Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. [http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](http://researchspace.auckland.ac.nz) and [Deposit Licence](http://researchspace.auckland.ac.nz).

Note: Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.
Speciation genes in native New Zealand Leafroller moths

Pia Langhoff

March 2010

Supervisors: Allen Rodrigo
Richard Newcomb
Thomas Buckley

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS OF
DOCTOR OF PHILOSOPHY IN BIOLOGICAL SCIENCE
Abstract

This thesis tests predicted characteristics of speciation genes made by the 'genic view' of speciation. The genic view states that the genome of a species is mosaic, and is divided into regions that contribute to a speciation event (speciation genes) and regions that do not (neutral genes). These speciation loci are characterized by the absence of ancestral polymorphism and diminished or no gene flow. This work investigates whether Pheromone Binding Protein 1 (PBP1) fulfils the criteria of a speciation gene in a complex of native New Zealand Leafroller moths. *Ctenopseustis* and *Planotortrix* comprise five and seven species, respectively. Intraspecific sampling resolved many allelic variants in four neutral genes (*COI*[barcoding region], *COI-COII*, *EF-1α* and *TPI*). These loci show evidence of purifying selection and typically are not reciprocally monophyletic for *P. octo* + *P. excessana* and *C. obliquana* + *C. herana*. PBP1 is no better at recovering monophyly of species groups than any of the neutral genes. The phylogenetic patterns are different among loci confirming a mosaic genome for recently speciated groups. PBP1 is under purifying selection or functional constraint, however there are indications of excess non-synonymous substitutions within PBP1 in *P. octo*, *P. excessana* and *P. avicenniae*. Some of these amino acid changes are in positions predicted to be involved in pheromone binding and receptor interactions. In addition the lineage leading to the species *C. fraterna* and *C. filicis* show evidence of positive selection in the PBP1 gene tree. In conclusion, PBP1 may be a speciation gene for some of the speciation events that gave rise to this species complex.
Acknowledgements

These past four years have been a journey and I have to thank many people that have helped me during this time. I would not have been able to finish this PhD without their help and I would like to show a little bit of appreciation by mentioning them in my thesis.

Firstly I have to thank my supervisors Allen Rodrigo for his enthusiasm and especially financial support in my last year of study. Richard Newcomb for his open door when it comes to discuss science and Thomas Buckley for his reliability and constant support throughout this whole time.

John Dugdale and Robert Hoare for helping me identifying and catching these little critters and their amazing lepidopteran knowledge. Dave Greenwood and Mindy Wang for their help on the GC and Daves expertise on the subsequent analysis.

Anne Barrington for providing me with endless amounts of insects and looking after them. Jill McLaren and Vanessa Mitchell for sending me moths specimen. Robyn Howitt from Landcare for giving me generous amounts of precious DNA samples from their collection. Tim Holmes and Martin Heffer for taking pretty pictures of my moths. Tamara Sirey for letting me use her PBP data.

My labmates Aidan, Andrew, Andi, Caroline, Cyril, Clinton, Colm, Doreen, Edwige, Nadee-sha and Robert of course for fun times in the lab and of course varies social activities that kept the happy-chappy mood in the lab. Special thanks to Yael for being enthusiastic about my thesis, that was much appreciated!

People from the fifth floor still residing or that used to live there, Elena (gracias amiga), Mark aka Mr Andersen, Pete (almost done mate), Elaine (miss bunny-ears), Wei (my yum-char teacher), Kim Plummer (for looking after me even though I wasn’t one of her numerous students) and Gianna (for her happy smile).
My family and friends that have always supported and encouraged me outside my PhD bubble and kept me sane during this time. Especially my dad and Barbara, my grandmas Oma Inge and Friede, my brothers Micky and Matti, my aunt Walli and my friends Heidi, Meik and Gregor. I appreciated your numerous phone calls and little care packages shipped over to me in time of need.

I also like to thank my friends here in New Zealand who have helped me feeling at home and as my second family. Edwige (ultimate-fun-factor), Clinton (very nice), Nikki (my sis), Pia-Pia (No.2) and Tam (aka Dr Tammy). My friends from taiko, Sam-san, Fumi-chan and Brendon and my asahi jins Uncle, Saho, Midori, Lindi and Melissa. My old flatmates at Shackleton, Sean and Yoko, from lovely Waterview, Steph, Ash, Cindy and Ruby and my current flatmates Lauren, Sarah and Scott for looking after me especially in the final days of writing up.

I have to specially thank some people again for their support and effort in getting me where I am today, which is FINISHED and ready to hand in: Edwige, Clinton, Nadeesha thanks guys for doing the final hours countdown with me and helping me with my final analyses and labwork and Tam, my heroine, for helping me with...everything really!

And finally the Royal Society of New Zealand for financing me during the first 3 years of my thesis through a Marsden Grant (HRT203) and Allen Rodrigo for financing the final year of my thesis.
1 General Introduction

1.1 Species Concepts .................................................. 2

1.1.1 Modes of Speciation ........................................... 4

1.2 The genic view of speciation .................................... 6

1.3 Speciation genes .................................................... 9

1.4 Lepidopteran sex pheromones .................................. 13

1.4.1 Sex pheromone biosynthesis in female moth ............... 13

1.4.2 Sex pheromone reception in male moth ..................... 15

1.4.3 Sex pheromone evolution ...................................... 16

1.5 A model system for speciation .................................. 17

1.5.1 Morphology of Planotortrix and Ctenopseustis moths .... 17

1.5.2 Phylogenetics .................................................. 19

1.5.3 Distribution and host plants .................................. 20

1.5.4 Sex pheromones in native Leafroller ....................... 20

1.5.5 Sex pheromone biosynthesis Planotortrix and Ctenopseustis ... 24
## Contents

1.5.6 Pheromone binding proteins in Leafroller moths .......................... 24
1.5.7 Genetic study on behaviour in male *Ctenopseustis* moths ............. 25
1.6 Thesis objectives ............................................................................. 26

2 Evolutionary patterns in the DNA barcoding region of the *Cytochrome c* oxidase subunit 1 gene .................................................. 29

2.1 Introduction ..................................................................................... 29
2.1.1 Aims ............................................................................................ 30
2.2 Materials and Methods ................................................................. 30
2.2.1 Specimens .................................................................................. 30
2.2.2 PCR amplification and sequencing ............................................. 33
2.2.3 Data analysis .............................................................................. 33
2.3 Results ........................................................................................... 34
2.4 Discussion ...................................................................................... 37

3 Evolutionary patterns in non-speciation genes .................................. 40

3.1 Introduction ..................................................................................... 40
3.2 Aims ............................................................................................... 42
3.3 Data limitations .............................................................................. 43
3.4 Materials and Methods ................................................................. 43
3.4.1 Insect collection ......................................................................... 43
3.4.2 Genomic DNA extraction from adult moths ............................... 44
3.4.3 PCR amplification and sequencing ............................................. 44
3.4.4 Sequence alignment .................................................................. 47
3.4.5 Phylogenetic analysis ................................................................. 47
3.4.6 PAML analysis for testing selection ........................................... 48
3.4.7 Constraint analysis and monophyly tests .................................... 49
3.5 Results ........................................................................................... 49
3.5.1 Gene tree phylogenies ............................................................... 49
3.5.2 Tree topology tests .................................................................... 54
3.5.3 Testing sites and lineages under selection .................................. 58
3.6 Discussion ...................................................................................... 60
3.6.1 Discordant tree topologies ........................................... 60
3.6.2 COI, COII and EF-1α are under purifying selection ............... 63

4 A Candidate Speciation Gene 64

4.1 Introduction ................................................................. 64
4.2 Aims ................................................................. 65
4.3 Materials and Methods ....................................................... 66
  4.3.1 Specimen dataset ....................................................... 66
  4.3.2 PCR amplification and sequencing .................................. 66
  4.3.3 Sequence analysis ................................................... 67
4.4 Results ................................................................. 69
  4.4.1 PBP phylogeny ....................................................... 69
  4.4.2 Tests for selection in PBP ......................................... 70
  4.4.3 PBP amino acid variation between Planotortrix and Ctenopseustis . . 73
  4.4.4 A 3D protein structure model of PBP for Planotortrix octo PBP1 ...... 75
  4.4.5 Mapping amino acid changes in C. fraterna and C. filicis onto P. octo model structure ................................................... 78
4.5 Discussion ................................................................. 80

5 General discussion 84

5.1 Sorting of alleles .......................................................... 85
  5.1.1 Taxon sampling and power ......................................... 85
  5.1.2 Effective population sizes ........................................... 86
  5.1.3 Hybridization and Introgression .................................... 87
  5.1.4 Taxonomy ........................................................... 87
5.2 Mosaic genome ........................................................... 88
  5.2.1 Species tree ........................................................ 89
5.3 Selection ................................................................. 89
5.4 Functional work and future directions .................................. 90

A Appendix A 93

B Appendix B 94
## List of Figures

1.1 Dobzhansky-Muller model of hybrid incompatibility ........................................... 6
1.2 The two horizontal bars represent the genomes of two diverging populations .... 7
1.3 Contrasting gene genealogies at two types of loci ............................................. 9
1.4 Adults and larva of *Planotortrix* and *Ctenopseustis* moths .......................... 18
1.5 Distribution maps of *Planotortrix* and *Ctenopseustis* moths .......................... 21

2.1 Species collection sites for the North Island. Each individual dot represents a specimen sampled. ................................................................. 31
2.2 Species collection sites for the South Island. Each individual dot represents a specimen sampled. ................................................................. 32
2.3 Maximum likelihood tree constructed from the barcoding region of the Cytochrome c oxidase subunit 1 (*COI*) gene ........................................... 35
2.4 Comparison of maximal within species divergence with minimal between species divergence at the *COI* ............................................................... 36
2.5 Comparison of maximal within species divergence with minimal between species divergence at the *COI* ............................................................... 37

3.1 Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on combined *COI* and *COII* sequences ........................................ 51
3.2 Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on *EF-1α* sequences ......................................................... 53
3.3 Maximum likelihood gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on *TPI* exon sequences ..................................................... 55
3.4 Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on *TPI* exon and intron sequences ........................................ 56

4.1 Bayesian inference gene phylogeny for *Planotortrix* and *Ctenopseustis* species based on *PBP* sequences ................................................................. 71

4.2 Alignment of PBP1 across *Ctenopseustis* and *Planotortrix* showing only variable amino acid sites ............................................................................. 74

4.3 Structure-based sequence alignment of *P.octo-2* PBP1 ........................................... 75

4.4 Ribbon diagram of the predicted structure of *P.octo-2* PBP1 (blue) and the structure of *A. polyphemus* PBP1 (orange). .................................................. 77

4.5 Ribbon diagram of the predicted structure of *P.octo-2* PBP1 (blue) and the structure of ApolPBP1 (orange) showing fixed amino acid changes between *C. fraterna* and *C. filicis*. .................................................. 79
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Reproductive isolation mechanism</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Distribution and Host plants of <em>Planotortrix</em> and <em>Ctenopseustis</em></td>
<td>22</td>
</tr>
<tr>
<td>1.3</td>
<td>Sex pheromone components and their ratios for <em>Planotortrix</em> and <em>Ctenopseustis</em> species</td>
<td>23</td>
</tr>
<tr>
<td>2.1</td>
<td>Shimodaira-Hasegawa test analysis for topological constraints</td>
<td>36</td>
</tr>
<tr>
<td>3.1</td>
<td>Non-speciation gene dataset</td>
<td>44</td>
</tr>
<tr>
<td>3.2</td>
<td>Additional dataset from (Newcomb &amp; Gleeson, 1998)</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Primer list</td>
<td>46</td>
</tr>
<tr>
<td>3.4</td>
<td>Species monophyly status</td>
<td>57</td>
</tr>
<tr>
<td>3.5</td>
<td>Constraint tree analysis using Shimodaira-Hasegawa test</td>
<td>58</td>
</tr>
<tr>
<td>3.6</td>
<td>Likelihood values and parameter estimates for mt genes <em>COI-COII</em></td>
<td>59</td>
</tr>
<tr>
<td>3.7</td>
<td>Likelihood values and parameter estimates for <em>EF-1α</em></td>
<td>60</td>
</tr>
<tr>
<td>3.8</td>
<td>Nested model results of PAML analysis</td>
<td>60</td>
</tr>
<tr>
<td>3.9</td>
<td>Branch-model results of PAML analysis</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>PBP1 gene dataset</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>Primer list for PBP1</td>
<td>68</td>
</tr>
<tr>
<td>4.3</td>
<td>Likelihood values and parameter estimates for PBP</td>
<td>72</td>
</tr>
<tr>
<td>4.4</td>
<td>Nested model results of PAML analysis</td>
<td>72</td>
</tr>
<tr>
<td>4.5</td>
<td>Species monophyly status</td>
<td>83</td>
</tr>
</tbody>
</table>
Explaining the diversity of species has and remains a major goal in biology. Since Darwin’s *Origin of Species* in (1859), others have contributed to the development of theory on how species might evolve (Grant, 1981; Mayr, 1963; White, 1978). In the early 20th century after the rediscovery of Gregor Mendel’s work on hybridization in peas, genetics became an important factor in understanding how species form (Bateson, 1922; DeVries, 1906). Use of genetics to understand speciation lead to the Modern Synthesis after Dobzhansky published his results on *Drosophila* species (1937b). During the 1980s, speciation and the *Origin of Species* became a more widely researched area due to the field of molecular systematics and the usage of molecular tools to study speciation. Today, scientists use a variety of disciplines to understand the forces that help generate species. These include mathematical theories in population genetics, ecological research on allopatric and sympatric species, molecular analysis such as locating and measuring the genetic effects of genes underlying reproductive isolation and comparative studies using phylogenetics (Coyne & Orr, 2004).
1.1 Species Concepts

In studying speciation it is critically important to consider the definition of a species that will be used. Currently there are more than 25 species concepts each with their own species definition (Coyne & Orr, 2004). The sheer multitude of species concepts suggests that defining a species is not straightforward with little consensus among biologists. Identifying the nature of one’s species problem first helps the choice of the appropriate species concept. Whether species are real or just arbitrary (Darwin, 1859) or “theoretical constructs of the human mind” ((Mayr, 1982)) or whether species are sexually or asexually reproducing organism can all be taken into account to help choose the appropriate species concept. Most species concepts either fall into the category of distinguishing species in populations on reproductive and phenotypic traits or using molecular and phenotypic based phylogenies to resolve a species. Below are examples representing both categories of species concepts.

The most widely used species concept is the Biological Species Concept (BSC) first proposed by Ernst Mayr (Mayr, 1942). Mayr defined species as “groups of actually or potentially inter-breeding natural populations which are reproductively isolated from other such groups” (Mayr, 1963). The critical factor in this concept is Reproductive Isolation (RI), therefore it is sometimes called the Isolation Species Concept. RI through a lack of gene flow between two nascent species is the main driver of a speciation event (Mayr, 1942). Dobzhansky (1951; 1937a) also contributed to the BSC by adding a list of potential barriers to gene flow or isolation mechanisms and also dividing them into three different types (Table 1.1). One of the reasons that the BSC is so widely used may be based on the fact that reproductive isolation in the form of postzygotic isolation is easy to observe in species resulting in sterile or inviable offspring (Hybrid-sterility or inviability).

<table>
<thead>
<tr>
<th>Reproductive isolation mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premating isolation barrier</td>
</tr>
<tr>
<td>Behavioral isolation</td>
</tr>
<tr>
<td>Ecological isolation</td>
</tr>
<tr>
<td>Mechanical isolation</td>
</tr>
<tr>
<td>Mating system ’isolation’</td>
</tr>
<tr>
<td>Temporal isolation</td>
</tr>
</tbody>
</table>

A further species concept is the Recognition Species Concept (RSC) by Paterson (1985)
1.1 Species Concepts

(Lambert & Paterson, 1984). Paterson describes species as groups of individuals that share a common fertilization system or a "Specific Mate Recognition System" (SMRS). Emphasis in this species concept is the sharing of characteristics that hold a population together through a common SMRS and not the absence of a shared mating system on isolation as stressed in the BSC. The recognition of an appropriate mating partner through the SMRS, either active (courtship signals or responses) or passive (as a biochemical process of gamete fusion), and the subsequent signalling during mating is essential to guarantee reproductive success. If recognition and signalling fails then mating is disrupted or prevented. This concept is only useful for sexually reproducing organisms.

The Phylogenetic Species Concept (PSC) is used by those attempting to reconstruct the history of life by identifying historically related groups. Therefore under the PSC “species are clusters that are diagnosably distinct from other such groups” (Cracraft, 1989), where factors such as Reproductive Isolation and SMRS take on less importance. The PSC is based on diagnostic traits such as single nucleotide differences and whether two or more species share a derived character (synapomorphies) relative to an outgroup, if so they are considered a monophyletic group (a group descended from one ancestral species). Speciation therefore consists of the fixation of a diagnostic character in a lineage. Under this phylogenetic concept, speciation occurs faster than biological speciation because the fixation of one new allele is undoubtedly faster than the evolution of RI, which usually requires changes at several loci (Coyne & Orr, 2004).

The Genealogical Species Concept (GSC) by Baum and Shaw (1995) states that “Species are mutually monophyletic in the genealogies at all (or at a consensus of) gene genealogies in the genome”. Shaw (1998) later defined a “genealogical species as a basal, exclusive group of organisms, whose members are all more closely related to each other than they are to any organisms outside the group, and that contains no exclusive group within it”. Exclusivity is generally specified genetically where a group of organism’s loci coalesce more recently within the group than between any member of the group and any organisms outside the group and recognizable as being reciprocal monophyletic in a species phylogeny (Baum & Shaw, 1995). One difficulty is to assess how many loci and alleles are necessary to be reciprocal monophyletic before a group is considered to be a genealogical species under the GSC. The proportion of examined loci allowing a species status is somewhat arbitrary and as Shaw (2001) suggests
1.1 Species Concepts

should include most loci or less than 100% reciprocal monophyletic. This arbitrariness however
is similar when it comes to identify biological species in the BSC when there is gene flow so
overall the GSC seems to provide a reasonable way to recognize species when using phylogenies.

Recently an attempt was made by Queiroz (Queiroz, 2007) to propose a ”Unified Species
Concept” separating species conceptualization from species delimitation. As a common ele-
ment it defines species as ’separately evolving metapopulation lineages’. A lineage refers to an
ancestor-descendant series of a metapopulation extended through time where species are seg-
ments of such a lineage, giving rise to other species. It is therefore the only necessary property
of a species. This general conceptual agreement among the numerous species concepts is con-
sidered as the primary defining property of the species category. The differences and diversity
of species concepts can be accounted for as the secondary species criteria. Attributes that char-
acterize a species concept such as phenetic distinguishability, reciprocal monophyly, pre- and
postzygotic reproductive isolation etc., are all properties that lineages acquire as they separate
and diverge from one another and arise at different times during the process of speciation. All
those properties are relevant to species delimitation and provide different lines of evidence (op-
erational criteria) for lineage separation. Under the unified species concept the presence of a
single property (e.g. monophyly, ecological divergence, reproductive isolation) serves as a line
of evidence or tests that a set of populations possessing that property represents a separate
lineage (i.e., a species). It is a methodological (rather than conceptual) evidence of inferring
the boundaries and numbers of species that is, species delimitation. Therefore any evidence of
lineage separation is sufficient to infer the existence of separate species.

1.1.1 Modes of Speciation

Speciation can be defined as ’the splitting of one species into two through the evolution of
any form of reproductive isolation between taxa’ (Presgraves, 2003). However depending on
the species concept used, a better criterion of whether speciation occurred might be the pres-
ence of stable genetic differentiation at multiple loci considering the possibility of gene flow
among species (Coyne & Orr, 2004; Mallet, 2006). A stable coexistence of one sexual species
with another without the fusion into a single gene pool is possible when species are somewhat
reproductively isolated and also show ecological differentiation (Jiggins et al., 2005).

Speciation in allopatry is characterized by geographical separation of one population into two
populations through the formation of geographic barriers such as a mountain or river. These
two populations develop differently from each other because selective forces such as mutation,
genetic drift and natural selection can create genetic divergence of their gene pools. If these two
groups come into contact again they are no longer compatible in terms of their genetic make-up,
behaviour, mate preference etc. Therefore they have become two different species under this
model. It is possible that hybridisation takes place but only at the contact zone between the
two species (Mayr, 1942).

Potentially speciation can also occur under sympatry (within the same geographical region)
with almost unrestricted gene flow between the two nascent populations, or parapatry where
populations are still connected by gene flow but usually are separated geographically (Mayr,
1963; Wu & Ting, 2004). Sympatric speciation takes place in the same habitat of two incipi-
ent species and often involves resource differentiation producing species with ecologically-based
isolation barriers. Populations of *Rhagoletis pomonella* (apple maggot fly) whose natural host
is the hawthorn tree of North America, for example, have acquired a taste for fruits introduced
from Europe, mainly apples (Bush, 1966; Feder *et al.*, 2005). Associated with this host change
has been a shift in courtship display and subsequently mating. Despite an increasing number of
examples, however, sympatric speciation remains controversial.

If species have been in allopatry and then come back into secondary contact postzygotic isolation
evolves through accumulation of genetic incompatibilities in their gene pools (Dobzhansky,
1940). Postzygotic isolation is first thought to occur geographically through extrinsic factors (e.g.
observed in phenotypes involved in ecological speciation in sticklebacks (Schluter & McPhail,
1992)) and later on biologically through intrinsic factors (e.g. observed as hybrid sterility and
inviability in Drosophila (Orr *et al.*, 2007)). Prezygotic isolation can evolve through reinforce-
ment, which is the evolution of increased prezygotic isolation as a result of selection against
hybrids (Butlin, 1989; Lemmon & Kirkpatrick, 2006)

Mayr also suggested that the whole genome or parts of the genome could be responsible
for reproductive isolation and that genetic changes within each species are strongly coadapted
forming a cohesive unit (Mayr, 1963; Dobzhansky, 1937b). But why, as Darwin already observed
when he wrote the *Origin of Species* (Darwin, 1859), would natural selection which actually acts
to increase fitness cause the evolution of hybrid sterility? Bateson, Dobzhansky (1909; 1937a)
and Muller (1942; 1940) devised a simple model (Dobzhansky-Muller model) for the loss of fitness
between recently diverged species that explains that hybrid inviability and sterility are caused by incompatible epistatic interactions between genes that have functionally diverged between species Figure (1.1). Hybrid incompatibilities therefore can evolve as incidental by-products of adaptive divergence.

Figure 1.1. Dobzhansky-Muller model of hybrid incompatibility. In the ancestral population, the genotype is $AABB$. When the population is split into two, $A$ evolves into $a$ in one population and $B$ evolves into $b$ in the other. $a$ and $b$ are mutually incompatible. In the far right (hybrid), the divergence process is indicated by the black double-headed arrows and the incompatibility is indicated by the green double-headed arrow. Figure (Box 2) taken from (Wu & Ting, 2004)

1.2 The genic view of speciation

Recently a new species concept has been proposed by Wu (2001), known as 'the genic view of speciation'. As the name suggests its major focus is on the genetic basis of speciation. In this concept, species are groups that are differentially adapted and, upon contact, are not able to share genes controlling these adaptive characters (speciation genes), by direct exchanges or through intermediate hybrid populations. These groups may not be differentiated elsewhere in the genome (non-speciation genes) (Wu, 2001). Under this view, the genome is thought to be divided into regions that code or determine aspects of differential adaptation, referred to as speciation loci, and regions that do not (neutral or non-speciation loci) (Wu, 2001).

Selection (sexual or natural) as a positive force favours the adaptation of regions in the genome responsible for a selective advantageous in diverging populations. These loci are differentially adapted while the rest of the genome is not. If speciation genes become linked to these regions they themselves can become differentially adapted resulting in incompatible gene products or alternative alleles and may result in hybrid sterility.

The genetic architecture of barriers to interspecific gene flow and patterns of introgression along with the concept that the genome is somewhat 'porous' with some loci being shared among
taxa and others not is not novel and has been previously addressed in the hybrid zone and plant literature (Barton & Hewitt, 1985; Butlin, 1998; Harrison, 1990; Harrison, 1993; Rieseberg et al., 1999; Rieseberg, 2001).

Speciation is the stage where two criteria are fulfilled: 1) not losing differentiation at specific loci (speciation genes) and 2) still being able to continue to diverge at the rest of the genome.

For example, take a population of grasshoppers living on the northern and southern slopes of a mountain where migration is still possible. The northern slope adapted population has A, B and C alleles whereas the southern slope adapted ones have a, b and c alleles. Due to local selection the two populations would differentiate at the three loci but the rest of the genome would be equally fit in both habitats see Figure 1.2.

![Figure 1.2](image)

**Figure 1.2.** The two horizontal bars represent the genomes of two diverging populations. When they start to differentiate, only a few loci (indicated by black lines) are differentially adapted and genes at such loci are not exchanged between populations (a). Gene flow continues in the rest of the genome (arrows). Although the regions of differential adaptation expand, the amount of gene flow between the two genomes is gradually reduced owing to linkage with such regions (indicated in red/purple) (b,c) until the two populations are completely reproductively isolated and are therefore considered to be separate species (d). Figure (Figure 1) taken from (Wu & Ting, 2004)

This could be considered step (a) in the process of speciation: populations have become differentiated at specific loci responsible for functional divergence e.g. the mate recognition system as in the Zimbabwe and Cosmopolitan behavioural races of *Drosophila* (Wu *et al.*, 1995).
1.2 The genic view of speciation

Upon secondary contact gene flow would be possible but decreased at the loci of functional divergence. As the population diverges more loci will get involved in differentiation and gene flow is prevented by local selection at and near the loci of differential adaptations (see Figure 1.2 (b)). In stage (c) the two populations have further diverged and will either coexist in sympatric environments or form a small hybrid zone in parapatric conditions. The two populations now have divergent gene complexes and are at least distinct at some aspects of reproductive, sexual behaviour and morphology. Gene sharing or gene flow through introgression is still possible and reproductive isolation is not complete yet. They are considered to be 'good species'. *Drosophila simulans* and its two sibling species *D. mauritiana* and *D. sechellia* are such examples showing still incomplete premating isolation but hybrid females are often fertile in F1 and later generations (Coyne, 1984; Wu & Palopoli, 1994). Finally at stage (d) the genomes of both populations are isolated from each other, reproductive isolation is complete. There are three predictions made by the genic view of speciation:

1. The genome is mosaic in terms of its speciation history
2. No gene flow or gene introgression at speciation loci
3. No sharing of ancient polymorphism at speciation loci

The genome is mosaic in terms of its speciation history for each gene. No ancient polymorphism will be shared at this locus because it is believed that this and adjacent loci have undergone a selective sweep. Therefore gene flow and gene introgression is no longer possible at a speciation locus leading to an observed reciprocal monophyly at this locus while gene flow still continues to be possible at non-speciation loci.

Figure 1.3 shows a comparison between speciation and non-speciation loci in the genome which summarizes these predictions made by the genic view of speciation. In Figure 1.3 two different gene genealogies are shown. One locus is referred to as a speciation locus while the other one is a random or non-speciation locus. At the beginning there is one stem species giving rise to three species in the end through two speciation events over time. Gene flow has diminished over time and gene introgression through secondary contact is no longer possible between different species because of incompatible allelic interactions. Due to an elimination of shared ancient polymorphism at this speciation locus the resulting phylogeny for speciation loci is characterized by a clear sorting by species (reciprocally monophyletic), which means all
Figure 1.3. Contrasting gene genealogies at two loci. Two speciation events occurred between species 1, 2 and 3. Gene flow between species boundaries diminished over time and is indicated by arrows. Figure 1 taken from (Ting et al., 2000)

members of a species cluster together and are each others closest relative. Therefore at non-speciation loci, gene introgression, ancient polymorphisms and genetic drift can still be expected to occur resulting in mixed genealogies and no reciprocal monophyly by species.

### 1.3 Speciation genes

Classical genetic studies are performed through backcrosses, F2 and introgression analysis that allowed the mapping and counting of chromosome regions causing reproductive isolation between taxa. These studies elucidated a variety of genetic problems, including Haldane’s rule, faster-male evolution and reinforcement (see review (Coyne & Orr, 2004)). The genetic study of speciation requires a combined analysis of reproductively isolated species with molecular evolution and requires the identification of DNA sequences causing reproductive isolation in species (Orr et al., 2007). Several genes that cause reproductive isolation, also called 'barrier genes', have been identified and are summarized below.

The identification of speciation genes in insects has been predominately conducted in *Drosophila* species due to the ease of genetics in *Drosophila* and the availability of the complete sequence of its genome.
A first candidate gene involved in chemosensory isolation was a desaturase gene (*desat2*) between Cosmopolitan (M-race) and African/Zimbabwe (Z-race) populations of *Drosophila*. The *desat2* locus encodes a Δ9 fatty acid desaturase responsible for cuticular hydrocarbon (CH) polymorphism between these two ecological races of *Drosophila* ([Dallerac et al., 2000](#)). CHs are non-volatile contact pheromones in females controlling male mate responses during mating. African-females seem to prefer to mate with African-males even in the presence of Cosmopolitan-males whereas Cosmopolitan-females readily mate with males from both races ([Takahashi & Ting, 2004](#); [Wu et al., 1995](#)). It appears that the *desat2* locus is one of the loci in *Drosophila* that governs African-female mating characteristics since CH differences have been known to govern female attractiveness in interspecific crosses ([Coyne et al., 1994](#)). Possible ecological adaptations such as lower cold resistance and decreased resistance to starvation were reported for carriers of the Cosmopolitan alleles ([Greenberg et al., 2003](#)). However a reinvestigation ([Coyne & Elwyn, 2006](#)) on those climatic adaptations leaves doubt whether the distributions of the African and Cosmopolitan alleles reflect natural selection involving ecological adaptations.

Another interesting gene, the *Pdd* gene responsible for postdiapause development, causing phenotypic variations in natural populations was recently mapped in the European corn borer moth (ECB), (*Ostrinia nubilalis*) ([Dopman et al., 2005](#)). Previous studies on ECB revealed two different sexually isolated pheromone strains (Z and E). Genetic studies revealed loci that code for pheromone production (*Pher*) as autosomal and loci that code for male behavioural response (*Resp*) as sex linked ([Dopman et al., 2004](#); [Roelofs et al., 1987](#)). Recent studies revealed that the sex-linked gene, *Tpi*, showed pheromone strain exclusivity as the only marker among five different genetic regions in a genealogical analysis ([Dopman et al., 2005](#)). *Resp* was mapped in close proximity to *Tpi* on the sex chromosome however *Tpi* mapped to the exact same position as the major factor *Pdd* affecting differences in postdiapause development time. Due to this linkage with *Tpi* and a possible recent selective sweep, *Pdd* provides a possible candidate speciation gene shaped by natural selection in the European corn borer moth.

The two cases are examples for premating isolation and specifically a form of chemosensory isolation in *Drosophila* since chemical communication during courtship is thought to be mediated by the CHs found in the adults epicuticle. In the case of *Ostrinia* pheromonal divergence is believed to cause partial behavioural isolation but also phenotypic variations by affecting the postdiapause development time.
However the majority of identified speciation genes were identified through detailed studies on hybrid sterility and inviability in *Drosophila*.

Examples of genes that cause postzygotic isolation in hybrids through functional incompatibilities are Hybrid male rescue (*Hmr*) gene which causes male lethality and female sterility and Lethal hybrid rescue (*Lhr*) that produces viable hybrid sons in hybrids of *D. melanogaster* and its sibling species *D. simulans* (Watanabe, 1979; Hutter & Ashburner, 1987). Hybrid inviability involves a recessive X-linked allele from *D. melanogaster* that is incompatible with a dominant autosomal allele from *D. simulans* (Sturtevant, 1920). *Hmr* encodes a protein with homology to a family of MYB-related DNA-binding transcriptional regulators. The recessive X-linked *Hmr* \textsuperscript{mel} allele is incompatible with the dominant autosomal *Lhr* \textsuperscript{sim} causing lethality of hybrid males. Hybrid daughters are viable because they are heterozygous *Hmr* \textsuperscript{mel}/*Hmr* \textsuperscript{sim} (Presgraves, 2006).

The speciation gene *Odysseus* (*OdsH*) is a homeobox gene that causes male sterility in backcross hybrids between *D. simulans* and *D. mauritiana* (Ting et al., 2000; Wu & Ting, 2004). In the coding region of *OdsH*, gene genealogies show species exclusivity between the sibling species of *D. simulans* and *D. mauritiana* (Ting et al., 2000). This species resolution is only achieved at this specific locus. Other gene loci, even adjacent ones, do not show a clear species resolution because of shared ancient polymorphisms (see Figure 1.3).

A Nuclear Core Protein (*Nup96*) is also responsible for hybrid inviability between *D. melanogaster* and *D. simulans* (Presgraves et al., 2003). The *D. simulans* allele of *Nup96* causes lethality in hybrids that are hemizygous for the *D. melanogaster* X chromosome but not for the *D. simulans* X chromosome (Presgraves, 2003; Presgraves et al., 2003). It seems likely that hybrid inviability is due to protein-protein interactions between *Nup96* and one of its eight known protein partners that form the Nup107 subcomplex or one of the three shuttling nucleoporins that dock at the subcomplex. It seems that there is only one nucleoporin encoded by a gene on the X chromosome of *D. simulans*, *Nup153*, that also shows evidence of adaptive evolution in this lineage, hence it might be incompatible with the ancestral *Nup153* from *D. melanogaster* (Presgraves & Stephan, 2007).

However, one of the first speciation genes identified was *Xmrk*-2 between the platyfish *Xiphophorus maculatus* and the swordtail *Xiphophorus helleri* which causes inviability in back-cross hybrids between these two species. Malignant melanomas are induced by a *Tu* gene in
hybrids between \textit{X. maculatus} and \textit{X. helleri} (Gordon, 1927) and controlled by an \textit{R} gene (autosomal), which is a suppressor gene and acts negatively on the expression of \textit{Tu} (X-linked). The \textit{Tu} locus was mapped to \textit{Xmrk-2} which encodes a transmembrane growth factor of the Receptor Tyrosin Kinase Superfamily (Wittbrodt, 1989). Mutations of \textit{Xmrk-2} showed no \textit{Tu} related phenotypes whereas overexpression leads to frequent tumour formations. Tumour formation is only induced in hybrids in the presence of \textit{Xmrk-2} and the absence of the \textit{R} gene.

All the examples given above can be explained by the Dobzhansky-Muller model of intrinsic postzygotic isolation through functional incompatibilities in hybrid species (see Fig 1.1).

However \textit{JYAlpha}, which encodes a male fertility-essential \textit{Na$^+/K^+$} ATPase, was found to cause sterility due to gene transposition from the fourth chromosome of \textit{D. melanogaster} to the third chromosome of \textit{D. simulans} rather than due to adaptive evolution. Certain hybrids between these two species are therefore sterile because they lack \textit{JYAlpha} (Masly \textit{et al.}, 2006). Also recently Machado \textit{et al.} (2007; 2002) have shown that introgression in hybrids of \textit{D. persimilis} and \textit{D. pseudoobscura} are limited to loci adjacent to chromosomal inversions. Noor and colleagues (Machado \textit{et al.}, 2007; Noor \textit{et al.}, 2007) provided further evidence that regions of restricted introgression can be substantially larger than the inverted chromosomal regions. These inverted regions restrict recombination and have been suggested to maintain differentiation between hybridizing species (Butlin, 2005) so that they may represent “islands of differentiation”.

The study of speciation genes involved in postzygotic isolation has revealed some striking patterns. Loci causing postzygotic isolation are not novel genetic factors or processes but ordinary genes. These genes do have a variety of functions including transcriptional factors and structural proteins and they are rapidly evolving (Orr \textit{et al.}, 2007). This rapid evolution is driven by positive natural selection supporting the traditional view of Darwinian adaptation (Coyne & Orr, 2004). Yet recent work on \textit{JYAlpha} has highlighted a different mechanism that can cause hybrid sterility other than the traditional Donzhansky-Muller incompatibility model. Here hybrid sterility is not due to functional incompatibilities but due to gene movement between chromosomes (Masly \textit{et al.}, 2006). The idea that genetic conflict plays a role in speciation through meiotic drive is not new (reviewed in (Coyne & Orr, 2004)). Numerous studies have been conducted that elucidated different forms of meiotic drive such as the role of ‘centromeric drive’ as an example of segregation distortion (Frank, 1991; Henikoff \textit{et al.}, 2001; Henikoff & Malik, 2002; Hurst & Pomiankowski, 1991; Malik & Henikoff, 2001; Montchamp-Moreau \textit{et al.}, 2002).
It is therefore likely that intrinsic postzygotic isolation in species has a variety of causes, including Dobzhansky-Muller incompatibilities, gene transpositions and inversions and further studies on speciation genes might reveal additional mechanisms that could be important in the evolution of reproductive isolation.

1.4 Lepidopteran sex pheromones

Insect species commonly use chemical cues to communicate. Within species these compounds are known as pheromones which, by definition, are substances secreted by an individual and received by a conspecific in whom it elicits a specific reaction (e.g. behavioural) (Karlson & Lüscher, 1959). Pheromone cues are used by many insects especially moths, to attract conspecifics. These compounds are known as sex pheromones. Sex pheromones are intraspecific communication signals between species and have been characterized chemically and behaviourally from over 1500 species of insects of the orders Coleoptera, Diptera, Blattodea and Lepidoptera within the last four decades (Arn et al., 1992).

In order to find a mate during mating, female moths release a plume of volatile sex pheromones produced in special glands in their abdomen. Male moths are able to detect these pheromones via specialised hair-like structures (sensilla) on their antennae, enabling them to locate the female and fly upwind towards the female using the sex pheromone as a cue. Chemical communication in insects is therefore thought to be under strong selective pressure (Löfstedt & Kozlov, 1997), especially in sexual selection which is much stronger in male moths than it is in females because males are generally the ones searching for the opposite sex (Löfstedt, 1990; Phelan, 1992).

Sex pheromones are often blends of volatile hydrocarbons such as acetates, aldehydes and alcohols. Chain length is typically 10-18 carbons and often includes 1-3 double bonds in different configurations Z (cis) or E (trans) (Roelofs & Bjostad, 1984). As well as the blend composition, the relative amounts of different components within the blend are also important for specificity (Tamaki, 1985).

1.4.1 Sex pheromone biosynthesis in female moth

The biosynthetic pathway of sex pheromone production has evolved in the female abdominal glands, between the 8th and the 9th abdominal segments (Bjostad et al., 1987) and includes two important enzymatic steps: desaturation and chain shortening of the pheromone (Roelofs &
1.4 Lepidopteran sex pheromones

Rooney, 2003). In many female Lepidoptera, pheromone components can either be synthesized de novo or derived from dietary (host) precursors. For the later case, for example fatty-acid is used as starting material for leucine derived sex pheromone biosynthesis in *Holomelina spp.* (Arctiidae) (Charlton & Roelofs, 1991). Examples of de novo production of pheromone components can be found in many tortricid species (Bjostad & Roelofs, 1981). However de novo production of sex pheromones is most prevalent within the Lepidoptera.

Key steps in de novo pheromone biosynthesis are:

1. combination of acetyl-CoA carboxylase and FAS (fatty acid synthase) to make 16- and 18-carbon fatty acid precursors in the cytoplasm

2. desaturation with species-specific desaturases of fatty acids with various chain lengths

3. formation of specific oxygenated functional groups via reductases, acetyl-transferases, alcohol oxidases and acetate esterases in the endoplasmic reticulum (ER)

Fatty acyl-CoA desaturases are highly substrate-, regio- and stereospecific and therefore contribute to the production of species-specific sex pheromones. The desaturases active in pheromone biosynthesis are membrane bound and belong to the largest group of desaturases containing three characteristic histidine-boxes (Sperling *et al.*, 2003). The name of the desaturase enzyme reflects its function, for example a ∆9 desaturase places a double-bond at position nine carbons along the carbon backbone of the pheromone from the functional group end of the molecule. Different combinations of different desaturases create the different blends produced by many moths. For example, the pink bollworm *Pectinophora gossypiella* (Gelechiidae) uses a combination of ∆9 and ∆11 desaturases to produce a Z7/Z11- and Z7/E11-16:OAc sex pheromone blend. Other moths such as the female almond moth, *Cadra cautella* (Pyralidae) and the beet armyworm *Spodoptera exigua* (Noctuidae) use a unique ∆12 desaturase to convert Z9-tetradecenoic acid to Z9/E12-tetradecenoic acid.

Insects possess a desaturase multigene family as a result of duplication events of genes involved in normal fatty acid metabolism (Dugdale, 1997; Roelofs & Rooney, 2003) that occurred before the split of Lepidoptera, Diptera and Orthoptera, approximately 350 million years ago (Gaunt & Miles, 2002). The desaturase multigene family consists of at least four gene clusters, all evolving at different evolutionary rates according to their function (Roelofs *et al.*, 2002). In Lepidoptera and Diptera six different gene lineages were characterized. Moth desaturases are
1.4 Lepidopteran sex pheromones

selectively expressed in the pheromone gland and probably active in pheromone biosynthesis except for \( \Delta^9 \) desaturases (Roelofs & Rooney, 2003). The ancestral form of this gene is believed to be a \( \Delta^9 \) desaturase (Roelofs & Rooney, 2003). A highly similar \( \Delta^9 \) desaturase was found in *Drosophila melanogaster* where it is involved in the production of female contact-sex pheromones (Wicker-Thomas *et al*., 1997).

1.4.2 Sex pheromone reception in male moth

Sex pheromones emitted by the female moth are detected via trichoid sensilla on the antennae of males. The sex pheromone diffuses through fine pores on the surface of the sensilla entering the sensillar lymph. In the lymph Pheromone Binding Proteins (PBPs) are thought to specifically bind the sex pheromone components so they are protected from pheromone-degrading enzymes present in the sensillar lymph (Vogt *et al*., 1985). Pheromones bound to PBPs are then transported across the lymph to the dendrites of pheromone sensory neurons of pheromone receptors (PRs) where they trigger a behavioural response. The specificity in PBPs of discriminating and binding components of the sex pheromones is achieved either by the ability of a single PBP to bind different components or isoforms of pheromones as in the case of the two closely related European (*Ostrinia nubilalis*) and Asian (*O. furnacalis*) corn borer, *O. nubilalis* uses a blend of (E,Z)-11-14:OAc while *O. furnicalis* uses a blend of (E,Z)-12-14:OAc (Willett & Harrison, 1999b). Or through the possession of multiple PBPs that each have different affinities for binding major or minor pheromone components as in the case of the giant silkmoth *Antheraea polyphemus*. Three PBPs have been identified ApolPBP1-3 that show different binding affinities for the main pheromone components of (E,Z)-6,11-hexadecadienyl acetate and the (E,Z)-6,11-hexadecadienal (Maida *et al*., 2000).

PBPs in insects were first characterised in the silkmoth, *Antheraea polyphemus* (Saturnidae), (Vogt & Riddiford, 1981) and subsequently in over 20 species of moths (Vogt, 2005). PBPs are a subclass of the Odorant Binding Protein (OBP) multi-gene family that also includes two classes of general OBPs (GOBP1 and GOBP2) and antennal binding protein X (ABPX). The pheromone sensitive sensilla trichodea express PBPs while most sensilla basicona, which respond to general odorants, mainly express GOBPs (Laue *et al*., 1994; Steinbrecht *et al*., 1992). OBPs are small hydrophilic proteins (15 kDa) expressed in sensilla support cells and secreted by these cells into the aqueous sensilla lumen (Vogt *et al*., 2002) where they bind and transport
small hydrophobic ligands. A common feature of these proteins is the positionally conserved presence of six cysteins (Breer et al., 1990) that form disulfide bonds which stabilize the three-dimensional structure of the protein (Leal et al., 1999). So far six insect OBP structures have been resolved (reviewed in (Pelosi et al., 2006)). The PBP of B. mori was the first crystal structure to be determined and revealed a six α-helical conformation. Four of these helices form a binding pocket where the sex pheromone, bombykol, is bound through hydrophobic interactions with aliphatic side chains that line the wall of the pocket (Sandler et al., 2000). Upon interaction with negatively charged sites on or near the pheromone receptor a pH-dependant major conformational change in the structure of PBP has been observed. The C-terminus of the protein folds into a seventh α-helix, enters the binding side of the pheromone and assists in releasing the molecule from the binding cavity (Leal et al., 2005a). The PBP structure of the giant silkmoth, Antheraea polyphemus, also exhibits a pH-induced conformational change with a similar mechanism for releasing the ligand. After a drop in pH, three residues in the binding pocket are protonated and cause a reorientation of three α-helices providing the driving force to release the pheromone from the cavity (Leal et al., 2005b; Zubkov et al., 2005).

Pheromone receptors have only been recently identified in B. mori (Sakurai et al., 2004) and Heliothis virescens and are expressed in the male antennae in cells located in close proximity with the long sensilla trichodea (Krieger et al., 2004). PRs are members of the insect olfactory receptor family and believed to contain seven putative membrane-spanning regions (Vosshall, 2003). The pheromone receptor, once activated by a bound pheromone, transmits a signal through the antennal nerve to the olfactory area of the insect brain (Anton & Homberg, 1999). This signal initiates a behavioural response in the male moths by orienting himself directly upwind to the female (Vickers & Baker, 1994).

1.4.3 Sex pheromone evolution

The production and release of different sex pheromone blends guarantees specificity between different species making it possible for male moths to differentiate between hundreds of different pheromone blends. It also implies that a change in this diverse and specific receptor system in male moths can lead to speciation events. A pheromone shift through differences in pheromone blend may have initiated a speciation process in nine European species of small ermine moths (Yponomeutidae) (Löfstedt et al., 1991). The question of how a new pheromone signal spreads
and eventually leads to speciation is addressed by different theories. One theory by Phelan (1992) is called 'asymmetric tracking' and assumes the selection on the signal to be weak whereas 'tracking' changes in the signal is thought to undergo positive selection. Asymmetric tracking refers to differential parental investment as a critical factor in the evolution of mate-signalling systems. The selection on signal and response will be asymmetric because the sex that invests more in offspring will invest less in mate-finding (Phelan, 1992). Females in moths are the signallers so selection will be weaker than in male moths since they track evolutionary changes in the signal. Liu and Haynes (1994) showed in mutants of the cabbage looper Trichoplusia ni, that expressed different ratio of pheromone blends than the wild type, that males were able to track down the female pheromone changes in the mutant strain. Mutant males initially responded poorly to the 'new' blend but showed an equally good response to the 'new' blend as to the 'normal' blend after 49 generations. Another theory postulates the importance of premating isolation mechanism as a cause of evolutionary change in moth sex pheromones (Löfstedt, 1993). The case of Yponomeuta shows that hybridizing in the laboratory produces viable offspring (Hendrikse, 1988), however, there is no cross-attraction between species in sympatry and therefore no hybridization because of different pheromone blends (Löfstedt et al., 1991).

1.5 A model system for speciation

1.5.1 Morphology of Planotortrix and Ctenopseustis moths

The New Zealand Leafroller moths (Tortricidae: Tortricinae, Archipini) are a large family within Lepidoptera, with more than 5000 species described worldwide for the family of Tortricidae (Dugdale, 1997). The two morphologically homogeneous genera Planotortrix and Ctenopseustis are both indigenous to New Zealand. The polyphaga larvae of some of their species have a major impact as horticultural pests in New Zealand (Wearing et al., 1991) (Figure 1.4). The genus Planotortrix, also known as the green-headed Leafroller (GHLR) contains seven species, while the genus Ctenopseustis or brown-headed Leafroller (BHLR) contains five species (Dugdale, 1990). Prior to their revaluation in 1990 by Dugdale, both Planotortrix and Ctenopseustis contained only three species, P. notophaea, P. flammea, P. excessana and C. obliquana, C. servana, C. fraterna, respectively. A morphological revision following extensive study of their sex pheromone system revealed that each genus contains sibling or cryptic species (Foster et al.,
1.5 A model system for speciation

1986; Foster & Roelofs, 1988; Foster & Roelofs, 1987; Foster & Dugdale, 1988; Foster et al., 1989; Foster et al., 1990; Galbreath et al., 1985).

Figure 1.4. Adults and larva of Planotortrix and Ctenopseustis. Adult moths are between 1-1.5cm in length, while larvae are approximately up to 2.5cm as 5th instars. Colour and wing colouration in adults vary according to distribution not due to sexual dimorphism between male and female moths. Females are generally larger than males.

Morphologically the genus Planotortrix falls into two main groupings, the 'notophaea' group containing (P. notophaea, P. flammea) and P. puffini and the 'excessana' group containing (P.
1.5 A model system for speciation

excessana, *P. octo*) and (*P. avicenniae, P. octoides*) (Dugdale, 1997). Within the 'notophaea' group three synapomorphies (shared derived characters) define their arrangement. Within the structurally homogenous 'excessana' group, *P. octoides* and *P. avicenniae* are defined on the reduction of synapomorphies, whereas *P. excessana* and *P. octo* are defined only by symplesiomorphies (shared ancestral characters). This leads to a rather uninformative or paraphyletic arrangement for the latter. As a result, *Planotortrix* consists of seven described species with *P. octo*, *P. octoides* and *P. puffini* as newly described. In *Ctenopseustis*, *C. servana* is sister group to all other species with *C. obliquana* and *C. herana* being sister taxa. *C. fraterna* and *C. filicis* are both fern-feeding segregates of *C. obliquana* and *C. herana* (John Dugdale, personal communication). These relationships are based on an intuitive phylogeny by John Dugdale (Dugdale, 1997).

1.5.2 Phylogenetics

So far only a few genetic studies have been conducted for this group mainly to elucidate phylogenetic relationships among the species. Isozyme markers were studied in *Planotortrix* and *Ctenopseustis* revealing substantial genetic differences among the taxa although differences were not fixed between many of the sibling taxa (White & Lambert, 1995). These findings did reflect the differences in sex pheromone use (White & Lambert, 1994). Restriction fragment length variations of ribosomal DNA distinguished between the sibling species pairs of *P. octo*, *P. excessana* and *C. herana*, *C. obliquana* using laboratory-reared specimen (Sin et al., 1995). Genetic studies on wild populations revealed a high degree of stability in distinct gene frequencies. Neither of these studies, however, provided sufficient characters to resolve all phylogenetic relationships within the two genera.

A molecular phylogeny for the group based on partial sequence of the mitochondrial cytochrome c oxidase subunit I gene (*COI*) was inferred for both genera *Planotortrix* and *Ctenopseustis* (Newcomb & Gleeson, 1998). The analysis of the *COI* sequences supports the monophyletic status of *Ctenopseustis* based on morphological characters, with *C. servana* being ancestral. Within the genus *Planotortrix*, the 'excessana' grouping, including *P. excessana*, *P. octo*, *P. avicenniae*, *P. octoides* was supported. However the 'notophaea' grouping including *P. notophaea*, *P. flammea* and *P. puffini* was not supported by this small dataset. A further two new genera, *Apoctena* and *Leucotenes*, using *Epiphyas postvittana* as an outgroup, were also described and thought to
be closely related to *Planotortrix* and *Ctenopseustis*. A study by Gleeson *et al.* (2000) used a larger dataset using most of the *COI* and *COII* genes to study a similar set of taxa. The monophyly of both *Planotortrix* and *Ctenopseustis* was confirmed. However the closely related genus, *Apoctena*, was found to be polyphyletic. Their research revealed 36 unique nucleotide synapomorphies within *Planotortrix* and 64 for *Ctenopseustis* and advocates the utility of a phylogenetic approach to develop a diagnostic key for this group in the future.

### 1.5.3 Distribution and host plants

The geographic distribution for *Planotortrix* and *Ctenopseustis* is quite diverse. *C. obliquana* are found all over New Zealand, including Stewart Island, while endemic *P. octoides* inhabits only the Chatham Islands and *P. avicenniae* is found only on the mangrove species, *Avicennia resinifera* on the North Island above 38° S latitude. One of the sibling species pairs, *Planotortrix octo* is found on both the North and South Islands and is particulary important in the eastern regions of Poverty Bay, Hawkes Bay, Marlborough, Canterbury and Central Otago. Whereas *Planotortrix excessana* is also found on both North and South islands, especially around the Nelson and Waikato area, but rare or infrequent in the eastern regions of the country. Both genera *Planotortrix* and *Ctenopseustis* include a range of polyphagous and monophagous species. Within *Planotortrix*, *P. excessana*, *P. octo*, *P octoides* and *P. notophaea* are polyphagous. *P. flammea* is found on *Scrophulariaceae* and *P. puffini* predominately on ‘muttonbird scrub’ *Senecio reinoldii*. Within *Ctenopseustis*, *C. obliquana* and *C. herana* are polyphagous, as is *C. servana* but only on coastal angiosperms. *C. fraterna* and *C. filicis* are fern-feeders and found only on *Pteridophyta* species (Dugdale, 1990) (Table 1.2) and (Figure 1.5).

### 1.5.4 Sex pheromones in native Leafroller

*Ctenopseustis* and *Planotortrix* are characterized by the use of unusual sex pheromones among Tortricinae (Table 1.3) but more importantly sex pheromones can be used as an important chemotaxonomic tool for this group. Certain morphologically indistinguishable populations of supposedly the same species of moths use distinct sex pheromones and are therefore considered sibling species as in the case of *C. obliquana* and *C. herana* (Dugdale, 1990; Foster *et al.*, 1991). A combination of morphological and pheromonal characters therefore has been proven to be suitable for reliable species identification.
Figure 1.5. Distribution maps of *Planotortrix* and *Ctenopseustis*. Green colouring refers to areas where species are commonly found. See also Table 1.2
Table 1.2. Distribution and Host plants of Planotortrix and Ctenopseustis

<table>
<thead>
<tr>
<th>Species</th>
<th>Host</th>
<th>Distribution*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ctenopseustis obliquana</em></td>
<td>Polyphagous</td>
<td>NI, SI and StI</td>
</tr>
<tr>
<td><em>Ctenopseustis herana</em></td>
<td>Polyphagous</td>
<td>SI, StI and CI</td>
</tr>
<tr>
<td><em>Ctenopseustis fraterna</em></td>
<td>Monophagous on Pteridophyta</td>
<td>NI</td>
</tr>
<tr>
<td><em>Ctenopseustis filicis</em></td>
<td>Monophagous on <em>Dicksonia squarrosa</em></td>
<td>SI and StI</td>
</tr>
<tr>
<td><em>Ctenopseustis servana</em></td>
<td>Polyphagous on coastal angiosperms</td>
<td>NI and 3KI</td>
</tr>
<tr>
<td><em>Planotortrix excessana</em></td>
<td>Polyphagous</td>
<td>NI, SI and StI</td>
</tr>
<tr>
<td><em>Planotortrix octo</em></td>
<td>Polyphagous</td>
<td>NI and SI</td>
</tr>
<tr>
<td><em>Planotortrix notophaea</em></td>
<td>Polyphagous</td>
<td>NI, SI, StI and 3KI</td>
</tr>
<tr>
<td><em>Planotortrix avicenae</em></td>
<td>Monophagous on <em>Avicennia resinifera</em></td>
<td>NI</td>
</tr>
<tr>
<td><em>Planotortrix octoides</em></td>
<td>Polyphagous</td>
<td>CI</td>
</tr>
<tr>
<td><em>Planotortrix puffini</em></td>
<td>Mainly on <em>Brachyglottis reinoldii</em></td>
<td>SI and StI</td>
</tr>
<tr>
<td><em>Planotortrix flammea</em></td>
<td>Mainly on <em>Scrophulariaceae</em></td>
<td>NI and SI</td>
</tr>
</tbody>
</table>

*NI = North Island, SI = South Island, StI = Stewart Island, CI = Chatham Islands, 3KI = Three-Kings Islands

*Ctenopseustis* and *Planotortrix* species use mono-unsaturated fourteen carbon acetates except for the fern-feeders *C. fraterna* (Bruce Morris, personal communication) and *C. filicis* which use sixteen carbon acetates, all pheromones are used either singly or as a blend (Foster et al., 1986). These components are distinct from the Δ11-unsaturated tetradecenyl acetates, alcohols or aldehydes commonly used by species of the tribe Archipini (Roelofs & Brown, 1982). The sex pheromones of *Ctenopseustis* and *Planotortrix* species differ only in the unusual position of double bonds placed by the desaturase enzymes. All double bonds are in cis (Z) configuration. Females of the genera *Planotortrix* and *Ctenopseustis* use Δ5, Δ7, Δ9, Δ10 and Δ11 desaturases (Foster & Roelofs, 1996) to produce sex pheromone components such as (Z)-5-tetradecenyl acetate (Z5-14:OAc), (Z)-7-tetradecenyl acetate (Z7-14:OAc), (Z)-8-tetradecenyl acetate (Z8-14:OAc), (Z)-9-tetradecenyl acetate (Z9-14:OAc) and (Z)-10-tetradecenyl acetate (Z10-14:OAc). *C. obliquana* is divided into three different pheromonal types. Type I and III are only distinct in the ratio of produced sex pheromones by the female and their distribution. Type I uses a Z8-14:OAc and Z5-14:OAc blend in a ratio of 80:20 and is located on the North Island while type III uses a 90:10 ratio blend and is only found on the South Island. *C. obliquana* Type II is exclusively found on the North Island, morphologically identical to Type I and III but the females produce the same sex pheromone, Z5-14:OAc, as *C. herana* females. Table 1.3 is a summary of sex pheromone components and ratios of different blends used by members of the genera *Ctenopseustis* and *Planotortrix*.
Table 1.3. Sex pheromone components and their ratios for *Planotortrix* and *Ctenopseustis* species

<table>
<thead>
<tr>
<th>Species</th>
<th>Sex pheromone components</th>
<th>Ratios</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ctenopseustis obliquana</em></td>
<td>Z5-14:OAc, Z8-14:OAc</td>
<td>4:1</td>
<td>(Young <em>et al.</em>, 1985)</td>
</tr>
<tr>
<td><em>Ctenopseustis obliquana</em></td>
<td></td>
<td>90:10</td>
<td>(Foster <em>et al.</em>, 1986)</td>
</tr>
<tr>
<td>Ctenopseustis obliquana$^2$</td>
<td>Z5-14:OAc</td>
<td>80:20</td>
<td>(Foster <em>et al.</em>, 1986)</td>
</tr>
<tr>
<td><em>Ctenopseustis herana</em></td>
<td>Z5-14:OAc</td>
<td>-</td>
<td>(Foster &amp; Roelofs, 1987)</td>
</tr>
<tr>
<td><em>Ctenopseustis fraterna</em></td>
<td>16 carbon acetate diene</td>
<td>-</td>
<td>(pers. comm. Bruce Morris)</td>
</tr>
<tr>
<td><em>Ctenopseustis filicis</em></td>
<td>Z10-16:OAc, Z8-14:OAc</td>
<td>-</td>
<td>(Foster &amp; Dugdale, 1988)</td>
</tr>
<tr>
<td><em>Ctenopseustis servana</em></td>
<td>Z5-14:OAc, Z7-14:OAc$^3$</td>
<td>32:68 to 35:65</td>
<td>(Foster &amp; Dugdale, 1988)</td>
</tr>
<tr>
<td><em>Ctenopseustis servana</em></td>
<td></td>
<td>27:73</td>
<td>(Foster <em>et al.</em>, 1990)</td>
</tr>
<tr>
<td><em>Planotortrix excessana</em></td>
<td>Z5-14:OAc, Z7-14:OAc$^1$</td>
<td>60:40</td>
<td>(Foster <em>et al.</em>, 1989)</td>
</tr>
<tr>
<td><em>Planotortrix excessana</em></td>
<td></td>
<td>3:97 to 71:29</td>
<td>(Galbreath <em>et al.</em>, 1985)</td>
</tr>
<tr>
<td><em>Planotortrix octo</em></td>
<td>Z8-14:OAc, 14:OAc$^1$</td>
<td>-</td>
<td>(Galbreath <em>et al.</em>, 1985)</td>
</tr>
<tr>
<td><em>Planotortrix octo</em></td>
<td>Z8-14:OAc, Z10-14:OAc$^1$</td>
<td>98:2</td>
<td>(Foster <em>et al.</em>, 1986)</td>
</tr>
<tr>
<td><em>Planotortrix notophaea</em></td>
<td>Z7-14:OAc$^3$ + ?</td>
<td>-</td>
<td>(Foster <em>et al.</em>, 1986)</td>
</tr>
<tr>
<td><em>Planotortrix avicenniae</em></td>
<td>Z5-14:OAc$^1$</td>
<td>-</td>
<td>(Foster &amp; Roelofs, 1987)</td>
</tr>
<tr>
<td><em>Planotortrix octoides</em></td>
<td>Z8-14:OAc$^3$</td>
<td>-</td>
<td>(Dugdale, 1990)</td>
</tr>
<tr>
<td><em>Planotortrix puffini</em></td>
<td>Z5-14:OAc, Z7-14:OAc, Z9-14:OAc$^3$</td>
<td>3:97:2</td>
<td>(Foster &amp; Dugdale, 1988)</td>
</tr>
<tr>
<td><em>Planotortrix flammea</em></td>
<td>Z5-14:OAc, Z7-14:OAc$^3$</td>
<td>52:48, 61:39</td>
<td>(Foster <em>et al.</em>, 1990)</td>
</tr>
</tbody>
</table>

$^1$ identified as active sex pheromone  
$^2$ Type II North Island variation  
$^3$ pheromone gland components only
1.5.5 Sex pheromone biosynthesis Planotortrix and Ctenopseustis

Sex pheromones in the native Leafroller are singly desaturated at various positions, except for C. fraterna that uses a diene, and only differ at the position of a single double-bond. To elucidate the biosynthetic pathways, labelling experiments with deuterium precursors were undertaken in P. octo and P. excessana and C. herana (Foster & Roelofs, 1988; Foster & Roelofs, 1996; Foster, 1998). The sex pheromone Z8-14:OAc in P. octo is synthesized via a Δ10 desaturation of palmitate to form (Z)-10-hexadecanoate that undergoes a single round of β-oxidation to form a (Z)-8-hexadecanoate sex pheromone precursor. This unusual Δ10 desaturase activity was confirmed by Hao et al. (2002). Both sex pheromone components of P. excessana, Z5-14:OAc and Z7-14:OAc, are synthesized by β-oxidation of the common Δ9-desaturated fatty acids, oleate and palmitoleate. In C. herana the sex pheromone Z5-14:OAc is synthesized via a Δ5-desaturation of myristic acid. It might be possible though that Z5-14:OAc components are synthesized via chain shortening of oleic acid, as it is proposed for C. obliquana which uses the same component as part of its sex pheromone blend. It appears that there are different possible biosynthetic routes to the same components in these Leafrollers. Either through unique desaturation, as in P. octo with a Δ10 desaturase, or alternative routes when it come to more commonly used components as in P. excessana (Foster, 1998). For example, to get a Z5-14:OAc sex pheromone three alternative biosynthetic pathways are possible. First the Δ5-desaturation of a 14C, or the Δ7-desaturation of a 16C and one round of β-oxidation or a Δ9-desaturation of a 18C and two rounds of β-oxidation.

A recent molecular study by Sirey et al. (2006) revealed at least six distinct groups of desaturase genes isolated from three species of Planotortrix and two species of Ctenopseustis in the native Leafroller moths. The desaturase genes showed differences in transcriptional regulation between species and specific desaturase expression was correlated with sex pheromone production. Preliminary results showed the presence of desaturase activities that were specific to sex pheromone production and not general lipid metabolism (Sirey, 2006).

1.5.6 Pheromone binding proteins in Leafroller moths

PBPs are hypothesized to play a role in pheromone reception including transport and discrimination, receptor activation and pheromone degradation. A molecular study on pheromone binding proteins (PBPs) in Epiphyas postvittana (Newcomb et al., 2002) and all twelve species
1.5 A model system for speciation

of Planotortrix and Ctenopseustis (Sirey, 2000) investigated the role of PBPs in pheromone discrimination. In Epiphyas postvittana two PBP genes have been described, PBP1 and PBP2. PBP1 has two different allelic versions and binds to the major component of the sex pheromone in E. postvittana. From antennal ESTs there appears to be a third PBP. It is assumed that there are therefore likely to be at least three pheromone binding proteins present in the native Leafrollers (Newcomb et al., 2002). PBPs in the native Leafrollers are capable of gross discriminatory abilities when it comes to bind the sex pheromone molecule but are not capable of discriminating between different double-bond positions of sex pheromone molecules. However a phylogenetic comparison based on PBP sequence data confirmed the monophyletic status of both genera and also revealed a grouping of C. fraterna and C. filicis which correlates with their sex pheromone blend of a 16-carbon acetate as opposed to 14-carbon acetate as in all the other species of both genera (Sirey, 2000) (see Table 1.3).

1.5.7 Genetic study on behaviour in male Ctenopseustis moths

In an attempt to understand how many genes are involved in sex pheromone-based mating systems in these indigenous Leafroller species an interbreeding approach of closely related species in Ctenopseustis has been conducted. C. herana and C. obliquana are sibling species and differ in the usage of their sex pheromones, Z8-14:OAc and Z5-14:OAC for C. obliquana and Z5-14:OAc in C. herana. Hybrids of crosses between both species have been analysed by male olfactory sensillum response to sex pheromones and female production of sex pheromones (Foster et al., 1997). In each of the species the male responded strongly to the main component of the pheromone (Z8-14:OAC in C. obliquana and Z5-14:OAc in C. herana) but also showed a weaker response to the other pheromone. In the hybrids the responses were more variable than those of the parental moths. The response pattern likely reflects a Z-linked inheritance with the C. herana type being dominant (Hansson et al., 1989). The male behavioural responses to the sex pheromone fits the pattern of Z-linked inheritance and C. herana type dominance, therefore the Z-linked male olfactory and behavioural responses are suggested to be linked (Foster et al., 1997). Sex pheromone production in the female revealed that C. herana only produces Z5-14:OAc and C. obliquana only Z5-14:OAc and Z8-14:OAc in a ratio of 20:80. The hybrid crosses gave conflicting results in the F2 generations where it appeared that the differences in the pheromone production could be autosomally linked to a single gene. The paternal backcross displayed a
sex-linked genetic factor and the maternal backcross didn’t fit either of them (Foster et al., 1997). It is therefore concluded that a single gene inheritance cannot explain the complexity of the pheromone production observed in hybrids between *C. herana* and *C. obliquana*.

### 1.6 Thesis objectives

The major focus of this study is to test ‘the genic view of speciation’ made by Wu (2001) for the genera *Planotortrix* and *Ctenopseustis* using a phylogenetic approach. These native New Zealand Leafroller moths are well-studied and represent a suitable system to test the ‘genic view of speciation’ because it is a system where taxa have diverged recently and the genes involved in pheromone production and reception are beginning to be well understood.

There are three hypotheses of the genic view that are to be tested in this thesis.

- **Hypothesis 1**: The genome is mosaic with respect to the species history

  Closely related species often reflect a low level of genetic divergence. Introgressive hybridization and lineage sorting of ancient polymorphisms are still in process especially if divergence occurred recently. Therefore assessing a species phylogenetic history should include the comparison of multiple loci and multiple sequences per species (Beltran et al., 2002; Dopman et al., 2005; Machado & Hey, 2003; Ting et al., 2000). Genealogical analysis of two mitochondrial marker (*COI-COII*) and two nuclear marker (*EF1α* and *TPI*) with multiple sequences as examples of non-speciation loci were used to access whether they reflect a mosaic genome through a gene tree phylogenetic approach (chapter 2 and 3). It is expected that all chosen genes show exclusivity for some species but not for all of them. A mosaic genome is reflected by showing monophyly in some species groups and the fact that different genes show different species groups as monophyletic because of different expected coalescence times among loci. It is also expected that, due to the genetic effective population size the mitochondrial genes will coalesce on average four times faster than nuclear genes, that mt genes possibly show more species as being an exclusive group than the non-speciation genes. Therefore, the prediction is that monophyly of species may be seen in *COI-COII*, but these groups may not be concordant with groups in *EF1α* and *TPI* (Chapter 3).

- **Hypothesis 2**: Genomic regions such as speciation loci should show closely related species as exclusive groups, that is they reflect a reciprocal monophyly in a genealogical tree
1.6 Thesis objectives

Genealogical tests in form of phylogenies are able to detect loci that contribute to rapid adaptive divergence and ultimately reproductive isolation due to the elimination of shared ancient polymorphisms and/or gene introgression (Ting et al., 2000). It is expected that a gene involved in speciation reflects a clear monophyletic pattern for all species in a gene tree. A candidate speciation locus, a gene involved in pheromone binding (PBP1) is sampled and i will test whether it recovers exclusivity of the group in form of monophyly by species better than a non-speciation locus (Chapter 4).

- Hypothesis 3: Speciation genes are under positive selection and have possibly undergone selective sweeps

Speciation genes have been found to evolve under positive natural selection (Barbash et al., 2004; Presgraves et al., 2003). Gene flow at a speciation locus is often decreased or absent and often characterized by 'selective sweeps' (Dopman et al., 2005; Ting et al., 2000).

Tests for selective pressure through synonymous vs. non-synonymous ratios are performed to investigate incidences of positive selection on the PBP1 gene (Chapter 4). Due to functional constraints a non-speciation gene is expected to be under purifying selection whereas speciation genes are expected to diverge through positive selection. It is important to note that these hypotheses are tested using a phylogenetic approach, rather than a population genetic one. I believe that this is justified because the predictions of the genic view of speciation are primary phylogenetic predictions that predict reciprocal monophyly as a necessary pattern for speciation genes. Population genetics on the other hand provides us with an insight of how genes sort. The principal role of population genetics in tests of our hypotheses in this thesis is the determination of whether reciprocal monophyly can be observed by chance alone, given the parameters of the population under study (i.e. effective size and sample size). I address these issues in the following chapters where appropriate.
This chapter has been published in Molecular Ecology Resources and is attached under Appendix A on the CD.

Evolutionary patterns in the DNA barcoding region of the Cytochrome c oxidase subunit 1 gene

2.1 Introduction

In an attempt to diagnose taxa and assess species diversity universally, a technique called DNA barcoding has been introduced that uses short DNA sequences as a molecular diagnostic for species-level identification (Blaxter, 2003; Marshall, 2005). DNA barcoding sequences are very short and for many animals stem from the mitochondrial gene region of cytochrome c oxidase subunit 1 (COI). A number of studies have shown that COI sequences have a low variability and closely related species differ by several % suggesting their suitability for species identification (Hebert et al., 2003b; Hebert et al., 2003a; Smith et al., 2006; Vences et al., 2005; Ward et al., 2005; Waugh, 2007). Lepidoptera is a diverse order, containing over one hundred thou-
sand species of moths alone (Roelofs & Rooney, 2003). Speciation within families and genera can be rapid, with morphologically identical (cryptic) species complexes often observed. DNA barcoding has shown that it is useful for identifying morphological identical or cryptic species complexes, for example, in families of neotropical skipper butterflies (Hebert et al., 2004), sphinx moths and wild silk moths (Hajibabaei et al., 2006).

Phylogenetic relationships for the Leafroller moths of the genera Planotortrix and Ctenopseustis have been estimated previously based on morphological, pheromonal and molecular characteristics (Dugdale, 1997; Gleeson et al., 2000). These studies have shown that molecular markers such as mitochondrial DNA (mtDNA) genes, in particular, cytochrome c oxidase subunit 1 and 2 (COI and COII) are able to separate Ctenopseustis and Planotortrix as monophyletic genera. However these studies have only used a single specimen in some taxa and whether species are reciprocal monophyletic with a larger number of individuals per species has to be tested.

2.1.1 Aims

Does the short DNA barcoding region show exclusivity or reciprocal monophyly for each species group? Can COI be used as a suitable marker for species exclusivity on a large dataset of multiple specimen per species?

To test these two hypotheses a phylogenetic analysis of 200 specimens of New Zealand Leafroller moths from the genera Ctenopseustis and Planotortrix (Tortricidae) for the barcoding region within mitochondrial cytochrome c oxidase subunit 1 gene (COI) was conducted.

2.2 Materials and Methods

2.2.1 Specimens

Moths of the genera Ctenopseustis and Planotortrix were supplied either by the insect rearing facility at HortResearch, collected using pheromone traps or collected as larvae, reared on artificial diet (Singh, 1974) to adults and identified morphologically by John Dugdale or Robert Hoare. Species ID details, such as collection site, langitude/latitude information and Genbank accession numbers are given in Appendix A (supplemented on a CD). Maps of the localities for all 200 collected specimen are shown in Figure 2.1 and 2.2.
Figure 2.1. Species collection sites for the North Island. Each individual dot represents a specimen sampled. Regional codes are taken from Crosby et al. (1998)
2.2 Materials and Methods

Figure 2.2. Species collection sites for the South Island. Each individual dot represents a specimen sampled. Regional codes are taken from Crosby et al. (1998)
2.2 Materials and Methods

2.2.2 PCR amplification and sequencing

Genomic DNA was extracted from whole moths using the DNeasy® Tissue Kit (Qiagen), following the manufacturer’s instructions.

A 468 bp fragment, the ‘barcoding’ region, of the COI gene was amplified by Polymerase Chain Reaction (PCR) using the primers CI-J-1718 (5'-GGAGGATTTGGAAATTGATTAGTTCC-3') and CI-N-2191 (5'-CCCGGTAAATTTAAAATAAACTTC-3') (Simon et al., 1994). This region encodes the third, fourth, fifth and sixth of twelve transmembrane helices, the first and the second internal and the second and third external loops of the COI protein (Gleeson et al., 2000). Amplification reactions of 50 µL consisted of 10 pmol of each primer, 20 mM Tris-HCl (pH 8.4), 1.5 mM MgCl₂, 50 mM KCl, 0.2 mM of each dNTP and 2 units of Taq polymerase (Invitrogen). The cycling conditions consisted of an initial denaturation at 94°C for 2 min, followed by 30 sec at 94°C, annealing at 40°C for 30 sec, and extension at 72°C for 1 min for 35 cycles. A final extension step at 72°C for 10 min was performed to ensure PCR products were full length.

The resulting PCR products were purified using the High Pure™ PCR Product Purification Kit (Roche), following the manufacturer’s instructions. Direct sequencing of purified products using the primers CI-J-1718 and CI-N-2191 was undertaken using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), according to the manufacturer’s instructions. Cycle sequencing products were resolved on an Applied Biosystems 3100 automated DNA sequencer.

2.2.3 Data analysis

Sequence data was analysed using Sequencher software (Gene Codes Corporation). Since the sequence data from the COI gene contained no insertions or deletions alignment algorithms were unnecessary. We used PAUP*4.0b (Swofford, 2003) and ModelTest version 3.7. (Posada & Crandall, 1998) to calculate Akaike Information Criterion (AIC) values for 56 different nucleotide substitution models. Maximum likelihood analyses and constrained topology analyses including Shimodaira-Hasegawa (SH) tests (Shimodaira & Hasegawa, 1999) were performed using PAUP*4.0b on the nucleotide alignment.
2.3 Results

All 188 amplified PCR products representing five species within the genus *Ctenopseustis* and seven species within the genus *Planotortrix* were sequenced, resulting in a region of 468bp of aligned sequences. Twelve COI sequences from the same region from (Newcomb & Gleeson, 1998) were added to the alignment. The combined dataset of 200 individuals contained 72 unique haplotypes, 102 variable sites and 80 parsimony-informative sites.

The GTR+I+Γ model was determined to be the best fit model under the AIC. A maximum likelihood phylogenetic tree was constructed from the above alignment using this model (Figure 2.3). The tree shows that the two genera *Ctenopseustis* and *Planotortrix* are reciprocally monophyletic. The single sample of the costal species *C. servana* is sister to the remaining *Ctenopseustis* species. None of the other four *Ctenopseustis* species are monophyletic, including the two fern feeders, *C. fraterna* and *C. filicis*, and the two polyphagus species *C. obliquana* and *C. herana*. Within the genus *Planotortrix* four species are monophyletic, *P. puffini, P. notophaea, P. flammea, P. octoides* well supported with bootstrap values of 100%, while *P. excessana, P. octo* and the mangrove *P. avicennia* are all polyphyletic.

Bootstrap support values for all the nodes in the tree were calculated for 500 replicates. However, of more interest than the robustness of individual nodes in the tree is whether the data are able to reject monophyly of the apparently polyphyletic species. To test this, constrained searches were performed such that each species that is not monophyletic in the unconstrained tree shown in Figure 2.3 was forced to be monophyletic. Again due to the large computational burden a TBR reconnection limit of four was used to build the constrained trees. We then compared the likelihood scores of these constrained trees to the unconstrained tree to examine whether they are significantly different using the SH test (Table 2.1). Both the topology and likelihood scores for the unconstrained maximum likelihood tree (-lnL 1803.92206) and the unconstrained maximum likelihood tree set with a reconnection limit of 4 (-lnL 1806.39639) were very similar, justifying the use of the lower reconnection limit to construct the constrained trees. The SH test rejected the null hypothesis of monophyly for the sibling species pairs of *C. herana* and *C. obliquana*, and *P. octo* and *P. excessana*, but failed to reject the null hypothesis of monophyly for *P. avicennia, C. filicis* and *C. fraterna* even though they are not monophyletic in the unconstrained tree.

Finally, levels of within species divergence compared to levels of between species divergence in the COI dataset were compared (Figures 2.4 and 2.5). Examples of species comparisons where
2.3 Results

Figure 2.3. Maximum likelihood tree constructed from the barcoding region of the Cytochrome c oxidase subunit 1 (COI) gene from 200 specimens of the genera *Ctenopseustis* and *Planotortrix*. Bootstrap support values over 50% are shown above the branches. The tree was rooted between the two genera.
Table 2.1. Shimodaira-Hasegawa test analysis for topological constraints

<table>
<thead>
<tr>
<th>Trees</th>
<th>-lnL</th>
<th>Diff -lnL</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>1803.92206</td>
<td>(best)</td>
<td></td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>1850.41907</td>
<td>46.49701</td>
<td>0.0018*</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>1849.88425</td>
<td>45.96219</td>
<td>0.0029*</td>
</tr>
<tr>
<td><em>P. avicennia</em></td>
<td>1806.27765</td>
<td>2.35559</td>
<td>0.8727</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>1816.55916</td>
<td>12.63710</td>
<td>0.3889</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td>1810.60370</td>
<td>6.68164</td>
<td>0.6504</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>1863.08761</td>
<td>59.16555</td>
<td>0.0011*</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>1862.31310</td>
<td>58.39104</td>
<td>0.0001*</td>
</tr>
</tbody>
</table>

*P ≤ 0.05

Figure 2.4. Comparison of maximal within species divergence with minimal between species divergence at the COI gene for Planotortrix species. Percentage divergence levels were calculated from p-distances. The diagonal of slope=1 is drawn as a dotted line.

levels of within species diversity exceeded between species diversity (points below the dotted line in Figures 2.4 and 2.5) were found. In Planotortrix this included comparisons among *P. octo*, *P. excessana* and *P. avicenniae* and in Ctenopseustis included comparisons among *C. obliquana*, *C. herana*, *C. fraterna* and *C. filicis*. Levels of within species diversity were especially high in *C. obliquana*, *C. herana* and *C. fraterna*, with maximum levels of within species diversity exceeding 3%. 
2.4 Discussion

DNA barcoding has been proposed as a rapid species identification tool yet significant controversy exists as to its utility (DeSalle & Egan, 2005; Hebert et al., 2003a; Meyer & Paulay, 2005; Will & Rubinoff, 2004). To fully assess the utility of DNA barcoding, analysis of species with a well-established taxonomic basis is required. The New Zealand tortricid genera *Ctenopseustis* and *Planotortrix* are ideal in this respect because the species boundaries are supported by morphological and behavioural data through the analysis of species-specific pheromones (Dugdale, 1990; Foster & Dugdale, 1988). The molecular phylogenetic study of *Ctenopseustis* and *Planotortrix* shows that two genera are monophyletic with respect to one another in agreement with the study of (Newcomb & Gleeson, 1998) that sampled the same two genera but with fewer individuals from each species. However, only four of the 11 species, in which multiple individuals were sampled, form monophyletic groupings in the optimal tree Figure (2.3). For four species that are not monophyletic, the SH tests suggest the data are significantly in favour of this.

There are many possible reasons for the lack of species monophyly in the *COI* dataset for *Ctenopseustis* and *Planotortrix*. Biological explanations include incomplete lineage sorting and
introgression (Funk & Omland, 2003). The most likely reason is that the species complexes have been derived recently and that the COI haplotypes are still in the process of sorting. Certainly it seems that the two genera arose recently: with a sequence divergence of approximately 10% and using a rate derived from Brower (1994), the estimated age of the most recent common ancestor between *Ctenopseustis* and *Planotortrix* is approximately 5 million years. However, the age of the unresolved species complexes within each genera are much more recent, with estimates of less than 500,000 years for the most recent common ancestor of *C. obliquana*, *C. herana*, *C. fraterna*, and *C. filicis*, and even more recently for the most recent common ancestor of *P. octo*, *P. octoides*, *P. avicenia*, and *P. excessana*. For several of the species, comparisons of levels of within-species divergence are higher than the minimum levels of between species divergence Figures (2.4 and 2.5). When within species divergence is as high or higher than between species divergence there is no “gap” to enable barcoding to distinguish some of the species from one another. This phenomenon of overlapping barcodes has also been observed in the Hesperiidae family of Lepidoptera (Hajibabaei et al., 2006).

Introgression is another possible reason that can lead to non-reciprocal monophyletic species trees (Funk & Omland, 2003). Introgression events are not uncommon in New Zealand insects, having been observed in cicadas (Buckley et al., 2006) and cockroaches (Chinn & Gemmell, 2004). If introgression events are recent then sympatric species are more likely to share haplotypes than allopatric populations of the two species. To access haplotype frequencies, the Nelson area (NN) of the South Island (see Figure 2.2) represents a suitable sympatric habitat in terms of geographic range and sampling depth for the sibling species pairs of *C. obliquana* and *C. herana* and also for *P. octo* and *P. excessana*. However a comparison to an allopatric population for any of the above mentioned sibling species pairs cannot be addressed with this current dataset in terms of sampling depth or breadth. To allow for a significant statistical test a new dataset from an allopatric population as in the case of *C. herana* for instance needs to be collected.

In contrast shared ancient polymorphism (non-sorting) is less likely to show an effect due to shared distributions (Funk & Omland, 2003). For within *Ctenopseustis*, for example, if introgression was responsible for the observed pattern of relationships, we would expect *C. obliquana* in the South Island of New Zealand to be polyphyletic with *C. herana* and *C. filicis* with which it is sympatric, while *C. obliquana* in the North Island should be polyphyletic with *C. fraterna*, which
is only found in the North Island. However this does not seem to be the case suggesting that incomplete lineage sorting is more likely to explain the polyphyly observed. Sampling nuclear genes may also help in distinguishing introgression from incomplete lineage sorting (Buckley et al., 2006). There are also non-biological reasons that could potentially explain the many non-monophyletic species groupings observed in the *Ctenopseustis* and *Planotortrix* COI dataset. Some possible explanations are that errors have been made in the identification of individual specimens or that the taxonomy of *Ctenopseustis* and *Planotortrix* is incorrect. Certainly some of the sibling species pairs are difficult to distinguish morphologically as adults. However, a majority of our samples were identified based on the current species descriptions by those involved in the most recent taxonomic revision (Dugdale, 1990). Of the 200 specimens sequenced for the COI gene, 117 were identified using fore-wing costal fold ratios for the male specimen and reproductive structures e.g. cestum length ratios for the female specimen, while the remainder were identified as larvae, through host plant affinity or through trapping using species-specific sex pheromone lures. Even if this is so, the analysis of the COI gene dataset does not provide a better hypothesis for the taxonomy of the group and those groups that seem good monophyletic groups in the phylogenetic tree represent groups of already described species (e.g. *P. notophaea* and *P. puffini*).

While the two groups of unresolved species in *Ctenopseustis* and *Planotortrix* in the COI gene tree can be difficult to distinguish morphologically, they use distinct blends of sex pheromones (Newcomb & Gleeson, 1998). Moreover, changes in sex pheromone biochemistry may be driving the speciation process in these moths. Therefore perhaps investigating the patterns of variation in genes encoding elements of the pheromone production system in females and reception in males may yield genes that have already sorted or cannot introgress between species. The length of sequence sampled is also very short. This might also be a factor for the power of phylogenetic information in this dataset. In the next chapter a longer region of the mitochondria is sampled, together with other non-speciation genes, however for a reduced number of taxa.
3.1 Introduction

Genetic sequence information has become a very important tool to infer species relationships and boundaries. For recently speciated taxa distinct loci can produce different phylogenetic patterns making the construction of a species tree problematic. This problem arises because each gene can have an independent history (Edwards & Beerli, 2000; Nei, 1989). Also in closely related species this problem can be confounded by the ability of the phylogenetic algorithms to return the correct tree. After a speciation event takes place it takes time for a gene tree to become reciprocally monophyletic due to the sharing of ancestral polymorphism. This problem is often described as “gene tree versus species tree” (Carstens & Knowles, 2007; Hudson, 1992; Hudson & Coyne, 2002; Maddison, 1997; Page & Charleston, 1997; Pamilo & Nei, 1988; Tajima, 1983; Takahata, 1989; Wu, 1991). If gene genealogies are discordant or display non-exclusive
3.1 Introduction

genealogical patterns they reflect recently diverged populations of species where reproductive isolation is not complete yet. Discordant phylogenies among closely related species is often observed when introgression takes place between different genomic regions as in the case of three closely related Drosophila species, D. pseudoobscura, D. persimilis and D. bogotana. Here an analysis of 16 loci scattered across different regions of the genome showed incidences of gene introgression at some loci which resulted in problems with phylogenetic reconstructions and resulted in discordant gene trees (Machado & Hey, 2003). Exclusivity at a specific gene region, however, may be indicative of genes involved in reproductive isolation, so called 'speciation genes'. Markers at or close to speciation loci are therefore able to exhibit exclusivity in a genealogical tree as in the case of the European corn borer moth (Ostrinia nubilalis). Out of five gene regions reflecting discordant gene phylogenies only one marker displayed pheromone strain exclusivity in a gene tree and was mapped in close proximity to a gene affecting postdiapause development as a candidate speciation gene involved in reproductive isolation (Dopman et al., 2005). Differences in expected coalescence times among loci and recent gene flow between species of Heliconius butterflies also demonstrates the need for multiple locus comparisons to resolve species phylogenies (Beltran et al., 2002). Comparisons of gene genealogies that rely on multiple loci should help to infer a more accurate picture of the genetic architecture of reproductive isolation in diverging groups of closely related species. In our native New Zealand Leafroller moths study system, genetic chromosome maps are not available. Therefore commonly used markers to infer gene phylogenies in Lepidoptera were selected.

In this study cytochrome c oxidase subunit 1 and 2 (COI-COII) were chosen as mitochondrial protein-coding loci. Mitochondrial markers have been widely used to infer relationships between recently diverged populations and species complexes because they evolve rapidly, are maternally inherited, have no recombination and are easy to align because of a lack of gaps, at least for protein coding genes (Avise, 1994). To compare different evolutionary rates among loci two nuclear genes, elongation factor-1α (EF-1α) an autosomal marker and triose-phosphatase isomerase (TPI) as a presumably sex-linked locus, were chosen. Both markers have been frequently used in phylogenetic studies (Beltran et al., 2002; Blum et al., 2003; Cho et al., 1995; Fang et al., 1997; Jiggins et al., 2001; Logsden et al., 1995) and have proven to be useful in inferring phylogenetic relationships within the Lepidoptera.

Evolutionary hypothesis testing through likelihood ratio tests on codon sequences have typi-
cally been performed using the software package Phylogenetic Analysis by Maximum Likelihood (PAML). Ratios of non-synonymous ($d_N$) to synonymous ($d_S$) substitution rates or omega values $\omega = d_N/d_S$ are used to test whether sites or lineages are under selective pressure (Kimura, 1983; Ohta, 1993). Values of $\omega < 1$, $= 1$, and $>1$ can be implied to infer negative purifying, neutral evolution and positive selection, respectively. First, the following models were run for each dataset: M0, M1a, M2a, M3, M7 and M8. The models assume different $d_N/d_S$ ratios or $\omega$ values for lineages or sites within the phylogeny. The simplest model M0 assumes one value for $\omega$ for all sites and is referred to as the “one-ratio” model (uniform selective pressure among sites). Models M1a and M2a or “two-ratio” models are slightly modified versions of the previous M1 and M2 models (Nielsen & Yang, 1998) and assume variable selective pressures, nearly neutral for M1a and positive selection for M2a. In these new models, $0 < \omega_0 < 1$ is estimated from the data while $\omega_1 = 1$ is fixed (Wong et al., 2004; Yang et al., 2005). M3 is a discrete model which assesses heterogeneous $\omega$ ratios amongst sites. M7 or the “beta” model assumes a beta distribution with a continuous distribution of $\omega$ values limited to the interval 0-1 while the M8 “beta plus omega” model adds an extra site class to M7 with a free $\omega$ value $>1$ (Yang et al., 2005). Nested models (M0-M3) were used to test heterogeneity of $\omega$, (M1a-M2a) and (M7-M8) to assess site specific selective pressure.

### 3.2 Aims

The first hypothesis that was tested is whether these four different genes (COI, COII, EF-1α and TPI) have discordant gene genealogies reflecting a mosaic genome? Do phylogenies based on mitochondrial loci result in species as exclusive groupings (reciprocal monophyletic) more often than autosomal and Z-linked markers? In addition tests for selection were performed to examine whether these genes are under different types of selection, as they may also explain any differences in apparent genealogies. A phylogenetic approach was chosen to test these three hypotheses since phylogenetic reconstructions using multiple loci and sequences per species were successfully used in other insects to reveal discordant gene genealogies. Four different genes that are commonly used to infer phylogenetic relationships among members of Lepidoptera have been sampled for 38 different individuals in 12 species within the genera of Planotortrix and Ctenopseustis to address the question of a mosaic genome. Two protein-coding mt genes cytochrome c oxidase subunit 1 and 2 (COI and COII), one protein-coding nuclear gene elongation factor-1α (EF-1α) and one
protein-coding nuclear gene, including non-coding intron regions, \textit{triose-phosphatase isomerase} (\textit{TPI}) were selected to compare differences in coalescence times among gene regions between these closely related taxa. It is expected that mitochondrial and sex-linked loci coalesce faster than autosomal genes. In addition tests for selection were performed to access what type of selection acts on each individual gene.

3.3 Data limitations

In Chapter 2, 200 specimens were used to access a region of the \textit{COI} gene as a suitable marker for species monophyly. Unfortunately due to the poor conservation of the genomic DNA samples, new specimens for all 12 species had to be collected from the field. Known sampling localities from Chapter 2 with the aim of covering good geographic representation for each species were chosen to collect new specimens for most of the species. In this chapter a very small set of specimens was used to collect sequence information at four loci. Ideally the set should consist of at least three specimens per species to decrease the chance outcome of random branching when observing monophyly in a gene tree. Due to the nature of acquiring samples from the field and the seasonal conditions of specimens availability, a total of only 38 specimens were retrieved.

Since Chapter 2 provided us with information of which species do not form exclusive groups I targeted these to collect larger numbers of specimens than for taxa that already show exclusivity in the \textit{COI} tree, such as \textit{P. notophaea} and \textit{P. puffini}.

Within this specimen set, one representative for \textit{C. servana} and \textit{P. flammea}, two specimens for \textit{P. avicenniae} and \textit{P. octoides}, three representatives for \textit{P. notophaea, puffini, excessana, C. filicis} and \textit{C. fraterna}, five specimens for \textit{C. herana} and \textit{P. octo} and a total of seven representatives for \textit{C. obliquana} were collected.

3.4 Materials and Methods

3.4.1 Insect collection

A dataset of 38 specimens from 12 species of the genera \textit{Planotortrix} and \textit{Ctenopseustis} was used in this study. Table 3.1 represents 22 specimens collected as larvae from North- and South-Island localities within New Zealand. Specimens were then reared through to adults and morphologically identified by John S. Dugdale and Robert Hoare based on reproductive
structures in female moths and costal wing fold ratios in male moths. In addition DNA from 16 specimen (Table 3.2) from a previous study by (Newcomb & Gleeson, 1998) were also used and combined to one dataset of 38 specimens.

Table 3.1. Non-speciation gene dataset

<table>
<thead>
<tr>
<th>Species name</th>
<th>Tree-ID</th>
<th>Location</th>
<th>Regional code</th>
<th>Collector</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. obliquana</td>
<td>C.obl-4</td>
<td>Wenderholm</td>
<td>AK</td>
<td>Langhoff, P.</td>
<td>30/12/2003</td>
</tr>
<tr>
<td>C. obliquana</td>
<td>C.obl-5</td>
<td>Whakapapa</td>
<td>TO</td>
<td>Hoare, R.</td>
<td>12/04/2003</td>
</tr>
<tr>
<td>C. obliquana</td>
<td>C.obl-6</td>
<td>Hamilton</td>
<td>WK</td>
<td>Langhoff, P.</td>
<td>04/04/2004</td>
</tr>
<tr>
<td>C. obliquana</td>
<td>C.obl-7</td>
<td>Mt Albert</td>
<td>AK</td>
<td>HortResearch*</td>
<td>2004</td>
</tr>
<tr>
<td>C. herana</td>
<td>C.her-2</td>
<td>Bluff</td>
<td>SL</td>
<td>Langhoff, P.</td>
<td>27/01/2004</td>
</tr>
<tr>
<td>C. herana</td>
<td>C.her-3</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Dugdale, J.S.</td>
<td>20/01/2004</td>
</tr>
<tr>
<td>C. herana</td>
<td>C.her-4</td>
<td>Coll Creek</td>
<td>BR</td>
<td>Dugdale, J.S.</td>
<td>07/12/2003</td>
</tr>
<tr>
<td>C. herana</td>
<td>C.her-5</td>
<td>Tai Tapu</td>
<td>MC</td>
<td>HortResearch*</td>
<td>2004</td>
</tr>
<tr>
<td>C. fraterna</td>
<td>C.fra-3</td>
<td>Tiritangi</td>
<td>AK</td>
<td>Newcomb, R.D.</td>
<td>01/02/2004</td>
</tr>
<tr>
<td>C. filicis</td>
<td>C.fil-1</td>
<td>Bluff</td>
<td>SL</td>
<td>Langhoff, P.</td>
<td>06/01/2004</td>
</tr>
<tr>
<td>P. avicenniae</td>
<td>P.avi-2</td>
<td>Wenderholm</td>
<td>AK</td>
<td>Langhoff, P.</td>
<td>12/12/2003</td>
</tr>
<tr>
<td>P. excessana</td>
<td>P.exc-2</td>
<td>Bluff</td>
<td>SI</td>
<td>Langhoff, P.</td>
<td>27/01/2004</td>
</tr>
<tr>
<td>P. excessana</td>
<td>P.exc-3</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Dugdale, J.S.</td>
<td>02/12/2003</td>
</tr>
<tr>
<td>P. octo</td>
<td>P.octo-2</td>
<td>Te Puke</td>
<td>BP</td>
<td>McKena, C.</td>
<td>30/12/2003</td>
</tr>
<tr>
<td>P. octo</td>
<td>P.octo-3</td>
<td>Whakapapa</td>
<td>TO</td>
<td>?</td>
<td>13/01/2004</td>
</tr>
<tr>
<td>P. octo</td>
<td>P.octo-4</td>
<td>Brooks Reservoir</td>
<td>NN</td>
<td>Langhoff, P.</td>
<td>19/02/2004</td>
</tr>
<tr>
<td>P. octo</td>
<td>P.octo-5</td>
<td>Okutti valley</td>
<td>MC</td>
<td>?</td>
<td>29/03/2004</td>
</tr>
<tr>
<td>P. puffini</td>
<td>P.puf-1</td>
<td>Lee Bay</td>
<td>SI</td>
<td>Langhoff, P.</td>
<td>21/01/2004</td>
</tr>
<tr>
<td>P. puffini</td>
<td>P.puf-2</td>
<td>Golden Bay Track</td>
<td>SI</td>
<td>Newcomb, R.D.</td>
<td>01/06/2004</td>
</tr>
<tr>
<td>P. notophaea</td>
<td>P.not-2</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Langhoff, P.</td>
<td>22/03/2004</td>
</tr>
<tr>
<td>P. notophaea</td>
<td>P.not-3</td>
<td>Murphy’s Bush</td>
<td>AK</td>
<td>Langhoff, P.</td>
<td>10/02/2004</td>
</tr>
</tbody>
</table>

*Moths were provided by the HortResearch Rearing facility; for regional codes see Figures (2.1) and (2.2)

3.4.2 Genomic DNA extraction from adult moths

Genomic DNA (gDNA) was extracted from single adult moths using the Qiagen DNeasy Tissue kit (QIAGEN) according to the manufacturer’s protocol.

3.4.3 PCR amplification and sequencing

Mitochondrial genes

Amplification reactions were prepared using genomic DNA (gDNA) as a template and insect specific primers. The mitochondrial genes cytochrome c oxidase subunit 1 and 2 (COI,COII)
3.4 Materials and Methods

Table 3.2. Additional dataset from (Newcomb & Gleeson, 1998)

<table>
<thead>
<tr>
<th>Species name</th>
<th>Tree-ID</th>
<th>Location</th>
<th>Regional code</th>
<th>Collector</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. obliquana</em></td>
<td><em>Cobl-1</em></td>
<td>Mt Albert</td>
<td>AK</td>
<td>HortResearch*</td>
<td>1982</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td><em>Cobl-2</em></td>
<td>Pollen Island</td>
<td>AK</td>
<td>Dugdale, J.S.</td>
<td>12/02/1999</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td><em>Cobl-3</em></td>
<td>Hamilton</td>
<td>WK</td>
<td>?</td>
<td>30/03/1999</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td><em>Chem-1</em></td>
<td>Tai Tapu</td>
<td>MC</td>
<td>HortResearch*</td>
<td>1990</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td><em>Cfra-1</em></td>
<td>Titirangi</td>
<td>AK</td>
<td>?</td>
<td>20/12/1998</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td><em>Cfra-2</em></td>
<td>Taurewa</td>
<td>TO</td>
<td>?</td>
<td>06/12/1998</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td><em>Cfil-3</em></td>
<td>Half Moon Bay</td>
<td>SI</td>
<td>Dugdale, J.</td>
<td>1999</td>
</tr>
<tr>
<td><em>C. servana</em></td>
<td><em>Cser</em></td>
<td>Muriwai</td>
<td>AK</td>
<td>Newcomb, R.D.</td>
<td>14/09/1999</td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td><em>Pavi-1</em></td>
<td>Pollen Island</td>
<td>AK</td>
<td>?</td>
<td>26/03/1999</td>
</tr>
<tr>
<td><em>P. exessana</em></td>
<td><em>Pexc-1</em></td>
<td>Rukuhia</td>
<td>WO</td>
<td>HortResearch*</td>
<td>1987</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td><em>Pocto-1</em></td>
<td>Canterbury</td>
<td>MC</td>
<td>HortResearch*</td>
<td>1982</td>
</tr>
<tr>
<td><em>P. octoides</em></td>
<td><em>Pocti-1</em></td>
<td>Pitt Island</td>
<td>CH</td>
<td>Marris, J.W.M.</td>
<td>1998</td>
</tr>
<tr>
<td><em>P. octoides</em></td>
<td><em>Pocti-2</em></td>
<td>South East Is.</td>
<td>CH</td>
<td>Marris, J.W.M.</td>
<td>1998</td>
</tr>
<tr>
<td><em>P. puffini</em></td>
<td><em>Ppuf-3</em></td>
<td>Lee Bay</td>
<td>SI</td>
<td>Dugdale, J.S.</td>
<td>1999</td>
</tr>
<tr>
<td><em>P. notophae</em></td>
<td><em>Pnot-1</em></td>
<td>Palmerston North</td>
<td>WI</td>
<td>?</td>
<td>1990</td>
</tr>
<tr>
<td><em>P. flammea</em></td>
<td><em>Pflam</em></td>
<td>Kaipakati Point</td>
<td>BR</td>
<td>Dugdale, J.S.</td>
<td>2001</td>
</tr>
</tbody>
</table>

*Moths were provided by the HortResearch Rearing facility; for regional codes see Figures (2.1) and (2.2)

and the leucine tRNA were amplified as a single fragment using the primer pair TY-J-1460 and TK-N-3785 from Simon et al. (1994). Products were amplified using the following cycling profile of 94°C for an initial 10 mins, followed by 35 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 2 mins followed by a 10 min final extension time at 72°C. Amplification reactions of 50 µL consisted of 10 pmol of each primer, 1.5 mM MgCl2, 0.2 mM of each dNTP and 2 units of *Taq* polymerase (Invitrogen) and included positive and negative PCR controls. The PCR products were electrophoretically separated on a 1% agarose gel and stained with ethidium bromide (1 mg/ml). PCR products were sequenced directly using the primers from the original PCR amplification and species-specific sequencing primers (Tort1-7) from Gleeson et al. (2000) and additional mt primers (mt8, 11 and 15) from Simon et al. (1994) (Table 3.3). Big Dye cycle-sequencing reactions were performed and resolved by either the HortResearch DNA-sequencing facility on an ABI310 capillary sequencer or by the Allan Wilson Centre Genome Service in Albany (AWCGS) on an ABI3730 Genetic Analyzer.
Table 3.3. Primer list

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’ to 3’)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M3</td>
<td>CACATYAACATTGTCGTSATYGG</td>
<td>Cho et al., 1995</td>
</tr>
<tr>
<td>rcM44.9</td>
<td>CTGTAGAATACCTCTGTCGCC</td>
<td>Cho et al., 1995</td>
</tr>
<tr>
<td>rcM52.6</td>
<td>GCCYTGGGTCTCATCTGTCAC</td>
<td>Cho et al., 1995</td>
</tr>
<tr>
<td>M52.7</td>
<td>CGTCAGAAYGTCGTCTG</td>
<td>Cho et al., 1995</td>
</tr>
<tr>
<td>rcM4</td>
<td>ACAGCVAACGTYYCTGCTCATRT</td>
<td>Cho et al., 1995</td>
</tr>
<tr>
<td>TPI-1</td>
<td>GGTCACCTCTGAAAGGAGAACCATCTTT</td>
<td>Beltran et al., 2002</td>
</tr>
<tr>
<td>TPI-2</td>
<td>CACAACATTTGCCCAGTGGTTCGCAA</td>
<td>Beltran et al., 2002</td>
</tr>
<tr>
<td>TPI for</td>
<td>CACAACATTGGCCAGTTGTTGCAATAGCCG</td>
<td>this study</td>
</tr>
<tr>
<td>TPI rev</td>
<td>GGTCACCTCTGAAAGGAGAACCATCTTTGGTGA</td>
<td>this study</td>
</tr>
<tr>
<td>TPI rev deg</td>
<td>GGTCACCTGAAAGGAGAACCATCTTGGYGA</td>
<td>this study</td>
</tr>
<tr>
<td>TY-J-1460</td>
<td>TACAATTATCCCTAAACTCCAGCC</td>
<td>Simon et al., 1994</td>
</tr>
<tr>
<td>Tort1</td>
<td>AAAGTGGGGGGAAGTACTCTGC</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort2</td>
<td>AATCTGGTTATATTTATTTTTTTTGAATT</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort3</td>
<td>AATTGAATTATATTATTTTTTATATT</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort4</td>
<td>GTTGCCATTATTTTAAAGAGAT</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort5</td>
<td>CTTTTTTTCTAAATGGCAAC</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort6</td>
<td>CAATGATATGTAAGTGCT</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>Tort7</td>
<td>CAATTTATACATTGGTCTCTTC</td>
<td>Gleeson et al., 2000</td>
</tr>
<tr>
<td>mt8</td>
<td>CAACCTAATTGGTGATTTTGG</td>
<td>Simon et al., 1994</td>
</tr>
<tr>
<td>mt11</td>
<td>ACCTGAAATATATGATGACTCA</td>
<td>Simon et al., 1994</td>
</tr>
<tr>
<td>mt15</td>
<td>TCATAAGTTTCTATCTGTC</td>
<td>Simon et al., 1994</td>
</tr>
<tr>
<td>TK-N-3785</td>
<td>GTTTAAGAGACCAGTACTTG</td>
<td>Simon et al., 1994</td>
</tr>
</tbody>
</table>

Nuclear genes

The elongation factor 1 alpha (EF-1α) gene was PCR amplified using the primers M3 and rcM4 from (Cho et al., 1995). PCR conditions were 3 min at 95°C, followed by 35 cycles of 30 sec at 95°C, 30 sec at 50°C, 30 sec at 72°C and a final elongation time at 72°C for 7 min. The triosephosphate isomerase (TPI) gene was amplified using the degenerate primers TPI-1 and TPI-2 from (Beltran et al., 2002) and using species-specific designed primers (TPI for, rev and rev deg) (see Table 3.3) for sequencing. The primers were situated in exons 3 and 4 of Heliothis (GenBank accession number U23080) and spanned intron 3 of the TPI gene. Amplification reactions of 50 µL consisted of 10 pmol of each primer, 5 mM buffer, 1.5 mM MgCl2, 0.2 mM of each dNTP and 2 units of Taq polymerase (INVITROGEN). The PCR products were electrophoretically separated on a 1% agarose gel and stained with ethidium bromide (1 mg/ml). Both TPI and EF-1α exhibited one or two bands on the agarose gel corresponding to different sized alleles and also exhibited ambiguous base calls in heterozygotes. Therefore PCR amplified products were shot-gun cloned into pGEM-T Easy Vector System (PRIMEGA) to obtain the sequence for
each individual allele. To minimize the problem of PCR errors through nucleotide substitution (Kobayashi et al., 1999), templates obtained from three to five colonies per individual were sequenced on an ABI3730 Genetic Analyzer at the Allan Wilson Centre in Albany (AWCGS) using vector-specific primer (M13 forward and M13 reverse) or primers used in the original amplification reaction (see Table 3.3). A consensus sequence was selected if the cloned sequences varied from the original PCR amplification product. For EF-1α additional primers for direct sequencing of the PCR product from Cho et al. (1995) were used (rcM44.9, rcM52.6 and M52.7).

### 3.4.4 Sequence alignment

Chromatograms of all sequences were edited and aligned using Sequencher 4.5 (Gene Codes Corporation, Inc.). Mitochondrial sequence data from COI and COII was aligned into a single contiguous sequence including leucine tRNA and compared to lepidopteran sequences on GenBank using Blast. Exon data for EF-1α and TPI and introns for TPI were aligned in ClustalX (Thompson et al., 1997). Introns from TPI could only be aligned within each genus (Planotorrix and Ctenopseustis) due to the existence of indels. Five datasets were constructed, one for the mitochondrial genes including COI, leucine tRNA and COII, one for EF-1α and three for the TPI gene. For TPI one dataset is based on exon data only and includes all sequences of both genera. The other two datasets are comprised of intron and exon data and are separated by genus since it was difficult to align both genera in one dataset due to differences in intron sizes.

### 3.4.5 Phylogenetic analysis

The nucleotide sequences of the protein-coding COI and COII genes of the mitochondria and nuclear sequences of TPI and EF-1α were checked manually for reading-frame errors and translated to peptide sequences in MacClade 4.06 (Maddison & Maddison, 1997) or Sequencher 4.5 (Gene Codes Corporation, Inc). Phylogenetic analyses were conducted with PAUP* version 4.0b8 (Swofford, 2003). Models of sequence evolution were estimated using ModelTest 3.7 (Posada & Crandall, 1998) and PAUP*. PAUP* was then used to search for the maximum likelihood (ML) tree, based on the best fit model and parameter estimates given by ModelTest 3.7 and PAUP* using a heuristic search with tree-bisection reconnection (TBR). For comparison Bayesian Inference tree analysis using MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001) was also
3.4 Materials and Methods

performed for all five datasets with the same evolutionary model settings used in the PAUP* analysis. Each individual gene dataset was run ten times starting from different random seeds for ten million generations sampling every 100 generations. The burn-in value was taken after all parameter values had stabilized and four Metropolis-Coupled Markov Chains (MCMC) chains were used at different temperatures optimized for each dataset. Since the ML trees obtained from PAUP* were the same as those obtained using MrBayes analysis and due to the large computational burden of estimating bootstrap support values, posterior probability scores were used as a confidence value for each node in the nuclear gene trees. The mt genes have bootstrap support values (1000 replicates, heuristic search with TBR branch swapping) as an additional confidence indicator.

3.4.6 PAML analysis for testing selection

All datasets were codon-aligned and non-protein coding regions and stop-codons removed, such as tRNA-Leu from the mt dataset. The ML trees and nucleotide datasets from above were used in CODEML from the PAML 3.15 package (Yang, 1997). Models M0, M1a, M2a, M3, M7 and M8 were run for each dataset to test if particular sites within the protein exhibit positive selection. The model that provided the best fit for the data was determined by implementing a likelihood-ratio test (LRT) for nested models, M0-M3, M1a-M2a, M7-M8. There the LRT is performed by twice the log-likelihood difference between the nested models and comparing this value to a chi-square ($\chi^2$) distribution with degrees of freedom being equal to the difference in the number of parameters between the two models. To test whether a specific lineage is under positive selection ($d_N/d_S$ ratio > 1) the two-ratio models (model=2) or branch-site model A can be used to construct a branch-site test 2 of positive selection (Zhang et al., 2005). This test divides branches a priori into foreground and background lineages (Yang & Nielsen, 2002). A LRT was performed between model M1a (nearly neutral) and the branch-site test 1 and 2. Test 1 is not recommended by itself (Zhang et al., 2005) because relaxed selective constraint can be mistaken for false positive selection. Test 2 appears to be robust and is recommended to be used as a test of positive selection along branches (Zhang et al., 2005). The Null hypothesis for branch-site model A fixes $\omega$ at 1 on the foreground branches whereas in the alternative hypothesis $\omega$ is estimated.
3.4.7 Constraint analysis and monophyly tests

To assess whether the observed tree topology pattern for the four non-speciation genes are reliable and to test the hypothesis of reciprocal monophyly for each species the Shimodaira-Hasegawa test (SH) (Shimodaira & Hasegawa, 1999) implemented in PAUP* 4.0b (Swofford, 2003) was performed. The SH test compares the likelihood values for the unconstrained ML tree versus a constrained tree topology and generates a p-value to obtain a confidence limit for the topology. One thousand bootstrap replicates were used with a default resampling of estimated log-likelihoods (RELL) setting.

Rosenberg (2007) developed a test of the Null hypothesis that monophyly is a chance outcome of random branching. In addition he calculated the sample size required of a set of lineages to lie below a prescribed tolerance for the probability of monophyly to occur by chance. Ten alternative hypotheses to Null were compared for each species being monophyletic. Using the Bonferroni correction for multiple comparison, \( \alpha/n \), \( \alpha = 0.005 \) was obtained as the significance level at which the null hypothesis of random branching can be rejected.

3.5 Results

3.5.1 Gene tree phylogenies

Mitochondrial DNA

A 2280bp fragment of the mitochondrial genome including the COI, tRNA and COII genes was sequenced from 36 individuals of the 38 specimen listed in Table 3.1. COI/COII sequence was not recovered for C.her-3 and P.octi-2. The alignment contained no gaps. Of 2280 aligned nucleotide sites, 212 were variable (9.3%) and 160 (7%) were parsimony informative. This represents 1525bp of the COI gene corresponding to positions 1484 to 3009 of the D. yakuba sequence (X03240), the complete leucine-tRNA gene (73bp), and 682bp representing the entire COII coding sequence (For nucleotide alignment see Appendix B on the CD). The Bayesian tree (Figure 3.1) had an identical topology to the ML tree. There is good support both by 100% bootstrap support values and a posterior probabilities value of 1.0 for the split that separates Planotortrix and Ctenopseustis. The tree was rooted between the genera. Within Ctenopseustis, C. servana is sister group to the other four species. Only C. fraterna forms a monophyletic or exclusive group with good support values of 99% bootstrap and a posterior probability of 1.0.
Within the *C. obliquana* clade all specimens group together as each others closest relatives with the exception of *C. her-4* which appears to be closely related to *C. obl-3* and not with any of the other *C. herana* specimens in the tree. However the branch leading to this grouping is poorly supported with bootstrap support of just 60% and a posterior probability of 0.6. The branch leading to the specimens of *C. obl-1, C. obl 3-7* and *Cher-4* has an even lower bootstrap support of <50%. Only *C. obl-2* comes out as sister group to the other *C. obliquana* specimens with support values of (100%/ 1.0). The three remaining *C. herana* specimens are paraphyletic with *C. filicis* with 100% bootstrap confidence and 0.9 posterior probability for the branch leading to *C. herana* and *C. filicis*. Within the *Planotortrix* clade *P. flammea* is sister group to *P. notophaea* and *P. puffini*. Both *P. puffini* and *P. notophaea* are well supported clades with 100% bootstrap support and a posterior probability score of 1.0. The second grouping includes *P. octoides* as the sister species for *P. octo, P. excessana* and *P. avicennia*. *P. excessana* and *P. octo* show a paraphyletic pattern. *P.octo-4*, which clusters with *P.exc-3*, shows a lower bootstrap confidence and posterior probability score (80%, 0.9) than the branch leading to *P.exc-2, P.exc-3* and *P.octo-4* with (100%, 1.0). Both specimens of *P. avicennia* are each others closest relative and group together, whereas *P.exc-1, P.octo-1 to 3* and *P.octo-5* are polyphyletic.

**EF-1α**

Aligned sequence data consisted of 1238bp characters from 35 specimens. *EF-1α* sequence data was not recovered for *C.her-3, P.octi-2* and *P.not-3*.

Out of the 35 specimens, 19 were heterozygous with 18 displaying more than one polymorphic site. The data contained no introns. Only a very small percentage of sites (71) were variable (4.6%) and 66 (4.2%) were parsimonomious informative. (For nucleotide alignments see Appendix B on the CD). Again the ML tree and Bayesian Inference trees were identical in their topology. Due to the large computational burden of calculating bootstrap support values, only posterior probability values are shown in Figure 3.2 as branch support values. The *EF-1α* ML tree shows *Planotortrix* and *Ctenopseustis* as well resolved genera (posterior probability of 1.0). Again the tree has been arbitrarily rooted between the two genera. Within *Ctenopseustis* three well supported groups are observed. *C. servana* is sister group to all other *Ctenopseustis* species. Both fern-feeders *C. fraterna* and *C. filicis* form an exclusive group with high posterior probability score of 1.0 for *C. filicis* and 0.8 for *C. fraterna*. *C. herana* reflects a paraphyletic pattern with
Figure 3.1. Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on combined COI and COII sequences. Bootstrap branch support values (1000 replicates) if > 50 are given above the branches. Posterior probability values from MrBayes are shown below the branches. Colours were used for four species in the tree to emphasize their non-exclusive status. # refers to branch numbering in PAML analysis. Tree is rooted between the genera.
C. filicis except for C. her-4 being heterozygous, where both alleles are more closely related to the C. obliquana grouping. C. obliquana reflects a paraphyletic pattern with a high posterior probability value of 0.9. A similar pattern to that observed in Ctenopseustis is apparent in the genus Planotortrix. Three species form exclusive groups, P. puffini, P. flammia and P. notophaea with very high posterior probability scores from 0.8-1.0. P. octoides, the Chatham-Island endemic is sister to the mangrove species P. avicenniae. P. avicenniae, P. excessana and P. octo form not-exclusive groupings.

**TPI**

The amplified products for the TPI gene yielded a 650bp fragment for 36 individuals. The specimens C.frat-2 and C.her-3 could not be amplified for the TPI gene and therefore could not be included in the analysis. Out of 36 individuals 22 were homozygous and 14 were heterozygous. Six of these heterozygous individuals displayed two different products in length, either a 400 and 650bp PCR product or a 650 and 800bp fragment. These length polymorphisms were due to different intron sizes within the species. The nucleotide alignments consists of exon 3, 4 and intron 3 corresponding to position 275 to 298 of exon 3, and 299 to 422 in exon 4 in Heliothis virescens (U23080). Three TPI datasets were created (Appendix B on CD). The first dataset consists of exon 3 and 4 only, containing 147 nucleotides that include 17 variable sites of which 16 are parsimonious informative and is referred to as 'TPI exon'. For the second and third dataset, TPI was split into both genera Planotortrix and Ctenopseustis and analysed separately due to length variations in the intron sites that made it difficult to align all the sequences together. The combined intron and exon dataset consists of 767 nucleotide bp for Ctenopseustis (including alignment gaps) with 35 parsimonious informative out of 43 variable sites, whereas Planotortrix yielded a 737 bp sequence dataset (including gaps) with 77 parsimonious informative out of 80 variable sites. Intron sizes varied within Ctenopseustis from 399 to 455pb, within Planotortrix between 260 and 423bp. Both datasets two and three containing intron and exon nucleotides for the genera are referred to in later chapters as the 'TPI intronP' for the Planotortrix and 'TPI intronC' for the Ctenopseustis dataset.

The inferred Bayesian trees based on the TPI exon data is shown in Figure 3.3, and both intron datasets TPI intronC and TPI intronP are shown in Figure 3.4 with posterior probability scores as support values. The phylogenetic tree based on the TPI exon dataset was rooted
3.5 Results

Figure 3.2. Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on EF-1α sequences. Posterior probability values from MrBayes are shown above the branches. Species numbering refers to those given in Table 3.1, if the same specimen appears twice in the tree, it is heterozygous. Colours were used for four species in the tree to emphasize their non-exclusive status. # refers to numbering in PAML analysis. Tree is rooted between the genera.
between the genera and shows polyphyly for most taxa. However some taxa including *C. filicis*, *C. fraterna*, *C. servana* and *P. flammea* are monophyletic. This poor resolution is probably due to the small amount of informative sites in the exon region with only 16 parsimony informative sites.

The inferred phylogenetic trees based on the *TPI* intron datasets are shown in Figure 3.4. For *TPI* intron C, *C. servana* was used to root the tree since it has shown to be sister group to *Ctenopseustis* for all other markers. Both fern-feeders, *C. filicis* and *C. fraterna*, are sister to each other as a well supported clade with a posterior probability of 1.0. Closest to the *C. fraterna/filicis* group are two *C. herana* individuals, both alleles of *C. her-5* and one allele of *C. her-1*. One allele of *C. her-1* groups with two more sister group alleles of *C. obliquana*, *C. obl-3* and one allele of *C. obl-5*. The remaining two alleles of *C. herana* appear within the second clade of *C. obliquana* which is sister to the above mentioned *C. fraterna*, *C. filicis* and *C. herana* grouping. *P. flammea* was used to root the tree for the *TPI* intron P dataset as it has been shown to be sister group to the clade including *P. octo*, *P. excessana* and *P. avicenniae* in *EF-1α* and *COI-COII*. Within the *Planotortrix* tree four species (*P. flammea*, *P. notophaea*, *P. puffini* and *P. octoides*) form exclusive groups as observed in the *EF-1α* and mitochondrial trees. For *TPI* a second individual of *P. octoides* was successfully amplified and corroborates the monophyletic pattern of *P. octoides* in the tree with a good posterior probability score of 1.0. The sibling species *P. octo* and *P. excessana* and the mangrove species of *P. avicenniae* display a polyphylectic pattern.

### 3.5.2 Tree topology tests

Table 3.4 gives a summary of whether species are reciprocal monophyletic or form exclusive groupings for all regions examined, excluding *P. flammea* and *C. servana*. In *Ctenopseustis* neither of the sibling species pair *C. obliquana* and *C. herana* show exclusivity for any of the four genes. *C. fraterna* is monophyletic when wider regions of the mt genes are used (*COI* and *COII*) and also for the nuclear genes *EF1α* and *TPI*. *C. filicis* is not monophyletic for the mt loci but for both nuclear loci. In *Planotortrix* the sibling species pair *P. octo* and *P. excessana* shows a paraphyletic pattern for all loci as does *P. avicenniae*, except for the wider mt region where *P. avicenniae* forms a monophyletic clade. The available number of alleles for *TPI* were 47 alleles (excluding *P. flammea* and *C. servana*) and 52 alleles for *EF-1α*.
Figure 3.3. Maximum likelihood gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on *TPI* exon sequences. Species numbering refers to those given in Table 3.1, if the same specimen appears twice in the tree, it is heterozygous. Colours were used for four species in the tree to emphasize their non-exclusive status.
Figure 3.4. Bayesian inference gene tree phylogeny for *Planotortrix* and *Ctenopseustis* species based on *TPI* sequences. Species numbering refers to those given in Table 3.1, if the same specimen appears twice in the tree, it is heterozygous. Colours were used for four species in the tree to emphasize their non-exclusive status. Tree is rooted between the genera.
3.5 Results

Table 3.4. Species monophyly status

<table>
<thead>
<tr>
<th>Species</th>
<th>'Barcoding-region'</th>
<th>COI-COII</th>
<th>EF-1α</th>
<th>TPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.herana</td>
<td>not-exclusive(52)*</td>
<td>not-exclusive(4)</td>
<td>not-exclusive(6)</td>
<td>not-exclusive(6)</td>
</tr>
<tr>
<td>C.obliquana</td>
<td>not-exclusive(73)</td>
<td>not-exclusive(7)</td>
<td>not-exclusive(11)</td>
<td>not-exclusive(10)</td>
</tr>
<tr>
<td>C.fraterna</td>
<td>not-exclusive(9)</td>
<td>exclusive(3)</td>
<td>exclusive(6)</td>
<td>exclusive(3)</td>
</tr>
<tr>
<td>C.filicis</td>
<td>not-exclusive(8)</td>
<td>not-exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
</tr>
<tr>
<td>P.octo</td>
<td>not-exclusive(22)</td>
<td>not-exclusive(5)</td>
<td>not-exclusive(8)</td>
<td>not-exclusive(7)</td>
</tr>
<tr>
<td>P.excessana</td>
<td>not-exclusive(14)</td>
<td>not-exclusive(3)</td>
<td>not-exclusive(6)</td>
<td>not-exclusive(3)</td>
</tr>
<tr>
<td>P.avicenniae</td>
<td>not-exclusive(3)</td>
<td>exclusive(2)</td>
<td>not-exclusive(3)</td>
<td>not-exclusive(3)</td>
</tr>
<tr>
<td>P.puffini</td>
<td>exclusive(7)</td>
<td>exclusive(3)</td>
<td>exclusive(5)</td>
<td>exclusive(5)</td>
</tr>
<tr>
<td>P.notophaea</td>
<td>exclusive(5)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
</tr>
<tr>
<td>P.octoides</td>
<td>exclusive(4)</td>
<td>-(1)</td>
<td>-(1)</td>
<td>exclusive(4)</td>
</tr>
</tbody>
</table>

| Species exclusive ∑ | 3 | 4 | 4 | 5 |
| Species not-exclusive ∑ | 7 | 5 | 5 | 5 |

* Number of alleles are given in brackets

Tests for monophyly

To test for monophyly of the different species, a comparison between the ML tree, referred to as “Unconstrained” tree and a constrained tree for each of the species that reflected a polyphyletic pattern was performed. Six alternative topologies (alternative hypothesis to Null) were compared to the unconstrained ML (Null hypothesis) tree using the Shimodaira-Hasegawa (SH)-test as implemented in PAUP*. Log likelihood differences, log likelihoods and P-values for all polyphyletic groups are given in Table 3.5. If the mitochondrial tree is constrained so that all C. obliquana specimens are forced to be monophyletic, the (SH)-test is significant (p < 0.5) suggesting that C. obliquana is not monophyletic. For the remaining taxa only C. filicis and P. avicenniae have non-significant p-values suggesting that even though both are not monophyletic in the ML tree the data cannot reject monophyly. In EF-1α none of the p-values for any species are significant suggesting that monophyly cannot be rejected. In TPI only C. obliquana and C. herana show significant p-values indicating no exclusivity for these two species.

However the possibility that monophyly has been observed due to insufficient sampling rather than by evolutionary processes is still likely since only a small number of individuals were available in this study. To address whether monophyly occurred by chance through random branching, can be tested statistically as in Rosenberg (2007). All species including the ones
### Table 3.5. Constraint tree analysis using Shimodaira-Hasegawa test

<table>
<thead>
<tr>
<th>Tree</th>
<th>-lnL</th>
<th>Diff(-lnL)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>1803.92</td>
<td>(best)</td>
<td></td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td>1806.27</td>
<td>02.355</td>
<td>0.872</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>1850.41</td>
<td>46.497</td>
<td>0.001*</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>1849.88</td>
<td>45.967</td>
<td>0.002*</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>1862.31</td>
<td>58.391</td>
<td>0.001*</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>1863.08</td>
<td>59.165</td>
<td>0.001*</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>1816.55</td>
<td>12.637</td>
<td>0.388</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tree</th>
<th>-lnL</th>
<th>Diff(-lnL)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>2492.38</td>
<td>(best)</td>
<td></td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td>2501.4</td>
<td>9.021</td>
<td>0.427</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>2492.38</td>
<td>0.000</td>
<td>0.903</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>2500.19</td>
<td>7.815</td>
<td>0.471</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>2494.33</td>
<td>1.950</td>
<td>0.818</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>2506.41</td>
<td>14.034</td>
<td>0.294</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>2492.38</td>
<td>0.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tree</th>
<th>-lnL</th>
<th>Diff(-lnL)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>2492.38</td>
<td>(best)</td>
<td></td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td>2075.89</td>
<td>10.442</td>
<td>0.356</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>2075.30</td>
<td>9.857</td>
<td>0.402</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>2076.93</td>
<td>11.483</td>
<td>0.383</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>2077.18</td>
<td>73.540</td>
<td>0.001*</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>2066.62</td>
<td>62.986</td>
<td>0.006*</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>2003.64</td>
<td>0.000</td>
<td>0.895</td>
</tr>
</tbody>
</table>

* significant if p < 0.5

with only three alleles available to study, for example *C. filicis* in *EF-1α* and *TPI*, are under the significance level at which the Null hypothesis of random branching can be rejected (Table 7, in (Rosenberg, 2007)). However small the number of alleles, it suggests that the exclusivity we see in the different gene genealogies is not likely a chance outcome.

### 3.5.3 Testing sites and lineages under selection

The PAML analysis of *EF-1α* and *COI-COII* revealed that none of the genes are under positive selection. On the contrary both genes are under strong purifying selection possibly due to functional constraints acting on these loci. An analysis for the *TPI* gene in PAML could not be achieved due to the number of polytomies in the phylogeny (see section 3.4.1).
3.5 Results

Site model analysis

The results of the PAML analysis for the site specific models are given in Table 3.6 for COI-COII and in Table 3.7 for EF-1α. The average ω ratio for the mt genes ranges from 0.006 to 0.012 among all models (Table 3.6), which is very low and suggests strong purifying selection (ω<<1). This ratio suggests that a non-synonymous mutation has an only 0.6 - 1.2% as much chance as a synonymous mutation of being fixed in the population. Estimates of transition/transversion rate ratio (κ) ranged from 6.37 to 6.76. (Model M3 suggests that a small proportion of sites (4.6%) are under diversifying selection with ω_1 = 0.154.) The parameters for the β-distribution (p and q) indicate L-shaped, highly left-skewed distributions. A nested model comparison of M0 versus M3 is a test of heterogeneity among sites and is significant in COI-COII implying that there is no uniform selective pressure acting on them (Table 3.8). The likelihood scores for model M1a and M2a are the same so none of these models fit the data better suggesting that no sites are under positive selection. Model M7 was a better fit to the data than M8 also implying that no positive selective pressure is detectable.

**Table 3.6.** Likelihood values and parameter estimates for mt genes COI-COII

<table>
<thead>
<tr>
<th>Model code</th>
<th>1</th>
<th>d_N/d_S</th>
<th>Estimates of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0 (one-ratio)</td>
<td>-3196.73</td>
<td>0.006</td>
<td>ω = 0.006</td>
</tr>
<tr>
<td>M1 (neutral)</td>
<td>-3190.90</td>
<td>0.012</td>
<td>p_0 = 0.987, (p_1 = 0.012)</td>
</tr>
<tr>
<td>M2 (selection)</td>
<td>-3190.90</td>
<td>0.012</td>
<td>p_0 = 0.987, p_1 = 0.012, (p_2 = 0.00), ω_2 = 10.403</td>
</tr>
<tr>
<td>M3 (discrete)</td>
<td>-3179.75</td>
<td>0.007</td>
<td>p_0 = 0.727, p_1 = 0.225, (p_2 = 0.046)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ω_0 = 0.000, ω_1 = 0.000, ω_2 = 0.154</td>
</tr>
<tr>
<td>M7 (beta)</td>
<td>-3181.86</td>
<td>0.008</td>
<td>p = 0.010, q = 0.364</td>
</tr>
<tr>
<td>M8 (beta&amp;ω)</td>
<td>-3181.87</td>
<td>0.008</td>
<td>p_0 = 1.000, (p_1 = 0.000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p = 0.011, q = 0.373, ω = 5.140</td>
</tr>
</tbody>
</table>

In EF-1α the model comparison yielded similar results compared with the mitochondrial genes (Table 3.8). The discrete model M3 was a better fit than M0 implying heterogeneous selective pressure and model M1a and M7 fit the data better than M2a and M8, respectively, suggesting that no positive selection is detectable. The ω values ranged from 0.031 to 0.035, again indicating strong purifying selection. The chance for a non-synonomous substitution of getting fixed in the population is 3.1% to 3.5%. Estimates of ω the transition/transversion ratio ranged from 3.34 to 3.42. Model M8 suggests that a small number of sites, 3.5%, are possibly under positive selection with ω of just 1.0. Again the β-distribution parameters (p and q) specify a highly left-skewed distribution indicating that most amino acids are conserved.
Table 3.7. Likelihood values and parameter estimates for EF-1α

<table>
<thead>
<tr>
<th>Model code</th>
<th>l</th>
<th>dN/dS</th>
<th>Estimates of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0 (one-ratio)</td>
<td>-2210.72</td>
<td>0.031</td>
<td>ω = 0.031</td>
</tr>
<tr>
<td>M1 (neutral)</td>
<td>-2200.42</td>
<td>0.035</td>
<td>p₀ = 0.964, (p₁ =0.035)</td>
</tr>
<tr>
<td>M2 (selection)</td>
<td>-2200.42</td>
<td>0.035</td>
<td>p₀ = 0.964, p₁ = 0.017, (p₂ = 0.018), ω₂ = 1.00</td>
</tr>
<tr>
<td>M3 (discrete)</td>
<td>-2200.34</td>
<td>0.032</td>
<td>p₀ = 0.321, p₁ = 0.639, (p₂ = 0.038)</td>
</tr>
<tr>
<td>M7 (beta)</td>
<td>-2202.11</td>
<td>0.034</td>
<td>p = 0.010, q = 0.226</td>
</tr>
<tr>
<td>M8 (beta&amp;ω)</td>
<td>-2200.42</td>
<td>0.035</td>
<td>p₀ = 0.964, (p₁ =0.035)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p = 0.005, q = 1.811, ω = 1.0</td>
</tr>
</tbody>
</table>

Table 3.8. Nested model results of PAML analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>2Δl</th>
<th>p-value for χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF-1α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0-M3</td>
<td>20.7638</td>
<td>0.0015</td>
</tr>
<tr>
<td>M1a-M2a</td>
<td>3.3716</td>
<td>0.00035*</td>
</tr>
<tr>
<td>M7-M8</td>
<td>0.99999</td>
<td>0.18528</td>
</tr>
<tr>
<td>COI-COII</td>
<td>33.9641</td>
<td>0</td>
</tr>
<tr>
<td>M0-M3</td>
<td>-0.0009</td>
<td>7.57E-07*</td>
</tr>
<tr>
<td>M1a-M2a</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

* p < 0.01

Branch model analysis

The analysis of lineage specific selection can reveal ancestral selection pressure on protein coding genes. Four lineages were tested in COI-COI leading to the species complexes of C. obliquana and C. herana and P. excessana and P. octo see Figure (3.1) and one lineage in EF-1α leading to the paraphyletic grouping of P. octo, P. excessana and P. avicenniae as presented in Figure (3.2). The original branch-site model A was split into two tests (test 1 and test 2) and compared to the nearly neutral model M1a. Table (3.9) summarizes the results for the branch site model in COI-COI and EF-1α. For the four lineages tested in COI-COI and one lineage in EF-1α none of the branches showed a significant p-value for the nested model comparison of M1a to model A (test 1) implying that no positive selection was detected along those lineages.

3.6 Discussion

3.6.1 Discordant tree topologies

Four different gene regions (COI, COII, EF-1α and TPI) have been examined for 12 species within the native Leafroller genera Planotortrix and Ctenopseustis. The gene trees show that each gene region reflects a different pattern of sorting or species exclusivity (Table 3.4). Three
Table 3.9. Branch-model results of PAML analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Branch</th>
<th>ln-M1a</th>
<th>ln-Model A</th>
<th>$2\Delta l$ (df=1)</th>
<th>$\chi^2$ (df=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF-1α</td>
<td>1</td>
<td>-2200.425073</td>
<td>-2200.424553</td>
<td>0.0010</td>
<td>0.9742</td>
</tr>
<tr>
<td>COI-COII</td>
<td>1</td>
<td>-3190.908745</td>
<td>-3190.908745</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-3190.908745</td>
<td>-3190.444359</td>
<td>0.9287</td>
<td>0.3351</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-3190.908745</td>
<td>-3190.908747</td>
<td>-0.000004</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-3190.9087</td>
<td>-3190.4817</td>
<td>-0.000018</td>
<td>-</td>
</tr>
</tbody>
</table>

* p < 0.01

species, *P. notophaea*, *P. puffini* and *P. octoides* constantly form a monophyletic grouping for all loci. Whereas *C. obliquana* and *C. herana*, *P. octo* and *P. excessana* never exhibit that kind of behaviour for any loci examined. The monophyletic status for the remaining five species is dependant on the gene, except for *P. flammea* and *C. servana* for which only one individual is sampled. Since there is only one specimen per species each in *P. flammea* and *C. servana* for COII, EF-1α and TPI it is not possible to determine whether these species exhibit exclusivity at these loci. Two specimen of *P. flammea* however were available in the COI-barcoding tree suggesting the monophyletic status of this species for at least a small number of individuals. *Planotortrix avicenniae* and *C. fraterna* only form an exclusive group for the larger mt region of COI-COII but not for the COI-barcoding region itself. This pattern was expected since the larger region of COII has more power in terms of phylogenetic sequence information than the COI region. Each gene tree shows a different pattern of species exclusivity within each genus, displaying discordant genealogies. All markers studied here provide well supported gene genealogies, but the lack of concordance in displaying exclusivity between these closely related species and the disagreement between loci clearly support the mosaic view of the genome (Figure 1.2). Different loci in the genome represent different stages of a species gene history that are influenced by e.g. introgression or lineage sorting which, when completed, reflects reciprocal monophyly and show species as exclusive groups in a genealogical tree.

*Ctenopseustis filicis* is not reciprocally monophyletic for the mitochondrial markers but it is for the nuclear markers. This can be explained by two non-exclusive processes, either lineage sorting or gene introgression. In case of lineage sorting the development of a monophyletic pattern among loci can happen more quickly in some species than in others simply by chance.
If that is the case no general patterns of species sorting should be detectable. However since some of the species such as *P. puffini* and *P. notophaea* are always monophyletic for all observed loci and some species like *C. obliquana* and *C. herana* never form an exclusive group, it reflects whether lineages have actually sorted already or are still in the process of sorting. The speciation event that gave rise to *P. notophaea* and *P. puffini* likely occurred before that for *C. obliquana* and *C. herana*. Therefore *C. obliquana* and *C. herana* still share ancestral polymorphisms while *P. notophaea* and *P. puffini* have diminished gene flow and do not share these ancestral polymorphisms anymore.

An alternative cause for the different patterns observed among loci is the capability of loci to introgress between species. *C. filicis* might have hybridized in the past and this is still reflected in the mitochondrial loci, since they do not undergo recombination, therefore they might retain the signature of introgressed genes longer than any of the nuclear genes. Another example might be *C. her-2* as it clusters with *C. filicis* in *COI-COII* and also in *EF-1α* but not in *TPI*. It is likely that *C. filicis* and *C. herana* have hybridized in the past and still share ancestral polymorphism with each other. These two species are sympatric in the region where they were collected raising the possibility of a hybridization.

In the mitochondrial and *EF-1α* trees, *C. her-4* seems to be the only individual very closely related to the *C. obliquana* grouping. A similar pattern is observed in *TPI* but here *C. her-1A* and *C. her-2* are also polyphyletic with *C. obliquana*. Because of this pattern is seems likely that *C. her-4* is a morphological misidentification based on wing costal-fold ratios in male moths. The location where *C. her-4* was collected from is the Buller region on the South-Island in New Zealand where *C. obliquana* and *C. herana* overlap sympatrically in their habitat distribution. If this individual is excluded from all analyses, *C. obliquana* would indeed be monophyletic for at least the mitochondrial *COI-COII* region and the nuclear *EF-1α* gene. For *P. octo* and *P. excessana* none of the genes reflects a better phylogenetic pattern within the genus *Planotortrix*. The mitochondrial genes *COI-COII* and *TPI*, proposed to be Z-linked, should reflect monophyly by species better than *EF-1α* as the autosomal representative. *P. avicenniae* confirms this pattern and shows reciprocal monophyly for *COI-COII* but not for any of the nuclear markers. However the mitochondrial or sex-linked marker used in this study do not show any other species as exclusive groups more often than autosomal marker in the gene phylogenies for the native Leafroller. All loci, considering the larger region for the mitochondrial genes, show an equal
number of reciprocally monophyletic groups (Table 3.4). This might be due to the fact that only one representative for a nuclear and sex-linked marker was available to study. Also the actual choice of the marker is of significance, for example TPI might not be sex-linked in the native New Zealand Leafroller moths.

### 3.6.2 COI, COII and EF-1α are under purifying selection

The site-model analysis revealed that both COI-COII and EF-1α are under strong purifying selection with only a small proportion of sites that might be under weak positive selection in EF-1α. Cytochrome oxidase c is one of the key energy-generating proteins in respiration in virtually all organism (Capaldi, 1990) whereas EF-1α regulates the rate of polypeptide elongation during translation in ribosomes and is the second most abundant protein in the cell after actin (Condeelis, 1995). They both belong to highly conserved protein families and their functional role might explain why they are under strong purifying selection. This is not surprising since purifying selection on proteins is often observed due to high functional constraints in proteins (Li, 1997) and is predicted for a non-speciation gene according to the genic view. A comparative study by Yang et al. (2000) on ten different datasets of genes revealed that mitochondrial protein-coding genes in hominoids are highly conserved with an average ω values of 0.4 to 0.5 and a small proportion of 0.6% under diversifying selection. This is very similar to the ω values observed in the native Leafrollers and confirms that cytochrome oxidase c exhibits similar evolutionary patterns among different taxa.

The site-models implemented in this study assume that selective pressure indicated by the ω ratio is identical among evolutionary lineages (Yang et al., 2000) therefore individual testing of particular branches of interest is recommended. The branch-model tests undertaken for COI-COII and EF-1α however did not indicate positive selection for any of the branches tested, again as expected for a non-speciation gene. The following chapter investigates whether pheromone binding protein (PBP) displays patterns of positive selection along sites or branches as predicted for a candidate speciation gene.
4

A Candidate Speciation Gene

4.1 Introduction

Olfaction is a very important sense in moths that enables conspecific mates to locate each other. In many species female moths release a volatile sex pheromone that can be detected by male moths over distances. The first step in the recognition process involves pheromone-binding proteins, (PBPs), localized in the male moth’s antennae. PBPs are believed to transport the hydrophobic pheromone components to putative sex pheromone receptor proteins in the dendritic membrane of the sensory neurons (see General Introduction, section 1.4.2). Closely related species of moths use species-specific sex pheromones that can differ in as little as the position of a double-bond along the hydrocarbon backbone of the sex pheromone or variation of the amount of different components of a blend. The specificity in discriminating and binding components of the sex pheromones is thought to be, at least in part, encoded by PBPs (see section 1.4.2). The native New Zealand Leafroller moths of the genera *Ctenopseustis* and *Planotortrix* use blends of heptadeca (14)- or hexadeca (16)- acetates, that vary in the position of double-bonds at either
the Z5, Z7, Z8 or Z9 position (see General Introduction, section 1.5.4). A pheromone-binding protein has been isolated and characterized in the related Leafroller *Epiphyas postvittana* and shown to bind the major component of the sex pheromone blend. More recently the orthologous PBPs have been isolated from *Ctenopseustis* and *Planotortrix* (see General Introduction, section 1.5.6). This therefore represents a suitable system to address different hypotheses on the evolution of PBPs and whether PBP might be a speciation gene in these native Leafroller moths.

Adaptive evolution acts through natural selection on mutations in genes that affect the fitness of their carrier. If a new allele confers a higher fitness it tends to increase in frequency over time until it gets fixed in a population. This process is called positive or directional selection. On the other hand a gene that is under negative or purifying selection, as it decreases the carrier’s fitness, tends to get lost from the population. An allele under balancing selection is called neutral driven by genetic drift and dependant on the effective population size.

The neutral theory on the other hand claims that the overwhelming majority of evolutionary changes at the molecular level are not caused by selection acting on advantageous mutants, but by random fixation of selectively neutral or very nearly neutral mutants through the cumulative effect of genetic drift (Kimura, 1991). The neutral theory serves as an underlying basis for running selection tests such as likelihood ratio tests using the software package Phylogenetic Analysis by Maximum Likelihood (PAML). Ratios of non-synonymous ($d_N$) to synonymous ($d_S$) substitution rates or omega values $\omega = d_N/d_S$ are used to test whether sites or lineages are under selective pressure (Kimura, 1983; Ohta, 1993). An excess of non-synonymous substitutions over synonymous ones is an important indicator of positive selection at the molecular level. Speciation genes are predicted to be under positive selective pressures. As a result likelihood ratio tests were conducted to find evidence for positive selection in *PBP* as a candidate speciation gene.

### 4.2 Aims

It is expected that a speciation gene could show reciprocal monophyly among species in a gene tree due to the absence of shared ancient polymorphism. Gene flow at speciation loci is reduced or completely absent and often characterized by a selective sweep (see General Introduction, section 1.2). In this chapter, pheromone binding protein 1 is tested as a candidate speciation gene. The three objectives for this chapter are:
1. test whether the \textit{PBP} gene shows reciprocal monophyly or exclusivity among species

2. test whether PBP1 is under selection or shows evidence of a selective sweep

3. relate differences in PBP1 among species to the different sex pheromones that they use through 3D protein modelling

### 4.3 Materials and Methods

#### 4.3.1 Specimen dataset

Thirty-four individuals from the datasets in Chapter 3 Tables (3.1) and (3.2) and three additional specimens were combined for the PBP1 dataset analysis (Table 4.1). All sequence data for specimens from Table 3.2 except \textit{C.fra}-1 and \textit{P.octi}-2 were amplified in a previous study on pheromone binding protein by Tamara Sirey (2000).

#### 4.3.2 PCR amplification and sequencing

PCR amplifications were setup using genomic DNA (gDNA) and insect-specific primers. The \textit{Pheromone Binding Protein 1 (PBP1)} was amplified with the primer pair nztort5 and nztort3’un (Sirey, 2000), Table 4.2. Products were amplified using the following cycling profile of 94°C for an initial 2mins, followed by 30 cycles of 94°C for 10sec, 55°C for 30sec and 72°C for 2mins. Amplification reactions of 50 µL consisted of 10pmol of each primer, 1.5mM MgCl$_2$, 0.2mM of each dNTP and 2 units of \textit{Taq} polymerase (INVITROGEN). PCR products were electrophoretically separated on a 1% agarose gel and stained with ethidium bromide (1 mg/ml). Where two PCR products of different lengths were observed on gels, both bands were sequenced. To obtain the sequence for each band, amplified products were then shot-gun cloned into pGEM-T Easy Vector System (PROMEGA). The templates obtained from three to five colonies per individual were sequenced on an ABI3730 Genetic Analyzer at the Allan Wilson Center in Albany (AWCGS) using vector specific primer (M13 forward and M13 reverse) and external primer as used in the PCR reaction along with additional species-specific primers for direct sequencing of the PCR product ((Sirey, 2000), see Table 4.2).
### 4.3 Materials and Methods

#### Table 4.1. PBP1 gene dataset

<table>
<thead>
<tr>
<th>Species name</th>
<th>Tree-ID</th>
<th>Location</th>
<th>Area code</th>
<th>Collector</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-1</td>
<td>Mt Albert</td>
<td>AK</td>
<td>HortResearch*</td>
<td>1982</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-2</td>
<td>Pollen Island</td>
<td>AK</td>
<td>Dugdale, J.S.</td>
<td>12/02/1999</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-3</td>
<td>Hamilton</td>
<td>?</td>
<td>?</td>
<td>30/03/1999</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-4</td>
<td>Wenderholm</td>
<td>AK</td>
<td>Langhoff, P.</td>
<td>30/12/2003</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-5</td>
<td>Whakapapa</td>
<td>TO</td>
<td>Hoare, R.</td>
<td>12/04/2003</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-6</td>
<td>Hamilton</td>
<td>WK</td>
<td>Langhoff, P.</td>
<td>04/04/2004</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-7</td>
<td>Mt Albert</td>
<td>AK</td>
<td>HortResearch*</td>
<td>2004</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>C.obl-8</td>
<td>Mt Albert</td>
<td>AK</td>
<td>Newcomb, R.</td>
<td>2004</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-1</td>
<td>Tai Tapu</td>
<td>MC</td>
<td>HortResearch*</td>
<td>1990</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-2</td>
<td>Bluff</td>
<td>SL</td>
<td>Langhoff, P.</td>
<td>27/01/2004</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-3</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Dugdale, J.S.</td>
<td>20/01/2004</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-4</td>
<td>Coll Creek</td>
<td>BR</td>
<td>Dugdale, J.S.</td>
<td>07/12/2003</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-5</td>
<td>Tai Tapu</td>
<td>MC</td>
<td>HortResearch*</td>
<td>2004</td>
</tr>
<tr>
<td><em>C. herana</em></td>
<td>C.her-6</td>
<td>Tai Tapu</td>
<td>MC</td>
<td>Newcomb, R.</td>
<td>2004</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td>C.fra-1</td>
<td>Tiritangi</td>
<td>AK</td>
<td>?</td>
<td>20/12/1998</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td>C.fra-2</td>
<td>Taurewa</td>
<td>TO</td>
<td>?</td>
<td>06/12/1998</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>C.fil-1</td>
<td>Bluff</td>
<td>SL</td>
<td>Langhoff, P.</td>
<td>06/01/2004</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>C.fil-2</td>
<td>Moonlight Ecol. Res.</td>
<td>BR</td>
<td>Dugdale, J.S.</td>
<td>06/12/2003</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>C.fil-3</td>
<td>Half Moon Bay</td>
<td>SI</td>
<td>Dugdale, J.</td>
<td>1999</td>
</tr>
<tr>
<td><em>C. servana</em></td>
<td>C.ser</td>
<td>Muriwai</td>
<td>AK</td>
<td>Newcomb, R.D.</td>
<td>14/09/1999</td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td>P.avi-1</td>
<td>Pollen Island</td>
<td>AK</td>
<td>?</td>
<td>26/03/1999</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>P.exc-1</td>
<td>Rukuhia</td>
<td>WO</td>
<td>HortResearch*</td>
<td>1987</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>P.exc-2</td>
<td>Bluff</td>
<td>SI</td>
<td>Langhoff, P.</td>
<td>27/01/2004</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>P.exc-3</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Dugdale, J.S.</td>
<td>02/12/2003</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-1</td>
<td>Canterbury</td>
<td>MC</td>
<td>HortResearch*</td>
<td>1982</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-2</td>
<td>Te Puke</td>
<td>BP</td>
<td>McKena, C.</td>
<td>30/12/2003</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-3</td>
<td>Whakapapa</td>
<td>TO</td>
<td>?</td>
<td>13/01/2004</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-4</td>
<td>Brooks Reservoir</td>
<td>NN</td>
<td>Langhoff, P.</td>
<td>19/02/2004</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-5</td>
<td>Okut valley</td>
<td>MC</td>
<td>?</td>
<td>29/03/2004</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>P.octo-6</td>
<td>Canterbury</td>
<td>MC</td>
<td>Newcomb, R.</td>
<td>2004</td>
</tr>
<tr>
<td><em>P. octoide</em></td>
<td>P.octi-1</td>
<td>Pitt Island</td>
<td>CH</td>
<td>Marris, J.W.M.</td>
<td>1998</td>
</tr>
<tr>
<td><em>P. puffini</em></td>
<td>P.puf-1</td>
<td>Lee Bay</td>
<td>SI</td>
<td>Langhoff, P.</td>
<td>21/01/2004</td>
</tr>
<tr>
<td><em>P. puffini</em></td>
<td>P.puf-2</td>
<td>Golden Bay Track</td>
<td>SI</td>
<td>Newcomb, R.D.</td>
<td>01/06/2004</td>
</tr>
<tr>
<td><em>P. puffini</em></td>
<td>P.puf-3</td>
<td>Lee Bay</td>
<td>SI</td>
<td>Dugdale, J.S.</td>
<td>1999</td>
</tr>
<tr>
<td><em>P. notophaea</em></td>
<td>P.not-1</td>
<td>Palmerston North</td>
<td>WI</td>
<td>?</td>
<td>1990</td>
</tr>
<tr>
<td><em>P. notophaea</em></td>
<td>P.not-2</td>
<td>Brook Reservoir</td>
<td>NN</td>
<td>Langhoff, P.</td>
<td>22/03/2004</td>
</tr>
<tr>
<td><em>P. flammea</em></td>
<td>P.flam</td>
<td>Kaipakati Point</td>
<td>BR</td>
<td>Dugdale, J.S.</td>
<td>2001</td>
</tr>
</tbody>
</table>

*Moths were provided by the HortResearch Rearing facility; for area codes see Figures (2.1) and (2.2)*

#### 4.3.3 Sequence analysis

Chromatograms of PBP1 gene sequence were edited and aligned using SEQUENCHER 4.5 (Gene Codes Corporation, Inc.). The PBP1 nucleotide alignment was manually adjusted by eye using
Table 4.2. Primer list for PBP1

<table>
<thead>
<tr>
<th>Name</th>
<th>5’ to 3’</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>nztor5</td>
<td>TCNAARCARGTNNTNGARGGNATG</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>nztor3’un</td>
<td>TACAGCGACAAGGTCCAATG</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Cintfor</td>
<td>TCAATGACAAATGGTACTGCCGC</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Cintrev</td>
<td>CGGTGACTATATGGAGACCGTG</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Cfintfor</td>
<td>CGGGAATACCCCTAGAATCGC</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Cfintrev</td>
<td>GTCTGACTGGCGATCAACGC</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Pintfor</td>
<td>ATTAAGTAACTGGGAAATCCAC</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Pintrev</td>
<td>TGCTTATACTATGAGCCAG</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Pnoint</td>
<td>CGCCGGAGATCAATTTGCATG</td>
<td>(Sirey, 2000)</td>
</tr>
<tr>
<td>Pexint</td>
<td>CCTTCAATGGATGTGCTAGC</td>
<td>(Sirey, 2000)</td>
</tr>
</tbody>
</table>

the PBP1 translated amino-acid alignment as a reference.

The PBP1 nucleotide alignment was first analysed using Modeltest 3.7 (Posada & Crandall, 1998) and PAUP* (Swofford, 2003) to obtain an appropriate model of sequence evolution. Phylogenetic trees were constructed using MrBayes (Huelsenbeck & Ronquist, 2001). See section 3.3.5 for details of analysis. To test whether sites are under selection $\omega = d_N/d_S$ ratios were estimated using CODEML from the PAML 3.15 package (Yang, 1997), see section 3.3.6 for details. The nucleotide dataset used for the PAML analysis was reduced due to technical problems in running the analysis. Identical haplotypes were removed and sequences with missing data were excluded. The remaining dataset consists of 32 sequences from 28 specimens. According to the neutral theory of evolution, rates of divergence between species should equal rates of polymorphism within species (Kimura, 1977). A simple method of testing for deviations from neutral expectations is the McDonald-Kreitmann test (McDonald & Kreitman, 1991). This test considers the relationship of synonymous (silent) and non-synonymous (replacement) nucleotide differences between and within populations using a contingency table and a $\chi^2$ test. Only variable sites are taken into account and divided into fixed and polymorphic sites. The Null hypothesis assumes that the ratio of non-synonymous fixed sites versus synonymous fixed sites ($n_F/s_F$) is the same as the ratio of non-synonymous polymorphic sites versus synonymous polymorphic sites ($n_P/s_P$). If the ratio of ($n_F/s_F$) is significantly higher than the ratio of ($n_P/s_P$) it might be indicative of adaptive evolution between two species. If the ratio is lower it might be due to purifying selection. Fixed and polymorphic sites were manually counted and an online Fishers Exact Probability test was used to calculate p-values (http://faculty.vassar.edu/lowry/webtext.html).

A 3D structure homology model of PBP1 for the native Leafroller, *P.octo-2* was generated
using the SWISS-MODEL modelling server. The \textit{PBP1} sequence from \textit{P.octo-2} to be modelled was loaded on to the SWISS-MODEL server and was subsequently used to identify the possible structure templates for homology modelling from the SWISS-MODEL server template library ExPDB (Guex & Peitsch, 1997). From the templates the programme identified the crystal structure of \textit{Bombyx mori} pheromone binding protein, BmorPBP (PDB entry 2FJY) and the NMR structure of \textit{Antheraea polyphemus} PBP, ApolPBP (PDB entry 1QWV) were selected as templates for structure modelling. The two template structures 2FJY and 1QWV were superimposed by iterative fitting and was used to construct the homology model of the \textit{P.octo-2} PBP1. The project file was then submitted to the SWISS-MODEL server for model building and model evaluation. The resulting PDB file was analysed for features such as the pheromone binding pocket and its relevant amino acid residues using the programme PyMOL (DeLano Scientific).

4.4 Results

4.4.1 PBP phylogeny

Thirty-seven \textit{PBP1} gene fragments from 12 species of native New Zealand Leafrollers were amplified. \textit{PBP1} sequence data was not recovered for \textit{C. obl-7}. The \textit{PBP1} gene varied in length depending on the species due to variation in intron size and protein coding lengths differences. \textit{PBP1} gene contains three exons, (1, 2 and 3) separated by two introns (1 and 2). The protein coding region yielded a 408bp nucleotide alignment consisting of 3 exons respectively. Exon one encodes the first 22 amino acids of the mature protein and is 66bp in length. Exon two encodes the next 59 amino acids and is 177bp long. Exon three has a length difference between the genera, \textit{Planotortrix} with a length of 186bp encoding 62 amino acids, in \textit{Ctenopseustis} it yields 180bp encoding 60 amino acids. Lengths for intron 1 varied from 340bp to 390bp in \textit{Ctenopseustis} and 818bp to 1717bp in \textit{Planotortrix}. The lengths of intron 2 varied from 457bp to 730bp in \textit{Ctenopseustis} and from 67bp to 533bp in \textit{Planotortrix}. In \textit{Planotortrix}, \textit{P. octo}, \textit{P. avicenniae} and \textit{P. notophaea} had more than one allele of \textit{PBP1} detected by different intron sizes of the amplified fragment. These individuals have two alleles coded as allele A and B respectively as shown in Figure 4.2. The exon dataset included 27 specimens with full length sequences of 408bp, eleven of them had 21bp to 45bp missing at the 5’ end and eight were missing 8bp to 188bp at
the 3’ end. The evolutionary relationships among the native Leafrollers for *PBPI* based on the exon nucleotide alignment is shown in Figure 4.1. Support values are posterior probability scores from the MrBayes analysis. The tree was rooted between the genera. The genera *Planotortrix* and *Ctenopseustis* are both monophyletic. *P. puffini* and *P. notophaea* form well supported monophyletic groups with a posterior probability of 1.0. *P.octo-3* is sister to these two species whereas *P. octo* and *P. excessana* display a paraphyletic pattern. Within this paraphyletic clade two groupings are observable. *P.octo-6* and one allele of *P.octo-1* form a subgroup and are sister to *P.octo-3, P.octo-5* and *P.exc-3*. This clade is well supported with a posterior probability of 0.8 and forms a sister clade to the second grouping of *P. octo* and *P. excessana* alleles. This clade consists of *P.octo-2, -4* and one allele of *P. octo-1* and the two individuals *P.exc-2* and -1.

Within this paraphyletic grouping both *P. avicenniae* representatives form an exclusive group with a posterior probability of 1.0.

*Ctenopseustis* is divided into two major clades (posterior probability of 1.0). The first clade includes all specimen of *C. herana* and *C. obliquana* displaying a non-exclusive pattern. The single representative of *C. servana, C.ser*, groups within *C. herana* and *C. obliquana*. *C.her-5* and -1 are more closely related to each other as they form a monophyletic subgroup with high support of 1.0 whereas *C.her-3* is more closely related to *C.obl-3* and *C.obl-2* with a posterior probability score of 0.8. The remaining six taxa of *C. obliquana, C.obl-4* to -7 and -3 to -1 are grouping with the remaining three individuals of *C. herana, C.her-2, -4* and -6. The second clade leads to the fern-feeding species *C. filicis* and *C. fraterna*, which form an exclusive group with high support of 1.0. The branch leading to *C. fraterna* and *C. filicis* is very long. Within this clade all specimen of both taxa form monophyletic subgroups and are sister to each other. *C. fraterna* and *C. filicis* are the only species that use a 16-carbon acetate as a sex pheromone whereas all the remaining species in the native Leafrollers use a 14-carbon acetate.

### 4.4.2 Tests for selection in PBP

Tests for selection using PAML revealed that *PBPI* is under purifying selection (Table 4.3). The average $\omega$ values ranges from 0.218 to 0.269. This suggest that a non-synonymous mutation has a 2.18 - 2.69% chance of being fixed in the population. Model 3 suggests that a small proportion of sites (1.8%) are under positive selection with an $\omega$ value of 3.621 and model 8 predicted that 2.3% are under positive selection with an $\omega$ of 3.25. One amino acid site at position 44,
4.4 Results

Figure 4.1. Bayesian inference gene phylogeny for *Planotortrix* and *Ctenopseustis* species based on PBP1 sequences. Posterior probability values from MrBayes are shown above the branches. Colours were used for four species in the tree to emphasize their non-exclusive status. Tree was rooted between the genera.
was predicted by model 3 and 8 to be under positive selection with an $\omega$ of 3.5 and 3.2 ($p > 0.95$). The parameters for the $\beta$-distribution ($p$ and $q$) indicate a L-shaped, highly left-skewed distribution this indicates that the $\omega$ values are very close to 0.

Table 4.3. Likelihood values and parameter estimates for PBP

<table>
<thead>
<tr>
<th>Model code</th>
<th>l</th>
<th>$d_N/d_S$</th>
<th>$\omega$ ESTIMATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0 (one-ratio)</td>
<td>-1066.12</td>
<td>0.218</td>
<td>$\omega = 0.218$</td>
</tr>
<tr>
<td>M1a (neutral)</td>
<td>-1051.14</td>
<td>0.221</td>
<td>$p_0 = 0.834$, ($p_1 = 0.165$)</td>
</tr>
<tr>
<td>M2a (selection)</td>
<td>-1049.74</td>
<td>0.269</td>
<td>$p_0 = 0.837$, $p_1 = 0.146$, ($p_2 = 0.015$), $\omega_2 = 3.812$</td>
</tr>
<tr>
<td>M3 (discrete)</td>
<td>-1049.70</td>
<td>0.262</td>
<td>$p_0 = 0.808$, $p_1 = 0.172$, ($p_2 = 0.018$) $\omega_0 = 0.065$, $\omega_1 = 0.829$, $\omega_2 = 3.621$</td>
</tr>
<tr>
<td>M7 (beta)</td>
<td>-1052.21</td>
<td>0.221</td>
<td>$p = 0.154$, $q = 0.545$</td>
</tr>
<tr>
<td>M8 (beta &amp; $\omega$)</td>
<td>-1049.84</td>
<td>0.258</td>
<td>$p_0 = 0.097$, ($p_1 = 0.023$) $p = 0.259$, $q = 1.122$, $\omega = 3.25$</td>
</tr>
</tbody>
</table>

The nested model comparison of M0 versus M3 is significant and implies that there is no uniform selective pressure acting on PBP1 (Table 4.4). The likelihood scores for model M2a fit the data better than M1a but is not significant. Model M8 was a slightly better fit than model M7 but again not significantly, implying that no positive selection pressure is detectable.

Table 4.4. Nested model results of PAML analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>2$\Delta l$</th>
<th>p-value for $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M0-M3 (df=4)</td>
<td>M1a-M2a (df=2)</td>
</tr>
<tr>
<td>PBP</td>
<td>16.4219</td>
<td>1.4067</td>
</tr>
</tbody>
</table>

* $p < 0.01$

One of the species of Planotortrix, *P. octo* showed some interesting amino acid changes in the amino acid alignment (see Appendix B on CD). The total number of polymorphic sites is six, where four are non-synonymous and two are synonymous or silent changes. The number of silent changes is very low and compared to a closely related species *Apoctena orthocopa* (see section 1.5.2) with 43 synonymous sites out of 74 polymorphic sites in total for the orthologous PBP1 gene. PBP1 of *A. orthocopa* has an average amino acid identity of 74.3% with *Ctenopseustis* and 74.7% with *Planotortrix* (Sirey, 2000). Comparing the ratios of replacement and silent changes at fixed sites versus replacement and silent changes at polymorphic sites, the MacDonald-Kreitman test was used to test whether the observed changes in *P. octo* are consistent with neutral expectations. The non-synonymous and synonymous changes for all members of *P. octo* compared to *A. orthocopa* as an outgroup were not significant, ($p=0.39$, two-tailed,
4.4 Results

N=7) so the null hypothesis of a neutral evolution at this loci cannot be rejected.

4.4.3 PBP amino acid variation between Planotortrix and Ctenopseustis

The translated amino acid alignment yielded 137 amino acids from 37 individuals of all 12 species of native Leafroller. The full mature PBP1 protein consists of 3 exons and was represented in 25 taxa, 11 individuals had 15 amino acids or less missing at the N-terminus and 5 individuals had 61 amino acids missing from the C-terminus. There is a six nucleotide INDEL between Ctenopseustis and Planotortrix species resulting in a two amino acid difference. This results in a total of 60 amino acids for Ctenopseustis PBP and 62 amino acids in Planotortrix. These two additional amino acids in Planotortrix are glycine and leucine at positions 105 and 106 in the amino acid alignment. Between the two genera the highest amino acid sequence identity was between P.octo-1A and C.obl-4 at 89.8% and the lowest was between P.octo-1A and both C. fraterna specimens at 35.3%. An alignment of variable sites among species is shown in Figure 4.2.

There are ten fixed amino acid sites between Ctenopseustis and Planotortrix at position 28, 57, 65, 67, 78, 88, 92, 114, 122 and 128. Within Ctenopseustis there are two species, C. filicis and C. fraterna that share nine differences between them (Lys13, Leu16, Leu24, Arg50, Asp76, Ser77, Gln99, Phe132 and Val133) and all other taxa. C. servana has four unique residues (Met3, Ala14, Gln59 and Asn37) that differ from the rest of the species within the genus. In Planotortrix, P. flammea (Ser38 and Ala92) and P. notophaea (Leu16 and Gly95) both contain two unique residues, whereas P. puffini (Phe76) has only one different amino acid change within the genus but shares (Leu34) with P. notophaea. Both species and P. flammea also have residues (Val24 and Lys59) in common. The P. excessana, P. octo and P. avicenniae complex contains four amino acid variable positions (6, 59, 85 and 120). Both alleles of P. avi-1 are identical and have the combination of Ser, Gln, Leu and Ser for the above mentioned residue changes. Two individuals of P. excessana, P. exc-2 and P. exc-1 and two specimen of P. octo, P.octo-2 and allele A of P.octo-1 share the same combination of amino acid changes as P. avicenniae. However P. octo has more than two combinations of different amino acids at these sites. The second combination has a possible Gly or Ser at position 6, a Thr at 59, a Val at 85 and a Asn at position 120 as represented by P.octo-3, P.octo-5 and also P.exc-3. A change to Gly at positions 6 and 59 ,while Val at 85 and Asn at 120 remain the same, display a third possible
**Figure 4.2.** Alignment of PBP1 across *Ctenopseustis* and *Planotortrix* showing only variable amino acid sites. Residues highlighted in yellow indicate changes within the *P. octo, P. excessana* and *P. avicenniae* complex. Green highlights residues that are fixed between *C. fraterna* and *C. filicis* and all other species. Amino-acid positions are given in the first line of this figure and correspond to the full AA alignment of PBP. Identical residues to *P. octo-6* PBP1 are represented as a dot and missing data as a question mark. The amino acids are shown in 1-letter code.
4.4 Results

combination of different residues in *P. octo* as represented in *P.octo-6* and allele B of *P.octo-1*. *P. octoides* shares a Gly at position 6 with *P.octo-6* and *P.octo-1B* but a Gln at 59, Leu at 85 and Ser at 120 for the remaining positions as represented by *P. avicenniae* and two individuals of *P. octo*, *P.octo-2* and -4, and also *P.exc-2* and -4.

4.4.4 A 3D protein structure model of PBP for *Planotortrix octo* PBP1

To determine the positions of various amino acid changes in the three dimensional (3D) structure of the PBP protein for *P. octo*, *C. fraterna* and *C. filicis*, a protein structure model of PBP1 for *Planotortrix* was created. *P.octo-2* was chosen as a representative and aligned with the amino acid sequence of BmoriPBP (2FJY) and ApolPBP (1QWV) see Figure 4.3. The sequence identity between *P.octo-2* and BmorPBP (2FJY) is 54% or 77 identical residues out of 142 and 52% or 75 identical residues with ApolPBP (1QWV).

Figure 4.3. Structure-based sequence alignment of *P.octo-2* PBP1, *Bombyx mori* PBP, BmorPBP (PDB entry 2FJY) and *Antheraea polyphemus* PBP, ApolPBP (PDB entry 1QWV). Residues that are fully conserved are shown in white lettering on a red background, while conservatively substituted residues are shown in red on a white background. An asterisk below the sequences indicates the position of amino acid changes among *P. octo*, *P. avicenniae* and *P. excessana* PBP1 sequences. The numbers in green underneath the sequences indicate the C-residues that form disulphide bonds. The secondary structure elements and the residue numbering for ApolPBP are displayed above the sequence. The figure was generated using ESPript (Gouet et al., 1999).

*P.octo-2* was predicted to have the same six α-helices and six conserved cysteine residues present in the structure of *B. mori* PBP. In *A. polyphemus* at physiological pH the helices 1 and 3 are subdivided into helix-α1a-c and helix-α3a-b, and together with helices α2, 4, 5 and 6 form a total of nine helices. *P.octo-2* has two additional amino acids, Gly and Leu on the loop.
between \( \alpha5 \) and \( \alpha6 \) at positions 105 and 106. *P. octo*-2 lacks a Pro at position 64 compared to both *B. mori* and *A. polyphemus*. The PBP1 structure of *A. polyphemus* in the 3D model (Figures 4.4 and 4.5) was chosen over the one of *B. mori* to analyse the active site because the natural pheromone, that is bound by PBP in *A. polyphemus* (ApolPBP), is a \((6E,11Z)\)-hexadeca-6,11-dienyl-1-acetate and more similar to the ones in the native Leafrollers that use hepta- \((14:OAc)\) or hexadeca- \((16:OAc)\) acetates. They are both acetates while BmorPBP binds an alcohol.

The structure of ApolPBP contains a large hydrophobic cavity that is lined by the side-chains of 29 residues (Mohanty *et al.*, 2004) that interact with the pheromone. Ten of these residues are identical in *P. octo*-2 at Phe12, Phe36, Ile52, Ala56, Ala73, Phe76, Ala77, Ile94, Ala115 and Ile122. The remaining 19 residues are different from ApolPBP based on comparison with the BmorPBP active site, nine residues in ApolPBP are predicted to be forming direct interactions with the pheromone (Klusák *et al.*, 2003; Sandler *et al.*, 2000). These are Leu8, Ser9, Phe12, Phe36, Trp37, Ser56, Met61, Phe76 and Phe118 in ApolPBP. Of these nine residues four residues are identical in *P. octo*-2 (Phe12, Phe36, Trp37 and Phe76). In *P. octo*-2 at positions 56 and 61 Ala and Glu replaced Ser and Met and Tyr replaced Phe at position 118, which is identical to it except for an additional OH-group. These five aromatic residues (Phe12, Phe36, Trp37, Phe76 and Phe118) are highly conserved in lepidopteran PBPs. In *B. mori* residues Phe12 and Phe118 are thought to stabilize the double-bonds in bombykol.

The four variable amino acid positions among *P. octo*, *P. excessana* and *P. avicenniae* PBP1s at positions 6, 59, 85 and 120 were mapped onto the model structure along with nine residues in ApolPBP that are predicted to be involved in binding the pheromone. There are five different combinations of amino acids for positions 6, 59, 85 and 120 from the sequence alignment (see Figure 4.2). Residue 6 (13 in ApolPBP) is located on the loop connecting \( \alpha \)-helix 1a with \( \alpha \)-helix 1b in the model structure, residue 59 (67 in ApolPBP) is on the loop region connecting \( \alpha \)-helix 3b with \( \alpha \)-helix 4 which form the acetate-binding cavity in ApolPBP. Residue 85 (93 in ApolPBP) is located on \( \alpha \)-helix 5 and close to Leu90 and Ile94 which are part of the 29 residues that line the cavity with their side-chains and interact directly with the pheromone in ApolPBP. Residue 120 (126 in ApolPBP) is the first residue forming the C-terminal tail of the protein. One combination of residues at these 4 positions is Ser, Thr, Leu and Ser (STLS) (see Figure 4.4). This allelic variant of PBP1 is found in *P. octo*-4 only. A second combination of Ser, Glu,
Figure 4.4. Ribbon diagram of the predicted structure of *P.octo*-2 PBP1 (blue) and the structure of *A. polyphemus* PBP1 (orange). The eight residues forming the binding pocket in ApoPBP (1QWV) are shown as green sticks while the four variable sites among *P.octo*, *P. excessana* and *P. avicenniae* are shown as red sticks. Amino acids are shown in 3-letter code and numbering corresponds to the positions shown in the alignment of Figure 4.3. Figure was generated using PyMol (DeLano Scientific).
4.4 Results

Leu and Ser (SQLS) is found in six individuals from three species with three different pheromone blends, *P. octo-2* and *P. octo-1A* that uses Z8-14:OAc, *P. exc-2* and *P. exc-1* that use a blend of Z5-, Z7-14:OAc and both alleles of *P. avi-1* that only uses Z5-14:OAc. A third combination is a Gly, Gly, Val and Asn (GGVN) used by two individuals of *P. octo* again, *P. octo-6* and -1B. There are two additional combinations that are found in *P. exc-3* with Ser, Thr, Val and Asn (STVN) and also a possible Thr, Val and Asn (XTVN) in *P. octo-5* where X represents missing data at residue 6.

Positions 6, 85 and 120 are on the surface of the predicted *P. octo-2* PBP1, only position 59 is inside the binding pocket close to Phe76. The residues at position 59, situated on the loop connecting α-helix 3b with helix 4 that form the narrow end of the binding cavity, are either Gly, Thr or Glu. Both Thr and Glu are polar whereas Gly represents a tiny nonpolar residue. Since these three residues also vary in size, Glu has 5 C-atoms followed by Thr with 4 and Gly with 2 the larger amino acids could possibly restrict the space in the pocket where the pheromone binds. Positions 6, 85 and 120 are on the surface of the protein and might be playing a role when PBP is interacting with the pheromone receptor and ligand release.

4.4.5 Mapping amino acid changes in *C. fraterna* and *C. filicis* onto *P. octo* model structure

In *C. fraterna* and *C. filicis* there are nine amino acid fixed differences between the native Leafroller PBP at position Lys13, Leu16, Leu24, Arg50, Asp76, Ser77, Glu99, Phe132 and Val133 (see Figure 4.5). These different residues are mapped onto the model structure of *P. octo-2* and compared to ApolPBP (1QWV). There are four identical amino acids out of the nine mentioned above between *C. fraterna*, *C. filicis* and ApolPBP, these residues are Lys13, Leu16, Asp76 and Val133. Lys13 at position 20 in ApolPBP and Leu16, at position 23 in ApolPBP, are both located on helix-α1c, Lys13 on the surface and Leu16 pointing into the cavity of the protein. Asp76 at position 83 in ApolPBP is located on the loop between helix-α4 and helix-α5 and Val133 at position 140 in ApolPBP is found on the C-terminus of the protein. At position 31 in ApolPBP Ala24 → Leu24 is located on the surface of the protein on helix-α2. Thr50 → Arg50, at position 57, found on helix-α3b and points into the binding pocket of ApolPBP protein. Glu77 → Ser77 is the first residue on helix-α5 in ApolPBP at position 84. Asp99 → Glu99 at position 106 in ApolPBP is found on the loop connecting helix-α5 with helix-α6 also.
Figure 4.5. Ribbon diagram of the predicted structure of *P. octo-2* PBP1 (blue) and the structure of ApolPBP1 (1QWV) (orange) showing fixed amino acid changes between *C. fraterna* and *C. filicis*. Eight residues forming the binding pocket in (1QWV) are shown as green sticks and nine amino acid residues of *C. fraterna* and *C. filicis* as magenta sticks. Amino acids are shown in 3-letter code and numbering corresponds to the positions of ApolPBP as shown in the alignment of Figure 4.3.
pointing into the binding cavity. Glu132 → Phe132 at position 139 in ApolPBP is located on
the C-terminus of the protein.

4.5 Discussion

Speciation genes are predicted to be under positive selective pressures measured as the ratio of
non-synonymous versus synonymous substitution rates.

However the PAML analysis, to test whether PBP1 is under positive selection, showed that
PBP1 is under purifying selection. Only a very small percentage of sites were predicted to be
under positive selection. The dataset used for this analysis has been pruned and was very small
due to technical difficulties running the program PAML. A more extensive sampling of individ-
uals and a larger dataset is necessary to confirm these results. The amino acid alignment for the
native Leafrollers showed that there is a low level of variation within P. octo specimen and many
haplotypes are identical. This high frequency of identical haplotypes in P. octo may be indica-
tive of a selective sweep. However to perform tests of haplotype diversity that test for selective
sweeps a bigger dataset is necessary to infer statistically relevant analyses. Four replacement
substitutions among the P. octo PBP sequences were found but only two synonymous substi-
tutions. Compared to the closely related species, A. orthocopa, the number of non-synonymous
changes is much higher than the number of synonymous changes (74 replacement versus 43
silent). The MacDonald-Kreitman test is not significant however. This is probably due to a low
sample size with only 7 sequences obtained from P. octo. Low synonymous substitution rates
could also be indicative of a recent selective sweep or a relaxed constraint in PBP however again
the sampling sizes are very small and more sampling is required to confirm the possibility of a
selective sweep. In most genes the number of synonymous or silent substitutions is much higher
than the number of non-synonymous or replacement substitutions due to purifying selection or
constraints on genes and proteins (Kimura, 1977).

The existence of multiple amino acid combinations at position 6, 59, 85 and 120 in P. octo
specimen as shown in Figure 4.2 implies that not only one version of a PBP exists in these moths.
This could be explained by multiple allelic versions of PBP1. This has been found in another
tortricid moth, Epiphyas postvittana where two major forms of PBP1 were identified by native
PAGE and attributed as allelic at a single PBP locus (referred to as a 'fast' and 'slow' form)
(Newcomb et al., 2002). If P. octo and P. excessana have hybridized in the past the combination
of (SQLS) at residues 6, 59, 85 and 120 might be ancestral as it is shared by individuals of *P. octo*, *P. excessana* and *P. avicenniae*. It is interesting that these different allelic versions of PBP are not correlating with the taxonomic identities, although it appears that the combinations (STLS), (GGVN) and (XTVN) are only found in members of *P. octo*, whereas (STVN) is unique for *P. excessana* and (SQLS) is shared by *P. octo*, *P. excessana* and *P. avicenniae*. Behavioural response studies on male moths in *C. herana* and *C. obliquana* have shown that males show weak responses to the opposite pheromone of the species (Foster et al., 1997). Males of *C. herana* responded to Z8-14:OAc and males of *C. obliquana* to Z5-14:OAc and 14:OAc as well. Hybrids between these two species are even more variable to responses of opposite pheromones. Multiple allelic versions of PBP could explain this behaviour. Unfortunately no such behavioural studies have been conducted in *Planotortrix*.

Among the species that form exclusive groupings most of them use a 14 carbon acetate (14:OAc) with different desaturation positions at Z5, Z7, Z8 and Z9 except for the two fern-feeding species of *C. fraterna* and *C. filicis* that also use a 16 carbon acetate (16:OAc) as part of their sex pheromone blend. PBP1 in *C. fraterna* and *C. filicis* species might have a higher binding specificity for a 16-carbon acetate than for a 14:OAc. Such differences have been demonstrated in binding studies of other moths where the presence of multiple PBPs in a species reflected different binding preferences and affinities for sex pheromone components (Maida et al., 2000). In *Antheraea polyphemus*, for example, the predominant ApoPBP1 binds much more tightly to the main pheromone component than to the minor alcoholic component (Du & Prestwich, 1995; Prestwich, 1996). The ability to recognize different chain-lengths in the pheromone has been hypothesized due to two residues Leu8 and Ser9, which are not conserved among PBPs (Mohanty et al., 2004). *A. polyphemus*, *B. mori* and *Heliothis virescens*, for example, use a 16-carbon acetate and their PBPs have serine and leucine at the above mentioned positions. Whereas moths that use a 14-carbon acetate such as *Ostrinia nubilalis*, *Choristoneura rosaceana* and most of the native Leafroller have Met8 and Thr9 residues. These residues are bigger than serine and leucine, therefore possibly restricting the available space in the cavity so shorter 14-carbon chain pheromones are more likely to be able to fit. However for the two species of *C. filicis* and *C. fraterna* that use a 16-carbon acetate both residues Met8 and Thr9 are identical to those using shorter 14-carbon based pheromones. The ability to discriminate between chain-lengths might be encoded by different amino acids in these species. Interestingly
both species have an amino acid difference at position Arg50 and Glu99 compared to the other native Leafrollers. These two residues point into the binding pocket and are in close proximity to Asn53 which is hypothesized to play a key role in binding the pheromone in ApolPBP (Mohanty et al., 2004). Alternatively the two residues Leu24 and Lys13 which are different again to the rest of the natives in *C. fraterna* and *C. filicis* are located on the surface of the protein. If the specificity to recognize different chain-lengths is not encoded in the PBPs but lies within the Pheromone receptor, residues on the surface of PBP that might interact with the receptor might be involved in chain length discrimination. Studies on the *Ostrinia nubilalis* and *O. furnicalis*, the european and asian corn borer, suggest that the PBP/pheromone complex is responsible for the discrimination between different E and Z isomers since the PBP itself is identical in both species (Willett & Harrison, 1999b).

The PBP1 gene tree shows that *C. fraterna* and *C. filicis* form a reciprocal monophyletic group. Previous studies have also shown that the lineage leading to both species in a genealogical tree is under positive selection (Sirey, 2000). PBP1 is also exclusive for *P. notophaea*, *P. avicenniae*, *P. octoides*, *P. flammea*, *P. puffini* and *C. servana* among the native Leafroller species (Figure 4.5). However similar as in chapter 3 the sample sizes are very small for some species. In case of a species represented by a single specimen, for example *C. servana*, *P. flammea* and *P. octoides*, it will require further sampling to corroborate an exclusivity status in these species even though they have shown exclusivity in the non-speciation genes. In the case of *P. avicenniae* and *C. fraterna* represented by only two alleles, the probability of monophyly by chance through random branching is under the significance level of $10^{-4}$ according to the statistical tests of Rosenberg (Table 7) in Rosenberg (2007)). Therefore it is very likely that the observed monophylies are reflecting the current status of these recently diverged Leafrollers for the PBP1 gene. This analysis also reveals that PBP1 is not reciprocal monophyletic for *P. octo*, *P. excessana*, *C. obliquana* and *C. herana* thus cannot resolve the non-exclusive gene patterns observed in the non-speciation genes for those species groups.

To conclude from this analysis it has been shown that PBP1 does display exclusivity for the species pairs *C. filicis* and *C. fraterna* and the occurrence of positive selection along the lineage leading to these two species may therefore indicate that PBP1 is a speciation gene for these two species. On the contrary, PBP1 does not exhibit the same pattern for *P. octo* and therefore is not likely to be a speciation gene for this species. PBP1 is under purifying selection, only a very
### Table 4.5. Species monophyly status

<table>
<thead>
<tr>
<th>Species</th>
<th>COI-COII</th>
<th>EF-1α</th>
<th>TPI</th>
<th>PBP</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. herana</em></td>
<td>not-exclusive(4)(^a)</td>
<td>not-exclusive(6)</td>
<td>not-exclusive(6)</td>
<td>not-exclusive(6)</td>
</tr>
<tr>
<td><em>C. obliquana</em></td>
<td>not-exclusive(7)</td>
<td>not-exclusive(11)</td>
<td>not-exclusive(10)</td>
<td>not-exclusive(8)</td>
</tr>
<tr>
<td><em>C. fraterna</em></td>
<td>exclusive(3)</td>
<td>exclusive(6)</td>
<td>exclusive(3)</td>
<td>exclusive(2)</td>
</tr>
<tr>
<td><em>C. filicis</em></td>
<td>not-exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
</tr>
<tr>
<td><em>P. octo</em></td>
<td>not-exclusive(5)</td>
<td>not-exclusive(8)</td>
<td>not-exclusive(7)</td>
<td>not-exclusive(7)</td>
</tr>
<tr>
<td><em>P. excessana</em></td>
<td>not-exclusive(3)</td>
<td>not-exclusive(6)</td>
<td>not-exclusive(3)</td>
<td>not-exclusive(3)</td>
</tr>
<tr>
<td><em>P. avicenniae</em></td>
<td>exclusive(2)</td>
<td>not-exclusive(3)</td>
<td>not-exclusive(3)</td>
<td>exclusive(2)</td>
</tr>
<tr>
<td><em>P. puffini</em></td>
<td>exclusive(3)</td>
<td>exclusive(5)</td>
<td>exclusive(5)</td>
<td>exclusive(3)</td>
</tr>
<tr>
<td><em>P. notophaea</em></td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(3)</td>
<td>exclusive(4)</td>
</tr>
<tr>
<td><em>P. octoides</em></td>
<td>-(1)</td>
<td>-(1)</td>
<td>exclusive(4)</td>
<td>-(1)</td>
</tr>
</tbody>
</table>

| Species exclusive \(\sum\) | 4 | 4 | 5 | 5 |
| Species not-exclusive \(\sum\) | 5 | 5 | 5 | 4 |

\(^a\)Number of alleles are given in brackets.

A small number of sites were predicted to be under positive selection. However, some interesting residues have been identified in *P. octo*, *P. excessana* and *P. avicenniae* and it might be worth to further investigate the causes for these different haplotypes.
The genic view of speciation is characterized by dividing the genome into genes or loci that are differentially adapted and contribute to reproductive isolation as a first step in incipient species and into other loci that are not involved in speciation. These differentially adapted regions are called speciation genes (Wu, 2001) and it is hypothesized that they fulfill certain criteria that have been tested in this thesis. The genic view of speciation carries with it a number of predictions. These are an absence of shared ancient polymorphism and reduced gene flow at these loci reflected in reciprocal monophyletic species trees. Associated with the differential adaptation positive selection is also predicted to be associated with speciation genes, often observed in the form of selective sweeps for speciation and adjacent regions. The genome is predicted to be mosaic, not assuming the whole genome itself is responsible for reproductive isolation but only a few loci are actually involved in instigating a speciation process within species (Ting et al., 2000; Wu, 2001).

In this PhD thesis it was tested whether the pheromone binding protein 1 gene, PBP1 fulfils
any of the criteria above for a speciation gene. That is, does it show reciprocal monophyly for species by comparing gene phylogenies of three non-speciation loci (COI, COII, EF-1α and TPI). This study shows that PBP1 may fulfil some of the criteria for some species, which was reflected in different sorting patterns within gene trees (see Table 4.5). A comparison between PBP1 and non-speciation loci confirmed the mosaic genome view. That is, the phylogenetic patterns observed in the different gene trees are not the same. A preliminary examination of amino acid changes in PBP1 suggests a selective sweep may have occurred or was occurring in the Planotortrix species P. octo, P. excessana and P. avicenniae. A further investigation of whether the amino acid changes among the species are at biochemically relevant positions within the PBP1 protein showed that some may possibly play a role in discriminating different length sex pheromone components used in C. fraterna and C. filicis.

5.1 Sorting of alleles

5.1.1 Taxon sampling and power

The comparison of multiple loci in Chapter 3 showed that different loci have different evolutionary histories. In the case of mitochondrial genes the analysis of the COI ‘barcoding-region’ allows to address an important issues, that of phylogenetic power in tree reconstruction. The COI dataset comprises 200 taxa from all 12 species of native Leafrollers and is therefore more powerful from the perspective of taxon sampling than the longer combined COI-COII region with fewer number of taxa. This comparison addresses the issue of whether the number of taxa are important when it comes to detecting species sorting in the native Leafrollers. Since COI has the bigger dataset in terms of number of individuals but shorter sequence data it should show less taxa being reciprocal monophyletic than the combined COI-COII dataset, followed by the nuclear markers. Table 3.4 in section 3.4 does show this behaviour for the native Leafrollers. It demonstrates that for the mt genes, that it is not the number of taxa that is important but the lengths of the sequence information available. For example in P. avicenniae and C. fraterna both species do not show an exclusive grouping in the COI region alone but do exhibit that pattern when the combined region of COI-COII is used. To conclude, a larger number of individuals is important when addressing questions of sorting in species but even more important is the lengths of sequence information as demonstrated for the mitochondrial genes. As the
number of taxa increases the chance of getting monophyly by chance decreases. The barcoding region is not long enough to show reciprocal monophyly by species and it is necessary to use a longer mt fragment to get more phylogenetic power.

5.1.2 Effective population sizes

A process called lineage sorting is a possible explanation for the observed non-sorting pattern observed for all of the genes for some taxa (see Table 3.4). Lineage sorting describes the disappearance of shared ancestral polymorphic allelic lineages between species over time. Mitochondrial genes are expected to complete lineage sorting faster or earlier than nuclear genes due to smaller effective population sizes (Nichols, 2001). Nonetheless incomplete lineage sorting can affect mitochondrial genes especially in recently radiated taxa where speciation events occurred before sorting was completed as in the example of cichlid fish (Moran & Kornfield, 1993). TPI is hypothesized to be sex-linked in other Lepidoptera and should therefore sort earlier than EF-1α. This is not observed among the native Leafrollers, TPI reflects exclusivity for the same number of species as EF-1α does. A possible explanation is that TPI is not sex-linked in the native Leafrollers and therefore has the same effective population size than EF-1α as an unlinked autosomal marker.

Another factor that is relevant to the issue of non-sorting is the extended time period in which the specimen samples have been acquired and the availability of the actual Leafroller species. Some of the specimens have been collected in the early eighties and some of them just recently a few years ago (see Tables (3.1) and (3.2)). That leaves a time period of over twenty years. The distribution and habitat for some populations for some species also changed reflected in the absence of specimens at previously used collection sites. Additionally some specimens have been kept at the Insect Rearing facility at HortResearch with just a couple of hundred specimens per species, equivalent of creating an artificial bottleneck for those species. Other specimens from the same species that were collected from the field were from a substantially bigger population in size. This difference in population size can account for differences in allele frequencies and ultimately in differences of displaying exclusivity in an allele gene tree.
5.1.3 Hybridization and Introgression

Alleles can be shared between species due to hybridization, a phenomenon also called interspecific gene flow or gene introgression (Lemmon & Kirkpatrick, 2006; Servedio & Noor, 2003). In *Ctenopseustis* one individual in *C. herana* is thought to have hybridized with *C. filicis* in the past (see Chapter 3 section 3.5.1). This conclusion is supported by crossing experiments in native Leafrollers that have demonstrated that these species can form hybrids in the laboratory, even down to the F2 generation with viable offspring (Foster *et al*., 1997). However it is unclear how much the native Leafroller hybridize in nature. The existence of a long range sex pheromone system usually suggests that without the right attractant mating will not occur between species. The possibility of two different species meeting at close proximity without the action of long range pheromones and then mating with each other is certainly low. The existence of distinct species-specific sex-pheromones in Leafroller moths suggests that hybridization is a rare event (see General Introduction, section 1.5.7). However introgression as a lack of concordance between mitochondrial genes and morphologically defines species in New Zealand cockroaches (Chinn & Gemmell, 2004) and incongruence among mitochondrial and nuclear gene phylogenies in New Zealand Cicadas (Buckley *et al*., 2006) has been hypothesized to explain these observed patterns. It is however difficult to distinguish incomplete lineage sorting from gene introgression since both phenomena display similar patterns of non-sorting in gene tree phylogenies. There are different methods to address this problem but no consensus on a general diagnostic on how to distinguish lineage sorting from introgression has been reported yet (Holder *et al*., 2001). Since the native Leafroller are recently speciated (see Chapter 2, section 2.4) as well lineage sorting might not been complete yet, however gene introgression cannot be ruled out as factor for the non-sorting pattern of the observed gene trees.

5.1.4 Taxonomy

A robust taxonomic system and accurate identification is crucial before one can start to infer phylogenetic relationships. However between morphologically homogenous or cryptic species it sometimes happens that the morphological identification is not correct. This can be detected in trees when single or multiple genetic markers are used to infer a species phylogeny. In the native Leafrollers there is one incident within *Ctenopseustis* that points to a possible miss-identification (see chapter 3 section 3.5.1) unfortunately this specimen was destructively sampled and cannot
be used for a taxonomical re-examination. It is very difficult to morphologically resolve some of these species pairs, especially as some of the characters involve ratios of quantitative characters. The overall taxonomy for this group is, however, not contradicted by the genetic DNA sequence information presented here. There are species groups that are consistently monophyletic and even species that are polyphyletic show a consistent pattern in the sense that they are always sorting with the same sister species. If the taxonomy is incorrect we would expect to see a consistent alternative pattern among the native Leafroller.

## 5.2 Mosaic genome

The mosaic genome is a term used by Wu (2001) used in his paper on the ‘genic view of speciation’ to emphasize the separation of the genome into regions that contributed to an incipient speciation process, speciation genes, and regions that are not involved in that process. Speciation is a genome based event whereas reproductive isolation is driven by single or a few gene loci. His idea is contradictory to Mayr’s idea of the genome as a whole ‘being a cohesive unit’ that needs to change for speciation to take place (Mayr, 1963). An appropriate way of testing whether the genome is mosaic is to infer gene phylogenies from different loci. This has been done numerous times in various organisms, for example in *Drosophila* (Machado & Hey, 2003). The inferred phylogenies can be either congruent or incongruent when multiple genetic markers are used. Usually different loci give different gene trees hence incongruence among multiple gene phylogenies is commonly observed. The native Leafrollers display an incongruent gene pattern for the observed loci (see section 3.5) and (section 4.5). For a species phylogeny to be congruent, all examined individuals within this phylogeny, would have to be reproductively isolated from each other as depicted as stage (d) in Figure (1.2). *P. notophaea* and *P. puffini* for example are probably at stage (d) since all genes give a congruent picture of reciprocal monophyly whereas *C. herana* and *C. obliquana* are probably at stage (b) or (c) in Figure (1.2). Suggesting that the speciation events leading to *P. notophaea* and *P. puffini* are older than for *C. herana* and *C. obliquana*. Using the analogy of a mosaic, if every locus in the genome is reflected by a mosaic ‘piece’ that has a different colour we end up having a multicoloured picture with areas that are similar in colour and property. On a close up we can see all the different pieces and their characteristics, when we step back the whole picture becomes visible as a complete mosaic (the whole genome in this analogy). This ‘mosaic’ is the essence of a species, the challenge is now
to find the 'pieces' that best resembles it, therefore finding genetic loci or speciation genes that defines a species is not an easy task.

5.2.1 Species tree

A common way of deducing a species tree from multiple gene trees is through the creation of a majority-rule consensus tree (Huelsenbeck et al., 1996). A majority consensus tree counts the number of times a species pair is observed in a set of gene trees. Another way of inferring a species phylogeny based on genetic data is from generating a supermatrix (Driskell et al., 2004). A supermatrix includes all available genetic sequences obtained from separate loci in a single concatenated alignment (Bapteste et al., 2002; Rokas et al., 2005). This practice was very popular in the 1990s where the phrase 'total evidence' originated and was believed to increase the power of inferring species phylogenies (Farias et al., 2000). However more recent studies have shown that this practice can be misleading in inferring a species history. A variety of other methods are available to estimate species trees such as gene tree parsimony (Takahata, 1989) and supertree methods (Bininda-Emonds, 2005). The gene tree parsimony method with a Bayesian posterior probability extension was used to reconstruct the species tree of recently radiated New Zealand alpine cicadas (Buckley et al., 2006). Recently another Bayesian based approach for jointly estimating the distribution of gene trees and species trees using a prior has been proposed by Edwards et al. (2007). This method is called Bayesian Estimation of Species Trees (BEST) and provides a mechanism where multiple gene trees can be reconciled into a single species history (Edwards et al.; 2007). A Bayesian adapted gene tree parsimony approach or the BEST method might be suitable to investigate incongruence patterns among gene phylogenies in native New Zealand Leafroller. New Zealand cicadas represent a similar model system to the Leafrollers, having recently speciated and their usage of species-specific courtship songs, which are similar to the species-specific sex-pheromone system in moths (Buckley & Simon, 2007).

5.3 Selection

In the genic view speciation loci are predicted to be under positive selection or become linked to loci that are differentially adapted and therefore likely to be subject to positive selection. Selection is often observed in forms of selective sweeps at speciation loci. Analyses to test for positive selection were conducted in PAML but did not show positive selection for either COI-
COII, EF-1α, TPI or PBP1. This is not surprising for genes that are under purifying selection due to high functional constraints like COI-COII, EF-1α and TPI. However, PBP1 showed some interesting amino acid changes within *P. octo*, *P. excessana* and *P. avicenniae* that might be indicative of an ancient selective sweep. An excess of non-synonymous substitutions within *P. octo*, *P. excessana* and *P. avicenniae* is interesting and a more extensive sampling would allow to do more tests and may confirm these observed results.

5.4 Functional work and future directions

This work has revealed interesting aspects of PBP1 as a speciation gene and also gives insights into areas that are worth pursuing further and areas that might be beneficial when it comes to study genes involved in mate recognition and potentially speciation.

There are a lot of areas that can be further investigated in PBP. One aspect would be to increase the sample size of specimen to look at sorting pattern in gene trees and selection again. Is the ratio of four non-synonymous to two synonymous substitutions in *P. octo* the same when more taxa are included? And if so, enable a better test whether PBP1 has undergone a recent selective sweep (see section 4.4.2). As a comparison, COI-COII, EF-1α and TPI could be used to see whether their nucleotide data sets indicate recent selective pressures or not. Populational sampling, including species from allopatric versus sympatric distributions would be interesting to look at because it might indicate whether allelic versions of one PBP1 protein or more than one PBP1 gene are necessary to distinguish between the different pheromone blends in the native Leafroller. Examples for both are found in *O. nubilialis* (Willett & Harrison, 1999a) having different allelic versions of one PBP or in *A. polyphemus* with the existence of three ApolPBPs (Maida *et al.*, 2000). Also extrapolating from *E. postvittana*, is it likely that the native Leafroller have three PBP genes?

This thesis used the ‘genic view of speciation’ and one of its predictions about reciprocal monophyly to test whether PBP1 is a speciation gene in the native Leafrollers. It was hypothesized that PBP1 is a speciation gene if it behaves like a speciation gene, as in the case of OsdH (Ting *et al.*, 2001). This argument itself has an aspect of circularity though. To circumvent it functional work needs to be to conducted.

For example in doing binding essays to show that PBP is actually capable of binding a 14- or 16-carbon acetate in the native Leafroller. This was shown in a study by Newcomb *et al.*
(2002) on the closely related Leafroller moths *Epiphyas postvittana* where *PBP1* was found to be capable of binding the major component of the species sex pheromone blend.

However if *PBP1* is only a speciation gene for some taxa, other candidate speciation genes are available to study in this system. Focusing on pheromone reception in the male moths, pheromone receptors are also likely candidates for speciation genes. As they specifically bind the *PBP* and the pheromone bound to it they are likely to be under selection to maintain this specificity. While no sex pheromone receptors are currently available for the native Leafrollers it may be just a matter of time before pheromone receptors are characterized and available to study in this system.

Looking for a candidate speciation gene involved in sex pheromone production has revealed some promising recent results. Isolation of desaturase genes has revealed the differential regulation of their gene expression (Sirey, 2006). Investigations are currently being done in isolating the promoter regions of desaurase genes in five native Leafroller species to investigate how desaturase genes are differentially expressed. Desaturases are responsible for placing double-bonds along the hydrocarbon backbone of a pheromone through desaturation at specific positions (see general introduction, section 1.5.5). In *P. octo* a Δ 10 desaturase has been identified that produces a Z8-14:OAc sex pheromone in female moths (Hao et al., 2002). *P. octo* mostly uses Z8-14:OAc singly. Functional work could reveal if males would still be attracted to females after this Δ10 desaturase gene has been knocked-out. To test this, I used a technique called RNA interference (RNAi) or gene silencing by injecting dsRNA into pupae of *P. octo* females. This was conducted to remove transcripts of the gene. After the adults emerged the pheromone glands were removed to extract total RNA. RNA was transcribed into cDNA and used in a quantative Realtime PCR to assess levels of desaturase gene expression. The comparison with controls, however, showed only a reduction in desaturase expression for a few female specimen but was not significant compared to the control specimen. A Gas Cromatograph analyses of pheromone glands from gene silenced females was conducted to see whether the pheromone titer of Z8-14:OAc is significantly lower compared to non-silenced females of *P. octo*. Again it showed that some females had a reduction in produced Z8-14:OAc but these results were not consistent in all female moths. Finally to analysis whether male moths are still attracted to silenced female moths, a behavioural experiment in a windtunnel had been conducted. Silenced females were placed into a windtunnel until they displayed calling behaviour to attract male moths.
If they did show this behaviour male moths were released into the windtunnel and recorded whether they are attracted by the female pheromones and fly upwind to her. This behaviour was displayed by some male moths but again was not significant between a silenced versus a non-silenced female. These experiments are only preliminary results and have to be repeated with more samples to see whether the observed patterns are consistent among females in *P. octo*. It shows that the native New Zealand Leafroller are an exciting system to further investigate genes involved in mate recognition. With the help of different techniques such as quantitative realtime PCR, RNAi and an insect-cell based functional assay system, further studies in the area of pheromone binding, receptor interactions and desaturase regulation will be accessible. These studies will hopefully help to elucidate the process of speciation through its underlying genes as an important aspect of evolution.
Appendix A

Specimen details for the COI gene dataset in Chapter 2
List of voucher specimen used for Chapter 2
Paper on DNA barcoding of the endemic New Zealand leafroller moth genera, *Ctenopseustis* and *Planotortrix*
DNA BARCODING

DNA barcoding of the endemic New Zealand leafroller moth genera, *Ctenopseustis* and *Planotortrix*

PIA LANGHOFF,*§ ASTRID AUTHIER,* THOMAS R. BUCKLEY,+ JOHN S. DUGDALE,+ ALLEN RODRIGOS and RICHARD D. NEWCOMB*

*Research Group, HortResearch, Private Bag 92169, Auckland 1003, New Zealand, †Landcare Research, Private Bag 92170, Auckland 1142, New Zealand, ‡Landcare Research, Private Bag 6, Nelson 7042, New Zealand, §Bioinformatics Institute, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

Abstract

Molecular techniques such as DNA barcoding have become popular in assisting species identification especially for cryptic species complexes. We have analysed data from a 468-bp region of the mitochondrial cytochrome oxidase subunit I (COI) gene from 200 specimens of 12 species of endemic New Zealand leafroller moths (Tortricidae) from the genera *Planotortrix* and *Ctenopseustis* to assess whether the DNA barcoding region can distinguish these species. Among the 200 sequences analysed, 72 haplotypes were recovered, with each genus forming a separate major clade. Maximum likelihood phylogenetic methods were used to test whether species fell into reciprocally monophyletic clades. The optimal phylogeny showed that four species within the genus *Ctenopseustis* (*C. obliqua*, *C. herana*, *C. filicis* and *C. fraterna*) and three within *Planotortrix* (*P. octo*, *P. excessana* and *P. avicenniae*) are polyphyletic. Shimodaira–Hasegawa tests rejected a null hypothesis of monophyly for the species *C. obliquana*, *C. herana*, *P. octo* and *P. excessana*. Comparisons of within and between species levels of sequence divergence for the same set of seven species showed cases where maximum levels of within-species divergence were greater than some levels of between-species divergence. DNA barcoding using this region of the COI gene is able to distinguish the two genera and some species within each genus; however, many species cannot be identified using this method. Finally, we discuss the possible reasons for this polyphyly, including incomplete lineage sorting, introgression, horizontal gene transfer and incorrect taxonomy.

Keywords: *Ctenopseustis*, DNA barcoding, Lepidoptera, *Planotortrix*, Tortricidae

Received 26 September 2008; revision accepted 3 December 2008

Introduction

Lepidoptera is a diverse order, containing over 100 000 species of moths alone (Roelofs & Rooney 2003). Common to many lepidopteran species is the use of a long-distance chemical communication system based on sex pheromones to attract mates (Byers 2006). Lepidopterans are often used as model systems for a wide range of evolutionary studies such as speciation (Beltrán et al. 2002; Sperling 2003). Speciation within families and genera can be rapid, with morphologically identical (cryptic) species complexes often observed, containing species that use distinct sex pheromone blends (Smadja & Butlin 2008).

The endemic New Zealand leafroller genera, *Ctenopseustis* (Meyrick) (brown-headed leafroller moth) and *Planotortrix* (Dugdale) (green-headed leafroller moth) (Tortricidae: Lepidoptera) are geographically widespread. Some species are specialized herbivores and feed on specific host plants. For example, *Planotortrix avicenniae* Dugdale feeds on mangrove (*Avicennia marina*), and *Ctenopseustis fraterna* (Walker) on silver fern (*Cyathea dealbata*). Other species are polyphagous, feeding on a wide range of plants and important horticultural crops such as apple, kiwifruit, avocados, strawberries and blueberries, as well as native plants such as Pteridophyta, *Coprosma* (Rubiaceae), various *Hebe* species, *Senecio reinoldii*, various *Olearia* species and native gymnosperms

Correspondence: Richard D. Newcomb, Fax: xxxxxxxxxxxxxx; E-mail: newcomb@hortresearch.co.nz

© 2009 The Authors
Journal compilation © 2009 Blackwell Publishing Ltd
2 DNA BARCODING

(Dugdale 1990). Before 1984, each genus was believed to contain only a single species, *Ctenopseustis obligna* Walker and *Planotortrix excessana* Walker. More recently, cryptic sibling species were discovered within each species based on studies of their sex pheromone composition (Foster & Dugdale 1988). The two genera were then revised by Dugdale (1990) based on morphological characters that corroborated the sex pheromone differences resulting in five species of *Ctenopseustis* (*C. obligna*, *C. herana*, *C. filicis*, *C. fratera* and *C. servana*) and seven species of *Planotortrix* (*P. octo*, *P. octoides*, *P. excessana*, *P. avicenniae*, *P. puffini*, *P. notophaeas*).

Details of the geographical distribution of these species within New Zealand are given in Dugdale (1990).

Leafroller moths of the genera *Planotortrix* and *Ctenopseustis* use pheromone blends of mono-unsaturated tetradecenyl (14:OAc) or hexadecenyl (16:OAc) acetates, containing double-bond positions at Δ5, Δ7, Δ8 and Δ10. The resulting sex pheromones consist of either single components or blends of Z5-14:OAc and Z8-14:OAc or Z5-14:OAc and Z7-14:OAc or Z10-16:OAc (Foster & Dugdale 1988).

Phylogenetic relationships for the genera *Ctenopseustis* and *Planotortrix* have been estimated previously based on morphological, pheromonal and molecular characteristics (Sin et al. 1995; Dugdale 1997; Newcomb & Gleeson 1998; Gleeson et al. 2000). These studies have shown that molecular and morphological markers are able to separate *Ctenopseustis* and *Planotortrix* into well-supported monophyletic groups, with some groups supported within genera (e.g. *P. flammea*, *P. notophaeas* and *P. puffini*) and that some species are likely sister to the rest of the species within the genera (e.g. *C. servana*). Furthermore, these studies have revealed that the use of certain pheromone blends has evolved more than once within the two genera (e.g. Z5-14:OAc).

Recently, species identification using short mitochondrial DNA fragments known as DNA barcodes, is becoming increasingly popular to diagnose animal taxa and assess biodiversity (Blaxter 2003; Marshall 2005). DNA barcoding has been used successfully to identify new Lepidoptera species either through barcoding data alone (Hebert 2003; Hajibabaei et al. 2006) or in combination with morphological and ecological studies (Hebert et al. 2004; Hulcr et al. 2007).

Here we analyse an extensive sample of 200 specimens of New Zealand leafroller moths from the genera *Ctenopseustis* and *Planotortrix* for the barcoding region within mitochondrial cytochrome oxidase I gene (COI) to determine whether this marker can be used to identify species from the two genera.

Materials and methods

Specimen sampling

Moth specimens from the genera *Ctenopseustis* and *Planotortrix* were supplied either by the insect-rearing facility at HortResearch, collected using pheromone traps or collected as larvae, reared on artificial diet (Singh 1974) to adults and identified morphologically by John S. Dugdale or Robert Hoare (Landcare Research) (see Fig. 1 for location details). A table of specimens and collection localities is available online (Table S1, Supporting information). Voucher specimens from each species included in this study have been accessioned into the New Zealand Arthropod Collection, Landcare Research and HortResearch (Table S2, Supporting information).

Polymerase chain reaction amplification and sequencing

Genomic DNA was extracted from whole moth specimens using the DNeasy Tissue Kit (QIAGEN), following the manufacturer’s instructions. A 468 bp fragment, the ‘barcoding’ region, of the COI gene was amplified by polymerase chain reaction (PCR) using the primers CI-J-1718 (5'-GGAGGATT TGGAATTGATTAGTTCC-3') and CI-N-2191 (5'-CCCCG TAAAATTTATATAACCTTC-3') (Simon et al. 1994). This region encodes the third, fourth, fifth and sixth of 12 trans-membrane helices, the first and the second internal and the second and third external loops of the COI protein (Gleeson et al. 2000). Amplification reactions (50 μL) consisted of 10 pmol of each primer, 20 mM Tris-HCl (pH 8.4), 1.5 mM MgCl2, 50 mM KCl, 0.2 mM of each dNTP and 2 U of Taq polymerase (Invitrogen). The cycling conditions consisted of an initial denaturation at 94 °C for 2 min, followed by 30 s at 94 °C, annealing at 40 °C for 30 s, and extension at 72 °C for 1 min for 35 cycles. A final extension step at 72 °C for 10 min was performed to ensure PCR products were full length.

The resulting PCR products were purified using the High Pure PCR Product Purification Kit (Roche), following the manufacturer’s instructions. Direct sequencing of purified products using the PCR primers was undertaken using the BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems), according to the manufacturer’s instructions. Cycle sequencing products were resolved on an Applied Biosystems 3100 automated DNA sequencer.

Data analysis

Sequence data were analysed using Sequencher software (Gene Codes Corporation). Because the COI gene contained no insertions or deletions, alignment algorithms were unnecessary. We used PAUP* 4.0b10 (Swoford 2003) and ModelTest version 3.7 (Posada & Crandall 1998) to calculate Akaike information criterion (AIC) values for 56 different nucleotide substitution models and to select the best-fit model. Maximum likelihood analyses and constrained topology analyses including Shimodaira–Hasegawa (SH) tests (Shimodaira & Hasegawa 1999) were performed using PAUP* 4.0b10 under the best-fit model. Tree searches were
performed using tree-bisection–reconnection (TBR) branch swapping starting from a single random addition tree. We also calculated within- and among-species genetic distances using \( p \)-distances.

**Results**

We obtained 468 bp of COI DNA sequence from 187 individuals representing all species of *Ctenopseustis* and *Planotortrix*. These data were combined with the COI sequences from Newcomb & Gleeson (1998) to produce a final alignment of 200 individuals. Sequences have been deposited at GenBank under accession nos FJ225474–FJ225660. The alignment contained 72 unique haplotypes, 102 variable sites and 80 parsimony-informative sites. The GTR + I + \( \Gamma \) model was estimated to be the best-fit model under the AIC and the maximum likelihood tree estimated under this model is shown in Fig. 2. When midpoint rooted, the tree shows the two genera *Ctenopseustis* and *Planotortrix* are reciprocally monophyletic. The single sample of the

![Fig. 1 Location information of samples of *Ctenopseustis* and *Planotortrix* species used in this study from the North Island (a) and South Island (b) of New Zealand.](image)
Fig. 1 Continued
coastal species Ctenopseustis servana is sister to the remaining Ctenopseustis species. None of the other four Ctenopseustis species are each monophyletic, including the two fern feeders, C. fraterna and C. filicis, and the two polyphagous species C. obliquana and C. herana. Within the genus Planotortrix, four species are each monophyletic, P. puffini, P. notophaea, P. flammea, P. octoides, whereas P. excessana, P. octo and the mangrove species P. avicennia are all polyphyletic. Using other tree-building algorithms such as parsimony, distance and Bayesian approaches also yielded similar patterns of monophyly and polyphyly among the species of the two genera (data not shown).

We then tested more formally whether these data are able to reject monophyly of the apparently polyphyletic species. To test this, we performed constrained searches such that each species which is not monophyletic in the unconstrained tree shown in Fig. 2 was forced to be monophyletic. Due to the large computational burden of an unconstrained TBR search, we used a TBR limit of four to build constrained trees. We then compared the likelihood scores of these constrained
trees to the unconstrained tree to examine whether they differed significantly using the SH test (Table 1). Both the topology and likelihood scores for the unconstrained maximum likelihood tree (−lnL 1803.92206) and the unconstrained maximum likelihood tree set with a reconnection limit of 4 (−lnL 1806.39639) were very similar, justifying the use of the lower reconnection limit to construct the constrained trees. The SH test rejected the null hypothesis of monophyly for the sibling species pairs of C. herana and C. obliquana, and P. octo and P. excessana, but failed to reject the null hypothesis of monophyly for P. avicennia, C. filics and C. fraterna even though they are not monophyletic in the unconstrained tree (Table 1).

We examined levels of within-species divergence by comparing them to levels of between-species divergence in the COI data set (Fig. 3) by calculating p-distances. We found examples of species comparisons where levels of within-species diversity exceeded between-species diversity (points below the dotted line in Fig. 3). In Planotortrix, this included comparisons among P. octo, P. excessana and P. avicenniae and in Ctenopseustis included comparisons among C. obliquana, C. herana, C. fraterna and C. filics. Levels of within-species diversity were especially high in C. obliquana, C. herana and C. fraterna, with maximum levels of within-species diversity exceeding 3%.

Discussion

DNA barcoding has been proposed as a rapid species identification tool yet significant controversy exists as to its utility (e.g. Hebert et al. 2003; Will & Rubinoff 2004; DeSalle et al. 2005; Meyer & Paulay 2005). To fully assess the utility of DNA barcoding, analysis of species with a well-established taxonomic basis is required. The New Zealand tortricid genera Ctenopseustis and Planotortrix are ideal in this respect because the species boundaries are supported by morphological and behavioural data through the analysis of species-specific pheromones (Foster & Dugdale 1988; Dugdale 1990).

Our molecular phylogenetic study of Ctenopseustis and Planotortrix shows that two genera are monophyletic with respect to one another in agreement with the study of Newcomb & Gleeson (1998) that sampled the same two genera but with fewer individuals from each species. However, only four of the 11 species, in which multiple individuals were sampled, form monophyletic groupings in our optimal tree. For four species that are not monophyletic, the SH tests suggest the data are significantly in favour of this.

There are many possible reasons for the lack of species monophyly in the COI data set for Ctenopseustis and Planotortrix. Biological explanations include incomplete lineage sorting, introgression and horizontal gene transfer (Funk et al. 2003). We believe that the most likely reason is that the species complexes have been derived recently and that the COI haplotypes are still in the process of sorting. Certainly, it seems that the two genera arose recently: with a sequence...
divergence of approximately 10% and using a rate derived from Brower (1994), the estimated age of the most recent common ancestor between \textit{Ctenopseustis} and \textit{Planotortrix} is approximately 5 million years. However, the age of the unresolved species complexes with each genus are much more recent, with estimates of less than 500,000 years for the most recent common ancestor of \textit{Ctenopseustis obliquana}, \textit{C. herana}, \textit{C. fraterna} and \textit{C. filicis}, and even more recently for the most recent common ancestor of \textit{Planotortrix octo}, \textit{P. octoïdes}, \textit{P. avicenia} and \textit{P. excessana}. For several of the species, comparisons of levels of within-species divergence are higher than the minimum levels of between-species divergence (Fig. 3). When within-species divergence is as high or higher than between-species divergence, there is no ‘gap’ to enable barcoding to distinguish some of the species from one another. This phenomenon of overlapping barcodes has also been observed in the Lepidoptera family Hesperiidae (Hajibabaei et al. 2006).

Introgression is another possible reason that can lead to nonreciprocal monophyletic species trees (Funk & Omland 2003). Introgression events are not uncommon in New Zealand insects, having been observed in cicadas (Buckley et al. 2006) and cockroaches (Chinn & Gemmell 2004). If introgression events are recent then sympatric species are more likely to share haplotypes than allopatric populations of the two species. In contrast shared ancient polymorphism (nonsorting) is less likely to show an effect due to shared distributions (Funk & Omland 2003). Within \textit{Ctenopseustis}, for example, if introgression was responsible for the observed patterns, we would expect \textit{C. obliquana} in the South Island of New Zealand to be polyphyletic with \textit{C. herana} and \textit{C. filicis} with which it is sympatric, whereas \textit{C. obliquana} in the North Island should be polyphyletic with \textit{C. fraterna}, which is only found in the North Island. Since this appears not to be the case, incomplete lineage sorting is more likely to explain the observed polyphyly. It is also possible that both introgression and incomplete lineage sorting may be acting at the same time and sampling of nuclear genes is required to differentiate between the two processes (Buckley et al. 2006). Another formal possibility is that genetic material has been transferred horizontally among species by a pathogen or symbiont, such as \textit{Wolbachia}, a symbiont known to inhabit some members of the Lepidoptera.

There are also nonbiological reasons why we may observe many nonmonophyletic species groupings in \textit{Ctenopseustis} and \textit{Planotortrix}. Some possible explanations are that errors have been made in the identification of individual specimens or that the taxonomy of \textit{Ctenopseustis} and \textit{Planotortrix} is incorrect. Certainly some of the sibling species pairs are difficult to distinguish morphologically as adults and also as larvae. However, a majority of our samples were identified based on the current species descriptions by those involved in the most recent taxonomic revision (Dugdale 1990). Of the 200 specimens sequenced for the COI gene, 117 were identified using forewing costal fold ratios for the male specimen and reproductive structures, for example, cestum length ratios for the female specimen, while the remainder were identified as larvae, through host plant affinity or through trapping using species-specific sex pheromone lures. Another possibility is that the taxonomy is incorrect. However, even if this is so, analysis of the COI data set does not provide a better hypothesis for the taxonomy of the group and seemingly, good monophyletic groups in the phylogenetic tree do represent already described species (e.g. \textit{P. notophaca} and \textit{P. puffini}).

Although the two groups of unresolved species in \textit{Ctenopseustis} and \textit{Planotortrix} in the COI gene tree can be difficult to distinguish morphologically, they use distinct blends of sex pheromones (Newcomb & Gleeson 1998). Moreover, changes in sex pheromone biochemistry may be driving the speciation process in these moths. Therefore perhaps investigating the patterns of variation in genes encoding elements of the pheromone production system in females and reception in males may provide better markers for species identification than mitochondrial DNA.

Acknowledgements

We wish to thank Anne Barrington for providing specimens from the HortResearch rearing facility, Robert Hoare for assistance in collection and identification, Leonie Clunie for data-basing, Clinton Turner for helpful comments on the manuscript, and Minna Pesonen for figure preparation. This work was supported by the Royal Society of New Zealand, Marsden Fund (HRT203) and the Foundation for Research, Science and Technology (contract number C09X0501).

References


Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 Location information including species name, sample ID number, collection location, lat/long, and GenBank Accession number of samples used in this study

Table S2 Vouchered material from some of the moths used in this study held at HortResearch (a) and representative voucher specimens from each species held within the New Zealand Arthropod Collection, Landcare Research (b)

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
<table>
<thead>
<tr>
<th>Species ID</th>
<th>Species name</th>
<th>Locality</th>
<th>GPS</th>
<th>Genbank Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_fil_1024</td>
<td>Ctenopseustis filicis</td>
<td>Raroa Track, SI</td>
<td>-468945 1680750</td>
<td>FJ225529</td>
</tr>
<tr>
<td>C_fil_1045</td>
<td>Ctenopseustis filicis</td>
<td>Moonlight Ecological Reserve, BR</td>
<td>-4652 16807</td>
<td>FJ225530</td>
</tr>
<tr>
<td>C_fil_487</td>
<td>Ctenopseustis filicis</td>
<td>Lee Bay, SI</td>
<td>-4652 16807</td>
<td>FJ225531</td>
</tr>
<tr>
<td>C_fil_489</td>
<td>Ctenopseustis filicis</td>
<td>Lee Bay, SI</td>
<td>-4652 16807</td>
<td>FJ225532</td>
</tr>
<tr>
<td>C_fil_490</td>
<td>Ctenopseustis filicis</td>
<td>Lee Bay, SI</td>
<td>-4652 16807</td>
<td>FJ225533</td>
</tr>
<tr>
<td>C_fil_491</td>
<td>Ctenopseustis filicis</td>
<td>Lee Bay, SI</td>
<td>-4652 16807</td>
<td>FJ225534</td>
</tr>
<tr>
<td>C_fil_640</td>
<td>Ctenopseustis filicis</td>
<td>Mt. Maude, D’Urville Island, SD</td>
<td>-408567 1738168</td>
<td>FJ225535</td>
</tr>
<tr>
<td>C_fil_genbank</td>
<td>Ctenopseustis filicis</td>
<td>Oban, SI</td>
<td>-4655 16810</td>
<td>AF016466</td>
</tr>
<tr>
<td>C_fra_1016</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225536</td>
</tr>
<tr>
<td>C_fra_1017</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225537</td>
</tr>
<tr>
<td>C_fra_998</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225538</td>
</tr>
<tr>
<td>C_fra_999</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225539</td>
</tr>
<tr>
<td>C_frat_496</td>
<td>Ctenopseustis fraterna</td>
<td>Whakapapa, TO</td>
<td>-3912 17612</td>
<td>FJ225540</td>
</tr>
<tr>
<td>C_frat_732</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225541</td>
</tr>
<tr>
<td>C_frat_733</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225542</td>
</tr>
<tr>
<td>C_frat_734</td>
<td>Ctenopseustis fraterna</td>
<td>Titirangi, AK</td>
<td>-369414 1746539</td>
<td>FJ225543</td>
</tr>
<tr>
<td>C_frat_genbank</td>
<td>Ctenopseustis fraterna</td>
<td>Wenderholme, AK</td>
<td>-365245 1747200</td>
<td>AF016467</td>
</tr>
<tr>
<td>C_her_041</td>
<td>Ctenopseustis herana</td>
<td>Takaka Hill, NN</td>
<td>-4102 17249</td>
<td>FJ225544</td>
</tr>
<tr>
<td>C_her_043</td>
<td>Ctenopseustis herana</td>
<td>Takaka Hill, NN</td>
<td>-4102 17249</td>
<td>FJ225545</td>
</tr>
<tr>
<td>C_her_088</td>
<td>Ctenopseustis herana</td>
<td>Takaka Hill, NN</td>
<td>-4102 17249</td>
<td>FJ225546</td>
</tr>
<tr>
<td>C_her_093</td>
<td>Ctenopseustis herana</td>
<td>Takaka Hill, NN</td>
<td>-4102 17249</td>
<td>FJ225547</td>
</tr>
<tr>
<td>C_her_1006</td>
<td>Ctenopseustis herana</td>
<td>Bluff, SL</td>
<td>-465749 1633174</td>
<td>FJ225548</td>
</tr>
<tr>
<td>C_her_1007</td>
<td>Ctenopseustis herana</td>
<td>Bluff, SL</td>
<td>-465749 1633174</td>
<td>FJ225549</td>
</tr>
<tr>
<td>C_her_1046</td>
<td>Ctenopseustis herana</td>
<td>Coll Creek, Perseverance Road, BR</td>
<td>-419934 1718004</td>
<td>FJ225550</td>
</tr>
<tr>
<td>C_her_1047</td>
<td>Ctenopseustis herana</td>
<td>Brook Reservoir, NN</td>
<td>-413459 1733601</td>
<td>FJ225551</td>
</tr>
<tr>
<td>C_her_146</td>
<td>Ctenopseustis herana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123 17351</td>
<td>FJ225552</td>
</tr>
<tr>
<td>C_her_153</td>
<td>Ctenopseustis herana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123 17351</td>
<td>FJ225553</td>
</tr>
<tr>
<td>C_her_157</td>
<td>Ctenopseustis herana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123 17351</td>
<td>FJ225554</td>
</tr>
<tr>
<td>C_her_158</td>
<td>Ctenopseustis herana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123 17351</td>
<td>FJ225555</td>
</tr>
<tr>
<td>C_her_168</td>
<td>Ctenopseustis herana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td>-4118 17334</td>
<td>FJ225556</td>
</tr>
<tr>
<td>C_her_186</td>
<td>Ctenopseustis herana</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225557</td>
</tr>
<tr>
<td>Code</td>
<td>Species</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------</td>
<td>---------------------------------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>C_her_191</td>
<td><em>Ctenopseustis herana</em></td>
<td>Riwaka Resurgence, NN</td>
<td>-410206</td>
<td>1725401</td>
</tr>
<tr>
<td>C_her_216</td>
<td><em>Ctenopseustis herana</em></td>
<td>Snowden's Bush, Brightwater, NN</td>
<td>-412216</td>
<td>1730612</td>
</tr>
<tr>
<td>C_her_219</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tobin's Bridge, Rahu, BR</td>
<td>-421736</td>
<td>1720350</td>
</tr>
<tr>
<td>C_her_220</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tobin's Bridge, Rahu, BR</td>
<td>-421736</td>
<td>1720350</td>
</tr>
<tr>
<td>C_her_222</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tobin's Bridge, Rahu, BR</td>
<td>-421736</td>
<td>1720350</td>
</tr>
<tr>
<td>C_her_226</td>
<td><em>Ctenopseustis herana</em></td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255</td>
<td>1725532</td>
</tr>
<tr>
<td>C_her_274</td>
<td><em>Ctenopseustis herana</em></td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255</td>
<td>1725532</td>
</tr>
<tr>
<td>C_her_520</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_521</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_522</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_523</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_524</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_525</td>
<td><em>Ctenopseustis herana</em></td>
<td>Tophouse Reserve, Buller, BR</td>
<td>-4143</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_528</td>
<td><em>Ctenopseustis herana</em></td>
<td>St. Arnaud, MB</td>
<td>-4145</td>
<td>17250</td>
</tr>
<tr>
<td>C_her_533</td>
<td><em>Ctenopseustis herana</em></td>
<td>Lee Bay, SI</td>
<td>-4652</td>
<td>16807</td>
</tr>
<tr>
<td>C_her_534</td>
<td><em>Ctenopseustis herana</em></td>
<td>Horseshoe Bay, SI</td>
<td>-4652</td>
<td>16808</td>
</tr>
<tr>
<td>C_her_552</td>
<td><em>Ctenopseustis herana</em></td>
<td>The Ave Res., Aitkens, AK</td>
<td>-368677</td>
<td>1747534</td>
</tr>
<tr>
<td>C_her_756</td>
<td><em>Ctenopseustis herana</em></td>
<td>Eglington Valley, FD</td>
<td>-4509</td>
<td>16755</td>
</tr>
<tr>
<td>C_her_792</td>
<td><em>Ctenopseustis herana</em></td>
<td>St. Arnaud, Mainland Island, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_her_793</td>
<td><em>Ctenopseustis herana</em></td>
<td>St. Arnaud, Mainland Island, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_her_822</td>
<td><em>Ctenopseustis herana</em></td>
<td>Stoke; abandoned orchard, NN</td>
<td>-411922</td>
<td>1731322</td>
</tr>
<tr>
<td>C_her_823</td>
<td><em>Ctenopseustis herana</em></td>
<td>Stoke; abandoned orchard, NN</td>
<td>-411922</td>
<td>1731322</td>
</tr>
<tr>
<td>C_her_824</td>
<td><em>Ctenopseustis herana</em></td>
<td>Stoke; abandoned orchard, NN</td>
<td>-411922</td>
<td>1731322</td>
</tr>
<tr>
<td>C_her_826</td>
<td><em>Ctenopseustis herana</em></td>
<td>Stoke; abandoned orchard, NN</td>
<td>-411922</td>
<td>1731322</td>
</tr>
<tr>
<td>C_her_838</td>
<td><em>Ctenopseustis herana</em></td>
<td>Riwaka Resurgence, NN</td>
<td>-410206</td>
<td>1725401</td>
</tr>
<tr>
<td>C_her_859</td>
<td><em>Ctenopseustis herana</em></td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255</td>
<td>1725532</td>
</tr>
<tr>
<td>C_her_866</td>
<td><em>Ctenopseustis herana</em></td>
<td>Lake Rotoroa, BR</td>
<td>-414749</td>
<td>1723545</td>
</tr>
<tr>
<td>C_her_867</td>
<td><em>Ctenopseustis herana</em></td>
<td>Lake Rotoroa, BR</td>
<td>-414749</td>
<td>1723545</td>
</tr>
<tr>
<td>C_her_880</td>
<td><em>Ctenopseustis herana</em></td>
<td>Lake Rotoroa, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_her_882</td>
<td><em>Ctenopseustis herana</em></td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255</td>
<td>1725532</td>
</tr>
<tr>
<td>C_her_891</td>
<td><em>Ctenopseustis herana</em></td>
<td>Ohau Point seal colony, KA</td>
<td>-422483</td>
<td>1738321</td>
</tr>
<tr>
<td>C_her_938</td>
<td><em>Ctenopseustis herana</em></td>
<td>Lake Rotoroa, BR</td>
<td>-414749</td>
<td>1723545</td>
</tr>
<tr>
<td>Code</td>
<td>Species</td>
<td>Location</td>
<td>Coordinates</td>
<td>GenBank Accession</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------</td>
<td>-------------------------------</td>
<td>----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>C_her_939</td>
<td>Ctenopseustis herana</td>
<td>Lake Rotoroa, BR</td>
<td>-414749 1723545</td>
<td>FJ225604</td>
</tr>
<tr>
<td>C_her_941</td>
<td>Ctenopseustis herana</td>
<td>Lake Rotoroa, BR</td>
<td>-414749 1723545</td>
<td>FJ225605</td>
</tr>
<tr>
<td>C_her_952</td>
<td>Ctenopseustis herana</td>
<td>Lake Rotoroa, BR</td>
<td>-414749 1723545</td>
<td>FJ225606</td>
</tr>
<tr>
<td>C_her_954</td>
<td>Ctenopseustis herana</td>
<td>Lake Rotoroa, BR</td>
<td>-414749 1723545</td>
<td>FJ225607</td>
</tr>
<tr>
<td>C_her_995</td>
<td>Ctenopseustis herana</td>
<td>Bluff, SL</td>
<td>-465749 1633174</td>
<td>FJ225608</td>
</tr>
<tr>
<td>C_her_genbank</td>
<td>Ctenopseustis herana</td>
<td>Tai Tapu, Lincoln, MC</td>
<td>-436614 1725492</td>
<td>AF016488</td>
</tr>
<tr>
<td>C_obl_005</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225566</td>
</tr>
<tr>
<td>C_obl_007</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225567</td>
</tr>
<tr>
<td>C_obl_008</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225568</td>
</tr>
<tr>
<td>C_obl_009</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225569</td>
</tr>
<tr>
<td>C_obl_014</td>
<td>Ctenopseustis obliquana</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225570</td>
</tr>
<tr>
<td>C_obl_015</td>
<td>Ctenopseustis obliquana</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225571</td>
</tr>
<tr>
<td>C_obl_019</td>
<td>Ctenopseustis obliquana</td>
<td>Kaipakati Pt., Truman Track, BR</td>
<td>-4202 17122</td>
<td>FJ225572</td>
</tr>
<tr>
<td>C_obl_022</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255 1725532</td>
<td>FJ225573</td>
</tr>
<tr>
<td>C_obl_023</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255 1725532</td>
<td>FJ225574</td>
</tr>
<tr>
<td>C_obl_024</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255 1725532</td>
<td>FJ225575</td>
</tr>
<tr>
<td>C_obl_026</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255 1725532</td>
<td>FJ225576</td>
</tr>
<tr>
<td>C_obl_029</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255 1725532</td>
<td>FJ225577</td>
</tr>
<tr>
<td>C_obl_1029</td>
<td>Ctenopseustis obliquana</td>
<td>Hamilton, WO</td>
<td>-377854 1752821</td>
<td>FJ225475</td>
</tr>
<tr>
<td>C_obl_103</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225476</td>
</tr>
<tr>
<td>C_obl_1031</td>
<td>Ctenopseustis obliquana</td>
<td>Hamilton, WO</td>
<td>-377854 1752821</td>
<td>FJ225477</td>
</tr>
<tr>
<td>C_obl_1032</td>
<td>Ctenopseustis obliquana</td>
<td>Hamilton, WO</td>
<td>-377854 1752821</td>
<td>FJ225478</td>
</tr>
<tr>
<td>C_obl_1033</td>
<td>Ctenopseustis obliquana</td>
<td>Hamilton, WO</td>
<td>-377854 1752821</td>
<td>FJ225479</td>
</tr>
<tr>
<td>C_obl_104</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225480</td>
</tr>
<tr>
<td>C_obl_108</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225481</td>
</tr>
<tr>
<td>C_obl_110</td>
<td>Ctenopseustis obliquana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020 173200</td>
<td>FJ225482</td>
</tr>
<tr>
<td>C_obl_116</td>
<td>Ctenopseustis obliquana</td>
<td>Grove Scenic Reserve, NN</td>
<td>-4117 17355</td>
<td>FJ225483</td>
</tr>
<tr>
<td>C_obl_118</td>
<td>Ctenopseustis obliquana</td>
<td>Grove Scenic Reserve, NN</td>
<td>-4117 17355</td>
<td>FJ225484</td>
</tr>
<tr>
<td>Code</td>
<td>Species</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>-------------------------------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>C_obl_119</td>
<td>Ctenopseustis obliquana</td>
<td>Grove Scenic Reserve, NN</td>
<td></td>
<td>-4117</td>
</tr>
<tr>
<td>C_obl_137</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_138</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_142</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_147</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_148</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_149</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_150</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_151</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_159</td>
<td>Ctenopseustis obliquana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td></td>
<td>-4123</td>
</tr>
<tr>
<td>C_obl_167</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_171</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_172</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_173</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_175</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_176</td>
<td>Ctenopseustis obliquana</td>
<td>Pelorus River Bridge, SD/MB</td>
<td></td>
<td>-4118</td>
</tr>
<tr>
<td>C_obl_192</td>
<td>Ctenopseustis obliquana</td>
<td>Riwaka Resurgence, NN</td>
<td></td>
<td>-410206</td>
</tr>
<tr>
<td>C_obl_193</td>
<td>Ctenopseustis obliquana</td>
<td>Riwaka Resurgence, NN</td>
<td></td>
<td>-410206</td>
</tr>
<tr>
<td>C_obl_197</td>
<td>Ctenopseustis obliquana</td>
<td>Riwaka Resurgence, NN</td>
<td></td>
<td>-410206</td>
</tr>
<tr>
<td>C_obl_199</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_200</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_201</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_202</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_205</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_206</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_209</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_210</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_212</td>
<td>Ctenopseustis obliquana</td>
<td>Snowden's Bush, Brightwater, NN</td>
<td></td>
<td>-412216</td>
</tr>
<tr>
<td>C_obl_283</td>
<td>Ctenopseustis obliquana</td>
<td>Moss's Scenic Reserve, NN</td>
<td></td>
<td>-410255</td>
</tr>
<tr>
<td>C_obl_536</td>
<td>Ctenopseustis obliquana</td>
<td>Pollen Island, AK</td>
<td></td>
<td>-3652</td>
</tr>
<tr>
<td>C_obl_538</td>
<td>Ctenopseustis obliquana</td>
<td>Tophouse Reserve, Buller, BR</td>
<td></td>
<td>-4143</td>
</tr>
<tr>
<td>Code</td>
<td>Species</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------</td>
<td>-----------------------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>C_obl_544</td>
<td>Ctenopseustis obliquana</td>
<td>St. Arnaud, MB</td>
<td>-4145</td>
<td>17250</td>
</tr>
<tr>
<td>C_obl_545</td>
<td>Ctenopseustis obliquana</td>
<td>Wenderholm, AK</td>
<td>-365245</td>
<td>1747200</td>
</tr>
<tr>
<td>C_obl_641</td>
<td>Ctenopseustis obliquana</td>
<td>Catherine Cove, D'Urville Island, SD</td>
<td>-4052</td>
<td>17354</td>
</tr>
<tr>
<td>C_obl_754</td>
<td>Ctenopseustis obliquana</td>
<td>Aporo Orchard, NN</td>
<td>-411214</td>
<td>1730410</td>
</tr>
<tr>
<td>C_obl_791</td>
<td>Ctenopseustis obliquana</td>
<td>St. Arnaud, Mainland Island, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_obl_794</td>
<td>Ctenopseustis obliquana</td>
<td>St. Arnaud, Mainland Island, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_obl_797</td>
<td>Ctenopseustis obliquana</td>
<td>St. Arnaud, Mainland Island, BR</td>
<td>-414830</td>
<td>1725049</td>
</tr>
<tr>
<td>C_obl_863</td>
<td>Ctenopseustis obliquana</td>
<td>St. Arnaud, BR</td>
<td>-4146</td>
<td>17250</td>
</tr>
<tr>
<td>C_obl_955</td>
<td>Ctenopseustis obliquana</td>
<td>Lake Rotoroa, BR</td>
<td>-414749</td>
<td>1723545</td>
</tr>
<tr>
<td>C_obl_968</td>
<td>Ctenopseustis obliquana</td>
<td>Wenderholm, AK</td>
<td>-365245</td>
<td>1747200</td>
</tr>
<tr>
<td>C_obl_991</td>
<td>Ctenopseustis obliquana</td>
<td>Wenderholm, AK</td>
<td>-365245</td>
<td>1747200</td>
</tr>
<tr>
<td>C_obl_993</td>
<td>Ctenopseustis obliquana</td>
<td>Wenderholm, AK</td>
<td>-365245</td>
<td>1747200</td>
</tr>
<tr>
<td>C_obl_genbank</td>
<td>Ctenopseustis obliquana</td>
<td>Mt Albert, AK</td>
<td>-368880</td>
<td>1747149</td>
</tr>
<tr>
<td>C_obl_man_genbank</td>
<td>Ctenopseustis obliquana</td>
<td>Pollen island, AK</td>
<td>-3652</td>
<td>17440</td>
</tr>
<tr>
<td>C_obl_typellNL_genbank</td>
<td>Ctenopseustis obliquana</td>
<td>Rukuhia, WO</td>
<td>-378488</td>
<td>1753033</td>
</tr>
<tr>
<td>C_ser_genbank</td>
<td>Ctenopseustis servana</td>
<td>Wenderholme, AK</td>
<td>-365245</td>
<td>1747200</td>
</tr>
<tr>
<td>P.avi_735</td>
<td>Planotortrix avicenniae</td>
<td>Pollen Island, AK</td>
<td>-3652</td>
<td>17440</td>
</tr>
<tr>
<td>P.avi_744</td>
<td>Planotortrix avicenniae</td>
<td>Pollen Island, Ak</td>
<td>-3652</td>
<td>17440</td>
</tr>
<tr>
<td>P.avi_genbank</td>
<td>Planotortrix avicenniae</td>
<td>Te Puna, Tauranga, BP</td>
<td>-376959</td>
<td>1760724</td>
</tr>
<tr>
<td>P.exc_012</td>
<td>Planotortrix excessana</td>
<td>Eve's Valley Bush, NN</td>
<td>-412020</td>
<td>173200</td>
</tr>
<tr>
<td>P.exc_094</td>
<td>Planotortrix excessana</td>
<td>Brian Petrie, NN</td>
<td>-411312</td>
<td>1730013</td>
</tr>
<tr>
<td>P.exc_095</td>
<td>Planotortrix excessana</td>
<td>Brian Petrie, NN</td>
<td>-411312</td>
<td>1730013</td>
</tr>
<tr>
<td>P.exc_097</td>
<td>Planotortrix excessana</td>
<td>Brian Petrie, NN</td>
<td>-411312</td>
<td>1730013</td>
</tr>
<tr>
<td>P.exc_1005</td>
<td>Planotortrix excessana</td>
<td>Bluff, SL</td>
<td>-465749</td>
<td>1633174</td>
</tr>
<tr>
<td>P.exc_1042</td>
<td>Planotortrix excessana</td>
<td>Brook Reservoir, NN</td>
<td>-413459</td>
<td>1733601</td>
</tr>
<tr>
<td>P.exc_1049</td>
<td>Planotortrix excessana</td>
<td>Pt. Chevalier, AK</td>
<td>-368679</td>
<td>1747003</td>
</tr>
<tr>
<td>P.exc_107</td>
<td>Planotortrix excessana</td>
<td>Eve's Valley, NN</td>
<td>-412020</td>
<td>173200</td>
</tr>
<tr>
<td>P.exc_156</td>
<td>Planotortrix excessana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123</td>
<td>17351</td>
</tr>
<tr>
<td>P.exc_160</td>
<td>Planotortrix excessana</td>
<td>Payne's Ford, Takaka Valley, NN</td>
<td>-4123</td>
<td>17351</td>
</tr>
<tr>
<td>P.exc_600</td>
<td>Planotortrix excessana</td>
<td>Flora Saddle, NN</td>
<td>-4111</td>
<td>17244</td>
</tr>
<tr>
<td>P.exc_842</td>
<td>Planotortrix excessana</td>
<td>Moss's Scenic Reserve, NN</td>
<td>-410255</td>
<td>1725532</td>
</tr>
<tr>
<td>P.exc_994</td>
<td>Planotortrix excessana</td>
<td>Bluff, SL</td>
<td>-465749</td>
<td>1633174</td>
</tr>
<tr>
<td>Code</td>
<td>Species</td>
<td>Location</td>
<td>Coordinates</td>
<td>Accession Numbers</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------</td>
<td>---------------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>P_exc_genbank</td>
<td>Planotortrix excessana</td>
<td>Rukuhia, Waikato, WO</td>
<td>-378488 1753033</td>
<td>AF016475</td>
</tr>
<tr>
<td>P_fl_017</td>
<td>Planotortrix flammae</td>
<td>Kaipakati Pt., Truman Track, BR</td>
<td>-4202 17122</td>
<td>FJ225625</td>
</tr>
<tr>
<td>P_fl_018</td>
<td>Planotortrix flammae</td>
<td>Kaipakati Pt., Truman Track, BR</td>
<td>-4202 17122</td>
<td>FJ225626</td>
</tr>
<tr>
<td>P_not_1041</td>
<td>Planotortrix notophaea</td>
<td>Brook Reservoir, NN</td>
<td>-413459 1733601</td>
<td>FJ225627</td>
</tr>
<tr>
<td>P_not_309</td>
<td>Planotortrix notophaea</td>
<td>Wakefield, NN</td>
<td>-4124 17303</td>
<td>FJ225628</td>
</tr>
<tr>
<td>P_not_313</td>
<td>Planotortrix notophaea</td>
<td>Eve's Valley, NN</td>
<td>-412020 173200</td>
<td>FJ225629</td>
</tr>
<tr>
<td>P_not_928</td>
<td>Planotortrix notophaea</td>
<td>Pelorus River Bridge, SD/MB</td>
<td>-4118 17334</td>
<td>FJ225630</td>
</tr>
<tr>
<td>P_not_genbank</td>
<td>Planotortrix notophaea</td>
<td>Palmerston North, WI</td>
<td>-403557 1756122</td>
<td>AF016477</td>
</tr>
<tr>
<td>P_octo_098</td>
<td>Planotortrix octo</td>
<td>Brian Petrie, NN</td>
<td>-411312 1730013</td>
<td>FJ225631</td>
</tr>
<tr>
<td>P_octo_100</td>
<td>Planotortrix octo</td>
<td>Brian Petrie, NN</td>
<td>-411312 1730013</td>
<td>FJ225632</td>
</tr>
<tr>
<td>P_octo_1040</td>
<td>Planotortrix octo</td>
<td>Brook Reservoir, NN</td>
<td>-413459 1733601</td>
<td>FJ225633</td>
</tr>
<tr>
<td>P_octo_1044</td>
<td>Planotortrix octo</td>
<td>Okuti Valley, Banks Peninsula, MC</td>
<td>-437868 1728161</td>
<td>FJ225634</td>
</tr>
<tr>
<td>P_octo_125</td>
<td>Planotortrix octo</td>
<td>McKee Memorial Reserve, NN</td>
<td>-4113 17305</td>
<td>FJ225635</td>
</tr>
<tr>
<td>P_octo_164</td>
<td>Planotortrix octo</td>
<td>Pelorus River Bridge, SD/MB</td>
<td>-4118 17334</td>
<td>FJ225636</td>
</tr>
<tr>
<td>P_octo_185</td>
<td>Planotortrix octo</td>
<td>Pupu Valley Walk, NN</td>
<td>-4084 17276</td>
<td>FJ225637</td>
</tr>
<tr>
<td>P_octo_187</td>
<td>Planotortrix octo</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225638</td>
</tr>
<tr>
<td>P_octo_190</td>
<td>Planotortrix octo</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225639</td>
</tr>
<tr>
<td>P_octo_196</td>
<td>Planotortrix octo</td>
<td>Riwaka Resurgence, NN</td>
<td>-410206 1725401</td>
<td>FJ225640</td>
</tr>
<tr>
<td>P_octo_211</td>
<td>Planotortrix octo</td>
<td>Snowdon's Bush, Brightwater, NN</td>
<td>-412216 1730612</td>
<td>FJ225641</td>
</tr>
<tr>
<td>P_octo_292</td>
<td>Planotortrix octo</td>
<td>Grove Scenic Reserve, NN</td>
<td>-4117 17355</td>
<td>FJ225642</td>
</tr>
<tr>
<td>P_octo_293</td>
<td>Planotortrix octo</td>
<td>Grove Scenic Reserve, NN</td>
<td>-4117 17355</td>
<td>FJ225643</td>
</tr>
<tr>
<td>P_octo_296</td>
<td>Planotortrix octo</td>
<td>Nelson Spelo Gp. Hut, Takaka Hill, NN</td>
<td>-4102 17249</td>
<td>FJ225644</td>
</tr>
<tr>
<td>P_octo_601</td>
<td>Planotortrix octo</td>
<td>Flora Saddle, NN</td>
<td>-4111 17244</td>
<td>FJ225645</td>
</tr>
<tr>
<td>P_octo_611</td>
<td>Planotortrix octo</td>
<td>Flora Saddle, NN</td>
<td>-4111 17244</td>
<td>FJ225646</td>
</tr>
<tr>
<td>P_octo_740</td>
<td>Planotortrix octo</td>
<td>Hapuku Res., KA</td>
<td>-4219 17345</td>
<td>FJ225647</td>
</tr>
<tr>
<td>P_octo_788</td>
<td>Planotortrix octo</td>
<td>Omahanui, HB</td>
<td>-411311 1725927</td>
<td>FJ225648</td>
</tr>
<tr>
<td>P_octo_802</td>
<td>Planotortrix octo</td>
<td>Ruby Bay, NN</td>
<td>-411407 1730509</td>
<td>FJ225649</td>
</tr>
<tr>
<td>P_octo_861</td>
<td>Planotortrix octo</td>
<td>Grovetown, MB</td>
<td>-412839 1735858</td>
<td>FJ225650</td>
</tr>
<tr>
<td>P_octo_862</td>
<td>Planotortrix octo</td>
<td>Wye River, NN</td>
<td>-413812 1731225</td>
<td>FJ225651</td>
</tr>
<tr>
<td>P_octo_genbank</td>
<td>Planotortrix octo</td>
<td>Canterbury, MC</td>
<td>-435259 1717070</td>
<td>AF016478</td>
</tr>
<tr>
<td>P_oi_643</td>
<td>Planotortrix octo</td>
<td>Pitt Island, CH</td>
<td>-4418 17612W</td>
<td>FJ225652</td>
</tr>
<tr>
<td>P_oi_644</td>
<td>Planotortrix octo</td>
<td>Pitt Island, CH</td>
<td>-4418 17612W</td>
<td>FJ225653</td>
</tr>
<tr>
<td>Code</td>
<td>Scientific Name</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>P_oi_646</td>
<td>Planotortrix octo</td>
<td>South East Island, CH</td>
<td>-4430</td>
<td>17600W</td>
</tr>
<tr>
<td>P_oi_genbank</td>
<td>Planotortrix octoides</td>
<td>South East Island, CH</td>
<td>-4430</td>
<td>17600W</td>
</tr>
<tr>
<td>P_puf_1012</td>
<td>Planotortrix puffini</td>
<td>Lee Bay, SI</td>
<td>-4652</td>
<td>16807</td>
</tr>
<tr>
<td>P_puf_1018</td>
<td>Planotortrix puffini</td>
<td>Golden Bay track, SI</td>
<td>-468945</td>
<td>1680750</td>
</tr>
<tr>
<td>P_puf_606</td>
<td>Planotortrix puffini</td>
<td>Lee Bay, SI</td>
<td>-4652</td>
<td>16807</td>
</tr>
<tr>
<td>P_puf_607</td>
<td>Planotortrix puffini</td>
<td>Lee Bay, SI</td>
<td>-4652</td>
<td>16807</td>
</tr>
<tr>
<td>P_puf_608</td>
<td>Planotortrix puffini</td>
<td>Ulva Island, SI</td>
<td>-4306</td>
<td>17020</td>
</tr>
<tr>
<td>P_puf_609</td>
<td>Planotortrix puffini</td>
<td>Ulva Island, SI</td>
<td>-4306</td>
<td>17020</td>
</tr>
<tr>
<td>P_puff_genbank</td>
<td>Planotortrix puffini</td>
<td>Stewart Island, SI</td>
<td>-468945</td>
<td>1680750</td>
</tr>
<tr>
<td>ID number</td>
<td>hort</td>
<td>species</td>
<td>gender</td>
<td>location</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>--------------------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>1007</td>
<td>1007</td>
<td>C.herana</td>
<td>female</td>
<td>Bluff</td>
</tr>
<tr>
<td>1005</td>
<td>1005</td>
<td>P.excessana</td>
<td>male</td>
<td>Bluff</td>
</tr>
<tr>
<td>1006</td>
<td>1006</td>
<td>C.herana</td>
<td>male</td>
<td>Bluff</td>
</tr>
<tr>
<td>999</td>
<td>999</td>
<td>C.fratera</td>
<td>male</td>
<td>Titirangi</td>
</tr>
<tr>
<td>991</td>
<td>991</td>
<td>C.obliquana</td>
<td>male</td>
<td>Wenderhom</td>
</tr>
<tr>
<td>970</td>
<td>970</td>
<td>P.avicennia</td>
<td>male</td>
<td>Wenderhom</td>
</tr>
<tr>
<td>994</td>
<td>994</td>
<td>P.excessana</td>
<td>female</td>
<td>Bluff</td>
</tr>
<tr>
<td>993</td>
<td>993</td>
<td>C.obliquana</td>
<td>female</td>
<td>Wenderhom</td>
</tr>
<tr>
<td>962</td>
<td>962</td>
<td>P.octo</td>
<td>male</td>
<td>Te Puke</td>
</tr>
<tr>
<td>965</td>
<td>965</td>
<td>P.octo</td>
<td>male</td>
<td>Whakapapa</td>
</tr>
<tr>
<td>995</td>
<td>995</td>
<td>C.herana</td>
<td>male</td>
<td>Bluff</td>
</tr>
<tr>
<td>Position in box</td>
<td>Genus</td>
<td>species</td>
<td>Country</td>
<td>Crosby code</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>L1</td>
<td>Ctenopseustis</td>
<td>obliquana</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>L1</td>
<td>Ctenopseustis</td>
<td>obliquana</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>L2</td>
<td>Ctenopseustis</td>
<td>herana</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>L2</td>
<td>Ctenopseustis</td>
<td>herana</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>L3</td>
<td>Ctenopseustis</td>
<td>obliquana</td>
<td>NZ</td>
<td>WO</td>
</tr>
<tr>
<td>L3</td>
<td>Ctenopseustis</td>
<td>obliquana</td>
<td>NZ</td>
<td>WO</td>
</tr>
<tr>
<td>L4</td>
<td>Ctenopseustis</td>
<td>servana</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>L4</td>
<td>Ctenopseustis</td>
<td>servana</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>L5</td>
<td>Ctenopseustis</td>
<td>fraterna</td>
<td>NZ</td>
<td>TK</td>
</tr>
<tr>
<td>L5</td>
<td>Ctenopseustis</td>
<td>fraterna</td>
<td>NZ</td>
<td>TK</td>
</tr>
<tr>
<td>L6</td>
<td>Ctenopseustis</td>
<td>filicis</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>L6</td>
<td>Ctenopseustis</td>
<td>filicis</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>R1</td>
<td>Planotortrix</td>
<td>octo</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>R1</td>
<td>Planotortrix</td>
<td>octo</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>R2</td>
<td>Planotortrix</td>
<td>octoides</td>
<td>NZ</td>
<td>CH</td>
</tr>
<tr>
<td>R2</td>
<td>Planotortrix</td>
<td>octoides</td>
<td>NZ</td>
<td>CH</td>
</tr>
<tr>
<td>R3</td>
<td>Planotortrix</td>
<td>excessana</td>
<td>NZ</td>
<td>CH</td>
</tr>
<tr>
<td>R3</td>
<td>Planotortrix</td>
<td>excessana</td>
<td>NZ</td>
<td>CH</td>
</tr>
<tr>
<td>R4</td>
<td>Planotortrix</td>
<td>avicenniae</td>
<td>NZ</td>
<td>BP</td>
</tr>
<tr>
<td>R4</td>
<td>Planotortrix</td>
<td>avicenniae</td>
<td>NZ</td>
<td>BP</td>
</tr>
<tr>
<td>R5</td>
<td>Planotortrix</td>
<td>puffini</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>R5</td>
<td>Planotortrix</td>
<td>puffini</td>
<td>NZ</td>
<td>SI</td>
</tr>
<tr>
<td>R6</td>
<td>Planotortrix</td>
<td>notophae</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>R6</td>
<td>Planotortrix</td>
<td>notophae</td>
<td>NZ</td>
<td>AK</td>
</tr>
<tr>
<td>R7</td>
<td>Planotortrix</td>
<td>flammea</td>
<td>NZ</td>
<td>FD</td>
</tr>
<tr>
<td>Day 1</td>
<td>Month 1</td>
<td>Year 1</td>
<td>Day 2</td>
<td>month 2</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Mar</td>
<td>1986</td>
<td>11 Sep</td>
<td>1986</td>
<td>J S Dugdale</td>
</tr>
<tr>
<td>Mar</td>
<td>1986</td>
<td>14 Sep</td>
<td>1986</td>
<td>J S Dugdale</td>
</tr>
<tr>
<td>Mar</td>
<td>1986</td>
<td>16 Sep</td>
<td>1986</td>
<td>J S Dugdale</td>
</tr>
<tr>
<td>Mar</td>
<td>1986</td>
<td>9 Sep</td>
<td>1986</td>
<td>J S Dugdale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1986</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1986</td>
<td></td>
</tr>
<tr>
<td>em</td>
<td>23 Feb</td>
<td>1976</td>
<td>J S Dugdale</td>
<td></td>
</tr>
<tr>
<td>16 Feb</td>
<td>1976</td>
<td></td>
<td>J S Dugdale</td>
<td></td>
</tr>
<tr>
<td>29 Mar</td>
<td>1987</td>
<td>em</td>
<td>30 Apr</td>
<td>1987</td>
</tr>
<tr>
<td>29 Mar</td>
<td>1987</td>
<td>em</td>
<td>30 May</td>
<td>1987</td>
</tr>
<tr>
<td>Mar</td>
<td>1987</td>
<td>em</td>
<td>20 Apr</td>
<td>1987</td>
</tr>
<tr>
<td>Mar</td>
<td>1987</td>
<td>em</td>
<td>20 Apr</td>
<td>1987</td>
</tr>
<tr>
<td>16 May</td>
<td>1980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Dec</td>
<td>1987</td>
<td>em</td>
<td>19 Jan</td>
<td>1988</td>
</tr>
<tr>
<td>15 Dec</td>
<td>1987</td>
<td>18 Dec</td>
<td>1987</td>
<td>J S Dugdale</td>
</tr>
<tr>
<td></td>
<td>1986-87</td>
<td></td>
<td></td>
<td>J R Clearwater</td>
</tr>
<tr>
<td></td>
<td>1986-87</td>
<td></td>
<td></td>
<td>J R Clearwater</td>
</tr>
<tr>
<td>em</td>
<td>30 Oct</td>
<td>1984</td>
<td>J R Clearwater</td>
<td></td>
</tr>
<tr>
<td>em</td>
<td>30 Oct</td>
<td>1984</td>
<td>J R Clearwater</td>
<td></td>
</tr>
<tr>
<td>4 May</td>
<td>1984</td>
<td>em</td>
<td>July</td>
<td>1984</td>
</tr>
<tr>
<td>4 May</td>
<td>1984</td>
<td>em</td>
<td>July</td>
<td>1984</td>
</tr>
<tr>
<td>26 Oct</td>
<td>1980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 May</td>
<td>1980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Feb</td>
<td>1976</td>
<td>em</td>
<td>17 Feb</td>
<td>1976</td>
</tr>
<tr>
<td>Code</td>
<td>Pheromone code</td>
<td>Notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------------</td>
<td>--------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFF(4)</td>
<td>Z 5, Z 8 14:OAc</td>
<td>Type II North Island/North Island herana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFF(4)</td>
<td>Z 5, Z 8 14:OAc</td>
<td>Type II North Island/North Island herana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFF(1)</td>
<td>Z 5 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5, Z 5 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 7, Z 5 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR coll MK</td>
<td>????????</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR coll QJ</td>
<td>Z 10 16:OAc</td>
<td>abdomen in vial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR coll SJ</td>
<td>Z 10 16:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 8, Z10 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 8, Z10 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR coll MM</td>
<td>Chathams Z 8'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chathams Z 8'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5 gen</td>
<td>Z 7, Z 5, Z 9 14:OAc</td>
<td>parents from DN and WO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5 gen</td>
<td>Z 7, Z 5, Z 9 14:OAc</td>
<td>parents from DN and WO</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5 14:OAc Avicennia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5 14:OAc Avicennia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5, Z 7, Z 9 14:OAc Asteraceae</td>
<td>coll as larva</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 5, Z 7, Z 9 14:OAc Asteraceae</td>
<td>coll as larva</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 7 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z 7 14:OAc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>no code</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

Nucleotide alignments for \( TPI, COI-COII, EF-1\alpha, PBP1 \)

Protein alignment for PBP1
#NEXUS
[MacClade 4.06 registered to Richard Newcomb, HortResearch]

BEGIN DATA;
    DIMENSIONS NTAX=36 NCHAR=2280;
    FORMAT DATATYPE=DNA MISSING=? GAP=-;
    MATRIX

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
</tr>
<tr>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
</tr>
<tr>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
<td>270</td>
</tr>
<tr>
<td>300</td>
<td>310</td>
<td>320</td>
<td>330</td>
<td>340</td>
<td>350</td>
</tr>
<tr>
<td>380</td>
<td>390</td>
<td>400</td>
<td>410</td>
<td>420</td>
<td>430</td>
</tr>
<tr>
<td>460</td>
<td>470</td>
<td>480</td>
<td>490</td>
<td>500</td>
<td>510</td>
</tr>
<tr>
<td>540</td>
<td>550</td>
<td>560</td>
<td>570</td>
<td>580</td>
<td>590</td>
</tr>
<tr>
<td>620</td>
<td>630</td>
<td>640</td>
<td>650</td>
<td>660</td>
<td>670</td>
</tr>
<tr>
<td>700</td>
<td>710</td>
<td>720</td>
<td>730</td>
<td>740</td>
<td>750</td>
</tr>
<tr>
<td>780</td>
<td>790</td>
<td>800</td>
<td>810</td>
<td>820</td>
<td>830</td>
</tr>
<tr>
<td>860</td>
<td>870</td>
<td>880</td>
<td>890</td>
<td>900</td>
<td>910</td>
</tr>
<tr>
<td>940</td>
<td>950</td>
<td>960</td>
<td>970</td>
<td>980</td>
<td>990</td>
</tr>
<tr>
<td>1010</td>
<td>1020</td>
<td>1030</td>
<td>1040</td>
<td>1050</td>
<td>1060</td>
</tr>
<tr>
<td>1080</td>
<td>1090</td>
<td>1100</td>
<td>1110</td>
<td>1120</td>
<td>1130</td>
</tr>
<tr>
<td>1150</td>
<td>1160</td>
<td>1170</td>
<td>1180</td>
<td>1190</td>
<td>1200</td>
</tr>
<tr>
<td>1220</td>
<td>1230</td>
<td>1240</td>
<td>1250</td>
<td>1260</td>
<td>1270</td>
</tr>
<tr>
<td>1290</td>
<td>1300</td>
<td>1310</td>
<td>1320</td>
<td>1330</td>
<td>1340</td>
</tr>
<tr>
<td>1360</td>
<td>1370</td>
<td>1380</td>
<td>1390</td>
<td>1400</td>
<td>1410</td>
</tr>
<tr>
<td>1430</td>
<td>1440</td>
<td>1450</td>
<td>1460</td>
<td>1470</td>
<td>1480</td>
</tr>
<tr>
<td>1500</td>
<td>1510</td>
<td>1520</td>
<td>1530</td>
<td>1540</td>
<td>1550</td>
</tr>
<tr>
<td>1570</td>
<td>1580</td>
<td>1590</td>
<td>1600</td>
<td>1610</td>
<td>1620</td>
</tr>
<tr>
<td>1640</td>
<td>1650</td>
<td>1660</td>
<td>1670</td>
<td>1680</td>
<td>1690</td>
</tr>
<tr>
<td>1710</td>
<td>1720</td>
<td>1730</td>
<td>1740</td>
<td>1750</td>
<td>1760</td>
</tr>
<tr>
<td>1780</td>
<td>1790</td>
<td>1800</td>
<td>1810</td>
<td>1820</td>
<td>1830</td>
</tr>
<tr>
<td>1850</td>
<td>1860</td>
<td>1870</td>
<td>1880</td>
<td>1890</td>
<td>1900</td>
</tr>
<tr>
<td>1920</td>
<td>1930</td>
<td>1940</td>
<td>1950</td>
<td>1960</td>
<td>1970</td>
</tr>
<tr>
<td>1990</td>
<td>2000</td>
<td>2010</td>
<td>2020</td>
<td>2030</td>
<td>2040</td>
</tr>
<tr>
<td>2060</td>
<td>2070</td>
<td>2080</td>
<td>2090</td>
<td>2100</td>
<td>2110</td>
</tr>
<tr>
<td>2130</td>
<td>2140</td>
<td>2150</td>
<td>2160</td>
<td>2170</td>
<td>2180</td>
</tr>
<tr>
<td>2200</td>
<td>2210</td>
<td>2220</td>
<td>2230</td>
<td>2240</td>
<td>2250</td>
</tr>
</tbody>
</table>
| 2270| 2280]
BEGIN CODONS;
CODONPOSSET * CodonPositions =
N: 1-2280;
CODESET * UNTITLED = mtDNA.dros: all;
END;

BEGIN CODONS;
CODONPOSSET * CodonPositions =
N: 1-2280;
CODESET * UNTITLED = mtDNA.dros: all;
END;

BEGIN ASSUMPTIONS;
OPTIONS DEFTYPE=unord PoyTcount=MINSTEPS;
END;

BEGIN MacClade;
Version 4.0 84;
LastModified -1026998770;
FileSettings editor '0' '0' '1' '1';
Singles 100;
Editor 0001100111111101001001001 '1' '12' Geneva '9' '100' '1' all;
EditorPosition '46' '6' '755' '1215';
TreeWindowPosition '46' '6' '959' '1230';
ListWindow Characters closed Geneva '9' '50' '10' '129' '406' 000;
ListWindow Taxa closed Geneva '9' '50' '10' '129' '490' 1000000;
ListWindow Trees closed Geneva '9' '50' '10' '276' '490';
ListWindow TypeSets closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ListWindow ExSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharSetSets closed Geneva '9' '50' '10' '276' '490';
ListWindow TaxSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartitions closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartNames closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ChartWindowPosition '52' '30' '942' '1220';
StateNamesSymbols closed Geneva '9' '10' '50' '30' '148' '220';
WindowOrder Data;
OtherSymbols 6/ 00 ?-=
Correlation '0' '0' '1000' '0' '0' 10011010;
Salmo 000000001;
EditorFile '2';
ExportHTML _ MOSS '100' 110000;
PrettyPrint 10;
EditorToolsPosition '722' '85' '115' '165';
TreeWindowProgram 10;
TreeWindow 0000;
Continuous '0' '3' 1;
Calculations 00000001;
SummaryMode '0' '0' 0;
Charts Geneva '9' (normal) 0010;
NexusOptions '0' '0' '48' 001011001;
TipLabel '1';
# NEXUS
[MacClade 4.06 registered to Richard Newcomb, HortResearch]

BEGIN DATA;
DIMENSIONS NTAX=70 NCHAR=1238;
FORMAT DATATYPE=DNA MISSING=? GAP=-;

MATRIX

```
[       10      20      30      40      50
  60   70  80  90  100   110  120  130
 140  150 160 170 180  190  200  210
 220  230 240 250 260  270  280  290
 300  310 320 330 340  350  360  370
 380  390 400 410 420  430  440  450
 460  470 480 490 500  510  520  530
 540  550 560 570 580  590  600  610
 620  630 640 650 660  670  680  690
 700  710 720 730 740  750  760  770
 780  790 800 810 820  830  840  850
 860  870 880 890 900  910  920  930
 940  950 960 970 980  990 1000
1010 1020 1030 1040 1050 1060 1070
1080 1090 1100 1110 1120 1130 1140
1150 1160 1170 1180 1190 1200 1210
1220 1230 ]
```

```
Pocito 962 A
ACACGTGACCTCGGCAATCACCACCCACCGCCACTTTGATCTTACAAATGCAGTTGATAGCTGACAAACGTACC
ATTCGAGAAGTTCTGAGAAGGAGCCAGCCAGAAATTGGTAATGACCTCCCTTCAAGTACCGCCCGTCTGAGTACAT
TGAGGCTGACGTGCAAGCGGTGATGTACGTTAGTCAGTGCCTTGAGTACCTGAGCTCTGAGCTTTAC
GCGGCAACATGCCTGACGTGCAAGCGGTGATGTACGTTAGTCAGTGCCTTGAGTACCTGAGCCTCAGC
CAGATGCCTCAGCAGCTACTGAGGCTGATGACTCTGAGCTTCTTCTCTTCTTTCTTAGACGTTGACGT
GACGGTGCCTGCTGATGACTCTGAGCTTCTTCTCTTCTTTCTTAGACGTTGACGT
TGCACGTGCCTGCTGATGACTCTGAGCTTCTTCTCTTCTTTCTTAGACGTTGACGT
AATCCGCTGCTGATGACTCTGAGCTTCTTCTCTTCTTTCTTAGACGTTGACGT
CTGAGTACGTTAGTCAGTGCCTTGAGTACCTGAGCTCTGAGCTTTAC
```

```
G TCTCCGAACCTCCCCTGGTC
CGTGCCGCTTTGCTGGCTCTCCCCCCTTGAGAACCTTCCTGAGGTGCGT

[1238]

Pcoti 1 A
ACACGTGCACTCCGGCAAGTCCACCAACACCCGGCAACTTGTAGTCTACAAATGGCGCTGATGACAAAGCTAC
ATCGGAAAGTTCCGAAAGGAGGCACGAAATAGGTGAAAGGCCTCTCAAGTACGCTGGTGGACAAAACGTTAC
CTGCAATCTGCAGCCCGGGCCACAGGAGATTCATCAAAAGACTAAGTCACTGGTGCTACTACGT
GACCATCTCGCCGCTTTGCTGGCTCTCCCCCCTTGAGAACCTTCCTGAGGTGCGT

[1238]
Pcoti 1 A
ACACGTGCACTCCGGCAAGTCCACCAACACCCGGCAACTTGTAGTCTACAAATGGCGCTGATGACAAAGCTAC
ATCGGAAAGTTCCGAAAGGAGGCACGAAATAGGTGAAAGGCCTCTCAAGTACGCTGGTGGACAAAACGTTAC
CTGCAATCTGCAGCCCGGGCCACAGGAGATTCATCAAAAGACTAAGTCACTGGTGCTACTACGT
GACCATCTCGCCGCTTTGCTGGCTCTCCCCCCTTGAGAACCTTCCTGAGGTGCGT

[1238]
CTGGTCCCTCCAAGCCTCTGTGGCTGAGTTCTCCAAAGTGTTCCCGCTCGTGCTGTTGCC????

[1238]

Cfii_3_A

ACACGACGACTCCGGGAAACGTCACCACACCGCCCCACTTGGATCTACAAATTGCGGCTGATCGAAACATGCC
ATCGAGAAGTTGGAAAGGAGGAGAGAGAAGAGGCTCTATCCAATGACGCGTATTGAGGAAAGGAGGACATCC
TGAAGGCTAGCGCCGCTGCTGATATCATCAGATGCCGCTGCTGAGTCTGAGACTGACGTACGCTGAGTCTGAG
GACCACACTCAGTCCTCCGGGACACAGAGTTTCATCAGAAGACATACGTAGTGAACTGCGGCTGCAACAGGCTG
GCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

[1238]

Cfrra_998_B

ACACGACGACTCCGGGAAACGTCACCACACCGCCCCACTTGGATCTACAAATTGCGGCTGATCGAAACATGCC
ATCGAGAAGTTGGAAAGGAGGAGAGAGAAGAGGCTCTATCCAATGACGCGTATTGAGGAAAGGAGGACATCC
TGAAGGCTAGCGCCGCTGCTGATATCATCAGATGCCGCTGCTGAGTCTGAGACTGACGTACGCTGAGTCTGAG
GACCACACTCAGTCCTCCGGGACACAGAGTTTCATCAGAAGACATACGTAGTGAACTGCGGCTGCAACAGGCTG
GCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

[1238]

Cfrra_998_A

ACACGACGACTCCGGGAAACGTCACCACACCGCCCCACTTGGATCTACAAATTGCGGCTGATCGAAACATGCC
ATCGAGAAGTTGGAAAGGAGGAGAGAGAAGAGGCTCTATCCAATGACGCGTATTGAGGAAAGGAGGACATCC
TGAAGGCTAGCGCCGCTGCTGATATCATCAGATGCCGCTGCTGAGTCTGAGACTGACGTACGCTGAGTCTGAG
GACCACACTCAGTCCTCCGGGACACAGAGTTTCATCAGAAGACATACGTAGTGAACTGCGGCTGCAACAGGCTG
GCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

[1238]
BEGIN CODONS;
CODONPOSSET * CodonPositions =
  N: 1-1238;
  CODESET * UNTITLED = mtDNA.dros: all;
END;

BEGIN ASSUMPTIONS;
OPTIONS DEFTYPE=unord PolyCount=MINSTEPS;
END;

BEGIN MacClade;
Version 4.0 84;
LastModified -1026999769;
FileSettings editor '0' '0' '1' '1';
Singles 100;
Editor 00011001111111010001001 11 12' Geneva '9' '100' '1' all;
EditorPosition '44' '53' '931' '1227';
TreeNodePosition '46' '6' '955' '1230';
ListWindow Characters closed Geneva '9' '50' '10' '129' '406' 000;
ListWindow Taxa closed Geneva '9' '50' '10' '129' '490' 100000;
ListWindow Trees closed Geneva '9' '50' '10' '276' '490';
ListWindow TypeSets closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ListWindow ExSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharSets closed Geneva '9' '50' '10' '276' '490';
ListWindow TaxSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartitions closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartNames closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ChartWindowPosition '52' '30' '942' '1220';
StateNamesSymbols closed Geneva '9' '10' '50' '30' '148' '220';
WindowOrder Data;
OtherSymbols &/ 00 ?-;
Correlation '0' '0' '1000' '0' '0' '10011010;
Salmo 00000001;
EditorFile '2';
ExportHTML MOSS '100' 110000;
PrettyPrint 10;
EditorToolsPosition '817' '93' '115' '165';
TreeNodeProgram 10;
TreeNodeW 0000;
Continuous '0' '3' 1;
Calculations 00000001;
SummaryMode '0' '0' 0;
Charts Geneva '9' (normal) 0010;
NexusOptions '0' '0' '48' 001011001;
TipLabel '1';
TreeFont Geneva '9' (normal);
TreeShape 1.0 1.0 0100;
TraceLabels 0101;
ChartColors '0' '0' '65535' '9' '0' 1;
ChartBiggestSpot 1;
ChartPercent 10;
ChartBarWidth '10' 1;
ChartVerticalAxis 10101;
ChartMinMax '0';
TraceAllChangesDisplay '1' 1;
BarsOnBranchesDisplay '0' '0' '60000' '10000' '10000' '10000'
'10000' '60000' '65000' '65000' '65000' '65000' '65000' '6'
'1' 0000101;
ContinuousBranchLabels 0;
AllStatesBranchLabels 1;
IndexNotation '2' 1;
PrintTree 10.00 '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2'
Geneva '9' (normal) Geneva '10' (normal) Geneva '9' (normal) Geneva '9' (normal)
Geneva '9' (bold) Geneva '9' (normal) Geneva '9' (normal) '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '1' '1' '1' '1' '1' '0' '15' '-39' '4' '-40' '0'
'2' '2' '2' '2' '2' '1' '1' '8' '0' '0' '0' '0' '2' 100011100000000000100000111000;
MatchChar 00 .:
EntryInterpretation 01;
ColorOptions 00;
TreeTools '0' '5' '4' '0' '10' '4' '0'
00100111111101110;
EditorTools '0' '0' '0' '1000' '0' '0' '6' '3'
100000101110001;
PairAlign '2' '3' '2' '3' '1' '1' '1'
1 1 1 1 1 1 1 1010;
BothTools '1';
END;
<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pnot1041_B</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Pnot1041_A</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Pcto1040</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Pcto1040</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Cobi_1032</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Cobi_1032</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Cfiil_1024</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Puf_1018</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Puf_1018</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Puf_1012</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Puf_1012</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
<tr>
<td>Pcto_962</td>
<td>ACACGTCGAC TCCGGCAAGT CCACCACAC CGGCCACTTG ATCTACAAAT</td>
</tr>
</tbody>
</table>

GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
GGTGTTGAT CGACAAAGCT ACCACGAGA AGTTCGAGAA CGAGGCCAG
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCGAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGACTGCTAG GTACTACGTG ACCATCATCG ATGCCCCCGG CCACAGGGAT
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
CATACTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATCCTCCTG TGAAGG AAGA ACTAAAGGAG CTT TGCCGAG
AGAAGAACT A CATGCTCCTG TGAAGG AAGA ACTAAAGGAG CTT TGCCGAG
TAAAGGAG CTT TGCCGAG
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA
AGAAGAACT A CATGCTCCTA TGAAGG AAAA ACTAAAGGAG CTT TGCCGAA

TCGTTACAAC CCCACTTGGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC
TCGTTACAAC CCCACTTTGG GTGCTTTTGG TTCCCATTTT TGGCTGGCAC

AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
AGGTCTGAA C TACCTGAATC CCAACA GGAG TCACCTAAAC GGC CATGCTG
Cobinth_Ta--------
MTLNFGKGLDECRTMLPESVTADFNFNTMKACMLKLELVSDGKLHGHNTLEFAKQHG
ADATTQAEIVDLHTCEKLPA-DDPCMVLEWAKCYKAIEIHKRNWVPMDVMAGELAEV

Cobl_Tam--------
MTLNFGKGLDECRTMLPESVTADFNFNTMKACMLKLELVSDGKLHGHNTLEFAKQHG
ADATTQAEIVDLHTCEKLPA-DDPCMVLEWAKCYKAIEIHKRNWVPMDVMAGELAEV

Cobl_Richa--------
MTLNFGKGLDECRTMLPESVTADFNFNTMKACMLKLELVSDGKLHGHNTLEFAKQHG
ADATTQAEIVDLHTCEKLPA-DDPCMVLEWAKCYKAIEIHKRNWVPMDVMAGELAEV

A_polyphemSPEIMKNLSNFGKMDQCKDELSLPSVADLYNFKDDDYMTDRLAGClACLTKLDVD
PDGLNHGNKADFMKHKADETMAQQLVDDIHGCSEKAPP-
NDDKCMKTDVMCFKKEIHKLNWVPMDLTVGEVLAEV
## Nexus Structure

**MacClade 4.06 registered to Richard Newcomb, HortResearch**

BEGIN DATA;

```plaintext
dimensions ntax=42 nchar=144;
format datatype=protein symbols = " 1 2 3 4" missing=? gap=-
charstatelabels

13 1.1, 14 1.2, 15 1.3, 16 1.4, 17 1.5, 18 1.6, 19 1.7, 20 1.8, 21 1.9,
22 1.10, 23 1.11, 24 1.12, 25 1.13, 26 1.14,
27 1.15, 28 1.16, 29 1.17, 30 1.18, 31 1.19,
32 1.20, 33 1.21, 34 1.22, 35 1.23, 36 1.24,
37 1.25, 38 1.26, 39 1.27, 40 1.28, 41 1.29,
42 1.30, 43 1.31, 44 1.32, 45 1.33, 46 1.34,
47 1.35, 48 1.36, 49 1.37, 50 1.38, 51 1.39,
52 1.40, 53 1.41, 54 1.42, 55 1.43, 56 1.44,
57 1.45, 58 1.46, 59 1.47, 60 1.48, 61 1.49,
62 1.50, 63 1.51, 64 1.52, 65 1.53, 66 1.54,
67 1.55, 68 1.56, 69 1.57, 70 1.58, 71 1.59,
72 1.60, 73 1.61, 74 1.62, 75 1.63, 76 1.64,
77 1.65, 78 1.66, 79 1.67, 80 1.68, 81 1.69,
82 1.70, 83 1.71, 84 1.72, 85 1.73, 86 1.74,
87 1.75, 88 1.76, 89 1.77, 90 1.78, 91 1.79,
92 1.80, 93 1.81, 94 1.82, 95 1.83, 96 1.84,
97 1.85, 98 1.86, 99 1.87, 100 1.88, 101 1.89,
102 1.90, 103 1.91, 104 1.92, 105 1.93, 106 1.94,
107 1.95, 108 1.96, 109 1.97, 110 1.98, 111 1.99,
112 1.100, 113 1.101, 114 1.102, 115 1.103, 116 1.104,
117 1.105, 118 1.106, 119 1.107, 120 1.108, 121 1.109,
122 1.110, 123 1.111, 124 1.112, 125 1.113, 126 1.114,
127 1.115, 128 1.116, 129 1.117, 130 1.118, 131 1.119,
132 1.120, 133 1.121, 134 1.122, 135 1.123, 136 1.124,
137 1.125, 138 1.126, 139 1.127, 140 1.128, 141 1.129,
142 1.130, 143 1.131, 144

matrix

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>141</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>142</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pocto_Richard

MTLNFGKGLDCRKMNLPSVTADFDFWKDDHVLTNRTGCAIMCLAKKLEVS-
EGGLHALLGDFAKQHGADATQAEOIVDVIHNECTLPLGALDDPCMIVLEAKYKMEIHKRNWAPSMDVLA
GELLSEV [136]

Pocto2_Tam

MTLNFGKGLDCRKMNLPSVTADFDFWKDDHVLTNRTGCAIMCLAKKLEVS-
EGGLHALLGDFAKQHGADATQAEOIVDVIHNECTLPLGALDDPCMIVLEAKYKMEIHKRNWAPSMDVLA
GELLSEV [136]

Pocto_965

MNLPSVTADFDFWKDDHVLTNRTGCAIMCLAKKLEVS-
EGTLHAGLDFAKQHGA---------
-------------------------------------------------------------------- [59]

Pocto_1044

MNLPSVTADFDFWKDDHVLTNRTGCAIMCLAKKLEVS-
```
EGTLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELL--- [118] 
Pexc_1042

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGTLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pocto_1040

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGTLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELL--- [133] 
Pocto_962

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Posto15_Tam

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pexc_994

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pexc_Tam

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pavi5_Tam

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pavi22_Tam

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pcti_Tam

MTLNFSDGLDECRKEMNLPSVTADFIDFWKDDHVLTNRTGCAIMCLAKKLELVS-
EGQLHGNALDFAKQHGADTAQEIVDVIHNCETLPGALDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Pnot_1041_2

MLPNLPSVTDFIDFWKDDHLTNRTGCAIMCLAKKLELVS-
EGLKHGNALDFAKQHGADTAQEIVDVIHNCETLIPGLDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [121] 
Pnot_1041_1

MTLNFSDGLDECRKELNLPDSVTDFIDFWKDDHLTNRTGCAIMCLAKKLELVS-
EGLKHGNALDFAKQHGADTAQEIVDVIHNCETLIPGLDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELL--- [131] 
Proto22_Tam

MTLNFSDGLDECRKELNLPDSVTDFIDFWKDDHLTNRTGCAIMCLAKKLELVS-
EGLKHGNALDFAKQHGADTAQEIVDVIHNCETLIPGLDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Proto18_Tam

MTLNFSDGLDECRKELNLPDSVTDFIDFWKDDHLTNRTGCAIMCLAKKLELVS-
EGLKHGNALDFAKQHGADTAQEIVDVIHNCETLIPGLDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [136] 
Ppuf_1012

YECRKEMNLPSVTDFIDFWKDDHLTNRTGCAIMCLAKKLELVS-
EGLKHGNALDFAKQHGADTAQEIVDVIHNCETLIPGLDPCMIVLEWAKCYKMEIHKNWAPMDVLA GELLSEV [127]
Ppuf_1018

Puff_Tam

Pflam_Tam

Cserv_Tam

Cher_R

Cobl_R

Cfrat_998

Cfrat_Tam

Cfil_1045

Cfil_1024

Cfil_Tam

Cher_1046

Cfrat_1047

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]

LDPCMVLEWAKCYKAEIHKRNWPSMDVAGELAEEV [133]
BEGIN PAUP;

Iset Base=equal Nst=6 Rmat=(1.0000 1.8424 0.4753 0.4753 1.8424)
Rates=gamma Shape=0.5570 Pinvar=0;
END;

BEGIN CODONS;
ContinuousBranchLabels 0;
AllStatesBranchLabels 1;
IndexNotation '2' '1';
PrintTree 10.00 '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' Geneva '9' (normal) Geneva '10' (normal) Geneva '9' (normal) Geneva '9' (normal)
Geneva '9' (bold) Geneva '9' (normal) Geneva '9' (normal) '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
'0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0' '0'
MatchChar 00 .
EntryInterpretation 01;
ColorOptions 00;
TreeTools '0' '5' '4' '0' '10' '4' '0' 00100111111101110;
EditorTools '0' '0' '0' '1000' '0' '0' '0' '6' '3' '0'
PairAlign '2' '2' '3' '2' '1' '1' '2' '1' '3' 1010;
BothTools '1';
END;
Pflam_Tam
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Cfrat_Tan
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Cofil_Tan
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Coblin_Tam
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Cobinl_Tan
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Cer_Tan
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Cser_Tan
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
Pcto_Richard
ATGACATGGACTTTTGATAAGGCCTGGACGAATGCAGGAAAGGAGTGAACCTCCCGAGACTCTGCAAGTGG
ACCTCCATGACTTCTGCGAGACAGCAAGCCTCTCTGACGACACGAGCCTGGTACACTTGGGCTG
CAAGAAGACGGACCTGCTGCTTCCAGAGCAAATCAGCTAGAGGAATGTCGAGTCGACATCAGAC
GTTGGATGATAGTACTGCTGACCTCGGGAGAACATATGAATGTGGATGATAGTACTGCCGAG
ACCTTCATGGAATGGTACACTGCTGCAAGTGGCAGAATGTCGAGTCGACATCAGAC
GGTCTTCAGCTACTTC A
C GGTCCGTGTAATAC CGTGTCGGACACAGGGACA ACCATTCCTGTACTAGCAG GAA
A GGTACAACAAAGCG CTTCAGATCCCGTAAGGGC ACTACATTAGGCGGAAGCC TCTGGTCAAGGTCGAAGAA C
C CGGCTCCCGACCGT CGCAGAAGAGCGTTAACAC CTAGTGCAGGTGATAAAGG ACCCGATGACAGCGTAGTC G
AGAGGTAAAACATCG TGAACCGGGTGAGGTCGTG CTAGTACGTCCCCAGCAGC T
C GGGTCAAGGCGAAC ACCT

ATGA AGTCTGTTGTTAAGAGGTC GATCGTGTAGGTAACTTCC A
drahciR_lbo C
G ACGTCACTGTCTAA GTCCCTCCAAGTAGAGGCA GGACGTAAGCAGGTCGGGG AATGGTTTCAAGTTACAGT A
C GGTCCGTGTAATAC CGTGTCGGACACAGGGATA ATCAGTCATGCACCAGCAG GAAGGTTTTTAACTACTTC A
C TCTGGTCAAGGTCAAAAAA C
A GGTACAACAAATCG CTTAAGATCCCATAAGGGC ACTACGTCAAACGGCAGCC TCTGGTCAAGGTCAAAAAA C
AAAGG ACCCGTCATCAGCGTAGTC G
TCCCGACCAT TACTGAAGAGCGTCCACAC CTAGTCCAGGTGATAAAGG ACCCGTCATCAGCGTAGTC G

END;

BEGIN PAUP;

Set Base=equal Nst=6 Rmat=(1.0000 1.8424 0.4753 0.4753 1.8424) Rates=gamma Shape=0.5570 Pinvar=0;
END;

BEGIN CODONS;
  CODONPOSSET * CodonPositions =
    1: 1-406\3, 2: 2-407\3, 3: 3-408\3 408;
  CODESET * UNTITLED = Universal: all ;
END;

BEGIN ASSUMPTIONS;
  OPTIONS DEFTYPE=unord PolyTcount=MINSTEPS ;
END;

BEGIN NOTES;

  TEXT TAXON=2 CHARACTER=300 TEXT= G;
END;

BEGIN MacClade;
EntryInterpretation 01;
ColorOptions 00;
TreeTools '0' '5' '4' '0' '10' '4' '0' 00100111111101110;
EditorTools '0' '0' '0' '1000' '0' '0' '6' '3' '0'
11010101110001;
PairAlign '2' '2' '3' '2' '1' '1' '2' '1' '3' 1010;
BothTools '1';
END;
<p>|-footnotes|</p>
<table>
<thead>
<tr>
<th>110</th>
<th>111</th>
<th>112</th>
<th>113</th>
<th>114</th>
<th>115</th>
<th>116</th>
<th>117</th>
<th>118</th>
<th>119</th>
<th>120</th>
<th>121</th>
<th>122</th>
<th>123</th>
<th>124</th>
<th>125</th>
<th>126</th>
<th>127</th>
<th>128</th>
<th>129</th>
<th>130</th>
<th>131</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
</tr>
<tr>
<td>G</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
#NEXUS

[MacClade 4.06 registered to Richard Newcomb, HortResearch]

BEGIN DATA;
    DIMENSIONS NTAX=70 NCHAR=149;
    FORMAT DATATYPE=DNA MISSING=? GAP=- ;
    MATRIX
        [  
          [ 0 10 20 30 40 50 60 70 80 90 100 110 120 130  
            [  
               . . . . . . . . . . . . . . . .  
               . . . . . . . . . . . . . . . .  
            ]
        ]
    Pavi_B
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto1040_A
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto1040_B
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto1044_A
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto962_A
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto962_B
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto965_A
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pcoeto965_B
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT
    GAAACCTTGGAAAGAGGGAGCCGAAAGACAGAAAGAGGTCTGCTTTCAGGCAGACTAAAGCTTTGTACCGG
    CTA [149]
    Pexco1_A
    AAAAAATGACTGTGTTGCGACAAAGGTGGGTAGTCTGGTCTGAAGATTGATTGCTTGGCGATCGGT

This document was created using Smart PDF creator.
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pexcl_B
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pavi1970_A
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pavi1970_B
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Poctil_B
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Poctil_A
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pexc994_A
A AAAAAGATGACTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Poctol_A
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pexc994_B
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Pavi_A
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGT
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Ppuf1018_B
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGC
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Ppuf1018_A
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGC
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
Ppuf4_B
A AAAAAGATGACTTTGGTTGGCCGACAAAATGGTCTACGCCCTAGAGTCATGGCTCTGACTCCTGACATCGGC
GAAACCTTTGGAAAGAGAGGAGGCCGAAAGACAGAAGAGGTTTGTCTTCAAGGCAAGACTAAAGCCCTTGATCCCGG
CTA [149]
This document was created using Smart PDF Converter.
GAAACTCTAGAAGAGAGAGGCCGCAAGACAGAAGAGTTGCTCTTCAAGCAGACTAAAGCTTTGTACCCAGCTA [149]
Cob12_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
Cob12_B
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
Cob11_B
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
Cob11_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
Cob11R_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
Cob11R_B
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
ChER_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
ChER_B
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
C?? [149]
ChER95_B
????????????????????????????????GTGGCTCAGCCTTATTAGGTTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
CTA [149]
ChER95_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
CTA [149]
Cob11032_B
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
CTA [149]
Cob11032_A
AAAAAGATGACTTGTGGTGCAGCAAAAGTTGGCTCAGCCTTAGAGTCTGCTGAAAGTGATCGCCCTGCCTGCTG
AAAACTCTGGAAAGAGAGAAGGCGGAAAGACAGAAGAGGCTGCTCTCCAGCAGACAAAGGCTTTGTACCCAG
CTA [149]
B_769lbo C

ACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

GGCT

?? CTTAGTTCCGAA ACCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGAGAGAGAAGGTCTCAAA G

?? ?

941[

A_769lbo C

T GGCTACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ACCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGGGAGAGAAGGTTTCAAA G

941[ AT C

A_869lbo C

T GGCTACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ACCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGGGAGAGAAGGTTTCAAA G

941[ AT C

B_869lbo C

GTGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

GTGGCTACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGGGAGAGAAGGTTTCAAA G

941[ AT C

A_6401reh C

T GGCTACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

AGTCCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGGGAGAGAAGGTTTCAAA G

G GCCCT

941[ AT C

B_6401reh C

T GGCTACGTCCGCTA ATGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGCCGG AGGGAGAGAAGGTTTCAAA G

941[ AT C

A_899arf C

GGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

941[ AT C

B_899arf C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

941[ AT C

A_1arf C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTTCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

941[ AT C

B_1arf C

TCCGCACTCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA T

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

941[ AT C

A_3lif C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTCCGCACCCGGTGAAACA GCCGTTGGTT????????? ?

G GCCCTAGTTCCGAA ATCAGACGGAC

941[ AT C

B_4201lif C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTCCGCACCCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GAGAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

941[ AT C

B_3lif C

GTGAAACA GCCGTTGGTT????????? ?

G GCCCTAGTTCCGAA ATCAGACGGAC

941[ AT C

A_5401lif C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTCCGCACCCGGTGAAACA GCCGTTGGTTCAGTAGAAA A

GAAGACAGAACGGTCGG AGGGAGAGAAGGTCTCAAA G

G GCCCTAGTTCCGAA ATCAGACGGACTTCTGCTG GA

941[ AT C

B_5401lif C

T GGTTACGTCCGTTA GTGAAAGTCTGGTCTGAGA TTCCGCACCCGGTGAAACA GCCGTTGGTTCAGTAGAAA A
GAAACTCTGGAGAGAGGGAGGCTGGCAAGACAGAAGAGGTGGTCTTCCAGGCAGACCTAAAGCCTCCTGATCCCGG
CTA [149]
Cfi11024_A
AAAAGATGACTTGTGTTGCCGACAAAGTGGCCACGCTTAGAACAGTTGGTTGTGCTGAAGTGGATTCCTGCTCAT
GAGACTCTGGAGAGAGGGAGGCTGGCAAGACAGAAGAGGTGGTCTTCCAGGCAGACCTAAAGCCTCCTGATCCCGG
CTA [149]
;
END;

BEGIN PAUP;

Lset Base=(0.2911 0.2047 0.3188) Nst=6 Rmat=(6.8684 6.3556 0.6211 1.7012 44.9603) Rates=gamma Shape=0.6180 Pinvar=0.5086;

BEGIN CODONS;
  CODONPOSSET * CodonPositions =
    N: 1-149;
  CODESET * UNTITLED = Universal: all;
END;

BEGIN SETS;
  CHARSET partition1 = 1-48;
END;

BEGIN ASSUMPTIONS;
  OPTIONS DEFTYPE=uncord PolyTcount=MINSTEPS;
END;

BEGIN MacClade;
  Version 4.0 84;
  LastModified -1026999677;
  FileSettings editor '0' '0' '1' '1';
  Singles 100;
  Editor 000110011111111110010001001 '1' '12' Geneva '9' '100' '1' 'all';
  EditorPosition '46' '6' '963' '1251';
  TreeWindowPosition '46' '6' '979' '1230';
  ListWindow Characters closed Geneva '9' '50' '10' '833' '399' '000';
  ListWindow Taxa closed Geneva '9' '50' '10' '129' '490' '100000';
  ListWindow Trees closed Geneva '9' '50' '10' '276' '490';
  ListWindowTypeDefs closed Geneva '9' '50' '10' '276' '490';
  ListWindow ExSets closed Geneva '9' '50' '10' '129' '225';
  ListWindow CharSets closed Geneva '9' '50' '10' '129' '181';
  ListWindow TaxSets closed Geneva '9' '50' '10' '276' '490';
  ListWindow CharPartitions closed Geneva '9' '50' '10' '276' '490';
  ListWindow CharPartNames closed Geneva '9' '50' '10' '276' '490';
  ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
  ChartWindowPosition '52' '30' '942' '1220';
  StateNamesSymbols closed Geneva '9' '10' '50' '30' '148' '220';
  WindowOrder Data;
  OtherSymbols &/ 00 ?-;

Pocti_B
AAAAAGATGACTTGGTGTTGCGGACAAGAGTTGCATCAAGGTTCTTTTAAATGTAATGATCGCTCTGCTGATCGT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pocti_A
AAAAAGATGACCTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pexc994_A
AAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Poctol1_A
AAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pexc994_B
AAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pavi_A
AAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Ppufl0138_BAAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
Ppufl0138_BAAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
Ppufl012_BAAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
Ppufl012_BAAAAAGATGACTTGTGGTCCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
Pnot1_A
AAAAAGATGACCTTGGTGTTGCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pnot1_B
AAAAAGATGACCTTGGTGTTGCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCGCTCTGCTGAT
GAAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTACCCGG
CTA
Pnot1041_BAAAAAGATGACCTTGGTGTTGCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
Pnot1041_BAAAAAGATGACCTTGGTGTTGCGGACAAGAGTTGGTCTACGCTTTAGAGATCTGTCTGAAAGTGGTAGATCG
CTGGCTACCGGAAACTTGGGAAAGAGAGGAGCCGAGAGACAGAGAGGAGGTTGCTCTCAGGAGAGCTAAGAGCCTTGTAC
CTA
GAAACTCTGGGAAGAGGAGGCTGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCCTTGATCCCAG
CTA
Cf1r11_B
AAAAGATGACCTTGTTGCCGACAAAGTGAGCGCTACGCCTAGAGCTGCTGCTGAAGTAGTCTGAGTGCATTTGT
GAAACTCTGGGAAGAGGAGGCTGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCCTTGATCCCAG
CTA
Cf1l3_A
????????????TTGGTTGCCGACAAGATGGGCACACGGCCTTAGAGCTGCTGCTGAAGTAGTCTGAGTGCATTTGT
GAAACTCTGGGAAGAGGAGGCTGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCCTTGATCCCAG
CTA
Cf2104_BAAAAGATGACCTTGTTGCCGACAAGATGGGCACACGGCCTTAGAGCTGCTGCTGAAGTAGTCTGAGTGC
CTGCAATGTGTTGAAACTCTGGGAGAGAGGCTGGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCC
TGATCCCAGCTA
Cf213_B
????????????TTGGTTGCCGACAAGATGGGCACACGGCCTTAGAGCTGCTGCTGAAGTAGTCTGAGTGCATTTGT
GAAACTCTGGGAAGAGGAGGCTGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCCTTGATCCCAG
CTA
Cf21045_BAAAAGATGACCTTGTTGCCGACAAGATGGGCACACGGCCTTAGAGCTGCTGCTGAAGTAGTCTGAGTGC
CTGCAATGTGTTGAAACTCTGGGAGAGAGGCTGGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCC
TGATCCCAGCTA
Cf21104_BAAAAGATGACCTTGTTGCCGACAAGATGGGCACACGGCCTTAGAGCTGCTGCTGAAGTAGTCTGAGTGC
CTGCAATGTGTTGAAACTCTGGGAGAGAGGCTGGGCAAGACAAGAAGGTCGTCTTGCAGCAAGACTAAAGCC
TGATCCCAGCTA

This document was created using SmartPDFCreator.com
To remove this message purchase the product at www.SmartPDFCreator.com
#NEXUS

[MacClade 4.06 registered to Richard Newcomb, HortResearch]

BEGIN DATA;
    DIMENSIONS NTAX=34 NCHAR=767;
    FORMAT DATATYPE=DNA MISSING=? GAP=-;
    MATRIX

    | 10 | 20 | 30 | 40 | 50 |
    |----|----|----|----|----|
    60 | 70 | 80 | 90 | 100| 110|
    140| 150| 160| 170| 180| 190|
    220| 230| 240| 250| 260| 270|
    300| 310| 320| 330| 340| 350|
    380| 390| 400| 410| 420| 430|
    460| 470| 480| 490| 500| 510|
    540| 550| 560| 570| 580| 590|
    620| 630| 640| 650| 660| 670|
    700| 710| 720| 730| 740| 750|

Cser1_B

GAAAA GATGACTTTGTTGCCGACAAAGTAATA TATGCAATGCAATTAAATTAACATCTCTGTATACAAAATA
TCGGAATATAATTGTTGCCGGTTGTTAGTGTGTGTGTTAGATATTTAATTATTCCA-
TTAAATTACCCTGCTGATATATAGATATATATACGAAATTACGTTGACATATCTGATGGAAGCTAAACATTGTG
TTCTTGCAAGGTTTGGTTTTATTAAATTTCTTTAAAGTTTCATCTGATGTGAGGAGAGCTTTGAAGCTATGCTTGA
GACAAGATCTCTACTGCTACAAGTTGAGCTAGTGAAGAGTTTTGAATCAGGCCCTTGCGGATGCTGAGCTTTGTG
GTTGGTGAATATGCACGCTGAGCTTTGGTTGGTTGAGCTTATGCTTTGTTTCGCGTGTTTGGTTGAGCTTATGCT

Cser1_A

GAAAA GATGACTTTGTTGCCGACAAAGTAATA TATGCAATGCAATTAAATTAACATCTCTGTATACAAAATA
TCGGAATATAATTGTTGCCGGTTGTTAGTGTGTGTGTTAGATATTTAATTATTCCA-
TTAAATTACCCTGCTGATATATAGATATATATACGAAATTACGTTGACATATCTGATGGAAGCTAAACATTGTG
TTCTTGCAAGGTTTGGTTTTATTAAATTTCTTTAAAGTTTCATCTGATGTGAGGAGAGCTTTGAAGCTATGCTTGA
GACAAGATCTCTACTGCTACAAGTTGAGCTAGTGAAGAGTTTTGAATCAGGCCCTTGCGGATGCTGAGCTTTGTG
GTTGGTGAATATGCACGCTGAGCTTTGGTTGGTTGAGCTTATGCTTTGTTTCGCGTGTTTGGTTGAGCTTATGCT

Cob1_3_A

GAAAA GATGACTTTGTTGCCGACAAAGTAATA TATGCAATGCAATTAAATTAACATCTCTGTATACAAAATA
TCGGAATATAATTGTTGCCGGTTGTTAGTGTGTGTGTTAGATATTTAATTATTCCA-

Cher1_A

AAAAGATGACTTGTTGCGCAAAAGTAAATATAATGACAGCAAATATAATTTAAACATCCTGTATACAAAATA
TGCTGACATAATTTGGGTTGTAAGTGTTTGGTTTGGTTAGATATATATATTATTTCAA-
TTAAATCTTTGCGCTATAGATCAATATACCAAGTACAGTATATACAGGATATATCGTGACATATCCTGATGAGACTAAACATTTTG
GTTCTCGTATAGGTGTGTTTTATTTATATTGAATATGCTGATTATGGGACGCCATATTAGATAGCTGTAAGCTGAGCAGTATCTATTTGGTT
GACAGATCTCCCTAAACAGGATACGATAGATACGATACGAGGCCTAAACCGGAAATCTCGGCTTTACCGCGCGAGGCCCCTA
TCTGTAAGATGTTGAGTAAATAATTACGCTTGAGGCTTGGTCTGCTCAAGGTTTACTGGAATCTAGTCGGCGCT
GTTGTGGTAAATCGCAG
---

Cher2_A

GTTGTTGTGTTTTATTTATATTGAATATGCTGATTATGGGACGCCATATTAGATAGCTGTAAGCTGAGCAGTATCTATTTGGTT
GACAGATCTCCCTAAACAGGATACGATAGATACGATACGAGGCCTAAACCGGAAATCTCGGCTTTACCGCGCGAGGCCCCTA
TCTGTAAGATGTTGAGTAAATAATTACGCTTGAGGCTTGGTCTGCTCAAGGTTTACTGGAATCTAGTCGGCGCT
GTTGTGGTAAATCGCAG
---

Cher95_B

GTTGTTGTGTTTTATTTATATTGAATATGCTGATTATGGGACGCCATATTAGATAGCTGTAAGCTGAGCAGTATCTATTTGGTT
GACAGATCTCCCTAAACAGGATACGATAGATACGATACGAGGCCTAAACCGGAAATCTCGGCTTTACCGCGCGAGGCCCCTA
TCTGTAAGATGTTGAGTAAATAATTACGCTTGAGGCTTGGTCTGCTCAAGGTTTACTGGAATCTAGTCGGCGCT
GTTGTGGTAAATCGCAG
---

Cher95_A

GTTGTTGTGTTTTATTTATATTGAATATGCTGATTATGGGACGCCATATTAGATAGCTGTAAGCTGAGCAGTATCTATTTGGTT
GACAGATCTCCCTAAACAGGATACGATAGATACGATACGAGGCCTAAACCGGAAATCTCGGCTTTACCGCGCGAGGCCCCTA
TCTGTAAGATGTTGAGTAAATAATTACGCTTGAGGCTTGGTCTGCTCAAGGTTTACTGGAATCTAGTCGGCGCT
GTTGTGGTAAATCGCAG
CTAATTCTTTACTTTCTTTCT TATACTTCAGT
TAATGA AAGTCTGGTCTGAGATTTC GCACTCGGTGGACCCTTTT T
AATCAGACGGACTT CTGCTGGAGAAGACAGAAC GGCCGGAGAGAGAGAAGGT CTCAAAGTGGCTACGTCCG C
CCTAGTTCC G

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G

ACATATGTCC TACAAATTAAATTAACGAC ACGTAATATAAATGAAACA GCCGTTGGTTCAGTAGAAA A

CCTAGTTCC G
Cfra1_A
GAAAGATGACTTGGTTGCCGACAAAGTAAATATAATGCACAGCAATTAAATTAAACATCCTGTATACAAAATA
TGTGAACATAATTGGGGTGTTAGTGTGTGTTGTGTTAAGATATTTAATTATTCATTAATTACCTTGCCTGATATCAGATATATATCAGCAATATCGGTGACATATCTGATGGAGACTAAAACATTTG
GTTCTTGCATAGGTTTGGTTTTATTTAAATTTAAGTTTATCTGATTGATGGAGACCGATTATCATAGCCTGTA
GACAAGATCCCTAATAGCGAATACGATAAGCGCTAAACGCGAAATTCTATCCCTTTCTCACGCGCACGCCTTA
CTCGTAAGAGTGGGATAAAATATCGGCGTTAGCGTTCGTCGTCAACGTTTATAGGCATCTTGTGTAGGGCCGT
GTTTGGGTAAATGCCAG-----------------------------------------------------------------------------------------------------------------------------------------------------------TAACTTCATATTCTTTCTTTCATTTCGTAACC-----TTTTTCCCAGGTGGCTCACGCTTTAGAGTCTGGTCTGAAAGTAATCGCCTGCATCGGTGAAACTCTAGAAGAGAGAGAGGCCGGCAAGACAGAAGAGGTCGTCTTCAGGCAGACTAAA
GCCTTGATCCCGGCTA
[602]
Cfra1_A
AAAAGATGACTTGGTTGCCGACAAAGTATGTATAAGGCGCAGCATTTAAATTAAACATCCTGAATACAAATATGTGAACATACTTGGGGTGTTAG------TGTGTTTAGATAATTAATTATTATTTTAATTACCTTGCCTGATATCAGATATAT-CAGCGATATCGGTGTTATATCTGATGGAGACTAAAACATTTGG-----------TTTGGTTTTATGTAAATTTAAGGTTATCTG---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ATTAAAGTGGACTGATGATCTGCACCTCTAGCACGTCGCGTCATTTTCTCGCTTGTCCCTTTTCTACCTGGTG
AGTCAGAGAAGGACGCACTGAGCGACAAGCGAGAACATTACGTGGCATGCTAAGCCCCCTGGTCGTGTTTGGA
CAGAAAGCCAGTGACTTAATTTTCTTTCTTTCA----ATACCTTGTTTTTTTTCCCAGGTGGCTCACGCCTTAGAGTCTGGTCTGAAAGTGATTGCCTGCATTGGTGAAACTCTGGAAGAGAGGGAGGCTGGCAAGACAGAAGAGGTCGTCTTCAGGCAGACTAAAG
CCTTGATCCCGGCTA
[546]
Cfil3_A
??????????TTGGTTGCCGACAAAGTATGTATAATGCGCAGCATTTAAATTAAACATCCTGTATACAAATATGTGAACATACTTGGGGTGTTAG------TGTGTTTAGATAATTAATTATTATTTTAATTACCTTGCCTGATATCAGATATAT-CAGCGATATCGGTGTTATATCTGATGGAGACTAAAACATTTGG-----------TTTGGTTTTATGTAAATTTAAGGTTATCTG---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ATTAAAGTGGACTGATGATCTGCACCTCTAGCACGTCGCGTCATTTTCTCGCTTGTCTCTTTTCTACCTGGTG
AGTCAGAGAAGGACGCACTGAGCGACAAGCGAGAACATTACGTGGCATGCTAAGCCCCCTGGTCGTGTTTGGA
CAGAAAGCCAGTGACTTAATTTTCTTTCTTTCA----ATACCTTGTTTTTTTTCCCAGGTGGCCCACGCCTTAGAGTCTGGTCTGAAAGTGATTGCCTGCATTGGTGAAACTCTGGAAGAGAGGGAGGCTGGCAAGACAGAAGAGGTCGTCTTCAGGCAGACTAAAG
CCTTGATCCCGGCTA
[546]
Cfil1024_A
AAAAGATGACTTGGTTGCCGACAAAGTATGTATAATGCGCAGCATTTAAATTAAACATCCTGTATACAAATATGTGAACATACTTGGGGTGTTAG------TGTGTTTAGATAATTAATTATTATTTTAATTACCTTGCCTGATATCAGATATAT-CAGCGATATCGGTGTTATATCTGATGGAGACTAAAACATTTGG-----------TTTGGTTTTATGTAAATTTAAGGTTATCTG---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ATTAAAGTGGACTGATGATCTGCACCTCTAGCACGTCGCGTCATTTTCTCGCTTGTCTCTTTTCTACCTGGTG
AGTCAGAGAAGGACGCACTGAGCGACAAGCGAGAACATTACGTGGCATGCTAAGCCCCCTGGTCGTGTTTGGA
CAGAAAGCCAGTGACTTAATTTTCTTTCTTTCA----ATACCTTGTTTTTTTTCCCAGGTGGCCCACGCCTTAGAGTCTGGTCTGAAAGTGATTGCCTGCATTGGTGAAACTCTGGAAGAGAGGGAGGCTGGCAAGACAGAAGAGGTCGTCTTCAGGCAGACTAAAG
CCTTGATCCCGGCTA
[546]


Cfl13_A

TTGTTGGCCCAAAAGATGATATGATAATGGCAGCACTTTAAATCAAACATCACTGTATAC-
AAATATGCAACATACATTTGGGTGTAG--------
TGTTTTAGATATTAATATTATATATTATTACCTTGACTGATACGATATAT---
CAGCGATATCTGCTATATCGTATGAGACTAAACATTTGG------------------
TTTTGTATGAAATTTAAGGTATCTG---------------------
-----------------------------------------------
-----
ATTAAGCTGACTGATCGATGACATCCACTCTACTGCTGCGTCATTCTTGCTCTTCCTTTCTACCTGGTG
AGTCAGAGAAGAGGCACGTAGCGCAAGCGAGAAACATTACGTGCGATGCTAGCGCTACCCCTGGGTGTCTGGTAGA
CAGAAAGCCAGTGACTTATTTCTCTTTCTCACA---------
ATACCTGTGTTTTTTCCAGGTGCGCCCAGCCCTTAGACTGCTGCTGAAAGTAGATT--
GCTCGATATGGTGAACCTCTGGAAGAGGAGGAGCTGCCACAGACAGAGAGGGTGCTCGTCTCAGCAACATAAAG
CCTGTGACCGCGCTA [546]

Cfl11045_A

AAAAAGATGACCTTGTTGGCCCAAAAGATGATATGATAATGGCAGCACTTTAAATCAAACATCACTGTATAC-
AAATATGCAACATACATTTGGGTGTAG--------
TGTTTTAGATATTAATATTATATATTATTTATACCTTGACTGATACGATATAT---
CAGCGATATCTGCTATATCGTATGAGACTAAACATTTGG------------------
TTTTGTATGAAATTTAAGGTATCTG---------------------
-----------------------------------------------
-----
ATTAAGCTGACTGATCGATGACATCCACTCTACTGCTGCGTCATTCTTGCTCTTCCTTTCTACCTGGTG
AGTCAGAGAAGAGGCACGTAGCGCAAGCGAGAAACATTACGTGCGATGCTAGCGCTACCCCTGGGTGTCTGGTAGA
CAGAAAGCCAGTGACTTATTTCTCTTTCTCACA---------
ATACCTGTGTTTTTTCCAGGTGCGCCCAGCCCTTAGACTGCTGCTGAAAGTAGATT--
GCTCGATATGGTGAACCTCTGGAAGAGGAGGAGCTGCCACAGACAGAGAGGGTGCTCGTCTCAGCAACATAAAG
CCTGTGACCGCGCTA [546]

Cfl11024_A

AAAAAGATGACCTTGTTGGCCCAAAAGATGATATGATAATGGCAGCACTTTAAATCAAACATCACTGTATAC-
AAATATGCAACATACATTTGGGTGTAG--------
TGTTTTAGATATTAATATTATATATTATTTATACCTTGACTGATACGATATAT---
CAGCGATATCTGCTATATCGTATGAGACTAAACATTTGG------------------
TTTTGTATGAAATTTAAGGTATCTG---------------------
-----------------------------------------------
-----
ATTAAGCTGACTGATCGATGACATCCACTCTACTGCTGCGTCATTCTTGCTCTTCCTTTCTACCTGGTG
AGTCAGAGAAGAGGCACGTAGCGCAAGCGAGAAACATTACGTGCGATGCTAGCGCTACCCCTGGGTGTCTGGTAGA
CAGAAAGCCAGTGACTTATTTCTCTTTCTCACA---------
ATACCTGTGTTTTTTCCAGGTGCGCCCAGCCCTTAGACTGCTGCTGAAAGTAGATT--
BEGIN PAUP;

Lset Base=(0.2739 0.2021 0.2265) Nst=2 TRatio=1.1808 Rates=gamma Shape=0.6820 Pinvar=0.4749;

END;

BEGIN CODONS;

CODONPOSSET * CodonPositions =
N: 1-767;
CODESET * UNTITLED = Universal: all ;

END;

BEGIN SETS;

 CHARSET partition1 = 41-88;

END;

BEGIN ASSUMPTIONS;

OPTIONS DEFTYPE=unord PolyTcount=MINSTEPS ;

END;
BEGIN MacClade:
Version 4.0 84;
LastModified `-1026999594';
FileSettings editor '0' '0' '1' '1';
Singles 100;
Editor 000110011111110010001001 '1' '12' Geneva '9' '100' '1' all;
EditorPosition '46' '6' '963' '1251';
TreeWindowPosition '46' '6' '971' '1230';
ListWindow Characters closed Geneva '9' '50' '10' '833' '399' 000;
ListWindow Taxa closed Geneva '9' '50' '10' '129' '490' 1000000; 
ListWindow Trees closed Geneva '9' '50' '10' '276' '490';
ListWindow TypeSets closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ListWindow ExSets closed Geneva '9' '50' '10' '129' '225';
ListWindow CharSets closed Geneva '9' '50' '10' '129' '181';
ListWindow TaxSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartitions closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartNames closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ChartWindowPosition '52' '30' '942' '1220';
StateNamesSymbols closed Geneva '9' '10' '50' '30' '148' '220';
WindowOrder Data:
OtherSymbols 6/00 ?-;
Correlation '0' '0' '1000' '0' '0' 10011010;
Salmo 00000000;
EditorFile '2';
ExportHTML _ MOSS '100' 110000;
PrettyPrint 10;
EditorToolsPosition '796' '739' '115' '165';
TreeWindowProgram 10;
TreeWindow 0000;
Continuous '0' '3' 1;
Calculations 0000001;
SummaryMode '0' '0' 0;
Charts Geneva '9' (normal) 0010;
NexusOptions '0' '0' '48' 001011001; 
TipLabel '1';
TreeFont Geneva '9' (normal); 
TreeShape 1.0 1.0 0100;
TraceLabels 0101;
ChartColors '0' '0' '65535' '9' '0' 1;
ChartBiggestSpot 1;
ChartPercent 10;
ChartBarWidth '10' 1;
ChartVerticalAxis 10101;
ChartMinMax '0';
TraceAllChangesDisplay '1' 1;
BarsOnBranchesDisplay '0' '0' '60000' '10000' '10000' '10000' 
'10000' '60000' '65000' '65000' '65000' '65000' '65000' '6' '1' 0000101;
ContinuousBranchLabels 0;
AllStatesBranchLabels 1;
IndexNotation '2';
PrintTree 10.00 '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' '2' 
Geneva '9' (normal) Geneva '10' (normal) Geneva '9' (normal) Geneva '9' (normal) 
Geneva '9' (bold ) Geneva '9' (normal) Geneva '9' (normal) '0' '0' '0'
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GATATAT--C</td>
<td>ACCTGATATCG</td>
<td>GTTTATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTGACATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATAC</td>
<td>ACCTGATATCG</td>
<td>GTGTATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTTATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATAC</td>
<td>ACCTGATATCG</td>
<td>GTGATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTGATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTGATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTGATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATATATATC</td>
<td>ACCTGATATCG</td>
<td>GTGATATATATC</td>
<td>TGATGGAGAC</td>
<td>TAAACATTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
<tr>
<td>TAGTTAGTC</td>
<td>TATTTGAATTT</td>
<td>AAATTT</td>
<td>ATTT</td>
<td>TGGTTTGGAT</td>
<td>ACG</td>
<td>TTCTTGG</td>
</tr>
</tbody>
</table>

**Note:** The above text seems to be a sequence alignment or a genetic code representation, possibly related to bioinformatics or genetics. The format and content suggest it might be part of a larger study or research document. The specific context would depend on additional information not provided in the image.
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------
--- -------

CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
CGTTGATGCA TAAGACCTCAGGGTTGGTCTAAAATCCGTAATTCGCTATGG
# NEXUS

[MacClade 4.06 registered to Richard Newcomb, HortResearch]

BEGIN DATA;
DIMENSIONS NTAX=36 NCHAR=737;
FORMAT DATATYPE=DNAA MISSING=? GAP=-;

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
</tr>
<tr>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
</tr>
<tr>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
<td>270</td>
</tr>
<tr>
<td>300</td>
<td>310</td>
<td>320</td>
<td>330</td>
<td>340</td>
<td>350</td>
</tr>
<tr>
<td>380</td>
<td>390</td>
<td>400</td>
<td>410</td>
<td>420</td>
<td>430</td>
</tr>
<tr>
<td>460</td>
<td>470</td>
<td>480</td>
<td>490</td>
<td>500</td>
<td>510</td>
</tr>
<tr>
<td>540</td>
<td>550</td>
<td>560</td>
<td>570</td>
<td>580</td>
<td>590</td>
</tr>
<tr>
<td>620</td>
<td>630</td>
<td>640</td>
<td>650</td>
<td>660</td>
<td>670</td>
</tr>
<tr>
<td>700</td>
<td>710</td>
<td>720</td>
<td>730</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pavi_B
GAAAAAGATGACCTTGGTTGGCCGACAAGGTATGTATAATCCGCGCAA------
TTAACATACTGTA---------AAATGACTAGGCGTTATAGGTC---------
GTTGAGATATTTAATTAGTAAATTGCAAATTGCGTTGCTATGATATCAGGATATAATCAAT---
TTTCTTGCATACTTTTGTTGTCCTGCTGAAATATCGCTTTAT---
TTTTAATTAGTAAATGTATAGGCTTGGTTAGTCATGCTTAGCTGAAAAATAGTTGC
TTGCTGCTGCAAAGCTTGGTGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TTGATCC [407]

Pavi_A
GAAAAAGATGACCTTGGTTGGCCGACAAGGTATGTATAATCCGCGCAA------
TTAACATACTGTA---------AAATGACTAGGCGTTATAGGTC---------
GTTGAGATATTTAATTAGTAAATTGCAAATTGCGTTGCTATGATATCAGGATATAATCAAT---
TTTCTTGCATACTTTTGTTGTCCTGCTGAAATATCGCTTTAT---
TTTTAATTAGTAAATGTATAGGCTTGGTTAGTCATGCTTAGCTGAAAAATAGTTGC
TTGCTGCTGCAAAGCTTGGTGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TTGATCC [407]

Postco1040_A
GAAAAAGATGACCTTGGTTGGCCGACAAGGTATGTATAATCCGCGCAA------
TTAACATACTGTA---------AAATGACTAGGCGTTATAGGTC---------
GTTGAGATATTTAATTAGTAAATTGCAAATTGCGTTGCTATGATATCAGGATATAATCAAT---
TTTCTTGCATACTTTTGTTGTCCTGCTGAAATATCGCTTTAT---
TTTTAATTAGTAAATGTATAGGCTTGGTTAGTCATGCTTAGCTGAAAAATAGTTGC
TTGCTGCTGCAAAGCTTGGTGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TTGATCC [407]
TTAATTAAGATCACTGATGATGTTGTAACAGTGC
---------TGGTCAGCTTTGGGATAGAAAGCCA--
GTCACCT---------
ATTTTCTTTCTTTAATTTCGTAATATTGTCCCCAGTTGGTCAGGCTTAACTAGATGTGTAATGATGC
TTGCATCGGTGAACCTCTTGAGAGGAAGAGGTGCTGCCTTCAGCAGACTAAAGGCC
TGTACGCC [450]
Pexcl A GAAAAGATGACTTTTGTGGCCGACAAGGTATGTATAATCCCCGCGCA----
TTAAACATACGTA------------------AACATGACTAAGGCCGTAGACTGTA
---
-----------------------------------------
GGTCGATATTTTAAATTAGTAATTCAATCCGCCTCAGGCTATGATACCGGATAAATCAGATAT----
---
CGGCATATACTCGATGATGATGTCGTAACAGTGC
---------TGGTCAGCTTTGGGATAGAAAGCCA--
GTCACCT---------
ATTTTCTTTCTTTAATTTCGTAATATTGTCCCCAGTTGGTCAGGCTTAACTAGATGTGTAATGATGC
TTGCATCGGTGAACCTCTTGAGAGGAAGAGGTGCTGCCTTCAGCAGACTAAAGGCC
TGTACGCC [450]
Pav1970 A GAAAAGATGACTTTTGTGGCCGACAAGGTATGTATAATCCCCGCGCA----
TTAAACATACGTA------------------AACATGACTAAGGCCGTAGACTGTA
---
-----------------------------------------
GGTCGATATTTTAAATTAGTAATTCAATCCGCCTCAGGCTATGATACCGGATAAATCAGATAT----
---
CGGCATATACTCGATGATGATGTCGTAACAGTGC
---------TGGTCAGCTTTGGGATAGAAAGCCA--
GTCACCT---------
ATTTTCTTTCTTTAATTTCGTAATATTGTCCCCAGTTGGTCAGGCTTAACTAGATGTGTAATGATGC
TTGCATCGGTGAACCTCTTGAGAGGAAGAGGTGCTGCCTTCAGCAGACTAAAGGCC
TGTACGCC [450]
Pav1970 A GAAAAGATGACTTTTGTGGCCGACAAGGTATGTATAATCCCCGCGCA----
TTAAACATACGTA------------------AACATGACTAAGGCCGTAGACTGTA
---
-----------------------------------------
GGTCGATATTTTAAATTAGTAATTCAATCCGCCTCAGGCTATGATACCGGATAAATCAGATAT----
---
CGGCATATACTCGATGATGATGTCGTAACAGTGC
---------TGGTCAGCTTTGGGATAGAAAGCCA--
GTCACCT---------
TTTTTCTTTCTTTAAATTTCGTGCTATTTCCCAGGTGGCTCAGGCCTTAGGCTGCTGCTGAAAGGTGATC
CTGCGATCGTGAGAACCTTGGAAAGAGAGGGCCGAAAAGCAAGAAGAGAGCTGCTCTCAGGCAAGACTAAAGG
CTTGGATCC [450]
Poctil_B GAAAAAGATGACTTTGGTGCGACAAGGTATGATAATAGCGCGAGCA-------
TTAACATACGTGA-----------AACATCTAGGGCGTTAAGGTGAT-------
--------------------------
GGTGGATATTTAAATTAGTAAATTTACGTTACATTTGCGGCGTATAGTATCCGATAATAATCAACCAGAGAT------
---
CAGCGATATATATCGTGAGACTTAAACTAAATTTTTGGTCTTAAAGAGTTTGGTTTGGGAATAATAGCTTT
AT-----------------------------------
--------------------------
TTAAATTAAAGATGACTTGATGATGCTGATGCAAGGCTC---------TGGTATGTTTTGGATAGAAAGCCA--
GTGACCTC------
ATTTTCTTTCTTTTTAATTTCGTGATATTTGCTCCAGGTGGCTCAGCGCTTAGGCTGCTGAAAGGTGATCGC
CTGCGATCGTGAGAACCTTGGAAAGAGAGGGCCGAAAAGCAAGAAGAGAGCTGCTCTCAGGCAAGACTAAAGG
CTTGGATCC [450]
Poctil_A GAAAAAGATGACTTTGGTGCGACAAGGTATGATAATAGCGCGAGCA-------
TTAACATACGTGA-----------AACATCTAGGGCGTTAAGGTGAT-------
--------------------------
GGTGGATATTTAAATTAGTAAATTTACGTTACATTTGCGGCGTATAGTATCCGATAATAATCAACCAGAGAT------
---
CAGCGATATATATCGTGAGACTTAAACTAAATTTTTGGTCTTAAAGAGTTTGGTTTGGGAATAATAGCTTT
AT-----------------------------------
--------------------------
TTAAATTAAAGATGACTTGATGATGCTGATGCAAGGCTC---------TGGTATGTTTTGGATAGAAAGCCA--
GTGACCTC------
ATTTTCTTTCTTTTTAATTTCGTGATATTTGCTCCAGGTGGCTCAGCGCTTAGGCTGCTGAAAGGTGATCGC
CTGCGATCGTGAGAACCTTGGAAAGAGAGGGCCGAAAAGCAAGAAGAGAGCTGCTCTCAGGCAAGACTAAAGG
CTTGGATCC [450]
Pocci_B GAAAAAGATGACTTTGGTGCGACAAGGTATGATAATAGCGCGAGCA-------
TTAACATACGTGA-----------AACATCTAGGGCGTTAAGGTGAT-------
--------------------------
GGTGGATATTTAAATTAGTAAATTTACGTTACATTTGCGGCGTATAGTATCCGATAATAATCAACCAGAGAT------
---
CAGCGATATATATCGTGAGACTTAAACTAAATTTTTGGTCTTAAAGAGTTTGGTTTGGGAATAATAGCTTT
AT-----------------------------------
--------------------------
TTAAATTAAAGATGACTTGATGATGCTGATGCAAGGCTC---------TGGTATGTTTTGGATAGAAAGCCA--
GTGACCTC------
ATTTTCTTTCTTTTTAATTTCGTGATATTTGCTCCAGGTGGCTCAGCGCTTAGGCTGCTGAAAGGTGATCGC
CTGCGATCGTGAGAACCTTGGAAAGAGAGGGCCGAAAAGCAAGAAGAGAGCTGCTCTCAGGCAAGACTAAAGG
CTTGGATCC [450]
Pocci_A GAAAAAGATGACTTTGGTGCGACAAGGTATGATAATAGCGCGAGCA-------
TTAACATACGTGA-----------AACATCTAGGGCGTTAAGGTGAT-------
--------------------------
GGTGGATATTTAAATTAGTAAATTTACGTTACATTTGCGGCGTATAGTATCCGATAATAATCAACCAGAGAT------
---
CAGCGATATATATCGTGAGACTTAAACTAAATTTTTGGTCTTAAAGAGTTTGGTTTGGGAATAATAGCTTT
AT-----------------------------------
--------------------------
TTAAATTAAAGATGACTTGATGATGCTGATGCAAGGCTC---------TGGTATGTTTTGGATAGAAAGCCA--
GTGACCTC------
ATTTTCTTTCTTTTTAATTTCGTGATATTTGCTCCAGGTGGCTCAGCGCTTAGGCTGCTGAAAGGTGATCGC
CTGCGATCGTGAGAACCTTGGAAAGAGAGGGCCGAAAAGCAAGAAGAGAGCTGCTCTCAGGCAAGACTAAAGG
CTTGGATCC [450]
TCAACTTAAATTTTCTTTTCTACTTCGTAATCTTTTTCTCCAGGTGGTCACGCCCCTAGAGTGTTGCTGCTGAAAGTGACTGAGCTCAGGCGGAAACATTGGAAGAGGAGGGAGGTGGAGAAAGACAGAAGAGGT

BEGIN PAUP;

Lset Base=(0.3003 0.1726 0.2048) Nst=6 Rmat=(1.0000 1.8015 1.0000 1.0000 2.8728) Rates=gamma Shape=0.3887 Pinvar=0;

BEGIN CODONS;
CODONPOSSET * CodonPositions =
N: 1-737;
CODESET * UNTITLED = Universal: all;

BEGIN SETS;
CHARSET partition1 = 1-48;

BEGIN ASSUMPTIONS;
OPTIONS DEFTYPE=unord PolyTcount=MINSTEPS;

BEGIN MacClade;
Version 4.0 84;
LastModifed 1026999548;
FileSettings editor '0' '0' '1' '1';
Singles 100;
Editor 000110011111110010001001 1' '12' Geneva '9' '100' '1' all;
EditorPosition '46' '6' '963' '1251';
TreeWindowPosition '46' '6' '955' '1230';
ListWindow Characters closed Geneva '9' '50' '10' '833' '399' 000;
ListWindow Taxa closed Geneva '9' '50' '10' '129' '490' 100000;
ListWindow Trees closed Geneva '9' '50' '10' '276' '490';
ListWindow TypeSets closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ListWindow ExSets closed Geneva '9' '50' '10' '129' '225';
ListWindow CharSets closed Geneva '9' '50' '10' '129' '181';
ListWindow TaxSets closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartitions closed Geneva '9' '50' '10' '276' '490';
ListWindow CharPartNames closed Geneva '9' '50' '10' '276' '490';
ListWindow WtSets closed Geneva '9' '50' '10' '276' '490';
ChartWindowPosition '52' '30' '942' '1220';
StateNamesSymbols closed Geneva '9' '10' '50' '30' '148' '220';
WindowOrder Data;
OtherSymbols &/ 00 ?--;
Correlation '0' '0' '1000' '0' '0' 10011010;
Salmo 00000001;
EditorFile '2';
36 737

Pavi_B  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Pavi_A  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto1040_ GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto1040_ GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto1B  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto1A  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto144_ GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto144_ GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto962_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto962_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto965_B GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posto965_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Pex1_B  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Pex1_A  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Pavi970_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GTGGCA----
Pavi970_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GTGGCA----
Posti1_B  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posti1_A  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posti_B  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Posti_A  GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAAGC GCAGCA----
Pex994_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAATC GCAGCA----
Pex994_A GAAAAAGATG ACTTGGTGGC CGACAAGGTA TGATATAATC GCAGCA----
Ppuf1018_A GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Ppuf1018_B GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Ppuf4_B  GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Ppuf4_A  GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Ppuf1012_A GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Ppuf1012_A GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1_A  GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1_A  GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1041_A ?AAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1041_A ?AAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1051_A ?AAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pnot1051_A ?AAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pf1am1_A GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT
Pf1am1_A GAAAAAGATG ACTTGGTGGC CGACAAGA GTGATATAGT GCACCAATT

---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
---TTAACAT ACTGTA---- ---------A ACATGACTAG GGCCTTACAG
| TGCTAGGTT | TGCTAGGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGCTAGGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGCTAGGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGCTAGGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |
| TGCTAGGTT | TGAGTTTCGTT | CAAAATTAAG | ATTTTTCTGA | CGAAGACCTAG |

| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |
| TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT | TGGCTTTGGT |

| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |
| TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT | TAAGCAGTT |

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>TGGTTGGTTA CTCATTCAT</td>
<td>GGGGCACCTC TTCGAGGTAC ATCATCACCG</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
</tr>
<tr>
<td>Pocto 962</td>
<td>A</td>
</tr>
<tr>
<td>Pocto 1040</td>
<td></td>
</tr>
<tr>
<td>Pocto Richard</td>
<td></td>
</tr>
<tr>
<td>Pocto 1044</td>
<td></td>
</tr>
<tr>
<td>Pocto 965</td>
<td></td>
</tr>
<tr>
<td>Pocto2 Tam</td>
<td></td>
</tr>
<tr>
<td>Pocto15 Tam</td>
<td>T</td>
</tr>
<tr>
<td>Pexc Tam</td>
<td></td>
</tr>
<tr>
<td>Pexc 994</td>
<td></td>
</tr>
<tr>
<td>Pexc 1042</td>
<td></td>
</tr>
<tr>
<td>Pavi15 Tam</td>
<td></td>
</tr>
<tr>
<td>Pavi22 Tam</td>
<td></td>
</tr>
<tr>
<td>Pocci Tam</td>
<td>G</td>
</tr>
<tr>
<td>Pnot 1041 2</td>
<td></td>
</tr>
<tr>
<td>Pnot 1041 1</td>
<td></td>
</tr>
<tr>
<td>Photo22 Tam</td>
<td></td>
</tr>
<tr>
<td>Photo18 Tam</td>
<td></td>
</tr>
<tr>
<td>Ppuf 1018</td>
<td></td>
</tr>
<tr>
<td>Ppuff Tam</td>
<td></td>
</tr>
<tr>
<td>Ppuf 1012</td>
<td></td>
</tr>
<tr>
<td>Pflam Tam</td>
<td>T</td>
</tr>
<tr>
<td>Aortho cDNA</td>
<td>T</td>
</tr>
<tr>
<td>Csev Tam</td>
<td>A</td>
</tr>
<tr>
<td>Cher 1046</td>
<td></td>
</tr>
<tr>
<td>Cher 1047</td>
<td></td>
</tr>
<tr>
<td>Cfil 1045</td>
<td>T</td>
</tr>
<tr>
<td>Cfil 1024</td>
<td>T</td>
</tr>
<tr>
<td>Cher 995</td>
<td></td>
</tr>
<tr>
<td>Cobl 968</td>
<td></td>
</tr>
<tr>
<td>Cobl 967</td>
<td></td>
</tr>
<tr>
<td>Cfrat Tam</td>
<td>T</td>
</tr>
<tr>
<td>Cfrat 998</td>
<td>T</td>
</tr>
<tr>
<td>Cfil Tam</td>
<td>T</td>
</tr>
<tr>
<td>Coblman Tam</td>
<td>T</td>
</tr>
<tr>
<td>Coblith Tam</td>
<td>T</td>
</tr>
<tr>
<td>Cobl Tam</td>
<td>T</td>
</tr>
<tr>
<td>Cher R</td>
<td>T</td>
</tr>
<tr>
<td>Cobl R</td>
<td>T</td>
</tr>
<tr>
<td>Cobl 1032</td>
<td></td>
</tr>
<tr>
<td>Cobl Richard</td>
<td>T</td>
</tr>
<tr>
<td>Cher Tam</td>
<td>T</td>
</tr>
<tr>
<td>Cher Richard</td>
<td>T</td>
</tr>
</tbody>
</table>

FOOTNOTES
<p>|   | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| C | G   | G   | G   | C   | A   | A   | T   | G   | G   | C   | G   | C   | C   | A   | T   | A   | A   | G   | G   |
|   |     |     |     |     | A   |     |     | A   |     |     | A   |     | A   | A   |     |     |     |     |     |
|   |     |     |     |     |     |     | A   |     |     |     |     |     |     |     |     |     |     |     |     |
|   |     |     |     |     |     |     |     |     |     | A   |     |     |     |     |     |     |     |     |     |
|   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| A | A   |     |     |     | T   | A   | C   |     |     |     |     |     |     |     |     |     |     |     |     |
| A | A   | T   | T   | G   | T   | A   | A   | G   | C   | G   | A   | A   |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |
| A | A   | T   | G   | C   | A   | T   | C   | A   | A   |     |     |     |     |     |     |     |     |     |</p>
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>T</th>
<th>T</th>
<th>G</th>
<th>T</th>
<th>G</th>
<th>T</th>
<th>T</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>124</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>125</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>126</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>127</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>128</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>129</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>130</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>131</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>132</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
</tbody>
</table>


DeVries, H. 1906. *Species and Varieties, Their Origin by Mutation*. Lectures delivered at the University of California, Open Court, Chicago, IL.


is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). *Systematic Biology*, **46**, 269–283.


REFERENCES


REFERENCES


Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. *Biological Symposium, 6*, 74–125.


REFERENCES


REFERENCES


