Developing Genetic Markers for Stick Insects from Pyrosequencing Data

For a revised phylogeny of New Zealand stick insects (Phasmatodea)

Allan Wilson Centre Summer Studentship Report 2011/2012

by Geoffrey Thomson

Supervisor: Thomas Buckley

Introduction

New Zealand (NZ) is home to ten endemic genera of stick insect (Phasmatodea) in the Australasian clade Lanceocercata, with 23 described and several undescribed species (Buckley, et al. 2010; Buckley & Bradler 2010). They inhabit a wide range of environments from lowland forest to high alpine zones (Salmon 1991). Previous phylogenetic and biogeographic studies show that the NZ stick insect fauna originated in New Caledonia (NC) (Buckley, et al. 2009; Buckley, et al. 2010). This common origin is shared with several key parts of the New Zealand biota (Wallis & Trewick 2009). Similar species compositions between the two landmasses facilitates colonisation without significant adaptation to new host plants.

Previous Oceanic Lanceocercata phylogenies were constructed using mitochondrial cytochrome oxidase subunits I and II (COI/COII), nuclear large subunit ribosomal RNA and histone subunit 3 genes. These have indicated that that the NZ stick insect fauna is polyphyletic. One clade consists of only *Spinotectarchus* Salmon which split from its NC sister clade 29.9 mya (19.79-41.16 mya). The other clade consists of the nine remaining genera which separated 25.6 mya (16.8-35.5 mya) (Buckley, et al. 2010).

Biogeograpic studies of stick insects in the Oceania region has highlighted the need for lower level taxonomic revision (Buckley, et al. 2009; Buckley, et al. 2010). Polyphyly, synonymies and undescribed species are known to exist (Buckley, et al. 2009; Buckley, et al. 2010). To differentiate the genera further genetic markers are needed to construct a phylogeny with greater resolution.

Project Scope

This project developed and sequenced genetic markers for multiple nuclear and mitochondrial genes with the intention of contributing to a revised phylogeny of NZ stick insects.

Markers were developed from existing 454 transcriptome alignments obtained from *Micrachus* Carl. The candidates selected were longer than 1000bp and had high homology to known genes in the Genbank (Blastx, Blastn) and swissprot databases. Seventeen markers had primers designed for PCR amplification using the Primer design function of the Geneious program (Biomatters Ltd., v5.5.6, NZ). In addition to the 454 transcriptome alignmants from *Micrachus*, 454 transcriptome alignments from *Niveaphasma* Jewell & Brock and *Clitarchus* Stål along with a Genbank hit with high homology were used to ensure primers could cross amplify genera. Individual primers were designed to be 20-21bp long and contain as much of the marker region as possible. Primers for five markers were

(ordered from Sigma-Aldrich Co. LLC, USA): calreticulin, filamin-C, ezrin radixin moesin, opsin-1 and protein kinase shaggy.

These markers were sequenced from cDNA samples extracted from 58 individual stick insects; several from each NZ genus (except for the rare *Tepakiphasma* Buckley & Bradler, for which only one sample was available for RNA extraction). Outgroups from NC and Australia were also included. Previously developed markers were also amplified and sequenced, these were: COI/COII, elongation factor 2 (EF2) and succinate dehydrogenase (SDH). Initially COI/COII were sequenced together but lack of coverage on several samples necessitated the later use of internal primers to sequence them individually. Furthermore, new primers were designed for SDH midway through the project due to poor amplification of some samples.

PCR amplifications were performed using 25µl reaction volumes containing 0.5-1µl of cDNA, 2.5µl PCR Buffer with $MgCl_2$ (Roche, USA), 200µM each of dNTP, 1µl BSA(10 mg/ml), 1µl of each 10µM primer and 1.5 U of Fast start *Taq* DNA polymerase (Roche, USA). PCR conditions were 95°C for five minutes, then 38 cycles of 94°C for one minute, an annealing stage of one minute at a primer specific temperature (see Table 1) and an extension period of one minute at 72°C. When COI/COII were sequenced together the extension period was extended to one minute 30 seconds. Lastly there was a hold period at 72°C for ten minutes and the reactions were cooled to 4°C.

PCR products were visualised using gel electrophoresis to ensure there was no non-specific amplification and whether further PCR optimisation was required. All products were run on 1.5% w/v agarose gel in 1x TAE buffer, and stained with ethidium bromide. Successful products were then sequenced using BigDye®Terminator v3.1 (Invitrogen, USA). Sequencing reactions were purified using BigDye® XTerminator™ Purification Kit (Invitrogen, USA), and loaded into the Applied Biosystems 3100-Avant Genetic Analyzer. Sequence analysis was performed using the Geneious (Biomatters Ltd., v5.5.6, NZ) software platform.

During the course of this project experience was also gained in RNA extraction from dissected leg muscle using the RNeasy Mini Kit (Qiagen, DE). cDNA libraries were created by reverse transcribing the RNA using the Superscript® III First-Strand kit (Invitrogen, USA). Furthermore, field collections of specimens was also carried out in the Waitakere Ranges Regional Park, specifically the Huia region.

This project will contribute to constructing a revised phylogeny of the NZ stick insect genera along with several markers from other genes such as those in the glycolysis cycle. This will further illuminate the biogeographic origins, diversification and speciation of NZ stick insects.

Summary of main results

A total of 6,343 bp were sequenced for the majority of individuals sampled across all the markers that successfully amplified. With the exception of the marker opsin-1 which failed to amplify. The PCR primers designed are presented in Table 1.

Rudimentary neighbour joining trees were constructed using Geneious for each marker (e.g. Fig. 1) which possessed broadly similar characteristics. In each tree the Australian and NC species formed a disparate range of outgroups much more diverse than the subsequent NZ species. Furthermore the NZ species fell into two clades with the three *Spinotectarchus* individuals forming one clade sister to the NC species. The other clade, comprised of the rest of the New Zealand individuals, showed some

limited variation among the markers. In this clade the individuals in each genus mostly grouped together but the topology varied.

Nevertheless some associations and patterns can be observed. For example, in all but filamin-C the second NZ clade divides again into the *Micrachus* and all others. While only *Micrarchus hystricuelus* Westwood is currently described there are thought to be at least four species in this genus (Pers. Comm. T.R. Buckley). The *Micrachus* themselves divide into two groups; the *Micrachus sp*. alpine and the *Micrachus sp.* stephens with the *Micrachus sp.* kaikoura. In some markers such as calreticulin, filamin-C, COI/COII and EF2 *Micrarchus hystricuelus* is part of the *Micrachus sp.* alpine group while in the other markers it bridges the two groups.

Table 1: PCR primers and PCR thermo-cycling conditions used in this study. Highlighted primers were designed for this study. Opsin-1, in red, failed to amplify.

Gene target	Primer Name	Primer Sequence 5'-3'	Annealing temperature (°C)
Calreticulin	CRC_F	ACCTGGTGGTGCAGTTCACGG	58
Calreticulin	CRC_R	GGCGTGTCTCCTGCTTCTGCC	58
COI	C1-J-2195	TTGATTTTTTGGTCATCCAGAAGT	54
COI	TL2-N-3014	TCCAATGCACTAATCTGGCATATTA	54
COII	TL2-J-3034	AATATGGCAGATTAGTGCA	54
COII	TK-N-3785	GTTTAAGAGACCAGTACTTG	54
EF2	EF2-795-F	CCACCTAGCCGTCCCACGGA	60-62
EF2	EF2-1496-R	GTCGGCAGGGCAGAGAGCTG	60-62
Ezrin Radixin Moesin	MER_F	GCGCATCCTGGCCTTGTGCAT	60
Ezrin Radixin Moesin	MER_R	GCGCTTCGTGTTGCCCTTGC	60
Filamin-C	Fil_C_F	TCGTACCTTCCCACTGCCCCT	60
Filamin-C	Fil_C_R	TGGCCCACCTGCATGCACTC	60
Opsin-1	Ops_F	TATGGCACGGCCTCCTGGGG	-
Opsin-1	Ops_R	GTCACGCCAGAAGCGACGGA	-
Protein kinase shaggy	PKS_F	GCGAAGCTATGCGACACTGGGG	64
Protein kinase shaggy	PKS_R	CCGGGGTTGTTGCTCGCAGT	64
SDH	SDH-36-F	TTGCCCTTGCCCGTCTTGGC	66
SDH	SDH-1213-rev	CCTGGTCAACGGCCGGATCG	66
SDH	SDH_F2	CATGGGACACGAGGCGAGCG	60
SDH	SDH_R2	AGTTTGAAGTTGTGGGTGAT	60

Primers highlighted green and red designed by Geoffrey Thomson.
Primers highlighted blue designed by Thomas Buckley.
Primers not highlighted sourced from Buckley, et al., 2008.

In the other New Zealand genera, *Niveaphasma* appears as the least derived in calreticulin, ezrin radixin moesin and protein kinase shaggy but in COI/COII and EF2 *Clitarchus* is the least derived. Other notable patterns include *Niveaphasma* appearing alongside *Tectarchus* Salmon in almost all markers except calreticulin and protein kinase shaggy and *Clitarchus* being commonly associated with Pseudoclitarchus Salmon, *Argosarchus* Hutton and *Tepakiphasma*. *Acanthoxyla* Uvarov and *Asteliaphasma* Jewell & Brock show more variation in their placement within the trees than other genera.

Discussion

The results obtained support the previous research on the biogeographical origins and higher order taxonomy of NZ stick insects (Buckley, et al. 2010). The NZ stick insect fauna is polyphyletic with two clades, both nested within the NC radiation.

At a higher taxonomic level; while the variation within the markers is not substantial (0.005-0.2 substitutions per base pair) the individuals of each genus group together in each tree. This suggests, at least provisionally, that these markers are appropriate for determining the evolutionary relationships between genera. The consistent sub-groupings within the *Micrachus* genera support the hypothesis that *Micrachus* contains four species within the genera, despite only one being currently described.

The tentative associations observed between the nine of the monophyletic NZ genera are consistent with the current classifications of evolutionary relationships. Nevertheless these are far from comprehensively supported and there is a high level of variation overall in the relative position of each genus. The especially high variation in the placement of *Acanthoxyla* is not entirely unsurprising because of the likely hybrid origin of *Acanthoxyla* and their triploid genomes. Since they were only represented with two sequences heterozygotes sites were problematic (Buckley, et al. 2008).

It is known that the speed of genetic substitution varies across the genome (Wu, 2001) and that therefore one marker alone is unlikely to represent an accurate representation of the evolutionary history of a species. The variation observed between the neighbour-joining trees therefore demonstrates the need for using multiple markers when creating a molecular phylogeny at the genus or species level, as has been found elsewhere (e.g. Wortley, et al. 2005).

From this data it can therefore be concluded that the construction of a revised phylogeny of the New Zealand stick insect genera using multiple markers not only increases the statistical support for a phylogeny but is required to produce a phylogeny with any accuracy. It is thought that at least 10000 base pairs are required for a fully resolved phylogeny (Wortley, et al. 2005). The combination of this data with markers from other genes such as those in the glycolysis cycle surpasses this and will generate close to 15000 sequenced base pairs for each of the 58 individual stick insects. It will consequently be possible to produce a phylogeny to a high resolution. This work therefore contributes to the further illumination of the biogeographic origins, diversification and speciation of NZ stick insects.

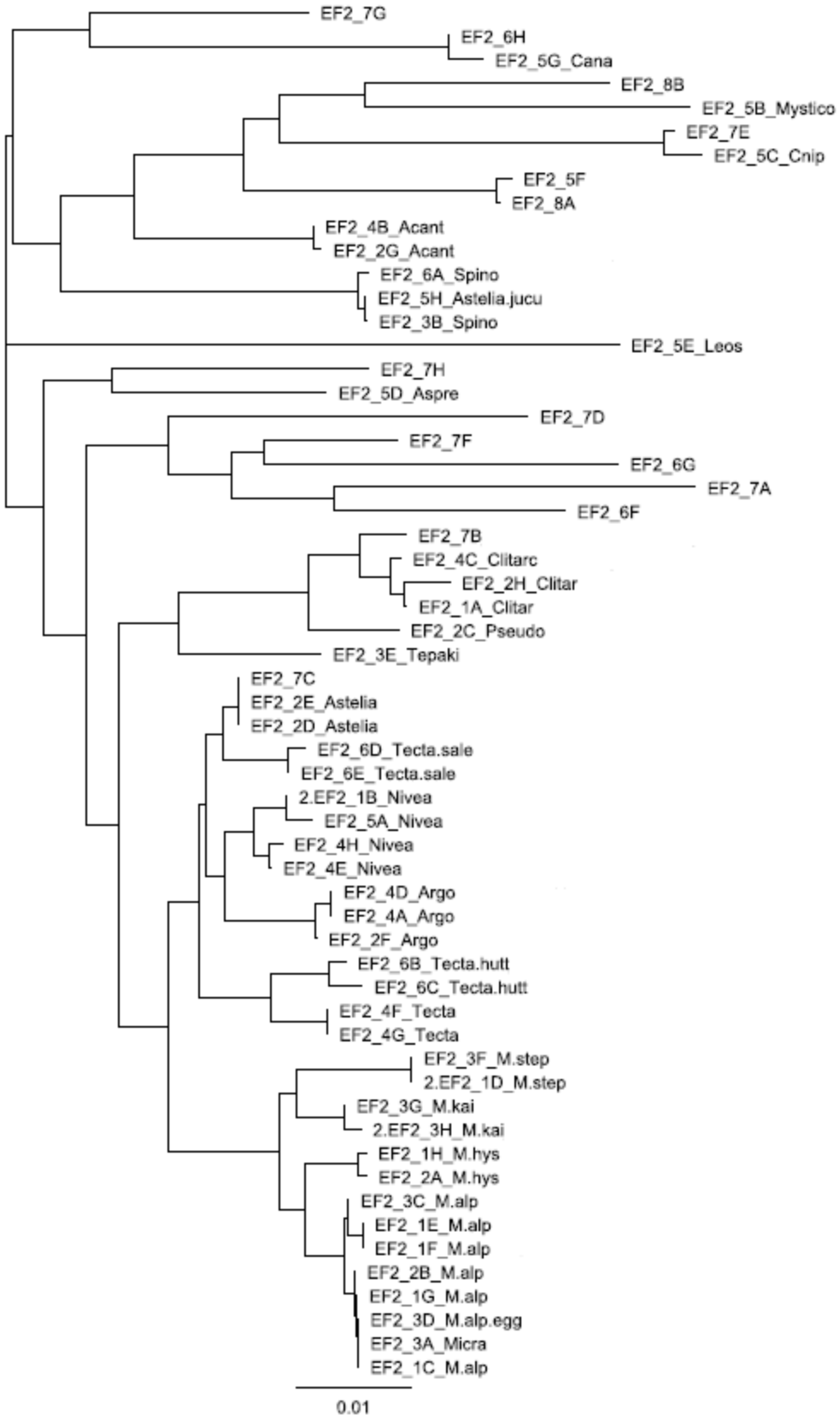

Figure 1: Rudimentary neighbour-joining tree for EF2 using default parameters and a Tamura-Nei genetic distance model with no outgroups, constructed using Geneious Tree Builder (Biomatters Ltd., v.5.5.6, NZ).

References

Buckley, T.R., D. Attanayake, D.-C. Park, S. Ravindran, T.R. Jewell, B.B. Normak (2008). Investigating hybridization in the parthenogenetic New Zealand stick insect *Acanthoxyla* (Phasmatodea) using single-copy nuclear loci. *Mol. Phylogen. Evol*, 48: 335–349. doi: 10.1016/j.ympev.2008.02.016,

Buckley, T.R., D. Attanayake, and S. Bradler (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B, 276: 1055-1062.* doi: 10.1098/rspb.2008.1552

Buckley, T.R., D. Attanayake, J.A.A. Nylander, and S. Bradler (2010). The phylogenetic placement and biogeographical origins of the New Zealand stick insects (Phasmatodea). *Systematic Entomology, 35: 207–225.* doi: 10.1111/j.1365-3113.2009.00505.x

Buckley, T.R., and S. Bradler (2010). *Tepakiphasma ngatikuri*, a new genus and species of stick insect (Phasmatodea) from the Far North of New Zealand**.** *New Zealand Entomologist, 33: 118-126.*

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious v5.5.6, Available from http://www.geneious.com

Salmon, J.T. (1991). *The Stick Insects of New Zealand*. Reed, Auckland.

Wallis, G. P. and Trewick, S. A. (2009), New Zealand phylogeography: evolution on a small continent. *Molecular Ecology*, 18: 3548–3580. doi: 10.1111/j.1365-294X.2009.04294.x

Wortley, A.H., Rudall, P.J., Harris, D.J., and Scotland, R.W. (2005). How much data are needed to resolve a difficult phylogeny? Case study in Lamiales. *Syst. Biol.* 54: 697–709. doi: 10.1080/10635150500221028

Wu, C.-I. (2001), The genic view of the process of speciation. *Journal of Evolutionary Biology*, 14: 851–865. doi: 10.1046/j.1420-9101.2001.00335.x