. https://doi.org/10.1038/s41598-021-98076-8

Hung, T. M., Shepard, D. S., Bettis, A. A., Nguyen, H. A., McBride, A., Clapham, H. E., & Turner, H. C. (2020). Productivity costs from a dengue episode in Asia: A systematic literature review. *BMC Infectious Diseases*, *20*(1), 1–18. https://doi.org/10.1186/S12879-020-05109-0/FIGURES/6

Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., Edwards, F., Figueroa, D., Jacob, U., Jones, J. I., Lauridsen, R. B., Ledger, M. E., Lewis, H. M., Olesen, J. M., van Veen, F. J. F., Warren, P. H., & Woodward, G. (2009). Review: Ecological networks - beyond food webs. *Journal of Animal Ecology*, *78*(1), 253–269. https://doi.org/10.1111/j.1365-2656.2008.01460.x

Iturbe-Ormaetxe, I., Walker, T., & O'Neill, S. L. (2011). Wolbachia and the biological control of mosquito-borne disease. *EMBO Reports*, *12*(6), 508. https://doi.org/10.1038/EMBOR.2011.84

Iwamura, T., Guzman-Holst, A., & Murray, K. A. (2020). Accelerating invasion potential of disease vector Aedes aegypti under climate change. *Nature Communications 2020 11:1*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-16010-4

J. Kweka, E., Baraka, V., Mathias, L., Mwang'onde, B., Baraka, G., Lyaruu, L., & M. Mahande, A. (2019). Ecology of Aedes Mosquitoes, the Major Vectors of Arboviruses in Human Population. In *Dengue Fever - a Resilient Threat in the Face of Innovation*. IntechOpen. https://doi.org/10.5772/intechopen.81439

Jansen, C. C., Webb, C. E., Graham, G. C., Craig, S. B., Zborowski, P., Ritchie, S. A., Russell, R. C., & Van Den Hurk, A. F. (2009). Blood sources of mosquitoes collected from urban and peri-urban environments in eastern Australia with species-specific molecular analysis of avian blood meals. *American Journal of Tropical Medicine and Hygiene*, *81*(5), 849–857. https://doi.org/10.4269/ajtmh.2009.09-0008

Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., McKeever, D., Mutua, F., Young, J., McDermott, J., & Pfeiffer, D. U. (2013). Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8399–8404. https://doi.org/10.1073/PNAS.1208059110/-/DCSUPPLEMENTAL

Joos, L., Beirinckx, S., Haegeman, A., Debode, J., Vandecasteele, B., Baeyen, S., Goormachtig, S., Clement, L., & De Tender, C. (2020). Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics*, *21*(1), 1–17. https://doi.org/10.1186/S12864-020-07126-4/FIGURES/6

Jordano, P. (1987). Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. *American Naturalist*, *129*(5), 657–677. https://doi.org/10.1086/284665

Kaiser-Bunbury, C. N., & Blüthgen, N. (2015). Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants*, *7*, plv076. https://doi.org/10.1093/aobpla/plv076

Karimkhani, C., Wanga, V., Coffeng, L. E., Naghavi, P., Dellavalle, R. P., & Naghavi, M. (2016). Global burden of cutaneous leishmaniasis: a cross-sectional analysis from the Global Burden of Disease Study 2013. *The Lancet Infectious Diseases*, *16*(5), 584–591. https://doi.org/10.1016/S1473-3099(16)00003-7

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., & Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature 2010 468:7324*, *468*(7324), 647–652. https://doi.org/10.1038/nature09575

Kocher, A., Gantier, J.-C., Gaborit, P., Zinger, L., Holota, H., Valiere, S., Dusfour, I., Girod, R., Bañuls, A.-L., & Murienne, J. (2017). Vector soup: high-throughput identification of Neotropical phlebotomine sand flies using metabarcoding. *Molecular Ecology Resources*, *17*(2), 172–182. https://doi.org/10.1111/1755-0998.12556

Koenraadt, C. J., Spitzen, J., & Takken, W. (2021). *Innovative strategies for vector control*. *6*. https://doi.org/10.3920/978-90-8686-895-7

Kolimenakis, A., Heinz, S., Wilson, M. L., Winkler, V., Yakob, L., Michaelakis, A., Papachristos, D., Richardson, C., & Horstick, O. (2021). The role of urbanisation in the spread of Aedes mosquitoes and the diseases they transmit—A systematic review. *PLOS*

Neglected Tropical Diseases, *15*(9), e0009631. https://doi.org/10.1371/JOURNAL.PNTD.0009631

Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., Nolan, T., & Crisanti, A. (2018). A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature Biotechnology 2018 36:11*, *36*(11), 1062–1066. https://doi.org/10.1038/nbt.4245

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology*, *28*(2), 420–430. https://doi.org/10.1111/MEC.14920

Landi, P., Minoarivelo, H. O., Brännström, Å., Hui, C., & Dieckmann, U. (2018). Complexity and stability of ecological networks: a review of the theory. *Population Ecology*, *60*(4), 319–345. https://doi.org/10.1007/s10144-018-0628-3

Lau, M. K., Borrett, S. R., Baiser, B., Gotelli, N. J., & Ellison, A. M. (2017). Ecological network metrics: opportunities for synthesis. *Ecosphere*, *8*(8), e01900. https://doi.org/10.1002/ecs2.1900

Lecomte, N., Ahlstrøm, Ø., Ehrich, D., Fuglei, E., Ims, R. A., & Yoccoz, N. G. (2011). Intrapopulation Variability Shaping Isotope Discrimination and Turnover: Experimental Evidence in Arctic Foxes. *PLOS ONE*, *6*(6), e21357. https://doi.org/10.1371/JOURNAL.PONE.0021357

Lim, A. Y., Cheong, H. K., Chung, Y., Sim, K., & Kim, J. H. (2021). Mosquito abundance in relation to extremely high temperatures in urban and rural areas of Incheon Metropolitan City, South Korea from 2015 to 2020: an observational study. *Parasites and Vectors*, *14*(1), 1–10. https://doi.org/10.1186/S13071-021-05071-Z/FIGURES/5

Liu, Y. X., Qin, Y., Chen, T., Lu, M., Qian, X., Guo, X., & Bai, Y. (2021). A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein and Cell*, *12*(5), 315–330. https://doi.org/10.1007/S13238-020-00724-8/TABLES/3

Logue, K., Keven, J. B., Cannon, M. V., Reimer, L., Siba, P., Walker, E. D., Zimmerman, P. A., & Serre, D. (2016). Unbiased Characterization of Anopheles Mosquito Blood Meals by Targeted High-Throughput Sequencing. *PLOS Neglected Tropical Diseases*, *10*(3), e0004512. https://doi.org/10.1371/journal.pntd.0004512

Maclachlan, N. J., Mayo, C. E., Daniels, P. W., Savini, G., Zientara, S., & Gibbs, E. P. J. (2015). Bluetongue. *Revue Scientifique et Technique (International Office of Epizootics)*, *34*(2), 329–340. https://doi.org/10.20506/RST.34.2.2360

Maheshwari, P., & Albert, R. (2020). Network model and analysis of the spread of Covid-19 with social distancing. *Applied Network Science*, *5*(1). https://doi.org/10.1007/S41109-020-00344-5

Martín González, A. M., Dalsgaard, B., & Olesen, J. M. (2010). Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, *7*(1), 36–43. https://doi.org/10.1016/j.ecocom.2009.03.008

MartíNez-Ibarra, J. A., Guillén, Y. G., Arredondo-Jiménez, J. I., & Rodríguez-López, M. H. (2002). Indigenous fish species for the control of Aedes aegypti in water storage tanks

in Southern México. *BioControl*, *47*(4), 481–486. https://doi.org/10.1023/A:1015691831489/METRICS

Mattah, P. A. D., Futagbi, G., Amekudzi, L. K., Mattah, M. M., de Souza, D. K., Kartey-Attipoe, W. D., Bimi, L., & Wilson, M. D. (2017). Diversity in breeding sites and distribution of Anopheles mosquitoes in selected urban areas of southern Ghana. *Parasites & Vectors*, *10*(1), 25. https://doi.org/10.1186/s13071-016-1941-3

Mayi, M. P. A., Foncha, D. F., Kowo, C., Tchuinkam, T., Brisco, K., Anong, D. N., Ravinder, S., & Cornel, A. J. (2019). Impact of deforestation on the abundance, diversity, and richness of Culex mosquitoes in a southwest Cameroon tropical rainforest. *Journal of Vector Ecology*, *44*(2), 271–281. https://doi.org/10.1111/jvec.12359

McCann, R. S., van den Berg, H., Diggle, P. J., van Vugt, M., Terlouw, D. J., Phiri, K. S., Di Pasquale, A., Maire, N., Gowelo, S., Mburu, M. M., Kabaghe, A. N., Mzilahowa, T., Chipeta, M. G., & Takken, W. (2017). Assessment of the effect of larval source management and house improvement on malaria transmission when added to standard malaria control strategies in southern Malawi: Study protocol for a cluster-randomised controlled trial. *BMC Infectious Diseases*, *17*(1), 1–15. https://doi.org/10.1186/S12879-017-2749-2/FIGURES/5

Melián, C. J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P., & Williams, R. J. (2011). Eco-evolutionary Dynamics of Individual-Based Food Webs. *Advances in Ecological Research*, *45*, 225–268. https://doi.org/10.1016/B978-0-12-386475-8.00006-X

Miot, E. F., Calvez, E., Aubry, F., Dabo, S., Grandadam, M., Marcombe, S., Oke, C., Logan, J. G., Brey, P. T., & Lambrechts, L. (2020). Risk of arbovirus emergence via bridge vectors: case study of the sylvatic mosquito Aedes malayensis in the Nakai district, Laos. *Scientific Reports 2020 10:1*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-64696-9

Miranda, P. N., Ribeiro, J. E. L. da S., Luna, P., Brasil, I., Delabie, J. H. C., & Dáttilo, W. (2019). The dilemma of binary or weighted data in interaction networks. *Ecological Complexity*, *38*, 1–10. https://doi.org/10.1016/j.ecocom.2018.12.006

Moreira, E. F., Ferreira, A., Lopes, L. E., Gomes, R., Soares, S., Boscolo, D., Moreira, E. F., Ferreira, P. A., Lopes, L. E., Soares, • R G S, & Boscolo, D. (2018). Ecological Networks in Changing Tropics. *Ecological Networks in the Tropics*, 155–169. https://doi.org/10.1007/978-3-319-68228-0_11

Morris, R. J., Gripenberg, S., Lewis, O. T., & Roslin, T. (2014). Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*, *17*(3), 340–349. https://doi.org/10.1111/ele.12235

Morrison, D., Bedinger, M., Beevers, L., & McClymont, K. (2022). Exploring the raison d'etre behind metric selection in network analysis: a systematic review. *Applied Network Science 2022 7:1*, *7*(1), 1–25. https://doi.org/10.1007/S41109-022-00476-W

Mougi, A. (2016). The roles of amensalistic and commensalistic interactions in large ecological network stability. *Scientific Reports*, *6*(1), 29929. https://doi.org/10.1038/srep29929

Näslund, J., Ahlm, C., Islam, K., Evander, M., Ran Bucht, G., & Lwande, O. W. (2021).

Emerging Mosquito-Borne Viruses Linked to Aedes aegypti and Aedes albopictus: Global Status and Preventive Strategies. *Vector-Borne and Zoonotic Diseases*, *Vol. 21*(No. 10). https://doi.org/10.1089/vbz.2020.2762

Nearing, J. T., Douglas, G. M., Comeau, A. M., & Langille, M. G. I. (2018). Denoising the Denoisers: An independent evaluation of microbiome sequence error- correction approaches. *PeerJ*, *2018*(8). https://doi.org/10.7717/PEERJ.5364/SUPP-18

Ngonghala, C. N., De Leo, G. A., Pascual, M. M., Keenan, D. C., Dobson, A. P., & Bonds, M. H. (2017). General ecological models for human subsistence, health and poverty. *Nature Ecology & Evolution 2017 1:8*, *1*(8), 1153–1159. https://doi.org/10.1038/s41559-017-0221-8

Nguyen-Tien, T., Lundkvist, Å., & Lindahl, J. (2019). Urban transmission of mosquito-borne flaviviruses – a review of the risk for humans in Vietnam. *Https://Doi.Org/10.1080/20008686.2019.1660129*, *9*(1). https://doi.org/10.1080/20008686.2019.1660129

Nielsen, S. S., Alvarez, J., Bicout, D. J., Calistri, P., Depner, K., Drewe, J. A., Garin-Bastuji, B., Gonzales Rojas, J. L., Schmidt, C. G., Michel, V., Miranda Chueca, M. A., Roberts, H. C., Sihvonen, L. H., Stahl, K., Velarde, A., Trop, A., Winckler, C., Cetre-Sossah, C., Chevalier, V., … Van der Stede, Y. (2020). Rift valley fever: Risk of persistence, spread and impact in Mayotte (France). *EFSA Journal*, *18*(4), 1–61. https://doi.org/10.2903/J.EFSA.2020.6093

Niquil, N., Chaumillon, E., Johnson, G. A., Bertin, X., Grami, B., David, V., Bacher, C., Asmus, H., Baird, D., & Asmus, R. (2012). The effect of physical drivers on ecosystem indices derived from ecological network analysis: Comparison across estuarine ecosystems. *Estuarine, Coastal and Shelf Science*, *108*, 132–143. https://doi.org/10.1016/J.ECSS.2011.12.031

Orta-Pineda, G., Abella-Medrano, C. A., Suzán, G., Serrano-Villagrana, A., & Ojeda-Flores, R. (2021). Effects of landscape anthropization on sylvatic mosquito assemblages in a rainforest in Chiapas, Mexico. *Acta Tropica*, *216*, 105849. https://doi.org/10.1016/j.actatropica.2021.105849

Ostrom, P. H., Colunga-Garcia, M., & Gage, S. H. (1996). Establishing pathways of energy flow for insect predators using stable isotope ratios: field and laboratory evidence. *Oecologia 1996 109:1*, *109*(1), 108–113. https://doi.org/10.1007/S004420050064

Oviedo-Pastrana, M., Méndez, N., Mattar, S., Arrieta, G., & Gomezcaceres, L. (2017). Epidemic outbreak of Chikungunya in two neighboring towns in the Colombian Caribbean: A survival analysis. *Archives of Public Health*, *75*(1), 1–8. https://doi.org/10.1186/S13690-016-0169-1/TABLES/2

Parmakelis, A., Russello, M. A., Caccone, A., Marcondes, C. B., Costa, J., Forattini, O. P., Sallum, M. A. M., Wilkerson, R. C., & Powell, J. R. (2008). Historical Analysis of a Near Disaster: Anopheles gambiae in Brazil. *The American Journal of Tropical Medicine and Hygiene*, *78*(1), 176–178. https://doi.org/10.4269/AJTMH.2008.78.176

Patterson, B. D., & Atmar, W. (1986). Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biological Journal of the Linnean Society*, *28*(1–2),

65–82. https://doi.org/10.1111/J.1095-8312.1986.TB01749.X

Patz, J. A., Daszak, P., Tabor, G. M., Aguirre, A. A., Pearl, M., Epstein, J., Wolfe, N. D., Kilpatrick, A. M., Foufopoulos, J., Molyneux, D., Bradley, D. J., Amerasinghe, F. P., Ashford, R. W., Barthelemy, D., Bos, R., Bradley, D. J., Buck, A., Butler, C., Chivian, E. S., … Zakarov, V. (2004). Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence. *Environmental Health Perspectives*, *112*(10), 1092. https://doi.org/10.1289/EHP.6877

Paz, S. (2015). Climate change impacts on West Nile virus transmission in a global context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1665), 1–11. https://doi.org/10.1098/RSTB.2013.0561

Peyre, M., Chevalier, V., Abdo-Salem, S., Velthuis, A., Antoine-Moussiaux, N., Thiry, E., & Roger, F. (2015). A Systematic Scoping Study of the Socio-Economic Impact of Rift Valley Fever: Research Gaps and Needs. *Zoonoses and Public Health*, *62*(5), 309–325. https://doi.org/10.1111/ZPH.12153

Piñol, J., San Andrés, V., Clare, E. L., Mir, G., & Symondson, W. O. C. (2014). A pragmatic approach to the analysis of diets of generalist predators: the use of next-generation sequencing with no blocking probes. *Molecular Ecology Resources*, *14*(1), 18–26. https://doi.org/10.1111/1755-0998.12156

Pollegioni, P., North, A. R., Persampieri, T., Bucci, A., Minuz, R. L., Groneberg, D. A., Nolan, T., Papathanos, P. A., Crisanti, A., & Müller, R. (2020). Detecting the population dynamics of an autosomal sex ratio distorter transgene in malaria vector mosquitoes. *Journal of Applied Ecology*, *57*(10), 2086–2096. https://doi.org/10.1111/1365-2664.13702

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, *21*(8), 1931–1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x

Poulin, R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, *26*(10), 492–498. https://doi.org/10.1016/J.PT.2010.05.008

Proulx, S. R., Promislow, D. E. L., & Phillips, P. C. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, *20*(6), 345–353. https://doi.org/10.1016/J.TREE.2005.04.004

Pybus, O. G., Tatem, A. J., & Lemey, P. (2015). Virus evolution and transmission in an ever more connected world. *Proceedings. Biological Sciences*, *282*(1821), 20142878. https://doi.org/10.1098/rspb.2014.2878

Quintero, E., Isla, J., & Jordano, P. (2022). Methodological overview and data-merging approaches in the study of plant–frugivore interactions. *Oikos*, *2022*(2). https://doi.org/10.1111/OIK.08379

Ranson, H., & 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), 187–196. https://doi.org/10.1016/J.PT.2015.11.010

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System: Barcoding. *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/10.1111/J.1471-8286.2007.01678.X

Reeves, L. E., Gillett-Kaufman, J. L., Kawahara, A. Y., & Kaufman, P. E. (2018). Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. *PLOS Neglected Tropical Diseases*, *12*(8), e0006767. https://doi.org/10.1371/journal.pntd.0006767

Reeves, L. E., Holderman, C. J., Gillett-Kaufman, J. L., Kawahara, A. Y., & Kaufman, P. E. (2016). Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasites & Vectors*, *9*(1), 503. https://doi.org/10.1186/s13071-016-1791-z

Reñones, O., Polunin, N. V. C., & Goni, R. (2002). Size related dietary shifts of Epinephelus marginatus in a western Mediterranean littoral ecosystem: an isotope and stomach content analysis. *Journal of Fish Biology*, *61*(1), 122–137. https://doi.org/10.1111/J.1095-8649.2002.TB01741.X

Robeson, M. S., Khanipov, K., Golovko, G., Wisely, S. M., White, M. D., Bodenchuck, M., Smyser, T. J., Fofanov, Y., Fierer, N., & Piaggio, A. J. (2018). Assessing the utility of metabarcoding for diet analyses of the omnivorous wild pig (Sus scrofa). *Ecology and Evolution*, *8*(1), 185–196. https://doi.org/10.1002/ECE3.3638

Rocklöv, J., & Dubrow, R. (2020). Climate change: an enduring challenge for vector-borne disease prevention and control. In *Nature Immunology* (Vol. 21, Issue 5, pp. 479–483). Nature Research. https://doi.org/10.1038/s41590-020-0648-y

Ross, R. (1897). On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood. *British Medical Journal*, *2*(1929), 1786. https://doi.org/10.1136/BMJ.2.1929.1786

Rushton, J., & Lyons, N. (2015). Economic impact of Bluetongue: a review of the effects on production. *Veterinaria Italiana*, *51*(4), 401–406. https://doi.org/10.12834/VETIT.646.3183.1

Ryan, S. J., Lippi, C. A., & Zermoglio, F. (2020). Shifting transmission risk for malaria in Africa with climate change: A framework for planning and intervention. *Malaria Journal*, *19*(1), 1–14. https://doi.org/10.1186/S12936-020-03224-6/FIGURES/7

Santiago, P. B., De Araújo, C. N., Motta, F. N., Praça, Y. R., Charneau, S., Bastos, I. M. D., & Santana, J. M. (2017). Proteases of haematophagous arthropod vectors are involved in blood-feeding, yolk formation and immunity - a review. *Parasites & Vectors 2017 10:1*, *10*(1), 1–20. https://doi.org/10.1186/S13071-017-2005-Z

Schirmer, M., Ijaz, U. Z., D'Amore, R., Hall, N., Sloan, W. T., & Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Research*, *43*(6), e37. https://doi.org/10.1093/NAR/GKU1341

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-source, platform-

independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09/ASSET/14428E0A-11AE-4C2E-8480-78064604726F/ASSETS/GRAPHIC/ZAM0230904840002.JPEG

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., Bertelsen, M. F., & Gilbert, M. T. P. (2012). Screening mammal biodiversity using DNA from leeches. *Current Biology*, *22*(8), R262–R263. https://doi.org/10.1016/J.CUB.2012.02.058

Sheppard, S. K., & Harwood, J. D. (2005). Advances in molecular ecology: Tracking trophic links through predator-prey food-webs. *Functional Ecology*, *19*(5), 751–762. https://doi.org/10.1111/J.1365-2435.2005.01041.X

Shroff, S., Mir, S., Naik, B., Baitharu, I., & Behera, A. K. (2020). Chemical Methods for Control of Mosquito Vector. *Molecular Identification of Mosquito Vectors and Their Management*, 35–50. https://doi.org/10.1007/978-981-15-9456-4_3

Simmons, B. I., Cirtwill, A. R., Baker, N. J., Wauchope, H. S., Dicks, L. V., Stouffer, D. B., & Sutherland, W. J. (2019). Motifs in bipartite ecological networks: uncovering indirect interactions. *Oikos*, *128*(2), 154–170. https://doi.org/10.1111/OIK.05670

Sousa, L. L., Silva, S. M., & Xavier, R. (2019). DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. *Environmental DNA*, edn3.27. https://doi.org/10.1002/edn3.27

Swain, A., MacCracken, S. A., Fagan, W. F., & Labandeira, C. C. (2022). Understanding the ecology of host plant–insect herbivore interactions in the fossil record through bipartite networks. *Paleobiology*, *48*(2), 239–260. https://doi.org/10.1017/PAB.2021.20

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For Biodiversity Research and Monitoring. *Environmental DNA: For Biodiversity Research and Monitoring*, 1–253. https://doi.org/10.1093/OSO/9780198767220.001.0001

Tatem, A. J., Hay, S. I., & Rogers, D. J. (2006). Global traffic and disease vector dispersal. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(16), 6242–6247. https://doi.org/10.1073/PNAS.0508391103/ASSET/9A4B21EF-0293-4852-ABF0-CF04479332F4/ASSETS/GRAPHIC/ZPQ0140618070003.JPEG

Tizifa, T. A., Kabaghe, A. N., McCann, R. S., van den Berg, H., Van Vugt, M., & Phiri, K. S. (2018). Prevention Efforts for Malaria. *Current Tropical Medicine Reports*, *5*(1), 41–50. https://doi.org/10.1007/s40475-018-0133-y

Townroe, S., & Callaghan, A. (2014). British Container Breeding Mosquitoes: The Impact of Urbanisation and Climate Change on Community Composition and Phenology. *PLoS ONE*, *9*(4), e95325. https://doi.org/10.1371/journal.pone.0095325

Tusting, L. S., Thwing, J., Sinclair, D., Fillinger, U., Gimnig, J., Bonner, K. E., Bottomley, C., & Lindsay, S. W. (2013). Mosquito larval source management for controlling malaria. *The Cochrane Database of Systematic Reviews*, *2013*(8). https://doi.org/10.1002/14651858.CD008923.PUB2

UN., D. (2019). World Urbanization Prospects: The 2018 Revision. *World Urbanization*

Prospects: The 2018 Revision. https://doi.org/10.18356/B9E995FE-EN

UNICEF. (2017). *Malaria in Africa - UNICEF DATA*. https://data.unicef.org/topic/child-health/malaria/

Valentine, M. J., Murdock, C. C., & Kelly, P. J. (2019). Sylvatic cycles of arboviruses in non-human primates. *Parasites & Vectors*, *12*(1). https://doi.org/10.1186/S13071-019-3732-0

Van Den Berg, H., Von Hildebrand, A., Ragunathan, V., & Das, P. K. (2007). Reducing vector-borne disease by empowering farmers in integrated vector management. *Bulletin of the World Health Organization*, *85*(7), 561. https://doi.org/10.2471/BLT.06.035600

Vasilakis, N., Cardosa, J., Hanley, K. A., Holmes, E. C., & Weaver, S. C. (2011). Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. *Nature Reviews Microbiology 2011 9:7*, *9*(7), 532–541. https://doi.org/10.1038/nrmicro2595

Vazquez-Prokopec, G. M., Chaves, L. F., Ritchie, S. A., Davis, J., & Kitron, U. (2010). Unforeseen Costs of Cutting Mosquito Surveillance Budgets. *PLoS Neglected Tropical Diseases*, *4*(10), 858. https://doi.org/10.1371/JOURNAL.PNTD.0000858

Walrant, A., & Loreau, M. (1995). Comparison of Iso-enzymeElectrophoresis and Gut Content Examination for Determining theNatural Diet of the Groundbeetle Species Abax ater (Coleoptera:Carabidae). *Entomologia Generalis*, *19*(4), 253–259. https://doi.org/10.1127/ENTOM.GEN/19/1995/253

Walter, C. B., & O'Neill, E. (1986). Electrophoresis in the study of diets and digestive rates of seabirds. *Comparative Biochemistry and Physiology Part A: Physiology*, *84*(4), 763–765. https://doi.org/10.1016/0300-9629(86)90402-0

Weaver, S. C., & Barrett, A. D. T. (2004). Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature Reviews. Microbiology*, *2*(10), 789. https://doi.org/10.1038/NRMICRO1006

Weaver, S. C., & Reisen, W. K. (2010). Present and future arboviral threats. In *Antiviral Research* (Vol. 85, Issue 2, pp. 328–345). Elsevier. https://doi.org/10.1016/j.antiviral.2009.10.008

Westcott, S. L., & Schloss, P. D. (2015). De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ*, *2015*(12), e1487. https://doi.org/10.7717/PEERJ.1487/FIG-5

White, R. J., & Razgour, O. (2020). Emerging zoonotic diseases originating in mammals: a systematic review of effects of anthropogenic land-use change. *Mammal Review*, mam.12201. https://doi.org/10.1111/mam.12201

WHO. (2014). *A global brief on vector-borne diseases*. www.who.int

WHO. (2020). *Vector-borne diseases*. https://doi.org/10.1002/9781118597361.ch7

Wilson, A. L., Courtenay, O., Kelly-Hope, L. A., Scott, T. W., Takken, W., Torr, S. J., & Lindsay, S. W. (2020). The importance of vector control for the control and elimination of

vector-borne diseases. *PLOS Neglected Tropical Diseases*, *14*(1), e0007831. https://doi.org/10.1371/journal.pntd.0007831

World Health Organization. (2015). *Investing to overcome the global impact of neglected tropical diseases: third WHO report on neglected diseases 2015*. https://apps.who.int/iris/handle/10665/152781

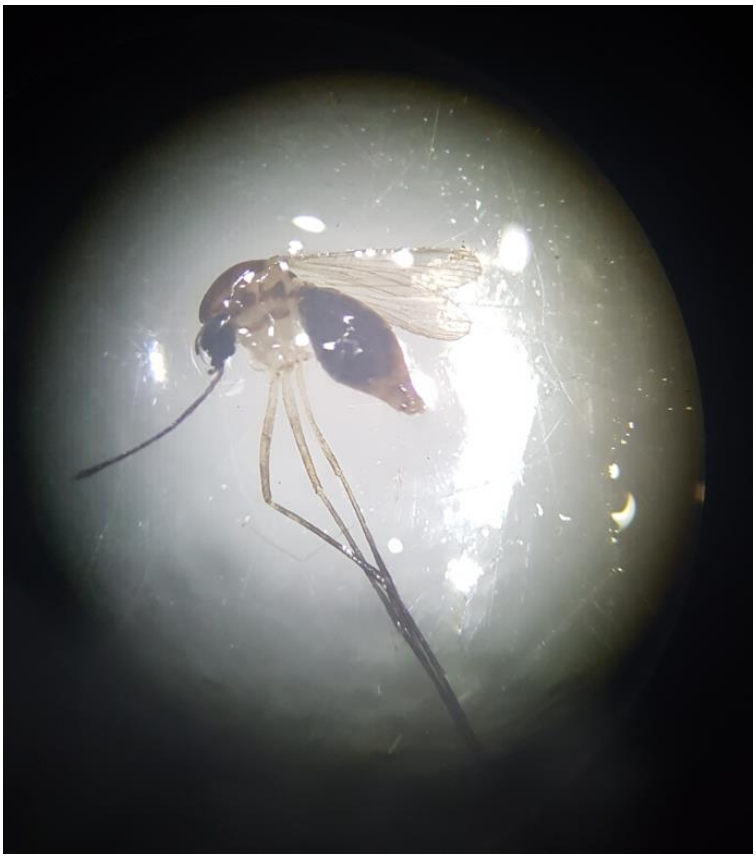
Yue, Y., Sun, J., Liu, X., Ren, D., Liu, Q., Xiao, X., & Lu, L. (2018). Spatial analysis of dengue fever and exploration of its environmental and socio-economic risk factors using ordinary least squares: A case study in five districts of Guangzhou City, China, 2014. *International Journal of Infectious Diseases : IJID : Official Publication of the International Society for Infectious Diseases*, *75*, 39–48. https://doi.org/10.1016/J.IJID.2018.07.023

Yukich, J. O., Lengeler, C., Tediosi, F., Brown, N., Mulligan, J. A., Chavasse, D., Stevens, W., Justino, J., Conteh, L., Maharaj, R., Erskine, M., Mueller, D. H., Wiseman, V., Ghebremeskel, T., Zerom, M., Goodman, C., McGuire, D., Urrutia, J. M., Sakho, F., … Sharp, B. (2008). Costs and consequences of large-scale vector control for malaria. *Malaria Journal*, *7*(1), 1–12. https://doi.org/10.1186/1475-2875-7-258/TABLES/6

Yun, S. I., & Lee, Y. M. (2017). Zika virus: An emerging flavivirus. *Journal of Microbiology 2017 55:3*, *55*(3), 204–219. https://doi.org/10.1007/S12275-017-7063-6

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L., Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. https://doi.org/10.1111/mec.15060

Chapter 2: A network perspective on the vectoring of human disease



Author contributions: BB: Data analysis and writing, OTL: Supervision, review, and editing, TDH: Supervision, review, and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

This chapter has been published as:

Bellekom, B., Hackett, T. D., & Lewis, O. T. (2021). A Network Perspective on the Vectoring of Human Disease. *Trends in Parasitology*, 37(5), 391–400. https://doi.org/10.1016/j.pt.2020.12.001

The full published version is found in Appendix 3.

Ben Bellekom*, Talya D. Hackett and Owen T. Lewis

Department of Zoology, 11a Mansfield Road, Oxford OX1 3SZ, UK

*Correspondence: ben.bellekom@Biology.ox.ac.uk (B. Bellekom)

Keywords: biting Diptera; interaction network; vector-borne disease; zoonoses

2.1 Abstract

Blood sucking insects are vectors of diseases with biting Diptera alone causing an estimated 700,000 human deaths each year. Insect vectors also bite non-human hosts, linking them into host-biting networks. While the major vectors of prominent diseases such as malaria, yellow fever, dengue and Zika are intensively studied, there has been limited focus on the wider interactions of biting Diptera with non-human hosts. Drawing on network analysis and visualisation approaches from food web ecology, we discuss the value of a network perspective for understanding host-insect-disease interactions. Potential applications include highlighting pathways of disease transmission, highlighting reservoirs of infection, and identifying emerging and previously unrecognised vectors.

Highlights

Insect vectors of human diseases are embedded within complex ecological communities. Their interactions with other species, particularly alternative hosts, have important consequences for disease transmission and dynamics.

Modern molecular methods of blood meal analysis provide the opportunity to document and quantify interactions between known or potential vectors and their host species accurately, and at an unprecedented scale.

Ecological interaction networks analysis and visualisation provides a useful tool to highlight the interconnectedness of these communities, the degree to which humans are embedded within them, and to identify and predict pathways of disease transmission.

2.2 The emergence and spread of zoonotic disease

More than 60% of known human infectious diseases and 75% of emerging infectious diseases are zoonotic, spreading from animals to humans [1,2]. Recent epidemics, including Zika fever, chikungunya and dengue fever, highlight how novel or apparently stable zoonoses can emerge, creating major challenges to human health and economic wellbeing [3]. Blood feeding arthropods have been implicated in the spread of at least a quarter of all such outbreaks during the last century [4].

Over the coming century, increased human population densities, greater movement of both people and insects, and the expansion of agriculture and habitation into new areas are likely to favour the emergence, spread and success of novel and existing insect-vectored zoonoses [5]. Here we argue that a multi-species analysis of vector-host interactions will help us understand disease epidemiology, minimise risk, and inform control strategies.

Moving beyond pairwise interactions between species

Most research on insect-vectored diseases has focused on the epidemic and **epizootic** (see Glossary) role of known vector species and their pairwise interactions with their hosts, particularly humans [3,6,7]. However, insect vectors are embedded within complex ecological communities [8] and their interactions with a wider set of species are likely to have important consequences for disease transmission. Ecological interaction network modelling [9] is a useful tool for understanding the interdependencies within these sets of interacting species, and for predicting the dynamics of the component species. Networks provide a visual and mathematical representation of interactions between **nodes** (e.g. species) connected by **edges** (e.g. feeding interactions)[10]. Networks may be unweighted, where edges are either present or absent, or weighted by the strength or frequency of interactions [11,12]. Both mutualistic interactions (such as interactions between plants and their pollinators) and antagonistic interactions (such as those between predators and their prey) can be represented in this way [13,14].

Using biting Diptera, which currently cause approximately 700,000 deaths annually [15], as a case study, we focus on networks of antagonistic interactions between biting Diptera and the

host animals from which they obtain blood meals, referred to here as ‘host-biting networks’. Network ecology can reveal how community structure and function change in response to a range of perturbations, with widespread applications in conservation and agriculture [9]. We suggest that increased application of network approaches in disease vector research will provide valuable insights for understanding vector-borne diseases.

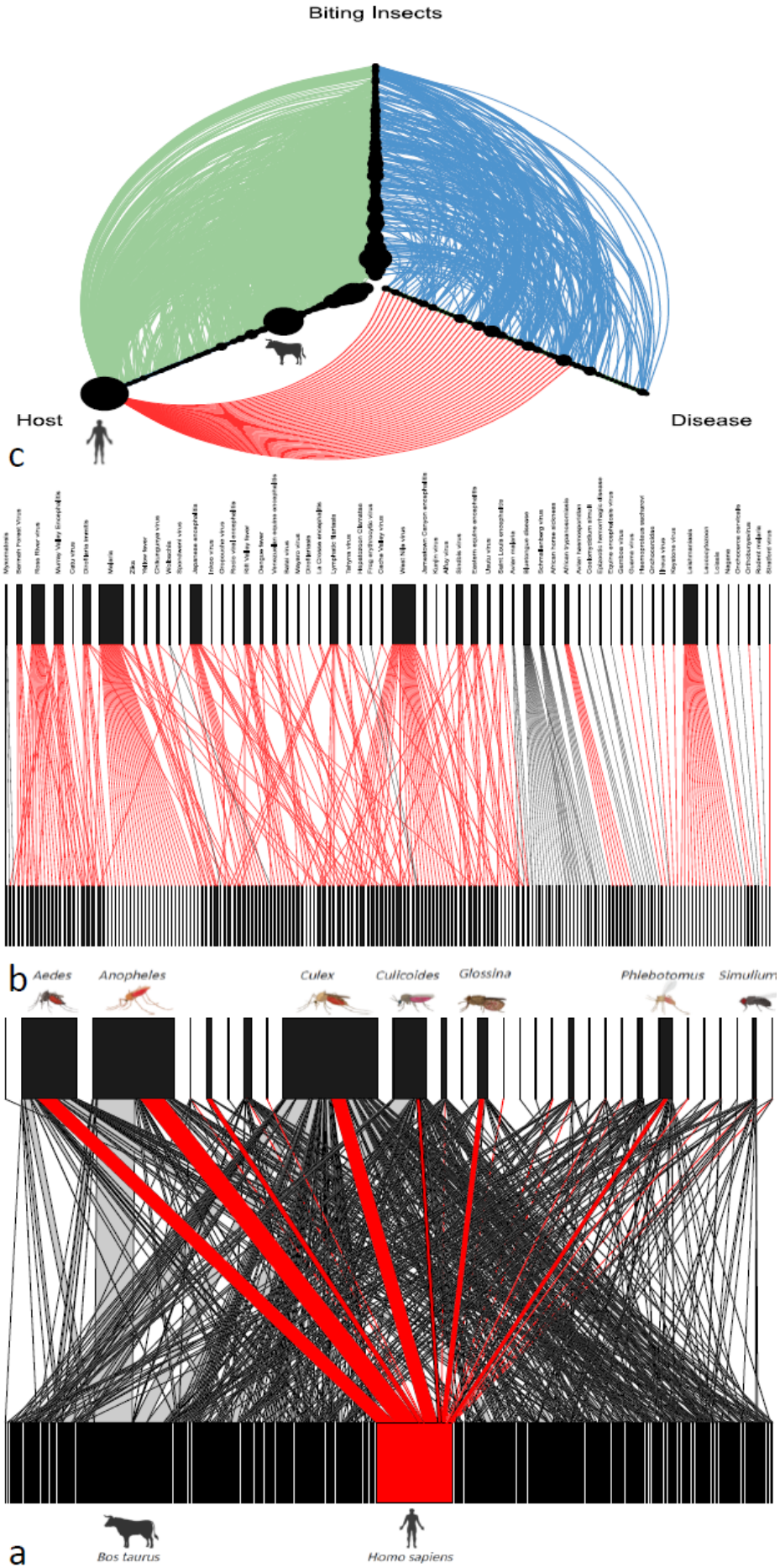

Biting Insects
Host
Disease
c
b
Aedes
Anopheles
Culex
Culicoides
Glossina
Phlebotomus
Simulium
a
Bos taurus
Homo sapiens

Figure 1. Host-biting networks and their associated disease transmission pathways. (Bottom) Global bipartite [91] network of host-biting interactions, highlighting how humans (red node and edges) are embedded into the community. Node and edge width is proportional to its frequency. Nodes are resolved to species for hosts and genus for biting-insects. (Middle) Bipartite network linking disease (upper) to vectors (lower). Data on disease-vector interactions were extracted from Vectorbase [19]. Red links represent diseases that are transmissible to humans. Mosquito genera (*Aedes, Culex* and *Ochlerotatus*) vector the largest number of human disease (23, 18 and 15 respectively). (Top) Hive plot [92] showing how hosts (left axis), biting insects (vertical axis) and their respective diseases (right axis), are globally interconnected. Interactions between hosts and biting Diptera are represented in green, vector-disease links in blue, and diseases that are transmissible to humans in red. Nodes along each axis scale with frequency. Taxon silhouettes are from BioRender.com. Data were compiled from published literature on blood meals of biting Diptera. To maximise the amount of relevant data, source publications were identified using ad hoc searches in Web of Science and Google Scholar with search terms including 'blood meal, 'mosquito ecology', 'metabarcoding', 'vector-host', and genus, family and subfamily names for biting insect taxa. Further, we scrutinised the citations within these publications to identify additional relevant literature. Data were extracted from peer-reviewed articles that conducted molecular blood meal analysis for >1 biting Diptera species. To maximise our data set, we included data generated using a variety of blood meal analysis methods including diagnostic PCR approaches using species-specific or general vertebrate primers, ELISA, gel diffusion immunoassays, precipitin tests, microsatellite analyses, and monoclonal antibodies.

The global network of biting interactions

Whilst host species may not directly interact in our network analysis, host species are connected indirectly through their shared parasites [16]: multiple parasite species can share the same host species [17], and individual parasite species can interact with multiple host species [18]. To illustrate how hosts and parasites are interconnected on a global level, we constructed networks of biting Diptera-host interactions using data compiled from published literature (Figure 1 Top& Bottom. In total, these data comprise 26,049 biting Diptera-host

interactions (as informed by blood meals) from 67 publications involving fieldwork in 32 countries. The resulting network has 263 biting Diptera species, 244 host species, and 1,331 links. Biting Diptera interact with host species from five classes: Amphibia, Insecta, Aves, Mammalia, and Reptilia. Cattle (*Bos taurus*) are the most commonly documented hosts, accounting for 6,902 (26.6%) of all interactions. Host species are highly interconnected through shared vector species with an unweighted network **connectance** of 0.02, meaning that 2% of all globally potential host-insect interactions have been observed. In reality, many potential links will not be possible because of geographic or temporal separation of hosts and biting Diptera, so the true connectance value will be higher. The average unweighted **degree centrality** (the number of links a species has in a network) for the 263 insect species is 5 (range: 1 to 86) host species. A higher degree centrality indicates that an insect may bite by a larger number of host species.

While the blood meal data summarised in Figure 1 do not document pathogens within blood meals, nor vector competence in transmitting them, we inferred disease vectoring potential from VectorBase, a database of insect vectors of different diseases [19]. Insects in our network are known to transmit 55 diseases, 39 of which are transmissible to humans (Figure 1 Top& Middle). Within our network, 190 insect species are known vectors of these 55 diseases, of which 154 are capable of transmission to humans; most of these (106) were indeed recorded as biting humans in the data compilation.

How embedded are humans?

There are clear epidemiological benefits to characterising the degree to which humans are embedded within networks of host-biting interactions. Increasing interaction of humans with livestock and wildlife populations is likely to facilitate a rise in the transmission of emerging diseases [20]. Examples of such vector-driven **disease spillover** events include Lyme disease and West Nile Virus [21]. Unrestricted and repeated interactions between human hosts and **competent vectors** can allow diseases to reach epidemic levels rapidly [22].

Our compilation of blood meal data reveals that humans are highly embedded in networks of blood-feeding, and interact frequently with well-connected vectors for a variety of

diseases (Figure 1Top &Bottom). After cattle, humans were the second most frequently-documented host species in the global network, accounting for 6,217 (23.9%) of the interactions involving 152 biting Diptera species. Diptera recorded as biting humans had a significantly wider variety of non-human host species (mean: 6, range 0 to 85) than those that did not (mean: 3, range 1 to 23; Mann-Whitney U = 6448.5, $p<0.05$). The number of insect-human interactions reported is likely to be skewed by the anthropocentric focus of many of the studies from which data were extracted. However, even mosquito species typically considered anthropophilic (e.g. *Anopheles gambiae* and *Aedes aegypti*) [23–27] interact with a range of non-human hosts (9 and 10 respectively) including, birds, reptiles and non-human mammals.

While biting Diptera can show specific feeding preferences [28,29], many species will feed opportunistically, with consequences for disease control measures [28,30]. Indeed, the five biting Diptera species with the most recorded interactions (*Anopheles arabiensis* n=2,280, 8.8%, *Culex annulirostris* n=1,986, 7.6%, *Cx*. *pipiens* n=1,698, 6.52, *Aedes camptorhynchus* n=1,159, 4.5%, and *Ae. aegypti* n=1039, 4%) had a range of non-human hosts (4, 25, 85, 14 and 10, respectively) which comprised 84.5%, 95.5%, 86.5%, 99%, and 14.8% of their interactions, respectively.

Interpolation (rarefaction) and extrapolation (prediction) curves [31,32] indicate how comprehensively the host ranges of these individual species have been documented (Figure 2A). Observed host ranges were underestimated relative to those predicted by extrapolation by 4, 61 and 3 species for *Cx. annulirostris, Cx*. *pipiens* and *Ae. aegypti,* respectively. This suggests that the known range of host use for even the most commonly sampled species is still not complete, and highlights the extent to which generalist species such as *Cx. pipiens* might interact with previously unsuspected hosts. No such differences between documented and predicted host species richness were found for *An. arabiensis* and *Ae. camptorhynchus*. The large number of undiscovered hosts for *Cx*. *pipiens* likely reflects their preference for feeding on birds [33], a particularly species-rich vertebrate class.

Rarefaction at the community level indicates that biting Diptera and host species are relatively well- resolved, with species accumulation curves in each case approaching an asymptote (Figure 2B). In contrast, documented interactions between vectors and hosts are much less complete. We therefore have an incomplete understanding of the full range of

vector-host interactions, and further sampling will be necessary to reveal these cryptic interactions and to increase the overall resolution of host-biting networks.

The network shown in Figure 1 Top & Bottom contains interactions from a global dataset. In reality, geographically-separated hosts are unlikely to co-occur spatially to the extent that a single Diptera individual can interact with them. However, some hosts (e.g. migratory birds) and humans may be highly mobile. Therefore, it is feasible that hosts that never co-occur could be linked indirectly by a single mobile human individual interacting with a variety of hosts and vectors in different locations. Inevitably, human movement, agricultural encroachment, illegal trafficking of hosts via the bushmeat and wildlife trade, and accidental introductions of invasive species result in increased opportunities for animal-human interactions and may facilitate zoonotic disease transmission [34].

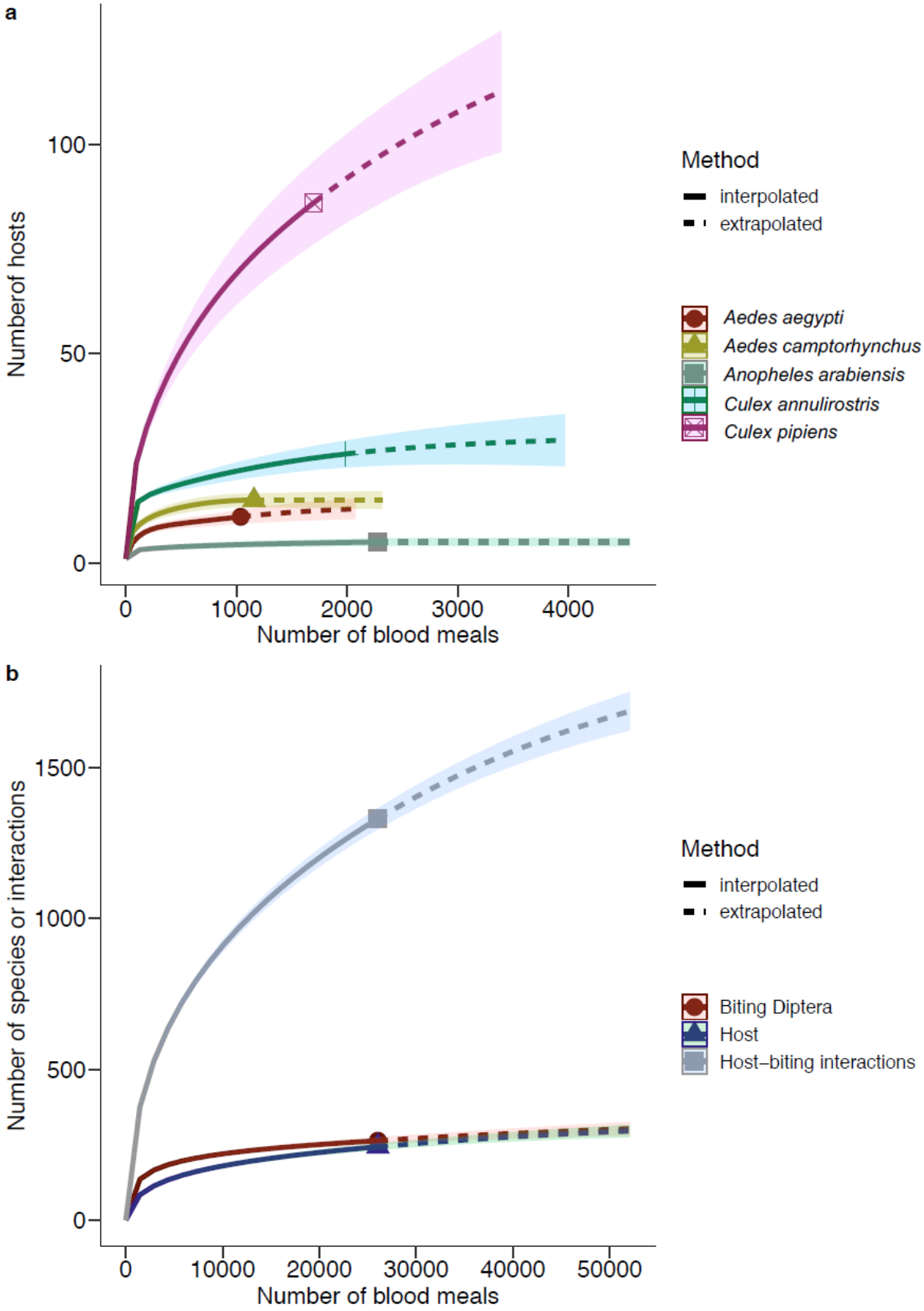

a
Numberof hosts
100
50
0
0
1000
2000
3000
4000
Number of blood meals
Method
interpolated
extrapolated
Aedes aegypti
Aedes camptorhynchus
Anopheles arabiensis
Culex annulirostris
Culex pipiens
b
Number of species or interactions
1500
1000
500
0
0
10000
20000
30000
40000
50000
Number of blood meals
Method
interpolated
extrapolated
Biting Diptera
Host
Host–biting interactions

Figure 2. Smoothed accumulation and extrapolation curves to assess the completeness of sampling. (a) Number of host species recorded for the five most frequent biting Diptera species in the compiled dataset, as a function of sampling effort (the number of blood meals screened). (b) Total numbers of hosts, biting insects and interactions recorded in the whole dataset, as a function of sampling effort (the number of blood meals screened). These curves suggest that the level of sampling achieved across studies is sufficient to record most biting Diptera (red) and host (blue) species, but that many interactions among these species remain undocumented. Curves were calculated and plotted using the R package iNEXT [93]. Shaded areas show 95% confidence intervals for the predicted values.

Using networks to identify emerging vectors and disease

As human populations expand into new areas, there is increased opportunity for people to interact directly with novel potential disease vectors, and indirectly, via shared biting insects, with other host species [35]. For example, deforestation for commodity crop trade and road development in the Peruvian Amazon has increased the frequency of human interaction with the malaria vector *Anopheles darlingi* 278 times relative to intact forested habitat, leading to an increase in malaria cases [36]. The integration of novel species into interaction networks is widely studied in community ecology and conservation biology, where it has broad implications for understanding the impact of invasive species. The influence of a new species depends on the degree to which it is a generalist or specialist and might also be conditioned on the diversity of available hosts [28], with generalists expected to have greater impacts on network structure and dynamics [37,38]. Quantitative network data, combined with information on biting behaviour and vector competence, will be useful to identify novel or encroaching vectors, and their likely disease transmission potential.

The emergence of novel diseases is hard to predict [39], but targeted vector surveillance at the interface of humans and wildlife, informed by data on species' biting preferences, has the potential to help predict and limit the emergence and spread of zoonotic disease. For example, following the rapid spread of *Culex* mosquito vectored West Nile Virus in the United States, intensive surveillance of mosquitoes and their hosts allowed outbreaks to be anticipated and effective vector control measures to be implemented [40]. Regular monitoring for pathogens of people, domestic animals [41] and wildlife [42] could provide an 'early warning system' for disease emergence. Such an approach could be complemented

by mining data associated with molecular analysis of blood meals, focusing on likely vector species identified using weighted network data. Key species to identify in this way will include those that are highly connected and which interact frequently with disease-susceptible hosts (e.g. those with high degree centrality)[43]. In the longer term, data from such a surveillance programme could provide valuable general insights into spatial patterns of disease spillover, temporal variations in vector-host contact patterns, and the relative importance of local and long-range transmission events [44].

Interaction networks could also help highlight unexpected or important hosts, as well as situations where vector populations have different host use in different geographic locations [45]. Hosts that are bitten by many biting Diptera species are candidates for the long-term persistence of disease and infection of sympatric **reservoir hosts** [46]. Identifying these species could therefore inform preventative measures, paralleling the situation where domestic dogs are vaccinated to reduce rabies transmission to humans [47]. Just as for novel vectors, the impact of novel hosts could be determined though network analysis. For example, the introduction of domesticated animal hosts into tick-vertebrate networks increased disease transmission throughout the community, as domesticated hosts acted as super-spreaders, linking nodes that would not otherwise interact [48]. Better understanding of vector-host interactions could help prevent resurgence in infection from an unknown reservoir. Localised elimination of malaria in Guiana was followed by an increase in rice farming and subsequent displacement of cattle; as a result the malarial vector *Anopheles aquasalis* swapped from livestock hosts to the more abundant humans, leading to a localised outbreak of malaria [49]. Marked consequences for transmission are expected if vector species or disease-vulnerable host species occupying a central position within a network are added or removed [12,17]. Moreover, certain vectors may act as a **hub species**, with many strong connections to other species both in the same and separated habitats, or **connector species**, that couple spatially, or temporally, distinct communities, potentially facilitating between-habitat disease transmission [43,50–53]. Such effects could be predicted if well-resolved weighted network data were available: network ecologists increasingly use such data to inform models predicting the occurrence and consequences of **re-wiring** events (where network interactions reconfigure as species are added or lost) [53,54].

Bridge vectors between humans and other animals could facilitate sustained transmission from a reservoir host to humans [55],[56]. Biting insect species that routinely interact with multiple well-connected hosts, particularly where individuals contain blood from multiple susceptible host species, are candidate bridge vectors [57]. The incidence and identity of mixed blood meals, where molecular analysis indicates a single insect contains blood meals from multiple host species, is highly relevant for predicting spillover events. However, mixed blood meals are rarely reported, either because the molecular methods used are unsuitable to detect them (**Box 1**), or because they are genuinely rare: only 10 (15%) of the papers included in our network compilation report mixed blood meals. When reported, the frequency of mixed blood meals varied, averaging 13.2% of total blood meals (range: 1% to 34.4%). Using molecular methods that detect mixed blood meals should be a priority (**Box 2**), and could ultimately allow network approaches to be applied on an individual basis, where nodes represent host individuals or insects, and edges represent feeding events that link them[58,59]. Modelling of such networks may identify those individuals that have interacted with multiple hosts and highlights potential disease transmission events.

Box 1. Integrating molecular approaches to determine interactions and pathways of disease transmission

Documenting potential and realised pathways of insect-vectored disease transmission has, until recently, relied on expensive and labour intensive techniques that are error-prone and often require prior identification of candidate hosts and host specific anti-sera [65–67]. Recent developments in molecular methods can overcome these limitations, generating data on feeding interactions more rapidly and comprehensively, and enabling the identification of multiple hosts from mixed blood meals [68].

Previously, species interactions were commonly identified from antibodies present in insect blood meals using enzyme-linked immunosorbent assays (ELISA) [69,70] and monoclonal antibodies [71]. These methods have now largely been superseded by DNA barcoding, diagnostic PCR (polymerase chain reaction) [69,72–74] and, most recently, metabarcoding. In metabarcoding, group-specific PCR primers are coupled with high-throughput sequencing to identify a broad range of taxa within bulk mixtures [72,75].

While the primary motivation behind the shift to metabarcoding has been to identify hosts of individual vectors, this approach also allows the construction of quantitative networks of host-biting interactions at a scale and resolution that was previously impossible. As the cost of metabarcoding continues to decrease, molecular analysis of blood meals offers excellent opportunities to identify and quantify previously cryptic species interactions. In future, shotgun sequencing holds considerable potential for blood meal analyses, with the advantage that it limits the negative effects of primer bias [62], although uptake of this approach is currently limited by high costs and a lack of genome data for many species.

Finally, quantitative (weighted) networks can highlight potential emerging vectors which occur at low abundance, but which might increase rapidly following a disturbance or environmental change such as habitat degradation or climate change [22], or competitive release following reduction in the abundance of another species [60]. In our network, 48.7% (n=129) of insect species had 10 or fewer blood meal records. Of these species, 61 (47.3%) are known vectors of human diseases, and 32 (24.8%) were recorded interacting with humans. The sparse data for these biting insect species is likely due to a combination of their rarity and the use of sampling methods or locations that underestimate their abundance (Box 2). Nonetheless, they may represent unrecognised sources of cross-species disease transmission, currently or in the future. Identification of host use and interaction frequency of understudied vectors is therefore of particular importance for predicting current and future disease transmission risk, and for informing control measures to minimise potential outbreaks [61].

Box 2. Improving the resolution and utility of host-biting networks

The network in Figure 1 represents an *ad hoc* compilation of published data. While it reveals how host-biting networks are interconnected and their potential applications in disease biology, a full realisation of the sorts of applications we suggest in the main text will require a more targeted approach to generate unbiased data on interactions at a local or regional scale.

Variation in trapping success for different insect taxa [76] will result in their uneven representation within networks. In particular, collection methods that use live hosts, such as human landing catches [77], will be heavily biased [78]. Widely-used and efficient insect sampling methods include USA Centre for Disease Control (CDC) miniature light traps and BGS traps with CO_2 bait, which capture a wide range of mosquitos [79], sand flies [80] and biting midges [81,82]. We recommend using a combination of sampling methods to maximise representation of the biting insect fauna. Biases will inevitably remain, however; for example, light traps will under-sample insects with diurnal and crepuscular activity. "Non-attractive" methods, such as malaise traps or suction sampling, will be a valuable supplement to attractive traps, potentially documenting a wider range of biting insects and providing less-biased data on their relative abundances [83]. A common approach to construct weighted interaction is to combine randomised sampling (to document abundances in an unbiased way) with more targeted sampling of component species (to generate data on interaction frequencies efficiently)[84,85].

The spatial location of sampling will influence both the species composition of biting insects, and the set of potential hosts with which they are interacting. For example, samples taken within or close to human habitation will over-represent both anthropophilic insects and human blood meals [27]. Humans are likely to be exposed to a much wider variety of biting insects (and will be connected indirectly to a much wider variety of their alternative hosts) when they spend time outside the domestic setting, for example, for farming or recreation [86,87]. Non-urban settings are also prime contexts for the emergence of novel and bridge vectors as well as novel pathogens, and should therefore be a priority for network-focused sampling.

The choice of **molecular analysis methods** – especially the use of species-specific techniques and primer sets - can lead to incomplete or biased blood meal identification [88]. Metabarcoding methods are preferable, since they are likely to document a wider set of potential hosts including many that may not have been recorded previously. Universal primer sets that are validated using *in silico* analysis, combined with high-throughput sequencing can accurately and rigorously characterise the species composition of blood meals and generate unbiased interaction data [74].

Limited **mixed blood meal** reporting in the literature may reflect methods of blood meal analysis that are not appropriate for their detection and identification [89,90]. Wherever possible, mixed blood meals should be screened for using metabarcoding methods, and reported, facilitating the identification of bridge vectors and informing disease transmission risks [74].

Concluding remarks

Available blood meal interaction data, compiled from a set of global studies, reveal the global pattern of insect-host community structure, and demonstrate the extent to which humans share potential vectors with other vertebrates. Many of the benefits of a network perspective that we have discussed will however, be best realised with dedicated biting insect-host community-level sampling on a local or regional scale; and we encourage researchers to collect such data in a way that minimises biases to facilitate network-based analyses (**Box 2**). The resulting data will allow researchers to address a wide variety of unanswered questions relevant to understanding and predicting disease transmission and dynamics (see **Outstanding Questions**), both for diseases of humans and more generally.

In the long-term, increasingly reliable sequencing techniques such as shotgun metagenomics [62] will offer unique opportunities to gather large-scale, long-term network data sets. Ultimately, these could allow global-scale biomonitoring to detect changes in community structure and identify emerging disease threats [63][64].

2.3 Outstanding Questions

- What factors predict host-biting interactions (e.g. phylogeny, traits, host abundance, abiotic factors), and can they be used to predict interactions for under-studied vectors?
- What influences the apparent host-specialisation of biting insects, and to what extent is this determined by encounter rates with hosts?
- Can network data be used to predict changes in biting frequency (and disease transmission) when the densities of hosts that share potential vectors change?

- How common are mixed blood meals, and is their frequency correlated with the degree centrality of individual biting insect species?
- To what extent does the frequency that a pair of hosts is detected in mixed blood meals match the biting frequencies on those hosts inferred from interaction networks?
- How does the structure of host-biting networks change seasonally, geographically, and across environmental gradients?
- Are there biting insect species that consistently fill the role of hub and connector nodes across different geographical locations?

2.4 Acknowledgements

We thank members of the Community Ecology Research Group Oxford for helpful discussions. We thank John Connolly, John Mumford, Silke Fuchs, and the three anonymous reviewers for providing their insightful comments and suggestions. This work was supported by a NERC studentship to Ben Bellekom through the Oxford University Environmental Research Doctoral Training Programme (NE/L002612/1). Owen Lewis and Talya Hackett were supported by supported by the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation, as part of the Target Malaria consortium.

2.5 References

1 White, R.J. and Razgour, O. (2020) Emerging zoonotic diseases originating in mammals: a systematic review of effects of anthropogenic land-use change. *Mamm. Rev.* DOI: 10.1111/mam.12201

2 Jones, K.E. *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451, 990–993

3 Braack, L. *et al.* (2018) Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasit. Vectors* 11, 29

4 Alcaide, M. *et al.* (2009) Disentangling vector-borne transmission networks: a universal DNA barcoding method to identify vertebrate hosts from arthropod bloodmeals. *PLoS One* 4, e7092

5 Gratz, N.G. (1999) Emerging and resurging vector-borne diseases. *Annu. Rev. Entomol.* 44, 51–75

6 Molaei, G. *et al.* (2013) Vector-host interactions and epizootiology of eastern equine encephalitis virus in massachusetts. *Vector-Borne Zoonotic Dis.* 13, 312–323

7 J. Kweka, E. *et al.* (2019) Ecology of Aedes Mosquitoes, the Major Vectors of Arboviruses in Human Population. In *Dengue Fever - a Resilient Threat in the Face of Innovation* IntechOpen

8 Ferguson, H.M. *et al.* (2010) Ecology: A Prerequisite for Malaria Elimination and Eradication. *PLoS Med.* 7, e1000303

9 Kaiser-Bunbury, C.N. and Blüthgen, N. (2015) Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants* 7, plv076

10 Proulx, S. *et al.* (2005) Network thinking in ecology and evolution. *Trends Ecol. Evol.* 20, 345–353

11 Blüthgen, N. (2010) Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic Appl. Ecol.* 11, 185–195

12 Poulin, R. (2010) Network analysis shining light on parasite ecology and diversity.

Trends Parasitol. 26, 492–498

13 Ings, T.C. *et al.* (2009) Review: Ecological networks - beyond food webs. *J. Anim. Ecol.* 78, 253–269

14 Bendel, C.R. *et al.* (2019) Plant–pollinator networks in grassland working landscapes reveal seasonal shifts in network structure and composition. *Ecosphere* 10, e02569

15 Müller, R. *et al.* (2019) Vector-Borne Diseases. In *Biodiversity and Health in the Face of Climate Change* pp. 67–90, Springer International Publishing

16 Espinaze, M.P.A. *et al.* (2018) Domestic mammals facilitate tick-borne pathogen transmission networks in South African wildlife. *Biol. Conserv.* 221, 228–236

17 Godfrey, S.S. *et al.* (2009) Network structure and parasite transmission in a group living lizard, the gidgee skink, Egernia stokesii. *Behav. Ecol. Sociobiol.* 63, 1045–1056

18 Pilosof, S. *et al.* (2015) Potential Parasite Transmission in Multi-Host Networks Based on Parasite Sharing. *PLoS One* 10, e0117909

19 Giraldo-Calderón, G.I. *et al.* (2015) VectorBase: An updated Bioinformatics Resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res.* 43, D707–D713

20 Grubaugh, N.D. *et al.* (2015) Xenosurveillance: A Novel Mosquito-Based Approach for Examining the Human-Pathogen Landscape. *PLoS Negl. Trop. Dis.* 9, e0003628

21 Kulkarni, M.A. *et al.* (2015) Major emerging vector-borne zoonotic diseases of public health importance in Canada. *Emerg. Microbes Infect.* 4, e33

22 Weaver, S.C. and Reisen, W.K. (2010) Present and future arboviral threats. , *Antiviral Research*, 85. , Elsevier, 328–345

23 Orsborne, J. *et al.* (2018) Using the human blood index to investigate host biting plasticity: a systematic review and meta-regression of the three major African malaria vectors. *Malar. J.* 17, 479

24 Pappa, V. *et al.* (2011) Estimation of the Human Blood Index in malaria mosquito vectors in Equatorial Guinea after indoor antivector interventions. *Am. J. Trop. Med.*

Hyg. 84, 298–301

25 Santiago-Alarcon, D. *et al.* (2012) Bloodmeal analysis reveals avian plasmodium infections and broad host preferences of culicoides (diptera: Ceratopogonidae) vectors. *PLoS One* 7,

26 Carvalho, G.M.L. *et al.* (2017) Bloodmeal Identification in Field-Collected Sand Flies From Casa Branca, Brazil, Using the Cytochrome b PCR Method. *J. Med. Entomol.* 54, 1049–1054

27 Orsborne, J. *et al.* (2020) Evidence of extrinsic factors dominating intrinsic blood host preferences of major African malaria vectors. *Sci. Rep.* 10, 741

28 Chaves, L.F. *et al.* (2010) Blood feeding patterns of mosquitoes: Random or structured? *Front. Zool.* 7, 1–11

29 Harrington, L.C. *et al.* (2009) Why Do Female Aedes aegypti (Diptera: Culicidae) Feed Preferentially and Frequently on Human Blood? *J. Med. Entomol.* 38, 411–422

30 Lassen, S.B. *et al.* (2011) Molecular identification of bloodmeals from biting midges (Diptera: Ceratopogonidae: Culicoides Latreille) in Denmark. *Parasitol. Res.* 108, 823–829

31 Vázquez, D.P. *et al.* (2009) Uniting pattern and process in plant–animal mutualistic networks: a review. *Ann. Bot.* 103, 1445–1457

32 Budka, A. *et al.* (2019) The use of rarefaction and extrapolation as methods of estimating the effects of river eutrophication on macrophyte diversity. *Biodivers. Conserv.* 28, 385–400

33 Kothera, L. *et al.* (2019) Bloodmeal, Host Selection, and Genetic Admixture Analyses of Culex pipiens Complex (Diptera: Culicidae) Mosquitoes in Chicago, IL. *J. Med. Entomol.* DOI: 10.1093/jme/tjz158

34 Johnson, C.K. *et al.* (2020) Global shifts in mammalian population trends reveal key predictors of virus spillover risk. *Proc. R. Soc. B Biol. Sci.* 287, 20192736

35 Santiago-Alarcon, D. *et al.* (2012) Bloodmeal analysis reveals avian plasmodium infections and broad host preferences of culicoides (diptera: Ceratopogonidae)

vectors. *PLoS One* 7, 31098

36 Vittor, A.Y. *et al.* (2006) The effect of deforestation on the human-biting rate of Anopheles darlingi, the primary vector of falciparum malaria in the Peruvian Amazon. *Am. J. Trop. Med. Hyg.* 74, 3–11

37 Russo, L. *et al.* (2014) Patterns of introduced species interactions affect multiple aspects of network structure in plant–pollinator communities. *Ecology* 95, 2953–2963

38 Henneman, M.L. and Memmott, J. (2001) Infiltration of a hawaiian community by introduced biological control agents. *Science (80-.).* 293, 1314–1316

39 Murphy, F.A. (1998) , Emerging zoonoses. , in *Emerging Infectious Diseases*, 4, pp. 429–435

40 Healy, J.M. *et al.* (2015) Comparison of the efficiency and cost of west nile virus surveillance methods in California. *Vector-Borne Zoonotic Dis.* 15, 147–155

41 Wolfe, N.D. *et al.* (2007) Origins of major human infectious diseases. *Nature* 447, 279–283

42 Morse, S.S. *et al.* (2012) Prediction and prevention of the next pandemic zoonosis. , *The Lancet*, 380. , Elsevier, 1956–1965

43 Graham, S.P. *et al.* (2009) Nestedness of ectoparasite-vertebrate host networks. *PLoS One* 4, 7873

44 Craft, M. and Caillaud, D. (2011) Network models: an underutilized tool in wildlife epidemiology? *Interdiscip Perspect Infect Dis* 2011,

45 Tirados, I. *et al.* (2006) Blood-feeding behaviour of the malarial mosquito Anopheles arabiensis: Implications for vector control. *Med. Vet. Entomol.* 20, 425–437

46 Streicker, D.G. *et al.* (2013) Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecol. Lett.* 16, 975–984

47 Lembo, T. *et al.* (2010) The Feasibility of Canine Rabies Elimination in Africa: Dispelling Doubts with Data. *PLoS Negl. Trop. Dis.* 4, e626

48 Estrada-Peña, A. *et al.* (2015) Interactions between tick and transmitted pathogens

evolved to minimise competition through nested and coherent networks. *Sci. Rep.* 5, 10361

49 Giglioli, G. (1963) Ecological change as a factor in renewed malaria transmission in an eradicated area. a localized outbreak of a. aquasalis-transmitted malaria on the demerara river estuary, british guiana, in the fifteenth year of *A. darlingi* and malaria eradication. *Bull. World Health Organ.* 29, 131–45

50 Olesen, J.M. *et al.* (2007) The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19891–19896

51 Guimerà, R. and Amaral, L.A.N. (2005) Cartography of complex networks: Modules and universal roles. *J. Stat. Mech. Theory Exp.* 2005, 1–13

52 Hackett, T.D. *et al.* (2019) Reshaping our understanding of species' roles in landscape-scale networks. *Ecol. Lett.* 22, 1367–1377

53 Bartley, T.J. *et al.* (2019) Food web rewiring in a changing world. , *Nature Ecology and Evolution*, 3. , Nature Publishing Group, 345–354

54 CaraDonna, P.J. *et al.* (2017) Interaction rewiring and the rapid turnover of plant-pollinator networks. *Ecol. Lett.* 20, 385–394

55 Shahhosseini, N. *et al.* (2018) Host-feeding patterns of Culex mosquitoes in Iran. *Parasites and Vectors* 11, 669

56 Mayi, M.P.A. *et al.* (2020) Habitat and Seasonality Affect Mosquito Community Composition in the West Region of Cameroon. *Insects* 11, 312

57 Kang, D.S. *et al.* (2019) Identification of Host Blood Meals of Mosquitoes (Diptera: Culicidae) Collected at the Aripo Savannas Scientific Reserve in Trinidad, West Indies. *J. Med. Entomol.* DOI: 10.1093/jme/tjz113

58 White, L.A. *et al.* (2017) Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biol. Rev.* 92, 389–409

59 Godfrey, S.S. (2013) Networks and the ecology of parasite transmission: A framework for wildlife parasitology. , *International Journal for Parasitology: Parasites and Wildlife*, 2. , Elsevier, 235–245

60 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. *Bull. Entomol. Res.* 51, 243

61 Faria, N.R. *et al.* (2017) Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* 546, 406–410

62 Boers, S.A. *et al.* (2019) Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. , *European Journal of Clinical Microbiology and Infectious Diseases*, 38. , Springer Verlag, 1059–1070

63 Bohan, D.A. *et al.* (2017) Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. , *Trends in Ecology and Evolution*, 32., 477–487

64 Olivier, P. *et al.* (2019) Exploring the temporal variability of a food web using long-term biomonitoring data. *Ecography (Cop.).* 42, 2107–2121

65 Evans, D.M. *et al.* (2016) Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Funct. Ecol.* 30, 1904–1916

66 Miyake, T. *et al.* (2019) Bloodmeal host identification with inferences to feeding habits of a fish-fed mosquito, Aedes baisasi. *Sci. Rep.* 9, 4002

67 Townzen, J.S. *et al.* (2008) Identification of mosquito bloodmeals using mitochondrial cytochrome oxidase subunit I and cytochrome b gene sequences. *Med. Vet. Entomol.* 22, 386–393

68 Clare, E.L. *et al.* (2018) Approaches to integrating genetic data into ecological networks. *Mol. Ecol.* DOI: 10.1111/mec.14941

69 Jansen, C.C. *et al.* (2009) Blood sources of mosquitoes collected from urban and peri-urban environments in eastern Australia with species-specific molecular analysis of avian blood meals. *Am. J. Trop. Med. Hyg.* 81, 849–857

70 Apperson, C.S. *et al.* (2002) Host-Feeding Habits of *Culex* and Other Mosquitoes (Diptera: Culicidae) in the Borough of Queens in New York City, with Characters and

Techniques for Identification of *Culex* Mosquitoes. *J. Med. Entomol.* 39, 777–785

71 Svobodová, M. *et al.* (2003) Short report: Distribution and feeding preference of the sand flies Phlebotomus sergenti and P. papatasi in a cutaneous leishmaniasis focus in Sanliurfa, Turkey. *Am. J. Trop. Med. Hyg.* 68, 6–9

72 Pompanon, F. *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21, 1931–1950

73 Omondi, D. *et al.* (2015) Unraveling Host-Vector-Arbovirus Interactions by Two-Gene High Resolution Melting Mosquito Bloodmeal Analysis in a Kenyan Wildlife-Livestock Interface. *PLoS One* 10, e0134375

74 Logue, K. *et al.* (2016) Unbiased Characterization of Anopheles Mosquito Blood Meals by Targeted High-Throughput Sequencing. *PLoS Negl. Trop. Dis.* 10, e0004512

75 Zinger, L. *et al.* (2019) DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862

76 Hesson, J.C. *et al.* (2015) Trapping biases of *Culex torrentium* and *Culex pipiens* revealed by comparison of captures in CDC traps, ovitraps, and gravid traps. *J. Vector Ecol.* 40, 158–163

77 Maliti, D. V. *et al.* (2015) Development and evaluation of mosquito-electrocuting traps as alternatives to the human landing catch technique for sampling host-seeking malaria vectors. *Malar. J.* 14, 502

78 Ponlawat, A. *et al.* (2017) Field evaluation of two commercial mosquito traps baited with different attractants and colored lights for malaria vector surveillance in Thailand. *Parasit. Vectors* 10, 378

79 Gorsich, E.E. *et al.* (2019) A comparative assessment of adult mosquito trapping methods to estimate spatial patterns of abundance and community composition in southern Africa. *bioRxiv* DOI: 10.1101/633552

80 Aklilu, E. *et al.* (2017) Comparative study on the nocturnal activity of phlebotomine sand flies in a highland and lowland foci of visceral leishmaniasis in north-western Ethiopia with special reference to Phlebotomus orientalis. *Parasit. Vectors* 10, 393

81 Hopken, M.W. *et al.* (2017) Picky eaters are rare: DNA-based blood meal analysis of Culicoides (Diptera: Ceratopogonidae) species from the United States. *Parasit. Vectors* 10, 169

82 Hoel, D.F. *et al.* (2010) Efficacy of Commercial Mosquito Traps in Capturing Phlebotomine Sand Flies (Diptera: Psychodidae) in Egypt. *J. Med. Entomol.* 47, 1179–1184

83 Vazquez, D.P. *et al.* (2005) Species abundance and the distribution of specialization in host-parasite interaction networks. *J. Anim. Ecol.* 74, 946–955

84 Downey, H. *et al.* (2020) Assessing the potential for indirect interactions between tropical tree species via shared insect seed predators. *Biotropica* 52, 509–520

85 Lewis, O.T. *et al.* (2002) Structure of a diverse tropical forest insect–parasitoid community. *J. Anim. Ecol.* 71, 855–873

86 Ferraguti, M. *et al.* (2016) Effects of landscape anthropization on mosquito community composition and abundance. *Sci. Rep.* 6,

87 Thongsripong, P. *et al.* (2013) Mosquito Vector Diversity across Habitats in Central Thailand Endemic for Dengue and Other Arthropod-Borne Diseases. *PLoS Negl. Trop. Dis.* 7, e2507

88 Hannon, E.R. *et al.* (2019) Bloodmeal Host Selection of Culex quinquefasciatus (Diptera: Culicidae) in Las Vegas, Nevada, United States. *J. Med. Entomol.* 56, 603–608

89 Brugman, V.A. *et al.* (2017) Blood-feeding patterns of native mosquitoes and insights into their potential role as pathogen vectors in the Thames estuary region of the United Kingdom. *Parasit. Vectors* 10, 163

90 Kent, R.J. (2009) Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. , *Molecular Ecology Resources*, 9. , John Wiley & Sons, Ltd, 4–18

91 Dormann, C.F. *et al.* (2008) Introducing the bipartite Package: Analysing Ecological Networks. *R News*, 8, 8-11

92 Krzywinski, M. *et al.* (2012) Hive plots-rational approach to visualizing networks. *Brief.*

Bioinform. 13, 627–644

93 Hsieh, T.C. *et al.* (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* 7, 1451–1456

2.6 Glossary

Bridge vector: A vector that acquires the causative agent of a disease from an infected reservoir host and transmits it to a human or secondary host.

Competent vector: An organism with the capacity to maintain and transmit a particular infectious agent, or with the ability to generate a new infection when interacting with a susceptible host.

Connectance: A network metric indicating the fraction of potential links that are realised, calculated as *L/HP*, where *L* is the number of links and *H* and *P* are the number of host and parasite nodes, respectively.

Connector species: Species in a network that link groups or habitats together, typically through high mobility or varied interactions that cross structural, temporal or behavioural boundaries.

Degree centrality: The number of links a species has in a network.

Disease spillover: Transmission of a pathogen from its reservoir host into a susceptible population of an alternative host.

Edge: A relationship or connection between nodes within a network; synonymous with link.

Epizootic: A disease outbreak event within a nonhuman animal population, analogous to an epidemic within a human population.

Hub species: Species in a network with strong connections to many other species in the network.

Node: The components of a network. In the context of ecological interactions, species, populations or individuals can be represented as nodes.

Reservoir host: An epidemiologically connected population in which a pathogen may be permanently maintained, and from which an infectious agent may be transmitted to a susceptible population.

Re-wiring: The reassembling of interactions within a network, typically as occurs following a perturbation to the community.

Chapter 3: Latitudinal and anthropogenic effects on the structuring of networks linking blood-feeding flies and their vertebrate hosts

Mafi Agove

Author contributions: BB: Data analysis and writing, OTL: Supervision, review, and editing, TDH: Supervision, review, and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Publication status: **Bellekom, B**., Lewis, O. T. & Hackett, T. D. (2023). Latitudinal and anthropogenic effects on the structuring of networks linking blood-feeding flies and their vertebrate hosts. *Medical and Veterinary Entomology*, 1–8.

https://doi.org/10.1111/mve.12671

The full published version is found in Appendix 4.

Ben Bellekom[1*], Owen T. Lewis[1] and Talya D. Hackett[1]

[1]University of Oxford, Department of Biology, 11a Mansfield Road, Oxford OX1 3SZ, UK

* Corresponding author: ben.bellekom@biology.ox.ac.uk

BB: ben.bellekom@biology.ox.ac.uk

OTL: owen.lewis@biology.ox.ac.uk

TDH: talya.hackett@biology.ox.ac.uk

3.1 Abstract

Biting flies (Diptera) transmit pathogens that cause many important diseases in humans as well as domestic and wild animals. The networks of feeding interactions linking these insects to their hosts, and how they vary geographically and in response to human land-use, are currently poorly documented but are relevant to understanding cross-species disease transmission. We compiled a database of biting Diptera-host interactions from the literature to investigate how key interaction network metrics vary latitudinally and with human land-use. Interaction evenness and H2' (a measure of the degree of network specificity) did not vary significantly with latitude. Compared to near-natural habitats, interaction evenness was significantly lower in agricultural habitats, where networks were dominated by relatively few species pairs, but there was no evidence that the presence of humans and their domesticated animals within networks led to systematic shifts in network structure. We discuss the epidemiological relevance of these results and the implications for predicting and mitigating future spill-over events.

Keywords: Blood meal, Bipartite, ecological interactions, biting Diptera, habitat modification

3.2 Introduction

Across taxa, species richness consistently decreases from equatorial to polar latitudes (Lawrence & Fraser, 2020). This macroecological pattern is linked to changes in the structure and organisation of ecological communities. For example, increased species richness and associated niche partitioning at low latitudes should increase both the frequency and specificity of interspecific interactions such as mutualism, competition, and parasitism (Morris et al., 2014; Schemske et al., 2009; Willig et al., 2003). Superimposed on these large-scale patterns, anthropogenic habitat modification for agriculture and other land-uses also has marked effects on diversity, community composition, and species interactions (Meyer Steiger et al., 2016), and can reshape interactions within a community, even in the absence of biodiversity loss (Morris, 2010; Tylianakis et al., 2007).

Analyses of quantitative ecological networks documenting the interactions among species, weighted by the interaction frequency (Schleuning et al., 2012; Xing & Fayle, 2021), provide an approach for understanding how communities are structured across space and time. A growing number of network studies involving diverse taxa, locations and contexts provide opportunities for synthetic analyses investigating large-scale ecological patterns (Xing & Fayle, 2021), such as trends in specialisation with latitude (Morris et al., 2014; Schleuning et al., 2012a) and in response to land-use intensification (Tylianakis et al., 2007; Weiner et al., 2014). Few studies, however, have explored the relative influence of large-scale macroecological gradients and more localised anthropogenic impacts on the structural properties of networks of interacting species (Pellissier et al., 2018; Tylianakis et al., 2007). Here, we investigate how macro-ecological and anthropogenic factors influence the properties of networks linking biting flies (Diptera) and their vertebrate hosts. Feeding interactions between Diptera and hosts can now be routinely established using molecular analysis of insect blood meals. These interactions are of particular interest, since biting Diptera are often vectors of a wide range of pathogens including malarial parasites (Sinka et al., 2012) , Bluetongue virus (Baker et al., 2021), West Nile Virus (Kilpatrick et al., 2005) and Leishmania (Killick-Kendrick, 1999). These may infect humans, domestic and non-domestic hosts, and cause significant damage to public and health economics (Barber et al., 2010; Rushton & Lyons, 2015; Sachs & Malaney, 2002). Moreover, their transmission potential can

co-vary with their interactions across land-use gradients (Meyer Steiger et al., 2016; Müller et al., 2019; Runghen et al., 2021). We analyse biting fly-vertebrate interaction data from a wide range of latitudes and across different habitat types to explore the relative importance of latitude and land-use in structuring interaction networks.

3.3 Methods

Data compilation

Biting Diptera-host interaction data were extracted from the literature on insect blood meals, using a subset of the data compiled by Bellekom et al., (2021). To limit bias, we restricted analyses to data generated using PCR and DNA sequencing (Logue et al., 2016), and excluded studies that used sampling methods and locations inappropriate for collecting a variety of biting Diptera species and subsequent host blood meals. For example, those that used live host baited trapping methods, such as livestock and cattle-baited tents, were excluded as blood meals would be heavily biased towards the bait. Studies that provided data for a single biting Diptera species were also excluded, as these do not provide network data. We also excluded studies lacking site location details. For each remaining study, we recorded site location and classified habitats into three broad categories of anthropogenic landscape modification. Sites where cultivated land or livestock were the dominant land-use were categorised as Agricultural; those that referenced natural vegetation with limited human presence were classified as Near-natural; and those where sampling took place within, or around human habitation were classified as Village/Urban. Where sampling was carried out in more than one habitat type, separate networks were generated for each reported habitat. Where habitat could not be determined reliably from the published information (four cases), we used satellite imagery (QGIS in combination with Google Earth) to infer the habitat, using the same categories used for studies where published information on habitat type was available. For example, where satellite imagery indicated that the location of the study was associated with a human settlement, the data were assigned to the Village/Urban category.

Biting Diptera and hosts were resolved to species level, where possible. Where a single species-level identification was missing for a biting Diptera or host, nodes were simply labelled with the relevant genus or family (e.g., *Culicoides* spp.). Where a genus contained multiple unknown species-level identifications, we checked whether sympatric congeneric or confamilial species were likely to occur at the focal location using field guides and online resources (*GBIF.org (2021)*). As before, where no sympatric congeneric or confamilial species occurred, nodes were resolved to the lowest level possible (e.g., *Anas* spp.). Interactions where either the host or the biting insect could not be resolved to genus or family level (or where more than one species could occur in a single node) were removed to prevent different species being combined into the same genus-level node. In total, 18 biting Diptera and 28 host records were resolved to genus or family level and 76 hosts, 14 biting insects and 119 interactions were removed.

Data analysis

For each included study, we analysed species interaction data as weighted antagonistic bipartite networks and calculated two network metrics, interaction evenness (IE) and network specialisation (H2'), chosen for their ecological and epidemiological relevance. IE is a weighted network metric based on Shannon diversity that describes the homogeneity of interaction frequencies across all links in the network ($E_2=H_2/\ln L$, where H is Shannon diversity, and L is the number of all links) (Blüthgen et al., 2008; Kaiser-Bunbury & Blüthgen, 2015). Low IE values indicate contexts in which a small number of species and their links dominate the community (Kaiser-Bunbury & Blüthgen, 2015) . This is particularly relevant if the dominant species are vectors or susceptible hosts. H2' is a weighted network metric that quantifies the deviation of observed interaction frequencies from those expected if interaction frequencies were random (Blüthgen et al., 2006). A high degree of generalism within our network would be expected to facilitate transmission between phylogenetically dissimilar hosts (Abella-Medrano et al., 2018). IE and H2' values were calculated using the *networklevel* function in the R package *bipartite* (Dormann et al., 2008).

Variation in host and biting Diptera richness across latitudes was explored using a linear model, including the total number of blood meals analysed per network (a proxy for sampling effort) as a covariate. We explored the extent to which variations in IE and H2' could be explained by habitat type and latitude using a generalized linear model (GLM) with a Gaussian error distribution. Absolute values for latitude (i.e., removing negative signs) were used to combine data from the northern and southern hemisphere. To control for the confounding effects of network size and species richness we included species richness (S= number of resource species + number of consumer species), and log transformed matrix size (sum of all interactions within the matrix) in the GLM (Galiana et al., 2019). Most of our networks (46 out of 47) comprised Diptera from multiple genera, but drawn from a single family; each network could therefore be classified by its dominant family. In the case where multiple families were represented in a network, we classified it by the dominant family (the family with the highest number of interactions). Our models tested for the effects of family as well as the two-way interaction terms between family and latitude and between family and habitat. This allowed us to examine the influence of Diptera family on IE and H2' values and how these changed across latitude and between habitat types. Residuals were visually inspected to check model assumptions. The statistical significance of habitat type and latitude was assessed by comparing simpler models to more complex models for variation in deviance based on a chi-square distribution (Mayi et al., 2020). Post-hoc analysis was conducted to identify intra-factor significant differences, using Tukey's HSD (honestly significant difference) tests.

We used a null model to evaluate whether domesticated hosts had a measurable effect on network interaction evenness. Within each network, we simulated targeted removal of humans and domestic animals (specifically chickens, dogs, cats, goats, cattle, horses, pigs, and sheep) and compared this to removal of an equal number of randomly selected host species, both domestic and non-domestic, replicated 100 times. We then calculated the z-scores and compared the observed network metric to the distribution of the simulated values.

The dataset used is not an exhaustive set of biting Diptera-host interactions. Therefore to assess sampling completeness for each habitat we drew species interpolation and extrapolation curves for hosts, biting Diptera and interactions as a function of sampling

effort (the number of blood meals analysed) using the iNEXT package (q=0, data type = incidence frequency) (Hsieh et al., 2016).

All data handling and analysis was conducted using R (version 4.01), and the *bipartite* (Dormann et al., 2008), *iNEXT* and *tidyverse* (Wickham et al., 2019) packages. Figures were plotted using *ggplot2, iNEXT* and *bipartite.*

3.4 Results

In total, we compiled data for 9,102 biting Diptera blood meals from 45 publications involving field sites in 27 countries (Figure 1A). An aggregated global network contained 227 host species, 202 biting Diptera species, and 1,121 links (Figure 1B). Based on these data, 47 quantitative bipartite networks were constructed from 14 Agricultural (338 links), 18 Near-natural (461 links) and 15 Village/Urban sites (322 links). Almost all (97%) of our networks comprised Diptera species from a single family: Culicidae (24 cases), Ceratopogonidae (12 cases), Glossinidae (4 cases), Psychodidae (4 cases) and Simuliidae (2 cases). Sampling sites had a wide latitudinal distribution, ranging from Sweden to South Africa (Figure 1A). Total host and Diptera species richness were lower in Agricultural (65 and 64, respectively) than Village/Urban (87 and 74), and Near-natural (140 and 115) habitats (Figure 1C, D, E). We found no significant trend in host and biting Diptera richness with latitude after controlling for sampling effort. Accumulation curves showed that host and biting Diptera species were well-resolved in Agricultural and Village/Urban habitats, with curves approaching an asymptote in each case; host richness in Near-natural habitats was less complete. Sampling of biting Diptera-host interactions was incomplete for all levels of anthropogenic landscape modification (Figure S1).

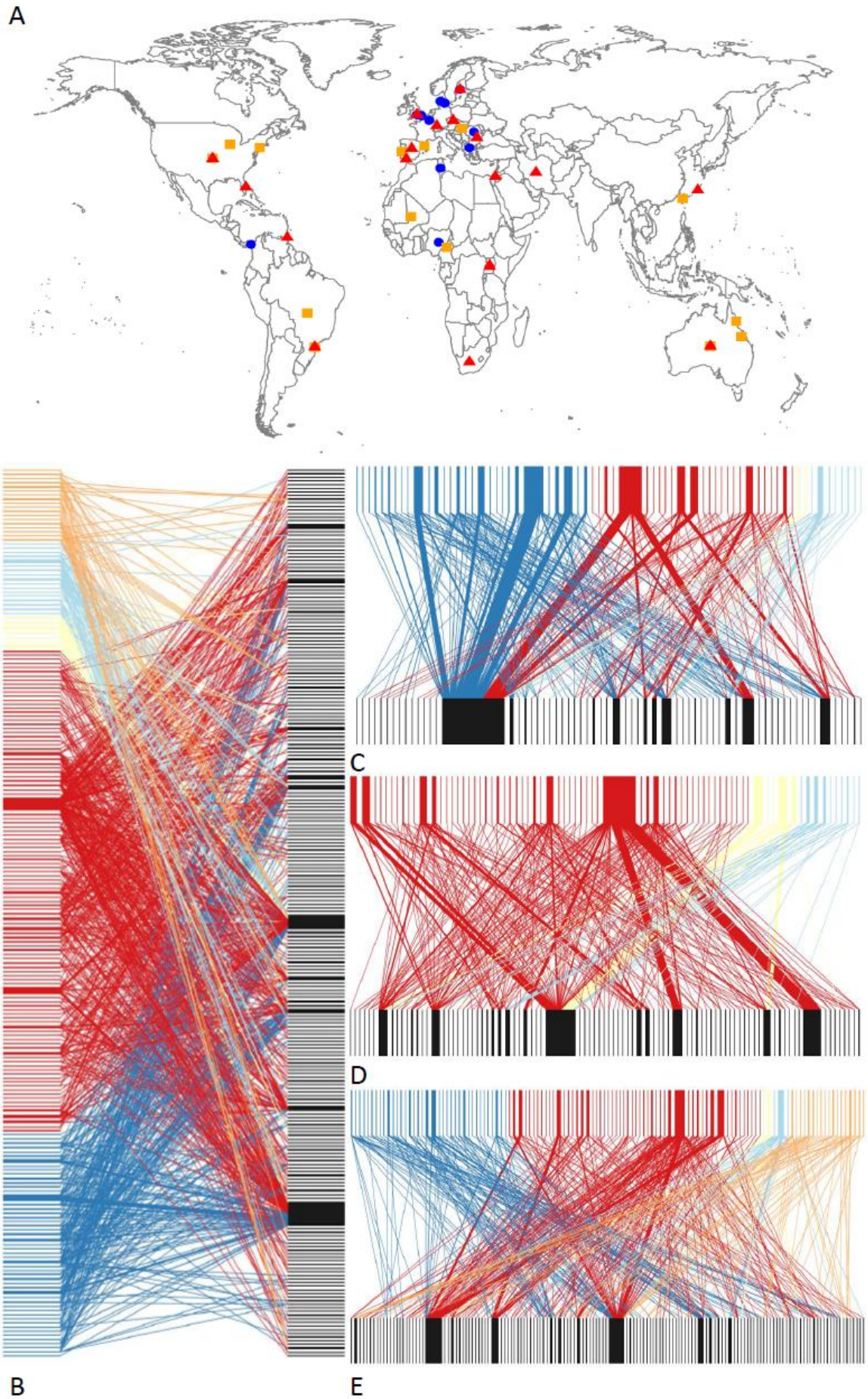

A
B
C
D
E

Figure 1 A**.** Global distribution of the studies included in the analysis. Agricultural sites are represented by blue circles, Near-natural sites by red triangles, and Village/Urban sites by yellow squares. Co-located symbols are studies that sampled in multiple habitat classifications and were treated as separate networks in our analyses (Map data: Google Maps 2020). B**.** The aggregated global network containing all host (right) and Diptera (left) interactions, and Diptera (top)-host (bottom) interactions separated by habitat classification: Agricultural(C), Village/Urban (D), and Near-natural (E). Node and edge widths are proportional to frequency of occurrence and are coloured by biting Diptera family (Ceratopogonidae= dark blue, Culicidae= red, Glossinidae= yellow, Psychodidae= light blue, and Simuliidae= orange).

Influence of geographical and anthropogenic factors on biting Diptera-host networks

Interaction evenness differed significantly among habitat types (χ^2=0.068, df=2, p=0.042), but did not show a latitudinal trend ($\chi^2$0. 001, df=1, p= 0.968) (Figure 2B). Mean interaction evenness was significantly lower in Agricultural habitats (mean=0.472, SE=0.026) than in both Village/Urban (mean=0.576, SE= 0.023, Tukey; p=0.027) and Near-natural habitats (mean= 0.558, SE= 0.027, Tukey; p=0.040), but did not differ significantly between Village/Urban and Near-natural habitats (Tukey; p=0.943) (Figure 2A). Species richness (χ^2=0.014, df=1, p=0.239), network size (χ^2=0.023, df=1, p=0.147), and family (χ^2=0.063, df=4, p=0.180) did not explain a significant amount of variance in interaction evenness and there were no significant interactions between family and habitat (χ^2=0.045, df=5, p=0.495) or family and latitude (χ^2=0.026, df=3, p=0.475).

The average network specialisation (H2') across our networks was 0.395 (SD=0.034). Neither latitude (χ^2= 0.114, df=1, p=0.115) nor habitat (χ^2=0.035, df=1, p=0.682) had a significant influence on H2' (Figure 2C, D). H2' differed significantly among Diptera families (χ^2=0.429, df=4, p=0.030). Simuliidae-dominated networks had the highest average H2' (mean=0.619), followed by Ceratopogonidae (mean=0.296) and Glossinidae (mean=0.127). There were no significant interactions between family and habitat (χ^2=0.207, df=5, p=0.391) or family and latitude (X^2=0.207, df=3, p=0.242). There was a highly significant decrease in H2' (t=-2.875, df=41, p=0.006) with increasing matrix size.

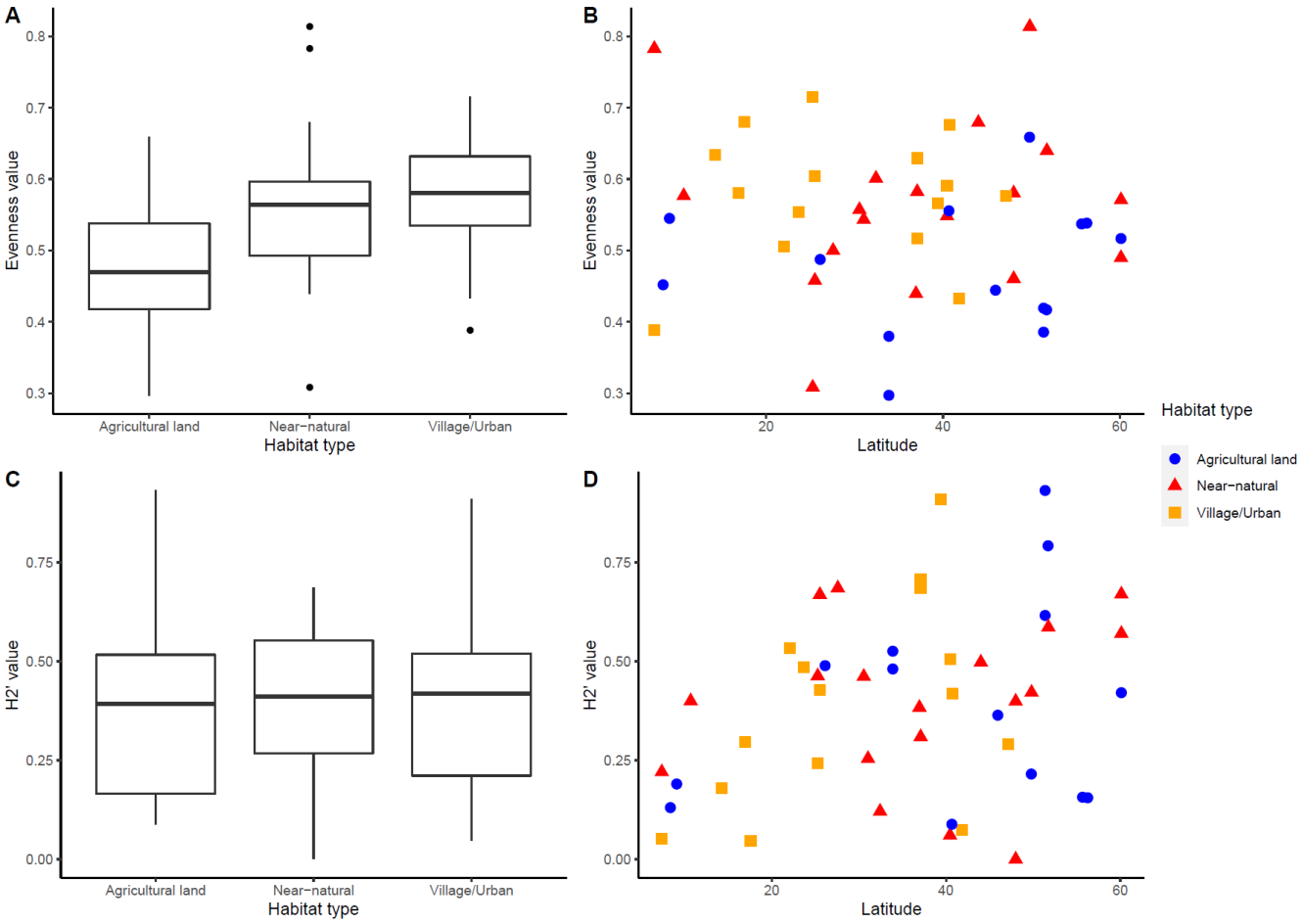


Figure 2. Interaction evenness by habitat type (A) and latitude (B), and H2' specialisation by habitat type (C) and latitude (D). Interaction evenness was significantly lower in Agricultural habitats. Each box displays the interquartile range and the solid line represents the median. Whiskers display the maximum and minimum interaction evenness and H2' values for each habitat type.

The influence of domesticated animals on interaction evenness

Humans and domestic animals were involved in 2,928 interactions across Agricultural networks, of which 2,653 were Diptera-bovine interactions, involving 87 Diptera species from 5 families (Ceratopogonidae, Glossinidae, Culicidae, Psychodidae, and Simuliidae). Near-natural networks contained 1,886 interactions involving humans and domestic animals,

and there were 1,560 such interactions in Village/Urban networks. There was no evidence that the presence of domesticated animals and humans within networks led to altered patterns of interaction evenness. Resampling networks to target removal of these species did not lead to mean interaction evenness values that differed from those generated when an equivalent number of host species selected at random was removed: Agricultural (mean= 0.721, n=10, SE= 0.472), Village/Urban (mean= 0.696, n=7, SE= 0.434), and Near-natural (mean=0.563, n=16, SE=0.292) (Figure S2).

3.5 Discussion

Overall, latitudinal trends in the structure of biting Diptera-host network properties were dwarfed by the impact of anthropogenic habitat modification. Neither network-level metric varied significantly with latitude, but agricultural habitats had significantly lower interaction evenness than Near-natural and Village/Urban habitats.

The absence of a latitudinal trend in feeding specialisation or interaction evenness is counter to the expectation that high species richness at low latitudes will be associated with increased dietary specialisation (Dyer et al., 2007). Turnover in genera of Diptera and hosts was low across latitudes, perhaps explaining the consistency of interspecific interactions. Whilst Diptera families differ in degree of network specialisation as judged by H2', there was no significant interaction between family and latitude. Consequently, the absence of clear latitudinal trends in network structure, as also documented for host-parasitoid networks (Morris et al., 2014), could result from underlying rules for how these antagonistic interactions are structured, regardless of the diversity and size of component networks, or their taxonomic composition. However, the apparent lack of a latitudinal trend may also result from local climatic differences (e.g. rainfall) among sites at similar latitudes masking latitudinal effects (Fischer et al., 2022; Zhu et al., 2014).

The low levels of network specialisation (Agricultural mean H2'= 0.397, Village/Urban mean H2'=0.390, Near-Natural mean H2'= 0.398) may reflect plasticity in host choice and the wide global distribution of suitable hosts. Host usage is characterised by a high degree of plasticity in biting insects and may be strongly influenced by host densities (Takken & Verhulst, 2013). For example, biting Diptera that are commonly described as anthropophilic such as *An.*

gambiae still interact with a range of domestic and non-domestic hosts (Bellekom et al., 2021).

The low interaction evenness observed within Agricultural habitats, in comparison to other habitat types, indicates that interactions are dominated by relatively few species pairs, with a long tail of infrequently observed interactions. In Agricultural habitats, interactions involving domestic animals and humans dominated the networks, accounting for 81% of interactions, with cattle (51%) the most frequent hosts. This may reflect the high biomass of domestic animals (Lassen et al., 2012) in these habitats, and potentially the success and dominance of anthropophilic and livestock-adapted biting Diptera species. Approximately 70% of biting Diptera species in the Agricultural networks fed predominantly (>50% of interactions) on humans and domestic animals, whilst the remaining Diptera fed on either a wider range of mammals and birds or had too few recorded blood meals to assess their diets with confidence (Bellekom et al., 2021). The minimal difference observed between our Agricultural model with targeted removal of domestic hosts and the null model with random removal may be explained by the high number of domestic hosts compared with non-domestic hosts; random removal of hosts inevitably results in the removal of domestic hosts.

We found little difference in interaction evenness between Village/Urban and Near-natural environments, although sampling was less complete within Near-natural habitats. Despite an expected high human host availability, the Village/Urban networks were not dominated by interactions between humans and biting Diptera to the same extent as Agricultural habitats were by biting Diptera-cattle interactions. In Village/Urban habitats, 26% of interactions were with humans, and we identified interactions with a wide range of other taxa, predominantly birds and domesticated animals. Therefore, high interaction evenness values within Village/Urban sites may result from higher than expected generalism of biting Diptera (Hassell et al., 2017) and the absence of dominating species pairs (pairs with high abundance and interaction fidelity). In contrast, network interaction structure in Agricultural habitats is strongly influenced by the super-abundance of a single suitable host species (cattle). Interaction evenness may therefore largely reflect host species evenness; independent data on host abundances would be required to test this.

For many taxa, specialist species are often more highly represented within pristine habitats, and are more susceptible to anthropogenic landscape modification and homogenisation

than generalist species (Devictor, Julliard, & Jiguet, 2008; Devictor, Julliard, Clavel, et al., 2008; Sverdrup-Thygeson et al., 2017). Despite this, we did not find differences in network specialism across levels of anthropogenic landscape modification, perhaps because the same genera and often the same species of biting Diptera were documented across habitat types, leading to similar feeding patterns.

Ecological networks are often asymmetric, composed of few strong interactions and a greater number of weak interactions (Poulin, 2010). Therefore, nodes likely affect each other with differing amounts of reciprocation, and the strength of an interaction is often determined by the distribution of abundance of the component species (Vázquez et al., 2007). Consequently, pairs of abundant species may exhibit more symmetric, and reciprocally strong, effects on the network than pairs of rare species (Dormann et al., 2017). Because independent abundance data were lacking for nodes within the networks, we considered all interactions to be inherently equivalent, with a normalised interaction strength of 1. This data limitation makes it impossible to identify the relative dynamic importance of different nodes. This is a common limitation in ecological network analyses of other interaction types, such as pollination (Novella-Fernandez et al., 2019) and herbivory (Neff et al., 2021), in which the influence of each pollination visit or herbivory damage by different species is considered of equivalent impact on a plant. Network metrics can be sensitive to network size and species richness, leading to a risk that trends in cross-network analyses reflect sampling differences, rather than genuine ecological patterns (Dormann et al., 2009). Heterogeneous data extracted from the literature are particularly susceptible to such biases as a result of variations in methods, sampling intensity and network dimensions (Prendergast & Ollerton, 2022; Xing & Fayle, 2021). Interaction evenness and network specialisation are relatively robust to network size differences (Blüthgen et al., 2006) and we included network size and species richness as explanatory variables in statistical models to control for this potential source of bias.

Agriculturally driven anthropogenic habitat modification, through its effects on biting Diptera-host interaction evenness and network specialism, could result in increased zoonotic disease transmission potential (McDaniel et al., 2014). There was a very high number of biting Diptera-bovine interactions in our Agricultural networks, involving a wide range of Diptera species from multiple families, many of which are vector-competent. This may be of

particular concern, since bovine-related diseases such as Rift Valley fever, Animal African Trypanosomosis (nagana), and Bluetongue disease have high morbidity and mortality rates (Lopes et al., 2020; Rushton & Lyons, 2015; Vreysen et al., 2013). The growing global demand for agriculture products will result in continued anthropogenic habitat modification, which will provide increasing opportunities for pairwise interactions between unfamiliar species, zoonotic transmission, and the emergence of novel zoonotic disease (Carlson et al., 2022). Surveillance of biting Diptera-host networks, particularly at the interface of humans, wildlife, and domestic animals, could help identify pathways of zoonotic disease transmission and help predict and mitigate future spill-over events. Surveillance may be conducted through the routine sampling of the Diptera community using a combination of trapping methods, such as malaise traps, USA Center for Disease Control (CDC) miniature light traps and Modified CDC Backpack Aspirators, Biogents' Sentinel (BGS), as well as trapping locations that limit accidental overrepresentation of a species in order to minimise sources of bias (Bellekom et al., 2021; Grubaugh et al., 2015; Gyawali et al., 2019; Rivera et al., 2021).

3.6 Acknowledgements

We thank John Connolly and the late John Mumford for their helpful comments and suggestions. We also thank three anonymous reviewers for their helpful comments. TDH and OTL were supported by Target Malaria which is supported by a grant from the Bill and Melinda Gates Foundation and the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation.

Authors' contributions

BB: Data analysis and writing, OTL: Supervision, review, and editing, TDH: Supervision, review, and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Data Availability Statement

The data that support the findings of this study are available in [Latitudinal-and-anthropogenic-effects-on-the-structuring-of-networks] at [https://github.com/Ben-Bellekom/Latitudinal-and-anthropogenic-effects-on-the-structuring-of-networks].

Competing interests

We declare we have no competing interests.

3.7 References

Abella-Medrano, C. A., Ibáñez-Bernal, S., Carbó-Ramírez, P., & Santiago-Alarcon, D. (2018). Blood-meal preferences and avian malaria detection in mosquitoes (Diptera: Culicidae) captured at different land use types within a neotropical montane cloud forest matrix. *Parasitology International*, *67*(3), 313–320. https://doi.org/10.1016/J.PARINT.2018.01.006

Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L. F., Wesolowski, A., & Metcalf, C. J. E. (2021). Infectious disease in an era of global change. *Nature Reviews Microbiology 2021 20:4*, *20*(4), 193–205. https://doi.org/10.1038/s41579-021-00639-z

Barber, L. M., Schleier, J. J., & Peterson, R. K. D. (2010). Economic Cost Analysis of West Nile Virus Outbreak, Sacramento County, California, USA, 2005. *Emerging Infectious Diseases*, *16*(3), 480. https://doi.org/10.3201/EID1603.090667

Bellekom, B., Hackett, T. D., & Lewis, O. T. (2021). A Network Perspective on the Vectoring of Human Disease. *Trends in Parasitology*, *37*(5), 391–400. https://doi.org/10.1016/j.pt.2020.12.001

Blüthgen, N., Fründ, J., Vázquez, D. P., & Menzel, F. (2008). What do interaction network metrics tell us about specialization and biological traits. *Ecology*, *89*(12), 3387–3399. https://doi.org/10.1890/07-2121.1

Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecology*, *6*(1), 9. https://doi.org/10.1186/1472-6785-6-9

Carlson, C. J., Albery, G. F., Merow, C., Trisos, C. H., Zipfel, C. M., Eskew, E. A., Olival, K. J., Ross, N., & Bansal, S. (2022). Climate change increases cross-species viral transmission risk. *Nature 2022*, 1–1. https://doi.org/10.1038/s41586-022-04788-w

Devictor, V., Julliard, R., Clavel, J., Jiguet, F., Lee, A., & Couvet, D. (2008). Functional biotic homogenization of bird communities in disturbed landscapes. *Global Ecology and Biogeography*, *17*(2), 252–261. https://doi.org/10.1111/j.1466-8238.2007.00364.x

Devictor, V., Julliard, R., & Jiguet, F. (2008). Distribution of specialist and generalist species

along spatial gradients of habitat disturbance and fragmentation. *Oikos*, *117*(4), 507–514. https://doi.org/10.1111/J.0030-1299.2008.16215.X

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal*, *2*(1), 7–24. https://doi.org/10.2174/1874213000902010007

Dormann, C. F., Fründ, J., & Schaefer, H. M. (2017). Identifying Causes of Patterns in Ecological Networks: Opportunities and Limitations. *Https://Doi.Org/10.1146/Annurev-Ecolsys-110316-022928*, *48*, 559–584. https://doi.org/10.1146/ANNUREV-ECOLSYS-110316-022928

Dormann, C. F., Gruber, B., & Fründ, J. (2008). *Introducing the bipartite Package: Analysing Ecological Networks*. https://www.semanticscholar.org/paper/Introducing-the-bipartite-Package%3A-Analysing-Dormann-Gruber/f23d90643ae91bd7f33650a69069b0221ca13b30

Dyer, L. A., Singer, M. S., Lill, J. T., Stireman, J. O., Gentry, G. L., Marquis, R. J., Ricklefs, R. E., Greeney, H. F., Wagner, D. L., Morais, H. C., Diniz, I. R., Kursar, T. A., & Coley, P. D. (2007). Host specificity of Lepidoptera in tropical and temperate forests. *Nature*, *448*(7154), 696–699. https://doi.org/10.1038/nature05884

Fischer, C., Gerstmeier, R., & Wagner, T. C. (2022). Seasonal and temporal patterns of rainfall shape arthropod community composition and multi-trophic interactions in an arid environment. *Scientific Reports 2022 12:1*, *12*(1), 1–9. https://doi.org/10.1038/s41598-022-07716-0

Galiana, N., Hawkins, B. A., & Montoya, J. M. (2019). The geographical variation of network structure is scale dependent: understanding the biotic specialization of host–parasitoid networks. *Ecography*, *42*(6), 1175–1187. https://doi.org/10.1111/ecog.03684

GBIF.org (2021), GBIF Home Page. (n.d.). Available from: Https://Www.Gbif.Org.

Grubaugh, N. D., Sharma, S., Krajacich, B. J., Fakoli III, L. S., Bolay, F. K., Diclaro II, J. W., Johnson, W. E., Ebel, G. D., Foy, B. D., & Brackney, D. E. (2015). Xenosurveillance: A Novel Mosquito-Based Approach for Examining the Human-Pathogen Landscape. *PLOS Neglected Tropical Diseases*, *9*(3), e0003628.

https://doi.org/10.1371/journal.pntd.0003628

Gyawali, N., Taylor-Robinson, A. W., Bradbury, R. S., Huggins, D. W., Hugo, L. E., Lowry, K., & Aaskov, J. G. (2019). Identification of the source of blood meals in mosquitoes collected from north-eastern Australia. *Parasites & Vectors*, *12*(1), 198. https://doi.org/10.1186/s13071-019-3455-2

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and Disease Emergence: Dynamics at the Wildlife–Livestock–Human Interface. In *Trends in Ecology and Evolution* (Vol. 32, Issue 1, pp. 55–67). Elsevier Ltd. https://doi.org/10.1016/j.tree.2016.09.012

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*(12), 1451–1456. https://doi.org/10.1111/2041-210X.12613

Kaiser-Bunbury, C. N., & Blüthgen, N. (2015). Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants*, *7*, plv076. https://doi.org/10.1093/aobpla/plv076

Killick-Kendrick, R. (1999). The biology and control of Phlebotomine sand flies. *Clinics in Dermatology*, *17*(3), 279–289. https://doi.org/10.1016/S0738-081X(99)00046-2

Kilpatrick, A. M., Kramer, L. D., Campbell, S. R., Alleyne, E. O., Dobson, A. P., & Daszak, P. (2005). West Nile Virus Risk Assessment and the Bridge Vector Paradigm. *Emerging Infectious Diseases*, *11*(3), 425. https://doi.org/10.3201/EID1103.040364

Lassen, S., Nielsen, S., & Kristensen, M. (2012). Identity and diversity of blood meal hosts of biting midges (Diptera: Ceratopogonidae: Culicoides Latreille) in Denmark. *Parasites & Vectors*, *5*(1), 143. https://doi.org/10.1186/1756-3305-5-143

Lawrence, E. R., & Fraser, D. J. (2020). Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity. In *Global Ecology and Biogeography* (Vol. 29, Issue 5, pp. 770–788). Blackwell Publishing Ltd. https://doi.org/10.1111/geb.13075

Logue, K., Keven, J. B., Cannon, M. V., Reimer, L., Siba, P., Walker, E. D., Zimmerman, P. A., &

Serre, D. (2016). Unbiased Characterization of Anopheles Mosquito Blood Meals by Targeted High-Throughput Sequencing. *PLOS Neglected Tropical Diseases*, *10*(3), e0004512. https://doi.org/10.1371/journal.pntd.0004512

Lopes, V. L., Costa, F. V., Rodrigues, R. A., Braga, M., Pichorim, M., & Moreira, P. A. (2020). High fidelity defines the temporal consistency of host-parasite interactions in a tropical coastal ecosystem. *Scientific Reports*, *10*(1). https://doi.org/10.1038/S41598-020-73563-6

Mayi, M. P. A., Bamou, R., Djiappi-Tchamen, B., Fontaine, A., Jeffries, C. L., Walker, T., Antonio-Nkondjio, C., Cornel, A. J., & Tchuinkam, T. (2020). Habitat and Seasonality Affect Mosquito Community Composition in the West Region of Cameroon. *Insects*, *11*(5), 312. https://doi.org/10.3390/insects11050312

McDaniel, C. J., Cardwell, D. M., Moeller, R. B., & Gray, G. C. (2014). Humans and Cattle: A Review of Bovine Zoonoses. *Vector Borne and Zoonotic Diseases*, *14*(1), 1. https://doi.org/10.1089/VBZ.2012.1164

Meyer Steiger, D. B., Ritchie, S. A., & Laurance, S. G. W. (2016). Mosquito communities and disease risk influenced by land use change and seasonality in the Australian tropics. *Parasites & Vectors*, *9*(1), 387. https://doi.org/10.1186/s13071-016-1675-2

Morris, R. J. (2010). Anthropogenic impacts on tropical forest biodiversity: a network structure and ecosystem functioning perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1558), 3709–3718. https://doi.org/10.1098/rstb.2010.0273

Morris, R. J., Gripenberg, S., Lewis, O. T., & Roslin, T. (2014). Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*, *17*(3), 340–349. https://doi.org/10.1111/ele.12235

Müller, R., Reuss, F., Kendrovski, V., & Montag, D. (2019). Vector-Borne Diseases. In *Biodiversity and Health in the Face of Climate Change* (pp. 67–90). Springer International Publishing. https://doi.org/10.1007/978-3-030-02318-8_4

Neff, F., Brändle, M., Ambarli, D., Ammer, C., Bauhus, J., Boch, S., Hölzel, N., Klaus, V. H., Kleinebecker, T., Prati, D., Schall, P., Schäfer, D., Schulze, E. D., Seibold, S., Simons, N. K.,

Weisser, W. W., Pellissier, L., & Gossner, M. M. (2021). Changes in plant-herbivore network structure and robustness along land-use intensity gradients in grasslands and forests. *Science Advances*, *7*(20), 18. https://doi.org/10.1126/SCIADV.ABF3985

Novella-Fernandez, R., Rodrigo, A., Arnan, X., & Bosch, J. (2019). Interaction strength in plant-pollinator networks: Are we using the right measure? *PLoS ONE*, *14*(12). https://doi.org/10.1371/JOURNAL.PONE.0225930

Pellissier, L., Albouy, C., Bascompte, J., Farwig, N., Graham, C., Loreau, M., Maglianesi, M. A., Melián, C. J., Pitteloud, C., Roslin, T., Rohr, R., Saavedra, S., Thuiller, W., Woodward, G., Zimmermann, N. E., & Gravel, D. (2018). Comparing species interaction networks along environmental gradients. *Biological Reviews*, *93*(2), 785–800. https://doi.org/10.1111/brv.12366

Poulin, R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, *26*(10), 492–498. https://doi.org/10.1016/J.PT.2010.05.008

Prendergast, K. S., & Ollerton, J. (2022). Spatial and temporal scale of analysis alter conclusions about the effects of urbanisation on plant–pollinator networks. *Arthropod-Plant Interactions*, *1*, 1–13. https://doi.org/10.1007/S11829-022-09925-W/FIGURES/2

Rivera, N. A., Varga, C., Ruder, M. G., Dorak, S. J., Roca, A. L., Novakofski, J. E., & Mateus-Pinilla, N. E. (2021). Bluetongue and Epizootic Hemorrhagic Disease in the United States of America at the Wildlife–Livestock Interface. *Pathogens*, *10*(8). https://doi.org/10.3390/PATHOGENS10080915

Runghen, R., Poulin, R., Monlleó-Borrull, C., & Llopis-Belenguer, C. (2021). Network Analysis: Ten Years Shining Light on Host–Parasite Interactions. In *Trends in Parasitology*. Elsevier Ltd. https://doi.org/10.1016/j.pt.2021.01.005

Rushton, J., & Lyons, N. (2015). Economic impact of Bluetongue: a review of the effects on production. *Veterinaria Italiana*, *51*(4), 401–406. https://doi.org/10.12834/VETIT.646.3183.1

Sachs, J., & Malaney, P. (2002). The economic and social burden of malaria. *Nature*, *415*(6872), 680–685. https://doi.org/10.1038/415680a

Schemske, D. W., Mittelbach, G. G., Cornell, H. V, Sobel, J. M., Roy, K., & Kellogg, W. K. (2009). *Is There a Latitudinal Gradient in the Importance of Biotic Interactions?* https://doi.org/10.1146/annurev.ecolsys.39.110707.173430

Schleuning, M., Fründ, J., Klein, A.-M., Abrahamczyk, S., Alarcón, R., Albrecht, M., Andersson, G. K. S., Bazarian, S., Böhning-Gaese, K., Bommarco, R., Dalsgaard, B., Dehling, D. M., Gotlieb, A., Hagen, M., Hickler, T., Holzschuh, A., Kaiser-Bunbury, C. N., Kreft, H., Morris, R. J., … Blüthgen, N. (2012a). Specialization of Mutualistic Interaction Networks Decreases toward Tropical Latitudes. *Current Biology*, *22*(20), 1925–1931. https://doi.org/10.1016/J.CUB.2012.08.015

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., Mbogo, C. M., Hemingway, J., Patil, A. P., Temperley, W. H., Gething, P. W., Kabaria, C. W., Burkot, T. R., Harbach, R. E., & Hay, S. I. (2012). A global map of dominant malaria vectors. *Parasites & Vectors*, *5*(1), 69. https://doi.org/10.1186/1756-3305-5-69

Sverdrup-Thygeson, A., Skarpaas, O., Blumentrath, S., Birkemoe, T., & Evju, M. (2017). Habitat connectivity affects specialist species richness more than generalists in veteran trees. *Forest Ecology and Management*, *403*, 96–102. https://doi.org/10.1016/J.FORECO.2017.08.003

Takken, W., & Verhulst, N. O. (2013). Host Preferences of Blood-Feeding Mosquitoes. *Annual Review of Entomology*, *58*(1), 433–453. https://doi.org/10.1146/annurev-ento-120811-153618

Tylianakis, J. M., Tscharntke, T., & Lewis, O. T. (2007). Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, *445*(7124), 202–205. https://doi.org/10.1038/nature05429

Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, *116*(7), 1120–1127. https://doi.org/10.1111/J.0030-1299.2007.15828.X

Vreysen, M. J. B., Seck, M. T., Sall, B., & Bouyer, J. (2013). Tsetse flies: Their biology and control using area-wide integrated pest management approaches. *Journal of Invertebrate Pathology*, *112*, S15–S25. https://doi.org/10.1016/J.JIP.2012.07.026

Weiner, C. N., Werner, M., Linsenmair, K. E., & Blüthgen, N. (2014). Land-use impacts on plant–pollinator networks: interaction strength and specialization predict pollinator declines. *Ecology*, *95*(2), 466–474. https://doi.org/10.1890/13-0436.1

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Willig, M. R., Kaufman, D. M., & Stevens, R. D. (2003). Latitudinal Gradients of Biodiversity: Pattern, Process, Scale, and Synthesis. In *Annual Review of Ecology, Evolution, and Systematics* (Vol. 34, pp. 273–309). Annual Reviews Inc. https://doi.org/10.1146/annurev.ecolsys.34.012103.144032

Xing, S., & Fayle, T. M. (2021). The rise of ecological network meta-analyses: Problems and prospects. *Global Ecology and Conservation*, *30*, e01805. https://doi.org/10.1016/j.gecco.2021.e01805

Zhu, H., Wang, D., Wang, L., Fang, J., Sun, W., & Ren, B. (2014). Effects of altered precipitation on insect community composition and structure in a meadow steppe. *Ecological Entomology*, *39*(4), 453–461. https://doi.org/10.1111/EEN.12120

3.8 Supplementary

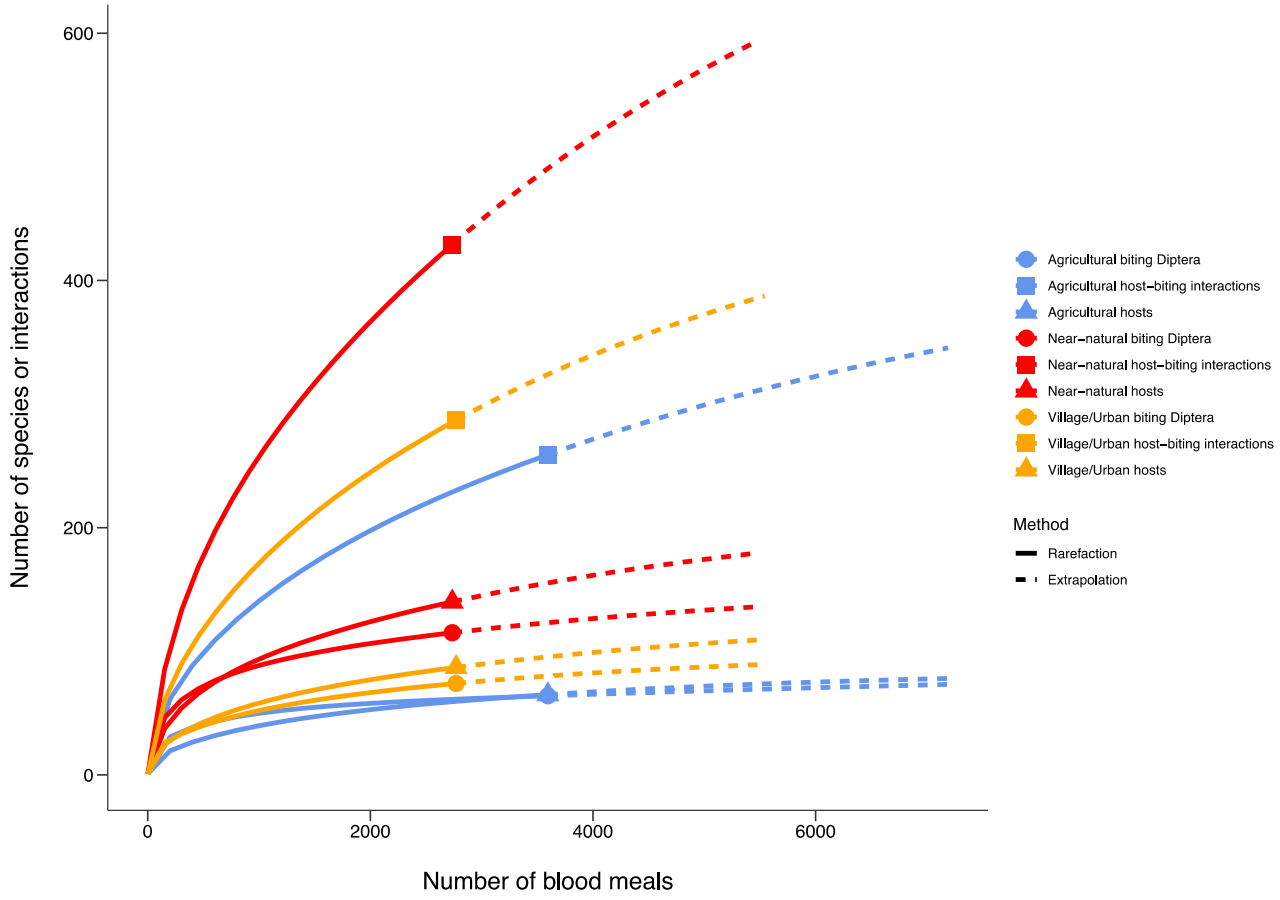


Figure S1. Smoothed accumulation and extrapolation curves to assess sampling completeness. Total numbers of hosts (triangle), biting Diptera (circle), and interactions recorded in the whole dataset (square), by habitat type: Agricultural (blue), Near-natural (red), and Village/Urban (orange), as a function of sampling effort (the number of blood meals screened).

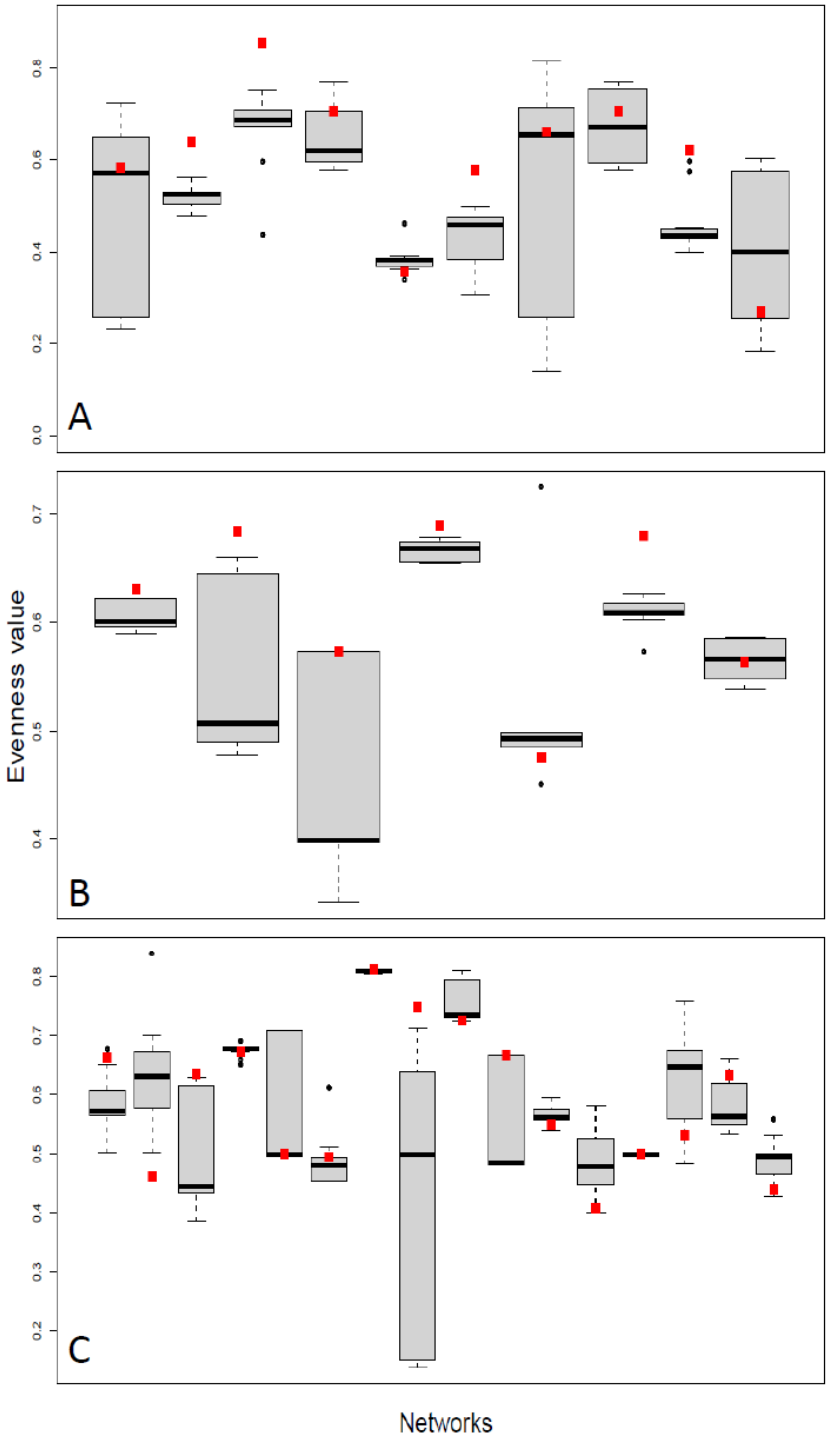
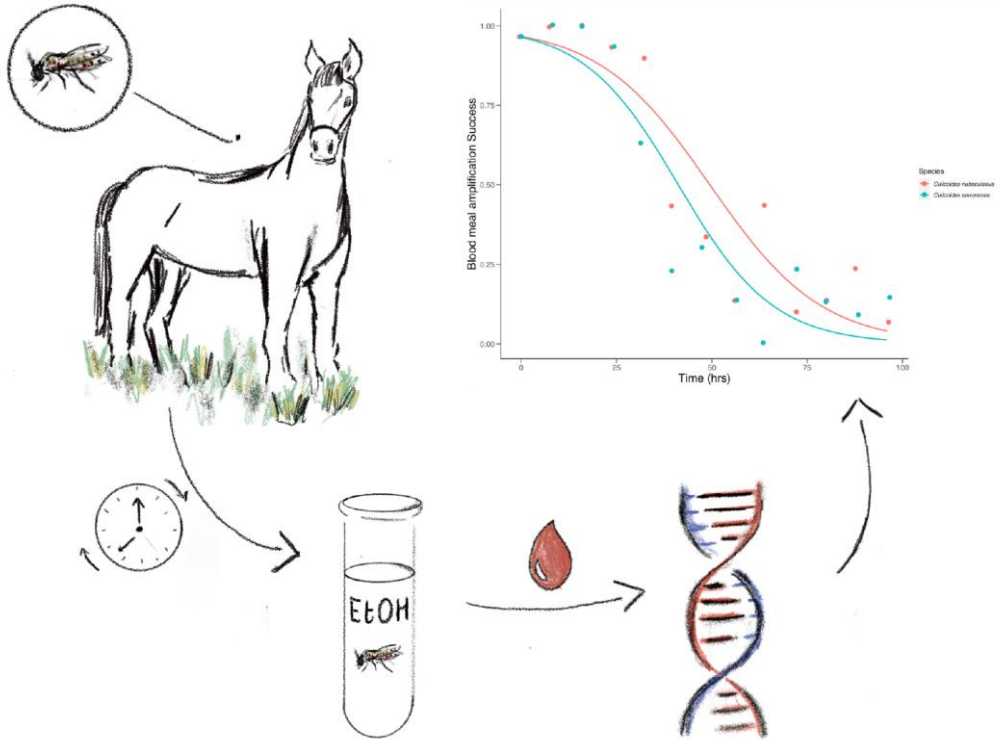


Figure S2. Null model interaction evenness for each component network, with empirical IE values (red squares), by habitat type: Agricultural (A), Village/Urban (B), and Near-natural (C). Each grey box displays the interquartile range, and the solid line represents the median values for interaction evenness. Whiskers display the maximum and minimum interaction evenness for each network.

Chapter 4: Effects of storage conditions and digestion time on DNA amplification of biting midge (*Culicoides*) blood meal



Author contributions: BB: conceptualisation, data analysis and writing, AB: review, technical advice, and guidance, ME: midge rearing, review, and editing, ZL: midge rearing, OTL and TDH: supervision, review, and editing.

This chapter has been published as:

Bellekom, B., Bailey, A., England, M., Langlands, Z., Lewis, O. T., & Hackett, T. D. (2023). Effects of storage conditions and digestion time on DNA amplification of biting midge (Culicoides) blood meals. *Parasites & Vectors*, 16(1), 1–7. https://doi.org/10.1186/S13071-022-05607-X

The full published version is found in Appendix 5**.**

Ben Bellekom[1*], Abigail Bailey[1], Marion England[2], Zoe Langlands[2], Owen T. Lewis[1] and Talya D. Hackett[1]

[1]University of Oxford, Department of Biology, 11a Mansfield Road, Oxford OX1 3SZ, UK

[2]The Pirbright Institute, Ash Road, Pirbright, Surrey, GU24 0NF, UK

* Corresponding author: ben.bellekom@biology.ox.ac.uk

BB: ben.bellekom@biology.ox.ac.uk

AB: abigail.bailey@biology.ox.ac.uk

ME: marion.england@pirbright.ac.uk

ZL: zoe.langlands@pirbright.ac.uk

OTL: owen.lewis@biology.ox.ac.uk

TDH: talya.hackett@biology.ox.ac.uk

4.1 Abstract

Background

Molecular analysis of blood meals is increasingly used to identify the hosts of biting insects such as midges and mosquitoes. Successful host identification depends on the availability of sufficient host DNA template for PCR amplification, making it important to understand how amplification success changes under different storage conditions and with different durations of blood meal digestion within the insect gut, before being placed into the storage medium.

Method

We characterised and compared the digestion profile of two species of *Culicoides* over a 96-hour period using a novel set of general vertebrate primers targeting the 16S rRNA gene. A set number of individuals from each species were sacrificed over 13 time points post-blood feeding and preserved in 95% ethanol. Samples were stored either at ambient room temperature or in a -20°C freezer, to examine the effect of storage condition on the PCR amplification success of host DNA.

Results

We found that amplification success across the 96-hour sampling period post-feeding was reduced from 96% to 6% and 96% to 14% for *Culicoides nubeculosus* and *Culicoides sonorensis,* respectively. We found no effect of storage condition on PCR amplification success, and storage in 95% ethanol was sufficient to maintain high rates of amplifiable host DNA for at least 9 months, even at room temperature.

Conclusions

These findings highlight the limited time frame during which an individual may contain amplifiable host DNA and demonstrate the importance of timely sample capture and processing post blood feeding. Moreover, storage in 95% ethanol alone is sufficient to limit host DNA degradation. These results are relevant to the design of studies investigating the biting behaviour and disease transmission potential of *Culicoides* and other biting Diptera.

Key words: *Culicoides,* DNA degradation, Metabarcoding, Blood meal, Biting Diptera, DNA digestion

4.2 Introduction

Biting (haematophagous) Diptera are vectors of a wide range of pathogens that cause disease in humans as well as domesticated and wild animals. Transmission dynamics of these diseases are, in part, mediated by the host-biting preferences of the vectors [1]. Thus, data on vertebrate blood meal sources and host preferences can provide valuable information for understanding disease transmission [2], including transmission across the livestock and wildlife interface. Blood meal data can also be used to populate ecological interaction networks, helping to identify transmission pathways and reservoirs of infection, and informing control strategies [3].

In the past, host blood meals were identified using serological methods which have now been largely superseded by molecular methods, diagnostic PCR, DNA barcoding and, most recently, metabarcoding [2,3]. These methods rely on extraction of DNA from the blood meal, PCR amplification of a taxonomically informative fragment of host DNA [1], followed by sequencing of the PCR product or diagnostic PCR using primers specific to a species or set of species. From as little as US $2/sample at 2022 prices, barcoding is now a relatively low-cost method for sequencing a single gene, but it is unsuitable for mixtures of DNA from multiple species. Metabarcoding coupled with Next Generation Sequencing allows a large number of sequences to be read in parallel, facilitating the analysis of multiple gene loci and individual samples, and the simultaneous identification of species within mixed samples such as gut contents [4], and blood meals [1]. Whilst sequencing costs have dramatically fallen, metabarcoding still remains more expensive than targeted DNA barcoding. Price per individual sample may vary greatly depending on the sequencing kit and platform used, the number of samples being processed, and the target read depth.

DNA degradation by digestion enzymes will reduce the success of PCR amplification, jeopardising successful blood meal diagnosis [2,5–7]. In mosquitoes (Culicidae), digestion occurs rapidly, with host DNA undetectable within 32-72 hours [5,8]. However, rates of DNA degradation attributable to digestion are likely to vary among haematophagous taxa, and

equivalent data for other biting Diptera families of medical, veterinary, and epidemiological importance are lacking. Prompt preservation of blood fed biting Diptera to halt digestive processes can increase the success of PCR amplification [9]. Preservation methods include storage in ethanol [10], desiccation [11], transferring blood meals onto filter paper [12], and cryopreservation [13,14]. However, field constraints and institutional limitations often limit the use of -20 °C and ultra-low (-80 °C) freezers and collected samples may be left for extended periods at room temperature, especially during transport. It is therefore of considerable practical interest to understand how amplification success varies for samples stored under different conditions.

Here we examine the effect of digestion time and the impact of two storage conditions (ambient room temperature and a -20°C freezer) on the PCR amplification success of host DNA in blood meals from two species of *Culicoides*. Biting midges in this genus have a wide distribution globally, bite a broad range of host species [10], and are vectors of several pathogens of veterinary and medical significance [10,15]. We used PCR amplification success to confirm the presence of sufficient quantities of non-degraded host template DNA to allow for further molecular analysis and host identification. In parallel, we assess the suitability of a novel 16S rRNA general vertebrate primer set for Diptera blood meal identification.

4.3 Methods

We experimentally examined the effect of blood meal digestion time and storage method on the PCR amplification success of host DNA in two species of biting midges (*Culicoides nubeculosus* and *C. sonorensis*) using novel vertebrate PCR primers. Individuals from each species were fed fresh horse blood, sacrificed, and preserved in 95% ethanol at 13 specific time points ranging from freshly fed to 96 hours post-feeding, in intervals of eight hours. At each time point, 30 individuals from each species were knocked down by freezing and stored in ethanol, to minimise DNA degradation. An additional 10 individuals at time point 0 were collected to investigate long term storage effects, thus resulting in 40 individuals at time point 0. In total, 800 midges were collected for blood meal analysis (Table 1).

To examine the effect of storage method, 15 samples from each species at every time point were preserved in 95% ethanol and stored in a -20°C freezer, whilst the remaining 15 were stored in ethanol and left at room temperature (22°C – 24°C). To minimise biases, samples

were extracted and underwent PCR amplification in a randomised order across all time points and storage methods. To examine the impact of long-term storage on DNA integrity, two sets of five individuals of each species from the first time point were preserved in ethanol and stored under the two temperature conditions (-20°C freezer and at ambient room temperature), respectively, and left for 9 months before extraction and amplification. Extraction and amplification of all samples, excluding the long-term storage sets, took place across a three-month period following sample preparation.

Midge rearing

Culicoides sonorensis and *C. nubeculosus* specimens were obtained from lines maintained in existing closed colonies at the Pirbright Institute. The *C. nubeculosus* colony was established in 1969 from pupae collected in Hertfordshire, UK. The *C. sonorensis* colony was established from eggs provided by Dr. H. Jones who initiated a laboratory colony in Colorado in 1957. The colonies were maintained following previously developed protocols [16], with adult females fed on commercially supplied horse blood (TCS Biosciences, UK) using a Hemotek blood feeder (Hemotek, UK). Midges were not sugar fed during the 96-hour period as a trial conducted prior to the commencement of the study indicated that unfed midges survived beyond the 96-hour period (unpublished data). Further, this most accurately represented the behaviour of females in the field following feeding, whereby they rest until oogenesis is complete and then find a suitable habitat to lay [17].

Blood meal analysis

DNA was extracted from individual blood meals using a Qiagen DNeasy Blood and Tissue kit, using the standard protocol with the following minor alterations. Prior to lysis, individuals were homogenised using an MP Biomedicals FastPrep-24 5G homogeniser, to release the blood meal from the abdomen. Prior to homogenisation, a single sterile 2.3mm zirconia/silica agitating ball was placed in each 1.5ml microcentrifuge tube with 180 µl of Buffer ATL. To minimise Buffer ATL foaming, which may reduce the homogenisation efficiency, 2µl of Reagent DX was added to each microcentrifuge tube. To increase final DNA concentration, prior to elution, 60µl of Buffer AE was pipetted onto the Dneasy spin column membrane and allowed to incubate at room temperature for 5 minutes. Final DNA concentration for all samples was quantified using a Qubit 3.0 fluorometer.

In the absence of *a priori* host assumptions and the potential presence of mixed blood meals (derived from different host species), the PCR amplification of host genomic DNA from blood meals of wild-caught *Culicoides* and other haematophagous Diptera require the use of general vertebrate primers [18]. To facilitate the identification of horse-derived blood meals used in this study and to determine the likelihood of successful amplification in future blood meal metabarcoding studies, we designed primers intended to amplify vertebrate templates on the 16S rRNA gene, whilst excluding invertebrate templates. We downloaded 128 mitogenomes belonging to the classes Aves and Mammalia from the NCBI Genbank database. As these primers are designed for blood meal analysis of a wide range of biting Diptera, we also downloaded five biting Diptera mitogenomes (*Anopheles gambiae, Aedes albopictus, Aedes aegypti, Culicoides arakawae* and *Culicoides imicola*). The sequences were MAFFT aligned using the bioinformatics software Geneious Prime (Biomatters Limited, New Zealand). Where DNA is highly degraded, primer pairs targeting short amplicons are preferable. Consequently, this alignment was used to identify two potential 18-22 bp primer binding sites that contained primer-invertebrate annealing site mismatches, which produced a 200-base pair (bp) amplicon. Initial binding site selection was informed by previous work [19]. To ensure binding across all mammal and bird mitogenomes, two degenerate bases were inserted into the forward primer to account for single base mismatches. Primer protocol selection and validation was conducted using gradient PCRs with a range of known mammalian and avian DNA templates (Additional file 1: Text S1.). As the PCR primers were designed for use with a wide range of biting Diptera species, examination of blood meal amplification success and identification of non-specific binding was conducted on genomic DNA from blood fed and unfed *C. sonorensis* and *An. gambiae*. I. Fed and unfed *C. sonorensis* were obtained from the colonies at the Pirbright Institute. *An. gambiae* individuals were wild caught, and stored in 95% ethanol, at field sites in Burkina Faso as part of ongoing research by members of the Target Malaria research consortium (see Acknowledgements). The PCR product of blood fed *C. sonorensis* and *An. gambiae* were Sanger sequenced (Source Bioscience) to confirm successful target amplification.

The presence of sufficient concentrations of non-degraded host DNA template for PCR amplification was assessed using end-point PCR and our novel primer set, 16smbF (5’-GGT TGG GGY GAC CTY GGA-3’) and 16sbbR (5’-CTG ATC CAA CAT CGA GGT CGT A-3’). PCR

amplification was carried out in 25µl reactions that contained, 12.5µl HotStarTaq Master mix (Qiagen, Germany), 10µM of each primer, 8.5µl H_2O, and 2µl of DNA template. The PCR protocol consisted of an initial denaturation step of 15 min at 95°C followed by 35 cycles of 94°C for 45 sec, 58°C for 45 sec, and 72°C for 30 sec, followed by a final extension step of 72°C for 10 min. Two negative controls, containing nuclease free water (ThermoFisher Scientific, United States), were included in every set of reactions to monitor for contamination.

PCR products were electrophoresed and visualised on a gel red (SYBR™ Safe, ThermoFisher Scientific, United States) stained 2% agarose gel. The presence of a band of the expected amplicon size was taken as a positive result, indicating that sufficient host DNA remained viable for blood meal identification. To confirm successful target amplification, a subset of PCR products from across the range of timesteps that yielded bands were Sanger sequenced (Source Bioscience, England) and the sequence's origin was identified using the basic alignment search tool (BLAST).

Statistical analysis

We modelled the effect of digestion time, storage method, and species on a binary measure of amplification success (1 or 0) using a multivariate binomial logistic regression. Model fit was evaluated with a likelihood ratio test and the significance of the overall effect of each variable was determined using the chi squared statistic [5]. Data were analysed using R (version 4.1.2) and visualised using the package *ggplot2* [20].

4.4 Results

Overall, 759 midge samples were used in this study. Some midges died during the later time steps (64 hours post-feeding onwards), so four time-steps had fewer than the planned 15 *C. sonorensis* individuals (Table 1).

Table 1. The number of individuals (and successful amplifications) for both species of *Culicoides* that underwent PCR amplification by time point (hrs) and storage method.

	Culicoides nubeculosus	*Culicoides nubeculosus*	*Culicoides sonorensis*	*Culicoides sonorensis*
Time point (hrs)	Ambient	-20°C Freezer	Ambient	-20°C Freezer
0	20 (20)	20 (19)	20 (19)	20 (20)
8	15 (15)	15 (15)	15 (15)	15 (15)
16	15 (15)	15 (15)	15 (15)	15 (15)
24	15 (14)	15 (14)	15 (14)	15 (14)
32	15 (14)	15 (13)	15 (10)	15 (9)
40	15 (4)	15 (9)	15 (4)	15 (3)
48	15 (4)	15 (6)	15 (4)	15 (5)
56	15 (2)	15 (2)	15 (3)	15 (1)
64	15 (9)	15 (4)	13 (0)	12 (0)
72	15 (2)	15 (1)	15 (3)	15 (4)
80	15 (3)	15 (1)	8 (0)	7 (2)
88	15 (4)	15 (3)	12 (2)	10 (0)
96	15 (0)	15 (2)	7 (1)	7 (1)

Digestion time significantly decreased amplification success of host DNA (χ^2= 402.48, df=1, p < 0.001), with amplification success approaching 6% 96 hours post-feeding (Figure 1A.). We found no significant effect of storage method on amplification success (χ^2= 0.024, df=1, p= 0.875; Figure 1B.). Amplification success was significantly higher for *Culicoides nubeculosus* than *C. sonorensis* (χ^2= 8.318, df=1, p= 0.004; Figure 1). All twenty long-term individuals retained amplifiable host DNA following the 9-month storage period.

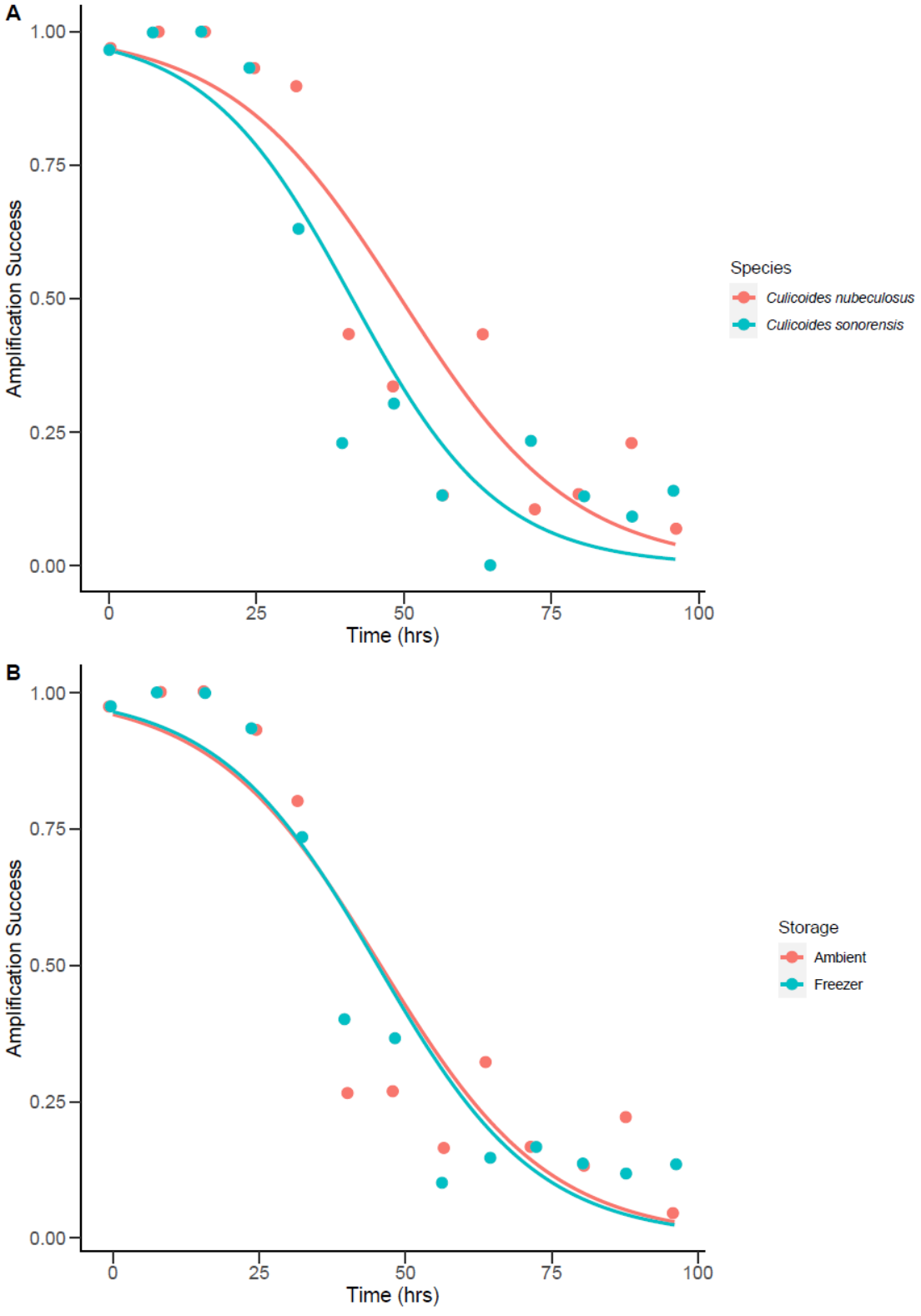


Figure 1. (A) PCR amplification success with time, by *Culicoides* species. Red and blue points represent the amplification success of *Culicoides nubeculosus* and *Culicoides sonorensis* respectively, at a given time point. (B) PCR amplification success with time, by storage

method. The red and blue points represent amplification success of ambient and -20°C freezer conditions respectively, at a given time point.

4.5 Discussion

We found that the duration of blood meal digestion, but not the storage condition (ambient or -20°C freezer) had a significant impact on the success of host DNA PCR amplification, and thus the likelihood of subsequent host identification. Whilst four time-steps post 64 hours had fewer *Culicoides sonorensis* individuals than planned, amplification during this period was limited in both species and the reduction in sample size is unlikely to have impacted the findings significantly.

Increased post-feeding digestion time significantly reduced amplification success of blood meals from both colony *Culicoides* species used in this study, with a rapid decrease after the 32-hour time step. Similar results have been documented in mosquitoes [7], suggesting that rates of DNA degradation in digested blood meals can be similar in hematophagous taxa that differ markedly in blood meal size. Consistency in the digestion profiles of these taxa may reflect similar production patterns of digestive enzymes such as late trypsin, a protease responsible for the endoproteolytic cleavage of protein in the blood meal [21–23].

Our findings highlight the limited time frame over which samples are usable for blood meal analysis and the importance of timely sample capture and processing, and may partially explain variations in *Culicoides* blood meal amplification success in previous studies [10,24]. Whilst it was possible to amplify host DNA successfully at the later time steps, success was significantly reduced and may bias later analyses, such as examination of host preferences, and species roles in networks. Selection of trapping methods and sampling protocols should account for this to maximise the proportion of amplifiable blood meals. Specifically, we show that storage in 95% ethanol limits blood meal degradation of freshly fed biting Diptera even after prolonged periods of storage. Further, individuals should ideally be killed as soon as possible after feeding to halt digestive processes, using trapping methods that catch individuals directly into ethanol. Traps commonly used for *Culicoides,* such as OVI (Onderstepoort Veterinary Institute) light traps and CDC (Centers for Disease Control) light traps, use ethanol-filled collection chambers. However, mosquitoes are often collected with designs such as the BG Sentinel Trap, where collections are kept dry to facilitate

morphological identification of specimens. Therefore, if the primary goal is to identify host blood meals, it may be advisable to modify these traps to include an ethanol-filled chamber, although this may necessitate molecular identification of the sampled insect if immersion in liquid causes damage to scales used for morphology-based identifications.

Halting digestion and minimising degradation of host DNA post-capture is of particular importance for downstream blood meal analysis. We found no significant difference in amplification success between samples stored in 95% ethanol in a -20°C freezer versus at room temperature, suggesting that storage in 95% ethanol is sufficient to limit host DNA degradation in *Culicoides*. These findings are consistent with previous work which suggested that ethanol alone is sufficient to maintain host DNA integrity in mosquito blood meals [5,25]. The limited effect of storage at ambient temperatures is relevant for the planning and logistics of field work. Collection of *Culicoides* and other biting Diptera often necessitates sampling in remote locations without access to freezers; our results provide reassurance that, given proper processing, samples can safely be stored and transported at ambient temperatures for extended periods without impacting amplification success. The range of ambient temperatures used in our study reflect those in temperate regions, where monitoring of *Culicoides* species and limiting the spread of *Culicoides*-borne diseases is a major concern [24,26–30]. Ethanol has also been shown to be an effective short to medium term mosquito blood-meal storage method under higher-temperature tropical conditions [5]. This suggests that our results might also apply to *Culicoides* samples stored under ambient conditions in the tropics, though further work is required to confirm this; there is potentially an upper temperature limit where 95% ethanol does not sufficiently halt DNA degradation in blood meals.

We found slightly greater amplification success for blood meals of *Culicoides nubeculosus* compared with *C. sonorensis*. These *Culicoides* species are of similar size, making it unlikely that blood meal size was a factor [31]. Differences between the species were most apparent at a time estimated to correspond to stage 3 of *Culicoides* digestion [32], during which secretory granules associated with the production of digestive enzymes are formed. Thus, the differences in amplification success between the species may reflect differences in granule and digestive enzyme production. Alternatively, the higher mortality of *C. sonorensis* during the later time steps (64 hours onwards) and the differing amplification success for the

two species, may reflect minor differences in the abiotic conditions under which the individuals were reared, or slight differences in life history traits. For example, minor variations in ambient temperature can impact longevity [33] and rates of blood meal digestion [7].

General vertebrate PCR primers, when combined with high-throughput sequencing, allow identification of the range of host DNA template contained in mixed blood meals (metabarcoding). Moreover, general primers that target small amplicons are advantageous in blood meal analysis because of the high degree of DNA degradation resulting in fragmentation of the DNA template [1]. We identified a 200bp region of the 16s gene that is suitable for the interrogation of blood meals from a wide range of biting Diptera, whilst avoiding co-amplification of Dipteran DNA. Our primers and PCR protocol were effective with highly degraded DNA, successfully amplifying host templates during later time steps.

Conclusions

Based on the findings of this work, we propose the following recommendations. We suggest the use of the primers and the PCR protocol described here to complement existing primer sets in future biting Diptera blood meal analysis. Moreover, because of the efficiency and versatility of ethanol for maintaining blood meal DNA integrity and the need to halt digestive processes to maximise amplification success, we recommend the use of trapping methods that contain an ethanol filled collection device, though potential ethanol evaporation should be accounted for by regular trap emptying and replenishment of ethanol. During field sampling of blood fed biting Diptera, digestion of the blood meal may occur in the time between initial feeding and successful trapping. However, the placement of traps near potential sources of blood meals to limit digestion time may be unadvisable, depending on the overall aims of the trapping, as this will bias any host selection data [3]. These findings are relevant to future work that aims to investigate the biting behaviour and disease transmission potential of *Culicoides* (and likely other small haematophagous Diptera) through blood meal analysis.

4.6 Acknowledgements

We thank John Connolly, John Mumford and Silke Fuchs for their helpful comments and suggestions. We also thank Samantha O'Loughlin for providing blood fed *Anopheles* samples. Additionally, we also thank Robbie Bellekom for the design and creation of the graphical abstract. Finally, we would like to thank the two anonymous reviewers for the comments and suggestions.

Funding

This work was supported by a Natural Environment Research Council studentship to B.B. though the NERC-Oxford University Environmental Research Doctoral Training Partnership (NE/L002612/1). O.T.L., A.B, and T.D.H. were supported by Target Malaria. Target Malaria is supported by a grant from the Bill and Melinda Gates Foundation and the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

BB: conceptualisation, data analysis and writing, AB: review, technical advice, and guidance, ME: midge rearing, review, and editing, ZL: midge rearing, OTL and TDH: supervision, review, and editing.

Ethics approval and consent to participate

Not applicable

Consent for publication

All authors contributed to and gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests

We declare we have no competing interests

4.7 References

1. Reeves LE, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. 2018 Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. *PLoS Negl. Trop. Dis.* **12**, e0006767. (doi:10.1371/journal.pntd.0006767)

2. Martínez-de la Puente J, Ruiz S, Soriguer R, Figuerola J. 2013 Effect of blood meal digestion and DNA extraction protocol on the success of blood meal source determination in the malaria vector Anopheles atroparvus. *Malar. J.* **12**, 109. (doi:10.1186/1475-2875-12-109)

3. Bellekom B, Hackett TD, Lewis OT. 2021 A Network Perspective on the Vectoring of Human Disease. *Trends Parasitol.* (doi:10.1016/j.pt.2020.12.001)

4. Nielsen JM, Clare EL, Hayden B, Brett MT, Kratina P. 2018 Diet tracing in ecology: Method comparison and selection. *Methods Ecol. Evol.* **9**, 278–291. (doi:10.1111/2041-210X.12869)

5. Reeves LE, Holderman CJ, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. 2016 Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasit. Vectors* **9**, 503. (doi:10.1186/s13071-016-1791-z)

6. Ngo KA, Kramer LD. 2003 Identification of Mosquito Bloodmeals Using Polymerase Chain Reaction (PCR) With Order-Specific Primers. *J. Med. Entomol.* **40**, 215–222. (doi:10.1603/0022-2585-40.2.215)

7. Oshaghi MA, Chavshin AR, Vatandoost H, Yaaghoobi F, Mohtarami F, Noorjah N. 2006 Effects of post-ingestion and physical conditions on PCR amplification of host blood meal DNA in mosquitoes. *Exp. Parasitol.* **112**, 232–236. (doi:10.1016/J.EXPPARA.2005.11.008)

8. Mukabana WR, Takken W, Seda P, Killeen GF, Hawley WA, Knols BGJ. 2002 Extent of digestion affects the success of amplifying human DNA from blood meals of Anopheles gambiae (Diptera: Culicidae). *Bull. Entomol. Res.* **92**, 233–239. (doi:10.1079/BER2002164)

9. Seutin G, White BN, Boag PT. 2011 Preservation of avian blood and tissue samples for DNA analyses. *https://doi.org/10.1139/z91-013* **69**, 82–90. (doi:10.1139/Z91-013)

10. Lassen S, Nielsen SA, Kristensen M. 2012 Identity and diversity of blood meal hosts of biting midges (Diptera: Ceratopogonidae: Culicoides Latreille) in Denmark. *Parasites and Vectors* **5**, 143. (doi:10.1186/1756-3305-5-143)

11. Valinsky L, Ettinger G, Bar-Gal GK, Orshan L. 2014 Molecular Identification of Bloodmeals From Sand Flies and Mosquitoes Collected in Israel. *J. Med. Entomol.* **51**,

678–685. (doi:10.1603/me13125)

12. Orsborne J *et al.* 2019 Investigating the blood-host plasticity and dispersal of Anopheles coluzzii using a novel field-based methodology. *Parasit. Vectors* **12**, 143. (doi:10.1186/s13071-019-3401-3)

13. Hernandez-Colina A *et al.* 2021 Blood-feeding ecology of mosquitoes in two zoological gardens in the United Kingdom. *Parasites Vectors 2021 141* **14**, 1–18. (doi:10.1186/S13071-021-04735-0)

14. Navia-Gine WG, Loaiza JR, Miller MJ. 2013 Mosquito-Host Interactions during and after an Outbreak of Equine Viral Encephalitis in Eastern Panama. *PLoS One* **8**, e81788. (doi:10.1371/journal.pone.0081788)

15. Travassos Da Rosa JF, De Souza WM, De Paula Pinheiro F, Figueiredo ML, Cardoso JF, Acrani GO, Teixeira Nunes MR. 2017 Oropouche Virus: Clinical, Epidemiological, and Molecular Aspects of a Neglected Orthobunyavirus. *Am. J. Trop. Med. Hyg.* **96**, 1019–1030. (doi:10.4269/AJTMH.16-0672)

16. Boorman J. 1974 The maintenance of laboratory colonies of Culicoides variipennis (Coq.), C. nubeculosus (Mg.) and C. riethi Kieff. (Diptera, Ceratopogonidae). *Bull. Entomol. Res.* **64**, 371–377. (doi:10.1017/S0007485300031254)

17. Rozo-Lopez P, Park Y, Drolet BS. 2022 Effect of Constant Temperatures on Culicoides sonorensis Midge Physiology and Vesicular Stomatitis Virus Infection. *Insects* **13**. (doi:10.3390/INSECTS13040372)

18. Logue K, Keven JB, Cannon M V., Reimer L, Siba P, Walker ED, Zimmerman PA, Serre D. 2016 Unbiased Characterization of Anopheles Mosquito Blood Meals by Targeted High-Throughput Sequencing. *PLoS Negl. Trop. Dis.* **10**, e0004512. (doi:10.1371/journal.pntd.0004512)

19. Taylor PG. 1996 Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Mol. Biol. Evol.* **13**, 283–285. (doi:10.1093/oxfordjournals.molbev.a025566)

20. Wickham H. 2016 ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*

21. Lehiy CJ, Drolet BS. 2014 The salivary secretome of the biting midge, Culicoides sonorensis. *PeerJ* **2**. (doi:10.7717/PEERJ.426)

22. Noriega FG, Wells MA. 1999 Mini review A molecular view of trypsin synthesis in the midgut of Aedes aegypti. *J. Insect Physiol.* **45**, 613–620.

23. Filimonova SA. 2005 Morphological Study of Digestive Cycle in Bloodsucking Biting Midges of Genus Culicoides. *Transl. from Zhurnal Evolyutsionnoi Biokhimii i Fiziol.* **41**, 176–185.

24. England ME *et al.* 2020 Culicoides species composition and molecular identification of host blood meals at two zoos in the UK. *Parasit. Vectors* **13**, 139.

(doi:10.1186/s13071-020-04018-0)

25. Srisawat R, Sungvornyothin S, Jacquet M, Komalamisra N. 2014 Preserving blood-fed aedes albopictus from field to laboratory for blood source determination. *JItMM2013 Proc.*

26. Ninio C, Augot D, Delecolle JC, Dufour B, Depaquit J. 2011 Contribution to the knowledge of Culicoides (Diptera: Ceratopogonidae) host preferences in France. *Parasitol. Res.* **108**, 657–663. (doi:10.1007/s00436-010-2110-9)

27. Carpenter S, Groschup MH, Garros C, Felippe-Bauer ML, Purse B V. 2013 Culicoides biting midges, arboviruses and public health in Europe. *Antiviral Res.* **100**, 102–113. (doi:10.1016/J.ANTIVIRAL.2013.07.020)

28. Mignotte A *et al.* 2020 Important Dispersal Capacity of One Main Species Vector of Bluetongue and Schmallenberg Viruses Culicoides Obsoletus (Diptera: Ceratopogonidae) Revealed by Landscape Genetic Analyses. (doi:10.21203/RS.3.RS-75603/V1)

29. Tomazatos A, Jöst H, Schulze J, Spînu M, Schmidt-Chanasit J, Cadar D, Lühken R. 2020 Blood-meal analysis of Culicoides (Diptera: Ceratopogonidae) reveals a broad host range and new species records for Romania. *Parasites and Vectors* **13**, 79. (doi:10.1186/s13071-020-3938-1)

30. Kasičová Z, Schreiberová A, Kimáková A, Kočišová A. 2021 Blood meal analysis: host-feeding patterns of biting midges (Diptera, Ceratopogonidae, Culicoides Latreille) in Slovakia. *Parasite* **28**. (doi:10.1051/PARASITE/2021058)

31. O'Gower AK. 1956 The Rate of Digestion of human Blood by certain Species of Mosquitoes. *Aust. J. Biol. Sci.* **9**.

32. Filimonova SA. 2005 Morphological Study of Digestive Cycle in Bloodsucking Biting Midges of Genus Cu1icoides. *J. Evol. Biochem. Physiol. 2005 412* **41**, 221–232. (doi:10.1007/S10893-005-0057-8)

33. Lysyk TJ, Danyk T. 2007 Effect of Temperature on Life History Parameters of Adult Culicoides sonorensis (Diptera: Ceratopogonidae) in Relation to Geographic Origin and Vectorial Capacity for Bluetongue Virus. *J. Med. Entomol.* **44**, 741–751. (doi:10.1093/JMEDENT/44.5.741)

4.8 Supplementary

Additional file 1: Text S1. Primer protocol selection method

To examine primer annealing temperature and binding success, DNA was extracted from seven meats, covering a range of mammalian and avian families: beef, venison, pork, goat, duck, chicken and pigeon. Additionally, DNA from a male mosquito and blood fed and unfed *C. sonorensis* and *A. gambiae* was extracted (see methods for extraction and PCR chemistry).

To determine ideal annealing temperature for the primers, a gradient PCR was carried out for each meat. Samples from each meat with similar starting DNA concentrations were selected for use in gradient PCRs. An initial gradient PCR with a temperature range of 60-70^{o}C was conducted using beef samples. Following examination of the PCR success, all subsequent gradient PCRs were carried out with a temperature range of 58-65^{o}C.

Following selection of an ideal annealing temperature, a PCR was repeated for each meat extract using the same PCR reaction volumes and refined PCR protocols: 15 min at 95^{o}C followed by 35 cycles of 94^{o}C for 45 sec, 58^{o}C for 45 sec, and 72^{o}C for 30 sec followed by a final extension step of 72^{o}C for 10 min.

Chapter 5: The effect of anthropogenic landscape modification on biting Diptera-host community structure

Abutia Amagame village

Author contributions: BB: Data analysis and writing, OTL: Supervision, review, and editing, TDH: Supervision, review, and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Publication status: In preparation

Ben Bellekom[1*], Owen T. Lewis[1] and Talya D. Hackett[1]

[1]University of Oxford, Department of Biology, 11a Mansfield Road, Oxford OX1 3SZ, UK

* Corresponding author: ben.bellekom@biology.ox.ac.uk

BB: ben.bellekom@biology.ox.ac.uk

OTL: owen.lewis@biology.ox.ac.uk

TDH: talya.hackett@biology.ox.ac.uk

5.1 Abstract

The rapid increase in anthropogenic landscape modification for agricultural and urban expansion has significant epidemiological consequences, in particular, for developing nations, which are disproportionately at risk. Anthropogenic landscape modification may alter the composition of biting Diptera and host communities and increase contact between humans, domestic and wild animals, and biting Diptera. The change in the network structure of biting Diptera-host communities has potential implications for the transmission of vector-borne pathogens. Here we investigate the effects of anthropogenic landscape modification on biting Diptera communities and their hosts. We sampled biting Diptera from two field sites in Ghana and used DNA metabarcoding of Diptera blood meals to establish Diptera-host interaction networks. We found a greater Diptera and host species richness in habitat with lower levels of anthropogenic modification and high degree of overlap in biting Diptera species composition. Species diversity and evenness did not change among levels of anthropogenic modification. Greater values of network specificity were found in less modified habitat and connectance was greatest in heavily modified habitats. Humans were highly embedded in biting Diptera-host networks across all levels of landscape modification. Moreover, we identified several potential disease transmission pathways, linking competent vectors with susceptible hosts. Our findings are of epidemiological concern, as they empirically place humans at risk of introduction to pathogen transmission cycles through the movement of opportunistic bridge vectors, with significant consequences for the transmission of emerging and established zoonotic disease.

5.2 Introduction

Biting Diptera are responsible for the transmission of a wide range of pathogens that affect humans and wild and domestic animals, and represent a significant threat to public and economic wellbeing (Braack et al., 2018). Predominant vector-borne diseases include malaria, dengue, yellow fever, leishmaniasis, and African trypanosomiasis (Müller et al., 2019). An estimated one-sixth of all worldwide illness and disability is linked to vector-borne diseases, with over half of the global population currently at risk of infection (Campbell-Lendrum et al., 2015). Developing nations, particularly those in Africa and the tropics, are disproportionately at risk due to extensive on-going changes in their natural landscapes for agricultural and urban expansion, and the persistence of poverty (Chaves et al., 2020; Jones et al., 2013). Indeed, Africa shoulders the burden of infection, with 593,000 malaria related deaths (95% of the estimated global malaria deaths)(W.H.O, 2022) and over 831 million people (70% of the African continent) at risk of infection from vector borne diseases (Agboli et al., 2021; Benelli & Beier, 2017).

The prevalence of vector-borne disease, the emergence of novel zoonoses, and their impact on populations in developing nations is intrinsically linked to limited levels of socioeconomic development and healthcare availability (Franklinos et al., 2019). Rural populations that rely on subsistence farming are at higher risk of infection, due to the positive relationship between their primary means of income (such as livestock) and vector abundance (Bonds et al., 2010). Further, such populations are often situated at the interface between rural or urban landscapes and sylvatic habitats (habitat in which pathogens cycle between wild animals and vectors), resulting in a higher risk of exposure to established and emerging zoonoses. Zoonotic pathogens predominantly circulate among communities of wild or domestic reservoir hosts and Diptera vectors, with spillover events occurring as a result of human encroachment into sylvatic environments, increased contact with wild and domestic hosts, and interactions with infectious vectors (Gibb, Franklinos, et al., 2020).

Increasing anthropogenic landscape modification may directly and indirectly drive the loss, homogenisation, and turnover of biting Diptera and host communities. Modification of the landscape increases contact between humans, domestic and wild animals, and associated vectors, and impacts epidemiological processes (Gibb, Redding, et al., 2020; Mayi et al., 2019; Patz et al., 2004). Landscape modification is increasing rapidly with large amounts of

species-rich tropical forest habitat being converted into species-poor agricultural land (Patz et al., 2004; Shah et al., 2019) with a parallel increase in human interactions with wildlife and novel pathogens.

Anthropogenic modification fragments the structure of the landscape, producing smaller patches of sylvatic habitat surrounded by agricultural or urban land (Chaves et al., 2010). The resulting greater proportion of edges promotes increased interaction frequencies between vectors, hosts, and pathogens across this interface. Indeed, such edges act as ecotones, in which shared communities of hosts and biting Diptera from the sylvatic habitat interior and exterior are present, increasing the potential for opportunistic biting Diptera species to act as bridge vectors, linking agricultural/urban with sylvatic habitats, resulting in novel disease transmission to humans living and farming in these habitats (Hoyos et al., 2021; Meyer Steiger et al., 2016)

Host species diversity is known to decline as human land-use intensifies (Abella-Medrano et al., 2018). Such decreased host diversity is expected to increase disease transmission potential by eroding the dilution effect, where high host diversity limits the probability of a vector encountering a competent host (Civitello et al., 2015; Miller & Huppert, 2013). Consequently, habitats with limited host diversity (such as urban or agricultural land) are expected to have higher zoonotic transmission and spillover, especially if these habitats are dominated by competent intermediate or amplifier hosts (organisms in which a pathogen may replicate rapidly and be maintained at a high concentration) (Franklinos et al., 2019; Jones et al., 2013).

Landscape modification can also have a significant effect on the diversity and composition of biting Diptera communities. Anthropogenic landscape change alters the conditions of the affected habitat through a variety of mechanisms including by increasing local temperatures, modifying the distribution and availability of standing water, and changing host availability and vegetation coverage (Chaves et al., 2021; Ferraguti et al., 2016; Li et al., 2014). As biting Diptera species distributions and abundances are linked to their habitat requirements, feeding behaviour, and larval habitat preferences, there is likely to be a large heterogeneity in the effects of such landscape changes on different biting Diptera species (Perrin et al., 2022). For example, increased local temperatures and the abundance of standing water in agricultural and village habitats decreases larval development time and provides breeding

opportunities outside of seasonal precipitation cycles, allowing Diptera that are adapted to anthropogenically modified landscapes to proliferate at the expense of other Diptera species (Gratz, 1999; Kesavaraju et al., 2008; Mattah et al., 2017). Moreover, increased landscape modification could be epidemiologically relevant, as the increased abundance of anthropogenically adapted Diptera in modified habitats favours increased interspecific interaction with abundant susceptible hosts, such as humans and cattle, with a resultant increase in potential pathogen transmission events (Chaves et al., 2021; Perrin et al., 2022; van Hoesel et al., 2019). Consequently, the dynamics of changing landscape conditions could homogenise and alter the ecology of the biting Diptera and host community, increase transmission and spillover potential, and alter trophic interactions (Chaves et al., 2021; Ferraguti et al., 2016; van Hoesel et al., 2019).

Biting Diptera-host feeding interactions can now be routinely assessed through the molecular analysis of insect blood meals using DNA metabarcoding. Metabarcoding uses high-throughput sequencing to facilitate the parallel identification of a large number of species in a mixed sample using short DNA amplicons, making the approach highly suitable for samples containing degraded DNA, such as Diptera blood meals (Borland & Kading, 2021). Ecological networks, populated with such interaction data, provide the opportunity to describe the interdependencies between sets of biting Diptera and their hosts (Evans et al., 2013; Fath et al., 2007; Proulx et al., 2005) and quantify the impact of potential shifts in interactions, and thus change in network structure, caused by altered species composition as a result of habitat modification (Evans et al., 2016).

We used bloodmeal metabarcoding to generate new data on interactions between biting Diptera and their hosts from two rural villages in Ghana. We assessed the effect of differing levels of anthropogenic landscape modification on biting Diptera species diversity, community dissimilarity, and species evenness. Further, we constructed biting Diptera-host interaction networks to examine how anthropogenic landscape modification and affected network connectance (the proportion of realised interactions among all potential ones), network specialisation (the deviation of observed interactions from that expected at random), and the centrality of humans (their relative importance within the host-parasite network).

5.3 Methods

Insect sampling and preliminary identification

Field sites

We sampled insect around two rural villages within the Volta region, Ghana. No active biting Diptera control initiatives were in place during the sampling period, although some residents sleep under insecticide treated bed nets.

Ghana has a wide diversity of biting Diptera capable of transmitting a range of vector-borne diseases, including malaria (Awine et al., 2017), Yellow Fever (Captain-Esoah et al., 2020), and leishmaniasis (Akuffo et al., 2021). Indeed, Ghana accounts for over 4% of the global malaria burden (Agyemang-Badu et al., 2023). Approximately 68% of the population of Ghana reside in rural habitats and over 52% of the country's labour force is engaged in agriculture (Food and Agriculture Organization of the United Nations., 2023). Expansion of agricultural land accounted for 78% of forest loss over the last 29 years, with over 85% of the Guinea Forest Region being degraded due to anthropogenic activity (Acheampong et al., 2019). Consequently, increasing anthropogenic landscape modification and human encroachment into forest habitat is likely to result in a significant proportion of the Ghanian population living and working in the interface between rural and sylvatic habitat, with implications for the community structure of biting Diptera, their associated host interactions, and disease transmission potential.

Abutia Amagame (AA) (Figure 1.) (6.459064°N, 0.316247 °E), is a small rural village situated on a hillside, surrounded by mostly seminatural vegetation. Cultivated land is primarily used for maize and cassava. There is a high representation of natural vegetation, and the village is surrounded by a mixture of forest and grassland. A dirt road is present, though traffic activity is minimal.

Figure 1. Satellite image of Abutia Amagame and surrounding habitat (Google Earth Pro 2023)

Mafi Agove (MA) (Figure 2.) (6.208310°N 0.442600 °E), is the larger of the two villages. It consists of a broadly flatland agricultural landscape matrix, fragmented by small clusters of housing and a large village. The surrounding cultivated landscape consists primarily of maize and cassava fields, intersected by dirt tracks and dirt and paved roads, with minimal traffic activity and disturbed forest and savannah grassland.

Figure 2. Satellite image of Mafi Agove and surrounding habitat (Google Earth Pro 2023)

Sampling timing and field collection methods

Sampling was initially scheduled to coincide with the southern rainy season, however, due to COVID-19 restrictions limiting access to the field sites, samples included in this study were collected between July 2019 and February 2020. We developed a sampling protocol in collaboration with University of Ghana and members of the Target Malaria research consortium to complement their existing sampling design, and conducted field sampling from July-August 2019. Sampling was then continued by Ghanian field technicians under close supervision from August 2019-February 2020.

In line with the existing sampling protocol, sites were divided into four quadrants (NE, NW, SE, SW) and random sampling locations were distributed across each quadrant within a 500m radius. The established sampling regime followed a one week on, two weeks off rotation, with a sampling week being comprised of a 24-hour collection period in AA, followed by an additional 24-hour collection period in MA. During a sampling period, we left traps to run from morning through to the following morning to sample diurnal, crepuscular, and nocturnal species.

For each sampling period, we placed traps at four randomly selected locations, one from each quadrant. As the primary aim was the analysis of blood meals, immediate preservation of the host template DNA was required to minimise DNA degradation that would impede further analysis. Consequently, we captured biting Diptera using four CDC (Centre for Disease Control and Prevention) light traps (Bioquip, Rancho Dominguez, CA) modified to include a 95% ethanol-filled collection chamber (Bellekom et al., 2023). Individual CDC light traps were powered by a single 6V battery (Bioquip Rancho Dominguez, CA) capable of powering the trap for approximately 37.5 hours. Each CDC light trap used a single incandescent light bulb with an approximate attraction range of between 5 and 15m, depending on the Diptera taxa (Kirkeby et al., 2013; Wilson et al., 2021). We hung each CDC trap from a custom-made stand 1.3 metres above the ground and supplemented it with a sugar fermented yeast CO_2 bait, consisting of 1L of water, 100g of sugar and 7g of yeast within a 1.5L bottle, attached to the trap entrance (Figure 3.) (Jerry et al., 2017; Smallegange et al., 2010). We carried out the preliminary preparation of each bait bottle ex-situ, prior to use; the yeast was added immediately prior to trap placement to ensure the high CO_2 flow rate. At each site, a broad habitat description was noted.

Figure 3. CDC light trap set up. This example uses a net collection chamber and was used only during the first sampling period due to logistical issues limiting the delivery of light traps with ethanol collection chambers. All further sampling was conducted using the ethanol collection chambers traps.

Sample processing

Following the 24-hour collection period, we retrieved the samples in the order in which traps were placed, to ensure equal collection time for each site. We stored the collected samples in a field refrigerator for the duration of the sampling week and transported in cold boxes, before storage within a -20°C freezer at the University of Ghana. We then shipped the samples in cold boxes to the University of Oxford Biology department to be sorted. Prior to molecular identification to species level, samples were preliminarily identified to Order and assigned a morphospecies. Up to five representatives of each morphospecies were placed in 96 well microplates containing 95% ethanol, and the total abundance of each

morphospecies per sampling event was counted. Blood fed (individuals that had a visible blood meal) and potentially blood fed (gravid and parous individuals, and those with almost entirely digested blood meals) individuals were visually identified and separated into their own plate for further blood meal analysis. All plated samples were shipped in cold boxes to the Centre for Biodiversity Genomics facility at the University of Guelph for sequencing on the PacBio Sequel II and Miseq platforms.

Blood meal analysis

All samples (morphospecies representatives, blood fed, and potentially blood fed) were processed using automated pipelines established by the Canadian Centre for DNA Barcoding. Diptera and blood meal DNA were handled by a Biomek FXP liquid handling system and extracted using a glass fiber extraction protocol, described by (Natalia Ivanova et al., 2006).

All samples were identified using the PacBio Sequel II platform, targeting the CO1 gene loci. Host identification was conducted using both the PacBio Sequel II, targeting the CO1 loci, and the Illumina MiSeq platform, targeting the 16S loci. Host identification using the Sequel II platform was experimental and guided by previous unpublished blood meal results (personal communication from Evgeny V. Zakharov, University of Guelph), which indicated the platform may be used for host discovery. Samples were processed on the Sequel II platform using single molecule real-time (SMRT) sequencing. Amplicons were generated using the SMRTbell Express Template kit 2.0, Sequel II Binding kit 2.1 and Sequel II DNA Internal Control 1.0, following the protocol described by (Hebert et al., 2018).

Amplicons for host identification created on the Illumina MiSeq platform were generated using a standard v3 chemistry, and the 16S primers and PCR protocol described by (Bellekom et al., 2023).

Bioinformatics

To identify the biting Diptera blood meal origin and establish Diptera-host interactions, PacBio Sequel II CO1 blood meal read data were initially processed using the mBRAVE platform. Sequences were trimmed and filtered using the default parameters and clustered using a 97% ID distance and OTU exclusion threshold and a 98% OTU threshold. Identified chimeric sequences, false reads, and controls were manually removed. A custom CO1

reference library was created using the python package NSDPY (Hebert & Meglécz, 2022) and blastn (Wheeler & Bhagwat, 2007). Search parameters included all families of expected taxa in our sampling region. Expected taxa were identified using a combination of the Global Biodiversity Information Facility (GBIF) (*GBIF.Org,* 2021) and the text resources 'Mammals of Ghana, Sierra Leone and the Gambia' (P. Grubb, 1998) and 'Field guide to the Birds of Ghana' (Demey & Borrow, 2010). Search terms 'CO1' and 'COX1' were included to provide as many reference sequences for the target loci as possible. The reference sequences were converted to a useable nucleotide database using the makeblastdb function. Sequences were assigned a taxonomic identify using the blastn function, with word size set to 28 and max target seq to 1, and a 98% sequence identity threshold.

During Sequel II sequencing, individual samples may generate multiple contigs. Where an individual sample was represented by multiple contigs with a single read, the singletons were removed as they are likely as result of contamination. Species that were represented by singletons that were not present in other samples in the SMRT- cell (i.e., found once in the dataset) were retained as they were likely to be true results and not introduced through contamination. Where a sample was represented by multiple contigs with different species assignments, contigs were retained only if they were not singletons (unless the singleton was not present in any other sample), were of expected fragment length (~180bp), and had a sequence identity match of over 98% (Hopken et al., 2021). Where a sample could not be identified to species level, the highest taxonomic resolution was assigned. Samples that retained contigs of different species origin after filtering were classified as having a mixed blood meal. Preliminary results indicated a moderate abundance of reads of human origin in our samples. Due to the inability to distinguish between true human positive blood meals and false positives resulting from potential human contamination, a stricter filtering process was employed for these samples. A threshold >10 reads was implemented for human positive samples. The threshold was selected based on the average read counts of true positive samples (~11), such as those positive for species likely to be in our study sites (e.g., *Philantomba maxwellii*).

Identification of the Sequel II blood fed (individuals with a blood meal) and non-blood fed biting Diptera (individuals collected that did not contain a blood meal) read data was conducted using the same mBRAVE protocol. Species were assigned during the mBRAVE

pipeline. Where a sample was not identified as result the parameters of the pipeline, the associated read was manually BLASTed. Where a sample could not be identified to species level, the highest taxonomic resolution was assigned with a unique identify to prevent multiple species being represented by a single family identification in downstream analysis. Abundances were assigned to each occurrence of a species using the previously estimated abundance for each morphospecies. Where a morphospecies was represented by multiple species, the abundance of each component species was assigned based on their overall relative proportion.

Illumina MiSeq Bioinformatics

All MiSeq data were handled in a custom Nextflow pipeline (DI Tommaso et al., 2017). For each sequencing run, files were demultiplexed (maximum error rate 0.15) and the 16S primers were trimmed using cutadapt. The resulting fastq files were filtered by their error scores (a threshold maximum of 4 errors per read). The error rates of each sequencing run was learned using the DADA2 package (Callahan et al., 2016), which was then used to merge the amplicons of each sequencing run into Amplicon Sequence variants (ASVs). A custom 16S reference library was created using the same protocol previously described, and ASVs were assigned a taxonomic identity using the blastn function, with word size set to 28 and max target seq to 1, and a 98% sequence identity threshold. Preliminary examination of the subsequent data set revealed a hyper abundance of human sequences, the presence of species unlikely to be found at our field sites, such as the Indian gecko species *Cnemaspis nilagirica*, and an unexpectedly high number of blood meals, indicating that contaminants had been introduced. Consequently, due to limited reliability of these data no further analysis was conducted on them.

Data analysis

To examine the effect of landscape modification on biting Diptera community and interaction structure, each sampling site was assigned a landscape using a mixture of the initial vegetation data and satellite images (QGIS with XYZ tile Google Map base layer from the appropriate sampling year and Google Earth) which, together, provided sufficient resolution to identify levels of landscape modification and vegetation structure. Sites where cultivated land or livestock were the dominant land-use were categorised as Agricultural;

those that contained natural vegetation with limited human presence were classified as Near-natural; and sampling locations in or around human habitation were classified as Village.

For each sampling site and classification, we calculated a range of descriptive and quantitative metrics and measures. Host species were assigned a classification based on whether they were human, domestic, or wild species. The total number and distribution across site and habitat of fed and non-blood fed Diptera species, and host species was examined. The overall proportion of blood fed individuals for each species was assessed; and the proportion of mixed blood meals was calculated. In addition, we explored the effect of habitat type on log transformed Diptera Shannon-Wiener diversity using a generalized linear model with a Gaussian distribution. The statistical significance of habitat type were assessed by comparing nested models for variation in deviance based on a chi-square distribution (Mayi et al., 2020). Post hoc analysis was conducted to identify intra-significant differences using a Tukey's HSD (honestly significant difference) test. Further, we examined the dissimilarity of the biting Diptera community composition within each habitat and site, using the adonis2 function which implements a Permutational Multivariate Analysis of Variance Using Distance Matrices (PERMANOVA) (Dixon, 2003).

Interaction data across sampling sites, for all habitat classifications, were compiled to create a weighted bipartite network which represented each habitat classification. For each habitat classification we calculated network and species level metrics chosen for their ecological and epidemiological relevance. To examine the complexity of our networks, we calculated their connectance, the proportion of realised ecological interactions among the potential interactions (Poisot & Gravel, 2014). Network specialisation for each habitat classification was examined using H2', which quantifies the deviation of observed interaction frequencies from expected frequencies if interactions were random (Blüthgen et al., 2006). As generalist species have the potential to act as hub species that promote cross-species disease transmission, networks with a high a degree of generalism may sustain a greater disease transmission potential between unrelated hosts (Su et al., 2022). To assess the risk of disease exposure potential for humans in each habitat type, we calculated their degree and closeness centrality for each network. Degree centrality calculates the number of connections the target node has in a network. As the within species prevalence of a disease

increases linearly with their degree centrality, nodes with a high degree are more likely to participate in disease transmission events (Su et al., 2022). Closeness centrality describes the mean path length from the target node to all other nodes in the network, with a high value indicating that the node may be rapidly affected by other nodes in the network and vice versa (Silk et al., 2017; Su et al., 2022). Further, to examine the relative importance of a node in connecting parts of each network and their potential to mediate transmission throughout the network, we calculated the betweenness centrality, of the biting Diptera in each of our networks. Betweenness centrality describes the proportion of shortest pathlengths in a network that go via the node, with a high value indicating that the node is important for the cohesiveness of the network (Llaberia-Robledillo et al., 2022).

While our data do not document pathogens present within the blood meals, we inferred disease vectoring potential of each blood fed Diptera species using Vectorbase (Giraldo-Calderón et al., 2015), and highlighted potential transmission events for susceptible hosts. Further, to assess sampling completeness, we constructed interpolation and extrapolation curves for hosts, biting Diptera and their interactions as a function of sampling effort (the number of blood meals collected) using the iNEXT package (q=0, data type = incidence frequency) (Hsieh et al., 2016). Finally, we estimated total interactions, and host and Diptera species richness using the ChaoRichness function (Hsieh et al., 2016), which is based on methods proposed by (Chao, 1987).

5.4 Results

Overview

In total we collected 7,095 biting Diptera from 30 collection sites (12 Agricultural, 12 Near-natural, and six Village sites) (Figure 4. & Supplementary 1.). We identified 42 individual Diptera species (84% of nodes) and resolved an additional eight Diptera to genus or family level (16% of nodes) (Supplementary 2.). Blood fed individuals were collected from 20 sampling locations: 6 Agricultural, 11 Near-natural, and 3 Village sites. Blood meal sequencing on the PacBio Sequel II platform generated an initial 189,240 reads. After filtering and removal of chimeras this was reduced to 184,938 reads. Approximately half (95,772) of our reads were assigned to a BIN, whilst 89,166 were placed in 11,016 OTUs.

Following progression through our BLAST pipeline, using our custom reference library for host species identification, this was reduced to 111,449 reads. Further, after implementing our strict filtering criteria we retained 3,151 reads which characterised the composition of 75 blood meals and identified 18 individual host species derived from blood meals of 29 biting Diptera species.

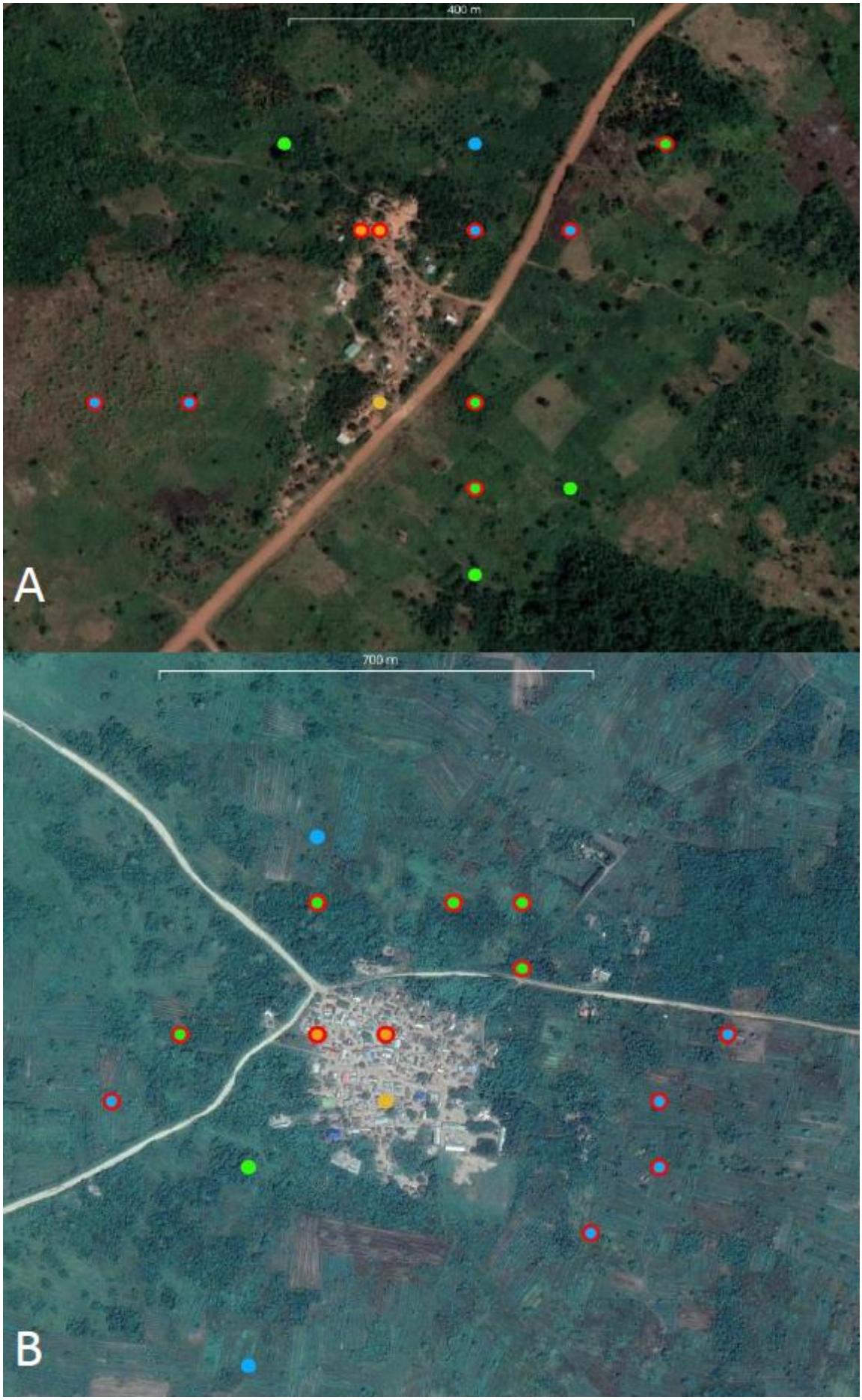

400 m
A
700 m
B

Figure 4. Satellite image sampling locations in Abutia Amagame (A) and Mafi Agove (B) separated by habitat classification: Village (orange), Agricultural (blue), and Near-natural (green). Points bordered by a red ring indicate sites in which blood fed Diptera were caught.

Table 1. Non-blood fed and blood fed Diptera richness, number of blood meals, and host richness by site.

Site/Habitat	Non-blood fed Diptera richness	Non-blood fed Diptera abundance	Blood fed Diptera richness	Number of blood meals	Host richness
Abutia Amegame	30	3,405	15	27	6
Mafi Agorve	33	3,690	16	48	14

Table 2. Non-blood fed and blood fed Diptera richness, number of blood meals, and host richness by level of anthropogenic landscape modification.

Site/Habitat	Non-blood fed Diptera richness	Non-blood fed Diptera abundance	Blood fed Diptera richness	Number of blood meals	Host richness
Village	11	1,382	4	8	2
Agricultural	29	1,454	11	28	9
Near-natural	34	4,259	22	39	11

Effect of habitat modification

Non-blood fed and blood fed Diptera richness and abundance was lower in Abutia Amegame than in Mafi Agorve (Table 1). Further, total non-blood fed and blood fed Diptera richness and abundance, and host richness was lower in the village habitats compared to Agricultural and Near-natural habitat (Table 2.). We found no significant difference in biting Diptera species evenness between Village (mean evenness= 0.109, SE= 0.04) and Agricultural (mean evenness= 0.272, SE= 0.06, Tukey; p= 0.207), or Near-natural (mean evenness= 0.220, SE= 0.04, Tukey; p= 0.374). Moreover, we found no difference in species evenness between Near-natural and Agricultural habitat (Tukey; p=0.882).

We found no significant effect of habitat modification (χ^2=0.29178, df=2, p=0.186) (Figure 5.) on biting Diptera diversity. Whist mean H diversity was lowest in Village habitats (mean=

0.158, SE= 0.068), it did not differ significantly to those in Agricultural (mean=0.427, SE= 0.097) and Near-natural habitat (mean= 0.461, SE= 0.083). There was no significant difference in H diversity between Agricultural and Near-natural habitat.

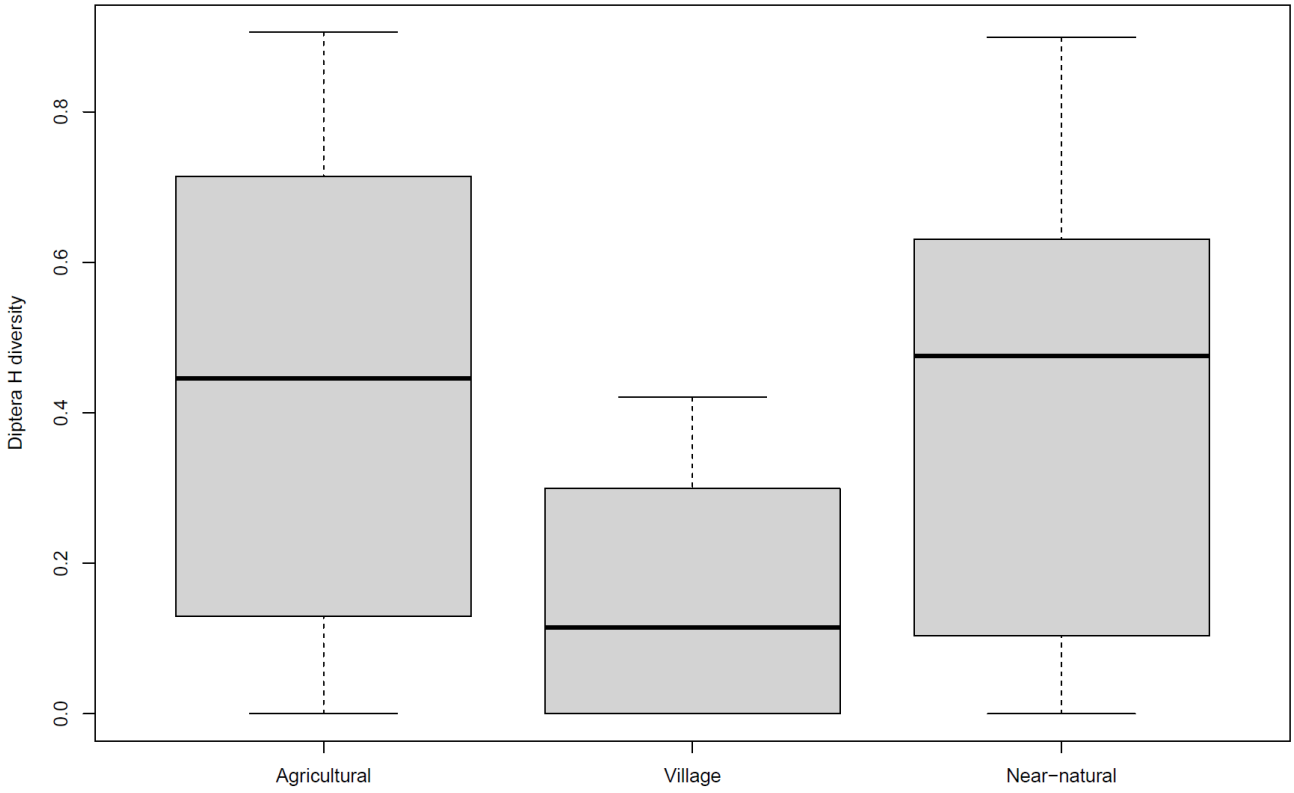


Figure 5. Diptera H diversity by habitat type. Each box displays the interquartile range, and the solid line represents the median. Whiskers display the maximum and minimum interaction evenness and H values for each habitat type.

The biting Diptera community composition did not differ significantly across habitat type ($F_{2,27}$=1.09, p=0.306) (Figure 6.) or site ($F_{1,28}$=0.862, p=0.611)(Figure 7.).

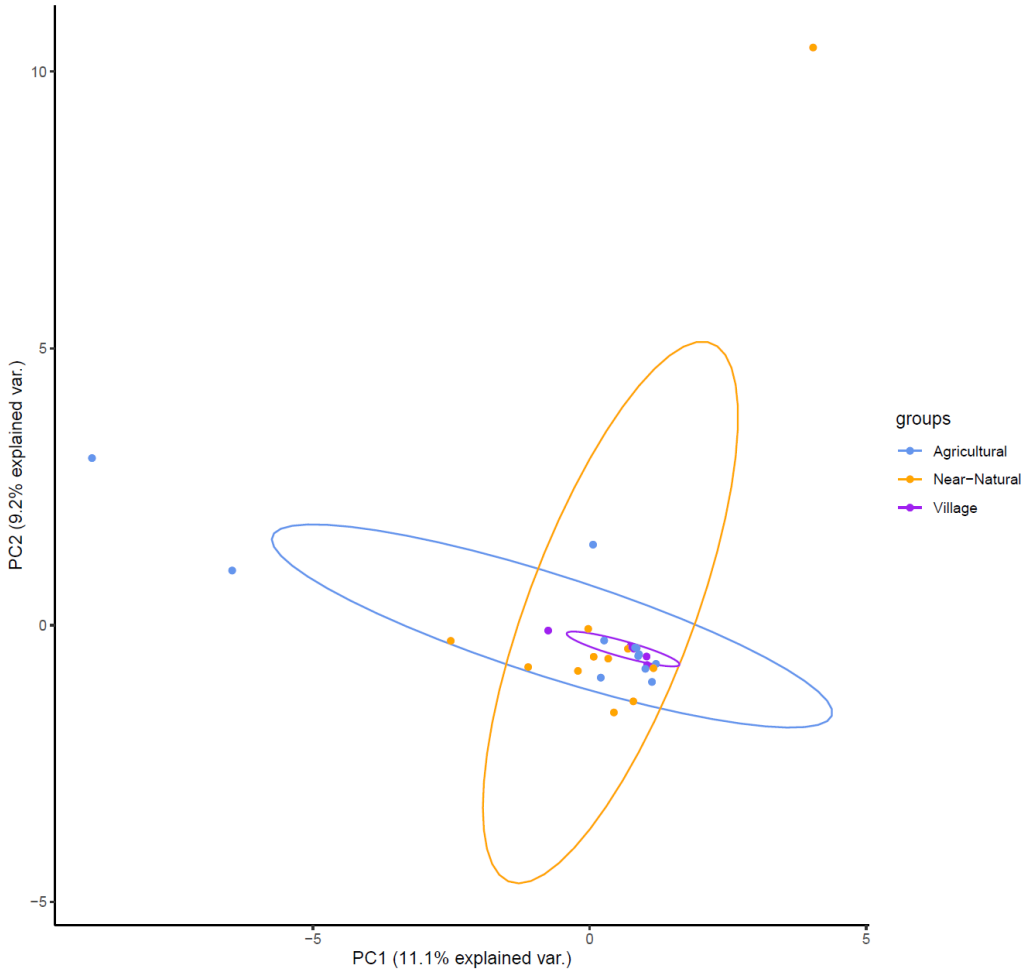


Figure 6. Ordination of the overall biting Diptera community composition by habitat category. Each point represents the biting Diptera diversity captured by a unique sampling event in the Village (purple) Agricultural (blue), and Near-natural (orange) habitat.

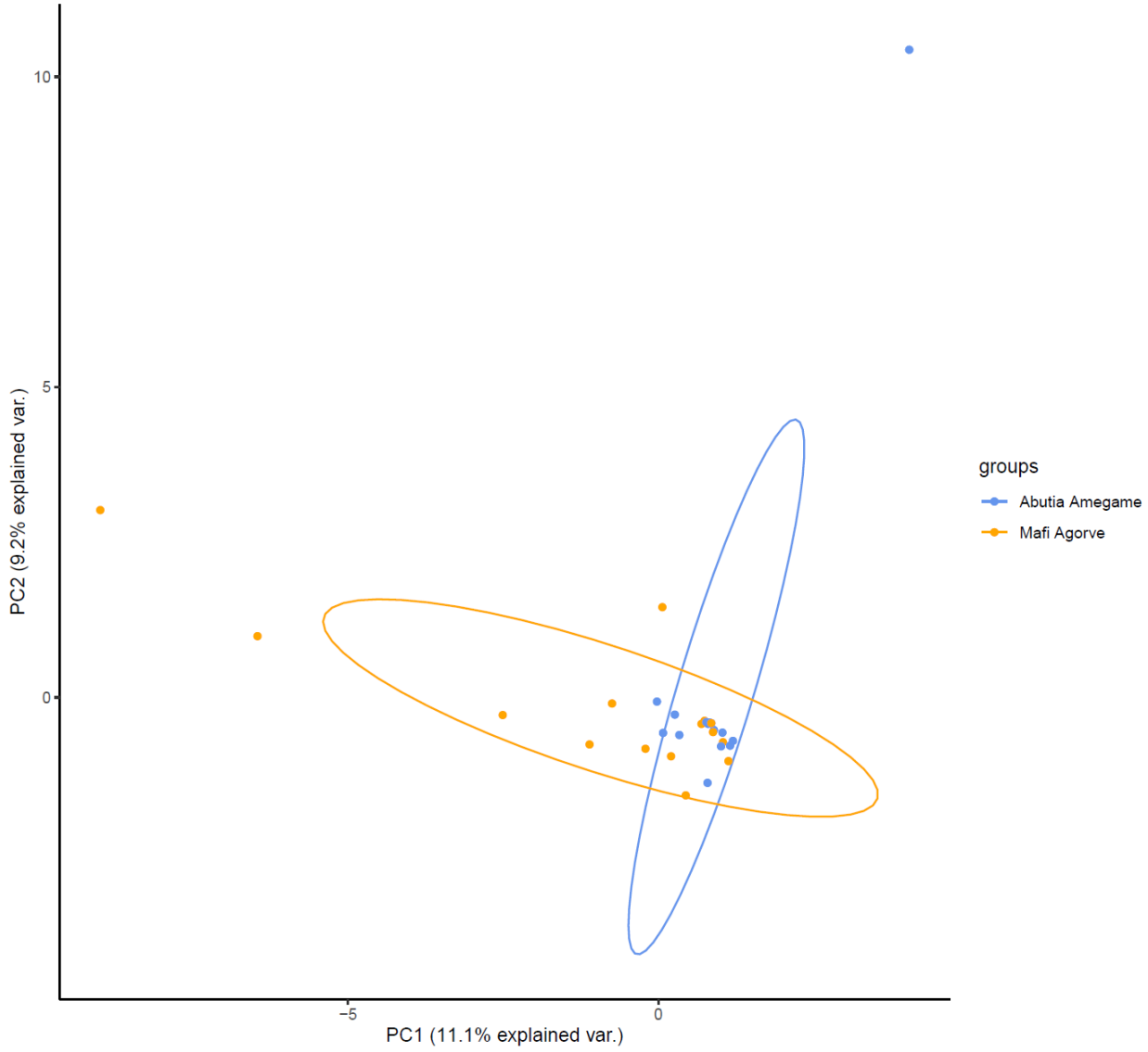


Figure 7. Ordination of the overall biting Diptera community composition by site category. Each point represents the biting Diptera diversity captured by a unique sampling event in Abutia Amagame (blue) and Mafi Agove (orange).

Blood meal origin

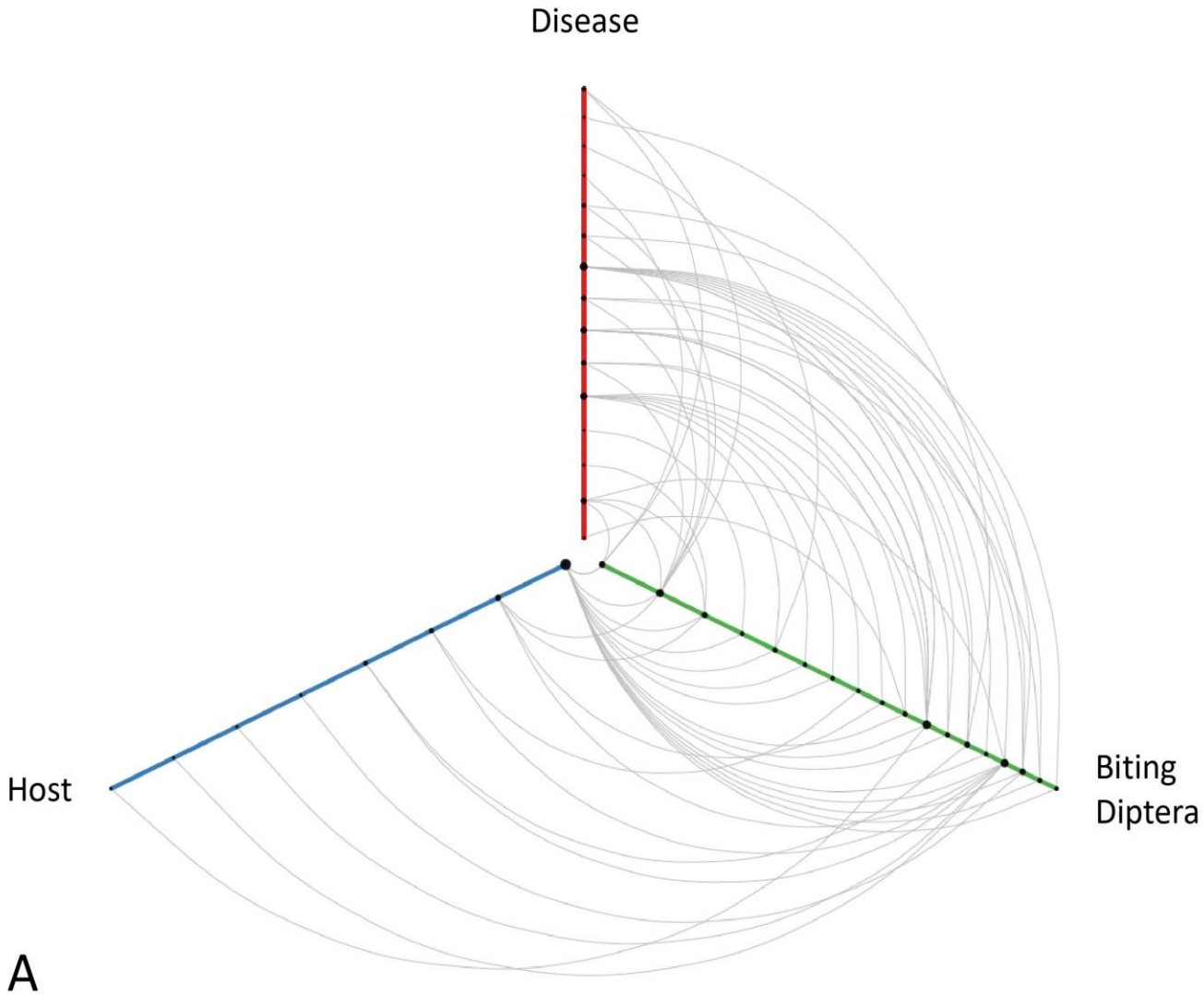
Humans were had the highest number of interactions across all sites and habitat classifications; human DNA was detected in 68% of all collected blood meals. The number of wild host species varied greatly between the two sampling sites, with 10 species in Mafi Agorve and just one in Abutia Amegame, with minimal difference in domestic species richness (3 and 4 species, respectively). We identified a slightly greater number of domestic

and wild hosts in Near-natural habitat (4 and 6) than in Agricultural habitat (3 and 5). We only identified interactions with human, and one mixed blood meal with containing host DNA from human and a wild species (*Coturnix* sp.) in Village habitats.

Mixed blood meals (those containing DNA from more than one host origin) accounted for 7.89% (6) of the total collected blood meals. The highest number of mixed blood meals were found in the Near-natural habitat (7.69%; 3), in comparison to Agricultural (7.14%; 2) and Village (12.5%; 1). The majority (83%) of mixed blood meals contained human host DNA. Additionally, we identified two mixed blood meals containing DNA from three host origins (Soricidae, Fringillidae, and Phasianinae) and (*Capra hircus*, *Ovis aries*, and *Homo sapiens*).

Potential disease transmission events

Of the 29-blood fed biting Diptera species collected, 18 (62%) were competent of vectoring one or more pathogen of medical or veterinary importance. Of the vector competent species, 16 (88%) were found to have interacted with at least one host that was susceptible to the pathogens they vector (Figure 8.).



A

Species	Diseases	Hosts
Sergentomyia distincta	Leishmaniasis	*Homo sapiens, Crotalinae spp.*
Culicoides imicola	Bluetongue disease, African Horse Sickness, Schmallenberg	*Bos taurus, Homo sapiens*
Culex watti	West Nile virus	*Homo sapiens, Coturnix spp., Bos taurus, Pternistis bicalcaratus, Gallus gallus*
Mansonia africana	Lymphatic filariasis, Rift Valley fever, West Nile virus	*Homo sapiens, Tyto alba, Sus scrofa*
Culex decens	Rift Valley fever, West Nile virus	*Homo sapiens*
Culex univittatus	Rift Valley fever, West Nile virus, Japanese encephalitis	*Turdus pelios*
Coquillettidia metallica	West Nile virus, Middelburg, Sindbis fever	*Homo sapiens, Milvus migrans*
Aedes bromeliae	Chikungunya, Yellow fever	*Homo sapiens*
Culicoides distinctipennis	Bluetongue disease	*Homo sapiens, Bos taurus*
Sergentomyia ingrami	Leishmaniases	*Homo sapiens*
Culex nebulosus	West Nile virus, Middelburg, Ntaya virus	*Homo sapiens*
Culex cinereus	Sindbis fever	*Homo sapiens*
Sergentomyia africana	Leishmaniases	*Homo sapiens*
Culex perexiguus	Avian malaria, West Nile virus, Rift Valley fever	*Homo sapiens*
Culex vansomereni	West Nile virus	*Homo sapiens*
Sergentomyia inermis	Leishmaniases	*Bos taurus*

B

Figure 8. (A) Hive plot showing how potential transmission events between hosts (blue), biting Diptera (green) and their respective diseases (red). Nodes along each axis scale with frequency. (B) Interaction and disease competence data used to produce the Hive plot. Each row represents a collected blood fed Diptera species, the diseases they are competent vectors for and the susceptible hosts they interacted with, as determined by molecular blood meal analysis. Row, disease, and host order are not representative of relative importance. As we did not collect data on pathogen occurrence, these data represent hypothetical transmission events based on interactions between competent vectors and hosts.

Blood fed Diptera, the diseases they are competent vectors of, and the hosts they were found to interact with. Those Diptera species highlighted in yellow represent an interaction by a competent vector with a host that is susceptible to pathogens they vector. As we did not collect data on pathogen occurrence, these data represent hypothetical transmission events based on interactions between competent vectors and hosts.

Community structure

Accumulation curves showed that host and biting Diptera species richness for Agricultural and hosts richness for Near-natural habitat were relatively well resolved, with curves approaching asymptote in each case. Sampling for Diptera species in Near-natural habitat and species interactions in all habitats were incomplete (Figure 9. & Supplementary 3.). Due to the lack of interaction data, Village habitats were excluded from sampling completeness analysis.

All interaction data were compiled into an overall summary network, and separate networks were also constructed for each habitat category (Figures 10-11.). The connectance of the overall network was 0.088. Whilst no statistical testing of network metrics was possible due to lack of replication, Village habitat had a higher connectance value (Connectance= 0.625) compared to Agricultural (Connectance= 0.191) and Near-natural (Connectance= 0.123) habitat categories.

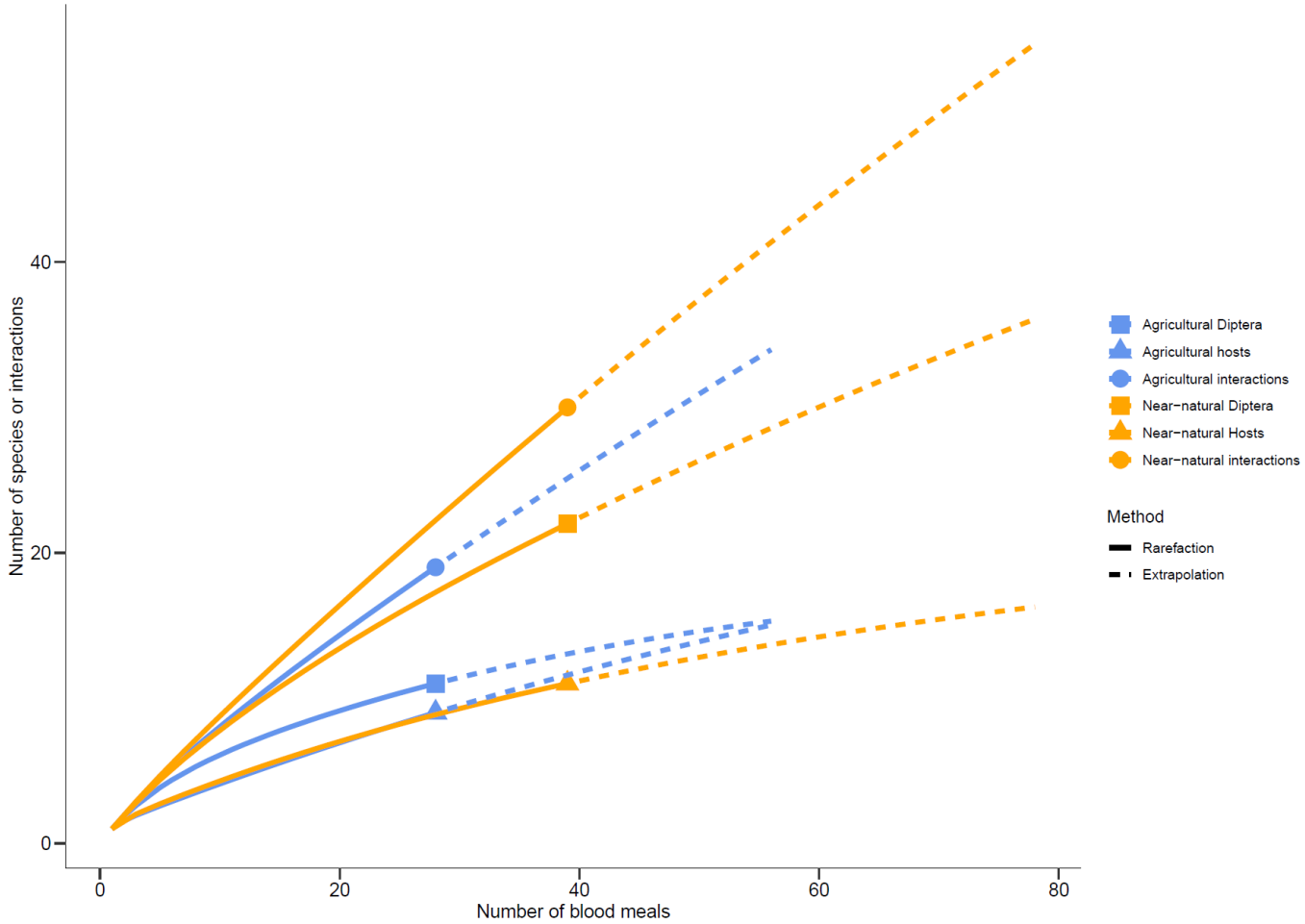


Figure 9. Smoothed accumulation and extrapolation curves to assess sampling completeness. Total numbers of hosts (triangle), biting Diptera (square), and interactions recorded in the whole dataset (circle), by habitat type: Agricultural (blue), Near-natural (orange), as a function of sampling effort (the number of blood meals screened). Village habitat was excluded from this analysis due to under sampling.

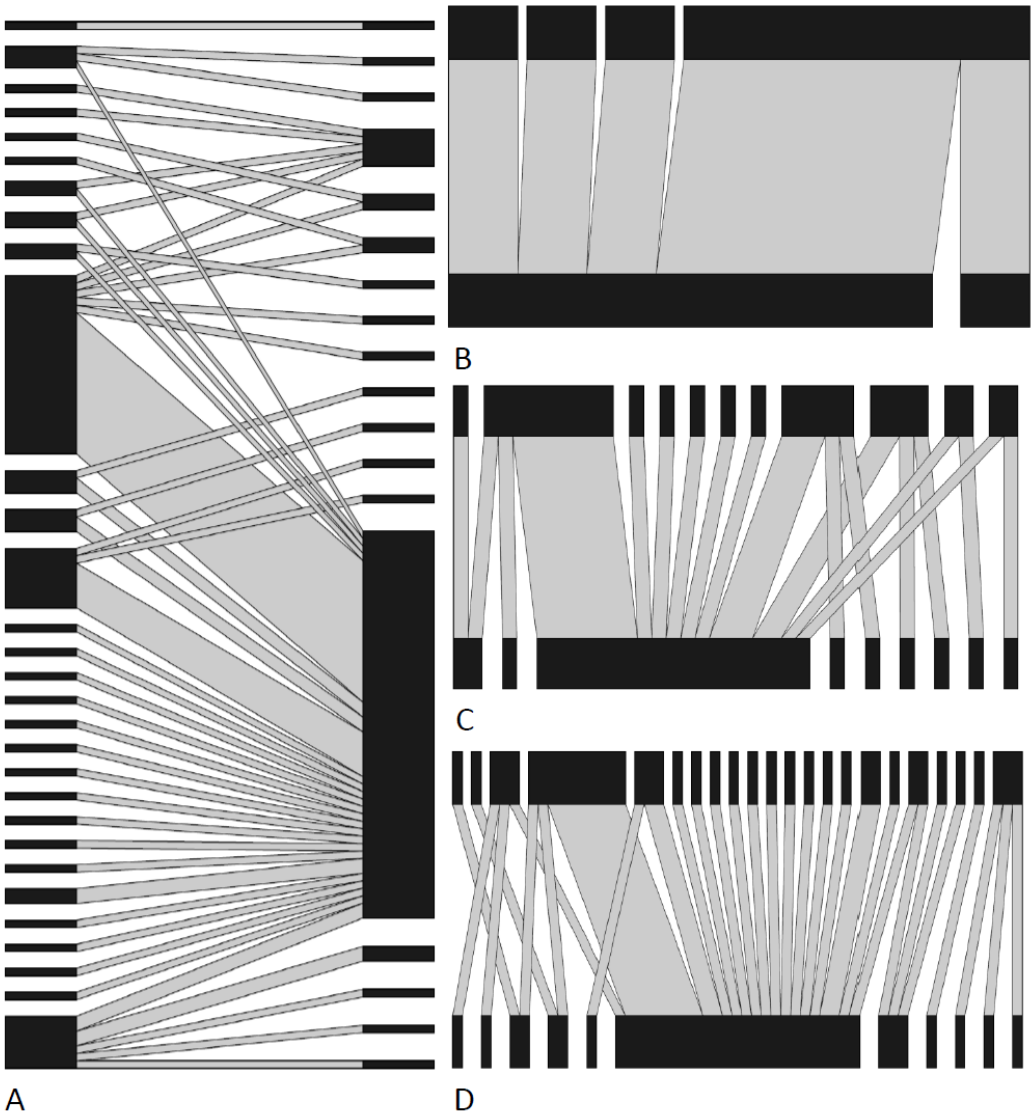


Figure 10. Combined habitat networks. (A) The aggregated whole network containing all host (right) and Diptera (left) interactions, and Diptera (top)-host (bottom) interactions separated by habitat classification: Village (B), Agricultural (C), and Near-natural (D). Node and edge widths are proportional to frequency of occurrence.

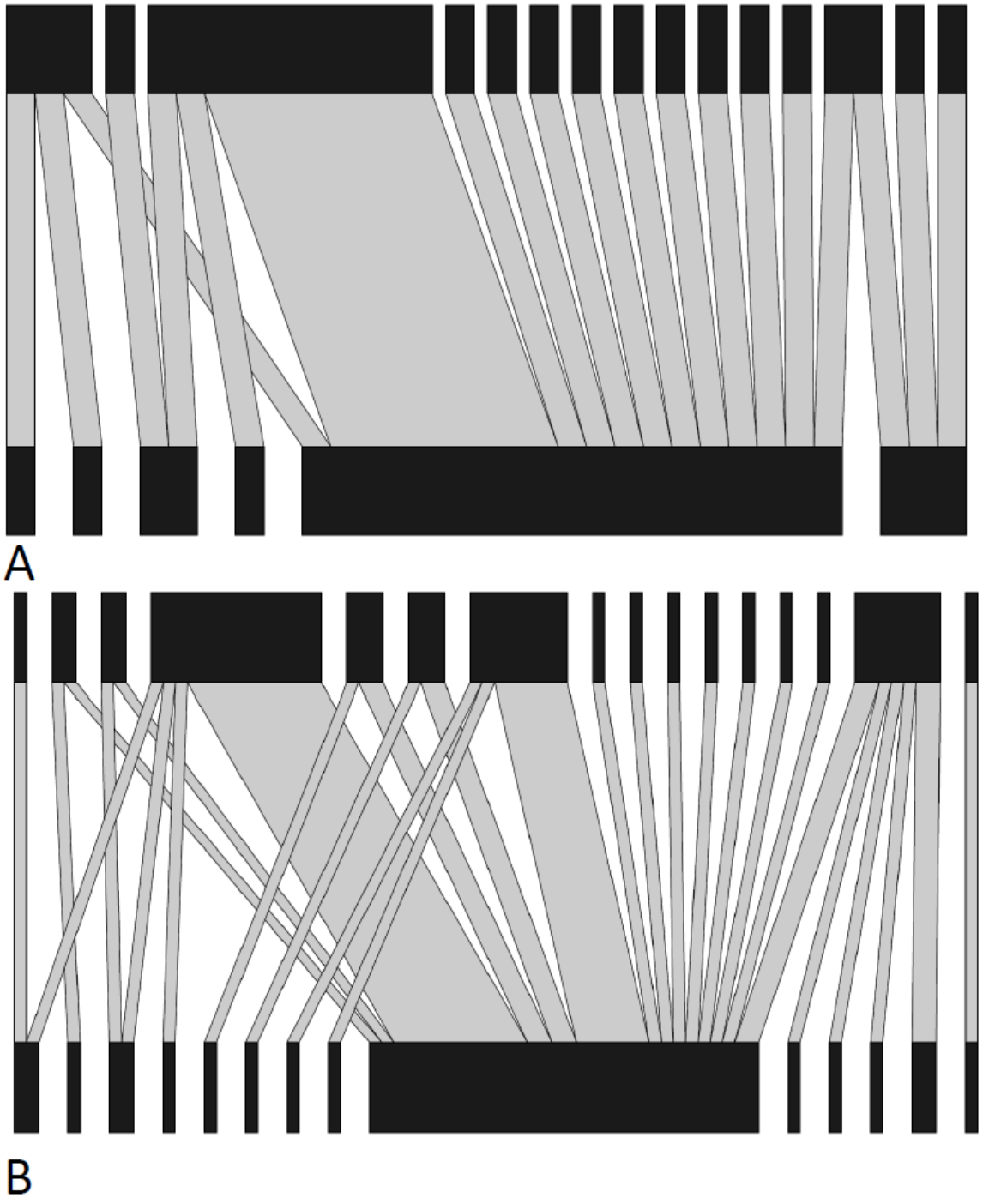


Figure 11. Biting Diptera-host network divided by site. Diptera (top)-host (bottom) interactions separated by field site: Within Abutia Amagame (A), Mafi Agove (B). Node and edge widths are proportional to frequency of occurrence.

The H2' specialisation of the combined network across both sites was 0.192, and network specialisation was higher in the Near-natural (H2'= 0.347) habitat than Agricultural (H2'= 0.148) habitat. There were insufficient data to reliably calculate the Village H2' value.

The closeness centrality and degree of humans was assessed for the overall network (closeness= 0.094; degree= 25) and each habitat category. Closeness centrality was higher in Village habitat (closeness = 0.5) than Agricultural (closeness= 0.170) and Near-natural habitat (closeness= 0.171). Further, the degree of humans was higher in Near-natural habitat (degree= 16) than in Agricultural (degree= 10) and Village habitats (degree=4).

Culex watti had the highest betweenness centrality of all the blood fed biting Diptera in our combined network (betweenness= 0.450). Moreover, *C. watti* had the highest betweenness centrality across both Near-natural (betweenness= 0.507) and Agricultural (betweenness = 1). Due to the lack of data for the Village habitat was excluded from this analysis.

MiSeq sequence data

Blood meal sequencing on the Illumina MiSeq platform using the 16S rRNA loci produced a higher number of blood meals than anticipated. In total we produced 3,324 unique ASVs and following read processing and filtering, we characterised the composition of 689 blood meals. Examination of these data found that 71% (492) were of human origin. Comparison with the dataset produced on the Sequel II platform indicated that 38 blood meals characterised as human were also positively identified as human by the MiSeq platform. After discounting the blood meals of human origin from the MiSeq data set, we identified 11 host species which broadly followed a similar trend of decreasing with anthropogenic landscape modification. However, due to the high likelihood of contamination, no further analysis was conducted on this dataset.

5.5 Discussion

Overall, we found no significant dissimilarity in community composition and no effect of habitat classification on H diversity. Biting Diptera and host species richness was higher in Near-natural and Agricultural habitats. Our selected network metrics varied with differing

levels of anthropogenic landscape modification. Moreover, we documented patterns of biting interactions that had the potential to lead to a wide variety of transmission events between a range of biting Diptera and competent hosts.

The impact of anthropogenic landscape modification on community composition

There were fewer hosts and biting Diptera with increasing landscape modification and a nearly threefold increase in biting Diptera abundance in Near-natural habitat compared to the other classifications. This could be because of the proximity and availability of breeding sites and hosts (Barrientos-Roldán et al., 2022; Ferraguti et al., 2016; Johnston et al., 2014; Pereira-Silva et al., 2021; Young et al., 2021). Indeed, despite highly modified land, such as Village habitat, providing breeding habitats outside of seasonal precipitation cycles, only those synanthropic species capable of propagating and feeding on hosts in these habitats are likely to be present (Barata et al., 2012; Johnston et al., 2014). Consequently, the greater Diptera species richness and abundance in Near-natural and Agricultural habitat are likely due to the habitat providing a greater number of hosts and more variable breeding sites that supports a greater number of Diptera species. Moreover, despite host sampling being incomplete, the observed pattern of increased host richness in Near-natural habitat may have implications for disease transmission potential in less anthropogenically modified habitat through the dilution effect (Shah et al., 2019).

Species evenness remained low across all habitat, though followed a pattern of being lowest in village contexts. Similar patterns in which mosquito species evenness was broadly lower in more anthropogenically modified habitat have previously been observed (Möhlmann et al., 2017; Petruff et al., 2020; Visa Shalini et al., 2022). Several highly abundant species were found across all habitat types, including *Culicoides imicola, Culicoides distinctipennis, Sergentomyia africana* and *Culex watti,* indicating they can dominate in a range of landscape contexts and likely display a degree of plasticity in blood feeding behaviour to utilise the range of host available across the landscapes (Stone & Gross, 2018; Villard & Metzger, 2014). Indeed, we recorded *Culex watti* as having blood meals from six host species. As these highly abundant species are competent vectors of a range of disease causative pathogens, their abundance and wide distribution has significant implications for the emergence of zoonotic diseases and disease transmission potential across all habitat classifications.

Different habitats are connected through overlapping species distributions

We found no significant dissimilarity in community composition across habitat classifications. However, we identified a range of biting Diptera species that were unique to each habitat. Indeed, 53% of Diptera species collected in the Near-natural classification were unique to this habitat, likely due to a preference for the Near-natural environmental characteristics (Vittor et al., 2009). Communities in transitional zones (ecotones) are often comprised of species from the adjacent habitats (Fortin et al., 2000). We observed a pattern of high Diptera species overlap across an anthropogenic landscape modification gradient from Agricultural to Near-natural (34% of Agricultural species were also present in Near-Natural habitat), indicating a shared community and movement between the two habitat types (Meyer Steiger et al., 2016). Overlapping biting Diptera communities link Agricultural with adjacent Near-natural habitat and, consequently, humans with the sylvatic cycle through the movement of opportunistic bridge vectors (Miot et al., 2020). We found several species that link the Village and Agricultural habitat, such as *Anopheles coustani,* a known vector of a range of diseases including malaria, and which displays anthropophilic behaviour (Ciubotariu et al., 2020). These findings are epidemiologically relevant as they indicate that existing and future Agricultural habitat may facilitate disease transmission across the landscape and provide a pathway through which emerging and established zoonotic disease are introduced into rural and urban transmission cycles.

The impact of anthropogenic landscape modification on community structure

Network connectance was greater within the Village habitat than Agricultural and Near-natural habitats. Although, these findings could be a result of the limited number of recorded interactions, interaction partners, and the abundance of interactions with humans (Valdovinos et al., 2009). Low network connectance in Agricultural and Near-natural habitat is likely a result of the increased host species richness (Rivera-Hutinel et al., 2012). Interpolation and extrapolation curves and demonstrated that biting Diptera-host interactions were highly under sampled, in comparison to richness estimates, in all habitat types. Consequently, connectance values are potentially not representative of the true Diptera-host interaction structure (Blüthgen, 2010; Heleno et al., 2012). However, whilst most networks are incomplete to a degree (Vizentin-Bugoni et al., 2016), limited sampling

effort, such as in this work, can be sufficient to capture the majority of functionally important species of the network (Hegland et al., 2010).

H2' network specialisation decreased with anthropogenic landscape modification, though each habitat classification displayed low values of network specialisation, likely reflecting plasticity in host choice. Host usage is broadly characterised by a high degree of plasticity in biting Diptera, which may be significantly influenced by local host densities (Takken & Verhulst, 2013). Increased H2' in Near-natural habitat could be reflective of the observed greater host and Diptera density, species richness, and potential niche partitioning (Chakravarty et al., 2023; Dalsgaard et al., 2011). Lower network specialism in Agricultural than Near-natural habitat is particularly epidemiologically relevant as generalist vectors have an increased capacity to transmit emerging and established zoonoses into novel hosts (Santiago-Alarcon et al., 2012). Therefore, habitat with a lower network specialism will likely express a greater disease transmission potential (Ellwanger & Chies, 2021). Indeed, agricultural drivers have previously been linked to over 50% of all zoonotic diseases emergence (Rohr et al., 2019). Each habitat was represented by a single network. Whilst each network was considered distinct and sampling conducted over an extended period provided a snapshot of community structure in each habitat classification, the contiguous nature of the landscape, and overlapping Diptera communities indicated a potential movement of individuals between habitats. Such movement is epidemiologically relevant and may act to link the spatially distinct networks and their associated disease transmission cycles, linking susceptible hosts in a network with competent vectors in another, therefore, facilitating transmission across the landscape.

The embeddedness of humans in biting Diptera-host networks

The high degree centrality of humans in Agricultural and Near-natural habitat demonstrates that they may act as hub species (Toju et al., 2018), which interact with, and potentially link, a wide range of biting Diptera species and hosts, including disease competent vectors. Indeed, the presence of mixed blood meals containing DNA of human and wild animal origin highlights the potential for transmission of established and emerging zoonotic disease via bridge vectors (Brackney et al., 2021; Miot et al., 2020).

The closeness centrality of humans was greatest in Village habitat and there was minimal difference between Agricultural and Near-natural habitat, suggesting that transmission potential to humans is higher in anthropogenically modified habitat (J. Grubb et al., 2021; Ribeiro et al., 2020). These findings are particularly epidemiologically relevant, as they demonstrate that humans are equally at risk of disease transmission working in Agricultural land when compared to encroachment into Near-natural habitat. The high centrality in Village habitats is expected due to the anthropophilic feeding behaviour of the sampled blood fed Diptera and their historical association with anthropogenic habitat (Bennett et al., 2015; Ughasi et al., 2012). However, the high centrality in the Village habitat may also be a product of the limited number of blood meals collected from this habitat (Costenbader & Valente, 2003). Indeed, given the observed presence of domestic hosts, such as goats, chickens, and dogs, there was a surprising lack of domestic animal host DNA in the Village blood meals. This could potentially be due to anthropophilic feeding behaviour of the biting Diptera, though it is more likely that this habitat was under-sampled due to concentration of dwellings only at the centre of the sampling areas (Figure 1, 2, & 4). Despite lower closeness centrality values in Agricultural and Near-natural habitat, high human degree centrality in these habitats remains an epidemiological concern (Bell et al., 1999).

Culex watti is consistently the most central Diptera species

Across the range of blood fed biting Diptera collected in all habitat classifications, *Culex watti*, a potential vector of West Nile Virus (Diarra et al., 2019), consistently had the highest betweenness centrality. A high betweenness value indicates that the species is important for the cohesion of the network and favours the circulation of their vectored pathogens throughout the network (Espinaze et al., 2018; Llaberia-Robledillo et al., 2022). Whilst a higher number of blood fed and non-blood fed *C. watti* were collected in Agricultural and Near-natural habitats, the presence of *C. watti* within all habitat classifications indicates a degree of generality in their habitat range (Stone & Gross, 2018). Moreover, the high betweenness values and presence across all levels of anthropogenic landscape modifications suggest that *C. watti* has the potential to act as a connector species, linking and facilitating disease transmission between spatially distinct communities (Bellekom et al., 2021; Llaberia-Robledillo et al., 2022). The high betweenness centrality and widespread distribution of *C. watti* in our sites is surprising given the limited available literature on the ecology of this

species. Future work should therefore examine if the patterns observed here are consistent with those across gradients of landscape modification in other regions of Africa.

Limitations

Ecological networks populated with interaction data generated through the DNA sequencing of biting Diptera blood meals provide an insight into the community and interaction structure of Diptera and their hosts in a sampled location. Whilst every effort was taken to reduce bias, there are several potential limitations to this approach. Interaction networks provide a snapshot of community and interaction structure, and multiple years of sampling are required to document temporal changes in network structure (Olivier et al., 2019). To minimise the effect of varying diel activity patterns, for each sampling event we sampled over a period of 24 hours. Under-sampling may alter conclusions drawn from several network metrics (Blüthgen, 2010); the rarefaction analysis demonstrated that interactions, and host and Diptera richness are indeed under sampled. Significantly increased sampling effort, beyond what was possible within time and travel limitations would be required to increase the confidence in the network metrics values (Ings et al., 2009).

The sequencing of host DNA was conducted on multiple sequencing platforms, however due to technical and time constraints a different gene region was used by each platform, limiting any comparison of host identification efficacy. As we lacked a priori host usage assumptions, we used general vertebrate primers targeting the 16S rRNA gene region when sequencing on the MiSeq platform (Bellekom et al., 2023). However, as high-throughput sequencing does not discriminate between endogenous and contaminating DNA, the hyper abundance of human sequences indicated that human contaminants had potentially been introduced during sample processing, limiting the reliability of the sequence data (Llamas et al., 2016). Consequently, we draw all conclusions in this work based on the PacBio Sequel II host identification data set.

Conclusions

Increasing anthropogenic landscape modification could drive a range of vector-borne disease outbreaks and zoonotic emergence events through its effect on biting Diptera and host community composition and vector-host relationships (Meyer Steiger et al., 2016; Patz et al., 2004). The characterisation of the biting Diptera-host community and interaction

structure provides opportunities to highlight and implement control initiatives to mitigate transmission potential. In the short term, such opportunities may include suggestions of behavioural modification for village inhabitants to limit exposure to transmission, including informing locals of potential risk factors, limiting encroachment into Near-natural habitat, using insect repellent when working at the interface between Agricultural and Near-natural habitat to minimise biting, and increase uptake of insecticide treated bed netting. Further, in the medium term, such interaction data may highlight vectors and hosts of particular importance to network cohesion and disease circulation. These data will inform localised control initiatives, for example, targeting synanthropic breeding habitat to suppress anthropophilic species, and the identification and removal of key domestic host species that may act as reservoirs of disease. In the long term, these interaction data may be used to inform wide scale control initiatives of the potential ecological consequences and network rewiring that may occur following population suppression, such as those that aim to eliminate malaria through reduction in Anopheles gambiae population using transgenic mosquitoes.

The high degree of overlap in community composition and shared host usage between Agricultural and Near-natural land places rural populations that rely on subsistence farming at the interface of disease transmission between rural and sylvatic habitat (Meyer Steiger et al., 2016). Closeness and degree centrality metrics indicated that there was a high disease transmission risk to humans. Incorporation of complementary quantitative susceptibility and vector competency rates may provide a more resolved assessment of vector-human transmission potential. Our interaction data, when combined with vector competency data, indicated a high number of potential transmission pathways. Sampling of biting Diptera and subsequent storage methods that maintain pathogen integrity, such as storage in dry ice, RNALater, or liquid nitrogen (Hoyos et al., 2021; Waldetensai et al., 2021) would provide opportunities to identify the origin of host DNA and interrogate pathogens present in Diptera blood meals. Subsequent interaction and pathogen data would thus allow for network approaches to be applied on an individual basis, in which nodes represent host or Diptera individuals and edges potential transmission events (Bellekom et al., 2021). Such data would be highly epidemiologically relevant and allow for the monitoring of the transmission of established and emerging zoonosis across the rural-sylvatic interface.

5.6 References

Abella-Medrano, C. A., Ibáñez-Bernal, S., Carbó-Ramírez, P., & Santiago-Alarcon, D. (2018). Blood-meal preferences and avian malaria detection in mosquitoes (Diptera: Culicidae) captured at different land use types within a neotropical montane cloud forest matrix. *Parasitology International*, *67*(3), 313–320. https://doi.org/10.1016/J.PARINT.2018.01.006

Acheampong, E. O., Macgregor, C. J., Sloan, S., & Sayer, J. (2019). Deforestation is driven by agricultural expansion in Ghana's forest reserves. *Scientific African*, *5*, e00146. https://doi.org/10.1016/J.SCIAF.2019.E00146

Agboli, E., Zahouli, J. B. Z., Badolo, A., & Jöst, H. (2021). Mosquito-Associated Viruses and Their Related Mosquitoes in West Africa. *Viruses 2021, Vol. 13, Page 891*, *13*(5), 891. https://doi.org/10.3390/V13050891

Agyemang-Badu, S. Y., Awuah, E., Oduro-Kwarteng, S., Dzamesi, J. Y. W., Dom, N. C., & Kanno, G. G. (2023). Environmental Management and Sanitation as a Malaria Vector Control Strategy: A Qualitative Cross-Sectional Study Among Stakeholders, Sunyani Municipality, Ghana. *Environmental Health Insights*, *17*. https://doi.org/10.1177/11786302221146890

Akuffo, R., Sanchez, C., Chicharro, C., Carrillo, E., Attram, N., Mosore, M. T., Yeboah, C., Kotey, N. K., Boakye, D., Ruiz-Postigo, J. A., Moreno, J., Wilson, M., Sarfo, B., & Anto, F. (2021). Detection of cutaneous leishmaniasis in three communities of Oti Region, Ghana. *PLOS Neglected Tropical Diseases*, *15*(5), e0009416. https://doi.org/10.1371/JOURNAL.PNTD.0009416

Awine, T., Malm, K., Bart-Plange, C., & Silal, S. P. (2017). Towards malaria control and elimination in Ghana: challenges and decision making tools to guide planning. *Global Health Action*, *10*(1), 1381471. https://doi.org/10.1080/16549716.2017.1381471

Barata, R. A., Ursine, R. L., Nunes, F. P., Morais, D. H., & Araújo, H. S. (2012). Synanthropy of mosquitoes and sand flies near the Aimorés hydroelectric power plant, Brazil. *Journal of Vector Ecology*, *37*(2), 397–401. https://doi.org/10.1111/J.1948-7134.2012.00243.X

Barrientos-Roldán, M. J., Abella-Medrano, C. A., Ibáñez-Bernal, S., & Antonio Sandoval-Ruiz, C. (2022). Landscape Anthropization Affects Mosquito Diversity in a Deciduous Forest in Southeastern Mexico. *Journal of Medical Entomology*, *59*(1), 248–256. https://doi.org/10.1093/JME/TJAB154

Bell, D. C., Atkinson, J. S., & Carlson, J. W. (1999). Centrality measures for disease transmission networks. *Social Networks*, *21*(1), 1–21. https://doi.org/10.1016/S0378-8733(98)00010-0

Bellekom, B., Bailey, A., England, M., Langlands, Z., Lewis, O. T., & Hackett, T. D. (2023). Effects of storage conditions and digestion time on DNA amplification of biting midge (Culicoides) blood meals. *Parasites & Vectors 2023 16:1*, *16*(1), 1–7. https://doi.org/10.1186/S13071-022-05607-X

Bellekom, B., Hackett, T. D., & Lewis, O. T. (2021). A Network Perspective on the Vectoring of Human Disease. *Trends in Parasitology*, *37*(5), 391–400. https://doi.org/10.1016/j.pt.2020.12.001

Benelli, G., & Beier, J. C. (2017). Current vector control challenges in the fight against malaria. *Acta Tropica*, *174*, 91–96. https://doi.org/10.1016/J.ACTATROPICA.2017.06.028

Bennett, K. L., Linton, Y. M., Shija, F., Kaddumukasa, M., Djouaka, R., Misinzo, G., Lutwama, J., Huang, Y. M., Mitchell, L. B., Richards, M., Tossou, E., & Walton, C. (2015). Molecular Differentiation of the African Yellow Fever Vector Aedes bromeliae (Diptera: Culicidae) from Its Sympatric Non-vector Sister Species, Aedes lilii. *PLOS Neglected Tropical Diseases*, *9*(12), e0004250. https://doi.org/10.1371/JOURNAL.PNTD.0004250

Blüthgen, N. (2010). Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic and Applied Ecology*, *11*(3), 185–195. https://doi.org/10.1016/J.BAAE.2010.01.001

Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecology*, *6*(1), 9. https://doi.org/10.1186/1472-6785-6-9

Bonds, M. H., Keenan, D. C., Rohani, P., & Sachs, J. D. (2010). Poverty trap formed by the ecology of infectious diseases. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1685), 1185–1192. https://doi.org/10.1098/RSPB.2009.1778

Borland, E. M., & Kading, R. C. (2021). Modernizing the Toolkit for Arthropod Bloodmeal Identification. *Insects 2021, Vol. 12, Page 37*, *12*(1), 37. https://doi.org/10.3390/INSECTS12010037

Braack, L., Gouveia de Almeida, A. P., Cornel, A. J., Swanepoel, R., & de Jager, C. (2018). Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasites & Vectors*, *11*(1), 29. https://doi.org/10.1186/s13071-017-2559-9

Brackney, D. E., LaReau, J. C., & Smith, R. C. (2021). Frequency matters: How successive feeding episodes by blood-feeding insect vectors influences disease transmission. *PLOS Pathogens*, *17*(6), e1009590. https://doi.org/10.1371/JOURNAL.PPAT.1009590

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581. https://doi.org/10.1038/NMETH.3869

Campbell-Lendrum, D., Manga, L., Bagayoko, M., & Sommerfeld, J. (2015). Climate change and vector-borne diseases: what are the implications for public health research and policy? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1665), 1–8. https://doi.org/10.1098/RSTB.2013.0552

Captain-Esoah, M., Kweku Baidoo, P., Frempong, K. K., Adabie-Gomez, D., Chabi, J., Obuobi, D., Kwame Amlalo, G., Balungnaa Veriegh, F., Donkor, M., Asoala, V., Behene, E., Adjei Boakye, D., & Dadzie, S. K. (2020). Biting Behavior and Molecular Identification of Aedes aegypti (Diptera: Culicidae) Subspecies in Some Selected Recent Yellow Fever Outbreak Communities in Northern Ghana. *Journal of Medical Entomology*, *57*(4), 1239–1245. https://doi.org/10.1093/jme/tjaa024

Chakravarty, R., Radchuk, V., Managave, S., & Voigt, C. C. (2023). Increasing species richness along elevational gradients is associated with niche packing in bat assemblages. *The Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.13897

Chao, A. (1987). Estimating the Population Size for Capture-Recapture Data with Unequal Catchability. *Biometrics*, *43*(4), 783. https://doi.org/10.2307/2531532

Chaves, L. F., Harrington, L. C., Keogh, C. L., Nguyen, A. M., & Kitron, U. D. (2010). Blood feeding patterns of mosquitoes: Random or structured? *Frontiers in Zoology*, *7*(1), 1–11. https://doi.org/10.1186/1742-9994-7-3

Chaves, L. S. M., Bergo, E. S., Conn, J. E., Laporta, G. Z., Prist, P. R., & Sallum, M. A. M. (2021). Anthropogenic landscape decreases mosquito biodiversity and drives malaria vector proliferation in the Amazon rainforest. *PLOS ONE*, *16*(1), e0245087. https://doi.org/10.1371/JOURNAL.PONE.0245087

Chaves, L. S. M., Fry, J., Malik, A., Geschke, A., Sallum, M. A. M., & Lenzen, M. (2020). Global consumption and international trade in deforestation-associated commodities could influence malaria risk. *Nature Communications 2020 11:1*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-14954-1

Ciubotariu, I. I., Jones, C. M., Kobayashi, T., Bobanga, T., Muleba, M., Pringle, J. C., Stevenson, J. C., Carpi, G., & Norris, D. E. (2020). Genetic Diversity of Anopheles coustani (Diptera: Culicidae) in Malaria Transmission Foci in Southern and Central Africa. *Journal of Medical Entomology*, *57*(6), 1782. https://doi.org/10.1093/JME/TJAA132

Civitello, D. J., Cohen, J., Fatima, H., Halstead, N. T., Liriano, J., McMahon, T. A., Ortega, C. N., Sauer, E. L., Sehgal, T., Young, S., & Rohr, J. R. (2015). Biodiversity inhibits parasites: Broad evidence for the dilution effect. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(28), 8667–8671. https://doi.org/10.1073/PNAS.1506279112/SUPPL_FILE/PNAS.1506279112.SD02.XLSX

Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks*, *25*, 283–307. https://doi.org/10.1016/S0378-8733(03)00012-1

Dalsgaard, B., Magård, E., Fjeldså, J., González, A. M. M., Rahbek, C., Olesen, J. M., Ollerton, J., Alarcón, R., Araujo, A. C., Cotton, P. A., Lara, C., Machado, C. G., Sazima, I., Sazima, M., Timmermann, A., Watts, S., Sandel, B., Sutherland, W. J., & Svenning, J. C. (2011). Specialization in Plant-Hummingbird Networks Is Associated with Species Richness, Contemporary Precipitation and Quaternary Climate-Change Velocity. *PLOS ONE*, *6*(10), e25891. https://doi.org/10.1371/JOURNAL.PONE.0025891

Demey, R., & Borrow, N. (2010). *Field Guide to the Birds of Ghana*. Christopher Helm Publishers, Incorporated.

DI Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology 2017 35:4*, *35*(4), 316–319. https://doi.org/10.1038/nbt.3820

Diarra, A. Z., Laroche, M., Berger, F., & Parola, P. (2019). Use of MALDI-TOF MS for the

Identification of Chad Mosquitoes and the Origin of Their Blood Meal. *The American Journal of Tropical Medicine and Hygiene*, *100*(1), 47. https://doi.org/10.4269/AJTMH.18-0657

Dixon, P. (2003). VEGAN, A Package of R Functions for Community Ecology. *Journal of Vegetation Science*, *14*(6), 927–930. https://www.jstor.org/stable/3236992

Ellwanger, J. H., & Chies, J. A. B. (2021). Zoonotic spillover: Understanding basic aspects for better prevention. *Genetics and Molecular Biology*, *44*(1 Suppl 1), 20200355. https://doi.org/10.1590/1678-4685-GMB-2020-0355

Espinaze, M. P. A., Hellard, E., Horak, I. G., & Cumming, G. S. (2018). Domestic mammals facilitate tick-borne pathogen transmission networks in South African wildlife. *Biological Conservation*, *221*, 228–236. https://doi.org/10.1016/J.BIOCON.2018.03.017

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., & Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology*, *30*(12), 1904–1916. https://doi.org/10.1111/1365-2435.12659

Evans, D. M., Pocock, M. J. O., & Memmott, J. (2013). The robustness of a network of ecological networks to habitat loss. *Ecology Letters*, *16*(7), 844–852. https://doi.org/10.1111/ELE.12117

Fath, B. D., Scharler, U. M., Ulanowicz, R. E., & Hannon, B. (2007). Ecological network analysis: network construction. *Ecological Modelling*, *208*(1), 49–55. https://doi.org/10.1016/J.ECOLMODEL.2007.04.029

Ferraguti, M., Martínez-De La Puente, J., Roiz, D., Ruiz, S., Soriguer, R., & Figuerola, J. (2016). Effects of landscape anthropization on mosquito community composition and abundance. *Scientific Reports 2016 6:1*, *6*(1), 1–9. https://doi.org/10.1038/srep29002

Food and Agriculture Organization of the United Nations. (2023). *Ghana at a glance | FAO in Ghana | Food and Agriculture Organization of the United Nations*. https://www.fao.org/ghana/fao-in-ghana/ghana-at-a-glance/en/

Fortin, M. J., Olson, R. J., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K., & Klemas, V. (2000). Issues related to the detection of boundaries. *Landscape Ecology*, *15*(5), 453–466. https://doi.org/10.1023/A:1008194205292/METRICS

Franklinos, L. H. V., Jones, K. E., Redding, D. W., & Abubakar, I. (2019). The effect of global change on mosquito-borne disease. In *The Lancet Infectious Diseases* (Vol. 19, Issue 9, pp. e302–e312). Lancet Publishing Group. https://doi.org/10.1016/S1473-3099(19)30161-6

GBIF.org, GBIF Home Page. (2021). Available from: Https://Www.Gbif.Org.

Gibb, R., Franklinos, L. H. V, Redding, D. W., & Jones, K. E. (2020). Ecosystem perspectives are needed to manage zoonotic risks in a changing climate. *BMJ*, *371*, m3389. https://doi.org/10.1136/bmj.m3389

Gibb, R., Redding, D. W., Chin, K. Q., Donnelly, C. A., Blackburn, T. M., Newbold, T., & Jones, K. E. (2020). Zoonotic host diversity increases in human-dominated ecosystems.

Nature, 1–5. https://doi.org/10.1038/s41586-020-2562-8

Giraldo-Calderón, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Emrich, S., Collins, F., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Madey, G., Collins, F. H., Lawson, D., Kersey, P., Allen, J., Christensen, M., Hughes, D., Koscielny, G., Langridge, N., … Wieck, R. (2015). VectorBase: An updated Bioinformatics Resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Research*, *43*(D1), D707–D713. https://doi.org/10.1093/nar/gku1117

Gratz, N. G. (1999). Emerging and resurging vector-borne diseases. *Annual Review of Entomology*, *44*(1), 51–75. https://doi.org/10.1146/annurev.ento.44.1.51

Grubb, J., Lopez, D., Mohan, B., & Matta, J. (2021). Network centrality for the identification of biomarkers in respondent-driven sampling datasets. *PLoS ONE*, *16*(8). https://doi.org/10.1371/JOURNAL.PONE.0256601

Grubb, P. (1998). *Mammals of Ghana Sierra Leone and the Gambia*. Trendrine.

Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., Janzen, D. H., Hallwachs, W., Naik, S., Sones, J. E., & Zakharov, E. V. (2018). A Sequel to Sanger: Amplicon sequencing that scales. *BMC Genomics*, *19*(1), 1–14. https://doi.org/10.1186/S12864-018-4611-3/FIGURES/10

Hebert, R., & Meglécz, E. (2022). NSDPY: A python package to download DNA sequences from NCBI. *SoftwareX*, *18*, 101038. https://doi.org/10.1016/J.SOFTX.2022.101038

Hegland, S. J., Dunne, J., Nielsen, A., & Memmott, J. (2010). How to monitor ecological communities cost-efficiently: The example of plant–pollinator networks. *Biological Conservation*, *143*(9), 2092–2101. https://doi.org/10.1016/J.BIOCON.2010.05.018

Heleno, R., Devoto, M., & Pocock, M. (2012). Connectance of species interaction networks and conservation value: Is it any good to be well connected? *Ecological Indicators*, *14*(1), 7–10. https://doi.org/10.1016/J.ECOLIND.2011.06.032

Hopken, M. W., Reyes-Torres, L. J., Scavo, N., Piaggio, A. J., Abdo, Z., Taylor, D., Pierce, J., & Yee, D. A. (2021). Temporal and Spatial Blood Feeding Patterns of Urban Mosquitoes in the San Juan Metropolitan Area, Puerto Rico. *Insects*, *12*(2), 1–21. https://doi.org/10.3390/INSECTS12020129

Hoyos, J., Carrasquilla, C., León, C., Montgomery, J. M., Salyer, S. J., Komar, N., & González, C. (2021). Host selection pattern and flavivirus screening of mosquitoes in a disturbed Colombian rainforest. *Scientific Reports |*, *11*(1), 18656. https://doi.org/10.1038/s41598-021-98076-8

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*(12), 1451–1456. https://doi.org/10.1111/2041-210X.12613

Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., Edwards, F., Figueroa, D., Jacob, U., Jones, J. I., Lauridsen, R. B., Ledger, M. E., Lewis, H. M., Olesen, J. M., van Veen, F. J. F., Warren, P. H., & Woodward, G. (2009). Review: Ecological networks - beyond food webs. *Journal of Animal Ecology*, *78*(1), 253–269. https://doi.org/10.1111/j.1365-2656.2008.01460.x

Jerry, D. C. T., Mohammed, T., & Mohammed, A. (2017). Yeast-generated CO2: A convenient source of carbon dioxide for mosquito trapping using the BG-Sentinel® traps. *Asian Pacific Journal of Tropical Biomedicine*, *7*(10), 896–900. https://doi.org/10.1016/J.APJTB.2017.09.014

Johnston, E., Weinstein, P., Slaney, D., Flies, A. S., Fricker, S., & Williams, C. (2014). Mosquito communities with trap height and urban-rural gradient in Adelaide, South Australia: implications for disease vector surveillance. *Journal of Vector Ecology*, *39*(1), 48–55. https://doi.org/10.1111/J.1948-7134.2014.12069.X

Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., McKeever, D., Mutua, F., Young, J., McDermott, J., & Pfeiffer, D. U. (2013). Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8399–8404. https://doi.org/10.1073/PNAS.1208059110/-/DCSUPPLEMENTAL

Kesavaraju, B., Damal, K., & Juliano, S. A. (2008). Do natural container habitats impede invader dominance? Predator-mediated coexistence of invasive and native container-dwelling mosquitoes. *Oecologia*, *155*(3), 631–639. https://doi.org/10.1007/S00442-007-0935-4/FIGURES/4

Kirkeby, C., Græsbøll, K., Stockmarr, A., Christiansen, L. E., & Bødker, R. (2013). The range of attraction for light traps catching Culicoides biting midges (Diptera: Ceratopogonidae). *Parasites & Vectors*, *6*(1), 67. https://doi.org/10.1186/1756-3305-6-67

Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., Zhou, Y., Yao, L., Yan, G., & Chen, X. G. (2014). Urbanization Increases Aedes albopictus Larval Habitats and Accelerates Mosquito Development and Survivorship. *PLOS Neglected Tropical Diseases*, *8*(11), e3301. https://doi.org/10.1371/JOURNAL.PNTD.0003301

Llaberia-Robledillo, M., Balbuena, J. A., Sarabeev, V., & Llopis-Belenguer, C. (2022). Changes in native and introduced host–parasite networks. *Biological Invasions*, *24*(2), 543–555. https://doi.org/10.1007/S10530-021-02657-7/FIGURES/2

Llamas, B., Valverde, G., Fehren-Schmitz, L., Weyrich, L. S., Cooper, A., & Haak, W. (2016). From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *Https://Doi.Org/10.1080/20548923.2016.1258824*, *3*(1), 1–14. https://doi.org/10.1080/20548923.2016.1258824

Mattah, P. A. D., Futagbi, G., Amekudzi, L. K., Mattah, M. M., de Souza, D. K., Kartey-Attipoe, W. D., Bimi, L., & Wilson, M. D. (2017). Diversity in breeding sites and distribution of Anopheles mosquitoes in selected urban areas of southern Ghana. *Parasites & Vectors*, *10*(1), 25. https://doi.org/10.1186/s13071-016-1941-3

Mayi, M. P. A., Bamou, R., Djiappi-Tchamen, B., Fontaine, A., Jeffries, C. L., Walker, T., Antonio-Nkondjio, C., Cornel, A. J., & Tchuinkam, T. (2020). Habitat and Seasonality Affect Mosquito Community Composition in the West Region of Cameroon. *Insects*, *11*(5), 312. https://doi.org/10.3390/insects11050312

Mayi, M. P. A., Foncha, D. F., Kowo, C., Tchuinkam, T., Brisco, K., Anong, D. N., Ravinder, S., & Cornel, A. J. (2019). Impact of deforestation on the abundance, diversity, and

richness of Culex mosquitoes in a southwest Cameroon tropical rainforest. *Journal of Vector Ecology*, *44*(2), 271–281. https://doi.org/10.1111/jvec.12359

Meyer Steiger, D. B., Ritchie, S. A., & Laurance, S. G. W. (2016). Mosquito communities and disease risk influenced by land use change and seasonality in the Australian tropics. *Parasites and Vectors*, *9*(1), 1–13. https://doi.org/10.1186/S13071-016-1675-2/FIGURES/8

Miller, E., & Huppert, A. (2013). The Effects of Host Diversity on Vector-Borne Disease: The Conditions under Which Diversity Will Amplify or Dilute the Disease Risk. *PLoS ONE*, *8*(11), e80279. https://doi.org/10.1371/journal.pone.0080279

Miot, E. F., Calvez, E., Aubry, F., Dabo, S., Grandadam, M., Marcombe, S., Oke, C., Logan, J. G., Brey, P. T., & Lambrechts, L. (2020). Risk of arbovirus emergence via bridge vectors: case study of the sylvatic mosquito Aedes malayensis in the Nakai district, Laos. *Scientific Reports 2020 10:1*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-64696-9

Möhlmann, T. W. R., Wennergren, U., Tälle, M., Favia, G., Damiani, C., Bracchetti, L., & Koenraadt, C. J. M. (2017). Community analysis of the abundance and diversity of mosquito species (Diptera: Culicidae) in three European countries at different latitudes. *Parasites and Vectors*, *10*(1), 1–12. https://doi.org/10.1186/S13071-017-2481-1/FIGURES/4

Müller, R., Reuss, F., Kendrovski, V., & Montag, D. (2019). Vector-Borne Diseases. In *Biodiversity and Health in the Face of Climate Change* (pp. 67–90). Springer International Publishing. https://doi.org/10.1007/978-3-030-02318-8_4

Natalia Ivanova, B., deWaard, J., & Hebert, P. (2006). An inexpensive, automation-friendly proto-col for recovering high-quality DNA. *Molecular Ecology Notes*, *6*, 998–1002. www.dnabarcoding.ca

Olivier, P., Frelat, R., Bonsdorff, E., Kortsch, S., Kröncke, I., Möllmann, C., Neumann, H., Sell, A. F., & Nordström, M. C. (2019). Exploring the temporal variability of a food web using long-term biomonitoring data. *Ecography*, *42*(12), 2107–2121. https://doi.org/10.1111/ecog.04461

Patz, J. A., Daszak, P., Tabor, G. M., Aguirre, A. A., Pearl, M., Epstein, J., Wolfe, N. D., Kilpatrick, A. M., Foufopoulos, J., Molyneux, D., Bradley, D. J., Amerasinghe, F. P., Ashford, R. W., Barthelemy, D., Bos, R., Bradley, D. J., Buck, A., Butler, C., Chivian, E. S., … Zakarov, V. (2004). Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence. *Environmental Health Perspectives*, *112*(10), 1092. https://doi.org/10.1289/EHP.6877

Pereira-Silva, J. W., Ríos-Velásquez, C. M., Lima, G. R. de, Marialva dos Santos, E. F., Belchior, H. C. M., Luz, S. L. B., Naveca, F. G., & Pessoa, F. A. C. (2021). Distribution and diversity of mosquitoes and Oropouche-like virus infection rates in an Amazonian rural settlement. *PLOS ONE*, *16*(2), e0246932. https://doi.org/10.1371/JOURNAL.PONE.0246932

Perrin, A., Glaizot, O., & Christe, P. (2022). Worldwide impacts of landscape anthropization on mosquito abundance and diversity: A meta-analysis. *Global Change Biology*, *28*(23), 6857–6871. https://doi.org/10.1111/GCB.16406

Petruff, T. A., McMillan, J. R., Shepard, J. J., Andreadis, T. G., & Armstrong, P. M. (2020). Increased mosquito abundance and species richness in Connecticut, United States 2001–2019. *Scientific Reports 2020 10:1*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-76231-x

Poisot, T., & Gravel, D. (2014). When is an ecological network complex? Connectance drives degree distribution and emerging network properties. *PeerJ*, *2014*(1), e251. https://doi.org/10.7717/PEERJ.251/FIG-4

Proulx, S. R., Promislow, D. E. L., & Phillips, P. C. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, *20*(6), 345–353. https://doi.org/10.1016/J.TREE.2005.04.004

Ribeiro, S. P., Castro e Silva, A., Dáttilo, W., Reis, A. B., Góes-Neto, A., Alcantara, L. C., Giovanetti, M., Coura-Vital, W., Fernandes, G. W., & Azevedo, V. A. C. (2020). Severe airport sanitarian control could slow down the spreading of COVID-19 pandemics in Brazil. *PeerJ*, *2020*(6), e9446. https://doi.org/10.7717/PEERJ.9446/SUPP-3

Rivera-Hutinel, A., Bustamante, R. O., Marín, V. H., & Medel, R. (2012). Effects of sampling completeness on the structure of plant-pollinator networks. *Ecology*, *93*(7), 1593–1603. https://doi.org/10.1890/11-1803.1

Rohr, J. R., Barrett, C. B., Civitello, D. J., Craft, M. E., Delius, B., DeLeo, G. A., Hudson, P. J., Jouanard, N., Nguyen, K. H., Ostfeld, R. S., Remais, J. V., Riveau, G., Sokolow, S. H., & Tilman, D. (2019). Emerging human infectious diseases and the links to global food production. *Nature Sustainability 2019 2:6*, *2*(6), 445–456. https://doi.org/10.1038/s41893-019-0293-3

Santiago-Alarcon, D., Havelka, P., Schaefer, H. M., & Segelbacher, G. (2012). Bloodmeal Analysis Reveals Avian Plasmodium Infections and Broad Host Preferences of Culicoides (Diptera: Ceratopogonidae) Vectors. *PLOS ONE*, *7*(2), e31098. https://doi.org/10.1371/JOURNAL.PONE.0031098

Shah, H. A., Huxley, P., Elmes, J., & Murray, K. A. (2019). Agricultural land-uses consistently exacerbate infectious disease risks in Southeast Asia. *Nature Communications 2019 10:1*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-12333-z

Silk, M. J., Croft, D. P., Delahay, R. J., Hodgson, D. J., Boots, M., Weber, N., & McDONALD, R. A. (2017). Using Social Network Measures in Wildlife Disease Ecology, Epidemiology, and Management. *BioScience*, *67*(3), 245–257. https://doi.org/10.1093/BIOSCI/BIW175

Smallegange, R. C., Schmied, W. H., van Roey, K. J., Verhulst, N. O., Spitzen, J., Mukabana, W. R., & Takken, W. (2010). Sugar-fermenting yeast as an organic source of carbon dioxide to attract the malaria mosquito Anopheles gambiae. *Malaria Journal*, *9*, 292. https://doi.org/10.1186/1475-2875-9-292

Stone, C., & Gross, K. (2018). Evolution of host preference in anthropophilic mosquitoes. *Malaria Journal*, *17*(1), 257. https://doi.org/10.1186/S12936-018-2407-1

Su, M., Jiang, Z., & Hui, C. (2022). How Multiple Interaction Types Affect Disease Spread and Dilution in Ecological Networks. *Frontiers in Ecology and Evolution*, *10*, 466. https://doi.org/10.3389/FEVO.2022.862986/BIBTEX

Takken, W., & Verhulst, N. O. (2013). Host Preferences of Blood-Feeding Mosquitoes. *Annual Review of Entomology*, *58*(1), 433–453. https://doi.org/10.1146/annurev-ento-120811-153618

Toju, H., Tanabe, A. S., & Sato, H. (2018). Network hubs in root-associated fungal metacommunities. *Microbiome 2018 6:1*, *6*(1), 1–16. https://doi.org/10.1186/S40168-018-0497-1

Ughasi, J., Bekard, H. E., Coulibaly, M., Adabie-Gomez, D., Gyapong, J., Appawu, M., Wilson, M. D., & Boakye, D. A. (2012). Mansonia africana and Mansonia uniformis are Vectors in the transmission of Wuchereria bancrofti lymphatic filariasis in Ghana. *Parasites and Vectors*, *5*(1), 1–5. https://doi.org/10.1186/1756-3305-5-89/TABLES/3

Valdovinos, F. S., Ramos-Jiliberto, R., Flores, J. D., Espinoza, C., & López, G. (2009). Structure and dynamics of pollination networks: the role of alien plants. *Oikos*, *118*(8), 1190–1200. https://doi.org/10.1111/J.1600-0706.2009.17364.X

van Hoesel, W., Marzal, A., Magallanes, S., Santiago-Alarcon, D., Ibáñez-Bernal, S., & Renner, S. C. (2019). Management of ecosystems alters vector dynamics and haemosporidian infections. *Scientific Reports 2019 9:1*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-45068-4

Villard, M. A., & Metzger, J. P. (2014). REVIEW: Beyond the fragmentation debate: a conceptual model to predict when habitat configuration really matters. *Journal of Applied Ecology*, *51*(2), 309–318. https://doi.org/10.1111/1365-2664.12190

Visa Shalini, P., Shriram, A. N., Elango, A., Natarajan, R., Vijayakumar, B., Raju, K. H. K., Dengel, L., Gunasekaran, K., & Kumar, A. (2022). Mosquito Diversity in an Experimental Township in Tamil Nadu, India. *Journal of Medical Entomology*, *59*(5), 1615–1624. https://doi.org/10.1093/JME/TJAC064

Vittor, A. Y., Pan, W., Gilman, R. H., Tielsch, J., Glass, G., Shields, T., Sánchez-Lozano, W., Pinedo, V. V., Salas-Cobos, E., Flores, S., & Patz, J. A. (2009). Linking Deforestation to Malaria in the Amazon: Characterization of the Breeding Habitat of the Principal Malaria Vector, Anopheles darlingi. *The American Journal of Tropical Medicine and Hygiene*, *81*(1), 5. /pmc/articles/PMC3757555/

Vizentin-Bugoni, J., Maruyama, P. K., Debastiani, V. J., Duarte, L. da S., Dalsgaard, B., & Sazima, M. (2016). Influences of sampling effort on detected patterns and structuring processes of a Neotropical plant–hummingbird network. *Journal of Animal Ecology*, *85*(1), 262–272. https://doi.org/10.1111/1365-2656.12459

W.H.O. (2022). *World malaria report 2022*. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022

Waldetensai, A., Gemechu, F., Kinfe, E., Amare, H., Hagos, S., Teshome, D., Hailemariam, A., Nigatu, W., Tadesse, A., Keneni, D., Dinku, Z., Telaye, M., Esmael, N., Kweka, E. J., Moges, B., & Massebo, F. (2021). Aedes mosquito responses to control interventions against the Chikungunya outbreak of Dire Dawa, Eastern Ethiopia. *International Journal of Tropical Insect Science*, 1–10. https://doi.org/10.1007/s42690-021-00430-w

Wheeler, D., & Bhagwat, M. (2007). BLAST QuickStart. *Methods in Molecular Biology*, *395*,

149–176. https://doi.org/10.1385/1-59745-514-8:149

Wilson, R., Wakefield, A., Roberts, N., & Jones, G. (2021). Artificial light and biting flies: the parallel development of attractive light traps and unattractive domestic lights. *Parasites and Vectors*, *14*(1), 1–11. https://doi.org/10.1186/S13071-020-04530-3/FIGURES/3

Young, K. I., Buenemann, M., Vasilakis, N., Perera, D., & Hanley, K. A. (2021). Shifts in mosquito diversity and abundance along a gradient from oil palm plantations to conterminous forests in Borneo. *Ecosphere*, *12*(4), e03463. https://doi.org/10.1002/ECS2.3463

Author contributions: BB: Data analysis and writing, OTL: Supervision, review, and editing, TDH: Supervision, review, and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

5.7 Acknowledgements

All sampling was conducted in collaboration with the University of Ghana and members of the Target Malaria research consortium, who were conducting field sampling at the sites and provided logistical support.

5.8 Supplementary

Supplementary 1. Collection site classifications

Lot.ID	Site	Habitat	Distance
955	Abutia Amegame	Agricultural	Far
867	Abutia Amegame	Village	Within Village
922	Mafi Agorve	Village	Within Village
779	Abutia Amegame	Near-Natural	Village edge
702	Abutia Amegame	Agricultural	Far
933	Mafi Agorve	Agricultural	Far
999	Mafi Agorve	Near-Natural	Far
1358	Abutia Amegame	Near-Natural	Village edge
911	Mafi Agorve	Village	Within Village
1193	Mafi Agorve	Near-Natural	Village edge
1021	Mafi Agorve	Near-Natural	Village edge
1010	Mafi Agorve	Agricultural	Far
713	Abutia Amegame	Near-Natural	Far
900	Abutia Amegame	Near-Natural	Far
834	Mafi Agorve	Village	Within Village
724	Abutia Amegame	Village	Within Village
1237	Mafi Agorve	Near-Natural	Far
1314	Abutia Amegame	Village	Village edge
1043	Abutia Amegame	Near-Natural	Far
1226	Mafi Agorve	Agricultural	Far
966	Abutia Amegame	Near-Natural	Far
1109	Mafi Agorve	Near-Natural	Far
988	Abutia Amegame	Agricultural	Far
889	Abutia Amegame	Near-Natural	Far
1303	Abutia Amegame	Agricultural	Far
1182	Mafi Agorve	Agricultural	Far
1032	Mafi Agorve	Agricultural	Far
878	Abutia Amegame	Agricultural	Far
1204	Mafi Agorve	Agricultural	Far
1248	Mafi Agorve	Agricultural	Far

Supplementary 2. The proportion of fed individuals for each biting Diptera species sampled across all sites.

Species	Total	Bloodfed	Proportion blood fed
Culex nebulosus	4	1	25.00%
Culex watti	175	24	13.71%
Aedes bromeliae	2	1	50.00%
Culex theileri	2	0	0.00%
Aedeomyia africana	2	0	0.00%
Culicoides imicola	1649	2	0.12%
Culicoides hortensis	130	3	2.31%
Sergentomyia ingrami	17	1	5.88%
Sergentomyia hamoni	18	1	5.56%
Sergentomyia africana	227	1	0.44%
Culex rima	9	3	33.33%
Mimomyia mimomyiaformis	12	7	58.33%
Coquillettidia metallica	20	2	10.00%
Culicoides distinctipennis	2766	2	0.07%
Mimomyia mediolineata	4	0	0.00%
Culex perexiguus	8	2	25.00%
Sergentomyia inermis	337	1	0.30%
Culex aurantapex	2	0	0.00%
Mansonia africana	25	8	32.00%
Culex decens	2	1	50.00%
Forcipomyia sp. 1	347	0	0.00%
Culicoides oxystoma	130	0	0.00%
Sergentomyia schwetzi	40	1	2.50%
Sergentomyia sp. 1	41	1	2.44%
Dasyhelea ludingensis	555	0	0.00%
Sergentomyia clydei	15	0	0.00%
Culex univittatus	7	1	14.29%
Psychoda sp. 1	140	0	0.00%
Culex trifoliatus	7	1	14.29%
Culex cinereus	1	1	100.00%
Aedes unilineatus	3	0	0.00%
Uranotaenia alboabdominalis	3	0	0.00%
Aedes hirsutus	4	0	0.00%
Culex duttoni	2	0	0.00%
Culex neavei	3	0	0.00%
Anopheles coustani	4	0	0.00%
Culex vansomereni	4	1	25.00%
Culex adersianus	2	0	0.00%
Aedes quasiunivittatus	2	0	0.00%
Culicoides neavei	3	0	0.00%
Uranotaenia bilineata	2	0	0.00%
Culex annulioris	1	0	0.00%
Culicoides enderleini	257	0	0.00%
Ceratopogonidae sp.1	4	1	25.00%
Sergentomyia distincta	12	3	25.00%
Ceratopogonidae sp.2	45	1	2.22%
Ceratopogonidae sp.3	42	1	2.38%
Uranotaenia micromelas	3	1	33.33%
Atrichopogon sp.1	4	1	25.00%
Allohelea sp.1	1	1	100.00%

Supplementary 3. Observed and estimated Diptera-host interactions and species numbers by habitat

Site interactions and species	Observed	Estimated	S.E.	95% Lower	95% Upper
Agricultural interactions	19	142.429	138.444	40.047	742.825
Agricultural hosts	9	32.625	30.012	12.460	170.331
Agricultural Diptera	11	19.679	9.826	12.463	62.475
Near-natural interactions	30	207.577	143.583	74.224	743.047
Near-natural hosts	11	22.936	12.829	13.143	77.490
Near-natural Diptera	22	84.359	54.579	36.222	295.420

Chapter 6: General Discussion

Leaving Abutia Amagame

6.1 Synopsis

In this thesis I provide insight into the community and interaction structure of biting Diptera and their hosts. I provide a rationale for the use of ecological interaction networks, demonstrate the interconnectedness of biting Diptera and their hosts on a global scale, and suggest fruitful applications for network analysis (Chapter 2). Further, through compiling interaction data extracted from the literature, I examined the latitudinal and anthropogenic effect on biting Diptera-host network structure (Chapter 3). Interaction data derived from blood meals requires the presence of sufficient quantities of non-degraded host DNA to allow for further molecular analysis and host identification. I examined the effect of digestion time and storage conditions on the PCR amplification success of biting midge blood meals (Chapter 4). Finally, using data collected from field sites in Ghana, I implemented a network approach to examine the effect of anthropogenic landscape modification on the community of biting Diptera and their hosts (Chapter 5), and address outstanding questions highlighted in Chapter 2.

6.2 Key findings

Biting Diptera research has, until now, primarily focused on the examination of pairwise interactions of a single Diptera species or small sets of interacting species (Bellekom et al., 2021) and has broadly overlooked their effect on the complex communities in which they are embedded (Ferguson et al., 2010), though recent work has started to move beyond this (Macias Torres & Naranjo Mayorga, 2022; Yee et al., 2022; Young et al., 2020). In Chapter 2, I highlight the applicability of a network-based approach to examine biting Diptera-host community and interaction structure and reveal important network properties that relate to disease transmission potential. Further, I highlight the embeddedness of humans in networks of biting Diptera, and their risk of exposure to a wide range of established and emerging vectored disease. My findings and suggested application of networks in the examination of biting-Diptera community structure in Chapter 2 laid the foundations for this thesis.

6.2.1 Variability in hosts and Diptera species richness between levels of anthropogenic landscape modification

Through use of a biting Diptera-host interaction dataset derived from a subset of data I compiled in Chapter 2 and refined with a strict exclusion criterion to minimise the sources of bias, I examined the effect of anthropogenic and latitudinal factors on network structure. In addition, I conducted an analysis of samples collected at two field sites in Ghana to examine empirically the effect of anthropogenic landscape modification on biting Diptera community composition and interaction structure. In agreement with previous findings in other systems (Chaves et al., 2021; Ferraguti et al., 2016; Li et al., 2014; Mayi et al., 2019; Pernat et al., 2021; Perrin et al., 2022; Steiger et al., 2016), biting Diptera and host species richness and diversity were found to be consistently highest in a near-natural habitat across both Chapters. However, in contrast to my findings in Chapter 5, host and Diptera species richness was lowest in the agricultural habitat classification. Moreover, relative host and Diptera species richness was far higher in the urban modification category in Chapter 3 than that identified in my field-based sampling (Chapter 5).

The consistently high Diptera species richness within near-natural habitats that I highlighted in both Chapter 3 and 5 are likely a result of the high host diversity, local environmental conditions, and favourable breeding environments for a wide range of Diptera (Evans et al., 2022). Indeed, natural phytotelma habitats have previously been found to contain a higher Diptera diversity than artificial containers, such as those found in urban habitats, which likely only provide suitable breeding environments for synanthropic species (Srisuka et al., 2022). My findings are particularly epidemiologically relevant as there was a high degree of potential movement of village inhabitants into near-natural environments, and a habitat which maintains a high host richness has a greater depth of hosts from which a novel zoonotic pathogen may emerge (Allen et al., 2017; Keesing et al., 2010). The lower agricultural host and Diptera richness in my compiled interaction database (Chapter 3) compared to the empirical field data (Chapter 5) could be due to the greater variability within the agricultural classification in Chapter 3, which ranged from cattle farms to cropland. Therefore, the high biomass of domestic animals in certain contexts, such as farms, may have limited the overall observed host species richness (Mayi et al., 2020; Thiemann et al., 2011). My findings indicated that there was a degree of overlap in host and

Diptera between agricultural and near-natural habitats at our Ghanian field sites, suggesting that agricultural land in this landscape represented a transitional zone between highly modified and unmodified habitat. Consequently, variation in agricultural species richness is likely to be driven by the specific landscape context (Möhlmann et al., 2017).

The greater species richness in urban habitats that I discussed in Chapter 3 could arise because the compiled dataset contained human dwellings of more variable size (e.g., cities, towns, and villages), contrasting only the isolated villages sampled in Chapter 5. Indeed, host species richness is higher in cities and towns than previously anticipated, and likely aids in the maintenance of regional biodiversity (Aronson et al., 2014; Spotswood et al., 2021). Moreover, the greater richness of biting Diptera is potentially linked to the high host species richness (Mayi et al., 2020), and greater presence of urban green spaces (Zhao et al., 2020). Whilst extrapolation and interpolation curves revealed that host and Diptera species richness were well resolved in the interaction dataset used for Chapter 3, they were highly unresolved in Chapter 5. Therefore, the true host and Diptera species richness of this category is likely greater than reported. The variation in host and Diptera richness in urban habitats I identified in Chapter 3 and five has implications for vectored disease transmission potential. Whilst a high host richness may potentially decrease zoonotic disease transmission through the dilution effect (Civitello et al., 2015), increased population has been positively linked to increased levels of arboviral disease transmission (Kolimenakis et al., 2021). Consequently, my findings suggest that vectored disease transmission potential may vary across different levels of urbanisation (Hay et al., 2005).

6.2.2 Network analysis reveals the effect of macro-ecological and anthropogenic factors on the specialisation of biting Diptera-host networks

In Chapter 2, I discuss the potential uses and implications of a network-based approach for the examination of interactions between biting Diptera and their hosts. In this thesis, I demonstrate the value of ecological network analysis in the examination of biting Diptera-host interactions derived from the metabarcoding of Diptera blood meals, at a macro-ecological and local scale. For a wide range of taxa, species richness increases towards the equator with an expected resultant change in the strength and dynamic consequences of their interspecific interactions (Morris et al., 2014). In Chapter 3, I investigated the effect of

latitude on the specialism of Diptera-host networks; I found little indication of the expected pattern of dietary specialisation (Dyer et al., 2007). These findings were in agreement with those documented for host-parasitoid networks (Morris et al., 2014), suggesting that the lack of variation in network specialism could be as a result of underlying rules for how antagonistic interactions are structured, regardless of the taxonomic composition and size of the network.

The effect of anthropogenic landscape modification on network specialism was quantified using a complied Diptera-host interaction dataset from published data (Chapter 3) and empirically examined using field collected samples (Chapters 5). There was no significant effect of landscape modification on H2' network specialisation using the Diptera-host interaction dataset, though a patten of decreasing H2' with increasing anthropogenic landscape modification was observed in our Ghanian field sites. H2' values of networks across both Chapters were consistently low, indicating a high degree of generality (Blüthgen et al., 2006). Landscapes that support high levels of generality are particularly epidemiologically relevant, as generalist vectors have an increased capacity to act as bridge vectors and facilitate sustained transmission from reservoir hosts to humans (Bellekom et al., 2021; Santiago-Alarcon et al., 2012). The variation in H2' values that I observed in Chapter 5 indicate that agricultural habitat in this context could have a greater risk of disease transmission than near-natural habitat. As much of the population of the sampled villages rely on subsistence farming, they are at particular risk of infection (van de Straat et al., 2022). Therefore, this result could help to inform future disease mitigations strategies in these locations.

6.2.3 Humans are highly embedded in biting Diptera-host networks across all levels of landscape modification

Characterisation of the degree to which humans are embedded within biting Diptera-host networks has clear epidemiological benefits (Bellekom et al., 2021). Across Chapters 2, 3, and 5 I found humans to be highly central to network structure. Indeed, in Chapter 2, using a compilation of over 26,000 interactions, I demonstrated that interactions between biting Diptera and humans accounted for nearly 24% of all total interactions and involved over 57% of Diptera species. A host that has a high degree of interactions with a variety of competent

vectors over multiple habitats could act as a hub species, linking Diptera and hosts, and facilitating disease transmission across a landscape (Poulin, 2010; Toju et al., 2018). However, the extent to which a host may act as a transmission hub for a pathogen will be governed by the relative susceptibility of a host to the pathogen and the vectoral competence of the Diptera they interact with, resulting in varying transmission potential between pathogens across the landscape. I found that interactions with humans were abundant across all habitat classifications and interacted with many biting Diptera species, potentially exposing them to a wide range of vectored disease (Chapter 3). Whilst host usage is suggested to be in part modulated by host density (Takken & Verhulst, 2013), surprisingly, the Village/Urban networks were not dominated by interactions with humans, despite their likely high abundance. I suggest that a greater than expected generalism of biting Diptera (Hassell et al., 2017), and an abundance of other taxa, such as birds and domesticated animals are potential causes for these findings.

The Ghanaian field sites sampled for Chapter 5 were characterised by a high degree of human movement across habitat classifications, in particular the presence of humans at the agricultural/near-natural interface and encroachment into near-natural habitat (Acheampong et al., 2019). The high degree of human embeddedness in agricultural and near-natural habitat is likely representative of this movement and is a particular epidemiological concern. Indeed, the highest number of biting Diptera-human interactions occurred in near-natural habitat, exposing humans to the sylvatic cycle and potential transmission of established and emerging zoonoses through opportunistic blood feeding by bridge vectors (Miot et al., 2020). Moreover, increasing human movement between farms and the encroachment of humans into near-natural habitat could also offer the potential for spill back events, in which pathogens confined to a human transmission cycle may form novel sylvatic cycles in areas in which they were previously absent (Hanley et al., 2013). Historic spill back events include the establishment of a Yellow Fever virus sylvatic transmission cycle in Brazil (De Abreu et al., 2020). An urban Yellow Fever cycle was introduced to the Americas as early as the 17th century, and following the adaptation of neotropical vectors, such as *Haemagogus spp,* and the pathogen to neotropical non-human primates, a sylvatic cycle then emerged (Klitting et al., 2018). Such spill back events have

prevented the eradication of the Yellow Fever virus from the continent despite widespread vector control efforts (De Abreu et al., 2020).

Consequently, the field sites sampled for Chapter 5, and those in a similar landscape context, could be good candidates for future vector surveillance initiatives (Carrasquilla et al., 2023).

6.2.4 The prevalence of mixed blood meals and their importance

Advances in biting Diptera blood meal host identification have enabled easier and more cost-effective methods to detect and identify mixed blood meals (those containing DNA from more than one host origin) (Bellekom et al., 2021, 2023). However, the occurrence of mixed blood meals is highly variable, both in my results and those published. I found that only 15% of included papers in the network compilation reported mixed blood meals, with an average of approximately 13% mixed blood meals when reported (Chapter 2). Moreover, only 7% of the total blood meals analysed in Chapter 5 contained DNA from more than one host; 66% of these contained human blood as one of the hosts. The proportion of mixed blood meals can be biased by sampling strategy. For example, previous work has identified relatively high proportions (22-37%) of mixed blood meals, however, sampling was conducted using human landing catches near cattle. Consequently, the presence of mixed human and cattle blood meals in this context is unsurprising (Tirados et al., 2006). Therefore, I suggest that the presence and quantity of mixed blood meals could be heavily dependent on sampling location and collection method.

The characterisation of mixed blood meal composition is epidemiologically relevant as competent vectors that feed on a diverse range of successive hosts are candidates to act as bridge vectors (Burkett-Cadena et al., 2021; Hendy et al., 2020; Kang et al., 2019). Further, the detection of mixed blood meals has the potential to inform the success of control strategies. Increased frequency of human-animal mixed blood meals could represent a vectors response to control initiatives, such as increased bed net coverage and the implementation of livestock-targeted interventions, with vectors altering their feeding behaviour from biting humans to including other readily available hosts, such as cattle (Ndenga et al., 2016; Tandina et al., 2020). However, the application of this approach may be dependent on the vector of interest. As indicated in Chapter 5, there is a high degree of heterogeneity in the proportion of captured blood fed individuals between biting Diptera

species. Moreover, blood meal source has previously been shown to influence a vector's microbiome composition and diversity, and the presence of a mixed blood meal is likely to a have a different impact on the microbiota than a single origin blood meal (Landesman et al., 2019; Muturi et al., 2019, 2021). For example, the presence of a mixed blood meal has a synergistic effect on an *Aedes aegypti*'s microbiota, increasing bacterial richness and diversity (Muturi et al., 2021). Resultant changes in composition and abundance of a vector's gut microbiota could alter their transmission potential, as a range of bacteria have been positively and negatively linked to a vector's susceptibility to transmissible pathogens (Muturi et al., 2021). Whilst mixed blood meals are indicators of a vector interacting with multiple host species, hosts may also be linked through consecutive blood meals across an adult vector's lifetime. Indeed, consecutive blood meals may also act to increase transmission potential through reduction of the extrinsic incubation period and increasing pathogen load (Pathak et al., 2023; Serafim et al., 2018). Consequently, the examination of gut microbiota, the abundance of mixed blood meals in different landscape contexts, and the change in a vector's life history traits as a result of a mixed blood meal, could be a promising avenue for future research.

6.2.5 The transmission potential of vector-borne disease

Vectored pathogen infections are established through the feeding of a biting Diptera on a vertebrate host (Huang et al., 2019). There has been an increased focus on novel emerging zoonoses and a range of studies have examined the role of a single species involved in a pathogens transmission (Graham et al., 2009). However, there has been less consideration of a pathogen's transmission potential in the context of biting Diptera-host interaction networks. To establish informed vectored disease control measures and predict transmission risk, data on vector distributions, habitat associations and interaction patterns are required (Kading et al., 2018). Whilst no data on pathogen presence in blood meals were collected due to logistical constraints, we can make informed assumptions about a vector's competence (Giraldo-Calderón et al., 2015). In Chapter 2, I highlight the wide range of pathogens that biting Diptera may vector and demonstrates the scale to which humans are embedded in pathogen-vector-host networks. However, as these data were interactions

from a global dataset, geographically separated hosts are unlikely to co-occur to the extent that a single Diptera may interact with them, though increased human movement, encroachment, and accidental introduction of species will likely increase future interaction opportunities (Bellekom et al., 2021). In Chapter 5 I empirically examine potential transmission events between sampled biting Diptera and their hosts. I found that of the 29 blood fed biting Diptera species collected, the majority were competent of vectoring a pathogen of medical or veterinary importance. Concerningly, from an epidemiological and healthcare perspective, nearly 90% of competent blood fed species had interacted with at least one susceptible host, predominantly human.

Through construction of interaction networks, I highlight the consistently high betweenness centrality values for *Culex watti* across the different levels of anthropogenic landscape modification and the species' relative importance to network structure (Chapter 5). A node with a high betweenness centrality value across all levels of modification is likely to act as a connector node, linking the species across the landscape and potentially facilitating movement of disease from the sylvatic environment into urban habitats (Espinaze et al., 2018; Hoyos et al., 2021). Moreover, I identified a large overlap in biting Diptera community composition between agricultural and near-natural, indicating that hosts at the interface between agricultural and near-natural habitat are likely being exposed to disease cycles typically confined to near-natural habitat (Johnson et al., 2020). These findings further demonstrate the importance of ecological interaction networks in the examination of biting Diptera-host interaction structure and disease transmission potential, as discussed in Chapter 2, and may be used to inform control initiatives targeting vector species with high relative importance to network structure.

6.2.6 The impact of DNA digestion and storage conditions on host identification success

Reliable host identification is paramount to the construction of accurate biting Diptera-host networks and the examination of vector-borne pathogen transmission dynamics (Martínez-de la Puente et al., 2013). Preservation of host DNA is of utmost importance for successful downstream molecular analysis. Examination of the effect of storage conditions and digestion time on host DNA integrity has previously been conducted using several mosquito

species (Martínez-de la Puente et al., 2013; Reeves et al., 2016). To complement these findings and inform and validate my empirical field sampling protocol (Chapter 5), I examined the effect of storage and digestion time on blood meal DNA integrity in midges (Chapter 4). I found that there was a significant negative effect of digestion time on PCR amplification success and storage condition had no effect on amplification success.

My findings in Chapter 4 are particularly relevant to the design and implementation of biting Diptera sampling for the purpose of host identification using a molecular approach. As digestion significantly decreases the likelihood of amplification success and the availability of blood fed individuals is often limited, the digestive process should be halted immediately on collection to maximise identification potential (Bellekom et al., 2023; Martínez-de la Puente et al., 2013). Previous work has often sampled overnight, only retrieving collected samples in the morning, which has allowed digestion and degradation of host DNA to continue over this period (Kang et al., 2019; Pimentel et al., 2022; Sang et al., 2017). Based on this and my findings in Chapter 4, I used a CDC light trap, modified to include an ethanol filled collection device to immediately halt digestive processes, to sample the biting Diptera community (Chapter 5). This sampling method was broadly effective; however, due to the high local temperatures, a large initial ethanol volume, replenishing when required, was necessary. Nevertheless, I recommended that future work aiming to sample biting Diptera blood meals adopt such trap modifications to increase downstream amplification success.

In Chapter 4, I demonstrated the limited difference between storage in ethanol and a -20°C freezer on the amplification success of biting Diptera blood meals, indicating that samples can safely be maintained at ambient temperatures for extended periods without negatively impacting amplification success if stored in 95% ethanol. Whilst I stored samples at ambient temperatures in Chapter 4, due to logistical issues and delays during shipping, samples collected during empirical field work (Chapter 5) potentially experienced a range of temperatures for an extended period outside of cold storage. These samples maintained sufficient blood meal DNA integrity to identify host species. These results are particularly encouraging and are highly relevant for the planning and logistics of field work in remote locations with limited access to cold storage.

6.3 Limitations

The construction of biting Diptera-host interaction networks using data extracted from the literature revealed the effect of macro-ecological and local factors on network structure, however, this approach is not without its limitations. Nodes may affect each other with differing levels of reciprocation; however, as we consider all interactions as inherently equal, the relative dynamic importance of an interaction may not be captured by my dataset. This issue is common in ecological network analysis, though independent measures of node abundance and measurement of functional traits may help resolve this issue (Sonne et al., 2020; Xing & Fayle, 2021).

Network metrics are sensitive to variation in species abundance and overall network size, resulting in the potential for observed variations in network metrics across networks being due to differences in sampling intensity rather than true ecological patterns (Dormann et al., 2009). Consequently, careful selection of network metrics that are robust to variation in network size is required. Indeed, to account for the potential biases present in my heterogeneous dataset in Chapter 3, I calculated metrics that are relatively robust to network size and included network size and species richness as explanatory variables in my analysis (Blüthgen et al., 2006).

A particular challenge with the compilation of interaction data from the literature is the frequent presence of binomial misnomers, either because of incorrect identification or out of date taxonomy. For example, *Ochlerotatus* was previously considered a subgenus of *Aedes*, but was upgraded to genus level in 2000 (Reinert, 2000). Therefore, careful curation of the dataset was required prior to analysis to ensure a single species was not represented by multiple binomials, inflating node abundance.

Sampling completeness is a major limitation in network ecology, as rare nodes will be frequently underrepresented or go undetected and may distort our understanding of network structure and dynamics (Rivera-Hutinel et al., 2012). However, promising tools have been developed using predictive structural and coverage-deficit models (which utilises information on the completeness of sampling for individual nodes) to identify and augment observed networks with potential missing interactions (Terry & Lewis, 2020). Large networks

in a species rich environment are inevitably more difficult to sample and rare interactions require a greater sampling effort.

In Chapter 5, I used CDC light traps complemented by sugar fermented yeast CO_2 bait because of their portability, low cost, and their ability to collect a range of Diptera species. However, as with the majority of trap types, there is inevitably a degree of bias introduced through trapping method. Certain Diptera species may be more attracted to a trap than others, because of the limited range of attraction, the variability in attraction to ambient light, and responses to different wavelengths of light (McDermott & Mullens, 2018). Moreover, the sampling for Chapter 5 was conducted during an extended period of out of season dryness, which may have limited the diversity and abundance of Diptera species due to the limited availability of breeding sites (Romero-Vega et al., 2023). Consequently, whilst my results were generated in a uniform manner, the findings presented here should be considered a potentially incomplete representation of the biting Diptera-host community structure.

The use of molecular analysis and metabarcoding in ecological network analysis is a highly promising advance; however, there are inherent issues with the approach that may introduce bias into the analyses. Laboratory and bioinformatic choices will likely have affected the data input into my network analysis. The selection of primer may introduce bias through intrinsic differences in amplification efficiency of template, potentially due to mismatches between the primer and their corresponding binding site, resulting in the overamplification and representation of certain species (Acinas et al., 2005). Consequently, PCR primer bias may alter network structure through the exclusion of compartments of Diptera hosts (Cuff et al., 2022). Further, gene loci may affect the overall taxonomic resolution due to limited availability of reference sequences (Clare et al., 2019). To limit introduced bias, after examining existing databases to ensure appropriate numbers of references sequences for species likely to be found in our sampling region, I designed and validated a set of general vertebrate PCR primers targeting the 16S gene region. Bioinformatic approaches may also profoundly impact the data output, and I implemented appropriate measures to remove the influence of sequencing errors and false positive artefacts from the data (Cuff et al., 2022). Indeed, due to the high potential of contamination in my MiSeq data, these data were excluded from any further analysis.

6.4 Future research

The findings of this thesis and the approaches used offer promising opportunities for informing control strategies and implementation in future research (Supplementary 1.). The biting Diptera-host interaction dataset I compiled in Chapter 2 provides the opportunity to examine a range of ecological phenomena in detail (as demonstrated in Chapter 3) and, through a collaboration with the Verena consortium, has already been incorporated into research examining and predicting tripartite networks of mosquito-borne disease (Appendix 1). As research examining biting Diptera-host interactions increases, it may be prudent to continue to update this dataset to increase resolution and facilitate a greater range of network analyses.

The decreasing sequencing costs and the development of new sequencing and bioinformatic techniques such as shotgun sequencing and metagenomics will allow future research to generate ever more comprehensive biting Diptera-host interaction datasets (Evans et al., 2016), and increase their statistical power. As control strategies, policy, and health care measures are guided by vector surveillance, increased network completeness and resolution is to be of great importance. The identification of pathogens in biting Diptera blood meals offers a unique opportunity to examine transmission events using ecological networks, in which each node is represented by an individual host or Diptera, and may be highly informative to current xenosurveillance and biosurveillance efforts (Grubaugh et al., 2015). However, careful consideration of the logistics of sample storage, such as the use of dry ice and liquid nitrogen, is required in order to maintain pathogen integrity, which may prohibit sampling in locations with limited access to the storage methods and sites. The miniaturisation and development of field laboratories and sequencing devices is perhaps the most universally beneficial advancement in sequencing technology and will likely significantly increase the ease of vector and pathogen surveillance (Borland & Kading, 2021). The pocket-sized MinION next generation sequencer has previously been used in pathogen discovery, such as monitoring the spread of Zika virus (Faria et al., 2016), and has also been effective in identification of wildlife samples. Consequently, future research should examine pocket sequencers' potential to facilitate the construction of biting Diptera-host transmission networks through the simultaneous field identification of the Diptera, blood

meal origin, and pathogen presence, whilst eliminating complications regarding sample storage (Mongan et al., 2019).

Understanding the movement and distribution of biting Diptera that are the target of control initiatives is fundamental to their success. Whilst anthropophilic species are typically associated with human habitation, human activities, such as agricultural development, extend out of urban habitat, potentially facilitating movement of key species across the landscape (Epopa et al., 2020). Moreover, anthropophilic Diptera species have been shown to display a degree of plasticity in their host feeding behaviour, indicating that populations could be maintained outside of human habitation (Orsborne et al., 2019; Tchouassi et al., 2012; Thomas et al., 2009). Consequently, the examination of biting Diptera community and interaction structure outside of commonly sampled urban areas may reveal the presence of unexpected Diptera species. Indeed, synanthropic *Anopheles* species have been found in areas between two villages in Burkina Faso, which may be indicative of migration between the villages and thus the potential for acting as connector species facilitating transmission between the two sites (Epopa et al., 2020). However, in their work the distance from human habitation was limited by the proximity of the two sampled sites, as the sites were only 6km apart. Future research should examine the effect of distance to human habitation on biting Diptera community composition and interaction structure over a much greater spatial scale.

Spatial variation in Diptera-borne pathogens transmission risk is strongly linked to the distribution of the competent vector species. Both biotic and abiotic interactions have a prominent role in influencing species distributions. In Chapters 3 and 5 I suggested that variation in Diptera species richness was as a result of host diversity, and local environmental conditions. The incorporation of data, such as species distribution, and trait data, will provide the opportunity for future work to integrate a range of complementary analyses with network data to quantify the extent to which these factors influence species distributions, Diptera-host interactions, and subsequent disease transmission potential. For example, joint species distribution models have demonstrated the effect of abiotic and inter-species interactions on the spatial distributions of mosquito vectors in the UK (Golding et al., 2015).

Further, interactions between biting Diptera and their hosts may be moderated through matching of their respective morphological traits (Benadi et al., 2022). Combining trait data,

such as body size, mouthpart and wing morphology, with ecological network analysis can quantify the effect of trait matching on interaction frequencies, the relative contribution to interaction network structure, and subsequent disease transmission potential (Peña et al., 2023). Moreover, such data may be incorporated into complex vector-borne transmission models to increase the reliability of their predictions (Cator et al., 2020).

6.5 Conclusion

This thesis provides new insights into the benefits of using molecular metabarcoding combined with ecological network analysis of biting Diptera-host interactions. I revealed the global pattern of insect-host community structure and demonstrated the extent to which humans share potential vectors with other vertebrates. I demonstrated the use of ecological networks to highlight the effect of macro-ecological and anthropogenic effects on biting Diptera-host interaction structure. Moreover, I highlighted the negative effect of digestion time and lack of an effect of storage conditions on the amplification success of *Culicoides* blood meals, with significant implications for future research. Finally, I characterised the biting Diptera community composition and variation in interaction structure at two field sites in Ghana, demonstrating the effect of anthropogenic modification and highlighting potential transmission events. By collecting data in such a way that facilitates network analysis, future work may reveal novel insights into the effect of landscape changes, perturbations, and climate change on biting Diptera community and interaction structure and inform control measures through the prediction of potential zoonotic spillover and identification of bridge vectors, and disease transmission events.

6.6 References

Acheampong, E. O., Macgregor, C. J., Sloan, S., & Sayer, J. (2019). Deforestation is driven by agricultural expansion in Ghana's forest reserves. *Scientific African*, *5*, e00146. https://doi.org/10.1016/J.SCIAF.2019.E00146

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Applied and Environmental Microbiology*, *71*(12), 8966. https://doi.org/10.1128/AEM.71.12.8966-8969.2005

Allen, T., Murray, K. A., Zambrana-Torrelio, C., Morse, S. S., Rondinini, C., Di Marco, M., Breit, N., Olival, K. J., & Daszak, P. (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications 2017 8:1*, *8*(1), 1–10. https://doi.org/10.1038/s41467-017-00923-8

Aronson, M. F. J., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., Warren, P. S., Williams, N. S. G., Cilliers, S., Clarkson, B., Dobbs, C., Dolan, R., Hedblom, M., Klotz, S., Kooijmans, J. L., Kühn, I., Macgregor-Fors, I., Mcdonnell, M., Mörtberg, U., … Winter, M. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1780). https://doi.org/10.1098/RSPB.2013.3330

Bellekom, B., Bailey, A., England, M., Langlands, Z., Lewis, O. T., & Hackett, T. D. (2023). Effects of storage conditions and digestion time on DNA amplification of biting midge (Culicoides) blood meals. *Parasites & Vectors 2023 16:1*, *16*(1), 1–7. https://doi.org/10.1186/S13071-022-05607-X

Bellekom, B., Hackett, T. D., & Lewis, O. T. (2021). A Network Perspective on the Vectoring of Human Disease. *Trends in Parasitology*, *37*(5), 391–400. https://doi.org/10.1016/j.pt.2020.12.001

Benadi, G., Dormann, C. F., Fründ, J., Stephan, R., & Vázquez, D. P. (2022). Quantitative Prediction of Interactions in Bipartite Networks Based on Traits, Abundances, and Phylogeny*. *Https://Doi.Org/10.1086/714420*, *199*(6), 841–854. https://doi.org/10.1086/714420

Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecology*, *6*(1), 9. https://doi.org/10.1186/1472-6785-6-9

Borland, E. M., & Kading, R. C. (2021). Modernizing the Toolkit for Arthropod Bloodmeal Identification. *Insects 2021, Vol. 12, Page 37*, *12*(1), 37. https://doi.org/10.3390/INSECTS12010037

Burkett-Cadena, N. D., Day, J. F., & Unnasch, T. R. (2021). Ecology of Eastern Equine Encephalitis Virus in the Southeastern United States: Incriminating Vector and Host Species Responsible for Virus Amplification, Persistence, and Dispersal. *Journal of Medical Entomology*. https://doi.org/10.1093/JME/TJAB076

Carrasquilla, M. C., Ortiz, M. I., Amórtegui-Hernández, D., García-Restrepo, S., León, C., Méndez-Cardona, S., & González, C. (2023). Pathogens, reservoirs, and vectors involved in the transmission of vector-borne and zoonotic diseases in a Colombian region. *Brazilian Journal of Microbiology*, 1–12. https://doi.org/10.1007/S42770-023-00903-9/TABLES/4

Cator, L. J., Johnson, L. R., Mordecai, E. A., El Moustaid, F., Smallwood, T. R. C., LaDeau, S. L., Johansson, M. A., Hudson, P. J., Boots, M., Thomas, M. B., Power, A. G., & Pawar, S. (2020). The Role of Vector Trait Variation in Vector-Borne Disease Dynamics. *Frontiers in Ecology and Evolution*, *8*, 507337. https://doi.org/10.3389/FEVO.2020.00189/BIBTEX

Chaves, L. S. M., Bergo, E. S., Conn, J. E., Laporta, G. Z., Prist, P. R., & Sallum, M. A. M. (2021). Anthropogenic landscape decreases mosquito biodiversity and drives malaria vector proliferation in the Amazon rainforest. *PLOS ONE*, *16*(1), e0245087. https://doi.org/10.1371/JOURNAL.PONE.0245087

Civitello, D. J., Cohen, J., Fatima, H., Halstead, N. T., Liriano, J., McMahon, T. A., Ortega, C. N., Sauer, E. L., Sehgal, T., Young, S., & Rohr, J. R. (2015). Biodiversity inhibits parasites: Broad evidence for the dilution effect. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(28), 8667–8671. https://doi.org/10.1073/PNAS.1506279112/SUPPL_FILE/PNAS.1506279112.SD02.XLSX

Clare, E. L., Fazekas, A. J., Ivanova, N. V., Floyd, R. M., Hebert, P. D. N., Adams, A. M., Nagel, J., Girton, R., Newmaster, S. G., & Fenton, M. B. (2019). Approaches to integrating

genetic data into ecological networks. *Molecular Ecology*, *28*(2), 503–519. https://doi.org/10.1111/MEC.14941

Cuff, J. P., Windsor, F. M., Tercel, M. P. T. G., Kitson, J. J. N., & Evans, D. M. (2022). Overcoming the pitfalls of merging dietary metabarcoding into ecological networks. *Methods in Ecology and Evolution*, *13*(3), 545–559. https://doi.org/10.1111/2041-210X.13796

De Abreu, F. V. S., Ferreira-De-Brito, A., De Souza Azevedo, A., Linhares, J. H. R., De Oliveira Santos, V., Miranda, E. H., Neves, M. S. A. S., Yousfi, L., Ribeiro, I. P., Dos Santos, A. A. C., Dos Santos, E., Dos Santos, T. P., Teixeira, D. S., Gomes, M. Q., Fernandes, C. B., Da Silva, A. M. V., Da Rocha Queiroz Lima, M., Paupy, C., Romano, A. P. M., … Lourenço-De-Oliveira, R. (2020). Survey on Non-Human Primates and Mosquitoes Does not Provide Evidences of Spillover/Spillback between the Urban and Sylvatic Cycles of Yellow Fever and Zika Viruses Following Severe Outbreaks in Southeast Brazil. *Viruses 2020, Vol. 12, Page 364*, *12*(4), 364. https://doi.org/10.3390/V12040364

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal*, *2*(1), 7–24. https://doi.org/10.2174/1874213000902010007

Dyer, L. A., Singer, M. S., Lill, J. T., Stireman, J. O., Gentry, G. L., Marquis, R. J., Ricklefs, R. E., Greeney, H. F., Wagner, D. L., Morais, H. C., Diniz, I. R., Kursar, T. A., & Coley, P. D. (2007). Host specificity of Lepidoptera in tropical and temperate forests. *Nature*, *448*(7154), 696–699. https://doi.org/10.1038/nature05884

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., & Diabaté, A. (2020). Anopheles gambiae (s.l.) is found where few are looking: assessing mosquito diversity and density outside inhabited areas using diverse sampling methods. *Parasites and Vectors*, *13*(1), 1–11. https://doi.org/10.1186/s13071-020-04403-9

Espinaze, M. P. A., Hellard, E., Horak, I. G., & Cumming, G. S. (2018). Domestic mammals facilitate tick-borne pathogen transmission networks in South African wildlife. *Biological Conservation*, *221*, 228–236. https://doi.org/10.1016/J.BIOCON.2018.03.017

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., & Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology*, *30*(12), 1904–1916. https://doi.org/10.1111/1365-2435.12659

Evans, M., Bhatnagar, S., Drake, J., Murdock, C., & Mukherjee, S. (2022). Socio-ecological dynamics in urban systems: An integrative approach to mosquito-borne disease in Bengaluru, India. *People and Nature*, *4*(3), 730–743. https://doi.org/10.1002/PAN3.10311

Faria, N. R., Sabino, E. C., Nunes, M. R. T., Alcantara, L. C. J., Loman, N. J., & Pybus, O. G. (2016). Mobile real-time surveillance of Zika virus in Brazil. *Genome Medicine*, *8*(1). https://doi.org/10.1186/S13073-016-0356-2

Ferguson, H. M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M. S., Gimnig, J. E., Fish, D., & Killeen, G. F. (2010). Ecology: A Prerequisite for Malaria Elimination and Eradication. *PLoS Medicine*, *7*(8), e1000303. https://doi.org/10.1371/journal.pmed.1000303

Ferraguti, M., Martínez-De La Puente, J., Roiz, D., Ruiz, S., Soriguer, R., & Figuerola, J. (2016). Effects of landscape anthropization on mosquito community composition and abundance. *Scientific Reports 2016 6:1*, *6*(1), 1–9. https://doi.org/10.1038/srep29002

Giraldo-Calderón, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Emrich, S., Collins, F., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Madey, G., Collins, F. H., Lawson, D., Kersey, P., Allen, J., Christensen, M., Hughes, D., Koscielny, G., Langridge, N., … Wieck, R. (2015). VectorBase: An updated Bioinformatics Resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Research*, *43*(D1), D707–D713. https://doi.org/10.1093/nar/gku1117

Golding, N., Nunn, M. A., & Purse, B. V. (2015). Identifying biotic interactions which drive the spatial distribution of a mosquito community. *Parasites and Vectors*, *8*(1), 1–10. https://doi.org/10.1186/S13071-015-0915-1/FIGURES/5

Graham, S. P., Hassan, H. K., Burkett-Cadena, N. D., Guyer, C., & Unnasch, T. R. (2009). Nestedness of ectoparasite-vertebrate host networks. *PLoS ONE*, *4*(11), 7873.

https://doi.org/10.1371/journal.pone.0007873

Grubaugh, N. D., Sharma, S., Krajacich, B. J., Fakoli III, L. S., Bolay, F. K., Diclaro II, J. W., Johnson, W. E., Ebel, G. D., Foy, B. D., & Brackney, D. E. (2015). Xenosurveillance: A Novel Mosquito-Based Approach for Examining the Human-Pathogen Landscape. *PLOS Neglected Tropical Diseases*, *9*(3), e0003628. https://doi.org/10.1371/journal.pntd.0003628

Hanley, K. A., Monath, T. P., Weaver, S. C., Rossi, S. L., Richman, R. L., & Vasilakis, N. (2013). Fever versus Fever: the role of host and vector susceptibility and interspecific competition in shaping the current and future distributions of the sylvatic cycles of dengue virus and yellow fever virus. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *0*, 292. https://doi.org/10.1016/J.MEEGID.2013.03.008

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and Disease Emergence: Dynamics at the Wildlife–Livestock–Human Interface. In *Trends in Ecology and Evolution* (Vol. 32, Issue 1, pp. 55–67). Elsevier Ltd. https://doi.org/10.1016/j.tree.2016.09.012

Hay, S. I., Guerra, C. A., Tatem, A. J., Atkinson, P. M., & Snow, R. W. (2005). Urbanization, malaria transmission and disease burden in Africa. *Nature Reviews. Microbiology*, *3*(1), 81. https://doi.org/10.1038/NRMICRO1069

Hendy, A., Hernandez-Acosta, E., Valério, D., Mendonça, C., Costa, E. R., Júnior, J. T. A., Assunção, F. P., Scarpassa, V. M., Gordo, M., Fé, N. F., Buenemann, M., de Lacerda, M. V. G., Hanley, K. A., & Vasilakis, N. (2020). The vertical stratification of potential bridge vectors of mosquito-borne viruses in a central Amazonian forest bordering Manaus, Brazil. *Scientific Reports*, *10*(1), 18254. https://doi.org/10.1038/s41598-020-75178-3

Hoyos, J., Carrasquilla, C., León, C., Montgomery, J. M., Salyer, S. J., Komar, N., & González, C. (2021). Host selection pattern and flavivirus screening of mosquitoes in a disturbed Colombian rainforest. *Scientific Reports* |, *11*(1), 18656. https://doi.org/10.1038/s41598-021-98076-8

Huang, Y.-J. S., Higgs, S., & Vanlandingham, D. L. (2019). Arbovirus-Mosquito Vector-Host

Interactions and the Impact on Transmission and Disease Pathogenesis of Arboviruses. *Frontiers in Microbiology*, *10*, 22. https://doi.org/10.3389/fmicb.2019.00022

Johnson, C. K., Hitchens, P. L., Pandit, P. S., Rushmore, J., Evans, T. S., Young, C. C. W., & Doyle, M. M. (2020). Global shifts in mammalian population trends reveal key predictors of virus spillover risk. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1924), 20192736. https://doi.org/10.1098/rspb.2019.2736

Kading, R. C., Golnar, A. J., Hamer, S. A., & Hamer, G. L. (2018). Advanced surveillance and preparedness to meet a new era of invasive vectors and emerging vector-borne diseases. *PLOS Neglected Tropical Diseases*, *12*(10), e0006761. https://doi.org/10.1371/journal.pntd.0006761

Kang, D. S., Martinez, R., Hosein, A., Shui Feng, R., James, L., Lovin, D. D., Cunningham, J. M., Miller, C. S. T., Eng, M. W., Chadee, D. D., & Severson, D. W. (2019). Identification of Host Blood Meals of Mosquitoes (Diptera: Culicidae) Collected at the Aripo Savannas Scientific Reserve in Trinidad, West Indies. *Journal of Medical Entomology*. https://doi.org/10.1093/jme/tjz113

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., & Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature 2010 468:7324*, *468*(7324), 647–652. https://doi.org/10.1038/nature09575

Klitting, R., Gould, E. A., Paupy, C., & de Lamballerie, X. (2018). What Does the Future Hold for Yellow Fever Virus? (I). *Genes*, *9*(6). https://doi.org/10.3390/GENES9060291

Kolimenakis, A., Heinz, S., Wilson, M. L., Winkler, V., Yakob, L., Michaelakis, A., Papachristos, D., Richardson, C., & Horstick, O. (2021). The role of urbanisation in the spread of Aedes mosquitoes and the diseases they transmit—A systematic review. *PLoS Neglected Tropical Diseases*, *15*(9). https://doi.org/10.1371/JOURNAL.PNTD.0009631

Landesman, W. J., Mulder, K., Allan, B. F., Bashor, L. A., Keesing, F., LoGiudice, K., & Ostfeld, R. S. (2019). Potential effects of blood meal host on bacterial community composition in Ixodes scapularis nymphs. *Ticks and Tick-Borne Diseases*, *10*(3), 523–527. https://doi.org/10.1016/J.TTBDIS.2019.01.002

Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., Zhou, Y., Yao, L., Yan, G., & Chen, X. G. (2014). Urbanization Increases Aedes albopictus Larval Habitats and Accelerates Mosquito Development and Survivorship. *PLOS Neglected Tropical Diseases*, *8*(11), e3301. https://doi.org/10.1371/JOURNAL.PNTD.0003301

Macias Torres, M., & Naranjo Mayorga, F. (2022). Characterization of resilience in Aedes aegypti mosquito networks. *Physica A: Statistical Mechanics and Its Applications*, *596*, 127114. https://doi.org/10.1016/J.PHYSA.2022.127114

Martínez-de la Puente, J., Ruiz, S., Soriguer, R., & Figuerola, J. (2013). Effect of blood meal digestion and DNA extraction protocol on the success of blood meal source determination in the malaria vector Anopheles atroparvus. *Malaria Journal*, *12*(1), 109. https://doi.org/10.1186/1475-2875-12-109

Mayi, M. P. A., Bamou, R., Djiappi-Tchamen, B., Fontaine, A., Jeffries, C. L., Walker, T., Antonio-Nkondjio, C., Cornel, A. J., & Tchuinkam, T. (2020). Habitat and Seasonality Affect Mosquito Community Composition in the West Region of Cameroon. *Insects*, *11*(5), 312. https://doi.org/10.3390/insects11050312

Mayi, M. P. A., Foncha, D. F., Kowo, C., Tchuinkam, T., Brisco, K., Anong, D. N., Ravinder, S., & Cornel, A. J. (2019). Impact of deforestation on the abundance, diversity, and richness of Culex mosquitoes in a southwest Cameroon tropical rainforest. *Journal of Vector Ecology*, *44*(2), 271–281. https://doi.org/10.1111/jvec.12359

McDermott, E. G., & Mullens, B. A. (2018). The Dark Side of Light Traps. *Journal of Medical Entomology*, *55*(2), 251–261. https://doi.org/10.1093/JME/TJX207

Miot, E. F., Calvez, E., Aubry, F., Dabo, S., Grandadam, M., Marcombe, S., Oke, C., Logan, J. G., Brey, P. T., & Lambrechts, L. (2020). Risk of arbovirus emergence via bridge vectors: case study of the sylvatic mosquito Aedes malayensis in the Nakai district, Laos. *Scientific Reports 2020 10:1*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-64696-9

Möhlmann, T. W. R., Wennergren, U., Tälle, M., Favia, G., Damiani, C., Bracchetti, L., & Koenraadt, C. J. M. (2017). Community analysis of the abundance and diversity of mosquito species (Diptera: Culicidae) in three European countries at different latitudes. *Parasites and Vectors*, *10*(1), 1–12. https://doi.org/10.1186/S13071-017-2481-

1/FIGURES/4

Mongan, A. E., Tuda, J. S. B., & Runtuwene, L. R. (2019). Portable sequencer in the fight against infectious disease. *Journal of Human Genetics 2019 65:1*, *65*(1), 35–40. https://doi.org/10.1038/s10038-019-0675-4

Morris, R. J., Gripenberg, S., Lewis, O. T., & Roslin, T. (2014). Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*, *17*(3), 340–349. https://doi.org/10.1111/ele.12235

Muturi, E. J., Dunlap, C., Ramirez, J. L., Rooney, A. P., & Kim, C. H. (2019). Host blood-meal source has a strong impact on gut microbiota of Aedes aegypti. *FEMS Microbiology Ecology*, *95*(1). https://doi.org/10.1093/FEMSEC/FIY213

Muturi, E. J., Njoroge, T. M., Dunlap, C., & Cáceres, C. E. (2021). Blood meal source and mixed blood-feeding influence gut bacterial community composition in Aedes aegypti. *Parasites & Vectors*, *14*(1), 83. https://doi.org/10.1186/s13071-021-04579-8

Ndenga, B. A., Mulaya, N. L., Musaki, S. K., Shiroko, J. N., Dongus, S., & Fillinger, U. (2016). Malaria vectors and their blood-meal sources in an area of high bed net ownership in the western Kenya highlands. *Malaria Journal*, *15*(1), 1–10. https://doi.org/10.1186/S12936-016-1115-Y/TABLES/6

Orsborne, J., Furuya-Kanamori, L., Jeffries, C. L., Kristan, M., Mohammed, A. R., Afrane, Y. A., O'Reilly, K., Massad, E., Drakeley, C., Walker, T., & Yakob, L. (2019). Investigating the blood-host plasticity and dispersal of Anopheles coluzzii using a novel field-based methodology. *Parasites & Vectors*, *12*(1), 143. https://doi.org/10.1186/s13071-019-3401-3

Pathak, A. K., Shiau, J. C., Freitas, R. C. S., & Kyle, D. E. (2023). Blood meals from 'dead-end' vertebrate hosts enhance transmission potential of malaria-infected mosquitoes. *One Health*, *17*, 100582. https://doi.org/10.1016/J.ONEHLT.2023.100582

Peña, R., Schleuning, M., Dalerum, F., Donoso, I., Rodríguez-Pérez, J., & García, D. (2023). Abundance and trait-matching both shape interaction frequencies between plants and birds in seed-dispersal networks. *Basic and Applied Ecology*, *66*, 11–21. https://doi.org/10.1016/J.BAAE.2022.11.008

Pernat, N., Kampen, H., Jeschke, J. M., & Werner, D. (2021). Buzzing homes: Using citizen science data to explore the effects of urbanization on indoor mosquito communities. *Insects*, *12*(5), 374. https://doi.org/10.3390/INSECTS12050374/S1

Perrin, A., Glaizot, O., & Christe, P. (2022). Worldwide impacts of landscape anthropization on mosquito abundance and diversity: A meta-analysis. *Global Change Biology*, *28*(23), 6857–6871. https://doi.org/10.1111/GCB.16406

Pimentel, A. C., Del Valle, Y., Uzcátegui, S., Stocco De Lima, A. C., Silveira, F. T., Vasconcelos, T., Santos, D., Aoba, E., & Ishikawa, Y. (2022). Blood Feeding Sources of Nyssomyia antunesi (Diptera: Psychodidae): A Suspected Vector of Leishmania (Kinetoplastida: Trypanosomatidae) in the Brazilian Amazon. *Journal of Medical Entomology*. https://doi.org/10.1093/JME/TJAC108

Poulin, R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, *26*(10), 492–498. https://doi.org/10.1016/J.PT.2010.05.008

Reeves, L. E., Holderman, C. J., Gillett-Kaufman, J. L., Kawahara, A. Y., & Kaufman, P. E. (2016). Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasites & Vectors*, *9*(1), 503. https://doi.org/10.1186/s13071-016-1791-z

Reinert, J. F. (2000). NEW CLASSIFICATION FOR THE COMPOSITE GENUS AEDES (DIPTERA: CULICIDAE: AEDINI), ELEVATION OF SUBGENUS OCHLEROTATUS TO GENERIC RANK, RECLASSIFICATION OF THE OTHER SUBGENERA. AND NOTES ON CERTAIN SUBGENERA AND SPECIES. *Journal of the American Mosquito Contol Association*, *16*(3), 175–188.

Rivera-Hutinel, A., Bustamante, R. O., Marín, V. H., & Medel, R. (2012). Effects of sampling completeness on the structure of plant–pollinator networks. *Ecology*, *93*(7), 1593–1603. https://doi.org/10.1890/11-1803.1

Romero-Vega, L. M., Piche-Ovares, M., Soto-Garita, C., Barantes Murillo, D. F., Chaverri, L. G., Alfaro-Alarcón, A., Corrales-Aguilar, E., & Troyo, A. (2023). Seasonal changes in the diversity, host preferences and infectivity of mosquitoes in two arbovirus-endemic regions of Costa Rica. *Parasites and Vectors*, *16*(1), 1–12. https://doi.org/10.1186/S13071-022-05579-Y/FIGURES/4

Sang, R., Arum, S., Chepkorir, E., Mosomtai, G., Tigoi, C., Sigei, F., Lwande, O. W., Landmann, T., Affognon, H., Ahlm, C., & Evander, M. (2017). Distribution and abundance of key vectors of Rift Valley fever and other arboviruses in two ecologically distinct counties in Kenya. *PLOS Neglected Tropical Diseases*, *11*(2), e0005341. https://doi.org/10.1371/journal.pntd.0005341

Santiago-Alarcon, D., Havelka, P., Schaefer, H. M., & Segelbacher, G. (2012). Bloodmeal Analysis Reveals Avian Plasmodium Infections and Broad Host Preferences of Culicoides (Diptera: Ceratopogonidae) Vectors. *PLOS ONE*, *7*(2), e31098. https://doi.org/10.1371/JOURNAL.PONE.0031098

Serafim, T. D., Coutinho-Abreu, I. V., Oliveira, F., Meneses, C., Kamhawi, S., & Valenzuela, J. G. (2018). Sequential blood meals augment vector infectiousness by promoting Leishmania replication and triggering amplification of metacyclics via a novel retroleptomonad developmental stage. *Nature Microbiology*, *3*(5), 548. https://doi.org/10.1038/S41564-018-0125-7

Sonne, J., Vizentin-Bugoni, J., Maruyama, P. K., Araujo, A. C., Chávez-González, E., Coelho, A. G., Cotton, P. A., Marín-Gómez, O. H., Lara, C., Lasprilla, L. R., Machado, C. G., Maglianesi, M. A., Malucelli, T. S., Martín González, A. M., Oliveira, G. M., Oliveira, P. E., Ortiz-Pulido, R., Rocca, M. A., Rodrigues, L. C., … Dalsgaard, B. (2020). Ecological mechanisms explaining interactions within plant–hummingbird networks: morphological matching increases towards lower latitudes. *Proceedings of the Royal Society B*, *287*(1922). https://doi.org/10.1098/RSPB.2019.2873

Spotswood, E. N., Beller, E. E., Grossinger, R., Grenier, J. L., Heller, N. E., & Aronson, M. F. J. (2021). The Biological Deserts Fallacy: Cities in Their Landscapes Contribute More than We Think to Regional Biodiversity. *Bioscience*, *71*(2), 148. https://doi.org/10.1093/BIOSCI/BIAA155

Srisuka, W., Sulin, C., Sommitr, W., Rattanarithikul, R., Aupalee, K., Saeung, A., & Harbach, R. E. (2022). Mosquito (Diptera: Culicidae) Diversity and Community Structure in Doi Inthanon National Park, Northern Thailand. *Insects*, *13*(9), 814. https://doi.org/10.3390/INSECTS13090814/S1

Steiger, D. B. M., Ritchie, S. A., & Laurance, S. G. W. (2016). Land Use Influences Mosquito Communities and Disease Risk on Remote Tropical Islands: A Case Study Using a Novel Sampling Technique. *The American Journal of Tropical Medicine and Hygiene*, *94*(2), 314. https://doi.org/10.4269/AJTMH.15-0161

Takken, W., & Verhulst, N. O. (2013). Host Preferences of Blood-Feeding Mosquitoes. *Annual Review of Entomology*, *58*(1), 433–453. https://doi.org/10.1146/annurev-ento-120811-153618

Tandina, F., Niare, S., Almeras, L., Davoust, B., Doumbo, O. K., Raoult, D., Parola, P., & Laroche, M. (2020). Identification of mixed and successive blood meals of mosquitoes using MALDI-TOF MS protein profiling. *Parasitology*, *147*(3), 329–339. https://doi.org/10.1017/S003118201900163X

Tchouassi, D. P., Quakyi, I. A., Addison, E. A., Bosompem, K. M., Wilson, M. D., Appawu, M. A., Brown, C. A., & Boakye, D. A. (2012). Characterization of malaria transmission by vector populations for improved interventions during the dry season in the Kpone-on-Sea area of coastal Ghana. *Parasites & Vectors*, *5*, 212. https://doi.org/10.1186/1756-3305-5-212

Terry, J. C. D., & Lewis, O. T. (2020). Finding missing links in interaction networks. *Ecology*, *101*(7), e03047. https://doi.org/10.1002/ECY.3047

Thiemann, T. C., Wheeler, S. S., Barker, C. M., & Reisen, W. K. (2011). Mosquito Host Selection Varies Seasonally with Host Availability and Mosquito Density. *PLoS Neglected Tropical Diseases*, *5*(12), e1452. https://doi.org/10.1371/journal.pntd.0001452

Thomas, F., Lefèvre, T., Renaud, F., Elguero, E., Fontenille, D., Gouagna, L.-C., Dabiré, K. R., & Costantini, C. (2009). Beyond Nature and Nurture: Phenotypic Plasticity in Blood-Feeding Behavior of Anopheles gambiae s.s. When Humans Are Not Readily Accessible. *The American Journal of Tropical Medicine and Hygiene*, *81*(6), 1023–1029. https://doi.org/10.4269/ajtmh.2009.09-0124

Tirados, I., Costantini, C., Gibson, G., & Torr, S. J. (2006). Blood-feeding behaviour of the malarial mosquito Anopheles arabiensis: Implications for vector control. *Medical and*

Veterinary Entomology, *20*(4), 425–437. https://doi.org/10.1111/j.1365-2915.2006.652.x

Toju, H., Tanabe, A. S., & Sato, H. (2018). Network hubs in root-associated fungal metacommunities. *Microbiome 2018 6:1*, *6*(1), 1–16. https://doi.org/10.1186/S40168-018-0497-1

van de Straat, B., Sebayang, B., Grigg, M. J., Staunton, K., Garjito, T. A., Vythilingam, I., Russell, T. L., & Burkot, T. R. (2022). Zoonotic malaria transmission and land use change in Southeast Asia: what is known about the vectors. *Malaria Journal 2022 21:1*, *21*(1), 1–13. https://doi.org/10.1186/S12936-022-04129-2

Xing, S., & Fayle, T. M. (2021). The rise of ecological network meta-analyses: Problems and prospects. *Global Ecology and Conservation*, *30*, e01805. https://doi.org/10.1016/J.GECCO.2021.E01805

Yee, D. A., Dean Bermond, C., Reyes-Torres, L. J., Fijman, N. S., Scavo, N. A., Nelsen, J., & Yee, S. H. (2022). Robust network stability of mosquitoes and human pathogens of medical importance. *Parasites & Vectors*, *15*(1), 216. https://doi.org/10.1186/S13071-022-05333-4

Young, K. I., Medwid, J. T., Azar, S. R., Huff, R. M., Drumm, H., Coffey, L. L., Pitts, R. J., Buenemann, M., Vasilakis, N., Perera, D., & Hanley, K. A. (2020). Identification of Mosquito Bloodmeals Collected in Diverse Habitats in Malaysian Borneo Using COI Barcoding. *Tropical Medicine and Infectious Disease*, *5*(2), 51. https://doi.org/10.3390/tropicalmed5020051

Zhao, J., Tang, T., & Wang, X. (2020). Effects of landscape composition on mosquito population in urban green spaces. *Urban Forestry and Urban Greening*, *49*, 126626. https://doi.org/10.1016/j.ufug.2020.126626

6.6 Supplementary

Supplementary 1. Best practice and considerations for well and poorly described faunal regions**.**

Best practice and considerations for well described faunal regions.

Trap choice

As discussed in Chapter 2, variation in trapping success between taxa as a result of inherent biases of trapping methods will result in an uneven representation within biting Diptera-host interaction networks. Methods that utilise live hosts should be avoided as they will heavily bias the interaction data. The use of a combination of trapping methods to maximise representation of Diptera fauna is advisable. The USA Centre for Disease Control (CDC) miniature light trap with CO_2 bait has an attraction range of ~15m and is efficient in capturing a wide range of Diptera including mosquitos, black flies, sand flies and biting midges. As biting Diptera identification will likely solely be through molecular identification in well described faunal regions, modification of the trap, to include an ethanol collection chamber to halt blood meal digestion and minimise blood meal degradation, is highly advisable to improve identification success in downstream molecular work. Complementing 'active trapping' with a passive method, such as malaise traps, will provide less biased data on relative abundances.

Trap location

As discussed in Chapter 2 and demonstrated in Chapter 5, sampling location will influence biting Diptera species composition and the set of potential hosts they interact with. For example, blood meals from the Village habitat in Chapter 5 were dominated by human, and due to the limited sample number in these sites, potentially provided an over representative view of Diptera-human interactions. However, trapping location should be guided by the overall purpose of the sampling effort. Sampling in non-urban settings, such as Agricultural and the interface with Near-natural habitats, is advisable to examine the wider community of biting Diptera and their interactions and identify vectors that may act to bridge transmission cycles. Indeed, sampling in non-urban areas in which humans are often present

may also provide a clearer picture of transmission risk in urban settings than sampling in urban habitat alone.

Sampling effort

A greater sampling effort will likely result in a greater network resolution and a more complete representation of the community composition. Therefore, where logistically possible, sampling effort should be maximised. However, whilst species richness is sensitive to sampling effort, most unique functionally important taxa sampled at sites in a similar landscape context were identified after 2-3 days of sampling effort (Gorsich et al., 2019). Further, smaller sampling efforts and limiting sampling to only in peak seasons has been shown to be sufficient to capture the majority of functionally important species of a network (Hegland et al., 2010).

For repeated sampling efforts, it is highly advisable to extrapolate Diptera and host richness and Diptera-host interactions with sampling effort, to estimate sampling completeness and predict the number of sampling days required to capture the full community composition and interaction structure. Moreover, tools have recently been developed to guide sampling regimes and determine the appropriate sampling effort required to maximise network resolution and statistical power whilst accounting for fluctuations in community composition as a result of seasonality (Sedda et al., 2023).

Sample storage

As indicated in Chapter 4, storage of blood fed biting Diptera for extended periods of time in 95% ethanol at ambient room temperature is sufficient to maintain blood meal integrity. It may be advisable to store samples in cold storage during shipping (if applicable) where possible to avoid temperature fluctuations that may negatively impact DNA quality. Where the aim is to interrogate blood meals for pathogen presence, the use of cold storage, such as dry ice and liquid nitrogen, is highly advisable. It is recognised though, that the logistics of such storage methods may prove problematic in sites with limited accessibility.

Molecular analysis

In well-described faunal regions, reference libraries for molecular identification of samples are likely to be well resolved. It is advisable to examine existing databases, such as GenBank, prior to sampling to confirm that reference sequences are available for the Diptera and host species in the sampling region. The species likely to be present may be identified using existing literature. Examination of existing sequence databases will also guide primer selection and indicate which gene loci has the highest number of reference sequences and will therefore, provide the greatest network resolution.

Best practice and considerations for poorly described faunal regions.

Trap choice

As discussed in Chapter 2, variation in trapping success between taxa as a result of inherent biases of trapping methods will result in an uneven representation within biting Diptera-host interaction networks. Methods that utilise live hosts should be avoided as they will heavily bias the interaction data. The use of a combination of trapping methods to maximise representation of Diptera fauna is advisable. The USA Centre for Disease Control (CDC) miniature light trap with CO_2 bait has an attraction range of ~15m and is efficient in capturing a wide range of Diptera including mosquitos, black flies, sand flies and biting midges.

Modification of the trap to include an ethanol collection chamber, to halt blood meal digestion and minimise blood meal degradation, is highly advisable to improve identification success in downstream molecular work. However, in poorly described faunal regions, a preliminary sampling effort is recommended to indicate the range of Diptera that may be present. This effort should use a trapping method that captures live samples to maintain the integrity of morphological features, and aid in non-molecular species identification. Complementing 'active trapping' with a passive method, such as malaise traps, will provide less biased data on relative abundances.

Trap location

As with well-described faunal regions, sampling location will influence biting Diptera species composition and the set of potential hosts they interact with. For example, blood meals from the Village habitat in Chapter 5 were dominated by humans, and due to the limited sample number in these sites, potentially provided an over representative view of Diptera-human interactions. However, trapping location should be guided by the overall purpose of the sampling effort. Sampling in non-urban settings, such as Agricultural and the interface with Near-natural habitats, is advisable to examine the wider community of biting Diptera and their interactions and identify vectors that may act to bridge transmission cycles. Indeed, sampling in non-urban areas, in which humans are often present, may also provide a clearer picture of transmission risk in urban settings than sampling in urban habitat alone.

Sampling effort

As with well described faunal regions, greater sampling effort will likely result in a greater network resolution and a more complete representation of the community composition. Therefore, where logistically possible, sampling effort should be maximised. In poorly described regions, a preliminary sampling effort is advisable to indicate the range of Diptera and hosts that may be present, and guide the primary sampling regime. Whilst species richness is sensitive to sampling effort, most unique functionally important taxa sampled at sites in a similar landscape context were identified after 2-3 days of sampling effort (Gorsich et al., 2019). Moreover, whilst a smaller sampling effort and limiting sampling to only in peak seasons has been shown to be sufficient to capture the majority of functionally important species of a network (Hegland et al., 2010), it is advisable to maximise sampling effort where possible in poorly described regions.

For repeated sampling efforts, it is highly advisable to extrapolate Diptera and host richness and Diptera-host interactions with sampling effort, to estimate sampling completeness and predict the number of sampling days required to capture the full community composition and interaction structure. Tools have recently been developed to guide sampling regimes and determine the appropriate sampling effort required to maximise network resolution and statistical power whilst accounting for fluctuations in community composition as a result of seasonality (Sedda et al., 2023).

Sample storage

As with well-described faunal regions, storage of blood fed biting Diptera for extended periods of time in 95% ethanol at ambient room temperature is sufficient to maintain blood meal integrity. If applicable, it may be advisable to store samples in cold storage during shipping where possible to avoid temperature fluctuations that may negatively impact DNA quality. Where the aim is to interrogate blood meals for pathogen presence, the use of cold storage, such as dry ice and liquid nitrogen, is highly advisable. It is recognised though, that the logistics of such storage methods may prove problematic in sites with limited accessibility.

Molecular analysis

In poorly described faunal regions it is likely that a well-resolved reference library will be unavailable for both Diptera and host identification. Where possible, it is highly advisable to conduct a preliminary sampling effort to collect Diptera and host samples for morphological identification and barcoding to generate a novel reference library. Whilst this step will aid in species identification, it is likely that rare species will be underrepresented in the library. Consequently, it may be appropriate to limit downstream identification to genus or family level for poorly described individuals using reference sequences of similar individuals from other well resolved locations. When integrating individuals identified to genus or family level into network analysis, care should be taken to ensure that nodes are not under or over-represented by separating conspecific samples into unique nodes or combining nodes that likely represent separate species into a single node. Identification to morphospecies prior to molecular identification may aid in separating similarly identified samples into unique nodes. The utilisation of portable sequencing devices, such as the Oxford Nanopore MinION, could allow for barcoding to be conducted in the field, expediting the creation of a reference library. In order to aid future research, barcodes should be uploaded to a public sequence repository, such as BOLD and GenBank.

Supplementary 1. References

Gorsich, E. E., Beechler, B., Bodegom, P. M. van, Govender, D., Guarido, M. M., Venter, M., & Schrama, M. (2019). A comparative assessment of adult mosquito trapping methods to estimate spatial patterns of abundance and community composition in southern Africa. *BioRxiv*, 633552. https://doi.org/10.1101/633552

Hegland, S. J., Dunne, J., Nielsen, A., & Memmott, J. (2010). How to monitor ecological communities cost-efficiently: The example of plant–pollinator networks. *Biological Conservation*, *143*(9), 2092–2101. https://doi.org/10.1016/J.BIOCON.2010.05.018

Sedda, L., Taylor, B. M., Cain, R., Vajda, É. A., Tatarsky, A., & Lobo, N. F. (2023). TIMESS a power analysis tool to estimate the number of locations and repeated measurements for seasonally and clustered mosquito surveys. *Annals of Operations Research*, 1–17. https://doi.org/10.1007/S10479-023-05491-3/TABLES/1

Appendices

Appendix 1. Journal preprint: Predicting the tripartite network of mosquito-borne disease

Work completed external to this thesis but are derived from data generated during its creation.

Predicting the tripartite network of mosquito-borne disease

Tad A. Dallas[a,b,†], Sadie J. Ryan[c,d], Ben Bellekom[e], Anna Fagre[f], Rebecca Christofferson[a] and Colin J. Carlson[g,h]

[a] *Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70802*
[b] *Department of Biological Sciences, University of South Carolina, Columbia, SC, 29208*
[c] *Department of Geography, University of Florida, Gainesville, FL 32611 USA*
[d] *Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611, USA*
[e] *Department of Zoology, 11a Mansfield Road, Oxford OX1 3SZ, UK*
[f] *Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523, US.A.*
[g] *Center for Global Health Science and Security, Georgetown University Medical Center, Washington, D.C. 20007*
[h] *Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, D.C. 20007*
[†] *Corresponding author: tad.a.dallas@gmail.com*

Running title: A tripartite model of vector-borne disease

Author contributions: TAD, CJC, and SJR conceived the study. TAD and CJC performed the analysis. All authors contributed to manuscript writing.

Acknowledgements: Lauren Holian provided comments and discussion on an earlier version of this work. This work has been supported by funding to the Viral Emergence Research Initiative (VERENA) consortium, including a grant from the U.S. National Science Foundation (NSF BII 2021909).

Conflict of interest: The authors have no conflicts of interest to declare.

Keywords: Arbovirus, Host-virus, Link prediction, Networks, Vector-borne

Predicting the tripartite network of mosquito-borne disease

Abstract

The potential for a pathogen to infect a host is mediated by traits of both the host and pathogen, as well as the complex interactions between them. Arthropod-borne viruses (arboviruses) require an intermediate arthropod vector, which introduces an additional layer of compatibility filters. Existing computational models for the prediction of host-virus networks rarely incorporate the unique aspects of vector transmission, instead treating vector biology as a hidden, unobserved layer. Here, we explore two possible extensions to existing approaches, to address this nuance: first, we added vector traits into predictions of the bipartite host-virus network; and second, we used host, vector, and virus traits to predict the tripartite host-vector-virus network. We tested both approaches on the most thoroughly characterized group of arboviruses; mosquito-borne flaviviruses of mammals, including dengue, yellow fever, and Zika virus. Using host-virus models, we find that the inclusion of vector traits may improve inference in some cases, while viral traits proved to be the most important for model performance. Further, we found that it was possible to predict full life cycles (host-vector-virus links), but the model only showed fair performance, and was heavily influenced by the geographic bias of component input datasets (especially the dipteran biting data). Both approaches are interesting avenues for further model development, but our results keenly underscore a need to collect more comprehensive datasets to characterize arbovirus ecology, across a wide geographic scope, especially outside of North America, and to better identify molecular traits that underpin host-vector-virus interactions.

Introduction

Emerging viruses continue to pose a threat to human and wildlife populations [1]. A growing set of computational tools have explored viral dynamics in the context of species interaction networks using a set of tools called *link prediction models*. Typically, these represent hosts and viruses as a bipartite network of either known interactions (that occur in nature [2, 3]) or all possible interactions (including, for example, experimental infections [4]), with both represented as links in the network [5]. Host-virus link prediction models are predominantly trained on the genomic, immunological, morphological, and ecological traits of hosts and viruses (e.g., [6, 7]), while some approaches also leverage information on the latent structure of the network instead of, or in addition to, these traits [8, 9]. The objective of these modeling exercises is to learn about the underlying biology, explain and reproduce patterns found in nature, and anticipate what future dynamics of viral emergence could look like. For example, many models use networks to understand why some viruses can infect humans but others cannot, with the objective of identifying animal viruses that could someday infect humans for the first time. In most cases, these models assume that any given "link" between a host and a virus could represent a self-contained transmission cycle (though not necessarily *onwards* transmission, e.g., West Nile virus in humans and horses [10]).

Vector-borne disease (VBD) transmission substantially complicates this conceptual framework. Vector-borne viruses require an additional species–usually an arthropod (hence arthropod-borne viruses, or *arboviruses*)–to move them between hosts, which adds complexity into their ecology, epidemiology, and evolution. For example, in the case of arboviruses, the presence of both virus and suitable hosts is not necessarily sufficient for transmission, and the presence or absence of suitable vectors (e.g., their geographic distributions or host preferences) may be a latent variable in ecological datasets [11]. Moreover, the "compatibility filters" that can be inferred from the host-virus network will be incomplete, as models will miss both the molecular and physiological determinants of vector-virus compatibility (i.e., vector competence) and the behavioral and ecological determinants

of vector-host compatibility (i.e., biting preferences, in the case of blood-feeding arthropods). If vectors are entirely omitted from the inference process, a model might therefore reach spurious conclusions about whether a given host and virus are incompatible based on their biology, or otherwise miss key drivers of network structure; for example, arboviruses have been shown repeatedly to have a higher-than-expected host breadth [12].

No one canonical approach exists to address vector transmission in link prediction studies. Vector transmission could be described as a binary trait of viruses, which may help make some distinctions (e.g., separating the ecology of mosquito-borne and tick-borne flaviviruses from counterparts like hepatitis C), but leaves much to be desired in terms of information content (e.g., not distinguishing the tick- and mosquito-borne flaviviruses). The possibility of incorporating more detailed information on vector-borne transmission into these models has been under-explored, likely because arboviruses are usually seen as a complicated exception to existing datasets, rather than a feature with significant impacts on network structure. Incorporating traits characterizing the life cycle of arboviruses might improve model performance, given that virus traits are often sparser than host traits, and their interactions usually have non-additive but positive effects on model performance. However, adding sparse traits that only describe some of the viruses in the network could also reduce accuracy if the network includes a mix of vector-borne and directly-transmitted viruses.

Alternately, vectors could be added directly into the network as an additional layer of nodes (Figure 1). While previous work has predicted vector-virus networks [13], none have predicted *host-vector-virus* networks. Existing network models have been used to predict undetected links in *tripartite* networks [14], but this has yet to be explored for ecological networks. This approach would be much more informative than the bipartite form, but also requires difficult-to-obtain data: sylvatic VBD cycles tend to be characterized one at a time in scientific literature (e.g., "*Culex quinquefasciatus* vectors West Nile virus in house finches"). While available datasets could be used to reconstruct these cycles from each of their component

parts (biting preferences, vector competence, and host-virus compatibility), to our knowledge, this has not previously been explored in predictive work.

To address this, we developed two new approaches and tested them on mosquito-borne flaviviruses, a well-studied group that includes important zoonoses like dengue, West Nile, yellow fever, and Zika viruses. Through a synthesis of existing data sources, we combined data on mammal-virus associations [12], vector-flavivirus associations [13], and diptera-mammal biting preferences [15]. We combined these data into one mammal-mosquito-flavivirus network, which can also be reduced down to a mammal-flavivirus network where viruses' mosquito communities are represented as node metadata. Using boosted regression trees (BRT; a machine learning method popular in ecological modeling, also sometimes called gradient boosting machines), we tested two approaches to predicting vector-borne transmission as an aspect of the host-virus network. First, we predicted the mammal-flavivirus network using every possible combination of host, vector, and virus traits, as metadata for any given host-virus association, assuming that additional data layers would enhance model performance. This was generally shown to be true, although the combination of host and vector trait data was not informative compared to the incorporation of viral trait data. Second, we developed a tripartite model of vector-borne disease transmission, in which each link represents a known host-vector-virus link and attempted to predict those complete cycles using traits of hosts, mosquito vectors, and viruses. We found that these models performed more poorly on average, but that they were able to make better than random predictions, including some of relevance to arboviral ecology and human health.

Methods

Host, vector, and virus data Host-virus interaction data were obtained from the CLOVER database [16], a manually- and programmatically-curated database of host-virus associations built by reconciling four disparate datasets (the Host-

Parasite Phylogeny Project, or HP3 [12]; the Global Mammal Parasite Database v2.0 [17]; the Enhanced Infectious Disease Database [18]; and an unnamed dataset curated by Shaw *et al.* [19]). We used CLOVER release 0.1.2, which includes data on 5,477 known interactions between 831 viruses of 1,085 mammal species. These data have been carefully cleaned for taxonomic quality control and include detailed metadata on interaction evidence. These data are also part of a larger open database called The Global Virome in One Network (VIRION), the largest open atlas of vertebrate-virus associations [20]. Although more data is available from this source, we restricted our analysis to the manually-curated data to prevent inclusion of spurious interactions.

Vector-virus association data were taken from a previous study that aimed to predict the mosquito-flavivirus network. [13] These data include 334 associations between 180 mosquito species and 37 flaviviruses. Host-vector association data were taken from a recent study of dipteran biting networks [15]. These data describe 1744 associations between 255 biting dipteran species and 214 hosts (including 67 mammals). Trait data for hosts, vectors, and viruses were assembled from published sources. Thirty-three traits on mosquito life history, ecology, and geography and 22 traits on viral features, were taken from the Evans *et al.* study of the mosquito-flavivirus network [13]. Finally, we used a total of 18 traits on mammal life history, ecology, and morphology from the PanTHERIA database [21].

Modeling approach Boosted regression tree (BRT) models were used to model host-virus and host-vector-virus associations. BRT models have previously been used to model species distributions [22], predict associations in bipartite networks [23, 24, 25, 5], and in other conservation and management settings e.g., [26]. Much of the diversity of applications can be attributed in part to the allowance for nonlinear responses and variable interactions in BRT models. Since the regression tree is hierarchical, "upstream" splits based on one variable influence "downstream" splits, which automatically models variable interactions. Further, the process of boosting enhances learning on complex data, as the process produces many regres-

sion trees with a small number of splits, each of these "weak learners" iteratively build on previous trees to account for the remaining variation. This approach removes the need to partition variance among submodels, as the goal is not to examine the components of variance explained, but to assess overall model performance with the inclusion or exclusion of particular variable sets. Models were trained in the *R* statistical programming language [27] using the *gbm* package [28].

Model 1: Modeling mammal-virus associations as a bipartite network We used the mammal and virus trait data as described above. However, mosquito vector "traits" were created by calculating the number of mosquito species in a given genus which were demonstrated to transmit a particular flavivirus [13]. This is because each host-virus association could be transmitted by any number of mosquito species, creating a range of trait values that may be less informative than simply knowing breadth and composition of the vector community. This resulted in a total of 19 mosquito vector covariates, ranging in value from 0 to 22 species. We removed covariates with less than 25% data coverage, resulting in 13 host traits, 19 mosquito covariates (as virus traits), and 17 virus traits.

The data were split into 80% training and 20% testing sets, where model performance was assessed on the 20% test set. A total of 20 models per covariate group were fit in order to account for the random train/test split. These same 20 train/test divisions were used across the different covariate models, as we trained every possible combination of host, vector, and virus trait data to predict host-virus associations. Together, this resulted in a dataset that allows the estimation of the relative influence of host traits, viral traits, and vector community data on resulting mammal-virus associations. We sampled background data by randomly combining host and virus species, resulting in 25% known positive associations and 75% background data.

We subset these data in two different ways, to explore how vector data may improve prediction of 1) flaviviruses for which we have some vector data (235 known host-virus associations) and 2) all vector-borne viruses (3016 host-virus

associations). This breakdown corresponds to data subsets of 1) only mosquito-borne flaviviruses present in [13] and 2) all viruses that were recorded as vector-borne (or unknown) in the Clover data [16]. We present the flavivirus-specific results here, which are qualitatively similar to the more general models for all vector-borne viruses, which are in the Supplemental Materials.

Model 2: Modeling mammal-mosquito-virus associations as a tripartite network Using the same data resource as used above on host-virus associations, we now considered the identity of the mosquito vector species, and the association between the vector and virus [13], and the feeding association between mosquito vector and mammal species [15]. While host and virus traits were largely the same as considered above, the mosquito vector traits consisted of a set of 33 mosquito vector traits from [13]. Host and virus traits must have 75% of data coverage – the same as in *Model 1* – to be included in this analysis. This resulted in 8 host traits, 29 vector traits, and 16 virus traits. A tripartite link – detailing the full host-vector-virus cycle – was only considered if there were all three associations; host-vector association, vector-virus association, and host-virus association. This creates a situation where a host and vector species may interact, and that vector may be infected by a virus, but this is not a confirmed link if there is no evidence that the host is infected by the virus.

A total of 135 full tripartite links were documented. We sampled background data by randomly combining host, vector, and virus species and then adding enough unique host-vector-virus background points to have 50% true tripartite links and 50% background data. Models were trained in the same manner as in `Model 1`.

Assessing model performance Model performance was quantified using two measures; accuracy and the area under the receiver operating characteristic (AUC). Accuracy was defined as the correctly estimated positives (true positives) and negatives (true negatives) over all the predictions, capturing the fraction of times the model correctly classified host-virus associations in the holdout data. Accuracy is

bounded between 0 and 1, where larger values correspond to higher model performance. AUC is a widely used metric of model discrimination that captures the ability of the classifier to rank positive instances higher than negative instances. AUC is bounded between 0 and 1, where a random model will perform with AUC of 0.5 on average, and values closer to 1 indicate higher model performance.

Data and code availability *R* code and data to reproduce the analyses is available on figshare at
`https://doi.org/10.6084/m9.figshare.17033309`.

Results

Model 1: The mammal-virus models Models trained only on host ($AUC = 0.57$) or vector ($AUC = 0.46$) traits consistently performed poorly at the task of host-virus link prediction (Figure 2), though the viral trait model performed well ($AUC = 0.95$). Generally, combinations of predictor features led to improved model performance. The full model including host, vector, and virus traits performed extremely well ($AUC = 0.96$). However, both the host-virus and vector-virus traits only models also performed extremely well (performance differences among these models were essentially indistinguishable; Figure 2). The inclusion of viral traits seems to have been particularly important; for comparison, the model using host and vector traits to predict host-virus associations barely performed better than random ($AUC = 0.59$).

Variables important for predicting host-virus associations were generally conserved across submodels considering all combinations of host, vector, and virus traits (Figure 3). In the full bipartite model, the most informative variable was whether a virus was found in the Pacific region (likely a proxy for Zika virus, which spread through Pacific islands preceding the epidemic in the Americas). Other important characteristics predictive of host-virus associations in bipartite models including virus traits were disease severity, genome length, year of virus

isolation, if the virus is found in Africa or Australia, and viral clade. In models that omitted virus traits, the top predictors represented host allometry (body mass and metabolic rate, an unsurprising axis of variation) and *Culex* association, which likely captures a latent split between some bird-reservoired viruses (e.g., West Nile virus) and primate-reservoired ones (e.g., dengue and Zika virus).

Overall, our results suggest that models learned from vector trait data, particularly in the full model, where the contribution of each individual variable is more diffuse. However, our findings also indicate that the inclusion of vector data only minimally improved performance after data on hosts and viruses was already available. As host-virus models are usually trained only on host and virus trait data, our findings suggest that the incorporation of vector data into a host-virus model is an imperfect way to explore the role of vectors in structuring the host-virus network. However, this also suggests that improved arthropod trait data could improve model performance, and thus the importance of the vector cannot be overlooked.

Finally, we investigated whether including vector trait data would improve performance even if only available for a subset of data informing the network. To test this, we trained the model on a network that included all the arboviruses present in the CLOVER dataset, even though viral trait data and vector associations were only known for flaviviruses. We found that the model using just host and virus traits performed substantially worse here ($AUC = 0.70$) than the flavivirus-only model with those traits ($AUC = 0.95$). We found that the best performing models were those that used vector and virus traits ($AUC = 0.98$) and those that included host, vector, and virus traits ($AUC = 0.99$; Figure 2). We suggest that this finding indicates that adding data on the vector aspect of transmission may be useful even when it only covers a subset of species in the network.

Model 2: The tripartite model Models trained on tripartite (i.e., host-vector-virus) associations had moderate explanatory power (mean $AUC = 0.64$ (0.065); mean Accuracy = 0.66 (0.046) out of 100 models trained on random subsets). This

lower model performance could simply be due to the smaller amount of data used for training (recall that only 135 full tripartite links were known), or the imbalance between the number of potential full tripartite links given host, vector, and virus diversity, and the small number of realized links (see the small number of red links in Figure 4). Although the model's performance was only fair, we found that the model still predicted higher suitability for tripartite links where one or two of the three possible components were confirmed (Figure 5), even though these would be recorded as a "0" outcome variable the same as if none of them were known. We suggest that this indicates the model was identifying and reproducing real biological signals of compatibility.

The top nine covariates to predicting tripartite (i.e., host-vector-virus) associations were host ($n = 5$) or virus ($n = 4$) traits (Figure 6). The top predictors mostly reflected the geography of transmission (host geographic range size, virus transmission in Asia, vector presence in Africa), the life history of the host (age at first birth, lifespan, weaning age, and neonate body mass), and aspects of viral transmission (genome length and transmission by non-mosquito arthropods).

The predictions made by the tripartite model suggest the model may be able to recover interesting or important biologically-plausible interactions. Both the top predicted "undiscovered" human-mosquito-virus links (Table 1) and mammal-mosquito-virus links (Table 2) heavily over-represent a small number of viruses, in particular Wesselbron virus and West Nile virus. This is driven by the existing level of sampling in the data: West Nile has the greatest number of known hosts ($n = 103$ species) and mosquito vectors ($n = 51$); Wesselbron has the second highest number of vectors ($n = 41$), though many fewer hosts ($n = 11$; ranked #13). This "rich-get-richer" has been previously debated as a strength or weakness for link prediction models; it may be that models are identifying a genuine biological signal of generality (which is known to be true for these viruses), but they may also be recapitulating sampling bias [5, 29] and underpredicting link probabilities for undersampled species. Indeed, the richness of flavivirus data available to us in this study is likely largely due to a discovery and data synthesis bump in the wake

of the Zika virus epidemic in the Americas. The mammal-mosquito-virus predictions also contain a visible signal of geographic bias: most of the top predictions either involve agricultural species (pigs, *Sus scrofa*; cows, *Bos taurus*; or sheep, *Ovis aries*), synanthropic species (black rats, *Rattus rattus*), or charismatic North American species (the opossum, *Didelphis virginiana*; the raccoon, *Procyon lotor*; the white-tailed deer, *Odocoileus virginianus*). These likely reflect a compounded bias between the host-virus association data and the biting data, the latter of which is particularly limited to North American and European species.

Despite the signal of data bias in these predictions, the models reveal several predictions of biological interest. For example, *Anopheles hyrcanus* is predicted as a possible vector of Kokobera virus in humans. The virus was implicated in an outbreak of acute polyarticular illness in Australia in the 1980s based on serology, but it remains poorly understood [30]. The virus was first isolated from *Culex annulirostris*, which also vectors Japanese encephalitis virus and a handful of others; *An. hyrcanus* is a European and Asian mosquito only currently known to vector Japanese encephalitis virus. Similarly, the model predicts that *Culex tritaeniorhynchus* – the main vector of Japanese encephalitis virus, found in southeast Asia – could transmit Murray Valley encephalitis virus in wallabies (*Macropus agilis*). Neither the Australian virus nor the host have been recorded in association with this vector, but as of 2021, the mosquito has been detected in Australia [31], indicating the possibility that this interaction could now emerge.

Discussion

In this study, we considered two approaches to incorporate arboviral life cycles into link prediction models of the mammal-flavivirus network. First, we used a host-virus (bipartite) framework, and assessed the relative influence of including different trait covariates. We found that viral traits were the strongest contributor to model performance, and the incorporation of host and vector traits into the bipartite models did little to improve model performance. Second, we explored how

these models could be extended to predict the entire host-vector-virus (tripartite) network. This framing is both inherently more complex than the host-virus predictive problem, and is massively limited by the availability of training data, but appears promising for future development.

Neither of these approaches provided a complete solution to the host-vector-virus prediction problem, though their limitations differ slightly, with different implications for next steps. Adding vector community data to the bipartite (host-virus) models may be useful where data allow, but may be less important when more detailed, biologically meaningful viral trait data are available. Compared to synthetic datasets of animal ecology, life history, and morphology, only a handful of viral traits (e.g., genome length or disease severity) are available in a standardized format, to the point that viral host range is itself often used as a viral trait (e.g., our "primate" or "bird" traits, or "host breadth" (see Table S1)). Recently, some studies have begun to use immunogenetic or genome composition variables to characterize host and virus compatibility more directly [32, 33, 34, 35, 36, 37]; comparable features for vectors are not yet available or tested in this framework. Shifting towards these kinds of predictors could help models identify more meaningful signals of virus-animal compatibility, and proportionally reduce the signal of bias in predictions.

In contrast, directly modeling the host-vector-virus tripartite network addresses the nuance of vector transmission head-on, but this problem is more severely data limited. As a result, these predictions are very visibly influenced by the geographic and taxonomic bias in the component datasets. However, these data limitations can be addressed by investment in future work characterizing arboviral life cycles in understudied areas [38]. Vector-virus combinations can be tested in the laboratory, including in model-experiment feedback designs that leverage existing predictions (e.g., [13]) much like model-guided fieldwork can be used to optimize viral discovery [25]. Similarly, further investigation of mosquito biting behavior will help resolve the host-vector component [15], highlighting the need for "basic" natural history research even on mosquitoes that are not known to be primary vectors of human

disease.

Our study is the first to attempt modeling the entire tripartite host-vector-virus network. This is a clear knowledge gap in existing approaches to modeling the host-virus network: identifying a suitable host-pathogen association that has no shared vector may not accurately estimate spillover risk. This may be particularly relevant to efforts to identify viruses with undiscovered zoonotic potential, as the presence or absence of human-biting mosquitoes will be a key contributor to their emergence risk [39]. Similarly, the tripartite framework can provide useful insights into the establishment of sylvatic cycles in interepidemic periods or upon expansion into new geographic areas. The ability of arboviruses to persist in non-human hosts may determine whether an epidemic ends as immunity grows (like Zika virus in the Americas, which was primarily transmitted human-to-human by *Aedes aegypti* and *Ae. albopictus*) or instead becomes a regular occurrence (e.g., yellow fever in the Americas, which is maintained by *Haemogogus spp.* and *Sabethes spp.* in non-human primates, between human epidemics driven by *Aedes aegypti*). These are likely to be particularly important nuances as arboviruses continue to spread around an increasingly globalized world in a changing climate [40, 41, 42, 43]

The broader question of "how should we model multi-layer ecological interaction networks" is also one that is likely to have broader implications in computational ecology. For example, there are other cases where researchers are interested the traits that structure tripartite networks, such as bat-bat fly-pathogen networks or plant-pest-parasitoid networks. Multilayer networks are also a topic of increasing interest in network science and mathematics, which will likely open doors for more advanced predictive approaches than the extensions we propose here. This is therefore a promising space for the development of future models, particularly if approached through the lens of iterative validation and data collection [25].

References

[1] Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L. & Daszak, P., 2008 Global trends in emerging infectious diseases. *Nature* **451**, 990–993.

[2] Kading, R. C., Kityo, R. M., Mossel, E. C., Borland, E. M., Nakayiki, T., Nalikka, B., Nyakarahuka, L., Ledermann, J. P., Panella, N. A., Gilbert, A. T. *et al.*, 2018 Neutralizing antibodies against flaviviruses, babanki virus, and rift valley fever virus in ugandan bats. *Infection Ecology & Epidemiology* **8**, 1439215.

[3] Fagre, A. C., Lewis, J., Miller, M. R., Mossel, E. C., Lutwama, J. J., Nyakarahuka, L., Nakayiki, T., Kityo, R., Nalikka, B., Towner, J. S. *et al.*, 2021 Subgenomic flavivirus rna (sfrna) associated with asian lineage zika virus identified in three species of ugandan bats (family pteropodidae). *Scientific Reports* **11**, 1–8.

[4] Malmlov, A., Bantle, C., Aboellail, T., Wagner, K., Campbell, C. L., Eckley, M., Chotiwan, N., Gullberg, R. C., Perera, R., Tjalkens, R. *et al.*, 2019 Experimental zika virus infection of jamaican fruit bats (artibeus jamaicensis) and possible entry of virus into brain via activated microglial cells. *PLoS Neglected Tropical Diseases* **13**, e0007071.

[5] Albery, G. *et al.*, 2021 The science of the host-virus network. *Nature Microbiology* p. in press.

[6] Dallas, T., Park, A. W. & Drake, J. M., 2017 Predicting cryptic links in host-parasite networks. *PLoS computational biology* **13**, e1005557.

[7] Elmasri, M., Farrell, M. J., Davies, T. J., Stephens, D. A. *et al.*, 2020 A hierarchical bayesian model for predicting ecological interactions using scaled evolutionary relationships. *Annals of Applied Statistics* **14**, 221–240.

[8] Ghasemian, A., Hosseinmardi, H., Galstyan, A., Airoldi, E. M. & Clauset, A., 2020 Stacking models for nearly optimal link prediction in complex networks. *Proceedings of the National Academy of Sciences* **117**, 23393–23400.

[9] Becker, D. J., Albery, G. F., Kessler, M. K., Lunn, T. J., Falvo, C. A., Czirják, G. Á., Martin, L. B. & Plowright, R. K., 2020 Macroimmunology: The drivers and consequences of spatial patterns in wildlife immune defence. *Journal of Animal Ecology* **89**, 972–995.

[10] Bowen, R. A. & Nemeth, N. M., 2007 Experimental infections with west nile virus. *Current Opinion in Infectious Diseases* **20**, 293–297.

[11] Huang, Y.-J. S., Higgs, S. & Vanlandingham, D. L., 2019 Arbovirus-mosquito vector-host interactions and the impact on transmission and disease pathogenesis of arboviruses. *Frontiers in microbiology* **10**, 22.

[12] Olival, K. J., Hosseini, P. R., Zambrana-Torrelio, C., Ross, N., Bogich, T. L. & Daszak, P., 2017 Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650.

[13] Evans, M. V., Dallas, T. A., Han, B. A., Murdock, C. C. & Drake, J. M., 2017 Data-driven identification of potential zika virus vectors. *Elife* **6**, e22053.

[14] Chen, C., Grady, S. K., Ellingson, S. R. & Langston, M. A., 2021 Gene-disease-drug link prediction using tripartite graphs. In *Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 1–1.

[15] Bellekom, B., Hackett, T. D. & Lewis, O. T., 2021 A network perspective on the vectoring of human disease. *Trends in Parasitology* pp. S1471–4922.

[16] Gibb, R., Albery, G. F., Becker, D. J., Brierley, L., Connor, R., Dallas, T. A., Eskew, E. A., Farrell, M. J., Rasmussen, A. L., Ryan, S. J. *et al.*, 2021 Data proliferation, reconciliation, and synthesis in viral ecology. *BioScience* p. in press.

[17] Stephens, P., Pappalardo, P., Huang, S., Byers, J., Farrell, M., Gehman, A., Ghai, R., Haas, S., Han, B., Park, A. *et al.*, 2017 Global mammal parasite database version 2.0. *Ecology* **98**, 1476–1476.

[18] Wardeh, M., Risley, C., McIntyre, M. K., Setzkorn, C. & Baylis, M., 2015 Database of host-pathogen and related species interactions, and their global distribution. *Scientific data* **2**, 1–11.

[19] Shaw, L. P., Wang, A. D., Dylus, D., Meier, M., Pogacnik, G., Dessimoz, C. & Balloux, F., 2020 The phylogenetic range of bacterial and viral pathogens of vertebrates. *Molecular ecology* **29**, 3361–3379.

[20] Carlson, C. J., Gibb, R. J., Albery, G. F., Brierley, L., Connor, R., Dallas, T., Eskew, E. A., Fagre, A. C., Farrell, M. J., Frank, H. K. *et al.*, 2021 The global virome in one network (virion): an atlas of vertebrate-virus associations. *bioRxiv* .

[21] Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C. *et al.*, 2009 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648.

[22] Elith, J., Leathwick, J. R. & Hastie, T., 2008 A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802–813.

[23] Dallas, T. A., Han, B. A., Nunn, C. L., Park, A. W., Stephens, P. R. & Drake, J. M., 2019 Host traits associated with species roles in parasite sharing networks. *Oikos* **128**, 23–32.

[24] Dallas, T. A. & Becker, D. J., 2021 Taxonomic resolution affects host- parasite association model performance. *Parasitology* **148**, 584–590.

[25] Becker, D. J., Albery, G. F., Sjodin, A. R., Poisot, T., Bergner, L. M., Dallas, T. A., Eskew, E. A., Farrell, M. J., Guth, S., Han, B. A. *et al.*, 2021 Optimizing predictive models to prioritize viral discovery in zoonotic reservoirs. *The Lancet Microbe* p. in press.

[26] Soykan, C. U., Eguchi, T., Kohin, S. & Dewar, H., 2014 Prediction of fishing effort distributions using boosted regression trees. *Ecological Applications* **24**, 71–83.

[27] R Core Team, 2020 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[28] Greenwell, B., Boehmke, B., Cunningham, J. & Developers, G., 2020 *gbm: Generalized Boosted Regression Models*. R package version 2.1.8.

[29] Gibb, R., Albery, G. F., Mollentze, N. F., Eskew, E. A., Brierley, L., Ryan, S. J., Seifert, S. N. & Carlson, C. J., 2021 Mammal virus diversity estimates are unstable due to accelerating discovery effort. *bioRxiv* .

[30] Boughton, C. R., Hawkes, R. A. & Nairn, H. M., 1986 Illness caused by a kokobera-like virus in south-eastern australia. *Medical journal of Australia* **145**, 90–92.

[31] Lessard, B. D., Kurucz, N., Rodriguez, J., Carter, J. & Hardy, C. M., 2021 Detection of the japanese encephalitis vector mosquito culex tritaeniorhynchus in australia using molecular diagnostics and morphology. *Parasites & vectors* **14**, 1–11.

[32] Babayan, S. A., Orton, R. J. & Streicker, D. G., 2018 Predicting reservoir hosts and arthropod vectors from evolutionary signatures in rna virus genomes. *Science* **362**, 577–580.

[33] Brierley, L. & Fowler, A., 2021 Predicting the animal hosts of coronaviruses from compositional biases of spike protein and whole genome sequences through machine learning. *PLoS Pathogens* **17**, e1009149.

[34] Mollentze, N., Babayan, S. & Streicker, D., 2021 Identifying and prioritizing potential human-infecting viruses from their genome sequences. *bioRxiv* pp. 2020–11.

[35] Poisot, T., Ouellet, M.-A., Mollentze, N., Farrell, M. J., Becker, D. J., Albery, G. F., Gibb, R. J., Seifert, S. N. & Carlson, C. J., 2021 Imputing the mammalian virome with linear filtering and singular value decomposition. *arXiv preprint arXiv:2105.14973* .

[36] Wardeh, M., Blagrove, M. S., Sharkey, K. J. & Baylis, M., 2021 Divide-and-conquer: machine-learning integrates mammalian and viral traits with network features to predict virus-mammal associations. *Nature Communications* **12**, 1–15.

[37] Young, F., Rogers, S. & Robertson, D. L., 2020 Predicting host taxonomic information from viral genomes: A comparison of feature representations. *PLoS computational biology* **16**, e1007894.

[38] Swei, A., Couper, L. I., Coffey, L. L., Kapan, D. & Bennett, S., 2020 Patterns, drivers, and challenges of vector-borne disease emergence. *Vector-Borne and Zoonotic Diseases* **20**, 159–170.

[39] dos Passos Cunha, M., Duarte-Neto, A. N., Pour, S. Z., Ortiz-Baez, A. S., Černỳ, J., de Souza Pereira, B. B., Braconi, C. T., Ho, Y.-L., Perondi, B., Sztajnbok, J. *et al.*, 2019 Origin of the são paulo yellow fever epidemic of 2017–2018 revealed through molecular epidemiological analysis of fatal cases. *Scientific reports* **9**, 1–10.

[40] Ryan, S. J., Carlson, C. J., Mordecai, E. A. & Johnson, L. R., 2019 Global expansion and redistribution of aedes-borne virus transmission risk with climate change. *PLoS neglected tropical diseases* **13**, e0007213.

[41] Ryan, S. J., Carlson, C. J., Tesla, B., Bonds, M. H., Ngonghala, C. N., Mordecai, E. A., Johnson, L. R. & Murdock, C. C., 2021 Warming temperatures could expose more than 1.3 billion new people to zika virus risk by 2050. *Global Change Biology* **27**, 84–93.

[42] Kraemer, M. U., Reiner, R. C., Brady, O. J., Messina, J. P., Gilbert, M., Pigott, D. M., Yi, D., Johnson, K., Earl, L., Marczak, L. B. *et al.*, 2019 Past

and future spread of the arbovirus vectors aedes aegypti and aedes albopictus. *Nature microbiology* **4**, 854–863.

[43] Messina, J. P., Brady, O. J., Golding, N., Kraemer, M. U., Wint, G. W., Ray, S. E., Pigott, D. M., Shearer, F. M., Johnson, K., Earl, L. *et al.*, 2019 The current and future global distribution and population at risk of dengue. *Nature microbiology* **4**, 1508–1515.

Figures

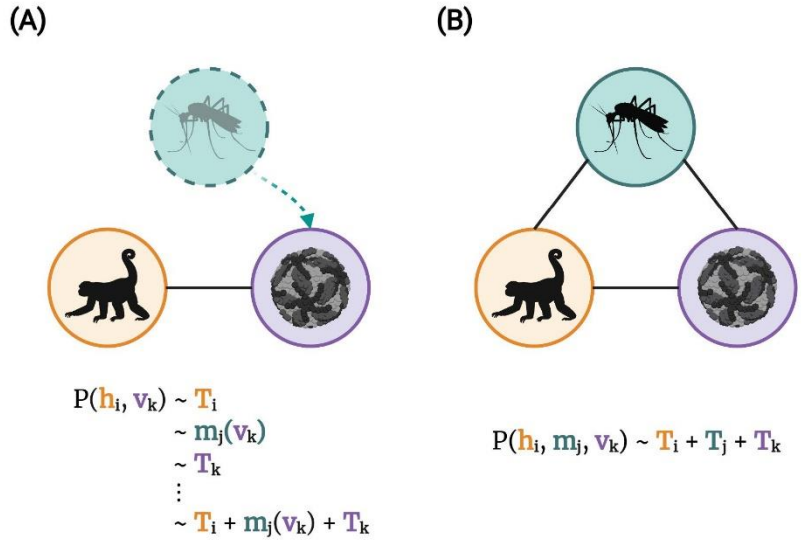


Figure 1: (A) Predicting host-virus associations (a bipartite network) based on host traits (T_h), virus traits (T_v), and vector communities ($m(v)$) associated with viruses, is a different problem than (B) predicting host-vector-virus associations (a tripartite network) based on host traits, vector traits, and virus traits. In this paper, we consider both solutions as approaches to end goals like forecasting potential novel associations or spillover scenarios

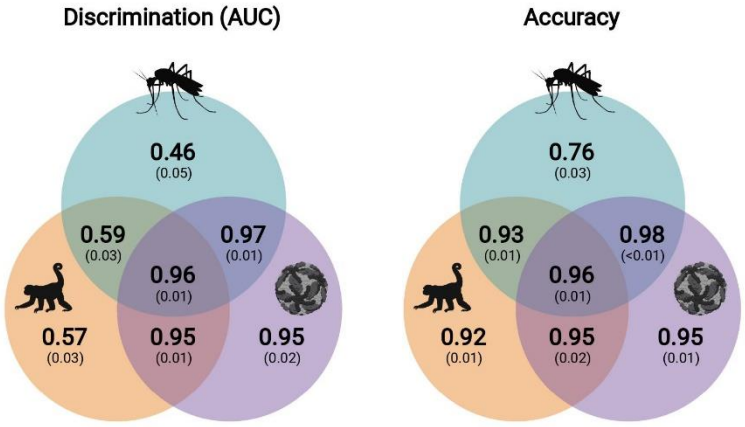


Figure 2: Model performance – quantified using AUC (left panel) and accuracy (right panel) – was highest when host, vector, and virus traits were included in the model (reported values are mean and standard deviation based on 20 model runs, assessing performance on a random 20% subset of the data). However, host-virus association model performance was not appreciably increased by the addition of 32 host trait covariates, suggesting that host-virus associations may be best predicted by considering information on the vector and the virus.

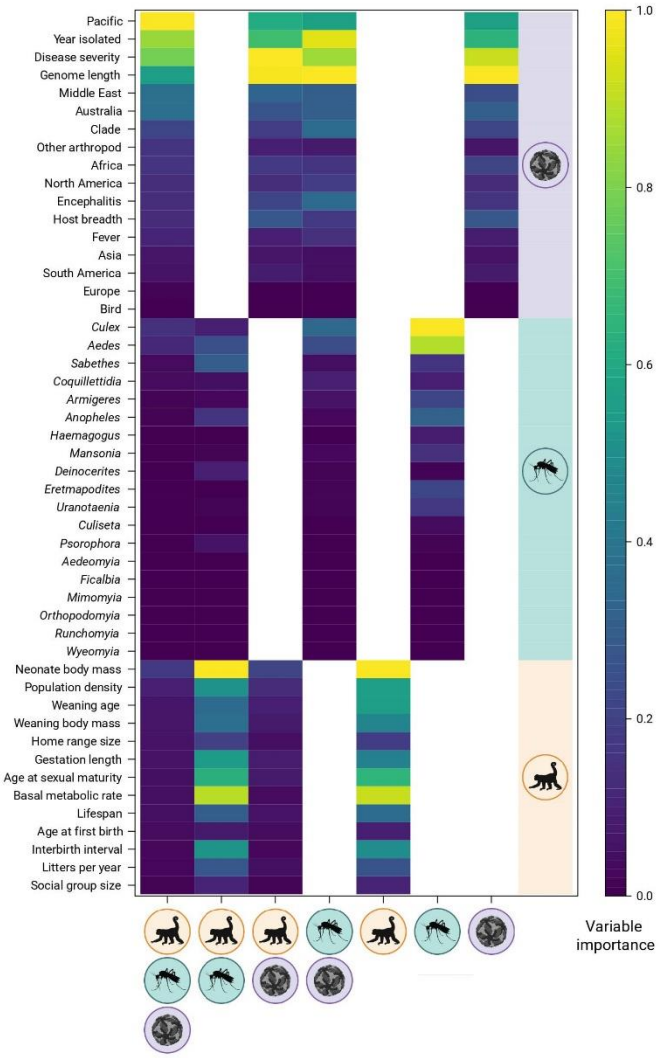


Figure 3: The relative importance of host (orange), vector (blue), and virus (purple) traits on predictive model performance. Each column corresponds to a different combination of these three trait groups, with the first column corresponding to the full model (as indicated at the bottom of each column using the glyphs). Variables are ordered based on the full model.

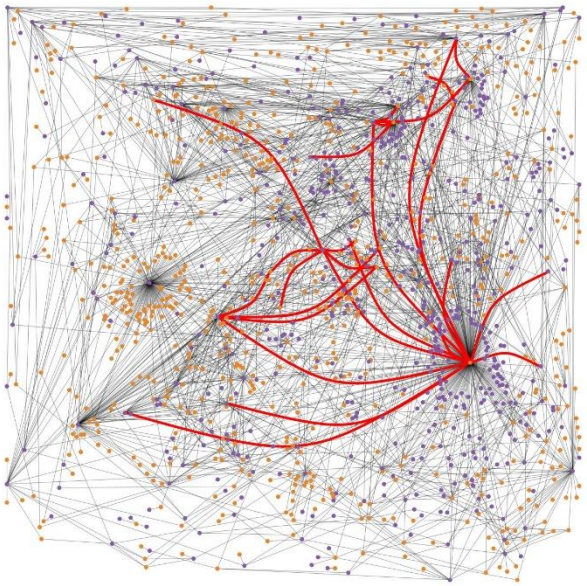


Figure 4: Full graph of host-virus associations (host species are in orange and viruses in purple), where links between host and virus species represent known associations. Red links are those which the full host-vector-virus cycle is known.

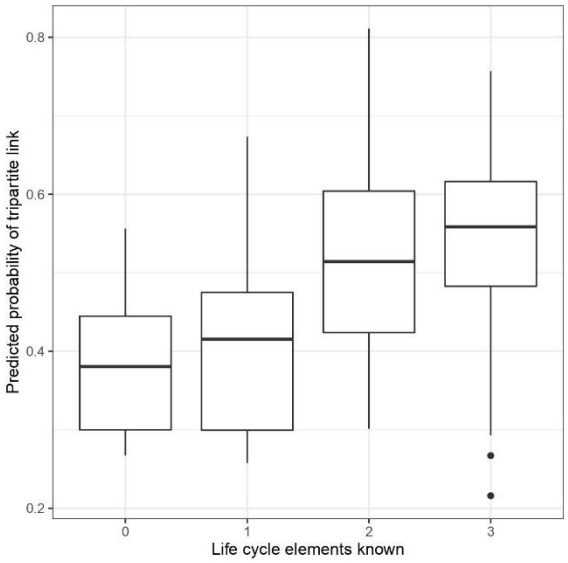


Figure 5: The tripartite model predicts a higher average probability for associations that have one or two links known (which are still not recorded as positive values in the training data) than those with no elements known to be possible. This suggests that the model is capable of more than just recapitulating the data, and is able to distinguish different levels of biological plausibility within unknown tripartite elements.

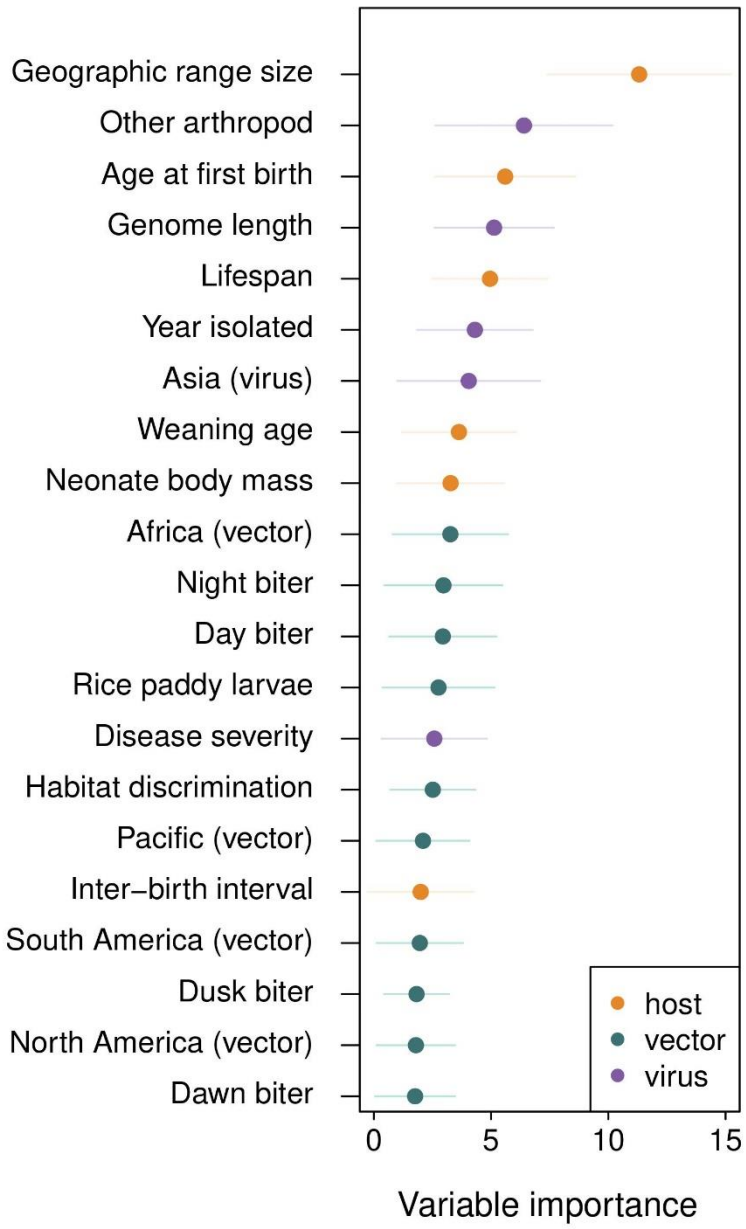


Figure 6: The relative importance of host (orange), vector (grey), and virus (blue) traits on predictive model performance in the tripartite model.

Tables

Table 1: **Top predicted epidemic cycles in humans.** All vectors are known to be human-biting; all viruses are known to be zoonotic based on either clinical or serological data.

Host	*Mosquito*	Virus	Prob
H. sapiens	*Culex pipiens*	Wesselsbron virus	0.81
H. sapiens	*Aedes aegypti*	West Nile virus	0.74
H. sapiens	*Aedes aegypti*	Japanese encephalitis virus	0.73
H. sapiens	*Culex pipiens*	Murray Valley encephalitis virus	0.73
H. sapiens	*Culex sitiens*	West Nile virus	0.72
H. sapiens	*Aedes scapularis*	West Nile virus	0.68
H. sapiens	*Mansonia uniformis*	West Nile virus	0.68
H. sapiens	*Anopheles coustani*	Wesselsbron virus	0.67
H. sapiens	*Culex pipiens*	Yellow fever virus	0.67
H. sapiens	*Aedes aegypti*	Ilheus virus	0.66
H. sapiens	*Aedes albopictus*	Ilheus virus	0.65
H. sapiens	*Anopheles hyrcanus*	Kokobera virus	0.62
H. sapiens	*Culex nigripalpus*	Wesselsbron virus	0.61
H. sapiens	*Aedes cantans*	Wesselsbron virus	0.61
H. sapiens	*Mansonia africana*	West Nile virus	0.61
H. sapiens	*Culex perexiguus*	Wesselsbron virus	0.60
H. sapiens	*Culex thalassius*	Wesselsbron virus	0.60
H. sapiens	*Culex gelidus*	West Nile virus	0.60
H. sapiens	*Culex annulirostris*	St. Louis encephalitis virus	0.59
H. sapiens	*Anopheles pharoensis*	West Nile virus	0.59

Table 2: **Top predicted enzootic cycles.** All mammals in the top 20 are either species found alongside humans (cows, sheep, pigs, and rats) or easily-sampled species from eastern North America (deer, raccoons, and possums).

Host	*Mosquito*	Virus	Prob
Sus scrofa	*Aedes albopictus*	West Nile virus	0.70
Sus scrofa	*Mansonia uniformis*	Wesselsbron virus	0.69
Didelphis virginiana	*Aedes aegypti*	Wesselsbron virus	0.67
Didelphis virginiana	*Aedes albopictus*	Wesselsbron virus	0.67
Sus scrofa	*Anopheles coustani*	Wesselsbron virus	0.66
Sus scrofa	*Culex quinquefasciatus*	Japanese encephalitis virus	0.65
Procyon lotor	*Culex tritaeniorhynchus*	West Nile virus	0.62
Odocoileus virginianus	*Anopheles pharoensis*	Wesselsbron virus	0.62
Procyon lotor	*Culex pipiens*	Japanese encephalitis virus	0.61
Didelphis virginiana	*Aedes aegypti*	West Nile virus	0.59
Odocoileus virginianus	*Aedes albopictus*	St. Louis encephalitis virus	0.56
Bos taurus	*Aedes vexans*	West Nile virus	0.56
Bos taurus	*Culex tritaeniorhynchus*	West Nile virus	0.55
Procyon lotor	*Culex annulirostris*	West Nile virus	0.55
Macropus agilis	*Culex tritaeniorhynchus*	Murray Valley encephalitis virus	0.54
Bos taurus	*Anopheles maculipennis*	Wesselsbron virus	0.54
Ovis aries	*Culex quinquefasciatus*	Ilheus virus	0.53
Procyon lotor	*Culex tarsalis*	West Nile virus	0.52
Rattus rattus	*Aedes aegypti*	Zika virus	0.51
Odocoileus virginianus	*Culex tritaeniorhynchus*	Banzi virus	0.51

Supplemental Material

Predicting the tripartite network of mosquito-borne disease

Trait data

Trait data were compiled from a variety of sources, with host trait data coming from PanTHERIA [21], and vector and virus trait data from Evans et al. 2017 [13].

Table S1: Host, vector, and virus covariates considered in the models of host-virus (`h-v` column) and host-vector-virus (`h-m-v` column) associations. See the Pantheria documentation (https://esapubs.org/archive/ecol/E090/184/metadata.htm) for more information on host trait variables.

Taxa	Variable	Units	Definition	h-v	h-m-v
Host					
	Lifespan	days	Maximum observed lifespan	☑	☑
	Age at sexual maturity	days	Age at which individual is sexually mature	☑	☑
	Home range size	km^2	Area used by individual for daily tasks on average	☑	
	Gestation length	days	Period of time young are gestated	☑	☑
	Neonate body mass	grams	Average neonate body mass	☑	☑
	Population density	n / km^2	Number of individuals per unit area, on average	☑	
	Age at first birth	days	Age at which females give birth to their first litter	☑	☑
	Litters per year	n / year	Average number of litters per year	☑	
	Max lifespan	months	Longest observed lifespan	☑	☑
	Basal metabolic rate	mLO_2 / hr	Individual metabolic rate	☑	

	Interbirth interval	months	Period in between reproductive bouts	☑	☑
	Age at eye opening	days	Time when neonates open eyes	☑	
	Social group size	count	Number of individuals per social group	☑	
	Adult forearm length	mm	Length of adult forearm	☑	
	Dispersal age	days	Age at which young leave parents	☑	
	Neonate head-body length	mm	Body length of neonates	☑	
	Weaning age	days	Period of time when young stop weaning	☑	☑
	Weaning body mass	grams	Mass of young during weaning	☑	☑
Vector					
	Mosquito genus	numeric	Number of mosquito species of genus that vector a given virus		☑
	Human biter	1/0	Vector bites humans	☑	
	Host breadth	count	Number of host species bitten	☑	
	Non-primate mammals	1/0	Are non-primate mammals bitten	☑	
	Geographic range	count	Number of countries species collected	☑	
	Geographic location	-	Could include any or all of the following; Africa, Middle East, Australia, Pacific, Asia, Europe, North America, South America	☑	
	Biting behavior	-	Timing of biting behavior. Can be; dawn, day, dusk, and/or night	☑	
	Artificial container	1/0	Vector breeds in artificial containers	☑	

	Oviposition site	-	Larval site. Could include one or all of; treehole, container, pond, rockhole, marsh, swamp, ground pool, or rice paddy	☑	
	Permanent habitat	1/0	Species uses permanent habitat	☑	
	Habitat discrimination	count	number of habitat types	☑	
	Urban preference	1/0	vector shows urban preference	☑	
	Indoor preference	1/0	vector shows indoor preference	☑	
	Viral range	count	Number species within genus to harbor virus	☑	
Virus					
	Average genome length	numeric	Length of viral genome	☑	☑
	Geographic location	-	Could include any or all of the following; Africa, Middle East, Australia, Pacific, Asia, Europe, North America, South America	☑	☑
	Clade	–	Viral clade (roman numerals)	☑	
	Year isolated	year	Virus isolation year	☑	☑
	Other arthropod	1/0	Vectored by other arthropods	☑	
	Host breadth	count	number of known hosts	☑	☑
	Encephalitis	1/0	Virus causes encephalitis	☑	☑
	Fever	1/0	Virus causes fever	☑	☑
	Disease severity	numeric	How severe is disease	☑	☑
	Bird host	1/0	Virus infects birds		☑

What if we consider all vector-borne viruses?

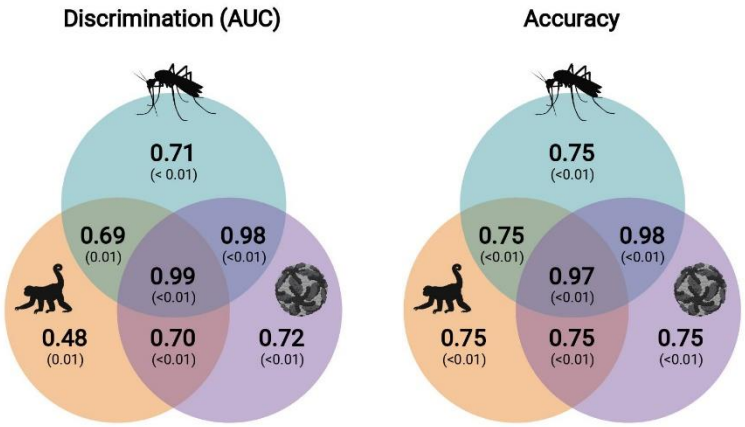


Figure S1: Model performance – quantified using AUC (left panel) and accuracy (right panel) – was highest when host, vector, and virus traits were included in the model (reported values are mean and standard deviation based on 20 model runs, assessing performance on a random 20% subset of the data). However, host-virus association model performance was not appreciably increased by the addition of vector data compared to just host and vector traits (AUC = 0.98).

Different models and similar predictions

When predicting host-virus associations, the different models had quite different variable importance values, apart from obviously having different explanatory variables. One question we had was whether models trained on different covariates would not only have similar overall performance, but identify the same likely host-virus associations as other models. To explore this graphically, we generated a correlation matrix (Figure S2), where we find strong positive relationships between different model predictions. Interestingly, the least positive correlation was from the full model, suggesting that the predictions from the full model differed from models which consisted of nested subsets of the same features as in the full model.

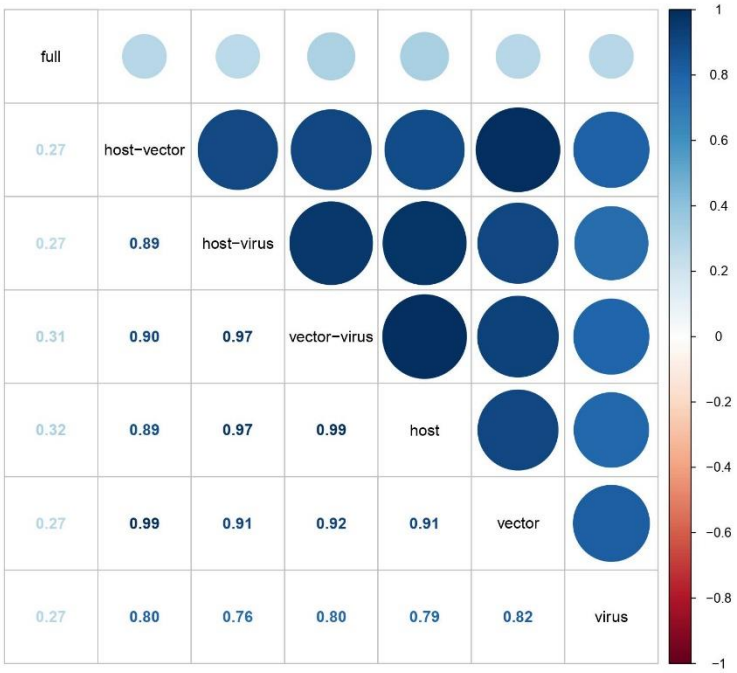


Figure S2: Correlation matrix between model predictions of the full set of subset models including different combinations of host, vector, and virus traits. The full model, including all traits, resulted in the predictions that were most weakly related to the other model predictions, though this model had similar performance as other models (see main text Figure 2). Lower triangle values and color scale correspond to Pearson's correlation coefficient values.

Appendix 2. Sampling protocol for excluded chapter. The sampling for this chapter has been completed, however due to COVID-19 related delays this work was not completed to a degree suitable for inclusion in the main body of this thesis**.**

Distance from Human Habitation Protocol

Goals

We will collect biting Diptera samples from sites located across a large spatial scale (6km) to examine the effect of distance from human habitation and vegetation complexity on biting insect community composition and biting Diptera-host interaction frequencies.

Sampling will take place within and around the Kalakpa nature reserve, a 325 km^2 region comprised of savannah grasslands and primary forest. Trapping will take place across 9 weeks (6 trapping weeks) to ensure a large amount of sample collection and increase the likelihood of successful capture of blood fed insects. Sampling will use CDC light traps, complemented with sugar fermented yeast CO_2 bait, BG Sentinel traps and malaise traps. The trapping regime will incorporate multiple trapping nights at each distance. There will be four sampling sites in the village and at each of four distances along a transect away from Pojive village (50m, 1,500m, 3,000m and 6,000m).

Sampling Site Locations (With detail of village and peri-village locations)

Week number	First site	Y (latitude)	X (longitude)	Second site	Y (latitude)	X (longitude)	Third Site	Y (latitude)	X (longitude)	Fourth Site	Y (latitude)	X (longitude)
Date												
1	Site 1. Forest Edge C	6.43407	0.32665	Site2. Forest Interior D	6.42979	0.34015	Site 3. Peri-Village B	6.44505	0.31766	Site 4. Village A	6.44846	0.31761
Date												
2	Site 5. Village B	6.44750	0.31726	Site 6. Forest Edge A	6.43330	0.32546	Site 7. Forest Interior C	6.42944	0.33986	Site 8. Far Forest A	6.40847	0.35646
Date												
3	Site 9. Forest Interior A	6.42830	0.33888	Site 10. Far Forest C	6.40951	0.35736	Site11. Peri-Village D	6.44584	0.31871	Site 12. Forest Edge D	6.43438	0.32702
Date												
4	Site 13. Village D	6.44633	0.31713	Site 14. Peri-Village A	6.44498	0.31720	Site 15. Forest Interior B	6.42869	0.33920	Site 16. Far Forest D	6.40987	0.35766
Date												

5	Site 17. Peri-Village C	6.44550	0.31843	Site 18. Village C	6.44685	0.31713	Site 19. Forest Edge B	6.43356	0.32587	Site 20. Far Forest B	6.40883	0.35678
Date												
6												

Sampling Location Schedule and Co-ordinates

Protocol

Sampling Regime

The sampling will follow a 2 week on and 1 week off schedule. During the one week off, materials should be replenished (ethanol, falcon tubes etc), in preparation for the following 2 weeks. A total of 6 weeks of sampling (within a 9-week period) will be carried out across the 20 collections sites. The 6th week of sampling will be used as a contingency week, to be used should issues arise during the preceding weeks that interrupt sampling.

Each week of sampling will consist of 2 trapping days at each of 4 sites, with trapping at 2 sites running concurrently (Figure 1.). The complete schedule and order of trapping sites is on the previous page (page 3).

Sites	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
First site	setup		collection				Rest
Second site	setup		collection				Rest
Third site				setup		collection	Rest
Fourth site				setup		collection	Rest

Figure 1. Sampling schedule, black boxes indicate that sampling at those sites will be carried out during those days.

All traps will run for 2 periods of 24 hours within 3 days at each sampling site. After the first 24-hour period, samples will be collected, batteries replaced, and the traps will be reset. After the second 24-hour period, samples will be collected, and the traps taken down in preparation for moving to a new site the following day.

General Equipment List

- 4x 6V batteries
- 4x 12V batteries
- 6V and 12V battery charger
- 4x CDC light traps + CDC light trap bulbs
- 4x BGS traps
- 4x Malaise traps
- GPS device + batteries + battery charger
- 2x Multiplug power strips
- Tape measure
- 4kg sugar
- 500g yeast
- 8x 1.5L bottles
- 4x hose bottle caps
- 4x CDC light trap stand
- Rubber mallet
- Pencils + rubber
- Falcon tubes
- Forceps
- Ethanol
- Sample labels
- Covered clipboard
- String
- Multitool
- Machete
- Sorting tray
- Funnel
- Icebox

The sampling checklist should be printed and carried into the field as a reminder of the steps outlined below

Nightly Protocol

Each night, the nightly checklist (see sampling checklist) should be completed.

Daily Protocol

Travelling to the sampling site

Prior to travelling to the sampling site, you will check-in with your ranger assistant and discuss the site locations for the coming day, identifying if access to the nature reserve will be required for the day's trapping. Equipment should be stored in the vehicle and driven between sampling locations. All sampling locations are either placed within the village or along a road leading into the Kalakpa nature reserve. Sampling sites are separated into 5 categories (village, peri-village, forest edge, forest interior, and far forest), with 4 sampling sites within each category. There is a distance of approximately 50 metres between each site, within each category, and there is a distance of approximately 50 metres from the road to the closest site (see sampling site location image and the sampling location schedule and co-ordinates (Page 3)).

If the road becomes inaccessible due to an obstruction that cannot be cleared and you are close to your next sampling site, then proceed to it on foot, and arrange for the obstruction to be cleared with the ranger assistant. If you are too far away from the sampling site (use your judgement to assess this), then note the site number and swap it with a closer unvisited site. On the scheduled visit to the substitute site, proceed instead to the site that had been temporarily unavailable.

Selection of sampling site

Travel to the sampling sites with all the required equipment, using the GPS device to identify the correct location. If a selected sampling location is unsuitable or inaccessible, then the traps should be placed in an appropriate alternative location, as close as possible to the designated sites, with the same habitat classification. For example, if a sampling site within a village habitat is inaccessible, then the relocated trapping site should also be within the village. Coordinates of the new sampling site should be recorded in the pre-printed data sheet (see Site data sheet).

Setting up the traps and collecting samples

At each sampling site you will set up three traps: (1) a CDC light trap with CO_2 bait, (2) Malaise trap, and (3) BGS trap. For detailed protocols for each trap and CO_2 bait protocol, see Appendices 1,2, 3 & 4. Traps should be positioned at least 15 meters away from one another, roughly centred on the GPS location for the site, to prevent attraction bias and reduced trapping efficiency. If the sampling site is on the border of two habitat types, ensure that all traps are placed within the same habitat.

All traps will be run for two 24-hour periods at each sampling site. Whilst there is no set time by which a trap should be set up, some sampling locations will be more inaccessible than others, so plenty of time should be given to allow for travel to each sampling location. The exact time the traps are set up at a site should be recorded in the site data sheet and traps should emptied and reset in the order that they were initially placed, to allow for each trap to run for 24 hours. The site data sheet should be completed, and site habitat information recorded (see Appendix 5 for detailed habitat instructions). Record the presence of any vertebrate within the 30-metre sampling area in the data sheet using a Yes (present) or No (absent) format. This process should then be repeated at the second sampling site. If a logistics issue prevents the visiting of two sites in a day, the site number should be noted, and traps should be run there in the contingency week (trapping week 6).

Following the first 24-hour trapping period, each trap should be emptied and reset (see Appendices 1, 2, 3 & 4 for detailed instructions). Sample labels (Appendix 7) should be completed in pencil and placed within each trapping cup or bottle. Each site, trap type and sampling day will have an associated 'lot' number. For example, the CDC trap on day 1 at the first site is site 1.1 and lot 0001 and CDC trap day 2 at site 1 is site 1.2 is lot 0004 (see Appendix 8 for full list of lot numbers). Lot numbers will be pre-printed on the sampling labels, so care should be taken to ensure that the correct lot number is on the sample label before placing it into the sampling tube.

If a trap has failed overnight (e.g., the battery has stopped powering the light) note the site number, trap type, and trap coordinates and revisit the site during the contingency week using the failed trap type only.

Removing the traps and storing samples

Following the second 24-hour trapping period, each trap will be taken down in the order they were put up to allow for them to run for the full 24-hour period (see Appendices 1, 2, 3 & 4 for detailed instructions). Care should be taken with all traps when repacking them to prevent any damage to

their netting, which would lower trap effectiveness. Traps and collected samples should be stored safely within the vehicle before travelling to the second sampling location.

If the collected samples have little or no ethanol remaining, then they should be topped up with fresh 95% ethanol. Therefore, ethanol should be carried to the sampling sites even on the final day of each sampling location.

Upon leaving the field, samples will be transferred from their collecting cups, bottles or beakers into large falcon tubes. A funnel may be used to transfer the contents. However, if there is too much ethanol in the collecting cup to fit in a falcon tube, then samples should be emptied into a sorting tray. In this situation, fill the falcon tube with ethanol and use forceps to transfer individual insects into their tube (care should be taken to not damage individual samples and ensure that all samples have been placed into their new tube).

Examine the sample label to ensure that it is has not been damaged; if damage has occurred replace the label using the blank sample labels, taking care to fill in the correct site number, lot number and trap type. Place the paper label inside the corresponding falcon tube. The filled falcon tube should then be placed securely within the freezer, and the process repeated for the remaining collected samples. If at any point a sample tube becomes frozen when placed in the freezer, allow the sample to thaw and replace the ethanol using a fresh bottle of 95% ethanol.

Returning to the Lab

Prior to returning to Accra, make a note of which materials you are lacking (e.g. ethanol, falcon tubes etc). During the off week, materials should be replenished and packed in preparation to return to the field.

After two weeks in the field, you will return to Accra with the collected samples. Ensure that all samples are securely placed within the transport icebox and their screw caps are tightened. Place ice blocks below and on top of the samples to ensure they remain cold during the journey. Upon returning to the University of Ghana, place all the samples into the pre-allocated position within the laboratory freezer.

Sampling checklist

The following checklist should be printed out and taken into the field.

Nightly checklist

1. Pre-mix CO_2 bait sugar solution and weigh yeast
2. Recharge used 6V, 12V & GPS batteries
3. Pre-measure 2L of 95% ethanol into a transport bottle
4. Ensure coordinates of site locations for the following day are correctly entered into the handheld GPS device.
5. Pack equipment required for the following day

Equipment list to take to field

1. 2x 6V and 2x 12V batteries
2. 2x CDC light traps, 2x BGS traps, 2x Malaise traps
3. 2x CO_2 bait bottles and 2x pre-weighed yeast
4. GPS device and spare fully charged batteries
5. Tape measure
6. 2L of 95% ethanol
7. Pre-printed site data sheet, sample labels, and clipboard
8. Pencil and rubber
9. Machete
10. Rubber mallet

Trap set up

For each site:

1. Set up 3 traps (CDC, BGS, and Malaise) with approximately 15m distance between each trap
2. Complete site data sheet (Appendix 6), recording time, habitat percentages, vertebrate presence and location, entering new site coordinates if required
3. Update site coordinates in GPS device if required

Trap emptying and collection

For each site:

1. Replace trapping bottles and cups (containing sufficient quantities of ethanol)
2. Place sample label within each respective collected bottle or cup
3. Replace CDC 6V and BGS 12V batteries
4. Replace CDC CO_2 bait bottles
5. Take down each of the 3 traps (on the final day of trapping at each site)

After returning from field actions/checklist

1. Check sample labelling and replace labels if needed
2. Transfer samples and sample labels from collections bottles and cups into long term storage falcon tubes, topping up ethanol if required
3. Transfer samples to freezer
4. Complete the nightly checklist

Appendices

Appendix 1. CDC light trap protocol

What is a CDC light trap?

A CDC light trap is an active (uses attractants) trapping method to collect a wide range of biting Diptera. It uses a light source placed above a fan to attract insects and suck them into the collection net. We supplement the light source with a CO_2 bait to increase the trapping potential.

Where to install?

As the light traps will be attached to a T shaped pole that is driven into the ground, and are an active trapping method, the trap can be placed in most locations, though preferably within a small clearing to allow the light to be seen. However, if possible, the traps should not be placed in direct sunlight, which will increase the rate of evaporation of ethanol within the collection cup and reduce the effectiveness of the trap. If the poles become damaged and cannot be easily replaced, traps may be hung from trees or other vegetation at around body or head height.

How to install

1. Prior to travelling to the sampling sites, test the light and fans of both CDC light traps by briefly connecting them to a 6V battery. If the light does not work, but the fan is spinning, unplug the battery and replace the light bulb. If neither the fan nor light bulb work, replace the battery. If the trap still does not work, then make a note of which trap is broken and use a different CDC light trap.
2. Assemble the T-shaped poles and drive the pole securely into the ground, using a rubber mallet if necessary. The pole should not be driven so far into the ground that attached traps will touch the ground.
3. Carefully attach the collection cup to the bottom of the CDC light by placing the bottom of the netting over the rim of the collection cup. Lower the collection cup outer rim over the netting and screw onto the collection cup, trapping netting between the cup and the outer rim (Figure 1.)

4. Ensure that the collection cup is held straight when the trap is lifted up. Once the cup is secure and straight, fill the cup three quarters of the way up with ethanol.
5. Attach the plastic rain protector (disc) to the top of the CDC light trap.
6. Securely attach the CDC light trap to one of the arms of the T bar using string, allowing the trap to hang off the arm (Figure 2.). It may be helpful to place the trap close to the upright pole to prevent the T bar from leaning to one side prior to attaching the CO_2 bait.
7. Attach the CO_2 bait bottle (see CO_2 bait bottle protocol for instructions) to the other arm of the T bar using string (Figure 2.).
8. Securely attach the plastic hose near the CDC light source using string, taking care to limit the blockage of the entrance to the CDC trap (Figure 3).
9. Place the 6V battery off the ground, using a rock or other raised area. Ensure that the battery is sheltered from the rain using the available vegetation as cover. If there is no suitable vegetation, a plastic bag can be placed over the battery.
10. Attach the power cables from the CDC trap to the 6V battery: red cable to red receiver and black cable to black receiver.
11. Ensure that the light on the trap has now been lit and the fan is working.

How to collect

1. Examine the trap to determine if any live insects remain within the net portion of the trap. If insects are present, then use a wash bottle containing ethanol to 'shoot' insects with ethanol so that they fall into the collection cup.
2. Once there are no living insects within the trap, carefully unscrew the outer rim of the collection cup and remove the collection cup. It may be helpful to remove the entire CDC trap from the T bar prior to this.
3. If there is no or a small amount of ethanol remaining within the collection cup, then top it up with fresh ethanol.
4. Place the preprepared sample label into the collection pot
5. Securely fasten the screw lid (note that the containers are not completely watertight and so should be transported upright with limited shaking to ensure minimal spillage).
6. Replace the collection cup with a fresh collection cup containing ethanol and replace the 6V battery (unless on the sites final collection day).

How to dismantle (on final day at collection site only)

1. Following the collection of the samples, work through the ***how to install*** instructions in reverse.
2. Take great care not to damage the CDC trap netting, as holes in the net will drastically reduce its efficiency.

Figure 1. Collection cup fastened to CDC light trap (note the white outer rim securing the cup to the black material)

Figure 2. T bar with attached CDC light trap (note that the light trap pictured will not be the one you use)

Figure 3. Hose attached to the CDC trap, below the rain protector and next to the light source (note that the light trap pictured will not be the one you use).

Appendix 2. CO_2 bait protocol

Equipment required per bottle:

- 1L of water
- 100g of sugar
- 7g of yeast
- 1.5L bottle
- Regular bottle lid
- Bottle lid with hose attached

How to create

1. The night before visiting each sampling location, create the sugar solution, by adding 100g of sugar to each 1L of water and mix thoroughly to ensure all sugar is dissolved. You will require two 1L bottles of sugar solution, one for each trapping site.
2. Measure out 7 grams of yeast twice, storing each 7 grams in an individual small screw top falcon tube.
3. Transport the bottles to the sampling sites with a regular bottle lid attached
4. At the sampling location, add 7g of yeast to the sugar solution and shake the bottle to mix the yeast.
5. Replace the regular bottle lid with the hose lid. Ensure that the hose within the bottle does not touch the liquid and there is plenty of space between the hose and the liquid to allow for the solution to foam.
6. Hang the bottle from the T bar using string.
7. Securely attach the plastic hose near the CDC light source using string, taking care to limit the blockage of the entrance to the CDC trap

How to collect

1. CO_2 bait bottles will be emptied upon collection. Whilst the contents are not harmful to the surrounding vegetation, take care not to empty the contents onto surrounding agricultural crops (if present).
2. Replace the bottle lids with regular bottle lids
3. Upon leaving the field, clean each bait bottle and tube lid thoroughly with water to remove undissolved sugar and yeast remnants.

Appendix 3. BG sentinel (BGS) trap protocol

What is a BGS trap?

The BGS trap uses a chemical lure and additional UV lights to attract a wide range of biting Diptera. The chemical lure mimics human scent and a fan distributes the scent into the air. Attracted Diptera and drawn into the trap via a funnel and collection pot using the suction created by the ventilator.

Where to install

The BGS trap should be positioned close to resting areas for Diptera, such as bushes, hedges, shrubs and other vegetation (Figure 1.). The trap may be placed directly under vegetation, but the top of the trap should be visible and not covered by leaves. Ensure that at least an area of 0.5 meters above the trap remains clear. The trap should not be placed within direct sunlight and ideally under vegetation, such as trees, to minimise the impact of rainfall.

Figure 1. Example positioning of the BGS trap.

How to install

1. Unhook the eyelet on the carrying handle and pop up the trap body
2. Place a plastic catch beaker within the beaker catch bag

3. Attach the funnel net and beaker catch bag over the protruding ring of the intake funnel (Figure 2.)
4. Carefully pour ethanol into the catch beaker, filling it three quarters of the way up.
5. Carefully insert the intake funnel into the opening on the main trap (Figure 3.)
6. Remove the plastic from the BG lure cartridge (or remove from plastic container on subsequent uses).
7. Remove the white disc cover and insert the lure into the trap (Figure 4.).
8. Place the UV light strip around the neck of the intake funnel.
9. Place the 12V battery off the ground, using a rock or other raised area. Ensure that the battery is sheltered from the rain using the available vegetation as a cover. If there is no suitable vegetation, a plastic bag can be placed over the battery.
10. Attach the 12V battery cable to the Y power cable (attached to the UV lights) and attach the remaining cable branch to the fan power cable (found on the side of the trap).

How to collect

7. Fix the shutter on the intake funnel by pressing together the intake funnel at the axis of the shutter and pressing the tab down (Figure 5.).
8. Remove the UV light strip from the neck of the intake funnel
9. Remove the intake funnel and catch bag by turning it counter-clockwise and pulling it out.
10. Examine the trap to determine if any live insects remain within the net portion of the trap. If insects are present, then use a wash bottle containing ethanol to 'shoot' insects with ethanol so that they fall into the collection beaker.
11. Remove the collection net from the intake funnel and carefully remove the collection beaker.
12. Place the preprepared sample label into the collection pot
13. Securely fasten the screw lid (note that the containers may not be completely watertight and so should be transported upright with limited shaking to ensure minimal spillage).
14. Replace the collection beaker with a fresh collection beaker containing ethanol and replace the 12V battery (unless on the sites final collection day).

How to dismantle (on final day at collection site only)

1. Following the collection of the samples, remove the BGS lure and place it in an airtight plastic container.
2. Work through the *how to install* instructions in reverse.

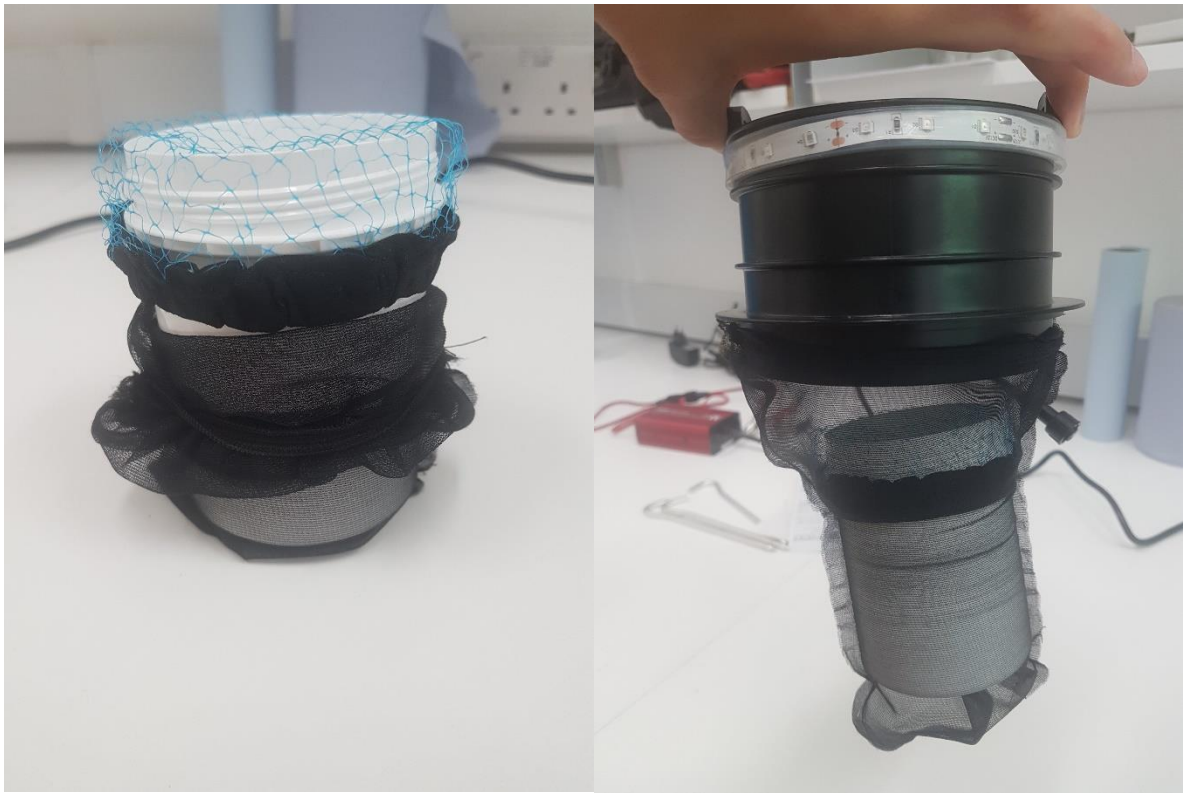


Figure 2. Beaker catch bag and funnel net placed over the intake funnel

Figure 3. Placing the intake funnel into the trap.

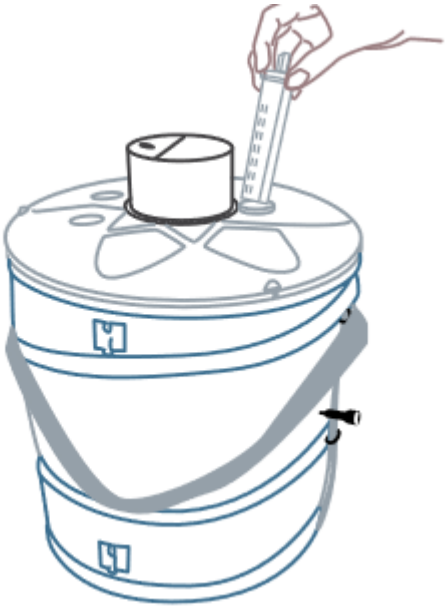


Figure 4. Inserting the lure into the trap.

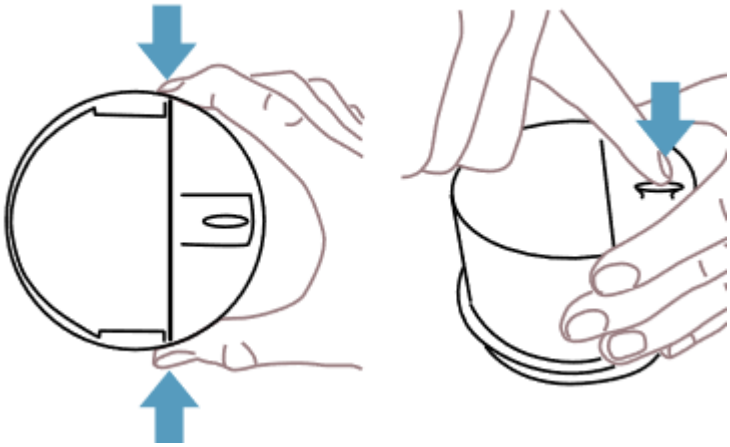


Figure 5. Closing the intake funnel prior to removal from the trap.

Appendix 4. Malaise trap protocol

What is a Malaise trap?

A malaise trap is made from fine netting, guy ropes, tent pegs, and a wooden or metal pole. Its purpose is to passively (rather than active trapping, see CDC and BGS traps) collect insects in a given area. It works by directing insects that land on the black netting, upwards towards the light material and trapping head located at the top of the trap.

Where to install

The Malaise trap should be erected, so it bisects a natural insect flight line, such as next to a forest opening, trail, bush, or dense patch of grass. The exact placement of the Malaise trap is based on your judgement at each site, so you should first examine the surrounding plant life for a likely flight path. The ground should be reasonably flat and clear and requires an area of around 2 meters by 1.5 meters.

How to install

The Malaise trap can typically set up in the order that you find easiest. Here is a basic outline of how to set up the trap.

1. Raise the main pole, with collection head, and secure it in the ground, ensure the trap is securely fastened to the main pole using the ribbon loops (Figure 1.)
2. Secure the corner poles into the ground using a rubber mallet.
3. Guy ropes from each corner pole and the main pole should be extended and secured by hammering the attached tent pegs into the ground. Rigidity in a malaise trap is critical so care should be taken to ensure that all guy lines are taught and secure (Figure 2.).
4. If a guy rope cannot be hammered into the ground, for example due to vegetation blocking it, it may be tied securely to surrounding vegetation (such as a tree)
5. Attach the collection bottle, filled three quarters with ethanol, by screwing it into the collection head (Figure 3.)

How to collect

1. To collect samples, simply unscrew the collection bottle from the collection head
2. Place the preprepared sample label into the tube
3. Securely fasten the screw lid
4. Replace the collection bottle with a fresh collection bottle containing ethanol (unless on the sites final collection day).

How to dismantle (on final day at collection site only)

1. Following the collection of the samples, work through the *how to install* instructions in reverse.
2. Take great care not to damage the Malaise trap netting, as holes in the net will reduce its efficiency.

Figure 1. Raising and securing the main pole

Figure 2. Fully erected trap with taught guy lines.

Figure 3. Collection head and attached collection bottle,

Appendix 5. Selecting habitat classifications

How to classify

To select the most appropriate habitat classification, stand at the centre (on top of the site GPS point) of the sampling location and examine the surrounding 30 metres of vegetation. Identify the vegetation types in your surroundings (grassland, forest, agriculture etc) and estimate, as a percentage, the total land coverage of each vegetation type within the 30m radius. Enter your estimations into the pre-printed site data sheet.

Habitat classification types

- Fallow – ploughed land with minimal above ground biomass,

Figure 1. Example of fallow land

- Scrubland – mixture of short grasses, shrubs and bushes. Limited number small isolated trees.

Figure 2. Example of scrub land

- Short grass Savanna (Sudan Grass Savanna) – mixture of short grassland and forest, characterised by an open canopy despite presence of trees. Absence/limited number of other flora.

Figure 3. Example of short grass Savanna

- Tall grass savanna (Guinea Grass Savanna)- mixture of grassland and forest, tall (potentially head height) grasses, few scattered deciduous trees (trees may predominantly be found on hillsides and along waterbodies)

Figure 4. Example of tall grass savanna

- Forest – dense forest, containing large density of diverse large trees, minimal grassland may be present. High degree of canopy cover.

Figure 5. Example of Forest habitat

- Forest edge- Edge of dense forest, containing large density of diverse large trees. High degree of canopy cover. Clear divide between forest edge and surrounding habitat.

Figure 6. Example of forest edge habitat

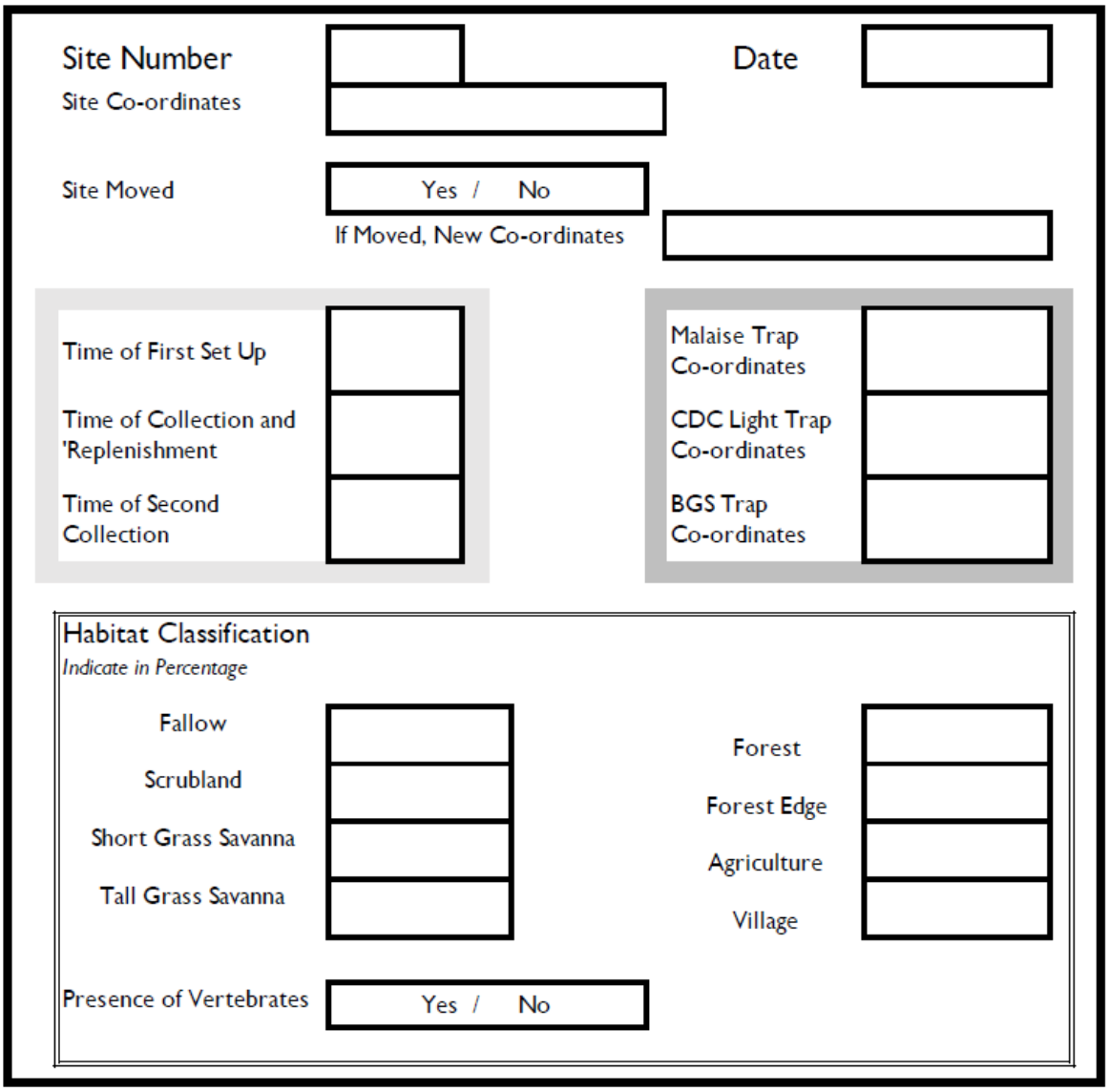
- Agriculture – Cropland (yam, cassava, maize, millet etc). Limited presence of other vegetation, open canopy.

Figure 7. Example of agricultural land (cassava)

- Village- Characterised by multiple houses and cleared land with minimal vegetation, restricted to trees and bushes around housing.

Figure 8. Example of a village interior

Appendix 6. Site data sheet



Appendix 7. Example of sample label

Site Number	1.1
Date	XX/XX/XX
Trap Type	BGS
Lot	0002

Appendix 8. List of lot numbers

Site Number	Trap Type	Lot	Y (latitude)	X (longitude)	Category
1.1	CDC	0001	6.43407	0.32665	Forest Edge (Inside)
1.1	BGS	0002	6.43407	0.32665	Forest Edge (Inside)
1.1	Malaise	0003	6.43407	0.32665	Forest Edge (Inside)
1.2	CDC	0004	6.43407	0.32665	Forest Edge (Inside)
1.2	BGS	0005	6.43407	0.32665	Forest Edge (Inside)
1.2	Malaise	0006	6.43407	0.32665	Forest Edge (Inside)
2.1	CDC	0007	6.42979	0.34015	Forest Interior
2.1	BGS	0008	6.42979	0.34015	Forest Interior
2.1	Malaise	0009	6.42979	0.34015	Forest Interior
2.2	CDC	0010	6.42979	0.34015	Forest Interior
2.2	BGS	0011	6.42979	0.34015	Forest Interior
2.2	Malaise	0012	6.42979	0.34015	Forest Interior
3.1	CDC	0013	6.44505	0.31766	Peri-Village
3.1	BGS	0014	6.44505	0.31766	Peri-Village
3.1	Malaise	0015	6.44505	0.31766	Peri-Village
3.2	CDC	0016	6.44505	0.31766	Peri-Village
3.2	BGS	0017	6.44505	0.31766	Peri-Village
3.2	Malaise	0018	6.44505	0.31766	Peri-Village
4.1	CDC	0019	6.44846	0.31761	Village
4.1	BGS	0020	6.44846	0.31761	Village
4.1	Malaise	0021	6.44846	0.31761	Village
4.2	CDC	0022	6.44846	0.31761	Village
4.2	BGS	0023	6.44846	0.31761	Village
4.2	Malaise	0024	6.44846	0.31761	Village
5.1	CDC	0025	6.44750	0.31726	Village
5.1	BGS	0026	6.44750	0.31726	Village
5.1	Malaise	0027	6.44750	0.31726	Village
5.2	CDC	0028	6.44750	0.31726	Village

5.2	BGS	0029	6.44750	0.31726	Village
5.2	Malaise	0030	6.44750	0.31726	Village
6.1	CDC	0031	6.43330	0.32546	Forest Edge (outside)
6.1	BGS	0032	6.43330	0.32546	Forest Edge (outside)
6.1	Malaise	0033	6.43330	0.32546	Forest Edge (outside)
6.2	CDC	0034	6.43330	0.32546	Forest Edge (outside)
6.2	BGS	0035	6.43330	0.32546	Forest Edge (outside)
6.2	Malaise	0036	6.43330	0.32546	Forest Edge (outside)
7.1	CDC	0037	6.42944	0.33986	Forest Interior
7.1	BGS	0038	6.42944	0.33986	Forest Interior
7.1	Malaise	0039	6.42944	0.33986	Forest Interior
7.2	CDC	0040	6.42944	0.33986	Forest Interior
7.2	BGS	0041	6.42944	0.33986	Forest Interior
7.2	Malaise	0042	6.42944	0.33986	Forest Interior
8.1	CDC	0043	6.40847	0.35646	Far Forest
8.1	BGS	0044	6.40847	0.35646	Far Forest
8.1	Malaise	0045	6.40847	0.35646	Far Forest
8.2	CDC	0046	6.40847	0.35646	Far Forest
8.2	BGS	0047	6.40847	0.35646	Far Forest
8.2	Malaise	0048	6.40847	0.35646	Far Forest
9.1	CDC	0049	6.42830	0.33888	Forest Interior
9.1	BGS	0050	6.42830	0.33888	Forest Interior
9.1	Malaise	0051	6.42830	0.33888	Forest Interior
9.2	CDC	0052	6.42830	0.33888	Forest Interior
9.2	BGS	0053	6.42830	0.33888	Forest Interior
9.2	Malaise	0054	6.42830	0.33888	Forest Interior
10.1	CDC	0055	6.40951	0.35736	Far Forest
10.1	BGS	0056	6.40951	0.35736	Far Forest
10.1	Malaise	0057	6.40951	0.35736	Far Forest
10.2	CDC	0058	6.40951	0.35736	Far Forest

10.2	BGS	0059	6.40951	0.35736	Far Forest
10.2	Malaise	0060	6.40951	0.35736	Far Forest
11.1	CDC	0061	6.44584	0.31871	Peri-Village
11.1	BGS	0062	6.44584	0.31871	Peri-Village
11.1	Malaise	0063	6.44584	0.31871	Peri-Village
11.2	CDC	0064	6.44584	0.31871	Peri-Village
11.2	BGS	0065	6.44584	0.31871	Peri-Village
11.2	Malaise	0066	6.44584	0.31871	Peri-Village
12.1	CDC	0067	6.43438	0.32702	Forest Edge (Inside)
12.1	BGS	0068	6.43438	0.32702	Forest Edge (Inside)
12.1	Malaise	0069	6.43438	0.32702	Forest Edge (Inside)
12.2	CDC	0070	6.43438	0.32702	Forest Edge (Inside)
12.2	BGS	0071	6.43438	0.32702	Forest Edge (Inside)
12.2	Malaise	0072	6.43438	0.32702	Forest Edge (Inside)
13.1	CDC	0073	6.44633	0.31713	Village
13.1	BGS	0074	6.44633	0.31713	Village
13.1	Malaise	0075	6.44633	0.31713	Village
13.2	CDC	0076	6.44633	0.31713	Village
13.2	BGS	0077	6.44633	0.31713	Village
13.2	Malaise	0078	6.44633	0.31713	Village
14.1	CDC	0079	6.44498	0.31720	Peri-Village
14.1	BGS	0080	6.44498	0.31720	Peri-Village
14.1	Malaise	0081	6.44498	0.31720	Peri-Village
14.2	CDC	0082	6.44498	0.31720	Peri-Village
14.2	BGS	0083	6.44498	0.31720	Peri-Village
14.2	Malaise	0084	6.44498	0.31720	Peri-Village
15.1	CDC	0085	6.42869	0.33920	Forest Interior
15.1	BGS	0086	6.42869	0.33920	Forest Interior
15.1	Malaise	0087	6.42869	0.33920	Forest Interior
15.2	CDC	0088	6.42869	0.33920	Forest Interior

15.2	BGS	0089	6.42869	0.33920	Forest Interior
15.2	Malaise	0090	6.42869	0.33920	Forest Interior
16.1	CDC	0091	6.40987	0.35766	Far Forest
16.1	BGS	0092	6.40987	0.35766	Far Forest
16.1	Malaise	0093	6.40987	0.35766	Far Forest
16.2	CDC	0094	6.40987	0.35766	Far Forest
16.2	BGS	0095	6.40987	0.35766	Far Forest
16.2	Malaise	0096	6.40987	0.35766	Far Forest
17.1	CDC	0097	6.44550	0.31843	Peri-Village
17.1	BGS	0098	6.44550	0.31843	Peri-Village
17.1	Malaise	0099	6.44550	0.31843	Peri-Village
17.2	CDC	0100	6.44550	0.31843	Peri-Village
17.2	BGS	0101	6.44550	0.31843	Peri-Village
17.2	Malaise	0102	6.44550	0.31843	Peri-Village
18.1	CDC	0103	6.44685	0.31713	Village
18.1	BGS	0104	6.44685	0.31713	Village
18.1	Malaise	0105	6.44685	0.31713	Village
18.2	CDC	0106	6.44685	0.31713	Village
18.2	BGS	0107	6.44685	0.31713	Village
18.2	Malaise	0108	6.44685	0.31713	Village
19.1	CDC	0109	6.43356	0.32587	Forest Edge (outside)
19.1	BGS	0110	6.43356	0.32587	Forest Edge (outside)
19.1	Malaise	0111	6.43356	0.32587	Forest Edge (outside)
19.2	CDC	0112	6.43356	0.32587	Forest Edge (outside)
19.2	BGS	0113	6.43356	0.32587	Forest Edge (outside)
19.2	Malaise	0114	6.43356	0.32587	Forest Edge (outside)
20.1	CDC	0115	6.40883	0.35678	Far Forest
20.1	BGS	0116	6.40883	0.35678	Far Forest
20.1	Malaise	0117	6.40883	0.35678	Far Forest
20.2	CDC	0118	6.40883	0.35678	Far Forest

20.2	BGS	0119	6.40883	0.35678	Far Forest
20.2	Malaise	0120	6.40883	0.35678	Far Forest

Appendix 3. Journal offprint: A Network Perspective on the Vectoring of Human Disease

Trends in Parasitology

CellPress

Opinion

A Network Perspective on the Vectoring of Human Disease

Ben Bellekom,[1,*] Talya D. Hackett,[1] and Owen T. Lewis[1]

Blood-sucking insects are important vectors of disease, with biting Diptera (flies) alone transmitting diseases that cause an estimated 700 000 human deaths a year. Insect vectors also bite nonhuman hosts, linking them into host-biting networks. While the major vectors of prominent diseases, such as malaria, yellow fever, dengue, and Zika, are intensively studied, there has been limited focus on the wider interactions of biting insects with nonhuman hosts. Drawing on network analysis and visualisation approaches from food-web ecology, we discuss the value of a network perspective for understanding host–insect–disease interactions, with a focus on Diptera vectors. Potential applications include highlighting pathways of disease transmission, highlighting reservoirs of infection, and identifying emerging and previously unrecognised vectors.

Highlights

Insect vectors of human diseases are embedded within complex ecological communities. Their interactions with other species, particularly alternative hosts, have important consequences for disease transmission and dynamics.

Modern molecular methods of blood-meal analysis provide the opportunity to document and quantify interactions between known or potential vectors and their host species accurately, and at an unprecedented scale.

Ecological interaction networks analysis and visualisation provides a useful tool to highlight the interconnectedness of these communities, the degree to which humans are embedded within them, and to identify and predict pathways of disease transmission.

The Emergence and Spread of Zoonotic Disease

More than 60% of known human infectious diseases and 75% of emerging infectious diseases are zoonotic, spreading from animals to humans [1,2]. Recent epidemics, including Zika fever, chikungunya and dengue fever, highlight how novel or apparently stable zoonoses can emerge, creating major challenges to human health and economic wellbeing [3]. Blood-feeding arthropods have been implicated in the spread of at least a quarter of all such outbreaks during the last century [4].

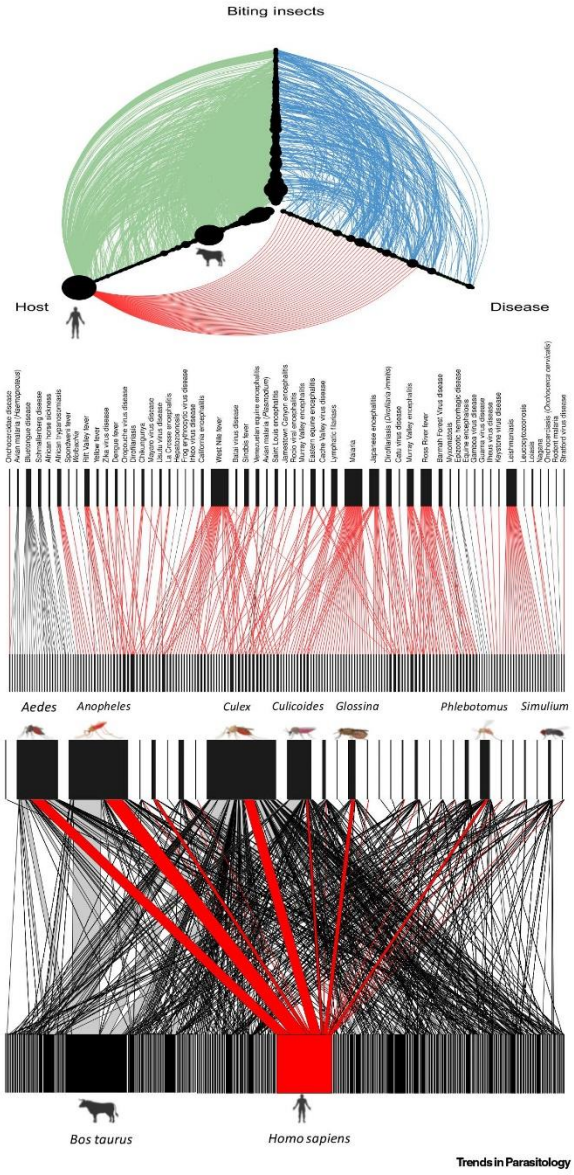
Over the coming century, increased human population densities, greater movement of both people and insects, and the expansion of agriculture and habitation into new areas are likely to favour the emergence, spread, and success of novel and existing insect-vectored zoonoses [5]. Here we argue that a multispecies analysis of vector–host interactions will help us to understand disease epidemiology, minimise risk, and inform control strategies.

Moving beyond Pairwise Interactions between Species

Most research on insect-vectored diseases has focused on the epidemic and **epizootic** (see Glossary) role of known vector species and their pairwise interactions with their hosts, particularly humans [3,6,7]. However, insect vectors are embedded within complex ecological communities [8] and their interactions with a wider set of species are likely to have important consequences for disease transmission. Ecological interaction network modelling [9] is a useful tool for understanding the interdependencies within these sets of interacting species, and for predicting the dynamics of the component species. Networks provide a visual and mathematical representation of interactions between **nodes** (e.g., species) connected by **edges** (e.g., feeding interactions) [10]. Networks may be unweighted, where edges are either present or absent, or weighted by the strength or frequency of interactions [11,12]. Both mutualistic interactions (such as interactions between plants and their pollinators) and antagonistic interactions (such as those between predators and their prey) can be represented in this way [13,14].

[1]Department of Zoology, 11a Mansfield Road, Oxford OX1 3SZ, UK

*Correspondence: ben.bellekom@zoo.ox.ac.uk (B. Bellekom).



Glossary

Bridge vector: a vector that acquires the causative agent of a disease from an infected reservoir host and transmits it to a human or secondary host.
Competent vector: an organism with the capacity to maintain and transmit a particular infectious agent, or with the ability to generate a new infection when interacting with a susceptible host.
Connectance: a network metric indicating the fraction of potential links that are realised, calculated as *L*/*HP*, where *L* is the number of links and *H* and *P* are the number of host and parasite nodes, respectively.
Connector species: species in a network that link groups or habitats together, typically through high mobility or varied interactions that cross structural, temporal, or behavioural boundaries.
Degree centrality: the number of links a species has in a network.
Disease spillover: transmission of a pathogen from its reservoir host into a susceptible population of an alternative host.
Edge: a relationship or connection between nodes within a network; synonymous with link.
Epizootic: a disease outbreak event within a nonhuman animal population, analogous to an epidemic within a human population.
Hub species: species with strong connections to many other species in a network.
Node: the components of a network. In the context of ecological interactions, species, populations, or individuals can be represented as nodes.
Reservoir host: an epidemiologically connected population in which a pathogen may be permanently maintained, and from which an infectious agent may be transmitted to a susceptible population.
Rewiring: the reassembling of interactions within a network, typically as occurs following a perturbation to the community.

(See figure legend at the bottom of the next page.)

Using biting Diptera, which currently cause approximately 700 000 deaths annually [15], as a case study, we focus on networks of antagonistic interactions between biting Diptera and the host animals from which they obtain blood meals, referred to here as 'host-biting networks'. Network ecology can reveal how community structure and function change in response to a range of perturbations, with widespread applications in conservation and agriculture [9]. We suggest that increased application of network approaches in disease vector research will provide valuable insights for understanding vector-borne diseases.

The Global Network of Biting Interactions

Host species are connected indirectly through their shared parasites [16]: multiple parasite species can share the same host species [17], and individual parasite species can interact with multiple host species [18]. To illustrate how hosts and parasites are interconnected on a global level, we constructed networks of biting Diptera–host interactions using data compiled from published literature (Figure 1). In total, these data comprise 26 049 biting Diptera–host interactions (where an interaction is defined as the blood of a host individual recorded in the blood meal of a biting insect) from 67 publications involving fieldwork in 32 countries. The resulting network has 263 biting Diptera species, 244 host species, and 1331 links between host species and biting Diptera species. Biting Diptera interact with host species from five classes: Amphibia, Insecta, Aves, Mammalia, and Reptilia. Cattle (*Bos taurus*) are the most commonly documented hosts, accounting for 6902 (26.6%) of all interactions. Host species are highly interconnected through shared vector species, with an unweighted network **connectance** of 0.02, meaning that 2% of all globally potential host–insect interactions have been observed. In reality, many potential links will not be possible because of geographic or temporal separation of hosts and biting Diptera, so the true connectance value will be higher. The average unweighted **degree centrality** (the number of links a species has in a network) for the 263 insect species is 5 (range: 1–86) host species. A high degree centrality indicates that an insect bites a large number of host species.

While the blood meal data summarised in Figure 1 do not document pathogens within blood meals, nor vector competence in transmitting them, we inferred disease vectoring potential from VectorBase, a database of insect vectors of different diseases [19]. Insects in our network are known to transmit 55 diseases, 39 of which are transmissible to humans (top and middle sections of Figure 1). Within our network, 190 insect species are known vectors of these 55 diseases, of which 154 can be transmitted to humans; most of these insect species (106) were indeed recorded as biting humans in the data compilation.

Figure 1. Host-biting Networks and Their Associated Disease Transmission Pathways. *Bottom*. Global bipartite [91] network of host-biting interactions, highlighting how humans (red node and edges) are embedded into the community. Node and edge widths are proportional to frequency. Nodes are resolved to species for hosts and genus for biting insects. *Middle*. Bipartite network linking disease (upper) to vectors (lower). Data on disease–vector interactions were extracted from Vectorbase [19]. Red links represent diseases that are transmissible to humans. Mosquito genera (*Aedes*, *Culex*, and *Ochlerotatus*) vector the largest number of human diseases (23, 18, and 15 respectively). *Top*. Hive plot [92] showing how hosts (left axis), biting insects (vertical axis) and their respective diseases (right axis), are globally interconnected. Interactions between hosts and biting Diptera are represented in green, vector–disease links in blue, and diseases that are transmissible to humans in red. Nodes along each axis scale with frequency. Taxon silhouettes are from BioRender.com. Data were compiled from published literature on blood meals of biting Diptera. To maximise the amount of relevant data, source publications were identified using ad hoc searches in Web of Science and Google Scholar with search terms including 'blood meal, 'mosquito ecology', 'metabarcoding', 'vector-host', and genus, family, and subfamily names for biting insect taxa. We also scrutinised the citations within these publications to identify additional relevant literature. Data were extracted from peer-reviewed articles that conducted molecular blood meal analysis for >1 biting Diptera species. To maximise our data set, we included data generated using a variety of blood meal analysis methods including diagnostic PCR approaches using species-specific or general vertebrate primers, enzyme-linked immunosorbent assay (ELISA), gel diffusion immunoassays, precipitin tests, microsatellite analyses, and monoclonal antibodies.

How Embedded Are Humans?

There are clear epidemiological benefits to characterising the degree to which humans are embedded within networks of host-biting interactions. Increasing contact of humans with livestock and wildlife populations is likely to facilitate a rise in the transmission of emerging diseases [20]. Examples of such vector-driven **disease spillover** events include Lyme disease and West Nile virus [21]. Unrestricted and repeated interactions between human hosts and **competent vectors** can allow diseases to reach epidemic levels rapidly [22].

Our compilation of blood-meal data reveals that humans are highly embedded in networks of blood-feeding and interact frequently with well-connected vectors for a variety of diseases (Figure 1). After cattle, humans were the second most frequently documented host species in the global network, accounting for 6217 (23.9%) of the interactions involving 152 biting Diptera species. Diptera recorded as biting humans had a significantly wider variety of nonhuman host species (mean: 6, range 0–85) than those that did not (mean: 3, range 1–23; Mann-Whitney U = 6448.5, $P<0.05$). The number of insect–human interactions reported is likely to be skewed by the anthropocentric focus of many of the studies from which data were extracted. However, even mosquito species typically considered anthropophilic (e.g., *Anopheles gambiae* and *Aedes aegypti*) [23–27] interact with a range of nonhuman hosts (nine and ten, respectively) including, birds, reptiles, and nonhuman mammals.

While biting Diptera can show specific feeding preferences [28,29], many species will feed opportunistically, with consequences for disease control measures [28,30]. Indeed, the five biting Diptera species with the most recorded interactions (*Anopheles arabiensis* n = 2280, 8.8%, *Culex annulirostris* n = 1986, 7.6%, *Cx. pipiens* n = 1698, 6.52, *Aedes camptorhynchus* n = 1159, 4.5%, and *Ae. aegypti* n = 1039, 4%) had a range of nonhuman hosts (4, 25, 85, 14 and 10, respectively) which comprised 84.5%, 95.5%, 86.5%, 99%, and 14.8% of their interactions, respectively.

Interpolation (rarefaction) and extrapolation (prediction) curves [31,32] indicate how comprehensively the host ranges of these individual species have been documented (Figure 2A). Observed host ranges were underestimated relative to those predicted through extrapolation by 4, 61, and 3 species for *Cx. annulirostris, Cx. pipiens*, and *Ae. aegypti,* respectively. This suggests that the known range of host use for even the most commonly sampled species is still not complete, and highlights the extent to which generalist species, such as *Cx. pipiens*, might interact with previously unsuspected hosts. No such differences between documented and predicted host species richness were found for *An. arabiensis* and *Ae. camptorhynchus*. The large number of undiscovered hosts for *Cx. pipiens* likely reflects its preference for feeding on birds [33], a particularly species-rich vertebrate class.

Rarefaction at the community level indicates that biting Diptera and host species are relatively well-resolved, with species accumulation curves in each case approaching an asymptote (Figure 2B). In contrast, documented interactions between vectors and hosts are much less complete. We therefore have an incomplete understanding of the full range of vector–host interactions, and further sampling will be necessary to reveal these cryptic interactions and to increase the overall resolution of host-biting networks.

The network shown in Figure 1 contains interactions from a global dataset. In reality, geographically separated hosts are unlikely to co-occur spatially to the extent that a single Diptera individual can interact with them. However, some hosts (e.g., migratory birds) and humans may be highly mobile. Therefore, it is feasible that hosts that never co-occur could be linked indirectly by a single mobile

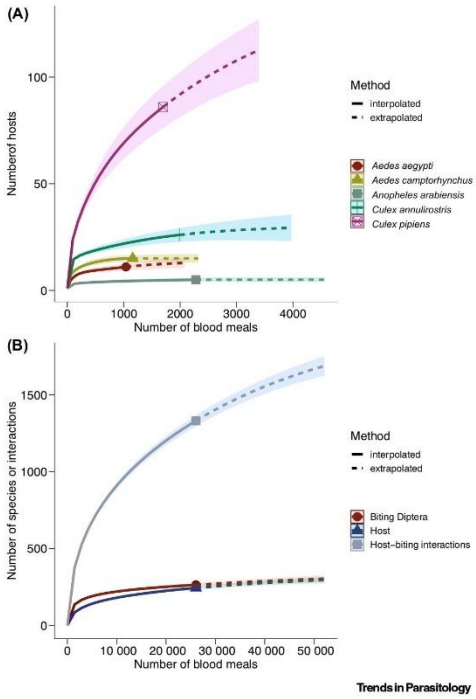


Figure 2. Smoothed Accumulation and Extrapolation Curves to Assess the Completeness of Sampling. (A) Number of host species recorded for the five most frequent biting Diptera species in the compiled dataset as a function of sampling effort (the number of blood meals screened). (B) Total numbers of hosts, biting insects, and interactions recorded in the whole dataset, as a function of sampling effort (the number of blood meals screened). These curves suggest that the level of sampling achieved across studies is sufficient to record most biting Diptera (red) and host (blue) species, but that many interactions among these species remain undocumented. Curves were calculated and plotted using the R package iNEXT [93]. Shaded areas show 95% confidence intervals for the predicted values.

human individual interacting with a variety of hosts and vectors in different locations. Inevitably, human movement, agricultural encroachment, illegal trafficking of hosts via the bushmeat and wildlife trade, and accidental introductions of invasive species result in increased opportunities for animal–human interactions and facilitate zoonotic disease transmission [34].

Using Networks to Identify Emerging Vectors and Disease

As human populations expand into new areas, there is increased opportunity for people to interact directly with novel potential disease vectors, and to interact indirectly, via shared biting insects, with other host species [35]. For example, deforestation and road development in the Peruvian Amazon has increased the frequency of human interaction with the malaria vector *Anopheles darlingi* 278 times relative to intact forested habitat, leading to an increase in malaria cases [36]. The integration of novel species into interaction networks is widely studied in community ecology and conservation biology, where it has broad implications for understanding the impact of invasive species. The influence of a new species depends on the degree to which it is a generalist or specialist and might also be conditioned on the diversity of available hosts [28], with generalists expected to have greater impacts on network structure and dynamics [37,38]. Quantitative network data, combined with information on biting behaviour and vector competence, will be useful to identify novel or encroaching vectors, and their likely disease transmission potential.

The emergence of novel diseases is hard to predict [39], but targeted vector surveillance at the interface of humans and wildlife, informed by data on species' biting preferences, has the potential to help predict and limit the emergence and spread of zoonotic disease. For example, following the rapid spread of *Culex* mosquito-vectored West Nile virus in the USA, intensive surveillance of mosquitoes and their hosts allowed outbreaks to be anticipated and effective vector control measures to be implemented [40]. Regular monitoring for pathogens of people, domestic animals [41], and wildlife [42] could provide an 'early warning system' for disease emergence. Such an approach could be complemented by mining data associated with molecular analysis of blood meals, focusing on likely vector species identified using weighted network data. Key species to identify in this way will include those that are highly connected and which interact frequently with disease-susceptible hosts (e.g., those with high degree centrality) [43]. In the longer term, data from such a surveillance programme could provide valuable general insights into spatial patterns of disease spillover, temporal variations in vector–host contact patterns, and the relative importance of local and long-range transmission events [44].

Interaction networks could also help to highlight unexpected or important hosts, as well as situations where vector populations have different host use in different geographic locations [45]. Hosts that are bitten by many biting Diptera species are candidates for the long-term persistence of disease and infection of sympatric **reservoir hosts** [46]. Identifying these species could therefore inform preventative measures, paralleling the situation where domestic dogs are vaccinated to reduce rabies transmission to humans [47]. Just as for novel vectors, the impact of novel hosts could be determined though network analysis. For example, the introduction of domesticated animal hosts into tick–vertebrate networks increased disease transmission throughout the community, as domesticated hosts acted as superspreaders, linking nodes that would not otherwise interact [48]. Better understanding of vector–host interactions could also help to prevent resurgence in infection from an unknown reservoir. Localised elimination of malaria in Guiana was followed by an increase in rice farming and subsequent displacement of cattle; as a result the malarial vector *Anopheles aquasalis* swapped from livestock hosts to the more abundant humans, leading to a localised outbreak of malaria [49]. Marked consequences for transmission are expected if vector species or disease-vulnerable host species occupying a central position within a network are added or removed [12,17]. Moreover, certain vectors may act as a **hub species**, with many strong connections to other species both in the same and separated habitats, or **connector species**, that couple spatially or temporally distinct communities, potentially facilitating between-habitat disease transmission [43,50–53]. Such effects could be predicted if well-resolved weighted network data were available: network ecologists increasingly use such data to inform models predicting the occurrence and consequences of **rewiring** events (where network interactions reconfigure as species are added or lost) [53,54].

Bridge vectors between humans and other animals could facilitate sustained transmission from a reservoir host to humans [55,56]. Biting insect species that routinely interact with multiple well-connected hosts, particularly where individuals contain blood from multiple susceptible host species, are candidate bridge vectors [57]. The incidence and identity of mixed blood meals, where molecular analysis indicates that a single insect contains blood meals from multiple host species, is highly relevant for predicting spillover events. However, mixed blood meals are rarely reported, either because the molecular methods used are unsuitable to detect them (Box 1), or because they are genuinely rare: only 10 (15%) of the papers included in our network compilation report mixed blood meals. When reported, the frequency of mixed blood meals varied, averaging 13.2% of total blood meals (range: 1–34.4%). Using molecular methods that detect mixed blood meals should be a priority (Box 2), and could ultimately allow network approaches to be applied on an individual basis, where nodes represent host or insect individuals, and edges represent

Box 1. Integrating Molecular Approaches to Determine Interactions and Pathways of Disease Transmission

Documenting potential and realised pathways of insect-vectored disease transmission has, until recently, relied on expensive and labour-intensive techniques that are error-prone and often require prior identification of candidate hosts and host-specific antisera [65–67]. Recent developments in molecular methods can overcome these limitations, generating data on feeding interactions more rapidly and comprehensively, and enabling the identification of multiple hosts from mixed blood meals [68].

Previously, species interactions were commonly identified from antibodies present in insect blood meals using enzyme-linked immunosorbent assays (ELISAs) [69,70] and monoclonal antibodies [71]. These methods have now largely been superseded by DNA barcoding, diagnostic PCR [69,72–74] and, most recently, metabarcoding. In metabarcoding, group-specific PCR primers are coupled with high-throughput sequencing to identify a broad range of taxa within bulk mixtures [72,75]. While the primary motivation behind the shift to metabarcoding has been to identify hosts of individual vectors, this approach also allows the construction of quantitative networks of host-biting interactions at a scale and resolution that was previously impossible. As the cost of metabarcoding continues to decrease, molecular analysis of blood meals offers excellent opportunities to identify and quantify previously cryptic species interactions. In future, shotgun sequencing holds considerable potential for blood meal analyses, with the advantage that it limits the negative effects of primer bias [62], although uptake of this approach is currently limited by high costs and a lack of genome data for many species.

feeding events that link them [58,59]. Analysis of such networks can reveal vector individuals that have interacted with multiple hosts, highlighting potential disease transmission events.

Finally, quantitative (weighted) networks can reveal potential emerging vectors which occur at low abundance, but which might increase rapidly following a disturbance or

Box 2. Improving the Resolution and Utility of Host-biting Networks

The network in Figure 1 represents an ad hoc compilation of published data. While it reveals how host-biting networks are interconnected and their potential applications in disease biology, a full realisation of the sorts of applications we suggest in the main text will require a more targeted approach to generate unbiased data on interactions at a local or regional scale.

Variation in trapping success for different insect taxa [76] will result in their uneven representation within networks. In particular, collection methods that use live hosts, such as human landing catches [77], will be heavily biased [78]. Widely used and efficient insect-sampling methods include USA Center for Disease Control (CDC) miniature light traps and BG-Sentinel traps with CO_2 bait, which capture a wide range of mosquitoes [79], sand flies [80], and biting midges [81,82]. We recommend using a combination of sampling methods to maximise representation of the biting insect fauna. Biases will inevitably remain, however; for example, light traps will undersample insects with diurnal and crepuscular activity. 'Nonattractive' methods, such as malaise traps or suction sampling, will be a valuable supplement to attractive traps, potentially documenting a wider range of biting insects and providing less-biased data on their relative abundances [83]. A common approach to construct weighted interaction is to combine randomised sampling (to document abundances in an unbiased way) with more targeted sampling of component species (to generate data on interaction frequencies efficiently) [84,85].

The spatial location of sampling will influence both the species composition of biting insects and the set of potential hosts with which they are interacting. For example, samples taken within, or close to, human habitation will over-represent both anthropophilic insects and human blood meals [27]. Humans are likely to be exposed to a much wider variety of biting insects (and will be connected indirectly to a much wider variety of their alternative hosts) when they spend time outside the domestic setting, for example, for farming or recreation [86,87]. Nonurban settings are also prime contexts for the emergence of novel and bridge vectors as well as novel pathogens and should therefore be a priority for network-focused sampling.

The choice of molecular analysis methods – especially the use of species-specific techniques and primer sets – can lead to incomplete or biased blood-meal identification [88]. Metabarcoding methods are preferable since they are likely to document a wider set of potential hosts, including many that may not have been recorded previously. Universal primer sets that are validated using *in silico* analysis, combined with high-throughput sequencing, can accurately and rigorously characterise the species composition of blood meals and generate unbiased interaction data [74].

Limited mixed blood meal reporting in the literature may reflect methods of blood meal analysis that are not appropriate for their detection and identification [89,90]. Wherever possible, metabarcoding methods should be used to screen samples to detect and report mixed blood meals, facilitating the identification of bridge vectors and informing disease transmission risks [74].

environmental change such as habitat degradation or climate change [22], or competitive release following reduction in the abundance of another species [60]. In our network, 48.7% (n = 129) of insect species had ten or fewer blood meal records. Of these species, 61 (47.3%) are known vectors of human diseases, and 32 (24.8%) were recorded interacting with humans. The sparse data for these biting insect species is likely a consequence of their relative rarity as well as sampling methods or locations that underestimate their abundance (Box 2). Nonetheless, they may represent unrecognised sources of cross-species disease transmission, now or in the future. Identification of host use and interaction frequency of understudied vectors is therefore of particular importance for predicting current and future disease transmission risk, and for informing control measures to minimise potential outbreaks [61].

Concluding Remarks

Available blood meal interaction data, compiled from a set of global studies, reveal the global pattern of insect–host community structure, and demonstrate the extent to which humans share potential vectors with other vertebrates. Many of the benefits of a network perspective that we have discussed will be best realised with dedicated biting insect–host community-level sampling on a local or regional scale. We encourage researchers to collect such data in a way that minimises biases to facilitate network-based analyses (Box 2). The resulting data will allow researchers to address a wide variety of unanswered questions relevant to understanding and predicting disease transmission and dynamics (see Outstanding Questions), both for diseases of humans and more generally.

In the long term, increasingly reliable sequencing techniques, such as shotgun metagenomics [62], will offer unique opportunities to gather large-scale, long-term network data sets. Ultimately, these could allow global-scale biomonitoring to detect changes in community structure and identify emerging disease threats [63,64].

Outstanding Questions

What factors predict host-biting interactions (e.g., phylogeny, traits, host abundance, abiotic factors), and can they be used to predict interactions for understudied vectors?

What influences the apparent host-specialisation of biting insects, and to what extent is this determined by encounter rates with hosts?

Can network data be used to predict changes in biting frequency (and disease transmission) when the densities of hosts that share potential vectors change?

How common are mixed blood meals, and is their frequency correlated with the degree centrality of individual biting insect species?

To what extent does the frequency that a pair of hosts is detected in mixed blood meals match the biting frequencies on those hosts inferred from interaction networks?

How does the structure of host-biting networks change seasonally, geographically, and across environmental gradients?

Are there biting insect species that consistently fill the role of hub and connector nodes across different geographical locations?

Acknowledgments

We thank members of the Community Ecology Research Group Oxford for helpful discussions, and John Connolly, John Mumford, Silke Fuchs, and three anonymous reviewers for their comments and suggestions. This work was supported by a Natural Environment Research Council studentship to B.B. though the NERC-Oxford University Environmental Research Doctoral Training Partnership (NE/L002612/1). O.L. and T.H. were supported by Target Malaria. Target Malaria is supported by a grant from the Bill and Melinda Gates Foundation and the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation.

References

1. White, R.J. and Razgour, O. (2020) Emerging zoonotic diseases originating in mammals: a systematic review of effects of anthropogenic land-use change. *Mammal Rev.* Published online June 2, 2020. https://doi.org/10.1111/mam.12201
2. Jones, K.E. *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451, 990–993
3. Braack, L. *et al.* (2018) Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasit. Vectors* 11, 29
4. Alcaide, M. *et al.* (2009) Disentangling vector-borne transmission networks: a universal DNA barcoding method to identify vertebrate hosts from arthropod bloodmeals. *PLoS One* 4, e7092
5. Gratz, N.G. (1999) Emerging and resurging vector-borne diseases. *Annu. Rev. Entomol.* 44, 51–75
6. Molaei, G. *et al.* (2013) Vector–host interactions and epizootiology of eastern equine encephalitis virus in massachusetts. *Vector-Borne Zoonotic Dis.* 13, 312–323
7. Kweka, E.J. *et al.* (2019) Ecology of *Aedes* mosquitoes, the major vectors of arboviruses in human population. In *Dengue Fever – A Resilient Threat in the Face of Innovation*, pp. 39–56, IntechOpen
8. Ferguson, H.M. *et al.* (2010) Ecology: a prerequisite for malaria elimination and eradication. *PLoS Med.* 7, e1000303
9. Kaiser-Bunbury, C.N. and Blüthgen, N. (2015) Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants* 7, plv076
10. Proulx, S. *et al.* (2005) Network thinking in ecology and evolution. *Trends Ecol. Evol.* 20, 345–353
11. Blüthgen, N. (2010) Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic Appl. Ecol.* 11, 185–195
12. Poulin, R. (2010) Network analysis shining light on parasite ecology and diversity. *Trends Parasitol.* 26, 492–498
13. Ings, T.C. *et al.* (2009) Review: Ecological networks – beyond food webs. *J. Anim. Ecol.* 78, 253–269
14. Bendel, C.R. *et al.* (2019) Plant–pollinator networks in grassland working landscapes reveal seasonal shifts in network structure and composition. *Ecosphere* 10, e02569
15. Müller, R. *et al.* (2019) Vector-borne diseases. In *Biodiversity and Health in the Face of Climate Change* (Marselle, M.R. *et al.*, eds), pp. 67–90, Springer International Publishing

16. Espinaze, M.P.A. *et al.* (2018) Domestic mammals facilitate tick-borne pathogen transmission networks in South African wildlife. *Biol. Conserv.* 221, 228–236
17. Godfrey, S.S. *et al.* (2009) Network structure and parasite transmission in a group living lizard, the gidgee skink, *Egernia stokesii*. *Behav. Ecol. Sociobiol.* 63, 1045–1056
18. Pilosof, S. *et al.* (2015) Potential parasite transmission in multi-host networks based on parasite sharing. *PLoS One* 10, e0117909
19. Giraldo-Calderón, G.I. *et al.* (2015) VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res.* 43, D707–D713
20. Grubaugh, N.D. *et al.* (2015) Xenosurveillance: a novel mosquito-based approach for examining the human-pathogen landscape. *PLoS Negl. Trop. Dis.* 9, e0003628
21. Kulkarni, M.A. *et al.* (2015) Major emerging vector-borne zoonotic diseases of public health importance in Canada. *Emerg. Microbes Infect.* 4, e33
22. Weaver, S.C. and Reisen, W.K. (2010) Present and future arboviral threats. *Antivir. Res.* 85, 328–345
23. Orsborne, J. *et al.* (2018) Using the human blood index to investigate host biting plasticity: a systematic review and meta-regression of the three major African malaria vectors. *Malar. J.* 17, 479
24. Pappa, V. *et al.* (2011) Estimation of the human blood index in malaria mosquito vectors in Equatorial Guinea after indoor antivector interventions. *Am. J. Trop. Med. Hyg.* 84, 298–301
25. Busula, A.O. *et al.* (2015) Mosquito host preference affect their reponse to synthetic and natural odour blends. *Malar. J.* 14, 133
26. Carvalho, G.M.L. *et al.* (2017) Bloodmeal identification in field-collected sand flies from Casa Branca, Brazil, using the cytochrome *b* PCR method. *J. Med. Entomol.* 54, 1049–1054
27. Orsborne, J. *et al.* (2020) Evidence of extrinsic factors dominating intrinsic blood host preferences of major African malaria vectors. *Sci. Rep.* 10, 741
28. Chaves, L.F. *et al.* (2010) Blood feeding patterns of mosquitoes: Random or structured? *Front. Zool.* 7, 1–11
29. Harrington, L.C. *et al.* (2009) Why do female *Aedes aegypti* (Diptera: Culicidae) feed preferentially and frequently on human blood? *J. Med. Entomol.* 38, 411–422
30. Lassen, S.B. *et al.* (2011) Molecular identification of bloodmeals from biting midges (Diptera: Ceratopogonidae: *Culicoides* Latreille) in Denmark. *Parasitol. Res.* 108, 823–829
31. Vázquez, D.P. *et al.* (2009) Uniting pattern and process in plant–animal mutualistic networks: a review. *Ann. Bot.* 103, 1445–1457
32. Budka, A. *et al.* (2019) The use of rarefaction and extrapolation as methods of estimating the effects of river eutrophication on macrophyte diversity. *Biodivers. Conserv.* 28, 385–400
33. Kothera, L. *et al.* (2020) Bloodmeal, host selection, and genetic admixture analyses of *Culex pipiens* Complex (Diptera: Culicidae) mosquitoes in Chicago, IL. *J. Med. Entomol.* 57, 78–87
34. Johnson, C.K. *et al.* (2020) Global shifts in mammalian population trends reveal key predictors of virus spillover risk. *Proc. R. Soc. B Biol. Sci.* 287, 20192736
35. Santiago-Alarcon, D. *et al.* (2012) Bloodmeal analysis reveals avian *Plasmodium* infections and broad host preferences of *Culicoides* (Diptera: Ceratopogonidae) vectors. *PLoS One* 7, 31098
36. Vittor, A.Y. *et al.* (2006) The effect of deforestation on the human-biting rate of *Anopheles darlingi*, the primary vector of falciparum malaria in the Peruvian Amazon. *Am. J. Trop. Med. Hyg.* 74, 3–11
37. Russo, L. *et al.* (2014) Patterns of introduced species interactions affect multiple aspects of network structure in plant–pollinator communities. *Ecology* 95, 2953–2963
38. Henneman, M.L. and Memmott, J. (2001) Infiltration of a hawaiian community by introduced biological control agents. *Science* 293, 1314–1316
39. Murphy, F.A. (1998) Emerging zoonoses. *Emerg. Infect. Dis.* 4, 429–435
40. Healy, J.M. *et al.* (2015) Comparison of the efficiency and cost of West Nile virus surveillance methods in California. *Vector-Borne Zoonotic Dis.* 15, 147–155
41. Wolfe, N.D. *et al.* (2007) Origins of major human infectious diseases. *Nature* 447, 279–283
42. Morse, S.S. *et al.* (2012) Prediction and prevention of the next pandemic zoonosis. *Lancet* 380, 1956–1965
43. Graham, S.P. *et al.* (2009) Nestedness of ectoparasite–vertebrate host networks. *PLoS One* 4, 7873
44. Craft, M. and Caillaud, D. (2011) Network models: an underutilized tool in wildlife epidemiology? *Interdiscip. Perspect. Infect. Dis.* 2011, 676949
45. Tirados, I. *et al.* (2006) Blood-feeding behaviour of the malarial mosquito *Anopheles arabiensis*: Implications for vector control. *Med. Vet. Entomol.* 20, 425–437
46. Streicker, D.G. *et al.* (2013) Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecol. Lett.* 16, 975–984
47. Lembo, T. *et al.* (2010) The feasibility of canine rabies elimination in africa: dispelling doubts with data. *PLoS Negl. Trop. Dis.* 4, e626
48. Estrada-Peña, A. *et al.* (2015) Interactions between tick and transmitted pathogens evolved to minimise competition through nested and coherent networks. *Sci. Rep.* 5, 10361
49. Giglioli, G. (1963) Ecological change as a factor in renewed malaria transmission in an eradicated area. A localized outbreak of *A. aquasalis*-transmitted malaria on the Demerara River estuary, British Guiana, in the fifteenth year of *A. darlingi* and malaria eradication. *Bull. World Health Organ.* 29, 131–145
50. Olesen, J.M. *et al.* (2007) The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19891–19896
51. Guimerà, R. and Amaral, L.A.N. (2005) Cartography of complex networks: Modules and universal roles. *J. Stat. Mech. Theory Exp.* 2005, 1–13
52. Hackett, T.D. *et al.* (2019) Reshaping our understanding of species' roles in landscape-scale networks. *Ecol. Lett.* 22, 1367–1377
53. Bartley, T.J. *et al.* (2019) Food web rewiring in a changing world. *Nat. Ecol. Evol.* 3, 345–354
54. CaraDonna, P.J. *et al.* (2017) Interaction rewiring and the rapid turnover of plant-pollinator networks. *Ecol. Lett.* 20, 385–394
55. Shahhosseini, N. *et al.* (2018) Host-feeding patterns of *Culex* mosquitoes in Iran. *Parasit. Vectors* 11, 669
56. Mayi, M.P.A. *et al.* (2020) Habitat and seasonality affect mosquito community composition in the West Region of Cameroon. *Insects* 11, 312
57. Kang, D.S. *et al.* (2019) Identification of host blood meals of mosquitoes (Diptera: Culicidae) collected at the Aripo Savannas Scientific Reserve in Trinidad, West Indies. *J. Med. Entomol.* 56, 1734–1738
58. White, L.A. *et al.* (2017) Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biol. Rev.* 92, 389–409
59. Godfrey, S.S. (2013) Networks and the ecology of parasite transmission: A framework for wildlife parasitology. *Int. J. Parasitol. Parasit. Wildl.* 2, 235–245
60. 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. *Bull. Entomol. Res.* 51, 243
61. Faria, N.R. *et al.* (2017) Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* 546, 406–410
62. Boers, S.A. *et al.* (2019) Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 1059–1070
63. Bohan, D.A. *et al.* (2017) Next-generation global biomonitoring: large-scale, automated reconstruction of ecological networks. *Trends Ecol. Evol.* 32, 477–487
64. Olivier, P. *et al.* (2019) Exploring the temporal variability of a food web using long-term biomonitoring data. *Ecography (Cop.)* 42, 2107–2121
65. Evans, D.M. *et al.* (2016) Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Funct. Ecol.* 30, 1904–1916
66. Miyake, T. *et al.* (2019) Bloodmeal host identification with inferences to feeding habits of a fish-fed mosquito, *Aedes baisasi*. *Sci. Rep.* 9, 4002
67. Townzen, J.S. *et al.* (2008) Identification of mosquito bloodmeals using mitochondrial cytochrome oxidase subunit I and cytochrome *b* gene sequences. *Med. Vet. Entomol.* 22, 386–393
68. Clare, E.L. *et al.* (2018) Approaches to integrating genetic data into ecological networks. *Mol. Ecol.* 28, 503–519
69. Jansen, C.C. *et al.* (2009) Blood sources of mosquitoes collected from urban and peri-urban environments in eastern

Australia with species-specific molecular analysis of avian blood meals. *Am. J. Trop. Med. Hyg.* 81, 849–857
70. Apperson, C.S. *et al.* (2002) Host-feeding habits of *Culex* and other mosquitoes (Diptera: Culicidae) in the Borough of Queens in New York City, with characters and techniques for identification of *Culex* mosquitoes. *J. Med. Entomol.* 39, 777–785
71. Svobodová, M. *et al.* (2003) Short report: Distribution and feeding preference of the sand flies *Phlebotomus sergenti* and *P. papatasi* in a cutaneous leishmaniasis focus in Sanliurfa, Turkey. *Am. J. Trop. Med. Hyg.* 68, 6–9
72. Pompanon, F. *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21, 1931–1950
73. Omondi, D. *et al.* (2015) Unraveling host–vector–arbovirus interactions by two-gene high resolution melting mosquito bloodmeal analysis in a Kenyan wildlife-livestock interface. *PLoS One* 10, e0134375
74. Logue, K. *et al.* (2016) Unbiased characterization of *Anopheles* mosquito blood meals by targeted high-throughput sequencing. *PLoS Negl. Trop. Dis.* 10, e0004512
75. Zinger, L. *et al.* (2019) DNA metabarcoding – need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862
76. Hesson, J.C. *et al.* (2015) Trapping biases of *Culex torrentium* and *Culex pipiens* revealed by comparison of captures in CDC traps, ovitraps, and gravid traps. *J. Vector Ecol.* 40, 158–163
77. Malti, D.V. *et al.* (2015) Development and evaluation of mosquito-electrocuting traps as alternatives to the human landing catch technique for sampling host-seeking malaria vectors. *Malar. J.* 14, 502
78. Ponlawat, A. *et al.* (2017) Field evaluation of two commercial mosquito traps baited with different attractants and colored lights for malaria vector surveillance in Thailand. *Parasit. Vectors* 10, 378
79. Gorsich, E.E. *et al.* (2019) A comparative assessment of adult mosquito trapping methods to estimate spatial patterns of abundance and community composition in southern Africa. *Parasit. Vectors* 12, 462
80. Aklilu, E. *et al.* (2017) Comparative study on the nocturnal activity of phlebotomine sand flies in a highland and lowland foci of visceral leishmaniasis in north-western Ethiopia with special reference to *Phlebotomus orientalis*. *Parasit. Vectors* 10, 393
81. Hopken, M.W. *et al.* (2017) Picky eaters are rare: DNA-based blood meal analysis of *Culicoides* (Diptera: Ceratopogonidae) species from the United States. *Parasit. Vectors* 10, 169
82. Hoel, D.F. *et al.* (2010) Efficacy of commercial mosquito traps in capturing phlebotomine sand flies (Diptera: Psychodidae) in Egypt. *J. Med. Entomol.* 47, 1179–1184
83. Vazquez, D.P. *et al.* (2005) Species abundance and the distribution of specialization in host–parasite interaction networks. *J. Anim. Ecol.* 74, 946–955
84. Downey, H. *et al.* (2020) Assessing the potential for indirect interactions between tropical tree species via shared insect seed predators. *Biotropica* 52, 509–520
85. Lewis, O.T. *et al.* (2002) Structure of a diverse tropical forest insect–parasitoid community. *J. Anim. Ecol.* 71, 855–873
86. Ferraguti, M. *et al.* (2016) Effects of landscape anthropization on mosquito community composition and abundance. *Sci. Rep.* 6, 29002
87. Thongsripong, P. *et al.* (2013) Mosquito vector diversity across habitats in central Thailand endemic for dengue and other arthropod-borne diseases. *PLoS Negl. Trop. Dis.* 7, e2507
88. Hannon, E.R. *et al.* (2019) Bloodmeal host selection of *Culex quinquefasciatus* (Diptera: Culicidae) in Las Vegas, Nevada, United States. *J. Med. Entomol.* 56, 603–608
89. Brugman, V.A. *et al.* (2017) Blood-feeding patterns of native mosquitoes and insights into their potential role as pathogen vectors in the Thames estuary region of the United Kingdom. *Parasit. Vectors* 10, 163
90. Kent, R.J. (2009) Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. *Mol. Ecol. Resour.* 9, 4–18
91. Dormann, C.F. *et al.* (2008) Introducing the bipartite package: analysing ecological networks. *R News* 8, 8–11
92. Krzywinski, M. *et al.* (2012) Hive plots-rational approach to visualizing networks. *Brief. Bioinform.* 13, 627–644
93. Hsieh, T.C. *et al.* (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* 7, 1451–1456

Appendix 4. Journal offprint: Latitudinal and anthropogenic effects on the structuring of networks linking blood-feeding flies and their vertebrate hosts

Received: 12 October 2022 | Accepted: 5 May 2023
DOI: 10.1111/mve.12671

ORIGINAL ARTICLE

Medical and Veterinary Entomology

Latitudinal and anthropogenic effects on the structuring of networks linking blood-feeding flies and their vertebrate hosts

Ben Bellekom | **Owen T. Lewis** | **Talya D. Hackett**

Department of Biology, University of Oxford, Oxford, UK

Correspondence
Ben Bellekom, Department of Biology, University of Oxford, 11a Mansfield Road, Oxford, OX1 3SZ, UK.
Email: ben.bellekom@biology.ox.ac.uk

Funding information
Bill and Melinda Gates Foundation; NERC-Oxford University Environmental Research Doctoral Training Partnership, Grant/Award Number: NE/L002612/1; Open Philanthropy Project

Abstract

Biting flies (Diptera) transmit pathogens that cause many important diseases in humans as well as domestic and wild animals. The networks of feeding interactions linking these insects to their hosts, and how they vary geographically and in response to human land-use, are currently poorly documented but are relevant to understanding cross-species disease transmission. We compiled a database of biting Diptera–host interactions from the literature to investigate how key interaction network metrics vary latitudinally and with human land-use. Interaction evenness and H2' (a measure of the degree of network specificity) did not vary significantly with latitude. Compared to near-natural habitats, interaction evenness was significantly lower in agricultural habitats, where networks were dominated by relatively few species pairs, but there was no evidence that the presence of humans and their domesticated animals within networks led to systematic shifts in network structure. We discuss the epidemiological relevance of these results and the implications for predicting and mitigating future spill-over events.

KEYWORDS
bipartite, biting Diptera, blood meal, ecological interactions, habitat modification

INTRODUCTION

Across taxa, species richness consistently decreases from equatorial to polar latitudes (Lawrence & Fraser, 2020). This macroecological pattern is linked to changes in the structure and organisation of ecological communities. For example, increased species richness and associated niche partitioning at low latitudes should increase both the frequency and specificity of interspecific interactions such as mutualism, competition and parasitism (Morris et al., 2014; Schemske et al., 2009; Willig et al., 2003). Superimposed on these large-scale patterns, anthropogenic habitat modification for agriculture and other land-uses also has marked effects on diversity, community composition and species interactions (Meyer Steiger et al., 2016), and can reshape interactions within a community, even in the absence of biodiversity loss (Morris, 2010; Tylianakis et al., 2007).

Analyses of quantitative ecological networks documenting the interactions among species, weighted by the interaction frequency (Schleuning et al., 2012; Xing & Fayle, 2021), provide an approach for understanding how communities are structured across space and time. A growing number of network studies involving diverse taxa, locations and contexts provide opportunities for synthetic analyses investigating large-scale ecological patterns (Xing & Fayle, 2021), such as trends in specialisation with latitude (Morris et al., 2014; Schleuning et al., 2012) and in response to land-use intensification (Tylianakis et al., 2007; Weiner et al., 2014). Few studies, however, have explored the relative influence of large-scale macroecological gradients and more localised anthropogenic impacts on the structural properties of networks of interacting species (Pellissier et al., 2018; Tylianakis et al., 2007).

Here, we investigate how macro-ecological and anthropogenic factors influence the properties of networks linking biting flies (Diptera) and their vertebrate hosts. Feeding interactions between Diptera and hosts can now be routinely established using molecular analysis of insect blood meals. These interactions are of particular

interest, since biting Diptera are often vectors of a wide range of pathogens including malarial parasites (Sinka et al., 2012), Bluetongue virus (Baker et al., 2021), West Nile virus (Kilpatrick et al., 2005) and Leishmania (Killick-Kendrick, 1999). These may infect humans, domestic and non-domestic hosts, and cause significant damage to public and health economics (Barber et al., 2010; Rushton & Lyons, 2015; Sachs & Malaney, 2002). Moreover, their transmission potential can co-vary with their interactions across land-use gradients (Meyer Steiger et al., 2016; Müller et al., 2019; Runghen et al., 2021). We analyse biting fly–vertebrate interaction data from a wide range of latitudes and across different habitat types to explore the relative importance of latitude and land-use in structuring interaction networks.

METHODS

Data compilation

Biting Diptera–host interaction data were extracted from the literature on insect blood meals, using a subset of the data compiled by Bellekom et al. (2021). To limit bias, we restricted analyses to data generated using PCR and DNA sequencing (Logue et al., 2016), and excluded studies that used sampling methods and locations inappropriate for collecting a variety of biting Diptera species and subsequent host blood meals. For example, those that used live host-baited trapping methods, such as livestock and cattle-baited tents, were excluded as blood meals would be heavily biased towards the bait. Studies that provided data for a single biting Diptera species were also excluded, as these do not provide network data. We also excluded studies lacking site location details. For each remaining study, we recorded site location and classified habitats into three broad categories of anthropogenic landscape modification. Sites where cultivated land or livestock were the dominant land-use were categorised as Agricultural; those that referenced natural vegetation with limited human presence were classified as Near-natural; and those where sampling took place within, or around human habitation were classified as Village/Urban. Where sampling was carried out in more than one habitat type, separate networks were generated for each reported habitat. Where habitat could not be determined reliably from the published information (four cases), we used satellite imagery (QGIS in combination with Google Earth) to infer the habitat, using the same categories used for studies where published information on habitat type was available. For example, where satellite imagery indicated that the location of the study was associated with a human settlement, the data were assigned to the Village/Urban category.

Biting Diptera and hosts were resolved to species level, where possible. Where a single species-level identification was missing for a biting Diptera or host, nodes were simply labelled with the relevant genus or family (e.g., *Culicoides* spp.). Where a genus contained multiple unknown species-level identifications, we checked whether sympatric congeneric or confamilial species were likely to occur at the focal location using field guides and online resources (GBIF.org, 2021). As before, where no sympatric congeneric or confamilial species occurred, nodes were resolved to the lowest level possible (e.g., *Anas* spp.). Interactions where either the host or the biting insect could not be resolved to genus or family level (or where more than one species could occur in a single node) were removed to prevent different species being combined into the same genus-level node. In total, 18 biting Diptera and 28 host records were resolved to genus or family level and 76 hosts, 14 biting insects and 119 interactions were removed.

Data analysis

For each included study, we analysed species interaction data as weighted antagonistic bipartite networks and calculated two network metrics, interaction evenness (IE) and network specialisation (H2'), chosen for their ecological and epidemiological relevance. IE is a weighted network metric based on Shannon diversity that describes the homogeneity of interaction frequencies across all links in the network ($E_2 = H_2/\ln L$, where H is Shannon diversity, and L is the number of all links) (Blüthgen et al., 2008; Kaiser-Bunbury & Blüthgen, 2015). Low IE values indicate contexts in which a small number of species and their links dominate the community (Kaiser-Bunbury & Blüthgen, 2015). This is particularly relevant if the dominant species are vectors or susceptible hosts. H2' is a weighted network metric that quantifies the deviation of observed interaction frequencies from those expected if interaction frequencies were random (Blüthgen et al., 2006). A high degree of generalism within our network would be expected to facilitate transmission between phylogenetically dissimilar hosts (Abella-Medrano et al., 2018). IE and H2' values were calculated using the *networklevel* function in the R package *bipartite* (Dormann et al., 2008).

Variation in host and biting Diptera richness across latitudes was explored using a linear model, including the total number of blood meals analysed per network (a proxy for sampling effort) as a covariate. We explored the extent to which variations in IE and H2' could be explained by habitat type and latitude using a generalised linear model (GLM) with a Gaussian error distribution. Absolute values for latitude (i.e., removing negative signs) were used to combine data from the northern and southern hemisphere. To control for the confounding effects of network size and species richness, we included species richness (S = number of resource species + number of consumer species), and log-transformed matrix size (sum of all interactions within the matrix) in the GLM (Galiana et al., 2019). Most of our networks (46 out of 47) comprised Diptera from multiple genera but drawn from a single family; each network could therefore be classified by its dominant family. In the case where multiple families were represented in a network, we classified it by the dominant family (the family with the highest number of interactions). Our models tested for the effects of family as well as the two-way interaction terms between

family and latitude and between family and habitat. This allowed us to examine the influence of Diptera family on IE and H2' values and how these changed across latitude and between habitat types. Residuals were visually inspected to check model assumptions. The statistical significance of habitat type and latitude was assessed by comparing simpler models to more complex models for variation in deviance based

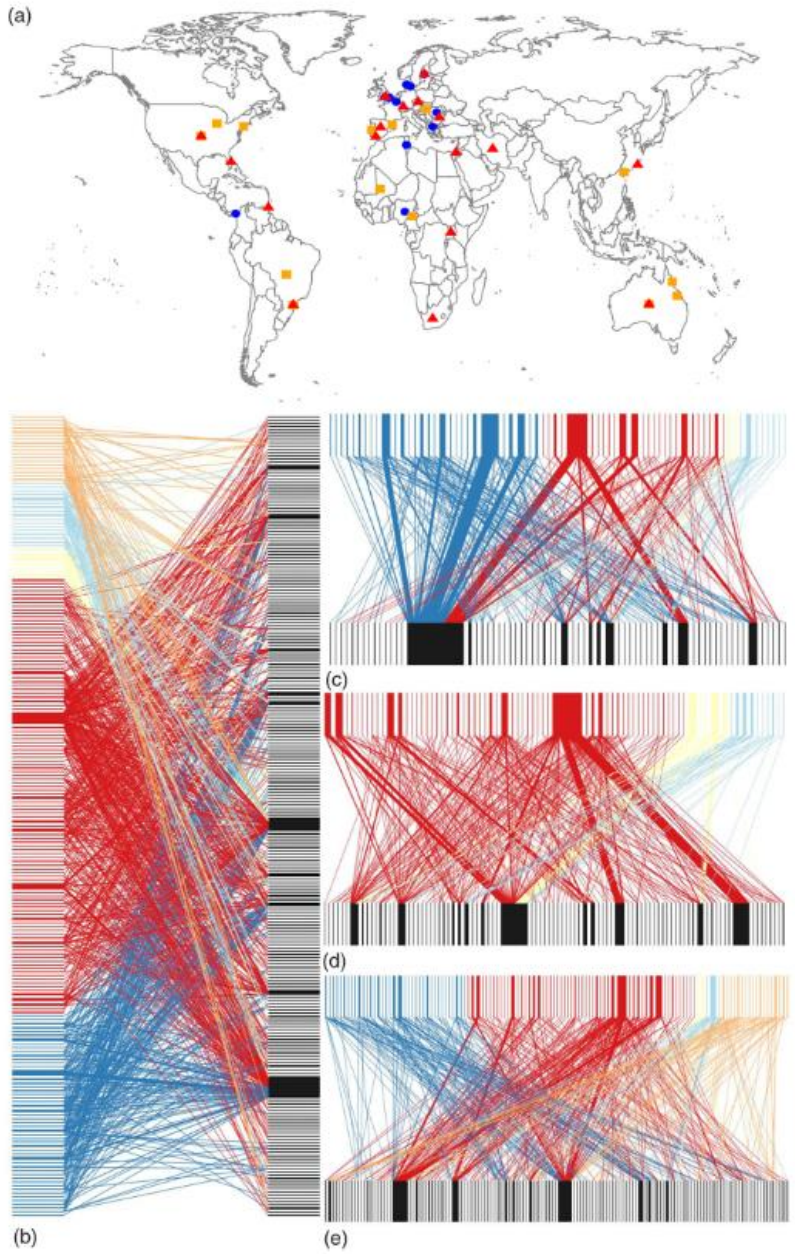


FIGURE 1 (a) Global distribution of the studies included in the analysis. Agricultural sites are represented by blue circles, Near-natural sites by red triangles, and Village/Urban sites by yellow squares. Co-located symbols are studies that sampled in multiple habitat classifications and were treated as separate networks in our analyses (Map data: Google Maps 2020). (b) The aggregated global network containing all host (right) and Diptera (left) interactions, and Diptera (top)–host (bottom) interactions separated by habitat classification: Agricultural(c), Village/Urban (d) and Near-natural (e). Node and edge widths are proportional to frequency of occurrence and are coloured by biting Diptera family (Ceratopogonidae = dark blue, Culicidae = red, Glossinidae = yellow, Psychodidae = light blue and Simuliidae = orange).

on a chi-square distribution (Mayi et al., 2020). Post-hoc analysis was conducted to identify intra-factor significant differences, using Tukey's HSD (honestly significant difference) tests.

We used a null model to evaluate whether domesticated hosts had a measurable effect on network interaction evenness. Within each network, we simulated targeted removal of humans and domestic animals (specifically chickens, dogs, cats, goats, cattle, horses, pigs, and sheep) and compared this to removal of an equal number of randomly selected host species, both domestic and non-domestic, replicated 100 times. We then calculated the z-scores and compared the observed network metric to the distribution of the simulated values.

The dataset used is not an exhaustive set of biting Diptera–host interactions. Therefore to assess sampling completeness for each habitat, we drew species interpolation and extrapolation curves for hosts, biting Diptera and interactions as a function of sampling effort (the number of blood meals analysed) using the iNEXT package ($q = 0$, data type = incidence frequency) (Hsieh et al., 2016).

All data handling and analysis was conducted using R (version 4.01), and the *bipartite* (Dormann et al., 2008), *iNEXT* and *tidyverse* (Wickham et al., 2019) packages. Figures were plotted using *ggplot2*, *iNEXT* and *bipartite*.

RESULTS

In total, we compiled data for 9102 biting Diptera blood meals from 45 publications involving field sites in 27 countries (Figure 1a). An aggregated global network contained 227 host species, 202 biting Diptera species and 1121 links (Figure 1b). Based on these data, 47 quantitative bipartite networks were constructed from 14 Agricultural (338 links), 18 Near-natural (461 links) and 15 Village/Urban sites (322 links). Almost all (97%) of our networks comprised Diptera species from a single family: Culicidae (24 cases), Ceratopogonidae (12 cases), Glossinidae (4 cases), Psychodidae (4 cases) and Simuliidae (2 cases). Sampling sites had a wide latitudinal distribution, ranging from Sweden to South Africa (Figure 1a). Total host and Diptera species richness were lower in Agricultural (65 and 64, respectively) than Village/Urban (87 and 74), and Near-natural (140 and 115) habitats (Figure 1c–e). We found no significant trend in host and biting Diptera richness with latitude after controlling for sampling effort. Accumulation curves showed that host and biting Diptera species were well-resolved in Agricultural and Village/Urban habitats, with curves approaching an asymptote in each case; host richness in Near-natural habitats was less complete. Sampling of biting

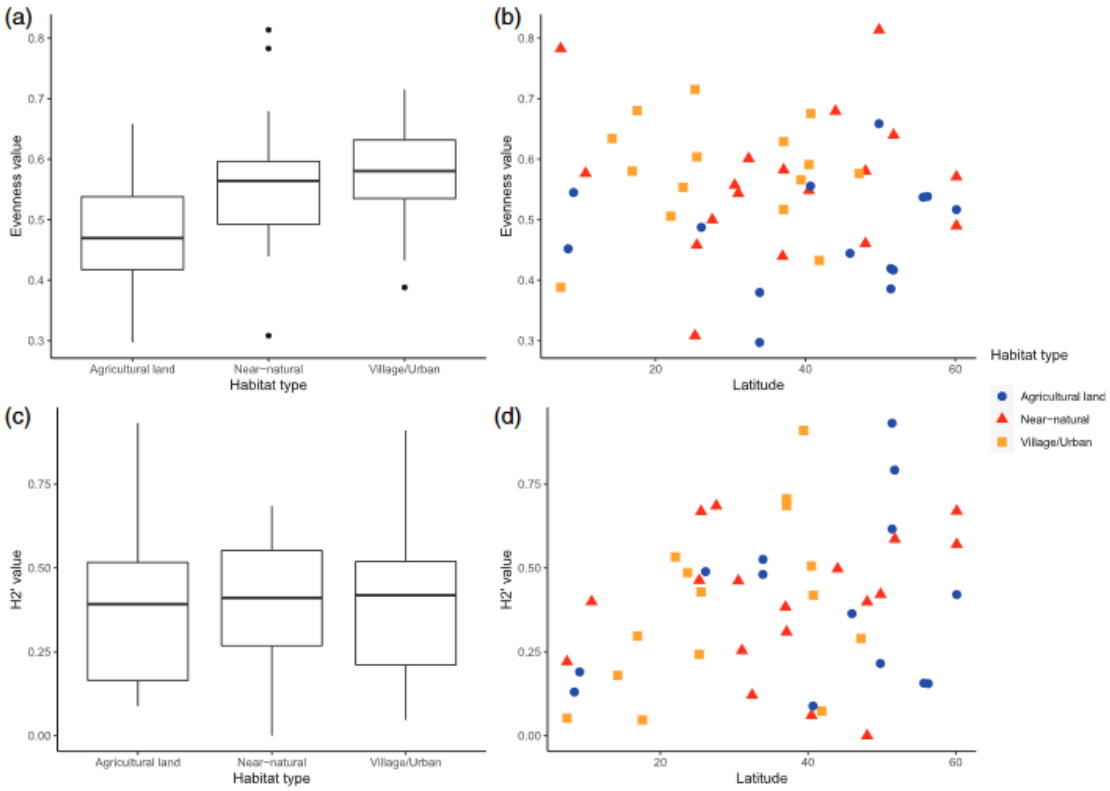


FIGURE 2 Interaction evenness by habitat type (a) and latitude (b), and H2' specialisation by habitat type (c) and latitude (d). Interaction evenness was significantly lower in Agricultural habitats. Each box displays the interquartile range and the solid line represents the median. Whiskers display the maximum and minimum interaction evenness and H2' values for each habitat type.

Diptera–host interactions was incomplete for all levels of anthropogenic landscape modification (Figure S1).

Influence of geographical and anthropogenic factors on biting Diptera–host networks

Interaction evenness differed significantly among habitat types ($X^2 = 0.068$, $df = 2$, $p = 0.042$), but did not show a latitudinal trend ($X^2 = 0.001$, $df = 1$, $p = 0.968$) (Figure 2b). Mean interaction evenness was significantly lower in Agricultural habitats (mean = 0.472, SE = 0.026) than in both Village/Urban (mean = 0.576, SE = 0.023, Tukey; $p = 0.027$) and Near-natural habitats (mean = 0.558, SE = 0.027, Tukey; $p = 0.040$), but did not differ significantly between Village/Urban and Near-natural habitats (Tukey; $p = 0.943$) (Figure 2a). Species richness ($X^2 = 0.014$, $df = 1$, $p = 0.239$), network size ($X^2 = 0.023$, $df = 1$, $p = 0.147$) and family ($X^2 = 0.063$, $df = 4$, $p = 0.180$) did not explain a significant amount of variance in interaction evenness and there were no significant interactions between family and habitat ($X^2 = 0.045$, $df = 5$, $p = 0.495$) or family and latitude ($X^2 = 0.026$, $df = 3$, $p = 0.475$).

The average network specialisation (H2') across our networks was 0.395 (SD = 0.034). Neither latitude ($X^2 = 0.114$, $df = 1$, $p = 0.115$) nor habitat ($X^2 = 0.035$, $df = 1$, $p = 0.682$) had a significant influence on H2' (Figure 2c,d). H2' differed significantly among Diptera families ($X^2 = 0.429$, $df = 4$, $p = 0.030$). Simuliidae-dominated networks had the highest average H2' (mean = 0.619), followed by Ceratopogonidae (mean = 0.296) and Glossinidae (mean = 0.127). There were no significant interactions between family and habitat ($X^2 = 0.207$, $df = 5$, $p = 0.391$) or family and latitude ($X^2 = 0.207$, $df = 3$, $p = 0.242$). There was a highly significant decrease in H2' ($t = -2.875$, $df = 41$, $p = 0.006$) with increasing matrix size.

The influence of domesticated animals on interaction evenness

Humans and domestic animals were involved in 2928 interactions across Agricultural networks, of which 2653 were Diptera–bovine interactions, involving 87 Diptera species from 5 families (Ceratopogonidae, Glossinidae, Culicidae, Psychodidae and Simuliidae). Near-natural networks contained 1886 interactions involving humans and domestic animals, and there were 1560 such interactions in Village/Urban networks. There was no evidence that the presence of domesticated animals and humans within networks led to altered patterns of interaction evenness. Resampling networks to target removal of these species did not lead to mean interaction evenness values that differed from those generated when an equivalent number of host species selected at random was removed: Agricultural (mean = 0.721, $n = 10$, SE = 0.472), Village/Urban (mean = 0.696, $n = 7$, SE = 0.434) and Near-natural (mean = 0.563, $n = 16$, SE = 0.292) (Figure S2).

DISCUSSION

Overall, latitudinal trends in the structure of biting Diptera–host network properties were dwarfed by the impact of anthropogenic habitat modification. Neither network-level metric varied significantly with latitude, but agricultural habitats had significantly lower interaction evenness than Near-natural and Village/Urban habitats.

The absence of a latitudinal trend in feeding specialisation or interaction evenness is counter to the expectation that high species richness at low latitudes will be associated with increased dietary specialisation (Dyer et al., 2007). Turnover in genera of Diptera and hosts was low across latitudes, perhaps explaining the consistency of interspecific interactions. Whilst Diptera families differ in degree of network specialisation as judged by H2', there was no significant interaction between family and latitude. Consequently, the absence of clear latitudinal trends in network structure, as also documented for host–parasitoid networks (Morris et al., 2014), could result from underlying rules for how these antagonistic interactions are structured, regardless of the diversity and size of component networks, or their taxonomic composition. However, the apparent lack of a latitudinal trend may also result from local climatic differences (e.g., rainfall) among sites at similar latitudes masking latitudinal effects (Fischer et al., 2022; Zhu et al., 2014).

The low levels of network specialisation (Agricultural mean H2' = 0.397, Village/Urban mean H2' = 0.390, Near-Natural mean H2' = 0.398) may reflect plasticity in host choice and the wide global distribution of suitable hosts. Host usage is characterised by a high degree of plasticity in biting insects and may be strongly influenced by host densities (Takken & Verhulst, 2013). For example, biting Diptera that are commonly described as anthropophilic such as *Anopheles gambiae* still interact with a range of domestic and non-domestic hosts (Bellekom et al., 2021).

The low interaction evenness observed within Agricultural habitats, in comparison to other habitat types, indicates that interactions are dominated by relatively few species pairs, with a long tail of infrequently observed interactions. In Agricultural habitats, interactions involving domestic animals and humans dominated the networks, accounting for 81% of interactions, with cattle (51%) the most frequent hosts. This may reflect the high biomass of domestic animals (Lassen et al., 2012) in these habitats, and potentially the success and dominance of anthropophilic and livestock-adapted biting Diptera species. Approximately 70% of biting Diptera species in the Agricultural networks fed predominantly (>50% of interactions) on humans and domestic animals, whilst the remaining Diptera fed on either a wider range of mammals and birds or had too few recorded blood meals to assess their diets with confidence (Bellekom et al., 2021). The minimal difference observed between our Agricultural model with targeted removal of domestic hosts and the null model with random removal may be explained by the high number of domestic hosts compared with non-domestic hosts; random removal of hosts inevitably results in the removal of domestic hosts.

We found little difference in interaction evenness between Village/Urban and Near-natural environments, although sampling was

less complete within Near-natural habitats. Despite an expected high human host availability, the Village/Urban networks were not dominated by interactions between humans and biting Diptera to the same extent as Agricultural habitats were by biting Diptera–cattle interactions. In Village/Urban habitats, 26% of interactions were with humans, and we identified interactions with a wide range of other taxa, predominantly birds and domesticated animals. Therefore, high interaction evenness values within Village/Urban sites may result from higher than expected generalism of biting Diptera (Hassell et al., 2017) and the absence of dominating species pairs (pairs with high abundance and interaction fidelity). In contrast, network interaction structure in Agricultural habitats is strongly influenced by the super-abundance of a single suitable host species (cattle). Interaction evenness may therefore largely reflect host species evenness; independent data on host abundances would be required to test this.

For many taxa, specialist species are often more highly represented within pristine habitats, and are more susceptible to anthropogenic landscape modification and homogenisation than generalist species (Devictor, Julliard, & Jiguet, 2008; Devictor, Julliard, Clavel, et al., 2008; Sverdrup-Thygeson et al., 2017). Despite this, we did not find differences in network specialism across levels of anthropogenic landscape modification, perhaps because the same genera and often the same species of biting Diptera were documented across habitat types, leading to similar feeding patterns.

Ecological networks are often asymmetric, composed of few strong interactions and a greater number of weak interactions (Poulin, 2010). Therefore, nodes likely affect each other with differing amounts of reciprocation, and the strength of an interaction is often determined by the distribution of abundance of the component species (Vázquez et al., 2007). Consequently, pairs of abundant species may exhibit more symmetric, and reciprocally strong, effects on the network than pairs of rare species (Dormann et al., 2017). Because independent abundance data were lacking for nodes within the networks, we considered all interactions to be inherently equivalent, with a normalised interaction strength of 1. These data limitations makes it impossible to identify the relative dynamic importance of different nodes. This is a common limitation in ecological network analyses of other interaction types, such as pollination (Novella-Fernandez et al., 2019) and herbivory (Neff et al., 2021), in which the influence of each pollination visit or herbivory damage by different species is considered of equivalent impact on a plant. Network metrics can be sensitive to network size and species richness, leading to a risk that trends in cross-network analyses reflect sampling differences, rather than genuine ecological patterns (Dormann et al., 2009). Heterogeneous data extracted from the literature are particularly susceptible to such biases as a result of variations in methods, sampling intensity and network dimensions (Prendergast & Ollerton, 2022; Xing & Fayle, 2021). Interaction evenness and network specialisation are relatively robust to network size differences (Blüthgen et al., 2006) and we included network size and species richness as explanatory variables in statistical models to control for this potential source of bias.

Agriculturally driven anthropogenic habitat modification, through its effects on biting Diptera–host interaction evenness and network specialism, could result in increased zoonotic disease transmission potential (McDaniel et al., 2014). There was a very high number of biting Diptera–bovine interactions in our Agricultural networks, involving a wide range of Diptera species from multiple families, many of which are vector-competent. This may be of particular concern, since bovine-related diseases such as Rift Valley fever, Animal African Trypanosomosis (nagana) and Bluetongue disease have high morbidity and mortality rates (Lopes et al., 2020; Rushton & Lyons, 2015; Vreysen et al., 2013). The growing global demand for agriculture products will result in continued anthropogenic habitat modification, which will provide increasing opportunities for pairwise interactions between unfamiliar species, zoonotic transmission and the emergence of novel zoonotic disease (Carlson et al., 2022). Surveillance of biting Diptera–host networks, particularly at the interface of humans, wildlife and domestic animals, could help identify pathways of zoonotic disease transmission and help predict and mitigate future spill-over events. Surveillance may be conducted through the routine sampling of the Diptera community using a combination of trapping methods, such as malaise traps, USA Center for Disease Control (CDC) miniature light traps and Modified CDC Backpack Aspirators, Biogents' Sentinel (BGS) as well as trapping locations that limit accidental overrepresentation of a species in order to minimise sources of bias (Bellekom et al., 2021; Grubaugh et al., 2015; Gyawali et al., 2019; Rivera et al., 2021).

AUTHOR CONTRIBUTIONS

Ben Bellekom: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; validation; visualization; writing – original draft; writing – review and editing. **Owen T. Lewis:** Conceptualization; funding acquisition; supervision; writing – review and editing. **Talya D. Hackett:** Conceptualization; funding acquisition; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

We thank John Connolly and the late John Mumford for their helpful comments and suggestions. We also thank three anonymous reviewers for their helpful comments. We also thank Robbie Bellekom for the design and creation of the graphical abstract. Talya D. Hackett and Owen T. Lewis were supported by Target Malaria, which is supported by a grant from the Bill and Melinda Gates Foundation and the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation. The first and corresponding author is an early career researcher, and the last author is a woman.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Latitudinal-and-anthropogenic-effects-on-the-structuring-of-networks at https://github.com/Ben-Bellekom/Latitudinal-and-anthropogenic-effects-on-the-structuring-of-networks.

REFERENCES

Abella-Medrano, C.A., Ibáñez-Bernal, S., Carbó-Ramírez, P. & Santiago-Alarcon, D. (2018) Blood-meal preferences and avian malaria detection in mosquitoes (Diptera: Culicidae) captured at different land use types within a neotropical montane cloud forest matrix. *Parasitology International*, 67(3), 313–320. Available from: https://doi.org/10.1016/J.PARINT.2018.01.006

Baker, R.E., Mahmud, A.S., Miller, I.F., Rajeev, M., Rasambainarivo, F., Rice, B.L. et al. (2021) Infectious disease in an era of global change. *Nature Reviews. Microbiology*, 20(4), 193–205. Available from: https://doi.org/10.1038/s41579-021-00639-z

Barber, L.M., Schleier, J.J. & Peterson, R.K.D. (2010) Economic cost analysis of West Nile virus outbreak, Sacramento County, California, USA, 2005. *Emerging Infectious Diseases*, 16(3), 480–486. Available from: https://doi.org/10.3201/EID1603.090667

Bellekom, B., Hackett, T.D. & Lewis, O.T. (2021) A network perspective on the vectoring of human disease. *Trends in Parasitology*, 37(5), 391–400. Available from: https://doi.org/10.1016/j.pt.2020.12.001

Blüthgen, N., Fründ, J., Vázquez, D.P. & Menzel, F. (2008) What do interaction network metrics tell us about specialization and biological traits. *Ecology*, 89(12), 3387–3399. Available from: https://doi.org/10.1890/07-2121.1

Blüthgen, N., Menzel, F. & Blüthgen, N. (2006) Measuring specialization in species interaction networks. *BMC Ecology*, 6(1), 9. Available from: https://doi.org/10.1186/1472-6785-6-9

Carlson, C.J., Albery, G.F., Merow, C., Trisos, C.H., Zipfel, C.M., Eskew, E.A. et al. (2022) Climate change increases cross-species viral transmission risk. *Nature*, 607, 555–562. Available from: https://doi.org/10.1038/s41586-022-04788-w

Devictor, V., Julliard, R., Clavel, J., Jiguet, F., Lee, A. & Couvet, D. (2008) Functional biotic homogenization of bird communities in disturbed landscapes. *Global Ecology and Biogeography*, 17(2), 252–261. Available from: https://doi.org/10.1111/j.1466-8238.2007.00364.x

Devictor, V., Julliard, R. & Jiguet, F. (2008) Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos*, 117(4), 507–514. Available from: https://doi.org/10.1111/J.0030-1299.2008.16215.X

Dormann, C.F., Frund, J., Bluthgen, N. & Gruber, B. (2009) Indices, graphs and null models: analyzing bipartite ecological networks. *The Open Ecology Journal*, 2(1), 7–24. Available from: https://doi.org/10.2174/1874213000902010007

Dormann, C.F., Fründ, J. & Schaefer, H.M. (2017) Identifying causes of patterns in ecological networks: opportunities and limitations. *Annual Review of Ecology Evolution and Systematics*, 48, 559–584. Available from: https://doi.org/10.1146/ANNUREV-ECOLSYS-110316-022928

Dormann, C.F., Gruber, B. & Fründ, J. (2008) *Introducing the bipartite Package: Analysing Ecological Networks*. https://www.semanticscholar.org/paper/Introducing-the-bipartite-Package%3A-Analysing-Dormann-Gruber/f23d90643ae91bd7f33650a69069b0221ca13b30

Dyer, L.A., Singer, M.S., Lill, J.T., Stireman, J.O., Gentry, G.L., Marquis, R.J. et al. (2007) Host specificity of lepidoptera in tropical and temperate forests. *Nature*, 448(7154), 696–699. Available from: https://doi.org/10.1038/nature05884

Fischer, C., Gerstmeier, R. & Wagner, T.C. (2022) Seasonal and temporal patterns of rainfall shape arthropod community composition and multi-trophic interactions in an arid environment. *Scientific Reports*, 12(1), 1–9. Available from: https://doi.org/10.1038/s41598-022-07716-0

Galiana, N., Hawkins, B.A. & Montoya, J.M. (2019) The geographical variation of network structure is scale dependent: understanding the biotic specialization of host–parasitoid networks. *Ecography*, 42(6), 1175–1187. Available from: https://doi.org/10.1111/ecog.03684

GBIF.org. (2021) *GBIF Home Page*. Https://Www.Gbif.Org

Grubaugh, N.D., Sharma, S., Krajacich, B.J., Fakoli, L.S., III, Bolay, F.K., Diclaro, J.W., II et al. (2015) Xenosurveillance: a novel mosquito-based approach for examining the human-pathogen landscape. *PLoS Neglected Tropical Diseases*, 9(3), e0003628. Available from: https://doi.org/10.1371/journal.pntd.0003628

Gyawali, N., Taylor-Robinson, A.W., Bradbury, R.S., Huggins, D.W., Hugo, L.E., Lowry, K. et al. (2019) Identification of the source of blood meals in mosquitoes collected from North-Eastern Australia. *Parasites & Vectors*, 12(1), 198. Available from: https://doi.org/10.1186/s13071-019-3455-2

Hassell, J.M., Begon, M., Ward, M.J. & Fèvre, E.M. (2017) Urbanization and disease emergence: dynamics at the wildlife–livestock–human Interface. *Trends Ecol Evol*, 32(1), 55–67. Available from: https://doi.org/10.1016/j.tree.2016.09.012

Hsieh, T.C., Ma, K.H. & Chao, A. (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, 7(12), 1451–1456. Available from: https://doi.org/10.1111/2041-210X.12613

Kaiser-Bunbury, C.N. & Blüthgen, N. (2015) Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants*, 7, plv076. Available from: https://doi.org/10.1093/aobpla/plv076

Killick-Kendrick, R. (1999) The biology and control of phlebotomine sand flies. *Clinics in Dermatology*, 17(3), 279–289. Available from: https://doi.org/10.1016/S0738-081X(99)00046-2

Kilpatrick, A.M., Kramer, L.D., Campbell, S.R., Alleyne, E.O., Dobson, A.P. & Daszak, P. (2005) West Nile virus risk assessment and the bridge vector paradigm. *Emerging Infectious Diseases*, 11(3), 425–429. Available from: https://doi.org/10.3201/EID1103.040364

Lassen, S., Nielsen, S. & Kristensen, M. (2012) Identity and diversity of blood meal hosts of biting midges (Diptera: Ceratopogonidae: Culicoides Latreille) in Denmark. *Parasites & Vectors*, 5(1), 143. Available from: https://doi.org/10.1186/1756-3305-5-143

Lawrence, E.R. & Fraser, D.J. (2020) Latitudinal biodiversity gradients at three levels: linking species richness, population richness and genetic diversity. *Global Ecol Biogeogr*, 29(5), 770–788. Available from: https://doi.org/10.1111/geb.13075

Logue, K., Keven, J.B., Cannon, M.V., Reimer, L., Siba, P., Walker, E.D. et al. (2016) Unbiased characterization of anopheles mosquito blood meals by targeted high-throughput sequencing. *PLoS Neglected Tropical Diseases*, 10(3), e0004512. Available from: https://doi.org/10.1371/journal.pntd.0004512

Lopes, V.L., Costa, F.V., Rodrigues, R.A., Braga, M., Pichorim, M. & Moreira, P.A. (2020) High fidelity defines the temporal consistency of host-parasite interactions in a tropical coastal ecosystem. *Scientific Reports*, 10(1), 16839. Available from: https://doi.org/10.1038/S41598-020-73563-6

Mayi, M.P.A., Bamou, R., Djiappi-Tchamen, B., Fontaine, A., Jeffries, C.L., Walker, T. et al. (2020) Habitat and seasonality affect mosquito community composition in the west region of Cameroon. *Insects*, 11(5), 312. Available from: https://doi.org/10.3390/insects11050312

McDaniel, C.J., Cardwell, D.M., Moeller, R.B. & Gray, G.C. (2014) Humans and cattle: a review of bovine zoonoses. *Vector Borne and Zoonotic Diseases*, 14(1), 1–19. Available from: https://doi.org/10.1089/VBZ.2012.1164

Meyer Steiger, D.B., Ritchie, S.A. & Laurance, S.G.W. (2016) Mosquito communities and disease risk influenced by land use change and seasonality in the Australian tropics. *Parasites & Vectors*, 9(1), 387. Available from: https://doi.org/10.1186/s13071-016-1675-2

Morris, R.J. (2010) Anthropogenic impacts on tropical forest biodiversity: a network structure and ecosystem functioning perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1558), 3709–3718. Available from: https://doi.org/10.1098/rstb.2010.0273

Morris, R.J., Gripenberg, S., Lewis, O.T. & Roslin, T. (2014) Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*, 17(3), 340–349. Available from: https://doi.org/10.1111/ele.12235

Müller, R., Reuss, F., Kendrovski, V. & Montag, D. (2019) Vector-borne diseases. In: *Biodiversity and health in the face of climate change*. Cham: Springer International Publishing, pp. 67–90. Available from: https://doi.org/10.1007/978-3-030-02318-8_4

Neff, F., Brändle, M., Ambarli, D., Ammer, C., Bauhus, J., Boch, S. et al. (2021) Changes in plant-herbivore network structure and robustness along land-use intensity gradients in grasslands and forests. *Science Advances*, 7(20), 18. Available from: https://doi.org/10.1126/SCIADV.ABF3985

Novella-Fernandez, R., Rodrigo, A., Arnan, X. & Bosch, J. (2019) Interaction strength in plant-pollinator networks: are we using the right measure? *PLoS One*, 14(12), e0225930. Available from: https://doi.org/10.1371/JOURNAL.PONE.0225930

Pellissier, L., Albouy, C., Bascompte, J., Farwig, N., Graham, C., Loreau, M. et al. (2018) Comparing species interaction networks along environmental gradients. *Biological Reviews*, 93(2), 785–800. Available from: https://doi.org/10.1111/brv.12366

Poulin, R. (2010) Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, 26(10), 492–498. Available from: https://doi.org/10.1016/J.PT.2010.05.008

Prendergast, K.S. & Ollerton, J. (2022) Spatial and temporal scale of analysis alter conclusions about the effects of urbanisation on plant–pollinator networks. *Arthropod-Plant Interactions*, 1, 1–13. Available from: https://doi.org/10.1007/S11829-022-09925-W/FIGURES/2

Rivera, N.A., Varga, C., Ruder, M.G., Dorak, S.J., Roca, A.L., Novakofski, J.E. et al. (2021) Bluetongue and epizootic hemorrhagic disease in The United States of America at the wildlife–livestock Interface. *Pathogens*, 10(8), 915. Available from: https://doi.org/10.3390/PATHOGENS10080915

Runghen, R., Poulin, R., Monlleó-Borrull, C. & Llopis-Belenguer, C. (2021) Network analysis: ten years shining light on host–parasite interactions. *Trends in Parasitology*, 37(5), 445–455. Available from: https://doi.org/10.1016/j.pt.2021.01.005

Rushton, J. & Lyons, N. (2015) Economic impact of bluetongue: a review of the effects on production. *Veterinaria Italiana*, 51(4), 401–406. Available from: https://doi.org/10.12834/VETIT.646.3183.1

Sachs, J. & Malaney, P. (2002) The economic and social burden of malaria. *Nature*, 415(6872), 680–685. Available from: https://doi.org/10.1038/415680a

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M., Roy, K. & Kellogg, W.K. (2009) Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology, Evolution, and Systematics*, 40, 245–269. Available from: https://doi.org/10.1146/annurev.ecolsys.39.110707.173430

Schleuning, M., Fründ, J., Klein, A.-M., Abrahamczyk, S., Alarcón, R., Albrecht, M. et al. (2012) Specialization of mutualistic interaction networks decreases toward tropical latitudes. *Current Biology*, 22(20), 1925–1931. Available from: https://doi.org/10.1016/J.CUB.2012.08.015

Sinka, M.E., Bangs, M.J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M. et al. (2012) A global map of dominant malaria vectors. *Parasites & Vectors*, 5(1), 69. Available from: https://doi.org/10.1186/1756-3305-5-69

Sverdrup-Thygeson, A., Skarpaas, O., Blumentrath, S., Birkemoe, T. & Evju, M. (2017) Habitat connectivity affects specialist species richness more than generalists in veteran trees. *Forest Ecology and Management*, 403, 96–102. Available from: https://doi.org/10.1016/J.FORECO.2017.08.003

Takken, W. & Verhulst, N.O. (2013) Host preferences of blood-feeding mosquitoes. *Annual Review of Entomology*, 58(1), 433–453. Available from: https://doi.org/10.1146/annurev-ento-120811-153618

Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007) Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, 445(7124), 202–205. Available from: https://doi.org/10.1038/nature05429

Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. & Poulin, R. (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, 116(7), 1120–1127. Available from: https://doi.org/10.1111/J.0030-1299.2007.15828.X

Vreysen, M.J.B., Seck, M.T., Sall, B. & Bouyer, J. (2013) Tsetse flies: their biology and control using area-wide integrated pest management approaches. *Journal of Invertebrate Pathology*, 112, S15–S25. Available from: https://doi.org/10.1016/J.JIP.2012.07.026

Weiner, C.N., Werner, M., Linsenmair, K.E. & Blüthgen, N. (2014) Land-use impacts on plant–pollinator networks: interaction strength and specialization predict pollinator declines. *Ecology*, 95(2), 466–474. Available from: https://doi.org/10.1890/13-0436.1

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R. et al. (2019) Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. Available from: https://doi.org/10.21105/joss.01686

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, 34, 273–309. Available from: https://doi.org/10.1146/annurev.ecolsys.34.012103.144032

Xing, S. & Fayle, T.M. (2021) The rise of ecological network meta-analyses: problems and prospects. *Global Ecology and Conservation*, 30, e01805. Available from: https://doi.org/10.1016/j.gecco.2021.e01805

Zhu, H., Wang, D., Wang, L., Fang, J., Sun, W. & Ren, B. (2014) Effects of altered precipitation on insect community composition and structure in a meadow steppe. *Ecological Entomology*, 39(4), 453–461. Available from: https://doi.org/10.1111/EEN.12120

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Smoothed accumulation and extrapolation curves to assess sampling completeness. Total numbers of hosts (triangle), biting Diptera (circle), and interactions recorded in the whole dataset (square), by habitat type: Agricultural (blue), Near-natural (red), and Village/Urban (orange), as a function of sampling effort (the number of blood meals screened).

Figure S2. Null model interaction evenness for each component network, with empirical IE values (red squares), by habitat type: Agricultural (a), Village/Urban (b), and Near-natural (c). Each grey box displays the interquartile range, and the solid line represents the median values for interaction evenness. Whiskers display the maximum and minimum interaction evenness for each network.

How to cite this article: Bellekom, B., Lewis, O.T. & Hackett, T.D. (2023) Latitudinal and anthropogenic effects on the structuring of networks linking blood-feeding flies and their vertebrate hosts. *Medical and Veterinary Entomology*, 1–8. Available from: https://doi.org/10.1111/mve.12671

Appendix 5. Journal offprint: Effects of storage conditions and digestion time on DNA amplification of biting midge (*Culicoides*) blood meals

Bellekom *et al. Parasites & Vectors* (2023) 16:13
https://doi.org/10.1186/s13071-022-05607-x

Parasites & Vectors

BRIEF REPORT Open Access

Effects of storage conditions and digestion time on DNA amplification of biting midge (*Culicoides*) blood meals

Ben Bellekom[1*], Abigail Bailey[1], Marion England[2], Zoe Langlands[2], Owen T. Lewis[1] and Talya D. Hackett[1]

Abstract

Background Molecular analysis of blood meals is increasingly used to identify the hosts of biting insects such as midges and mosquitoes. Successful host identification depends on the availability of sufficient host DNA template for PCR amplification, making it important to understand how amplification success changes under different storage conditions and with different durations of blood meal digestion within the insect gut before being placed into the storage medium.

Method We characterised and compared the digestion profile of two species of *Culicoides* over a 96-h period using a novel set of general vertebrate primers targeting the 16S rRNA gene. A set number of individuals from each species were killed over 13 time points post-blood feeding and preserved in 95% ethanol. Samples were stored either at ambient room temperature or in a − 20 °C freezer to examine the effect of storage condition on the PCR amplification success of host DNA.

Results We found that amplification success across the 96-h sampling period post-feeding was reduced from 96 to 6% and 96% to 14% for *Culicoides nubeculosus* and *Culicoides sonorensis*, respectively. We found no effect of storage condition on PCR amplification success, and storage in 95% ethanol was sufficient to maintain high rates of amplifiable host DNA for at least 9 months, even at room temperature.

Conclusions These findings highlight the limited time frame during which an individual may contain amplifiable host DNA and demonstrate the importance of timely sample capture and processing post-blood feeding. Moreover, storage in 95% ethanol alone is sufficient to limit host DNA degradation. These results are relevant to the design of studies investigating the biting behaviour and disease transmission potential of *Culicoides* and other biting Diptera.

Keywords *Culicoides*, DNA degradation, Metabarcoding, Blood meal, Biting diptera, DNA digestion

Introduction

Biting (haematophagous) Diptera are vectors of a wide range of pathogens that cause disease in humans as well as domesticated and wild animals. Transmission dynamics of these diseases are, in part, mediated by the host-biting preferences of the vectors [1]. Thus, data on vertebrate blood meal sources and host preferences can provide valuable information for understanding disease transmission [2], including transmission across the livestock and wildlife interface. Blood meal data can also be used to populate ecological interaction networks, helping to identify transmission pathways and reservoirs of infection and informing control strategies [3].

In the past, host blood meals were identified using serological methods which have now been largely superseded

*Correspondence:
Ben Bellekom
ben.bellekom@biology.ox.ac.uk
[1] Department of Biology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, UK
[2] The Pirbright Institute, Ash Road, Pirbright, Surrey GU24 0NF, UK

by molecular methods, diagnostic PCR, DNA barcoding and, most recently, metabarcoding [2, 3]. These methods rely on extraction of DNA from the blood meal and PCR amplification of a taxonomically informative fragment of host DNA [1] followed by sequencing of the PCR product or diagnostic PCR using primers specific to a species or set of species. From as little as US $2/sample at 2022 prices, barcoding is now a relatively low-cost method for sequencing a single gene, but it is unsuitable for mixtures of DNA from multiple species. Metabarcoding coupled with next generation sequencing allows many sequences to be read in parallel, facilitating the analysis of multiple gene loci and individual samples and the simultaneous identification of species within mixed samples such as gut contents [4] and blood meals [1]. Whilst sequencing costs have dramatically fallen, metabarcoding remains more expensive than targeted DNA barcoding. Price per individual sample may vary greatly depending on the sequencing kit and platform used, the number of samples being processed and the target read depth.

DNA degradation by digestion enzymes will reduce the success of PCR amplification, jeopardising successful blood meal diagnosis [2, 5–7]. In mosquitoes (Culicidae), digestion occurs rapidly, with host DNA undetectable within 32–72 h [5, 8]. However, rates of DNA degradation attributable to digestion are likely to vary among haematophagous taxa, and equivalent data for other biting Diptera families of medical, veterinary and epidemiological importance are lacking. Prompt preservation of blood-fed biting Diptera to halt digestive processes can increase the success of PCR amplification [9]. Preservation methods include storage in ethanol [10], desiccation [11], transferring blood meals onto filter paper [12] and cryopreservation [13, 14]. However, field constraints and institutional limitations often limit the use of − 20 °C and ultra-low (− 80 °C) freezers and collected samples may be left for extended periods at room temperature, especially during transport. It is therefore of considerable practical interest to understand how amplification success varies for samples stored under different conditions.

Here we examine the effect of digestion time and the impact of two storage conditions (ambient room temperature and a − 20 °C freezer) on the PCR amplification success of host DNA in blood meals from two species of *Culicoides*. Biting midges in this genus have a wide distribution globally, bite a broad range of host species [10] and are vectors of several pathogens of veterinary and medical significance [10, 15]. We used PCR amplification success to confirm the presence of sufficient quantities of non-degraded host template DNA to allow for further molecular analysis and host identification. In parallel, we assess the suitability of a novel 16S rRNA general vertebrate primer set for Diptera blood meal identification.

Methods

We experimentally examined the effect of blood meal digestion time and storage method on the PCR amplification success of host DNA in two species of biting midges (*Culicoides nubeculosus* and *C. sonorensis*) using novel vertebrate PCR primers. Individuals from each species were fed fresh horse blood, killed and preserved in 95% ethanol at 13 specific time points, ranging from freshly fed to 96 h post-feeding, in 8-h intervals. At each time point, 30 individuals from each species were knocked down by freezing and stored in ethanol to minimise DNA degradation. An additional 10 individuals at time point 0 were collected to investigate long-term storage effects, thus resulting in 40 individuals at time point 0. In total, 800 midges were collected for blood meal analysis (Table 1).

To examine the effect of storage method, 15 samples from each species at every time point were preserved in 95% ethanol and stored in a − 20 °C freezer, whilst the remaining 15 were stored in ethanol and left at room temperature (22–24 °C). To minimise biases, samples were extracted and underwent PCR amplification in a randomised order across all time points and storage methods. To examine the impact of long-term storage on DNA integrity, two sets of five individuals of each species from the first time point were preserved in ethanol and stored under the two temperature conditions (− 20 °C freezer and at ambient room temperature), respectively, and left for 9 months before extraction and amplification. Extraction and amplification of all samples, excluding the long-term storage sets, took place across a 3-month period following sample preparation.

Midge rearing

Culicoides sonorensis and *C. nubeculosus* specimens were obtained from lines maintained in existing closed colonies at the Pirbright Institute. The *C. nubeculosus* colony was established in 1969 from pupae collected in Hertfordshire, UK. The *C. sonorensis* colony was established from eggs provided by Dr. H. Jones who initiated a laboratory colony in Colorado in 1957. The colonies were maintained following previously developed protocols [16] with adult females fed on commercially supplied horse blood (TCS Biosciences, UK) using a Hemotek blood feeder (Hemotek, UK). Midges were not sugar fed during the 96-h period as a trial conducted prior to the commencement of the study indicated that unfed midges survived beyond the 96-h period (unpublished data). Furthermore, this most accurately represented the behaviour

Table 1 The number of individuals (and successful amplifications) for both species of *Culicoides* that underwent PCR amplification by time point (h) and storage method

Time point (h)	*Culicoides nubeculosus* Ambient	*Culicoides nubeculosus* − 20 °C Freezer	*Culicoides sonorensis* Ambient	*Culicoides sonorensis* − 20 °C Freezer
0	20 (20)	20 (19)	20 (19)	20 (20)
8	15 (15)	15 (15)	15 (15)	15 (15)
16	15 (15)	15 (15)	15 (15)	15 (15)
24	15 (14)	15 (14)	15 (14)	15 (14)
32	15 (14)	15 (13)	15 (10)	15 (9)
40	15 (4)	15 (9)	15 (4)	15 (3)
48	15 (4)	15 (6)	15 (4)	15 (5)
56	15 (2)	15 (2)	15 (3)	15 (1)
64	15 (9)	15 (4)	13 (0)	12 (0)
72	15 (2)	15 (1)	15 (3)	15 (4)
80	15 (3)	15 (1)	8 (0)	7 (2)
88	15 (4)	15 (3)	12 (2)	10 (0)
96	15 (0)	15 (2)	7 (1)	7 (1)

of females in the field following feeding, whereby they rest until oogenesis is complete and then find a suitable habitat to lay [17].

Blood meal analysis

DNA was extracted from individual blood meals using a Qiagen DNeasy Blood and Tissue kit, using the standard protocol with the following minor alterations. Prior to lysis, individuals were homogenised using an MP Biomedicals FastPrep-24 5G homogeniser to release the blood meal from the abdomen. Prior to homogenisation, a single sterile 2.3-mm zirconia/silica agitating ball was placed in each 1.5-ml microcentrifuge tube with 180 µl of Buffer ATL. To minimise Buffer ATL foaming, which may reduce the homogenisation efficiency, 2 µl of Reagent DX was added to each microcentrifuge tube. To increase final DNA concentration, prior to elution, 60 µl of Buffer AE was pipetted onto the Dneasy spin column membrane and allowed to incubate at room temperature for 5 min. Final DNA concentration for all samples was quantified using a Qubit 3.0 fluorometer.

In the absence of a priori host assumptions and the potential presence of mixed blood meals (derived from different host species), the PCR amplification of host genomic DNA from blood meals of wild-caught *Culicoides* and other haematophagous Diptera requires the use of general vertebrate primers [18]. To facilitate the identification of horse-derived blood meals used in this study and to determine the likelihood of successful amplification in future blood meal metabarcoding studies, we designed primers intended to amplify vertebrate templates on the 16S rRNA gene, whilst excluding invertebrate templates. We downloaded 128 mitogenomes belonging to the classes Aves and Mammalia from the NCBI Genbank database. As these primers are designed for blood meal analysis of a wide range of biting Diptera, we also downloaded five biting Diptera mitogenomes (*Anopheles gambiae, Aedes albopictus, Aedes aegypti, Culicoides arakawae* and *Culicoides imicola*). The sequences were MAFFT aligned using the bioinformatics software Geneious Prime (Biomatters Limited, New Zealand). Where DNA is highly degraded, primer pairs targeting short amplicons are preferable. Consequently, this alignment was used to identify two potential 18–22-bp primer binding sites that contained primer-invertebrate annealing site mismatches, which produced a 200-base pair (bp) amplicon. Initial binding site selection was informed by previous work [19]. To ensure binding across all mammal and bird mitogenomes, two degenerate bases were inserted into the forward primer to account for single base mismatches. Primer protocol selection and validation were conducted using gradient PCRs with a range of known mammalian and avian DNA templates (Additional file 1: Text S1.). As the PCR primers were designed for use with a wide range of biting Diptera species, examination of blood meal amplification success and identification of non-specific binding was conducted on genomic DNA from blood fed and unfed *C. sonorensis* and *An. gambiae*. I. Fed and unfed *C. sonorensis* were obtained from the colonies at the Pirbright Institute. *Anopheles gambiae* individuals were wild caught and stored in 95% ethanol at field sites in Burkina Faso as part of ongoing research by members of the Target Malaria research consortium (see Acknowledgements). The PCR product of blood-fed *C. sonorensis* and

An. gambiae were Sanger sequenced (Source Bioscience) to confirm successful target amplification.

The presence of sufficient concentrations of non-degraded host DNA template for PCR amplification was assessed using end-point PCR and our novel primer set, 16smbF (5′-GGT TGG GGY GAC CTY GGA-3′) and 16sbbR (5′--CTG ATC CAA CAT CGA GGT CGT A-3′-). PCR amplification was carried out in 25 µl reactions that contained 12.5 µl HotStarTaq Master mix (Qiagen, Germany), 10 µM of each primer, 8.5 µl H_2O and 2 µl DNA template. The PCR protocol consisted of an initial denaturation step of 15 min at 95 °C followed by 35 cycles of 94 °C for 45 s, 58 °C for 45 s and 72 °C for 30 s, followed by a final extension step of 72 °C for 10 min. Two negative controls, containing nuclease-free water (ThermoFisher Scientific, USA), were included in every set of reactions to monitor for contamination.

PCR products were electrophoresed and visualised on a gel red (SYBR™ Safe, ThermoFisher Scientific, USA) stained 2% agarose gel. The presence of a band of the expected amplicon size was taken as a positive result, indicating that sufficient host DNA remained viable for blood meal identification. To confirm successful target amplification, a subset of PCR products from across the range of time steps that yielded bands were Sanger sequenced (Source Bioscience, England) and the sequence's origin was identified using the basic alignment search tool (BLAST).

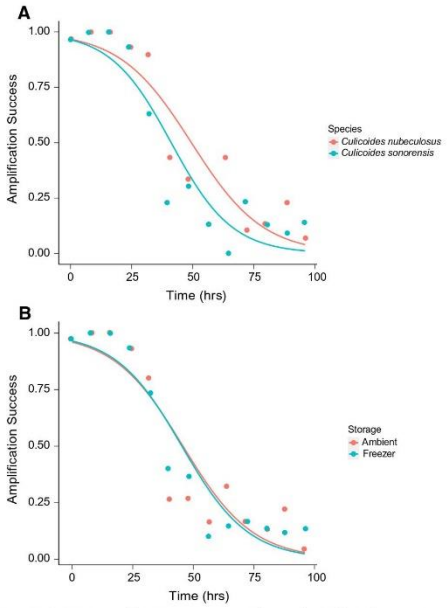


Fig. 1 (**A**) PCR amplification success with time by *Culicoides* species. Red and blue points represent the amplification success of *Culicoides nubeculosus* and *C. sonorensis*, respectively, at a given time point. (**B**) PCR amplification success with time by storage method. The red and blue points represent amplification success of ambient and − 20 °C freezer, conditions respectively, at a given time point

Statistical analysis

We modelled the effect of digestion time, storage method and species on a binary measure of amplification success (1 or 0) using a multivariate binomial logistic regression. Model fit was evaluated with a likelihood ratio test and the significance of the overall effect of each variable was determined using the chi-squared statistic [5]. Data were analysed using R (version 4.1.2) and visualised using the package *ggplot2* [20].

Results

Overall, 759 midge samples were used in this study. Some midges died during the later time steps (64 h post-feeding onwards), so four time steps had fewer than the planned 15 *C. sonorensis* individuals (Table 1).

Digestion time significantly decreased amplification success of host DNA ($X^2 = 402.48$, df = 1, $P < 0.001$), with amplification success approaching 6% 96 h post-feeding (Fig. 1A). We found no significant effect of storage method on amplification success ($X^2 = 0.024$, df = 1, $P = 0.875$; Fig. 1B). Amplification success was significantly higher for *Culicoides nubeculosus* than for *C. sonorensis* ($X^2 = 8.318$, df = 1, $P = 0.004$; Fig. 1). All 20 long-term individuals retained amplifiable host DNA following the 9-month storage period.

Discussion

We found that the duration of blood meal digestion, but not the storage condition (ambient or − 20 °C freezer), had a significant impact on the success of host DNA PCR amplification and thus the likelihood of subsequent host identification. Whilst four time steps post 64 h had fewer *C. sonorensis* individuals than planned, amplification during this period was limited in both species and the reduction in sample size is unlikely to have impacted the findings significantly.

Increased post-feeding digestion time significantly reduced amplification success of blood meals from both colony *Culicoides* species used in this study, with a rapid decrease after the 32-h time step. Similar results have been documented in mosquitoes [7], suggesting that rates

of DNA degradation in digested blood meals can be similar in hematophagous taxa that differ markedly in blood meal size. Consistency in the digestion profiles of these taxa may reflect similar production patterns of digestive enzymes such as late trypsin, a protease responsible for the endoproteolytic cleavage of protein in the blood meal [21–23].

Our findings highlight the limited time frame over which samples are usable for blood meal analysis and the importance of timely sample capture and processing and may partially explain variations in *Culicoides* blood meal amplification success in previous studies [10, 24]. Whilst it was possible to amplify host DNA successfully at the later time steps, success was significantly reduced and may bias later analyses, such as examination of host preferences and species roles in networks. Selection of trapping methods and sampling protocols should account for this to maximise the proportion of amplifiable blood meals. Specifically, we show that storage in 95% ethanol limits blood meal degradation of freshly fed biting Diptera even after prolonged periods of storage. Furthermore, individuals should ideally be killed as soon as possible after feeding to halt digestive processes, using trapping methods that catch individuals directly into ethanol. Traps commonly used for *Culicoides*, such as OVI (Onderstepoort Veterinary Institute) light traps and CDC (Centers for Disease Control) light traps, use ethanol-filled collection chambers. However, mosquitoes are often collected with designs such as the BG Sentinel Trap, where collections are kept dry to facilitate morphological identification of specimens. Therefore, if the primary goal is to identify host blood meals, it may be advisable to modify these traps to include an ethanol-filled chamber, although this may necessitate molecular identification of the sampled insect if immersion in liquid causes damage to scales used for morphology-based identifications.

Halting digestion and minimising degradation of host DNA post-capture is of particular importance for downstream blood meal analysis. We found no significant difference in amplification success between samples stored in 95% ethanol in a − 20 °C freezer versus at room temperature, suggesting that storage in 95% ethanol is sufficient to limit host DNA degradation in *Culicoides*. These findings are consistent with previous work, which suggested that ethanol alone is sufficient to maintain host DNA integrity in mosquito blood meals [5, 25]. The limited effect of storage at ambient temperatures is relevant for the planning and logistics of field work. Collection of *Culicoides* and other biting Diptera often necessitates sampling in remote locations without access to freezers; our results provide reassurance that, given proper processing, samples can safely be stored and transported at ambient temperatures for extended periods without impacting amplification success. The range of ambient temperatures used in our study reflects those in temperate regions, where monitoring of *Culicoides* species and limiting the spread of *Culicoides*-borne diseases is a major concern [24, 26–30]. Ethanol has also been shown to be an effective short- to medium-term mosquito blood meal storage method under higher temperature tropical conditions [5]. This suggests that our results might also apply to *Culicoides* samples stored under ambient conditions in the tropics, though further work is required to confirm this; there is potentially an upper temperature limit where 95% ethanol does not sufficiently halt DNA degradation in blood meals.

We found slightly greater amplification success for blood meals of *C. nubeculosus* compared with *C. sonorensis*. These *Culicoides* species are of similar size, making it unlikely that blood meal size was a factor [31]. Differences between the species were most apparent at a time estimated to correspond to stage 3 of *Culicoides* digestion [32], during which secretory granules associated with the production of digestive enzymes are formed. Thus, the differences in amplification success between the species may reflect differences in granule and digestive enzyme production. Alternatively, the higher mortality of *C. sonorensis* during the later time steps (64 h onwards) and the differing amplification success for the two species may reflect minor differences in the abiotic conditions under which the individuals were reared or slight differences in life history traits. For example, minor variations in ambient temperature can impact longevity [33] and rates of blood meal digestion [7].

General vertebrate PCR primers, when combined with high-throughput sequencing, allow identification of the range of host DNA template contained in mixed blood meals (metabarcoding). Moreover, general primers that target small amplicons are advantageous in blood meal analysis because of the high degree of DNA degradation resulting in fragmentation of the DNA template [1]. We identified a 200-bp region of the 16 s gene that is suitable for the interrogation of blood meals from a wide range of biting Diptera whilst avoiding co-amplification of Dipteran DNA. Our primers and PCR protocol were effective with highly degraded DNA, successfully amplifying host templates during later time steps.

Conclusions

Based on the findings of this work, we propose the following recommendations. We suggest the use of the primers and the PCR protocol described here to complement existing primer sets in future biting Diptera blood meal analysis. Moreover, because of the efficiency and versatility of ethanol for maintaining blood meal DNA integrity and the

need to halt digestive processes to maximise amplification success, we recommend the use of trapping methods that contain an ethanol-filled collection device, though potential ethanol evaporation should be accounted for by regular trap emptying and replenishment of ethanol. During field sampling of blood-fed biting Diptera, digestion of the blood meal may occur in the time between initial feeding and successful trapping. However, the placement of traps near potential sources of blood meals to limit digestion time may be unadvisable, depending on the overall aims of the trapping, as this will bias any host selection data [3]. These findings are relevant to future work that aims to investigate the biting behaviour and disease transmission potential of *Culicoides* (and likely other small haematophagous Diptera) through blood meal analysis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13071-022-05607-x.

Additional file 1: Text S1. Primer protocol selection method.

Acknowledgements
We thank John Connolly, John Mumford and Silke Fuchs for their helpful comments and suggestions. We also thank Samantha O'Loughlin for providing blood-fed *Anopheles* samples. Additionally, we also thank Robbie Bellekom for the design and creation of the graphical abstract. Finally, we thank the two anonymous reviewers for the comments and suggestions.

Author contributions
BB: conceptualisation, data analysis and writing, AB: review, technical advice, and guidance, ME: midge rearing, review, and editing, ZL: midge rearing, OTL and TDH: supervision, review, and editing. All authors read approved the final manuscript.

Funding
This work was supported by a Natural Environment Research Council studentship to B.B. though the NERC-Oxford University Environmental Research Doctoral Training Partnership (NE/L002612/1). O.T.L., A.B and T.D.H. were supported by Target Malaria. Target Malaria is supported by a grant from the Bill and Melinda Gates Foundation and the Open Philanthropy Project Fund, an advised fund of Silicon Valley Community Foundation.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors contributed to and gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests
We declare we have no competing interests.

Received: 6 October 2022 Accepted: 2 December 2022
Published online: 13 January 2023

References

1. Reeves LE, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. PLoS Negl Trop Dis. 2018;12:e0006767. https://doi.org/10.1371/journal.pntd.0006767.
2. Martínez-de la Puente J, Ruiz S, Soriguer R, Figuerola J. Effect of blood meal digestion and DNA extraction protocol on the success of blood meal source determination in the malaria vector *Anopheles atroparvus*. Malar J. 2013;12:109. https://doi.org/10.1186/1475-2875-12-109.
3. Bellekom B, Hackett TD, Lewis OT. A network perspective on the vectoring of human disease. Trends Parasitol. 2021. https://doi.org/10.1016/j.pt.2020.12.001.
4. Nielsen JM, Clare EL, Hayden B, Brett MT, Kratina P. Diet tracing in ecology: method comparison and selection. Methods Ecol Evol. 2018;9:278–91. https://doi.org/10.1111/2041-210X.12869.
5. Reeves LE, Holderman CJ, Gillett-Kaufman JL, Kawahara AY, Kaufman PE. Maintenance of host DNA integrity in field-preserved mosquito (Diptera: *Culicidae*) blood meals for identification by DNA barcoding. Parasit Vectors. 2016;9:503. https://doi.org/10.1186/s13071-016-1791-z.
6. Ngo KA, Kramer LD. Identification of mosquito bloodmeals using polymerase chain reaction (PCR) with order-specific primers. J Med Entomol. 2003;40:215–22. https://doi.org/10.1603/0022-2585-40.2.215.
7. Oshaghi MA, Chavshin AR, Vatandoost H, Yaaghoobi F, Mohtarami F, Noorjah N. Effects of post-ingestion and physical conditions on PCR amplification of host blood meal DNA in mosquitoes. Exp Parasitol. 2006;112:232–6. https://doi.org/10.1016/J.EXPPARA.2005.11.008.
8. Mukabana WR, Takken W, Seda P, Killeen GF, Hawley WA, Knols BGJ. Extent of digestion affects the success of amplifying human DNA from blood meals of Anopheles gambiae (Diptera: *Culicidae*). Bull Entomol Res. 2002;92:233–9. https://doi.org/10.1079/BER2002164.
9. Seutin G, White BN, Boag PT. Preservation of avian blood and tissue samples for DNA analyses. Can J Zool. 2011;69:82–90. https://doi.org/10.1139/Z91-013.
10. Lassen S, Nielsen SA, Kristensen M. Identity and diversity of blood meal hosts of biting midges (Diptera: Ceratopogonidae: *Culicoides* Latreille) in Denmark. Parasit Vectors. 2012;5:143. https://doi.org/10.1186/1756-3305-5-143.
11. Valinsky L, Ettinger G, Bar-Gal GK, Orshan L. Molecular identification of bloodmeals from sand flies and mosquitoes collected in Israel. J Med Entomol. 2014;51:678–85. https://doi.org/10.1603/me13125.
12. Orsborne J, et al. Investigating the blood-host plasticity and dispersal of Anopheles coluzzii using a novel field-based methodology. Parasit Vectors. 2019;12:143. https://doi.org/10.1186/s13071-019-3401-3.
13. Hernandez-Colina A, et al. 2021 Blood-feeding ecology of mosquitoes in two zoological gardens in the United Kingdom. Parasites Vectors. 2021;141:1–18. https://doi.org/10.1186/S13071-021-04735-0.
14. Navia-Gine WG, Loaiza JR, Miller MJ. Mosquito-host interactions during and after an outbreak of equine viral encephalitis in Eastern Panama. PLoS ONE. 2013;8:e81788. https://doi.org/10.1371/journal.pone.0081788.
15. Travassos Da Rosa JF, De Souza WM, De Paula PF, Figueiredo ML, Cardoso JF, Acrani GO, et al. Oropouche virus: clinical, epidemiological, and molecular aspects of a neglected *Orthobunyavirus*. Am J Trop Med Hyg. 2017;96:1019–30. https://doi.org/10.4269/AJTMH.16-0672.
16. Boorman J. The maintenance of laboratory colonies of *Culicoides* variipennis (Coq.), C. nubeculosus (Mg.) and C. riethi Kieff. (Diptera, Ceratopogonidae). Bull Entomol Res. 1974;64:371–7. https://doi.org/10.1017/S0007485300031254.
17. Rozo-Lopez P, Park Y, Drolet BS. Effect of constant temperatures on *Culicoides* sonorensis midge physiology and vesicular stomatitis virus infection. Insects. 2022. https://doi.org/10.3390/INSECTS13040372.
18. Logue K, Keven JB, Cannon MV, Reimer L, Siba P, Walker ED, et al. Unbiased Characterization of Anopheles Mosquito Blood Meals by Targeted High-Throughput Sequencing. PLoS Negl Trop Dis. 2016;10:e0004512. https://doi.org/10.1371/journal.pntd.0004512.
19. Taylor PG. Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. Mol Biol Evol. 1996;13:283–5. https://doi.org/10.1093/oxfordjournals.molbev.a025566.
20. Wickham H. ggplot2 elegant graphics for data analysis. New York: Springer-Verlag; 2016.
21. Lehiy CJ, Drolet BS. The salivary secretome of the biting midge *Culicoides*. sonorensis PeerJ. 2014. https://doi.org/10.7717/PEERJ.426.

22. Noriega FG, Wells MA. Mini review a molecular view of trypsin synthesis in the midgut of *Aedes* aegypti. J Insect Physiol. 1999;45:613–20.
23. Filimonova SA. Morphological study of digestive cycle in bloodsucking biting midges of genus *Culicoides*. Transl Zhurnal Evolyutsionnoi Biokhimii Fiziol. 2005;41:176–85.
24. England ME, et al. *Culicoides* species composition and molecular identification of host blood meals at two zoos in the UK. Parasit Vectors. 2020;13:139. https://doi.org/10.1186/s13071-020-04018-0.
25. Srisawat R, Sungvornyothin S, Jacquet M, Komalamisra N. Preserving blood-fed aedes albopictus from field to laboratory for blood source determination. JItMM2013 Proc. 2014;3:29.
26. Ninio C, Augot D, Delecolle JC, Dufour B, Depaquit J. Contribution to the knowledge of *Culicoides* (Diptera: Ceratopogonidae) host preferences in France. Parasitol Res. 2011;108:657–63. https://doi.org/10.1007/s00436-010-2110-9.
27. Carpenter S, Groschup MH, Garros C, Felippe-Bauer ML, Purse BV. *Culicoides* biting midges, arboviruses and public health in Europe. Antiviral Res. 2013;100:102–13. https://doi.org/10.1016/J.ANTIVIRAL.2013.07.020.
28. Mignotte A et al. High dispersal capacity of Culicoides obsoletus (Diptera: Ceratopogonidae), vector of bluetongue and Schmallenberg viruses, revealed by landscape genetic analyses. Parasites and Vectors. 2021;14:1–14. https://doi.org/10.1186/s13071-020-04522-3. Accessed 2 Feb 2022.
29. Tomazatos A, Jöst H, Schulze J, Spînu M, Schmidt-Chanasit J, Cadar D, et al. Blood-meal analysis of *Culicoides* (Diptera: Ceratopogonidae) reveals a broad host range and new species records for Romania. Parasit Vectors. 2020;13:79. https://doi.org/10.1186/s13071-020-3938-1.
30. Kasičová Z, Schreiberová A, Kimáková A, Kočišová A. Blood meal analysis: host-feeding patterns of biting midges (Diptera, Ceratopogonidae, *Culicoides* Latreille) in Slovakia. Parasite. 2021. https://doi.org/10.1051/PARASITE/2021058.
31. O'Gower AK. The rate of digestion of human blood by certain species of mosquitoes. Aust J Biol Sci. 1956;9:125.
32. Filimonova SA. 2005 Morphological study of digestive cycle in Bloodsucking biting midges of genus Cu1icoides. J Evol Biochem Physiol. 2005;412:221–32. https://doi.org/10.1007/S10893-005-0057-8.
33. Lysyk TJ, Danyk T. Effect of temperature on life history parameters of Adult *Culicoides* sonorensis (Diptera: Ceratopogonidae) in relation to geographic origin and vectorial capacity for bluetongue virus. J Med Entomol. 2007;44:741–51. https://doi.org/10.1093/JMEDENT/44.5.741.

Publisher's Note