Multiple Incursions and Putative Species Revealed using a Mitochondria and Nuclear Phylogenetic approach to the Warehouse Beetle (Coleoptera: Dermestidae) Trapping Program in Australia.

Mark A Castalanelli[1,2,5]*, Katarina M Mikac[3], Andrew M Baker[4], Kylie Munyard[2], Mike Grimm[5] and David M Groth[2],

Cooperative Research Centre for National Plant Biosecurity, Deakin, ACT, Australia[1]

Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, Curtin University of Technology, Perth, Australia[2]

Institute for Conservation Biology and Environmental Management, University of Wollongong, Wollongong, NSW Australia[3]

Faculty of Science and Technology, Queensland University of Technology, Brisbane, QLD, Australia[4]

Division of Biosecurity and Research, Department of Agriculture Western Australia, WA, 6151, Australia[5]

Running Title: Phylogenetic analysis of Australian Warehouse Beetle Populations

Key words: Dermestidae, population, cryptic, molecular, eradication, invasive, correct diagnosis

Abstract: 168 words

25 ## Abstract

26 The Warehouse beetle *Trogoderma variabile* (Coleoptera: Dermestidae) is an internationally
27 significant invasive pest of grain storage structures, packed goods and stored grain. When it
28 was first documented in Australia at Griffith, New South Wales, in 1977 an eradication
29 campaign was initiated. After several years and considerable effort the eradication campaign
30 was abandoned. To monitor the presence and spread of *T. variabile,* surveys were carried
31 out by government agencies, in 1992 and 2002. Comparing data from these surveys, it was
32 concluded that the distribution of *T. variabile*, identified via morphological characters alone,
33 had doubled in most Australian states. We used samples from the 2002 survey to conducted
34 a phylogenetic study using partial sequences of mitochondrial genes Cytochrome Oxidase I
35 and Cytochrome B, and the nuclear gene 18S, to examine the distribution and dispersal of
36 *T. variabile* and detect the presence of misidentified species . Based on our molecular results
37 we have shown that only 47% of the samples analysed were *T. variabile* and the remaining
38 were a mixture of six putative species. In addition, *T. variabile* was found in only 78% of the
39 trapping sites. We discuss the importance of correct diagnosis in relation to the eradication
40 campaign.
41

42 ## Introduction

43 The incursion of economically destructive arthropods into Australia is becoming more
44 frequent due to the increase of international trade in goods, particularly from developing
45 countries (Mack, Simberloff et al. 2000). Once established, invasive pests can cause
46 considerable environmental and economic damage as they generally proliferate and spread
47 due to factors such as lack of natural enemies and climate or edaphic constraints (Mack,
48 Simberloff et al. 2000). One such invasive arthropod is *Trogoderma variabile* (Ballion), the
49 warehouse beetle (Coleoptera: Dermestidae), which infests grain storage structures, stored
50 grain, and packed goods (Wright 1994). In Australia *T. variabile* is regarded as a minor and

51 persistent pest. However, it is of considerable concern because it could mask the presence
52 of the more damaging and invasive Khapra beetle, *Trogoderma granarium* (Everts), due to
53 high morphological similarity between the species (Rees, Starick et al. 2003). It has been
54 estimated that an introduction of the Khapra beetle (*Trogoderma granarium* Everts) into
55 Australia could cause losses approaching AU$2 billion per annum (Cook 2003).
56
57 When *T. variabile* was first documented in Australia at Griffith, New South Wales, in 1977
58 and then again in 1979 on the other side of the continent at Morawa, Western Australia,
59 Australian State and Federal agricultural departments attempted to control and eradicate this
60 pest (Hartley and Greening 1983; Wright 1994; Rees, Starick et al. 2003). After several
61 years, the eradication campaign was terminated when surveys showed that the beetle had
62 spread to surrounding areas (Hartley and Greening 1983). Although the eradication program
63 was cancelled, surveys have been routinely carried out to monitor the distribution of the pest.
64 An extensive survey, conducted by The Commonwealth Scientific and Industries Research
65 Organization (CSIRO) from 2001 to 2003, discovered that in Victoria and South Australia
66 abundance of *T. variabile* had doubled since 1990, suggesting that little progress has been
67 made in limiting its spread (Rees, Starick et al. 2003).
68
69 In order to understand the apparently uncontrollable dispersal of *T. variabile* throughout
70 Australia there is a need to better understand the mechanisms underlying their invasion
71 biology. Without a detailed understanding of these mechanisms, efforts to control the
72 dispersal of an organism can be rendered useless (Tsutsui, Suarez et al. 2000; Tsutsui,
73 Suarez et al. 2001; Tsutsui, Suarez et al. 2003; Schutze, Mather et al. 2006). Studies on
74 other invasive insect species have shown that molecular markers can provide valuable
75 information about population structure, gene flow and dispersal pathways. In turn, this can
76 provide an insight into the dispersal of species and determine the presence of cryptic species
77 (Loxdale and Lushai 1998; Mikac and Clarke 2006; Mikac and FitzSimmons 2010).
78

79 In this study we selected three genes, two mitochondrial and one nuclear, to investigate the
80 evolutionary history of *T. variabile* and determine the presence of any cryptic species. The
81 mitochondrial gene Cytochrome Oxidase I (COI) is frequently used to understand
82 mechanisms underlying invasion biology (Jenkins, Jones et al. 2007; Nadel, Slippers et al.
83 2009). Therefore, COI was selected to resolve patterns of geographical distribution and
84 determine if the dispersal of *T. variabile* throughout Australia has been the result of multiple
85 incursions and/or subsequent human-aided dispersal. We predicted this because the
86 distance between grain storage facilities within Australia ranges from 9 to 3760km, which far
87 exceeds the average dispersal distance for individual *T. variabile* of only 75m (Campbell and
88 Mullen 2004). We also wanted to test the possibility that samples morphologically as *T.*
89 *variabile* could in fact be *Trogoderma* spp. endemic to Australia, because there are currently
90 52 described endemic species and at least another 100 not yet formally described. While
91 COI is well-suited for examining intra- and inter-specific variation at both the species and
92 genus level, we also included Cytochrome B (Cyt b), which evolves more slowly than COI, as
93 well as the nuclear gene 18S, which together should help resolve deeper nodes and more
94 clearly differentiate among evolutionary trajectories.
95

96 Materials and Methods

97 *Collection*

98 One hundred and forty one specimens were collected from 27 grain storage sites throughout
99 Australia (Table 1.). These samples were collected between 2001-2003 using baited sticky
100 flight traps as described by Rees and Colleagues (Rees, Starick et al. 2003). The grain
101 storage sites (Figure 2 C) were located in Western Australia (Varley), South Australia
102 (Balaklava, Crystal Brook, Garrah and Port Adelaide), New South Wales (Ariah Park,
103 Booroowa, Bowman, Caracabal, Coleambally, Coolamon, Cowra, Griffith, Lockhart, Narrabri
104 West, Quirindi, Tamworth, The Rock, West Wyalong and Willow Tree) and Victoria
105 (Dimboola, Geelong, Little Desert National Park (LDNP), Manangatang, Nhill, Rainbow, and

106 Swan Hill). Ten additional specimens were collected from nine Western Australian sites
107 (Carnamah, Dalwallinu, Geraldton, Katanning, Perth, Pithara, Three Springs, Wagin and
108 Wyalkatchem) in 2007 as part of an ongoing trapping program. These ten samples were the
109 only samples that were verified as *T. variabile* using traditional morphological keys (Szito
110 2007). These *T. variabile* specimens were the reference samples used for this study. Single
111 specimens of *T. granarium* and *Anthrenus verbasci*, both non native pest species, were
112 taken from the Department of Agriculture and Food Western Australian insect collection.
113
114 ***DNA extraction***
115 The number of specimens used in this study ranged from 1 to 8 per site, dependent on the
116 number of specimens collected. DNA was isolated from the whole individual by crushing
117 them in 5% Chelex beads (Biorad; (Walsh, Metzger et al. 1991).
118
119 DNA was extracted from the ten *T. variabile* samples collected in Western Australia were and
120 the single specimens of *T. granarium* and *A. verbasci* using ANDE (Castalanelli, Severtson
121 et al. 2010) (www.ande.com.au) as per manufactures instructions.
122
123
124 ***Amplification of COI***
125 A semi-nested Polymerase Chain Reaction (PCR) was used to amplify a fragment of COI
126 using the primers UAE5 (5'AGTTTTAGCAGGAGCAATTACTAT 3') and UAE10 (5'
127 TCCAATGCACTAATCTGCCATATTA 3') for the first round of PCR (PCR1) followed by
128 amplification of first round PCR product using the primer combination UEA7: (5'
129 TACAGTTGGAATAGACGTTGATAC 3') and UAE 10 in a second PCR amplification (PCR2)
130 (Lunt , Zhang et al. 1996). PCR were carried out in 12.5 µl volumes using 25-50 ng, which
131 was measured using a NanoDrop spectrophotometer (Thermo Fisher Scientific Inc), of DNA
132 (PCR1) or 1 µl of PCR product (PCR2), 0.3 µM each primer (PCR1: UAE5 and UAE10;

133 PCR2: UAE7 and UAE10), 0.2 mM each dNTP, 1x PCR buffer (New England Biolabs), 3 mM
134 $MgCl_2$, and 1.5 units of Taq DNA polymerase (New England Biolabs). PCR1 cycling
135 conditions consisted of an initial denaturation step of 94 ºC for 3 min, followed by 30 cycles of
136 94 ºC for 40 s, 54 ºC for 1 min 40 s and 72 ºC for 60 s, with a final extension temperature of
137 72 ºC for 8 min. PCR2 cycling conditions consisted of an initial denaturation step of 94 ºC for
138 3 min followed by 40 cycles of 94 ºC for 40 secs, 52 ºC for 1 min 40 s and 72 ºC for 60 s, with
139 a final extension temperature of 72 ºC for 8 mins. Ten percent of the final volume of the
140 PCR2 products were electrophoresed on 2% w/v agarose containing 1 x TBE buffer and
141 visualized with SYBR safe (Molecular Probes Inc.). Unpurified PCR products were sent to
142 Macrogen (Korea) for cleanup and sequencing (see the 'sequencing' section below for
143 further details).
144
145 ***Amplification of Cytochrome B and 18S***
146 Amplification of the Cytochrome B gene was performed using PCR primers CB-J-10933 and
147 CB-N-11367 (Simon, Frati et al. 1994). The nuclear gene 18S was amplified using PCR
148 primers 18SF2 (5'TACCACATCCAAGGAAGG'3) and 18SR2
149 (5'CCTCTAACGTCGCAATAC'3). Amplification of the Cytochrome B and 18S genes was
150 conducted in a reaction mix comprising 1× polymerase buffer (Roche), 3mM $MgCl_2$, 0.2μM of
151 each primer and 0.5 Units of Faststart Hi Fidelity Taq polymerase (Roche). PCR cycling
152 conditions consisted of 45 cycles at 95 °C for 10 s, 45 °C for 10s and 72 °C for 30 s. Finally,
153 all PCR products were purified by the addition of 10U of Exonuclease I (NEB) and 2.5U of
154 Antarctic Phosphatase (NEB) and incubation for 30 min at 37°C followed by inactivation by
155 heating the reaction to 80°C for 20 min. Twenty percent of the final volume of the PCR
156 products were electrophoresed on 1.5% w/v agarose gels containing 1 x TBE buffer and
157 visualized with ethidium bromide.
158

159 ***Sequencing***

160 DNA sequencing was performed by Macrogen Inc (Korea) using an Applied Biosystems ABI
161 3730 48-capillary DNA analyser using Big Dye Terminator Technology according to the
162 manufacturer's protocols (Applied Biosystems).

163

164 ***Data Analysis***

165 Sequences were edited using CodonCode Aligner 3.0.3 (CodonCode Corporation) and
166 aligned using a built-in version of ClustalW (Thompson, Higgins et al. 1994). Paup 4.0
167 (Swofford 2003) was used to perform a Partition Homogeneity Test (PHT) to determine the
168 level of congruence between the three genes. The three genes were concatenated and
169 PAUP 4.0 (Swofford 2003) used to generate Parsimony and Maximum Likelihood gene trees.
170 In addition, PAUP 4.0 (Swofford 2003) was used to generate a COI parsimony gene tree
171 which was compared to a Bayesian COI gene tree generated by Mr Bayes 3.1.2 (Ronquist
172 and Huelsenbeck 2003). MEGA 4 (Tamura, Dudley et al. 2007) was used to calculate net
173 divergences between clades, TCS v1.21 (Clement, Posada et al. 2000) for network analysis
174 and an R- statistics (Team 2007) function that we designed to plot the distribution of
175 haplotypes. The Partition Homogeneity Test of the three genes was performed using the
176 parameters hsearch, randomseed=0 and nrep=1000. The parsimony and ML analysis of the
177 three concatenated genes was calculated under the following conditions: hsearch,
178 addseq=random, nrep=1000, swap=TBR, MaxTrees=1000, with the additional parameters
179 added for the bootstrapping analysis nchuck=5 and chuckscore=1. The ML tree construction
180 used likelihood settings that were selected by Modeltest 3.7 (Posada and Crandall 1998)
181 from the best-fit model (GTR+I+G) and were Lset Base=(0.2884 0.1771 0.2319); Nst=6;
182 Rmat=(0.1618 10.5497 2.0353 0.3522 1.6956); Rates=gamma; Shape=1.8432; and
183 Pinvar=0.6063. Bayesian analysis was performed on the COI sequence data using the
184 evolution model 4by4, gen=1,000,000, sample freq=100, sump burnin=2,500, and sumt
185 burnin=250, the bootstrapping values were compared to the parsimony results. Parsimony

186 analysis of the COI data used the same parameters as concatenated gene analysis. *A.*
187 *verbasci* was the out group for the generation of all trees. The R statistics program was used
188 to plot the distribution of each clade as well as the *T. variabile* haplotypes.
189

190 Results

191 *Sequence Data*

192 To reconstruct phylogenetic relationships of individuals collected during the 2001-2003
193 CSIRO trapping program, sections of two mitochondrial genes and one nuclear gene were
194 sequenced. The mitochondrial gene COI (620 bp, of which 264 were polymorphic) had 46
195 haplotypes from 87 individuals, with an average sequence divergence of 16.12% (S.E
196 0.87%). In comparison, the mitochondrial gene Cyt b (434 bp of which 171 were
197 polymorphic) produced fewer haplotypes (24) from 112 individuals, but a similar average
198 sequence divergence of 14.73% (S.E. 0.95%). For both mitochondrial genes, the
199 polymorphisms caused non-synonymous changes in the amino acid sequence. The 18S
200 nuclear fragment length was 518bp and as expected slower evolving, with only 7 nucleotide
201 sites were polymorphic from 31 individuals. The average sequence divergence for this gene
202 was only 3.9% (S.E. 0.17%). No insertions or deletions were observed in any gene fragment.
203

204 *Partition Homogeneity Test*

205 Partition Homogeneity Tests (PHT) were performed using Paup 4.0 (Swofford 2003) to
206 determine the level of congruence between each gene fragment. The PHT showed
207 congruence between the nuclear gene 18S and the two mitochondrial genes COI and Cyt b,
208 (p = 1 and 0.995 respectively). In comparison, the two mitochondrial genes were less
209 congruent (p = 0.643). With no significant difference between the genes trees, all sequences
210 were concatenated to create total evidence parsimony and ML trees.
211

212 ***All Genes Parsimony and ML Tree***

213 Because no significant difference was detected between the trees of individual genes, all
214 sequences were concatenated to create a total evidence parsimony tree (Fig 1). The
215 parsimony tree revealed 9 distinct clades (depicted using a unique letter or reference name)
216 that were separated by deep branch lengths and supported by high net divergence between
217 clades (from 7.56% to 18.44%; Table 2). Most ML bootstrapping values were greater than
218 84%, apart from the divergence of clade d from e and clade f from g (67% and 59%,
219 respectively). The *T. variable* reference specimens that we identified using classical
220 morphological techniques formed a distinct clade with only 47% of the samples. The pairwise
221 divergence between *T. variabile* and the other clades ranged from 12.93 – 17.28 % (ca.
222 mean 14.46 % S.E. 1.37), the lowest between it and the *T. granarium* reference sample. No
223 samples clustered with the *T. granarium* reference sample.
224

225 ***Gene Analysis***

226 The nine clades that were generated by the all gene tree analysis (Fig 1) were used to
227 examine the intra- and inter-pairwise divergence for both COI and Cyt b (Table 3). Inter clade
228 divergence values for COI and Cyt b ranged from 10.72 to 26.11% and 10.61 to 26.11%,
229 respectively. The mean divergence between clades was 1.61%, and in two cases Cyt b was
230 more divergent than COI (clades c and g). The mean intra -clade variation between the two
231 genes was identical (0.5%), with COI and Cyt b ranging from 0 to 0.19% and 0 to 0.4%,
232 respectively. Interestingly, Cyt b intra-clade variation for the *T. variabile* clade was zero
233 (n=54), while the divergence for COI ranged from 0 to 0.17%. Analysis of the 18S gene
234 fragment revealed no variation within each clade but variation between. The high inter clade
235 divergence values calculated for both mitochondrial genes suggest that 6 putative species
236 plus *T. variabile* were collected.
237

238 The distribution of each of these species was visualised on a geospatial distribution map (Fig
239 2). The majority of specimens that were placed in clades d and f (95% and 91%,
240 respectively) were observed in New South Wales, mainly in the southern part of the state.
241 Similarly, 73% of clade c was found in Victoria. Clade b was only observed at Port Adelaide
242 in South Australia and clade g was only located at Varley in Western Australia. The
243 specimens that clustered with the *T. variabile* reference specimens were found throughout
244 Australia, but, only at 78% of the sites sampled (Table 1).
245

246 ***COI parsimony tree***

247 While both the mitochondrial genes provided similar intra- and inter-specific divergence
248 values, we only used the COI data to generate a parsimony gene tree. This was because the
249 COI gene showed variation within the *T. variabile* clade and had 46 unique sequences,
250 compared to only 24 for Cyt b thus providing greater resolution. The COI parsimony tree
251 revealed the same distinct clades as were observed in the all gene parsimony tree (Fig 1).
252 Within the T. variabile clade, we observed 10 unique haplotypes, of which 4 were distantly
253 related. Divergence values for the 4 distantly related haplotypes ranged from 0.089 to
254 0.17%, which was evident from the long branch lengths (Fig 3).
255

Comment [ab1]: which haps are you referring the reader to in that tree?

256 **Genetic Structure of *T. variabile using COI***

257 COI was the only gene to resolve divergence within the *T. variabile* clade (Table 3), revealing
258 10 unique haplotypes (Table 5). Network analysis of these haplotypes with the program TCS
259 v1.21 (Clement, Posada et al. 2000) revealed one main haplotype (H4) with four direct
260 descendents (H2, H3, H6 & H9), and one derived haplotype (H7) separated by one missing
261 mutational step. One haplotype (H5) was distantly related to H4 (missing multiple mutational
262 steps) and three haplotypes (H1, H8, H10) were not connected in the network because their
263 divergence ranged outside the 95% parsimony limit imposed by the program (Fig 4 D). With
264 no connectivity, we are unable to determine the relatedness of these haplotypes to the rest of

265 the haplotypes. Haplotype 4 was found at all the eastern states collection sites (where we
266 identified the presence of *T. variabile* Table 4), except for Willow Tree that had only
267 haplotype 5. Tamworth and Swan Hill both contained additional haplotypes that showed no
268 connectivity. All Western Australian sites contained that same haplotype (H7) that was
269 unique to Western Australia (Table 5).
270

271 Discussion

272 The invasion of pest species not only causes adverse economic effects for agricultural
273 industries but threatens efforts to conserve native biodiversity and sustain natural
274 ecosystems (Mack, Simberloff et al. 2000; Armstrong and Ball 2005). To limit the effect of an
275 invasion, it may be deemed feasible to try and eradicate or limit the range of the pest species
276 through control measures and management programmes (Mack, Simberloff et al. 2000). The
277 success of an eradication program largely depends on understanding the ecological and
278 evolutionary pathways of the organism, as well as ensuring correct identification (Mack,
279 Simberloff et al. 2000; Schutze, Mather et al. 2006; Suarez, Holway et al. 2008). However,
280 the presence of cryptic species makes correct identification difficult, thus placing pressure on
281 management efforts, which could reduce the effectiveness of an eradication program
282 (Schutze, Mather et al. 2006). This study aimed to investigate the population structure of *T.*
283 *variabile* in Australia to determine what factors may have influenced the unsuccessful
284 outcome of the eradication campaign. The results revealed seven deeply distinct clades in
285 both the all gene parsimony trees and COI gene trees (Figs 1 and 3 respectively) with
286 average pairwise divergence between clades of 18.5% (S.E. 2.27) and 17.5% (S.E. 1.87%),
287 COI and Cyt b, respectively. These data, consistent at both nuclear and mitochondrial gene
288 loci, strongly suggest that seven distinct species were collected, of which only 47% were
289 clustered with the *T. variabile* reference sample. This is further supported by the fact that a
290 range of these deeply divergent lineages were found to occur in sympatry in south-eastern
291 Australia, suggesting they are presently reproductively isolated (see Figure 2). It is likely that

292 *T. variabile* specimens are comprised of one main population that was observed throughout
293 south-eastern Australia, with subsequent incursions at Tamworth (NSW), Willow Tree (NSW)
294 and Swan Hill (Vic). The haplotypes that were observed at Tamworth (H8 & H10), Willow
295 Tree (H5) and Swan Hill (H1) were too distantly related (multiple missing mutation steps or
296 no connectivity between haplotypes) to have been derived from the main population. Both
297 misidentification and multiple incursions would have a direct negative effect on the success
298 of the eradication problem.
299
300 Of the 153 individuals that were genotyped, only 46.7% were able to be grouped with the
301 specimens that we morphologically verified as *T. variabile.* Since the specimens were
302 macerated to liberate the DNA we can only theorise that other clades are distinct species.
303 However it is probable that the other specimens were not *T. variabile,* primarily due to the
304 large nucleotide percentage differences that were observed between clades (Tables 2 and
305 3). The average inter- and intra-specific divergence for the COI locus in this study was 18.5%
306 and 0.05%, respectively. In comparison, the inter and intra specific divergence levels for the
307 COI loci, previously determined for a multitude of insect pests (Cognato 2006), was 7.4% (2
308 – 24%) and 1.75% (ranging from 0.077 – 26%) respectively. The mean inter-clade
309 divergence observed in our study was more than double and the intra-clade divergence only
310 a fraction of that calculated by Cognato (2006), which further supports the presence of a
311 number of deeply divergent cryptic species in the present study. Furthermore, *T. variable* and
312 *T. granarium*, which can be differentiated using morphological keys, had a COI divergence of
313 18.8% (Table 3), the lowest between *T. variable* and any other clade. This offers promise
314 that with further study, morphological characters may be resolved that discriminate between
315 the other cryptic species discovered here making recognition and management on the
316 ground more feasible.
317
318 The six putative species accounted for 53% (n=82) of the individuals we analysed, which
319 could be attributed to both high trapping yields and the difficulty encountered when

320 identifying samples collected from a trapping program. The traps were placed outside of the
321 grain facilities, possibility to increase the number of *T. variabile* collected (Campbell and
322 Mullen 2004), but this also increases the number of native Dermestids collected since some
323 natives are attracted to the pheromone used in the traps (unpublished data). The number of
324 specimens collected in each trap was likely to have exceeded a few hundred (unpublished
325 data) making identification particularly difficult, especially when adopting a quick visual
326 screening approach. Morphological identification of *T. variabile* from other closely related
327 *Trogoderma* requires dissection and examination of the genitalia (Szito 2007), not just the
328 examination of external characteristics. Dissection and preparation of the genitalia for
329 examination requires time and skill and would be impractical to perform on the number of
330 samples collected at each site. If the eradication campaign adopted the same quick
331 screening methods the likelihood of misidentification would be high and potentially resulting
332 in the management of a non target species.
333
334 We are unable to resolve morphological characters separating the new species because the
335 method used for DNA extraction from the 2001 to 2003 trapping specimens required the
336 samples to be macerated, such that subsequent morphological examinations were
337 impossible. In the future, for *Trogoderma* species we recommend that a non-destructive
338 method, such as the recently developed ANDE (Castalanelli, Severtson et al. 2010), be
339 used. Use of such a non-destructive DNA extraction technique when this study was
340 commenced would have allowed us to revisit the samples and identify each specimen using
341 morphological techniques after grouping them into genetic lineages. Two years after the
342 initial Chelex extractions, additional DNA samples (from the same trapping program) were
343 extracted using the non-destructive ANDE method. However, the extracted DNA was
344 subsequently shown to be severely fragmented resulting in an unsuccessful amplification.
345 We believe that this can be attributed to the use of harsh chemicals, such as hexane or
346 limoene used to remove the specimens from the sticky traps (Szito 2007), combined with
347 long term storage in ethanol at room temperature. (sentence on extraction from sticky tapes)

Comment [BS2]: It does because it address the reviewers comment of why we didn't go back and reanalysis more samples and match the genetic profiles to a physical specimen

348 Only minimal genetic analysis has been conducted on the native and international
349 Dermestids, so until these sequences are detected again, with accompanying whole
350 specimens, the six new species identified in this work are putative.
351
352 In addition to the discovery of the six putative species, this study was able to investigate the
353 genetic structure of *T. variabile* populations throughout Australia. The analysis of *T. variabile*
354 populations suggested large-scale movement, a lack of genetic structure at Tamworth (NSW)
355 and Swan Hill (Vic), and a distantly related haplotype found at Willow Tree (NSW). The main
356 *T. variabile* haplotype (H4) was found at all eastern Australian sites where *T. variabile* was
357 present, except for Willow Tree in NSW. The distance between each of these sites ranged
358 from 9 – 1613 km, and considering that *T. variabile* has a limited dispersal range based on its
359 ecology (Campbell and Mullen 2004), human-aided transport is the only plausible dispersal
360 mechanism. With the continual movement of grain around the country (as feed, seed stock
361 and export commodity) the dispersal of *T. variabile* into new areas would be difficult to
362 control. Moreover, as the population grows, reinfestation of previously eradicated sites would
363 be almost inevitable.
364
365 The only example of isolation by distance was when we compared the Western Australian
366 specimens to those collected from the Eastern States. TCS analysis (Fig 4 d) showed that
367 the Western Australian haplotype was only two mutational steps from haplotype 4. While
368 TCS results suggest that the West and East populations are connected, we are unable to
369 confirm if the main population is the result of one incursion that has been spread across the
370 continent east to west or vice versa. Such confirmation would be dependent on finding the
371 missing haplotype, which may have been overlooked in this study due the small sample size.
372
373 The three haplotypes observed at Tamworth (NSW) and Swan Hill (Vic), which were not
374 related to the rest of population, indicate a lack of population structure at these sites (Fig 4
375 d). Since these haplotypes were too distantly related, TCS was unable to connect them to

376 any other haplotypes, suggesting that specimens from these sites are not derived from the
377 main population. Rather, the data suggest they are the result of subsequent incursions into
378 Australia because not enough time has elapsed since the initial incursion (1977) to account
379 for such diversity at these mitochondrial loci. First, the main haplotype (H4) was observed at
380 all other eastern state sites (where *T. variabile* was confirmed) as well as Tamworth and
381 Swan Hill where as the unconnected haplotypes are localised and yet to disperse,
382 suggesting that they are newer – such is the expectation based on coalescent theory (Page
383 and Holmes 1998). Second, the distance between Tamworth and Swan Hill is >1000km
384 suggesting that at least two incursions are likely to have occurred for the observed
385 haplotypes to be present. However, we are unable to confirm if the two unconnected
386 haplotypes (H8 & H10) at Tamworth are the result of multiple incursions or a single incursion
387 with multiple haplotypes. Haplotype 5, which was observed at Willow Tree (NSW), is likely to
388 be another separate incursion due to 3 missing mutational steps between itself and
389 haplotype 4. The estimated time taken for haplotype 4 to evolve into haplotype 5 is ~80,000
390 years (based on 1 million year per 1% divergence; (Brower 1994) therefore clearly
391 insufficient time has passed for this to occur since *T. variabile* was first documented in
392 Australia in 1977. With the lack of connectivity between haplotypes 1, 5, 8 and 10, and
393 haplotype 4, it is plausible that the *T. variabile* population in Australia is the result of at least 4
394 separate incursions.
395
396 With increasing international trade the likelihood of serious pests invading Australia will
397 inevitably increase. When an invasive species is detected, the succuss of control or
398 eradication is dependent on accurate identification (Walter 2003; Schutze, Mather et al.
399 2006). This study revealed that 53% of the specimens collected during the 2001-2003 *T.*
400 *variabile* trapping program were misidentified, and only 78% of the sites actually contained *T.*
401 *variabile*. We believe this to be the result of dealing with several hundred samples per site as
402 well as the use of a quick screening method to identify *T. variabile*. Without keying-out each
403 specimen fully, misidentification was inevitable. But regardless of the method of

404 morphologically-based identification adopted in the eradication program, our study strongly
405 suggests that the presence of several cryptic species would have thwarted successful
406 eradication or containment of the incursion. Misidentification would have seriously
407 compromised the eradication campaign through the calculation of incorrect estimates of
408 population size and location, which would have resulted in misdirection of resources (Walter
409 2003). Our work clearly shows that a combined approach using morphology and genetics is
410 required and work continues on a technique that allows genetic information to be discovered
411 whilst preserving the dermestid specimen for subsequent morphological screening and
412 identification of specie specific morphological characters. Until this is resolved, we believe
413 management of *T. variabile* will not be possible. Furthermore, analysis of *T. variabile*
414 populations suggested large-scale movement aided by human dispersal (up to approx.
415 1600km), but, more concerning, were the numerous haplotypes that were not derived from
416 main population, suggesting that at least 4 incursions occurred. Evidence supporting the
417 occurrence of multiple incursions of *T. variabile* into Australia is of high concern considering
418 that the highly invasive and destructive Khapra Beetle, *Trogoderma granarium* (Everts) is so
419 physically similar to *T. variable.*
420
421

422 ## Acknowledgements

423 This work was jointly funded by the CRC for National Plant Biosecurity and Australian
424 Government by the 2007 Science and Innovation Awards for Young People in Agriculture,
425 Fisheries and Forestry as awarded to K. M. Mikac. We would like to thank the two
426 anonymous reviewers for their comments that improved the quality of the manuscript. Thanks
427 to Jane Wright, CSIRO, for providing the specimens used in this study.

428

429

References

Armstrong, K. F. and S. L. Ball (2005). "DNA barcodes for biosecurity: invasive species identification " Phil. Trans. R. Soc. B 360: 1813–1823.

Brower, A. V. Z. (1994). "Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution." Proceedings of the National Academy of Science USA 91: 6491-6495.

Campbell, J. F. and M. A. Mullen (2004). "Distribution and Dispersal Behavior of Trogoderma variabile and Plodia interpunctella Outside a Food Processing Plant." Journal of Economic Entomology 97(4): 1455-1464.

Castalanelli, M. A., D. L. Severtson, et al. (2010). "A Rapid non-Destructive DNA extraction method for insects and other arthropods." Journal of Asia-Pacific Entomology. In press doi:10.1016/j.aspen.2010.04.003

Clement, M., D. Posada, et al. (2000). "TCS: a computer program to estimate gene genealogies." Molecular Ecology Notes 9(10): 1657-1660.

Clement, M., D. Posada, et al. (2000). "TCS: a computer program to estimate gene genealogies." Molecular Ecology 9(10): 1657-1660.

Cognato, A. I. (2006). "Standard Percent DNA Sequence Difference for Insects Does Not Predict Species Boundaries." Journal of Economic Entomology 99(4): 1037-1045.

Cook, D. (2003). Prioritising Exotic Pest Threats to Western Australia Plant Industries. D. o. A. W. Australia, Department of Agriculture Western Australia

Hartley, D. J. and H. G. Greening (1983). Review of the warehouse beetle control programme. D. o. Agriculture, New South Wales, Sydney. 43.

Jenkins, T. M., S. C. Jones, et al. (2007). "Phylogeography illuminates maternal origins of exotic Coptotermes gestroi (Isoptera: Rhinotermitidae)." Molecular Phylogenetics and Evolution 42: 612-621.

Loxdale, H. D. and G. Lushai (1998). "Molecular markers in entomology." Bulletin of Entomological Research 88: 577-600.

Lunt , D. H., D. X. Zhang , et al. (1996). "The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies." Insect Molecular Biology 5: 153-165.

Mack, R. N., D. Simberloff, et al. (2000). "Biotic invasions: causes, epidemiology, global consequences, and control." Ecological Applications 10: 689-710.

Mikac, K. M. and G. C. Clarke (2006). "Tracing the geographic origin of the cosmopolitan parthenogenetic insect pest Liposcelis bostrychophila (Psocoptera: Liposcelididae)." Bulletin of Entomological Research 96: 1-8.

467 *Mikac, K. M. and N. F. FitzSimmons (2010). "Genetic structure and dispersal patterns*
468 *of the invasive psocid Liposcelis decolor (Pearman) in Australian grain storage*
469 *systems." Bulletin of Entomological Research in press.*

470 *Nadel, R. L., B. Slippers, et al. (2009). "DNA bar-coding reveals source and patterns of*
471 *Thaumastocoris peregrinus invasions in South Africa and South America."*
472 *Biological Invasions 12.*

473 *Page, R. D. M. and E. C. Holmes (1998). Molecular Evolution A Phylogenetic Approach.*
474 *Victora, Blackwell Publishing.*

475 *Posada, D. and K. A. Crandall (1998). "Modeltest: testing the model of DNA*
476 *substitution." Bioinformatics 14(9): 817-818.*

477 *Rees, D. P., N. Starick , et al. (2003). "Current status of the warehouse beetle*
478 *Trogoderma variabile (Coleoptera: Dermestidae) as a pest of grain storage in*
479 *Australia." Stored grain in Australia.*

480 *Ronquist, F. and J. P. Huelsenbeck (2003). "MRBAYES 3: Bayesian phylogenetic*
481 *inference under mixed models." Bioinformatics 19: 1572-1574.*

482 *Schutze, M. K., P. B. Mather, et al. (2006). "Species status and population structure of*
483 *the Australian Eucalyptus pest Paropsis atomaria Olivier (Coleoptera:*
484 *Chrysomelidae)." Agricultural and Forest Entomology 8: 323-332.*

485 *Simon, C., F. Frati, et al. (1994). "Evolution, Weighting, and Phylogenetic Utility of*
486 *Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase*
487 *Chain Reaction Primers." Ann. Ent. Soc. Am 87(6): 651-701.*

488 *Suarez, A. V., D. A. Holway, et al. (2008). "Genetics and Behavior of a Colonizing*
489 *Species: The Invasive Argentine Ant." The american naturalis 172: 72-84.*

490 *Swofford, D. L. (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other*
491 *Methods). Sunderland, Massachusetts, Sinauer Associates.*

492 *Szito, A. (2007). Diagnostic Protocols for Regulated Pests Trogoderma granarium. D.*
493 *o. A. a. Food: 32.*

494 *Tamura, K., J. Dudley, et al. (2007). "MEGA4: Molecular Evolutionary Genetics*
495 *Analysis (MEGA) software version 4.0." Molecular Biology and Evolution 24:*
496 *1596-1599.*

497 *Team, R. D. C. (2007). R: A Language and Environment for Statistical Computing.*
498 *Vienna, Austria, R Foundation for Statistical Computing.*

499 *Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of*
500 *progressive multiple sequence alignment through sequence weighting,*
501 *position specific gap penalties and weight matrix choice." Nucleic Acids*
502 *Research 22(22): 4673-4680.*

503 *Tsutsui, N. D., A. V. Suarez, et al. (2003). "Genetic diversity, asymmetrical aggression,*
504 *and recognition in a widespread invasive species." Proceedings of the National*
505 *Academy of Sciences 100: 1078-1083.*

506 *Tsutsui, N. D., A. V. Suarez, et al. (2000). "Reduced Genetic Variation and the Success*
507 *of an Invasive Species." Proceedings of the National Academy of Sciences of*
508 *the United States of America 97: 5948-5953.*

509 *Tsutsui, N. D., A. V. Suarez, et al. (2001). "Relationship among native and introduced*
510 *populations of the Argentine ant (Linepithema humile) and the source of*
511 *introduced populations." Molecular Ecology 10: 2151-2161.*

512 *Walsh, P. S., D. A. Metzger, et al. (1991). " Chelex-100 as a medium for simple*
513 *extraction of DNA for PCR based typing from forensic material." BioTechniques*
514 *10: 506-513.*

515 *Walter , G. H. (2003). Insect Pest Management and Ecological Research. New York,*
516 *Cambridge University Press.*

517 *Wright, E. J. (1994). Warehouse beetle, final report. R. I. R. a. D. Corporation, CSIRO*
518 *Entomology: 196.*

519

520

521

522 ## *Appendix*

523 ## Tables

524
525 Table 1. Sample details.
526

Location	n	B	C	D	E	F	G	*T.variabile*	State	Latitude	Longitude
Ariah Park	**6**	0	1	3	0	0	0	2	NSW	-34.332	147.218
Balaklava	**6**	0	0	0	0	0	0	6	SA	-34.147	138.416
Booroowa	**2**	0	2	0	0	0	0	0	NSW	-34.482	148.751
Bowman	**1**	0	0	0	0	0	0	1	NSW	-33.15	151.292
Caracabal	**5**	0	0	3	0	0	0	2	NSW	-33.795	147.65
Carnanmah	**3**	0	0	0	0	0	0	3	WA	-29.69	115.884
Colleambally	**4**	0	0	3	0	1	0	0	NSW	-34.799	145.785
Coolaman	**8**	0	0	1	0	0	0	7	NSW	-34.848	147.185
Cowra	**7**	0	2	3	2	0	0	0	NSW	-33.819	148.658
Crystal Brook	**4**	0	0	0	0	0	0	4	SA	-33.353	138.205
Dalwallinu	**1**	0	0	0	0	0	0	1	WA	-30.249	116.669
Dimboola	**8**	0	5	0	2	0	0	1	Vic	-36.454	142.028
Garrah	**2**	0	0	0	0	0	0	2	SA	-35.483	140.3
Geelong	**3**	0	3	0	0	0	0	0	Vic	-38.149	144.357
Geraldton	**1**	0	0	0	0	0	0	1	WA	-28.78	114.613
Griffith	**7**	0	0	0	0	4	0	3	NSW	-34.282	146.035
Katanning	**1**	0	0	0	0	0	0	1	WA	-33.415	117.333
Little desert NP	**8**	0	8	0	0	0	0	0	Vic	-36.544	141.84
Lockhart	**6**	0	0	2	0	4	0	0	NSW	-35.215	146.768
Manangatang	**8**	0	1	0	0	0	0	7	Vic	-35.051	142.882
Narrabri West	**3**	0	0	0	0	0	0	3	NSW	-30.333	149.767
Nhill	**6**	0	6	0	0	0	0	0	Vic	-36.333	141.651
Pithara	**1**	0	0	0	0	0	0	1	WA	-30.389	116.667
Port Adeliade	**10**	5	4	0	0	0	0	1	SA	-34.846	138.503
Quirindi	**1**	0	0	0	0	0	0	1	NSW	-31.5	150.683
Rainbow	**4**	0	1	1	0	1	0	1	Vic	-35.9	141.993
Swan Hill	**8**	0	0	0	0	0	0	8	Vic	-35.339	143.555
Tamworth	**6**	0	0	0	0	0	0	6	NSW	-31.091	150.932
The Rock	**4**	0	0	1	0	1	0	2	NSW	-31.865	149.685
Three Springs	**1**	0	0	0	0	0	0	1	WA	-29.536	115.761
Varley	**5**	0	0	0	0	0	5	0	WA	-32.706	119.641
Wagin	**1**	0	0	0	0	0	0	1	WA	-33.187	117.208
West Wyalong	**7**	0	0	3	2	0	0	2	NSW	-33.932	147.201
Willow Tree	**2**	0	0	0	0	0	0	2	NSW	-31.649	150.735
Wyalkatchem	**1**	0	0	0	0	0	0	1	WA	-31.179	117.382
Total	**151**	**5**	**33**	**20**	**6**	**11**	**5**	**71**			

527
528

Table 2. Net divergence between the eight clades calculated on the three gene regions (18S, COI & CB), the values are nucleotide percentage differences (%).

	A.v	b	c	d	e	f	g	T.g
A.v								
b	18.44							
c	15.97	15.17						
d	15.34	15.18	8.35					
e	15.55	14.71	7.56	7.91				
f	15.62	15.76	10.12	10.44	10.06			
g	15.83	15.34	10.27	10.06	10.49	9.08		
T.g	16.89	14.85	14.17	14.73	13.93	14.5	13.72	
T.v	17.28	14.47	14.24	15.51	14.07	13.24	13.99	12.93

Table 3. Inter- and Intra-Specific Pairwise Divergence percentages for COI and Cyt B

		COI Intra				CB Inter								
		n	min	max	mean	A.v	b	c	d	e	f	g	T.g	T.v
CB Intra	n					1	3	23	14	6	7	5	1	54
	min					-	0.00	0.00	0.00	0.00	0.00	0.00	-	0.00
	max					-	0.13	0.18	0.08	0.07	0.07	0.40	-	0.00
	mean					-	0.06	0.06	0.02	0.03	0.03	0.13	-	0.00
COI Inter	A.v	1	-	-	-		26.11	22.6	21.39	20.16	21.3	21.66	20.83	22.58
	b	3	0.00	0.13	0.06	26.11		23.7	23.92	24	22.91	21.35	23.55	22.27
	c	25	0.00	0.19	0.08	23.71	24.57		15.28	13.43	18.43	23.11	13.65	23.11
	d	11	0.00	0.11	0.06	23.14	25.48	10.72		10.61	16.16	20.85	15.68	19.99
	e	2	0.00	0.10	0.05	25.89	22.66	11.34	13.15		15.65	19.93	15.23	19.93
	f	4	0.00	0.08	0.05	24.44	23.43	13.90	16.28	16.21		21.66	15.54	18.76
	g	3	0.00	0.04	0.02	24.52	23.47	14.85	15.69	16.61	12.34		20.61	20.18
	T.g	1	-	-	-	25.65	21.56	21.18	23.43	22.61	23.25	22.04		18.28
	T.v	34	0.00	0.17	0.05	24.84	21.81	21.59	25.68	23.39	23.22	23.65	18.88	

Note. COI inter- and intra-pairwise divergence (%) are below the Diagonal and Cyt B are above. Abb. A.v: *A. verbasci*, T.g: *T. granarium*, T.v: *T. variabile*

Table 4. The distribution of each species per state.

	A.verbasci	B	C	D	E	F	G	*T.variabile*
NSW*	0	0	15	95	67	91	0	44
SA	0	100	12	0	0	0	0	18
Vic	0	0	73	5	33	9	0	24
WA	100	0	0	0	0	0	100	14
Total Frequency	0.7	3.3	21.7	13.2	3.9	7.2	3.3	46.7

*NSW: new South Wales; SA: South Australia; Vic: Victoria; WA: Western Autsralia.

Table 5. Number of *T. variabile* haplotypes and frequency of each haplotype per collection site for the mtDNA markers Cytochrome Oxidase 1 (CO1) ~~LDNP: Little Desert National Park; n: sample size.~~

Site	**N***	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
Ariah Park	**2**				2						
Balaklava	**1**				1						
Carnanmah	**3**							3			
Coolaman	**2**				2						
Crystal Brook	**3**		1		1		1				
Dalwallinu	**1**							1			
Dimboola	**1**				1						
Garrah	**2**				2						
Geraldton	**1**							1			
Griffith	**1**				1						
Katanning	**1**							1			
Manangatang	**1**				1						
Narrabin West	**3**				3						
Pithara	**1**							1			
Port Adeliade	**1**				1						
Quirindi	**1**				1						
Swan Hill	**4**	1		1	1					1	
Tamworth	**5**				3				1		1
Three_springs	**1**							1			
Wagin	**1**							1			
West Wylong	**1**				1						
Willow Tree	**1**					1					
Wyalkatchem	**1**							1			
Total	**39**	1	1	1	21	1	1	10	1	1	1

* n: sample size.

554 ## Figures

555 Figure 1. Consensus Parsimony Tree combining the partial sequences of the two
556 mitochondrial gene regions COI and Cyt B and the nuclear 18S gene. The numbers above
557 the branches are maximum likelihood bootstrap values. Each letter depicts a clade. The
558 outgroup was *Anthrenus verbasci* an exotic pest to Australia. The external branches of the
559 three gene parsimony tree are labelled with the location of the specimen and the sample
560 number for that site. These labels can be cross-referenced with the sample collection data
561 found in Table 1.

562

563 Figure 2. Distribution of each species determined by parsimony and maximum likelihood
564 phylogenetic analyses. a) Spatial distribution of Dermestid spp. within Western Australia. b)
565 Spatial distribution of Dermestid spp. on the east coast of Australia. c) Sampling areas within
566 Australia. d) Parsimony and ML phylogenetic tree. e) Species wheel, each colour
567 representing a species determined by the phylogenetic trees.

568

569 Figure 3. Cytochrome Oxidase I parsimony tree. The value above the branch represents the
570 parsimony bootstrapping support, the Bayesian bootstrap values are shown below the
571 branch.

572

573 Figure 4. Spatial distribution of the ten *T. variabile* haplotypes. a) Spatial distribution of *T.*
574 *variabile* within Western Australia. b) Spatial distribution of *T. variabile* on the east coast of
575 Australia. c) Sampling areas within Australia where the presence of *T. variabile* was
576 confirmed. d) TCS network of the ten haplotypes. e) Haplotype colour key.