

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.

GENETIC STRUCTURE OF
PLEUROBRANCHAEA MACULATA IN
NEW ZEALAND

A thesis presented in partial fulfilment of the
requirements for the degree of **Doctor of Philosophy (PhD)** in Genetics

The New Zealand Institute for Advanced Study
Massey University, Auckland, New Zealand

YEŞERİN YILDIRIM

2016

ACKNOWLEDGEMENTS

I have a long list of people to acknowledge, as my PhD project would not have been possible without their support.

Firstly, I would like to thank my supervisor, Professor Paul B. Rainey (New Zealand Institute of Advanced Study, Massey University), for providing me with the opportunity to join his research group. He gave me constant support, insightful guidance, valuable input, and showed me how to think like a scientist. I would also like to thank my co-supervisor, Dr Craig D. Millar, who welcomed me into his genetics laboratory at the Department of Biological Sciences at the University of Auckland whenever I needed help. He guided me patiently right from the beginning of my PhD project, was generous with his time, encouraging, and taught me how to troubleshoot where necessary. I would like to extend a special mention to Selina Patel, a very talented technician at Dr Millar's Lab who shared her extensive technical and theoretical knowledge with me generously, but also allocated considerable time to help me progress with my research. My study was kicked off by Selina Patel and Dr Millar's willingness to sequence my study organism's genome in the facilities of the "Centre for Genomics, Proteomics and Metabolomics" at the University of Auckland. Thank you also to my co-supervisor, Professor Marti Anderson (Massey University), for her guidance, her helpful contributions, for not only teaching me how to analyse my data but also for being involved in the analysis herself, and patiently answering my very naive questions about statistical methods. Her explanations helped me to understand that these very sophisticated methods are not something to be scared of – they are actually fun. Many thanks to my other co-supervisor, Professor Nigel French, (Massey University), for supporting me whenever I asked for help, and encouraging and backing me in the Allan Wilson Centre meetings.

My project was very much dependent on the divers and researchers who collected *Pleurobranchaea maculata* samples for me. I am grateful to Dr Susanna Wood and Dr David Taylor (Cawthron Institute, Nelson) for supplying me with numerous sea slug samples, sharing their extensive knowledge of *P. maculata*, discussing my findings, providing access to unpublished manuscripts, and introducing me to the people who

helped me to obtain more samples. I would have been lost without them. I would like to acknowledge the following people and/or institutes who either collected samples, put me into contact with people who provided samples or helped with the shipment of the samples: Mike McMurtry (Auckland Regional Council); Lauren Salvitti, Dudley Bell, Warrick Powrie and David Culliford (University of Waikato, Hamilton); Richard Huges and Bakhti Patel (Leigh Marine Laboratory, the University of Auckland); Severine Hannam and Dr Wilma Blom (Auckland Museum); Shane Genage (Victoria University of Wellington), Steve Journee (The Dive Guys, Wellington); Don Morrissey, Matthew Smiths and Stephen Brown (The National Institute of Water and Atmospheric Research); Dr Mark Stevens (South Australian Museum) and Te Papa Museum (Wellington).

Dr Paul McNabb (Cawthron Institute) conducted the tetrodotoxin assay for some of the samples. Serena Khor (The University of Waikato) extracted tetrodotoxin from some samples for the toxicity assay. I thank them both.

I would like to extend my gratitude to everyone who helped me with the data analyses and interpretation. I need to mention the contribution of my supervisors and Selina Patel again in this area. Virginia Moreno-Puig (Massey University), who has just submitted her thesis in which she utilised a population genetics tool, was a fount of knowledge and took a great deal of time to reflect on my questions with me. I would like to thank Dr Peter Ritchie (Victoria University of Wellington) and my colleagues Dr Stephen Ritchie, Jenny Herzog, Christina Straub, Dr Katrin Hammerschmidt, Dr Libby Liggins, Dr Peter Deines, Oliver Hannaford and Luca Bütikofer for their help and suggestions.

Many people helped me while I was writing my thesis. Many thanks once more to my main supervisor Professor Rainey for his patience while reading my thesis, his guidance, his outstanding comments and his corrections. I am immensely grateful to Dr Honour McCann (NZIAS, Massey University) as she critically reviewed, proofread and edited Chapter 1. I thank Dr Gayle Ferguson (INMS, Massey University) as she edited Chapter 2, but also Jacque Mackenzie (Institute of Veterinary, Animal & Biomedical Sciences, MU) as she proofread an earlier version of the same chapter. Thank you to Adam Bedford, Tatyana Pichugina (the University of Auckland), and my colleagues

Chhavi Chawla and Elena Colombi for editing and proofreading some parts of my thesis. I am very grateful to Joanna Niederer as well for copy-editing my whole thesis.

On a more personal level, I thank my mother Çakır Ayşe Selçuk and my brother Övgün Yıldırım, who have never failed to support me from the other side of the world. I know that I will always have a warm home waiting for me. I am very grateful to Adam Bedford, who encouraged me and gave me sound advice whenever I fell into despair. He was the one who suffered my whims, but he was also at my side during the many happy days with lots of laughs and adventures. I would also like to thank Rashmi Ramesh Iyer and Setareh Mokhtari, who were my first flatmates and my first friends in New Zealand. I would not have been able to survive in New Zealand without their support. I am grateful to Jenny Herzog for her friendship and cheerfulness, her wise observations and encouragement. My special thanks goes to one of the most wonderful and sophisticated people I have ever met, Marleen Bailing. She has always been there without hesitation whenever I needed help. She also showed me many tricks to improve my climbing skills. Thank you to Chhavi Chawla for being such a kind-hearted, sincere, wise and supportive friend. I wish you had joined our lab earlier. Thank you to my office mates and friends Yuriy Pichugin and Yunhao Liu. They always listened to my complaints, understood me very well and offered me practical advice and valuable opinions on my work. Thank you to all the Rainey Lab members as well as to Claire-Marie Loudon, Tatyana and Luca for making my life enjoyable. I have to stop here otherwise this list will never end.

My study was funded by the Allan Wilson Centre, Institute for Advanced Study at Massey University, and the Auckland Regional Council. I am very grateful to these institutes, as this research could not have happened without their support.

ABSTRACT

AIMS

The grey side-gilled sea slug (*Pleurobranchaea maculata*), which is native to the western and south Pacific, is known to contain high concentrations of tetrodotoxin (TTX). *P. maculata* populations around New Zealand exhibit individual, spatial and temporal differences in TTX concentration, but the origin of TTX in *P. maculata* is not fully understood. The main goal of my PhD project was to examine the genetic structure and demographic history of *P. maculata* populations from different regions in New Zealand and to clarify whether there is a correlation between variability in TTX concentrations and genetic structure.

METHODS

A sample of 146 *P. maculata* individuals were collected from three populations from the north-eastern North Island (Ti Point-TP, Auckland-AKL and Tauranga-TR), one population from the southern North Island (Wellington-WL) and one population from the northern South Island (Nelson-NL). TP, AKL and TR were designated the “Northern cluster”, whereas the WL and NL population were labelled as the “Southern cluster” due to the relative geographical locations of these clusters. Twelve nuclear microsatellite markers that were developed based on shotgun sequence were obtained from the genome of *P. maculata*. The markers were used to analyse the genetic structure of *P. maculata* populations. The mitochondrial cytochrome c oxidase subunit I (1153 bp) and cytochrome b (1060 bp) genes were also partially sequenced in *P. maculata* individuals.

RESULTS AND MAIN CONCLUSIONS

The microsatellite data reveal high genetic diversity and lead to the rejection of the hypothesis of panmixia: populations from the Northern cluster are highly connected but significantly differentiated from the Southern cluster. A weak differentiation was also observed between the WL and NL populations. The two populations correlate with regional variations in TTX concentrations: the Northern cluster populations contain highly toxic individuals, whereas the Southern cluster (WL and NL) populations

harbour either slightly toxic or non-toxic populations. The disjunction between the Northern and Southern clusters can be explained by biogeographical barriers specific to New Zealand but also with a stepping stone model. The geographical gap between the sampling locations made it impossible to draw firm conclusions as to the origin of the disjunction. The mtDNA sequence data reveal high haplotype diversity, low nucleotide diversity and a star-shaped haplotype network. These data can be explained by a population expansion dating back to the Pleistocene era. All the sampling locations are significantly differentiated from each other according to mtDNA data. Given that microsatellite and mitochondrial sequences evolve at different rates, incomplete linkage sorting is expected to be completed for mtDNA before, which should be reflected in a more pronounced structure for mtDNA markers where members of the populations have diverged recently. Although this may explain the geographical conflict between the microsatellite and mtDNA data, it is necessary to consider the possibility that the discordance between microsatellite markers and mtDNA may be in part attributable to the relatively small sample size.

TABLE OF CONTENTS

| | |
|--|----|
| CHAPTER 1: GENERAL INTRODUCTION | 1 |
| 1.1 <i>P. MACULATA</i> AND ITS ASSOCIATION WITH TTX | 4 |
| 1.1.1 Variability in TTX levels | 4 |
| 1.1.2 Ecological role of TTX in <i>P. maculata</i> | 4 |
| 1.1.3 Is the source of TTX dietary? | 5 |
| 1.1.4 Bacterial origin of TTX | 7 |
| 1.1.5 TTX resistance in animals | 9 |
| 1.2 POPULATION GENETICS: A TOOL TO STUDY POPULATIONS | 10 |
| 1.2.1 Factors shaping populations | 10 |
| 1.2.1.1 Mutation | 10 |
| 1.2.1.2 Natural selection | 11 |
| 1.2.1.3 Recombination and mating systems | 11 |
| 1.2.1.4 Genetic drift | 12 |
| 1.2.1.5 Migration | 13 |
| 1.2.2 Genetic markers | 13 |
| 1.2.3 Measurement of genetic variation | 14 |
| 1.3 COASTAL MARINE CONNECTIVITY IN NZ | 15 |
| 1.3.1 Isolation by distance | 17 |
| 1.3.2 Hydrodynamics | 17 |
| 1.3.3 Panmixia or localised structures | 19 |
| 1.3.4 Habitat requirement | 19 |
| 1.3.5 Developmental stage | 20 |
| 1.3.6 Patterns of sexual reproduction | 20 |
| 1.4 THESIS SUMMARY | 21 |
| 2 CHAPTER 2: DEVELOPMENT OF THE MICROSATELLITE MARKERS | 23 |
| 2.1 ABSTRACT | 23 |
| 2.2 INTRODUCTION | 23 |
| 2.3 MATERIALS AND METHODS | 25 |
| 2.3.1 Sampling, DNA extraction and genome sequencing | 25 |
| 2.3.2 Isolation and development of the primers | 26 |
| 2.3.3 Data analysis | 27 |

| | | |
|---------|--|----|
| 2.4 | RESULTS AND DISCUSSION | 27 |
| 3 | CHAPTER 3: MATERIALS AND METHODS | 32 |
| 3.1 | Sampling, DNA extraction | 32 |
| 3.2 | Tetrodotoxin assay | 34 |
| 3.3 | Genotyping | 34 |
| 3.3.1 | Microsatellite markers | 34 |
| 3.3.2 | mtDNA markers | 35 |
| 3.4 | MegaBlast analysis of the mtDNA sequences | 38 |
| 3.5 | Statistical analysis | 38 |
| 3.5.1 | Genetic diversity | 38 |
| 3.5.1.1 | Microsatellite data | 38 |
| 3.5.1.2 | mtDNA data | 40 |
| 3.5.2 | Genetic differentiation and population structure | 40 |
| 3.5.2.1 | Microsatellite data | 40 |
| | Exact differentiation test | 40 |
| | Fixation index and related estimators | 41 |
| | Power test | 41 |
| | Multivariate approach | 41 |
| | Population assignment with the Bayesian approach | 43 |
| | Analysis of molecular variance | 43 |
| 3.5.2.2 | mtDNA data | 44 |
| | Haplotype networks | 44 |
| | Multivariate analysis | 44 |
| | F-statistics | 45 |
| | Analysis of molecular variance | 45 |
| 3.5.3 | Migration | 45 |
| 3.5.4 | Spatial analysis | 46 |
| 3.5.5 | Demographic analysis, neutrality test and genealogy | 46 |
| 3.5.5.1 | Microsatellite data | 46 |
| 3.5.5.2 | mtDNA data | 47 |
| 4 | CHAPTER 4: GENETIC DIVERSITY AND STRUCTURE OF <i>P. MACULATA</i> POPULATIONS AS REVEALED BY MICROSATELLITE DATA | 48 |

| | | |
|---------|---|----|
| 4.1 | INTRODUCTION | 48 |
| 4.2 | RESULTS | 51 |
| 4.2.1 | Tetrodotoxin concentrations in the Wellington samples | 51 |
| 4.2.2 | Genetic diversity | 52 |
| 4.2.2.1 | Genetic diversity estimators | 52 |
| 4.2.2.2 | Analysis of the Hardy-Weinberg equilibrium | 54 |
| 4.2.2.3 | Investigation of null alleles | 54 |
| 4.2.3 | Linkage disequilibrium | 56 |
| 4.3 | Genetic differentiation and structure | 56 |
| 4.3.1 | Exact differentiation test | 56 |
| 4.3.2 | Fixation index and related estimators | 57 |
| 4.3.3 | Analysis of molecular variance | 60 |
| 4.3.4 | Power analysis | 61 |
| 4.3.5 | Individual distance matrices and the multivariate approach | 61 |
| 4.3.6 | Bayesian population assignment | 65 |
| 4.3.7 | Migration | 67 |
| 4.3.8 | Spatial analysis | 68 |
| 4.3.9 | Demographic analysis | 68 |
| 4.4 | DISCUSSION | 70 |
| 4.4.1 | Genetic diversity | 70 |
| 4.4.2 | Population differentiation | 71 |
| 4.4.3 | Biogeographical barriers | 72 |
| 4.4.4 | Pelagic larval duration | 74 |
| 4.4.5 | Correlation with toxicity | 74 |
| 5 | CHAPTER 5: DIVERSITY AND STRUCTURE OF <i>P. MACULATA</i> POPULATIONS AS REVEALED BY MITOCHONDRIAL DATA | 78 |
| 5.1 | INTRODUCTION | 78 |
| 5.2 | RESULTS | 80 |
| 5.2.1 | Identification of polymorphic regions | 80 |
| 5.2.2 | Genetic diversity | 80 |
| 5.2.3 | MegaBlast results | 82 |
| 5.2.4 | Genetic differentiation | 85 |

| | | |
|---------|---|-----|
| 5.2.4.1 | Haplotype networks | 85 |
| 5.2.4.2 | Individual distance matrices and the multivariate approach | 88 |
| | Abbreviations of the sampling locations are in parentheses: Whangarei (WH), Ti Point (TP), Auckland (AKL), Coromandel (CR), Tauranga (TR), Wellington (WL), Picton (PC), Nelson (NL) and Kaikoura (KK). | 92 |
| 5.2.4.3 | Differentiation based on F-statistics | 92 |
| 5.2.4.4 | Haplotype-based F_{ST} | 93 |
| 5.2.4.5 | Sequence-based Φ_{ST} | 94 |
| 5.2.4.6 | Molecular variance and the search for the north-south disjunction | 95 |
| 5.2.5 | Spatial analysis | 96 |
| 5.2.6 | Neutrality tests and demographic changes | 96 |
| 5.3 | DISCUSSION | 100 |
| 5.3.1 | Genetic diversity | 100 |
| 5.3.2 | Population structure | 100 |
| 5.3.3 | Incongruence between mtDNA and microsatellite data | 102 |
| 5.3.4 | A population expansion dating back to last glacial maxima | 103 |
| 5.3.5 | Human impact on the distribution of <i>P. maculata</i> | 107 |
| 6 | CHAPTER 6: CONCLUSION AND PERSPECTIVES | 109 |
| 7 | REFERENCES | 113 |
| 8 | APPENDIX ONE | 135 |
| 9 | APPENDIX TWO: SUPPLEMENTARY FIGURES AND TABLES | 139 |

List of Figures

| | |
|--|-----|
| Figure 1.1 Images of <i>P. maculata</i> | 1 |
| Figure 1.2 The localities where <i>P. maculata</i> has been spotted..... | 2 |
| Figure 1.3 The structure of TTX (Figure taken from Noguchi and Arakawa, 2008). | 3 |
| Figure 1.4 Map of NZ with major regional surface currents and biogeographic regions | 18 |
| Figure 3.1 Sampling locations for the <i>P. maculata</i> individuals..... | 33 |
| Figure 4.1 Rarefaction curves for the microsatellite data | 55 |
| Figure 4.2 MDS graph representing the pairwise R_{ST} distances between the five sampling locations..... | 59 |
| Figure 4.3 Ordination analyses of the multivariate genetic distances between the <i>P.</i> <i>maculata</i> individuals obtained from 12 microsatellite markers..... | 63 |
| Figure 4.4 Results of Bayesian clustering analysis where the sampling location was introduced for the calculations. | 66 |
| Figure 4.5 First-generation migrants in the <i>P. maculata</i> individuals from five sampling locations..... | 67 |
| Figure 4.6 Test of isolation by distance, based on the microsatellite data..... | 68 |
| Figure 5.1 Rarefaction curves for the Cytb sequences..... | 83 |
| Figure 5.2 Median joining network of the Cytb haplotypes. | 87 |
| Figure 5.3 Graphical visualisation of the standard nucleotide differences between individuals based on the mtDNA data. | 89 |
| Figure 5.4 The frequencies of Cytb haplotypes at each location | 93 |
| Figure 5.5 Test of isolation by distance, based on the Cytb data..... | 96 |
| Figure 5.6 Mismatch distributions of pairwise base pair differences between the <i>P.</i> <i>maculata</i> mtDNA haplotypes | 99 |
| Supplementary Figure 1 Allele frequencies for each population at each locus. | 152 |
| Supplementary Figure 2 Results of Bayesian clustering analysis where the sampling location has not been introduced as a priori for the calculations | 156 |
| Supplementary Figure 3 Median joining network of the COI haplotypes from 156 individuals | 174 |

| | |
|--|-----|
| Supplementary Figure 4 Median joining network of the concatenated COI haplotypes from 156 samples | 175 |
| Supplementary Figure 5 Median joining networks drawn for quality error control or saturation at the third codons | 176 |
| Supplementary Figure 6 The frequencies of the <i>P. maculata</i> COI haplotypes at each location..... | 177 |
| Supplementary Figure 7 Mismatch distribution of pairwise base pair differences between the concatenated COI and Cytb haplotypes..... | 178 |

List of Tables

| | |
|--|----|
| Table 2.1 PCR annealing temperature and elongation time for each primer pair..... | 27 |
| Table 2.2 Characteristics of 24 polymorphic microsatellite loci in <i>P. maculata</i> isolated using 454 high-throughput sequencing..... | 30 |
| Table 3.1 Sampling locations and abbreviations for the NZ <i>P. maculata</i> populations... | 32 |
| Table 3.2 PCR annealing temperature and elongation time for microsatellite markers. | 35 |
| Table 3.3 Primers used to amplify the mtDNA genes. | 37 |
| Table 4.1 Summary of the genetic diversity statistics at 12 microsatellite loci across five <i>P. maculata</i> populations..... | 53 |
| Table 4.2 Summary of the genetic diversity statistics for 146 <i>P. maculata</i> individuals sampled from five locations. | 54 |
| Table 4.3 Pairwise population differentiation estimates based on the microsatellite data. | 58 |
| Table 4.4 AMOVA results for the <i>P. maculata</i> populations based on the microsatellite data | 60 |
| Table 4.5 Results of the PERMANOVA analysis of the microsatellite data..... | 64 |
| Table 4.6 Assignment of 146 individuals to the a priori groups by CAP analysis using the microsatellite data | 66 |
| Table 4.7 Summary results of population expansion and decline analysis for the NZ <i>P. maculata</i> populations based on microsatellite data..... | 70 |
| Table 5.1 Summary genetic diversity statistics for the mtDNA data..... | 81 |
| Table 5.2 MegaBlast results for the COI and Cytb sequences of the <i>P. maculata</i> individuals | 84 |
| Table 5.3 Results of saturation tests for COI and Cytb sequences. | 88 |
| Table 5.4 Results of the statistical tests performed to compare the distribution of the genetic variation in mtDNA between five <i>P. maculata</i> populations | 91 |
| Table 5.5 Results of the cross-validation test for the COI data. CAP was used to calculate leave-one-out allocation rates of individuals to populations | 92 |
| Table 5.6 AMOVA results of concatenated sequences based on F and Θ -statistics..... | 95 |
| Table 5.7 Results of the neutrality and demographic tests for the mtDNA data | 98 |

| | |
|--|-----|
| Supplementary Table 1 Primer pairs for candidate microsatellite loci that were selected for amplification and polymorphism analysis..... | 139 |
| Supplementary Table 2 Summary statistics for the 12 microsatellites and 5 populations of <i>P. maculata</i> | 141 |
| Supplementary Table 3 Estimating the null allele frequency using the EN algorithm. | 143 |
| Supplementary Table 4 The estimation of the possible effect of null alleles on population structure analysis..... | 144 |
| Supplementary Table 5 Results of the linkage equilibrium test for 12 microsatellite markers..... | 147 |
| Supplementary Table 6 Allele frequencies at 12 microsatellite loci by population | 149 |
| Supplementary Table 7 Pairwise population matrix of the Dest values at each locus.. | 155 |
| Supplementary Table 8 Detection of first generation migrants | 157 |
| Supplementary Table 9 Partial mtDNA sequence data for one individual, 17B | 160 |
| Supplementary Table 10 Partial concatenated mitochondrial COI and Cytb alignment for 170 <i>P. maculata</i> haplotypes | 162 |
| Supplementary Table 11 Haplotype frequencies for the mtDNA genes by mere counting in populations..... | 171 |