PLOS ONE

RESEARCH ARTICLE

Using NextRAD sequencing to infer movement of herbivores among host plants

Zhen Fu[1]*, Brendan Epstein[2¤a], Joanna L. Kelley[2], Qi Zheng[3], Alan O. Bergland[4¤b], Carmen I. Castillo Carrillo[1¤c], Andrew S. Jensen[5], Jennifer Dahan[6], Alexander V. Karasev[6], William E. Snyder[1]

1 Department of Entomology, Washington State University, Pullman, Washington, United States of America, **2** School of Biological Sciences, Washington State University, Pullman, Washington, United States of America, **3** Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, Kentucky, United States of America, **4** Department of Biology, Stanford University, Stanford, California, United States of America, **5** Northwest Potato Research Consortium, Lakeview, Oregon, United States of America, **6** Department of Plant, Soil and Entomological Sciences, University of Idaho, Moscow, Idaho, United States of America

¤a Current address: College of Biological Sciences, University of Minnesota, Saint Paul, Minnesota, United States of America
¤b Current address: Department of Biology, University of Virginia, Charlottesville, Virginia, United States of America
¤c Current address: Departamento de Protección Vegetal, Estación Experimental Santa Catalina, Instituto Nacional de Investigaciones Agropecuarias (INIAP), Panamericana Sur km 1, Quito, Ecuador
* zhen.fu@wsu.edu

Check for updates

OPEN ACCESS

Citation: Fu Z, Epstein B, Kelley JL, Zheng Q, Bergland AO, Castillo Carrillo CI, et al. (2017) Using NextRAD sequencing to infer movement of herbivores among host plants. PLoS ONE 12(5): e0177742. https://doi.org/10.1371/journal.pone.0177742

Editor: Daniel Doucet, Natural Resources Canada, CANADA

Received: November 20, 2016

Accepted: May 2, 2017

Published: May 15, 2017

Copyright: © 2017 Fu et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: DNA sequences were deposited into NCBI Short Read Archive with BioProject Accession No. PRJNA292336.

Funding: This research was funded by the National Institute of Food and Agriculture, U.S. Department of Agriculture, Specialty Crop Research Initiative (https://nifa.usda.gov/funding-opportunity/specialty-crop-research-initiative-scri) award number 2015-09273, along with grants from the Northwest Potato Research Consortium (http://www.nwpotatoresearch.com/) and the Washington

Abstract

Herbivores often move among spatially interspersed host plants, tracking high-quality resources through space and time. This dispersal is of particular interest for vectors of plant pathogens. Existing molecular tools to track such movement have yielded important insights, but often provide insufficient genetic resolution to infer spread at finer spatiotemporal scales. Here, we explore the use of Nextera-tagmented reductively-amplified DNA (NextRAD) sequencing to infer movement of a highly-mobile winged insect, the potato psyllid (*Bactericera cockerelli*), among host plants. The psyllid vectors the pathogen that causes zebra chip disease in potato (*Solanum tuberosum*), but understanding and managing the spread of this pathogen is limited by uncertainty about the insect's host plant(s) outside of the growing season. We identified 1,978 polymorphic loci among psyllids separated spatiotemporally on potato or in patches of bittersweet nightshade (*S. dulcumara*), a weedy plant proposed to be the source of potato-colonizing psyllids. A subset of the psyllids on potato exhibited genetic similarity to insects on nightshade, consistent with regular movement between these two host plants. However, a second subset of potato-collected psyllids was genetically distinct from those collected on bittersweet nightshade; this suggests that a currently unrecognized source, i.e., other nightshade patches or a third host-plant species, could be contributing to psyllid populations in potato. Oftentimes, dispersal of vectors of pathogens must be tracked at a fine scale in order to understand, predict, and manage disease spread. We demonstrate that emerging sequencing technologies that detect genome-wide SNPs of a vector can be used to infer such localized movement.

State Commission on Pesticide Registration (http://wscpr.org/), all to WES. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Herbivores often move among host plant species, driven by their need to evade and detoxify plant defenses, balance nutritional requirements that cannot be met by single plants, and/or track spatiotemporal variation in plants' resource quality [1, 2]. At the broadest scale, herbivores may traverse thousands of kilometers, tracking host-plant availability across seasons (e.g. [3–6]) or due to varying rainfall and wind patterns (e.g. [7, 8]). At a finer scale, herbivores often move among host-plant species within a habitat while tracking host-plant phenology, as different host-plant species go through seasonal changes in nutritional value and/or ability to physically or chemically defend themselves [9–11]. When herbivores act as vectors of plant pathogens, these movements can have particularly dramatic effects on host plants; herbivores can initiate pathogen outbreaks even when herbivore densities are too low to inflict appreciable direct damage [12–14]. Oftentimes, a detailed understanding of movement of vectors among host plant species, or within stands of the same species, is critical for predicting patterns of disease spread (e.g. [14–16]).

When herbivores are relatively large, or the distances covered are relatively small, physically marking and tracking individual herbivores can be an effective way to unravel patterns of host-plant switching [17–19]. However, when this is impossible or impractical, patterns of interrelatedness among herbivores can be used to infer likely movement patterns. Molecular techniques, including protein and microsatellite DNA markers, were among the first genetic tools used to infer gene flow and thus herbivore dispersal [20–22]. However, developing a sufficiently large set of markers to delineate localized movement can be time consuming and expensive, or even impossible when there are few microsatellites in the genome [23–25]. Recently, restriction-site associated DNA (RAD) markers have been used to overcome these limitations by allowing quick detection of single nucleotide polymorphisms (SNPs) across focal organisms' entire genomes. RAD-based approaches have proven powerful in tracking genetic differentiation across landscapes (e.g. [26, 27]), but the relatively high DNA-volume inputs required has thus far limited their use to larger-bodied organisms. Because of their small body sizes, many herbivorous insects that feed heavily on plants (and/or vector key plant pathogens) have thus far been outside the reach of these approaches.

Here, we explore the use of Nextera-tagmented reductively-amplified DNA ("NextRAD") sequencing to infer movement among host plant species by a winged, small-bodied insect, the potato psyllid (*Bactericera cockerelli*). The psyllid is the vector of the bacterium (*Candidatus* Liberibacter solanacearum) that causes zebra chip disease in cultivated potato (*Solanum tuberosum*) [28], whose spread has endangered potato production in several parts of the United States of America (e.g. [29]). In the northwestern U.S., it has been proposed that potato psyllids transmit the zebra chip pathogen as the insects migrate from the perennial solanaceous weed bittersweet nightshade, *Solanum dulcumara*, to annually-cultivated *S. tuberosum* fields each year [30]. However, movement of potato psyllids from bittersweet nightshade to potato has never been directly demonstrated, hindering any ability to understand, predict, or manage zebra chip outbreaks [30]. Indeed, a wide variety of plant species other than bittersweet nightshade have been proposed to be the true source of psyllids (and perhaps also the zebra chip pathogen) that colonize potato fields [31]. Existing molecular tools for dividing psyllids into geographically-separated genetic groups, based on sequence variation within the cytochrome *c* oxidase I (COI) gene, are too limited to reveal genetic subpopulations at a fine-enough scale to identify gene flow among host plants [30]. NextRAD sequencing overcomes these limitations by fragmenting and ligating adaptor sequences to genomic DNA via engineered transposomes. Critically, NextRAD requires less than 50 ng of DNA [32, 33], making it possible to generate sequence data from organisms far smaller than was possible with the original RAD sequencing

approaches. Using this technique in combination with high-throughput sequencing generates a large number of markers, greatly facilitating examination of the genomic variation of psyllid populations and individuals. Thus, we could assess whether bittersweet nightshade could be the sole source of potato psyllids colonizing potato fields, or whether instead other non-crop host species might need to be identified.

Materials and methods

Our project included regional sampling of spatially-dispersed herbivore populations on two host plant species, followed by sequencing the insects to infer population interrelatedness. First, over two years, we collected potato psyllids from bittersweet nightshade patches located throughout much of the potato-growing region of east-central Washington State (USA); in one of these years, we also collected psyllids from production potato fields across this same region (Fig 1). Additionally, we collected psyllids from a nightshade patch located in southern Idaho (Fig 1) to serve as a geographically-distinct outgroup. A subsample of the psyllids

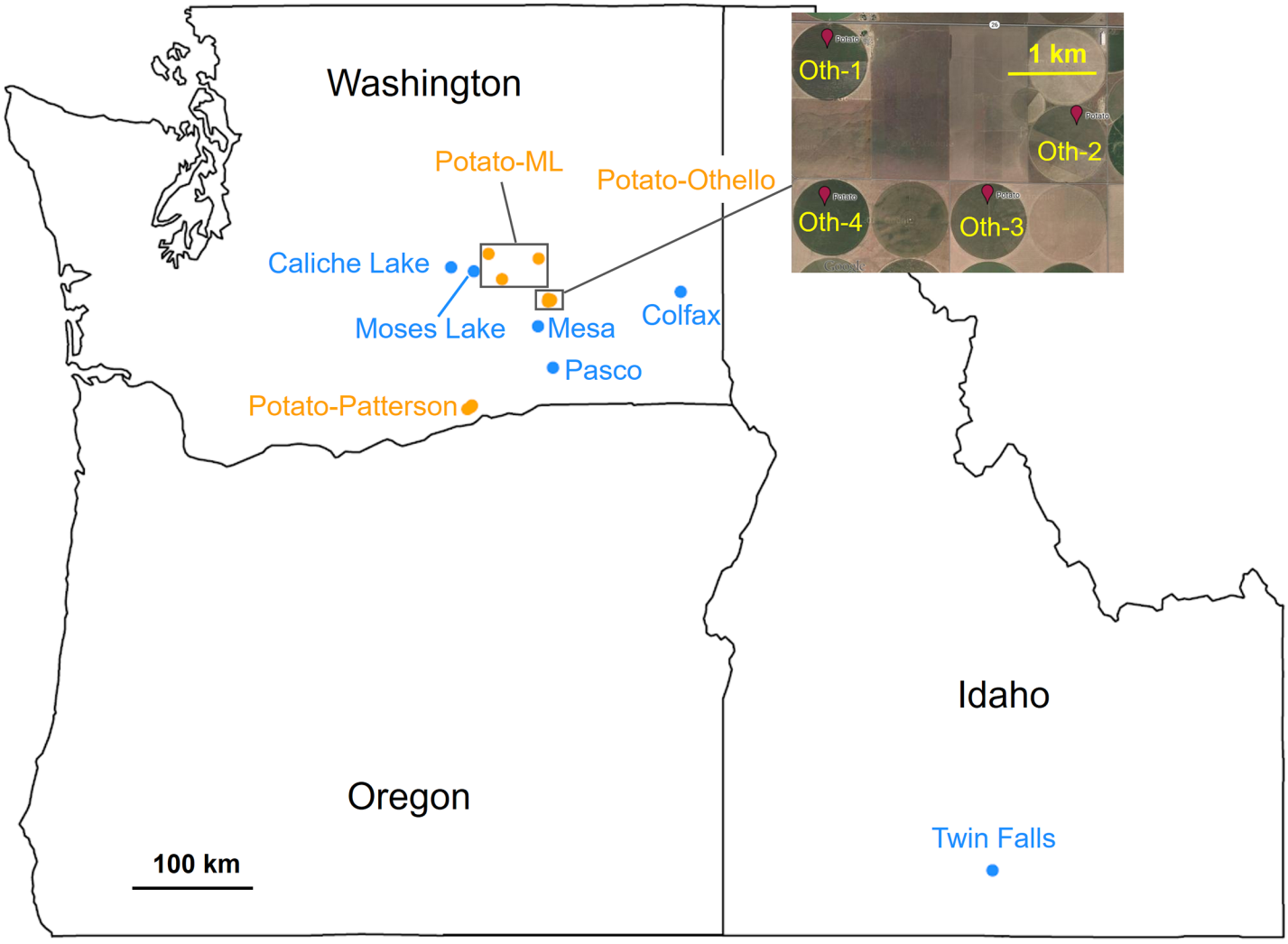

Fig 1. Potato psyllid collection sites across the states of Washington and Idaho, United States of America. Potato psyllids were collected from bittersweet nightshade patches (blue circles) or potato fields (yellow circles). See Table 1 and S1 Table for detailed information about each population.

https://doi.org/10.1371/journal.pone.0177742.g001

collected from nightshade patches (up to 10 psyllids per sampling date), and all psyllids collected from potato fields, were then sequenced using the NextRAD approach; this allowed us to identify variant sites throughout the psyllid genome. We then used multiple population-genetic approaches to determine population structure among psyllids collected from the two host plants, to infer whether they are composed of a single interbreeding population or instead include members of genetically-distinct sub-populations. Each of these project sub-components are detailed below.

Potato psyllid sampling and sequencing

Our study did not involve any endangered or protected insect species, and no specific permits were required. For the psyllids collected from private land, we obtained the permission from the land owners. Potato psyllids were collected from six bittersweet nightshade patches and ten potato fields in the U. S. states of Washington and Idaho (Fig 1 and Table 1), over two growing seasons, using a suction sampling device (see [34, 35]). Sites were chosen to cover the majority of the potato-growing region in east-central Washington, with the Idaho site serving as a geographically-distant outgroup, and were sampled periodically over the 2012 and 2013 growing seasons (Fig 1 and S1 Table). Psyllids were placed on dry ice immediately following collection, and were stored in 95% ethanol upon arrival in the laboratory. For each sampling date, on each host plant and at each location, four to ten intact adult psyllids were randomly selected for DNA extraction and sequencing (Table 1 and S1 Table). In total, we processed 285 psyllids for NextRAD sequencing.

To begin DNA extraction, individual psyllid adults were placed into separate microcentrifuge tubes with 150 μl tissue lysis buffer (10mM Tris pH = 8; 50mM Ethylenediaminetetraacetic acid; 200mM NaCl; 1% (w/v) Sodium dodecyl sulfate; we found this homemade lysis buffer to be more efficient in breaking down the insect exoskeleton than the buffer in the commercial kit), and ground for 1 minute using a pestle driven by a handheld electric mixer. Thereafter, DNA extraction was conducted following the instructions of the Qiagen DNeasy Blood &

Table 1. Sampling locations and sampling dates of potato psyllids *Bactericera cockerelli* in the states of Washington and Idaho, United States of America.

Sampling ID	Host plants	Coordinates	# of psyllids sequenced	# of sampling dates
Caliche Lake	Nightshade	47˚ 1'54''N, 119˚55'40"W	18	2
Colfax*	Nightshade	46˚50'51"N, 117˚28'44" W	31	4
Mesa*	Nightshade	46˚35'18"N, 119˚ 0'1"W	66	7
Moses Lake*	Nightshade	47˚ 0'7"N, 119˚41'5"W	60	6
Twin Falls	Nightshade	42˚29'57"N, 114˚ 9'15"W	50	5
Pasco	Nightshade	46˚16'39"N, 118˚50'29"W	10	1
ML-1	Potato	46˚ 56' 31"N, 119˚ 23' 4" W	5	2
ML-2	Potato	47˚ 7' 58"N, 119˚ 31' 43"W	1	1
ML-3	Potato	47˚ 5' 50"N, 118˚ 59' 41" W	1	1
Othello-1	Potato	46˚ 47' 40"N, 118˚ 53' 30"W	11	2
Othello-2	Potato	46˚ 47' 12"N, 118˚ 51' 41"W	10	2
Othello-3	Potato	46˚ 46' 40"N, 118˚ 52' 20"W	6	1
Othello-4	Potato	46˚ 46' 29"N, 118˚ 53' 44"W	10	1
Patterson-1	Potato	45˚ 57' 57"N, 119˚ 45' 15"W	4	1
Patterson-2	Potato	45˚ 58' 55"N, 119˚ 42' 58"W	1	1
Patterson-3	Potato	45˚ 59' 36"N, 119˚ 42' 19"W	1	1

* indicated that location was sampled for both 2012 and 2013, the other locations were sampled either in 2012 or 2013.

https://doi.org/10.1371/journal.pone.0177742.t001

Tissue Kit (Qiagen, Valencia, CA). DNA was eluted in 100 μl 10 mM Tris-HCl buffer (pH = 8). The quantity of DNA extracted from each insect was measured using a Qubit 2.0 Fluorometer (Life Technologies, Grand Island, NY).

DNA samples were sent to SNPsaurus LLC (Eugene, OR) to generate NextRAD libraries and perform sequencing. To construct DNA libraries, genomic DNA (~10 ng) was first fragmented with the Nextera reagent (Illumina, San Diego, CA), which also ligated short adapter sequences to the ends of the fragments (S1 Fig). DNA fragments were then amplified with two primers matching adaptor sequences, with one of the primers extending an additional nine nucleotides (GTGTAGAGC) as the selective sequence at the 3' end (S1 Fig). Thus, only fragments that could be hybridized to the selective sequence were efficiently amplified. The libraries were sequenced on an Illumina HiSeq2000 with 1x100 bp configuration to generate ~65X coverage.

For each insect, an aliquot of the same DNA sent for NextRAD sequencing was used for COI "haplotyping" through high-resolution melting analysis [36]. This is an approach used to delineate genetic subgroups of potato psyllids at a coarser, continental scale (i.e., the "Central", "Western", and "Northwestern" haplotypes typical of different sections of North America; [36]) relative to the fine-scale genetic differentiation within our study region that NextRAD provided (see below).

Sequence alignment, variant calling, and filtering

Quality trimming of raw reads and variant calling were performed by SNPsaurus. Trimmomatic [37] was used to remove the Nextera adapters and low quality reads (Phred quality score < 20). Thereafter, reads of all psyllids were pooled and aligned to each other to form allelic clusters ($> 95\%$ identity) using custom scripts, and the read with the highest count in the population was chosen as a reference contig. In order to identify these contigs based on gene homology, all 23,191 contigs were queried to the NCBI reference sequence (RefSeq) database [38] using BLASTN [39] with an e-value cutoff of 0.0001. Subsequently, reads from each sample were aligned to the reference using the BWA-mem algorithm (parameters: -B 3 -O 4 -k 13) of the Burrows-Wheeler Aligner [40]. Variant sites were called by SNPSaurus using the mpileup and bcftools algorithms in SAMtools [41].Bcftools uses a statistical approach to call variants. With the conservative parameters used in this study, it was more likely to call a heterozygote with low read numbers as a homozygote than call a sequencing error an allele. Thereafter, we employed PLINK (v1.90; [42]) to calculate heterozygosity, allele frequencies and the missing data rate. Loci that were missing in $> 5\%$ of individuals (2,400 loci) and individuals with $> 10\%$ missing loci (four individuals) were excluded from the dataset. In addition, we removed loci with observed heterozygosity > 0.5, excluded loci with minor allele frequency (MAF) < 0.05, and removed all the indels. We randomly sampled one variant from each contig to assure the loci were mostly independent. We used PLINK [42] to report the p-value of Hardy-Weinberg Equilibrium (HWE) tests for each locus within sampled populations (with samples grouped by site and date of collection). HWE tests were applied only to the loci with no missing genotypes, and we only analyzed populations with sample size ≥ 8. We tested the outlier loci using three programs: OutFLANK [43], LOSITAN [44] and BayeScan [45]. For LOSITAN and BayeScan, the default setup was used, and for OutFLANK, q-value was set to 0.05.

Neighbor-joining tree construction

First, we used the neighbor-joining clustering technique in order to visually describe interspersion and/or separation of psyllids based on the host plant, site, and date from which the insects

were collected. We calculated genetic similarity (proportion of shared alleles) between all pairs of individuals using the "—distance square 1-ibs" option in PLINK (v1.90; [42]). We then constructed an unrooted neighbor-joining tree [46] from the pairwise genetic distances using the "nj" function in the R package ape (v3.3;[47]) and obtained support values by randomly resampling variants with replacement 500 times (bootstrapping) in R and the R package ape [47].

Clustering analysis and population structure

We next used ADMIXTURE [48] to delineate genetically-distinct groups within our psyllid collections, searching for the number of genetic lineages that best described the data. We increased the pre-defined number of ancestral populations (K) from $K = 1$ to $K = 20$. Ancestry coefficient matrices from 50 replicated runs were aligned and averaged using the program CLUMPAK [49]. Because gene flow among sites could occur throughout our study region, and because the ancestry and lineage of all psyllid samples was unknown, we conducted ADMIXTURE analysis in the "unsupervised" mode without providing any sampling information, and we identified the best K value as the run with the lowest cross-validation error [48].

As a complementary approach examining the genetic population structure of psyllids separated by hosts, time and space, we also conducted principal component analysis (PCA) using the smartpca algorithm from EIGENSOFT (v6.0.1; [50]).

F-statistics and spatiotemporal separation

Third, as a means of inferring how patterns of psyllid interrelatedness differ through time and space, we calculated fixation indices (F_{ST}) among pairs of collection sites and dates. We first estimated the inbreeding coefficient (F_{IS}), based on Nei [51], of psyllid collections that included ≥ 4 psyllids (S1 Table). Next we estimated population differentiation (F_{ST}) between pairs of collections that included ≥ 4 psyllids (S1 Table), based on the equations described in Weir and Cockerham [52]. First, we employed linear models to examine the relationship between geographic distance separating psyllid populations and their degree of genetic interrelatedness, working with insects collected between August, 2012 and October, 2013 (this was the time window during which the most sites were sampled roughly synchronously). Second, within sites sampled repeatedly through time, we examined the relationship between degree of temporal separation and the degree of genetic divergence.

Analysis of Molecular Variance (AMOVA)

Fourth, to determine the relative importance of host, time, ADMIXTURE clusters, and spatial separation among the Washington samples, we ran two AMOVAs [53] using the poppr [54, 55] and ade4 [56] R packages. Collection population (the combination of location and time) was nested within either the ADMIXTURE population assignment or the host plant species. We used the $K = 3$ run to determine the ADMIXTURE assignment and assigned samples based on the population with the greatest ancestry fraction. We tested significance for the host AMOVA using 1000 random permutations; significance was not assessed for the ADMIXTURE cluster run because testing the significance of clusters defined by exploratory analyses on the same dataset is circular, and produces meaningless p-values [57].

Results

Sequencing

On average 400 Megabases of sequence data, equivalent to ~ 2.7 million 100 bp reads, were obtained from each psyllid NextRAD library. As no reference genome was available, a *de novo*

assembly was constructed using a custom script from our sequencing service provider, SNPsaurus. We searched for homologs of each contig in the assembly in the RefSeq database [38]. However, only 2.7% (643 of 23,191) of the contigs returned homologies and most of these contigs aligned with sequences of the Asian citrus psyllid, *Diaphorina citri*; *D. citri* is the most-closely-related species to the potato psyllid that has been sequenced [58]. We identified 8,443 variants by aligning cleaned reads of each sample back to the *de novo* assembly. After removing indels, loci with a high missing data rate, loci with high heterozygosity (> 0.5), and randomly sampling independent loci (described in Material and Methods), we included 1,978 loci in the downstream analyses (S2 Table). Within the 1,978 loci, no locus failed the HWE test (p-value < 0.001) and none of the loci were detected to be under selection by all three programs.

Our COI-haplotyping revealed that all 285 potato psyllids belonged to the "Northwestern" COI-haplotype that is typical of the region where our work was conducted (e.g., [36]). As described below, we used several methods, including neighbor-joining clustering, ADMIXTURE, PCA, and AMOVA, to characterize finer-scale population structure using the NextRAD variants.

Neighbor-joining tree

Clustering using the neighbor-joining method indicated that samples taken from the geographically isolated bittersweet nightshade patch near Twin Falls, Idaho, formed a cluster separate from all other psyllids that we collected from either of the two host plants in Washington (Fig 2). Otherwise, psyllids collected from potato fields in Washington (ML and Patterson populations) were generally interspersed with insects collected from bittersweet nightshade patches in that same state, particularly the Colfax and Moses Lake sites (Fig 2). An exception to this broader pattern was a group of psyllids collected from a suite of potato fields near Othello, Washington (enlarged circles in Fig 2, Table 1 and S1 Table); this group of potato-collected psyllids fell out in a distinct cluster separate from any psyllids collected from any other potato field (Fig 2). Note that the terminal branch lengths were quite long for this potato-collected "Othello" cluster, indicating substantial genetic variation among individual psyllids in the cluster (Fig 2). Potato psyllids collected across two years at the Moses Lake bittersweet nightshade site clustered separately, suggesting genetic differentiation between years. In contrast, the insects from the Mesa site with multi-year collections clustered together (Fig 2). Psyllids from Colfax of two years were placed into different clades, but the separation was not temporal related, suggesting genetic divergence among psyllid individuals within a year.

Admixture and PCA

In general, the population structure identified by ADMIXTURE (Fig 3) was consistent with the results of the neighbor-joining method: Idaho and Washington samples were differentiated, there was a group of psyllids, mostly from potatoes at the Othello site, that were also distinct, and there was otherwise little evidence for distinct potato and nightshade populations. In particular, At $K = 2$ (i.e. ADMIXTURE was constrained to split the samples into two groups), we saw separation of the psyllids of the most geographically-distinct population, the single nightshade patch in southern Idaho, from the majority of the psyllids collected in Washington. However, we noticed that in half of the 50 runs, some psyllids collected from Othello were grouped with the populations in Idaho (Fig 3). Interestingly, in 13 out of 50 runs, these Othello psyllids showed the opposite pattern, and they shared more similarity with the other psyllid populations in Washington (S3 Fig). The next group to separate from the others, at $K = 3$, did not reflect host plant species; rather, these were the same group of potato-collected psyllids, from potato fields near Othello, WA, identified by the neighbor-joining method as being

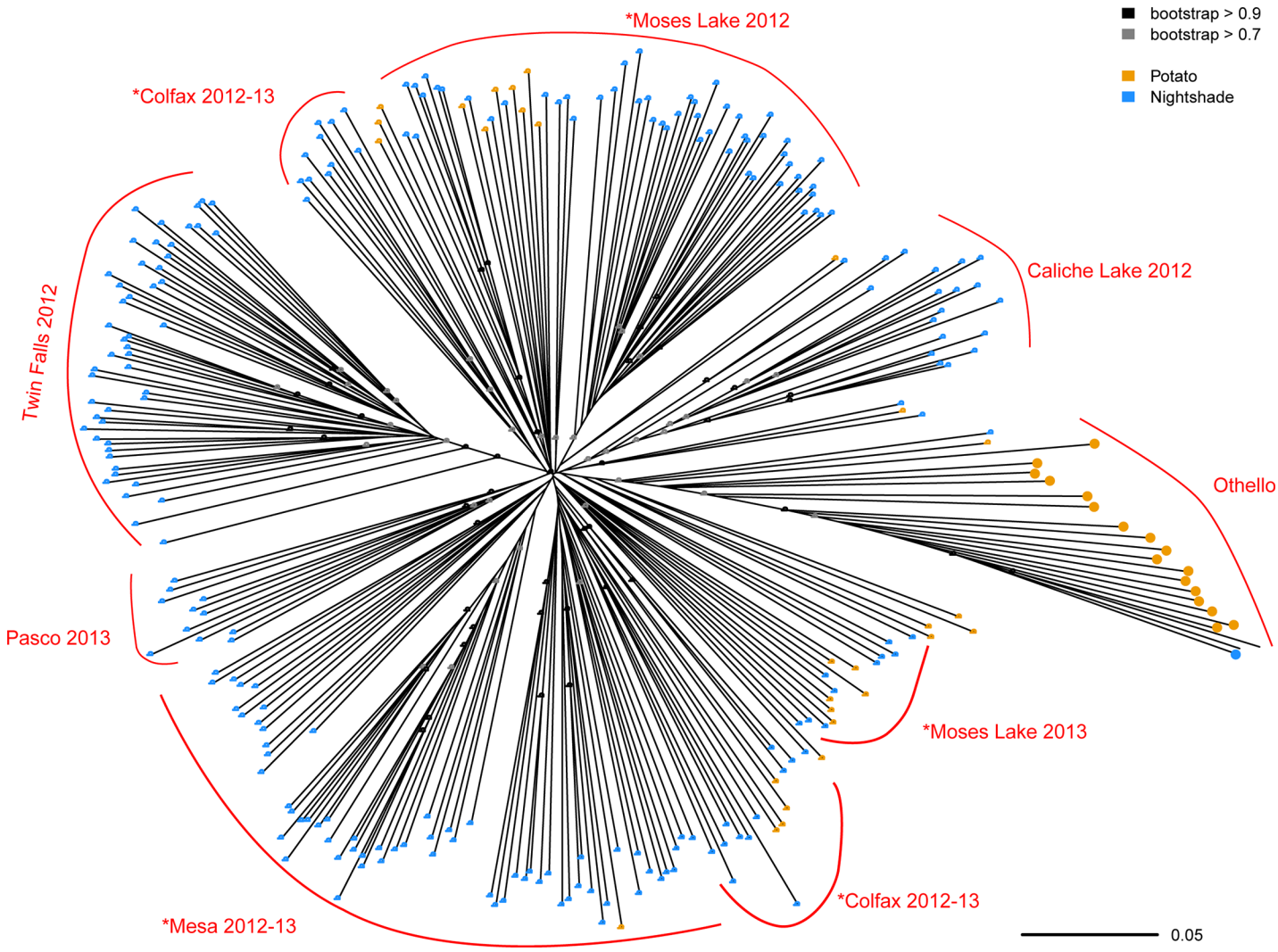

Fig 2. Unrooted neighbor-joining tree for potato psyllids collected from bittersweet nightshade patches (blue circles) or potato fields (yellow circles). The neighbor-joining tree was constructed using proportion of shared alleles. Psyllids from Potato-ML and Potato-Patterson were not labeled as they intermixed with psyllids from nightshade sites. An asterisk (*) indicates three nightshade locations with samples spanning two years. Samples with enlarged circles were grouped into a genetically distinct group in ADMIXTURE (shown in purple bars in Fig 3).

https://doi.org/10.1371/journal.pone.0177742.g002

genetically distinct (Figs 2 and 3). This grouping persisted through higher K values. At $K = 4$ some of the psyllids collected from bittersweet nightshade at Mesa and Caliche Lake were placed in a separate group; there was also potential genetic turnover for potato psyllids collected at the Moses Lake and Mesa sites between the two years during which those bittersweet nightshade patches were sampled, whereas the Colfax site exhibited constant genetic makeup across the two years (Fig 3). K values of 5 through 7 identified relatively modest genetic divisions within sites and host-plant species. We found that running ADMIXTURE with $K = 9$ minimized the cross-validation error (S2 Fig). However, there were three and five grouping patterns among the 50 runs at $K = 8$ and $K = 9$, respectively, none of the patterns represented the majority of the runs, and many individuals were highly admixed (S3 Fig). Thus, the biological interpretation of the higher K values was not obvious and may not be very informative for understanding contemporary psyllid population structure.

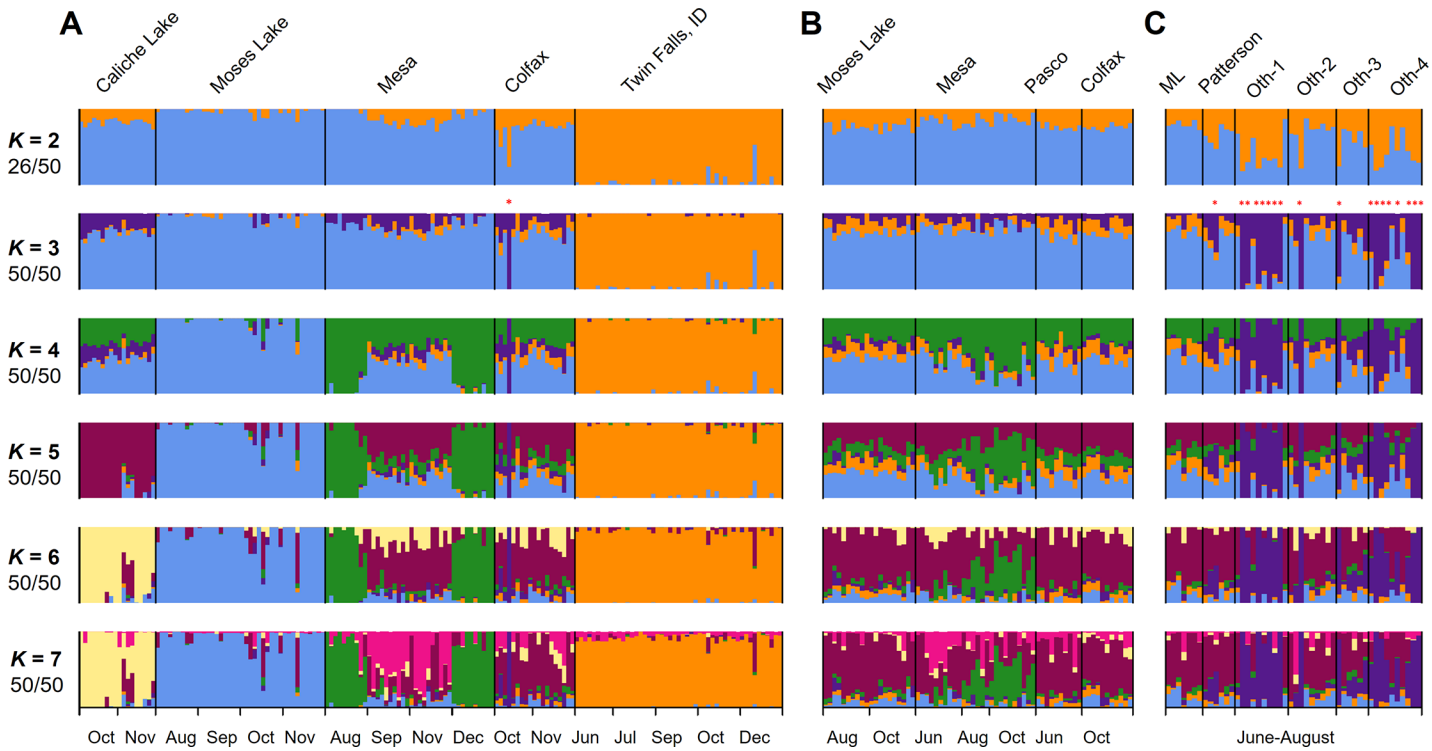

Fig 3. Estimated ancestry of potato psyllids collected from bittersweet nightshade patches and potatoes. Potato psyllids were collected from bittersweet nightshade patches in (A) 2012, (B) 2013, and from potato fields in (C) 2013. The number of ancestral populations (*K*) ranged from *K* = 2 to *K* = 7, and only the grouping pattern that represents the majority (> 50%) of the runs were presented. Numbers below each *K* indicate the number of runs (of 50 runs) showed the representative grouping. Each vertical bar represents a psyllid individual. * indicate "purple cluster" (assigned by majority-rule at *K* = 3), a genetically distinct group which was further analyzed in AMOVA.

https://doi.org/10.1371/journal.pone.0177742.g003

PCA revealed patterns similar to those detected with the other methods. Principal component (PC) 1 separated psyllids from Othello (dark blue triangles, Fig 4A) from those collected in other locations; this further supports the genetic distinctiveness of these psyllids (Fig 4A). Psyllids from Washington clustered together regardless their host plant species, and separated from Twin Falls populations collected in Idaho along PC2 (Fig 4B). Patterns of temporal variation for psyllids from the three nightshade patches with multi-year samples were more or less consistent with ADMIXTURE (Figs 3 and 4); specifically, samples from Mesa collected in different years largely overlapped, and samples from Colfax did not show temporal related separation. Conversely, the 2012 and 2013 samples from Moses Lake exhibited clear separation. Because the separation between the "third" admixture group and the other WA samples seemed much stronger than the separation between potato and nightshade psyllids, or among other sampling sites, we performed additional analyses on the samples assigned to that "third" cluster.

F-statistics and geographic separation

The inbreeding coefficients (F_{IS}) of psyllids collected from potatoes (mean F_{IS} = 0.196) were markedly higher than the F_{IS} of psyllids from bittersweet nightshade patches (mean F_{IS} = 0.08, Wilcoxon signed-rank test, *p*-value = 0.0004, S4 Fig), suggesting smaller effective psyllid populations in potato fields. We then investigated the correlation of population differentiation (F_{ST}) and geographic distance. As described above, we included in these analyses the potato psyllids collected from August 2012 through October 2013, when insects were collected roughly

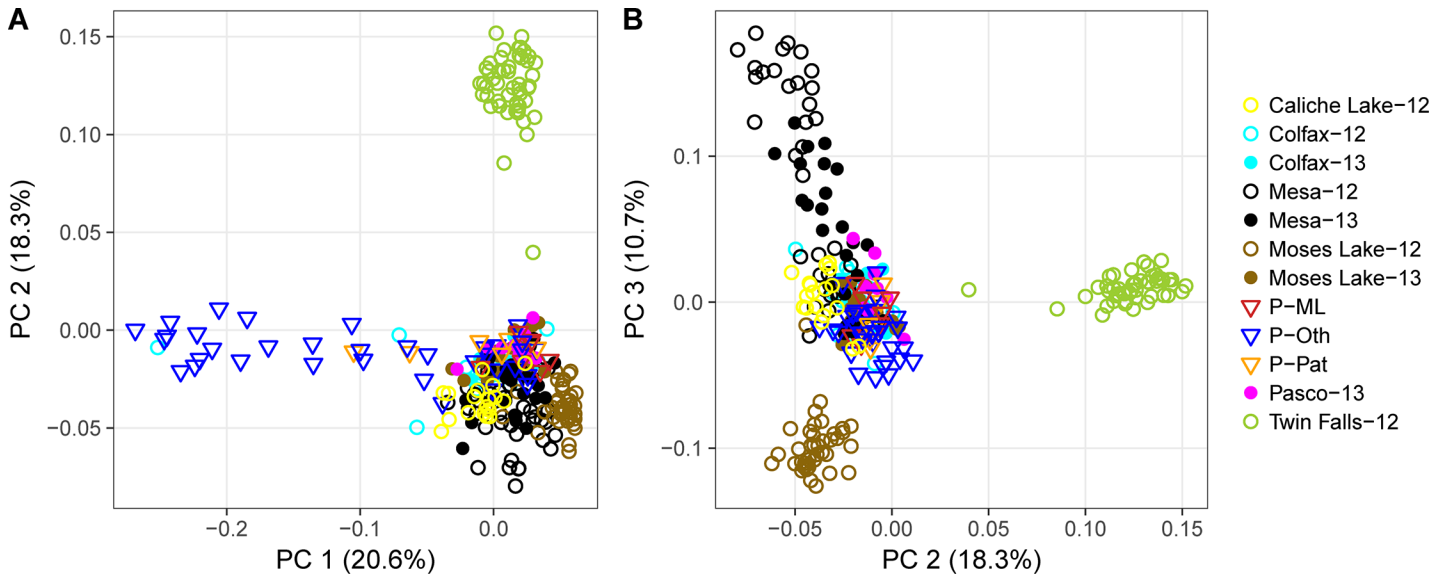

Fig 4. Principal component analysis (PCA) of genotypes of potato psyllids from nightshades and potatoes. (A). Principal component (PC) 1 and PC2. (B). PC2 and PC3. Psyllids were grouped by sampling sites (same color indicate same site), sampling year (non-filled symbol: 2012; filled symbol: 2013) and host plants (circle: nightshades, triangle: potatoes). Percentage accounted for overall genotype variability of each PC was indicated in the axis labels.

https://doi.org/10.1371/journal.pone.0177742.g004

synchronously across all sampled sites. We found a statistically significant relationship between increasing degree of geographic separation and increasingly-large F_{ST} when the single Idaho bittersweet nightshade patch, the most-distant site, was included in the analysis (R^2 = 0.46, df = 40, p-value = 8.8e-07; Fig 5A). However, when that single-most-distant site was dropped from the analysis, this significant relationship disappeared ($R^2 < 0.1$, df = 34, p-value = 0.978; Fig 5B).

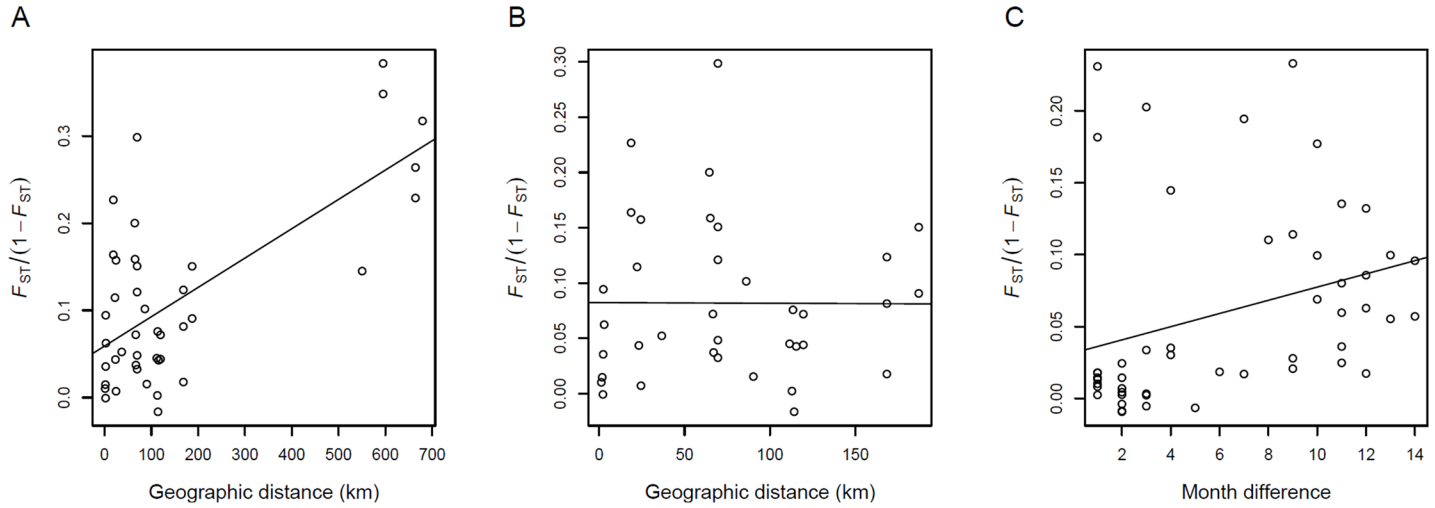

Fig 5. F_{ST} of potato psyllid populations separated spatiotemporally. (A) Regression of F_{ST} versus degree of geographic separation between pairs of psyllid populations differing in distance (but collected during the same month) for all population pairs across sampling dates from August, 2012 through October, 2013, and (B) for the same pairs of populations without the Twin Fall, ID, outgroup. (C) Regression of F_{ST} values for pairwise populations only differing in time (but collected at the same locations) across months.

https://doi.org/10.1371/journal.pone.0177742.g005

We found a weak, but statistically-significant, correlation between within-site F_{ST} values and the duration of time between collections ($R^2 = 0.09$, df = 48, p-value = 0.031; Fig 5C); this suggests increasing genetic differentiation within sites across time. In contrast to the relatively weak overall trend, F_{ST} of populations separated temporally at the Mesa and Moses Lake nightshade sites exhibited relatively large genetic changes through time. At the Mesa site within 2012, the F_{ST} between August and September populations was 0.154, compared to an F_{ST} between September and November of only 0.002 (S3 Table). At the Moses Lake site, the F_{ST} of any pairwise comparison between sampling dates within the same year was ≤ 0.022. In contrast, the F_{ST} of pairwise populations spanning two years was much higher (S4 Table), in agreement with the results of the neighbor-joining tree and ADMIXTURE analyses. Psyllids at the Colfax nightshade site were more residential, as the F_{ST} of temporally-separated populations was consistently low (between 0.013 and 0.024 across the two years).

Psyllids from Othello were evidently distinct from psyllids collected elsewhere (more psyllids were assigned to purple group in Fig 3 $K = 3$). Four populations from this region demonstrated different patterns when compared to psyllids from nightshade populations (S5 Fig). Specifically, F_{ST} of psyllids from Oth-1 versus psyllid populations across all the nightshade sites were generally greater than F_{ST} of Oth-2 versus all nightshade populations (the color was warmer of the Oth-1 column, S5 Fig), though psyllids of Oth-1 and Oth-2 were collected at the same month and two potato fields were only a few kilometers apart. Intriguingly, while comparing four psyllid populations from the Othello region to each other, the divergence was minor except Oth-1 versus Oth-2 (S5 Table).

AMOVA

Because our sampling was hierarchical in space, and because we have several potential sources of genetic structure (space, time, and host plant), we performed AMOVAs to quantify how genetic variation is partitioned by each of these factors. Among the many Washington sites, regardless of whether the top level was host or ADMIXTURE cluster, the largest component of genetic variability was explained at the individual level (Table 2), with little genetic differentiation among sites within genetic clusters and among individuals within sites. Consistent with

Table 2. Results of two Analyses of Molecular Variance (AMOVAs) for samples within Washington. Populations were separated by sampling location and time.

% Variance	Φ-statistics	Component
With host as top level:		
3.8	$\Phi_{host\text{-}total} = 0.04^*$	Between hosts
9.0	$\Phi_{pops\text{-}host} = 0.09^*$	Within hosts, among populations
7.9	$\Phi_{inds\text{-}pop} = 0.09^*$	Within populations, among individuals
79.3	$\Phi_{inds\text{-}total} = 0.21$	Within individual
With ancestry as top level:		
18.3	$\Phi_{ancestry\text{-}total} = 0.18$	Between ADMIXTURE ancestry clusters
7.4	$\Phi_{pops\text{-}ancestry} = 0.09$	Within ancestry clusters, among populations
5.4	$\Phi_{inds\text{-}pop} = 0.07$	Within populations, among individuals
68.8	$\Phi_{inds\text{-}total} = 0.31$	Within individual

Individuals were assigned to the ADMIXTURE cluster ($K = 3$) with the greatest proportion of ancestry.
* indicates a significantly greater variance than the expectation of randomly distributed variation at $p < 0.001$; Significance was not assessed for partitioning by ADMIXTURE cluster because the clusters were identified using the same data the AMOVA was performed on, the p-values would not be meaningful (e.g. [57]).

https://doi.org/10.1371/journal.pone.0177742.t002

the clustering analyses, relatively little genetic variation (~4% of the total) was explained by host plant species, although the variation was significantly greater than zero (note that potato and nightshade populations were not always geographically adjacent). However, the clustering analyses (neighbor-joining, ADMIXTURE, and PCA) indicated that the strongest signal of population structure in WA comes from differentiation between some of the Othello (and one Colfax) samples—the "purple" group that appears at $K = 3$ in the ADMIXTURE analysis—and the other samples. When we made the top level in AMOVA the ADMIXTURE cluster (assigned by majority-rule at $K = 3$), we found that a moderate amount of the genetic variation (~18%) was partitioned by the separation between the "purple" or "third" cluster and the other WA cluster (see Materials and Methods). These results indicate that while overall there is little genetic differentiation between psyllids based on the host plant from which they were collected, the "third" genetic cluster identified by ADMIXTURE (Fig 3) is somewhat differentiated from the other Washington cluster.

Discussion

We took advantage of emerging NextRAD technology to examine interrelatedness of potato psyllids, vectors of a bacterium that causes zebra chip disease [59], collected from two host plant species. The insect has been suggested to overwinter on the perennial weed bittersweet nightshade before colonizing potato crops each summer, although this migratory linkage has never been demonstrated and many other putative non-crop hosts have been proposed [30, 31]. Multiple analyses indicated that the psyllids from bittersweet nightshade and potato crops formed regularly interbreeding populations not clearly separated by host plant (Table 2, Figs 2–4). For example, within our neighbor-joining tree (Fig 2) potato-collected psyllids were generally interspersed among psyllids collected from bittersweet nightshade patches in the same region. Likewise, our ADMIXTURE analysis showed potato- and nightshade-collected psyllids assigned to the same ancestral populations (Fig 3), and our PCA did not show substantial separation by host-plant species (Fig 4). All of these analyses are consistent with the small amount of genetic variance partitioned among host species in an AMOVA (Table 2). Overall, these results suggest that the psyllids found on potatoes during the growing season are very likely persisting on nightshade during the winter. Bittersweet nightshade is common in the Pacific Northwestern US, it grows near bodies of water and along fence lines in large stands [30]. From an applied perspective, this suggests that removal of weedy, invasive bittersweet nightshade plants from the landscape might reduce a key source of potato psyllids eventually colonizing, and perhaps bringing the zebra chip pathogen to, potatoes.

At the same time, there was a second, genetically distinct group of potato psyllids found in four potato fields near Othello, Washington (Fig 1, Table 1 and S1 Table), that strongly differed from these overall patterns. Insects from those fields fell out as a unique clade in our neighbor-joining tree (Fig 2), while clustering analysis suggested that some, though not all, of the Othello samples were genetically distinct from the other psyllids on both potato and nightshade in nearby fields and populations (purple bars in Fig 3). We assessed the magnitude of the genetic differentiation using AMOVAs and F_{ST}: A comparison of AMOVAs run with either host or genetic cluster as the top-level population indicated that about five times more of the genetic variation could be explained by the separation between this unusual group and the other WA psyllids than could be explained by separation between host plant species and about twice as much variation as between sampling sites (Table 2). Furthermore, F_{ST} between the "purple cluster" in ADMIXTURE ($K = 3$–7, Figs 2 and 3) and the other Washington psyllids (0.18–0.20) was greater than that between the other Washington psyllids and those from the distant Idaho site (0.12–0.13, S6 Table), clearly indicating that this group is quite different

from the other psyllids. There are at least two possible explanations for these findings. One is that there is a genetically-isolated sub-population of potato psyllids on bittersweet nightshade plants outside of our sampling network or that the distinct psyllids were moved in from a distant location. An intriguing, second possibility is that a third host plant species is the source of the unique potato-collected insects, with insects on that as-yet-unidentified plant species genetically isolated from those on bittersweet nightshade. It is unlikely that the differentiation we see is simply the result of a barrier to gene flow between Othello and the other sites because Othello is quite close to other sampling sites (Fig 1) and some of the Othello samples (Oth-2 and Oth-3 samples) cluster with the other WA psyllids (Figs 3 and 4, S5 Fig). We note that empirically testing these hypotheses will require sampling additional sites and plant species from Othello and the surrounding region. Many other putative potato psyllid host plant species have been suggested (e.g., other solanaceous weeds species, field bindweed *Convolvulus arvensis*, and matrimony vine *Lycium barbarum*; [31]). From an applied perspective, in turn, pest managers might consider the possibility that suppressing the exotic weed bittersweet nightshade might not entirely suppress regional potato-psyllid populations.

Several lines of evidence suggest that, despite the apparent stability of perennial bittersweet nightshade patches that may persist for decades, psyllids regularly move across the landscape. For example, we observed genetic turnover between (and even within) years at our Moses Lake and Mesa sites, as evidenced by genetic differentiation seen in the ADMIXTURE analysis from $K = 4$ through $K = 7$ (Fig 3). As a more general pattern, for pairings of collections within single sites but separated in time, we found that F_{ST} increased with increasing time between collection dates (Fig 5C). This suggests a general, although relatively modest, turnover in genetic makeup across sites through time that would be consistent with gene flow among sites (although micro-evolutionary adaptation to particular sites could also explain this result; e.g., [60]). Furthermore, we noted no relationship between degree of genetic divergence and geographic distance between sites within Washington (Fig 5B), consistent with a lack of strong barriers to gene flow among these sites. Geographic separation often strongly predicts genetic differentiation (e.g., [61, 62]), as indeed was the case when the most-distant Idaho site was included in analyses (Fig 5A). Perhaps these potato psyllids move readily among sites in the absence of significant physical (e.g., the Blue and Bitterroot mountain ranges) and biological (e.g., the relative dearth of irrigated agriculture) barriers separating the Washington and Idaho sites. It remains unclear if the insects are moving for nutritional reasons (e.g., [10,17,63]), perhaps related to the seasonal drought typical of the region that could render irrigated potato crops more attractive than water-stressed bittersweet nightshade plants. Of course, a wide variety of other biotic (e.g., [64]) and abiotic (e.g., [65]) factors are known to trigger dispersal of pathogen-vectoring herbivores in other systems.

While the findings we present here are specific to one particular plant-pathogen vector and a pair of its host plants, our approach could be widely applicable in other systems. Our study community is typical of many insect-vectored plant diseases, where a detailed understanding of population structure of vectors is critical for understanding, predicting and managing plant disease dynamics. For example, outbreaks of bean leaf roll viruses damaging to leguminous crops often depend upon movement of pea aphid (*Acyrthosiphon pisum*) vectors from alfalfa (*Medicago sativa* L.), a perennial host of both aphid and virus, onto peas [66]. This general movement of aphid between these two host plants has been documented using microsatellite markers in the pea aphid (e.g., [67]). However, sequencing approaches that detail SNPs across the genome, such as the NextRAD, could reveal fine-scaled population structure and thus infer aphid and virus movement among particular fields within a growing region (e.g., Figs 2 and 3). In turn, this degree of resolution could provide field-specific predictions of disease risk that benefit individual land-managers weighing treatment options. It is notable that fine-scale

population structure of vectors is important not just for predicting movement of plant pathogens, but also when highly-mobile vectors spread vertebrate pathogens (e.g., [68]). Indeed, SNPs generated through RAD sequencing already hold promise for understanding relatively small-scale movement patterns of the mosquito *Aedes aegypti*, the vector of Dengue fever and other arboviruses [69]. This demonstrates the broad utility of these approaches for understanding the ecology of vector-transmitted diseases across diverse pathosystems.

The field of population genetics is increasingly making use of "Genotyping by Sequencing" which provides detailed information on genomic variation among individuals and populations. This approach has wide applicability in ecology and evolution, improving our understanding of site-specific adaptive evolution within species [70,71] and evolutionary origins and dispersal patterns of migratory species [72,73], while helping to associate loci with particular phenotypic traits [74,75]. As a powerful and the most-commonly-used approach, RAD sequencing has limitations. Key among these is the reliance on restriction enzyme digestion in the workflow, which limits the approach to use with vertebrates, or relatively large arthropods, from which a sufficiently-large quantity of DNA can be extracted from individuals (e.g., stickleback fish, [70]; land snails, [74]; butterflies, [76]). NextRAD substitutes transposomes for restriction enzymes, necessitating less DNA per sample and thus allowing the approach to be used with small amounts of DNA [32, 33]. This is critical, because relatively small-bodied insects make up a majority of the most injurious herbivores of plants in many natural and agricultural settings (e.g., aphids, fruit flies, and whiteflies), while small arthropods serve as key vectors of some of the most damaging animal and plant pathogens (e.g., mosquitoes, ticks, fleas, aphids, and thrips). Using the NextRAD approach, we were able to describe relatively detailed patterns of population structures of insects in a region. In turn, these patterns suggested local movement patterns of the vectors. We suggest that our work provides a model that may be of value in the many other systems where small-bodied insects move among host plants, and/or vector plant or animal pathogens.

Supporting information

S1 Fig. Workflow of NextRAD sequencing.
(PDF)

S2 Fig. Cross-validation (CV) error and standard error of ADMIXTURE runs.
(PDF)

S3 Fig. Estimated ancestry of potato psyllids from bittersweet nightshade patches and potatoes (K = 2, 8 and 9).
(PDF)

S4 Fig. Inbreeding coefficient of psyllid populations separated by plant hosts.
(PDF)

S5 Fig. Pairwise F_{ST} of psyllids separated on potato (Othello only) and nightshades.
(PDF)

S1 Table. Details of potato psyllid sampling.
(PDF)

S2 Table. Number of SNPs in each filtering step.
(PDF)

S3 Table. Pairwise F_{ST} of potato psyllids from nightshades at the Mesa site.
(PDF)

S4 Table. Pairwise F_{ST} of potato psyllids from nightshades at the Moses Lake site.
(PDF)

S5 Table. Pairwise F_{ST} of potato psyllids from potatoes at Othello.
(PDF)

S6 Table. Median pairwise genetic distance and F_{ST} among groups of psyllids designated following ADMIXTURE $K = 3$.
(PDF)

Acknowledgments

We thank Carmen Blubaugh (Washington State University) and Sanford Eigenbrode (University of Idaho) for providing helpful comments on the manuscript, and Randall Svancara for assistance with cluster computing. The Minnesota Supercomputing Institute (MSI) at the University of Minnesota also provided computing resources for this paper.

Author Contributions

Conceptualization: ZF CICC WES.

Data curation: ZF BE.

Formal analysis: ZF BE QZ AOB.

Funding acquisition: WES ASJ ZF.

Investigation: ZF BE CICC JD WES.

Methodology: ZF BE JLK AOB WES.

Project administration: ZF WES.

Resources: CICC ZF.

Validation: BE JLK AOB ASJ AVK.

Visualization: ZF BE JLK AOB.

Writing – original draft: ZF BE WES.

Writing – review & editing: ZF WES BE JLK QZ AOB JD ASJ AVK.

References

1. Fryxell JM, Sinclair ARE. Causes and consequences of migration by large herbivores. Trends Ecol Evol. 1988; 3: 237–241. https://doi.org/10.1016/0169-5347(88)90166-8 PMID: 21227239
2. Loxdale HD, Lushai G. Slaves of the environment: the movement of herbivorous insects in relation to their ecology and genotype. Philos Trans R Soc B Biol Sci. 1999; 354: 1479–1495.
3. Despland E, Rosenberg J, Simpson SJ. Landscape structure and locust swarming: a satellite's eye view. Ecography. 2004; 27: 381–391.
4. Bischof R, Loe LE, Meisingset EL, Zimmermann B, Van Moorter B, Mysterud A. A migratory northern ungulate in the pursuit of spring: jumping or surfing the green wave? Am Nat. 2012; 180: 407–424. https://doi.org/10.1086/667590 PMID: 22976006
5. Chasen EM, Dietrich C, Backus EA, Cullen EM. Potato Leafhopper (Hemiptera: Cicadellidae) ecology and integrated pest management focused on alfalfa. J Integr Pest Manag. 2014; 5: A1–A8.
6. Shariatinajafabadi M, Wang T, Skidmore AK, Toxopeus AG, Kölzsch A, Nolet BA, et al. Migratory herbivorous waterfowl track satellite-derived green wave index. PLoS ONE. 2014; 9: e108331. https://doi.org/10.1371/journal.pone.0108331 PMID: 25248162

7. Rose DJW. The significance of low-density populations of the African armyworm *Spodoptera exempta* (Walk.). Philos Trans R Soc Lond B Biol Sci. 1979; 287: 393–402.
8. Zhu M, Radcliffe EB, Ragsdale DW, MacRae IV, Seeley MW. Low-level jet streams associated with spring aphid migration and current season spread of potato viruses in the U.S. northern Great Plains. Agric For Meteorol. 2006; 138: 192–202.
9. Peterson MA. Host plant phenology and butterfly dispersal: causes and consequences of uphill movement. Ecology. 1997; 78: 167–180.
10. Wilmshurst JF, Fryxell JM, Farm BP, Sinclair A, Henschel CP. Spatial distribution of Serengeti wildebeest in relation to resources. Can J Zool. 1999; 77: 1223–1232.
11. Monteith KL, Bleich VC, Stephenson TR, Pierce BM, Conner MM, Klaver RW, et al. Timing of seasonal migration in mule deer: effects of climate, plant phenology, and life-history characteristics. Ecosphere. 2011; 2: 1–34.
12. Nault LR. Arthropod transmission of plant viruses: a new synthesis. Ann Entomol Soc Am. 1997; 90: 521–541.
13. Redak RA, Purcell AH, Lopes JRS, Blua MJ, Mizell RF, Andersen PC. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol. 2004; 49: 243–270. https://doi.org/10.1146/annurev.ento.49.061802.123403 PMID: 14651464
14. Weintraub PG, Beanland L. Insect vectors of phytoplasmas. Annu Rev Entomol. 2006; 51: 91–111. https://doi.org/10.1146/annurev.ento.51.110104.151039 PMID: 16332205
15. Power AG. Plant Community Diversity, Herbivore movement, and an insect-transmitted disease of maize. Ecology. 1987; 68: 1658–1669.
16. McElhany P, Real LA, Power AG. Vector preference and disease dynamics: A study of barley yellow dwarf virus. Ecology. 1995; 76: 444–457.
17. Belovsky GE. Food plant selection by a generalist herbivore: the moose. Ecology. 1981; 62: 1020–1030.
18. Illius AW, Clark DA, Hodgson J. Discrimination and patch choice by sheep grazing grass-clover swards. J Anim Ecol. 1992; 61: 183–194.
19. Hagler JR, Jackson CG. Methods for marking insects: current techniques and future prospects. Annu Rev Entomol. 2001; 46: 511–543. https://doi.org/10.1146/annurev.ento.46.1.511 PMID: 11112178
20. Loxdale H d., Lushai G. Molecular markers in entomology. Bull Entomol Res. 1998; 88: 577–600.
21. Behura SK. Molecular marker systems in insects: current trends and future avenues. Mol Ecol. 2006; 15: 3087–3113. https://doi.org/10.1111/j.1365-294X.2006.03014.x PMID: 16968257
22. Wink M. Use of DNA markers to study bird migration. J Ornithol. 2006; 147: 234–244.
23. Zhang D- X. Lepidopteran microsatellite DNA: redundant but promising. Trends Ecol Evol. 2004; 19: 507–509. https://doi.org/10.1016/j.tree.2004.07.020 PMID: 16701315
24. Meglécz E, Anderson SJ, Bourguet D, Butcher R, Caldas A, Cassel-Lundhagen A, et al. Microsatellite flanking region similarities among different loci within insect species. Insect Mol Biol. 2007; 16: 175–185. https://doi.org/10.1111/j.1365-2583.2006.00713.x PMID: 17298557
25. Coates BS, Sumerford DV, Hellmich RL, Lewis LC. Repetitive genome elements in a European corn borer, *Ostrinia nubilalis*, bacterial artificial chromosome library were indicated by bacterial artificial chromosome end sequencing and development of sequence tag site markers: implications for lepidopteran genomic research. Genome Natl Res Counc Can Génome Cons Natl Rech Can. 2009; 52: 57–67.
26. Barley AJ, Monnahan PJ, Thomson RC, Grismer LL, Brown RM. Sun skink landscape genomics: assessing the roles of micro-evolutionary processes in shaping genetic and phenotypic diversity across a heterogeneous and fragmented landscape. Mol Ecol. 2015; 24: 1696–1712. https://doi.org/10.1111/mec.13151 PMID: 25763872
27. Szulkin M, Gagnaire P- A, Bierne N, Charmantier A. Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. Mol Ecol. 2016; 25: 542–558. https://doi.org/10.1111/mec.13486 PMID: 26800038
28. Liefting LW, Sutherland PW, Ward LI, Paice KL, Weir BS, Clover GRG. A New "*Candidatus* Liberibacter" species associated with diseases of Solanaceous Crops. Plant Dis. 2009; 93: 208–214.
29. Greenway G. Economic impact of zebra chip control costs on grower returns in seven US states. Am J Potato Res. 2014; 91: 714–719.
30. Horton DR, Cooper WR, Munyaneza JE, Swisher KD, Echegaray ER, Murphy AF, et al. A new problem and old questions: potato psyllid in the Pacific Northwest. Am Entomol. 2015; 61: 234–244.
31. Horton DR, Cooper WR, Munyaneza J, Thinakaran J, Wohleb CH, Waters TD. Non-potato host plants of potato psyllid in the Pacific Northwest: a year-round complication? Potato Prog Res Ext Potato Ind

Ida Or Wash. 2015;XV. Available from: https://idahopotato.com/uploads/media/potato-pulse/2015-02-23-potatoprogressxv2.pdf

32. Emerson KJ, Conn JE, Bergo ES, Randel MA, Sallum MAM. Brazilian *Anopheles darlingi* root (Diptera: Culicidae) clusters by major biogeographical region. PLOS ONE. 2015; 10: e0130773. https://doi.org/10.1371/journal.pone.0130773 PMID: 26172559

33. Russello MA, Waterhouse MD, Etter PD, Johnson EA. From promise to practice: pairing non-invasive sampling with genomics in conservation. PeerJ. 2015; 3: e1106. https://doi.org/10.7717/peerj.1106 PMID: 26244114

34. Koss AM, Jensen AS, Schreiber A, Pike KS, Snyder WE. Comparison of predator and pest communities in Washington potato fields treated with broad-spectrum, selective, or organic insecticides. Environ Entomol. 2005; 34: 87–95.

35. Crowder DW, Northfield TD, Strand MR, Snyder WE. Organic agriculture promotes evenness and natural pest control. Nature. 2010; 466: 109–112. https://doi.org/10.1038/nature09183 PMID: 20596021

36. Swisher KD, Munyaneza JE, Crosslin JM. High resolution melting analysis of the cytochrome oxidase I gene identifies three haplotypes of the potato psyllid in the United States. Environ Entomol. 2012; 41: 1019–1028.

37. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma Oxf Engl. 2014; 30: 2114–2120.

38. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 2007; 35: D61–D65. https://doi.org/10.1093/nar/gkl842 PMID: 17130148

39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712

40. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009; 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinforma Oxf Engl. 2009; 25: 2078–2079.

42. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–575. https://doi.org/10.1086/519795 PMID: 17701901

43. Whitlock MC, Lotterhos KE. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST}. The American Naturalist. 2015; 186: S24–S36. https://doi.org/10.1086/682949 PMID: 26656214

44. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics. 2008; 9: 323. https://doi.org/10.1186/1471-2105-9-323 PMID: 18662398

45. Foll M, Gaggiotti O. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. Genetics. 2008; 180: 977–993. https://doi.org/10.1534/genetics.108.092221 PMID: 18780740

46. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4: 406–425. PMID: 3447015

47. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 2004; 20: 289–290. PMID: 14734327

48. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009; 19: 1655–1664. https://doi.org/10.1101/gr.094052.109 PMID: 19648217

49. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour. 2015; 15: 1179–1191. https://doi.org/10.1111/1755-0998.12387 PMID: 25684545

50. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38: 904–909. https://doi.org/10.1038/ng1847 PMID: 16862161

51. Nei M. F-statistics and analysis of gene diversity in subdivided populations. Ann Hum Genet. 1977; 41: 225–233. PMID: 596830

52. Weir BS, Cockerham CC. Estimating F-Statistics for the analysis of population structure. Evolution. 1984; 38: 1358–1370.

53. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992; 131: 479–491. PMID: 1644282

54. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Plant Genet Genomics. 2015; 6: 208.

55. Kamvar ZN, Tabima JF, Grünwald NJ. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014; 2: e281. https://doi.org/10.7717/peerj.281 PMID: 24688859

56. Dray S, Dufour AB. The ade4 package: implementing the duality diagram for ecologists. J Stat Softw. 2007; 22: 1–20.

57. Meirmans PG. Seven common mistakes in population genetics and how to avoid them. Mol Ecol. 2015; 24: 3223–3231. https://doi.org/10.1111/mec.13243 PMID: 25974103

58. Reese J, Christenson MK, Leng N, Saha S, Cantarel B, Lindeberg M, et al. Characterization of the Asian citrus psyllid transcriptome. J Genomics. 2014; 2: 54–58. https://doi.org/10.7150/jgen.7692 PMID: 24511328

59. Munyaneza JE. Zebra chip disease of potato: biology, epidemiology, and management. Am J Potato Res. 2012; 89: 329–350.

60. Watt WB, Wheat CW, Meyer EH, Martin J-F. Adaptation at specific loci. VII. Natural selection, dispersal and the diversity of molecular–functional variation patterns among butterfly species complexes (*Colias*: Lepidoptera, Pieridae). Mol Ecol. 2003; 12: 1265–1275. PMID: 12694289

61. Leebens-Mack J, Pellmyr O. Patterns of genetic structure among populations of an oligophagous pollinating yucca moth (*Tegeticula yuccasella*). J Hered. 2004; 95: 127–135. PMID: 15073228

62. Massonnet B, Weisser WW. Patterns of genetic differention between populations of the specialized herbivore *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). Heredity. 2004; 93: 577–584. https://doi.org/10.1038/sj.hdy.6800559 PMID: 15329663

63. Fraser D, Chavez ER, Palohelmo JE. Aquatic feeding by moose: selection of plant species and feeding areas in relation to plant chemical composition and characteristics of lakes. Can J Zool. 1984; 62: 80–87.

64. Long EY, Finke DL. Predators indirectly reduce the prevalence of an insect-vectored plant pathogen independent of predator diversity. Oecologia. 2015; 177: 1067–1074. https://doi.org/10.1007/s00442-014-3194-1 PMID: 25561170

65. Canto T, Aranda MA, Fereres A. Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. Glob Change Biol. 2009; 15: 1884–1894.

66. Clement SL. Pea aphid outbreaks and virus epidemics on peas in the US Pacific Northwest: histories, mysteries, and challenges. Plant Health Prog. 2006;

67. Eigenbrode SD, Davis TS, Adams JR, Husebye DS, Waits LP, Hawthorne D. Host-adapted aphid populations differ in their migratory patterns and capacity to colonize crops. J Appl Ecol. 2016; 53: 1382–1390.

68. Thomas CJ, Cross DE, Bøgh C. Landscape movements of *Anopheles gambiae* malaria vector mosquitoes in rural Gambia. PloS One. 2013; 8: e68679. https://doi.org/10.1371/journal.pone.0068679 PMID: 23874719

69. Rašić G, Filipović I, Weeks AR, Hoffmann AA. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. BMC Genomics. 2014; 15: 275. https://doi.org/10.1186/1471-2164-15-275 PMID: 24726019

70. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. 2010; 6: e1000862. https://doi.org/10.1371/journal.pgen.1000862 PMID: 20195501

71. Keller I, Wagner CE, Greuter L, Mwaiko S, Selz OM, Sivasundar A, et al. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. Mol Ecol. 2013; 22: 2848–2863. https://doi.org/10.1111/mec.12083 PMID: 23121191

72. Larson WA, Utter FM, Myers KW, Templin WD, Seeb JE, Guthrie CM III, et al. Single-nucleotide polymorphisms reveal distribution and migration of Chinook salmon (*Oncorhynchus tshawytscha*) in the Bering Sea and North Pacific Ocean. Can J Fish Aquat Sci. 2012; 70: 128–141.

73. Zhan S, Zhang W, Niitepõld K, Hsu J, Haeger JF, Zalucki MP, et al. The genetics of monarch butterfly migration and warning colouration. Nature. 2014; 514: 317–321. https://doi.org/10.1038/nature13812 PMID: 25274300

74. Richards PM, Liu MM, Lowe N, Davey JW, Blaxter ML, Davison A. RAD-Seq derived markers flank the shell colour and banding loci of the *Cepaea nemoralis* supergene. Mol Ecol. 2013; 22: 3077–3089. https://doi.org/10.1111/mec.12262 PMID: 23496771

75. Takahashi T, Sota T, Hori M. Genetic basis of male colour dimorphism in a Lake Tanganyika cichlid fish. Mol Ecol. 2013; 22: 3049–3060. https://doi.org/10.1111/mec.12120 PMID: 23176589

76. Ebel ER, DaCosta JM, Sorenson MD, Hill RI, Briscoe AD, Willmott KR, et al. Rapid diversification associated with ecological specialization in Neotropical *Adelpha* butterflies. Mol Ecol. 2015; 24: 2392–2405. https://doi.org/10.1111/mec.13168 PMID: 25809206