Variation of flour colour in Western Australia adapted wheat: comparative genomics, molecular markers and QTL analysis

Karon Magdalene Leanne Ryan BSc (Hons)

Sponsored by

Valued Added Wheat Cooperative Research Centre Limited
New South Wales, Australia.

Candidature at

State Agricultural and Biotechnology Centre
Murdoch University, Western Australia.

Supervised by

Dr Michael Francki[1-3] and Prof Rudi Appels[1,3]

1 State Agricultural and Biotechnology Centre, Murdoch University, Western Australia.
2 Value Added Wheat Co-operative Research Centre Limited, New South Wales.
3 Department of Agriculture, Western Australia.

Submission Date: 31st October 2005

Declaration & List of Papers Published

I declare that this thesis is my own account of my research and contains as its main content, work that has not previously been submitted for a degree at any tertiary education institution.

Papers that have been published from the research described in this thesis are:

Francki M, Carter M, Ryan K, Hunter A, Bellgard M and Appels R (2004). Comparative organization of wheat homoeologous groups 3S and 7L using wheat-rice synteny and identification of potential markers for genes controlling xanthophyll content in wheat. Functional & Integrative Genomics 4, 118-130.

……..

Karon Magdalene Leanne Ryan

Abstract

The yellowness of flour colour ranges is an important quality trait in wheat for end-use products and is determined by the accumulation of carotenoids in the endosperm. The aims of this study were to develop EST-based molecular markers for genes encoding enzymes of the carotenoid biosynthetic pathway leading to xanthophyll accumulation and identify quantitative trait loci for flour colour (b*) and xanthophyll content in Western Australian adapted germplasm.

A novel bioinformatic strategy was developed to identify rice genes encoding key enzymes of the carotenoid biosynthetic pathway and to predict wheat orthologues on the short arm of chromosome 3 or long arm of chromosome 7. The bioinformatic strategy involved the identification of rice carotenoid genes on BAC/PAC contigs aligned to wheat mapped ESTs. Rice genes predicted to have wheat orthologues were selected based on ESTs mapping to regions on wheat homoeologous chromosomes 3 and 7 known to be involved in flour colour. The rice genes predicted to have wheat orthologues were *Geranylgeranyltransferase* I *β*–subunit (*GGT*-I*beta*) and *Rab geranylgeranyltransferase* component A (*RGGT*-A) on the short arm of chromosome 3, *Lycopene β–cylcase* (*LBC*) on the long arm of chromosome 3 and *Lycopene ε–cylcase* (*LEC*) on the long arm of chromosome 7.

The prediction of these wheat orthologues provided the basis for development of EST-based molecular markers for detecting variation in xanthophyll content. Wheat ESTs with unknown chromosomal locations and having the highest similarity to *GGT*-I*beta*, *RGGT*-A and *LBC* were selected for the development of molecular markers. No EST homologues were identified for LEC and therefore this gene was not further considered. Orthology was confirmed by sequencing and deletion lines were used to confirm chromosomal locations. Two partial orthologues of *GGT*-I*beta* were identified on the short arms of chromosomes 3B and 3D. A partial orthologue of *RGGT*-A was mapped to the proximal regions of the short and long arms of chromosome 3B. At least two or more orthologues of *LBC* were identified from nullisomic-tetrasomic lines. An EST-based molecular marker for *GGT*-I*beta* was found to be involved in minor variation of xanthophyll content in a Westonia*2/Janz doubled haploid population.

QTL analysis from three doubled haploid populations indicated variation in WA-adapted germplasm may be due to different alleles controlling flour colour. QTLs for

b* and xanthophyll content were found to coincide on the short arms of chromosomes 3A, 4D, and 7B and the long arm of chromosomes 7A and 7B in WA-adapted germplasm. Homoeologous expression of regions controlling variation in b* and xanthophyll content on the long arm of chromosomes 7A and 7B suggests the shut-down of genes in the same region on chromosome 7D. The main outcome of this study is flour colour and identification of gene orthologues in wheat controlling variation in xanthophyll content is complex most likely because of the interaction of the carotenoid biosynthetic pathway with other pathways.

Acknowledgements

There are so many people that I must thank who have contributed throughout my PhD candidature which has been a rewarding experience, both professionally and personally. I sincerely thank my supervisors, Dr Michael Francki and Prof Rudi Appels, for their constant guidance and support over the years. I am especially grateful to Michael for his endless patience and constant encouragement during the last four months because I would not have got through it without him. I am also grateful to Rudi for his invaluable advice and infinite wisdom on molecular genetics. He has been a wealth of knowledge for me. Michael and Rudi, through their kindness and belief in me, gave me the courage and strength to continue and I will always appreciate what both have done for me.

Thankyou to the Value Added Wheat CRC for the opportunity and I particularly thank Clare Johnson for her kindness and support over the years. I owe special thanks to Dr Katia Stefanova for the statistical modelling of phenotypic data and I am very grateful to Katia for her constant assistance and encouragement. I also thank Dr Ari Verbyla for his help and kindness. I am also grateful to Dr Bill Lambe and staff for milling and phenotyping activities. I would also like to thank the breeders, Iain Barclay, Robin Wilson and Robyn McLean and their teams from the Department of Agriculture Western Australia (DAWA) for the management of field trials.

Thank you to Paula Moolhuijzen and Adam Hunter from CBBC for their bioinformatic support. I also thank Dr Rob Trengove for the HPLC work in association with Prof Rudi Appels. Thank you to Dr Wujun Ma for kindly providing technical advice on QTL Cartographer. I especially thank everyone from the DAWA lab in the SABC at Murdoch for their endless support and help over the years. I particularly thank Danielle for listening to me and making me laugh. I also thank my fellow PhD students, Dan and Yumi, and I wish them all the best for their remainder of their candidature.

I have been so lucky to have the best support from my family and all my friends. I am grateful to them all because they gave me the emotional strength to see it through. I especially thank my Mum, my brother Shaun and sister-in-law Kerry for their guidance and understanding. To my friends Debs and Shane, I wouldn't have made it without your help and I will always be truly grateful.

Glossary

ABA abscisic acid.

ABA abscisic acid locus.

ADR Australian Durum wheat.

AFLP Amplified fragment length polymorphisms.

allele One of the different forms of a gene that can exist a single locus.

ASW Australian Standard White.

AWB Australian Wheat Board

BAC Bacterial artificial chromosome; an F plasmid engineered to act as a cloning vector that can carry large inserts.

BCH *β-Carotene hydroxylase*.

BLAST Basic Local Alignment Search Tool.

BLASTN comparison of nucleotide sequence databases.

BLASTP comparison of protein sequence databases.

BLASTX comparison of nucleotide query sequence translated in all reading frames against a protein sequence database.

BSA Bulked segregant analysis.

candidate gene A sequenced gene of previously unknown function that, because of its chromosomal position or some other property, becomes a candidate for a particular function such as disease determination.

Ccs *Capsanthin-capsorubin synthase* gene.

cDNA (complementary DNA) Synthetic DNA transcribed from a specific RNA through the action of the enzyme reverse transcriptase.

CIE b* Commission Internationale l'Eclairage whiteness/yellowness value.

CIE L* Commission Internationale I'Eclairage L* (brightness) value.

CIELAB Commission Internationale I'Eclairage L* a* b* colour space.

CO_2 Carbon dioxide

crtE *Lycopene ε-cyclase* gene from *Erwinia uredovora.*

CrtL-b *Lycopene β-cyclase* gene.

CrtL-e *Lycopene ε-cyclase* gene.

CrtR *β-Carotene hydroxylase* gene.

DAWA Department of Agriculture Western Australia.

dbEST Expressed sequence tag database.

deletion Removal of a chromosomal segment from a chromosome set.

DH Doubled haploid.

DMADP dimethylallyl diphosphate.

DNA (deoxyribonucleic acid) A double chain of linked nucleotides (having deoxyribose as their sugars); the fundamental substance of which genes are composed.

DPA Days post anthesis.

environment The combination of all the conditions external to the genome that potentially affect its expression and its structure.

EST Expressed sequence tag(s).

FAD Flavin adenine dinucleotide.

FAO Food and Agricultural Organization of the United Nations

FDP Farnesyl diphosphate.

ha Major locus controlling the recessive grain hardness trait.

GDP Geranyl diphosphate.

GDPS Rab Geranylgeranyl transferase-like protein.

gene The fundamental physical and functional unit of heredity, which carries information from one generation to the next; a segment of DNA, composed of a transcribed region and a regulatory sequence that makes possible transcription.

gene family A set of genes in one genome all descended from the same ancestral gene.

genome The entire complement of genetic material in a chromosome set.

genotype The specific allelic composition for a certain gene or set of genes.

GGH Geranylgeranyl hydrogenase (reductase).

GGPP Geranylgeranyl pyrophosphate.

GGPP Geranylgeranyl pyrophosphate synthase.

GGT-Ibeta Geranylgeranyltransferase I β–subunit.

GRDC Grains Research and Development Corporation of Australia

HIF heterogenous inbred families.

HMW-GS high-molecular-weight glutenin subunit.

IDP Isopentyl diphosphate.

ITMI International Triticeae Mapping Initiative.

L Long arm of chromosome.

LBC *Lycopene β–cyclase*.

LEC *Lycopene ε–cyclase*.

LHC light harvesting complex.

library A collection of DNA clones obtained from one DNA donor.

linkage The association of genes on the same chromosome.

linkage group A group of genes known to be linked; a chromosome.

linkage map A chromosome map; an abstract map of chromosome loci, based on recombinant frequencies.

LMW-GS low-molecular-weight glutenin subunit.

locus (plural **loci**) The specific place on a chromosome where a gene is located.

MAS Marker assisted selection.

molecular markers genetic markers; alleles used as experimental probes to keep track of an individual, a tissue, a cell, a nucleus, a chromosome, a gene or a trait.

NIL Near isogenic line.

NS *Neoxanthin synthase*.

Nulli-tetra Nullisomic-tetrasomic line.

NWMMP National Wheat Molecular Marker Project.

PAC Phage P1 artificial chromosome.

PCR *See* **Polymerase chain reaction**.

PDS *Phytoene desaturase*.

phenotype (1) The form taken by some character (or group of characters) in a specific individual. (2) The detectable outward manifestations of a specific genotype.

Polymerase chain reaction A method for amplifying specific DNA segments which exploits certain features of DNA replication.

polymorphism The occurrence in a population (or among populations) of several phenotypic forms associated with alleles of one gene or homologs of one chromosome.

PPDP Prephytoene diphosphate.

PPO Polyphenol oxidase.

PHST Pre-harvest sprouting tolerance.

PSII Photosystem II.

PYS *Phytoene synthase*.

Quantitative trait loci loci controlling genetic variation of a quantitative trait.

QTL *See* **Quantitative trait loci**.

RAPD Randomly amplified polymorphic DNA.

RFLP Restriction fragment length polymorphisms.

***RGGT*-A** *Rab geranylgeranyltransferase* component A.

Rht Semi-dwarfing (restricted height) gene.

RIL Recombinant inbred line.

S Short arm of chromosome.

SNP Single nucleotide polymorphisms.

SSR Simple sequence repeats (or microsatellites).

STS Sequence tagged site.

TBLASTN comparison of a protein query sequence against a nucleotide sequence database translated in all reading frames.

TBLASTX comparison of six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.

TIGR The Institute for Genomic Research

VDE *Violaxanthin deepoxidase*.

Vp14 *Violaxanthin-cleavage* gene (*Zea mays*).

WA Western Australia.

WSN White salted noodles.

YAC yeast artificial chromosome.

YAN Yellow alkaline noodles.

ZDS *ζ-Carotene desaturase*.

ZE *Zeaxanthin epoxidase*.

Table of Contents

Page #

List of Figures

Chapter 1

Chapter 2

List of Tables