# Pest risk assessment of light brown apple moth, *Epiphyas postvittana* (Lepidoptera: Tortricidae) using climate models and fitness-related genetic variation

---

**A thesis**

**submitted in partial fulfillment**

**of the requirements for the Degree of Master of Science**


**By**
**Shu-qi He**


---


**Supervisor:**

- **Associate Professor Susan P. Worner**
  **Project Leader Ecological Informatics for Biosecurity**
  **Bio-protection Research Centre**
  **P.O. Box 84**
  **Lincoln University 7647**
  **Christchurch, New Zealand**

**Associate supervisor:**

- **Dr. Karen F. Armstrong**
  **Project Leader Molecular Diagnostics**
  **Bio-protection Research Centre**
  **P.O. Box 84**
  **Lincoln University 7647**
  **Christchurch, New Zealand**



**Bio-protection Research Center**
**Lincoln University**
**March, 2010**

**Abstract & Key words**

**Abstract of a thesis submitted in partial fulfillment of the requirements for the Degree of M.S.**

# Pest risk assessment of light brown apple moth, *Epiphyas postvittana* (Lepidoptera: Tortricidae) using climate models and fitness-related genetic variation

**By**

**Shu-qi He**

**Abstract**: Light brown apple moth (LBAM), *Epiphyas postvittana*, is a leafroller pest native to Australia and currently has limited distribution around the world. The moth is an extremely polyphagous species which can feed on more than 250 host plants, including many important economic crops such as fruit, vegetable and ornamental species. *Epiphas postvittana* is also well-known for its polymorphism with respect to its morphology. Its physiology and biology have been well studied compared with many other pest species. For the past 10 years, however, this moth has been gradually expanding its territory around the world. This has resulted in increasing calls for targeted risk assessment so that appropriate quarantine measurements can be put in place to prevent its entry and spread into new regions. Three important aspects related to risk of establishment by *E. postvittana* were investigated in this study, 1) *E. postvittana* potential global distribution based on climate, 2) the genetic variation of a target gene for its potential to assess population fitness and invasibility, and 3) the potential use of barcoding as molecular tool for *E. postvittana* identification at the border.

By comparing the climatic conditions of its native (Australia) and long-established (New Zealand) ranges to the rest of the world using CLIMEX, it is

suggested that *E. postvittana* has potential to establish in countries in Central and South America, southern Africa, west Europe and South-east Asia. However, the predicted global distribution of *E. postvittana* using a range of other types of species distribution models suggested that there are additional climatically suitable areas around the world where this species could potentially survive and establish.

Morphological identification of *E. postvittana* has been problematic which increased the risk of it esacaping detection at the border of countries that wish to regulate this pest. In this study, we sequenced the COI gene from 26 samples of *E. postvittana* from four populations in New Zealand. We found that the intraspecies variation of *E. postvittana* is less than 3%, while interspecies variation between *E. postvittana* and other tortricid species available in the barcode of life database (BOLD) system is much greater than 3%. This result confirms that using barcodes for identification of *E. postvittana* for biosecurity purposes is practical. The COI gene sequences have been submitted to GenBank as reference sequences.

Genetic analysis of the phosphoglucose-isomerase (*Pgi*) gene in *E. postvittana* was investigated based on its association with various characteristics of fitness in other Lepidoptera. Using novel PCR primers developed in this study, a comparison of 957 bp of the *Pgi* coding region amongst four *E. postvittana* populations revealed 70 segregating sites including 61 synonymous and nine non-synonymous sites. Introns of *Pgi* gene also show a great variation in length among populations and between alleles within the same locus. The significant variation of *Pgi* gene in *E. postvittana* populations indicates the *Pgi* gene as a useful target gene to assess fitness factors associated with invasibility of *E. postvittana*.

Modeling species distributions and pest identification are both key components in pest risk assessment. The study of fitness-associated genetic variation is currently a novel approach additional approach to risk assessment of invasive species but has much potential in this area. Our study provides basic but important information for further assessment of the establishment capacity of this species in new habitats.

Further research in these areas will provide the knowledge required to make science-based decisions in biosecurity.

Key words: biosecurity, quarantine pest, invasive species, CLIMEX, ecological model, modeling, distribution prediction, machine learning, SVM, openModeller, support-vector machine, phosphoglucose isomerase, *Pgi* gene, barcode, COI

# Acknowledgements

My utmost gratitude goes to my supervisors, Associate Professor Sue Worner and Dr Karen Armstrong, for their support, tenacity and invaluable suggestions and guidance throughout the process and for making it possible for me to accomplish my study.

I would like to extend my special thanks to Rob Cruickshank, Andrew Holyoake, Hayley Ridgeway, Anastasija Chomic, Takayoshi Ikeda and many other staff and postgraduates in the Bio-protection Research Centre and Ecology Department in Lincoln University, who helped me with my research and study.

Sincere thanks to the New Zealand Government Development Scholarship (NZAID) for financial support, which allowed for me to have such a wonderful experience at Lincoln University. Many thanks to Angela Parry for her support during my study here.

Finally, I would like to express my gratitude to my best friend, PhD student Emily Fountain, and my parents, for their encouragement, help, and understanding throughout the years. Without all your support, I would be without the final accomplishment of my master's degree.

Thank you all!

Shu-qi He, 1054752
Christchurch, March 2010

# Contents

## Contents Pages

# List of Tables

# List of Figures

# Chapter 1 General Introduction

## Light brown apple moth, *Epiphyas postvittana* (Lepidoptera: Tortricidae) – how serious a pest is it and could it be?

### 1.1 Invasive Alien Species and Light brown apple moth (LBAM)

With the rapid increase in world trade and tourism, exotic species are more likely to arrive in a non-native habitat. Additionally, global climate change increases the potential establishment and spread capability of these exotic species and many become invasive alien species (IAS) (Dukes & Mooney 1999) that can cause considerable economic and environmental damage. The environmental damage and economic loss caused by IAS can be enormous. For example, the environmental damage and control cost of a single species, the red imported fire ant in the United State, has been estimated as between 400 and 600 million US dollars per year respectively (Pimentel *et al*. 2005). In New Zealand, non-indigenous invasive invertebrate pests have been estimated to cause about 800 million to 2 billion NZ dollars of economic impact on pasture and forage production (Goldson *et al*. 2005). Invasive species may also have negative effects on an ecosystem by endangering native species or driving them to extinction through competition, either directly or indirectly through the food chain (Howarth 2000). It is estimated that about 42% of threatened or endangered species in the United States are at risk due to the introduction of exotic invasive species (Pimentel *et al*. 2005).

Williamson (1996) estimated that only about 1% of exotic species that arrive at the border are likely to become IAS. However, because many thousands of interceptions are made at the border of any country this small percentage may still pose a significant risk. Risk assessments are, therefore, critical to provide policy makers with scientific evidence so that rapid and effective actions can be made to reduce the potential negative impacts of IAS. A comprehensive understanding of the biology and ecology of the potential target IAS is the most basic and essential step for its risk assessment.

One such example of an IAS that has had significant economic impact in its invaded area is the light brown apple moth (LBAM), *Epiphyas postvittana* (Walker) (Lepidoptera; Tortricidae). This species belongs to the group commonly known as

leafrollers, due to their leafrolling behaviour. LBAM is native to Australia where it was first recorded in 1863, and named *Teras postvittana* (Danthanarayana 1975a). Bradley (1956) re-assigned the species to a new genus *Austrotortirx*, which is still cited in some early references. Common (1961) classified the moth along with another 31 Australian related species into genus *Epiphyas*, including *E. pulla*, *E. xylodes* and *E. liadelpha*, which also occur in orchards. All Australian native *Epiphyas sp.*, have been referred to as light brown apple moth (LBAM) due to their morphological similararity and their similar damage on horticultural host plants (Geier & Springett 1976). In this thesis, LBAM refers only to *E. postvittana*.

Despite being an economically important pest in its native and established ranges, LBAM had a very limited distribution worldwide but was considered as an "actionable quarantine pest" by most LBAM-free countries such as Mexico, Canada and Japan (Varela *et al*. 2008). During the previous 10 years, the moth has continuously expanded its territory in Europe and more recently to North America. The most recent confirmed establishment was in 2007 in California USA, and it was also detected in Sweden in 2008 (Suckling & Brockerhoff 2010). The new invasion of LBAM into California drew the attention of scientists with respect to its potential risk globally. Despite this renewed attention and many years of study, information about LBAM remains deficient in some countries and may result in either an over- or under-estimate of its quarantine status. For example, quarantine pest lists of some countries includes Mexico as one of the countries in which LBAM is distributed, but to my knowledge there has been no reference of the moth having established in Mexico. While the quarantine status of this moth is still debated, further risk assessment with respect to its potential global distribution and environmental adaptation capability is urgently needed. A pest risk assessment underpinned by a good understanding of the diagnosis, biology and physiology of *E. postvittana* would be particularly useful.

## 1.2 Features relevant to LBAM as a quarantine species

### 1.2.1 Current geographic distribution

*Epiphyas postvittana* is native to south-eastern Australia and is distributed naturally in Queensland, New South Wales, Victoria and South Australia including Tasmania. It was introduced to Western Australia in 1968 (Danthanarayana 1975a;

Geier & Springett 1976; Geier & Briese 1981). So far no LBAM have been recorded in the Northern Territory of Australia. LBAM was introduced to New Zealand in 1891 (Evans 1952), and is currently widely spread over the country (Shaw *et al*. 1994). Its distribution in the British Isles has been restricted to the south-west in Devon and Cornwall where it was first introduced in 1953 (Baker 1968). However in the 1990s, LBAM spread rapidly and in 2000s, it became a common moth throughout England, Wales, and Southeast Ireland (Porter 2001, Fountain & Cross 2007). In 1956, the moth was introduced into Hawaii and New Caledonia (Danthanarayana 1975a). In March 2007, the US Department of Agriculture's Animal and Plant Health Inspection service (APHIS) confirmed that LBAM was present in California. In a matter of nine months 15,594 male LBAM were caught in pheromone traps (Varela *et al*. 2008), suggesting that this species was well established there. By the 2$^{nd}$ of December 2009, 18 counties in California reported LBAM infestations, nine of which applied quarantine measurements. The most affected areas were located in the middle of the west-coast areas along the coast of California (APHIS 2009; USAD 2009a). An eradication program was launched in response to the first detection of the moth, but its effectiveness has been limited (USDA 2007a). Since then, LBAM was detected in Sweden in 2008 (Suckling & Brockerhoff 2010).

A quarantine pest, defined by IPPC (2009), is an organism that: 1) has potential to cause economic or environmental damage in an area, 2) is currently not present or not widely distributed in the area or 3) is under official control. Despite a relatively limited current distribution, LBAM meets all conditions of being a "quarantine pest" in a risk area (more in below). Considering its continuous range expansion, especially with increasing international horticulture and vegetable product trading, the potential risk of the moth should not be underestimated.

### 1.2.2 Host plant range

LBAM is a highly polyphagous species. Initially, it was known to feed on native shrubs such as genus *Acacias* (Danthanarayana 1975a). Cultivation activity and the natural adaptability of the moth have potentially allowed it to adjust to new environmental conditions and to feed on additional exotic species brought into Australia and New Zealand (Geier & Briese 1981). Danthanarayana (1975a) recorded 73 host plant species belonging to 28 families in Australia, including both native and

exotic plant species. Geier & Briese (1981) updated its host range to 123 species from 55 families (in Australia). Wearing *et al*. (1991) estimated the moth could feed on more than 250 plant species over 55 families, which is close to the number of 265 recorded host plants in New Zealand noted by Thomas (1989). The newest host records for LBAM stated the moth has been found on more than 500 plant species over 363 genera and 121 families (Suckling & Brockerhoff 2010). According to the United States Department of Agriculture (USDA), it is estimated that LBAM could potentially cause damage to more than 2000 plant species in USA, including the most economically important fruit and ornamental crops and young conifer forests (APHIS Factsheet 2008). More scientific evidence is required to justify this number, however. The most common host plants for LBAM include fruit and vegetables, such as apple, pear, grape, citrus, kiwifruit, cucumber, capsicum, cut flowers and ornamental species, such as chrysanthemum, daisy, lily and orchid, and young exotic forests such as conifers (Thomas 1974; Danthanarayana *et al*. 1995).

This wide range of host plants that are both native and exotic species indicates, 1) the potential capability of the moth to feed on unrecorded host plant species is relatively high, indicating a high invasive risk, and 2) given the climate within Australia and New Zealand is quite variable resulting in the patchy spatial and temporal distribution of host plant species the occurrence of LBAM is also patchy (Danthanarayana *et al*. 1995; Geier & Briese 1981). Both factors reflect the capability of LBAM to adjust to variable conditions. Also, Peacock & Worner (2008) examined factors such as host range, developmental temperature and propagule pressure, which might potentially affect the invasibility of insects using two groups: one group as well-established invasive insects which are continuously intercepted at the border and the other group that are not-established but are often intercepted. They suggested that given suitable climatic conditions, species with a wide host plant range were more likely to establish in a new habitat. Clearly, LBAM with such a wide host range and potential for host expansion might pose a relatively high risk in any current LBAM-free areas or countries.

### 1.2.3 Identification and biosecurity

Morphological identification of *E. postvittana* is very difficult. This is largely due to the scale pattern on the adult wings (the main identification feature of Lepidoptera)

being extremely variable within the species, making it difficult to distinguish from related leafroller species. Ideally adult male genitalia need be dissected to accurately identify to species (Dugdale *et al*. 2005). However, it is the immature life stages that are mainly intercepted at border and these also look very similar between species. This morphological ambiguity is compounded by the fact that, although *E. postvittana* is usually the dominant leafroller species in crop fields in Australia and New Zealand, it co-occurs with indigenous leafroller species such as *Ctenopseustis obliquana* (brownheaded leafroller), *Planotortrix excessana* (greenheaded leafroller) and *P. notophaea* (blacklegged leafroller) in New Zealand and with *E. pulla* and *E. xylodes* in Australia (Brockerhoff *et al*. 2002; Geier & Springett 1976; Suckling *et al*. 1998). Moreover, *E. postvittana* can hybridize with *E. pulla* and *E. xylodes* in the field, although the offspring may not survive more than three generations (Geier & Springett 1976). This co-occurrence and potential hybridization make the morphological identification of LBAM intercepted at the border more complicated.

Potential IAS intercepted at the border need to be accurately identified as quickly as possible to minimize the risk. It became well-established in California and wide-spread in UK causing serious economic loss, because the moth was misidentified as some other moth species rather than LBAM (Fountain & Cross 2007; Suckling & Brockerhoff 2010). Other data worth to mentioning is that, in the 20 years from 1984 to 2003, *Epiphyas spp.* were intercepted in the US only 55 times (Venette *et al* 2003), while in a single year from July 2000 to June 2001 *E. postvittana* was intercepted in Japan 63 times (Takahashi 2002). Both Japan and USA are key export markets for New Zealand and Australia fruit and vegetables. The interception rate difference, on one hand, might be due to the relative volume of trade; however, it may also reflect to some extent the ability to detect the moth correctly.

Given the difficulty identifying LBAM morphologically, molecular methods have been developed. Most recently DNA barcoding using the cytochrome *c* oxidase subunit I (COI) gene has been employed. This method is widely used for rapid and accurate species diagnosis, not only for adults but also for immature stages of IAS such as eggs, larvae and pupa intercepted at the border (Armstrong & Ball 2005; Ball and Armstrong 2006). The variation of the COI gene has been found to be less than 2.3% within LBAM populations (Barr unpublished data) compared to more than 7%

between *E. postvittana* and other native moth species in California (Barr 2007). Such results indicate using barcodes for LBAM identification at border is practical. Large samples of immature LBAM were also discriminated from other moth species in California using ITS2 (internal transcribed spacer 2) locus for less expensive diagnosis of LBAM (Barr *et al*. 2009). Integrating the use of molecular tools together with morphological characters will greatly improve the rapid, cost-effective and correct identification of LBAM.

The extreme polymorphism of adult LBAM might be an adaptation to environmental impacts such as temperature variation as well as variety, availability and quality of food plants. These factors are likely to have significant effects on population dynamics of LBAM. However, currently, no research has been done that relates highly polymorphic phenomena and what it may indicate with respect to the moth's biology and ecology.

### 1.2.4 Economic damage

Economic loss by a potential invasive species includes direct damage such as the crop-yield loss, and indirect economic consequence such as the potential export market loss, extensive control cost and post-harvest treatment for export. A precise and accurate estimate of the crop-yield loss is the most essential element in risk assessment for a crop pest because all other economic losses may be the consequence of significant crop-yield loss caused by the species. Also, potential environmental damages should be considered, such as pesticide application and competition with native leafroller species. If the moth poses no economic or environmental threat, then it would be worthless to spend millions of dollars to exclude or eradicate.

The larvae of LBAM cause direct damage by feeding on buds, leaves, shoots and fruits, particularly the spring and summer generations, and indirectly by transmitting fungal disease. Larvae mainly feed on the surface of the fruit, but sometimes they can bore into the fruit (Geier & Briese 1981). Crop loss from larval feeding can range from 5 to 20% in Australia, sometimes exceeding 30%, and is about 12~70% in New Zealand, and commonly 50% without insecticide application (Wearing *et al*. 1991). Indirect damage of the fungal disease *Botrytis cinerea* transmission by larvae can cause an estimated 16.5% of mean weight loss of fruit products (Bailey *et al*. 1997).

When the larval populations are high, the yield loss might be up to NZ $2,000 per ha in New Zealand (Bailey *et al*. 1996). However, Lo & Murrell (2000) give very different estimates where they found that maximum direct damage was 12% in December in Hawke's Bay, New Zealand, and another 12% of indirect damage was in March, which corresponds to approximately NZ$60-360 per ha. Although in the USA, it has been estimated that LBAM could cause an annual crop and control loss of US$160 to $640 million in 15 infested counties (APHIS 2009), no actual crop yield loss by LBAM has been reported so far.

LBAM is considered a major pest in south-eastern Australia and New Zealand and it is ranked the second after the codling moth, *Cydia pomonella*, in Australia (Wearing *et al*. 1991). In the UK, *E. postivittana* caused serious damage to cherry fruit in 2005, and it is now considered that serious damage on cherry in 2003 and 2004 was caused by LBAM but that it was misidentified as some other tortricid species (Fountain & Cross 2007). In California, the crop loss has not been accurately determined per se, with the main economic loss being due to the strict management cost associated with confining its distribution and eradication (USDA 2007b). However, whether the moth should be treated as an "eradicable pest" or "manageable minor pest" is still under debate because evidence on its economic damage is currently insufficient (USDA 2009b). Apparently, further investigation is urgent and worthwhile on this aspect of LBAM.

### 1.2.5 Life-cycle

LBAM occurs in the field continuously throughout the year in suitable climatic conditions. Generally, the moth goes through three generations in southern Australia and central New Zealand, such as Nelson and Hawkes Bay (Danthanarayana 1975a; Wearing *et al*. 1991; Hortnet 2000). In the first generation, named the summer generation, eggs are laid in December or early January, larvae develop quickly during March and April, and adult moths emerge during April. The second or autumn-winter generation, develops from April to September, and larvae overwinter on weeds or the fallen leaves of host plants without diapause, but with a developmental rate much slower in winter season than in summer (Geier & Briese 1981). For the spring generation, eggs are laid in October and adults emerge in December. Four overlapping generations occur in northern New Zealand such as Auckland and Waikato, and two

generations occur in more southern regions such as Canterbury, Otago and Southland because of the colder temperatures (Thomas 1974; Hortnet 2000). However, populations are found more abundant in Canterbury than Nelson and Hawkes Bay, and are the least abundant in central Otago (Suckling *et al*. 1998).

Eggs are deposited in masses on the upper surfaces of leaves or the surface of the fruit of the host plants where the mean size of an egg mass varies significantly ranging from 4-96 eggs normally (Dantharayana 1975a; Powell & Common 1985; Wearing *et al*. 1991). Fecundity of female adults varies significantly depending on the temperature and the host plant, but the maximum potential fecundity was found to be 1492 eggs per female, which indicates that the moth has the potential capability to build up large populations in a short time with lower initial propagule pressure (Danthanarayana 1983; Worner 2003). However, mortality of eggs and first instar larvae was considerably high (ranging from 23~92% for eggs and 27~90% for first instar larvae) due to a high predation rate, which constrains population abundance (Danthanarayana 1983).

Larvae hatch from eggs after one or two weeks and disperse immediately to the underside of a leaf settling near the midrib or a vein, to construct a nest (leaf roll) after the first moult (Danthanarayana 1975a). Larvae generally pass through six instars, but sometimes four, five or seven (Geier & Briese 1981). Eggs are mature in the pupal stage and adults are therefore able to oviposit immediately after emergence (Gu & Danthanarayana 1990b). Adults fed with only water or honey water, especially for three days post emergence, have significantly increased fecundity and longevity (Gu & Danthanarayana 1990a).

### 1.2.6 Dispersal capacity

Flight activity contributes to the dispersal and migration capacity of an invasive alien pest, and strong flight capability allows a species to spread to new habitats and survive even if environmental conditions are unfavorable (Tauber *et al*. 1986). Therefore, consideration of flight capability is an essential and important element in risk assessment of a potential invasive species (IPPC 2004).

However, distant dispersal of LBAM relies on human activity such as international transportation of host plant commodities. The moth itself is generally

considered to be a sedentary flier (Gu & Danthanarayana 1992a). Study by Suckling *et al*. (1994) found the mean dispersal distance of adult *E. postvittana* measured by pheromone traps was less than 54.8 m, while the majority of the moths reached less than 100 m but up to 600 m were recorded. Additionally, they found variations between females and males. For example, females on average could fly further than males, but males flew over longer periods than females (Suckling *et al*. 1994; Gu & Danthanarayana 1990b). Flight speed for females was approximately $2.37m^{-s}$ and $2.16m^{-s}$ for males (Danthanarayana 1976b). Flight activity appears to be associated with light intensity, occurring mainly during the night with the major peak 2-3 hours after sunset (at about 2100h) and a very small peak 3-4 hours before sunrise (Danthanarayana 1976a). The maximum length of flight occurred between 3~5 days after emergence in females and 5~9 days in males, during which a prolonged duration of flight activity is apparent (flying more hours during the peak flight activity period) (Gu & Danthanarayana 1990b). Mating behavior also peaked during the peak flight period, suggesting that flight activities are primarily related to reproductive behaviour (Gu & Danthanarayana 1990b).

Flight activity of LBAM is affected by environmental conditions. The lower flight temperature threshold has been found to be about 8~11$^{o}$C, with an upper threshold ranging from 20 to 28$^{o}$C. However, these thresholds can vary among different generations in a year (Danthanarayana 1976b). Danthanarayana (1976b) has shown that humidity has no significant influence on flight activity, neither did flight activity occur on rainy days nor when wind speed is greater than 2.8m/s. However, a study by Danthanarayana & Gu (1992) showed flight activity peaked at relative humidity around 60%. Wing-loading (ratio of body length to wing length) has been used as a variable to indicate the dispersal ability of LBAM. When experiencing adverse environmental conditions (e.g. hot, dry conditions and/or scarce food) wing-loading was found to be low (smaller body length with larger wings) and resulted in more efficient flight (Danthanarayana 1976c; Gu & Danthanarayana 1992b). Wing-loading is also affected by larval population density. At times when the larval population was high, a so-called “low-density larval crowding” happened in the field where 2-3 larvae share one leafroll nest (Danthanarayana *et al*. 1982). Wing-loading significantly declined when larval crowding occurred, which also negatively correlated with declines in other life-history traits such as adult body weight and

fecundity (Danthanarayana *et al*. 1982; Gu & Danthanarayana 1990b). The flexibility of the flight capability in variable environmental conditions is therefore apparently an adaptation to survive in unfavorable seasons. Gu & Danthanarayana (1990a, 1992b) also showed that water content and food quality (young leaves) of larval diet and water supply for adult moths after three days of emergence could significantly affect flight activities of female moths.

Suckling *et al* (1994) in a study of dispersal ability of *E. postvittana* estimated that the moths tended to fly upwind, since the majority of trapped moths were found upwind of the trap. Nevertheless, Danthanarayana (1976a, b) suggested that the diurnal and lunar flight periodicity of LBAM indicates that the moth flies upwards towards light and migrates downwind, which has also been demonstrated in other small insects such as aphids and fruit flies (Taylor 1974). Adult moths have been caught at a range of heights in 5, 10 and 20 meters high suction traps (Gu & Danthanarayana 1990b), indicating that wind-assisted migration is possible. Most research on LBAM flight activity has used either pheromone traps or 48-cm-high suction traps, but such traps might not be able to detect the wind-assisted long distance migration trend. Further study on the dispersal (migration) ability of this species would be worthwhile to understand the potential spread rate of LBAM in a new habitat.

### 1.3 Population dynamics and adaptability to new environments

#### 1.3.1 External factors

The population abundance and distribution of LBAM is known to fluctuate significantly, and is considered to be affected by several environmental factors as are many herbivores such as climatic conditions especially temperature and precipitation, food plant variety, quality, quantity and availability, natural enemies, and larval density. Among these, temperature and host plant are known to have significant and predictable effects on population dynamics of *E. postvittana*. Such variables are essential for predicting the potential distribution of LBAM in a new habitat.

##### 1.3.1.1 Climatic conditions

Insects are ectothermic animals and their development is affected by climatic

factors, particularly by temperature (Atkinson 1994). Here I discuss particular climatic factors such as temperature and humidity that significantly and directly affect the population abundance, distribution and biological behaviors of LBAM.

A warm and wet climate favours moth development, while hot and dry weather reduces its population abundance. For example, in two hot and dry years of 1973 and 1976 in Victoria Australia, populations of LBAM remained at a low number such that the relative abundance of the species was only about 10% of that in 1974 and 1975 (Danthanarayana 1983). Seasonal climatic variation also has a strong effect on population development within any one year, with spring and summer generations more abundant and causing the main damage to crops compared with the autumn-winter generation (Danthanarayana 1975b). Temperature is the main climatic factor affecting population abundance by regulating the fitness of LBAM, such as larval and pupal development time, body weight of different development stages, adult longevity, fecundity (eggs laid), fertility (eggs hatched out) and oviposition period (Danthanarayana 1975b; Danthanarayana 1976c; Gu and Danthanarayana 1992b; Danthanarayana *et al*.1995).

According to laboratory studies, the developmental zero of LBAM is 7-7.5$^{o}$C and upper temperature threshold between 30.7-33$^{o}$C, the favourable temperature range is between 15~25$^{o}$C with the optimum temperature for development around 20$^{o}$C (Danthanarayana 1975a; Geier & Briese 1981; Danthanarayana *et al*. 1995). The shortest generation time (in terms of degree-days) is around 620.5 with an average of 673.6 (Danthanarayana 1975a).

Research on the effect of moisture and relative humidity on the development of LBAM is rare. Based on the current geographic distribution of LBAM, the species is able to survive in areas with an annual rainfall of no less than 500 mm (Geier & Briese 1981; Wearing *et al*. 1991; Danthanarayana *et al*. 1995). Studies by Danthanarayana & Gu (1992) & Danthanarayana *et al*. (1995) found that the population rate of increase was optimal at a mean annual relative humidity of 60~71%, annual rainfall of about 600~800mm and relative higher altitude (500-1000m). Water supply to an adult moth after 3 days of its emergence could significantly affect the longevity, body weight maintenance and fecundity. This partly explains high population abundance during wet conditions at the time of emergence (Gu &

Danthanarayana 1990a).

LBAM prefers relatively cool and wet conditions. As a result, high temperature stress and low air moisture levels appear to limit the distribution of the moth from large areas of Australia. However, rainfall in New Zealand appears very favourable for development of LBAM throughout most of the country and might be the reason that LBAM is more abundant and injurious in New Zealand than in Australia. It is only in southern New Zealand that this species experiences really cold temperatures that suppress development. Altitude was considered as an environmental factor affecting population dynamics of LBAM (Danthanarayana *et al*. 1995), which may be due to correlation with temperature, but very few studies have been done so far to verify this assumption.

Climatic conditions may also influence LBAM populations indirectly by affecting the abundance and distribution of its host plants. Climatic conditions in Australia and New Zealand are highly heterogeneous, which affect the distribution and availability of host plant species spatially and temporally (Danthanarayana 1983; Danthanarayana and Gu 2000).

#### 1.3.1.2 Host plants

Variety, quality, quantity and availability (succession) of larval host plants significantly affect the fitness of LBAM, such as fecundity, adult body size, and pupal weight, although only few host species have been tested despite the wide host range of the moth. For example, Danthanarayana (1975a) tested four host species (dock, plantain, capeweed and apple). Larvae fed on dock (*Rumex*) had shorter duration life stage, heavier pupae and significantly higher fecundity of female adults compared with those fed on the other species. Dock also caused a significant difference in weight of adult moths (Gu & Danthanarayana 1992b), developmental periods, optimal developmental temperature (Tomkins *et al*. 1989; Danthanarayana *et al*. 1995), fecundity (Danthanarayana 1975b), flight ability (Danthanarayana 1976c), larval resistance to chemical azinphosmethyl (Robertson. *et al*. 1990) and parasitism rate (Suckling *et al*. 2001). Plant nutrients are likely to explain the important role host plants play on LBAM fitness, but this has not been studied in detail.

Danthanarayana (1983) tested the preference of larval LBAM on 13 host species

under laboratory conditions. Dock is the most preferred host, and also succession (availability) of the host plants significantly affected population abundance at a particular location. Individuals developed better on young leaves than senescent leaves (Geier & Briese 1980a; Gu & Danthanarayana 1992b; Mo *et al*. 2006a). However, LBAM develops continuously throughout the year, and interestingly, larvae feeding on defoliated or even freeze-dried leaves during the winter did not have a significantly different pre-imaginal development compared with those fed fresh leaves (Geier & Briese 1980a; Tomkins *et al*. 1989).

#### 1.3.1.3 Natural enemies

There are numerous natural enemies and entomopathogens of *E. postvittana*, including parasitoids, predators, bacteria and viruses, which constantly regulate the populations of LBAM in Australia and New Zealand (Geier & Briese 1981; Thomas 1989; Wearing *et al*. 1991). Among these, predators and parasitoids are the most important. For example, predators (mainly spiders and earwigs) have been shown to contribute up to 92.1% of egg mortality and 90.7% of first instar mortality respectively (Danthanarayana 1983). A study by Suckling *et al*. (1987) showed more than 50% of the autumn-winter generation larvae of LBAM were killed by the parasitoid *Apanteles tasmanicus*. Unfortunately, frequent pesticide application causes the exclusion of the natural enemies from cultivated fields. It has been suggested that the low temperature preference shown by LBAM might be another adaptation to escape from many of the more generalist natural enemies (Geier & Briese 1981). However, research on biological interaction between LBAM and its natural enemies is rare.

#### 1.3.1.4 Population density

When the population density is high, low-density larval crowding occurs such that more than one larva (normally 2-3) are found feeding in the same nest. This behavior produces smaller individuals with more dispersive ability, but is unlikely to affect the population size of LBAM moth (Danthanarayana *et al*. 1982).

### 1.3.2 Intrinsic genetic variation associated with life-history traits of LBAM

Considerable intra-specific variation in LBAM morphology and flight activity

has previously been discussed and LBAM is also noted for its considerable physiological and demographic variation. Moreover, given the same rearing conditions, (e.g. the same temperature, humidity, and food plants), there has been shown to be variation amongst geographic populations in their sex ratio, resistance to a nuclear polyhedrosis virus (NPV), and important life history traits such as development time, adult body weight, adult lifespan, age at first reproduction, total fecundity and flight capacity (e.g. Geier & Briese 1980b; Gu & Danthanarayana 2000a). This therefore suggests that there is a genetic component to this variation, and to population dynamics as it relates to pest status, irrespective of external influences.

There has been little research published on the genetics of variation in LBAM. A quantitative genetic analysis on flight capacity, that was found to be significantly correlated with life-history traits (Gu & Danthanarayana 1992a), shows evidence of heritability in flight capacity in populations from Canberra and Melbourne (Gu & Danthanarayana 1992c). Further investigation of six life-history traits for these two populations also indicated significant heritability (Gu & Danthanarayana 2000b). Furthermore, LBAM populations from both Australia and New Zealand tested for five life-history traits and flight capacity, showed that phenotypic plasticity existed in LBAM with all measured traits were significantly different among populations which might be due to environmental conditions such as temperature and all these traits were genetically correlated (Gu & Danthanarayana 2000a). These genetic correlations suggested there may be a genetic interaction between environmental variables and fitness traits, i.e. environmental conditions may influence fitness traits through single or more common genes (Via 1983).

Other variation implicated having a genetic basis includes sex ratio and resistance to nuclear polyhedrosis virus (NPV). Sex ratio in LBAM field populations is normally 1:1. However, Geier & Briese (1977) revealed a heritable “Q-condition” whereby “carrier” females produce all-female or predominantly-female progenies, and proportion of the Q-condition moths was different among different geographic populations in Australia. This may have resulted in the notable sex ratio bias toward females in mainland of Australia but toward males in Tasmania (Geier & Briese 1978). It also suggested that the Q-condition occurs more frequently in long-established populations compared with new-entry ones (Geier & Briese 1978). NPV is a

ubiquitous virus in Australian where LBAM is widely distributed. Larvae of laboratory reared “Can strain” and Tasmania field population were found to be genetically more resistant to NPV than larvae of other field populations from various localities in Australia (Geier & Briese 1979).

The variation among natural populations is likely to be an evolutionary adaptation of LBAM to cope with local environmental heterogeneity (Gu & Danthanarayana 2000a). This together with its wide host plant range and phenotypic plasticity could facilitate expansion to unpredictable habitats, and the potential establishment of populations in new regions of the world (Booy *et al*. 2000).

### 1.4 Conclusion

My particular research on LBAM was suggested because of the recent establishment of this species in California, USA in 2007 and its apparently expanding geographic range in Europe. International plant health authorities are concerned about the potential risk the moth imposes on countries that are currently free of infestation of this species, and which require further phytosanitary regulations to prevent its further spread.

LBAM is considered a quarantine pest in many LBAM-free countries because of the potentially significant economic loss via direct damage and indirect effects on market loss and management cost. Its wide host plant range, multiple generations per year with high potential fecundity, compounded by identification difficulties at the border increase the risk of LBAM as an invasive species capable of serious impact. Moreover, LBAM has evolved with human modification of natural habitats and has shown a ready ability to adapt to novel habitats (Geier and Briese 1981). However, how serious an invasive pest it could be globally has yet to be determined. The natural dispersal ability of LBAM is considered to be relatively weak, but it still can easily arrive at a new habitat assisted by human activity, particularly trade of its plant hosts. After arrival, its ability to establish is then decided by host availability and local climatic conditions, especially temperature. Fitness of the founding population, which is genetically determined, may also be a key factor.

Danthanarayana *et al*. stated in 1995 that the possibility of LBAM establishing in North America is “remote”, but 12 years later, in 2007, the moth actually established

in California. A deeper understanding of the biological and physiological characters of the moth is therefore required to predict and prevent its further distribution around the world. The research here therefore will investigate the potential global distribution of LBAM based on climate and evaluate intraspecific variation of the *Pgi* gene, which has been shown to be related to fitness and dispersal in other insects such as butterflies and beetles (e.g. Dahlohff & Rank 2000; Hanski & Saccheri 2006; Wheat *et al.* 2006; Karl *et al.* 2009).

## 1.5 Aims, objective and Hypotheses

### 1.5.1 Aims

The general aim of this research was to generate and analyze novel information that could contribute towards a more rigorous risk assessment of LBAM with respect to its potential global spread. In particular this will include:

- Use of models based on climate to produce maps that indicate areas with likely high and low risk of LBAM establishment, particularly so that phytosanitary measures can be better formed.
- Generation of DNA barcode (COI gene) data to enable accurate identification of the moth to be used in this study as well as to add population data to that available for identification of various life stages intercepted at country border.
- Investigation of variation in the *Pgi* gene as a potential genetic component, and possible predictor, of establishment and spread in relation to environmental conditions.

Both modeling and molecular information provided here could be used by any country to develop a pest risk analysis for phytosanitary policy development.

### 1.5.2 Hypotheses

To assess the utility of these approaches the hypotheses to be explored are:

- Models based on climate are effective tools to indicate the potential distribution of LBAM.
- The *Pgi* gene, as linked to fitness traits and dispersal capability in other species, shows variation between LBAM populations and therefore may be a useful target gene to assess fitness factors associated with invasibility of LBAM.

### 1.5.3 Objectives

Specific **objectives** of the study were to: 1) investigate and predict the potential global distribution of LBAM using CLIMEX; 2) compare the result from CLIMEX with other models to improve the precision of the prediction on LBAM potential distribution; 3) generate DNA barcodes as a rapid and effective tool for identification of LBAM at borders; 4) design a novel PCR priming system to access the *Pgi* gene of LBAM and 5) to investigate variation amongst geographic LBAM populations using that system as a possible explanation and predictor of the fitness and dispersal ability of insects in particular environment.

# Chapter 2 Modeling the potential global distribution of light brown apple moth, *Epiphyas postvittana* (Lepidoptera: Tortricidae), using CLIMEX

## 2.1 Introduction

Ever-increasing global trade and tourism increases the potential pathways for exotic species to a new habitat (Levine & D'antonio 2003). Climate change associated with global warming can extend the range of suitable habitats for species to establish, spread and become invasive. Such species are often referred to as invasive alien species (IAS) (Dukes & Mooney 1999). IAS can cause significant economic loss and ecosystem damage (e.g. Pimentel *et al*. 2005; Goldson *et al*. 2005), therefore, rapid and effective actions are required to minimize potential negative effects.

Despite perceived threats, Williamson (1996) estimated IAS account for only around 1% of total exotic species that arrive at borders. Hence, official control measures are not aimed at all exotic species, but only at those species with potential to seriously impact the endangered area. Such species, particularly insects and plant pathogens, are defined as quarantine pests (IPPC 2009).

To determine which exotic species should be considered a quarantine pest and to decide the sort of measures that should be taken to prevent their entry and establishment to a new habitat, pest risk assessment is often carried out. Pest risk assessment is essential to provide the scientific evidence that will indicate the most effective measures to prevent the IAS entering, establishing and/or spreading in the endangered area. In standard pest risk assessment, it is critical to evaluate the susceptibility of a given area to a certain pest according to the environmental conditions in that area (IPPC 2004). Some determination of habitat suitability helps to make predictions about whether the pest will be able to establish and spread in that area.

Environmental conditions for consideration in pest risk assessment include climate variables, host plant availability, presence of natural enemies as well as any abiotic or biotic variable considered to be relevant to a species establishment. Climate variables, such as temperature, rainfall, humidity, flooding, storms and seasonal

changes, can have an important, and often, dominant influence on the behavior, abundance and distribution of insects which are ectotherms (Worner 1994; Musolin 2007). Therefore, climate parameters are widely used to predict the distribution and abundance of invasive species with the idea of comparing their fundamental and realized niches (Sutherst *et al*. 2007).

To provide policy makers with reliable evidence to support rapid quarantine decisions scientific information must be readily available. However, specific scientific research on the biology of a quarantine pest is costly and very time consuming, especially to conduct the required experiments that measure insect response to temperature and other biotic variables. As a result, models are widely used as an alternative approach. If applied appropriately, they give useful, rapid predictions of the potential distribution of the target species.

CLIMEX is one of a range of modeling systems that may provide insights into the climatic factors that limit the geographical distribution of a species in different parts of its range. It has been broadly used for biological control agent selection (e.g. Senaratne *et al*. 2006; Robertson *et al*. 2008; Poutsma *et al*. 2008) as well as determining the potential successful establishment and geographical distribution of exotic species (e.g. Worner 1988; Vera *et al*. 2002; Kriticos *et al*. 2003;Sutherst & Maywald 2005; Stephens *et al*. 2007).

Climatic parameters and the climate-matching function of CLIMEX enable the risks of an exotic species to be assessed by directly comparing the climatic condition of a given location with any number of other locations without knowing the full distribution of a species (Sutherst & Maywald 1991; Sutherst *et al.* 2007). The "Growth Index" (GI) describes the potential growth of a population during the favorable seasons, and four Stress Indices (Cold, Hot, Wet and Dry Stresses) describe the survival possibility of the population through unfavorable seasons. The combination of GI and Stress Indices, "Ecoclimatic Index" (EI), gives an overall assessment of the potential risks of pest establishment in the tested locations (Sutherst *et al*. 2007).

Distribution and population abundance of *E. postvittana* is significantly affected by environmental factors especially temperature (see Chapter 1). Although much

research has been done on its biological and physiological features, no formal global risk assessment has been done for this species. The lack of a formal global risk assessment results in uncertainty about the pest status of the moth and therefore it is hard to take cost-effective protective measures to control the moth or prevent its further spread and establishment. Improper control actions might result in high control cost and also prevent normal international trade by contravening important phytosanitary agreements. An eradication campaign has been carried out in California USA, which has cost billions of dollars but has not yet been shown to work. Scientists and farmers have begun to question if the eradication is necessary and even if it is effective. In New Zealand where the moth is considered a serious pest because its impact on fruit production and its quarantine status in other parts of the world (with trade implications), pesticides are relied on heavily to keep populations at/or under an extremely low economic threshold level (Suckling & Brockerhoff 1999).

Pest risk assessment for *E. postvittana* is urgent given that it seems to be on the move internationally especially during the last 20 years (Suckling & Brockerhoff 2010). The first question that comes to mind is under what climatic conditions the moth can survive, and where those conditions are in the world indicating where it can potentially establish and spread. The aim of the research here is to investigate the potential global distribution of light brown apple moth using CLIMEX version 3 to provide a global overview of relative high risk and low risk areas for *E. postvittana* establishment.

## 2. 2 Methods

### 2.2.1 Distribution data

A distribution map of *E. postvittana* is available from CABI which was last revised in 1992, and does not indicate sites recording the most recent spread of the moth, such as UK, California USA and Sweden (Appendix 1a). Sixty-nine locations recording the presence of the moth were collected from publications, and coordinates of these locations were estimated using Google Earth (Appendix 2). Another 76 coordinates were downloaded directly from Global Biodiversity Information Faculty (GBIF, http://www.gbif.org), and a final of 30 locations remains after combining overlapping coordinates as one location. Parameter calibration is mainly using the

distribution of *E. postvittana* according to CABI distribution map (1992), and presence location points were used to determine how well the parameters fit after calibration.

Parameters for CLIMEX were calibrated according to the climatic requirements of the moth and its current distribution in Australia, the native range of the moth, and its distribution in New Zealand, where the moth has been established for more than 100 years. Climate conditions in Australia and New Zealand are heterogeneous, which is sufficient for being used in CLIMEX model (Sutherst 2003).

### 2.2.2 Climate dataset

CLIMEX contains a meteorological database of ~3000 selected locations worldwide as well as interpolated climatic grids provided by the global change community through the CRU in Norwich UK for non-commercial use. New Zealand climate data, such as the mean daily/monthly maximum/minimum temperature and mean monthly rainfall for 1971 to 2000 period, is freely available from the website of National institute of Water & Atmospheric Research (NIWA, http://www.niwa.co.nz). Climate data of Australia is also available from Bureau of Meteorology website (http://www.bom.gov.au/climate). Such data helped to calibrate some indices particularly the stress indices by attempting to fit CLIMEX to *E. postvittana* distributed area and to calibrate parameters more objectively.

### 2.2.3 Parameters

Parameters were calibrated without considering human activities such as irrigation. However, such factors will be discussed later in the discussion. All the parameters after calibration are listed in Appendix 3.

#### 2.2.3.1 Template

In CLIMEX, there are different templates for species with different climate requirements. The range of parameters that can be calibrated from each template varies according to the templates chosen. When setting up for a new species, choosing an appropriate template to start the model is important. The distribution of *E. postvittana* mainly ranges between latitudes 11~47$^{o}$S, and longitudes 129~167$^{o}$E in

Australia and New Zealand, which belongs to a temperate zone (Danthanarayana 1975a). Extensive research has shown that the species favors relatively low temperatures and wet conditions (e.g. Danthanarayana 1976c). As a result, we chose the temperate-template in CLIMEX to start our analysis. The initial parameter values in the temperate-template "predicted" the distribution of *E. postvittana* in Australia and New Zealand and were reasonably close to the laboratory data (see below). All parameters and their abbreviation are listed in Appendix 3.

### 2.2.3.2 Temperature index

The temperature index in CLIMEX consists of four parameters: DV0 represents the limiting low temperature (lowest temperature threshold above which the species can develop) and DV3 represents the limiting high temperature (upper temperature threshold above which development of the species ceases) (e.g. Figure 2.1). Beyond DV0 and DV3, no population growth takes place, in other words, the species will not be able to survive if populations are exposed continuously to temperatures outside of this range. DV1 and DV2 define the optimal temperature range for the species, that is, population growth decreases when temperature is below DV1 (lower optimal temperature) or above DV2 (upper optimal temperature). PDD represents the number of the degree-days above DV0 required to complete an entire generation.

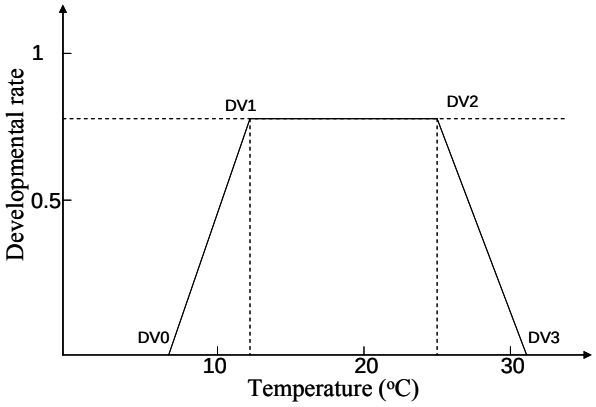



Figure 2.1. Temperature parameters in CLIMEX. DV0 and DV3 represent the temperatures from which the insect starts and stop to develop. DV1 to DV2 represent the optimal developmental temperature range.

Laboratory data has shown that the developmental zero for *E. postvittana* is 7.1-7.5°C (DV0), the upper temperature threshold is between 30.7~33°C (DV3), the optimum temperature is around 20-25°C with the favorable temperature range between 15~25°C (DV1 & DV2), and population growth rate is highest at lower mean annual temperature of 13-14°C (Danthanarayana 1975a; Danthanarayana *et al.* 1995).

The total degree-days (PDD) for the shortest generation time is about 620.5 and the mean generation time (time to 50% egg-laying or cohort generation) is 673.6 (Danthanarayana 1975a). Geier & Briese (1981) found the developmental temperature of CAN strain of *E. postvittana* to be 7$^{o}$C, which is 0.5$^{o}$C lower than that determined by Danthanarayana (1975a).

Using the distribution map of Australia and New Zealand (CABI 1992), and the initial parameters of the temperate template, the core distribution of *E. postvittana* in the model was found to be around south-eastern Australia and the North Island of New Zealand (Figure 2.2). Lowering DV0 made regions such as Otago in the South Island of New Zealand favorable for *E. postvittana*, which in reality is slightly too cold for optimal population growth of this species (Suckling *et al*. 1998). Increasing DV3 moved the margin of *E. postvittana* distribution to the very North of Australia including the Northern Territory where currently *E. postvittana* is not established (Suckling & Brockerhoff 2010). The lowest developmental temperature (DV0) was set at 7$^{o}$C and 30$^{o}$C as the highest developmental temperature (DV3).

DV1 and DV2 were set to 13$^{o}$C and 23$^{o}$C respectively, slightly lower than the laboratory results (15-25$^{o}$C, Danthanarayana 1975a) (see Appendix 3). The reasons for setting a lower DV1 and DV2 are: 1) laboratory data shows that DV1 and DV2 of LBAM development are 15-25$^{o}$C respectively, and the population growth rate of LBAM is highest at mean annual temperature of 13-14$^{o}$C (Danthanarayana 1975a; Danthanarayana *et al.* 1995). However, LBAM populations are more abundant in Canterbury than in Nelson and Auckland, of which the annual mean temperature is about 11~12$^{o}$C (NIWA 2010), which may in turn indicate that optimal developmental temperature range of LBAM may be lower than the laboratory data 2) Laboratory reared LBAM is living under constant artificial temperature and more subjective to death after released to natural environmental conditions (Kye person. comm.). This may be partially due to the better cold resistant of LBAM natural populations.

The visual fit of the resulting CLIMEX distribution map corresponded to the actual distribution map published by CABI (1992). However, the EI value of most locations in South Island of New Zealand, while suitable for *E. postvittana* development, for example, Christchurch, were still a bit low because of high cold stress, which has to be further calibrated in stress indices.

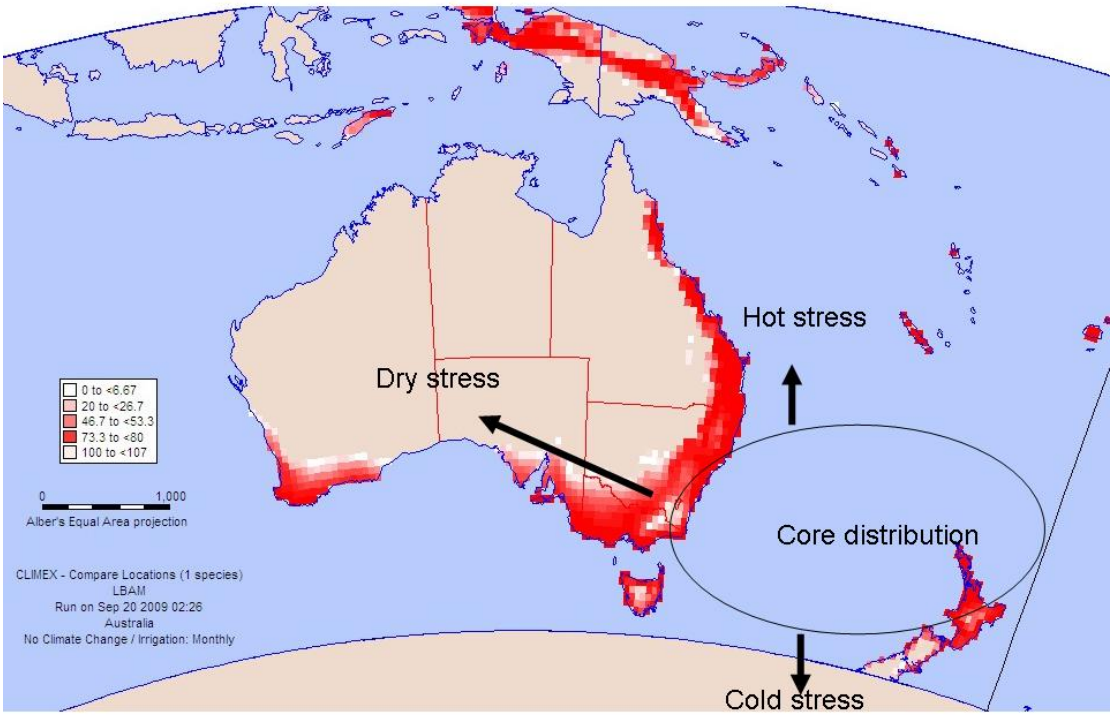



Figure 2.2: The modeled core distribution of *E. postvittana* in Australia and New Zealand. Most coastal regions of the South Island of New Zealand and Southeastern Australia had EI values of 100. Hot stress increases gradually to the north of the core, cold stress increases to the south, and dry stress increases to the west of the core distribution along direction of each arrow.

A degree-day of 673.6 was used to define the generation time and hence the number of generations for most of the known areas where *E. postvittana* occurs. For example, the moth has two generations in Southern New Zealand especially in Otago and Southland, three generations in Nelson and Hawkes Bay and in southern Australia, and four generations in Auckland (Wearing *et al*. 1991; Suckling *et al*. 1998; Hortnet 2000).

### 2.2.3.3 Moisture index

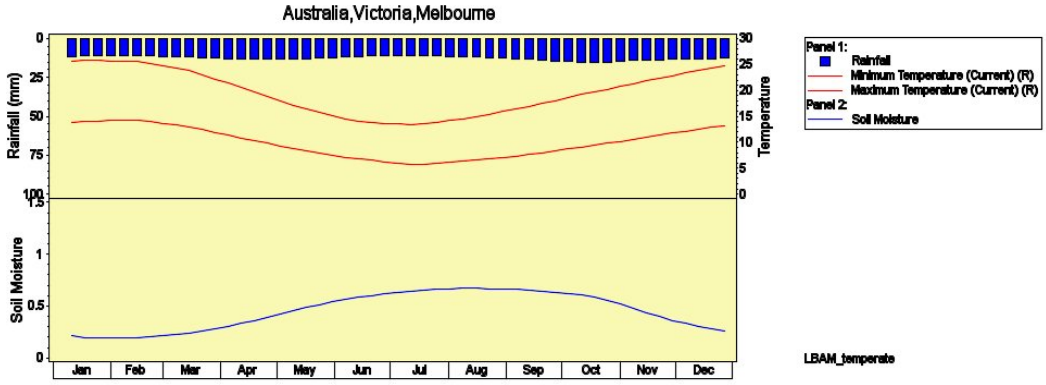
The moisture index also consists of four parameters. SM0 and SM3 represent the lower and upper limiting soil moisture thresholds respectively. No population growth occurs beyond this range. SM1 and SM2 define the optimal soil moisture range in which population growth is the fastest.

Evidence about the relationship between moisture and development of the moth is rare. It was estimated that areas with an annual rainfall above 500 mm (SM0) (range from 500-800 mm) would be suitable for the development of the species (Geier

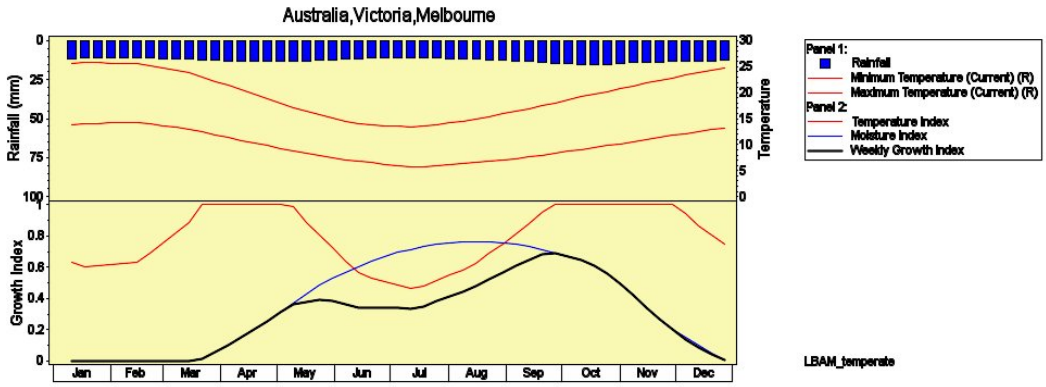
& Springett 1976), which means the average rainfall per week would be about 9.6 mm or so (between 0.3~0.4 of soil moisture value in CLIMEX). Relative humidity could affect the fecundity and flight activities of *E. postvittana* but not significantly. Adult moths feeding on water after 3 days of emergence laid more eggs than those with no water supply. As a result, the optimum humidity was estimated about 60-70% (Danthanarayana *et al*. 1995).

Soil moisture varies with the seasonal changes of rainfall and is related to the weekly rainfall in a particular area (Figure 2.3). Locations presented in Appendix 2a were checked with respect to the moisture index. Most areas in New Zealand have a annual rainfall over 500 mm. Areas such as Auckland, Nelson, and Christchurch where populations of the moth are most abundant and were where the soil moisture index is greater than 0.2. Soil moisture is similar in areas such as Invercargill, Dunedin and Auckland, and *E. postvittana* populations appear to change mainly according to the temperature. The initial values of the soil moisture parameters in CLIMEX temperate-template were adopted because they are close to most of the areas when the corresponding values of soil moisture and growth were examined (Appendix 3).

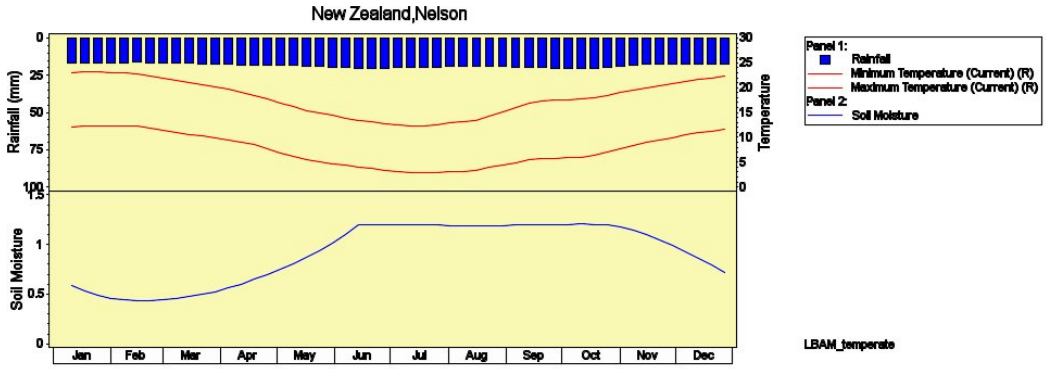
a.



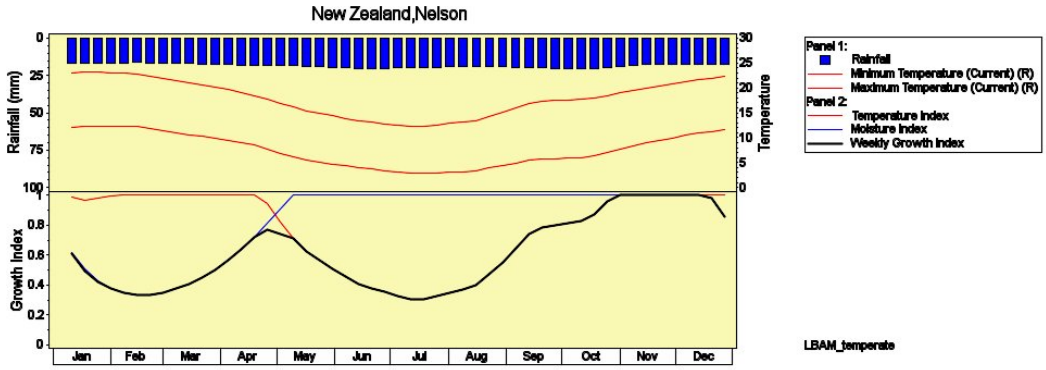

b.




c.




d




Figures 2.3 a & c show that rainfall can be converted to soil moisture. They also show seasonal changes in the soil moisture values so that the relationship between the species requirement and soil moisture can be visualized. Figures 2.3 b & d show that the growth of *E. postvittana* populations varies according to the seasonal changes of temperature and soil moisture. The growth index combines the temperature index and moisture index, and varies more in response to temperature than to moisture in this case.

#### 2.2.3.4 Diapause index and light index

Light brown apple moth does not have a diapause and overwinters at a slow developmental rate (Geier & Briese 1981). No data is available relating the development of *E. postvittana* to light, only flight activities of the moth are known to be affected by the intensity of sunlight (Danthanarayana 1976a, b). As a result, both CLIMEX diapause index and light index were not included in the current study.

#### 2.2.3.5 Stress indices

Both the temperature index and moisture index indicate the developmental capability of a species in favorable environmental conditions. On the contrary, the stress indices demonstrate the species' response to unfavorable environmental conditions. In other words the conditions are too hot, too cold, too wet and/or too dry for the species to survive. All the stress parameters, therefore, are assumed to be outside the ranges of DV0 and DV3 (lower and upper developmental temperature limits), and of SM0 and SM3 (soil moisture limits). Accordingly, there are four different stresses: cold stress, heat stress, dry stress, and wet stress. The final value of each index is listed in Appendix 3.

##### 2.2.3.5.1 Cold stress

Cold stress starts to accumulate when temperature is too low for a species to survive. There are three ways to set parameter values for cold stress in CLIMEX. First, which we adopted in current research, is the assumption that a species cannot survive if exposed to an extremely cold temperature and given a Cold Stress Temperature Threshold (TTCS) ($^{o}$C) the cold stress accumulates at a given rate (THCS) below this temperature. Second, a species may die if the daily thermal accumulation is too low to maintain metabolism and cold stress accumulates at a given rate (DHCS) if the weekly number of degree days does not reach the Cold Stress Degree-day Threshold (DTCS). Third, if a species lives in locations that can buffer it from extreme minima and maxima, then cold stress will accumulate at a given rate (THCSA) when the weekly average temperature drops below the Cold Stress Average Temperature Threshold (TTCSA). In the current model, the first alternative was chosen represent the effect of cold stress on *E. postvittana* to estimate its potential distribution. TTCS was set lower than DV0.

While *E. postvittana* is exposed to relatively low temperatures, the species overwinters by reducing its developmental rate. However, Danthanarayana (1975a) found that no eggs or larvae could survive at 5$^{o}$C. THCS were tested at a range from -1 (maximum accumulation rate of cold stress) to 0 (no cold stress). For locations such as Auckland, where the average temperature is relatively high, *E. postvittana* would not be affected by cold stress. Only localities with a relatively low temperature, such as locations in South Island of New Zealand, were *E. postvittana* distribution affected by cold stress. Cold stress in turn affected the eco-climatic index (EI) which reflects in general, the fitness of the moth in a given location.

When there is no cold stress, THCS should be set to 0. New Zealand is in general suitable for the development of *E. postvittana* with cold stress most likely to occur in the Central Otago area (Suckling *et al*. 1998). When THCS was set to 0, there were 13 locations in New Zealand that were all suitable for *E. postvittana* development where in 12 out of the 13 locations the EI values were above 24. When THCS was set to -1, only one location was suitable. The number of suitable locations increased as the value of THCS was gradually increased from -1 to 0. When THCS was set to -0.00075, 12 locations were suitable and finally 13 locations when the THCS was set to -0.0007. The EI value differs slightly as THCS is increased from -0.0007 to 0 in Christchurch, Timaru, and Milford Sound, but relatively more in other locations such as Gore, Invercargill and Dunedin. *Epiphyas postvittana* is widespread in New Zealand. Thus, it was assumed that the moth could survive most areas in New Zealand although the exact presence/absence data was not available. However it is known that populations are significantly less abundant in central Otago than in Canterbury and Nelson (Suckling *et al*. 1998). The EI variation of all locations in South Island of New Zealand was compared as THCS was varied. THCS was finally set as -0.0005. The template value of THCS is -0.0001, the value for *E. postvittana* was set lower because the moth is affected by high temperatures significantly more than low temperatures (Danthanarayana *et al*. 1995).

#### 2.2.3.5.2 Heat stress

Heat stress can be set in CLIMEX in two ways. First, the lethal temperature assumes that a species fails to survive when exposed to excessively high temperature. When temperature is higher than the heat stress temperature threshold (TTHS), heat

stress starts to accumulate at the stress accumulation rate (THHS). Second, the degree-day method assumes that a species is not able to persist because the daily heat load is too high for essential physiological processes. Heat stress begins to accumulate when the temperature greater than DV3 or the weekly degree-day threshold (DTHS) is multiplied by the stress accumulation rate (DHHS). In current research, the first option was chosen to represent the effect of heat stress on *E. postvittana*. TTHS is required be set higher than DV3.

Development of *E. postvittana* is known to be significantly affected by high temperature. It was shown that no eggs could develop fully at 31.3$^{o}$C and development of larvae and pupa ceased at temperatures between 32~33$^{o}$C (Danthanarayana 1975a, Danthanarayana *et al*. 1995). *Epiphyas postvittana* is not able to tolerate high temperatures as much as low temperatures (Geier & Briese 1981), as a result, heat stress was set to accumulate faster than the cold stress.

The mean monthly maximum temperature in New Zealand seldom achieves 31$^{o}$C (NIWA), which makes it possible for *E. postvittana* to maintain a widespread distribution in New Zealand. Therefore, the heat stress was calibrated mainly according to its distribution in northern Australia (see Figure2.2, the core distribution). TTHS was set as 31ºC, one degree centigrade higher than DV3 (set as 30 ºC, see above). If the TTHS was set higher than 31ºC, the distribution would cover the Northern Territory where currently no *E. postvittana* have been detected yet. THHS was tested within its range of between 0 (no heat stress) and 1 (maximum accumulation rate of heat stress) and finally set as 0.01 according to the visually calibrated distribution of the moth in the Northern Territory Australia.

**2.2.3.5.3 Dry stress**

Moisture can cause stress for a species when it is too dry or too wet. Dry stress for some species accumulates when it experiences consecutively low soil moisture levels below the dry stress threshold (SMDS). That stress will accumulate at a given rate or dry stress rate (HDS). SMDS is required to be set lower than SM0.

Population abundance of *E. postvittana* significantly decreased in hot and dry weather (Danthanarayana 1983). It is rare in New Zealand that the annual rainfall in any location is below 500 mm which was considered as the threshold for *E.*

*postvittana* development (NIWA 2009; Geier & Briese 1981; Danthanarayana *et al*. 1995). Dry stress was mainly tested and calibrated with the *E. postvittana* distribution in Australia (Figure 2.2). SMDS was set to 0.2 which is 0.05 less than SM0, the same value as in the template. HDS was tested over its range from -1 (when dry stress accumulates extremely fast) to 0 (no dry stress). At -1 no *E. postvittana* were distributed in Queensland and Western Australia, and when set to 0, its distribution only increased to several locations in Queensland, one additional location in Western Australia, and several in South Australia around the presence/absence margin area of the distribution map. At -0.015 the distribution approximated the actual distribution map. The dry stress and eco-climatic index varied slightly for those locations with extremely low rainfall. Most dry and hot areas are excluded by the temperature and moisture indices anyway. When the resulting map was compared to the actual distribution map and the variation of dry stress and eco-climatic indices were examined, HDS was set as -0.01 (-0.005 in template).

#### 2.2.3.5.4 Wet stress

Wet stress accumulates when the soil moisture exceeds the wet stress threshold (SMWS) consecutively with a wet stress rate (HWS). The SMWS parameter is required to be set lower than SM3.

No data is available for wet stress, and only research on flight activities shows that no flight activities happen in rainy weather (Danthanarayana 1976a, b). Moreover, it is hard to detect the effect of wet stress from the examination of the original distribution of *E. postvittana* in Australia and New Zealand, because weekly average rainfall does not cause the soil moisture to increase more than 1.5 in most of the areas in these two countries. As a result, the value for SMWS and HWS were set at 2.5 and 0.002 respectively, the same as the values in the template (Appendix 3).

### 2.2.4 Potential global distribution

The potential global distribution of *E. postvittana* was predicted using the calibrated parameters (Appendix 3). Locations in light brown colour without showing EI values were not suitable for the species to survive. The EI values are ranging from 0 to 100, which represent the possibilities of a species' establishment in certain locations. A location where the EI value close to 0 is a location not favourable for a

species' establishment, and EI value of 100 is only possible under constant and ideal conditions. In practical, locations with EI values less than 10 are considered as marginal locations for a species that population fluctuations are likely to occur and locations with EI value greater than 20 are considered potentially be able to build up substantial populations (Sutherst 2003). Localities with an EI value greater than 30 (range from 30 to 100) were considered as very suitable for *E. postvittana* to survive and establish (Suthurst *et al*. 2007).

Also, all locations with an EI value greater than 30 within Australia, New Zealand, and UK were plotted in a map using coordinates extracted from CLIMEX, to check if they correspond to (within) the current distribution as indicated in CABI (1992). We selected EI values greater than 30 because it gives more confident of the potential establishment of the moth in a location. The 69 presence locations extracted from references and the 76 GBIF presence locations were used to test the fit of the parameters after calibration by determining if their EI values were greater than 10.

## 2.3 Results

### 2.3.1 Distribution of *E. postvittana* in Australia and New Zealand

The distribution of *E. postvittana* in Australia and New Zealand generated by CLIMEX after all parameters have been calibrated was restricted mainly to Southern and Eastern Australia, Western Australia and New Zealand (Figure 2.4), which was similar to the actual distribution of the moth as shown in the CABI distribution map (Appendix 1).In most areas in Australia *E. postvittana* is limited by heat and dry stress, which is not the case in New Zealand (Appendix 4c, d). On the other hand, cold stress in Australia is not a limiting factor as much as in the South Island of New Zealand (Appendix 4b). *Epiphyas postvittana* is currently established in New Caledonia with a high predicted EI value ranging from 56 to 79 (Figure 2.4).

We plotted locations with EI values greater than 30, using coordinates extracted from CLIMEX and found there are about three points located right in the middle of Australia, in which *E. postvittana* does not occur according to its actual distribution (Appendix 5c). This might either indicate that the parameters in CLIMEX have not been perfectly calibrated, or some coordinates in CLIMEX might be wrong. After plotting all CLIMEX coordinates into a map, we found that coordinates of Asia in

CLIMEX are very inaccurate (Appendix 5a & b). This does not affect the distribution maps predicted by CLIMEX, but only when these coordinates were used to generate a map other than CLIMEX. After excluding coordinates of Asia, no unexpected points with an EI value greater than 30 were found in the middle of Australia (Appendix 5d).

The predicted distribution of *E. postvittana* are similar to the actual distribution after calibration of parameters in CLIMEX, which indicated that these parameters are suitable for predicting the potential global distribution of the species.

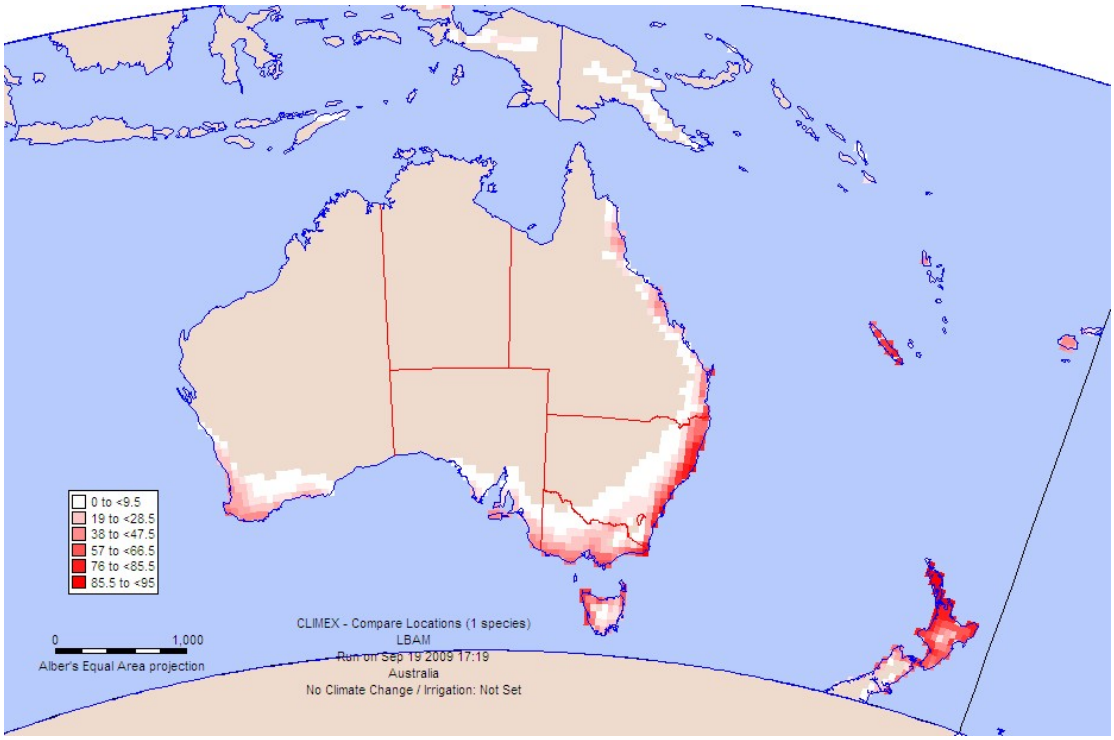



Figure 2.4. Modeled distribution of *E. postvittana* in Australia and New Zealand. *Epiphyas postvittana* is currently established in New Caledonia where the red colour indicates a high EI value. Climatic parameters in CLIMEX were calibrated according to the CABI distribution map (1992) of *E. postvittana*.

### 2.3.2 Potential distribution of *E. postvittana* in North America

*Epiphyas postvittana* is currently established in California USA and countries such as Canada and Mexico consider the moth as an actionable quarantine pest. As a result, the potential distribution of the species in North America was predicted particularly to test modeled results against the actual distribution, at least in California.

The current actual distribution of *E. postvittana* in California is mainly along the

west coast of America (Appendix 1b). The distribution of *E. postvittana* in this region was regulated by government to prevent further expansion of the moth. Predicted potential distribution in USA is mainly in areas along the west coast (EI value ranging from 1 to 44) and in the eastern states such as Virginia, North/South Carolina, Florida, Georgia and Tennessee with EI value slightly higher (ranging from 1 to 47) in some areas than on the west coast (Figure 2.5a). *Epiphyas postvittana* is known established in Hawaii, and our prediction shows high risk EI values ranging from 41 to 100. Central American countries, from Mexico, Guatemala, Honduras, to Panamá, and Cuba, and the Dominican Republic to Puerto Rico, have areas more likely to be suitable for *E. postvittana* establishment than in USA with EI value ranging from 1 to 84. Canada was shown by the model to be free from *E. postvittana* because of the high cold stress (Appendix 6b).

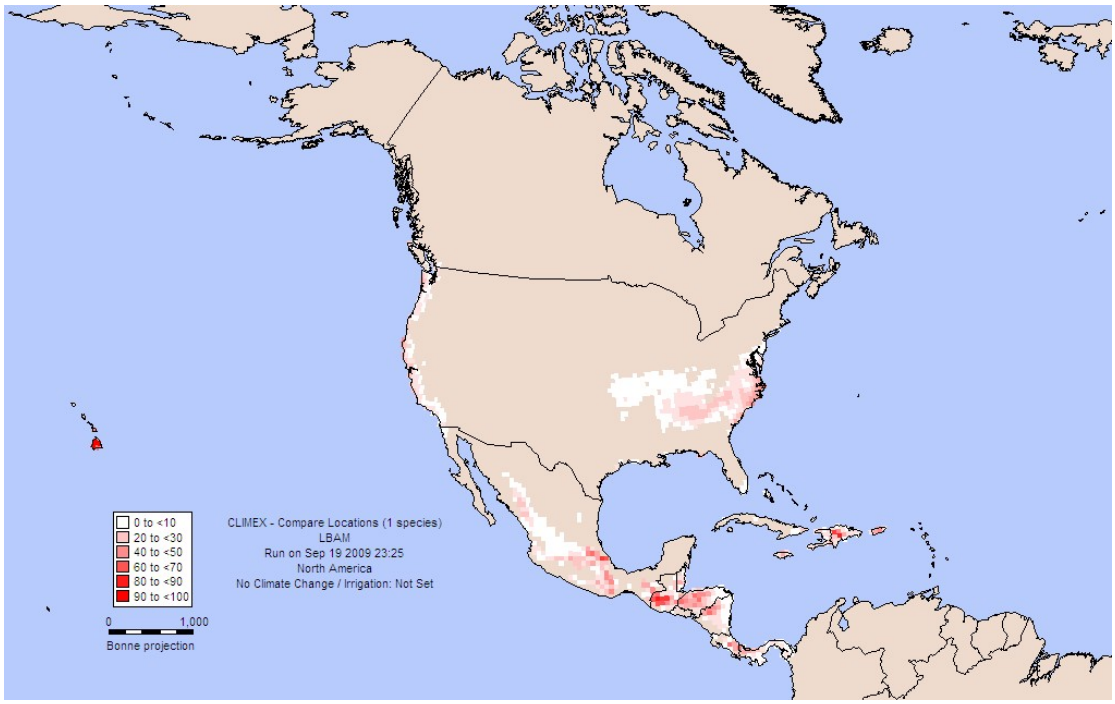
EI values of most areas in California USA were less than 20, and an irrigation parameter was set up as in the template in CLIMEX, which is defaulted as a 3.6 mm/day irrigation rate from May to October. The range of EI values in California increased ranging from 5 to 88 (previously ranging from 1 to 44 without irrigation, see above), and the values decreased slightly in Hawaii ranging from 38 to 97, which is still considered suitable for *E. postvittana* establishment (Figure 2.5b). EI values also decreased in Central American countries mentioned above ranging from 1 to 80 but the range of the area which is suitable for the moth to survive increased.

### 2.3.3 Predicted potential distribution in European countries

Southwestern Europe is relatively more suitable for *E. postvittana* establishment than north and east Europe (Figure 2.6). Most areas in UK are predicted suitable for the moth, as *E. postvittana* is currently widespread all over that country (Suckling & Brockerhoff 2010). Countries such as Portugal, France, Italy, and Greece are predicted under high risk of *E. postvittana* establishment with EI values up to 69. Heat and dry stress indices showed no general effects in Europe. Northwestern Europe appears not to be favorable for *E. postvittana* establishment because of high cold stress in that area (Appendix 7a, b & c). European countries such as Germany, the Netherlands, Belgium, France, Italy, United Kingdom, Ireland and Portugal might potentially be suitable for establishment of *E. postvittana* in a few locations where the EI values range from 50 to 69. The British Isles in which *E. postvittana* is currently distributed,

was shown suitable for *E. postvittana* establishment with EI value around 2-54 and the lowest values are those in the northern area.

a.




b.

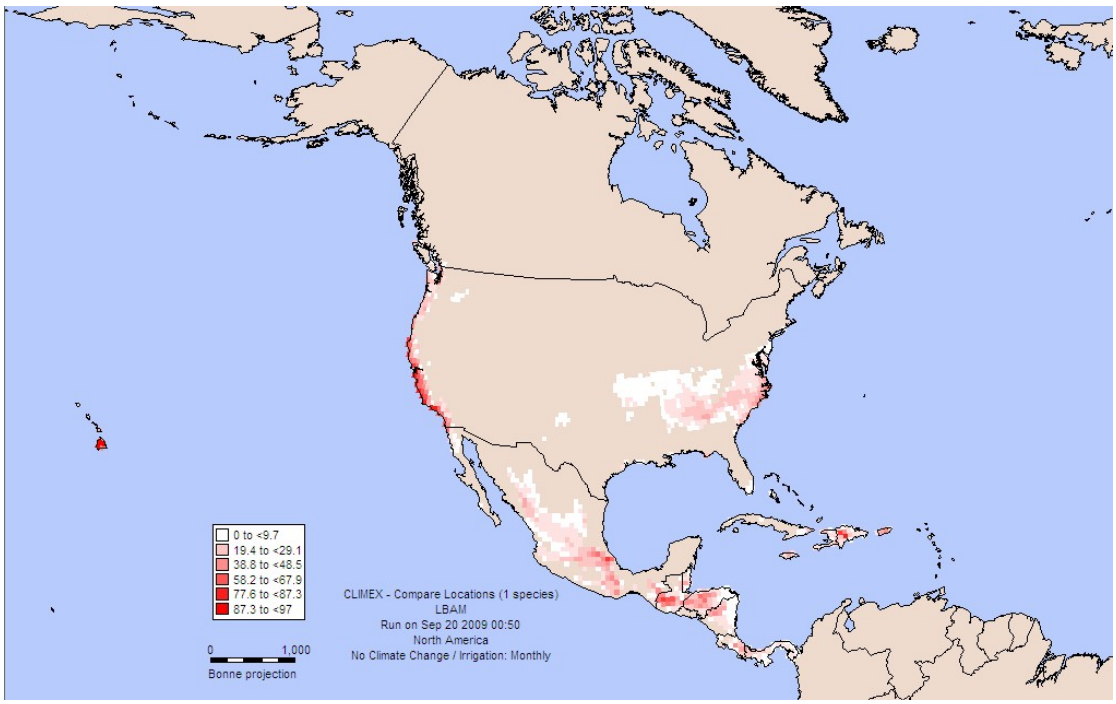



Figure 2.5: Potential distribution of *E. postvittana* in North America. Figure 2.5a shows the potential distribution without considering irrigation and Figure 2.5b shows the potential distribution with irrigation set as 3.6mm/per from May to October.

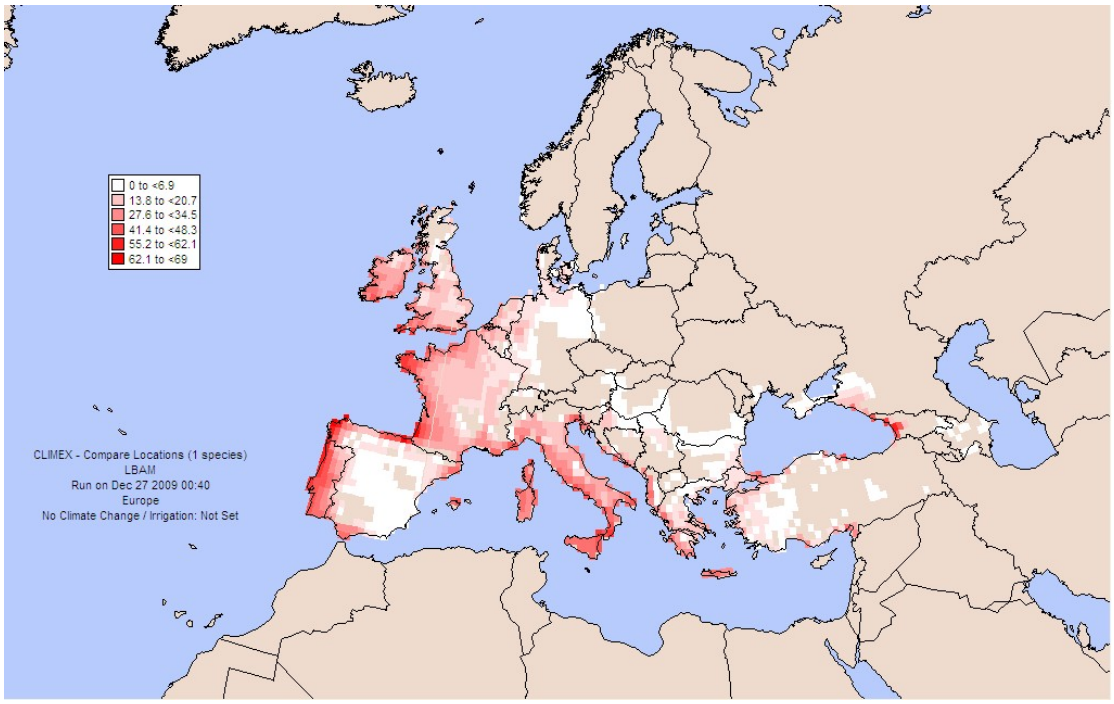



Figure 2.6. Predicted potential distribution of *E. postvittana* in Europe.

### 2.3.4 Predicted potential global distribution of *E. postvittana*

Potential distribution of *E. postvittana* in Asia is predicted mainly in southern Japan (EI up to 53), south-eastern mainland China (EI up to 64), while Taiwan with the highest EI value in Asia up to 89, Vietnam (EI up to 66), Laos (EI up to 33), Burma (EI up to 46), most eastern parts of India (between Burma and Bangladesh with EI up to 62) and small areas within Malaysia, Philippines and Indonesia (EI up to 46) (Figure 2.7).

The potential distribution in Africa is predicted to be mainly in central Africa, with EI values greatest in Ethiopia (up to 80), Madagascar (up to 87), western Kenya (up to 96), areas around Rwanda and Burundi (up to 92) and east coast of South Africa (up to 74) while most areas had an EI less than 20.

The potential distribution in South America concentrates in Colombia (EI up to 100), Ecuador (EI up to 100), central area of Peru (EI up to 91), part of Argentina and parts of Brazil close to Uruguay (EI up to 92). Climate in Uruguay is highly suitable for *E. postvittana* establishment with EI values ranging from 50 to 80 all over the country.

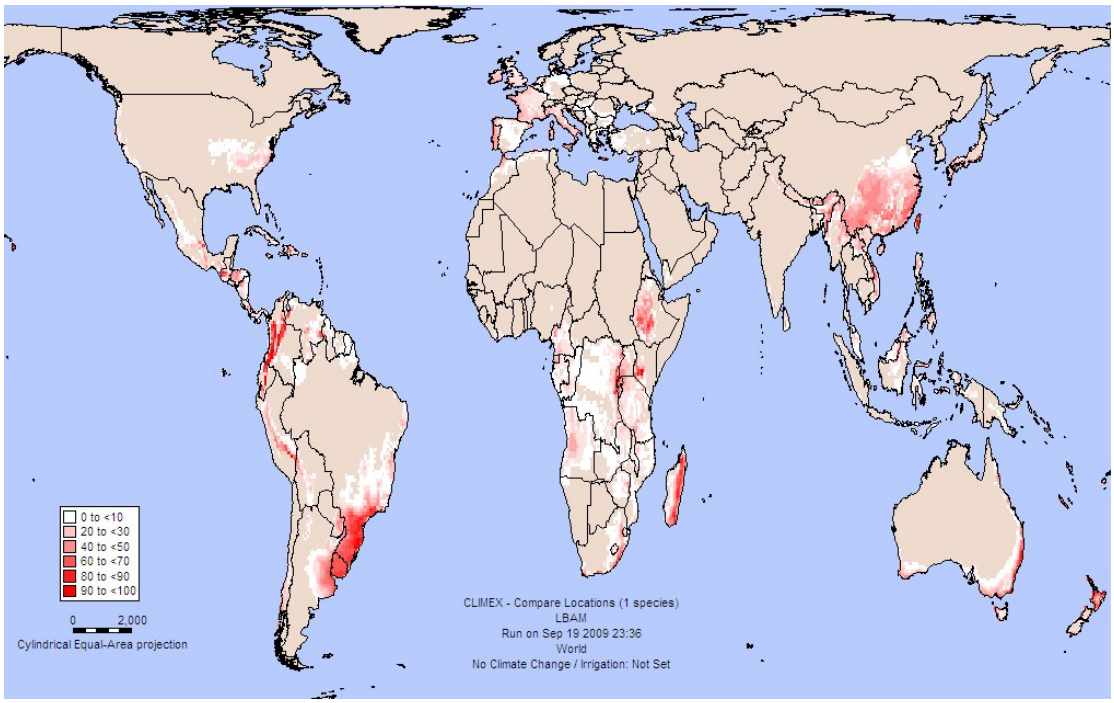



Figure 2.7: Predicted potential global distribution of *E. postvittana*.

*E. postvittana* distribution is limited by cold stress mainly in Asia and North America, and by heat stress mainly in North Africa, south-east Asia and part of South America, and by dry stress in North Africa and Asia (Appendix 8b, c & d).

## 2.4 Discussion

The potential global distribution of *E. postvittana* was estimated using CLIMEX based on its known climatic requirements and current distribution in Australia and New Zealand. The moth has potential to establish in countries mainly in southeastern Asia, southwestern Europe, Central and South America and Central Africa (Figure 2.7). Cold stress limits the potential distribution of *E. postvittana* in North America, Asia, and northwestern Europe, and heat and dry stresses are limiting factors in Africa and Central and South America (Appendix 8b, c, & d).

It is clear that climatic factors, especially temperature, could significantly affect the fitness and activities of the moth, and restrict its distribution in many areas of the world. For example, temperature is known to significantly affect the developmental rate, body size, fecundity and adult longevity of *E. postvittana* (Danthanarayana 1975b; Danthanarayana 1976c; Gu and Danthanarayana 1992b; Danthanarayana *et al*.1995). Dispersal capability, or in other words, flight activity, is also affected by

temperature and humidity (Danthanarayana 1976b). Research by Danthanarayana (1975a), Geier & Briese (1981), and Danthanarayana *et al*. (1995) on the biology of *E. postvittana* has provided the necessary information on the temperature and moisture conditions for development of *E. postvittana* that made this study less subjective. Additionally, the distribution map of *E. postvittana* produced by CABI (1992) provided useful information for calibrating the parameters in CLIMEX. Therefore, the estimated distribution of the moth can provide a reliable indication of high and low risk areas for *E. postvittana* establishment.

A rapid and reliable estimate of the potential environmental suitability for establishment is essential for decision makers when making a risk assessment of potential invasive species. The more extensive information available for modeling using CLIMEX, the more accurate and reliable the result would be. Although research on *E. postvittana* has been carried out over many decades, it has focused more on moth management, and detailed presence/absence information is largely lacking. The lack of good presence/absence information made the calibration of stress indices relatively subjective. Temperature and moisture indices were calibrated using laboratory data, and restricted the potential distribution map of *E. postvittana* to a great extent. However, laboratory data may depart from real field development data of *E. postvittana*, which needs to be further, investigated. The stress indices did not affect the distribution critically, only at those marginal locations with an EI value approximately in the range from 10-30. The high risk areas (with high EI value) were relatively less sensitive to parameter changes.

The distribution of *E. postvittana* is limited by cold, heat and dry stresses in different areas, however, micro-climatic conditions in the crop fields can be significantly affected by human activity (e.g. irrigation and greenhouse production). The distribution of *E. postvittana* in California is apparently a good example of such effect of human cultivate activities. The current model does not include the irrigation parameter, assuming that the distribution of *E. postvittana* in Australia and New Zealand is affected mainly by natural environmental conditions. However, if we assume that irrigation activity significantly decreases dry stress in field crops, and the current distribution of *E. postvittana* is a result of both natural environmental conditions and irrigation activity, irrigation needs be considered. An estimated

irrigation of 3.6 mm/day from May to October (Northern hemisphere) or November to April (Southern hemisphere) as default value in the template was added to the model, requiring the dry stress to be reset as -0.05, which is far more stringent than the template value (-0.0001). The potential distribution of *E. postvittana* in most of the world remains the same after irrigation has been added into the model. However, if irrigation is removed from the model but the dry stress index is unchanged (at -0.05), Western Australia becomes unsuitable for from *E. postvittana* establishment. That makes sense because *E. postvittana* only became established in Western Australia late in 1960s (Wearing *et al*. 1991). When human activities such as irrigation altered the moisture conditions, large areas of field crops became suitable for the moth to establish. Cultivate activities vary around the world, and when interpreting the modeled distribution of *E. postvittana*, such factors should be carefully considered.

Despite the effect of climatic conditions, host plant distribution and natural enemies are also important factors that would affect the potential distribution of a species (Baker *et al*. 2000). For example, *E. postvittana* initially only fed on native shrub species such as genus *Acacias*, and after many exotic plant species were brought in Australia, the species readily adapted (or was pre-adapted) to this change, becoming more polyphagous and a serious pest on cultivated crops such as fruit, vegetable and ornamental plants (Geier & Breise 1981). Also, the variety, quality, quantity, availability and composition of host plants significantly affect *E. postvittana* fitness with respect to fecundity and body size (Danthanarayana 1975a; Gu & Danthanarayana 1992b).

Climate conditions in Australia and New Zealand are highly heterogeneous, which affects the distribution and growth of the host plant species of *E. postvittana* both spatially and temporally. Heterogeneous distribution of host plant species in turn will shape the distribution and abundance of the moth (Danthanarayana 1983; Danthanarayana & Gu 2000). Tomkins *et al*. (1989) and Danthanarayana *et al*. (1995) showed how host plants could affect moth distribution when they showed that populations fed with different food plants developed at different optimum temperatures. While host plant influence is not directly involved in the CLIMEX model although clearly climate will also influence host plant distribution and the potential distribution of the moth should be interpreted with regard to the spatially

explicit distribution of the host plant species.

Moreover, *E. postvittana* is well known for its highly variable morphology, demography, physiology (e.g. Dugdale *et al*. 2005; Danthanarayana *et al*.1995), and genetic variation, which would significantly affect the species adaptability in new habitats (Gu & Danthanarayana 1992a; Gu & Danthanarayana 1992c; Gu & Danthanarayana 2000a; Gu & Danthanarayana 2000b). These factors should be considered along with climatic conditions and host plants when assessing the potential risk of establishment and economic impact of the moth.

Extensive laboratory research has been done on the developmental requirements of *E. postvittana*, however, field data on population abundance and variation are lacking. Presence/absence data is still rare, and economic loss due to *E. postvittana* damage remains uncertain, all of which make the risk assessment of the moth much more than difficult. The reliability of different prediction models varies depending on the quality form and availability of the data, for each pest species (Worner *et al*. 2010) and because there is no perfect model, assembling or comparing the results of more than one model is highly recommended to have a more precise prediction (Worner *et al*. 2010). Most models require not only climate information but good presence/absence data for the species. Moreover, absence data is very difficult to obtain when dealing with IAS.

Danthanarayana *et al.* (1995) stated that "the likelihood of *E. postvittana* attaining a foothold in North America is remote", but these authors highlighted that a full understanding of the ability of the moth to subsist across a range of temperatures is essential to estimate the geographical distribution of *E. postvittana* influenced by climatic boundaries. Ten years later, the moth ***did*** arrive and establish in USA and still the potential distribution has not been fully determined. The CLIMEX model presented here suggests that there are similar climates in other parts of the world that are suitable for *E. postvittana* establishment. Geier and Briese (1981) estimated that the species might be undergoing certain evolutionary changes to adapt in increasingly diverse habitats, thus, the possibility of the moth invading unexpected locations maybe high. The moth may have developed a capability to adapt to new habitats given its wide phenotypic variation in morphology, biology and physiology. A full risk assessment of the moth not only needs to account for the potential global distribution

of the species but also the impact of its genetic variation should also be investigated.

# Chapter 3 Potential distribution prediction of *Epiphyas postvittana* using a new multiple model system and the openModeller

## 3.1 Introduction

In previous work (Chapter 2), we predicted the potential global distribution of *E. postvittana* based on climate using CLIMEX. The advantage of the CLIMEX model is that relatively rapid prediction can be made based on scarce information. Typically, that information is the current distribution of the species in its native range and/or areas where it has been established for a long time (Sutherst *et al*. 2007). Georeferenced presence/absence data (comprising coordinates of each presence or absence location) are not required to be put into the model. Information on the species biological features derived from laboratory or field studies such as the developmental temperature threshold and degree-days for development, greatly facilitates calibrating the parameters of the model and result in more reliable output (e.g. *E. postvittana* prediction in Chapter 2).

However, parameter calibration in CLIMEX is decided by visualization of the correspondence between a distribution map and model output. This process is entirely subjective especially with respect to the calibration of stress indices. The difficulty calibrating such a model is mainly due to the lack of a definite boundary between present and absent areas within the distribution of an organism. Subjective parameter fitting is also required when there is no labotatory or field data on a species' biological features to refer to and/or the presence of a species comprises only several point locations. Furthermore, there is no performance measurement that can indicate the reliability of the final output (the potential distribution map). In other words, statistical evaluation of the resulting distribution map is difficult (or impossible).

There are many other modeling systems similar to CLIMEX that also use the ecological niche concept (Hutchinson 1957) to predict the potential distribution of a species. Examples of such models that have been applied to modelling species potential distribution are artificial neutral networks (ANN) (Gevrey & Worner 2006; Worner & Gevrey 2006; Watts & Worner 2008), support-vector machines (SVM)

(Cortes & Vapnik 1995; Giovanelli *et al*. 2010) and Genetic Algorithm for Rule-set Production (GARP) (Stockwell & Peters 1999; Engler *et al*. 2004). These modeling tools combine occurrence points of a species into environmental layers that enable the computer algorithm to "learn" the ecological or habitat requirements (niche) of the species, and then predict the potential of the species to establish a viable population in a defined area. The performance of these models is usually measured with statistical criteria such as area under the curve (AUC) and accuracy. Different models vary with respect to their capability of prediction (e.g. Elith *et al.* 2006), and the process of using and comparing these models can be complex. Software packages such as openModeller (http://openmodeller.sourceforge.net, Santana *et al*. 2008; Nativi *et al*. 2009; Muňoz *et al*. 2009) and R (http://www.r-project.org/, Leday 2008) provide a platform (framework) to perform and compare the most popular ecological niche modeling algorithms using the same species occurrence data and environmental variables.

In this study we tested the performance of different algorithms in openModeller 1.0 and a new multiple model system (Worner *et al*. 2010) to predict the potential global distribution of *E. postvittana*, and came up with the best model. We expected that the different modelling systems to give different projected distributions of *E. postvittana* worldwide based on the differing data processing capability of each model.

### 3.2 Methods

#### 3.2.1 Data collection

Presence data for *E. postvittana* were collected in two ways:

- **Publication data**. *Epiphyas postvittana* has been extensively studied over the past 30 years, and locations where these studies were carried out and that indicated the presence of the moth were used. Coordinates (longitudes & latitudes) of 69 such locations were obtained using Google-earth (Appendix 2a).

- **GBIF data**. Seventy-six additional presence locations in UK and New Zealand were also found on GBIF (Global Biodiversity Information Facility, http://www.gbif.org, 2010). The data were downloaded directly from the website. Thirty locations remained after combining overlapping coordinates as one

location (Appendix 2b).

A total of 99 point locations (coordinates) were used for further prediction after combining overlapping coordinates. Figure 3.1 shows the current distribution of collected data points.

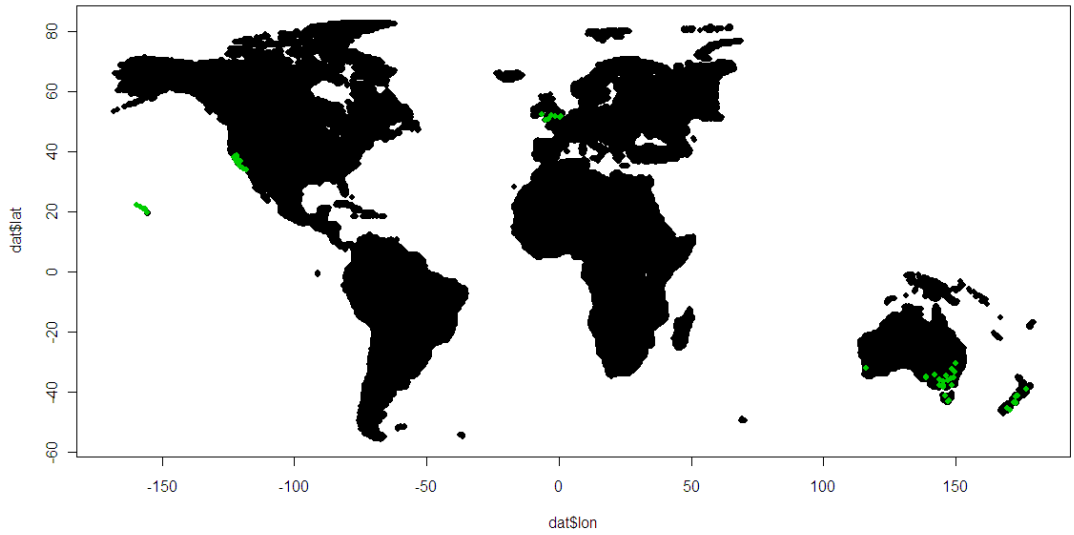



Figure 3.1: Locations of the occurrence data points (green dots) from publications and GBIF.

The same data was used in CLIMEX prediction, however, in different way. In CLIMEX, no coordinates are required for running the program, and parameters were calibrated using mainly the distribution map of CABI (1992). Location data points were used after parameter calibration to confirm that the values of each index were calibrated well. In current study, all prediction models required the actual point locations (coordinates) to be input into the models, which means the data are geo-referenced. The distribution map of *E. postvittana* available in CABI (1992) (Appendix 1) was used to visually confirm the accuracy of modeling results.

### 3.2.2 Prediction with openModeller 1.0

#### 3.2.2.1 Environmental variables

Population dynamics of *E. postvittana* are clearly significantly affected by temperature and precipitation (e.g. Geier & Briese 1981; Wearing *et al*. 1991; Danthanarayana *et al*. 1995), but other variables such as wind speed, diurnal variation

and frost days are rarely recorded as important factors. As a result, the 22 environmental variables available in openModeller 1.0 were reduced to 17 environmental variables based on temperature and precipitation and were used in the current analysis (Table 3.1)

Table 3.1: Variables used in openModeller 1.0 for distribution prediction of *E. postvittana*.

| **Variables in openModeller 1.0** | |
|---|---|
| meanPrecip | meanTemp |
| meanPrecipOverCoolestQ | meanTempOverCoolestM |
| meanPrecipOverDriestM | meanTempOverCoolestQ |
| meanPrecipOverDriestQ | meanTempOverFrostFreeM |
| meanPrecipOverFrostFreeM | meanTempOverWarmestM |
| meanPrecipOverWarmestM | meanTempOverWarmestQ |
| meanPrecipOverWarmestQ | stdevMeanPrecip |
| meanPrecipOverWettestM | stdevMeanTemp |
| MeanPrecipOverWettestQ | |

#### 3.2.2.2 Models

There are nine algorithms available in openModeller 1.0, and seven of them were used to predict the potential distribution of *E. postvittana*: Bioclim, Climate Space Model (CSM), Envelope Score, Environmental distance, OM-GARP (new openModeller implementation GARP), DT-GARP (desktopGARP implementation), and SVM (http://openmodeller.sourceforge.net). ANN was not used because of the limitation of the computer, besides it was used in the other multi-model approach (see below). For GARP, we also ran both best subsets and single run algorithms for OM-GARP and DT-GARP. Therefore, a total of nine models of seven algorithms were performed in this study. The presence data and environmental variables used for running all algorithms were the same. Default parameter values were applied for all tested models.

#### 3.2.2.3 Model evaluation

The performance of each model was evaluated using percentage accuracy and area under the ROC curve (AUC) calculated automatically by openModeller.

### 3.2.3 A new comparative multi-model approach programmed in R

A new multi-model approach recently developed by Worner *et al* (2010) relies on 1) objective, representative extraction of absence data, 2) reduction of environmental

variables to those that are highly relevant and, 3) weighting performance measures such as percentage accuracy, AUC and specificity of each model to choose the best one for prediction of the distribution of the species of interest. This approach was used to, 1) predict *E. postvittana* potential distribution using models not included in openModeller and 2) to compare model performance.

#### 3.2.3.1 Absence data

True absence data for a species is difficult to find. A species that has not been recorded in a particular location does not mean that the climatic conditions in that location are not suitable for the species to establish. The "absence" may be because the species has not had a chance to reach that location because of geographic barriers, or it simply has not been detected yet. Many invasive species are deliberately excluded from certain countries or areas to prevent human activities allowing them to "overcome" natural barriers.

Using one-class support vector machines (OCSVMs) areas that were environmentally dissimilar from areas where *E. postvittana* is currently present, were determined. The OCSVMs calculate an index of environmental suitability between 0 and 1. Environmental suitability less than 0.1 were considered to represent areas where *E. postvittana* is most likely to be absent (Figure 3.2). However, this results in *E. postvittana* potentially absent from large areas of the world over many thousands of locations. To balance the presence/absence data, K-means clustering was used to reduce the data to the same number of presence locations, by choosing 76 clusters (Figure 3.3). Coordinates of these 76 pseudo-absences were estimated using the location with the closest Euclidean distance to the K-means centroid (Figure 3.4). These absence locations were then combined with the presence data for further model comparison and prediction.

Each time, a bootstrap of 63.2% of presence and pseudo-absence data was chosen for training and the rest of 36.7% was set aside for testing.

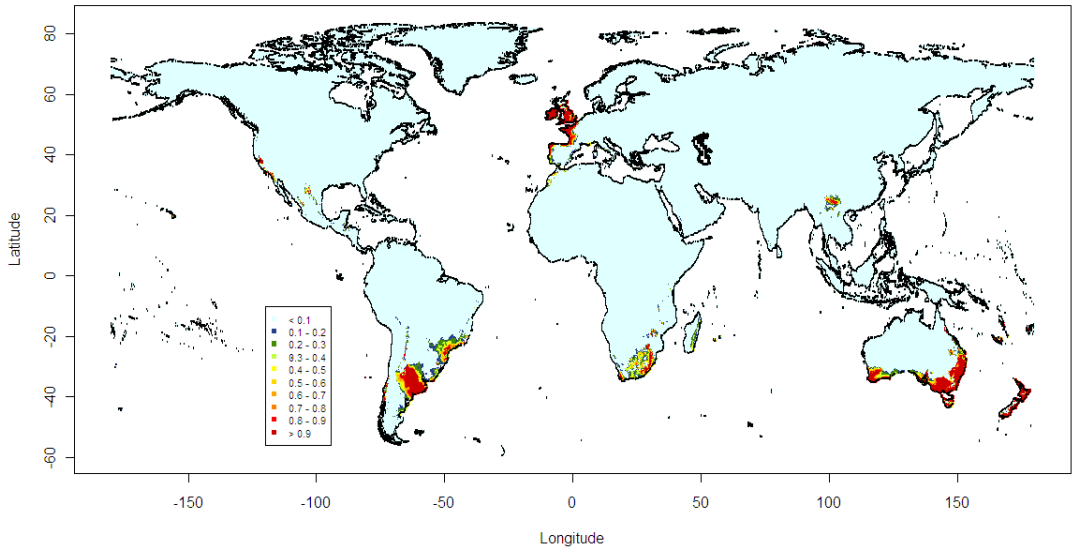



Figure 3.2: Habitat suitability for *E. postvittana* using one-class support vector machines. Warmer colors indicate suitable habitat. Light blue (index less than 0.1) indicates area of unsuitable habitat.

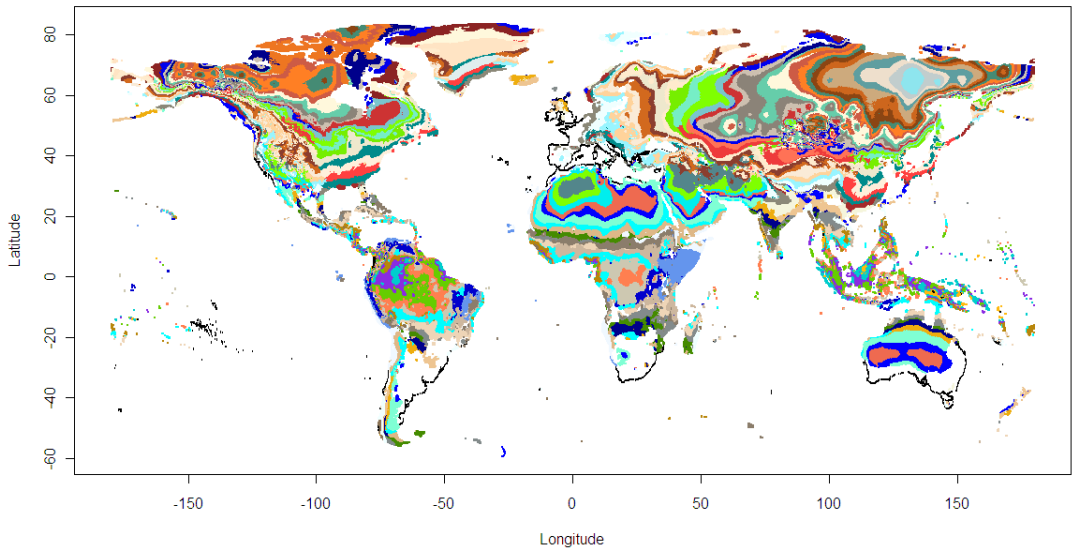



Figure 3.3: Clustered locations determined as unsuitable for *E. postvittana* establishment using OCSVMs. Each colour represents a cluster where environmental conditions are similar.

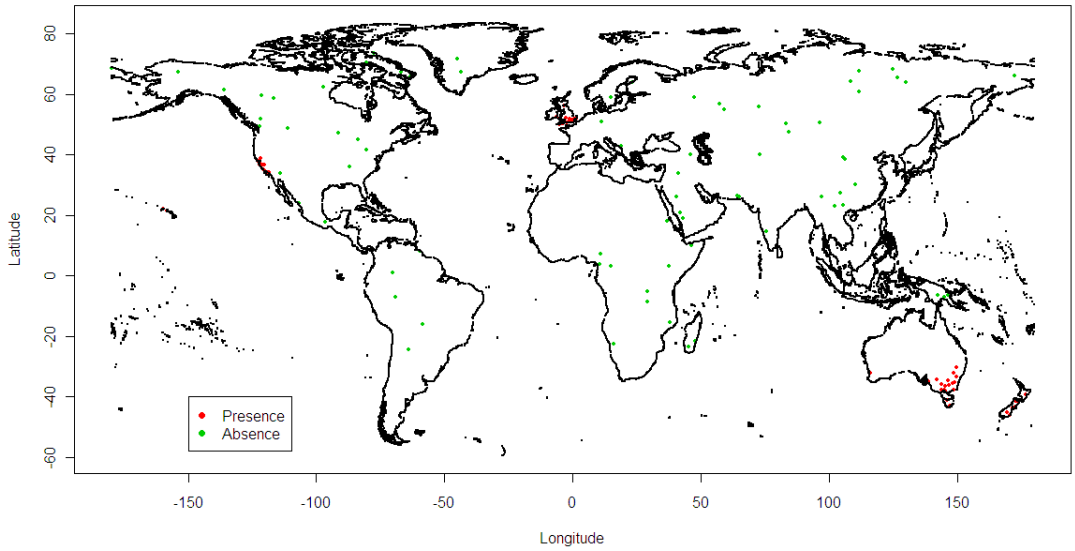



Figure 3.4: Pseudo-absence locations determined using K-means clustering. Each green dot represents an absence derived from the centroid of a K-means cluster. Red dots represent presence data derived from publications and GBIF records.

### 3.2.3.2 Environmental variable selection

Proper selection of environmental variables often reduces the computation time with similar or improved performance of models (Worner *et al.* 2010). The environmental variables considered most relevant for *E. postvittana* establishment were identified using a random forest analysis (Leday 2009; Worner *et al*. 2010; Table 3.2).

Table 3.2: Environmental variables selected for modeling *E. postvittana* potential distribution using random forest analysis.

| Variables |
|---|
| BIO1 Annual Mean Temperature |
| BIO 3 Isothermality |
| BIO 4 Temperature of Warmest Month |
| BIO 5 Min Temperature of Coldest Month |
| BIO 6 Min Temperature of Coldest Month |
| BIO 11 Mean Temperature of Coldest Quarter |
| BIO 13 precipitation of wettest month |
| BIO 15 Precipitation Seasonality |

### 3.2.3.3 Models used in the multi-model approach

Nine models that predict *E. postvittana* distribution were constructed using relevant packages in R (Table 3.3): Linear discriminant analysis, Quadratic discriminant analysis, logistic regression, classification & regression tree, naïve Bayes,

conditional tree, K-nearest neighbors, SVM and ANN. Each model was programmed and parameterized within R (R Development Core Team 2008, http://www.R-project.org), and the performance measured using 10 criteria that included: percentage accuracy, precision, recall, F-score, kappa index, specificity, true skill statistic (TSS), uncertainty, .632+error and AUC. Prediction performance was measured on samples of data independent from the model fitting process, and created by bootstrapping or cross-classification. Each performance criteria was given a rank (numbered 1 to 9 with 1 as best and 9 allocated to the worst performance) for all models. The ranks for the 10 criteria were added up for each model and the one with the least number was judged the best performing model.

Table 3.3: Packages and functions of R used in the current distribution analysis.

| Classifier | abbreviation | Package | Function |
|---|---|---|---|
| Linear Discriminant Analysis | LDA | **MASS** | lda() |
| Quadratic Discriminant Analysis | QDA | **MASS** | qda() |
| Logistic Regression | LOG | **stats** | glm() |
| Naive Bayes | BNET | **e1071** | naiveBayes() |
| Decision Tree | CART | **rpart** | rpart() |
| Conditional Tree | CTREE | **party** | ctree() |
| K-Nearest Neighbors | KNN | **class** | knn() |
| Support Vector Machine | SVM | **kernlab** | ksvm() |
| Artificial Neural Networks | ANN | **nnet** | nnet() |

#### 3.2.3.4 A comparison prediction using presence data after California has been excluded

In Chapter 3, potential global distribution of E. postvittana has been predicted using a multi-model approach, using presence data from both publications and GBIF. The result map predicted a large area of potentially suitable habitat for E. postvittana to survive and establish. In the analysis, the presence data included current distribution of E. postvittana in California. However, we considered that presence data in California might not be suitable to be included in the analysis because E. postvittana in California spread mainly among nurseries, and environmental conditions might be changed due to cultivation activities. It has been also previously indicated using CLIMEX in Chapter 2. To investigate the potential influence of California data on the distribution prediction of E. postvittana, we tested using the same method in Chapter 2 of the new multi-model approach. All the methods are the same as above but excluded presence data from California (Appendix 2 with presence data from California being excluded; Figure 3.5).

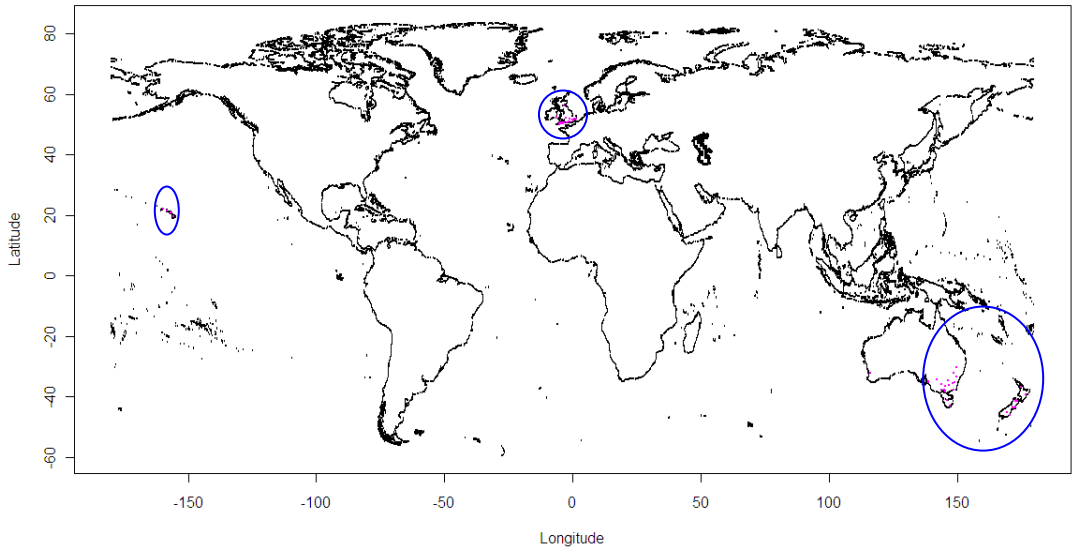



Figure 3.5: Presence data of *E. postvittana* collected from publications and GBIF (red dots, and the blue circles to make them obvious) (details see Chapter 2 and Appendix 2).

## 3.3 Results

### 3.3.1 Prediction of *E. postvittana* potential distribution using openModeller 1.0

Performance criteria applied to the result for all the models in openModeller showed that nine models performed well with high percentage accuracy and AUC, except for a low accuracy for Climate Space Model with a value of 0.681319 (Table 3.4). Although the Envelope Score (ES) ranks the third highest sum of accuracy (=1) and AUC (= 0.97625), the prediction map of ES appears unreliable with the moth distributed all over Australia (Appendix 9d & e), which is not consistent with the current *E. postvittana* distribution as shown by CABI (1992).

Among four different GARP models, the single run–new openModeller implementation has the lowest accuracy with AUC values as 0.802198 and 0.880049 respectively. Both single run and best subset DT-GARP models have better results than of OM-GARP models overall in this study, and GARP (single run) has relatively higher accuracy and AUC values as 1 and 0.9469 respectively than GARP (with best subsets) model.

Table 3.4: Performance criteria for nine models used in openModeller. The Rank of each model is determined by the sum of accuracy and AUC values, and number 1 to 8 represents the highest and the lowest performing model.

| Models | Abbreviation | Accuracy | AUC | Rank |
|---|---|---|---|---|
| Bioclim | Bioclim | 1 | 0.977789 | 2 |
| Climate Space Model | CSM | 0.681319 | 0.922004 | 9 |
| Envelop Score | ES | 1 | 0.97625 | 3 |
| Environmental Distance | ED | 1 | 0.9999 | 1 |
| GARP with best subsets – desktopGARP implementation | DT-GARP (best subset) | 0.967033 | 0.969156 | 5 |
| GARP with best subsets – new openModeller implementation | OM-GARP (best subset) | 0.945055 | 0.95554 | 7 |
| GARP (single run) – desktopGARP implementation | DT-GARP (single run) | 1 | 0.9469 | 4 |
| GARP (single run) – new openModeller implementation | OM-GARP (single run) | 0.802198 | 0.880049 | 8 |
| Support Vector Machine | SVM | 0.989011 | 0.932336 | 6 |

Bioclim performed with the second highest accuracy and AUC values of all nine models in openModeller for *E. postvittana*. However, the prediction map has only two possibilities for the world map, either high risk or no risk (Appendix 9a). It might be considered to over-fit the data, because very few locations other than the presence locations were reproduced.

Environmental Distance (ED) produced highest accuracy and AUC for *E. postvittana* in openModeller with values at 1 and 0.9999 respectively. As a result, Environmental Distance model was considered as the best model in openModeller predicting the potential distribution of *E. postvittana* in this study.

According to prediction map of Environmental Distance, southern South America and Africa and Eastern Europe are climatically susceptible to *E. postvittana* establishment. A small area in south-eastern Asia and areas between North America and Central America are at high risk of *E. postvittana* establishment as well.

### 3.3.2 Prediction of *E. postvittana* potential distribution based on the multi-model approach

Compared with the other eight models in the multi-model system, SVM performed the best predicting the potential distribution of *E. postvittana* on independent data and based on performance criteria ranking (Table 3.5).

Table 3.5: Values for the performance criteria for SVM, the best performing of the eight models.

| Performance criteria | Values |
|---|---|
| Accuracy | 0.968838 |
| Precision | 0.967565 |
| Recall | 0.970694 |
| F-score | 0.968444 |
| Kappa | 0.936825 |
| Specificity | 0.967145 |
| TSS | 0.937839 |
| Uncertainty | 0.039474 |
| .632+ error | 0.020823 |
| AUC | 0.995672 |

The predicted potential distribution of *E. postvittana* using this new multi-model approach shows that Southern North American and South America, most areas in Central America, Eastern Europe, Southern Africa and very small proportion of Asia are significantly at risk of *E. postvittana* establishment (Figure 3.6). The prediction map is very similar to that of Environmental Distance especially locations with high risk (indicated as warmer colors). However, SVM in the new approach predicted more at risk areas in South-eastern Asia, Northern North America, and many islands in each Ocean which seem to have been ignored by models in the openModeller program.

SVM is the only model used in both methods. Compared with the SVM result using presence-only data in openModeller, the SVM accuracy (0.968838) in the new multi-model approach is slightly lower than of in openModeller SVM (0.989011), but AUC (0.995672) is much higher than that of openModeller SVM (0.932336). In general, the new method produced a better result according to its performance over a range of performance criteria that include both accuracy and AUC.

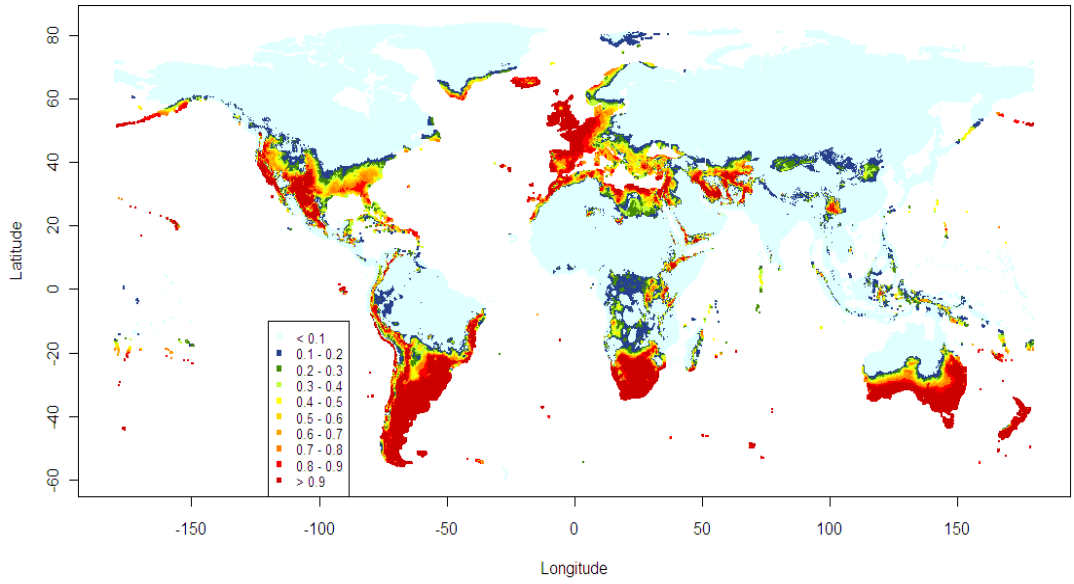



Figure 3.6: Prediction map of *E. postvittana* potential global distribution using SVM in the new comparative multi-model approach.

Prediction with non-California presence dataset produced a better outcome of the potential distribution of *E. postvittana* than that of presence dataset including California. The SVM still performed the best among all the models have been applied. And the performance criteria are all better than using presence data including California (Table 3.6). Distribution map of E. postvittana using non-California presence dataset shows less risk areas, especially in South America and Africa (Figure 3.7).

Table 3.6: Values for the performance criteria for SVM, the best performing of the eight models using none-California presence dataset. The models performed much better without California presence data.

| Performance criteria | Values (without California data) | Values (with California data) |
|---|---|---|
| Accuracy | 0.974285 | 0.968838 |
| Precision | 0.994846 | 0.967565 |
| Recall | 0.954813 | 0.970694 |
| F-score | 0.973702 | 0.968444 |
| Kappa | 0.947891 | 0.936825 |
| Specificity | 0.994876 | 0.967145 |
| TSS | 0.949689 | 0.937839 |
| Uncertainty | 0.025424 | 0.039474 |
| .632+ error | 0.016255 | 0.020823 |
| AUC | 0.999713 | 0.995672 |

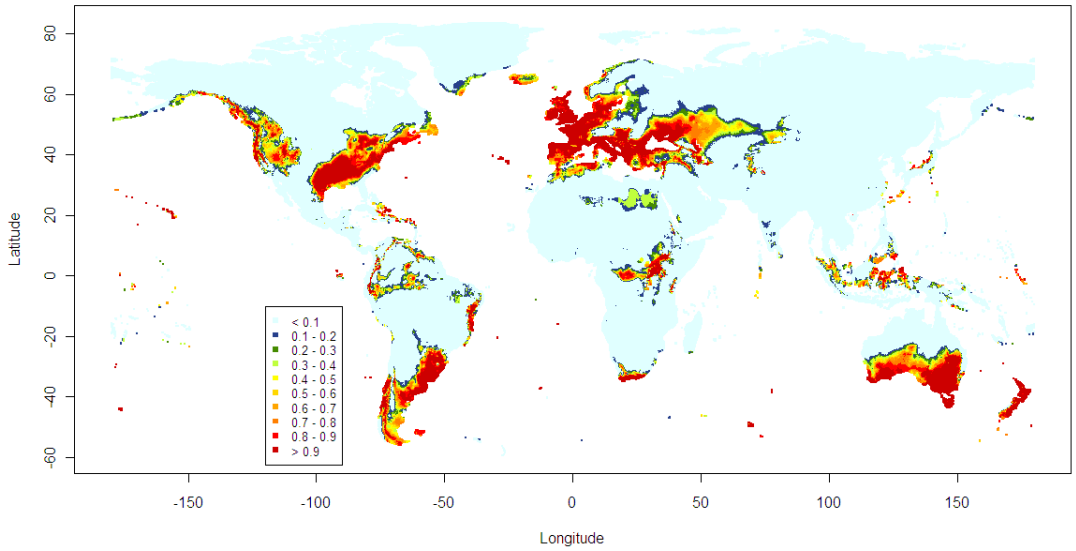



Figure 3.7: Predicted distribution map of *E. postvittana*. Potential suitable area for *E. postvittana* to survive and establish reduced especially in South America and Africa after excluded presence data from California.

## 3.4 Discussion

Potential global distribution of *E. postvittana* has been predicted using two modeling programs, openModeller 1.0 and the new multi-model approach (Worner *et al.* 2010). Environmental Distance using presence-only data in openModeller and SVM using presence/pseudo-absence data in the multi-model approach performed better than other models, where both gave high values over all the performance criteria.

### 3.4.1 Occurrence data

Quality and quantity of occurrence data are essential for accurate distribution prediction using models (Kadmon *et al*. 2003). Stockwell & Peterson (2002) tested effects of sample size on accuracy of four models for distribution prediction of several species and found that models such as GARP and coarse surrogate method (CSM) could achieve 90% of maximum accuracy on about ten data points, nearly maximum accuracy with 50 data points and little increase in accuracy from 50 to 100 data points. In this study, a total of 99 presence data points of *E. postvittana* located within and outside of its native range were used to fit or train the models. Lobo (2008) stated that accuracy differences among models might be diminished when good data is used. The quality of occurrence data can sometimes be problematic especially for new invasive

species where they are often difficult to obtain. Occurrence data are often from museum specimen records and more recently, increasingly available through websites such as the Global Biodiversity Information Framework (GBIF, http://www.gbif.org). However, such data can be less reliable because of unknown collection methods and the identification skills of the recorders. Publication records are more reliable, which is the main data resource used in the current study. The quantity and quality of the occurrence data points in this study was very much guaranteed because *E. postvittana* has been extensively studied for decades due to its economic importance to Australia and New Zealand.

Several studies have suggested that models perform better for species with a defined climate range, or, in other words, a narrow climatic tolerance range) (e.g. Kadmon *et al*. 2003; Tsoar *et al*. 2007). *Epiphyas postvittana* currently has limited distribution and is more confined to cold and wet areas (Danthanarayana 1976; Suckling & Brockhoff 2010), which might contribute to the high accuracy of the model performance shown in this study.

### 3.4.2 Model performance in openModeller

All models tested in openModeller performed very well according to the accuracy and AUC performance metric used, except for Climate Space Model and OM-GARP (single run) both with low accuracy (Table 3.4). Accuracy in openModeller, as used here, may not be the only suitable criterion for judging model performance. Models such as Bioclim, Envelope Score, DT-GARP (single run) and Environmental Distance all apparently use presence-only data points to train the model and the performance of models was tested using the same data points. Therefore, these models gave perfect percentage accuracy "1". While models such as Climate Space Model (CSM), GARP (other than DT-GARP (single run)) and SVM, use a subset of presence-only data and produce a complementary "predicted absence" dataset (the number of predicted absence data equal to the number of excluded presence data in the training process) as the training data. Model predictions are evaluated on the subset of the data that is not used for training. That data is new to the model such that accuracy of these models would never achieve a perfect 1. CSM uses the least number of presence-only data as training data and produced more predicted absence data. As a result, the accuracy was the lowest.

The other performance criterion, AUC is extensively used for model evaluation. AUC ranges from 0 to 1. A score of 1 represents a perfect discrimination between presence and absence of a species, and 0.5 represents the discrimination is equal to random guess (Hanley & McNeil 1982; Phillips *et al*. 2003; Elith *et al*. 2006). With presence-only data, the maximum AUC would be less than 1 (Phillips *et al*. 2003). AUC value of all models in openModeller, except OM-GARP (single run), are greater than 0.92. Although presence-only data is considered suitable for distribution prediction, models such as Bioclim and DT-GARP performed worse than other presence/absence models such as MAXENT and OM-GARP (Elith *et al*. 2006; Tsoar *et al*. 2007; Giovanelli *et al*. 2010). Moreover, Bioclim performed worse than DT-GARP in above mentioned studies, while in the current study, Bioclim ranks the second and much better than GARP and SVM in openModeller even without considering percentage accuracy. This seems very unusual compared with other studies.

In general, distribution prediction of *E. postvittana* using models such as Envelope Score, Bioclim and Environmental Distance in openModeller could not be evaluated statistically for their accuracy. Using AUC only to evaluate the performance of the rest of models is very unconvincing.

Other problems encountered while using the different models in openModeller were: 1) both presence and absence data points can be recognized in openModeller. However, some data points may end up being discarded because they lack matching environmental data. Environmental layers from Worldclim (http://www.worldclim.org) could not be used openModeller, which made the presence/absence modeling impossible within this system. 2) Parameter fitting is also time-consuming in openModeller, because each parameter combination comparison required manual adjustment.

### 3.4.3 The Multi-model approach

Locations where a species is present represent environmental conditions that are suitable for a species to survive, while absence data, if available, would indicate environmental conditions that are not suitable for a species to survive. Therefore, presence-absence data were considered as unbiased reflection of "ecological niche" of

a species (Elith & Leathwick 2009). As absence data is difficult to obtain and identify, pseudo-absence data is now commonly used in distribution models. Examples of model comparison between using presence-only data and presence/pseudo-absence data include Elith *et al*. (2006) and Giovanelli *et al*. (2010). The discussion however, continues on this subject, but it seems clear that, models using a combination of presence/pseudo-absence data, if that data is appropriately selected and used, would significantly outperform models using presence-only data.

Pseudo-absence data are usually extracted using presence-only data models from habitat areas where the species is not known to be present, and the main objection of using pseudo-absence data may be because the data extraction method may not result in representative absence data. For example, pseudo-absence data were randomly sampled from the calibration area of OM-GARP and MAXENT respectively in Giovanelli *et al*. (2010). Chefaoui & Lobo (2008) sampled 10 times as many pseudo-absence data as presence data, using ENFA (Ecological Niche Factor Analysis) modeled locations with habitat suitability threshold (HS) less than 0.1 and 0.2 which performed best in GLM (Generalized Linear Models). Engler *et al.* (2004) applied ENFA to produce pseudo-absence data sample within HS less than 0.3. Such data are likely at risk of being biased especially when a large area of unsuitable habitat is predicted. The new method applied in this study by Worner *et al*. (2010) extracted pseudo-absence data by grouping areas with environmental suitability predicted less than 0.1 to match the number of presence data points, and measured the centroid of each group, which guarantees the quality (objectiveness and representativeness) of the sampled data.

Model comparison is very complicated and problematic. Single performance criterion, such as AUC and kappa, has often been used for evaluation of model performance (e.g. Tsoar *et al*. 2007; VanderWal *et al*. 2009). Some criteria such as MPA (minimal predicted area) are rarely used but considered as a suitable criterion for model performance evaluation of rare/endanger species prediction (Engler *et al*. 2004). However, deviation among performance criteria could be significant (Elith *et al*. 2006). Therefore, using many different criteria can best describe how well a model performs as well as, make model comparison well balanced. Currently, the new method includes 10 most commonly used criteria giving an evaluation capability that

is superior to any other study.

Also, the advantage of the multi-model system is that all models are programmed in R, and the extremely complicated and repetitive calculation and parameter adjustment can be programmed to be automated thus significantly shortening model running time.

### 3.4.4 Potential distribution prediction of *E. postvittana*

The result of SVM predictions using the multi-model tool was adopted as the final distribution prediction of *E. postvittana*. The distribution map predicted a large area around the world that is under high risk of *E. postvittana* establishment with the exception of Asian countries (Figure 3.6). The presence locations in California may be one important factor that accounts for the large potential distribution areas of *E. postvittana*. *Epiphyas postvittana* is mainly spread among nurseries in California, where human cultivation activities such as irrigation and artificial climate conditions (e.g. greenhouse) might affect the establishment of *E. postvittana* in this area. Prediction using presence data excluding of California may clarify whether these factors are important. Comparison of this result with that of CLIMEX will be further addressed in Chapter 6.

# Chapter 4 Identification of *Epiphyas postvittana* for biosecurity using DNA barcoding

## 4.1 Introduction

With the rapid escalation of global trade and tourism, the risk posed by invasive alien species (IAS) on the economy and environment of most nations has also increased. Rapid and accurate identification of the newly arrived exotic species intercepted at borders is most important to prevent their establishment and spread (Armstrong & Ball 2005; Ball & Armstrong 2006). Traditional morphological taxonomy is still the main method for arthropod identification at the border. However, identification of arthropods, the most species-rich animal phylum in terrestrial ecosystems, is often extremely time-consuming. Many specimens must be sent to specialist taxonomists for identification. This is a slow process and not practical at the border. Moreover, there are four significant limitations of traditional morphological taxonomy (Hebert *et al*. 2003; Ball and Armstrong 2006). First, phenotypic plasticity and genetic variability in taxonomically important characters can lead to incorrect identification. Second, morphologically cryptic species (species that morphologically similar or appear to be identical but genetically often distinct), which are common in many pest groups (Jarman & Elliott 2000), cannot be identified by standard morphological methods. Third, morphological keys are often only available for certain life stages (e.g. adults) or gender, and it is difficult to identify other life stages (e.g. eggs, larvae, and pupae) to species level. Fourth, even if keys are available, years of experience is often required to use them effectively.

Light brown apple moth (LBAM), *Epiphyas postvittana*, adults are well-known for their morphological polymorphism (Dugdale *et al*. 2005). Often it is very difficult to morphologically identify the moth accurately without genitalic dissection. It is even more difficult to identify the immature stages of *E. postvittana* for which there are limited identification keys (Dugdale *et al*. 2005). Misidentification of *E. postvittana* might be one of the main reasons for the latest establishment of the moth in California and its establishment and wide-spread occurence in the UK (Takahashi 2002; Venette *et al.* 2003; Fountain & Cross 2007).

DNA barcoding is a recently promoted alternative method for rapid species

identification compared with traditional morphological taxonomy. The DNA barcode consists of a ~650 bp fragment of the mitochondrial gene, cytochrome *c* oxidase I (COI) existing in all aerobic life including animals and is widely used to determine the identification for most insect taxa and other arthropods (Hebert *et al*. 2003). The COI gene is involved with oxidative metabolism and energy generation, and it evolves at an appropriate rate to distinguish most well-defined animal taxa (Behura 2006), i.e. variation within a species is significantly less than it is between species. Based on the standardization of the method, including the same gene region, same molecular sequencing approach and demand for voucher specimens linked to the barcode sequence, this approach has been proposed as a standard method for exotic insect identification in biosecurity (Armstrong and Ball 2005). Within *E. postvittana* populations, variation in the COI gene has been found to be less than 2.3% (Barr *et al*. 2009 unpub. data), but it is more than 7% between *E. postvittana* and other native moth species in California (Barr 2007). Such results indicate that barcodes are a promising method for *E. postvittana* identification for specimens intercepted at the border if the other, potentially confounding species in the country are known.

Currently only one *E. postvittana* COI sequence is publicly available from GenBank (http://www.ncbi.nlm.nih.gov/genbank), which is not sufficient as a reference dataset for this species. Barr (2007) tested barcoding for discriminating *E. postvittana* from native moth species in California, but that data is not available. Therefore, in the current study, the COI of *E. postvittana* from four New Zealand populations was sequenced. The potential ability to distinguish these *E. postvittana* from other tortricid species represented in BOLD (barcode of life data systems, http://boldsystems.org) was determined using the standard neighbour joining analysis as is commonly used to make (Ratnasingham & Hebert 2007). We expected, in keeping with previous studies (Barr 2007), the intraspecies variation among *E. postvittana* populations to be far less than the interspecies variation between *E. postvittana* and other Tortricidae in BOLD. Therefore, barcoding can be used at the border for rapid and accurate *E. postvittana* identification.

## 4.2 Methods

### 4.2.1 *Epiphyas postvittana* specimens

Colony *E. postvittana* specimens were provided by Plant & Food Research Ltd. Auckland, reared under laboratory conditions on artificial diet for several generations. Field *E. postvittana* specimens were provided by Plant & Food Research Ltd. collected with LBAM-specific pheromone sticky traps from Hawke's Bay, Lincoln, and Clyde (Appendix 10). These specimens were morphologically identified by providers (although their genitalia have not been checked). One Lincoln specimen was collected using light trap and identified using Dugdale *et al*. (2005). All specimens were frozen at -20$^{o}$C for further analysis.

### 4.2.2 DNA extraction

The upper part of the left mid-leg femur of each adult *E. postvittana* was used for DNA extraction, and the remainder of each specimen was preserved as voucher in separated containers frozen at -20$^{o}$C. The tissue was cut into small pieces and placed into an extraction solution of prepGEM$^{TM}$ kit (ZyGEM, Inc., Solana Beach, CA): 40 μL of 1×buffer and 1 μL of Enzyme, and incubated at 75$^{o}$C for 1 h, then 99$^{o}$C for 5mins.

### 4.2.3 Polymerase chain reaction (PCR)

The mitochondrial gene, COI, was amplified by using the universal primer pair LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer *et al*. 1994) for ~650 bp of the barcode region. The reaction mixture contained: 2.5 mM of $MgCl_2$, 0.2 mM of dNTPs, 0.06 μM/L of each primer, 0.35 units of polymerase, 1 μL of the DNA extraction and deionized water to bring the total reaction volume to 10μL. The PCR thermal regime consisted of: one step at 94$^{o}$C for 2 min, 40 cycles of 94$^{o}$C for 15 s, 52$^{o}$C for 30 s, and 72$^{o}$C for 1 min, and with a final step at 72$^{o}$C for 5 min. PCR products were visualized on 1% agarose gel stained with Sybr Safe$^{TM}$ DNA Gel Stain (Invitrogen Inc., CA) and submerged gel electrophoresis at 80 volts for 30 min.

### 4.2.4 Sequencing

PCR products were sequenced in both directions using 0.8 μM of each primers, 0.5 ul Big Dye version 3.1 (Applied Biosystems, Warrington, Cheshire, UK) and the following thermal regime: one step at 96$^{o}$C for 1 min, 25 cycles of 96$^{o}$C for 10 s,

50°C for 5 s, and 60°C for 4 min. Sequence products were resolved in an AVANT 3100 (ABI) capillary sequencer.

### 4.2.6 Data analysis

Sequences were pruned to ~619 bps, then aligned and edited using Sequencher 4.0. All sequences were submitted to the BOLD identification engine to determine their similarity compared to tortricid sequences including *E. postvittana*.

To calculate the divergence within *E. postvittana* and between species of Tortricidae, a total of 1487 Tortricidae sequences currently available in BOLD were downloaded; this included 179 species from 60 genera of four subfamilies. Short sequences (less than ~350 bp) were excluded from the analysis, leaving a total of 1435 sequences for 173 species. Pairwise divergence within *E. postvittana* and among the tortricid species was calculated and an unrooted neighbor-joining (NJ) tree of all Tortricinae species available in BOLD was built using MEGA 4.1 (Kumar *et al*. 2008) based on the K-2-P parameter.

## 4.3 Results

Five COI sequences of ~ 658bp long from each of four populations were sequenced plus an extra Lincoln sample collected with light trap. All sequences have been submitted to GenBank (accession numbers: GU827562 – GU827587).

All 26 COI sequences grouped into the same cluster as 31 other *Epiphyas postvittana* sequences in BOLD (Appendix 11a) (these 31 sequences are only available for species identification in BOLD system but not publicly available from BOLD, since only one sequence of *E. postvittana* is accessable in GenBank as we stated in Introduction section above). These 31 sequences were collected from Australia (n=16), United States (n=3), United Kingdom (n=14), and one specimen from an unspecified location. There are three other *Epiphyas* species in BOLD system, *E. ashworthana* (n=14), *E. peloxythana* (n=1) and *E. xylodes* (n=3). The variation within *E. postvittana* is less then 1%. There was greater than 3% divergence between *E. postvittna* and *E. xylodes*, but close to 3% from *E. ashworthana* and *E. peloxythana*. The divergence between *E. postvittana* and the other seven species of six genera from Tortricidae is much greater than 3%.

The intraspecies variation of *E. postvittana* (26 sequences) calculated using MEGA 4.1 is about 0.2%; there was no variation among 24 individuals, only one Lincoln individual with 0.48% and one Hawkes Bay individual with 0.23% variation from the others. Divergence between *E. postvittana* and other Tortricidae species in BOLD ranges from 7 to 16%. An unrooted neighbor-joining (NJ) tree of all Tortricinae species in BOLD is attached (Appendix 11b), and illustrates *E. postvittana* can be readily distinguished by this method from all Tortricinae species publicly available in BOLD.

**4.4 Discussion**

Identification of species using DNA barcoding is based on the principle that intraspecies variation is greater than interspecies variation (Hebert *et al*. 2003*b*). Variation within four *E. postvittana* populations including one colony population is about 0.2% in this study and less than 1% within *E. postvittana* sequence data in BOLD. The four populations in our study were all collected from New Zealand, which might account for the lower divergence than that of BOLD populations collected from a broader geographic range. Barr *et al*. (2009, unpub. data) found a maximum 2.3% divergence between populations from California, southern England, New Zealand, Australia, and Hawaii. The South Island of New Zealand in their study was found to be the most genetically variable region besides Australia, which was used to identify the original of *E. postvittana* in California. One Lincoln specimen collected from apple orchard using sticky trap had the "greatest" variation in this study.

Barr and his colleagues also investigated Tortricidae species native to the US in the genera *Choristoneura*, *Archips*, *Platynota*, *Amorbia*, *Pandemis*, *Decodes*, *Cnephasia*, *Deidra*, *Argyrotaenia*, *Clespsis*, *Sparganothis* and *Eulia* and found a greater than 7% divergence of these species from *E. postvittana*, which is consistent with our analysis using BOLD data. This further indicates that the use of barcoding can be effective to distinguish *E. postvittana* identification from at least 173 other Tortricidae species. Some species in genus *Epiphyas* might be more difficult to discriminate from *E. postvittana* due to small divergence between these species (Appendix 11). *Epiphyas* is a leafroller genus native to Australia and consists of 32 species including *E. postvittana* (Common 1961). Species such as *E. pulla* and *E.*

*xylodes* have been reported to hybridize with *E. postvittana* in the field (Geier & Springett 1976). However, the variation between *E. xylodes* and *E. postvittana* COI is greater than 3% (Appendix 11) which is greater than variation between *E. postvittana* and the other two species (*E. peloxythana* and *E. ashworthana*). Regardless of the relatively lower variation, there was no single *E. postvittana* sequence that has been classified into other species of *Epiphyas*. Moreover, *Epiphyas* species are all native to Australia, therefore the relative lower variation between *E. postvittana* and other *Epiphyas* species would not become problematic since all species in *Epiphyas* are quarantine target species at the border.

For two other leafroller genera, *Ctenopseustis* and *Planotortri*, that are endemic to New Zealand, it was found that whilst DNA barcoding can distinguish between the two genera and some species within each genus, it was unable to distinguish many species within each genus (Langhoff *et al*. 2009). However, the variation between *E. postvittana* and species of these two genera ranges from 9% to 15% (Appendix 12b), which makes distinguishing *E. postvittana* from species in these two genera possible.

Identification for surveillance using barcoding sometimes can be time-consuming when a large number of specimens are submitted (Barr *et al*. 2009). Barr *et al* (2009), however, have applied the ITS2 multiplex markers of *E. postvittana* that has successfully identified a large number of moth specimens within California. However, ITS2 profiles of amplicon patterns are not available for species of Tortricidae from regions other than California. It was suggested that programs integrate all these three methods to achieve the most cost-effective, rapid and accurate identification of *E. postvittana* for biosecurity purposes (Barr *et al*. 2009).

# Chapter 5 Phosphoglucose isomerase: a gene with links to fitness and potential invasive capability of *Epiphyas postvittana*

## 5.1 Introduction

The potential global distribution of *E. postvittana* (light brown apple moth, LBAM) has been predicted by comparing climate conditions of its native ranges to the rest of the world (Chapter 2 & 3). However, whether the species can actually establish and further spread in a new "potentially-suitable" environment depends not only on the suitability of climatic conditions in that new habitat, but also on the adaptability of the species to the new environmental conditions. Adaption of an organism to a new habitat means a population shifts to a phenotype that best fits the current environmental conditions (Orr 2005). Genetic variation, in response to environmental change, affects performance and fitness of a species, which may lead to population reduction or extinction of the species (Watt 2000). Therefore, analysis of the genetic variation of a species in different environmental conditions would contribute to predicting the performance and fitness of a species in new habitat.

*Epiphyas postvittana* is known for high variability in its morphology (Dugdale *et al*. 2005), demography (Geier & Briese 1980a, b) and physiology (Gu & Danthanarayana 2000a). Environmental conditions, especially temperature, significantly affect the life-history parameters of the moth, such as fecundity, development time and adult longevity, and these in turn determine the population fitness of the moth (Danthanarayana 1975a; Danthanarayana 1976c; Gu and Danthanarayana 1992b; Danthanarayana *et al*.1995). Moreover, given the same environmental conditions, populations collected from different localities also show significant variation in their life history traits and dispersal ability (Geier & Briese 1980a, b; Gu & Danthanarayana 2000a).

Quantitative genetic analysis of *E. postvittana* has shown that life-history traits, plus dispersal based on flight capacity, are genetically related and significantly influenced by environmental factors and that there may also be one or several common genes involved (Gu & Danthanarayana 1992a, c; Gu & Danthanarayana

2000a & b; Via 1983). However, no research has been done to date to investigate the gene(s) that potentially relate to the fitness and flight capability of the moth, which, at the same time, may be affected by environmental factors such as temperature and humidity.

One potential genetic mechanism could involve the phosphoglucose-6-isomerase (*Pgi*) gene. The product of this gene is an enzyme (PGI), which catalyses the second step in glycolysis and is also involved in several other metabolic reaction sequences (Watt 1977). Accumulated evidence in other insect species has shown that PGI is related to the fitness (e.g. fecundity) and dispersal capacity of the species and is subject to natural selection (e.g. Watt 1983; Watt *et al.* 1983; Riddoch 1993; Hagg *et al*. 2005; McMillan *et al.* 2005). For example, for the butterfly species *Colias eurytheme* (Lepidoptera: Pieridae) and *Melitaea cinxia* (Lepidoptera: Nymphalidae), individuals with different dominant PGI genotypes, show significant differences with respect to their fecundity (Watt 1992), male mating success (Watt *et al*. 1986), flight metabolic rate (Haag *et al*. 2005), population growth rate (Hanski & Saacheri 2006) and lifespan (Saastamoinen *et al.* 2009). In another butterfly species, *Lycaena tityrus* (Lepidoptera: Lycaeidae), different PGI genotypes were shown to associate with life-history trait variation, such as larval and pupal development time, growth rate, and pupal mass, and species resistance to extreme temperatures (Karl *et al* 2008; Karl *et al.* 2009). Different PGI genotypes have also been shown to affect the running speed in a leaf beetle *Chrysomela aeneicollis* (Coleoptera: Chrysomelidae) (Dahlhoff & Rank 2000).

The expression of different PGI forms is linked to different environmental conditions, particularly temperature. For example, 10 PGI genotypes of a butterfly species, *L. tityrus*, were detected according to their different allozyme mobility. Of these, PGI genotype 2-2 was the dominant genotype in alpine populations, which also had increased cold stress resistance and a relatively long development time (Karl *et al.* 2009). In another butterfly species, *M. cinxia* (Lepidoptera: Nymphalidae), egg clutches of individuals with *Pgi*-f allele were 32% larger than *Pgi*-non-f individuals (Saatamoinen & Hanski 2008).

These studies suggest that the *Pgi* can be considered as a candidate gene that reflects the effects of environmental conditions on insect population dynamics (Zheng

*et al.* 2009). However, most of these organisms, such as butterflies (e.g. *C. eurytheme*, and *M cinxia*) and the leaf beetle (*Chrysomela aeneicollis*) are bio-indicators and no PGI studies have been applied to answer questions in pest control or the biosecurity domain. However, Hanski & Saccheri (2006) point out that such environment-*Pgi*-fitness effects might be strong in "the expanding front of invasive species".

Recently, nucleotide sequences of the *Pgi* gene has been successfully accessed in several butterfly species. The gene in *Colias* butterfly species is a large (about 10 kb) single copy locus composed of 12 exons with 11 intervening introns (Wheat *et al.* 2006). Knowledge at the nucleotide level can make the more precise investigation of this genotype-environment-phenotype relationship possible (Orsini *et al.* 2009; Wheat *et al*. 2010). This study collected *E. postvittana* populations from the North and South Islands of New Zealand, where the climatic conditions are warmer in the North Island than in the South Island (NIWA, 2009). It aims to develop the first step in the investigation of the genetic variation of *Pgi* in *E. postvittana* with respect to different geographic populations from different climatic conditions. Specifically, this study will first involve the development of a new PCR-based system to isolate the gene so that the exons and introns can be sequenced, and second, determine if there is any nucleotide variation that can be subsequently screened at the population level.

## 5.2 Methods

### 5.2.1 *Epiphyas postvittana* specimens

Colony *E. postvittana* specimens were provided by Plant & Food Research Ltd. Auckland, reared under laboratory conditions on artificial diet for several generations. Field *E. postvittana* specimens were provided by Plant & Food Research Ltd collected with LBAM-specific pheromone sticky traps from Hawke's Bay, Lincoln and Clyde (Appendix 10). All specimens were morphologically identified by providers (genitalia have not been checked). One specimen from Lincoln was collected using light trap and morphologically identified according to Dugdale *et al*. (2005). The COI barcode region gene was sequenced and morphological identification was further confirmed using BOLD Identification (see Chapter 4). Specimens were frozen at -20$^{o}$C degree for further analysis.

### 5.2.2 DNA and RNA extraction

DNA was extracted from the whole abdomen of each adult moth using DNeasy Blood & Tissue kit (Qiagen, Inc., Valencia, CA), following the manufacturer's instructions. The quantity of DNA in each sample was measured using nanodrop 3.0 (NanoDrop Technologies, Inc, Wilmington, DE).

Total RNA was extracted from two whole moths from the laboratory colony population using the RNeasy mini kit (Qiagen, Inc., Valencia, CA) following the manufacturer's instructions. RNA concentration was measured using nanodrop 3.0 (NanoDrop Technologies, Inc, Wilmington, DE) and diluted at a ratio of 1:5 with di$H_2O$. cDNA synthesis used 5µl of the diluted RNA and primers PGI-12_R (Table 5.1) and oligodT(20) respectively, with Superscrip III reverse transcriptase (Invitrogen Inc., Foster City, USA) according to the manufacturer's instructions.

### 5.2.3 Primer design

#### 5.2.3.1 Degenerate primers

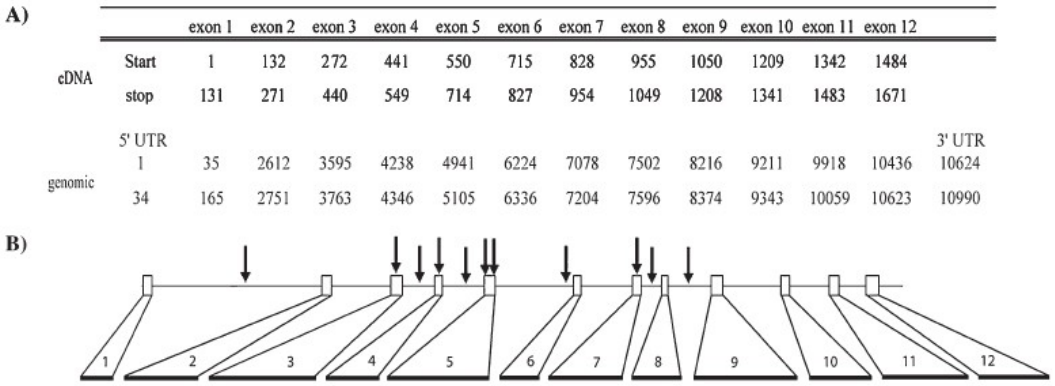



Figure 5.1: cDNA and genomic DNA structure of *Colias* PGI with start and stop position in both cDNA and genomic DNA. Boxes represent 12 exons and dark lines represent 11 introns (diagram source: Wheat *et al*. 2006).

*Pgi* sequences of mainly Lepidopetera insect species, including *C. eurytheme*, *Bombyx mori*, *Colia meadii*, *Cydia pomonella*, *Antheraea paukstadtorum*, plus several *Drosophila* species were downloaded from GenBank. These were aligned in MEGA 4.0 (Kumar *et al*. 2008) using the exon-intron map of Wheat *et al.* (2006) (Figure 5.1) as a backbone template from which the boundaries of exon-introns had been defined. Degenerate primer pairs were designed according to the most conservative regions, which is located in exons 11 & 12, exons 5 & 6 and exon 4.

#### 5.2.3.2 LBAM-*Pgi*-specific primers

Degenerate primers were used to amplify corresponding regions between exon 11 to 12 and exon 5 to 6 on two samples from each of four *E. postvittana* populations. Based on the successful sequences exon 5 of *E. postvittana Pgi*, an LBAM-*Pgi*-specific primer was designed for amplification of exon 4 to 5. Since the GenBank-sequence alignment used to design degenerate primers shows few constant conservated areas, it was time-consuming and non-cost-effective to design effective degenerate primers from those sequences. Therefore, we designed LBAM-*Pgi*-specific primer pairs in exon 4 and 12, to amplify the cDNA of *Pgi*. More LBAM-*Pgi*-specific EPIC primer pairs were designed to amplify different regions of *Pgi* in *E. postvittana* based on cDNA sequences, and *E. postvittana Pgi* sequences amplified using degenerate primers.

#### 5.2.3.3 TAIL-PCR (thermal asymmetric interlaced PCR) primers

Based on *Pgi* exon sequences of *E. postvittana* (section 5.2.4), two TAIL-PCR primer sets were designed to amplify the 5' and 3' end of the *Pgi* respectively using the first strand of cDNA.

Optimal primer features such as melting temperature and primer dimmers, were determined with online software (Integrated DNA technology, www.idtdna.com). Optimal annealing temperatures of the different primer sets were tested using gradient PCR ranging from 48 - 52$^{o}$C; 50$^{o}$C was a suitable annealing temperature for most primer sets. Annealing temperature for TAIL-PCR primers are higher around 65$^{o}$C because of the TAIL-PCR requirements.

### 5.2.4 PCR amplification and sequencing of *Pgi*

Genomic DNA was amplified using the designed degenerate primers in 20 µl PCR reactions containing final concentrations of 2.5 mM $MgCl_2$, 0.06 µM of each primer (Table 5.1), 0.2 mM of each dNTP, 0.05 U of HiFi Taq DNA polymerase (Invitrogen, CA), plus 2 µl of genomic DNA. An initial denaturing step of 5 min at 95$^{o}$C was followed by 40-42 cycles of amplification with 30 sec at 95$^{o}$C, 30 sec at 50$^{o}$C and 1.5-3 min (depending on the length of target DNA) at 72$^{o}$C, and a final extension step of 7 min at 72$^{o}$C. Sometimes, PCR amplification was difficult and

likely due to the poor quality of genomic DNA (specimens collected with sticky traps produced far less DNA than colony specimens or light-trap fresh specimens). Usually, 5% DMSO, touch down PCR or less stringent PCR conditions (lower annealing temperature) were used to produce better results. Variation in exons and introns of *Pgi* gene among populations was investigated using the designed LBAM-specific EPIC primers and the same PCR protocol and thermocycle regime as that for the degenerate primers.

cDNA was amplified using 1 µl of the first strand cDNA in a 10 µl PCR reactions containing with final concentration of 2.5 mM $MgCl_2$, 0.06 µM of each designed specific primer, 0.2 mM of each dNTP, and 0.05 U of HiFi Taq DNA polymerase (Invitrogen, California), with a cycle regime of: 95 $^oC$ for 5min, then 40 cycles of 95 $^oC$ for 30 sec, 50 $^oC$ for 30 sec, 72 $^oC$ for 2min, with a final extension at 72 $^oC$ for 7 min.

The TAIL-PCR PCR protocol to amplify the 5' end and 3' end of the *Pgi* using the first strand synthesized cDNA was modified from Liu *et al*. (1995) and Mullins *et a*l. (2001). However, the experiment failed and this was not pursued due to the research time and budget limitation.

All PCR products were visualized on 1% agarose gel stained with Sybr Safe$^{TM}$ DNA Gel Stain (Molecular Probes, Oregon) and submerged gel electrophoresis at 80 volts for 30 min. For *Pgi* amplification using EPIC primers from genome DNA, 18µl of PCR product were loaded to the gel wells in case gel extraction was required. Where two PCR products of different length were amplified from the same sample, these were excised separately off the agarose gel and extracted using Zymoclean gel DNA recover kit (Zymo research, Orange, CA) following the manufacturer's introduction. After dilution at 1:50, 2 µl of the gel extract was re-amplified using the same PCR protocol and gel electrophoresis.

Final PCR products with bright single band on the agarose gel were sequenced in both directions using 0.8 µM of each primer, BigDye version 3.1 (Applied Biosystems, UK) and the following thermal regime: one step at 96 $^oC$ for 1 min, 25 cycles of 96 $^oC$ for 10 sec, 50 $^oC$ for 5 sec, and 60 $^oC$ for 4 min. Sequence products were resolved in an AVANT 3100 (ABI) capillary sequencer (Applied Biosystems Inc., Foster City, CA,

USA).

### 5.2.5 Data analysis

Sequences were pruned and edited in Sequencher 4.0 and aligned in MEGA 4. Segregating sites were calculated with all sequenced exon fragments in MEGA 4. Each intron was aligned with sequences from different populations.

## 5.3 Results

### 5.3.1 Primers designed

#### 5.3.1.1 Degenerate primers

Six degenerate primers were successfully applied to amplify the genomic DNA of *Pgi* gene in *E. postvittana* (Table 5.1 marked with "*"). Primer Exon 11 to 12 and exon 5 to 6 including introns between the exons were successfully amplified from two specimens from each population, showing one or two bands for each sample (Figure 5. 3). Nucleotide sequences of partial exon 11 plus 12 and exon 5 plus 6 (excluded introns) were blasted in GenBank respectively and they are 86-87% & 80% and similar to *Pgi* sequences of *M. cinxia* and *C. eurytheme* correspondingly.

*Pgi* sequences downloaded from GenBank used to design the degenerate primers have very conservative areas in exon 11 & 12 and exon 5 & 6, but not in other exons. Primers designed from other exon regions either failed to produce PCR products or produce multiple weak PCR products.

Table 5.1: Primers designed for *Pgi* amplification. Amplicon size of each amplification region (excluding the length of primers) is estimated based on *Pgi* sequences of *C. eurytheme* (Wheat *et al.* 2006). Primer set 4, PGI-19_F & PGI-12_R, was used in cDNA amplification PCR.

| No. of primer set | Amplification region | Degenerate primers (5'-3') | Tm ($^{o}$C) | Amplicon size estimated |
|---|---|---|---|---|
| 1 | exon11-12 | *PGI-1_F: CCSCACAAAGTGTTCAA | 50.3 | 485bp |
| | | *PGI-2_R: TCGAASGAGTTGATGTCC | 51.6 | |
| 2 | exon5-6 | *PGI-5_F: ACTTTCACMACHCARGARAC | 52.3 | 1261bp |
| | | *PGI-6_R: CCVACCCARTCCCARAA | 53.8 | |
| 3 | exon4-5 | *PIG-17_F: GGHCCDCTSATGGTCAC | 54.6 | 781bp |
| | | PGI-20_R: AGACTTACATCCTTAGCAGCG | 54.3 | |

| No. of primer set | Amplification region | Degenerate primers (5'-3') | Tm ($^{o}$C) | Amplicon size estimated |
|---|---|---|---|---|
| 4 | exon4-11 **(for cDNA amplification)** | PGI-19_F: CACAGGCAAAAGCATCACTG86 | 54.8 | 996bp |
| | | PGI-12_R: AGTGCTCCGAGGGTGAAAG | 57.1 | |
| 5 | exon 4-5 | PGI-19_F: CACAGGCAAAAGCATCACTG | 54.8 | 832bp |
| | | PGI-20_R: AGACTTACATCCTTAGCAGCG | 54.3 | |
| 6 | exon 5-6 | PGI-30_F: CACTTCGTCTCCAACATCG 143 | 53.2 | 1301bp |
| | | PGI-6_R: CCVACCCARTCCCARAA 40 | 53.8 | |
| 7 | exon 6-7 | PGI-31_F: CKAAGCACTTYGTCGCTC 82 | 53.7 | 931bp |
| | | PGI_32_R: TTTCTCCAGAGGCGCG 108 | 55.1 | |
| 8 | exon 7-8 | PGI-33_F: TACTCCCTCTGGTCGGC 110 | 56.2 | 483bp |
| | | PGI-41_R: CCTGTGAAGATACTGGTCGT 76 | 53.9 | |
| 9 | exon 8-9 | PGI-42_F: CGCCYGTRATYYTGGCT 76 | 55.1 | 827bp |
| | | PGI-34_R: CCTGGTGTATGAGCTGGTAG132 | 54.6 | |
| 10 | exon 9-10 | PGI-35_F: TTCGCCGCGTACTTCCA 140 | 57.1 | 1084bp |
| | | PGI-43_R: CCGAAGACTTKCCCTTCAT 108 | 53.5 | |
| 11 | exon 10-11 | PGI-11_F: GCTRATYCCNTGCGAYTT 116 | 52.9 | 809bp |
| | | PGI-12_R: AGTGCTCCGAGGGTGAAAG 119 | 57.1 | |
| 12 | exon 3-4 | *PGI-15_F: GCTSGMMCAYATGAAGGA -36 | 53.4 | 571bp |
| | | PGI-28_R:GCCTCAGTCACCATAAGAGG 61- | 54.8 | |

Note: F = forward primer; R = reverse primer; degenerate base symbols: R = A/G, S = C/G, M = A/C, H = A/C/T, V = A/C/G. Primers designed from sequences of other Lepidoptera downloaded from GenBank are indicated with *. Other primers were designed from *E. postvittana Pgi* sequences. Estimated amplicon sizes do not include the size of primers.

### 5.3.1.2 LBAM-*Pgi*-specific EPIC primers

Seven EPIC primers were designed to amplify regions from one exon to the next exon from exon 4 to 11 (Table 5.1; Figure 5.2). Most of the time the PCR amplification produced only one or two bands; seldom did an EPIC primer pair produce more than two bands of PCR products.

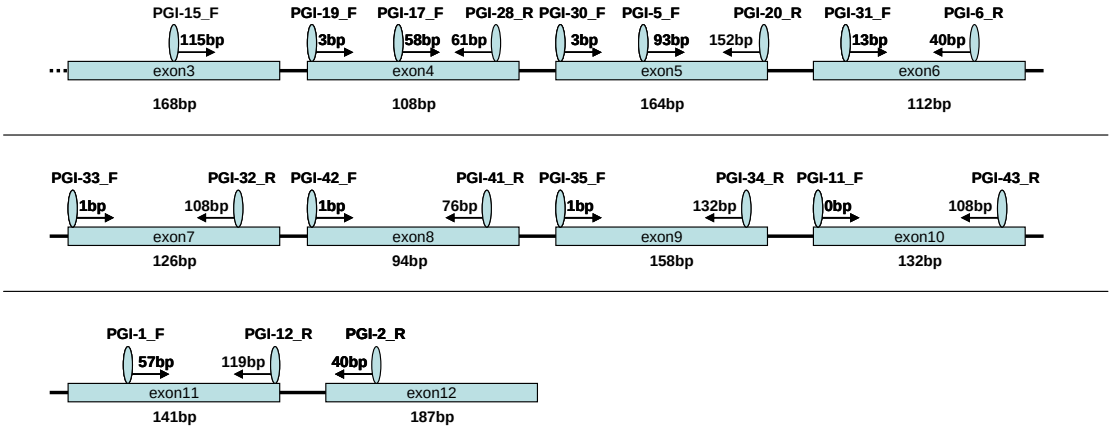



Figure 5.2: Location of each EPIC primer along the coding region of *Pgi* gene. Numbers indicate the location of the first base pair (5') of each primer from the first base pair (5') its exon. Not drawing to the scale.

### 5.3.1.3 TAIL-PCR primers


Figure 5.3: Locations of TAIL-PCR primers along the coding region of *Pgi*.

Two TAIL-PCR primers for amplification of 5' and 3' ends of *Pgi* were designed (Table 5.2). Figure 5.3 shows the location of each TAIL-PCR primers along the coding region of *Pgi*. TAIL-PCR primers failed to produce any visible PCR product on an agarose gel. This method has been successfully applied to a gene of fungi species (Hayley Ridgway person. comm.), but these were not used as a positive control in this study, which might have helped to determine if was only an issue with primer design.

Table 5.2: Tail PCR primer sets: (the experiment has FAILED).

| Location of the primers | Primers (5'-3') | Tm ($^{o}$C) |
|---|---|---|
| exon10~11 | LBAM_3'_Tail_F1: CGCAGACGCACAACCCCATCTC | 62.8 |
| | LBAM_3'_Tail_F2: AACTTCCTGGCKCAGACTGAGGC | 61.0 |
| | LBAM_3'_Tail_F3: GGCAACCGRCCCACCAACTCCA | 64.3 |
| exon4~5 | LBAM_5'_Tail_R1: CCGCCAACCCAGTCCCAGAATCC | 64.2 |
| | LBAM_5'_Tail_R2: GGTGGCGTTGGTGATGGTCTCCTG | 63.9 |
| | LBAM_5'_Tail_R3: TGGGTGCCGTCGATGTTGGAGAC | 63.5 |

## 5.3.2 Coding region of *Pgi* in *E. postvittana*

A ~957 bp coding region of *Pgi* cDNA (two colony specimens) between exon 4 to 11 was successfully sequenced using LBAM-specific primer pair PGI_19_F & PGI_12_R. The exon-intron boundaries are identical to those in *C. euritheme* and *M. cinxia*. The percent sequence identities for the coding region compared with GenBank blast hits of *M. deione*, *Bombyx mori*, *C. euritheme*, *M. cinxia* and *Drosophila ananassae* are 77-78% in general for nucleotide sequences, and 90%, 87%, 87%, 90%, and 78%, respectively, for amino acid sequences. The G+C content of this region is 59.2%. There are 31 segregating sites within the cDNA coding region, with two sites being different between the two individuals (Appendix 12a). All these sites are synonymous, therefore, the amino acid sequences of these two specimens are identical (Appendix 12b). No segregating sites were found in exon 9.

The coding region from exon 4 to 12 in 17 *E. postvittana* specimens (eight colony and nine wild, from Clyde n=5, Lincoln n=3 and Hawkes Bay n=1) was sequenced. No part of the *Pgi* gene from exon 4 to 12 (including introns) has been able to be completely amplified from any *E. postvittana* specimen, because different primer sets worked differently with different samples. A total of 70 segregating sites were found amongst the four populations, including 61 synonymous and nine nonsynonymous (Appendix 13). Five of the synonymous sites and one nonsynonymous site were found in the exon 12 partial sequences, a region which had not been sequenced in previous colony cDNA sequences. Eight synonymous and two nonsynonymous sites were found at exon 9.

### 5.3.3 Introns of *Pgi* in *E. postvittana*

Introns 3-11 (excluding intron 10 where no sequence was obtainable) were successfully sequenced for only 13 individuals from four populations (colony n = 4; Hawkes Bay n = 1; Lincoln n = 3; Clyde n = 5). The PCR products normally showed one or two bands on the agarose gel by electrophoresis. Most of these were sequenced easily, but sometimes the sequence of those with a single bright band was too messy (with multiple overlapping peaks) to analyze. In general, the length of introns of *Pgi* in *E. postvittana*, with an average length of 789bp, is much longer than the average intron length in *C. eurytheme*, with an average length of 622 bp (introns 3-12 excluding 10).

The length of introns varied in several samples, both within the same specimen and among populations. Figure 5.4 shows the amplification of genomic DNA between exon 11and 12. Two bands were amplified from the colony sample, with introns different by 335 base pairs, while only one band was amplified from each wild sample from Lincoln and Hawkes Bay as confirmed by clean DNA sequence. The intron length of the shorter allele of colony specimen was 27 base pairs longer than of Hawkes Bay and 119 base pairs longer than of Lincoln.

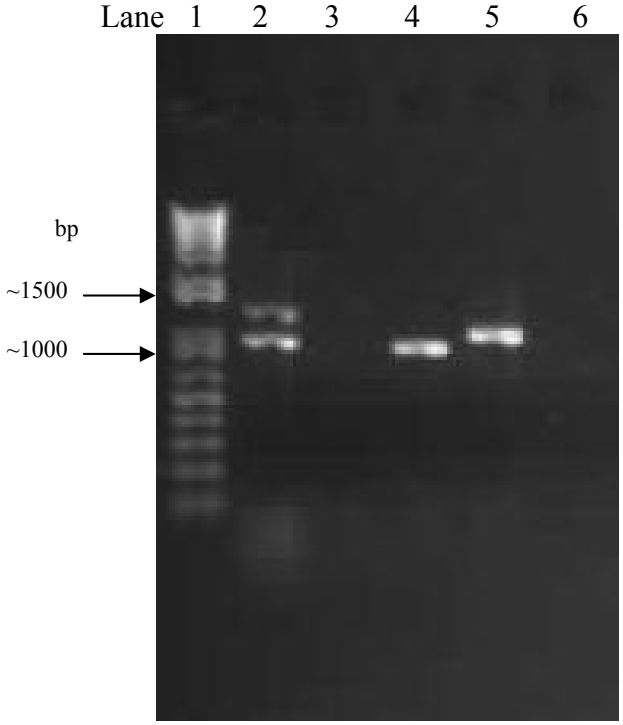



Figure 5.4: PCR products of exon 11-12, using degenerate primers PGI-1_F & PGI-2_R. Lane 1 is 1 kb plus DNA ladder, lane 2-5 are LBAM genomic DNA samples from colony, Clyde, Lincoln & Hawkes Bay respectively, lane 6 is a negative control.

From their sequences the introns varied significantly in length among populations as well as alleles within an individual (Table 5.3). For example, intron 11 of two alleles of the same colony sample has a 337 bp difference, while the same intron for a Clyde sample was much shorter than for the colony specimens, which had a 258 bp difference between the two alleles. Intron 11 for two Lincoln samples each produced two single bands with a 556 bp length difference between them. The most significant difference was at intron 5 for two Clyde samples with a 793 bp difference. Apart from the base insertions and deletions influencing the length, most introns are very similar and identifiable with respect to their nucleotide sequence. Intron 9 was the most variable in terms of nucleotide sequence among populations and allele length.

Table 5.3: Intron variations of *Pgi* gene in *E. postvittana*. Length of each intron in the butterfly *C. eurytheme* is indicated under each intron number (Wheat *et al*. 2006). There are single or double bands from the same sample; single bands from different samples can have significantly different length (indicated with short/long band in "single band" column). The dashed line indicates that no sequence was obtainabed either because of failed PCR or messy sequence.

| Intron | Population | No. of samples sequenced | Length of intron (bp) | | | |
|---|---|---|---|---|---|---|
| | | | Double bands | | Single band | |
| | | | Short band | Long band | Short band | Long band |
| 3 (475 bp) | Colony | 3 | 843 (identical) | - | - | - |
| 4 (505 bp) | Colony | 1 | - | - | 953 | |
| | Hawkes Bay | 1 | - | - | 947 | |
| | Clyde | 1 | - | - | 947 | |
| 5 (1119 bp) | Colony | 4 | 899 (identical) | - | - | - |
| | Hawkes Bay | 0 | - | - | - | - |
| | Lincoln | 0 | - | - | - | - |
| | Clyde | 4 | ~745/910 | - | 749 | 1542 |
| 6 (742 bp) | Colony | 2 | - | - | 460 | |
| | Hawkes Bay | 1 | - | 926 | - | - |
| | Lincoln | 2 | - | - | 460 | 738 |
| | Clyde | 5 | 460/599/738 | 926 | - | - |
| 7 (298 bp) | Colony | 2 | - | - | 606 | |
| | Hawkes Bay | 1 | - | - | - | - |
| | Lincoln | 1 | - | - | 624 | |
| | Clyde | 1 | - | 540 | 600 | |
| 8 (620 bp) | Colony | 1 | - | - | 419 | |
| | Hawkes Bay | 1 | - | - | 660 | |
| | Lincoln | 1 | - | - | 755 | |
| | Clyde | 2 | 421 | - | 370 | |
| 9 (837 bp) | Colony | 1 | - | 1149 | - | - |
| | Hawkes Bay | 1 | 849 | - | - | - |
| | Lincoln | 2 | 1051 | - | 886 | |
| | Clyde | 2 | 1204 | - | 1206 | |
| 11 (377 bp) | Colony | 4 | ~775 | ~1112 | - | - |
| | Hawkes Bay | 1 | - | - | ~748 | |
| | Lincoln | 3 | - | - | ~1112 | ~656 |
| | Clyde | 2 | ~490 | ~748 | ~748 | |
| Total No. of sequences | | 50 | | | | |

## 5.4 Discussion

### 5.4.1 Novel primers for *Pgi* gene amplification in *E. postvittana*

Primers designed in this study successfully amplified part of the *Pgi* gene of *E. postvittana*. Degenerate primers (PGI-1_F & PGI-2_R and PGI-5_F & PGI-6_R) were designed from the most conservative areas of *Pgi* exon sequences of several Lepidoptera species. These therefore might be suitable to amplify the *Pgi* gene of other Lepidoptera species, although we have not tested the assumption yet.

However amplification across the introns was not always successful; even when a single band was produced the sequences were messy and other times no band was produced. This inconsistency might be due to the LBAM-specific primers being designed from limited *E. postvittana Pgi* sequences of the cDNA libraries for only two colony specimens, but variation being present at the *Pgi* priming sites in the other populations.

### 5.4.2 Variation of coding region (exons) of *Pgi* gene in *E. postvittana*

#### 5.4.2.1 Full length of cDNA sequence of *Pgi* in *E. postvittana*

The coding region of *Pgi* gene of *E. postvittana* has been partially sequenced to include ~ 957 bp of nucleotides across exons 4 to 11, while the full length remains unknown because of the un-sequenced 5' and 3' ends of the gene. The full length of coding region of *Pgi* shows a small difference between *C. eurytheme*, which is 1668 bp long (556 amino acids) and *M. cinxia* 1671 bp (557 amino acids) (Wheat *et al.* 2006; Orsini *et al*. 2009). Additional work to design primers for the 5' and 3'ends plus exon 10 will need to be achieved before the complete gene can be sequenced in *E. postvittana*. A system to amplify the two ends of *Pgi* coding region can be accomplished by RACE (Rapid Amplification of cDNA Ends) (supporting material, Orsini *et al*. 2009), but the resources required for this, e.g., by SMARTer RACE (Clontech Laboratories, Inc. Mountain View, CA), were beyond the scope of this project.

#### 5.4.2.2 Single copy or multiple copies of *Pgi* gene in *E. postvittana*

From the cDNA and EPIC sequences here, it is most likely that the *Pgi* gene in *E. postvittana* is a single copy (one locus) gene consisting of two different alleles, as has been described in other insects (He & Haymer 1997). PGI in all invertebrates investigated to date possess only a single *Pgi* gene locus (e.g. Watt 1977; Wheat *et al.* 2010). However, in a study of 53 species of bony fish belonging to 38 families, two independent structural gene loci were shown to exist, most likely resulting from gene duplication (Avise & Kitto 1973). In *E. postvittana*, using only DNA analysis, it is difficult to confirm whether the variation observed on occasion within an individual is a result of amplifying alleles of different loci or different alleles of the same locus. Further investigation using enzyme electrophoresis may help determine the number of

copies (locus/loci) of the gene in *E. postvittana*. Since it has largely been mobility differentiating allozymes and allozyme frequencies that have been shown to be significantly related to environmental conditions, similar allozyme analysis may be necessary to further this study on *E. postvittana*. Such a study will be required in any attempt to predict PGI phenotypes related to certain environmental conditions, or to study the evolutionary mechanism of PGI with respect to environmental variation (Riddoch 1993). Unfortunately, the quality of tissue for allozyme analysis is very important and the tissue available from trap caught specimens in this study was not of that quality.

#### 5.4.2.3 Homozygosity of colony populations

The two cDNA sequences from exon 4 to 11 from the *E. postvittana* colony population have the same amino acid sequences (Appendix 12b). The specimens had been freshly frozen and the DNA extraction quality was better than of specimens from the sticky traps. This resulted in the colony population showing the most *Pgi* exon sequences compared with other populations. However, of the nine nonsynonymous polymorphic sites found amongst the other populations only two were found in only one colony sample (Appendix 13). The PGI enzyme is known to be highly polymorphic in many taxa (Watt 2004). For example, seven alleles have been distinguished in the butterfly *M. cinxia*, and 10 genotypes have been identified in *L. tityrus* (Haag *et al*. 2005; Karl *et al.* 2009). Therefore, we assume that homozygous PGI may be dominant in the colony population, which have to be further tested. This *E. postvittana* colony has been reared for several generations under constant "optimal" conditions, and if released into the natural environment they die more readily than natural populations (Chung pers. commun.). This is consistent with heterozygosity usually being associated with survival in environmental uncertainty (Nevo & Beiles 1988). Accordingly, PGI heterozygotic advantage in natural *Colia* butterfly populations was reported with respect to their survival rate, flight activity, male mating success and female fecundity (Watt 1977; Watt 2003). The likely PGI homozygosity in the *E. postvittana* colony might result from the invariant laboratory conditions, compared with the heterozygotic natural populations studied here.

The complete coding region of *Pgi* in *E. postvittana* was not sequenced in this study, therefore, it is not yet possible to determine which population has the most

variation. Further study on the coding region of *Pgi* in different *E. postvittana* populations from known variant environmental conditions would greatly facilitate making associations with respect to *E. postvittana* fitness.

#### 5.4.2.4 More variation in *E. postvittana Pgi* is expected

The most extensively studied lepidoptera species with respect to PGI are the butterflies *Colias eurytheme* and *M. cinxia*. The coding region of *Pgi* is 1668 bp and 1671 bp long in *C. eurytheme* and *M. cinxia* respectively, and 119 synonymous and 17 nonsynonymous sites have been found in *C. eurytheme* and 46 synonymous and 13 nonsynonymous sites in *M. cinxia* (Wheat *et al.* 2006; Orsini *et al.* 2009). Currently, *C. eurytheme* has the highest nucleotide diversity. In our research generating the partial coding region of *Pgi* in *E. postvittana*, 61 synonymous and nine nonsynonymous sites were found along the ~1,000 bp sequence. More variation than this in *E. postvittana Pgi* may be expected for the following reasons: 1) about 500 bp long coding region from exon 1 to 4 in *E. postvittana Pgi* has not been accessed yet, and half of nonsynonymous sites were found within that region in *C. eurytheme*. Therefore if the same pattern applies to *E. postvittana*, variation of *Pgi* in *E. postvittana* may be much more than in *C. eurytheme* and become the most diverse Lepidoptera taxon with respect to the *Pgi* gene. Besides, most exon sequences analyzed in this study are not completed yet and some field populations have yet to be successfully sequenced. 2) PGI polymorphism is considered to be maintained by habitat heterogeneity, which greatly increases the variation of *Pgi* in nature (Rank & Dahlhoff 2002; Haag *et al.* 2005; Zheng *et al.* 2009). *Epiphyas postvittana* is known to have highly variable life-history parameters and dispersal capacity among populations due to heterogeneous environmental conditions, such as spatial and temporal distribution of host plant species and unpredictable climatic factors, especially temperature (Geier & Briese 1981; Danthanarayana *et al*. 1995; Gu & Danthanarayana 2000a). Thus, it seems clear that *E. postvittana* may have more variation at the *Pgi* gene. Further sequencing of the two tails of the *Pgi* coding region in *E. postvittana* and full length cDNA sequences of different populations might support this prediction.

Also, because of high variation at the *Pgi* gene, if we have a better understanding of the variation in nucleotide sequences, more specific primers can be designed to

sequence each haplotype of PGI, which will enable the number of PGI genotypes (haplotypes) to be examined more accurately than by electrophoresis analysis (Wheat *et al.* 2010). By comparing nucleotide and amino acid changes and their association with differences among PGI genotypes along with the effects of each genotype on fitness traits, it may be possible to explore the fitness evolution in *E. postvittana* (Wheat *et al.* 2010).

### 5.4.3 Variation of introns of *Pgi* gene in *E. postvittana*

The coding region in the *Pgi* is composed of 12 exons that are separated by long introns (Wheat *et al*. 2006). Intron variation was significant in the *E. postvittana Pgi*, the same as that found in *C. eurytheme* (Wheat *et al.* 2006), where two alleles at the same locus have as many as ~790 bp differences in *E. postvittana*. How these intron variations contribute to the variation of the PGI needs to be investigated.

Unlike the PGI exons which are under natural selection (Watt *et al*. 1983; Watt 2003), introns are selectively neutral and considered to evolve faster (Bulmer 1987; He & Haymer 1997). Despite this, the function of introns is still not entirely understood. Introns are excised from mature RNA and not translated into proteins. The number and size of introns in *Pgi* are believed to have contributed to intra-genetic recombination by expanding the coding regions by about more than 10 times across its chromosomal locus (Wheat *et al.* 2006).

### 5.4.4 Variaton of *Pgi* gene and with respect to environmental factors

In the current study, reasonable variation in exons and significant variation in introns of *E. postvittana Pgi* gene was found. The genetic variation of *Pgi* in *E. postvittana* found here might be related to population fitness under variable environmental conditions, however, the scope of the current study meant that such correlation was not addressed.

The *Pgi* gene is known to be affected by environmental conditions such as temperature either over long-term or short-term time scales. For instance, an annual and seasonal temperature decline over a relatively long-term variation of climatic conditions from 1988 to 1996, resulted in an increase of PGI 1-1 in *Chrysomela aeneicollis* (Rank & Dahlhoff 2002). PGI allele frequency also can vary within a

single summer in response to temperature changes and the most recent environmental conditions affecting the genetic variation at PGI would be reflected through its phenotypic expression (Dahlhoff & Rank 2007; Rank *et al*. 2007). Not only temperature affects on PGI, other environmental factors such as the isolation of populations, the area of habitat patches, and newly established habitat also affect PGI diversity (Hanski & Saccheri 2006; Zheng *et al*. 2009).

This research focused only on *Pgi* genetic variation at the DNA level. No biological correlates were studied for the same *E. postvittana* populations. Further research on PGI genotypes, nucleotides and amino acid changes, together with biological characteristics study of *E. postvittana* under both field and laboratory conditions may provide a much better understanding of the potential invasive capacity of *E. postvittana* in a new habitat. Further discussion on this topic will be addressed in Chapter 6.

# Chapter 6 Discussion and future research

The risk of establishment of a potential global invasive species, *E. postvittana* has been assessed in this study. Not only has its potential global distribution been estimated but also its genetic variation with respect to the environment-affected *Pgi* gene has been determined. Three experimental hypotheses were explored: 1) Models based on climate have potential as tools to indicate the potential distribution of *E. postvittana*; 2) the potential capability of COI as molecular marker for accurate and rapid identification of *E. postvittana* intercepted at borders; and 3) the *Pgi* gene that has been linked to fitness traits and dispersal capability in other species, shows variation between *E. postvittana* populations, has potential as a target gene to assess fitness factors associated with invasibility of *E. postvittana*.

## 6.1 Modeled potential global distribution of *E. postvittana*

CLIMEX and several other ecological models, all based on characterizing the ecological niche of a species (Hutchinson 1956) predicted that there are areas around world that are climatically suitable for *E. postvittana* to survive and establish and therefore are high risk sites for its establishment.

### 6.1.1 CLIMEX prediction

Prediction of the potential distribution of *E. postvittana* by CLIMEX shows that the species has potential to establish and survive mainly in Central and South America, Western Europe, South Africa and South-eastern Asia. The distribution of *E. postvittana* is mainly restricted by cold stress in North America, Eastern Europe and North Asia, by heat stress in Central and South America, North Africa and south-west of Asia, and by dry stress in North Africa.

We calibrated the parameters in CLIMEX, using a restricted range of biological requirements obtained from laboratory studies (Danthanarayana 1975a; Geier & Briese 1981; Danthanarayana *et al.* 1995), such as developmental temperatures and degree-days, and its distribution in its native and long-established habitat. We avoided analyzing and extracting its environmental requirements from only its current invasive distribution ranges. Fortunately in our study, the relatively wide distribution of *E.*

*postvittana* in both its native and long-established range (CABI 1996), and laboratory data of its biological features (e.g. Danthanarayana 1975a) are available. Such information greatly increased the objectivity and reliability of the final predictions. For example, developmental temperatures from a lower threshold of 7$^{o}$C to and upper threshold of 31$^{o}$C (Danthanarayana 1975a; Geier & Briese 1981) excluded *E. postvittana* establishment from large areas. Degree-days predicted the number of generations that *E. postvittana* can complete in a certain climatic areas. The number of generations indicates a relative abundance of *E. postvittana*. However, there are problems with using such data. First, laboratory data may not be consistent with data derived from field conditions. The latter however, are extremely difficult, even impossible to obtain. Second, the distribution range of *E. postvittana* lacks a definite boundary between present and absent areas. Therefore, subjectivity of the resulting predictions is unavoidable.

Another issue highlighted in this study is that habitats can be modified by human activities, such as introduction of new host plants and cultivation. *Epiphyas postvittana* are thought to be gradually evolving with and adapting to these changes (Geier & Briese 1981). Micro-environmental conditions experienced by *E. postvittana* can be significantly different from natural field environmental conditions caused by human activities such as irrigation in dry areas, and the use of greenhouses in cold areas. Such modification may result in an area that is predicted to have a low probability of *E. postvittana* establishment becoming suitable for *E. postvittana* survival when environmental stress is reduced. Human activities, therefore, should be taken into account when certain regions are analyzed, such as California in this study. The assumption was proved in other ecological models in this study as well.

### 6.1.2 Statistical ecological models

Ten models of seven algorithms in openModeller all performed well, with high percentage accuracy and AUC values. However, calculation of these criteria may be inaccurate in openModeller with respect to the independence of the data used to evaluate their performance compared with similar studies (Elith *et al*. 2006; Tsoar *et al*. 2007; Giovanelli *et al*. 2010). Therefore, modeled map of the multi-model approach (Worner *et al*. 2010) was adopted for final result of ecological model prediction other than CLIMEX.

SVM in the new multi-model approach performed the better than other nine models in the same approach. The multi-model approach used 63.2% of occurrence points as training data and the 36.7% remaining data (or sub-set of the total) as testing data. Therefore, the resulting calculated criteria represented the performance of the models predicting new data. Moreover, it is very difficult to compare the performance of models, and using a single criterion such as accuracy or AUC may be biased if used alone (Elith *et al*. 2006). The method used in this study solved this problem by evaluating models using 10 commonly used criteria which maximally avoiding the bias of using different performance criteria.

The modeled potential distribution of *E. postvittana* by SVM using the multi-model approach predicted that *E. postvittana* can potentially survive and establish in North America mainly in the US, Central America, southern South America, Western Europe and southern Africa. The predicted distribution area is much wider than that of CLIMEX. That does not necessarily mean that CLIMEX is a better model than SVM, because there are no statistical criteria available in the CLIMEX program that allows it to be compared with other models.

### 6.1.3 Future study

One notable advantage of CLIMEX is that it is designed to include species' biological information such as abundance (approximated by degree-days) into modeling, which adds extra information for distribution prediction. Apparently, statistical ecological models are more convincing in our study, but these models may be improved if more information about species' biological requirements can be incorporated during modeling.

Other models in the multi-model approach performed not as well as SVM, however, the differences between SVM and some of them with respect to the performance criteria were not significant. The maps produced by different models in openModeller, indicating the potential distribution of *E. postvittana* have overlapping areas where there is high risk of *E. postvittana* establishment. Similarly, the models in the multi-model approach that have similar performance the, predicted distribution maps can be overlaid, and overlapping areas weighted by how many models predicted the same risk for specific areas. As such, predicted high risk areas of *E. postvittana*

establishment from several different well performing models can be easily discriminated visually (Worner *et al*. 2010). Details of this method and its feasibility could be tested in future study.

The population dynamics of *E. postvittana* is affected not only by climatic conditions but also by other factors such as host plant species, and their quality and availability (Danthanayana 1975a; Gu & Danthanarayana 1992b; Danthanarayana *et al*. 1995), abundance of natural enemies (Danthanarayana 1983; Suckling *et al*. 1987), and their population density (Danthanarayana *et al*. 1982). These factors should be considered along with climate conditions to produce a more comprehensive risk assessment of *E. postvittana* especially with respect to its potential distribution.

### 6.2. Accurate identification of all stages of *E. postvittana* at borders

Accurate identification of *E. postvittana* at the border would significantly reduce the risk of *E. postvittana* entrance into a country. Due to the high morphological variation of *E. postvittana*, developing molecular techniques is crucial for accurate and rapid identification for quarantine purposes.

In this study, 26 *E. postvittana* samples of four populations from New Zealand were correctly identified using COI barcodes, although variation between *E. postvittana* and several closely related species in *Epiphyas* genus is less than 3%, but still greater than intraspecies variation of *E. postvittana*. Variation within and between species of *E. postvittana* and other 173 tortricid species in BOLD has been calculated and it is indicated that interspecies variation between *E. postvittana* and these tortricid species is much greater (range from 7% to 16%) than within species variation (less than 1% in this study).

Sequencing the COI gene of related species can be expensive when a large number of samples has been found from surveillance and it requires a high standard of facilities, which may not be affordable at the border of some countries. Further identification using tools that are more time and cost effective, and easier to use would be another research focus.

### 6.3 Variation of *Pgi* gene in *E. postvittana*

The variation in both coding region and introns of the *Pgi* gene in *E. postvittana* is relatively high, despite that only part of the gene has been successfully sequenced. Genetic variation may reflect the adaptability of species to new environmental conditions (Booy *et al*. 2000). The variation found in *E. postvittana* may indicate that this species has the potential to survive and establish in highly variable environmental conditions. Clarification of this observation would add more information to the distribution prediction of this species.

*Epiphyas postvittana* prefers a warm and wet climate. The survivable temperature of *E. postvittana* ranges from 7-7.5 $^{o}$C to 30.7-33 $^{o}$C according to laboratory data, with an optimal developmental temperature around 20 $^{o}$C (Danthanarayana 1975a; Geier & Briese 1981; Danthanarayana *et al.* 1995). Population fitness and flight activities of *E. postvittana* are significantly affected by environmental factors, especially temperature (e.g. Danthanarayana 1975b; Danthanarayana 1976b; Gu and Danthanarayana 1992b; Danthanarayana *et al*.1995). Additionally, quantitative genetic analysis on fitness traits of *E. postvittana* showed that these fitness traits were correlated with climatic conditions such as temperature (e.g. Gu & Danthanarayana 2000a, b). *Epiphyas postvittana* populations were collected from northern (Hawkes Bay), central (Lincoln, Christchurch) and southern (Clyde) New Zealand with one colony population (from Auckland and raised under artificial and constant climate conditions). Climatic conditions especially the temperature and moisture conditions are very different in these areas. For example, average temperatures in Hawkes Bay are relatively warmer than in Lincoln (Christchurch) and are the lowest in Clyde (NIWA 2009). Significant variation of *Pgi* gene in *E. postvittana* might be due to the environmental differences such as climate among various geographic populations. Nevertheless, populations are more abundant in Canterbury than in Hawkes Bay and are least abundant in Central Otago (Suckling *et al.* 1998). Further study on potential *Pgi* related fitness traits of populations may be capable to reveal insight relationship of variation in *Pgi* gene and population abundance of the species.

Different genotypes of PGI (allozymes) and frequency of genotypes, which were considered a response to environmental conditions, have effects on different fitness traits. Alleles with faster mobility (close to anode) appear more active in low temperatures and those with less mobility are more tolerant to high temperatures

(Watt 1977; Riddoch 1993; Watt 2003). Homozygote genotypes consisting of alleles with either high or low mobility appear more resistant to extreme conditions while heterozygote genotypes appear more variable with high adaptation in their fitness traits. Heterozygote advantage happens when a species is in its optimal environmental conditions, while the homozygote occurs more often when a species experiences extreme environmental conditions, such as cold or heat. For example, life-history traits, such as larval and pupal development, pupal mass, and growth rate of the butterfly *Lycaena tityrus*, have significant advantages for individuals with PGI genotype 2-3 (Karl *et al* 2008; Karl *et al.* 2009). However, genotype PGI 2-2 was found to be the dominant genotype (90%) among high-altitude populations with significantly more resistant characteristics to cold stress (Karl *et al* 2008; Karl *et al.* 2009). Apparently, to achieve high fitness in a cold environment, the butterfly has to balance survival rate and high population development rate. There may be a balance between heterozygote advantage with respect of life-history traits and thermal tolerance in PGI. A similar pattern has been observed in the leaf beetle *Chrysomela aeneicollis*, where among three common genotypes, PGI 1-1, 1-4 and 4-4, PGI 1-1 was the predominant genotype in cold north habitats of the beetle, PGI 4-4 in warm south habitats and PGI 1-4 was found from an area somewhere between (Dahlohff & Rank 2000). In contrast, larval development rate in the field is slowest for PGI 1-1 and PGI 4-4 showed high larval growth rate at 27 $^{o}$C (McMillan *et al.* 2005). PGI genotype 1-1 also induced higher expression of Hsp70 in PGI 1-1 predominant individuals, which contributed to population survival in stressful conditions (Dahlhoff & Rank 2000; Rank & Dahlhoff 2002).

With respect to the PGI genotype-thermal tolerance relationship mentioned above, we may predict that populations of *E. postvittana* from a cold region such as Central Otago might exhibit a predominance of PGI homozygote genotypes with relatively fast kinetic property and more resistant to cold stress. Similarly warmer regions such as Auckland populations would be dominated by homozygote genotypes with slower kinetic property, which are resistant to heat stress; while populations from areas such as Canterbury and Hawkes Bay would have more variable genotypes with heterozygote advantage occurring in these populations for faster population growth as well as highly diverse PGI polymorphism. Whether this variation in PGI genotypes among populations from different climatic conditions would affect the invasibility of

the moth needs to be further investigated.

High polymorphism in PGI may be partly due to environmental heterogeneity. Environmental conditions also affect PGI genotype composition, which changes according to changes in an environmental variable such as temperature either over long-term or short-term time scales. For instance, annual and seasonal temperature decline representing relatively long-term variation of climatic conditions from 1988 to 1996, resulted in an increase of PGI 1-1 in *Chrysomela aeneicollis* (Rank & Dahlhoff 2002). PGI allele frequency also can vary within a single summer in response to temperature changes and the most recent environmental conditions affecting the genetic variation at PGI would be reflected through its phenotype expression (Dahlhoff & Rank 2007; Rank *et al*. 2007). Not only temperature affects PGI, other environmental factors such as the isolation of populations, the area of habitat patches, and newly established habitat also affect PGI diversity (Hanski & Saccheri 2006; Zheng *et al*. 2009). Clearly, the adaptation of *E. postvittana* to new environmental conditions is likely to depend on its phenotypic plasticity to adapt to short-term environmental change, and genetic diversity and genotypic response to long-term environmental changes (Dahlhoff & Rank 2007). The climatic conditions in Australia and New Zealand are very variable, which results in spatially and temporally patchy distributions of host plants. This would result in *Pgi* in *E. postvittana* remaining highly variable through natural selection, allowing the species to establish relatively easily in a new habitat.

It is interesting to note that fecundity and flight capacity were positively related in adults of both *C. eurytheme* and *M. cinxia*, which individuals with a higher flight metabolic rate also have higher fecundity (Watt *et al.* 1983; Watt 1992; Saastamoinen 2007; Saatamoinen & Hanski 2008). The female adults lay eggs either as singles or in a clutch, and they fly frequently to find host plants for laying eggs, which can make their flight capability highly related to their fecundity (Watt *et al.* 1983; Haag *et al.* 2005). Individuals with *Pgi*-f allele in *M. cinxia* laid on average 32% larger clutches with higher flight metabolic performance than in *Pgi*-non-f individuals (Haag *et al.* 2005; Saastamoinen & Hanski 2008). However, in *E. postvittana*, a trade-off appears to exist between life history traits and flight capability, in other words, slow development rate, small body size and low fecundity during unfavorable

environmental conditions produce moths with strong flight capacity (Gu & Danthanarayana 1992a). The reason(s) why there is such difference between *E. postvittana* and other Lepidoptera species in their flight acitivity and fecundity relationship require(s) further investigation.

*Pgi* variation seems to affect the fitness of the immature larvae of the beetle *Chrysomela aeneicollis*, the butterflies *L. tityrus* and *M. cinxia* (Rank *et al.* 2007; Karl *et al.* 2008; Orsini 2009), but not in the larvae of the butterfly *C. eurytheme* (Orsini 2009). When experiencing unfavorable environmental conditions, the larvae of *E. postvittana* start to prolong their development time and produce adults with high dispersal capability (Gu & Danthanarayana 1992a). This might indicate that the *Pgi* variation in *E. postvittana* may also be obvious in the immature stages. When sampling different populations of *E. postvittana*, potential *Pgi* gene variation in larvae should be considered, together with the effects of long-term (annual and seasonal) and short-term (daily) environmental changes on the *Pgi* genotype composition and frequency.

This research focuses only on *Pgi* genetic variation at the DNA level. No biological correlates were studied on the same *E. postvittana* populations. Further research on PGI genotypes, nucleotides and amino acid changes, together with the study of *E. postvittana* biological characteristics under both field and laboratory conditions may provide a much better understanding of the potential invasive capacity of *E. postvittana* in a new habitat.

### 6.4 Implications of this study for biosecurity

The current study has added to the available and valuable information on invasibility of *E. postvittana* to assist further risk assessment of the moth. Although the actual crop damage caused by the larval feeding is not considered serious, the main economic loss is high when based on protection input, entrance restriction on host plant products and the loss of potential market (Bailey *et al*. 1996; Lo & Murrell 2000). Due to the current limited distribution and wide host plant range, keeping the species from crossing the border is very important for those countries with high risk of *E. postvittana* establishment. This is especially so when it has been shown that eradication of the moth is fairly difficult (USDA 2007a).

*Epiphyas postvittana* is not an insect with strong dispersal capability, and its flight activities are mainly related to reproduction (Gu & Danthanarayana 1990b; Gu & Danthanarayana 1992a). The spread of *E. postvittana* in California is considered mainly due to the transportation of nursery host plants among counties (Varela *et al.* 2008). Long distance spread of *E. postvittana* is believed to occur mainly through trading of fruit, vegetable and ornamental host plants. To prevent the further spread of the moth, strategies such as develop new pest management techniques and post-harvest treatment before exportation of plants/plant products, should be put in place, ideally without affecting the economic activities. Rapid and correct identification of each stage of the species at borders would reduce the risk posed by *E. postvittana*. Moreover, proper biosecurity measures would be the most effective at preventing the further spreading of the species to other LBAM-free countries, and should be based on scientific information such as potential distribution and genetic variation related to fitness traits of the species in current study.

This study has identified two important factors that could contribute to a more accurate assessment of risk of invasion by a pest species. The techniques applied here should provide biosecurity agencies such as Biosecurity New Zealand with an opportunity to be better equipped to determine risk management strategies towards minimum economic loss by such potential invasive species as *E. postvittana*.

## References


APHIS (Animal and Plant health inspection Service, USDA) factsheet 2008. Questions and answers: light brown apple moth. Available online: http://www.aphis.usda.gov/publications/plant_health/content/printable_version/faq_lbam.pdf.

APHIS (Animal and Plant health inspection Service, USDA) 2009. Weekly situation report: light brown apple moth (LBAM). Available online: http://www.aphis.usda.gov/plant_health/plant_pest_info/lba_moth/downloads/updates/2009/report12-2-09.pdf.

Armstrong KF & Ball SL. 2005. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Science*. 360: 1813-1823.

Atkinson D. 1994. Temperature and organism size – a biological law for ectotherms? *Advances in Ecological Research*. 25: 1-58.

Avise JC & Kitto GB. 1973. Phosphoglucose isomerase gene duplication in the bony fishes: an evolutionary history. *Biochemecal Genetics*. 8(2): 113-132.

Bailey P, Baker G & Caon G. 1996. Field efficacy and persistence of *Bacillus thuringiensis* var. *kurstaki* against *Epiphyas postvittana* (Walker) (Lepidoptera: Tortricidae) in relation to larval behaviour on grapevine leaves. *Australian Journal of Entomology*. 35: 297-302.

Bailey PT, Ferguson KL, McMahon R & Wicks TJ. 1997. Transmission of *Botrytis cinerea* by lightbrown apple moth larvae on grapes. *Australian Journal of Grape and Wine Research*. 3: 90-94.

Baker DRB. 1968. Notes on *Epiphyas* (= *Austrotortrix*) *postvittana* (Walker), (Lep. Tortricidae). *Entomologist's Gazette.* 19: 167-72.

Baker RHA, Sansford CE, Jarvis CH, Cannon RJC, Macleod A & Walters KFA. 2000. The role of climate mapping in predicting the potential geographical distribution of non-indigenous pests under current and future climates. *Agriculture, Ecosystems and Environment*. 82: 57-71.

Ball SL & Armstrong KF. 2006. DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera: Lymantriidae). *Canadian Journal of Forest Research*. 36: 337-350.

Barr N. 2007. ID of light brown apple moths using DNA barcodes. USDA APHIS PPQ CPHST News. Volume IV (3): 1-2. Available online: http://www.aphis.usda.gov/plant_health/cphst/downloads/newsletterfall2007.pdf.

Barr NB, Ledezma LA, Vasquez JD, Epstein M, Kerr PH, Kinnee S, Sage O & Gilligan TM. 2009. Molecular identification of the light brown apple moth (Lepidoptera: Tortricidae) in California using a polymerase chain reaction assay of the internal transcribed spacer 2 locus. *Journal Economic Entomology.* 102(6): 2333-2342.

Behura SK. 2006. Molecular marker systems in insects: current trends and future avenues. *Molecular Ecology*. 15: 3087-3113.

Booy G, Hendriks RJJ, Smulders MJM, Van Groenendael JM & Vosman B. 2000. Genetic

diversity and the survival of populations. *Plant Biology.* 2: 379-395.

Bradley JD. 1956. A new genus for *Tortrx postvittana* (Walker) and certain other Australian and New Zealand species (Lepidoptera: Tortricidae). *Bulletin of Entomological Research*. 47: 101-105.

Brockerhoff EG, Jactel H, Leckie AC & Suckling DM. 2002. Species composition and abundance of leafrollers in a Canterbury pine plantation. *New Zealand Plant Protection*. 55: 85-89.

Buchanan GA. 1977. The seasonal abundance and control of light brown apple moth, *Epiphyas postvittana* (Walker) (Lepidoptera: Tortricdae), on grapevines in Victoria. *Australian Journal of Agricultural Research.* 28: 125-32.

Bulmer M. 1987. A statistical analysis of nucleotide sequnces of introns and exons in human genes. *Molecular Biology and Evolution.*. 4: 395-405.

CABI 1992. Distribution maps of pests, *Epiphyas postvittana* (Walker). Map No. 82 (2nd revision). Available online: http://www.cabi.org/dmpp/default.aspx?LoadModule=Review&ReviewID=101679&site=164&page=1155.

Chefaoui RM & Lobo JM. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*. 210: 478-486.

Common, IFB. 1961. The generic position of the Australian light-brown apple moth (Lepidoptera: Tortricidae). *Proceedings of the Linnean Society of New South Wales.* 84: 177-82.

Cortes C & Vapnik V. 1995. Support-vector Networks. *Machine Learning*. 20: 273-297.

Dahlhoff EP & Rank NE. 2000. Functional and physiological consequences of genetic variation at phosphoglucose isomerase: heat shock protein expression is related to enzyme genotype in a mountain beetle. *Proceedings of the National Academy of Sciences USA*. 97: 10056-10061.

Dahlhoff EP & Rank NE. 2007. The role of stress proteins in responses of a montane willow leaf beetle to environmental temperature variation. *Journal of Biosciences.* 32: 477-488.

Danthanarayana W. 1975a. The bionomics, distribution and host range of the light brown apple moth, *Epiphyas postvittana* (Walk.) (Tortricidae). *Australian Journal of Zoology.* 23: 419-437.

Danthanarayana W. 1975b. Factors determining variation in fecundity of the light brown apple moth, *Epiphyas postvittana* (Walker) (Tortricidae). *Australian Journal of Zoology.* 23: 439-451.

Danthanarayana W. 1976a. Diel and lunar flight periodicities in the light brown apple moth, *Epiphyas postvittana* (Walker) (Tortricidae) and their possible adaptive significance. *Australian Journal of Zoology.* 24: 65-73.

Danthanarayana W. 1976b. Flight thresholds and seasonal variations in flight activity of the light-brown apple moth, *Epiphyas postvittana* (Walk.) (Tortricidae), in Victoria, Australia. *Oecologia* (Berl.). 23: 271-282.

Danthanarayana W. 1976c. Environmental cue size variation in light brown apple moth, *Epiphyas postvittana* (Walk.) (Tortricidae), and its adaptive value. *Oecologia* (Berl.) 26: 121-132.

Danthanarayana W. 1983. Population ecology of the light brown apple moth, *Epiphyas postvittana* (Lepidoptera: Tortricdae). *Journal of Animal Ecology*. 52: 1-33.

Danthanarayana W. Hanilton JG & Khoul SP. 1982. Low-density larval crowding in the light brown apple moth, *Epiphyas postvittna* and its ecological significance. *Entomologia Experimentalis et Applicata.* 31: 353-358.

Danthanarayana W & Gu H. 1992. Influence of environmental conditions on flight duration of *Epiphyas postvittna* (Walker) (Lepidoptera: Tortricidae). *Austrlian Journal of Zoology*. 40: 477-484.

Danthanarayana W & Gu H. 2000. Variation in life history traits and flight capacity among populations of the light brown apple moth *Epiphyas postvittna*, (Lepidoptera: Tortricidae). *Austral Ecology*. 25: 571-579.

Danthanarayana W, Gu H & Ashley S. 1995. Population growth potential of *Epiphyas postvittna*, the light brown apple moth (Lepidoptera: Tortricidae) in relation to diet temperature and climate. *Australian Journal of Zoology*. 43: 381-394.

Muňoz MES, Giovanni R, Siqueira MF, Sutton T, Brewer P, Pereira RS, Canhos DAL & Canhos VP. 2009. openModeller: a generic approach to species' potential distribution modelling". *GeoInformatica*. DOI: 10.1007/s10707-009-0090-7.

Dugdale JS, Gleeson D, Clunie LH & Holder PW. 2005. *A diagnostic guide to Tortricidae encountered in field surveys & quarantine inspections in New Zealand: morphological and molecular characters*. Minisry of Agriculture and Forestry.

Dukes JS & Mooney HA. 1999. Does global change increase the success of biological invaders? *Trends in Ecology and Ecolution.*14(4): 135-139.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijimans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS & Zimmermann NE. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. 29: 129-151.

Elith J & Leathwick JR. 2009. species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*. 40: 677-697.

Engler R, Guisan A & Rechsteiner L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*. 41: 263-274.

Evans JW. 1952. The injurious insects of the British Commonwealth (except the British Isles, India and Pakistan). Commonwealth Institute of Entomology: London.

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. *Molecular Marine Biology and Biotechnology.* 3: 294-299.

Fountain MT and Cross JV. 2007. The light brown apple moth, *Epiphas postvittana* (Walker)

(Lepidoptera: Tortricidae), in UK pome and stone fruit orchards. *IOBC*. 30 (4): 51-60.

GBIF (Global Biodiversity Information Facility). Accessed through GBIF data portal, Lepidoptera Observations from UK and New Zealand. http://data.gbif.org/species/14157366/. Latest accessed on 20-Jan-2010.

Geier PW & Briese DT. 1977. Predominantly female progeny in the light-brown apple moth. *Search*. 8: 83-85.

Geier PW & Briese DT. 1978. The light-brown apple moth, *Epiphyas postvittana* (Walker): 2. Unevern sex ratios and a condition contributiong to them in the field. *Australian Journal of Ecology*. 3: 467-488.

Geier PW & Briese DT. 1979. The light-brown apple moth, *Epiphyas postvittana* (Walker): 3. diffenrences in susceptibility to a nuclear polyhedrosis virus. *Australian Journal of Ecology*. 4: 187-194.

Geier PW & Briese DT. 1980a. The light-brown apple moth, *Epiphyas postvittana* (Walker): 4. Studies on population dynamics and injuriousness to apples in the Australian Capital territory. A*ustralian Journal of Ecology.* 5: 63-93.

Geier PW & Briese DT. 1980b. The light-brown apple moth, *Epiphyas postvittana* (Walker): 5. Variability of demographic characteristics in field populations of southeastern Australia. A*ustralian Journal of Ecology.*5: 135-142.

Geier PW & Briese DT. 1981. The light-brown apple moth, *Epiphyas postvittana* (Walker): a native leafroller fostered by European settlement. In. *The Ecology of pests: some Australian case histories*. Eds: Kitching PL & Jones RE. CSIRO. pp130-155.

Geier PW & Springett BP. 1976. Population characteristics of Australian leafrollers (*Epiphyas* spp., Lepidoptera) infesting orchards. *Australian Journal of Ecology*. 1: 127-144.

Gevrey M & Worner SP. 2006. Prediction of global distribution of insect pest species iin relation to climate by using an ecological informatics method. *Journal of Economic Entomology.* 99 (3): 979-986.

Giovanelli JGR, de Siqueira MF, Haddad CFB & Alexandrino J. 2010. Modeling a spatially restricted distribution in the Neotropics: how the size of calibration area affects the performance of five presence-only methods. *Ecological Modelling*. 221: 215-224.

Goldson SL, Rowarth JS & Caradus JR. 2005. The impact of invasive invertebrate pests in pastoral agriculture: a review. *New Zealand Journal of Agricultural Research*. 48: 401-415.

Gu H & Danthanarayana W. 1990a. The role of availability of food and water to the adult *Epiphyas postvittana*, the light brown apple moth, in its reproductive performance. *Entomolgia experimentalis et Applicata*. 54: 101-108.

Gu H & Danthanarayana W. 1990b. Age-related flight and reproductive performance of the light brown apple moth, *Epiphyas postvittana*. *Entomolgia experimentalis et Applicata*. 54: 09-115.

Gu H & Danthanarayana W. 1992a. Quantitative genetic analysis of dispersal in *Epiphyas postvittana*. I. Genetic variation in flight capapcity. *Heredity*. 68: 53-60.

Gu H & Danthanarayana W. 1992b. Influence of laval rearing conditions on the body size and flight capacity of *Epiphyas postvittana* Moths. *Australian Journal of Zoology.* 40: 573-581.

Gu H & Danthanarayana W. 1992c. Quantitative genetic analysis of dispersal in *Epiphyas postvittana*. II. Genetic covariations between flight capacity and life-history traits. *Heredity*. 68: 61-69.

Gu H & Danthanarayana W. 2000a. Variations in life history traits and flight capacity among populations of the light brown apple moth *Epiphyas postvittana* (Walker) (Lepidoptera: Tortricidae). *Autral Ecology. 25*: 571-579.

Gu H & Danthanarayana W. 2000b. Genetic variation in the life-history traits of *Epiphyas postvittana*: population structure and local adaptation. *Austral Ecology*. 25: 394-401.

Haag CR, Saastamoinen M, Marden JH & Ilkka Hanski. 2005. A candidate locus for variation in dispersal rate in a butterfly metapopulaiton. *Proceedings of the Royal Society B.* 272: 2449-2456.

Hanley JA & McNeil BJ. 1982. The meaning and use of the area under a receiver operating Characteristic (ROC) Curve. *Radiology*. 143: 29-36.

Hanski I & Saccheri I. 2006. Molecular-level variation affects population growth in a butterly metapopulation. *PLoS Biology*. 4(5): 719-726.

He M & Haymer DS. 1997. Polymorphic intron sequences detected within and between populations of the oriental fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America*. 90(6): 825-831.

Hebert PDN, Cywinska A, Ball SL & de Waard, JR. 2003. Barcoding animal life: cytochrome C oxidase 1 divergences among closely related species. *Proceedings of the Royal Society B. Biological Science.* 270: 596-599

Hortnet (by HortResearch) 2000. http://www.hortnet.co.nz/key/keys/info/lifecycl/lba-life.htm

Howarth FG. 2000. Non-target effects of biological control agents. In *Biological Control: Measures of Success.* Eds: by Gurr, G. & Wratten, S.pp369-403

Hutchinson GE. 1957. Concluding remarks. *Cold Spring harbor Symposia on Quantitative Biology*. 22: 415-427.

IPPC (the International Plant Protection convention). 2004. Pest risk analysis for quarantine pests including analysis of environmental risks and living modified organisms. International Standards for Phytosanitary measures. ISPM No. 11

IPPC, 2009. International standard for phytosanitary measures No. 5: Glossary of phytosanitary terms. FAO-UNEP. ISPM No. 5. Rome, Italy. Available online at: https://www.ippc.int/servlet/BinaryDownloaderServlet/184195_E.pd.pdf?filename=1241701271714_ISPM05_2009_E.pdf&refID=184195

Jarman SN & Elliott NG. 2000. DNA evidence for morphological and cryptic Cenozoic speciations in the Anaspididae, "living fossils" from the Triassic. *Journal of Evolutionary Biology.* 13: 624-633.

Kadmon R, Farber O & Danin A. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*. 13(3): 853-867.

Karl I, Schmitt T & Fischer K. 2008. Phosphoglucose isomerase genotype affects life-history traits and cold stress resitstance in a Copper butterfly. *Functional Ecology*. 22: 887-894.

Karl I, Schmitt T & Fischer K. 2009. Genetic differentiation between alpine and lowland populations of a butterfly is related to PGI enzyme genotype. *Ecogrphy*. 32: 488-496.

Kriticos DJ, Sutherst RW, Brown JR, Adkins SW & Maywald GF. 2003. Climate change and the potential distribution of an invasive alien plant: Acacia nilotica ssp. indica in *Australia. Journal of Applied Ecology*. 40: 111-124.

Kumar S, Nei M, Dudley J & Tamura K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*. 9: 299-306.

Langhoff P, Authier A, Buckley TR, Dugdale JS, Rodrigo A & Newcomb RD. 2009. *Molecular Ecology Resources*. 9: 691-698.

Leday G. 2008. Modelling invasive species distributions: a comparative approach. Work Placement Report. Lincoln University, New Zealand.

Levine JM & D'Antonio CM. 2003. Forecasting biological invasions with increasing international trade. *Conservation Biology*. 17(1): 322-326.

Liu Y, Mitsukawa N, Oosumi T & Whittier RF. 1995. Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *The Plant Journal*. 8(3): 457-463.

Lo PL & Murrell VC. 2000. Time of Leafroller infestation and effect on yield in grapes. *New Zealand Plant Protection*. 53: 173-178.

LoBo JM. 2008. More complex distribution models or more representative data? *Biodiversity Informatics*. 5: 14-19.

McMillan DM, Fearnley SL, Rank NE & Dahlhoff EP. 2005. Natural temperature variation affects larval survival development and Hsp70 expression in a leaf beetle. *Functional Ecology*. 19: 844-852.

Mo J, Glover M, Munro S & Beattie GAC. 2006a. Development of *Epiphyas postvittana* (Lepidoptera: Tortricidae) on leaves and fruit of orange trees. *Journal of Economic Entomology*. 99 (4): 1321-1326.

Mullins ED, Chen X, Romaine P, Raina R, Geiser DM & Kang S. 2001. *Agrobacterium*-Mediated Transformation of *Fusarium oxysporum*: an efficient tool for insertional mutagenesis and gene transfer. *The American Phytopathological Society*. 91(2):173-180.

Musolin DL. 2007. Insects in a warmer world: ecological, physiological and life-history responses of true bugs (Heteroptera) to climate change. *Global Change Biology*. 13: 1565~1585.

Nativi S, Mazzetti P, Saarenmaa H, Kerr J, Tuama EO. 2009. Biodiversity and climate change use scenarios framework for the GEOSS interoperability pilot process. *Ecological Informatics*. 4: 23-33.

Nevo E & Beeles A. 1988. genetic parallelism of protein polymorphism in nature: ecological test of the neutral theory of molecular evolution. *Biological Journal of the Linnean Society*. 35: 229-245.

NIWA (National Institute of Water & Atmospheric Research). 2009. http://www.niwa.co.nz/

Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nature reviews. Genetics.* 6: 119-127.

Orsini L, Wheat CW, Haag CR, Kvist J, Frilander MJ & Hanski I. 2009. Fitness differences associated with *Pgi* SNP genotypes in the Glanville fritillary butterfly (*Melitaea cinxia*). *Journal of Evolutionary Biology*. 22: 367-375.

Peacock L & Worner SP. 2008. Biological and ecological traits that assist establishment of alien invasive insects. *New Zealand Plant Protection*. 61: 1-7.

Phillips DM & Schapire RE. 2003. A maximum entropy approach to species distribution modeling. *Physics Review Letters.* 91.

Pimentel D, Zuniga R & Morrison D. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*. 52: 273-288.

Porter J. 2001. Range expansion in the light brown apple moth *Epiphyas postvittana* (Walk.). *Atropos*. 14: 1-4.

Poutsma J, Loomans AJM, Aukema B & Heijerman T. 2008. Predicting the potential geographical distribution of the harlequin ladbird, *Harmonia axyridis*, using the CLIMEX model. *BioControl.* 53: 103-125.

Powell JA & Common IFB. 1985. Oviposition patterns and egg characteristics of Australian Tortricine Moths (Lepidoptera: Tortricidae). *Australian Journal of Zoology.* 33: 179-216.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. URL http://www.R-project.org .

Rank NE & Dahlhoff EP. 2002. Allele frequency shifts in response to climate change and physiological consequences of allozyme variation in a montane insect. *Evolution*. 56: 2278-2289.

Rank NE, Bruce DA, McMilan DM, Barclay C & Dahlhoff EP. 2007. Phosphoglucose isomerase genotype affects running speed and heat shock protein expression after exposure to extreme temperatures in a montane willow beetle. *Journal of Experimental Biology*. 210: 750-764.

Ratnasingham S & Hebert PDN. 2007. BOLD: the barcode of life data system (www.harcodinglife.org). *Molecular Ecology Notes*. 7: 355-364.

Riddoch BJ. 1993. The adaptive significance of electrophoretic mobility in phosphoglucose isomerase (PGI). *Biological Journal of the Linnean Society.* 50: 1-17.

Robertson JL, Armstrong KF, Suckling DM & Preisler HK. 1990. Effects of host plants on the Toxicity of azinphosmethyl susceptible and resistant light brown apple moth (Lepidoptera:

Tortricidae). *Journal of Economic Entomology.* 83: 2124-2129.

Robertson MP, Kriticos DJ & Zachariades C. 2008. Climate matching techniques to narrow the search for biological control agents. *Biological Control*. 46: 442-452.

Saastamoinen M. 2007. Life-history, genotypic, and environmental correlateds of clutch size in the Glanville fritillary butterfly. *Ecological Entomology*. 32: 235-242.

Saastamoinen M & Hanski I. 2008. Genotypic and environmental effects on flight activity and oviposition in the Galnville fritillary butterfly. *American Naturalist*. 171: 701-712.

Saastamoinen M, Ikonen S & Hanski I. 2009. Significant effects of Pgi genotype and body reserves on lifespan in the Glanville fritillary butterfly. *Proceedings of the Royal Society B.* 276: 1313-1322.

Santana FS, de siqueira MF, Saraiva AM, & Correa PLP. 2008. A reference business process for ecological niche modelling. *Ecological Informatics*. 3: 75-86.

Senaratne KAD, Palmer WA & Sutherst RW. 2006. Use of CLIMEX modeling to identify prospective areas for exploration to find new biological control agents for prickly acacia. *Australian Journal of Entomology.* 45: 298-302.

Shaw PW, Cruickshank VM & Suckling DM. 1994. Geographic changes in leafroller species in Nelson orchards. *New Zealand Journal of Zoology*. 21: 289-294.

Stephens, A.E.A., Kriticos, D.J. & Leriche, A. 2007. The current and future potential geographical distribution of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *Bulletin of Entomological Research.* 97: 369-378.

Stockwell D & Peters D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*. 13(2): 143-158.

Stockwell DRB & Peterson AT. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*. 148: 1-13.

Suckling DM & Brockerhoff EG. 1999. Control of light brown apple moth (Lepidoptera: Tortricidae) using an attracticide. *Journal of Economic Entomology*. 92: 367-72.

Suckling DM & Brockerhoff EG. 2010. Invasion biology, ecology and management of the light brown apple moth (Tortricidae). *Annual Review of Entomology*. 55: 285-306.

Suckling DM, Brunner JF, Burnip GM & Walker JTS. 1994. Dispersal of *Epiphyas postvittana* (Walker) and *Planotortrix octo* Dugdale (Lepidoptera: Tortricidae) at a Canterbury, New Zealand orchard. *New Zealand Journal of Crop and Horticultrual Science*. 22: 225-234.

Suckling DM, Burnip GM, Gibb AR, Daly JM & Armstrong KF. 2001. Plant and host effects on the leafroller parasitoid *Dolichogenidia tasmanica*. *Entomolgia Experimentalis et Applicata*. 100: 253-260.

Suckling DM, Burnip GM, Walker JTS, Shaw PW, McLaren GF, Howard CR, Lo P, White V & Fraser J. 1998. Abundance of leafrollers and their parasitoids on selected host plants in New Zealand. *New Zealand Journal of Crop and Horticultural Science*.26:193-203.

Suckling DM, Rogers DJ, Shaw PW, Wearing CH, Penman DR & Chapman RB. 1987. Monitoring Azinphosmethyl resistance in the light brown apple moth (Lepidoptera, Tortricidae) in New Zealand. *Journal of Economic Entomology*. 80(4): 733-738.

Sutherst RW & Maywald GR. 1991. Climate modeling and pest establishment. *Plant Protection Quarterly*. 6: 3-7.

Sutherst RW & Maywald GR. 2005. A climate model of the red imported fire ant, Solenopsis invicta Buren (hymenoptera: Formicidae): implications for invasion of new regions, particularly Oceania. *Environmental Entomology*. 34(2): 317-335.

Sutherst RW, Maywald GF & Kriticos D. 2007. *CLIMEX version 3 user's guide*. CSIRO. Australia.

Takahashi G. 2002. Notes on the light brown apple moth *Epiphyas postvittana* (Walker) (Lepidoptera: Tortricidae) intercepted at import plant quarantine of Narita Airport in Japan. *Research Bulletin of the Plant Protection Service, Japan.* 38: 99~103.

Tauber MJ, Tauber CA & Masaki S. 1986. *Seasonal adaptation of insects*. Oxford University Press, Oxford.

Taylor LR. 1974. Insect immigration, flight periodicity and the boundary layer. *Journal of Animal Ecology*. 43: 225-238.

Thomas WP. 1974. Light brown apple moth. Department of Scientific and Industrial Research, Wellington, New Zealand. Information Series No. 105/3.

Thomas WP. 1989. *Epiphyas postvittana* (Walker), lightbrown apple moth (Lepidoptera: Tortricidae). In *A Review of Biological Control of Invertebrate Pest and Weeds In New Zealand 1974 to 1987*, eds. P Cameron, R Hill, J Bain, WP Thomas, pp. 187-96. Wallingford, UK/Auckland, NZ: Cab int. DSIR Entomol. Div.

Tomkins AR, Penman DR & Chapman RB. 1989. Effect of temperature and host plant on development of three species of leafrollers (Lepidoptera: Trotricidae). *New Zealand Entomologist*. 12: 48-54.

Tsoar A, Allouche O, Steinitz O, Rotem D & Kadmon R. 2007. A comparative evaluation of presence-only methods for modeling species distribution. *Diversity and Distributions*. 13: 397-405.

USDA (United States Department of Agricultural) 2007a. Eradication of isolated populations of light brown apple moth in California. Available online: http://www.aphis.usda.gov/plant_health/ea/downloads/lbam-ea-07-18-07.pdf.

USDA 2007b. Economic analysis: risk to U.S. apple, grape, organge and pear production from the light brown aplle moth, *Epiphyas postvittana* (Walker).

USDA 2009a. Light brown apple moth distribution& detection levels in California. Available online: http://www.aphis.usda.gov/plant_health/plant_pest_info/lba_moth/downloads/maps/dist-detection-11-2009.pdf.

USDA 2009b. Review of the U.S. Department of Agriculture's Animal and Plant Health

Inspection ServiceResponse to Petitions to Reclassify the Light Brown Apple Moth as a Non-Actionable Pest. http://www.nap.edu/catalog/12762.html.

VanDerVal J, Shoo LP, Graham C & Williams SE. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling*. 220: 589-594.

Varela LG, Johnson MW, Strand L, Wilen CA & Pickel C. 2008. Light brown apple moth's arrival in California worries commodity groups. *California Agriculture*. 62(2) Http://Californiaagriculture.ucop.edu.

Venette RC, Davis EE, DaCosta M, Heisler H & Larson M. 2003. Mini-risk assessment: light brown apple moth, *Epiphyas postvittana* (Walker) (Lepidoptera: Trotricidae). Rep USDA-APHIS.

Vera MT, Rodriguez R, Segura DF, Cladera JL & Sutherst RW. 2002. Potential geographical distribution of the Mediterranean fruit fly, Ceratitis capitata (Diptera: Tephritidae), with emphasis on Argentina and Australia. *Environmental Entology*. 31(6): 1009-1022.

Via S. 1983. The quantitative genetics of polyphagy in an insect herbivore: II. Genetic correlations in larval performance within and among host plants. *Evolution*. 38 (4): 896-905.

Watt WB. 1977. Adaptation at species loci. I. Natural selection on phosphoglucose isomerase of colias butterflies: biochemical and population aspects. *Genetics*. 87:177-194.

Watt WB. 1983. Adaptation at specific cloci. II. Demographic and biochemical elements in the maintenance of the colias PGI polymorphism. *Genetics*. 103: 691-724.

Watt WB. 1992. Eggs, enzymes, and evlution: natural genetic variants change insect fecundity. *Proceedings of the National Academy of Sciences.* 89: 10608-10612.

Watt WB. 2000. Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. *Annual Rewiew of Genetics*. 34: 593-622.

Watt WB. 2003. Mechanistic stidies of butterfly adaptations. In: Ecology and volution taking flight: butterflies as model systems. Ed: Boggs CL, Watt WB, Ehrlich PR. Chicago (IL): University of Chicago Press.

Watt WB. 2004. Adaptation, constraint, and neutrality: mechanistic case studies with butterflies and their general implications. In *Evolution of population Biology*. Eds: Singh RS, Uyenoyama MK & Jain SK. Cambridge University Press, NY USA. pp 275-296.

Watt WB, Carter PA, Donohue K. 1986. Females' choice of "good genotypes" as mates is promoted by an insect mating system. *Science.* 233: 1187-1190.

Watt WB, Cassin RC & Swan MS. 1983. Adaptation at specific loci. III. Field behavior and survivorship differences among colias PGI genotypes are predictable from in vitro biochemistry. *Genetics*. 103: 725-739.

Watts MJ & Worner SP. 2008. Using artificial neural networks to determine the relative contribution of abiotic factors influencing the establishment of insect pest species. *Ecological Informatics*. 3: 64-74.

Wearing CH, Thomas WP, Dugdale JS & Danthanarayana W. 1991. Tortricid pests of pome and stone-fruits, Australian and New Zealand species. In: *Tortricid pests: their biology, natural enemies, and control world crop pests*. Vol. 5. Elsevier: Amsterdam. pp453-472.

Wheat CW, Haag CR, Marden JH, Hanski I & Frilander MJ. 2010. Nucleotide polymorphism at a Gene (*Pgi*) under balancing selection in a butterfly Metaopulation. *Molecular Biology and Evolution.* 27(2): 267-281.

Wheat CW, Watt WB, Pollock DD & Schulte PM. 2006. From DNA to fitness differences: sequences and structures of adaptive variants of *Colias* phosphoglucose isomerase (PGI). *Molecular Biology and Evolution.* 23(3): 499-512.

Williamson, M. 1996. *Biological invasions*. Chapman & Hall, London. p33.

Worner SP. 1988. Ecoclimatic assessment of potential establishment of exotic pests. *Journal of Economic Entomology* 81: 973-983.

Worner SP. 1994. Predicting the invasive potential of exotic insects. In: *Invasive arthropods in agriculture problems and solutions*. Eds: Hallman GJ & Schwalbe CP. Science Publishers. UK.

Worner SP & Gevrey M. 2006. Modelling global insect pest species assemblages to determine risk of invasion. *Journal of Applied Ecology.* 43: 858-867.

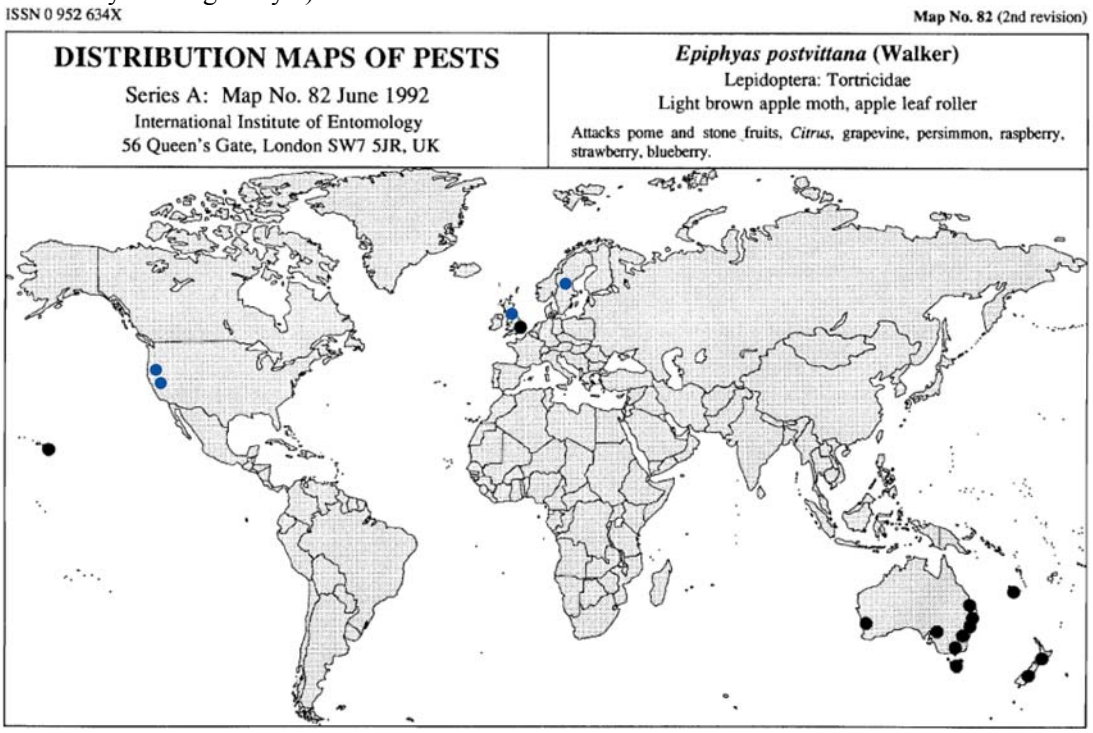
Worner SP, Ikeda T, Leday G, Joy M. 2010. Surveillance tools for freshwater invertebrates. MAF Biosecurity Technical Paper (In Press). Ministry of Agriculture and Forestry, New Zealand.

Zheng, C. Ovaskainen O & Hanski I. 2009. Modelling single nucleotide effects in phosphoglucose isomerase on dispersal in the Glanville fritillary butterfly: coupling of ecological and evolutionary dynamics. *Philosophical Transactions of the Royal Society B*. 364: 1519-1532.

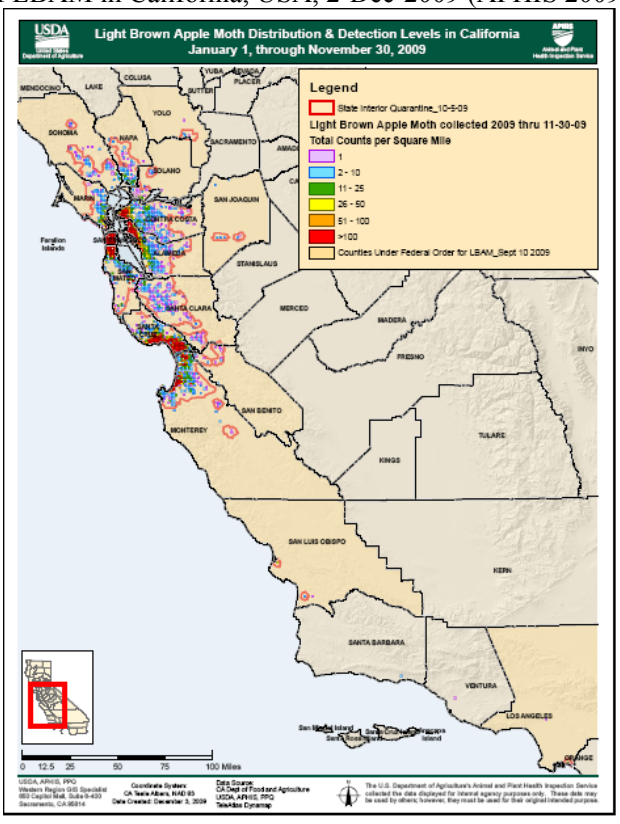
# Appendices

## Appendix 1: Current distribution of LBAM

a. World distribution map of light brown apple moth revised from CABI 1992 (black dots). The blue dots represent the presence of LBAM in all over UK, California USA, and Sweden (detected but not fully investigated yet).




b. Distribution map of LBAM in California, USA, 2-Dec-2009 (APHIS 2009).



## Appendix 2: Presence data of LBAM

a. Localities have been proved with LBAM presence by references.

| Regions | Locations | Longitudes | Latitudes | References |
|---|---|---|---|---|
| UK | Newquay, Cornwall | -5.083364 | 50.413608 | Fountain & Cross 2007 |
| | Devon | -3.99946 | 50.777213 | |
| | Yalding, Kent | 0.432 | 51.226146 | |
| | Hereford | -2.717547 | 52.05599 | |
| | Essex | 0.667366 | 51.765908 | |
| | Oxfordshire | -1.246467 | 51.761206 | |
| Southeast Ireland | Wexfor | -6.460125 | 52.339161 | Suckling & Brockerhoff 2010 |
| Hawaii | Kauai | -159.52612 | 22.09644 | |
| | Oahu Molkai | -158.00005 | 21.438912 | |
| | Lanai | -156.92731 | 20.861472 | |
| | Maui | -156.331925 | 20.79836 | |
| | Hawaii(Big Island) | -155.665857 | 19.542915 | |
| California | Alameda | -122.241636 | 37.765206 | USDA-APHIS(Dec-2-2009) |
| | Contra costa | -121.901795 | 37.853409 | |
| | Los Angeles | -118.243685 | 34.052234 | |
| | Marin | -122.763304 | 38.0834 | |
| | Napa | -122.2898 | 38.305031 | |
| | Monterey | -121.89467 | 36.60023 | |
| | San Benito | -121.08186 | 36.509685 | |
| | San Francisco | -122.41941 | 37.77493 | |
| | San Joaquin | -120.1890447 | 36.6066162 | |
| | San Luis Obispo | -120.6596156 | 35.2827524 | |
| | San Mateo | -122.3255254 | 37.5629917 | |
| | Santa Barbara | -119.6981901 | 34.4208305 | |
| | Santa Clara | -121.9552356 | 37.3541079 | |
| | Santa Cruz | -122.0307963 | 36.9741171 | |
| | Solano | -121.901795 | 38.310497 | |
| | Sonoma | -122.4580356 | 38.291859 | |
| | Ventura | -119.2931676 | 34.2783352 | |
| | Yolo | -121.8077431 | 38.7318481 | |
| Australia | Canberra | 149.12858 | -35.28204 | Gu & Danthanarayana 2000 |
| | Melbourne | 144.963169 | -37.814251 | |
| | Lenswood | 138.828222 | -34.920014 | Bailey *et al*. 1997 |
| | Merbein | 142.060861 | -34.167929 | Buchanan 1977 |

Appendix 2a continued:

| Regions | Locations | Longitude | Latitude | References |
|---|---|---|---|---|
| Australia | Narrabri | 149.782833 | -30.324835 | Geier & Briese 1979 |
| | Dubbo | 148.60421 | -32.245192 | |
| | Bathurst | 149.574258 | -33.41978 | |
| | Yanco | 146.410003 | -34.603418 | |
| | Batlow | 148.144623 | -35.522019 | |
| | Albury | 146.910174 | -36.082137 | |
| | Kerang | 143.918869 | -35.735049 | |
| | Tatura | 145.227847 | -36.440085 | |
| | Ballarat | 143.863715 | -37.563318 | |
| | Melbourne | 148.454887 | -37.707336 | |
| | Orbost | 147.048437 | -43.030849 | |
| | Huonville | 145.228757 | -37.847915 | |
| | Wantirna | 145.07196 | -37.70132 | |
| | Bundoora, Victoria | 145.190507 | -38.30771 | Danthanarayana *et al*. 1995 |
| | Hastings | 145.906337 | -41.052874 | Geier & Briese 1980b |
| | Burnie | 147.330234 | -42.882743 | |
| | Hobart | 146.406444 | -34.5522 | |
| | Leeton | 145.189829 | -38.261302 | Geier *et al*. 1978 |
| | Tyabb | 138.616756 | -35.021155 | |
| | Blackwood | 116.069584 | -32.11287 | |
| | Roleystone | 147.096403 | -42.9818 | |
| | Grove | 169.310859 | -45.21606 | |
| New Zealand | Earnscleugh | 169.326609 | -45.189837 | McLaren & Suckling 1993 |
| | Clyde | 169.388507 | -45.251666 | |
| | Alexandra | 173.011121 | -41.110107 | |
| | Motueka | 173.099638 | -41.25612 | Suckling *et al*. 1987 |
| | Mapua | 172.651684 | -43.515414 | |
| | Richmond | 172.484281 | -43.640254 | |
| | Lincoln | 173.284 | -41.270786 | Suckling *et al*. 1998 |
| | Nelson | 176.7416374 | -39.1089867 | |
| | Hawke's Bay | 170.503388 | -45.87456 | |
| | Central Otago | 171.7070007 | -43.5259018 | |
| | Canterbury | 172.997403 | -41.07648 | |
| | Riwaka | 173.046028 | -41.265947 | Shaw *et al*. 1994 |
| | Mahana | 172.699997 | -41.4017982 | |

Appendix 2 continue:
b. Present data of LBAM downloaded from GBIF. Duplicate data has been removed.

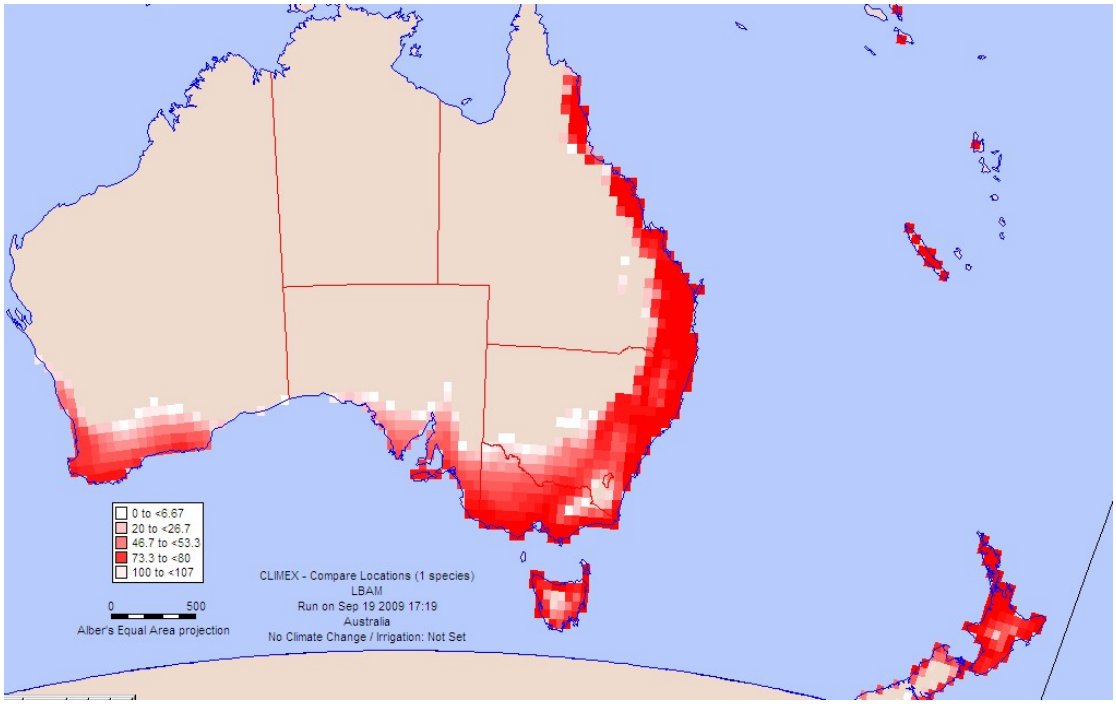
| Longitude | Latitude | Longitude | Latitude |
|---|---|---|---|
| 172.62 | -43.49 | -3.27339 | 50.6127 |
| 174.74 | -36.74 | -3.13209 | 50.6142 |
| 174.8 | -36.82 | -3.97971 | 50.6029 |
| 174.74 | -36.72 | -3.83847 | 50.6052 |
| -5.34635 | 49.9419 | -3.55006 | 50.4295 |
| -5.3526 | 50.0317 | -3.13425 | 50.7041 |
| -4.82663 | 50.5854 | -3.694 | 50.5174 |
| -4.80006 | 50.1363 | -3.41468 | 50.6111 |
| -4.25353 | 50.418 | -3.69079 | 50.4276 |
| -4.10491 | 50.2409 | -2.00136 | 50.5298 |
| -3.553 | 50.5194 | -1.57486 | 50.8887 |
| -3.82118 | 50.1558 | -1.28918 | 50.9772 |
| -3.55596 | 50.6093 | -0.27575 | 51.4163 |
| -3.83498 | 50.5153 | 1.265 | 51.849 |
| -3.68442 | 50.2478 | -3.44481 | 56.0045 |

**Appendix 3:** Parameters set for *E. postvittana* in CLIMEX.

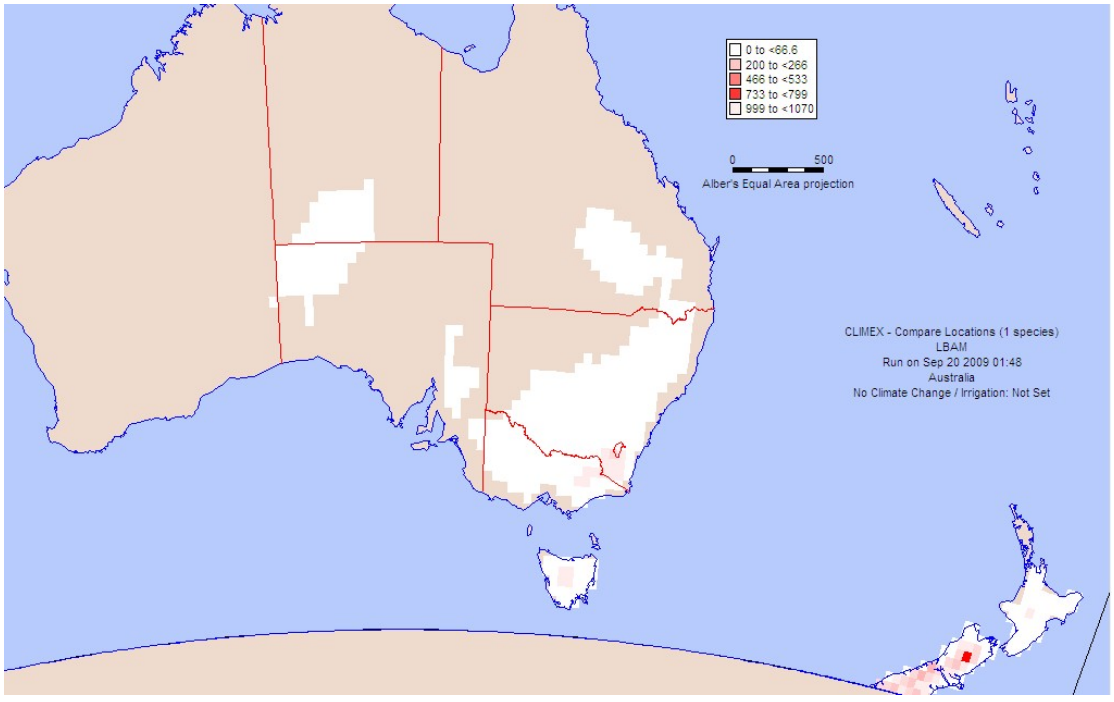
| Parameters | Mnemonic | Values in template | Values after calibration |
|---|---|---|---|
| Temperature | | | |
| Limiting low temperature | DV0 | 8 | **7** |
| Lower optimal temperature | DV1 | 18 | **13** |
| Upper optimal temperature | DV2 | 24 | **23** |
| Limiting high temperature | DV3 | 28 | **30** |
| Degree-days per generation | PDD | 600 | **673.6** |
| Moisture | | | |
| Limiting low soil moisture | SM0 | 0.25 | **0.25** |
| Lower optimal soil moisture | SM1 | 0.8 | **0.8** |
| Upper optimal soil moisture | SM2 | 1.5 | **1.5** |
| Limiting high soil moisture | SM3 | 2.5 | **2.5** |
| Cold stress | | | |
| Cold stress temperature threshold | TTCS | 0 | **5** |
| Cold stress temperature rate | THCS | 0 | **-0.0005** |
| Cold stress degree-day threshold | DTCS | 15 | **0** |
| Cold stress degree-day rate | DHCS | -0.0001 | **0** |
| Cold stress temperature threshold (average) | TTCSA | 0 | **0** |
| Cold stress temperature rate (average) | THCSA | 0 | **0** |
| Heat stress | | | |
| Heat stress temperature threshold | TTHS | 30 | **31** |
| Heat stress temperature rate | THHS | 0.005 | **0.01** |
| Heat stress degree-day threshold | DTHS | 0 | **0** |
| Heat stress degree-day rate | DHHS | 0 | **0** |
| Dry stress | | | |
| Dry stress threshold | SMDS | 0.2 | **0.2** |
| Dry stress rate | HDS | -0.005 | **-0.01** |
| Wet stress | | | |
| Wet stress threshold | SMWS | 2.5 | **2.5** |
| Wet stress rate | HWS | 0.002 | **0.002** |

**Appendix 4:** Modeled core distribution of LBAM and cold/heat/dry index in Australia and New Zealand.

a. Core distribution of LBAM in Australia and New Zealand.

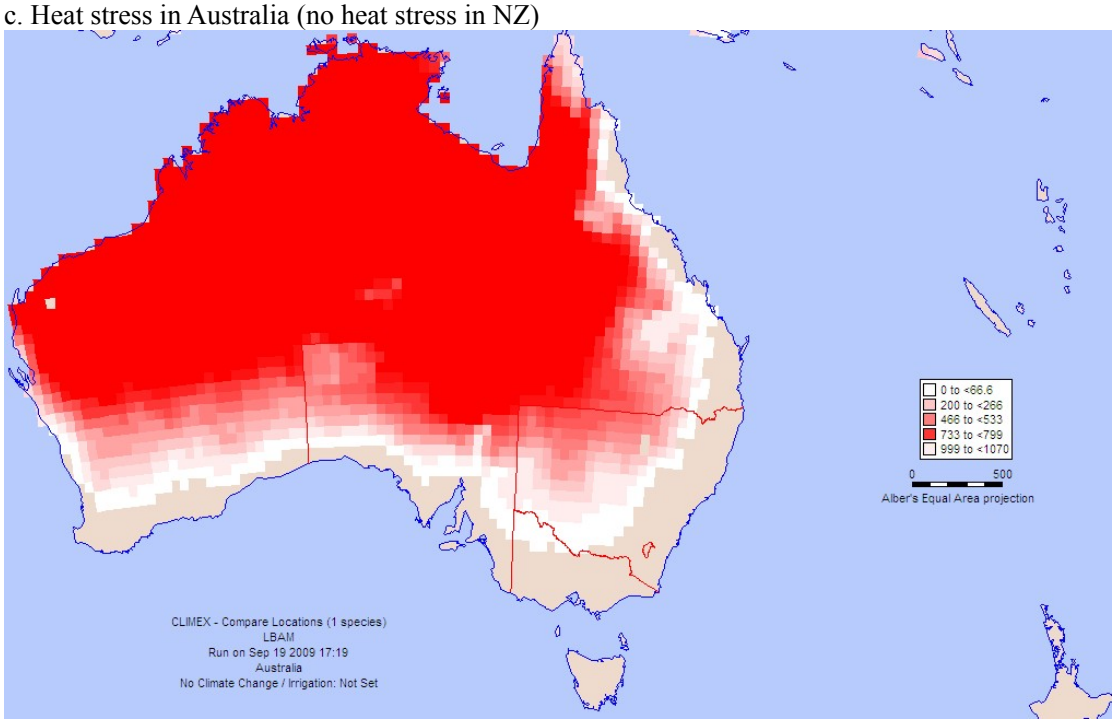



b. Cold stress in Australia and New Zealand

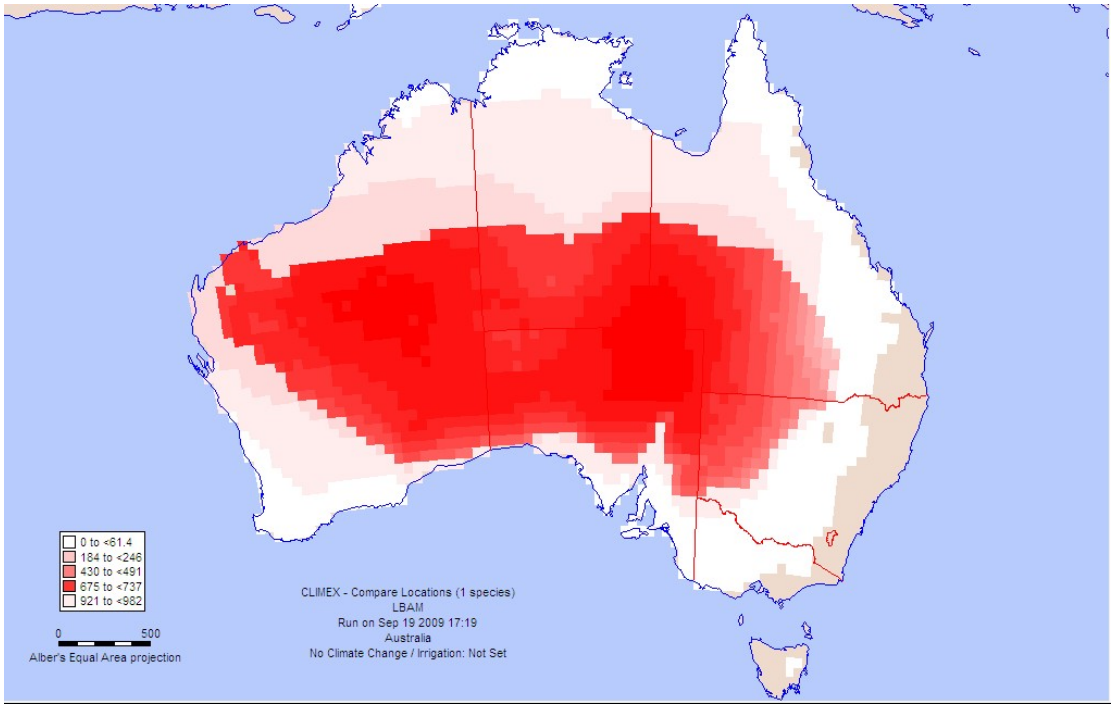


Appendix 4 continued:

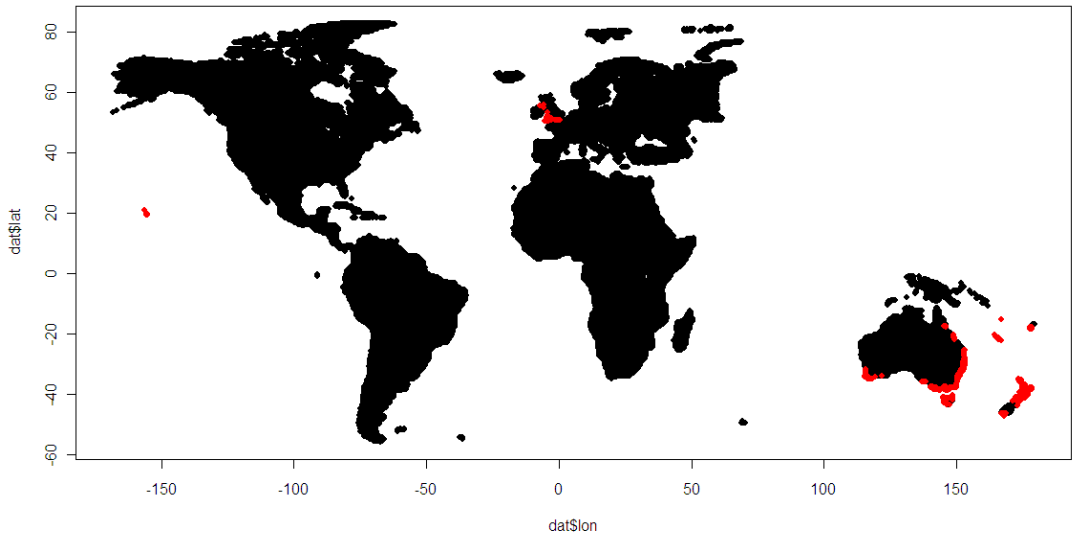
c. Heat stress in Australia (no heat stress in NZ)




d. Dry stress in Australia (no dry stress in New Zealand) shows where LBAM might not be able to survive.

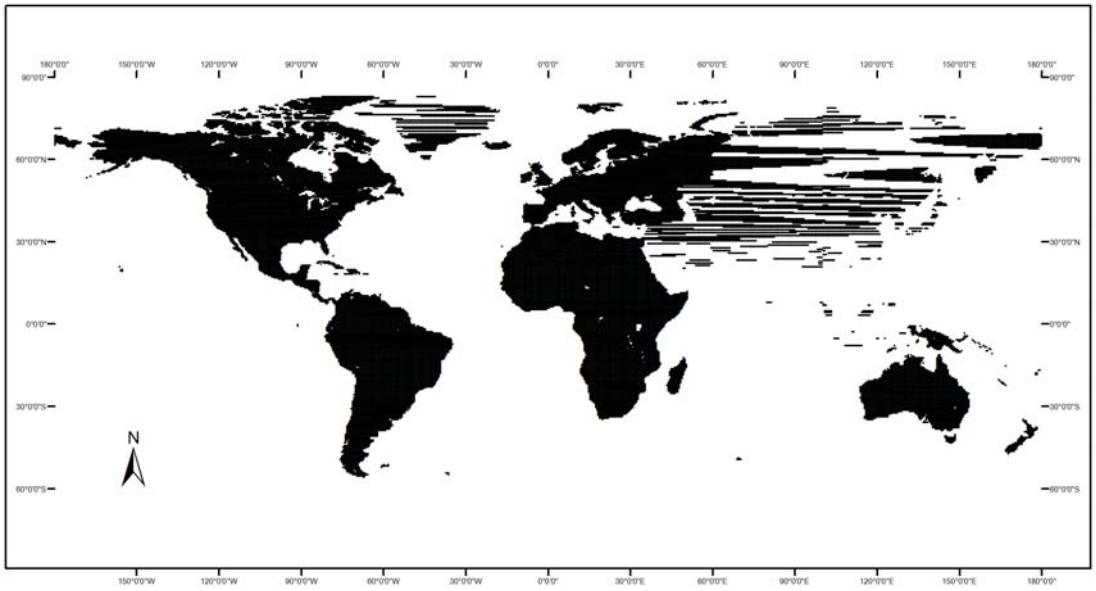


Appendix 4 continued:
e. Predicted distribution of LBAM in areas that currently with LBAM occurrence (red dots represent predicted EI value greater than 30).

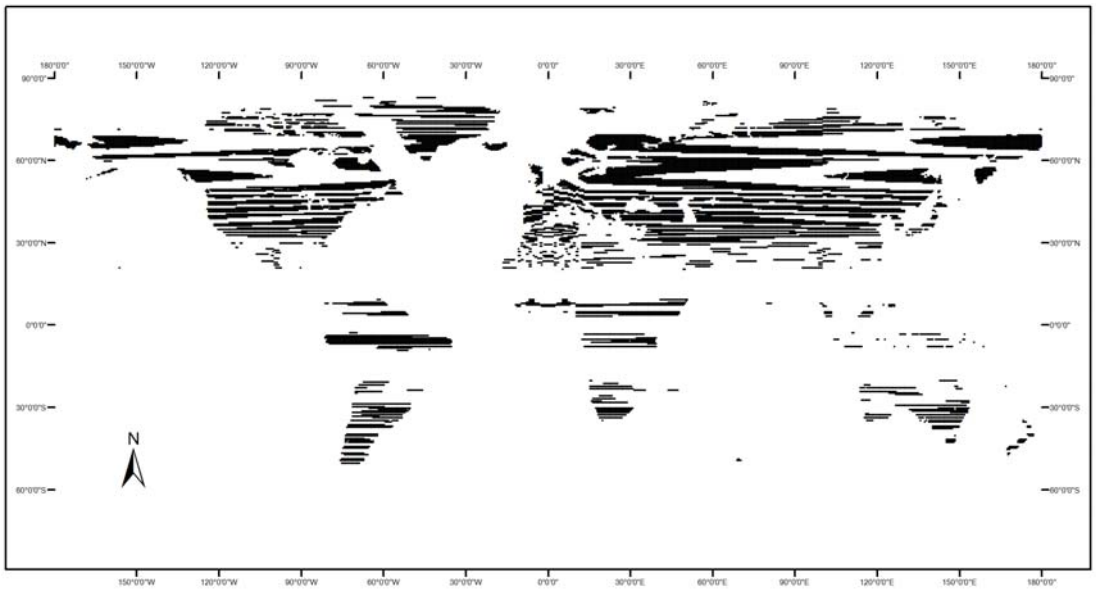


**Appendix 5: Plotted coordinates extracted from CLIMEX**

a. A world map plotted with 61,076 coordinates from CLIMEX. As we noticed that some parts of Asia is missing.

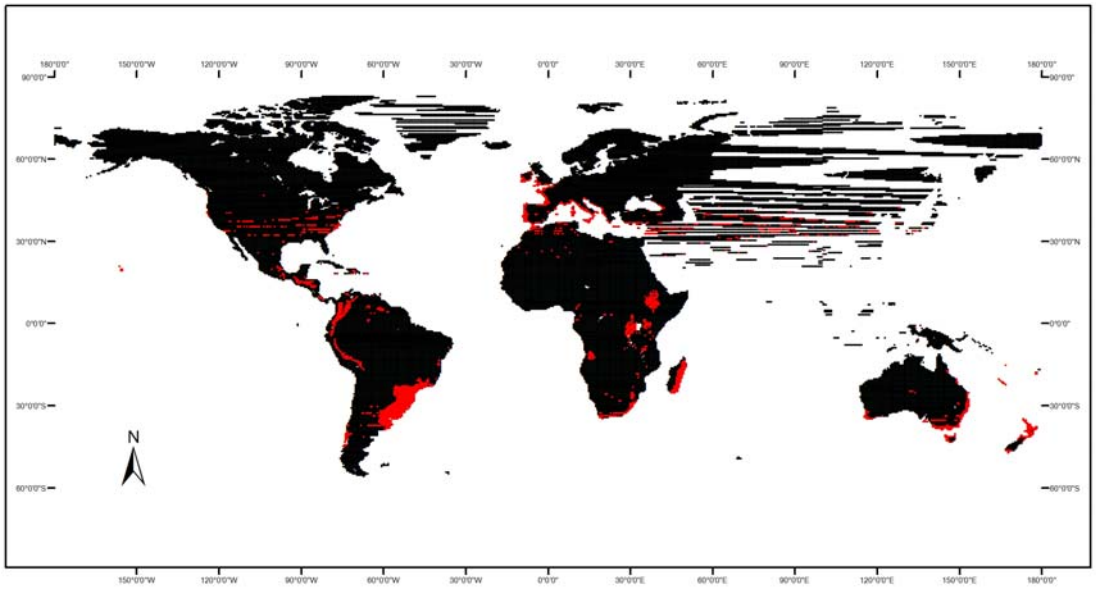


b. A map of Asia using coordinates of Asia in CLIMEX, which looks more like a world map.

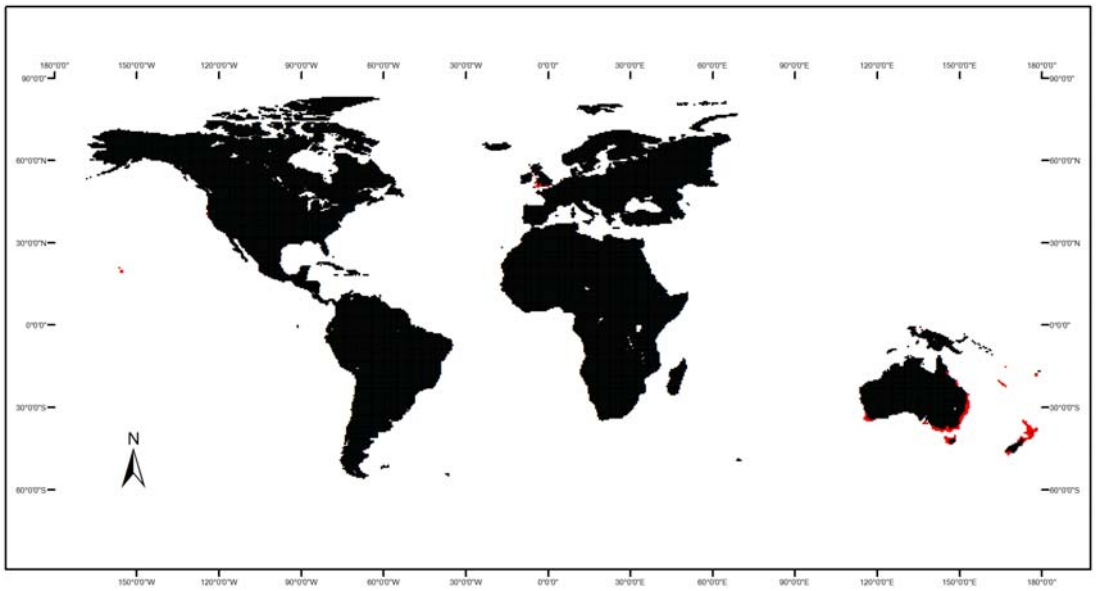


Appendix 5 continued:

c. A world map plotted with coordinates from CLIMEX, and locations with EI value greater than 30 are highlighted (red dots). However, several dots in the central of Australia were plotted in the map, which should not have been there due to the intensive hot and dry stresses at that location. This indicates either the prediction was made by setting the wrong parameters in CLIMEX, or coordinates of Asia in CLIMEX are wrong.

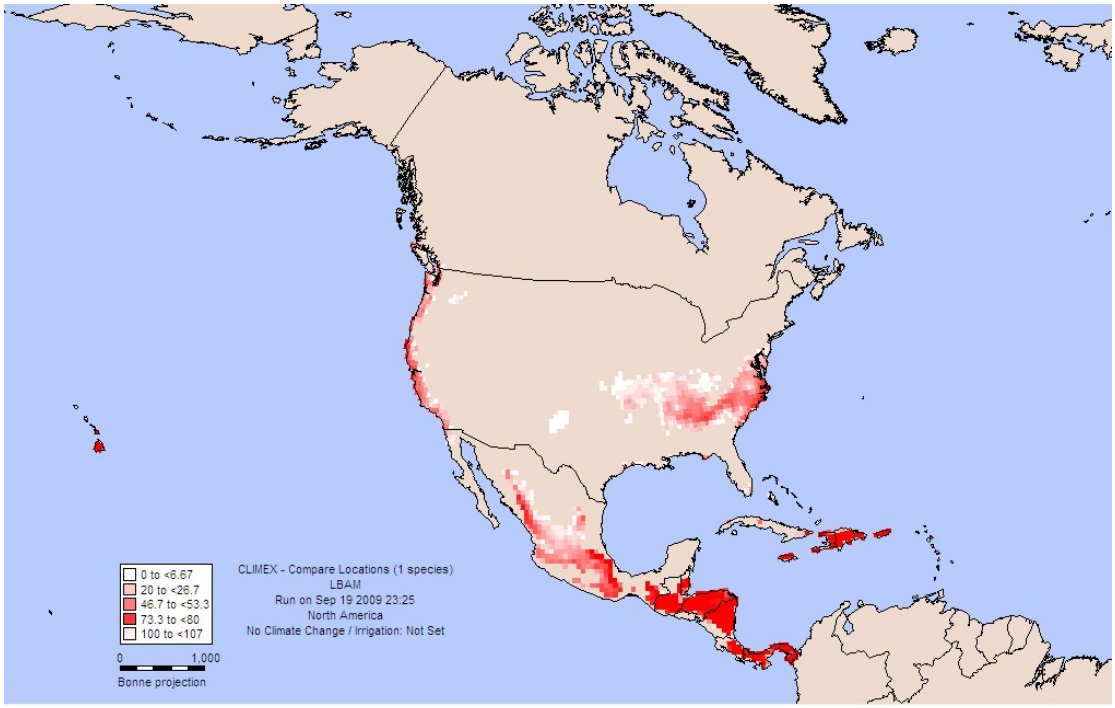


d. World map excluded Asia, with EI value greater than 30 in Australia and New Zealand highlighted (red dots). As we can tell that all red dots are fall into the current distribution map (CABI 1992). No unexpected EI-value-greater-than-30 points were plotted after the exclusion of Asia coordinates.

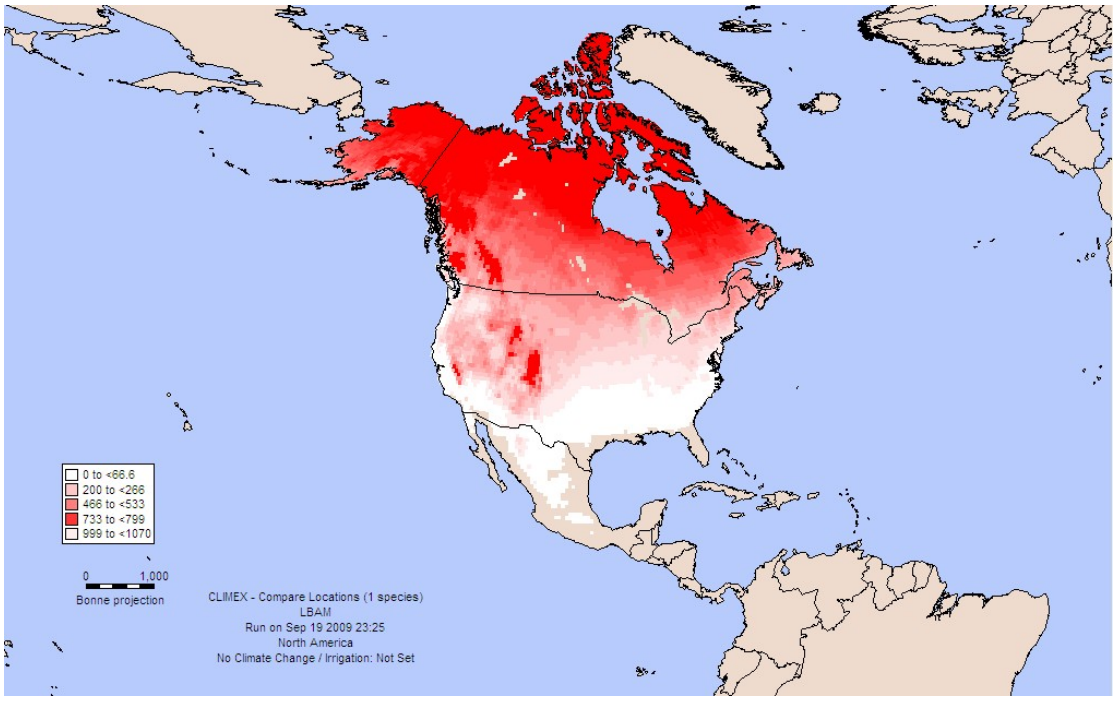


**Appendix 6**: Predicted core distribution and estimated cold/heat/dry stress of LBAM in North America. Figure from a to d were modeled without considering irrigation, figure e and f were remodeled by adding irrigation set automatically in the temperate template.

a. Predicted core distribution of LBAM in North America

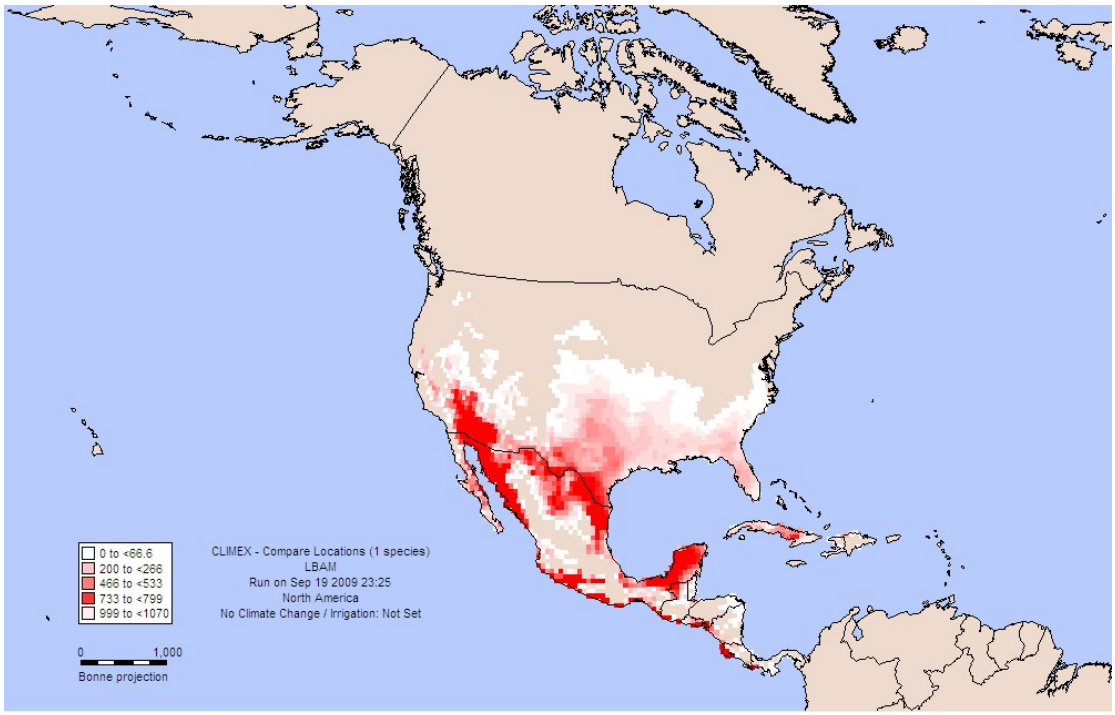



b. Estimated cold stress of LBAM in North America

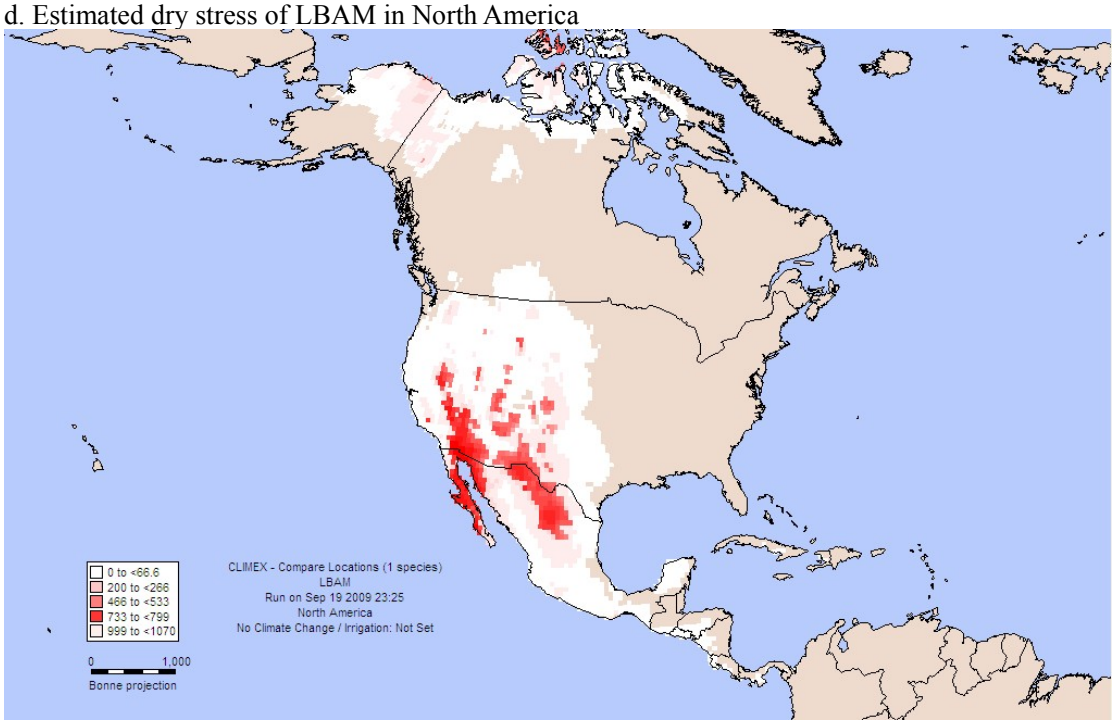


Appendix 6 continued:

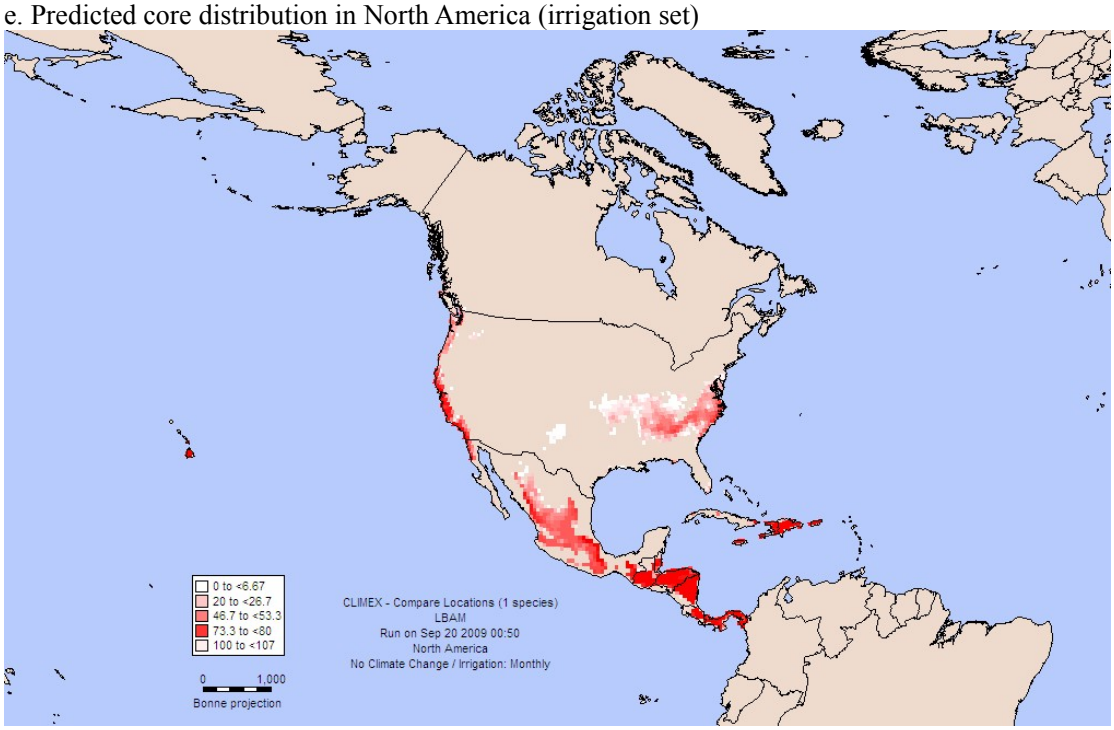
c. Estimated heat stress of LBAM in North America




d. Estimated dry stress of LBAM in North America

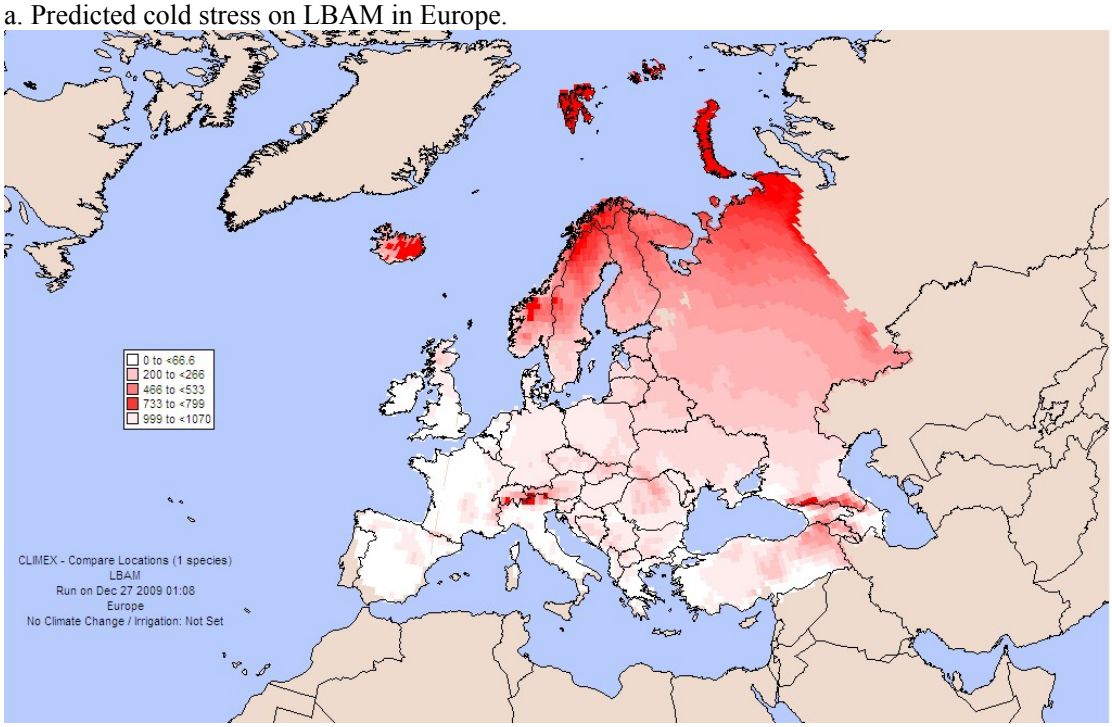


Appendix 6 continued:
e. Predicted core distribution in North America (irrigation set)



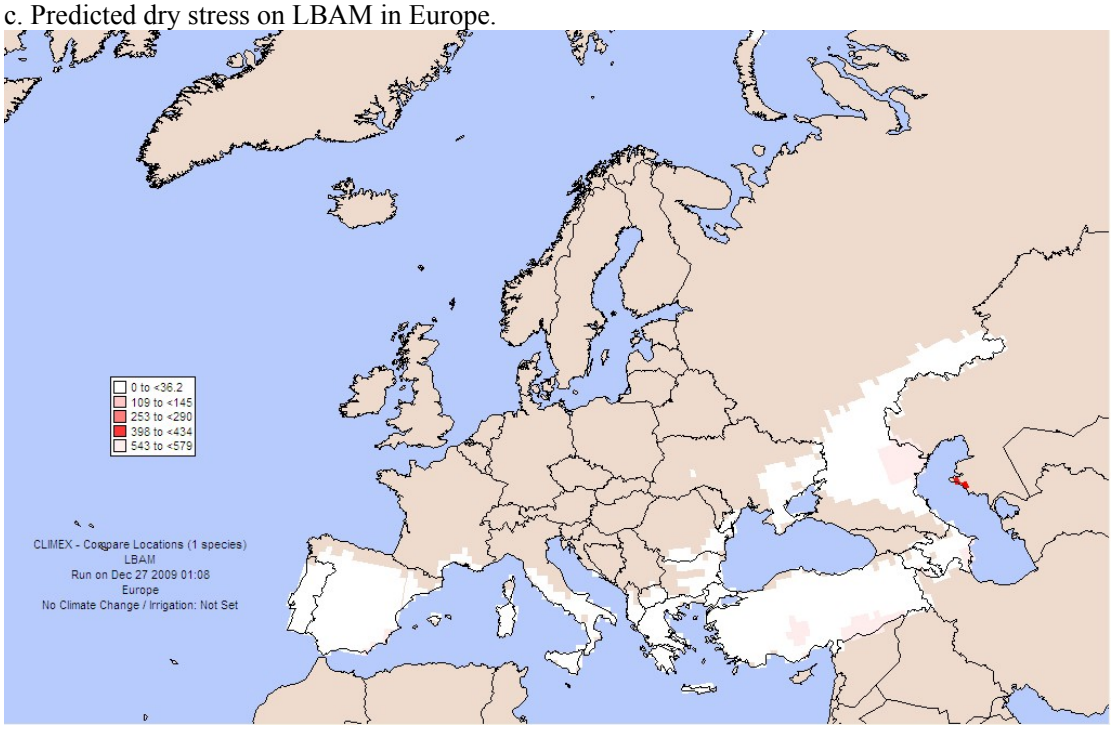
**Appendix 7**. Predicted potential core distribution and cold/heat/dry stress of LBAM in Europe.

a. Predicted cold stress on LBAM in Europe.

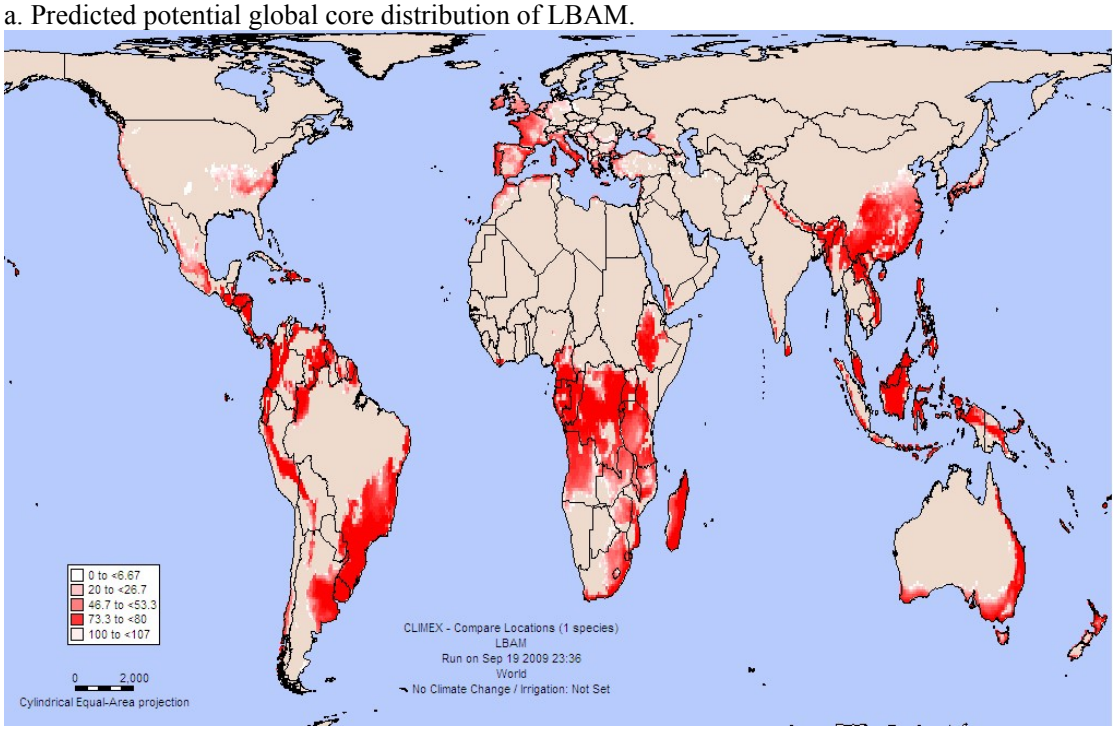



b. Predicted heat stress on LBAM in Europe.

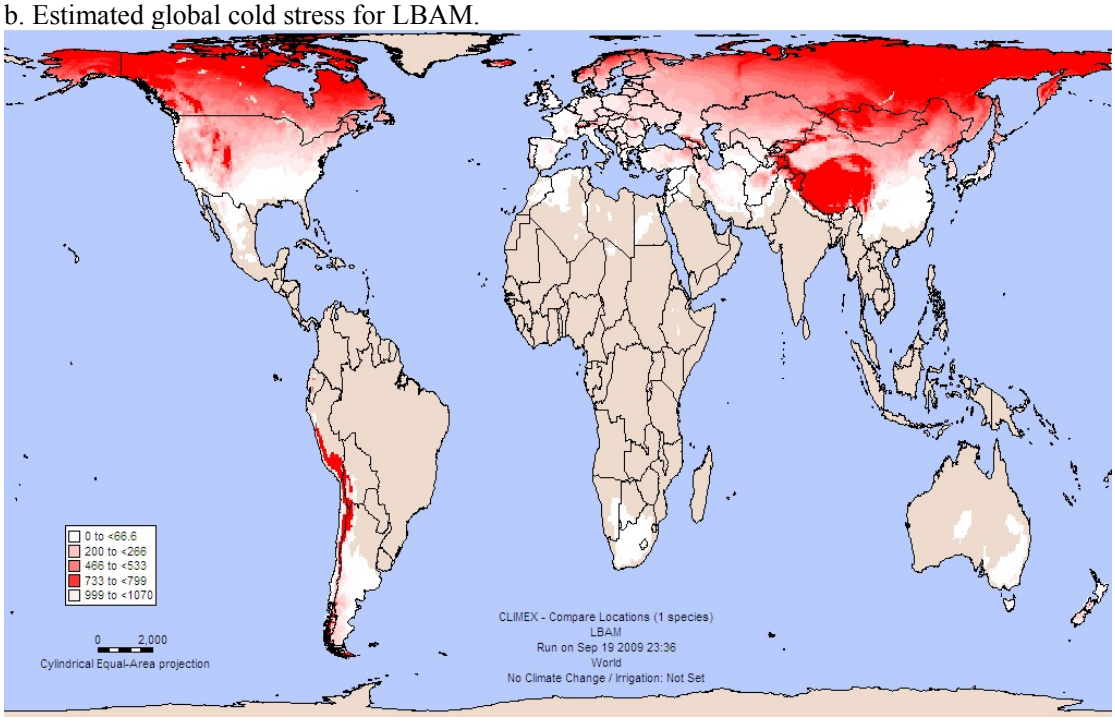
Appendix 7 continued:
c. Predicted dry stress on LBAM in Europe.



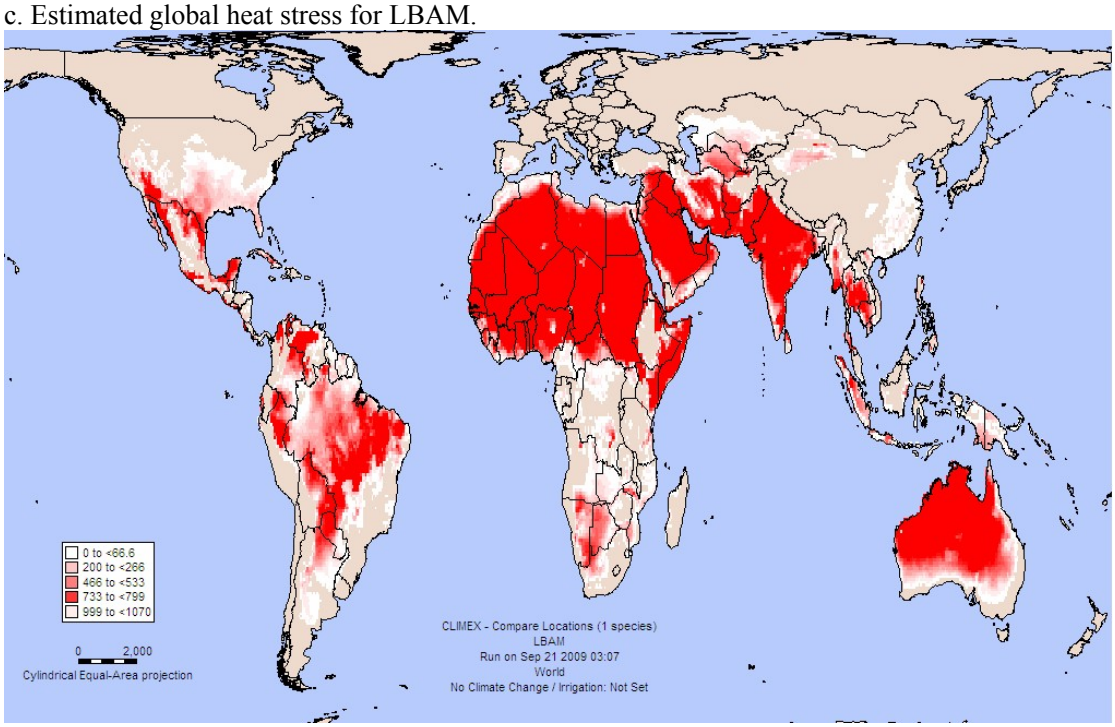
**Appendix 8**. Predicted potential core distribution and cold/heat/dry stress of LBAM worldwide.
a. Predicted potential global core distribution of LBAM.



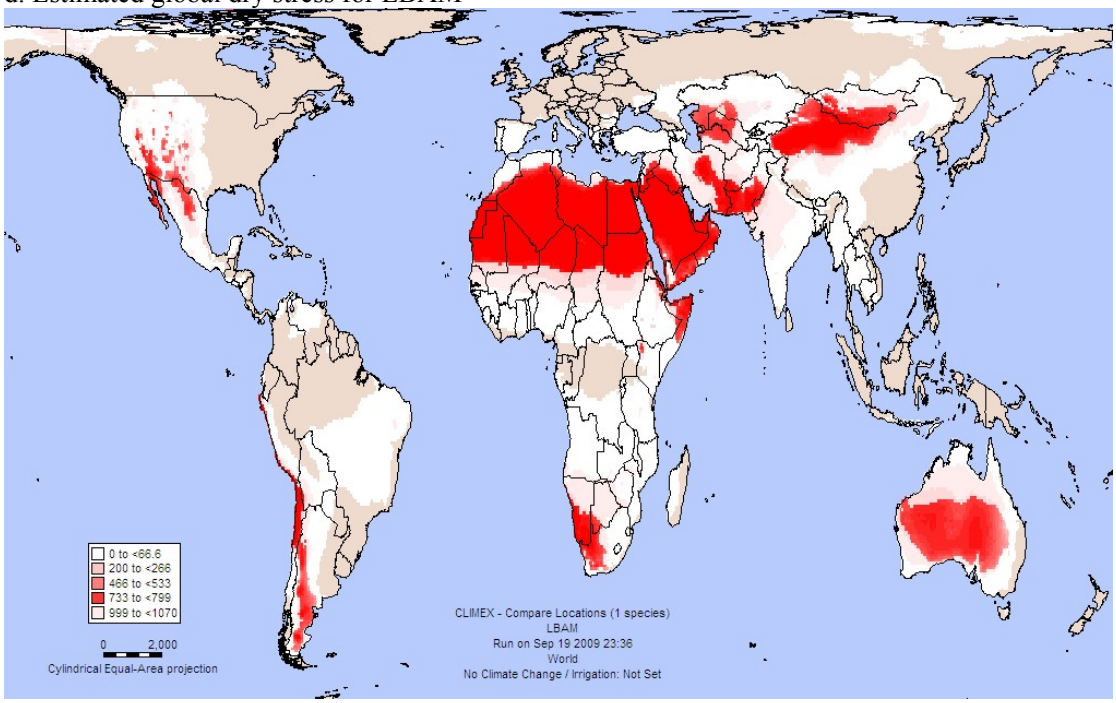

b. Estimated global cold stress for LBAM.



Appendix 8 continued:
c. Estimated global heat stress for LBAM.

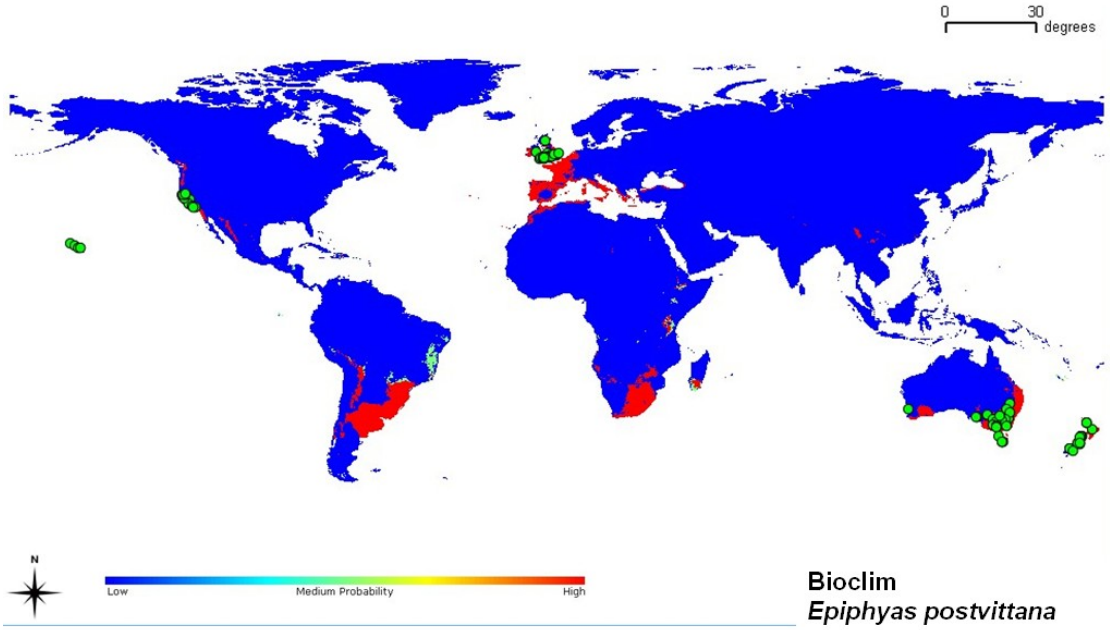



d. Estimated global dry stress for LBAM

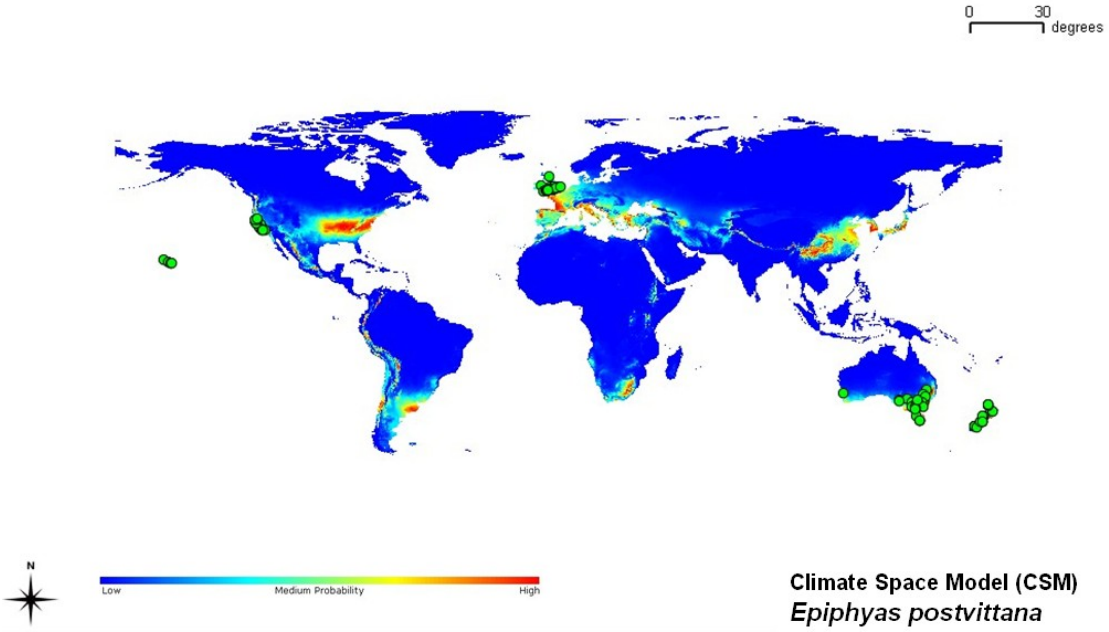


## Appendix 9: Prediction of LBAM potential distribution using openModeller 1.0

a. Prediction map of LBAM potential distribution using Bioclim (Accuray = 1; AUC = 0.977789).

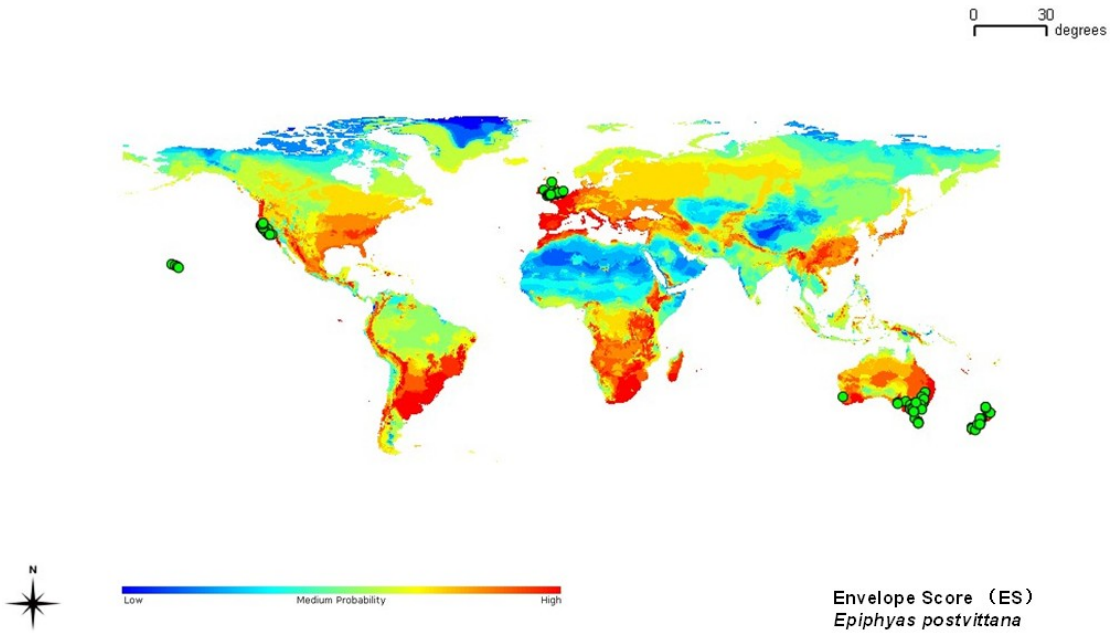



b. Prediction map of LBAM potential distribution using Climate Space Model (CSM) (Accuracy = 0.681319; AUC = 0.922004).

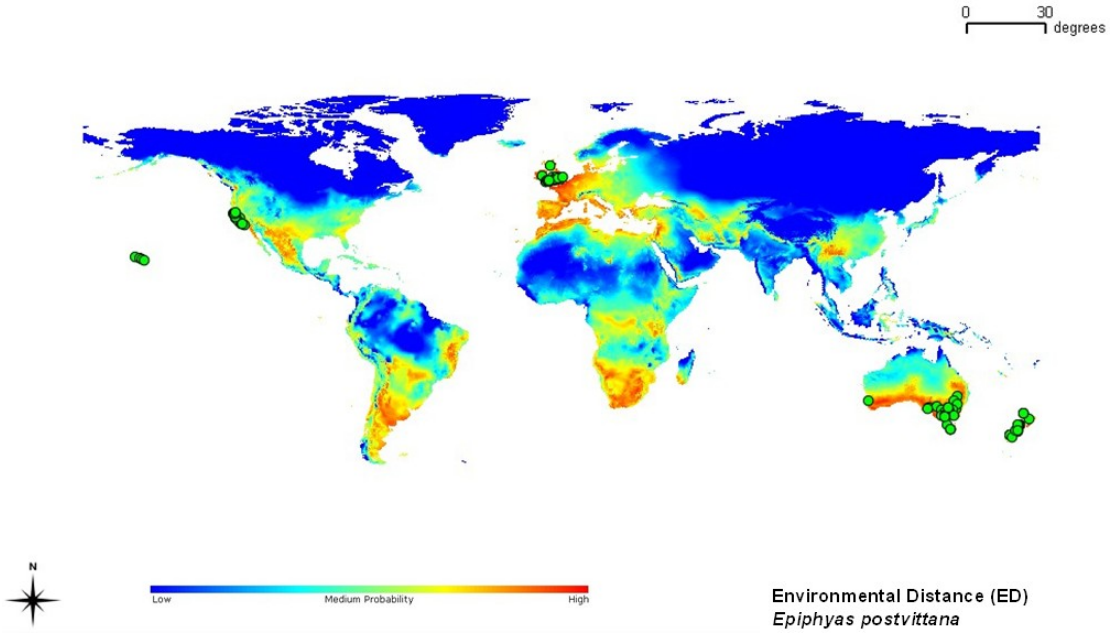


Appendix 9 continued:

c. Prediction map of LBAM potential distribution using Envelop Score (Accuracy = 1; AUC = 0.97625).

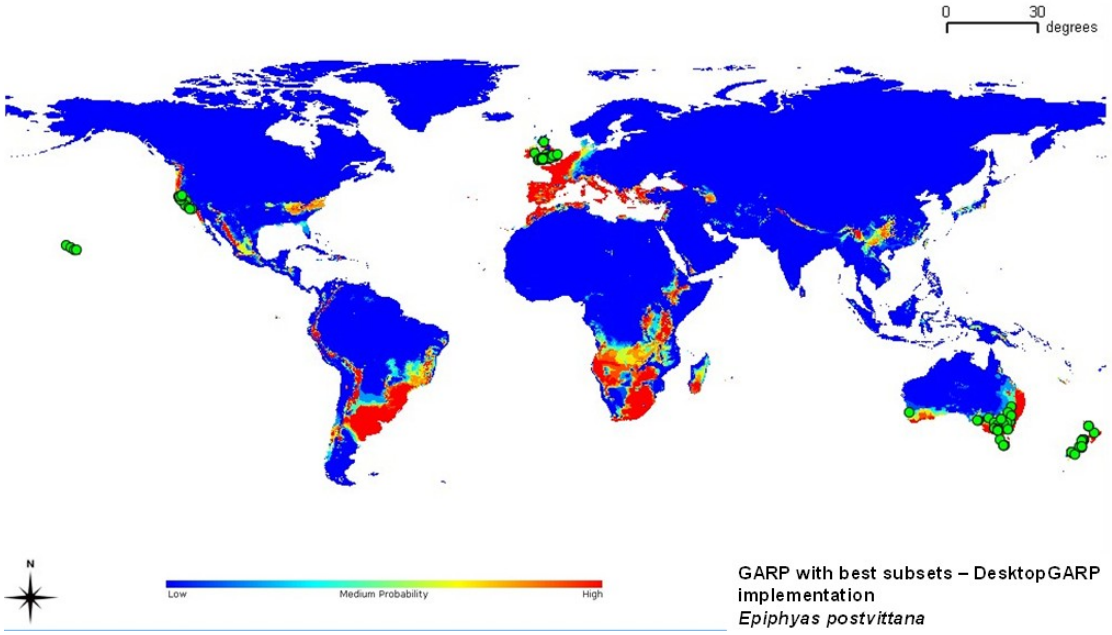



d. Prediction map of LBAM potential distribution using Environmental Distance (Accuracy = 1; AUC = 0.9999).

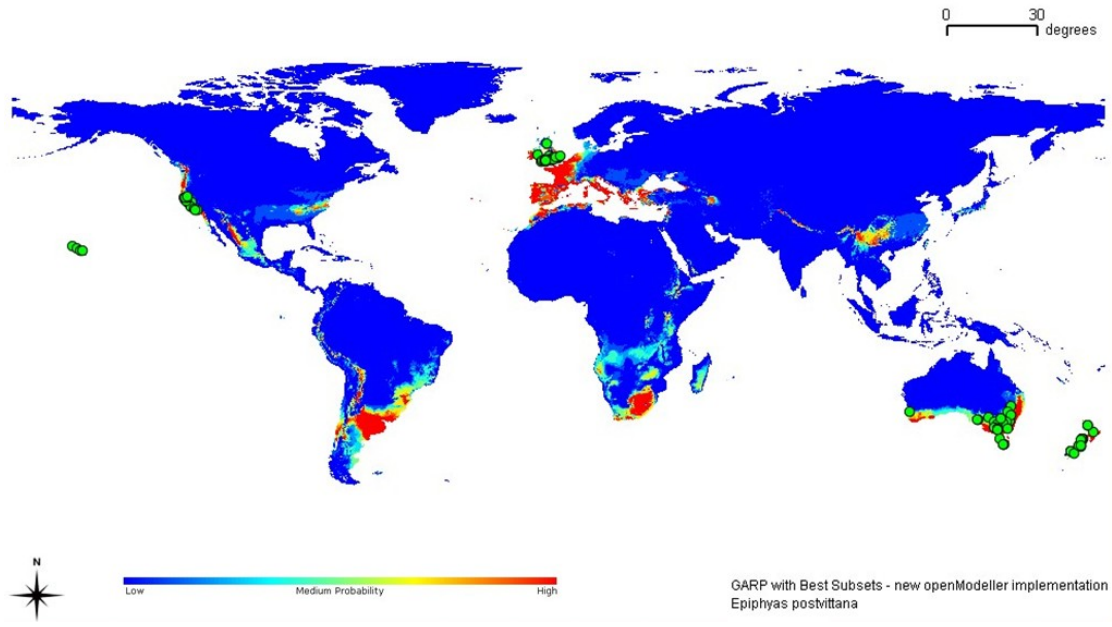


Appendix 9 continued:

e. Prediction map of LBAM potential distribution using GARP with best subsets-desktopGARP implementation (Accuracy = 0.967033; AUC = 0.969156).

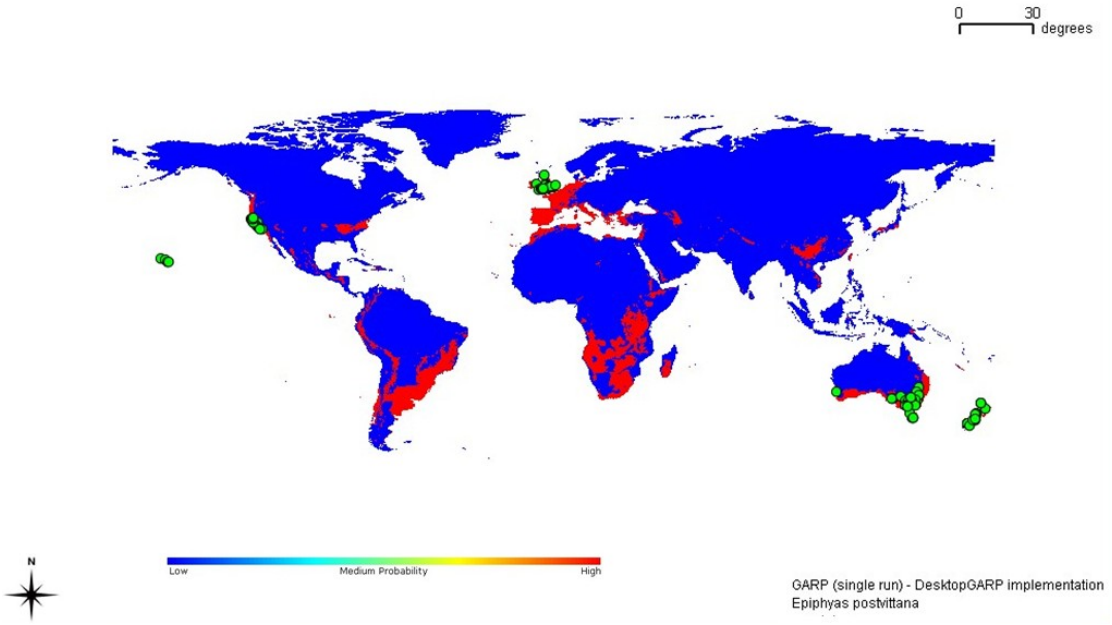



f. Prediction map of LBAM potential distribution using GARP with best subsets – new openModeller implementation (Accuracy =0.945055; AUC =0.95554).

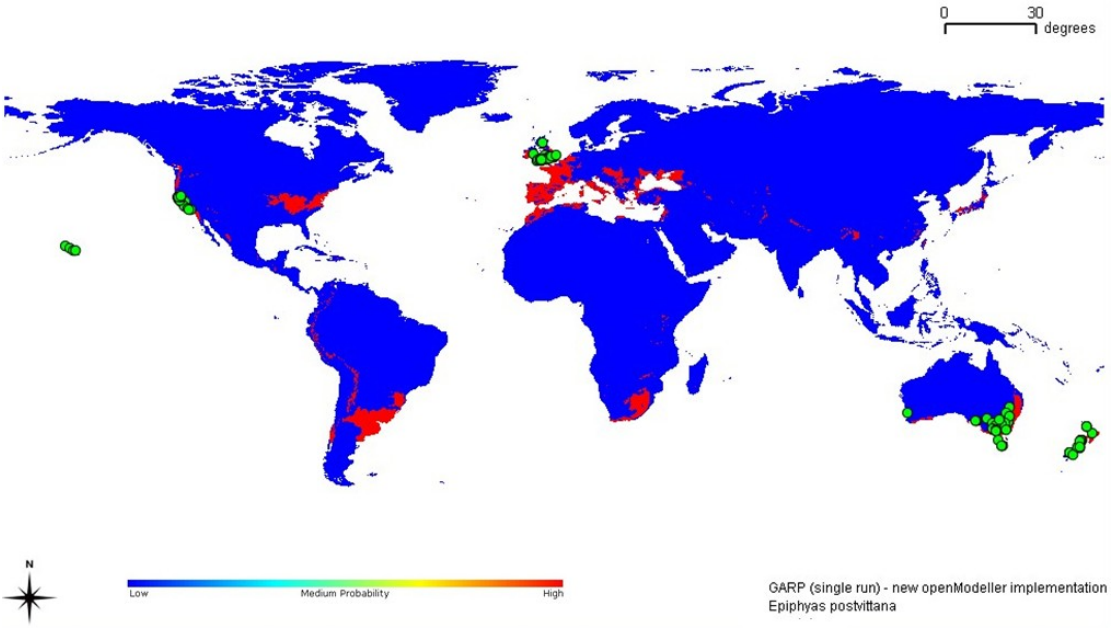


Appendix 9 continued:

g. Prediction map of LBAM potential distribution using GARP (single run) – desktopGARP implementation (Accuracy = 1; AUC = 0.9469).

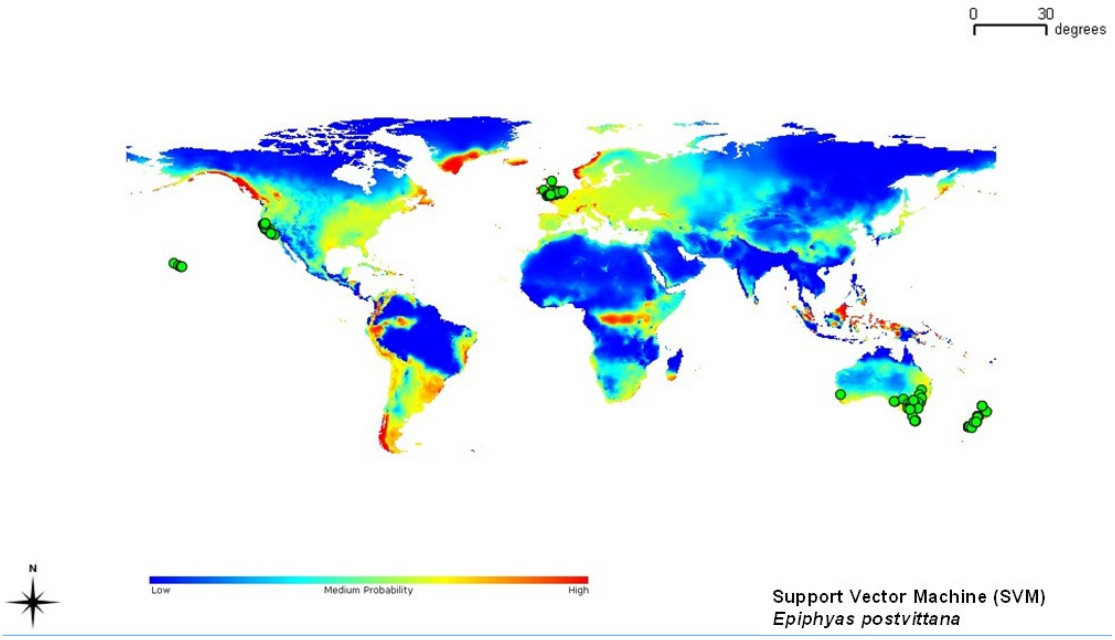



h. Prediction map of *E. postvittana* potential distribution using GARP (single run) – desktopGARP implementation (Accuracy =0.802198; AUC =0.880049).



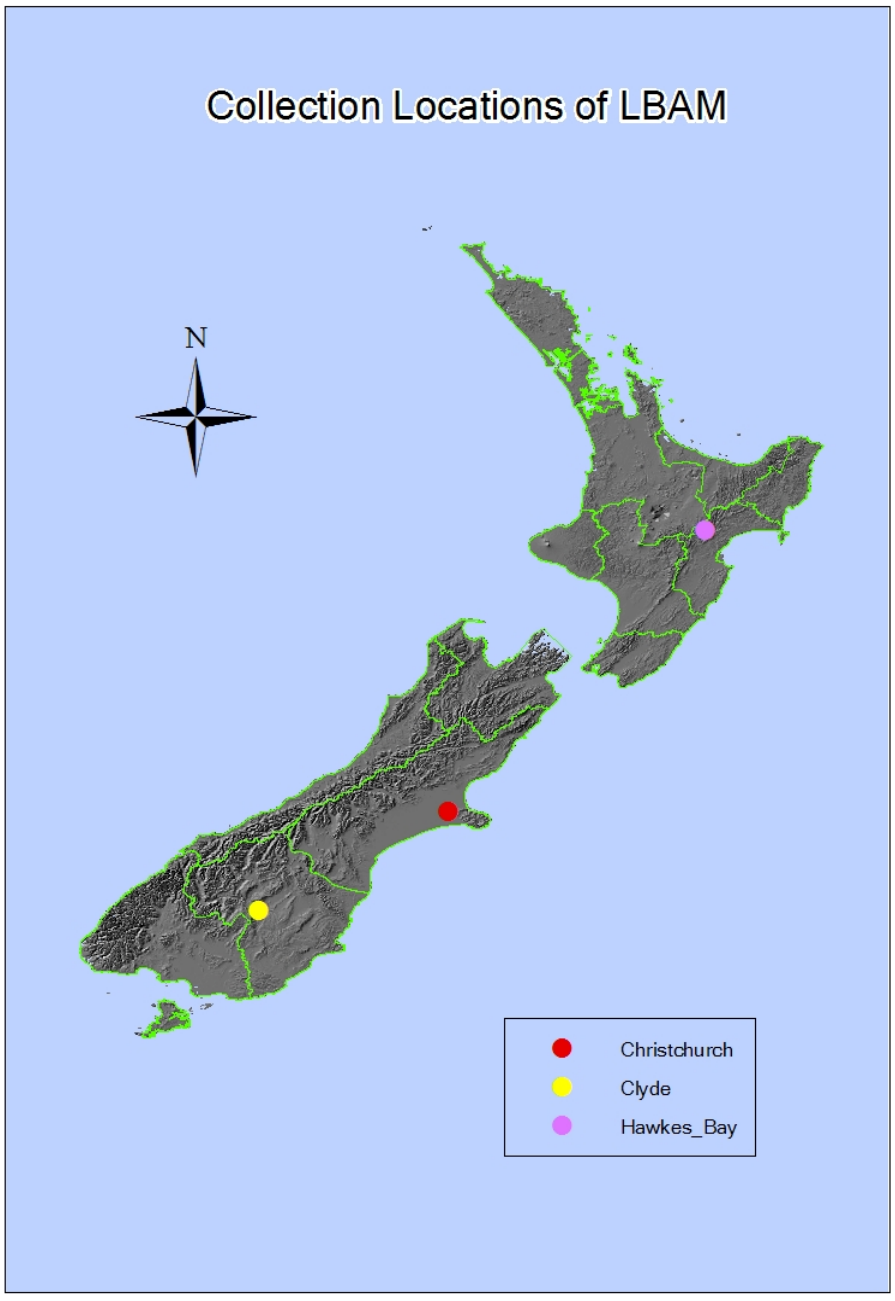
Appendix 9 continued:

i. Prediction map of *E. postvittana* potential distribution using SVM. (SVM type = C-SVM; kernel type = polynomial. Accuracy = 0.9890011; AUC = 0.93).



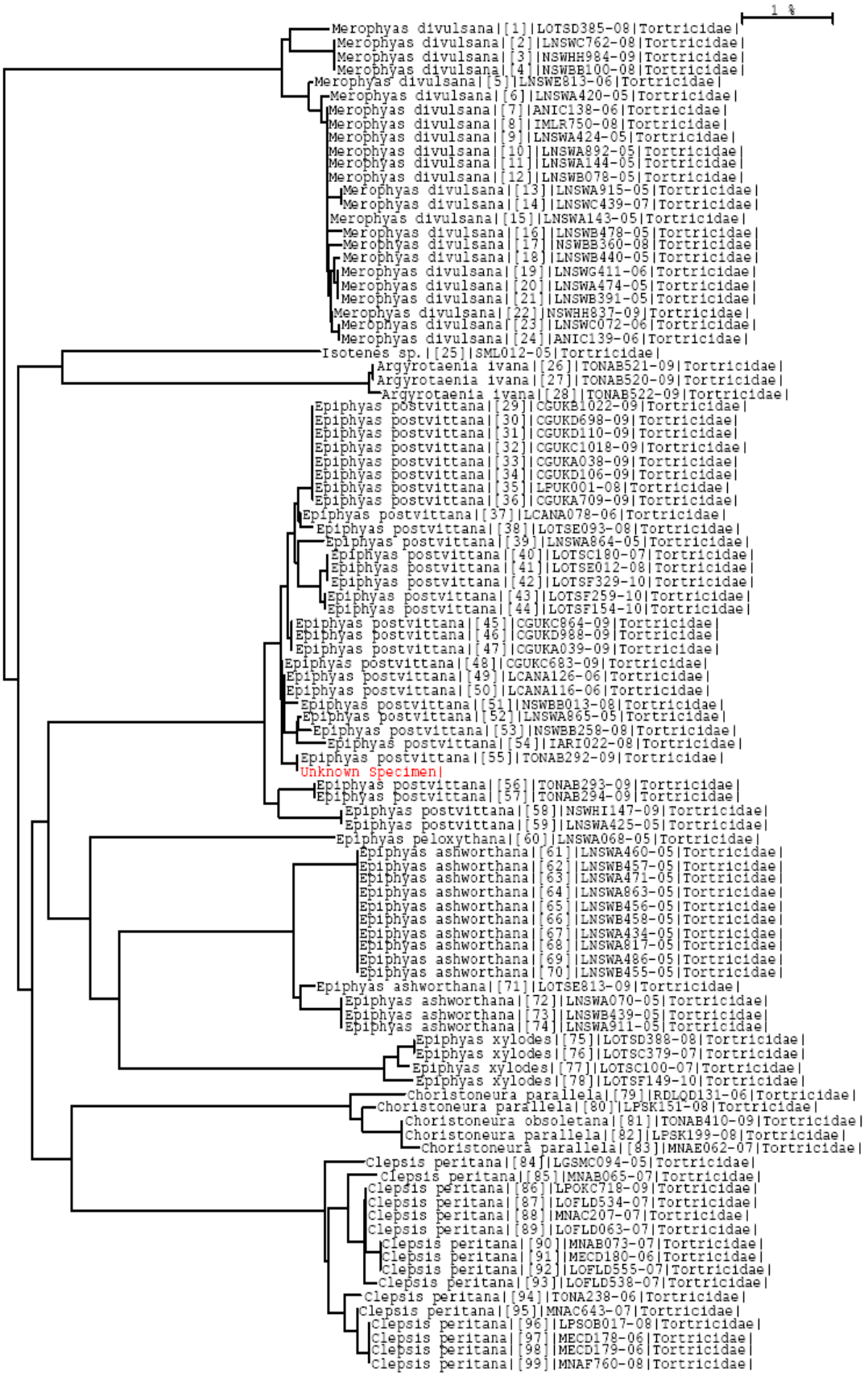
**Appendix 10**: Collection locations of *E. postvittana*.

Three populations collected from Hawkes Bay, Christchurch and Clyde respectively. Another population is from colony reared in Auckland Landcare Research Ltd, under controlled climate-chamber with artificial diet.



**Appendix 11: Identification of *E. postvittana* using barcodes**

a. Identification result from BOLD system (March 2010, http://boldsystems.org).



Appendix 11 continued:
b. Unrooted neighbour joint (NJ) tree of Tortricinea from BOLD (March 2010).

## Appendix 12

a. Thirty-one synonymous segregating sites (no nonsynonymous site) in cDNA sequences of two *E. postvittana* colony specimens. Each row is an individual with segregating variation indicated by nucleotide positions at the top of the column (the length of the whole cDNA of *E. postvittana* is unknown, however, the exon-intron boundaries are the same as the *C. euritheme* and *M. cinxia*, therefore, nucleotide positions of these two species was applied for the *E. postvittana* cDNA temporarily). Dots represent nucleotides that are identical to the first entry in the column, vertical lines delimit exons boundaries (the sequences are from the end of exon 4 to the beginning of the exon 11, however, the exon 9 does not have any segregating site).

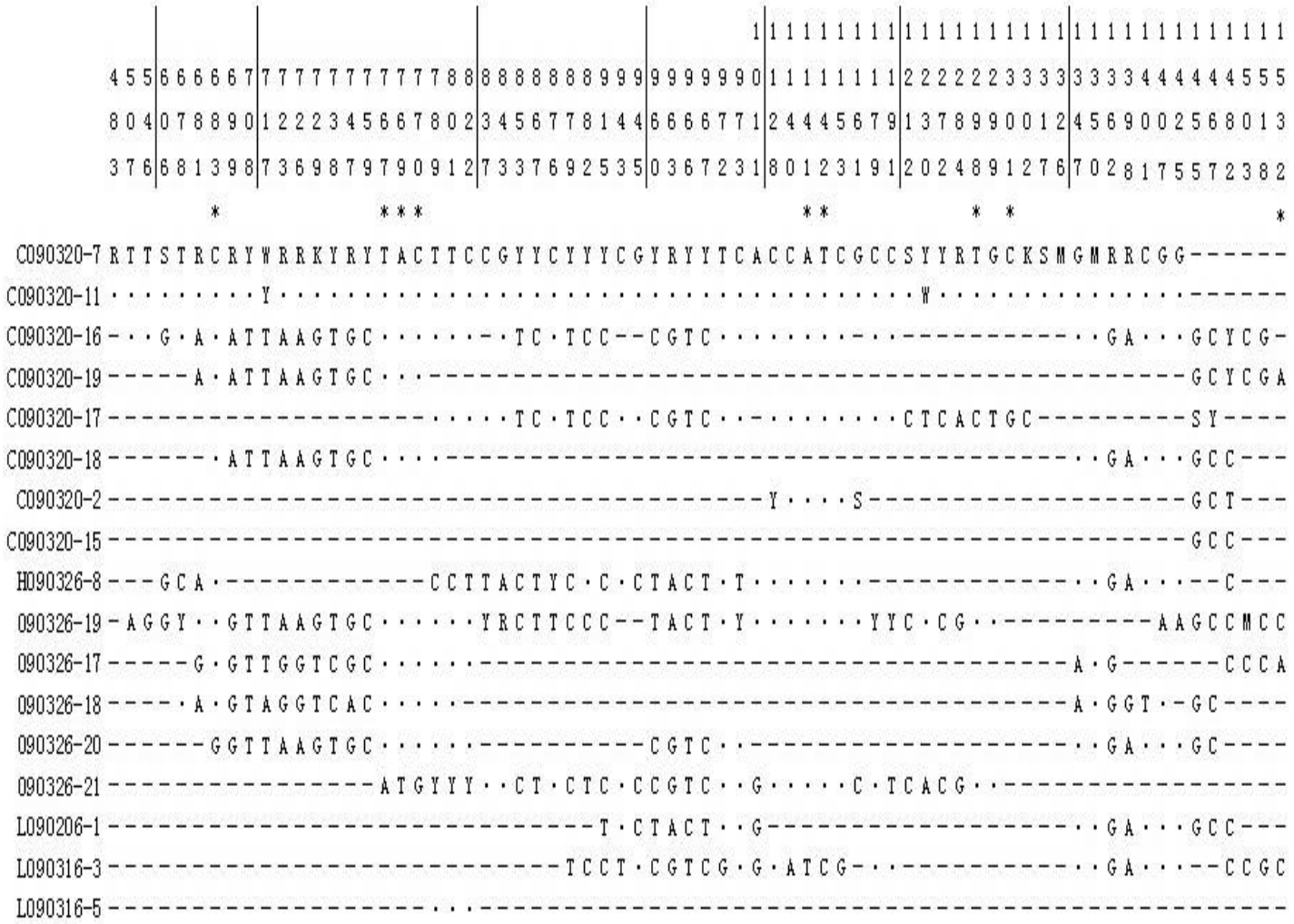
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | | | | | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 4 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 9 | 9 | 9 | 9 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| | 8 | 0 | 8 | 9 | 0 | 1 | 2 | 2 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 8 | 1 | 6 | 6 | 6 | 6 | 1 | 3 | 7 | 8 | 0 | 1 | 2 | 5 | 6 | 9 | | |
| | 9 | 3 | 1 | 9 | 8 | 7 | 3 | 6 | 9 | 8 | 7 | 9 | 3 | 7 | 9 | 2 | 5 | 0 | 3 | 6 | 7 | 2 | 0 | 2 | 4 | 2 | 7 | 6 | 0 | 2 | 8 | | |
| LBAM_cDNA_colony1 | R | S | R | R | Y | W | R | R | K | Y | R | Y | Y | Y | Y | Y | Y | Y | R | Y | Y | S | Y | Y | R | K | S | M | M | R | R | | |
| LBAM_cDNA_colony2 | | | | | | Y | | | | | | | | | | | | | | | | | W | | | | | | | | | | |

b. Amino acid sequence translated from the above cDNA sequences of two colony samples. Since there is no nonsynonymous site, the amino acid sequence of them is the same. Green letters represent locations where the nonsynonymous site in the nucleotide sequences. Red letters represent that the nucleotide changes are not the same in the two cDNA sequences.

1-50 GGSDLGPLMV TEALKPYANH LKVHFVSNID GTHLAEVLKR LDPETALFIV
51-100 ASKTYTTQET ITNATSAKNW FLDAAKDPSA VAKHFVALST NAEKVTAFGI 101-150 DPKNMFGFWD WVGGSYSLWS AIGLSIALHV GYDNFEKLLE GASYMDEHYT 151-200 TAPLEKNAPV ILALLGIWYS NFHGAETHAL LPYDQYLHSF AAYFQQGDME 201-250 SNGKYVTRGG EQVNYSTGPI VWGEPGTNGQ HAFYQLMHQG TSLIPCDFIA 251-300 PAQTHNPISG GQHHKILLAN FLAQTEALMK GKSSEEAKAE LEKSGMAPEA 301-319 VAKILPHKVF KGNRPTNSI

## Appendix 13

A total of 70 segregating sites were found in partial coding region of Pgi gene in *E. postvittana*. Each row is an individual with segregating variation indicated by nucleotide positions at the top of the column (the length of the whole cDNA of *E. postvittana* is unknown, however, the exon-intron boundaries are the same as the *C. eurithem*e and *M. cinxia*, therefore, nucleotide positions in these two species were applied for the *E. postvittana* cDNA temporarily). Dots represent nucleotides that are identical to the first entry in the column, dash lines represent nucleotides at those positions have not been sequences, and vertical lines delimit exons boundaries. Stars indicate the positions of nonsynonymous sites. The name of each sample starts with a capital letter, C = Colony specimen, H = Hawkes Bay specimen, L = Lincoln specimen, those without a capital letter but only number = Clyde specimen.


4556666677777777777788888888999999901111111112222223333333334444445555
8040788901222345667802345677814466666771244456791378990012456900256801337
3766813987369879790912733769253503672318012319120248912767028175572382
C090320-7 RTTSTRCRYWRRKYRYTACTTCCGYYCYYYCGYRYYTCACCATCGCCSYYRTGCKSMGMRRCGG------
C090320-11 ··········Y·····································W·····················------
C090320-16 ···G·A·ATTAAGTGC·····TC·TCC--CGTC·····-·····-·····GA···GCYCG-
C090320-19 -----A·ATTAAGTGC·····-----------------------------------------GCYCGA
C090320-17 ----------------------TC·TCC··CGTC··-·····CTCACTGC----------SY----
C090320-18 -------·ATTAAGTGC·····-----------------------------------------GA···GCC---
C090320-2 ----------------------------------------Y····S-------------------GCT---
C090320-15 ---------------------------------------------------------------GCC---
HD90326-8 ---GCA·----------CCTTACTYC·C·CTACT·T·····-----------GA·····C---
090326-19 -AGGY··GTTAAGTGC·····YRCTTCCC--TACT·Y·····YYC·CG·----------AAGCCMCC
090326-17 -----G·GTTGGTCGC·····-----------------------------------A·G------CCCA
090326-18 -----·A·GTAGGTCAC·····-----------------------------------A·GGT··GC----
090326-20 ------GGTTAAGTGC·····----------CGTC·-----------·····GA···GC----
090326-21 ----------------ATGYYY··CT·CTC·CCGTC··G·····C·TCACG·----------
L090206-1 ----------------------------T·CTACT··G-----------·····GA···GCC---
L090316-3 ----------------------------TCCT·CGTCG·G·ATCG-·----------GA·····CCGC
L090316-5 ----------------------..-----------------------------------------