ddRAD Sequencing: A Novel Arsenal Added to the Biosecurity Toolbox to Trace the Origin of Brown Marmorated Stink Bug, Halyomorpha Halys (Hemiptera: Pentatomidae)

Juncong Yan  
MPI  https://orcid.org/0000-0001-5074-8795

Gábor Vétek  
Szent Istvan university

Chandan Pal  
MPI

Jinping Zhang  
CABI China

Rania Gmati  
Szent Istvan university

Qinghai Fan  
MPI

Disna Gunawardana  
MPI

Allan Burne  
MPI

Diane Anderson  
MPI

Rebijith Kayattukandy Balan  
MPI

Sherly George  
MPI

Péter Farkas  
Szent Istvan university

Dongmei Li  (Dongmei.Li@mpi.govt.nz)  
Ministry for Primary Industries  https://orcid.org/0000-0002-0509-3221

Research article

Keywords: BMSB, SNP, population genetics, invasion, biosecurity, ddRADSeq, restriction digestion

DOI: https://doi.org/10.21203/rs.3.rs-63068/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Brown marmorated stink bug (BMSB), *Halyomorpha halys* (Hemiptera: Pentatomidae) is native to East Asia but has invaded many countries in the world. It is a polyphagous insect pest and causing significant economic losses to agriculture worldwide. Knowledge on the genetic diversity among BMSB populations is scarce but is essential to understand the patterns of colonization and invasion history of local populations. Efforts have been made to assess the genetic diversity of BMSB using partial mitochondrial DNA sequences but genetic divergence on mitochondrion is not high enough to precisely identify and distinguish various BMSB populations. Therefore, in this study, we applied a ddRAD (double digest restriction-site associated DNA) sequencing approach to ascertain the genetic diversity of BMSB populations collected from 12 countries (2 native and 10 invaded) across four continents with the ultimate aim to trace the origin of BMSBs intercepted during border inspections and post-border surveillance.

Result

A total of 1775 high confidence single nucleotide polymorphisms (SNPs) were identified from ddRAD sequencing data collected from 389 BMSB individuals. Principal component analysis (PCA) of the identified SNPs indicated the existence of two main distinct genetic clusters representing individuals sampled from regions where BMSB is native to, China and Japan, respectively, and one broad cluster comprised individuals sampled from countries which have been invaded by BMSB. The population genetic structure analysis further discriminated the genetic diversity among the BMSB populations at a higher resolution and distinguished them into five potential genetic clusters.

Conclusion

The study revealed hidden genetic diversity among the studied BMSB populations across the continents. The BMSB populations from Japan were genetically distant from the other studied populations. Similarly, the BMSB populations from China were also separated from the Japanese and other populations. Further genetic structure analysis revealed the presence of at least three genetic clusters of BMSB in the invaded countries, possibly originating from multiple invasions. Furthermore, this study has produced novel set of SNP markers to enhance knowledge of genetic diversity among BMSB populations and demonstrate a great potential to trace the origin of BMSB individuals for future invasion events.

Background

The brown marmorated stink bug (BMSB), *Halyomorpha halys* (Stål, 1855) (Hemiptera: Pentatomidae) is a polyphagous pest with a wide host range [1]. It can cause extensive damage to agricultural crops worldwide [2, 3], and in 2010 alone was responsible for a loss of more than 37 million USD in agricultural products in North America [4]. The native range of BMSB is China (including Taiwan), Japan, and the Korean peninsula [5-7]. To date, BMSB has been reported from more than 30 countries [8], including almost all states in the USA [2, 4], multiple countries in Europe [9-16] and Chile [17]. Climate modelling studies indicates its potential range to develop could even expand further that includes South and Central America, Southern Africa, Southern Australia, and the North Island of New Zealand [9, 18].

BMSB was first intercepted at the New Zealand border in 2005 [19], and since then the frequency of interceptions have been increasing, due to the rise of international travelling and trade [20]. There have been 1620 recorded interceptions of BMSB since 2005 at the New Zealand border [19]. In the past decades, BMSB has been invaded and established in a range of countries irrespective of the environmental conditions [2, 4, 10-17, 21]. Adaptive evolutionary changes and/or ecological adaptation in a new region has made this pest a successful global invader and its recent invasion history can shed light on that. However, an in-depth genetic information of BMSB at the population level is scarce. Such information on the genetic diversity of BMSB can enhance our understanding on their population structure and global invasion history. This could also assist in constructing a global genetic population structure of BMSB and develop a potential strategy to trace the country of origin for BMSB individuals intercepted at the border or in post-border scenarios in biosecurity settings.

Up to date, nearly all published studies for tracing the origin of BMSB utilised PCR based molecular methods and focused on small regions on mitochondrial DNA (mtDNA), such as the COI (Cytochrome c oxidase I) and/or COII (Cytochrome c oxidase II) genes [16, 21-24]. mtDNA is highly variable between species and can potentially provide sufficient resolution to identify genetic differences between species [25]. Since mtDNA is inherited maternally and lacks recombination, the resolution of mitochondria-derived genetic divergence is not sufficient to differentiate between individuals in a population [26]. Therefore, there is the need to study the genome-wide, high-resolution markers among BMSB populations from their native and invaded regions. The study will be able to discern genetically distinct populations thus allowing us to trace the geographical origin of BMSBs within an interception scenario. This calls for an innovative method to explore the genetic diversity within BMSB populations on a whole genome scale.
The detection of genetic markers is crucial for genetic diversity study. Recently, a high-throughput sequencing-based method (HTS) has replaced traditional gel-based experiment to discover genetic marker [27]. RADseq (Restriction-site associated DNA sequencing), is often applied for genome-wide SNP (single nucleotide polymorphism) identification in large genomes because of its relatively low cost and high-throughput [28]. The RADseq technique utilises one (or more) restriction enzyme to digest the whole genome into short genomic fragments that are then subjected to high-throughput DNA sequencing [28]. Restriction site-associated DNA markers provide a well-established basis for population genetics, as they are sensitive to both SNPs and insertion or deletion events (indels) in genomes [29]. So far, RADseq has been widely used in population genetic studies for many taxa including plants [30], and animals [28]. Double digest restriction-site associated DNA (ddRAD) sequencing is using two restriction enzymes to allow greater control of the genomic regions sampled for sequencing and allowing more reproducible recovery of sequenced regions [31]. Therefore, in this study, we applied ddRADseq to explore the genetic diversity among BMSB specimens collected from 41 populations across 12 countries.

Results

EcoR I-Msp I restriction enzyme pair was suitable for ddRAD sequencing

To select the most suitable restriction enzyme (RE) pairs for digesting BMSB genomes, in silico test using 15 combinations of RE against the BMSB scaffold were conducted. The prediction revealed that more than 100K fragments produced from most of the RE pairs selected except the pairs, Msel-MluCl, Msel-MluCl and EcoRI-PstI (Table 1). Since the genome used for the test was a scaffold, thus the simulation results might not reflect the real situation. A pilot ddRADseq in vitro experiment was conducted with genomic DNA samples derived from two BMSB individuals. Of the 15 pairs of REs used for the in-silico test, nine different pairs of RE were selected for ddRADseq. After the HiSeq run, approximately, 2 Gb of raw RADseq sequencing data were generated for each individual. EcoR I-Msp I restriction enzyme pair recovered the highest number of genetic variances (i.e. SNPs) after SNP high standard quality control (QC) filtering, thus selected as the most suitable pairs of restriction enzymes for digesting the BMSB genome via ddRAD sequencing (Table 1).

ddRAD sequencing statistics and SNPs estimation

In total, 399 ddRAD sequencing datasets were obtained from the BMSB individuals, which yielded 3.6 billion of raw paired end reads (2x150 bp) (min: 4 million, max: 40 million and median: 7.6 million). On average, 9 million reads were generated for each individual. Using quality-trimming of the sequence data, 387,629 SNPs were estimated from 399 BMSB individuals. A high standard QC criterion was applied for filtering the SNPs, and only those loci that were shared by all the individuals were retained. As a result, the sequences from 389 individuals passed QC with 1775 high confidence biallelic SNPs that were used for the subsequent analysis of genomic diversity and population structure.

Genetic clusters were observed among the BMSB populations

At least three genetic clusters comprising China, Japan, and the invaded countries (Austria, Chile, Georgia, Hungary, Italy, Romania, Serbia, Slovenia, Turkey, and the USA) were revealed by Principal component analysis (PCA) (Figure 1). All BMSB individuals from Japan formed an isolated cluster, whereas BMSBs collected from the invaded countries were genetically closer to those of China. Interestingly, one BMSB individual from Chile was clustered together with the Chinese populations whereas the rest of the individuals were mixed into the individuals from Europe/USA populations (Figure 1).

Individuals from the same geographical region were genetically linked

To further emphasise the outcome of genetic clustering pattern via principal component analysis, minimum spanning networks (MSN) were constructed using SNPs profile of each individual, and their genetic variability were visualised among the population lineages (Figure 2). The MSN showed that all the individuals from China are genetically linked together in the network, which also applies for the individuals from Japan (Figure 2). There was a genetic divergence among the BMSB individuals from native regions of China and Japan, while those of invaded countries were more closely related in the network. Similar to the PCA results, one individual from Chile was found in the same clade of the Chinese samples, suggesting that this BMSB specimen might have originated from a recent invasion from China. The rest of the Chilean sample were distantly related from those in China and Japan but were more closely related to the samples from the European/USA groups, indicating that those possibly originated from secondary invasions from European/USA regions. The MSN also showed that one individuals from Italy and three from Slovenia were genetically linked to the Chinese populations, whereas the rest from these two countries were more closely related to those from European and the USA, suggesting multiple invasion might have been occurred (Figure 2).

Genetic distance between native populations of China and Japan was relatively higher

Population genetic divergence in the form of pairwise $F_{ST}$ revealed significant ($p<0.05$) genetic difference (except for that between China and Serbia) among 12 geographical groups or countries, with $F_{ST}$ value ranging from 0.00055 between BMSB populations from Hungary and Serbia, to 0.20842 between BMSB populations from Japan and Romania (Table 2). We also observed that the genetic distance between native
populations of China and Japan was moderately higher ($F_{ST}=0.08$) than that between the populations of China and many other BMSB-invaded countries, such as Slovenia ($F_{ST}=0.03792$). Similarly, the genetic distance between the invaded populations in the USA and Chile was relatively low ($F_{ST}=0.0393$) compared to genetic distance between BMSB populations in Chile and the native regions, China ($F_{ST}=0.09844$) and Japan ($F_{ST}=0.17654$). Moreover, the $F_{ST}$ between the BMSB populations from the neighbouring countries was very small, for example, Turkey and Georgia ($F_{ST}=0.01651$), Austria and Slovenia ($F_{ST}=0.02033$), Hungary and Serbia ($F_{ST}=0.00055$) (Table 2). A Neighbour-Joining (NJ) tree constructed using the $F_{ST}$ pairwise values among the individuals from the 12 countries revealed the similar relationships among the BMSB populations from the 12 countries (Figure 4). The tree depicted the overall relationships of the populations and showed that Chinese and Japanese populations were clustered together while the others were formed into a clade (Figure 4).

**Five genetic clusters exist in the BMSB populations**

Furthermore, insights into the BMSB genetic diversity were achieved by population genetic structure analysis using fastSTRUCTURE. This analysis expanded the results of PCA (Figure 1) and provided more in-depth clustering for the BMSB populations from the invaded countries. It predicts the presence of at least three genetic clusters within the BMSB-invaded countries (Figure 3). The first is of those populations from the USA, Italy, Chile, Turkey, Georgia, and Hungary, the second is formed by Romania, and the third cluster is formed by Slovenia, Austria, and Serbia. The BMSB populations from China and Japan were clearly separated from the invasive populations (Figure 3).

**Discussion**

To the best of our knowledge, this is the most comprehensive population genomic study so far to unravel the genetic diversity and population structure of BMSB. The study utilised ddRAD sequencing to enhance the knowledge of global BMSB diversity and invasion history. We identified a suitable restriction enzyme pair for genomic digestion of BMSB genome for ddRAD sequencing study, which will be useful in future application. The ddRAD data were analysed using a combination of approaches, including principal component analysis (PCA), phylogenetic analysis and population structure analysis to elucidate the population structure and genetic diversity among the BMSB populations. The present study unambiguously proved that the BMSB populations in the two native regions of China and Japan were genetically distinct. Many BMSB populations from the invaded countries were genetically closer to those of China. Conversely, the Japanese BMSB populations were isolated and showed genetically less similarity to those from the invaded countries. Overall, this study has provided a remarkable resolution in unravelling the population structure and estimating the genetic relatedness among the BMSB populations from a set of native and invaded regions of the world.

**Genetics isolation between the BMSB populations from China and Japan**

The PCA and MSN analyses showed that the Japanese BMSB populations were isolated from the other populations studied here. The fastSTRUCTURE analysis showed that population structures of the BMSB samples from China and Japan were derived predominantly from different genetic components (Figure 3). These results suggested that there are genetic isolations for the BMSB populations between the two countries. Xu *et al.* [22] also observed genetic divergence between the BMSB populations from China and Japan via haplotype analysis of COII and 12S/CR gene regions. One possible explanation for such phenomenon could be the geographical distance between the isolated populations. The width of Tsushima Strait is from 41.6 km to 222 km which is the distance between mainland and Japan [34]. However, the average non-stop flight distance of BMSB is less than 1 km [35]. Therefore, active spread of BMSB individuals from China to Japan is unlikely.

**Estimation of the BMSB invasion pathway**

BMSB populations from the invaded countries are more closely related to China than that of Japan. For example, the USA populations with $F_{ST}$ of 0.084 to the Chinese populations and 0.16 to the Japanese populations, indicating that the USA populations originated more likely from China, but Japan. The results are consistent with the findings of a previous study [36]. Moreover, there were large genetic differences ($0.15 < F_{ST} < 0.25$) between the BMSB populations in Japan and the invaded countries of Chile, Georgia, Italy, Romania, Turkey, and the USA (Table 2). Therefore, we can conclude that the BMSB populations in the six invaded countries might not have been originated from Japanese populations.

Moreover, PCA plot (Figure 1, Additional file 3) indicated that the established populations in Chile, Hungary, Georgia, Italy, Turkey and the USA belonged to one genetic cluster. The fastSTRUCTURE analysis supported this point that those populations had similar genetic components (Figure 3). The neutrality test (Fu's $F_s$) in a previous study (Yan *et al.* unpub.) for these populations was negative ($p < 0.05$), indicating that these established populations have been under population expansion stage [37].

One individual from the Italy cohort was genetically closely related to the Chinese population (Figure 2), suggesting the possible recent arrival of that specimen from China. Besides, the rest of the BMSB samples from Italy showed their close relatedness to the USA and other European populations, suggesting that multiple BMSB invasions have occurred in Italy in the past [16, 38]. Such phenomenon was also observed in Chile.
Novel genetic analysis for the BMSB populations from the invade countries

Nevertheless, this study also provided novel genetic information on the BMSB populations for some of the European countries, such as Austria, Serbia, Slovenia, and Georgia. Since BMSB was detected in Austria [34] and Serbia [35] in 2015 and Slovenia [36] in 2017, Turkey [37] in 2019 and Georgia [38] in 2016, to the best of our knowledge, no genetic studies have been reported on those populations, and this study shows the first effort to characterise the BMSB populations for these countries. The results showed that the BMSB populations from Georgia and Turkey were more likely the next generation of the BMSB from the USA whereas those populations from Austria, Serbia and Slovenia were the hybrids of Chinese and USA populations (Figure 3). The results obtained here will be useful for future monitoring of the pest and assessing the potential pathways of invasion in those countries.

Finally, in compassion with the mitochondrial based (i.e. COI and COII) haplotypes analyses, here we show that ddRADseq can provide high resolution in identifying BMSB genetic populations. As an alternative, analysis of genome-wide SNPs via whole genome sequencing can be applied for deeper understanding the BMSB genomes and genetic diversity, however, it is still more expensive due to the large size of the BMSB genome (~1 Gbp) (Hhal_2.0). Therefore, cost effective approach such as the RAD-based approach with restriction enzymes (ddRAD) used in this study can show a higher resolution in revealing genomic diversity and population structure among BMSB populations. Taken together, we recommend applying genome-wide SNP markers-based study using ddRAD to explore the genomic diversity in insects and other eukaryotic organisms in future applications. Most importantly, our study will help to trace the geographic origin of the BMSB samples, irrespective of their life stages, and sex, that are intercepted at the New Zealand border. The advancement and findings resulted from our study will not only help in claiming the BMSB freedom status to New Zealand but also will greatly assist in the decision making during an incursion and response situation.

Conclusions

Most of the BMSB specimens from Chile, except one individual, were genetically close to each other (Figure 2), implying that there were at least two separated invasions of BMSB in Chile.
This study demonstrates that the restriction enzymes pair of EcoRI and MspI is suitable for generating restriction-associated DNA fragments for identifying SNPs within the BMSB genome through ddRAD sequencing. Via population analyses using ddRADseq data we detected genetic differentiations among different BMSB populations studied in both the native and invaded countries. We believe that this study could also assist in studying the invasion history of BMSB populations and tracing the potential geographical origin of unknown BMSB intercepted at the border or in post-border scenarios in biosecurity settings using the population structure model. Thus, ddRAD sequencing going to be an invaluable arsenal added in to the biosecurity toolbox for tracing the geographic origin of insects at the border/post-border in the near future.

**Methods**

**BMSB specimen collection and DNA extraction**

BMSB samples were collected from 41 regions in 12 countries, namely Austria, Chile, China, Georgia, Hungary, Italy, Japan, Romania, Serbia, Slovenia, Turkey, and the USA (Table 3, Additional file: 1). The field-collected samples were preserved in plastic vials containing 95% ethanol and imported into New Zealand in accordance to the Import Health Standard, Section 22 of the Biosecurity Act 1993. All the specimens were identified by the collectors and confirmed by the Entomologists at PHEL. No specific permits were required for all the sample collected from all the countries covered in this study. Ethics approval was not required as insects are not classified as animals for the purposes of the Animal Welfare Act, 1999, New Zealand Legislation.

Total genomic DNA was extracted from each individual using QIAGEN DNeasy® Blood & Tissue Kit with QIAGEN RNase A treatment (Qiagen, Valencia, CA, USA). Briefly, head and thorax or abdomen was ground using sterile plastic pestle in 1.5 mL tube with ATL buffer and Proteinase K added, and then incubated at 56 °C overnight. The DNA extraction followed the manufacturer's instructions and the DNA was eluted in 150 µl AE buffer. DNA quality was assessed using NanoDrop™ spectrophotometer (ThermoFisher Scientific, USA) and quantified using QuantiFluor™ system (ThermoFisher Scientific, USA). It was followed by an assessment for DNA shearing on a 1% agarose gel against 1Kb Plus DNA ladder (Invitrogen™, CA, USA) in TAE buffer stained with SYBR safe (Life Technologies, CA, USA) and visualised using a Gel Doc Software system (BioRad, Hercules, CA, USA).

**Selection of restriction enzymes in silico for RAD sequencing**

The selection of restriction enzyme (RE) is the crucial step for experimental ddRAD-seq analysis and discovery of the SNP across the genome for genetic diversity study, thus initial tests on the RE used for the library was conducted. To select the best RE, a simulation based on the BMSB reference genome (submitted as Hhal_2.0 by The i5k Initiative on 15/12/2017) and 15 combinations of 9 enzymes (AvaII, EcoRI, MSeI, MspI, NlaIII, PstI, SbfI, SphI, MluCI). A Python script was developed to calculate the number of segments after using different combination of restriction enzymes to cut reference genome.

**Selection of restriction enzymes in vitro: RAD sequencing and bioinformatics analysis**

Nine pairs of enzymes which had higher numbers of segments in in silico analysis or had been used in other published studies [31] were selected. To further confirm the suitability of the RE pairs, two DNA samples were selected to be used for ddRADseq as parallels to eliminate the effects from individual difference. Eighteen RADSeq libraries (9 pairs of enzymes for 2 samples) were prepared using Illumina® TruSeq DNA Nano following the manufacturer’s instructions and ToBo lab eZRAD Protocol v3.2 [49]. Before sequencing, all RADSeq libraries were assessed using a Fragment Analyzer for quality control. The sequencing was conducted by Annoroad Gene Technology Co., Ltd (Beijing, China) on a HiSeqX ten sequencing platform (paired end, 2x150 bp). The 3’end adaptors of raw reads were trimmed using AdaptorRemoval v2 [50] and reads with phred quality score of less than 20 and length of less than 50 bp were discarded using an in-house script. Then, quality-trimmed sequence datasets, were mapped to the BMSB reference genome, Hhal_2.0, using BWA version 0.7.17 [51] default setting and files with the mapping information (i.e. in SAM format) were converted to BAM format, sorted and indexed using SAMtools version 1.9 [52] before identifying the SNPs. SNPs were called using GATK v4.1 [53]. A hard filtering was conducted to discard the false positive SNPs as recommended by GATK developers (QD (QualByDepth) < 2.0, FS (FisherStrand) > 200.0, SOR (StrandOddsRatio) > 10.0, MQRankSum (MappingQualityRankSumTest) < -12.5, ReadPosRankSum (ReadPosRankSumTest) < -8.0). Population level quality control was done using Plink 1.9 (parameters used: call rate 0.8) [54] for subsequent analysis. All bioinformatics analysis was performed on the high-performance computing platform of the New Zealand eScience Infrastructure (NeSI), Auckland, New Zealand.

**Library preparation and ddRAD sequencing**

After selecting the most suitable pairs of restriction enzymes (EcoRI and MspI) for digestion of BMSB genome via RADseq approach, 399 BMSB individuals were further subjected to double-digest restriction-associated DNA (ddRAD) sequencing. ddRAD libraries were prepared using 500 ng of genomic DNA from each sample following the protocol described by Peterson et al [31] with some modifications. Briefly, genomic DNA was digested at 37 °C for 5 h using 10 Units of the two selected restriction enzymes, namely EcoRI and MspI (NEB, MA, USA), and deactivated at 65
°C for 20 mins, then followed by the ligation of Illumina adapter sequences and unique 8 bp barcodes, which varied by at least three bases. Sets of 24 differentially barcoded individuals were pooled and run on a 1% agarose gel, where 220~450 bp fragments were manually excised and purified using a Zymoclean™ Gel DNA recovery kit (Zymo Research, CA, USA). Each pool was amplified by PCR reactions, which were carried out at a total volume of 25 μl, each containing 5 μl 5×Reaction buffer, 5 μl 5×High GC enhancer, 0.25 μl Q5 polymerase, 5nM of library DNA and a unique indexing primer for each pool that corresponds to the standard Illumina multiplexed sequencing protocol. Temperature cycling for PCR comprised one initial denaturation at 98 °C for 30 s, followed by 14 cycles of denaturation at 98 °C for 15 s, annealing at 65 °C for 30 s, and extension at 72 °C for 30 s, followed by a final extension at 72 °C for 5 mins. The PCRs were carried out in a Veriti 96-well thermal cycler (Life Technologies). DNA libraries were quantified using the Agilent high-sensitivity DNA kit in a 2100 Bioanalyzer (Agilent Technologies, CA, USA). Library pools were combined in equimolar concentration to form a single genomic library and sequenced in one lane of a HiSeqX ten Illumina sequencer (paired-end, 2×150 bp) by Personalbio co. (Shanghai, China)

**Sequencing data pre-processing and Quality control of ddRAD data**

To pre-process the raw sequencing data generated via ddRAD, a similar pipeline and data quality filtering approach was applied as described in the section of processing of RADSeq data. In order to call SNPs from ddRADseq data, the raw fastq files were mapped to reference genome, Hhal_2.0, to produce a mapped read file (sam file) using Burrows-Wheeler Aligner v 0.7 [51]. The result was converted into binary format (BAM files) and sorted to increase calculation speed for further analysis using Samtools v1.9 [54]. SNPs were called from BAM files using GATK4 [53]. The quality control parameter used here was the same as that used for the selection of restriction enzymes above. In addition, the obtained SNPs were further filtered using SNP call rate of 1 and MAF (minor allele frequency) of 0.03 using Plink 1.9 [54]. As a result, 10 samples were discarded due to lower SNP call rate.

**Population Genomic Analysis**

Using the SNPs identified from the 389 BMSB individuals, we conducted principal component analysis (PCA), neighbour-joining tree (NJ tree) and population structure analysis using fastSTRUCTURE to elucidate the population structure and genetic diversity among the BMSB populations. PCA was conducted based on with 1775 SNPs obtained from 389 individuals using Plink 1.9 [54] (parameters used: –allow-extra-chr –pca) and plotted using Tableau 2019 [55]. To further explore the genetic relatedness among BMSB individuals, a Minimum Spanning Network based on the SNP results of each individual was also reconstructed using R package poppr [56].

To test for the presence of population structure, pairwise F\textsubscript{ST} values were generated from the ddRAD data by combining the individual data from the same country as one geographic group. This analysis was implemented in Arlequin 3.5 [57] by converting the SNP profile (in Plink format) to Arlequin format using PGDSpider 2.1.1.5 [58]. The obtained population average pairwise F\textsubscript{ST} values were further used to construct a Neighbour-Joining tree in MEGA X 10.1 [32] to reveal the population level’s genetic distance.

To provide additional insight into the genetic variation and population differences, population genetic structure analysis was conducted. The population genetic structure was inferred using the SNPs profile of individual samples. In this process, fastSTRUCTURE 1.0 [59] was used to conduct model-based clustering of all individuals. The best K value (i.e. the number of populations or clusters that the samples are best divided into) was determined as 5 using a python script, chooseK.py within the fastSTRUCTURE software.

**Abbreviations**

**BMSB**: brown marmorated stink bug

**mtDNA**: mitochondrial DNA

**COI**: Cytochrome c oxidase I

**COII**: Cytochrome c oxidase II

**RADseq**: Restriction-site associated DNA sequencing

**ddRAD**: double-digest Restriction-Associated DNA

**SNPs**: Single Nucleotide Polymorphisms

**PCA**: principal component analysis

**MSN**: Minimum spanning networks

**NJ tree**: Neighbour-Joining tree
**Declarations**

**Ethics approval and consent to participate**

No specific permits were required for all the sample collected for this study. No samples were collected in national parks and no endangered or threatened insects were included in this study, thus collection permits were not required for all the collections. All specimens imported into New Zealand were in accordance to the Import Health Standard, Section 22 of the Biosecurity Act 1993. Ethics approval was not required as insects are not classified as animals for the purposes of the Animal Welfare Act, 1999, New Zealand Legislation.

**Consent for publication**

Not applicable

**Availability of data and materials**

The SNP data can be found using the link below

https://1drv.ms/u/s!AvRFcQuxR5slgpcYffCZ3C3pSAa1Xg?e=L7PmSe

**Competing interests**

The authors declared that they have no competing interests.

**Funding**

The research project (number 405731) was funded by the Operational Research programme from the Ministry for Primary Industries (MPI), New Zealand. The funding body provided the financial supports for conducting the research, in which the results were produced. The costs for publishing this paper is also covered by the research fund. The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in the preparation of the manuscript.

**Author’s contributions**

DL, CP, DA and SG conceived the study. DL, JY, CP, AB, and RKB designed the study. GV, PF, JZ and RG collected and provided the BMSB specimens. QF and DG performed the morphological identification. JY conducted all the laboratory experiments and sequence analysis. CP provided the bioinformatics support. JY drafted the manuscript with input from DL and CP. All authors have read, edited and approved the final version of the manuscript.

**Acknowledgments**

We would like to thank Dr David Waite, and Dr Luciano Rigano from Plant Health and Environment Laboratory (PHEL), MPI for the critically reviewing the manuscript. We would like to thank Zhidong Yu, Claire McDonald and Sue Escott-Brown from the Operational Research Team, MPI for their support, suggestions and providing feedback during the project planning phases. Our special thanks go to the Entomology team members of the PHEL of MPI for supplying the BMSB specimens intercepted at the New Zealand border or from the post-border surveillance and assisting in morphological identification of the specimens. Finally, we would like to thank the following researchers for providing the BMSB specimens from a range of different countries:

Dr Kyo Itoyama from Meiji University, Japan; Valerio Mazzoni from Edmund Fond Mach Foundation, Iris Bernardinelli from Servizio fitosanitario e chimico, ricerca, sperimentazione e assistenza tecnica, Italy; Dr Gonzalo Avila from Plant & Food Research, New Zealand; Dr Laura Nixon and Kim Hoelmer from US Department of Agriculture, USA; Dr Catherine Duthie from Readiness Programmes, MPI, New Zealand; Xinyao Gu from Institute of Entomology, Guizhou University; Jianyun Wang from Environment and Plant Protection Institute, Chinese Academy of Tropical Agriculture sciences; Dr Shu-Jun Wei from Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences; Dr Jun Ma from Institute of Plant Quarantine, Guangzhou Customs Technology Center; Dr Yongliang Fan from Northwest A & F University, China; Dr Roxana Ciceoi from University of Agricultural Sciences and Veterinary Medicine of Bucharest, Romania; Dr Ilania Astorga Leiva from Subdepto Vigilancia y control de plagas agrícolas, División Protección Agrícola y Forestal, Servicio Agrícola y Ganadero, Gobierno de Chile; Dr Attila Torma from University of Szeged; Dr Ákos Varga from Government Office of Pest County, Hungary; Dr Temel Göktürk from Artvin Coruh University, Turkey; Dr Maka Murvanidze from Agricultural University of Georgia, Georgia; Dr Wolfgang Rabitsch from University of Vienna, Austria.
References

1. MacLellan R: Plants and environment brown marmorated stink bug: a potential risk to New Zealand. Surveillance (Wellington) 2013, 40(1).

2. Hoebeke ER, Carter ME: Halyomorpha halys (Stål) (Heteroptera: Pentatomidae): a polyphagous plant pest from Asia newly detected in North America. Proc Entomol Soc Wash 2003, 105(1):225-237.

3. Nielsen AL, Hamilton GC: Life history of the invasive species Halyomorpha halys (Hemiptera: Pentatomidae) in northeastern United States. Ann Entomol Soc Am 2009, 102(4):608-616.

4. Leskey TC, Nielsen AL: Impact of the invasive brown marmorated stink bug in North America and Europe: history, biology, ecology, and management. Annu Rev Entomol 2018, 63:599-618.

5. Josifov M, IM K: heteroptera aus korea. ii. Aradidae, berytidae, lygaeidae, pyrrhocoridae, rhopalidae, alydidae, coreidae, urostylidae, acanthosomatidae, sautelleridae, pentatomidae, cydnidae, platuspidae. 1978.

6. Rider DA, Zheng LY, Kerzchner IM: Checklist and nomenclatural notes on the Chinese Pentatomidae (Heteroptera). II. Pentatominae. Zoosystematica Rossica 2002, 11(1):135-153.

7. Lee D, Short BD, Joseph SV, Bergh JC, Leskey TC: Review of the biology, ecology, and management of Halyomorpha halys (Hemiptera: Pentatomidae) in China, Japan, and the Republic of Korea. Environ Entomol 2013, 42(4):627-641.

8. EPPO: Halyomorpha halys (HALYHA). EPPO Global database, Paris, France: EPPO 2020. https://gd.eppo.int/taxon/HALYHA/distribution (accessed 3 August 2020)

9. Zhu G, Bu W, Gao Y, Liu G: Potential geographic distribution of brown marmorated stink bug invasion (Halyomorpha halys). PLoS One 2012, 7(2):e31246.

10. Macavei L, Baetan R, Oltean I, Florian T, Varga M, Costi E, Maistrello L: First detection of Halyomorpha halys Stål, a new invasive species with a high potential of damage on agricultural crops in Romania. Lucrări Științifice, Universitatea de Științe Agricole și Medicină Veterinară “Ion Ionescu de la Brad” Iași, Seria Agronomie 2015, 58:105-108.

11. Heckmann R: First evidence of Halyomorpha halys (Stål, 1855) (Heteroptera: Pentatomidae) in Germany. Heteropteron 2012, 36:17-18.

12. Vetek G, Papp V, Haltrich A, Redeì D: First record of the brown marmorated stink bug, Halyomorpha halys (Hemiptera: Heteroptera: Pentatomidae), in Hungary, with description of the genitalia of both sexes. Zootaxa 2014, 3780(1):194-200.

13. Maistrello L, Dioli P, Vaccari G, Nannini R, Bortolotti P, Caruso S, Costi E, Montemini A, Casoli L, Bariselli M: First records in Italy of the Asian brown marmorated stink bug Halyomorpha halys, a new threat for fruit crops. Atti, Giornate Fitopatologiche, Chianciano Terme (Siena), 18-21 marzo 2014, Volume primo 2014, 283-288.

14. Wermelinger B, Wyniger D, Forster B: First records of an invasive bug in Europe: Halyomorpha halys Stål (Heteroptera: Pentatomidae), a new pest on woody ornamentals and fruit trees? Mitteilungen-Schweizerische Entomologische Gesellschaft 2008, 81(1/2):1.

15. Milonas PG, Partsinevelos GK: First report of brown marmorated stink bug Halyomorpha halys Stål (Hemiptera: Pentatomidae) in Greece. EPPO Bulletin 2014, 44(2):183-186.

16. Cesari M, Maistrello L, Ganzneri F, Dioli P, Rebecchi L, Guidetti R: A pest alien invasion in progress: potential pathways of origin of the brown marmorated stink bug Halyomorpha halys populations in Italy. Journal of Pest Science 2015, 88(1):1-7.

17. Faúndez EI, Rider DA: The brown marmorated stink bug Halyomorpha halys (Stål, 1855) (Heteroptera: Pentatomidae) in Chile. Arquivos Entomológicos 2017, (17):305-307.

18. Kriticos DJ, Kean JM, Phillips CB, Senay SD, Acosta H, Haye T: The potential global distribution of the brown marmorated stink bug, Halyomorpha halys, a critical threat to plant biosecurity. Journal of Pest Science 2017, 90(4):1033-1043.

19. Labware LIMS. Ministry for Primary Industries. 2020. (accessed June 2020).

20. Vandervoet TF, Bellamy DE, Anderson D, MacLellan R: Trapping for early detection of the brown marmorated stink bug, Halyomorpha halys, in New Zealand. New Zealand Plant Protection 2019, 72:36-43.

21. Gariepy TD, Haye T, Fraser H, Zhang J: Occurrence, genetic diversity, and potential pathways of entry of Halyomorpha halys in newly invaded areas of Canada and Switzerland. Journal of pest science 2014, 87(1):17-28.

22. Xu J, Fonseca DM, Hamilton GC, Hoelmer KA, Nielsen AL: Tracing the origin of US brown marmorated stink bugs, Halyomorpha halys. Biol Invasions 2014, 16(1):153-166.

23. Zhu G, Ye Z, Du J, Zhang D, Zhen Y, Zheng C, Zhao L, Li M, Bu W: Range wide molecular data and niche modeling revealed the Pleistocene history of a global invader (Halyomorpha halys). Scientific reports 2016, 6(1):1-10.
24. Morrison WR, Milonas P, Kapantaidaki DE, Cesari M, Di Bella E, Guidetti R, Haye T, Maistrello L, Moraglio ST, Piemontese L: Attraction of Halyomorpha halys (Hemiptera: Pentatomidae) haplotypes in North America and Europe to baited traps. Scientific reports 2017, 7(1):1-11.

25. Hebert PD, Cywinska A, Ball SL, Dewaard JR: Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B: Biological Sciences 2003, 270(1512):313-321.

26. Kang J, Ma X, He S: Population genetics analysis of the Nuijiang catfish Creuchtheloglanis macropterus through a genome-wide single nucleotide polymorphisms resource generated by RAD-seq. Scientific reports 2017, 7(1):2813.

27. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML: Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 2011, 12(7):499-510.

28. Lemopoulos A, Prokkola JM, Uusiäälä Heikilä Ä, S. Vaseemäätgä A, Huusko A, Hyvääärinen P, Koljonen M, Koskinemi J, Vainikka A: Comparing RADseq and microsatellites for estimating genetic diversity and relatedness†”Implications for brown trout conservation. Ecology and evolution 2019, 9(4):2106-2120.

29. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA: Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res 2007, 17(2):240-248.

30. Lexer C, Wüst RO, Mangili S, Heuertz M, Stöltgen KN, Pearman PB, Forest F, Salamin N, Zimmermann NE, Bossolini E: Genomics of the divergence continuum in an African plant biodiversity hotspot, I: drivers of population divergence in Restio capensis (Restionaceae). Mol Ecol 2014, 23(17):4373-4386.

31. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE: Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one 2012, 7(5).

32. Kumar S, Stecher G, Li M, Knyaz C, Tamura K: MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 2018, 35(6):1547-1549.

33. Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987, 4(4):406-425.

34. Imamura K: Prehistoric Japan: new perspectives on insular East Asia: University of Hawaii Press; 1996.

35. Wiman NG, Walton VM, Shearer PW, Rondon SI, Lee JC: Factors affecting flight capacity of brown marmorated stink bug, Halyomorpha halys (Hemiptera: Pentatomidae). Journal of pest science 2015, 88(1):37-47.

36. Valentin RE, Nielsen AL, Wiman NG, Lee D, Fonseca DM: Global invasion network of the brown marmorated stink bug, Halyomorpha halys. Scientific reports 2017, 7(1):9866.

37. Fu Y: Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 1997, 147(2):915-925.

38. Cesari M, Maistrello L, Piemontese L, Bonini R, Dioli P, Lee W, Park C, Partsinevelos GK, Rebecchi L, Guidetti R: Genetic diversity of the brown marmorated stink bug Halyomorpha halys in the invaded territories of Europe and its patterns of diffusion in Italy. Biol Invasions 2018, 20(4):1073-1092.

39. Leskey TC, Hamilton GC, Nielsen AL, Polk DF, Rodriguez-Saona C, Bergh JC, Herbert DA, Kuhar TP, Pfeiffer D, Dively GP: Pest status of the brown marmorated stink bug, Halyomorpha halys in the USA. Outlooks on Pest Management 2012, 23(5):218-226.

40. Haye T, Wyniger D, Gariepy T (2014) Recent range expansion of brown marmorated stink bug in Europe. In: Müller G, Pospischil R, Robinson WH (eds) (2014) Proceedings of the 8th International Conference on Urban Pests. Zurich, Zurich, Switzerland, pp 309–314

41. Cianferoni F, Graziani F, Dioli P, Ceccolini F: Review of the occurrence of Halyomorpha halys (Hemiptera: Heteroptera: Pentatomidae) in Italy, with an update of its European and World distribution. Biologia 2018, 73(6):599-607.

42. Claerebout S, Haye T, Olafsson E, Pannier E, Baltot J: Première occurrence de Halyomorpha halys (Stål, 1855)(Hemiptera: Heteroptera: Pentatomidae) pour la Belgique et actualisation de sa distribution en Europe. Bull Soc R Belg Entomol 2018, 154:205-227.

43. Ciceoi R, Bolocan IG, Dobrin I: The spread of brown marmorated stink bug, Halyomorpha halys, in Romania. Journal of Horticulture, Forestry and Biotechnology 2017, 21(3):15-20.

44. Rabitsch W, Friebe GJ: From the west and from the east? First records of Halyomorpha halys (Stål, 1855)(Hemiptera: Heteroptera: Pentatomidae) in Vorarlberg and Vienna, Austria. Beiträge zur Entomofaunistik 2015, 16:115-139.

45. Šeát J: Halyomorpha halys (Stål, 1855)(Heteroptera: Pentatomidae) a new invasive species in Serbia. Acta Entomol Serbica 2015, 20:167-171.

46. Gapon DA: First records of the brown marmorated stink bug Halyomorpha halys (Stål, 1855)(Hemiptera, Pentatomidae) in Russia, Abkhazia, and Georgia. Entomol Rev 2016, 96(8):1086-1088.

47. Rot M, Devetak M, Carlevaris B, Žežlina J, Žežlina I: First record of brown marmorated stink bug (Halyomorpha halys Stål, 1855)(Hemiptera: Pentatomidae) in Slovenia. Acta Entomol Sloven 2018, 26(1):5-12.
48. Gönçan A, Gümüş E: Brown marmorated stink bug, *Halyomorpha halys* (Stål, 1855)(Hemiptera: Heteroptera, Pentatomidae), a new and important pest in Turkey. Entomol News 2019, 128(2):204-210.

49. Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR, Bird CE: ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ 2013, 1:e203.

50. Schubert M, Lindgreen S, Orlando L: AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC research notes 2016, 9(1):1-7.

51. Li H, Durbin R: Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009, 25(14):1754-1760.

52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The sequence alignment/map format and SAMtools. Bioinformatics 2009, 25(16):2078-2079.

53. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010, 20(9):1297-1303.

54. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ: Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015, 4(1):7.

55. Deardorff A: Tableau (version. 9.1). Journal of the Medical Library Association 2016, 104(2):182-183.

56. Kamvar ZN, Tabima JF, Grünwald NJ: Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2014, 2:e281.

57. Excoffier L, Lischer HE: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources 2010, 10(3):564-567.

58. Lischer HE, Excoffier L: PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 2011, 28(2):298-299.

59. Raj A, Stephens M, Pritchard JK: fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 2014, 197(2):573-589.

Tables

Table 1 Summary of the *in silico* and *in vitro* tests of RE pairs for ddRADseq.

| RE pairs     | *In silico*: numbers of Segment * | *In vitro*: numbers of SNPs ** |
|--------------|----------------------------------|--------------------------------|
| MspI-NlaIII  | 556,040                          | 969                            |
| EcoRI-NlaIII | 508,739                          | 311                            |
| PstI-NlaIII  | 481,590                          | Null                           |
| MseI-PstI    | 303,697                          | 17                             |
| EcoRI-MseI   | 296,739                          | 20,136                         |
| EcoRI-MluCI  | 267,087                          | Null                           |
| MspI-MseI    | 265,748                          | 27,871                         |
| PstI-MluCI   | 259,942                          | Null                           |
| MspI-MluCI   | 222,963                          | 16,369                         |
| MseI-NlaIII  | 182,338                          | 3,909                          |
| NlaIII-MluCI | 148,911                          | 2,135                          |
| EcoRI-MspI   | 118,867                          | 28,328                         |
| MseI-MluCI   | 86,829                           | Null                           |
| MspI-Pstl    | 86,012                           | Null                           |
| EcoRI-Pstl   | 31,417                           | Null                           |

Note: * Prediction of the DNA segments of 300-500 bp against the BMSB genome scaffold

** Two replicates were used for each RE pairs. The number showed the shared SNPs between two replicates. Null indicates not tested.
Table 2. The group pairwise F<sub>ST</sub> (Fixation index) between the BMSB populations from 12 countries

|         | Japan | China | Hungary | Serbia | the USA | Chile | Slovenia | Italy | Romania | Turkey | Austria |
|---------|-------|-------|---------|--------|---------|-------|----------|-------|---------|--------|---------|
| China   | 0.08473 |      |         |        |         |       |          |       |         |        |         |
| Hungary | 0.14372 | 0.0639 |         |        |         |       |          |       |         |        |         |
| Serbia  | 0.1338  | 0.04756* | 0.00055 |        |         |       |          |       |         |        |         |
| the USA | 0.16267 | 0.08428 | 0.01295 | 0.01226 |         |       |          |       |         |        |         |
| Chile   | 0.17654 | 0.09844 | 0.03937 | 0.0405 | 0.0393 |       |          |       |         |        |         |
| Slovenia| 0.10275 | 0.03792 | 0.02487 | 0.01052 | 0.04199 | 0.06715 |         |       |         |        |         |
| Italy   | 0.15174 | 0.07859 | 0.02627 | 0.02892 | 0.03944 | 0.04412 | 0.03764 |       |         |        |         |
| Romania | 0.20842 | 0.13334 | 0.0609 | 0.06616 | 0.05928 | 0.09133 | 0.09115 | 0.07734 |         |        |         |
| Turkey  | 0.17471 | 0.09696 | 0.03322 | 0.02977 | 0.01641 | 0.05674 | 0.06398 | 0.05738 | 0.0937 |         |         |
| Austria | 0.13983 | 0.06013 | 0.04398 | 0.03359 | 0.06301 | 0.0944 | 0.02033 | 0.0657 | 0.11817 | 0.09175 |         |
| Georgia | 0.16825 | 0.0901 | 0.02413 | 0.02107 | 0.01465 | 0.05951 | 0.05157 | 0.04188 | 0.07368 | 0.01651 | 0.07778 |

Note: Asterisk (*) indicates not statistically significant difference (p> 0.05). FST value ranges from 0 to 1, where 0 means no genetic difference (i.e. similar) and 1 means high difference (isolated populations). Values close to zero indicate the populations are sharing their genetic structure and has minimal difference between them.

Table 3 Sample metadata of BMSB collection. N represents the number of individuals.
| Country    | Locality                                   | N   |
|------------|--------------------------------------------|-----|
| Austria    | Vienna City                                | 16  |
| Chile      | Santiago City                              | 19  |
| China      | Anhui Province                             | 6   |
|            | Beijing City                               | 26  |
|            | Guizhou Province                           | 1   |
|            | Hainan Province                            | 7   |
|            | Hebei Province                             | 8   |
|            | Jilin Province                             | 3   |
|            | Shaanxi Province                           | 27  |
|            | Shanxi Province                            | 9   |
| Georgia    | Eki City                                   | 30  |
|            | Samegrelo-Zemo Svaneti Region              | 3   |
| Hungary    | Budapest City                              | 39  |
|            | Debrecen City                             | 10  |
|            | Pécs City                                  | 6   |
|            | Szeged City                                | 3   |
|            | Szombathely City                          | 12  |
| Italy      | Codroipo (UD) Comune                       | 4   |
|            | Mantua City                                | 6   |
|            | Pozzuolo Del Friuli (UD) Comune            | 11  |
|            | Trentino-Alto Region                       | 1   |
| Japan      | Akita Prefecture                           | 3   |
|            | Chiba Prefecture                           | 2   |
|            | Gifu City                                  | 2   |
|            | Ibaraki Prefecture                         | 2   |
|            | Ishikawa Prefecture                        | 2   |
|            | Iwate Prefecture                           | 5   |
|            | Kagoshima City                            | 8   |
|            | Kanagawa Prefecture                        | 3   |
|            | Kyoto City                                 | 10  |
|            | Mie Prefecture                             | 5   |
|            | Miyagi Prefecture                          | 3   |
|            | Saga Prefecture                            | 3   |
|            | Shizuoka City                             | 7   |
|            | Yamanashi Prefecture                       | 12  |
| Romania    | Bucharest City                             | 22  |
| Serbia     | Senta Town                                 | 10  |
| Slovenia   | Ljubljana City                             | 16  |
| Turkey     | Arhavi Town                                | 12  |
Figures

**Figure 1**

Principal component analysis (PCA) plot using 1775 SNPs of 389 individuals. Each point represents an individual. The colour represents the country where the individuals were collected from. X axis represents the variance explained by PC2 (10.3%), and Y axis represents the variance explained by PC1 (28.7%). The picture was created using Tableau 2019 (https://www.tableau.com/)[55]
Figure 2

Minimum spanning networks (MSN) of BMSB individuals. The analysis was based on 775 SNPs derived from 389 individuals of 12 geographical groups comprising Austria, Chile, China, Georgia, Hungary, Italy, Japan, Romania, Serbia, Slovenia, Turkey, and the USA. Each node represents an individual specimen and the edge indicates the genetic distance (dissimilarity: fast pairwise distances) between the individuals. The colour in each circle represents the countries where the samples were collected from. The picture was created using R package poppr [56].
Figure 3

The fastSTRUCTURE bar plot showing the genetic clusters. The assignment probabilities (K=5; the number of populations or clusters that the samples are best divided into) of each genotyped individual of BMSB from 12 different geographical regions were shown. Each pixel column is one individual sample. Different colours in a bar represent proportion of the genomic component in a sample derived from different populations. Dotted line was used to divide BMSB populations from different countries.
Figure 4

The Neighbour-Joining (NJ) tree of 12 geographical groups. The phylogenetic tree was constructed using MEGA X [32] based on genetic distances of population pairwise FST values. The tree shows the evolutionary history of each BMSB population [33]. The optimal tree with the sum of branch length = 0.30043098 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The branch lengths are showed in each node. Scale bar shows the evolutionary time between two nodes and represents the substitutional per sequence site.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- additionalfile2.xlsx
- additionalfile1.xlsx