

Genomic changes in the biological control agent *Cryptolaemus montrouzieri* associated with introduction

Hao-Sen Li¹ | Gerald Heckel² | Yu-Hao Huang¹ | Wei-Jian Fan³ | Adam Ślipiński^{1,4} | Hong Pang¹ 

¹State Key Laboratory of Biocontrol, Ecology and Evolution, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China

²Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

³College of Life Sciences, Tianjin Normal University, Tianjin, China

⁴Australian National Insect Collection, National Research Collections, CSIRO, Canberra, Australian Capital Territory, Australia

Correspondence

Hong Pang, State Key Laboratory of Biocontrol, Ecology and Evolution, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China.
Email: lsshspang@mail.sysu.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 31572052 and 31702012; Open Project of the State Key Laboratory of Biocontrol, Grant/Award Number: 2018-04; National Key R&D Program of China, Grant/Award Number: 2017YFD0201000; Science and Technology Planning Project of Guangdong Province, China, Grant/Award Number: 2017B020202006

Abstract

Biological control is the main purpose of intentionally introducing non-native invertebrate species. The evolutionary changes that occur in the populations of the introduced biological control agents may determine the agent's efficiency and the environmental safety. Here, to explore the pattern and extent of potential genomic changes in the worldwide introduced predatory ladybird beetle *Cryptolaemus montrouzieri*, we used a reduced-representation sequencing method to analyze the genome-wide differentiation of the samples from two native and five introduced locations. Our analyses based on a total of 53,032 single nucleotide polymorphism loci showed that beetles from the introduced locations in Asia and Europe exhibited significant reductions in genetic diversity and high differentiation compared with the samples from the native Australian range. Each introduced population belonged to a unique genetic cluster, while the beetles from two native locations were much more similar. These genomic patterns were also detected when the dataset was pruned for genomic outlier loci (52,318 SNPs remaining), suggesting that random genetic drift was the main force shaping the genetic diversity and population structure of this biological control agent. Our results provide a genome-wide characterization of polymorphisms in a biological control agent and reveal genomic differences that were influenced by the introduction history. These differences might complicate assessments of the efficiency of biological control and the invasion potential of this species but also indicate the feasibility of selective breeding.

KEYWORDS

biological control introduction, *Cryptolaemus montrouzieri*, evolution, genomics, population genetics

1 | INTRODUCTION

Beneficial organisms, especially invertebrate species, are used in biological control to suppress populations of pests. The intentional introduction of invertebrate species is generally used as a form of biological control (Kumschick, Richardson, & Kueffer, 2013), including

classical and augmentative types of control. Classical biological control involves the introduction of non-native natural enemies into an area for the long-term control of non-native pests, whereas the augmentative form is the basis of commercial biological control activities and involves the periodic release of mass-produced natural enemies that may or may not be native (De Clercq, Mason, &

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

Babendreier, 2011). Similar to other biological introductions or invasions, unexpected evolutionary changes can occur in populations of biological control agents, which might decrease the efficiency of pest control or increase the risk to local environments (Fauvergue, Vercken, Malausa, & Hufbauer, 2012; Guillemaud, Ciosi, Lombaert, & Estoup, 2011). Recently, there have been more reports of undesired side effects on local biodiversity caused by introduced biological control agents through rapid expansion, nontarget attack, and competition, and the invasiveness of these agents is usually related to population genetic changes (e.g., *Harmonia axyridis* in North and South America and Europe, Roy et al., 2016 and *Cactoblastis cactorum*, Zimmermann, Bloem, & Klein, 2004 and *Coccinella septempunctata*, Losey et al., 2012 in North America). Thus, understanding the population genetics of biological control agents can aid in the exploration of their patterns of potential evolutionary changes and may be beneficial for biological control programs.

The mealybug destroyer, *Cryptolaemus montrouzieri* Mulsant (Coleoptera, Coccinellidae), is native to Australia and is a predator specialist of mealybug (Ślipiński, 2007). Over the last century, this predatory ladybird has been introduced to at least 64 countries or regions for classical or augmentative biological control purposes (Kairo, Paraiso, Gautam, & Peterkin, 2013). There is currently only a single available report of the side effects of such introductions on local environments, which showed interference with other biological control agents (Anneck, Karny, & Burger, 1969). However, *C. montrouzieri* populations with different introduction histories can show changes in several life history traits, such as development time, oviposition number, and performance under adverse conditions (Li, Zou, De Clercq, & Pang, 2018). We also detected range expansion in introduced populations in southern China (mentioned by Li et al. 2017). Furthermore, host range tests in a laboratory population revealed nontarget attack abilities (Maes, Grégoire, & De Clercq, 2014). These potential risks, including range expansion and nontarget attack by ladybird populations, might result from evolutionary changes that occurred during and after their introduction.

The evolution of introduced biological control agents is usually explained by processes such as population founding or admixture events, rather than being attributed to phenotypic plasticity or selection (Andraca-Gómez et al., 2014; Kajita, O'Neill, Zheng, Obyrcki, & Weisrock, 2012; Lombaert et al., 2011; Retamal et al., 2016). Furthermore, high genetic differentiation is often observed among introduced populations (Fauvergue et al., 2012). We previously conducted several genetic analyses of native and introduced populations of *C. montrouzieri*; using 12 simple sequence repeat (SSR) markers, we found high and significant genetic differentiation between native and introduced populations (Li et al., 2017). Additionally, we detected signals of positive selection by scanning for nonsynonymous mutations across mitochondrial protein-coding genes (Li, Liang et al., 2016).

Improvements in next-generation sequencing (NGS) and bioinformatic tools have spurred the development of genome-wide genetic markers for studying the ecology and evolution of nonmodel organisms (Davey et al., 2011). Recently, this genomic methodology

has been applied for the detection of genetic variation in field and mass-reared populations of biological control agents and to elucidate the genetic basis of target traits (Lommen, Jong, & Pannebakker, 2017; van de Zande et al., 2014). In this work, we aim to explore the pattern and extent of potential evolutionary changes in worldwide introduced *C. montrouzieri*. We conducted reduced-representation sequencing of samples from two native and five introduced locations of *C. montrouzieri* and performed a genome-wide survey of intra-specific polymorphism. All of the non-native beetle populations analyzed here result from artificial transfers, and their introduction histories at the studied locations are well documented. Analyses of genomic diversity and structural differences between the populations partially revealed the evolutionary forces functioning during human-mediated species introduction. Finally, the contributions of these genomic findings to biological control applications are discussed.

2 | MATERIALS AND METHODS

2.1 | Sampling and DNA extraction

A total of 128 individuals of *C. montrouzieri* from two native and five introduced locations were analyzed in this study (Figure 1 and Supporting Information Table S1). Individuals from the native locations Brisbane, Queensland, Australia (QL), and Canberra, ACT, Australia (BM), and the introduced locations Shenzhen, mainland China (SZ), and Taipei, Taiwan (TP), were collected in the field. Individuals of SY, SA, and GT were obtained from three cultures in two laboratories at Sun Yat-sen University (Guangzhou, mainland China) (SY and SA) and Ghent University (Belgium) (GT), respectively. The SY and SA beetles were fed mealybugs, whereas the GT beetles were fed both mealybugs and moth eggs. Laboratory populations are usually mass-reared in cages under summer conditions (23–27°C and long days) in populations with large numbers (thousands) of individuals, with more than 10 generations generally being produced in 1 year. The introduction histories of the studied locations are shown in Figure 1, which are mainly based on the work of Kairo et al. (2013) and personal communications with the three laboratories. The native sources were recorded for only four studied locations, including Brisbane (QL), Canberra (BM), the Bugs for Bugs biocompany (one of the SY sources), and Queensland (one of the SA sources). GT and TP are assumed to have been founded without population admixture, whereas SY, SA, and SZ appear to have originated from several source populations that were mixed. The dates of introduction ranged from 1909 to 2016. Population sizes were not recorded during or after introduction.

For each individual, total genomic DNA was extracted using a CTAB-based procedure (Milligan, 1988). DNA quality and quantity were determined using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA) and via electrophoretic separation in an agarose gel. Only samples with a DNA concentration >20 ng/μl and slight degradation were used.

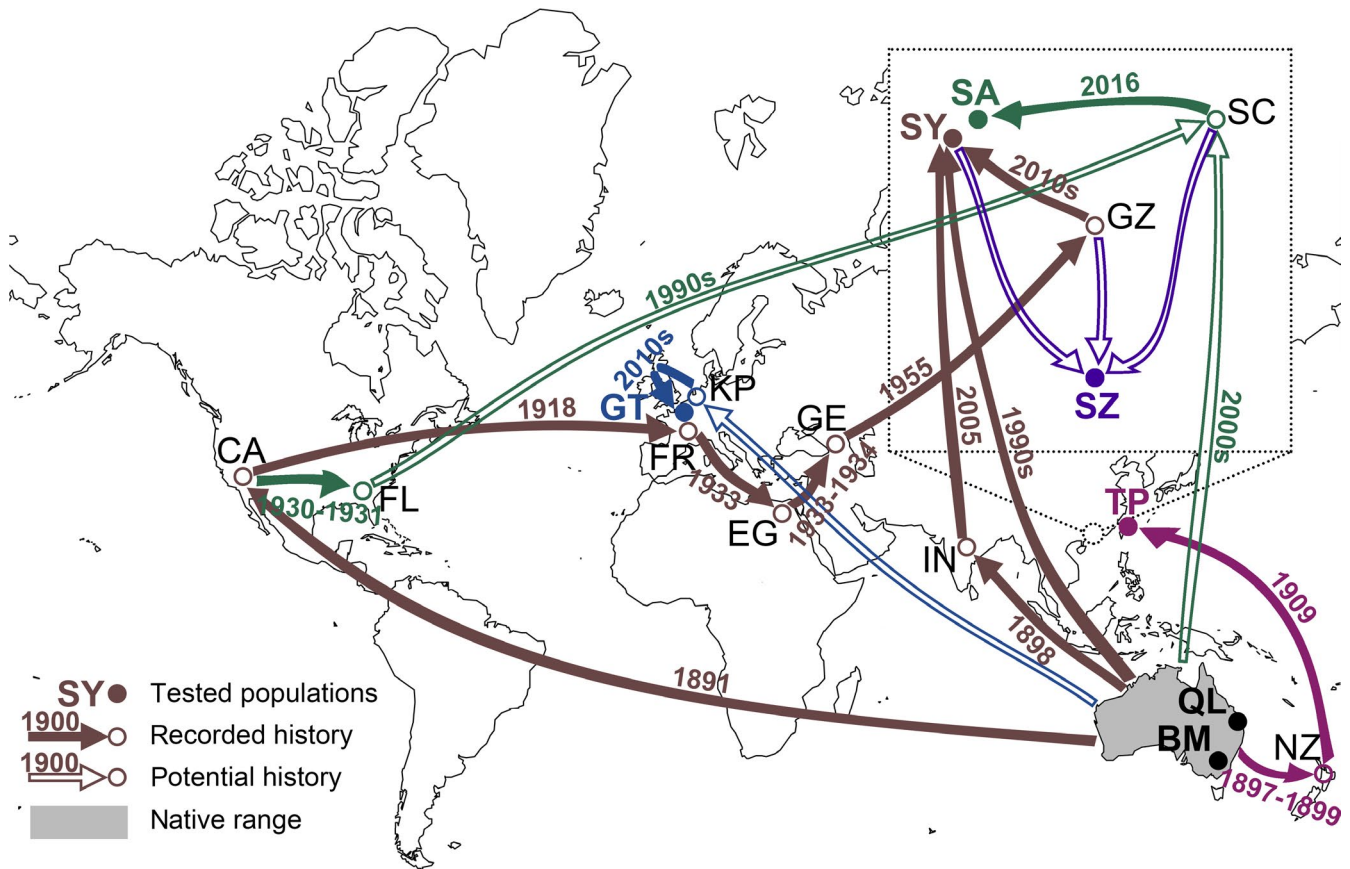


FIGURE 1 The recorded and potential introduction histories at the studied locations of *Cryptolaemus montrouzieri*. These locations included BM: Canberra, Australia; CA: California; EG: Egypt; FL: Florida; FR: France; GE: Georgia; GT: Ghent University; GZ: Guangzhou; IN: India; KP: Koppert B.V.; NZ: New Zealand; QL: Brisbane, Australia; SC: South China Agricultural University; SY/SA: Sun Yat-sen University; SZ: Shenzhen; TP: Taipei. Colors represent the histories of introductions including red: SY, green: SA, blue: GT, yellow: SZ, and pink: TP

2.2 | Reduced-representation genome sequencing

Approximately 500 ng of genomic DNA from each sample was processed for reduced-representation genome sequencing using the specific-locus amplified fragment sequencing (SLAF-seq) technique (Sun et al., 2013), with slight modifications. Briefly, in a pilot SLAF experiment, the restriction enzymes Hpy166II and EcoRV-HF (NEB, USA) were selected because they generate abundant and high-quality nonrepetitive SLAFs that are relatively evenly distributed across the genome of the model beetle species *Tribolium castaneum* (Richards et al., 2008). In addition, the DNA of a plant species, *Oryza sativa* ssp. *japonica*, was used as a control to assess the normal rate of enzyme digestion. Genomic DNA from each sample was completely digested with these two enzymes, and a single-nucleotide A overhang was added to the digested fragments. The A-tailed fragments were then ligated to dual-index sequencing adapters (Kozich, Westcott, Baxter, Highlander, & Schloss, 2013). PCR was performed in a mixture of diluted restriction–ligation samples including *Taq* DNA polymerase (NEB), dNTPs, and adapter primers. The PCR products were purified using Agencourt AMPure XP (Beckman, UK) and subjected to electrophoresis through a 2% agarose gel. Fragments in the size range of 250–350 bp were excised and purified using a

QIAquick Gel Extraction Kit (QIAGEN, Germany). The samples with indexes were allocated to a single lane and subjected to 150 bp paired-end sequencing on the HiSeq 2500 platform (Illumina, San Diego, USA) at Biomarker Technologies Corporation (China). Real-time monitoring was performed for each cycle during sequencing. The ratio of high-quality reads exhibiting Q30 (representing a quality score of 30, indicating a 0.1% chance of an error and, thus, 99.9% confidence) and the guanine-cytosine (GC) content of the raw reads were calculated for quality control.

2.3 | SNP calling

Full-length, high-quality reads with clear index information were clustered based on their sequence similarity using BLAT v1.9.4 (Kent, 2002), and those sharing >90% identity were grouped to form one SLAF tag. Polymorphic SLAF tags were employed for single nucleotide polymorphism locus (SNP) calling in the GATK v3.7 (McKenna et al., 2010) and SAMtools v1.4.1 (Li et al., 2009) routines. Only SNPs called by both routines were considered to be of high-quality. The SNPs were further filtered to exclude SNPs present in <80% of individuals and those with a minor allele frequency (MAF) <5% using PLINK v1.90 (Purcell et al., 2007). We excluded

potential paralogs by filtering loci with observed heterozygosities (H_O) higher than 0.75 (White, Perkins, Heckel, & Searle, 2013). Only one SNP per SLAF tag was retained to remove the effects of physical linkage. Deviations from Hardy–Weinberg equilibrium (HWE, $p < 0.05$) for each SNP were calculated using PLINK, and those loci that significantly deviated from HWE at more than three studied locations were excluded. The filtered SNPs were mapped to the annotated transcriptome of the species (Li, Pan, De Clercq, Ślipiński, & Pang, 2016; Zhang et al., 2012) using BLASTn, with a cut-off E -value of $10e-5$ and a minimum length of 70 bp. Finally, the filtered SNPs were converted into Arlequin, BayeScan, FASTA, and PED formats for subsequent analyses using PGDSpider v2.1.0.0 (Lischer & Excoffier, 2012).

2.4 | Detection of F_{ST} outliers

We detected F_{ST} outliers among the total SNPs using methods implemented in Arlequin 3.5.1.1 (Excoffier & Lischer, 2010) and BayeScan 2.1 (Foll & Gaggiotti, 2008). We scanned for outlier loci deviating from the expected null distribution of F_{ST} in a system of hierarchically structured populations (Arlequin's hierarchical island model, AH, Excoffier, Hofer, & Foll, 2009), which was identified as the best structure in the following analyses (six-group structure: QL+BM, SY, SA, GT, SZ, and TP). In addition, we used Arlequin's nonhierarchical island model (ANH) to test for F_{ST} outliers. In both the AH and ANH analyses, we ran 20,000 simulations, with 10 simulated groups in AH, 100 demes per group, and a maximum expected heterozygosity of 0.5. We also employed the Bayesian F_{ST} outlier method implemented in BayeScan to test whether loci were highly differentiated when parameterizing a nonhierarchical island model. BayeScan analysis was performed with a burn-in period of 50,000, followed by 100,000 Markov chain Monte Carlo (MCMC) iterations. The q -values for each locus in AH and ANH were subsequently calculated from the p -value with $p.adjust$ in R to monitor the false discovery rates (FDRs). Those loci with q -values below 0.05 were considered F_{ST} outliers. Some simulation studies have suggested that potential error rates must be evaluated by interpreting the results with a variety of methods (de Villemereuil, Fricot, Bazin, Francois, & Gaggiotti, 2014; Narum & Hess, 2011). Thus, we drew Venn diagrams to determine the intersection of high/low F_{ST} outliers detected by AH, ANH, and BayeScan. We also drew heatmaps to present the allele frequencies of the high F_{ST} outliers. The F_{ST} outliers detected by the three methods were used for further analysis.

2.5 | Genetic diversity and population structure

To detect the contribution of different loci to the diversity of this biological control agent, the total SNPs were further separated into three datasets, which included the high and low F_{ST} outliers and the remaining loci (assumed as neutral loci). Genetic diversity and population structure were analyzed based on the total loci, the neutral loci, and the high F_{ST} outlier loci.

We employed Arlequin to calculate genetic diversity parameters at each of the studied locations, including the percentage of polymorphic loci (Poly%), nucleotide diversity (Π), observed (H_O), and expected heterozygosity (H_E). Comparisons of these parameters between native and introduced locations were performed using t tests in SPSS 21 (IBM SPSS Statistical, Chicago, USA).

We calculated overall and pairwise F_{ST} values using Arlequin, with significance estimated based on 10,000 permutations. Then, population structure was studied through (a) the reconstruction of phylogenetic trees, (b) the estimation of ancestry proportions, and (c) an analysis of molecular variation (AMOVA). Neighbor-joining (NJ) trees were reconstructed in MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) using the p -distance between individuals and 1,000 bootstrap replicates. The software ADMIXTURE (Alexander, Novembre, & Lange, 2009) was employed to estimate ancestry proportions using different numbers of ancestral clusters (K). We explained the genetic structure by obtaining the lowest cross-validation error from $K = 2$ to 10. Then, AMOVA was performed in Arlequin with 10,000 permutations to further test for the population structure identified by clustering with ADMIXTURE.

3 | RESULTS

3.1 | SNP calling

A total of 504.25 million reads were obtained from the raw sequencing data. The sequencing quality of each sample is shown in Supporting Information Table S2. The number of reads per sample ranged from 1,142,115 to 7,711,651, with an average of 3,920,825 and a standard deviation of 1,396,936. The Q30 of each sample ranged from 93.53% to 96.56%, with an average of 95.38% and a standard deviation of 0.77%. The GC% of each sample ranged from 35.44% to 43.56%, with an average of 36.55% and a standard deviation of 1.13%. The above parameters suggested that all the sequencing data were of high-quality for further analysis.

TABLE 1 Procedures and results of SNP calling and filtering

Parameter	Value
SLAF tags	1,423,906
Poly SLAF tags	430,841 (30.26% of total SLAF tags)
Initial SNPs	3,670,447
SNPs present in >80% of individuals	689,925 (18.80% of initial SNPs)
SNPs with MAF >5%	285,397 (7.78% of initial SNPs)
SNPs with $H_O < 0.75$	285,072 (7.77% of initial SNPs)
One SNP per SLAF tag	53,181 (1.45% of initial SNPs)
SNPs in HWE	53,032 (1.44% of initial SNPs)
Mapped to the annotated transcriptome	10,507

Note. H_O : observed heterozygosity; HWE: Hardy–Weinberg equilibrium; MAF: minor allele frequency; Poly SLAF tags: polymorphic SLAF tags; SLAF: specific-locus amplified fragment.

Initially, 3,670,447 SNPs were detected among the 1,423,906 SLAF tags, and 53,032 of these loci were retained after filtering for integrity (excluding SNPs present in <80% of individuals), MAF, H_O , a single SNP per tag, and HWE (Table 1). These 53,032 SNPs were further employed for genetic analyses. Among these SNPs, 10,507 were mapped to the annotated transcriptome (Table 1) and were detected in every EuKaryotic Orthologous Groups (KOG) category (Supporting Information Figure S1).

3.2 | Detection of F_{ST} outliers

The three genome scanning methods detected a total of 521 high and 193 low F_{ST} outliers (Figure 2). Venn diagrams of the results (Figure 2d) showed that most of the high F_{ST} outliers detected by BayeScan were unique, whereas AH and ANH shared many high F_{ST} outliers. Among the three methods, 14 high F_{ST} outliers were shared. In contrast, 28 of the low F_{ST} outliers detected were common to the three methods.

The allele frequencies of the high F_{ST} outlier loci at the seven studied locations showed that most of the high F_{ST} outliers were caused by the drastically different allele frequencies in the five introduced populations (Figure 3a). A total of 114 high F_{ST} outliers were mapped to genes of the annotated transcriptome (Figure 3b), some of which

were related to the food source, including *fructose-2,6-bisphosphatase* (carbohydrate metabolism, with a dramatically different allele frequency in SA), *acetyl-CoA acyltransferase* (lipid metabolism, GT), *glycine-tRNA ligase* (amino acid metabolism, SY/SZ), and the *calcium-activated potassium channel slowpoke* (inorganic ion transport, SA).

3.3 | Genetic diversity and population structure

The genetic diversity parameters of each studied location are shown in Table 2 (all 53,032 loci) and Supporting Information Table S3 (52,318 neutral loci and 521 high F_{ST} outlier loci). The results based on the datasets containing all loci and the neutral loci only were similar, while the results based on the high F_{ST} outlier loci were quite different. Significant reductions in the number of polymorphic loci and nucleotide diversity were found when the introduced populations were compared to the native populations (p values ranged from 0.002 to 0.003 in Poly% and from 0.007 to 0.026 in Pi in the three datasets according to t tests). There was no significant difference in H_O and H_E values between the native and introduced locations.

The inferred population structure among *C. montrouzieri* samples was very similar between the dataset with all loci and only the neutral loci. In the dataset with all loci, we detected high differentiation among global populations (overall $F_{ST} = 0.323$). The pairwise F_{ST}

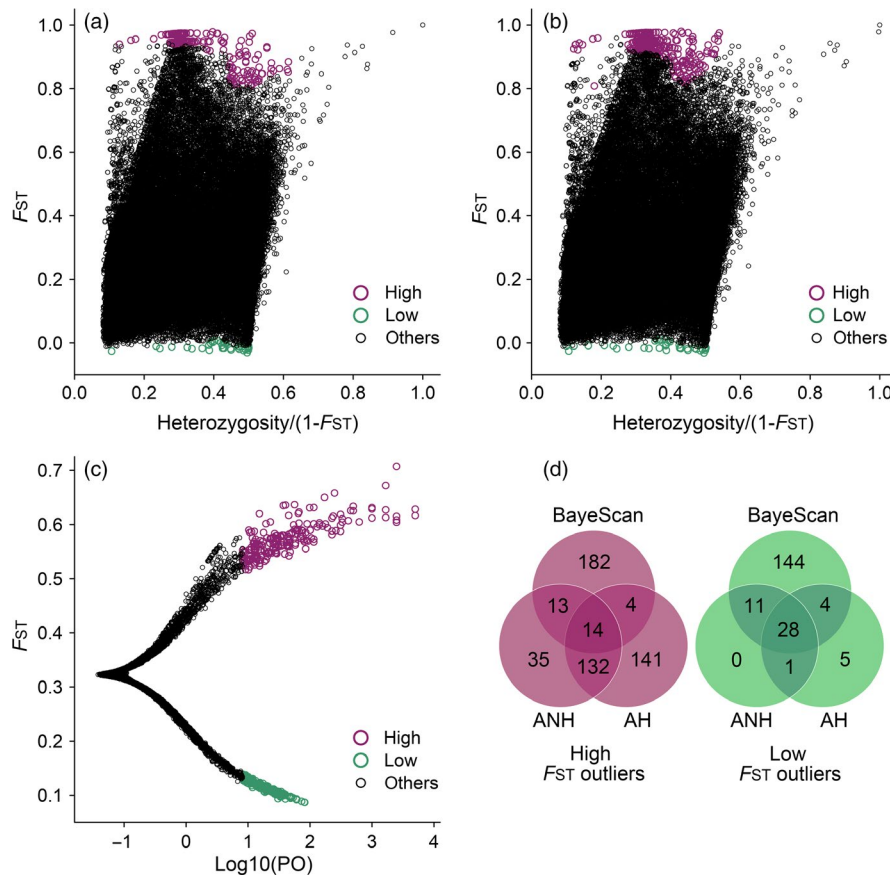


FIGURE 2 Detection of F_{ST} outliers based on three methods. The presented results are locus-specific F_{ST} values against heterozygosity from (a) the hierarchical (AH) and (b) nonhierarchical (ANH) island models implemented in Arlequin, (c) the posterior odds of the selection model ($\text{Log}_{10}(\text{PO})$) and the locus-specific F_{ST} value for each SNP according to BayeScan results. (d) Venn analysis of the above three methods for detecting high and low F_{ST} outliers

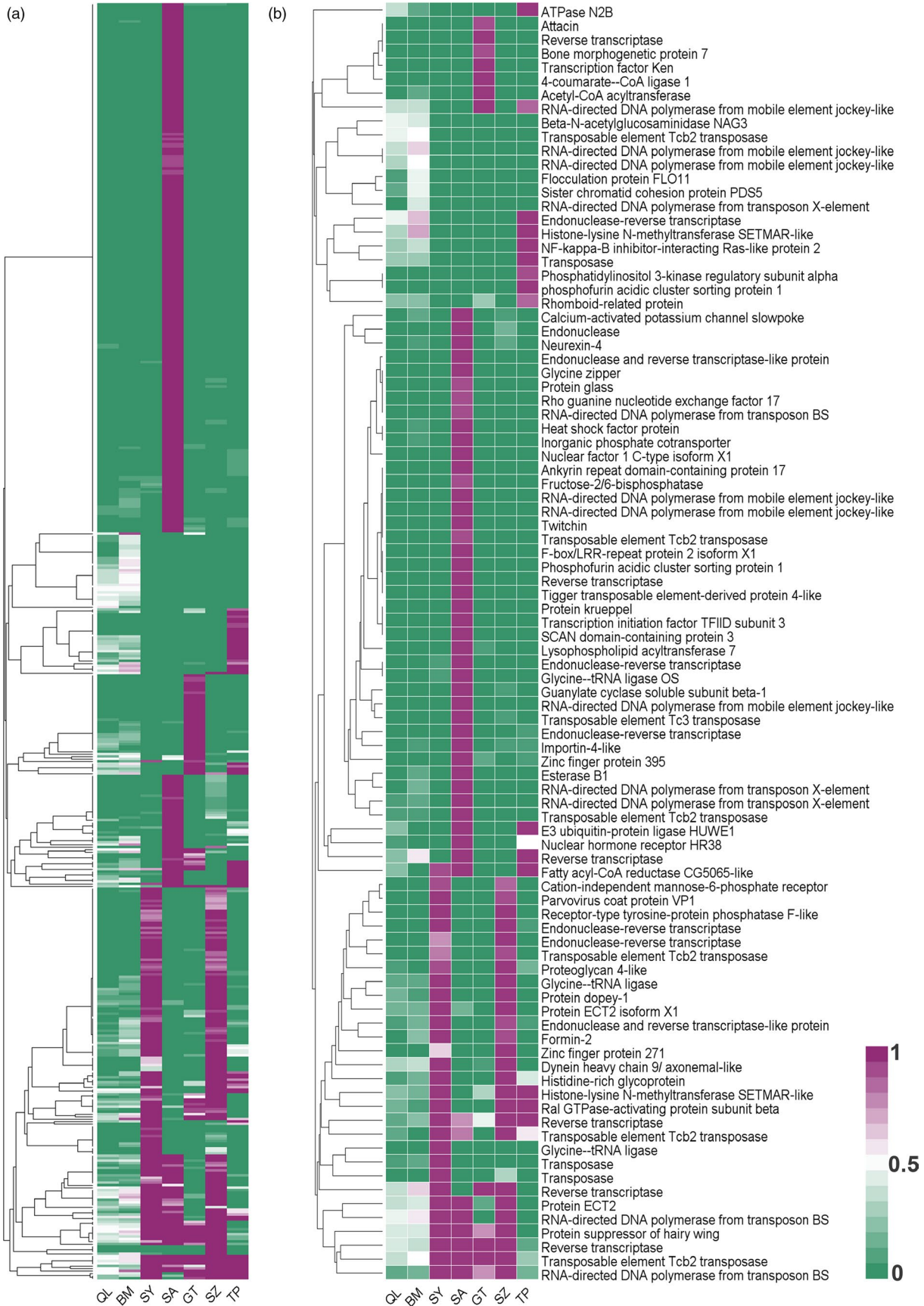


FIGURE 3 Heatmaps of the normalized allele frequencies of (a) the total 521 high F_{ST} outliers and (b) the 92 annotated high F_{ST} outliers. Color represents the percentage of minor allele of QL in each population from 0% (blue) to 100% (red)

TABLE 2 Genetic diversity at seven studied locations based on all 53,032 loci, including the percentage of polymorphic loci (Poly%), nucleotide diversity (Pi), observed heterozygosity (H_O), expected heterozygosity (H_E)

Location	Poly%	Pi	H_O	H_E
QL	89.1	0.244	0.229	0.274
BM	89.3	0.284	0.368	0.319
SY	61.4	0.195	0.372	0.318
SA	45.3	0.125	0.322	0.276
GT	63.9	0.205	0.297	0.320
SZ	59.8	0.189	0.309	0.315
TP	51.0	0.194	0.382	0.380
p in t -test	0.002*	0.026*	0.452	0.442

Note. BM: Canberra, ACT, Australia; GT: Ghent University (Belgium); QL: Brisbane, Queensland, Australia; SY/SA: Sun Yat-sen University; SZ: Shenzhen; TP: Taipei, Taiwan.

*Significant differences in means between native QL/BM and introduced SY/SA/GT/SZ/TP according to t tests.

values among the studied locations are shown in Table 3. The two native samples from Australia exhibited a low level of differentiation ($F_{ST} = 0.017$, $p < 0.001$), while all introduced populations showed much higher differentiation compared to the native populations (F_{ST} values from 0.113 to 0.319, $p < 0.001$). All of the introduced populations were also highly differentiated from each other (F_{ST} values from 0.207 to 0.495, $p < 0.001$). The NJ tree clustered individuals largely by population of origin, placing individuals from the native Australian range (QL, BM) in the middle of separate branches leading to the introduced populations SA, GT, and TP as well as the common branch to SY and SZ (Figure 4a).

ADMIXTURE analysis suggested that most of the tested individuals mainly comprised one ancestry. We obtained the lowest cross-validation error value for $K = 6$ (Figure 5a), with individuals from each introduced population belonging to a unique ancestral cluster, while those from the two native locations shared one genetic cluster (Figure 5b). However, the cross-validation error values for $K = 4$ – 7 were quite close. For $K = 5$, SY and SZ shared the same ancestral cluster, while for $K = 4$, QL+BM and TP were also grouped together. For $K = 7$, SA appeared to be an admixed population of two ancestral clusters. However, a too-large K value may have led

to an overfitting effect, which forced the algorithm to subdivide the populations. The percentage of variation associated with differences among groups (F_{CT}) was higher and significant for $K = 6$ structuring (26.94%, $p = 0.0489$, Supporting Information Table S4) than for $K = 5$ (1.92%, $p = 0.4360$, Supporting Information Table S4) and $K = 4$ (4.29%, $p = 0.2776$, Supporting Information Table S4). However, these AMOVAs had limited power since some of the groups were formed by one location only.

The high F_{ST} outlier loci generated a population structure that was in many aspects qualitatively similar but quantitatively stronger. The NJ tree showed that SA and SY+SZ were extremely divergent from the other populations (Figure 4c), with F_{ST} values ranging from 0.884 to 0.966 between SA and the other populations and from 0.613 to 0.960 between SY+SZ and the other populations (Supporting Information Table S5). The ADMIXTURE analysis showed that $K = 7$ resulted in the lowest cross-validation error (Figure 5e), with the individuals from the same locations belonging to a unique cluster. For $K = 6$, the subdivision of SA was possibly an artefactual effect. The AMOVAs showed that the $K = 5$ and $K = 4$ structuring explained nearly none of the total variance among populations (Supporting Information Table S4).

4 | DISCUSSION

4.1 | Genomic changes in biological control agents arising from introduction

In this study, we demonstrated very high genetic differentiation among introduced populations of the ladybird beetle *C. montrouzieri*. Our results based on a large number of genome-wide SNP markers are overall consistent with the differentiation patterns in this system detected with a few SSR markers (Li et al., 2017). More generally, high genetic differentiation based on traditional genetic markers (e.g., SSRs) is commonly detected among different introduced populations of biological control agents (Lombaert et al., 2014; Retamal et al., 2016; Sethuraman, Janzen, & Obrycki, 2015). In theory, genetic differentiation among populations can be largely attributed to genetic drift or selection, and the repeated process of establishment of introduced populations from native ones or other introduced populations may contribute to the strength of these effects (Fauvergue et

TABLE 3 Pairwise F_{ST} based on all 53,032 loci. All F_{ST} values were significantly different for zero with $p < 0.001$

	QL	BM	SY	SA	GT	SZ	TP
QL							
BM	0.017						
SY	0.213	0.188					
SA	0.319	0.315	0.465				
GT	0.181	0.172	0.348	0.453			
SZ	0.225	0.194	0.207	0.495	0.361		
TP	0.126	0.113	0.309	0.456	0.288	0.327	

Note. BM: Canberra, ACT, Australia; GT: Ghent University (Belgium); QL: Brisbane, Queensland, Australia; SY/SA: Sun Yat-sen University; SZ: Shenzhen; TP: Taipei, Taiwan.

al., 2012). Because environmental conditions (e.g., temperature and diet) differ from native conditions of *C. montrouzieri*, we considered that not only genetic drift but also selection might have contributed to the observed differences between populations. Our reduced-representation genomic sequencing approach provides a relatively powerful dataset for a better understanding of the extent and the forces contributing to the genetic changes in the introduction history of this biological control agent.

4.2 | Role of genetic drift and selection in genomic population changes

The studied introduced populations were artificially transferred from the native areas in Australia or introduced populations on other continents. This intentional introduction process shows some similarities to biological invasions (Blackburn, Lockwood, & Cassey, 2015) but could potentially exhibit even more extreme founder effects due to the total isolation of the introduced populations from the native populations. In this study, a significant decrease in genetic diversity was detected in the introduced populations, which

is consistent with founder effects together with strong genetic drift during the early stages of introduction. Additionally, our pairwise F_{ST} results, NJ trees, and clustering results supported the notion that all the introduced populations are deeply divergent from the native populations. Since these patterns are very similar in the datasets of all loci and only neutral loci, we suggest that genetic drift—in addition to allele frequency changes during the founding event—was the main force shaping the genetic diversity and population structure of the introduced populations of this biological control agent. Only ~100 years has passed since the introduction of *C. montrouzieri* for biological control was initiated (Kairo et al., 2013), but the production of more than 10 generations per year can contribute to rapid genomic changes. We also found that some alleles in the introduced populations were not present in the native locations. Apart from probably rare de novo mutations, it is possible that these alleles are too rare to have been picked up in the limited sampled individuals within a population. An alternative explanation is that some of these genomic differences go back to naturally existing differentiation in the native range of this species, rather than being caused by human activities. Our analyses of native samples from two distinct locations

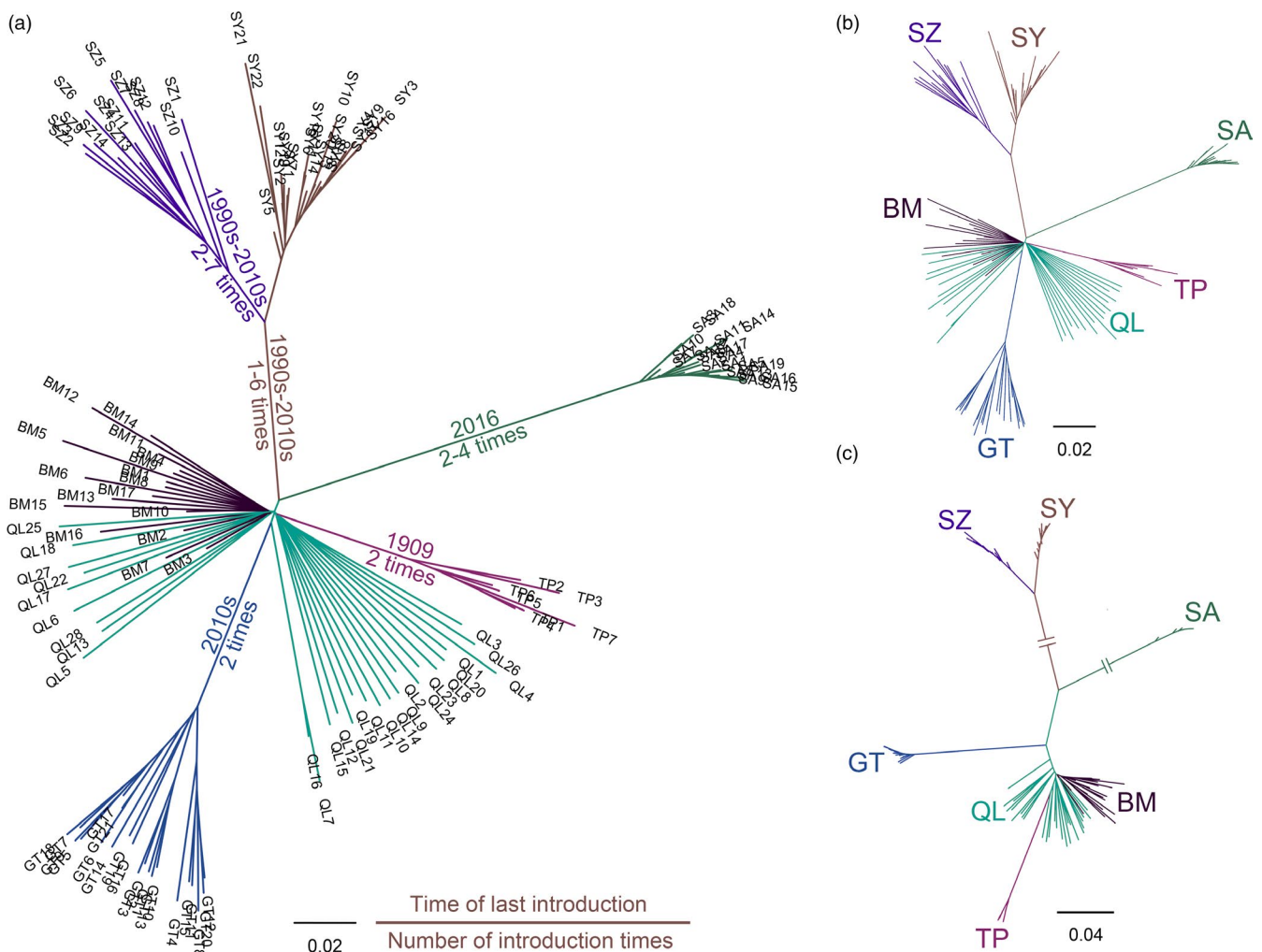


FIGURE 4 Neighbor-joining trees of *Cryptolaemus montrouzieri* individuals based on (a) all 53,032 loci, (b) 52,319 neutral loci, and (c) 521 high F_{ST} outlier loci. Information including the time of the last introduction and the number of introduction events for each introduced population is provided on the tree branches

approximately 1,240 km apart showed only very slight genomic differentiation—particularly in relation to the introduced populations. It would require a detailed survey of the natural range (Zemanova, Broennimann, Guisan, Knop, & Heckel, 2018; Zemanova, Knop, & Heckel, 2016) to determine the full extent of natural genetic variation and differentiation but given all other factors, most of the genetic differences detected here arose most likely during and/or after introduction.

Individuals introduced into new areas are usually subject to different environmental conditions and selection pressures than in their native ranges. Thus, signals of selection in fitness-related traits/genes should be expressed during the local adaptation process. Recently, F_{ST} outlier methods have been widely used in population genomic analyses for detecting candidate loci under selection (Carreras et al., 2017; Jeffery et al., 2017; Pais, Whetten, & Xiang, 2017; Van Wyngaarden et al., 2017). In this study, we detected several high and low F_{ST} outliers through three alternative methods. High F_{ST} outliers are usually considered candidate loci or linked to loci under positive selection, while low F_{ST} outliers might indicate negative or balancing selection (Beaumont & Nichols, 1996; Fischer, Foll, Heckel, & Excoffier, 2014). However, with the

increase in differentiation caused by genetic drift, loci with higher F_{ST} values can be mistaken for evidence of positive selection, since the two mechanisms produce similar signals in genomic variation (Cruickshank & Hahn, 2014; Currat et al., 2006; Vera, Díez-del-Molino, & García-Marín, 2016). In our case, the detected high F_{ST} outliers should be considered as only very tentative evidence for positive selection because of the confounding effects of very high differentiation among the studied locations. Nevertheless, some of the high F_{ST} outliers were mapped to genes related to functions such as nutrient transport and metabolism. Irrespective of whether the extreme changes in allele frequencies at these loci were caused by genetic drift or positive selection or both, the number and extent of genetic differences between populations make it unlikely that phenotypic traits of these beetles remained completely unaffected.

4.3 | Implications for biological control

The extent of heritable phenotypic differences between the introduced or native populations of *C. montrouzieri* has not been characterized at a larger scale, but there is likely variation in

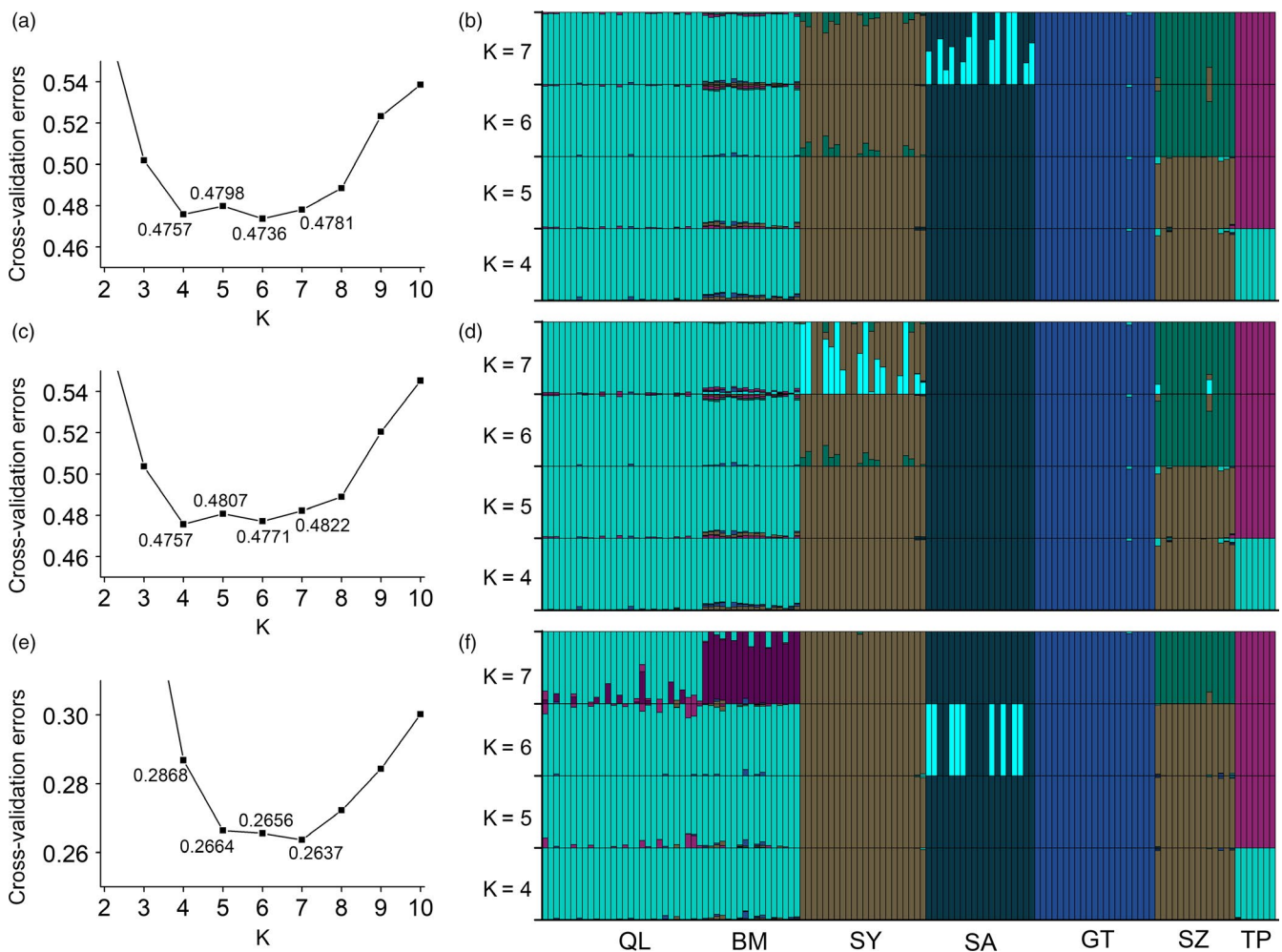


FIGURE 5 Ancestral clustering using ADMIXTURE based on all 53,032 loci (a, b), 52,318 neutral loci (c, d), and 521 high F_{ST} outlier loci (e, f). The cross-validation error of different numbers of ancestral clusters (K) and the individual ancestry bar plots for $K = 4-7$ are shown

the performance as a biological control agent. In fact, ladybirds from the SY population have a significantly longer developmental time and perform better under cold stress but worse under starvation than beetles from SA when tested under controlled laboratory conditions (Li et al., 2018). The large genetic differentiation between these and among the other populations indicates several potential issues related to biological control applications.

First, intra-specific variation might complicate the assessment of biological control efficiency and invasion potential. The evolution of biological control agents not only influences their effectiveness in pest control but also generates potential side effects on local biodiversity (Roderick, Hufbauer, & Navajas, 2012). Such changes could certainly result from phenotypic plasticity that enables ladybirds to respond to varying environments (Murren et al., 2015). However, relatively often, the successful establishment of introduced species has been associated with rapid genetic changes. These genetic changes could be caused by genetic drift, as most likely in the present study, and/or adaptive responses to novel selection pressures (Fauvergue et al., 2012). For example, in a parasitoid wasp introduced to New Zealand from South America for the control of weevil pests, changes in biotype frequency and better performance were detected after 10 years, indicating local adaptation (Phillips et al., 2008). Further studies should take the potential for intra-specific variation and different reactions within *C. montrouzieri* into account.

Second, the extent of the detected genetic changes might indicate the feasibility of selective breeding for improvement of traits of interest (Lommen et al., 2017). Using methods such as genome-wide association studies (GWAS), it is possible to identify and characterize loci affecting target traits for further genomic-based selection (Visscher, Brown, McCarthy, & Yang, 2012). For example, similar population genomic techniques are being applied to parasitoid wasps, predatory bugs, and predatory mites in the Breeding Invertebrates for Next Generation BioControl Training Network (BINGO-ITN, <http://www.bingo-itn.eu/en/bingo.htm>). Our genome-wide description of polymorphisms in the ladybird *C. montrouzieri* demonstrates ample genetic variation and thus suggests that such approaches might be promising avenues for further improvements of the performance of this biological control agent.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 31572052, 31702012), Science and Technology Planning Project of Guangdong Province, China (Grant No. 2017B020202006), National Key R&D Program of China (Grant No. 2017YFD0201000), and Open Project of the State Key Laboratory of Biocontrol (Grant No. 2018-04). We would like to thank Shang-Jun Zou for assistance with the experiments. We would also like to thank Mary Finlay-Doney and Cody Murray for the help of collecting samples.

CONFLICT OF INTEREST

None declared.

DATA ARCHIVING STATEMENT

All raw sequences are available in the NCBI SRA databases (BioProject: PRJNA445430, BioSample: SAMN08784023–SAMN08784122 and SAMN09702120–SAMN09702148). The filtered 53,032 SNPs used in this study are available at Dryad (<https://doi.org/10.5061/dryad.qg215qn>).

ORCID

Hong Pang  <https://orcid.org/0000-0002-0585-3577>

REFERENCES

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Andraca-Gómez, G., Ordano, M., Boege, K., Domínguez, C. A., Piñero, D., Pérez-Ishiwara, R., ... Fornoni, J. (2014). A potential invasion route of *Cactoblastis cactorum* within the Caribbean region matches historical hurricane trajectories. *Biological Invasions*, *17*, 1397–1406. <https://doi.org/10.1007/s10530-014-0802-2>
- Annecke, D. P., Karny, M., & Burger, W. A. (1969). Improved biological control of the prickly pear, *Opuntia megacantha* Salm-Dyck, in South Africa through the use of an insecticide. *Phytophylactica*, *1*, 9–13.
- Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B: Biological Sciences*, *263*, 1619–1626.
- Blackburn, T. M., Lockwood, J. L., & Cassey, P. (2015). The influence of numbers on invasion success. *Molecular Ecology*, *24*, 1942–1953. <https://doi.org/10.1111/mec.13075>
- Carreras, C., Ordonez, V., Zane, L., Kruschel, C., Nasto, I., Macpherson, E., & Pascual, M. (2017). Population genomics of an endemic Mediterranean fish: Differentiation by fine scale dispersal and adaptation. *Scientific Reports*, *7*, 43417. <https://doi.org/10.1038/srep43417>
- Cruikshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*, 3133–3157. <https://doi.org/10.1111/mec.12796>
- Curat, M., Excoffier, L., Maddison, W., Otto, S. P., Ray, N., Whitlock, M. C., & Yeaman, S. (2006). Comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in *Homo sapiens*" and "Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans". *Science*, *313*, 172; author reply 172.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*, 499–510. <https://doi.org/10.1038/nrg3012>
- De Clercq, P., Mason, P. G., & Babendreier, D. (2011). Benefits and risks of exotic biological control agents. *BioControl*, *56*, 681–698. <https://doi.org/10.1007/s10526-011-9372-8>
- de Villemereuil, P., Frichot, E., Bazin, E., Francois, O., & Gaggiotti, O. E. (2014). Genome scan methods against more complex models: When and how much should we trust them? *Molecular Ecology*, *23*, 2006–2019.

- Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*, 285–298. <https://doi.org/10.1038/hdy.2009.74>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Fauvergue, X., Vercken, E., Malausa, T., & Hufbauer, R. A. (2012). The biology of small, introduced populations, with special reference to biological control. *Evolutionary Applications*, *5*, 424–443. <https://doi.org/10.1111/j.1752-4571.2012.00272.x>
- Fischer, M. C., Foll, M., Heckel, G., & Excoffier, L. (2014). Continental-scale footprint of balancing and positive selection in a small rodent (*Microtus arvalis*). *PLoS ONE*, *9*, e112332. <https://doi.org/10.1371/journal.pone.0112332>
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993. <https://doi.org/10.1534/genetics.108.092221>
- 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. <https://doi.org/10.1016/j.crvi.2010.12.008>
- Jeffery, N. W., DiBacco, C., Van Wyngaarden, M., Hamilton, L. C., Stanley, R. R. E., Bernier, R., ... Bradbury, I. R. (2017). RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. *Ecology and Evolution*, *7*, 2513–2524.
- Kairo, M. K. T., Paraiso, O., Gautam, R. D., & Peterkin, D. D. (2013). *Cryptolaemus montrouzieri* (Mulsant) (Coccinellidae: Scymninae): A review of biology, ecology, and use in biological control with particular reference to potential impact on non-target organisms. *CAB Reviews*, *8*, 005. <https://doi.org/10.1079/PAVSNNR20138005>
- Kajita, Y., O'Neill, E. M., Zheng, Y., Obrycki, J. J., & Weisrock, D. W. (2012). A population genetic signature of human releases in an invasive ladybeetle. *Molecular Ecology*, *21*, 5473–5483. <https://doi.org/10.1111/mec.12059>
- Kent, W. J. (2002). BLAT – The BLAST-like alignment tool. *Genome Research*, *12*, 656–664. <https://doi.org/10.1101/gr.229202>
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, *79*, 5112–5120. <https://doi.org/10.1128/AEM.01043-13>
- Kumschick, S., Richardson, D. M., & Kueffer, C. (2013). Species-based risk assessments for biological invasions: Advances and challenges. *Diversity and Distributions*, *19*, 1095–1105. <https://doi.org/10.1111/ddi.12110>
- Li, H.-S., Liang, X.-Y., Zou, S.-J., Liu, Y., De Clercq, P., Ślipiński, A., & Pang, H. (2016). Episodic positive selection at mitochondrial genome in an introduced biological control agent. *Mitochondrion*, *28*, 67–72. <https://doi.org/10.1016/j.mito.2016.03.005>
- Li, H.-S., Liang, X.-Y., Zou, S.-J., Liu, Y., De Clercq, P., Ślipiński, A., & Pang, H. (2017). New EST-SSR markers reveal strong genetic differentiation in native and introduced populations of the mealybug destroyer *Cryptolaemus montrouzieri*. *Biological Control*, *109*, 21–26. <https://doi.org/10.1016/j.biocontrol.2017.03.008>
- Li, H.-S., Pan, C., De Clercq, P., Ślipiński, A., & Pang, H. (2016). Variation in life history traits and transcriptome associated with adaptation to diet shifts in the ladybird *Cryptolaemus montrouzieri*. *BMC Genomics*, *17*, 281. <https://doi.org/10.1186/s12864-016-2611-8>
- Li, H.-S., Zou, S.-J., De Clercq, P., & Pang, H. (2018). Population admixture can enhance establishment success of the introduced biological control agent *Cryptolaemus montrouzieri*. *BMC Evolutionary Biology*, *18*, 36. <https://doi.org/10.1186/s12862-018-1158-5>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Proc, G. P. D. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- 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>
- Lommen, S. T., Jong, P. W., & Pannebakker, B. A. (2017). It is time to bridge the gap between exploring and exploiting: Prospects for utilizing intraspecific genetic variation to optimize arthropods for augmentative pest control—a review. *Entomologia Experimentalis Et Applicata*, *162*, 108–123. <https://doi.org/10.1111/eea.12510>
- Losey, J., Perlman, J., Kopco, J., Ramsey, S., Hesler, L., Evans, E., ... Smyth, R. (2012). Potential causes and consequences of decreased body size in field populations of *Coccinella novemnotata*. *Biological Control*, *61*, 98–103. <https://doi.org/10.1016/j.biocontrol.2011.12.009>
- Maes, S., Grégoire, J.-C., & De Clercq, P. (2014). Prey range of the predatory ladybird *Cryptolaemus montrouzieri*. *BioControl*, *59*, 729–738. <https://doi.org/10.1007/s10526-014-9608-5>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*, 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Milligan, B. G. (1988). Total DNA isolation. In A. R. Hoelzel (Ed.), *Molecular genetic analysis of populations* (pp. 29–60). Oxford, UK: Oxford University Press.
- Murren, C. J., Auld, J. R., Callahan, H., Ghalambor, C. K., Handelsman, C. A., Heskell, M. A., ... Maughan, H. (2015). Constraints on the evolution of phenotypic plasticity: Limits and costs of phenotype and plasticity. *Heredity*, *115*, 293. <https://doi.org/10.1038/hdy.2015.8>
- Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*(Suppl 1), 184–194.
- Pais, A. L., Whetten, R. W., & Xiang, Q.-Y.-J. (2017). Ecological genomics of local adaptation in *Cornus florida* L. by genotyping by sequencing. *Ecology and Evolution*, *7*, 441–465.
- Phillips, C. B., Baird, D. B., Iline, I. I., McNeill, M. R., Proffitt, J. R., Goldson, S. L., & Kean, J. M. (2008). East meets west: Adaptive evolution of an insect introduced for biological control. *Journal of Applied Ecology*, *45*, 948–956. <https://doi.org/10.1111/j.1365-2664.2008.01461.x>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575. <https://doi.org/10.1086/519795>
- Retamal, R., Zaviezo, T., Malausa, T., Fauvergue, X., Le Goff, I., & Toleubayev, K. (2016). Genetic analyses and occurrence of diploid males in field and laboratory populations of *Masturus ridens* (Hymenoptera: Ichneumonidae), a parasitoid of the codling moth. *Biological Control*, *101*, 69–77. <https://doi.org/10.1016/j.biocontrol.2016.06.009>
- Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., Beeman, R. W., ... Grimmelikhuijzen, C. J. (2008). The genome of the

- model beetle and pest *Tribolium castaneum*. *Nature*, 452, 949–955. <https://doi.org/10.1038/nature06784>
- Roderick, G. K., Hufbauer, R., & Navajas, M. (2012). Evolution and biological control. *Evolutionary Applications*, 5, 419–423. <https://doi.org/10.1111/j.1752-4571.2012.00281.x>
- Roy, H. E., Brown, P. M. J., Adriaens, T., Berkvens, N., Borges, I., Clusella-Trullas, S., ... Zhao, Z. (2016). The harlequin ladybird, *Harmonia axyridis*: Global perspectives on invasion history and ecology. *Biological Invasions*, 18, 997–1044. <https://doi.org/10.1007/s10530-016-1077-6>
- Sethuraman, A., Janzen, F. J., & Obrycki, J. (2015). Population genetics of the predatory lady beetle *Hippodamia convergens*. *Biological Control*, 84, 1–10. <https://doi.org/10.1016/j.biocontrol.2015.01.002>
- Ślipiński, A. (2007). *Australian ladybird beetles (Coleoptera: Coccinellidae): Their biology and classification*. Canberra, Australia: Australian Biological Resources Study.
- Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., ... Zheng, H. (2013). SLAF-seq: An efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE*, 8, e58700. <https://doi.org/10.1371/journal.pone.0058700>
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- van de Zande, L., Ferber, S., de Haan, A., Beukeboom, L. W., van Heerwaarden, J., & Pannebakker, B. A. (2014). Development of a *Nasonia vitripennis* outbred laboratory population for genetic analysis. *Molecular Ecology Resources*, 14, 578–587.
- Van Wyngaarden, M., Snelgrove, P. V., DiBacco, C., Hamilton, L. C., Rodriguez-Ezpeleta, N., Jeffery, N. W., ... Bradbury, I. R. (2017). Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RADseq-derived SNPs. *Evolutionary Applications*, 10, 102–117.
- Vera, M., Díez-del-Molino, D., & García-Marín, J.-L. (2016). Genomic survey provides insights into the evolutionary changes that occurred during European expansion of the invasive mosquitofish (*Gambusia holbrooki*). *Molecular Ecology*, 25, 1089–1105.
- Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, 90, 7. <https://doi.org/10.1016/j.ajhg.2011.11.029>
- White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, 22, 2971–2985.
- Zemanova, M. A., Broennimann, O., Guisan, A., Knop, E., & Heckel, G. (2018). Slimy invasion: Climatic niche and current and future biogeography of Arion slug invaders. *Diversity and Distributions*, 24, 1627–1640.
- Zemanova, M. A., Knop, E., & Heckel, G. (2016). Phylogeographic past and invasive presence of Arion pest slugs in Europe. *Molecular Ecology*, 25, 5747–5764.
- Zhang, Y., Jiang, R., Wu, H., Liu, P., Xie, J., He, Y., & Pang, H. (2012). Next-generation sequencing-based transcriptome analysis of *Cryptolaemus montrouzieri* under insecticide stress reveals resistance-relevant genes in ladybirds. *Genomics*, 100, 35–41. <https://doi.org/10.1016/j.ygeno.2012.05.002>
- Zimmermann, H., Bloem, S., & Klein, H. (2004). *Biology, history, threat, surveillance and control of the cactus moth, Cactoblastis cactorum*. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).

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: Li H-S, Heckel G, Huang Y-H, Fan W-J, Ślipiński A, Pang H. Genomic changes in the biological control agent *Cryptolaemus montrouzieri* associated with introduction. *Evol Appl*. 2019;12:989–1000. <https://doi.org/10.1111/eva.12774>