ABSTRACT

Title of Dissertation: URBANIZATION AND LANDSCAPE HETEROGENEITY INFLUENCE *CULEX* SPECIES ECOLOGY AND GENETICS IN EASTERN NORTH AMERICA

Arielle Arsenault-Benoit,
Doctor of Philosophy, 2023

Dissertation directed by: Dr. Megan Fritz,
Department of Entomology
University of Maryland

Vector-borne disease is an important facet of public health, as they account for nearly 20% of global disease burden. Multiple species, including at least one vector, at least one host, and a pathogen, must interact in vector-borne disease transmission cycles, and thus understanding human risk of vector-borne disease and public health outcomes requires a community ecology framework. Members of the *Culex* genus, including *Cx. pipiens*, *Cx. quinquefasciatus*, and *Cx. restuans* are sympatric in eastern North America and are vectors of West Nile virus. This dissertation explores the roles of habitat use, community ecology, phenology, and landscape heterogeneity on *Culex* spatiotemporal dynamics and genetics along urban to rural gradients in eastern North America.

Through surveillance of belowground structures in Washington, D.C. over two years, I found that mosquito species of public health importance, including *Aedes aegypti*, *Aedes albopictus,* and members of the *Culex pipiens* assemblage, use these structures for breeding and development. Belowground structures may serve as refugia against extreme climatic conditions and facilitate overwintering survival for non-diapausing taxa and/or taxa with thermal tolerance limitations, potentially expanding their suitable ranges.

On an urban to rural gradient in greater Washington, D.C. and Maryland, a complex of cryptic *Culex* WNv vectors co-occur on the landscape. Using molecular techniques and constrained ordination, I found that these cryptic *Culex* species were differently distributed at fine spatial scales, likely due to the impacts of urbanization on vector habitat and subsequent niche segregation. *Culex pipiens* were cosmopolitan and dominant across sites in greater Washington D.C. and Maryland. However, individuals with *Cx. quinquefasciatus* ancestry were limited to urban and peri-urban sites closest to the city center, and *Cx. restuans* were most abundant in rural and suburban sites furthest from the city center with dense and heterogeneous canopy cover. Previous work suggested that phenology has a considerable impact on *Culex* species dynamics; *Cx. restuans* was thought to be an early season species that cedes to *Cx. pipiens* over the course of the season. Initially, I did not detect an effect of season on *Culex* spatiotemporal dynamics when collections were undertaken from June through October, but when I expanded the collection season to include the months of April and May, the influence of season was evident. Therefore, the hallmark "crossing-over" point that is common in the *Culex* literature happens prior to the local mosquito abatement season in Washington and D.C. and Maryland. During the active surveillance and management period, season has little impact on

Culex species abundance as compared to environmental factors measured along our urban to rural gradient.

A replicated comparison of the abundance and relative frequency of *Cx. pipiens* and *Cx. restuans* along urbanization gradients in Washington D.C., greater Philadelphia, PA and greater Chicago, IL, using gradient forests demonstrated that phenology was consistently the most important predictor of the shift between a *Cx. restuans*-dominant community and a *Cx. pipiens*-dominant community. This crossing-over point trended later in the season with increasing latitude. Turnover in species abundance tended to occur at intermediate points along environmental gradients associated with urbanization, like percent impervious surface, percent tree cover, distance to city center, and vegetation index. Results of two analytical approaches (ordination and regression trees) and from three metropolitan areas support *Cx. restuans* as an early season species that is otherwise associated with sites with cooler temperatures, less impervious surface, more tree cover, a shallower water table, and increased distance from city center. Conversely, *Cx. pipiens* is more abundant than *Cx. restuans* in sites that are more characteristic of urbanization.

Culex pipiens is globally ubiquitous and was common across site classes in the three localities in this study. This species comprises two bioforms, *pipiens* and *molestus*, which are characterized by divergent ecological, physiological, and behavioral traits. These bioforms can interbreed in the field and the lab. However, at all sites analyzed across three northeastern metropolitan areas, analysis of genotypes at a single neutral locus violated assumptions of Hardy Weinberg Equilibrium, suggesting that there is not unrestricted geneflow between bioforms across the landscape. The proportion of *molestus* alleles increased with increasing percent impervious surface and decreased vegetation, two environmental correlates of urbanization.

Molestus alleles may confer an advantage in urban environments because they can leverage human infrastructure to overcome thermal limitations and persist in isolated belowground populations via autogeny and use of mammalian hosts.

Overall, *Culex* WNv vectors are differentially distributed across urban to rural gradients in the northeastern United States. These aspects are influenced by a heterogeneous land use and landscape-level changes associated with urbanization. A clear understanding of vector life history, genetics, interspecies interactions, and distribution across the landscape can improve practitioners' power and precision in predicting and managing vector borne disease transmission. While some patterns in species distribution and composition were universal across metropolitan areas, there was variation between localities that could significantly contribute to WNv transmission and human disease risk. Therefore, I conclude that modeling, as well as development of surveillance and management strategies for WNv vectors should be implemented locally to have the greatest impact on public health outcomes.

URBANIZATION AND LANDSCAPE HETEROGENEITY INFLUENCE *CULEX* SPECIES ECOLOGY AND GENETICS IN EASTERN NORTH AMERICA

by

Arielle Arsenault-Benoit

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Dr. Megan Fritz, Chair
Dr. Paul Leisnham
Dr. Maile Neel
Dr. Matt Fitzpatrick
Dr. Anahí Espíndola

Dedication

For my daughter, GGB.

Acknowledgements

I am unendingly grateful for the guidance, patience, insight, and mentorship of my advisor, Dr. Megan Fritz. You've taught me a lot about mosquitoes, but more important you've taught me to think, be a teacher, and be a scientist, all while maintaining balance. I am thankful for the members of the Fritz lab, especially Josh Yeroshevsky's help getting this project off the ground in the early days, and Dr. Kate Bell and Dr. Kate Taylor for endless advice, technical help, analysis and writing assistance, and optimism and reassurance that I would eventually graduate on the days when I wasn't so sure.

A giant thank-you to my dissertation committee for thoughtful critiques and questions, for encouraging me to push myself, and for always making my work better.

This dissertation was a huge collaborative effort covering many years and thousands of miles. Thank you to Dr. Patrick Irwin and his team at NWMAD for years of specimen collection, offering friendly reviews on my writing, and inviting me to share my work more broadly. Thank you to Dr. Keith Price and folks at the PA DEP for continuing to collect and store mosquitoes for us, especially when COVID made it even tougher, Dr. Al Greene for giving us a tour of dark, dirty, and smelly D.C. basements, and all of those who granted us site access: Arnaud Martin, Alex and Annie Kersey, Patuxent Research Refuge (USFWS), Smithsonian Research and Education Center (SERC), the Smithsonian Institution, and University of Maryland Research and Education Centers at Beltsville, Upper Marlboro, and Clarksville. My graduate studies and these chapters were funded by NIH R01AI125622A to MLF.

This work could not have been completed without the hard work and dedication of the undergraduate researchers who contributed to mosquito collection and identification over the past five years, including Linette Kingston, Jamia Mason, Lily Sheehan, Josh Levy, Helen Craig, and Casey Lanham. Christine Johnson was particularly invaluable for the success of this dissertation due to her years of field work, especially during the summer of 2020. Sommer Steven's pilot work made Ch. 4 possible. Two of these students were funded by the AGNR SROP/SOARE program.

I have had such a positive experience in the UMD Entomology department, and I need to thank faculty, staff, and students for making the day-to-day run smoothly, for developing a culture of open communication where students can succeed, and for always striving to make the department better.

Now, to my personal life. Thank you to my husband Patrick for years of support, listening, and enthusiasm. And to my daughter Genevieve for being my reason to leave the lab bench or computer every day so we can joyfully figure out this world together. Thank you to my parents for a lifetime of support and understanding, and more recently the countless 12-hour drives over the past three years to help me piece together this student-parent life. And to KB and LY for childcare- possibly the most valuable reagent expended in the development of this dissertation.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Table 1. Counts and percentages of *Cx. pipiens* genotype frequencies by locality and site class at the CQ11 locus. Hardy Weinberg Equilibrium (HWE) statistics were calculated for each population, and all were out of HWE according to our statistical significance threshold of $\alpha = 0.05$.

List of Figures

Chapter 2

Chapter 3

Figure 3. CCA tri-plots and variance partitioning for June-October, 2019-2021. Site names have been suppressed to improve readability. Two tri-plots are presented due to collinearity between predictors. Seasonal ordiellipses projected over CCA tri-plots show how responses cluster in ordination space. The centroid of each ellipse is designated by a color-coordinated point. A. CCA tri-plot for model 1: Primary Site Class+ Distance to City Center (DCC) + Latitude (Y). B. CCA tri-plot for model 2: Secondary Site Class + Percent Impervious Surface +Distance to City Center (DCC) +Latitude (Y). C. Variation in the 2021 community assemblage represented as the percent explained by season, geographic space (latitude and longitude), environment. Values less than zero (<0) are not shown.

Figure 4. Summary of April-October 2021 *Culex* molecular identifications. Approximately n=10 individuals per collection event per year were molecularly identified for this analysis. A: Composition of *Culex* by primary site class is represented as mean percent of specimens (% ± SE) identified. B: Mean species counts identified over the course of the season, aggregated across years. The start of each 2-week collection window is indicated by week of the year (WOY). C: Mean relative frequency of species identified in early, mid, and late season, separated by site class.

Figure 5. CCA tri-plots and variance partitioning for May-October 2021. Site names have been suppressed for readability. Two tri-plots are presented due to collinearity between predictors. Seasonal ordiellipses projected over CCA tri-plots show how responses cluster in ordination space. The centroid of each seasonal ellipse is designated by a color-coordinated point. A. CCA tri-plot for model 1: Primary Site Class+ Distance to City Center (DCC) + Latitude (Y). B. CCA tri-plot for model 2: Secondary Site Class + Percent Impervious Surface +Distance to City Center (DCC) +Latitude (Y). C. Variation in the 2021 community assemblage represented as the percent explained by season, geographic space (latitude and longitude), environment. Values less than zero (<0) are not shown.

Chapter 4

Figure 1. Collection sites in three metropolitan localities in the eastern United States: Chicago, Illinois (upper left), Philadelphia, Pennsylvania (upper right), and Washington, District of Columbia, including Maryland suburbs (lower right). Rural sites are in blue, suburban sites are in purple, and urban sites are in orange.

Figure 2. Summary of mean species counts (±S.E.M.) per trap by site class (left) and season (right) represented from week of the year (WOY) 15-40 and separated by locality. Localities are arranged from southernmost (MD) to northernmost (IL), and aggregate data is in the bottommost panel. In each panel, *Cx. pipiens* are in blue, *Cx. restuans* are in gold, and *Cx. quinquefasciatus* are in black.

Figure 3. Ranked conditional importance of predictor variables on composition of *Cx. pipiens* and *Cx. restuans* for all localities aggregated. A. Importance of each predictor to model accuracy. B. Relative importance of each predictor variable weighted by species. Detailed descriptions of predictors are in Appendix II, Table 4.2.

Figure 4. Gradient Forest (GF) results for full dataset aggregated across localities. A. Location, density, and importance of GF splits along gradients of the top-ranked 16 environmental predictors. Density of split points are in black, and density of observations are in red. The ratio of these densities are in blue. Locations where the ratio of densities is greater than 1 (blue dotted line) indicates regions along the predictor gradient with greatest relative change in species assemblage. B. Cumulative distributions of standardized splits importance for each species. Monotonic functions for *Cx. restuans* are in teal and *Cx. pipiens* are in red. Shallow slopes indicate low rates of change for each species along the gradient, and steep slopes indicate higher rates of change.

Figure 5. Species abundances in response to a subset of highly ranked environmental variables. *Culex pipiens* abundances are on the left, and *Cx. restuans* abundances are on the right. Points represent species abundances across each environmental gradient. Trends are shown using local regression with locally estimated scatterplot smoothing (*loess*) and 95% confidence interval. A-B. Species trends along week of the year (WOY). C-D. Species trends in response to mean daily maximum temperatures, 1-7 days prior to collection date. E-F. Species trends along the percent impervious surface gradient. G-H. Species trends along the percent tree cover gradient.

Chapter 5

Figure 1. Collection sites in three metropolitan localities in the eastern United States: Chicago, Illinois, Philadelphia, Pennsylvania, and Washington, District of Columbia, including Maryland suburbs. Rural sites are in blue, suburban sites are in purple, and urban sites are in orange.

Figure 2. Counts of *pipiens* homozygotes (black), *molestus* homozygotes (white), and heterozygotes (patterned) based on genotyping of the CQ11 locus. Panels represent each collection locality: A. Maryland and Washington D.C. (N=271), B. NWMAD in Illinois (N=183), C. Greater Philadelphia, PA (N=113). Panel D shows aggregate values of all sites combined (N=567).

Figure 3. Mean ± standard error of proportion *molestus* alleles in rural, suburban, and urban populations in metropolitan Chicago, Illinois (IL), Washington D.C. and Maryland (MD) and Philadelphia, Pennsylvania (PA).

Figure 4. Linear trends for *molestus* allele frequencies by trap with A. increasing impervious surface and B. vegetation index (NDVI), for each locality. Only NDVI values greater than 50 were included. Shaded regions represent 95% confidence intervals around the lines of best fit.

List of Abbreviations

ANOVA	Analysis of Variance
CCA	Constrained Correspondence Analysis
DCC	Distance to (focal) city center [Washington D.C., Philadelphia, Chicago]
GF	Gradient Forest
GLM	Generalized Linear Model
HWE	Hardy Weinberg Equilibrium
IL	Locality characterizing collection sites in greater Chicago, IL
MD	Locality characterizing collection sites in Washington D.C. and Maryland
MPIS	Mean Percent Impervious Surface
MSO	Multiscale Ordination
NDVI	Normalized Difference Vegetation Index
NLCD	National Landcover Database
NMDS	Nonmetric multidimensional scaling
NWMAD	Northwest Mosquito Abatement District
s.l.	*Sensu lato*
PA	Locality characterizing collection sites in greater Philadelphia, PA
PA DEP	Pennsylvania Department of Environmental Protection
PCR	Polymerase Chain Reaction
WNv	West Nile virus
WOY	Epidemiological Week of the Year

Chapter 1: *Life history of cryptic Culex vectors of West Nile virus in Eastern North America*

Ecology of vector-borne disease transmission

Human land use changes influence both the composition of species and their interactions within an ecosystem (Grimm et al. 2008). Urbanization, the process by which a location becomes densely populated and developed over time (Tisdale 1942, Johnson and Munshi-South 2017), is one of the major ways humans modify the landscape (McDonnell et al. 1997, Palumbi 2001, McDonnell and Hahs 2008). Although the human activities associated with urbanization are somewhat predictable, including construction of housing, transportation, and community infrastructure, the biotic consequences of these changes are not as consistent or predictable. Human-made environmental changes are unique to each location (Combs et al. 2018), promote heterogeneity, and impact individual species, the landscape, and the nature of their interactions. My dissertation aims to understand how urbanization impacts the composition of mosquito vectors of human and animal diseases in an ecosystem.

Globally, up to 17% of human infectious diseases are vector-borne (Rinker et al., 2016), including those transmitted by mosquitoes. Mosquito-borne zoonoses, or diseases that are transmitted between animals and humans by mosquitoes, require at least three organisms: the pathogen, the vector, and the host (Johnson et al. 2015a, Karesh et al. 2012). Some vectored pathogens have complex life cycles involving multiple vertebrate hosts. In addition, a community of mosquitoes may vector a pathogen, but each vary in their capacity for transmission (Ferraguti et al. 2016). This

makes understanding the ecology of vector-borne disease critical to preventing transmission (Karesh et al. 2012, Gomes et al. 2013). More specifically, knowledge of vector interactions with potential habitat (Kilpatrick et al. 2008, Norris et al. 2004), as well as interactions among the vectored pathogen, vertebrate hosts, and the vectors themselves can be exploited to develop approaches that prevent mosquito-borne illness.

Mosquito taxa are both ubiquitous and diverse (Ferraguti et al. 2016), and different species often have distinct requirements for ovipositional sites and vertebrate blood hosts. In general, females oviposit eggs into standing water, where they develop through the larval and pupal stages. Upon eclosion, adults mate and females must then seek out blood hosts to obtain proteins and minerals required for oogenesis. Mosquito niche requirements therefore include standing water for oviposition, suitable habitat for their primary blood hosts, as well as resting locations, and refugia for diapause and/or overwintering in temperate latitudes. Optimal conditions vary by species, allowing mosquitoes to thrive across heterogeneous environments.

Because of their diverse niche requirements, alteration of the environment can expand viable habitat for certain mosquito species and destroy habitat for others. Urbanization generally lowers overall biodiversity (Crispo et al. 2011), including the species richness and abundance of mosquitoes (Ferraguti et al. 2016, Goodman et al. 2018) due to environmental homogenization and reduction of green space. Human activities can create new resources for oviposition, however, including irrigation, impervious surfaces, containers, runoff, sewage, or below ground infrastructure (Norris et al. 2004, Townroe and Callaghan 2014). Urbanization can also help mosquitoes overcome thermal tolerance thresholds by generating novel habitats like insulated

structures, or by creating suitable overwintering conditions within an urban heat island. Some species in the genera *Culex, Aedes,* and *Anopheles* thrive in urban environments because they can exploit novel human-generated habitats (LaDeau et al. 2015, Ferraguti et al. 2016, Medeiros-Sousa et al. 2017). These so-called anthropophilic species tend to contribute *most directly* to human disease and therefore are most heavily scrutinized with respect to their disease transmission potential to humans (Brady et al. 2016).

Vectorial capacity

Vectorial capacity is a metric used to quantify the potential impacts that each vector population has on disease transmission to humans within an ecosystem. By definition, vectorial capacity is the potential number of new infectious bites per human per day by a vector population (Garrett-Jones 1964a). Its mathematical formula uses aspects of vector physiology and life history (e.g. vector competence, or the physiological ability of certain species to become infected with and transmit a particular pathogen, human biting rates and behavior, and extrinsic incubation time) to predict or describe the potential for infectious bites under ideal conditions (Brady et al. 2016, Feraguitti et al. 2016, Kramer and Ciota 2015). Human biting rate is an important predictor of vectorial capacity (Garrett-Jones 1964b).

Species that bite humans infrequently have lower vectorial capacity, but enzootic vectors of a pathogen can still act critically to propagate a disease transmission cycle. The ecology of dengue virus and yellow fever are illustrative examples, as these diseases have both an enzootic sylvatic cycle where mosquitoes primarily infect non-human primates in natural habitats, and an epizootic urban cycle where humans are infected in populated areas by anthropophilic species. Each cycle relies on mosquito

species that are associated with each habitat type (Figueiredo 2019), but the vectorial capacity of the enzootic vectors would be quite low compared to that of the epizootic vectors due to significant differences in human biting rates across taxa. As areas become more urbanized, the risk of spillover of a primarily zoonotic pathogen into human host populations increases due to human encroachment into natural areas and creates concern for public health. Perturbations in the arboviral cycles or changes to the landscape can have downstream impacts on the entire system leading to increased unpredictability of disease risk over space and time (Bowdon et al. 2011, Gorris et al. 2021). These points illustrate the importance of environment and ecological relationships in vector borne disease, and emphasize the importance of integrating ecological variables, including relative abundance of competent species and environmental characteristics into disease transmission models (Kilpatrick et al. 2008, Townroe and Callaghan 2014, Johnson et al. 2015b).

Ecology of West Nile virus transmission in North America

West Nile virus (WNv) is one such disease with an ecologically complex transmission cycle. It is widespread across much of the globe, including Africa, Southern Asia, Northern Australia, and is intermittently present in other regions (Kilpatrick et al. 2011). West Nile virus is zoonotic, and cycles enzootically between bird-biting mosquitoes and avian amplifying hosts. Humans are exposed when a mosquito first feeds on an infected bird, then later feeds on a human, resulting in epizootic transmission, or "spillover." Mosquitoes that connect the enzootic cycle to dead-end mammalian hosts, including humans, are called "bridge vectors." For human

cases to occur, enzootic vector species and bridge vector species, or a single vector species that serves as both must be present in the environment (Brown et al. 2008).

West Nile virus was first introduced to the United States in 1999 and quickly infected local mosquito vectors, some of which were novel vectors for WNv (Kilpatrick et al. 2011). Species contributing to WNv transmission in the USA, include but are not limited to *Cx. pipiens*, *Cx. restuans, Cx. quinquefasciatus.* The importance of each species to transmission depends on geography, abundance (and relative abundance) of each species, local conditions, and other factors (Kilpatrick et al. 2005, Tokarz and Smith 2019). For instance, between 1999-2010 in the Northeast United States, WNv was isolated from 33 species in 8 genera, predominated by *Culex species* (96%) including *Cx. pipiens* (66.0%) and *Cx. restuans* (27.8%). In the Southeast and Southwest, the majority of WNv was isolated from *Cx. quinquefasciatus* (64.6% and 82.1% respectively, Andreadis 2012 and CDC Arbonet). Further, in the mid-Atlantic region, it is suspected that 80% of human cases are caused by *Cx. pipiens* or *Cx. restuans* (Andreadis 2012). Members of the *Culex* genus are recognized as important vectors of WNv due to high vector competence, high infection rate in the field, and local abundance in areas of high case incidence (Fonseca et al. 2004, Andreadis 2012). *Culex* species likely also play a role in WNv overwintering (Jackson and Paulson 2006, Andreadis 2012), and a community assemblage of co-occurring *Culex* species may increase and extend WNv transmission across a season due to differences in phenology, habitat use, and host preferences (Reiskind and Wilson 2008). *Culex* species communities can amplify WNv across the season, connect the pathogen to hosts across a heterogeneous landscape, and act as both enzootic and bridge vectors.

Worldwide, *Cx. pipiens* is a major arboviral vector for a number of pathogens of both animals and humans (Andreadis 2012, Gomes et al. 2013). It was identified as a primary epidemic bridge vector to humans after a foundational study documented WNv-positive *Cx. pipiens* with mixed human and avian bloodmeals (Hamer et al. 2009). *Culex pipiens* comprises two bioforms, *Culex pipiens* form *pipiens* and *Culex pipiens* form *molestus,* which share a geographic distribution, are morphologically indistinguishable, and can interbreed, but demonstrate ecological, behavioral, and physiological differences that may contribute to spatial and genetic isolation between the forms (Chevillon et al. 1995, Fonseca et al. 2004). *Culex pipiens* f. *pipiens* primarily use avian hosts, require a bloodmeal for the first gonotrophic cycle, breed in above-ground aquatic water sources, and can use cosmopolitan habitats (Bahnck and Fonseca 2006, Di Luca et al. 2016). Conversely, *Cx. pipiens* f. *molestus* prefer mammalian hosts, do not require a bloodmeal prior to laying the first egg raft, and are often found using isolated belowground structures and human-dominated areas (Byrne and Nichols 1999, Fonseca et al. 2004, Fritz et al. 2015).

Culex pipiens mosquitoes are not native to North America but likely became established during European colonization centuries ago (Reiskind and Wilson 2008), allowing for US populations to undergo local adaptation and novel gene flow patterns. Studies suggest that in most of their range, *pipiens* and *molestus* are genetically divergent but populations in the United States form an independent cluster, with up to 40% of individuals demonstrating hybrid ancestry (Fonseca et al. 2004, Yurchenko et al. 2020). Yet, the frequency of contemporary hybridization is uncertain. Hybridization between bioforms has been shown to cause indiscriminate host use (Fritz et al. 2015),

creating a potential “bridge” between birds and mammals, including humans (Gomes et al. 2009, Farajollahi et al. 2011, Gomes et al. 2013). Switching between the use of birds and mammals as blood hosts causes epizootic transmission, or spillover, of zoonoses, including WNv to human populations. Landscape heterogeneity across an ecological gradient from urban to rural regions in a temperate metropolitan area could play a role in the differentiation or isolation of bioforms due to niche segregation (Chevillon et al. 1995). Allele frequency distributions and hybridization rates across environmental gradients may influence the risk of human WNv cases across the landscape.

Congeners of *Cx. pipiens* can also contribute significantly to WNv, including its sister species, *Cx. quinquefasciatus* (Gomes et al. 2009). As members of the same species complex, these species can hybridize, and hybrid zones occur worldwide (Fonseca et al. 2009). *Culex quinquefasciatus* is primarily found in tropical and subtropical regions with a well-established hybrid zone in more temperate regions, generally between 36° and 39° latitude (Barr 1957, Sanogo et al. 2008), but there are records of *Cx. quinquefasciatus* ancestry as far north as 42° latitude in the United States (Kothera et al. 2020). *Culex quinquefasciatus* is common in the southeastern United States and is a competent vector of WNv (Andreadis 2012), but there is some uncertainty about the importance of this species in the WNv system in temperate North America. To transmit WNv to humans, they would need to feed on humans, which was once thought to be rare (Zinser et al. 2004). However, *Cx. quinquefasciatus* feeding may be highly opportunistic, plastic, and dependent on the geographic origin of the *Cx.*

quinquefasciatus population or the local abundance of certain hosts (Zinser et al. 2004, Takken and Verhulst 2013). Humans might act as blood hosts in urban areas.

Culex restuans is endemic to North America (Johnson et al. 2015b), co-occurs with *Cx. pipiens* (Harrington and Poulson 2008), and has similar WNv vector competence to that of *Cx. pipiens* (Ebel et al. 2005). WNv is regularly isolated from field-collected *Cx. restuans* (Jackson and Paulson 2006, Jackson et al. 2005, Andreadis 2012), highlighting their critical role in the WNv transmission cycle. *Culex restuans* may be as important, if not more important, than *Cx. pipiens* for WNv amplification and transmission throughout the season because they are considered to be ornithophilic (Johnson et al 2015b), strongly contributing to the enzootic WNv amplification cycle. Interestingly, *Cx. restuans* is less often described as a WNv bridge vector due to their ornithophilic blood feeding tendencies (Kilpatrick 2005, Hamer et al. 2008). Yet *Cx. restuans* can, and do, engage in mammalian blood feeding, including on humans (Ebel et al. 2005, Hamer et al. 2008). It has been suggested that their impact to human health is indirect, because they amplify WNv infections in avian hosts, but rarely act as epizootic vectors (Kilpatrick 2005, Hamer et al. 2008, Harrington and Poulson 2008). Humans are dead-end hosts for WNv, meaning human to human transmission does not occur. Rather, the virus must spillover from infected birds through mosquito bites, so prolonged transmission and amplification, in part by *Cx. restuans*, within the avian population are consequential to the risk of human disease (Andreadis et al. 2001, Andreadis 2012).

This contradiction underscores the need to clarify the role of *Cx. restuans* in WNv transmission across the landscape. *Cx. restuans* tend to be dominant in natural

areas, but studies show they can exploit disturbed and human-dominated habitats (Ebel et al. 2005, Johnson et al. 2015b) and sometimes show substantial niche overlap with *Cx. pipiens* (Reiskind and Wilson 2008). Unique from *Cx. pipiens*, however, is the phenology of *Cx. restuans*. In the northeastern United States, *Cx. restuans* is considered an early-season mosquito, emerging sooner and declining before *Cx. pipiens* reaches peak abundance (Lee and Rowley 2000, Andreadis et al. 2001, Johnson et al. 2015b). A number of surveys of *Culex* vectors demonstrate that *Cx. restuans* were abundant earlier in the season than *Cx. pipiens* (Lee and Rowley 2000, Andreadis et al. 2001, Jackson and Paulson 2006). It is thought that species is important for early season WNv amplification within avian populations, which allows for increased transmission to mammals when *Cx. pipiens* populations peak (Andreadis et al. 2001, Johnson et al. 2015b). Morphological similarities between *Cx. pipiens, Cx. restuans,* and other cryptic taxa create challenges for accurate vector surveillance, predictive disease modeling, and research into species' life histories and roles in WNv transmission (Harrington and Poulson 2008, Johnson et al. 2015b).

Conclusion

Considerable labor and financial resources are dedicated to surveilling for and treating potential WNv hotspots in United States metropolitan areas, yet cryptic *Culex* species are often treated as a single taxon for surveillance and management because they are morphologically similar and are difficult to distinguish via microscopy. Although these species have overlapping geographic ranges and can co-occur in space and time, they also have different optimal niches: they use different hosts, ovipositional

resources, and refugia, and thus distribute themselves differently over space and time. Within *Cx. pipiens*, phenotypic differences may create broad sympatry but microhabitat spatial segregation between bioforms analogous to that we see between species.

Urbanization creates a complex ecological matrix with heterogeneous habitat across space. The heterogeneous nature of the landscape offers a variety of habitat types, hosts, and biotic and abiotic features, that influence how both the *species within a community* and the *alleles within populations* are distributed through space and time (Bierne et al. 2013, Vinatier et al. 2011). Linking ecological processes to spatial patterns is the basis of landscape ecology and offers insights into community dynamics, species interactions, and population genetics (Vinatier et al. 2011).

A community of co-occurring *Culex* species may increase and extend WNv transmission across a season due to differences in phenology, habitat use, and host preferences and the complement of important vectors likely differs geographically (McMillan et al. 2020, DeGroote et al. 2008, Reiskind and Wilson 2008, Tokarz and Smith 2019). The abundance and seasonality of members of the *Culex* species assemblage likely differ over space, time, and scale, leading to spatiotemporally dynamic changes in the network of vectors and the resulting disease burden. It remains unclear how seasonality, climate, and landscape factors, including characteristics of urbanization, mediate *Culex* community assemblages and ultimately their role in the WNv transmission cycle. A better understanding of the dynamics of the vector community and interactions with the environment across a heterogeneous landscape could aid us in describing the potential impact of ecological and genetic forces on WNv

risk. The role of *Culex* species in WNv transmission makes them an important subject for disease transmission research; additionally, this group is an optimal case study, or a "living laboratory," for the influence of the landscape on concepts of habitat use, community ecology, and hybridization on an urban to rural gradient.

Dissertation Aims

In this dissertation I explore the influence of urbanization on *Culex* species population dynamics by:

Aim 1: *Examining use of below ground structures by mosquitoes in Washington, D.C. and describing the conditions that make these structures suitable habitat for oviposition and refugia.*

Aim 2: *Identifying features of the landscape and seasonal trends that drive the* Culex *community composition on an urban to rural gradient in greater Washington, D.C.,*

Aim 3: *Extending the results of Aim 2 to other Northeastern metropolitan areas: Chicago, IL and Philadelphia, PA, to describe how compositional turnover of* Culex *assemblages occurs over environmental and climatic gradients.*

Aim 4: *Characterizing the influence of landscape on* Cx. pipiens *bioforms, hybridization rates, and allele frequencies along an urban to rural gradient in eastern North America.*

Chapter 2: *Paved Paradise: Below ground parking structures sustain urban mosquito populations in Washington, D. C.*

Abstract

After notification of mosquitoes within federal buildings in Washington, District of Columbia, we surveyed below ground levels of parking structures for mosquitoes and standing water in the summer months of 2018 and 2019. *Aedes aegypti* (Linnaeus), *Aedes albopictus* (Skuse), and members of the *Culex pipiens* Assemblage (Linnaeus) were found using below ground habitat as oviposition sites. Genotyping of *Culex* mosquitoes revealed *pipiens*, *molestus*, and *quinquefasciatus* ancestry, and allele frequency comparisons within and between collection years indicated a stable, resident population. Temperature measurements at one below ground level of a parking structure and above ground site in an adjacent plaza were made from December through February and April to June, 2018. Above ground temperatures ranged from -11°C to 35°C during this period, while below ground temperatures never dropped below 5°C or exceeded 30°C. Temperatures during December-February were significantly higher below ground compared to above ground. Moderated winter conditions suggest that below ground urban structures could act as refugia for warmer climate species, like *Ae. aegypti* and *Cx. quinquefasciatus*, allowing them to overcome assumed thermal barriers. Surveys of parking structures should be incorporated into an integrated vector management program in urban areas.

Introduction

Risk of vector-borne disease is of increasing concern in urban temperate regions (LaDeau et al. 2015), in part due to vector range expansion (Osland et al. 2021). In Washington D.C., for example, vectors of public health importance include but are not limited to *Culex pipiens s.l.* (L), a primary vector of West Nile virus in North America (Spielman 2001, Mutebi and Savage 2009, Wu et al. 2014), and *Aedes aegypti* (L), a vector of yellow fever virus, dengue virus, chikungunya, and Zika virus (Kay et al. 2000, Lima et al. 2016).

Environmental temperature influences the physiology, ecology, and behavior of mosquitoes, is one of the most important abiotic factors for their development and survival (Reinhold et al. 2018), and limitations on range expansion (Osland et al. 2021). Increased temperature leads to shorter generation times, more generations per season, a lengthened active season, faster digestion of bloodmeals and reduction of the extrinsic incubation period for pathogens (Townroe and Callaghan 2014, LaDeau et al. 2015). For some species, the relatively lower environmental temperatures in temperate regions serve as a barrier to their range expansion. Other species have evolved behavioral and physiological mechanisms that enable them to tolerate temperate latitudes, including physiological changes like the synthesis of cytoprotective proteins (Toxopeus and Sinclair 2018), overwintering diapause (Kaur Gill et al. 2017), and exploitation of protective habitats as refugia or hibernacula (Nasci eat al. 2001, Bugbee and Forte 2004, Norris 2004).

Winter temperatures were thought to limit the range of *Ae. aegypti* to latitudes lower than 33°N and regions that remain above 10°C (Gloria-Soria et al. 2018). *Culex quinquefasciatus* (Say) is a member of the *Culex pipiens* Assemblage commonly found

between 30° N and 30° S and has similar limited tolerance for low temperatures (Samy et al. 2016), yet both species are found to persist and recur in Washington D.C. (~38° N) (Lima et al. 2016, Farajollahi et al. 2011). Although urban Washington D.C. is found at a temperate latitude, the urban heat island effect elevates daily minimum temperatures in the city above what would be found in the surrounding natural areas (Kim 1992). Furthermore, there is an abundance of subterranean human-made structures that are permeable to vector species. Studies speculate that subterranean habitat is used for *Ae. aegypti* overwintering (Lima et al. 2016, Gloria-Soria et al. 2018), and this may be true for *Cx. quinquefasciatus*, as well. Mosquito control practitioners and researchers have known for decades that some important vector mosquito species use subterranean habitat for resting and oviposition. In Washington D.C., populations of *Cx. pipiens* form *molestus* have been found and exterminated from government buildings, where they used below ground water-collecting sumps for resting and oviposition (Albert Green, personal communication).

Urban environments can offer protective habitats when human-made structures are permeable and allow for mosquito entry. Urban temperate regions also provide an abundance of larval habitats. Anthropophilic mosquitoes, those which live in close association with humans, have adapted to exploit a variety of larval breeding sites within the urban built environment. They can tolerate smaller puddles and containers, dirty water or water with high organic matter content (Mattingly 1963, Irwin et al. 2008), and even take advantage of water collecting on impervious surfaces or in refuse (Becker et al. 2014). Human-made below ground structures such as stormwater conveyance pipes, basements, and crawl spaces are naturally insulated, covered, and

sometimes have systems to moderate temperature and humidity. Containers, drainage systems, sumps, and conduits, often subterranean, are prime locations for water to pool, creating potential breeding habitat. During the warmer months, these structures provide breeding sites like sumps, blocked drains, and puddles, as well as available vertebrate hosts for blood feeding, including rats, other commensal rodents, humans, and birds (Mallis 2011, Goodman et al. 2018). Therefore, a clear understanding of how these below ground structures promote mosquito persistence in urban environments, and documentation of the species likely to exploit them, is critical for public health efforts to reduce mosquito-borne disease transmission.

We were alerted to nuisance mosquitoes in a federal building and our investigation led to a wider survey of nearby below ground structures for breeding habitat, primarily parking garages. Here, we establish that some, but not all, of these structures offer larval habitat and adult refugia that could allow persistence of vector mosquito species more associated with warmer climates.

Methods

Site Surveys and Mosquito Collection

In spring of 2018, we were directed to indoor populations of mosquitoes on federal property in Southwest Washington D.C., then identified and surveyed accessible below ground structures within 1 km for standing water, adult mosquitoes, and juvenile forms. The third subterranean level(s) (*ca.* 11m below grade) of two below ground parking structures contained adult and juvenile mosquitoes and became our index sites. These are indicated as sites “A” and “B” (Figure 1, Table 1). We surveyed the lowest levels of eight unique parking structures and a dry moat surrounding a

federal building (*ca.* 2.5m below grade), for a total of nine sites, including visits to the two index sites on every date, over five dates between June and September 2018 (Figure 1, Table 1). Adult mosquitoes were collected by mechanical aspiration (BioQuip InsectaZooka, Item #2888A, BioQuip Products Inc. Rancho Dominguez, CA) of stairwells, walls, corners, and other potential resting sites. Any standing water was measured and surveyed for mosquito eggs, larvae, and pupae visually and by dipping with a standard 13 cm, 350 ml larvae dipper (BioQuip, Item #1132BQ, BioQuip Products Inc. Rancho Dominguez, CA). Individuals were collected via modified transfer pipette, clipped at the end to enlarge the opening to approximately ½ cm in diameter, or with the dipper. Where accessible, we dipped in below ground groundwater sumps for eggs, larvae, and pupae. When standing water was located but no mosquitoes were visible, we dipped into the water five to seven times before categorizing the water as "negative". The two index garages were also surveyed monthly between December 2018 and June 2019 for standing water but no mosquitoes were observed or collected during this time. In the summer of 2019, index sites were checked on a biweekly basis for standing water and immature mosquitoes, but collections were limited due to considerably drier conditions. During this time, standing water was rarely found in parking structures, except following one severe rainstorm that caused significant flooding, then subsided within three days. To determine whether gravid female mosquitoes were present in the garage even when standing water was absent, we deployed a CDC gravid trap (John W. Hock Co, Gainesville, FL) at site "A" biweekly from June through September. Additionally, on six dates between May 15,

2019 and September 9, 2019 we collected adults via mechanical aspiration both below ground and immediately above ground at the same index site.

All live insects collected were transported back to the University of Maryland, College Park, MD. Adults were sacrificed at 20°C and morphologically identified to species or species complex. Juveniles were reared as described in Fritz et al. (2015), and emerging adults were morphologically identified. Specimens were then retained at -80°C for future molecular analysis. Samples identified as cryptic *Culex* spp., including *Cx. pipiens* and *Cx. quinquefasciatus* were categorized as members of the *Culex pipiens* complex for purposes of measuring species richness and diversity indices. Mosquitoes that could not be identified due to damage (N=3) were excluded from measures of richness, diversity, and evenness.

To assess the genetic ancestry of the below ground *Cx. pipiens* complex population, we quantified the *Cx. pipiens* form *molestus, Cx. pipiens* form *pipiens,* and *Cx*. *quinquefasciatus* alleles present in the population. A random sample of 105 cryptic specimens *of the Cx. pipiens* complex that could not be identified to species morphologically from July and September of 2018 (30 from garage A, 43 from garage B, collected either as adults or as juveniles and reared to adult), plus 32 individuals collected in garage A in July and September 2019 were identified using molecular approaches. Genomic DNA was extracted with a Zymo Quick-DNA Miniprep Plus kit (Cat #11-397B, Zymo Research, Irvine, CA) according to standard protocol. We first confirmed that all specimens were *Cx. pipiens* Assemblage members*,* not *Cx. restuans,* or *Cx. salinarius* using a polymerase chain reaction (PCR) by Crabtree et al. (1995). Then a pair of PCRs amplified the ACE2 and CQ11 loci of these samples to identify

quinquefasciatus, *pipiens*, and *molestus* alleles following previously described protocols (Smith and Fonseca, 2004, Bahnck and Fonseca 2006) but with slight modifications described in Fritz et al. (2015). We used a χ^2 test with a simulated p-value based on 2000 replications of a Monte Carlo simulation to compare *Cx. pipiens* Assemblage allele frequencies across years, sites, and early to late season (July and September), where α=0.05. All statistical analyses were performed in R version 3.2.6 (R Foundation for Statistical Computing, Vienna, Austria).

Abiotic Conditions

To determine whether below ground parking structures offer refuge from extreme temperatures, we placed four temperature loggers (Onset HOBO 001-08, Onset Computer Corps, Bourne MA) dispersed through the third subterranean floor of parking garage A, and four loggers dispersed through the above ground courtyard adjacent to the parking structure. Three of the four loggers placed above ground in the courtyard were lost or removed during the study, so analysis includes data from only one above ground logger. To supplement our above ground temperature data, we retrieved daily mean temperature readings for our study dates from the NOAA weather station at Reagan International Airport (DCA), Arlington, VA (NOAA NCEI). Loggers recorded temperature every 20 minutes between December 20, 2018-February 27, 2019 and we re-deployed the single remaining above ground logger and below ground loggers 1 and 2 to record spring temperatures between April 1, 2019-June 18, 2019. To enable comparison between our logger data and the DCA NOAA weather station, we calculated daily mean temperatures for each logger, then completed a sliding window analysis of these daily means with a two week window

and a one week step using the evobiR package in R (v1.1, Blackmon and Adams 2015). We then used a set of linear models to determine whether the change in temperature over time from the sliding window analysis is better described by including or excluding above or below ground location in the model using analysis of variance (ANOVA). We also compared the mean, minimum, and maximum temperatures above and below ground, as well as percent of time and total time that temperatures remained below the 10°C threshold for *Ae. aegypti* overwintering, which also encompasses suspected *Cx. quinquefasciatus* overwintering tolerance. For this, we used a Welsh's two-sample t-test for unequal variances to directly compare 20 min recording intervals from the single above ground logger to the mean temperature of all below ground loggers for that same interval.

Results

We found mosquitoes in three of nine below ground sites surveyed in 2018, and on four of five dates (Table 1). This survey revealed six taxa using below ground structures as adult resting habitat in 2018, including 71 *Cx. pipiens*, 10 *Cx. salinarius,* 32 *Ae. aegypti,* 4 *Ae. vexans, 1 Anopheles spp.,* and 5 specimens in the *Aedes* genus that could not be morphologically identified, for a total of 123 adults. However, only members of the *Cx. pipiens* complex, *Ae. aegypti,* and *Ae. albopictus* were found to use these structures as ovipositional habitat. In 2019, 39 *Cx. pipiens* complex and 1 *Ae. albopictus* were collected below ground via mechanical aspiration across six dates in May through September. Two hundred sixty-six *Cx. pipiens* complex, 1 *Cx. erraticus*, 2 *Ae. vexans,* 4 *Ae. albopictus* and 2 *Ae. aegypti* adults collected throughout the 2019

season in biweekly gravid trapping at a single below ground site, for a two-year species richness of S=7. Specimens collected in gravid traps were not included in diversity index calculations, however, because the oviposition attractant used is designed for *Culex,* and thus biases collections (Reiter 1983). In 2018, standing water was observed covering variable surface areas and depths. Larvae were collected from puddles ranging from 0.3 to 3.5 cm deep, and from 0.5 to 3 m in diameter. Summer of 2019 was much drier than 2018. Standing water was observed once in one index site after a heavy storm but was dry upon re-check three days later and no juveniles were found.

To further investigate the *Culex* community within the garages, we identified the population-level genetic ancestry for 105 randomly selected cryptic *Culex* complex mosquitoes collected across years, dates, and locations using ACE2 and CQ11 locus genotypes. In total, genotyping revealed 29 individuals with *Cx. quinquefasciatus* ancestry, 30 *Cx. pipiens f. pipiens*, 37 *Cx. pipiens f. molestus* and 9 individuals with unresolved identity, meaning the results of the two assays were not in agreement, leading to uncertain ancestries for these individuals (Bahnck and Fonseca 2006). Unresolved individuals were excluded from further statistical analyses. We found no difference in genotype distributions between mosquitoes collected in garage A and garage B (χ^2=2.083, p=0.397), those collected in 2018 and 2019 (χ^2=0.031, p=1), and those collected in July and September (χ^2=0.245, p=0.901) (Table 2).

Temperature below ground never dropped below 5°C, and 17.9% of the logging period was spent below 10°C, compared to 41.8% above ground. Above ground temperatures also reached extreme highs (35°C) and lows (-11°C) between December and June, whereas the temperature range below ground was narrower (5.4°C to 29.5°C,

Figure 2). We also compared winter temperatures, limited to between December 20, 2018 and February 27, 2019. Compared to a mean winter temperature of 5.12°C immediately above ground, mean winter temperature below ground was 10.74°C, and below ground temperatures were significantly warmer (Welsh's two sample t-test, t=-69.874, df=6588.9, $p<0.001$). The more complex linear model including the above or below location parameter significantly improved model fit over the simpler model ($F=25.195$, df=1, $p<0.001$). This suggested that there were significant differences in the temperature changes over time between the above ground and below ground locations (Figure 2), and considering above ground and below ground locations separately better describes the variation in the data.

Discussion

A multi-species mosquito community of both adults and juveniles was documented in some, but not all, below ground structures surveyed in Washington D.C. in the summer of 2018. For the most productive garage surveyed in 2018 (site "A"), mosquito presence was also documented in 2019. The discovery of mosquitoes at index sites demonstrated that an absence of mosquitoes at other sites was not due to unsuitable regional conditions, but whether the site itself provided suitable conditions. This indicates the infrastructure of some subterranean garages have more habitat for resting and oviposition than others. Mosquitoes may arrive at below ground structures via wind, through ventilation fans, by vehicular movement, or they may seek them out. Over 1000 larvae were collected and identified in 2018, compared to just over 100 adult specimens, suggesting that oviposition in below ground standing water was not

incidental or rare. Species found as both adults and juveniles, namely, members of the *Cx. pipiens* complex, *Ae. aegypti* and *Ae. albopictus*, likely seek out and exploit these structures, compared to species found occasionally, whose presence may be incidental. Our study documented that mosquitoes also use these structures for breeding habitat. Standing water does not have to be deep or large to support oviposition. For example, the puddle from which 60 juveniles were collected on September 20, 2018 was approximately 2m x 1.2m, and 0.6 cm deep at its deepest point, but much of this puddle was less than 0.3 cm deep. A single collection from a single puddle yielded over 900 individuals, collected as eggs, larvae, and pupae (Table 1). In addition, immatures of both *Culex* and *Aedes* species and at different stages of development were collected together from the same puddles on multiple occasions. This suggests that gravid females share ovipositional sites. Mosquito larvae species can coexist within containers (Gilbert et al. 2008, Laporta and Sallum 2014), likely due to partitioning of water depth combined with morphological differences in feeding structures that allow for dispersal through the container's depth (Yee et al. 2004, Laporta and Sallum 2014). For example, *Culex* larvae have a siphon that allows them to feed deeper in the water table compared to *Aedes* species, while still breathing at the surface. However, the authors were struck by the coexistence of these larvae in a single puddle with a depth of less than 0.5 cm.

Culex pipiens Assemblage mosquitoes dominated our summer subterranean collections but use of below ground structures by *Culex* spp. is not localized to Washington D.C. (Mutebi and Savage 2009, Harbison et al. 2011). Instead, these habitats may provide opportunities for interbreeding among below ground *molestus* and above ground *pipiens* and *quinquefasciatus* members of the *Cx. pipiens* Assemblage

throughout their range. In Wuhan, China, a similar survey revealed ovipositional sites and juvenile mosquitoes, later identified as *Cx. quinquefasciatus* and *Cx. pipiens*. Rearing showed these *Cx. pipiens* to be both autogenous and stenogamous, indicating they may be *Cx. pipiens* form *molestus* (Wu et al. 2014). Genotyping results from the present study suggest a similar complement of Cx. pipiens Assemblage ancestries in below ground habitat, and we were able to rear a subset of the larvae we collected in culture for 3 generations without a blood meal, supporting our genetic findings that some, but not all, display the traits of *Cx. pipiens* form *molestus*.

Allele frequencies were consistent across the season (July-September) and between years, suggesting year-round resident *Culex* populations in the garages, rather than transient populations. Such genetic admixture fits expectations for North American *Culex.* Some *Cx. quinquefasciatus* ancestry in this population was expected, as Washington D.C. is well within the *Cx. pipiens*-*Cx. quinquefasciatus* hybrid zone, which is historically bounded between 36° and 39° latitude (Barr 1957, Sanogo et al. 2008). This zone may be expanding or not fully described, as individuals with *Cx. quinquefasciatus* ancestry have recently been documented as far north as the Chicago, IL area (~ 42° latitude, Samy et al. 2016, Kothera et al. 2020). Nine specimens could not be genotyped so they were categorized as "unresolved." Unresolved identities are likely due to local recombination events within the loci, causing disagreement between assays (Bahnck and Fonseca 2006). As a more southerly species, the conditions below ground may confer the same climatic benefit to *Cx. quinquefasciatus* or those with *Cx. quinquefasciatus* genetic admixture as for other cold-intolerant species. Beyond serving to buffer above ground temperatures, these below ground habitats may provide

opportunities for contact and interbreeding among the below and above ground forms of the *Cx. pipiens* Assemblage. Use of below ground structures as refugia for *Culex* spp. is not localized to Washington D.C., but a more widespread phenomenon in urban areas throughout their range.

Temperature recordings demonstrated that extreme fluctuations were moderated by below ground structures in Washington D.C. These urban structures were more likely to maintain temperatures within the suspected thermal tolerance ranges of the warmer climate species, *Cx. quinquefasciatus* (Morin and Comrie 2010) and *Ae. aegypti* (Rozeboom 1960), even in winter months. Daily temperature ranges below ground were much narrower than above ground, and this difference in variability could have an impact on mosquito fitness and vectorial capacity (Carrington et al. 2013). Extreme fluctuations in daily temperature also are known to negatively affect *Ae. aegypti* survival (Lambrechts et al. 2011), and because they feed and rest indoors (Dzul-Manzanilla et al. 2017), they are well adapted to exploit urban infrastructure when available. In a study of below ground stormwater systems, *Cx. quinquefasciatus* egg rafts were collected below ground even when adjacent above ground temperatures ranged from 2.8°C to 42.2°C (Harbison et al. 2011), so temperature moderation below ground may aid in survival beyond reported thermal tolerance ranges.

We collected both cryptic *Culex* with diagnostic *quinquefasciatus* genotypes, as well as *Ae. aegypti* adults and juveniles in below ground structures over several dates and locations through the summer. Previous reports limit the established population of *Ae. aegypti* to the Capitol Hill neighborhood of Washington, D.C. (Lima et al. 2016, Gloria-Soria et al. 2018), yet we found *Ae. aegypti* using subterranean larval habitat in

Southwest D.C (~1.5km from Capitol Hill). Additionally, we collected *Ae. aegypti* adults below ground in this survey, above ground immediately adjacent to these sites, and by gravid trap and mechanical aspiration in a broader geographic area including Foggy Bottom (4.7km), Columbia Heights (5.3km), Anacostia (4.0km), and on the National Mall (2.3km) in 2018, 2019, and 2020 (AAB, unpublished data). These observations further confirm that *Ae. aegypti* populations are well established and their local range has expanded beyond that of previous reports in Washington D.C. This is likely made possible by moderate temperatures in the below ground, urban environment, where temperatures are warmer during cooler periods, as the exploitation of suitable microhabitats is one of the primary pathways to pole-ward range expansion for organisms across taxa (Osland et al. 2021).

Utilization of below ground urban infrastructure may allow populations to persist for longer and build up earlier in the spring, plus warmer temperatures can shorten the development and generational time of individuals. These factors can extend the breeding season, increase population size, and potentially have implications for disease transmission. Populations of *Culex* spp. and *Aedes* spp. breeding in urban below ground structures may act as a source for above ground populations. For anthropophilic species that are particularly well adapted to human-dominated landscapes, urbanization may reduce isolation and increase connectivity among local populations (Crispo et al. 2011, Ferraguti et al., 2016), skewing the mosquito community toward those that vector human disease. Notably, the three taxa were found using subterranean garages as ovipositional habitats (*Cx. pipiens s.l., Ae. aegypti,* and *Ae. albopictus*) are of primary concern for potential to transmit disease and are known to be well adapted to human

dominated habitats. While we identified larvae and pupae in two parking structures in 2018, our 2019 results suggest they may not serve as reliable ovipositional sites every year. Consistent heavy rainfall in urban areas may allow these below ground structures to flood, creating partially concealed ovipositional habitats. The consistent presence of adult females across years and locations highlights the value of surveying and targeting these structures for vector control, however. A thorough integrated pest management approach for vectors and nuisance mosquito populations should include surveillance of subterranean habitats such as parking structures, tunnels, sumps, and subway systems; removal of standing water when possible and larvicidal treatments when necessary may significantly contribute to reduction of mosquitoes above ground. In years with persistent wet weather, pest control operators and local public health officials should incorporate surveillance of these structures into their programs, if possible.

Chapter 2: Tables

Table 1. 2018 Surveillance and Collection. Collections represented a subsample of observed mosquitoes to determine the relative proportions of vector species. For each species at each site and date, the relative proportion of adults and larvae are presented as a percentage of the total, and raw counts are in parentheses. Site identifiers correspond to labels on the map in Figure 1. No eggs or larvae were found in index sites via surveys in 2019. All adults collected were collected through either gravid trapping or mechanical aspiration in Garage A in 2019 and are not included in this table.

Date	**Site**		**N**	**Species Distributions by Site and Date**			
				Cx. pipiens Assemblage	*Ae. aegypti*	*Ae. albopictus*	Other
June 29, 2018	A	Adults:	24	100% (24)			
		Juveniles:	167	83% (149)	8% (14)	2% (4)	
July 26, 2018	A	Adults:	4	100% (4)			
		Juveniles:	0				
	B	Adults:	8	55% (5)			44% (3)
		Juveniles:	925	100% (925)			
	C	Adults:	38	100% (38)			
		Juveniles:	0				
August 10, 2018	A	Adults:	31	13% (4)	48% (15)		39% (12)
		Juveniles:	73	100% (73)			
	B	Adults:	14		100% (14)		
		Juveniles:	7		100% (7)		
Sept. 20, 2018	A	Adults:	5				100% (5)
		Juveniles:	0				
	B	Adults:	3		100% (3)		
		Juveniles:	0				
Total Collection		**Adults:**	***123***	**58% (71)**	**26% (32)**	**0% (0)**	**16% (20)**
		Juveniles:	***1171***	**98% (1147)**	**2% (20)**	**0.3% (4)**	**0% (0)**

Table 2: Genotypic identifications for a subsample of *Cx. pipiens* individuals collected in two subterranean garages in 2018 and 2019, based on the ACE2 and CQ11 loci.

	Year		**Site**		**Season**	
	2018	*2019*	*Garage A*	*Garage B*	*July*	*September*
***Culex* Ancestry**						
pipiens	20	10	20	10	14	16
molestus	25	12	22	15	17	20
quinquefasciatus	19	10	19	10	15	14
N*	64	32	61	35	46	50
Test Statistics	χ^2=0.031, p=1.00		χ^2=2.083, p=0.397		χ^2=0.245, p=0.901	

*This excludes the 9 unresolved individuals collected from Garage A in 2018, where two individuals were unresolved for our July collection and seven were unresolved in our September collection.

Chapter 2: Figures

Figure 1. 2018 Survey Sites. Index sites are sites A and B and were visited on each date. We surveyed, but did not find, standing water or mosquitoes at site D or E on July 30th, F or G on August 10th, or H or I on September 20th, 2018. For details about sites and dates where mosquitoes were present see Table 1.

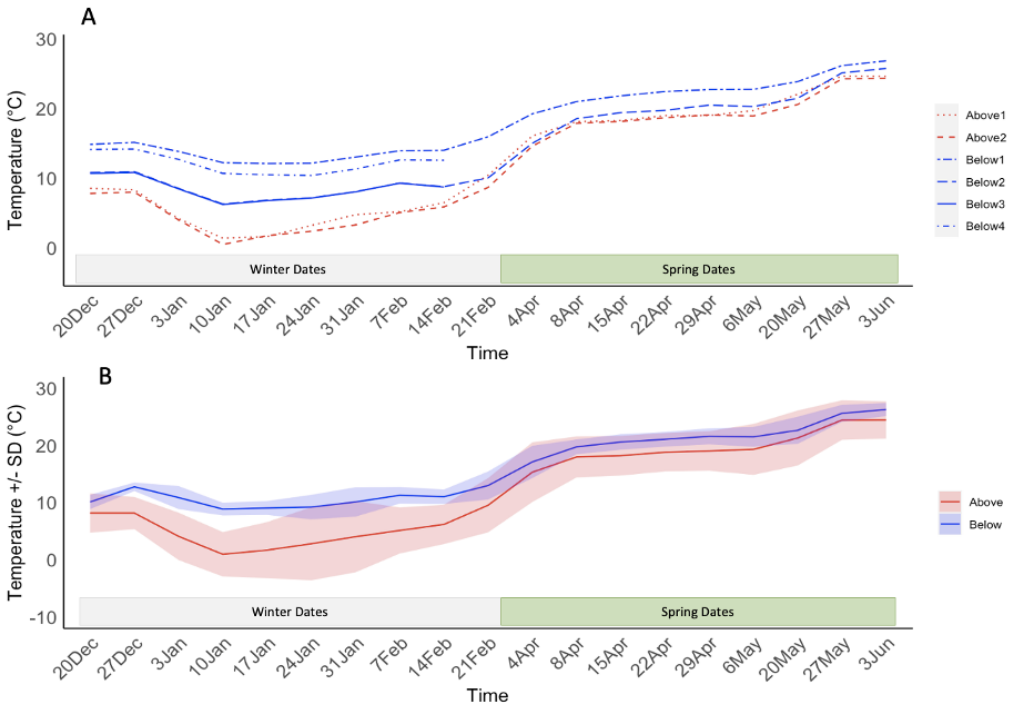

Figure 2. Moving average of above and below ground temperatures over the season. Temperatures were averaged over a two-week interval, then the analysis window slid one week forward to average the next two-week interval, from 20Dec2018-27Feb2019 (winter), and 1Apr2019-10Jun2019 (spring). A) Data from a single above ground Hobo logger (Above1), daily mean temperatures from a NOAA weather station at Reagan National Airport (DCA, Above2), and data from four below ground Hobo loggers dispersed on P3 of the garage (Below1-Below4). Below ground loggers 3 and 4 only recorded temperatures during winter dates. B) Means and standard deviations for sliding windows of daily mean temperatures, aggregated for all above and below ground loggers, respectively.

Chapter 3: *Spatiotemporal organization of cryptic North American* Culex *species along an urbanization gradient*

Abstract

Landscape heterogeneity driven by human land use creates diverse habitat and resources for mosquito vectors of disease. A consequence may be variation in the distribution and abundance of vector species over space and time dependent on niche requirements. We tested the hypothesis that landscape heterogeneity driven by urbanization influences the distribution and relative abundance of *Culex pipiens*, *Cx. restuans*, and *Cx. quinquefasciatus,* three vectors of West Nile virus (WNv) in the eastern North American landscape. We collected 9,803 cryptic *Culex* from urban, suburban, and rural sites in metropolitan Washington, District of Columbia during the months of June- October, 2019-2021. In 2021, we also collected mosquitoes in April and May to measure early-season abundance and distribution. Molecular techniques were used to identify a subset (n = 2,461) of collected *Culex* to species. Ecological correlates of the spatiotemporal distribution of these cryptic *Culex* were examined using constrained and unconstrained ordination. Seasonality was not associated with *Culex* community composition in June-October over three sampled years, but introducing April and May collection data influenced community composition in the final year of our study. *Culex pipiens* was dominant across site types and season. *Culex quinquefasciatus* was associated with urban environments, whereas *Cx. restuans* was associated with rural and suburban sites, primarily those sites with dense vegetation

and limited impervious surface. All three species rarely coexisted, and only did so in suburban environments nearest to the urban center.

Synthesis and applications: Mosquito communities maintain and amplify vector-borne pathogen transmission across space and time. Our work demonstrates that human-mediated land-use changes influence the distribution and relative abundance of *Culex* vectors of WNv, even on fine geospatial scales. In Maryland and Washington D.C., site classification, percent impervious surface, distance to city center, and latitude predicted *Culex* community composition. Moreover, we documented active *Culex* months before vector surveillance typically commences in this region, with *Culex restuans* being most abundant during April and May. Knowledge of habitat use and timing of activity of cryptic *Culex* in the mid-Atlantic USA will enhance surveillance and vector management aimed at reducing WNv transmission.

Introduction

West Nile virus (WNv) is a mosquito-borne zoonotic *Flavivirus* causing mild to severe neuroinvasive disease in humans (Ciota 2017). It is globally widespread and the most prevalent mosquito-borne disease in the continental United States. Between 1999-2010 in the northeast United States, WNv was isolated from 33 mosquito species in 8 genera, predominated by *Culex species* (96%) including *Cx. pipiens* (66.0%) and *Cx. restuans* (27.8%). In the southeast, the majority of WNv was isolated from *Cx. quinquefasciatus* (64.6%, Andreadis 2012). These species' ranges overlap across much of North America, and they may coexist within habitats across that range. There are phenological and spatial differences in the abundances of these species, yet patterns

may not be universal across land cover types or regions (Johnson et al. 2015, McMillan et al. 2020).

Human-generated land use change results in a complex ecological matrix with heterogeneous habitat across space (McDonnell and Hahs 2008). Urbanization influences both the composition of species present as well as their interactions (Grimm et al. 2008), but not all species are impacted by urbanization in the same ways. Urbanization can homogenize the landscape, typically leading to habitat loss and loss of biodiversity (Ferraguti et al. 2016). For some species, including mosquitoes, human infrastructure can create novel and suitable habitat, and impact community composition in urban centers. Understanding the effects of urbanization on these communities is important because mosquito-borne pathogen transmission dynamics pose risks to human health (Ferraguti et al. 2016).

Mosquitoes require moist conditions or standing water for oviposition, suitable habitat for their vertebrate hosts, resting locations, and refugia. Optimal conditions vary by species, however. Variability in the conditions across a landscape as well as resource availability influences how members of vector communities are distributed through space (Vinatier et al. 2011). For example, standing irrigation waters, small containers, stormwater infrastructure, and flooded subterranean concrete surfaces produced through human activity may serve as suitable ovipositional habitat for some, but not all, vector species (Norris et al. 2004, Townroe and Callaghan 2014). Human-made habitats, such as insulated structures, may serve as overwintering (Arsenault-Benoit et al. 2021) or oversummering (Thereja et al. 2022) refugia that help to maintain mosquito populations in climatic conditions otherwise outside of their natural thermal

tolerance limits. Vegetation density and heterogeneity creates varied resting sites and habitat for vectors and their hosts, from cityscapes to parks, manicured lawns, forests, or agricultural fields (Ferraguti et al. 2016, Gorris et al. 2021). Together, these natural and human-made features create an environmental mosaic across a landscape that can underpin the distribution of mosquito vectors.

Communities of vectors can amplify, distribute, and transmit arboviruses as a network, so interactions between species and with the landscape impact disease burden. In the northeastern United States, *Cx. restuans* is considered an early-season mosquito, emerging sooner and declining before *Cx. pipiens* reaches peak abundance (Lee and Rowley 2000, Andreadis et al. 2001, Johnson et al. 2015). Because *Cx. restuans* is a competent vector of WNv, ornithophilic, and abundant in the spring, a common inference is that this species is important for early season WNv amplification within the avian population. This allows for increased transmission to mammals when *Cx. pipiens* populations peak (Andreadis et al. 2001, Johnson et al. 2015). At mid-latitudes, non-diapausing *Cx. quinquefasciatus* are also present in the environment (Smith and Fonseca, 2004). *Culex pipiens* and *Cx. quinquefasciatus* are closely related species that may hybridize where their ranges overlap (Smith and Fonseca 2004, Kothera et al. 2009). A stable hybrid zone between 36°N and 39°N latitude is well-described (Barr et al. 1957, Kothera et al. 2009, Kothera et al. 2012). Defining species boundaries within this hybrid zone is challenging, as secondary contact and interbreeding between taxa has been widespread, suggesting the potential for gene flow between *Cx. pipiens* and *Cx. quinquefasciatus* (Kothera et al. 2012). Although *Cx. quinquefasciatus* are also likely vectors of WNv, they are not regularly included in studies of the spatiotemporal

dynamics of *Culex.* There is a need for the systematic study of *Culex* population dynamics that considers environmental features, as well as phenology.

Governments and municipalities in WNv-affected areas contribute significant resources to collecting, identifying, testing, and treating mosquitoes, but time and resource constraints cause pooling of the three *Culex* taxa described here during surveillance (Kilpatrick et al. 2005, Jackson and Paulson 2006). Routine pooling of *Culex* species limits the utility of surveillance data to understand species' distribution and contribution to WNv transmission across the landscape. Better understanding of the spatiotemporal relationships between these *Culex* vectors can be leveraged to improve surveillance to better predict human WNv cases and potentially develop approaches that prevent mosquito-borne illness (Kilpatrick et al. 2008, Norris et al. 2004). We contribute to this understanding by assessing the role of seasonality and landscape features on the spatiotemporal organization of the community assemblage of *Culex* species along an urban to rural gradient in a major North American urban center, Washington, District of Columbia.

Methods

Mosquito Collection and Identification

Field Sites

Fifteen collection sites were established radiating ~50km north and east from downtown Washington, D.C., (Figure 1). Sites were initially categorized as urban, suburban, and rural using characteristics of primary housing type, zoning, and the design and extent of nearby green space. Urban sites (n = 5) were within Washington,

D.C. proper, including four above- and one belowground site in commercial areas and residential areas. Suburban sites (n = 5) were within Montgomery and Prince George's County, MD, and were primarily residential. Rural sites (n = 5) were in Prince George's, Howard, and Anne Arundel Counties, MD. Rural site land management ranged from unmanaged natural forested areas to active farms. Ecological variables for each site were quantified within a buffer zone with a 0.5 km radius using GIS tools in ArcMap 10.7. Variables included land cover class, percent impervious surface, and percent tree canopy cover (NLCD 2016); population density and housing unit density (2010 US Census); linear distance to the geographic center of the city, soil water depth, elevation, and Normalized Differential Vegetative Index (NDVI). The buffer zone was approximately equal to the typical flight range of 0.8km^2 for adult *Cx. pipiens*, as measured by mark-release-recapture (Hamer et al. 2014). The mean for continuous variables, and majority for categorical variables, of pixels within the buffer for each layer were calculated for each site (Appendix I, Table 3.1). After preliminary data collection, we refined our classification categories to reflect alternative site characteristics: natural, agricultural, residential, and commercial sites. These categories were named "secondary site classes" to represent combinations of site characteristics that were obscured in the initial site classification. For example, all urban sites had high impervious surface, but residential sites had high population density whereas commercial sites had low population density (Appendix I, Table 3.1).

Mosquito collection

In 2019-2021, CDC gravid traps (John Hock Co., Gainesville FL, USA) baited with a hay-infused oviposition attractant and powered by 6V battery were set to run for 18-24 hours at each site, encompassing both dusk and dawn (Reiter 1983). Traps were set at one randomly assigned urban, suburban, and rural site on a given night. Collected mosquitoes were transported in an ice-chest to the University of Maryland, where they were cold-anesthetized at -20°C for 30 minutes or more, morphologically identified to the lowest taxon possible via microscopy, and stored at -80°C until molecular identification to species.

Timing of collections

Eight collection events were undertaken at each of the 15 sampling sites per year in 2019-2021, yielding 120 trap-nights per year. Collections were made once per site within a two-week block in random order between June and October. Data were divided into early (June 1-July 10), mid (July 10-August 20), and late (August 20-October 1) season blocks for analysis. These dates were selected to align with periods when local vector control practitioners survey for mosquitoes and vector-borne disease (Brian Prendergast, Maryland Department of Agriculture, personal communication). No changes to typical mosquito management were made at any site. Preliminary data suggested that some seasonal trends may exist in the very early season. To capture these dynamics, we began collections six weeks earlier in 2021, beginning on April 4 and ending on May 25, to sample three additional times in the “very early” season block. Overall in 2021, we sampled 11 collection events for each of the 15 sites for a total of

165 trap-nights. Of the 405 trap-nights intended for the study (120 trap nights in 2019 and 2020, and 165 trap nights in 2021), we collected mosquitoes on 390 of the nights. There were 12 trap-nights without data collection due to trap malfunction or tampering. Trap failures were randomly spread across sites, dates, and years. In 2020, we could not access Patuxent Wildlife Refuge from June 1- July 20th due to COVID-19 restrictions, so the first three collections from that site and year are also missing.

Culex *species identification*

Culex species are prone to misidentification (Harrington and Poulson 2008), but molecular approaches are available to identify cryptic specimens. Intact *Culex* specimens, including those suspected to be *Cx. pipiens*, *Cx. restuans*, or *Cx. quinquefasciatus* were separated from blood fed individuals, damaged individuals, and members of other taxa. Subsets of up to 30 cryptic *Culex* females were randomly selected per site, per season block (early, mid, late) for molecular identification. Where possible, we included 10 individuals per trap-night. Genomic DNA was extracted using one of three methods: from the abdomen using a rapid alkaline-based extraction (Rudbeck and Dissing 1998) or from the abdomen (excluding spermathecae) using commercially available DNA isolation kits in 2019 and 2020 (Qiagen DNeasy Blood and Tissue Kit (Cat. No. 69506, Qiagen Inc., Valencia CA, USA or Zymo Quick-DNA Miniprep Plus kit (Cat #11-397B, Zymo Research, Irvine, CA)); or a rapid macerating procedure in 2021 (Gloor and Engels 1991). All methods had high-yield amplification via polymerase chain reaction (PCR, 89%, 97%, and 95% respectively). DNA was amplified with a multiplexed PCR targeting the 28S ribosomal subunit to identify *Cx.*

pipiens (698 bp), from *Cx. restuans* (506 bp) (Crabtree et al. 1995). This PCR also identifies *Culex salinarius.* Because gravid traps are biased against this species (Williams and Gingrich 2007), we likely did not collect a representative sample of this species across the landscape. Therefore, any detected were excluded from analysis (n=16 over 3 years). Specimens identified as *Cx. pipiens* underwent a second PCR targeting their *acetylcholinesterase-2* (ACE2) gene which distinguished *Cx. pipiens* (610 bp) from *Cx. quinquefasciatus* (274 bp) (Smith and Fonseca 2004). The PCR conditions for the each of the assays were run as described in Arsenault-Benoit et al. (2021). We molecularly identified 712 individuals from 2019, 814 from 2020, and 935 for 2021. Six hundred seventy one of these 2021 samples were from early, mid, and late season blocks (June-October), and 264 were from the very early season block (April-May).

Statistical Analysis

Total Culex *catch*

All statistical analyses were completed in R version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria). To assess the influence of landscape and season on total cryptic *Culex* catch, we compared generalized linear models including primary site class (urban, suburban, or rural) and season (early, mid, or late). We began with the full model Catch=Class × Season+(1|Year) for 2019-2021, and Catch=Class × Season for 2021 with a negative binomial error structure. A model reduction approach was used to determine whether including or excluding predictors explained more of the variance in total catch. We tested the impact of temporal autocorrelation by applying

an autocorrelation function (`acf`) on the residuals of the generalized linear model to look for significant autocorrelation across six time lags, as suggested for seasonally collected data (two*period of seasonality (early, mid, late), Hyndman and Athanasopoulos 2018).

Culex pipiens *assemblage genetic analysis*

To determine whether *Cx. pipiens* and *Cx. quinquefasciatus* could be considered separate taxa for this analysis, following molecular identification, we tested whether populations deviated from expectations of heterozygosity under Hardy Weinberg Equilibrium (HWE) at the ACE2 locus by tallying *Cx. pipiens* homozygotes, *Cx. quinquefasciatus* homozygotes, and heterozygotes by site, season, and year (n=135). The `HWChisqMat` function in the package `HardyWeinberg` (Graffelman 2015) was applied to analyze deviance from HWE in each instance.

Species count distributions by site class and season

Contingency table analyses were used to compare species counts across season, site class, and the interaction between season and site class. We used a Cochran-Mantel-Haenszel test for three-way analysis (Species Counts ~ Class × Season) with an *a priori* α value of 0.05 to determine statistical significance.

Community Analyses

For community analysis, individuals that were homozygous for *Cx. quinquefasciatus* alleles at the ACE2 locus were designated *Cx. quinquefasciatus*, while the few heterozygotes that we identified (n=7) were considered *Cx. pipiens.* We used non-metric multidimensional scaling (NMDS) with Bray-Curtis distance as the dissimilarity value to visualize the cryptic *Culex* spp. community assemblages in relation to site, season, and environmental variables for two through five dimensions (Appendix I, Table 3.3, Figure 3.3). We then used the `envfit` function in `vegan` (Oksasan et al. 2020) to test the correlation between environmental and temporal predictor variables to the ordination axes and tested for significance using 999 permutations at each k value (Appendix I, Table 3.3). Variables that were consistently correlated with the ordination axes were considered as predictors in a constrained ordination.

We applied canonical correspondence analysis (CCA) to identify predictors most likely to correlate to variation in community assemblages (see supplementary materials, Appendix I), and variance partitioning to determine how much variance in community assemblage was explained by season, geographic space, and environment. We plotted the variance of residuals of final CCA models with increasing geographic distances between sites to assess the spatial autocorrelation using the `mso` function in `vegan`. This method uses permutation tests to compare mean inertia in each distance class to the overall mean inertia (Wagner 2004). We did not correct for spatial autocorrelation due to spatial independence of residuals for CCA models (Appendix I, Figure 3.6).

Our primary analysis consisted of data collected from June 1-October 1 across all three years of the study. We then used the same analytical approaches to capture the impact of the very early season block (April 4-May 25) using only 2021 data.

Results

Mosquito Collection and Identification

We collected 4,103 unfed *Culex* individuals in 2019, 1,777 in 2020, and 2,891 in 2021 for a total of 8,771 over 345 trap-nights from June-October. Including the interaction between site class and season explained more variation in total collections than season alone (χ^2=12.868, df=6, p=0.0452), and marginally more variation than including primary class alone (χ^2=12.337, df=6, p=0.0549, Appendix I, Figure 3.1). The residuals of the generalized linear model did not demonstrate temporal autocorrelation using `acf` (α<0.05, ACF= -0.2 to 0.2), so we did not correct for temporal autocorrelation.

Molecular identification revealed a total of 650 *Cx. restuans* and 1,811 *Cx. pipiens* assemblage mosquitoes. Of the *pipiens* assemblage members, 1,703 had homozygous *Cx. pipiens* genotypes at the ACE2 locus and 108 had homozygous *Cx. quinquefasciatus* genotypes. Seven individuals out of 1,811 originally identified as *Cx. pipiens* assemblage were then identified as heterozygotes, or *pipiens-quinquefasciatus* hybrids, over three years (Appendix I, Table 3.2), although a single-locus assay likely underestimates the number of potential hybrids in a population (Kothera et al., 2009). Of the 135 site by season by year occurrences, 53 had both *pipiens* and *quinquefasciatus* alleles represented, and of those, 37 significantly deviated from HWE

expectations, indicating violation of HWE assumptions for a single population (Appendix I, Table 3.2). Hybrids were present but rare, and HWE analysis suggested a lack of gene flow among *Cx. pipiens* and *Cx. quinquefasciatus* at each site, perhaps due to assortative mating between taxa. Accordingly, homozygous specimens were considered separate species according to genotype and the 7 heterozygotes were categorized as *Cx. pipiens,* the most abundant genotype.

2019-2021 June through October collections

Counts of *Cx. pipiens*, *Cx. restuans*, and *Cx. quinquefasciatus* were influenced by both site class and season (Cochran-Mantel-Haenszel Test: M^2=30.781, df=4, $p<0.001$, Figure 2). Using unconstrained ordination, we found that latitude, NLCD land cover class, distance to city center (DCC), normalized differential vegetation index (NDVI), and site were significantly associated with community structure (Appendix I, Tables 3.3 and 3.4, Figure 3.4). Season did not have a significant effect on community composition by site class.

`Ordistep` revealed that several environmental variables impacted *Culex* distribution across the landscape, and significant variables were used to create constrained ordination (CCA) models ($\alpha<0.01$; Table 1, Appendix I, Table 3.4). After accounting for multicollinearity, we report two models consisting of environmental predictors that were associated with the distribution of cryptic *Culex* along an urbanization gradient. In Model 1, DCC, latitude, and site class explained 39.7% of the total inertia, which is significant based on permutation tests (adjusted R^2=0.338, p=0.002, Figure 3A). Constraining the response by primary site class explained a

significant amount of variation in this model (χ^2=0.0894, F=12.585, df=2, p=0.001). *Culex quinquefasciatus* was positively associated with urban environments and shorter DCC, whereas *Cx. restuans* was positively associated with rural environments and increasing latitude, but negatively associated with DCC (Figure 3A). In Model 2, secondary site class, DCC, percent impervious surface, and latitude explained 44% of the total inertia (adjusted R^2=0.352, p=0.001). Both secondary site class (χ^2=0.0763, F=7.309, df=3, p=0.001) and percent impervious surface (χ^2=0.0252, F=7.244, df=1, p=0.006) were independently significant. *Culex quinquefasciatus* was positively associated with commercial environments, increasing impervious surface and shorter DCC. *Culex restuans* was negatively associated with commercial environments, increasing percent impervious surface, and shorter DCC. *Culex pipiens* was not strongly associated with any environmental predictor in either model. Early and late season communities encompass mixed-species assemblages. In the mid-season, however, the centroid of responses shifted toward a *Cx. pipiens*-dominated community, closer to the origin of the tri-plot, suggesting more homogeneity in community composition between sites in mid-season (Figure 3B).

Variance partitioning of the three-year dataset via the `varpart` function in `vegan` demonstrated that about 12% of inertia in the community assemblage was explained by season, and 5% was explained by spatial distribution. We described environment by using all linearly independent continuous variables in our site by environment matrix: DCC, NDVI, percent impervious surface, population density, water table depth, and elevation (Appendix I, Table 3.1). Environment alone only

explained 9% of inertia, but 34% of inertia was shared between latitude-longitude and environment (Figure 3C).

2021 April through October collections

April and May collections added 1,032 *Culex* to the previously described 2,891 collected in June through October 2021 (Appendix I, Figure 3.2). During April and May, mean *Culex* catch was 24±5.851 per trap, congruent with the June through October surveillance season (early=28.72±8.075, mid=20.02±4.340, late=30.20±8.779). The interaction between site class and season explained more variation in *Culex* catch than if primary class (χ^2=31.997, df=8, p<0.0001), or season (χ^2=29.095, df=9, p=0.0006) were modeled alone.

We molecularly identified and analyzed 264 *Culex* from the very early season alongside the 611 June-October 2021 samples described above, revealing 337 *Cx. restuans* and 578 *Cx. pipiens* assemblage mosquitoes. Twenty-one had homozygous *Cx. quinquefasciatus* genotypes and 557 had homozygous *Cx. pipiens* genotypes at the ACE2 locus. None were heterozygotes.

Species counts differed by season and site class (Cochran-Mantel-Haenszel Test: M^2=24.917, df=6, p=0.0003, Figure 4). Fewer collections included all three species in 2021 compared to the three-year June-October dataset (Figure 6). In 2021, collection event, week of the year, season, latitude, percent impervious surface, percent tree cover, DCC, site, primary class, secondary class, and NLCD land cover class were significantly associated with community structure, as determined by NMDS (Appendix I, Table 3.3, α=0.05).

We executed the same CCA models for April-October data from 2021 and found associations between species and environmental variables were similar to the 2019-2021 data (see Appendix I). *Culex quinquefasciatus* was positively associated with commercial environments with increasing impervious surface near the city center. *Culex restuans* was positively associated with all but commercial environments and increased DCC, and negatively associated with increased impervious surface. *Culex pipiens* did not have strong relationships to any predictor, although may be loosely associated with residential and agricultural environments (Figure 5).

Season significantly predicted community assemblages for April-October collections in both unconstrained and constrained analysis (p=0.005, Figure 5, Appendix I, Table 3.3, Figure 3.4). In 2021, season explained 36% of inertia in the community assemblage, latitude-longitude explained 1% of inertia, and environment explained 9% of inertia. Eighteen percent of variance was shared between geographic space and environment, and 5% was explained by all three partitions (Figure 5C).

Discussion

We sought to describe the spatiotemporal distribution and assemblages of cryptic *Culex* species across an urbanization gradient using molecular identification and community modeling approaches. Although these cryptic *Culex* species are sympatric, we found that most collections comprised 1-2 *Culex* species and collections containing all three species were uncommon. When all three did occur together, they were limited to suburban areas or urban areas early in the season.

Culex pipiens is commonly considered the primary bridge vector of WNv to humans due to high vector competence, high infection rate in the field and the laboratory, high abundance, and willingness to feed on both birds and mammals (Andreadis 2012, Hamer et al. 2008). *Culex pipiens* dominated our collections across sites and seasons and were not closely associated with any season, site class, or landscape feature, perhaps due to their ability to use varied habitats. Site class and environmental features were important in determining which cryptic *Culex* species coexisted with *Cx. pipiens* at each site. Collections in rural and suburban environments often contained *Cx. pipiens* and *Cx. restuans,* while *Cx. pipiens* and *Cx. quinquefasciatus* were primarily collected together in urban environments.

We demonstrated that *Cx. restuans* were associated with rural, natural, and agricultural sites with low impervious surface furthest from the city center. They were rarely collected in urban environments, suggesting they may use these environments but are either ill-suited to them, or outcompeted by congeners that are highly adapted to anthropogenic features, including *Cx. pipiens* and *Cx. quinquefasciatus*. For instance, the fewest *Cx. restuans* (n=7) and most *Cx. quinquefasciatus* (n=27) were identified from the belowground site, which demonstrated the most pronounced anthropogenic and urban characteristics (see Appendix I, Table 3.1).

Culex quinquefasciatus were found at low but consistent levels across the season. We identified *Cx. quinquefasciatus* across site classes but most commonly in urban environments, commercial environments and at non-urban sites closest to the city center, suggesting they may disperse from a primary population in the city center.

There was some evidence of hybridization between *Cx*. *quinquefasciatus* and *Cx. pipiens*, but we speculate that there is limited gene flow between these taxa in the field based on our HWE analysis, even though they overlap geographically and seasonally in the greater Washington D.C. region.

Culex quinquefasciatus were most abundant in early and late season, possibly because this species does not diapause and is thus able to persist when conditions are suitable and remain in the environment until conditions require them to seek refuge. High impervious surface, urban heat, and dense subterranean networks are hallmarks of commercial city centers, as are small containers in which *Cx. quinquefasciatus* are likely to oviposit (Leisnham et al. 2014). We propose that a combination of the urban heat island effect and insulated and linked subterranean habitat in commercial areas support this species better than urban residential areas due to overwintering thermal limitations (Harbison et al. 2011, Arsenault-Benoit et al. 2021).

Natural and agricultural sites hosted different species assemblages; we speculate that plants in monoculture do not offer the same resources for mosquitoes as diverse vegetation and heterogeneous canopy cover in forests do. Rural *Cx. quinquefasciatus* collections tended to be at agricultural sites closest to the city, with high overall vegetation but not necessarily a diverse canopy. Conversely, *Cx. restuans* were found at rural agricultural sites less than they were at rural natural sites. Normalized Differential Vegetation Index (NDVI) was high across rural sites and some suburban sites (Appendix I, Table 3.1), yet was not independently found to be an important factor for species assemblages in any ordination (Table 1). The types of

vegetation comprising NDVI may be more important than the abundance of vegetation. Species distribution models suggest that the most important variables for *Cx. restuans* habitat are areas with high humidity, evergreen forests, and herbaceous vegetation (Gorris et al. 2021). They may not thrive in agricultural fields with no overstory, and consequently high heat and low humidity. *Culex quinquefasciatus* rely less on environmental humidity as they exploit human-generated water sources (Valentine et al. 2020, Gorris et al. 2021), which are common at both urban and agricultural sites. *Culex restuans* is the only truly native species in our study and were prevalent in natural environments. Natural sites are least disturbed and may be most resistant to establishment of species undergoing range expansion and establishment, like *Cx. quinquefasciatus* and, to a lesser extent, *Cx. pipiens*. Use of natural, forested areas by *Cx. restuans* may enhance their role in WNV transmission in the early season at higher latitudes due to wet conditions in natural areas after snowmelt or in regions with a particularly rainy spring season.

Our findings support *Cx. restuans* as an early season species in greater Washington, D.C. prior to the local mosquito management season, but assemblages do not change with season from June-October when local management occurs. The importance of early season *Cx. restuans* populations to WNv transmission is uncertain. In the eastern USA, human WNv cases generally appear in August and September (week of the year (WOY) 30-35), and infected mosquitoes are rarely collected before late July (WOY 28; McMillan et al. 2020, Karki et al. 2020). In a survey of *Culex* vectors in Connecticut, *Cx. restuans* were abundant much earlier in the season than *Cx. pipiens*, but WNv positive *Cx. restuans* were not collected until mid-July (Andreadis

et al. 2001). In Virginia, a single WNv positive *Cx. restuans* pool was identified in early May, but WNv was not commonly identified in *Cx. restuans* until late June (Johnson et al. 2015). We observed a high abundance of *Cx. restuans* in rural environments early in the collection season, and consequently at the very beginning of vector surveillance and abatement in our region. Surveillance may not capture mosquitoes at sites where early season amplification occurs, like natural areas. Low risk of human cases in these areas cause resources to be focused elsewhere.

Studies speculate that a community of co-occurring *Culex* species may increase and extend WNv transmission across a season due to differences in phenology, habitat use, and host preferences (Reiskind and Wilson 2008) and the complement of important vectors likely differs geographically (McMillan et al. 2020, DeGroote et al. 2008, Tokarz and Smith 2019). Our work was limited to a few species but highlights how vector communities can change across the landscape. The existence of networks of vectors necessitates a community approach to surveillance and management. West Nile virus is a highly connected pathogen that is associated with many vectors; the suite of competent vectors is projected to be up to 60 species worldwide (Yee et al. 2022). A likely consequence is that dominant or primary vectors differ regionally and across the landscape based on environmental, climatic, and biotic factors (Chandrasegaran et al. 2020, McMillan 2020). Phenology may also be of major or minor importance to disease transmission depending on geographic region and local species assemblages.

In this study, we found that season is important to *Culex* dynamics when mosquitoes first emerge. However, from June to October, when WNv is detected in

mosquitoes and human cases occur, environmental features, and especially characteristics representative of urbanization, are the primary predictors of the *Culex* community assemblage across the landscape. Future studies analyzing community structure over space in other US metropolitan areas may elucidate the role of both landscape and phenology in assemblage dynamics and establish whether the patterns we observed are universal.

Chapter 3: Tables

Table 1. Significant predictors that were included in the final constrained ordination, and associated p-values calculated by the ordistep function. The full set of environmental predictors tested is in Appendix I, Table 3.1, and measures of significance for each predictor via both envfit on NMDS and ordistep for constrained analysis can be found in Appendix I, Table 3.3 and 3.4.

Predictor	**Value Range**	**p-values**	
		2019-2021, June-October	***2021, April-October***
Latitude	38.847°N-39.256°N	0.010*	NS
NLCD Land Cover Class	*Categorical, see App. I, Table 3.1*	0.005*	0.005*
Percent Impervious Surface (Mean)	0.4-75.8%	0.005*	0.005*
Distance to City Center	1.979-41.351	0.005*	0.005*
NDVI	68.69-120.9	0.025*	NS
Primary Class	*Categorical, see App. I, Table 3.1*	0.005*	0.005*
Secondary Class	*Categorical, see App. I, Table 3.1*	0.005*	0.010*
Season	(Very) Early-Late	NS	0.005*
Percent Tree Cover	0-74%	NS	0.005*

* denotes statistical significance assuming an *a priori* $\alpha = 0.05$. NS=not significant.

Chapter 3: Figures

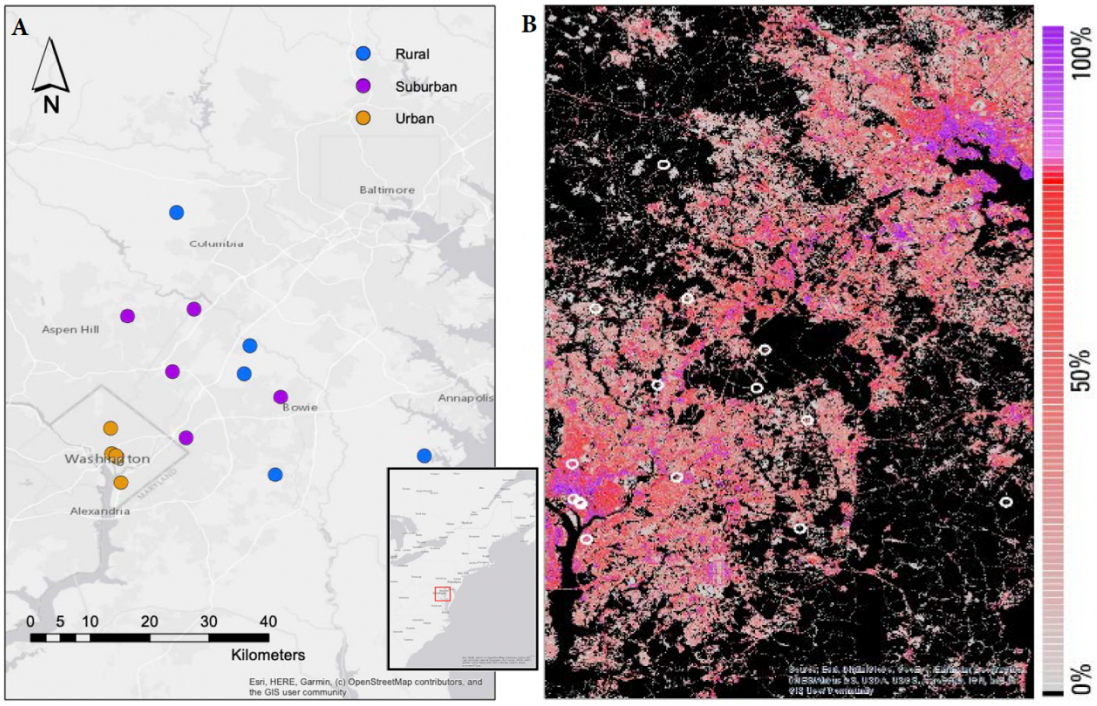

Figure 1. Collection sites. A. A map of the greater Washington D.C. area with collection sites marked by dots. Urban sites are in orange, suburban sites are in purple, and rural sites are in blue. B. Sites are overlayed on NLCD percent impervious surface map (National Land Classification Database, 2016) to demonstrate environmental heterogeneity of the study sites.

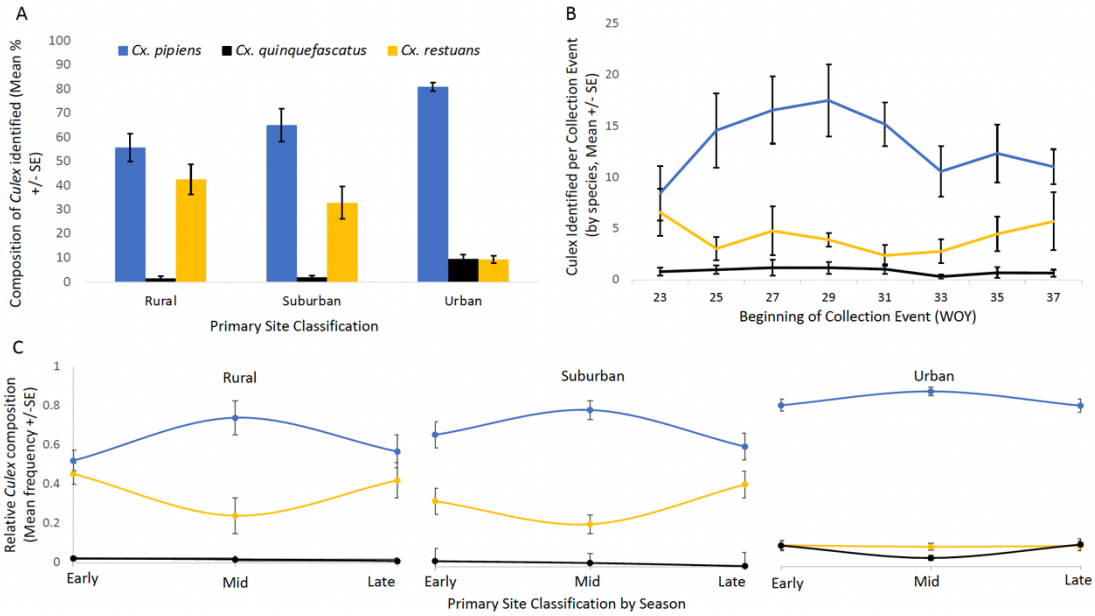

Figure 2. Summary of 2019-2021 June-October *Culex* molecular identifications (N=2,197). Approximately n=10 individuals per collection event per year were molecularly identified for this analysis. A: Composition of *Culex* by primary site class is represented as mean percent of specimens (% ± SE) identified. B: Mean species counts identified over the course of the season, aggregated across years. The start of each 2-week collection window is indicated by week of the year (WOY). C: Mean relative frequency of species identified in early, mid, and late season, separated by site class.

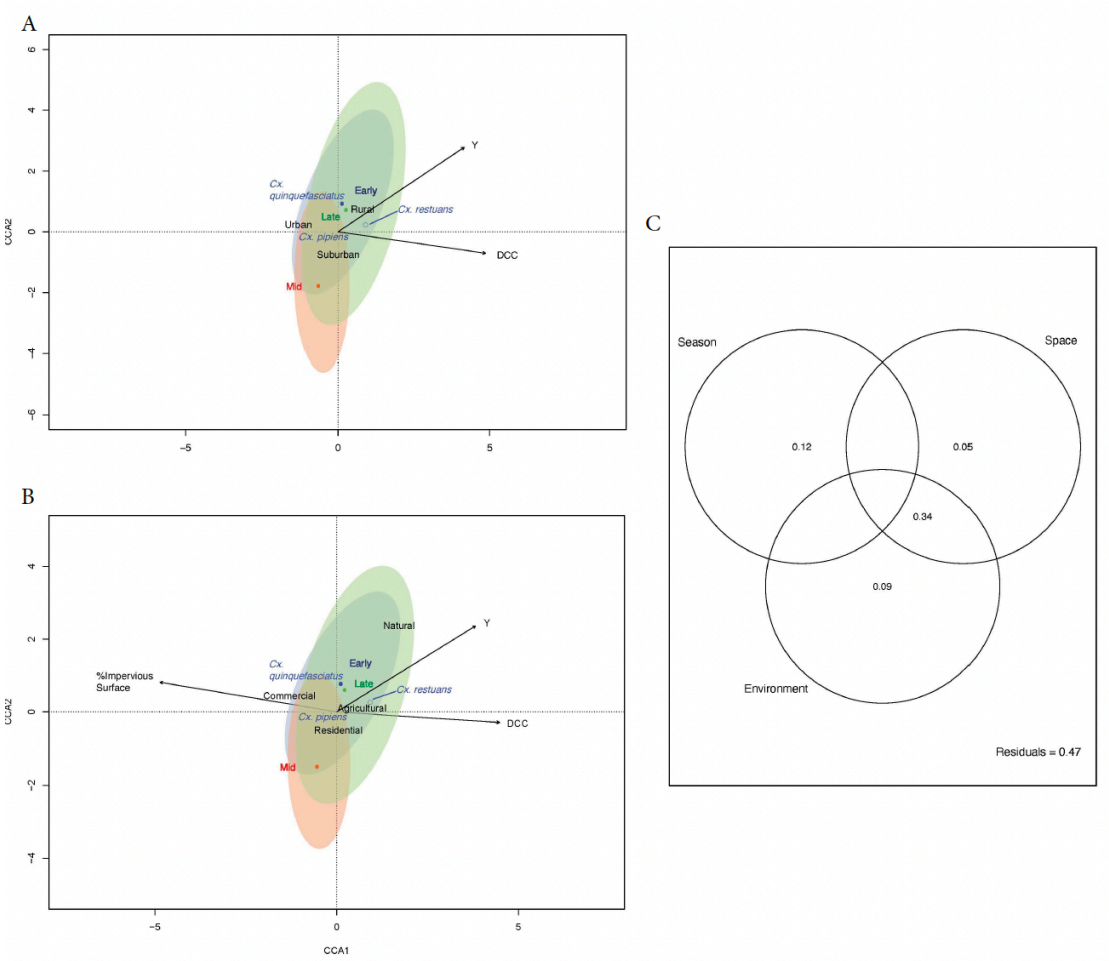

Figure 3. CCA tri-plots and variance partitioning for June-October, 2019-2021. Site names have been suppressed to improve readability. Two tri-plots are presented due to collinearity between predictors. Seasonal ordiellipses projected over CCA tri-plots show how responses cluster in ordination space. The centroid of each ellipse is designated by a color-coordinated point. A. CCA tri-plot for model 1: Primary Site Class+ Distance to City Center (DCC) + Latitude (Y). B. CCA tri-plot for model 2: Secondary Site Class + Percent Impervious Surface +Distance to City Center (DCC) +Latitude (Y). C. Variation in the 2021 community assemblage represented as the percent explained by season, geographic space (latitude and longitude), environment. Values less than zero (<0) are not shown.

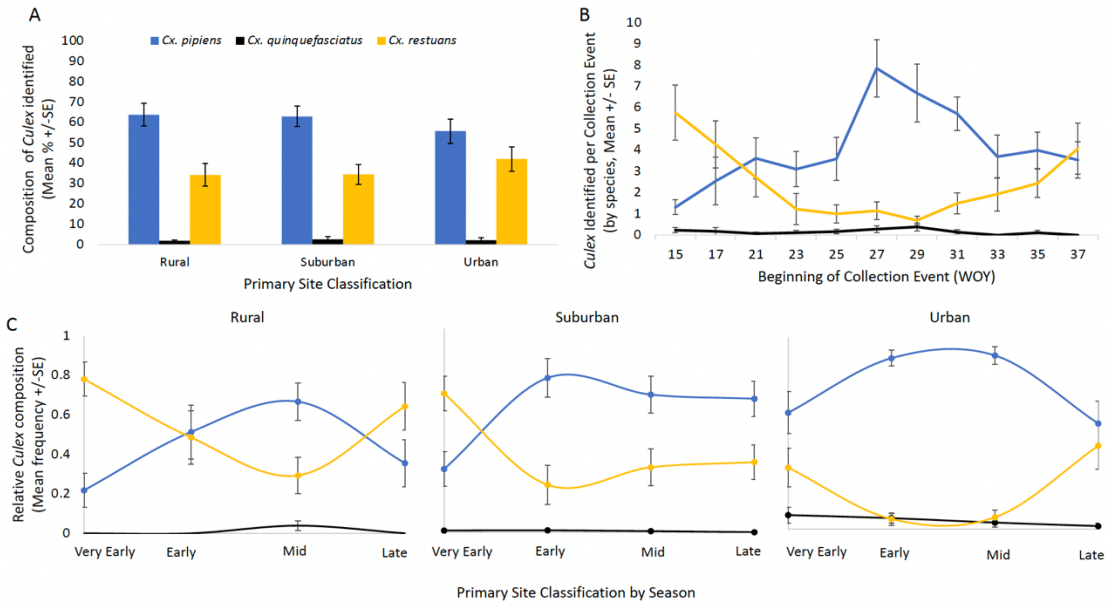

Figure 4. Summary of April-October 2021 *Culex* molecular identifications. Approximately n=10 individuals per collection event per year were molecularly identified for this analysis. A: Composition of *Culex* by primary site class is represented as mean percent of specimens (% ± SE) identified. B: Mean species counts identified over the course of the season, aggregated across years. The start of each 2-week collection window is indicated by week of the year (WOY). C: Mean relative frequency of species identified in early, mid, and late season, separated by site class.

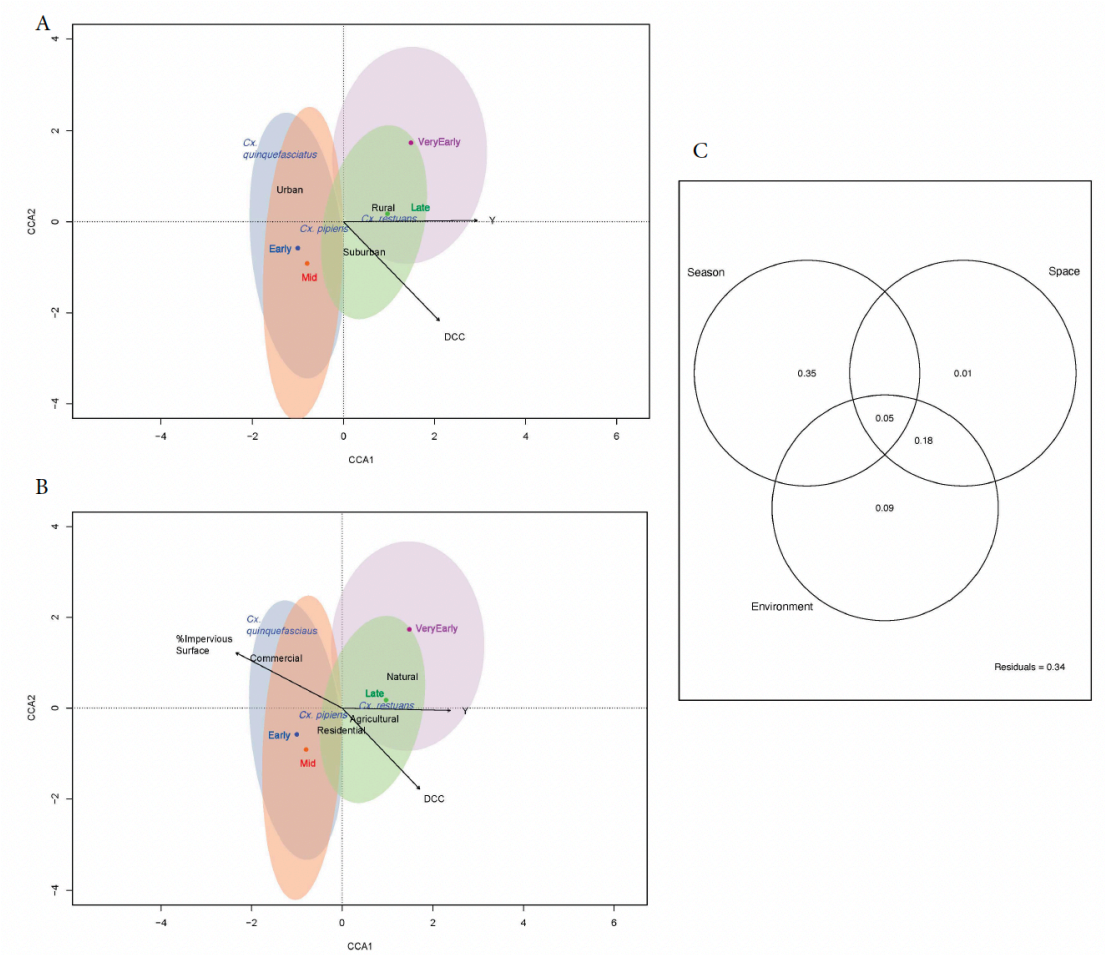

Figure 5. CCA tri-plots and variance partitioning for May-October 2021. Site names have been suppressed for readability. Two tri-plots are presented due to collinearity between predictors. Seasonal ordiellipses projected over CCA tri-plots show how responses cluster in ordination space. The centroid of each seasonal ellipse is designated by a color-coordinated point. A. CCA tri-plot for model 1: Primary Site Class+ Distance to City Center (DCC) + Latitude (Y). B. CCA tri-plot for model 2: Secondary Site Class + Percent Impervious Surface +Distance to City Center (DCC) +Latitude (Y). C. Variation in the 2021 community assemblage represented as the percent explained by season, geographic space (latitude and longitude), environment. Values less than zero (<0) are not shown.

Chapter 4: Landscape factors, phenology, and climate affect relative cryptic Culex abundance along urban to rural gradients in Eastern North America

Abstract

Cryptic *Culex* mosquito species can transmit West Nile virus (WNv) and co-occur across the landscape, as part of a larger network of WNv vectors in eastern North America. These species, including *Cx. pipiens*, *Cx. restuans,* and *Cx. quinquefasciatus* form a dynamic community assemblage, in which they interact with each other, with the climate, and with environmental factors across the landscape at broad and fine spatiotemporal scales. To explore these relationships, we collected and molecularly identified cryptic *Culex* from urban to rural gradients in three eastern United States metropolitan areas: greater Washington, D.C. and Maryland, greater Philadelphia, PA, and greater Chicago, IL. North of Maryland, *Cx. quinquefasciatus* were rare or absent, so they were excluded from further analysis. We used a gradient forest (GF) approach to discover important environmental predictors for *Culex* species abundance and to predict where along temporal and environmental gradients turnover in species abundance occurs. Climate, phenology, and landscape characteristics all played a role in determining the community assemblage between *Cx. pipiens* and *Cx. restuans.* Crossing-over from *Cx. restuans*-dominant to *Cx. pipiens*-dominant assemblages was evident at all sites, but occurred two to five weeks later in more northerly localities. This shift is likely due to a temperature gradient associated with latitude, as *Cx. restuans* and *Cx. pipiens* are negatively and positively associated with warmer

temperatures, respectively. Landscape factors associated with urbanization also predict the community composition. Turnover between *Cx. pipiens* and *Cx. restuans* dominance occurred at intermediate values along gradients of percent impervious surface, distance to city center, and percent tree cover; *Cx. pipiens* was more abundant in human-dominated areas while *Cx. restuans* were more abundant in less disturbed, more natural environments. Our results demonstrate that the composition of cryptic *Culex* assemblages is mediated by phenological, climatic, and environmental factors. A diversity of species plays a role in maintaining, amplifying, and transmitting WNv over space and time, including members of the *Culex* genus explored here. An understanding of the community ecology of these vectors and how they use the landscape as we present here, can aid in predicting, managing and mitigating WNv risk.

Introduction

Mosquito-borne disease is a primary public health concern, as it accounts for 17% of worldwide illness (Rinker et al. 2016, Chandrasegaran et al. 2020). A single mosquito-borne pathogen can be transmitted by a complement of competent vectors, forming a network between pathogens, vectors, and hosts (Yee et al. 2022). This vector network facilitates pathogens' covering broad geographic and climatic ranges and rapid invasion of novel environments. Unique ecologies and life histories among vectors, including different habitat use, host preferences, and physiological needs influence their distribution on broad and fine scales, and impact each population's vectorial capacity for a given pathogen (Ferraguti et al. 2016, Dunphy et al. 2019, McMillian et al. 2020). Vector networks can span broad geographic ranges connecting pathogens

among allopatric vectors, but species can also be sympatric and form community assemblages that undergo compositional turnover as the environmental and climatic variables shift over time and space. One such highly networked pathogen is West Nile virus (WNv).

West Nile virus is a neuroinvasive flavivirus that was first recognized in the United States in 1999 (Andreadis 2012). and has established across much of the continental United States by exploiting a network of endemic vectors (Kilpatrick et al. 2011). It has been isolated from at least 33 species in eight genera, including several members of the *Culex* genus (Andreadis 2012). A review of regional WNv vectors indicates that in the northeast 96% of WNv positive mosquito pools were *Culex*, including *Cx. pipiens* and *Cx. restuans,* and *Cx. quinquefasciatus* is a common vector in the southeast (Andreadis 2012). These *Culex* species have broad and overlapping ranges in North America, including many major metropolitan areas.

Culex species coexist and form assemblages in the eastern United States, yet factors like land cover and phenology influence how they are distributed across a heterogeneous landscape, yet these patterns may not be universal across land cover types or regions (Ebel et al. 2005, Lee and Rowley 2000, Johnson et al. 2015, Gorris et al. 2021, Arsenault-Benoit and Fritz, *in review*). The roles and contributions of each species to WNv transmission, and therefore their vectorial capacity for WNv in eastern North America, differ. Vectorial capacity is the potential number of new infectious bites per human per day by a vector population (Garrett-Jones 1964), and is mathematically defined by a vector population's abundance, longevity, human biting rate, and extrinsic incubation period and vector competence for a pathogen. Each of

these variables is affected by the vector's life history and physiology, as well as abiotic and biotic features of the environment. For a particular population in a specific location, abundance is dependent on phenology and habitat suitability (Ciota et al. 2011, Gorris et al. 2021), longevity is related to ambient temperature and available refugia from predation or intolerable conditions (Norris et al. 2004, Kramer and Ciota 2015), human biting rate depends on host preference and availability (Fritz et al. 2015, Hamer et al. 2008), and extrinsic incubation period changes with climate (Kramer and Ciota 2015). Vector competence for a particular pathogen is a physiological trait that varies among taxa (Hardy et al. 1983, Ferraguti et al. 2016). Therefore, even if all species in an assemblage are capable of transmitting the pathogen, vectorial capacities for populations within the community differ across space and time.

In the eastern United States, including metropolitan areas, *Cx. pipiens sensu lato* (*s.l.*) is commonly acknowledged as the primary vector of WNv due to competence in the laboratory and the field, willingness to feed on both birds and humans, isolation of WNv from field-collected samples, and spatial and temporal co-occurrence with disease outbreaks (Fonseca et al. 2004, Hamer et al. 2008, Johnson et al. 2015). They are implicated as bridge vectors for WNv because they are opportunistic feeders and will host-switch between birds and mammals, resulting in epizootic transmission of the virus to mammals including humans (Hamer et al. 2008, Fritz et al. 2015). They have a broad geographic distribution globally, and can be found in diverse habitats, including urban and suburban areas (Diuk-Wasser et al. 2006, Gorris et al. 2021).

Culex quinquefasciatus is a member of the *Cx. pipiens* complex that is also globally abundant in southern latitudes (Ciota and Kramer 2013). A stable *Cx. pipiens*

- *Cx. quinquefasciatus* hybrid zone between 36°N and 39°N latitude is well-described (Barr et al. 1957, Kothera et al. 2009, Kothera et al. 2012), and specimens with *Cx. quinquefasciatus* ancestry have been identified as far North as Chicago, IL (~42°N, Kothera et al. 2020), but *Cx. quinquefasciatus* are not regularly included in studies about the spatiotemporal dynamics of *Culex* WNv vectors.

Culex restuans is endemic to North America, has a similar WNv vector competence to *Cx. pipiens* and co-occurs with members of the *Cx. pipiens* complex in the eastern United States (Johnson et al. 2015, Ebel et al. 2005, Harrington and Poulson 2008). As a primarily ornithophilic species using forested habitats, *Cx. restuans* is likely responsible for enzootic transmission of WNv among birds but may not commonly act as a bridge vector (Kilpatrick et al. 2005, Harrington and Poulson 2008, Dunphy et al. 2019). Although they tend to be dominant in natural areas, studies show they have substantial niche overlap with *Cx. pipiens* (Reiskind and Wilson 2008, Arsenault-Benoit, *in review*), and may use disturbed or urbanized habitat as well (Johnson et al. 2015).

Culex restuans is considered an early-season mosquito, with populations emerging early and declining before *Cx. pipiens* reaches peak density, leading to a "crossing-over" point (Lee and Rowley 2000, Andreadis et al. 2001, Johnson et al. 2015). A common inference is that *Cx. restuans* initiates WNv amplification within birds early in the season, allowing for increased transmission to mammals later in the season by *Cx. pipiens* (Andreadis et al. 2001, Johnson et al. 2015). *Culex restuans* may be equally, if not more, important than *Cx. pipiens* for WNv amplification and transmission throughout the season (Johnson et al. 2015), but it is widely held that their human health impact is indirect due to limited mammalian host use (Kilpatrick 2005, Harrington and Poulson 2008). Consequently, studies focusing on the physiology, life history, and habitat use

of *Cx. restuans* are less common than for *Cx. pipiens* (Ebel et al. 2005, Harrington and Poulson 2008). The true interspecific dynamics are likely dependent on geographic location, land use, relative abundance of the two species, and annual fluctuations (Lee and Rowley 2000, Jackson and Paulson 2006, Andreadis and Armstrong 2007, Tokarz and Smith 2019).

Understanding the role of individual *Culex* species in WNv transmission in the eastern United States is further complicated due to the morphological similarities among these taxa, which are amplified if specimens were damaged in trapping. Often, these taxa are pooled in mosquito surveillance efforts or likely to be misidentified (Jackson and Paulson 2006, DeGroote et al. 2008, Johnson et al. 2015), which can obscure important differences in the traits between species that determine their roles in WNv cycling and vectorial capacity. Misidentification and pooling can also lead to unresolved or potentially incorrect assumptions of spatial and temporal relationships between species and the environment (Johnson et al. 2015). For instance, in an analysis of decades of mosquito population data correlated with landscape and climate variables, all cryptic *Culex* species were considered members of the "*Cx. pipiens* group" (DeGroote et al. 2008) so species-specific conclusions are lost. Further, occupancy maps used to describe habitat use of *Culex* have been generated based almost entirely on morphological identification, thus potentially suffering from misidentification. The authors of that study suggest that their habitat models could be improved with molecular identification and investigation of finer-scale community structure (Gorris et al. 2021). Given these limitations, we need spatiotemporal studies of confidently identified *Culex* community interactions.

In this study, we used molecular tools to identify *Culex* specimens over urban to rural gradients from three eastern North American metropolitan areas to assess species composition and turnover over the season, as well as across the landscape, at broad and fine spatial scales. If *Culex* vector networks are not uniform over space and time, enhanced resolution and specificity of habitat use and phenology may enable practitioners to target management efforts and improve prediction of disease outbreaks.

Methods

Mosquitoes were collected by local investigators over the course of the active mosquito season in each of three localities: Maryland (MD), Pennsylvania (PA), and Illinois (IL, Figure 1). Specimens were collected via gravid traps with a hay-infused oviposition attractant (*see supplemental text, Appendix II*) at urban, suburban, and rural sites in Washington D.C. and Prince George's, Montgomery, Anne Arundel, and Howard counties in Maryland ("MD"), in the Northwest Mosquito Abatement District (NWMAD) in Cook County, northwest of Chicago, IL ("IL"), and in Bucks, Delaware, Philadelphia, and Montgomery counties in Pennsylvania ("PA", Table 2). Date and GPS coordinates for each collection were recorded. To coarsely examine the impact of season, we categorized our collection dates into three seasonal periods. Collections from the start of the local season to July 11 were designated as "early season", collections made between July 11 and August 15 were designated as "mid season", and collections made between August 15 and October 1 were designated as "late season." Adult mosquitoes suspected to be *Culex pipiens, Culex restuans,* or *Culex*

quinquefasciatus were separated from non-*Culex* species morphologically and stored at -80°C until molecular identification at the University of Maryland, College Park.

Study Sites

Greater Washington, D.C. and Maryland

The greater Washington D.C. and Maryland specimens used in this study are the same specimens described in Arsenault-Benoit and Fritz (*in review).* Briefly, fifteen collection sites were established in 2019, within and approximately 50 km from downtown Washington, D.C. Sites were categorized as urban (N=5), suburban (N=5), and rural (N=5) using characteristics of primary housing type, zoning, and features and extent of nearby green space. During the summers of 2019 through 2021 traps were set at one randomly selected urban, suburban, and rural site on a given night. Collections occurred between June 1- October 1 (week of the year (WOY) 22-40) for eight collections per year in 2019-2020, and April 4-October 1 (WOY 14-40) for eleven collections in 2021. Mosquitoes were transported in an ice-chest to the University of Maryland, where they were cold-anesthetized at -20°C for 30 minutes or more, morphologically identified to the lowest taxon possible via microscopy, and stored at -80°C until molecular identification to species.

We used quantitative environmental data from these site classifications (Appendix II, Table 4.1, Arsenault-Benoit and Fritz *in review*), along with satellite imagery to assign IL and PA sites to urban, suburban, and rural categories reflective of those developed for MD. Site classifications were used for exploratory analysis, but the

majority of analyses was performed using only quantitative continuous variables that could be extracted for all sites.

Greater Chicago, IL

In IL, nine sites regularly surveyed by the NWMAD, including three urban, three suburban, and three rural sites, were selected and established for this study. Sites were selected and classified according to GIS data, satellite imagery, and visual inspection by NWMAD personnel. Sites were approximately 50 km north and west of downtown Chicago, IL, but encompassed a 350 km^2 region in the northwestern Cook County, IL. All sites were at least 30 km from Chicago. Specimens were collected biweekly in gravid traps over eight occasions per year from June to October in 2020 and 2021 (WOY 22-40).

Greater Philadelphia, PA

In PA, specimens were collected as part of vector surveillance and management activities by the Pennsylvania Department of Environmental Protection Division of Vector Management (PA DEP). These individuals were set aside from normal WNv surveillance collections as permitted by the DEP protocol. In 2019, individuals collected from gravid traps in Bucks, Montgomery, Delaware, and Philadelphia Counties, PA were retained on a weekly basis from May 23, 2019- August 20, 2019 (WOY 21-34). The number of individuals within each collection varied, with a minimum of ten per trap. In 2020 and 2021, COVID-19 restrictions and changes to resource allocation limited collections to Philadelphia and Bucks counties, where up to

10 samples were retained per county per month from May through September (WOY 18-35), for a total sample size of approximately 100 individuals each of those two years.

Molecular Identification

Specimens were molecularly identified as described in Arsenault-Benoit and Fritz (*in review*). Briefly, genomic DNA from MD samples was extracted using one of three methods: a rapid alkaline-based extraction from the abdomen or legs (Rudbeck and Dissing 1998) or commercially available DNA isolation kits (Qiagen DNeasy Blood and Tissue Kit (Cat. No. 69506, Qiagen Inc., Valencia CA, USA or Zymo Quick-DNA Miniprep Plus kit (Cat #11-397B, Zymo Research, Irvine, CA)) from the abdomen (excluding female spermathecae) in 2019 and 2020; or a rapid macerating procedure from the abdomen in 2021 (Gloor and Engels 1991). Genomic DNAs from IL and PA were all isolated using the rapid macerating procedure. Extracted DNAs were amplified with a multiplexed polymerase chain reaction (PCR) targeting the 28S ribosomal subunit to identify *Culex pipiens s.l.* and *Culex restuans* as described by Crabtree et al. (1995) and Rochlin et al. (2007). This PCR also identifies *Culex salinarius,* but because gravid traps are biased against this species (Williams and Gingrich 2007), we likely did not collect a representative sample of *Cx. salinarius* across the landscape. Therefore, any detected were excluded from analysis (N=18). Specimens identified as *Cx. pipiens s.l.* underwent a second PCR reaction to identify *Cx. quinquefasciatus* ancestry at the ACE2 locus (Smith and Fonseca 2004).

PCRs were conducted according to the previously described protocols (Crabtree et al. 1995, Smith and Fonseca 2004), with slight modifications (see Supplemental Methods, Appendix I). Two negative controls, a PCR negative in which purified water replaced genomic DNA, and an extraction negative, where the DNA extraction process was conducted without a mosquito carcass, were always PCR amplified and electrophoresed alongside mosquito samples. Thermocycler conditions followed Crabtree et al. (1995) and Smith and Fonseca (2004), respectively, a using Bio-Rad T100 Thermal Cycler (Bio-Rad Life Sciences, Hercules, CA). Amplicons and controls were visualized on an agarose gel (2% w/v) alongside a 1Kb ladder (GeneRuler 1Kb Plus, ThermoScientific, Waltham MA) following gel electrophoresis at 120 v for 60 minutes, and identifications were made according to taxon-specific amplicon sizes.

Because the PA dataset was considerably smaller than the other two localities, we simulated ten datasets of PA size from the MD data by random draw to assess the risk of rare species in the environment going undetected in our study.

Quantification of Landscape Variables

Ecological variables for each site were measured within a buffer zone with a 0.5 km radius using GIS tools in ArcMap 10.7. This buffer size was selected because the mean mark-recapture distance for this species is 0.9 km, (Hamer et al. 2014), suggesting a typical range of nearly 0.8km^2, and this buffer size results in a coverage area of similar size. Environmental variables include percent impervious surface and percent tree canopy cover (NCLD 2016), as well as population density and housing

unit density (2010 US Census), linear distance to the geographic center of the focal city (Washington D.C., Philadelphia, or Chicago, km), water table depth (cm), and Normalized Differential Vegetative Index (NDVI). The mean of values for pixels within the buffer for each layer were calculated for each site (Table 2).

Quantification of Climatic Variables

Daily climate data was extracted from the DAYMET database using R package `daymetr` (Hufkens et al. 2018) for each individual trap's geographic coordinates and date. Then, climate variables were aggregated across weekly time lags to determine the mean for each variable one to seven days (one week) and eight to 14 days (2 week) preceding the collection date. Climate data included day length (seconds/day), daily precipitation (mm), and maximum temperature (°C). Climatic conditions over these time lags have been cited as reliable predictors of *Cx. pipiens-restuans* populations and WNv cases (Chuang et al. 2011, Lebl et al. 2013).

Analysis

All statistical analyses were completed in R (v.4.2.2). Counts of *Cx. pipiens, Cx. restuans*, and *Cx. quinquefasciatus* identified were compared across site classes (urban, suburban, or rural) and season (early, mid, late) at each locality separately then overall using either a Fisher Exact test when cell counts were low (IL and PA), or Pearson's chi-squared test when expected cell counts were greater than 10 (MD and aggregated).

We then used gradient forest (GF), a regression tree-based analysis, to model community assemblage turnover across environmental gradients using the landscape, social, and climatic variables in Appendix II, Table 4.2. However, specimens identified as *Cx. quinquefasciatus* homozygotes were rare or absent in northern localities (IL: N=15, PA: N=0, Table 2, Figure 2) so they were excluded from GF analysis. Gradient forest is an extension of random forests that generates an ensemble of regression trees for each species in a community, then aggregates results to explore associations between environmental gradients and community assemblages (Ellis et al. 2012, Pitcher et. al. 2012). We first modeled species abundance turnover, or where, along environmental gradients, changes in relative abundance occur, of cryptic *Culex* over gradients of environmental variables using the aggregated dataset, then at each locality independently using the R package `gradientForest` (Ellis et al. 2012). We assessed relative abundance and species turnover at each locality and the overall dataset to determine which environmental variables are the most important predictors of species abundance, as well as where along each gradient compositional turnover occurs.

Results

Species distributions by site class and season

We identified 2,461 specimens from MD, 706 from IL, and 303 from PA (total N=3,650). In MD and IL, *Cx. pipiens* were dominant across all three site classes, but in PA *Cx. restuans* were more abundant than *Cx. pipiens* in rural and suburban sites (Table 2, Figure 2). *Culex quinquefasciatus* were more common in MD than IL, and none were identified in PA. Ten iterations of simulated random samples from the MD

dataset of equivalent size to the PA dataset suggests that if *Cx. quinquefasciatus* were present in the environment, we would likely have detected them. From each random draw of 300 individuals, between 7 and 16 *Cx. quinquefasciatus* were projected to be in the sampled population (mean±SEM: 13.1±0.82).

Season, site class, and locality affect cryptic *Culex* species abundance. Counts of individuals of each species were significantly different across site classes in MD (X^2=258.21, df=4, p<0.0001) and PA (Fisher's Exact Test, p=0.0006), but not IL (Fisher's Exact Test, p=0.5214), suggesting that site class has limited influence on the community assemblage at our IL sites. Species counts were also unevenly distributed across site classes when aggregated across localities (X^2=154.26, df=4, p<0.0001, Figure 2). In all localities, *Cx. restuans* were more abundant than *Cx. pipiens* in the early season, and crossing-over in species abundance occurred at all localities but at different time points that correspond to latitude (Week of the Year (WOY) 20 in MD, WOY 26 in PA, and WOY 28 in IL, Figure 2). Analysis of species counts across early, mid, and late season suggests that abundances differ across the season in all localities individually and combined (MD: X^2=99.852, df=4, p<0.0001, IL: Fisher's Exact Test, p<0.0001, PA: Fisher's Exact Test, p<0.0001, aggregated: X^2=235.73, df=4, p<0.0001).

Gradient Forest Analysis

Because *Cx. quinquefasciatus* were rare or absent in the northern localities, we excluded them from the gradient forest analyses. We created a forest of 500 regression trees for each of the two species. Underlying random forest models to build the gradient

forest explained about 23% of variation for *Cx. pipiens* (R^2_f=0.23) and 28% of variation for *Cx. restuans* (R^2_f=0.28).

Climatic variables

Phenology, represented as week of the year (WOY), was consistently one of the most important predictors of compositional change between *Cx. pipiens* and *Cx. restuans* when modeled for the aggregated dataset and for each locality independently (Figure 3, Table 1). The locations and density of decision tree splits, standardized by observation density, in the overall dataset suggested that compositional turnover occurred between weeks 22-28 (June 1-July 11, Figure 4), but when the GF model was run independently for each locality, the window for compositional turnover tended to narrow and shift depending on latitude (MD: WOY 21-25, PA: WOY 22-28, IL: WOY 25-31, Figures S1, S2, S3). In each case, these compositional shifts reflected and were consistent with the crossing-over points of mean species counts per trap night observed at each locality (Figure 2, Figure 5A-B). These results were consistent with compositional changes along a gradient of mean daylength both one week and two weeks prior to collection dates. Changes occurred between 14.5 hours and 15 hours of daylight in each of the preceding two weeks (Figure 4). The abundance of *Cx. restuans* decreased considerably during the compositional turnover periods in each locality, whereas *Cx. pipiens* increased, but the rate of change in abundances was more gradual than for *Cx. restuans* (Figures S1, S2, S3).

The mean daily maximum temperature days 8-14 prior to collection was more important to species composition than days 1-7 prior to collection (Figure 3). Many

important splits occurred between 22-27℃ two weeks prior to collection; abundance of *Cx. pipiens* changed most under those conditions (Figure 4). One week prior to collection, compositional turnover occurred between 18-24℃. *Culex restuans* abundance decreased when the temperatures were between 18-20℃, and *Cx. pipiens* increased in abundance at 24℃ and above (Figure 5C-D). The abundance of *Cx. pipiens* was mediated by precipitation more than *Cx. restuans*, and the mean precipitation 1-7 days prior to collection was more important than mean precipitation 8-14 prior to collection (Figures 3&4). During periods of limited mean daily precipitation (0-5 mm) one or two weeks prior to collection, abundance of *Cx. pipiens* was highest, and abundance decreased when lagged precipitation in the week(s) preceding collection was greater (Appendix II, Figure 4.5). *Culex restuans* populations remained consistent regardless of precipitation. Both species experienced changes in abundance over the gradient of water table depths (Appendix II, Figure 4.5). The most important splits on this gradient are between 35-65 cm and 75-85 cm (Figure 4). *Culex restuans* abundance increased at water table depths of 35-45 cm and decreased at higher depths. Conversely, *Cx. pipiens* abundances increased when water table depths were above 75 cm (Appendix II, Figure 4.5). This result may be influenced by the fact that only MD sites had water table depths greater than 50 cm (Appendix II, Table 4.1).

Landscape Variables

Relative abundance of *Cx. pipiens* and *Cx. restuans* shifted in consistent and predictable ways along gradients that are characteristic of urbanization, including percent tree cover, percent impervious surface, distance to city center, population

density, and vegetation index (NDVI). The greatest compositional change occurred between 45-60% tree cover and 20-50% impervious surface (Figure 4). *Culex pipiens* abundance decreased with increasing tree cover and increased with increasing impervious surface (Appendix II, Figure 4.5E-H). *Culex restuans* had the inverse relationship with both variables. Compositional changes occurred at two distances from the center of each focal city (DCC). *Cx. pipiens* were abundant close to city center and decreased slightly in abundance between 0-15 km DCC, whereas *Cx. restuans* increased in abundance 20-30 km DCC. Vegetation index (NDVI) had low relative importance to compositional turnover compared to other variables in the overall dataset and in the MD and IL localities. In PA, however, NDVI was the 6th most important predictor, likely because a broader range of NDVI values was represented at that locality (Table 2, Appendix II, Figure 4.2). In PA, the abundance of *Cx. restuans* increased with increasing NDVI and *Cx. pipiens* decreased with increasing NDVI, with significant compositional change occurring at values of 80-100 (Appendix II, Figure 4.2). *Culex restuans* increased with increasing NDVI at all localities, but in MD and IL, *Cx. pipiens* abundance was consistent across values of NDVI (Figures S1, S3, S5). Compositional changes occurred at NDVI of 85-100 in MD and 90-105 in IL, but rates of change were gradual (Appendix II, Figure 4.3). The abundances of both species also changed at the highest population and housing densities, with *Cx. pipiens* associated with high human density and *Cx. restuans* associated with low human density (Appendix II, Figure 4.5).

Discussion

Across urban to rural gradients in metropolitan areas in the eastern United States, cryptic *Culex* species abundances and compositional turnover were influenced by season, climate, and characteristics of the landscape. Generally, *Cx. pipiens* was more abundant in our collections than *Cx. restuans*, except in rural and suburban sites in PA. *Culex quinquefasciatus* was rare in our collections overall, and exceedingly rare or undetected outside of MD. Collections of *Cx. quinquefasciatus* in IL were sporadic, and we posit individuals with *Cx. quinquefasciatus* alleles as detected by the ACE2 locus are not present in the greater Philadelphia metropolitan area. In Arsenault-Benoit and Fritz (*in review*) we found that *Cx. quinquefasciatus* were present in low abundances but in consistent populations at the most urban sites in Washington D.C. and MD, but we did not identify enough individuals in IL or PA to gain a deeper understanding of the spatiotemporal dynamics of that species across a broader spatial scale. Future studies should include additional collections from localities within and south of the established hybrid zone to determine whether *Cx. quinquefasciatus* is limited to urban sites elsewhere or if the lack of thermal limitations in southerly latitudes enables them to be as cosmopolitan as *Cx. pipiens*.

Early in the collection season, *Cx. restuans* was more abundant than *Cx. pipiens* at all localities, and later increases in the relative abundance of *Cx. pipiens* resulted in a crossing-over point that is similarly well-documented in the literature (Lee and Rowley 2000, Andreadis et al. 2001, Johnson et al. 2015, Tokarz and Smith, 2019). Recognizing and documenting this crossover period is important to public health because the timing and duration of crossing over between *Cx. pipiens* and *Cx. restuans* precedes and predicts the timing and intensity of human WNv cases (Lampman et al.

2006, Tokarz and Smith 2019, Bondo et al. 2023). Unsurprisingly, the timing of crossing over differed based on latitude, likely due to the inherent associations between latitude and critical environmental cues such as temperature. *Culex pipiens* and *Cx. restuans* both undergo diapause as an overwintering strategy in temperate latitudes, which is common among exothermic organisms (Denlinger and Armbruster 2014). Timing of emergence from diapause by each species is likely responsible for the phenological trends observed. Diapause in *Culex* is initiated in the autumn by a combination of short daylength and cooler ambient temperatures (Yu et al. 2018), but conditions that trigger emergence from diapause in the spring are less understood. Emergence may be associated with depletion of nutritional resources in overwintering adults, but the primary predictor is likely temperature (Ciota et al. 2011). Crossing over occurred in mid-May in MD, early to mid-June in PA, and mid- to late June in IL. Between the vernal equinox and the summer solstice, photoperiod is longer at progressively higher latitudes, but we observed that crossing over occurs later at higher latitudes, suggesting that longer days are not independently responsible for emergence. Since *Culex* use protected hibernacula for overwintering, exposure to daylight is limited (Ciota et al. 2011). This is especially true for *Cx. pipiens* that use below ground refugia or human infrastructure for overwintering compared to *Cx. restuans* that use natural features (Ciota et al. 2011, Arsenault-Benoit et al. 2021). Daylength was consistently an important predictor, but the daylength range in which turnover occurred was the same across localities (14.5-15 hours). Thus, daylength was probably an important predictor due to its direct correlation to latitude but photoperiod may not inherently affect species composition in the early season.

Temperature 8-14 days preceding collections had similar importance to compositional turnover as phenology (WOY). Maximum temperature 1-7 days prior to collection had less, but still high-ranking, importance. Ambient temperature is important to mosquito survival, development, longevity, and emergence from diapause (Townroe and Callaghan 2014, Ciota et al. 2014, Gorris et al. 2021). Cooler temperatures at higher latitudes may contribute to higher *Cx. restuans* relative abundance at northerly localities generally. *Culex restuans* develops more quickly at lower temperatures than other *Culex* species and have higher mortality and reduced longevity at temperatures over 25°C (Buth et al. 1990, Ciota et al. 2014). Conversely, *Cx. pipiens* immature stages develop more quickly and have improved longevity at higher temperatures. At temperatures over 24°C, *Cx. pipiens* and *Cx. restuans* develop through aquatic stages at similar rates, but *Cx. restuans* experience higher mortality and reduced longevity compared to *Cx. pipiens.* Additionally, *Cx. pipiens* can continue to produce eggs at exceedingly high temperatures (>32°C) which leads to rapid population growth at the height of the summer (Ciota et al. 2014). Therefore, *Cx. restuans* has a negative relationship to rising summer temperatures, and *Cx. pipiens* has a positive relationship to rising temperatures, which aligns to our findings across latitudes. Our model suggests the transition in dominant species occurs approximately two weeks after mean daily maximum ambient temperatures reach 24°C. Allowing for development through aquatic stages, this means crossing over in oviposition and survival rates between species happens when maximum average daily spring and summer conditions rise to the temperatures between 20°C and 25°C.

The emphasis on temperature for spring species emergence, habitat use, and crossing over suggests that warming due global climate change could have major implications on WNv dynamics that are challenging to predict. Because the timing of crossing over is a predictor of the timing and intensity of human WNv cases, earlier emergence and population growth of *Cx. pipiens* could extend the transmission season and increase disease risk. However, higher ambient temperatures could extirpate *Cx. restuans* or lead to even starker habitat delineation between species. Consequently, overall avian WNv amplification in the environment could be reduced, or less spatial and temporal overlap between *Cx. pipiens* and *Cx. restuans* could reduce the risk of epizootic spillover.

Species also had divergent responses to mean precipitation in the weeks preceding adult collection. Once eggs are oviposited, *Culex* mosquitoes require persistent standing water for the duration of their development through aquatic stages, which takes several days. We found that *Cx. restuans* was most abundant when precipitation was moderate, and this species was associated with a shallower water table. This observation is congruent with *Cx. restuans'* use of forested habitats and preference for natural rather than anthropogenic ovipositional resources (Gorris et al. 2021). Moist and permeable ground conditions may be important for their development and survival through aquatic stages. Conversely, *Cx. pipiens* abundance was high when mean daily precipitation one to two weeks before collection was between 0-5 mm, suggesting that heavy rains are not required for *Cx. pipiens* oviposition and development success. They are also more abundant when the water table is deeper, meaning groundwater is less accessible. Taken together, these results are in accordance

with a life history strategy based on using containers, artificial water sources, and impervious surfaces as breeding and development habitat. Willingness to oviposit in anthropogenic water sources, such as landscape irrigation or puddles formed on uneven concrete could allow *Cx. pipiens* populations to persist in urban areas, where we did observe them in higher relative abundance. Conversely, *Cx. restuans* require more naturally occurring water sources. Forested canopy may allow for natural water sources to persist longer due to high relative humidity and reduced evapotranspiration shallow, still, water sources they prefer for oviposition and development of immature stages.

Habitat use is determined by suitable conditions for both adult and aquatic stages; although each life stage requires different resources, they cannot be completely decoupled from each other (Chandrasegaran et al. 2020). Across localities, characteristics related to urbanization were linked to compositional turnover between *Cx. pipiens* and *Cx. restuans*, although species abundances were relatively consistent across IL sites. We note that IL sites were more environmentally homogeneous than the other localities across site classes, so this observation may be a result of sampling design rather than actual local variation in species distributions in IL. IL sites were initially selected to fit within the criteria for site classes established in MD, but when we described the landscape quantitatively using GIS tools, we discovered that the range of differences in quantitative variables between site classes was narrower in IL and broader in PA and MD. So, we may not have fully described the variation in *Culex* species abundances across the landscape in IL, as we may not have fully captured the heterogeneity of the landscape.

Relative abundances of these species turned over at intermediate impervious surface, tree cover, distance to city center, human population, and NDVI values, suggesting that suburban landcover serves as an interface for compositional turnover between species (Figure 3). *Culex pipiens* also benefits from higher temperatures, which are common at more southerly latitudes, but on a finer scale, also tend to occur in city centers due to the urban heat island effect (Ladeau et al. 2015).

In Arsenault-Benoit and Fritz (*in review*) we found that in MD, *Cx. restuans* were associated with rural and suburban areas, and especially those with limited disturbance and heterogeneous canopy cover, whereas *Cx. pipiens* were cosmopolitan and could reach relatively high abundances in a variety of habitats. Here, we found that *Cx. restuans* are associated with cooler temperatures and naturally occurring ovipositional resources, which is congruent with rural and suburban forested sites. *Culex pipiens* appear to have broad habitat tolerances, which likely contributes to their global ubiquity and their role as disease vectors. Their use of human infrastructure for diapause and refugia (Ciota et al. 2011, Arsenault-Benoit et al. 2021), impervious surface and containers for breeding and development (Duik-Wasser et al. 2006, Townroe and Callahan 2014), warmer temperatures, and a diversity of bloodmeal hosts (Kilpatrick et al. 2005, Hamer et al. 2009) means they are well-suited to urban environments, especially in comparison to *Cx. restuans*.

Using molecular identification and collections across an urban to rural gradient in disparate localities, we were able to describe phenology and habitat preference of these cryptic *Culex* species in eastern North America. We determined that *Cx. restuans* are indeed an earlier season species, and that crossing over between *Cx. restuans* and

Cx. pipiens occurs throughout their sympatric range, but the timing of crossing over differed according to latitude and associated ambient conditions. However, niche segregation between species across the landscape is apparent. Natural areas with high tree cover, low impervious surface, and low human population density are most suitable for *Cx*. *restuans*, and *Cx. pipiens* tend to dominate more developed landscapes. This finding is consistent with the results of a community assemblage analysis undertaken in MD (Arsenault-Benoit and Fritz, *in review*), as well as occupancy models for individual species (i.e. Gorris et al. 2021).

A community ecology approach to understanding and managing WNv is imperative due to the importance of interspecies dynamics to its transmission cycle. Each of the species discussed in this study are competent WNv vectors and contribute to transmission, but populations' vectorial capacity for WNv varies. If practitioners intend to target *Cx. pipiens* specifically, management should be focused in urban and suburban areas after the crossing over point in their region, which can be estimated using mean temperatures and daylength. However, WNv is a highly networked pathogen, and a diversity of species play a role in maintaining, amplifying, and transmitting the pathogen over space and time (Andreadis et al. 2012, Yee et al. 2021), so designing management efforts around a single species may not be optimally effective. Further, abatement districts are required to satisfy a number of interests, including vector borne disease prevention and management of nuisance mosquitoes. Due to resource constraints and a responsibility to local communities, local abatement districts often begin their seasonal management work with the onset of nuisance mosquito complaints and focus on areas of high human density, and some localities

face legal or social limitations regarding where and when they can use insecticide products (Brain Prendergast and Patrick Irwin, personal communication). We propose that management strategies that reflect the ecology and life history of all species in a network can lead to efficient and effective mitigation. For example, even a small population of WNv infected *Cx. restuans* emerging in the early spring far from human population density could have a major downstream impact of annual WNv cases due to their initiation of viral transmission and amplification in the environment. *Culex* surveillance and WNv testing in natural areas early in the season and beginning abatement efforts at or before the local crossing over point may aid in predicting and mitigating subsequent local cases later in the summer when human cases are common (McMillan et al. 2020, Karki et al. 2020).

Chapter 4: Tables

Table 1. Overall ranked conditional importance of each environmental predictor variable for predicting distributions of *Cx. pipiens* and *Cx. restuans* across gradients. Descriptions of abbreviations used are in Appendix II, Table 4.2

Rank	PA	IL	MD	Overall
1	Tmax.1	WOY	WOY	WOY
2	Tmax.2	Tmax.2	DayLength.2	X
3	Y	DayLength.1	X	Tmax.2
4	WOY	DayLength.2	Tmax.2	DayLength.1
5	DCC	prcp.2	DayLength.1	DayLength.2
6	NDVI	Tmax.1	WaterTbl	Tmax.1
7	%Imp	prcp.1	DCC	%Tree
8	%Tree	Hous	Y	WaterTbl
9	prcp.1	Y	Tmax.1	prcp.1
10	prcp.2	NDVI	%Tree	%Imp
11	DayLength.2	X	%Imp	Y
12	DayLength.1	%Imp	prcp.1	DCC
13	WaterTbl	%Tree	NDVI	Hous
14	X	Pop	Hous	Pop
15	Hous	WaterTbl	Pop	NDVI
16	Pop	DCC	prcp.2	prcp.2

Table 2. Species counts and environmental predictor summaries by locality and site classification. Continuous environmental predictors are represented as either variable ranges (minimum-maximum) or mean ± standard error. WOY=Week of the Year, DCC=Distance to city center, NDVI= Normalized Differential Vegetation Index.

Locality	**Class**	***Cx. pipiens***	***Cx. quinque-fasciatus***	***Cx. restuans***	**Latitude (°N)** (Range)	**Longitude (°W)** (Range)	**WOY** (Range)	**DCC (km)** (Mean±SE)	**Impervious Surface (%)** (Mean±SE)	**Tree Cover (%)** (Mean±SE)	**Housing Units** (Mean±SE)	**Population Density** (Mean±SE)	**Water Table (cm)** (Mean±SE)	**NDVI** (Mean±SE)	**Daily Max Temp (°C)** (Range)	**Daily Precipitation (mm)** (Range)
MD	Urban	704	80	82	38.85; 38.93	-77.03; -77.01	15-40	3.4 ± 0.17	60.5 ± 0.81	0.99 ± 0.1	3241 ± 505.6	7312 ± 1177	76.2 ± 1.2	75.8 ± 0.6	18.9-35.5	0-43.0
	Suburban	611	18	299	38.91; 39.11	-77.00; -76.77		19.4 ± 0.54	19.4 ± 0.57	35.5 ± 0.2	209 ± 17.9	340 ± 15.8	53.1± 2.0	76.6 ± 0.5	17.3-35.5	0.0-35.1
	Rural	388	10	269	38.86; 39.26	-76.93; -76.55		29.7 ± 0.86	1.8 ± 0.73	37.3 ± 2.4	24.23 ± 3.6	80.7 ± 12.1	52.4 ± 1.6	106 ± 1.2	18.2-35.2	0.0-30.5
PA	Urban	46	N/A	26	39.93; 40.25	-75.66; -75.12	19-36	8.4 ± 3.00	67.0 ± 4.02	5.6 ± 2.3	58.6 ± 10.8	156 ± 47.0	13.8 ± 0.8	18.3 ± 5.5	24.5-35.2	0.0-26.4
	Suburban	58	N/A	84	39.84; 40.45	-75.45; -74.82		25.8 ± 1.84	28.2 ± 1.61	19.8 ±1.3	83.4 ± 13.9	255 ± 49.1	17.8 ± 1.9	73.3 ± 2.1	19.8-34.9	0.0-157.4
	Rural	31	N/A	57	39.89; 40.37	-75.48; -74.91		29.9 ± 3.27	13.4 ± 2.39	31.3 ±3.5	32.9 ± 10.8	74.6 ± 23.7	12.5 ± 2.6	94.4 ± 3.5	18.6-35.0	0.0-46.8
IL	Urban	173	4	35	42.01; 42.14	-87.99; -87.92	23-38	36.7 ± 0.75	70.2 ± 1.38	0.35 ±0.02	0.66 ± 0.16	1.00 ± 0.23	16.3 ± 1.3	71.2 ± 1.7	20.2-33.6	0.0-25.6
	Suburban	235	7	64	42.03; 42.14	-88.14; -87.92		40.5 ± 0.62	42.5 ± 1.02	5.7 ± 0.5	77.0 ± 6.47	196 ± 19.2	16.5 ± 1.6	88.7 ± 1.9	20.5-33.6	0.0-35.1
	Rural	167	4	51	42.05; 42.1	-88.18; -88.05		48.3 ± 0.45	7.6 ± 1.33	17.4 ± 0.5	64.5 ± 13.3	142 ± 29.2	18.4 ± 0.2	104 ± 1.3	20.4-33.0	0.0-25.6

Chapter 4: Figures

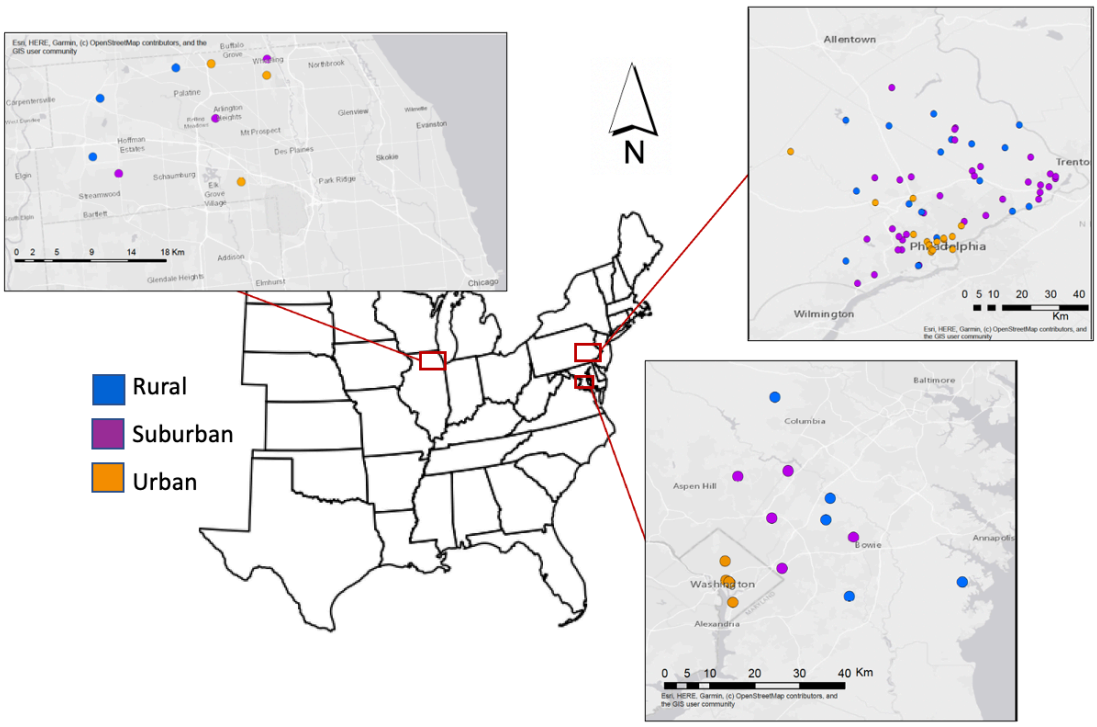

Figure 1. Collection sites in three metropolitan localities in the eastern United States: Chicago, Illinois (upper left), Philadelphia, Pennsylvania (upper right), and Washington, District of Columbia, including Maryland suburbs (lower right). Rural sites are in blue, suburban sites are in purple, and urban sites are in orange.

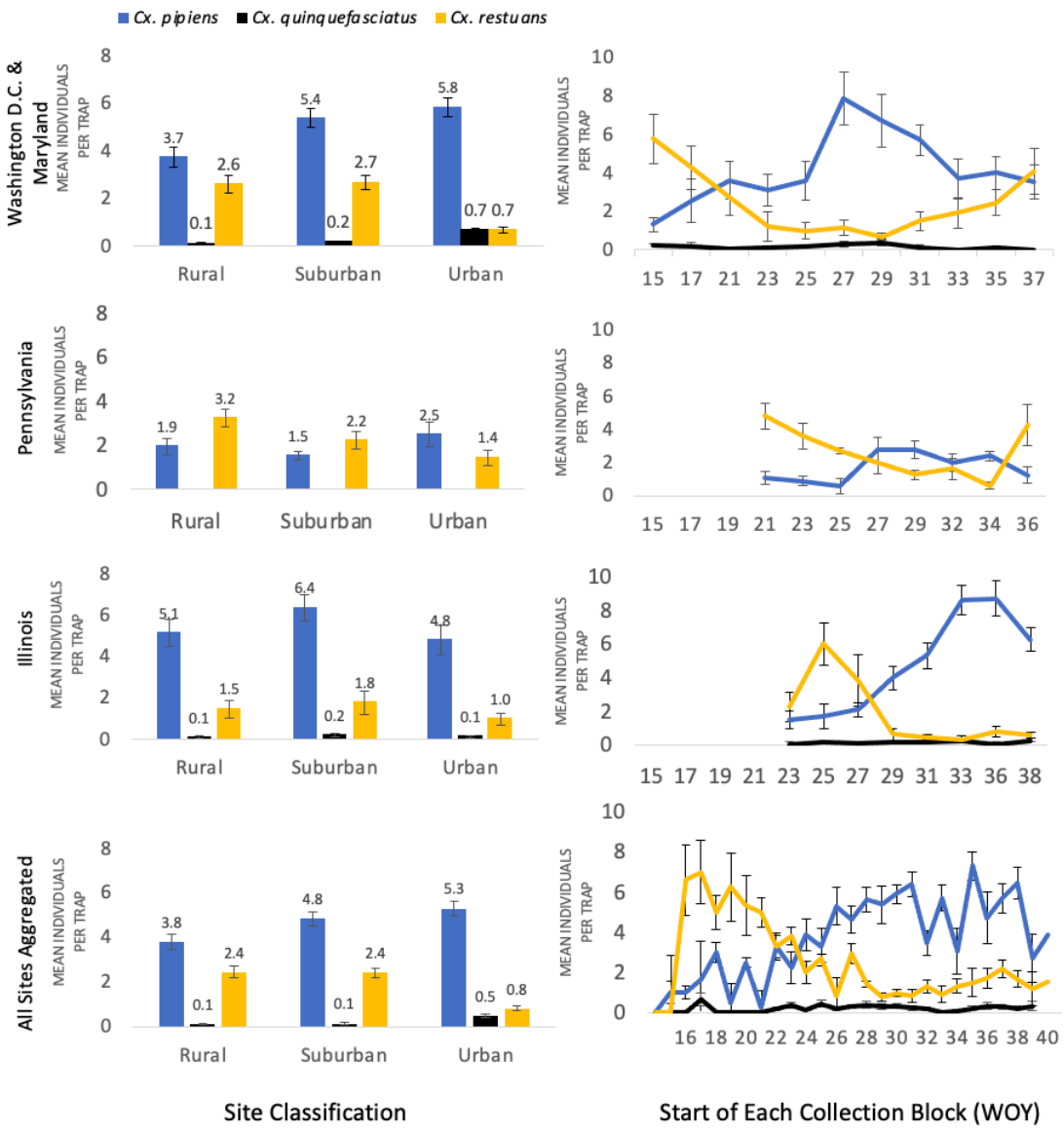

Figure 2. Summary of mean species counts (±S.E.M.) per trap by site class (left) and season (right) represented from week of the year (WOY) 15-40 and separated by locality. Localities are arranged from southernmost (MD) to northernmost (IL), and aggregate data is in the bottommost panel. In each panel, *Cx. pipiens* are in blue, *Cx. restuans* are in gold, and *Cx. quinquefasciatus* are in black.

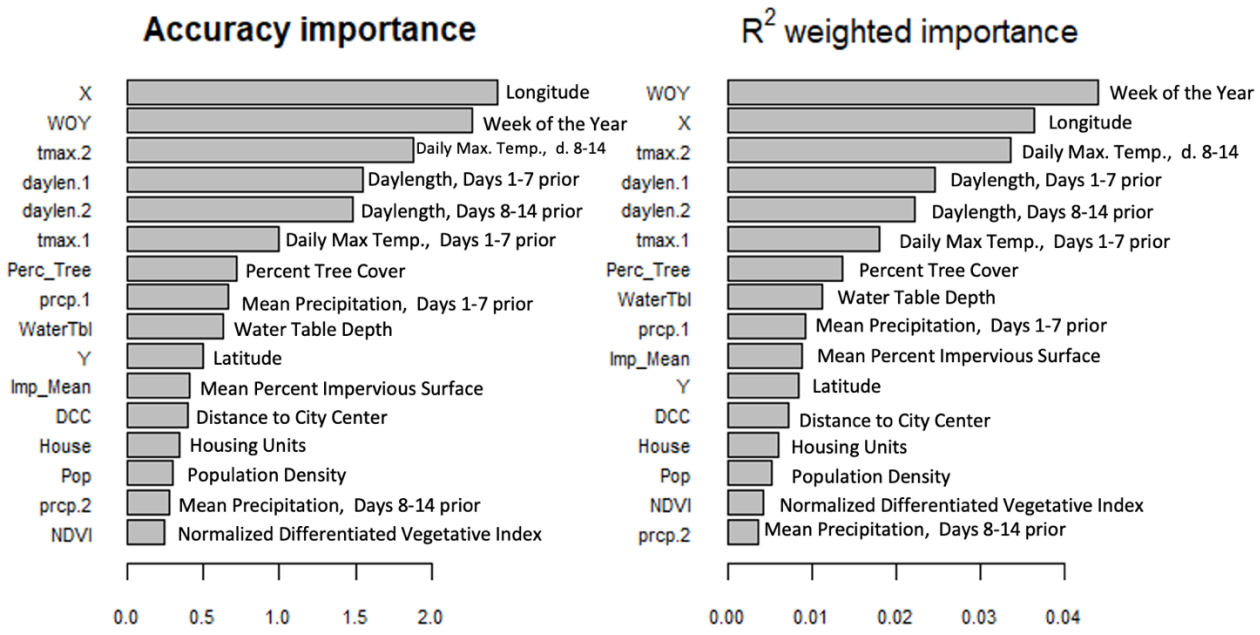

Figure 3. Ranked conditional importance of predictor variables on composition of *Cx. pipiens* and *Cx. restuans* for all localities aggregated. A. Importance of each predictor to model accuracy. B. Relative importance of each predictor variable weighted by species. Detailed descriptions of predictors are in Appendix II, Table 4.2.

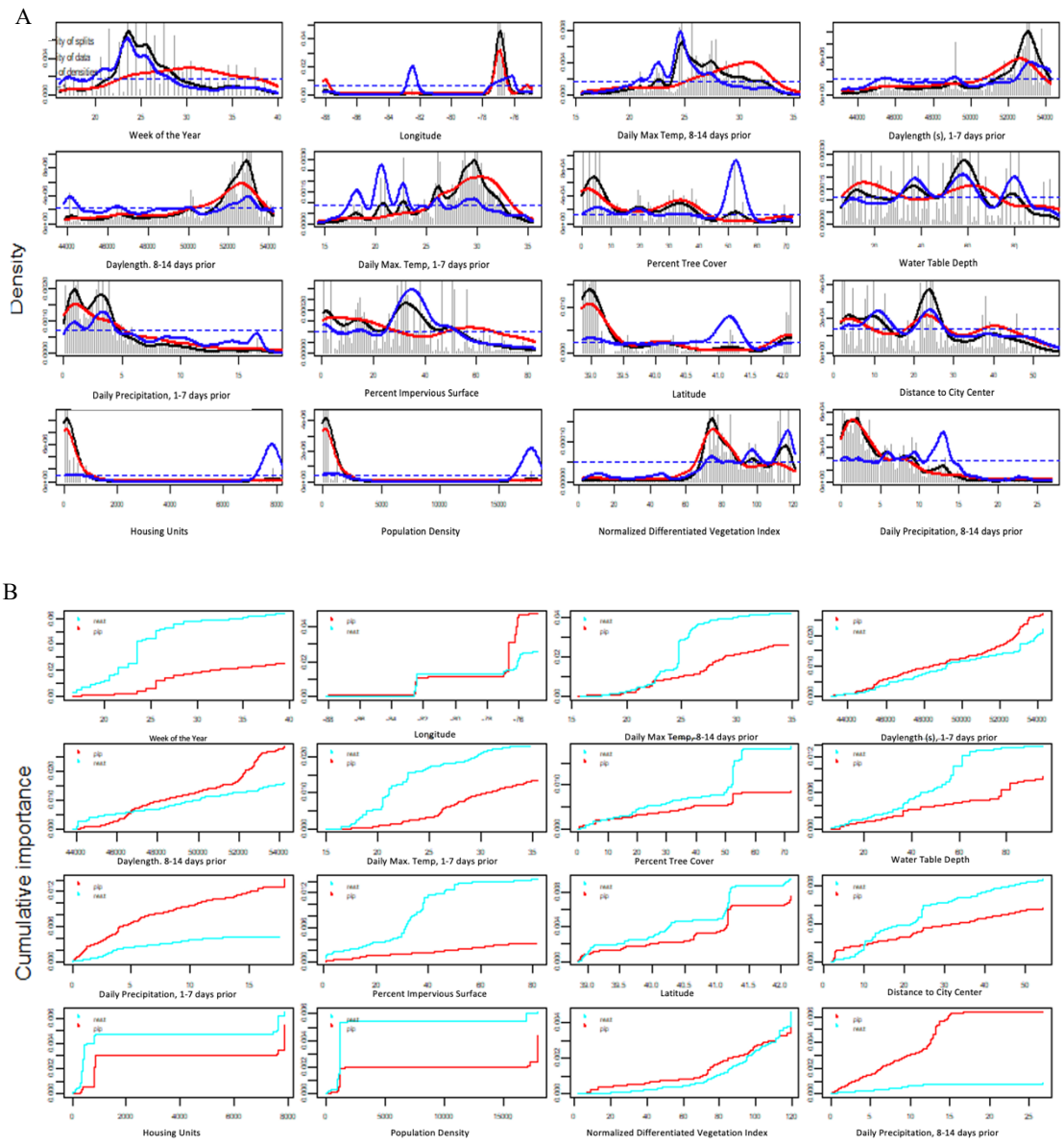

Figure 4. Gradient forest (GF) results for full dataset aggregated across localities. A. Location, density, and importance of GF splits along gradients of the top-ranked 16 environmental predictors. Density of split points are in black, and density of observations are in red. The ratio of these densities are in blue. Locations where the ratio of densities is greater than 1 (blue dotted line) indicates regions along the predictor gradient with greatest relative change in species assemblage. B. Cumulative distributions of standardized splits importance for each species. Monotonic functions for *Cx. restuans* are in teal and *Cx. pipiens* are in red. Shallow slopes indicate low rates of change for each species along the gradient, and steep slopes indicate higher rates of change.

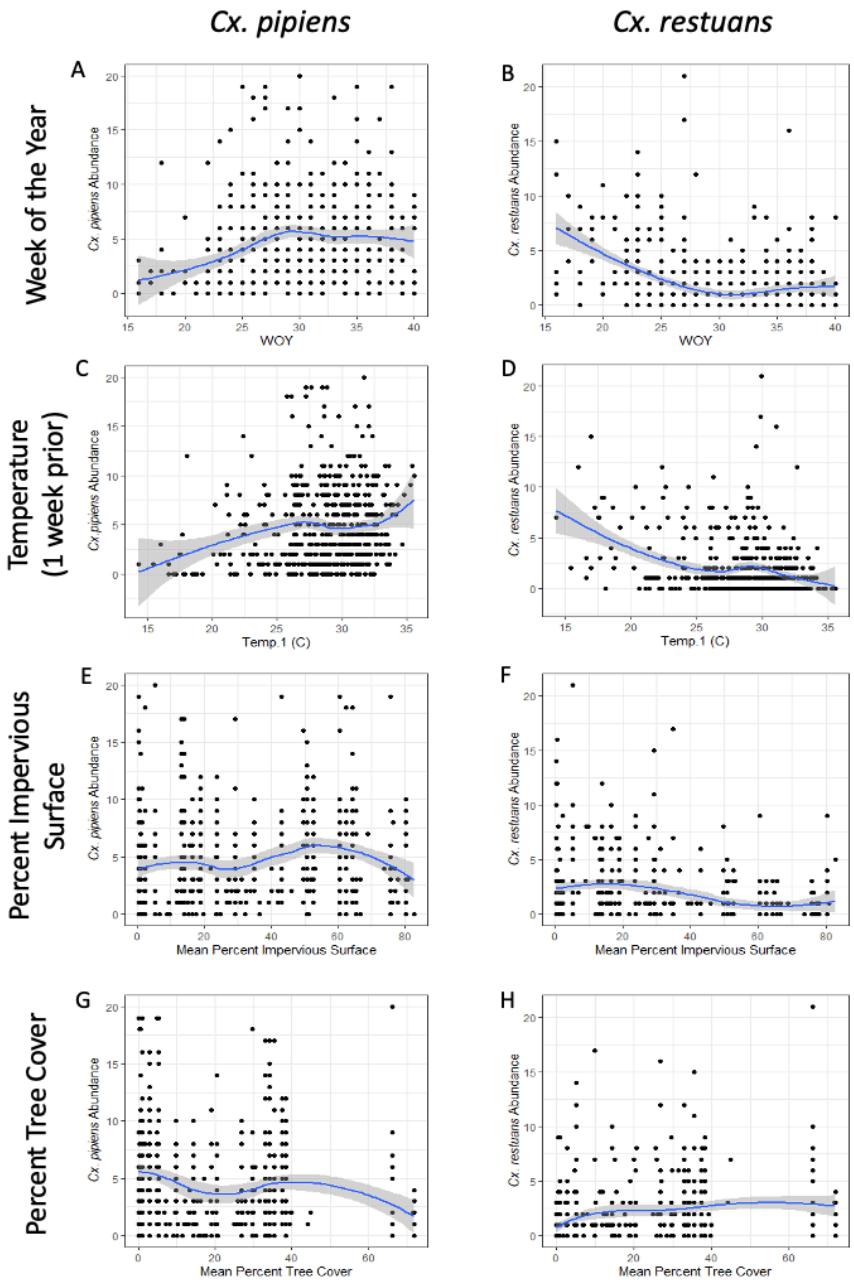

Figure 5. Species abundances in response to a subset of highly ranked environmental variables. *Culex pipiens* abundances are on the left, and *Cx. restuans* abundances are on the right. Points represent species abundances across each environmental gradient. Trends are shown using local regression with locally estimated scatterplot smoothing (*loess*) and 95% confidence interval. A-B. Species trends along week of the year (WOY). C-D. Species trends in response to mean daily maximum temperatures, 1-7 days prior to collection date. E-F. Species trends along the percent impervious surface gradient. G-H. Species trends along the percent tree cover gradient.

Chapter 5: *Allele frequencies of Culex pipiens bioforms vary across an urban to rural gradient in metropolitan areas of the Northeastern United States*

Abstract

Culex pipiens is a primary vector of West Nile virus (WNv) across the northeastern United States, and is comprised of two sympatric bioforms, *pipiens* and *molestus*, that exhibit contrasting ecological, behavioral, and physiological phenotypes. These traits may be differentially advantageous across a heterogeneous landscape, and the relative frequency of these bioforms in a population, as well as their hybridization rates, can impact the likelihood of human cases on WNv. We collected *Cx. pipiens* along urban to rural gradients at 73 rural, suburban, and urban sites in three localities: Washington D.C. and Maryland (N=271), greater Philadelphia, PA (N=113), and greater Chicago, IL (N=183). We genotyped individuals at the CQ11 locus and summarized allele frequencies within populations and assessed genotype frequencies and hybridization rates to determine whether bioforms are in Hardy Weinberg Equilibrium (HWE) at rural, suburban and urban sites. We used generalized linear models to assess the probability of an individual possessing a *molestus* allele at the CQ11 locus along an urban to rural gradient using percent impervious surface and normalized differential vegetation index (NDVI) as contrasting proxies for the gradient. The likelihood of having *molestus* alleles increased with increasing impervious surface and decreased with increasing vegetation cover, but these patterns were not consistent across localities. We found all populations violated the assumptions of HWE. Random mating between

bioforms is not supported by our results. *Culex pipiens molestus* may prefer, or have higher fitness in, urban environments because of their use of belowground habitat, preference for mammalian bloodmeals, and ability to breed and persist in isolated spaces. Our results support non-random distribution and assortative mating of *Cx. pipiens* bioforms across urban to rural gradients in the eastern United States, likely due to niche partitioning and spatial segregation over a heterogeneous landscape.

Introduction

Culex pipiens Linnaeus (Diptera, Culicidae) is a mosquito vector of human and animal pathogens whose current geographic distribution spans the globe (Fonseca et al. 2004, Ciota and Kramer 2013, Gomes et al. 2013). The species was introduced to the North American continent as a result of European colonization (Reiskind and Wilson 2008) where it now plays a critical role in the transmission of multiple parasites and arboviral pathogens, including West Nile encephalitis virus (WNv) (reviewed in Farajollahi et al. 2011). In eastern North America, *Cx. pipiens* serves as primary bridge vector of WNV from birds to humans, and WNv-positive *Cx. pipiens* with mixed human and avian bloodmeals have been documented (Hamer et al. 2009).

Comparative genetic analyses of European, African, and North American *Cx. pipiens* has revealed that some North American populations appear genetically distinct from their European conspecifics (Fonseca et al. 2004, Yurchenko et al. 2020). Globally, the species comprises two bioforms, *Cx. pipiens* form *pipiens* and *Cx. pipiens* form *molestus,* which overlap in geographic distribution and are morphologically indistinguishable (Bahnck and Fonseca 2006). Yet these two bioforms (hereafter *pipiens* and *molestus*) demonstrate ecological, behavioral, and physiological differences. Form *pipiens* primarily feeds upon avian hosts, requires a bloodmeal in their first

gonotrophic cycle (anautogeny), and breeds in swarms in aboveground habitats (eurygamy). Conversely, *molestus* prefers to feed upon mammalian hosts, does not require a bloodmeal prior to laying their first egg raft (autogeny), and breeds belowground in enclosed spaces near human-dominated areas (stenogamy) (Harbach et al. 1984, Chevillon et al. 1995, Byrne and Nichols 1999, Fritz et al. 2015).

Niche partitioning likely plays a role in dividing the two forms spatially (Chevillon et al. 1995). For example, the ability of *molestus* to breed in enclosed spaces belowground enables them to exploit urban infrastructure and associate closely with humans, while *pipiens* is distributed throughout the landscape, ranging from urban to natural and agricultural areas (Mutebi and Savage 2009, Gomes et al. 2009, Arsenault-Benoit et al. 2021). Yet, individuals of the *molestus* form can cross the above-belowground interface and interbreed with aboveground *pipiens* (Kasai et al. 2008), allowing for hybridization between the bioforms.

Genetic distinction between North American *Cx. pipiens* and European conspecifics is thought to result, in part, from interbreeding that occurs between the bioforms in North America. In North America, up to 40% of *Cx. pipiens* have hybrid ancestry in urban regions (Fonseca et al. 2004, Ciota and Kramer 2013, Yurchenko et al. 2020). Recent whole genome sequencing revealed that field-collected aboveground anautogenous *Cx. pipiens* and belowground autogenous *Cx. pipiens* from Washington D.C. were found to be more genetically similar to each other than their respective bioforms from Eurasia (Yurchenko et al. 2020). Hybridization or mixed ancestry between *molestus* and *pipiens* may typical form-specific phenotypic traits, such as autogeny, to be inconsistently expressed by progeny from a single family (Spielman 1957), or inconsistently expressed over the course of a female's lifetime, such as indiscriminate feeding behaviors and loss of blood host specificity (Gomes et al. 2013, Fritz et al. 2015). Reduced host specificity may

contribute to the role of *Cx. pipiens* as a primary bridge vector for WNv in North America, where willingness to feed on both avian and mammalian hosts results in epizootic transmission of WNv to humans (Gomes et al. 2013, Farajollahi et al. 2011).

Molestus can and do persist in urban environments, due in part to available belowground habitat, but it is unclear whether all urban environments offer sufficient, high-quality belowground habitat and resources to support *molestus* populations. *Molestus'* ability to survive and reproduce over several generations in confined, isolated spaces poses challenges for practitioners and researchers to locate and manage populations of *molestus*. As such, few "true" *molestus* populations have been located and documented worldwide, including populations in London, UK, Calumet, IL, Egypt, and Russia (Haba and McBride 2022, Huang et al. 2009, Harbach et al. 1984, Vinogradova and Shaikevich 2007).

Co-occurrence of *pipiens* and *molestus* bioforms on the landscape leads to opportunities for hybridization and interbreeding, theoretically resulting in increased direct risk to human health by catholic feeding hybrids. Suppression of the total *molestus* genetic signature on the landscape by effective and targeted management strategies could lead to reduced opportunities for interbreeding between bioforms overall, if those *molestus* source populations can be located. A thorough assessment of *molestus* gene flow and interbreeding across urbanization gradients in distinct geographic regions may elucidate whether *molestus* ancestry exists universally in urban ecosystems, as well as which characteristics are associated with increased *molestus.*

In this study, we quantified the distribution of allelic variation at the CQ11 locus in wild *Cx. pipiens* populations collected along urban to rural gradients in three replicated metropolitan regions of the eastern United States. Insertion-deletion polymorphisms flanking the CQ11 microsatellite marker were previously leveraged for development of a rapid polymerase chain

reaction (PCR) diagnostic assay (Bahnck and Fonseca 2006). The insertion-deletion polymorphisms result in differing form-specific PCR amplicon sizes that can be visualized using gel electrophoresis. Bioform-specific CQ11 genotypes have been shown to correlate well with the diagnostic behavioral and physiological traits of *molestus* and *pipiens* (Bahnck and Fonseca 2006, Gomes et al. 2009, Di Luca et al. 2016). While the CQ11 assay cannot describe the full genetic background of an individual, it has been found to perform better than other loci to differentiate bioforms (Di Luca et al. 2016) and can be used approximate the relative proportion of a population comprised of *pipiens* and *molestus* alleles for between-population comparisons (Bahnck and Fonseca 2006, Gomes et al. 2009, Di Luca et al. 2016, and additional references summarized in Haba and McBride 2022).

We postulated that the frequency of *molestus* alleles would be highest in urban environments, consistent with the ability of *molestus* to exploit belowground human infrastructure and would decline across an urban to rural gradient due to fitness constraints associated with *molestus*-specific traits where that infrastructure is absent (Hewitt et al. 1990, Shastry et al. 2021). Frequency of *molestus* alleles may differ by locality as well, due to the availability of belowground refugia, or the severity of overwintering conditions. Due to local differences in city planning, design, maintenance, and history, urban ecosystem services vary among cities (McPhearson et al. 2022), as does the amount and quality of *molestus* habitat across urban areas. Identifying biotic and abiotic factors that influence the distribution of *molestus* and their alleles across the landscape may better inform WNv transmission models, improve our predictions about bridge transmission and human risk, and better target surveillance and management in North America (Rinker et al. 2016, Guillot et al. 2005).

Methods

Mosquito collections

Mosquitoes were collected with hay-infused CDC gravid traps (*see supplemental text*, Appendix III) at urban, suburban, and rural sites in 3 replicated metropolitan regions of the eastern U.S.: 1) Washington, District of Columbia (D.C.), and Prince George's, Montgomery, Anne Arundel, and Howard counties in Maryland (hereafter "MD"), 2) in the Northwest Mosquito Abatement District (NWMAD), in Cook County, northwest of Chicago, Illinois ("IL"), and 3) in Bucks, Delaware, Philadelphia, and Montgomery counties in Pennsylvania ("PA") (Figure 1). All specimens were collected from between 38.84°N and 42.14°N latitude (Appendix III, Table 5.1). In MD, fifteen sites, evenly distributed into urban, suburban, and rural classifications, were established in 2019. Mosquitoes were collected on eight occasions per site per year from June to October in 2019 and 2020. In IL, nine sites regularly surveyed by the Northwest Mosquito Abatement District (NWMAD), including three urban, three suburban, and three rural sites, were selected and established for this study. Specimens were collected in gravid traps over eight occasions from June to October in 2020 and 2021. In PA, specimens were collected as part of vector surveillance and management activities by the PA Department of Environmental Protection (PA DEP). Surplus specimens that were not required for pathogen testing were reserved for this study from 49 unique sites from May-October, 2019-2021 (12 rural sites, 25 suburban sites, and 12 urban sites, Figure 1, Appendix III, Table 5.1). Each PA collection contained 10 individuals that were caught in the same trap. Across localities we trapped at 73 unique locations (20 rural, 33 suburban, 20 urban), but there were totals of 120 collection events at rural sites, 145 at suburban sites, and 125 at urban sites due to multiple sampling events at trapping sites in MD and IL. Collection date and GPS coordinates for each collection were recorded at the time of collection.

Molecular identification of Cx. pipiens

Unfed cryptic *Culex spp.* were separated from non-*Culex* using morphological characters and stored at -80°C. Specimens from IL and PA were shipped overnight to University of Maryland, College Park, on dry ice for molecular analysis. Genomic DNA was extracted from the abdomens of MD specimens after spermatheca were removed using one of two approaches: 1) a rapid alkaline-based extraction (Rudbeck and Dissing 1998), 2) commercially available DNA isolation kits (Qiagen DNeasy Blood and Tissue Kit (Cat. No. 69506, Qiagen Inc., Valencia CA, USA or Zymo Quick-DNA Miniprep Plus kit (Cat #11-397B, Zymo Research, Irvine, CA)). Genomic DNA from IL and PA specimens were all isolated from the abdomen using a rapid macerating procedure (Gloor and Engels 1991). All methods had high-yield amplification via polymerase chain reaction (PCR, 89%, 97%, and 84%, respectively). Species and bioform identifications were completed using a series of PCRs following previously described protocols (Crabtree et al. 1995, Smith and Fonseca 2004, Bahnck and Fonseca 2006) with minor modifications as specified below.

Differentiation of Cx. pipiens *assemblage from* Cx. restuans *and* Cx. salinarius

Specimens were identified as *Cx. pipiens* via a multiplexed PCR, targeting the 28S ribosomal subunit to differentiate *Cx. pipiens s.l., Cx. restuans* and *Cx. salinarius* (Crabtree et al. 1995). Each 20µl PCR reaction included 1µl of template DNA, 0.1µl of GoTaq polymerase (Promega GoTaq, cat. #PRD4005), 11.7 µl of molecular grade water, 4.0µl of 5X GoTaq reaction buffer containing 1.5mM $MgCl_2$, 1.6µl of 2.5mM dNTP mix, and 0.4µl of each of the following primers at 10 mM concentration: PQ10 (5'- CCTATGTCCGCGTATACTA -3'), R6 (5'-CCAAACACCGGTACCCAA-3'), S20 (5'-TGAGAATACATACCACTGCT-3'), and the

reverse, CP16 (5'-GCGGGTACCATGCTTAAATTTAGGGGGTA-3'). The thermocycler program was set to one cycle of 94°C for 5 minutes, 33 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, with a final extension step of 72°C for 7 minutes (Crabtree et al. 1995, Rochlin et al. 2007).

Identification of Cx. quinquefasciatus *ancestry*

A second PCR was applied to *Cx. pipiens s.l.* individuals to identify *Cx. quinquefasciatus* homozygotes or *pipiens-quinquefasciatus* heterozygotes from *Cx. pipiens* homozygotes at the acetylcholinesterase (ACE2) locus (Smith and Fonseca 2004). Each 20µl PCR included 1µl of template DNA, 0.1µl of GoTaq polymerase, 10.9 µl of molecular grade water, 4.0µl of 5X GoTaq reaction buffer containing 1.5mM $MgCl_2$, 1.6µl of 2.5mM dNTP mix, and 0.8 µl of each of the following primers at 10 mM concentration: ACEpip (5'-GGAAACAACGACGTATGTACT-3'), ACEquin (5'-CCTTCTTGAATGGCTGTGGCA-3'), and the reverse B1246s (5'-TGGAGCCTCCTCTTCACGG-3'). Thermalcycler conditions included initial denaturation at 94°C for 5 minutes, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, and a final extension step of 5 minutes at 72°C. Individuals with one or two *Cx. quinquefasciatus* alleles at the ACE2 locus were excluded.

Specimen selection and identification of pipiens *and* molestus *bioforms*

Samples identified as *Cx. pipiens* in the two preceding reactions were scored for *molestus* or *pipiens* alleles based on PCR amplification of the CQ11 locus (Bahnck and Fonseca 2006). For MD specimens, a random sample of up to 20 *Cx. pipiens* per site were included (n=271), and for IL up to 25 specimens per site were included (n=183) (Table 1, Appendix III, Table 5.1).

Specimens were distributed evenly across years and seasons (June-October). For PA samples, all specimens identified as *Cx. pipiens* were included (n=113) (Table 1, Appendix III, Table 5.1). Each PCR included 1µl of template DNA, 0.1µl of taq polymerase (Promega GoTaq, cat. # PRD4005), 12.5 µl of molecular grade water, 4.0µl of Promega 5X GoTaq reaction buffer containing 1.5mM MgCl, 1.6µl of 2.5mM dNTP mix, and each of the following primers at 10 mM concentration: 0.2µl of pipCQ11R (5′-CATGTTGAGCTTCGGTGAA-3'), 0.3µl of molCQ11R (5′-CCCTCCAGTAAGGTATCAAC-3'), and 0.3µl CQ11F2 (5'-GATCCTAGCAAGCGAGAAC-3'). For the CQ11 locus reaction, the conditions were as follows: 94°C for 5 minutes, 40 cycles of 94°C for 30 seconds, 54°C for 30 seconds, 72°C for 40 seconds, and a final 5-minute extension step at 72°C. All PCRs were run using a Bio-Rad T100 Thermal Cycler (Bio-Rad Life Sciences, Hercules, CA).

Amplicon visualization by gel electrophoresis

For all PCRs, one negative PCR control, in which purified water replaced genomic DNA, and one negative extraction control, where an elution from a DNA extraction process was conducted without a mosquito carcass, were run alongside mosquito samples. Amplicons were visualized on an agarose gel (2% w/v) alongside a 1Kb ladder (GeneRuler 1Kb Plus, ThermoScientific, Waltham MA) following gel electrophoresis at 120 v for 60 minutes, and identifications were made according to taxon-specific amplicon sizes (Appendix III, Table 5.2).

Analysis of molestus *allele frequency shifts over time*

Allele frequencies at the CQ11 locus were calculated for each collection site and date in each metropolitan region. To examine the impact of season on *molestus* allele frequency, we

initially separated our collection dates into three seasonal periods. One hundred thirteen collections that took place between May 25 to July 11 were designated as "early", 133 collections made between July 11 and August 15 were designated as "mid", and 144 collections made between August 15 and October 1 were designated as "late" collections. Analysis of variance (ANOVA) was used to test for associations between arcsine square root transformed *molestus* allele frequencies and these three seasonal periods using an *a priori* α value of 0.05 as the threshold for statistical significance. Statistical analysis of allele frequency differences over time was performed for each locality, as well as with the combined dataset for a multi-locality analysis.

Landscape-level predictors of molestus *allele frequency*

In Arsenault-Benoit and Fritz (*in review*), site classification (urban, suburban, rural), mean percent impervious surface (MPIS), and canopy density as measured by a normalized differential vegetation index (NDVI) varied along urbanization gradients and correlated with shifts in *Culex* community composition. Sites in each locality were classified as "urban", "suburban", or "rural" using characteristics of primary housing type, zoning, and the design and extent of nearby green space (Arsenault-Benoit and Fritz *in review*). To calculate MPIS and NDVI at each trapping site, we used the National Land Classification Database (NLCD 2016) and NDVI calculator tool based on satellite imagery in ArcMap 10.7 to estimate the mean values for these metrics for pixels encompassed by a 0.5 km radius buffer around each trap location with GIS tools (ArcMap 10.7, Appendix III, Table 5.1). Prior to testing for their associations with *molestus* allele frequency, we examined associations between each continuous environmental variable, MPIS and NDVI, with site classification and locality using a GLM (see Appendix III, Supplemental Methods). Using a likelihood ratio test to compare the complex models with interactions to reduced models, both

MPIS (df = 4, χ^2 = 63.03, p < 0.001) and NDVI (df = 4, χ^2 = 173.54, p < 0.001), were strongly associated with the interaction of site class and locality, suggesting multicollinearity among our landscape-level predictors. MPIS and NDVI means calculated by site class at each locality differed, but trends were similar across localities (Appendix III, Figure 5.1A & B). Variance inflation factors calculated based on a logistic regression model of *molestus* allele frequencies on all three predictors further revealed moderate to high correlations among variables (Site Class VIF = 7.03, MPIS VIF = 7.25, NDVI VIF = 2.12). Due to the: 1) multicollinearity among our landscape-level predictors, 2) differences in how each predictor measures degree of urbanization, and 3) the impact of locality on both MPIS and NDVI, further regression models to examine *molestus* allele frequencies along our urbanization gradient only included one landscape-level predictor and its interaction with locality at a time, but we modelled each landscape-level predictor and its interaction with locality to assess the effect size of the likelihood of having a *molestus* allele with incremental changes in each predictor. Ranges of MPIS and NDVI for each site are in Appendix III, Table 5.2.

Statistical analysis of molestus *allele frequencies along an urbanization gradient*

To examine whether *molestus* allele frequencies varied according to site classification, we used logistic regression model with a binomial error structure. The full model used to examine site-specific *molestus* allele frequencies at all three of our replicate metropolitan localities was:

$$Model\ 1:\ P(mol,\ 1\text{-}mol) \sim (1 + Loc \times Class + Loc + Class + \varepsilon)$$

where Loc = locality, Class = site classification. Test statistics for the effect of a model term on *molestus* allele frequency was calculated by comparing the full model above with a reduced model

(without one model term), as implemented with by the `lrtest` function (lmtest package v.0.9-40; Zeileis and Hathorn 2002).

A second logistic regression model with a binomial error structure was used to examine whether *molestus* allele frequencies varied according to MPIS surrounding a site. The full model used to examine site-specific *molestus* allele frequencies at all 3 of our replicate metropolitan localities was:

$$\textit{Model 2: } P(\text{mol, 1-mol}) \sim (1 + Loc \times MPIS + Loc + MPIS + \varepsilon)$$

where Loc = locality, MPIS = MPIS.

A third logistic regression model with a binomial error structure was used to examine whether *molestus* allele frequencies varied according to NDVI surrounding a site, where NDVI is measured on a scale of 0 (no vegetation) to 200 (full vegetation cover). No IL or MD sites had NDVI values less than 50; NDVI values less than 50 were recovered for only one locality (PA) and were excluded from our final analysis. The full model used to examine site-specific *molestus* allele frequencies was:

$$\textit{Model 3: } P(\text{mol, 1-mol}) \sim (1 + Loc \times NDVI + Loc + NDVI + \varepsilon)$$

where Loc = locality, and NDVI = NDVI. As above, test statistics for the effect of Model 2 and 3 terms on *molestus* allele frequency were calculated by comparing models with `lrtest`.

In our above analysis of NDVI across localities, only one PA urban site remained after removal of sites for which NDVI values were less than 50. To fully examine the impact of NDVI on *molestus* allele frequency, we also performed our NDVI modeling procedure independently for each locality using the following full model, but we found that in IL and PA, the intercept-only model explained more variation than including NDVI as a predictor (see Appendix III, Supplementary Methods and Results).

Model 4 [For each Loc]: P(mol, 1-mol)~(1+NDVI+ε)

Estimated slopes and contrasts

To account for unbalanced sample sizes and outliers across localities, and to represent the average proportion of *molestus* alleles adjusted for other predictors, we also present estimated marginal means (± 95% CI) generated from our two best fitting logistic regression models. The proportion of *molestus* alleles at each locality and site class were compared using a contrast analysis for the site class model using `estimate_contrasts` function in the `emmeans` package (v 1.8.4, Lenth et al. 2023). We also estimated the marginal effect of changes to each continuous variable to *molestus* allele frequency at each locality using the `estimate_slopes` function. All analyses were performed in R v4.2.2 (R Statistical Computing, Vienna Austria, 2022).

Heterozygosity and Hardy Weinberg Equilibrium

Previous studies suggested hybridization occurs between bioforms in North America (Fonseca et al. 2004, Yuchenko et al. 2020). Therefore, we tested whether populations from each locality and site class deviated from expectations of heterozygosity under Hardy Weinberg Equilibrium (HWE) to test whether this is a panmictic population with free interbreeding between bioforms, as a high rate of hybridization would suggest. To test for HWE, we applied the `HWChisqMat` function from the package `HardyWeinberg` using an *a priori* α value of 0.05 to determine statistical significance (Graffleman, 2015). `HWChisqMat` is an extension of `HWChisq` that can be applied to multiple populations (rows) simultaneously. The `HWChisq` procedure generates a contingency table from given genotypes counts, performs a Pearson's X^2 test to

determine the goodness of fit of observed genotypes to expected values under the assumptions of HWE, and generates a simulated p-value from 10,000 simulations.

Results

Collections occurred over the active mosquito surveillance season (May-October) in each locality. For each location and for the aggregated dataset, the relative frequency of *pipiens* and *molestus* alleles did not differ among early, mid and late season collections as determined by analysis of variance (ANOVA) of arcsine square root-transformed *molestus* allele frequencies (IL: $F=0.424$, $df=2$, $p=0.656$, MD: $F=0.563$, $df=2$, $p= 0.572$, PA: $F=0.751$, $df=2$, $p=0.477$, aggregated: $F=0.499$, $df=2$, $p=0.608$, Appendix III, Figure 5.2).

Molestus *allele frequencies along an urbanization gradient*

The interaction between site class and locality only modestly influence variation in *molestus* allele frequency ($\chi2=8.112$, $df=4$, $p=0.088$, Figure 3). Both site class ($\chi2=68.101$, $df=3$, $p<0.001$) and locality ($\chi2=9.488$, $df=3$, $p=0.008$) were independently significant predictors as well. In MD and PA, the mean proportion of *molestus* alleles increased from rural to suburban to urban environments, but in IL, we found relatively fewer *molestus* alleles in suburban environment than in the rural environment (Figure 3). The marginal contrasts analysis demonstrated that there was a significant reduction in *molestus* allele frequency between urban and rural sites (contrast: -0.48, 95% CI:[-0.85, -0.10], $p=0.007$). There were fewer *molestus* alleles in rural environments compared to suburban environments, as well as fewer *molestus* alleles in suburban environments compared to urban environments, but these contrasts were not significant (rural to suburban

contrast: -0.21, 95% CI:[-0.58, 0.15], p=0.167, suburban to urban contrast: -0.26, 95% CI:[-0.63, 0.10], p=0.167).

The interaction between MPIS and locality significantly influenced variation in *molestus* allele frequency ($\chi 2$=6.279, df=2, p=0.043). In MD, there was a small but significant increase in *molestus* allele frequency with increasing impervious surface (β=0.01, 95%CI [0.00, 0.02], p=0.002). A similar, but non-significant trend was observed in PA (β=0.009, 95%CI [0.00, 0.02], p=0.113). In IL, increasing impervious surface had no relationship to *molestus* allele frequency ((β= -0.0013, 95%CI [-0.01, 0.01], p=0.721).

When we examined sites where NDVI was greater than 50, we found that the interaction between NDVI and Locality significantly influenced variation in *molestus* allele frequencies ($\chi 2$=6.997, df=2, p=0.030). In MD, there was a weak but significant decrease in *molestus* allele frequency with increasing NDVI, (β=-0.02, 95%CI [-0.01, -0.03], p<0.001). In PA, we observed a similar negative but only marginally significant relationship between *molestus* allele frequency and NDVI (β=-0.02, 95%CI [-0.05, 0.00], p=0.094), but in IL, there was no relationship (β=0.0017, 95%CI [-0.01, 0.01], p=0.785). These results were consistent with the outcomes of modeling NDVI for each locality individually (see Appendix III, Supplemental Methods and Results). When adjusted for both landscape-level predictors, marginal means for proportion of *molestus* alleles was consistently greatest in MD and lowest in PA (Appendix III, Table 5.3).

Heterozygosity and Hardy Weinberg Equilibrium

Counts and relative frequencies of *pipiens* homozygotes, *molestus* homozygotes, and heterozygotes varied by locality and site class (Table 1, Figure 2). In MD, *pipiens* and *molestus* homozygotes were identified in nearly equal proportions in rural environments, and *molestus* were

dominant in suburban and urban environments. We found the prevalence of heterozygotes was highest in suburban areas (N=11, 12% of population). In IL, counts of *pipiens* and *molestus* homozygotes were nearly equal across site classes. *Pipiens* homozygotes outnumbered *molestus* homozygotes in rural and suburban areas, but not urban areas. The relative proportion of heterozygotes was greatest in the suburban environment in IL (N=13, 23% of the population), which reflected our findings in MD. In PA, *molestus* alleles were much rarer than in the other two localities. *Pipiens* homozygotes were dominant across all site classes, but heterozygotes were more common in all PA site classes relative to MD or IL, and outnumbered *molestus* homozygotes in urban environments (Table 1, Figure 2), which was unique to PA. When all localities were pooled, the rate of heterozygosity was highest in suburban environments (17%, compared to 10% in rural environments and 11% urban environments). *Molestus* homozygotes were most prevalent in urban environments, accounting for 52% of the specimens identified in urban environments in the study (Figure 2, Table 1).

Discussion

Here, we examined the distribution and gene flow of *Cx. pipiens* bioforms and their alleles across urban to rural gradients in eastern North America. In all study locations, we documented the presence of *molestus* alleles at sites spanning our urbanization gradient. Yet, differences in *molestus* allele frequency between site classes were evident: the frequency of *molestus* alleles was significantly higher in urban environments compared to rural environments, and characteristics of urbanization, including increased impervious surface and decreased vegetation, were correlated with increasing *molestus* allele frequencies. In MD and PA, trends in changes to *molestus* allele frequency along urbanization gradients were notable, but weak. In IL we found no relationship

between *molestus* allele frequency and NDVI or MPIS. IL sites exhibited reduced environmental heterogeneity overall, but there were recognizable differences in NDVI and MPIS across site classes. It is evident that there was locality-specific variation in genetic processes and local adaptation that were not fully described by our models.

The use of urban habitats by the *molestus* bioform has been well-documented in the literature (Chevillon et al. 1995, Byrne and Nichols 1999, Haba and McBride 2022). The built environment offers isolated and enclosed locations for breeding and refuge from climatic extremes (Brown et al. 2008, Arsenault-Benoit et al. 2021). Propagation in isolation through autogeny over many generations allows for persistence of *molestus* populations with limited contact with aboveground *pipiens*, resulting in isolation between bioforms in regions where they appear broadly sympatric. In this way, urbanization can lead to genetic isolation between bioforms. Here, we established that bioforms are differentially distributed across landscape gradients in eastern North America, with a trend of increasing *molestus* ancestry at increasingly urbanized sites. Both bioforms were present in all three metropolitan areas studied as measured by a single genetic marker, but there was considerable local variation in the distribution of alleles, suggesting that not all urban areas offer suitable conditions for true *molestus* populations or gene flow along urbanization gradients.

We found that alleles of a single neutral locus used to distinguish between *Cx. pipiens* bioforms were not in Hardy Weinberg Equilibrium (HWE) in the northeastern USA. Our results suggest that one or more of the assumptions of HWE, including random mating within the population, were violated. Heterozygotes were underrepresented across localities and site classes, and all populations deviated from HWE expectations, likely due to heterozygote deficit and a lack of interbreeding between bioforms in the populations analyzed. A deficit of observed

heterozygotes compared to expected values in a population may be an indication of underlying population structure, and an example of phenomenon known as the Wahlund Effect (Kothera et al. 2009, Di Luca et al. 2016). We posit that non-random mating between bioforms is the basis for site classes to be out of HWE, however there are a number of other assumptions for HWE, the violation of which could also lead to the results presented here. For instance, the HWE equation assumes no selection within a population, but phenotypic traits associated with bioforms likely have adaptive value across a landscape gradient. Therefore, disruptive selection leading to speciation among bioforms, in tandem with non-random mating, could also explain the results observed, as selection for *molestus* traits under urban conditions would lead to an unequal distribution of alleles across the urbanization gradient. Although there were significantly more *molestus* alleles in urban environments than rural, both bioforms were observed across all site classes. Our results show some support for selection playing a role in allelic frequency distributions, yet there is still gene flow occurring across the landscape.

Generally, heterozygotes were rare, but in all three localities, they were observed most often in suburban habitats with primarily single-family homes and well-managed, fragmented green space. This environment may serve as an interface between above-ground natural environments that support the *pipiens* bioform, and densely populated and altered environments suitable for the *molestus* bioform. Both *pipiens* and *molestus* homozygotes were commonly observed in urban environments, but rates of hybridization were reduced in urban than suburban environments.

Culex pipiens are naturalized in the United States and became established in North America from Northern Africa and Europe during the time of European exploration (Reiskind and Wilson 2008). In the native range, *Cx. pipiens* bioforms' relative frequency and rates of hybridization vary

across a latitudinal gradient (Arich et al. 2022). At the northernmost extreme of the range, *molestus* are less common, interbreeding is rare, and bioforms maintain distinct populations and traits, like habitat use, mating behaviors, and autogeny (Fonseca et al. 2004, Bahnck and Fonseca 2006, Haba and McBride, 2022). It is plausible that individuals with *molestus* ancestry, including heterozygotes, are less fit at higher latitudes because they do not diapause and cannot overcome thermal limits for overwintering without sufficient insulative belowground habitat in concert with the urban heat island effect (Yurchenko et al. 2020, Arsenault-Benoit et al. 2021). In the eastern United States, we would expect reduced *molestus* alleles in localities traveling north on a latitudinal gradient, with fewer *molestus* overall in PA and IL compared to MD. However, traits associated with *molestus* alleles could confer an advantage in highly urbanized areas, subsequently "rescuing" *molestus* populations in northern localities. We observed evidence of this trend, as MD had the highest frequency of *molestus* alleles overall.

Hybridization is more common in Mediterranean areas of southern Europe, but generally populations still deviate from HWE due to a deficit of heterozygotes (Gomes et al. 2009, Di Luca et al. 2016, Arich et al. 2022, Haba and McBride 2022). Conversely, in Northern Africa, populations are panmictic, and landscape features do not appear to influence the distribution of *Cx. pipiens* bioforms (Arich et al. 2022). Collections for the current study occurred between 38°N and 42°N, which is latitudinally similar to North Africa and the Mediterranean, but climatically more similar to central and northern Europe. Geographic origins of the initial introductions of *Cx. pipiens* from the native range into North America likely influence the underlying contemporary population structure, but latitudinal effects and habitat features may also play a role in bioform distribution and hybridization. Compared to the native range, populations in North America tend to have considerably higher rates of hybridization, including previously studied populations from

Washington D.C., Chicago, IL, New York, NY (Fonseca et al. 2004, Kent et al. 2007, Yurchenko et al. 2020).

Relative frequency of *pipiens*, *molestus*, and heterozygotes were similar in MD and IL, but *molestus* alleles were identified less frequently in PA samples. Notably, PA sites are intermediate in latitude compared to MD and IL. Populations of autogenous, belowground *Cx. pipiens molestus* have been identified in the Chicago IL region (Mutebi and Savage 2009), Philadelphia, PA (Bahnck and Fonseca 2006, Strickman and Fonseca 2012), and Washington D.C. (Arsenault-Benoit et al. 2021). Underrepresentation of *molestus* ancestry in a population is likely due to reduced adaptive value of *molestus*-like traits. In temperate regions, non-diapausing mosquitoes, including *Cx. pipiens molestus* require refuge to overcome winter thermal limitations. Although belowground *molestus* can be located in Philadelphia, it is possible that belowground infrastructure is insufficient in the broader geographic region to support large populations with detectable *molestus* ancestry beyond the inner city, in comparison to MD and IL.

Across the landscape, more interbreeding between bioforms and the abundance of individuals with hybrid ancestry could have significant implications for disease transmission (Huang et al. 2008, Arich et al, 2022). Hybrids can demonstrate intermediate characteristics, including catholic feeding behaviors (Gomes et al. 2013, Di Luca et al. 2016). Admixed mosquitoes are more likely to be indiscriminate feeders and cause spillover of WNv from birds to mammals, including humans (Hamer et al. 2009, Gomes et al. 2013). Evidence suggests that the risk of WNv transmission to humans is greatest in urban and suburban environments (Ruiz et al. 2007, Kilpatrick et al. 2011, Ciota and Kramer 2013, Karki et al. 2020), and the majority of human-fed *Cx. pipiens* have been recovered from residential properties (Hamer et al. 2009). Both human WNv cases and human-fed mosquitoes may be simply correlated with the high density of humans

in those regions, but due to the epizootic nature of WNv, the biology of the mosquito vectors likely also plays a role. West Nile virus cycles enzootically between bird-biting mosquitoes, including *Cx. pipiens pipiens* and avian amplifying hosts. Humans are exposed when a mosquito first feeds on an infected bird, then later feeds on a human, resulting in epizootic transmission, or "spillover." Bioforms are characterized in part by their host preferences, to the extent that they are considered to have different vectorial capacity for WNv and different epidemiological importance (Bahnck and Fonseca 2006, Di Luca et al. 2016).

We found that in eastern North America, *Cx. pipiens* do not experience unrestricted gene flow and interbreeding in metropolitan areas, as all populations deviated from HWE expectations. However, *Cx. pipiens* heterozygotes, which are more likely to have catholic host use, were most common in suburban environments, followed by urban environments. *Culex pipiens molestus,* which are phenotypically more likely to use human hosts, were also associated with characteristics of urbanization such as increased impervious surface and decreased vegetation. In these habitats, where human population densities are also greatest, risk of WNv transmission via epizootic spillover is also highest. Vector management practitioners often focus their efforts on areas of high population density due to nuisance complaints. Our results support the continuation of this practice, as the genetic background of *Cx. pipiens* found in suburban and urban habitats, including individuals with increased *molestus* ancestry and increased heterozygosity at CQ11 may result in populations with higher vectorial capacity for human WNv in these habitats. However, our data show that there are locality-specific processes ongoing across the broader landscape that are not fully described by our models. Therefore, surveillance and management efforts must consider local environmental heterogeneity, habitat availability, WNv cases, and *Culex* population genetics to best manage local populations. Non-diapausing and human biting *molestus* individuals are active

year-round and difficult to target, and in tandem with hybrid individuals with intermediate traits, pose the greatest risk to human health. Our results suggest that hybrid individuals that can cause spillover are likely found at the interface of urban and suburban environments as the availability of belowground infrastructure wanes. Identifying below ground refugia and treating those water sources, especially those that are available through the winter, should be incorporated into urban integrated vector management programs to target populations demonstrating *molestus* traits.

Chapter 5: Tables

Table 1. Counts and percentages of *Cx. pipiens* genotype frequencies by locality and site class at the CQ11 locus. Hardy Weinberg Equilibrium (HWE) statistics were calculated for each population, and all were out of HWE according to our statistical significance threshold of $\alpha = 0.05$.

Locality	Site Class	Total Sample size	Number of *pipiens* homozygotes (%)	Number of heterozygotes (%)	Number of *molestus* homozygotes (%)	HWE Statistics	
						χ^2	p-value
MD	Rural	83	37, (45%)	7, (9%)	38, (46%)	53.929	<0.001*
	Suburban	96	28, (29%)	11, (11%)	57, (59%)	51.176	<0.001*
	Urban	93	25, (27%)	6, (6%)	62, (66%)	63.537	<0.001*
IL	Rural	77	36, (47%)	9, (12%)	32, (42%)	42.861	<0.001*
	Suburban	54	24, (44%)	13, (24%)	17, (31%)	12.545	<0.001*
	Urban	51	22, (43%)	4, (8%)	25, (49%)	33.718	<0.001*
PA	Rural	24	16, (67%)	3, (13%)	5, (21%)	8.623	0.003*
	Suburban	53	32, (60%)	10, (19%)	11, (21%)	14.174	<0.001*
	Urban	36	20, (56%)	9, (25%)	7, (19%)	0.094	0.002*
Total	Rural	183	89, (49%)	19, (10%)	75, (41%)	112.158	<0.001*
	Suburban	203	84, (41%)	34, (17%)	85, (42%)	87.793	<0.001*
	Urban	180	67, (37%)	19, (11%)	94, (52%)	108.237	<0.001*

* denotes a statistically significant p-value.

Chapter 5: Figures

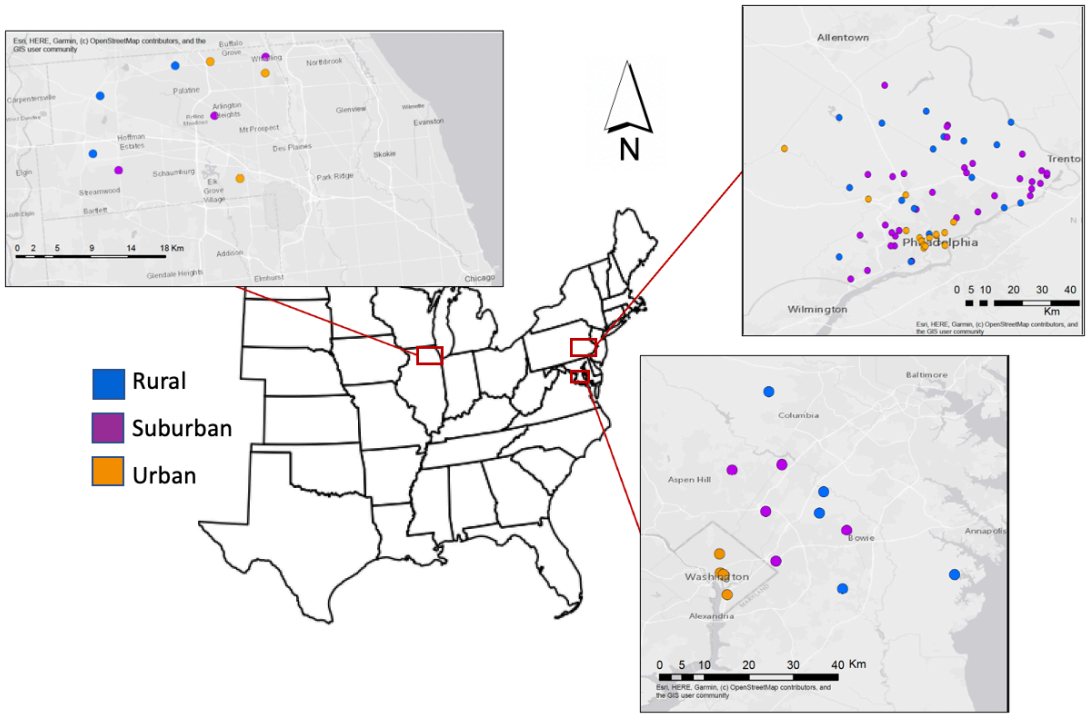

Figure 1. Collection sites in three metropolitan localities in the eastern United States: Chicago, Illinois, Philadelphia, Pennsylvania, and Washington, District of Columbia, including Maryland suburbs. Rural sites are in blue, suburban sites are in purple, and urban sites are in orange.

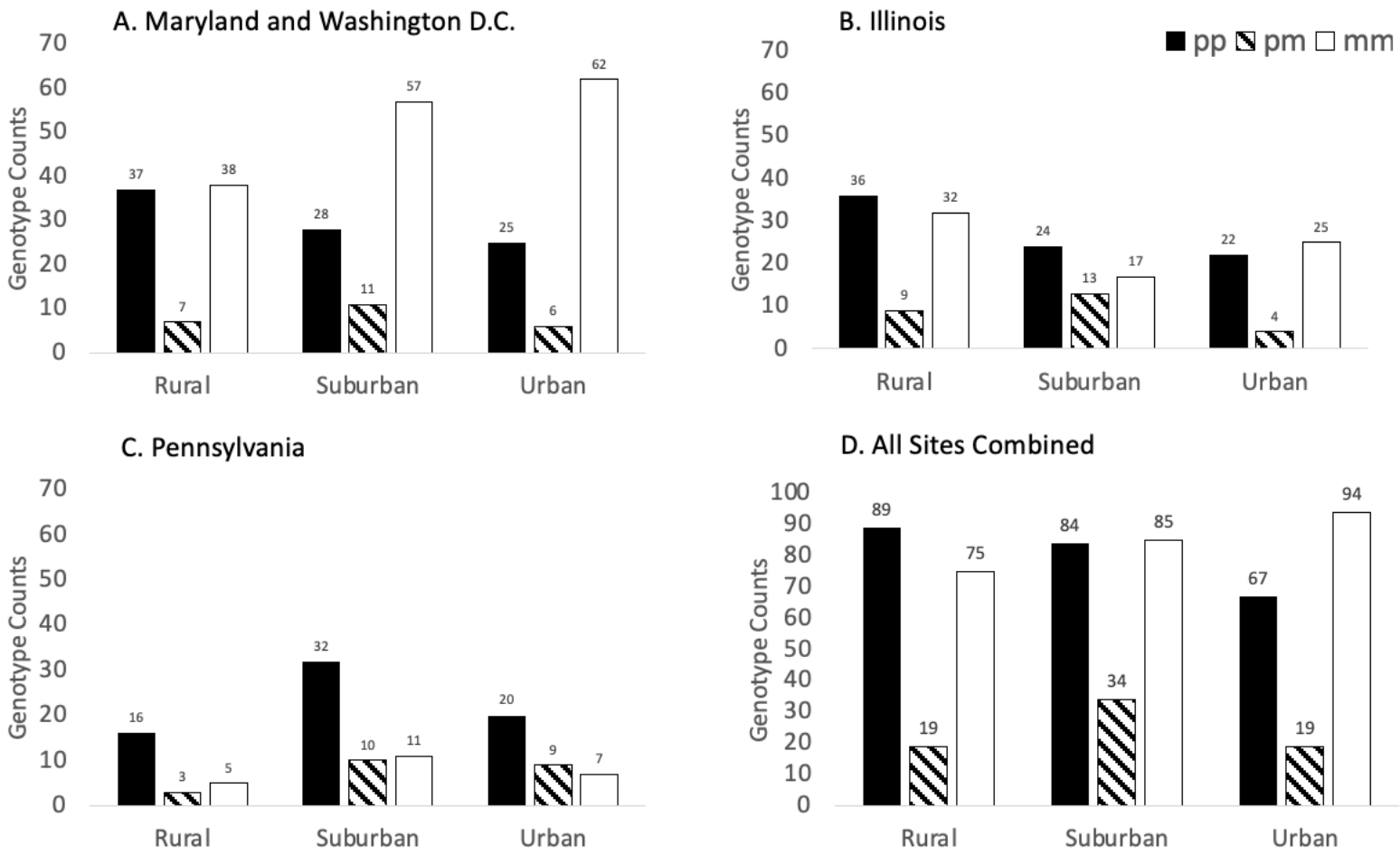

Figure 2. Counts of *pipiens* homozygotes (black), *molestus* homozygotes (white), and heterozygotes (patterned) based on genotyping of the CQ11 locus. Panels represent each collection locality: A. Maryland and Washington D.C. (N=271), B. NWMAD in Illinois (N=183), C. Greater Philadelphia, PA (N=113). Panel D shows aggregate values of all sites combined (N=567).

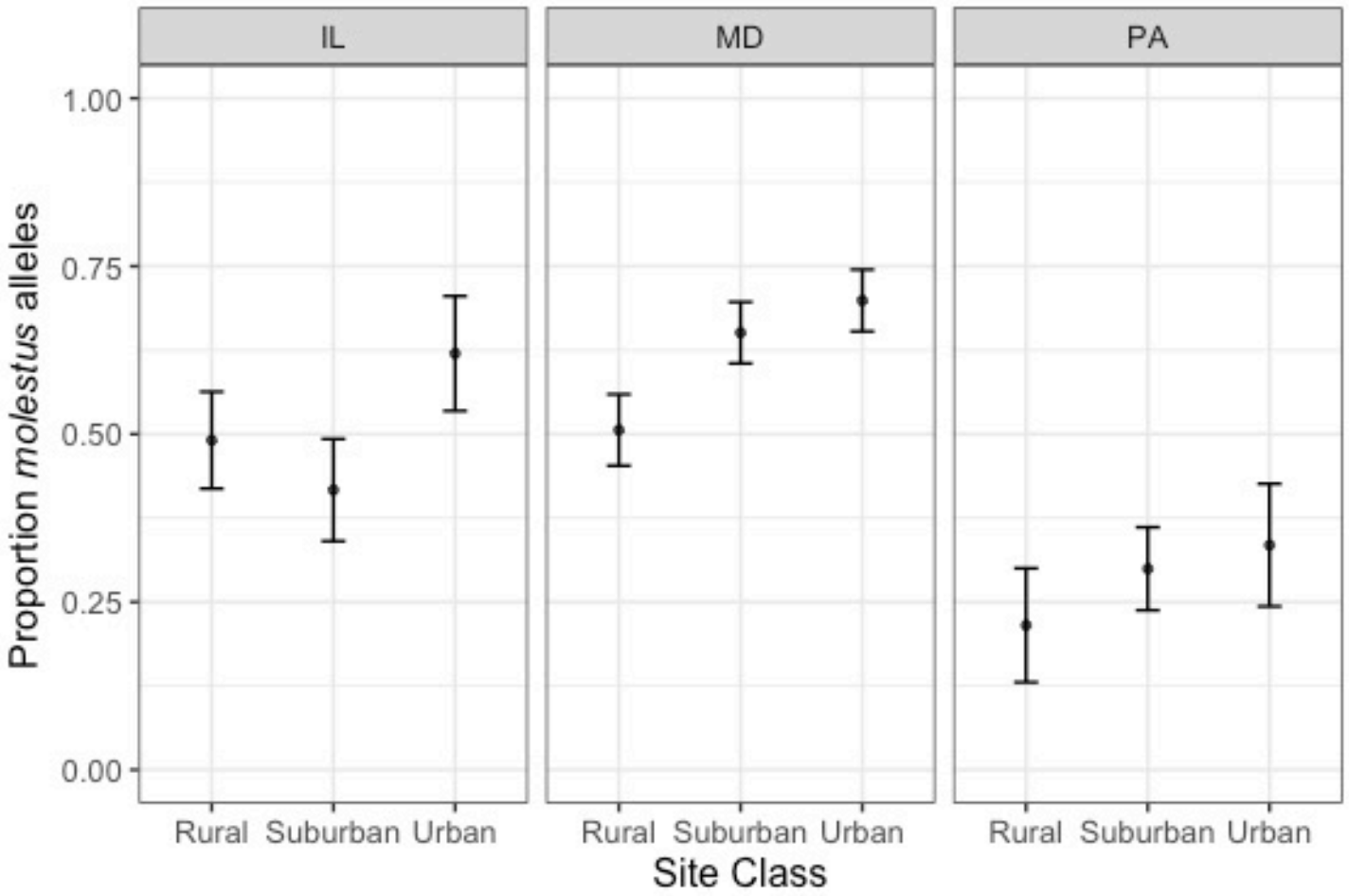

Figure 3. Mean ± standard error of proportion *molestus* alleles in rural, suburban, and urban populations in metropolitan Chicago, Illinois (IL), Washington D.C. and Maryland (MD) and Philadelphia, Pennsylvania (PA).

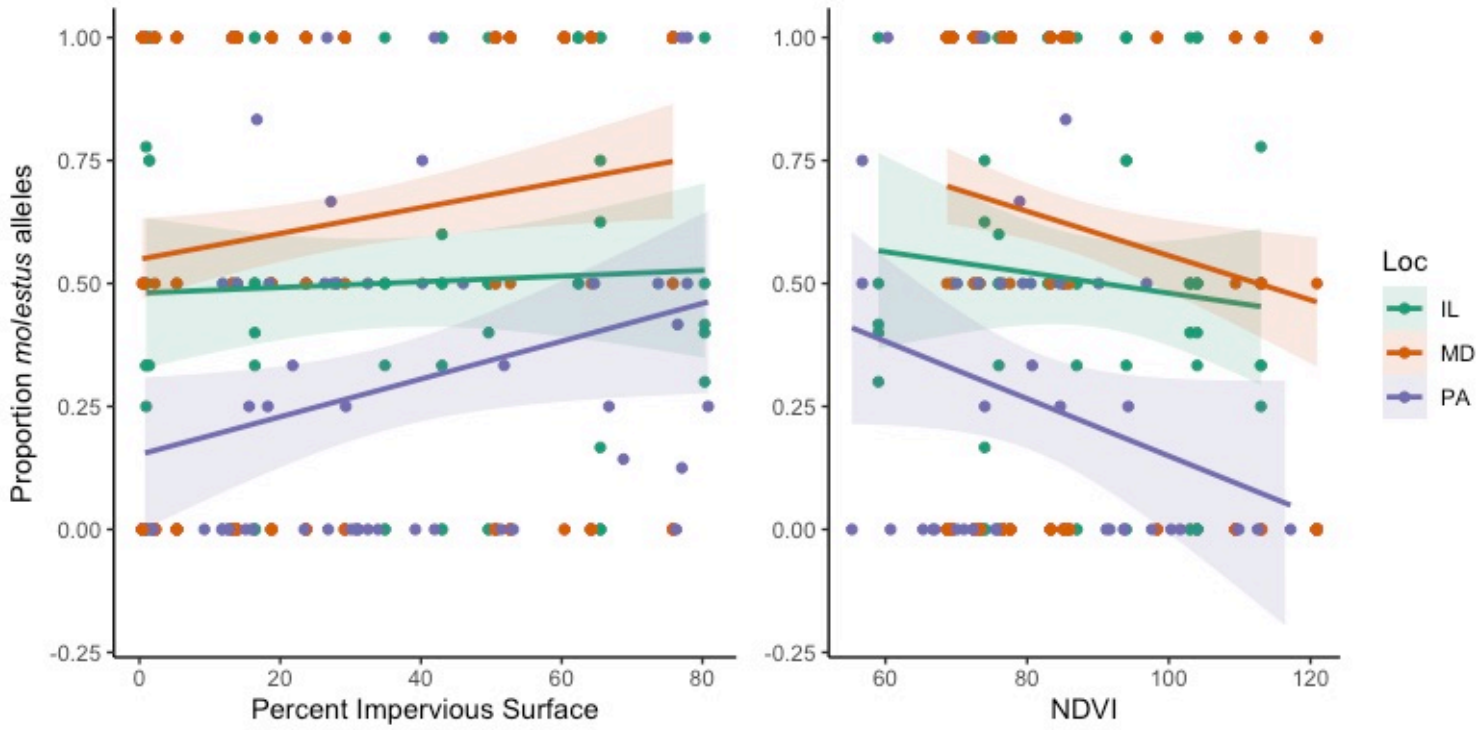

Figure 4. Linear trends for *molestus* allele frequencies by trap with A. increasing impervious surface and B. vegetation index (NDVI), for each locality. Only NDVI values greater than 50 were included. Shaded regions represent 95% confidence intervals around the lines of best fit.

Chapter 6: *Synthesis and Future Directions*

Landscape heterogeneity generated by the process of urbanization impacts mosquito habitat availability in North America, promoting some vectors, while suppressing others. Trait variation within and between cryptic *Culex* vectors of West Nile virus (WNv)(Andreadis 2012, Johnson et al. 2015, Bahnck and Fonseca 2006, Fritz et al. 2015), may influence habitat use across heterogenous North American landscapes (Ferraguti et al. 2016, Gorris et al. 2021). My studies explored the impact of land use on the spatiotemporal dynamics and genetics of four *Culex* taxa that contribute to human risk of WNv transmission in eastern North America.

I demonstrated that members of the *Cx. pipiens* assemblage use belowground urban structures in Washington, D.C. for breeding, development, and these structures have potential to serve as potential refugia against extreme climatic conditions. I then sampled *Culex* mosquitoes along an urbanization gradient in greater Washington, D.C. and Maryland during the months of April through October. *Culex restuans* was the first to become abundant on the landscape, and by mid-May *Cx. pipiens* assemblage mosquitoes co-occurred with them across the gradient. The seasonal shift from *Cx. restuans*-dominant communities in the spring months to *Cx. pipiens*-dominant communities in the summer months was previously reported in the literature (Lee and Rowley 2000, Andreadis et al. 2001, Ebel et al. 2005, Johnson et al. 2015). I observed this shift in three different metropolitan regions in different US states, indicating this phenomenon is widespread, and its timing appeared to be driven by latitudinal differences in early spring temperatures.

By June, *Cx. pipiens*, *Cx. restuans*, and suspected *Cx. quinquefasciatus* were differentially distributed at fine spatial scales, likely due to the impacts of urbanization on vector

habitat. *Culex pipiens* were cosmopolitan and the dominant species across the gradient, while individuals with *Cx. quinquefasciatus* ancestry were limited to urban and peri-urban sites, and *Cx. restuans* were most abundant in rural and suburban sites furthest from the city center. Urban landscape characteristics like percent impervious surface, percent tree cover, and distance of collection to city center were most strongly associated with differences in species distributions between June-October.

Within the species *Cx. pipiens,* two bioforms known as *pipiens* and *molestus* can interbreed, but have divergent ecological, physiological, and behavioral traits (Fonseca et al. 2004). I quantified the frequencies of *molestus* alleles at a single genetic locus for *Cx. pipiens* populations sampled from across MD, PA, and IL urbanization gradients. Increasing percent impervious surface and reduced vegetation was correlated with an increase in the frequency of *molestus* alleles. *Pipiens-molestus* heterozygotes primarily occurred in suburban environments, but were rare, and no populations were in Hardy Weinberg Equilibrium. I propose that rather than forming a panmictic population, *pipiens* and *molestus* are differentially distributed on the landscape due to non-random mating, selection, or these factors working in concert.

We can best mitigate WNv risk by following epidemiological principles of evidence-based and early intervention. For cryptic *Culex*, at both the genus and species level, we found that seasonality and landscape-level predictors affect the abundance and relative frequency of the taxa studied here. Changes in abundance in response to spatiotemporal correlates can lead to corresponding changes in a vector population's ability to transmit a pathogen, which is mathematically described using an equation called vectorial capacity (C). Overall disease risk is a function of the composite C(s) for all populations. When all other parameters are equal, C is

positively correlated with changes in abundance, so the patterns of seasonality, habitat use, and abundance I documented throughout this dissertation can inform where and when populations of cryptic *Culex* are likely to contribute most to disease. For example, *Cx. restuans* are at their highest densities soon after spring emergence, then in forested rural and suburban sites, so the value of C for *Cx. restuans* should be greatest among populations under those conditions. *Culex quinquefasciatus* were limited to warmer urban areas due to their need for refugia from overwintering thermal limitations (Samy et al. 2016). Consequently, the values of C for populations residing outside of those optimal conditions are likely to be comparatively low. We found that after the summer seasonal shift, *Cx. pipiens* were cosmopolitan and dominant across the landscape, so they have a large contribution to overall C for *Culex.* However, since *molestus* or heterozygote females are more likely to take bloodmeals from human hosts, populations with high *molestus* allele frequencies, like those we documented in urban environments, have higher vectorial capacity for human hosts, leading to increased epizootic spillover.

While considering the contributions individual species among shifting vector communities is complex, my work on changes in *Culex* abundance can inform WNv transmission management recommendations. First, since WNv in the environment undergoes exponential growth throughout the season, and the initiation of transmission relies on virus that overwinters in adult female mosquitoes (Nasci et al. 2001, Reisen and Wheeler 2019), interventions aimed at suppressing early spring populations, should reduce WNv amplification throughout the season, and ultimately human cases. *Culex restuans* emerge first, and are most common in natural, forested sites. These types of sites are not currently prioritized for management due to low human population density, but early intervention at these sites could have significant downstream effects. We also found that in cities, members of the *Cx. pipiens*

assemblage, including those with *quinquefasciatus* and *molestus* genetic markers, use belowground infrastructure as refugia and for oviposition. This could enable early season emergence and viral amplification among these non-diapausing taxa that are also likely to feed on humans. Source reduction, larvicide application, and surveillance at these types of sites throughout the year, including dates beyond what is conventionally considered the active mosquito season could have positive outcomes on disease transmission at the landscape level.

Future Directions

I emphasized the potential importance of *Cx. restuans* spring emergence on the overall WNv burden in the environment due to exponential growth of the pathogen over the season. It is documented that WNv overwinters in diapausing *Cx. pipiens* females, but these studies are limited by potential misidentification or pooling of specimens into a "*Culex* group" (Nasci et al. 2001, Andreadis et al. 2010). Although we know that *Cx. restuans* are competent vectors of WNv (Ebel et al. 2005, Harrington and Poulson 2008), there is limited empirical evidence of *Cx. restuans* harboring WNv infection through the winter. To accurately determine the early season impact of *Cx. restuans* abundance on disease burden overall, I suggest a study assessing the presence of WNv in *Cx. restuans* at the time of emergence or collected from overwintering hibernacula and identified using molecular methods.

In this dissertation, I aimed to describe the spatiotemporal dynamics of cryptic *Culex* species across the landscape, yet very few of our collection sites yielded *Cx. quinquefasciatus* specimens, and single locus assays limited our ability to describe the rate of heterozygosity or degree of introgression and backcrossing among members of the *Cx. pipiens* assemblage, including individuals with potential *molestus* or *quinquefasciatus* ancestry. Individuals with *Cx.*

quinquefasciatus and/or *molestus* background are likely to be important in WNv epizootic spillover due to behavior and host preferences (Zinser et al. 2004, Ciota et al. 2013, Haba and McBride 2022). I found evidence of both taxa in urban environments, but a genomic analysis of the *Cx. pipiens* assemblage across environmental gradients, with a focus on admixture, will enable deeper understanding of gene flow and interbreeding across the landscape and may further improve predictive WNv modeling. This work could also be of benefit to public health practitioners by providing insight into the future composition of the *Cx. pipiens* assemblage under regimes of climate change and increasing urbanization.

Appendices

Appendix I: Chapter 3 Supplement

Supplemental Text

Methods:

Specification of CCA models

We determined which explanatory variables to keep in the CCA model based on variables having $p<0.01$ based on 999 permutations in the forward-model selection function `ordistep` from the `vegan` package on a constrained ordination combined with having $p\leq0.05$ from the `envfit` function on an NDMS (Appendix I, Table 3.4). We then used the VIF function in the `car` package (Fox and Weisberg 2019) to test for multicollinearity between variables. When variable inflation factors for any pair of variables were greater than 10, we eliminated the member of the pair that had the lower significance in the ordistep function, until we had a satisfactory model including the maximum predictor variables with limited linear dependence (Alves et al. 2017).

Results:

Summary of annual trap catch

Average *Culex* catch was lowest in 2020 (23.58±2.153, mean±SEM), relative to 2019 (47.16±4.305) or 2021 (43.967±4.014). Rural sites produced fewer *Culex* (19.06±3.341) than suburban (24.04±3.548) and urban (29.30±4.085) sites. *Culex* collections were largest in the early season (29.48±2.358), declined mid-season (24.35±2.088), and were lowest in the late season (20.94±2.220).

Details of 2021 CCA outcomes

We report the outcomes of the same two top performing models as June-October 2019-2021 (Figure 7). In model 1 the total inertia was 35.3% (adjusted R^2=0.303, p=0.001, Figure 7A), and primary site class (χ^2=0.123, F=11.396, df=2, p=0.001), and distance to city center (χ^2=0.0244, F=4.587, df=1, p=0.023) were independently significant. Species associations with site class and DCC were similar to June-October Model 1.

The total inertial for model 2 was 35.7% (adjusted R^2=0.285, p=0.002). Percent impervious surface (χ^2=0.101, F=18.26, df=1, p=0.001), distance to city center (χ^2=0.025, F=4.552, df=1, p=0.028), and latitude (χ^2=0.020, F=3.560 df=1, p=0.047) all independently accounted for a significant amount of inertia in the model.

Supplemental Tables

Table 3.1. Quantitative description of Greater DC sites, generated in ArcMap 10.7 based on a 0.5km radius around the location where traps are set.

Site	Classifi-cation	X	Y	Land Cover (Maj)	Land Cover (Name)	% Imperv Mean	% Tree Canopy Cover	Popula-tion	Housing units	Distance to City Center (km)	Water Table Depth (cm)	Elevation	NDVI (0-200)
Beltsville	Rural	39.01231	-76.8258	82	Cultivated Crops	0.6	26.8	0	0	21.239	57	45.9	113.07
Clarksville	Rural	39.25637	-76.9276	81	Pasture	0.4	5.1	0	0	39.903	25.4	109.7	120.9
Patuxant	Rural	39.05431	-76.817	41	Deciduous Forest	5.3	66.19	0	0	24.457	57.8	45.4	83.33
SERC	Rural	38.88785	-76.5534	41	Deciduous Forest	0.7	72	65	20	41.351	50.7	11.3	98.35
Upper Marlboro	Rural	38.85905	-76.7786	82	Cultivated Crops	2.2	29.6	301	90	22.526	71.5	26.7	109.41
Bowie	Suburban	38.97728	-76.7705	22	Developed Low Intensity	18.8	37.5	592	208	24.135	90.9	55.1	85.15
Cheverly	Suburban	38.91504	-76.9142	22	Developed Low Intensity	29.2	35.4	256	95	10.185	42.9	13.6	73.17
College Park	Suburban	39.01578	-76.9341	90	Woody Wetlands	23.7	38.4	213	632	14.738	44	34.4	72.51
Jerald	Suburban	39.1101	-76.9016	21	Developed Open	13.9	33	453	159	24.779	55	114.1	77.56
Old Barn	Suburban	39.0992	-77.0018	21	Developed Open	13.2	34.2	146	55	21.561	29.9	143.4	73.3
Air and Space	Urban	38.88816	-77.0198	23	Developed Medium Intensity	60.4	0.9	117	134	2.542	65	5	77.68
Columbia Heights	Urban	38.92955	-77.0274	23	Developed Medium Intensity	75.8	0	33889	14656	2.590	69	56.1	69.45
Cst	Urban	38.88587	-77.0177	23	Developed Medium Intensity	64.2	0.4	1846	1095	2.880	100	4.9	68.69
Anacostia	Urban	38.84722	-77.0118	23	Developed Medium Intensity	50.6	2.7	1899	526	7.047	84	8.1	85.92
Natural History	Urban	38.89126	-77.0260	23	Developed Medium Intensity	52.7	0.9	573	549	1.979	65	5.2	76.61

Table 3.2. Counts of *Cx. pipiens* homozygotes (pp), *Cx. quinquefasciatus* homozygotes (qq) and heterozygotes (pq) at the ACE2 locus by site, season, and year, used for Hardy-Weinberg Equilibrium calculations. Chi-square statistics and p-values were generated with the HWChisqMat function in the `HardyWeinberg` R package. Instances where no *Cx. quinquefasciatus* alleles were present were excluded from analysis (n=74).

Year	*Site*	*Season*	*Primary Class*	*pp*	*pq*	*qq*	X^2	*p value*
2019	AirSpace	Early	Urban	24	0	1	6.349392	0.011742*
2019	AirSpace	Late	Urban	10	0	4	10.3975	0.001262*
2019	AirSpace	Mid	Urban	15	0	1	4.084444	0.04328*
2019	Bowie	Early	Suburban	14	0	1	3.831633	0.050294
2019	Bowie	Late	Suburban	20	0	1	5.344375	0.020789*
2019	ColHeights	Early	Urban	14	0	4	13.55867	0.000231*
2019	ColHeights	Late	Urban	11	0	4	11.19434	0.00082*
2019	ColHeights	Mid	Urban	12	0	3	10.24479	0.001371*
2019	CollPark	Mid	Suburban	9	0	2	6.075617	0.013706*
2019	Cst	Early	Urban	15	0	6	17.29833	3.19E-05*
2019	Cst	Late	Urban	6	0	1	1.777778	0.182422
2019	Cst	Mid	Urban	11	0	4	11.19434	0.00082*
2019	Anacostia	Early	Urban	17	0	1	4.5891	0.032176*
2019	Anacostia	Late	Urban	33	0	1	8.60629	0.00335*
2019	Anacostia	Mid	Urban	11	1	1	1.1297	0.287839
2019	Jerald	Early	Suburban	17	0	2	10.65614	0.001097*
2019	NatHist	Early	Urban	12	1	2	4.636	0.031308*
2019	NatHist	Late	Urban	15	1	1	1.588681	0.207515
2019	UM	Early	Rural	6	0	1	1.777778	0.182422
2020	AirSpace	Early	Urban	25	0	4	22.09098	2.60E-06*
2020	AirSpace	Mid	Urban	12	0	5	13.41826	0.000249*
2020	Beltsville	Early	Rural	11	0	2	7.231405	0.007164*
2020	Beltsville	Mid	Rural	23	0	2	14.05529	0.000178*
2020	Bowie	Early	Suburban	14	0	3	11.66355	0.000637*
2020	Clarksville	Early	Rural	23	1	0	23.25532	1.42E-06*
2020	ColHeights	Early	Urban	5	0	2	3.695	0.054576
2020	ColHeights	Mid	Urban	6	0	1	1.777778	0.182422
2020	CollPark	Mid	Suburban	24	0	2	14.62066	0.000131*
2020	Cst	Early	Urban	26	0	2	15.75074	7.23E-05*
2020	Cst	Late	Urban	15	0	5	15.92222	6.60E-05*
2020	Cst	Mid	Urban	23	0	3	17.98451	2.23E-05*
2020	Anacostia	Early	Urban	24	0	3	18.68359	1.54E-05*
2020	Anacostia	Mid	Urban	25	1	3	13.93155	0.00019*
2020	Anacostia	Early	Urban	29	0	1	7.603746	0.005825*

2020	Anacostia	Mid	Urban	25	1	2	9.4211	0.002145*
2020	NatHist	Early	Urban	22	0	1	5.847107	0.015603*
2020	NatHist	Mid	Urban	5	0	2	3.695	0.054576
2020	OldBarn	Early	Suburban	30	1	1	3.27848	0.070194
2020	UM	Early	Rural	19	0	1	5.092798	0.024025*
2020	UM	Late	Rural	5	0	1	1.51	0.219139
2021	AirSpace	Early	Urban	17	0	2	10.65614	0.001097*
2021	AirSpace	Mid	Urban	15	0	1	4.084444	0.04328*
2021	Beltsville	Mid	Rural	11	0	1	3.070248	0.079737
2021	ColHeights	Early	Urban	20	0	1	5.344375	0.020789*
2021	CollPark	Mid	Suburban	14	0	1	3.831633	0.050294
2021	Cst	Early	Urban	14	0	2	8.94898	0.002776*
2021	Cst	Late	Urban	10	0	1	2.815	0.093387
2021	Anacostia	Late	Urban	9	0	1	2.558642	0.109693
2021	NatHist	Early	Urban	21	0	1	5.595805	0.018004*
2021	NatHist	Mid	Urban	10	0	2	6.655	0.009888*
2021	OldBarn	Early	Suburban	22	0	1	5.847107	0.015603*
2021	Pax	Mid	Rural	13	0	1	3.578402	0.058535
2021	UM	Mid	Rural	8	0	1	2.300781	0.129309

Table 3.3. Correlations of environmental variables with NMDS ordination of sites and species at increasing number of dimensions (k=2 through 5). The stress of the ordination with each number of dimensions is included. Variables that were consistently correlated with the NMDS ordination axes (in bold) were preliminarily included in the constrained ordination analyses.

Predictor	***EnvFit (k=2)***	***EnvFit (k=3)***	***EnvFit (k=4)***	***EnvFit (k=5)***
STRESS	0.0298	0.0215	0.0158	0.0128
Longitude	0.959	0.980	0.990	0.990
Latitude	**0.023***	**0.029***	**0.031***	**0.028***
NLCD Land Cover Class	**0.033***	**0.037***	**0.030***	**0.033***
Percent Impervious Surface (Mean)	0.128	0.132	0.111	0.131
Percent Tree Cover	0.287	0.318	0.282	0.300
Population	0.693	0.738	0.710	0.725
Housing Density	0.657	0.697	0.675	0.692
Distance to City Center	**0.038***	**0.028***	**0.022***	**0.028***
Water Table Depth	0.808	0.816	0.868	0.854
Elevation	0.636	0.604	0.540	0.566
NDVI	**0.024***	**0.028***	**0.013**	**0.016***
Site	**0.003***	**0.003***	**0.004***	**0.001***
Primary Class	0.518	0.481	0.439	0.425
Secondary Class	0.112	0.100*	0.099*	0.082
Season	0.486	0.506	0.513*	0.450
Collection Event	0.917	0.895	0.906	0.900

Table 3.4. Significance values for predictors of environmental and temporal variables used in analyses, using the envfit (at k=2) and ordistep functions, respectively. Values are shown for both the three-year June to October dataset, as well as the one-year April-October dataset. P-values are simulated with 999 permutations. Values in bold were included in the constrained analysis.

Predictor	***EnvFit (p-value)***	***OrdiStep (p-value)***	***EnvFit (p-value)***	***OrdiStep (p-value)***
	2019-2021, June-October		*2021, April-October*	
Longitude	0.959	0.005*	0.182	0.005*
Latitude	**0.023***	**0.010***	0.001*	0.090
NLCD Land Cover Class	**0.033***	**0.005***	**0.001***	**0.005***
Percent Impervious Surface (Mean)	0.128	0.005*	**0.001***	**0.005***
Percent Tree Cover	0.287	0.005*	**0.001***	**0.005***
Population	0.693	0.045*	0.465	0.185
Housing Density	0.657	0.040*	0.433	0.250
Distance to City Center	**0.038***	**0.005***	**0.001***	**0.005***
Water Table Depth	0.808	0.010*	0.176	0.070
Elevation	0.636	0.020*	0.366	0.335
NDVI	**0.024***	**0.025***	0.082	0.155
Site	**0.003***	**0.005***	**0.001***	**0.005***
Primary Class	**0.518**	**0.005***	**0.001***	**0.005***
Secondary Class	**0.112**	**0.005***	**0.001***	**0.010***
Season	0.486	0.095	**0.001***	**0.005***
Collection Event	0.917	--	0.011*	--

Supplemental Figures

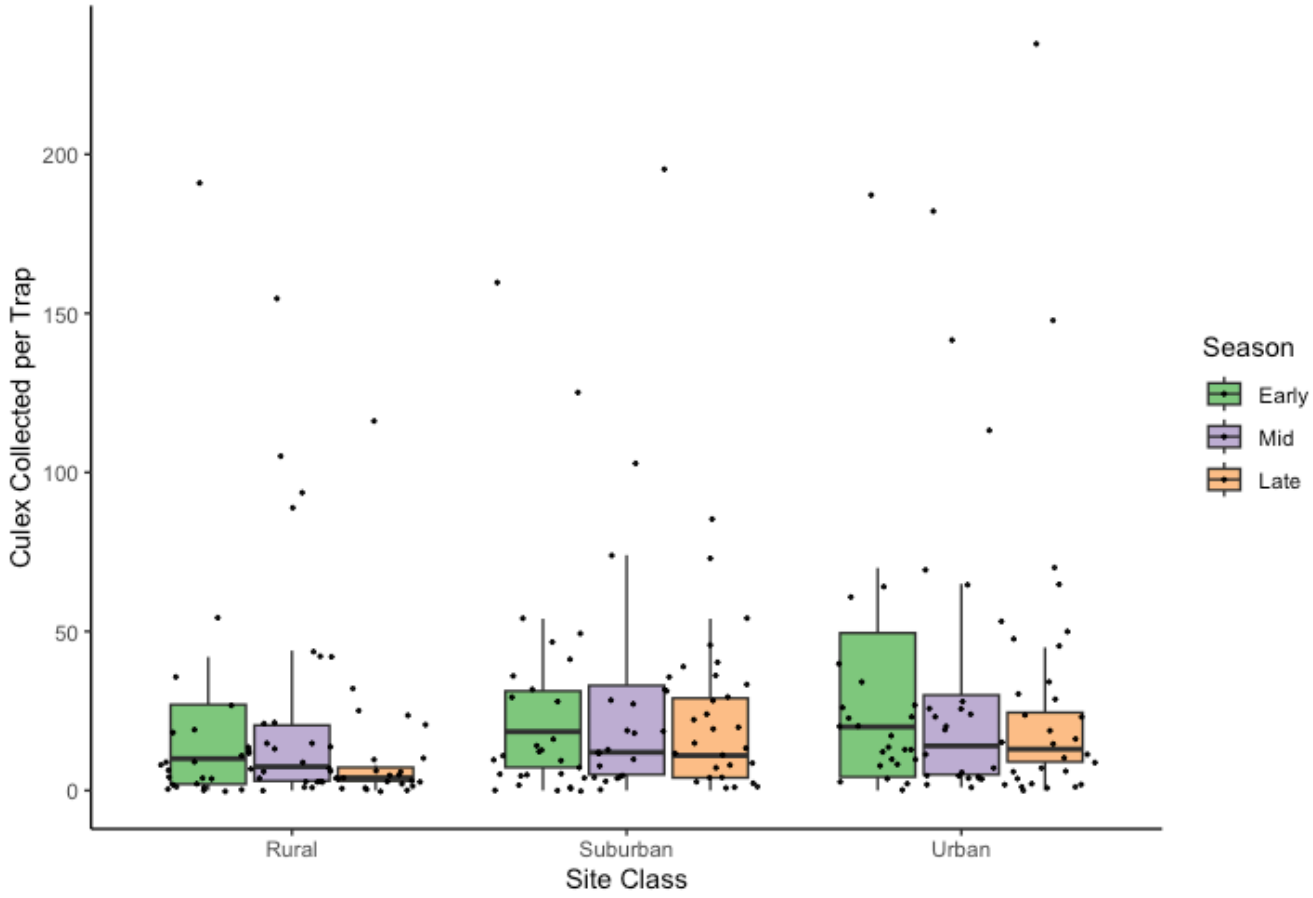

Figure 31. Mean and standard error of unfed cryptic *Culex* collected per trap by site and season in June-October 2019-2021.

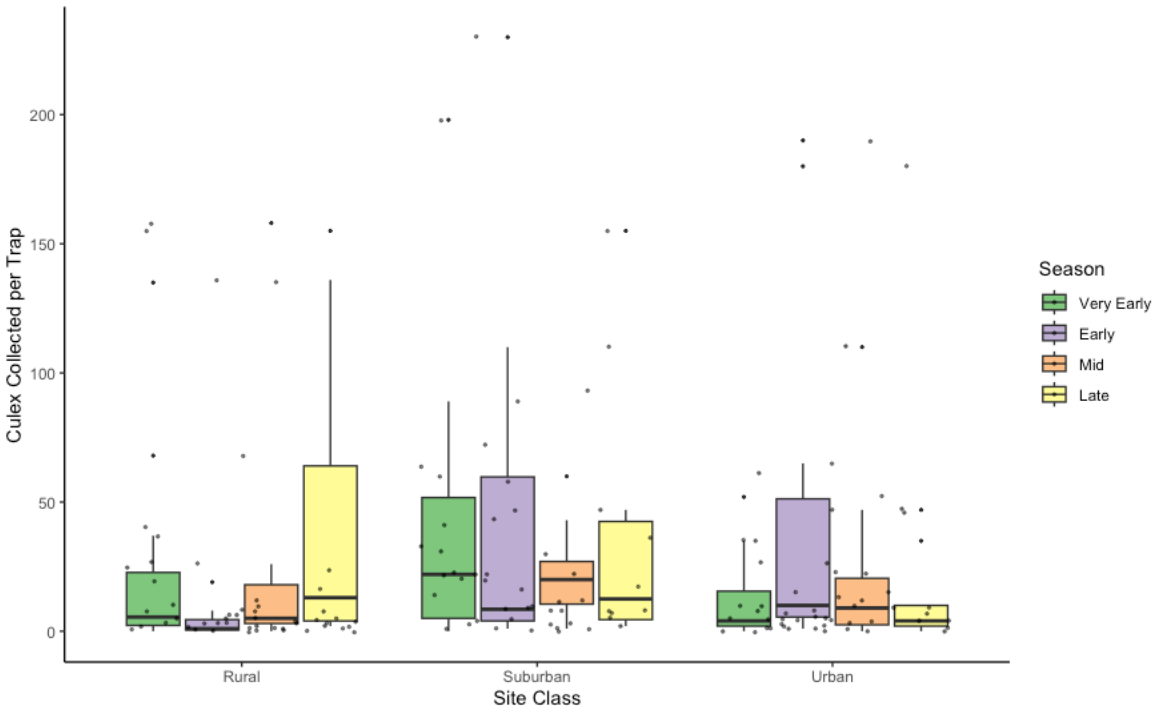

Figure 3.2. Mean and standard error of unfed cryptic *Culex* collected per trap by site and season in April-October 2021.

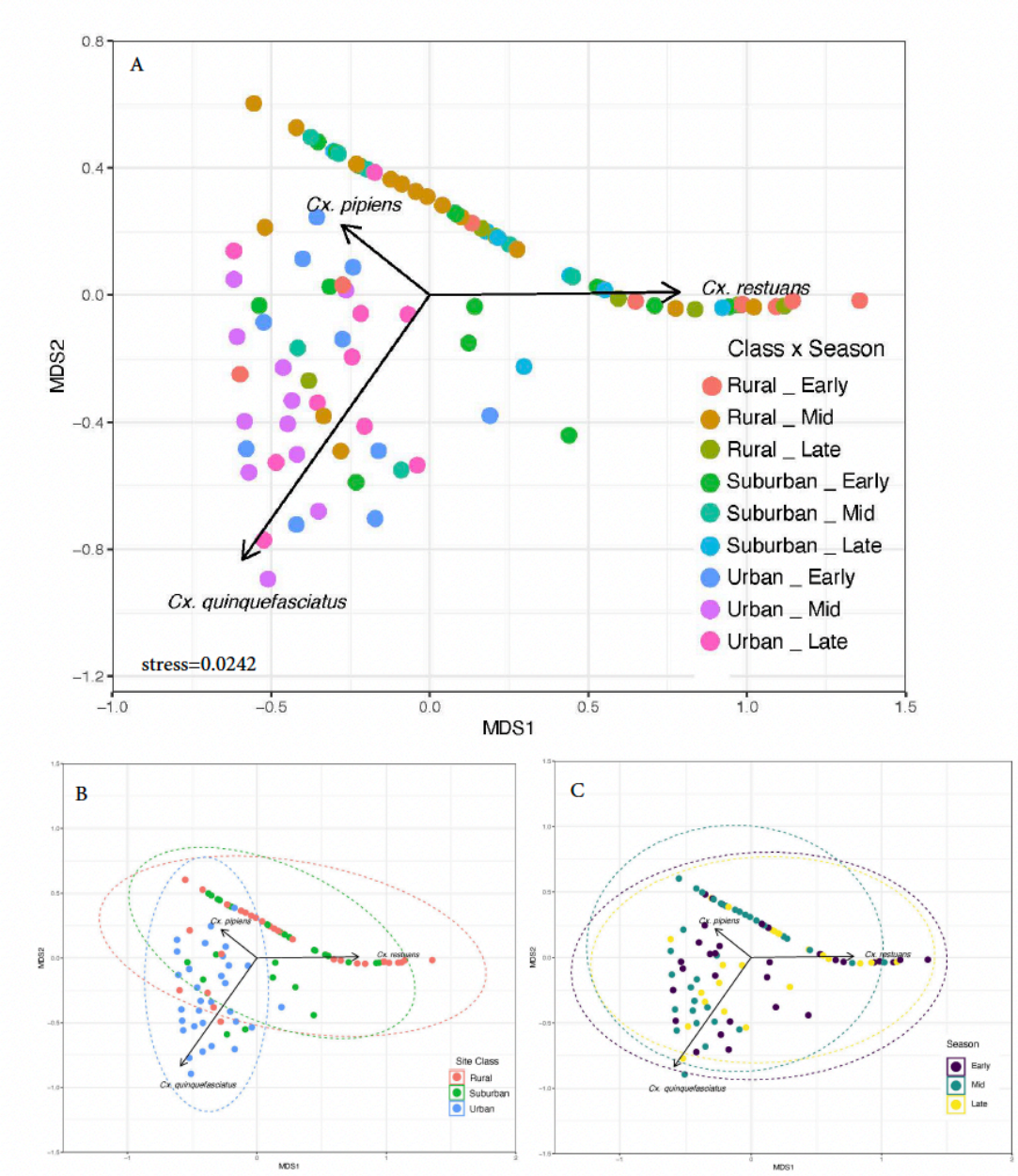

Figure 3.3. NMDS ordination plots for June-October 2021. The function `ordiellipse` was used to overlay group-level polygons on the NMDS ordinations. The generated polygons and branches represent the centroids plus 95% confidence intervals for each group. A. NMDS of community assemblage dissimilarity points, colored by site and season. B. Ellipses show clustering of points by site class. C. Ellipses show clustering by season.

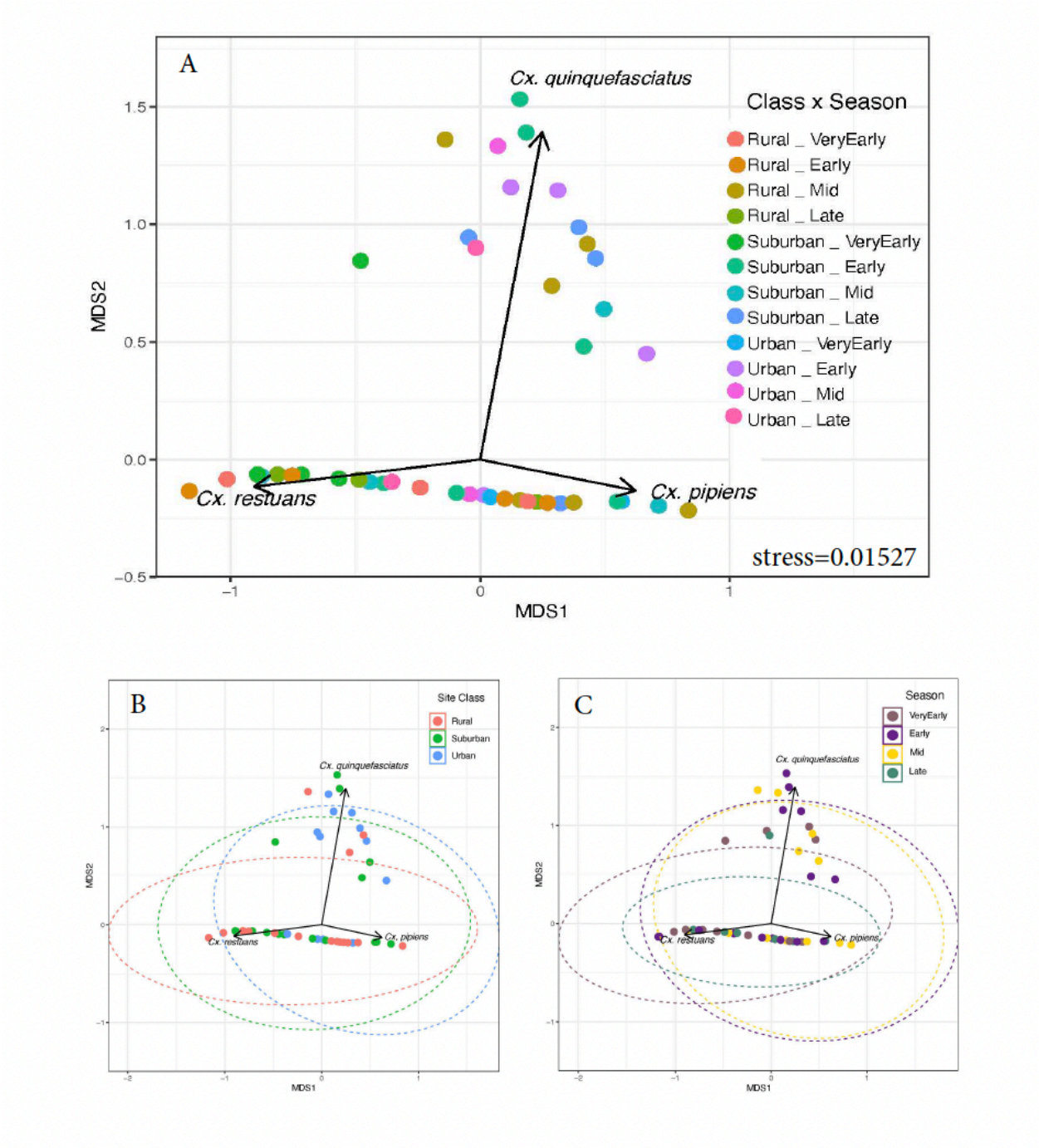

Figure 3.4. NMDS ordination plots for April-October 2021. The function `ordiellipse` was used to overlay group-level polygons on the NMDS ordination. The generated polygons and branches represent the centroids plus 95% confidence intervals for each group. A. NMDS of community assemblage dissimilarity points, colored by site and season. B. Ellipses show clustering of points by site class. C. Ellipses show clustering by season.

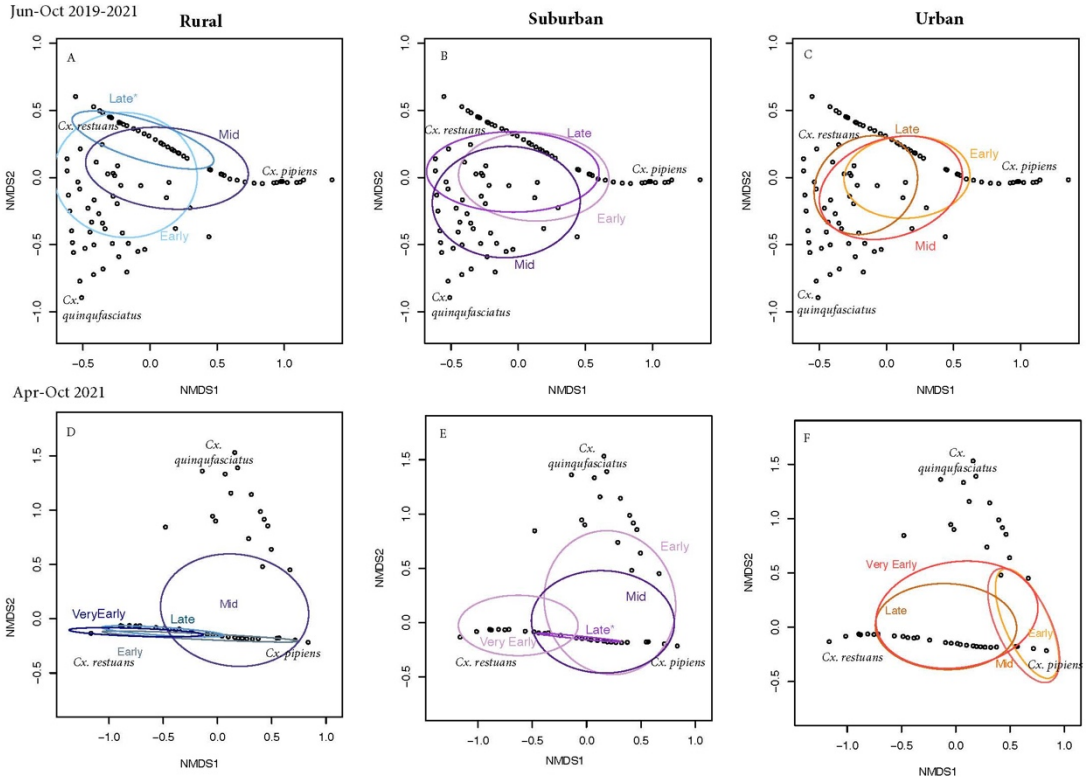

Figure 3.5. NMDS ordinations with 95% CI ellipses drawn for seasons within each site class. A-C: June-October 2019-2021. D-F: April-October 2021. Ellipses indicated with an asterisk (*) indicate points are clustered significantly more than by random chance.

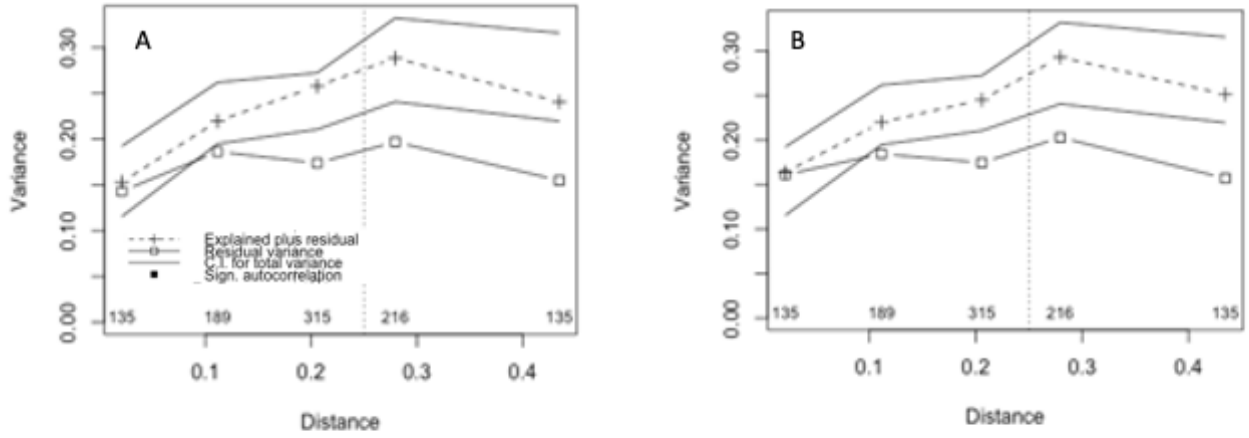

Figure 3.6. Multiscale ordination (MSO) plots for 2019-2021 CCA models. Distance is measured by differences in geographic coordinates between sites. Filled points indicate that model residuals are significantly autocorrelated at any distance lags. None are pictured. A. Model 1: Primary Site Class+ Distance to City Center (DCC) + Latitude (Y). B. Model 2: CCA tri-plot for model 2: Secondary Site Class + Percent Impervious Surface +Distance to City Center (DCC) +Latitude (Y).

Appendix II: Chapter 4 Supplement

Supplemental Text

Supplemental Methods

Trap specifications for each locality:

All collections were undertaken using gravid traps with an external battery source and a hay-infused oviposition attractant. Minor modifications were used according to the normal procedures for each team of investigators at each locality. In MD, we used CDC gravid traps (John Hock Co., Gainesville FL, USA), and the oviposition attractant comprised of 83g of alfalfa hay in 19L of tap water in a five-gallon bucket and generally brewed for two weeks, but at least seven days. In IL, the same CDC gravid trap was used. The oviposition attractant was made in 9.5 L batches infused with 75g of rabbit pellets, 58 g of fresh grass, and 2.5g Whey protein, brewed outside for two weeks. In PA, we used the BioQuip 2800 series gravid traps (Rancho Domingues, CA, USA) with a hay and lactalbumin infusion.

Differentiation of Cx. pipiens assemblage, Cx. restuans and Cx. salinarius

Cryptic *Culex* specimens were differentiated via a multiplexed PCR, targeting the 28S ribosomal subunit to differentiate *Cx. pipiens* s.l., *Cx. restuans* and *Cx. salinarius* (Crabtree et al. 1995). Each 20µl PCR reaction included 1µl of template DNA, 0.1µl of GoTaq polymerase (Promega GoTaq, cat. #PRD4005), 11.7 µl of molecular grade water, 4.0µl of 5X GoTaq reaction buffer containing 1.5mM $MgCl_2$, 1.6µl of 2.5mM dNTP mix, and 0.4µl of each of the following primers at 10 mM concentration: PQ10 (5'- CCTATGTCCGCGTATACTA -3'), R6 (5'-CCAAACACCGGTACCCAA-3'), S20 (5'-TGAGAATACATACCACTGCT-3'), and the reverse, CP16 (5'-GCGGGTACCATGCTTAAATTTAGGGGGTA-3'). The thermocycler

program was set to one cycle of 94°C for 5 minutes, 33 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, with a final extension step of 72°C for 7 minutes (Crabtree et al. 1995, Rochlin et al. 2007).

Identification of Cx. quinquefasciatus *ancestry*

A second PCR was applied to *Cx. pipiens* s.l. individuals to identify *Cx. quinquefasciatus* homozygotes or *pipiens-quinquefasciatus* heterozygotes from *Cx. pipiens* homozygotes at the acetylcholinesterase (ACE2) locus (Smith and Fonseca 2004). Each 20µl PCR included 1µl of template DNA, 0.1µl of GoTaq polymerase, 10.9 µl of molecular grade water, 4.0µl of 5X GoTaq reaction buffer containing 1.5mM $MgCl_2$, 1.6µl of 2.5mM dNTP mix, and 0.8 µl of each of the following primers at 10 mM concentration: ACEpip (5'-GGAAACAACGACGTATGTACT-3'), ACEquin (5'-CCTTCTTGAATGGCTGTGGCA-3'), and the reverse B1246s (5'-TGGAGCCTCCTCTTCACGG-3'). Thermalcycler conditions included initial denaturation at 94°C for 5 minutes, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, and a final extension step of 5 minutes at 72°C. Any specimens identified as *Cx. pipiens-Cx. quinquefasciatus* heterozygotes were classified as *Cx. pipiens*, the dominant taxon in the Northeast region.

Supplemental Tables

Table 4.1. Ranges generated for each environmental feature in Greater Washington DC for site classes were used to generalize the site classifications for collections from greater Chicago, IL and greater Philadelphia, PA. (*-Percent Impervious Surface Range includes one belowground site in Washington DC, so was forced to be 100%. **-Highly commercialized sites have low residential population but high human density on a daily basis.)

Site Classification	% Impervious Mean	% Impervious Range	NDVI (0-200)	Mean Population
Rural	<10%	<10%	>80	<300
Suburban	10%-50%	10-30%	70-80	150-1000
Urban	>50%	50-100%*	60-80	>1000**

Table 4.2. Abbreviations, descriptions, and data sources for variables used in the gradient forest analysis.

Variable	Variable Description
WOY	Week of the Year during which collection occurred
X	Longitude (°W)
Y	Latitude (°N)
DayLength.1	Mean daily daylight (sec) in days 1-7 prior to the collection date (Daymet Database)
DayLength.2	Mean daily daylight (sec) in days 8-14 prior to the collection date (Daymet Database)
Tmax.1	Mean daily temperature (°C) for days 1-7 prior to the collection date (Daymet Database)
Tmax.2	Mean daily temperature (°C) for days 8-14 prior to the collection date (Daymet Database)
prcp.1	Mean daily precipitation (mm) for days 1-7 prior to the collection date (Daymet Database)
prcp.2	Mean daily precipitation (mm) for days 8-14 prior to the collection date (Daymet Database)
%Imp	Mean percent impervious surface in 0.5km^2 buffer around collection site (NLCD 2016)
%Tree	Mean percent tree cover in 0.5km^2 buffer around collection site (NLCD 2016)
NDVI	Mean Normalized Differential Vegetative Index (0-200) in 0.5km^2 buffer around collection site
DCC	Linear distance from collection site to focal city (Washington, D.C., Philadelphia, Chicago)
Hous	Number of housing units in in 0.5km^2 buffer around collection site (2010 US Census)
Pop	Population residing in 0.5km^2 buffer around collection site (2010 US Census)
WaterTbl	Water Table Depth at collection site (cm)

Supplemental Figures

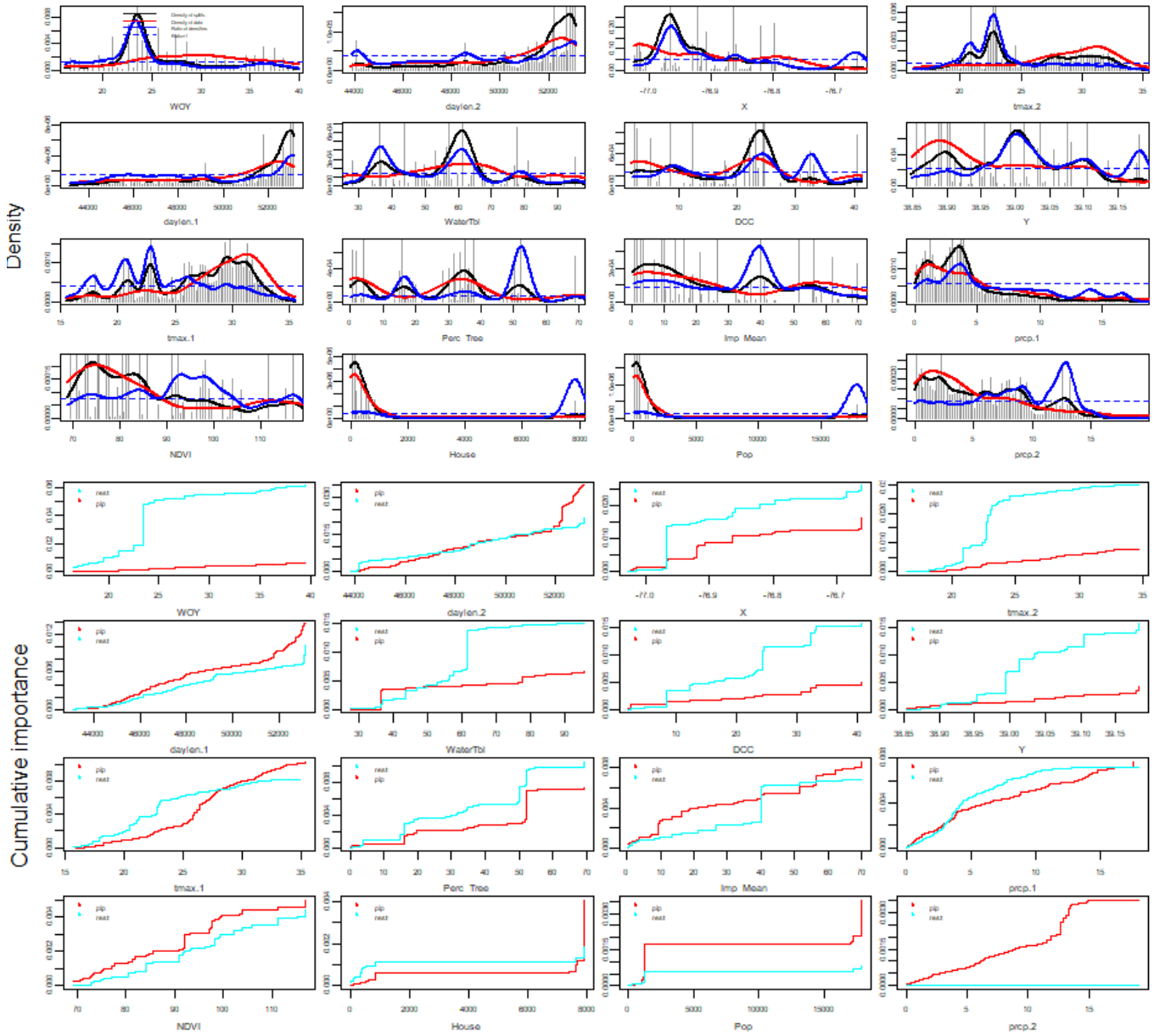

Figure 4.1. Gradient Forest (GF) results for the MD dataset. A. Location, density, and importance of GF splits along gradients of the top-ranked 16 environmental predictors. B. Cumulative distributions of standardized splits importance for each species. Monotonic functions for *Cx. restuans* are in teal and *Cx.pipiens* are in red. Descriptions of abbreviations used are in Appendix II, Table 4.2.

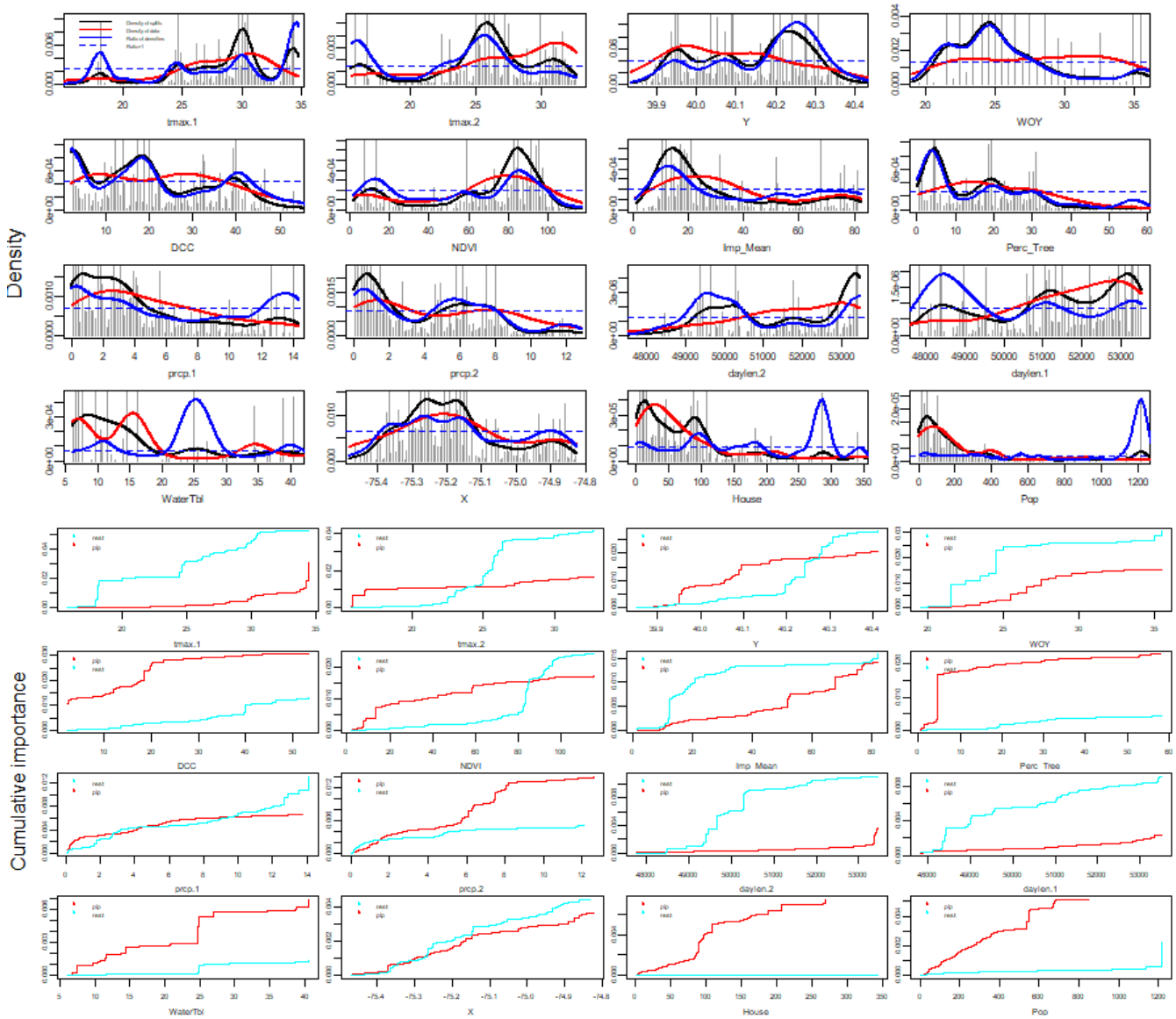

Figure 4.2. Gradient forest (GF) results for the PA dataset. A. Location, density, and importance of GF splits along gradients of the top-ranked 16 environmental predictors. B. Cumulative distributions of standardized splits importance for each species. Monotonic functions for *Cx. restuans* are in teal and *Cx. pipiens* are in red. Descriptions of abbreviations used are in Appendix II, Table 4.2.

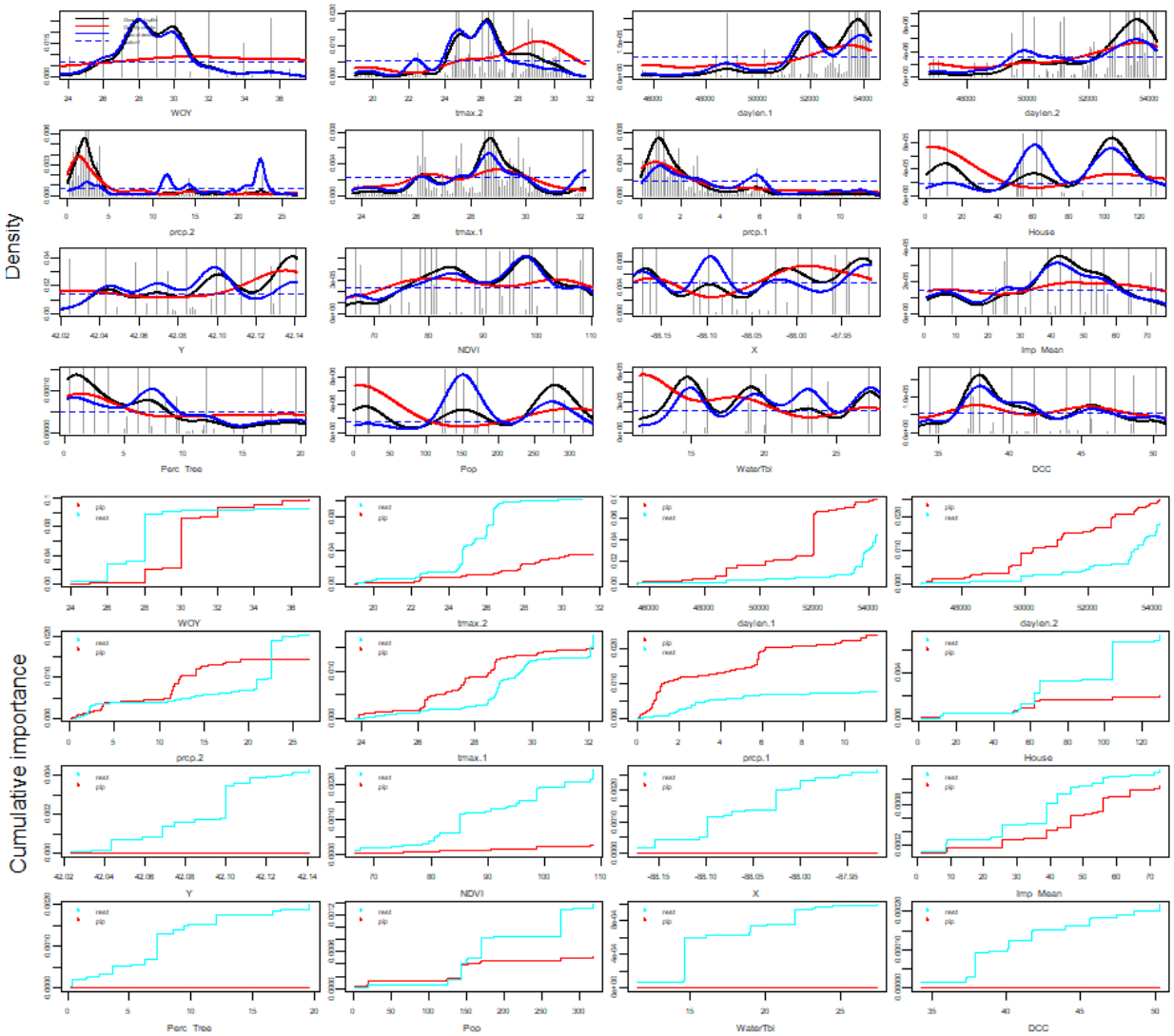

Figure 4.3. Gradient forest (GF) results for the IL dataset. A. Location, density, and importance of GF splits along gradients of the top-ranked 16 environmental predictors. B. Cumulative distributions of standardized splits importance for each species. Monotonic functions for *Cx. restuans* are in teal and *Cx. pipiens* are in red. Descriptions of abbreviations used are in Appendix II, Table 4.2.

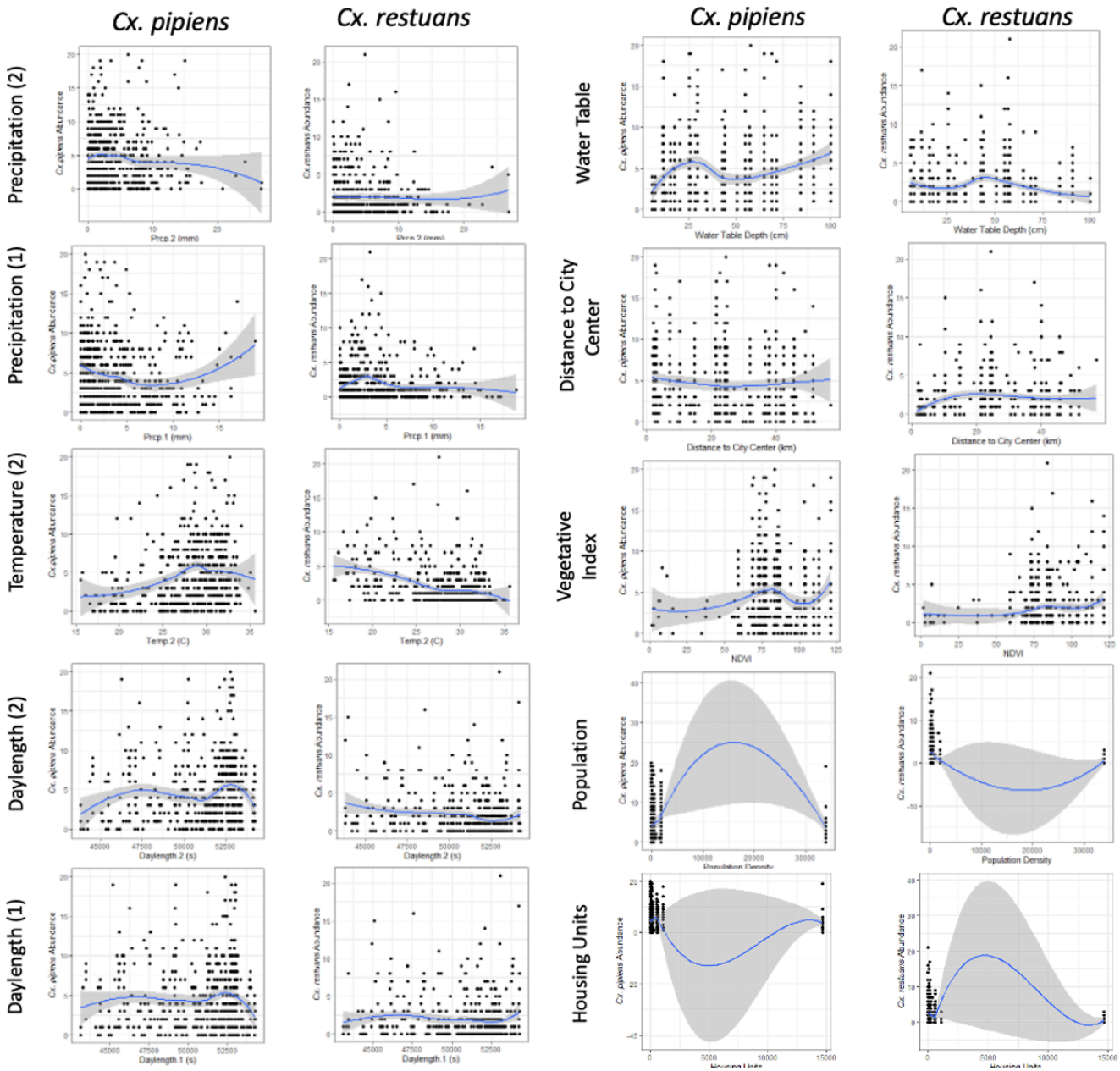

Figure 4.4. Species abundance responses to environmental variables not included in Figure 5. *Culex pipiens* responses are on the left of each group and *Cx. restuans* are on the right. Trends are represented by regression between species abundance and environmental gradient with locally estimated scatterplot smoothing plus 95% confidence intervals for each predictor.

Appendix III: Chapter 5 Supplement

Supplemental Text

Supplemental Methods

Trap specifications for each locality:

All collections were undertaken using gravid traps with an external battery source and a hay-infused oviposition attractant. Minor modifications were used according to the normal procedures for each team of investigators at each locality. In MD, we used CDC gravid traps (John Hock Co., Gainesville FL, USA), and the oviposition attractant comprised of 83g of alfalfa hay in 19L of tap water in a five-gallon bucket and generally brewed for two weeks, but at least seven days. In IL, the same CDC gravid trap was used. The oviposition attractant was made in 9.5 L batches infused with 75g of rabbit pellets, 58 g of fresh grass, and 2.5g Whey protein, brewed outside for two weeks. In PA, we used the BioQuip 2800 series gravid traps (Rancho Domingues, CA, USA) with a hay and lactalbumin infusion.

Statistical associations between landscape-level predictors of molestus *allele frequency*

To examine whether site classification, mean percent impervious surface surrounding a site, and NDVI were correlated with one another, we initially used linear regression models with gaussian error structures, after checking for heteroscedasticity, skewedness, and kurtosis of residuals. The full models used to test for associations between primary site classification and each continuous landscape-level predictor were as follows:

$$Model\ S1:\ ImSur \sim (Loc \times Class + \varepsilon)$$

$$Model\ S2:\ NDVI \sim (Loc \times Class + \varepsilon)$$

where Loc = locality, ImSur = mean percent impervious surface, NDVI = NDVI, and Class = site classification. Test statistics for associations between site classification, locality and their interactions were calculated by comparing the full model above with a reduced model (without one model term), as implemented with by the `lrtest` function (lmtest package v. 0.9-40; Zeileis and Hathorn 2002).

We also examined the variance inflation factors (VIF) for terms in the following logistic regression model with a binomial error structure.

$$Model\ S3:\ P(mol,\ 1\text{-}mol) \sim (ImSur + NDVI + Class + \varepsilon)$$

We considered VIF estimates between 1 - 5 for a model term to indicate a moderate correlation with predictors, and VIF estimates between 5 - 10 for a model term to indicate a strong correlation with other model terms.

NDVI models for individual localities

NDVI values less than 50 were underrepresented in MD and IL, so we developed NDVI models for each locality independently as to include all sites. approach described above, but the initial complex model included the NDVI*Class interaction.

$$PA:\ P(mol,\ 1\text{-}mol) \sim 1 + NDVI + \varepsilon$$

$$IL:\ P(mol,\ 1\text{-}mol) \sim 1 + NDVI + \varepsilon$$

$$MD:\ P(mol,\ 1\text{-}mol) \sim 1 + NDVI + \varepsilon$$

Supplemental Results

Statistical associations between landscape-level predictors of molestus *allele frequency*

Using a likelihood ratio test, we found that both MPIS and NDVI were impacted by the interaction of site class and locality (MPIS: $\chi 2=63.028$, df=4, $p<<0.001$, NDVI: $\chi 2=173.54$, df=4, $p<<0.001$, Appendix III, Figure 5.2). We also found strong correlations between continuous predictors in the full model (Model S3).

NDVI models for individual localities

In PA and IL, the model including NDVI as a predictor did not explain more variation in the data than the intercept-only model, (PA: $\chi 2=2.091$, df=1, p=0.148, IL: $\chi 2=0.075$, df=1, p=0.785). For MD, the model including NDVI explained more variation than the intercept-only model ($\chi 2=12.251$, df=1, p=0.0004). In MD, there was a small but highly significant decrease in odds of having a *molestus* allele will increasing NDVI (estimated coefficient=-0.02, 95%CI [-0.03, -0.01], p<0.001.

Supplemental Tables

Table 5.1. Coordinates, site classification, mean percent impervious surface and mean NDVI for each collection sites.

Locality	Site	Latitude	Longitude	Site Class	Mean Percent Impervious Surface	Mean NDVI
IL	R1	-88.18128	42.05359	Rural	1.4	94
IL	R2	-88.16373	42.11548	Rural	0.95	113
IL	R3	-88.05042	42.14014	Rural	16.39	104
IL	S1	-87.91722	42.14106	Suburban	42.99	76
IL	S2	-87.99897	42.08368	Suburban	34.88	87
IL	S3	-88.14585	42.03345	Suburban	49.62	103
IL	U1	-87.91975	42.12402	Urban	65.49	74
IL	U2	-87.96958	42.013858	Urban	80.32	59
IL	U3	-87.99876	42.14134	Urban	62.37	83
MD	Beltsville	-76.82588	39.01231	Rural	0.6	113.07
MD	Clarksville	-76.92767	39.25637	Rural	0.4	120.9
MD	Patuxent	-76.817	39.05431	Rural	5.3	83.33
MD	SERC	-76.55342	38.88785	Rural	0.7	98.35

MD	UM	-76.77863	38.85905	Rural	2.2	109.41
MD	Bowie	-76.77058	38.97728	Suburban	18.8	85.15
MD	Cheverly	-76.91429	38.91504	Suburban	29.2	73.17
MD	CollPark	-76.93413	39.01578	Suburban	23.7	72.51
MD	Jerald	-76.90164	39.1101	Suburban	13.9	77.56
MD	OldBarn	-77.0018	39.0992	Suburban	13.2	73.3
MD	AirSpace	-77.01986	38.88816	Urban	60.4	77.68
MD	ColHeights	-77.02741	38.92955	Urban	75.8	69.45
MD	Cst	-77.01776	38.88587	Urban	64.2	68.69
MD	JBAB	-77.01188	38.84722	Urban	50.6	85.92
MD	NatHist	-77.02606	38.89126	Urban	52.7	76.61
PA	PA_21_10	-75.25547	39.89228	Rural	13.867	96.86
PA	PA_19_12	-75.48336	39.90705	Rural	1.895	117.17
PA	PA_19_49	-75.197836	39.9806142	Rural	12.62	100.33
PA	PA_21_01	-75.24534	40.06053	Rural	11.719	101.59
PA	PA_19_41	-74.959635	40.0627527	Rural	1.518	109.88
PA	PA_19_36	-75.28527	40.0863	Rural	15.584	94.27
PA	PA_19_54	-75.45026	40.12671	Rural	41.958	55.27
PA	PA_19_42	-75.062495	40.15859	Rural	26.36	76.24
PA	PA_19_52	-75.151978	40.288458	Rural	12.687	93.74
PA	PA_19_50	-75.347784	40.3314491	Rural	18.426	84.56
PA	PA_21_05	-74.93625	40.33281	Rural	0.788	112.53
PA	PA_19_10	-75.207164	40.3686667	Rural	16.692	85.452
PA	PA_19_17	-75.44673	39.83711	Suburban	32.46	70.096
PA	PA_19_18	-75.39371	39.86372	Suburban	30.691	71.082
PA	PA_20_06	-75.25318	39.89322	Suburban	23.43	75.524
PA	PA_19_48	-75.31944	39.9414	Suburban	28.077	80.511
PA	PA_19_39	-75.30783	39.94208	Suburban	26.827	75.641
PA	PA_19_15	-75.305529	39.9731919	Suburban	46.005	46.637*
PA	PA_19_11	-75.31649	39.98378	Suburban	41.953	60.294
PA	PA_19_51	-75.2936	39.9905	Suburban	27.21	78.941
PA	PA_19_46	-75.33729	40.00813	Suburban	11.789	93.656
PA	PA_20_09	-75.110804	40.0315008	Suburban	51.823	36.995
PA	PA_21_08	-75.04382	40.04947	Suburban	31.006	65.305
PA	PA_19_20	-75.239283	40.057798	Suburban	26.673	73.598
PA	PA_21_04	-74.99026	40.10062	Suburban	11.839	90.12
PA	PA_20_03	-74.872358	40.1232174	Suburban	21.783	80.713
PA	PA_20_05	-74.844888	40.1406778	Suburban	31.088	72.35
PA	PA_19_38	-74.871615	40.1456878	Suburban	28.007	79.382
PA	PA_19_03	-74.910132	40.1548073	Suburban	33.891	66.976

PA	PA_19_33	-75.31387	40.16142	Suburban	15.111	91.188
PA	PA_20_04	-74.840818	40.1799563	Suburban	39.205	60.72
PA	PA_19_47	-75.085756	40.1900279	Suburban	40.197	56.722
PA	PA_21_02	-75.06053	40.20303	Suburban	29.327	74.023
PA	PA_19_06	-74.902274	40.2324967	Suburban	30.084	66.705
PA	PA_21_09	-75.14123	40.28656	Suburban	9.222	97.589
PA	PA_20_01	-75.141436	40.3208204	Suburban	27.502	73.126
PA	PA_21_06	-75.3386	40.45173	Suburban	18.237	84.669
PA	PA_19_43	-75.214253	39.9351508	Urban	64.638	24.475*
PA	PA_20_10	-75.21044	39.9410049	Urban	51.402	37.286*
PA	PA_19_40	-75.147737	39.9437179	Urban	76.481	8.315*
PA	PA_20_08	-75.214171	39.9443412	Urban	68.776	11.359*
PA	PA_19_24	-75.221745	39.9567908	Urban	80.813	2.807*
PA	PA_19_53	-75.197285	39.9675677	Urban	66.71	15.222*
PA	PA_19_25	-75.228242	39.967911	Urban	73.719	6.755*
PA	PA_19_23	-75.175254	39.9747089	Urban	77.077	5.909*
PA	PA_21_03	-75.17567	39.97868	Urban	76.285	1.703*
PA	PA_19_26	-75.148683	39.9839593	Urban	77.843	2.608*
PA	PA_19_37	-75.270922	39.9910532	Urban	15.975	91.673
PA	PA_20_07	-75.120046	40.0170057	Urban	53.118	38.32*

*Denotes sites with mean NDVI less than 50. These sites were excluded from the multi-locality NDVI model.

Table 5.2. Amplicon sizes for each taxon in the PCR reactions. Amplicon sizes were used to identify individuals via gel electrophoresis as described in the text.

PCR Reaction	Taxon	Fragment Size (bp)
28S *(Crabtree et al. 1995)*	*Cx. pipiens* assemblage	698
	Cx. restuans	506
	Cx. salinarius	175
ACE2 *(Smith and Fonseca 2004)*	*Cx. pipiens*	610
	Cx. quinquefasciatus	247
CQ11 *(Bahnck and Fonseca 2006)*	*Cx. pipiens pipiens*	258-266
	Cx. pipiens molestus	284

Table 5.3. Calculated mean proportion *molestus* for site classes at each locality and estimated marginal means with 95% confidence intervals for localities from each landscape-level predictor. Estimated marginal means were generated from the best fitting generalized linear models for characteristics of urbanization, MPIS and NDVI. Only sites with NDVI>50 are included. Estimates were generated using `estimate_means` in R package `modelbased` (Makowski et al. 2020).

Loc	Site Class	N	Raw Mean	Mean proportion *molestus* (MPIS Model)	95% CI	Mean proportion *molestus* (NDVI Model)	95% CI
MD	Rural	83	0.51	0.63	[0.59, 0.67]	0.62	[0.58, 0.66]
	Suburban	96	0.65				
	Urban	93	0.70				
IL	Rural	77	0.49	0.47	[0.42, 0.53]	0.48	[0.43, 0.54]
	Suburban	54	0.41				
	Urban	51	0.62				
PA	Rural	24	0.22	0.29	[0.24, 0.36]	0.26	[0.19, 0.35]
	Suburban	53	0.30				
	Urban	36/1*	0.33/0.0				

*Only one urban PA site had an NDVI>50, so only one instance is included for the NDVI model.

Supplemental Figures

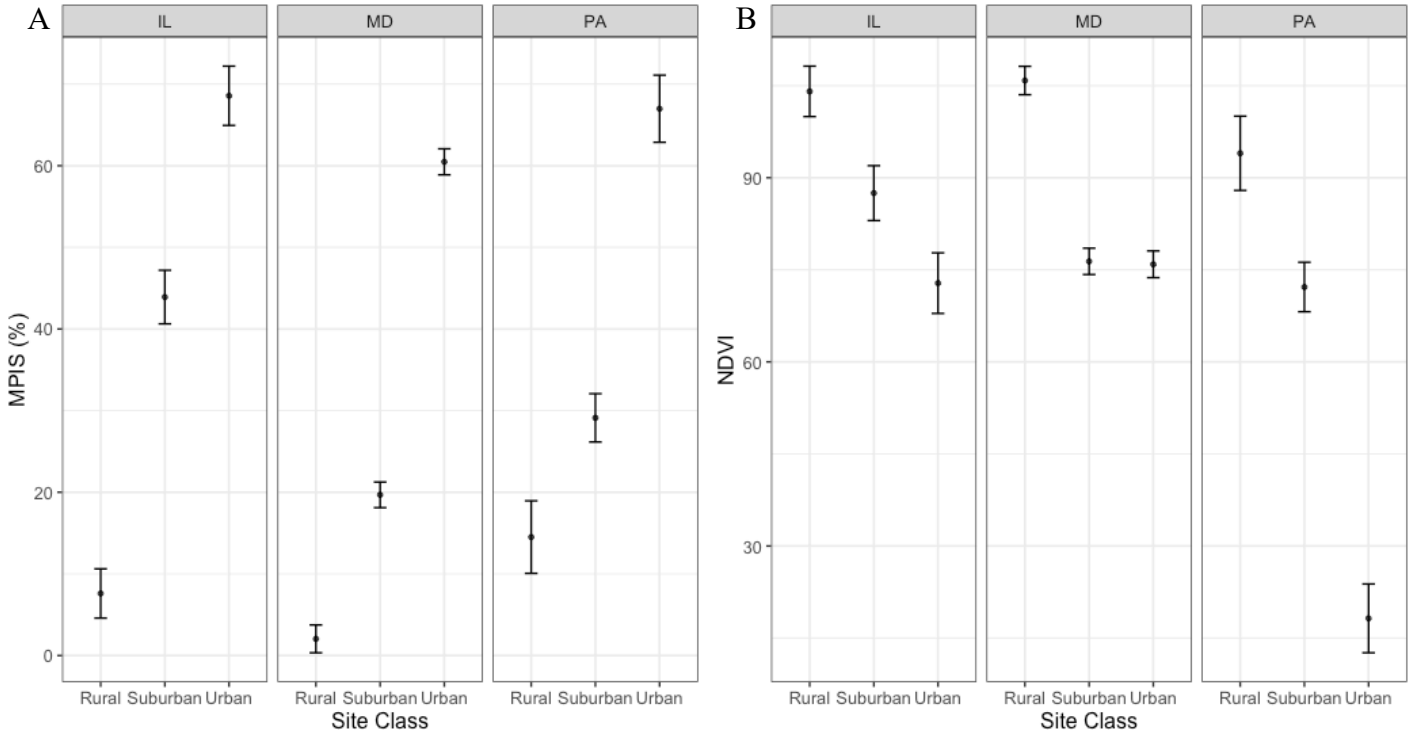

Figure 5.1. Mean and 95% confidence intervals for continuous environmental variable values by site class for each locality. IL=Illinois, MD=Washington, D.C. and Maryland, PA= Pennsylvania. A. Mean percent impervious surface (MPIS). B. Normalized Differential Vegetative Index (NDVI).

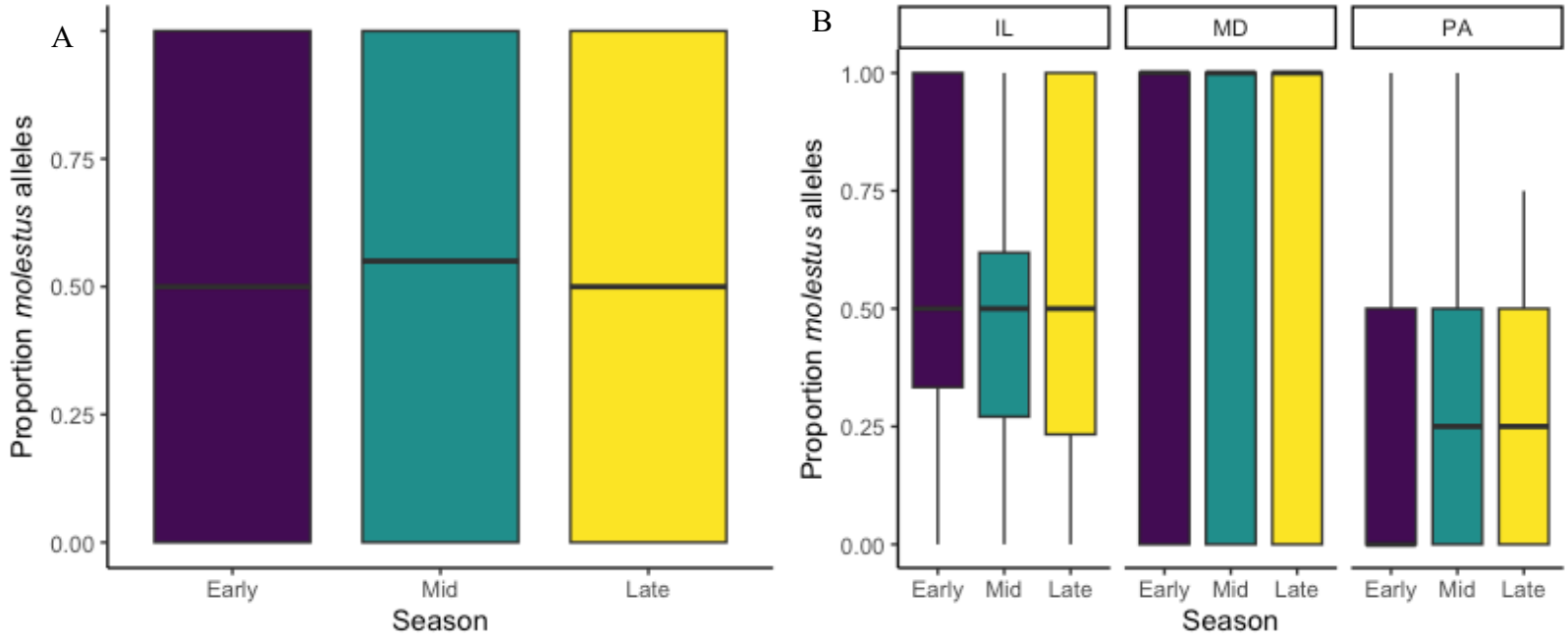

Figure 5.2. *Molestus* allele frequencies in early (purple), mid (teal), and late (yellow) season. A. Median *molestus* allele frequencies plus the first through third interquartile ranges over the course of the season aggregated across localities. B. Median *molestus* allele frequencies (plus the first through third interquartile ranges) over the course of the season, separated by locality.

References

Alves, B.M., Cargnelutti Filho, A. and Burin, C. (2017) Multicollinearity in canonical correlation analysis in maize. *Genetics and Molecular Research*, *16*(1), 1-14.

Andreadis, T.G. (2012) The contribution of *Culex pipiens* complex mosquitoes to transmission and persistence of West Nile Virus in North America. *Journal of the American Mosquito Control Association*, 28, 137–151.

Andreadis, T.G., Armstrong, P.M. (2007) A two-year evaluation of elevated canopy trapping for *Culex* mosquitoes and West Nile virus in an operational surveillance program in the northeastern United States. *Journal of the American Mosquito Control Association,* 23, 137–48.

Andreadis, T. G., Armstrong, P. M., Bajwa, W. I. (2010). Studies on hibernating populations of *Culex pipiens* from a West Nile virus endemic focus in New York City: parity rates and isolation of West Nile virus. *Journal of the American Mosquito Control Association*, 26(3), 257-264.

Andreadis, T.G., Anderson, J.F., Vossbrinck, C.R. (2001) Mosquito surveillance for West Nile virus in Connecticut, 2000: Isolation from *Culex pipiens, Cx. restuans, Cx. salinarius*, and *Culiseta melanura*. *Emerging Infectious Diseases*, 7, 670–674.

Arich, S., Haba, Y., Assaid, N., Fritz, M.L., McBride, C.S., Weill, M., Taki, H., Sarih, M.H. and Labbé, P. (2022) No association between habitat, autogeny and genetics in Moroccan *Culex pipiens* populations. *Parasites and Vectors*, *15*(1), 405.

Arsenault-Benoit A, Fritz, M. *In review*. Spatiotemporal organization of cryptic North American *Culex* species along an urbanization gradient. *Ecological Solutions and Evidence.*

Arsenault-Benoit, A., Greene, A., Fritz, M. (2021) Paved Paradise: Belowground parking structures sustain urban mosquito populations in Washington DC. *Journal of the American Mosquito Control Association*, 37, 291–295.

Bahnck, C.M., Fonseca, D.M. (2006) Rapid assay to identify the two genetic forms of *Culex (Culex) pipiens* L. (Diptera: Culicidae) and hybrid populations. *American Journal of Tropical Medicine and Hygiene*, 75, 251–255.

Barr, A.R. (1957) The distribution of *Culex p. pipiens* and *Cx. p. quinquefasciatus* in North America. *The American Journal of Tropical Medicine and Hygiene*, 6, 153–165.

Becker B, PT Leisnham, SL LaDeau. 2014. A tale of two city blocks: Differences in immature and adult mosquito abundances between socioeconomically different urban blocks in Baltimore (Maryland, USA). *International Journal of Environmental Research and Public Health.* 11, 3256–3270.

Bierne, N., Gagnaire, P.A., David, P. (2013) The geography of introgression in a patchy environment and the thorn in the side of ecological speciation. *Current Zoology*, 59, 72–86.

Blackmon H., Adams, R.H. (2015) evobiR: Comparative and population genetic analyses. R package version 1.1. https://CRAN.R-project.org/package=evobiR

Bondo, K. J., Montecino-Latorre, D., Williams, L., Helwig, M., Duren, K., Hutchinson, M. L., Walter, W. D. (2023). Spatial modeling of two mosquito vectors of West Nile virus using integrated nested Laplace approximations. *Ecosphere*, *14*(1), e4346.

Bowden, S.E., Magori, K., Drake, J.M. (2011) Regional differences in the association between land cover and West Nile virus disease incidence in humans in the United States. *American Journal of Tropical Medicine and Hygiene*, 84, 234–238.

Brady, O.J., Godfray, H.C.J., Tatem, A.J., Gething, P.W., Cohen, J.M., Ellis McKenzie, F., *et al.* (2016) Vectorial capacity and vector control: Reconsidering sensitivity to parameters for malaria elimination. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 110, 107–117.

Brown, H., Diuk-Wasser, M., Andreadis, T., Fish, D. (2008) Remotely-sensed vegetation indices identify mosquito clusters of West Nile virus vectors in an urban landscape in the Northeastern United States. *Vector-Borne and Zoonotic Diseases*, 8, 197–206.

Bugbee LM, Forte LR. 2004. The discovery of West Nile virus in overwintering *Culex pipiens* (Diptera: Culicidae) mosquitoes in Lehigh County, Pennsylvania. *Journal of the American Mosquito Control Association*, 20(3), 326-327.

Buth, J.L., Brust, R.A., Ellis, R.A., 1990. Development time, oviposition activity and onset of diapause in *Culex tarsalis*, *Culex restuans* and *Culiseta inornata* in southern Manitoba. *Journal of the American Mosquito Control Association*, *6*(1), 55-63.

Byrne, K., Nichols, R.A. (1999) *Culex pipiens* in London Underground tunnels: Differentiation between surface and subterranean populations. *Heredity*, 82, 7–15.

Chandrasegaran, K., Lahondère, C., Escobar, L.E., Vinauger, C. (2020) Linking mosquito ecology, traits, behavior, and disease transmission. *Trends in Parasitology*, 36, 393–403.

Chevillon, C., Eritja, R., Pasteur, N., Raymond, M. (1995) Commensalism, adaptation and gene flow: Mosquitoes of the *Culex pipiens* complex in different habitats. *Genetical Research*, 66, 147–157.

Chuang, T.W., Hildreth, M.B., Vanroekel, D.L., Wimberly, M.C. (2011) Weather and land cover influences on mosquito populations in Sioux Falls, South Dakota. *Journal of Medical Entomology*, *48*(3), 669-679.

Ciota, A.T., Kramer, L.D. (2013) Vector-virus interactions and transmission dynamics of West Nile virus. *Viruses*, 5, 3021–3047.

Ciota, A.T. (2017) West Nile virus and its vectors. *Current Opinion in Insect Science*, *22*, 28-36.

Ciota, A.T., Drummond, C.L., Drobnack, J., Ruby, M.A., Kramer, L.D., Ebel, G.D. (2011) Emergence of *Culex pipiens* from Overwintering Hibernacula. *Journal of the American Mosquito Control Association*, 27, 21–29.

Ciota, A.T., Matacchiero, A.C., Kilpatrick, A.M., Kramer, L.D., 2014. The effect of temperature on life history traits of *Culex* mosquitoes. *Journal of Medical Entomology*, *51*(1), 55-62.

Combs, M., Byers, K.A., Ghersi, B.M., Blum, M.J., Caccone, A., Costa, F., Himsworth, C.G., Richardson, J.L., Munshi-South, J. (2018) Urban rat races: spatial population genomics of brown rats (Rattus norvegicus) compared across multiple cities. *Proceedings of the Royal Society B*, *285*(1880), 20180245.

Crabtree, M.B., Savage, H.M., Miller, B.R. (1995) Development of a species-diagnostic polymerase chain reaction assay for the identification of *Culex* vectors of St. Louis encephalitis virus based on interspecies sequence variation in ribosomal DNA spacers. *The American Journal of Tropical Medicine and Hygiene*, 53, 105–109.

Crispo, E., Moore, J.-S., Lee-Yaw, J.A., Gray, S.M., Haller, B.C. (2011) Broken barriers: Human-induced changes to gene flow and introgression in animals. *BioEssays*, 33, 508–518.

DeGroote, J.P., Sugumaran, R., Brend, S.M., Tucker, B.J., Bartholomay, L.C. (2008) Landscape, demographic, entomological, and climatic associations with human disease incidence of West Nile virus in the state of Iowa, USA. *International Journal of Health Geographics*, 7, 1–16.

Di Luca, M., Toma, L., Boccolini, D., Severini, F., Rosa, G. La, Minelli, G., *et al.* (2016) Ecological distribution and CQ11 genetic structure of *Culex pipiens* complex (Diptera: Culicidae) in Italy. *PLoS ONE*, 11, 146476.

Duik-Wasser, M., Brown, H.E., Andreadis, T.G., Fish, D. (2006) Modeling the Spatial Distribution of Mosquito Vectors for West Nile Virus in Connecticut, USA. *Vector-Borne and Zoonotic Diseases*, 6, 283–295.

Dunphy, B.M., Kovach, K.B., Gehrke, E.J., Field, E.N., Rowley, W.A., Bartholomay, L.C., *et al.* (2019) Long-term surveillance defines spatial and temporal patterns implicating *Culex tarsalis* as the primary vector of West Nile virus. *Scientific Reports*, 9.

Dzul-Manzanilla F, Ibarra-López J, Marín WB, Martini-Jaimes A, Leyva JT, Correa-Morales F, Huerta H, Manrique-Saide P, Vazquez-Prokopec GM, Day J. (2017) Indoor resting behavior of *Aedes aegypti* (Diptera: Culicidae) in Acapulco, Mexico. *Journal of Medical Entomology*, 54, 501–504.

Ebel, G.D., Rochlin, I., Longacker, J., Kramer, L.D. (2005) *Culex restuans* (Diptera: Culicidae) relative abundance and vector competence for West Nile Virus. *Journal of Medical Entomology*, 42, 838–43.

Ellis, N., Smith, S.J., Pitcher, C.R., 2012. Gradient forests: calculating importance gradients on physical predictors. *Ecology*, *93*(1), 156-168.

Farajollahi, A., Fonseca, D.M., Kramer, L.D., Marm Kilpatrick, A. (2011) "Bird biting" mosquitoes and human disease: A review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infection, Genetics and Evolution*, 11, 1577-1585.

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*(1), 1-9.

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, 1–6.

Fonseca, D.M., Keyghobadi, N., Malcolm, C.A., Mehmet, C., Schaffner, F., Mogi, M., *et al.* (2004) Emerging vectors in the *Culex pipiens* complex. *Science*, 303, 1535–1538.

Fonseca, D.M., Smith, J.L., Kim, H.C., Mogi, M. (2009) Population genetics of the mosquito *Culex pipiens pallens* reveals sex-linked asymmetric introgression by *Culex quinquefasciatus*. *Infection, Genetics and Evolution*, 9, 1197–1203.

Fox and Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fritz, M.L., Walker, E.D., Miller, J.R., Severson, D.W., Dworkin, I. (2015) Divergent host preferences of above- and below-ground *Culex pipiens* mosquitoes and their hybrid offspring. *Medical and Veterinary Entomology*, 29, 115–123.

Garrett-Jones C. (1964a) Prognosis for Interruption of Malaria Transmission Through Assessment of the Mosquito's Vectorial Capacity. *Nature* 204, 1173–1175.

Garrett-Jones C. (1964b) The human blood index of malaria vectors in relation to epidemiological assessment. *Bulletin of the World Health Organization,* 30, 241-261.

Gilbert B, Srivastava DS, Kirby K.R. (2008) Niche partitioning at multiple scales facilitates coexistence among mosquito larvae. *Oikos*. 117, 944–950.

Gloor, G.B., Engels, W. (1992) Technique for preparation of fly DNA for use in PCR. *Drosophila Information Service, 71*, 148-149.

Gloria-Soria A, Lima A, Lovin DD, Cunningham JM, Severson DW, Powell JR. 2018. Origin of a high-latitude population of *Aedes aegypti* in Washington, DC. *American Journal of Tropical Medicine and Hygiene*, 98, 445–452.

Gomes, B., Sousa, C.A., Novo, M.T., Freitas, F.B., Alves, R., Côrte-Real, A.R., *et al.* (2009) Asymmetric introgression between sympatric *molestus* and *pipiens* forms *of Culex pipiens* (Diptera: Culicidae) in the Comporta region, Portugal. *BMC Evolutionary Biology*, 9.

Gomes, B., Sousa, C.A., Vicente, J.L., Pinho, L., Calderón, I., Arez, E., *et al.* (2013) Feeding patterns of *molestus* and *pipiens* forms of *Culex pipiens* (Diptera: Culicidae) in a region of high hybridization. *Parasites and Vectors*, 6.

Goodman H, Egizi A, Fonseca DM, Leisnham PT, LaDeau SL. (2008) Primary blood-hosts of mosquitoes are influenced by social and ecological conditions in a complex urban landscape. *Parasites and Vectors*, 11, 1–10.

Gorris, M.E., Bartlow, A.W., Temple, S.D., Romero-Alvarez, D., Shutt, D.P., Fair, J.M., *et al.* (2021) Updated distribution maps of predominant *Culex* mosquitoes across the Americas. *Parasites and Vectors*, 14, 547.

Graffelman, J. (2015). Exploring Diallelic Genetic Markers: The HardyWeinberg Package. Journal of Statistical Software, 64(3), 1-23. URL https://doi.org/10.18637/jss.v064.i03

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., *et al.* (2008) Global change and the ecology of cities. *Science*, 319, 756–760.

Guillot, G., Estoup, A., Mortier, F., Cosson, J.F. (2005) A spatial statistical model for landscape genetics. *Genetics*, 170, 1261–1280.

Haba, Y., McBride, L. (2022). Origin and status of *Culex pipiens* mosquito ecotypes. *Current Biology*, *32*(5), R237-R246.

Hamer, G.L., Anderson, T.K., Donovan, D.J., Brawn, J.D., Krebs, B.L., Gardner, A.M., *et al.* (2014) Dispersal of adult *Culex* mosquitoes in an urban West Nile virus hotspot: A mark-capture study incorporating stable isotope enrichment of natural larval habitats. *PLoS Neglected Tropical Diseases*, 8, e2768.

Hamer, G.L., Kitron, U.D., Brawn, J.D., Loss, S.R., Ruiz, M.O., Goldberg, T.L., *et al.* (2008) *Culex pipiens* (Diptera: Culicidae): A bridge vector of West Nile virus to humans. *Journal of Medical Entomology*, 45, 125–128.

Hamer, G.L., Kitron, U.D., Goldberg, T.L., Brawn, J.D., Loss, S.R., Ruiz, M.O., *et al.* (2009) Host Selection by *Culex pipiens* Mosquitoes and West Nile Virus Amplification. *American Journal of Tropical Medicine and Hygiene*, 80(2), 268–278.

Harbach, E., Harrison, B.A., Gad, A.M. (1984) *Culex (Culex) molestus* Forskal (Diptera: Culicidae): neotype designation, description, variation and taxonomic status. *Proceedings of the Entomological Society Washington*, 86, 521–542.

Harbison, J.E., Metzger, M.E., Hu, R. (2011) Seasonal oviposition of *Culex quinquefasciatus* in proprietary belowground stormwater treatment systems in an urban area of southern California. *Journal of Vector Ecology*, 36, 224–226.

Hardy, J. L., Houk, E. J., Kramer, L. D., Reeves, W. C. (1983). Intrinsic factors affecting vector competence of mosquitoes for arboviruses. *Annual Review of Entomology*, *28*(1), 229-262.

Harrington, L.C., Poulson, R.L. (2008) Considerations for accurate identification of adult *Culex restuans* (Diptera: Culicidae) in field studies. *Journal of Medical Entomology*, 45, 1–8.

Hewitt, G.M. (1990) Divergence and speciation as viewed from an insect hybrid zone. *Canadian Journal of Zoology*, 68, 1701–1715.

Hope AC. (1968) A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society Series B.* 30, 582–598.

Huang, S., Molaei, G., Andreadis, T.G. (2008) Genetic insights into the population structure of *Culex pipiens* (Diptera: Culicidae) in the northeastern United States by using microsatellite analysis. *American Journal of Tropical Medicine and Hygiene*, 79, 518–527.

Hufkens, K., Basler, D., Milliman, T., Melaas, E.K., Richardson, A.D. (2018) An integrated phenology modelling framework in R. *Methods in Ecology and Evolution*, 9(5), 1276-1285.

Hyndman, R.J., Athanasopoulos, G. (2018) Forecasting: principles and practice. *OTexts.*

Irwin P, Arcari C, Hausbeck J, Paskewitz S. (2008) Urban wet environment as mosquito habitat in the upper Midwest. *Ecohealth* 5, 49–57.

Jackson, B., Paulson, S.L. (2006) Seasonal abundance of *Culex restuans* and *Culex pipiens* in Southwestern Virginia through ovitrapping. *Journal of the American Mosquito Control Association*, 22, 206–212.

Jackson, B.T., Paulson, S.L., Youngman, R.R., Scheffel, S.L., Hawkins, B. (2005) Oviposition Preferences of *Culex restuans* and *Culex pipiens* (Diptera: Culicidae) for selected infusions in oviposition traps and gravid traps. *Journal of the American Mosquito Control Association*, 21, 360–365.

Johnson, B.J., Robson, M.G., Fonseca, D.M. (2015b) Unexpected spatiotemporal abundance of infected *Culex restuans* suggest a greater role as a West Nile virus vector for this native species. *Infection, Genetics and Evolution*, 31, 40–47.

Johnson, M.T.J., Munshi-South, J. (2017) Evolution of life in urban environments. *Science*, 358, 1–11.

Johnson, P.T.J., De Roode, J.C., Fenton, A. (2015a) Why infectious disease research needs community ecology. *Science*, 349.

Karesh, W.B., Dobson, A., Lloyd-Smith, J.O., Lubroth, J., Dixon, M.A., Bennett, M., *et al.* (2012) Ecology of zoonoses: Natural and unnatural histories. *The Lancet*, 380, 1936–1945.

Karki, S., Brown, W.M., Uelmen, J., Ruiz, M.O.H., Smith, R.L. (2020) The drivers of West Nile virus human illness in the Chicago, Illinois, USA area: Fine scale dynamic effects of weather, mosquito infection, social, and biological conditions. *PLoS ONE*, *15*(5), e0227160.

Kasai, S., Komagata, O., Tomita, T., Sawabe, K., Tsuda, Y., Kurahashi, H., *et al.* (2008) PCR-based identification of *Culex pipiens* complex collected in Japan. *Japanese Journal of Infectious Diseases*, 61, 184–191.

Kaur Gill H, Goyal G, Chahil G. (2017) Insect Diapause: A Review. *Journal of Agricultural Science and Technology,* 7, 454–473.

Kay B.H., Ryan P.A., Russell B.M., Holt J.S., Lyons S.A., Foley P.N. (2000) The importance of subterranean mosquito habitat to arbovirus vector control strategies in North Queensland, Australia. *Journal of Medical Entomology*, 37, 846–853.

Kent, R.J., Harrington, L.C., Morris, D.E. (2007) Genetic differences between *Culex pipiens* f. *molestus* and *Culex pipiens pipiens* (Diptera: Culicidae) in New York. *Journal of Medical Entomology*, 44(1), 50-59.

Kilpatrick, A.M. (2011) Globalization, land use, and the invasion of West Nile virus. *Science*, 334, 323–327.

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, 425–429.

Kilpatrick, A.M., Meola, M.A., Moudy, R.M., Kramer, L.D. (2008) Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathogens*, 4(6).

Kothera, L., Godsey, M.S. Jr., Doyle, M.S., Savage, H.M. (2012) Characterization of Culex pipiens Complex (Diptera: Culicidae) Populations in Colorado, USA Using Microsatellites. *PLoS ONE,* 7(10), e47602.

Kothera, L., Mutebi, J.P., Kenney, J.L., Saxton-Shaw, K., Ward, M.P., Savage, H.M., *et al.* (2020) Bloodmeal, host selection, and genetic admixture analyses of *Culex pipiens* complex (Diptera: Culicidae) mosquitoes in Chicago, IL. *Journal of Medical Entomology*, 57, 78–87.

Kothera, L., Zimmerman, E.M., Richards, C.M., Savage, H.M. (2009) Microsatellite characterization of subspecies and their hybrids in *Culex pipiens* complex (Diptera: Culicidae) mosquitoes along a north-south transect in the Central United States. *Journal of Medical Entomology*, 46, 236–248.

Kramer, L.D., Ciota, A.T. (2015) Dissecting vectorial capacity for mosquito-borne viruses. *Current Opinion in Virology*, 15, 112–118.

Kramer, L. D., Ciota, A. T., Kilpatrick, A. M. (2019). Introduction, spread, and establishment of West Nile virus in the Americas. *Journal of Medical Entomology*, 56(6), 1448-1455.

LaDeau, S.L., Allan, B.F., Leisnham, P.T., Levy, M.Z. (2015) The ecological foundations of transmission potential and vector-borne disease in urban landscapes. *Functional Ecology*, 29, 889–901.

Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW. (2011) Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proceedings of the National Academy of Science U.S.A.* 108, 7460–7465.

Lampman, R., Slamecka, M., Krasavin, N., Kunkel, K., Novak, R. (2006). *Culex* population dynamics and West Nile virus transmission in east-central Illinois. *Journal of the American Mosquito Control Association*, *22*(3), 390-400.

Laporta, G. Z., Sallum, M. A. M. (2014). Coexistence mechanisms at multiple scales in mosquito assemblages. *BMC Ecology*, *14*, 1-10.

Lebl, K., Brugger, K., Rubel, F. (2013) *Culex pipiens/restuans* population dynamics by interval lagged weather data. *Parasites and Vectors*, 6, 129.

Lee, J., Rowley, W. (2000) The abundance and seasonal distribution of *Culex* mosquitoes in Iowa during 1995-1997. *Journal of the American Mosquito Control Association*, 16, 275–278.

Leisnham, P.T., LaDeau, S.L., Juliano, S.A. (2014) Spatial and temporal habitat segregation of mosquitoes in urban Florida. *PloS ONE*, *9*(3), e91655.

Lima A., Lovin D.D., Hickner P.V., Severson D.W. (2016) Evidence for an overwintering population of *Aedes aegypti* in Capitol Hill neighborhood, Washington, DC. *American Journal of Tropical Medicine and Hygiene*, 94, 231–235.

Lenth R (2023). _emmeans: Estimated Marginal Means, aka Least-Squares Means_. R package version 1.8.4, <https://github.com/rvlenth/emmeans>.

Makowski, D., Ben-Shachar, M. S., Patil, I., Lüdecke, D. (2020). Estimation of Model-Based Predictions, Contrasts and Means. CRAN.

Mallis A. (2011) Handbook of Pest Control: The Behavior, Life History, and Control of Household Pests (Tenth Edition). Edited by Hedges, S. A. and D. Moreland (eds.). Mallis Handbook, LLC, Richfield, Ohio. 1599 pp.

Mattingly PF (1963) The urban mosquito hazard today. *Bulletin of the World Health Organization,* 29 Suppl: 135–139.

McDonnell, M.J., Hahs, A.K. (2008) The use of gradient analysis studies in advancing our understanding of the ecology of urbanizing landscapes: Current status and future directions. *Landscape Ecology*, 23, 1143–1155.

McDonnell, M.J., Pickett, Steward T.A Groffman, P., Bohlen, P., Pouyat, R. V, Zipperer, W.C., Parmelee, R.W., *et al.* (1997) Ecosystem processes along urban-to-rural gradient. *Urban Ecosystems*, 21–36.

McMillan, J.R., Armstrong, P.M., Andreadis, T.G. (2020) Patterns of mosquito and arbovirus community composition and ecological indexes of arboviral risk in the Northeast United States. *PLoS Neglected Tropical Diseases*, 14.

McPhearson, T., Cook, E. M., Berbés-Blázquez, M., Cheng, C., Grimm, N. B., Andersson, E., ... Troxler, T. G. (2022). A social-ecological-technological systems framework for urban ecosystem services. *One Earth*, 5(5), 505-518.

Medeiros-Sousa, A.R., Fernandes, A., Ceretti-Junior, W., Wilke, A.B.B., Marrelli, M.T. (2017) Mosquitoes in urban green spaces: Using an island biogeographic approach to identify drivers of species richness and composition. *Scientific Reports*, 7, 17826.

Morin C.W., Comrie A.C. (2010) Modeled response of the West Nile virus vector *Culex quinquefasciatus* to changing climate using the dynamic mosquito simulation model. *International Journal of Biometeorology,* 54, 517–529.

Mutebi, J.P., Savage, H.M. (2009) Discovery of *Culex pipiens pipiens* form *molestus* in Chicago. *Journal of the American Mosquito Control Association*, 25, 500–503.

Nasci R.S., Savage H.M., White D.J., Miller J.R., Cropp B.C., Godsey M.S., Kerst A.J., Bennett P., Gottfried K., Lanciotti R.S. (2001) West Nile virus in overwintering *Culex mosquitoes*, New York City, 2000. *Emerging Infectious Di*seases, 7(4), 742-4.

Norris, D.E. (2004) Mosquito-borne diseases as a consequence of land use change. *EcoHealth*, 1, 19–24.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. Wagner, H. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Osland M.J., Stevens P.W., Lamont M.M., Brusca R.C., Hart K.M., Waddle J.H., Langtimm C.A., Williams C.M., Keim B.D., Terando A.J., Reyier E.A., Marshall K.E., Loik M.E., Boucek R.E., Lewis A.B., Seminoff J.A. (2021) Tropicalization of temperate ecosystems in North America: The northward range expansion of tropical organisms in response to warming winter temperatures. *Global Change Biology,* 00, 1-26.

Palumbi, S.R. (2001) Humans as the world's greatest evolutionary force. *Science*, 293, 1786–1790.

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*, 10(1), 19287.

Pitcher, R. Lawton, P., Ellis, N., Smith, S.J., Incze, L.S., Wei, C.L., *et al*. (2012) Exploring the role of environmental variables in shaping patterns of seabed biodiversity composition in regional-scale ecosystems. *Journal of Applied Ecology*, 49, 670–679.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reinhold J.M., Lazzari C.R., Lahondère C. (2018) Effects of the environmental temperature on *Aedes aegypti* and *Aedes albopictus* mosquitoes: A review. *Insects.* 9.

Reisen, W. K., Wheeler, S. S. (2019). Overwintering of West Nile Virus in the United States. *Journal of Medical Entomology*, 56(6), 1498-1507.

Reiskind, M.H., Wilson, M.L. (2008) Interspecific competition between larval *Culex restuans* Theobald and *Culex pipiens* L. (Diptera: Culicidae) in Michigan. *Journal of Medical Entomology*, 45, 20–7.

Reiter, P. (1983) A portable battery-powered trap for collecting gravid *Culex* mosquitoes. *Mosquito News*, 43(4), 496-498.

Rinker, D.C., Pitts, R.J., Zwiebel, L.J. (2016) Disease vectors in the era of next generation sequencing. *Genome Biology*, 17, 1–11.

Rochlin, I., Santoriello, M.P., Mayer, R.T., Cambell, S.R. (2007) Improved high-throughput method for molecular identification of *Culex* mosquitoes. *Journal of the American Mosquito Control Association*, 23, 488–491.

Rozeboom L.E. (1960) *Aedes aegypti* (L.). The Yellow Fever Mosquito. Its life history, bionomics and structure. Rickard Christophers. *Quarterly Review of Biology,* 35, 237–237.

Rudbeck, L., Dissing, J. (1998) Rapid, simple alkaline extraction of human genomic DNA from whole blood, buccal epithelial cells, semen and forensic stains for PCR. *BioTechniques*, 25, 588–592.

Ruiz, M.O., Walker, E.D., Foster, E.S., Haramis, L.D., Kitron, U.D. (2007) Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *International Journal of Health Geographics*, *6*, 1-11.

Russell B.M., McBride W.J.H., Mullner H., Kay B.H. (2002) Epidemiological significance of subterranean *Aedes aegypti* (Diptera: Culicidae) breeding sites to dengue virus infection in Charters Towers, 1993. *Journal of Medical Entomology* 39, 143–145.

Samy A.M., Elaagip A.H., Kenawy M.A., Ayres C.F.J., Peterson A.T., Soliman D.E. (2016) Climate change influences on the global potential distribution of the mosquito *Culex quinquefasciatus*, vector of West Nile virus and lymphatic filariasis. *PLoS ONE*. 11.

Sanogo, Y.O., Kim, C.H., Lampman, R., Halvorsen, J.G., Gad, A.M., Novak, R.J. (2008) Identification of male specimens of the *Culex pipiens* complex (Diptera: Culicidae) in the hybrid zone using morphology and molecular techniques. *Journal of Medical Entomology*, 45, 203–209.

Shastry, V., Adams, P., Lindtke, D., Mandeville, E., Parchman, T., Gompert, Z., Buerkle, A. (2021) Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources*. 21, 1434-1451.

Smith, J.L., Fonseca, D.M. (2004) Rapid assays for identification of members of the *Culex* (*Culex*) *pipiens* complex, their hybrids, and other sibling species (Diptera: Culicidae). *American Journal of Tropical Medicine and Hygiene*, 70, 339–345.

Spielman A. (1964) Studies on autogeny in *Culex pipiens* populations in nature. I. Reproductive isolation between autogenous and anautogenous populations. *American Journal of Hygiene*, 80, 175–183.

Spielman A. (2001) Structure and seasonality of Nearctic *Culex pipiens p*opulations. *Annals of the New York Academy of Sciences,* 951, 220–234.

Strickman D., Fonseca D.M. (2012) Autogeny in *Culex pipiens* complex mosquitoes from the San Francisco Bay Area. *American Journal of Tropical Medicine and Hygiene*, 87(4), 719-26.

Takken, W., Verhulst, N.O. (2013) Host preferences of blood-feeding mosquitoes. *Annual Review of Entomology*, *58*, 433-453.

Thareja V., Basotra, R., Sareen, A. (2022) Concrete urban refugia for *Culex quinquefasciatus*, impact of summer stressors, and strategies adopted to oversummer. *Journal of American Mosquito Control Association*, 38(3), 165-174.

Tisdale, H. (1942) The process of urbanization. *Social Forces*, 20, 311–316.

Tokarz, R.E., Smith, R.C. (2019) Crossover Dynamics of *Culex* (Diptera: Culicidae) vector populations determine WNV transmission intensity. *Journal of Medical Entomology*, 1–8.

Townroe, S., Callaghan, A. (2014) British container breeding mosquitoes: The impact of urbanisation and climate change on community composition and phenology. *PLoS ONE*, 9.

Toxopeus J, Sinclair BJ. (2018) Mechanisms underlying insect freeze tolerance. *Biological Rev*iews, 93: 1891–1914.

Valentine, M.J., Ciraola, B., Jacobs, G.R., Arnot, C., Kelly, P.J. Murdock, C.C. (2020) Effects of seasonality and land use on the diversity, relative abundance, and distribution of mosquitoes on St. Kitts, West Indies. *Parasites and Vectors*, 13(1), 1-14.

Vinogradova, E. B., Shaikevich, E. V. (2007). Morphometric, physiological and molecular characteristics of underground populations of the urban mosquito *Culex pipiens* Linnaeus f. *molestus* Forskål (Diptera: Culicidae) from several areas of Russia. *European Mosquito Bulletin*, 22, 17-24.

Vinatier, F., Tixier, P., Duyck, P.F., Lescourret, F. (2011) Factors and mechanisms explaining spatial heterogeneity: A review of methods for insect populations. *Methods in Ecology and Evolution*, 2, 11–22.

Wagner, H.H. (2004) Direct multi-scale ordination with canonical correspondence analysis. *Ecology* 85, 342–351.

Weaver S, Reisen W. (2010) Present and future arboviral threats. *Antiviral Research,* 85(2), 328-245.

Williams, G.M., Gingrich, J.B. (2007) Comparison of light traps, gravid traps, and resting boxes for West Nile virus surveillance. *Journal of Vector Ecology*, 32, 285.

Wu T.P., Hu Q., Zhao T.Y., Tian J.H., De Xue R. (2014) Morphological studies on *Culex molestus* of the *Culex pipiens* complex (Diptera: Culicidae) in underground parking lots in Wuhan, Central China. *Florida Entomologist*, 97, 1191–1198.

Yee DA, Kesavaraju B, Juliano SA. (2004) Larval feeding behavior of three co-occurring species of container mosquitoes. *Journal of Vector Ecology*, 29, 315–22.

Yee, D.A., Bermond, C.D., Reyes-Torres, L.J., Fijman, N.S., Scavo, N.A., Nelsen, J., Yee, S. (2022) Robust network stability of mosquitoes and human pathogens of medical importance. *Parasites and Vectors,* 15, 216.

Yu, D., Madras, N., Zhu, H. (2018) Temperature-driven population abundance model for *Culex pipiens* and *Culex restuans* (Diptera: Culicidae). *Journal of Theoretical Biology*, *443*, 28-38.

Yurchenko, A.A., Masri, R.A., Khrabrova, N. V, Sibataev, A.K., Fritz, M.L., Sharakhova, M. V. (2020) Genomic differentiation and intercontinental population structure of mosquito vectors *Culex pipiens pipiens* and *Culex pipiens molestus*. *Scientific Reports*, 10.

Zeileis, A., Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Zinser, M., Ramberg, F., Willott, E. (2004) *Culex quinquefasciatus* (Diptera: Culicidae) as a potential West Nile virus vector in Tucson, Arizona: Blood meal analysis indicates feeding on both humans and birds. *Journal of Insect Science*, 4, 1–3.