THE EFFECTS OF INVASION DYNAMICS ON POST-ESTABLISHMENT EVOLUTION, GLOBAL SPREAD, AND POPULATION DETECTABILITY

by

RAFAEL VALENTIN

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate program in Ecology and Evolution

Written under the direction of

Dr. Julie L. Lockwood

Dr. Dina M. Fonseca

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2018

ABSTRACT OF THE DISSERTATION

The effects of invasion dynamics on post-establishment evolution, global spread, and population detectability

By RAFAEL VALENTIN

Dissertation Directors:

Dr. Julie L. Lockwood

Dr. Dina M. Fonseca

Research on non-native species has broad implications, ranging from understanding evolutionary adaptation of species in novel environments to ameliorating the negative ecological and economic impacts dangerous non-natives (i.e. invasive) promulgate. Non-native species are those that have been transported either intentionally or by accident by humans out of their native geographical range and released (introduced) into a new non-native location. Although they have long been a part of human society, non-native species numbers have grown exponentially over the last few decades with the rapid growth of global trade. Species are moved via trade as the commodity itself (e.g., ornamental plants or aquarium fish) or as stow-a-ways within the traded commodity or within the packing material and vessels used to transport commodities. Given the ubiquity of invasive species, there is an urgent need to understand the effects of invasion dynamics and build tools that can aid in our efforts to slow their spread or limit their impact.

My research contributes to this understanding at both biogeographical and local scales by addressing: 1) the influence of transport and release on post-establishment morphological evolution within a non-native bird population on Hawaii; 2) the global transportation and release pathways of a rapidly spreading invasive insect that harms agricultural interests worldwide; and 3) the implementation of novel molecular techniques to rapidly detect incipient invasive populations of agricultural pests when at low abundance. To achieve these goals, I use genetic information from the target non-native species. Genetic data is advantageous in this context as it avoids over-reliance on historical records or real-time surveillance to document transportation pathways, does not require *a priori* assumptions regarding the current distribution of the target species, and is naturally deposited and detectable within the environment for prolonged periods of time. My results inform efforts to manage the global spread of invasive species, and highlight the importance of introduction history on how non-native's evolutionarily respond to the conditions prevalent in their non-native location.

Acknowledgements

First and foremost, I would like to thank my advisors, Julie Lockwood and Dina Fonseca. I would not be where I am today without the both of them (and I really did need them both), and for that I am forever grateful. I know it was a very long road, and looking back I wonder how either of them had the patience to deal with me in those first few years. I was lost, without any sense of what I was doing, where I was going, what I did or didn't know, and what an actual question was. But worse still, I think, was my writing. How either of them read my earliest drafts and proposals and didn't give up on me then and there is unimaginable to me. But they both stuck with me, and I cannot put into words how indebted I am to them both. I'm sure this has been said by students many times in the past, but I really did have the best advisors anyone could have asked for.

Next, I would like to thank my committee members, both past and present, who provided me guidance throughout my time in graduate school. Specifically, I'd like to thank Peter Smouse for always lending his ear when I needed to talk about population genetics, and reciprocating with some of the best stories to match any discussion point. Ed Green, for teaching me the ways of Bayesian statistics and always having a comedic remark after asking "you're still here?" at Rutgers any time he saw me in passing. Nina Fefferman, for demonstrating time and again how to logically work through problems. She was probably the most intelligent person I had ever spoken with (so much so it was intimidating), and who I swear to this day must have lived multiple lives given her countless life experiences (seriously, ask her about anything and you will understand). Anne Nielsen and Tracy Leskey for teaching me how to work within agricultural systems, and for coming to my rescue when I was starting my field experiments and

needed a research site, supplies, and guidance. Malin Pinsky, for teaching me how to carry out Approximate Bayesian Computation analyses, which ended up being one of the more useful statistical analyses I had learned to date.

I'd be remiss if I did not take this time to thank Marsha Morin, who without I'm confident I would have been kicked out of Rutgers for filling some sort of paperwork incorrectly or missing some critical deadline.

Finally, I need to thank my family for supporting me through this and understanding what I was doing, without really understanding what I was doing. Most important of all, I must thank my wife, Jessica Valentin, for starting me on this path. If not for her, I would not have fought my initial rejection into Rutgers as an undergraduate, which ultimately set the wheels in motion for everything that has happened to me in my academic career to date. She has been my rock when I needed support, my greatest advocate when I swore I was worthless, and my best friend. I love you, always.

Table of Contents

List of Tables

Chapter 1

choice. Variations in scenarios refer to no enforced reductions in the exotic population's effective population size (i.e. no bottleneck – variations 1 and 4), or enforced reductions (i.e. bottleneck) followed by a change in effective population that was free to either increase or decrease (variations 2 and 3).

Table S1. PCA outputs for the combined native range and Hawaii (all islands in one group) dataset. The top table gives the eigenvalues and amount of variation explained for each PC. The second table gives the coordinate variables per physical northern cardinal attribute for each PC. The final table provides the PC loadings per physical attribute for each PC.

Table S2. MrBayes specific mutation models found to be best via the program PartitionFinder given the specific codon partitioning schemes. The model(s) for each partitioning scheme were selected and scored via AIC.

Chapter 2

Table 1. Illustration of the collections from (A.) native countries and (B.) non-native countries, as well as the specific localities within those countries, of BMSB with the number of specimens for each in parenthesis. Haplotypes shown are from the CO1 locus, with the number of specimens representing each haplotype in parenthesis as well. Locations marked with an asterisk indicate a lack of specific location data per haplotype. Haplotypes in bold indicate new haplotypes found in the US post-2008. We inserted

Haplotypes H24 through H29, when we realized that previous studies (Gariepy et al., 2015) jumped from H23 to H30.

Table 2. Probability and 95% credible interval for all Approximate Bayesian Computation scenarios used throughout the study, along with confidence in scenario choice. A downward arrow indicates a subsequent analysis with scenarios with probabilities below 0.1 excluded.

Table 3. The four principal questions being asked throughout this study. These questions shape the smaller questions that become the scenarios we model throughout the paper.

Table 4. Mitochondrial DNA loci used in population genetic studies of brown marmorated stink bug by country, with the relevant authors conducting them on the right. The asterisk designates loci that have five or fewer samples sequenced for that respective country.

Table 5. Prior distributions used for all ABC analyses. Mutation parameters refer to selected DNA mutation model, distributions used, and bounds for said distributions within the model validation screen.

Table S1. BMSB CO1 haplotype table sorted by country, with total number of samples from each country below the name, and accession numbers to the Genbank sequences

provided. Samples corresponding to specific haplotypes in bold represent new haplotypes found in the US post-2008 analysis.

Chapter 3

Table 1. Universal primer sets utilized throughout the project.

Table 2. List of twelve species from the United States tested from the family Pentatomidae, the same family as *Halyomorpha halys*.

Table 3. Summary of the samples used throughout the project. Sources of samples, and amount of material used by weight, are shown as pools if from a common homogenate, or as individual pellets if analyzed independently. Testing method indicates whether samples were subjected to cloning and sequencing (Cloned + Seq.) or directly sequenced without cloning (Dir. Seq.), tested with the real-time PCR assay (qPCR), or tested by amplifying with just the assay primers and direct sequencing (Assay Dir. Seq.). Number of species detected from sequencing is shown per sample. In sample 4, the cloned sequence matched equally a few closely related species, some already on the list, which is why the count is "6 or 7". "Y" indicates detection of BMSB. "N" indicates no detection. Samples 12 through 24 were not directly measured, thus their weight is approximated (~) based on the sampling method used for samples 4 and 8.

Supplemental table 1. List of the species sequences detected from guano amplifications using CO1 primers and cloning or direct sequencing. Species listings with more than one

species in a line, or a lack of a distinct species name, indicates inability to discern the correct species from the sequence provided using a 98% selection criteria. R1 refers to a ~650 bp amplicon, while R2 refers to a ~350bp amplicon. Percentages indicate the prevalence of each species among the clones successfully sequenced.

Chapter 4

WebTable 1. (a) Detection data for the four trees surveyed at the New Jersey site (RAREC). Columns are sorted by tree, while rows are arranged by survey day (first number) and survey method (D = eDNA, T = pheromone trap). (b) Detection data for the four fields in the New Hampshire farm. Columns are sorted by field, while rows indicate survey days (first number), and then the survey method (D = eDNA, T = pheromone traps) second. 1 represents positive detection, while 0 represents negative detection. Dots (•) represent days where a field was not surveyed with the eDNA method, due to lack of harvest. ‘Greens’ refer to the crop varieties chard, kale, and arugula.

List of Figures

Chapter 1

representing a clade-specific trait space. Large symbols within each oval depict the mean PC scores for each clade. We depict all individual cardinals captured and measured in Hawaii in light blue (with different shapes for each island) to visually show the distribution of their morphology (light blue oval) relative to native clades.

Figure 4. A phylogenetic breakdown of all native northern cardinal sequences analyzed by Smith *et al.* (2011), with the Hawaiian Islands samples we sequenced intermixed within the dataset. Monophyletic groups were categorized (and color-coded) to their respective mitochondrial DNA clade, while the branches representing the Hawaiian individuals were color coded in light blue. All Hawaiian samples grouped with the *C.c. cardinalis* native range clade.

Figure S1. Three scenarios describing the introduction of northern cardinals to the main Hawaiian Islands. Scenario I (two gray lines) = individual cardinals are collected from two clades, and the individuals of one clade are introduced to a subset of all islands whereas the individuals of the other clade are introduced to a separate set. Scenario II (black line) = all individuals collected from the native range come from a single clade and they are released across all islands. Scenario III (blue line) = individuals are collected from more than one native clade but all are introduced to every island, effectively creating an admixed population on Hawaii. We have depicted specific clades as sources, and islands as sets where introduction occurred, for illustration only. We have no reason to *a priori* restrict the set of clades, or the introduction scenarios, considered.

Figure S2. Haplotype network for *C.c.cardinalis* and Hawaiian cardinals. Network nodes in black indicate haplotypes found in *C.c.cardinalis,* while nodes in blue are those found in Hawaii. Numbers in bold (i.e. 22, 24, 47, 27, and 42) indicate haplotypes seen both in Hawaii and the *C.c.cardinalis* native range, while black nodes with a blue numbers (i.e. 22, 24, 47) indicate haplotypes that are seen across both regions of the *C.c.cardinalis* clade and in Hawaii. Numbers in red indicate haplotypes found in the western region of *C.c.cardinalis,* while numbers in black (not next to blue nodes) indicate haplotypes found in the eastern region of *C.c.cardinalis,* and numbers in purple (i.e. 31 and 36) indicate haplotypes seen in both regions.

Chapter 2

Figure 1. A CO1 haplotype network generated for BMSB, with geographic representation for each haplotype.

Figure 2. Map of the most likely BMSB invasion pathways connecting native and established populations across the globe, summarizing the results of our ABC analyses. Red dots on the map indicate the relevant native or established populations that are part of separate invasion pathways. The map's pathways are directional and labeled with the source population(s) listed, along with the calculated probabilities. The abbreviation NW refers to the northwestern population within the United States (US), CA refers to the US state of California, and the small arrow within the text provides additional clarity regarding the direction of the pathway. The base map, titled Blank Map Pacific World, was created by Dmthoth

(https://commons.wikimedia.org/wiki/File:Blank_Map_Pacific_World.svg) and altered to show the invasion pathways.

Figure 3: Approximate Bayesian Computation scenario outputs for the six significant scenarios with confidence values above 0.75: a. native source population for the pre-2008 US dataset; b. native source population for Canada; c. native source population for the Western US; d. introduction scenario for EU; e. native source population for Greece; f. bridgehead from US to Emilia-Romagna, Italy. E. US and W. US represent Eastern and Western United States, respectively. Can is an abbreviation for Canada. Em. R. is an abbreviation for Emilia-Romagna, Italy. All scenarios shown here are ordered based on experiment order in Table 2.

Figure S1. A visual representation of the scenarios tested for each of the questions asked while carrying out our ABC analyses. The populations in question are located at the center of each box, while the possible sources surround it. Each of the surrounding boxes represents a unique scenario that was modeled, with exception to distinct native populations (i.e. CN, JP, and KR) that also had every possible admixture option modeled as well, and the boxes connected by a dashed red arrow indicating another alternative scenario. The two letter abbreviations were used for each country, with the lowercase letters “s”, “w”, and “e” representing the cardinal directions of south, west, and east respectively. The abbreviation “Nat” indicates native range, meaning all scenarios for what the native source could be were tested prior to continuing with the displayed question.

Chapter 3

Figure 1. Polymorphic sites and BMITS1 primer/probe locations within the internal transcribed spacer 1. Position numbers refer to the basepair position along the *Halyomorpha halys* sequence (Accession # KU594609-KU594612), with position one being the start of ITS1 (after the stop codon at the end of the 18S exon). Below the line are the five BMSB ITS1 variants we sequenced. The colon symbol indicates a deletion. Compared with the other three variants, variants SB9 and SB6 have a large 83 bp deletion between positions 92 and 176 that we did not include in the figure (only the first and last bp are shown) to conserve space. SB9 also contains an ambiguous site in position 372, marked as "S", which corresponds to either a "C" or "G". qPCR primers and probe are represented by light and dark grey shading, respectively, with the associated basepairs bolded black and white, respectively. The thin vertical line indicates the start of the 5.8S exon. The * symbol indicates amplification with the ITS1A&B primer set, while the remaining variants resulted from amplifications with the BD1/4S primer set.

Chapter 4

Figure 1. Illustration of (a.) the brown marmorated stink bug (BMSB) and (b.) the damage it causes to fruits it has fed upon. The photo on the left (a.) was taken from StopBMSB.org, and the right (b.) taken by Brett Blaauw, Rutgers University.

Figure 2. Terrestrial surveillance efforts using eDNA require several steps (i) assessment of likely sources of target DNA (fruit, bark, soil, etc.), (ii) implementation of strategies to aggregate the target DNA (e.g. fruit washing), (iii) concentration of the target DNA (e.g.

by using filtration), (iv) sub-sampling of the concentrated sample (when needed), (v) DNA extraction (vi) testing for the presence or quantity of target DNA via qPCR, NextGen sequencing, or any other applicable method, (vii) data analysis (e.g. occupancy modeling, spatial analyses).

Figure 3. Map of the distribution of the brown marmorated stink bug within the continental US, with both study locations shown. RAREC stands for Rutgers Agricultural Research and Extension Center. Range map provided by StopBMSB.org.

Introduction

Species have been moved from one location to the next over the course of human history, either for food, personal desire and comfort, or inadvertently (Hodkinson and Thompson, 1997, Mack and Lonsdale, 2001). These species are thus considered non-native species, as they are species that have been transported from their native range to a novel environment where they share no evolutionary history. While the transportation of non-native species continues to this day, in recent decades the rate of inadvertent movement resulting in species invasions has increased rapidly on a global scale (Lockwood et al., 2013). This increase has inspired research to better understand the processes behind intentional and accidental invasions, as well as its effects on both the invaded natural systems and the introduced species itself. By studying these invasions researchers can gain insight to questions of the natural world, such as those related to community dynamics and evolution, which would otherwise be infeasible or unethical to carry out experimentally. This concept of using species invasions to address such natural questions is also known as taking advantage of accidental natural experiments (HilleRisLambers et al., 2013). While invasions can provide insight to such questions, often they are associated with negative effects that can be detrimental to the invaded system, thus requiring management.

Before non-native species establish in a novel habitat they must first go through different phases of the invasion process. Blackburn et al. (2011) defined these in four phases: transport, introduction, establishment, and spread; and the invasion process can fail at any of these phases. However, if the non-native species is successful in passing through all four phases, then it will begin to spread throughout the landscape. Invasion

spread (the last phase) is essentially repetition of the four step process as non-natives move throughout the novel environment (Pauchard and Shea, 2006, Blackburn et al., 2011), since all incipient populations must arrive to a new location either on their own power or through human transport and must successfully establish viable new populations. This process repeats as the species expands its non-native range throughout the landscape.

Once a non-native species has successfully transitioned across all four phases, surveillance and identification of its source range(s) are essential to answer specific questions. At large biogeographical scales, source identification determines the location(s) of the founding individuals, whether they be native or non-native sources. The latter case represents a bridgehead effect (Lombaert et al., 2010, Estoup and Guillemaud, 2010, Ascunce et al., 2011), where individuals from an established population are subsequently transported and introduced into another novel habitat. This is not to be confused with a dispersal event (either natural or human-mediated), which would spread the non-native throughout the landscape of an already invaded range. When executed correctly, the information acquired from source identification allows researchers and managers to identify the most likely invasion pathways, as well as evaluate whether there had been introductions from either a single source or multiple sources (Cristescu, 2015, Facon et al., 2006).

Source identification of non-natives also plays an important role in understanding their post-establishment evolution. The existence of post-establishment evolution in species invasions is well documented, with several published examples showing marked divergence in genes or traits between a non-native population and its putative native

source (Panarari-Antunes et al., 2012, Guggisberg et al., 2013, Kolbe et al., 2007), or between non-native populations residing within a single region (Mathys and Lockwood, 2011). Although there is an almost knee-jerk tendency to regard these differences as reflecting adaptation during or after initial establishment, such differences can emerge from several other processes (Keller and Taylor, 2008, Kolbe et al., 2007). For instance, founder effects such as whether multiple introductions events took place through space or time (i.e., multiple introductions all at one time, or at different times), whether they were sourced from multiple distinct native populations, and how many individuals were likely part of an introduction event all directly influence the genetic diversity of non-native species. Thus, they can produce stark genetic and trait differences between populations, and are important to consider when generating ecological or evolutionary hypotheses that involve such populations.

While source identification is important for assessing invasions at biogeographical scales, at more local scales the problem of how to rapidly identify an invasion and survey for incipient non-native populations is one of detection. Initial surveillance and rapid detection of nascent populations is a critical component in understanding and managing the spread of non-natives that may cause ecological or economic damage (becomes invasive) (Jerde et al., 2011). For instance, it is through surveillance that the presence of a newly established population is initially documented. Once the population has been identified in a specific locale, and if detected early enough where populations have not grown uncontrollably, management action can be taken to eradicate it or keep abundance low to minimize impact. Surveillance also allows for accurate documentation of the geographical pattern of spread (Jerde et al., 2011). From

this information researchers and managers can then deduce the pattern with which the invasive species is spreading across the landscape (e.g. directional or unidirectional spread), and how they are moving (i.e., dispersing under their own power or human mediated long-distance dispersal). Based on this information, policies can be implemented to stop invasion vectors from continuing the spread and allow for more focused control action.

Though the concept of surveillance to detect invasions and nascent populations early is promising, in practice it is problematic since initial population densities will likely be so low they will avoid physical detection. This is a result of the invasion process, as the barriers to establishment will likely reduce the number of founders to a subset of the introduced individuals. As a result, non-native populations are typically discovered only when population abundance levels have reached a point where they are clearly observable in the introduced range, indicating they have established and are likely spreading throughout the landscape (Crooks, 2005, Aikio et al., 2010). Fortunately, new molecular tools and techniques have become available to circumvent this problem in an effort to rapidly detect non-native populations.

My research highlights and contributes to understanding invasion dynamics at both biogeographic and local scales. The specific questions I ask address the influence of introduction history on post-establishment morphological evolution of a non-native bird, the global pathways of introduction for a rapidly spreading non-native species that harms agricultural interests worldwide, and the implementation of novel molecular techniques to rapidly detect incipient invasive populations of agricultural pests when at low abundance. The data I use primarily throughout this proposal is genetic. Genetic data is

advantageous in invasion ecology as it does not rely on historical records, and in some cases does not require any *a priori* assumptions regarding the current distribution of the non-native species in question. Below I briefly outline each of the four chapters within this dissertation.

Influence of invasion history on rapid morphological divergence across island populations of an exotic bird

For my first chapter we explore an example of post-establishment evolution of a non-native bird species, the northern cardinal (*Cardinalis cardinalis*), introduced to the Hawaiian Islands in the 1930s. From written historical records we know approximately 300-350 of these cardinals were shipped from the port of San Francisco (USA) and released onto Kauai, Oahu, and Hawaii Island between 1920 and 1931 (Pyle and Pyle, 2009). Since its introduction the northern cardinal has successfully established on these three islands, and has since spread and established to other islands within the archipelago. Most interestingly though is the documented evidence of morphological trait divergence among northern cardinals throughout the Hawaiian Islands, as noted by Mathys and Lockwood (2011).

Northern cardinals are native to North America, with populations spanning the eastern half of the continent through Texas and down into Central America, and consists of six mitochondrial DNA clades (Smith et al., 2011). The closest native population of northern cardinals to San Francisco is over 600km to the southeast, and there is no written record indicating whether they were acquired and shipped from this population, or another one located further away.

Here we addressed the role founder effects may have had in shaping this observed divergence among island populations (Keller and Taylor, 2008). We used a combination of morphological and genetic analyses to determine the likely native source population(s) and elucidate the invasion history of these cardinals. Our results show the cardinals were sourced from a single mitochondrial DNA clade, specifically the eastern half of the *C.c.cardinalis* clade consisting of states east of the Mississippi river. Thus, founder events were not likely an influential component in the observed morphological divergence among populations. Additionally, though circumstantial, our results suggest the Hawaiian cardinal's morphology appears to have diverged away from their native counterparts and are beginning to resemble more insular populations of northern cardinal.

This chapter was formatted for *Ecology and Evolution*, and was accepted for publication there (Valentin RE, Lockwood JL, Mathys BA, Fonseca DM – 2018)

Global invasion network of the brown marmorated stink bug, Halyomorpha halys

For my second chapter we were presented with a scenario where a non-native agricultural pest, the brown marmorated stink bug (BMSB; *Halyomorpha halys*), had become established in several countries around the world (Rice et al., 2014, Milonas and Partsinevelos, 2014, Vetek et al., 2014). To assess possible BMSB source populations, Xu *et al.* (2014) genetically analyzed US populations collected between 2001-2008, as well as specimens from several populations across China, Republic of Korea, and Japan. They analyzed sequence variation in portions of the mitochondrial DNA (mtDNA) cytochrome oxidase subunit II (CO2) and the control region (CR), and were able to

pinpoint the source of US populations to the northern region of China around Beijing (Xu et al., 2014).

Later, similar genetic analyses were carried out in Switzerland, Italy, France, Greece, and Hungary using mtDNA cytochrome oxidase subunit I (CO1) (Gariepy et al., 2014, Cesari et al., 2014, Gariepy et al., 2015), though with limited success. In these three studies the absence of BMSB specimens from putative source populations, and small sample sizes from many those they had, hampered the analyses. In an effort to remedy the situation and ascertain the global invasion pathways of BMSB, we combined the data from all prior studies(Xu et al., 2014, Cesari et al., 2014, Gariepy et al., 2014, Gariepy et al., 2015), as well as resampled across the US, and explored potential pathways using Approximate Bayesian Computation (ABC). We found China to be the most likely native source for all introductions around the globe. Additionally, we found China was also the source population for all introductions globally, with the exception of two bridgehead events from the eastern United States to Italy and Greece to Hungary.

This chapter was formatted for *Nature Scientific Reports*, and was accepted for publication there (Valentin RE, Nielsen AL, Wiman NG, Doo-Hyung L, Fonseca DM – 2017)

Real-time PCR assay to detect brown marmorated stink bug, Halyomorpha halys (Stål), in environmental DNA (eDNA)

For my third chapter we shift our focus to the detection of incipient non-native species populations. Specifically, we focus on the need to surveil for nascent BMSB populations

as it continues to spread throughout its introduced range in the United States. Current efforts aimed at detecting the presence of BMSB require direct observation or trapping with sweep-nets, black light traps, and pheromone traps followed by visual taxonomic identification (Khrimian et al., 2007). As is the case for many other invasive species, surveillance through direct sampling and observation can be costly due to the significant investment of labor, which may not even detect the presence of the target species until it is highly abundant (Rees et al., 2014, Dejean et al., 2012, Harvey et al., 2009).

An emerging surveillance tool that can prove highly useful in this regard is environmental DNA. Environmental DNA (eDNA) are DNA molecules released from various biological byproducts (e.g. skin cells, saliva, excreta, etc.) that have been deposited into the surrounding environment (Bohmann et al., 2014, Rees et al., 2014, Herder et al., 2014). This abundant source of DNA allows for indirect sampling that can then be used to identify the presence of one or more target species (Bohmann et al., 2014, Rees et al., 2014). However, before we could use eDNA to detect BMSB however we first needed a reliable and specific method for identifying trace amounts of often highly degraded DNA in terrestrial settings. Our primary objective was to develop such a tool.

We designed a set of primers (e.g. BMITS1F and BMITS1R) and a fluorescent TaqMan probe (BMITS1TM) from the internal transcribed spacer 1 (ITS1) for use with a real-time PCR (qPCR) system. The probe was designed to operate with the selected primer set, is just 4 base pairs from the BMITS1F primer, and contains a 5' VIC reporter dye and an MGB moiety attached to the 3' non-fluorescent quencher. The fragment amplified by the BMITS1 assay is just 96bp long. We found the BMITS1 assay's sensitivity was capable of detecting concentrations of BMSB genomic DNA as low as

20fg. We also evaluated its specificity against 12 species from the family Pentatomidae, the same family as BMSB. Further, we evaluated specificity by using guano from the big brown bat (*Eptesicus fuscus*), as a source from a generalist insectivore. These guano samples previously unknown to contain BMSB DNA, and thus were unknown whether they would indicate positive. We assessed samples that did present positive by sequencing them using only the assay's primers and examining whether there was sequence variation or indication of multiple sequences present at once. Of the samples we tested in this way we found no evidence of multiple sequences present, and no sequence variation was observed (i.e. all sequences matched the target sequence of BMSB). These results together indicate that the assay was in fact highly sensitive and species specific. Finally, we used the assay to identify potential sources of BMSB eDNA (e.g. exuviae and BMSB fecal material) and found both were viable sources.

> This chapter was formatted for *Pest Management Science*, and was accepted for publication there (Valentin RE, Maslo B, Lockwood JL, Pote J, Fonseca DM – 2016)

Early detection of terrestrial invasive insect infestations by using eDNA from crop surfaces

For my fourth and final chapter we used the newly designed BMSB assay to develop field protocols for surveillance of BMSB using eDNA. The purpose of this project was to provide a proof of concept and a protocol for using eDNA for early detection of BMSB across individual farms in the northeastern US. Early detection is a critical component of

best management practices since detecting and controlling the invader early before it can establish within a novel landscape, or for nascent populations, tend to be more successful and cost effective (Jerde et al., 2011, Harvey et al., 2009, Simberloff et al., 2013).

As previously mentioned, BMSB were typically detected by direct observation or trapping with sweep-nets, black light traps, and pheromone traps (Khrimian et al., 2007), thus requiring the capture of exemplary specimens. The use of eDNA would circumvent this requirement and allow for surveillance without any *a priori* assumptions about distribution or requiring the visualization of specimens (Thomsen et al., 2012a, Thomsen et al., 2012b, Foote et al., 2012, Dejean et al., 2012, Simberloff et al., 2013, Bohmann et al., 2014, Rees et al., 2014). To utilize eDNA surveillance we developed a framework to guide its use within terrestrial systems and adapted the sampling techniques used in aquatic surveillance methods to meet the requirements of this framework. We tested the use of BMSB surveillance using eDNA in a known high abundance population and an unknown population that would be at low abundance (if present), and compared its effectiveness against blacklight and pheromone traps in both sites. We found the high abundance site had a positive detection for all methods, while the unknown (low abundance site) only indicated positive for BMSB using the eDNA technique, except for the last day of sampling where a single individual was found in a pheromone trap. Our results indicate eDNA surveillance can indeed be utilized in terrestrial systems, and has a higher probability of detecting nascent populations compared to conventional practices.

This chapter was formatted for *Frontiers in Ecology and the Environment*, and was accepted for publication there (Valentin RE, Fonseca DM, Nielsen AL, Leskey TC, Lockwood JL – 2018)

Literature Cited

AIKIO, S., DUNCAN, R. P. & HULME, P. E. 2010. Lag‐phases in alien plant invasions: Separating the facts from the artefacts. *Oikos,* 119**,** 370-378.

ASCUNCE, M. S., YANG, C.-C., OAKEY, J., CALCATERRA, L., WU, W.-J., SHIH, C.-J., GOUDET, J., ROSS, K. G. & SHOEMAKER, D. 2011. Global invasion history of the fire ant *Solenopsis invicta*. *Science,* 331**,** 1066-1068.

BLACKBURN, T. M., PYŠEK, P., BACHER, S., CARLTON, J. T., DUNCAN, R. P., JAROŠÍK, V., WILSON, J. R. & RICHARDSON, D. M. 2011. A proposed unified framework for biological invasions. *Trends in Ecology & Evolution,* 26**,** 333-339.

BOHMANN, K., EVANS, A., GILBERT, M. T. P., CARVALHO, G. R., CREER, S., KNAPP, M., YU, D. W. & DE BRUYN, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution,* 29**,** 358-367.

CESARI, M., MAISTRELLO, L., GANZERLI, F., DIOLI, P., REBECCHI, L. & GUIDETTI, R. 2014. A pest alien invasion in progress: potential pathways of origin of the brown marmorated stink bug Halyomorpha halys populations in Italy. *Journal of Pest Science***,** 1-7.

CRISTESCU, M. E. 2015. Genetic reconstructions of invasion history. *Molecular ecology,* 24**,** 2212-2225.

CROOKS, J. A. 2005. Lag times and exotic species: the ecology and management of biological invasions in slow-motion. *Ecoscience,* 12**,** 316-329.

DEJEAN, T., VALENTINI, A., MIQUEL, C., TABERLET, P., BELLEMAIN, E. & MIAUD, C. 2012. Improved detection of an alien invasive species through

environmental DNA barcoding: the example of the American bullfrog Lithobates catesbeianus. *Journal of applied ecology,* 49**,** 953-959.

ESTOUP, A. & GUILLEMAUD, T. 2010. Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology,* 19**,** 4113-4130.

FACON, B., GENTON, B. J., SHYKOFF, J., JARNE, P., ESTOUP, A. & DAVID, P. 2006. A general eco-evolutionary framework for understanding bioinvasions. *Trends in ecology & evolution,* 21**,** 130-135.

FOOTE, A. D., THOMSEN, P. F., SVEEGAARD, S., WAHLBERG, M., KIELGAST, J., KYHN, L. A., SALLING, A. B., GALATIUS, A., ORLANDO, L. & GILBERT, M. T. P. 2012. Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PloS one,* 7**,** e41781.

GARIEPY, T., BRUIN, A., HAYE, T., MILONAS, P. & VÉTEK, G. 2015. Occurrence and genetic diversity of new populations of Halyomorpha halys in Europe. *Journal of Pest Science,* 1-10.

GARIEPY, T., FRASER, H. & SCOTT-DUPREE, C. 2014. Brown marmorated stink bug (Hemiptera: Pentatomidae) in Canada: recent establishment, occurrence, and pest status in Southern Ontario. *The Canadian Entomologist,* 1-4.

GUGGISBERG, A., LAI, Z., HUANG, J. & RIESEBERG, L. H. 2013. Transcriptome divergence between introduced and native populations of Canada thistle, Cirsium arvense. *New Phytologist,* 199**,** 595-608.

HARVEY, C. T., QURESHI, S. A. & MACISAAC, H. J. 2009. Detection of a colonizing, aquatic, non‐indigenous species. *Diversity and Distributions,* 15**,** 429-437.

HERDER, J. E., VALENTINI, A., BELLEMAIN, E., DEJEAN, T., DELFT, J. J. C. W. V., THOMSEN, P. F. & TABERLET, P. 2014. Environmental DNA - a review of the possible applications for the detection of (invasive) species. *In:* STICHTING RAVON, N. (ed.).

HILLERISLAMBERS, J., ETTINGER, A. K., FORD, K. R., HAAK, D. C., HORWITH, M., MINER, B. E., ROGERS, H. S., SHELDON, K. S., TEWKSBURY, J. J. & WATERS, S. M. 2013. Accidental experiments: ecological and evolutionary insights and opportunities derived from global change. *Oikos,* 122**,** 1649-1661.

HODKINSON, D. J. & THOMPSON, K. 1997. Plant dispersal: the role of man. *Journal of Applied Ecology***,** 1484-1496.

JERDE, C. L., MAHON, A. R., CHADDERTON, W. L. & LODGE, D. M. 2011. "Sight‐unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters,* 4**,** 150-157.

KELLER, S. R. & TAYLOR, D. R. 2008. History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters,* 11**,** 852-866.

KHRIMIAN, A., SHEARER, P. W., ZHANG, A., HAMILTON, G. C. & ALDRICH, J. R. 2007. Field trapping of the invasive brown marmorated stink bug, Halyomorpha halys, with geometric isomers of methyl 2, 4, 6-decatrienoate. *Journal of agricultural and food chemistry,* 56**,** 197-203.

KOLBE, J. J., LARSON, A. & LOSOS, J. B. 2007. Differential admixture shapes morphological variation among invasive populations of the lizard Anolis sagrei. *Molecular Ecology,* 16**,** 1579-1591.

LOCKWOOD, J. L., HOOPES, M. F. & MARCHETTI, M. P. 2013. *Invasion ecology*, John Wiley & Sons.

LOMBAERT, E., GUILLEMAUD, T., CORNUET, J.-M., MALAUSA, T., FACON, B. & ESTOUP, A. 2010. Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *PLoS ONE,* 5**,** e9743.

MACK, R. N. & LONSDALE, W. M. 2001. Humans as Global Plant Dispersers: Getting More Than We Bargained For: Current introductions of species for aesthetic purposes present the largest single challenge for predicting which plant immigrants will become future pests. *BioScience,* 51**,** 95-102.

MATHYS, B. A. & LOCKWOOD, J. L. 2011. Contemporary morphological diversification of passerine birds introduced to the Hawaiian archipelago. *Proceedings of the Royal Society of London B: Biological Sciences***,** rspb20102302.

MILONAS, P. & PARTSINEVELOS, G. 2014. First report of brown marmorated stink bug *Halyomorpha halys* Stål (Hemiptera: Pentatomidae) in Greece. *EPPO Bulletin,* 44**,** 183-186.

PANARARI‐ANTUNES, R., PRIOLI, A., PRIOLI, S., GOMES, V., JÚLIO, H., AGOSTINHO, C., SILVA FILHO, J., BONI, T. & PRIOLI, L. 2012. Genetic divergence among invasive and native populations of Plagioscion squamosissimus (Perciformes, Sciaenidae) in Neotropical regions. *Journal of fish biology,* 80**,** 2434-2447.

PAUCHARD, A. & SHEA, K. 2006. Integrating the study of non-native plant invasions across spatial scales. *Biological Invasions,* 8**,** 399-413.

PYLE, R. & PYLE, P. 2009. The birds of the Hawaiian islands: occurrence, history, distribution, and status. *BP Bishop*.

REES, H. C., MADDISON, B. C., MIDDLEDITCH, D. J., PATMORE, J. R. & GOUGH, K. C. 2014. REVIEW: The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology,* 51**,** 1450-1459.

RICE, K. B., BERGH, C. J., BERGMANN, E. J., BIDDINGER, D. J., DIECKHOFF, C., DIVELY, G., FRASER, H., GARIEPY, T., HAMILTON, G. & HAYE, T. 2014. Biology, Ecology, and Management of Brown Marmorated Stink Bug (Hemiptera: Pentatomidae). *Journal of Integrated Pest Management,* 5**,** A1-A13.

SIMBERLOFF, D., MARTIN, J.-L., GENOVESI, P., MARIS, V., WARDLE, D. A., ARONSON, J., COURCHAMP, F., GALIL, B., GARCÍA-BERTHOU, E. & PASCAL, M. 2013. Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution,* 28**,** 58-66.

SMITH, B. T., ESCALANTE, P., BAÑOS, B. E. H., NAVARRO-SIGÜENZA, A. G., ROHWER, S. & KLICKA, J. 2011. The role of historical and contemporary processes on phylogeographic structure and genetic diversity in the Northern Cardinal, Cardinalis cardinalis. *BMC evolutionary biology,* 11**,** 136.

THOMSEN, P., KIELGAST, J., IVERSEN, L. L., WIUF, C., RASMUSSEN, M., GILBERT, M. T. P., ORLANDO, L. & WILLERSLEV, E. 2012a. Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology,* 21**,** 2565-2573.

THOMSEN, P. F., KIELGAST, J., IVERSEN, L. L., MØLLER, P. R., RASMUSSEN, M. & WILLERSLEV, E. 2012b. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS one,* 7**,** e41732.

VETEK, G., PAPP, V., HALTRICH, A. & REDEI, D. 2014. First record of the brown marmorated stink bug, Halyomorpha halys (Hemiptera: Heteroptera: Pentatomidae), in Hungary, with description of the genitalia of both sexes. *Zootaxa*, 3780**,** 194–200.

XU, J., FONSECA, D. M., HAMILTON, G. C., HOELMER, K. A. & NIELSEN, A. L. 2014. Tracing the origin of US brown marmorated stink bugs, Halyomorpha halys. *Biological invasions*, 16**,** 153-166.

Influence of invasion history on rapid morphological divergence across island populations of an exotic bird

Rafael E. Valentin[1,4], Julie L. Lockwood[1], Blake A. Mathys[2] and Dina M. Fonseca[1,3]

[1]Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ USA
[2]Division of Mathematics, Computer and Natural Sciences, Ohio Dominican University, Columbus, OH, USA
[3]Department of Entomology, Rutgers University, New Brunswick, NJ USA

[4]**Corresponding author**: Rafael E. Valentin, Ecology and Evolution Graduate Program, 14 College Farm Rd., New Brunswick, NJ 08901, USA;

Rafael.Valentin@Rutgers.edu

Acknowledgements

Funding was provided by Ohio Dominican University Summer Research Grant, Fonseca lab startup funds from Rutgers University, and National Geographic Society Committee for Research and Exploration (grant number 8261-07). Thanks to Carla Kishinami (Bishop Museum, Hawaii), James Dean and Robert Fleischer (Smithsonian National Museum of Natural History, US), Marc Adams (Natural History Museum, Tring, U.K.), Van Remsen (Louisiana State University), Jason Omick (State of Hawaii Division of Forestry and Wildlife, Department of Land and Natural Resources), Ariane Le Gros, Eric VanderWerf, Lindsay C. Young, Kim & Forest Starr, Kurt Pohlman, Jim Denny, Jean

Olbert, Monte Tudor-Long, Pete Donaldson, Joe Kern, Dan Lindsay, Randi Rhodes, Abby Brown-Watson, Jack Jeffrey, Andre Raine, Caleb Slemmons, and Judy Bird for their assistance.

Abstract

Aim

There is increasing evidence that exotic populations may rapidly differentiate from those in their native range, and that differences also arise among populations within the exotic range. Using morphological and DNA-based analyses we document the extent of trait divergence among native North American and exotic Hawaiian populations of northern cardinal (*Cardinalis cardinalis*). Furthermore, using a combination of historical records and DNA-based analyses we evaluate the role of founder effects in producing observed trait differences.

Location

North America and the Hawaiian archipelago

Methods

We measured and compared key morphological traits across northern cardinal populations in the native and exotic ranges to assess whether trait divergence across the Hawaiian Islands, where this species was introduced between 1929-1931, reflected observed variation across native phylogeographic clades in its native North America. We used and added to prior phylogenetic analyses based on a mitochondrial locus to identify the most likely native source clade(s) for the Hawaiian cardinal populations. We then used Approximate Bayesian Computation (ABC) to evaluate the role of founder effects in producing the observed differences in body size and bill morphology across native and exotic populations.

Results and Main Conclusions

Phylogeographic analysis identified the eastern North American clade (*C. cardinalis cardinalis*) as the most likely and sole native source for all the Hawaiian cardinal populations. Cardinal populations on the Hawaiian Islands had morphological traits that diverged substantially across islands and overlapped the trait space of all measured native North American clades. The ABC analyses supported written accounts of the cardinal's introduction that indicate the original 300 cardinals shipped to Hawaii were simultaneously and evenly released across Hawaii, Kauai, and Oahu. Populations on each island likely experienced bottlenecks followed by expansion, with cardinals from the island of Hawaii eventually colonizing Maui unaided. Overall, our results suggest that founder effects had limited impact on morphological trait divergence of exotic cardinal populations in the Hawaiian archipelago, which instead reflect post-introduction events.

Keywords: Founder effects, *Cardinalis cardinalis*, Hawaii, morphology, Approximate Bayesian Computation, exotic species, evolution, invasion history

Introduction

The recognition that biological invasions provide unique insight into the mechanics of evolutionary divergence has led to a spike in published research on post-establishment evolution of exotic species (e.g. Dlugosch & Parker, 2008a; Dlugosch & Parker, 2008b; Suarez & Tsutsui, 2008). There are now several examples of marked divergence in genetic or phenotypic traits between two or more exotic populations (Freed *et al.*, 1987; Phillimore *et al.*, 2008; Xu *et al.*, 2010; Westley *et al.*, 2012; Lucek *et al.*, 2014; Egizi *et al.*, 2015). Such differences can be explained by *in situ* adaptation to local biological and environmental conditions, or from events that occurred within the species' invasion history (e.g. Allendorf & Lundquist, 2003; Dlugosch & Parker, 2008a; Keller & Taylor, 2008). In particular, founder effects can result in divergence of traits across exotic populations if colonizing individuals are derived from two or more genetically and/or phenotypically structured native sub-populations and introduced in such a way where these features are structured across the exotic range (Keller & Taylor, 2008). Here we deduce the invasion history of northern cardinals (*Cardinalis cardinalis*) established across the main Hawaiian Islands, and using this history, evaluate the role of founder effects in producing previously observed morphological divergence of these populations (Mathys & Lockwood, 2011). In the process we also elucidate the degree to which cardinals on Hawaii have diverged from their native source population(s), and provide insight into their post-establishment population dynamics.

From written records we know that between 1929 and 1931, 300–350 northern cardinals were purposefully transported and released onto Hawaii (Pyle & Pyle, 2009). These cardinals were shipped from the port of San Francisco (USA) and released onto

Kauai, Oahu, and Hawaii Island (Pyle & Pyle, 2009). Northern cardinals are native to North America, with populations spanning the eastern half of the continent through to New Mexico and down into Mexico (**Figure 1**). There are six mitochondrial clades present in North America (**Figure 1**), with considerable morphological differences between them (Smith *et al.*, 2011). The closest native population of northern cardinals to San Francisco is over 600km to the south representing the *C.c. igneous* clade. There are no written records telling us whether the cardinals shipped from San Francisco came from this clade, or another one located further away but perhaps more connected to the city via train or other transportation mechanisms typical of this era. Thus, we do not know whether the cardinals on Hawaii were derived from one or more source clades; and, if more than one clade was involved, if a single or multiple clades founded the exotic populations on each island. The records also do not tell us how the 300–350 individual cardinals were divided across release events, or how (or if) they were divided between shipments across years.

What we do know is this was the only introduction of northern cardinals to the archipelago, and they rapidly increased in population size after establishment, eventually colonizing all of the main Hawaiian Islands by the 1950s. We also know that current island cardinal populations are statistically different from each other in several morphological traits (e.g., wing and bill sizes - Mathys & Lockwood, 2011). These morphological traits are known to be heritable among birds (Badyaev & Martin, 2000a; Badyaev & Martin, 2000b; Jensen *et al.*, 2003), and Mathys and Lockwood (2011) show that observed across-island differences are of such a magnitude that *in situ* genetic drift is not a likely causal mechanism (Mathys & Lockwood, 2011).

There are three ways invasion history could have produced the morphological divergence seen in cardinal populations across Hawaii. First, populations on each island may have been founded by individuals from genetically and morphologically distinct native source clades, and the morphological differences observed today recapitulate these across-clade differences (**Figure S1-Scenario1**). Second, one or more island populations may represent an admixture of individuals sourced from different native cardinal clades (**Figure S1-Scenario 3**). Any observed differences across islands today thus evolved in response to island-specific selective forces enabled by the increases in genetic diversity that accompany admixture. Third, the cardinals on Hawaii may have been derived from one native source clade, which would suggest that current morphological differences arose after establishment from the existing genetic variation found within these founders (**Figure S1-Scenario 2**).

We examine these possibilities by updating and expanding the between-island morphological trait analysis from Mathys and Lockwood (2011). We then, for the first time, compare the distribution of traits across the Hawaiian Islands to traits typical of cardinal clades in the native range. Finally, we determine the most likely native source population(s) for the exotic island populations, and deduce their post-establishment population dynamics using phylogeographic and Approximate Bayesian Computation (ABC) analyses. By combining these analyses, we assess which of the above three introduction and differentiation scenarios most likely occurred among northern cardinals in Hawaii, and we shed light on the post-establishment evolutionary dynamics of this species.

Methods

Morphological Analysis

In this analysis we sought to establish the magnitude and direction of morphological differences in cardinals between the five main Hawaiian Islands, between the cardinals associated with each native range clade, and between the native clades and Hawaii. We used the following morphological traits: tail length, wing chord, culmen length, bill depth (at anterior margin of nares), and bill width (also at anterior margin of nares); all measured in millimetres. These traits are commonly used metrics for evaluating evolutionary divergence between bird populations due to their known associations with life history and foraging adaptations (Ricklefs & Travis, 1980; Lockwood *et al.*, 1993).

We visited Kauai, Oahu, Maui and Hawaii Island in the summer of 2008, and again in the summer of 2013, to obtain morphological measurements of 74 live-caught northern cardinals. Mist nets were placed in areas that experience regular bird activity. No lures or baits were used in order to prevent bias in the sex ratio of captured individuals. Captured individuals were fitted with USGS numbered bands before release, allowing us to avoid measuring the traits of any one individual multiple times. All morphological measurements on field-captured individuals were taken in the same season, thus avoiding systematic bias in morphological traits that vary with season (e.g., wing chord - Arendt & Faaborg, 1989). Only adults were measured, as young individuals are still growing and do not provide accurate measures of adult body dimensions. Culmen length, bill depth, bill width, and tarsus length were measured with a Mitutoyo dial calliper (Mitutoyo America Corporation, Aurora, IL, USA) to one-hundredth of a

millimetre precision. Tail length and wing chord were measured with a 15 cm wing rule accurate to one millimetre (Avinet, Inc., Dryden, NY, USA).

In addition to live individuals, we measured specimens housed in the Bishop Museum (Hawaii, USA), American Museum of Natural History (New York, NY, USA), and the National Museum of Natural History (Washington, D.C., USA). In total we measured 130 specimens collected across four of the six native range clades; and 106 specimens collected on Kauai, Oahu, Maui and Hawaii Island. All museum specimens of Hawaiian cardinals measured were collected between 1980 and 1999, and included both males and females. We did not have enough specimens measured from the *C.c. carneus* native clade, and none were available for the *C.c. mariae* clade, to include them in the analysis. We purposefully selected individual specimens that came from across the full geographic expanse of each of the four remaining native clades. Thus, for example in the case of the very wide-spread *C.c. cardinalis* clade (**Figure 1**), we measured individuals from Virginia, South Carolina, North Carolina, Montana, Missouri, Maryland Washington D.C., Ohio, Florida, Georgia, Texas, Michigan, New York, Kansas, and Mexico. This effort allowed us to capture a representative portion of the morphological trait variation within each native clade. All measurements were taken in the same way as for live specimens.

Data from live-captured individuals and museum specimens were combined for all morphological analyses. We measured only museum specimens that were captured at the same time of year as the live-caught individuals to reduce any systematic bias between the two data sources, and combined measurements for males and females to maximize sample sizes. Northern cardinals show very little differences between sexes in

the traits we measured, however to ensure that across-population comparisons were not biased by sex-specific differences we kept sex ratios across clades and islands as close to 50:50 as possible.

Finally, it is well documented that bird specimens experience changes in some mensural characters (e.g., wing chord) after museum preparation (Haftorn, 1982; Bjordal, 1983; Winker, 1993) due to drying of the skin. In order to combine the measurements from live individuals with museum specimens, we multiplied field (live-caught) measurements of tail length and wing chord by taxon and character-specific correction factors following Winker (1993) and Mathys and Lockwood (2011). In order to correct for individuals with missing measurements due to condition of the specimen or inability to take all measurements in the field, we approximated missing trait values using the data imputation MICE package in R (Buuren & Groothuis-Oudshoorn, 2011). This method was preferred as it has little impact on the observed population mean, uses the dataset itself to generate imputed data values, and does not reduce the variation in the dataset. We imputed trait information for less than 2% of the full dataset.

Recognizing that morphological traits are often inter-correlated we collapsed the six measured traits from live-caught and museum cardinals into two principal components using Principal Component Analysis (PCA; Lockwood *et al.* 1993) in R statistical software with the factoextra package (Team, 2014; Kassambara & Mundt, 2016). Prior to conducting the PCA, we log transformed all variables and then centred and scaled the means. The first two dimensions of the PCA (PC1 and PC2) explained 75% of the observed variation in morphological traits, with PC1 capturing overall size of individuals and PC2 reflecting the ratio of the bill to body size (**Table S1**). We retained

the PC1 and PC2 scores for each measured individual so we could compare morphological differences across populations and clades.

We initially updated and expanded the between-island morphological analysis of Mathys and Lockwood (2011) by increasing the number of individuals measured across islands, and adding specimens from Maui to the comparisons. Using individual PC1 and PC2 scores, we evaluated differences in cardinal morphology between islands using multivariate analysis of variance (MANOVA) in R. If the overall MANOVA resulted in statistical significance, we followed that test with a series of pair-wise MANOVA tests between islands.

Next we compared the morphologies of Hawaiian cardinals to the four native cardinal clades for which we had sufficient data. To aid in visualizing quantitative differences in morphology across island populations and native clades, we plotted PC1 and PC2 for all measured cardinals in 2-dimensional space. We visually identified individuals from each native-range clade using color-coding, adding an ellipse that contained 95% of all individuals from these clades to clearly identify the range of morphologies present within each. We designated cardinals from Hawaii with a unique color code as well as designated individuals according to their island of residence using island-specific symbols. This graph allows one to visualize the morphological ‘map’ of native range cardinals, where each clade occupies a relatively distinct position in the 2-dimensional space, and then visually evaluate where the Hawaiian cardinals ‘fit’ onto this map.

Using these data we evaluated the following scenarios: (1) the Hawaiian cardinals fall entirely into the trait space of only one native range clade, indicating all Hawaiian

cardinals were derived from this single native source and any divergence they show across islands is typical of the range of morphologies seen in that clade; (2) Hawaiian cardinals span two (or more) native range clade spaces, and that cardinals from one island clearly fall within one clade and cardinals from another island fall in the other clade, indicating that inter-clade morphological differences are being recapitulated across islands (founder effect); or (3) the Hawaiian cardinals do not neatly fit into any single native clade's morphological space, indicating potential admixture at the time of introduction, post-establishment divergence, or both. We quantitatively evaluated differences in PC1 and PC2 between clades and Hawaii with MANOVA followed by pairwise MANOVA.

Sequence data generation

In order to determine the native source(s) of cardinal populations across Hawaii, we combined the *C. cardinalis* native range genetic data from Smith *et al.* (2011) and genetic information from the live-caught individuals to create a merged northern cardinal dataset. Smith *et al.* (2011) used the sodium dehydrogenase subunit-2 (ND2) mitochondrial locus to establish discrete genetic boundaries for six native range clades. In order to compare Hawaii cardinals with this dataset, we used the same locus. We collected feathers from 46 of the measured individuals caught in the Hawaiian Islands in 2013. Feathers were placed in small envelopes, and upon return to the lab, the calamus from multiple feathers were clipped to obtain a biological sample for each individual. These samples were placed in 1.5ml Eppendorf tubes in order to extract genomic material from cells found on the feather calamus. We extracted DNA using a DNeasy blood and tissue kit under standard

protocols (Qiagen reference), with Proteinase K incubation taking place overnight (minimum of 8 hours) to ensure complete digestion.

We amplified 1,042 base pairs of ND2 via polymerase chain reaction (PCR) using primers L5215 (5′- TATCGGGCCCATACCCCGAAAAT-3′) and HTrpC (5'-CGGACTTTACGACAAACTAAGAG-3'), identical to those used by Smith *et al.* (2011). Amplification was accomplished with 20 µl reactions consisting of 1× PCR buffer (10 mM Tris-HCl, pH 8.3, and 50 mM KCl), 2.25mM $MgCL_2$, 150 µM each dNTP, 200nM of each primer, 1 unit of Amplitaq Gold DNA Polymerase and 3µl of genomic DNA. The protocol was optimized to run at an initial denaturing temperature of 96°C for 10 minutes, followed by 40 cycles of the following steps: denaturing at 96°C for 45 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 45 seconds. Final extension was completed at 72°C for 5 minutes. All PCRs were run on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA). We visualized reactions in a 1% agarose gel with Ethidium Bromide, and selected DNA fragments of appropriate size for sequencing. Successful amplicons were cleaned using ExoSAP-IT (Affymetrix, OH), and mixes of 25pmoles of primer and 40ng of template DNA were sent for cycle sequencing and sizing (Genscript, Piscataway, NJ, USA). Sequences were obtained using both primers to create a consensus of the full 1,042bp ND2 sequence after chromatograms were cleaned and aligned in Sequencer 5.1 (GeneCodes, Ann Arbor, MI). All sequences were evaluated for insertions and deletions, as well as translated to amino acids to check for stop codons and the presence of nuclear copies (Sorenson & Fleischer, 1996).

Phylogeographic analysis

We executed a phylogeographic analysis using the merged northern cardinal dataset to determine which native source clades were associated with each exotic island population. We ran the dataset through the program PartitionFinder 1.1.1 (Lanfear *et al.*, 2012) in Python v2.7 under two model schemes: unpartitioned whole gene ND2 sequences, and partitioned by codon. We implemented a MrBayes model filter to select only the twenty-four DNA evolutionary models that were compatible with the MrBayes program. PartitionFinder generated model schemes for both partitioned and unpartitioned data, and ranked them by Akaike Information Criteria (AIC). We then constructed a phylogenetic tree with MrBayes v3.2.2 (Ronquist & Huelsenbeck, 2003) under the selected best scheme for both unpartitioned and partitioned data. The program was allowed to run for 10 million generations, while being sampled every 1000, with a relative burn-in of 0.25. We visually inspected MCMC chains using the program Tracer v1.6 (Rambaut *et al.*, 2014) to confirm adequate burn-in and convergence of chains, and used FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree) for final tree assembly and inspection.

Approximate Bayesian Computation analysis

We used Approximate Bayesian Computation (ABC) to test a suite of possible introduction and range expansion scenarios. Briefly, ABC is a Bayesian analysis that allows for direct comparison of multiple introduction hypotheses (known as scenarios) and provides relative probabilities for each, given the data provided. This comparison is accomplished by performing inference computations from simulated pseudo-observed datasets (PODs) that take into consideration the putative introduction histories modeled,

moving backwards through time from the observed data. The PODs most similar to the observed dataset are then selected (with replacement) via a Euclidean distance measure (Estoup & Guillemaud, 2010; Lombaert *et al.*, 2010; Cornuet *et al.*, 2014; Valentin *et al.*, 2017). The selected PODs have relative posterior probabilities calculated for their respective scenarios via a logistic regression estimate, allowing the user to select a significantly different scenario as being most likely to have occurred (Cornuet *et al.*, 2014; Valentin *et al.*, 2017).

We framed testable scenarios around three main questions: 1) can we identify from which of the source clade(s) the cardinals brought to Hawaii (the founding cardinals) were sourced; 2) can we assess if the 300+ cardinals that reached Hawaii were effectively divided into three evenly distributed groups of founders and released simultaneously across all three islands, or were approximately 100 founders introduced to a single island during each introduction event over three years; and 3) can we identify which island population(s) provided the founders of the Maui population. For each question we modeled two or more scenarios, and then compared these against each other in order to quantify their relative probabilities. We used the program DIYABC to conduct these analyses (Cornuet *et al.*, 2008; Cornuet *et al.*, 2014), and used the following summary statistics to conduct our analyses: One Sample statistics - Number of Haplotypes and Number of Segregating sites, Two Sample statistics – Number of Haplotypes.

To address the first question (source clade), we evaluated four variations of three scenarios. The first scenario supposed that the source of Hawaii cardinals was the western region of the source clade (see Results for clade analysis below; **Figure 1**). The second

scenario supposed the source individuals were derived from the eastern region of the source clade (**Figure 1**). The third scenario supposed that the Hawaii population was a mix of both regions. For these three scenarios, the first variation evaluated which region was the likely source of the Hawaii introduction without enforcing a change in effective population size (i.e. no genetic bottleneck). The second variation reduced the effective population size after initial introduction into the Hawaiian archipelago (genetic bottleneck – conditioned to be less than both native sources), then allowed the population to change (no condition set to Hawaiian populations). The last two variations (three and four) considered the possibility that each source region contained an unsampled population that was the source of Hawaii founders, and contains genetic haplotypes not present in our dataset. Variations three and four were identical to the above second and first variations, respectively, except an unsampled population for each region was used rather than the region data itself. The variation with the highest confidence in scenario choice (i.e. contained the least amount of error) and contained a statistically significant scenario was considered the most probable, given our data.

To address the second question (pattern of release events) we evaluated two variations of two scenarios. The first scenario supposed the 300+ cardinals transported from the mainland were equally divided among the three islands, but equal subsets were released in 1929, 1930, and 1931 resulting in smaller founding population sizes. The second scenario supposed that of the 300+ founding cardinals, roughly 100 were acquired and introduced to one island per year. We again evaluated whether there was evidence of a population bottleneck with our scenario variations. For the first variation there was no change in effective population size enforced after founders were introduced to Hawaii

(i.e. no enforced bottleneck – no restrictions placed on Hawaii parameters). For the second variation we did enforce an initial reduction in effective population size (i.e. bottleneck – restricted Hawaii parameters to be less than native range and fit scenario) then allowed the population to increase.

To address the third question (source of Maui cardinals), we evaluated three different scenarios: a) colonizers to Maui came from Hawaii Island; b) colonizers came from Oahu; c) colonizers were derived from both islands.

In all ABC scenarios, we set parameter priors to fit a uniform distribution (under default bounds), and placed conditions on parameter priors only to fit the intention of each scenario as defined above. We chose the HKY mutation model, based on the results from PartitionFinder during the phylogenetic analysis (see results below, **Table S2**), and set it identically for all scenarios evaluated (**Table 3**). We ran all experiments for three million computations prior to conducting any analyses.

Results

Morphological Analyses

Reinforcing the findings of Mathys and Lockwood (2011), we found that northern cardinal populations showed substantial morphological divergence across the main Hawaiian Islands (**Table 1**, **Figure 2**). In particular, cardinals from Hawaii Island differ from those on all other islands except Maui (**Table 1**). Cardinals resident on Hawaii Island and Maui tend to be larger than their counterparts on Oahu and Kauai, especially in tail length (**Figure 2**). We also find residents of Maui have significantly larger wings than all other Hawaiian island populations (**Figure 2**).

Our evaluation of morphological differences among the four evaluated phylogeographic clades confirms the existence of substantial morphological variation between northern cardinal clades across their native range (**Table 1**, **Figure 3**). In particular, we found that the native populations differ substantially in body size with *C.c. igneous* being the largest of the set, *C.c. cardinalis* moderately large sized, and the two lower-latitude clades in Mexico the smallest (**Figure 3**). Bill dimensions also vary across clades, with cardinals exhibiting somewhat shorter and pointier bills (relative to body size) in the southern Mexican clades as compared to the two clades that cover sections of the US (**Figure 3**). We found very little differentiation in morphology between the two southern clades *C.c. saturatus* and *C.c. coccineus*. This result agrees with ongoing research that indicates that the island clade of *C.c. saturatus* (located just off the Yucatan peninsula) is a recently derived population established via colonization of nearby *C.c. coccineus* individuals (Smith *et al.* 2011).

Collectively, the cardinals of Hawaii do not fall neatly into the morphological trait space of any single native clade (**Figure 3**). Hawaiian cardinals overlap in trait space with all four native clades for which we were able to obtain measurements. In addition, there is no clear pattern whereby the morphology of cardinals resident on an island correspond to the morphology of cardinals from any one clade (**Figure 3**). Thus, there is no indication from this analysis that the pattern of morphological divergence observed on the islands matches any observed pattern of morphological differentiation among clades across the native range.

Sequence Generation and Phylogeographical Analysis

After amplifying and sequencing the ND2 mtDNA locus for the 41 northern cardinal samples obtained from Hawaii, we found a total of 19 haplotypes (**Table 2**). We observed six, ten, and seven haplotypes for the cardinals present on Oahu, Hawaii Island, and Kauai, respectively, while Maui had just three haplotypes. These sequence data can be found in Genbank under accession numbers MH010209-MH010303.

After combining our sequence data with that of Smith *et al.* (2011), PartitionFinder 1.1.1 produced a single model scheme for unpartitioned data, with the GTR+I+G DNA evolutionary model. In contrast, data partitioned by codon position produced 5 model schemes, with the best scheme keeping the start codon for all three reading frames separated. For this scheme, the first and second codon positions were assigned the HKY + I evolutionary model, while the third codon position was assigned the GTR + G model (**Table S2**). After tree construction in MrBayes v3.2.2, and final assembly in FigTree v1.4.2, we found the unpartitioned scheme produced a tree showing a similar topology to that produced by Smith *et al.* (2011) (**Figure 4**). While of the 19 haplotypes found in the Hawaiian archipelago, 14 (74%) were not observed by Smith *et al.* (2011), all cardinal sequences from Hawaii fell into the *C. c. cardinalis* clade providing strong evidence that this was the single source clade for cardinals on Hawaii. The Hawaiian haplotypes were evenly distributed across the range of haplotypes in the *C. c. cardinalis* clade (**Figure 4)**, which also did not show geographical assortment across the wide sampled range (**Figure S2**).

Approximate Bayesian Computation Analysis

The first question was aimed at identifying the region within the native source clade (*C.c. cardinalis*) from which the Hawaiian cardinals were derived. Without an enforced bottleneck we did not find a significant difference in the relative probabilities among any of the three tested scenarios (**Table 4**). When we enforced a population bottleneck (variation two), however, we found the scenario where cardinals on Hawaii were derived from the eastern region of the *C.c. cardinalis* clade to be significantly more likely (**Table 4**). However, both of these variations had confidence scores below the remaining two, which included unsampled populations from each region. Of the remaining two variations, the third had the highest confidence score (0.643, **Table 4**), with the scenario where the source population came from an unsampled population within the eastern region of the *C.c. cardinalis* clade being the only one to be significant between the two variants (three and four) (0.8971 [0.6177, 1.000], **Table 4**).

Regarding the second question, results also indicated little genetic support for the transported founders having been released on each island across several years, regardless of the presence of bottlenecks (0.0015 [0.000, 0.0777] – effective population held static, and 0.0018 [0.000, 0.0780] – effective population bottleneck followed by increase). The scenario where cardinals were released simultaneously and evenly across islands proved most likely, and the scenario that included bottlenecks (prob = 0.5879 [0.5344, 0.6415]) was significantly more likely than the scenario where a bottleneck was not enforced (prob = 0.4088 [0.3322, 0.4853]).

Regarding the third question, we found the scenario where individuals from Hawaii Island colonized Maui had a higher relative probability (prob = 0.5086 [0.4636,

0.5537]) than the scenario were Maui colonizers were derived from Oahu (prob = 0.1508 [0.1226, 0.1789]), or from both Oahu and Hawaii Island (prob = 0.3406 [0.2982, 0.3831]).

Discussion

A species' invasion history can profoundly influence the degree of divergence observed, including via founder effects whereby phenotypic and genetic spatial structure in the species' native range is captured and then recapitulated across exotic populations (Keller & Taylor, 2008). Here we combined a suite of morphological and genetic analyses to deduce the invasion history of northern cardinals on Hawaii, including directly testing for the presence of founder effects in producing between-island phenotypic variation. In total, we found evidence of an intricate history of colonization, spread, and post-invasion morphological differentiation.

Our phylogenetic tree indicated that the *C.c. cardinalis* native clade was the only source of individuals introduced to the Hawaiian Islands. Furthermore, the ABC analyses gave the highest likelihood to the scenario where founding individuals were derived from populations in the eastern half of that clade. San Francisco was an active commercial center in the 1920 and 1930s so we suspect that the 300 to 350 founder cardinals were captured near a large city in the eastern portion of the US and shipped by train to San Francisco.

Furthermore, we found that only about half the measured individuals from Hawaii fell within the morphological variation we documented across the *C.c. cardinalis* clade. Since all cardinals now resident on Hawaii were likely derived from individuals sourced

from that native clade, the morphological analyses make clear that the observed morphological divergence of cardinals on Hawaii is not the result of founder effects. The cardinals in Hawaii that exceeded the *C.c. cardinalis* trait space fell mostly within the southeastern clades of *C.c. coccineus* (a clade within the southeastern peninsula of Mexico) and *C.c. saturatus* (an island population derived from *C.c. coccineus*). There is no evidence that northern cardinal populations in eastern North America have evolved over the time span that cardinals have been resident in Hawaii. Thus, although circumstantial, this evidence suggests that cardinals on Hawaii have diverged in morphology away from their native continental source population toward body sizes and bill shapes that are more typical of island and peninsular cardinal populations.

Further, we found that cardinal populations on Kauai, Oahu and Hawaii Island were likely all simultaneously founded by equal numbers of transported individuals, and that all of these founder populations experienced a bottleneck. While many of the haplotypes present among Hawaii cardinals were not present in the *C.c. cardinalis* clade sequences, the ND2 locus sampled exhibited very high levels of diversity (48 haplotypes in 78 specimens, Smith *et al.* 2011). Therefore, the ABC analysis suggested that these haplotypes likely originated from unsampled haplotypes within the eastern region of the source clade. While it is possible some of the haplotype variants could have emerged post-introduction through mutations (e.g. Agrawal & Wang, 2008; Kaňuch *et al.*, 2014; Vandepitte *et al.*, 2014), such mutations would likely have added only a few new haplotypes and not likely the 74% new haplotypes we detected. With a substantial increase in sampling across the *C.c. cardinalis* clade we suspect a number of these

haplotypes would be found, and from this information it may be possible to further resolve the source population of Hawaiian cardinals.

Finally, the ABC analyses demonstrated a strong likelihood that the Hawaii Island cardinal population was the source of cardinals now resident on Maui. This scenario is supported by the fact that cardinals on Maui do not differ in overall morphology from those found on Hawaii Island. However, we do find that Maui cardinals have larger wings than cardinals on the other islands. The larger wing size in Maui cardinals could have resulted from selection on founders, as there is no record that humans mediated the expansion of cardinals to Maui. If so, this might be the only evidence of founder effects in Hawaiian cardinals.

Based on this collection of evidence, the story of the establishment and divergence of northern cardinals on Hawaii seems to be as follows. In the early 1930s, 300 to 350 cardinals were captured in the eastern US and shipped to San Francisco likely by train. These individuals were then shipped by boat to the Hawaiian archipelago, and released simultaneously and in approximately equal numbers on Hawaii Island, Kauai and Oahu. A subset of these individuals founded viable exotic populations on all three islands. At a later date, individuals from Hawaii Island colonized Maui. Since these initial founding events cardinals have substantially diverged in morphology from their native source clade, and within the islands, cardinals on Hawaii Island and Maui show particularly divergent morphologies compared to the other islands.

Our results add to a growing number of studies that demonstrate evolution within an invasive species' new range (Whitney & Gabler, 2008; Egizi *et al.*, 2015). Most questions now center on deducing the mechanisms driving these patterns, using these

examples to inform our broader understanding of evolutionary diversification processes. Relative to the evolution of morphological traits amongst birds colonizing islands, likely mechanisms of divergence center on factors such as thermoregulation, competition, and predation all of which can vary substantially on islands as compared to a mainland (Moulton & Lockwood, 1992; Moulton *et al.*, 2001; Duncan & Blackburn, 2002; Luther & Greenberg, 2014). Our approach combining detailed historical records, comprehensive phenotypic analysis, and rigorous phylogenetic and population genetic techniques allowed us to reveal insights into the mechanisms that have produced post-invasion divergence in this exotic bird. Aside from conducting a full genomic or transcriptomic analysis of northern cardinals on Hawaii, however, we cannot at this point determine which of these potential mechanisms has driven the evolution of exotic cardinals in Hawaii.

References

Agrawal, A.F. & Wang, A.D. (2008) Increased transmission of mutations by low-condition females: evidence for condition-dependent DNA repair. *PLoS biology*, **6**, e30.

Allendorf, F.W. & Lundquist, L.L. (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, **17**, 24-30.

Arendt, W.J. & Faaborg, J. (1989) Sources of Variation in Measurements of Birds in a Puerto Rican Dry Forest (Fuentes de Variación en las Medidas de Aves Estudiadas en un Bosque Xerofítico de Puerto Rico). *Journal of Field Ornithology*, 1-11.

Badyaev, A. & Martin, T. (2000a) Individual variation in growth trajectories: phenotypic and genetic correlations in ontogeny of the house finch (Carpodacus mexicanus). *Journal of Evolutionary Biology*, **13**, 290-301.

Badyaev, A.V. & Martin, T.E. (2000b) Sexual dimorphism in relation to current selection in the house finch. *Evolution*, **54**, 987-997.

Bjordal, H. (1983) Effects of deep freezing, freeze-drying and skinning on body dimensions of house sparrows (*Passer domesticus*). *Cinclus*, **6**, 105-108.

Buuren, S. & Groothuis-Oudshoorn, K. (2011) mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, **45**

Cornuet, J.-M., Santos, F., Beaumont, M.A., Robert, C.P., Marin, J.-M., Balding, D.J., Guillemaud, T. & Estoup, A. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713-2719.

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M. & Estoup, A. (2014) DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187-1189.

Dlugosch, K. & Parker, I. (2008a) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular ecology*, **17**, 431-449.

Dlugosch, K.M. & Parker, I.M. (2008b) Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology letters*, **11**, 701-709.

Duncan, R.P. & Blackburn, T.M. (2002) Morphological over-dispersion in game birds (Aves: Galliformes) successfully introduced to New Zealand was not caused by interspecific competition. *Evolutionary Ecology Research*, **4**, 551-561.

Egizi, A., Fefferman, N.H. & Fonseca, D.M. (2015) Evidence that implicit assumptions of 'no evolution' of disease vectors in changing environments can be violated on a rapid timescale. *Philos Trans R Soc Lond B Biol Sci*, **370**

Estoup, A. & Guillemaud, T. (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113-4130.

Freed, L.A., Conant, S. & Fleischer, R.C. (1987) Evolutionary ecology and radiation of Hawaiian passerine birds. *Trends in Ecology & Evolution*, **2**, 196-203.

Haftorn, S. (1982) Variation in body measurements of the Willow Tit Parus montanus, together with a method for sexing live birds and data on the degree of shrinkage in size after skinning. *Cinclus*, **5**, 16-26.

Jensen, H., Sæther, B.E., Ringsby, T.H., Tufto, J., Griffith, S. & Ellegren, H. (2003) Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (Passer domesticus). *Journal of evolutionary biology*, **16**, 1296-1307.

Kaňuch, P., Berggren, Å. & Cassel‐Lundhagen, A. (2014) Genetic diversity of a successful colonizer: isolated populations of Metrioptera roeselii regain variation at an unusually rapid rate. *Ecology and evolution*, **4**, 1117-1126.

Kassambara, A. & Mundt, F. (2016) Factoextra: extract and visualize the results of multivariate data analyses. *R package version*, **1**

Keller, S.R. & Taylor, D.R. (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, **11**, 852-866.

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and evolution*, **29**, 1695-1701.

Lockwood, J.L., Moulton, M.P. & Anderson, S.K. (1993) Morphological assortment and the assembly of communities of introduced passeriforms on oceanic islands: Tahiti versus Oahu. *The American Naturalist*, **141**, 398-408.

Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B. & Estoup, A. (2010) Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *PLoS ONE*, **5**, e9743.

Lucek, K., Sivasundar, A. & Seehausen, O. (2014) Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. *Evolution*, **68**, 2619-2632.

Luther, D. & Greenberg, R. (2014) Habitat type and ambient temperature contribute to bill morphology. *Ecology and evolution*, **4**, 699-705.

Mathys, B.A. & Lockwood, J.L. (2011) Contemporary morphological diversification of passerine birds introduced to the Hawaiian archipelago. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20102302.

Moulton, M.P. & Lockwood, J.L. (1992) Morphological dispersion of introduced Hawaiian finches: evidence for competition and a Narcissus effect. *Evolutionary Ecology*, **6**, 45-55.

Moulton, M.P., Sanderson, J.G. & Labisky, R.F. (2001) Patterns of success in game bird (Aves: Galliformes) introductions to the Hawaiian islands and New Zealand. *Evolutionary Ecology Research*, **3**, 521-535.

Phillimore, A., Owens, I., Black, R., Chittock, J., Burke, T. & Clegg, S. (2008) Complex patterns of genetic and phenotypic divergence in an island bird and the consequences for delimiting conservation units. *Molecular Ecology*, **17**, 2839-2853.

Pyle, R. & Pyle, P. (2009) The birds of the Hawaiian islands: occurrence, history, distribution, and status. *BP Bishop*,

Rambaut, A., Suchard, M., Xie, W. & Drummond, A. (2014) Tracer v. 1.6. Institute of Evolutionary Biology, University of Edinburgh. In:

Ricklefs, R.E. & Travis, J. (1980) A morphological approach to the study of avian community organization. *The Auk*, 321-338.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Smith, B.T., Escalante, P., Baños, B.E.H., Navarro-Sigüenza, A.G., Rohwer, S. & Klicka, J. (2011) The role of historical and contemporary processes on phylogeographic structure and genetic diversity in the Northern Cardinal, Cardinalis cardinalis. *BMC evolutionary biology*, **11**, 136.

Sorenson, M.D. & Fleischer, R.C. (1996) Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Sciences*, **93**, 15239-15243.

Suarez, A.V. & Tsutsui, N.D. (2008) The evolutionary consequences of biological invasions. *Molecular Ecology*, **17**, 351-360.

Team, R.C. (2014) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. In:

Valentin, R.E., Nielsen, A.L., Wiman, N.G., Lee, D.-H. & Fonseca, D.M. (2017) Global invasion network of the brown marmorated stink bug, Halyomorpha halys. *Scientific Reports*, **7**

Vandepitte, K., Meyer, T., Helsen, K., Acker, K., Roldán‐Ruiz, I., Mergeay, J. & Honnay, O. (2014) Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular Ecology*, **23**, 2157-2164.

Westley, P.A., Conway, C.M. & Fleming, I.A. (2012) Phenotypic divergence of exotic fish populations is shaped by spatial proximity and habitat differences across an invaded landscape. *Evolutionary Ecology Research*, **14**, 147-167.

Whitney, K.D. & Gabler, C.A. (2008) Rapid evolution in introduced species,'invasive traits' and recipient communities: challenges for predicting invasive potential. *Diversity and Distributions*, **14**, 569-580.

Winker, K. (1993) Specimen Shrinkage in Tennessee Warblers and" Traill's" Flycatchers (Se Encojen Especímenes de Vermivora peregrina y Empidonax traillii). *Journal of Field Ornithology*, 331-336.

Xu, C.Y., Julien, M.H., Fatemi, M., Girod, C., Van Klinken, R.D., Gross, C.L. & Novak, S.J. (2010) Phenotypic divergence during the invasion of Phyla canescens in Australia and France: evidence for selection‐driven evolution. *Ecology Letters*, **13**, 32-44.

Table 1. Results from the MANOVA analysis of northern cardinal morphological features taken across populations. Global results are the overall MANOVA testing for differences in PC1 and PC2 between the clades within the native range and Hawaii (grouped together), and the five main Hawaiian Islands. P-values for MANOVA tests indicate overall significance across both PC1 and PC2, with individual PCs found significant highlighted in bold. Effect sizes for each PC are calculated using partial Eta^2.

Hawaii (whole) & Native Range						
Source	*n*	*df*	*approx. F*	*P*	PC1 effect size	PC2 effect size
Global (PC1 & PC2)	229	4	31.91	< 2.2e-16	**0.32**	**0.40**
cardinalis x igneus	108	1	39.16	1.74E-13	**0.41**	0.0094
cardinalis x coccineus	92	1	39.93	3.57E-13	0.0040	**0.47**
cardinalis x saturatus	89	1	35.54	4.85E-12	0.00014	**0.45**
igneus x coccineus	34	1	31.83	1.99E-08	**0.44**	**0.55**
igneus x saturatus	31	1	26.04	2.78E-07	**0.35**	**0.56**
coccineus x saturatus	15	1	0.28	0.7634	0.011	0.029
Hawaii (whole) x cardinalis	189	1	40.89	1.81E-15	**0.048**	**0.29**
Hawaii (whole) x igneus	131	1	53.36	< 2.2e-16	**0.36**	**0.14**
Hawaii (whole) x coccineus	115	1	11.28	3.39E-05	**0.043**	**0.10**
Hawaii (whole) x saturatus	112	1	8.99	2.41E-04	0.020	**0.10**
Hawaii only (by island)						
Source	*n*	*df*	*approx. F*	*P*	PC1 effect size	PC2 effect size
Global (PC1 & PC2)	103	3	2.95	0.00882	**0.093**	0.041
Hawaii Island x Kauai	56	1	4.57	0.01461	0.063	0.041
Hawaii Island x Maui	58	1	1.13	0.329	0.038	0.003
Hawaii Island x Oahu	65	1	8.19	6.83E-04	**0.14**	0.010
Kauai x Maui	38	1	1.75	0.1878	0.0023	0.083
Kauai x Oahu	45	1	0.7388	0.4835	0.021	0.022
Maui x Oahu	47	1	2.15	0.1281	0.033	0.034

Table 2. Summary of number of samples (*n.*) used in the genetic analyses conducted, with localities sorted by mtDNA clade for the native range (with the west and east regions for *C.c.cardinalis* identified) and the Hawaiian archipelago. Each clade, and Hawaii, is further subdivided by locality, while providing the number of haplotypes per location (*n.* Haps) and the haplotypes observed. Any haplotypes followed by a number in parentheses indicates multiple specimens observed with said haplotype. Haplotypes in bold are those found only in Hawaii.

Locality	***n.***	***n.* Haps**	**Haplotypes**
***C.c.cardinalis* (Total)**	**78**	**48**	
***C.c.cardinalis* (West)**	**51**	**32**	
Coahuila	7	6	4, 33(2), 34, 47, 48, 52
Kansas	5	4	24(2), 37, 48, 60
Louisiana	9	8	6, 19, 24, 26, 27(2), 30, 31, 46, 47
Oklahoma	10	7	22(2), 28, 41, 43, 45, 47 (2), 48 (2)
Tamaulipas/Nuevo Leon	8	8	5, 22, 32, 33, 38, 44, 50, 71
Texas/New Mexico	12	10	15, 22, 23, 27, 33, 47, 48(3), 53, 54, 61
Queretaro	2	2	36, 55
Veracruz	1	1	40
***C.c.cardinalis* (East)**	**27**	**20**	
New York	9	8	13, 18 (2), 20, 24, 31, 36, 47, 51
Florida/Georgia	9	7	10, 14, 17, 21, 22(2), 24, 49(2)
Minnesota/Wisconsin	9	7	11, 16, 25(3), 47, 48, 56, 62
C.c.igneus	**49**	**20**	
Arizona/New Mexico	11	2	73, 77(10)
Baja California	13	9	73, 74, 75, 76, 77, 83(5), 84, 88, 92
Sinaloa	19	13	72, 76, 77(6), 78, 80, 81, 82, 83, 85, 86(2), 87, 93, 95
Tiburón Island	4	3	79(2), 94, 95
C.c.coccineus	**11**	**3**	
Campeche	1	1	57
Yucatán	10	3	57(4), 58(5), 59
C.c.carneus	**8**	**3**	
Michoacán	6	1	2
Guerrero	1	1	1
Oaxaca	1	1	3
C.c.satturatus	**8**	**2**	
Cozumel Island			63(7), 64

C.c.mariae	**6**	**3**	
Tres Marías Islands			89(2), 90, 91(2)
Hawaiian archipelago	**41**	**19**	
Hawaii Island	14	10	**7**(2), 22, 27, **29**, **39**(2), 47(2), 48, **66**, **67**(2), **68**
Oahu	8	6	**12**(2), 24, **35**(2), **42**, 48, **69**
Kauai	8	7	**8**, **9**(2), **12**, 24, **35**, 48, **70**
Maui	11	3	22(4), 24(4), **65**(3)

Table 3. Prior distributions used for all ABC analyses. Mutation parameters refer to selected DNA mutation model, distributions used, and bounds for said distributions within the model validation screen.

Description	Prior distribution			
Mutation Parameters				
Mutation model	HKY	10% invariant sites		Shape (2)
Mean mutation rate	Uniform	(min) 1.00E-7	(max) 1.00E-5	
Indiv. locus mutation rate	Gamma	(min) 1.00E-7	(max) 1.00E-5	Shape (2)
Mean coefficient (k C/T)	Uniform	(min) 1.5	(max) 20	
Indiv. locus coefficient (k C/T)	Gamma	(min) 1.5	(max) 20	Shape (2)

Table 4. Probability and 95% credible interval for all Approximate Bayesian Computation scenarios used throughout the study, along with confidence in scenario choice. Variations in scenarios refer to no enforced reductions in the exotic population's effective population size (i.e. no bottleneck – variations 1 and 4), or enforced reductions (i.e. bottleneck) followed by a change in effective population that was free to either increase or decrease (variations 2 and 3).

Experiment	Prob.	95% CI	Conf.
C.c cardinalis source region (variation 1)			0.511
1: Western source region	0.3587	[0.3042, 0.4133]	
2: Eastern source region	0.2807	[0.2292, 0.3321]	
3: Western source + Eastern source	0.3606	[0.3106, 0.4105]	
C.c cardinalis source region (variation 2)			0.501
1: Western source region	0.2576	[0.2203, 0.2949]	
2: Eastern source region	0.4451	[0.4048, 0.4854]	
3: Western source + Eastern source	0.2973	[0.2509, 0.3437]	
C.c cardinalis source region (variation 3)			0.643
1: Western unsampled source region	0.0638	[0.0000, 0.2523]	
2: Eastern unsampled source region	0.8971	[0.6177, 1.0000]	
3: Western + Eastern unsampled source	0.0391	[0.0000, 0.1722]	
C.c cardinalis source region (variation 4)			0.596
1: Western unsampled source region	0.4016	[0.0000, 1.0000]	
2: Eastern unsampled source region	0.0000	[0.0000, 1.0000]	
3: Western + Eastern unsampled source	0.5984	[0.0000, 1.0000]	
Hawaii introduction scheme			0.604
1: Introduced to Hawaii evenly (no bottleneck enforced)	0.4088	[0.3322, 0.4853]	
2: Introduced to Hawaii structured (no bottleneck enforced)	0.0018	[0.0000, 0.0780]	

3: Introduced to Hawaii evenly (bottleneck enforced)	0.5879	[0.5344, 0.6415]	
4: Introduced to Hawaii structured (bottleneck enforced)	0.0015	[0.0000, 0.0777]	
Maui introduction scheme			0.491
1: Colonized from Hawaii island	0.5369	[0.5059, 0.5679]	
2: Colonized from Oahu	0.1501	[0.1309, 0.1692]	
3: Colonized from both	0.3130	[0.2845, 0.3415]	

Figure 1. Map depicting the six northern cardinal (*Cardinalis cardinalis*) mitochondrial clades in their native range. Map adapted from Smith *et al.* (2011). Clades represented as follows: blue = *C.c. cardinalis*, green = *C.c. igneous*, orange = *C.c. miriae*, red = *carneus*, brown = *C.c. coccineus*, and yellow = *C.c. saturatus*. The textured portion of the blue clade represents the eastern region of the *C.c. cardinalis* clade, while the non-textured portion represents the western region.

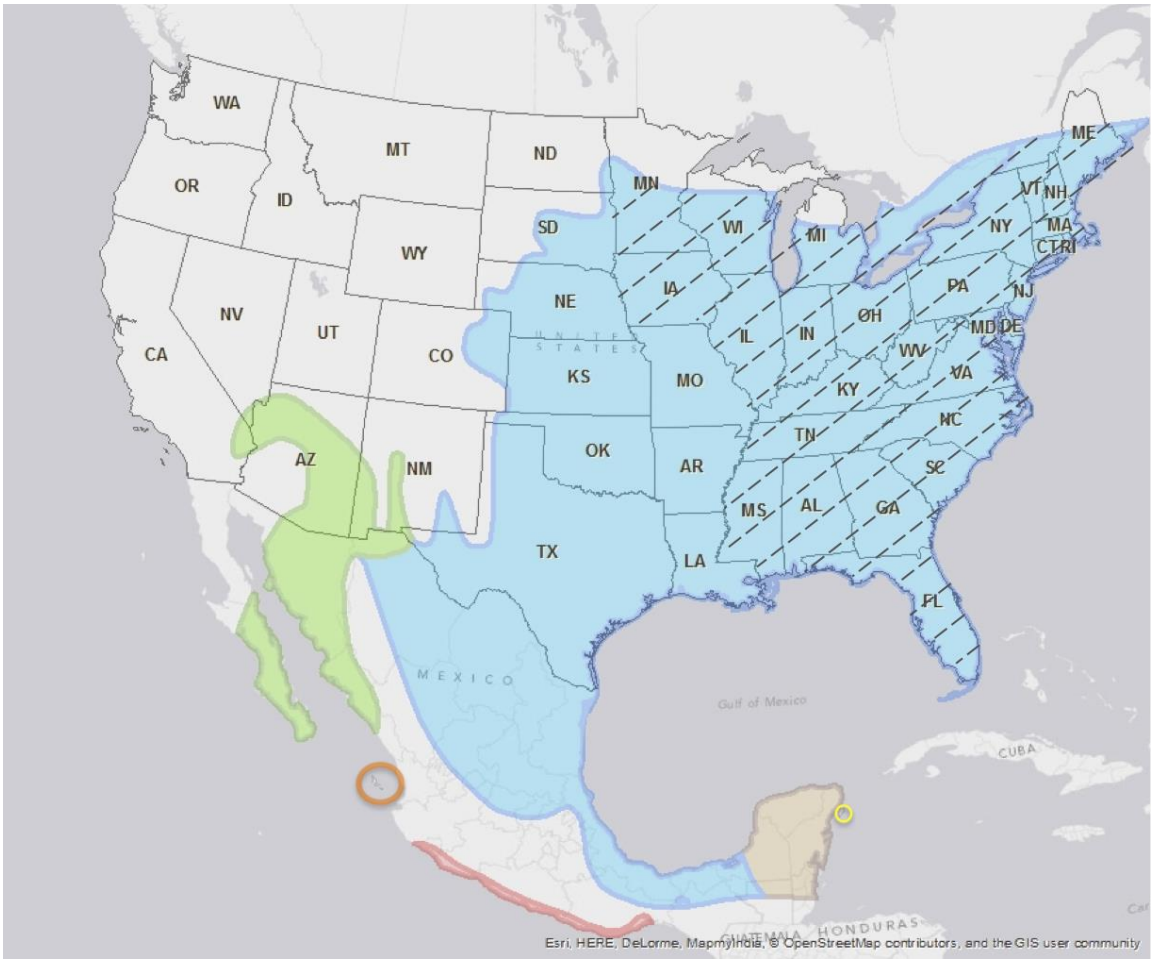

Figure 2. Morphological characteristics of northern cardinals sampled across five main Hawaiian Islands. The bars represent morphological traits read from left to right as: tail length (red), wing length (green), bill depth (blue), and bill width (purple). The zero-line is the mean trait value calculated across all individuals and all islands. Deviations away from this value per island are depicted as bars, including calculated within-island standard deviation for this trait. A large deviation from the island-wide mean suggests that northern cardinal individuals sampled on that island have a divergent morphology. Most large differences across islands were due to tail and wing lengths.

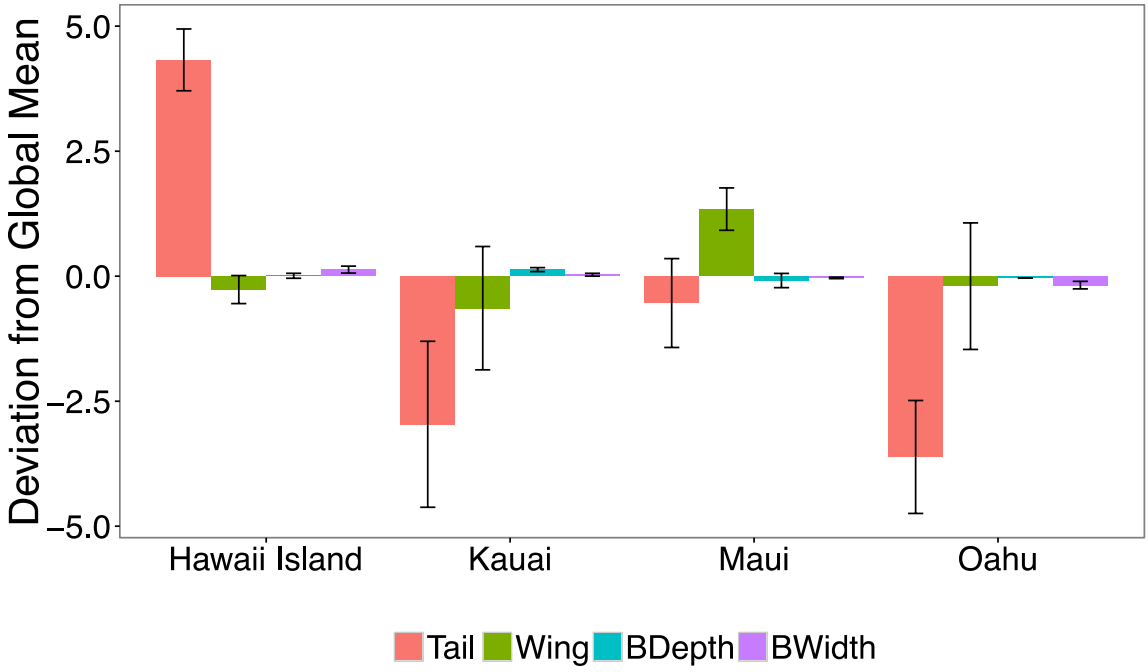

Figure 3. Two-dimensional representation of northern cardinal morphological trait space using PC1 and PC2. PC1 reflects overall body size, whereas PC2 measures how bill depth and width change relative to a change in body size. We only include four of the native range clades in this analysis due to low sample size in two clades. Each oval encapsulates 95% of the variation in morphology between the individuals we measured, representing a clade-specific trait space. Large symbols within each oval depict the mean PC scores for each clade. We depict all individual cardinals captured and measured in Hawaii in light blue (with different shapes for each island) to visually show the distribution of their morphology (light blue oval) relative to native clades.

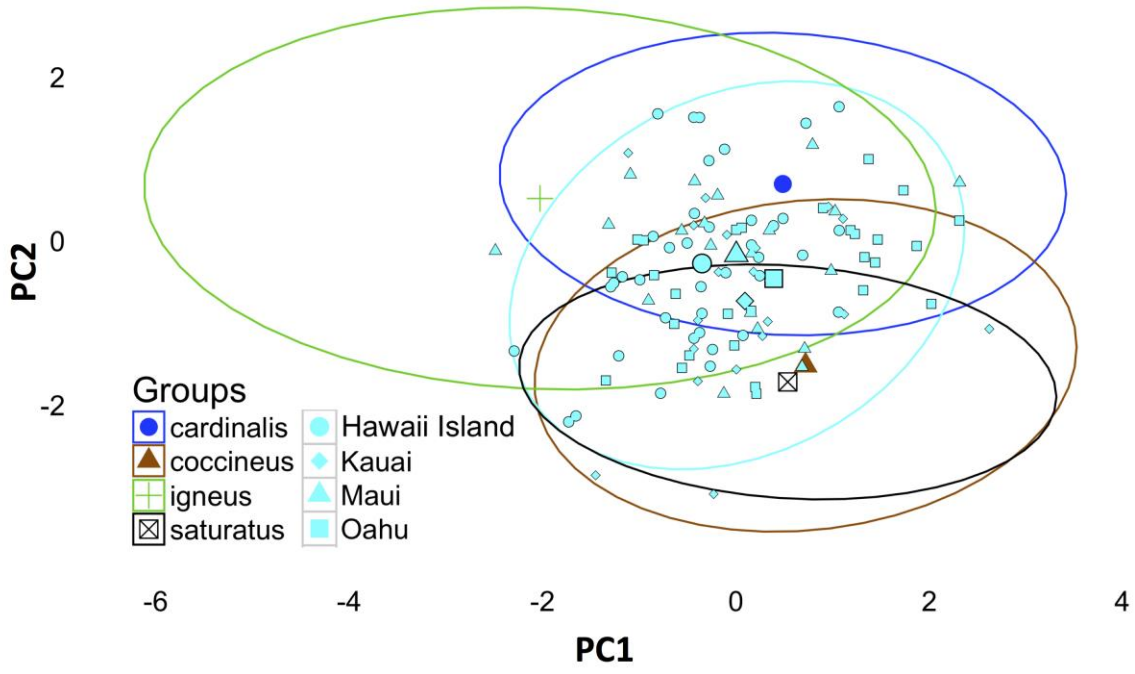

Figure 4. A phylogenetic breakdown of all native northern cardinal sequences analyzed by Smith *et al.* (2011), with the Hawaiian Islands samples we sequenced intermixed within the dataset. Monophyletic groups were categorized (and color-coded) to their respective mitochondrial DNA clade, while the branches representing the Hawaiian individuals were color coded in light blue. All Hawaiian samples grouped with the *C.c. cardinalis* native range clade.

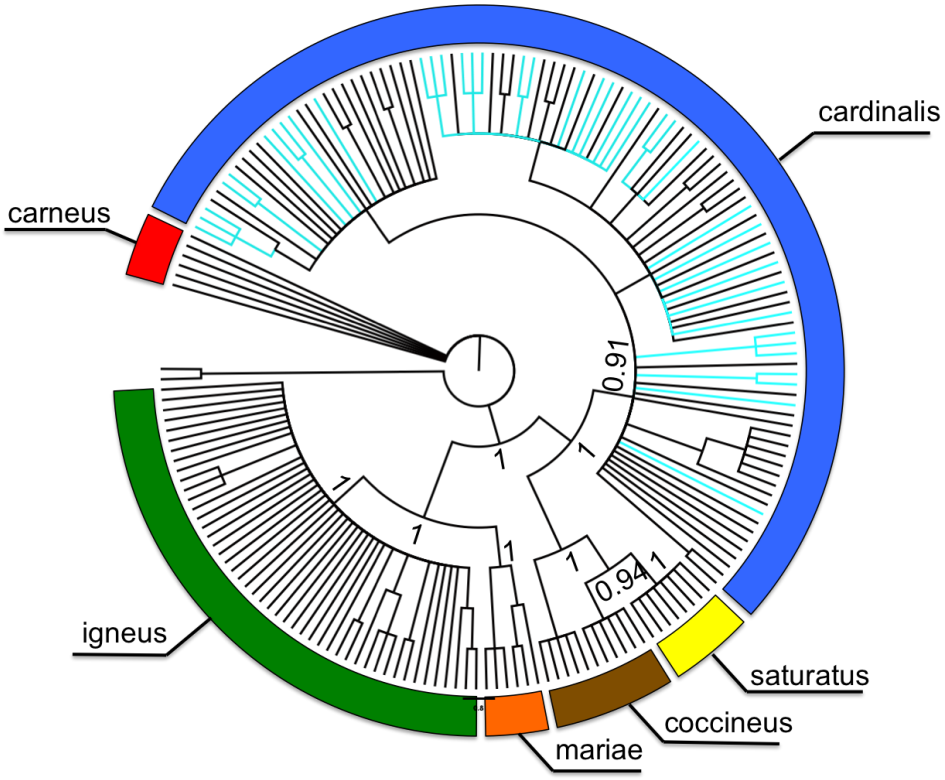

Figure S1. Three scenarios describing the introduction of northern cardinals to the main Hawaiian Islands. Scenario I (two gray lines) = individual cardinals were collected from two clades, and the individuals of one clade were introduced to a subset of all islands whereas the individuals of the other clade were introduced to a separate set. Scenario II (black line) = all individuals collected from the native range come from a single clade and they were released across all islands. Scenario III (blue line) = individuals were collected from more than one native clade but all were introduced to every island, effectively creating an admixed population on each island. We have depicted specific clades as sources, and islands as sets where introduction occurred, for illustration only. We have no reason to *a priori* restrict the set of clades, or the introduction scenarios, considered within our analyses.

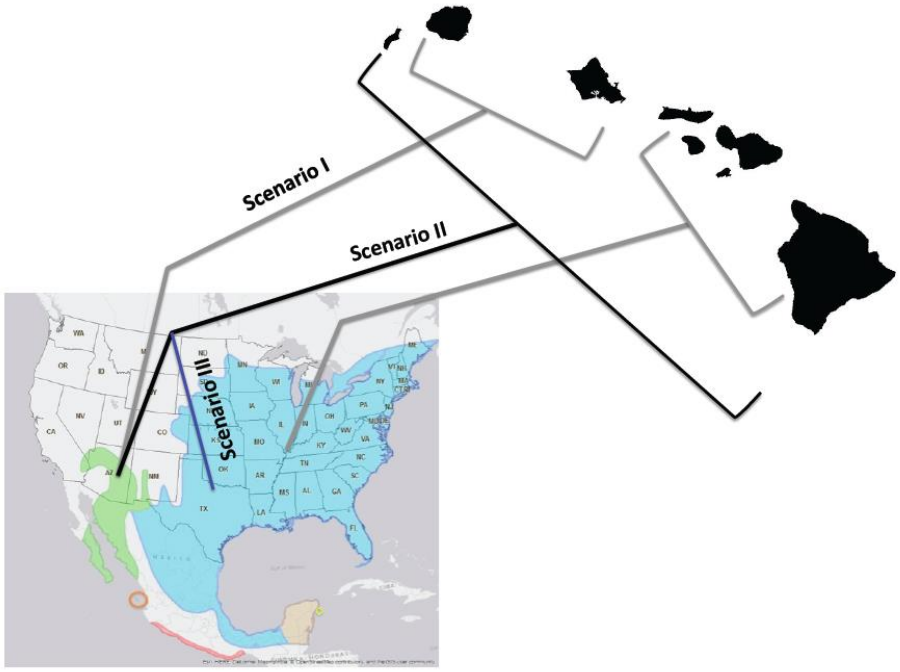

Figure S2. Haplotype network for *C.c.cardinalis* and Hawaiian cardinals. Network nodes in black indicate haplotypes found in *C.c.cardinalis*, while nodes in blue are those found in Hawaii. Numbers in bold indicate haplotypes seen both in Hawaii and across the *C.c.cardinalis* native range, while black nodes with a blue numbers (i.e. 22, 24, 47) indicate haplotypes that are seen in either eastern and western regions of the *C.c.cardinalis* clade and in Hawaii. Numbers in red indicate haplotypes found in the western region of the *C.c.cardinalis* range, while numbers in black (not next to blue nodes) indicate haplotypes found in the eastern region of *C.c.cardinalis* range. Numbers in purple indicate haplotypes seen in both the eastern and western regions.

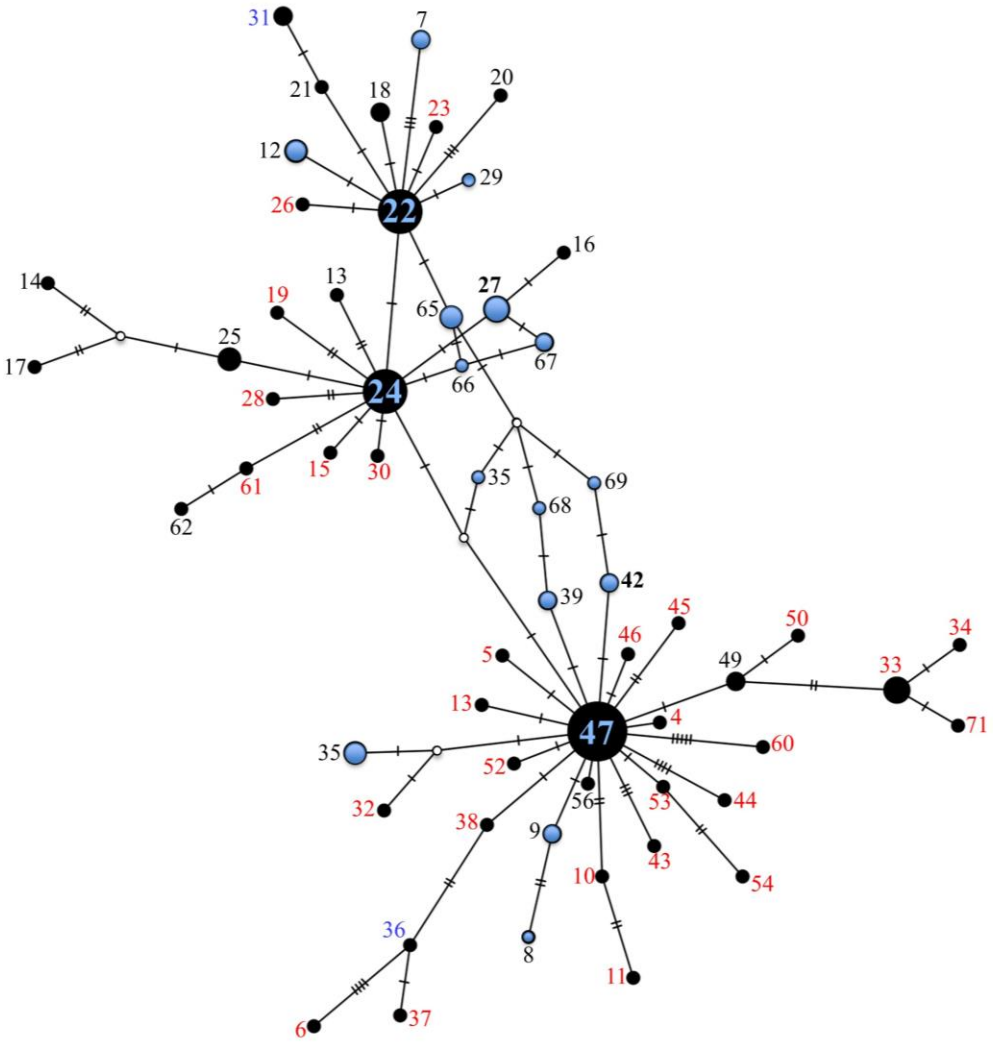

Table S1. Principle Component Analysis (PCA) outputs for the combined native range and Hawaiian northern cardinal (all islands in one group) morphological dataset. The top table provides eigenvalues and the amount of morphological variation explained by each principle component (PC). The bottom table provides the PC loadings (correlation) for each morphological trait measured.

PCA Summary		
	Eigenvalue	**% Variance**
PC.1	1.79692	44.923
PC.2	1.19142	29.7855
PC.3	0.57145	14.2864
PC.4	0.44021	11.0052

PC Loadings				
	PC.1	**PC.2**	**PC.3**	**PC.4**
Tail Length	-0.49524	0.46083	0.71721	-0.16728
Wing Length	-0.4038	0.60789	-0.68166	-0.05252
Bill Depth	-0.58893	-0.34461	-0.01477	0.73089
Bill Width	-0.49482	-0.54713	-0.14398	-0.65959

Table S2. MrBayes specific mutation models for the phylogeographic analysis found to be best via the program PartitionFinder given the specific codon partitioning schemes. Models were partitioned by codon scheme, and the best model(s) for each were selected and scored via AIC.

PartitionFinder Schemes					
Scheme	**Subset Partitions**	**Best Model**	**Scheme lnL**	**Scheme AIC**	**Number of params**
1	(Gene1_codon1, Gene1_codon2, Gene1_codon3)	GTR+I+G	-2785.2437	6324.48744	377
2	(Gene1_codon1, Gene1_codon2)	GTR+I	-2671.9739	6115.94782	386
	(Gene1_codon3)	GTR+G			
3	(Gene1_codon1, Gene1_codon3)	GTR+I+G	-2703.2277	6172.45536	383
	(Gene1_codon2)	HKY+I			
4	(Gene1_codon1)	HKY+I	-2742.921	6251.8419	383
	(Gene1_codon2, Gene1_codon3)	GTR+I+G			
5	(Gene1_codon1)	HKY+I	-2634.4855	6044.97104	388
	(Gene1_codon2)	HKY+I			
	(Gene1_codon3)	GTR+G			

Global invasion network of the brown marmorated stink bug, *Halyomorpha halys*

Rafael Valentin[1,5], Anne L. Nielsen[2], Nik G. Wiman[3], Doo-Hyung Lee[4], Dina M. Fonseca[1,2]

Author affiliations:
[1]Department of Ecology, Evolution & Natural Resources, Rutgers University, 14 College Farm Rd., New Brunswick, NJ 08901, USA

[2]Department of Entomology, Rutgers University, 93 Lipman Drive, New Brunswick, NJ 08901, USA

[3]Department of Horticulture, Oregon State University, 4017 ALS Building, Corvallis, OR 97330, USA

[4]Department of Life Sciences, Gachon University, Seongnam-si, Gyeonggi-do, South Korea

[5]**Corresponding author**: Rafael Valentin, Ecology and Evolution Graduate Program, 14 College Farm Rd., New Brunswick, NJ 08901, USA; Rafael.Valentin@Rutgers.edu

Abstract

Human mediated transportation into novel habitats is a prerequisite for the establishment of non-native species that become invasive, so knowledge of common sources may allow prevention. The brown marmorated stink bug (BMSB, *Halyomorpha halys*) is an East Asian species now established across North America and Europe, that in the Eastern United States of America (US) and Italy is causing significant economic losses to agriculture. After US populations were shown to originate from Northern China, others have tried to source BMSB populations now in Canada, Switzerland, Italy, France, Greece, and Hungary. Due to selection of different molecular markers, however, integrating all the datasets to obtain a broader picture of BMSB's expansion has been difficult. To address this limitation we focused on a single locus, the barcode region in the cytochrome oxidase I mitochondrial gene, and analyzed representative BMSB samples from across its current global range using an Approximate Bayesian Computation approach. We found that China is the likely source of most non-native populations, with at least four separate introductions in North America and three in Europe. Additionally, we found evidence of one bridgehead event: a likely Eastern US source for the central Italy populations that interestingly share enhanced pest status.

Introduction

Over the past few decades the introduction of non-native species, which are species that have been moved by human action into novel habitats beyond their natural geographic range [1], has increased rapidly and on a global scale [2,3]. However, in order for any non-native species to become established it must first successfully navigate the first

phase of the invasion process, which is transportation into a novel habitat [1,3]. Inadvertent human-mediated transportation of organisms often occurs due to global trade of goods [4]. Therefore, the physical and energetic limitations of long distance dispersal are bypassed and can result in unintended rapid and widespread establishment of non-native species [5,6]. If non-natives become destructive or dangerous in their novel habitats, and spread rapidly beyond their introduced location, they are considered invasive[1]. Prior to becoming invasive, newly introduced non-native species can go unnoticed for some time until populations reach levels that result in significant economic impacts and high management costs [7,8]. This danger is inflated for countries that trade regularly, as trade has been positively correlated with introduction rates of non-native species [3,9]. Research on invasive species has led to recommendations that preventive actions be taken to limit the impacts of new invaders[10], such as a better understanding of the source and likelihood of introductions, followed by novel strategies that allow early detection of incipient non-native populations [11,12].

Approaches to identify the source of an invasive species often make use of genetic methods that compare haplotypes or allele frequencies from potential native ranges with those in the introduced population(s) [13]. While that strategy works when there are clear genetic discontinuities among the native populations being evaluated, it can still be difficult to assign formal probabilities to alternatives (e.g. whether the introduced population came from a native source or secondarily from another introduced population (i.e. bridgehead effect) [9,13,14]. This is especially the case when the analysis is based on a single or a few genetic loci, which has been common in the analyses of non-model organisms with worldwide distributions [15-18]. Approximate Bayesian Computation (ABC)

is a statistically robust Bayesian analysis that allows the direct comparison of multiple introduction hypotheses (known as scenarios) providing distinct likelihoods for each. ABC performs inference computations under a Bayesian framework [13,14,19] that takes into consideration putative evolutionary histories, or in this case introduction histories, by quantifying support for modeled scenarios given the data provided. ABC accomplishes this by generating simulated data, known as pseudo-observed datasets (PODs), and randomly selecting (with replacement) the PODs closest to the observed data (i.e. the genetic data collected) by a Euclidean distance measure. These PODs then have relative posterior probabilities calculated via a logistic regression estimate, which allows the user to determine the most probable invasion scenario [20]. Of course, ABC does have its drawbacks. First, the computational time and power necessary to run moderately complex invasion models can be fairly demanding, at times taking several days to complete one analysis. Second, the scenarios to be evaluated are created by the modeler, and can be subject to bias unless all possible alternative hypotheses are included. Finally, the data must be of sufficient quality to address the desired questions being modeled in the scenario. Without appropriate data, estimates can be biased and inappropriate conclusions can be drawn. When correctly executed, however, ABC is an excellent analysis method for determining the most probable invasion pathways of unintended introductions [13].

Here we incorporated an ABC approach to unravel the pathways of the worldwide expansion of the brown marmorated stink bug (BMSB; *Halyomorpha halys* (Stål)). BMSB is native to Northeast Asia, but non-native populations of BMSB were first detected in the United States of America (US hereafter) in Allentown, Pennsylvania in 1996 [21]. Since then the species has been detected in at least 40 US states, Canada, and

several European countries [22-24]. BMSB can cause significant damage both to agricultural crops and ornamental plants [25], such as the documented damage to tree fruits in New Jersey and Pennsylvania in 2006 and 2007 [26] and WV and MD in 2010 and 2011 [27]. BMSB feeding injury has resulted in significant economic impacts to growers; a one-year loss in excess of 37 million USD across the mid-Atlantic in apples alone, as well as 100% losses to peaches in Maryland [25], and 60-90% losses of peaches in New Jersey (the 4th national peach producer) [28] during a population outbreak in 2010. Since the outbreak in 2010, established BMSB populations have been detected as far south as Georgia and as far west as Michigan, which experienced high populations in 2016 and injury in apples (J Wilson personal communication); and in the Pacific Northwest, specifically Oregon and Washington, where injury to hazelnuts and small fruits has been documented [29,30]. To evaluate possible BMSB source populations, Xu et al. (2014) [31] analyzed genetic variation at two mitochondrial loci in US specimens collected between 2004-2008, as well as in specimens from several populations across the native range of China, Republic of Korea, and Japan. They found best match to populations from the northern region of China, around Beijing, and low mtDNA haplotype diversity in the US populations relative to the native range, possibly indicating a single introduction of a small number of individuals [31].

In Europe, there are currently known established BMSB populations in Switzerland, Italy, France, Greece, Hungary, Serbia, and Romania [22-24,32,33], with very recent detections also in Bulgaria, Russia, Georgia, and the Autonomous Region of Abkhazia (ARA) [34,35]. The first detection of BMSB in Europe was in Zurich, Switzerland in 2007 [36], and soon after it was collected in several locations throughout Switzerland [37].

To find the source population(s), and determine the invasion pathways, genetic analyses were carried out in Switzerland, Italy, France, Greece, and Hungary using mtDNA [38-40]. The likely source(s) for many of the established European populations was not determined due to insufficient power of the analyses to match or reject the few native populations examined [38,39]. Of note, Cesari et al. (2015) hypothesized that BMSBs in Lombardy, Italy were likely the result of southward spread from Switzerland (either natural or human mediated), while the population in Emilia-Romagna, Italy was an independent introduction into Italy possibly from the US [40]. Again, due to insufficient samples from the native range, the authors were unable to determine the likely source.

To address the limitations of the individual studies and provide an updated analysis of the worldwide expansion of BMSB, we developed a meta-analysis using existing data [31,38-40] and an expanded sampling in the US and native ranges. By combining these datasets, we aimed to shore the power of the analyses and increase the likelihood of ascertaining the source(s) of all non-native populations. We therefore amassed and analyzed the largest BMSB dataset to date, with more than 900 individual DNA sequences (both existing and newly generated) including 214 from its native range in China, the Republic of Korea, and Japan. Although we are using a single mtDNA locus, we make use of ABC as our analysis framework to reach robust conclusions about invasion pathways of this global pest species.

Results

Sequence data and haplotypes

We amplified and sequenced 685bp of cytochrome oxidase I (CO1) for 110 of the 139 specimens examined by Xu et al. (2014) [31], as well an additional 80 specimens collected in 2016 in the US (48), Japan (28), and the Republic of Korea (4). Combined with preexisting data from the European studies, we amassed a total of 916 individual DNA sequences for our global dataset (**Table 1**). In the pre-2008 specimens from the US we found one new haplotype (H7), previously unreported, likely because Xu et al. (2014) only sequenced the CO1 locus from four specimens [31]. We found this new haplotype was restricted to California, and all remaining specimens across the Eastern US had the one haplotype (H1) previously reported (**Table 1**). Among the post-2008 US specimens sequenced all eastern specimens had the same haplotype (H1), while the western specimens again displayed additional genetic variation (**Table 1**). The new specimens from the western states (i.e. California, Oregon, and Washington) had five different haplotypes, two of them being haplotypes we had previously seen (H1 and H3) but three were new for the US (H7, H23, and H47).

Data analysis

The analysis using CO1 sequences from specimens used in Xu et al. (2014) [31] supported their conclusion that China was the source population for the initial introduction of BMSB into the US (probability (p) = 0.89) (**Table 2**). When we tested our first question, which was whether there was more than one introduction event into North America (**Table 3**), we found that Canadian populations were likely also sourced from China, with

a probability of 0.76, instead of Japan or the Republic of Korea (**Table 2**). The haplotype network further supported this finding, with the Canada haplotypes found only within the China cluster of the network (**Figure 1**). When we tested scenarios with admixture to see if it was more probable that the US, or a mix of the US and China, colonized Canada rather than China alone we found that the US only colonization scenario was the only unlikely scenario of the four, with the other three scenarios very similar in probability (**Table 2**).

Tests for multiple introductions to the US indicated that Northwestern US populations (in Washington and Oregon) were also likely separate introductions from China ($p = 0.88$) (**Figure 2, Table 2**). We observed no overlap of haplotypes between the United States and Korea and a single haplotype overlap (H23) between the United States and Japan within the observed haplotype network (**Figure 1**), and found a low probability that this population was a result of admixture from these different native populations via ABC analysis (**Table 2**). Given the disparity in haplotype make-up and diversity between the western states (5 haplotypes) and eastern states (1 haplotype, **Table 2**), we decided it was unnecessary to conduct a formal test to see if the western populations of BMSB were the result of a dispersal event from the east. Instead, we focused on the seven possible scenarios for the introduction to California (**Table 2**). Of these, the scenario of California being a mix of the northwestern states and China (question 1g in Methods) was significantly different from all except the scenario of California being a mix of the eastern states and China (question 1f), with the latter not being significantly different from the remaining five (**Table 2**). When we re-ran the analysis after excluding all scenarios below a 0.10 probability (e.g. scenarios for questions 1a, 1b, and 1d), we found

the scenario for question 1g to be significantly more likely than the rest (**Table 2**, **Figure 2**). However, confidence for this result was below 0.4 (**Table 2**).

Regarding questions two and three, which were to identify the source(s) of the introduction into Europe (excluding Greece) and to assess the likelihood of a bridgehead from the US to Europe (**Table 3**), our results indicated that the European populations of BMSB examined were sourced from China with a high probability of 0.96 (**Table 2**). When we tested for the possibility of a bridgehead from either the US' eastern or western populations into Europe we found there was little support for this ($p = 0.10$ and $p = 0.13$ respectively) and that China was again the likely source, with a 0.77 probability (**Table 2**, **Figure 2**). China was also the likely source of the Greece population with a probability of 0.60 (**Table 2**, **Figure 2**), which was significantly different from a source in Korea ($p = 0.37$) or Japan ($p = 0.03$).

Next, we addressed question four, which was to determine whether there were multiple introductions into Europe and identify the sources (**Table 3**). We did so by first testing the hypothesis that the Italian Emilia-Romagna population was the result of a bridgehead from the US ($p = 0.71$) (**Figure 2**) rather than natural dispersal from Northern Italy or a separate introduction from China ($p = 0.07$ and $p = 0.22$, respectively) (**Table 2**). Second, we tested the possibilities of Greece populations being a separate introduction from only China, a mix from China and nearby Hungary, and a separate introduction from China to Greece that subsequently became the source of BMSB to Hungary. We found that Greece was likely sourced from China and then became the source of the introduction to Hungary ($p = 0.52$) (**Table 2**). This scenario was significantly more likely than an introduction from both China and Hungary to Greece ($p = 0.25$) or from only

China to Greece without spreading to Hungary ($p = 0.23$), with a confidence just over 0.5 (**Table 2**).

Discussion

We developed an analysis that included representative infestations of non-native brown marmorated stink bugs (BMSB – *Halyomorpha halys*) from across the world, as well as multiple populations from the native range, to assess the most probable worldwide invasion pathways. To accomplish this, we combined published genetic data from four studies of BMSB in non-native regions using the CO1 marker [31,38-40], added new sequences from both the introduced and native range as needed and feasible, and performed a Templeton, Crandall, and Sing (TCS) haplotype network analysis[41] and Approximate Bayesian Computation (ABC) analyses to quantify pathway probabilities. Sadly, our analyses excluded published sequences that did not match the CO1 barcoding region (e.g. Zhu et al, 2016 [42]) as well as specimens from the most recent discovered populations in Serbia, Romania, Russia, ARA, Georgia, and Bulgaria due to lack of access to specimens of BMSB from those populations. Though ABC does have its drawbacks (i.e., computational power, potential user scenario bias, and some limitations on the questions that can be effectively asked), we carefully tried to minimize bias in our study by modeling all scenarios under numerous alternative hypotheses to prevent forcing a desired outcome. Furthermore, we kept many questions at the country level due to low sample size in several parts of the native range and lack of genetic differentiation at the barcode CO1 locus. Case in point, unlike Xu et al (2014) [31] we did not attempt to identify

the specific location(s) of origin within China. Instead, the worldwide analyses provide insights into broader patterns of expansion of this economically important pest species.

Although there was a stark contrast in the genetic makeup of eastern and Western US populations (**Table 1**), this alone was not conclusive evidence of multiple introductions since the eastern population could have been a result of dispersal from the western states. We found this possibility unlikely, and our results instead indicate that the Eastern and Western US experienced separate introduction events of BMSB from China. In addition, we also found that populations in California appear to be a mix from both China and populations in the northwestern states (Washington and Oregon), indicating that at least three separate introduction events from China may have occurred to the US. However, likely due to the lack of variation at the CO1 locus, confidence in the specific scenarios underlying the expansion into California was relatively low (**Table 2**) and warrants further exploration with additional and more variable loci (e.g. microsatellites) and/or more extensive sequencing of nuclear regions.

We also found that Canada was likely sourced from a separate introduction from China. However, our analyses could not provide a single clear scenario for the complete introduction history of BMSB into Canada, since the China only scenario was not significantly different from the China and US mixed population scenarios (**Table 2**). Given the proximity of Canada's invaded range to the northern range found in the US, it is not completely surprising that this could be the case.

In contrast, our results clearly indicated that China was the source of the introduction into Switzerland that subsequently spread to neighboring European countries. We also found support for a bridgehead event from the Eastern US into Emilia-

Romagna, Italy, reinforcing the hypothesis made by Cesari et al. 2015 [40] regarding the occurrence of multiple introductions into Italy. This is a particularly interesting finding given the documented extensive economic injury to tree fruits and nuts in locations with high proportions of the H1 haplotype in the US and in Emilia-Romagna, Italy [22,43]. A high frequency of the H1 haplotype may be indicative of a phenotype either more prone to become invasive or adapted to tree fruits as a primary food resource - hypotheses that warrant further behavioral, physiological, and genomic analyses.

Greece was analyzed separately from the remainder of Europe due to its observed higher level of genetic diversity. Again, we found significant evidence of a direct introduction from China to Greece that excluded other European populations as sources. Additionally, we found that the population in Hungary was most likely due to a dispersal event from Greece rather than a continued dispersal event from Western Europe. This indicates at least three separate introduction events from China into Europe, two in Western Europe and one in Eastern Europe.

Although we answered the four primary questions (**Table 3**) our analyses are based on a single maternally inherited locus, CO1, which may underrepresent existing genetic variation. The addition of nuclear data, such as microsatellites or a NextGen based array of Single Nucleotide Polymorphisms (SNPs), should provide much needed information from both sides of the parental lineage that can be paired with the CO1 data for more in depth analyses.

While we recommend evaluating the invasion pathways of the more recent detections into other parts of Europe (i.e. Serbia, Romania, Russia, ARA, Georgia, and Bulgaria), and found evidence of one long-distance bridgehead event, most critically our

analyses indicate that the expansion of BMSB across the world is still primarily sourced from China (**Figure 2)**. This is an important result because numerous invasion events all from China, possibly even from the same location, in a relatively short amount of time, indicate the existence of an export pathway that if identified could be closed leading to major decreases in the spread of BMSB across the World. The next step then is to identify the invasion vectors, and while careful vetting of cargo potentially harboring BMSB may be onerous, it may prevent the much higher economic losses associated with this agricultural pest continuing to spread globally.

Methods

Global dataset assembly

To direct our approach and ascertain which loci had been used, we first reviewed the literature and downloaded sequence data from Genbank for the four genetic analyses of the expansion of BMSB in North America and Europe that contained both native and non-native BMSB populations [31,38-40]. The study performed in the US primarily sequenced regions of the cytochrome oxidase 2 (CO2) locus and the control region (CR), with very limited sequencing in the cytochrome oxidase 1 (CO1) [31]. In contrast, Canadian and European studies examined sections of CO1, CO2, and Cytochrome b (Cyt b) [38-40], though CO1 was the only common locus between these studies (**Table 4**). Since the most often sequenced locus was CO1 we decided to focus on it for the analysis.

We started by sequencing CO1 for many of the specimens from the US and Asia in Xu et al. 2014 [31], but since the most recent US samples in the dataset were from 2008, new specimens from across the US range were added (**Table 1**). Additionally, we

obtained samples from two new populations in Japan and one new population from the Republic of Korea in order to have a better representative sample of the native range. Newly acquired samples were received dry (with a desiccant present) or in 95%-100% molecular grade ethanol. All dry specimens were stored at -20°C while those in ethanol were stored at room temperature. To extract the genomic DNA (gDNA) we used flame sterilized tweezers to pull one leg with the underlying thoracic muscle tissue connected from each specimen [31]. We then extracted total gDNA with a DNeasy blood and tissue kit using the provided protocol (Qiagen Sciences, Germantown, MD, USA), with Proteinase K incubation overnight (minimum of 8 hours) to ensure complete digestion.

We amplified the 685bp of the CO1 standard barcode region [44] used by the European studies [38-40] with primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG -3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA -3'). Amplifications were accomplished in 20µl reactions consisting of 1× PCR buffer (10 mM Tris-HCl, pH 8.3, and 50 mM KCl), 2.5mM $MgCl_2$, 100µM of each dNTP, 200nM of each primer, 1 unit of Amplitaq Gold DNA polymerase (Applied Biosystems, Life Technologies, Carlsbad, CA) and approximately 20ng gDNA. The protocol was optimized to run at an initial denaturing temperature of 96°C for 10 minutes, followed by 45 cycles of the following steps: denaturing at 96°C for 20 seconds, annealing at 50°C for 30 seconds, and extension at 72°C for 30 seconds. Final extension was completed at 72°C for 2 minutes. All PCRs were run on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA). We visualized amplifications in a 1% agarose gel with Ethidium Bromide, and selected DNA fragments of appropriate size for sequencing. Successful

amplicons were cleaned using ExoSAP-IT (Affymetrix, OH), and mixes of 25pmoles of primer and 20ng of template DNA were sent for cycle sequencing and sizing (Genscript, Piscataway, NJ, USA). Cycle sequencing was performed using both the forward and reverse primers to create a consensus sequence and increase haplotype reliability. Chromatograms were cleaned and aligned in Sequencer 5.1 (GeneCodes, Ann Arbor, MI). All sequences were evaluated for insertions and deletions, as well as translated to amino acids to check for stop codons and limit the presence of nuclear copies [45]. We then exported the full CO1 contig in nexus format and created a TCS haplotype network using the program PopArt [46] (**Figure 1**).

ABC performance evaluation

Once we obtained a global dataset of variants of CO1 in BMSB we conducted several analyses using Approximate Bayesian Computation (ABC). Specifically, through our analyses we sought to address four major questions regarding the spread of BMSB globally (**Table 3).** These questions were chosen not only to check assumptions made in the literature regarding the invaded range in prior studies [31,38-40], but also to address hypotheses only testable with a worldwide dataset. We modeled these invasion pathway scenarios based on the initial findings of the genetic data, possible pathways hypothesized in other papers [31,38-40], and other *a priori* factors that can affect introductions (e.g. proximity and natural dispersal, bridgehead effects, etc.). We then compared appropriate scenarios against each other to quantitatively determine which had the highest likelihood of having taken place based on logistic posterior probabilities.

We selected the program DIYABC to carry out our analyses [19,20]. Before addressing our four major questions, we sought to evaluate the performance of this analysis method, based on a single mitochondrial locus, by testing the hypothesis that China was the native source population of the US introduction, as proposed by Xu et al. [31] using two loci. For parity, we limited the analysis to the same specimens analyzed at the time. To accomplish this, we developed three simple scenarios and tested them against each other: (1) US haplotypes were sourced from the China population; (2) US haplotypes were sourced from the Japan population; and (3) US haplotypes were sourced from the Republic of Korea population. All three scenarios were weighted equally, with priors set to fit a uniform distribution and the US effective population size restricted to be lower than the native populations. We determined the evolutionary model to be used via the program PartitionFinder 1.1.1 [47] in Python v2.7 under an unpartitioned whole gene model scheme with a MrBayes filter. We did so to ensure the evolutionary models selected by PartitionFinder would be limited to the model options supported by the DIYABC software. Once complete, we used the recommended Hasegawa, Kishino and Yano (HKY) model [48] and set priors as shown in **Table 5.** We set population parameters to uniform, with default settings, set the condition for the US population (N4) to be less than all other native populations (e.g. N1, N2, and N3) and ran the program for 1 million iterations.

Once the computations were complete, we performed a pre-evaluation of the scenarios and prior combinations using a principal component analysis (PCA) approach within the program. We inspected both the PCA plot and the numerical values and proceeded only when the observed dataset was within the cloud of pseudo-observed

datasets (PODs) and the numerical values displayed summary statistics for simulated data with very few values below the observed dataset (indicated by having very few or no three-star – i.e. highly significant - exceptions across most or all scenarios). We then computed the posterior probabilities for each of the scenarios to assess which was most likely to have occurred given our dataset, and evaluated confidence in scenario choice using posterior based error.

Testing invasion pathway hypotheses with ABC

After evaluating the performance of this method by comparing our initial findings against previously proposed results [31], we began developing models for our four questions regarding the global invasion of BMSB. To address question one (**Table 3**) we first examined possible native sources for the population in Canada, then tested to see if it was more likely to have come from the US, a native source, or a mix of the two (**Figure S1a**). Next, we assessed if the haplotype identity and diversity found in the western portion of the US could have been attributed to natural dispersal from the eastern portion, or if another introduction from the native range was more likely. To do so we first determined the likely native source(s) of the western populations, then tested if that native source had a higher probability of having occurred than a dispersal event from the eastern half, or if it was a combination of the two (**Figure S1b**). However, it is possible that the reverse is true, and the eastern population was the result of a dispersal event from California. Therefore, after determining the likely native source of the western states we split them into two groups, consisting of the northwestern states (Washington and Oregon) and California, and created seven scenarios (**Table 2**, **Figure S1c**): 1) California

was a separate introduction sourced from China; 2) California was sourced from a separate introduction event to the northwestern states; 3) California was sourced from the eastern states; 4) California was the site of the initial introduction site of BMSB to the US and then spread to the eastern states; 5) California is a mix of Northwest and Eastern US populations; 6) California is a mix of an introduction from China and a dispersal event from the Eastern US; and 7) California is a mix of a separate introduction event from China and a dispersal event from the Northwestern US. The above experiments had their priors set identical to our first US experiment (**Table 5**), with the non-native populations (US and Canada) effective population again restricted to being lower than all native populations, run for 1 million iterations, and scenario confidence calculated using posterior based error.

We addressed both questions two and three (**Table 3**) by first developing three simple scenarios, each testing the probability of a native population being the source of the introduction into Europe (excluding Greece), to see what the most likely native source could have been, again following the same prior parameters throughout (**Table 5**). Once the likely native source was determined we tested the possibility of a bridgehead from the US to the early introduction of Europe by comparing the most likely modeled scenarios against the native source scenario (**Figure S1d**).

We addressed question four (**Table 3**) by first testing a hypothesis from Cesari et al. [40] that the southern-most population in Italy (Emilia-Romagna) was a bridgehead event from the US, while the northern most population (Lombardi) was a dispersal from broader Europe. We tested this with two scenarios, one involving a dispersal event from Northern Italy to Southern Italy and the other a bridgehead event from the Eastern US

into Southern Italy (**Figure S1e**). Lastly, due to the very different haplotype identity and diversity detected in Greece, we tested multiple scenarios for its colonization, first by determining the most likely native source(s). Once we determined the native source(s) we created four scenarios to evaluate: 4a) the possibility of the Greek population having been a separate introduction from the native range; 4b) a dispersal event from its closest European country with an established population (Hungary); 4c) a mixture of scenarios 4a and 4b; or 4d) a separate introduction from the native range to Greece that then spread to Hungary (**Figure S1f**).

Data Accessibility

Accession numbers to all downloaded and generated sequence data analyzed are provided within a supplementary table (**Table S1**), sorted by haplotype. Sequence accession numbers are for submissions to Genbank.

Author Contributions

REV and DMF contributed to the original idea; REV, ALN, and DMF designed the research; REV performed the laboratory work; ALN, NGW, and DHL provided specimens; REV, ALN, and DMF analyzed the data; REV, ALN, DMF wrote the manuscript. All authors read and approved the manuscript.

Competing Interests

We have no competing interests.

Funding

This work was supported in part by USDA #2011-51181-30937 and the USDA National Institute of Food and Agriculture Hatch Multistate Project 1001717 through the New Jersey Agricultural Experiment Station, Hatch Multistate Project NJ08225. RV was funded by a Department of Education Graduate Fellowship (GAANN).

Acknowledgements

Celeste Welty (Ohio State University), Kristen Poley (Michigan State University), Chuck Ingels (University of California Cooperative Extension), Ted Cottrell (USDA, Byron, GA) for specimens.

References

1 Blackburn, T. M. *et al.* A proposed unified framework for biological invasions. *Trends in Ecology & Evolution* **26**, 333-339 (2011).

2 Roman, J. & Darling, J. A. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**, 454-464 (2007).

3 Lockwood, J. L., Hoopes, M. F. & Marchetti, M. P. *Invasion ecology*. (John Wiley & Sons, 2013).

4 Pimentel, D., Zuniga, R. & Morrison, D. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological economics* **52**, 273-288 (2005).

5 Hänfling, B., Carvalho, G. R. & Brandl, R. mt-DNA sequences and possible invasion pathways of the Chinese mitten crab. *Marine Ecology Progress Series* **238**, 307-310 (2002).

6 Brown, J. E. & Stepien, C. A. Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. *Molecular Ecology* **18**, 64-79 (2009).

7 Ghabooli, S. *et al.* Invasion Pathway of the Ctenophore *Mnemiopsis leidyi* in the Mediterranean Sea. *PLoS ONE* **8**, e81067, doi:10.1371/journal.pone.0081067 (2013).

8 Kaufman, M. G. & Fonseca, D. M. Invasion biology of *Aedes japonicus japonicus* (Diptera: Culicidae). *Annual review of entomology* **59**, 31-49 (2014).

9 Ascunce, M. S. *et al.* Global invasion history of the fire ant *Solenopsis invicta*. *Science* **331**, 1066-1068 (2011).

10 Leung, B. *et al.* An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London B: Biological Sciences* **269**, 2407-2413 (2002).

11 Valentin, R. E., Maslo, B., Lockwood, J. L., Pote, J. & Fonseca, D. M. Real‐time PCR assay to detect brown marmorated stink bug, *Halyomorpha halys* (Stål), in environmental DNA (eDNA). *Pest management science* (2015).

12 Simberloff, D. *et al.* Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution* **28**, 58-66 (2013).

13 Estoup, A. & Guillemaud, T. Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113-4130 (2010).

14 Lombaert, E. *et al.* Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *PLoS ONE* **5**, e9743, doi:10.1371/journal.pone.0009743 (2010).

15 Navarro, J.-C., Quintero, L., Zorrilla, A. & González, R. Molecular tracing with mitochondrial ND5 of the invasive mosquito *Aedes* (Stegomyia) *albopictus* (Skuse) in Northern South America. *J Entomol Zool Stud* **1**, 32-38 (2013).

16 Darling, J. A., Bagley, M. J., Roman, J., Tepolt, C. K. & Geller, J. B. Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Molecular Ecology* **17**, 4992-5007 (2008).

17 Costa-da-Silva, A. L. d., Capurro, M. L. & Bracco, J. E. Genetic lineages in the yellow fever mosquito *Aedes* (Stegomyia) *aegypti* (Diptera: Culicidae) from Peru. *Memórias do Instituto Oswaldo Cruz* **100**, 539-544 (2005).

18 Mousson, L. *et al.* Phylogeography of *Aedes* (Stegomyia) *aegypti* (L.) and *Aedes* (Stegomyia) *albopictus* (Skuse)(Diptera: Culicidae) based on mitochondrial DNA variations. *Genetical research* **86**, 1-11 (2005).

19 Cornuet, J.-M. *et al.* Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713-2719 (2008).

20 Cornuet, J.-M. *et al.* DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187-1189 (2014).

21 Hoebeke, E. R. & Carter, M. E. *Halyomorpha halys* (Stal)(Heteroptera: Pentatomidae): A polyphagous plant pest from asia newly detected in North America. *Proceedings of the Entomological Society of Washington* **105**, 225-237 (2003).

22 Rice, K. B. *et al.* Biology, Ecology, and Management of Brown Marmorated Stink Bug (Hemiptera: Pentatomidae). *Journal of Integrated Pest Management* **5**, A1-A13 (2014).

23 Milonas, P. & Partsinevelos, G. First report of brown marmorated stink bug *Halyomorpha halys* Stål (Hemiptera: Pentatomidae) in Greece. *EPPO Bulletin* **44**, 183-186 (2014).

24 Vetek, G., Papp, V., Haltrich, A. & Redei, D. First record of the brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Heteroptera: Pentatomidae), in Hungary, with description of the genitalia of both sexes. *Zootaxa* **3780**, 194–200 (2014).

25 Leskey, T. C. *et al.* Pest status of the brown marmorated stink bug, *Halyomorpha halys* in the USA. *Outlooks on Pest Management* **23**, 218-226 (2012).

26 Nielsen, A. L. & Hamilton, G. C. Seasonal occurrence and impact of *Halyomorpha halys* (Hemiptera: Pentatomidae) in tree fruit. *Journal of Economic Entomology* **102**, 1133-1140 (2009).

27 Leskey, T. C., Short, B. D., Butler, B. R. & Wright, S. E. Impact of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stål), in mid-Atlantic tree fruit orchards in the United States: case studies of commercial management. *Psyche: A Journal of Entomology* **2012** (2012).

28 Polk, D. F. in *Proc. 86th Cumberland-Shenandoah Fruit Workers Conference* Vol. Winchester, VA November 2010 (2011).

29 Hedstrom, C., Shearer, P., Miller, J. & Walton, V. The effects of kernel feeding by *Halyomorpha halys* (Hemiptera: Pentatomidae) on commercial hazelnuts. *Journal of economic entomology* **107**, 1858-1865 (2014).

30 Wiman, N. G., Parker, J. E., Rodriguez-Saona, C. & Walton, V. M. Characterizing damage of brown marmorated stink bug (Hemiptera: Pentatomidae) in blueberries. *Journal of economic entomology*, tov036 (2015).

31 Xu, J., Fonseca, D. M., Hamilton, G. C., Hoelmer, K. A. & Nielsen, A. L. Tracing the origin of US brown marmorated stink bugs, *Halyomorpha halys*. *Biological invasions* **16**, 153-166 (2014).

32 Macavei, L. I. *et al.* First detection of *Halyomorpha halys* Stål, a new invasive species with a high potential of damage on agricultural crops in Romania. *Lucrări Ştiinţifice, Universitatea de Ştiinţe Agricole şi Medicină Veterinară" Ion Ionescu de la Brad" Iaşi, Seria Agronomie* **58**, 105-108 (2015).

33 Šeat, J. *Halyomorpha halys* (Stål, 1855)(Heteroptera: Pentatomidae) a new invasive species in Serbia. *Acta entomologica serbica* **20**, 167-171 (2015).

34 Gapon, D. First records of the brown marmorated stink bug *Halyomorpha halys* (Stål, 1855)(Heteroptera, Pentatomidae) in Russia, Abkhazia, and Georgia. *Entomological Review* **96**, 1086-1088 (2016).

35 Simov, N. The invasive brown marmorated stink bug *Halyomorpha halys* (Stål, 1855)(Heteroptera: Pentatomidae) already in Bulgaria. *Ecologica Montenegrina* **9**, 51-53 (2016).

36 Wermelinger, B., Wyniger, D. & Forster, B. First records of an invasive bug in Europe: *Halyomorpha halys* Stål (Heteroptera: Pentatomidae), a new pest on

woody ornamentals and fruit trees? *Mitteilungen der Schweizerischen entomologischen Gesellschaft* **81**, 1-8 (2008).

37 Arnold, K. *Halyomorpha halys* (Stal, 1855), a stink bug species newly detected among the European fauna (Insecta: Heteroptera, Pentatomidae, Pentatominae, Cappaeini). *Mitteilungen des Thuringer Entomologenverbandes* **16** (2009).

38 Gariepy, T., Bruin, A., Haye, T., Milonas, P. & Vétek, G. Occurrence and genetic diversity of new populations of Halyomorpha halys in Europe. *Journal of Pest Science*, 1-10 (2015).

39 Gariepy, T., Haye, T., Fraser, H. & Zhang, J. Occurrence, genetic diversity, and potential pathways of entry of *Halyomorpha halys* in newly invaded areas of Canada and Switzerland. *Journal of pest science* **87**, 17-28 (2014).

40 Cesari, M. *et al.* A pest alien invasion in progress: potential pathways of origin of the brown marmorated stink bug *Halyomorpha halys* populations in Italy. *Journal of Pest Science* **88**, 1-7 (2015).

41 Templeton, A. R., Crandall, K. A. & Sing, C. F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619-633 (1992).

42 Zhu, G.-P. *et al.* Range wide molecular data and niche modeling revealed the Pleistocene history of a global invader (*Halyomorpha halys*). *Scientific reports* **6** (2016).

43 Bariselli, M., Bugiani, R. & Maistrello, L. Distribution and damage caused by *Halyomorpha halys* in Italy. *EPPO Bulletin* **46**, 332-334 (2016).

44 Ekrem, T., Willassen, E. & Stur, E. A comprehensive DNA sequence library is essential for identification with DNA barcodes. *Molecular phylogenetics and evolution* **43**, 530-542 (2007).

45 Egizi, A., Healy, S. P. & Fonseca, D. M. Rapid blood meal scoring in anthropophilic *Aedes albopictus* and application of PCR blocking to avoid pseudogenes. *Infection, Genetics and Evolution* **16**, 122-128 (2013).

46 Leigh, J. W. & Bryant, D. popart: full‐feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110-1116 (2015).

47 Lanfear, R., Calcott, B., Ho, S. Y. & Guindon, S. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and evolution* **29**, 1695-1701 (2012).

48 Hasegawa, M., Kishino, H. & Yano, T.-a. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160-174, doi:10.1007/bf02101694 (1985).

Table 1. Illustration of the collections from (A.) native countries and (B.) non-native countries, as well as the specific localities within those countries, of BMSB with the number of specimens for each in parenthesis. Haplotypes shown are from the CO1 locus, with the number of specimens representing each haplotype in parenthesis as well. Locations marked with an asterisk indicate a lack of specific location data per haplotype. Haplotypes in bold indicate new haplotypes found in the US post-2008. We inserted Haplotypes H24 through H29, when we realized that previous studies [38] jumped from H23 to H30.

A.

Native specimens				
Country (no. of specimens)	Locality (no. of specimens)	Haplotype (no. per haplotype)		
China (158)	Heibei/Beijing (106)*	H1 (75)	H10 (1)	H17 (1)
		H2 (1)	H11 (1)	H18 (1)
		H3 (14)	H12 (1)	H19 (1)
		H4 (1)	H13 (3)	H20 (1)
		H5 (1)	H14 (1)	H21 (1)
		H6 (1)	H15 (1)	
		H7 (1)	H16 (1)	
	Xi'an (6)	H1 (1)		
		H33 (2)		
		H45 (1)		
		H53 (2)		
	Nanjing (12)	H2 (5)	H26 (2)	
		H3 (1)	H34 (1)	
		H22 (2)	H55 (1)	
	Anhui Prov. (9)	H1 (3)		
		H2 (1)		
		H3 (4)		
		H33 (1)		
	Fuzhou (7)	H1 (7)		
	Haidain (6)	H1 (5)		
		H46 (1)		
	Hefei (7)	H22 (4)		
		H54 (3)		
	Kunming (5)	H1 (1)		

B.

Non-Native specimens		
Country (no. of specimens)	Locality (no. of specimens)	Haplotype (no. per haplotype)
United States (108)	New Jersey (14)	H1 (14)
	Maryland (14)	H1 (14)
	Georgia (6)	H1 (6)
	Delaware (3)	H1 (3)
	Massachusetts (2)	H1 (2)
	Mississipii (2)	H1 (2)
	New York (2)	H1 (2)
	Pennsylvania (6)	H1 (6)
	Virginia (6)	H1 (6)
	West Virginia (6)	H1 (6)
	Ohio (4)	H1 (4)
	Michigan (5)	H1 (5)
	California (13)	H1 (11)
		H3 (1)
		H7 (1)
	Oregon (21)	H1 (1)
		H3 (10)
		H23 (1)
		H47 (9)
	Washington (4)	**H3 (2)**
		H47 (2)
Canada (51)	Canada (51)*	H1 (49)
		H6 (1)

		H17 (4)					H14 (1)
Japan (44)	Tsubuka (16)	H23 (1)	H45 (5)		Switzerland (223)	North Switzerland (195)*	H3 (164)
		H24 (1)	H56 (1)				H8 (30)
		H27 (1)					H9 (1)
		H39 (1)				Lugano (28)	H1 (2)
		H41 (4)					H3 (25)
		H44 (2)					H8 (1)
	Yokote (13)	H27 (1)	H41 (1)	H50 (1)		Ticino (2)	H3 (2)
		H39 (2)	H42 (1)	H51 (3)			
		H40 (2)	H49 (1)	H57 (1)	Italy (40)	Emilia-Romagna (31)	H1 (31)
	Yuzawa (15)	H1 (1)	H40 (4)	H51 (5)		Lombardy (9)	H1 (1)
		H23 (1)	H43 (1)	H52 (1)			H3 (7)
		H39 (1)	H48 (1)				H8 (1)
Republic of Korea (12)	Yangpyeong (1)	H22 (1)			France (139)	Schiltigheim (139)	H1 (1)
	Suwon (4)	H2 (1)					H3 (136)
		H22 (1)					H8 (2)
		H25 (1)					
		H38 (1)			Hungary (84)	Budapest (84)	H1 (83)
	Chungcheong Province (4)	H28 (1)					H3 (1)
		H35 (1)					
		H36 (1)			Greece (57)	Athens (57)	H1 (18)
		H37 (1)					H3 (4)
	East Seoul (2)	H22 (1)					H22 (2)
		H29 (1)					H30 (1)
	Anyang (1)	H22 (1)					H31 (1)
							H32 (8)
							H33 (23)

Table 2. Probability and 95% credible interval for all Approximate Bayesian Computation scenarios used throughout the study, along with confidence in scenario choice. A downward arrow indicates a subsequent analysis with scenarios with probabilities below 0.1 excluded.

Experiment	Prob.	95% CI	Conf.
Pre-2008 NE US source determination			0.869
1: China source	0.8910	[0.7925, 0.9894]	
2: Japan source	0.0625	[0.0000, 0.1429]	
3: Korea source	0.0466	[0.0000, 0.1018]	
Canada source determination			0.796
1: China source	0.7561	[0.6415, 0.8708]	
2: Japan source	0.1150	[0.0373, 0.1926]	
3: Korea source	0.1289	[0.0452, 0.2126]	
Canada source w/ admixture			0.551
1: China only source	0.3087	[0.2592, 0.3583]	
2: US only source	0.1284	[0.0934, 0.1634]	
3: China + East US	0.3103	[0.2677, 0.3528]	
4: China + Northwest US	0.2526	[0.2078, 0.2974]	
Northwestern US source determination			0.776
1: China source	0.8841	[0.8292, 0.9389]	
2: Japan source	0.0364	[0.0133, 0.0596]	
3: Korea source	0.0795	[0.0347, 0.1243]	
Northwestern US source w/ admixture			0.509
1: China source	0.5538	[0.4574,	

		0.6502]	
2: China + Japan source	0.1262	[0.0697, 0.1828]	
3: China + Korea source	0.3200	[0.2261, 0.4138]	
Introduction to California			0.352
1: Separate introduction from China	0.0764	[0.0000, 0.1874]	
2: Dispersal from Northwestern US	0.0890	[0.0308, 0.1471]	
3: Dispersal from Eastern US	0.1071	[0.0484, 0.1657]	
4: Introduced to California then spread to Eastern US	0.0891	[0.0299, 0.1484]	
5: Mixture of Northwestern and East US	0.1345	[0.0758, 0.1933]	
6: Mixture of separate China introduction and Eastern US	0.1988	[0.1267, 0.2709]	
7: Mixture of separate China introduction and Northwestern US	0.3051	[0.2180, 0.3922]	
↓			
3: Dispersal from Eastern US	0.1258	[0.0726, 0.1791]	
5: Mixture of Northwestern and East US	0.1894	[0.1217, 0.2570]	
6: Mixture of separate China introduction and Eastern US	0.2433	[0.1537, 0.3329]	
7: Mixture of separate China introduction and Northwestern US	0.4415	[0.3345, 0.5485]	
Europe source determination (minus Greece)			0.746
1: China source	0.9570	[0.9326, 0.9814]	
2: Japan source	0.0179	[0.0047, 0.0312]	
3: Korea source	0.0251	[0.0077, 0.0426]	
Likelihood of bridgehead from US to Europe			0.758
1: China source	0.7742	[0.6404, 0.9081]	
2: Bridgehead from Eastern US	0.0980	[0.0000, 0.5300]	

3: Bridgehead from Western US	0.1278	[0.0418, 0.2137]	
Greece source determination			0.762
1: China source	0.5948	[0.5170, 0.6726]	
2: Japan source	0.0332	[0.0000, 0.1435]	
3: Korea source	0.3720	[0.2544, 0.4896]	
Bridgehead from US to Emilia-Romagna, Italy			0.779
1: Dispersal from Europe	0.0731	[0.0222, 0.1240]	
2: Bridgehead from US	0.7078	[0.6119, 0.8037]	
3: Separate introduction from China	0.2191	[0.1310, 0.3071]	
Introduction history of Greece and Hungary			0.530
1: Greece from China only	0.2265	[0.1813, 0.2716]	
2: Mixture from Hungary and China	0.2539	[0.2060, 0.3019]	
3: China source that spread to Hungary	0.5196	[0.4583, 0.5809]	

Table 3. The four principal questions being asked throughout this study. These questions shape the smaller questions that become the scenarios we model throughout the paper.

1	Were there multiple introductions into North America, and if so what were the sources?
2	What was the source population for the initial introduction (i.e. through Switzerland) to Europe?
3	Was there a bridgehead event from the United States to Europe?
4	Were there multiple introductions into Europe, and if so what were the sources?

Table 4. Mitochondrial DNA loci used in population genetic studies of brown marmorated stink bug by country, with the relevant authors conducting them on the right. The asterisk designates loci that have five or fewer samples sequenced for that respective country.

Country (no. of specimens)	Loci CO1	CO2	CR	Cty - b	References
China	X	X	X	X	Xu *et al.* 2014[31], Gariepy *et al.* 2014[39]
Japan	X*	X	X	X*	Xu *et al.* 2014[31], Gariepy *et al.* 2014[39]
Republic of Korea	X*	X	X		Xu *et al.* 2014[31], Gariepy *et al.* 2014[39]
United States	X*	X	X		Xu *et al.* 2014[31]
Canada	X			X	Gariepy *et al.* 2014[39]
Switzerland	X	X		X	Gariepy *et al.* 2014[39], Cesari et al. 2015[40]
Italy	X	X			Cesari *et al.* 2015[40]
France	X				Gariepy *et al.* 2015[38]
Hungary	X				Gariepy *et al.* 2015[38]
Greece	X				Gariepy *et al.* 2015[38]

* Designates sequences with five or few samples sequenced.

Table 5. Prior distributions used for all ABC analyses. Mutation parameters refer to selected DNA mutation model, distributions used, and bounds for said distributions within the model validation screen.

Description	Prior distribution			
Mutation Parameters				
Mutation model	HKY	10% invariant sites		Shape (2)
Mean mutation rate	Uniform	(min) 1.00E-7	(max) 1.00E-5	
Indiv. locus mutation rate	Gamma	(min) 1.00E-7	(max) 1.00E-5	Shape (2)
Mean coefficient (k C/T)	Uniform	(min) 1.5	(max) 20	
Indiv. locus coefficient (k C/T)	Gamma	(min) 1.5	(max) 20	Shape (2)

Table S1. BMSB CO1 haplotype table sorted by country, with total number of samples from each country below the name, and accession numbers to the Genbank sequences provided. Samples corresponding to specific haplotypes in bold represent new haplotypes found in the US post-2008 analysis.

	China (158)	Japan (44)	Korea (12)	East US (70)	West US (38)	Canada (51)	Switzerland (223)	Italy (40)	France (139)	Hungary (84)	Greece (57)	Accession number
H1	92	1		70	12	49	2	32	1	83	18	KF273380.1
H2	7		1									KF273381.1
H3	19				**13**		191	7	136	1	4	KF273382.2
H4	1											KF273383.2
H5	1											KF273384.1
H6	1					1						KF273385.1
H7	1				1							KF273386.2
H8							31	1	2			KF273387.1
H9							1					KF273388.1
H10	1											KF273389.1
H11	1											KF273390.1
H12	1											KF273391.1
H13	3											KF273392.1
H14	1					1						KF273393.1
H15	1											KF273394.1
H16	1											KF273395.1
H17	5											KF273396.1
H18	1											KF273397.1
H19	1											KF273398.1
H20	1											KF273399.1
H21	1											KF273400.1
H22	6		4								2	KF273401.1
H23		2			**1**							KF273402.1
H24		1										MF537219
H25			1									MF537220
H26	2											MF537221
H27		2										MF537222
H28			1									MF537223
H29			1									MF537224
H30											1	KR070749.1
H31											1	KR070748.1
H32											8	KR070750.1
H33	3										23	KR070751.1

H34	1					MF537225
H35			1			MF537226
H36			1			MF537227
H37			1			MF537228
H38			1			MF537229
H39		4				MF537230
H40		6				MF537231
H41		5				MF537232
H42		1				MF537233
H43		1				MF537234
H44		2				MF537235
H45	1	5				MF537236
H46	1					MF537237
H47					**11**	MF537238
H48		1				MF537239
H49		1				MF537240
H50		1				MF537241
H51		8				MF537242
H52		1				MF537243
H53	2					MF537244
H54	3					MF537245
H55	1					MF537246
H56		1				MF537247
H57		1				MF537248

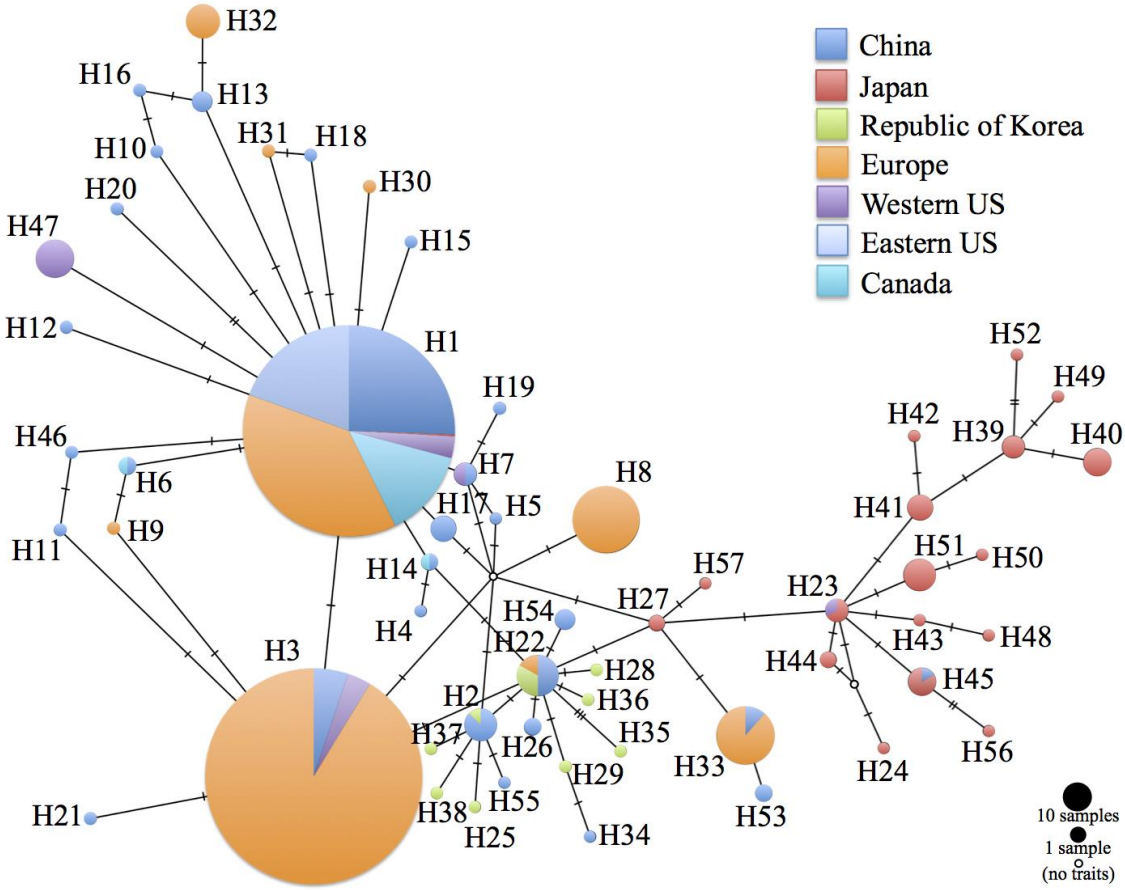

Figure 1. A CO1 haplotype network generated for BMSB, with geographic representation for each haplotype.

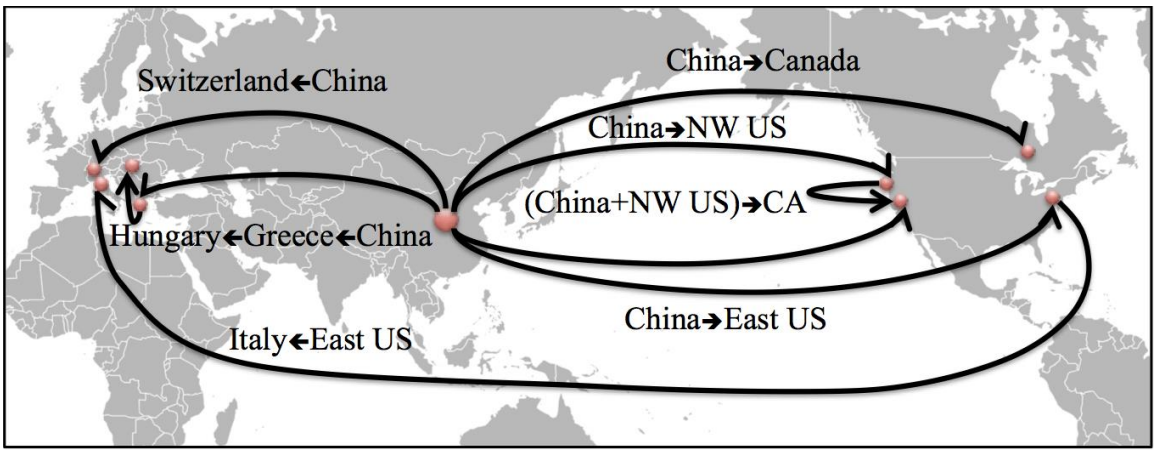

Figure 2. Map of the most likely BMSB invasion pathways connecting native and established populations across the globe, summarizing the results of our ABC analyses. Red dots on the map indicate the relevant native or established populations that are part of separate invasion pathways. The map's pathways are directional and labeled with the source population(s) listed, along with the calculated probabilities. The abbreviation NW refers to the northwestern population within the United States (US), CA refers to the US state of California, and the small arrow within the text provides additional clarity regarding the direction of the pathway. The base map, titled Blank Map Pacific World, was created by Dmthoth (https://commons.wikimedia.org/wiki/File:Blank_Map_Pacific_World.svg) and altered to show the invasion pathways.

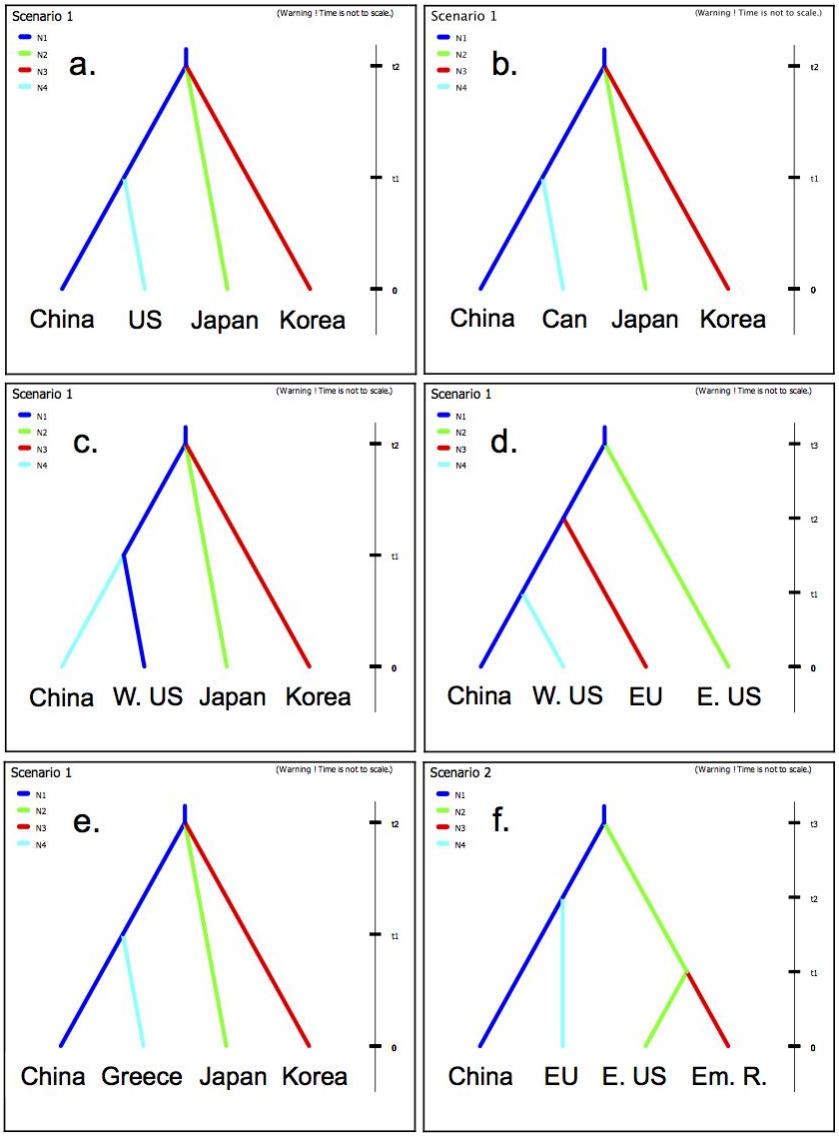

Figure 3: Approximate Bayesian Computation scenario outputs for the six significant scenarios with confidence values above 0.75: a. native source population for the pre-2008 US dataset; b. native source population for Canada; c. native source population for the Western US; d. introduction scenario for EU; e. native source population for Greece; f. bridgehead from US to Emilia-Romagna, Italy. E. US and W. US represent Eastern and Western United States, respectively. Can is an abbreviation for Canada. Em. R. is an abbreviation for Emilia-Romagna, Italy. All scenarios shown here are ordered based on experiment order in Table 2.

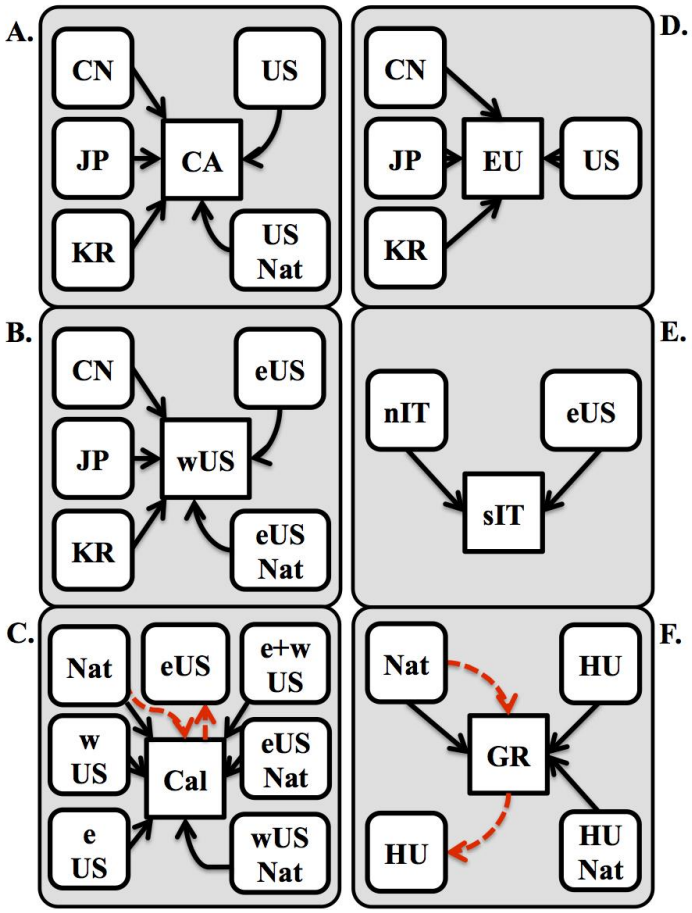

Figure S1. A visual representation of the scenarios tested for each of the questions asked while carrying out our ABC analyses. The populations in question are located at the center of each box, while the possible sources surround it. Each of the surrounding boxes represents a unique scenario that was modeled, with exception to distinct native populations (i.e. CN, JP, and KR) that also had every possible admixture option modeled as well, and the boxes connected by a dashed red arrow indicating another alternative scenario. The two letter abbreviations were used for each country, with the lowercase letters "s", "w", and "e" representing the cardinal directions of south, west, and east respectively. The abbreviation "Nat" indicates native range, meaning all scenarios for what the native source could be were tested prior to continuing with the displayed question.

Real-time PCR assay to detect brown marmorated stink bug, *Halyomorpha halys* (Stål), in environmental DNA

Rafael E. Valentin[1], Brooke Maslo[1,2], Julie L. Lockwood[1], John Pote[3], and Dina M. Fonseca[1,2,4]

[1]Department of Ecology, Evolution and Natural Resources and the Graduate Program in Ecology and Evolution, Rutgers, the State University of New Jersey, 14 College Farm Rd. New Brunswick, NJ 08901, USA

[2]Rutgers Cooperative Extension, New Jersey Agricultural Experiment Station, Rutgers, the State University of New Jersey, 88 Lipman Dr. New Brunswick, NJ 08901, USA

[3]Department of Entomology, Rutgers, The State University of New Jersey, 93 Lipman Drive, New Brunswick, NJ 08901, USA

[4]Department of Entomology, Rutgers, the State University of New Jersey, 180 Jones Ave. New Brunswick, NJ 08901, USA

Corresponding author: Dina M. Fonseca, Center for Vector Biology, 180 Jones Ave. New Brunswick, NJ 08901; dinafons@rci.rutgers.edu

Running title: eDNA tools for an invasive agricultural pest, *Halyomorpha halys*

1. ABSTRACT

BACKGROUND: Early detection before establishment and identification of key predators are time-honored strategies towards effective eradication or control of invasive species. The brown marmorated stink bug (BMSB; *Halyomorpha halys*) is a recent exotic pest of several important crops in North America and Europe. Resulting widespread applications of insecticides have countered years of careful integrated pest management and are leading to the resurgence of other agricultural pests. Environmental DNA (eDNA) has been used effectively to detect aquatic invasives. RESULTS: We developed a real-time PCR (qPCR) assay for BMSB in a conserved region of the ribosomal DNA interspacer 1 (ITS1). We validated this assay on worldwide populations of BMSB and tested its specificity and sensitivity against other US Pentatomidae species and on guano of big brown bat, *Eptesicus fuscus*, which we confirmed is a BMSB predator in NJ, USA. We also detected BMSB DNA after rapid (and inexpensive) HotSHOT DNA extractions of soiled paper from cages briefly holding BMSB, as well as from discarded *exuviae*. CONCLUSION: Due to the demonstrated high-sensitivity of our assay to BMSB environmental DNA (eDNA) in terrestrial samples this tool should become a cost-effective approach for using eDNA to detect terrestrial invasive species and their key predators.

Keywords: surveillance, invasive species, environmental DNA, big brown bat, *Eptesicus fuscus*, guano, molecular scatology, xenosurveillance.

2. INTRODUCTION

Early detection is a critical component to reducing the ill effects of invasive species. [1-3] Because newly established exotic populations often have very low and patchy abundance within the landscape,[1, 4] immediate and focused control increases the likelihood of eradication.[2, 4] A delay in the implementation of management due to a lack of early detection can mean the difference between executing an offensive eradication strategy on a small population, and a defensive control approach on an exotic population that has become invasive.[5] In addition, early detection of exotic incipient populations can prevent their spread into new habitats, thus containing them to specific locations where resources aimed toward eradication can be focused.[2] Unfortunately, the lower the abundance of a species, the harder it is to see or capture exemplary specimens.[1, 2, 4] However, every individual in a landscape releases fragments of itself (e.g. skin flakes, hair, feathers, scales, setae, *exuviae*, fecal matter, rotting bodies, etc.) containing DNA, known as environmental DNA (eDNA), which persists and may accumulate for some time, and if correctly identified will reveal its presence.[6-8] In particular, eDNA provides information on species' presence without any *a priori* assumptions about distribution, or the need to directly see or catch the target species.[1, 9-12] This sampling strategy can prove extremely beneficial for the detection of exotic species, where high sensitivity, early detection, and fast interventions are key.[2, 9, 13, 14] While this form of surveillance has been predominantly used to detect aquatic exotics,[7] here we explore the utility of eDNA for terrestrial surveillance of invasive agricultural pests.

Native to northern Asia, the brown marmorated stink bug (BMSB), *Halyomorpha halys* (Heteroptera; Pentatomidae) (Stäl, 1855), was originally detected in the United

States in Allentown, Pennsylvania in 1996.[15, 16] Since then the species has spread to many US states and has also become established in Canada and several European countries.[16-25] The BMSB is transported into novel habitats by hitch-hiking on agricultural and horticultural plants moved between markets, or in packing crates, cargo, machinery, or personal luggage,[20] and once established it can cause significant damage both to agricultural crops and ornamental plants.[15, 16, 26, 27] The economic impact of BMSB has been extensive, with reports showing a one-year loss in excess of 37 million USD to mid-Atlantic apples and 100% losses to peaches in Maryland.[28] Furthermore, insecticide applications aimed at BMSB control have upset years of carefully optimized integrated pest management strategies and reduced natural insect predators and parasitoids of other important agricultural pests.[28] Given these economic and management costs, there is a clear need to develop tools that enable the surveillance of potentially BMSB infested commodities and rapid detection of incipient BMSB populations in high-risk locations.

Current efforts aimed at detecting the presence of BMSB require direct observation or trapping with sweep-nets, black light traps, and pheromone traps followed by visual taxonomic identification.[29] As is the case for many other invasive species, surveillance through direct sampling and observation of BMSB can be costly because it requires significant investment of labor and may not be able to detect its presence until it is relatively abundant.[4, 8, 9] Furthermore, once an exotic species becomes abundant in a location, management efforts switch from surveillance and eradication to the implementation of strategies to reduce population sizes to a level below which the species imparts little economic harm.[4] This goal may be accomplished by enhancing the

populations or per capita kill rates of predators. However, current methods require direct observation to confirm a species as a BMSB predator, which often takes so much time that it limits the ability of managers to use predators as tools in controlling it. Clearly the detection of trace amounts of DNA of BMSB, in the environment or in the gut or excreta of putative predators would be useful. However, we must first have a reliable and specific method for identifying trace amounts of often very degraded DNA in terrestrial settings. Thus, our primary objective was to provide the necessary tool to develop such an approach.

We designed and optimized a real-time PCR (qPCR) assay consisting of primers and a probe matching sequences within the internal transcribed spacer 1 (ITS1), an intron region of the ribosomal DNA (rDNA). rDNA is found in multiple copies in each cell and the exons have relatively low intraspecific variability, but in contrast interspecific differences in the intronic interspacers can be highly pronounced often involving large insertions/deletions and/or sequence differences.[30, 31] This combination of characteristics allows the development of stable species-specific rapid assays.[30, 32] We tested our assay in twelve related pentatomid species native to North America and in multiple specimens of BMSB both from invasive (USA) and native populations (China, Korea and Japan). We know from previous and ongoing studies, such as Xu *et al.* (2013)[33] and Valentin RE (unpublished data), that these specimens exhibit significant genetic variability and we reasoned that by testing specimens from across the native range we would cover most of the sequence variation likely to occur in the invasive range. Further, to assess the performance of the assay in degraded and mixed environmental samples we tested BMSB

exuviae, BMSB fecal/urine samples, and guano of the big brown bat, *Eptesicus fuscus*, an insectivore we confirmed as a BMSB predator.

3. EXPERIMENTAL METHODS

3.1. Real-time PCR assay design and optimization

3 . 1 . 1. Sample collection

We obtained 10 specimens of BMSB, consisting of two individuals chosen from colonies kept by the Hamilton lab at Rutgers University that are derived from a mix of specimens from across the Eastern US, six specimens collected from houses in New Jersey (NJ), USA and two other NJ specimens from a prior study.[33] To obtain genomic DNA (gDNA) we used flame sterilized tweezers to dig into the BMSB thorax and pull one leg with the underlying muscle tissue connected.[33] We then extracted total gDNA with a DNeasy blood and tissue kit (Qiagen Sciences, Germantown, MD, USA) using standard protocols.

3 . 1 . 2. Amplification and sequencing of BMSB ITS1

Because there were no sequences of ITS1 for BMSB available to us we amplified and sequenced the entire BMSB ITS1 region, using the universal invertebrate primer sets ITS1A/ITS1B[34] and BD1/4S[35, 36] (**Table 1**) which match sequences in the flanking exons (18S and 5.8S). Amplifications were accomplished in 20µl reactions consisting of 1× PCR buffer (10 mM Tris-HCl, pH 8.3, and 50 mM KCl), 2.5mM $MgCl_2$, 150µM of each dNTP, 200nM of each primer, 1 unit of Amplitaq Gold DNA polymerase (Applied Biosystems, Life Technologies, Carlsbad, CA) and approximately 20ng gDNA. We optimized the PCR protocol to run as a four stage touchdown, with an initial denaturing temperature of 96°C for 10 min to activate the AmplitaqGold. The touchdown stage used a five cycle protocol of 30 s at 96°C, 30 s at the annealing temperature, and 45 s at 72°C

for extension. The touchdown annealing temperature began at 57°C, with a 1°C decrease in temperature per cycle concluding at 53°C. The following stage consisted of 50 cycles of the following steps: denaturing at 96°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 45 s. Final extension was completed at 72°C for 5 min. We ran all PCRs on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA). PCR reactions visualized in a 1% agarose gel with ethidium bromide yielded amplicons either ~500bp or ~600bp in size. Successful amplicons were cleaned using ExoSAP-IT (Affymetrix, OH), and mixes of 25pmoles of primer and 20ng of template DNA were sent for cycle sequencing and sizing (Genscript, Piscataway, NJ, USA). Sequences were obtained in both directions in order to have a consensus of the full ITS1 sequence after chromatograms were cleaned and aligned in Sequencer 5.1 (GeneCodes, Ann Arbor, MI).

3 . 1 . 3. Development and Optimization of the BMITS1 assay

Despite some sequence variation, upon examination of the sequences (Accession# KU594609-KU594612), we chose 211 bp in the ITS1 region that were clearly distinct from those of another Pentatomidae obtained from GenBank (accession#: AB725684) but conserved across all ten BMSB ITS1 sequences obtained (**Figure 1**). We used Primer Express v3 (Applied Biosystems, Life Technologies, Carlsbad, CA) to design qPCR primers and a probe. We selected the primers and probe based on a strict melting temperature criteria of 60°C for the primers, and 70°C for the probe. Additionally, following Applied Biosystems recommended guidelines (Publication 117PB24-01), we further narrowed down the list of potential primer/probe candidates by selecting sequences that contained minimal disruptions to the probe itself (e.g. probe hairpins and

dimers with the associated primers) that would negatively affect its performance while in solution.

We assessed the limits of detectable concentrations of DNA by diluting a BMSB DNA extraction initially of 10ng/µl, measured using a Nanodrop 2000 (Nanodrop, Wilmington, DE), into a dilution series that was subsequently used as a concentration standard. We created the dilution series via six successive 1:10 dilutions with final concentrations ranging from 1ng/µl to 10fg/µl of BMSB DNA. We carried out the qPCR amplifications in replicates of three in 20µl reactions, with 500nM of each primer, 250nM of the probe, 1x TaqMan® Universal Master Mix II with no UNG, and 1-2µl of DNA ranging in concentration from 2ng to 10fg of gDNA. The optimized reaction protocol was conducted at an initial denaturing step of 96°C for 10 min, followed by 45 cycles of denaturing for 15 s and annealing and extension at 60°C for 1 min. All reactions were run on an Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Life Technologies, Carlsbad, CA).

To validate the BMITS1 assay across a broad range of BMSB populations we used the assay on 12 specimens from the native range: Hefei, China (1); Nanjing, China (1); road between Dangshan and Xuzhou, China (2); Fuzhou, China (1); Kunming, China (2); Xi'an, China (1); Gyeonggi Province, South Korea (2); and Tsubuka, Japan (2). We performed the tests in 20µl reactions with 500nM of each primer, 250nM of the probe, 1x TaqMan® Universal Master Mix II with no UNG, and 2µl of DNA, following the same optimized protocol.

3. 2. Detection of BMSB in environmental DNA

In an effort to both identify BMSB eDNA sources and determine the developed BMITS1 assay's performance on such material, we tested two potential sources of BMSB eDNA: BMSB nymphal *exuviae* and BMSB excreta. Three BMSB *exuviae* at least a week old were collected from the Hamilton colony. In addition, a segment of the paper covering the floor of the colony containing evidence of defecation from BMSB was torn off with clean gloves and placed in a glass vial for processing. The heads of the *exuviae* were removed with a razorblade that was flame sterilized beforehand and placed using flame sterilized tweezers into labeled 0.2 ml eppendorf tubes. Three locations on the paper sample that contained a single defecation spot were torn off using flame sterilized tweezers and a razorblade and placed in separate labeled 0.2 ml eppendorf tubes. All samples were then extracted via a HotSHOT extraction protocol [37]. Briefly, HotSHOT extractions are a very rapid and cost effective method where tissue is placed in 50µl of an alkaline lysis reagent (25 mM NaOH, 0.2 mM EDTA, pH = 12.0), heated to 96°C for 1 h, cooled to 4°C, then mixed with 50µl of a neutralization buffer (40 mM Tris–HCl, pH = 5.0), all at a cost of less than 0.08 USD per sample, mostly in plastic supplies.[37] HotSHOT extracted gDNAs, along with two extraction negative controls, were tested with the developed BMITS1 assay using the optimized protocol.

3. 3. Detection of BMSB from predator-derived sources

To test the performance of the developed BMITS1 assay in degraded environmental samples we used excrement from a generalist insectivore, the big brown bat (*Eptesicus fuscus*) as a blind sample. We reasoned that by testing guano samples we would assess the efficacy and sensitivity of the BMITS1 assay in degraded DNA

containing multiple invertebrate species. We obtained guano from a plastic sheet placed on the ground at man-made big brown bat summer roosting sites. Bats were not approached or handled during collection activities and therefore IACUC oversight was not required. Indeed, because their roosts were inside bat houses or high in building's rafters, the bats were not disturbed. The guano was allowed to accumulate for one week and then samples of approximately 100 pellets were taken and brought to the laboratory where they were placed at -20°C. To extract the DNA from guano we first used a Qiagen, TissueLyser (Retsch MM301, Qiagen Sciences, Germantown, MD, USA) to homogenize 16-40 pellets from 15 separate collection sites and dates, thus representing the digestive products of several individual bats. For more details pertaining to the pooled samples, their collection sites, and collection methods please refer to Maslo *et al.* (in review).

To examine the effect of the amount of starting material (by weight) on the likelihood of finding BMSB and different invertebrate species in the guano, samples one through four (**Table 3)** were all derived from one pooled homogenate by preparing increasing amounts of material for extraction (**Table 3**). This procedure was repeated in samples five through eight obtained from a different pooled homogenate collected from a different bat roosting site at a different date (**Table 3**). To explore the possibility of sequencing individual pellets directly without cloning,[38] we also randomly selected three guano pellets (samples nine through eleven, **Table 3**) and extracted them individually. We extracted DNA from both pooled and single pellets with a QiaAmp fast DNA stool mini kit (Qiagen Sciences, Germantown, MD, USA) and re-suspended the DNA in 200µl of ATE buffer. We took specific eDNA precautions[6-8] for extraction of DNA from the

guano by using a dedicated DNA extraction laboratory away from the main lab, with water extractions as negative controls.

To confirm the presence of multiple insect species within the guano, we used cytochrome c oxidase 1 (CO1) primers LEP-F1, LEP-R1, and LEP-R2[39] (**Table 1**) to amplify that locus from the DNA in guano samples (**Table 3**). The LEP-R1 primer produces a ~650bp amplicon, while the LEP-R2 primer produces a smaller ~350bp amplicon that presumably works better with more degraded DNA.[39] The CO1 locus was chosen due to the extensive amount of data available from the barcoding initiative,[40, 41] with multiple sequences from many insect species representing all orders and many families. We used a standard three stage PCR in 20µl reactions consisting of 1× PCR buffer, 2.5mM $MgCl_2$, 150µM of each dNTP, 200nM of each primer, 1 unit of Amplitaq Gold DNA polymerase (Applied Biosystems, Life Technologies, Carlsbad, CA) and 1µl gDNA. The optimized protocol ran with an initial denaturing temperature of 96°C for 10 min to activate the AmplitaqGold. The protocol then proceeded for 50 cycles as follows: denaturing at 96°C for 45 s, annealing at 45°C for 30 s, and extension at 72°C for 30 s. Final extension was completed at 72°C for 3 min. All PCRs were run on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA).

In order to identify the species present in the guano, we cloned and sequenced CO1 amplicons from samples one through eight using TOPO TA cloning kits under standard guidelines (Invitrogen, Life Technologies, Carlsbad, CA). We picked 40 colonies for each amplicon size: the larger LEP-R1 fragment size and the smaller LEP-R2 fragment size. Colonies were re-amplified using the same primers used in the original amplifications and cleaned with ExoSAP-IT (Affymetrix, Santa Carla, CA, USA).

Additionally, we cleaned PCR products from the extractions of single pellets (samples 9-11) with ExoSAP-IT and prepped them for sequencing without cloning. Mixes of 25 pmoles of the LEP-F1 primer and 20ng of cleaned template DNA were sent for cycle sequencing (Genscript, Piscataway, NJ, USA). Sequences were assessed and cleaned in Sequencer 5.1 (GeneCodes, Ann Arbor, MI, USA) and compared to sequences on the Barcode of Life Data Systems (BOLD).[41] All species classifications were made to a 98% or higher sequence similarity criteria.

Once we had established the presence of multiple insect species in the guano, and the amount of starting material needed, we used the BMITS1 assay to screen the gDNA from the 11 guano samples, as well as 13 additional samples of big brown bat guano from New Jersey. For the latter tests, gDNA was extracted from homogenates of approximately 16 guano pellets (please refer to Maslo et al *in review* for more details). All qPCR experiments were performed in replicate using the optimized protocol and were repeated if the two replicates did not conform.

3. 4. Species specificity tests

We used the presence of multiple insect species in the guano to also assess the specificity of the BMITS1 assay to only BMSB DNA. To accomplish this, we chose seven guano DNA extractions from pooled pellets that were found positive for BMSB, amplified the small DNA fragment between the developed assay's primers by standard PCR, and sequenced directly (i.e. without cloning). We performed all amplifications in 20µl reactions consisting of 1× PCR buffer, 2.5mM $MgCl_2$, 100µM of each dNTP, 250nM of each primer, 1 unit of Amplitaq Gold DNA Polymerase, and approximately 2µl of gDNA from the extracted guano containing a mix of bat DNA and multiple

insects. The protocol was run at an initial denaturing temperature of 96°C for 10 min, followed by 40 cycles of the following steps: denaturing at 96°C for 20 s, annealing at 60°C for 30 s, extension at 72°C for 30 s and a final extension at 72°C for 1 min. We assessed amplifications of the small DNA fragment between the assay's two primers in a 1% agarose gel with ethidium bromide visualized under UV-light. Positive amplifications were cleaned using ExoSAP-IT (Affymetrix, OH) then prepared for cycle sequencing (Genscript, Piscataway, NJ, USA) with 25pmols of the forward primer and 20ng of the cleaned template DNA. Sequences were cleaned and aligned using Sequencer 5.1 (GeneCodes, Ann Arbor, MI) to check for polymorphisms. A messy or polymorphic sequence would indicate that there was likely a non-specific amplification, thus indicating that the BMSB primers were non-specific.

To further test the specificity of the assay we ran it on twelve other species within the same insect family as BMSB (Pentatomidae, **Table 2**) that are common in the northeastern USA. Because we had already validated this DNA extraction method we extracted gDNA from individual specimens of the 12 species, using the HotSHOT protocol. All qPCR tests were performed using the optimized reaction protocol with a series of BMSB standards (acting also as positive controls) plus extraction and PCR negative controls.

To ensure that negative results from the test on the related pentatomids were not due to low quality DNA, from the same HotSHOT extractions (of the 12 species) we also amplified a 350 bp section of the CO1 mtDNA locus with the LepF1 and LepR2 primers. The PCR protocol for this test was conducted as a standard three stage PCR in 20µl reactions consisting of 1× PCR buffer, 2.5mM $MgCl_2$, 150µM of each dNTP, 200nM of

each primer, 1 unit of Amplitaq Gold DNA polymerase (Applied Biosystems, Life Technologies, Carlsbad, CA) and 1µl gDNA of specimens of each species. The optimized protocol ran with an initial denaturing temperature of 94°C for 10 min to activate the AmplitaqGold, then proceeded for 50 cycles as follows: denaturing at 94°C for 45 s, annealing at 45°C for 30 s, and extension at 72°C for 45 s. Final extension was completed at 72°C for 3 min. All PCRs were run on a Veriti 96-Well Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA).

4. RESULTS

4. 1. BMITS1 assay performance, validation, and specificity

We designed primers BMITS1F (5'- CGAGGCCGCCGATGA-3') and BMITS1R (5'-CCCACGAGCCGAGTGATC-3') (**Figure 1**) from the rDNA ITS1 region. The primers were purchased HPLC cleaned because this effectively increases the melting temperature (T_m) for shorter sequences, allowing for an overall shorter amplicon while remaining within temperature requirements (i.e., 10°C below the temperature of the probe, or 60°C). The probe BMITS1TM (5'-CAGGCAATGAAGCACA-3') is just 4 base pairs away from the BMITS1F primer (**Figure 1**) in accordance with guidelines for optimal probe performance during amplification (Applied Biosystems Publication 117PB24-01). The BMITS1TM probe was designed with a 5' VIC reporter dye and an MGB moiety attached to the 3' non-fluorescent quencher. The MGB moiety results in more stable probe hybridization and increases T_m allowing for a shorter probe while remaining within temperature guidelines (70°C). The fragment amplified by the two BMITS1 qPCR primers is just 96bp long.

Serial dilutions of BMSB DNA ranging from the 2ng to 20fg levels of DNA returned positive results, with the 10fg dilution never returning a positive result. All three replicates of the serial dilutions tested yielded C_T values nearly identical, varying by < 0.05 points with a standard curve of -3.38 (97.6% efficient) and an R^2 of 0.99. Tests of BMSB specimens from the native range as well as from the USA were all positive, validating the assay.

The attempt to amplify DNA of specimens from 12 additional species of Pentatomidae (**Table 2**) with the BMITS1 assay resulted in no amplifications aside from the BMSB positive controls. The presence of usable DNA from those Pentatomidae species was confirmed since we obtained the appropriate amplicon size with the CO1 primers. We also confirmed specificity when we tested the seven guano DNA samples found positive for BMSB using a standard PCR with only the BMITS1 qPCR primers. Only five of the seven samples were able to produce a DNA fragment visible in an agarose gel likely due to very low initial BMSB DNA concentration. However, in all five that produced a visible DNA fragment, direct sequencing resulted in sequences that aligned easily, contained no polymorphisms, and were identical to the target ITS1 sequence from BMSB. Thus we were able to confirm that, at least in US BMSB, the primers themselves are specific to BMSB although the probe further increases specificity due to its annealing stringency. In addition, a GenBank BLAST search with the BMITS1 assay's primer and probe sequences resulted in no hits, indicating a lack of comparable sequences within the vast database.

4. 2. BMSB detection in environmental DNA

Despite the likely considerable degradation of DNA in BMSB *exuviae* and fecal matter, the BMITS1 qPCR assay successfully detected BMSB DNA in both cases. Furthermore, the BMITS1 qPCR assay also identified BMSB DNA in 16 out of the 24 guano samples tested (**Table 3**). Of note, samples one through four that were derived from the same stock were all positive, while samples five through eight derived from a different stock were all negative indicating no cross-contamination. We also confirmed the guano contained numerous insect species since we found evidence of 17-18 different insect species from 12 different taxonomic families **(Table S1)**, via Sanger sequencing after cloning. These were not species being studied within the laboratory, supporting the absence of contamination. Of note, direct sequencing of amplifications of the CO1 locus from gDNA extracted from single pellets resulted in clean sequences in two out of the three pellets tested (**Table 3**). However, none of the sequences recovered matched BMSB CO1 sequences in GenBank, indicating the BMSB DNA present, which was detected with the BMITS1 qPCR assay (**Table 3**), may have been considerably degraded. It is also possible that the CO1 regions of the BMSB DNA were not amplified due to different affinities and competition for the primers among sequences from the different species present[42].

5. DISCUSSION

We designed and validated a real-time PCR assay (BMITS1 qPCR) that is specific and highly sensitive to brown marmorated stink bug DNA even when the DNA is significantly degraded. As a way to quickly test and validate the BMITS1 qPCR assay on degraded environmental DNA (eDNA) of a large number of insect species, we tested

DNA from bat guano. This strategy combines two of the conditions for assay validation outlined by Bohmann *et al.* (2014)[6]: specificity and sensitivity to small DNA fragments. The presence of multiple insect species was expected from previous studies of bat diet,[43-45] including big brown bats,[46] and we confirmed the presence of many species of insects in the guano of the big brown bats we surveyed here. Although our sampling was very limited, we also confirmed the possibility of obtaining clean sequences after direct sequencing of cytochrome c oxidase 1 (CO1) amplifications from single pellets. However, our optimized BMITS1 qPCR assay showed those pellets also contained BMSB DNA (**Table** 3), underscoring that a species-specific and highly sensitive test such as probe-based qPCR is most appropriate for our objective of targeting eDNA for surveillance.

We found that our BMITS1 qPCR assay is sensitive to even very degraded eDNA from BMSB, which allowed us to conclude that native big brown bats are consuming invasive BMSB in New Jersey farms (Maslo *et al.* in review). We found that if used across multiple unknown and degraded DNA samples, the BMITS1 assay can be both time and cost effective in surveying for BMSB. This advantage persists despite the added expense of the probe and the need to use a qPCR machine, especially when used in conjunction with the inexpensive HotSHOT DNA extractions that we validated for qPCR. However, perhaps the greatest advantage of utilizing qPCR for surveillance comes from the speed with which results can be obtained. Because qPCR relies on fluorescence detection there is no need to sequence and so results are obtained often within minutes. While Next Generation sequencing can provide a similar level of detection sensitivity and

extraordinary specificity[47], data acquisition and analyses commonly require several days or even weeks.

We propose that the BMITS1 qPCR assay can be a part of rapid-response comprehensive eDNA based surveillance for BMSB. Though most current uses of eDNA practices have focused on aquatic surveillance, researchers have shown that sufficient DNA can remain for long periods in particulates and sediments in terrestrial settings. [6, 7] Here we show that degraded *exuviae* and fecal remnants are positive for BMSB DNA, indicating that our assay will detect the presence of small traces of BMSB eDNA. If used strategically, the positive occurrence of BMSB eDNA within agricultural commodities or other transportation vectors can provide actionable information for quarantine efforts, and may act as a means of quality control to prevent further spread of this invasive pest species. Our assay should also prove useful for surveying locations still devoid of BMSB but likely to become invaded (i.e. just outside the current distribution). In this context, surveillance allows the detection of incipient BMSB populations that can then be targeted for quick and relatively low cost eradication. Rapid detection of such incipient populations will be extremely beneficial since it allows managers to more effectively allocate resources for further monitoring and control efforts. It can also inform focused treatments with pesticides rather than repeated broad applications, minimizing high collateral damage and preserving effective integrated pest management programs for other agricultural pests.

The ability to detect small amounts of degraded BMSB DNA can also be helpful in efforts to control well-established exotic populations of this species. We have shown that BMSB DNA can be detected in the excrement of big brown bats, indicating these

insectivorous mammals, often common on farms, may act as important predators. Although we do not yet know the extent to which bats prey on BMSB, and thus whether the bats can provide a means of population control, they may act as additional surveillance tools since they consume prey across a broad area and accumulate guano centrally, facilitating sampling (Maslo *et al.* in review). This concept is not exclusive to bats, as it may be applicable to other BMSB predators that may act as a natural surveillance tool (e.g., birds, amphibians, even predatory insects).

6. CONCLUSIONS

This assay enables relatively inexpensive, high-throughput, sensitive and specific testing of eDNA from multiple terrestrial sources for the presence of brown marmorated stink bug DNA. We propose that this tool should be incorporated into current efforts to identify the spread and manage the damage due to invasive BMSB in the US and Europe. Although other studies have developed qPCR tests for pest insect identification,[48, 49] to the best of our knowledge this constitutes a first attempt at using eDNA to detect invasive species in a terrestrial setting. Further studies should develop best practices for sampling and testing using this assay in order to optimize reproducibility and sensitivity.

7. ACKNOWLEDGMENTS

We would also like to thank George Hamilton for providing specimens, and Anne Nielsen, Dave Schmitt, and Dean Polk for useful technical guidance and support. We acknowledge Karen Leu, Kathleen Kerwin, Amanda Bevan, Danielle Podmayersky, and Brittany Chrans for their efforts in collection and preparation of guano samples. Finally,

we thank three anonymous reviewers for their detailed suggestions and comments that have significantly improved this final version. Funding for this project was provided by a USDA-NRCS Conservation Innovation Grant (69-2B29-12-176) and the Rutgers Cooperative Extension Program. RV was funded by a US Department of Education Graduate Fellowship (GAANN).

8. REFERENCES

1. Berry O, Sarre SD, Farrington L and Aitken N, Faecal DNA detection of invasive species: the case of feral foxes in Tasmania. *Wildlife Research*; **34**(1): 1-7 DOI Electronic Resource Number (2007).

2. Jerde CL, Mahon AR, Chadderton WL and Lodge DM, "Sight‐unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*; **4**(2): 150-157 DOI Electronic Resource Number (2011).

3. Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E and Pascal M, Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*; **28**(1): 58-66 DOI Electronic Resource Number (2013).

4. Harvey CT, Qureshi SA and MacIsaac HJ, Detection of a colonizing, aquatic, non‐indigenous species. *Diversity and Distributions*; **15**(3): 429-437 DOI Electronic Resource Number (2009).

5. Rejmánek M, Invasive plants: approaches and predictions. *Austral ecology*; **25**(5): 497-506 DOI Electronic Resource Number (2000).

6. Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW and de Bruyn M, Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution*; **29**(6): 358-367 DOI Electronic Resource Number (2014).

7. Herder JE, Valentini A, Bellemain E, Dejean T, Delft JJCWv, Thomsen PF and Taberlet P. Environmental DNA - a review of the possible applications for the detection of (invasive) species, ed. by Stichting RAVON N (2014).

8. Rees HC, Maddison BC, Middleditch DJ, Patmore JR and Gough KC, REVIEW: The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*; **51**(5): 1450-1459 DOI Electronic Resource Number (2014).

9. Dejean T, Valentini A, Miquel C, Taberlet P, Bellemain E and Miaud C, Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. *Journal of applied ecology*; **49**(4): 953-959 DOI Electronic Resource Number (2012).

10. Foote AD, Thomsen PF, Sveegaard S, Wahlberg M, Kielgast J, Kyhn LA, Salling AB, Galatius A, Orlando L and Gilbert MTP, Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PloS one*; **7**(8): e41781 DOI Electronic Resource Number (2012).

11. Thomsen P, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, Orlando L and Willerslev E, Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*; **21**(11): 2565-2573 DOI Electronic Resource Number (2012).

12. Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M and Willerslev E, Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS one*; **7**(8): e41732 DOI Electronic Resource Number (2012).

13. Mahon AR, Jerde CL, Galaska M, Bergner JL, Chadderton WL, Lodge DM, Hunter ME and Nico LG, Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. *PloS one*; **8**(3): e58316 DOI Electronic Resource Number (2013).

14. Piaggio AJ, Engeman RM, Hopken MW, Humphrey JS, Keacher KL, Bruce WE and Avery ML, Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular ecology resources*; **14**(2): 374-380 DOI Electronic Resource Number (2014).

15. Nielsen AL and Hamilton GC, Seasonal occurrence and impact of *Halyomorpha halys* (Hemiptera: Pentatomidae) in tree fruit. *Journal of economic entomology*; **102**(3): 1133-1140 DOI Electronic Resource Number (2009).

16. Rice KB, Bergh CJ, Bergmann EJ, Biddinger DJ, Dieckhoff C, Dively G, Fraser H, Gariepy T, Hamilton G and Haye T, Biology, Ecology, and Management of Brown Marmorated Stink Bug (Hemiptera: Pentatomidae). *Journal of Integrated Pest Management*; **5**(3): A1-A13 DOI Electronic Resource Number (2014).

17. Arnold K, *Halyomorpha halys* (Stal, 1855), a stink bug species newly detected among the European fauna (Insecta: Heteroptera, Pentatomidae,

Pentatominae, Cappaeini). *Mitteilungen des Thuringer Entomologenverbandes*; **16**(10)2009).

18. Callot H and Brua C, *Halyomorpha halys* (Stål, 1855), the marmorated stink bug, new species for the fauna of France (Heteroptera Pentatomidae). *LÕEntomologiste*; **69**(69-71 DOI Electronic Resource Number (2013).

19. Cesari M, Maistrello L, Ganzerli F, Dioli P, Rebecchi L and Guidetti R, A pest alien invasion in progress: potential pathways of origin of the brown marmorated stink bug *Halyomorpha halys* populations in Italy. *Journal of Pest Science*; **88**(1): 1-7 DOI Electronic Resource Number (2015).

20. Gariepy T, Bruin A, Haye T, Milonas P and Vétek G, Occurrence and genetic diversity of new populations of *Halyomorpha halys* in Europe. *Journal of Pest Science*: 1-10 DOI Electronic Resource Number (2015).

21. Gariepy T, Fraser H and Scott-Dupree C, Brown marmorated stink bug (Hemiptera: Pentatomidae) in Canada: recent establishment, occurrence, and pest status in Southern Ontario. *The Canadian Entomologist*: 1-4 DOI Electronic Resource Number (2014).

22. Heckmann R, First evidence of *Halyomorpha halys* (Stål, 1855) (Heteroptera: Pentatomidae) in Germany. *Heteropteron H*; **36**(17-18 DOI Electronic Resource Number (2012).

23. Maistrello L, Dioli P, Vaccari G, Nannini R, Bortolotti P, Caruso S, Costi E, Montermini A, Casoli L and Bariselli M, First records in Italy of the Asian stinkbug *Halymorpha Halys*, a new threat for fruit crops. *ATTI Giornate Fitopatologiche*; **1**(283-288 DOI Electronic Resource Number (2014).

24. Vetek G, Papp V, Haltrich A and Redei D, First record of the brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Heteroptera: Pentatomidae), in Hungary, with description of the genitalia of both sexes. *Zootaxa*; **3780**(1): 194–200 DOI Electronic Resource Number (2014).

25. Wermelinger B, Wyniger D and Forster B, First records of an invasive bug in Europe: *Halyomorpha halys* Stål (Heteroptera: Pentatomidae), a new pest on woody ornamentals and fruit trees? *Mitteilungen der Schweizerischen entomologischen Gesellschaft*; **81**(1-8 DOI Electronic Resource Number (2008).

26. Hamilton G, Shearer P and Nielsen A, Brown marmorated stink bug: a new exotic insect in New Jersey. *FS002 Rutgers University Cooperative Extension*: 2 DOI Electronic Resource Number (2008).

27. Leskey TC, Short BD, Butler BR and Wright SE, Impact of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stål), in mid-Atlantic tree fruit orchards in the United States: case studies of commercial management. *Psyche: A Journal of Entomology*; **2012**(2012).

28. Leskey TC, Hamilton GC, Nielsen AL, Polk DF, Rodriguez-Saona C, Bergh JC, Herbert DA, Kuhar TP, Pfeiffer D and Dively GP, Pest status of the brown marmorated stink bug, *Halyomorpha halys* in the USA. *Outlooks on Pest Management*; **23**(5): 218-226 DOI Electronic Resource Number (2012).

29. Khrimian A, Shearer PW, Zhang A, Hamilton GC and Aldrich JR, Field trapping of the invasive brown marmorated stink bug, *Halyomorpha halys*, with geometric isomers of methyl 2, 4, 6-decatrienoate. *Journal of agricultural and food chemistry*; **56**(1): 197-203 DOI Electronic Resource Number (2007).

30. Collins F and Paskewitz S, A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species. *Insect molecular biology*; **5**(1): 1-9 DOI Electronic Resource Number (1996).

31. Hill LA, Davis JB, Hapgood G, Whelan PI, Smith GA, Ritchie SA, Cooper RD and van den Hurk AF, Rapid identification of *Aedes albopictus, Aedes scutellaris, and Aedes aegypti* life stages using real-time polymerase chain reaction assays. *The American journal of tropical medicine and hygiene*; **79**(6): 866-875 DOI Electronic Resource Number (2008).

32. Crabtree M, Savage H and Miller B, 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**(1): 105-109 DOI Electronic Resource Number (1995).

33. Xu J, Fonseca DM, Hamilton GC, Hoelmer KA and Nielsen AL, Tracing the origin of US brown marmorated stink bugs, *Halyomorpha halys*. *Biological invasions*; **16**(1): 153-166 DOI Electronic Resource Number (2014).

34. Beebe NW, Foley DH and Ellis JT, Populations of the south‐west Pacific malaria vector *Anopheles farauti ss* revealed by ribosomal DNA transcribed spacer polymorphisms. *Heredity*; **84**(2): 244-253 DOI Electronic Resource Number (2000).

35. Remigio E and Blair D, Relationships among problematic North American stagnicoline snails (Pulmonata: Lymnaeidae) reinvestigated using nuclear ribosomal DNA internal transcribed spacer sequences. *Canadian journal of zoology*; **75**(9): 1540-1545 DOI Electronic Resource Number (1997).

36. von der Schulenburg JHG, Hancock JM, Pagnamenta A, Sloggett JJ, Majerus ME and Hurst GD, Extreme length and length variation in the first ribosomal internal transcribed spacer of ladybird beetles (Coleoptera: Coccinellidae). *Molecular Biology and Evolution*; **18**(4): 648-660 DOI Electronic Resource Number (2001).

37. Johnson BJ, Robson MG and Fonseca DM, 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* 2015).

38. Dodd LE, Chapman EG, Harwood JD, Lacki MJ and Rieske LK, Identification of prey of *Myotis septentrionalis* using DNA-based techniques. *Journal of Mammalogy*; **93**(4): 1119-1128 DOI Electronic Resource Number (2012).

39. Hebert PD, Penton EH, Burns JM, Janzen DH and Hallwachs W, Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*; **101**(41): 14812-14817 DOI Electronic Resource Number (2004).

40. Doo-Sang P, Foottit R, Maw E and Hebert PDN, Barcoding Bugs: DNA-Based Identification of the True Bugs (Insecta: Hemiptera: Heteroptera). *PLoS ONE*; **6**(4): 1-9 DOI Electronic Resource Number (2011).

41. Ratnasingham S and Hebert PD, BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular ecology notes*; **7**(3): 355-364 DOI Electronic Resource Number (2007).

42. Kanagawa T, Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of bioscience and bioengineering*; **96**(4): 317-323 DOI Electronic Resource Number (2003).

43. Bohmann K, Monadjem A, Noer CL, Rasmussen M, Zeale MR, Clare E, Jones G, Willerslev E and Gilbert MTP, Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PLoS One*; **6**(6): e21441 DOI Electronic Resource Number (2011).

44. Clare EL, Fraser EE, Braid HE, Fenton MB and Hebert PD, Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular approach to detect arthropod prey. *Molecular Ecology*; **18**(11): 2532-2542 DOI Electronic Resource Number (2009).

45. Kruger F, Clare EL, Symondson WO, Keiss O and Petersons G, Diet of the insectivorous bat *Pipistrellus nathusii* during autumn migration and summer residence. *Mol Ecol* 2013).

46. Clare EL, Symondson WO and Fenton MB, An inordinate fondness for beetles? Variation in seasonal dietary preferences of night-roosting big brown bats (*Eptesicus fuscus*). *Mol Ecol*; **23**(15): 3633-3647 DOI Electronic Resource Number (2014).

47. Bartram AK, Lynch MD, Stearns JC, Moreno-Hagelsieb G and Neufeld JD, Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Applied and environmental microbiology*; **77**(11): 3846-3852 DOI Electronic Resource Number (2011).

48. Huang KS, Lee SE, Yeh Y, Shen GS, Mei E and Chang CM, Taqman real-time quantitative PCR for identification of western flower thrip (*Frankliniella occidentalis*) for plant quarantine. *Biology letters*; **6**(4): 555-557 DOI Electronic Resource Number (2010).

49. Naaum AM, Foottit RG, Maw HEL and Hanner R, Real-time PCR for identification of the soybean aphid, *Aphis glycines* Matsumura. *Journal of Applied Entomology*; **138**(485-489 DOI Electronic Resource Number (2014).

Data Accessibility

Sequences generated during this study were submitted to Genbank (Accession# KU594609-KU594612)

Author Contributions

B.M. contributed the original need for the assay; R.E.V., B.M. and D.F.M. designed the research; R.E.V. performed the laboratory tests; B.M. and D.M.F. contributed reagents and analytical tools; R.E.V., B.M. and D.M.F. analyzed the data; JP obtained, identified, extracted, and tested the quality of the DNA of the 12 related Pentatomidae species; R.E.V., B.M., J.L.L. and D.M.F. wrote the manuscript.

Table 1. Universal primer sets utilized throughout the project.

Primers	Sequence	Reference
ITS1		
BD1	5'-GTCGTAACAAGGTTTCCGTA-3'	Schulenburg *et al.* 2001
4S	5'-TCTAGATGCGTTCGAAGTGTCGATG-3'	Schulenburg *et al.* 2001
ITS1A	5'-CCTTTGTACACACCGCCCGTCG-3	Beebe *et al.* 1999
ITS1B	5'-ATGTGTCCTGCAGTTCACA-3'	Beebe *et al.* 1999
CO1		
LEP-F1	5'-ATTCAACCAATCATAAAGATAT-3'	Herbert *et al.* 2004
LEP-R1	5'-TAAACTTCTGGATGTCCAAAAA-3'	Herbert *et al.* 2004
LEP-R2	5'-CTTATATTATTTATTCGTGGGAAAGC-3'	Herbert *et al.* 2004

Table 2. List of twelve species from the United States tested from the family Pentatomidae, the same family as *Halyomorpha halys*.

Family	**Genus**	**Species**
Pentatomidae	*Acrosternum*	*hilaris*
	Podisus	*maculiventris*
	Banasa	*calva*
	Perillus	*bioculatus*
	Oebalus	*pugnax*
	Brochymena	*carolinensis*
	Murgantia	*histrionica*
	Euschistus	*variolarus*
		tristigmus
		servus
	Cosmopepla	*lintneriana*
	Chlorochroa	*senilis*

Table 3. Summary of the samples used throughout the project. Sources of samples, and amount of material used by weight, are shown as pools if from a common homogenate, or as individual pellets if analyzed independently. Testing method indicates whether samples were subjected to cloning and sequencing (Cloned + Seq.) or directly sequenced without cloning (Dir. Seq.), tested with the real-time PCR assay (qPCR), or tested by amplifying with just the assay primers and direct sequencing (Assay Dir. Seq.). Number of species detected from sequencing is shown per sample. In sample 4, the cloned sequence matched equally a few closely related species, some already on the list, which is why the count is "6 or 7". "Y" indicates detection of BMSB. "N" indicates no detection. Samples 12 through 24 were not directly measured, thus their weight is approximated (~) based on the sampling method used for samples 4 and 8.

ID	Source	guano mg	Testing method	Non-assay Sequencing		Real-time PCR
				# Species	BMSB	BMSB
1	Pool 1	6	Cloned + Seq. / rtPCR	4	N	Y
2	Pool 1	22	Cloned + Seq. / rtPCR	6	N	Y
3	Pool 1	29	Cloned + Seq. / rtPCR	5	N	Y
4	Pool 1	37	Cloned + Seq. / rtPCR / Assay Dir. Seq.	6 or 7	N	Y
5	Pool 2	13	Cloned + Seq. / rtPCR	3	N	N
6	Pool 2	19	Cloned + Seq. / rtPCR	3	N	N
7	Pool 2	32	Cloned + Seq. / rtPCR	5	N	N
8	Pool 2	38	Cloned + Seq. / rtPCR	6	N	N
9	Single Pellet	20	Dir. Seq. / rtPCR	1	N	Y
10	Single Pellet	25	Dir. Seq. / rtPCR	1	N	Y
11	Single Pellet	19	Dir. Seq. / rtPCR	NA	NA	Y
12	Pool 3	~37	rtPCR / Assay Dir. Seq.	-	-	Y
13	Pool 4	~37	rtPCR / Assay Dir. Seq.	-	-	Y
14	Pool 5	~37	rtPCR / Assay Dir. Seq.	-	-	Y
15	Pool 6	~37	rtPCR / Assay Dir. Seq.	-	-	Y
16	Pool 7	~37	rtPCR	-	-	Y
17	Pool 8	~37	rtPCR	-	-	Y
19	Pool 10	~37	rtPCR	-	-	Y
20	Pool 11	~37	rtPCR	-	-	Y
21	Pool 12	~37	rtPCR	-	-	N
22	Pool 13	~37	rtPCR	-	-	N
23	Pool 14	~37	rtPCR	-	-	Y
24	Pool 15	~37	rtPCR	-	-	N

Figure 1. Polymorphic sites and BMITS1 primer/probe locations within the internal transcribed spacer 1. Position numbers refer to the basepair position along the *Halyomorpha halys* sequence (Accession # KU594609-KU594612), with position one being the start of ITS1 (after the stop codon at the end of the 18S exon). Below the line are the five BMSB ITS1 variants we sequenced. The colon symbol indicates a deletion. Compared with the other three variants, variants SB9 and SB6 have a large 83 bp deletion between positions 92 and 176 that we did not include in the figure (only the first and last bp are shown) to conserve space. SB9 also contains an ambiguous site in position 372, marked as "S", which corresponds to either a "C" or "G". qPCR primers and probe are represented by light and dark grey shading, respectively, with the associated basepairs bolded black and white, respectively. The thin vertical line indicates the start of the 5.8S exon. The * symbol indicates amplification with the ITS1A&B primer set, while the remaining variants resulted from amplifications with the BD1/4S primer set.

Assay target region (96bp)

Variant	17	31	68	69	73	92	176	214	215	216	217	218	219	220	221	298	**337**	**351**	**356**	**371**	372	401	**415**	**432**	433	464
HhNJ	C	A	:	:	C	A	C	G	G	T	:	:	:	:	:	A	**C**	**A**	**C**	**A**	G	A	**G**	**G**	C	T
SB5	T	A	:	:	C	A	C	G	G	T	:	:	:	:	:	A	**C**	**A**	**C**	**A**	C	A	**G**	**G**	C	T
SB3	T	A	:	:	C	A	C	G	G	T	:	:	:	:	:	A	**C**	**A**	**C**	**A**	C	A	**G**	**G**	C	T
SB9	C	G	T	T	A	:	:	C	T	C	G	A	G	G	C	G	**C**	**A**	**C**	**A**	S	C	**G**	**G**	T	T
SB6*	C	G	T	T	A	:	:	C	T	C	G	A	G	G	C	G	**C**	**A**	**C**	**A**	G	C	**G**	**G**	T	C

85bp deletion

^ 5.8S Start

Supplemental table 1. List of the species sequences detected from guano amplifications using CO1 primers and cloning or direct sequencing. Species listings with more than one species in a line, or a lack of a distinct species name, indicates inability to discern the correct species from the sequence provided using a 98% selection criteria. R1 refers to a ~650 bp amplicon, while R2 refers to a ~350bp amplicon. Percentages indicate the prevalence of each species among the clones successfully sequenced.

Order	Family	Genus/species	R1	R2	Percent occurrence
Coleoptera	Pyrochriodae	*Dendroides canadensis*	4	-	3.25%
Hemiptera	Acanaloniidae	*Acanalonia conica*	-	1	0.81%
Ephemeroptera	Caenidae	*Caenis amica* (Mayfly)	3	-	2.44%
Ephemeroptera	Caenidae	*Caenis diminuta* (Mayfly)	-	32	26.02%
Ephemeroptera	Caenidae	*Caenis punctata* (Mayfly)	1	-	0.81%
Coleoptera	Silphidae	*Calosilpha brunneicollis*	3	-	2.44%
Diptera	Ceratopogonidae	*Culicoides austropalpalis*	-	1	0.81%
Diptera	-	*Diptera sp.*	1	1	1.63%
Diptera	Drosophilidae	*Drosophila punjabiensis*	-	1	0.81%
Diptera	Drosophilidae	*Drosophila flavopilosa*	2	-	1.63%
Diptera	Drosophilidae	*Drosophila subquinaria*	19	13	26.02%
Diptera	Drosophilidae	*Drosophila punjabiensis/vulcana/recens/jambulina/subquinaria*	-	1	0.81%
Coleoptera/Lepidoptera/Coleoptera	Elateridae/Oecophoridae/Elateridae	*Megathous dauricus/Oligoloba severa/Oedostethus sp.*	-	1	0.81%
Coleoptera	Elateridae	*Melanotus testaceus*	-	2	1.63%
Coleoptera	Elateridae	*Melanotus similis* (Click beetle)	1	14	12.20%
Diptera	Tipulidae	-	-	1	0.81%
Coleoptera	Scarabaeidae	*Nipponoserica peregrina*	10	7	13.82%
Lepidoptera	Oecophoridae	*Philobota hylophila/haplogramma/Cirromitra tetratherma*	-	1	0.81%
Coleoptera	Carabidae	*Trichotichnus vulpeculus* (Ground beetle)	2	-	1.63%
Hemiptera	Pentatomidae	*Acrosternum hilare* (Green stink bug)	1	-	0.81%

Early detection of terrestrial invasive insect infestations by using eDNA from crop surfaces

Rafael Valentin[1,4], Dina M. Fonseca[1,2], Anne L. Nielsen[2], Tracy C. Leskey[3], Julie L. Lockwood[1]

Author affiliations:
[1]Department of Ecology, Evolution & Natural Resources, Rutgers University, New Brunswick, NJ, USA

[2]Department of Entomology, Rutgers University, New Brunswick, NJ, USA

[3]USDA-ARS Appalachian Fruit Research Station, Kearneysville, WV, USA

[4]**Corresponding author**: Rafael Valentin, Ecology and Evolution Graduate Program, 14 College Farm Rd., New Brunswick, NJ 08901, USA; Rafael.Valentin@Rutgers.edu

Abstract

As the number of invasive exotic species has increased over recent decades, so too has the ecological harm and economic burdens they impose. Rapid-response eradication of nascent exotic populations is a viable approach to minimizing damage, however implementation is limited by the difficulty in detecting rare individuals. In aquatic ecosystems, the use of environmental DNA (eDNA) has helped address this issue, however, to our knowledge eDNA has not been trialed for surveillance of terrestrial exotic species. Using a high-resolution real time PCR assay for the invasive brown marmorated stink bug (*Halyomorpha halys*), we developed a highly efficient protocol to survey agricultural fields for infestations of BMSB. We compared eDNA to conventional monitoring traps and document significantly increased sensitivity and detection effectiveness. Our methodology is transferable to terrestrial situations where target species' DNA can be aggregated suggesting eDNA can potentially transform our ability to survey for exotic insects in terrestrial settings.

Introduction

Early detection of exotic populations followed by rapid management response has resulted in successful eradication of several species known to cause ecological or economic harm (Mehta et al., 2007). Eradication requires implementation of lethal control measures, many of which have unwanted secondary effects (e.g., harm to non-target species leading to loss of ecosystem services). Any delay in detection and eradication allows time for exotic populations to increase in number and geographical extent, and as both of these increase, the economic cost of eradication quickly grows while the probability of successful eradication substantially decreases (Simberloff et al., 2013). Furthermore, when exotic populations are left unmanaged for long periods, efforts shift from eradication to protection of valued assets, which is often accomplished through the continual application of control methods (Simberloff et al., 2013). Taken together, these management realities place a large emphasis on detecting the presence of individuals of unwanted and harmful exotic species when they are very rare. However, achieving this goal has proven extraordinarily difficult due to the low likelihood of detecting individuals when they are still rare in the landscape (Simberloff et al., 2013). In response, researchers have invested in improving survey design and statistical analysis, and in devising more sensitive surveillance tools (Mehta et al., 2007;Jerde et al., 2011). Here, we describe the use of environmental DNA (eDNA) to substantially improve the detection of an agricultural insect pest, and in doing so provide precedence for the use of eDNA for surveillance within other terrestrial invasion scenarios.

Environmental DNA is freely available DNA, or biological material containing DNA, that has been shed or dropped by individuals as they move through their

environment (e.g. skin flakes, hair, feathers, scales, setae, *exuviae*, fecal matter) (Bohmann et al., 2014). This source of DNA can persist and accumulate within non-living environmental material, which is then tested using high-resolution processing techniques in order to detect trace amounts of DNA (Rees et al., 2014;Barnes and Turner, 2016). Environmental DNA has been used successfully to surveil for invasive aquatic organisms (Jerde et al., 2011;Jerde et al., 2013), and is considered a burgeoning field of investigation within invasion science (Ricciardi et al., 2017). However, as of this writing, efforts to use eDNA for exotic species surveillance have not made the jump to terrestrial ecosystems. Extraction and analysis of DNA within soil is used extensively to characterize microbial and other communities, clearly indicating that the technical issues associated with using eDNA in terrestrial settings are minor. However, in the context of surveillance, eDNA approaches must be capable of detecting individuals of focal species when they are very rare. Aquatic systems are bounded and can more readily mix, and sampling approaches that filter large amounts of water accumulate DNA making detection of exotic species, even when very rare, more likely. The same may not always be true for terrestrial systems, perhaps limiting the usefulness of eDNA approaches to exotic species surveillance.

However, terrestrial systems could benefit greatly from the use of eDNA techniques in invasive exotic species surveillance. Successful development of eDNA surveillance for exotic insects alone could translate into rapid response eradications of species known to be harmful to valuable crops and forests before operations require massive applications of chemical insecticides, tree removal programs, or regulatory restrictions (McClure et al., 2001;Kovacs et al., 2010). The challenge to reaching this

goal is addressing the inherent heterogeneity in terrestrial systems that prevent dispersed eDNA from being easily detected. Our goal was to adapt current eDNA strategies to develop and test a highly-sensitive surveillance framework for use in detecting a terrestrial exotic insect.

We developed our approach within an agricultural system, focusing on an expanding invasive insect, the brown marmorated stink bug (BMSB - *Halyomorpha halys*) (**Figure 1a)**. The BMSB is native to northeast Asia, and was first found in the United States in Allentown, Pennsylvania in 1996 (Hoebeke and Carter, 2003). It has since been found in at least 40 US states, Canada, and several European countries (Valentin et al., 2017), and has caused significant damage to agricultural crops and ornamental plants (**Figure 1b)** resulting in millions of USD in economic losses (Nielsen and Hamilton, 2009;Leskey et al., 2012a). The range of crops BMSB attacks is extensive, making it a threat to farmers around the world. Due to the severity of the damage it causes, farmers have controlled populations by increasing their frequency and intensity of broad-spectrum insecticide applications (Leskey et al., 2012b). Such applications are known to be disruptive to natural ecosystems, and undermine integrated pest management efforts (Leskey et al., 2012b). We aimed to develop a novel way to conduct surveillance of nascent populations of BMSB using eDNA within an agricultural setting, and test whether such a method would be more effective than current practices.

Methods

Target eDNA collection

We used a genetic tool we previously designed for BMSB that is very sensitive to trace amounts of degraded DNA and exclusively targets BMSB (Valentin et al., 2016;Maslo et al., 2017). The BMSB is a sap-feeder that remains on the host plant for extended periods of time (Leskey et al., 2012a), potentially leaving a detectable level of DNA as they feed, defecate, or molt. The crops they feed on are often harvested by farmers and brought to centralized locations for rinsing to remove soil and other detritus, and for boxing to sell. We posited that rinsing harvested crops in water, and then concentrating, extracting and testing existing DNA (**Panel 1** and **Figure 2**) could be used as a viable surveillance technique.

To test whether BMSB DNA could be collected from water we placed individual BMSB excreta and *exuvia* in replicates of six and two, respectively, in a liter of deionized water, with two water-only samples acting as negative controls. Following one methodology in the eDNA literature (e.g. Rees et al., 2014;Turner et al., 2014), we used a peristaltic pump (Pegasus Alexis, Pegasus Pump Company, Bradenton, Florida 34211, USA) and 10µm PCTE filter membrane (GVS North America, Sanford, ME 04073, USA) combination to remove the DNA from the water. Once DNA collection was complete, we handled filter membranes with flame sterilized tweezers, cut pieces approximately 14mm^2 from the center of the filter with flame sterilized scissors, and extracted DNA using an affordable and readily available HotSHOT extraction (Johnson et al., 2015) (**Panel 1**). To assess the presence of BMSB DNA within these samples we used a TaqMan qPCR assay designed specifically for BMSB (for details regarding

assessments of specificity and sensitivity please refer to Valentin et al., 2016). Briefly, we used 20µl reactions with 500nM of each primer, 250nM of the probe, 1x TaqMan® Environmental Master Mix 2.0, and 2µl of DNA, following a reaction protocol with an initial denaturing step of 96°C for 10 min, followed by 45 cycles of denaturing for 15 s and annealing and extension at 60°C for 1 min. All reactions were run on an Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Life Technologies, Carlsbad, CA) in replicates of two within a laminar flow hood with a UV light for surface sterilization prior to qPCR setup to ensure a clean working environment.

BMSB DNA deposition rate

To document how much time an individual BMSB must be present and feeding on a fruit before a detectable level of eDNA was deposited, we conducted a time series experiment placing a single BMSB adult (from a colony maintained at Rutgers) in a small cage containing a single tomato. We allowed individual BMSB to feed on single tomatoes for a period of two, four, six, or eight hours with four replicates of each treatment (in all 16 BMSB and 16 tomatoes were used). While wearing nitrile gloves, we rinsed each tomato in a bucket containing a liter of deionized water (changing gloves between tomatoes), pumped the water to collect the eDNA, then processed and tested the filters as described above. In addition, as controls, we rinsed and filtered water from tomatoes kept in cages without BMSB (two replicates) and from two tomatoes that were not placed in cages. Filter extraction and qPCR processing were identical to the previous experiment.

Development and testing of field protocol

To examine the efficiency of this protocol in locations varying in levels of BMSB infestation we sampled crops from two farms. The first was in New Jersey (NJ) where BMSB is prevalent, and the second in New Hampshire (NH) where BMSB have not been confirmed as present in agricultural fields but that sits near the edge of the species' current known range (**Figure 3**). At both farms, we performed eDNA based surveillance in conjunction with a blacklight trap (Old Boys Enterprises Inc., Oregon, WI) and four Dead-Inn 4-ft black pyramid traps (AgBio Inc., Westminster, CO) with Trécé PHEROCON® BMSB (low dose) pheromone lures (Trécé Inc., Adair, OK) (Weber et al., 2017) so that we could directly compare effectiveness at detecting BMSB. We trapped for BMSB and filtered one to two liters of rinse water at both sites in July and August when they are naturally most abundant.

Field-testing at a high BMSB abundance site

In New Jersey we worked in a peach orchard in the Rutgers Agricultural Research Extension Center (RAREC) in Bridgeton, NJ known to harbor large populations of BMSB (**Figure 3**). While wearing nitrile gloves, we collected five to seven peaches from four different peach trees and washed them in buckets with one liter of deionized water, while still in the field. All peaches from each tree were washed in the same bucket, and each tree had a pyramid trap and pheromone lure directly next to it that had been placed at the start of the season (with lures regularly replaced). Since each tree had its own trap and was considered a separate location within the site, gloves were changed between trees to prevent cross contamination, and buckets representing each tree were kept isolated from each other to assess positive or negative detections by location within each

site. The water in each bucket was processed using the pump and filter combination as in the laboratory experiments. Once filtration was completed, the filter membranes were removed from their housing and placed in 1.5mL microcentrifuge tubes containing molecular grade 100% ethanol for storage and transport to the lab. Filters were handled using flame-sterilized tweezers and processed as previously described immediately upon return to the lab. Trees were chosen by proximity to the four pheromone traps deployed in the site, which were deployed > 50m from each other. The New Jersey site was visited twice during the first and third weeks of July, and all four trees were tested once per visit.

Field-testing at an unknown BMSB abundance site

We further tested the performance of the eDNA field surveillance protocol against conventional monitoring methods (i.e. blacklight traps and pheromone traps) at Heron Pond Farm (NH), a diversified vegetable farm that sits near the expanding front of the BMSB geographical range but was not known to be infested. We visited the New Hampshire farm twice, during the first and third weeks of August. We set one blacklight trap powered from a 120v wall outlet, and four Dead-Inn 4-ft black pyramid traps with Trécé PHEROCON® BMSB (low dose) lures spread throughout four fields containing anywhere from one to three different crop varieties each (cucurbits, chard, kale, arugula, tomatoes, and peppers) (**WebTable 1**). All traps were run continuously throughout the sampling period, while blacklight traps were inspected each morning, and pheromone traps inspected both in the morning and at various points throughout the harvesting period each day during each week. To ensure the farm's wash containers were not contaminated with BMSB DNA prior to contact with crops, after each container was

filled with the farm's local water supply (river and well water) and readied for use, we filtered one liter of water and tested the filter paper for the presence of BMSB DNA. This effort ensured that any positive identification of BMSB from using these containers was not due to contamination of BMSB DNA from anywhere else except the crops being washed that day. For each water container, after crops were harvested and thoroughly washed, approximately one to two liters of water, depending on amount of suspended materials and subsequent filter saturation, were pumped through the filtration system and processed for eDNA collection. This resulted in seven to thirteen filter samples per day, from nine different crops over eight days (two four-day sampling periods). While some crops harvested and washed were directly next to the pheromone trap in their respective field, most were over fifty meters away. Samples were processed in the field and lab in the same way as during the experiments in New Jersey.

Detection probability calculation

We calculated detection probabilities using multi-method occupancy modeling for both eDNA surveillance and pheromone traps for the four fields surveyed at the New Hampshire farm. We used multi-method occupancy modeling because it corrects for the fact that the two surveillance methods (eDNA and pheromone traps) are not independent of each other as they are sampling areas that overlap one another (Schmelzle and Kinziger, 2016). We collapsed replicate filter samples for sampled containers, as well as experimental qPCR replicates, into a single detection event. We considered the container, and thus the crops washed in that container, positive for BMSB DNA if at least one replicate produced a positive result. We binned the surveyed crops by field, in

conjunction to the placement of the pheromone traps, and treated each day as a separate survey period (**WebTable 1**). All calculations were run using the program PRESENT v12.1 (Hines, 2006).

Results

We found that all water samples spiked with BMSB were qPCR positive, and all negative controls were negative. The time series experiment resulted in positive detections in the rinse water across all time ranges (i.e. two, four six, and eight hours) indicating that detectable levels of BMSB DNA were deposited after only a few hours, at least under cage conditions.

Rinse water of peaches from the four trees on the New Jersey farm tested positive for BMSB DNA during both visits (**WebTable 1a**). Pheromone traps located next to each of the trees were also positive, and on a few occasions we observed BMSB nymphs crawling on peaches just before the fruit was collected for processing. All negative controls were negative.

We found that the eDNA strategy was both effective in the field and more sensitive to smaller populations than the blacklight and the pheromone traps. At the New Hampshire farm, we found evidence of BMSB eDNA on all eight days sampled (**WebTable 1b**). Tests of the wash containers prior to washing harvested crops yielded no positive detections, indicating no pre-contamination. The blacklight trap, collected a number of different insect species, but not BMSB. The pheromone traps caught a few native stink bug species throughout the sampling period (e.g. green stink bug – *Chinavia hilaris*), but only one BMSB, a nymph collected on the last day of sampling (**WebTable**

1b). Physical detection in the New Hampshire farm provided a visual confirmation of the presence of BMSB. We note that this nymph was found near the end of August, after BMSB populations had the opportunity to grow throughout the season.

We found that, in contrast to the pheromone traps, our eDNA surveillance protocol detected BMSB across all sampling periods. Our multi-method occupancy model for the surveying efforts conducted on the New Hampshire farm yielded detection probabilities of 0.03 (0.038 standard error) and 1.0 (0.00 standard error) for the pheromone traps and the eDNA method, respectively.

Discussion

We report here on the development and testing of a novel surveillance tool for a terrestrial insect invader using eDNA, and our results provide strong evidence that such an approach can be used successfully within an agricultural setting. Our approach provides much greater sensitivity to the presence or absence of brown marmorated stinkbug (BMSB) than the blacklight and pheromone-baited traps evaluated here. Although these traps were originally designed to monitor population abundance within established or spreading BMSB populations and not for surveillance (Nielsen et al., 2013;Short et al., 2017), they are currently the best option available for either purpose. The key to our eDNA approach is recognizing that individual BMSB naturally gather on fruit and vegetable crops and regularly deposit their DNA as they feed. This DNA would still be difficult to sample effectively for use in surveillance protocols if we did not have a way to aggregate it and test for target species' presence. We solved this problem by recognizing that harvested crops are gathered before they are sold by farmers, and thus

can easily be sampled in bulk using rinse water. At the New Hampshire site, a fully operational vegetable farm, we showed that our eDNA surveillance protocol could be worked seamlessly into on-farm protocols by using their wash containers and water sources. This demonstration was paramount for our method to be considered fully implementable on working farms.

Our approach has the potential to revolutionize agricultural pest surveillance, although there remains much needed research regarding the 'ecology' of eDNA on working farms (Barnes and Turner, 2016), the cost effectiveness of eDNA for surveillance, as well as a better understanding of when an eDNA detection heralds an infestation since not all introductions result in establishment (Blackburn et al., 2011). Nevertheless, the growing number of exotic insects that are known to be harmful to agricultural crops makes such research investments worthwhile. In particular, once optimized, the possibility of stratified sampling of eDNA in agricultural settings to pinpoint infested areas has the potential to substantially reduce the need to apply chemical insecticides over an entire landscape (**Panel 1**) with follow-on positive impacts on native species and ecosystems (Kremen et al., 2002). Stratifying sampling across farms or discrete fields may also be useful from a biosecurity perspective if it allows nascent infestations to be tightly spatially delineated, and appropriate quarantine or control methods strategically applied so that they minimize disruption to normal farm operations.

Acknowledgments

We would like to thank Andre Cantelmo Jr. and Greg Blalog and their staff at Heron Pond Farm, NH, the staff at RAREC, USDA SARE (2014-38640-22161) and USDA NIFA Pre-doctoral Fellowship (2017-67011-26068) for funding, Dr. George Hamilton and John Cambridge for the use of their BMSB colony.

References

Barnes, M.A., and Turner, C.R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics* 17**,** 1-17.

Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R., and Richardson, D.M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology & Evolution* 26**,** 333-339.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., and De Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution* 29**,** 358-367.

Hines, J. (2006). "PRESENCE. Software to estimate patch occupancy and related parameters. USGS, Patuxent Wildlife Research Center, Laurel, Maryland, USA".).

Hoebeke, E.R., and Carter, M.E. (2003). *Halyomorpha halys* (Stal)(Heteroptera: Pentatomidae): A polyphagous plant pest from asia newly detected in North America. *Proceedings of the Entomological Society of Washington* 105**,** 225-237.

Jerde, C.L., Chadderton, W.L., Mahon, A.R., Renshaw, M.A., Corush, J., Budny, M.L., Mysorekar, S., and Lodge, D.M. (2013). Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. *Canadian Journal of Fisheries and Aquatic Sciences* 70**,** 522-526.

Jerde, C.L., Mahon, A.R., Chadderton, W.L., and Lodge, D.M. (2011). "Sight‐unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters* 4**,** 150-157.

Johnson, B.J., Robson, M.G., and Fonseca, D.M. (2015). 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.

Kovacs, K.F., Haight, R.G., Mccullough, D.G., Mercader, R.J., Siegert, N.W., and Liebhold, A.M. (2010). Cost of potential emerald ash borer damage in US communities, 2009–2019. *Ecological Economics* 69**,** 569-578.

Kremen, C., Williams, N.M., and Thorp, R.W. (2002). Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences* 99**,** 16812-16816.

Leskey, T.C., Hamilton, G.C., Nielsen, A.L., Polk, D.F., Rodriguez-Saona, C., Bergh, J.C., Herbert, D.A., Kuhar, T.P., Pfeiffer, D., and Dively, G.P. (2012a). Pest status of the brown marmorated stink bug, *Halyomorpha halys* in the USA. *Outlooks on Pest Management* 23**,** 218-226.

Leskey, T.C., Lee, D.-H., Short, B.D., and Wright, S.E. (2012b). Impact of insecticides on the invasive *Halyomorpha halys* (Hemiptera: Pentatomidae): analysis of insecticide lethality. *Journal of Economic Entomology* 105**,** 1726-1735.

Maslo, B., Valentin, R., Leu, K., Kerwin, K., Hamilton, G.C., Bevan, A., Fefferman, N.H., and Fonseca, D.M. (2017). Chirosurveillance: The use of native bats to detect invasive agricultural pests. *PloS one* 12**,** e0173321.

Mcclure, M.S., Salom, S.M., and Shields, K.S. (2001). *Hemlock woolly adelgid.* US Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team.

Mehta, S.V., Haight, R.G., Homans, F.R., Polasky, S., and Venette, R.C. (2007). Optimal detection and control strategies for invasive species management. *Ecological Economics* 61**,** 237-245.

Nielsen, A.L., and Hamilton, G.C. (2009). Seasonal occurrence and impact of *Halyomorpha halys* (Hemiptera: Pentatomidae) in tree fruit. *Journal of Economic Entomology* 102**,** 1133-1140.

Nielsen, A.L., Holmstrom, K., Hamilton, G.C., Cambridge, J., and Ingerson–Mahar, J. (2013). Use of Black Light Traps to Monitor the Abundance, Spread, and Flight Behavior of *Halyomorpha halys* (Hemiptera: Pentatomidae). *Journal of Economic Entomology* 106**,** 1495-1502.

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R., and Gough, K.C. (2014). REVIEW: The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51**,** 1450-1459.

Ricciardi, A., Blackburn, T.M., Carlton, J.T., Dick, J.T., Hulme, P.E., Iacarella, J.C., Jeschke, J.M., Liebhold, A.M., Lockwood, J.L., and Macisaac, H.J. (2017). Invasion Science: A Horizon Scan of Emerging Challenges and Opportunities. *Trends in Ecology & Evolution*.

Schmelzle, M.C., and Kinziger, A.P. (2016). Using occupancy modelling to compare environmental DNA to traditional field methods for regional‐scale monitoring of an endangered aquatic species. *Molecular ecology resources* 16**,** 895-908.

Short, B.D., Khrimian, A., and Leskey, T.C. (2017). Pheromone-based decision support tools for management of *Halyomorpha halys* in apple orchards: development of a trap-based treatment threshold. *Journal of Pest Science***,** 1-14.

Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D.A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., and Pascal, M. (2013). Impacts of

biological invasions: what's what and the way forward. *Trends in Ecology & Evolution* 28**,** 58-66.

Turner, C.R., Barnes, M.A., Xu, C.C., Jones, S.E., Jerde, C.L., and Lodge, D.M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution* 5**,** 676-684.

Valentin, R.E., Maslo, B., Lockwood, J.L., Pote, J., and Fonseca, D.M. (2016). Real‐time PCR assay to detect brown marmorated stink bug, Halyomorpha halys (Stål), in environmental DNA. *Pest management science* 72**,** 1854-1861.

Valentin, R.E., Nielsen, A.L., Wiman, N.G., Lee, D.-H., and Fonseca, D.M. (2017). Global invasion network of the brown marmorated stink bug, Halyomorpha halys. *Scientific Reports* 7.

Weber, D.C., Morrison, W.R., Khrimian, A., Rice, K.B., Leskey, T.C., Rodriguez-Saona, C., Nielsen, A.L., and Blaauw, B.R. (2017). Chemical ecology of *Halyomorpha halys*: discoveries and applications. *Journal of Pest Science***,** 1-20.

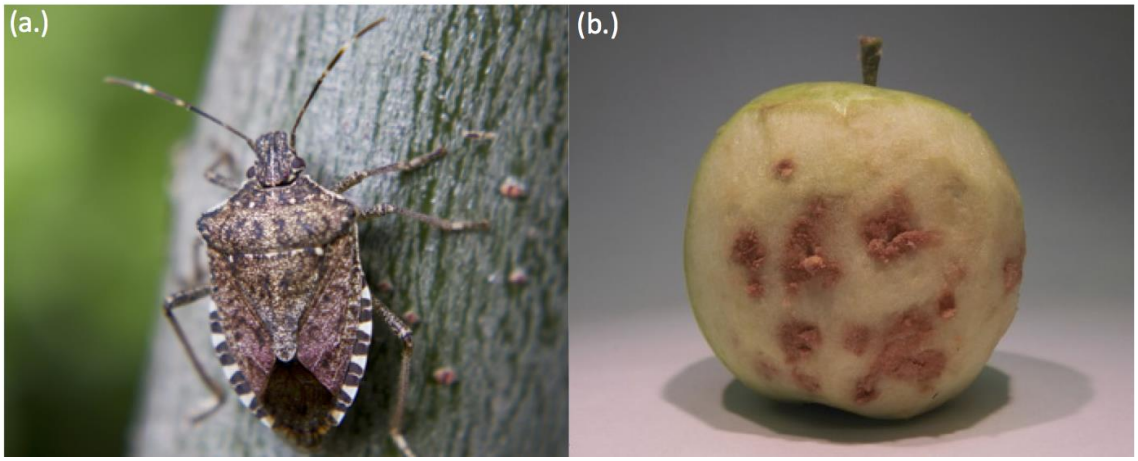

Figure 1: Illustration of (a.) the brown marmorated stink bug (BMSB) and (b.) the damage it causes to fruits it has fed upon. The photo on the left (a.) was taken from StopBMSB.org, and the right (b.) taken by Brett Blaauw, Rutgers University.

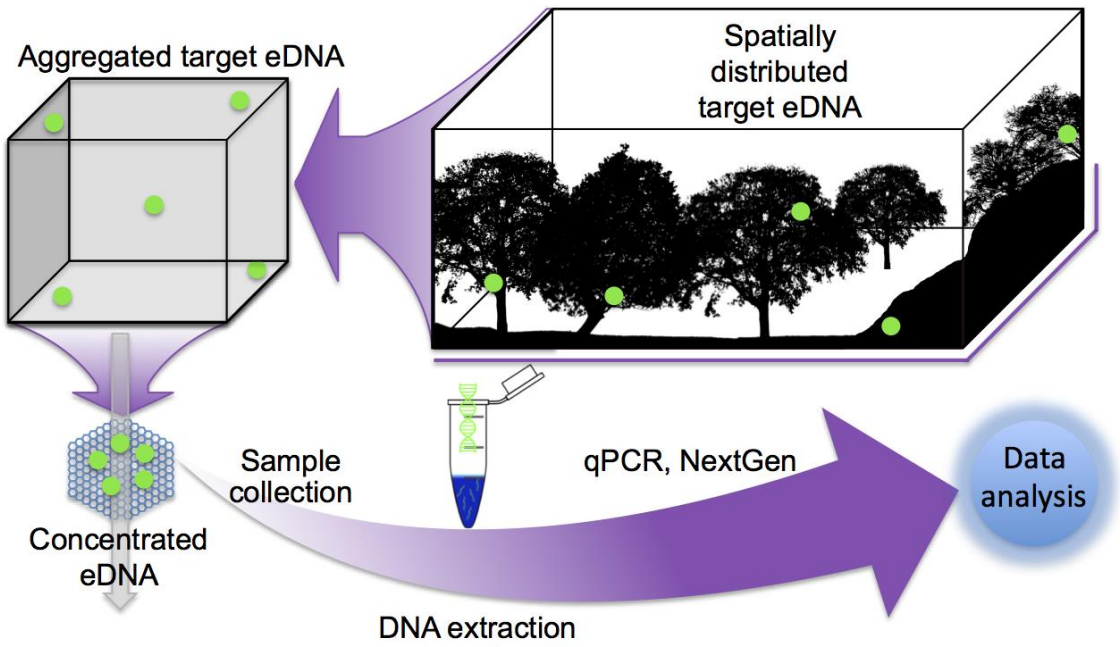

Figure 2: Terrestrial surveillance efforts using eDNA require several steps (i) assessment of likely sources of target DNA (fruit, bark, soil, etc.), (ii) implementation of strategies to aggregate the target DNA (e.g. fruit washing), (iii) concentration of the target DNA (e.g. by using filtration), (iv) sub-sampling of the concentrated sample (when needed), (v) DNA extraction (vi) testing for the presence or quantity of target DNA via qPCR, NextGen sequencing, or any other applicable method, (vii) data analysis (e.g. occupancy modeling, spatial analyses).

Panel 1: Terrestrial eDNA surveillance

The use of eDNA in terrestrial systems differs from its use in aquatic systems primarily because DNA from terrestrial species can remain localized in the environment. In contrast, DNA from aquatic species is commonly naturally dispersed, making surveillance a matter of concentrating water from a location and testing it for the presence of the target species' DNA.

To maximize probability of detection of eDNA in terrestrial systems it is necessary to first assess likely sources and develop cost-effective strategies for bulk sampling (**Figure 2**). In the case of an agricultural pest, the brown marmorated stinkbug, we made use of the aggregation of crops during harvest. Harvested crops are brought to a centralized location for sorting, washing, and packaging. The washing process is where the aggregation takes place, as all the material that was once on the surface of many individual crops is then transferred into a single bulk material (the water). From here, sampling, extraction and testing can be carried out using standard eDNA strategies (**Figure 2**). A further benefit of this strategy is that, depending on expected levels of infestation, sampling can be stratified (analyses can be performed separately) at multiple scales (by crop, crop variety, farm, latitude, state) in order to gain insight of the presence and spatial distribution of the of the target species at a multitude of spatial scales to fit information needs. This strategy would allow for more focused management where only infested locations would need treatment to control, or eradicate, nascent populations.

While we found sampling crop wash water a viable strategy for an agricultural pest, other approaches will certainly be needed for invasive species with different life histories and habitat preferences. Additionally, an analysis of the cost-to-benefit ratio for

using eDNA as opposed to traditional surveillance methods is a necessary step for any real-world applications.

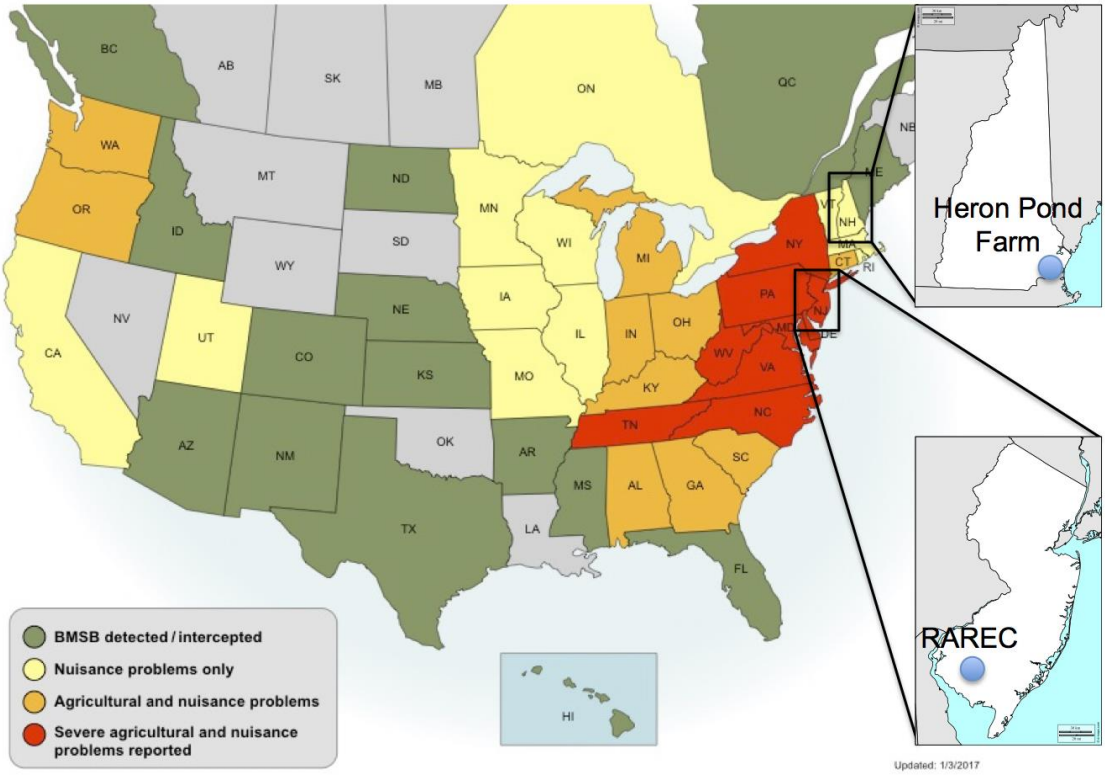

Figure 3: Map of the distribution of the brown marmorated stink bug within the continental US, with both study locations shown. RAREC stands for Rutgers Agricultural Research and Extension Center. Range map provided by StopBMSB.org.

WebTable 1: (a) Detection data for the four trees surveyed at the New Jersey site (RAREC). Columns are sorted by tree, while rows are arranged by survey day (first number) and survey method (D = eDNA, T = pheromone trap). (b) Detection data for the four fields in the New Hampshire farm. Columns are sorted by field, while rows indicate survey days (first number), and then the survey method (D = eDNA, T = pheromone traps) second. 1 represents positive detection, while 0 represents negative detection. Dots (•) represent days where a field was not surveyed with the eDNA method, due to lack of harvest. 'Greens' refer to the crop varieties chard, kale, and arugula.

(a)

	Tree 1 (peaches)	Tree 2 (peaches)	Tree 3 (peaches)	Tree 4 (peaches)
1-D	1	1	1	1
1-T	1	1	1	1
2-D	1	1	1	1
2-T	1	1	1	1

(b)

	Field 1 (cucurbits)	Field 2 (greens)	Field 3 (tomatoes)	Field 4 (peppers)
1-D	1	1	1	•
1-T	0	0	0	0
2-D	1	1	1	1
2-T	0	0	0	0
3-D	1	1	1	•
3-T	0	0	0	0
4-D	•	•	1	•
4-T	0	0	0	0
5-D	•	1	•	•
5-T	0	0	0	0
6-D	1	1	•	•
6-T	0	0	0	0
7-D	•	•	•	1
7-T	0	0	0	0
8-D	•	•	1	•
8-T	0	0	1	0

Conclusion

In examining facets of invasion ecology, we gain insight into the role invasion dynamics plays across multiple ecological systems. Specifically, by studying how non-native species are moved across biogeographic scales we can begin to understand how events may have unfolded during their invasion history, and how this can inform observations in the present day. Additionally, by studying these movements and unravelling their invasion history we see the extent even a single source population can have in the introduction of non-natives across the globe. Finally, by understanding the invasion process non-native species must undergo when moved across biogeographic scales, not only do we see how difficult it can really be to successfully navigate the many barriers, but also the problem this causes when attempting to determine an invasion during the earliest stages of establishment. In this dissertation we explored each of these individually, and for the latter point attempted to find a resolution to the problem at hand for terrestrial systems.

The results of this dissertation demonstrate the significance in assessing each of these points within the context of invasion ecology. Without these results, only assumptions of a non-native's invasion history can be made in an effort to understand the post-establishment evolution, or the global spread, of an introduced species. In my first two chapters I specifically explored each of these individually for two different species, the northern cardinal (*Cardinalis cardinalis*) and the brown marmorated stink bug (BMSB – *Halyomorpha halys*). For cardinals, while the observed divergence among islands was indeed significant, this variation was likely not driven by its introduction history. This was counter to our original assumption, which was the observed variation

may have been a result of how they were introduced. With respect to BMSB, assumptions again were made regarding its spread across the globe. It was not until we explored these likely pathways simultaneously that the reality of their spread was clear.

Additionally, the difficulty of detecting BMSB as it invaded landscapes, at both biogeographic and local scales, prompted a means to overcome this dilemma. The application of environmental DNA to detect them (eDNA), while appealing given its success in aquatic systems, presented several challenges in order to transfer its use within a terrestrial environment. Fortunately, the framework we developed allowed us to be successful in this endeavor, and led to the development of highly sensitive new approach that was significantly superior in detecting nascent populations than conventional methods already in place. In taking all chapters taken together, we present a contribution to the understanding of invasion dynamics that can benefit researchers and managers alike.