DISCOVERY, ANALYSIS, AND UTILITY OF BOVINE SINGLE NUCLEOTIDE POLYMORPHISMS

Gabrielle Simone Sellick

July, 2002

A thesis submitted in partial fulfilment of the award of Ph.D. in the Department of Animal Science at The University of Adelaide, Australia

TABLE OF CONTENTS

ABSTRACT

DNA sequence variation provides the fundamental material for improving livestock through selection. To date, few studies in cattle have concentrated on the isolation of single nucleotide polymorphisms (SNPs) to describe rates of nucleotide diversity across the bovine genome. In an attempt to determine the future value of SNP markers in livestock genomics, the screening of 97 bovine gene fragments amplified from a diverse panel of individuals was undertaken. Over 200 bovine SNPs were isolated from 70 known genes (1 SNP every 231bp), allowing the first large-scale analysis of SNP occurrence and comparison of the rates of bovine nucleotide polymorphism with other mammalian species. To further investigate the properties of bovine SNPs, pooled DNA sequencing analysis of 48 bovine gene fragments containing 187 SNPs were analyzed for minor allele frequency in four samples constructed from 10 common *Bos taurus* and *Bos indicus* cattle breeds. Analysis showed 28% of SNPs were common (≥ 20%) in all four pools providing vital evidence for the existence of informative bovine SNPs shared across many diverse cattle populations.

To clarify the level of assay flexibility, ease of implementation, and overall rates of genotyping success, two distinct SNP genotyping technologies based on either primer extension or primer hybridization were assessed. Analysis of different SNP base variants and in/dels coupled with the genotyping of over 1000 individuals per SNP clearly showed that primer extension-based SNP genotyping was a robust technology suitable for generating confident SNP genotypes. Furthermore, resultant SNP genotypes were used to highlight the successful application of SNPs in the genetic linkage mapping of type I markers for enhanced comparative map resolution and the development of candidate gene SNP haplotypes for linkage (QTL) and linkage disequilibrium mapping of 11 complex traits in cattle. Compelling evidence for a QTL on BTA2 affecting economically important carcass traits was found in the region harbouring the *MSTN* gene. Linkage disequilibrium analysis confirmed the association of *MSTN* haplotypes with total meat and fat percentage traits in Limousin cattle. In total, almost 290 bovine SNPs located within 70 known genes were described, resulting in a resource which will continue to provide key DNA markers for future comparative mapping strategies and association studies of candidate genes with complex traits in cattle.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

A	adenine
Abs	absorbance
AP	alkaline phosphatase
ASA	allele specific amplification
ASH	allele specific hybridization
BAC	bacterial artificial chromosome
bp	base pair
BSA	bovine serum albumin
BTA	*Bos taurus*
C	cytosine
°C	degrees Celcius
cm	centimetre
cM	centiMorgan
CpG	C and G dinucleotide pair (running 5' to 3')
cSNP	coding single nucleotide polymorphism
dATP	2'-deoxyadenosine 5'-triphophate
dCTP	2'-deoxycytosine 5'-triphophate
ddF	dideoxy fingerprinting
dGTP	2'-deoxyguanosine 5'-triphophate
dH_2O	distilled water
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded deoxyribonucleic acid
dTTP	2'-deoxythymidine 5'-triphophate
EDC	1-ethyl-3-(3-dimethyl-aminopropyl) carbodiimide hydrochloride
EDTA	ethylenediamine tetra acetic acid
EST	expressed sequence tag
EtBr	ethidium bromide
FISH	fluorescence *in situ* hybridisation
FITC	fluorescein isothiocyanate
fmol	fentamole
g	gram
G	guanine
hr	hour

HRP	horse radish peroxidase
HSA	*Homo sapiens*
in/del	insertion/deletion
kb	kilobase pair
L	litre
LB	Luria broth
LD	linkage disequilibrium
LOD	Logarithm of odds score
M	molar
MAS	marker assisted selection
mA	milliampere
Mb	million base pairs
MFI	mean fluorescence intensity
min	minute
mM	millimolar
mRNA	messenger ribonucleic acid
ng	nanograms
nm	nanometre
OAR	*Ovis aries*
OD	optical density
ODA	*N,N*-dimethyloctylamine
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PE	phycoerythrin
P_E	probability of exclusion
pers. comm.	personal communication
pg	picograms
pmol	picomole
pNPPD	p-Nitrophenyl phosphate solution
PPD11	ρ-phenylenediamine dihydrochloride, pH 11
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rpm	revolutions per minute
SA-HRP	streptavidin-conjugated horse radish peroxidase

SA-PE	streptavidin-R-phycoerythrin
SDS	sodium dodecyl sulphate
sec	second
SNP	single nucleotide polymorphism
SSC	sodium chloride sodium citrate
SSCP	single stranded conformational polymorphism
ssDNA	single stranded deoxyribonucleic acid
STR	short tandem repeat
STS	sequence tagged site
T	thymine
TAE	tris acetate ethylenediaminetetra-acetic acid
Taq	*Thermus aquaticus*
TBE	tris borate ethylenediaminetetra-acetic acid
TE	tris ethylenediaminetetra-acetic acid
TEMED	*N,N,N',N'*-tetramethylethylenediamine
T_m	melting temperature
TMAC	tetramethylammonium-chloride
TSA	tyramide signal amplification
µg	microgram
µl	microlitre
µM	micromolar
U	restriction endonuclease unit
UTR	untranslated region
UV	ultraviolet light
V	volts
VDA	variation detection array
v/v	volume for volume
w/v	weight for volume

DECLARATION

I certify that this thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

G. S. Sellick
23/7/02

DEDICATION

For my father, who always wanted to see me wear a "black cap".

Dad, I know you will always be there close, somewhere.

For my mother, whose enduring love, strength and support

has been forever inspirational.

ACKNOWLEDGMENTS

I would sincerely like to thank my supervisor Dr. Cynthia Bottema for her resolute support and encouragement throughout my candidature and for the vast opportunities presented to me to work, learn and experience the many different facets of science both nationally and internationally. I would also like to thank my co-supervisors, Professor Herman Raadsma and Dr. Wayne Pitchford for their helpful editorial comments on my thesis.

I am indebted to Orchid BioSciences Inc. for their contributions of reagents and materials for the primer-extension based SNP genotyping implemented in this project. Great thanks must also go to the J.S. Davies Bequest (University of Adelaide) and the Foundation for Research, Science and Technology (New Zealand) for access to the many thousands of microsatellite genotypes and phenotypes generated as part of the highly successful Adelaide/AgResearch beef cattle QTL mapping trial.

Thanks also to John Rader (GeneScreen, Dallas, TX), Jim Jacobson and Santosh Arcot (Luminex, Austin, TX), Mark Kunkel, Mark Pohl and Minh Xu (Orchid BioSciences, Princeton, NJ) for their hospitality and helpful discussions on the many intricacies of SNP genotyping assay development and analysis; Marilyn Jones and Dr. William Blackhall (University of Sydney) for extraction of the cattle DNA pools and verification of the pooled SNP allele frequency sequencing strategy; Dr. Christopher Smyle, Bozena Kruk and Marilyn Jones for their assistance with the pooled SNP sequencing; Dr. Graham Webb for his help with FISH gene assignments; Adam Kister for generating the combined genotype databases and guidance in using the programs implemented in the linkage/linkage disequilibrium studies; Dave Rutley for developing the saleable meat yield equations and all the staff and students of the animal molecular genetics group at the University of Adelaide and at AgResearch for their tireless participation in phenotypic data collection from the QTL mapping animals.

Importantly, this study would not have been the same without the past and present staff and students of the Department of Animal Science and the Department of Molecular

Biosciences, University of Adelaide especially Jan Cook and Helena Kojevnikoff who have allowed me to experience and learn so much more than they will ever know.

Finally, my most heart-felt thanks must go to my dearest friends (Adam, Andrew, Karen, Kathryn, Kirsten and Trudy) and family who have stuck by me through it all! Words can not describe how much your encouragement, love, support and guidance have meant to me. The endless phone calls and e-mail messages, you have made me whom I am today.

Thankyou with all my heart.