GENETIC LINKAGE MAPS AND POPULATION GENETICS OF MACROPODS

KYALL RICHARD ZENGER

B.App.Sc. 1993 (UWS)
M.Sc. (Hons) 1996 (UWS)

Department of Biological Sciences,
Division of Environmental and Life Sciences.

November 2001

A thesis submitted to Macquarie University, in fulfilment of the regulations for the degree of Doctor of Philosophy.

Copyright in relation to this Thesis

ACKNOWLEDGMENTS

This project would not have been possible without the contribution of many people. I would like to thank those who provided financial support, samples and the like, and also the wonderful group of people who were my guides, colleges and companions:

- to my supervisors Professor Desmond Cooper for his constant support, enthusiasm, and valued advice, and Dr Jenny Donald for her encouragement and constructive criticism;
- to Professor Trevor Tansley, Dean of Graduate Students, Macquarie University and Professor D. Yerbury, Vice Chancellor, Macquarie University for financial support;
- to the many people who helped collect samples, especially Tony English, Graeme Coulson, Peter Johnson and Steve Cronin;
- to Dr Mark Eldridge and Dr Peter Johnston for reviewing many drafts of my thesis and providing much valued advice and support;
- to my many friends and colleges who have not only contributed immensely to this project through innovative ideas and problem solving, but who have also kept a constant smile on my face for the past few years – Adam Stow, Anne-Marie Vachot, Andrea Taylor, Cathy Adderton, Cushla Metcalfe, Julia Harlfinger, Kathy Belov, Lee Webley, Luciano Beheregaray, Mary Lam, Mathew Sloane, Paul Sunnucks and Valma Cahill;
- to Dr Bill Poole for locating the original tammar wallaby hybrid pedigree data buried within archives for more than ten years;
- to Michael Baxter for compiling and debugging many of the programs used throughout this thesis;
- to Dr Lisa Pope and Professor Craig Moritz for providing many of the microsatellite loci used in this study;
- to Dr Peter Sharp for providing a copy of his PhD thesis;
- and last but not least my family and loved ones, especially Sandra and Buddy for the constant support and encouragement, and putting up with me over the last few years through these tenuous times.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Ascertainment bias[1] – The hypothesis that a microsatellite selected in a focal species will differ systematically from its orthologues in related species due to the criteria used to isolate it in the focal species.

Cross species amplification – Primer binding sites sufficiently conserved in a related species allows amplification using primers designed in a different species.

Effective population size – The number of individuals that actually reproduce and contribute genes to the next generations.

Genic analysis – allele and / or haplotype frequencies are population statistics that can be changed by genetic drift, founder effect, gene flow and selection. They are estimated most accurately from multiple, separate, nuclear loci and mtDNA.

Genotypic analysis – consist of composite genotypes of multiple loci. Individual genotypes are labile – a single round of sexual recombination usually destroys a genotype. They are quantified most rigorously using multiple single-locus nuclear markers (typically microsatellites).

Gene synteny – Where loci are located on the same chromosome.

Homoplasy – Similarities in character states for reasons other than inheritance from a common ancestor. At a microsatellite locus, homoplasy occurs when two allelic lineages converge to the same size but have different histories of mutations. Thus, identity by state does not always entail identity by descent.

Likelihood ratio – The ratio of probabilities of obtaining the observed data under different hypotheses concerning the assumed model used to generate the expected values.

Linkage disequilibrium – Departure from the predicted frequencies of multiple locus gamete types assuming alleles of different loci are randomly associated.

Meiotic drive – Aberrant segregation ratios among the gametes of heterozygotes.

Monophyletic group – Set of species containing common ancestor and all its descendants.

Nucleolar organiser region – A region on a chromosome that contains the ribosomal RNA genes and associated spacers.

Null allele – An allele that fails to be expressed under the conditions analysed.

Orthologous loci – Loci in two or more species where sequences are similar because of their common derivation from a common ancestor.

Paraphyletic group – Set of species containing an ancestral species together with some, but not all, of its descendants. The species included in the group are those that have continued to resemble the ancestor; the excluded species have evolved rapidly and no longer resemble their ancestor.

Parsimony – Principle of phylogenetic reconstruction in which the phylogeny of a group of species is inferred to be the branching pattern requiring the smallest number of evolutionary changes.

Polyphyletic group – Set of species descended from more than one common ancestor. The ultimate common ancestor of all the species in the group is not a member of the polyphyletic group.

$\Delta T_m H$ – The difference between the median melting temperature of homoduplex DNA and heteroduplex DNA formed in a DNA-DNA hybridisation reaction.

[1] In the context of this thesis.

ABBREVIATIONS

AR	androgen receptor
CASA	alpha - casein
cDNA	complementary DNA
cM	centimorgan
DBB	MHC class II B-chain
ESUs	evolutionary significant units
ETL	economic trait locus
FISH	fluorescence *in situ* hybridisation
G6PD	glucose-6-phosphate dehydrogenase
GI	'Garden Island'
HBB	haemoglobin beta chain
H_E	expected heterozygosity
H_O	observed heterozygosity
HPRT	hypoxanthine phosphoribosyltransferase
ISHR	*in situ* hybridisation radiolabeled probes
KI	'Kangaroo Island'
LALBA	alpha - lactalbumin
LD	linkage disequilibrium
LLP	late lactation protein
LPL	lipoprotein lipase
MFL	mean fragment length
MHC	multiple histocompatability complex
MLE	maximum likelihood estimation
MP	maximum parsimony
MPI	mannose phosphate isomerase
mtDNA	mitochondrial DNA
MUs	management units
Mya	million years ago
N_e	effective population size
NJ	neighbour joining
NOR	nucleolar organiser region
NP	nucleoside phosphorylase
NSW	New South Wales
PGK1	phosphoglycerate kinase
PE	probability of exclusion
PIC	polymorphic information content
PCR	polymerase chain reaction
QLD	Queensland
QTL	quantitative trait locus
r	recombination fraction
RFLP	restriction fragment length polymorphism
RNR	ribosomal RNA
SDL	segregation distorting locus
SNPs	single nucleotide polymorphisms
SNuPE	single nucleotide primer extension analysis
SSCP	single strand conformation polymorphism
STRs	short tandem repeats
STRPs	short tandem repeat polymorphisms
SA	South Australia
TAS	Tasmania
TF	transferrin
tRNA	transfer RNA
VIC	Victoria

DECLARATION

I declare that this submission is my own work and that, to the best of my knowledge it contains no material written by another persons nor material which has been submitted for a higher degree to this or any other institution, except where due acknowledgement has been made in the text.

Kyall R. Zenger

November 2001

SUMMARY

The analysis of DNA using molecular techniques is an important tool for studies of evolutionary relationships, population genetics and genome organisation. The use of molecular markers within marsupials is primarily limited by their availability and success of amplification. Within this study, 77 macropodid type II microsatellite loci and two type I genetic markers were characterised within *M. eugenii* to evaluate polymorphic levels and cross-species amplification artifacts. Results indicated that 65 microsatellite loci amplified a single locus in *M. eugenii* with 44 exhibiting high levels of variability. The success of cross-species amplification of microsatellite loci was inversely proportional to the evolutionary distance between the macropod species. It is revealed that the majority of species within the Macropodidae are capable of using many of the available heterologous microsatellites. When comparing the degree of variability between source-species and *M. eugenii*, most were significantly higher within source species ($P < 0.05$). These differences were most likely caused by ascertainment bias in microsatellite selection for both length and purity.

The production of a marsupial genetic linkage map is perhaps one of the most important objectives in marsupial research. This study used a total of 353 informative meioses and 64 genetic markers to construct a framework genetic linkage map for *M. eugenii*. Nearly all markers (93.7%) formed a significant linkage (LOD > 3.0) with at least one other marker. More than 70% (828 cM) of the genome had been mapped when compared with chiasmata data. Nine linkage groups were identified, with all but one (LG7; X-linked) allocated to the autosomes. Theses groups ranged in size from 15.7 cM to 176.5 cM, and have an average distance of 16.2 cM between adjacent markers. Of the autosomal linkage groups, LG2 and LG3 were assigned to chromosome 1 and LG4 localised to chromosome 3 based on physical localisation of genes. Significant sex-specific distortions towards reduced female recombination rates were revealed in 22% of comparisons. Positive interference was observed within all the linkage groups analysed. When comparing the X-chromosome data to closely related species it is apparent that it is conserved both in synteny and gene order.

The investigation of population dynamics of eastern grey kangaroos has been limited to a few ecological studies. The present investigation provides analysis of mtDNA and microsatellite data to infer both historical and contemporary patterns of population structuring and dispersal. The average level of genetic variation across sample locations was exceedingly

high (h = 0.95, H_E = 0.82), and is one of the highest observed for marsupials. Contrary to ecological studies, both genic and genotypic analyses reveal weak genetic structure of populations where high levels of dispersal may be inferred up to 230 km. The movement of individuals was predominantly male-biased (average N_em = 22.61, average N_fm = 2.73). However, neither sex showed significant isolation by distance. On a continental scale, there was strong genetic differentiation and phylogeographic distinction between southern (TAS, VIC and NSW) and northern (QLD) Australian populations, indicating a current and / or historical restriction of geneflow. In addition, it is evident that northern populations are historically more recent, and were derived from a small number of southern eastern grey kangaroo founders. Phylogenetic comparisons between *M. g. giganteus* and *M. g. tasmaniensis*, indicated that the current taxonomic status of these subspecies should be revised as there was a lack of genetic differentiation between the populations sampled.