

# **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<sup>1-3</sup> and Prof Rudi Appels<sup>1,3</sup>**

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: 31<sup>st</sup> 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  $\beta$ -subunit (GGT-Ibeta)* and *Rab geranylgeranyltransferase component A (RGGT-A)* on the short arm of chromosome 3, *Lycopene  $\beta$ -cylcase (LBC)* on the long arm of chromosome 3 and *Lycopene  $\epsilon$ -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-Ibeta*, *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-Ibeta* 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-Ibeta* 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 at 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** Bulk 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 l'Eclairage L\* (brightness) value.

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

**CO<sub>2</sub>** 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  $\beta$ -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  $\beta$ -cyclase*.

**LEC** *Lycopene  $\epsilon$ -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 #

<b>Declaration.....</b>	<b>i</b>
<b>Acknowledgements.....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>iv</b>
<b>Glossary.....</b>	<b>v</b>
<b>Table of Contents.....</b>	<b>8</b>
<b>List of Figures.....</b>	<b>12</b>
<b>List of Tables.....</b>	<b>17</b>

## **Chapter 1. Introduction and Literature Review**

1.1. Quality attributes in wheat.....	1
1.1.1. Grain morphology and seed storage products.....	5
1.1.1.1 Morphology of wheat grain.....	5
1.1.1.2 Biochemical components of wheat grain.....	8
1.1.2. Flour colour.....	9
1.1.3. Carotenoid and xanthophyll biosynthetic pathways of higher plants.....	10
1.1.3.1 Biosynthesis of carotenoids.....	12
1.1.3.2 Characterisation of carotenoid biosynthetic enzymes.....	13
1.2. Wheat genome organisation.....	15
1.2.1. General organisation.....	15
1.2.2. Genetic analysis utilising chromosomal aberrations in wheat.....	16
1.3. Genetic linkage mapping.....	18
1.3.1. Types of molecular markers.....	19
1.4. Applications of genetic linkage mapping.....	22
1.4.1. Marker trait associations for grain quality.....	22
1.4.2. QTLs controlling flour colour and xanthophyll content.....	23
1.4.3. Comparative genetic mapping.....	24
1.4.4. Rice as a model species for grass genomes.....	26
1.4.5. Wheat genomics.....	26
1.5. <i>In silico</i> comparative genomics of rice and wheat.....	27

## **Chapter 2. A novel bioinformatic strategy for predicting wheat orthologues of variation: rice carotenoid genes**

2.1.	Introduction.....	30
2.2.	Materials and Methods.....	34
2.2.1.	General strategy for <i>in silico</i> alignment of rice genes to wheat.....	34
2.2.2.	Details of processes involved in predicting wheat orthologues on groups 3 or 7 using rice genome sequences.....	35
2.2.2.1	Identifying annotated genes in rice related to enzymes of the xanthophyll pathway.....	35
2.2.2.2	Nearest neighbour contig analysis: alignment of rice contigs to wheat chromosomes and identification of wheat orthologues on groups 3 and 7.....	37
2.2.2.2.1	Rice BAC/PAC contigs aligned to wheat chromosomes 3 or 7.....	37
2.2.2.2.2	Selection of rice genes aligning to homoeologous wheat chromosomes 3 or 7 based on mapped wheat EST homology.....	37
2.2.2.3	Reannotation of selected genes predicted to have orthologues on wheat groups 3 or 7.....	39
2.2.2.3.1	Confirmation of gene structure for selected BAC/PAC clones using full length-cDNA clones.....	39
2.2.2.3.2	Analysis of GenBank rice annotations for selected candidate genes related to carotenoid pathway and alignments with wheat ESTs.....	39
2.3.	Results.....	40
2.3.1.	Identification of annotated carotenoid genes on rice chromosomes and confirmation of likely wheat orthologues.....	40
2.3.2.	Alignment of rice genes on wheat chromosomes.....	42
2.3.2.1	Selection of rice BAC/PAC clones targeting homoeologous wheat chromosomes 3 and 7.....	42
2.3.2.2	Mapped wheat EST alignments to BAC/PAC contigs for selected rice genes.....	44
2.3.3.	Verification of selected rice gene annotations.....	46
2.3.3.1	Confirmation of gene structure for selected BAC/PAC clones using full length-cDNA clones.....	46
2.3.3.2	Amino acid similarity with enzymes from other species.....	47
2.4.	Discussion.....	52

## **Chapter 3. Development of EST-based molecular markers for detecting variation in xanthophyll content**

3.1.	Introduction.....	56
3.2.	Materials and Methods.....	59
3.2.1.	Plant material.....	59

3.2.2.	Development of EST-based markers for xanthophyll content in wheat.....	59
3.2.2.1	Wheat EST alignments to selected rice genes for xanthophyll content.....	59
3.2.2.2	DNA extraction.....	60
3.2.2.3	Primer design and optimisation.....	61
3.2.2.4	Sequencing of partial candidate genes for xanthophyll content.....	62
3.2.2.4.1	DNA gel extraction and purification.....	62
3.2.2.4.2	Sequencing reaction, purification and analysis.....	62
3.2.3.	Genetic map construction.....	63
3.2.3.1	Linkage analysis.....	63
3.2.4.	Phenotyping of quality traits in WA-adapted germplasm.....	63
3.2.4.1	Field trials, milling design and prediction modelling.....	63
3.2.4.2	Milling.....	65
3.2.4.3	Flour colour evaluation.....	65
3.2.4.4	Xanthophyll content assay.....	65
3.2.5.	Major component of xanthophylls: lutein and lutein esters.....	66
3.2.6.	QTL analysis.....	67
3.2.7.	Statistical analysis.....	68
3.3.	Results.....	69
3.3.1.	Wheat EST alignments to selected rice genes for xanthophyll content.....	69
3.3.2.	Confirmation of wheat orthologues, chromosome location and detection of sequence variation between cultivars.....	72
3.3.2.1	Deletion mapping of BJ305709 and confirmation of orthology to rice <i>RGGT-A</i> .....	72
3.3.2.2	Deletion mapping of BE445261 and confirmation of orthology to rice <i>GGT-Ibeta</i> .....	74
3.3.2.3	Deletion mapping of BE591408 and CD887154, and confirmation of orthology to rice <i>LBC</i> .....	77
3.3.3.	Identification of polymorphisms between Australian wheat varieties.....	79
3.3.4.	Flour colour and xanthophyll content in groups 3 and 7 of Westonia*2/Janz doubled haploid population.....	80
3.3.4.1	Phenotypic analysis of b* and xanthophyll content in Westonia*2/Janz doubled haploid population.....	80
3.3.4.2	Single marker regression analysis of b* and xanthophyll content in Westonia*2/Janz doubled haploid population.....	82
3.4.	Discussion.....	85

## **Chapter 4. QTL analysis of b\* and xanthophyll content in WA adapted germplasm**

4.1.	Introduction.....	89
4.2.	Materials and Methods.....	92
4.2.1.	Plant material.....	92
4.2.2.	Field trials, milling design and prediction modelling.....	92
4.2.3.	Phenotyping of quality traits in WA adapted germplasm.....	92
4.2.4.	Genetic map construction and linkage analysis.....	93
4.2.5.	QTL analysis.....	94
4.2.6.	Statistical analysis.....	94
4.3.	Results.....	95
4.3.1.	Phenotyping of quality traits.....	95
4.3.1.1	Ajana/WAWHT2074 population.....	95
4.3.1.2	Carnamah/WAWHT2046 population.....	97
4.3.2.	QTL analysis.....	99
4.3.2.1	Ajana/WAWHT2074 population.....	99
4.3.2.2	Carnamah/WAWHT2046 population.....	102
4.4.	Discussion.....	105

## **Chapter 5. General Discussion**

5.1.	Overview.....	108
5.2.	Analysis of full length genes and associated molecular processes.....	108
5.3.	Gene networks and regulation.....	110
5.4.	Gene duplication, divergence and homoeologous expression.....	111
5.5.	Degradation of flour colour.....	113
5.6.	Future work.....	114

<b>References.....</b>	<b>116</b>
------------------------	------------

<b>Appendix A.....</b>	<b>134</b>
------------------------	------------

## List of Figures

### Chapter 1

- 1.1. Summary of wheat processing and the significant grain features.
- 1.2. Schematic diagram showing the main components of a cereal grain.
- 1.3. Transverse section of a developing wheat grain showing the major components.
- 1.4. Summary of the isoprenoid biosynthetic pathway.
- 1.5. The carotenoid biosynthetic pathway.

### Chapter 2

- 2.1. Summary flow chart of processes involved in identifying candidate genes for xanthophyll content on homoeologous wheat chromosomes 3 and 7
- 2.2. Schematic diagram of criteria used to select BAC/PAC clones that were syntenic to wheat groups 3 or 7 including the consideration of nearest neighbour
- 2.3. Summary of chromosomal locations of carotenoid genes in rice and alignment of mapped wheat ESTs to flanking BAC/PAC clones.
- 2.4. Alignment of mapped wheat ESTs to rice BAC/PAC contig containing *Rab Geranylgeranyltransferase* component A.
- 2.5. Alignment of mapped wheat ESTs to rice BAC/PAC contig containing *Geranylgeranyltransferase I  $\beta$ -subunit*.
- 2.6. Alignment of mapped wheat ESTs to rice BAC/PAC contig containing *Lycopene  $\varepsilon$ -cyclase*.
- 2.7. Alignment of mapped wheat ESTs to rice BAC/PAC contig containing *Lycopene  $\beta$ -subunit*.
- 2.8. Intron-exon structure of rice (AP002522) and *Arabidopsis GGT-Ibeta* (AC004218) orthologues from cDNAs.
- 2.9. Intron-exon structure of rice *RGGT-A* (AP001073) and *Arabidopsis* putative Rab

escort protein (AC020580) from cDNAs.

- 2.10. Intron-exon structure of rice (AP003332) and *Arabidopsis LEC* (At5g57030) orthologues from cDNAs.
- 2.11. Amino acid sequence alignment of protein *geranylgeranyltransferase* I beta chain of rice (GenBank accession AP002522), *Arabidopsis* (AF311225), yeast (*Saccharomyces cerevisiae*) (M74109) and human (L25441).
- 2.12. Results from TBLASTX search of *Arabidopsis* (AF311225) showing conserved region with rice (AP002522) indicating incorrect prediction of rice amino acid sequence.
- 2.13. Amino acid sequence alignment of protein RGGT-A for rice (AP001073), *Arabidopsis* (AC020580\_9), human (CAA5011) and yeast MSI4 (*Saccharomyces cerevisiae*) (CAA99701).
- 2.14. Amino acid sequence alignment of protein LEC for rice (AP003332) and *Arabidopsis* (AAB53336)
- 2.15. Amino acid sequence alignment of protein LBC for rice (AP005849) and *Arabidopsis* (AAB53337).

### **Chapter 3**

- 3.1. HPLC analysis of xanthophylls extracted from flour of Blanco, non-saponified (a.) and saponified (b.).
- 3.2. Summary of wheat ESTs that showed highest homology to the four candidate genes from rice.
- 3.3. Predicted protein sequence alignment of a portion of rice RGGT-A (AP001073) to portion of wheat EST (BJ305709) with 81% identity in 113 amino acid residues.
- 3.4. Predicted protein sequence alignment of a portion of rice GGT-Ibeta (AP002522) exons 3-6 to portion of wheat EST (BE445261) showing 75% identity in 99 amino acid residues.
- 3.5. Predicted protein sequence alignment of a portion of rice LBC (AP005849) to

- portion of wheat EST (CA637200) showing 85% identity in 165 amino acid residues.
- 3.6. Predicted protein sequence alignment of a portion of rice LBC (AP005849) to portion of wheat EST (BE591408) showing 85% identity in 140 amino acid residues.
  - 3.7. Predicted protein sequence alignment of a portion of rice LBC (AP005849) to portion of wheat EST (CD887154) showing 82% identity in 115 amino acid residues.
  - 3.8. Physical mapping of wheat *RGGT-A* (intron 1-2, exon 2, intron 2-3) to the centromeric region of chromosome 3B.
  - 3.9. Nucleotide sequence alignment of *RGGT-A* in wheat cultivar Chinese Spring and wheat EST (BJ305709).
  - 3.10. Sequencing of wheat *RGGT-A* (intron 1-2, exon 2, intron 2-3) in cultivars Chinese Spring, Sunco, Tasman, Westonia and Janz using BJ305709 E1F-E3R.
  - 3.11. Physical mapping of *GGT-Ibeta* using primers E3F-E5R designed from wheat EST BE445261.
  - 3.12. Nucleotide sequence alignment of wheat EST BE445261 and *GGT-Ibeta* in cultivars Westonia and Janz.
  - 3.13. Sequencing of wheat *GGT-Ibeta* (intron 3-4, exon 4, intron 4-5) from BE445261 E3F-E5R in cultivars Westonia and Janz.
  - 3.14. Analysis of wheat *LBC* using a series of Chinese Spring nullisomic-tetrasomic lines.
  - 3.15. Nucleotide sequence alignment of wheat EST CD887154 and *LBC* in cultivar Chinese Spring.
  - 3.16. Sequencing of wheat *LBC* in cultivars Chinese Spring, Sunco, Tasman, Westonia and Janz using BE591408 E1F-E2R and CD887154 E7F-E10R.
  - 3.17. Polymorphic marker of *GGT-Ibeta* using primers E3F-E5R designed from wheat

EST BE445261.

- 3.18. Polymorphic markers on groups 3 and 7 of *Westonia\*2*/Janz, doubled haploid population.
- 3.19. Histograms of frequency distribution of  $b^*$  values and xanthophyll in *Westonia\*2*/Janz doubled haploid populations from the 2002 and 2003 sites of Wongan Hills and Merridan.

#### **Chapter 4**

- 4.1 Histograms of frequency distribution of  $b^*$  values and xanthophyll in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.2 Histograms of frequency distribution of  $b^*$  values and xanthophyll in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.3 Summary of composite interval mapping for all chromosomes of  $b^*$  values and xanthophyll in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.4 QTL of  $b^*$  and xanthophyll content on chromosome 7A in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.5 QTL of  $b^*$  and xanthophyll content on chromosome 3A in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.6 QTL of  $b^*$  and xanthophyll content on chromosome 2D in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.7 QTL of  $b^*$  and xanthophyll content on chromosome 3B in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.



- 4.8 QTL of b\* and xanthophyll content on chromosome 4D in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.9 Summary of composite interval mapping for all chromosomes of b\* values and xanthophyll in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.10 QTL of b\* and xanthophyll content on the short arm of chromosome 7B in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.11 QTL of b\* and xanthophyll content on the long arm of chromosome 7B in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.
- 4.12 QTL of b\* and xanthophyll content on chromosome 4D in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan.

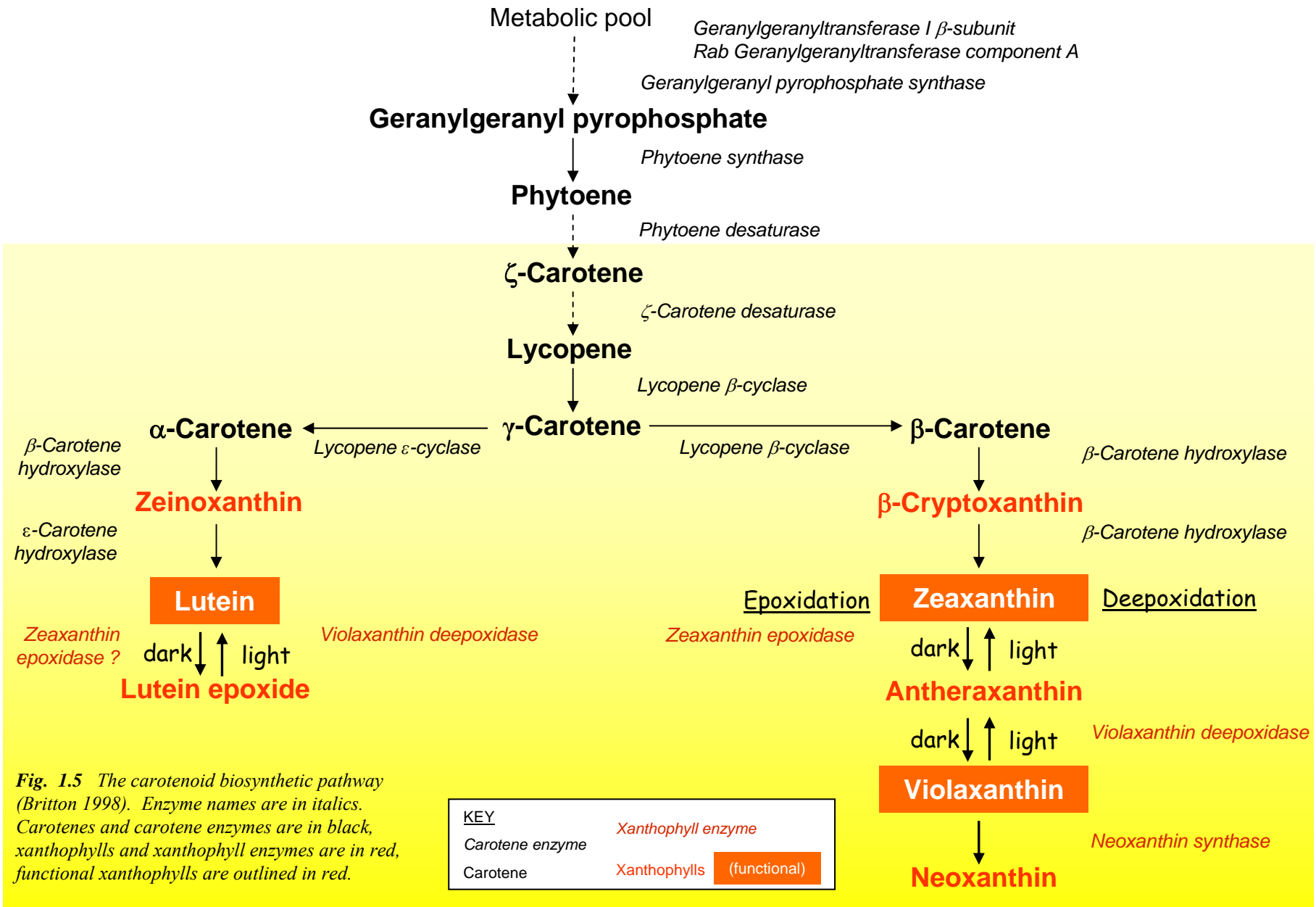
## **Chapter 5**

- 5.1. Summary of the processes used in this thesis and additional processes that can be used to investigate the regulation of pathways such as carotenoid biosynthesis and their role in variation of quality traits.
- 5.2. The carotenoid biosynthetic pathway leading to the xanthophyll accumulation illustrated by colour complementation in *Escherichia coli*.

## List of Tables

1.1. Description of Australian wheat grades defined by the Australian Wheat Board and their end-product uses.....	3
1.2. Protein families of carotenoids and summary of well-characterised genes identified in a variety of plant species for some of the carotenoid biosynthetic enzymes.....	14
1.3. Summary of QTLs controlling quality traits from various quality traits.....	22
2.1. Summary of EST loci mapped in each chromosome of the wheat genomes.....	31
2.2. Chromosomal location of well-characterised carotenoid genes and their possible orthologues in rice chromosomes.....	41
3.1. Summary of cDNA libraries from Chinese Spring germplasm used to generate ESTs produced by the NSF Project.....	60
3.2. Details of primer sequences that successfully amplified regions within each candidate gene.....	61
3.3. Summary of Westonia*2/Janz doubled haploid population field trials in 2002 and 2003 at Wongan Hills and Merredin, Western Australia.....	64
3.4. Details of wheat ESTs that aligned to portions of selected rice genes and their level of homology from TBLASTX results.....	70
3.5. b* and xanthophyll content in Westonia and Janz.....	81
3.6. Pearson's correlation coefficient, $r$ , for phenotypic values of b* and xanthophyll content.....	81
3.7. Single marker regression for b* and xanthophyll content on homoeologous groups 3 and 7.....	83
3.8. Regression analysis of BE445261 marker on chromosome 3 for predicted b* and xanthophyll content.....	83
3.9. Single marker regression for b* and xanthophyll content on homoeologous groups	

other than 3 and 7.....	84
4.1. Summary of the number of doubled haploid lines and duplications in Ajana/WAWHT2074 and Carnamah/WAWHT2046 populations at field trials in 2002 and 2003 at Wongan Hills and Merredin, Western Australia.....	92
4.2. Summary of genetic maps of Ajana/WAWHT2074 and Carnamah/WAWHT2046 doubled haploid populations.....	94
4.3. $b^*$ and xanthophyll content in Ajana and WAWHT2074.....	95
4.4. Pearson's correlation coefficient, $r$ , for phenotypic values of $b^*$ and xanthophyll content for the Ajana/WAWHT2074 population.....	96
4.5. $b^*$ and xanthophyll content for in Carnamah and WAWHT2046.....	97
4.6. Pearson's correlation coefficient, $r$ , for phenotypic values of $b^*$ and xanthophyll content for the Carnamah/WAWHT2046 population.....	99
4.7. Summary of QTL analysis using composite interval mapping of $b^*$ and xanthophyll content in Ajana/WAWHT2074 doubled haploid population.....	101
4.8. Summary of QTL analysis using composite interval mapping of $b^*$ and xanthophyll content in Carnamah/WAWHT2046 doubled haploid population....	103



**Fig. 1.5** The carotenoid biosynthetic pathway (Britton 1998). Enzyme names are in italics. Carotenes and carotene enzymes are in black, xanthophylls and xanthophyll enzymes are in red, functional xanthophylls are outlined in red.

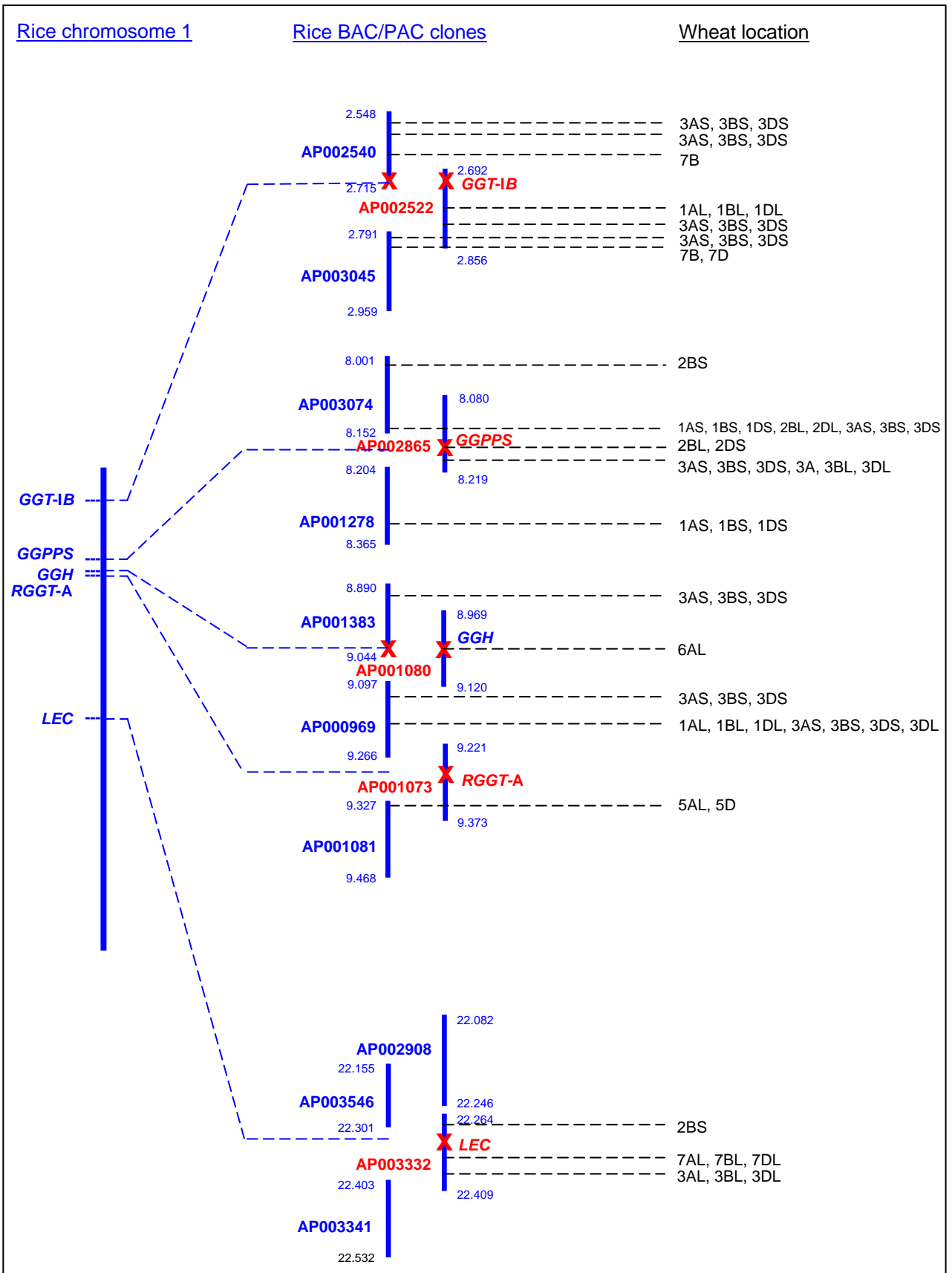


Fig. 2.3 Summary of chromosomal locations of carotenoid genes in rice and alignment of mapped wheat ESTs to flanking BAC/PAC clones.

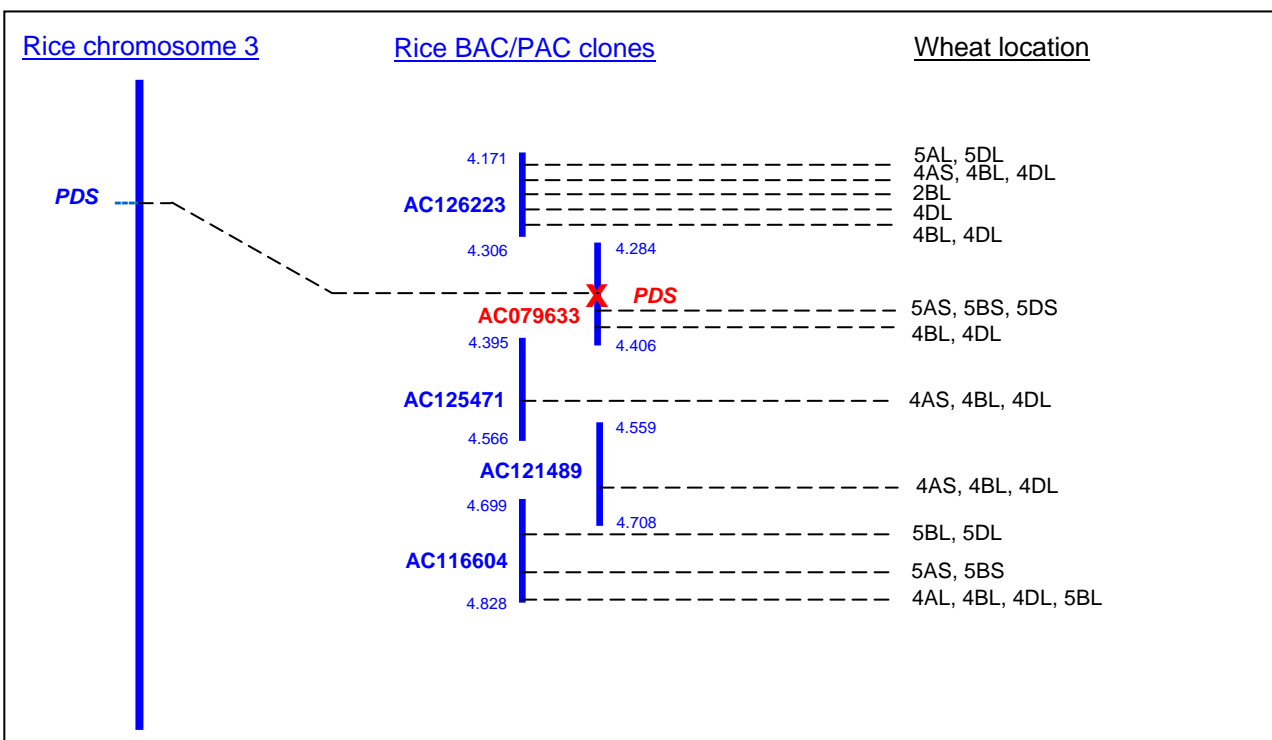
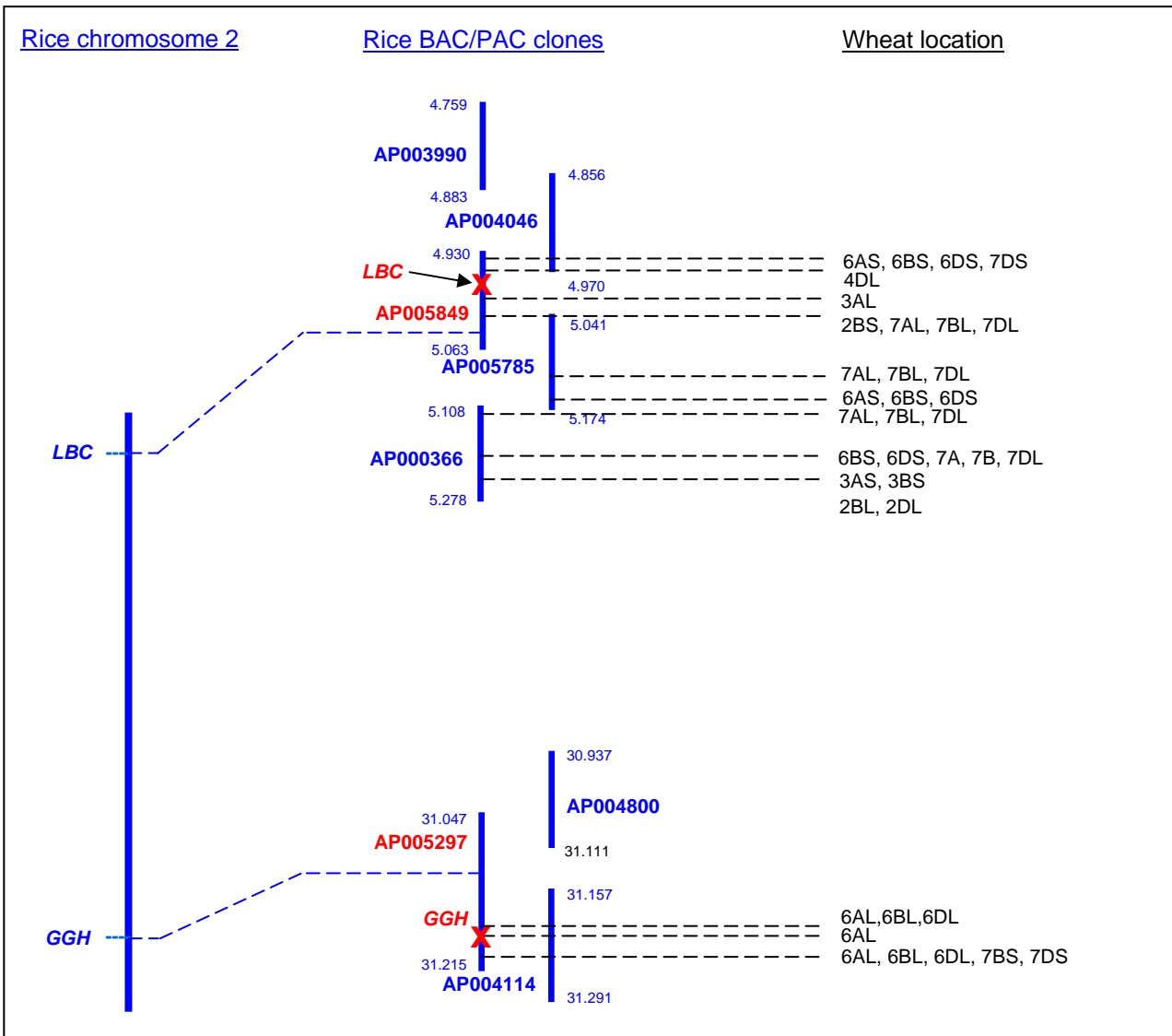


Fig. 2.3 continued.

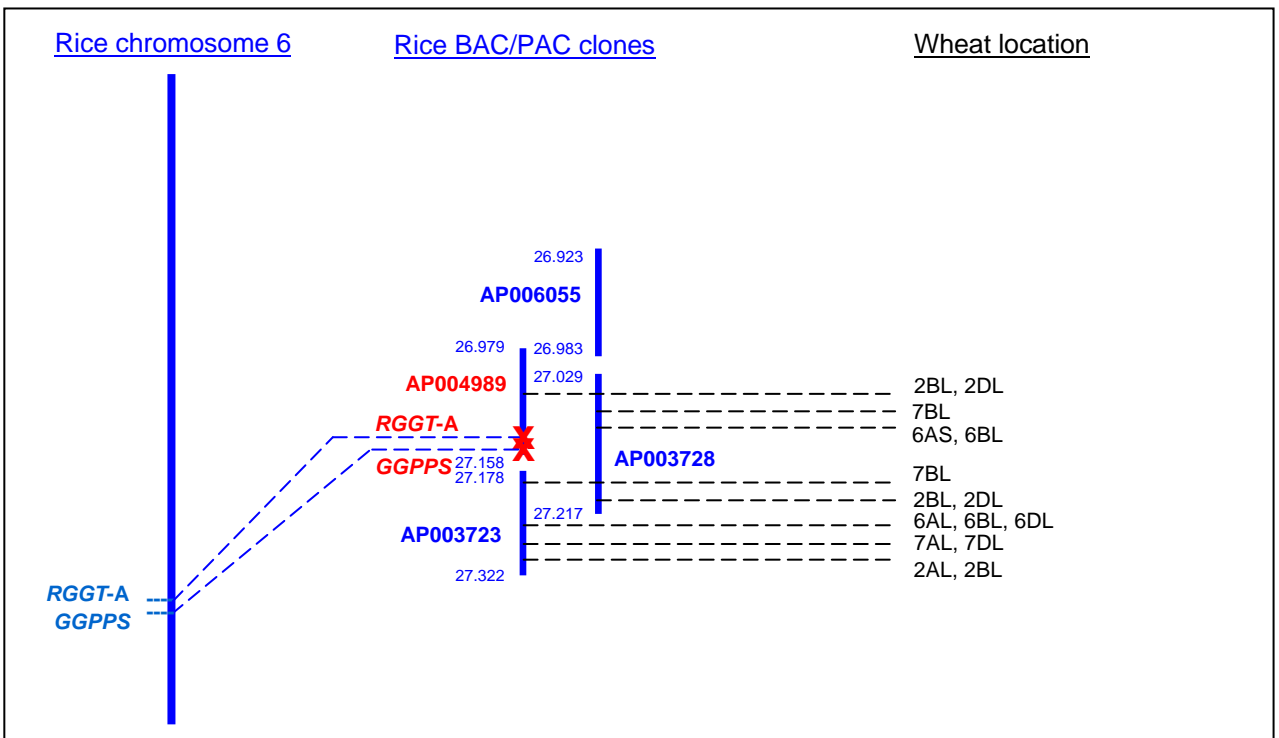
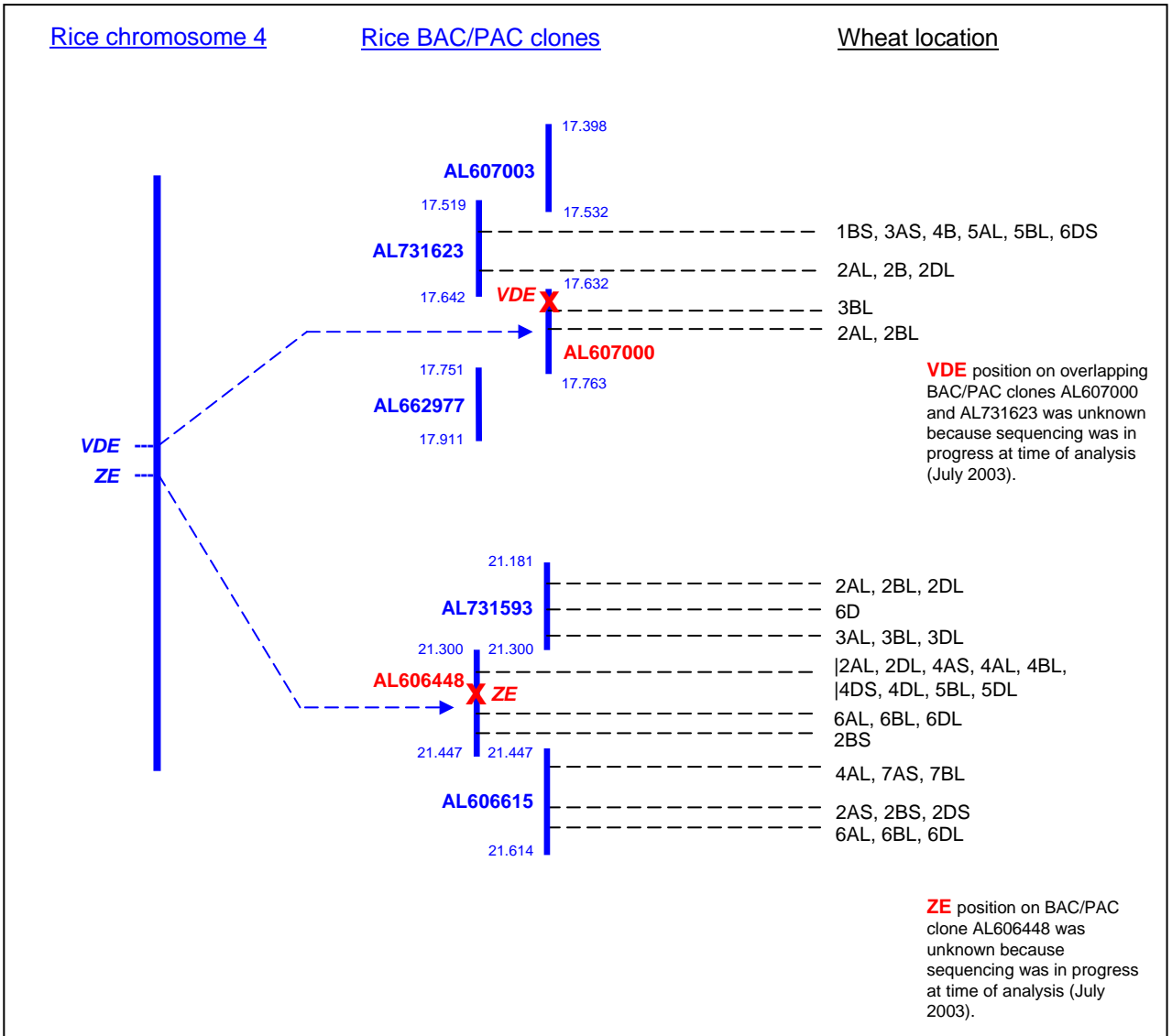


Fig. 2.3 continued.

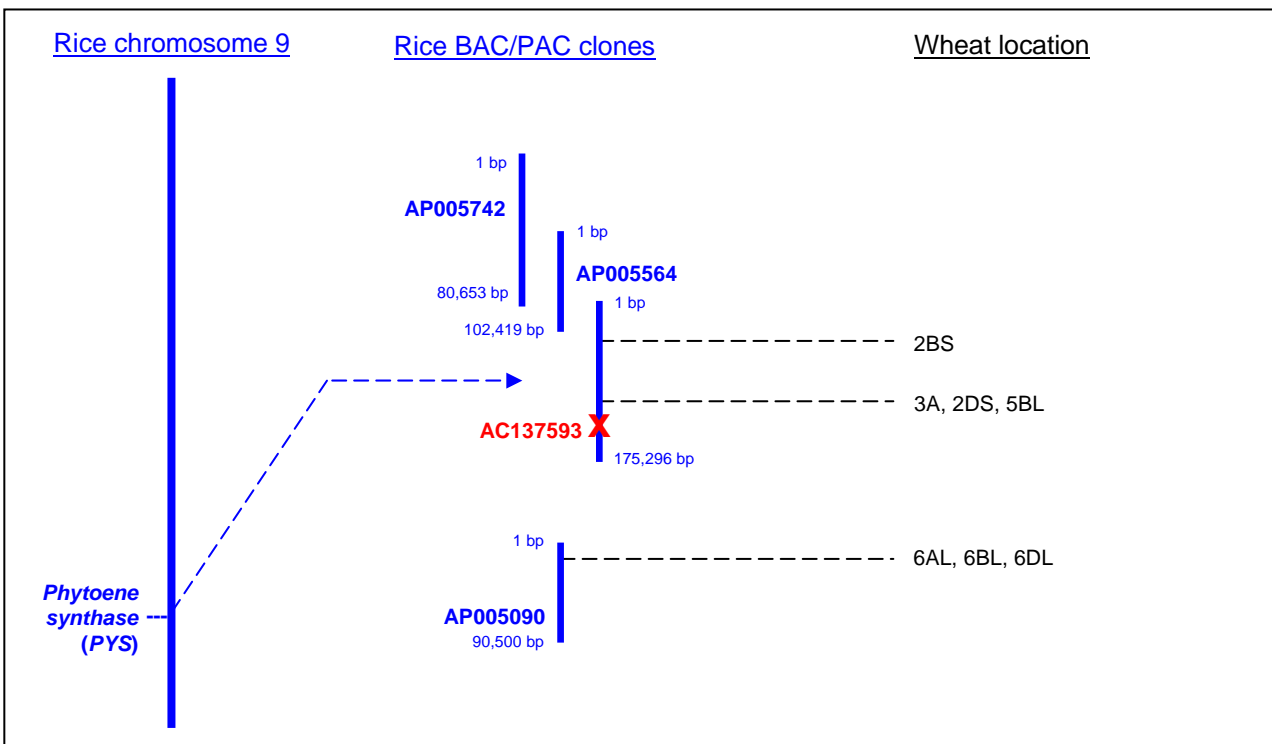
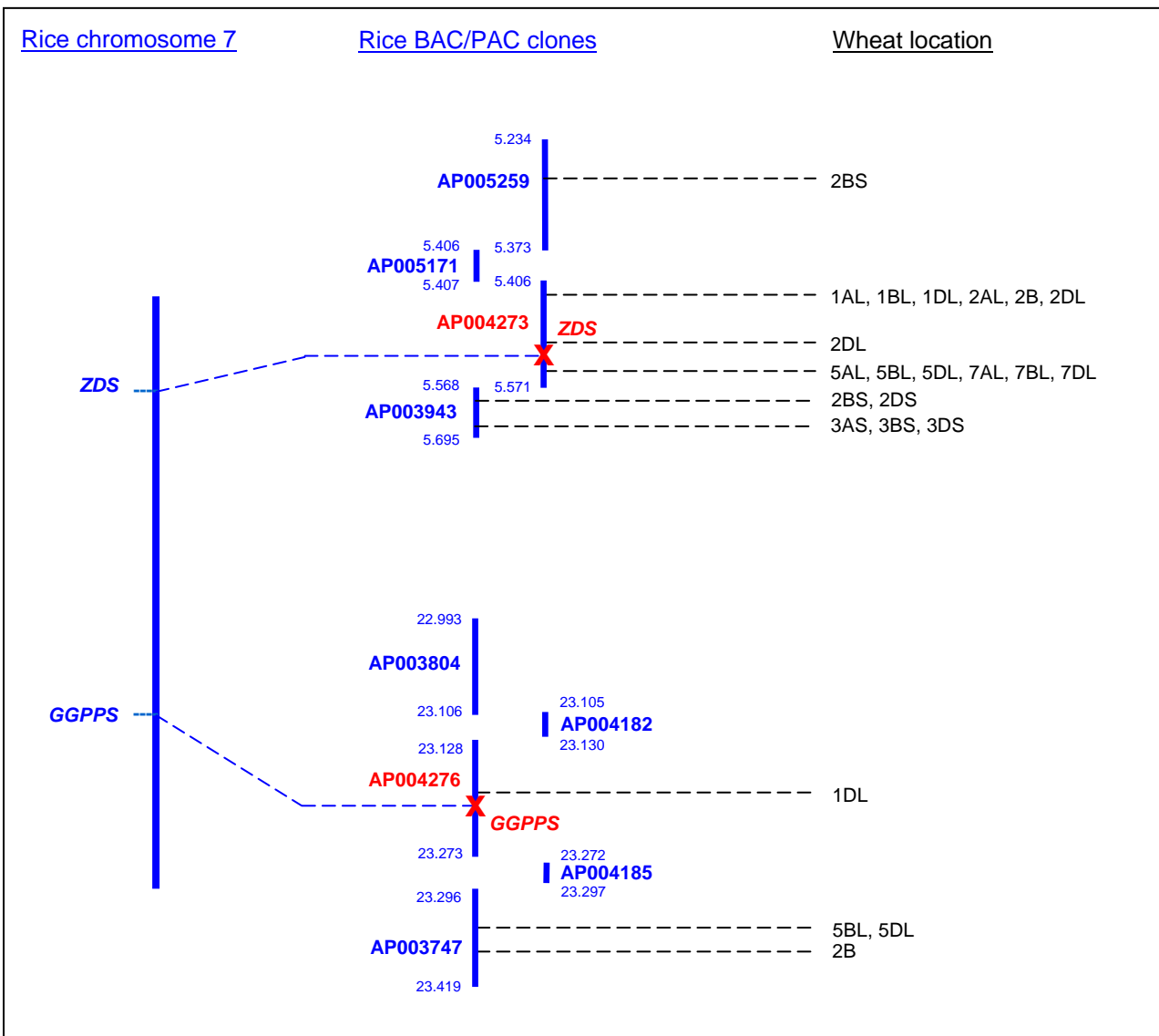


Fig. 2.3 continued.



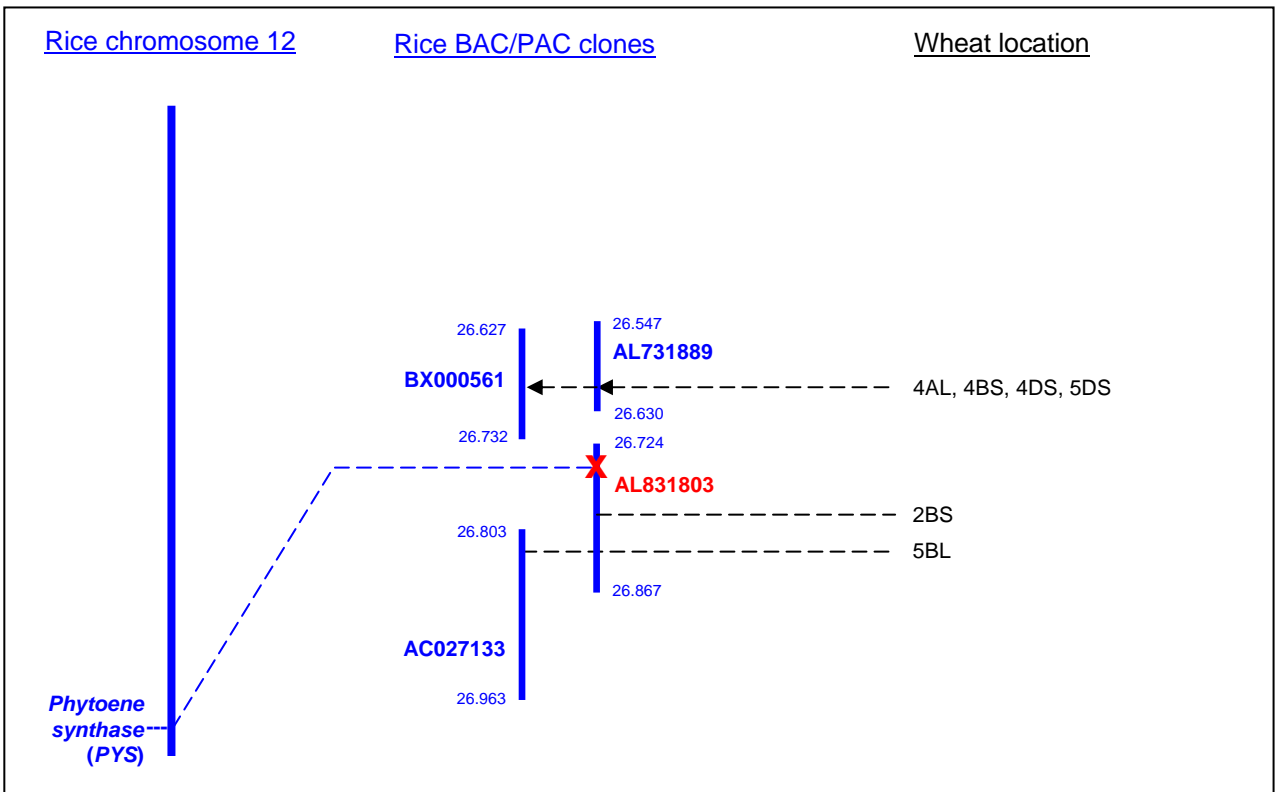
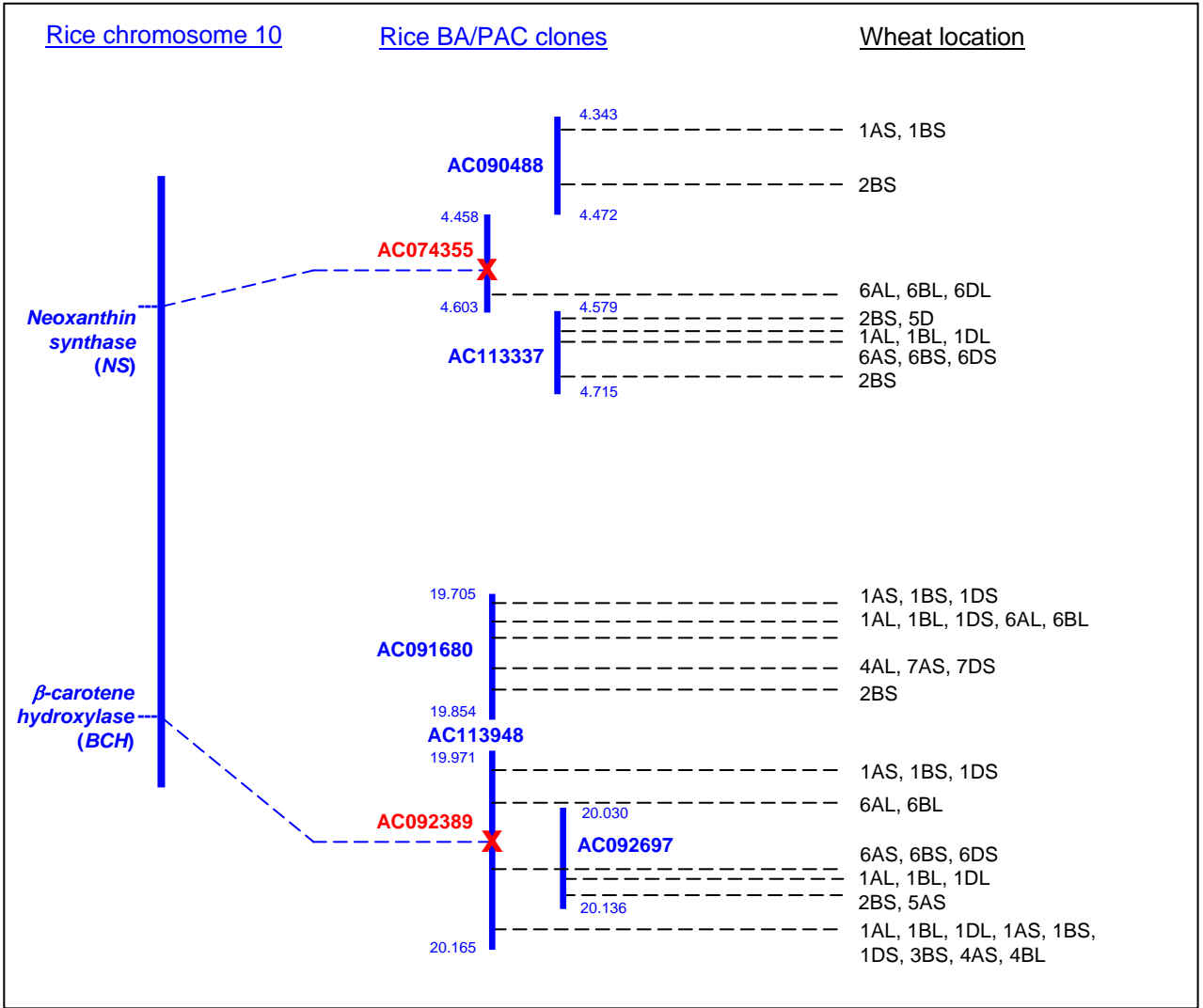
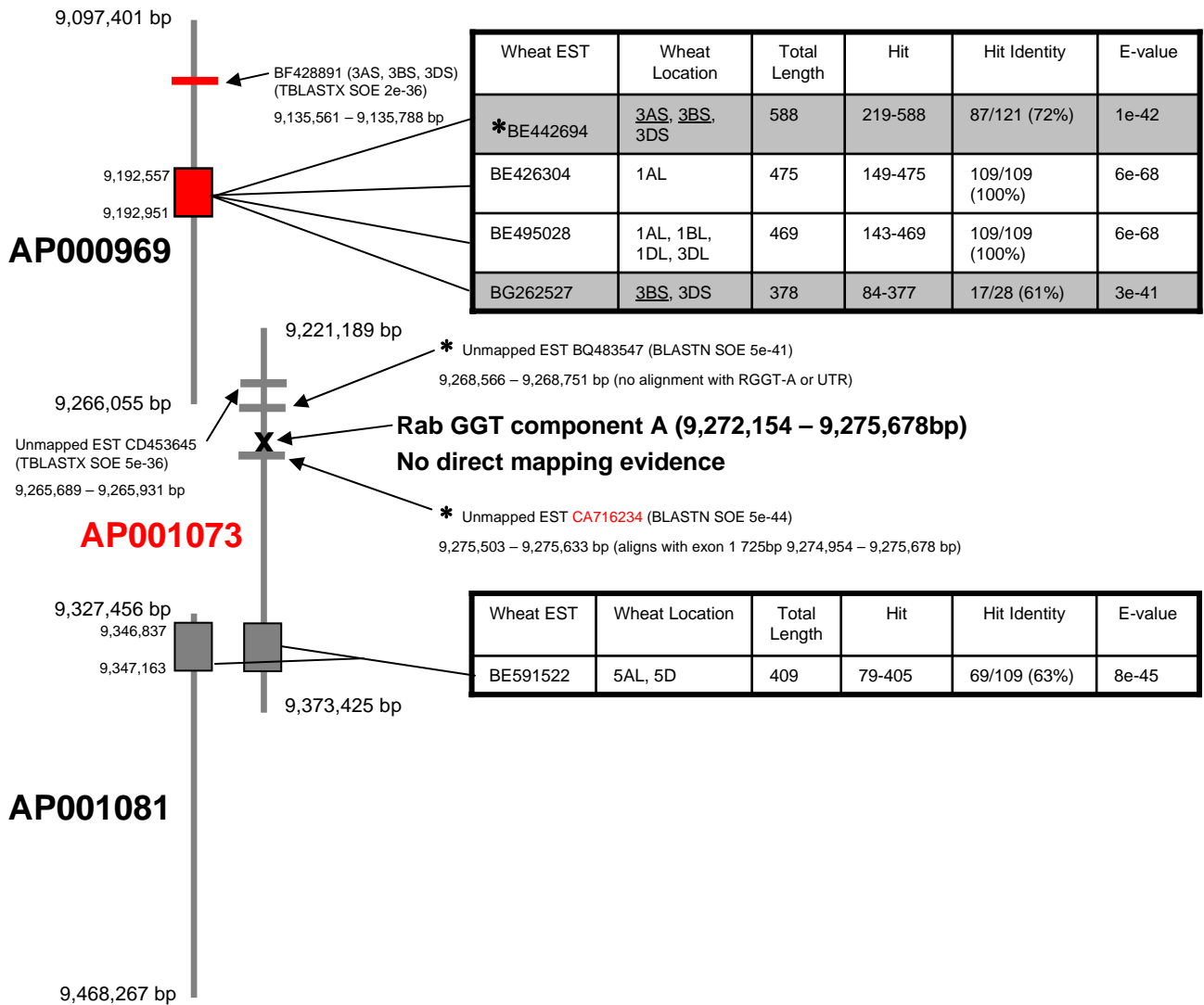
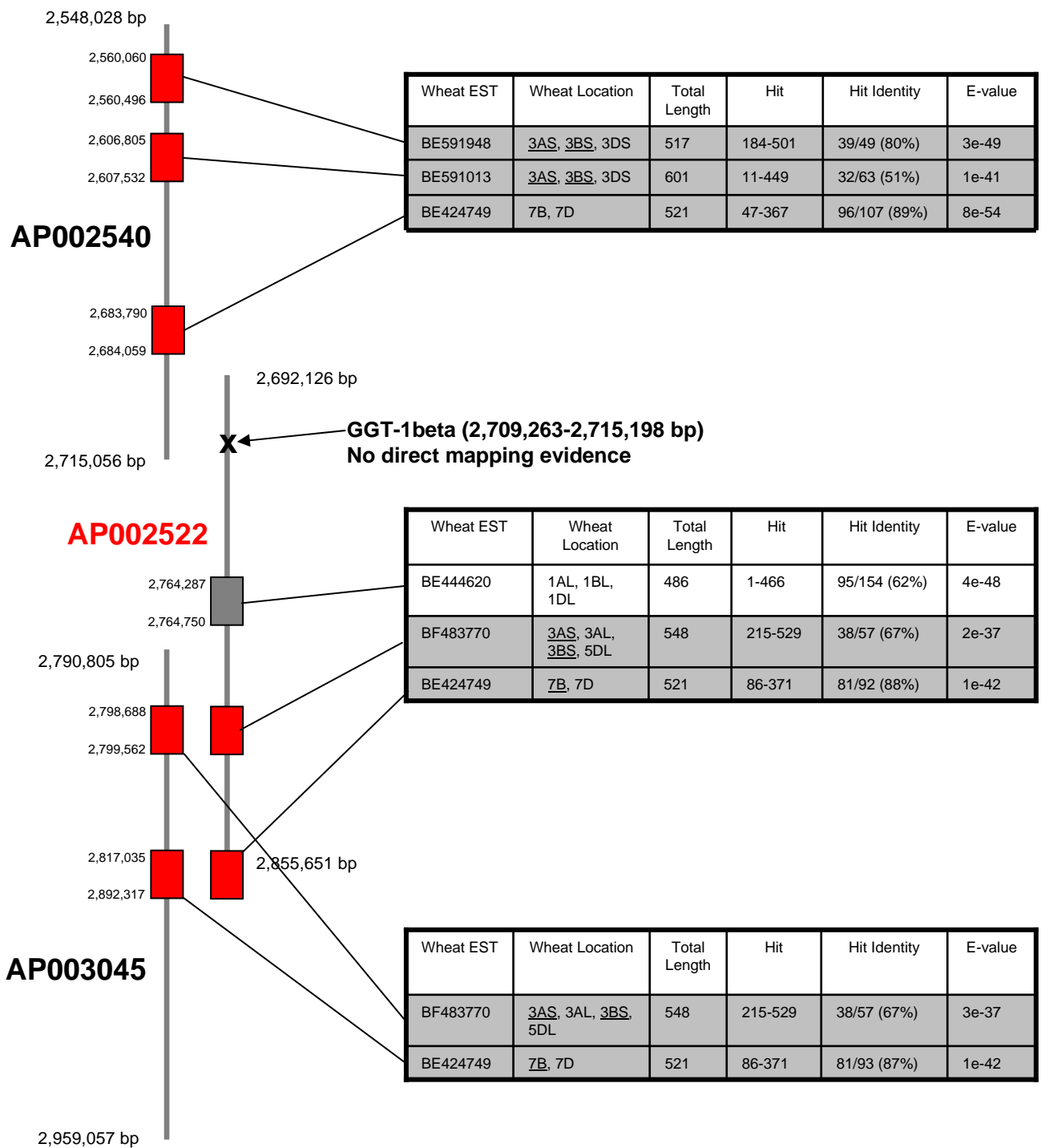


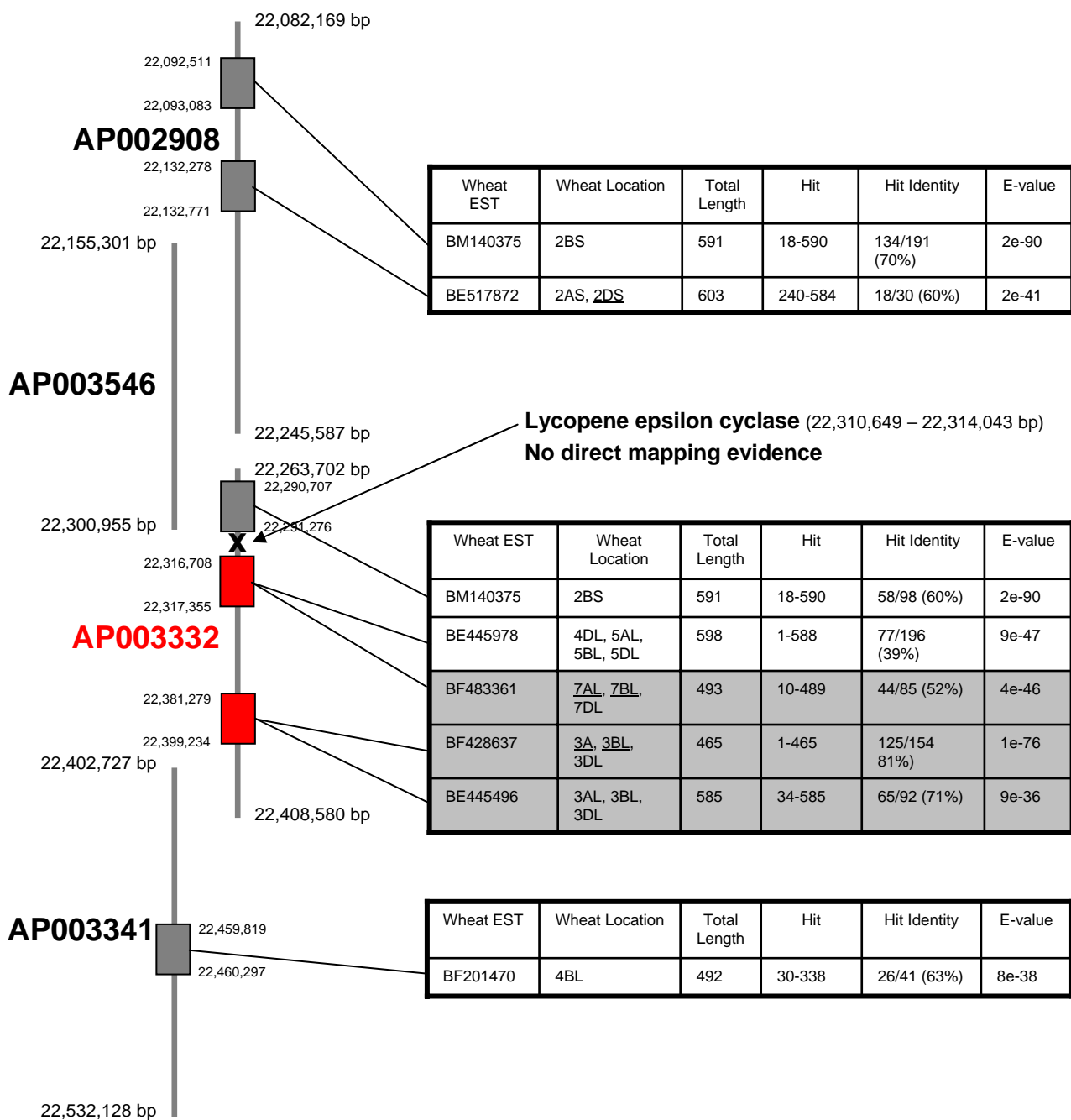
Fig. 2.3 continued.



**Fig. 2.4** Alignment of mapped wheat ESTs to rice Rab Geranylgeranyltransferase component A (AP001073) and flanking BAC/PAC clones on chromosome 1. The homology of the rice BAC/PAC clone and the aligned EST is detailed. Wheat ESTs mapped to groups 3 or 7 are shown in red.

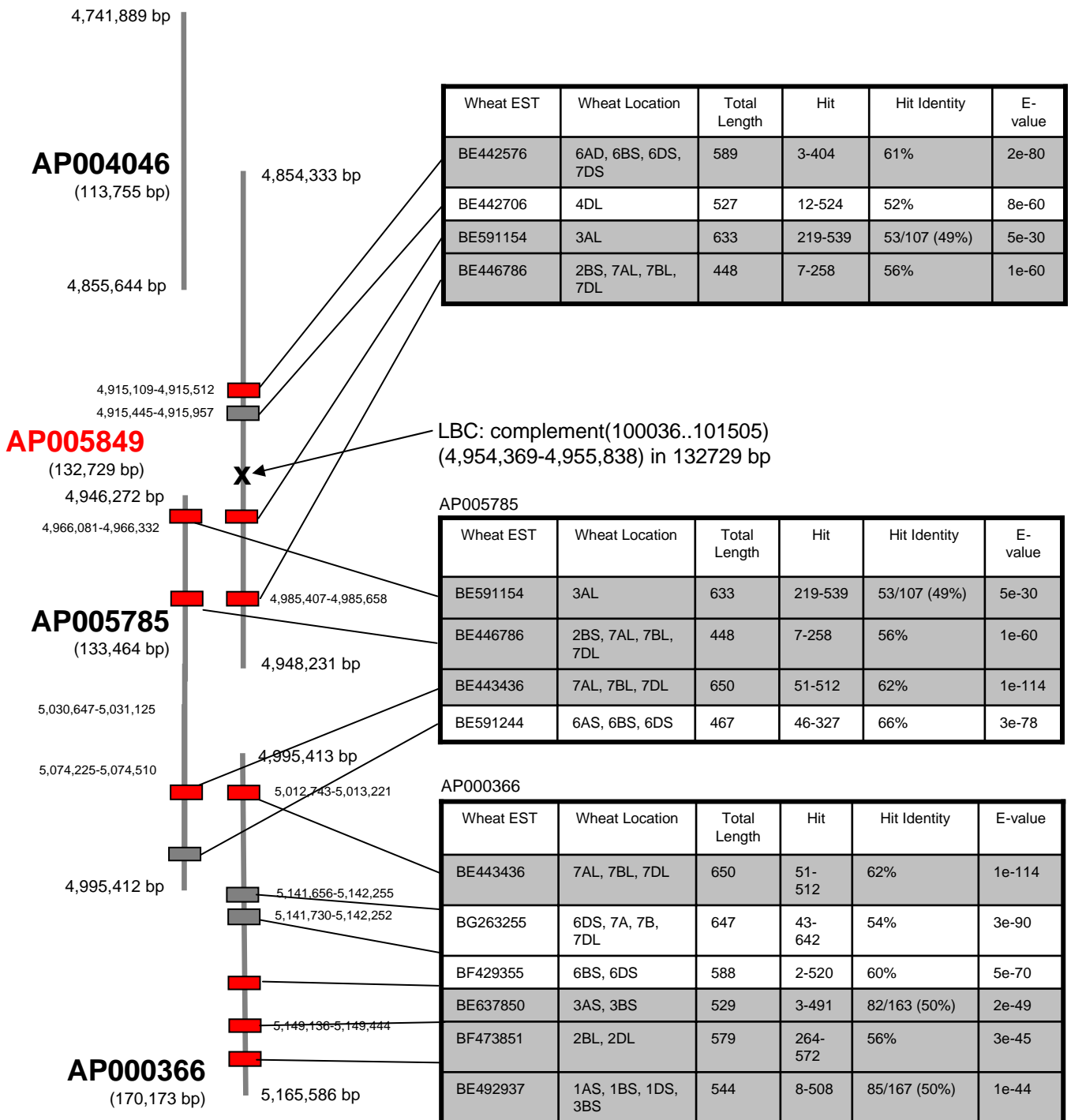


**Fig. 2.5** Alignment of mapped wheat ESTs to rice Geranylgeranyltransferase I  $\beta$ -subunit (AP002522) and flanking BAC/PAC clones on chromosome 1. The homology of the rice BAC/PAC clone and the aligned EST is detailed. Wheat ESTs mapped to groups 3 or 7 are shown in red.

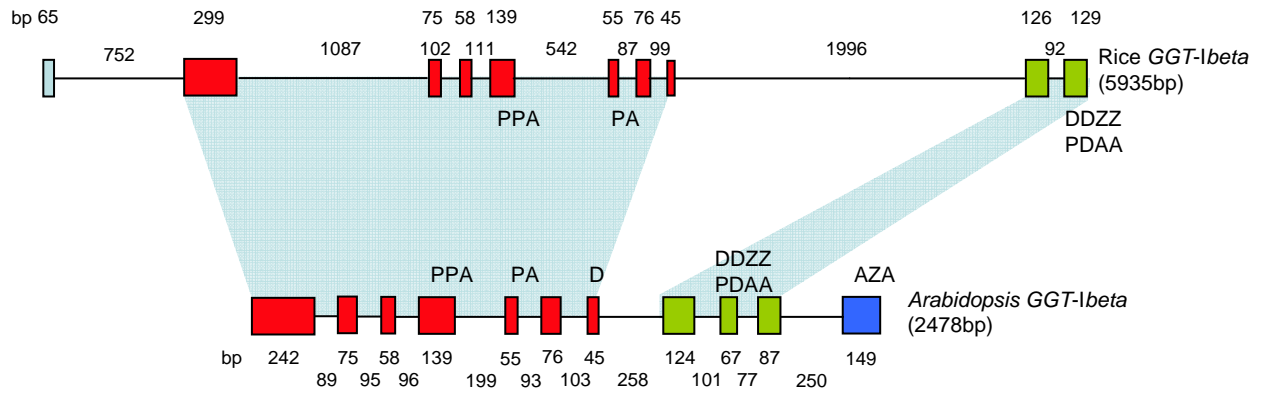


**Fig. 2.6** Alignment of mapped wheat ESTs to rice Lycopene  $\epsilon$ -cyclase (AP003332) and flanking BAC/PAC clones on chromosome 1. The homology of the rice BAC/PAC clone and the aligned EST is detailed. Wheat ESTs mapped to groups 3 or 7 are shown in red.

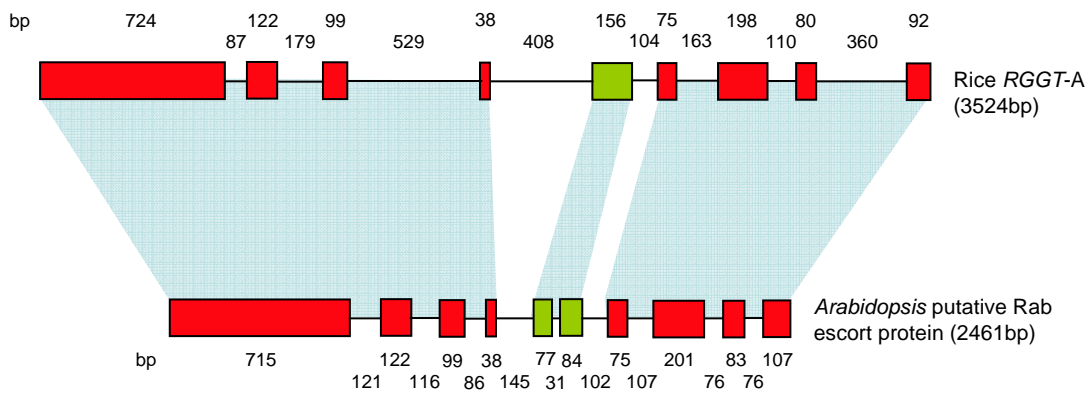
## Lycopene beta-cyclase (rice 2)



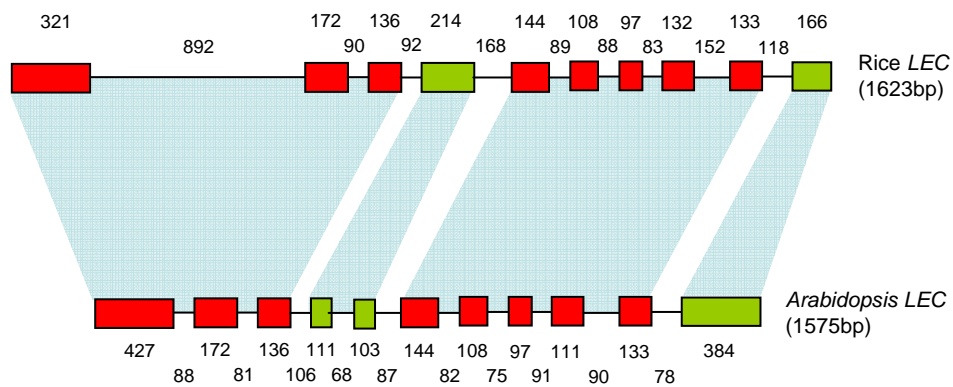
**Fig. 2.7** Alignment of mapped wheat ESTs to rice Lycopene  $\beta$ -cyclase (AP005849) and flanking BAC/PAC clones on chromosome 2. The homology of the rice BAC/PAC clone and the aligned EST is detailed. Wheat ESTs mapped to groups 3 or 7 are shown in red.



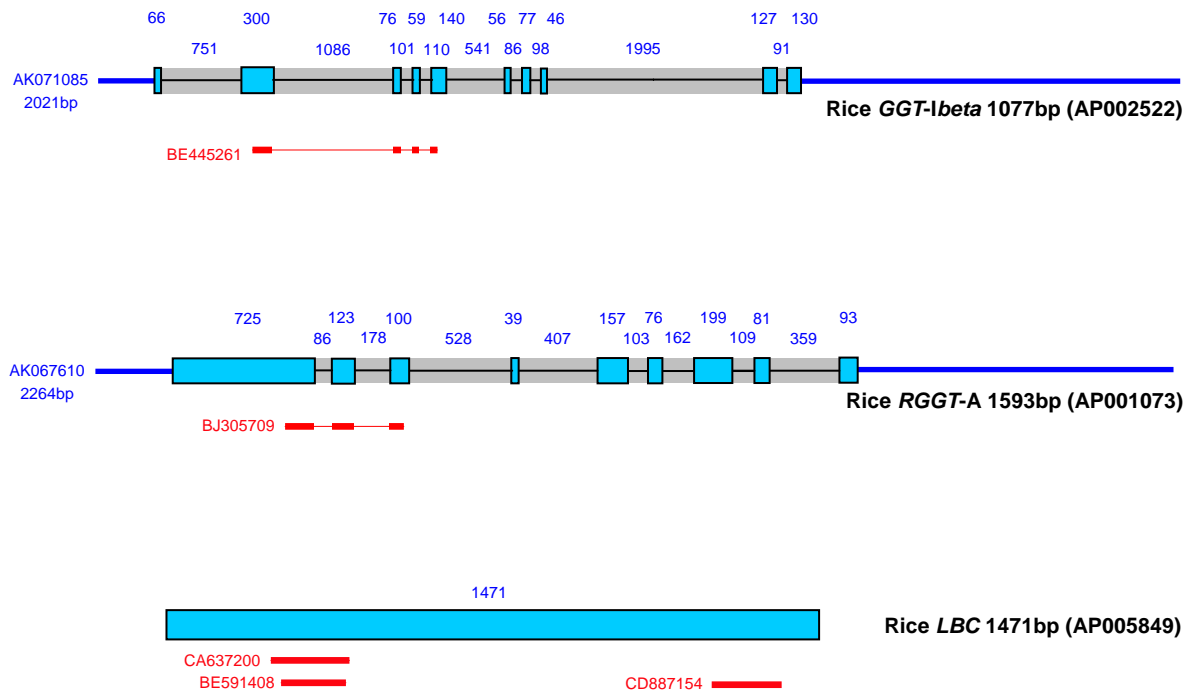
**Fig. 2.8** Intron-exon structure of rice (AP002522) and Arabidopsis GGT-Ibeta (AC004218) orthologues from cDNAs. Exons coloured in red are conserved in each species. Green exons represent a variable region between rice and Arabidopsis. Exons that are not conserved are shown in blue. Symbols indicated are A, residues in hydrophobic pocket; P, residues that interact with the CAAL peptide; D, residues that interact with the diphosphate group of GGPP; and Z, ligands of the catalytic zinc atom.



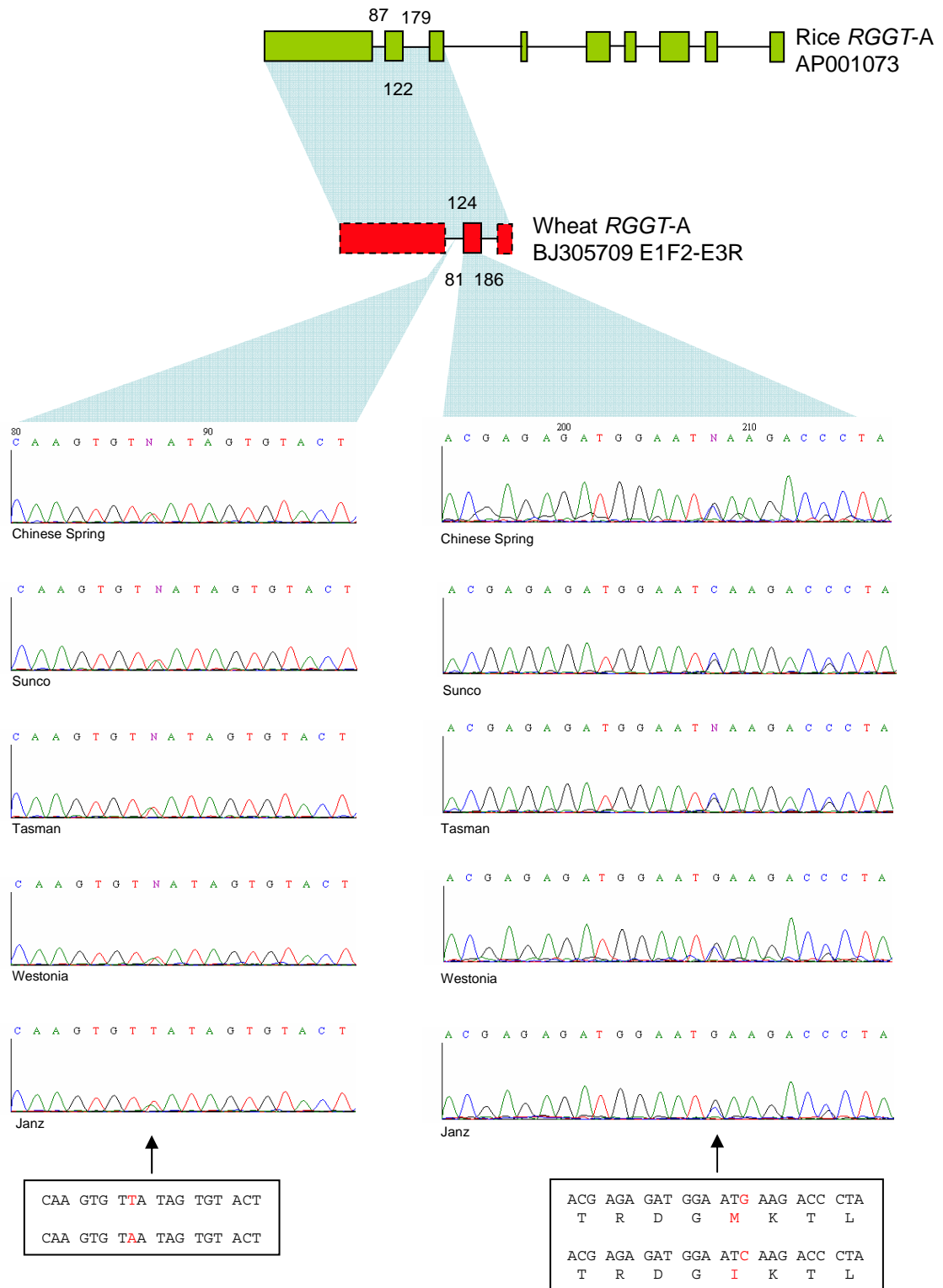
**Fig. 2.9** Intron-exon structure of rice RGGT-A (AP001073) and Arabidopsis putative Rab escort protein (AC020580) from cDNAs. Exons coloured in red are conserved in each species. Green exons represent a variable region between rice and Arabidopsis.



**Fig. 2.10** Intron-exon structure of rice (AP003332) and Arabidopsis LEC (At5g57030) orthologues from cDNAs. Exons coloured in red are conserved in each species. Green exons represent a variable region between rice and Arabidopsis.

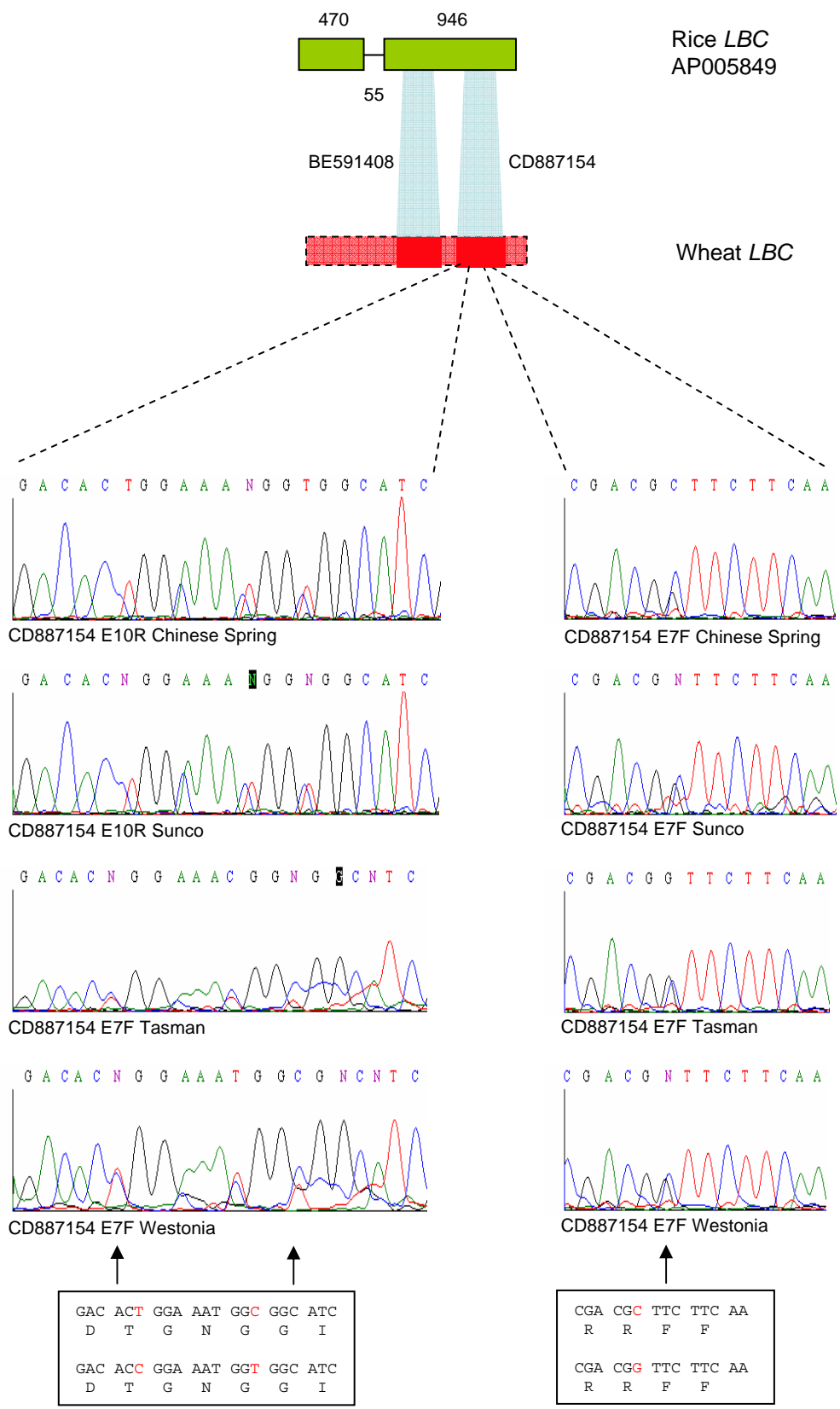


**Fig. 3.2** Summary of wheat ESTs that showed highest homology to the four candidate genes from rice. The wheat ESTs (red) are shown relative to the gene structure with thin red line indicating the EST spans an intron of the rice gene sequence. Exons are shown in cyan, introns are black lines and full-length cDNA clones are shown as dark blue lines.

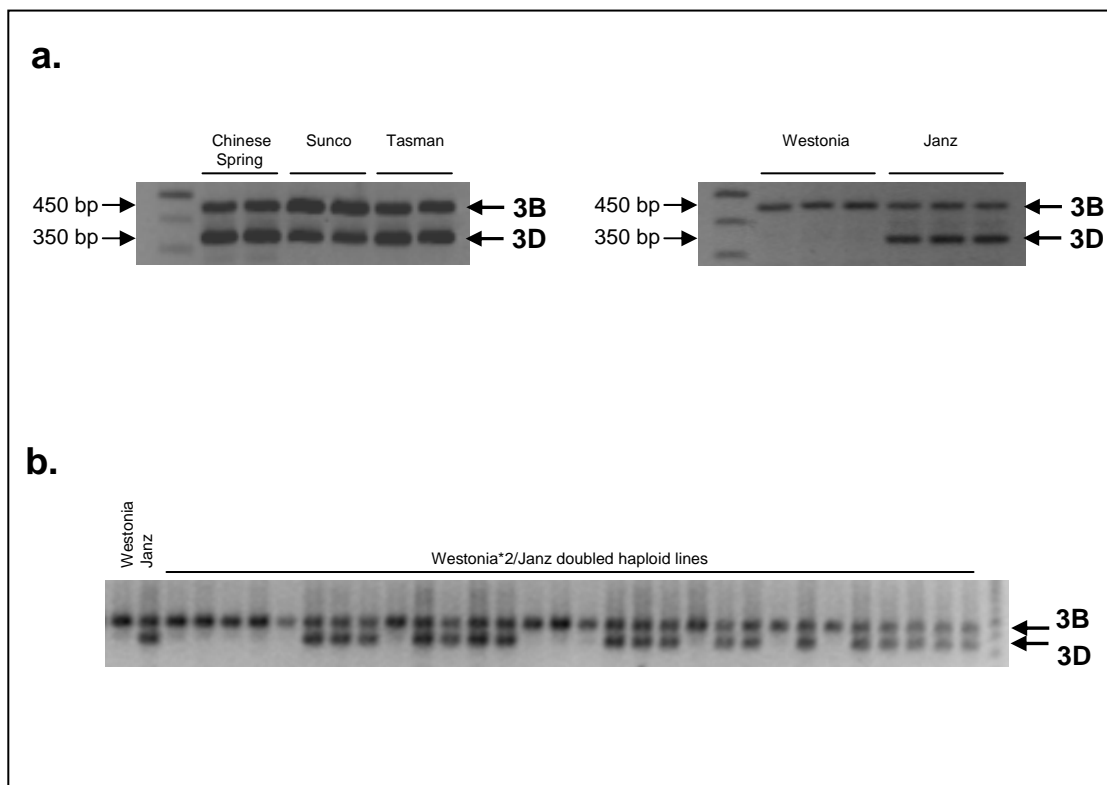


**Fig 3.10** Sequencing of wheat RGGT-A (intron 1-2, exon 2, intron 2-3) in cultivars Chinese Spring, Sunco, Tasman, Westonia and Janz using BJ305709 EIF-E3R. The gene structure of rice RGGT-A is shown at the top of the figure and below is EST BJ305709. Chromatograms of each cultivar are shown, arrows indicate genome-specific SNPs with translated protein sequence below.

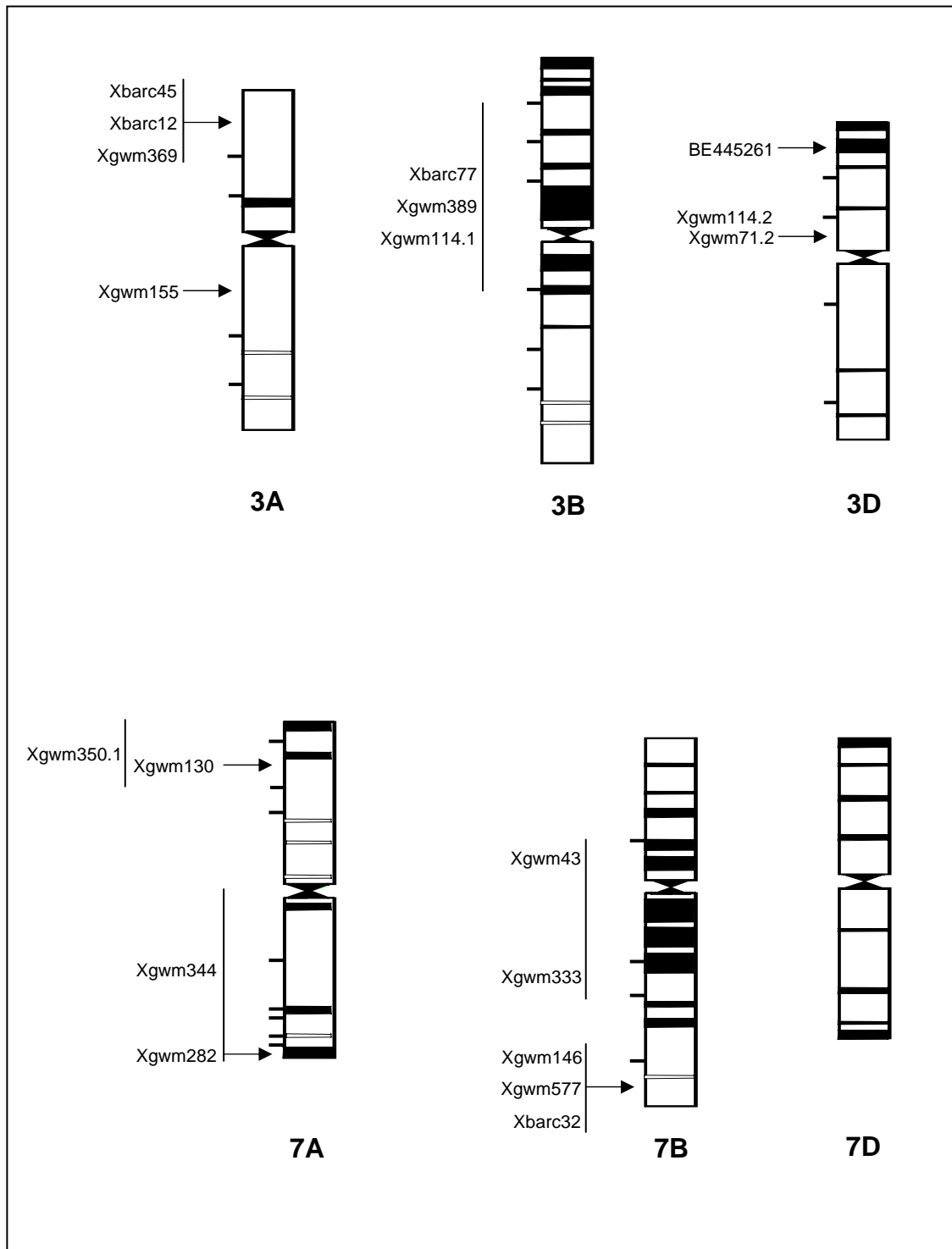




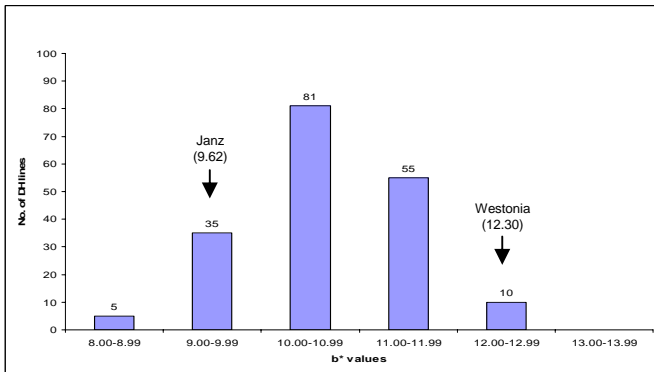
**Fig 3.17** Sequencing of wheat *LBC* in cultivars Chinese Spring Sunco, Tasman, Westonia and Janz using BE591408 E1F-E2R and CD887154 E7F-E10R. The intronless *LBC* gene of rice is shown at the top of the figure and below is the predicted location of the wheat orthologs BE591408 and CD887154 within the wheat gene sequence. Chromatograms of each cultivar are shown, arrows indicate genome-specific SNPs with translated protein sequence below.



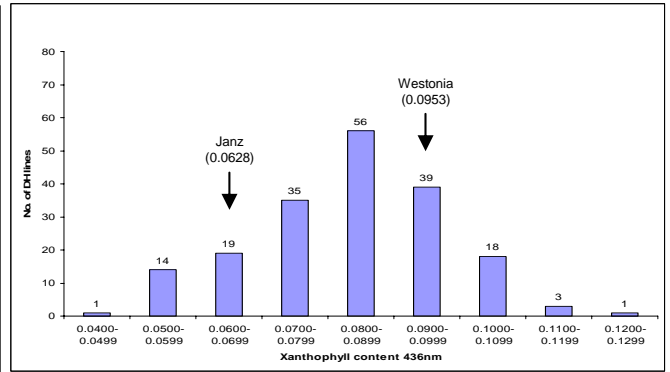
**Fig. 3.18** Polymorphic marker of GGT-Ibeta using primers E3F-E5R designed from wheat EST BE445261. (a.) Cultivars Chinese Spring, Sunco, Tasman and Janz contained both copies of GGT-Ibeta previously mapped to 3B and 3D. Westonia only contained the 3D copy of GGT-Ibeta as indicated by the amplification of the 450 bp band only. (b.) Screening of the Westonia\*2/Janz doubled haploid population showing segregation of the EST-based molecular marker. A subset of the population containing 30 of a total of 180 doubled haploid lines is shown.



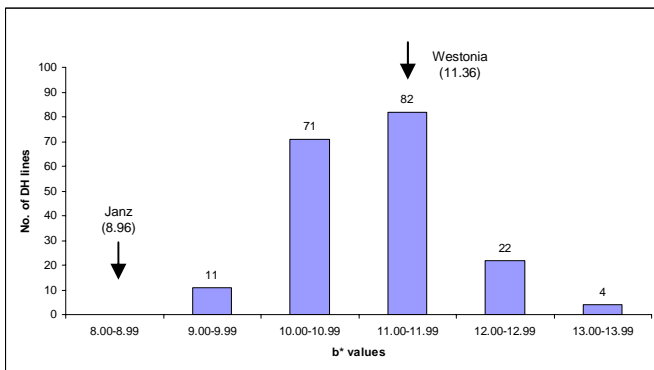
**Fig. 3.19** Polymorphic markers on groups 3 and 7 of *Westonia\*2/Janz* doubled haploid population. Vertical lines indicate linked markers and the location of the linkage groups. Arrows indicate the chromosomal bin location of a molecular marker.



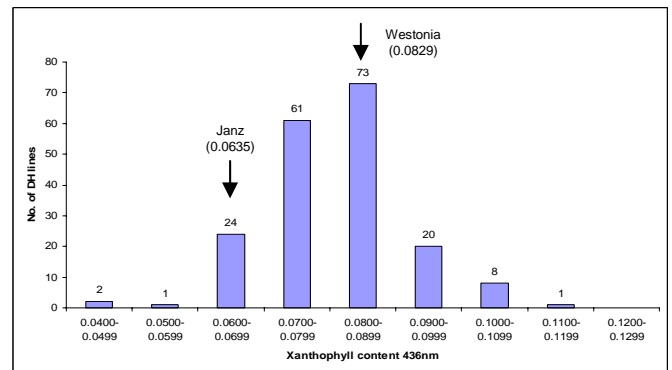
2002 Wongan Hills



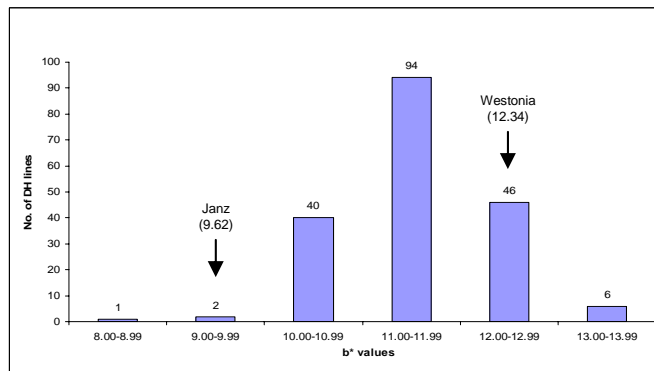
2002 Wongan Hills



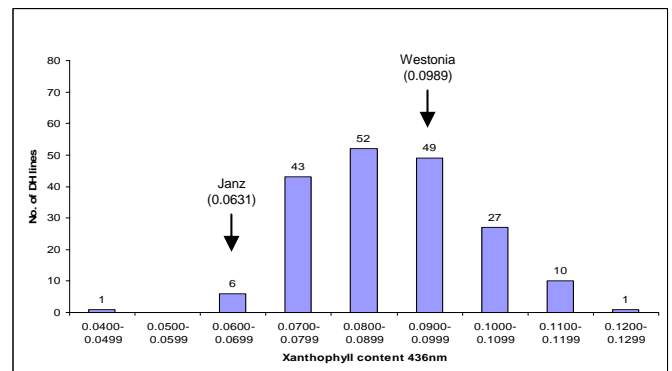
2003 Wongan Hills



2003 Wongan Hills

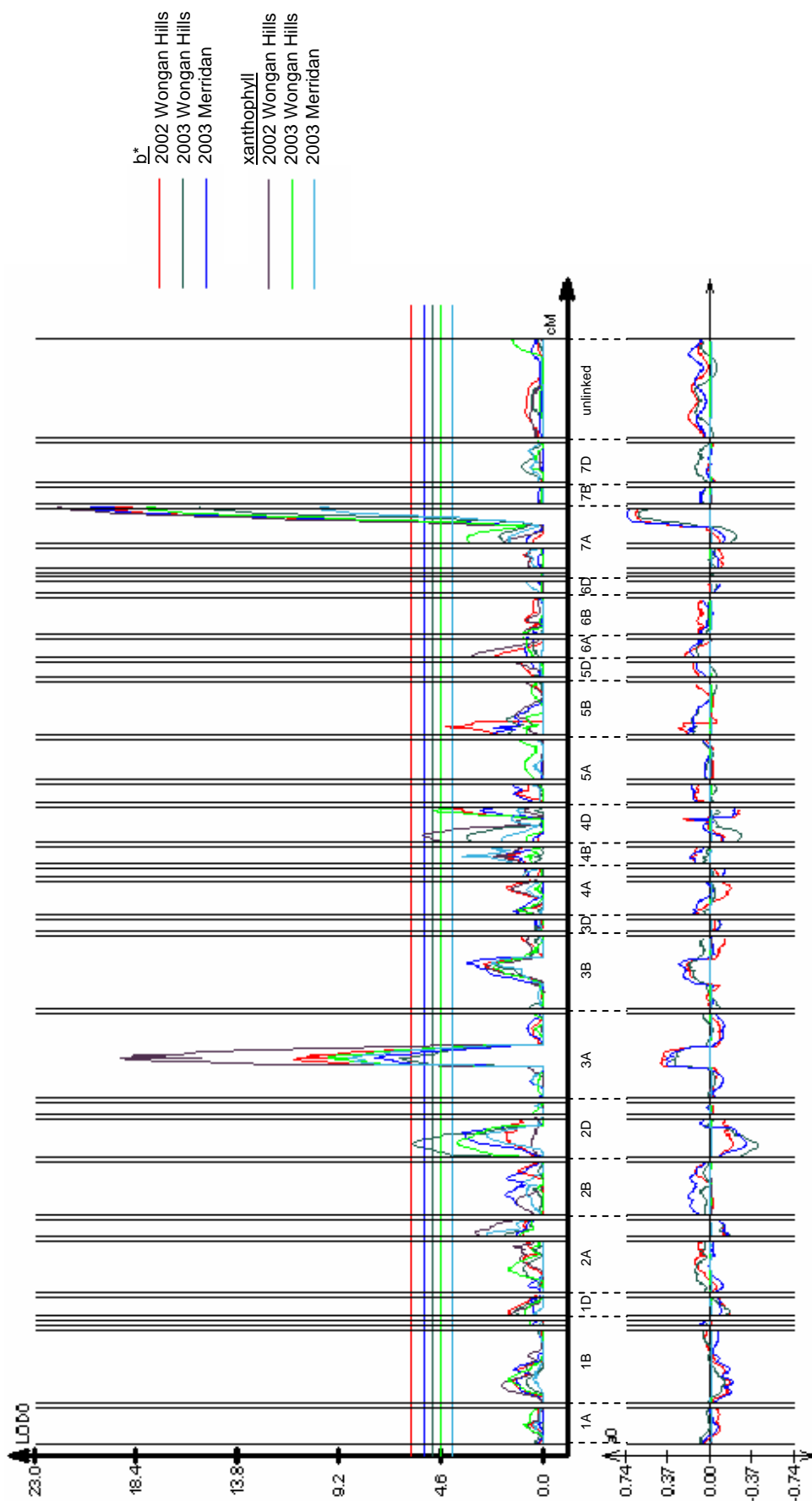


2003 Merridan

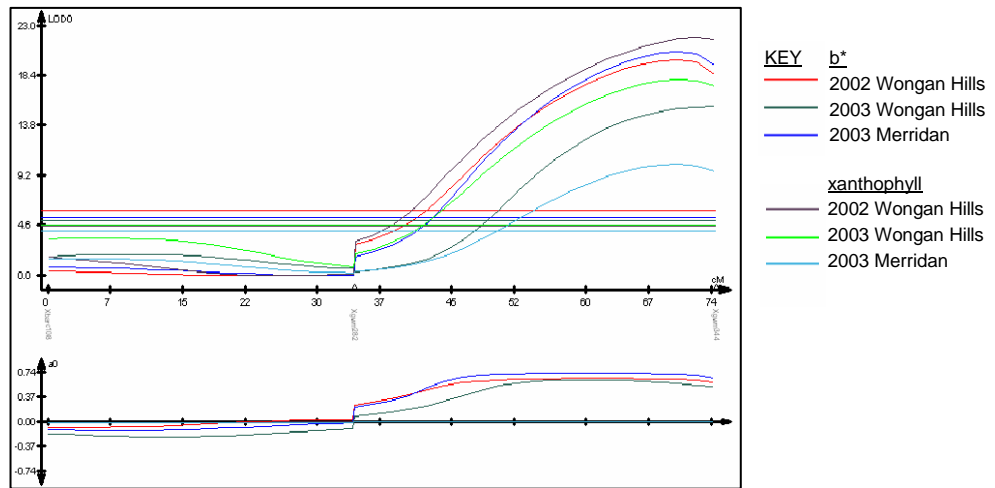


2003 Merridan

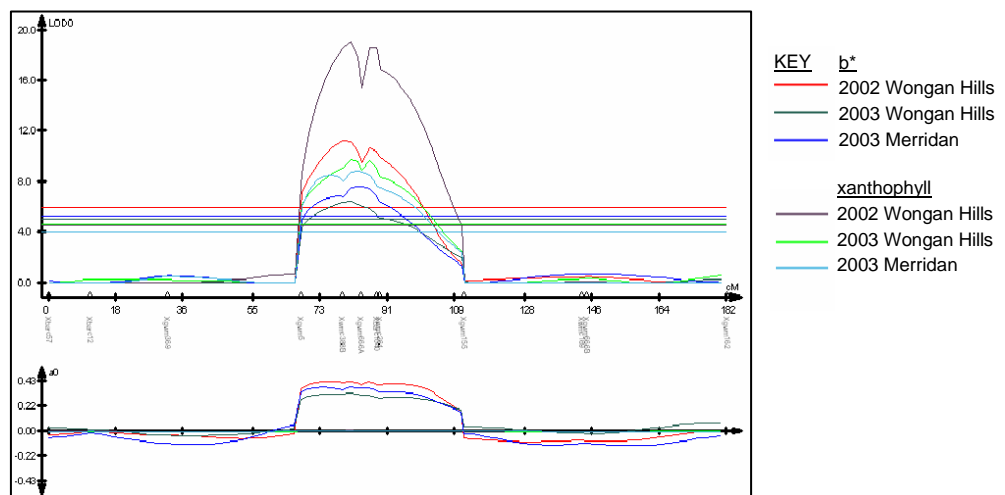
**Fig. 3.20** Histograms of frequency distribution of  $b^*$  values and xanthophyll in *Westonia*\*2/*Janz* doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. The  $b^*$  and xanthophyll values for *Westonia* and *Janz* at each site are shown and arrows indicate their distribution within the population.



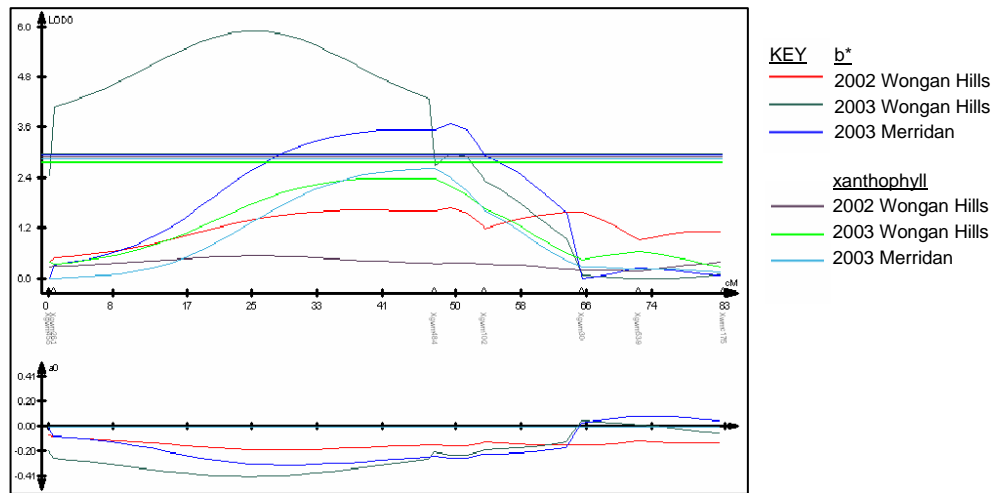
**Fig. 4.3** Summary of composite interval mapping for all chromosomes of *b\** values and xanthophyll in Ajana/WA WHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Walk speed 2cM, 1000 permutations and significance level of  $P > 0.001$ . The additive effect is shown below.



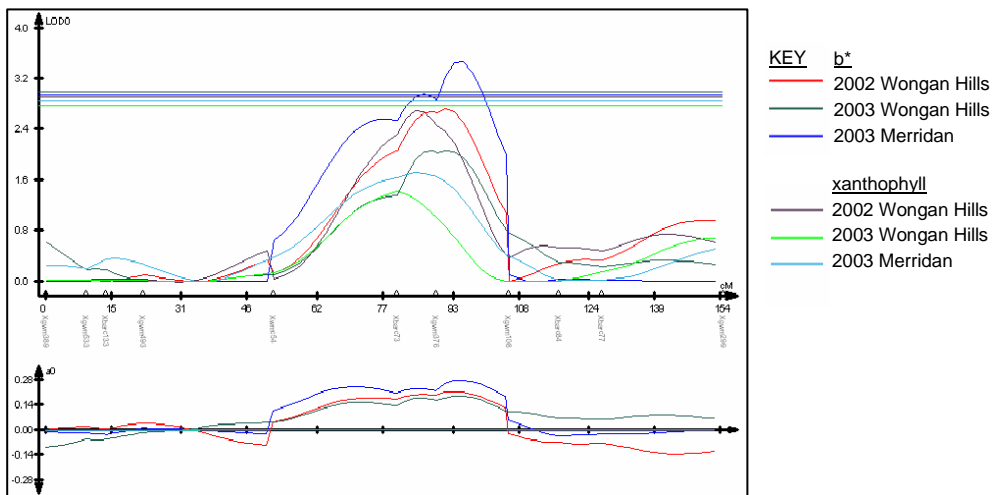
**Fig. 4.4** QTL of  $b^*$  and xanthophyll content on chromosome 7A in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.001$ , LOD thresholds shown as horizontal line. The additive effect is shown below.



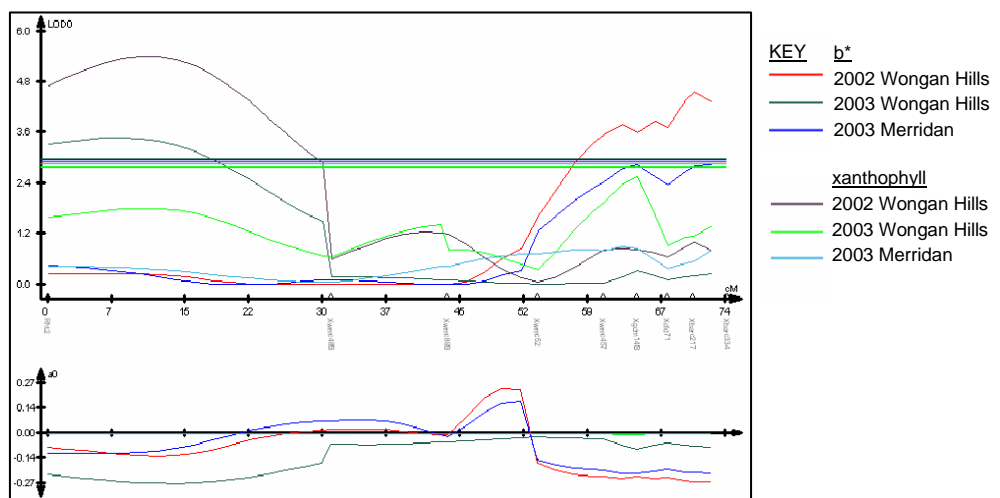
**Fig. 4.5** QTL of  $b^*$  and xanthophyll content on chromosome 3A in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.001$ , LOD thresholds shown as horizontal lines. The additive effect is shown below.



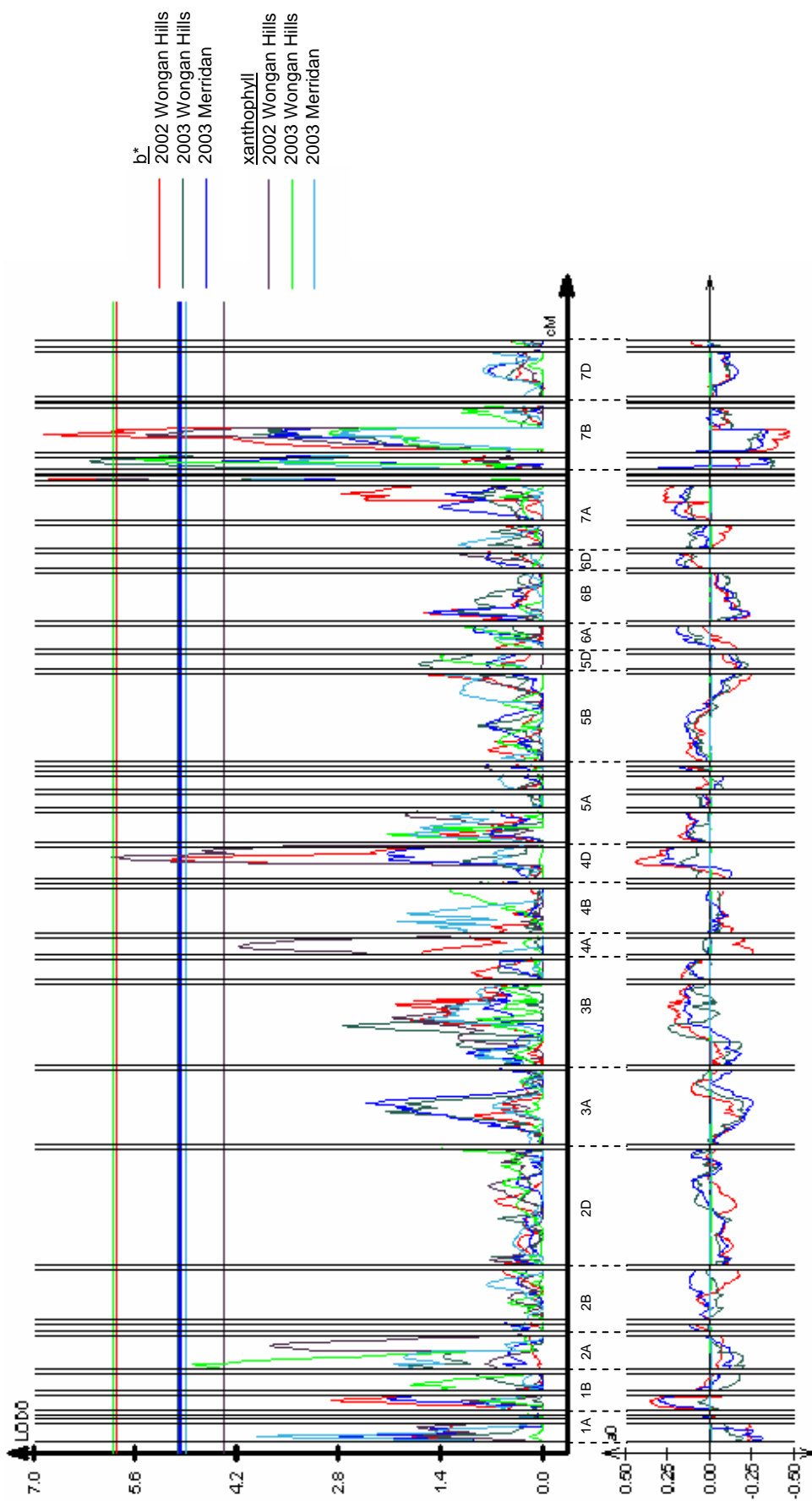
**Fig. 4.6** QTL of  $b^*$  and xanthophyll content on chromosome 2D in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal line. The additive effect is shown below.



**Fig. 4.7** QTL of  $b^*$  and xanthophyll content on chromosome 3B in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal line. The additive effect is shown below.

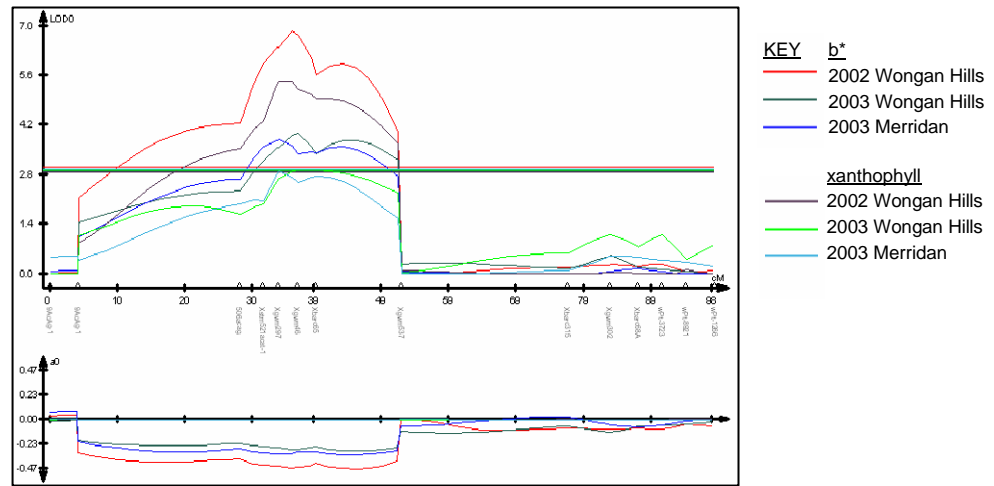


**Fig. 4.8** QTL of  $b^*$  and xanthophyll content on chromosome 4D in Ajana/WAWHT2074 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal lines. The additive effect is shown below.

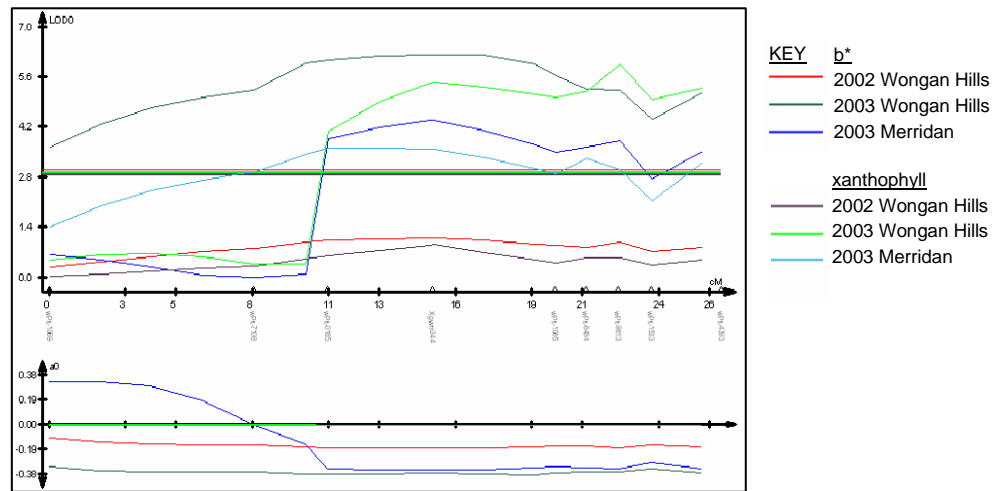


**Fig. 4.9** Summary of composite interval mapping for all chromosomes of  $b^*$  values and xanthophyll in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Walk speed 2cM, 1000 permutations and significance level of  $P > 0.001$ . The additive effect is shown below.

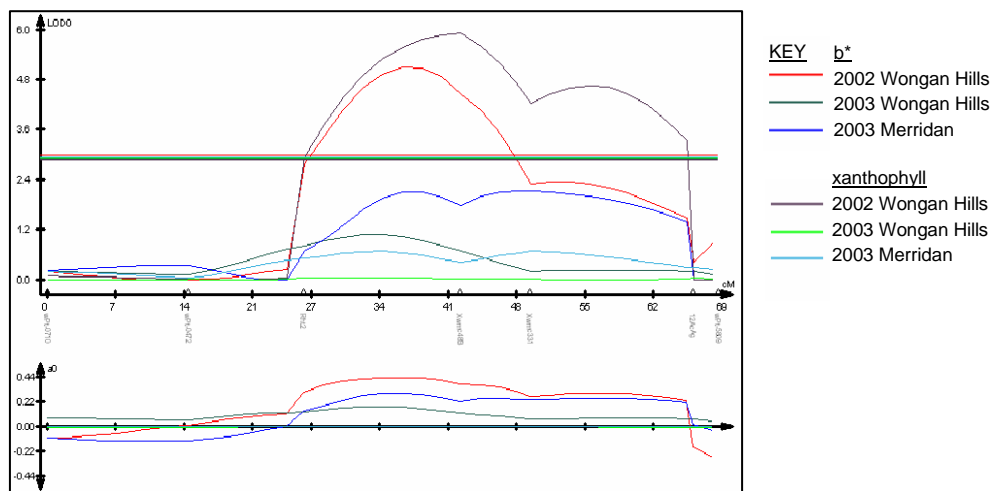




**Fig. 4.10** QTL of  $b^*$  and xanthophyll content on the short arm of chromosome 7B in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal lines. The additive effect is shown below.



**Fig. 4.11** QTL of  $b^*$  and xanthophyll content on the long arm of chromosome 7B in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal line. The additive effect is shown below.



**Fig. 4.12** QTL of  $b^*$  and xanthophyll content on chromosome 4D in Carnamah/WAWHT2046 doubled haploid population from the 2002 and 2003 sites of Wongan Hills and Merridan. Composite interval mapping: Walk speed 2cM, 1000 permutations and significance level of  $P > 0.05$ , LOD thresholds shown as horizontal lines. The additive effect is shown below.