Received: 26 April 2018 | Revised: 17 January 2019 | Accepted: 17 January 2019

DOI: 10.1111/mec.15030

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Deciphering the worldwide invasion of the Asian long-horned beetle: A recurrent invasion process from the native area together with a bridgehead effect

Marion Javal[1] | Eric Lombaert[2] | Tetyana Tsykun[3] | Claudine Courtin[1] | Carole Kerdelhué[4] | Simone Prospero[3] | Alain Roques[1] | Géraldine Roux[1,5]

[1]INRA UR633 Zoologie Forestière, Orléans Cedex 2, France

[2]INRA, Université Côte d'Azur, CNRS, ISA, Sophia Antipolis, France

[3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[4]CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, Université Montpellier, Montpellier, France

[5]COST, Université d'Orléans, Orléans, France

Correspondence
Marion Javal, INRA UR633 Zoologie Forestière, Orléans Cedex 2, France.
Email: marion.javal@gmail.com

Present Address
Marion Javal, Centre for Invasion Biology, Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

Funding information
Region Centre Val de Loire, Grant/Award Number: 2014 00095480; French Ministry of Agriculture, Food and Forestry DGAL, Grant/Award Number: E07/2014 ; COST Action FP 1401 - Global Warning

Abstract

Retracing introduction routes is crucial for understanding the evolutionary processes involved in an invasion, as well as for highlighting the invasion history of a species at the global scale. The Asian long-horned beetle (ALB) *Anoplophora glabripennis* is a xylophagous pest native to Asia and invasive in North America and Europe. It is responsible for severe losses of urban trees, in both its native and invaded ranges. Based on historical and genetic data, several hypotheses have been formulated concerning its invasion history, including the possibility of multiple introductions from the native zone and secondary dispersal within the invaded areas, but none have been formally tested. In this study, we characterized the genetic structure of ALB in both its native and invaded ranges using microsatellites. In order to test different invasion scenarios, we used an approximate Bayesian "random forest" algorithm together with traditional population genetics approaches. The strong population differentiation observed in the native area was not geographically structured, suggesting complex migration events that were probably human-mediated. Both native and invasive populations had low genetic diversity, but this characteristic did not prevent the success of the ALB invasions. Our results highlight the complexity of invasion pathways for insect pests. Specifically, our findings indicate that invasive species might be repeatedly introduced from their native range, and they emphasize the importance of multiple, human-mediated introductions in successful invasions. Finally, our results demonstrate that invasive species can spread across continents following a bridgehead path, in which an invasive population may have acted as a source for another invasion.

KEYWORDS

Anoplophora glabripennis, approximate Bayesian computation, Asian long-horned beetle, biological invasion, invasion routes, microsatellites, random forest

1 | INTRODUCTION

The number of human-mediated introductions of non-native species has increased considerably over the last decades, and the establishment rate of such species does not seem to be stabilizing (Liebhold, Brockerhoff, & Kimberley, 2017; Seebens et al., 2017). This increase is substantial for insects (Roques et al., 2016), which benefit from the intensification of international trade to settle out of their native range. Harmful consequences of such introductions are diverse and may affect all kinds of ecosystems as some of these species become invasive (Hulme, 2009; Kenis et al., 2008; Simberloff et al., 2013). Identifying patterns of introduction and potential pathways of spread of an invasive species is crucial for preventing further introductions and managing established populations (Essl et al., 2015; Estoup & Guillemaud, 2010). Additionally, retracing invasion routes is essential for understanding the evolutionary and ecological processes that have led to the invasion success (Estoup & Guillemaud, 2010).

Biological invasions can differ in the number of introduction events involved and in the spread pathway of the introduced species. The specificity of invasion scenarios strongly influences the genetic structure of invasive populations and their invasion success (Garnas et al., 2016). Newly founded populations often suffer from a drastic decrease in genetic diversity due to the introduction of a limited number of genotypes from the source population (founder effect; e.g., Auger-Rozenberg et al., 2012). However, multiple introductions seem to be rather common (Dlugosch & Parker, 2008) and have been reported for many invasive insect species (Garnas et al., 2016). They can help to partly restore the original genetic diversity and thus limit the severity of bottlenecks (Dlugosch & Parker, 2008). Nevertheless, single introductions may also occur, like in the case of the Asian yellow-legged hornet (*Vespa velutina*, Arca et al., 2015; Monceau, Bonnard, & Thiéry, 2014). Additionally, bridgehead scenarios, in which an invasive population acts as the source for another invasive population (Hurley et al., 2016; Lombaert et al., 2010; Yang, Sun, Xue, Li, & Hong, 2012), may complicate the invasion history considerably, and such a pattern seems to be much more frequent than previously estimated (Garnas et al., 2016).

Approximate Bayesian computation (ABC) is a method for forming model-based inferences for complex scenarios on the basis of simulations and summary statistics (Beaumont, Zhang, & Balding, 2002; Bertorelle, Benazzo, & Mona, 2010; Csilléry, Blum, Gaggiotti, & Francois, 2010). This method makes it possible to take into account complex demographic processes such as changes in effective population size, admixture events and the involvement of unsampled populations (i.e., ghost populations; Estoup & Guillemaud, 2010). ABC has been used successfully to retrace the invasion routes of numerous alien species (Lesieur et al., 2018; Lippens et al., 2017; Lombaert et al., 2010; Miller et al., 2005; Pascual et al., 2007). The number of simulated data sets increases steadily with the number of tested scenarios, which can lead to time-consuming computations. Additionally, in some cases, estimated posterior probabilities of scenarios based on polychotomous logistic regression give inaccurate predictions despite involving a large number of simulations (Robert, Cornuet, Marin, & Pillai, 2011). To compare scenarios, the use of a "random forest" algorithm (Breiman, 2001) was proposed recently by Pudlo et al. (2016). This application of a "random forest" process in an ABC context (ABC-RF) is expected to discriminate efficiently among scenarios while being computationally less costly, as well as less sensitive to the choice of summary statistics compared to standard ABC. ABC-RF is not routinely used for the reconstruction of invasion routes, but it has already produced some promising results in discriminating among complex invasion scenarios (Fraimout et al., 2017; Lombaert et al., 2018).

In the present study, we aimed to describe the major events of the worldwide invasion history of *Anoplophora glabripennis* (Motschulsky), the Asian long-horned beetle (ALB), with a specific focus on the European invasion. ALB is a xylophagous Cerambycidae native to China and Korea (Lingafelter & Hoebeke, 2002). Invasive populations were reported for the first time in the United States in 1996 (Haack, Cavey, Hoebeke, & Law, 1996), and in Canada in 2003 (Turgeon, Orr, Grant, Wu, & Gasman, 2015). Europe was the second continent to be invaded. Since the first European record in Austria in 2001 (Hérard et al., 2006), additional local outbreaks have been successively observed in nine other European countries, namely France (2003), Germany (2004), Italy (2007), Belgium (2008), the Netherlands (2010), Switzerland (2011), the United Kingdom (2012), Finland (2015) and Montenegro (2015), and new interceptions are frequently reported (EPPO Global Database, 2016). ALB introductions were favoured by the increase in global trade, and larvae may have been transported from their native range, hidden in wooden packing material (Haack, Hérard, Sun, & Turgeon, 2010). Despite ongoing eradication programmes, there are still active invasive ALB populations across Europe (EPPO Global Database, 2016).

Insect larvae damage, and ultimately kill, deciduous tree species (Sjöman, Östberg, & Nilsson, 2014) by boring galleries within trunks. In China, ALB infestations became an issue after implementation of a large-scale reforestation programme in the early 1980s (Pan, 2005). The species is estimated to cause about 10 billion yuan of economic loss per year (USD 1.5 billion), corresponding to 12% of economic losses from forest pests in the country (Hu, Angeli, Schuetz, Luo, & Hajek, 2009). In North America and Europe, *A. glabripennis* poses a significant threat to urban trees. In the United States, for example, its maximum potential impact is estimated at $669 billion, with a loss of almost 35% of the canopy and of 30% tree mortality (Nowak, Pasek, Sequeira, Crane, & Mastro, 2001). Between 1998 and 2006, the United States invested nearly $249 million in eradication programmes (Smith, Turgeon, Groot, & Gasman, 2009).

Previous studies involved investigations of the genetic structure of invasive ALB populations in both the native (Carter, Smith, & Harrison, 2009) and invaded range (Carter, Smith, & Harrison, 2010; Javal, Roques et al., 2017). Their findings highlighted a complex genetic

structure in the native zone that is dominated by a few major genetic clusters. In the invaded area, specimens belonging to these genetic clusters were still the most abundant in the sampled populations. However, most invasive populations appeared genetically distinct, although some geographically close outbreaks seemed genetically identical. These results suggested that multiple introductions followed by secondary movements within the invaded areas might be responsible for the current repartition of the species worldwide. However, this scenario has not been formally tested. The goal of this study was to define the pathways and sequence of main events that have resulted in the current distribution of ALB. Specifically, using population genetics analysis and ABC-RF, we investigated: (a) whether introductions from the native area were recurrent, and (b) whether the invasion in Europe could have originated from the first invaded range, that is North America, by focusing on specific sampled populations.

2 | MATERIALS AND METHODS

2.1 | Sampling and DNA extraction

A total of 633 ALB specimens (i.e., beetles, larvae and eggs) were sampled in the field between 1999 and 2016 in 43 locations worldwide, including 19 localities in Europe, 7 in North America and 17 in the native area (China and Korea; Table 1 and Figure 1). According to Meirmans (2015), particular attention was given to the sampling strategy. Usually, samples are collected when a newly settled population is identified and infested trees are felled, which makes it easy to collect several specimens from a single tree or from trees that are close together. ALB females do not disperse far and tend to lay most of their eggs within the same tree or in the vicinity (Sawyer, 2009). Therefore, the number of sampled trees was maximized in order to avoid collecting closely related individuals from a single family.

From the full set of samples, specimens from four of the North American populations (populations 20–23; Table 1) and from two of the Asian populations (populations 31 and 40) were kept in laboratory colonies of the US Forest Service Quarantine Laboratory in Ansonia, CT (USA). The colonies were reared for two to eight generations (depending on the population) prior to sampling for genetic analysis. The number of specimens used to start each colony is indicated in Table 1.

All specimens were collected alive and preserved in absolute ethanol at −24°C until DNA extraction. Genomic DNA was extracted from a leg fragment for adults and from a small abdominal fragment for immature stages. DNA extraction was performed using a Nucleospin kit (Macherey-Nagel, Düren, Germany) according to the manufacturer instructions, with a final elution volume of 100 µl.

2.2 | Microsatellite genotyping

All samples were genotyped at the 15 microsatellite loci developed for ALB by Carter, Casa, Zeid, Mitchell, and Kresovich (2008). Amplifications were organized into six multiplex reactions as follows: 1. ALB38/ALB77, 2. ALB10/ALB53/ALB59, 3. ALB9/ALB14, 4. ALB15/ALB35/ALB44, 5. ALB40/ALB24 and 6. ALB43/ALB19/ALB30. Multiplexed PCR was performed on a Veriti 96-well Fast Thermal Cycler (Applied Biosystems). Amplifications of the two first multiplexes (ALB38/ALB77 and ALB10/ALB53/ALB59) were run using the following settings: a first denaturation step at 94°C (5 min); 30 cycles at 94°C (30 s), 56°C (45 s) and 72°C (45 s); eight cycles at 94°C (30 s), 53°C (45 s) and 72°C (45 s); and a final elongation step at 72°C for 10 min (Carter et al., 2008). Amplifications of the four last multiplexes (ALB9/ALB14, ALB15/ALB35/ALB44, ALB40/ALB24 and ALB43/ALB19/ALB30) were run with a first denaturation step at 95°C (5 min) followed by 35 cycles at 95°C (1 min), 58°C (1 min) and 72°C (1 min) and a denaturation step at 72°C for 4 min (Carter et al., 2008). One µL of PCR product was denatured in a mix of 10 µl formamide and 0.24 µL 600 Liz marker before being run on an ABI PRISM 3500 sequencer (Life Technologies).

2.3 | Data cleaning and initial population genetics

Alleles were scored using GENEMAPPER V4.1 (Applied Biosystems). Ambiguous genotypes were re-amplified and re-scored once and, if still unclear, were excluded from further analyses. The frequency of null alleles at each locus and for each population was calculated using FREENA (Chapuis & Estoup, 2007). Loci for which the mean estimate of null allele frequency was higher than 10% were discarded. Thus, out of the 15 loci used for the genotyping, 2 were discarded (ALB35 and ALB40) and 13 were kept for the analyses. Expected (H_E) and observed heterozygosity (H_O), as well as the mean number of alleles per locus in a population (A_N), were computed with GENALEX 6.501 (Peakall & Smouse, 2012). The mean allelic richness per locus in each population (A_R) and the private allelic richness were computed by rarefaction using HP-RARE to correct bias due to unequal sample sizes (Kalinowski, 2005). Deviation from Hardy–Weinberg equilibrium for each locus and each population, as well as the fixation index F_{IS}, was calculated using GENEPOP 4.2 (Raymond & Rousset, 1995). The difference between native and invasive F_{IS} was tested using a Student *t* test. Linkage disequilibrium between loci for each population was tested with GENETIX 4.05.2 (Belkhir, Borsa, Chikhi, Raufaste, & Bonhomme, 2004), and a sequential Bonferroni procedure was applied to the results, as suggested by Verhoeven, Simonsen, and McIntyre (2005). All analyses using FREENA, GENALEX, HP-RARE and GENEPOP were conducted on populations with more than 10 sampled individuals (29 populations in total).

2.4 | Population genetic structure and clustering

Pairwise F_{ST} values (Weir & Cockerham, 1984) were calculated in FREENA with or without the Excluding Null Allele (ENA) method. This method corrects for the positive bias induced by the presence of null alleles on F_{ST} estimation and provides accurate estimates of F_{ST} in the presence of null alleles. An exact test for genotypic differentiation was computed with GENEPOP 4.2 for each pair of populations.

The genetic structure of the sampled populations was inferred with the Bayesian model-based cluster analysis implemented in STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) on the global data set, as well as on each continental subset (Asia, North America and Europe). The admixture model with correlated allele frequencies was used, and the geographic origin of populations was not considered as prior. Each run consisted of a burn-in period of 200,000 iterations, followed by 1,000,000 Markov Chain Monte Carlo (MCMC) iterations. Twenty independent runs were performed for each *K* value, the number of genetic clusters, testing from *K* = 1 to 13 for the sub-data sets and from 1 to 30 for the global data set. The most likely values of *K* were determined as suggested by Janes et al. (2017): first, the maximal mean and small standard deviation of the posterior probability of *K* among runs were considered (Pritchard et al., 2000); Δ*K* (Evanno, Regnaut, & Goudet, 2005) was then computed using STRUCTURE HARVESTER (Earl & von Holdt, 2012), and a decision was made considering the alterations of individual assignments with increasing *K* values. CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) was used to check for multimodality of the results given by the 20 runs of STRUCTURE. CLUMPAK is based on a Markov clustering algorithm that enables the generation of consensus solutions for each *K* by grouping runs on the basis of their similarity. For a given value of *K*, runs are either all very similar, resulting in one single consensus solution, or they differ, resulting in one majority mode (i.e., the consensus solution for the majority of the runs) and one or several minority modes (i.e., the consensus solution for the remaining runs). Spatial distribution of the defined genetic clusters in populations of each continent was mapped with ARCGIS 10.4.1.

Individuals were probabilistically assigned to their continent of origin (Asia, North America or Europe) using a discriminant analysis of principal components (DAPC) with the ADEGENET v. 2.0 (Jombart, 2008) package of R v. 3.2.2 (R Core Team, 2015). First, multivariate genetic data were transformed into principal components (PCs) and the optimal number of PCs was determined with cross-validation (Jombart, Devillard, & Balloux, 2010). Thereafter, ALB groups were predefined according to the continent of origin of the specimens and were then plotted along the first two discriminant functions of the analysed PCs. In addition, the DAPC was performed without prior assignments to the geographic groups, so that the number of clusters (*K*) that best describe the data was chosen based on the minimal Bayesian Information Criterion (BIC). Finally, an unrooted Neighbour-Joining (NJ) tree (Saitou & Nei, 1987) was built with the software TREEMAKER using 100,000 bootstrap replicates and the Cavalli-Sforza and Edwards (1967) chord distance.

2.5 | ABC-based inferences about colonization history

Approximate Bayesian computation (ABC; Beaumont et al., 2002) analyses were used to infer the invasion history of ALB. All populations used in ABC scenarios were chosen according to: (a) the number of sampled individuals (between 6 and 42 specimens); (b) for the native populations, their genetic proximity to sampled invasive populations (Table 2; Supporting information Table S1); (c) their heterozygosity level, in order to maximize diversity within samples (Table 1); (d) their geographic location, in order to cover most of the native and invaded ranges (Table 1 and Figure 1). In the native area, populations also had to belong to distinct genetic clusters according to the STRUCTURE membership probabilities and NJ tree results (Figure 2; Supporting information Figure S1).

Based on the genetic structure observed in the native range (see Section 3), three populations from Asia were used as potential sources for the invaded range (populations 28 and 32 in China and population 41 in South Korea). These populations comprised large numbers of specimens, originated from geographically distant areas and were assigned to different genetic clusters (Figure 2; Supporting information Figure S1). This made it possible to cover most of the geographic range in Asia and to consider a large part of the native genetic diversity. North America was the first invaded zone, and one population from that continent was selected (population 21) to estimate its most probable origin and to test for a bridgehead scenario for European invasion. This population was among the first detected in the USA in 1998 (EPPO Global Database, 2016), and it was sampled a year after its detection (Table 1). Even though it was kept in captivity for five generations (Supporting information Table S2), the reared colony was based on a large number of specimens (Table 1) that must have contributed to the preservation of the genetic diversity of the population. The other North American populations were not included in the analysis, either because they had too few samples or, in cases where the population was sampled in a laboratory colony, because too few individuals were used to initially form the colony. The focus was set on European invasion history, and four populations were chosen, that is Gien (France, population 19; first recorded in 2003), Marly (Switzerland, population 12; first recorded in 2014), Divonne-les-Bains (France, population 9; first recorded in 2016) and Furiani (France, Corsica, population 1; first recorded in 2015). These European populations were chosen based on field observations of their possible establishment and the substantial duration of their presence in Europe. They were assigned to different genetic clusters according to the STRUCTURE analysis (Figure 2) and were also scattered in the NJ tree (Supporting information Figure S1). When it was impossible to avoid the use of laboratory-reared populations, populations with the largest number of founder individuals sampled in the natural population were selected.

Six nested sets of invasion scenarios were analysed sequentially: after each analysis, a new population was added to the most likely scenario (Supporting information Figure S2), following the methodology described by Fraimout et al. (2017). The scenario topologies were based on the dates of first observation. However, as the date of first observation does not necessarily correspond to the actual date of introduction, a five-year uncertainty (Favaro, Battisti, & Faccoli, 2013) was included by the choice of priors (*ti* in Supporting information Table S3). In all scenarios, the three native populations were derived from an ancestral unsampled population. Each introduction event was followed by a bottleneck that ranged in duration from 0

TABLE 1 Origin of the *Anoplophora glabripennis* samples analysed in this study and the main genetic diversity parameters of the single populations

N pop	ID	Country	City/State	Year of collection (year of detection)	Long	Lat	Sample size	A_N	A_R	H_O	H_E	F_{IS}
Asia												
27	Harbin	China	Harbin	2016	126.53	45.80	15	2.31	2.06	0.29	0.36	0.23
28	Yanji	China	Yanji	2016	129.51	42.89	26	3.08	2.34	0.30	0.41	0.28
29	Chengd	China	Chengde	2016	117.96	40.95	26	3.46	2.54	0.35	0.43	0.20
30	Shijia	China	Shijiazhuang	2015	114.51	38.04	26	3.38	2.51	0.32	0.44	0.29
31	Huhhot	China	Huhhot	2001	111.63	40.80	15 (138)	2.31	2.09	0.29	0.36	0.22
32	Qingto	China	Qingtongxia	2015	106.08	38.02	25	3.38	2.59	0.32	0.45	0.30
33	Yanchi	China	Yanchi	2015	107.41	37.78	25	3.31	2.55	0.35	0.45	0.23
34	Jinan	China	Jinan	2016	117.12	36.65	11	2.62	2.26	0.38	0.38	0.04
35	Taian	China	Taian	2016	117.09	36.20	11	3.15	2.73	0.27	0.47	0.45
36	Hefei	China		2016	117.23	31.82	15	3.15	2.66	0.30	0.46	0.38
37	Cixi	China	Cixi	2015	121.26	30.16	20	3.00	2.3	0.26	0.38	0.34
38	Yunnan	China		2016	102.71	25.05	14	2.92	2.77	0.30	0.44	0.38
39	Tongli	China	Tongliao	2015	122.24	43.65	24	3.00	2.18	0.28	0.38	0.27
40	Kangwo	South Korea	Kangwon Nat Park	2001	128.42	38.13	14 (11)	1.23	1.1	0.01	0.02	0.50
41	Pocheo	South Korea	Pocheon	2015	127.20	37.89	6	–	–	–	–	–
42	Incheo	South Korea	Incheon	2013/2014	126.70	37.45	1	–	–	–	–	–
43	Ulsan	South Korea	Ulsan	2015	129.31	35.54	3	–	–	–	–	–
Mean value							16.29	2.88	2.33	0.29	0.39	0.29
North America												
20	Baysid	USA	New York City	1999 (1996)	−73.77	40.76	15 (332)	2.23	1.98	0.31	0.33	0.11
21	Ravens	USA	Illinois	1999 (1998)	−87.68	41.97	15 (1,064)	2.62	2.26	0.35	0.41	0.16
22	Worces	USA	Massachusetts	2009 (2008)	−71.80	42.31	15 (42)	1.85	1.8	0.29	0.30	0.07
23	Bethel	USA	Bethel, Ohio	2011 (2011)	−84.08	38.97	15 (48)	1.92	1.72	0.25	0.22	−0.13
24	Northv	Canada	Ontario	2005 (2003)	−79.55	43.79	4	–	–	–	–	–
25	Shepph	Canada	Ontario	2007 (2003)	−79.87	44.17	4	–	–	–	–	–
26	Tallgr	Canada	Ontario	2005 (2003)	−79.57	45.27	11	1.46	1.45	0.24	0.18	−0.31
Mean value							10.67	2.39	2.05	0.26	0.32	0.19
Europe												
1	Arenau	France	Furiani	2015 (2013)	9.43	42.68	20	1.92	1.82	0.23	0.28	0.19
2	Colast	France	Furiani	2016 (2013)	9.43	42.66	20	1.69	1.63	0.21	0.24	0.15

(Continues)

TABLE 1 (Continued)

N pop	ID	Country	City/State	Year of collection (year of detection)	Long	Lat	Sample size	A_N	A_R	H_O	H_E	F_{IS}
3	Conouv	France	Furiani	2015 (2013)	9.41	42.66	5	1.23	–	–	–	–
4	Costad	France	Furiani	2015 (2013)	9.44	42.65	6	1.92	–	–	–	–
5	Mcarlo	France	Furiani	2014 (2013)	9.43	42.65	7	–	–	–	–	–
6	Cornud	Italy	Cornuda	2014 (2009)	12.00	45.83	10	1.69	1.56	0.20	0.19	−0.02
7	Rapagn	Italy	Rapagnano	2015 (2013)	13.59	43.16	2	–	–	–	–	–
8	Austri	Austria	Gallspach	2014 (2013)	13.80	48.21	4	–	–	–	–	–
9	Divonn	France	Divonne les bains	2016 (2016)	6.14	46.36	42	2.69	1.98	0.26	0.32	0.20
10	Beriko	Switzerland	Berikon	2015 (2015)	8.37	47.35	21	1.85	1.75	0.19	0.28	0.33
11	Brunis	Switzerland	Brünisried	2013/2014 (2011)	7.28	46.76	23	1.92	1.67	0.18	0.22	0.20
12	Marly	Switzerland	Marly	2014 (2014)	7.14	46.80	25	2.23	1.87	0.25	0.28	0.11
13	Winter	Switzerland	Winterthur	2012 (2012)	8.52	47.39	32	1.85	1.67	0.25	0.25	0.00
14	Ebersb	Germany	Ebersberg		11.97	48.08	1	–	–	–	–	–
15	Feldk	Germany	Feldkirchen	(2012)	11.73	48.15	5	–	–	–	–	–
16	Neubib	Germany	Neubiberg	(2014)	11.67	48.07	12	2.15	1.88	0.18	0.26	0.32
17	Strasb	France	Strasbourg	2010/2011/2014 (2008)	7.75	48.57	7	–	–	–	–	–
18	Stann	France	St-Anne/Brivet	2004/2005 (2004)	−2.00	47.45	9	–	–	–	–	–
19	Gien	France	Gien	2015 (2003)	2.62	47.69	25	2.15	1.54	0.14	0.28	0.28
Mean value							14.53	1.97	1.74	0.21	0.26	0.18

Notes. Sample size: for bred specimens, the number of individuals used to found the colony is given in brackets. A_N, mean number of alleles per locus; A_R, mean allelic richness over loci; F_{IS}, fixation index; H_O, observed heterozygosity; H_E, expected heterozygosity.

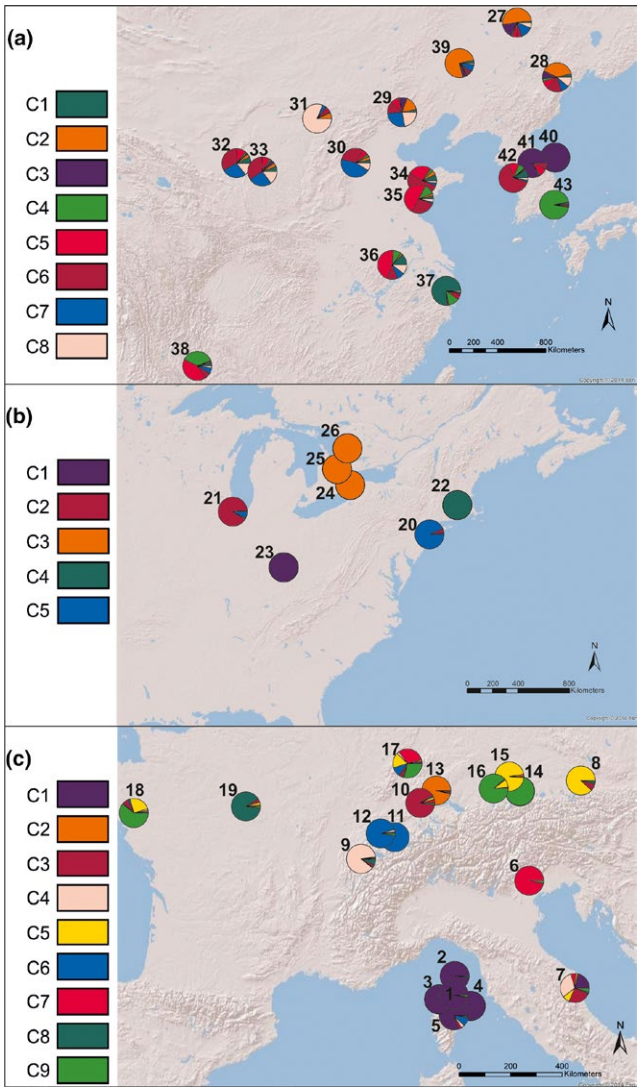

FIGURE 1 Distribution of ALB populations and their defined genetic clusters inferred by STRUCTURE analysis within each sampled continent: in Asia (a), North America (b) and Europe (c), for each most relevant value of *K*. The numbers correspond to populations used in the study

to 10 years (Supporting information Table S3). In our ABC analyses, historical, demographic and mutational parameter values for simulations were drawn from prior distributions, as described in Supporting information Table S3, and various conditions were set in order to avoid impossible topologies; for example, the introduction event of an invasive population could not be anterior to the formation of its source population. Generation time was set at one generation per year, as suggested by Meng, Hoover, and Keena (2015). As proposed by Lombaert et al. (2011), substructuring was simulated in the native area in order to consider the risk that the true source population was not sampled precisely using ghost populations. The scenarios are described in detail in Supporting information Figure S2.

The first analysis (Analysis 1; Supporting information Figure S2) questioned the native origin of the North American population (population 21). Once the most likely scenario was inferred, the population from the oldest sampled outbreak in Europe (population 19: Gien, France) was added for the second ABC analysis (Analysis 2; Supporting information Figure S2). This population could have originated either from one of the three native populations in Asia or from the invasive North American population. Similarly, population 12 (Marly, Switzerland; detected in 2014) was added in the third ABC analysis (Analysis 3; Supporting information Figure S2). Even though this population was officially detected in 2014, there is indirect evidence that it is actually the source population for population 11 (Brünisried, Switzerland; detected in 2011). Indeed, the EPPO report states that "trace-back studies revealed that during winter 2010/2011 firewood was transferred from Marly to Brünisried where adult beetles were subsequently captured in autumn 2011" (EPPO Global Database). Our results support this hypothesis because the two populations are strongly related (the population differentiation is not significant (Supporting information Table S1), the STRUCTURE results indicate that the two populations cluster together for all *K* values (Figure 2), and the F_{ST} between the two populations is low (Table 2; Supporting information Table S4)). Therefore, for the construction of our scenarios we assumed that population 12 was introduced at the latest in 2011. The two most recently invasive European populations (population 1 in Furiani, France and population 9 in Divonne-les-Bains, France) were added independently (Analyses 4a and 4b; Supporting information Figure S2) because their dates of first detection were very close together. As several years may be necessary to detect a new ALB population (Favaro et al., 2013), it was indeed not possible to formally determine which of these most recent invasions occurred first. The aim of the final ABC analysis (Analysis 5) was therefore to check whether these two European populations (1 and 9) were introduced independently. Moreover, it tested the independence of the introductions of populations 12 (Marly, Switzerland) and 1 (Corsica, France) (Supporting information Figure S2). Indeed, as these last two populations were assumed to have been introduced four years apart, they did not meet our five-year uncertainty criterion.

Although admixture in the recent invasion history of this species with a low dispersal ability can only happen under very specific conditions (human-mediated dispersal or geographically close outbreaks), we carried out an additional set of analyses that included the same scenarios and populations as described above, as well as admixture events between all possible pairs of source populations.

Summary statistics used in each ABC analysis included one-sample summary statistics (the mean number of alleles per locus, the mean genic diversity, the mean size variance and the mean Garza-Williamson's M; Garza & Williamson, 2001), two-sample summary statistics (the mean number of alleles per locus, the mean genic diversity, the mean size variance, the pairwise F_{ST}, the classification index, the shared allele distance and the $(d\mu)^2$ distance), and an admixture summary statistic (the maximum likelihood; Choisy, Franck, & Cornuet, 2004). A random forest algorithm was used to compare the scenarios, as described by Pudlo et al. (2016). This method enables efficient discrimination among models and gives an estimate of posterior probabilities of the best model,

TABLE 2 Pairwise F_{ST} values between the populations of ALB (only populations with 10 or more sampled specimens were considered) with the ENA correction

		Europe										North America				Asia												
		Arenau (1)	Colast (2)	Cornud (3)	Divonn (9)	Beriko (10)	Brunis (11)	Marly (12)	Winter (13)	Neubib (16)	Gien (19)	Baysid (20)	Ravens (21)	Worces (22)	Bethel (23)	Tallgr (26)	Harbin (27)	Yanji (28)	Chengd (29)	Shijia (30)	Huhhot (31)	Qingto (32)	Yanchi (33)	Jinan (34)	Taian (35)	Hefei (36)	Cixi (37)	Yunnan (38)	Tongli (39)
Europe	Colast (2)	0.056397																										
	Cornud (3)	0.580045	0.600842																									
	Divonn (9)	0.287529	0.296826	0.554968																								
	Beriko (10)	0.476213	0.450542	0.452793	0.46349																							
	Brunis (11)	0.494459	0.491014	0.611112	0.376153	0.545884																						
	Marly (12)	0.420504	0.413938	0.540209	0.313404	0.496926	0.078349																					
	Winter (13)	0.388649	0.34253	0.582101	0.435652	0.347695	0.55154	0.495437																				
	Neubib (16)	0.34532	0.371458	0.525169	0.376384	0.336639	0.544309	0.451997	0.386058																			
	Gien (19)	0.528659	0.55737	0.616688	0.465321	0.400402	0.643121	0.585466	0.529805	0.387867																		
North America	Baysid (20)	0.441574	0.47677	0.486644	0.393829	0.385224	0.515453	0.457247	0.488082	0.399458	0.408902	0.254416																
	Ravens (21)	0.231371	0.261443	0.42486	0.226996	0.285164	0.455966	0.385739	0.324424	0.18463	0.250032	0.375121	0.295779															
	Worces (22)	0.454825	0.481199	0.537783	0.401529	0.410853	0.472733	0.408866	0.484475	0.30275	0.466057	0.407362	0.310683	0.450345														
	Bethel (23)	0.408366	0.4411	0.623822	0.382377	0.474359	0.484298	0.408649	0.479525	0.458407	0.545283	0.534928	0.387238	0.497375	0.515571													
	Tallgr (26)	0.54989	0.561929	0.689173	0.527334	0.448459	0.647345	0.585406	0.501105	0.529161	0.628495																	
Asia	Harbin (27)	0.258312	0.25627	0.501389	0.195004	0.399489	0.396889	0.283268	0.351848	0.317177	0.48574	0.348499	0.219613	0.392072	0.403022	0.477746												
	Yanji (28)	0.204831	0.185846	0.388229	0.178388	0.249042	0.362303	0.279722	0.246451	0.180214	0.33487	0.257623	0.109256	0.297087	0.310421	0.354195	0.029444											
	Chengd (29)	0.225585	0.249208	0.446277	0.113397	0.348159	0.324165	0.253902	0.328356	0.260269	0.349142	0.24551	0.136711	0.335172	0.245018	0.427856	0.092058	0.077796										
	Shijia (30)	0.308506	0.31787	0.400482	0.187201	0.240056	0.368545	0.305172	0.309472	0.197278	0.277463	0.246315	0.121102	0.236247	0.253098	0.340323	0.195711	0.112947	0.099183									
	Huhhot (31)	0.214528	0.229887	0.478523	0.294083	0.356645	0.447529	0.361712	0.308538	0.31488	0.383156	0.274811	0.135556	0.390359	0.300184	0.467672	0.170244	0.125677	0.097629	0.185092								
	Qingto (32)	0.245954	0.256174	0.362343	0.166351	0.218084	0.313905	0.243916	0.26713	0.165432	0.269287	0.226403	0.087513	0.226363	0.212411	0.3405	0.155641	0.068135	0.072462	0.000558	0.138715							
	Yanchi (33)	0.220217	0.234875	0.378656	0.159073	0.242928	0.326488	0.253243	0.264853	0.151224	0.255865	0.225418	0.064427	0.224439	0.229706	0.349525	0.12391	0.048402	0.046388	0.017443	0.108805	-0.00351						
	Jinan (34)	0.255195	0.295169	0.458164	0.254509	0.323458	0.3726	0.298408	0.346359	0.19807	0.395412	0.294488	0.139649	0.253501	0.196417	0.357009	0.236516	0.148024	0.156767	0.104248	0.213395	0.068041	0.092462					
	Taian (35)	0.227521	0.270378	0.383308	0.233078	0.29149	0.331176	0.260508	0.29934	0.216408	0.395218	0.26707	0.129275	0.189292	0.158732	0.317847	0.197666	0.126353	0.125657	0.09932	0.163553	0.052032	0.06887	0.046276				
	Hefei (36)	0.275029	0.283334	0.332715	0.268585	0.193543	0.354414	0.293063	0.291881	0.179485	0.2782	0.198066	0.108729	0.163726	0.225702	0.288747	0.2156	0.117713	0.155025	0.07517	0.162402	0.049619	0.071514	0.067842	0.035647			
	Cixi (37)	0.38346	0.42826	0.433005	0.342501	0.378661	0.502736	0.440981	0.469865	0.321556	0.46361	0.406226	0.190478	0.347796	0.462957	0.474999	0.332791	0.248762	0.290094	0.239588	0.346831	0.203249	0.22114	0.272631	0.221481	0.175952		
	Yunnan (38)	0.285515	0.307437	0.315215	0.261229	0.240849	0.335376	0.26054	0.322692	0.214435	0.320015	0.222356	0.162368	0.206658	0.252799	0.377375	0.173076	0.117513	0.127562	0.129056	0.165783	0.090647	0.095196	0.133824	0.045468	0.041849	0.213777	
	Tongli (39)	0.283599	0.292011	0.512449	0.210075	0.398353	0.433297	0.370467	0.366965	0.354707	0.487911	0.293663	0.265198	0.405302	0.382841	0.486894	0.117736	0.127398	0.132891	0.248084	0.229122	0.206769	0.205127	0.237061	0.208096	0.246644	0.375346	0.1982
	Kangwo (40)	0.686198	0.727893	0.867434	0.507034	0.761931	0.64862	0.554249	0.742955	0.788407	0.829845	0.692514	0.644162	0.739713	0.79313	0.879707	0.476968	0.483093	0.437489	0.552086	0.636688	0.52625	0.503178	0.699539	0.6118	0.622733	0.669356	0.578209	0.503813

Notes. Bold and framed values indicate populations that were not significantly different ($p > 0.001$). F_{ST} values between 0 and 0.15 are marked in green, and values between 0.5 and 1 are marked in red. Numbers in brackets indicate the population number.

while being computationally less intensive than polychotomous logistic regression of ABC algorithms. Briefly, it creates a "forest" of bootstrapped decision trees to classify scenarios on the basis of the summary statistics of the simulated data sets. Some simulations are not used to build the trees and can thus be used to cross-validate the analysis by computing a "prior error rate." The robustness of this method for reconstructing invasion routes has been demonstrated in previous studies (Fraimout et al., 2017; Pudlo et al., 2016).

A total of 20,000 microsatellite data sets were simulated for each scenario using DIYABC 2.1.0 (Cornuet et al., 2008). The relevance of prior settings was checked by comparing the distributions of simulated statistics and the observed values. Based on the summary statistics of the simulated data sets, 1,000 decision trees were constructed using the R package ABCRF (v1.1.0; Pudlo et al., 2016), and the observed data set was submitted to each of the trees. The analysis then provided a classification vote for each scenario, which stands for the number of times the considered scenario is selected among the 1,000 decision trees. The scenario with the highest classification vote was selected as the most likely scenario. The posterior probability of this scenario was then estimated based on a second random forest procedure of 1,000 trees, as described by Pudlo et al. (2016).

Two final steps were conducted to evaluate the robustness of the ABC scenario choice. First, the prior error rate was computed based on the out-of-bag simulations. Using DIYABC, 500,000 data sets were re-simulated, using the most probable model according to the ABC-RF analysis. The resulting observed values of summary statistics were compared to 1,000 simulated values from data sets for which parameters were drawn from their posterior distribution (posterior model checking). Finally, posterior distributions of parameters were estimated for the final selected scenario using 10,000 simulated data sets under a regression by random forest methodology (Raynal et al. 2018, arXiv), with classification forests of 1,000 trees.

Second, all ABC analyses were performed a second time with another set of native sample sites (the alternative data set; Supporting information Table S2), representative of the genetic clusters (Figure 2; Supporting information Figure S1), that corresponded to the first selected sample sites (the core data set), as suggested by Lombaert et al. (2014). In most cases, it was impossible to choose alternative samples for the invasive populations. In the USA, each population was assigned to a specific cluster in the STRUCTURE analysis (Figure 2), which was confirmed by the NJ tree (Supporting information Figure S1). In Europe, only two out of the four populations could have replaced the ones used in the core data set (population 1 belonged to the same cluster as populations 2 to 5, and population 12 grouped with population 11; Figure 2; Supporting information Figure S1).

Finally, posterior distributions of parameters were estimated for the most probable scenario. Durations are given in number of generations, and all parameters are summarized in Supporting information Table S5.

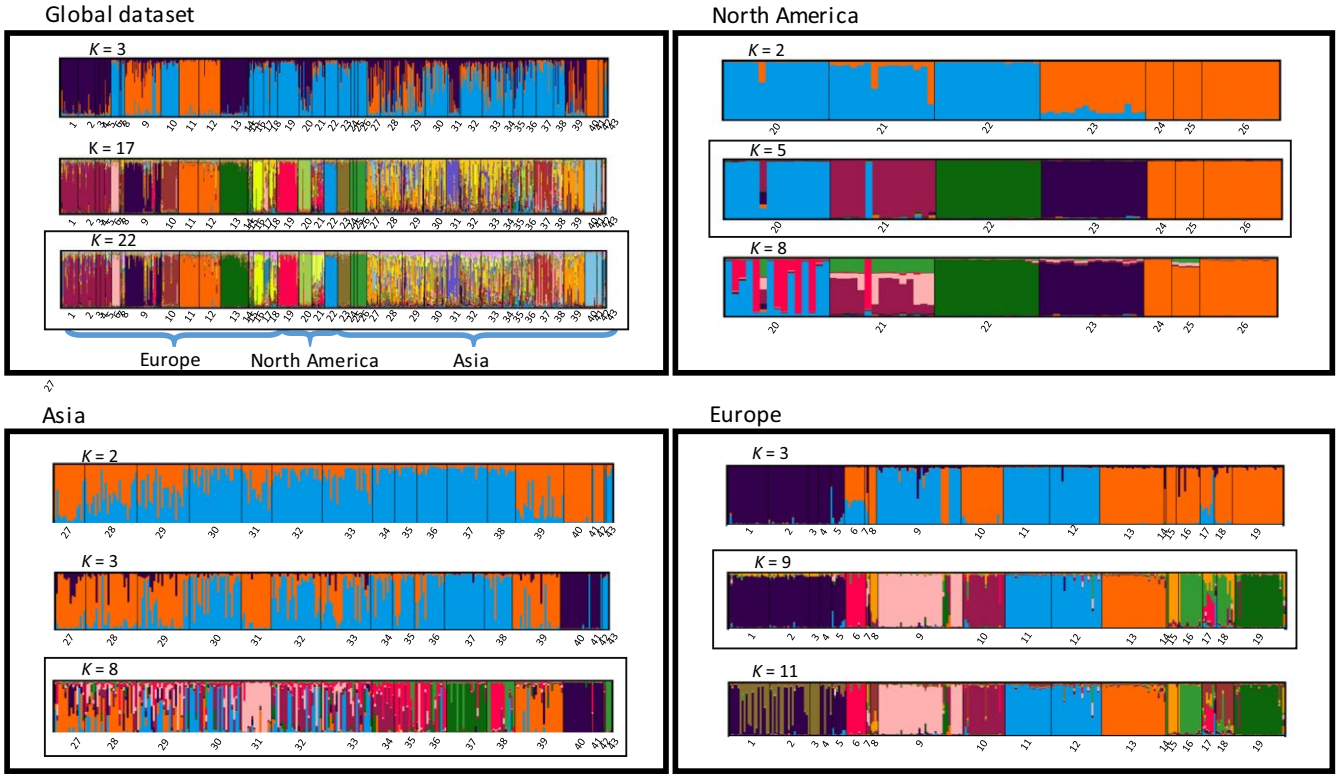

FIGURE 2 Representation of STRUCTURE membership probabilities of individuals belonging to a specific cluster (marked by different colours). Only the *K* values having the highest Δ*K* values are shown. For each sub-data set, the barplot at the most likely *K* value is framed in black, the others correspond to the other highest values of Δ*K*. Various simulations for each *K* value were summarized using CLUMPAK (only the major mode is shown)

3 | RESULTS

3.1 | Preliminary analyses

The expected heterozygosity ranged from 0.071 to 0.406 for invasive populations and from 0.021 to 0.465 for native populations. Average F_{IS} values were higher in the native area than in the two invaded continents (Table 1, $p < 0.01$). A heterozygote deficit was detected in 22 populations, including 13 in Asia, 2 in North America and 7 in Europe, and all loci but three (ALB 10, ALB 19 and ALB 44) showed significant deviation from Hardy–Weinberg equilibrium. The corrected allelic richness for all loci and per population ranged from 1.45 to 2.26 in the invasive populations and from 1.10 to 2.77 in the native populations. A total of 19 populations (12 Asian, 2 North American and 5 European) showed private alleles, with a richness ranging from 0.01 to 0.26 after rarefaction. Linkage disequilibrium was observed for three pairs of loci, but statistical support for their nonrandom associations was not consistent across populations. Detailed summary statistics for each population are listed in Table 1.

3.2 | Population genetic structure and clustering

Pairwise F_{ST} values were high overall but were lower within the native range than within the invaded range (Table 2). Pairwise F_{ST} values estimated with (Table 2) or without (Supporting information Table S4) correcting for the presence of null alleles were similar. Population pairs showing the lowest pairwise F_{ST} values (<0.15) always included at least one Asian population (Table 2), except for two Swiss populations (populations 11 and 12) and two Corsican populations (populations 1 and 2). Population differentiation was statistically significant ($p < 0.05$) for 820 out of 903 (~90%) of the pairwise comparisons (nonsignificant results are given in Supporting information Table S1). However, the cluster analysis of the DAPC also clearly showed that specimens from Asia largely overlapped with North American and European specimens, confirming the genetic proximity of the populations from the different continents (Supporting information Figure S3).

The NJ tree (Supporting information Figure S1) revealed a weak pattern of association between the clustering of populations and their geographic origin, and the majority of bootstrap values were less than 50. In the native zone, populations from NorthEast China grouped together (populations 27, 28, 29 and 39). Strong bootstrap support (72%) was observed for a common clade for two geographically close populations from Switzerland (populations 11 and 12) and for South Korean populations (40 and 41). The Corsican populations (1 to 5) also grouped together.

Replicate STRUCTURE runs led to one major solution for all *K* values and to minor solutions for values from *K* = 3 (i.e., genuine multimodality; Jakobsson & Rosenberg, 2007; Supporting information Figure S4). The high likelihood, but still small standard deviation, of the posterior probability of *K* among runs increased up to the maximum tested number of

clusters in the global ALB population data set (Supporting information Figure S5). The first largest difference in log-likelihood among different *K* numbers (i.e., Δ*K* peak) revealed that the strongest signal of genetic subdivision in the global population of ALB was already apparent at *K* = 3. However, analysis of alterations in individual assignment probabilities with increasing *K* values suggested that there were numerous nested clusters within each continent, which in turn made it difficult to determine the number of genetic clusters for the global ALB population (Figure 2a). The STRUCTURE analysis for each continent separately (Figure 2; Janes et al., 2017) showed that in the Asian sub-data set, the high mean likelihood and small standard deviation of the posterior probability of *K* among ran up to the maximum tested cluster number (13) in our data. However, meaningful assignments to clusters with increasing *K* values were observed up to *K* = 11, whereas higher cluster numbers showed rather partition of probabilities among different individuals and no confident representation of new cluster by at least one whole individual (Supporting information Figure S4a). In addition, considering the second highest difference of log-likelihood among different *K* values (Δ*K* peak at *K* = 8) we assumed presence of 8 clusters in Asian populations. In the North American sub-data set, the maximum likelihood and Δ*K* was at *K* = 5. In addition, alterations of individual assignment probabilities showed that no individual was fully assigned to additional cluster with increasing *K* values above this number. Thus, we assumed presence of 5 genetic clusters in North American populations. Finally, the high mean likelihood with small standard deviation among runs and pattern of individual assignments with increasing *K* in the European data set showed meaningful clustering in some populations, that is in Corsican populations (France, from 1 to 5) and in Marly and Brünisried (Switzerland, populations 11 and 12), up to 12 nested clusters in Europe. However, considering the highest difference of log-likelihood among different *K* (Δ*K* peak at *K* = 9) and in order not to overestimate the presence of nested clusters, we assumed 9 genetic clusters were reasonable to describe the genetic structure in European populations. For each continent, we further sampled from these genetic clusters three representative populations from Asia, one from North America and four from Europe to be included in the ABC analyses. Similar to the STRUCTURE analysis, the DAPC analysis showed consistently low BIC values for cluster numbers higher than *K* = 22 (Supporting information Figure S6). The pattern of probabilistic assignments to the defined clusters was overall congruent with STRUCTURE results and confirmed a complex genetic population structure of the global ALB population.

In the native area, except for differentiation along a north–south gradient between the northeastern populations (South Korean populations 40 and 41, and to a lesser extent Chinese populations 27, 28, 29 and 39) and all other populations supported by the STRUCTURE analyses, no clear pattern of geographical distribution of the genotypes could be identified (Figure 2; Supporting information Figure S4). In North America, the STRUCTURE analysis assigned the three Canadian populations (24, 25 and 26) to a single cluster, regardless of the level of *K*. In Europe, the same analyses revealed that some geographically close populations were interconnected, specifically the two Swiss populations in Marly and Brünisried (11 and 12) and the five Corsican populations (1 to 5). Notably, all but four individuals from the French population in Divonne-les-Bains (9) were grouped together in a single cluster, with high posterior probabilities (Figure 2; Supporting information Figure S4).

3.3 | Invasion scenarios

Comparisons of the distribution of simulated summary statistics with values from the observed core data set showed that the combination of scenarios and priors that we chose was relevant: for the six analyses, between 3.4% and 7.1% of the observed statistics were in the tails of the probability distribution of statistics calculated from prior simulations (at a 5% significance threshold). Random forest votes, posterior probabilities for the most likely scenarios and details of the ABC-RF analysis results are given in Table 3 (and in Supporting information Table S6 for scenarios including admixture). The consistency between results based on the core and alternative data sets for all but one analysis, as well as the moderate prior error rates and the large differences in terms of RF votes (Table 3), indicates a reasonable level of confidence in the choice of scenarios. Note that high F_{IS} values should not challenge ABC analyses, as demonstrated by Lippens et al. (2017).

Analysis 1, which investigated the origin of the North American population (population 21), clearly indicated that the species most likely originated from NorthCentral China (NC China in Table 3, population 32). Analysis 2 suggested that the oldest sampled invasive French population (Gien, detected in 2003) most likely originated from North America. The Swiss population in Marly (population 12), which was detected in 2014 but was probably introduced in 2011 or earlier (Section 2.5; EPPO Global Database, 2016), likely has a Korean origin; this finding was strongly supported by NJ clustering results (Supporting information Figure S1). Analyses 4a and 4b, which focused on two very recently identified French populations (1 and 9), suggested that they both originated from a native population in Asia. Analysis 5 showed that these two new populations were probably introduced independently, as were the Corsican and the Swiss populations (populations 1 and 12). Population 9 (Divonne-les-Bains, France) may have originated from NorthEast China, whereas the Corsican populations, depending on the data set, were found to be connected to either NorthEast China (core data set) or NorthCentral China (alternative data set). The posterior model checking performed on the final scenario (i.e., the one from analysis 5) indicated that the chosen model and posteriors fitted the observed genetic data closely. None of the 424 observed values of summary statistics were within the tail of the probability distribution of statistics calculated from the posterior simulation at a 5% significance threshold (Supporting information Table S7 and Figure S7). The global invasion routes inferred from our ABC-RF analyses are summarized in Figure 3.

The set of analyses that included admixture events yielded the same results for the ABC analyses 1, 2, 3 and 5 (Supporting information Table S6). In the cases of analyses 4a and 4b, admixture scenarios were selected but posterior probabilities were low and associated with markedly strong prior error rates; this finding was further confirmed by discrepancies in the results depending on the data set used. In these analyses 4a and 4b, it is noteworthy that the

TABLE 3 Results of the ABC-RF analyses

Analysis	Target population	Potential source areas	Number of compared scenarios	Number of summary statistics	Random forest votes		Prior error rates		Most probable origin		Posterior probability of the selected scenario	
					Core data set	Alternative data set	Core data set	Alternative data set	Core data set	Alternative data set	Core data set	Alternative data set
1	21 (Ravens, USA)	NE China	3	76	174	72	30.995	34.437	NC China	NC China	0.726	0.834
		NC China			**733**	**838**						
		Korea			93	90						
2	19 (Gien, France)	USA (21)	4	130	**796**	**610**	25.711	28.411	USA	USA	0.734	0.645
		NE China			45	92						
		NC China			111	208						
		Korea			48	90						
3	12 (Marly, Switzerland)	USA (21)	5	204	7	10	23.646	25.425	Korea	Korea	0.823	0.691
		NE China			60	146						
		NC China			121	135						
		Korea			**812**	**708**						
		Gien (19)			0	1						
4a	1 (Furiani, Corsica, France)	USA (21)	6	301	121	151	22.453	23.823	NE China	NC China	0.674	0.651
		NE China			**447**	196						
		NC China			268	**355**						
		Korea			131	248						
		Gien (19)			9	13						
		Switzerland (12)			24	37						
4b	9 (Divonne, France)	USA (21)	6	301	29	30	18.513	19.963	NE China	NE China	0.670	0.751
		NE China			**411**	**439**						
		NC China			346	426						
		Korea			196	90						
		Gien (19)			2	0						
		Switzerland (12)			16	15						
5	12 (Switzerland) versus 1 (Corsica) versus 9 (Divonne)	Divonne (9), Switzerland (12) and Corsica (1) were introduced independently	4	424	**642**	**483**	30.174	15.739	Independent introductions	Independent introductions	0.653	0.674
		Divonne (9) comes from Corsica (1)			99	254						
		Corsica (1) comes from Divonne (9)			214	204						
		Switzerland (12) comes from Corsica (1)			45	59						

Notes. Results are provided for both the core and the alternative data set but are only given for the most probable scenario of each analysis. NE China represents NorthEast China (populations 28 and 27 for the core and the alternative data set, respectively) and NC China represents NorthCentral China (population 32 for the core data set and population 33 for the alternative data set). For Korea, population 41 was used for the core data set and population 43 for the alternative data set.
Random forest votes are given in bold.

parental populations always included at least one population from the Asian native range, which is consistent with the results without admixture. Given the very high prior error rates of all these analyses (from 54.1% to 64.4%), and considering their general agreement with results obtained from analyses without admixture, we will not discuss these results further.

Estimations of posterior distributions of parameters for the most probable scenario are given in Supporting information Table S5. The mean time for all three native populations to merge with their native ancestral population was between 145 and 161 generations. The estimated mean effective size was comparable across native and invasive populations (between 2,235 and 2,813 individuals), except for the Korean population, for which mean effective size was estimated at 221 specimens (population 41, Pocheon). All five invasive populations experienced a mean bottleneck of 4.5 to 6.9 generations, with a mean effective number of founders ranging from 5 specimens (population 19, Gien and population 1, Corsica) to 30 specimens (population 12, Marly). Bottleneck severity (i.e., [duration of bottleneck × parental population size]/effective number of founders) ranged from 114 (Marly, Switzerland, population 12) to 4,422 (Divonne, France, population 9).

4 | DISCUSSION

The global invasion history of the Asian long-horned beetle *A. glabripennis* shows one of the possible patterns of invasive quarantine pest spread across continents. Our study is the first to shed light on the worldwide invasion scenario of ALB by analysing populations from both native and invaded ranges. We used recently developed ABC-RF procedures to retrace the main invasion routes, by analysing data in the light of field observations and previous phylogeographic studies (Carter et al., 2010; Javal, Roques et al., 2017). Although we could not infer the specific history of each population with sufficient statistical support, we were able to provide an overall outline of the global invasion scenario. Our findings suggest the occurrence of multiple introductions from the native range coupled with a bridgehead event, both contributing to the worldwide spread of ALB.

4.1 | Genetic structure in the native range

In the native range of ALB (mainly China and Korea), population genetics analyses revealed a rather low diversity within populations and a strong differentiation among most of them. Clustering analyses indicated differentiation along a north–south gradient between the northeastern populations (27–29, 39–41) and all other populations. This geographic structuring was highlighted in several other studies on ALB in its native range (An et al., 2004; Carter et al., 2009; Javal, Roques et al., 2017), and it might have resulted from the low natural dispersal capability of ALB adults (Smith, Bancroft, Li, Gao, & Teale, 2001; Smith, Tobin, Bancroft, Li, & Gao, 2004). However, demographic events and human-mediated translocations seem to have attenuated the putative global ancestral structure. Indeed, no clear geographic pattern was identified in the STRUCTURE analysis, although F_{IS} values were particularly high, indicating a high level of inbreeding in the native populations. This blurred geographic genetic structure is uncommon for populations in their native ranges, whose structure is usually shaped by multiple factors that constrain gene flow, such as biogeography and natural barriers (Goldberg & Lande, 2007; Hewitt, 2000). In previous studies, this particular genetic structure has been considered to reflect human-mediated population translocations at a very large scale, mainly across China (Carter et al., 2009; Javal, Roques et al., 2017). The considerable reforestation efforts undertaken by the Chinese government in the early 1980s resulted in the movement and plantation, mainly in the northern part of the country (Pan, 2005), of trees (mainly *Populus* spp.) that were possibly

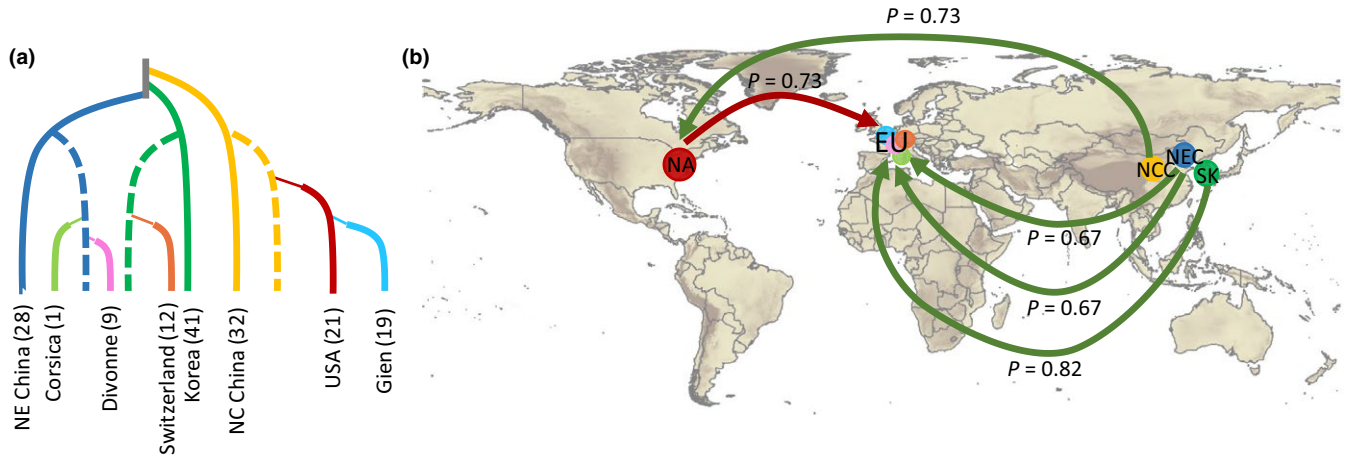

FIGURE 3 (a) Schematic representation of the most probable final scenario based on Analysis 5 (posterior probability = 0.653). Dashed lines correspond to unsampled (ghost) populations, and the thin line on each introduced population represents a bottleneck. Population numbers are given in brackets. (b) Summary of the global invasion scenario inferred from our analysis. Green arrows indicate invasions from the native area, and the red arrow represents the bridgehead hypothesis. *p*-values for each analysis are given with the corresponding arrow. The coloured circles indicate the populations used for the ABC-RF analysis. Invasive and native areas are indicated by codes (NCC: North Central China; NEC: North East China; SK: South Korea; NA: North America; EU: Europe)

already infested by ALB. Conversely, both clustering and population genetics analyses have shown that in South Korea, which—in contrast to China—was not involved in a reforestation programme, native ALB populations form a differentiated cluster.

Estimation of posterior distributions of parameters of the selected invasion scenario enabled the formation of several hypotheses that can be discussed in the light of the species' biology, as well as its history in the native area. First, the very low effective size of the Korean populations used in the ABC analysis distinguishes them from the other native populations and supports the singularity of the Korean sampling area. The Korean samples originated from a natural forest that was most likely dominated by nonhost species (e.g., conifers and oaks; Lee, Kang, & Park, 2004) for ALB (Sjöman et al., 2014). This might have influenced the population genetic structure of ALB by affecting spread and reproduction. Second, the merge events between the studied native populations and their native ancestral populations were estimated to have occurred about 150 generations ago (which corresponds to roughly 150 years, according to Meng et al., 2015). Therefore, a reasonable hypothesis is that ALB population expansion in Asia really started at that date. Interestingly, despite the high uncertainty associated with parameter estimations, these assumptions are supported by historical events. Indeed, records of ALB in China date back to the Qing Dynasty, which lasted from the end of the 17th until the beginning of the 20th century. The country remained relatively isolated from international trade until the 19th century and started to open its borders during the mid-1850s. This gradual opening was most likely associated with increased movement of goods and people across the country, which could have favoured human-mediated dispersal of ALB. The timing of this shift matches the expansion time estimated by our analysis. All of these results need further investigation, but they indicate a complex history prior to the first invasion, most likely including recent contact within the native range between historically isolated populations.

4.2 | Evidence of a complex worldwide invasion scenario

In North America, ALB was detected for the first time in 1996 in the USA (Haack et al., 1996) and then in 2003 in Canada (Turgeon et al., 2015). The invasion of this area is thought to have resulted from distinct introductions, followed in some cases by a secondary spread (Carter et al., 2010). The invasion history in North America could not, however, be specifically tested in this study, owing to limited sampling in the USA and Canada. Europe was the second continent to be invaded by ALB, and, following the first detection of an established population in Austria in 2001, infestations and interceptions at entry points have been reported regularly (EPPO Global Database). The high frequency of these reports, as well as the distribution of mitochondrial haplotypes (Javal, Roques et al., 2017), has suggested that the European infestation was most likely due to multiple independent introductions. The high F_{ST} values in our study support this hypothesis. Most importantly, the ABC analyses confirm that most of the studied European populations have different origins and that three out of the four European populations analysed were likely introduced independently from the native area.

Unexpectedly, the population in Gien, France (population 19) is an exception: it may have originated from an invasive population from the USA, rather than from the Asian native range as hypothesized in previous publications (Cocquempot, Prost, & Carmignac, 2003; EPPO Global Database). Some field observations further support this possibility: the Gien population that is thought to have received the bridgehead population from North America was first detected in a location hosting an American company (Cocquempot et al., 2003). This bridgehead scenario, in which a North American invasive population would be a source for a European infestation, has been described for several other invasive insect species (Garnas et al., 2016; Hurley et al., 2016; Lombaert et al., 2010; Yang et al., 2012). The identified pathway illustrates the possibility of secondary human-mediated spread among invaded areas, even for quarantine pests. This scenario, together with the multiple introductions from Asia, may have implications in terms of management because it could help direct early detection efforts towards material coming from all infested areas. Indeed, the identification of such scenarios emphasizes the need to consider not only the native range, but also other invaded areas to prevent or at least reduce the risk of further ALB introductions.

Interestingly, this bridgehead scenario could not be inferred from previous studies (Javal, Roques et al., 2017). Indeed, based on a single mtDNA marker, the French population of Gien (population 19) appeared genetically homogeneous and distinct from any other population, including the North American population. In some cases, it can be more challenging to infer an evolutionary history from mitochondrial DNA data than from nuclear DNA data, owing to the occurrence of selective sweeps, inherited symbionts or genetic drift affecting the whole mtDNA genome, and because mitochondrial DNA is only maternally inherited and nonrecombinant (Galtier, Nabholz, Glémin, & Hurst, 2009). For these reasons, mtDNA-based phylogeography can hardly be used to infer invasion pathways. Overall, our results illustrate the value of a multilocus approach to tracking invasion routes.

Secondary displacements of ALB within invaded zones have already been suggested by the distribution of haplotypes in certain areas (Javal, Roques et al., 2017). In Canada, more specifically, clustering analyses based on nuclear markers support the hypothesis of strong connections between the different outbreaks and confirm the results based on mitochondrial markers showing that the same haplotype (or for one specimen, a very close one) was present in all populations (Javal, Roques et al., 2017). Therefore, the Canadian populations may have originated either from a single introduction (Carter et al., 2010) followed by local spread (Turgeon et al., 2015), or from multiple introductions of closely related founders. Similarly, in Europe, pairs of populations in Corsica (populations 1 and 2) and in Switzerland (11 and 12) had low F_{ST} values and shared the same mitochondrial haplotype (Javal, Roques et al., 2017). These results suggest either a human-mediated or a natural stepping-stone dispersal in the invaded area at a small spatial scale, even though multiple introductions from the same source population cannot be excluded. In Switzerland, secondary dispersal has been documented, and a

trace-back field investigation suggested that beetles from Marly (population 12) were transported to Brünisried (population 11) in infested firewood (EPPO Global Database; Meier et al., 2015). The low genetic differentiation between these two Swiss populations supports the hypothesis of small-scale, human-mediated dispersal. Similarly, anthropogenic factors, such as transfer via cars, might have facilitated the establishment of different infestations in Corsica. The Corsican populations are, however, geographically very close (<5 km apart), and the possibility of natural dispersal cannot be ruled out (Javal, Roux, Roques, & Sauvard, 2017; Lopez, Hoddle, Francese, Lance, & Ray, 2017). To conclude, these concrete examples suggest that ALB is easily transported and dispersed. As the physiological flying capabilities of the species are known (Javal, Roux et al., 2017; Lopez et al., 2017), it would be interesting to understand the natural dispersal behaviour of ALB in its invaded environment, and how it could, combined with vector-mediated dispersal, affect the genetic structure of the invasive populations. More specifically, natural dispersal by flying could effectively produce genetic isolation by distance and human-mediated dispersal could produce a more random pattern, resulting in a complex pattern at a local scale. These additional analyses would require more intensive sampling but could provide detailed scenarios for each invaded area.

4.3 | How did the invasion scenario shape ALB's genetic structure and invasion success?

Invasion history can strongly influence the success of long-term establishment of invasive populations (Dlugosch & Parker, 2008; Estoup et al., 2016). When introduced into a new area outside their native range, invasive populations may experience a severe bottleneck (Dlugosch & Parker, 2008; Edmonds, Lillie, & Cavalli-Sforza, 2004; Estoup, Wilson, Sullivan, Cornuet, & Moritz, 2001; Peischl & Excoffier, 2015) that jeopardizes their adaptive abilities by decreasing their genetic diversity (Dlugosch & Parker, 2008; Facon et al., 2006). Multiple introductions can maintain or even increase the genetic diversity of invasive populations and, consequently, their ability to face new environmental constraints (Dlugosch & Parker, 2008). Our analyses show that the European ALB infestation most likely originated from different source populations, located in the native range of the species (China, Korea) but also in a previously invaded area (North America). The invasion history of ALB has thus maximized the likelihood of introducing genetically different specimens. As a result, allelic richness is similar in invasive and native ranges, indicating that the multiple introductions may have effectively helped maintain diversity in the invaded zone at the same level as in the native zone.

In the case of ALB, however, our findings show that multiple introductions occurred at a continental scale and only rarely within the same outbreak. Indeed, within single outbreaks, population genetic analyses revealed low diversity levels. Only one (population 9, Divonne-les-Bains, France) out of the 19 European populations analysed showed a genetic pattern suggesting an admixture of individuals from different source populations, according to the STRUCTURE results. As most existing ALB outbreaks in Europe are separated by large geographic distances that cannot be covered by natural dispersal (Javal, Roux et al., 2017; Lopez et al., 2017), and as human-mediated dispersal seems to be quite rare, admixture between genetically distinct populations is not expected, at least for now. Thus, invasive ALB populations on the European continent cannot take advantage of the aforementioned effects of introductions from genetically different source populations.

Our analyses showed a strong population structuration in the invaded area and, to a lesser extent, in the native area. This may indicate limited gene flow between populations, mostly due to the biology and the low dispersal of the species, both in its natural environment (Sawyer, 2009; Smith et al., 2001, 2004) and in its invaded range (Fragnière, Forster, Hölling, Wermelinger, & Bacher, 2018). High F_{IS} values and deviation from Hardy–Weinberg suggested a high level of inbreeding within populations in both the native and the invaded areas, resulting from mating primarily among closely related individuals.

Although low genetic diversity and a high level of inbreeding within invasive populations would rather suggest a reduced ability of ALB to rapidly face new environments, numerous introduced populations have become successfully established in the new environment. Indeed, our set of results [low level of genetic diversity, large inbreeding coefficient values within native populations, low number of founders and large differences in the severity of bottlenecks encountered by introduced populations] suggest that the species is generally only weakly affected by the adverse effects of inbreeding. Overall, its particular demography and its apparent high natural tolerance to inbreeding may be partly responsible for the success of ALB invasions. It can also be hypothesized that the reproductive pattern and the resulting inbreeding could have favoured the purging of some deleterious alleles by selecting individuals with a favourable association of alleles, potentially in the native area before the invasion process (Facon et al., 2011). This particular situation could also result from the fact that the introduced populations have not yet faced adaptive challenges in their invasive range (Estoup et al., 2016). The introduced populations might also be pre-adapted to their invaded environment because of similar host availability and climatic conditions. Indeed, host availability is not compromised in the invaded range, as most Asian host tree species are also planted at high densities in North America and in Europe. However, further long-term experiments and observations are needed to test these hypotheses. The large range of bottleneck severities computed for the invasive population reinforces the hypothesis of a high tolerance to genetic diversity loss, but it also highlights the fact that the conditions of introduction can be very different from one population to another in terms of numbers of individuals introduced but also regarding the intensity and duration of bottlenecks.

The success of ALB invasions may also be partially due to the human-mediated displacement of populations in the native area during the reforestation programme in China (Pan, 2005). These displacements may have contributed to the dispersal of particularly resilient populations, as suggested by the abundance of certain haplotypes in the native area (Javal, Roques et al., 2017). They may also have enabled the formation of new genotypes by admixture between populations that were historically genetically isolated. The admixed genotypes may have

subsequently been accidentally introduced into North America and Europe, enabling, for instance, the species to cope with a wide range of climatic conditions regardless of its precise native zone. Overall, our findings open up many new perspectives and underline the importance of a parsimonious study of the native range of an invasive alien species to better understand the success of its invasion worldwide.

ACKNOWLEDGMENTS

We warmly thank Jing Tao, Shi-xiang Zong, Shan-chun Yan, Jiang-hua Sun, Jian-ting Fan, Jacob Wickham, Melody Keena, Franck Hérard, Beat Forster and all the local observers who provided samples. We also thank Thomas Guillemaud for input on the analyses and Marie-Pierre Chapuis for comments on the manuscript. Financial support was provided by the French Ministry of Agriculture, Food and Forestry and its Forest Health Department (Convention DGAL E07/2014 "Estimation de l'impact écologique du Capricorne Asiatique *Anoplophora glabripennis* sur les populations xylophages natives et leurs organismes associés, et traçage génétique de l'origine de ses populations invasives en France continentale et en Corse"), by the Région Centre Val de Loire (Convention 2014 00095480), and by the COST Action FP 1401 "Global warning." We are also grateful to Melissa Dawes for English revision of the manuscript.

AUTHOR CONTRIBUTIONS

M.J., A.R. and G.R. designed the research. M.J. and S.P. provided samples. M.J. and C.C. conducted the experiments. M.J., E.L., C.K. and T.T. analysed the data. M.J. wrote the paper. All authors contributed to earlier drafts of the manuscript, and they all read and approved of the final version.

DATA ACCESSIBILITY

Microsatellite data: https://figshare.com/s/2387940ec60f27a586ef.

ORCID

Marion Javal https://orcid.org/0000-0001-7878-2936
Eric Lombaert https://orcid.org/0000-0003-0949-6690

REFERENCES

An, Y.-L., Wang, B.-D., Yang, X.-J., Lin, X.-J., Chen, J.-D., Huang, X.-M., & Mastro, V. (2004). Characterizing populations of *Anoplophora glabripennis* and related taxa with RAPD. *Acta Entomologica Sinica*, *47*, 229–235.

Arca, M., Mougel, F., Guillemaud, T., Dupas, S., Rome, Q., Perrard, A., … Silvain, J. (2015). Reconstructing the invasion and the demographic history of the yellow-legged hornet, *Vespa velutina*, in Europe. *Biological Invasions*, *17*, 2357–2371. https://doi.org/10.1007/s10530-015-0880-9

Auger-Rozenberg, M.-A., Boivin, T., Magnoux, E., Courtin, C., Roques, A., & Kerdelhué, C. (2012). Inferences on population history of a seed chalcid wasp: Invasion success despite a severe founder effect from an unexpected source population. *Molecular Ecology*, *21*, 6086–6103. https://doi.org/10.1111/mec.12077

Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*, 2025–2035.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (2004). GENETIX *4.05, logiciel sous Windows TM pour la génétique des populations*. Montpellier, France: Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II.

Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: Some cons, many pros. *Molecular Ecology*, *19*, 2609–2625. https://doi.org/10.1111/j.1365-294X.2010.04690.x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Carter, M., Casa, A. M., Zeid, M., Mitchell, S. E., & Kresovich, S. (2008). Isolation and characterization of microsatellite loci for the Asian longhorned beetle, *Anoplophora glabripennis*. *Molecular Ecology Resources*, *9*, 925–928.

Carter, M. E., Smith, M. T., & Harrison, R. G. (2009). Patterns of genetic variation among populations of the Asian longhorned beetle (Coleoptera: Cerambycidae) in China and Korea. *Annals of the Entomological Society of America*, *102*, 895–905. https://doi.org/10.1603/008.102.0516

Carter, M., Smith, M., & Harrison, R. (2010). Genetic analyses of the Asian longhorned beetle (Coleoptera, Cerambycidae, *Anoplophora glabripennis*), in North America, Europe and Asia. *Biological Invasions*, *12*, 1165–1182. https://doi.org/10.1007/s10530-009-9538-9

Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis: Models and estimation procedures. *The American Journal of Human Genetics*, *19*, 233–257.

Chapuis, M.-P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, *24*, 621–631. https://doi.org/10.1093/molbev/msl191

Choisy, M., Franck, P., & Cornuet, J. M. (2004). Estimating admixture proportions with microsatellites: Comparison of methods based on simulated data. *Molecular Ecology*, *13*, 955–968. https://doi.org/10.1111/j.1365-294X.2004.02107.x

Cocquempot, C., Prost, M., & Carmignac, D. (2003). Interceptions et introductions en France de Longicornes asiatiques : Cas des *Anoplophora glabripennis* (Motschulsky) et *chinensis* (Forster) (Coleoptera Cerambycidae). *Bulletin Mensuel De La Société Linnéenne De Lyon*, *72*, 273–278. https://doi.org/10.3406/linly.2003.13483

Cornuet, J.-M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J.-M., Balding, D. J., … Estoup, A. (2008). Inferring population history with DIYABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics*, *24*, 2713–2719. https://doi.org/10.1093/bioinformatics/btn514

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & Francois, O. (2010). Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*, 410–418. https://doi.org/10.1016/j.tree.2010.04.001

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*, 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Earl, D. A., & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Edmonds, C. A., Lillie, A. S., & Cavalli-Sforza, L. L. (2004). Mutations arising in the wave front of an expanding population. *PNAS*, *101*, 975–979. https://doi.org/10.1073/pnas.0308064100

Essl, F., Bacher, S., Blackburn, T. M., Booy, O., Brundu, G., Brunel, S., … Jeschke, J. M. (2015). Crossing frontiers in tackling pathways of biological invasions. *BioScience*, *65*, 769–782. https://doi.org/10.1093/biosci/biv082

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

Estoup, A., Ravign, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics*, *47*, 51–72.

Estoup, A., Wilson, I. J., Sullivan, C., Cornuet, J., & Moritz, C. (2001). Inferring population history from microsatellite and enzyme data in serially introduced cane toads, *Bufo marinus*. *Genetics*, *159*, 1671–1687.

European and Mediterranean Plant Protection Organization (EPPO) Global Database. (2016). https://gd.eppo.int/taxon/ANOLGL/distribution

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Facon, B., Genton, B. J., Shykoff, J., Jarne, P., Estoup, A., & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, *21*, 130–135. https://doi.org/10.1016/j.tree.2005.10.012

Facon, B., Hufbauer, R., Tayeh, A., Loiseau, A., Lombaert, E., Vitalis, R., … Estoup, A. (2011). Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. *Current Biology*, *21*, 424–427. https://doi.org/10.1016/j.cub.2011.01.068

Favaro, R., Battisti, A., & Faccoli, M. (2013). Dating *Anoplophora glabripennis* introduction in North-East Italy by growth-ring analysis. *Journal of Entomological and Acarological Research*, *45*, 35.

Fragnière, Y., Forster, B., Hölling, D., Wermelinger, B., & Bacher, S. (2018). A local risk map using field observations of the Asian longhorned beetle to optimize monitoring activities. *Journal of Applied Entomology*, *142*, 578–588. https://doi.org/10.1111/jen.12491

Fraimout, A., Debat, V., Fellous, S., Hufbauer, R. A., Foucaud, J., Pudlo, P., … Estoup, A. (2017). Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. *Molecular Biology and Evolution*, *34*, 980–996.

Galtier, N., Nabholz, B., Glémin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, *18*, 4541–4550. https://doi.org/10.1111/j.1365-294X.2009.04380.x

Garnas, J. R., Auger-Rozenberg, M.-A., Roques, A., Bertelsmeier, C., Wingfield, M. J., Saccaggi, D. L., … Slippers, B. (2016). Complex patterns of global spread in invasive insects: Eco-evolutionary and management consequences. *Biological Invasions*, *18*, 935–952. https://doi.org/10.1007/s10530-016-1082-9

Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, *10*, 305–318. https://doi.org/10.1046/j.1365-294x.2001.01190.x

Goldberg, E. E., & Lande, R. (2007). Species' borders and dispersal barriers. *The American Naturalist*, *170*, 297–304.

Haack, R. A., Cavey, J. F., Hoebeke, E. R., & Law, K. (1996). *Anoplophora glabripennis*: A new tree-infesting exotic cerambycid invades New York. *Michigan Entomological Society Newsletter*, *4*, 11–13.

Haack, R. A., Hérard, F., Sun, J., & Turgeon, J. J. (2010). Managing invasive populations of Asian longhorned beetle and citrus longhorned beetle: A worldwide perspective. *Annual Review of Entomology*, *55*, 521–546. https://doi.org/10.1146/annurev-ento-112408-085427

Hérard, F., Ciampitti, M., Maspero, M., Krehan, H., Benker, U., Boegel, C., … Bialooki, P. (2006). *Anoplophora* species in Europe: Infestations and management processes. *EPPO Bulletin*, *36*, 470–474. https://doi.org/10.1111/j.1365-2338.2006.01046.x

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913. https://doi.org/10.1038/35016000

Hu, J., Angeli, S., Schuetz, S., Luo, Y., & Hajek, A. E. (2009). Ecology and management of exotic and endemic *Anoplophora glabripennis* (Coleoptera: Cerambycidae). *Agricultural and Forest Entomology*, *11*, 359–375.

Hulme, P. E. (2009). Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, *46*, 10–18. https://doi.org/10.1111/j.1365-2664.2008.01600.x

Hurley, B. P., Garnas, J., Wingfield, M. J., Branco, M., Richardson, D. M., & Slippers, B. (2016). Increasing numbers and intercontinental spread of invasive insects on eucalypts. *Biological Invasions*, *18*, 921–993. https://doi.org/10.1007/s10530-016-1081-x

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K=2 conundrum. *Molecular Ecology*, *26*, 3594–3602.

Javal, M., Roques, A., Haran, J., Hérard, F., Keena, M., & Roux, G. (2017). Complex invasion history of the Asian longhorned beetle: Fifteen years after first detection in Europe. *Journal of Pest Science*, *92*(1), 173–187. https://doi.org/10.1007/s10340-017-0917-1

Javal, M., Roux, G., Roques, A., & Sauvard, D. (2017). Asian Long-horned Beetle dispersal potential estimated in computer-linked flight mills. *Journal of Applied Entomology*, *142*(1–2), 282–286. https://doi.org/10.1111/jen.12408

Kalinowski, S. T. (2005). HP-RARE: A computer program for per- forming rarefaction on measures of allelic diversity. *Molecular Ecology Notes*, *5*, 187–189.

Kenis, M., Auger-Rozenberg, M.-A., Roques, A., Timms, L., Péré, C., Cock, M. J. W., … Lopez-Vaamonde, C. (2008). Ecological effects of invasive alien insects. *Biological Invasions*, *11*, 21–45. https://doi.org/10.1007/s10530-008-9318-y

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191.

Lee, D. K., Kang, H. S., & Park, Y. D. (2004). Natural restoration of deforested woodlots in South Korea. *Forest Ecology and Management*, *201*, 23–32. https://doi.org/10.1016/j.foreco.2004.06.019

Lesieur, V., Lombaert, E., Guillemaud, T., Courtial, B., Strong, W., Roques, A., & Auger-Rozenberg, M. (2018). The rapid spread of *Leptoglossus occidentalis* in Europe: A bridgehead invasion. *Journal of Pest Science*, https://doi.org/10.1007/s10340-018-0993-x

Liebhold, A. M., Brockerhoff, E. G., & Kimberley, M. (2017). Depletion of heterogeneous source species pools predicts future invasion rates. *Journal of Applied Ecology*, *54*, 1968–1977. https://doi.org/10.1111/1365-2664.12895

Lingafelter, S. W., & Hoebeke, E. R. (2002). *Revision of* Anoplophora *(Coleoptera: Cerambycidae)*. Washington, D.C.: Entomological Society of Washington.

Lippens, C., Estoup, A., Hima, M. K., Loiseau, A., Tatard, C., Dalecky, A., … Brouat, C. (2017). Genetic structure and invasion history of the house mouse (*Mus musculus domesticus*) in Senegal, West Africa: A legacy of colonial and contemporary times. *Heredity*, *119*, 64–75. https://doi.org/10.1038/hdy.2017.18

Lombaert, E., Ciosi, M., Miller, N. J., Sappington, T. W., Blin, A., & Guillemaud, T. (2018). Colonization history of the Western Corn Rootworm (*Diabrotica virgifera virgifera*) in North-America: Insights from random forest ABC using microsatellite data. *Biological Invasions*, https://doi.org/10.1101/117424

Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of

the biocontrol harlequin ladybird. *PLoS ONE*, *5*, e9743. https://doi.org/10.1371/journal.pone.0009743

Lombaert, E., Guillemaud, T., Lundgren, J., Koch, R., Facon, B., Grez, A., … Estoup, A. (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, *23*, 5979–5997.

Lombaert, E., Guillemaud, T., Thomas, C. e., Lawson handley, L. j., Li, J., Wang, S., … Estoup, A. (2011). Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: Case study of the invasive ladybird *Harmonia axyridis*. *Molecular Ecology*, *20*, 4654–4670. https://doi.org/10.1111/j.1365-294X.2011.05322.x

Lopez, V. M., Hoddle, M. S., Francese, J. A., Lance, D. R., & Ray, A. M. (2017). Assessing flight potential of the invasive Asian Longhorned Beetle (Coleoptera: Cerambycidae) with computerized flight mills. *Journal of Economic Entomology*, *110*, 1070–1077. https://doi.org/10.1093/jee/tox046

Meier, F., Engesser, R., Forster, B., Odermatt, O., Angst, A., &Hölling, D. (2015). Forstschutz-Ueberblick 2014. WSL Berichte, Heft 23. 32 p.

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, *24*, 3223–3231. https://doi.org/10.1111/mec.13243

Meng, P. S., Hoover, K., & Keena, M. A. (2015). Asian longhorned beetle (Coleoptera: Cerambycidae), an introduced pest of maple and other hardwood trees in North America and Europe. *Journal of Integrated Pest Management*, *6*, 4. https://doi.org/10.1093/jipm/pmv003

Miller, N., Estoup, A., Toepfer, S., Bourguet, D., Lapchin, L., Derrigj, S., … Guillemaud, T. (2005). Multiple transatlantic introductions of the western corn rootworm. *Science*, *310*, 992. https://doi.org/10.1126/science.1115871

Monceau, K., Bonnard, O., & Thiéry, D. (2014). *Vespa velutina* : A new invasive predator of honeybees in Europe. *Journal of Pest Science*, *87*, 1–16. https://doi.org/10.1007/s10340-013-0537-3

Nowak, D. J., Pasek, J. E., Sequeira, R. A., Crane, D. E., & Mastro, V. C. (2001). Potential effect of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) on urban trees in the United States. *Journal of Economic Entomology*, *94*, 116–122.

Pan, H. Y. (2005). Review of the Asian Longhorned Beetle research, biology, distribution and management in China Forest Resources Development Service, Working Paper, FBS/6E.

Pascual, M., Chapuis, M. P., Mestres, F., Mestres, J., Balanyà, R., Huey, R. B., … Estoup, A. (2007). Introduction history of *Drosophila subobscura* in the New World: A microsatellite-based survey using ABC methods. *Molecular Ecology*, *16*, 3069–3083.

Peakall, R., & Smouse, P. E. (2012). GENALEX 6.5: Genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peischl, S., & Excoffier, L. (2015). Expansion load: Recessive mutations and the role of standing genetic variation. *Molecular Ecology*, *24*, 2084–2094. https://doi.org/10.1111/mec.13154

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, *32*, 859–866. https://doi.org/10.1093/bioinformatics/btv684

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Raymond, M., & Rousset, F. (1995). GENEPOP (version. 1.2), a population genetics software for exact tests and ecumenicism. *Journal of Heredity*, *86*, 248–249.

Raynal, L., Marin, J.-M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2018). ABC random forests for Bayesian parameter inference. arXiv:1605.05537v5.

Robert, C., Cornuet, J.-M., Marin, J.-M., & Pillai, N. (2011). Lack of confidence in ABC model choice. *PNAS*, *108*, 15112–15117.

Roques, A., Auger-Rozenberg, M. A., Blackburn, T. M., Garnas, J. R., Pyšek, P., Rabitsch, W., … Duncan, R. P. (2016). Temporal and interspecific variation in rates of spread for insect species invading Europe during the last 200 years. *Biological Invasions*, *18*, 907–920.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstruction of phylogenetic trees. *Molecular Biology and Evolution*, *4*, 406–425.

Sawyer, A. (2009). Expected dispersal of Asian longhorned beetles from preferred host trees as a function of infestation level and data of removal during the flight season. Report from the USDA APHIS PPQ Otis Laboratory to the ALB Technical Working Group.

Seebens, H., Blackburn, T. M., Dyer, E., Genovesi, P., Hulme, P. E., Jeschke, J. M., … Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*, 14435. https://doi.org/10.1038/ncomms14435

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., … Vilà, M. (2013). Impacts of biological invasions: What's what and the way forward. *Trends in Ecology and Evolution*, *28*, 58–66. https://doi.org/10.1016/j.tree.2012.07.013

Sjöman, H., Östberg, J., & Nilsson, J. (2014). Review of host trees for the wood-boring pests *Anoplophora glabripennis* and *Anoplophora chinensis*: An urban forest perspective. *Arboriculture & Urban Forestry*, *40*, 143–164.

Smith, M. T., Bancroft, J., Li, G., Gao, R., & Teale, S. (2001). Dispersal of *Anoplophora glabripennis* (Cerambycidae). *Environmental Entomology*, *30*, 1036–1040.

Smith, M. T., Tobin, P. C., Bancroft, J., Li, G., & Gao, R. (2004). Dispersal and spatiotemporal dynamics of Asian longhorned beetle (Coleoptera: Cerambycidae) in China. *Environmental Entomology*, *33*, 435–442. https://doi.org/10.1603/0046-225X-33.2.435

Smith, M., Turgeon, J., De Groot, P., & Gasman, B. (2009). Asian longhorned beetle *Anoplophora glabripennis* (Motschulsky): Lessons learned and opportunities to improve the process of eradication and management. *American Entomologist*, *55*, 21–25.

Turgeon, J. J., Orr, M., Grant, C., Wu, Y., & Gasman, B. (2015). Decade-old satellite infestation of *Anoplophora glabripennis* Motschulsky (Coleoptera: Cerambycidae) found in Ontario, Canada outside regulated area of founder population. *The Coleopterists Bulletin*, *69*, 674–678.

Verhoeven, K. J. F., Simonsen, K. L., & McIntyre, L. M. (2005). Implementing false discovery rate control: Increasing your power. *Oikos*, *108*, 643–647. https://doi.org/10.1111/j.0030-1299.2005.13727.x

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Yang, X. M., Sun, J. T., Xue, X. F., Li, J. B., & Hong, X. Y. (2012). Invasion genetics of the western flower thrips in China: Evidence for genetic bottleneck, hybridization and bridgehead effect. *PLoS ONE*, *7*, e34567. https://doi.org/10.1371/journal.pone.0034567

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Javal M, Lombaert E, Tsykun T, et al. Deciphering the worldwide invasion of the Asian long-horned beetle: A recurrent invasion process from the native area together with a bridgehead effect. *Mol Ecol*. 2019;28:951–967. https://doi.org/10.1111/mec.15030