Biol Invasions
https://doi.org/10.1007/s10530-018-1709-0

ORIGINAL PAPER

Population genetic analyses of complex global insect invasions in managed landscapes: a *Leptocybe invasa* (Hymenoptera) case study

G. Dittrich-Schröder · T. B. Hoareau · B. P. Hurley · M. J. Wingfield · S. Lawson · H. F. Nahrung · B. Slippers

Received: 5 June 2017 / Accepted: 13 March 2018

Abstract Increased rates of movement and the accumulation of insects establishing outside their native range is leading to the 'global homogenization' of agricultural and forestry pests. We use an invasive wasp, *Leptocybe invasa* (Hymenoptera: Eulophidae), as a case study to highlight the rapid and complex nature of these global invasions and how they can complicate management options. To trace the invasion history of *L. invasa* globally, we characterised the genetic diversity within and between populations from its origin and invaded regions using mitochondrial and nuclear markers. Three mitochondrial Haplogroups were identified, of which two are likely different species that appear to have been independently introduced into different parts of the world. One type (Mitochondrial Haplogroup 1) occurs globally, and is the exclusive type found in Europe, the Middle East, South America and most of Africa. The second type (Mitochondrial Haplogroup 2) co-occurs with the first-type in Laos, South Africa, Thailand and Vietnam, while a third type (Mitochondrial Haplogroup 3) occurs exclusively in Australia, its native range. The distinction of the two invasive Haplogroups was supported by analysis of newly developed simple sequence repeat (microsatellite) markers in populations from 13 countries. Further analyses using clustering methods and approximate Bayesian computation suggested the occurrence of hybridisation in the Laos population and revealed that an unsampled population was the origin of Mitochondrial Haplogroup 1. The analyses also showed little genetic differentiation within the invasive populations, suggesting a limited original introduction from a very small population followed by rapid, global range expansion in a stepwise fashion. Results of this study should provide some guidelines for characterizing invasion pathways of new invasive insect pests.

Keywords Invasive insect pest · Gall wasp · Forest entomology · Complex invasion pathways

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10530-018-1709-0) contains supplementary material, which is available to authorized users.

G. Dittrich-Schröder (✉) · B. P. Hurley · M. J. Wingfield
Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Cnr of Lynnwood and University Roads, Hatfield, Private Bag X20, Pretoria 0028, South Africa
e-mail: Gudrun.Dittrich@fabi.up.ac.za

T. B. Hoareau · M. J. Wingfield
Department of Genetics, University of Pretoria, Cnr of Lynnwood and University Roads, Hatfield, Private Bag X20, Pretoria 0028, South Africa

S. Lawson · H. F. Nahrung
Forest Industries Research Centre, The University of the Sunshine Coast, Queensland, Australia

B. Slippers
Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Cnr of Lynnwood and University Roads, Hatfield, Private Bag X20, Pretoria 0028, South Africa

Published online: 24 March 2018

Introduction

The worldwide prevalence of invasive pests is primarily attributed to increasing global travel and trade, and especially the increasing importation of live plants and plant material (Meyerson and Mooney 2007; Liebhold et al. 2012). Of the invasive organisms recorded, insects constitute the largest portion (Liebhold et al. 2012). This abundance of invasive pests and frequent records of new invasive pests is especially threatening to monocultures, including plantation forestry where uniform genetic material is deployed (Hoffmann et al. 2008; Wingfield et al. 2015). In addition, the global homogenization of pests in plantation forestry is accelerated due to small number of tree species planted in ever-growing areas, creating global "neighbourhoods" (Garnas et al. 2013).

Garnas et al. (2016) analysed global invasion patterns of insect pests and found a general pattern of multiple introductions into some areas, and complex patterns of secondary movement in the adventive range of pests. Accordingly, the authors predicted an increase in genetic diversity in adventive populations over time, leading to admixture of individuals from different populations. This results in rapidly evolving, unique populations, which are consequently difficult to manage (Garnas et al. 2016). At the same time, Hurley et al. (2016) showed that the rate of invasion of insect pests in *Eucalyptus* plantation forestry has increased five-fold since the mid-1980's. From these predictions, we should expect that the composition of pests in forestry areas globally will become increasingly similar and that these pests would show an increased genetic diversity due to multiple introductions and admixture of different types. In this study we use *Leptocybe invasa* in a case study to test and illustrate these predictions.

Leptocybe invasa (Hymenoptera: Eulophidae) is amongst the most threatening insect pests to *Eucalyptus* forestry plantations worldwide. This minute gall-forming wasp is native to Australia and was first found outside its range of origin in Israel in 2000 (Mendel et al. 2004). It oviposits in the stems, midribs and petioles of young leaves, leading to gall formation, stunting of plant growth, and, in severe cases, tree death (Nyeko 2005). Damage is especially evident in small seedlings, coppice and young leaves.

Host specificity of *L. invasa* appears to be restricted to *Eucalyptus* and *Corymbia* species belonging to the sections *Adnataria*, *Exsertaria*, *Latoangulata*, *Maidenaria* and *Transversaria*, *Renantheria*, and *Blakella*, *Ochraria* and *Rufaria*, respectively (Mendel et al. 2004; Thu et al. 2009). Previous studies showed that *Eucalyptus* species suitable for oviposition by *L. invasa* include *E. botryoides*, *E. bridgesiana*, *E. camaldulensis*, *E. camaldulensis simulata*, *E. camaldulensis* var *obtusa*, *E. coolabah*, *E. globulus*, *E. globulus maidenii*, *E. microcorys*, *E. gunii*, *E. grandis*, *E. molucana*, *E. pellita*, *E. pilularis*, *E. robusta*, *E. saligna*, *E. smithii*, *E. tereticornis*, *E. tereticornis tereticornis*, *E. urophylla*, *E. viminalis* and *Corymbia polycarpa*, as well as various genotypes and hybrids of *Eucalyptus* (Mendel et al. 2004; Thu et al. 2009; Dittrich-Schröder et al. 2012b).

Since its first record *L. invasa* spread rapidly to most areas of the world where *Eucalyptus* spp. are planted, including Africa, Asia, the Mediterranean region of Europe, India, China and South America (Dittrich-Schröder et al. 2012b). The wasp completes its development within a gall and is, therefore, generally well protected from biotic and abiotic threats. Control of the wasp relies on the use of specific natural enemies in biological control programs and the planting of tolerant or resistant *Eucalyptus* spp. and clones (Mendel et al. 2004; Dittrich-Schröder et al. 2012b, 2014).

The rate of invasion and spread of *L. invasa* worldwide has in part been attributed to its reproductive mechanism. *Leptocybe invasa* is thought to reproduce thelytokously, where offspring are all female and genetically identical to their mother (Mendel et al. 2004). Some *Leptocybe* individuals harbour different strains of endosymbiotic *Rickettsia*, which are bacteria known to induce thelytokous parthenogenesis in arthropods (Merçot and Poinsot 2009; Nugnes et al. 2015). Further, invasive species may alter their reproductive mode from sexual reproduction to parthenogenetic reproduction in the invaded area (Caron et al. 2013). Parthenogenetic reproduction in an invaded area has advantages, such as the absence of the Allee effect, which potentially increases the probability of establishment (Tobin et al. 2011; Kronauer et al. 2012). Recent reports have, however, confirmed the presence of males in invaded populations (Akhtar et al. 2012; Sangtongpraow et al. 2011). It is not known how prevalent these males are or what role, if any, they have in populations of the pest. The presence of males in a population suggests

the possibility of sexual reproduction that may lead to increased genetic diversity in a population (Schrey et al. 2010). Genetic diversity could facilitate adaptation to changing environmental conditions, as well as overcoming resistance in planting material (de Meeûs et al. 2007; Roderick and Navajas 2003). Understanding this aspect of the biology of *L. invasa* as well as other invasive organisms is critical for management programs.

Molecular data are increasingly used to understand aspects of the origin, biology, host range, cryptic species identity and genetic diversity of invasive species, providing essential information for the effective management of these pests (Valade et al. 2009; Caron et al. 2013). For example, one needs to determine whether an invasion is due to one or a number of closely related species. Knowledge of the origin or source and biology of an invasive species is important so that routes of invasion can be identified to prevent further introductions (Valade et al. 2009). Furthermore, knowing the genetic diversity of the invasive species can help to identify the correct biotype of the biocontrol agent or understand likely durability of resistant plant material (Chown et al. 2015). In this context and considering its high priority as a forestry pest, it is surprising that no comprehensive molecular study has been conducted on global populations of *L. invasa*.

In this study, COI sequence data and microsatellite markers were used to test three hypotheses, namely (1) populations of *L. invasa* comprise cryptic species, as suggested by Nugnes et al. (2015), (2) routes of invasion are linked to historical records of first reports, and (iii) admixture is apparent in the invasive range. To address these hypotheses, we determined the genetic diversity and structure within and between invasive populations of *L. invasa*. These data were then used to determine the route and history of the invasion of the pest globally. The overall aim was to provide a base of knowledge important not only for efforts to prevent further invasions but to inform biological control and resistance breeding programs for the management of the invasive *L. invasa* populations. Furthermore, we hoped to provide information relevant to the management of other invasive pests and some guidelines on how to unravel complex invasion pathways of insect pests.

Materials and methods

Sample collection and DNA extraction

A total of 511 female *Leptocybe invasa* specimens from 18 countries and five continents were included in this study (Table 1). Specimens from Australia were assumed to represent populations from the region of origin, if not the actual source populations, for the invasion. Comprehensive sampling was conducted in Australia from Tumoulin in northern Queensland to Kenmore in southern Queensland. Samples from eight countries (Ghana, Italy, Israel, Malaysia, Thailand, Uganda, Vietnam Zimbabwe) were from a single locality per country and these were collected between 2008 and 2013. From five countries (Brazil, Kenya, Laos, Mozambique and South Africa) sampling was conducted in the major forestry areas from between two to five localities per country, from 2008 to 2015.

Five countries (Argentina, China, Italy, Tunisia and Turkey) were represented from molecular data available in GenBank.

Specimens used for molecular analyses either emerged or were dissected from field-collected galls from various susceptible *Eucalyptus* species and hybrids. Due to one gall containing many individuals, galls were collected from different branches of the same tree, as well as neighbouring trees.

DNA was extracted using whole specimens and either the Nucleospin Tissue XS (Macherey–Nagel) kit or the ZyGEM DNA extraction using *prep*GEM^{TM} Insect (ZyGEM) kit following the methods described by Dittrich-Schröder et al. (2012a).

Mitochondrial diversity

Mitochondrial DNA sequencing

Polymerase chain reaction (PCR) was conducted using the universal barcoding region primers of the cytochrome oxidase I region of the mitochondrial DNA with modification of the forward primer (LCO1490 (C1-J-1514) (5′-GGTCAACAAATCATAAAGATATTGG-3′) (Folmer et al. 1994)) for improved amplification LiLCO1490 (5′-ATTTGATCTGGAATTTTAGG-3′) and HCO2198 (C1-N-2173) (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer et al. 1994) to yield a 605 bp fragment of the COI region of the mitochondrial DNA. A 25 µl reaction

Table 1 Collection localities of the *Leptocybe* spp. specimens used in this study

Continent	Country	Locality	Sampling year	Number of samples included in mtDNA COI analysis	Number of samples included in SSR analysis (per country)	GenBank accession numbers for mtDNA COI sequences
Africa	Ghana		2013	8	8	MH093117–MH093119, MH093168–MH093172
	Kenya	Maragua	2008–2013	24	44	MH093190–MH093193, MH093226–MH093228, MH093410–MH093426
		Kiamie	2013	6		MH092440–MH093445
		Nyeri	2013	6		MH093446–MH093451
		Gede	2013	13		MH093427–MH093439
	Mozambique	Ifloma	2010	13	34	MH093197–MH093199, MH093234–MH093238, MH093305–MH093309
		Manica, Bandula	2010	23		MH093453–MH094375
	South Africa	Bela-Bela	2009	15	31	MH093310–MH093324
		Pretoria	2010–2015	19		MH093186, MH093241–MH093243, MH093325–MH093339
		Piet Retief	2015	4		MH093111, MH093112, MH093180, MH093181
	Tunisia[a]			3		
	Uganda	Bugambe Tea Estate, Hoima	2009	23	20	MH093239, MH093240, MH093252–MH093259, MH093394–MH093402
	Zimbabwe	Nyabira	2010	1	1	MH093476
Asia	China[a]			15		
	Israel	Bet-Dagan	2007–2009	33	50	MH093187–MH093189, MH093207–MH093210, MH093218–MH093223, MH093260–MH093279
	Laos	Savannakhet	2013	16	36	MH093200–MH093204, MH093280–MH093289, MH093452
		Dan Hi	2013	7		MH093146, MH093147, MH093149–MH093152, MH093355
		Namdeua	2013	7		MH093148, MH093349–MH093354
		Thungnai	2013	4		MH093345–MH093348
		Nonmivai	2013	4		MH093340–MH093343
	Malaysia		2013	9	8	MH093113–MH093116, MH093175–MH093179
	Vietnam			10	9	MH093139–MH093145, MH093173, MH093174, MH093344

Table 1 continued

Continent	Country	Locality	Sampling year	Number of samples included in mtDNA COI analysis	Number of samples included in SSR analysis (per country)	GenBank accession numbers for mtDNA COI sequences
	Thailand	Panomtuan district, Kanchanaburi Province	2008	65	63	MH093120–MH093138, MH093153–MH093167, MH093211, MH093212, MH093217, MH093245, MH093247–MH093251, MH093372–MH093393
Australasia	Australia	Mareeba, North Queensland	2005	1	111	MH093063
		Hervey Bay	2013	5		MH093048–MH093052
		Nanango	2013	12		MH093053, MH093064–MH093074
		Walligan	2013	6		MH093054, MH093055, MH093182–MH093185
		Takura	2013	5		MH093056, MH093075–MH093078
		Tumoulin	2013	5		MH093082–MH093086
		Maleny	2013	2		MH093087, MH093088
		Marcella Creek	2013	6		MH093057–MH093062
		Kenmore	2012	33		MH093010–MH093042
		Maroochydore	2013	5		MH093001–MH093005
		Sunshine University	2015	4		MH093006–MH093009
		Toolara	2015	5		MH093043–MH093047
		Noosa	2014	2		MH093094, MH093095
		Miva	2014	6		MH093079, MH093089–MH093093
		Ingham	2010	2		MH093080, MH093081
		Gardens Point	2014	15		MH093096–MH093110
Europe	Italy	Aretina, Reggio, Calabria	2009	30 8[a]	29	MH093205, MH093206, MH093224, MH093225, MH093290–MH093304, MH093361–MH093371
	Turkey[a]			1		
South America	Brazil	Bahia Province	2012	17	21	MH093213–MH093216, MH093229–MH093233, MH093244, MH093403–MH093409
		Batucatu	2013	10		MH093356–MH093360, MH093477–MH093481
	Argentina[a]			3		
			Total	511	**465**	

[a]Sequences from GenBank

volume was used which included 2 µl DNA (≈ 2 ng) of diluted genomic DNA, 12.9 µl distilled water, 2.5 µl 10 × PCR Buffer (Roche, Roche Diagnostics, Mannheim, Germany), 2.5 µl dNTP's (10 µM of each dNTP) (Roche), 3 µl 25 mM $MgCl_2$ (Roche), 0.2 µl FastStart taq polymerase, 1 µl of each primer diluted to 10 µM (30 pmol). The PCR protocol included an initial denaturation of 7 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 56 °C and 1 min at 72 °C, followed by a final extension of 10 min at 72 °C. PCR products were visualised on a 2% agarose gel using a BioRad Gel Doc™ Ez Imager and the software Image Lab v 4.0 build 16. PCR products were purified using the QIAquick® PCR Purification Kit (QIAGEN) following the manufacturer's instructions. Purified PCR products were visualised on a 2% agarose gel as described above. Cycle sequencing was conducted using the ABI Prism Big Dye Terminator v3.1 Cycle Sequencing Kit 5000 reaction using 6 µl distilled water, 0.5 µl Big Dye 3.1, 2 µl sequencing buffer, 0.5 µl primer and 2 µl purified PCR product. Cycle sequencing conditions included initial denaturation of 2 min at 96 °C followed by 30 cycles of 30 s at 96 °C, 15 s at 55 °C and 4 min at 60 °C. Cycle sequencing products were cleaned using the Ethanol/NaAC precipitation of BigDye Terminator v3.1 DNA sequencing reactions protocol from the ABI manual. Samples were sequenced using an ABI Prism™ 3100 Genetic Analyzer (Applied Biosystems).

Mitochondrial DNA sequence analysis

Forward and reverse sequence reads were aligned and checked for base calling accuracy using the Staden package (Staden 1996). Resulting chromatograms were aligned and edited using ClustalX version 2.1 (Larkin et al. 2007) and BioEdit version 7.0.9 (Hall 1999), respectively. DNA sequences were translated into amino acids in order to check for the presence of stop codons using AliView version 1.17.1 (Larsson 2014). Thirty sequences of *L. invasa* were available on GenBank and these were added to the dataset (GenBank numbers JQ289999–290005, KP233954, KP233982–KP233989, KP233990–KP233993, KP233972–KP233981) before analysis. A neighbour-joining tree was constructed using the Kimura 2-parameter substitution model with 1000 bootstrap replicates, and the uncorrected percentage pairwise sequence divergence was calculated using MEGA version 5.0 (Tamura et al. 2011). A COI haplotype network was constructed using Network 4.6.1.1 (Polzin and Daneschmand 2003; www.fluxusengineering.com). Haplotype diversity and nucleotide diversity (Nei 1987) were calculated using DnaSP version 5.10.01 (Rozas et al. 2003).

Microsatellite diversity

Microsatellite development

Due to the small size of *L. invasa* specimens, ten adult specimens were used to obtain sufficient DNA for genome sequencing. DNA extraction of each individual specimen was conducted using the Nucleospin Tissue XS (Macherey–Nagel) kit. The genome of *L. invasa* was sequenced using Illumina HiSeq paired-end reads to obtain coverage of approximately 40x. Quality control on raw data was performed using FastQC. A preliminary assembly was prepared using Velvet Optimiser with an optimal kmer value of 77. Completeness of the genome assembly was assessed using CEGMA (Parra et al. 2007) and by mapping the raw reads back to the assembled genome using Bowtie2 (Langmead and Salzberg 2012).

Msatcommander 1.0.8 (Faircloth 2008) was used to search the assembly for microsatellites. Primer 3 version 0.4.0 (Untergrasser et al. 2012) was used to develop 14 primer pairs to span 14 randomly selected microsatellites. Primers were developed to amplify predominantly tetranucleotides, as well as a few trinucleotides and dinucleotides, and needed to contain at least seven repeats. These primers were tested using specimens from South Africa, Israel and Thailand to ensure that they were polymorphic. Two panels with seven primer pairs each were designed for use in multiplex combinations using Multiplex Manager v1.2 (Holleley and Geerts 2009). The forward primer in each primer pair was fluorescently labelled (Table 4, Online Appendix A).

Microsatellite amplification

DNA samples from 465 specimens included in the mtDNA analysis were used in an analysis of the microsatellite markers. PCR was performed for every specimen with each of the 14 primer pairs in a 10 µl reaction volume using 0.5 µl DNA (≈ 0.5 ng) of diluted genomic DNA, 7.5 µl distilled water, 1 µl 10×

PCR Buffer (Roche), 0.07 µl dNTP's (10 µM of each dNTP) (Roche), 0.6 µl 25 mM $MgCl_2$ (Roche), 0.12 µl FastStart taq polymerase, 0.1 µl of each primer diluted to 10 µM (30 pmol). The PCR protocol followed included initial denaturation of 7 min at 95 °C, 35 cycles of 30 s at 95 °C, 30 s at the appropriate annealing temperature and 1 min at 72 °C, and a final extension of 45 min at 60 °C. PCR products were visualised on a 2% agarose gel as described above.

Specimen genotyping

A 1:100 DNA:water dilution was prepared for all samples taking into account the intensity of the PCR product on the 2% agarose gel. The amount of DNA added per primer pair was either 1 µl of DNA for strong bands or 2 µl of DNA for faint bands with the amount of water adjusted such that the DNA: water dilution remained 1:100. One microliter of the 1:100 DNA:water dilution, containing PCR products of seven primer pairs (Panel 1 = 7 primer pairs; Panel 2 = 7 primer pairs) (Online Appendix A), was added to a formamide and Liz600 (Applied Biosystems) size standard mixture (14 µl Liz600: 1000 µl formamide). The products were run on an ABI Prism™ 3100 Genetic Analyzer (Applied Biosystems) and subsequently allele sizes determined using GeneMapper® software version 4.1.

Microsatellite analyses

Analysis of population structure

STRUCTURE version 2.3.4 (Pritchard et al. 2000), which applies a Bayesian clustering method, was used to indicate the number of populations or clusters (*K*) based on genetic identity of the individuals and without geographic prior. Based on the distinct groupings generated by the mitochondrial DNA results (see results section) as well as F_{st} values, Haplogroup A (Israel, South Africa, Kenya, Uganda, Italy, Mozambique, Zimbabwe, Laos, Vietnam and Brazil) was referred to as Lineage A and Haplogroup B (Laos, Thailand, Ghana, Malaysia and Vietnam) and Haplogroup C (Australia) were referred to as Lineage B. The Haplogroups referred to the groupings supported by mitochondrial data and lineages referred to groupings supported by the nuclear data. Three analyses were performed, including (1) the whole data set (N = 465), (2) only with Lineage A (N = 234), and (3) only with Lineage B (N = 231). An admixture model with correlated allele frequencies was implemented because it was more likely to detect subtle population structure. Tests included 20 repetitions for a range of *K*-values from one to 10. For the Markov chain Monte Carlo (MCMC) procedure, the 'burn in' period was set to 100 000 iterations followed by 100 000 samples. The optimal *K*- value, which indicates the most likely clustering of individuals, was verified by determining the variation of likelihood values and ΔK (Evanno et al. 2005) using STRUCTURE Harvester (Earl and von Holdt 2012). Results were visualised using CLUMPAK (Kopelman et al. 2015).

A principle coordinate analysis (PCoA) was conducted in GenAlEx version 6.501, using the average pairwise genetic distance among populations to determine the source of most of the variation (Peakall and Smouse 2006). A distance matrix, using the genetic distance for co-dominant data, was used to generate a covariance matrix by dividing the genetic distance by the square root of n-1.

Calculation of genetic diversity indices

Genetic diversity measures were calculated and included the total number of alleles, mean number of alleles, observed heterozygosity and expected heterozygosity using Arlequin version 3.5.1.3 (Excoffier et al. 2005). Genetic variation was quantified among regions (native vs invasive), among populations, and within populations using analysis of molecular variance (AMOVA) with 10 000 permutations using Arlequin version 3.5.1.3 (Excoffier et al. 2005). Pairwise F_{st} values were calculated for specimens grouping into Lineage A and Lineage B using Arlequin version 3.5.1.3 (Excoffier et al. 2005).

Approximate Bayesian Computation

Approximate Bayesian Computation is a bayesian method which can be used to infer the likelihood of a population or species by comparing the "probability of observed data given the values of the model parameters" (Cornuet et al. 2014). This method compares different models (referred to as scenarios) which could describe the history of a species or lineage

from its ancestral population to its current (observed) population taking into account model parameters such as the influence of population size, bottlenecks and admixture, to name a few, in shaping these populations (Cornuet et al. 2014). These analyses mostly use miscrosatellite data, which are popular markers used in population genetics, although DNA sequence data can also be used (Cornuet et al. 2014). Approximate Bayesian Computation can be summarized into three main steps (1) Generation of simulated data (2) Selection of a single simulated data set which most closely represents the observed data set, and (3) estimation of the posterior distribution of the selected parameters by a local linear regression (Cornuet et al. 2014).

Hypothesis testing of various invasion scenarios was conducted with the nuclear data using the software DIYABC version 2.0.4 (Cornuet et al. 2014). Based on results from the STRUCTURE clustering analysis, F_{st} values (see Results section), and historical records of first reports of the wasp, separate analyses were performed for Lineage A and Lineage B. In all cases, scenarios tested were kept as simple as possible as it is not possible to test all scenarios. To provide results that were sufficiently robust, it was necessary to select the most likely, yet contrasting scenarios, in combination with a stepwise approach that would improve the efficiency of the computer resources. The starting point for the analyses was a very broad global perspective with the idea of searching for general trends before identifying more specific trends. The above mentioned approach, compared to a location-oriented approach is advantageous due to its efficiency and potential to account for patterns of admixture between populations (following a similar approach as Brouat et al. 2014). The analysis was undertaken in four smaller analyses, thereby increasing computational efficiency. For each analysis 1,000,000 simulated data sets were used per scenario. For all four analyses "one sample summary statistics" included mean size variance and mean Garza-Williamson's M and "two sample summary statistics" included mean number of alleles, F_{st} and classification index. Some of the scenarios included an unsampled population to improve accuracy of the results when the origin is unknown, as suggested by Lombaert et al. (2011). For each of the four analyses a principle coordinate analysis (PCoA) was used to ensure that the simulated scenarios were within the 1% closest likely scenarios and a logistic regression approach was used to determine the posterior probability of each scenario.

Test of the origin of the invasive species of Lineage A. The Test#1 analysis intended to test the hypothesis that the origin of the worldwide invasion (Lineage A) was Australia (details of hypotheses available in Online Appendix). Results from STRUCTURE plots showed no substructure within Lineage A from the invasive range and we therefore made the assumption that all populations within Lineage A could be grouped together to detect their origin. This assumption is especially true for invasive populations that can have divergences that can be seen very rapidly, because of the combination of the strong genetic bottleneck and associated genetic drift related to multiple colonizations. Three scenarios were compared. The first scenario assumed that the Australian population diverged from an unsampled population, which gave rise to two populations, the Israel population and the Global population comprising the remaining populations of Lineage A (namely Brazil, Israel, Italy, Kenya, Mozambique, South Africa, Uganda, Zimbabwe). The second scenario assumed that the Israel and Global populations emerged directly from the Australian population. The third scenario assumed that the Australian population diverged from an unsampled population, which experienced a bottleneck before giving rise to the Israel population and before the colonization of other areas (Global population). This scenario was the same as the first scenario, but included the bottleneck option. Default prior parameters were used with additional priors of tb $>$ ta and tc $>$ tb, where 't' refers to the time which is calculated in generations and increases towards the past (i.e. 'a' represents the most recent generation whereas 'c' represents the oldest generation).

Test of the movement of invasive species of Lineage A. Subsequent to determining the origin of Lineage A, Test#2 analysis was conducted to test the hypothesis that the movement of Lineage A, from the area where it was first detected (Israel), was in a stepwise manner in the assumed newly colonized locations. Three scenarios were compared (Online Appendix B). Historical dates of first report (Fig. 1) were used to determine whether they could be supported by the molecular data or whether a colonisation event was the most likely scenario. The first scenario investigated the feasibility of a step-wise colonization of Africa

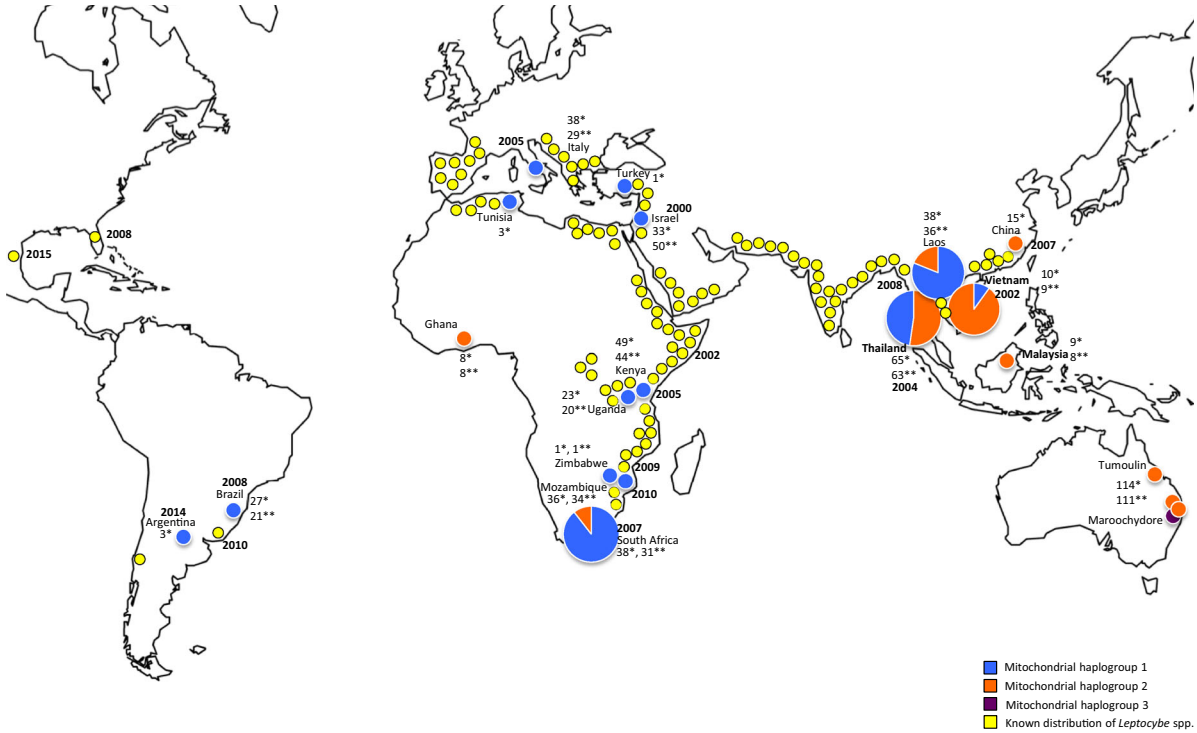


Fig. 1 World wide distribution of *Leptocybe invasa* and *Leptocybe* sp. indicated by yellow circles. Collection localities of the *Leptocybe* spp. samples used in this study are indicated by blue (Mitochondrial haplogroup A), orange (Mitochondrial haplogroup B) or purple (Mitochondrial haplogroup C) circles. Pie charts indicate the proportion of samples at a locality represented by each haplogroup. Dates indicate first reports of *Leptocybe*. A single asterisk indicates the number of samples using mitochondrial DNA (N = 511) whereas two asterisks indicate the number of samples using nuclear DNA (N = 465)

from Israel from north to south using the most direct path of invasion between countries (Kenya, Uganda, Mozambique and South Africa, respectively). The second scenario investigated the feasibility of a stepwise colonization of Africa from Israel using insights gained from the historical dates of invasion (i.e. Israel, Kenya, Uganda, South Africa, Mozambique, respectively). The third scenario assumed the occurrence of a colonisation event from Israel to all countries in Africa. Default prior parameters were used with additional priors of tb >=ta, td >=tc and tc >=tb, where 't' refers to the time.

Test of the origin of the Laos population. Results from STRUCTURE plots showed the presence of admixture in the Laos population and therefore the hypothesis tested various combinations of populations which could have given rise to the Laos population. The Test#3 analysis used six scenarios to elucidate the source of the admixed Laos population (Online Appendix B). The first scenario assumed that the Global population (Lineage A) and the Laos population emerged from an unsampled population and that the Laos population subsequently gave rise to the Australian and South East Asian populations. The second scenario assumed that the Global population and the Australian population emerged from an unsampled population with the Laos population and then the South East Asia population emerged from the Australian population. The third scenario assumed that the Global population and the Australian population emerged from an unsampled population and that admixture between the South East Asian population, which emerged from the Australian population, and the Global population gave rise to the Laos population. The fourth scenario assumed that the Australian and Global population emerged from an unsampled population and that the Laos and South East Asian populations directly emerged from the Global and Australian populations, respectively. The fifth scenario assumed that the Laos and South East Asian population emerged from an unsampled population and subsequently the Global and Australian

population arose from the Laos and South East Asian populations, respectively. The sixth scenario assumed that the Global and Australian populations emerged from an unsampled population with the Australian population, giving rise to the South East Asian population, and the Laos population consequently resulting from admixture between the Australian and South East Asian population. Default prior parameters were used with additional priors of tb >=ta, tc >=tb and tc >=ta, where 't' refers to the time.

Test of the origin of the invasive species of Lineage B. The Test#4 analysis intended to test the hypothesis that the origin of the Lineage B was Australia. Three scenarios were compared (Online Appendix B). Results from STRUCTURE plots showed no substructure within Lineage B from the invasive range and we therefore made the assumption that all populations within Lineage B could be grouped together to detect their origin. This is especially true for invasive populations that can have divergences that can be seen very rapidly, because of the combination of the strong genetic bottleneck and associated genetic drift related to multiple colonizations. The first scenario assumed that populations from Ghana, Vietnam, Malaysia and Thailand all emerged from the Australian population due to a colonisation event. The second scenario examined a similar option, but with all five populations (Ghana, Vietnam, Malaysia, Thailand and Australia) emerging from an unsampled population. The third scenario assumed that an unsampled population gave rise to an Australian population while the Ghana, Vietnam, Malaysia and Thailand populations arose from the unsampled population. Default prior parameters were used with an additional prior of tb > ta, where 't' refers to the time. All competing scenarios for each of the four hypotheses tested are presented in Online Appendix B.

Results

Mitochondrial DNA sequencing and analysis

Translated DNA sequences indicated the absence of stop codons. A total of 511 sequences (Table 1, Fig. 1) of a 605 bp region of the cytochrome oxidase I (COI) region of the mtDNA were generated and used to construct a NJ tree and a haplotype network (Fig. 2A, B). Both analyses gave rise to three distinct groups, which were thus named Mitochondrial Haplogroups 1, 2 and 3, respectively. Mitochondrial Haplogroup 1 included sequences from 311 specimens from 14 countries. The percent sequence divergence within Mitochondrial Haplogroups 1, 2 and 3 was from 0, 0–1.5 and 0–0.5% respectively. The between group mean distance for Mitochondrial Haplogroups 1 and 2 was 3.63%; 1 and 3 was 4.23%, and, 2 and 3 was 4.52%. The haplotype network (Fig. 2B) indicated that there were 24 Australian haplotypes and two exotic haplotypes (designated Haplotype 1 (Mitochondrial Haplogroup A) and Haplotype 2 (Mitochondrial Haplogroup B)). Haplotype 1 represented 60% of the total number of individuals and 78.3% of the individuals in the invasive range. Haplotype 2 represented 17% of the total number of individuals and 21.7% of the individuals in the invasive range. All the remaining haplotypes were present only for specimens from the native range, Australia. Haplotype diversity in the native range was higher than in the invasive range (Table 2).

Microsatellite development

The 40× coverage of the *L. invasa* genome using Illumina HiSeq paired-end reads produced a relatively fragmented assembly that was nevertheless valuable for microsatellite discovery. The assembly indicated a genome size of 350 Mb, containing 237 396 contigs, 474 433 microsatellite regions, had an N50 of 12364 bp and a GC content of 37%. Characterization of microsatellites constituting the *L. invasa* genome showed that dinucleotides (58.8%) and mononucleotides (32.4%) were the most abundant motifs (Table 3). Trinucleotides constituted 4.3%, while tetranucleotides, pentanucleotide and hexanucleotides combined constituted 4.5% of the microsatellites in the *L. invasa* genome. Of the 14 primer pairs developed, only one was discarded due to ambiguities, making scoring difficult, and the presence of null alleles (Table 4).

Microsatellite amplification

Data for the remaining 13 microsatellite markers were obtained for 465 specimens (Table 1). We were able to demonstrate structure in our dataset, indicating robustness of markers, and therefore the number of markers used in this study were considered adequate.

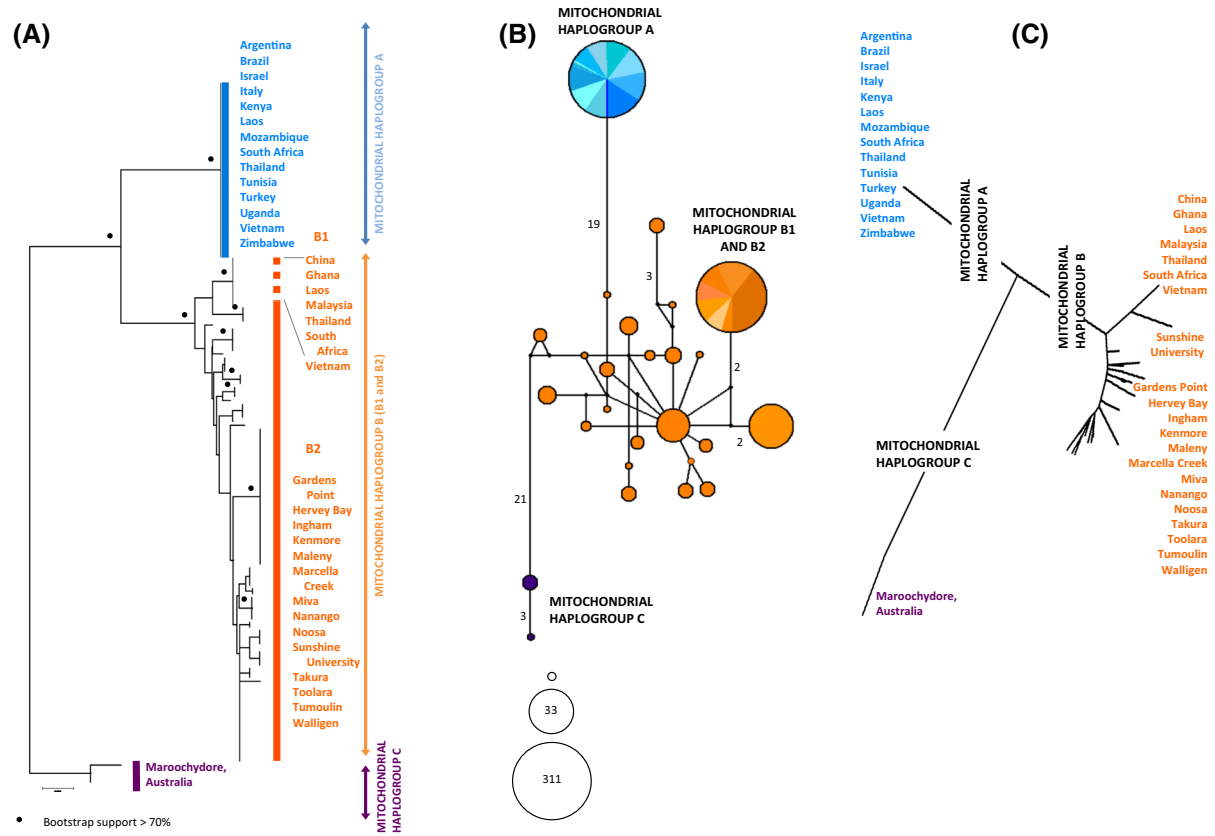


Fig. 2 **a** Neighbour joining tree – Bootstrap values above 70% are indicated by filled circles (**b**) Haplotype network—Circles are proportional to the number of individuals represented. The number of individuals are shown in the empty circles beneath the haplotype network. Each circle represents a single haplotype. Different shades of the same colour indicates the number of different geographical localities which share the same haplotype, with slices proportional to the number of individuals per geographic location (see Table 2 for further details). Each line connecting a circle represents one mutational step, unless otherwise indicated) (**c**) Unrooted tree of native and invasive *Leptocybe* populations using the COI data. Colours relate to the STRUCTURE grouping (Fig. 3) of specimens

For further analyses the single Zimbabwean specimen was grouped with specimens from Mozambique as the sample regions are geographically close.

Microsatellite analyses

Analysis of population structure

The analysis using STRUCTURE divided the collection into two clusters ($K = 2$) with Lineage A being highly uniform and separate from Lineage B (Fig. 3). Further analysis of Lineage B only divided the specimens into two further clusters ($K = 2$), separating the specimens from Laos, Thailand, Ghana, Vietnam, Malaysia and Maroochydore from the remaining Australian specimens (Kenmore, Tumoulin, Walligan, Marcella Creek, Nanango, Hervey Bay, Takura, Gardens Point, Toolara, Noosa, Sunshine University, Miva, Ingham and Maleny) (Fig. 3). Separate analysis of Lineage A indicated no further differentiation ($K = 1$).

The principle coordinate analysis showed a clear separation of Lineages A and B with 52.7% and 9.0% of the percentage of variation explained by coordinate 1 and coordinate 2, respectively (Fig. 4).

Genetic diversity indices

The mean number of alleles per locus and total number of alleles (A) was higher for Australia than any other

Table 2 The number of individuals, number of haplotypes, haplotype diversity and nucleotide diversity for the native and invasive ranges of *L. invasa* and *Leptocybe* sp. using mtDNA COI sequence data

Geographical region	No. of individuals	No. of localities/ countries	No. of haplotypes	Haplotype diversity (Hd)	Nucleotide diversity (Pi)
Native (Australia)					
Haplogroup B	109	14	22	0.8675	0.00588
Haplogroup C	5	1	2	0.4000	0.00198
Invasive					
Haplogroup A	311	14	1	0.0000	0.00000
(Argentina, Brazil, Italy, Israel, Kenya, Laos, Mozambique, South Africa, Thailand, Tunisia, Turkey, Uganda, Vietnam, Zimbabwe)					
Haplogroup B	86	7	1	0.0000	0.00000
(China, Ghana, Laos, Malaysia, South Africa, Thailand, Vietnam)					

Table 3 Characterization of the SSR's constituting the *L. invasa* genome

Microsatellite class	Total motifs (% of all microsatellites)		Abundant motifs	Number of motifs	% of microsatellite class	% of all microsatellites
Mononucleotides	153,541	(32.36%)	A	74,869	48.76	15.78
			T	75,219	48.99	15.85
			G	1727	1.12	0.36
			C	1726	1.12	0.36
Dinucleotides	279,078	(58.82%)	AG/GA	61,808	22.15	13.03
			CT/TC	61,361	21.99	12.93
			AT/TA	40,666	14.57	8.57
			AC/CA	48,188	17.27	10.16
			GT/TG	42,011	15.05	8.85
			CG/GC	25,044	8.97	5.28
Trinucleotides	20,575	(4.34%)	ACG/AGC/CAG/CGA/GAC/ GCA	7914	38.46	1.67
			CGT/CTG/GCT/GTC/TCG/ TGC	7900	38.40	1.67
			AAG/AGA/GAA	2381	11.57	0.50
			CTT/TCT/TTC	2380	11.57	0.50
Tetranucleotides, Pentanucleotides and Hexanucleotides	21,239	(4.48%)		21,239		4.48
Total of all microsatellites	474,433	100%				

location (Table 5). The AMOVA indicated that the among populations within regions differentiation (26.99%; $P < 0.001$) was less than the within populations differentiation (66.67%; $P < 0.001$), but greater than the among groups differentiation (6.33%; $P = 0.055$) (Table 6). Population pairwise F_{ST} values within Lineage A were much smaller than F_{ST} values within Lineage B. Population pairwise F_{ST}

Table 4 SSR primers used, their fluorescent label, repeat motif, annealing temperature and number of alleles

Locus		Primers (5′–3′)	Fluorescent label	Repeat	T_a	Range of allele sizes (bp)	Number of alleles	Alleles
LiSS1	F	TGTGTTGTGTTTGTGAAGGTG	VIC	$(TG)_{12}$	62	153–155	2	153, 155,
	R	CCATAAACAAACGTGCACTGA						
LiSS2	F	CCATATTGGGTCCACCTACC	VIC	$(AC)_{12}$ or $(AC)_{20}$	62	179–255	15	177, 179, 183, 187, 189, 193, 195, 197, 199, 201, 205, 213, 215, 231, 235, 237, 240, 241, 243, 245, 247, 249, 255
	R	ACCGTCCTTGCGTATACAGG						
LiSS3	F	CCGCTTTACAATACCCGAAA	FAM	$(AT)_{12}$	60	295–321	9	295, 297, 299, 301, 303, 305, 307, 313, 319, 321
	R	TCTATTGAAGAGAAATACCGAGCA						
LiSS4	F	GGTGTGCATGAAGACAGCAG	PET	$(AGC)_{10}$	62	218–239	4	215, 218, 224, 227, 233, 236, 239
	R	AAAGCTTCCTCGGTGTCTGA						
LiSS5	F	TCGTGTTTACCACCTGACCA	NED	$(AGC)_9$	62	351–360	3	351, 354, 360
	R	AGAGTGCTCAGGCTCGACAT						
LiSS6*	F	CGATACAAGGGTGTGCATGA	PET	$(AGC)_{11}$	58	120–150	–	
	R	AATATGACGTGCACGAATGAA						
LiSS7	F	TTGACGCTCTTTTCACATGC	NED	$(ACGC)_8$	60	213–225	2	213, 225, 229
	R	TGTCGGTGTGTGGGTGTATC						
LiSS8	F	TTGCCAGCTCGTGACTATTG	FAM	$(AAAT)_{10}$	60	397–401	2	390, 393, 397, 401
	R	GCAAACTGCATGTACGAAAAA						
LiSS9	F	CGACCAAAAGTCCTAATCCTTTC	FAM	$(AAAC)_{10}$	64	102–130	3	75, 102, 106, 114, 130
	R	GGTCGTTCGACACGAGCTTA						
LiSS10	F	ATCGCTGCAGCTCTGTCTCT	NED	$(ACAG)_7$	64	156–162	4	152, 156, 158, 160, 162
	R	AGCGAGGCTAATTGTCAAGG						
LiSS11	F	CTGGCGAGTTGAGTTCCTTC	VIC	$(AGGC)_{11}$	62	342–366	4	338, 342, 346, 350, 354, 358, 366
	R	TCGGGGCTAAGTCATTCAAG						
LiSS12	F	CGTGTGTATGTGCGAGAACC	PET	$(AACG)_8$	63	168–188	3	168, 180, 184, 186, 188
	R	GTCACAGTACCGGCCAAAGT						
LiSS13	F	TGGTACAAATCCCGTCTATGG	FAM	$(ACGC)_7$	65	141–149	2	141, 149
	R	CGCAACGGTACAGAAATTCA						
LiSS14	F	TTTTTCCCTACCGTGCGTAA	VIC	$(AAAT)_9$	61	107–115	3	103, 107, 111, 113, 115
	R	CGAAAGTTTCAATTTGCCAGT						

*The discarded primer pair due to amplification and scoring difficulties

values for Lineage A (Table 7) were significant only for comparisons between South Africa and Italy and South Africa and Uganda. For all other comparisons for Lineage A F_{ST} values were zero. For Lineage B population pairwise F_{ST} values were significant in most comparisons with the exception of Thailand and Vietnam, and Malaysia and Vietnam (Table 8).

Approximate Bayesian computation

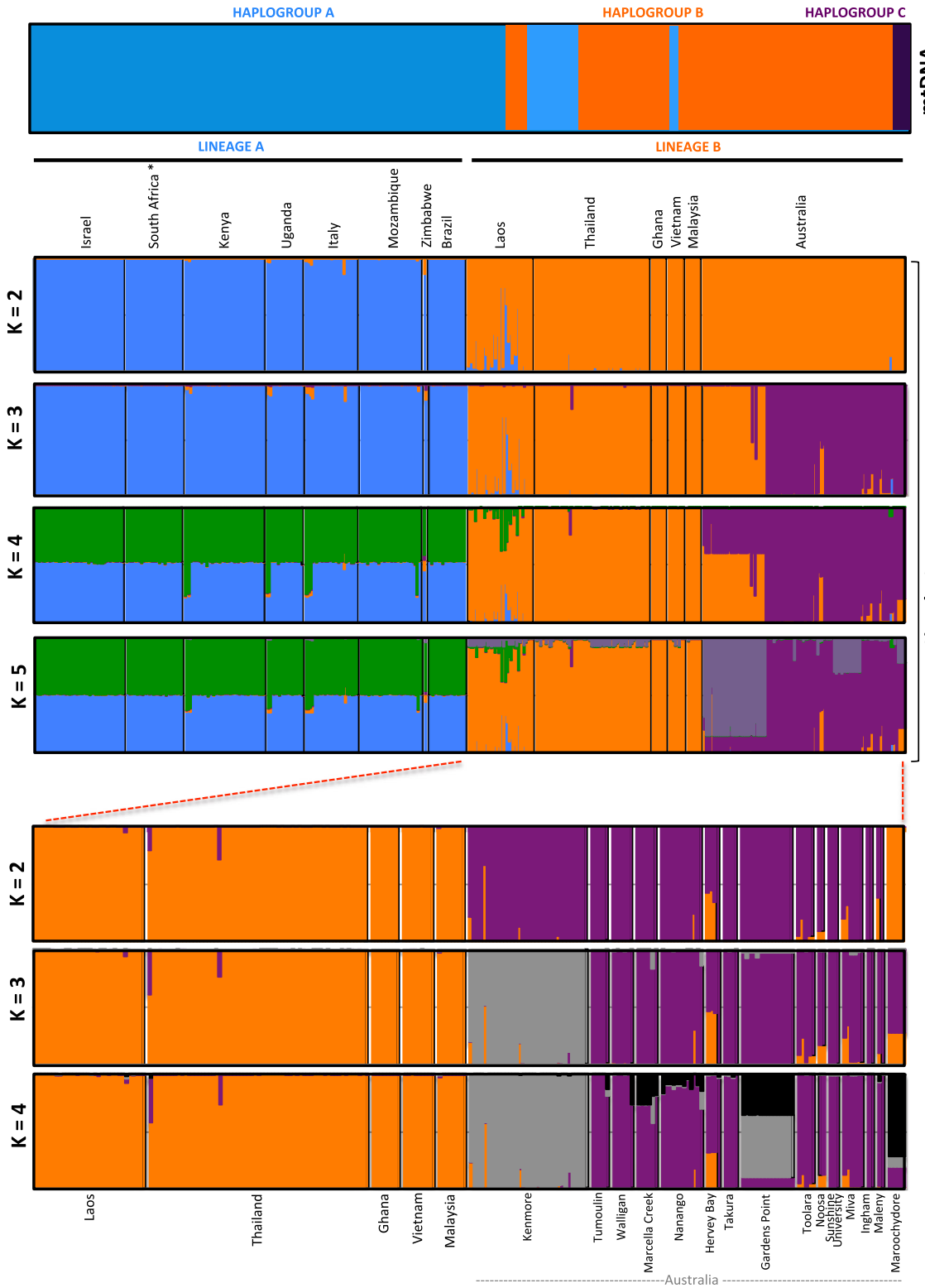
Scenario testing for Test#1 intended to test the hypothesis that the origin of the worldwide invasion (Lineage A) was Australia. Scenario number three was supported with the highest posterior probability. This scenario assumed that the Asutralian population diverged from an unsampled population. After divergence, the unsampled population experienced a bottleneck before giving rise to the Israel population and the global population (Global population (P = 0.7316 [0.7244; 0.7388]) (Fig. 5). Test#2 tested the hypothesis that the movement of Lineage A, from the area where it was first detected (Israel), was in a stepwise manner in the assumed newly colonized locations. High posterior probabilities supported the hypothesis that within Lineage A, a colonisation event occurred from Israel rather than a step-wise divergence from north to south in Africa (P = 0.9998 [0.9998; 0.9999]) (Fig. 5). Test#3 tested the hypothesis that admixture between various populations gave rise to the Laos population. The hypothesis that the Laos population originated as a result of admixture between Lineage A and the South East Asian populations (Vietnam, Malaysia) was highly supported (P = 0.8834 [0.8694; 0.8974]) (Fig. 5). Test#4 tested the hypothesis that the origin of the Lineage B was Australia. Scenario three, which assumed that an unsampled population gave rise to an Australian population while the Ghana, Vietnam, Malaysia and Thailand populations arose from the unsampled population (P = 0.9051, [0.8981; 0.9121]) (Fig. 5) was the best supported scenario. In all cases the chosen scenarios were strongly supported by high posterior probabilities.

Discussion

Cryptic species of *Leptocybe invasa*

An important conclusion from the COI sequence data was that *L. invasa* specimens group into three distinct Haplogroups. Both the neighbour joining tree and the haplotype network yielded the same three groups. The haplotype network separated the main groups by a substantial number of mutations (19 and 21 mutations) and the three groupings of the neighbour joining tree were similarly well supported. These three mitochondrial groupings were designated Haplogroup A, B and C. Nugnes et al. (2015) suggested the possibility of two cryptic *Leptocybe* species, with similar distributions to specimens in this study collected from Haplogroups A and B (Mediterranean and South America, and China, respectively). The percentage sequence divergence (3.63%) in our study between these groups supports the notion that they represent two cryptic species. The first group (Haplogroup A) is a highly genetically homogenous group that included samples from Israel where the wasp was first taxonomically described in 2004 (Mendel et al. 2004), as well as Argentina, Brazil, Italy, Kenya, Laos, Mozambique, South Africa, Thailand, Tunisia, Turkey, Uganda, Vietnam and Zimbabwe. The second group (Haplogroup B) included samples from China, Ghana, Laos, Malaysia, Thailand, South Africa, Vietnam and Australia, and is homogenous outside Australia. The third group, Haplogroup C, occured exclusively in Maroochydore, Australia, and is sufficiently different to Haplogroups A and B to be considered a separate species (sequence divergence between A and C is 4.23%, sequence divergence between B and C is 4.52%). Furthermore, Kim (2008) found nine different *Leptocybe* spp. in Australia based on morphological and molecular analyses, but none of which matched the invasive population.

Genetic variation within and between groups can also be used to indicate species boundaries. The genetic variation within and between the three *Leptocybe* Haplogroups was a maximum of 1.5 and 3.7% respectively, which is similar to intraspecific and interspecific sequence divergence values observed for other Hymenoptera (Fig. 6) (Hastings et al. 2008; Ács et al. 2010; Rehan and Sheffield 2011; Smith et al. 2013; Turcinaviciene et al. 2016). Percent sequence divergence can be a useful indicator of species boundaries, although it should not be used in isolation for this purpose (Cognato 2006). In most instances, the sequence divergence between specimens of the same species should be less than the sequence divergence observed between congeneric species (Hebert et al. 2003). However, not all Haplogroups evolve at the same rate; thus making it difficult to suggest a single value that can be used to predict species boundaries (Zhou et al. 2007). Furthermore, it may not be possible to differentiate species that have recently diverged using only a single gene (Zhou et al. 2007). Therefore, while the COI data strongly suggest that the three



* Mitochondrial data for haplogroup A and B exist for South Africa but SSR data only exists for haplogroup A. To avoid confusion only haplogroup A is indicated in the mitochondrial profile

◀**Fig. 3** The output from a STRUCTURE analysis (lower half of page) indicating the grouping of specimens, by country, belonging to Lineage A and B. For comparative purposes, the same groupings based on mitochondrial data are given in the same format above (Haplogroups A, B and C)

Haplogroups for *Leptocybe* could be distinct species, additional biological studies should be used to test the species boundaries.

While mitochondrial data suggested the presence of three *Leptocybe* Haplogroups, nuclear data provided support for only two groups, namely Haplogroup A and Haplogroups B and C combined into a single group. The combined results from the mitochondrial and nuclear data support two distinct groups of *Leptocybe*, which we hereafter refer to as Lineage A (Mitochondrial Haplogroup A) and Lineage B (Mitochondrial Haplogroup B and C). In four localities (Laos, Thailand, South Africa and Vietnam) mitochondrial data support the presence of both Lineage A and B. Nuclear data for Laos indicates introgression of genetic material between Lineage A and B. In Thailand and Vietnam no trace of introgression was observed in the nuclear data, although both lineages were represented by the mitochondrial data. This pattern suggests the occurrence of directional introgression, where not all alleles are retained after hybridization, as well as backcrossing to the haplotype having the highest level of fitness. This could suggest that the introgression observed in Laos is more recent than in Thailand and Vietnam, where sufficient time has passed for this characteristic signature of introgression to be lost.

The Laos and Israel locations may serve as "bridgehead populations" for further introductions. Bridgehead populations are thought to aid invasive organisms in becoming established by serving as "stepping-stones" for populations to move elsewhere (Lombaert et al. 2011). This phenomenon has been observed in many successful invaders making management of the global spread of these invaders difficult (Garnas et al. 2016). Areas where bridgehead populations co-occur may also result in admixture of different lineages or closely related species (Lombaert

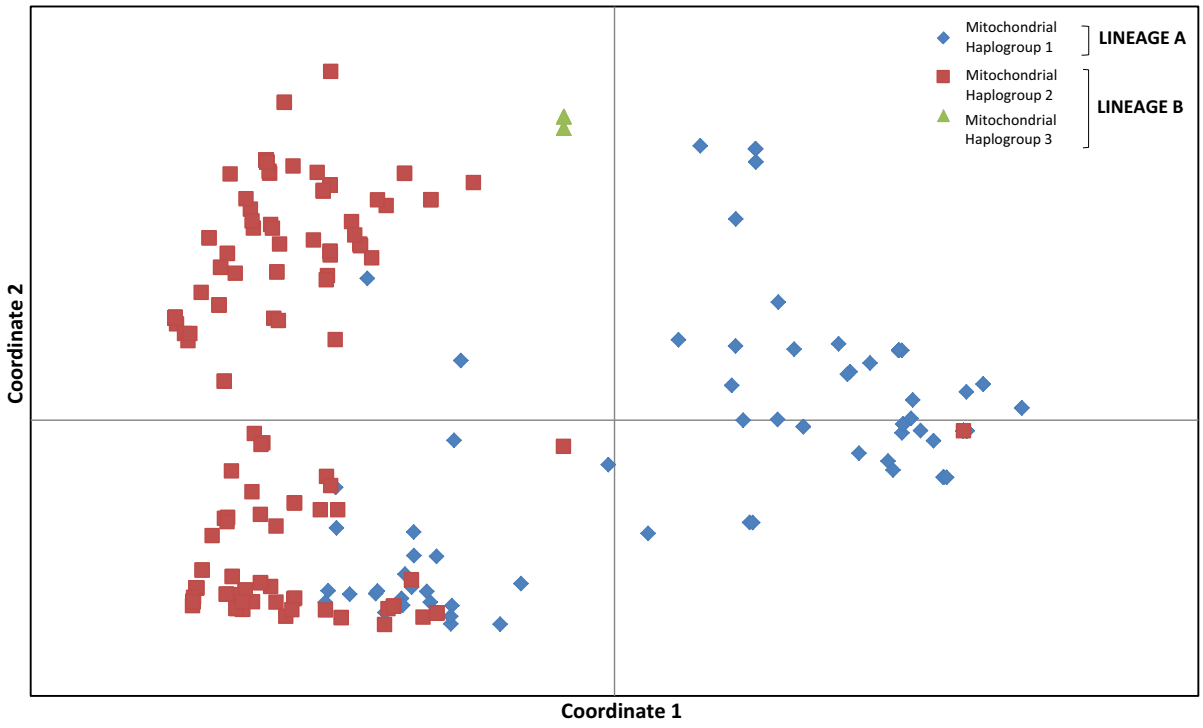


Fig. 4 A principle co-ordinate analysis using the mean population genetic distance among populations indicating the separation of Lineages A and B of *Leptocybe invasa* and *Leptocybe* sp.

Table 5 Measures of genetic diversity using SSR markers of *Leptocybe* spp. from native and introduced populations

Regions	Number of individuals	Total number of alleles (A)	Mean number of alleles (± SD) by locus	Number of polymorphic loci	Number of usable loci	Mean expected heterozygosity (± SD)	Mean observed heterozygosity (± SD) for polymorphic loci	Average gene diversity over loci (± SD)
Native								
Australia	112	77	5.923 ± 5.155	10	10	0.428 ± 0.275	0.180 ± 0.355	0.452 ± 0.247
Invasive								
Brazil	21	18	1.385 ± 0.506	4	8	0.197 ± 0.259	0.990 ± 0.021	0.244 ± 0.157
Ghana	8	15	1.154 ± 0.689	2	12	0.089 ± 0.220	1.000 ± 0.00	0.097 ± 0.074
Israel	45	27	2.077 ± 0.277	13	13	0.221 ± 0.232	0.367 ± 0.482	0.218 ± 0.129
Italy	29	20	1.538 ± 0.519	6	10	0.217 ± 0.243	0.669 ± 0.468	0.264 ± 0.159
Kenya	44	19	1.462 ± 0.519	4	9	0.203 ± 0.250	0.772 ± 0.370	0.179 ± 0.119
Laos	37	27	2.077 ± 1.038	2	5	0.308 ± 0.254	0.369 ± 0.381	0.231 ± 0.167
Malaysia	8	15	1.154 ± 0.689	0	5	0.084 ± 0.266	0.857 ± 0.202	0.000
Mozambique	35	25	1.923 ± 0.494	6	7	0.231 ± 0.229	0.447 ± 0.508	0.237 ± 0.156
South Africa	37	18	2.000 ± 0.000	4	10	0.194 ± 0.255	0.939 ± 0.030	0.197 ± 0.125
Thailand	64	31	2.385 ± 0.961	10	11	0.249 ± 0.236	0.299 ± 0.396	0.216 ± 0.131
Uganda	21	24	1.846 ± 0.376	6	7	0.256 ± 0.219	0.419 ± 0.482	0.251 ± 0.165
Vietnam	9	17	1.308 ± 0.630	4	9	0.131 ± 0.198	0.422 ± 0.487	0.164 ± 0.116
Mean ± SD			**1.970 ± 0.951**	**5.462 ± 3.666**	**8.923 ± 2.465**	**0.216 ± 0.237**	**0.336 ± 0.393**	**0.373 ± 0.206**
Lineage A[a]	231	32	2.462 ± 0.967	69.23		0.213 ± 0.237	0.365 ± 0.479	0.186 ± 0.121
Lineage B[a]	208	81	6.231 ± 5.525	84.62		0.419 ± 0.240	0.253 ± 0.398	0.474 ± 0.228

Values in bold indicate average values for invasive populations

[a]Excludes samples from Laos as nuclear data indicated admixture within this group

et al. 2011) resulting in unique population diversity, which is different to either of the parental lineages (Garnas et al. 2016).

Routes of invasion

Origin of Leptocybe invasa

Analyses of mitochondrial COI sequence data and nuclear microsatellite markers used to characterize the diversity of *L. invasa* showed that populations from Australia, its likely region of origin, were highly diverse. This was in contrast to the very low diversity in its invasive range, with repeated bottleneck effects during the process of invasion. Importantly, we identified three distinct Haplogroups, A, B and C, potentially representing cryptic species, of which two have been separately introduced around the world. We treat Lineage A as *Leptocybe invasa* because it includes specimens from Israel where the wasp was first detected and subsequently described (Mendel et al. 2004). All other groups should be referred to as *Leptocybe* sp.

The origin of Haplogroup A remains unknown as it was not linked back to the purported origin, based on mitochondrial DNA data. This is despite the fact that substantial collections across its Australian

Fig. 5 Most supported scenarios and corresponding logistic regression for analyses in DIYABC. Time is indicated on the axis on the right hand side and is not to scale

Table 6 Analysis of the molecular variance (AMOVA) of the native (Australia) and introduced populations of *Leptocybe* spp. using SSR data

Source of variation	df	Sum of squares	Variance components	Percentage of variation	*P* value
Among regions (native vs invasive)	1	52190.113	47.16396	6.33440	0.055
Among populations within regions	11	129791.415	200.99733	26.99512	0
Within populations	927	452305.089	496.40772	66.67048	0
Total	939	2336.047	3.02617	100	

Table 7 Pairwise genetic differentiation (F_{ST}) for populations of *Leptocybe invasa* (Lineage A)

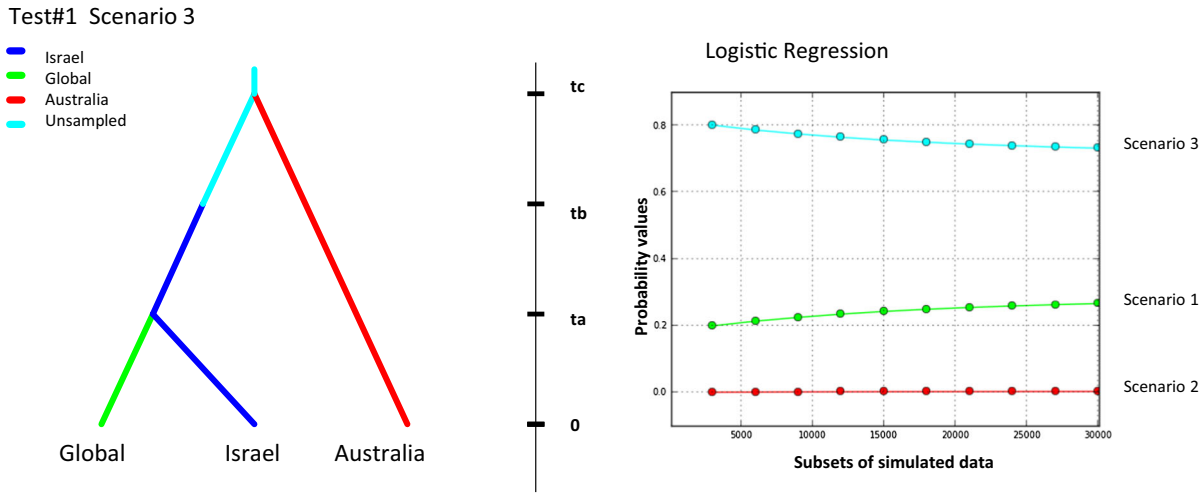
	Israel	South africa	Uganda	Mozambique	Italy	Brazil	Kenya
Israel	–						
South africa	0	–					
Uganda	0	0.00240*	–				
Mozambique	0	0	0	–			
Italy	0	0.00392*	0	0	–		
Brazil	0	0	0	0	0	–	
Kenya	0	0	0	0	0	0	–

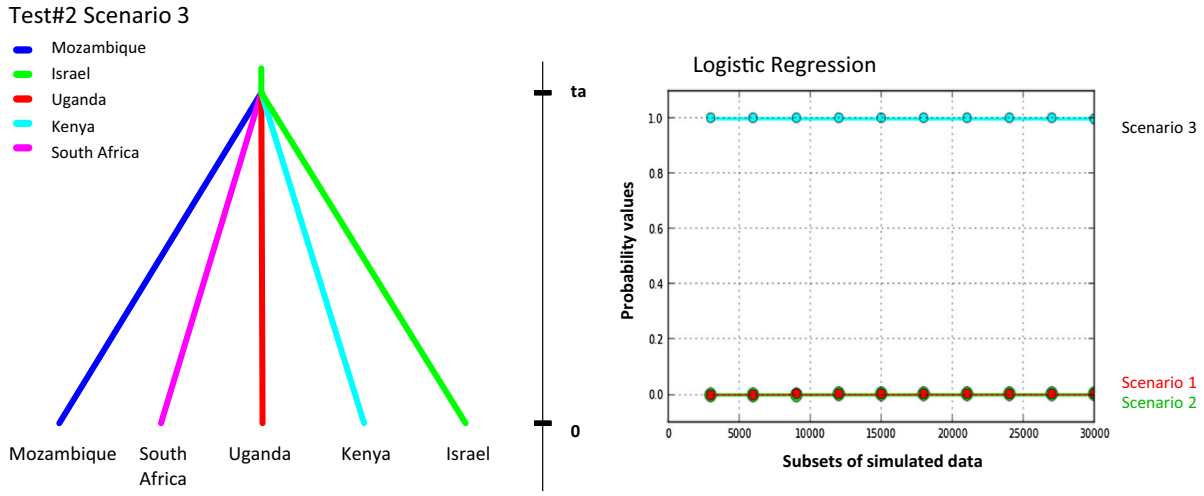
*Denotes statistically significant values at the 0.05 level

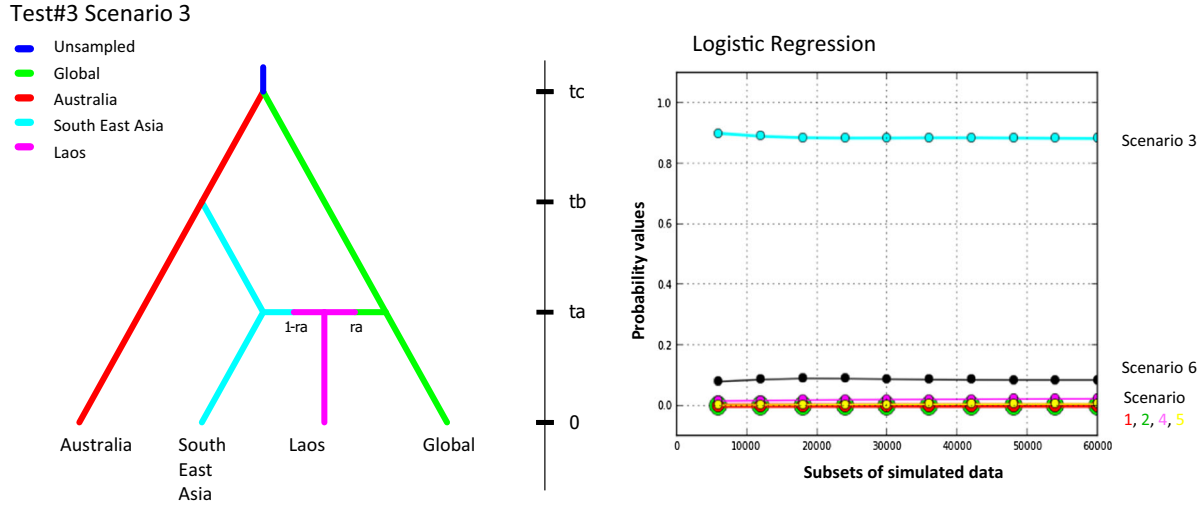
Table 8 Pairwise genetic differentiation (F_{ST}) for populations of *Leptocybe* sp. (Lineage B)

	Malaysia	Laos	Ghana	Thailand	Vietnam	Australia
Malaysia	–					
Laos	0.20170*	–				
Ghana	0.03759*	0.24463*	–			
Thailand	0.05861*	0.07144*	0.13894*	–		
Vietnam	0	0.16879*	0.13627*	0.01724	–	
Australia	0.18157*	0.21832*	0.19823*	0.21384*	0.19806*	–

*Denotes statistically significant values at the 0.05 level


Test#1 Scenario 3
Israel
Global
Australia
Unsampled
Global
Israel
Australia
tc
tb
ta
0
Logistic Regression
Probability values
Subsets of simulated data
Scenario 3
Scenario 1
Scenario 2


Test#2 Scenario 3
Mozambique
Israel
Uganda
Kenya
South Africa
Mozambique
South Africa
Uganda
Kenya
Israel
ta
0
Logistic Regression
Probability values
Subsets of simulated data
Scenario 3
Scenario 1
Scenario 2


Test#3 Scenario 3
Unsampled
Global
Australia
South East Asia
Laos
1-ra
ra
Australia
South East Asia
Laos
Global
tc
tb
ta
0
Logistic Regression
Probability values
Subsets of simulated data
Scenario 3
Scenario 6
Scenario 1, 2, 4, 5

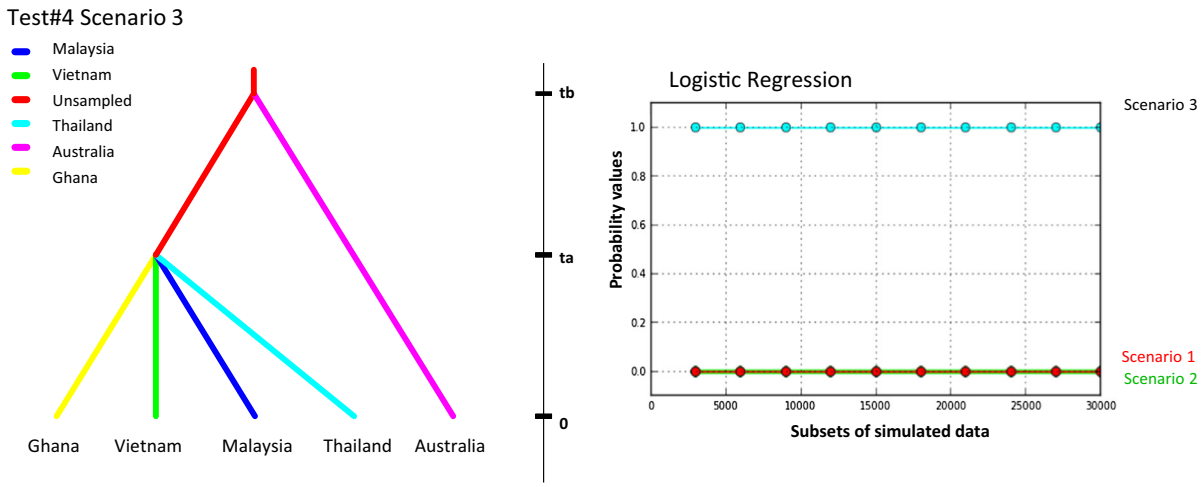


Fig. 5 continued

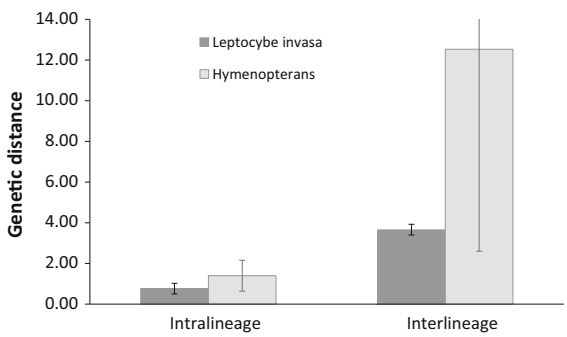


Fig. 6 A bar graph showing the intralineage and interlineage sequence divergence observed from *Leptocybe invasa* haplogroups and from other Hymenopteran species (Ács et al. 2010; Hastings et al. 2008; Rehan and Sheffield 2011; Smith et al. 2013; Turcinaviciene et al. 2016). Error bars indicate the standard deviation

distribution from the origin were included. Scenario testing supported the hypothesis that Lineage A (*L. invasa*) originated from an unsampled population. Although extensive sampling has been conducted from Queensland, Australia for over a decade (2005–2015), the source population for this lineage was not identified. This has significant implications for management of the pest, because there could be advantages to collecting natural enemies from the specific lineages and areas from which the invasive populations have originated (Brodeur 2012; Yek and Slippers 2014).

Leptocybe populations from Australia had a high degree of diversity. Microsatellite data showed a reduction in the mean number of alleles and gene diversity outside Australia (2.923 ± 1.754 and 0.373 ± 0.206, respectively) compared to the native range (6.333 ± 5.069 and 0.453 ± 0.247, respectively). There were also more COI haplotypes in samples from the native range (n = 24) than from the introduced range (n = 2). The higher genetic diversity observed for *L. invasa* in its assumed native range (Australia) compared to its introduced range was as expected, and is consistent with observations for other insect pests such as the Asian longhorn beetle (*Anoplophora glabripennis*; Carter et al. 2010), the Eurasian woodwasp (*Sirex noctilio*; Boissin et al. 2012), the bronze bug (*Thaumastocoris peregrinus;* Nadel et al. 2009), the Argentine ant (*Linepithema humile;* Tsutsui et al. 2001) and the tropical fire ant (*Solenopsis geminata*; Gotzek et al. 2015). The data from the present study, together with the origin of the host plant, *Eucalyptus*, support the assumption that Australia is the native region of *L. invasa*.

Lineage A

Both mtDNA COI sequence data and microsatellite data indicated the presence of a bottleneck in Lineage A, which was also confirmed by Approximate Bayesian Computation analysis of the simple sequence repeat (SSR) data. This was evident in the lack of diversity within the mtDNA COI sequence data regardless of its geographic collection locality. In addition, analyses of population allele frequencies

indicated that all populations in Lineage A show reduced heterozygosity accompanied by the loss of rare alleles, which is the characteristic signature of populations which have experienced a recent bottleneck (Roderick and Navajas 2003; Arca et al. 2015).

The fact that *L. invasa* reproduces clonally (at least a large part of the population) could help to overcome the effects of inbreeding and explain the genetic uniformity in Lineage A. This genetic uniformity could be due to an initial bottleneck caused by low genetic variation present when a small number of individuals are introduced (Caron et al. 2013). In some cases, however, bottlenecks of an intermediate nature may be favourable for invasive organisms because deleterious alleles, which may cause or lead to inbreeding, could be purged in the process (Facon et al. 2011). Endosymbionts, such as *Rickettsia*, have been observed in *L. invasa* (Nugnes et al. 2015) and these have also been shown to be drivers of genetic bottlenecks (Gotzek et al. 2015). Clonal reproduction may however also lead to the accumulation of deleterious mutations in the population over time. Therefore, the prevention of new introductions is likely to be a good long term management strategy in combination with breeding for resistance.

Since *L. invasa* was first recorded in 2000, historical records have shown a rapid movement of *L. invasa* through east Africa in a southerly direction; Kenya (2002), Uganda (2004), South Africa (2007), Zimbabwe (2010) and Mozambique (2010) (Dittrich-Schröder et al. 2012b). The high level of similarity and low genetic diversity of the specimens recorded from these countries support the assumption that they share a historical connection. Approximate Bayesian Computation indicated that further introductions, at least in the African continent, were as a result of a colonisation event from Israel.

The genetically uniform and clonal Lineage A has been remarkably successful at spreading globally. Genetic variation in recently invaded populations is usually low in organisms reproducing by asexual or parthenogenetic means (Caron et al. 2013). This has also been observed in other invasive hymenopterans (Auger-Rozenberg et al. 2012; Leach et al. 2012). Asexuality (thelytoky), the production of only female offspring and the possibility to generate high population numbers rapidly in a new environment are believed to be some of the important traits responsible for the success and establishment of various invasive species (Holway and Suarez 1999; Heimpel and Lundgren 2000; Reitz and Trumble 2002; Rabeling and Kronauer 2013). Asexuality allows favourable alleles to be conserved, which is of importance in stable environments, such as in plantation forestry, where predominantly clonal material is deployed. This is in contrast to sexual reproduction that would result in favourable allelic combinations being lost due to recombination and segregation (Stouthamer 1993).

Lineage B

It is unclear from historical records when or where Lineage B of *Leptocybe* sp. first appeared outside its native range. However, its dominance in South-East Asia suggests that this is most likely the region which it was first introduced. The wasp was first reported from China at the border of Vietnam in 2007 (Zheng et al. 2014). The present study has shown that Lineage B has now also been introduced into Africa, where it is present in Ghana and South Africa. This pattern of multiple independent introductions mirrors that of other plantation pests in the Southern Hemisphere, such as *Bradysia difformis* (Hurley et al. 2007), *Thaumastocoris peregrinus* (Nadel et al. 2009) and *Sirex noctilio* (Boissin et al. 2012). Other invasive forest pests for which patterns of multiple introductions have been recorded include the horse chestnut leafminer, *Cameraria ohridella* (Valade et al. 2009), as well as the Asian longhorn beetle, *Anoplophora glabripennis* (Carter et al. 2010), illustrating the extent of the problem regarding the global movement of forest pests (Liebhold et al. 1995; Wingfiel et al. 2008) and other invasive species (Arca et al. 2015; Auger-Rozenberg et al. 2012).

The presence of Lineage B in Ghana is intriguing. In its invasive range, Lineage B was first reported in Africa in 2013 from Ghana and later in 2015 from South Africa. Beforehand, Lineage B had been collected only from Southeast Asia. Its presence in two countries in Africa is interesting considering that Lineage A is dominant in all other regions in Africa. In the past, exotic tree species such as *Tectona grandis*, *Gmelina arborea*, *Eucalyptus* spp. and *Cedrela odorata* have been planted on a small scale in Ghana (Anonymous, 2009). There has been considerable expansion of *Eucalyptus* planting (amongst other exotic tree species) occurring in Ghana and Sierra Leone since 2010 (http://www.miroforestry.com,

accessed May 2017). *Eucalyptus* plantations in Ghana are relatively isolated in terms of major surrounding *Eucalyptus* plantations, with the closest minor plantations being in Togo and Nigeria (www.git-forestry.com, accessed May 2017). The phylogenetic grouping of the Ghana and other Lineage B *Leptocybe* sp. specimens, as well as the absence of Lineage A suggests that plant material from Asia was imported into Ghana, possibly linked to the start of the recent expansion of *Eucalyptus* plantations. The presence of Lineage B of *Leptocybe* sp. in *Eucalyptus* plantations in Ghana and South Africa potentially threatens control efforts in the rest of Africa. This is especially because forest pests have been shown to move very effectively throughout Africa after they have initially become established (Nadel et al. 2009; http://www.fao.org/forestry/fisna/26061/en/, accessed May 2017).

The presence of Lineage B of *Leptocybe* sp. was first noticed in South Africa in 2015. Previously, Lineage A was the only form of the insect type present among extensive collections since its first record in the country in 2007. The appearance of the two lineages at different time points suggests that multiple introductions have taken place into South Africa. Furthermore, the presence of both lineages in the same geographical area highlights the possibilities for admixture (Keller and Taylor 2010). Further investigation is required to determine the efficacy and effectiveness of the current control methods that have been developed for *L. invasa* Lineage A and how they might apply to Lineage B. This will be especially pertinent when considering the deployment of resistant clonal planting material.

Further work, examining the host preference, host range, presence and success of natural enemies of Lineage B is necessary to understand the niche of each lineage and to determine the extent of competitive exclusion between the lineages.

Mitochondrial Haplogroup B of *Leptocybe* sp. can be subdivided into Haplogroup B1 (consisting of samples from China, Ghana, Laos, Malaysia, Thailand, South Africa and Vietnam) and B2 (consisting of samples from various localities in Australia). Haplogroup B1 in the invaded range is more closely related to the currently sampled native populations (Haplogroup B2) than to Haplogroup A that represents samples from Israel and other non-native populations. This suggests that at least the same cryptic species, or evolutionary Haplogroup, of *Leptocybe* sp. has been identified in Australia, even though the source population of Haplogroup B1 has yet to be located. In the case of Haplogroup A (*L. invasa*), neither the source population, nor potentially the species, has been identified from the native range and this is despite the fact that extensive sampling has been conducted over at least a decade. The difficulty to obtain specimens of *Leptocybe* Lineage A and B in Australia can be attributed to its natural enemies keeping populations low, or, the natural range of this species may not have been sampled yet.

Admixture in the invasive range

ABC scenario testing using nuclear data indicated that admixture between Lineage A and the south East Asian part of Lineage B was occurring in Laos. The two distinct invasive lineages (A and B) of *Leptocybe* emerging from this study most likely represent two distinct introductions from distinct source populations. It seems highly likely that there has been a second introduction from the origin (i.e. not from another invaded population) into Asia as indicated by the grouping of specimens in the neighbour joining tree. This would have been subsequent to the Israel introduction and subsequent spread of that population.

Admixture in the invasive range holds many possible benefits for invasive populations. Keller and Taylor (2010) showed that multiple introductions resulted in admixture between distinct lineages of the cosmopolitan weed, *Silene vulgaris*. They further showed that admixture was a key factor driving the success of an invasive species due to improved fitness as a result of hybrid vigour. Verhoeven et al. (2011) suggest that invasive species, especially in the early stages of invasion, benefit considerably from admixture. Their work showed that native populations select for local adaptation, which ensures a higher level of fitness, although it carries the cost of inbreeding. Invasive populations are able to overcome the disadvantages associated with inbreeding depression by admixing in the invaded range (Verhoeven et al. 2011).

Admixture in the invasive range can also be unfavourable. This is especially the case in countries such as Israel where two biological control agents, *Quadrastichus mendeli* Kim & La Salle (Hymenoptera: Eulophidae) and *Selitrichodes kryceri* Kim & La Salle (Hymenoptera: Eulophidae), have

managed to effectively lower the level of damage caused by *Leptocybe invasa* (Kim et al. 2008). The success of this biological control programme could be at risk if the second lineage of *Leptocybe* were to spread to Israel because the potential exists for the two lineages to interbreed. The parasitoids have been sourced from Lineage B in Australia, and should therefore be well-adapted to that lineage, however no data on relative parasitism exists supporting this.

Due to the gall-forming nature of this wasp, feasible control measures are limited to two approaches, the planting of resistant and tolerant plant material, and the use of biological control (Dittrich-Schröder et al. 2012b). Both approaches require substantial time to develop and implement, and may only be effective for one lineage. The results suggest that improved quarantine measures and the prevention of new introductions and further spread of the pest, would also be valuable for future management. As the reproductive mechanism of *Leptocybe* spp. is not yet fully understood, the introduction of further alleles, potentially more suitable and successful in the invasive environment, and then subsequent clonal reproduction could lead to a rapid increase in pest populations.

Conclusions

The increase in worldwide travel and trade has been highlighted recently as a key factor aiding the introduction of invasive species (Liebhold et al. 2012). These data and observations of the *L. invasa* and *Leptocybe* sp. global invasion revealed in this study highlight the complex patterns of movement of invasive insect pests and their ability to move around the world rapidly, often resulting in admixture of different lineages. In this sense, the *Leptocybe* species serve as a model illustrating a number of important factors. Firstly, the success of an invasive pest, which is genetically homogeneous in its invasive range, is not necessarily hampered by an asexual mode of reproduction nor by a lack of genetic diversity. Secondly, the global movement of species invariably results in 'bridgehead populations', which may lead to admixtures of lineages and emergence of genotypes more suitable to other previously uninhabited environments. Thirdly, there appear to be two lineages of *Leptocybe* species moving globally, which could make current biological control efforts based on highly specific natural enemies, and the use of resistant planting stock, ineffective over time.

The results of this study highlight the danger of moving plant material and potentially further increasing genetic diversity of invasive populations. Even movement between areas where the pest already occurs carries this danger. Quarantine clearly needs to consider populations and cryptic species diversity. Secondly, this work serves as a guideline, systematically highlighting steps to be followed and analyses that should be conducted in order to characterize invasion pathways of newly invasive organisms for which this information is not yet available.

Data accessibility

- Mitochondrial Sequences were submitted to GenBank (Accession numbers MH093001-MH093481)

Acknowledgements We thank the following individuals for supplying specimens without which this study would not have been possible: Dr. Philip Nyeko, Dr. Zvi Mendel, Dr. Eston Mutitu, Prof. Jolanda Roux, Dr. Orlando Campolo, Ms. Titiporn Saimanee, Dr. John LaSalle, Ms. Nicole Fisher, Prof. Stefan Neser and Dr. Carlos Wilcken. We are grateful to Mrs. Alisa Postma Smidt for assembling the *L. invasa* genome, Dr. Irene Barnes for assistance with microsatellite development as well as to Mrs. Renate Zipfel and Dr. Kerry Reid for valuable discussions and assistance with microsatellite scoring. Members of the Tree Protection Co-operative Programme (TPCP), the THRIP Initiative of the Department of Trade and Industry and the National Research Foundation (NRF) (NRF Grant Number 88227) are acknowledged for providing funding.

Author contributions GDS, TBH, BS, BH and MJW designed research; GDS performed research; GDS, BS and TBH analysed data; BPH, MJW, SL and HFN provided material and contributed to interpretation of the results; all authors contributed to the writing of the manuscript.

References

Ács Z, Challis RJ, Bihari P, Blaxter M, Hayward A, Melika G, Csóka G, Pénzes Z, Pujade-Villar J, Nieves-Aldrey J, Schönrogge K, Stone GN (2010) Phylogeny and DNA barcoding of inquiline oak gallwasps (Hymenoptera: Cynipidae) of the Western Palearctic. Mol Phylogenet Evol 55:210–225

Akhtar MS, Patankar NV, Gaur A (2012) Observations on the biology and male of Eucalyptus gall wasp *Leptocybe*

invasa Fisher and La Salle (Hymenoptera: Eulophidae). Indian J Entomol 74:173–175

Arca MA, Mougel F, Guillemaud T, Dupas S, Rome Q, Perrard A, Muller F, Fossoud A, Capdevielle-Dulac C, Torres-Leguizamon M, Chen XX, Tan JL, Jung C, Villemant C, Arnold G, Silvain JF (2015) Reconstructing the invasion and the demographic history of the yellow-legged hornet, *Vespa velutina*, in Europe. Biol Invasions. https://doi.org/10.1007/s10530-015-0880-9

Auger-Rozenberg MA, Boivin T, Magnoux E, Courtin C, Roques A, Kerdelhue C (2012) Inferences on population history of a seed chalcid wasp: invasion success despite a severe founder effect from an unexpected source population. Mol Ecol 21:6086–6103

Boissin E, Hurley B, Wingfield MJ, Vasaitis R, Stenlid J, Davis C, De Groot P, Ahumada R, Carnegie A, Goldarazena A, Klasmer P, Wermelinger B, Slippers B (2012) Retracing the routes of introduction of invasive species: the case of the *Sirex noctilio* woodwasp. Mol Ecol 21:5728–5744

Brouat C, Tollenaere C, Estoup A, Loiseau A, Sommer S, Soanandrasana R, Rahalison L, Rajerison M, Piry S, Goodman SM, Duplantier J-M (2014) Invasion genetics of a human commensal rodent: the black rat *Rattus rattus* in Madagascar. Mol Ecol 23:4153–4167

Brodeur J (2012) Host specificity in biological control: insights from opportunistic pathogens. Evolutionary Applications 5:470–480

Caron V, Norgate M, Ede FJ, Nyman T, Sunnucks P (2013) Novel microsatellite DNA markets indicate strict parthenogenesis and few genotypes in the invasive willow sawfly *Nematus oligospilus*. Bull Entomol Res 103:74–88

Carter M, Smith M, Harrison R (2010) Genetic analyses of the Asian longhorn beetle (Coleoptera, Cerambycidae, *Anoplophora glabripennis*), in North America, Europe and Asia. Biol Invasions 12:1165–1182

Chown SL, Hodgins KA, Griffin PC, Oakeshott JG, Byrne M, Hoffmann AA (2015) Biological invasions, climate change and genomics. Evol Appl 8:23–46

Cognato AI (2006) Standard percent DNA sequence divergence for insects does not predict species boundaries. J Econ Entomol 99:1037–1045

Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M, Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism. DNA sequence and microsatellite data. Bioinformatics 30:1187–1189

de Meeûs T, Prugnolle F, Agnew P (2007) Asexual reproduction: genetics and evolutionary aspects. Cell Mol Life Sci 64:1355–1372

Dittrich-Schröder G, Wingfield MJ, Klein H, Slippers B (2012a) DNA extraction techniques for DNA barcoding of minute gall-inhabiting wasps. Mol Ecol Resour 12:109–115

Dittrich-Schröder G, Wingfield MJ, Hurley BP, Slippers B (2012b) Diversity in *Eucalyptus* susceptibility to the gall-forming wasp *Leptocybe invasa*. Agric For Entomol 14:419–427

Dittrich-Schröder G, Harney M, Neser S, Joffe T, Bush S, Hurley BP, Wingfield MJ, Slippers B (2014) Biology and host preference of *Selitrichodes neseri*: a potential biological control agent of the Eucalyptus gall wasp, *Leptocybe invasa*. Biol Control 78:33–41

Earl D, von Holdt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCRURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. Evolut Bioinform Online 1:47–50

Facon B, Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, Vitalis R, Guillemaud T, Lundgren JG, Estoup A (2011) Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. Curr Biol 21:424–427

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8:92–94

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Garnas JR, Hurley BP, Slippers B, Wingfield MJ (2013) Biological control of forest plantation pests in an interconnected world requires greater international focus. Int J Pest Manag 58:211–223

Garnas JR, Auger-Rozenberg M, Roques A, Bertelsmeier C, Wingfield MJ, Saccaggi DL, Roy HE, Slippers B (2016) Complex patterns of global spread in invasive insects: eco-evolutionary and management consequences. Biol Invasions 18:935–952

Gotzek D, Axen HJ, Suarez AV, Helms Cahan S, Shoemaker D (2015) Global invasion history of the tropical fire ant: a stowaway on the first global trade routes. Mol Ecol 24:374–388

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Available from http://www.mbio.ncsu.edu/BioEdit/bioedit.html

Hastings JM, Schultheis PJ, Whitson M, Holliday CW, Coelho JR, Mendell AM (2008) DNA barcoding of new world cicada killers (Hymenoptera: Crabronidae). Zootaxa 1713:27–38

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Ser B 270:313–321

Heimpel GE, Lundgren JG (2000) Sex ratios of commercially reared biological control agents. Biol Control 19:77–93

Hoffmann AA, Reynolds KT, Nash MA, Week AR (2008) A high incidence of parthenogenesis in agricultural pests. Proc R Soc B 275:2473–2481

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. Biotechniques 46:511–517

Holway DA, Suarez AV (1999) Animal behavior: an essential component of invasion biology. Trends Ecol Evol 14:328–330

Hurley BP, Govender P, Coutinho TA, Wingfield BD, Wingfield MJ (2007) Fungus gnats and other Diptera in South African

forestry nurseries and their possible association with the pitch canker fungus. S Afr J Sci 103:43–46

Hurley BP, Garnas J, Wingfield MJ, Branco M, Richardson DM, Slippers B (2016) Increasing numbers and intercontinental spread of invasive insects on eucalypts. Biol Invasions 18:921–933

Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. J Evol Biol 8:1720–1731

Kim I (2008) Evolution of gall inducing Eulophidae (Hymenoptera: Chalcidoidea) on Myrtaceae in Australia. PhD thesis, Australian National University

Kim I, Mendel Z, Protasov A, Blumberg D, La Salle J (2008) Taxonomy, biology, and efficacy of two Australian parasitoids of the eucalyptus gall wasp, *Leptocybe invasa* Fisher & La Salle (Hymenoptera: Eulophidae). Zootaxa 1910:1–20

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179–1191

Kronauer DJC, Pierce N, Keller L (2012) Asexual reproduction in introduced and native populations of the ant *Cerapachys biroi*. Mol Ecol 21:52221–55235

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large data sets. Bioinformatics 30:3276–3278

Leach IM, Ferber S, van de Zande L, Beukeboom LW (2012) Genetic variability of arrhenotokous and thelytokous *Venturia canescens* (Hymenoptera). Genetica 140:53–63

Liebhold AM, MacDonald WL, Bergdahl D, Mastro VC (1995) Invasion by exotic forest pests: a threat to forest ecosystems. For Sci Monogr 30:1–49

Liebhold AM, Brockerhoff EG, Garrett LJ, Parke JL, Britton KO (2012) Live plant imports: the major pathways for forest insect and pathogen invasions of the US. Front Ecol Environ 10:135–143

Lombaert E, Guillemand T, Thiams CE, Lawson Handley LJ, Li J, Wang S, Pang H, Goryacheva I, Zakharov IA, Jousselin E, Poland RL, Migeon A, van Lenteren J, de Clercq P, Berkvens N, Jones W, Estoup A (2011) Inferring the origin of populations introduced from a genetically structures native range by approximate Bayesian computation: a case study of the invasive ladybird *Harmonia axyridis*. Mol Ecol 20:4654–4670

Mendel Z, Protasov A, Fisher N, La Salle J (2004) Taxonomy and biology of *Leptocybe invasa* gen. & sp. n. (Hymenoptera: Eulophidae), an invasive gall inducer on Eucalyptus. Aust J Entomol 43:101–113

Merçot H, Poinsot D (2009) Infection by *Wolbachi*a: from passenger to residents. C R Biol 332:284–297

Meyerson LA, Mooney HA (2007) Invasive alien species in an era of globalization. Front Ecol Environ 5:199–208

Nadel RL, Slippers B, Scholes MC, Lawson SA, Noack AE, Wilcken CF, Bouvet JP, Wingfield MJ (2009) DNA barcoding reveals source and patterns of *Thaumastocoris peregrinus* invasions in South Africa and South America. Biol Invasions 12:1067–1077

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nugnes F, Gebiol M, Monti MM, Gaultieri L, Giorgini M, Wang J, Bernardo U (2015) Genetic diversity of the invasive Gall Wasp *Leptocybe invasa* (Hymenoptera: Eulophidae) and its Rickettsia endosymbiont, and associated sex-ratio differences. PLoS ONE 10(5):e0124660

Nyeko P (2005) The cause, incidence and severity of a new gall damage on eucalyptus species at Oruchinga refugee settlement in Mbarara district, Uganda. Ugandan J Agric Sci 11:47–50

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061–1067

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Polzin T, Daneschmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett 31:12–20

Pritchard JK, Stephens M, Donnely P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rabeling C, Kronauer DJC (2013) Thelytokous parthenogenesis in eusocial Hymenoptera. Annu Rev Entomol 58:273–292

Rehan SM, Sheffield CS (2011) Morphological and molecular delineation of a new species in the *Ceratina dupla* species-group (Hymenoptera: Apidae: Xylocopinae) of eastern North America. Zootaxa 2873:35–50

Reitz SR, Trumble JT (2002) Competitive displacement among insects and arachnids. Annu Rev Entomol 47:435–465

Roderick GK, Navajas M (2003) Genes in new environments: genetics and evolution in biological control. Nat Rev Genet 4:889–899

Rozas J, Sánchez-DelBarrio JC, Messegeur X, Rozas R (2003) DnaSP, DNA polymorphism nalyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Sangtongpraow B, Charernsom K, Siripatanadilok S (2011) Longevity, fecundity and development time of Eucalypyus Gall Wasp, *Leptocybe invasa* Fisher & La Salle (Hymenoptera: Eulophidae) in Kanchanaburi province, Thailand. Thai J Agric Sci 44:15–163

Schrey NM, Schrey AW, Heist EJ, Reeve JD (2010) Genetic heterogeneity in a cyclical forest pest, the southern pine beetle, *Dendroctonus frontalis*, is differentiated into east and west groups in the southern unites States. J Insect Sci 11:110

Smith MA, Hallwachs W, Janzen DH, Segura RB (2013) DNA barcoding a collection of ants (Hymenoptera: Formicidae) from Isla del Coco, Costa Rica. Fla Entomol 96:1500–1507

Staden R (1996) The Staden sequence analysis package. Mol Biotechnol 5:233–241

Stouthamer R (1993) The use of sexual versus asexual wasps in biological control. Entomophaga 38:3–6

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis

using likelihood, distance, and parsimony methods. Mol Biol Evol 10:2731–2739

Thu PQ, Dell B, Burgess TI (2009) Susceptibility of 18 eucalypt species to the gall wasp *Leptocybe invasa* in the nursery and young plantations in Vietnam. ScienceAsia 35:113–117

Tobin PC, Berec L, Liebhold AM (2011) Exploiting Allee effects for managing biological invasions. Ecol Lett 14:615–624

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2001) Relationships among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. Mol Ecol 10:2151–2161

Turčinavičiene J, Radzevičiute R, Budriene A, Budrys E (2016) Species identification and genetic differentiation of European cavity-nesting wasps (Hymenoptera: Vespidae, Pompilidae, Crabronidae) inferred from DNA barcoding data. Mitochondrial DNA Part A 27:476–482

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. Nucleic Acids Res 40:e115

Valade R, Kenis M, Hernandez-Lopez A, Augustin S, Mari Mena N, Magnoux E, Rougerie R, Lakatos F, Roques A, Lopez-Vaamonde C (2009) Mitochondrial and microsatellite DNA markers reveal a Balkan origin for the highly invasive horse-chestnut leaf miner *Cameraria ohridella* (Lepidoptera, Gracillariidae). Mol Ecol 18:3458–3470

Verhoeven KJF, Macel M, Wolfe LM, Bierre A (2011) Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. Proc R Soc B 278:2–8

Wingfiel MJ, Slippers B, Hurley BP, Coutinho TA, Wingfield BD, Roux J (2008) Eucalypt pests and diseases: growing threats to plantation productivity. South For 70:139–144

Wingfield MJ, Brockerhoff EG, Wingfield BD, Slippers B (2015) Planted forest health: the need for a global strategy. Science 349:832–836

Yek SH, Slippers B (2014) Biocontrol opportunities to study microevolution in invasive populations. Trends Ecol Evol 29:429–430

Zheng XL, Yang ZD, Li J, Xian ZH, Yang J, Liu JY, Su S, Wang XL, Lu W (2014) Rapid identification of both sexes of *Leptocybe invasa* Fisher & La Salle (Hymenoptera: Eulophidae: Tetrastichinae): a morphological perspective. Afr Entomol 22:643–650

Zhou X, Kjer KM, Morse JC (2007) Associating larvae and adults of Chinese *Hydropsychidae* caddisflies (Insecta: Trichoptera) using DNA sequences. J North Am Benthol Soc 26:719–742