Molecular tools to study historic and recent pathways of entry and spread for the invasive Japanese beetle

DIPARTIMENTO SCIENZE DELLA VITA

DOTTORATO DI RICERCA IN SCIENZE DELLA VITA
LIFE SCIENCES

CICLO XXXVII

COORDINATRICE Prof. ssa Simona Maccherini

MOLECULAR TOOLS TO STUDY HISTORIC AND RECENT PATHWAYS OF ENTRY AND SPREAD FOR THE INVASIVE JAPANESE BEETLE

SETTORE SCIENTIFICO-DISCIPLINARE: *BIOS-03/A*

TUTOR: Prof. Antonio Carapelli

DOTTORANDA: Dr. ssa Rebecca Funari

A.A. 2023-2024

A Eva, Marta e Jan

Table of contents

Abstract

Invasive species are an increasing global threat given their ability to rapidly spread and adapt to novel environments. The adverse ecological and economic impacts of invasive species highlight the critical need for understanding the mechanisms that underpin invasion processes and success. The Japanese beetle, *Popillia japonica* is an invasive pest of remarkable interest, as it feeds on hundreds of plant species of economic value. It has been expanding outside of its native range in Japan since the early decades of the 20th century, colonizing large areas of North America and, more recently, Europe.

The availability of a high-quality genomic sequence for this species is poised to significantly advance fundamental research on its ecology and evolution, as well as enable the development of genetically-informed and biotechnologically-oriented control strategies. Compared to previous draft versions, the *P. japonica* sequence, here presented, offers greater completeness and contiguity of genomic data, as well as comprehensive structural and functional annotations. This new genomic data may provide insights into the species' invasive potential and serve as a foundation for identifying potential biotechnological applications aimed at its control.

Subsequently, whole-genome resequencing data from individuals across the species' entire distribution were analysed to investigate the geographic differentiation of *P. japonica* populations and to reconstruct the expansion routes from Japan to the USA and Europe. The analysis identified six genomically distinct clusters, which correspond to approximate colonization areas on a continental scale. Coalescent simulations further uncovered two independent bridgehead events originating from the USA to the Azores and Italy, highlighting complex patterns of secondary spread.

In addition, an investigation into selection signals was conducted to better understand the adaptive mechanisms underpinning *P. japonica*'s invasion success. However, the lack of strong selection signatures in invasive specimens suggested that the beetle's adaptive capabilities may be rooted in pre-existing genomic features. The comprehensive genome-wide dataset enabled detailed inference

of the invasion process and may aid in identifying the origins of *P. japonica* samples in future invasion events.

Riassunto

Data la loro capacità di diffondersi e adattarsi rapidamente a nuovi ambienti, le specie invasive rappresentano una crescente minaccia a livello globale. Le conseguenze ecologiche ed economiche delle specie invasive evidenziano la necessità di comprendere i meccanismi alla base dei processi di invasione e del loro successo. Il coleottero giapponese, *Popillia japonica*, è un insetto invasivo infestante di notevole interesse, poiché si nutre di centinaia di specie vegetali di valore economico. Espandendosi al di fuori del suo areale nativo in Giappone fin dai primi decenni del XX secolo, questo insetto ha colonizzato ampie aree del Nord America e, più recentemente, dell'Europa.

La disponibilità di una sequenza genomica di alta qualità permette di far progredire significativamente la ricerca sull'ecologia e sui processi evolutivi di questa specie, oltre a consentire lo sviluppo di strategie di controllo basate su conoscenze genetiche e orientate verso applicazioni biotecnologiche. Rispetto alle precedenti versioni, la sequenza genomica di *P. japonica* qui presentata offre una maggiore completezza e contiguità dei dati, nonché annotazioni strutturali e funzionali complete, dando la possibilità di fornire informazioni sull'invasività della specie e servire come base per sviluppare strategie di controllo integrato.

L'analisi dei dati di whole genome resequencing, da campioni raccolti nell'intera distribuzione della specie, ha permesso di studiare la differenziazione geografica delle popolazioni e ricostruire le rotte di espansione dal Giappone agli Stati Uniti e all'Europa. Il presente studio ha identificato sei gruppi geneticamente distinti, che corrispondono alle aree di colonizzazione su scala continentale, ed ha evidenziato due eventi di invasione indipendenti dagli Stati Uniti verso le Azzorre e verso l'Italia, mostrando modelli complessi di diffusione secondaria.

Infine, uno studio sulla presenza di possibili segnali di selezione, nelle popolazioni non native, è stato condotto per comprendere meglio i meccanismi adattativi che sostengono il successo dell'invasione di *P. japonica*. Tuttavia, l'assenza di forti segnali di selezione negli esemplari invasivi suggerisce che le capacità di adattamento del coleottero potrebbero derivare da caratteristiche genomiche preesistenti. Il set di dati genomici presentato in questo lavoro ha permesso una comprensione dettagliata dei processi di invasione e potrebbe contribuire all'identificazione delle origini degli esemplari di *P. japonica* in futuri eventi di invasione.

1. Introduction

1.1 Invasive species

Invasive species are one of the utmost important environmental challenges of contemporary times. Over the past few centuries, increased human activities have greatly facilitated the transport of organisms, resulting in a higher frequency of biological invasions (Hulme et al., 2009; Seebens et al., 2018). This process entails the intentional or unintentional introduction of an organism into a habitat where it did not previously exist, leading to subsequent detrimental effects on native organisms and ecosystems (Alpert, 2006). Indeed, species that are moved beyond their natural limits by human actions, establishing themselves outside their native ranges, and becoming invasive, can disrupt the delicate balance of ecosystems, causing extensive damage to biodiversity, human health, and the global economy (Kolar & Lodge, 2001; Simberloff, 2009). Over the past two centuries, alongside increased import volumes and human mobility, alien and invasive species have reshaped the boundaries of biogeography, decreasing the differences in species composition among distant regions (Capinha et al., 2015; Hulme et al., 2009). This can profoundly alter ecosystems, affecting the availability of resources, the structure of trophic webs, and interactions within communities (Hoffmeister et al., 2005). The global financial toll of biological invasions has reached at least US$1.288 trillion, with an annual average cost of US$26.8 billion. However, these estimates might be conservative, as the true costs are significantly underestimated and have been escalating threefold each decade without any indication of slowing down (Diagne et al., 2021). Among animals, insects are the group with the highest number of reported establishments outside their native range, outnumbering invasions of all other taxa (Pyšek et al., 2020; Seebens et al., 2017). Moreover, in the list of the 100 world's worst invasive alien species by the IUCN (International Union for the Conservation of Nature, available at: http://www.iucngisd.org/gisd/100_worst.php), twelve are

insects. This list includes *Aedes albopictus* and *Anopheles quadrimaculatus*, two mosquito species associated with the transmission of many human diseases. The former, also known as the Asian tiger mosquito, was spread via the international trade of tires and is a vector for Chikungunya, Dengue, yellow fever, and Zika flaviviruses (Näslund et al., 2021). The latter is a vector of many animal and human pathogens, the most important of which are the malaria-causing *Plasmodium* spp. (Carpenter & LaCasse, 1955). A recent example of a destructive invasive species is the brown marmorated stink bug *Halyomorpha halys*. Originally from Asia, it quickly spread across the United States in the mid-1990s and was subsequently detected in Canada and various regions in Europe, including Italy, making it a global invasive species (Lee, 2015). Besides being a nuisance for invading homes and buildings in large numbers during the winter, this bug has become a significant pest, causing severe damage to both annual and perennial crops and resulting in substantial economic losses (Lee, 2015).

It is essential to consider factors that contribute to a species becoming invasive and how these factors interact within different ecosystems. Despite a paucity of comprehensive information on intrinsic traits that confer invasiveness to a species, previous studies have highlighted the significant potential of invasive populations for rapid adaptation (Hulme et al., 2009; Prentis et al., 2008). However, others suggest that many invasive species exhibit a predisposition to invasiveness, possessing traits in their genetic makeup that facilitate spread into new environments—traits that existed in native populations before transportation and are not a consequence of rapid adaptation following colonization (Estoup et al., 2016; North et al., 2021). Shared characteristics among diverse invasive species might include the ability to tolerate a wide range of climatic and environmental conditions, as well as effective dispersal and hitchhiking capabilities (Borner et al., 2023a; Whitney & Gabler, 2008). Introduced species may prevail in competition with native species or exploit vacant niches in the new ecosystem, behaving as opportunists (Alpert, 2006; Daly et al., 2023; Vall-llosera et al., 2016). Furthermore, the absence of natural predators and pathogens in the new environment provides a significant advantage for invaders (Alpert, 2006). Another factor influencing the

192 establishment and invasiveness of introduced populations is the magnitude of the introduction effort,
193 also referred to as propagule pressure, which depends on the number of released individuals and the
194 number of introduction attempts (Simberloff, 2009). High propagule pressure increases the likelihood
195 of successful introductions by mitigating the negative effects associated with small initial population
196 sizes and by increasing genetic diversity (Alpert, 2006; Bock et al., 2016), thereby promoting
197 admixture and enhancing the adaptive potential of these populations (Bock et al., 2015). Biological
198 invasions are also driven by the susceptibility of the recipient environment. This invasibility results
199 from factors such as human disturbance, habitat fragmentation, and the presence of unsaturated
200 ecosystems—characterized either by a short evolutionary history or specialized interactions between
201 organisms—with empty niches (Alpert, 2006; Daly et al., 2023; Hoffmeister et al., 2005; Richardson
202 & Pyšek, 2006). Environmental heterogeneity can increase invasibility, although it may also mitigate
203 impacts on native species due to the coexistence of multiple interactions absent in homogeneous
204 environments (Melbourne et al., 2007).
205
206 Despite challenges posed by low genetic diversity characterizing bottlenecked populations
207 with limited evolutionary potential, invasive alien species can establish and succeed in new
208 environments, giving rise to the genetic paradox of invasion (Frankham, 2005). Numerous studies
209 have sought an explanation for this paradox by illustrating how mechanisms dependent on high
210 propagule pressure, reproductive traits, and genetic mechanisms could counteract the negative effects
211 of depleted genetic diversity (Estoup et al., 2016; Schrieber & Lachmuth, 2017). For instance, the
212 reduction in diversity at a small set of neutral genetic loci does not necessarily correlate with genetic
213 variation in ecologically significant traits (Estoup et al., 2016). Moreover, genetic mechanisms such
214 as epigenetic processes, already existent preadaptive traits, or *de novo* mutations, along with the purge
215 of deleterious alleles, might increase the population fitness and compensate for the loss in genetic
216 diversity (Estoup et al., 2016; Schrieber & Lachmuth, 2017). In addition, invasion scenarios can be
217 particularly complex due to the potential for multiple introductions from various source areas (Kolar

& Lodge, 2001). For example, invasions can originate from successfully introduced populations, especially in highly connected hubs, which serve as sources for subsequent invasions through bridgehead events (Bertelsmeier & Keller, 2018; Lombaert et al., 2010). Bridgehead introductions can result in significant genetic depletion in introduced populations due to repeated bottleneck events. However, while some species simply tolerate a reduction in genetic diversity, others benefit from purging deleterious alleles during recurrent founder effects (Estoup et al., 2016; Roman & Darling, 2007). Consequently, bridgehead events may also enhance invasiveness through the acquisition of traits that increase the likelihood of successful establishment and spread, thereby acting as a stepping-stone for further invasion (Bertelsmeier and Keller, 2018; Blumenfeld et al., 2021).

1.2 Molecular approaches to invasion biology

The availability of a high-quality genomic sequence for a species can enhance research on its ecology and evolution and assist biotechnologically and genetically oriented control measures. However, the application of biotechnological solutions for control, or the use of molecular genetic data to guide non-biotechnological methods, remains limited, with few exceptions (Carroll et al., 2023). This limitation may arise from the unavailability or incompleteness of genomic data for the species. Nevertheless, recent advances in sequencing technology, data analysis, and progressively decreasing costs have significantly improved the sequencing and assembling of complex genomes, both for model and non-model species, including invasive ones (Matheson & McGaughran, 2022; McCartney et al., 2019).

Whole genomes enable various direct applications, such as providing qualitative information about genome structure, contributing to systematic and evolutionary studies, and examining gene family expansions to assess the role of preinvasion selection in invasion success (McCartney et al., 2019; North et al., 2021). In a globalized world, the number of invasive species is expected to

243 increase, and the availability of genomic data can help to develop biotechnological control strategies,
244 such as the sterile insect technique and gene drives (McGaughran et al., 2024). The sterile insect
245 technique involves the mass release of sterile individuals and has successfully eradicated the invasive
246 pink bollworm *Pectinophora gossypiella* in Arizona in 2013 (Tabashnik et al., 2021). Gene drives
247 were developed though the CRISPR-Cas9 gene-editing system. Being inherited at higher rates
248 compared with other alleles, selected genes "drives" rapidly and spread throughout a population and
249 can be used to target invasive species (Champer et al., 2016). A key example includes a CRISPR–
250 Cas9 gene drive targeting the *doublesex* gene in caged *Anopheles gambiae* mosquitoes, vectors of
251 human malaria (Kyrou et al., 2018). In these insects, the *doublesex* gene controls sex differentiation
252 through female- and male-spliced transcripts. Targeted disruption aimed at preventing the functional
253 formation of the female transcript resulted in homozygous females developing intersex phenotypes
254 and complete sterility. The gene drive construct spread rapidly in caged mosquitoes, achieving 100%
255 prevalence within a few generations, progressively reducing egg production and leading to complete
256 population suppression (Kyrou et al., 2018).
257
258 Another powerful application focuses on using genomes as references, which allows
259 remapping of raw reads from Whole Genome Resequencing (WGR) of multiple samples of the same
260 or of related species. This facilitates the identification of genetic variants among individuals and
261 populations (Chown et al., 2016). Population genomic approaches, particularly using whole-genome
262 resequencing data, have been integrated into invasion science in recent decades (Matheson and
263 McGaughran, 2022; North et al., 2021). They have proven effective in identifying distinct genetic
264 groups or clusters, studying genetic relationships across populations, inferring demographic patterns,
265 and most importantly, in reconstructing invasion routes (North et al., 2021). Investigating global
266 genomic diversity and population structure of invasive species can provide insights into the
267 evolutionary mechanisms underlying invasions and, along with the inference of invasion pathways,
268 facilitate the development of effective strategies for invasive species prevention, management, and

control (Estoup & Guillemaud, 2010; North et al., 2021). Finally, given that most human-mediated invasions occurred over just decades, WGR has contributed, through whole-genome scans, to identifying genetic loci that undergo selection and studying the rapid adaptive evolution during invasion (McCartney et al., 2019; North et al., 2021).

1.3 The invasive Japanese beetle *Popillia japonica*

The Japanese beetle, *Popillia japonica* Newman (Coleoptera: Scarabaeidae: Rutelinae; Figure 1), is an invasive insect native to Japan, reported on all four major islands (Fleming, 1972). At the beginning of the 20th century, it started to expand its distribution (Figure 2, occurrences downloaded from GBIF: https://doi.org/10.15468/dl.2hx5s2). The species was first reported outside its native range in Riverton, New Jersey (USA) in 1916, with more recent investigations tracing its introduction to six years earlier, via a shipment of iris bulbs from Japan (Fleming, 1972; Frank, 2016). In 1944, *P. japonica* was detected in Canada near Halifax, Nova Scotia, likely due to a range expansion from the

Figure 1. Adult specimen of *Popillia japonica*.

USA (Althoff & Rice, 2022; Nardi et al., 2024; Strangi et al., 2024). In 1970, the beetle was found on Terceira Island (Azores) near the Lages American military airbase, followed by further spread to other Azorean islands, such as São Miguel in 2003 and São Jorge in 2007 (Simões, 1984; EPPO, 2019; Teixeira et al., 2024). The first record of *P. japonica* in continental Europe was in 2014 in Turbigo, Milano (Italy), near both the Milano-Malpensa international airport and the Cameri military airport (Pavesi, 2014). Rapidly expanding its range, the Italian population reached the Swiss border, with the earliest report in Stabio, Canton Ticino, in 2017 (Servizio fitosanitario cantonale, 2017). Since its first report in continental Europe, the Japanese beetle has expanded its range to other regions in Northern Italy, such as Valle d'Aosta, Emilia-Romagna, and, more recently, Friuli-Venezia Giulia, as well as into Switzerland (Figure 3; Poggi et al., 2023; ERSA, 2023). In Switzerland, the beetle has increased its distribution in the Ticino Canton, and a new outbreak was recently reported in Kloten, near Zurich, in 2023 (Agroscope, 2023). The Japanese beetle is also present in the Far East of Russia; however, it is known to occur not on mainland, but only on the Russian island of Kunashir, less than 30 km from the Japanese island of Hokkaido (EFSA Panel on Plant Health (PLH) et al., 2018). The interceptions of a total of four specimens in Germany and the Netherlands in 2018, 2021, and 2022 were likely due to occasional human-mediated transport, while reports in other countries (e.g., China, Taiwan, South Korea, and India) are considered invalid or misidentifications (EPPO, 2022).

Figure 2. Global distribution of *Popillia japonica.*

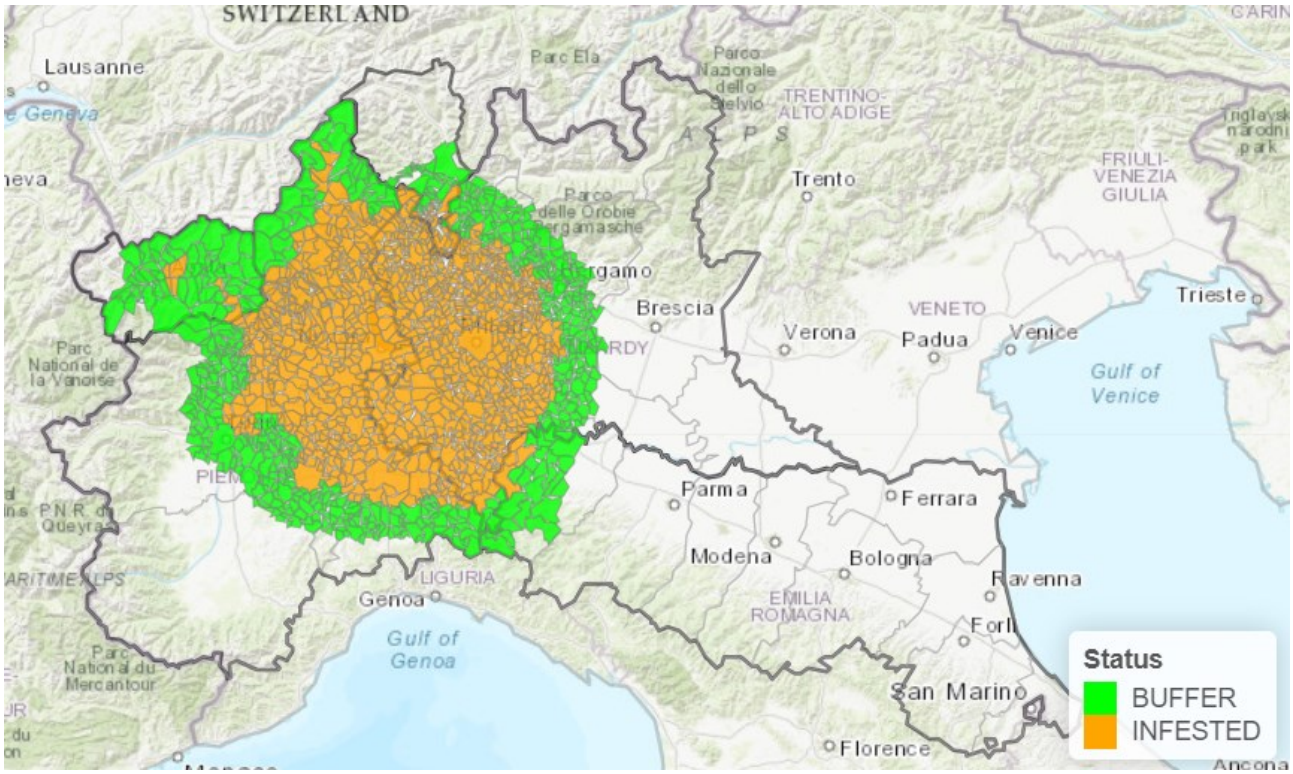


Figure 3. Distribution of *Popillia japonica* in Italy and Switzerland updated to 2022.

The species is univoltine across most of its range (Fleming, 1972; Potter & Held, 2002). In colder regions of the United States, Southern Ontario in Canada, and the Hokkaido Islands in Japan, the life cycle spans two years (Clausen et al., 1927). However, increasing global temperatures might cause a shift from biannual to annual life cycles in many areas (Kistner-Thomas, 2019), posing the risk of more frequent infestations. As reported by Potter and Held (2002), the emergence of adults typically begins by June. Males appear a few days earlier than females, which are inseminated as they first emerge from the soil, likely before feeding. Females alternate between feeding and oviposition, typically mating more than ten times. They enter the soil multiple times, depositing 40 to 60 eggs in the upper 7.5 cm of soil over a four-to-six-week adult lifespan. Eggs hatch after ten to 14 days. The first and second instars develop in about two to three and three to four weeks, respectively, reaching the third instar stage during September and feeding on herbaceous plant roots. To protect themselves from freezing, the grubs move deeper into the soil in late autumn, overwintering five to 15 cm below the surface. As soil temperatures rise above 10°C, larvae move upwards and feed for four to eight

315 weeks before pupation. The prepupal and pupal stages usually occur from April to May, lasting around
316 two to three weeks.
317
318 The European Commission has added *Popillia japonica* to the list of priority pests since 2019.
319 It is a major agricultural pest due to its highly polyphagous feeding behaviour on over 400 species of
320 wild and cultivated plants (Althoff & Rice, 2022; Poggi et al., 2022; Tayeh et al., 2023). Both larvae
321 and adults act as destructive plant pests, potentially affecting native biota (Baker & Potter, 2018).
322 Larvae feed on plant roots of herbaceous plants, causing extensive damage in lawns, parks, play
323 fields, golf courses, and pastures (Potter & Held, 2002). Adults mainly consume foliage, flowers, and
324 fruits of tree and shrub plants, with notable yield reductions in highly infested zones (Potter & Held,
325 2002). *P. japonica* can rapidly expand from newly invaded areas, with an estimated range increase
326 rate of approximately 10 km per year (Fleming, 1972; Mondino et al., 2022). In the Northern
327 hemisphere, habitat suitability modelling suggests that the Eastern USA and the Azores archipelago
328 are highly suitable areas for this species, and Central Europe, including currently invaded areas, faces
329 a high risk of further invasion across the continent (Borner et al., 2023b). In addition, increasing
330 global temperatures are expected to shift the suitable range for *P. japonica* northwards (Kistner-
331 Thomas, 2019). The reported great hitchhiking ability of this beetle also poses a high risk for further
332 unintentional human-mediated spread from currently infested areas, as it can hitch rides via road, rail,
333 and air (Borner et al., 2023a; Poggi et al., 2022).
334
335 In the USA, the economic impact of damage and control exceeds $460 million annually, while
336 in Europe, future damage to crops in the absence of management is estimated to range from €30
337 million to €7.8 billion annually (Straubinger et al., 2022). *P. japonica* continues to constitute a
338 significant threat to agriculture and horticulture due to the diverse damage caused by larvae and adults
339 and increasing restrictions on insecticide use, with only a few successful cases of local eradication
340 (Althoff & Rice, 2022; Gotta et al., 2023; Potter & Held, 2002). Several approaches have been

employed to control its expansion and population size, such as (a) biological methods (e.g., entomopathogenic nematodes, fungi, and natural predators), (b) chemical methods (e.g., α-cypermethrin, deltamethrin), and (c) physical methods (e.g., attract-and-kill devices or nets) (Althoff and Rice, 2022; Gotta et al., 2023). However, despite numerous treatments tested, identifying the most effective methodological protocol to control this insect in all its life stages in Europe has remained elusive. In this context, the potential for biotechnological applications or the integration of molecular genetic information to guide the development of both biotechnological and non-biotechnological approaches may play a crucial role in designing control strategies for *P. japonica*. However, this potential has been limited, with some exceptions (Carroll et al., 2023), by the absence of a high-quality, annotated genome for the species. For comparison, among the 20 quarantine pests identified as high-priority targets by the EU in 2019, predominantly insects, approximately 60% have at least one available genome sequence, 25% are currently undergoing genome sequencing projects, and only 15% appear to lack any genomic data (data retrieved from: https://www.ncbi.nlm.nih.gov/datasets/genome/). Also noteworthy, despite Coleoptera being a well-known and studied group of insects, there are only 319 reference genomes available in the National Center for Biotechnology Information (NCBI) data base, compared to nearly 390,000 described species. Going into more detail, for Scarabaeidae, there are currently only 38 reference genomes, with annotations available for only three, including the one described in this study. Significant studies on the biology and control of *Popillia japonica* have emerged in recent decades, focusing on its biology, phenology, and possible chemical or biological control methods (Potter & Held, 2002; Althoff & Rice, 2022). The first complete mitochondrial genome of *Popillia japonica* was published by Yang et al. (2018), and available on NCBI (accession number: NC_038115.1). An improved version of the mitochondrial genome, assembled with short and long reads, was deposited in NCBI within BioProject PRJNA860365, by Nardi et al., (2024) from the University of Siena, and used as new reference. Early genomic data generation efforts for *P. japonica* include transcriptome sequencing at the University of Southern Mississippi in 2013 and genome sequencing by Iridian Genomes in 2016.

The previous reference genome for the species was produced by Canada's Genomic Enterprise (CGE) within the CanSeq150 initiative and released in 2019 under BioProject PRJNA514809 (NCBI).

1.4 Aims

Reference genome of Popillia japonica

Advances in sequencing technology and data analysis, coupled with a continuous reduction in costs, have increased the rate of genome releases. Moreover, updating previously published genomes with higher quality versions is becoming more frequent, with the latest sequencing of the human genome serving as a prime example (Nurk et al., 2022). Insect genomes are also experiencing the same expected phenomenon. The genome produced by Canada's Genomic Enterprise (CGE), obtained from a male specimen of *P. japonica* collected in Vancouver and representative of the North American invasive population, has limitations: it was assembled solely with Illumina short reads, lacks structural or functional annotations, and has not been formally documented in a published paper. The present research, conducted in collaboration with other members of the research group, provides a new reference genome of *P. japonica* sampled in Cameri (Northern Italy), close to the site of the original Italian introduction. Compared to previous work, this new reference genome is of higher quality due to its enhanced completeness, detailed structural analysis and description, and annotation, providing valuable data that can be utilized for future applications in the design of control strategies, evolutionary studies, insect genomics, and invasion biology.

Population genetics, invasion pathways and signals for selection

Previous studies have explored the invasion pathways of *P. japonica* through the analysis of complete mitochondrial genome sequences, microsatellite loci, and cytochrome oxidase subunit I (*cox1*) and

cytochrome B (*cytb*) mitochondrial genes (Nardi et al., 2024; Strangi et al., 2024). However, they do not account for the full variability of *P. japonica* populations. The present study employs high-resolution genome-wide single nucleotide polymorphisms (SNPs) based on WGR data. Such data significantly enhance population genomic analysis by increasing resolution in reconstructing origins and routes of expansion of invasive species, particularly in systems with recent introduction events (North et al., 2021). The goal is to provide insights into the global genomic diversity and population structure of *P. japonica* and infer the pathways of its more than 100-year history of human-mediated invasion through demographic modelling. Finally, for the first time in this species, the present research investigates the presence of selection signatures by utilizing the aforementioned annotated genome and performing genome-scan analyses to identify divergent loci and putative genes of interest between native and invasive populations.

2. Materials and methods

2.1 The genome of *Popillia japonica*: sampling, extraction and sequencing

Adult specimens of *Popillia japonica* were collected in 2020 and 2021 during the summer from the outskirts of Cameri (NO, Italy; 45.5128460° N, 8.6908590° E), approximately 4 km from the site of the first Italian report in 2014 (Turbigo, MI, Italy; Pavesi, 2014). Additional samples of larvae and pupae, intended for RNA sequencing, were gathered from nearby locations. Specifically, these additional locations were Novara (NO, Italy; 45.44316° N, 8.66766° E) for larvae, and Briona (NO, Italy; 45.541974° N, 8.531064° E) for pupae. Samples were frozen in liquid nitrogen and stored at -80°C until processing. As discussed by Nardi et al. (2024), these specimens originate from a population with extremely low genetic variability, thereby reducing the drawbacks of using multiple individuals for genome assembly. Genome sequencing was performed on males, as they are the heterogametic sex, carrying both the X and Y chromosomes (Yosida, 1949).

High molecular weight DNA was extracted from dissected male testes, minimizing foreign contaminants, using the Wizard® Genomic DNA Purification Kit, following the manufacturer's instructions. Three males (DMR4, DMR5, and DMR6 in NCBI) were processed individually for the creation of three libraries with the Illumina DNA Prep Tagmentation kit. These libraries were sequenced using a 250 bp paired-end (PE) layout on an S4 flow cell in a NovaSeq6000 machine at the Department of Medical Biotechnologies of the University of Siena. For long-read sequencing, two libraries, from one male (DMR5 in NCBI) and a pool of three males (DMR28 in NCBI), were prepared using the Ligation sequencing kit and sequenced with a FLO153 MIN106 flow cell on an Oxford Nanopore MinION machine at Bio-Fab Research (Rome, Italy).

Total RNA for transcriptome sequencing was extracted from the whole bodies of individual specimens using the QIAGEN RNeasy Micro kit, including QIAshredder and DNase treatments. Two replicates for each of four different sexes/stages were used to enhance representativeness and diversity in the transcriptomic output. Specifically, the samples included two larvae, two pupae, two adult females, and two adult males for both total RNA and mRNA sequencing (respectively identified in NCBI as RL11, RL14, RP6, RP15, RF9, and RF15). Libraries were prepared with the TruSeq Stranded Total RNA with Ribo-Zero kit for total RNA sequencing and the TruSeq Stranded mRNA kit for mRNA sequencing. The 16 libraries were sequenced using a 150 bp PE layout on a NovaSeq6000 machine at Macrogen (Amsterdam, Netherlands).

The comprehensive sequencing strategy employed ensured high-quality genome and transcriptome assemblies, providing a robust foundation for subsequent genetic and functional analyses.

2.2 Sequence pre-processing

Illumina raw reads were checked using *FastQC* v0.11.9 (Andrews, 2010) and trimmed with *Cutadapt* v3.4 (Martin, 2011) and *Trimmomatic* v0.33 (Bolger et al., 2014) under default settings to remove adapter sequences, primers, poly-A tails, and other unwanted sequences. Nanopore raw reads were checked using the same approach and trimmed with *NanoFilt* v2.8.0 (De Coster et al., 2018) with a minimum quality score of 8 and a minimum read length of 1500 bp. To identify reads of mitochondrial origin, sequences were preliminarily mapped to the mitochondrial genome available in NCBI (accession number: NC_038115; Yang et al., 2018) using *Minimap2* (Li, 2018), and mapping reads were removed prior to the assembly.

2.3 Genome assembly

The three Illumina DNA libraries were assembled separately using a *de novo* approach in *ABySS* v2.3.7 (Jackman et al., 2017) with k=96 and default settings. The two Nanopore libraries were combined to increase read coverage and assembled *de novo* using *Flye* v2.9.1 (Kolmogorov et al., 2019), performing six rounds of polishing. The resulting assemblies were merged into a single, high-quality assembly using *MAC* v2.0 (Tang et al., 2020). Redundancies and contaminants were removed with *purge_haplotigs* v1.1.2 (parameters: low = 25 mid = 150 high = 275; (Roach et al., 2018), yielding a primary assembly for further analysis. RNA sequencing data were aligned to the primary assembly using *HISAT2* (Kim et al., 2019), and scaffolding was carried out with *P_RNA_scaffolder* (Zhu et al., 2018). A second *purge_haplotigs* run was then conducted to remove any remaining duplicates or contaminants identified during scaffolding. Finally, the *Foreign Contamination Screen* (Astashyn et al., 2024) was used to detect and eliminate any additional contaminant sequences, including those from bacteria, fungi, protists, viruses, and technical sequences. This process resulted in the final assembled genome.

2.4 Genome structural and functional annotation

A custom library of repetitive elements was generated by scanning the primary *de novo* assembly with *RepeatModeler* v2.0.2a (Flynn et al., 2020). This library was then used in *RepeatMasker* (Smit et al., 2013) to soft mask both repetitive elements and low-complexity regions within the assembly. The mRNA sequence data were remapped onto the genomic sequence using *HISAT2*, and structural annotation was performed with *BRAKER* (pipeline B, RNA-Seq guided; Hoff et al., 2016). Functional annotation was carried out using *InterProScan* (release 5.60-92.0), with Pfam, InterPro, and Ontology terms assigned to the transcripts. The final masked genomic sequence, complete with structural and

functional annotations, was deposited in GenBank under accession number JASPKY000000000 and published in BMC genomics (Cucini et al., 2024).

2.5 Genome completeness and statistics

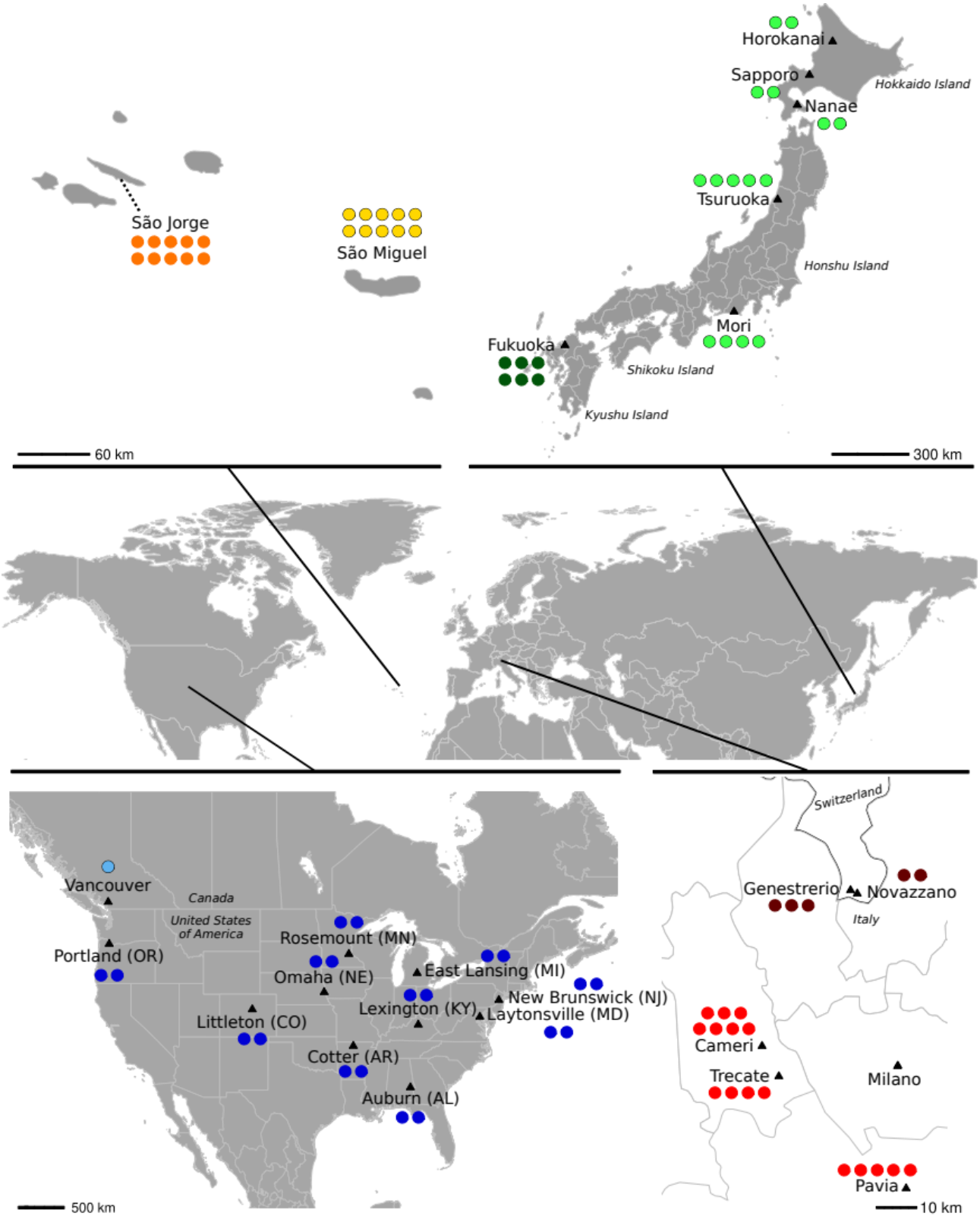
Genome statistics, such as N50, the number of contigs/scaffolds, and remapped read counts, were estimated and compared between the Italian and Canadian genome of *P. japonica* through *QUAST-LG* v5.0.2 (Mikheenko et al., 2018). Due to the absence of a true chromosome level reference, the Italian and the Canadian genomes were in turn used as reference into two parallel analyses. Assembly accuracy was assessed trough a k-mer spectrum analysis in *Merqury* v1.3 (Rhie et al., 2020) using Illumina reads (DMR4-6). To evaluate the quality of the annotation, trimmed RNA-seq reads were mapped back to the annotation with *HISAT2* and relative metrics were extracted and visualized from the BAM file through *samtools* v1.13 (commands: view, stat; Li et al., 2009). Moreover, a *BUSCO* analysis was performed to evaluate the completeness at the annotation level (v5.4.4; Endopterygota lineage; Waterhouse et al., 2018). Results were visualized with *ggplot* in the R environment. Annotation statistics, like the number of genes/transcripts, and the number of introns and exons, were produced using in-house python scripts employing the packages *GFFExaminer* (available at: https://biopython.org/wiki/GFF_Parsing) and *pandas* (McKinney, 2010) in R v4.3.1 (R Core Team, 2023). Results were visualized with *ggplot2* v3.3.5 (Wickham, 2016).

In the newly assembled genome, completeness, fragmentation, and duplication were assessed using *BUSCO* (Endopterygota lineage). This analysis included comparisons with all Coleoptera RefSeq genomes available in March 2023 in NCBI, as well as with *Apis mellifera* (NCBI accession: GCF_003254395.2) and the *Popillia japonica* genome produced by the CanSeq150 project (NCBI accession: GCA_004785975.1). Single-copy orthologs of the Endopterygota collection (n=2,124)

were extracted from the aforementioned genomes using *BUSCO*, individually aligned with *mafft* (parameter: --auto; Katoh et al., 2019) and quality-checked with *AMAS* (Alignment manipulation and summary statistics; Borowiec, 2016). High quality alignments (i.e., <10% of missing info) were retained and combined in a single matrix. A phylogenomic tree was constructed using *IQtree* v2.0.3 (parameters: -B 2000 -m MFP; Minh et al., 2020). The final tree was visualized and edited using the *ggtree* library in R (Yu et al., 2017).

2.6 Sample collection and whole genome resequencing for phylogeography analyses

Specimens of *P. japonica* were collected from its entire range to ensure a balanced representation of the four major regions where the species is found: Japan, North America, the Azores, and Italy+Ticino. Within each major region, samples were collected from multiple locations to maximize geographic coverage (Figure 4). To obtain high-quality DNA while minimizing potential gut and integument contaminants, total DNA was extracted from dissected male testes of 81 individuals using Wizard® Genomic DNA Purification kits (Promega). Sequencing was performed at Macrogen Europe (The Netherlands) using TruSeq DNA PCR-free libraries (Illumina) with a paired-end (PE) 150 bp strategy, targeting 20 GB of sequence data per individual (n=81). Sequences from three individuals (DMR5, DMR4, and DMR6) previously obtained for genome sequencing, as well as data from the Canadian genome draft (NCBI accession SRR8479473, unpublished), were included. Raw data from these latter libraries were subsampled based on initial remapping to 11.5%, 10.5%, 12.0%, and 21.0%, respectively, to achieve coverage comparable to other samples. After removing two individuals with sub-optimal sequencing data, a total of 83 individuals were successfully processed. This included samples from the ancestral area of the species (Japan, including six locations across three islands, n=21) and invasive populations in North America (11 locations in the USA and Canada, n=21), the Azores (two locations on two islands, n=20), and Italy and Ticino (five locations, n=21).



530 **Figure 4.** Map of *Popillia japonica* collection sites. Map showing labelled locations where samples of *Popillia*
531 *japonica* were collected for genomic analysis. Coloured dots indicate the number of specimens collected at
532 each location.

2.7 Variant calling and filtering

Quality control of raw sequence data was conducted using *FastQC*, followed by trimming with *fastp* v0.23.2 (head and tail trimming if Q <20, sliding window trimming if Q <24 over 4 bp; Chen et al., 2018). Trimmed sequences were aligned to the reference genome of *Popillia japonica* using *BBMap* v35 with a maximum allowed distance between paired-end reads of 500 bp, a maximum length of insertions or deletions of 200 bp, and other settings at default (Bushnell, 2014). Duplicated reads were removed using *Picard* v2.2.4 (available at: http://broadinstitute.github.io/picard). Non-matching read pairs were filtered out using *Samtools* v1.11 (Danecek et al., 2021).

Variant calling was performed using *BCFtools* v1.13 with the multiallelic caller, targeting SNPs and insertions/deletions (indels), and setting for diploidy (Danecek et al., 2021). After raw variant filtering, SNPs more than 3 bp away from an indel were retained. The filtering criteria included excluding SNPs located in repeat regions, contigs shorter than 1 kb, and contigs with median coverage greater than 3 standard deviations from the mean across all contigs.

Further analyses were conducted at both the SNP and individual levels using *VCFtools* v0.1.16 (Danecek et al., 2011). SNPs were retained if they were biallelic, had a site quality greater than 50, a mean depth between 15 and 52 across individuals, less than 5% missing data per site, and were more than 5 bp away from adjacent SNPs. Individual-level analyses included calculating the mean depth across sites, the number of missing sites per individual, and heterozygosity, with no identified outliers. A second dataset, referred to as snps_3p, was created by applying a minimum allele frequency (MAF) of 0.015 and requiring variants to be present in at least three chromosomes or two individuals. A third dataset, snps_3p_unlinked, was derived from the snps_3p dataset by pruning for linkage using *PLINK* v1.90b6.21 (parameters: window size of 50 kb, window step of 10 kb, r^2 threshold of 0.1; Purcell et al., 2007) to retain a subset of unlinked markers present in at least three chromosomes.

2.8 Genome-wide diversity

The population module of *STACKS* v2.64 (Rochette et al., 2019) was used to calculate average observed and expected heterozygosity (H_o and H_e), and average inbreeding coefficient per population (FIS; values ranging from -1 to 1; Wright, 1949). *VCFtools* was used for sliding windows analysis within 5 kb non-overlapping genomic windows, to evaluate overall nucleotide diversity (π) per population, and Tajima's D (Tajima, 1989). The significance of π was assessed using the Kruskal-Wallis test to evaluate differences among multiple groups and for pairwise comparisons, the Wilcoxon rank-sum test was employed with Bonferroni correction. In addition, pairwise nucleotide diversity among individuals from each area was also calculated using *VCFtools* as the average of site-π over sites. The possibility of differences in overall nucleotide diversity (π) among groups was tested using the Wilcoxon test with Bonferroni correction for multiple comparisons. Linkage disequilibrium (LD) decay patterns were evaluated using *PopLDdecay* v3.42 (Zhang et al., 2019), by calculating the R^2 between SNPs in different population subsets for six randomly selected individuals per population (South Japan, North/Central Japan, USA+Canada, São Jorge, São Miguel, Italy+Ticino; see Results), with a maximum distance between SNPs of 100 kb. All genomic diversity analyses were conducted using the snps_3p dataset.

2.9 Population structure and differentiation

To evaluate the distribution of genetic variation, a Principal Component Analysis (PCA) was performed using *plink* on the snps_3p_unlinked dataset and plotted the first two principal components using the package *ggplot2* in R.

To infer individual admixture proportions and structure within and among populations, the sparse non-negative matrix factorization (sNMF) algorithm implemented in the R package *LEA* v3.2.0 (Frichot et al., 2014; Frichot & François, 2015) was applied on the snps_3p_unlinked dataset to analyse an assumed number of ancestral populations (K), ranging from 1 to 10, with 10 repetitions per tested K value. The cross-entropy criterion was used to evaluate the optimal number of ancestral populations, with a smaller value suggesting a more optimal K value (Figure 5). The optimal number of clusters was identified at K=5, however, a range of K values was explored to examine possible hierarchical population structure (Lawson et al., 2018).

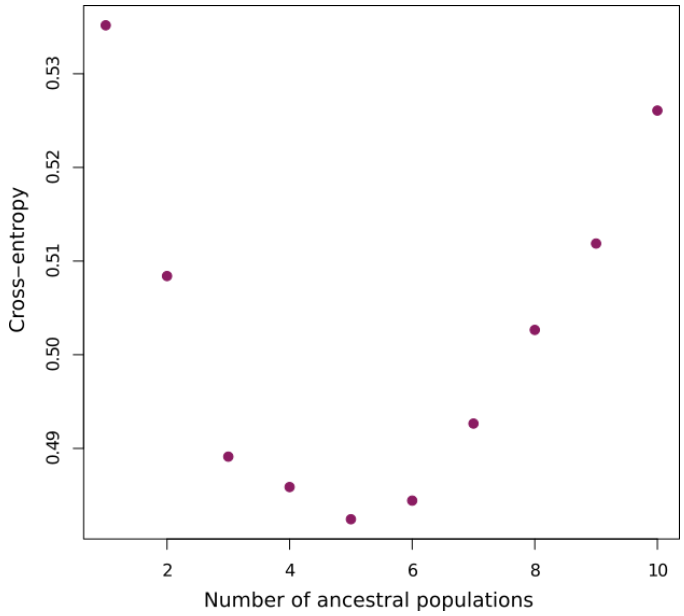


Figure 5. Cross-entropy results of sNMF runs for *Popillia japonica*. The run with the lowest cross-entropy value (K=5) represents the optimal population clustering scenario.

A maximum likelihood phylogeny was inferred with 2,000 bootstrap replicates using *IQtree* v2.0.3 and the snps_3p_unlinked dataset. The ModelFinder option in *IQtree* was used to identify the best-fit model of nucleotide substitution according to a Bayesian information criterion, which was identified to be K3Pu+F+R4.

Population pairwise fixation indices (FST) [Weir & Cockerham, 1984] were calculated for the snps_3p dataset, using the R package *StAMPP* v1.6.3 (Pembleton et al., 2013) with 100 bootstraps.

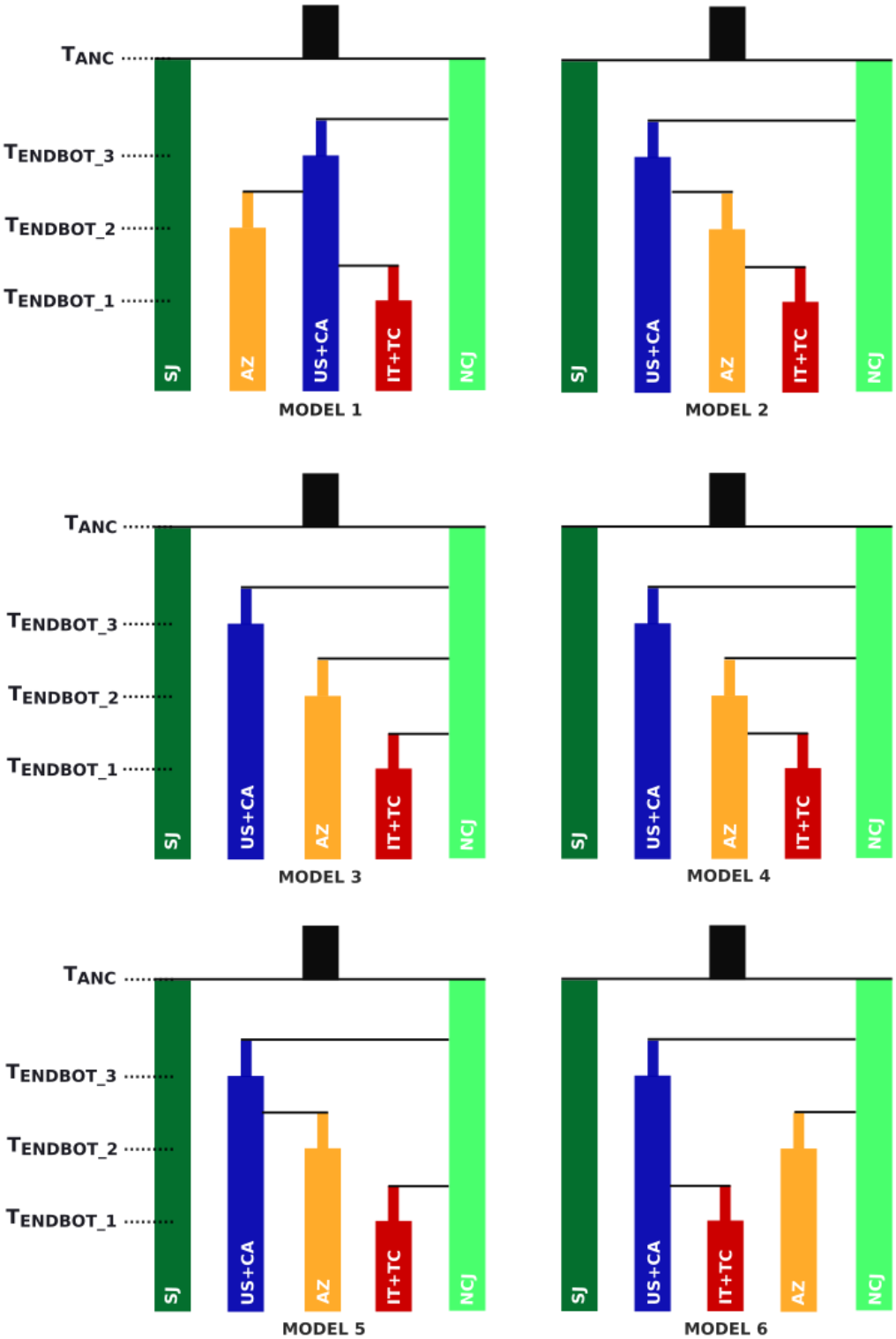
598 Covariance structure among population allele frequencies, resulting from shared demographic
599 histories among populations (Olazcuaga et al., 2020), was explored by estimating the scaled
600 covariance matrix of the population allele frequencies (Ω) using the *BayPass* v2.3 core model
601 (Gautier, 2015) with the snps_3p_unlinked dataset.

602

603 **2.10 Demographic inference**

604 To understand the colonization history of *P. japonica* invasive populations, demographic modelling
605 was performed based on the site frequency spectrum (SFS) using *Fastsimcoal2* (Excoffier et al., 2013,
606 2021). The snps_3p_unlinked dataset was converted to an SFS using *easySFS* (available at:
607 https://github.com/isaacovercast/easySFS), projecting the number of samples down in each
608 population to maximize the number of retained SNPs while minimizing missing data. Invasion
609 scenarios (Figure 6) assumed five main populations based on the population structure patterns (South
610 Japan, North/Central Japan, USA+Canada, Azores, Italy+Ticino; see Results). The time of ancestral
611 divergence between South Japan and North/Central Japan was set to 5,000-25,000 years ago, based
612 on the 95% higher posterior density (95% HPD) of molecular dating analysis estimates in Nardi et al.
613 (2024). Following the population structure patterns (see Results) and previous studies (Strangi et al.,
614 2023; Nardi et al., 2024), it was assumed that the invasive lineage originated from North/Central
615 Japan, excluding the possibility of South Japan as a source of invasion. Invasion dates were fixed at
616 111, 57, and 13 generations ago for USA+Canada (Fleming, 1972), the Azores (Simões, 1984) and
617 Italy+Ticino (Pavesi, 2014), respectively, by subtracting six years from official first reports to align
618 with the from-first-detection-to-invasion time gap suggested by Frank (2016) in the USA and
619 assuming one generation per year (Potter and Held, 2002). All demographic scenarios were
620 characterized by a bottleneck event at the time of invasion for invasive lineages. Bottleneck duration
621 was set to 1-6 generations and the number of individuals during bottlenecks was set to 10-500. An
622 absence of gene flow was hypothesized for all scenarios. Since the mutation rate for this species was

not available, we used *Fastsimcoal2* for modelling invasion scenarios without parameter estimation. Each model was simulated with 50 independent runs, 40 conditional maximization algorithm cycles, and 500,000 simulations for likelihood maximization. To determine the model with the highest support, best runs from all six models were compared using the Akaike information criterion (AIC).



636 **Figure 6.** *Fastsimcoal2* models testing invasion pathways of *Popillia japonica* in invaded areas, assuming
637 North/Central Japan as the origin of the invasive lineage. All demographic scenarios were characterized by a
638 bottleneck event at the time of invasion for invasive lineages and an absence of gene flow. Abbreviations:
639 T_{ANC}, divergence time of non-uplift lineages; T_{ENDBOT}, bottleneck end time; SJ, South Japan, NCJ:
640 North/Central Japan; US+CA, USA+Canada; AZ, Azores; IT+TC, Italy+Ticino.

2.11 Genome-wide scans for detection of outlier SNPs and SNPs annotation

Using the snps_3p_unlinked dataset, two approaches were applied to study putative genomic signatures of selection: outlier detection with the R package *PCAdapt* v4.3.3 (Privé et al., 2020) and an FST outlier calculation approach, contrasting North/Central Japan with all invasive population. After running *PCAdapt* using Mahalanobis distances and default parameters, optimal principal components (K) were selected through a scree plot test following Cattell's rule (Cattell, 1966): K=5 for North/Central Japan vs all invasive populations. The False Discovery Rate (FDR) with a 0.1% threshold was applied after testing different methods (including q-value, Benjamini-Hochberg procedure, and Bonferroni procedure), after which we applied the more stringent Bonferroni procedure to identify a final set of reliable candidate SNPs. FST outlier loci were identified through sliding windows analysis using *VCFtools*, within 5 kb non-overlapping genomic windows, selecting the top 0.1% of weighted FST values. A Manhattan plot of the genome-wide weighted FST values was plotted using R *qqman* package v0.1.9 (Turner, 2018).

Only those SNPs detected by both methods were considered as outliers and these were plotted in a Venn diagram using the R *ggvenn* package v0.1.10 (Yan, 2023). *SnpEff* v5.1 (Cingolani et al., 2012a) was used to annotate outlier SNPs and investigate potential functional downstream effects on genes and proteins. A *SnpEff* database was manually built from the *P. japonica* GFF file and the aforementioned reference genome (Cucini et al., 2024), which provided functional annotations from Pfam and Interpro. *SnpEff* was then run following the software's documentation. For a better understanding of obtained results, *SnpSift* v5.1 (Cingolani et al., 2012b) was applied to filter annotated variants and find putatively relevant ones, and additional information was searched through InsectBase.

3. Results

3.1 Genome sequencing, assembly and annotation

DNA Illumina sequencing generated between 333.2 and 383.5 million read pairs across three libraries. ONT long-read sequencing produced a total of 1.8 million reads. Given an expected genome size of approximately 0.6 Gb, this corresponds to a theoretical coverage of about 300x for Illumina reads and 20x for ONT long reads. The mRNA sequencing generated between 41.0 and 52.0 million read pairs across eight libraries, while total RNA sequencing yielded between 20.6 and 21.2 million read pairs across eight libraries. The final polished assembly was 578.347 Mb in length, fragmented into 6,164 scaffolds with an N50 of 0.895 Mb. The k-mer spectrum plot showed a monomodal distribution with a single major peak, indicating that only one haplotig was retained (Figure 7). BUSCO analysis showed 94.3% of complete BUSCOs, 1.1% of duplicated BUSCOs, 2.6% of fragmented BUSCOs, and 2.0% of missing BUSCOs (Figure 8).

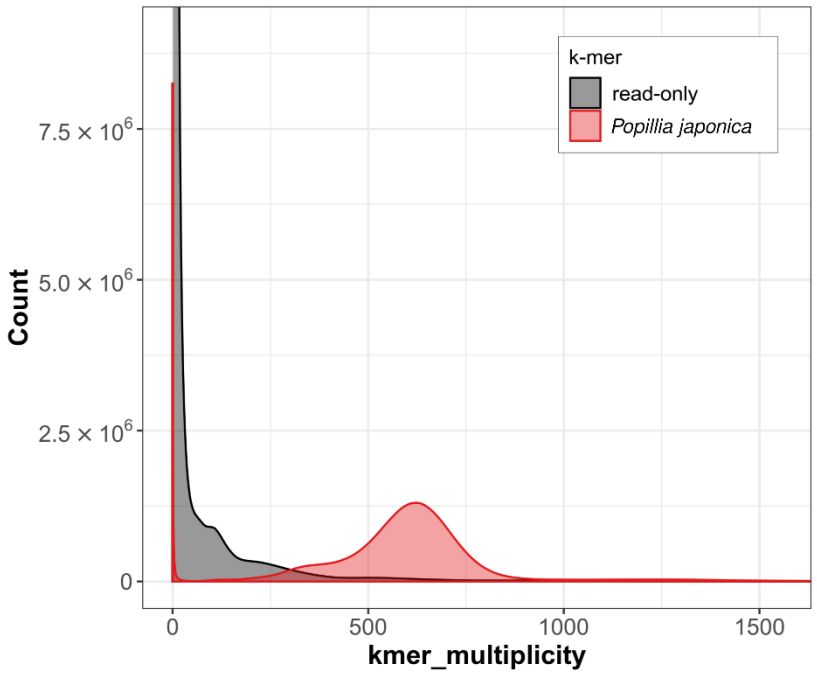


Figure 7. K-mer spectrum plot produced with Merqury on the final genome assembly. The curve displays a monomodal distribution with a prominent peak, indicating the presence of a distinct single haplotig.

The new *P. japonica* assembly, when compared to the previously available genome from the CanSeq150 program (NCBI GCA_004785975), demonstrates several improvements: it is slightly longer (578 Mb vs. 531 Mb), shows greater completeness (94.3% vs. 78.1%), has fewer duplications (1.1% vs. 2.4%), is markedly less fragmented (2.6% vs. 12.5%), and is significantly more contiguous (6,164 scaffolds vs. 75,144; N50 = 0.895 Mb vs. 0.015 Mb). Additionally, BUSCO statistics comparison with other available RefSeq genomes of Coleoptera (Figure 8) indicates that the new genome has one of the lowest duplication rates and matches well with others in terms of completeness and fragmentation. The phylogenetic analysis, constructed from a supermatrix encompassing 1,572,906 amino acid positions across 19 coleopteran taxa, with *Apis mellifera* as the outgroup, yielded the phylogenetic tree depicted in Figure 8. The inferred phylogenetic relationships are consistent with the current taxonomic classification of the group. The newly sequenced genome, together with the previously sequenced *P. japonica* genome, forms a sister clade to *Onthophagus taurus* within Scarabeiformia (Scarabaeidae), as illustrated in Figure 8.

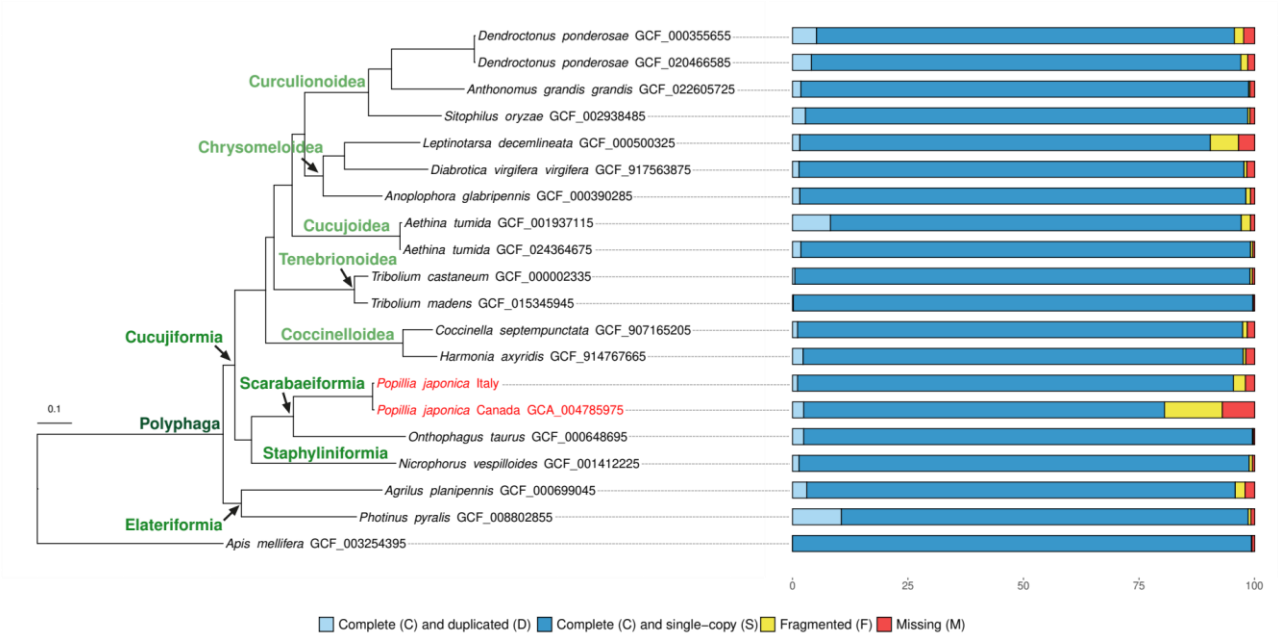
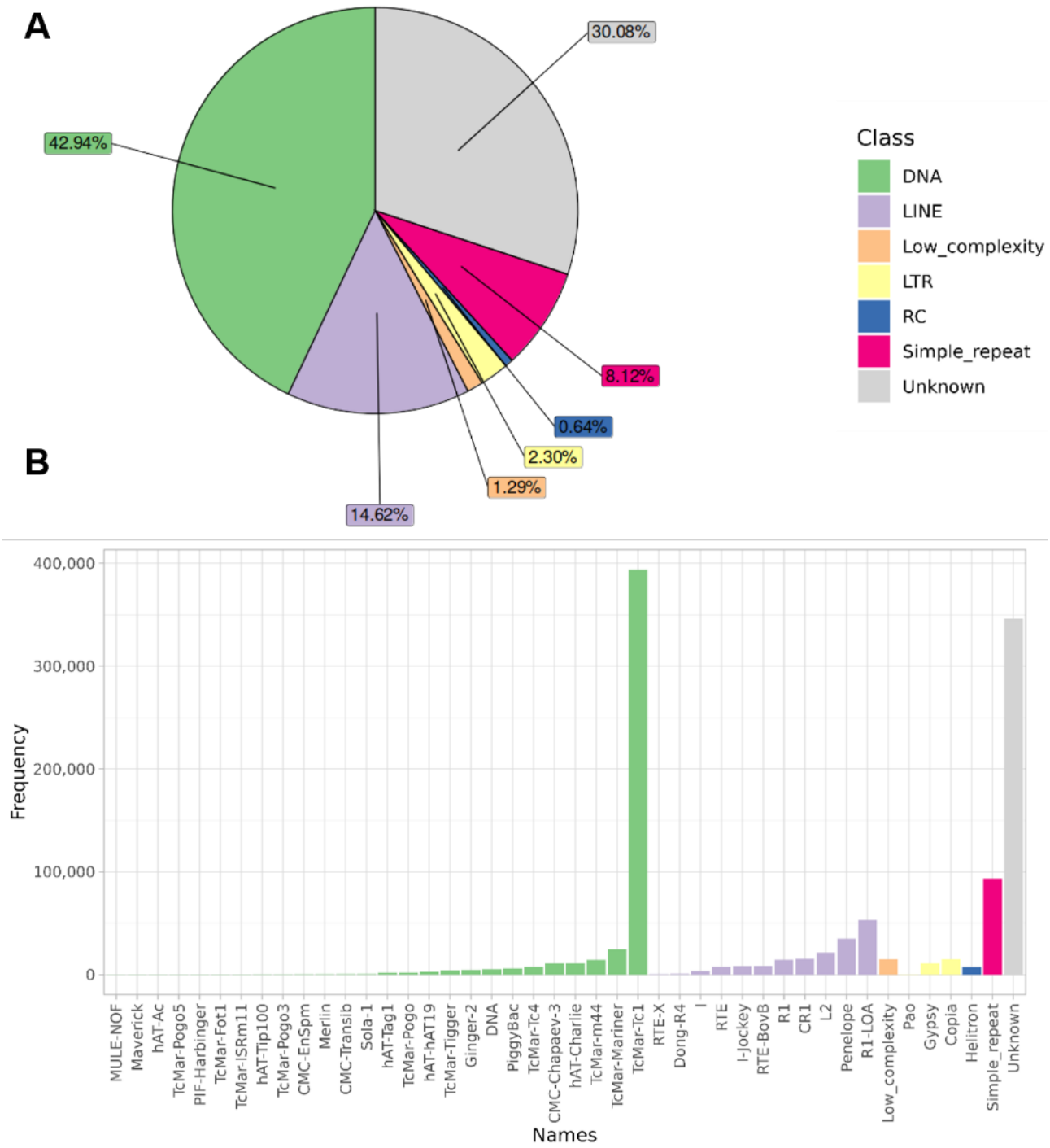


Figure 8. Phylogenetic relationships among Coleoptera based on high-quality aligned single-copy orthologs derived from the BUSCO analysis. All the nodes have a bootstrap value of 1. Each species is coupled with the corresponding BUSCO graphic. The two available genomes of *P. japonica* (from the CanSeq150 project and this study) are highlighted in red. Higher ranking categories are indicated at relevant nodes with green shades.

A substantial portion of the genome (36.7%) consists of interspersed repeats and low complexity regions. Among these, DNA elements are the most prevalent (42.9%), followed by Long Interspersed Nuclear Elements (LINEs) (14.6%) and simple repeats (8.1%). In contrast, Long Terminal Repeats (LTRs), low complexity regions, and Rolling Circles (RCs) are present in smaller proportions (Figure 9). The most frequent repeats are TcMariner transposons, with TcMar-Tc1 comprising approximately 34% of all annotated Transposable Elements (TEs; Figure 9b). Despite the presence of numerous outliers characterized by large dimensions (up to 6 kb in DNA elements), the majority of TEs are short elements with an average length of ~200 bp (Figure 10), which may indicate vestigial remnants of TEs that have lost their distinctive domains.



707 **Figure 9.** Relative abundances and count of TEs in the *P. japonica*'s genome, based on RepeatMasker
708 annotation. (A) Relative abundances of TEs. (B) Counts of Transposable Elements divided per type. TE classes
709 are coloured according to the provided key.

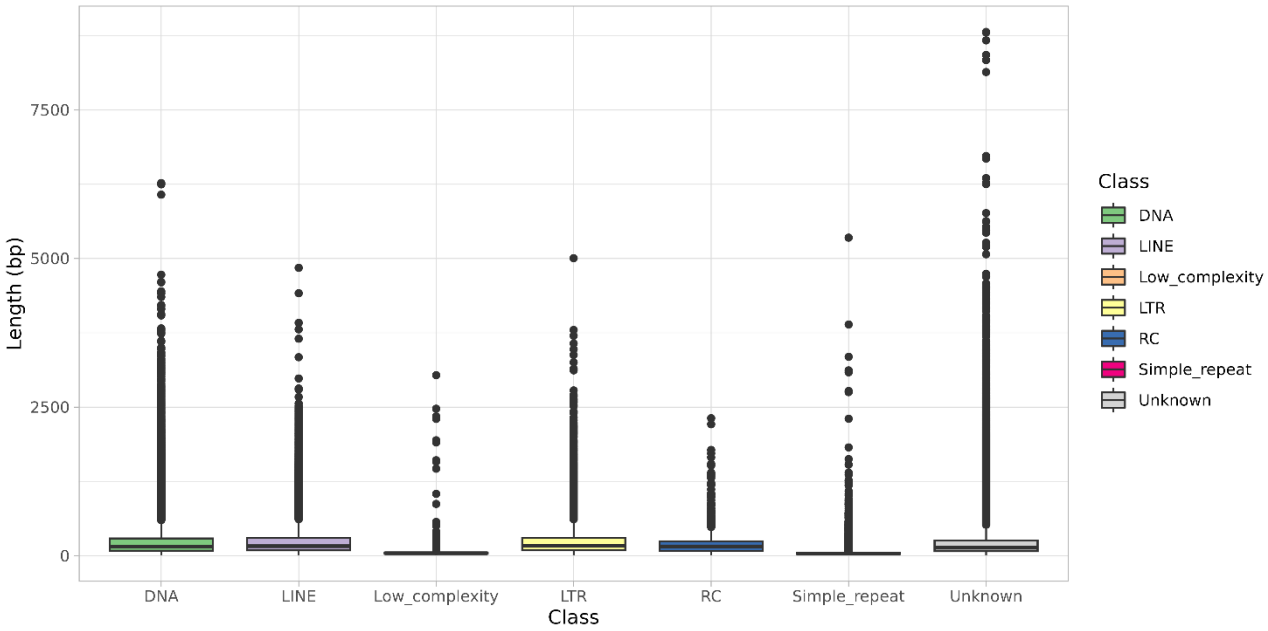
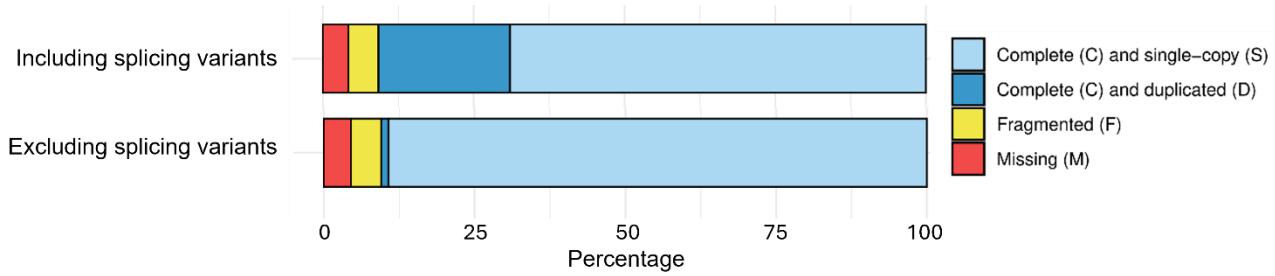


Figure 10. Length of TEs in *Popillia japonica*'s genome divided per class. Lengths were extracted from the RepeatMasker GFF annotation file through an in-house python script. Average length among TEs is approximately 200 bp.

The automatic structural annotation identified 31,668 genes, which are transcribed into 36,495 different transcripts (splice variants). Notably, the majority of these genes (89.1%) produce a single transcript. The BUSCO Endopterygota assessment of single-gene annotations highlighted a high level of completeness, with only a small fraction showing duplication (Figure 11). As regards the exon distribution per transcript, most genes contain a single exon (27.9%) or a few exons (five or fewer; 74.2%). Functional annotation analysis yielded viable annotations for 18,725 transcripts (51.3%) in at least one database. Specifically, all annotated transcripts received a Pfam annotation, covering 51.3% of total transcripts, while 12,228 transcripts (33.5%) were annotated in InterPro, and 10,668 (29.2%) were assigned Gene Ontology terms. Overall, 14,958 genes have at least one transcript with functional annotations in one or more databases, underscoring the comprehensive nature of the functional characterization achieved in this study.



724 **Figure 11.** BUSCO analysis through the Endopterygota lineage on the *Popillia japonica*'s genome at the
725 annotation level. The annotation completeness was evaluated on the full set of annotations (i.e., including
726 splice variants) and on single-genes only.

727

728 **3.2 Datasets from whole-genome resequencing**

729 From the 83 sequenced *P. japonica* individuals, we obtained 3,666,428 and 295,396 SNPs in the
730 snps_3p (hard-filtered with linked SNPs) and snps_3p_unlinked (hard-filtered with unlinked SNPs)
731 datasets, respectively.

732

733 **3.3 Genomic diversity**

734 *Popillia japonica* populations showed low levels of nucleotide diversity (π) across native and invasive
735 populations in sliding window analysis, with π ranging from 0.00083 for the Azorean population from
736 São Miguel to 0.00144 for North/Central Japan. On average, ancestral populations displayed higher
737 diversity (mean π=0,00134) than invasive ones (mean π=0,00105) (Figure 12, top panel). The same
738 pattern was confirmed for all populations through pairwise nucleotide diversity among individuals
739 (Figure 13). Similarly, native populations showed higher observed heterozygosity (mean Ho = 0.1912
740 and 0.1548 in native and invasive populations, respectively) and a slight heterozygote loss was
741 observed in North/Central Japan and the USA (Table 1). The inbreeding coefficient was negative for
742 the majority of populations, with the exception of North/Central Japan and the USA, which exhibited
743 slightly positive estimates (Table 1). Across populations, Tajima's D showed positive median

genome-wide estimates, ranging from 0.44630 to 1.33399 (Figure 12, bottom panel). Linkage disequilibrium (LD) decay across 100 kb genomic windows was slowest for Azorean populations from São Jorge and São Miguel (Figure 16). Italy displayed an intermediate value, while LD decay was more rapid for both Japanese populations and for USA+Canada (Figure 14).

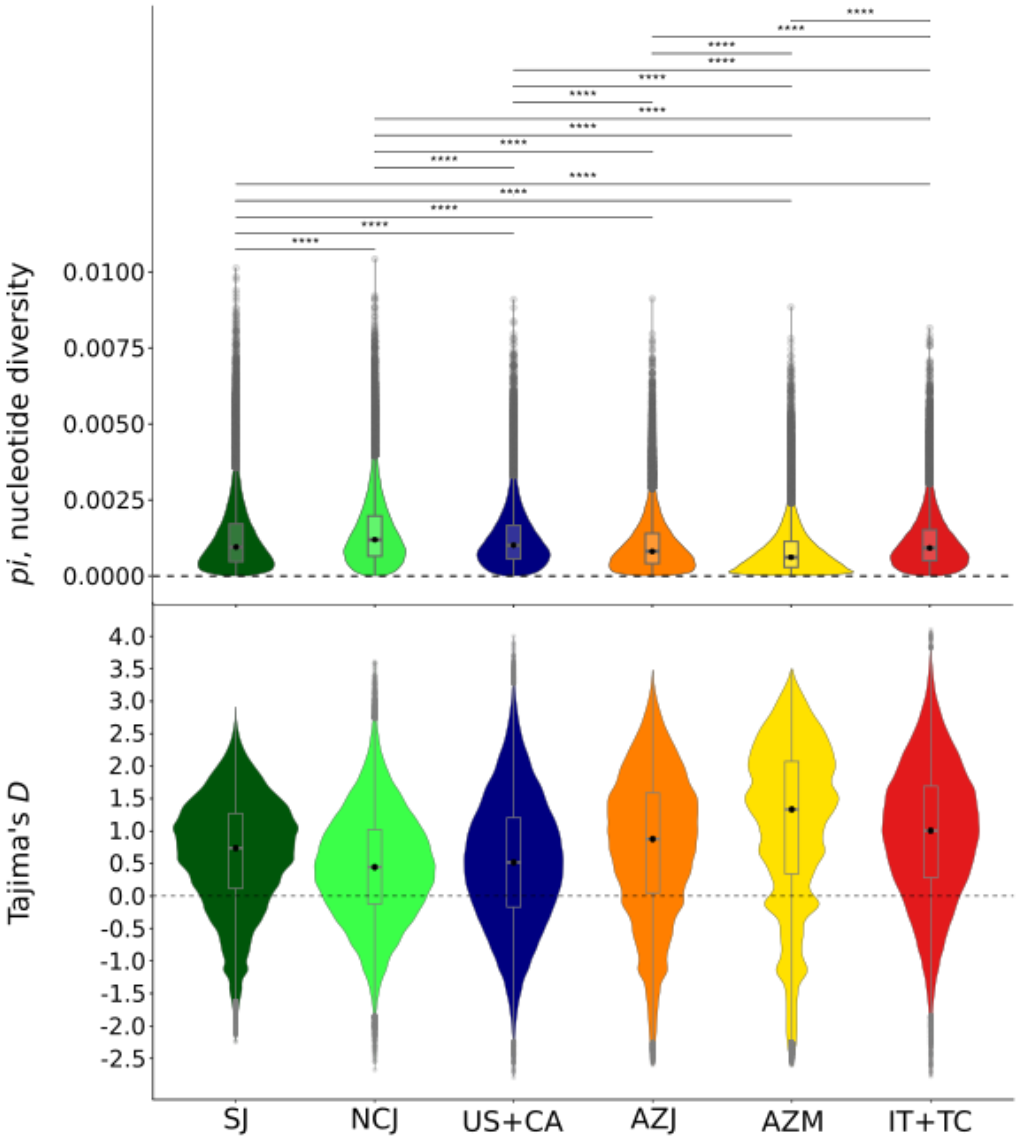


Figure 12. Per locus distribution of nucleotide diversity (top panel) and Tajima's D (bottom panel) for each population of *Popillia japonica*, summarized across 5 kb genomic windows. Significant differences were identified using the Kruskal-Wallis test and Wilcoxon rank-sum test with Bonferroni correction (p-adj<0.05). Asterisks indicate statistical significance. SJ, South Japan; NCJ, North/Central Japan; US+CA, USA+Canada; AZJ, São Jorge (Azores); AZM, São Miguel (Azores); IT+TC, Italy+Ticino.

Populations	Ho	S.E.	He	S.E.	F_{IS}	S.E.
South Japan	0.1943	0.0001	0.1722	0.0001	-0.0112	0.0003
North/Central Japan	0.1881	0.0001	0.2078	0.0001	0.0757	0.0003
USA+Canada	0.1785	0.0001	0.179	0.0001	0.0181	0.0003
São Jorge (Azores)	0.1534	0.0001	0.1408	0.0001	-0.0116	0.0002
São Miguel (Azores)	0.1126	0.0001	0.1011	0.0001	-0.013	0.0002
Italy+Ticino	0.1746	0.0001	0.1665	0.0001	-0.0091	0.0004

Table 1. Estimates of population genomic indices in native and invasive populations of *Popillia japonica*, using 3,666,428 SNPs (snps_3p dataset) from 83 individuals. Native and invasive populations are coloured in green and yellow, respectively. H_o: average observed heterozygosity; H_e: average expected heterozygosity; F_{IS}: inbreeding coefficient; S.E.: standard error. Results obtained using STACKS software.

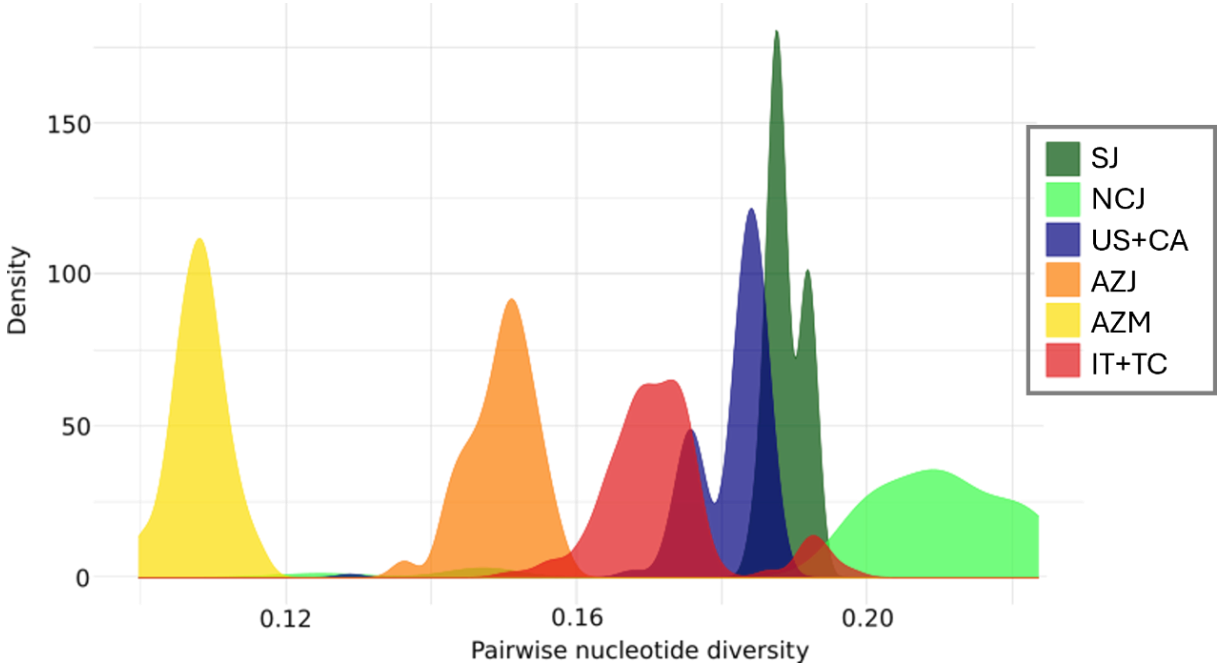
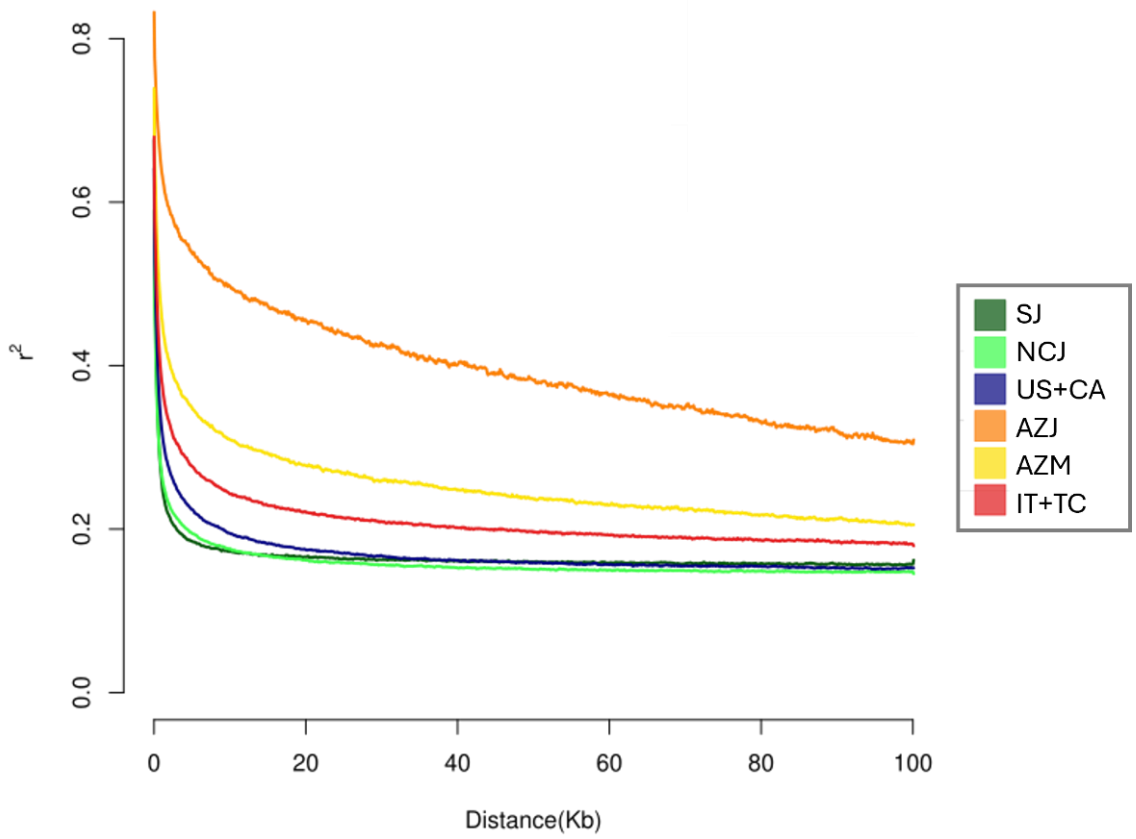


Figure 13. Density plot of pairwise nucleotide diversity among individuals, showing the distribution of average nucleotide diversity of values within each *Popillia japonica* population. Significant differences in overall nucleotide diversity among groups was tested using the Wilcoxon test with Bonferroni correction for multiple comparisons (p-adj<0.05). Populations are colour coded according to the provided key. SJ, South Japan; NCJ, North/Central Japan; US+CA, USA+Canada; AZJ, São Jorge (Azores); AZM, São Miguel (Azores); IT, Italy; TC, Italy+Ticino.



768 **Figure 14.** LD decay patterns in *Popillia japonica* populations, obtained by plotting LD decay estimates (r2)
769 within 100 kb genomic window. SJ, South Japan; NCJ, North/Central Japan; US+CA, USA+Canada; AZJ, São
770 Jorge (Azores); AZM, São Miguel (the Azores); IT+TC, Italy+Ticino; LD, linkage disequilibrium.

771

772 **3.4 Population structure and differentiation**

773 The PCA identified six clearly differentiated genetic groups: South Japan, North/Central Japan,
774 USA+Canada, São Jorge (the Azores), São Miguel (the Azores) and Italy+Ticino, with the first PC,
775 accounting for 18.3% of the total variation, showing differentiation between invasive and native
776 clusters (Figure 15). Some additional structure was observed within North/Central Japan, with
777 locations Horokanai and Nanae (Hokkaido Island) clustering together, whereas Sapporo (Hokkaido)
778 clustered with Tsuruoka (Honshu Island) (Figure 15). Mori (Honshu) was recovered as an
779 independent cluster. No additional structure was visible within Italy+Ticino and USA+Canada
780 (Figure 15).

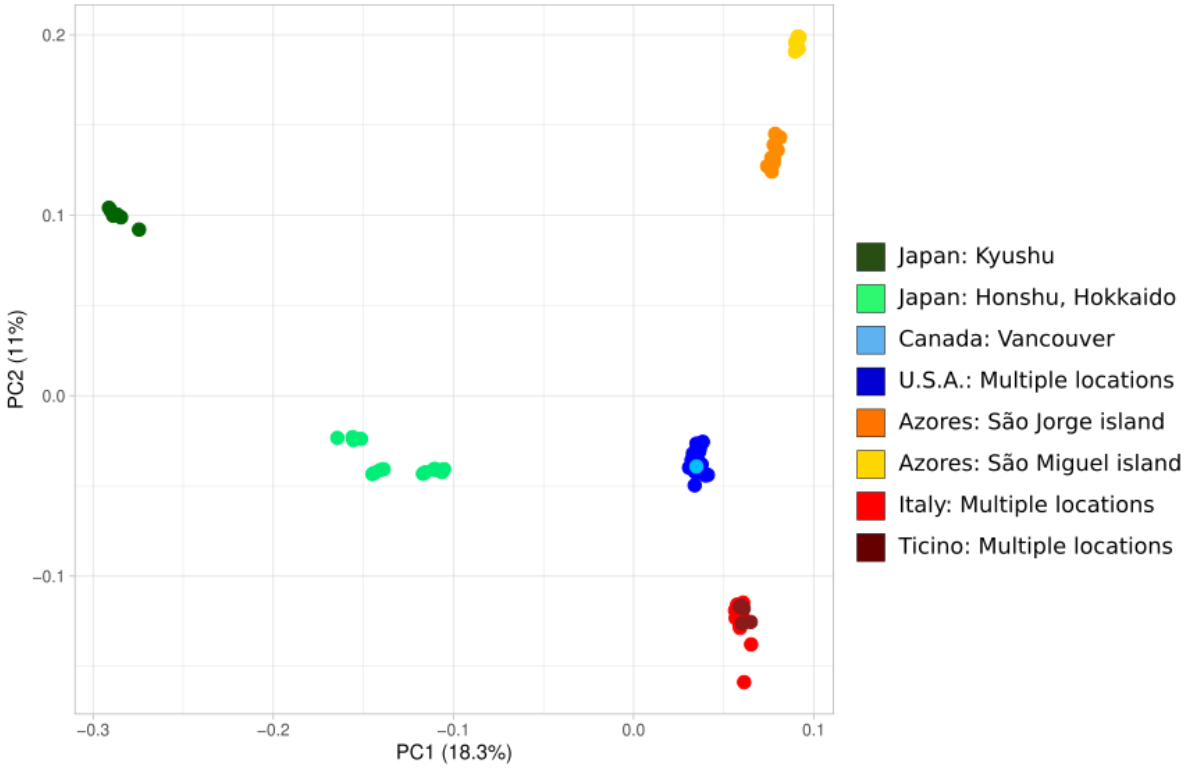


Figure 15. Principal component analysis showing variation in allele frequencies across populations of *Popillia japonica*, based on 295,396 SNPs in 83 individuals. Populations are colour coded according to the key.

Admixture analysis (sNMF) plots identified the same overall subdivision between ancestral (Japan) and invasive (all others) populations as the PCA. At K=2, North/Central Japan and USA+Canada displayed some intermixing (Figure 16). At K=4-6, North/Central Japan shared some ancestry with South Japan and the USA (Figure 16). South Japan consistently stands out as well-differentiated from the other populations, indicating a strong genetic separation. Despite some weak signals for USA+Canada and Italy+Ticino at K=10, North/Central Japan appeared to be the only population exhibiting internal genetic structure (Figure 16). This observation aligns with the results of other analyses conducted in this study. In detail, as already pinpointed in the PCA analysis, at K=4-10, the locations Horokanai and Nanae (Hokkaido Island) exhibited similar admixture patterns. Sapporo (Hokkaido) shared similar patterns with Tsuruoka (Honshu Island), whereas Mori (Honshu) differentiates from the other Japanese populations, showcasing a distinct genetic profile (Figure 16). Considering the invasive cluster, Azorean populations, especially São Miguel, were the most

differentiated (a single discrete cluster with limited admixture), while Italy+Ticino, with Ticino never diverging from the Italian population, showed higher admixture with USA+Canada (Figure 16).

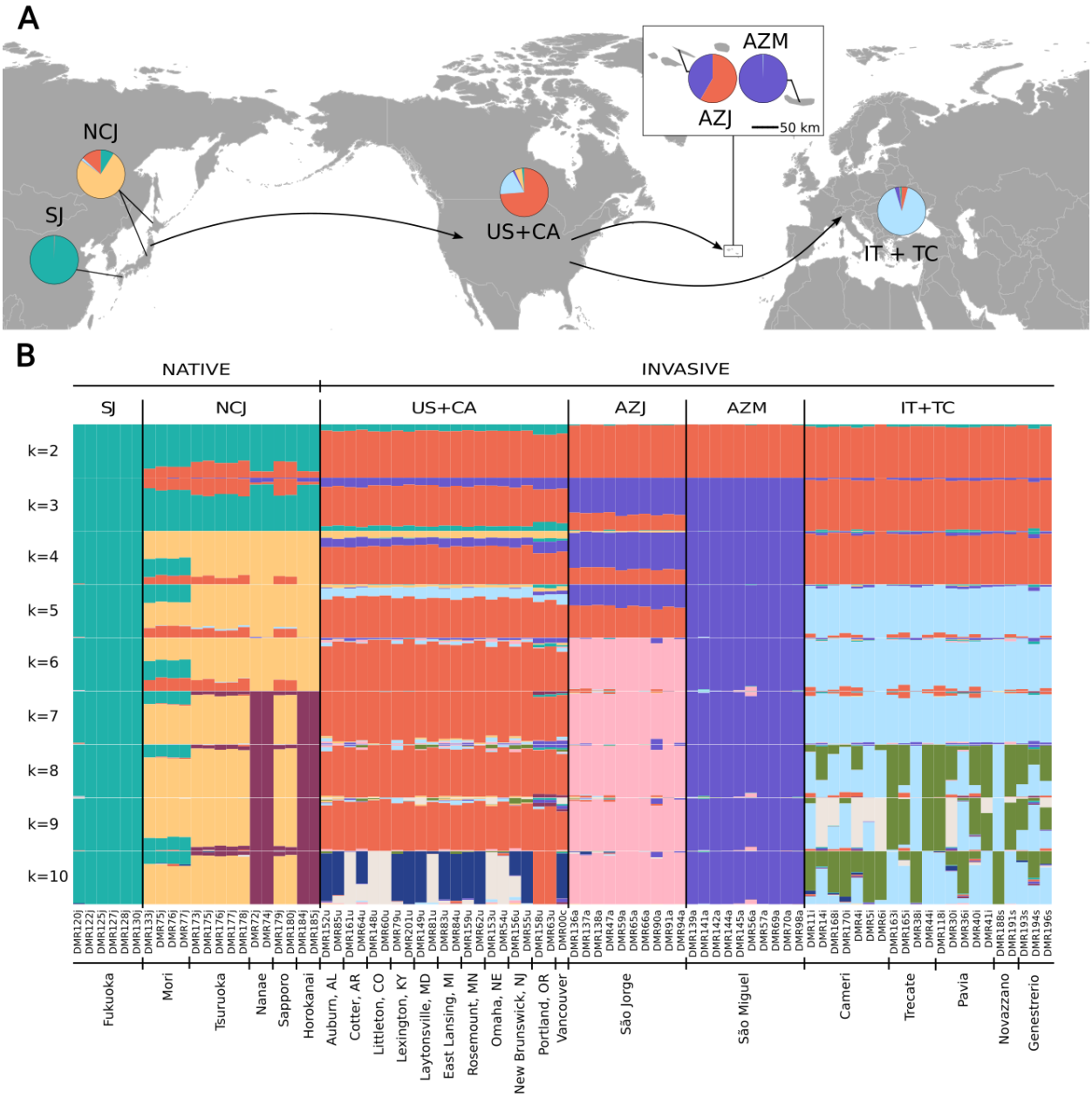
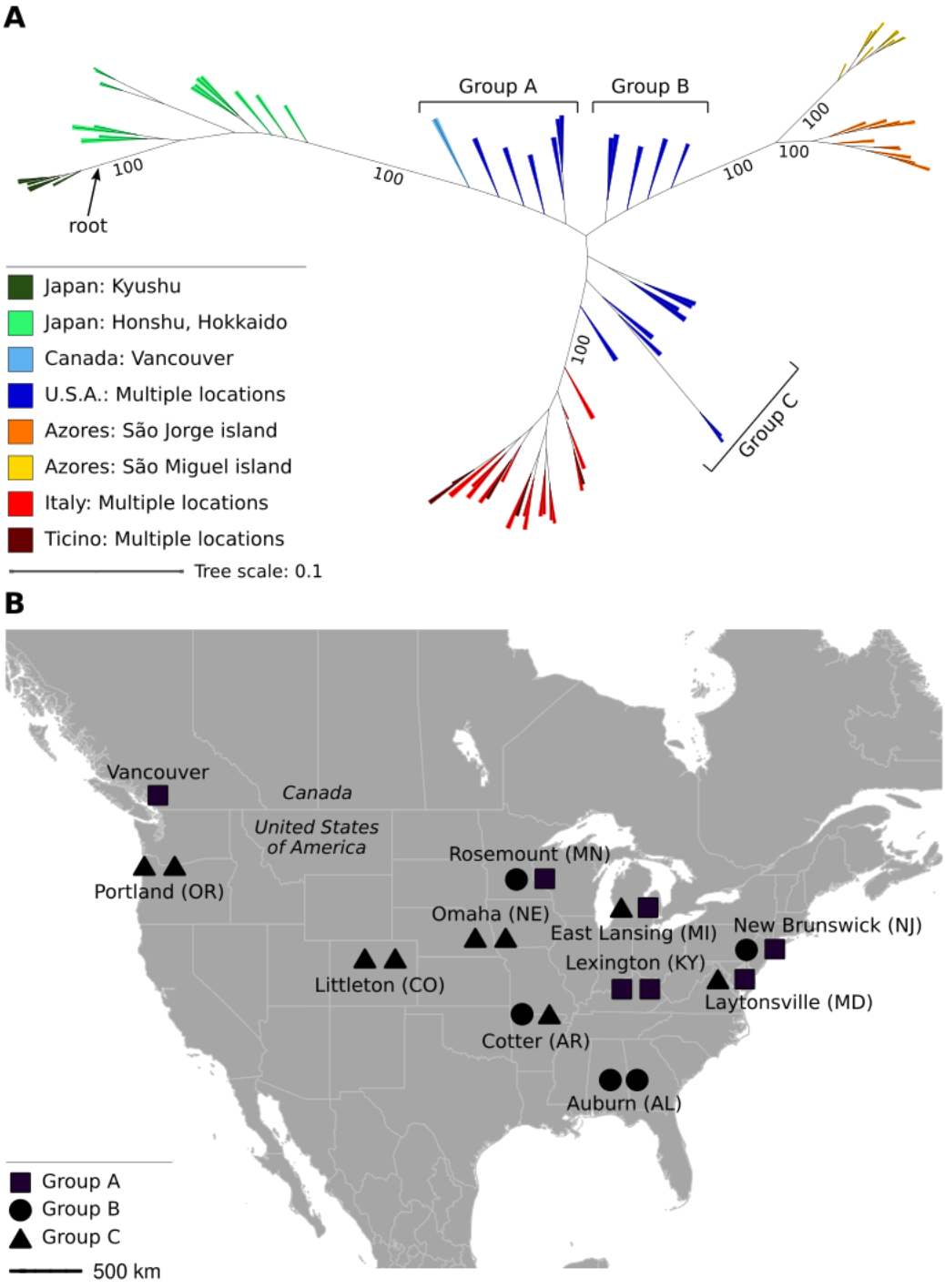


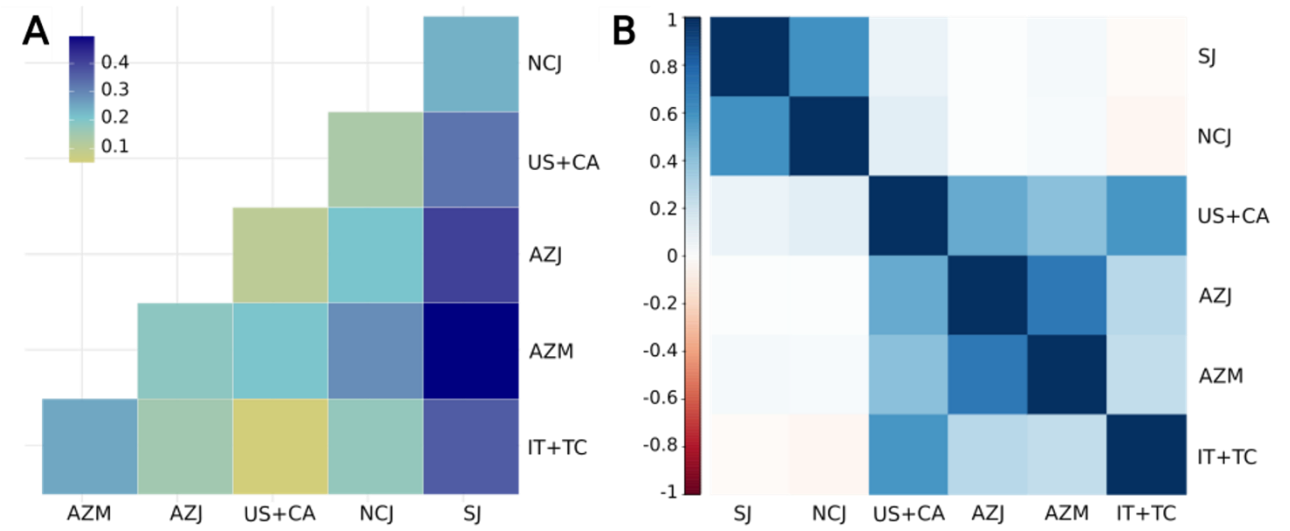
Figure 16. Genetic admixture of *Popillia japonica*, based on 295,396 SNPs in 83 individuals. (A) Map of the *P. japonica* distribution range. Pie charts of admixture proportions within all populations derived from the sNMF analysis at K=5. Black arrows indicate invasion pathways. (B) Admixture plots of hierarchical population structure across different K values using the sNMF estimates. Each individual is represented by a thin vertical bar partitioned into K-coloured segments representing estimated membership fractions of the individual in K clusters. SJ, South Japan; NCJ, North/Central Japan; US+CA, USA+Canada; AZJ, São Jorge (Azores); AZM, São Miguel (Azores); IT+TC, Italy+Ticino; sNMF: sparse non-negative matrix factorization.

804 The phylogenetic tree of specimens clearly identified the same six genetic groups outlined
805 above, with South Japan diverging at the base of the tree (Figure 17a). The branch leading to USA
806 samples emerged from a paraphyletic group in North/Central Japan, with the two closest nodes
807 representing individuals from Tsuruoka on Hokkaido Island (Figure 17a). From within the USA clade,
808 two clearly identifiable nodes branched off, one representing invasive individuals in the Azores
809 (further subdivided among the two islands of São Jorge and São Miguel), and one representing
810 invasive individuals in Italy+Ticino (Figure 17a). The distribution of samples from the USA was
811 subdivided in the tree, suggesting that the basal-most individuals were localized in North-Eastern
812 USA, individuals at the base of the Azorean node were localized in Central/Eastern USA, while
813 individuals rooting the Italy+Ticino node were mostly localized in Central/Western USA (Figure
814 17b).
815
816 The highest genomic differentiation was observed between South Japan and invasive
817 populations (FST ranging from 0.329 with the USA to 0.498 with São Miguel; Figure 18a, Table 2).
818 The USA displayed lower genetic distance values with both São Jorge and Italy+Ticino (FST=0.102
819 and FST=0.051, respectively), while between the latter a higher genomic distance was exhibited
820 (FST=0.152). Among invasive populations, São Miguel was the most differentiated (FST ranging
821 from 0.185 with the other Azorean Island São Jorge to 0.252 with Italy+Ticino; Figure 18a, Table 2).
822 A correlation matrix of population allele frequencies (Ω), summarizing shared population history,
823 confirmed the pairwise FST results. For example, South Japan was poorly correlated with invasive
824 populations (Ω ranging from 0.083 with USA+Canada to -0.022 with Italy+Ticino; Figure 18b; Table
825 3), and invasive populations revealed correlation patterns similar to those identified using FST
826 estimates. Indeed, USA+Canada exhibited high correlations with São Jorge and Italy+Ticino
827 (Ω=0.508 and Ω=0.585, respectively), while São Jorge and Italy+Ticino were less correlated

828 (Ω=0.270). However, the matrix also clustered populations according to their origin, highlighting a
829 strong division between native and invasive groups (Figure 18b; Table 3).



830 **Figure 17.** Phylogenetic relationships between native and invasive populations of *Popillia japonica*, based on
831 295,396 SNPs in 83 individuals. (A) Maximum likelihood unrooted tree of individuals. Populations are colour
832 coded throughout according to the key. (B) Map showing the distribution of different groups in the USA and
833 Canada [see (A)]. Groups are symbol-coded according to the key.



834 **Figure 18.** Heat maps of genomic distance between different pairs of *Popillia japonica* populations. (A) Heat
835 map of pairwise F_{ST} values among populations. Dark blue shades indicate higher genomic distance in
836 accordance to provided colour key. (B) Heat map of scaled covariance Ω matrix estimating shared population
837 history of *P. japonica,* comparing pairs of populations. Dark blue shades indicate higher correlations between
838 populations in accordance to provided colour key. SJ, South Japan; NCJ, North/Central Japan; US+CA,
839 USA+Canada; AZJ, São Jorge (The Azores); AZM, São Miguel (The Azores); IT+TC, Italy+Ticino.

840

841

842

843

844

845

846

847

848

Populations	South Japan	North/ Central Japan	Usa+ Canada	São Jorge (Azores)	São Miguel (Azores)
North/Central Japan	0.2374				
Usa+Canada	0.3291	0.1352			
São Jorge (Azores)	0.4083	0.2081	0.1016		
São Miguel (Azores)	0.4976	0.2909	0.2057	0.1849	
Italy+ticino	0.3657	0.1737	0.0506	0.1521	0.2524

849 **Table 2.** Pairwise population F_{ST} indicating genetic distance estimates between all populations of *Popillia*
850 *japonica*, using 3,666,428 SNPs from 83 individuals. Native and invasive populations are coloured in green
851 and yellow, respectively.

852

Populations	South Japan	North/ Central Japan	USA+ Canada	São Jorge (Azores)	São Miguel (Azores)	Italy+ Ticino
South Japan	1.0000					
North/Central Japan	0.6052	1.0000				
USA+Canada	0.0831	0.1238	1.0000			
São Jorge (Azores)	0.0190	0.0169	0.5078	1.0000		
São Miguel (Azores)	0.0469	0.0354	0.4106	0.7115	1.0000	
Italy+Ticino	-0.0223	-0.0426	0.5851	0.2700	0.2490	1.0000

853 **Table 3**. Scaled covariance Ω matrix, estimating shared population history of *Popillia japonica,* using 295,396
854 SNPs from 83 individuals. Native and invasive populations are coloured in green and yellow, respectively.

855

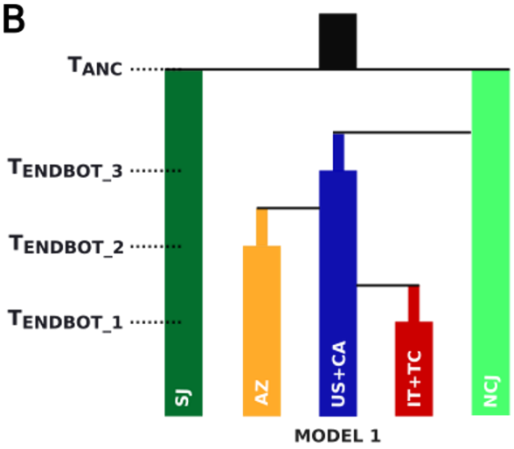
856 **3.5 Demographic history**

857 Among the six proposed demographic models (Figure 6), the strongest support (lowest AIC values;
858 Figure 19a) was shown for Model 1 (Figure 16a and Figure 19b), indicating that the USA was
859 colonized from North/Central Japan and that the Azores and Italy+Ticino were subsequently
860 colonized from the USA through two independent invasion events. This pattern was consistent with
861 the phylogenetic tree and all the others analyses of population structure.

A

MODEL	AIC
Model 1	**6645.299341**
Model 2	6918.326066
Model 3	6665.534459
Model 4	6948.139938
Model 5	6657.843825
Model 6	6659.524712

B

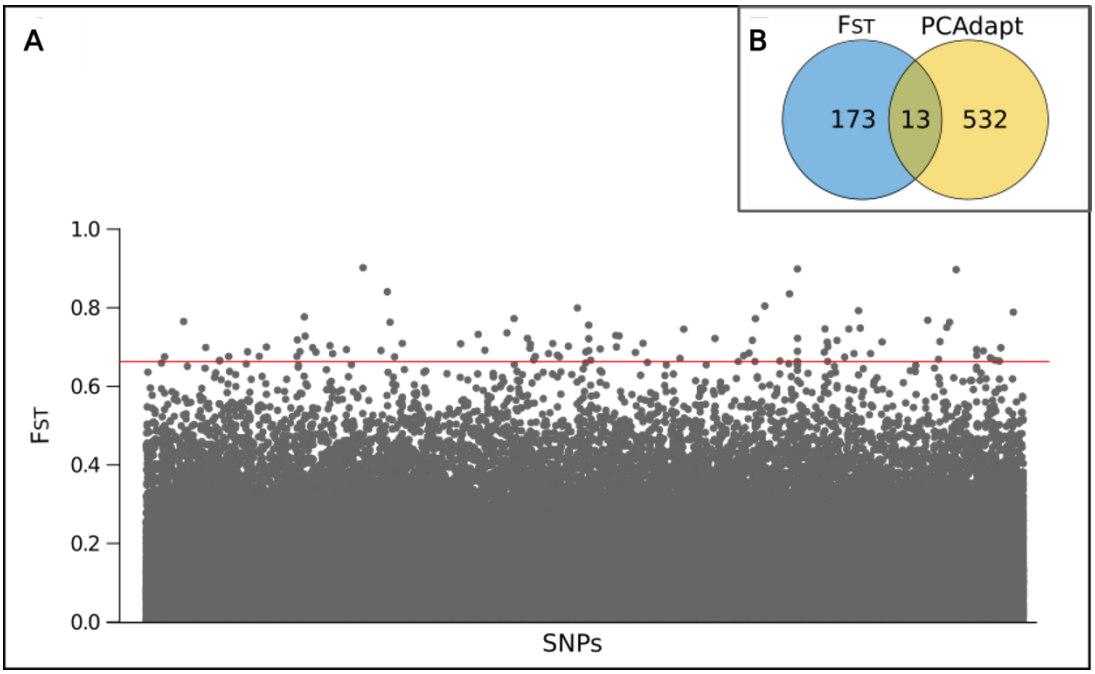


862 **Figure 19.** Best demographic model based on Fastsimcoal2 analysis. (A) AIC values for each tested model.
863 Best model and relative AIC value are highlighted in bold characters. (B) Best demographic model, showing
864 the pathways of invasion of *Popillia japonica* from native to invaded ranges. AIC, Akaike Information
865 Criterion; TANC, divergence time of non-uplift lineages; TENDBOT, bottleneck end time; SJ, South Japan,
866 NCJ: North/Central Japan; US+CA, USA+Canada; AZ, Azores; IT+TC, Italy+Ticino.

867

868 **3.6 Outlier SNPs and annotated genes**

869 The PCAdapt and FST-based genome-wide scans contrasting all invasive populations with
870 North/Central Japan identified 545 and 186 outliers, respectively, with 13 outliers in common (Figure
871 20). Variant annotation analysis showed that the majority of outlier SNPs were present in intergenic
872 regions (31.58%) and introns (31.58%), while the remaining SNPs were in downstream gene regions
873 (10.53%), upstream gene regions (15.79%), and exons (as synonymous variants, 10.53%). The latter
874 exon-based outliers were classified as low impact variants, and all others were categorized as
875 modifiers by SnpEff and SnpSift. Seven outlier SNPs had functional variant annotations for different
876 protein-coding gene families, including: ATP synthase alpha/beta, acetyl-CoA hydrolase/transferase,
877 pipsqueak, phosphoinositide 3-kinase, aldehyde dehydrogenase, and EF hand (Table 4).



878 **Figure 20.** Genome-scan analysis contrasting North/Central Japan (native) with all invasive populations
879 (USA+Canada, São Jorge, São Miguel, Italy+Ticino) of *Popillia japonica*, based on FST and PCAdapt. (A)
880 Manhattan plot of genome-wide weighted FST values, identified through sliding windows analysis within 5
881 kb non-overlapping genomic windows. The top 0.1% of SNPs are those above the red horizontal line,
882 indicating putative loci under selection. (B) Venn diagram showing outlier SNPs in common between the two
883 genome-scan analysis methods.

884

885

886

887

888

889

890

891

SNP position	Gene Annotation	Gene description	Gene function
000045F:432624	ATP synthase alpha/beta family, nucleotide-binding domain ATP synthase alpha/beta family, beta-barrel domain	V-type proton ATPase subunit B	Transmembrane ion transport
000071F:1412535	Acetyl-CoA hydrolase/transferase N-terminal domain Acetyl-CoA hydrolase/transferase C-terminal domain	4-hydroxybutyrate coenzyme A transferase	CoA-transferase activity
000406F:350909	helix-turn-helix, Psq domain	*Uncharacterized*	DNA binding
000445F:99440	Phosphatidylinositol 3- and 4-kinase Phosphoinositide 3-kinase family, accessory domain (PIK domain) Phosphoinositide 3-kinase C2	Phosphatidylinositol 3-kinase catalytic subunit type 3	Signal transduction
000463F:370370	Aldehyde dehydrogenase family	Aldehyde dehydrogenase, dimeric NADP-preferring	Oxidoreductase activity
000495F:207946	Aldehyde dehydrogenase family		
000706F:9597	Cytoskeletal-regulatory complex EF hand		Protein binding

892 **Table S5**. List of candidate genes (putatively under selection) derived from variant annotation analysis of seven
893 outlier SNPs, identified by F_{ST} and PCAdapt analyses contrasting North/Central Japan (native) with all
894 invasive populations (USA+Canada, São Jorge, São Miguel, Italy+Ticino) of *Popillia japonica*. SNP position
895 indicates genomic scaffold and locus. Annotations (gene annotation, gene description, and gene function) were
896 based on pfam, InterPro and InsectBase databases.
897
898
899
900
901
902
903

4. Discussion

Despite the availability of a draft sequence of the *P. japonica* genome since 2019 and the considerable interest generated by this pest's human-mediated expansion in the USA and Europe, studies on its biology and potential control methods have not fully exploited the available genetic data. Notable efforts aside, the application of molecular genetics remains limited. The new assembly presented here surpasses the existing draft in completeness and contiguity, featuring structural and functional annotations. This enhanced information is poised to illuminate molecular processes such as the extent of natural selection during the invasion process and the changes in gene expression associated with different control measures. This new genomic insight forms the basis for our investigation into the invasion history and geographic expansion routes of *P. japonica*.

In invasion biology, it is crucial to understand a species' history of invasion and geographic expansion routes. Here, genome-wide SNP data, from specimens that cover the entire species distribution, were used to obtain high resolution genomic patterns for *P. japonica* and reconstruct invasion pathways at a continental scale. Distinguishable genomic clusters were identified associated with independent invasion pathways. Population structure analysis, confirmed by demographic modelling, pinpointed USA as the source of all European populations through separate e distinct events of colonization in space and time, highlighting the increasingly common occurrence of bridgehead events in biological invasions. Finally, leveraging the new reference genome of *P. japonica*, possible signals of selection across various invasion routes of the species were also investigated. However, weak signals of selection in invasive populations were observed, suggesting that invasion success in *P. japonica* has likely been underlain by standing genetic variation.

4.1 *Popillia japonica*'s reference genome

Reference genomes play a pivotal role in advancing genetics, evolutionary biology, and molecular ecology by providing a foundational framework for comprehensive understanding and research, facilitating the study of genetic diversity, adaptation mechanisms, and species evolution. The data generated in this study could serve as a robust basis for future investigations into this pest, enabling targeted pest management strategies.

While the primary focus was not on resolving coleopteran systematic conflicts, the phylogenomic tree presented in Figure 8 contributes significantly to this field. This tree not only enhances our understanding of coleopteran phylogeny, but also underscores the evolutionary relationships among various beetle species. Notably, BUSCO single-copy orthologs have proven highly effective for exploring phylogenetic relationships at the whole genome level (Cicconardi et al., 2023). Despite the limited taxon sampling of 16 species, restricted to those with available complete genome sequences, the phylogenomic reconstruction strongly supports the previously proposed framework of Coleoptera systematics by Mckenna et al., (2015).

Regarding genome size and the presence of transposable elements (TEs), research suggests that TEs play an essential role in genome size evolution among eukaryotes, albeit with considerable variation. For instance, insects with smaller genomes, such as the midge *Belgica antarctica* (0.89 Mb), exhibit a modest TE repertoire (2.5%), whereas those with larger genomes, like the grasshopper *Locusta migratoria* (5.75 Gb), show expanded TE content (63.5%). Excluding these extremes, insect species with regular genome sizes (0.1 Gb to 2.7 Gb) typically have TE contents ranging from 17% to 30%. In Coleoptera, which includes genomes ranging from 0.19 Gb to 1.1 Gb, TE content varies from 12.9% to 41.4%. In this context, the genome of *P. japonica* displays a substantial yet not extreme expansion of TE repertoire (36.7%), ranking 13th out of 74 insect species studied by Petersen et al.,

(2019), alongside *Nicrophorus vespilloides* (Cunningham et al., 2015), and 2nd out of the 6 coleopterans (Figure 21). The detailed distribution of different TE classes observed in *P. japonica* (Figure 9, Figure 21) aligns broadly with other coleopterans investigated by Cunningham et al. (2015) and Petersen et al. (2019), with DNA elements and LINEs predominating among TE classes.

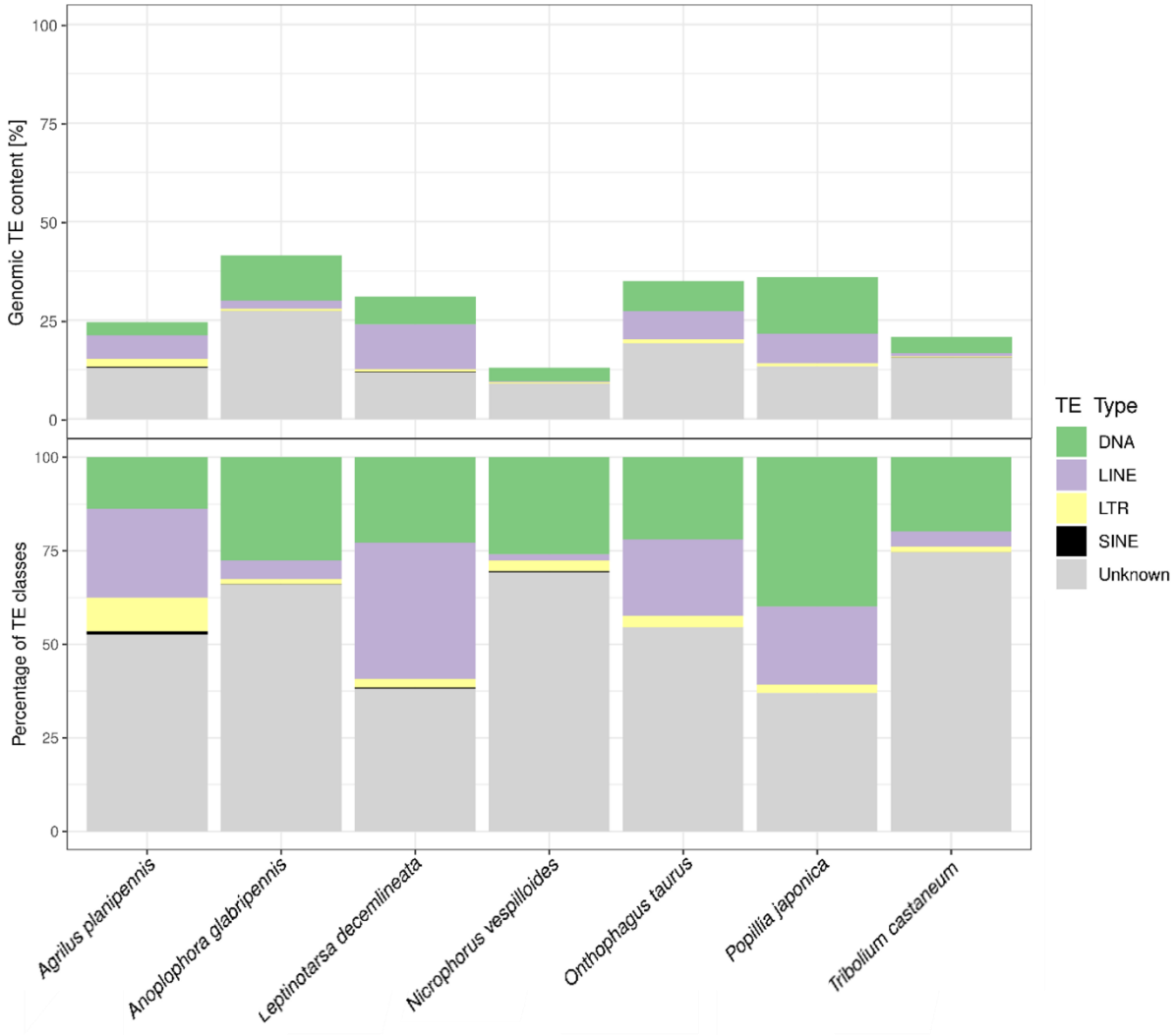


Figure 21. TE abundances across different species of Coleoptera. TE count resources are based on Petersen et al. (2019) and Cunningham et al. (2015). TE classes are coloured according to the provided key. TE, transposable element.

962 **4.2 Spatial genomic patterns and invasion pathways**

963 Population structure analyses clearly identified six genetic groups, that correspond to separate
964 geographic areas of the species distribution. These groups showed clear patterns of differentiation
965 and diversity, with invasive populations (i.e., USA+Canada, São Jorge, São Miguel, and Italy+Ticino)
966 diverging from native Japanese populations. Invasive species are expected to display lower genetic
967 diversity than native populations due to invasion-associated demographic events, such as founder
968 effects and genetic bottlenecks (Comeault et al., 2020; Michaelides et al., 2018). Tajima's D values
969 were positive for all populations, indicating an excess of intermediate-frequency alleles and
970 suggesting balancing selection in the native range. For the invasive populations, Tajima's D coupled
971 with lower π hinted at recent population size contractions and a loss of genetic diversity, as well as
972 depletion of rare alleles during colonization (Figure 12; Carlson et al., 2005; Comeault et al., 2020;
973 Yang et al., 2022). São Miguel had the slowest LD decay, indicating that the introduction of *P.*
974 *japonica* on the island was probably characterized by a very small founder population. Indeed, in both
975 invasive Azorean populations, higher LD, lower π, and higher Tajima's D, might be associated with
976 severe single-generation bottlenecks (Chen et al., 2018; Flanagan et al., 2021). In contrast, more rapid
977 decay for Italy+Ticino and USA+Canada suggests milder bottleneck events (Flanagan et al., 2021).
978 In addition, slow LD decay (for São Miguel, in particular), could be associated with populations of
979 constant size, while more rapid decay patterns, like Italy+Ticino and USA+Canada, might be a
980 signature of ongoing population expansion resulting in higher levels of recombination (Figure 14;
981 Shan et al., 2023; Slatkin, 1994).
982
983 Regarding the two native Japanese populations, our analyses confirmed previous research
984 (mitochondrial genomes [Nardi et al. 2024], and microsatellites and two mitochondrial loci [Strangi
985 et al. 2023]), showing ancestral divergence between the Southern population of Kyushu and Northern
986 and Central groups. As suggested in Nardi et al. (2024), this separation probably occurred between

5,274-25,532 years ago, after the last glaciation when the previously connected Islands of Kyushu and Honshu (Central Japan) were divided by the formation of the Seto Sea, isolating the Southern population. This reconstruction also explains the strong genetic distance of this population with respect to all others that was observed in the present study. Indeed, population structure analyses confirmed North/Central Japan as the most likely source of the beetle's global invasion. Specifically, our results suggested Honshu Island as the original source, supporting the observations made by Strangi et al. (2023). However, given the structure observed within this population (see Results), it is worth noting that individuals from Tsuruoka in the Central Island that are genetically closer to the invasive group clustered with individuals from Sapporo in the Northern Island, obscuring the reliable identification of the invasion source at finer spatial scales.

Bridgehead invasions have significantly increased during the last decades. Bertelsmeier et al., (2018) demonstrated that many ant species likely spread through bridgehead events, utilizing a century worth data on airports and seaports interceptions. Recent studies of insect invasion histories underscore the critical impact of international trade networks in driving biological invasions (Sherpa & Després, 2021). Numerous invasive insects originating from Eastern Asia, such as mosquitoes *A. albopictus*, *Drosophila suzukii*, the harlequin ladybird, *Harmonia axyridis*, and the brown marmorated stink bug, *Halyomorpha halys*, have established themselves in North America and subsequently spread to Europe, reflecting a shared bridgehead invasion model (Lombaert et al., 2010; Parvizi et al., 2023; Sherpa & Després, 2021). Our results unambiguously supported that, after the invasion of the USA from Japan, the present distribution of non-native populations of the Japanese beetle has been characterised by two independent bridgehead events. Specifically, our findings indicated that the European invasions in the Azores and Italy originated from already invaded areas of the USA, with evident admixture (at K=5) between the USA and São Jorge and the USA and Italy (Figure 16).

While our genomic analyses agreed with those of previous studies, we found some discrepancies regarding the origin of European populations in the USA at finer spatial scales. For example, Nardi et al. (2024) observed a certain association between one mitochondrial genome in Maryland (East USA) and the Italian group, whereas Strangi et al. (2023), based on two mitochondrial loci, suggested the East coast of the USA and New Jersey as sources of invasion to the Azores and Italy, respectively. Our phylogenetic analysis based on SNPs from WGR suggested Central-East USA as the most likely area of origin of the Azorean invasion, while little information could be obtained for the origin of the Italian outbreak, apart from a mild preference for Central-West USA. We suggest that this lack of resolution, more than from a limitation in the data, stems from the extensive intermixing and lack of structure within populations, which was also found in other studies, that does not allow a confident localization of the source of European individuals at finer geographic scales. Lastly, our PCA, phylogeny, and admixture analyses supported the previously described hypothesis that the Japanese beetle has reached Canada and Ticino (Switzerland) from neighbouring countries (i.e., USA and Italy) without post-invasion genetic differentiation. This is plausible given the great hitchhiking and flight capacity of *P. japonica* (Fleming, 1972; Allsopp et al. 1996; Poggi et al., 2022; Borner et al., 2023a).

4.3 Genomic signatures of selection and adaptation

The majority of outlier SNPs we identified lay in close proximity to genes associated with diverse metabolism and cellular functions. However, to the best of our knowledge, their involvement in invasion and in situ adaptation of *P. japonica* and other insect pests remains unclear. An exception to this was an outlier SNP located in close proximity to an ATP synthase-encoding gene involved in transmembrane ion transport, which has known functions in insect acclimation and water homeostasis (Enriquez & Colinet, 2019; Galikova et al., 2018).

Genome scan analysis, supported by Tajima's D results, suggested an overall weak signal of in situ adaptation in invasive populations (Figure 4). This pattern might be a consequence of the beetle's recent invasion history, since local adaptation may not be an essential attribute at the beginning of a successful invasion and may instead occur long after species establishment and spread (North et al., 2021). Additionally, it is well known that climatic and environmental conditions similar to those in the native distribution, extensive land use, and the highly-generalist herbivore feeding habit of P. japonica have together favoured the spread of the beetle (Fleming, 1972; Hamilton et al., 2007). Borner et al. (2023b) suggested that the majority of the American and European present infestation areas are moderately-to-highly suitable for the Japanese beetle. In addition, the recent investigation in Cucini et al. (2024) on gene families involved in chemoreception and detoxification reports paralogous expansion in different subfamilies/clans of odorant receptors, ionotropic receptors, and cytochrome p450 encoding genes. This may suggest that some invasion-facilitating characters, linked with the species' ability to feed on numerous plants and resist selective pressures from control and management measures, are already embedded in the Japanese beetle's genome. Altogether, the habitat suitability of currently invaded areas, in addition to the species' polyphagy and existence of pre-invasion adaptive genomic traits, may explain the invasion success of *P. japonica* in the absence of strong signatures of selection.

5. Conclusion

Despite some limitations in the newly described assembly, such as sequencing from a pool of multiple individuals and the lack of chromosome-level scaffolding, this genome provides a significant contribution to understanding the biology of *Popillia japonica*. The availability of a high-quality genome, complete with structural and functional annotations, provides numerous opportunities for further research into the species' biology and adaptations. Furthermore, genome-wide SNP datasets and genomic tools allowed a detailed inference of the global invasion process of *P. japonica*. The identified geographic lineages, their genome-wide variability and distribution world-wide, the relationships among populations, and the individual events of invasion of European territories from the USA, were well-resolved with the presented comprehensive dataset. This work provides a robust genomic databank against which genetic variation of future incursions can be compared to confidently pinpoint their origin. Moreover, the availability of a newly assembled genome and the SNPs datasets might help the design of surveillance strategies for integrated and targeted pest management practices in this species through the integration of molecular genetic information. Additionally, it also facilitates the development of biotechnological applications for pest control. Results of the present research found an absence of strong selection taking place on invasive populations, suggesting that invasion success might be attributable to the presence of pre-adaptive traits in *P. japonica*. However, to better understand the drivers of invasion success, future studies on *P. japonica* might focus on identification and analysis of transposable elements and structural variants, since a growing body of literature has delved into their potential roles on fitness and adaptation during invasion (Bertolotti et al., 2020; Carareto et al., 2020; Mérot et al., 2020).

Acknowledgements

Firstly, I would like to thank my supervisors, Prof. Antonio Carapelli, Prof. Francesco Nardi, and Prof. Francesco Frati, for their scientific guidance, feedback, and expertise. I am also grateful to the members of the ESZlab group with whom I have shared projects, ideas, and collaborations, providing me with both moral and scientific support. In particular, I would like to thank Dr. Claudio Cucini for his constant help, as well as for being part of my professional formation. Words cannot express my gratitude to Prof. Angela McGaughran and Dr. Elahe Parvizi, who supervised me during my research stay at the Invasomics Lab at the University of Waikato (Hamilton, New Zealand). Finally, I would like to extend my sincere thanks to the reviewers and the dissertation committee.

References

Allsopp, P. G. (1996). Japanese beetle, *Popillia japonica* Newman (Coleoptera: Scarabaeidae): rate of movement and potential distribution of an immigrant species. *The Coleopterists' Bulletin*, 81–95. http://www.jstor.org/stable/4009259

Agroscope. (2023). Japankäfer in Kloten entdeckt – erste Population nördlich der Alpen. Available at: https://www.agroscope.admin.ch/agroscope/de/home/aktuell/newsroom/2023/07-25_japankaefer_kloten.html

Alpert, P. (2006). The advantages and disadvantages of being introduced. *Biological Invasions*, 8, 1523–1534. https://doi.org/10.1007/s10530-005-5844-z

Althoff, E. R., & Rice, K. B. (2022). Japanese beetle (Coleoptera: Scarabaeidae) invasion of North America: history, ecology, and management. *Journal of Integrated Pest Management*, 13(1), 2. https://doi.org/10.1093/jipm/pmab043

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

1110 Astashyn, A., Tvedte, E. S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strope,
1111 P. K., Sylla, P. M., & Wagner, L. (2024). Rapid and sensitive detection of genome contamination at
1112 scale with FCS-GX. *Genome biology*, 25(1), 60. https://doi.org/10.1101/2023.06.02.543519

1113 Baker, A. M., & Potter, D. A. (2018). Japanese beetles' feeding on milkweed flowers may
1114 compromise efforts to restore monarch butterfly habitat. *Scientific Reports*, 8(1), 12139.
1115 https://doi.org/10.1038/s41598-018-30731-z

1116 Bertelsmeier, C., & Keller, L. (2018). Bridgehead effects and role of adaptive evolution in invasive
1117 populations. *Trends in Ecology & Evolution*, 33(7), 527–534.
1118 https://doi.org/10.1016/j.tree.2018.04.014

1119 Bertelsmeier, C., Ollier, S., Liebhold, A. M., Brockerhoff, E. G., Ward, D., & Keller, L. (2018).
1120 Recurrent bridgehead effects accelerate global alien ant spread. *Proceedings of the National*
1121 *Academy of Sciences*, 115(21), 5486–5491. https://doi.org/10.1073/pnas.1801990115

1122 Bertolotti, A. C., Layer, R. M., Gundappa, M. K., Gallagher, M. D., Pehlivanoglu, E., Nome, T.,
1123 Robledo, D., Kent, M. P., Røsæg, L. L., Holen, M. M., et al. (2020). The structural variation
1124 landscape in 492 Atlantic salmon genomes. *Nature Communications*, 11(1), 1-16.
1125 https://doi.org/10.1038/s41467-020-18972-x

1126 Blumenfeld, A. J., Eyer, P.-A., Husseneder, C., Mo, J., Johnson, L. N., Wang, C., Kenneth Grace, J.,
1127 Chouvenc, T., Wang, S., & Vargo, E. L. (2021). Bridgehead effect and multiple introductions shape
1128 the global invasion history of a termite. *Communications Biology*, 4(1), 196.
1129 https://doi.org/10.1038/s42003-021-01725-x

1130 Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., Turner, K. G.,
1131 Whitney, K. D., & Rieseberg, L. H. (2016). What we still don't know about invasion genetics.
1132 *Invasion genetics: The baker and stebbins legacy*, 346–370. https://doi.org/10.1111/mec.13032

1133 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina
1134 sequence data. *Bioinformatics*, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

1135 Borner, L., Martinetti, D., & Poggi, S. (2023b). A new chapter of the Japanese beetle invasion saga:
1136 predicting suitability from long-infested areas to inform surveillance strategies in Europe.
1137 *Entomologia Generalis*, 43(5), 951-960. https://doi.org/10.1101/2022.11.14.515960

1138 Borner, L., Martinetti, D., & Poggi, S. (2023a). A hitchhiker's guide to Europe: mapping human-
1139 mediated dispersal of the invasive Japanese beetle. *bioRxiv,* 2023–10.
1140 https://doi.org/10.1101/2023.10.06.561184

1141 Borowiec, M. L. (2016). AMAS: a fast tool for alignment manipulation and computing of summary
1142 statistics. *PeerJ*, 4, e1660. https://doi.org/10.7717/peerj.1660

1143 Bushnell, B. (2014). BBMap: A fast, accurate, splice-aware aligner. Lawrence Berkeley National
1144 Lab.(LBNL), Berkeley, CA (United States).

1145 Capinha, C., Essl, F., Seebens, H., Moser, D., & Pereira, H. M. (2015). The dispersal of alien
1146 species redefines biogeography in the Anthropocene. *Science*, 348(6240), 1248–1251.
1147 https://doi.org/10.1126/science.aaa8913

1148 Carareto, C. M. A., Hernandez, E. H., & Vieira, C. (2014). Genomic regions harbouring insecticide
1149 resistance-associated Cyp genes are enriched by transposable element fragments carrying putative
1150 transcription factor binding sites in two sibling *Drosophila* species. *Gene*, 537(1), 93–99.
1151 https://doi.org/10.1016/j.gene.2013.11.080

1152 Carlson, C. S., Thomas, D. J., Eberle, M. A., Swanson, J. E., Livingston, R. J., Rieder, M. J., &
1153 Nickerson, D. A. (2005). Genomic regions exhibiting positive selection identified from dense
1154 genotype data. *Genome research*, 15(11), 1553–1565. https://doi.org/10.1101/gr.4326505

1155 Carpenter, S.J., LaCasse, W.J. (1955). Mosquitoes of North America (North of Mexico). Berkeley:
1156 University of California Press.

Carroll, E., Kunte, N., McGraw, E., Gautam, S., Range, R., Noveron-Nunez, J. A., Held, D. W., & Avila, L. A. (2023). Gene silencing in adult Popillia japonica through feeding of double-stranded RNA (dsRNA) complexed with branched amphiphilic peptide capsules (BAPCs). *Frontiers in Insect Science*, 3, 1151789. https://doi.org/10.3389/finsc.2023.1151789

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Champer, J., Buchman, A., & Akbari, O. S. (2016). Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nature Reviews Genetics*, 17(3), 146–159. https://doi.org/10.1038/nrg.2015.34

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chown, S. L., Hodgins, K. A., & Griffin, P. C. (2016). Biological invasions, climate change, and genomics. *Crop Breeding*, 59–114. https://doi.org/10.1111/eva.12234

Cicconardi, F., Milanetti, E., Pinheiro de Castro, E. C., Mazo-Vargas, A., Van Belleghem, S. M., Ruggieri, A. A., Rastas, P., Hanly, J., Evans, E., & Jiggins, C. D. (2023). Evolutionary dynamics of genome size and content during the adaptive radiation of Heliconiini butterflies. *Nature Communications*, 14(1), 5620. https://doi.org/10.1038/s41467-023-41412-5

Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., & Lu, X. (2012b). Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Frontiers in genetics*, 3, 35. https://doi.org/10.3389/fgene.2012.00035

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012a). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6(2), 80–92. https://doi.org/10.4161/fly.19695

1181 Clausen, C. P., King, J. L., & Teranishi, C. (1927). The parasites of *Popillia japonica* in Japan and
1182 Chosen (Korea), and their introduction into the United States (Fascicolo 1429). US Department of
1183 Agriculture.

1184 Comeault, A. A., Wang, J., Tittes, S., Isbell, K., Ingley, S., Hurlbert, A. H., & Matute, D. R. (2020).
1185 Genetic diversity and thermal performance in invasive and native populations of African fig flies.
1186 *Molecular biology and evolution*, 37(7), 1893–1906. https://doi.org/10.1093/molbev/msaa050

1187 Cucini, C., Boschi, S., Funari, R., Cardaioli, E., Iannotti, N., Marturano, G., Paoli, F., Bruttini, M.,
1188 Carapelli, A., & Frati, F. (2024). De novo assembly and annotation of Popillia japonica's genome
1189 with initial clues to its potential as an invasive pest. *BMC genomics*, 25(1), 275.
1190 https://doi.org/10.1186/s12864-024-10180-x

1191 Cunningham, C. B., Ji, L., Wiberg, R. A. W., Shelton, J., McKinney, E. C., Parker, D. J., Meagher,
1192 R. B., Benowitz, K. M., Roy-Zokan, E. M., & Ritchie, M. G. (2015). The genome and methylome
1193 of a beetle with complex social behavior, *Nicrophorus vespilloides* (Coleoptera: Silphidae). *Genome*
1194 *Biology and Evolution*, 7(12), 3383–3396. https://doi.org/10.1093/gbe/evv194

1195 Daly, E. Z., Chabrerie, O., Massol, F., Facon, B., Hess, M. C., Tasiemski, A., Grandjean, F.,
1196 Chauvat, M., Viard, F., & Forey, E. (2023). A synthesis of biological invasion hypotheses associated
1197 with the introduction–naturalisation–invasion continuum. *Oikos*, 2023(5), e09645.
1198 https://doi.org/10.1111/oik.09645

1199 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E.,
1200 Lunter, G., Marth, G. T., & Sherry, S. T. (2011). The variant call format and VCFtools.
1201 *Bioinformatics*, 27(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

1202 Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A.,
1203 Keane, T., McCarthy, S. A., & Davies, R. M. (2021). Twelve years of SAMtools and BCFtools.
1204 *Gigascience*, 10(2), giab008. https://doi.org/10.1093/gigascience/giab008

De Coster, W., D'hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, 34(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J. A., & Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, 592(7855), 571–576. https://doi.org/10.1038/s41586-021-03405-6

EFSA Panel on Plant Health (PLH), Bragard, C., Dehnen-Schmutz, K., Di Serio, F., Gonthier, P., Jacques, M., Jaques Miret, J. A., Justesen, A. F., Magnusson, C. S., & Milonas, P. (2018). Pest categorisation of *Popillia japonica. Efsa Journal*, 16(11), e05438. https://doi.org/10.2903/j.efsa.2018.5438

Enriquez, T., & Colinet, H. (2019). Cold acclimation triggers major transcriptional changes in *Drosophila suzukii*. *BMC genomics*, 20, 1–17. https://doi.org/10.1186/s12864-019-5745-7

EPPO *Popillia japonica* (2019). EPPO global database. Available at: https://gd.eppo.int/taxon/POPIJA/distribution/PT_az

EPPO *Popillia japonica* (2022). EPPO datasheets on pests recommended for regulation. Available at: https://gd.eppo.int/taxon/POPIJA

ERSA. 2023. Piano d'azione per gestire l'emergenza *Popillia japonica* Newman in Friuli-Venezia Giulia. Available at: https://fitosanitario.regione.lombardia.it/wps/wcm/connect/e90054cb-a4da-4ac9-a603-885c69912b46/MASAF-2023-0529990-Allegato-ODSFCn5_2023Popilliajaponica_signed.pdf?MOD=AJPERES

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

1228 Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a
1229 genetic paradox of biological invasion? *Annual review of ecology, evolution, and systematics*, 47(1),
1230 51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

1231 Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust
1232 demographic inference from genomic and SNP data. *PLoS genetics*, 9(10), e1003905.
1233 https://doi.org/10.1371/journal.pgen.1003905

1234 Excoffier, L., Marchi, N., Marques, D. A., Matthey-Doret, R., Gouy, A., & Sousa, V. C. (2021).
1235 Fastsimcoal2: demographic inference under complex evolutionary scenarios. *Bioinformatics*,
1236 37(24), 4882–4885. https://doi.org/10.1093/bioinformatics/btab468

1237 Flanagan, B. A., Krueger-Hadfield, S. A., Murren, C. J., Nice, C. C., Strand, A. E., & Sotka, E. E.
1238 (2021). Founder effects shape linkage disequilibrium and genomic diversity of a partially clonal
1239 invader. *Molecular Ecology*, 30(9), 1962–1978. https://doi.org/10.1111/mec.15854

1240 Fleming, W.E. (1972). Biology of the Japanese beetle (Technical bulletin no. 1449). Washington
1241 DC: US Department of Agriculture.

1242 Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020).
1243 RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of*
1244 *the national academy of sciences*, 117(17), 9451–9457. https://doi.org/10.1073/pnas.1921046117

1245 Frank, K. D. (2016). Establishment of the Japanese beetle (*Popillia japonica* Newman) in North
1246 America near Philadelphia a century ago. *Entomological News*, 126(3), 153–174.
1247 https://doi.org/10.3157/021.126.0302

1248 Frankham, R. (2005). Resolving the genetic paradox in invasive species. *Heredity*, 94(4).
1249 https://doi.org/10.1038/sj.hdy.6800634

1250 Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association
1251 studies. *Methods in ecology and evolution*, 6(8), 925–929. https://doi.org/10.1111/2041-210X.12382

1252 Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient
1253 estimation of individual ancestry coefficients. *Genetics*, 196(4), 973–983.
1254 https://doi.org/10.1534/genetics.113.160572

1255 Galikova, M., Dircksen, H., & Nässel, D. R. (2018). The thirsty fly: Ion transport peptide (ITP) is a
1256 novel endocrine regulator of water homeostasis in *Drosophila. PLoS genetics*, 14(8), e1007618.
1257 https://doi.org/10.1371/journal.pgen.1007618

1258 Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-
1259 specific covariates. *Genetics*, 201(4), 1555–1579. https://doi.org/10.1534/genetics.115.181453

1260 Gotta, P., Ciampitti, M., Cavagna, B., Bosio, G., Gilioli, G., Alma, A., Battisti, A., Mori, N., Mazza,
1261 G., & Torrini, G. (2023). *Popillia japonica*–Italian outbreak management. *Frontiers in Insect*
1262 *Science*, 3, 1175138. https://doi.org/10.3389/finsc.2023.1175138

1263 Hamilton, R., Foster, R., Gibb, T., Sadof, C., Holland, J., & Engel, B. (2007). Distribution and
1264 dynamics of Japanese beetles along the Indianapolis airport perimeter and the influence of land use
1265 on trap catch. *Environmental entomology*, 36(2), 287–296. https://doi.org/10.1093/ee/36.2.287

1266 Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2016). BRAKER1:
1267 unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS.
1268 *Bioinformatics*, 32(5), 767–769. https://doi.org/10.1093/bioinformatics/btv661

1269 Hoffmeister, T. S., Vet, L. E., Biere, A., Holsinger, K., & Filser, J. (2005). Ecological and
1270 evolutionary consequences of biological invasion and habitat fragmentation. *Ecosystems*, 8, 657–
1271 667. https://doi.org/10.1007/s10021-003-0138-8

1272 Hulme, P. E., Nentwig, W., Pyšek, P., & Vilà, M. (2009). Common market, shared problems: time
1273 for a coordinated response to biological invasions in Europe. *Neobiota*, 8, 3–19.

1274 Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., Jahesh, G.,
1275 Khan, H., Coombe, L., & Warren, R. L. (2017). ABySS 2.0: resource-efficient assembly of large

genomes using a Bloom filter. *Genome research*, 27(5), 768–777. https://doi.org/10.1101/gr.214346.116

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in bioinformatics*, 20(4), 1160–1166. https://doi.org/10.1093/bib/bbx108

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology*, 37(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kistner-Thomas, E. J. (2019). The potential global distribution and voltinism of the Japanese beetle (Coleoptera: Scarabaeidae) under current and future climates. *Journal of Insect Science*, 19(2), 16. https://doi.org/10.1093/jisesa/iez023

Kolar, C. S., & Lodge, D. M. (2001). Progress in invasion biology: Predicting invaders. *Trends in ecology & evolution*, 16(4), 199–204. https://doi.org/10.1016/s0169-5347(01)02101-2

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature biotechnology*, 37(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8

Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., Nolan, T., & Crisanti, A. (2018). A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nature biotechnology*, 36(11), 1062–1066. 1062–1066. https://doi.org/10.1038/nbt.4245

Lawson, D. J., Van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature communications*, 9(1), 3258. https://doi.org/10.1038/s41467-018-05257-7

Lee, D.-H. (2015). Current status of research progress on the biology and management of *Halyomorpha halys* (Hemiptera: Pentatomidae) as an invasive species. *Applied entomology and zoology*, 50(3), 277–290. https://doi.org/10.1007/s13355-015-0350-y

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PloS one*, 5(3), e9743. https://doi.org/10.1371/journal.pone.0009743

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10–12. https://doi.org/10.14806/ej.17.1.200.

Matheson, P., & McGaughran, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports*, 12(1), 13987. https://doi.org/10.1038/s41598-022-17937-y

McCartney, M. A., Mallez, S., & Gohl, D. M. (2019). Genome projects in invasion biology. *Conservation Genetics*, 20(6), 1201–1222. https://doi.org/10.1007/s10592-019-01224-x

McGaughran, A., Dhami, M. K., Parvizi, E., Vaughan, A. L., Gleeson, D. M., Hodgins, K. A., Rollins, L. A., Tepolt, C. K., Turner, K. G., & Atsawawaranunt, K. (2024). Genomic tools in biological invasions: current state and future frontiers. *Genome biology and evolution*, 16(1), evad230. https://doi.org/10.1093/gbe/evad230

1322 Mckenna, D. D., Wild, A. L., Kanda, K., Bellamy, C. L., Beutel, R. G., Caterino, M. S., Farnum, C.
1323 W., Hawks, D. C., Ivie, M. A., & Jameson, M. L. (2015). The beetle tree of life reveals that
1324 Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial
1325 revolution. *Systematic Entomology*, 40(4), 835–880. https://doi.org/10.1111/syen.12132

1326 McKinney, Wes. (2010). Data structures for statistical computing in Python. 56-61.
1327 https://doi.org/10.25080/Majora-92bf1922-00a.

1328 Melbourne, B. A., Cornell, H. V., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A. L., Hall,
1329 R. J., Harrison, S., Hastings, A., & Holland, M. (2007). Invasion in a heterogeneous world:
1330 resistance, coexistence or hostile takeover? *Ecology letters*, 10(1), 77–94.
1331 https://doi.org/10.1111/j.1461-0248.2006.00987.x

1332 Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A roadmap for understanding the
1333 evolutionary significance of structural genomic variation. *Trends in Ecology & Evolution*, 35(7),
1334 561–572. https://doi.org/10.1016/j.tree.2020.03.002

1335 Michaelides, S. N., Goodman, R. M., Crombie, R. I., & Kolbe, J. J. (2018). Independent
1336 introductions and sequential founder events shape genetic differentiation and diversity of the
1337 invasive green anole (*Anolis carolinensis*) on Pacific Islands. *Diversity and Distributions*, 24(5),
1338 666–679. https://doi.org/10.1111/ddi.12704

1339 Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome
1340 assembly evaluation with QUAST-LG. *Bioinformatics*, 34(13), i142–i150.
1341 https://doi.org/10.1093/bioinformatics/bty266

1342 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., &
1343 Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in
1344 the genomic era. *Molecular biology and evolution*, 37(5), 1530–1534.
1345 https://doi.org/10.1093/molbev/msaa015

Mondino, E. B., Lessio, F., Bianchi, A., Ciampitti, M., Cavagna, B., & Alma, A. (2022). Modelling the spread of *Popillia japonica* Newman (Coleoptera: Scarabaeidae) from a recently infested area. *Entomologia Generalis*, 42(5), 713–721. https://doi.org/10.1127/entomologia/2022/1370

Nardi, F., Boschi, S., Funari, R., Cucini, C., Cardaioli, E., Potter, D., Asano, S.-I., Toubarro, D., Meier, M., & Paoli, F. (2024). The direction, timing and demography of *Popillia japonica* (Coleoptera) invasion reconstructed using complete mitochondrial genomes. *Scientific Reports*, 14(1), 7120. https://doi.org/10.1038/s41598-024-57667-x

Näslund, J., Ahlm, C., Islam, K., Evander, M., Bucht, G., & Lwande, O. W. (2021). Emerging mosquito-borne viruses linked to *Aedes aegypti* and *Aedes albopictus*: Global status and preventive strategies. *Vector-Borne and Zoonotic Diseases*, 21(10), 731–746. https://doi.org/10.1089/vbz.2020.2762

North, H. L., McGaughran, A., & Jiggins, C. D. (2021). Insights into invasive species from whole-genome resequencing. *Molecular Ecology*, 30(23), 6289–6308. https://doi.org/10.1111/mec.15999

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., & Gershman, A. (2022). The complete sequence of a human genome. *Science*, 376(6588), 44–53. https://doi.org/10.1126/science.abj6987

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., Diepenbrock, L. M., Kenis, M., Zhang, J., & Chen, X. (2020). A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. *Molecular biology and evolution*, 37(8), 2369–2385. https://doi.org/10.1093/molbev/msaa098

Parvizi, E., Dhami, M. K., Yan, J., & McGaughran, A. (2023). Population genomic insights into invasion success in a polyphagous agricultural pest, *Halyomorpha halys*. *Molecular ecology*, 32(1), 138–151. https://doi.org/10.1111/mec.16740

1370 Pavesi, M. (2014). *Popillia japonica* specie aliena invasiva segnalata in Lombardia. *L'informatore*
1371 *Agrario*, 32, 53–55.

1372 Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: an R package for calculation of
1373 genetic differentiation and structure of mixed-ploidy level populations. *Molecular ecology*
1374 *resource*s, 13(5), 946–952. https://doi.org/10.1111/1755-0998.12129

1375 Petersen, M., Armisén, D., Gibbs, R. A., Hering, L., Khila, A., Mayer, G., Richards, S., Niehuis, O.,
1376 & Misof, B. (2019). Diversity and evolution of the transposable element repertoire in arthropods
1377 with particular reference to insects. *BMC Ecology and Evolution*, 19, 1–15.
1378 https://doi.org/10.1186/s12862-018-1324-9

1379 Poggi, S., Borner, L., Roche, J., Tayeh, C., Martinetti, D. (2023). Biological invasion of the
1380 Japanese beetle in Continental Europe at a glance. *Recherche Data Gouv*, V3.
1381 https://doi.org/10.57745/R18NGL

1382 Poggi, S., Desneux, N., Jactel, H., Tayeh, C., & Verheggen, F. (2022). A nationwide pest risk
1383 analysis in the context of the ongoing Japanese beetle invasion in Continental Europe: The case of
1384 metropolitan France. *Frontiers in Insect Science*, 2, 1079756.
1385 https://doi.org/10.3389/finsc.2022.1079756

1386 Potter, D. A., & Held, D. W. (2002). Biology and management of the Japanese beetle. *Annual*
1387 *review of entomology*, 47(1), 175–205. https://doi.org/10.1146/annurev.ento.47.091201.145153

1388 Prentis, P. J., Wilson, J. R., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive
1389 evolution in invasive species. *Trends in plant science*, 13(6), 288–294.
1390 https://doi.org/10.1016/j.tplants.2008.03.004

1391 Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. (2020). Performing highly efficient genome
1392 scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*,
1393 37(7), 2153–2154. https://doi.org/10.1093/molbev/msaa053

1394 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P.,
1395 De Bakker, P. I., & Daly, M. J. (2007). PLINK: a tool set for whole-genome association and
1396 population-based linkage analyses. *The American journal of human genetics*, 81(3), 559–575.
1397 https://doi.org/10.1086/519795

1398 Pyšek, P., Hulme, P. E., Simberloff, D., Bacher, S., Blackburn, T. M., Carlton, J. T., Dawson, W.,
1399 Essl, F., Foxcroft, L. C., & Genovesi, P. (2020). Scientists' warning on invasive alien species.
1400 *Biological Reviews*, 95(6), 1511–1534. https://doi.org/10.1111/brv.12627

1401 Rhie, A., Walenz, B. P., Koren, S., & Phillippy, A. M. (2020). Merqury: reference-free quality,
1402 completeness, and phasing assessment for genome assemblies. *Genome biology*, 21, 1–27.
1403 https://doi.org/10.1186/s13059-020-02134-9

1404 Richardson, D. M., & Pyšek, P. (2006). Plant invasions: merging the concepts of species
1405 invasiveness and community invasibility. *Progress in physical geography*, 30(3), 409–431.
1406 https://doi.org/ 10.1191/0309133306pp490pr

1407 Roach, M. J., Schmidt, S. A., & Borneman, A. R. (2018). Purge haplotigs: allelic contig
1408 reassignment for third-gen diploid genome assemblies. *BMC bioinformatics*, 19, 1–10.
1409 https://doi.org/10.1186/s12859-018-2485-7

1410 Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: analytical methods for
1411 paired-end sequencing improve RADseq-based population genomics. *Molecular ecology*, 28(21),
1412 4737–4754. https://doi.org/10.1111/mec.15253

1413 Roman, J., & Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic
1414 invasions. *Trends in ecology & evolution*, 22(9), 454–464.
1415 https://doi.org/10.1016/j.tree.2007.07.002

1416 Schrieber, K., & Lachmuth, S. (2017). The Genetic paradox of invasions revisited: the potential role
1417 of inbreeding × environment interactions in invasion success. *Biological reviews*, 92(2), 939–952.
1418 https://doi.org/10.1111/brv.12263

1419 Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S.,
1420 Pyšek, P., van Kleunen, M., & Winter, M. (2018). Global rise in emerging alien species results from
1421 increased accessibility of new source pools. *Proceedings of the national academy of sciences*,
1422 115(10), E2264–E2273. https://doi.org/10.1073/pnas.1719429115

1423 Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S.,
1424 Pyšek, P., Winter, M., & Arianoutsou, M. (2017). No saturation in the accumulation of alien species
1425 worldwide. *Nature communications*, 8(1), 14435. https://doi.org/10.1038/ncomms14435

1426 Servizio fitosanitario cantonale. (2017). Bollettino fitosanitario n°25 del 26-VI-2017. Repubblica e
1427 Cantone Ticino. Available at: https://www4.ti.ch/fileadmin/DFE/DE-
1428 SA/comunicati/fito/2017/bof2517.pdf

1429 Shan, B., Yu, G., Wang, L., Liu, Y., Yang, C., Liu, M., & Sun, D. (2023). Genetic signature of
1430 *Pinctada fucata* inferred from population genomics: source tracking of the invasion in Mischief
1431 Reef of Nansha Islands. *Biology*, 12(1), 97. https://doi.org/10.3390/biology12010097

1432 Sherpa, S., & Després, L. (2021). The evolutionary dynamics of biological invasions: a multi-
1433 approach perspective. *Evolutionary applications*, 14(6), 1463–1484.
1434 https://doi.org/10.1111/eva.13215

1435 Simberloff, D. (2009). The role of propagule pressure in biological invasions. *Annual review of*
1436 *ecology, evolution, and systematics*, 40(1), 81–102. https://doi.org/10.1016/j.tree.2012.07.013

1437 Simões, A. M. M. A. (1984). Observações acerca de *Popillia japonica* Newman na Ilha Terceira.
1438 *Arquipelago, ciencias da natureza*. 5, 129-156.

1439 Slatkin, M. (1994). Linkage disequilibrium in growing and stable populations. *Genetics*, 137(1),
1440 331–336. https://doi.org/10.1093/genetics/137.1.331

1441 Smit, A., Hubley, R., & Green, P. (2013). 2015. RepeatMasker Open-4.0. Available at:
1442 http://www.repeatmasker.org

1443 Strangi, A., Paoli, F., Nardi, F., Shimizu, K., Kimoto, T., Iovinella, I., Bosio, G., Roversi, P. F.,
1444 Carapelli, A., & Marianelli, L. (2024). Tracing the dispersal route of the invasive Japanese beetle
1445 *Popillia japonica*. *Journal of Pest Science*, 97(2), 613–629. https://doi.org/10.1007/s10340-023-
1446 01653-1

1447 Straubinger, F. B., Benjamin, E. O., Venus, T. E., & Sauer, J. (2022). The economic importance of
1448 early pest control: new insights from potential *Popillia japonica* infestation in Europe. *AgriRxiv*,
1449 2022, 20220316785. https://doi.org/10.31220/agriRxiv.2022.00151

1450 Tabashnik, B. E., Liesner, L. R., Ellsworth, P. C., Unnithan, G. C., Fabrick, J. A., Naranjo, S. E., Li,
1451 X., Dennehy, T. J., Antilla, L., & Staten, R. T. (2021). Transgenic cotton and sterile insect releases
1452 synergize eradication of pink bollworm a century after it invaded the United States. *Proceedings of*
1453 *the national academy of sciences*, 118(1), e2019115118. https://doi.org/10.1073/pnas.2019115118

1454 Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA
1455 polymorphism. *Genetics*, 123(3), 585–595. https://doi.org/10.1093/genetics/123.3.585

1456 Tang, L., Li, M., Wu, F.-X., Pan, Y., & Wang, J. (2020). MAC: merging assemblies by using
1457 adjacency algebraic model and classification. *Frontiers in genetics*, 10, 1396.
1458 https://doi.org/10.3389/fgene.2019.01396

1459 Tayeh, C., Poggi, S., Desneux, N., Jactel, H., & Verheggen, F. (2023). Host plants of *Popillia*
1460 *japonica*: a review. *Recherche Data Gouv*, V2. https://doi.org/10.57745/SXZNQF

1461 Teixeira, M. B., Soares, A. O., Lamelas-López, L., Lopes, D. H., Mota, J. A., Borges, P. A. V.,
1462 Simões, N. (2024). Monitoring populations of *Popillia japonica* (Newman, 1838) over 16 years in

1463 the Azorean Islands. Version 1.3. Universidade dos Açores. Occurrence dataset
1464 https://doi.org/10.15468/gk6p48

1465 Turner, (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots.
1466 *Journal of Open Source Software*, 3(25), 731. https://doi.org/10.21105/joss.00731.

1467 Vall-llosera, M., Llimona, F., De Cáceres, M., Sales, S., & Sol, D. (2016). Competition, niche
1468 opportunities and the successful invasion of natural habitats. *Biological Invasions*, 18(12), 3535–
1469 3546. https://doi.org/10.1007/s10530-016-1246-7

1470 Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G.,
1471 Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to
1472 gene prediction and phylogenomics. *Molecular biology and evolution*, 35(3), 543–548.
1473 https://doi.org/10.1093/molbev/msx319

1474 Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population
1475 structure. *Evolution*, 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

1476 Whitney, K. D., & Gabler, C. A. (2008). Rapid evolution in introduced species, 'invasive traits' and
1477 recipient communities: challenges for predicting invasive potential. *Diversity and Distributions*,
1478 14(4), 569–580. https://doi.org/10.1111/j.1472-4642.2008.00473.x

1479 Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Springer-Verlag, New York.
1480 https://doi.org/10.1007/978-0-387-98141-3

1481 Wright, S. (1949). The genetical structure of populations. *Annals of eugenics*, 15(1), 323–354.
1482 https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

1483 Yan, L. (2023). ggvenn: draw Venn diagram by 'ggplot2'. R package version 0.1.10

Yang, F., Crossley, M. S., Schrader, L., Dubovskiy, I. M., Wei, S., & Zhang, R. (2022). Polygenic adaptation contributes to the invasive success of the Colorado potato beetle. *Molecular Ecology*, 31(21), 5568–5580. https://doi.org/10.1111/mec.16666

Yang, W., Zhang, Y., Feng, S., Liu, L., & Li, Z. (2018). The first complete mitochondrial genome of the Japanese beetle *Popillia japonica* (Coleoptera: Scarabaeidae) and its phylogenetic implications for the superfamily Scarabaeoidea. *International journal of biological macromolecules*, 118(B), 1406–1413. https://doi.org/10.1016/j.ijbiomac.2018.06.131

Yosida, T. (1949). Chromosome studies in the Coleoptera. IV. Comparison of the chromosomes in five species of the Scarabaeidae. *Transactions of the Sapporo Natural History Society*.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., & Yang, T.-L. (2019). PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, 35(10), 1786–1788. https://doi.org/10.1093/bioinformatics/bty875

Zhu, B.-H., Xiao, J., Xue, W., Xu, G.-C., Sun, M.-Y., & Li, J.-T. (2018). P_RNA_scaffolder: a fast and accurate genome scaffolder using paired-end RNA-sequencing reads. *BMC genomics*, 19, 1–13. https://doi.org/10.1186/s12864-018-4567-3