Genomic approach to understanding variation in bovine fat colour

Presented by

Rugang Tian

A thesis submitted to the University of Adelaide in fulfilment
of the requirement of the degree of
PhD of Science

THE UNIVERSITY
of ADELAIDE

The University of Adelaide
School of Animal and Veterinary Sciences
Discipline of Animal Science
September 2011

Table of Contents

Index of Figures vi

Index of Tables ix

Index of Appendices xi

Declaration xiii

Dedications xiv

Acknowledgements xv

Abstract xviii

Chapter 1 Literature Review 1

1.1 Introduction 2
1.2 Fat colour and market specification 2
1.3 Factors affecting fat colour 3
1.3.1 Non-genetic factors 3
1.3.1.1 Diet 3
1.3.1.2 Season 6
1.3.1.3 Management strategies 6
1.3.1.4 Age 8
1.3.1.5 Sex 8
1.3.1.6 Growth rate 9
1.3.2 Genetic factors 10
1.4 Carotenoids and fat colour 13
1.4.1 Carotenoid function 13
1.4.2 Ruminal digestion 15
1.4.3 Carotenoid absorption 16
1.4.4 Carotenoid transportation 18
1.4.5 Carotenoid deposition 19
1.4.6 Enzymes in metabolism of β-carotene 20
1.4.6.1 β-carotene cleavage enzymes 20
1.4.6.2 Other enzymes 27
1.5 Summary 30

Chapter 2 General methods and materials 32

2.1 Animal resources 33
2.2 Phenotypic data collection 35
2.3 DNA preparation 37
2.4 Candidate gene selection 38
2.5 Single nucleotide polymorphisms experiments 38
2.5.1 Primer design and selection 38
2.5.2 Optimization of PCR conditions 39
2.5.3 Automated cycle sequencing of PCR products 41
2.5.4 SNP discovery 42
2.5.5 Genotyping 43
2.5.5.1 Genotyping by PCR-RFLP 43
2.5.5.2 Genotyping by High Resolution Melting (HRM) 44
2.6 Real-time PCR 45
2.6.1 Tissue collection 45
2.6.2 RNA isolation 45
2.6.3 cDNA production 46
2.6.4 Real-time PCR conditions 46
2.7 Statistical analysis 47
2.7.1 Animal Group 1: 47
2.7.2 Groups 2 to 6 : 51
2.7.3 Group 7: 52

Chapter 3 Candidate gene selection 54

3.1 Introduction 55
3.2 Methods 59
3.3 Results 60
3.3.1 Candidate gene selection 60
3.3.2 SNP identification and selection for association studies 60
3.3.3 Genotype and allele frequencies 64
3.4 Discussion: 67

Chapter 4 Association between candidate gene SNPs and fat colour traits 73

4.1 Introduction 74
4.2 Materials and methods 74
4.3 Results 76
4.3.1 Cohort, breed, and sire effects 76
4.3.2 Individual SNP genotype effects 79

4.3.3 SNP additive and dominance effects 83
4.3.4 Non-candidate gene SNP associations 85
4.3.5 *BCO2* W80X effects on significance of other SNPs 87
4.3.6 QTL analysis 89
4.3.6.1 BCO2 W80X effects on QTL 91
4.3.6.2 ALDH8A1 SNP16 effects on QTL 93
4.3.6.3 PPARGC1A SNP12 effects on QTL 95
4.3.6.4 RDHE2 SNP2 effects on QTL 98
4.4 Discussion 101
4.4.1 Breed effects 101
4.4.2 Multiple and individual SNP testing 104
4.4.3 Non-coding polymorphisms 104
4.4.4 Functional significance of candidate genes 105
4.4.4.1 ALDH8A1 105
4.4.4.2 APOM 106
4.4.4.3 RARA 107
4.4.4.4 RDHE2 107
4.4.4.5 PPARGC1A 108
4.4.4.6 SCARB1 109
4.4.4.7 BCMO1 110
4.4.4.8 BCO2 111
4.4.4.9 Other non-candidate genes 112
4.4.5 Additional SNP effects on QTL 113

Chapter 5 Interactions between candidate gene SNPs 115

5.1 Introduction 116
5.2 Materials and methods 116
5.3 Results 117
5.3.1 Interaction effects between SNPs within genes 117
5.3.2 Haplotype associations 119
5.3.3 Epistatic interactions of major SNPs 122
5.4 Discussion 125
5.4.1 Haplotype effects 125
5.4.2 Epistatic effects 126

Chapter 6 *BCO2* and *RDHE2* gene validation studies 128

6.1 Introduction 129

6.2 Materials and methods 129
6.2.1 Animal resources 129
6.2.2 Statistical analysis 130
6.3 Results 130
6.3.1 *BCO2* validation 130
6.3.1.1 Australian Jersey dams 130
6.3.1.2 Validation data from New Zealand 131
6.3.2 *RDHE2* validation 133
6.3.3 Interaction effects between *RDHE2* and *BCO2* 133
6.4 Discussion 135
6.4.1 *BCO2* effects 135
6.4.2 *RDHE2* effects 136

Chapter 7 Gene expression studies 139

7.1 Introduction 140
7.2 Materials and methods 141
7.2.1 Animal resources and methods 141
7.2.2 Statistical analysis 141
7.3 Results 142
7.3.1 RNA extraction 142
7.3.2 Optimization of PCR 142
7.3.3 Evaluation of reference genes 143
7.3.4 Gene expression analysis 144
7.3.4.1 Gene expression and fat colour 144
7.3.4.2 Genotype effects on gene expression 146
7.3.4.3 Gene expression effects on β-carotene concentration 149
7.3.4.4 Relationships between relative gene expression levels 149
7.4 Discussion 151
7.4.1 *BCMO1* gene expression analysis 151
7.4.2 *BCO2* gene expression analysis 152
7.4.3 *RDHE2* gene expression analysis 157
7.5 Conclusion 158

Chapter 8 General Discussion 159

8.1 Introduction 160
8.2 SNP association studies 160
8.3 Non-synonymous SNP effects 163

8.4 Candidate pathways 168
8.4.1 Vitamin A synthesis via BCMO1 pathway 170
8.4.1.1 BCMO1 direct effects 170
8.4.1.2 Regulation of BCMO1 by transcription factors 171
8.4.1.3 Regulation of BCMO1 by negative feedback mechanisms 173
8.4.2 Vitamin A synthesis via BCO2 pathway 176
8.4.2.1 Asymmetric cleavage of β-carotene 176
8.4.2.2 Asymmetric cleavage of lutein 177
8.4.3 Absorption and transportation of carotenoids 178
8.4.4 Dilution of carotenoids in the fat 182
8.5 Epistasis 187
8.6 Future studies 189
8.7 Conclusion 191

Appendices 195

References 225

Index of Figures

Figure 1-1 β-carotene and lutein content of forages, cereals and root crops used in cattle diets 5
Figure 1-2 Three forms of vitamin A 14
Figure 1-3 Schematic overview of β-carotene metabolism 22
Figure 2-1 Backcross design of the Davies Gene Mapping Herd 34
Figure 2-2 AUS-MEAT standard colour chips 37
Figure 4-2 Effects of cohort, breed and sire on biopsy sample β-carotene concentration 78
Figure 4-1 Effects of cohort, breed and sire on carcass samples fat colour score .. 78
Figure 4-3 Effects of cohort, breed and sire on biopsy samples fat colour score .. 78
Figure 4-4 BCO2 W80X effects on BTA 15 QTL across families 92
Figure 4-5 *BCO2* W80X effects on BTA 15 QTL for sire 398 family 92
Figure 4-6 *ALDH8A1* SNP16 effect on BTA 9 QTL across families 94
Figure 4-7 *ALDH8A1* SNP16 effect on BTA 9 QTL for sire 368 family 94
Figure 4-8 BTA 6 QTL across families without *PPARGC1A* SNP12 in the model 97
Figure 4-9 BTA 6 QTL across families with *PPARGC1A* SNP12 in the model 97
Figure 4-10 *PPARGC1A* SNP12 effect on BTA 6 QTL in sire 361 family 97
Figure 4-11 *RDHE2* SNP2 effect on BTA 14 QTL across families 99
Figure 4-12 *RDHE2* SNP2 effect on BTA 14 QTL in sire 361 family 99
Figure 4-13 RDHE2 SNP3 effect on BTA 14 QTL across families 100
Figure 4-14 RDHE2 SNP3 effect on BTA 14 QTL in sire 361 family 100
Figure 5-1 Interaction between *PPARGC1A* SNP 12 and *BCO2* W80X on β-carotene concentration 124
Figure 5-2 Interaction between *PPARGC1A* SNP 12 and SCARB SNP1 on biopsy fat colour 124
Figure 5-3 Interaction between *PPARGC1A* SNP 12 and RDHE2 SNP3 on carcass fat colour 124
Figure 7-1 Average expression stability values (M) of the reference genes with stepwise exclusion of the least stable expressed reference genes 144
Figure 7-2 *BCMO1* mRNA expression in the two fat colour groups 145
Figure 7-3 *BCO2* mRNA expression in the two fat colour groups 145

Figure 7-4 *RDHE2* mRNA expression in the two fat colour groups..... 145
Figure 7-5 Relative normalised *BCMO1* gene mRNA levels for different gene genotypes..... 147
Figure 7-6 Relative normalised *BCO2* gene mRNA levels for different gene genotypes..... 147
Figure 7-7 Relative normalised *RDHE2* gene mRNA levels for different gene genotypes..... 148
Figure 7-8 Correlation between *BCMO1* mRNA expression and *BCO2* mRNA expression..... 150
Figure 7-9 Correlation between *BCMO1* mRNA expression and *RDHE2* mRNA expression 150
Figure 7-10 Correlation between *BCO2* mRNA expression and *RDHE2* mRNA expression 150
Figure 7-11 Regression of *BCO2* mRNA level and beta carotene concentration 155
Figure 7-12 Regression of *BCO2* mRNA level and log beta carotene concentration 155
Figure 8-1 ALDH8A1 protein tertiary structure as predicted by the $(PS)^2$ program 165
Figure 8-2 ALDH8A1 protein secondary structure as predicted by the $(PS)^2$ program 165
Figure 8-3 PPARGC1A protein tertiary structure as predicted by the $(PS)^2$ program 166
Figure 8-4 PPARGC1A protein secondary structure as predicted by the $(PS)^2$ program 166
Figure 8-5 RDHE2 protein tertiary structure as predicted by the $(PS)^2$ program. 167
Figure 8-6 RDHE2 protein secondary structure as predicted by the $(PS)^2$ program 167
Figure 8-7 Three forms of vitamin A 168
Figure 8-8 Retinoic acid formation from β-carotene and retinol 169
Figure 8-9 Putative transcription factor binding sites in the promoter region of *BCMO1* 173
Figure 8-10 Crosstalk between RAR and ISX signalling controls β-carotene absorption..... 175

Figure 8-11 Correlation between beta-carotene concentration and fat depth..... 184
Figure 8-12 Correlation between biopsy fat colour and fat depth..... 184
Figure 8-13 Correlation between carcass fat colour and fat depth..... 184
Figure 8-14 Interactions between genes..... 188

Index of Tables

Table 2-1 PCR reagent concentrations for *AmpliTaq Gold* DNA polymerase 39
Table 2-2 PCR reagent concentrations for *Kapa* DNA polymerase 40
Table 3-1 QTL detected by interval mapping for fat colour traits 57
Table 3-2 QTL detected by multiple-trait multiple marker for fat colour traits 58
Table 3-3 QTL on BTA and orthologous regions in human genomes and putative candidate genes 62
Table 3-4 DNA variants of the candidate genes 63
Table 3-5 Genotyped DNA variants 63
Table 3-6 Genotypic and allelic frequencies of SNPs 65
Table 4-1 Least squares means with standard errors for cohort, breed, and sire 77
Table 4-2 Association of candidate gene SNPs with fat colour traits 80
Table 4-3 SNP size of effect on fat colour traits as percentage of total SNP variation (%) 82
Table 4-4 Additive and dominance effects for significant trait association 84
Table 4-5 Additional gene SNP variance as percentage of total SNP variance (%) 86
Table 4-6 Comparison of individual SNP significance P values when *BCO2* W80X was excluded or included in the model for fat colour traits in Group 1 animals 88
Table 4-7 Major SNP effects on QTL across sire families 90
Table 4-8 Major SNP effects on QTL for individual sire families 90
Table 4-9 SNP allele frequency in XL and XJ backcross 103
Table 5-1 Interaction effects between SNPs within genes 118
Table 5-2 Diplotype association with significance and trait means 120
Table 5-3 Size of interaction effects between SNPs as percentage of total SNP variation (%) 123
Table 6-1 Means (SE) for year, BCO2 W80X genotype, additive and dominant effects in Australian Jersey dams 131
Table 6-2 Substitution effects of *BCO2* W80X alleles for fat colour traits in the New Zealand cattle 132
Table 6-3 Effects of RDHE2 SNPs on fat colour traits 134
Table 6-4 RDHE2 allele frequencies in experimental cattle populations 137
Table 7-1 Optimal annealing temperature for primer pairs 142

Table 7-2 M values of reference genes.....143
Table 7-3 Relative normalised gene mRNA expression[a] 144
Table 7-4 Relative normalised gene mRNA levels among *RDHE2* SNP2 genotypes and SNP *BCO2* W80X genotypes 146
Table 7-5 Regression co-efficients (slope of linear regression analysis) for the relationship between β-carotene concentration and relative gene expression level.....149
Table 8-1 Proportion of the sum of squares accounted for all SNP additive effects.162

Index of Appendices

Appendix 1: Protocols 196

Appendix 1.1: Extraction and quantification of carotenoids 196

Appendix 1.2: DNA Purification 197

Appendix 1.3: Purification of PCR products for sequencing 198

Appendix 1.4: Ambion TURBO DNA-*free*™ DNase Treatment 198

Appendix 1.5: TRI Reagent® DNA Isolation 199

Appendix 2: Chemicals and solutions 200

Appendix2.1: TE Buffer 200

Appendix 2.2: TAE buffer 200

Appendix 2.3: Formamide Loading Buffer 200

Appendix 2.4: 1M Tris-HCl 200

Appendix 3: PCR conditions 201

Appendix 3.1: PCR conditions for the candidate genes 202

Appendix 3.2: PCR conditions for the genotyped SNPs 206

Appendix 3.3: Real-time PCR conditions 208

Appendix 4: Identified DNA variants 209

Appendix 5: Candidate gene haplotypes 212

Appendix 6: Effects of candidate gene SNPs on fat colour traits with covariates 213

Appendix 7 QTL 214

Appendix 7.1: Most likely position and F-statistic values of putative QTL detected by individual family analysis (Sire 361 family) 214

Appendix 7.2: Most likely position and F-statistic values of putative QTL detected by individual family analysis (Sire 368 family) 216

Appendix 7.3: Most likely position and F-statistic values of putative QTL detected by individual family analysis (Sire 398 family) 218

Appendix 7.4: Position and F-statistic values of putative QTL detected by across family analysis 220

Appendix 8: Sequence alignment 222

Appendix 8.1: *RDHE2* and *RDHE2* similar genes from bovine and human 222

Appendix 8.2: Pairwise alignments score for bovine, human *RDHE2* and *RDHE2 similar* genes 222

Appendix 8.3: Bovine, human *RDHE2* and *RDHE2 similar* genes amino acid sequence alignments 223

Appendix 9: RNA quality 224

Declaration

I declare that this thesis is a record of original work and contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Rugang Tian except where as stated in chapter 3, some sequencing (*BCO2* gene) was completed as part of a research project for a Masters by Coursework degree at the University of Adelaide in 2006. To the best of my knowledge and belief, this thesis contains no material previously published or written by any other person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University‘s digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Rugang Tian

September, 2011

Dedications

I dedicate this work to my parents and my wife, Yuan Li, for their great support with love during the period of this study.

Acknowledgements

During my PhD study, I have learnt a lot from many people. It is with great pleasure that I acknowledge the assistance provided by people who supported and helped me to complete this thesis.

I would sincerely like to thank my supervisors, Dr. Cynthia D.K. Bottema and Assoc. Prof. Wayne Pitchford for their excellent supervision, guidance, trust and criticism over the course of my years of study. In particular, I would like to thank Dr. Bottema for encouraging me to continue the research journey. Dr Bottema is an excellent model as a multi-skilled, very knowledgeable and kind scientist, who led me to become a researcher and encouraged me to always strive to improve the quality of my project. Without her help, my life would be in a different path.

I am very grateful for the help from Assoc. Prof. Wayne. He was always willing to understand, discuss, listen, and encourage. The discussions on the analytical models were very beneficial in my understanding the complicated statistics. From his critical attitude, I gradually learned to criticise the research statistically.

To the laboratory members and my friends, Andrew Egarr, Irida Novianti, Nadiatur Akmar Zulkifli, David Lines, Leiyao Chang, Ruidong Xiang and Dr Graham Webb. Thank you all for the wonderful time we shared. My time in the laboratory would be boring without all your support, friendship and sharing the stories. I will always value your friendship and will remember the great and joyful times we have had forever.

Many thanks are extended to our collaborators, Chris Morris and Neil Cullen, for their useful comments on manuscripts and for their kind assistance with the animal sampling in New Zealand abattoirs. Thanks to Dr. Madan Naik for his assistance in learning various techniques in the molecular laboratory and his valuable advice during the course of my study. I would also like to acknowledge Dr. Zbigniew Kruk for his excellent pioneering work. The experience in the abattoir with Dr. Kruk was very valuable to help me to do the animal sampling in the abattoir by myself. My gratitude goes to Dr. Brian Siebert who dedicated much time and effort teaching me the techniques of β-carotene extraction.

Great thanks must go to the J.S. Davies Bequest for funding the cattle gene mapping project in Australia. My sincere thanks also extend to AgResearch in New Zealand for access to their data. They made a great contribution by providing the phenotype and genotype data for Groups 2, 4, 5 and 6. I also extend my sincere gratitude to the University of Adelaide for providing the PhD scholarship, without which this work was not possible.

My deepest gratitude goes to my mum and dad for being generous supporters over the past years. It was your love, support and strength that got me through the toughest time. Thank you for always being proud of my achievements. The moment when I finished typing the last word for my thesis, I can hear your voice, I can see your smile, and I can feel your love from thousands of miles away. This long distance cannot keep my love from you.

Most importantly, my special thanks must be given to my dear wife Yuan, who has

constantly motivated me and encouraged me in every way. I will never forget that you sacrificed your sleeping time to prepare breakfast for me every morning with your sleepy eyes. Thank for all your support even during the most emotionally challenging time. It is your understanding and your love that enabled me to complete this work.

Abstract

Subcutaneous fat is important not only in the live animal but also in the carcass, as it prevents the rapid chilling of the underlying muscle tissues, thereby reducing weight loss during chilling. However, beef with yellow fat is considered undesirable by consumers in most European and Asian markets. Beta-carotene is the major carotenoid deposited in adipose tissue, which results in the yellowness. Genes involved in the metabolism of β-carotene in the cattle are likely to regulate beef fat colour. Therefore, DNA variants in candidate genes related to β-carotene metabolism were examined for association with beef fat colour.

Based on their location in fat colour quantitative trait loci (QTL) and function in the metabolism of β-carotene, *ALDH8A1*, *APOM*, *BCMO1*, *BCO2*, *RARA*, *RDHE2*, *PPARGC1A* and *SCARB1* were chosen as candidate genes. One hundred eleven (111) DNA variants were identified from the direct sequencing of 3 F1 sires for these eight genes, of which, 27 DNA single nucleotide polymorphisms (SNPs) were selected for association studies (3-5 SNPs per gene). Most of these genotyped SNPs and their interactions were associated with fat colour related traits (biopsy fat colour (Fc-bio), carcass fat colour (Fc-car), beta-carotene concentration (Bc-bio)), although the size of the effects was relatively low for many of the variants. However, among the DNA variants, a nonsense mutation in the *BCO2* gene (*BCO2* W80X) accounted for a large proportion (12-16%) of the total SNP variation in fat colour related traits in Jersey-Limousin backcross progeny. Validation of this SNP in other independent herds (Group 2-7) confirmed the *BCO2* W80X genotype has a large effect on beef fat colour and milk colour. The individual genotypic effects of *RDHE2* SNP2 and SNP3 were also large.

However, these effects were greater in the New Zealand abattoir samples than from pedigreed Jersey-Limousin backcross progeny, amounting to 8-17% of the variance in one population. There was a significant interaction between the *BCO2* W80X and the *RDHE2* SNP2, which accounted for 1.8% of the total SNP variance in milk fat colour in a New Zealand Holstein cow population, and 4.0% of the total SNP variance in carcass fat colour in New Zealand Jersey-Limousin backcross progeny.

In addition to the individual SNP effects, the effects of the haplotypes formed for each gene were also investigated. Only haplotypes of *BCMO1*, *PPARGC1A*, *RDHE2* and *SCARB1* genes had effects on beef fat colour.

The most likely pathways involved in the beef fat colour were clarified. The association studies showed that the *BCMO1*, *BCO2*, *RARA*, *RDHE2*, *PPARGC1A* and *ALDH8A1* genes and their interactions account for a large proportion of the variation in beef fat colour. These genes have roles involving retinol or retinoic acid synthesis. Therefore, the retinol/retinoic acid synthesis pathway appears to be the most important in terms of the contribution to the β-carotene concentration in adipose tissue. The effects of APOM and SCARB1 indicate that transportation of β-carotene is also important in the regulation of the β-carotene concentration in fat.

Differences in the expression of *BCMO1*, *BCO2* and *RDHE2* genes were investigated. The *RDHE2* gene mRNA transcript level was significantly different between yellow fat and white fat samples. The gene expression of *BCO2* was highly correlated with β-carotene concentration. The results further support the role of *BCO2* in cleaving β-carotene eccentrically and the association of *RDHE2* with β-carotene concentration.

The results also indicate that the control of the retinol/retinoic acid pathway at the gene expression level is important for the β-carotene concentration in subcutaneous adipose tissue and consequently, for beef fat colour.

The study conducted herein contributes to the understanding of the metabolism of carotenoids and their numerous derivatives. BCO2-mediated conversion of β-carotene to vitamin A is confirmed in cattle. Epistatic effects accounted for much of the beef fat colour and β-carotene concentration variation. DNA variants that have a large influence on fat colour, such as the *BCO2* W80X, can be used in marker selection systems to rapidly reduce the incidence of yellow fat colour in beef.