Biol Invasions (2026) 28:249
https://doi.org/10.1007/s10530-026-03951-y

ORIGINAL PAPER

Tracing the genetic footprint and invasion pathway of *Drosophila suzukii* in Chile: from Asia to the Andes and beyond

Francisco Pinilla · Tamara Vera-Tobar · Christian C. Figueroa · Paula Irles · Luis Devotto · Julián Mensch · Anne Loiseau · Antoine Fraimout · Guillaume Le Mailloux · Paul Bastide · François-David Collin · Jean-Michel Marin · Moisés A. Valladares · Mathieu Gautier · Arnaud Estoup · Luis E. Castañeda

Received: 6 January 2026 / Accepted: 25 September 2026
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2026

Abstract The spotted wing drosophila (*Drosophila suzukii*), a major invasive pest of soft fruits native to Asia, has recently expanded across South America. Reconstructing its invasion history is critical for understanding the demographic and evolutionary processes shaping biological invasions. Here, we infer the origin, invasion pathways and demographic dynamics of *D. suzukii* in South America, with a particular focus on Chile. We used mitochondrial COI sequences and an updated global microsatellite dataset (25 loci, 27 populations) to reconstruct invasion routes and estimate demographic parameters by applying Approximate Bayesian Computation (ABC) and ABC-Random-Forest frameworks. Mitochondrial data support a sequential colonisation through the Americas and reveal a strong reduction of genetic diversity in Chile. Microsatellite-based analyses provide higher resolution and uncover weak differentiation among eastern and central US populations but marked divergence among western US populations, consistent with multiple introductions. In

Associate Editor - Andres Martinez

Arnaud Estoup and Luis E. Castañeda are joint senior authors on this work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10530-026-03951-y.

F. Pinilla · T. Vera-Tobar · M. A. Valladares · L. E. Castañeda (✉)
Núcleo Interdisciplinario de Biología y Genética (NiBG), Instituto de Ciencias Biomédicas (ICBM), Facultad de Medicina, Universidad de Chile, Santiago, Chile
e-mail: luis.castaneda@uchile.cl

F. Pinilla · C. C. Figueroa · P. Irles · M. A. Valladares · L. E. Castañeda
Research Ring in Pest Insects and Climate Change (PIC2), Santiago, Chile

C. C. Figueroa
Instituto de Ciencias Biológicas, Universidad de Talca, Talca, Chile

P. Irles
Instituto Agri-Food, Ciencias Animales y Ambientales (ICA3), Universidad de O'Higgins, Rancagua, Chile

L. Devotto
Instituto de Investigaciones Agropecuarias (INIA), Chillán, Chile

J. Mensch
Departamento de Ecología, Genética y Evolución, Universidad de Buenos Aires, Buenos Aires, Argentina

A. Loiseau · A. Fraimout · G. Le Mailloux · M. Gautier · A. Estoup (✉)
CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Université Montpellier, Montpellier, France
e-mail: arnaud.estoup@inrae.fr

G. Le Mailloux · F.-D. Collin · J.-M. Marin
IMAG, CNRS, Université Montpellier, Montpellier, France

Published online: 01 November 2026

South America, strong genetic continuity links Brazil, Argentina, and Chile, with Chilean populations forming a tight cluster derived from Argentina. Our analyses also support a single introduction into Chile followed by rapid, stepwise latitudinal expansion. Demographic inference indicates recurrent bottlenecks of variable intensity and heterogeneous admixture during the invasion of the Americas, highlighting the role of serial founder events and secondary spread in shaping genetic structure. Post-inference goodness-of-fit analyses indicate that the inferred scenario cannot be rejected given the observed data. Overall, our results refine the reconstruction of invasion routes at a continental scale and illustrate how large, standardized genotype datasets analysed with ABC-RF advance inference on large-scale invasion histories and support pest management.

Resumen La mosca del alas manchada (*Drosophila suzukii*), una importante plaga invasora de frutos de pulpa blanda y originaria de Asia, que ha colonizado recientemente Sudamérica. La reconstrucción de la historia de invasión es fundamental para comprender los procesos demográficos y evolutivos que dan forma a las invasiones biológicas. En este estudio inferimos el origen, las rutas de invasión y la dinámica demográfica de *D. suzukii* en Sudamérica, con un enfoque particular en Chile. Para ello, utilizamos secuencias mitocondriales del gen COI y un conjunto de datos actualizado de microsatélites a nivel mundial (25 loci, 27 poblaciones). Estos datos fueron utilizados para reconstruir las rutas de invasión y estimar parámetros demográficos mediante enfoques de cálculo bayesiano aproximado (ABC) y ABC combinado con bosques aleatorios (ABC-RF). Los datos mitocondriales respaldan una colonización secuencial a través de las Américas y revelan una fuerte reducción de la diversidad genética en Chile. Los análisis basados en microsatélites ofrecen una mayor resolución. Estos revelan una débil diferenciación entre las poblaciones del este y el centro de Estados Unidos, pero una marcada divergencia entre las poblaciones del oeste de Estados Unidos, lo cual concuerda con múltiples introducciones en esta región. En Sudamérica, una fuerte continuidad genética conecta a Brasil, Argentina y Chile, y las poblaciones chilenas forman un grupo homogéneo que proviene de Argentina. Nuestros análisis también respaldan una única introducción en Chile, seguida de una rápida expansión latitudinal y en etapas sucesivas. La inferencia demográfica indica cuellos de botella recurrentes de intensidad variable y niveles heterogéneos de mezcla genética durante la invasión de las Américas. Estos resultados resaltan el papel de sucesivos eventos fundadores y eventos secundarios de propagación en la estructuración genética. Los análisis de bondad de ajuste (*goodness-of-fit*) realizados después de la inferencia indican que el escenario inferido no puede rechazarse dados los datos observados. En general, nuestros resultados refinan la reconstrucción de las rutas de invasión a una escala continental e ilustran cómo se pueden utilizar grandes conjuntos de datos estandarizados de genotipos, analizados con ABC-RF, para realizar inferencias sobre la historia de invasión a gran escala, así como en qué medida estos análisis pueden contribuir al manejo de plagas.

Résumé Originaire d'Asie, la drosophile à ailes tachetées (*Drosophila suzukii*), ravageur invasif majeur des petits fruits, a récemment colonisé l'Amérique du Sud. Reconstruire l'histoire des invasions est essentiel pour comprendre les processus démographiques et évolutifs qui façonnent les invasions biologiques. Nous inférons ici l'origine, les routes d'invasion et la dynamique démographique de D. suzukii en Amérique du Sud, avec un focus sur le Chili. Pour ce faire, nous avons utilisé des séquences mitochondriales du gène COI et un jeu de données microsatellites mondial mis à jour (25 locus, 27 populations). Ces données ont servi à reconstruire les routes d'invasion et à estimer des paramètres démographiques clefs, au moyen d'approches de calcul bayésien approché (ABC) et d'ABC adossé à des forêts aléatoires (ABC-RF). Les données mitochondriales soutiennent une colonisation séquentielle à travers les Amériques et révèlent une forte réduction de la diversité génétique au Chili. Les analyses fondées sur les microsatellites offrent une meilleure résolution. Elles mettent en évidence une faible différenciation entre les populations de l'est et du centre des États-Unis, mais une divergence marquée entre les populations de l'ouest des États-Unis

P. Bastide
MAP5, CNRS, Université Paris Cité, Paris, France

A. Estoup
ORCAe, Off-Site Research Center, INRAE, Montpellier, France

cohérente avec des introductions multiples dans cette dernière zone. En Amérique du Sud, une forte continuité génétique relie le Brésil, l'Argentine et le Chili, les populations chiliennes formant un groupe très homogène issu de l'Argentine. Nos analyses soutiennent également une introduction unique au Chili, suivie d'une expansion latitudinale rapide et par étapes successives. L'inférence démographique indique des goulots d'étranglement récurrents d'intensité variable et des niveaux de mélange génétique hétérogènes au cours de l'invasion des Amériques. Ces résultats soulignent le rôle d'effets fondateurs successifs et d'évènements de dispersion secondaire dans la structuration génétique. Les analyses d'adéquation (*goodness-of-fit*) menées à la fin du processus d'inférence indiquent que le scénario final inféré ne peut pas être rejeté au vu des données observées. Globalement, nos résultats affinent la reconstruction des routes d'invasion à une large échelle continentale, et montrent comment de grands jeux de génotypes standardisés, analysés par ABC-RF, peuvent être utilisés pour des inférences sur des histoires d'invasion à large échelle et dans quelle mesure ces analyses peuvent contribuer à la gestion des ravageurs.

Keywords Demographic processes · Invasion genetics · Pest insects · Pest management · Population genetics · Spotted wing drosophila

Introduction

Biological invasions are a major driver of global change, with important ecological, evolutionary, and economic consequences (Lockwood et al. 2013). Invasive insects are of particular concern because of their high mobility, rapid reproduction, and ability to exploit human-modified environments, often leading to severe impacts on agroecosystems (Estoup et al. 2026). The spotted wing drosophila (*Drosophila suzukii* Matsumura 1931) has become one of the most damaging invasive insect species worldwide. Unlike most drosophilids, *D. suzukii* females lay their eggs in intact ripening fruit with a serrated ovipositor, and the resulting larvae inflict major economic losses in berry and stone fruits (Cini et al. 2014; Asplen et al. 2015). Since its first report outside Asia in Hawaii in the early 1980s (Kaneshiro, 1983), *D. suzukii* has been detected in continental North America (Hauser 2011), Europe (Calabria et al. 2012), and more recently in South America (Deprá et al. 2014; Santandino et al. 2015; González et al. 2015; Bizama 2020) and Africa (Boughdad et al. 2021). Several traits can explain the success of *D. suzukii* as invasive species, notably its high reproductive rate and broad host range, which limit rapid and effective control measures following introduction (Cini et al. 2012).

Understanding invasion processes requires reconstructing the routes of invasion, i.e., the introduction pathways and demographic events underlying establishment. Population genetics offers powerful tools to understand biological invasions, providing insights into source populations, bottlenecks, admixture events, and bridgehead effects (Estoup and Guillemaud 2010; Cristescu 2015; Estoup et al. 2016). For *D. suzukii*, several studies have clarified aspects of its global expansion (Adrion et al. 2014; Fraimout et al. 2017; Lewald et al. 2021). In North America, genetic data point to contributions from native Asian populations as well as Hawaiian introduced lineages. On the other hand, European populations are more genetically homogeneous and genetic data suggest independent introductions, with China being the most likely source with a limited gene flow from the eastern USA. Such results highlight the fact that invasion dynamics are rarely straightforward, often involving recurrent introductions—sometimes from genetically differentiated sources—and secondary spread. South American populations of *D. suzukii* remain comparatively less studied (but see Ferronato et al. 2019; de la Vega et al. 2020), despite its relevance as both a fruit-exporting region and a more recently colonized area. After initial detections in Brazil (Deprá et al. 2014) and Argentina (Cichón et al. 2015), *D. suzukii* was first reported in Chile in 2017 (Bizama 2020), where it is currently expanding its latitudinal distribution range across more than 2,000 km (García et al. 2022; SAG 2025). Chile is of particular interest in the context of invasion biology due to its long latitudinal range, strong climatic gradients and relative geographic isolation (notably imposed by the Andes to the east), as well as its heavy reliance on fruit production for export. These features provide an opportunity to study how invasive species become established, spread, and potentially adapt in different environments.

From a methodological perspective, mitochondrial DNA (mtDNA) has been widely used in invasion genetics due to its extensive availability in genetic databases and its usefulness in making broad

phylogeographic inferences (Hurst and Jiggins 2005). However, its maternal inheritance and single-locus nature constrain its resolution, limiting its ability to discriminate between competing invasion scenarios, especially when recent or complex demographic processes are involved. By contrast, nuclear microsatellite markers provide a higher level of detail for inferring fine-scale population structure, bottlenecks, and admixture, particularly when analysed within a global comparative framework (Estoup et al. 2016; Fraimout et al. 2017). Together, analyses based on mtDNA and microsatellites provide a robust foundation for reconstructing invasion histories. Importantly, for *D. suzukii*, a large, standardized microsatellite dataset already exists, covering worldwide populations with markers well distributed across the genome. This approach makes individual-level genotyping straightforward, cost-effective, and highly informative for reconstructing recent demographic events, advantages that remain more difficult to achieve with whole-genome sequencing for large sample sizes. Although combining datasets can be more challenging with microsatellites, their scalability and resolution for contemporary invasion processes make them especially valuable in this context.

In addition to their relevance for evolutionary biology, genetic approaches play an increasingly important role in pest management (Sentis et al. 2022). Molecular markers can be integrated into surveillance programmes to detect new introduction events, track spread, and assess the adaptive potential of invasive pests (Estoup et al. 2016; McGaughran et al. 2024). This is particularly relevant for *D. suzukii*, whose rapid global expansion is strongly driven by fruit trade and human-mediated dispersal. Monitoring genetic diversity over time can inform demographic changes and emerging adaptive responses, including the evolution of resistance to control measures (e.g., insecticides, biological control), thereby supporting quarantine policies and integrated pest management strategies. Effective implementation, however, relies on rigorous and reproducible protocols to link newly generated data to existing reference datasets, as well as on efficient computational frameworks capable of updating invasion scenarios as additional populations are sampled and genotyped. These methodological requirements are increasingly critical as genetic datasets expand and management agencies seek timely, actionable information.

In this study, we used mitochondrial COI sequences and an updated global microsatellite dataset to reconstruct the invasion history of *D. suzukii* in South America, with particular emphasis on its recent establishment and spread in Chile. Because inference on South American invasion routes depends on the accurate placement of these populations within the broader worldwide expansion of the species, we also reassessed selected upstream components of the global invasion scenario previously proposed for *D. suzukii*.

More specifically, we addressed four main questions: (i) How do the levels and patterns of genetic diversity and differentiation observed in South American, and particularly Chilean, populations compare with those found in the native range and in other invaded regions? (ii) Which native-range populations provide the most appropriate genetic proxies for the early Hawaiian and western US introductions, and is the low differentiation among central and eastern US populations better explained by independent introductions or by descent from a shared unsampled ancestral population? (iii) Which source populations, introduction routes, and demographic processes, including bottlenecks, admixture, and secondary spread, best explain the successive establishment of *D. suzukii* in Brazil, Argentina and Chile? and (iv) Did Chilean populations originate from a single or multiple external sources, and are their genetic relationships consistent with a serial latitudinal expansion within Chile? Beyond these questions, we emphasised the practical value of molecular tools for pest management by demonstrating how newly generated genotyping data can be reliably integrated into an existing global reference dataset and how advanced statistical frameworks can be efficiently applied to increasingly large datasets to update invasion scenarios and support genetic pest surveillance.

Materials and methods

Sampling

Adult individuals of *D. suzukii* were collected from three locations in Chile and one location in Argentina using a commercial attractant (Suzukiitrap®) and preserved in 95% ethanol at 4 °C. Sampling took place between May and July 2023 in Santiago

(CL-Rm; 33°34′23.05″ S, 70°37′52.28″ W), Chillán (CL-Bc; 36°36′22.54″ S, 72°6′8.41″ W), and Villarrica (CL-Ar; 39°17′26.211″ S, 71°59′57.82″ W) in Chile, and in February 2023 in Golondrinas (AR-Gol; 42°01′00.0″ S, 71°31′20.0″ W) in Argentina. In addition, the Agricultural and Livestock Service of Chile (SAG) provided ethanol-preserved flies that were collected in Villarrica, Chile, in 2017 (CL-Ar2017, 39°17′26.211″ S, 71°59′57.82″ W), approximately one month after the species was officially recorded in Chile for the first time. All specimens were examined under a stereoscope, with species identification confirmed based on morphological characteristics (i.e., the black spot on wings for males and the serrated ovipositor for females).

Molecular methods

For mtDNA sequencing, genomic DNA was extracted from whole individual flies using the salting-out method (Sunnucks and Hales 1996). The mitochondrial COI gene was PCR amplified using M13-tailed primers CO1490 and HCO2198 (Folmer et al. 1994). The PCR mix volume was 12 µl and included 1X MasterMix Platinum Hot Start II (Thermo Fisher Scientific), 0.8 µM forward and reverse primers, and 1–5 ng/µl DNA template. The amplification conditions were as follows: initial denaturation at 95 °C × 2 min, 40 cycles of denaturation at 95 °C × 30 s, annealing at 53 °C × 30 s, extension at 72 °C × 30 s, and a final extension at 72 °C × 5 min. Amplification was evaluated on a 2% agarose gel, where expected PCR products of ~ 800 bp were observed. PCR amplicons were sequenced by AUSTRAL-omics (Universidad Austral de Chile, Valdivia, Chile) using the Sanger method. Chromatograms were manually inspected and edited in ProSeq v2.91 (Filatov 2002), resulting in final sequences of 601 bp. Microsatellite genotyping (DNA extraction, PCR amplification, and allele scoring) followed the protocol described in Fraimout et al. (2015), using the same 25 microsatellite loci genotyped in Fraimout et al. (2017). To verify that these markers were sufficiently distant from one another to be treated as statistically independent, we mapped the 25 loci onto the latest chromosome-level assembly of the *D. suzukii* genome (Camus et al. 2025) using the UCSC In-Silico PCR tool (Kent et al. 2002) (see Table S1 for details).

Linking a new microsatellite dataset to the existing reference database

It is crucial that new datasets can be reliably linked to existing reference databases when large-scale and long-term genetic monitoring studies are being carried out. This task is challenging when microsatellites are used because allele scoring at such molecular markers can vary across runs, laboratories, or platforms. A rigorous experimental design is therefore essential. We carefully integrated our new microsatellite dataset into the large dataset (i.e., 23 populations sampled worldwide in the species' native and invaded areas) from Fraimout et al. (2017). First, we selected 24 individuals for whom DNA had previously been stored from Fraimout et al. (2017) dataset to obtain genotypes covering the full range of microsatellite allele sizes detected in *D. suzukii* to date. Second, we re-genotyped these 24 individuals, known as "anchor-genotyped individuals", alongside genotyping the four new South American populations: AR-Gol, CL-Ar, CL-Bc, and CL-Rm. Finally, following this integration, we re-read all the inferred genotypes and attempted to obtain genotypes for individuals with unreadable ones by processing a second PCR on their DNA. This was done in both the four newly genotyped South American populations and the 23-population reference database of Fraimout et al. (2017), for which DNA extractions were still available. The objective was to minimise the number of undetermined genotypes in the updated microsatellite dataset. Our careful rereading process sometimes led us to identify new alleles, which were usually rare, and remove certain ones. Ultimately, we obtained an updated and refined global dataset comprising 779 individuals, representative of 27 populations and genotyped at 25 microsatellite loci. This dataset contained only 0.56% null genotypes, compared to 0.85% in the original dataset of Fraimout et al. (2017). The complete updated microsatellite dataset, coded in a 3-digit GENEPOP format (Raymond and Rousset 1995), is publicly available through Recherche Data Gouv: Estoup & Loiseau (2026), https://doi.org/10.57745/VXLA6U. https://entrepot.recherche.data.gouv.fr/dataverse/inrae

We compared a broad range of descriptive metrics of within- and among-population genetic variation between the original dataset of Fraimout et al. (2017) and the updated microsatellite dataset, and found only minor differences (results not shown). Accordingly, the ABC-RF analyses used by Fraimout et al. (2017) to reconstruct invasion routes yielded highly consistent qualitative and quantitative results across the two datasets, particularly regarding the posterior probabilities of the best-supported scenarios (results not shown). These results supported the use of the updated microsatellite dataset, comprising 779 individuals from 27 populations, while allowing us to rely primarily on the invasion-route inferences reported by Fraimout et al. (2017).

Genetic diversity and population structure

We used our new mtDNA haplotype dataset to place Chilean populations within a broader phylogeographic context. To this end, we downloaded all available *D. suzukii* COI sequences from GenBank using the keywords "*Drosophila suzukii*" and "COI". A total of 441 sequences were retrieved, representing populations from the United States, Canada, Europe, Japan, South Korea, China, Brazil and Argentina. Sequence processing and filtering were performed using the DECIPHER v2.28.0 package for R (Wright 2016). The final datasets included Chilean and GenBank-derived sequences (511 bp), which were standardized using ProSeq (Filatov 2002).

Genetic diversity indices for COI were estimated using DnaSP v6.12.03 (Rozas et al. 2017), including the number of haplotypes (K), haplotype diversity (H_d), number of polymorphic sites (S), and average number of differences between pairs of sequences (π). For comparative purposes, populations were grouped into invasive (United States, Canada, Europe, Brazil, Argentina, and Chile) and native (China, Japan, and South Korea) categories. Diversity indices were compared between these invasive and native populations using the Wilcoxon rank-sum test for independent samples implemented in RStudio v4.3.1 (Posit Team 2023). We also estimated the F_{ST} among pairs of populations and performed an analysis of molecular variance (AMOVA) using Arlequin v3.5 (Excoffier and Lischer 2010). For both analyses, 10,000 permutations were performed and a significance P-value < 0.05 was defined. For the AMOVA, populations were grouped into four regions: Region 1—South America (Argentina, Brazil, and Chile), Region 2—North America (Canada and USA), Region 3—Europe, and Region 4—Asia (China, Japan, and South Korea). Finally, a haplotype network was constructed using PopART v1.7 software (Leigh and Bryant 2015).

Regarding the global updated microsatellite dataset, we used the adegenet R package (v2.1.11; Jombart 2008) to estimate (i) the level of genetic variability within samples by calculating the mean observed and expected heterozygosity (H_o and H_e; Nei 1987) and the corrected allelic richness (A_c), and (ii) the genetic structure between sample sites by computing pairwise F_{ST} values. A_c, H_o and H_e were compared between invasive and native populations (see Table S5) using the Wilcoxon rank-sum test for independent samples using RStudio v4.3.1 (Posit Team 2023). We also performed Discriminant Analysis of Principal Components (DAPC; Jombart et al. 2010) as implemented in the R package adegenet (v2.1.11). We further investigated population genetic structure using the Bayesian clustering approach implemented in STRUCTURE v2.3.4 (Pritchard et al. 2000). To examine hierarchical structure at different geographic scales, we performed three complementary analyses: (i) all 27 worldwide populations; (ii) the 13 American populations, comprising eight populations from the US and five from South America; and (iii) a balanced subset comprising five populations from the continental US (US-Wat, US-Sdi, US-Nca, US-Gen, and US-Col) and the five South American populations. The five US populations were selected to represent the major components of genetic variation identified within the continental US, including two differentiated western populations and three representatives of the weakly differentiated central/eastern population group. This balanced analysis was used to assess whether the inferred structure between the continental US and South America was sensitive to unequal regional representation. For each dataset, K ranged from 1 to 10, with 10 independent runs per K, using the admixture model with correlated allele frequencies. Each run consisted of a burn-in of 50,000 iterations followed by 1,000,000 MCMC iterations. Replicate runs were aligned and averaged, and individual membership coefficients were visualised using the R package pophelper v2.3.1 (Francis 2017). The ΔK

method of Evanno et al. (2005) was used to identify the main hierarchical levels of genetic structure rather than a single definitive number of clusters.

Invasion routes reconstructed using mtDNA data

To determine the most probable origin of the *D. suzukii* populations in Chile, we performed an Approximate Bayesian Computation (ABC) analysis using DIYABC v2.0 software (Cornuet et al. 2014). For this analysis, 30 South Korean samples were excluded from the dataset because they were associated with low-frequency haplotypes exclusively found in South Korea; their inclusion did not qualitatively affect the inferred results (not shown). The remaining samples from South Korea were grouped into a single unit with China and Japan representing the native range, and the US and Canadian samples were grouped into a single unit representing North America. Following Fraimout et al. (2017), we modelled an introduction event as a divergence without subsequent gene flow from the source population in the ABC analyses. The introduction event was set to occur at a time corresponding to the date of the first record (allowing an additional 0–4 years as the upper bound for introduction time) in the invaded area and was translated as the number of generations assuming 12 generations per year (Lin et al. 2014). Each introduction event was immediately followed by a bottleneck period, characterized by a smaller effective population size, after which the population size returned to a stable level (Table S2). We evaluated the following five alternative introduction scenarios into Chile, each scenario having 23 demographic and mutational parameters (Fig. S1): Scenario 1 = sequential introduction via North America, Brazil, and Argentina; Scenario 2 = direct introduction from the native range; Scenario 3 = direct introduction from North America; Scenario 4 = direct introduction from Europe; and Scenario 5 = sequential introduction via North America and Brazil. These scenarios were constructed based on fruit trade routes between Chile and other countries from 2013 to 2017 (ODEPA 2024). For the ABC simulations, we used demographic parameter prior distributions similar to the prior set 1 proposed in Fraimout et al. (2017) (see also Table S2). As there is no species-specific COI mutation rate estimation available for *D. suzukii*, a uniform mutation rate ranging from 1×10^{-8} to 3×10^{-8} substitutions per site per generation was used, in line with estimates for *Drosophila melanogaster* (Satta et al. 1987). A Kimura 2-parameter substitution model of sequence evolution was applied (Kimura 1980).

The most likely introduction scenario and its posterior probability estimate were determined using a logistic regression based on the Euclidean distance between observed and simulated summary statistics. All single-sample and two-sample summary statistics proposed by DIYABC v2.0 for mtDNA sequences were used (total = 123 summary statistics). For the regression step, 50,000 simulations were retained, corresponding to 1% of the full set of 5×10^{6} simulations (i.e. 10^{6} simulations per scenario). Following Cornuet et al. (2014), we evaluated the confidence in identifying the best-supported scenario for our observed mtDNA dataset (i.e., scenario 1 as we will see later) by computing the posterior predictive errors of type I and type II for such scenario. This involved simulating 1,000 pseudo-observed datasets (PODs) and drawing the scenario ID and parameter values with replacement from the 100 simulated datasets closest to the observed dataset, out of the total of 5×10^{6} simulated datasets recorded in the reference table. We used the same procedure as described above to assess for each POD which scenario had the highest posterior probability. We then calculated the type I and type II posterior predictive error rates. Type I errors occur when scenario 1 is the true model (H_o is true), but another scenario is selected (H_o is rejected). Type II errors occur when scenario 1 is not the true model (H_o is false), but scenario 1 is nevertheless selected (H_o is not rejected).

Invasion routes revealed by the microsatellite data

For this dataset, we used the Approximate Bayesian Computation Random Forest (ABC-RF) methodology (Collin et al. 2021) to infer invasion routes. Compared to more traditional ABC methods, such as those used for the mtDNA dataset, ABC-RF enables the estimation of the probability of the best scenario and the distribution of parameters, while substantially reducing the computational load (Collin et al. 2021). ABC-RF approaches are therefore particularly well suited to analyzing large datasets, such as our updated microsatellite dataset. The method enables robust

discrimination among complex invasion routes and estimation of key demographic parameters, including the intensity of introduction bottlenecks and levels of genetic admixture among multiple sources. Readers are invited to consult Pudlo et al. (2016), Raynal et al. (2019), and Collin et al. (2021) for detailed statistical descriptions and testing of the ABC-RF method.

In line with Fraimout et al. (2017), we employed historical data (i.e., the dates on which each invasive population was initially observed) to conduct six sets of ABC-RF analyses. The details of these analyses are presented in Table S3. Since the main results of the ABC-RF analyses carried out by Fraimout et al. (2017) were confirmed using our updated microsatellite dataset, the six sets of new analyses focused on four objectives: (i) to refine the identification of the most appropriate genetic proxies for the origin of the native source populations in Japan and China; (ii) to test whether a single, unsampled ancestral population could explain the low genetic differentiation observed among the central and eastern US populations (US-Nca, US-Wis, US-Gen, and US-Col); (iii) to infer the genetic origins of the five South American populations represented in the microsatellite dataset, namely BR-Poa (Brazil), AR-Gol (Argentina), and CL-Ar, CL-Bc, and CL-Rm (Chile); and (iv) to test whether the Chilean populations originated from single or multiple external sources and whether their genetic relationships were compatible with a continuous latitudinal expansion across Chile (i.e. CL-Bc derived from CL-Ar and CL-Rm from CL-Bc).

In our ABC-RF analyses, we modelled introduction events from single or multiple source populations and bottleneck periods following Fraimout et al. (2017). The parameterization of the scenarios and the prior distributions of the demographic and mutational parameters were identical to those in Fraimout et al. (2017) (see detailed prior distribution set 1 in Table S8 of Fraimout et al. 2017). The only exceptions were the lower bounds of the priors for the divergence time between native populations (set to 2000 generations instead of 0) and for the stable effective population sizes (set to 1000 individuals instead of 10,000). These adjustments were made to simplify the simulation setup (i.e., to avoid unnecessary time-event conditioning in DIYABC coding) and to broaden the range of plausible long-term population sizes, respectively. Preliminary tests indicated that these minor changes had no impact on the ABC-RF inferences (results not shown). The datasets were summarized using the full set of microsatellite statistics implemented in DIYABC, identical to those used in Fraimout et al. (2017).

Two types of ABC-RF analysis were conducted. The first type involves situations in which fewer than 10 scenarios are compared. In this case, we performed a standard ABC-RF analysis on a training set comprising 10,000 simulations for each scenario. This RF analysis was repeated 10 times (on the same training set) to better evaluate the stability of the results and estimate the average probability and standard deviation for the best scenario (i.e., the one with the highest number of votes), as well as determine how many times this best scenario was identified across all 10 analyses. The second type of ABC-RF analysis involves comparing more than ten scenarios. In this case, we first performed a pruning procedure using a lightweight ABC-RF analysis. The aim was to reduce the number of computationally expensive simulations required to build the training set and to prevent oversizing of Random Forests, which can lead to excessive memory usage and prohibitive computation times. The pruning ABC-RF analysis is a heuristic statistical method in which only 2,000 simulations are produced for each scenario using the command-line software DIYABC v1.1.54, which is available at https://diyabc.github.io/cli/ (Collin et al. 2021). We then used the R package abcrf v1.9 (Pudlo et al. 2016; Raynal et al. 2019) to analyze those simulations using RF algorithms with only 500 trees per forest. The pruning RF analysis was repeated 10 times (on the same training set), and the average number of votes was calculated for each scenario. Only scenarios associated with an average number of votes $>5\%$ (i.e., >25 votes) were retained as potential candidates for a second, standard, and thorough ABC-RF analysis. In the second ABC-RF analysis, we used a training set of 10,000 simulations per retained scenario, analysed using Random Forests with 1000 trees. The analysis was repeated 10 times (on the same training set) to assess result stability and to estimate the mean probability, standard deviation, and selection frequency of the best-supported scenario (i.e. the scenario receiving the highest number of votes). A detailed description of the statistical treatment applied to each ABC-RF analysis performed within the scope of this study is presented in Table S3.

Our parameter estimation focused on two types of demographic events that are considered key potential drivers of invasion success and that leave strong signatures in genetic data: the severity of bottlenecks during the formation of new populations, and the extent of genetic admixture (Estoup et al. 2016). Bottleneck severity was defined as the ratio of bottleneck duration to the number of effective individuals during this period (Pascual et al. 2007). ABC-RF estimations were processed in accordance with the methodology outlined by Raynal et al. (2019) using the R package *abcrf* v1.9 on a training set comprising 40,000 simulations generated under the final scenario, with 1,000 trees per forest.

As the final step in the analysis process, we conducted an objective evaluation of the DIYABC-RF inferences using a recently developed statistical method known as Post-Inference Goodness of Fit (Post-GoF; Le Mailloux et al. 2026). This method estimates how well a statistical model fits the observed data. It is useful for models with intractable likelihood, as was the case here. We used the R package *abcgof* v0.0.1 (Le Mailloux et al. 2026) to perform the post-GoF test on the final inferred scenario. This test determines whether the scenario omits any significant evolutionary features that could explain the observed dataset. If this is the case, then the multi-parameter posterior distribution conditional on the observed dataset has a poor propensity to generate multi-dimensional summary statistics close to those corresponding to the observed dataset, and the post-GoF test is associated with a low *p*-value. More specifically, we computed the post-GoF *p*-value from a dataset of 100,000 simulations generated under the final scenario. With 25 loci available, we kept the full dataset rather than splitting it as recommended by Le Mailloux et al. (2026) (see that reference for discussion of the rationale and implications of this choice). We further retained the 1,000 best simulations (corresponding to a threshold of 1%), keeping 500 of these for calibration of the test. We used the rejection method and the Local Outlier Factor (LOF) score for this propose. Readers are invited to consult Le Mailloux et al. (2026) for detailed statistical descriptions of the post-GoF test. The median and 95% Highest Density Intervals (HDIs) of the post-GoF *p*-value were calculated from 500 bootstrapped iterations of calibration particles.

Results

Genetic diversity

In the global mtDNA dataset (Table S4), native populations exhibited significantly higher numbers of haplotypes, haplotype diversity and polymorphic sites than invasive populations ($W = 0.000$, $P = 0.020$ for all three tests). However, no significant differences were observed for nucleotide diversity (π) ($W = 7.000$, $P = 0.710$). Notably, Chilean populations exhibited particularly low diversity indices relative to both native and other invasive populations (see Table S4),

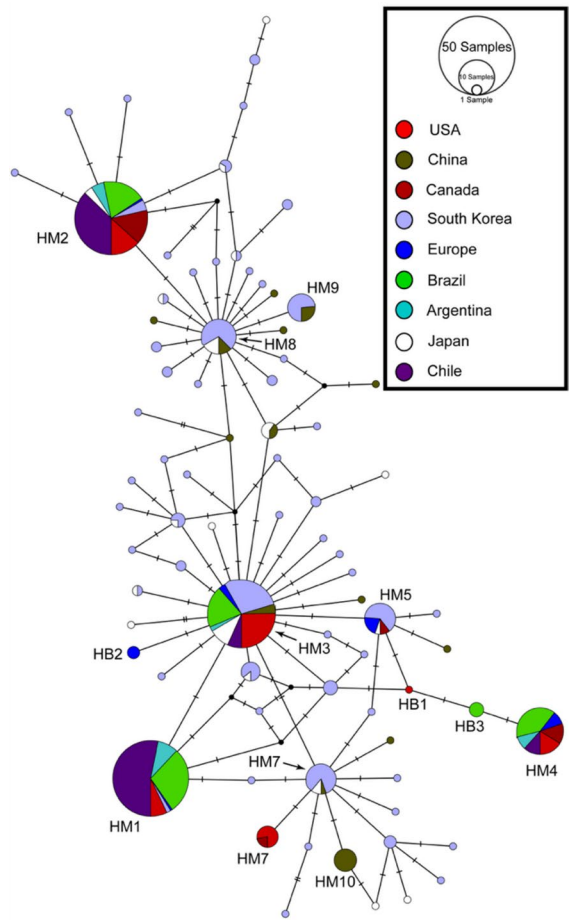


Fig. 1 MtDNA COI haplotype network of worldwide *D. suzukii* populations. Each mtDNA haplotype is represented by a circle. The size of each circle represents the haplotype frequency, and the colour of each circle corresponds to the samples included in the analysis. HM and HB labels indicate high- and low-frequency haplotypes, respectively

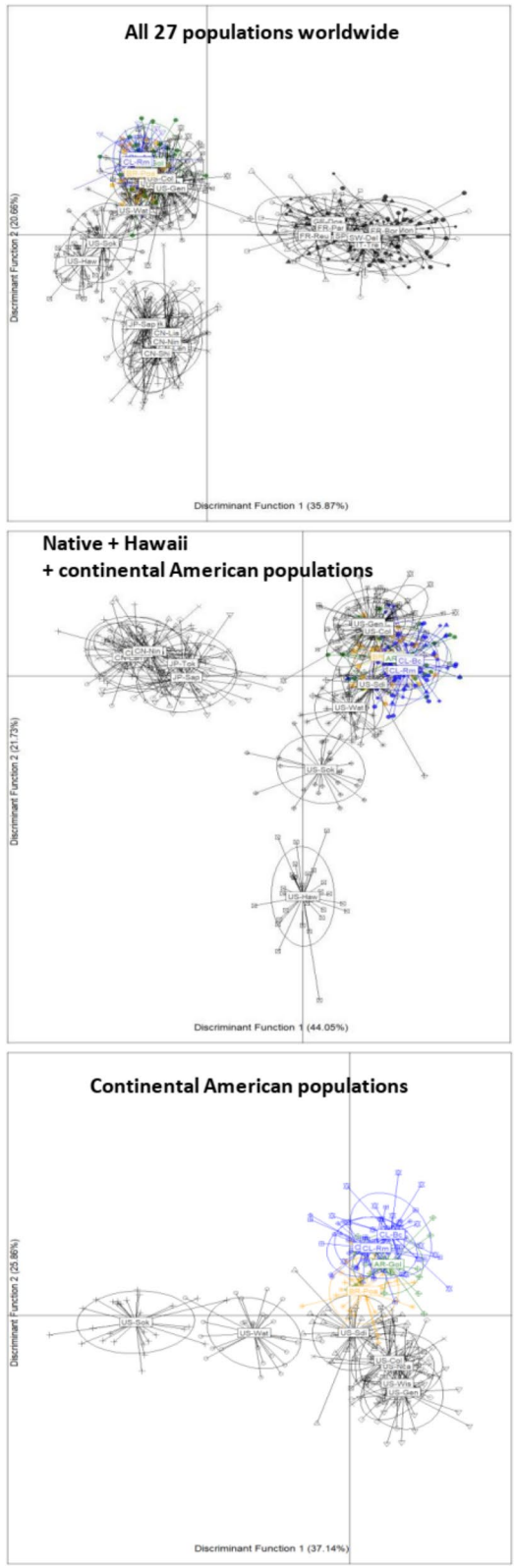


Fig. 2 DAPC for various sets of *D. suzukii* populations genotyped at 25 microsatellite loci. Populations from Brazil, Argentina and Chile are in orange, blue and green colour, respectively

which is consistent with a recent secondary introduction. Among Chilean populations, CL-Ar exhibited the highest haplotype diversity ($H_d = 0.622$), while CL-Ar2017 exhibited the lowest value ($H_d = 0.343$). In terms of the number of polymorphic sites, the highest values were observed in CL-Ar2017, CL-Ar, and CL-Bc, whereas the lowest value was observed in CL-Rm. Overall, pairwise sequence differences were close to one across populations, with CL-Ar exhibiting the highest values.

Regarding the microsatellite dataset (Table S5), estimates of genetic diversity within populations, based on allelic richness (A_c), observed heterozygosity (H_o), and expected heterozygosity (H_e), revealed significant differences between native and invasive populations ($W = 126$, $P = 6.8 \times 10^{-6}$ for A_c, H_o and H_e). This is consistent with demographic bottlenecks occurring during the invasion process. Native Asian populations displayed the highest levels of genetic diversity. For example, CN-Lan, CN-Lia and CN-Shi exhibited a mean (corrected) allelic richness of over 8, and H_o and H_e values of around 0.82 (CN-Lan: $A_c = 8.58$, $H_o = 0.815$, and $H_e = 0.828$). Japanese samples showed slightly lower values (JP-Tok: $A_c = 7.87$, $H_o = 0.765$, and $H_e = 0.793$; JP-Sap: $A_c = 7.69$, $H_o = 0.766$, and $H_e = 0.790$), though these were still higher than those of all invasive populations. In contrast, invasive populations exhibited reduced diversity. The oldest invasive sample, collected on the island of Hawaii (US-Haw), exhibited the lowest diversity, with $A_c = 3.65$, $H_o = 0.617$, and $H_e = 0.617$, consistent with severe demographic reduction during introduction and/or persistently low effective population size. European invasive populations also exhibited reduced diversity relative to those of the native range (e.g., SP-Bar: $A_c = 5.55$, $H_o = 0.692$, and $H_e = 0.717$; IT-Tre: $A_c = 5.24$, $H_o = 0.699$, and $H_e = 0.705$). Western and eastern continental US populations retained intermediate levels of diversity (e.g., US-Wat: $A_c = 6.13$, $H_o = 0.743$, and $H_e = 0.761$; US-Sdi: $A_c = 5.84$, $H_o = 0.751$, and $H_e = 0.745$; US-Wis: $A_c = 5.61$, $H_o = 0.738$, and $H_e = 0.750$). South American populations displayed a further erosion of genetic diversity, with BR-Poa (Brazil) at $A_c = 5.64$,

$H_o = 0.76,3$ and $H_e = 0.747$ and AR-Gol (Argentina) at $A_c = 5.49$, $H_o = 0.732$, and $H_e = 0.736$. The Chilean populations showed the lowest values among the South American samples, with a slight decrease from south to north (CL-Ar: $A_c = 5.33$, $H_o = 0.715$, and $H_e = 0.723$; CL-Bc: $A_c = 4.92$, $H_o = 0.704$, and $H_e = 0.702$; CL-Rm: $A_c = 5.10$, $H_o = 0.692$, and $H_e = 0.702$). This pattern is consistent with the expansion front suggested by observational records and the serial colonization pathway inferred from ABC-RF (see below).

Population structure

At the global geographic scale, most comparisons of mtDNA data between populations revealed significant differentiation, except for three comparisons: Chile—Argentina, Argentina—Brazil, and South Korea—Japan (Fig. S2). The AMOVA based on mtDNA data showed that 10.37% of the total variation was explained by differences among the four geographic regions (North America, South America, Europe, and Asia). The remaining variation was due to differences among populations within regions (4.76%) and within populations (84.87%). Fixation indexes showed a moderate genetic differentiation among regions ($F_{CT} = 0.104$) and low differentiation among populations within regions ($F_{SC} = 0.053$).

Within Chilean populations, the population CL-Ar2017 exhibited significant mtDNA differentiation from CL-Ar ($F_{ST} = 0.281$, $P = 0.0002$), CL-Bc ($F_{ST} = 0.586$, $P = 1 \times 10^{-6}$), and CL-Rm ($F_{ST} = 0.520$, $P = 1 \times 10^{-6}$). This result is noteworthy because CL-Ar2017 was sampled in 2017 near the site of the first reported occurrence of *D. suzukii* in Chile and likely represents one of the earliest established populations in the country. Among more recently sampled Chilean populations, only the comparison between CL-Ar and CL-Bc showed significant genetic differentiation ($F_{ST} = 0.099$, $P = 0.019$). In contrast, CL-Rm did not differ significantly from either CL-Ar ($F_{ST} = 0.038$, $P = 0.144$) or CL-Bc ($F_{ST} = -0.007$, $P = 0.438$).

A total of 78 mtDNA haplotypes were identified through the haplotype network analysis (Fig. 1). Overall, the mitochondrial network did not display clear geographic clustering by country. The most frequent haplotypes (HM1 to HM4) were widely distributed across both the native and invasive ranges and were found in all South American populations. In contrast, several low-frequency haplotypes exhibited more restricted distributions: four (HM7 to HM10) were found exclusively in the native range; HM5 was present in both the native and invasive populations; and HM6 was found only in North America (Canada and the USA). Three additional haplotypes were found only in invasive regions: HB1 in the US, HB2 in Europe, and HB3 in Brazil. Although Chilean populations exhibited low haplotype diversity (Table S4), consistent with a recent introduction, their main haplotypes were shared with those from other South American sites. This suggests that they all originate from similar, if not the same, invasion source(s).

Microsatellite data revealed a stronger spatial genetic structure than mtDNA, with clear differentiation between native and invasive populations and among geographic regions within the invasive range. Pairwise F_{ST} values (Table S6; Fig. S3), DAPC ordination methods (Fig. 2), STRUCTURE analysis (Fig. S4), and neighbor-joining trees (NJ trees; Fig. S5) consistently grouped the native Asian populations as a tightly cohesive cluster, strongly differentiated from all invasive populations. The first low-diversified invasive population, the Hawaiian island population, formed an isolated cluster that was highly differentiated from all the others. In continental North America, the eastern and central US populations (US-Nca, US-Wis, US-Gen, and US-Col) were weakly differentiated (with very low F_{ST} values) and consistently grouped together in the DAPC and NJ tree analyses. By contrast, western US populations (US-Wat, US-Sok, and US-Sdi) appeared more differentiated—a striking pattern given their status as some of the earliest invasive populations on the American continent. In South America, the Brazilian population (BR-Poa) clustered closely with the Argentinian population (AR-Gol), with a low pairwise F_{ST} value of 0.020 and tight grouping in DAPC, STRUCTURE, and NJ-tree analyses. The three Chilean populations (CL-Ar, CL-Bc, and CL-Rm) were also tightly grouped with low differentiation ($F_{ST} \leq 0.014$) and were clearly proximate to AR-Gol (and secondarily to BR-Poa) in both analyses. Within Chile, nuclear genetic differentiation increased slightly among populations, consistent with a latitudinal gradient of divergence.

STRUCTURE recovered hierarchical patterns that were broadly consistent with the pairwise FST,

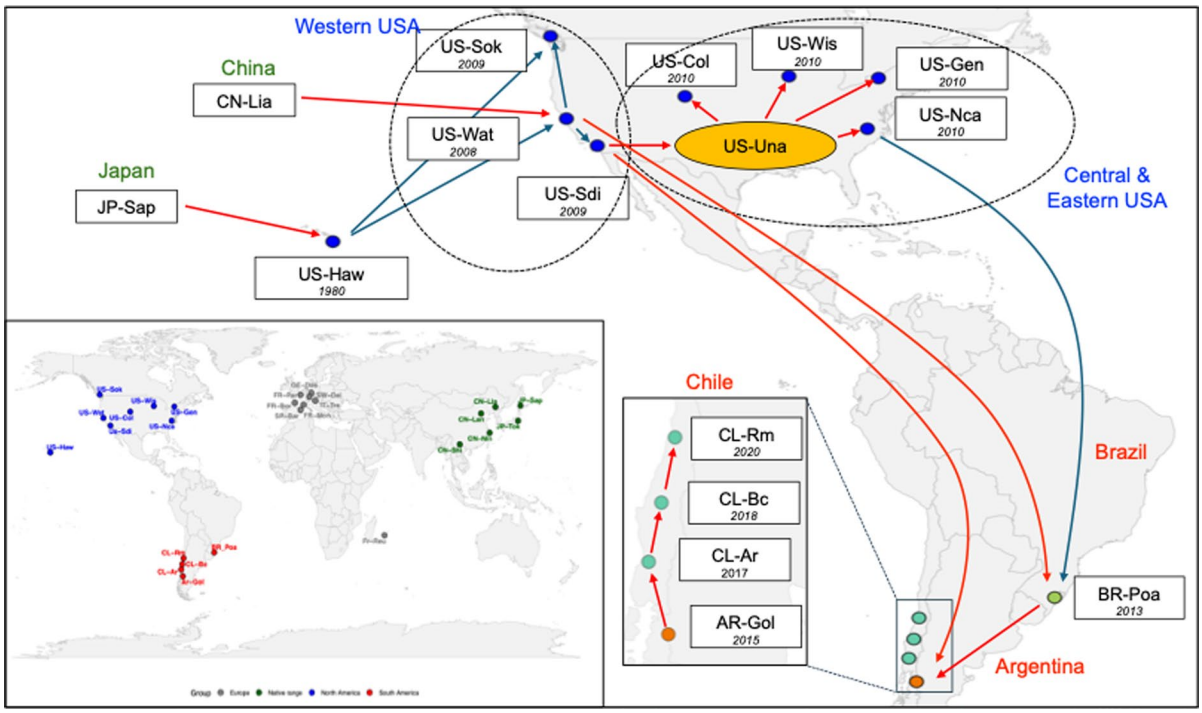


Fig. 3 Final scenario of invasion of North and South America inferred from ABC-RF analyses using the updated microsatellite dataset. New findings relative to the invasion scenario proposed by Fraimout et al. (2017) are shown in light red. These results include the identification of the most appropriate genetic proxies from the native range (JP-Sap in Japan and CN-Lia in China), as well as the inference of a common ancestral population (US-Una) for populations from the central and eastern United States (US-Nca, US-Wis, US-Gen, and US-Col). They further indicate contributions of US populations, notably US-Wat (and US-Nca), to the Brazilian population (BR-Poa), an admixed origin of the Argentine population (AR-Gol) from BR-Poa and US-Sdi, the role of AR-Gol as the sole source of the initial Chilean population, and a continuous latitudinal expansion of *D. suzukii* within Chile (i.e., CL-Bc derived from CL-Ar and CL-Rm from CL-Bc).

DAPC, and NJ tree analyses (Fig. S4). In the worldwide analysis, ΔK showed a dominant peak at $K=3$ and a secondary peak at $K=4$. At $K=3$, the principal structure distinguished European populations, most continental American and South American populations, and a component predominant in native Asian populations and US-Haw, with transitional assignments in US-Sok and US-Wat. At $K=4$, US-Haw and US-Sok formed an additional component, US-Wat showed mixed membership, and US-Sdi grouped predominantly with the central and eastern US populations. When the analysis was restricted to the Americas, the dominant $K=2$ signal separated US-Haw and US-Sok, with US-Wat showing an intermediate assignment, from the remaining populations. At the weaker $K=3$ level, continental US and South American populations became differentiated. BR-Poa showed the clearest mixed assignment, whereas AR-Gol and the three Chilean populations were predominantly assigned to the same South American component. In the balanced analysis, ΔK showed a dominant peak at $K=3$ and a secondary peak at $K=2$. At $K=2$, continental US and South American populations were clearly differentiated, although BR-Poa and, to a lesser extent, AR-Gol displayed mixed membership. At $K=3$, the two western US populations (US-Wat and US-Sdi) were further differentiated from the central/eastern US populations (US-Nca, US-Gen, and US-Col), while the South American populations formed a third major component. BR-Poa retained the strongest mixed assignment, AR-Gol was predominantly assigned to the South American component with a smaller US contribution, and the three Chilean populations were almost entirely assigned to the South American component. Thus, the differentiation

between the continental USA and South America, together with the mixed assignment of BR-Poa, was not driven by unequal regional representation (Fig. S4). Overall, the STRUCTURE analyses corroborated the major genetic discontinuities identified by the other descriptive analyses, including the complex structure of the western US populations and the genetic continuity between Brazil, Argentina, and Chile. However, they did not resolve the direction or detailed sequence of the introduction and admixture events, which were subsequently evaluated using explicit invasion scenarios within the ABC-RF framework.

Together, these findings from standard descriptive analyses of genetic structure across populations provide an independent baseline for reconstructing invasion patterns. These patterns were subsequently evaluated and refined using scenario-based ABC inference, allowing the inferred relationships to be rigorously tested within an explicit statistical framework.

Routes of invasion

Regarding the mtDNA dataset, the standard ABC analysis indicated that scenario 1 (USA → Brazil → Argentina → Chile) was the most likely scenario for the introduction of *D. suzukii* into Chile, with a posterior probability of 0.553 (95% CI: 0.452–0.654). Scenario 5 (Brazil → Chile) received substantial but lower support, with a probability of 0.345 (95% CI: 0.203–0.487). The probabilities of the invasion routes as described in scenario 3 (USA → Chile), scenario 4 (Europe → Chile), and scenario 2 (native → Chile) were much lower (i.e., 0.066 with CI: 0.000–0.185,

Table 1 Microsatellite-based estimation under the final invasion scenario and using Random Forest of admixture rates for introduced populations with multiple sources and bottleneck intensities at the time of each introduction.

	Parameter	Mean	Median	5% quantile	95% quantile
Admixture rates	Watsonville-USA (Liaoyuan, China contrib.)	0.769	0.768	0.646	0.894
	Dayton-USA (Hawaii-USA contrib.)	0.240	0.234	0.066	0.441
	Porto Alegre-Brazil (North Carolina-USA contrib.)	0.510	0.502	0.267	0.763
	Golondrinas-Argentina (Porto Alegre-Brazil contrib.)	0.732	0.741	0.555	0.907
Bottleneck intensity	Barcelona-Spain	0.496	0.457	0.268	0.847
	Hawaii-USA	0.521	0.517	0.368	0.697
	Watsonville-USA	0.120	0.109	0.052	0.226
	Dayton-USA	0.206	0.180	0.065	0.430
	San Diego-USA	0.093	0.076	0.025	0.223
	US-Una*	0.343	0.197	0.020	1.018
	Fort Collins-USA	0.048	0.027	0.002	0.181
	Geneva-USA	0.092	0.069	0.009	0.244
	Wisconsin-USA	0.170	0.145	0.033	0.377
	North Carolina-USA	0.051	0.028	0.002	0.189
	Porto Alegre-Brazil	0.248	0.203	0.047	0.573
	Golondrinas-Argentina	0.072	0.066	0.023	0.147
	Villarrica-Chile	0.058	0.028	0.002	0.202
	Santiago-Chile	0.110	0.085	0.021	0.247
	Chillán-Chile	0.045	0.034	0.005	0.115

See Fig. 3 for a description of the final invasion scenario. The source population whose genetic contribution is being estimated is indicated in parentheses. US-Una* is an unsampled common ancestral population that originates from San Diego (US-Sdi) and is shared by the populations in the central and eastern USA, i.e. North Carolina (US-Nca), Wisconsin (US-Wis), Geneva (US-Gen) and Fort Collins (US-Col)

0.036 with CI: 0.000–0.160, and 0.0001 with CI: 0.000–0.128, respectively), clearly ruling them out. To evaluate confidence in identifying the most probable scenario in the ABC analysis, we estimated the type I and type II predictive errors associated with scenario 1. The type I error (i.e., the probability of rejecting scenario 1 when it is true) was 0.477, and the type II error (i.e., the probability of accepting scenario 1 when another scenario is true) was 0.612 (Table S7). These high error rates mainly underscore the difficulty of distinguishing between scenarios 1 and 5. This pattern is consistent with the limited mtDNA differentiation observed between Argentinian and Brazilian populations.

ABC-RF analyses of the updated microsatellite dataset provided a clearer and more consistent picture, giving quantitative evidence of the invasion history of *D. suzukii* in the Americas, including Chile, compared to standard ABC analyses of the mtDNA dataset. Table S3 provides details of the characteristics and numerical results for each ABC-RF analysis. Figure 3 depicts the final invasion scenario inferred from our microsatellite-based ABC-RF analyses.

When using native-range sources as genetic references, the Hawaiian population (US-Haw) showed the strongest alignment with the Japanese population (JP-Sap) (analysis 1a in Table S3; $P = 0.610$, 10/10 runs), whereas the western US population (US-Wat) showed a clear genetic contribution from the Chinese population (CN-Lia) (analysis 1b in Table S3; $P = 0.840$, 10/10 runs).

For the origins of the central and eastern US populations (US-Nca, US-Wis, US-Gen, and US-Col) conducted separately, the scenario-pruning procedure based on a lightweight ABC-RF analysis reduced the initial 28 hypothetical scenarios to seven plausible scenarios (analysis 2a, Table S3). See the results of the other pruning analyses that we conducted (analyses 3a, 4a, 5a, and 6a) to visualize how effective the procedure can be when applied to an even greater number of scenarios. When we focused our analysis on these seven scenarios, it was inferred that each population derived from the same source population (US-Sdi; analysis 2b; $P = 0.596$, 10/10 runs). When analyzed jointly rather than independently, the four central and eastern US populations were inferred to originate from an unsampled ancestral (meta-)population founded from US-Sdi (analysis 2c; $P = 0.864$, $SD = 0.013$, 10/10 runs; see Fig. S6).

In South America, the Brazilian population (BR-Poa) was best explained by a two-source admixture involving the US western (US-Wat) and eastern (US-Nca) populations (analysis 3b; $P = 0.612$, $SD = 0.016$, 9/10 runs). Focusing on the two closest scenarios provided further support for this conclusion (analysis 3c; $P = 0.843$, $SD = 0.009$, 10/10 runs). The Argentinian population (AR-Gol) was inferred to originate from the Brazilian population (BR-Poa) and the western US population (US-Sdi), with consistent support across runs (analysis 4b; $P = 0.591$, $SD = 0.016$, 10/10 runs). Focusing on the two closest scenarios provided further support for this conclusion (analysis 4c; $P = 0.790$, $SD = 0.015$, 10/10 runs). Separate analyses of the CL-Ar, CL-Bc, and CL-Rm populations in Chile identified the AR-Gol population as their unique source (analysis 5b; $P = 0.610$, $SD = 0.011$, 10/10 runs). Restricting the analysis to the two best-supported scenarios reinforced this result (analysis 5c; $P = 0.812$, $SD = 0.010$; 10/10 runs). When all three populations were considered jointly, the most likely invasion pathway corresponded to a serial colonization originating from Argentina (AR-Gol → CL-Ar → CL-Bc → CL-Rm), with moderate statistical support (analysis 6b; $P = 0.517$, 10/10 runs). Restricting the ABC-RF comparison to the two closest scenarios further confirmed this inference (analysis 6c; $P = 0.668$, 10/10 runs). This invasion scenario supports a continuous latitudinal expansion of *D. suzukii* across Chile.

Under the final invasion scenario, RF-based parameter estimation provided quantitative insights into the admixture rates and bottleneck intensities associated with successive introductions (analysis 7a in Table S3, further detailed in Table 1). Admixture events revealed substantial, albeit heterogeneous, contributions from source populations. For example, the western US population (US-Wat) received a significant contribution from the CN-Lia population in China (median = 0.768, 95% CI: 0.646–0.894), whereas the direct contribution of the Hawaiian population US-Haw to the population US-Sok was more limited (median = 0.234, 95% CI: 0.066–0.441). In South America, the Brazilian population BR-Poa showed balanced admixture from the eastern US population US-Nca (median = 0.502, 95% CI: 0.267–0.763) and the western US population US-Wat. The AR-Gol population in Argentina was

predominantly founded by the BR-Poa population in Brazil (median = 0.741, 95% CI: 0.555–0.907), with a smaller amount of gene content originating from the US-Sdi population in the southwestern US.

Bottleneck intensity estimates revealed significant differences among introduction events. Low bottleneck intensity (i.e. low reductions in effective population size and/or short duration) was inferred for several secondary introduction populations, including US-Col (median bottleneck intensity = 0.027) and the South American populations from Argentina and Chile (medians = 0.066, 0.028, 0.085 and 0.045 for AR-Gol, CL-Ar, CL-Rm, and CL-Bc, respectively). By contrast, high bottleneck intensity was inferred for primary introduction populations, such as the US-Haw island population (median = 0.517) and the initial population in Europe SP-Bar (median = 0.457). Other secondary introduction populations exhibited mild bottleneck signatures, including US-Wat (median = 0.109) and US-Wis (median = 0.145), the Brazilian population BR-Poa (median = 0.203), and the unsampled ancestral population US-Una (median = 0.197) at the origin of central and eastern US populations.

Post-inference goodness-of-fit (post- GoF) indicated an acceptable, albeit moderate, fit of the final invasion scenario depicted in ***Please change by Fig. 3 2*** to the microsatellite data (analysis 8a in Table S3). Using SP-Bar (Barcelona, Spain) as a proxy for the early European introduction of *D. suzukii* yielded a median post GoF *p*-value of 0.132 (95% HDI: 0.026–0.234). Using IT-Tre (Trento, Italy) as European reference yielded a slightly higher *p*-value of 0.218 (95% HDI: 0.024–0.384). Hence, based on median *p*-values, the post GoF tests did not reject the null assumption that the observed data could be generated under these scenarios at the 5% level. Highest Density Intervals (HDIs) were nevertheless wide, and although their lower bounds approached the 1% rejection threshold, they did not fall below it.

Discussion

Using mitochondrial DNA and microsatellite markers, our study discovered patterns of genetic variation in invasive *D. suzukii* across South America, with a particular focus on Chile, and provided insights into invasion dynamics at both continental and regional scales. Our results revealed consistent patterns across datasets, while taking advantage of the finer resolution provided by microsatellites.

Genetic diversity and demographic processes

Both marker systems revealed a sharp reduction in genetic diversity in invasive populations compared to populations in the native range. Chilean populations displayed particularly low mtDNA haplotype diversity, which is consistent with the effects of founding events and bottlenecks that are commonly associated with secondary and recent colonisation (e.g., Lockwood et al. 2013). Microsatellite analyses confirmed these trends, with allelic richness and heterozygosity decreasing progressively from North to South America, reaching their lowest values in Chile. These losses in variability are comparable to patterns reported in other invasive insects undergoing secondary introductions or colonisation events in succession, such as the introduction of the solitary bee *Lasioglossum leucozonium* and the ants *Tetramorium immigrans* and *T. tsushimae* to North America, and the introduction of the pea aphid *Acyrthosiphon pisum* to Chile (Zayed et al. 2007; Peccoud et al. 2008; Flucher et al. 2021). Interestingly, the earliest Chilean sample (CL-Ar2017) collected in 2017 exhibited lower diversity than those collected more recently (i.e., in 2023). This temporal pattern of diversity may reflect ongoing introductions of propagules from Argentina into the Araucanía region (Dlugosch and Parker 2008), where the low altitude of the Andes would represent a less effective geographical barrier to the movement of this fly.

Population structure and geographic patterns

Genetic structure analyses consistently revealed cohesive regional groups. In North America, populations in the eastern and central US showed extremely low differentiation, whereas populations in the western US were more divergent despite being among the earliest invasive populations to be established. This unexpected pattern suggests a complex introduction history in western North America, potentially involving multiple genetically differentiated sources, as

previously noted by Fraimout et al. (2017). In South America, both mtDNA and microsatellites demonstrated genetic continuity between Brazil, Argentina, and Chile. BR-Poa and AR-Gol populations exhibited minimal differentiation, while the three Chilean populations formed a tightly clustered group with F_{ST} values ≤ 0.014, clearly connected to Argentina and, to a lesser extent, Brazil. Within Chile, we observed a slight south-to-north gradient in genetic differentiation, consistent with a pattern of continuous latitudinal expansion after introduction. Similar dynamics of serial expansion and low internal structuring have been reported in other invasive drosophilids (Pascual et al. 2007; Adrion et al. 2014) and other insect invaders (Estoup and Guillemaud 2010; Zepeda-Paulo et al. 2010; Correa et al. 2019).

The STRUCTURE analyses provided a complementary, descriptive view of population structure without specifying alternative invasion routes. In particular, the balanced analysis recovered three major genetic components corresponding broadly to the western US, the central/eastern US and South America, while revealing marked mixed membership in BR-Poa and, to a lesser extent, AR-Gol. Together with the mixed assignment of US-Wat in the worldwide analysis, these patterns were qualitatively consistent with a complex and admixed invasion history. However, STRUCTURE membership coefficients depend on both *K* and the set of populations analysed and should not be interpreted as direct quantitative estimates of historical admixture contributions. In a hierarchical and serial invasion such as that of *D. suzukii*, mixed assignments may reflect admixture, shared ancestry, or genetic drift and do not identify the source, direction, or timing of gene flow. We therefore used STRUCTURE as descriptive corroboration, whereas the specific admixture events, source populations and invasion pathways were inferred by comparing explicit demographic scenarios (i.e., different routes of invasion) within the ABC-RF framework.

Routes of invasion and source populations

Mitochondrial DNA analysis supported a sequential invasion route from North America through South America (US → Brazil → Argentina → Chile), but the resolution was limited due to high posterior predictive error rates when discriminating between scenarios. Microsatellite-based ABC-RF analyses provided a more detailed and quantitative reconstruction of this route. These analyses further supported Argentina as the sole source of the Chilean populations and indicated a sequential colonisation pathway within Chile (AR-Gol → CL-Ar → CL-Bc → CL-Rm).

Interestingly, at a broader scale, the best genetic proxy for the primary native source population was consistently CN-Lia (located in Jilin Province in China) for both the North American and European invasions. Notably, Jilin Province is one of the leading areas for blueberry production in China (blueberries being a preferred host for *D. suzukii*), particularly in the southeast (e.g., the Changbai region), though there is no evidence of intensive strawberry or cherry production in this region (Produce Report, 2021). This result suggests a potential role for specific regions, particularly those characterised by disturbed, intensive agricultural habitats, as recurrent sources of invasive lineages from northeastern Asia. This is consistent with the AIAI (Anthropogenically Induced Adaptation to Invade) framework (Hufbauer et al. 2012; Estoup et al. 2026), which states that populations adapted to human-disturbed environments can repeatedly act as initial or secondary bridgeheads for global invasions, including within the native range of a species.

Limitations and perspectives

Although post-inference goodness-of-fit tests indicated that the retained final invasion scenario provides an acceptable description of the data, the low values of the lower bounds of the confidence intervals suggest that further refinement is possible. Incorporating unsampled populations into invasion models—as illustrated by analysis 2c in this study—or relaxing temporal constraints, particularly in North America where observation dates are tightly clustered, could improve model fit. Advances will also come from denser genomic datasets such as whole-genome sequencing, which offer greater resolution for evaluating alternative invasion routes. Likewise, SNP-based admixture graph approaches (e.g., Gautier et al. 2022), which explicitly account for unsampled populations, provide a powerful complement to

microsatellite DIYABC analyses. Nevertheless, the large and standardised microsatellite dataset used here remains a highly effective and computationally efficient resource, especially for reconstructing recent demographic events at continental and regional scales.

Importantly, unpublished exploratory analyses in which we modified the invasion history in North America indicate that three major conclusions are robust to such modifications: (i) a sequential expansion from North America through South America (US → Brazil → Argentina → Chile); (ii) Argentina as the most likely single source of the Chilean populations; and (iii) a serial colonisation process within Chile. The pronounced genetic homogeneity observed across the sampled area along Chile (i.e., > 600 km of latitudinal range) further suggests a strong role for anthropogenic dispersal (e.g., fruit trade) in shaping the species' rapid spread (Estay et al. 2023). This combination of a single genetic source and rapid spatial expansion offers a unique opportunity to investigate latitudinal selection gradients and climate-related trait evolution across the marked environmental heterogeneity of Chile (Castañeda et al. 2013, 2015).

Genetic monitoring as a tool for pest management

Our study emphasises the importance of genetic monitoring as a tool for pest management. Molecular markers, including microsatellites and, increasingly, single nucleotide polymorphisms (SNPs), can serve as early warning systems by detecting novel introductions, tracking invasion fronts, and identifying emerging genetic lineages with higher adaptive potential (e.g., Estoup et al. 2016, 2026; McGaughran et al. 2024). Evidence from our analyses suggests that a single introduction from Argentina and the low genetic diversity of *D. suzukii* populations in Chile may constrain the adaptive potential of established populations, potentially limiting their ability to evolve resistance to control measures. However, the rapid geographic expansion of *D. suzukii* along Chile can be explained by its broad environmental tolerance and dispersal capacity (Asplen et al. 2015; Fraimout and Monnet 2018). A key risk remains the possibility of secondary introductions from genetically divergent sources, which could restore genetic diversity and increase adaptive potential, which has been reported in other invasive species (e.g., Van Boheemen et al. 2017). Preventing such introductions should therefore be a priority for quarantine and surveillance programmes. Our analyses also suggest that monitoring for the emergence of undesirable genetic variants (e.g. variants that are resistant to control measures) in Chile should primarily focus on South America, specifically Argentina and the USA, as these are the main sources of berry and cherry imports (de la Vega et al. 2020).

Ensuring reliable integration of new genotyping data and efficient statistical methodologies for increasingly massive datasets

In practice, real-time genetic monitoring requires reliable linking of new datasets to existing reference databases. While genotyping can be scaled up at low cost with microsatellites, microsatellite allele scoring can vary across runs, laboratories, or platforms, making this task challenging. A rigorous experimental design is therefore essential. This involves using numerous reference individuals that have already been genotyped as anchors and systematically re-genotyping a subset of old and new individuals to verify concordance. We rigorously employed this strategy to connect new data to existing data and enhance the reliability of the updated global *D. suzukii* microsatellite dataset, which now encompasses 27 populations worldwide.

Another methodological challenge relates to updating invasion scenarios when newly invaded populations are incorporated into analyses. With Approximate Bayesian Computation (ABC), the number of scenarios that need to be compared increases exponentially with the number of potential sources. This makes the process increasingly computationally demanding. This constraint motivated us to explore light and efficient pruning strategies using DIYABC-RF. To this aim, we implemented configurations involving a relatively small number of simulations per scenario (i.e., 2,000) and a limited number of trees (i.e., 500) in the random forests. We found that this heuristic scenario pruning procedure was effective in detecting (very) poorly supported scenarios at a low computational cost and hence retaining only those scenarios that were sufficiently supported (i.e., those that obtained an average number of votes greater

than 5%) for further analysis. For instance, in Analysis 5a, we narrowed down an initial set of 55 scenarios to five realistic alternatives, which were then examined in greater detail. Importantly, when this scenario pruning procedure was applied to numerous population genetics models using both simulated datasets with known true scenarios and real observed datasets, it was found that the best scenario among those compared was never eliminated by the pruning analysis. This remained the case even with fewer simulations per scenario (1,000; results not shown). We found that, despite being less formally developed, this lightweight ABC-RF pruning approach required less computation and outperformed the pre-inference GoF pruning method recently proposed by Le Maillloux et al. (2026); see Figure S7 for details. The lightweight ABC-RF pruning procedure, combined with standard ABC-RF analyses on the reduced set of retained scenarios, therefore provides a practical and efficient way to update invasion-route reconstructions in real time as new invasive populations arise, while preserving computational tractability even as population numbers—and hence candidate scenarios—expand markedly.

Conclusion

Beyond its broader-scale findings, our study sheds new light on the invasion history of *D. suzukii* in South America, particularly in Chile, and makes methodological contributions relevant to pest surveillance and management. Overall, our results suggest that the invasion process through the Americas was complex and sequential. This process was shaped by successive colonisation events involving demographic bottlenecks of various intensities and episodes of admixture events between native and invasive source populations. In Chile, however, the invasion history appears comparatively straightforward, involving a single introduction from Argentina, followed by rapid latitudinal expansion. Despite reduced genetic diversity, *D. suzukii* has successfully colonised a wide range of environments in South America, including a long latitudinal range in Chile. This highlights the importance of demographic history, phenotypic plasticity, and human-mediated dispersal in determining invasion dynamics.

Acknowledgements We thank Ruth Castro of Servicio Agrícola y Ganadero (SAG) for providing samples of *D. suzukii* from La Araucanía. F. Pinilla acknowledges the ANID master scholarship 22231618. We thank to the Associate Editor and reviewers for their constructive and invaluable comments and suggestions.

Author contributions Conceptualization: FP, AE, LEC. Conducting research: FP, TVT, AL, AE, and LEC. Data analysis: FP, AE, and LEC. Interpretation and writing: FP, CCF, PI, LD, JM, AF, GL, PB, FDC, JMM, MAV, MG, AE, and LEC. Funding acquisition: AE, JMM, PB, MG, CCF, PI, and LEC. All authors revised and approved the final version of the manuscript.

Funding This work was funded by the Universidad de Chile through the projects Puente-ICBM and VID-AYV056/01-22 (LEC), and the Agencia Nacional de Investigación y Desarrollo (ANID–Chile) through the projects ANID-Anillo-ATE230025 (LEC, CCF, and PI), ANID-FOVI 230149 (PI and LEC), and ANID-FONDECYT-1210713 (CCF and LEC). This work was also supported by the INRAE Off-site Research Center for the Anticipation of Emerging Pests (ORCAe) (AE and MG) and the France 2030 Maths-VivES programme (AgroStat project, ANR-23-EXMA-0002) handled by the French Agence Nationale de la Recherche (AE, JMM, PB, and MG).

Data availability COI sequences were deposited to GenBank (https://www.ncbi.nlm.nih.gov/) with accession numbers from PX794609 to PX794718. The complete updated microsatellite dataset, coded in a 3-digit GENEPOP format, is publicly in https://doi.org/10.57745/VXLA6UT.

Declarations

Conflict of interest The authors declare no competing interests.

References

Adrion JR, Kousathanas A, Pascual M, Burrack HJ, Haddad NM, Bergland AO, Machado H, Sackton TB, Schlenke TA, Watada M, Wegmann D, Singh ND (2014) *Drosophila suzukii*: the genetic footprint of a recent, worldwide invasion. Mol Biol Evol 31:3148–3163. https://doi.org/10.1093/molbev/msu246

Asplen MK, Anfora G, Biondi A, Choi D-S, Chu D, Daane KM, Gibert P, Gutierrez AP, Hoelmer KA, Hutchison WD, Isaacs R, Jiang Z-L, Kárpáti Z, Kimura MT, Pascual M, Philips CR, Plantamp C, Ponti L, Vétek G, Vogt H, Walton V, Yu Y, Zappalà L, Desneux N (2015) Invasion biology of spotted wing drosophila (*Drosophila suzukii*): a global perspective and future priorities. J Pest Sci 88:469–494. https://doi.org/10.1007/s10340-015-0681-z

Bizama G (2020) Invasión de *Drosophila suzukii* (Matsumura) en Chile: Utilizando los modelos de distribución de especies como herramienta de bioseguridad. Rev Chil Entomol 46:61–71. https://doi.org/10.35249/rche.46.1.20.08

Boughdad A, Haddi K, El Bouazzati A, Nassiri A, Tahiri A, El Anbri C, Eddaya T, Zaid A, Biondi A (2021) First record of the invasive spotted wing drosophila infesting berry crops in Africa. J Pest Sci 94:261–271. https://doi.org/10.1007/s10340-020-01280-0

Calabria G, Máca J, Bächli G, Serra L, Pascual M (2012) First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe. J Appl Entomol 136:139–147. https://doi.org/10.1111/j.1439-0418.2010.01583.x

Camus L, Rode N, Serga S, Loiseau A, Chen X, Iampetro C, Kenis M, Marande W, Mensch J, Parinello H, Savic Veselinovic M, Valerie S, Zhang J, Estoup A, Boitard S, Gautier M (2025) Adaptive challenges of past and future invasion of *Drosophila suzukii*: insights from novel genomic resources and statistical methods combining individual and pool sequencing data. Mol Ecol 72:e70192. https://doi.org/10.1111/mec.70192

Castañeda LE, Balanyà J, Rezende EL, Santos M (2013) Vanishing chromosomal inversions in *Drosophila subobscura* from Chile: is behavioral thermoregulation to blame? Am Nat 182:249–259. https://doi.org/10.1086/671057

Castañeda LE, Rezende EL, Santos M (2015) Heat tolerance in *Drosophila subobscura* along a latitudinal gradient: contrasting patterns between plastic and genetic responses. Evolution 69:2721–2734. https://doi.org/10.1111/evo.12757

Cini A, Ioriatti C, Anfora G (2012) A review of the invasion of *Drosophila suzukii* in Europe and a draft research agenda for integrated pest management. Bull Insectol 65:149–160

Cini A, Anfora G, Escudero-Colomar LA, Grassi A, Santosuosso U, Seljak G, Papini A (2014) Tracking the invasion of the alien fruit pest *Drosophila suzukii* in Europe. J Pest Sci 87:559–566. https://doi.org/10.1007/s10340-014-0617-z

Collin F-D, Durif G, Raynal L, Lombaert E, Gautier M, Vitalis R, Marin J-M, Estoup A (2021) DIYABC random forest v1.0: extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms. Mol Ecol Resour 21:2598–2613. https://doi.org/10.1111/1755-0998.13413

Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M, Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics 30:1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Correa MCG, Palero F, Malausa T, Crochard D, Zaviezo T, Lombaert E (2019) European bridgehead effect in the worldwide invasion of the obscure mealybug. Biol Invasions 21:123–136. https://doi.org/10.1007/s10530-018-1809-x

Cristescu ME (2015) Genetic reconstructions of invasion history. Mol Ecol 24:2212–2225. https://doi.org/10.1111/mec.13117

de la Vega GJ, Corley JC, Soliani C (2020) Genetic assessment of the invasion history of *Drosophila suzukii* in Argentina. J Pest Sci 93:63–75. https://doi.org/10.1007/s10340-019-01149-x

Deprá M, Poppe JL, Schmitz HJ, De Toni DC, Valente VLS (2014) The first records of the invasive pest *Drosophila suzukii* in the South American continent. J Pest Sci 87:379–383. https://doi.org/10.1007/s10340-014-0591-5

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Estay S, Silva CP, López DN, Labra FA (2023) Disentangling the spread dynamics of insect invasions using spatial networks. Front Ecol Evol 11:1124890. https://doi.org/10.3389/fevo.2023.1124890

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what?: reconstructing routes of invasion. Mol Ecol 19:4113–4130. https://doi.org/10.1111/j.1365-294X.2010.04773.x

Estoup A, Ravigné V, Hufbauer RA, Vitalis R, Gautier M, Facon B (2016) Is there a genetic paradox of biological invasion? Annu Rev Ecol Evol Syst 47:51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

Estoup A, Hulme PE, Hodgins KA, Gautier M, Lombaert E, Dhami MK, McGaughran A, Hufbauer RA (2026) Can habitat modification in the native range promote invasion? Trends Ecol Evol. https://doi.org/10.1016/j.tree.2026.06.006

Estoup, A., & Loiseau, A. Worldwide (2026) microsatellite genotypes of Drosophila suzukii at 25 loci across 27 native and invasive populations. Recherche Data Gouv. https://doi.org/10.57745/VXLA6U

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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Lischer HEL (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. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Ferronato P, Woch AL, Soares PL, Bernandi D, Botton M, Andreazza F, Oliviera EE, Corrêa AS (2019) A phylogeographic approach to the *Drosophila suzukii* (Diptera: Drosophilidae) invasion in Brazil. J Econ Entomol 112:425–433. https://doi.org/10.1093/jee/toy321

Filatov DA (2002) Proseq: a software for preparation and evolutionary analysis of DNA sequence data sets. Mol Ecol Notes 2:621–624. https://doi.org/10.1046/j.1471-8286.2002.00313.x

Flucher SM, Krapf P, Arthofer W, Suarez AV, Crozier RH, Steiner FM, Schlick-Steiner BC (2021) Effect of social structure and introduction history on genetic diversity and differentiation. Mol Ecol 30:2511–2527. https://doi.org/10.1111/mec.15911

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Fraimout A, Monnet AC (2018) Accounting for intraspecific variation to quantify niche dynamics along the invasion routes of *Drosophila suzukii*. Biol Invasions 20:2963–2979. https://doi.org/10.1007/s10530-018-1750-z

Fraimout A, Loiseau A, Price DK, Xuereb A, Martin JF, Vitalis R, Fellous S, Debat V, Estoup A (2015) New set

of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): a promising molecular tool for inferring the invasion history of this major insect pest. Eur J Entomol 112:855–859. https://doi.org/10.14411/eje.2015.079

Fraimout A, Debat V, Fellous S, Hufbauer RA, Foucaud J, Pudlo P, Marin J-M, Price DK, Cattel J, Chen X, Deprá M, François Duyck P, Guedot C, Kenis M, Kimura MT, Loeb G, Loiseau A, Martinez-Sañudo I, Pascual M, Richmond MP, Shearer P, Singh N, Tamura K, Xuéreb A, Zhang J, Estoup A (2017) Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. Mol Biol Evol 34:980–996. https://doi.org/10.1093/molbev/msx050

Francis RM (2017) Pophelper: an R package and web app to analyse and visualize population structure. Mol Ecol Resour 17:27–32. https://doi.org/10.1111/1755-0998.12509

García FRM, Lasa R, Funes CF, Buzzetti K (2022) *Drosophila suzukii* management in Latin America: current status and perspectives. J Econ Entomol 115:1008–1023. https://doi.org/10.1093/jee/toac052

Gautier M, Vitalis R, Flori L, Estoup A (2022) F-statistics estimation and admixture graph construction with Pool-Seq or allele count data using the R package poolfstat. Mol Ecol Resour 22:1394–1416. https://doi.org/10.1111/1755-0998.13557

González G, Mary AL, Goñi B (2015) Drosophila suzukii (Matsumura) found in Uruguay. Dros Info Serv 98:103-107

Hauser M (2011) A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. Pest Manag Sci 67:1352–1357. https://doi.org/10.1002/ps.2265

Hufbauer RA, Facon B, Ravigné V, Turgeon J, Foucaud J, Lee CE, Rey O, Estoup A (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. Evol Appl 5:89–101. https://doi.org/10.1111/j.1752-4571.2011.00211.x

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inhereted symbionts. Proc R Soc B 272:1525–1534. https://doi.org/10.1098/rspb.2005.3056

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94. https://doi.org/10.1186/1471-2156-11-94

Kaneshiro KY (1983) Drosophila (Sophophora) suzukii (Matsumura). Proc Hawaii Entomol Soc 24:179

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. Genome Res 12:996–1006. https://doi.org/10.1101/gr.229102

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:11–120. https://doi.org/10.1007/BF01731581

Le Mailloux G, Bastide P, Marin J-M, Estoup A (2026) Goodness of fit for Bayesian generative models with applications in population genetics. Ann Appl Stat, In presse. https://arxiv.org/abs/2501.17107.

Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype networkconstruction. Methods Ecol Evol 6:1110–1116. https://doi.org/10.1111/2041-210X.12410

Lewald KM, Abrieux A, Wilson DA, Lee Y, Conner WR, Andreazza F, Beers EH, Burrack HJ, Daane KM, Diepenbrock L, Drummond FA, Fanning PD, Gaffney MT, Hesler SP, Ioriatti C, Isaacs R, Little BA, Loeb GM, Miller B, Nava DE, Rendon D, Sial AA, SilvaBezerra da CS, Stockton DG, Van Timmeren S, Wallingford A, Walton VM, Wang X, Zhao B, Zalom FG, Chiu JC (2021) Population genomics of *Drosophila suzukii* reveal longitudinal population structure and signals of migrations in and out of the continental United States. G3 (Bethesda) 11:jkab343. https://doi.org/10.1093/g3journal/jkab343

Lin QC, Zhai YF, Zhang AS, Men XY, Zhang XY, Zalom FG, Zhou CG, Yu Y (2014) Comparative developmental times and laboratory life tables for *Drosophila suzukii* and *Drosophila melanogaster* (Diptera: Drosophilidae). Fla Entomol 97:1434–1442. https://doi.org/10.1653/024.097.0418

Lockwood J, Hoopes M, Marchetti M (2013) Invasion ecology, 2nd ed. Wiley-Blackwell, Oxford

Matsumura S (1931) 6000 illustrated insects of Japan-empire. Toko Shoin, Japan

McGaughran A, Dhami MK, Parvizi E, Vaughan AL, Gleeson DM, Hodgins KA, Rollins LA, Tepolt CK, Turner KG, Atsawawaranunt K, Battlay P, Congrains C, Crottini A, Dennis TPQ, Lange C, Liu XP, Matheson P, North HL, Popovic I, Rius M, Santure AW, Stuart KC, Tan HZ, Wang C, Wilson J (2024) Genomic tools in biological invasions: current state and future frontiers. Genome Biol Evol 16:evad230. https://doi.org/10.1093/gbe/evad230

ODEPA (2024) Bases de datos comercio exterior. ODEPA | Oficina de Estudios y Políticas Agrarias. Ministerio de Agricultura. https://www.odepa.gob.cl/estadisticas-del-sector/bases-de-datos-comercio-exterior. Accessed 13 Oct 2025

Pascual M, Chapuis MP, Mestres F, Balanyà J, Huey RB, Gilchrist GW, Serra L, Estoup A (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite based survey using ABC methods. Mol Ecol 16:3069–3083. https://doi.org/10.1111/j.1365-294X.2007.03336.x

Peccoud J, Figueroa CC, Silva AX, Ramirez CC, Mieuzet L, Bonhomme J, Stoeckel S, Plantegenest M, Simon JC (2008) Host range expansion of an introduced insect pest through multiple colonizations of specialized clones. Mol Ecol 17:4608–4618. https://doi.org/10.1111/j.1365-294x.2008.03949.x

Posit team (2023) RStudio: integrated development environment for R. Posit Software, PBC, Boston, MA.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. https://doi.org/10.1093/genetics/155.2.945

Pudlo P, Marin JM, Estoup A, Cornuet JM, Gautier M, Robert CP (2016) Reliable ABC model choice via random

forests. Bioinformatics 32:859–866. https://doi.org/10.1093/bioinformatics/btv684

Raymond M, Rousset F (1995) Genepop (version. 1.2), a population genetics software for exact tests and ecumenicism. J Hered 86:248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Raynal L, Marin J-M, Pudlo P, Ribatet M, Robert CP, Estoup A (2019) ABC random forests for Bayesian parameter inference. Bioinformatics 35:1720–1728. https://doi.org/10.1093/bioinformatics/bty867

Produce Report. 2021. China Blueberry Report 2020: Industry Growing Rapidly. https://www.producereport.com/article/china-blueberry-report-2020-industry-growing-rapidly. Accessed 13 Oct 2025

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34:3299–3302. https://doi.org/10.1093/molbev/msx248

SAG (2025) Drosófila de alas manchadas (*Drosophila* suzukii): Estatus fitosanitario de la plaga en las regiones del territorio chileno. https://www.sag.gob.cl/ambitos-de-accion/estatus-fitosanitario-de-la-plaga-en-las-regiones-del-territorio-chileno. Accessed 13 Oct 2025

Santandino MV, Riquelme Virgala MB, Ansa MA, Bruno M, Di Silvestro G, Lunazzi EG (2015) Primer registro de Drosophila suzukii (Diptera: Drosophilidae) asociado al cultivo de arándanos (Vaccinium spp.) de Argentina. Rev Soc Entomol Argent 74:183–185

Satta Y, Ishiwa H, Chigusa I (1987) Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila melanogaster* and its sibling species. Mol Biol Evol 4:638–650. https://doi.org/10.1093/oxfordjournals.molbev.a040464

Sentis A, Hemptinne JL, Magro A, Outreman Y (2022) Biological control needs evolutionary perspectives of ecological interactions. Evol Appl 15:1537–1554. https://doi.org/10.1111/eva.13457

Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol Biol Evol 13:510–524. https://doi.org/10.1093/oxfordjournals.molbev.a025612

Van Boheemen LA, Lombaert E, Nurkowski KA, Gauffre B, Rieseberg LH, Hodgins KA (2017) Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. Mol Ecol 26:5421–5434. https://doi.org/10.1111/mec.14293

Wright ES (2016) Using DECIPHER v2.0 to analyze big biological sequence data in R. R J 8:352. https://doi.org/10.32614/RJ-2016-025

Zayed A, Constantin ŞA, Packer L (2007) Successful biological invasion despite a severe genetic load. PLoS ONE 2:e868. https://doi.org/10.1371/journal.pone.0000868

Zepeda-Paulo FA, Simon JC, Ramirez CC, Fuentes-Contreras E, Margaritopoulos JT, Wilson ACC, Sorenson CE, Briones LM, Azevedo R, Ohashi DV, Lacroix C, Glais L, Figueroa CC (2010) The invasion route for an insect pest species: the tobacco aphid in the New World. Mol Ecol 19:4738–4752. https://doi.org/10.1111/j.1365-294x.2010.04857.x

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.