Mitochondrial cytochrome *c* oxidase subunit I sequence variation in New Zealand and overseas specimens of *Pieris brassicae* (Lepidoptera: Pieridae)

E. Hiszczynska-Sawicka^{1,2} and C.B. Phillips^{1,2}

¹AgResearch, Lincoln Science Centre, Private Bag 4749, Christchurch, New Zealand

²Better Border Biosecurity, www.b3nz.org

Corresponding author: Ela.Sawicka@agresearch.co.nz

Abstract Great white butterfly, *Pieris brassicae* (Lepidoptera: Pieridae), was first detected in Nelson, New Zealand, during 2010 and is the target of a current eradication programme. Knowledge of genetic variation in introduced species can be useful for various reasons, such as inferring the geographic source of the invading population and recognising when the invaded area is reinvaded by new immigrants. Genetic variation in specimens of *P. brassicae* collected at Nelson, New Zealand, was assessed by sequencing a region of mitochondrial cytochrome *c* oxidase subunit I gene (COI). COI sequences were obtained from 27 Nelson specimens and three haplotypes were found. Comparison to other sequences of *P. brassicae*, retrieved from GenBank and the Barcode of Life Data System, revealed that one New Zealand haplotype was identical to a sequence from Spain, Romania and Pakistan, and another was identical to a sequence from Germany. A third haplotype did not match any of the other available sequences.

Keywords *Pieris brassicae*, COI, barcode, eradication, reinvasion.

INTRODUCTION

Pieris brassicae (L.) (Lepidoptera: Pieridae), great white butterfly, is a cosmopolitan species inhabiting most of Europe and Asia. It has also been accidentally introduced to South Africa (Geertsema 1996) and Chile (Gardiner 1974).

Pieris brassicae has been recorded on more than 100 plant species, mainly from the families Brassicaceae and Tropaeolaceae (Feltwell 1978). It is one of the most destructive pests of cultivated brassicas (Bonnemaïson 1965; Lal & Ram 2004). Eggs are laid in batches on upper or lower surfaces of host plant leaves, and larvae feed gregariously until the plant is completely defoliated (Davies & Gilbert 1985; Le Masurier 1994). Larvae then crawl away to search for

additional hosts (Le Masurier 1994). Mature larvae disperse to pupate, usually under shelter on vertical surfaces such as fences, walls and tree trunks (Feltwell 1982).

Pieris brassicae was first reported in New Zealand in May 2010 in a Nelson home garden. Its distribution has since remained restricted mostly to residential gardens between Richmond, 12 km south of central Nelson, to Glenduan, 10 km to the north (Phillips et al. 2013). It is a risk to commercial and domestic brassica crops and to New Zealand native cresses, and is thus the target of an eradication programme being led by the Department of Conservation (Phillips et al. 2013).

Molecular biology can provide insights useful for understanding and managing biological invasions. Elucidating and comparing genetic variation in an introduced species can indicate the geographic source of the invading population (Eastoup & Guillemaud 2010; Xu et al. 2014), and signal when an area is reinvaded by new immigrants (Robertson & Gemmell 2004; Veale et al. 2013).

Mitochondrial DNA (mtDNA) is a non-recombining maternally-inherited genome and has often been used to investigate intra- and inter-specific variation (Hebert et al. 2003; Pachau et al. 2012). A 648 base pair region of mtDNA corresponding to a 5' segment of mitochondrial cytochrome oxidase subunit I (COI) was proposed as a 'DNA barcode' (Hebert et al. 2003). This paper describes variations in the COI barcode region of *P. brassicae* collected from Nelson, New Zealand, and compares them with overseas *P. brassicae* sequences available from Genbank and BOLD.

MATERIALS AND METHODS

Sample collection and DNA extraction

Larvae (n=26) and eggs (n=1) of *P. brassicae* were collected from Nelson in June and October 2013 from different locations between 8 km south of central Nelson to 4 km north. DNA was extracted using Quick-gDNATM Miniprep (Zymo Research).

PCR amplification

The barcode region of the mitochondrial COI gene was amplified using one pair of primers: LCO1490-JJ (5'-CHACWAAYCATAAAGATATYGG-3') – forward and HCO2198-JJ (5'-AWACTTCVGGRTGVCCAAARAATCA-3') – reverse (Astrin & Stüben 2008). DNA was amplified using the standard PCR protocol for i-StarTaqTM DNA polymerase (iNtRON Biotechnology, Inc.). The cycling profile was 40 cycles of 94°C denaturation (30 s), 42°C annealing (30 s), 72°C extension (1 min), with an initial denaturation of 3 min and a final extension of 5 min. All reactions were performed in a Mastercycler thermocycler (Eppendorf). PCR products were sequenced in both directions at MacroGen Inc. (Seoul, South Korea).

Analysis of *Pieris brassicae* sequences

DNA sequences were assembled and analysed using Geneious 6.1.5 software (Biomatters Ltd, New Zealand; Kears et al. 2012). Sequences were aligned using the default parameters of ClustalW within Geneious. Differences of one or more nucleotide bases between individuals were considered different haplotypes. New Zealand *P. brassicae* sequences were compared with others available in GenBank and the Barcode of Life Data System (BOLD; Ratnasingham & Hebert 2007) by neighbour-joining using the Tamura-Nei parameter. *Pieris rapae* (GenBank accession NC015895) and *Pieris napi* (AF170861) were used as outgroups.

Estimating if all *Pieris brassicae* COI haplotypes in Nelson were sampled

The haplotype frequencies were analysed using the nonparametric method of Chao et al. (2009) to estimate the total number of *P. brassicae* COI haplotypes present in Nelson. The simple calculations assume sampling with replacement, but give nearly identical results to the more complex calculations required for sampling without replacement when only a small fraction of the population is sampled (Chao & Lin 2012). They yield a simple stopping rule for sampling: sampling is complete when every haplotype is represented by at least two individuals (Chao et al. 2009).

RESULTS

The barcode fragment of the COI gene was obtained from 27 New Zealand *P. brassicae* specimens. No insertions or deletions were found, and three COI haplotypes were detected. One haplotype occurred in 16 of the 27 specimens (59%), one in seven specimens (26%) and the other in four (15%); these were arbitrarily designated as haplotypes 1, 2 and 3. Haplotypes 1 and 2 varied by one base substitution within the COI barcode region, haplotypes 1 and 3 by five, and haplotypes 2 and 3 by four. None of the substitutions changed the amino acid sequence. The method of Chao et al. (2009) indicated that the presence of additional *P. brassicae* COI haplotypes in Nelson was extremely unlikely.

Data were obtained from GenBank and BOLD for specimens collected in Spain (1 sequence), Germany (3), Switzerland (2), Romania (7), Russia (2), Kyrgyzstan (2), Pakistan (3), India (2) and Japan (1). Accession numbers for the New Zealand and overseas specimens are shown in a neighbour joining tree (Figure 1), which is based on the 638 bp barcode region. The two most dissimilar sequences were from Japan and Romania, and differed by nine base substitutions. Compared to the outgroups, *P. rapae* and *P. napi*, the New Zealand sequences clustered with the *P. brassicae* sequences obtained from Genbank and BOLD (Figure 1). New Zealand haplotype 2 was identical to a sequence found in Spain, Romania and Pakistan, and New Zealand haplotype 3 was identical to a sequence from Germany. New Zealand haplotype 1 differed slightly from all others.

DISCUSSION

The New Zealand sequences matched *P. brassicae* sequences from overseas, thus corroborating previous morphological identification of *P. brassicae* in Nelson. Three COI haplotypes were found amongst the 27 *P. brassicae* sequenced. Any

future discovery of additional COI haplotypes in Nelson would strongly suggest new immigrants had invaded Nelson after sampling for the present study during June and October 2013. However, it would be possible for new immigrants possessing one of three COI haplotypes already identified to invade Nelson and remain undetectable as new arrivals via analysis of the COI barcode region. Thus, studies of additional *P. brassicae* genetic markers are desirable.

The presence of three *P. brassicae* COI haplotypes in New Zealand indicates the founding population in Nelson comprised a minimum of three females plus an unknown number of males.

Variation between *P. brassicae* COI haplotypes available from five European and four Asian countries showed a weak geographical pattern (Figure 1). A set of four similar haplotypes that included New Zealand haplotype 3 occurred across Europe from Germany to Russia, while an 'Asian' haplotype occurred in Pakistan, India and Kyrgyzstan. However, the samples that included New Zealand haplotype 2 showed no geographical pattern, occurring in Pakistan,

Figure 1 Neighbour-joining tree for the COI barcode region of *Pieris brassicae*. Non-New Zealand sequences were obtained from GenBank and BOLD. Scale bar shows uncorrected genetic distance.

Romania and Spain (Figure 1). *Pieris brassicae* adults fly strongly and sometimes migrate hundreds of kilometres (Spieth & Cordes 2012). In addition, humans clearly also contribute to *P. brassicae* dispersal (Gardiner 1974; Geertsema 1996; Phillips et al. 2013). Thus, there is probably considerable gene flow across the Northern Hemisphere range of *P. brassicae*. The New Zealand COI haplotypes were similar or identical to those sampled from disparate Northern Hemisphere locations, and by themselves are insufficient for deducing the foreign origin of Nelson's *P. brassicae* population. Unfortunately, few *P. brassicae* sequences are available from GenBank for gene regions other than COI, so deducing the Northern Hemisphere source of *P. brassicae* would probably require obtaining specimens from the Northern Hemisphere for analysis in New Zealand.

Analysing genetic relationships between *P. brassicae* sampled from its invaded range in South Africa, Chile and New Zealand would indicate how many *P. brassicae* populations have independently crossed the equator and become established in the Southern Hemisphere. We have already obtained *P. brassicae* specimens from Chile, and are endeavouring to source specimens from South Africa.

Defining the spatial and temporal occurrence of different *P. brassicae* genotypes in Nelson could provide information useful to the eradication programme in several ways. First, *P. brassicae* sometimes reoccurs in outlying areas of Nelson after a control operation, and it would be helpful to know if this was due to either control failure or reinvasion from central Nelson. Second, when treating areas where *P. brassicae* is present in very low densities, different batches of eggs and broods of larvae are sometimes found on nearby properties at similar times. It would be helpful to know if these are the progeny of one or several females because evidence of several females would support more widespread and prolonged search effort. Third, identifying egg batches laid by a single female would provide information about the individual female's spatial and temporal patterns of movement and oviposition, which could help predict the best areas to search. Finally, estimating trends in *P. brassicae* population size currently relies on calculating

detections per unit of search effort and, for verification, additional independent estimation methods would be valuable. Small declining populations are more susceptible to genetic drift (Fitzpatrick et al. 2012), so all *P. brassicae* genotypes will have extinction probabilities that increase with declining population size. If the *P. brassicae* population is being diminished by the eradication programme, and no reinvasion is occurring, then the haplotypes should progressively disappear from the population.

Current knowledge of *P. brassicae* genetic variation in Nelson is insufficient to resolve the population beyond three broad groups defined by COI haplotype. However, even the coarsely probabilistic answers that could be provided from this basis should be of some value to the eradication. Ongoing research will develop quicker and cheaper methods of identifying the COI haplotypes, such as melt curve analysis (Winder et al. 2011), and investigate further *P. brassicae* geographic populations and additional genetic markers.

ACKNOWLEDGEMENTS

We thank Richard Toft (Entecol, Nelson) for providing us *P. brassicae* specimens from Nelson and Department of Conservation staff, particularly Kerry Brown, Bruce Vander Lee and Jo Rees, for their assistance. Nicky Richards, Sean Marshall (both AgResearch) and an anonymous reviewer provided helpful suggestions. The research was funded both by the Ministry for Primary Industries and by AgResearch via the Better Border Biosecurity research collaboration.

REFERENCES

- Astrin JJ, Stüben PE 2008. Phylogeny in cryptic weevils: molecules, morphology and new genera of western Palaearctic Cryptorhynchinae (Coleoptera:Curculionidae). *Invertebrate Systematics* 22: 503–522.
- Bonnemaison L 1965. Insect pests of crucifers and their control. *Annual Review of Entomology* 10: 233–256.
- Chao A, Colwell RK, Lin C-W, Gotelli NJ 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* 90: 1125–1133.

- Chao A, Lin C-W 2012. Nonparametric lower bounds for species richness and shared species richness under sampling without replacement. *Biometrics* 68: 912–921.
- Davies CR, Gilbert N 1985. A comparative study of the egg-laying behaviour and larval development of *Pieris rapae* L. and *P. brassicae* L. on the same host plants. *Oecologia* 67: 278–281.
- Estoup A, Guillemaud T 2010. Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology* 19: 4113–4130.
- Feltwell J 1978. The depredations of the large white butterfly. (*Pieris brassicae*) (Pieridae). *The Journal of Research on the Lepidoptera* 17: 218–225.
- Feltwell J 1982. Large white butterfly: The biology, biochemistry, and physiology of *Pieris brassicae*. Kluwer Academic Publishers Group, Dordrecht, Netherlands.
- Fitzpatrick BM, Fordyce JA, Niemiller ML, Reynolds RG 2012. What can DNA tell us about biological invasions? *Biological Invasions* 14: 245–253.
- Gardiner BOC 1974. *Pieris brassicae* (L.) established in Chile: another palearctic pest crosses the Atlantic (Pieridae). *Journal of the Lepidopterists' Society* 28: 269–277.
- Geertsema H 1996. The large cabbage white, *Pieris brassicae*, an exotic butterfly of potential threat to cabbage growers in the Western Cape, South Africa. *Journal of the Southern African Society of Horticultural Sciences* 6: 31–34.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London - Series B: Biological Sciences* 270: 313–321.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meinthes P, Drummond A 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Lal MN, Ram B 2004. Cabbage butterfly, *Pieris brassicae* L. - an upcoming menace for Brassica oilseed crops in Northern India. *Cruciferae Newsletter* 25: 83.
- Le Masurier A 1994. Costs and benefits of egg clustering in *Pieris brassicae*. *Journal of Animal Ecology* 63: 677–685.
- Pachau L, Vanlaruati C, Kumar SN 2012. Morphological versus molecular characterization of three similar Pierid species of butterflies. *International Journal of Pharma and Bio Sciences* 3: 1091–1102.
- Phillips CB, Brown K, Green C, Walker G, Broome K, Toft R, Vander Lee B, King M 2013. Great white butterfly interim report. Prepared for Ministry for Primary Industries External Technical Advisory Group, December 2013. Department of Conservation, Nelson. 32 p.
- Ratnasingham S, Hebert PDN 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* 7: 355–364 (DOI: 10.1111/j.1471-8286.2006.01678.x).
- Robertson BC, Gemmell NJ 2004. Defining eradication units to control invasive pests. *Journal of Applied Ecology* 41: 1042–1048.
- Spieth HR, Cordes R 2012. Geographic comparison of seasonal migration events of the large white butterfly, *Pieris brassicae*. *Ecological Entomology* 37: 439–445.
- Veale AJ, Edge K-A, McMurtrie P, Fewster RM, Clout MN, Gleeson DM 2013. Using genetic techniques to quantify reinvasion, survival and in situ breeding rates during control operations. *Molecular Ecology* 22: 5071–5083.
- Winder LM, Phillips CB, Richards NK, Ochoa-Corona F, Hardwick S, Vink CJ, Goldson SL 2011. Evaluation of DNA melting analysis as a tool for species identification. *Methods in Ecology and Evolution* 2: 312–320.
- Xu J, Fonseca DM, Hamilton GC, Hoelmer KA, Nielsen AL 2014. Tracing the origin of US brown marmorated stink bugs, *Halyomorpha halys*. *Biological Invasions* 16: 153–166.