

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

New Zealand Passerines: a contribution to Passerine phylogeny



A thesis presented in partial fulfilment of the requirements for the Degree of

Master of Science (MSc) in Genetics

Massey University,

Palmerston North, New Zealand

RYAN ENGLAND

2013

Abstract

The passerines are the largest avian order, with over 6000 species. There is strong evidence to suggest that this group arose in Australasia, with most deep lineages located in Australia, New Zealand and Papua New Guinea. The deepest lineage is the New Zealand wrens, which diverged from the main passerine lineage, possibly around 80 mya. The second split between the suboscines (mainly new world distribution) and oscines (Australasian origin) is well established. Within the oscines there are a number of small basal lineages, all located within Australasia, then around 45 mya the large division into the Passerida and the Core Corvoidea occurred. The Core Corvoidea have undergone many rapid radiations early on in their history, which has made resolving the relationships within the group complicated. The Passerida are classified into three main superfamilies, but the relationships of these three are unresolved, and the monophyly of each has been questioned.

Next generation Illumina sequencing was used to sequence the mitochondrial genomes of six native passerine species. We report a sequenced mitochondrial genome from a representative of each New Zealand passerine family, apparently a first for any country. For four new species (the fernbird, tomtit, pipit and waxeye) the mitochondrial genomes have been fully sequenced while there is partial mitochondrial genome sequence for two other species (the browncreeper and bellbird). These have been combined with the mitochondrial genome sequences of another 72 passerines including seven previously unpublished genomes. Phylogenetic trees have been produced using both maximum likelihood and Bayesian analyses, and these have been used to address a number of questions surrounding the phylogeny of the passerines.

The placements of the native New Zealand species are highlighted, and in many cases it confirms the results of earlier studies. The results suggest consideration needs to be given for formally classifying the Petroicidae, Callaeidae and Notiomystidae as basal Passerida, but whether these three families form a separate monophyletic group is still unresolved. Part of the polytomy at the base of the Core Corvoidea has been resolved. The monophyly of each of the three Passerida superfamilies have been confirmed, but it is still unclear which of the three superfamilies branched off first, and there is no support for the suggestion that the Paridae are their own fourth superfamily. A need for

the formation of an extensive collection of tissue/DNA samples from New Zealand's vertebrates has been identified, and a number of suggestions for the use of mitochondrial genome sequence when studying passerine phylogeny have been made.

Acknowledgments

Firstly I thank my co-supervisors David Penny and Gillian Gibb. David I appreciate all the guidance and support over the last two years. Thank you Gill for originally showing me the ropes and then how to build trees. To Trish McLenachan thanks for all the help in the lab, organising samples and sequencing, and proofreading my work. To Bennet McComish for the assembly of the fernbird mitochondrial genome. Lesley Collins for teaching me the basics of bioinformatics.

Thanks to Lara Shepherd and Gillian Stone from Te Papa, Sandra Burles from the Wairapapa DoC office, and Phil Casey and John Ewen, for providing tissue and blood samples.

Thanks to the George Mason Scholarship and the Harwood Farm Trust for the financial support.

To my family and friends thank you for all the support over the six years I have been at Massey. I have made a lot of good friends and have some brilliant memories. To Mum and Dad thanks for always making sure we were well fed, the help with shifting flats, and all the years of making sure I pushed myself. I would not be the person I am today without you two. To Ethan for being my brother, friend and fellow front rower. You have always been there with me, either helping me out of or escalating the havoc we tend to get ourselves into. And to Alicia thank you so much, you are my best friend and I cannot wait to begin the next part of our journey together.

Table of Contents

Abstract	iii
Acknowledgments	v
Table of Contents	vi
List of Figures.....	viii
List of Tables.....	x
Abbreviations.....	xi
1 Introduction.....	1
1.1 Passerines	1
1.1.1 Origin	1
1.1.2 The New Zealand wrens and suboscines.....	2
1.1.3 The basal oscines	8
1.1.4 The Core Corvoidea	11
1.1.5 Basal Passerida?.....	13
1.1.6 The Passerida.....	14
1.2 New Zealand’s passerines	18
1.3 Use of molecular data in phylogeny	22
1.4 Project outline	26
1.5 Overview of results.....	27
2 Materials and Methods	29
2.1 DNA extraction	29
2.2 DNA amplification.....	29
2.3 Capillary sequencing and confirmation of species	30
2.4 Illumina Sequencing	30
2.5 Genome assembly	32
2.5.1 Pipit and tomtit.....	32
2.5.2 Fernbird	33
2.5.3 Waxeye, browncreeper and bellbird	34
2.5.4 Completing the mitochondrial genomes.....	35
2.5.5 Transcriptome data	36
2.6 Phylogenetic analysis.....	37
3 Results	39

3.1	DNA extraction and next generation sequencing sample preparation.	39
3.2	Illumina sequencing results	42
3.3	Genome assembly.....	48
3.3.1	Pipit	48
3.3.2	Tomtit.....	51
3.3.3	Fernbird.....	55
3.3.4	Waxeye.....	56
3.3.5	Brown creeper	59
3.3.6	Bellbird	61
3.4	Transcriptome reads	64
3.5	Phylogenetic analysis	64
4	Discussion.....	73
4.1	DNA extraction and Illumina sequencing.....	73
4.2	Phylogeny.....	75
4.3	Molecular data.....	82
4.4	Phylogeny an insight into physical characteristics.....	87
4.5	Conclusions	89
5	References	91
6	Appendices.....	99
6.1	Supplementary tables	99
6.2	Supplementary figures.....	102

List of Figures

<i>Figure 1: Simplified view of phylogeny of the passerines showing the main lineages in the order</i>	<i>3</i>
<i>Figure 2: Network of Bayesian analysis of the passerines, based on mitochondrial data.</i>	<i>10</i>
<i>Figure 3: The differing opinions on the phylogeny of the superfamilies of the Passerida.....</i>	<i>15</i>
<i>Figure 4: Electrophoresis of DNA extractions.</i>	<i>40</i>
<i>Figure 5: Fast Q screen results for the brown creeper, bellbird, and waxeye reads.....</i>	<i>46</i>
<i>Figure 6: Probability of base calling error along reads in fernbird mix.</i>	<i>47</i>
<i>Figure 7: Length of fernbird mix reads after quality trim.</i>	<i>47</i>
<i>Figure 8: Pipit Mitochondrial Genome.....</i>	<i>51</i>
<i>Figure 9: Tomtit mitochondrial genome.....</i>	<i>54</i>
<i>Figure 10: Fernbird mitochondrial genome.....</i>	<i>56</i>
<i>Figure 11: Waxeye mitochondrial genome.....</i>	<i>58</i>
<i>Figure 12: Brown creeper partial mitochondrial genome.</i>	<i>60</i>
<i>Figure 13: Bellbird and Brown creeper Blast2GO results from Velvet contigs</i>	<i>62</i>
<i>Figure 14: Bellbird partial mitochondrial genome.....</i>	<i>63</i>
<i>Figure 15: Maximum likelihood phylogenetic analysis tree from all 79 passerine species using mitochondrial genome data (13588bp).....</i>	<i>65</i>
<i>Figure 16: Bayesian analysis phylogenetic tree from all 79 passerine species using mitochondrial genome data (13588bp).....</i>	<i>66</i>

Supplementary Figures

<i>Supplementary Figure 1: Full Maximum Likelihood tree of 83 species.....</i>	<i>102</i>
<i>Supplementary Figure 2: Maximum Likelihood consensus network showing split of the three suborders of the Passerines.</i>	<i>103</i>
<i>Supplementary Figure 3: Maximum Likelihood consensus network of the Core Corvoidea.</i>	<i>103</i>
<i>Supplementary Figure 4: Maximum Likelihood phylogenetic tree from 76 passerine species.</i>	<i>104</i>
<i>Supplementary Figure 5: Maximum Likelihood consensus network showing split of the basal Passerida (Australasian robins, New Zealand wattlebirds and hihi) from the Passerida.</i>	<i>105</i>
<i>Supplementary Figure 6: Maximum Likelihood phylogenetic tree from all 63 Passerida species (including basal Passerida).</i>	<i>106</i>
<i>Supplementary Figure 7: Maximum Likelihood phylogenetic tree from 61 Passerida species (including basal Passerida).</i>	<i>107</i>
<i>Supplementary Figure 8: Maximum Likelihood consensus network of the three superfamilies of the Passerida.....</i>	<i>108</i>
<i>Supplementary Figure 9: Maximum Likelihood consensus network of the superfamily Sylvioidea..</i>	<i>109</i>
<i>Supplementary Figure 10: Maximum Likelihood consensus network of the superfamily Sylvioidea.</i>	<i>109</i>
<i>Supplementary Figure 11: Maximum Likelihood consensus network of the superfamily Muscicapoidea.</i>	<i>110</i>
<i>Supplementary Figure 12: Maximum Likelihood consensus network of the superfamily Passeroidea.</i>	<i>111</i>

List of Tables

<i>Table 1: Passerines with sequenced mitochondrial genomes.</i>	5
<i>Table 2: All native New Zealand Passerine populations and their conservation status.</i>	19
<i>Table 3: The species extracted for this study.</i>	40
<i>Table 4: Samples used for Illumina HiSeq sequencing at BGI.</i>	41
<i>Table 5: Samples used for Illumina MiSeq sequencing at NZGL.</i>	42
<i>Table 6: Amount of data received back from BGI and NZGL after Illumina sequencing.</i>	42
<i>Table 7: Fast QC quality scores of pipit, tomtit, waxeye, browncreeper and bellbird reads.</i>	44
<i>Table 8: K-mer sweep of the pipit reads.</i>	48
<i>Table 9: Pipit contigs using K-mers 67 bases in length.</i>	49
<i>Table 10: Pipit contigs using K-mers 41 bases in length.</i>	49
<i>Table 11: K-mer sweep of the tomtit reads.</i>	52
<i>Table 12: Tomtit contigs using K-mers 71 bases in length.</i>	52
<i>Table 13: K-mer sweep of the waxeye reads.</i>	57
<i>Table 14: K-mer sweep of the browncreeper reads.</i>	59
<i>Table 15: K-mer sweep of the bellbird reads.</i>	61
<i>Table 16: Mitochondrial genome genes extracted from the brain transcriptome reads.</i>	64

Supplementary Tables

<i>Supplementary Table 1: List of all primers used during study.</i>	99
<i>Supplementary Table 2: Estimation of DNA concentration using Nanodrop for the long range PCR products of the waxeye.</i>	101

Abbreviations

BGI	-	Beijing Genome Institute
BLAST	-	basic local alignment search tool
bp	-	base pairs
BWA	-	Burrows Wheeler Alignment
CI	-	Chatham Islands of New Zealand
DNA	-	Deoxyribonucleic acid
dNTP	-	Deoxyribonucleotide triphosphate
DoC	-	Department of Conservation (New Zealand)
kb	-	Kilo base (1000bp)
K/PG	-	Cretaceous Paleogene boundary
MGS	-	Massey University Genome Services
mtDNA	-	mitochondrial DNA
mya	-	million years ago
NI	-	North Island of New Zealand
NZ	-	New Zealand
PCR	-	Polymerase chain reaction
PNG	-	Papua New Guinea
RNA	-	Ribonucleic acid
rRNA	-	Ribosomal RNA
SapExo	-	Shrimp alkaline phosphatase and exonuclease
SI	-	South Island of New Zealand
tRNA	-	Transfer RNA

