**Worldwide invasion genetics of the Asian ladybird *Harmonia axyridis*: new insights from intense sampling and complementary statistical treatments**

Eric Lombaert1,2,3, Thomas Guillemaud1,2,3, Jonathan Lundgren4, Robert Koch5, Benoît Facon6, Audrey Grez7, Antoon Loomans8, Thibaut Malausa1,2,3, Oldrich Nedved9, Lidwien Raak-van den Berg10, Emma Rhule11, Goody Sprinsloo12, Arnstein Staverlokk13, Tove Steenberg14 and Arnaud Estoup6

1 Inra, UMR 1355 ISA, 06903 Sophia-Antipolis, France

2 Université de Nice Sophia Antipolis, UMR ISA, 06903 Sophia-Antipolis, France

3 CNRS, UMR 7254 ISA, 06903 Sophia-Antipolis, France

4 USDA-ARS, North Central Agricultural Research Laboratory, Brookings, SD 57006, USA

5 Department of Entomology, University of Minnesota, Saint Paul, MN 55108, USA

6 Inra, UMR CBGP (INRA/IRD/CIRAD/Montpellier SupAgro), 34988 Montferrier-sur-Lez, France

7 Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Casilla 2, Correo 15, La Granja, Santiago, Chile

8 National Reference Centre, Netherlands Food and Consumer Product Safety Authority, 6706 EA, Wageningen, The Netherlands

9 University of South Bohemia, Ceske Budejovice 37005, Czech Republic

10 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

11 Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK

12 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

13 Norwegian Institute for Nature Research, NO-7485 Trondheim, Norway

14 Aarhus University, Department of Agroecology, DK-4200 Slagelse, Denmark

**Keywords:** biological invasion, microsatellite, harlequin ladybird, invasion routes, Approximate Bayesian Computation.

**Corresponding author:**

Eric Lombaert

INRA, UMR 1301 IBSV (INRA / Université de Nice Sophia Antipolis / CNRS). 400 Route des Chappes. BP 167 - 06903 Sophia Antipolis cedex. FRANCE

E-mail: [lombaert@sophia.inra.fr](mailto:lombaert@sophia.inra.fr)

Tel: +33 4 92 38 65 06

Fax: +33 4 92 38 64 01

**Running title:** Worldwide invasion genetics of *Harmonia axyridis*

**Abstract**

Inferences about invasion histories remain challenging because of the stochastic demographic processes involved. Approximate Bayesian computation (ABC) can help to overcome these problems, but such methods of inference require a prior understanding of population structure over the study area, necessitating the use of alternative methods and an intense sampling design. In this study, we made inferences about the worldwide invasion history of the ladybird *Harmonia axyridis* by various population genetics statistical methods, using a large set of sampling sites distributed over most of its native and invaded areas. We evaluated the complementarity of the statistical methods and the consequences of using different sets of population samples for ABC inferences. We found that the worldwide *H. axyridis* invasion has involved two bridgehead invasive populations in North America, which have served as the source populations for at least six independent introductions into other continents. We also identified several situations of genetic admixture between differentiated sources. Our results highlight the importance of coupling ABC methods with more traditional statistical approaches. We found that the choice of population samples could affect the conclusions of ABC analyses comparing possible scenarios. Approaches involving independent ABC analyses on several sample sets constitute a sensible solution, making it possible to avoid erroneous conclusions. This study provides biologists without expertise in this area with detailed methodological and conceptual guidelines for making inferences about invasion routes when dealing with a large number of sampling sites and complex population genetic structures.

**Introduction**

Inferences about the introduction routes of invasive species make it possible to describe the geographical pathways followed by propagules, between the source and invading populations. This is an important task in the exploration of fundamental eco-evolutionary aspects of colonization success or failure (Keller & Taylor 2008) and for preventing future invasions (Mack *et al.* 2000; Simberloff *et al.* 2013). Historical and observational data for invasive species are often sparse and incomplete. Indirect methods based on molecular genetic markers have therefore been increasingly used for the inference of invasion routes over the last 15 years, and have proved effective (Estoup & Guillemaud 2010). One of the main features that have emerged from the growing number of studies based on molecular methods is that the real histories of invasions are often far more complex than initially imagined, due in particular to multiple introductions (e.g. Ciosi *et al.* 2008; Darling *et al.* 2008; Dutech *et al.* 2012; Facon *et al.* 2003; Hoos *et al.* 2010; Kolbe *et al.* 2004; Lombaert *et al.* 2010; Rius *et al.* 2012). Multiple population sources and/or stochastic demo-genetic processes (such as founder effects and/or genetic admixture) may lead to a genetic structure within invaded areas that is difficult to predict. As a result, genetic studies of invasive species have become a methodological and analytical challenge in themselves (reviewed in Estoup & Guillemaud 2010).

The approximate Bayesian computation (ABC) method (Beaumont 2010; Beaumont *et al.* 2002; Bertorelle *et al.* 2010; Csillery *et al.* 2010) combines the use of simulations and summary statistics. This approach allows model-based inferences in a Bayesian setting for complex scenarios with intractable likelihoods, such as those relating to the introduction histories of invasive species (Box 1 and Estoup & Guillemaud 2010). ABC can take sampling, demographic and genetic stochasticity into account. Thanks to these features and increasing computer power, ABC has been widely used in recent years, to retrace the invasion routes of a growing number of invasive species (Ascunce *et al.* 2011; Auger-Rozenberg *et al.* 2012; Barres *et al.* 2012; Dilmaghani *et al.* 2012; Dutech *et al.* 2012; Keller *et al.* 2012; Konecny *et al.* 2013; Lombaert *et al.* 2010; Miller *et al.* 2005; Pascual *et al.* 2007; Rius *et al.* 2012; Yalcindag *et al.* 2012; Zepeda-Paulo *et al.* 2010). Studies based on controlled simulated datasets have shown that, in most cases, ABC is more powerful in this context than more traditional population genetics methods, such as the interpretation of neighbor-joining tree or F-statistics analysis (Estoup & Guillemaud 2010; Guillemaud *et al.* 2010; Lombaert *et al.* 2010). However, ABC is not an alternative to other population genetics statistical approaches, instead being complementary to such approaches (see Box 1).

The use of the ABC method can be complex and risky, for several reasons. It may be difficult to choose the most appropriate priors (e.g. Templeton 2010) and summary statistics (e.g. Robert *et al.* 2011), and this approach is also subject to what can be called the “population sampling curse”. Sampling is a general limitation in population genetics studies. It can have a strong impact on the power of statistical analyses, which depend on sample size and the spatial sampling scheme (Muirhead *et al.* 2008; Schwartz & McKelvey 2009). More generally, the definition of a “population” is a real issue that must be dealt with (Harwood 2009; Templeton 2004; Waples & Gaggiotti 2006). Decisions about what and how to sample are even more key issues in ABC analyses of invasion routes, in which (i) population units must be defined at the outset (what is the target population and what are its putative sources?) and (ii) computational power and scenario complexity limit the number of population samples that can be analyzed simultaneously (see Lombaert *et al.* 2011). Widely used population genetics approaches, such as clustering methods, are often used to define population genetic units. Different samples from the same defined population genetic units are then pooled together before ABC to infer invasion routes (Box 1, e.g. Auger-Rozenberg *et al.* 2012; Boissin *et al.* 2012; Boubou *et al.* 2012; Dilmaghani *et al.* 2012; Keller *et al.* 2012; Konecny *et al.* 2013; Rius *et al.* 2012). This makes sense theoretically, but the methods used to define consistent genetic groups are known to perform poorly for determining the precise number and identity of the populations (Box 1, Evanno *et al.* 2005; Pritchard *et al.* 2000; Waples & Gaggiotti 2006). However, while most ABC users are aware of the possibility that the sampling scheme may be incomplete (i.e. some important populations may have been missed), only a few take into account the possibility of the clustering of samples being not entirely realistic and the potential qualitative impact on the results of the choice of samples used for ABC analyses (but see Ascunce *et al.* 2011; Pascual *et al.* 2007; Yalcindag *et al.* 2012).

We used a set of statistical methods (including those mentioned above) to make inferences about the invasion history of the Asian ladybird *Harmonia axyridis* (HA) from a large number of genotyped samples distributed over most of the areas invaded by this species. The native area of HA covers a large part of Asia (Kazakhstan, southern Siberia, Mongolia, China, Korea and Japan, reviewed in Brown *et al.* 2011), with two distinct geographic East/West clusters, as demonstrated by genetic and phenotypic analyses (Blekhman *et al.* 2010; Dobzhansky 1933; Lombaert *et al.* 2011). After a long period of use as a biocontrol agent against aphids during the 20th century, this species has recently become invasive on most continents. Invasive feral populations were first recorded in Eastern (Louisiana, USA, Chapin & Brou 1991) and Western (Oregon, USA, LaMana & Miller 1996) North America, in 1988 and 1991, respectively. Populations were then recorded in Europe (Belgium, Adriaens *et al.* 2003), South America (Argentina, Saini 2004) and Africa (South Africa, Stals 2010) in 2001. This species has spread over large areas and is still spreading in these regions, in which it has become a nuisance (Koch 2003; Roy *et al.* 2012). By applying ABC methods to microsatellite and historical data, Lombaert *et al.* (2010) have shown that the two North American outbreaks originated from two independent introductions from the native area. They also found that the eastern North American population acted as a bridgehead for the worldwide invasion, serving as the source population for the European, South American and African outbreaks, and displaying admixture with a European biocontrol strain in Europe. However, these results were obtained with a limited number of population samples, particularly for the invaded areas (i.e. only one population sample per invaded area was used). This is unfortunate, because the invaded areas are large and their spatial genetic structure remains unknown.

In this study, we made inferences about the invasion history of HA from a large set of sampling sites distributed over most of the native and invaded areas of the species. We had three goals in mind. First, we aimed to provide new insight into the invasion routes and processes associated with the worldwide spread of HA, focusing particularly on detection of the presence of bridgehead populations and genetic admixture between differentiated sources, two potential important drivers of colonisation success (Guillemaud *et al.* 2011; Rius & Darling 2014). Second, we assessed the complementarity of a number of statistical methods and the consequences of using different sets of population samples for ABC inferences. Third, we aimed to provide biologists without expertise in this field with detailed methodological and conceptual guidelines about how to make inferences about invasion routes, using recently developed Bayesian methods, when dealing with a large number of sampling sites and a complex population genetic structure (hence the presence of Box 1 as well as of the extensive Supplementary Materials section).

**Methods**

*Sampling and genotyping*

We sampled HA at a total of 47 sites. Samples were collected from 33 sites in the invaded areas (five of these samples were previously used by Lombaert *et al.* 2011): 9 sites in North America (30 to 42 individuals per sample), 4 sites in South America (30 to 42 individuals per sample), 16 sites in Europe (20 to 42 individuals per sample) and 4 sites in Africa (31 individuals per sample). We also included in our analyses the nine samples collected in the native range (26 to 36 individuals per sample) and five European biocontrol samples (18 to 29 individuals per sample) previously used by Lombaert *et al.* (2011). Complete information about the samples is provided in Fig. 1 and Table S1. Genotyping at 18 microsatellite markers was carried out for all 47 samples, as described by Loiseau *et al.* (2009).

*Genetic variation within and between populations, and tree construction*

Genetic variation within samples was quantified by calculating the mean expected heterozygosity *He* (Nei 1987) and the mean allelic richness (*AR*), by the rarefaction method of Leberg (2002), with FSTAT (version 2.9.3.2, Goudet 2002). Three of the five biocontrol samples were not used in this analysis because they were originally stored dry and amplification was, therefore, difficult for eight of the 18 microsatellite loci (Lombaert *et al.* 2011).

The level of genetic variation between populations was measured by calculating pairwise *F*ST estimates as described by Weir and Cockerham (1984) and carrying out exact tests for population genotypic differentiation (Raymond & Rousset 1995a) for all pairs of populations, with Genepop (Raymond & Rousset 1995b). As these tests involved non-orthogonal and multiple comparisons, we corrected significance levels by the false discovery rate procedure (Benjamini & Hochberg 1995). We built a neighbor-joining (NJ) tree (Saitou & Nei 1987) with the pairwise genetic distances between populations described by Cavalli-Sforza & Edwards (1967), and Populations 1.2.30 software (<http://bioinformatics.org/~tryphon/populations/>). The robustness of the tree topology was evaluated by carrying out 1000 bootstrap replicates over loci.

*Inferences about worldwide population structure based on clustering methods*

We assessed the worldwide genetic structure of HA populations, with the Bayesian clustering method implemented in STRUCTURE v2.3.3 software (Falush *et al.* 2003; Pritchard *et al.* 2000). We chose the admixture model with correlated allele frequencies and we used the sampling location as prior information (Hubisz *et al.* 2009; Schwartz & McKelvey 2009). We used default values for all the other parameters of the software. Each run consisted of a burn-in period of 105 Markov chain Monte Carlo (MCMC) iterations, followed by 105 MCMC iterations. We analyzed the whole dataset, consisting of 47 population samples from both the native and invasive ranges of the species and from the five biocontrol strains. In total, we analyzed 1442 individuals. We carried out 20 replicate runs for each prior value of the number *K* of genetic clusters, set between 1 and 15. The STRUCTURE outputs were processed with CLUMPP (Jakobsson & Rosenberg 2007), using the LargeKGreedy algorithm, with 10,000 random permutations. A similarity coefficient (G’-statistic) greater than 90% was used to assign groups of runs to a common clustering pattern (i.e. to a single modal pattern), as proposed by Wang *et al.* (2007). The run with the highest likelihood pertaining to the most frequent clustering pattern at each *K* was used for plotting with DISTRUCT (Rosenberg 2004). We then empirically observed the assignment of individuals to the various clusters with increasing *K*. We also determined the highest level of population structure, by the Δ*K* method, which is based on the second-order rate of change in the log probability of data ln(P(X|*K*)) between successive *K* values (Evanno *et al.* 2005). As a complement to the STRUCTURE analyses, we used the BAPS clustering method (Corander *et al.* 2003), which is based on groups of individuals (i.e. population samples) rather than individuals. Separate BAPS analyses were carried out for each continent independently (see Appendix S1 for details).

*ABC-based inferences about invasion scenarios*

We used an approximate Bayesian computation framework (ABC, Beaumont *et al.* 2002) to make inferences about the relationships between the HA populations between and within the four invaded continents and the native area. It was not possible, in terms of computer capacity, to analyze all the target population samples together with all the putative source population samples (Box 1). The population units (clusters) considered in the ABC analyses therefore corresponded to the genetic clusters identified by the BAPS clustering method. These clusters were similar to those identified with STRUCTURE (see the results section). All ABC analyses were carried out in DIYABC v2 software (Cornuet *et al.* 2014).

In all ABC analyses, the statistics used to summarize the data were those used by Lombaert *et al.* (2011). These statistics are detailed in Appendix S2. The ABC analyses were performed with parameter values drawn from the prior distributions described in Table 1, and by simulating 5x105 microsatellite datasets for each competing scenario. For each analysis, we estimated the posterior probabilities of the competing scenarios by polychotomous logistic regression (Cornuet *et al.* 2008) on the 1% of the simulated datasets closest to the observed dataset. We used summary statistics transformed by linear discriminant analysis (LDA) for this analysis (Estoup *et al.* 2012). The selected scenario was that with the highest posterior probability value. When an admixed scenario was selected, we estimated the posterior distributions of the admixture rate parameter by local linear regression on the 1% of the 5x105 simulated datasets closest to the observed dataset (Beaumont *et al.* 2002; Cornuet *et al.* 2008). The robustness and relevance of our inferences were assessed with methods based on the analysis of pseudo-observed simulated datasets (Cornuet *et al.* 2010; Robert *et al.* 2011): (i) type I and type II error rates were calculated, to evaluate the robustness of scenario choice, and (ii) posterior model checking was performed on the final scenario of every analysis, to evaluate the goodness of fit between the inferred evolutionary history and the real data. These analyses are detailed in Appendix S2.

*1st set of ABC analyses*

We first used eight relatively old established population clusters (i.e. first observation of HA in 2001 or before) as a putative source of each of the target clusters analyzed. These potential source clusters included the two native clusters (west and east) identified by Lombaert *et al.* (2011), the European biocontrol cluster and the clusters associated with each of the five invasive populations used in the previous studies by Lombaert *et al.* (2010; 2011) (see Fig. 1). We used putative source clusters observed before or, at the latest, in the same year as the target cluster. The target clusters were those observed earlier or, at the latest, in the same year. We added the European biocontrol cluster as a putative source when the target cluster was located in Europe, South America or Africa because European biocontrol strains were used for control purposes in these regions (REF). Other than for North America, we did not use any cluster from the same continent as a putative source in this first set of analyses. The ABC analysis involved comparing introduction scenarios in which the target cluster originated from one of the source clusters or from an admixture between two source clusters (see Fig. 2 for an illustration and Appendix S2 for more details). We had a maximum of seven putative sources and, thus, 28 competing scenarios per ABC analysis.

We tested the impact on our ABC inferences of using different samples to represent each of the target clusters. We selected two samples from each cluster. If at least three sites had been sampled for a given target cluster, we chose (i) the sampling site having the lowest mean *F*ST value with the other sampled sites from the same cluster (i.e. the “most representative” sample) and (ii) the one with the highest mean *F*ST value (i.e. the “least representative” sample).

We also assessed the impact of using different sets of population samples within each source cluster. In a first sample set (referred to hereafter as the “reference” set), we used the same population samples as in the study by Lombaert *et al.* (2011) (Table 2). In the second and third sample sets, we used the population samples displaying the highest and lowest *F*ST values with the “reference” sample (as defined above) from the same predefined cluster, respectively. The second and third sample sets are thus referred to as the “high-*F*ST” and “low-*F*ST” sets, respectively. Finally, in order to evaluate the effect of pooling different samples from the same cluster, we tested two additional sets of population source samples constructed by pooling the samples from two sites together: the first set corresponded to the pooling of the “reference” and “high-*F*ST” samples (hereafter referred to as the “pool-high-*F*ST” set), and the second corresponded to the pooling of the “reference” and “low-*F*ST” samples (hereafter referred to as the “pool-low*-F*ST” set).

Overall, because we considered five different sample sets representing the different population sources (Table 2) to analyse 15 target population samples, we performed a total of 75 independent ABC analyses (see Appendix S2 for details).

*2nd set of ABC analyses*

In the second set of ABC analyses, we defined specific competing invasion scenarios, to test the historical independence of two target clusters from the same continent for which the same extracontinental sources were identified in the first set of ABC analyses. Our aim here was to determine whether the presence of several genetic clusters within a continent was due to multiple independent intercontinental introductions or to the secondary intracontinental foundation of these clusters. See the results section and Appendix S2 for details.

**Results**

*Genetic variation and tree construction*

The complete worldwide dataset, including a total of 1442 individuals from 47 population samples (Table S1), displayed substantial polymorphism, with a mean of 14.94 alleles per locus, over all samples. Allelic richness corrected for 19 individuals per sample (and heterozygosity He) ranged from 2.37 alleles per locus (He=0.20) in a biocontrol sample (EB-INRA06) to 6.48 (He=0.56) in a native sample (N-Japan1). See Figure S1 for a concise presentation of diversity measurements for each population sample.

We found significant genotypic differentiation, in 1045 of a total of 1081 pairwise comparisons between sampling sites (Table S2). For 34 of the 36 pairs of samples displaying no significant genotypic differentiation, both samples of the pair were located on the same continent. The two remaining pairs involved the Chilean sample (I-SA-Chi), which was not significantly different from two North American samples (I-NA-Geo and I-NA-Mic). As previously described by Lombaert *et al.* (2011), there was a low level of genetic differentiation in the native area, with a mean pairwise *F*ST of 0.013. The level of genetic differentiation between all invasive samples was moderate, with a mean *F*ST of 0.052. Within continent mean *F*ST values differed substantially: 0.004, 0.023, 0.040 and 0.091 for Africa, North America, Europe and South America, respectively. European biocontrol samples displayed a high degree of differentiation from all feral (i.e. native and invasive) population samples (mean *F*ST between biocontrol and feral samples = 0.219, Table S2), but the lowest *F*ST values were systematically those obtained with European invasive population samples.

Most of the population clusters identified in the unrooted NJ tree were geographically consistent (Fig. 3). Some distinctive patterns emerge from the tree. In North America, the nine samples form two groups, with two intermediate samples (I-NA-Col and I-NA-Uta). These two groups clearly correspond to an east versus west geographical pattern (Fig. 1), probably reflecting the spatial expansion of the two independent introductions of HA into North America, as previously inferred from two samples (I-NA-Lou and I-NA-Was, Lombaert *et al.* 2010). In South America, the Chilean sample clustered within the set of Eastern North American samples, whereas the Argentinean and Brazilian samples grouped together, albeit on long branches. In Europe, most samples (13 of 16) were located in a similar position in the tree, between all the other feral populations and the biocontrol samples. Interestingly, one sample from the South of France (I-EU-Opi) grouped with the biocontrol samples, and the two Italian samples (I-EU-Cun and I-EU-Ale) clustered with the most western North American population samples. Finally, all African samples belonged to a single population cluster. However, it should be stressed that the low bootstrap values at most tree nodes make it difficult to draw any robust conclusions about the relationships between the different population clusters observed in the NJ tree.

*Worldwide population structure on the basis of Bayesian clustering methods*

STRUCTURE analyses of the worldwide HA dataset yielded consistent results (i.e. low variance and high overall similarity coefficient) over the 20 runs tested, especially from *K* = 1 to *K* = 5. The natural logarithm of the likelihood of the data lnP(X|*K*) was highest for *K* = 14, but began to level off well before this *K* value was reached, and the Δ*K* statistic was higher for *K* = 2 (Fig. S2).

Detailed STRUCTURE clustering results (Fig. 4) gave qualitative outcomes consistent with NJ tree and F-statistics analyses and with previous knowledge about the invasion history of HA (Lombaert *et al.* 2010; Lombaert *et al.* 2011). European biocontrol samples were clearly differentiated from the Asian, American and African samples at *K* = 2, whereas most European invasive population samples displayed admixture between the two clusters. The individual proportions of ancestry from each cluster (*Q* values) were between 0.1 and 0.9 in 68.3% of all European individuals, whereas such intermediate Q values were found in only 2.8% of all other individuals (i.e. from the biocontrol, Asia, America and Africa samples). This pattern is consistent with most European invasive populations being admixtures between European biocontrol and East North American (ENA) invasive populations. This admixed origin was previously demonstrated with a single European sample from Belgium (I-EU-Bel, Lombaert *et al.* 2010). One notable exception is the sample from the South of France with a very high degree of ancestry from the biocontrol cluster (I-EU-Opi; mean biocontrol cluster *Q* value = 0.98). The 3rd cluster (appearing at *K* = 3) distinguished between invasive ENA populations from the native area and invasive Western North American (WNA) populations, and confirmed the link between the ENA outbreak and South America, Africa and Europe (e.g. Lombaert *et al.* 2010). Drift pulses due to demographic bottleneck events (cf. lower diversity, see Fig S1) probably account for the emergence of a new cluster in South America at *K* = 4 (except for the Chilean sample, which did not at any point differ significantly from ENA samples) and in Africa at *K* = 6. The 5th cluster (appearing at *K* = 5) distinguished WNA from the native area. Two admixed ENA/WNA samples could be identified (I-NA-Col and I-NA-Uta) in a geographic region consistent with the notion of a contact zone between the two independent North American introductions (Fig. 1). Interestingly, the STRUCTURE pattern at *K* = 5 indicates a moderate to high level of WNA ancestry in Europe: high in Italy (mean WNA cluster *Q* values above 0.98 in I-EU-Ale and I-EU-Cun) and moderate in several of the most Eastern samples (e.g. WNA cluster *Q* values of 0.45 and 0.54 in I-EU-Ber and I-EU-Den, respectively). The genetic structuring of the native area into two population clusters (previously identified by Lombaert *et al.* 2011) appeared at *K* = 8. Finally, a last cluster corresponding to the invasive Italian population samples appeared at *K* = 9.

The results of the intracontinental BAPS clustering analyses were very similar to those of the STRUCTURE analyses (Fig. 4 and Appendix S1). We used these clustering results to define the population units for analysis by ABC. The 47 population samples were grouped into 15 population clusters (two in Asia, four in North America, three in South America, one in South Africa, four in Europe and one European biocontrol cluster; see Appendix S1 and Appendix S2). We then defined the source and target population clusters and the various sample sets used in the ABC analyses (see Table 2, Table 3 and Appendix S2).

*Prime (i.e. extra-continental) origin of target population clusters, as assessed by ABC*

The results of this first set of ABC analyses are summarized in Table 3 (see more detailed results in Table S3). Our evaluation of confidence in scenario choice for two ABC analyses (Appendix S2) revealed moderately high type I errors (0.39 and 0.20, Table S4), but markedly low mean type II errors (0.035 and 0.051, Table S4), suggesting that our model choice analyses were reliable overall. Model checking analyses indicated that the chosen evolutionary scenarios and associated posterior distributions of parameters produced simulated data that correctly fitted the observed genetic data for HA (Table S5).

Regardless of the set of population samples used as putative sources, analyses of the two North American target clusters (i.e. the I-NA-Col and I-NA-Uta samples) always gave the highest posterior probability for an admixture scenario, with the ENA and WNA clusters as parental sources. This finding is consistent with the observed NJ tree topology (Fig. 3) and STRUCTURE results (Fig. 4). ABC estimates of the genetic contribution (i.e. admixture rate) of the ENA parental source were also consistent with geographic patterns (Fig. 1), as we obtained a higher rate for the more eastern Colorado sample I-NA-Col (mean of 71% over the five ABC analyses) than for the more western Utah sample I-NA-Uta (mean of 36% over the five ABC analyses).

All ABC analyses for the African target cluster gave similar results, regardless of the samples used as the target or source populations: the ENA cluster was unambiguously the source of the African cluster (Table 3).

The results were less clear-cut for the three South American target clusters. However, an ENA origin was considered to be the most likely scenario overall for each target cluster (i.e. Chile, Brazil and Argentina; Table 3). We obtained a small number of discrepant results for the I-SA-Cur and I-SA-Arg target samples, probably due to the high type I error risks associated with other scenarios. For example, we considered the Brazilian sample I-SA-Cur to have probably originated from the ENA cluster, despite two of the five ABC analyses indicating possible admixture between the ENA and African clusters. This admixture scenario is genealogically similar to the simpler scenario of an ENA origin, because the African population itself originates from the ENA population (Table 3). It is, therefore, not surprising that the observed posterior probabilities were rather low (below 0.5) and similar for the three most likely scenarios in this case (Table S3). Moreover, analyses of the I-SA-Gon sample, from the same Brazilian cluster, systematically suggested an ENA origin and the worldwide STRUCTURE clustering analysis suggested no direct link between South America and Africa (Fig. 4). Similar arguments led us to draw the same conclusion (i.e. an ENA origin) for the Argentinian cluster represented by the I-SA-Arg sample.

Three of the four European target clusters gave clear-cut ABC results (Table 3) consistent with NJ and STRUCTURE analyses (Fig. 3 and Fig. 4). We found that the WNA cluster was the source of the Italian target cluster (I-EU-Ale and I-EU-Cun samples), whereas the South of France cluster (I-EU-Opi) was probably directly derived from the European biocontrol strain. Finally, the two samples used to determine the origin of the western European cluster (I-EU-Bel and I-EU-Cas) clearly suggested a scenario involving an admixture between the ENA cluster and the European biocontrol strain. By contrast, the ABC analyses of the two samples representing the Eastern European cluster (I-EU-Hun and I-EU-Nor) gave different results. Depending on the target sample and the putative population source set, various admixture scenarios (involving the Eastern native, ENA, WNA, South American or European biocontrol clusters) were selected, often with low to moderate posterior probabilities (Table 3). This makes it difficult to draw any firm conclusions about this specific invasive cluster at this point. The possibility of admixture between three different sources (ENA, WNA and European Biocontrol) suggested by the STRUCTURE analyses (e.g. *K* = 7; Fig. 4) was not tested formally in these ABC analyses.

*Intracontinental relationships between target population clusters, as determined by ABC*

The invasion histories of South America and Eastern Europe remained unclear. The first set of ABC analyses indicated that all three South American target clusters had the ENA cluster as their source. We investigated whether these target clusters were established independently from the ENA source population (i.e. two or three ENA introductions into South America) or whether their establishment was not independent (i.e. a single ENA introduction responsible for all the South American clusters), by defining a second set of ABC analyses including five competing scenarios (see Appendix S2 for details). Whatever the set of population source samples used, we found that the Chilean cluster was independent of the other two South American clusters (Argentina and Brazil). By contrast, the South American clusters from Argentina and Brazil both originated from the same introduction event from the ENA source population (Table 4). The low type I and II error rates indicated that these results were robust (Table S4) and there was a good fit between the model and the observed data (see model checking analyses summarized in Table S6). This scenario is consistent with the observed NJ tree topology (Fig. 3) and with STRUCTURE results (e.g. Argentina and Brazil appeared to be associated with the same private cluster at *K* = 4; Fig. 4). Thus, the HA populations in South America originated from two independent introductions from Eastern North America.

The Eastern European cluster may have resulted from admixture events between ENA, WNA and European biocontrol populations, as suggested by STRUCTURE analyses (Fig. 4). The first set of ABC analyses indicated that several clusters within Europe may have originated from similar sources. In particular, the Italian and Eastern European clusters both have a WNA origin, whereas the Western and Eastern European clusters both originate at least partly from the ENA and biocontrol clusters. We defined a second set of ABC analyses with four competing scenarios, corresponding to all the possible combinations of one single or two independent introductions from ENA and WNA (see Appendix S2 for details). Whatever the set of putative population source samples used, we found that a single introduction from ENA and two independent introductions from WNA were involved in Europe (Table 4). However, despite a good fit between the selected scenario and the data (model checking analyses, Table S6), type I and type II error rates were found to be high (0.400 and 0.136, respectively; Table S4), suggesting that these results should be interpreted with caution.

**Discussion**

*New insights into the worldwide population structure and invasion routes of* H. axyridis

Since 1988, HA has spread worldwide at a surprisingly rapid rate (Brown *et al.* 2011). Previous population genetics studies of HA have focused on the native area, the European biocontrol strains and a limited number of samples representative of the main historical invasive outbreaks (Lombaert *et al.* 2010; Lombaert *et al.* 2011). In their study, Lombaert *et al.* (2011) used a total of 19 population samples, only five of which were sampled from invaded areas. Here, we have added another 28 invasive population samples, increasing the number of individuals from 561 to 1442. Through this major additional sampling effort and the use of different methodological tools, we provide new insight into the worldwide spatial structure and invasion routes of HA populations (summarized in Fig. 5).

North America was the first continent to be invaded, with two independent introductions from the Asian native area (Lombaert *et al.* 2010; Lombaert *et al.* 2011), one in the East, first observed in 1988 (ENA outbreak), and one in the West, first observed in 1991 (WNA outbreak). HA has since spread throughout North America and is now present in almost all states and jurisdictions of the USA, Canada and Mexico (Brown *et al.* 2011; Koch *et al.* 2006). Our survey of the genetic structure of North American HA populations is highly consistent with a spatial expansion of these two initial invasive populations, with no additional introduction event. A contact zone between both outbreaks with substantial genetic admixture was identified in Utah and Colorado. This result is consistent with historical spatial establishment data for this species in North America (Koch *et al.* 2006) and raises new unresolved questions about the evolutionary and practical consequences of such genetic admixture between two already successful invasive populations.

The previously identified predominant role of North America in the worldwide invasion was confirmed by our results. We found that the Eastern North American (ENA) outbreak was the probable source of at least four introductions into other continents. One ENA introduction was responsible for the African outbreak. In South America, ladybirds from the ENA outbreak have independently founded at least two invasive populations. The population found in Brazil/Argentina has been described before (Lombaert *et al.* 2010), but the Chilean invasive population was found to have originated from a second independent introduction from ENA never before described. The previously described introduction of ENA propagules into the Western part of Europe (Belgium, France, Holland) followed by admixture with the European biocontrol strain (Lombaert *et al.* 2010) was confirmed by our analyses. We found that the Western North American (WNA) population was the source of at least two independent introductions in Europe. One of these introductions led to the establishment of a feral population in northern Italy, whereas the second probably resulted in admixture with the Western European invasive population in the more recently invaded eastern and northern parts of the continent (Germany, Poland, Czech Republic, Hungary, Denmark, Norway). The inferred history of this species in Eastern Europe should be considered with caution, as indicated by the high type II error rates (Table S4) and the greater heterogeneity of STRUCTURE Q-plots than for all other target clusters (Fig. 4). Additional introduction events may have been involved in this area. Further analyses with additional samples and/or markers are required for a finer assessment of the invasion history of HA in Eastern Europe.

We identified a single population in South East of France that appeared to have originated exclusively from the European biocontrol (EBC) strain introduced into Europe in 1982. At first sight, this is surprising, because EBC individuals have long been thought to be unable to survive in the wild (Ferran *et al.* 1997; Turgeon *et al.* 2011). This feral population was first observed in 2005 in the town of Opio (France), in an area in which many attempts at introduction for biocontrol purposes occurred in the 1990s. HA has since repeatedly been observed in this area, in which it appears to have established a feral population. This population does not seem to have expanded spatially, unlike most of the other known HA outbreaks. This locally established population attests to the ability of the European biocontrol strain to found small overwintering populations in the wild, in an area with clement winters. It may also account for this species occasionally being recorded in France for brief periods during the 1990s, before the introduction and expansion of the highly invasive Eastern North American population (Brown *et al.* 2008; Coutanceau 2006).

*Two invasive bridgehead populations and several genetic admixture situations*

Our results highlight the complexity of the worldwide invasion history of HA. Two independent bridgehead invasive populations have been the source of at least six successful secondary introductions on three other continents. A bridgehead population can be defined as an invasive population serving as a source of colonists invading other areas (Estoup & Guillemaud 2010; Guillemaud *et al.* 2011; Lombaert *et al.* 2010). The findings reported here confirm the previously described bridgehead status of the ENA population (Lombaert *et al.* 2010). The role of the WNA population as a second bridgehead population was previously unknown, but is not entirely unexpected, given the invasive success of this population in Western North America. After several attempts at the acclimation of HA for biocontrol purposes during the 20th century (Krafsur *et al.* 1997; Tedders & Schaefer 1994), the WNA and ENA outbreaks correspond to the only two populations known to have been established from the native area and to have spread. The eco-evolutionary characteristics of the ENA outbreak have been studied in detail (e.g. Facon *et al.* 2011; Labrie *et al.* 2006; Tayeh *et al.* 2013; Tayeh *et al.* 2012; Turgeon *et al.* 2011), whereas those of the WNA outbreak have yet to be investigated. Comparisons of the life-history traits of the two North American bridgehead populations with those of noninvasive populations (native or biocontrol populations) might prove a fruitful source of knowledge about the respective role of adaptation and chance in invasion success. It would be particularly interesting to determine whether inbreeding depression has been purged in the WNA population, as in the ENA population (Facon *et al.* 2011).

Several situations of genetic admixture were identified in our dataset in North America and Europe. Besides, the ENA outbreak was previously shown to have probably originated from admixture between the two native clusters (Lombaert *et al.* 2011). Other admixture events have probably already occurred or will occur in the near future, due to the presence of several expanding independent outbreaks on several continents, such as South America and Europe. Genetic admixture events between differentiated sources are known to play a crucial role in shaping the levels of genetic variation in introduced populations. They can counterbalance bottlenecks and promote high levels of genetic diversity (e.g. Bossdorf *et al.* 2005; Kolbe *et al.* 2004; Roman & Darling 2007), and they may also directly increase or decrease the mean fitness of individuals in a population, depending on the importance of both genetic load and local adaptations (Edmands 1999; Lynch 1991; Marr *et al.* 2002; Rius & Darling 2014). The ENA propagules introduced into Western Europe may have beneﬁted from admixture with the European biological control strain (Turgeon *et al.* 2011), but no heterosis or outbreeding depression has been observed in laboratory crosses between individuals from various invasive HA populations (Tayeh *et al.* 2013). However, the full consequences of admixture in the wild remain to be explored, and the involvement of the WNA population in several admixture situations in North America and Europe has not yet been investigated.

*Various methods providing congruent and complementary results*

Despite the impressive complexity of the inferred invasion routes, most of the population genetics methods that we used provided congruent results. Basic measurements and representations of genetic variation within and between samples (e.g. allelic richness, *F*ST, NJ tree) and the use of Bayesian clustering methods implemented in STRUCTURE or BAPS software provided a first set of meaningful qualitative insights into the invasion routes of HA. Clustering methods also made it possible to classify samples into sensible genetic units. The subsequent use of ABC methods then made it possible to carry out rigorous tests for evolutionary relationships between these predefined genetic units. We argue that the combined use of several methods, as in this study, is crucial, given the known difficulties in making robust inferences about complex genetic structure from data for a large set of sampling sites (e.g. Evanno *et al.* 2005; Kalinowski 2011; Konecny *et al.* 2013; Lombaert *et al.* 2011; Waples & Gaggiotti 2006).

ABC is a powerful method, but it can give questionable results if misused. It is important, therefore, to check the quality of the results obtained thoroughly (Bertorelle *et al.* 2010; Cornuet *et al.* 2010; Robert *et al.* 2011). Our estimates of type I and type II error rates in several ABC analyses suggest that, overall, we had sufficient power to discriminate between the sets of scenarios studied. Interestingly, we found that, despite the robustness of our first set of ABC analyses, the choice of datasets for the analysis did, in some cases, affect the selection of the most likely scenario. This was particularly true for analyses of the prime origin of the Eastern European cluster, and to a lower extent forMore surprisingly analyses focusing on the population clusters of Brazil, Argentina, Western Europe and the South of France. However, in all the replicate analyses carried out, a single scenario predominated, and this scenario was, in each case, consistent with previous population genetics analyses. Closer analysis of the cases in which the “wrong scenario” was selected indicated that type I errors may have been responsible for these incorrect selections. For example, when the ENA-origin scenario was simulated for the analysis of the Brazilian population I-SA-Cur, most of the type I errors (the overall type I error was 0.39, Table S4) were associated with scenarios identifying the source population as the African population (0.19) or an admixture between the ENA and African populations (0.07). Consistent with these findings, these two scenarios were those systematically presenting the highest posterior probabilities, together with the ENA-origin scenario, in analyses of real datasets (Table S3), leading to erroneous conclusions in some cases. We hence show here that the choice of population samples can have non-negligible consequences for the conclusions drawn from a single ABC analysis. The common practice of pooling differentiated population samples may also give misleading results, as illustrated by the erroneous result obtained with the “pool-high-*F*ST” sample set for analysis of the invasion history of the Argentinian sample I-SA-Arg (Table 3). We argue that the use of different sample sets and comparisons of the results obtained in these different comparisons, together with the calculation of type I and type II errors, represents a sensible solution, making it possible to minimize the misinterpretation of ABC analyses (see also Box 1).

*Conclusion*

The worldwide invasion routes of HA are known to be complex (Lombaert *et al.* 2010; Lombaert *et al.* 2011), but the results of this study highlight the huge gap between previous historical data and the true invasion history of this species. Multiple invasion bridgeheads, introductions and admixtures have served as cornerstone events in the invasion history of this species. The actual invasion history of this species may be even more complex than that inferred here. Although intensive, our sampling scheme remains incomplete, and we therefore cannot exclude the possibility of other as yet unelucidated introduction events. Overall, our findings confirm that accidental introductions have probably played a major role in the current distribution of HA (Brown *et al.* 2011; Day *et al.* 1994; Evans *et al.* 2011; Koch 2003; Obrycki & Kring 1998), as the invasion history described here does not correspond to knowledge of the history of biocontrol attempts. Our findings also show that this species is a very good model for exploring the role of multiple introduction and admixture events in the evolutionary potential of invasive species and the potential occurrence of evolutionary and/or environmental shifts in bridgehead populations (Facon *et al.* 2006; Keller & Taylor 2008; Lee 2002; Wares *et al.* 2005). Finally, from a broader methodological perspective, this study highlights the importance of coupling ABC methods with more traditional approaches and suggests that carrying out independent ABC analyses on several sample sets, when possible, is a sensible solution, making it possible to avoid erroneous model choices and conclusions.

**Acknowledgments**

We thank J.-M. Cornuet, A. Dehne-Garcia, M. Kenis, L. J. Lawson Handley, R. Ware and I. A. Zakharov for fruitful discussions. We also thank our colleagues N. Berkvens, C. Borges, P. Brown, A. Coeur d’Acier, P. De Clercq, I. Goryacheva, L. Hautier, W. Jones, E. Jousselin, G. Kergoat, S. Kögel, J. Li, A. Migeon, H. Pang, R. Stals, V. Ternois, C. E. Thomas, J. van Lenteren and S. Wang for additional *H. axyridis* samples. We also thank A. Blin and A. Auguste for technical and administrative assistance. This work was funded by grants from the French ANR Biodiversite #ANR-06-BDIV-008-01 and the French Agropolis Fondation (RTRA, Montpellier, BIOFIS project #1001-001).

**References**

Adriaens T, Branquart E, Maes D (2003) The Multicoloured Asian Ladybird *Harmonia axyridis* Pallas (Coleoptera : Coccinellidae), a threat for native aphid predators in Belgium? *Belgian Journal of Zoology* **133**, 195-196.

Ascunce MS, Yang CC, Oakey J*, et al.* (2011) Global Invasion History of the Fire Ant *Solenopsis invicta*. *Science* **331**, 1066-1068.

Auger-Rozenberg MA, Boivin T, Magnoux E*, et al.* (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.

Barres B, Carlier J, Seguin M*, et al.* (2012) Understanding the recent colonization history of a plant pathogenic fungus using population genetic tools and Approximate Bayesian Computation. *Heredity* **109**, 269-279.

Beaumont MA (2010) Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics* **41**, 379-406.

Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* **57**, 289-300.

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.

Blekhman AV, Goryacheva II, Zakharov IA (2010) Differentiation of *Harmonia axyridis* Pall. according to polymorphic morphological traits and variability of the mitochondrial COI gene. *Moscow University Biological Sciences Bulletin* **65**, 174-176.

Boissin E, Hurley B, Wingfield MJ*, et al.* (2012) Retracing the routes of introduction of invasive species: the case of the Sirex noctilio woodwasp. *Molecular Ecology* **21**, 5728-5744.

Bossdorf O, Auge H, Lafuma L*, et al.* (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**, 1-11.

Boubou A, Migeon A, Roderick GK*, et al.* (2012) Test of Colonisation Scenarios Reveals Complex Invasion History of the Red Tomato Spider Mite *Tetranychus evansi*. *Plos One* **7**, 13.

Brown P, Thomas C, Lombaert E*, et al.* (2011) The global spread of *Harmonia axyridis* (Coleoptera: Coccinellidae): distribution, dispersal and routes of invasion. *Biocontrol* **56**, 623-641.

Brown PMJ, Adriaens T, Bathon H*, et al.* (2008) Harmonia axyridis in Europe: spread and distribution of a non-native coccinellid. *Biocontrol* **53**, 5-21.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics* **19**, 233-257.

Chapin J, Brou V (1991) *Harmonia axyridis* (Pallas), the third species of the genus to be found in the United States (Coleoptera: Coccinellidae). *Proceedings of the Entomological Society of Washington* **93**, 630-635.

Ciosi M, Miller NJ, Kim KS*, et al.* (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology* **17**, 3614-3627.

Corander J, Waldmann P, Marttinen P, Sillanpaa MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20**, 2363-2369.

Corander J, Waldmann P, Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* **163**, 367-374.

Cornuet J-M, Pudlo P, Veyssier J*, et al.* (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. 10.1093/bioinformatics/btt1763.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *Bmc Bioinformatics* **11**, 401.

Cornuet JM, Santos F, Beaumont MA*, et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713-2719.

Coutanceau J-P (2006) *Harmonia axyridis* (Pallas, 1773) : une Coccinelle asiatique introduite, acclimatée et en extension en France. *Bulletin de la société entomologique de France* **111**, 395-401.

Csillery K, Blum MGB, Gaggiotti OE, Francois O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* **25**, 410-418.

Darling JA, Bagley MJ, Roman J, Tepolt CK, Geller JB (2008) Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Molecular Ecology* **17**, 4992-5007.

Day WH, Prokrym DR, Ellis DR, Chianese RJ (1994) The known distribution of the predator *Propylea-Quattuordecimpunctata* (Coleoptera, Coccinellidae) in the United-States, and thoughts on the origin of this species and 5 other exotic lady beetles in Eastern North-America. *Entomological News* **105**, 244-256.

Dilmaghani A, Gladieux P, Gout L*, et al.* (2012) Migration patterns and changes in population biology associated with the worldwide spread of the oilseed rape pathogen Leptosphaeria maculans. *Molecular Ecology* **21**, 2519-2533.

Dobzhansky T (1933) Geographical variation in lady-beetles. *The American Naturalist* **67**, 97-126.

Dutech C, Barres B, Bridier J*, et al.* (2012) The chestnut blight fungus world tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. *Molecular Ecology* **21**, 3931-3946.

Edmands S (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* **53**, 1757-1768.

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

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**, 1591-1604.

Estoup A, Lombaert E, Marin JM*, et al.* (2012) Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources* **12**, 846-855.

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.

Evans EW, Soares AO, Yasuda H (2011) Invasions by ladybugs, ladybirds, and other predatory beetles. *Biocontrol* **56**, 597-611.

Facon B, Genton BJ, Shykoff J*, et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* **21**, 130-135.

Facon B, Hufbauer RA, Tayeh A*, et al.* (2011) Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. *Current Biology* **21**, 424-427.

Facon B, Pointier J-P, Glaubrecht M*, et al.* (2003) A molecular phylogeography approach to biological invasions of the New World by parthenogenetic Thiarid snails. *Molecular Ecology* **12**, 3027-3039.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Ferran A, Giuge L, Brun J, Gambier J, Kabiri F (1997) Coccinelle *Harmonia axyridis* pallas : mise au point sur son introduction et son utilisation en lutte biologique. *Adalia* **36**, 21-24.

Goudet J (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Updated from Goudet (1995).

Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity* **104**, 88-99.

Guillemaud T, Ciosi M, Lombaert E, Estoup A (2011) Biological invasions in agricultural settings: Insights from evolutionary biology and population genetics. *Comptes Rendus Biologies* **334**, 237-246.

Harwood TD (2009) The circular definition of populations and its implications for biological sampling. *Molecular Ecology* **18**, 765-768.

Hoos PM, Whitman Miller A, Ruiz GM, Vrijenhoek RC, Geller JB (2010) Genetic and historical evidence disagree on likely sources of the Atlantic amethyst gem clam Gemma gemma (Totten, 1834) in California. *Diversity and Distributions* **16**, 582-592.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**, 1322-1332.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. **23**, 1801-1806.

Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity* **106**, 625-632.

Keller SR, Gilbert KJ, Fields PD, Taylor DR (2012) Bayesian inference of a complex invasion history revealed by nuclear and chloroplast genetic diversity in the colonizing plant, Silene latifolia. *Molecular Ecology* **21**, 4721-4734.

Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* **11**, 852-866.

Koch RL (2003) The multicolored Asian lady beetle, *Harmonia axyridis*: A review of its biology, uses in biological control, and non-target impacts. *Journal of Insect Science* **3**, 1-16.

Koch RL, Venette RC, Hutchison WD (2006) Invasions by *Harmonia axyridis* (Pallas) (Coleoptera: Coccinellidae) in the Western Hemisphere: Implications for South America. *Neotropical Entomology* **35**, 421-434.

Kolbe JJ, Glor RE, Schettino LRG*, et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177-181.

Konecny A, Estoup A, Duplantier JM*, et al.* (2013) Invasion genetics of the introduced black rat (*Rattus rattus*) in Senegal, West Africa. *Molecular Ecology* **22**, 286-300.

Krafsur ES, Kring TJ, Miller JC*, et al.* (1997) Gene flow in the exotic colonizing ladybeetle *Harmonia axyridis* in North America. *Biological Control* **8**, 207-214.

Labrie G, Lucas E, Coderre D (2006) Can developmental and behavioral characteristics of the multicolored Asian lady beetle Harmonia axyridis explain its invasive success? *Biological Invasions* **8**, 743-754.

LaMana ML, Miller JC (1996) Field observations on *Harmonia axyridis* Pallas (Coleoptera: Coccinellidae) in Oregon. *Biological Control* **6**, 232-237.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology* **11**, 2445-2449.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**, 386-391.

Loiseau A, Malausa T, Lombaert E, Martin JF, Estoup A (2009) Isolation and characterization of microsatellites in the harlequin ladybird, *Harmonia axyridis* (Coleoptera, Coccinellidae), and cross-species amplification within the family Coccinellidae. *Molecular Ecology Resources* **9**, 934-937.

Lombaert E, Guillemaud T, Cornuet JM*, et al.* (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *Plos One* **5**, e9743.

Lombaert E, Guillemaud T, Thomas CE*, et al.* (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.

Lynch M (1991) The Genetic Interpretation of Inbreeding Depression and Outbreeding Depression. *Evolution* **45**, 622-629.

Mack RN, Simberloff D, Lonsdale WM*, et al.* (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689-710.

Marr AB, Keller LF, Arcese P (2002) Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (Melospiza melodia). *Evolution* **56**, 131-142.

Miller N, Estoup A, Toepfer S*, et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**, 992-992.

Muirhead JR, Gray DK, Kelly DW*, et al.* (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology* **17**, 1020-1035.

Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.

Obrycki JJ, Kring TJ (1998) Predaceous Coccinellidae in biological control. *Annual Review of Entomology* **43**, 295-321.

Pascual M, Chapuis MP, Mestres F*, et al.* (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**, 3069-3083.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution* **49**, 1280-1283.

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

Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution* **29**, 233-242.

Rius M, Turon X, Ordonez V, Pascual M (2012) Tracking Invasion Histories in the Sea: Facing Complex Scenarios Using Multilocus Data. *Plos One* **7**, 13.

Robert CP, Cornuet JM, Marin JM, Pillai NS (2011) Lack of confidence in approximate Bayesian computation model choice. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15112-15117.

Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**, 454-464.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137-138.

Roy HE, Adriaens T, Isaac NJB*, et al.* (2012) Invasive alien predator causes rapid declines of native European ladybirds. *Diversity and Distributions* **18**, 717-725.

Saini E (2004) Presencia de *Harmonia axyridis* (Pallas) (Coleoptera: coccinellidae) en la provincia de Buenos aires. Aspectos biologicos y morfologicos. *RIA* **33**, 151-160.

Saitou N, Nei M (1987) The neighbor-joining method - a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.

Simberloff D, Martin JL, Genovesi P*, et al.* (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution* **28**, 58-66.

Stals R (2010) The establishment and rapid spread of an alien invasive lady beetle: Harmonia axyridis (Coleoptera: Coccinellidae) in southern Africa, 2001-2009. *IOBC/WPRS Bulletin* **58**, 125-132.

Tayeh A, Estoup A, Hufbauer RA*, et al.* (2013) Investigating the genetic load of an emblematic invasive species: the case of the invasive harlequin ladybird *Harmonia axyridis*. *Ecology and Evolution* **3**, 864-871.

Tayeh A, Estoup A, Laugier G*, et al.* (2012) Evolution in biocontrol strains: insight from the harlequin ladybird Harmonia axyridis. *Evolutionary Applications* **5**, 481-488.

Tedders WL, Schaefer PW (1994) Release and establishment of *Harmonia Axyridis* (Coleoptera, Coccinellidae) in the Southeastern United States. *Entomological News* **105**, 228-243.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**, 789-809.

Templeton AR (2010) Coherent and incoherent inference in phylogeography and human evolution. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6376-6381.

Turgeon J, Tayeh A, Facon B*, et al.* (2011) Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *Journal of Evolutionary Biology* **24**, 1044-1052.

Wang S, Lewis CM, Jakobsson M*, et al.* (2007) Genetic variation and population structure in Native Americans. *Plos Genetics* **3**, 2049-2067.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**, 1419-1439.

Wares JP, Hughes AR, R.K. G (2005) Mechanisms that drives evolutionary change: insights from species introductions and invasions. In: *Species Invasions: Insights into Ecology, Evolution and Biogeography* (eds. Sax DF, Stachowicz JJ, Gaines SD), pp. 229-257. Sinauer Associates Inc, Sunderland, MA U.S.A.

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

Yalcindag E, Elguero E, Arnathau C*, et al.* (2012) Multiple independent introductions of *Plasmodium falciparum* in South America. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 511-516.

Zepeda-Paulo FA, Simon JC, Ramirez CC*, et al.* (2010) The invasion route for an insect pest species: the tobacco aphid in the New World. *Molecular Ecology* **19**, 4738-4752.

**Figure Legends**

**Figure 1: Geographic locations of the genotyped population samples of *Harmonia axyridis*.**

Note: In the central world map, the green and red areas roughly correspond to the native and invasive distribution of the species, respectively. Samples with code names in blue correspond to the “reference” population source sample set (see Table 2) previously used in the ABC analyses processed by Lombaert et al. (2011). More details about each population samples can be found in Table S1.

**Figure 2: Graphic representation of four ABC scenarios among the 28 compared to infer the prime origin of the Western European target population cluster.**

Note:The Western European target population cluster is represented by the sample I-EU-Bel (framed sample) and putative sources are those of the “reference” set (see Table 2). Population code names are as in Fig. 1 and Table S1. All parameters with associated prior distributions are described in Table 1. Time 0 is the sampling year 2010 and time 58 is the sampling year 1987 (we consider 2.5 generations per year). Time is not at scale. Scenario 1 corresponds to an Eastern native origin (Pop 1) of the target population cluster (Pop 8). Scenario 5 corresponds to a Western North American origin (Pop 5). In scenario 11, the target population cluster is the result of an admixture between individuals from the Eastern native area (Pop 1) at a rate *ar1,5,8* and from the Western North American area (Pop 5) at a rate 1 – *ar1,5,8*. In scenario 20, the target population cluster is the result of an admixture between individuals from the European biocontrol population (Pop 3) at a rate *ar3,5,8* and from the Western North American area (Pop 5) at a rate 1 – *ar3,5,8*. The scenarios presented here are part of the first set of ABC analyses (see Appendix S2 for more information).

**Figure 3: Neighbor-joining tree for the studied *Harmonia axyridis* population samples based on the distance of Cavalli-Sforza & Edwards (1967).**

Note: Population code names are as in Fig. 1 and Table S1. Different colors correspond to different continents or status: native area in green, North America in red, South America in orange, Africa in brown, Europe in purple and European biocontrol in blue. Bootstrap values calculated over 1,000 replications are given as percentages (only values > 20% are shown).

**Figure 4: Genetic clustering of the population samples of *Harmonia axyridis***

Note: Ancestry estimation based on the Bayesian individual clustering method implemented in STRUCTURE (Pritchard *et al.* 2000) are given for *K* = 2 to 9 population units (left). Each vertical line represents an individual and each color represents a genetic cluster. Individuals are grouped by population sample (names at the bottom of the figure are as in Fig. 1 and Table S1). The black text at the top of the figure groups the sampled populations by continent. The blue text at the top of the figure groups the sample by the genetic clusters inferred at the intra-continental scale with the Bayesian group clustering method implemented in BAPS (Corander *et al.* 2004). See Appendix S1 for details.

**Figure 5: Worldwide invasion scenario of *Harmonia axyridis*.**

Note: Most likely scenario of invasions of *Harmonia axyridis* into eastern North America (ENA), Western North America (WNA), South America (SA), Europe (EU) and Africa (AF) deduced from a set of complementary population genetics analyses. For each outbreak, the arrow indicates the most likely invasion pathway. Bicolor arrows indicate admixture events. Years of first observation of invasive populations are indicated for each continent. Initially collected from the native area in 1982, the European biocontrol strain is represented by a blue arrow (EBC). The ranges of the two native clusters are roughly drawn and colored in clear and dark green. The invasion routes of the most early WNA and ENA outbreaks were inferred by Lombaert *et al.* (2011).

**Tables**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameters | Distribution | Mean | Median | Mode | Quantile  2.5% | Quantile 97.5% |
| *NW; NA* | Uniform  [100 – 20,000] | 10,050 | 10,050 | NA | 590 | 19,510 |
| *NB* | Loguniform  [10 – 1,000] | 213 | 97 | 10 | 11 | 891 |
| *NFi* | Loguniform  [2 – 1,000] | 154 | 39 | 2 | 2 | 850 |
| *BDi* | Uniform  [0 – 5] | 2.5 | 2.5 | NA | 0 | 5 |
| *arj,k,i* | Uniform  [0.1 – 0.9] | 0.5 | 0.5 | NA | 0.12 | 0.88 |
| *ti* | Uniform  [*xi* – *xi*+5] | DV | DV | NA | DV | DV |
| *tun,i* | Loguniform  [*ti* – 3000] | DV | DV | DV | DV | DV |
| *tanc* | Uniform  [100 – 3000] | 1,550 | 1,550 | NA | 172 | 2,928 |
| mean *µ* | Uniform  [10-5 – 10-3] | 5.0x10-4 | 5.0x10-4 | NA | 3.4x10-5 | 9.8x10-4 |
| mean *P* | Uniform  [0.1 – 0.3] | 0.2 | 0.2 | NA | 0.10 | 0.30 |
| mean *µ*SNI | Uniform  [10-8 – 10-4] | 5.0x10-5 | 5.0x10-5 | NA | 2.5x10-6 | 9.7x10-5 |

**Table 1: Prior distributions of demographic, historical and genetic parameters used in all ABC analyses processed to retrace the worldwide routes of invasion of *H. axyridis*.**

Note: *NW*, *NA* and *NB* = stable effective population size (number of diploid individuals) of wild (native or invasive: *NW*), ancestral (*NA*) or biocontrol (*NB*) populations; *NFi* = effective number of founders during an introduction step lasting *BDi* generation(s) for invasive population *i*; *arj,k,i* = genetic contribution (admixture rate) of population *j* when the invasive population *i* is formed by the admixture between the putative source populations *j* and *k*; *ti* = introduction time of invasive populations *i* with bounds *xi* fixed from dates of first observation of established population (as in Brown *et al.* 2011); *tun,i* = merging time of the source unsampled native population into the sampled native population *n*, when this native population is or is supposed to be the source of the invasive or biocontrol population *i* (see Lombaert *et al.* 2011); *tanc* = merging time of the two native populations into an ancestral unsampled one (with condition *tuj* ≤ *tanc*). All times were expressed in numbers of generation assuming 2.5 generations per year and running back in time. The microsatellite loci were assumed to follow a generalized stepwise mutation model (Estoup *et al.* 2002) with three parameters: the mean mutation rate (mean *µ*), the mean parameter of the geometric distribution (mean *P*) of length in terms of the number of repeats of mutation events and the mean mutation rate for single nucleotide instability (mean *µ*SNI). Each locus has a possible range of 40 contiguous allelic states and is characterized by individual *µ*loc drawn from a Gamma (mean = mean *µ* and shape = 2), *P*loc drawn from a Gamma (mean = mean *P* and shape = 2) and *µ*SNIloc drawn from a Gamma (mean = mean *µ*SNI and shape = 2) distribution. For a loguniform[x;y] distribution, Log(x) and Log(y) are the bounds of a uniform distribution. All prior quantities presented were computed from 106 values. NA = not applicable; DV = can take different values. See Fig. 2 for a graphical representation of several evolutionary scenarios with associated parameters considered in the ABC analyses.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Different sets of samples representative of potential source population clusters | | | | |
| Potential source population cluster | “Reference” | “High-*F*ST” | “Low-*F*ST” | “Pool-high-*F*ST” | “Pool-low-*F*ST” |
| Eastern Asia | N-China2 | N-China1 | N-Japan1 | N-China2 + N-China1 | N-China2 + N-Japan1 |
| Western Asia | N-Kazak | N-Russia1 | N-Russia2 | N-Kazak + N-Russia1 | N-Kazak + N-Russia2 |
| Eastern North America | I-NA-Lou | I-NA-Dak | I-NA-Geo | I-NA-Lou + I-NA-Dak | I-NA-Lou + I-NA-Geo |
| Western North America | I-NA-Was | I-NA-Ida | I-NA-Was | I-NA-Was + I-NA-Ida | I-NA-Was + I-NA-Ida |
| European Biocontrol | EB-INRA87 | EB-Biobest | EB-Biotop | EB-INRA87 + EB-Biobest | EB-INRA87 + EB-Biotop |
| South America | I-SA-Cur | I-SA-Arg | I-SA-Gon | I-SA-Cur + I-SA-Arg | I-SA-Cur + I-SA-Gon |
| Europe | I-EU-Bel | I-EU-Cas | I-EU-Bou | I-EU-Bel + I-EU-Cas | I-EU-Bel + I-EU-Bou |
| Africa | I-AF-Som | I-AF-not | I-AF-Bet | I-AF-Som + I-AF-not | I-AF-Som + I-AF-bet |

**Table 2: Description of the five different sample sets that were chosen to represent the potential source population clusters in the ABC analyses.**

Note: Sample names are as in Fig.1 and Table S1.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Most likely scenario (i.e. origin) when using different sets of samples representative of source population clusters (with its posterior probability) | | | | |  |
| Continent | Target population cluster | Target population sample | “Reference” | “High-*F*ST” | “Low-*F*ST” | “Pool-high-*F*ST” | “Pool-low-*F*ST” | Final selected scenario |
| North America | Colorado | I-NA-Col | ENA + WNA  (0.36) | ENA + WNA  (0.83) | ENA + WNA  (0.55) | ENA + WNA  (0.77) | ENA + WNA  (0.81) | ENA + WNA (5/5) |
|  | Utah | I-NA-Uta | ENA + WNA  (0.77) | ENA + WNA  (0.42) | ENA + WNA  (0.84) | ENA + WNA  (0.68) | ENA + WNA  (0.81) | ENA + WNA (5/5) |
| South America | Brazil | I-SA-Cur | ENA  (0.41) | ENA + AF  (0.49) | ENA + AF  (0.42) | ENA  (0.45) | ENA  (0.44) | ENA (3/5) |
|  |  | I-SA-Gon | ENA  (0.81) | ENA  (0.40) | ENA  (0.50) | ENA  (0.77) | ENA  (0.83) | ENA (5/5) |
|  | Argentina | I-SA-Arg | ENA  (0.48) | WNA + AF  (0.59) | ENA  (0.57) | ENA + WNA  (0.41) | ENA  (0.43) | ENA (3/5) |
|  | Chile | I-SA-Chi | ENA  (0.51) | ENA  (0.41) | ENA  (0.58) | ENA  (0.71) | ENA  (0.71) | ENA (5/5) |
| Europe | West Europe | I-EU-Bel | EBC + ENA  (0.78) | EBC + ENA  (0.51) | EBC + ENA  (0.72) | EBC + ENA  (0.65) | EBC + ENA  (0.70) | EBC + ENA (5/5) |
|  |  | I-EU-Cas | EBC + ENA  (0.54) | EBC + AF  (0.63) | EBC + ENA  (0.76) | EBC + ENA  (0.41) | EBC + ENA  (0.79) | EBC + ENA (4/5) |
|  | South France | I-EU-Opi | EBC  (0.88) | EBC + WNA  (0.46) | EBC  (0.72) | EBC  (0.98) | EBC  (0.97) | EBC (4/5) |
|  | Italy | I-EU-Cun | WNA  (0.69) | WNA  (0.78) | WNA  (0.72) | WNA  (0.97) | WNA  (0.90) | WNA (5/5) |
|  |  | I-EU-Ale | WNA  (0.66) | WNA  (0.58) | WNA  (0.53) | WNA  (0.75) | WNA  (0.89) | WNA (5/5) |
|  | East Europe | I-EU-Hun | EBC + ENA  (0.42) | EBC + ENA  (0.44) | EBC + ENA  (0.37) | EBC + SA  (0.41) | EBC + ENA  (0.54) | EBC + ENA (4/5) |
|  |  | I-EU-Nor | ENat + EBC  (0.24) | EBC + WNA  (0.67) | EBC + WNA  (0.50) | EBC + WNA  (0.66) | EBC + WNA  (0.55) | EBC + WNA (4/5) |
| Africa | South Africa | I-EU-Bet | ENA  (0.55) | ENA  (0.76) | ENA  (0.70) | ENA  (0.68) | ENA  (0.77) | ENA (5/5) |
|  |  | I-EU-Som | ENA  (0.71) | ENA  (0.64) | ENA  (0.89) | ENA  (0.73) | ENA  (0.88) | ENA (5/5) |

**Table 3: Results of the first set of ABC analyses processed to make inferences about the prime origin of different target invasive population clusters.**

Note: the most likely scenario (i.e. the source cluster(s)) with its posterior probability between brackets is given for each target invasive population cluster and sample (names as in Fig. 1 and Table S1) and for each putative population source sample set (see Table 2). For each target population sample, the last column indicates the scenario that was found the most frequently among the five ABC analyses (the frequency is given between brackets). More detailed results are provided in Table S3. ENat = Eastern native area; ENA = Eastern North America; WNA = Western North America; SA = South America; AF = Africa; EBC = European Biocontrol.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Continent  (target clusters) | Number of competing scenarios | Source sample set | Selected scenario | Posterior probability of selected scenario with 95% credibility intervals between brackets |
| South America  (Brazil / Argentina / Chile) | 5 | “Reference” | Two independent ENA introductions: [Chile] + [Brazil, Argentina] | 0.747 [0.733,0.762] |
|  |  | “High-*F*ST” | Two independent ENA introductions: [Chile] + [Brazil, Argentina] | 0.788 [0.778,0.798] |
|  |  | “Low-*F*ST” | Two independent ENA introductions: [Chile] + [Brazil, Argentina] | 0.692 [0.676,0.708] |
|  |  | “Pool-high-*F*ST” | Two independent ENA introductions: [Chile] + [Brazil, Argentina] | 0.877 [0.867,0.886] |
|  |  | “Pool-low-*F*ST” | Two independent ENA introductions: [Chile] + [Brazil, Argentina] | 0.674 [0.661,0.688] |
| Europe  (East Europe / West Europe / Italy) | 4 | “Reference” | - Two independent WNA introductions: [East Europe] + [Italy]  - One ENA introduction: [West Europe] | 0.521 [0.510,0.532] |
|  |  | “High-*F*ST” | - Two independent WNA introductions: [East Europe] + [Italy]  - One ENA introduction: [West Europe] | 0.737 [0.724,0.749] |
|  |  | “Low-*F*ST” | - Two independent WNA introductions: [East Europe] + [Italy]  - One ENA introduction: [West Europe] | 0.508 [0.498,0.518] |
|  |  | “Pool-high-*F*ST” | - Two independent WNA introductions: [East Europe] + [Italy]  - One ENA introduction: [West Europe] | 0.524 [0.507,0.542] |
|  |  | “Pool-low-*F*ST” | - Two independent WNA introductions: [East Europe] + [Italy]  - One ENA introduction: [West Europe] | 0.646 [0.634,0.657] |

**Table 4: Results of the second set of ABC analyses processed to make inferences about the intra-continental origin of a subset of target invasive population clusters.**

Note: ENA = Eastern North America population cluster; WNA = Western North America population cluster. See main text (Materials and Methods sections) and Table 2 for details regarding the different source sample sets.