# Population genetics of the blueberry gall midge, Dasineura oxycoccana (Diptera: Cecidomyiidae), on blueberry and cranberry, with testing invasion scenarios

**Hyojoong Kim** ( hkim@kunsan.ac.kr )
Kunsan National University https://orcid.org/0000-0002-1706-2991
**Cesar Rodriguez-Saona**
Rutgers New Jersey Agricultural Experiment Station
**Heung-Sik Lee**
Animal and Plant Quarantine Agency

**Research Article**

**Keywords:** Cecidomyiidae, Cranberry tipworm, Exotic pests, Host-associated differentiation, Invasion route, Molecular ecology

**Posted Date:** April 12th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1340690/v1

# Abstract

Native to North America, the blueberry gall midge, *Dasineura oxycoccana* (Johnson) (Diptera: Cecidomyiidae), is an economically important insect pest of blueberry and cranberry in its native range (USA and Canada) as well as in other parts of the world. Although two genetically distinct *D. oxycoccana* populations associated with blueberry and cranberry were previously reported, this needs to be revised through further population genetics approaches and a large sample size. Moreover, this pest was recently introduced and spread through other countries in North America, Europe, and Asia. In Korea, *D. oxycoccana* was first found about a decade ago and rapidly spread throughout the country. In this study, we compared the population genetic structure between *D. oxycoccana* populations from blueberry and cranberry and determined genetic relationships among geographical subgroups by genotyping 632 individuals from 31 different populations from the USA (New Jersey, Michigan, and Georgia) and Korea using 12 microsatellite loci. Our population genetic analyses showed a clear separation between the two host-associated *D. oxycoccana* populations from blueberry and cranberry. Using data from only the blueberry-associated *D. oxycoccana* populations, five genetically isolated subgroups were identified. Analysis of the approximate Bayesian computation suggests that the invasive *D. oxycoccana* population from Korea appears to have been introduced from an unsampled source population. Our findings will allow for an easier identification of the source of *D. oxycoccana* into newly invaded regions, as well as to determine their association with blueberry and cranberry, which based on our results can be considered as two distinct species.


# Introduction

The blueberry gall midge, *Dasineura oxycoccana* (Johnson) (Diptera: Cecidomyiidae), is native to central and eastern North America (United States of America and Canada), where the ancestors of cultivated blueberry species, such as highbush blueberry (*Vaccinium corymbosum* L.), lowbush blueberry (*Vaccinium angustifolium* Aiton), and rabbiteye blueberry (*Vaccinium virgatum* Aiton), grow in the wild (Collins and Drummond 2019; Cook et al. 2012; Lyrene and Payne 1992; Roubos and Liburd 2010a, b). In 2004, *D. oxycoccana* was found in eastern US states (Oregon and Washington), likely transferred from infested blueberry nursery plants originated from other blueberry-producing states (Cook et al. 2012; Mathur et al. 2012; Yang 2005). In Canada, *D. oxycoccana* is found in British Columbia, New Brunswick, and Nova Scotia (EPPO 2021), although this species is only considered a secondary pest in the northeastern Canada (Reekie et al. 2009). Like in the United States, movement of this pest within Canada likely occurred through commercial trade of blueberry nursery plants (Cook et al. 2012; Mathur et al. 2012).

*Dasineura oxycoccana* is an economically important pest in its native range (Collins and Drummond 2019), affecting most cultivated blueberry (*Vaccinium* spp.) species, including rabbiteye and highbush blueberries, and causing substantial yield loss by injuring the flower and leaf buds (Lim et al. 2016). In North America, *D. oxycoccana* larval feeding can kill up to 80–90% of the flower buds (Lyrene and Payne 1992; Sampson et al. 2002). In addition, *D. oxycoccana* injury to the leaf buds can result in leaf curl,

stunted growth, and blackened leaf tips (Collins et al. 2010; Hahn and Isaacs 2012). In particular, *D. oxycocanna* is a major pest in Southern US states, such as Florida, where it can injure the flower buds because they are active earlier in the growing season than in the Northern states (Lyrene and Payne 1992). As a result, annual economic losses by *D. oxycoccana* to the Florida blueberry industry can be up to US $20 million (Dernisky et al. 2005).

In North America, there are two host races of *D. oxycoccana.* In addition to blueberries, this pest also feeds on cranberries (*Vaccinium macrocarpon* Aiton) and can be found in most cranberry-growing regions such as Massachusetts and Wisconsin (Collins and Drummond 2019), where it is referred to as the cranberry tipworm (Fitzpatrick et al. 2013; Sampson et al. 2006). Although *D. oxycoccana* populations from cranberry are morphologically similar those from blueberry, they are thought to be two distinct species based on their mating behaviors, genetics, and pheromone composition (Cook et al. 2011; Fitzpatrick et al. 2013; Mathur et al. 2012). These previous studies indicate that *D. oxycoccana* populations from blueberry and cranberry are likely genetically isolated and have likely developed host specificity due to their use of two different host-plant species, resulting in incipient speciation (Cook et al. 2011; Fitzpatrick et al. 2013; Mathur et al. 2012). However, the existence of host-associated differentiation (HAD) between *D. oxycoccana* populations from blueberry and cranberry needs to be further confirmed.

Due to growth in global trade and production of blueberries (Rodriguez-Saona et al. 2019), *D. oxycoccana* is becoming a threat of blueberries across the world. Within North America, it was recently reported in Mexico (Toledo Hernández et al. 2021). Outside of North America, it was first introduced into Europe around 1996 (EPPO 2021); with the first European invasion confirmed in Northern Italy, although the blueberry nursery plants carrying *D. oxycoccana* originated from Germany (Bosio et al. 1998; Prodorutti et al. 2007). In Europe, *D. oxycoccana* is currently distributed in The Czech Republic, Germany, France, Italy, Latvia, Lithuania, The Netherlands, Poland, and Romania (CABI 2022). Its distribution is limited to some regions in the United Kingdom, but it appears widespread in England (EPPO 2021). In Asia, *D. oxycoccana* was confirmed in Japan in 2015 (Yoshida et al. 2017), while no official record has been reported so far in China. In Korea, *D. oxycoccana* was found about a decade ago causing injury to blueberries (Kang et al. 2012; Kim et al. 2015). In recent years, as blueberry cultivation in Korea expanded rapidly from 30 ha in 2005 to 3,369 ha in 2020 (Kang et al. 2012; MAFRA 2022), *D. oxycoccana* has become a serious pest of this crop (Kang et al. 2012; Lim et al. 2016). Since its initial invasion, *D. oxycoccana* has rapidly spread throughout Korea (Kang et al. 2012; Lim et al. 2016). Since this pest is not considered a migratory species that can move through long distances for reproduction (Steck et al. 2000), *D. oxycoccana* populations are expected to be isolated and fragmented regionally in cultivated blueberry farms. Still, little is known on the population genetic structure of this pest in native and invaded regions.

Tracing the source region of an invasive pest and a better understanding of its population genetic structure are important to prevent continuous and multiple introductions, and to develop control strategies such as the importation of the pest's natural enemies (Kim et al. 2021a; Lombaert et al. 2010; Torchin et al. 2003). For instance, several *D. oxycoccana* biotypes are expected to occur in Korea because of the diverse global trade routes of imported blueberry plants (APQA 2022); yet, it remains unknown the origin

and number of biotypes of *D. oxycoccana* in this country. Although *D. oxycoccana* is native to the USA and Canada, it has been in Europe for over two decades and most blueberry nursery plants imported into Korea come from China and Japan (APQA 2022); thus, *D. oxycoccana* populations from Korea could have originated from various sources.

The main objectives of this study were to compare the population genetic structure of *D. oxycoccana* between 1) populations collected from cranberries and blueberries; and 2) a region of origin, i.e., USA, and an invaded region, i.e., Korea. Specifically, we first compared *D. oxycoccana* samples collected from blueberries and cranberries to test for HAD in this species. Second, we characterized the genetic diversity of *D. oxycoccana* populations from USA and Korea and tested for genetic differentiation among the regional populations. For these studies, *D. oxycoccana* samples were collected at the beginning of the Korean invasion and analyzed using 12 microsatellite loci previously developed (Kim et al. 2015). Polymorphic microsatellite loci are an effective tool for studying fundamental questions regarding the population genetics of invasive pests (Ascunce et al. 2011; Behura 2006; Miller et al. 2005). Finally, we inferred the most suspected source of *D. oxycoccana* invasive population from Korean using the standard analysis, approximate Bayesian computation (ABC).

# Materials And Methods

Collection sites

Since none of the *D. oxycoccana* sample collections in this study were carried out in restricted areas, national parks, etc., where permits are required, it is clearly stated that there is no content regarding collection permits. We examined a total of 632 *D. oxycoccana* individuals obtained in 2011–2013 from 31 different population collections from Korea (22 collections) and USA (9 collections) (Fig. 1; Supplementary Material Table S1). Of these, 28 collections were obtained from highbush, lowbush, or rabbiteye blueberries (*Vaccinium* spp.) in Korea and USA, while three collections were obtained from cranberries (*V. macrocarpon*) in USA. Samples from US populations were collected from the central (Michigan), southern (Georgia), and eastern (New Jersey) regions, which comprise the native range of *D. oxycoccana* and possible sources of invasion into Korea. In the invaded region, samples from Korean populations were collected from areas where *D. oxycoccana* occurred during the period of initial invasion (2011–2013). All samples consisted of *D. oxycoccana* larvae, which is the damaging stage living inside the plant tissues and thus having a strong association with its host plant. To avoid sampling related (sibling) *D. oxycoccana* individuals, specimens used in the molecular analyses were collected from different host plants that were distantly located. All freshly (live) collected *D. oxycoccana* larvae used for molecular analyses were carefully removed from infested buds and preserved in 95% or 99% ethanol, and stored at -70°C.

Microsatellite genotyping

A total of 632 *D. oxycoccana* individuals were genotyped using 12 microsatellite loci (Dox08, Dox09, Dox10, Dox11, Dox12, Dox22, Dox23, Dox25, Dox30, Dox33, Dox41, Dox42), which were previously

isolated from this species (Kim et al. 2015). In a preliminary test, all loci developed in the previous study (Kim et al. 2015) were polymorphic among most population samples, and were thus included in the henceforth analyses. Total genomic DNA was extracted from single individuals using LaboPass ™ Tissue Genomic DNA mini Kit (COSMOGENETECH, Daejeon, Korea) according to the manual protocol. All genomic DNA templates were extracted from the whole body of *D. oxycoccana* larvae, which were mostly at the final instar stage. The tissues were left in the lysis buffer with protease *K* solution at 55°C for 24 hours, and then the cleared cuticle was dehydrated. Microsatellite amplifications were performed using AccuPower® PCR PreMix K-2037 (BIONEER, Daejeon, Korea) in 20 µl reaction mixtures containing 0.5 µM forward labeled with a fluorescent dye (6-FAM, HEX, or TAMRA), reverse primers and 0.05 µg of DNA template. Polymerase chain reaction (PCR) was performed using a GS482 thermo-cycler (Gene Technologies, Essex, UK) according to the following procedure: initial denaturation at 95°C for 5 minutes, followed by 34 cycles of 95°C for 30 seconds; annealing at 56°C for 40 seconds; extension at 72°C for 45 seconds, and a final extension at 72°C for 5 minutes. PCR products were visualized by electrophoresis on a .5 % agarose gel with a low range DNA ladder to check for positive amplifications. Automated fluorescent fragment analyses were performed on the ABI PRISM 377 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA), and allele sizes of PCR products were calibrated using the molecular size marker, ROX labeled-size standard (GenScan$^{TM}$ ROX 500, ABI, Waltham, MA, USA). Raw data on each fluorescent DNA products were analyzed using GeneMapper® version 4.0 (ABI, Waltham, MA, USA).

Data analyses

For the 632 individual samples, the results of allele data analyses were processed in GENALEX 6.503 (Peakall, Smouse 2012) through Microsoft office Excel 2019 (Microsoft). We used GENCLONE 2.0 (Arnaud-Haond, Belkhir 2007) to identify multilocus genotypes (MLGs) among *D. oxycoccana* populations (Dorken, Eckert 2001). Observed ($H_O$) and expected heterozygosity ($H_E$) values were estimated on multiple loci using GENEPOP 4.0.7 (Raymond, Rousset 1995) among the population datasets as well as between the regional datasets. Levels of significance for Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium tests were adjusted using sequential Bonferroni correction for all tests involving multiple comparisons (Rice 1989). Deviations from HWE were tested for heterozygote deficiency or excess. MICRO-CHECKER (Oosterhout et al. 2004) was used to test for null alleles (Brookfield 1996) and identify possible scoring errors because of the large-allele dropout and stuttering. The program FSTAT 2.93 (Goudet 2002) was used to estimate the gene diversity ($H_S$), a mean number of alleles ($N_A$), allelic richness ($R_S$) and inbreeding coefficient ($F_{IS}$).

Different groupings were tested independently on the (1) ecological basis (cranberry-associated versus blueberry-associated; ‘case 1’), (2) geographical basis (source versus invasion; *D. oxycoccana* from blueberry only; ‘case 2’), and (3) genetic structure-based groups (A, B, C, D, E; *D. oxycoccana* from blueberry only; ‘case 3’) with analysis of molecular variance (AMOVA) in ARLEQUIN 3.5.1.2 (Excoffier, Lischer 2010), with significance determined using the non-parametric permutation approach described by Excoffier et al. (1992). We also used ARLEQUIN for calculations of pairwise genetic differentiation ($F_{ST}$) values (Weir, Cockerham 1984), in which 31 populations were assigned by each local collection. Exact test of

population differentiation was done as optioned by 100,000 Markov chains, 10,000 Dememorisation steps, and 0.05 significance level.

The program BOTTLENECK 1.2.02 (Piry et al. 1999) was used to identify in our samples the possible effect of a recent bottleneck, separately for each population. Two mutation models, considered appropriate for microsatellites (Cornuet, Luikart 1996; Piry et al. 1999), were applied as the strictly stepwise mutational model (SMM) and the two-phase model (TPM). For the TPM, a model that includes both 90% SMM and 10% TPM was used for 20,000 iterations. Significant deviations in observed heterozygosity over all loci were tested using a nonparametric Wilcoxon signed-rank test (Cornuet, Luikart 1996; Piry et al. 1999).

To examine genetic relationships among 31 *D. oxycoccana* populations, we used principal coordinate analysis (PCoA) on a genetic distance matrix based on codominant-genotypic distance (Peakall et al. 1995; Smouse, Peakall 1999) provided in GENALEX 6.503 (Peakall, Smouse 2012). The PCoA is a multivariate technique that allows to plot and visualize major patterns within a multivariate dataset (e.g., multiple loci and multiple samples) (Peakall et al. 1995; Smouse, Peakall 1999). Plots of the PCoA were independently estimated and drawn based on both individual and population distances.

STRUCTURE 2.3.4 (Pritchard et al. 2000) was used to analyze the genetic structure of 31 *D. oxycoccana* populations using a Bayesian clustering approach. We set the number of clusters (*K*) from 1 to 11 and conducted five independent runs for each value of *K*. Each run consisted of a burn-in period of 30,000 steps, followed by 500,000 Markov chain Monte Carlo (MCMC) repetitions with a model allowing admixture. The $\Delta K$ value calculated as '$\Delta K$ = m(|$L$"($K$)|) / s[$L$($K$)]' was obtained using the ad hoc quantity, which is calculated based on the second-order rate of change of the likelihood (Evanno et al. 2005). To correctly perform this process, $\Delta K$ was calculated using the online resource STRUCTURE HARVESTER 0.6.94 (Earl 2012) that explained the data structure. Visualization of the STRUCTURE results was conducted using DISTRUCT 1.1 (Rosenberg 2004).

In addition, GENECLASS 2 (Piry et al. 2004) was used to perform the assignment/exclusion tests, which were used for the detection of genetic signatures of dispersal and immigration. For each individual of a population, the program estimated the probability of belonging to any other reference population or to be a resident of the population where it was sampled. The sample with the highest probability of assignment was considered as the most likely source for the assigned genotype. We used a Bayesian method of estimating population allele frequencies (Rannala, Mountain 1997) with Monte-Carlo resampling computation (10,000 simulated individuals) to infer the significance of assignments (type I error, alpha = 0.01) (Paetkau et al. 2004).

Approximate Bayesian computation analysis

To estimate the relative likelihood of alternative scenarios of the *D. oxycoccana* invasion, an approximate Bayesian computation (ABC) was performed for microsatellite data, as implemented in DIYABC 2.1.0 (Cornuet et al. 2014). DIYABC allows for the comparison of complex scenarios involving bottlenecks,

serial or independent introductions, and genetic admixture events in introduced populations (Estoup and Guillemaud 2010). The parameters for modeling scenarios are the times of split or admixture events, the stable effective population size, the effective number of founders in introduced populations, the duration of the bottleneck during colonization, and the rate of admixture (Cornuet et al. 2010). The software generates a simulated dataset used to estimate the posterior distribution of parameters to select the most likelihood scenario (Cornuet et al. 2010). DIYABC generates a simulated dataset that is then used to select those most similar to the observed dataset, and so-called selected dataset ($n_\delta$), which are finally used to estimate the posterior distribution of parameters (Cornuet et al. 2008).

The DIYABC analysis was conducted on the purpose of inferring (1) the serial divergence between blueberry and cranberry host races in Analysis #1; (2) the initial introduction process of *D. oxycoccana* from the source (North American) to the invaded (Korean) regions in Analysis #2. Considering the results of PCoA, STRUCTURE and GENECLASS2 (see Results), several populations could be selectively used to estimate the scenarios based on their relationships. In the first ABC analysis that tested the divergence between blueberry and cranberry *D. oxycoccana* populations, Analysis #1, we set one cranberry-associated group (CR) with populations US-C-NJ4, US-C-MA, US-C-WC, and three blueberry-associated groups; New Jersey (USA) group (BBNJ) with populations US-B-NJ1, US-B-NJ2, US-B-NJ3, Georgia (USA) group (BBGA) with populations US-B-GA1, US-B-GA2, and Michigan (USA) group (BBMG) with population US-B-MG. Three scenarios (1—3) were estimated with comparison to each other in the DIYABC (Supplementary Material Fig. S1). In the second ABC analysis (Analysis #2) that tested the introduction process of *D. oxycoccana* from the source (USA) to the invaded region (Korea), we set one source blueberry-associated group (SCNJ) with populations US-B-NJ1, US-B-NJ2, US-B-NJ3, and two invasive blueberry-associated groups; one invasive group A (INVA) with populations KR-B-UW, KR-B-HS1, KR-B-KY, KR-B-CW, KR-B-YD, KR-B-DA, KR-B-IS, KR-B-DJ, KR-B-SC, KR-B-HW, KR-B-NH, KR-B-JJ1, and another invasive group B (INVB) with populations KR-B-GJ, KR-B-HS2, KR-B-PT, KR-B-YS, KR-B-SJ, KR-B-BH1. SCNJ was set as the source group because all populations within this group came from New Jersey (USA), where *D. oxyccocana* originates and thus a likely source of invasion into Korea. Three scenarios (1—3) also were estimated with comparison to each other in the DIYABC (Supplementary Material Fig. S2).

We produced 1,000,000 simulated datasets for each scenario. We used a generalized stepwise model (GSM) as the mutational model for microsatellites, which assumes increases or reductions by single repeat units (Cornuet et al. 2008). To identify the posterior probability of these three scenarios, the $n_\delta$ = 30,000 (1%) simulated datasets closest to the pseudo-observed dataset were selected for the logistic regression, which were similar to the $n_\delta$ = 300 (0.01%) ones for the direct approach (Cornuet et al. 2010). The summary of statistics was calculated from the simulated and observed data for each of the tested scenarios such as the mean number of alleles per locus ($A$), mean genetic diversity for each group and between group, genetic differentiation between pairwise groups ($F_{ST}$), classification index, shared alleles distance ($D_{AS}$), and Goldstein distance.

# Results

## Population genetics analyses

In this study, we genotyped 632 *D. oxycoccana* samples using 12 microsatellite loci, which all were discovered to be non-clonal MLGs, i.e., non-identical genotypes estimated by multiple loci (Table 1). Therefore, in all groups, the number of MLGs were the same as those of individuals in each population. The observed ($H_O$) and expected ($H_E$) heterozygosity values from all 31 populations ranged from 0.617 to 0.898 (averaging 0.764) and from 0.538 to 0.789 (averaging 0.705), respectively (Table 1). According to HWE, there were significant deviations in the KR-B-HE, KR-B-DJ, and KR-B-NH populations by heterozygote excess (Table 1), likely due to heterosis or over-dominance related to selection preference toward heterozygous combination or fixation of heterozygous genotypes. Whereas there were significant deviations in the KR-B-JJ1 and US-B-MG populations by heterozygote deficit (Table 1), likely due to retaining numerous unique genotypes with private alleles within a population related to their relatively high $H_E$ (Delmotte et al. 2003). Gene diversity ($H_S$), the mean number of alleles ($N_A$), and allelic richness ($R_S$) averaged 0.70, 6.44 and 1.70, respectively, and the inbreeding coefficient ($F_{IS}$) was mostly a negative value (-0.10 ± 0.13, mean ± SD). Generally, positive $F_{IS}$ values indicate that some amount of heterozygous offspring in the population decreased, usually due to inbreeding, whereas negative $F_{IS}$ values indicate an increase in heterozygosity due to random mating or outbreeding (Reichel et al. 2016).

Table 1
Summary statistics for microsatellite data from *Dasineura oxycoccana* populations.

| Pop. ID | No. | *Ho* (SD) | *He* (SD) | HWE | $H_S$ | $N_A$ | $R_S$ | $F_{IS}$ |
|---|---|---|---|---|---|---|---|---|
| KR-B-UW | 20 | 0.788(0.051) | 0.740(0.024) | Ns | 0.74 | 6.83 | 1.74 | -0.07 |
| KR-B-GJ | 17 | 0.692(0.068) | 0.727(0.024) | Ns | 0.73 | 6.17 | 1.73 | 0.06 |
| KR-B-HS1 | 23 | 0.803(0.052) | 0.698(0.018) | Ns | 0.70 | 6.75 | 1.70 | -0.15 |
| KR-B-HS2 | 12 | 0.788(0.051) | 0.740(0.024) | Ns | 0.70 | 6.69 | 1.72 | -0.13 |
| KR-B-KY | 20 | 0.792(0.059) | 0.686(0.026) | Ns | 0.68 | 5.25 | 1.69 | -0.16 |
| KR-B-PT | 2 | 0.708(0.114) | 0.653(0.075) | Ns | 0.68 | 2.42 | 1.65 | -0.13 |
| KR-B-HE | 18 | 0.880(0.032) | 0.697(0.024) | *excess | 0.69 | 5.75 | 1.70 | -0.27 |
| KR-B-CW | 20 | 0.817(0.059) | 0.706(0.022) | Ns | 0.70 | 6.00 | 1.71 | -0.16 |
| KR-B-YD | 39 | 0.738(0.052) | 0.705(0.017) | Ns | 0.70 | 8.08 | 1.70 | -0.05 |
| KR-B-DA | 15 | 0.769(0.061) | 0.699(0.022) | Ns | 0.70 | 5.25 | 1.70 | -0.10 |
| KR-B-CA | 20 | 0.821(0.042) | 0.732(0.023) | Ns | 0.73 | 6.17 | 1.73 | -0.13 |
| KR-B-DJ | 9 | 0.898(0.040) | 0.619(0.027) | *excess | 0.60 | 3.67 | 1.62 | -0.49 |
| KR-B-YS | 8 | 0.891(0.035) | 0.719(0.040) | Ns | 0.68 | 4.50 | 1.72 | -0.30 |
| KR-B-IS | 20 | 0.721(0.064) | 0.701(0.026) | Ns | 0.70 | 6.17 | 1.70 | -0.03 |
| KR-B-SC | 20 | 0.809(0.049) | 0.721(0.025) | Ns | 0.72 | 6.67 | 1.72 | -0.13 |
| KR-B-HW | 20 | 0.713(0.072) | 0.649(0.033) | Ns | 0.65 | 3.92 | 1.65 | -0.10 |
| KR-B-BH1 | 13 | 0.788(0.062) | 0.699(0.051) | Ns | 0.69 | 6.08 | 1.70 | -0.14 |
| KR-B-BH2 | 7 | 0.833(0.060) | 0.782(0.026) | Ns | 0.78 | 5.33 | 1.78 | -0.07 |
| KR-B-SJ | 31 | 0.805(0.049) | 0.734(0.029) | Ns | 0.73 | 7.83 | 1.73 | -0.10 |
| KR-B-NH | 20 | 0.825(0.064) | 0.647(0.036) | *excess | 0.64 | 4.83 | 1.65 | -0.29 |
| KR-B-JJ1 | 19 | 0.626(0.081) | 0.720(0.061) | *deficit | 0.72 | 7.08 | 1.72 | 0.13 |
| KR-B-JJ2 | 11 | 0.659(0.077) | 0.714(0.028) | Ns | 0.72 | 5.58 | 1.71 | 0.08 |
| US-B-GA1 | 8 | 0.762(0.072) | 0.729(0.028) | Ns | 0.73 | 4.58 | 1.73 | -0.05 |
| US-B-GA2 | 9 | 0.731(0.059) | 0.689(0.038) | Ns | 0.69 | 5.08 | 1.69 | -0.07 |
| US-B-NJ1 | 4 | 0.681(0.069) | 0.788(0.054) | Ns | 0.81 | 4.42 | 1.79 | 0.16 |
| US-B-NJ2 | 32 | 0.784(0.048) | 0.779(0.030) | Ns | 0.78 | 11.17 | 1.78 | -0.01 |
| US-B-NJ3 | 35 | 0.848(0.046) | 0.789(0.019) | Ns | 0.79 | 9.92 | 1.79 | -0.08 |

| Pop. ID | No. | *Ho* (SD) | *He* (SD) | HWE | $H_S$ | $N_A$ | $R_S$ | $F_{IS}$ |
|---|---|---|---|---|---|---|---|---|
| US-B-MG | 40 | 0.759(0.077) | 0.776(0.033) | *deficit | 0.78 | 11.08 | 1.78 | 0.02 |
| US-C-NJ4 | 40 | 0.697(0.094) | 0.681(0.062) | Ns | 0.68 | 10.83 | 1.68 | -0.02 |
| US-C-MA | 40 | 0.655(0.119) | 0.596(0.090) | Ns | 0.60 | 8.17 | 1.60 | -0.10 |
| US-C-WC | 40 | 0.617(0.119) | 0.538(0.084) | Ns | 0.54 | 7.42 | 1.54 | -0.15 |

No. = number of individuals; *Ho* = observed heterozygosity; *He* = expected heterozygosity; HWE = Hardy–Weinberg Equilibrium; $H_S$ = gene diversity; $N_A$ = mean number of alleles; $R_S$ = allelic richness; $F_{IS}$ = estimates of inbreeding coefficient (Weir and Cockerham (1984); SD = standard deviation; Ns = non-significance ($P > 0.05$); * = significant differences for heterozygote excess or deficit ($P < 0.0001$).

We estimated pairwise genetic differentiation ($F_{ST}$) among the 31 blueberry and cranberry *D. oxycoccana* populations (Supplementary Material Table S2). Pairwise comparisons of $F_{ST}$ values showed that the blueberry populations were largely different genetically from the cranberry populations, of which mean $F_{ST}$ between them was 0.276. The mean $F_{ST}$ within the blueberry populations was 0.120, whereas the mean $F_{ST}$ within the cranberry populations was 0.106. $F_{ST}$ values among the Korean *D. oxycoccana* populations averaged 0.106, while $F_{ST}$ values among the USA populations averaged 0.143. Eight pairwise $F_{ST}$ values, such as KR-B-PT versus KR-B-GJ, were not significant and with very low or negative values (-0.033—0.080), which indicates that genetically they are very similar to each other.

To confirm the molecular variance among the preordained groups, three cases were tested using AMOVA implemented in ARLEQUIN (Excoffier, Lischer 2010; Excoffier et al. 1992). Genetic variance among groups in ‘case 1’ was 14.58%, which suggests that there are relatively large differences between the blueberry and cranberry *D. oxycoccana* populations (Table 2). Excluding cranberry populations, genetic variance among groups in ‘case 2’ was smaller (7.49%) than in ‘case 3’ (10.57%), while the genetic variance among populations within groups in ‘case 2’ (8.34%) was larger than in ‘case 3’ (4.32%). Therefore, STRUCTURE-based groups have more genetic differences among the preordained groups within all blueberry *D. oxycoccana* populations than the otherwise case. These results support the notion that some ‘invasive’ groups in Korea are genetically close to the ‘native’ groups in the USA, regardless of geographic distance.

Table 2
Analysis of molecular variance (AMOVA) of 632 individuals in 31 populations of *Dasineura oxycoccana* in Korea and USA. Case 1: cranberry-associated versus blueberry-associated. Case 2: source (USA) versus invasive (KOR). Case 3: genetic structure-based groups (A, B, C, D, E).

| | **Among groups** | | | **Among populations within groups** | | | **Within populations** | | |
|---|---|---|---|---|---|---|---|---|---|
| Case | $V_a$ | PV | *P* | $V_b$ | PV | *P* | $V_c$ | PV | *P* |
| 1 | 0.82 | 14.58 | <0.0001 | 0.59 | 10.62 | <0.0001 | 4.18 | 74.80 | <0.0001 |
| 2* | 0.38 | 7.49 | <0.0001 | 0.42 | 8.34 | <0.0001 | 4.31 | 84.17 | <0.0001 |
| 3* | 0.54 | 10.65 | <0.0001 | 0.22 | 4.32 | <0.0001 | 4.31 | 85.03 | <0.0001 |
| * only *D. oxycoccana* populations from blueberry included in the analysis. | | | | | | | | | |

Based on the results from BOTTLENECK (Piry et al. 1999), a significant observed heterozygosity excess ($p < 0.05$, one tail) from the Wilcoxon sign-rank tests (both SMM and TPM) was detected in only two *D. oxycoccana* populations, KR-B-HS1 and KR-B-HW, and the shifted mode was observed in the four populations, KR-B-DJ, KR-B-HW, US-B-GA1, US-B-NJ1 (Supplementary Material Table S3). Among them, the population KR-B-HW has apparently undergone genetic bottleneck since it was significance in all analyses; although the bottleneck test should be cautiously interpreted because the sample size for some populations was less than 30 individuals (Cornuet, Luikart 1996).

PCoA was used to create three independent plots of the datasets with and without the cranberry populations. The first PCoA plot based on individual distances included the cranberry populations and show that the blueberry populations, located mostly in three quadrants, and the cranberry populations, located mostly in one quadrant, separated perfectly from each other (Fig. 2). The second PCoA plot, that included the cranberry populations but was based on population distances, also shows that the blueberry populations were largely different from the cranberry ones based on codominant-genotypic distance of multilocus (Supplementary Material Fig. S3). Our third PCoA plot excluded the cranberry populations from the dataset and showed the blueberry populations structured into five genetically isolated subgroups, which are labeled as subgroups A, B, C, D and E (Fig. 3).

In all STRUCTURE analyses, the most likely number of clusters was estimated using the $\Delta K$ calculation based on the Evanno et al. (2005) method (Supplementary Materials Fig. S4). We found that the best value was 985.98 on $\Delta K = 2$ and the second-best value was 64.23 on $\Delta K = 6$ after testing from $K = 1$ to $K = 10$ (Supplementary Materials Table S4). Although $K = 2$ provided the best estimate, the results from $K = 3$ to $K = 6$ were further considered because they most effectively displayed the relationships among *D. oxycoccana* populations when compared with other analyses such as PCoA (Fig. 3). The STRUCTURE result of $K = 2$ for all samples resulted in two clusters (green and red), showing that some of the blueberry populations fall into the red together with the cranberry populations or had mixed assignments of green and red (Fig. 4). Interestingly, the red cluster was almost all converted to blue in the result of $K = 3$, except for the cranberry populations (Fig. 4). A light blue cluster occurred for some US and Korean *D. oxycoccana*

populations in the result of $K$ = 4, and then, a yellow cluster occurred for some Korean populations in the result of $K$ = 5 (Fig. 4). Finally, the STRUCTURE result of $K$ = 6 showed five distinct clusters for the *D. oxycoccana* populations, partitioned into five subgroups (A, B, C, D and E) (Fig. 4), which are consistent with those from the PCoA (Fig. 3).

Based on the assignment test using GENECLASS 2 (Supplementary Materials Table S4), that shows the average probability with which samples were destined to the most likely reference population, the mean values of self-assignment probability in *D. oxycoccana* were 0.397 ± 0.132 (mean ± SD) for the blueberry populations, 0.526 ± 0.047 for the cranberry populations, 0.328 ± 0.140 in the US populations, and 0.415 ± 0.126 for the Korean populations. The highest values of non-self-assignment probability from residence to the expected source between *D. oxycoccana* populations were detected as being most likely in the first KR-B-UW (from five populations), the second KR-B-GJ (from four populations), and the third US-B-MG (from three populations).

Hypothetical invasion scenarios by ABC analysis

Most of the 31 *D. oxycoccana* populations were assigned to three or four subgroups according to the results from the PCoA and STRUCTURE (Figs. 3 and 4; Supplementary Material Fig. S3), in which each of the subgroups formed a larger group containing several populations with similar genetic structure. In addition, we included an 'unsampled' subgroup in the scenarios because one or more introductions could occur from an undetected (= unsampled) population (Guillemaud et al. 2010).

Analysis #1 tested for the simulated comparison of three scenarios to infer the serial divergence between populations from blueberry and cranberry in the native range (Fig. 5; Supplementary Material Fig. S1), and showed that the BBNJ (New Jersey blueberry), BBGA (Georgia blueberry), and BBMG (Michigan blueberry) groups swap positions with each other in a same clade, which indicates that one at the most basal position diverged from the CR (cranberry-associated) group and then the two remaining subgroups diverged from the 'unsampled' group. As a result of Analysis #1, a scenario 2 was estimated to be the most likely of the three scenarios, showing a posterior probability ranging from 0.475 ($n_\delta$ = 3,000) to 0.517 ($n_\delta$ = 30,000) with a 95% CI of 0.404—0.547 and 0.491—0.543, respectively, which predicts that BBGA first diverged from CR, and then BBNJ and BBMG later diverged from an 'unsampled' group (Fig. 5).

Analysis #2 tested for the simulated comparison of the three scenarios to infer the initial introduction process of *D. oxycoccana* from a source (USA) into an invaded (Korea) region (Fig. 5; Supplementary Material Fig. S2). Scenario 1 predicted that INVA (invasive group A) and INVB (invasive group B) subgroups serially branched off from the 'unsampled' group, which diverged from the SCNJ (source) group. Scenario 2 predicted that INVA first diverged from the 'unsampled' group and later the SCNJ and INVB groups serially branched off. Scenario 3 predicted that the SCNJ and INVA groups first diverged from an 'unsampled' group, and later the INVB group arose from an admixture event of both the SCNJ and INVA groups. As a result of Analysis #2, scenario 1 was estimated to be the most likely of the three,

showing a posterior probability ranging from 0.901 ($n_\delta$ = 3,000) to 0.881 ($n_\delta$ = 30,000) with a 95% CI of 0.886–0.927 and 0.873–0.889, respectively (Fig. 5).

# Discussion

Ecological speciation between the two host races in blueberry and cranberry

This study reexamined the possible ecological speciation between two races of *D. oxycoccana* in blueberry and cranberry. We found that two subgroups, consisting of different host-associated populations collected from blueberry and cranberry, clearly separated by our population genetic analyses, i.e., PCoA and STRUCTURE (Figs. 2 and 4; Supplementary Material Fig. S3). These results strongly corroborate previous studies that show ecological speciation between cranberry-associated and blueberry-associated *D. oxycoccana* populations (Cook et al. 2011; Fitzpatrick et al. 2013; Mathur et al. 2012). In an earlier study, Cook et al. (2011) found that *D. oxycoccana* individuals from cranberry and blueberry hosts display complete assortative mating, showing the potential for host race formation or cryptic speciation. In British Columbia, highbush blueberry (*V. corymbosum*) plants bloom several weeks before cranberry (*V. macrocarpon*) plants (Cook et al. 2011), which could lead to different *D. oxycoccana* life cycles and/or behaviors, such as differentiation in phenology, courtship, and pupation sites, resulting in assortative mating among populations from these host plants (Cook et al. 2011). Due to these ecological barriers, *D. oxycoccana* has likely diverged into two distinct species that specialize on two congeneric host plants, such as blueberry and cranberry. Mathur et al. (2012) also revealed that, based on mitochondrial *cytochrome c oxidase subunit I* (*COI*) sequence differences, cryptic speciation occurred between *D. oxycoccana* populations on cranberry and highbush blueberry. This study revealed 10.7–13.1% divergence between cranberry and blueberry *D. oxycoccana* samples on *COI* sequences, whereas little divergence was observed within cranberry (0–1.2%) or blueberry (0–1.3%) sample sequences (Mathur et al. 2012). This genetic difference between the two *D. oxycoccana* populations from blueberry and cranberry seems relatively large when compared to the genetic difference between the two host plants (blueberry versus cranberry), which is ca. 2.7% based on six chloroplast and two mitochondrial SSR loci (Schlautman et al. 2017). Earlier studies by Fitzpatrick et al. (2013) further support HAD in *D. oxycoccana* by showing that populations from blueberry and cranberry produce and respond to different sex pheromones. Thus, based on these previous studies and our current study, it is apparent that cryptic speciation has likely occurred between populations of *D. oxycoccana* on blueberry and cranberry.

Early on, Walsh (1864) proposed an scenario to understand the occurrence of sympatric HAD that could lead to ecological speciation. This scenario proposes that, by switching to new host plants, phytophagous insects exploit novel ecological niches that can result in genetic isolation, subsequently leading to speciation (Matsubayashi et al. 2010). HAD has been shown to cause host shifts mainly in monophagous insects, i.e., individuals of a population switch to a new, related host and then adapt and evolve through ecological isolation to utilize this newly acquired host (Matsubayashi et al. 2010). It is proposed that ecological specialization of host-associated populations can result in species diversification when individuals with greater fitness on one resource preferentially mate with other individuals on that same

resource (Dieckmann, Doebeli 2004), which is often correlated with oviposition site selection (Via 1999; Via et al. 2000). Therefore, mating and oviposition on the same host plant can facilitate genetic differentiation and lead to reproductive isolation among insect populations (Hawthorne and Via 2001; Via and Hawthorne 2002). Once formed, host-associated insect populations can be maintained by ecologically mediated reproductive isolation (Medina et al. 2017), which could potentially account for the origin of species (Matsubayashi et al. 2010). Moreover, certain characteristics within agricultural ecosystems may increase the probability for HAD to occur in insect herbivores (Medina et al. 2014); these include relatively long-standing evolutionary relationships between herbivores and their host plants, which is likely the case for *D. oxycoccana* associations with blueberry and cranberry in its native range.

Interestingly, a previous study also revealed possible HAD between blueberry and cranberry populations of the cranberry fruitworm, *Acrobasis vaccinii* Riley, a frugivorous pest species native to North American that feeds on blueberries and cranberries (Medina et al. 2014). However, although *A. vaccinii* shows allochronic isolation based on distinct phenologies of populations from blueberries and cranberries, genetic differences between host-related populations were not clearly detected in population genetics analyses (Medina et al. 2014). Thus, unlike the previous study by Medina et al. (2014), our results show a clear genetic differentiation between the two host-associated *D. oxycoccana* populations from blueberry and cranberry, which indicates ecological speciation occurred on these host plants. Based on previous studies and this study, we conclude that *D. oxycoccana* from blueberry and cranberry should be considered as two distinct species. Future taxonomic studies are needed to describe *D. oxycoccana* from blueberry and cranberry as separate species.

Genetic structure and fragmentation within *D. oxycoccana*

Although *D. oxycoccana* populations from blueberries and cranberries separated from each other as indicated above, the STRUCTURE analysis of $K = 2$ showed that some blueberry populations appear to be genetically close to the cranberry populations. Among the blueberry subgroups, subgroup D (Figs. 3 and 4) was the closest to the cranberry populations. In particular, two blueberry populations (US-B-GA1 and KR-B-JJ2) possess an intermediate genetic signature that aligns between the two host-associated populations (Fig. 4; Supplementary Material Fig. S3). The DIYABC test also inferred that the blueberry populations from Georgia were closely related to the cranberry populations (Fig. 5). In our analysis, we did not consider the different species of cultivated blueberry or their variety (Supplementary Materials Table S1); thus, future studies need to better explain the intermediate divergence process of *D. oxycoccana* between the two populations from blueberry and cranberry.

Within blueberry populations, five distinct subgroups were detected according to our PCoA and STRUCTURE analyses (Figs. 3 and 4). Although these groups were not separated geographically, they are genetically different due possibly to some ecological isolation factors. In the United States, New Jersey, Georgia, and Michigan populations are separated by considerable genetic distances from each other (Figs. 3 and 4), which appear to be associated to differences in blueberry species or varieties rather than to geographical isolation. Indeed, commercial blueberries are composed of multiple species and their

interspecific hybrids from the *Vaccinium* section *Cyanococcus* A. Gray, including *V. angustifolium*, *V. corymbosum*, *V. virgatum*, and *Vaccinium darrowii* Camp (Schlautman et al. 2017). Blueberries were first domesticated in New Jersey (USA) in 1908 (Coville 1937), and since then several varieties have been developed particularly for larger fruit size, increased concentrations of phytochemicals for improved human health and flavor, broadened phenological adaptations such as reduced chilling requirements, and increased yield (Schlautman et al. 2018). Therefore, in addition to separating populations associated with blueberry or cranberry, *D. oxycoccana* populations could be further separated based on blueberry species and variety, despite its relatively short domestication period. In addition, the possible existence of host specific biotypes needs to be tested in future studies.

*Dasineura oxycoccana*, like other cecidomyiids, has a short adult lifespan (2–3 days) (Roubos and Isaacs 2013; Roubos and Liburd 2010b); thus, it is not considered a migratory species. This suggests that diversity in the genetic structure of populations and geographic isolation can be explained mostly by spread of infested host plant material by humans. The genetic diversity of *D. oxycoccana* populations can also be explained by differences in blueberry species or variety cultivated in a specific area. As in the case of HAD between cranberry and blueberry, our results suggest the possibility of ecological isolation based on their association with specific blueberry species and variety, due to crop features (i.e., chemistry) and phenology, and/or geographical fragmentation due to differences in cultivation environments.

Inferring introductions from source to invaded regions

The recent increase in import and export of goods due to international trade has likely caused the unintentional introductions of invasive insect pests (Hulme 2021; Seebens et al. 2017), as was the case of the *D. oxycoccana* invasion into Korea. In fact, there have been many study cases of trans-pacific introductions by unintentional transportation of exotic insects (Ascunce et al. 2011; Carter et al. 2010; Kim et al. 2017; Kim et al. 2021b), such as the introduction of the imported fire ant, *Solenopsis invicta* (Buren), from the United States to East Asia (Ascunce et al. 2011). On the other hand, the soybean aphid, *Aphis glycines* Matsumura, originated from East Asia and spread to the eastern and central regions of the United States, and was shown to exhibit low genetic variation and diversity in the invaded regions compared to its native range (Kim et al. 2017). Other species that originated in Asia and invaded the United States include the Asian long-horned beetle, *Anoplophora glabripennis* (Motschulsky) (Carter et al. 2010) and more recently the spotted lanternfly, *Lycorma delicatula* (White) (Kim et al. 2021b). In most of these cases, the exotic insects spread through infested host plants or soil into other regions, and there is a tendency to have multiple introductions due to the bridgehead effect, where the initial invasive populations serve as the source of additional invasions via secondary introductions (Ascunce et al. 2011; Kim et al. 2021a; Kim et al. 2021b; Lombaert et al. 2010). Therefore, population genetics studies on invasive species are useful for inferring an introduction from the source to the invaded regions.

In this study, *D. oxycoccana* populations collected from the invaded Korea segregated into the subgroups A, B, D and E (Figs. 3 and 4). In particular, because many individuals from Korea have similar dominant genotypes found in the subgroups A and B, it is likely that blueberry plants of a specific species or variety

were introduced from the same US source into Korea (Figs. 3 and 4). On the other hand, some populations such as KR-B-JJ2 (subgroup D) and KR-B-CA (subgroup E) were apparently close to US populations from Georgia and Michigan, respectively, based on the results from PCoA and STRUCTURE (Figs. 3 and 4). In addition, the average $F_{st}$ value among the *D. oxycoccana* populations from Michigan, New Jersey, and Georgia was 0.136, whereas those between KR-B-JJ2 and Georgia populations or between KR-B-CA and Michigan populations were rather lower (0.129 and 0.036, respectively) (Supplementary Materials Table S2), which suggests a highly regional association between the source and the invasive *D. oxycoccana* populations.

Although the *D. oxycoccana* populations of subgroups A and B from Korea are most similar to those from New Jersey in the eastern United States, our data are insufficient to conclusively say that the genetic structure of the invasive populations in Korea is related to a specific US region. Nevertheless, based on the DIYABC analysis, it was determined that the possibility of independent introductions of subgroups A and B was high, estimating a scenario that included an unsampled population (Fig. 5), indicating that *D. oxycoccana* from Korea likely came from an unknown region (unsampled population), in the United States or from a country in Europe or Asia, and that the invasion of subgroup B occurred earlier than that of subgroup A (Fig. 5; Supplementary Material Fig. S2). These results suggest that *D. oxycoccana* populations in Korea spread by multiple introductions with genetic origins from at least four independent genotypes (subgroups A, B, D, and E) (Figs. 4 and 5), which strongly suggests that they were introduced by humans through imported blueberry nursery plants. In fact, even within North America, this species is likely moved across regions with the transport of infested host-plant nursery plants (Cook et al. 2012; Mathur et al. 2012). Moreover, the lack of a regionally based genetic structure indicates that *D. oxycoccana* populations in the invaded Korea are genetically similar.

# Conclusion

Our results demonstrate: 1) a clear separation between the two host-associated *D. oxycoccana* populations from blueberry and cranberry, which could be considered distinct species; 2) the occurrence of five genetically isolated *D. oxycoccana* subgroups from blueberry; and 3) that the invasive *D. oxycoccana* population from Korea was likely introduced from an unsampled source population. These findings will help to better identify *D. oxycoccana* populations associated with their host plant (i.e., blueberry or cranberry). They will also facilitate the identification of the source of *D. oxycoccana* into newly invaded regions, which may help prevent a bridgehead effect in these invaded regions.

# Declarations

**Acknowledgments** We thank the assistance of Vera Kyryczenko-Roth, who helped with the collection of *Dasineura oxycoccana*. Special thanks to Dr. Taek-Jun Kang for his dedication and hard work, rest in peace in heaven.

**Funding** This study was supported by a grant (Project Code No. Z-1541785-2012-12-01) from Animal and Plant Quarantine Agency, Ministry of Agriculture, Food and Rural Affairs, Republic of Korea in 2012-13. This research was also supported by a grant from the Korea Environment Industry & Technology Institute (KEITI) through Exotic Invasive Species Management Program (2018002270005), funded by Korea Ministry of Environment (MOE), and by a Hatch project no. NJ08140 to CR-S

**Authors' contributions** Conceptualization: HK; Methodology: HK; Software: HK; Funding acquisition: HK, CR-S; Validation: HK, H-SL; Formal analysis: HK; Investigation: HK, CR-S; Resources: HK, CR-S, H-SL; Data curation: HK; Writing—original draft preparation: HK; Writing—review and editing: HK, CR-S, H-SL; Visualization: HK; Supervision: HK; Project administration: HK, CR-S; Funding acquisition: HK, CR-S, H-SL. All authors read and approved the final manuscript.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Availability of data and material** Data are available upon request from the authors.

# References

1. APQA (2022) Animal and Plant Quarantine Agency, Republic of Korea, Plant quarantine online complaint system (https://okminwon.pqis.go.kr/minwon/information/statistics.html) [access date on: 01/14/22]
2. Arnaud-Haond S, Belkhir K (2007) GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. Mol Ecol Notes 7:15–17
3. Ascunce MS, Yang C-C, Oakey J et al (2011) Global Invasion History of the Fire Ant *Solenopsis invicta*. Science 331:1066–1068
4. Behura SK (2006) Molecular marker systems in insects: current trends and future avenues. Mol Ecol 15:3087–3113
5. Bosio G, Bogetti C, Brussino G et al (1998) *Dasineura oxycoccana*, a new pest of highbush blueberry in Italy. Informatore Fitopatol 11:36–41
6. Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455
7. CABI (2022) Invasive Species Compendium. Wallingford, UK: CAB International. . [access date on: 01/11/22]
8. Carter M, Smith M, Harrison R (2010) Genetic analyses of the Asian longhorned beetle (Coleoptera, Cerambycidae, *Anoplophora glabripennis*), in North America, Europe and Asia. Biol Invasions 12:1165–1182
9. Collins DW, Harris KM, Glover RH (2010) *Dasineura oxycoccana* (Johnson, 1899) (Diptera, Cecidomyiidae), a pest of cultivated blueberres, new to Britain. Dipterists Digest 17:53-59
10. Collins JA, Drummond FA (2019) The Blueberry Gall Midge (Diptera: Cecidomyiidae): A Recent Pest of Wild Blueberry (*Vaccinium angustifolium*; Ericales: Ericaceae) and Its Impact on Potential Yield. J

Econ Entomol 112:1151–1161

11. Cook MA, Fitzpatrick SM, Roitberg BD (2012) Phenology of *Dasineura oxycoccana* (Diptera: Cecidomyiidae) on cranberry and blueberry indicates potential for gene flow. J Econ Entomol 105:1205–1213
12. Cook MA, Ozeroff SN, Fitzpatrick SM et al (2011) Host-associated differentiation in reproductive behaviour of cecidomyiid midges on cranberry and blueberry. Entomol Exp Appl 141:8–14
13. Cornuet J-M, Pudlo P, Veyssier J et al (2014) DIYABCv2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism, DNA sequence and microsatellite data. Bioinformatics (Oxford, England) 30
14. Cornuet JM, Luikart G (1996) Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data. Genetics 144:2001–2014
15. Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). Bmc Bioinformatics 11
16. Cornuet JM, Santos F, Beaumont MA et al (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. Bioinformatics 24:2713–2719
17. Coville FV (1937) Improving the wild blueberry. Yearbook of the United States Department of Agriculture. U.S. Dept. Agri, Washington, D.C., pp. 559–574
18. Delmotte F, Sabater-Munoz B, Prunier-Leterme N et al (2003) Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species. Evolution 57:1291–1303
19. Dernisky AK, Evans RC, Liburd OE et al (2005) Characterization of early floral damage by cranberry tipworm (*Dasineura oxycoccana* Johnson) as a precursor to reduced fruit set in rabbiteye blueberry (*Vaccinium ashei* Reade). Int J Pest Manage 51:143–148
20. Dieckmann U, Doebeli M (2004) Adaptive dynamics of speciation: sexual populations. In: Dieckmann U, Doebeli M, Metz JAJ, Tautz D (eds) Adaptive Speciation. Cambridge University Press, Cambridge, UK, pp 76–111
21. Dorken ME, Eckert CG (2001) Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodonverticillatus* (Lythraceae). J Ecol 89:339–350
22. Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361
23. EPPO (2021) EPPO, 2021. EPPO Global database. In: EPPO Global database, Paris, France: EPPO. https://gd.eppo.int/ [access date on: 01/11/22]
24. Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? Mol Ecol 19:4113–4130
25. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620
26. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

27. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491
28. Fitzpatrick SM, Gries R, Khaskin G et al (2013) Populations of the Gall Midge *Dasineura oxycoccana* on Cranberry and Blueberry Produce and Respond to Different Sex Pheromones. J Chem Ecol 39:37–49
29. Goudet J (2002) FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices (version 2.9.3.2). Institute of Ecology, University of Lausanne, Switzerland. Available from http//www.unil.ch/ izea/softwares/fstat.html
30. Guillemaud T, Beaumont MA, Ciosi M et al (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. Heredity 104:88–99
31. Hahn NG, Isaacs R (2012) Distribution and phenology of *Dasineura oxycoccana* (Diptera: Cecidomyiidae) in Michigan blueberries. Environ Entomol 41:455–462
32. Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412:904–907
33. Hulme PE (2021) Unwelcome exchange: International trade as a direct and indirect driver of biological invasions worldwide. One Earth 4:666–679
34. Kang T-J, Cho M-R, Ahn S-J et al (2012) Monitoring and Environment-friendly Management of Blueberry Gall Midge, *Dasineura oxycoccana* (Diptera: Cecidomyiidae), on Blueberry. Korean J Organic Agri 20:607–618
35. Kim H, Hoelmer KA, Lee S (2017) Population genetics of the soybean aphid in North America and East Asia: test for introduction between native and introduced populations. Biol Invasions 19:597–614
36. Kim H, Kim S, Kim S et al (2021a) Population Genetics for Inferring Introduction Sources of the Oriental Fruit Fly, *Bactrocera dorsalis*: A Test for Quarantine Use in Korea. Insects 12:851
37. Kim H, Kim S, Lee Y et al (2021b) Tracing the Origin of Korean Invasive Populations of the Spotted Lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae). Insects 12:539
38. Kim H, Rodriguez-Saona C, Kwon DH et al (2015) Development and characterization of 12 microsatellite loci from the blueberry gall midge *Dasineura oxycoccana* (Diptera: Cecidomyiidae). Applied Entomology and Zoology 50:415-418
39. Lim J-R, Kim E, Moon H-C et al (2016) Patterns of Insect Pest Occurrences and *Dasineura oxycoccana* Johnson in Blueberry Farms in Jeonbuk Province. Korean J Appl Entomol 55:45–51
40. Lombaert E, Guillemaud T, Cornuet J-M et al (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. PLoS ONE 5:e9743–e9743
41. Lyrene PM, Payne JA (1992) Blueberry gall midge; a pest on rabbiteye blueberry in Florida. Proc. Fla. State Hort. Soc. 105:297-300
42. MAFRA (2022) Ministry of Agriculture, Food and Rural Affairs, Republic of Korea, Agricultural business registration information service (https://edu.agrix.go.kr/uni_docs7/biOlap/dashBoard.do)

[access date on: 01/11/22]

43. Mathur S, Cook MA, Sinclair BJ et al (2012) DNA Barcodes Suggest Cryptic Speciation in *Dasineura Oxycoccana* (Diptera: Cecidomyiidae) on Cranberry, *Vaccinium Macrocarpon*, and Blueberry, *V. Corymbosum*. Fla Entomol 95:387–394
44. Matsubayashi KW, Ohshima I, Nosil P (2010) Ecological speciation in phytophagous insects. Entomol Exp Appl 134:1–27
45. Medina RF, Dickey AM, Harrison K et al (2017) Host-associated differentiation in a pecan and water hickory Aphidomorpha community. Entomol Exp Appl 162:366–378
46. Medina RF, Szendrei Z, Harrison K et al (2014) Exploring host-associated differentiation in the North American native cranberry fruitworm, Acrobasis vaccinii, from blueberries and cranberries. Entomol Exp Appl 150:136–148
47. Miller N, Estoup A, Toepfer S et al (2005) Multiple transatlantic introductions of the western corn rootworm. Science 310:992–992
48. Oosterhout Cv, Hutchinson WF, Wills DPM et al (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538
49. Paetkau D, Slade R, Burden M et al (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol 13:55–65
50. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539
51. Peakall R, Smouse PE, Huff DR (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss Buchloë dactyloides. Mol Ecol 4:135–148
52. Piry S, Alapetite A, Cornuet J-M et al (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. J Hered 95:536–539
53. Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90:502–503
54. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
55. Prodorutti D, Pertot I, Giongo L et al (2007) Highbush Blueberry: Cultivation, Protection, Breeding and Biotechnoligy the European. J Plant Sci Biotechnol 1:44–56
56. Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences 94:9197-9201
57. Raymond M, Rousset F (1995) GENEPOP (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. J Hered 86:248–249
58. Reekie M, Mackenzie K, Lees B (2009) The biology and pest potential of cranberry tipworm (Diptera, Cecidomyiidae) on lowbush blueberry. Acta Hort 810:401–410

59. Reichel K, Masson J-P, Malrieu F et al (2016) Rare sex or out of reach equilibrium? The dynamics of *FIS* in partially clonal organisms. BMC Genet 17:76–76
60. Rice WR (1989) Analyzing Tables of Statistical Tests. Evolution 43:223–225
61. Rodriguez-Saona C, Vincent C, Isaacs R (2019) Blueberry IPM: Past Successes and Future Challenges. Annu Rev Entomol 64:95–114
62. Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Resour 4:137–138
63. Roubos CR, Isaacs R (2013) Blueberry Gall Midge. Michigan Bluebery Facts, Department of Entomology, Michigan State University. Extension Bulletin E3191
64. Roubos CR, Liburd OE (2010a) Evaluation of emergence traps for monitoring blueberry gall midge (Diptera: Cecidomyiidae) adults and within field distribution of midge infestation. J Econ Entomol 103:1258–1267
65. Roubos CR, Liburd OE (2010b) Pupation and Emergence of Blueberry Gall Midge, *Dasineura oxycoccana* (Diptera: Cecidomyiidae), Under Varying Temperature Conditions. Fla Entomol 93:283–290
66. Sampson BJ, Rinehart TA, Liburd OE et al (2006) Biology of Parasitoids (Hymenoptera) Attacking *Dasineura oxycoccana* and *Prodiplosis vaccinii* (Diptera: Cecidomyiidae) in Cultivated Blueberries. Ann Entomol Soc Am 99:113–120
67. Sampson BJ, Stringer SJ, Spiers J (2002) Integrated Pest Management for *Dasineura oxycoccana* (Diptera: Cecidomyiidae) in Blueberry. Environ Entomol 31:339–347
68. Schlautman B, Covarrubias-Pazaran G, Fajardo D et al (2017) Discriminating power of microsatellites in cranberry organelles for taxonomic studies in *Vaccinium* and Ericaceae. Genet Resour Crop Evol 64:451–466
69. Schlautman B, Diaz-Garcia L, Covarrubias-Pazaran G et al (2018) Comparative genetic mapping reveals synteny and collinearity between the American cranberry and diploid blueberry genomes. Mol Breeding 38:9
70. Seebens H, Blackburn TM, Dyer EE et al (2017) No saturation in the accumulation of alien species worldwide. Nat Commun 8:14435
71. Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82:561–573
72. Steck G, Lyrene P, Payne J (2000) Blueberry Gall Midge. *Dasineura oxycoccana* (Johnson) (Insecta: Diptera: Cecidomyiidae). IFAS EDIS Ext. Publ. EENY-136. University of Florida, Gainesville, FL
73. Toledo Hernández RA, Mikery O, Ibañez S et al (2021) First Record of Invasive Pest Blueberry Gall Midge, *Dasineura oxycoccana* (Johnson) in Mexico: Molecular and Morphological Confirmation.46%J Southwestern Entomologist:963–970, 8
74. Torchin ME, Lafferty KD, Dobson AP et al (2003) Introduced species and their missing parasites. Nature 421:628–630

75. Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. Evolution 53:1446–1457
76. Via S, Bouck AC, Skillman S (2000) Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. Evolution 54:1626–1637
77. Via S, Hawthorne DJ (2002) The Genetic Architecture of Ecological Specialization: Correlated Gene Effects on Host Use and Habitat Choice in Pea Aphids. Am Nat 159:S76–S88
78. Walsh BD (1864) On phytophagic varieties and phytophagic species. Proc. Entomol. Soc. Phila. 3:403-430
79. Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution 38:1358–1370
80. Yang WQ (2005) Blueberry gall midge. A possible new pest in the Northwest.Oregon State Univ. Extension Service Fact Sheet EM8889
81. Yoshida K, Elsayed AK, Matsuura H et al (2017) Occurrence of *Dasineura oxycoccana* (Diptera: Cecidomyiidae) infesting cultivated blueberry in Japan. Japanese J Appl Entomol Zool 61:25–28

# Figures

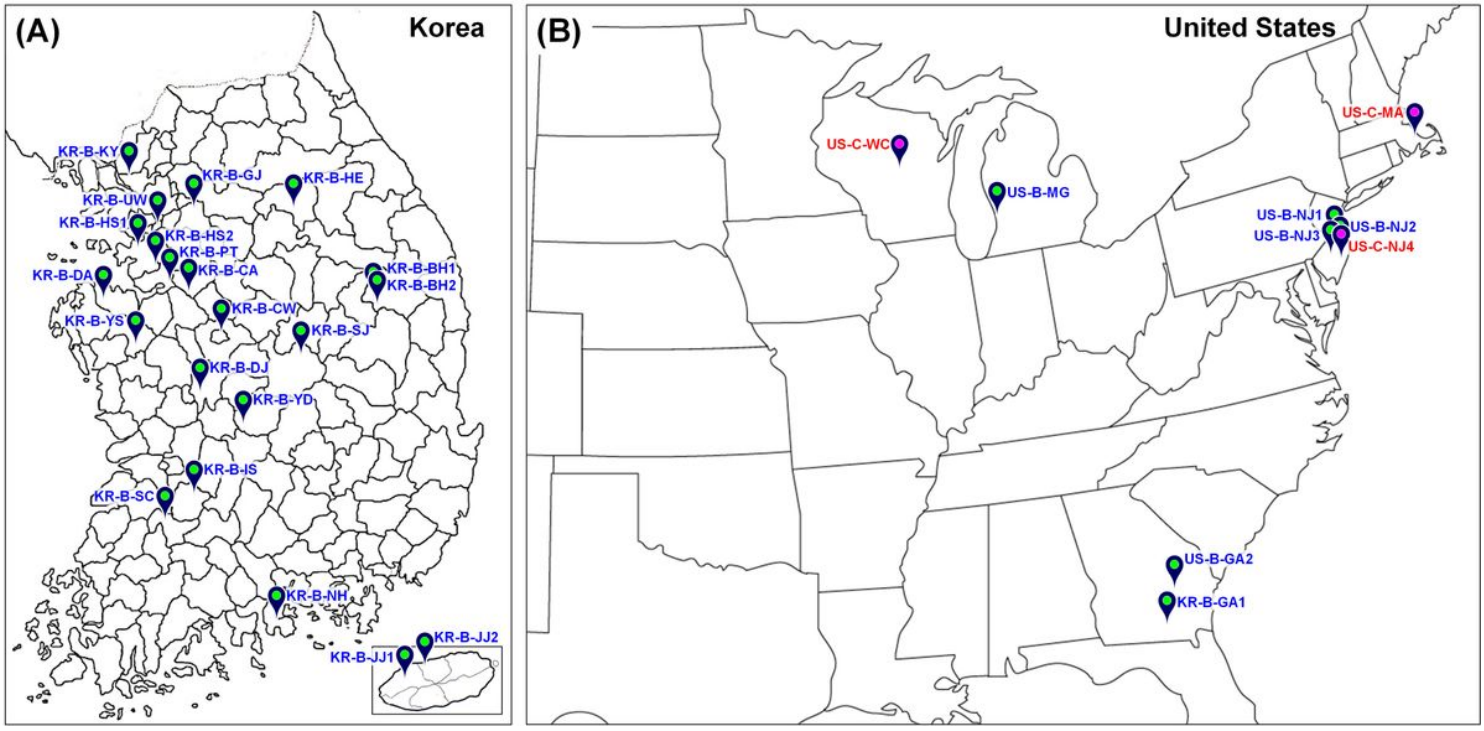
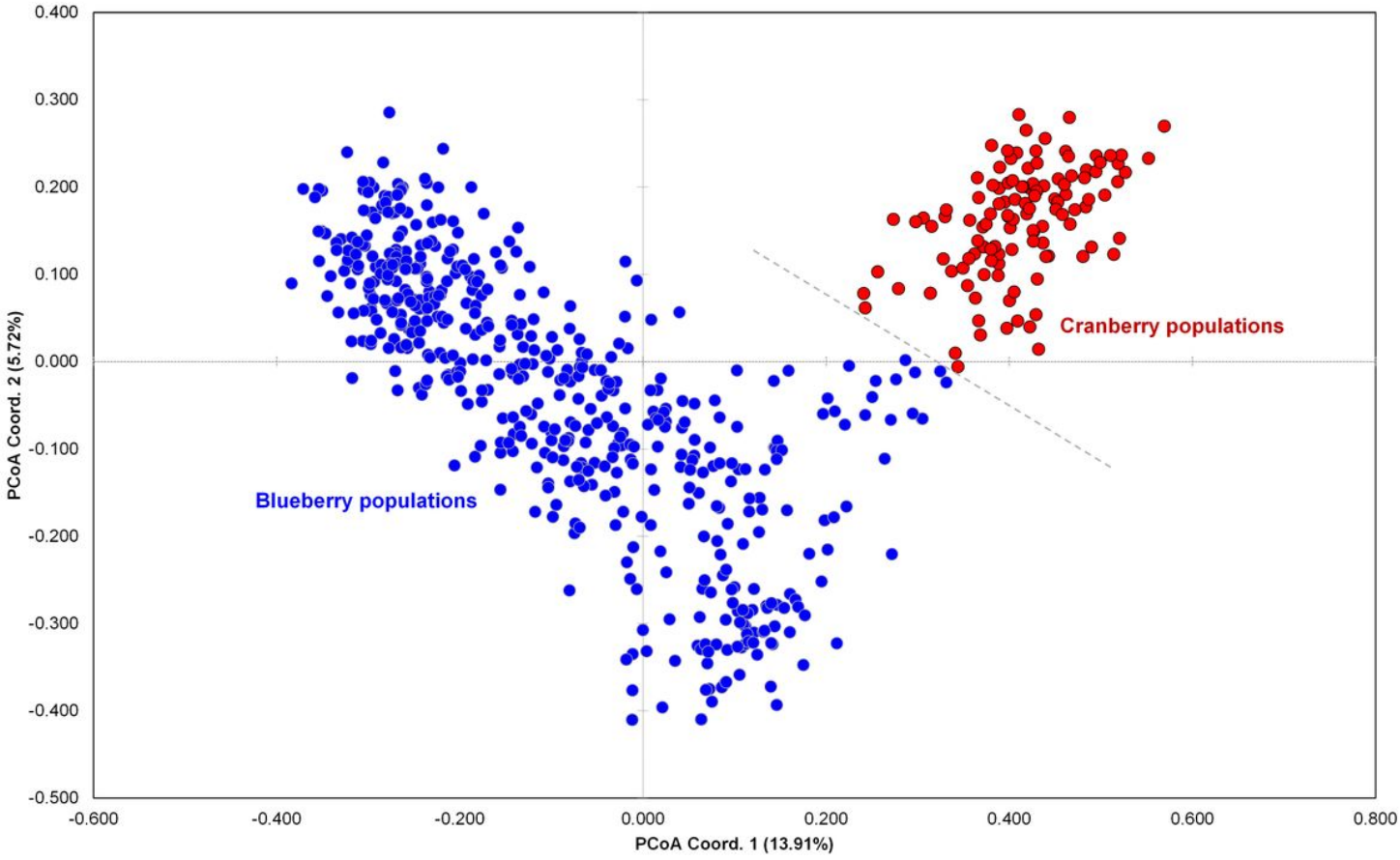


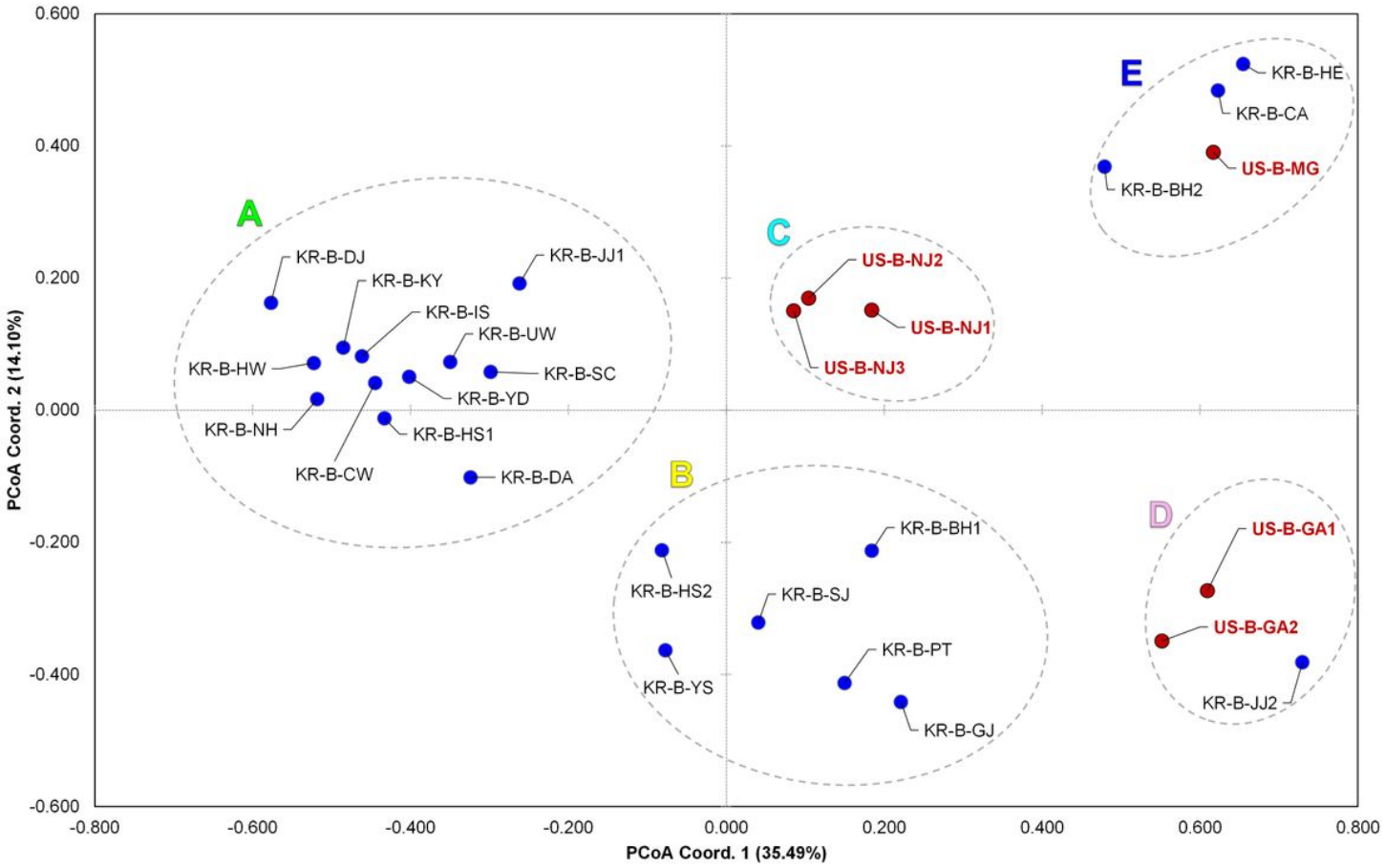

**Figure 1**

Sampling collection sites of *Dasineura oxycoccana.* (A) South Korea (B) USA. Detailed information of locations is described in Supplementary Materials 2, Table S1.



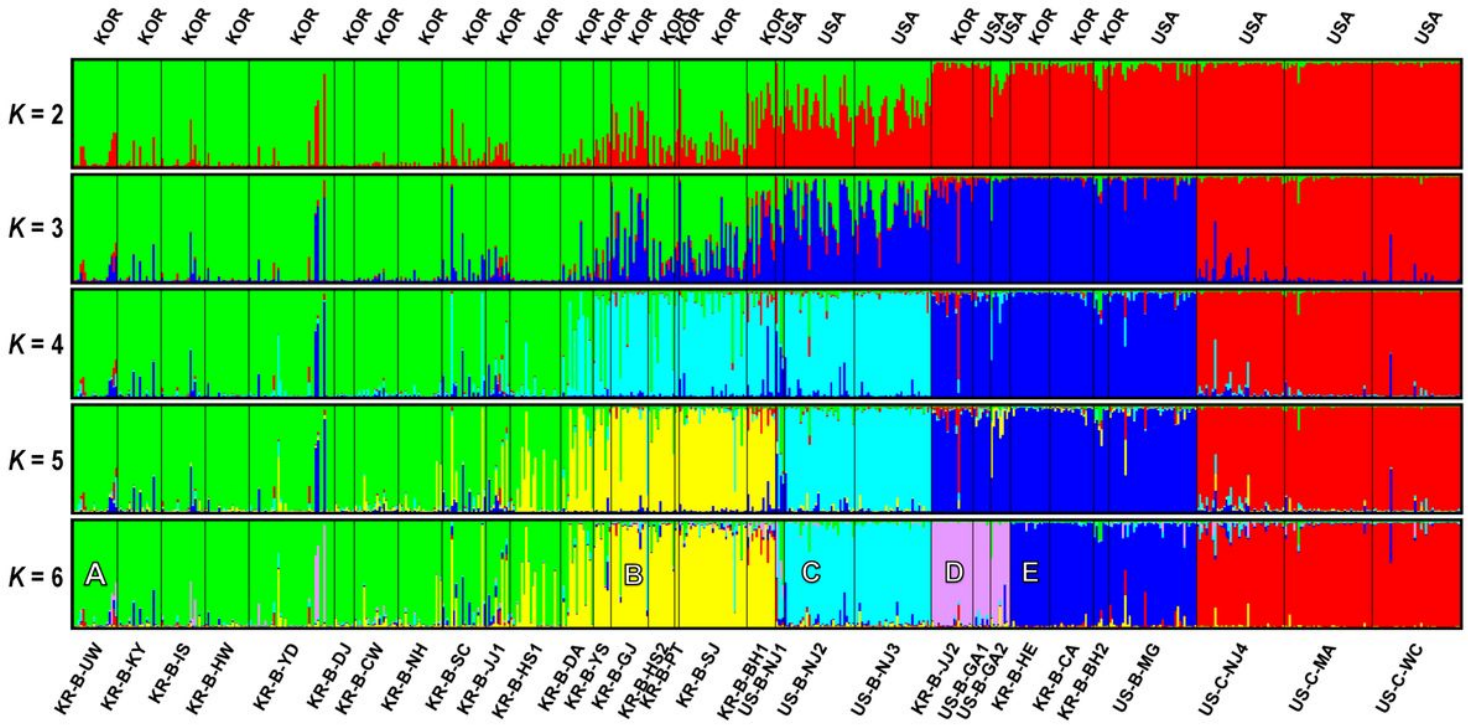

## Figure 2

Principal Coordinates Analysis (PCoA) plotted using microsatellite data from 632 *Dasineura oxycoccana* individuals collected from 31 blueberry and cranberry populations from USA and Korea. The X-axis is coordinate 1, ranging from -0.60 to 1.80, and the Y-axis coordinate 2, ranging from -0.50 to 0.40. Red circles correspond to cranberry populations, while blue circles correspond to blueberry populations.




## Figure 3

Principal Coordinates Analysis (PCoA) plotted using microsatellite data from 28 *Dasineura oxycoccana* populations from blueberry from USA and Korea. The X-axis is coordinate 1, ranging from -0.80 to 0.80, and the Y-axis coordinate 2, ranging from -0.60 to 0.60. Red circles correspond to the USA populations, while blue circles correspond to populations from Korea.




## Figure 4

Bayesian clustering (STRUCTURE) for the 31 *Dasineura oxycoccana* populations collected from blueberry and cranberry in the USA and Korea. Individual assignment plots for $K$ = 2, 3, 4, 5, and 6. Different colors indicate different clusters (e.g., A–E).

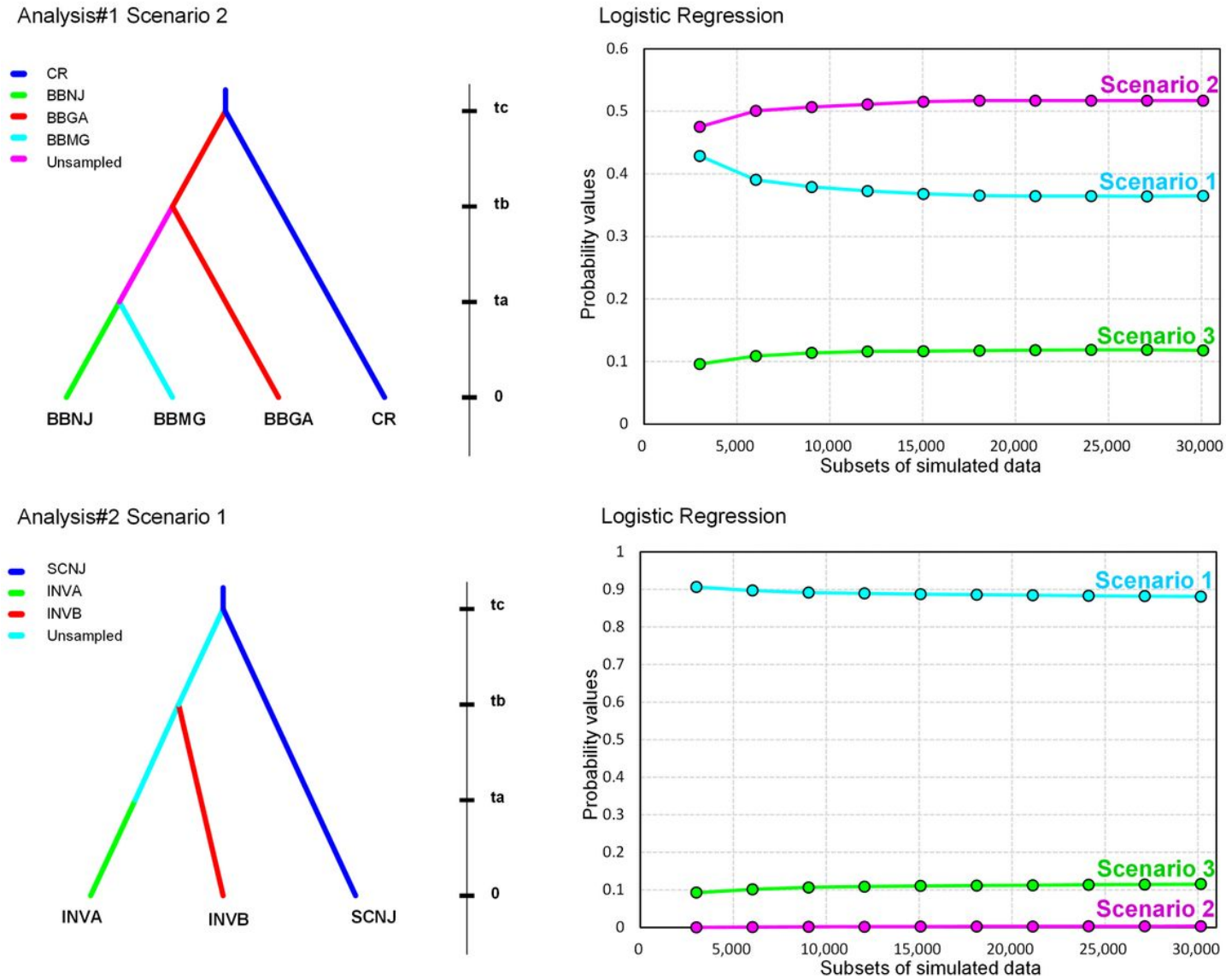



**Figure 5**

Scenarios (left panels) most supported by two independent analyses (Analysis #1 and Analysis #2) and probability values (right panels) from corresponding logistic regressions. Analysis #1 tested for serial divergence between *Dasineura oxycoccana* populations from blueberry and cranberry in the native range, while Analysis #2 tested for a divergence between a source (USA) of *D. oxycoccana* and the invaded region (Korea). Time (not to scale) is indicated on the right side of each scenario.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 3Supplementarymaterial2TablesS1S4.docx
- 4Supplementarymaterial1Figs.S1S4.docx