DOCTOR OF PHILOSOPHY

THESIS

Identifying and dissecting novel QTL regulating seed fatty acid composition in *Arabidopsis thaliana* for the improvement of rapeseed oil

Chie Hattori

University of East Anglia

John Innes Centre

Crop Genetics Department

September 2012

To Yoshitaka

ABSTRACT

Genes encoding desaturases in the polyunsaturated fatty acid (PUFA) biochemical pathway have been identified by mutation analyses and utilized in oilseed rape *(Brassica napus*) breeding programs. However the regulation mechanisms of these genes are as yet unknown and pursued by the oil industry for further manipulation of seed oil quality.

In order to identify loci involved in such regulatory mechanisms, *Arabidopsis thaliana*, a close relative of *B. napus,* was used as a model plant. Seed oil characteristics such as PUFA composition and oil content are regarded as adaptative traits (Linder, 2000) which are typically the result of complex genetic components controlled by multiple loci. Therefore QTL analysis was performed using six recombinant inbred populations derived from wild accessions rich in natural variations.

QTL analysis identified loci involved in PUFA composition and oil content and 13 QTL were selected for the development of Near Isogenic lines (NIL). NILs were used to test the robustness of QTL by ANOVA and 11 QTL were successfully validated for their effect on phenotype in this work. Finally a QTL detected in the NG population for 18:2/18:1 was chosen for fine mapping and the region containing the causative locus for the QTL was narrowed down to 2Mb. Seven candidate genes within that region were identified which could be targeted in further fine mapping in the future.

In addition, a comparative mapping of QTL between *A. thaliana* and *B. napus* was performed through a unigene based *B. napus* map developed in the Bancroft group. This map facilitated the integration of results found in *A. thaliana* with *B. napus* and allowed the detection of coincident QTL regions between these two species and candidate gene analysis within these regions. The genes *MTACP2, PKP-BETA1, DGAT1* and *LTP4* were identified repeatedly during the analysis and are likely candidate genes affecting seed oil and could be target genes for future breeding.

Thesis total words count =43723

ACKNOWLEDGEMENTS

I would like to thank my primary supervisor Ian Bancroft for letting me join his group and giving me the opportunity to work on this project. He has a good sense of humour, and is generous and very supportive throughout the whole four years. His door is always opened to everyone for discussion. He also is a master of carrot and stick approach. What else could one ask from their boss!

My very special thanks go to Carmel O'Neill who is my daily base supervisor. She generously spent her time for me in spite of the fact that she works part time at the moment. She disciplined me as a scientist in every manner. Discussion with an enthusiastic scientist like her was very inspiring and made me feel positive. During the long PhD period, everyone has occasional ups and downs. When I was in the bottom of down, she cheered me up saying "Chie, the bottom line is nobody is dead!". I will take this message with me for the rest of my life.

I also thank Colin Morgan and Rachel Wells. Since they were supervisors of my MSc project, I took advantage of their subliminal feeling that looking after me is still their duty. They were very helpful in every aspect and gave me lots of useful advice and contributions, especially in *Brassica napus* mapping.

Martin Trick, Andrea Harper and Antony Colles supported this project from the bioinformatics side and comparative mapping. I learned a lot from them and their contributions were also fundamental for this project.

Fiona Fraser and Charlotte Miller helped me when I needed it in my experiments in the lab and the glasshouse.

Lesley Phillips facilitated my three year glasshouse experiment greatly. Richard Goram kindly worked even weekends to get the job done in time, especially when screening large scale mapping populations. Frederic Beaudoin performed GC FID for my FAME samples to perfection which was critical for this project.

I also would like to thank my other supervisors, Simon Griffiths and Peter Werner for their sincere attention to my project at the supervisory meetings. Peter also provided me with the opportunity to experience work in the breeding industry at KWS.

I would like to thank sponsors of my CASE studentship for their generosity, which is provided by BBSRC, KWS, Saaten Union and Monsanto.

I would like to mention my friends at JIC, especially lunch mates Anastasia Gardner, Cintia Kawashima, Evelyn Koerner and Kate Bailey (alphabetical order) who amused me by their wide range of topics everyday.

The coffee room on Biffen first floor is a precious place where I could relax and where I could also seek scientific advice casually amongst the chitchat from someone like Judith Irwin.

I would like to thank our parents in Japan who understand the importance of what I am doing and have been supportive although I couldn't go to see them as often as a normal daughter or daughter-in-law should do because of the work.

And last of all I truly appreciate the encouragement by Yoshitaka Sawada, my dedicated husband who always believed that I can do it. He was supportive mentally and practically both at home and at JIC literally, which is a well known fact among our friends. I am looking forward to embarking on the next step of my life with him after my days at JIC.

TABLE OF CONTENTS

Abstract 2
Acknowledgements 3
List of tables 9
List of figures 10
Chapter 1. Introduction 11
1.1 Oilseed rape 11
1.1.1 *Brassica napus* 11
1.1.2 History of cultivation 11
1.1.2.1 Ancient history 11
1.1.2.2 Early history in Europe 11
1.1.2.3 History of use in the UK 12
1.1.2.4 The current situation 13
1.1.3 Use of *B. napus* 13
1.2 *Arabidopsis thaliana* 15
1.2.1 Use of *A. thaliana* as a model plant 15
1.2.2 Natural variation in *A. thaliana* 16
1.2.2.1 Adaptive variation 16
1.2.2.2 The advantages of using a natural variation approach over a mutation approach 17
1.2.3 Origin of plant material used for this work 20
1.3 Fatty acids 22
1.3.1 Fatty acid biosynthesis 22
1.3.1.1 16:0-ACP and 18:1- ACP synthesis in the plastid 23
1.3.1.2 Glycerolipid synthesis and desaturation activity 24
1.3.1.3 Formation of TAGs for seed storage in the endoplasmic reticulum 26
1.4 Key genes involved in fatty acid desaturation 27
1.4.1 *FATTY ACID BIOSYNTHESIS 1 (FAB1)* 28
1.4.2 *FATTY ACID BIOSYNTHESIS 2 (FAB2)* 28
1.4.3 *FATTY ACID DESATURATION 2 (FAD2)* 28
1.4.4 *FATTY ACID DESATURATION 3 (FAD3)* 29
1.4.5 *FATTY ACID ELONGATION 1(FAE1)* 29
1.5 QTL Mapping 31
1.5.1 Linkage maps 31
1.5.2 QTL analysis 32
1.5.3 Mapping populations 33
1.5.3.1 Recombinant inbred lines 33
1.5.3.2 Near isogenic lines 33

Chapter 2. Hypothesis aims and objectives 35
2.1 Hypothesis 35
2.2 Aims 35
2.3 Objectives 35
Chapter 3. General materials and methods 36
3.1 Plant material 36
3.2 Plant husbandry 36
3.3 Crossing 38
3.4 Harvesting seeds 39
3.4.1 Silique collection 40
3.4.2 Bulk collection 40
3.4.3 Cleaning of seeds for phenotyping 41
3.5 Sampling leaves for DNA extraction 42
3.6 Fluorescent labelled high throughput genotyping 43
3.6.1 PCR method 43
3.6.2 Preparation of samples for genotyping down the ABI 3730xl sequencer 45
3.6.3 Scoring the genotype results 46
3.7 Measuring fatty acid composition 47
3.7.1 Fatty acid methyl ester preparation 47
3.7.2 Gas chromatography analysis of FAMEs 50
3.7.3 Preparation of data 50
3.8 Measuring oil contents 51
Chapter 4. Identification and refinement of novel QTL controlling fatty acid profile in seeds of *A. thaliana* 53
4.1 Introduction 53
4.2 Material and methods 56
4.2.1 Plant material 56
4.2.2 Phenotypic data 56
4.2.3 Increasing marker density 57
4.2.4 Linkage mapping 57
4.2.5 QTL analysis 57
4.2.6 Selection of target QTL for NIL development 58
4.3 Results 59
4.3.1 Statistical analysis of phenotype 59
4.3.1.1 Parental lines 59
4.3.1.2 RI Populations 61
4.3.2 Linkage analysis 61
4.3.3 Interval mapping 64
4.3.3.1 18:1 64

4.3.3.2 18:2 66
4.3.3.3 18:2/18:1 68
4.3.3.4 18:3 69
4.3.3.5 18:3/18:2 70
4.3.3.6 oil content 71
4.3.4 MQM mapping 71
4.3.5 Sequence variations at major gene loci 72
4.3.6 Evaluation of phenotype associated with QTL close to major genes 74
4.4 Conclusion 75
4.4.1 Phenotypic data 75
4.4.2 Identification and examination of QTL close to major genes 75
4.4.3 Selection of target QTL 78
Chapter 5. Validation of QTL by near isogenic lines 82
5.1 Introduction 82
5.2 Materials and methods 85
5.2.1 Selection of RIL for NIL development 85
5.2.2 First backcrossing (B_1 development) 87
5.2.3 Second backcrossing (B_2 development) 88
5.2.4 First selfing (B_2S_1development) 89
5.2.5 Final stage of NIL development ($B_2S_1B_1$ development) 90
5.2.5.1 Experimental design 91
5.2.6 Measuring fatty acid compositions 96
5.2.7 Measuring seed oil content 97
5.2.8 Statistical analysis to test the robustness of QTL 97
5.2.9 Criteria for selection of NIL from NIL candidates 97
5.2.10 Selection of QTL for fine mapping 98
5.3 Results 99
5.3.1 Selection of RILs as starting materials 99
5.3.2 Examining the quality of data 100
5.3.2.1 G x E interaction 100
5.3.2.2 Block effects 101
5.3.2.3 Experimental errors 101
5.3.3 Examining the effect of introgressed segments 103
5.3.4 Selection of NIL from NIL candidates 104
5.3.4.1 18:2 (TJ25) 104
5.3.4.2 18:2/18:1 (NG28, NG32 and TJ79) 105
5.3.4.3 18:3 (TJ17) 106
5.3.4.4 18:3/18:2 (NG32, NG93 and SG10) 106
5.3.4.5 Oil content (KB55, TJ79 Ch2, TJ79 Ch5 and TJ83) 107

5.3.5 Selection of final NIL for fine mapping 109
5.4 Conclusion 111
5.4.1 NIL and HIF 111
5.4.2 Block effects 111
5.4.3 Three purposes of NIL developments 112
Chapter 6. relative content of linoleic acid /oleic acid and the determination of candidate genes 114
6.1 Introduction 114
6.2 Material and methods 116
6.2.1 Plant material 116
6.2.2 Plant husbandry 116
6.2.3 First screening 119
6.2.4 Binning 119
6.2.4.1 Primer design for binning 120
6.2.4.1.1 Sequencing of parental lines 120
6.2.4.1.2 InDel markers and primer design 120
6.2.4.2 Selection of markers 121
6.2.4.3 PCR and genotyping 122
6.2.4.4 Sorting into bins 122
6.2.5 Measuring fatty acid composition of mapping populations 123
6.2.6 Statistical analysis of phenotypic data 125
6.2.6.1 Checking quality of data 125
6.2.6.2 Identifying the position of the QTL causative locus 125
6.3 Results 126
6.3.1 First screening 126
6.3.2 Second screening for sorting plants into 48 bins 126
6.3.3 Phenotype data of mapping populations 127
6.3.3.1 Experimental error and environmental effects 127
6.3.3.2 Statistical summary of 18:2/18:1 129
6.3.3.3 Identifying the bin containing the QTL causative locus 130
6.3.3.4 Dissecting the phenotype of population(34) 132
6.3.4 Investigation of the interaction between *FAD8* 134
6.3.5 Candidate gene analysis 136
6.4 Conclusion 137
6.4.1 Population(33) and (34) 137
6.4.2 Fine mapping 137
6.4.3 Causative of QTL 139
6.4.4 Candidate genes 140

Chapter 7. Comparative analysis of QTL for seed lipid traits in *A. thaliana* and *B. napus* 142
7.1 Introduction 142
7.2 Material and methods 147
7.2.1 Mapping *A. thaliana* QTL onto the *B. napus* unigene based map 147
7.2.2 Comparison with the *B. napus* TNDH population 147
7.2.3 Comparison with previous works identified in the literature 148
7.3 Results 149
7.3.1 Alignment of *A thaliana* QTL with the unigene based *B. napus* map 149
7.3.1.1 Alignment of all AGI genes found within the confidence intervals 150
7.3.1.2 Alignment of AGI genes related to lipid synthesis and fatty acid desaturation pathway 150
7.3.2 Comparison of *A. thaliana* QTL with QTL detected in the *B. napus* TNDH population 151
7.3.2.1 QTL analysis for 18:1, 18:2, 18:3 as carried out by Colin Morgan (unpublished results) 151
7.3.2.2 QTL analysis for oil content by Qiu et al. 2006 154
7.3.3 Comparison of *A. thaliana* QTL with *B. napus* QTL mapped in different populations 155
7.3.3.1 QTL for oil content using the SGDH population developed by Zhao et al. 2012 155
7.4 Conclusion 159
7.4.1 Alignment of QTL 159
7.4.2 Comparative mapping 159
Chapter 8. General discussion 164
8.1 Summary of thesis 164
8.2 General discussion 166
8.3 Future work 172
References and bibliography 174
Appendices 182

LIST OF TABLES

Page	Table no.	Title of table
21	1-1	Origin and lipid characteristics of parental lines
57	4-1	Number of markers in population
60	4-2	The means and standard deviations of phenotype of the twelve parental lines and heritabilities of the six populations
63	4-3	Marker density of linkage group for three populations
64	4-4	List of the significance thresholds of LOD scores for six traits in six populations
65	4-5-a	List of QTL for 18:1
67	4-5-b	List of QTL for 18:2
68	4-5-c	List of QTL for 18:2/18:1
69	4-5-d	List of QTL for 18:3
70	4-5-e	List of QTL for 18:3/18:2
71	4-5-f	List of QTL for oil content
80	4-6	List of target QTL
96	5-1	Example of a FAME analysis sample list
100	5-2	Details of target QTL and selected RILs for NIL development
100	5-3	Gxblock interaction
101	5-4	*P* values between blocks
102	5-5	Experimental errors
103	5-6	Statistical differences between NIL candidates and MH plants
119	6-1	Four types of recombination patterns to be selected at the first screening
122	6-2	List of selected InDel markers and their PCR products sizes
124	6-3	Sets of samples of mapping populations for FAME analysis
126	6-4	Number of plants in four recombination types
127	6-5	Numbers of the different recombination type plants across the bins for both mapping populations
129	6-6	Average of 18:2/18:1 for each bin
131	6-7	Summary of identified shift differences
136	6-8	Regions covered by bins 6,7and 8
141	6-9	AGI gene code and the physical positions of candidate genes
146	7-1	Details of *A thaliana* QTL used for comparative mapping
149	7-2	List of AGI codes related to seed lipid biosynthesis and FA desaturation for seven QTL detected in *A. thaliana*
158	7-3	Summary of *Arabidopsis* gene model found within coincident region of B. napus and A. thaliana QTL

LIST OF FIGURES

Page	Figure no.	Title of figure
11	1-1	Genetic relationships of the *Brassica* species (modified from U 1935)
14	1-2	Diversity of end uses of rapeseed oil and desired oil profiles
16	1-3	Evolutional history of *Brassica* and *Arabidopsis* (Yang et al., 1999; O'Neill and Bancroft, 2000; Schmidt et al., 2001; Inaba and Nishio, 2002; Rana et al., 2004)
19	1-4	Schematic of natural null alleles (NNA) in *A. thaliana* genome in a natural variation approach and a mutation approach using Col.
24	1-5	Reaction of saturated fatty acid biosynthesis (Ohlrogge and Browse, 1995)
25	1-6	Diagram of FA synthesis and glycerolipid assembly in A. thaliana leaves showing the downstream effects of mutations (Browse, 1991)
27	1-7	Major genes controlling PUFA compositions. (Taken from Browse, 1991)
62	4-1-a	Linkage map of the CA population
62	4-1-b	Linkage map of the SG population
63	4-1-c	Linkage map of the WC population
81	4-2	Target QTL and major lipid genes physically positioned (Mb)
84	5-1	Schematic outlining NIL development
86	5-2	Example of charts for RIL selection for NIL development
92	5-3-a	Glasshouse bench positions of NIL candidates, MH plants and parental lines of the first group prior to the first screening
93	5-3-b	Glasshouse bench positions of NIL candidates, MH plants and parental lines of the first group after the first screening
94	5-4-a	Glasshouse bench positions of NIL candidates, MH plants and parental lines of the second group prior to the first screening
95	5-4-b	Glasshouse bench positions of NIL candidates, MH plants and parental lines of the second group after the first screening
114	6-1	Schematic images of NIL mapping populations development
117	6-2	Tray layouts of mapping populations in Room 2 of the West 1 glasshouse
118	6-3	Tray layouts of mapping populations in Room 3 of the West 1 glasshouse
120	6-4	Schematic of binning strategy in four different recombination types
123	6-5	Examples of binning in four different recombination types
128	6-6	Average values of 18:2/18:1 for the control material of both mapping populations
130	6-7	Identified shift differences between bins in population(33) for 18:2/18:1
131	6-8	Identified shift differences between bins in population(34) for 18:2/18:1
132	6-9	Identified shift differences between bins in population(34) for 18:1
133	6-10	Identified shift differences between bins in population(34) for 18:2.
134	6-11	Identified shift differences between bins in population(34) for 18:3.
144	7-1	Example of Brassica unigenes aligned to A. thaliana AGI gene models (Harper et al., 2012)
144	7-2	Flow chart showing the development procedures of the unigene based map of *B. napus* (pseudomolecules map) constructed from (Trick et al., 2009; Bancroft et al., 2011; Harper et al., 2012)
145	7-3	Schematic image of the association between QTL identified in *A. thaliana* (circled in red) and QTL detected in *B. napus* (circled in pale orange) with the unigene based *B. napus* map
151	7-4	Example of an *A. thaliana* QTL (18:2) mapped onto the unigene based *B. napus* map
152	7-5-a	QTL for 18:1 in the *B. napus* TNDH population (markers with LOD>2.0)
153	7-5-b	QTL for 18:2 in the *B. napus* TNDH population (markers with LOD>2.0)
153	7-5-c	QTL for 18:3 in the *B. napus* TNDH population (markers with LOD>2.0)
155	7-6	QTL for oil in the *B. napus* TNDH population ($P<0.05$)
156	7-7	Two QTL on C8 in SGDH population as taken from (Zhao et al., 2012)
157	7-8	QTL detected in the SGDH population mapped onto the unigene based *B. napus* map constructed from the TNDH population

CHAPTER 1. INTRODUCTION

1.1 OILSEED RAPE

1.1.1 *BRASSICA NAPUS*

The *Brassicaceae* family comprises approximately 338 genera and 3709 species, which include many economically valuable species. The *Brassica* genus belongs to this large family and there are 35 species, most are annual or biannual herbaceous plants. All parts of the plant are economically useful including roots, stems, leaves, inflorescence and seeds (Weiss, 1983; Lysak and Koch, 2011). Cultivated *Brassica* species are represented by six inter-related species, three diploids; *B. rapa*, *B. nigra*, and *B. oleraceae* and three amphidiploids; *B. juncea*, *B. napus* and *B. carinata*. These amphidiploids are derived from the hybridization of two of the diploid species as shown in Fig.1-1.

B. napus is the most common rapeseed plant grown in Europe, Canada, China and Australia. It is the same species as swede but is a non-bulbing form. Winter *B. napus* requires a prolonged period of low temperature to activate flowering, while spring types do not (Kimber and McGregor, 1995).

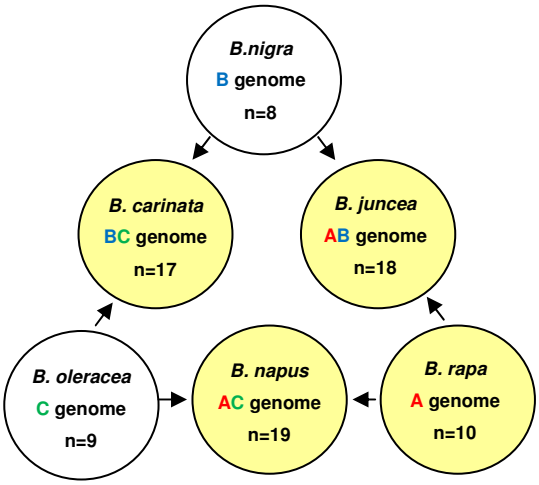


Fig.1-1 Genetic relationships of the *Brassica* species (modified from U 1935)

Filled yellow circles highlight those species with some cultivars grown for oil.

1.1.2 HISTORY OF CULTIVATION

1.1.2.1 ANCIENT HISTORY

The origin of rapeseed is still uncertain, however the Eurasian region is the most likely area of origin as the oldest known references to its cultivation have been discovered in India, China and Japan as early as 2000 BC (Weiss, 1983).

1.1.2.2 EARLY HISTORY IN EUROPE

Cultivation of rapeseed in Europe is thought to have begun in the thirteenth century and rapeseed oil was one of the important sources of lamp oil in northern Europe at that time. During the latter part of the middle ages, it was also used for manufacturing soap. Production in the Netherlands appears to have been restricted up until the sixteenth century at which time it was used as a pioneer crop in land reclamation projects (Weiss, 1983).

Seeds were subsequently introduced from the Netherlands to Germany, France and other European countries (Scarisbrick and Daniels, 1986). Practical instructions for the cultivation of rapeseed were published in Germany in 1743. During this period the oil was still used predominantly for lighting purposes and only the poor used it for edible purposes (Weiss, 1983).

1.1.2.3 HISTORY OF USE IN THE UK

The introduction of rapeseed into England is thought to have occurred in a similar manner to that of the Netherlands i.e. as part of land reclamation schemes in the fenland areas of eastern England. The first reference to rapeseed in England is a bill introduced in parliament in May 1572 detailing the preparation of oil from seeds grown in England (Scarisbrick and Daniels, 1986). There is a record of an indigenous seed crushing industry in the seventeenth century which satisfied the domestic demand and also produced surplus oil for export (Weiss, 1983).

However, due to competition with other oilseeds, such as Russian linseed and different oilseeds from Africa and Asia, the production of rapeseed oil declined and by 1850, mineral oil replaced rapeseed oil for lighting and lubrication purposes (Scarisbrick and Daniels, 1986).

In recent times formal agronomic trials started in 1948, although the scale and acreage of production was limited. In 1960 the European Economic Community (EEC), which the UK joined in 1973, was established. The EEC invested funds into research for breeding of improved cultivars with herbicide and pesticide resistance, training farmers in modern specialised cultivation techniques, and also supported farmers financially in the event of price fluctuation. The breeding strategy at that time concentrated on developing varieties with low levels of erucic acid, less than 5% in seed for edible oil use and this objective was achieved by 1974 (Scarisbrick and Daniels, 1986).

Over the period 1973 to 1979 the area of cultivated oilseed rape in the UK expanded from 14,000 ha to 74,000 ha and production increased from 31,000 to 198,000 tonnes taking advantage of milder weather and the introduction of winter rape. This increase in area and production has continued since then (Department for Environment Food and Rural Affair, 2006).

1.1.2.4 THE CURRENT SITUATION

Oilseed rape, *B. napus*, is grown throughout the U.K. for the production of vegetable oils, animal feeds and biodiesel. It is also an important element of arable rotation farming which breaks the life cycles of pests and diseases, and is beneficial for the cereal crops grown in the following seasons.

According to a survey conducted by the Farm Business Survey for the Department for Environment, Food and Rural Affair (DEFRA) in 2006, the total area of the UK for *B. napus* production increased from 400,000 ha to 600,000 ha between 1996 and 2005. In 2005, 13.3% of the total area used to grow crops in the UK was for the production of oilseed rape, which has had a significant impact on the land use in agriculture. However, rape contributes just 5.35% (average from 2001 to 2006) of the total economic output derived from total crop production in the UK (DEFRA 2006). This means that *B. napus* has a lower economic value return in terms of the unit area of land used to grow this crop. An important aim of growers, breeders, scientists, and agronomists is to increase the productivity of oilseed rape so that it will achieve a similar economic return as other crops thus matching the economic value of the seed to the area grown. The value of the crop is, however, subject to factors that are both complex and difficult to control such as market price, climate change and yield etc., which makes it difficult to plan future strategies.

1.1.3 USE OF *B. NAPUS*

As mentioned in 1.1.2.2, rapeseed oil was used for lighting because of its slow burning and fairly odourless properties. Additionally, people in Europe avoided using rapeseed oil for food as its high erucic acid content affected the taste of the oil. Therefore, the reduction of erucic acid levels became a key objective and was successfully achieved through a combination of improved industrial processing techniques and the development of new cultivars through breeding (Weiss, 1983). The discovery of cultivars which had wide variability in its erucic acid content by

Steffansson in 1961 was a breakthrough for his development of erucic acid free oilseed rape hence the development of rapeseed for edible uses (Robbelen et al., 1989).

Today, the oilseed rape seeds are processed to extract oil, and the residue (called meal) is used for animal feed. Oil is the most economically valuable part of the crop and it is used both for human consumption and industrial uses such as lubrication or bio-diesel (Kimber and McGregor, 1995; Oliver and Roland, 2004). As the uses of the end product are so diverse, it is natural that the desired quality of oil should also be different according to requirements as shown in Fig.1-2 below.

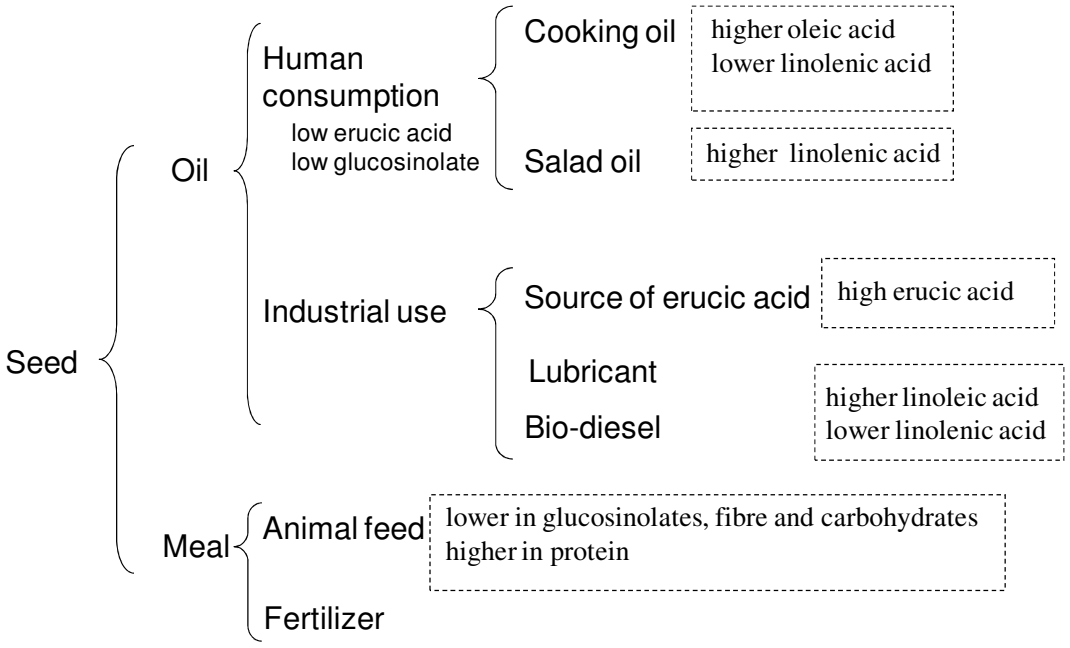


Fig.1-2 Diversity of end uses of rapeseed oil and desired oil profiles

1.2 *ARABIDOPSIS THALIANA*

1.2.1 USE OF *A. THALIANA* AS A MODEL PLANT

A. thaliana is a diploid plant (2n=10) and has a much smaller genome size (*ca*.145Mb) than *B. napus* (2n=38, *ca*.1200Mb) (Arumuganathan and Earle, 1991). Due to this small genome and other beneficial characteristics described later, its genome was the first plant to be sequenced in 2000 (The Arabidopsis Genome Initiative, 2000). The accession used in The Arabidopsis Genome Initiative 2000 project was Columbia. This completed genomic sequence provides a reference against which sequences from other *Arabidopsis* accessions or even species can be aligned. Hence it enabled the rise of next generation DNA sequencing technology which generates high throughput but short read DNA sequencing. This technology significantly both reduces the cost of genomic sequencing and increases the scale of output (Shendure and Ji, 2008).

Subsequently the 1001 Genomes project was launched and its aim was to identify whole genome sequence variation across 1001 *A. thaliana* accessions. To date genomic sequences have been released for 471 accessions, 57 accessions are fully sequenced and are in the release process with the sequencing of a further 178 accessions still planned as of 13th of February 2012. The resulting information enables identification of alleles underpinning phenotypic diversity across the entire genome of the large numbers of accessions within the species.

Arabidopsis and *Brassica* spp. are both in the *Brassica*ceae family and are closely related genera which split 14.5–20.4 million years ago (Yang et al., 1999). At the nucleotide level *ca*. 86% of sequences are conserved in homologous genes between *A. thaliana* and *B. napus* (Cavell et al., 1998).

As Fig.1-3 shows, due to the genome triplication observed in *B. oleracea* (O'Neill and Bancroft, 2000) and *B. rapa* (Lysak et al., 2005; Yang et al., 2006)and hybridisation events (U, 1935) the genetic components of *B. napus* may be the equivalent of six *A. thaliana* genomes, and this complicates the genetic research. Therefore studying the *A. thaliana* genome greatly facilitates the subsequent research on *B. napus*.

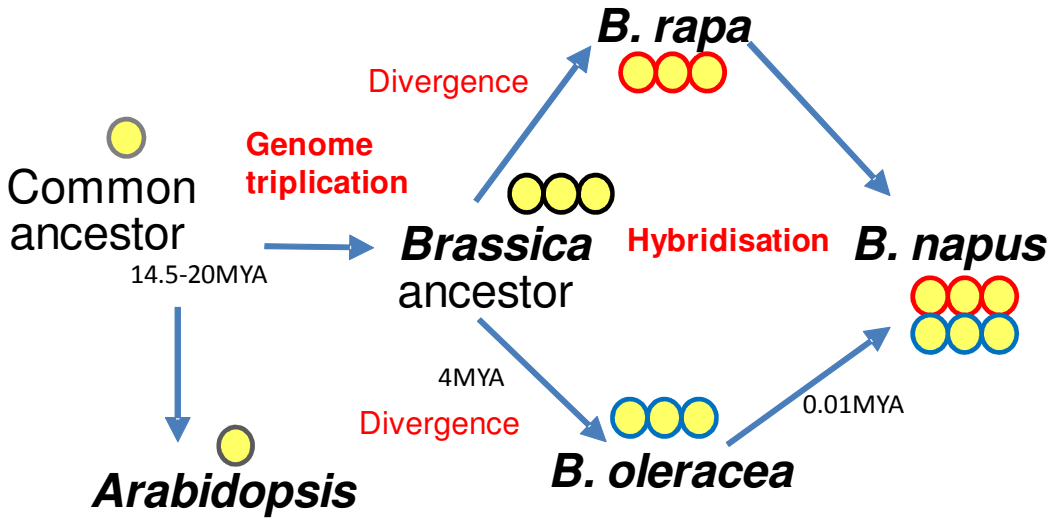


Fig.1-3 Evolutional history of *Brassica* and *Arabidopsis* (Yang et al., 1999; O'Neill and Bancroft, 2000; Schmidt et al., 2001; Inaba and Nishio, 2002; Rana et al., 2004)

MYA=million years ago, ○ represents a genetic component of the ancestral genome

The smaller physical size, the shorter life cycle together with *A. thaliana*'s simple growth requirements provide a great economical efficiency and expand the feasibility of experiments compared to *B. napus* which grows up to 180cm and takes more than 6 months from sowing until seed harvest. *A. thaliana* is an inbreeding plant type however it can be artificially out-bred and produces large numbers of seeds. *A. thaliana* can also be transformed by *Agrobacterium tumefaciens* which is advantageous for molecular genetic analysis work. Moreover, physiological and biochemical information are also widely available as well as a variety of mutants, recombinant inbred lines (RIL) and numerous genetic markers. As mentioned previously, sequence data of nearly 500 accessions are available in the 1001 Genomes and gene expression data in The Arabidopsis Information Resource (TAIR). These characters make *A. thaliana* suitable as a model plant for the investigation of *B. napus*.

1.2.2 NATURAL VARIATION IN *A. THALIANA*

1.2.2.1 ADAPTIVE VARIATION

A. thaliana is a wild species broadly distributed whose native range is from Europe to central Asia and is now naturalized worldwide through human activities. The

plants prefer open or disturbed habitats, such as sandy soil, river banks, roadsides, rocky slopes, waste places, cultivated ground, meadows, slightly alkaline flats, under shrubs and open areas with the sea level to 4250 m (Al-Shehbaz and OKane, 2002).

The wide range of habitats suggests that *A. thaliana* can locally adapt and part of the species' variation observed in many traits is presumed to reflect adaptive variation (Koornneef et al., 2011). Adaptative variation usually displays complex quantitative variations due to a wide range of integrated responses to cope with change of environmental conditions and adapt to the new environment to expand their habitat. Therefore these traits tend to be caused by polygenic effects, hence quantitative trait loci (QTL) analysis has been widely used to dissect this type of traits and recently linkage disequilibrium mapping has been used (Trontin et al., 2011).

Natural genetic variation has been observed and exploited for the understanding of the regulation of adaptative traits such as flowering time (Mitchell-Olds, 1996; Koornneef et al., 1998), seed and plant size (Krannitz et al., 1991; Alonso-Blanco et al., 1999; Ungerer et al., 2002), glucosinolates (Mithen et al., 1995), epicuticular wax composition (Rashotte et al., 1997) and tolerance to abiotic factors (Kobayashi and Koyama, 2002; Quesada et al., 2002; Hoekenga et al., 2003).

Seed lipid characteristics studied in this project are also regarded as adaptative traits because different combinations of FA affect the germination temperature, hence the timing of germination in nature. This could be due to the change of melting temperature determined by the degree of desaturation and length of carbon chain (Stryer, 1988; Linder, 2000; Sanyal and Linder, 2013). The characteristics of FA will be mentioned in section 1.3.

Therefore multiple populations developed from wild accessions from a range of habitats and geographical origins were appropriate material for the purpose of this study.

1.2.2.2 THE ADVANTAGES OF USING A NATURAL VARIATION APPROACH OVER A MUTATION APPROACH

Mutation approaches have been powerful tools for the functional analysis of genes, however the definition of gene functions was limited due to the small number of genetic backgrounds. The accessions historically used for mutagenesis are: Columbia (Col), Landsberg *erecta (*L*er*) and Wassilewskija (Ws). As they have been

mutagenised intensively by many researchers the genetic variation artificially induced within each accession is enormous. However, if Col, L*er* and Ws carried a natural null allele or weak effect alleles originally, it would be very difficult to detect such loci in a mutant population as it derived from a single accession. Hence by having additional accessions, the chance of detecting such loci could be increased.

In fact, the common lab strains Col and L*er* and Ws carry loss of function mutation in *FRIGIDA (FRI)* and/or *FLOWERING LOCUS C (FLC),* both of which are required for late flowering, therefore using other accessions is very important (Koornneef et al., 2004). According to the study by Borevitz, potential deletions in 111 genes were identified between Col and L*er,* these included transposons, disease resistance genes and genes responsible for secondary metabolism (Borevitz et al., 2003). A study of 20 diverse accessions, including Col, revealed that 9.4% of protein-coding genes in some accessions were affected by large effect intraspecific variation at the sequence level and the magnitude of variation was equivalent to a knock-out mutation (Clark et al., 2007). Therefore studying only one accession could limit access to genes.

The natural variation approach uses a wider range of accessions and multiple populations derived from them, therefore even if one of the accessions has a null allele or mutation at a gene for instance, other accessions used in the study may have wild type alleles, allowing the gene to still be studied. The schematic images of natural null alleles in the natural variation approach and mutation approach are shown in Fig.1-4.

Regarding the seed lipid characteristics, genes which have major impacts on the fatty acid composition of seed lipids were found through mutation approaches in *A. thaliana* (James and Dooner, 1990; Lemieux et al., 1990). Since then these genes have been intensely studied and used to manipulate the fatty acid profiles of the crop plant, *B. napus*. These works have lead to altered fatty acid composition for diverse end uses. However, following exploitation of these genes, further manipulation of individual fatty acids is becoming more challenging. Therefore in order to extend the current limits of breeding capability, breeders are seeking new approaches.

Although the lipid biochemical pathways are well studied and genes causing major impact have been identified (Beisson et al., 2003) the regulation of the process which affects in the final FA compositions is not known. The observation of a coordinated transcription of multiple genes involved in the pathway implies the

existence of loci functioning as regulatory mechanisms in addition to the main component of the biochemical pathway already identified (Ruuska et al., 2002).

Studies using multiple populations derived from a diverse array of accessions enable the exploitation of the natural variation acquired by the adaptative responses to the environment. Seed lipid characteristics are also regarded as adaptive traits (Linder, 2000), therefore this approach will facilitate the study of regulatory mechanisms that fine tune the expression of key genes, and contribute to the progress towards an understanding of the global architecture of seed lipids characteristics (Tonsor et al., 2005).

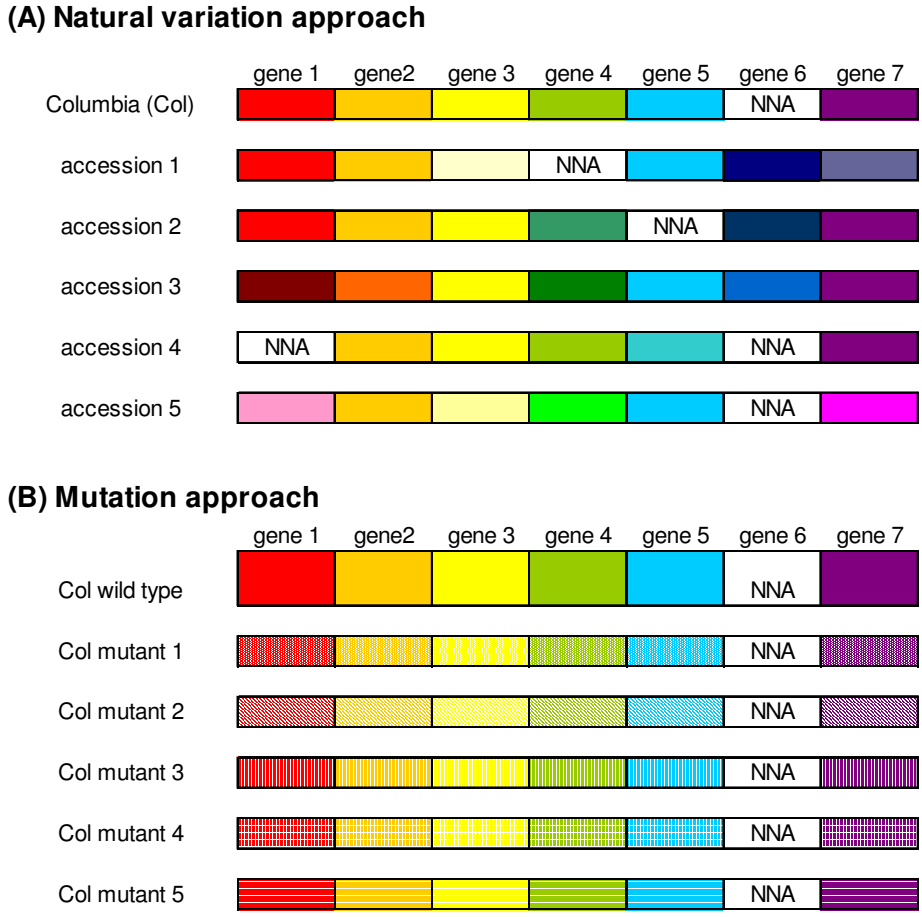


Fig.1-4 Schematic of natural null alleles (NNA) in *A. thaliana* genome in a natural variation approach (A) and a mutation approach using Col (B).

The different hues represent different genes (gene 1 to 7). Allelic differences are shown in different intensity and tone of the colour. The patterns represent induced mutation.

The natural variation approach uses multiple accessions which compensates for the presence of null alleles in one of the accessions. On the other hand, the mutation approach tends to use a single accession, if a natural null allele occurred in Col (for example gene 6 in the figure), the gene could not be investigated in mutant population

1.2.3 ORIGIN OF PLANT MATERIAL USED FOR THIS WORK

In a past experiment (O'Neill et al., 2003) 360 different *A. thaliana* accessions were screened for their total seed oil contents and fatty acid profiles. From these results 12 lines (Table 1-1) were chosen on the basis of differences in seed lipid composition and geographical origins to develop six RI populations. Each population consisted of 94 independent lines and was developed from F_2 lines by single seed descent up to the F_8 generation.

A bi-parental cross can only detect those loci which are polymorphic between a particular pair of accessions; therefore many pair-wise combinations are required for the study of allelic variation for a particular trait within a species.

The six RI populations in Table 1-1 were examined for their suitability for QTL analysis by using flowering time as an example of a complex trait (O'Neill et al., 2008). The QTL identified were a combination of previously known and others that were novel. For example, the QTL at the top of the linkage group (LG) 4 appeared in two populations, CA and SG. These QTL were coincident with the *FRI* locus a gene known to be a down regulator of *FLC* and suppressed by vernalization. In total 11 QTL were detected across six populations and eight were coincident with known genes while three were identified as novel loci. These results demonstrated that QTL analysis across multiple populations is not only reliable but is also capable of bearing innovative results. These RIL populations are used within the group and are also publicly available (http://www.jic.ac.uk/staff/ian-bancroft/arabidopsis.htm).

Table 1-1 Origin and lipid characteristics of parental lines

Recombinant Inbred population.	Accession name	Geographic origin	Oil characteristics			Cat.No
			Oil content	*Chain length ratio	Desaturation ratio	
CA	Cvi-0	Cape Verde Is.	low	low	high	N902
	Ag-0	France	medium	medium	low	N936
KB	Kondara	Tadjikistan	medium	low	high	N916
	Br-0	Czech	medium	high	medium	N994
NG	Nok-3	Netherlands	low	medium	medium	N1404
	Ga-0	Germany	high	high	low	N1180
SG	Sorbo	Tadjikistan	medium	low	medium	N931
	Gy-0	France	medium	high	high	N1216
TJ	Ts-5	Spain	medium	medium	high	N1558
	Mz-0	Germany	medium	medium	low	N9623
WC	Wt-5	Germany	low	medium	medium	N16120
	Ct-1	Italy	high	medium	low	N1092

* "Chain length ratio" shows the sum of all 20- and 22-carbon fatty acids contents divided by the sum of all 16- and 18-carbon fatty acid contents. "Desaturation ratio'' is the sum of polyunsaturated 18-carbon fatty acids divided by the sum of saturated and monounsaturated 18-carbon fatty acids. These ratios indicate the result of chain extension activity (e.g. by the product of the *FAE1* locus) and desaturation activity (e.g. by the products of the *FAD2* and *FAD3* loci). Modified from O'Neill 2008 and 2003.

1.3 FATTY ACIDS

Fatty acids (FAs) are a group of compounds that consist of a long hydrocarbon chain and a terminal carboxylate group. Symbols used to denote FAs use the number of carbons in its chain and also the number of double bonds. For example, 18:1 indicates that it has an 18 carbon chain and one double bond within that chain. To specify the position of the double bond, two alternative systems are in widespread use. Delta is used counting from the carboxyl terminals and the type of isomer is described as *cis*-delta 9. Alternatively, the position of the double bond can be denoted by starting by counting the methyl carbon as omega 1 (Stryer, 1988).

Together the chain length and degree of desaturation determine the properties of FAs and that of lipids derived from them. Unsaturated FAs have a lower melting point than saturated FAs although they have the same length of carbon chain. For example, the melting point of stearic acid (18:0) is 69.6°C, while oleic acid (18:1) melts at 13.4°C. A FA which has a single double bond is called a mono-unsaturated fatty acid (MUFA) and a FA with more than two double bonds is called a polyunsaturated fatty acid (PUFA). Chain length also affects the melting temperature, the shorter the length of chain, the lower the melting point (Stryer, 1988).

FAs are predominantly membrane components, while the remainder is allocated to the seeds for carbon storage, epidermal cells as cuticular lipids, serve as hormone precursors or are involved in acylation of certain membrane proteins (Ohlrogge and Browse, 1995). FAs for seed storage form triacylglycerols (TAG) which are uncharged esters of glycerol therefore they can be highly concentrated stores of metabolic energy (Stryer, 1988).

1.3.1 FATTY ACID BIOSYNTHESIS

The process of FA synthesis is complex and occurs in different plant cell organelles through many catalytic reactions and the lipids produced are distributed within the plant for a variety of uses as described above.

The seed lipid biosynthesis process can be categorised into four steps and is explained in the designated sections.

1.16:0-acyl carrier protein (ACP) and 18:1- ACP synthesis in the plastid (Section 1.3.3.1)

2. Glycerolipid synthesis (Section 1.3.1.2)

3. Further desaturation in the plastid and the endoplasmic reticulum (ER) (Section 1.3.1.2)

4. Formation of triacylglycerol (TAG) for seed storage in the endoplasmic reticulum (Section 1.3.1.3)

1.3.1.1 16:0-ACP AND 18:1- ACP SYNTHESIS IN THE PLASTID

The fatty acid biosynthesis pathway is a primary metabolic pathway found in every plant cell and it is vital for growth therefore loss of function in fatty acid biosynthesis is lethal to the cells. Acetyl-coenzyme A (CoA) is a fatty acid biosynthesis substrate in the plastid which is produced by a number of alternative pathways to supply the immediate need (Ohlrogge and Browse, 1995). In the first reaction of the fatty acid biosynthesis pathway, acetyl-CoA and CO_2 are catalysed by acetyl CoA carboxylase which contains a biotin prosthetic group (Stryer, 1988) and are converted into malonyl-CoA. 16:0 or 18:1 fatty acid are produced from the conversion of acetyl-CoA to malonyl-CoA through at least 30 enzymatic reactions (Ohlrogge and Browse, 1995).

The initial product of each condensation reaction is a 3-ketoacyl-acyl carrier protein (ACP). The first condensation step of acetyl-CoA and malonyl-ACP to form a product with carbon chain length of 4 (C4) is catalyzed by ketoacyl ACP synthase (KAS) III (Jaworski, 1989). A second condensing enzyme, KAS I, is believed to be responsible for extending carbon chain lengths from 6 to 16. Elongation of C16 palmitoyl-ACP to C18 stearoyl-ACP requires the condensing enzyme KAS II. Three additional reactions: reduction of 3-ketoacyl-ACP, a dehydration stage and a reduction of a double bond, occur after each condensation step to form a saturated fatty acid and the cycle continues as shown in Fig.1-5 (Ohlrogge and Browse, 1995).

When 3-ketoacyl-ACP is reduced at the carbonyl group by the enzyme, 3-ketoacyl-ACP reductase, NADPH is used as the electron donor. The next reaction is dehydration by hydroxyacyl-ACP dehydratase. Each round of fatty acid synthesis is completed by enoyl-ACP reductase, which uses NADH or NADPH to reduce the trans-2 double bond to form a saturated fatty acid. The combination of these four

reactions results in the lengthening of the precursor fatty acid by two carbons while it is still attached to ACP as a thioester (Ohlrogge and Browse, 1995).

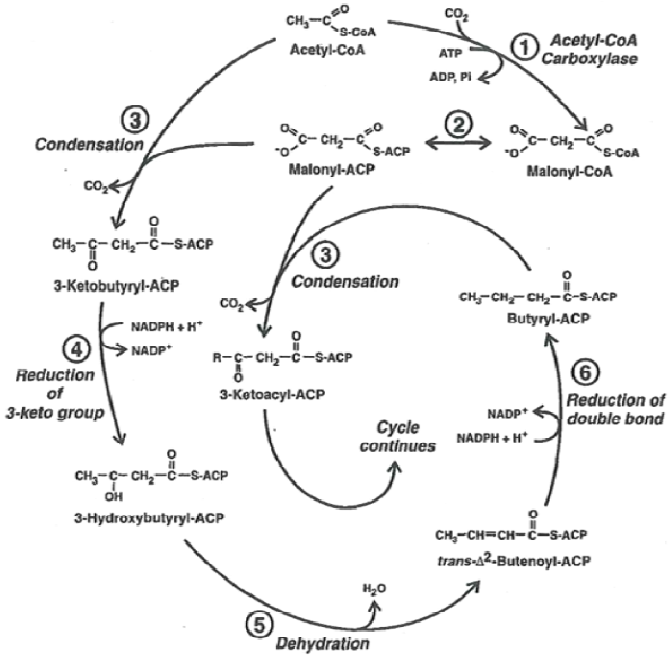


Fig.1-5 Reaction of saturated fatty acid biosynthesis (Ohlrogge and Browse, 1995)

The enzyme which catalyzes the saturated fatty acids to form a first double bond i.e. 18:0 to 18:1 is the plastidial stearoyl-ACP desaturase. This soluble fatty acid desaturase is unique in the plant kingdom because all the other known desaturases are proteins integrated in the membrane (Ohlrogge and Browse, 1995). Termination of plastidial fatty acid chain elongation is catalyzed by acyl-ACP thioesterases, which hydrolyze acyl chains from ACP. After termination, free fatty acids are activated to CoA esters, exported from the plastid, and assembled into glycerolipids in the ER (Thelen and Ohlrogge, 2002).

1.3.1.2 GLYCEROLIPID SYNTHESIS AND DESATURATION ACTIVITY

Free fatty acids produced in the plastid move through the plastid envelope and are converted into acyl-CoA thioesters by an acyl-CoA synthetase in the outer envelope. They are used for the synthesis of phosphatidylcholine (PC) in the ER which is the major substrate for 18:1 and 18:2 desaturations (Browse and Somerville, 1991).

The glycerol-3P acyl transferase transfers both C16 and C18 FAs to the sn-1. “sn” stands for stereochemical numbering and indicates the position of carbon atom in glycerol from 1 to 3 (Gurr, 2002). The selectivity of acylation by this enzyme was observed in safflower microsomes in the order 16:0>18:1=18:2>18:0 (Ichihara, 1984). The second acylation catalyzed by acyl-CoA:lysophosphatidic acid acyltransferase is selective for 18:2 over 18:1 and almost excludes 16:0 and 18:0 (Griffiths et al., 1985). These results reflect membrane phospholipid composition (Browse and Somerville, 1991).

The characterization of a series of *A. thaliana* mutants for individual desaturation steps, from *fad*2 to *fad 8* (James and Dooner, 1990; Lemieux et al., 1990) made an enormous contribution to the understanding of the mechanisms of plant desaturation. Among them mutations at two loci; *fad2* and *fad3,* were shown to affect the lipid desaturation in the ER, while others affect chloroplast lipid desaturation as seen in Fig.1-6.

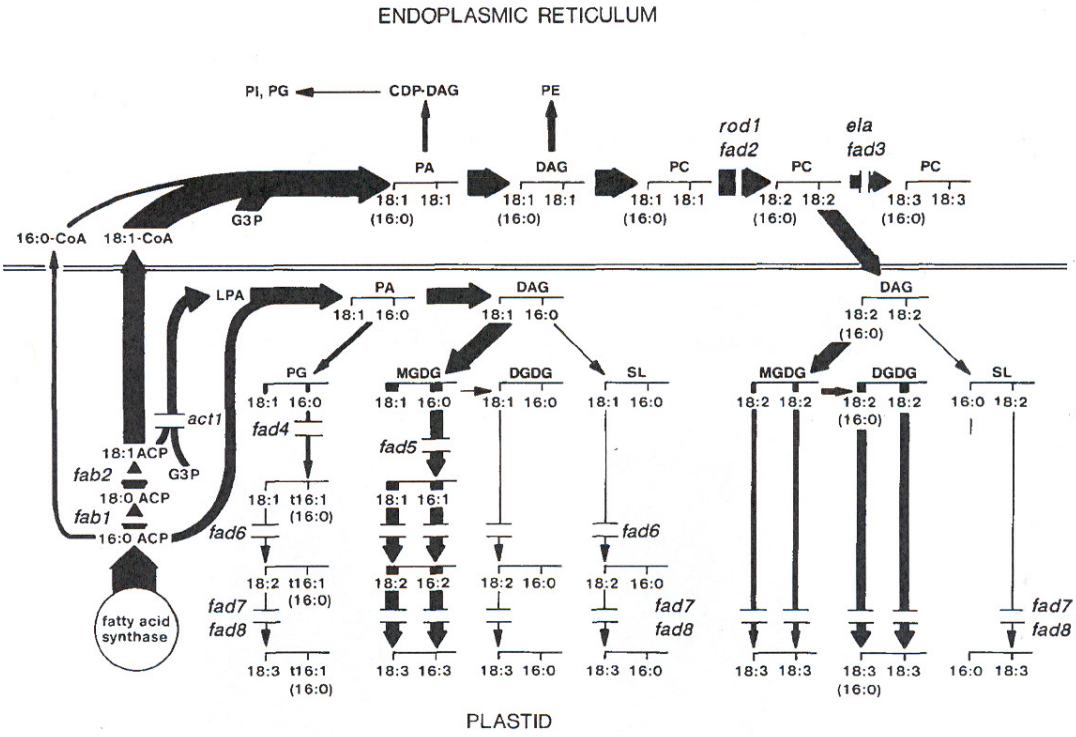


Fig.1-6 Diagram of FA synthesis and glycerolipid assembly in *A. thaliana* leaves showing the downstream effects of mutations (Browse and Somerville, 1991)

ACP: acyl carrier protein, PI: phosphatidylinositol, PG: phosphatidylglycerol, MGDG: monogalactosyldiacylglycerol, DGDG: digalactosyldiacylglycerol, SL: sulfoquinovosyldiacylglycerol, G3P: Glycerol-3-phosphate, CDP: cytidine diphosphate, DAG: diacylglycerol, PA: phosphatidic acid, PE:phosphatidylethanolamine, PC: phosphatidylcholine

1.3.1.3 FORMATION OF TAGS FOR SEED STORAGE IN THE ENDOPLASMIC RETICULUM

FAs are successively converted from CoA to sn-1 and sn-2 of glycerol-3-phosphate, resulting in the formation of the central metabolite, PA which is dephosphorylated to release diacylglycerol (DAG). In the final step of TAG synthesis a third FA is transferred to the vacant position sn-3 of DAG. This step is catalysed by diacylglycerol acyltransferase (DGAT), the only enzymatic reaction unique to TAG synthesis.(Ohlrogge and Browse, 1995). TAGs are highly concentrated stores of metabolic energy because they contain reduced carbon, and are anhydrous due to their non-polar character. The energy that TAGs can store is more than six times that of the equivalent weight of glycogen (Stryer, 1988).

Finally TAGs are stored as oil bodies in seed embryo cotyledons. Each oil body consists of a triacylglycerol matrix surrounded by a single layer of phospholipids embedded with structural oleosin proteins (Huang, 1992).

1.4 KEY GENES INVOLVED IN FATTY ACID DESATURATION

Difficulties in purification to homogeneity level limited the characterization and cloning of the membrane-bound FA desaturases and instead a genetic approach was taken (Ohlrogge and Browse, 1995). A series of EMS-induced mutants of *Arabidopsis thaliana* with altered seed lipid composition was isolated, such as the mutants which were deficient in 18:1 desaturation (*fad2),* 18:2 desaturation *(fad3)*, 18:1 elongation *(fae1)*, and others which had increased16:0 by two to three fold and decreased 18:1 by half *(fab1)* and increased 18:0 *(fab2)* (James and Dooner, 1990; Lemieux et al., 1990).

Using these mutants, genes encoding membrane-bound desaturases were cloned. The homology of these genes with the gene products of other organisms which have been characterized at the biochemical level revealed that the *FAD* loci represent the structural genes for desaturases rather than proteins that control the desaturation reaction (Ohlrogge and Browse, 1995).

A schematic image of the PUFAs desaturation pathway is shown in Fig.1-7. Major genes involved in the pathway are indicated as blue arrows and described in the following sections.

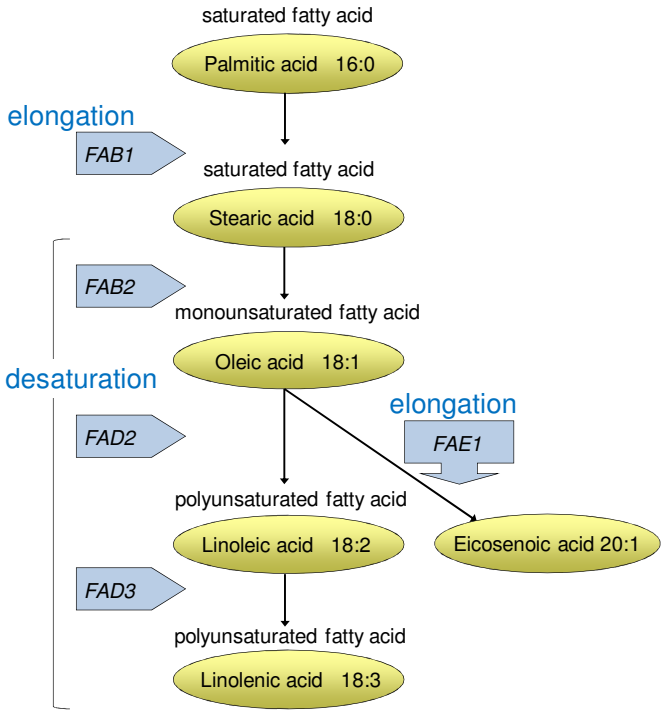


Fig.1-7 Major genes controlling PUFA compositions
(Taken from Browse and Somerville 1991)

1.4.1 *FATTY ACID BIOSYNTHESIS 1 (FAB1)*

The *fab1* mutant has a two to three fold increased 16:0 content, 18:1 is reduced by half, has a small decrease in 18:2, with increased 16:1 and 20:0 and a sensitivity to low temperatures which can result in damage or death (James and Dooner, 1990). The mutant phenotype was the result of a partial loss of function in *KAS2* activity caused by a single nucleotide change which led to a Leu337Phe substitution. A complementation test was carried out by transforming *fab*1 plants with a cDNA from *B. napus* encoding a KASII enzyme. This resulted in both the altered FA profile and insensitivity to low temperature being abolished (Carlsson et al., 2002).

1.4.2 *FATTY ACID BIOSYNTHESIS 2 (FAB2)*

FAB2 encodes a STEAROYL-ACYL CARRIER PROTEIN DESATURASE (S-ACP-DES) and catalyses conversion of 18:0 to 18:1 in plants (Kachroo et al., 2007). The *fab2* T-DNA insertion mutant of *A. thaliana* was isolated during a screen for seed fatty acid composition. It showed increased levels of stearate 18:0 (8%) in seed lipids compared with the wild type (2 to 4%).This increased 18:0 level was also found in the leaves. The altered fatty acid composition affected numerous cell and organ specific changes and the mutant exhibited reduced growth (Lightner et al., 1994).

1.4.3 *FATTY ACID DESATURATION 2 (FAD2)*

The *fad2* mutant showed an increased level of 18:1 with a decreased level of 18:2 in leaves and seeds due to the absence of ER desaturase, however no distinct morphological differences were observed relative to the wild type (James and Dooner, 1990; Miquel and Browse, 1992).

The *FAD2* allele was identified in an *A. thaliana* population in which mutations were generated by T-DNA insertions. Genomic DNA flanking the T-DNA was cloned by plasmid rescue and used to isolate cDNA and genomic clones of *FAD2*. By transforming the *fad2* mutant with a cDNA containing the entire *FAD2* coding sequence complemented the mutant fatty acid phenotype. The deduced amino acid sequence from the cDNA revealed that there were three histidine-rich sequences that showed homology to other membrane-bound desaturases, and this confirmed that *FAD2* is the structural gene for desaturase. However the homologies between

the open reading frame of *FAD2* and other desaturases are low at both the nucleotide and protein levels (Okuley et al., 1994).

1.4.4 *FATTY ACID DESATURATION 3 (FAD3)*

FAD3 encodes the ER LINOLEATE DESATURRASE and is responsible for the synthesis of 18:3 fatty acids. Ethyl methanesulfonate (EMS) induced mutants with a reduced accumulation of linolenic acid (18:3) and a corresponding increase in the amount of linoleic acid (18:2) in the extra chloroplast membrane and storage lipids, were isolated and termed *fad3* (Lemieux et al., 1990). The gene was cloned by map based cloning using an F_2 mapping population derived from a hybrid between the mutant and wild type plants (Arondel et al., 1992) and also by T-DNA tagging of screened T-DNA transformants (Yadav et al., 1993). The cDNA obtained was tested for its homology with soybean and rapeseed and a high degree of sequence similarity was observed (Yadav et al., 1993).

1.4.5 *FATTY ACID ELONGATION 1(FAE1)*

There are four different reactions in the elongation of carbon chain length of FAs in plants. (1) Condensation of 18:l CoA with malonyl CoA to form a P-ketoacyl derivative, (2) Reduction and (3) Dehydration of the 0-ketoacyl derivative, and (4) Reduction of the double bond (Fehling and Mukherjee, 1991).The first reaction affects the composition of C18 PUFA in the seed due to the competition for 18:1 CoA between the desaturation pathway and the elongation pathway.

The *A. thaliana FAE1* gene encodes a seed-specific condensing enzyme, 3-KETOACYL-CoA SYNTHASE (KCS) and mutations in the *FAE1* gene led to reduced levels of seed very-long-chain fatty acids (VLCFAs) (James and Dooner, 1990; Lemieux et al., 1990).

It is the first of four enzyme activities that constitute the microsomal fatty acid elongase involved in the biosynthesis of VLCFAs (James Jr et al., 1995). *FAE1* is the rate-limiting enzyme for VLCFA biosynthesis in *Arabidopsis* seed and also determines the acyl chain length of the VLCFAs produced (Millar and Kunst, 1997).

FAE1 was cloned by directed transposon tagging with the maize element *Activator (Ac).* A line carrying *Ac* linked to the *FAE1* locus on chromosome 4 was isolated,

and a DNA segment flanking *Ac* was cloned by inverse PCR to obtain *FAE1* genomic clones and a cDNA clone from a library prepared from immature siliques. Protein sequencing of FAE1 showed homology with other condensing enzyme, which suggested that *FAE1* was a structural gene for the enzyme (James Jr et al., 1995).

1.5 QTL MAPPING

Many agronomically important traits such as yield, quality and some forms of resistance against abiotic stress are controlled by multiple genes. They are characterized by continuous variation and are termed quantitative traits (Collard et al., 2005; Keurentjes et al., 2007). Natural variation typically appears as continuous quantitative variation while the mutation phenotypes often show discontinuous qualitative variation in diploids (Borevitz et al., 2003). The regions within genomes that contain genes associated with a particular quantitative traits are known as quantitative trait loci (QTL) (Collard et al., 2005).

The detection of loci associated with these quantitative traits, is not as straightforward as that of major genes due to their multi-genic inheritance and strong interaction with the environment (Keurentjes et al., 2007), therefore the technique of QTL mapping was developed with the help of genetic marker analysis and statistical analysis.

QTL mapping is based on the genotyping of progeny derived from a cross of divergent phenotype for the trait of interest. A linkage map is developed to indicate the order of markers and relative genetic distances between markers along linkage groups from the genotyping data. Phenotypic data for the quantitative trait is then compared with the molecular marker genotypes of the progeny. The linkage map is searched for genomic regions showing statistically significant associations with trait variation, these particular regions are the QTL (Collard et al., 2005; Slate, 2005).

1.5.1 LINKAGE MAPS

The purpose of producing linkage maps is to order the genetic markers and identify chromosomal locations containing genes and QTL associated with traits of interest. QTL mapping relies on the segregation of genes and markers via chromosome recombination during meiosis. The frequency of recombinant genotypes can be utilized for the calculation of recombination fractions, which indicates the genetic distance between markers. The lower the frequency of recombination between two markers, the closer they are situated on a chromosome. The relative order and distances between markers can then be determined (Collard et al., 2005).

1.5.2 QTL ANALYSIS

Basic QTL analysis utilizes an association between phenotype and the genotype of markers. Markers are used to partition the mapping population into different genotypic groups based on the presence or absence of a particular marker allele and to determine whether significant differences exist between groups with respect to the trait being measured (Young, 1996). A significant difference between phenotypic means of the groups indicates that the marker locus being used to partition the mapping population is linked to a QTL controlling the trait (Collard et al., 2005).

This analysis is carried out using computer software such as MapQTL5.0 which is designed to calculate the position of QTL on genetic linkage maps in experimental populations of diploid species (Van Ooijen, 2004).

The method used in this project is interval mapping which makes use of linkage maps to analyse the intervals between adjacent pairs of linked markers along linkage groups. As a linkage map gives marker order and additional genetic information, it is considered statistically more powerful compared to single marker analysis (Lander and Botstein, 1989).

The output from the interval mapping is a profile of the likely sites for a QTL between adjacent linked markers using a logarithmic of odds (LOD) score. The significance of a LOD score is calculated by the permutation test (Van Ooijen, 2004).

There are two types of error which could occur in QTL analysis, a false positive (type Ⅰ) and a false negative (type Ⅱ). A Type Ⅰ error leads to a conclusion that there is a segregating QTL which does not exist in reality, while a type Ⅱ error does not detect a QTL although it is present. The probability of type Ⅰ error is controlled by choosing the appropriate significance threshold and that of type Ⅱ error is determined by the experimental set-up and the size of the genetic effects of the QTL (Van Ooijen, 1999).

To predict a false positive, LOD significance thresholds were calculated through permutation test. Genome-wide significance in QTL analysis is equivalent to the experiment-wise significance level of 5%, this is the probability of obtaining a LOD above the threshold anywhere in the whole genome just by chance is 5% (Van Ooijen, 1999).

According to Stratton, there is always a possibility of type I error, therefore it is probably most useful to view QTL mapping experiments as exploratory studies that yield hypotheses which should be confirmed by the construction of near-isogenic lines (NIL) that differ only in the region of interest (Stratton, 1998) or by using heterozygous inbred families (HIF) (Tuinstra et al., 1997).

1.5.3 MAPPING POPULATIONS

1.5.3.1 RECOMBINANT INBRED LINES

Certain genetic variation present among *A. thaliana* accessions is not detectable by simple comparison of accessions, therefore accessions are required to be crossed and then the offspring could show segregation outside of the range of parental value which is known as transgressive segregation (Koornneef et al., 2004). Recombinant inbred (RI) populations are produced by crossing parents chosen to provide maximum polymorphism as both phenotypic variation and genotypic variation. The F_1 progenies are selfed repeatedly for several generations until their genomes become homozygous. RI populations have been widely recognized as a useful mapping resource due to their homozygosity, which enables repeat experiments and multiple analysis of the same population, especially important because quantitative traits are prone to interaction with the environment (Koornneef et al., 2004; Keurentjes et al., 2007). Doubled haploid (DH) populations are also homozygous, however the species which can be generated are limited and DH lines have less recombination than RIL (Keurentjes et al., 2007) which is a disadvantage when detecting QTL.

As a predominantly self fertile species (Koornneef et al., 2004), and together with its short life cycle *A. thaliana* is a useful plant in which to develop RIL populations . The populations used in this project were produced by single seed descent to the F_8 generation and the average frequency of heterozygosity for all alleles in the six populations was approximately 1.0% (O'Neill et al., 2008).

1.5.3.2 NEAR ISOGENIC LINES

Another type of population which also consists of homozygous individuals is near isogenic lines (NILs) (Monforte and Tanksley, 2000). NILs are generated through repeated backcrossing and genotyping in order to introduce a small fragment of

donor parent into a genetically homogeneous recipient parent background (Keurentjes et al., 2007) resulting in the elimination of other QTL effects. NILs have been proven to be extremely useful to identify target genes in many crops providing a starting point for fine mapping and cloning (Young et al., 1988; Paran and Zamir, 2003; Kim et al., 2008).

CHAPTER 2. HYPOTHESIS AIMS AND OBJECTIVES

2.1 HYPOTHESIS

Variation in seed oil profile cannot be fully explained by the current identified FA desaturation pathway genes. Therefore we hypothesise that there could be additional mechanisms controlling quantitative PUFA content.

2.2 AIMS

The aims of this project are to search for uncharacterised loci involved in the quantitative control of a PUFA composition of seed oil in *A. thaliana* and to take a comparative genomics approach to identify the corresponding loci in *B. napus*, oilseed rape.

2.3 OBJECTIVES

1), Using QTL analysis, search for genes with a significant effect in determining PUFA composition and oil content in *A. thaliana*. Six populations each consisting of 94 RILs; CA, SG, WC, KB, NG and TJ populations to provide maximum natural variation in PUFA composition and oil content will be used for this analysis. The primary QTL analysis for these traits was carried out by O'Neill in 2007. QTL analysis using genetic maps with increased marker density will be performed followed by fine mapping of target QTL.

2), NILs will be produced to test the robustness of selected QTL and a mapping population will be produced from a selected NIL for positional cloning to narrow down the region of interest and identify candidate genes within that.

3), The results of QTL analysis obtained in the model plant (*A. thaliana*) will be compared with data from the crop plant, *B. napus* in the literature, databases and studies within the Bancroft group to contribute a further understanding of the genetic control of FA desaturation.

CHAPTER 3. GENERAL MATERIALS AND METHODS

The materials and methods used twice or more throughout the studies are described in this chapter. Materials and methods used only once in specific experiments are described individually in Chapters 4, 5 and 6.

3.1 PLANT MATERIAL

This thesis reports research into *Arabidopsis thaliana* RI populations developed within the Bancroft group (O'Neill et al., 2008). Each population consists of 94 independent lines and was developed from F_2 plants by single seed descent up to the F_8 generation.

Parental lines were originally obtained from Nottingham Arabidopsis Stock Centre (NASC; http://www.arabidopsis.info/) and maintained as a lab stock by O'Neill.

3.2 PLANT HUSBANDRY

Equipment

- Forceps
- Breathable transparent plastic bags 150 X 600mm (P360 Lap seal bags Focus Packaging & Design Ltd)
- Bamboo stakes
- Labels
- Clingfilm or transparent plastic lid
- 7cm square pots or seed trays
- Arabidopsis soil mix: 6 parts Scott's Levington F2 with one part 4mm Grit and Intercept 5 GR (280g/ m^3). Intercept is a pesticide used to prevent root damage caused by larvae of Scarid fly

Materials

- Nemasys: Biological control targeting Scarid fly and Western flower thrips,
- Systhane: Systemic fungicide Bayer

Method

1. To reduce the risk of thrips infection, seeds should be kept in -20°C for more than 16 hours prior to sowing.

2. Irrigate soil in pots with water containing Nemasys as a biological control of other pests and diseases.

3. Spread seeds thinly onto *Arabidopsis* soil mix in 7cm square pots or seed tray using forceps if necessary.

4. Cover the pots or trays with transparent plastic lids or cling film to keep humid.

5. Material to be used for phenotyping of their seeds lipids or for crossing: vernalize the sown seeds in a 4°C chamber with an 8 hr photoperiod for 6 weeks. This is to encourage simultaneous flowering.

6. Plant material to be used for DNA extraction: place sown seeds in a 4°C chamber for 5 days for seed stratification to achieve uniform germination.

7. Transfer the pots or trays from the vernalization chamber to a temperature controlled glasshouse with 22°C daytime and 18°C night time and supplementary sodium light to give a 16hr photoperiod.

8. Keep cling film or covers on the pots to protect seedlings from drying out. Over a few days make increasing numbers of slits for ventilation. The cover is removed within a week.

9. When the seedlings produce their first true leaves, transplant into a 7cm square pot with *Arabidopsis* soil mix, the same compost used for sowing. Cover the pots with cling film. Over a few days make slits in the cling film and remove completely within a week.

10. Locate pots on pot trays which hold 30 pots (5 rows x6 columns) without gaps at the juvenile stage.

11. Spray with the systemic fungicide “Systhane” to prevent powdery mildew.

12. When plants start to bolt cover with transparent, breathable plastic bags and use canes to support the plants.

13. Spread the plants out within the trays to make room for them to expand their branches.

14. When half of the siliques are brown, put the plants for drying off for a minimum of 3 weeks.

3.3 CROSSING

Equipment

- Microscope

- Crossover forceps

- Fine forceps

- Blue roll

- Thread

- Cling film

Materials

-100% Ethanol

Timing

In the morning to obtain good viable pollen from freshly opened flowers.

Method

1. Before starting a new crossing clean the bench and sterilize forceps and scissors with ethanol each time.

2. Choose a raceme on the maternal plant and remove all siliques and opened flowers. To distinguish un-opened buds from once opened but closed, observe the stigma under the microscope. If the stigma projects above the closed petals, it has opened already, therefore remove it.

3. Choose three to four buds for pollination and remove the rest of the buds carefully with the fine forceps. The selected buds should not have opened or be immature.

4. Using the fine forceps carefully, remove sepals, petals and anthers from the bud to expose the stigma. Wetting fingers with clean water when handling plants helps prevent the exposed plant parts from drying out and decreases the risk of bruising the delicate pedicel.

5. Select a fresh flower from the paternal plant to act as pollen donor. The opened flowers next to the closed buds are the best for this.

6. To expose the anthers pinch the pollen donor flower mid-way along the sepals with crossover forceps. Remove sepals, petals and the stigma with fine forceps to further expose the anthers.

7. Powder the stigma of maternal plant with the paternal pollen until the entire receptacle is powdered with yellow. Use as many pollen donor flowers as necessary to obtain enough pollen.

8. Loosely, but securely, tie a thread on the pollinated raceme to label it. Use threads with different colours according to the date of pollination if necessary.

9. Roll out cling film and discard the first piece to avoid contamination and only use clean parts carefully. Cut out a $6cm^2$ of cling film to gently cover the pollinated buds and protect them from drying out and contamination with other pollen.

10. Place pollinated plants in half-shade to rest. Cling film can be removed the next day.

11. After crossing, pollinated siliques will be ready for harvesting in four weeks.

3.4 HARVESTING SEEDS

There are two ways to collect seeds from plants: individual silique collection method and bulk collection method.

The silique collection method was applied to collect seeds from the pollinated siliques or to collect seeds for use in future crossing as it guarantees the seeds are not contaminated with other seeds.

The bulk collection method is used to collect seeds from the whole plants for phenotyping and also as back up stocks. Both methods can be applied at the same time.

3.4.1 SILIQUE COLLECTION

Equipment

- Forceps

- Scissors

- Labelled eppendorf tubes

-Syringe needle

Methods

1. Prepare eppendorf tube. Pierce the lid of the tube with the needle of a syringe in two positions for ventilation purposes to prevent siliques from being mouldy and damaging seeds.

2. Label the tube and prepare a box to hold the tubes. Seeds can come out from the holes in the lids, therefore keep upright.

3. Insert the silique for harvesting into an eppendorf tube without touching the silique itself as it can shatter.

4. Cut the pedicel inside the tube to release the silique from the plant.

5. Close the lid and store in the box until the seeds are completely dry. Transfer the seeds into an eppendorf tube without holes if necessary.

3.4.2 BULK COLLECTION

Equipment

- Scissors

- Glassine ungummed paper bags (330 x 180mm)

- Kristal ungummed paper bags (135 x 108mm)

- Paper clips

Method

1. Cut the plant at the base of the stem.

2. Insert the whole plant including plastic bag and cane into a glassine paper bag upside down. If the seeds cannot be lost a paper bag is not necessary.

3. Store the harvested material in a cardboard box to dry completely.

4. Press the outer bag to shed seeds into the bag. Remove plastic bag and cane at this stage.

5. Collect seeds into one corner of the bag and cut off this corner.

6. Empty the collected seeds into a Kristal ungummed paper bag.

7. Fold over the top of the seed bags and close it with paper clip together with plant label.

3.4.3 CLEANING OF SEEDS FOR PHENOTYPING

Seeds used for the analysis of fatty acid composition and oil content need to be clean and free of plant debris to obtain accurate results. Therefore, seeds for phenotyping collected by the bulk collection method described in the previous Section 3.4.2 should be cleaned to minimize any contamination.

Equipment

- Scissors

- Fine sieve (55mm diameter)

- Sheet of chromatography paper

- Kristal ungummed paper bag (73X41mm)

Method

1. Sieve the seeds twice to remove relatively large silique debris and other parts of the plants.

2. Blow carefully on the paper to remove fine debris and dust.

3. Spread the seeds on a rough surfaced paper such as chromatography paper and pour into a seed bag. The lighter and finer debris stays on the paper.

5. Fold over the opening of the bag twice and close with a paper clip together with the plant label to store.

3.5 SAMPLING LEAVES FOR DNA EXTRACTION

Leaves were used to extract DNA because genotyping should be carried out as soon as possible to select plants. Plants to be sampled should be at the juvenile stage, but sufficiently developed to withstand the loss of leaves. It is usually two to three weeks after transplanting.

Leaf samples were submitted to Richard Goram in the Genome Laboratory Services and later DNA prep service both in John Innes Centre for DNA preparation. Qiagen DNeasy 96 Plant Kits (Cat. No. 69181) were used following manufacture's instructions.

Equipment

- Scissors

- Forceps

- Gloves

- Ice in an ice bucket or dry ice in a polystyrene box

- Collection micro tubes (Qiagen Cat. No. 19560)

Materials

-100% Ethanol

Method

1. Prior to sampling prepare in Excel a collection plate spreadsheet.

2. Sterilise scissors and forceps with ethanol and wear gloves to avoid contamination.

3. Place tubes or tubes in box on ice tray to keep it cool.

4. Cut the base of leaves with scissors.

> 50mg of leaf sample is the optimum amount of material for this DNA extraction method, which is approximately two newly expanded leaves.

5. Roll the leaves between the fingers to fit the micro tube and insert them into collection micro tube with a forceps.

6. Submit the samples immediately for DNA extraction or alternatively keep them at -80℃ freezer until they are needed.

3.6 FLUORESCENT LABELLED HIGH THROUGHPUT GENOTYPING

Genotyping was done throughout the work detailed in Chapters 5, 6 and 7 to test for heterozygosity, for population screening and to sort out plants according to their genotypes. The process of genotyping consists of PCR amplification using markers at the loci of interest and the size of PCR products was determined using an ABI 3730XL sequencer. The sequencer was run by the Genotyping Service at JIC and the results were returned in fsa file format. Scoring was carried out using Gene Mapper (v.3.7).

3.6.1 PCR METHOD

This PCR method is based on the protocol used by the Genotyping Service at JIC.

Equipment

- PCR machine

- THERMO-FAST® 96 NON-SKIRTED PCR plates (Thermo Scientific, Cat. No. AB-0600)

- Adhesive PCR film (Thermo Scientific, Cat. No. AB-0580)

- Strips of 8 Thermo-Tubes & Domed Caps (Thermo Scientific, Cat. No. AB-0266)

Materials

- Qiagen HotStarTaq Master Mix Kit (Cat. No. 203445)

- Forward primer with M13 tail (5' TGTAAAACGACGGCCAGT-XXX 3') at 100µM concentration

- Reverse primers at 100µM concentration

- AB PRISM primer (6FAM, NED, PET and VIC) with M13 sequence (TGTAAAACGACGGCCAGT) at 100µM concentration

- DNA from sample (20ng/µl)

- Sterile Elga water

Method

1. Prepare the Primer mix by mixing together and store at -20℃.

    -18.75µl AB PRISM primer (100µM). 500µl of 1xTE (pH7.5) was added on receipt to suspend dried contents

    -18.75µl Reverse primer (100µM)

    -1.25µl Forward primer (100µM)

    -211.25µl Sterilized Elga water

2. Into each well pipette:

    - 3.125µl Qiagen HotStarTaq Master Mix

    - 0.625µl of the Primer mix prepared in step1

    - 2.5µl sample DNA

3. Use the adhesive PCR seal to cover each PCR plate or set of microtubes to avoid evaporation during the PCR operation.

4. Vortex the sealed plate or tubes and centrifuge prior to PCR.

5. Place the plate or tubes in the PCR machine and run the program “ABI genotyping HSR60, HSR55, HSR52, or HSR50” depending on the annealing temperature required ie 60℃, 55℃, 52℃ and 50℃ respectively.

ABI genotyping HSR program:

1). 95℃ for 15 minutes

2). 95℃ for 1 minute

3). 0.5℃/s increasing temperature up to the annealing temperature

4). Annealing temperature for 1 minute

5). 0.5℃/s increasing temperature up to 72℃

6). 72℃ for one minute

7). Repeat step 2) to step 6) 34 times

8). 72℃ for 10 minutes

9). 12℃ to hold

6. After PCR, centrifuge the plate or tubes and store at -20℃.

3.6.2 PREPARATION OF SAMPLES FOR GENOTYPING DOWN THE ABI 3730XL SEQUENCER

Equipment

- THERMO-FAST® 96 NON-SKIRTED PCR plates (Thermo Scientific, Cat. No. AB-0600)

- Adhesive PCR film (Thermo Scientific, Cat. No. AB-0580)

- Strips of 8 Thermo-Tubes & Domed Caps (Thermo Scientific, Cat. No. AB-0266)

Materials

- LIZ 500 size standard (ABI, Cat. No. 4322682.)

- HiDi formamide (Applied Biosystems)

- Sterile Elga water

Method

1. Plan multiplexing of samples if the product sizes are already known.

2. Pipette 1µl of each sample into a tube or a well of a 96 well plate according to the multiplexing plan. Where 6 samples are multiplexed, there are 6 µl of PCR product from each sample per tube or well.

2. Add water to give 40x dilution of the sample. If 6 samples are multiplexed, 34µl of water is required.

3. Vortex and centrifuge the tube or plate.

4. Prepare a new tube and pipette 8.9µl of HiDi formamide and 0.1µl of LIZ 500 size standards. HiDi formamide and LIZ 500 can be mixed before and stored at -20℃.

5. Aliquot 1µl of the dilution mix made in step 2 into the tube prepared in step 4.

6. Vortex and centrifuge the tube or plate before submission to the Genotyping service at JIC.

3.6.3 SCORING THE GENOTYPE RESULTS

The sizes of the PCR products were measured using an ABI 3730XL sequencer and were visualized as electropherograms using Gene Mapper (v.3.7) software. Manual scoring was carried out at the developing of NIL when the number of markers and samples were manageable.

Automated scoring was introduced for mapping populations to deal with the large number of samples. 3000 individuals were genotyped with 14 markers. These markers were designed for the convenience of multiplication of 14 markers.

The method is developed from the JIC SOP CG-IB-013 version 003 “ SSR and SSCP genotyping and analysis using multiple fluorescence dyes and automated genotype scoring using ABI Genemapper(V4)” written by Kwangsoo Cho and Rachel Wells.

The technical details of the scoring methods are described in Appendix 7.

3.7 MEASURING FATTY ACID COMPOSITION

Nine FAs were used for tor QTL analysis: 16:0, 18:0, 18:1, 18:2, 18:3, 20:0, 20:1, 20:2 and 22:1 and their percentage of the total was calculated. Of them, 18:1, 18:2 and 18:3 were focused on in this project to detect loci involved in the desaturation pathway. The ratio of 18:2/18:1 and 18:3/18:2 were also calculated.

3.7.1 FATTY ACID METHYL ESTER PREPARATION

Seed lipid methanolysis was performed on seeds of *A. thaliana* in order to measure the composition of fatty acids by Gas Chromatography (CG) Flame Ionization Detector (FID). The method is based on the standard operation procedure at JIC (Reference No. CG-IB-019) written by Carmel O'Neill.

The procedure was altered twice over the period of the project. The first and second methods were used during NIL development. The original protocol was written for use of the GC Mass Spectrometry (MS) machine at JIC, however to achieve accurate Fatty Acid Methyl Ester (FAME) quantification, GC FID was adopted. Therefore the protocol was adapted for the use of GC FID at Rothamsted Research Institute. The third method was developed to aid management of a large number of samples during the fine mapping stage.

Here the third method is explained as this is the most efficient method of the three. The other two methods are described in Appendix 8.

<u>Equipment</u>

- TUB1202 Trident Vials, 3.5ml Tall with Screwcap (SCIENTIFIC LABORATORY SUPPLIES LTD Cat# 652050-1)

- Hot block (5x4) x 2

- Vortex

- Pipettor

- Glass Pasteur pipettes

- Glass beaker (100ml)

- Set of vials for GC analysis

    2 ml vials (Chromacol Cat# 2-SVW)

    200 µl glass inserts (Chromacol Cat# 02-MTVWG)

    Blue Screw Caps (Chromacol Cat# 9-SC(B)-ST1)

    Plastic foot/adaptor (Chromacol Cat# MTS-1)

Alternatively

    Pre-fixed EPA Screw Top Vials (Chromacol Cat# 03-FISV)

    Blue Screw Caps (Chromacol Cat# 9-SC(B)-ST1)

- GC vials and crimp caps for storage of backup samples

- Crimper

- Self adhesive label 8mm diameter

- Vinyl tape

- Heat proof gloves

Materials

- 3N Methanolic HCl (Supelco 33050-U 400ml, store @ 4°C)

- Iso-Hexane high purity reagent (Fischer Scientific H/0408/17 2.5l store @ room temperature)

- Methanol Analytic reagent (Fischer Scientific M/4000/PP17 2.5l store @ room temperature)

- NaCl 0.9% w.v.

- FAME Mix Rapeseed oil standard (Supelco, cat# 07756)

-cis-11,14-Eicosadienoic acid methyl ester (Sigma Chemical cat# E7877)

Method

Preparation:

- Rinse three glass beakers with Chloroform and air dry

- Assemble vials for GC analysis and label them

- Label TUB1202 trident vial with vinyl tape

- Label vials for GC analysis and for back up with self adhesive label

- Measure 25mg of seeds into pre-labelled trident vials.

FAME preparation

1. Place hot block in the fume hood cabinet and preheat to 80℃.

2. Dilute methanolic HCl (3N) to 1N using methanol in a cleaned glass beaker in the fume hood.

3. Using a glass Pasteur pipette add 0.5ml of 1N methanolic HCl into each trident vial containing seeds and screw the lid tightly to avoid evaporation.

4. Place the tubes in the preheated hot block and check the lids after 5 minutes, as they can loosen when heated. Wear heat proof gloves when tightening lids. Check again after a further 5 minutes.

5. Heat the samples for two hours in total.

6. Remove samples from the hot block and leave them to cool completely for approximately 15minutes.

7. To each add 0.5ml of 0.9% NaCl.

8. Add 0.5ml of iso-hexane and close the lid.

9. Vortex for 20 seconds.

10. To separate the phases allow to stand on the bench for approximately 15 minutes.

11. Aliquot 50µl of the clear, colourless supernatant solution into a vial for GC analysis and the remainder into the prepared backup vial.

12. To dilute samples for GC analysis add 50µl of iso-hexane.

13. Close the lids tightly and store at -20℃.

FAME samples were submitted to Rothamsted Research (Harpenden, UK) via next day deliver courier service together with FAME Mix Rapeseed oil standard (Supelco, cat# 07756) on dry ice.

3.7.2 GAS CHROMATOGRAPHY ANALYSIS OF FAMES

Gas liquid chromatography was performed at Rothamsted Research Institute by Frederic Beaudoin on a Hewlette-Packard 6890 series gas chromatography machine using an Alltech AT-225 (30m x 0.32mm x 0.3µm) capillary column. The details of the settings are described in Appendix 8.

The chromatograms were analysed on the Agilent ChemStation software. Retention times and identity of each FAME peak was calibrated using the FAME Mix Rapeseed oil standard supplemented with 2% w/w methyl 11,14-eicosadienoate (C20:2 n-6; Fluka Cat# 17272). 1 mM Methyl heptadecanoate (C17:0; Sigma cat# H4515) was added to some samples as an internal standard.

19 peaks were identified in the Arabidopsis seed FAME samples. Only 16 peaks, which had more than 0.1% of the total peak area after integration, were selected. The corresponding FAs were: 16:0, 16:1, 18:0, 18:1(n-9), 18:1(n-7): 18:2, 18:3, 20:0, 20:1(n-9), 20:1(n-7), 20:2, 20:3, 22:0, 22:1, 24:0 and 24:1.

The data were processed prior to QTL analysis as explained in the next section.

3.7.3 PREPARATION OF DATA

The peak areas of 18:1(n-9) and 18:1(n-7), 20:1(n-9) and 20:1(n-7) were added together for 18:1 and 20:1 totals respectively.

The FAs typically found in *Arabidopsis* are 16:0, 18:0, 18:1, 18:2, 18:3, 20:0, 20:1, 20:2 and 22:1. Other minor FAs were excluded and contents of these nine FAs were re-calculated without them.

3.8 MEASURING OIL CONTENTS

The seed oil contents were measured by nuclear magnetic resonance spectroscopy (NMR). The methodology is based on the standard operation procedure CG-IB-017 written by Carmel O'Neill.

Equipment

- Oxford Instruments MQC bench top NMR analyser with 10mm diameter probe

- 10mm diameter glass tubes

- Tuning sample containing mineral oil

- Canadian Canola seed standards

- Top loading balance with reads to two decimal places

- USB memory stick to transfer data

Method

1. Generate a calibration curve before taking sample readings by following the instructions written in the SOP. This calibration should be used within an same experiment.

2. Open the analysis session and tune with the tuning sample by following the instruction shown by the machine.

3. Start analysis with three standards, cl1, cl3, cl8 whose oil contents were 39.76%, 43.67%, 52.07% respectively and check if the measurements also show the similar figures. If not close figures, generate new a new calibration and recheck.

4. Weigh approximately 200mg of seeds into a glass tube and record the weight. Prepare three samples at a time. Make sure the tube is labelled 1,2,3 etc. to avoid errors.

5. Do three measurements for each sample. Type the name and the weight of the sample into the spread sheet on the screen and insert the glass tube with sample 1 into the probe.

6. Remove the tube immediately from the probe following instructions on the machine to avoid heating seeds. The temperature of the seeds affects the readings,

therefore after each measurement, samples should be left to cool before the second measurement.

7. Repeat Step 3 every two hours and tune the machine with the tuning sample as directed by the machine.

8. When finished analyzing, transfer the data from the Log File onto a USB memory stick. The data can then be opened in Excel.

CHAPTER 4. IDENTIFICATION AND REFINEMENT OF NOVEL QTL CONTROLLING FATTY ACID PROFILE IN SEEDS OF *A. THALIANA*

4.1 INTRODUCTION

The traits studied in this project were quantitative and are likely to be regulated by multiple loci within the genome, therefore in order to identify these loci, quantitative trait locus (QTL) analysis was applied.

QTL mapping comprised two steps, the construction of linkage maps using molecular markers and the association of phenotype with these markers. Mapping the markers required genotyping of markers on the population initially, followed by the linkage analysis of markers to produce genetic linkage maps based on the calculation of recombination frequencies, which indicate the position and relative genetic distances between markers along chromosomes (Collard et al., 2005). QTL analysis then utilizes the associations between the quantitative value of a trait and marker alleles that are segregating in the population (Kearsey and Farquhar, 1998).

There are three methods of QTL mapping available in MapQTL5.0, non-parametric mapping (Kruskal-Wallis analysis), interval mapping and multiple-QTL model (MQM) mapping. Nonparametric mapping uses the rank sum test of Kruskal-Wallis. The test is carried out individually on each locus and the linkage map is only used for aligning the loci. The interval mapping method calculates the likelihood of the presence of a segregating QTL for every centiMorgan (cM). The genetic effect of the QTL and the residual variance are taken into account. This likelihood is then compared to the likelihood of the absence of QTL with genetic effect by calculating the likelihood ratio statistic, which is termed the logarithm of odds (LOD) score. The position showing the largest LOD score on the linkage group is estimated as the location of the QTL on the map. MQM introduces the idea of a cofactor to the data produced by interval mapping. When a QTL detected in interval mapping accounts for a large percentage of variation, a linked marker to that QTL is used as a cofactor (with additive and dominant gene action) to facilitates the search for other segregating QTLs (Van Ooijen, 2004).

Interval mapping was used for the work in this project to exploit the linkage maps already developed, which were relatively robust as the physical positions of markers were known. However, the physical coordinates of the genetic markers used are

based on the Columbia genome, therefore there is a risk of positioning error of markers due to genomic variation between different accessions.

Subsequently, MQM mapping was carried out on those populations and traits which had large effect QTL close to the loci of known major effect genes, such as *FAE1, FAB2, FAD2* and *FAD3.* MQM was used to identify other segregating QTL which could have been masked by these large effect QTL during interval mapping.

Six new recombinant inbred (RI) populations were developed by Carmel O'Neill and their utility for quantitative traits was tested using flowering time in vernalized plants. New loci were identified in addition to those coincident with QTL detected in previous studies (O'Neill et al., 2008).

Having proven their usability in complex quantitative trait analysis, these RI populations were used for the mapping of oil content and FA composition by O'Neill in 2008 as a pilot research, and the result revealed the existence of novel loci. To search these novel loci in detail, the marker density of linkage maps was increased by adding further markers to all six populations.

Previous QTL mapping studies for seed lipid characteristics in *Arabidopsis* used a single population derived from two lab strains (Hobbs et al., 2004) or four populations crossed between four wild accession and a lab strain (Sanyal and Randal, 2012). The genetic variation in the six populations used in this study could be more diverse and these populations showed better phenotype distribution within each population, therefore more stringent results can be expected.

Genotyping of lines and linkage analysis for the KB, NG and TJ populations were performed by O'Neill. I conducted this work on the populations: CA, SG and WC.

QTL analysis for 18:1, 18:2, 18:3, 18:2/18:1, 18:3/18:2 and oil content were performed with these improved genotype files and linkage maps. The phenotype data was the same as that used in the previous study measured by O'Neill. For consistency, only the replicates grown under glasshouse conditions were used throughout the work described here.

Later the results of QTL analysis for 17 lipid related traits for all six populations used both the glasshouse and controlled environment grown replicates, were published for the understanding of global genetic architecture of lipid biosynthesis (O'Neill et al., 2012)

After QTL analysis, the identified loci were screened to check if they were close to known major genes and those deemed potential novel loci were selected for the next stage. The development of the near isogenic lines for the QTL is described in Chapter 5.

4.2 MATERIAL AND METHODS

4.2.1 PLANT MATERIAL

Six *A. thaliana* RI populations, CA, SG, WC, KB, NG and TJ populations (O'Neill et al., 2008) were used for QTL analysis. Each population consisted of a set of 94 RI lines (RIL).

The CA and WC populations were grown together in the glasshouse in a randomized trial with two blocks.

SG, KB, NG and TJ populations were grown together in the glasshouse in three randomized blocks.

The temperature of the glasshouse was kept at 22°C day/18°C night and overhead sodium lamps were used to maintain a 16hr photoperiod.

4.2.2 PHENOTYPIC DATA

Oil content and FA compositions were measured by O'Neill using the methods described in Section 3.8 for oil content and Appendix 8 for FA compositions(O'Neill et al., 2012). The Shapiro-Wilk test for Normality was performed for each population for each trait to examine if their distribution was non normal ($P<0.05$). Histograms were produced to identify outliers. Non normal distributions were corrected using $\log_{10}$ transformation or square root of the figures.

Phenotypic data of parental replicates within blocks were used to assess Gxblock interactions by ANOVA using Genstat 13.

Populations and traits showing a G x block interaction or non-normal distribution which could not be corrected even with transformation were noted and taken into account at the final selection of target QTL.

Heritability of each trait was calculated from the variance assessed through ANOVA. Variation of the phenotype was derived from two sources, genetic and environmental. The residual represents the environmental variance , therefore the genetic effect on the phenotypic measurement can be calculated from the following equations:

variety ms=σ^2+ nσ_g^2, residual ms=σ^2

σ_g^2= (var.ms - residual ms)/n

Proportion of Heritable variation= 100 (σ_g^2 / (σ_g^2+σ^2))

4.2.3 INCREASING MARKER DENSITY

The new Insertion/Deletion (InDel) markers were designed using the MSQT website (http://www.msqt.weigelworld.org/) and the RILs were genotyped with these markers by the Genome Laboratory Services (TGL), John Innes Centre, which is now called Genome Analysis centre (TGAC), as described in Section 3.6. The numbers of new markers are shown in the Table 4-1.

The genotype results received from TGL were checked and prepared in genotype file format for linkage analysis. The linkage analysis of the populations: CA, SG and WC were performed in this project and the KB, NG and TJ populations were analysed by O'Neill in 2008.

Table 4-1 Number of markers in population

Population	Previous	New	Total
CA	47	24	**71**
SG	48	26	**74**
WC	67	15	**82**
KB	52	24	**76**
NG	58	19	**77**
TJ	52	42	**94**

4.2.4 LINKAGE MAPPING

Joinmap version 3.0 was used to produce linkage maps for the CA, SG and WC populations.

Locus files were prepared for each chromosome separately. Heterozygous alleles were converted into unclassified (U) and default settings were applied.

If the default settings did not allow a single linkage group (LG) to be formed, the calculation options were adjusted step by step until it successfully formed one linkage group. Details of the adjustment of settings can be found in Appendix 1.

4.2.5 QTL ANALYSIS

QTL analysis was performed for all six populations using linkage maps of CA, SG and WC populations analysed in this project and also those of KB, NG and TJ population produced by O'Neill. Phenotype data was also provided by O'Neill.

Interval mapping was carried out for the traits 18:1, 18:2, 18:3, 18:2/18:1, 18:3/18:2 and oil content using MapQTL5. Permutation tests (10,000 times) were performed to determine the significance threshold (*P*=0.05) of the LOD score.

Subsequently, for large effect QTL close to major loci identified in interval mapping, MQM mapping was performed to separate the large effect QTL from other QTL which could have been masked by the large effect QTL in interval mapping.

4.2.6 SELECTION OF TARGET QTL FOR NIL DEVELOPMENT

The purpose of this QTL analysis was to identify novel QTL involved in the regulation of the PUFA desaturation pathway. Therefore, identifying already known major genes was important in order to avoid them at the selection stage of target QTL. Anticipated major genes which could appear as QTL were *FAB2, FAD2*, *FAD3 and FAE1* and their physical positions in the Columbia genome were know such as *FAB2* at 18.12Mb on chromosome (Ch)2*, FAD2* at 3.86Mb on Ch3, *FAD3* at 12.78Mb on Ch2 *and FAE1* at 16.49Mb on Ch4. Because the physical positions of both the major genes and the genetic markers used for the mapping were known, the bases of QTL close to major genes were considered to be due to major genes. QTL close to the major genes tended to have high LOD scores and affected across several markers adjacent to them, therefore not only loci with the highest LOD scores but also those within the influence of QTL were avoided in the selection of QTL for further study.

The reliability of those newly identified QTL, which were not coincident with the loci of known major genes, was further examined by checking whether they appeared across multiple related traits.

4.3 RESULTS

The phenotypic data was statistically analysed and normalized prior to QTL analysis. Linkage maps were produced from the improved genotype files.

The phenotype data, genotype data and linkage maps were used together to carry out QTL interval mapping within MapQTL5.0 and the results were summarised according to the traits: 18:2, 18:3, 18:2/18:1, 18:3/18:2 and oil content.

The detected QTL were evaluated whether they were coincident with known major genes of the PUFA desaturation pathway using their physical position. When the QTL had large effects and extended to several markers, MQM mapping was performed to identify QTL which could be masked by the large effect QTL. The expected major genes were *FAB2, FAE1, FAD2 and FAD3,* and QTL associated with these were avoided to select novel QTL for this project.

Following consideration of all the information from the analysis above, QTL were chosen to be taken forward for NIL development.

4.3.1 STATISTICAL ANALYSIS OF PHENOTYPE

4.3.1.1 PARENTAL LINES

The means and standard deviations of the phenotypes of parental lines were shown in Table 4-2.

Sorbo had far more 18:1 content (19.75%) than other lines and Cvi-0 the least (12.13%). This might reflect the 18:2/18:1 which was the reverse; Sorbo had the lowest ratio at 1.56 and Cvi-0 had the largest ratio with 2.56. The SG population showed the largest difference of 18:1 (5.77%) between the parental lines.

For 18:2 Br-0 had the lowest content with 22.79% and Mz-0 the largest at 32.28%. The difference in 18:2 content between the parental pairs was greatest for Kondara and Br-0 at 8.03% followed by the Sorbo and Gy-0 at 6.15%.

Wt-5 had the highest 18:3 content with 25.07% and an 18:3/18:2 ratio of 0.99, while Cvi-0 had the lowest with 19:97% and 0.62, respectively.

The CA population had the parents with the lowest oil contents (35.04% for Cvi-0 and 38.32% for Ag-0), while the NG population had the two highest oil containing lines as parents (51.28% for Nok-3 and 51.43% for Ga-0).

Table 4-2 The phenotypic means and standard deviations of the twelve parental lines and heritabilities of the six populations

		18:1	18:2	18:2/18:1	18:3	18:3/18:2	oil
Cvi-0	mean	12.13	30.87	2.56	19.97	0.62	35.04
	SD	1.05	0.75	0.26	0.23	0.04	2.64
Ag-0	mean	15.58	25.32	1.63	20.13	0.78	38.32
	SD	0.63	0.20	0.07	0.57	0.03	4.98
heritability h^2%		57.1	79.6	55.8	89.2	84.1	20.7
Sorbo	mean	19.75	30.73	1.56	20.55	0.67	44.80
	SD	0.52	1.61	0.11	0.97	0.03	0.98
Gy-0	mean	13.98	24.58	1.77	21.97	0.90	45.25
	SD	1.18	2.16	0.17	1.44	0.10	0.84
heritability h^2%		67.5	19.2	23.1	8.2	64.4	54.9
Wt-5	mean	12.98	25.45	1.97	25.07	0.99	41.18
	SD	0.91	0.89	0.20	0.86	0.03	1.85
Ct-1	mean	16.49	28.88	1.77	21.74	0.75	40.81
	SD	1.28	1.84	0.25	1.02	0.03	1.54
heritability h^2%		42.8	0.4	19.8	27.5	51.6	34.6
Kondara	mean	16.36	30.82	1.89	20.57	0.67	47.24
	SD	0.74	1.18	0.14	0.62	0.03	1.34
Br-0	mean	15.88	22.79	1.44	21.19	0.93	42.75
	SD	0.76	0.66	0.09	0.57	0.02	1.03
heritability h^2%		83.3	55.5	60.7	31.0	65.4	60.5
Nok-3	mean	16.43	27.26	1.67	23.71	0.88	51.28
	SD	1.40	1.75	0.10	1.57	0.05	2.32
Ga-0	mean	14.63	27.65	1.90	25.06	0.91	51.43
	SD	1.35	0.82	0.15	1.68	0.06	2.93
heritability h^2%		56.9	37.1	58.2	14.8	43.4	4.5
Ts-5	mean	13.82	27.34	1.98	22.56	0.83	42.53
	SD	0.47	1.06	0.10	1.02	0.03	0.83
Mz-0	mean	14.97	32.28	2.16	20.32	0.63	44.22
	SD	0.56	0.72	0.06	0.78	0.03	1.12
heritability h^2%		43.0	21.5	21.4	37.7	31.3	38.7

ANOVA was carried out to examine if there were G x block interactions, using replicate plants of parental lines. G x block interactions were observed in the KB population for 18:1, 18:3 and 18:3/18:2, and NG population for 18:3, 18:3/18:2 and oil content. The details of ANOVA results are given in Appendix 2.

4.3.1.2 RI POPULATIONS

The Normality test revealed that there were nine non-normal distributions within 5 populations for 5 traits, which were 18:2, 18:3 and 18:3/18:2 in CA population, 18:1 and oil content in SG population, 18:3 in WC population, oil content in NG population and 18:1 and 18:3 in KB population.

Oil content in SG population and 18:3 in WC population were corrected by removing outliers. 18:1 in KB population was normalized by $\log_{10}$ transformation. The results of Shapiro-Wilk test for Normality of original data and transformed data can be found in Appendix 2. Histograms found in Appendix 2 are produced from original data.

The rest of the data could not be normalized, although QTL analysis was performed using these data. All of the non-normally distributed traits in CA population and 18:1 in SG population had bimodal distribution. Oil content in NG population was skewed towards the larger values while 18:3 in KB population slightly plateaued in the middle.

The proportion of heritable variation calculated from the variance through ANOVA can be seen in Table 4-2. It revealed that the CA population had higher heritability across FA traits studied from 55.8% in 18:2/18:1 to 89.2% in 18:3, followed by the KB population which had from 31.0% in 18:3 to 83.3% in 18:1. Regarding oil content, the CA population had only 20.7% while the KB population had 60.5%. The WC population had the least heritability of the six populations where the highest 51.6% in 18:3/18:2 and lowest 0.4 for 18:2. The table with the values for variety mean squares, residual mean squares, variances from genotype and proportions of heritable variation is at the top of Appendix 2.

4.3.2 LINKAGE ANALYSIS

Linkage maps were produced by JoinMap3.0, using the physical position of markers as a reference. The resulting linkage map charts are shown in Fig.4-1 and the physical positions of markers are indicated in blue font in brackets. Some markers did not link in the expected physical order: CA population on LG1, LG4 and LG5; SG population on LG4 and LG5; WC population on LG2 and LG5. Therefore the order of markers in the pale blue boxes on Fig.4-1 was fixed according to the Columbia physical order. One marker on WC LG4, t19f6ind3-3, was removed because it did not align with its physical position and fixing it hugely disturbed the integrity of the other markers.

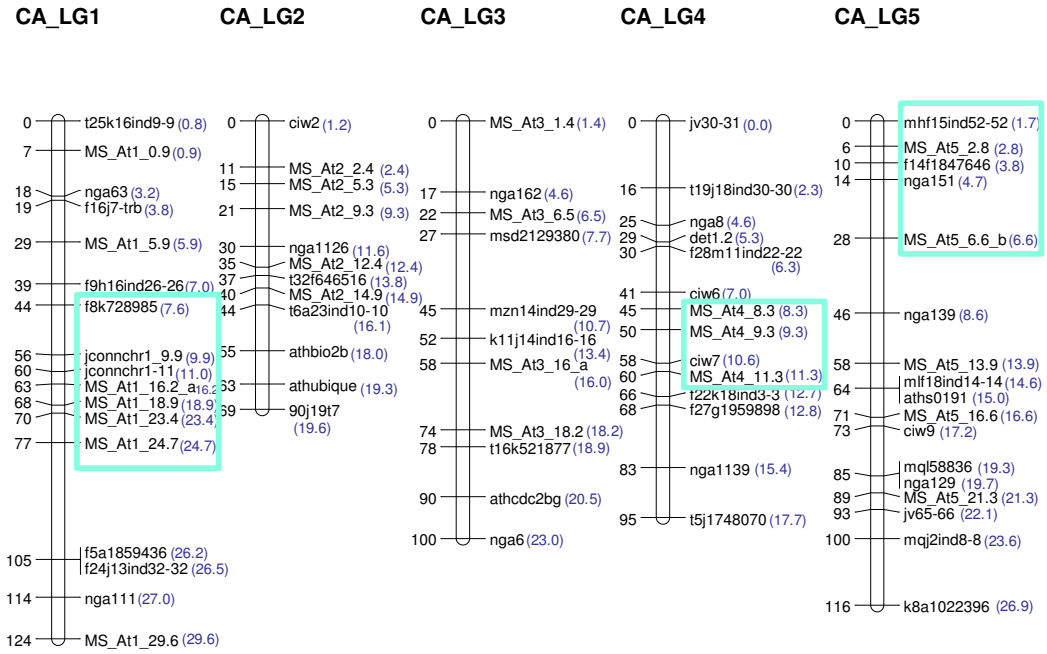


Fig.4-1-a Linkage map of the CA population

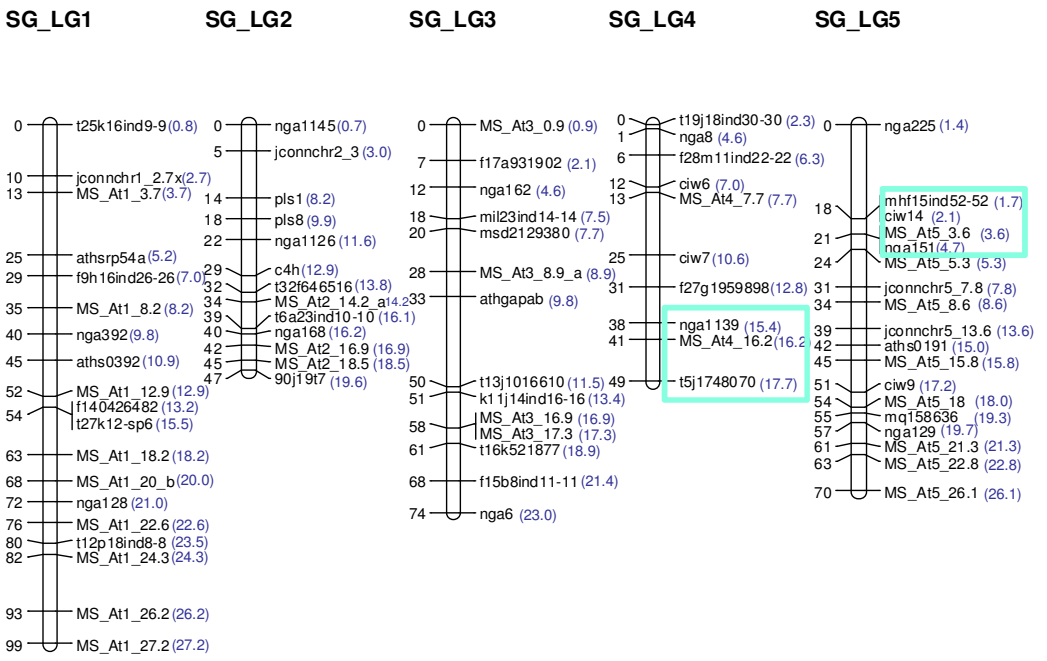


Fig.4-1-b Linkage map of the SG population

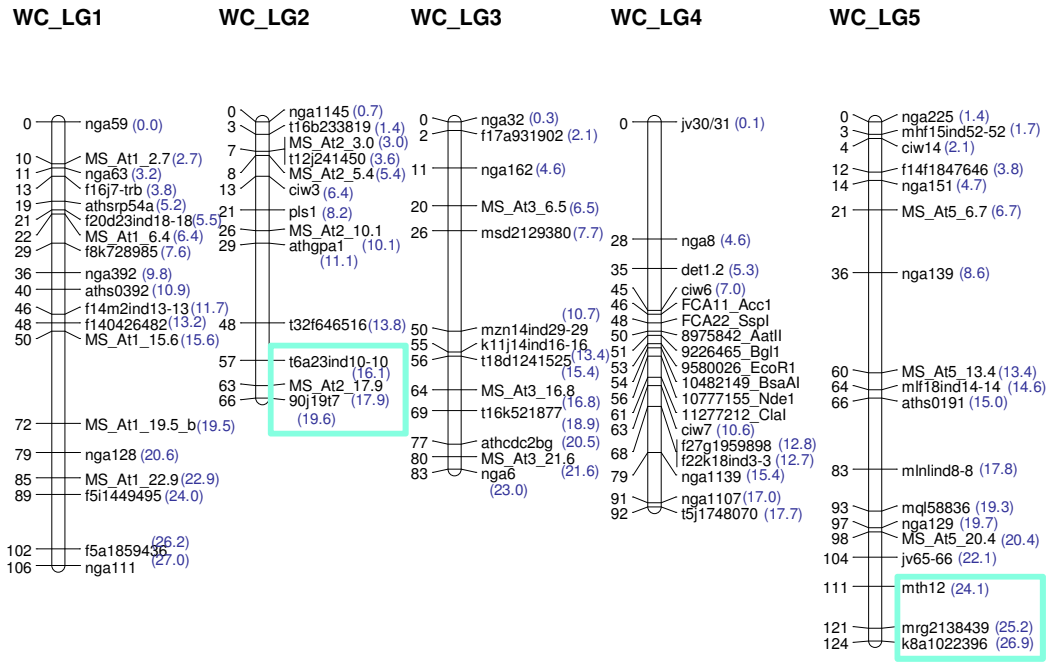


Fig.4-1-c Linkage map of the WC population

By increasing the numbers of markers, the marker density of most of the linkage groups was increased as shown in Table 4-3 below with the exception of LG4 in the WC population. That is because WC LG4 already had markers more frequently than the other LGs, therefore no new markers were added. Moreover one marker was removed during the linkage analysis as explained previously.

Table 4-3 Marker density of linkage group for three populations

Population	LG	Marker density (cM/marker)	
		Past	Now
CA	1	11.50	7.29
	2	9.14	5.75
	3	11.86	9.09
	4	8.36	6.79
	5	8.25	6.82
SG	1	7.50	5.21
	2	4.50	3.62
	3	6.40	5.29
	4	6.13	4.90
	5	6.40	3.89
WC	1	7.64	5.58
	2	7.56	5.08
	3	8.40	6.38
	4	4.84	5.11
	5	9.31	6.89

4.3.3 INTERVAL MAPPING

QTL analysis for the six populations was performed using linkage maps of CA, SG and WC population as described in the previous section. Linkage maps of KB, NG and TJ population and phenotype data were provided by O'Neill in 2008.

Genomewide permutation tests were carried out to determine the significance threshold of LOD score (*P*=0.05) for each trait in each population as given below in Table 4-4.

Table 4-4 List of the significance thresholds of LOD scores for six traits in six populations

	Significant threshold of LOD score(p=0.05)					
population	CA	SG	WC	KB	NG	TJ
18:1	2.6	2.6	2.6	2.6	2.6	2.5
18:2	2.7	2.5	2.6	2.6	2.7	2.6
18:2/18:1	2.6	2.5	2.6	2.6	2.6	2.5
18:3	2.7	2.5	2.6	2.4	2.6	2.5
18:3/18:2	2.7	2.6	2.6	2.6	2.6	2.6
oil	2.6	2.6	2.6	2.5	2	2.5

QTL with LOD scores above the threshold of significance level (*P*=0.05) were extracted and examined in the following sections.

A total 55 QTL were detected from six populations for the traits of interest, 19 QTL were found close to the loci of known major genes.

The QTL map charts are found in Appendix 2.

4.3.3.1 18:1

QTL were found in all six populations for 18:1 as shown in Table.4-5-a.

The QTL found in the CA population was at the end of LG4, as there were no adjacent markers to support the results, the effect of QTL was not as certain as other QTL which spanned multiple markers.

A large QTL with a LOD score of 17.62 was detected in the SG population on LG4 at 40.77cM. This was coincident with the locus of the known major gene *FAE1* which is involved in elongation of the FA carbon chain from 18:1 to 20:1.

On LG1 in the WC population, there was a minor QTL (LOD=2.96) which accounted for 13.6% of the variation.

Two QTL found in the KB populations were both close to the locus of major genes: *FAD2* and *FAE1*.

There was a large QTL (LOD 12.48) which accounted for 45.7% of the variation in the NG population on LG5, notably it did not coincide with other major genes.

The TJ population had five QTL and none of them were close to major genes. Among them one on LG1 had a relatively stronger QTL (LOD=4.15) in the middle of the linkage group than the second QTL identified within this population.

Table 4-5-a List of QTL for 18:1

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
18:1	CA	LG4	jv30-31	3.51	15.8	0.00	end of LG
	SG	LG4	f27g1959898	5.12	22.7	31.43	*FAE1 non normal distribution*
			nga1139	10.36	40.2	37.75	
			MS_At4_16.2	17.62	60.3	40.77	
			t5j1748070	8.13	36.8	48.50	
	WC	LG1	MS_At1_6.4	2.96	13.6	21.92	
	KB	LG3	f17a931902	3.09	15	9.12	*FAD2*
			nga162	4.58	20.7	16.02	
		LG4	MS_At4_13.2	2.71	12.9	57.19	*FAE1*
			nga1139	9.27	37.6	72.25	
			MS_At4_16.1	7.41	31.3	76.18	
	NG	LG5	nga225	8.8	35.3	0.00	
			mhf15ind52-52	8.22	33.2	3.52	
			ciw14	8.2	33.2	4.96	
			f14f1847646	12.41	45.7	9.05	
			nga151	12.48	45.7	11.98	
			nga139	3.78	17.2	33.96	
	TJ	LG1	nga392	3.1	14.1	31.39	
			aths0392	4.5	19.8	35.62	
			MS_At1_11.2	4.13	18.4	36.64	
			MS_At1_12.9	4.15	19.1	44.67	
			t27k12-sp6	2.89	13.4	48.40	
		LG1	f5i1449495	2.57	11.8	67.93	
		LG4	MS_At4_1.1	2.56	12.0	7.10	
		LG5	f20d23ind18-18b	3.32	15.0	65.03	
			nga129	3.5	15.8	66.50	
			MS_At5_20.4	3.46	15.6	66.55	
		LG5	mth12	3.68	16.5	84.66	

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.3.2 18:2

As seen in Table 4-5-b, the QTL found in the CA population on the lower part of LG1 had a LOD score of 5.03. The large QTL found in the same population on LG2 was close to the known major effect gene *FAD3* which is responsible for the desaturation of 18:2 to 18:3.

Two very minor QTL were detected in the SG population on LG1 and LG2 with LOD scores of 2.73 and 2.65 respectively. The position of the QTL on LG2 overlapped the *FAB2* locus which is responsible for the desaturation from 18:0 to 18:1 and it could affect composition of 18:2 content which is in the downstream in the pathway. The large QTL on the top half of LG3 was close to the major gene *FAD2* which is responsible for desaturation of 18:1 to 18:2. The QTL on the bottom half was identified as a different QTL. The QTL detected in the bottom of LG4 overlapped with the *FAE1* locus.

A large QTL was detected on LG3 in the KB population which was close to the locus of *FAD2*.

The positions of all QTL found in the NG and TJ populations did not coincide with major genes.

Table 4-5-b List of QTL for 18:2

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
18:2	CA	LG1	f5a1859436	5.03	21.8	105.17	
			f24j13ind32-32	5.03	21.8	105.18	
		LG2	MS_At2_5.3	2.73	13.0	14.85	*FAD3 non normal distribution*
			MS_At2_9.3	4.13	18.8	21.30	
			nga1126	13.06	47.5	29.85	
			MS_At2_12.4	13.47	48.3	34.88	
			t32f646516	21.59	65.7	37.37	
			MS_At2_14.9	11.4	43.1	40.47	
			t6a23ind10-10	8.24	33.6	44.08	
	SG	LG1	MS_At1_18.2	2.73	12.5	62.85	
		LG2	MS_At2_16.9	2.65	12.2	42.27	*FAB2*
		LG3	MS_At3_0.9	3.11	14.2	0.00	*FAD2*
			f17a931902	4.06	19.3	7.15	
			nga162	4.73	20.8	12.06	
			mil23ind14-14	2.7	12.5	18.04	
			msd2129380	4.26	18.9	20.13	
			MS_At3_8.9_a	3.85	17.2	28.48	
			athgapab	3.25	14.8	32.77	
		LG3	k11j14ind16-16	2.51	11.6	51.04	
			MS_At3_16.9	3.13	14.2	58.13	
			MS_At3_17.3	3.13	14.2	58.30	
		LG4	nga1139	2.98	14.1	37.75	*FAE1*
			MS_At4_16.2	3.04	14.3	40.77	
	WC	LG2	90j19t7	3.15	14.4	66.07	end of LG, skewed distribution
	KB	LG2	90j19t7	2.64	12.6	57.09	end of LG
		LG3	nga32	5.47	24.4	0.00	*FAD2*
			f17a931902	8.68	36.0	9.12	
			nga162	11.13	43.6	16.02	
			msd2129380	4.24	20.5	29.58	
	NG	LG3	t13j1016610	3.03	13.8	45.83	
			MS_At3_18.3	3.27	14.8	55.41	
			t16k521877	3.89	17.4	57.67	
	TJ	LG1	MS_At1_8.2	4.47	19.7	23.64	
			nga392	2.98	13.6	31.39	
		LG3	MS_At3_6.1	4.34	19.3	15.41	
			msd2129380	4.8	20.9	20.50	
		LG5	nga225	3.5	15.8	0.00	end of LG

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.3.3 18:2/18:1

QTL detected in the CA, SG and KB populations spanned regions containing known major genes as shown in Table 4-5-c.

A large QTL was identified in the NG population on LG5 with a LOD score of 11.82 and accounted for 44 % of the variation.

One of the QTL found in the TJ population on the lower part of LG5 was relatively significant with a LOD score 4.81 than other QTL found within the population.

Table 4-5-c List of QTL for 18:2/18:1

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
18:2/18:1	CA	LG2	t32f646516	2.79	12.9	37.37	*FAD3 non normal distribution*
			MS_At2_14.9	3.58	16.1	40.47	
		LG3	nga162	2.64	12.2	16.77	*FAD2*
	SG	LG3	MS_At3_0.9	4.14	18.4	0.00	*FAD2*
			f17a931902	3.22	15.4	7.15	
			nga162	6.2	27.0	12.06	
			mil23ind14-14	4.93	22.2	18.04	
			msd2129380	3.76	16.9	20.13	
			MS_At3_8.9_a	2.93	13.7	28.48	
		LG4	f27g1959898	2.55	12.0	31.43	*FAE1*
			nga1139	3.4	15.5	37.75	
			MS_At4_16.2	6.04	26.2	40.77	
			t5j1748070	3.61	18.7	48.50	
	KB	LG2	90j19t7	3.02	14.6	57.09	end of LG
		LG3	nga32	4.67	21.8	0.00	*FAD2*
			f17a931902	9.05	38.4	9.12	
			nga162	13.13	48.9	16.02	
			msd2129380	2.7	14.0	29.58	
	NG	LG5	nga225	7.64	31.5	0.00	
			mhf15ind52-52	6.88	28.6	3.52	
			ciw14	6.73	28.1	4.96	
			f14f1847646	11.82	44.0	9.05	
			nga151	11.17	42.2	11.98	
			nga139	4.35	19.6	33.96	
	TJ	LG1	MS_At1_12.9	2.62	12.7	44.67	
		LG1	f5i1449495	2.51	11.6	67.93	
		LG5	f20d23ind18-18b	4.5	19.8	65.03	
			nga129	4.81	21.0	66.50	
			MS_At5_20.4	4.72	20.6	66.55	
			MS_At5_21.3	2.62	12.0	69.76	
			jv65-66	2.56	11.8	75.24	
		LG5	MS_At5_22.8	3.23	14.8	77.52	
			mth12	3.54	16.0	84.66	
			MS_At5_26.1	3.22	14.6	93.08	

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.3.4 18:3

The large QTL found in the CA population with a LOD score of 32.68 spanned the locus of the major gene *FAD3*.

There were three QTL detected in the SG population on LG3, LG4 and LG5. The one on LG4 was close to the *FAE1* locus.

A small QTL was found in the WC population on LG1 (LOD=3.04).

A relatively large QTL was found in the TJ population on LG1 with a LOD value of 6.32.and it was not close to a major gene.

Table 4-5-d List of QTL for 18:3

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
18:3	CA	LG2	ciw2	2.92	13.3	0.00	*FAD3 non normal distribution*
			MS_At2_9.3	4.72	21.7	21.30	
			nga1126	14.7	51.4	29.85	
			MS_At2_12.4	15.31	52.8	34.88	
			t32f646516	32.68	80.1	37.37	
			MS_At2_14.9	10.93	42.0	40.47	
			t6a23ind10-10	9.31	37.7	44.08	
	SG	LG3	MS_At3_16.9	2.89	13.2	58.13	
			MS_At3_17.3	2.9	13.2	58.30	
			t16k521877	2.87	13.5	60.94	
		LG4	MS_At4_16.2	2.6	12.1	40.77	*FAE1*
		LG5	aths0191	3.08	14.2	42.16	
			MS_At5_15.8	2.81	12.9	45.36	
	WC	LG1	MS_At1_6.4	3.04	14.0	21.92	
	KB	LG5	MS_At5_20.0	2.6	12.4	53.89	
	NG	LG1	MS_At1_22.6	2.66	12.2	71.46	
	TJ	LG1	f9h16ind26-26	2.81	12.9	18.14	
			MS_At1_8.2	4.01	18.1	23.64	
			nga392	5.42	23.3	31.39	
			aths0392	6.32	26.6	35.62	
			MS_At1_11.2	5.66	24.3	36.64	

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.3.5 18:3/18:2

QTL close to *FAD2* were detected in the SG population and KB population on LG3 as shown in Table 4-5-e. A very large QTL observed in the CA population which had the LOD score of 32.44 was on LG2 and lay over the *FAD3* locus.

A QTL which had a LOD score of 6.11 mapped to the lower part of SG LG3 and accounted for 25.9% of the variation. Another significant QTL with a LOD score of 5.74 in the TJ population lay on LG1. Neither QTL had major gene loci under them.

Table 4-5-e List of QTL for 18:3/18:2

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
18:3/18:2	CA	LG1	f5a1859436	3.09	14.0	105.17	
			f24j13ind32-32	3.09	14.0	105.18	
		LG2	ciw2	2.83	13.0	0.00	end of LG
		LG2	MS_At2_9.3	4.75	21.6	21.30	*FAD3 non normal distribution*
			nga1126	15.13	52.4	29.85	
			MS_At2_12.4	16.49	55.4	34.88	
			t32f646516	32.44	79.9	37.37	
			MS_At2_14.9	12.16	45.4	40.47	
			t6a23ind10-10	9.62	39.7	44.08	
	SG	LG3	nga162	2.56	11.9	12.06	
			mil23ind14-14	2.56	11.9	18.04	
			msd2129380	3.61	16.3	20.13	
			MS_At3_8.9_a	3.53	16.0	28.48	
			athgapab	3.17	14.5	32.77	
		LG3	t13j1016610	4.03	17.9	50.50	
			k11j14ind16-16	3.79	17.1	51.04	
			MS_At3_16.9	6.1	25.8	58.13	
			MS_At3_17.3	6.11	25.9	58.30	
			t16k521877	5.1	22.4	60.94	
			f15b8ind11-11	3.86	17.5	67.97	
	WC	LG1	MS_At1_2.7	2.72	12.6	9.67	
	KB	LG2	athubique	3.32	15.5	53.88	near end
		LG2	90j19t7	3.43	16.2	57.09	
		LG3	nga32	7.38	31.8	0.00	*FAD2*
			f17a931902	11.01	43.8	9.12	
			nga162	13.02	48.7	16.02	
			msd2129380	4.55	21.4	29.58	
	NG	LG1	MS_At1_22.6	2.97	13.6	71.46	
		LG5	nga139	3	13.9	33.96	
	TJ	LG1	f9h16ind26-26	2.91	13.3	18.14	
			f8k728985	2.87	13.1	19.31	
			MS_At1_8.2	5.42	23.4	23.64	
			nga392	5.51	23.7	31.39	
			aths0392	5.74	24.5	35.62	
			MS_At1_11.2	4.96	21.6	36.64	
		LG3	MS_At3_6.1	3.45	15.7	15.41	
			msd2129380	3.53	15.9	20.50	
		LG5	nga225	3	13.7	0.00	end of LG

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.3.6 OIL CONTENT

Two oil QTL were found in the CA and WC populations at the similar marker locus on LG5 with similar LOD scores of 3.83 and 3.63 respectively. There was also a minor QTL on LG5 in the SG population however in a different position from these mentioned above.

In the KB population there was a large QTL (LOD=6.46) detected on LG2 and a small QTL (LOD=2.74) on LG5.

For the TJ population two QTL were mapped, on LG1 (LOD=3.52) and on LG2 (LOD=3.90).

Table 4-5-f List of QTL for oil content

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)	Note
oil	CA	LG5	k8a1022396	3.83	17.4	116.20	
	SG	LG5	MS_At5_21.3	2.79	13.0	60.716	
			MS_At5_22.8	2.66	12.5	62.895	
	WC	LG5	mth12	3.63	16.3	111.07	
			k8a1022396	3.18	15.2	124.46	
	KB	LG1	nga63	3.99	18.5	0.00	end of LG
		LG2	MS_At2_14.2	2.65	12.6	41.90	
			t6a23ind10-10	4.01	18.8	47.23	
			nga168	4.01	18.8	47.26	
			MS_At2_17.5	4.32	19.7	50.40	
			athubique	6.01	26.2	53.88	
			90j19t7	6.46	28.2	57.09	
		LG5	jv65-66	2.74	13.0	56.68	
	TJ	LG1	f5a1859436	2.61	12.1	83.25	near end
			nga111	3.52	15.9	87.80	
		LG2	MS_At2_9.3	3.45	15.6	28.57	
			pls8	3.9	17.4	30.63	
			nga1126	3.84	17.2	35.28	

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

4.3.4 MQM MAPPING

MQM mapping was performed for QTL which encompass loci of the known major genes which account for a large proportion of the total variance. Markers close to the major gene loci were used as cofactors to search for other QTL. Map charts are given in Appendix 2.

A QTL for 18:2 in the CA population on LG2 had a LOD score 21.59 and it spanned 23cM, within which *FAD3* locus (12.7Mb) lay. Marker t32f646516 (at 13.8Mb) was

used as a cofactor. The result of MQM mapping revealed that the adjacent loci did not have significant effects, indicating there were no other QTL masked by the effect of *FAD3* on LG2. The significance level of the QTL on LG1 became lower while two small QTL on LG3 were slightly improved.

A QTL for 18:3 and 18:3/18:2 in the CA population had a similar profile both in interval mapping and MQM mapping. The markers MS_At2_12.4 for *FAD3* and t32t646516 were used as cofactors. The results showed no other QTL near the QTL tested, however a further QTL was identified towards the end of LG2. Also a QTL on LG4 became significant after MQM mapping.

In the SG population, four QTL were studied. Following MQM mapping where MS_At4_16.2 (close to *FAE1)* was used as a cofactor on LG4, LOD scores of the adjacent markers for the QTL for 18:1 on LG4 became insignificant except the one at the bottom end. It also increased the significance level of QTL on LG3 and LG5. In the same population, 18:2 and 18:3/18:2 QTL were analysed using nga162 on LG3 close to *FAD2* locus as a cofactor. The results for 18:2 showed there were no significant QTL adjacent to the *FAD2* loci and proved that the QTL on the end of LG3 was another segregating QTL. The same results were observed for 18:3/18:2. For 18:2/18:1, MS_At4_16.2 on LG4 for the *FAE1* locus was used with nga162 on LG3 for *FAD2*, and no other significant QTL were observed.

In the KB population a QTL for 18:1 was examined using nga162 on LG3 near the *FAD2* locus, and MS_At4_16.1 and nga1139 on LG4 for the *FAE1* locus as cofactors. The result showed there was another minor but significant QTL on LG5.

QTL for 18:2, 18:2/18:1 and 18:3/18:2 from the KB population were analysed using nga162 as a cofactor and all showed there were no other QTL adjacent to the major gene loci. In 18:2 and 18:2/18:1, QTL in the end of LG4, close to the *FAE1* locus, had an increased LOD score.

4.3.5 SEQUENCE VARIATIONS AT MAJOR GENE LOCI

QTL corresponding to the major genes did not always appear in all six populations. To examine the reason behind it, sequences of these genes in each parental line were compared. Genomic sequencing of Kondara, Gy-0, Br-0, Ga-0, Mz-0, Ts-0, Sorbo and Nok-3 was carried out by TGAC. Sequences for Cvi-0 and Ag-0 were retrieved by Martin Trick from the sites below,

ftp://ftp.arabidopsis.org/home/tair/Sequences/Polymorphism_datasets/- and http://1001genomes.org/data/Salk/releases/current/TAIR10/strains, respectively.

Sequences of major genes were aligned with cDNA of the genes and only exons were extracted.

QTL close to the *FAD2* locus were identified across four traits in the KB population and three traits in the SG population and for a single trait in the CA population (Table 4-5a, b, c and e). Among them the most significant QTL was found in the KB population for 18:2/18:1 with a LOD score of 13.13. A region with no consensus sequence was identified in the Kondara *FAD2* sequence (from 1329 to 1396 bp) and also in Br-0 (from 1387 to 1410) both of which are the parental lines of the KB population. Sorbo, the maternal parental line of SG population had two SNPs, one of which was at the same position as Kondara and the second was unique. The SNP in common with Kondara resulted in one amino acid change in protein sequence. In the CA population a minor QTL was present close to *FAD2* and Cvi-0, the maternal parent, had two SNPs in *FAD2*, however the amino acid sequence was unchanged. Nok-3, the maternal parental line of NG population had two SNPs at 913bp and 1328 bp and the latter was shared with Kondara. Only the SNP at 913 resulted in an amino acid change. Ga-0, the paternal line of NG population had one SNP at 1025bp and caused one amino acid change. However, no QTL was identified in the NG population corresponding to the *FAD2* locus

The CA population was the only population with a QTL close to the *FAD3* locus. Consensus sequences for 1807 to 1958 bp, with 166bp of exon left un-aligned, could only be generated from three accessions,. Therefore, only where sequences were successfully aligned with cDNA was discussed. Sequence variations were observed only in Cvi-0, and four SNPs were identified which resulted in three amino acid changes.

QTL close to the *FAE1* locus were identified for four traits in the SG population and one trait in the KB population (Table 4-5-a, b, c and d) of which the QTL for 18:1 in the SG population had the highest LOD at 17.62 accounting 60.3 % of the phenotypic variation. Gy-0, the paternal parental line of the SG population had three SNPs. However Nok-3, the maternal parental line of the NG population, also had SNPs at the same positions meaning they shared the same allele but no QTL corresponding to *FAE1* were identified in the NG population. Only one SNP was identified as common to Gy-0 and Nok-3 this resulted in one amino acid change at the end of the protein sequence. A QTL identified in the KB population had a LOD

score of 9.2 and accounted for 37.4% of the phenotypic variation, however no sequence variation was identified between Kondara and Br-0. By comparison, Ts-5 was shown to have two SNPs however no QTL corresponding to the *FAE1* locus was mapped in the TJ population.

4.3.6 EVALUATION OF PHENOTYPE ASSOCIATED WITH QTL CLOSE TO MAJOR GENES

QTL mapping process generated average phenotypic values at each locus for each parental allele. From this the positive or negative effect of each parental line could be ascertained ie. the one with higher average value is positive and one with a lower value is negative. To assess if these agreed with the expectations based on gene function, values at QTL close to major genes across the traits were examined between two parental lines. Values given in brackets are the proportion of the FA concerned amongst all the other FAs in the seed lipid.

For *FAD2*, the KB, SG and CA populations all had QTL close to the *FAD2* locus and were examined for 18:1 and 18:2. At the genetic marker nga162, Kondara negatively (15.41%) and Br-0 positively (17.16%) controlled 18:1 content, but acted inversely for 18:2, where Kondara was the positive parent (27.63%) and Br-0 was the negative (24.76%). Both Sorbo and Cvi-0 showed the same pattern as Kondara in their populations.

For *FAD3*, the CA population was examined for 18:2 and 18:3. At the genetic marker t32f646516, Cvi-0 was the positive parent (30.08%) and Ag-0 was the negative parent (26.25%) for 18:2, while the opposite was the case for 18.3 with values of 17.70% and 22.30%, respectively.

FAE1 is involved in very long chain FA synthesis, therefore 18:1 and mono- or poly unsaturated very long chain FAs (those with more than 20 carbon atoms) were examined. The phenotypic data for 20:1, 20:2 and 22:1 contents were available, therefore QTL analysis was performed although these traits were not studied as part of this project. The KB and SG population were examined at nga1139 and MS_At4_16.2, respectively. For 18:1, Kondara was the positive parent (17.65%) and Br-0 was the negative parent (15.28%). For 20:1, 20:2 and 22:1, the average phenotypic values of Kondara were 17.94%, 1.77% and 1.37% and those of Br-0 were 22.02%, 2.18% and 1.59%. Kondara and Sorbo shared the same pattern; they positively affected 18:1 and negatively affected all very long FA.

4.4 CONCLUSION

4.4.1 PHENOTYPIC DATA

G x block interactions were observed in the KB population for 18:1, 18:3 and 18:3/18:2, and the NG population for 18:3, 18:3/18:2 and oil content, therefore these traits were studied in other populations. Exception was made for 18:3/18:2 in NG population as the NG population was also studied for 18:2/18:1, therefore FA composition required for the analysis of 18:3/18:2 was available.

Non normal distributions of the trait data were identified by statistical analysis. To normalise the data, outliers were removed, otherwise $\log_{10}$ of the data values were taken. The removal of outliers improved the normality of data dramatically for example, from $P<0.001$ to $P=0.176$ for oil contents in the SG population. However there were lines which were difficult to normalise and these were taken into account when selecting the target QTL. Those populations that could not be normalised were: the CA population for 18:2, 18:3 and 18:3/18:2, the SG population for 18:1, the NG population for oil, the KB population for 18:3

Of them, the CA QTL for 18:2, 18:3 and 18:3/18:2, and the SG QTL for 18:1 showed bimodal distributions which were difficult to normalise, however interval mapping was carried out using the original data. The results revealed very significant QTL (LOD scores ranged from 17 to 32) which acted as a major single gene rather than a QTL, therefore the populations were segregated into two phenotypic patterns. The heritabilities of these traits were also high ranging between 89% and 67%.

4.4.2 IDENTIFICATION AND EXAMINATION OF QTL CLOSE TO MAJOR GENES

The aim of this project was to identify uncharacterised loci, therefore it was important to distinguish newly discovered QTL from QTL caused by known major genes.

After the interval mapping, QTL lying close to the positions of major genes were investigated using MQM mapping. In some populations, the effect of these QTL were outstandingly strong, therefore several adjacent markers also had high LOD scores and they could look like part of a single, large and long expanse of a QTL.

However they could also be another QTL. To investigate this, MQM mapping was carried out to distinguish the effect of loci close to the major genes from other loci. The result was that the large LOD scores of neighbouring markers dropped below the significance level. Therefore there were no other QTL within the span of large effect QTL.

There is also the possibility that the causative basis of these large effect QTL and major gene could be different genes because there were still approximately 0.2 to 2 Mb distance between them and the marker used as cofactors. In this 50 to 500 genes could be present. However considering the fact that these QTL were found across populations and some of them had exceptionally high LOD scores, there were high possibilities that they are caused by major genes. To investigate this further, phenotypic data generated by QTL analysis was evaluated if these QTL caused phenotypic effect in agreement with the expected gene function of these major genes. If a phenotype was affected by *FAD2* the effect of a parent for these two traits should antagonistic, because *FAD2* desaturates 18:1 to 18:2, therefore it reduces the amount of 18:1 and increases that of 18:2. The KB, SG and CA populations were examined and their average phenotypes at the locus which scored the highest LOD score within the QTL showed expected pattern, if a allele was a positive allele for 18:1, it was a negative allele for 18:2 and vice versa. Therefore the effect of QTL close to *FAD2* followed that expected on the gene function. *FAD3* desaturates 18:2 to 18:3, therefore the phenotype should be antagonistic as *FAD2*. The CA population was examined for 18:2 and 18:3, the results showed that Cvi-0 was a positive parent for 18:2 and negative parent for 18:3, while Ag-0 was the opposite which agreed the expected function of *FAD3*.

FAE1 is involved in very long chain FA synthesis, therefore the antagonistic relation could be seen between 18:1 and mono- or poly unsaturated very long chain FAs. Both the KB and SG populations were examined and strong antagonistic relations were seen between 18:1 and very long FAs. Therefore QTL close to *FAE1* locus had effects agreeing with the expected gene function of *FAE1*.

Although QTL close to the major gene loci were excluded for the further study, they provided an opportunity to evaluate the usefulness of multiple populations for QTL analysis.

Out of the six RI populations 19 QTL were closely located to the loci of major genes. A single QTL was found on SG LG2 that overlapped the *FAB2* locus. QTL close to

the *FAE1* locus were detected in the SG population for four traits on LG4: 18:1, 18:2, 18:3 and 18:2/18:1. QTL associated with the *FAD2* locus were detected in the KB population for: 18:1, 18:2, 18:2/18:1 and 18:3:18:2; in the SG population for 18:2, 18:2/18:3 and 18:3/18:2; and once in CA population for 18:2/18:1. Three QTL close to the *FAD3* locus were found in the CA population for 18:2, 18:3 and 18:3/18:2.

As a result, QTL close to the loci of all four anticipated major genes were detected at least in one of the populations. This suggests that use of multiple populations could enable the elucidation of the main structure of the genetic architecture of the PUFA desaturation pathway.

The absence of QTL corresponding to major gene loci in some populations facilitated the discovery of other QTL for the associated trait, however the reasons why they didn't appear in all population remained unresolved. Therefore sequences of parental lines at the major gene positions were compared to check if the sequence variation between the two parental lines could result in QTL near the major genes loci.

Regarding *FAD2*, very significant QTL close to the *FAD2* locus were identified in four traits in the KB population and there were large segments which did not have consensus sequence for *FAD2* in Kondara and Br-0, the parental lines of the KB population. These non-consensus readings occurred when the quality of sequence reads was judged as insufficient or when there were too many variations from the reference Colombia sequence. For Kondara it was the latter case; there were many variations between Columbia, and for Br-0 it was the former case; the number of reads for the region was four, much fewer than other parts, therefore the total reads could be regarded as low quality, despite the quality of each read being as good as other parts. Hence the sequences were retrieved manually and they were the same as Columbia. Therefore there was a large sequence difference between the parental lines in the KB population and this might enable the detection of the QTL in this region. The SG and CA populations had QTL corresponding to the *FAD2* locus, whilst the NG population did not. However, the NG parents, Nok-3 and Ga-0 had *FAD2* sequence variation that resulted in amino acid changes. This raises a question as to the association between sequence variation and QTL detection. To dissect the problem, the region directly 1Kb upstream of the gene which was expected to cover a promoter region were observed, however the results were the same, the NG population had more variation than the SG and CA populations. The other possible explanation could be that the amino acid change in the NG population did not alter the protein structure, hence the function, or it required interaction with

other genes. More detailed study is needed to dissect this problem. However except the NG population, the KB, SG and CA had more sequence variation than the other populations and they were the ones which detected QTL associated to *FAD2*, therefore it could be said for these that the sequence variation could affect the detection of QTL at the locus.

For *FAD3*, only those sequence fragments which were successfully aligned with cDNA of *FAD3* were examined. Sequence variations were found only in Cvi-0 (four SNPs which resulted in three amino acid changes) and QTL close to the *FAD3* locus were also identified only in the CA population where Cvi-0 is the maternal parent. Therefore it could be that a unique allele of Cvi-0 at the *FAD3* locus enabled the detection of the QTL near the *FAD3* locus.

Analysis of *FAE1* sequences of parental lines revealed that Gy-0 (the paternal line of the SG population) and Nok-3 (the maternal line of the NG population) could share the same allele having three identical SNPs, whilst Ts-5 had unique SNPs. However QTL near the *FAE1* locus were identified only in the SG and KB populations. The fact that there were no sequence variations between Kondara and Br-0, and the fact that QTL near *FAE1* locus were not identified in the NG population were contradictory with the result above. Therefore the sequence 1Kb upstream of the gene were also observed, however the results did not resolve the apparent contradiction, showing more sequence variation in the Nok-3 – Ga-0 combination than the other parental pairs. There is the possibility that expression of *FAE1* may require epistasis or other mechanism, however further study is needed to understand this phenomenon.

4.4.3 SELECTION OF TARGET QTL

A number of QTL detected were in common for multiple traits across multiple populations. This was to be expected as these traits are connected to each other in the desaturation pathway. For example QTL found in the same position at the bottom of LG3 in the SG population were for 18:2, 18:3 and 18:3/18:2. Similarly QTL on LG1 in the WC population were detected for 18:1 and 18:3. Oil content QTL mapped at similar positions in the CA and WC populations. As these QTL were identified in more than one population there was a greater level of confidence in them than other QTL which were found in a single population.

The selection of target QTL was carried out as described below and a summary of the results is given in Table 4-6.

A QTL (LOD=4.8) for 18:2 in the TJ population on LG3 was selected, which also appeared for 18:3/18:2 (LOD=3.53) in this population.

Two QTL were selected for the trait 18:2/18:1, one in the NG population on LG5 and another in the TJ population on the LG5 which showed the LOD score (=4.81) at 66.5cM. The NG QTL was one of the largest QTL detected across the six populations of those not close to major gene loci. The TJ QTL was the strongest among the QTL found in TJ population. These two QTL were also observed in 18:1.

For 18:3, a QTL detected in the TJ population on LG1 and a second in the SG population on LG5 were selected. QTL for all five traits of interest were mapped in the TJ population, the QTL for18:3 had the highest LOD score. The SG QTL was chosen because it had a LOD score greater than the QTL for the other traits related with 18:3.

For 18:3/18:2, three QTL were selected. A strong QTL was found in the SG population on LG3 with LOD score of 6.11. QTL were also detected at the same position for 18:2 and 18:3. Two small QTL were selected from the NG population on LG1 and LG5 despite observed G x block interactions (P=0.037) because the QTL on LG1 was also detected in the WC population and the one on LG5 closely mapped to a QTL for 18:2/18:1 also in NG which had a high LOD score of 11.82. Two QTL detected in the TJ population for 18:3/18:2 were more significant than those selected from the NG population, but the QTL on LG1 (LOD= 5.74) was used for 18:3 (LOD=6.32) and the second on LG3 (LOD=3.53) was studied for 18:2 (LOD=4.8) as they showed higher LOD scores for these traits.

Five QTL were chosen for oil content as it is a very important agronomic trait. They were: a large QTL in the KB population at the end of LG2; two on LG5 one from the KB population and the second from the TJ population at MS_At5_22.8 which had a LOD score of 2.40, as it was below the significance threshold it isn't listed on the Table 4-5-f in the Results section. Both of these QTL on LG5 were chosen as they were close to a QTL from unpublished studies by Jitao Zou at the Plant Biotechnology Institute in Canada, the information was kindly provided. The one at the end of LG2 in the KB population coincident with a QTL identified by Hobbs (Hobbs et al., 2004). The final two QTL were chosen from the TJ population, one on LG1 near the end and the second on LG2.

Using the markers' physical coordinates, all markers from all populations were integrated to generate a single map as shown in Fig.4-2.

Table 4-6 List of target QTL

Trait	Population	Linkage group	Locus	LOD	* % Expl.	Position (cM)
18:2	TJ	LG3	MS_At3_6.1	4.34	19.3	15.41
			msd2129380	4.8	20.9	20.50
18:2/18:1	NG	LG5	nga225	7.64	31.5	0.00
			mhf15ind52-52	6.88	28.6	3.52
			ciw14	6.73	28.1	4.96
			f14f1847646	11.82	44	9.05
			nga151	11.17	42.2	11.98
			nga139	4.35	19.6	33.96
	TJ	LG5	f20d23ind18-18b	4.5	19.8	65.03
			nga129	4.81	21	66.50
			MS_At5_20.4	4.72	20.6	66.55
			MS_At5_21.3	2.62	12	69.76
			jv65-66	2.56	11.8	75.24
18:3	SG	LG5	aths0191	3.08	14.2	42.16
			MS_At5_15.8	2.81	12.9	45.36
	TJ	LG1	f9h16ind26-26	2.81	12.9	18.14
			MS_At1_8.2	4.01	18.1	23.64
			nga392	5.42	23.3	31.39
			aths0392	6.32	26.6	35.62
			MS_At1_11.2	5.66	24.3	36.64
18:3/18:2	NG	LG1	MS_At1_22.6	2.97	13.6	71.46
		LG5	nga139	3	13.9	33.96
	SG	LG3	t13j1016610	4.03	17.9	50.50
			k11j14ind16-16	3.79	17.1	51.04
			MS_At3_16.9	6.1	25.8	58.13
			MS_At3_17.3	6.11	25.9	58.30
			t16k521877	5.1	22.4	60.94
			f15b8ind11-11	3.86	17.5	67.97
Oil	KB	LG2	MS_At2_14.2	2.65	12.6	41.90
			t6a23ind10-10	4.01	18.8	47.23
			nga168	4.01	18.8	47.26
			MS_At2_17.5	4.32	19.7	50.40
			athubique	6.01	26.2	53.88
			90j19t7	6.46	28.2	57.09
		LG5	jv65-66	2.74	13	56.68
	TJ	LG1	f5a1859436	2.61	12.1	83.25
			nga111	3.52	15.9	87.80
		LG2	MS_At2_9.3	3.45	15.6	28.57
			pls8	3.9	17.4	30.63
			nga1126	3.84	17.2	35.28
		LG5	MS_At5_22.8	2.4	11.1	77.52

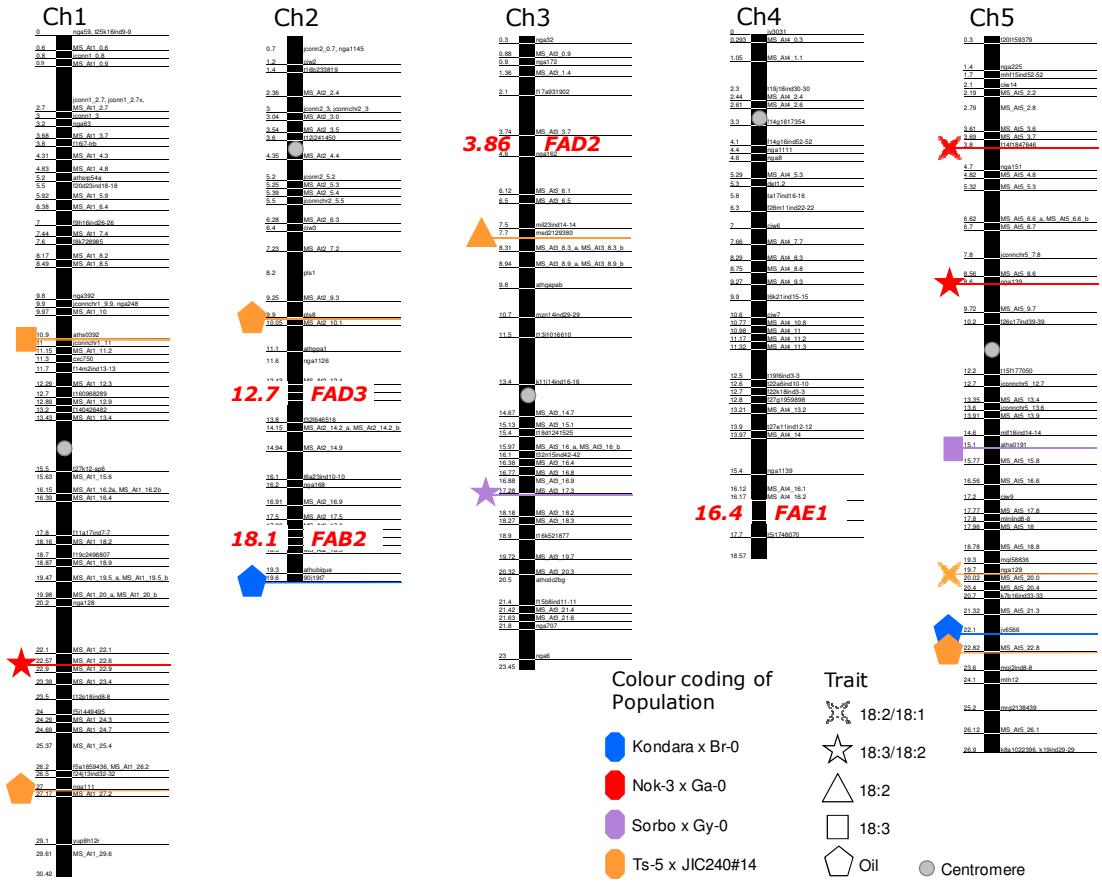


Fig.4-2 Target QTL and major lipid genes physically positioned (Mb)

CHAPTER 5. VALIDATION OF QTL BY NEAR ISOGENIC LINES

5.1 INTRODUCTION

The novel QTL discovered in Chapter 4 were expected to show relatively small effects on the phenotype due to lower LOD scores of between 2.4 and 11.8 compared to the effect of well known major genes such as *FAD3* which had a LOD score of 32 in the CA population as described in Section 4-3-1.

To dissect these relatively minor QTL, more specific plant materials were required as other QTL of the same trait could mask the effect of these QTL or cause an interaction between them (Keurentjes et al., 2007). To prevent these risks, the decision was made to produce near isogenic lines (NIL) which contain only isolated regions close to putative QTL and exclude other QTL for the same trait from the genome.

The procedure for developing NIL varies between researchers due to different types of starting materials and methodology. According to the work carried out by Alonso-Blanco, he developed *Arabidopsis thaliana* NILs both from RIL for flowering time and seed dormancy studies. For the flowering time research three backcrosses were applied to generate starting material followed by one backcross and one selfing to select the final NILs, while for the seed dormancy study the approach taken used two backcrosses followed by two selfings (Alonso-Blanco et al., 2003). The NILs produced in 1998 by Alonso-Blanco were widely exploited by researchers (Swarup et al., 1999; Bentsink et al., 2003; Edwards et al., 2005; Juenger et al., 2005) and has shown that developed NILs achieved their objectives which underlines the relevancy of the NIL development method.

For this project, RILs were used as the starting plant material for NIL development from populations in which putative QTL had previously been identified. Selected RILs were backcrossed twice and selfed once. This method was one generation (selfing) less than Alonso-Blanco's method in 2003. By examining the results of genomewide genotyping and phenotyping, the validity of this method will be assessed. Additional backcrossing or selfing were to be carried out if necessary.

The direction of crossing between parents is important because seed lipids are stored in the embryo whose nuclear phase is 3n, where 2n is derived from the maternal genome, and 1n from the paternal genome, hence there could be a biased

expression of the phenotype towards the maternal genome. According to the study of cytoplasmic effects on oil content conducted by Hobbs in 2004 using reciprocal crosses between Cvi-0 and L*er*, the effect on oil synthesis was small, falling within the range of the normal distribution, however 13 lines of the 15 lines which showed exceptionally low oil contents had Cvi-0 cytoplasm which has a lower oil content than L*er* (Hobbs et al., 2004).

Therefore the maternal parents of the RI populations were used as maternal parents throughout the NIL development to avoid potential effects due to cytoplasmic differences.

In total, 14 NILs were developed by the introgression of seed lipid characteristics of paternal origin into maternal backgrounds through two backcrossings and a selfing. Genotyping was conducted throughout the development for testing and selection. Because the phenotyping of FA compositions is labour intensive and costly it was carried out only once, at the final NIL selection stage as seen in Fig.5-1.

The aims of NIL development were first to provide materials for fine mapping, however it was also an opportunity to examine the robustness of identified QTL. Also this study could be used to test the methodology which was one step shorter than the previous study (Alonso-Blanco et al., 2003) for NIL development.

The details of each process are described in the following sections.

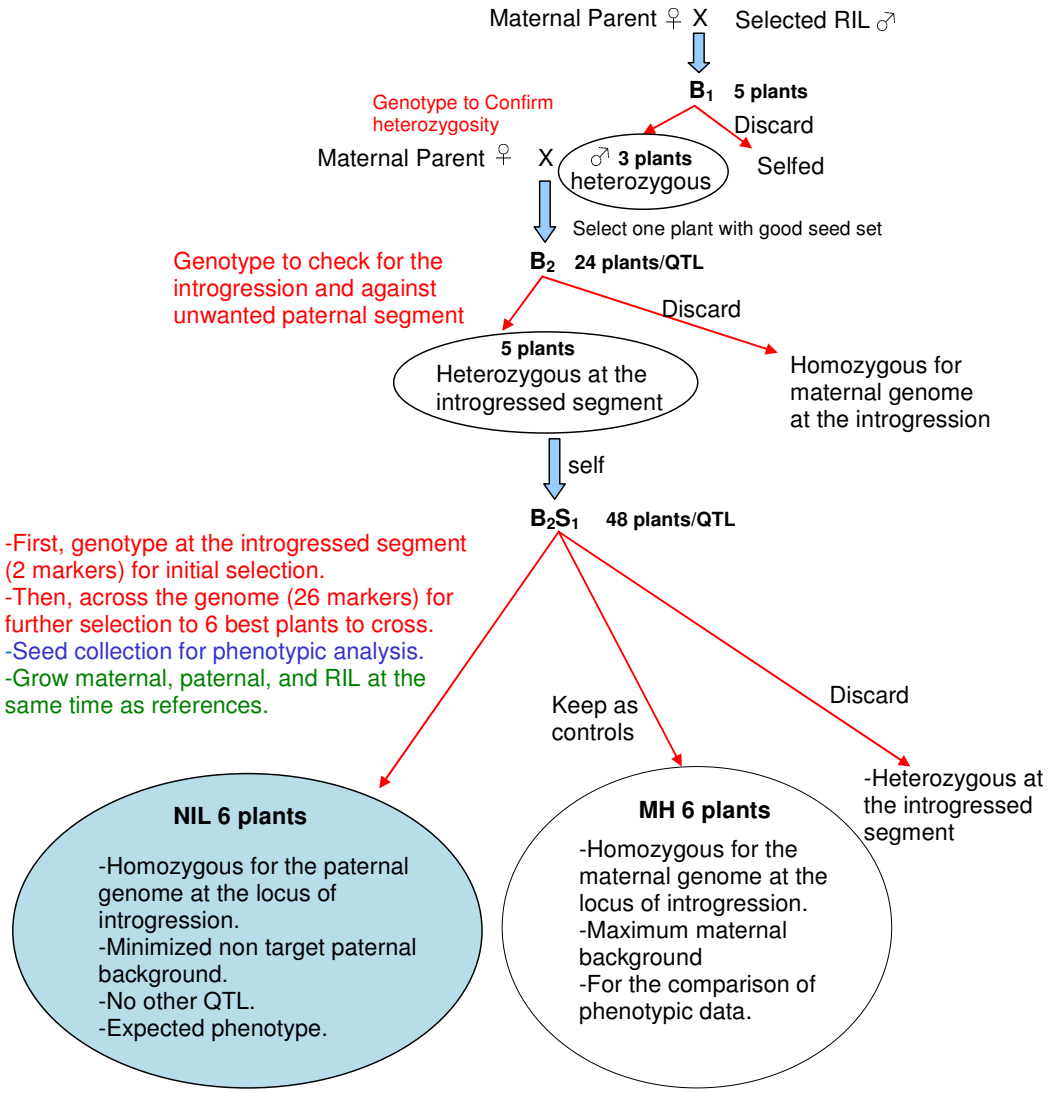


Fig.5-1 Schematic outlining NIL development

MH: maternal homozygous, **B**: backcross, **S**: selfing, $\mathbf{B_2S_1}$: backcrossed twice and selfing once

5.2 MATERIALS AND METHODS

5.2.1 SELECTION OF RIL FOR NIL DEVELOPMENT

The key features that selected RIL should have were cleaner backgrounds of predominantly maternal origin, strong phenotypes, appropriate spans of paternal alleles at the regions of interest and no paternal alleles at the other QTL for the same trait within the genome. The details of criteria and methods are listed below. Charts used for RIL selection of NG32 are shown in Fig 5-2 as an example. Charts used for all other populations are in Appendices 3 and 4.

Selection criteria of the best RIL material for NIL development

1. A line which did not contain other QTL for that trait.
2. A line which had less paternal alleles in the background.
3. A line with the desired phenotype

   -The position in the population was assessed by looking at the phenotypes across the population displayed in a histogram so as to avoid the outliers.

   -The line showing a stronger phenotype was selected. However, due to removal of other QTL within the genome, the candidate RIL could show moderate phenotype compared to those with multiple QTL.

   -The phenotypes of all three replicates from previous experiments were checked for reproducibility of the phenotype.

4. The line with adequate length of introgressed segment

   -The span of the introgressed segment should consist of only paternal alleles (B allele).

   -The ideal span between two markers is 2 LOD score less than the highest scored marker. The line most closely matching to this requirement was chosen.

;Nok-3 x Ga-0 RI F8 population, 94 individuals, genotype file,
;A = Nok-3 B = Ga-0 U = unclassified & heterozygotes
B is increasing allele
18:2/18:1

rep1	2.02	2.04	2.19	2.39	1.63	*	2.13	2.18	2.18	2.39
rep2	1.91	1.90	2.08	*	1.78	1.81	1.66	1.85	2.04	*
rep3	2.00	2.06	2.14	2.34	1.85	1.87	2.21	1.97	2.08	2.21
ave.	**1.97**	**2.00**	**2.13**	**2.37**	**1.75**	**1.84**	**1.98**	**1.99**	**2.09**	**2.30**
no. of A	55	54	54	51	50	50	49	48	48	47

markers	LOD	LG	9	36	59	32	6	72	29	17	41	13
nga59	0.1	1	A	A	A	B	A	A	B	A	A	A
MS_At1_0.6	0.4	1	B	B	A	B	A	A	B	A	A	A
MS_At1_2.7	0.2	1	B	B	A	B	A	B	B	A	A	A
nga63	0.0	1	B	A	A	B	A	A	B	A	A	A
f16j7-trb	0.0	1	U	A	A	B	A	B	B	U	A	A
athsrp54a	0.1	1	U	A	A	A	A	B	B	A	A	A
MS_At1_6.4	0.0	1	B	A	A	A	A	B	A	A	A	A
f9h16ind26-26	0.0	1	U	A	A	A	A	A	B	U	A	A
MS_At1_7.4	0.1	1	A	A	A	A	A	A	B	A	A	B
MS_At1_8.2	0.7	1	A	A	A	A	A	A	B	B	A	A
nga392	0.4	1	A	A	A	A	A	A	B	B	A	A
ath s0392	0.5	1	A	A	A	A	A	A	B	B	A	A
MS_At1_11.2	1.5	1	A	B	A	A	A	A	B	B	A	A
f140426482	1.6	1	A	A	A	A	A	A	A	A	A	B
t27k12-sp6	2.1	1	A	A	A	A	A	A	A	A	A	B
~~MS_At1_16.4~~		1	A	A	A	A	A	A	A	A	A	U
MS_At1_20_a	0.1	1	A	A	A	A	A	A	A	A	A	A
nga128	0.4	1	A	B	A	A	A	A	A	A	A	A
MS_At1_22.6	0.5	1	A	A	A	A	A	A	A	A	A	B
t12p18ind8-8	0.0	1	U	A	A	A	A	A	A	A	A	B
f5i1449495	0.1	1	A	B	A	A	A	A	A	A	A	B
f5a1859436	0.1	1	A	B	A	A	A	A	A	A	A	B
nga111	0.1	1	A	B	A	A	U	A	A	A	A	B
yup8h12r	0.3	1	A	B	A	A	A	A	A	A	A	B
nga1145	0.0	2	A	A	B	B	B	A	A	U	B	B
ciw2	0.0	2	A	A	U	B	B	A	A	B	B	B
t12j241450	0.1	2	A	A	A	B	B	A	A	U	B	B
ciw3	0.0	2	A	A	A	B	B	A	A	B	B	A
pls1	0.2	2	A	A	A	B	B	A	A	U	A	A
pls8	0.3	2	A	A	A	A	A	A	B	U	B	A
athgpa1	0.0	2	A	A	A	A	A	A	A	B	A	A
nga1126	0.0	2	A	A	A	A	A	A	A	B	A	A
t32f646516	0.2	2	A	A	B	A	A	A	A	A	B	A
MS_At2_14.2	0.2	2	A	A	B	A	A	A	A	A	B	A
nga168	0.4	2	A	A	U	A	A	A	B	A	B	A
90j19l7	0.1	2	B	B	U	A	A	B	B	A	B	A
nga32	1.3	3	A	A	A	A	A	B	A	A	A	B
f17a931902	0.8	3	A	A	A	A	A	B	B	A	A	B
nga162	0.4	3	A	B	A	A	A	B	B	A	A	B
msd2129380	0.9	3	A	A	A	U	U	B	A	A	A	A
MS_At3_8.3	0.8	3	A	A	A	A	A	A	A	A	A	A
athgapab	1.5	3	A	B	U	A	B	A	A	A	A	A
t13j1016610	1.4	3	A	B	A	A	U	A	A	A	A	A
k11j14ind16-16	1.3	3	A	A	A	A	B	A	A	A	A	A
MS_At3_18.3	1.9	3	U	A	A	A	B	A	A	A	A	A
t16k521877	2.1	3	A	B	A	A	B	A	A	A	A	A
athcdc2bg	1.9	3	A	A	A	A	B	A	A	A	A	U
nga6	1.6	3	A	A	A	A	B	A	A	A	B	A
jv30/31	0.1	4	B	A	U	A	B	B	B	B	B	B
MS_At4_0.3	0.1	4	B	A	U	A	B	B	B	B	B	B
t19j18ind30-30	0.0	4	U	A	B	A	B	B	B	U	B	B
nga8	0.1	4	A	A	A	A	A	B	B	B	A	B
MS_At4_5.3	0.2	4	B	A	A	A	A	B	B	A	A	B
ciw6	1.5	4	A	A	A	B	A	B	A	B	A	A
MS_At4_10.8	0.9	4	A	A	U	A	A	B	A	B	A	A
t22a6ind10-10	1.4	4	A	B	A	A	A	B	A	B	B	A
f27g1959898	1.4	4	A	A	U	A	A	B	A	B	B	A
~~t27e11ind12-12~~		4	A	B	A	A	A	B	A	B	B	A
nga1139	1.9	4	A	A	A	A	A	B	A	U	A	A
nga1107	1.3	4	A	A	A	B	A	A	A	A	A	A
t5j1748070	0.6	4	A	A	A	B	A	A	B	A	A	A
nga225	7.6	5	B	B	U	B	B	B	A	B	B	B
mhl15ind52-52	6.9	5	B	B	B	B	B	B	A	B	B	B
ciw14	6.7	5	B	B	B	B	B	B	A	B	B	B
f14f1847646	11.8	5	B	B	B	B	B	B	B	B	B	B
nga151	11.2	5	B	B	B	B	B	B	B	B	B	B
nga139	4.4	5	A	B	U	B	B	B	B	A	B	B
MS_At5_9.7	0.3	5	A	A	U	B	A	A	A	A	B	B
MS_At5_13.9	0.1	5	A	U	B	B	A	U	A	A	B	A
mlf18ind14-14	0.0	5	A	A	B	B	A	A	A	A	B	A
ath s0191	0.0	5	A	B	B	A	A	A	A	A	B	U
MS_At5_17.8	0.6	5	A	A	A	A	A	A	A	A	B	A
mq158836	0.6	5	A	A	A	U	U	A	A	A	B	A
MS_At5_20.4	0.6	5	A	A	A	A	B	A	A	A	B	A
jv65-66	0.5	5	U	A	A	A	B	A	B	A	A	A
mrg2138439	0.4	5	B	A	A	B	B	A	B	A	A	A
k8a1022396	0.3	5	B	A	U	B	A	A	A	A	A	A

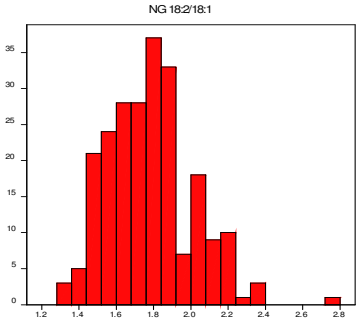


Fig.5-2 Example of charts for RIL selection for NIL development

The target putative QTL was on the top of LG5, highlighted in a gradation of magenta. The unwanted QTL were in the middle of LG1 and the end of LG3 with the marker highlighted in pale blue. RIL NG32 met the genotypic criteria. The phenotypic value of 2.37 for 18:2/18:1 was one of the strongest within distribution of the NG population and was not an outlier.

Charts for all target lines are in Appendices 3 and 4

5.2.2 FIRST BACKCROSSING (B_1 DEVELOPMENT)

The ideal NIL should have homozygous paternal alleles at the introgressed segment and the rest of the genome should be maternal homozygous. Selected RILs might still contain paternal alleles in their backgrounds. Therefore the first step was to replace the paternal alleles in the background with maternal alleles, to achieve this RILs were backcrossed with their respective maternal parents. As a result, alleles within the introgressed segment became heterozygous and the rest of the genome was a mixture of homozygous and heterozygous maternal alleles.

This step was followed by successive backcrossing for further incorporation of maternal alleles into the genome background as explained in the next Section 5.2.3.

Method

Five plants each of the selected RILs, ten plants for maternal lines and two plants for paternal lines were grown following the general plant husbandry described in Section 3.2. The RILs and maternal parental lines were used in crossing and paternal lines were used as reference material when scoring the genotyping data.

Leaf samples were taken, following the procedure outlined in Section 3.5 from RILs and maternal plants used in crossing and paternal plants, for DNA extraction so as to check their genotype later.

RILs were crossed with their maternal lines as soon as the crossing pairs started flowering, using the methods described in Section 3.3.Siliques derived from successful crosses were collected in eppendorf tubes as described in Section 3.4.1.

Crossing was repeated until plants set sufficient seeds to grow more than five plants.

5.2.3 SECOND BACKCROSSING (B_2 DEVELOPMENT)

In this section, the seeds produced from the backcrossing in the previous section were grown to maturity and then backcrossed again to replace unwanted heterozygous paternal alleles in the background with homozygous maternal alleles, while keeping the introgressed segment heterozygous.

Method

Five plants derived from B_1 seeds and ten maternal parents were grown as described in Section 3.2 and leaf samples of those B_1 plants and maternal plants used for crossing were taken following the methods described in Section 3.5. These samples were submitted for DNA extraction together with the leaf samples collected from the original RILs, maternal parents used for crossing and the paternal parents from the previous stage.

All five B_1 plants were backcrossed with their respective maternal parents following the method described in Section 3.3 to generate a minimum of 30 seeds although the results of heterozygosity test were not available at that point.

Genotyping was performed following the procedure described in Section 3.6, to examine the heterozygosity of B_1 plants using the flanking markers of the introgressed segment. Flanking markers are highlighted in green in the genotype file in Appendix 3. The genotypes of parental lines were also examined to confirm they had the correct genotypes.

Seeds from the crosses between confirmed hybrid B_1 plants and maternal plants showing the correct genotypes were taken through to the next stage.

5.2.4 FIRST SELFING (B_2S_1DEVELOPMENT)

Seeds harvested after two backcrossings segregated into maternal homozygous and heterozygous at the loci which had been heterozygous after the first backcrossing, such as the introgressed segment and unwanted paternal alleles still in the background. Only plants which were heterozygous at the introgressed segments and maternal homozygous at the unwanted paternal alleles in the background were kept for selfing.

Method

For each NIL 22 plants from the B_2 seeds were grown and allowed to self pollinate. Three parental lines were grown as DNA reference. Leaf samples were taken from all plants for DNA extraction and genotyping of the two flanking markers was performed. B_2 plants heterozygous at both flanking markers and maternal homozygous at the other QTL were selected while the remainder were discarded. Also the genotypes of the loci of other QTL were checked to confirm they were not paternal.

Plants were bagged with transparent breathable plastic bag to prevent cross pollination with other plants and also to retain the seeds. Seeds were harvested in two ways, the silique collection method and the bulk collection method described in Section 3.4.1 and Section 3.4.2 respectively.

5.2.5 FINAL STAGE OF NIL DEVELOPMENT ($B_2S_1B_1$ DEVELOPMENT)

During the process of backcrossing, the introgressed segments are maintained in the heterozygous state. Through selfing and the subsequent selection the paternal homozygous state is restored. Plants with paternal homozygous at the introgressed segments were NIL candidates and those with maternal homozygous at the introgressed segments were kept as control material. Genomewide genotyping with additional 24 markers was carried out in order to select a final NIL for fine mapping.

At the same time, all the NIL candidates were backcrossed to the maternal parent line for the development of a large scale mapping population for fine mapping at the next step.

Method

Seeds of selfed plants which were heterozygous at the introgressed segments segregated into three genotypes: paternal homozygotes, heterozygotes and maternal homozygotes. 50 plants grown from these seeds were genotyped at the flanking markers and all plants showing paternal homozygous inheritance at both markers were kept as NIL candidates while six plants with the maternal homozygous (MH) alleles at both markers were kept as control material. This process was called “first screening”.

All NIL candidate plants were backcrossed with the maternal parental line after the first screen to produce mapping populations suitable for fine mapping. Siliques bearing backcrossed seeds were collected for the fine mapping and then plants were left to dry off before harvesting the selfed seeds to be used in phenotyping.

Subsequently all plants taken through the first screen were genotyped with 24 markers, these were evenly spread across the whole genome so that six NIL candidates with the least amount of paternal alleles in their backgrounds could be selected. Phenotyping of this material was carried out together with MH plants and parental lines.

This was the most important stage of NIL development in terms of plant husbandry because it was the only occasion when plants were tested for their phenotype. Great care was taken to achieve uniform growing conditions as well as taking account of experimental design to detect any block effects. Experimental design is discussed in the next Section 5.2.5.1.

5.2.5.1 EXPERIMENTAL DESIGN

The experiment was designed in such a way to reduce environmental effect on the phenotype of interest and also to enable separation of the residuals from block and genotype effects through statistical analysis.

The development of the 14 NILs was divided into two occasions, the first group consisted of: NG28 for 18:2/18:1, NG32 for 18:2/18:1 and 18:3/18:2, TJ79 for 18:2/18:1, TJ17 for 18:3, TJ79 for oil content on chromosome (Ch) 2, TJ79 for oil content on Ch5 and TJ83 for oil content. This material was sown in December 2008. The second group had: TJ25 for 18:2, SG93 for 18:3, NG93 for 18:3/18:2, SG10 for 18:3/18:2, KB55 for oil content and KB91 for oil content. These were sown in August 2010.

The experimental design for the first group was to divide lines into three and gather them into three blocks as seen in Fig.5-3-a. Randomization was carried out within the block at the transplanting stage and after the first screen (Fig.5-3-b). The number of blocks was reduced to two after the first selection because the number of plants was decreased to a third and randomization with such a small number of plants is not efficient. Therefore the block two was divided and half was merged into block one and half into block three. Parental lines were also randomized and treated similarly. This design aimed to reduce block effects on the phenotype within a block.

A different approach was taken for the second group because the results from the first group showed that the phenotypic difference between NIL candidates and MH plants in each line was more subtle than anticipated. Since the NIL candidates and MH plants had slightly different genotypes from each other, comparison between blocks did not work as well as if they were biological replicates. Therefore it became more important to measure the environmental effect rather than to reduce it, so that the environmental effect could be taken into account at analysis for selection.

To measure the strength of the environmental effect, those parental lines with four biological replicates were excluded from randomization and scattered within the block systematically as seen in Fig.5-4-a. The number of blocks was also reduced to two from the beginning. After the first screen, plants were relocated to fill the vacated gaps of discarded plants, whilst maintaining their relative positions within a block (Fig.5-4-b).

Entrance

	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
1	2_3	RIL NG32_2	7_12	7_3	1_12	5_1	8_12	2_5	7_14	1_14	4_14	2_10	5_15	8_2	3_3	1
2	3_15	RIL NG32_3	1_16	Ts-5_2	3_2	6_13	6_14	5_7	6_3	5_13	Mz-0_2	7_10	3_5	5_12	6_7	2
3	RIL TJ83_3	RIL TJ17_2	7_16	4_13	8_11	8_14	6_4	4_5	Mz-0_3	3_12	RIL NG28_2	3_8	1_1	6_6	4_3	3
4	8_7	Mz-0_1	4_6	6_16	2_13	2_8	4_4	5_11	8_5	7_6	7_15	3_7	Ts-5_3	Nok-3_2	4_1	4
5	2_14	3_9	7_9	2_1	1_4	6_15	1_6	7_8	5_6	8_13	8_1	1_2	8_4	4_11	4_15	5
6	8_10	2_15	6_8	3_13	6_9	5_4	8_3	5_14	RIL NG28_3	Ga-0_3	RIL NG28_1	5_16	4_10	1_11	Ga-0_2	6
7	7_4	5_10	6_1	3_14	1_13	RIL TJ79_1	1_8	4_8	2_11	1_7	5_5	6_10	4_7	2_7	6_5	7
8	6_2	RIL NG32_1	1_3	RIL TJ17_1	2_4	Ga-0_1	5_8	Nok-3_1	8_16	4_2	6_11	Ts-5_1	1_15	6_12	7_1	8
9	7_7	3_6	7_13	1_9	8_6	5_2	1_5	8_15	RIL TJ79_2	4_12	2_16	3_4	3_16	3_11	7_11	9
10	RIL TJ79_3	3_1	RIL TJ83_2	2_6	5_9	3_10	RIL TJ83_1	2_9	2_2	RIL TJ17_3	7_5	5_3	4_16	2_12	7_2	10
11	8_8	4_9	Nok-3_3	8_9	1_10	Mz-0_4	7_19	Ga-0_5	1_28	2_28	3_20	3_25	1_24	RIL TJ17_5	RIL TJ79_4	11
12	4_25	5_24	Mz-0_6	RIL NG28_6	2_29	5_31	2_17	8_22	7_25	5_22	3_23	8_31	2_32	Ts-5_5	7_29	12
13	5_20	7_23	6_22	3_24	4_18	4_31	6_27	8_23	2_21	8_17	RIL NG32_5	8_19	2_18	5_19	5_25	13
14	2_23	2_31	4_29	6_28	8_18	5_27	4_22	Nok-3_4	5_21	Nok-3_6	4_24	4_27	7_24	8_27	1_19	14
15	1_26	1_29	RIL NG28_5	6_20	5_30	Ts-5_6	2_25	6_26	1_23	4_30	4_21	RIL TJ83_6	3_29	6_21	6_18	15
16	5_18	3_21	6_19	RIL TJ83_5	7_26	Nok-3_5	1_32	5_26	6_32	3_32	2_27	1_18	2_19	8_20	2_20	16
17	7_27	5_17	8_26	3_18	6_29	8_24	7_21	8_32	4_32	8_21	Ts-5_4	Mz-0_5	8_28	RIL NG32_6	5_23	17
18	6_25	7_17	RIL TJ79_6	4_26	RIL TJ17_4	7_28	3_19	6_23	6_17	3_30	7_32	1_30	6_31	1_27	1_22	18
19	2_30	8_29	2_26	RIL TJ83_4	8_30	7_31	1_17	8_25	1_21	1_31	2_24	4_28	3_28	4_20	1_25	19
20	3_26	Ga-0_4	6_30	7_18	3_22	RIL NG32_4	RIL TJ17_6	2_22	7_30	4_17	3_31	4_19	RIL NG28_4	Ga-0_6	7_22	20
21	RIL TJ79_5	5_32	5_29	5_28	3_27	1_20	7_20	6_24	3_17	4_23	4_33	4_45	4_46	7_36	1_47	21
22	RIL TJ79_9	6_42	1_33	4_36	2_41	2_36	RIL NG28_7	5_43	8_35	RIL TJ79_7	2_47	8_36	Mz-0_9	6_43	6_34	22
23	7_37	7_34	RIL NG32_7	8_33	6_35	3_42	1_37	6_36	2_34	4_40	RIL TJ83_8	2_33	Ts-5_8	2_40	5_36	23
24	1_38	6_33	1_43	Ga-0_7	Nok-3_8	RIL TJ17_8	8_44	5_38	3_33	Mz-0_8	RIL NG28_9	7_35	RIL TJ17_9	8_48	3_38	24
25	6_39	RIL NG32_8	6_40	1_42	2_39	Nok-3_9	4_34	1_36	4_35	7_39	3_39	4_48	3_37	4_41	6_37	25
26	1_44	Nok-3_7	5_47	7_46	1_46	2_48	RIL TJ83_7	6_46	1_39	2_45	4_47	7_43	2_46	6_48	8_42	26
27	5_40	6_44	Ts-5_7	5_37	1_34	Mz-0_7	8_41	3_48	3_45	5_44	7_42	Ts-5_9	5_46	RIL TJ83_9	7_47	27
28	3_34	1_35	2_35	7_45	7_41	8_47	3_43	RIL TJ17_7	RIL TJ79_8	7_38	4_37	1_45	3_40	3_47	8_37	28
29	4_44	8_46	3_41	5_35	5_42	8_34	1_41	7_33	7_48	4_39	2_42	7_44	6_38	RIL NG28_8	6_41	29
30	5_39	5_41	7_40	8_38	2_38	8_39	1_48	RIL NG32_9	3_46	5_48	8_40	3_44	2_43	6_45	5_45	30
31	5_33	2_37	4_43	6_47	8_45	8_43	Ga-0_9	4_42	5_34	3_36	2_44	Ga-0_8	3_35	1_40	4_38	31
	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	

Window side

Block1	Block3
Block2	Guard

Fig.5-3-a Glasshouse bench positions of NIL candidates, MH plants and parental lines of the first group prior to the first screening

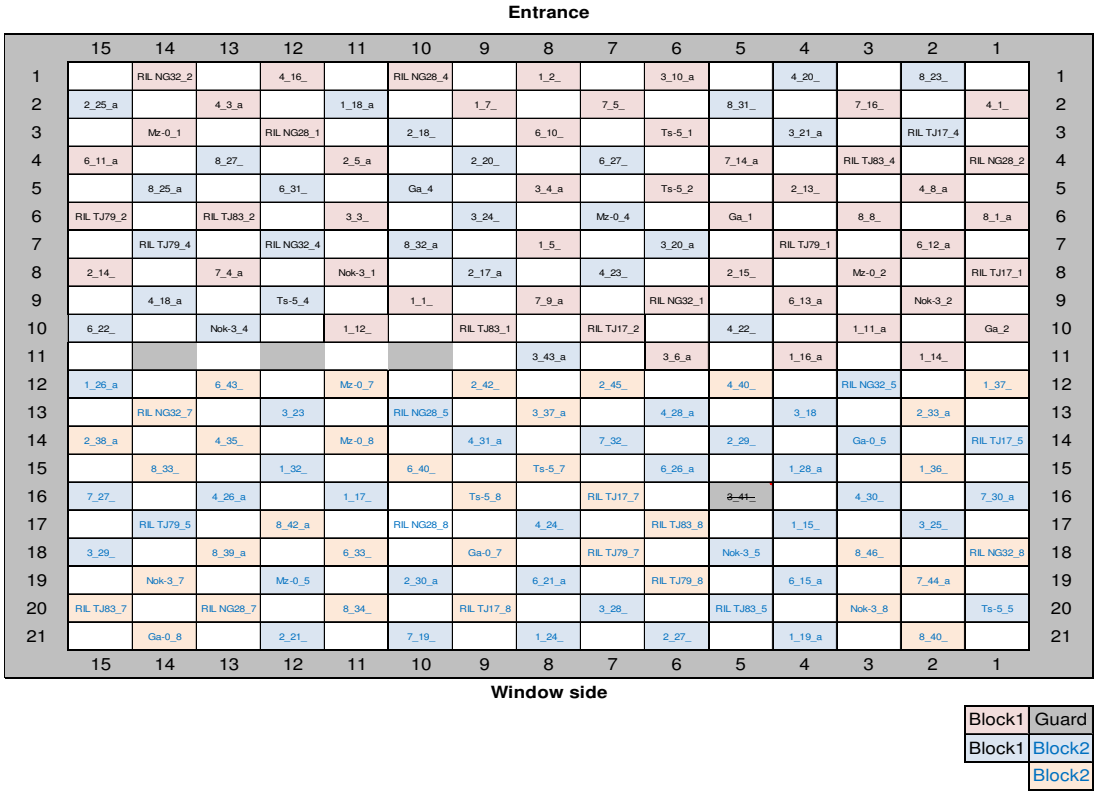


Fig.5-3-b Glasshouse bench positions of NIL candidates, MH plants and parental lines of the first group after the first screening

After the first screen, five plants for each NIL candidate and five MH plants were selected for each of the eight lines. Block 2 was divided into two; numbers 17 to 24 were merged into Block 1, numbers 25 to 32 were merged into Block 3. Plants were then randomized within Blocks 1 and 3.

Entrance

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	13_22	11_14	12_8	16_8	11_6	12_10	12_14	12_3	12_11	16_23	11_17	16_21	12_1	13_2	13_19	1
2	16_13	11_15	14_11	16_22	13_20	12_16	15_20	15_21	16_18	12_7	15_11	14_12	14_3	16_7	14_7	2
3	15_14	Sorbo (1)	Gy-0 (1)	Kondara (1)	Br-0 (1)	Nok-3 (1)	Ga-0 (1)	Ts-5 (1)	Mz-0 (1)	SG10 (1)	SG93 (1)	KB55 (1)	KB91 (1)	NG93 (1)	TM25 (1)	3
4	11_13	12_4	12_18	14_17	12_23	11_16	14_15	12_17	15_3	12_13	16_9	15_8	15_10	14_10	11_21	4
5	11_1	13_10	13_17	15_6	14_9	16_2	12_9	13_8	14_18	15_9	13_7	15_2	12_19	14_19	16_5	5
6	13_15	11_20	16_17	13_18	12_15	13_16	14_16	16_14	15_16	11_18	14_5	14_2	11_5	12_2	11_4	6
7	15_5	11_11	13_23	11_12	13_14	15_22	11_9	16_20	16_3	16_6	13_9	11_10	16_4	14_8	11_2	7
8	14_23	14_1	13_11	13_5	16_1	11_23	15_7	14_4	13_3	13_4	12_20	16_16	14_21	13_12	13_1	8
9	13_6	Sorbo (2)	Gy-0 (2)	Kondara (2)	Br-0 (2)	Nok-3 (2)	Ga-0 (2)	Ts-5 (2)	Mz-0 (2)	SG10 (2)	SG93 (2)	KB55 (2)	KB91 (2)	NG93 (2)	TM25 (2)	9
10	15_15	11_3	16_10	11_8	15_18	15_19	15_13	12_6	16_19	15_4	11_7	14_13	14_14	16_12	15_23	10
11	12_5	15_12	15_1	16_15	14_6	14_22	13_21	11_19	11_22	12_22	14_20	16_11	15_17	12_21	13_13	11
12	12_12	13_36	16_42	12_38	12_37	11_34	16_35	14_38	16_32	13_26	13_44	13_42	12_25	11_37	16_29	12
13	15_38	15_41	11_46	15_39	12_24	15_29	11_30	13_34	12_26	11_25	13_37	15_42	11_26	14_42	11_45	13
14	15_37	14_37	16_26	13_40	11_43	13_46	16_44	14_30	15_32	11_35	12_31	16_31	12_32	15_35	14_33	14
15	16_27	Sorbo (3)	Gy-0 (3)	Kondara (3)	Br-0 (3)	Nok-3 (3)	Ga-0 (3)	Ts-5 (3)	Mz-0 (3)	SG10 (3)	SG93 (3)	KB55 (3)	KB91 (3)	NG93 (3)	TM25 (3)	15
16	12_28	16_37	16_40	14_25	12_30	15_44	12_35	14_26	13_33	16_24	16_28	14_32	11_42	14_43	12_33	16
17	11_41	16_30	12_34	12_29	11_38	14_39	13_41	11_32	11_28	15_34	11_27	16_43	15_24	15_46	15_27	17
18	16_45	12_39	13_31	16_39	12_45	15_36	13_43	14_35	15_40	12_27	15_25	14_41	16_46	16_41	13_35	18
19	11_24	13_32	15_30	15_28	13_30	11_44	12_36	13_38	16_33	11_40	13_25	13_24	16_38	14_27	11_31	19
20	12_46	13_27	16_25	14_29	14_36	14_28	16_36	11_33	11_39	15_26	13_45	12_43	12_41	15_33	12_40	20
21	13_29	Sorbo (4)	Gy-0 (4)	Kondara (4)	Br-0 (4)	Nok-3 (4)	Ga-0 (4)	Ts-5 (4)	Mz-0 (4)	SG10 (4)	SG93 (4)	KB55 (4)	KB91 (4)	NG93 (4)	TM25 (4)	21
22	12_42	14_31	16_34	11_36	15_45	15_31	11_29	14_46	14_40	13_39	12_44	15_43	14_24	14_34	14_44	22
23	13_28	14_45														23
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Window side

Block1
Block2
Guard

Fig.5-4-a Glasshouse bench positions of NIL candidates, MH plants and parental lines of the second group prior to the first screening

Entrance

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	13_22		13_20		15_14		12_10		12_14		16_21		12_1		13_2	1
2		Sorbo (1)		Gy-0 (1)		Nok-3 (1)		Ga-0 (1)		12_7		(16_7)		14_12		2
3	Kondara (1)		Br-0 (1)		12_4		Ts-5 (1)		SG10 (1)		SG93 (1)		KB55 (1)		14_3	3
4		14_17		13_17		Mz-0 (1)		14_15		12_17		16_9		TM25 (1)		4
5	15_5		11_11		16_2		12_9		14_18		11_4		KB91 (1)		NG93 (1)	5
6		Sorbo (2)		13_23		14_16		15_16		15_9		11_10		15_2		6
7	Kondara (2)		Br-0 (2)		Gy-0 (2)		14_4		13_4		14_21		KB91 (2)		TM25 (2)	7
8		16_10		15_15		Nok-3 (2)		Ts-5 (2)		Mz-0 (2)		SG93 (2)		NG93 (2)		8
9	16_15		15_1		14_6		Ga-0 (2)		SG10 (2)		KB55 (2)		14_14		15_23	9
10		15_38		12_12		14_22		16_19		15_4		11_7		14_13		10
11	14_37		16_26		Br-0 (3)		11_34		16_35		14_20		13_42		12_21	11
12		Sorbo (3)		Gy-0 (3)		Kondara (3)		13_26		11_36		11_26		11_45		12
13	16_45		12_34		16_39		15_29		11_30		12_31		KB91 (3)		14_33	13
14		12_39		13_32		12_26		16_44		11_25		KB55 (3)		NG93 (3)		14
15	11_24		14_29		14_36		11_39		Nok-3 (3)		SG93 (3)		12_33		TM25 (3)	15
16		13_29		Sorbo (4)		Ga-0 (3)		Ts-5 (3)		Mz-0 (3)		14_41		11_31		16
17	13_28		Gy-0 (4)		Br-0 (4)		14_26		SG10 (3)		13_25		13_24		14_27	17
18		~~14_45~~		Kondara (4)		13_33		13_41		11_32		NG93 (4)		12_43		18
19	13_39		12_36		11_28		13_38		11_40		SG93 (4)		KB55 (4)		TM25 (4)	19
20		(16_24)		14_28		Ga-0 (4)		15_26		SG10 (4)		KB91 (4)		14_24		20
21					Nok-3 (4)		Mz-0 (4)		Ts-5 (4)							21
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Window side

Block1
Block2
Guard

Fig.5-4-b Glasshouse bench positions of NIL candidates, MH plants and parental lines of the second group after the first screening

After the first screen, six plants for each NIL candidate and six MH plants for each of the six lines were selected. The original locations of plants were maintained as much as possible.

5.2.6 MEASURING FATTY ACID COMPOSITIONS

Fatty acids (FAs) were extracted chemically from seeds and suspended in iso-hexane as fatty acid methyl esters (FAME) ready for GC analysis. The extraction methods varied slightly to optimise the concentration of the sample suitable for the GC machene. NILs derived from NG28 were extracted using the first method described in Appendix 8. The remainder of the lines were extracted using the second method which is also described in Appendix 8.

Four biological replicates were prepared for each NIL candidate plant and each MH plant. Four plants from each parental line were phenotyped and for these two biological replicates were prepared as shown in Table 5-1 as an example.

Table5-1 Example of a FAME analysis sample list

	SG10-1
1	NIL 11(4) rep1
2	NIL 11(7) rep1
3	NIL 11(10) rep1
4	NIL 11(30) rep1
5	NIL 11(36) rep1
6	NIL 11(39) rep1
7	MH 11(11) rep1
8	MH 11(24) rep1
9	MH 11(31) rep1
10	MH 11(32) rep1
11	MH 11(40) rep1
12	MH 11(45) rep1
13	Sorbo (1) rep1
14	Sorbo (2) rep1
15	Gy (1) rep1
16	Gy (2) rep1
17	RIL SG10 (1)rep1
18	RIL SG10 (2) rep1
19	NIL 11(4) rep2
20	NIL 11(7) rep2
21	NIL 11(10) rep2
22	NIL 11(30) rep2
23	NIL 11(36) rep2
24	NIL 11(39) rep2
25	MH 11(11) rep2
26	MH 11(24) rep2
27	MH 11(31) rep2
28	MH 11(32) rep2
29	MH 11(40) rep2
30	MH 11(45) rep2
31	Sorbo (3) rep1
32	Sorbo (4) rep1
33	Gy (3) rep1
34	Gy (4) rep1
35	RIL SG10 (3) rep1
36	RIL SG10 (4) rep1

	SG10-2
1	NIL 11(4) rep3
2	NIL 11(7) rep3
3	NIL 11(10) rep3
4	NIL 11(30) rep3
5	NIL 11(36) rep3
6	NIL 11(39) rep3
7	MH 11(11) rep3
8	MH 11(24) rep3
9	MH 11(31) rep3
10	MH 11(32) rep3
11	MH 11(40) rep3
12	MH 11(45) rep3
13	Sorbo (1) rep2
14	Sorbo (2) rep2
15	Gy (1) rep2
16	Gy (2) rep2
17	RIL SG10 (1) rep2
18	RIL SG10 (2) rep2
19	NIL 11(4) rep4
20	NIL 11(7) rep4
21	NIL 11(10) rep4
22	NIL 11(30) rep4
23	NIL 11(36) rep4
24	NIL 11(39) rep4
25	MH 11(11) rep4
26	MH 11(24) rep4
27	MH 11(31) rep4
28	MH 11(32) rep4
29	MH 11(40) rep4
30	MH 11(45) rep4
31	Sorbo (3) rep2
32	Sorbo (4) rep2
33	Gy (3) rep2
34	Gy (4) rep2
35	RIL SG10 (3) rep2
36	RIL SG10 (4) rep2

5.2.7 MEASURING SEED OIL CONTENT

Oil contents were measured with NMR at JIC using the method described in Section 3.8. No replicate sampling was performed as 200mg of Arabidopsis seeds represents a very large number of seeds.

5.2.8 STATISTICAL ANALYSIS TO TEST THE ROBUSTNESS OF QTL

Statistical analysis of the phenotype data was carried out using Genstat 13.

Block effects and experimental errors were examined to check the quality of phenotypic data by ANOVA.

The phenotypes of NIL candidates and MH plants were compared using ANOVA to investigate the significance of different between the two groups. If the difference was significant ($P<0.05$), the putative QTL above the introgressed segment was confirmed in affecting the phenotype. However, those lines which did not show a significant difference between NIL candidates and MH plants were excluded from further analysis.

5.2.9 CRITERIA FOR SELECTION OF NIL FROM NIL CANDIDATES

Those lines which passed the statistical tests demonstrating the robustness of their QTL were examined further if their phenotypes were sufficiently strong and different from the NIL candidates. Because the selected NIL would be used for fine mapping, the phenotypic difference between NIL candidates and MH plants needed to be more distinctive. To this end the phenotypes of a group of NIL candidates were compared with the group of its associated MH plants. The differences were statistically assessed using standard deviation of each group. If the expected variation (standard deviation) of the NIL exceeded that of MH plants, the differences between the two groups were clearly observed.

The strongest phenotype and least amount of paternal alleles in the background were key features for the final NIL selection.

5.2.10 SELECTION OF QTL FOR FINE MAPPING

One QTL was selected to be taken through to fine mapping from amongst other QTL which also had NILs that passed the criteria described in the previous Section 5.2.9.

The final QTL with its derived NIL was chosen from the first NIL development group which was brought through six months ahead of the second group.

Selection of the final target QTL and its NIL for fine mapping required complex and thorough criteria, because NIL material must maintain the desired phenotype through another two generations ie., one backcross and one selfing, to produce a large scale mapping population.

A primary factor was the phenotype of the line, which must be robust and similar to that of the putative QTL. However, most of the lines studied at this stage displayed similar strength of phenotypes. Therefore the accuracy of phenotypic measurements was very important in comparing these successful NILs. Also the overall fitness of a line was crucial from a practical point of view.

5.3 RESULTS

The results of RIL selection for each target QTL at the beginning of NIL development is shown below in Section 5.3.1. Prior to the comparisons described below, quality of phenotypic data was checked statistically by ANOVA in Chapter 5.3.2. Developed NIL candidates were measured for their phenotypes such as FA composition and oil content and the data were compared with that of MH plants in different depths and perspectives according to the objectives. The first comparison was to examine the effect of QTL on the phenotype statistically as outlined in Section 5.3.3. The second comparison was to test if the QTL was strong enough to be fine mapped and one NIL for each QTL was selected (Section 5.3.4). Finally one QTL and its NIL were selected for fine mapping in Section 5.3.5.

5.3.1 SELECTION OF RILS AS STARTING MATERIALS

At the start of NIL development both RIL phenotypes and genotypes were examined when selecting a RIL as starting material for each of the target QTLs. Those lines with strong phenotypes and ideal genotypes were carefully chosen. Charts used for selection of RILs can be found in Appendices 3 and 4.

As listed in Table 5-2, two lines from the NG population for the ratios 18:2/18:1 and 18:3/18:2 along with five lines from the TJ population for the traits 18:3, 18:2/18:1 and oil content (three lines) were selected for the first NIL development group.

Whilst for the second group two lines were selected from the KB population for oil content, one from the NG population for 18:2/18:1, two lines were chosen from the SG population for 18:3 and 18:3/18:2 and a further line from the TJ population for 18:2. Between both development groups a total of 13 QTL were included for study with 14 planned NILs for development.

Table 5-2 Details of target QTL and selected RILs for NIL development

Trait	Linkage Group	LOD	%Expl*	Population	Selected RIL no.	Maternal parent	Group
18:2	3	4.8	20.9	TJ	25	Ts-5	2
18:2/18:1	5	11.82	44.0	NG	28	Nok-3	1
				NG	32		
18:2/18:1	5	4.81	21.0	TJ	79	Ts-5	1
18:3	5	3.08	14.2	SG	93	Sorbo	2
18:3	1	6.32	26.6	TJ	17	Ts-5	1
18:3/18:2	5	3	13.9	NG	32	Nok-3	1
18:3/18:2	1	2.97	13.6	NG	93	Nok-3	2
18:3/18:2	3	6.11	25.9	SG	10	Sorbo	2
Oil	2	6.01	26.4	KB	91	Kondara	2
Oil	5	3.2	14.9	KB	55	Kondara	2
Oil	1	3.52	15.9	TJ	83	Ts-5	1
Oil	2	3.9	17.4	TJ	79	Ts-5	1
Oil	5	2.4	11.1	TJ	79	Ts-5	1

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

5.3.2 EXAMINING THE QUALITY OF DATA

5.3.2.1 G X E INTERACTION

G x block interaction was assessed from parental lines by ANOVA using Genstat 13. Significant interactions (P<0.001) were observed in the TJ population for the trait for oil content as seen in the Table5-3.

Table 5-3 G x block interaction

Trait	Population	*P*-value	Significance
18:2	TJ	0.675	ns
18:2/18:1	NG	0.378	ns
18:3/18:2	SG	0.249	ns
18:2/18:1	TJ	0.113	ns
18:3	SG	0.216	ns
18:3	TJ	0.813	ns
18:3/18:2	NG	0.291	ns
Oil	TJ	<.001	***
Oil	KB	0.064	ns

5.3.2.2 BLOCK EFFECTS

Block effects were assessed by ANOVA using Genstat 13. The phenotypic data of two replicates from each parental line in two blocks were examined.

For FA content, as seen in Table 5-4, no significant differences were observed between blocks for 18:1, 18:2, 18:3, 18:2/18:1 or 18:2/18:3.

The block effects on oil contents were higher than that of FA contents according to Table 5-4. There was no significant difference in the first development group, however a significant difference (*P*=0.006) was observed for the KB population in the second group and the oil content of Block 1 was about 1% higher on average than Block 2, 44.12% and 43.18% respectively. This can be due to the fact that plants in group two were not randomized after the screening at B_2F_1 stage, while the plants in group 1 were randomized within the block.

As described in Section 5.2.5.1, the aim behind not randomizing plants in the second development group was to evaluate the impact of block effects on phenotype. This environmental effect must be taken account at the selection of NIL from NIL candidates derived from KB55 and KB91. Plant numbers 1 to 23 belong to Block 1 and numbers 24 to 46 to Block 2.

Table 5-4 *P* values between blocks

Group	First group		Second group			
Population	NG	TJ	KB	NG	SG	TJ
18:1	0.688	0.806	N/A	0.514	0.655	0.361
18:2	0.413	0.889	N/A	0.393	0.902	0.573
18:2/18:1	0.532	0.971	N/A	0.399	0.259	0.317
18:3	0.122	0.581	N/A	0.827	0.33	0.399
18:3/18:2	0.129	0.714	N/A	0.891	0.856	0.462
oil	N/A	0.245	0.006**	N/A	N/A	N/A

5.3.2.3 EXPERIMENTAL ERRORS

For FA content analysis, each sample set consisted of four replicates for a single line. This was split into two occasions due to the limiting capacity of the experimental instruments. Therefore the difference between the four replicates and also the difference between the two occasions were examined by ANOVA. As shown in Table 5-5, no significant differences were observed between replicates and occasions for all of the lines.

For oil content, replicates meant three measurements were taken on the same sample on the same occasion. ANOVA showed there were no significant differences

between the three measurements. The same sample sets were measured twice on different occasions that are on 26th of May 2010 and 16th of June 2010 to check if measurements on different occasions gave different results. No statistically significant differences were seen between these occasions (*P*=0.898) as seen at the end of the Table 5-5, therefore lines in the second group were measured on only one occasion.

Table 5-5 Experimental errors

NIL population	Trait	*P*-value between reps		*P*-value between occasions	
TJ25	18:2	0.634	ns	0.232	ns
NG28	18:2/18:1	0.348	ns	0.632	ns
NG32	18:2/18:1	0.852	ns	0.96	ns
TJ79	18:2/18:1	0.144	ns	0.3	ns
SG93	18:3	0.249	ns	0.122	ns
TJ17	18:3	0.609	ns	0.296	ns
NG32	18:3/18:2	0.091	ns	0.746	ns
NG93	18:3/18:2	0.901	ns	0.631	ns
SG10	18:3/18:2	0.967	ns	0.94	ns
KB55 ch5	Oil	0.902	ns		
KB91 ch2	Oil	0.986	ns		
TJ79 ch2	Oil	0.974	ns		
TJ79 ch5	Oil	0.993	ns		
TJ83	Oil	0.982	ns		
	Oil			0.898	ns

5.3.3 EXAMINING THE EFFECT OF INTROGRESSED SEGMENTS

ANOVA was performed to examine differences between groups of NIL candidates and their associated MH plants.

The groups of NIL derived from TJ25, NG28, NG32 (for 18:2/18:1); TJ17, NG93, KB55, TJ79 (for oil content on Ch2) TJ79 (for oil content on Ch5) and TJ83 showed highly significant differences from their MH plants.

The NIL material derived from TJ79 (for 18:2/18:1), NG32 (for 18:3/18:2) and SG10 showed significant differences from their MH plants as seen in Table 5-6.

The SG93 and KB91 derived material did not show differences between NIL candidates and their associated MH plants, therefore both of these lines were excluded from further analysis.

Table 5-6 Statistical differences between NIL candidates and MH plants

NIL population	Trait	*P*-value NIL-MH	
TJ25	18:2	<.001	***
NG28	18:2/18:1	<.001	***
NG32	18:2/18:1	<.001	***
TJ79	18:2/18:1	0.001	**
SG93	18:3	0.058	ns
TJ17	18:3	<.001	***
NG32	18:3/18:2	0.009	**
NG93	18:3/18:2	<.001	***
SG10	18:3/18:2	0.014	*
KB55 ch5	Oil	<.001	***
KB91 ch2	Oil	0.578	ns
TJ79 ch2	Oil	<.001	***
TJ79 ch5	Oil	<.001	***
TJ83	Oil	<.001	***

5.3.4 SELECTION OF NIL FROM NIL CANDIDATES

The number of NILs for study was reduced from 14 to 12 because the material developed from SG93 and KB91 did not show significant differences in their phenotype between the NIL candidates and MH plants as outlined in the previous section.

In this section, NILs are discussed according to the traits, and their strengths were examined. If they had a distinctive phenotype from that of their MH plants, one NIL was selected from the NIL candidates following the criteria used for the selection of NIL for fine mapping.

Charts used for the selections are in Appendix 5. Figures in Appendix 5 are prefixed with Fig. A5.

5.3.4.1 18:2 (TJ25)

Phenotype

There was a highly significant difference (P<0.001) between NIL candidates and MH plants as seen in Table 5-4.

According to the bar charts in Fig. A5-2 which show 18:2 content of the grouped NIL candidates and the grouped MH plants, the standard deviation (SD) of NIL candidates and that of MH does not overlap. Therefore the phenotypes of these NILs were regarded as sufficiently distinct from the phenotype of the MH plants and hence the NILs derived from RIL TJ25 were selected for further analysis.

Genotype

As seen in Fig. A5-1, RIL TJ25 contained four large unwanted segments of homozygous B alleles in the middle of LG1, in the top half of LG2, towards the end of LG4, in the middle of LG5, and a locus in the bottom of LG5 which were to be substituted with A alleles during the NIL development. The loci such as MS_At2_9.3 and nga1126 in the middle of LG2 and t5j1748070 in the end of LG4 were successfully changed to homozygous A alleles0 for all the NIL candidates and MH lines. The remainder of the regions continued as B homozygous or heterozygous.

Suggested plants to be fine mapped in the future

As shown in Fig. A5-3, all except NIL(7) were above the error bar of MH(26) which showed the highest 18:2 content of the MH plants. Among the NILs, NIL(10),

NIL(35) and NIL(39) were stronger lines. NIL(10) had a cleaner background with four B homozygous loci and six heterozygous loci, while NIL(35) had six B homozygous and four heterozygous, and NIL(39) had four B homozygous and eight heterozygous. Therefore (10) was selected as NIL for the QTL for 18:2 on LG3.

5.3.4.2 18:2/18:1 (NG28, NG32 AND TJ79)

Phenotype

There was a highly significant difference between the NIL candidates and the MH plants derived from NG28 (P<0.001), however the error bars of the grouped NIL candidates and the grouped MH plants overlapped (Fig. A5-8 in Appendix 5). Therefore this QTL was not taken forward.

Within the NIL candidates derived from NG32, NIL(33) had an exceptionally high ratio of 18:2/18:1 which raised a suspicion that this phenotype was caused by non-genetic factors, therefore this line is discussed separately. A highly significant difference (P<0.001) was observed between the NIL candidates and the MH plants derived from NG32 both when NIL(33) was included and excluded. Error bars slightly overlapped the grouped NILs and grouped MH plants when NIL(33) was included but it did not when NIL(33) was excluded. When examining individual plants, NIL(33), NIL(34) and NIL(40) clearly outperformed the highest MH line number (25) (Fig. A5-9 in Appendix 5).

The TJ79 derived material showed a significant difference between NIL candidates and MH plants (P=0.001), however the error bars of these two groups overlapped (Fig. A5-11), therefore this line was not considered for further analysis.

Genotype

The original RIL NG32 had two large segments of B allele on the top of LG1 and top of LG2 and three small segments in the middle and end of LG4 and top and end of LG5 (Fig. A5-7). The segment on the upper arm of LG2 and the single locus on LG4 were cleared during the NIL developing process.

Suggested plants to be fine mapped in the future

Of the NIL candidates derived from NG32, NIL(34) was chosen as it had a reasonably strong phenotype compared with the other NIL candidates and a clearer background than NIL(40).

NIL(33), on the other hand, showed an outstandingly strong phenotype which was outside that of the parental lines, however it had the cleanest genomic background of the three mentioned above, therefore it was worth considering as a candidate of NIL for fine mapping.

5.3.4.3 18:3 (TJ17)

Phenotype

A highly significant difference was observed between the NILs and the associated MH plants (P<0.001). According to Fig A5-14 in Appendix 5, the error bars of NIL candidates and MH slightly overlapped with a margin of 0.0758, however if NIL(1) which had a large SD was excluded, the margin of overlap was reduced to just 0.0185 which is not significant, hence this line was kept to be considered for further analysis.

Genotype

The original RIL TJ17 had a large segment of B alleles on LG1 adjacent to the region of interest. Another large segment of B alleles was present on the bottom half of LG2. A medium sized B allele segment lay on the upper part of LG3. LG4 and LG5 consisted of homozygous A alleles only.

Suggested plants to be fine mapped in the future

For TJ17, unlike other QTL, the A allele was the increasing allele, therefore the line with the lower content of 18:3 was chosen. NIL(7), NIL(12), NIL(15) and NIL(24) had less 18:3 than the lower error bar of the lowest MH line (26) as seen in Fig. A5-15.

NIL(7) was chosen as it had the strongest phenotype i.e. least amount of 18:3 and the cleanest genotype of the four NILs.

5.3.4.4 18:3/18:2 (NG32, NG93 AND SG10)

Phenotype

There was a significant difference between the grouped NILs and the grouped MH plants derived from NG32 (P=0.009), however the error bars of these two groups greatly overlapped, therefore no further analysis was carried out on this material.

Unlike majority of other traits, maternal (A) allele was increasing allele of the trait. A highly significant difference (P<0.001) was observed between the grouped NIL candidates and the grouped MH plants derived from NG93. The error bar of the grouped MH plants was clearly above that of the NIL candidates. There were five lines of NIL candidate lines, NIL(2), NIL(14), NIL(16), NIL(23) and NIL(26). None of the error bars for these overlapped the error bar of the lowest MH plant MH(15).

There was a significant difference (P=0.014) between the NIL candidates and the MH plants of SG10, however the overlapping error bars showed that the difference was not distinctive enough to be useful for further analysis.

Genotype

NG93 had four small to medium sized paternal homozygous segments on the top of LG2, the bottom of LG3, the top of LG4 and the bottom of LG5 (Fig A5-19). Those on the top of LG2 and LG4 were successfully replaced with A alleles following the two backcrosses. The B alleles flanking the region of interest on the bottom of LG1 were also substituted with A alleles.

Suggested plants to be fine mapped in the future

Of the NIL candidates derived from the RIL NG93, NIL(23) and NIL(26) had very clean backgrounds which were almost entirely A alleles across the genome. As the decreasing allele was B, NIL(26) which had a lower value for 18:3/18:2 than (23) was chosen.

5.3.4.5 OIL CONTENT (KB55, TJ79 CH2, TJ79 CH5 AND TJ83)

Phenotype

There was a highly significant statistical difference (P<0.001) between the NIL candidates and the MH material across all of the different RIL derived material. Also the error bars of the NIL candidates and the MH plants derived from KB55 did not overlap (Fig. A5-26 in Appendix 5). Those NIL candidates which displayed a phenotype outside the error bar of highest MH line (39) were NIL(22), NIL(26), NIL(41) and NIL(42) as seen in Fig. A5-27. Taking into consideration the block effects identified by the statistical analysis of controls in Section5.3.2.1, 1% of oil content was deducted from those plants in Block 1 such as NIL(2), NIL(4), NIL(17), NIL(22), MH(20) and MH (23). Following this the difference between NIL candidates and MH remained very significant (P<0.001). As shown in Fig. A5-27, the chosen

lines also remained the same after the 1% deduction of oil contents from plants grown in Block 1.

A highly significant difference (P<0.001) was observed between NIL candidates and MH plants derived from TJ79 for the QTL on chromosome 2, and they were distinctively different without overlap of error bars (Fig. A5-29). The values of NIL(18) and NIL(23) were both above the error bar of the strongest MH line (20) as seen in Fig. A5-30.

There was a highly significant difference (P<0.001) between the NIL candidates and MH plants derived from TJ79 for the QTL on chromosome 5. However the error bars between the two groups overlapped as seen in Fig. A5-32, therefore this line was not included in further analysis.

A highly significant difference (P<0.001) was observed between NIL candidates and MH plants derived from TJ83 and the respective error bars did not overlap (Fig. A5-35). Four of the five NILs: NIL(22), NIL(31), NIL(33) and NIL(43) had stronger phenotypes than the strongest MH line (11) as shown in Fig. A5-36.

Genotype

In the RIL KB55 half of LG3 and most of LG4 was paternal homozygous (Fig. A5-25). As a result of backcrossing, the bottom half of LG4 was exchanged into maternal homozygous.

The RIL line TJ79 had three small segments of homozygous B alleles at the bottom of LG2, the top of LG3 and at the middle of LG5 as seen in Fig. A5-28. Two large B homozygous segments were observed on the top half of LG4 and the lower part of LG5. After the two backcrossings and single selfing, all of the loci in LG4 and LG5 and also those at the bottom of LG2 were likely to be replaced with A alleles.

According to Fig. A5-34, RIL TJ83 had a total of six segments of homozygous B alleles one at the end of LG2, two on the upper part of LG3, one on the bottom of LG4 and two in the middle of LG5. The regions at the end of LG2 and LG4 were presumably cleared however the rest remained B homozygous or heterozygous.

Suggested plants to be fine mapped in the future

From the RIL KB55, NIL(41) and NIL(42) had the cleanest genetic backgrounds of all the high phenotype candidates (Fig. A5-25 and Fig. A5-27). NIL(41) was chosen as it showed a stronger phenotype than NIL(42).

NIL(18) and NIL(23) from TJ79 (Ch2) had the same genotype according to Fig. A5-28, therefore NIL(23) was chosen as it showed a much stronger phenotype than NIL(18) as seen in Fig. A5-30.

For the NIL candidates derived from TJ83, NIL(33) had the cleanest genotype of the high phenotype lines. It was also free from B alleles at the end of chromosome 5 where another QTL for the same trait lay.

5.3.5 SELECTION OF FINAL NIL FOR FINE MAPPING

As mentioned in the methodology Section 5.2.10, the final NIL was selected from the first development group as this material was developed six months ahead of the second group. From the first group the material derived from: NG32 for 18:2/18:1, TJ17 for 18:3, TJ79 Ch2 for oil content and TJ83 for oil content were all short listed as described in the previous section.

Selection of the final NIL for fine mapping required a combination of complex criteria, such as

- Strength of the original putative QTL.

- The accuracy of phenotypic measurement.

- Growing characteristics of the population.

- The observed association of genotype and phenotype characteristics during NIL development.

NG32 had one of the largest QTL LOD score at 11.82 followed by TJ17 with 6.32. TJ79 (Ch2) and TJ83 scored 3.9 and 3.52, respectively.

From the point of experimental accuracy, the experimental error in the measurement of FA composition, between replicates was smaller for NG32 (P=0.8520) than TJ17 (P=0.609) and also the error between occasion also showed the same results, P=0.96 and P=0.296 respectively. Therefore when comparing NG32 and TJ17, the phenotype of NG32 was more accurately measured.

For oil content as measured by NMR, comparing TJ79 (Ch2) and TJ83, the accuracy of phenotypic measurements was very similar, P=0.974 and P=0.982, respectively.

Regarding plant husbandry, the NG population was more vigorous and larger in size, hence they produced more seeds compared to the TJ population lines. The accession Nok-3 was the maternal backcrossing parent and it set seeds more readily than Ts-5 which was the maternal parent of the TJ population. The environmental effect on phenotype was smaller for FA composition than for total oil contents.

The NG28 derived material showed a significant difference between the NIL candidates and the MH plants, however it was not distinctive enough to proceed with for further analysis while NG32 showed a stronger phenotype. NG32 also had longer span of introgressed segment possibly capturing the QTL causative locus better than NG28.

In conclusion, the NIL material derived from NG32 for the trait 18:2/18:1 was chosen for fine mapping.

Within NIL derived from NG32 NIL(34) was chosen for fine mapping, however the extraordinarily strong phenotype displayed by NIL(33) with a cleaner genotype was difficult to ignore despite the risk that the cause of the strong phenotype could be non-genetic. Therefore both lines NIL(33) and NIL(34) were selected to be fine mapped.

5.4 CONCLUSION

5.4.1 NIL AND HIF

At the final step of NIL development, plants which were heterozygous at the region of interest were successfully identified. These plants were equivalent to HIF(Tuinstra et al., 1997) and their potential uses were considered. If a QTL is dominant, the heterozygosity at the loci associated with the QTL should be sufficient to cause phenotypic differences to be seen from recessive homozygous lines. Also because HIF material is heterozygous at the region of interest, subsequent development of a mapping population for fine mapping can be done only by selfing them, while NIL material needs to be backcrossed prior to the selfing which adds a further generation to the process. However, in the absence of prior knowledge on the dominant or co-dominant effects of the QTL , the first group of plants, NG28 for 18:2/18:1, NG32 for 18:2/18:1 and 18:3/18:2, TJ79 for 18:2/18:1, TJ17 for 18:3, TJ79 for oil content on Ch2, TJ79 for oil content on Ch5 and TJ83 for oil content, were taken through and developed as NILs as planned in Fig. 5-1. The result showed that although their phenotypic differences were statistically significant from their controls, these were smaller than anticipated. By using heterozygous plants the analysis would be more complex, therefore the second group of study lines, TJ25 for 18:2, SG93 for 18:3, NG93 for 18:3/18:2, SG10 for 18:3/18:2, KB55 for oil content and KB91 for oil content, also followed the same procedure as that with the first group.

5.4.2 BLOCK EFFECTS

The different experimental designs between the first group and the second group demonstrated that total seed oil content is environmentally more sensitive than is FA composition. In fact, although no statistically significant difference was observed between blocks in the first group, the *P* value of oil content was relatively smaller than that of FA composition, indicating that therefore there was also more variation between blocks for total oil content than FA composition in the first group as well.

The average block effect of all traits in the first group which was randomized within a block was $P = 0.641$ and that of the second group which was not randomized within a block was $P = 0.398$. This result shows that both of them did not have significant difference between blocks, however plants in the first group were grown in more uniformed environment than the second group. Therefore the importance of

application of randomization was reaffirmed in order to obtain better quality of the data for statistical analysis. Therefore frequent randomization within a block is recommended, especially for the phenotype of oil content as it was shown to be more sensitive to block effects.

5.4.3 THREE PURPOSES OF NIL DEVELOPMENTS

NILs were developed for three purposes, the first was to test QTLs identified in previous experiments and the second purpose was to examine the validity of the methods of NIL development. The third purpose was to generate material suitable for fine mapping.

Other than the NILs derived from KB91 and SG93, all NILs were significantly different from their associated MH lines. Especially nine NILs showed highly significant differences between their NIL candidates and MH plants ($P<0.001$), therefore these 11 QTL had successfully proven to effect the phenotype.

The process of NIL development in this project was designed to be as quick and efficient as possible, therefore selection of plant material at each step was crucial. The phenotyping of FA composition was especially laborious and costly, therefore it was carried out only once during the NIL development. Instead, genotyping was done throughout the process to ensure the capture of those segments over which the putative QTLs were lying and also to avoid unwanted regions such as those associated with other QTL of the same trait. Although two backcrossings did not completely remove paternal alleles from the genomic background, it did generate material that was sufficiently clean to distinguish the effects of target QTL as described in the results above.

The third purpose was to select a single NIL for fine mapping from all the NILs developed successfully. Although significant differences were observed between NIL candidates and their MH plants, for a plant to be used for fine mapping it needed to be examined with strict multifaceted criteria and must be able to hold the desired phenotype during the development of the mapping population.

As a result a further four QTL were dropped however all target traits retained at least one NIL which was strong enough for further analysis. Those selected QTL and their NIL for further analysis were TJ25 NIL(10) for 18:2, NG32 NIL(33) or (34) for 18:2/18:1, TJ17 NIL(7) for 18:3, NG93 NIL(26) for 18:3/18:2 and KB55 NIL(41), TJ79 Ch2 NIL(23) and TJ83 NIL(33) for oil content.

Among them, the NG32 derived NIL material for 18:2/18:1 was taken forward for fine mapping because the target QTL had the highest LOD score of all the target QTL, it was therefore felt that there was a strong possibility of discovering the basis of the QTL. Also the experimental error at FAME preparation and GC analysis was the smallest amongst the first group, meaning its data was more accurate and reliable. The vigour of its maternal line, Nok-3, would be beneficial at the plant husbandry stages of the next two generations during the development of the large scale mapping population.

Moreover the trait, 18:2/18:1, fits nicely with a project on *Brassica napus* within the group. Rachel Wells is currently engaged in the optimization of FA composition in *Brassica napus* for the biolubricant industry where 18:1 was favoured over 18:2. Therefore the findings within this PhD project could be useful for future breeding programs.

Two NIL candidates derived from NG32 for 18:2/18:1 were selected for fine mapping. Those lines were NG32 NIL(33) which had an outstandingly strong phenotype and a background that contained less B alleles compared to the other NIL candidates. The second line was NG32 NIL(34) which had a strong phenotype compared with the other NIL candidates but was within the expected range according to the parental phenotypes. If the phenotype of NG32 NIL(33) was caused by genetic factors, it could give clear and distinctive results during the fine mapping. NG32 NIL(34) might produce less distinctive results, however it was more likely that the stronger phenotype exhibited by this line was genetically determined.

CHAPTER 6. RELATIVE CONTENT OF LINOLEIC ACID /OLEIC ACID AND THE DETERMINATION OF CANDIDATE GENES

6.1 INTRODUCTION

Large scale mapping populations were developed for fine mapping of the selected NILs, NG32 NIL(33) and NG32 NIL(34) as described in the previous chapter. The schematic charts in Fig. 6-1 explain the development processes of these populations. NG32 NIL(33) and NG32 NIL(34) were called NIL(33) and NIL(34) in this chapter. NILs were backcrossed to their maternal line, Nok-3, followed by a single round of selfing in order to allow recombination events within the introgressed segment. As a result plants in the population segregated for the parental alleles at the polymorphic loci. The genetic differences in the introgressed segment between plants in the mapping population would enable the detection of the region affecting the phenotype.

Lines NIL(33) and NIL(34) selected in Chapter 5 showed different strength in this phenotypes. NIL(33) had an unexpectedly strong phenotype while NIL(34) had a reasonably strong phenotype. Mapping populations derived from NIL(33) and NIL(34) were called population(33) and population(34), respectively.

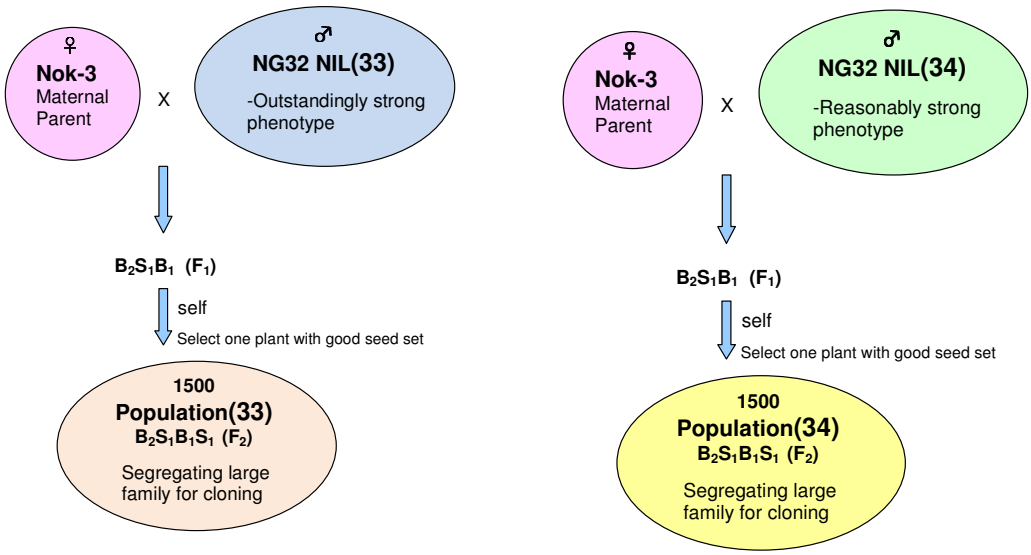


Fig.6-1 Schematic images of NIL mapping populations development

Mapping populations experiments were designed primarily to screen plants for genotypes and phenotyping was performed only once at the end to characterize those plants selected after repeated rounds of genotyping. This was due to the fact that phenotyping of FA composition is a lengthy and complex process requiring chemical preparation of FAMEs followed by GC analysis and the data processing of GC results as described in Section 3.7. Genotyping is much more cost and time efficient therefore it was used to lead the selection during population development.

The initial strategy of fine mapping was to select those plants which contained informative recombination within the interval defined by markers flanking the QTL. The second stage was to delineate the position of recombination points within the interval by classifying the plants into twelve bins defined by eleven markers equally positioned between the two flanking markers. Both objectives are dealt with in this chapter.

The size of the mapping population, 1500 plants for each NIL, was calculated to be necessary to ensure at least five plants in each bin type for phenotyping. The genetic distance between the two flanking markers was 41cM, therefore 615 plants were expected to have recombination within the interval. These 615 plants would then be sorted into 48 bins (12 bins x 4 types of recombination pattern), hence each bin could in theory contain 12 plants.

The phenotypes of each bin type were compared across the recombination types and these were examined to identify a bin set that influenced the phenotype.

Subsequently candidate genes were searched within the region covered by each bin and the adjacent bins using The Arabidopsis Information Resource (TAIR). TAIR is a widely used database with information of gene structure, gene product, metabolism, gene expression, genome maps, genetic and physical markers, publications, and the Arabidopsis research community.

The candidate genes listed in this chapter will be compared with the *B. napus* genome in Chapter 7, together with other QTL which successfully showed their distinctive phenotype against the control material in Chapter 5.

6.2 MATERIAL AND METHODS

6.2.1 PLANT MATERIAL

F_2 seeds of NIL(33) and NIL(34) together with seeds of the parental lines and the original RIL, NG32, were used.

6.2.2 PLANT HUSBANDRY

Seeds were sown in seed trays as described in Section 3.2.

1500 plants for each population and 10 plants each for parental lines and RIL NG32 were transplanted individually into 7cm square pots which were arranged in trays which hold 30 pots in a 5 x 6 format.

Plants were transferred to Room 2 in the W1 glasshouse initially and the first rotation was carried out after a week. Ten sets of three plants consisting of maternal parent, paternal parent and RIL NG32 plants were positioned lengthways down the centre of the bench to allow systematic examination of block effects as seen in Fig.6-2.

A week after the first rotation, due to a problem with the glasshouse, plants were moved to glasshouse Room 3 which had equivalent growth conditions and the layout of trays was altered as seen in Fig.6-3. Rotation was subsequently carried out every week. Leaf samples were taken 28 days after transplanting and submitted to the DNA extraction service.

The first genotyping, staking and bagging were carried out simultaneously. Plants were gathered to fill the gaps of discarded plants but alternately spaced to allow expansion of their branches. As a result, the number of plants per tray was reduced by half after the first screen. Details of the first screening are described in Section 6.2.3.

Plants were rotated twice after they were bagged. Plants were harvested and stored until further screening for binning was complete.

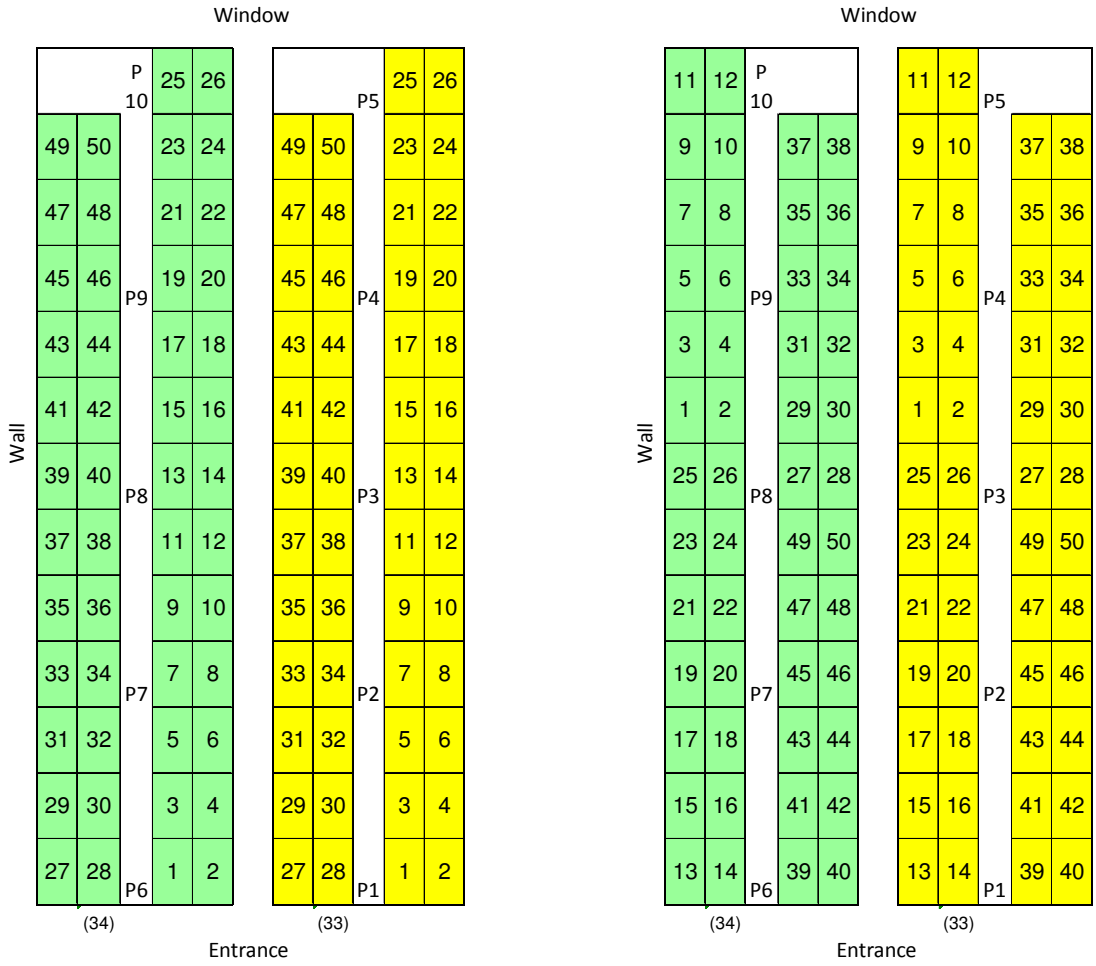


Fig.6-2 Tray layouts of mapping populations in Room 2 of the West 1 glasshouse

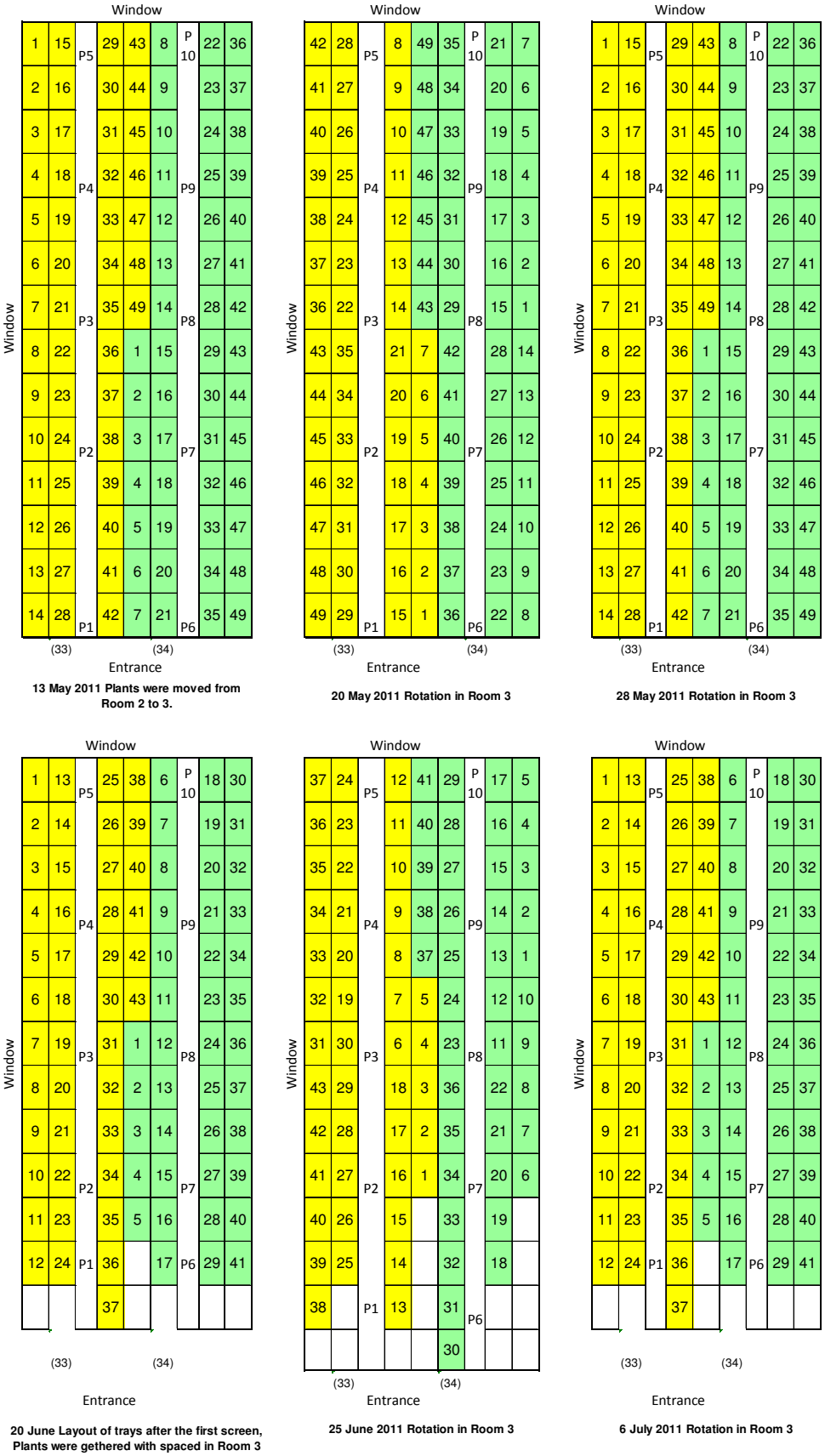


Fig.6-3 Tray layouts of mapping populations in Room 3 of the West 1 glasshouse

6.2.3 FIRST SCREENING

The first screening was carried out using two flanking markers, mhf15ind52-52 at 1.7Mb and MS_At5_9.7 at 9.7Mb on chromosome 5.

A total of 32 sample plates from the two populations were genotyped with these two markers using the methods described in Section 3.6.

Selection of plants and bagging were carried out as soon as the genotype results became available. Plants with informative recombination were categorized into four types as shown in Table 6-1.

As controls, 16 plants with paternal homozygous alleles at both flanking markers and also 16 plants, maternal homozygous at these markers, were kept from each population for comparison of phenotype results.

Table 6-1 Four types of recombination patterns to be selected at the first screening

	Physical position	Type1	Type2	Type3	Type4	Control	
mhf15ind52-52	1.7Mb ch5	AB	AA	BB	AB	AA	BB
MS_At5_9.7	9.7Mb ch5	AA	AB	AB	BB	AA	BB

6.2.4 BINNING

Those plants classified according to the four recombination types at the first screen were then genotyped again for sorting into twelve bins which together span the region 1.7Mb to 9.7Mb. These bins were at approximately 0.67Mb intervals (Fig.6-4).

As seen in Fig.6-4, plants of Type 1 show increased heterozygosity in the direction top-down replacing homozygous A in doing so. By comparison Type 2, worked in the opposite direction ie bottom-up but also started with increased heterozygous alleles. Type 3 was B homozygous at the top of the interval and became increasingly heterozygous going down. Whilst Type 4 was the opposite of Type 3 these were B homozygous at bottom of the segment and became increasingly heterozygous going up.

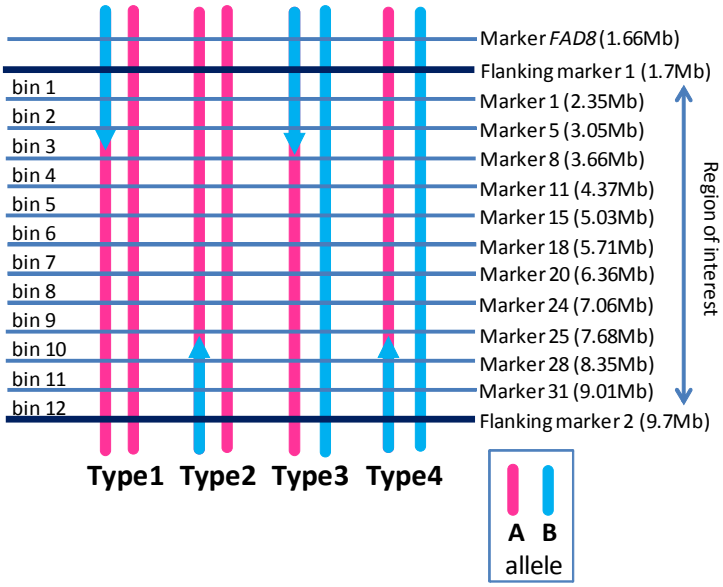


Fig.6-4 Schematic of binning strategy in four different recombination types

6.2.4.1 PRIMER DESIGN FOR BINNING

6.2.4.1.1 SEQUENCING OF PARENTAL LINES

To identify polymorphisms suitable for marker development, genomic sequencing of the parental lines was carried out. All eight parental lines of the NILs developed, i.e. Kondara, Br-0, Nok-3, Ga-0, Sorbo, Gy-0, Ts-5 and Mz-0 were sequenced by TGAC.

Quality checks of the samples were performed and the best samples from each parental line were put for sequencing. Selected lines, their concentrations and DNA amounts are detailed in Appendix 9.

Libraries for each parental line were constructed by TGAC with an average insert size 297bp. Libraries for the eight parental lines were pooled and run on one Illumina HiSeq 2000 lane with 100bp paired end reads.

6.2.4.1.2 INDEL MARKERS AND PRIMER DESIGN

The raw sequence data were aligned using the Colombia sequence as a reference by Martin Trick. The individual parental sequences were viewed using the software Xming and Mapview. Trick also detected 12944 SNP markers and 5288 InDel

markers between Nok-3 and Ga-0 on chromosome 5 from 1.7Mb to 9.7Mb and designed an automated method for primer design using the web-based software Primer 3. An annealing temperature of 60 ^{o}C was chosen for all markers. The frequency of both the SNP markers and the InDel markers was high enough to separate the region of interest into 12 bins as planned.

Due to the large number of samples for genotyping in a limited period, the InDel markers were preferred over the SNP markers as they required a much shorter reaction cycle during PCR amplification compared to SNPs. Also InDel PCR products can be multiplexed up to 40 at the genotyping stage if the products' sizes and fluorescent labels are distinctive enough from each other. Therefore it is a very cost effective method. Additionally the output of genotyping can be scored automatically using Gene Mapper software whilst manual scoring of the sequenced amplicons is necessary for SNP markers.

6.2.4.2 SELECTION OF MARKERS

InDel markers with a minimum of a 5bp difference between Nok-3 and Ga-0 amplicons were selected within each 0.67Mb interval. Parental sequences were aligned with Colombia sequence as a reference, therefore insertions or deletions in parental lines appeared as mismatch alerts when reading from the reference sequence causing disruption of the alignments. These alerts were used to indicate the presence of InDels. Subsequently the InDel markers identified were confirmed by checking the sequence data of both parental lines to see if they were real polymorphisms. These insertions or deletions could result in size differences of the PCR products between parental lines.

Three markers each for all 11 positions were selected and the primers were ordered from Sigma-Aldrich. These primers were tested with the parental DNA, Nok-3 and Ga-0, as to whether they could provide single and different sized amplicons for each marker, and if they did, the actual product sizes were scored. Of the 33 markers designed and tested, 29 were successfully genotyped and eleven of these were chosen for the colour of their fluorescent label and the size of PCR product for the convenience of multiplication of all eleven markers applied to a single plant. The selected markers are listed in Table 6-2.

Table 6-2 List of selected InDel markers and their PCR products sizes

No.marker	Bin	Chr:position (bp)	Label	Average size Nok-3 (bp)	Average size Ga-0 (bp)
1	1	Chr5:2354639	vic	182.005	197.105
5	2	Chr5:3048468	vic	310.26	318.64
8	3	Chr5:3659351	fam	242.025	229.78
11	4	Chr5:4367144	ned	344.35	329.98
15	5	Chr5:5031708	ned	254.98	245.085
18	6	Chr5:5708080	fam	344.2	337.3
20	7	Chr5:6361746	pet	231.155	212.22
24	8	Chr5:7063782	pet	421.755	417.845
25	9	Chr5:7679646	vic	150.775	165.835
28	10	Chr5:8353051	pet	106.875	101.095
31	11	Chr5:9009404	ned	475.71	470.64

6.2.4.3 PCR AND GENOTYPING

In order to have a minimum of five plants in each bin, four complete sample plates (376 samples in total) were extracted for each population from younger numbered plants. If the resulting number of plants was insufficient, additional plants were to be genotyped. Genotyping was carried out with the eleven markers described in the previous section. The PCR products of the eleven markers were multiplexed and submitted for genotyping following the same methodology described in Section 3.6. The genotyping results were scored automatically using Gene Mapper v.3.7 and subsequently manually checked.

6.2.4.4 SORTING INTO BINS

The hypothesis was that there could be a causative locus of the QTL within the interval defined by the two flanking markers. The purpose of the binning system was to identify the genomic region which positively affected the 18:2/18:1 ratio through the paternal (B) allele within the interval. The interval was divided into 12 bins and plants were sorted according to their positions of recombination. Examples of binning system of the four recombination types are shown in Fig.6-5. Type 1 bin 1 (1-1) contained B allele only up to the margin between bin 1 and bin 2 from the top of the introgressed segment, while Type 2 bin 1 contained B alleles all the way from the bottom of the segments up to bin 1. Types 3 and 4 were similar to Types 1 and 2, respectively, however they had B alleles along a complete haploid chromosome. They were included so as to check if the effect of the QTL was a dominant one.

Five plants from each of the 48 bins (12 bins x 4 types) were selected and phenotyped for their FA compositions.

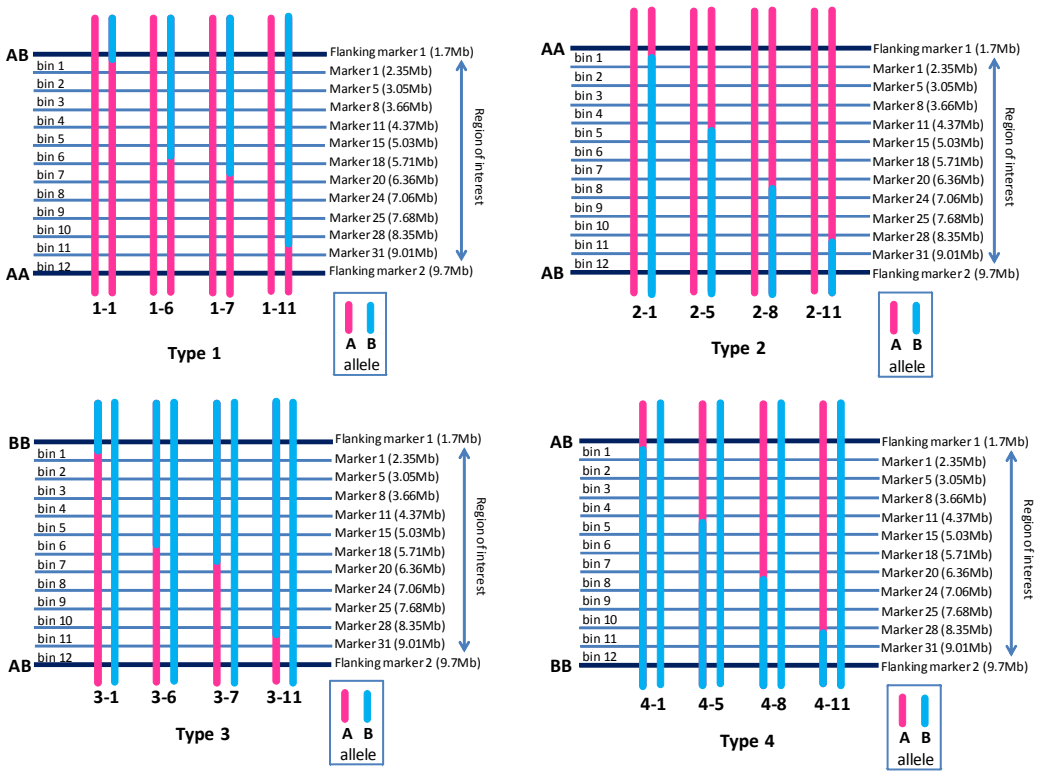


Fig.6-5 Examples of binning in four different recombination types

6.2.5 MEASURING FATTY ACID COMPOSITION OF MAPPING POPULATIONS

Five plants from each of the 48 bins were to be phenotyped, hence the number of samples required per population was 240 plus controls. As 40 samples can be processed at one time, seven processing sessions would be required to complete one population. Technical replicas were also required because samples from the mapping population did not have biological replicates due to the large number of plants. Therefore a set of parental lines was used as technical control material for all seven processing sessions for each population, these were Nok-3(1), Ga-0(1) and RIL NG32(1) for population(33), and Nok-3(6), Ga-0(6) and RIL NG32(6) for population(34). Two sets of parental lines Nok-3(1), Ga-0(1), RIL NG32(1), Nok-3(6), Ga-0(6) and RIL NG32(6) were included in the seventh processing session of both populations to be used for calculating errors across the populations. Additional parental samples were spread throughout the processing occasions as shown in Table 6-3. Seed FAME preparation was carried out as explained in Section 3.7.1. GC analysis was performed at Rothamsted Research Institute as described in Section 3.7.2.

Table 6-3 Sets of samples of mapping populations for FAME analysis

Population (33)

1		2			3			4			5			6			7		
	sample. no		sample. no			sample. no			sample. no			sample. no			sample. no			sample. no	
1	41	41	125	1	81	71	1	121	455	1	161	222	1	201	9	1	241	3	1
2	67	42	333	2	82	191	2	122	499	2	162	224	2	202	445	2	242	269	2
3	318	43	404	3	83	217	3	123	642	3	163	361	3	203	525	3	243	428	3
4	324	44	489	4	84	420	4	124	672	4	164	434	4	204	564	4	244	468	4
5	536	45	624	5	85	729	5	125	711	5	165	831	5	205	961	5	245	503	5
6	162	46	42	6	86	214	6	126	119	6	166	21	6	206	382	6	246	35	6
7	375	47	155	7	87	326	7	127	129	7	167	29	7	207	496	7	247	55	7
8	814	48	185	8	88	438	8	128	279	8	168	539	8	208	618	8	248	365	8
9	880	49	461	9	89	451	9	129	613	9	169	573	9	209	676	9	249	459	9
10	54	50	73	10	90	699	10	130	650	10	170	584	10	210	833	10	250	467	10
11	112	51	160	11	91	393	11	131	265	11	171	28	11	211	84	11	251	74	11
12	343	52	207	12	92	235	12	132	475	12	172	221	12	212	508	12	252	115	12
13	443	53	421	13	93	410	13	133	500	13	173	374	13	213	1018	13	253	145	13
14	534	54	645	14	94	567	14	134	701	14	174	446	14	214	81	14	254	154	14
15	2	55	33	15	95	616	15	135	703	15	175	604	15	215	92	15	255	156	15
16	114	56	239	16	96	27	16	136	304	16	176	198	16	216	172	16	256	79	16
17	136	57	482	17	97	296	17	137	480	17	177	755	17	217	229	17	257	147	17
18	218	58	575	18	98	368	18	138	630	18	178	984	18	218	353	18	258	278	18
19	259	59	795	19	99	619	19	139	761	19	179	78	19	219	66	19	259	412	19
20	171	60	429	20	100	149	20	140	169	20	180	178	20	220	371	20	260	418	20
21	478	61	602	21	101	295	21	141	195	21	181	497	21	221	640	21	261	388	21
22	718	62	674	22	102	367	22	142	309	22	182	666	22	222	180	22	262	510	22
23	131	63	710	23	103	370	23	143	327	23	183	820	23	223	315	23	263	576	23
24	359	64	839	24	104	651	24	144	678	24	184	234	24	224	550	24	264	834	24
25	433	65	Nok1-2	25	105	44	25	145	23	25	185	246	25	225	132	25	265	909	25
26	799	66	Nok2	26	106	103	26	146	272	26	186	377	26	226	414	26	266	Nok1-7	26
27	860	67	Nok3	27	107	133	27	147	294	27	187	542	27	227	552	27	267	Nok6-8	27
28	1	68	Nok4	28	108	287	28	148	331	28	188	692	28	228	995	28	268	Ga1-7	28
29	17	69	Nok5	29	109	349	29	149	380	29	189	126	29	229	1052	29	269	Ga6-8	29
30	201	70	Ga1-2	30	110	50	30	150	37	30	190	177	30	230	104	30	270	NG32 -1-7	30
31	558	71	Ga2	31	111	128	31	151	243	31	191	310	31	231	298	31	271	NG32 -6-8	31
32	745	72	Ga3	32	112	161	32	152	523	32	192	385	32	232	406	32	272	1189	32
33	25	73	Ga5	33	113	643	33	153	531	33	193	486	33	233	679	33	273	1335	33
34	113	74	NG32 -1-2	34	114	709	34	154	596	34	194	Nok1-5	34	234	857	34	274	1344	34
35	372	75	NG32 -2	35	115	Nok1-3	35	155	Nok1-4	35	195	Ga1-5	35	235	Nok1-6	35	275	blank	35
36	398	76	NG32 -3	36	116	Ga1-3	36	156	Ga1-4	36	196	NG32 -1-5	36	236	Ga1-6	36			
37	511	77	NG32 -4	37	117	NG32 -1-3	37	157	NG32 -1-4	37	197	607	37	237	NG32 -1-6	37			
38	Nok1-1	78	NG32 -5	38	118	168	38	158	417	38	198	633	38	238	882	38			
39	Ga1-1	79	49	39	119	165	39	159	493	39	199	714	39	239	891	39			
40	NG32 -1-1	80	39	40	120	319	40	160	559	40	200	828	40	240	997	40			

Population (34)

8		9			10			11			12			13			14		
	sample. no		sample. no			sample. no			sample. no			sample. no			sample. no			sample. no	
1	15	41	369	1	81	106	1	121	262	1	161	109	1	201	688	1	241	31	1
2	132	42	507	2	82	125	2	122	428	2	162	263	2	202	842	2	242	130	2
3	191	43	669	3	83	445	3	123	489	3	163	314	3	203	209	3	243	225	3
4	648	44	569	4	84	471	4	124	538	4	164	906	4	204	336	4	244	494	4
5	661	45	804	5	85	482	5	125	938	5	165	947	5	205	1047	5	245	678	5
6	297	46	840	6	86	19	6	126	2	6	166	226	6	206	1170	6	246	469	6
7	420	47	14	7	87	153	7	127	37	7	167	533	7	207	288	7	247	523	7
8	596	48	36	8	88	459	8	128	241	8	168	555	8	208	615	8	248	621	8
9	626	49	162	9	89	852	9	129	556	9	169	618	9	209	815	9	249	923	9
10	4	50	434	10	90	265	10	130	951	10	170	862	10	210	1240	10	250	941	10
11	66	51	509	11	91	325	11	131	346	11	171	207	11	211	359	11	251	176	11
12	68	52	13	12	92	892	12	132	402	12	172	528	12	212	388	12	252	316	12
13	427	53	97	13	93	12	13	133	458	13	173	586	13	213	435	13	253	631	13
14	495	54	281	14	94	237	14	134	43	14	174	685	14	214	645	14	254	874	14
15	320	55	530	15	95	333	15	135	272	15	175	918	15	215	937	15	255	897	15
16	368	56	610	16	96	496	16	136	446	16	176	348	16	216	59	16	256	403	16
17	544	57	5	17	97	543	17	137	680	17	177	356	17	217	136	17	257	527	17
18	608	58	572	18	98	232	18	138	790	18	178	376	18	218	184	18	258	663	18
19	886	59	575	19	99	286	19	139	46	19	179	483	19	219	487	19	259	943	19
20	221	60	578	20	100	395	20	140	88	20	180	1000	20	220	859	20	260	1027	20
21	269	61	592	21	101	429	21	141	178	21	181	26	21	221	259	21	261	Nok1-8	21
22	279	62	76	22	102	890	22	142	253	22	182	50	22	222	490	22	262	Nok6-7	22
23	360	63	231	23	103	129	23	143	649	23	183	155	23	223	865	23	263	Ga1-8	23
24	644	64	719	24	104	168	24	144	28	24	184	606	24	224	984	24	264	Ga6-7	24
25	312	65	131	25	105	211	25	145	60	25	185	952	25	225	177	25	265	NG32 -1-8	25
26	Nok6-1	66	140	26	106	439	26	146	201	26	186	75	26	226	255	26	266	NG32 -6-7	26
27	Nok7	67	334	27	107	619	27	147	452	27	187	133	27	227	570	27	267	792	27
28	Nok8	68	568	28	108	52	28	148	593	28	188	170	28	228	855	28	268	917	28
29	Nok9	69	770	29	109	84	29	149	137	29	189	305	29	229	966	29	269	936	29
30	Nok10	70	390	30	110	436	30	150	287	30	190	480	30	230	32	30	270	1064	30
31	Ga6-1	71	411	31	111	504	31	151	350	31	191	72	31	231	113	31	271	1072	31
32	Ga7	72	438	32	112	520	32	152	713	32	192	332	32	232	144	32	272	1242	32
33	Ga8	73	442	33	113	152	33	153	1016	33	193	822	33	233	180	33			
34	Ga9	74	524	34	114	194	34	154	Nok6-4	34	194	939	34	234	219	34			
35	Ga10	75	539	35	115	468	35	155	Ga6-4	35	195	979	35	235	Nok6-6	35			
36	NG32 -6-1	76	Nok6-2	36	116	566	36	156	NG32 -6-4	36	196	Nok6-5	36	236	Ga6-6	36			
37	NG32 -7	77	Ga6-2	37	117	896	37	157	343	37	197	Ga6-5	37	237	NG32 -6-6	37			
38	NG32 -8	78	NG32 -6-2	38	118	Nok6-3	38	158	337	38	198	NG32 -6-5	38	238	668	38			
39	NG32 -9	79	230	39	119	Ga6-3	39	159	437	39	199	515	39	239	682	39			
40	NG32 -10	80	223	40	120	NG32 -6-3	40	160	415	40	200	549	40	240	798	40			

Type1
Type2
Type3
Tyoe4
AA
BB

6.2.6 STATISTICAL ANALYSIS OF PHENOTYPIC DATA

6.2.6.1 CHECKING QUALITY OF DATA

Parental lines and controls were examined for environmental effects or experimental error using ANOVA within Genstat.

6.2.6.2 IDENTIFYING THE POSITION OF THE QTL CAUSATIVE LOCUS

Hypothetically, 12 bins were to be divided into two groups, a group with maternal (A) allele at the causative of QTL and another group with paternal (B) allele. The phenotype of the same group should remain constant and a difference between groups should occur only once between the marginal bins of groups. Therefore in Types 1 and 3 the phenotype was expected to remain low at the earlier bins and shift upwards after the bin which contained the causative locus of the QTL as B allele, while in Types 1 and 4, it should start with higher values and drop just after the bin which contained A allele at the causative locus of the QTL.

To identify a shift of phenotype occurring between bins, T-tests ($P \leqq 0.05$) were performed between samples in a bin and samples in its adjacent cluster of bins. For example, if no significant difference was observed between samples in bin 1 and samples in bin 2 by T-test, bin 1 and bin 2 were combined as a cluster of bins and plants in bin 1 and bin 2 were compared with plants in bin 3 by T-test. As long as the T-test did not show the significant difference ($P \leqq 0.05$) between them, plants in an adjacent bin were combined and compared to plants in the next adjacent bin. This process was continued until a significant difference was observed between plants in a cluster of bins and plants in the bin being tested.

If a significant difference was observed, this indicated the position where a shift of phenotype occurred. However to confirm there was no other shift in the reminder of the bins, the analysis were repeated from the bin showing significant difference and its adjacent bin. The process was continued until bin 12.Then the analysis was repeated in reverse from bin 12 to bin 1.

6.3 RESULTS

Plants were screened at two markers flanking the genomic interval that lay underneath the QTL in order to select plants with recombination in that region. The presence of multiple recombination events facilitates positioning of the causative genomic region underlying the QTL. Selected plants were then genotyped with eleven markers and categorized according to the binning system described. Five plants from each bin were phenotyped for their FA content and differences between the bins in each recombination type were examined. T-tests were carried out to identify which of the binned regions was associated with the target QTL.

6.3.1 FIRST SCREENING

The first screening was performed to select plants with informative recombination, with 602 plants being selected for the population(33) and 559 plants for the population(34). Sixteen plants each of those which were paternal homozygous at both markers and maternal homozygous at both markers were kept as controls. The detailed numbers of plants for the four recombination types are shown in the Table 6-4 below.

Table 6-4 Number of plants in four recombination types

Type	1	2	3	4	Total	Control	
mhf15ind52-52	AB	AA	BB	AB		AA	BB
MS_At5_9.7	AA	AB	AB	BB		AA	BB
Number of plants (33)	156	161	160	125	602	16	16

6.3.2 SECOND SCREENING FOR SORTING PLANTS INTO 48 BINS

Plants selected in the first screen and categorized into the four recombination types were genotyped with a further eleven markers for further classification. A total of 376 samples for each population were genotyped and sorted into 12 bins for each of the four types of recombination patterns with the aim of identifying five plants for each bin.

For some of the bins five lines could not be selected for phenotyping as seen in Table 6-5, however the majority of the bins contained more than five lines, therefore the phenotyping was carried out.

Table 6-5 Numbers of the different recombination type plants across the bins for both mapping populations

	Population (33)			
(33)	Type1	Type2	Type3	Type4
1	6	5	4	6
2	4	5	10	3
3	6	9	8	7
4	11	1	6	3
5	3	4	9	3
6	5	4	8	8
7	7	7	11	6
8	11	10	3	7
9	8	8	6	8
10	4	16	8	11
11	8	8	7	14
12	6	7	8	9
Total	79	84	88	85

	Population (34)			
(34)	Type1	Type2	Type3	Type4
1	7	1	3	2
2	4	8	8	4
3	11	8	7	4
4	9	4	14	5
5	6	3	5	10
6	4	10	9	4
7	3	7	6	6
8	10	12	7	9
9	9	9	6	7
10	12	8	8	15
11	4	11	7	8
12	5	7	5	10
Total	84	88	85	84

Detail of plants in bins can be seen in Appendix 6.

6.3.3 PHENOTYPE DATA OF MAPPING POPULATIONS

6.3.3.1 EXPERIMENTAL ERROR AND ENVIRONMENTAL EFFECTS

As the FAME preparation was limited to 40 samples at a time due to the instrument capacity, it took 14 occasions over five days to complete both populations. Therefore experimental error between occasions and populations had to be considered.

To test experimental error between preparation occasions, a set of parental lines was included and spread throughout all seven occasions of FAME preparation for both populations to serve as biological replicates. The lines used were Nok-3(1), Ga-0(1) and RIL NG32(1) for population (33), Nok-3(6), Ga-0(6), and RIL NG32(6) for population (34). The results of ANOVA showed there were no differences between occasions (P=1), therefore samples prepared and analyzed on different occasions could be considered as a single data set.

To assess experimental error between the two populations two sets of parental lines were used for both populations (33) and (34) on the seventh and fourteenth occasions, respectively. ANOVA was carried out to compare the data sets to see if

there was a significant difference between the two populations. No significant difference was observed (*P*=0.779).

Environmental effects were examined using ten sets of parental lines, each set consisted of Nok-3, Ga-0 and RIL NG32. Sets numbering 1 to 5 were spread throughout the bench where population (33) was grown and those numbering 6 to 10 with population (34). The orientation of sets 1 to 5 and 6 to 10 were the same, from the glasshouse entrance to the window side as seen in Fig.6-2 and 6-3 as shown in Section 6.2.2.No significant differences were observed between the locations (*P*=0.68) indicating there was not a significant environmental effect along the bench for the trait 18:2/18:1.

Plants which were A homozygous at both flanking markers or those which were B homozygous were kept as controls for phenotype comparison with the mapping population. Phenotypic variations amongst the controls are shown in Fig.6-6 below.

A more distinctive difference was observed between A homozygous type and B homozygous type in the population (33) than in the population (34).

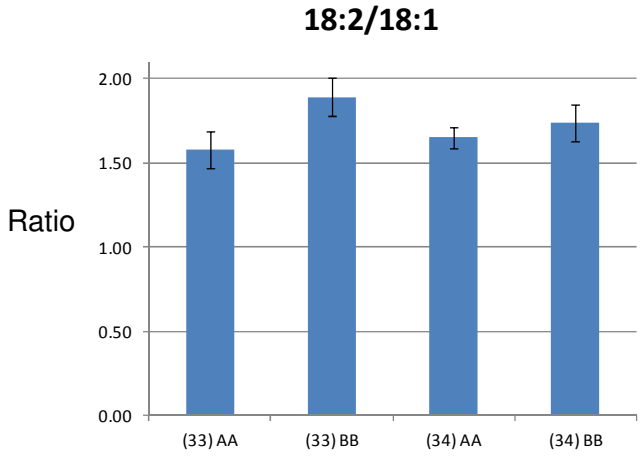


Fig.6-6 Average values of 18:2/18:1 for the control material of both mapping populations

Error bar represents standard deviations

6.3.3.2 STATISTICAL SUMMARY OF 18:2/18:1

The average values of 18:2/18:1 and the standard deviations of the 48 bins of the mapping populations and the five types of controls are shown in Table 6-6.

Controls consisted of maternal parents (MP), plants with homozygous A alleles at both flanking markers (AA), plants with homozygous B alleles at both flanking markers (BB), RIL NG32(RIL) and paternal parents (PP).

Of the controls, Ga-0, the paternal lines, showed unexpectedly low phenotype. According to the previous research, Ga-0 had higher ratio than that of Nok-3, 2.7 and 2.21, respectively (O'Neill et al., 2003).

In both populations, the average of bins within each type showed the same trend that Type 3 showed the highest average followed by Type 4, Type 1 and Type 2 being the lowest.

The ratios of 18:2/18:1 in the population(33) were approximately 0.04 to 0.06 lower than that of population(34).

Table 6-6 Average of 18:2/18:1 for each bin

18:2/18:1 (33)														
Bin	Type 1	SD	Bin	Type 2	SD	Bin	Type 3	SD	Bin	Type 4	SD	Parental lines		SD
1	1.62	0.09	1	1.61	0.10	1	1.71	0.04	1	1.74	0.08	MP	1.47	0.03
2	1.67	0.06	2	1.62	0.06	2	1.71	0.06	2	1.71	0.11	AA	1.58	0.11
3	1.58	0.06	3	1.69	0.08	3	1.73	0.07	3	1.73	0.03	BB	1.89	0.11
4	1.64	0.06	4	1.57	na	4	1.77	0.09	4	1.67	0.12	RIL	1.80	0.24
5	1.64	0.03	5	1.52	0.13	5	1.68	0.06	5	1.72	0.06	PP	1.51	0.13
6	1.64	0.06	6	1.64	0.03	6	1.75	0.09	6	1.76	0.05			
7	1.64	0.12	7	1.57	0.08	7	1.80	0.04	7	1.69	0.15			
8	1.66	0.04	8	1.56	0.07	8	1.74	0.05	8	1.72	0.13			
9	1.68	0.06	9	1.58	0.07	9	1.69	0.08	9	1.67	0.06			
10	1.67	0.09	10	1.55	0.08	10	1.80	0.03	10	1.62	0.05			
11	1.66	0.07	11	1.49	0.07	11	1.79	0.11	11	1.64	0.07			
12	1.69	0.10	12	1.55	0.06	12	1.83	0.11	12	1.65	0.06			
ave	1.65		ave	1.58		ave	1.75		ave	1.69				

18:2/18:1 (34)														
Bin	Type 1	SD	Bin	Type 2	SD	Bin	Type 3	SD	Bin	Type 4	SD	Parental lines		SD
1	1.69	0.06	1	1.79	na	1	1.76	0.12	1	1.83	0.11	MP	1.48	0.03
2	1.60	0.10	2	1.64	0.05	2	1.81	0.07	2	1.79	0.03	AA	1.65	0.06
3	1.64	0.03	3	1.68	0.08	3	1.78	0.07	3	1.75	0.09	BB	1.74	0.11
4	1.61	0.06	4	1.66	0.10	4	1.77	0.14	4	1.80	0.12	RIL	1.77	0.12
5	1.63	0.04	5	1.63	0.01	5	1.73	0.07	5	1.73	0.07	PP	1.49	0.04
6	1.64	0.05	6	1.61	0.05	6	1.79	0.14	6	1.74	0.09			
7	1.77	0.04	7	1.62	0.07	7	1.85	0.06	7	1.81	0.05			
8	1.77	0.07	8	1.65	0.04	8	1.79	0.08	8	1.71	0.08			
9	1.74	0.09	9	1.60	0.09	9	1.82	0.04	9	1.66	0.12			
10	1.77	0.04	10	1.60	0.03	10	1.82	0.05	10	1.73	0.07			
11	1.71	0.07	11	1.58	0.06	11	1.80	0.06	11	1.72	0.04			
12	1.72	0.08	12	1.58	0.08	12	1.86	0.08	12	1.70	0.08			
ave	1.69		ave	1.64		ave	1.80		ave	1.75				

MP=maternal parental line, AA=plants A homozygous at both flanking markers, BB= plants B homozygous at both flanking markers, RIL=RIL NG32, PP= paternal parental line

6.3.3.3 IDENTIFYING THE BIN CONTAINING THE QTL CAUSATIVE LOCUS

The binning system was developed so that the position of the causative locus of the target QTL could be detected as a shift difference of phenotype. If there was a single QTL within the interval, the shift should occur only once.

To identify shift differences between the bins, T-tests were carried out as described in Section 6.2.6. Significant differences ($P \leqq 0.05$) between a bin and a cluster of bins were observed and the points at which the differences occurred are shown in Figs.6-7 and 6-8 below as coloured vertical bars. The colour of the bar indicates the direction of analysis, hence bars were also positioned at the starting point of the analysis. Orange bars indicate analysis from bin 1 to bin 12 while green bars show the opposite direction from bin 12 to bin1.

The number of shift differences detected in populations (33) and (34) were 13 and 6, respectively.

However only a single shift was expected therefore results which showed multiple positions were excluded from further analysis.

The results of shift differences observed in both populations are summarised in Table 6-7.

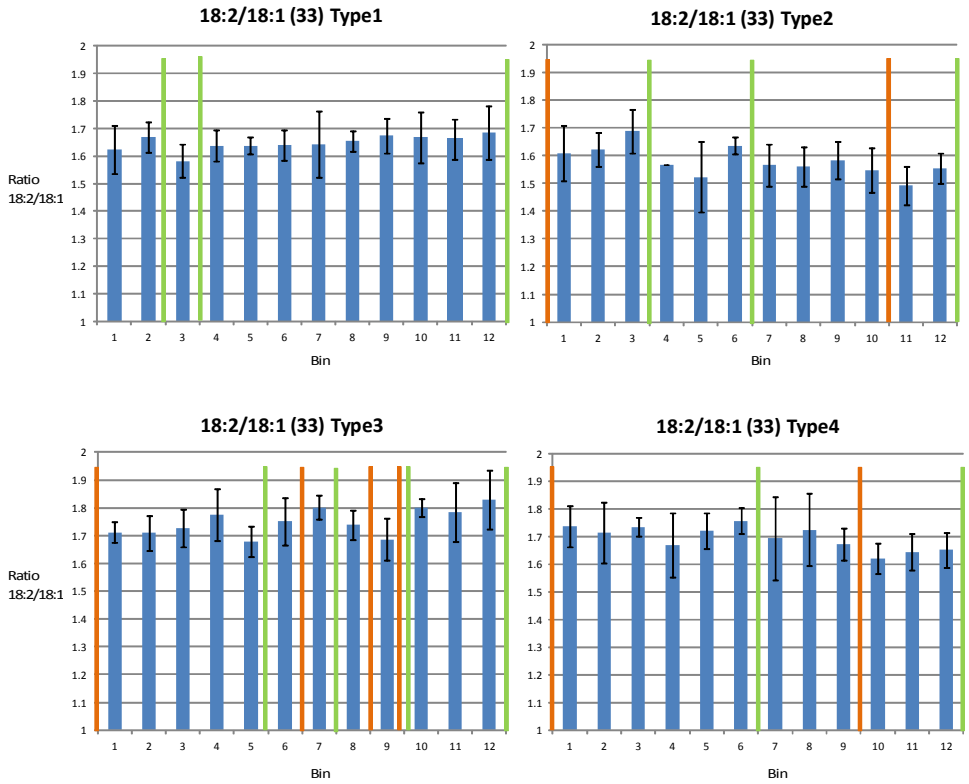


Fig.6-7 Identified shift differences between bins in population(33) for 18:2/18:1
Orange bars were analysed from bin 1 to bin 12and green bars from bin 12 to bin 1

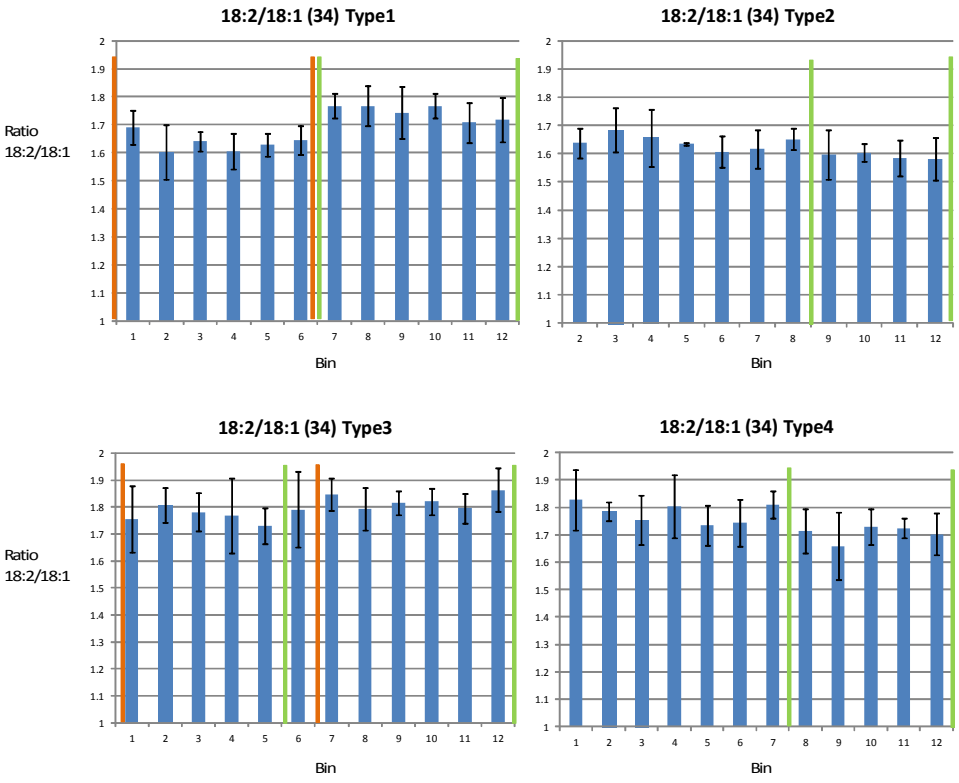


Fig.6-8 Identified shift differences between bins in population(34) for 18:2/18:1
Orange bars were analysed from bin 1 to bin 12 and green bars from bin 12 to bin 1

Table 6-7 Summary of identified shift differences

Population	Type	Position of shift difference	Indicated bin to contain QTL
(33)	2	10-11	10
	4	9-10	9
	4	6-7	6
(34)	1	6-7	7
	1	6-7	7
	2	8-9	8
	3	6-7	7
	3	5-6	6
	4	7-8	7

Population(33) showed three positions of single shift differences, at bin 10 in Type 2, and bin 9 and bin 6 in Type 4. As these were all different, the results were not informative.

The results of population(34) were clearer than population(33) because all four types showed only a single shift difference. The bins identified as possibly containing the causative locus of the QTL were bins 6, 7 and 8. Bin 7 was identified twice in Type 1 and once each in Types 3 and 4, so a total four times in the population. Therefore there is a high possibility that bin 7 contains the causative locus of the target QTL.

6.3.3.4 DISSECTING THE PHENOTYPE OF POPULATION(34)

Population(34) showed clearer and more consistent results than population(33), therefore this population was used to dissect the characteristics of phenotype of 18:2/18:1. FA contents of 18:1, 18:2 and 18:3 were studied to see how their values influenced the ratio of 18:2/18:1. T-tests were performed as for 18:2/18:1 to detect the positions where shift differences occurred.

The position of the most pronounced shift difference was found in Type 1 for 18:2/18:1 and was coincident with that found for 18:1 as seen in Fig.6-9 below. The shift difference position observed in Type 2 between bins 8 and 9 and also in Type 3 between bins 5 and 6 for 18:1 were also coincident with the results found for 18:2/18:1 in Fig. 6-8.

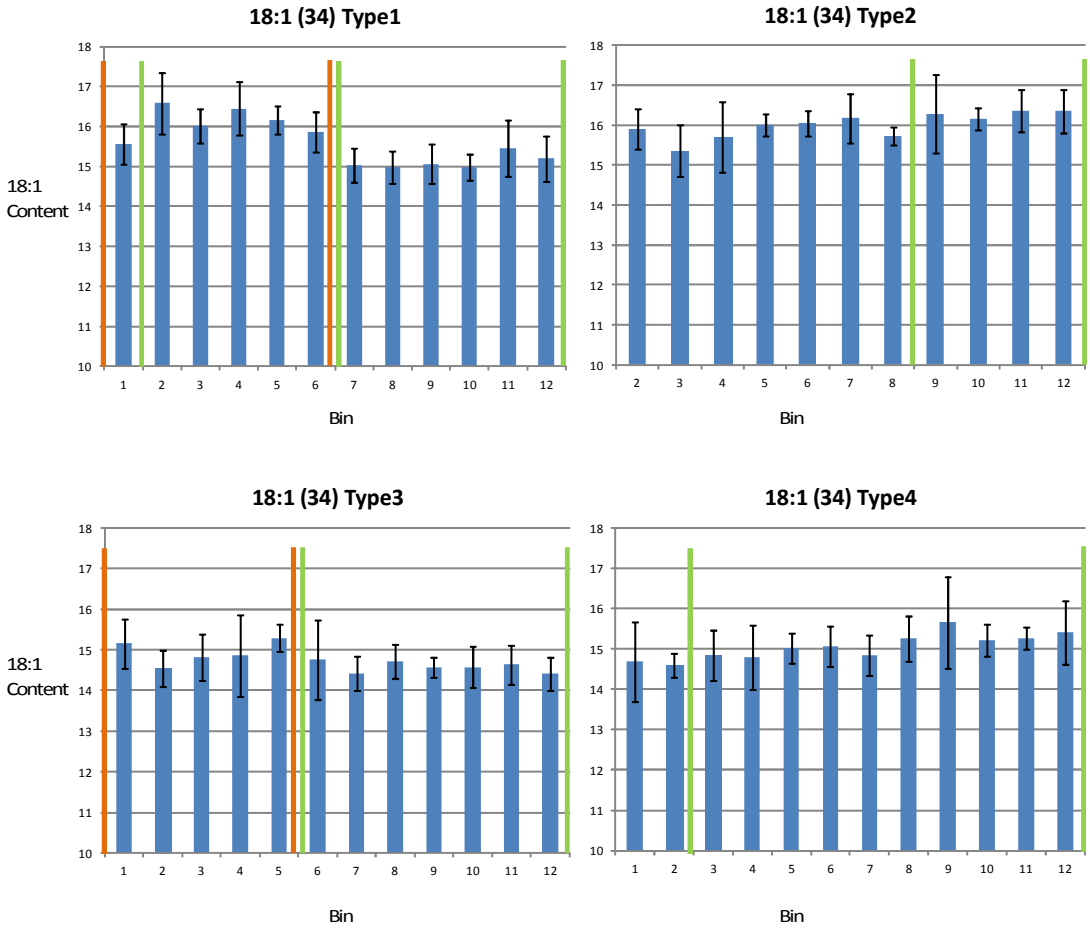


Fig.6-9 Identified shift differences between bins in population(34) for 18:1
Orange bars were analysed from bin 1 to bin 12 and green bars from bin 12 to bin 1

In 18:2 no significant shift differences were observed in Types 1, 2 and 3. Type 4 showed multiple positions, therefore no clear single shift difference could be observed in this trait.

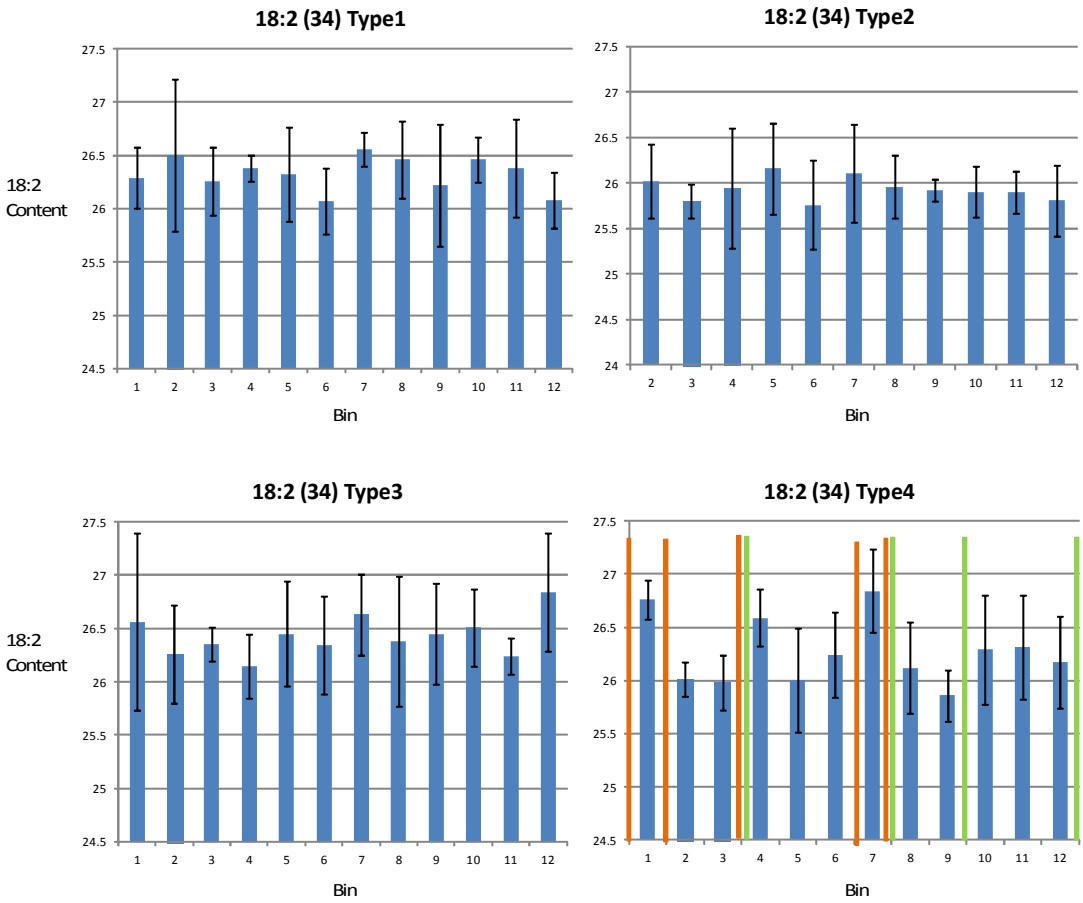


Fig.6-10 Identified shift differences between bins in population(34) for 18:2
Orange bars were analysed from bin 1 to bin 12 and green bars from bin 12 to bin

For trait 18:3, a single shift difference was observed in Type 1 between bins 6 and 7, Type 2 between bins 3 and 4 and Type 3 between bins 4 and 5 as seen in the chart below. The one in Type1 was coincident with that of 18:2/18:1.

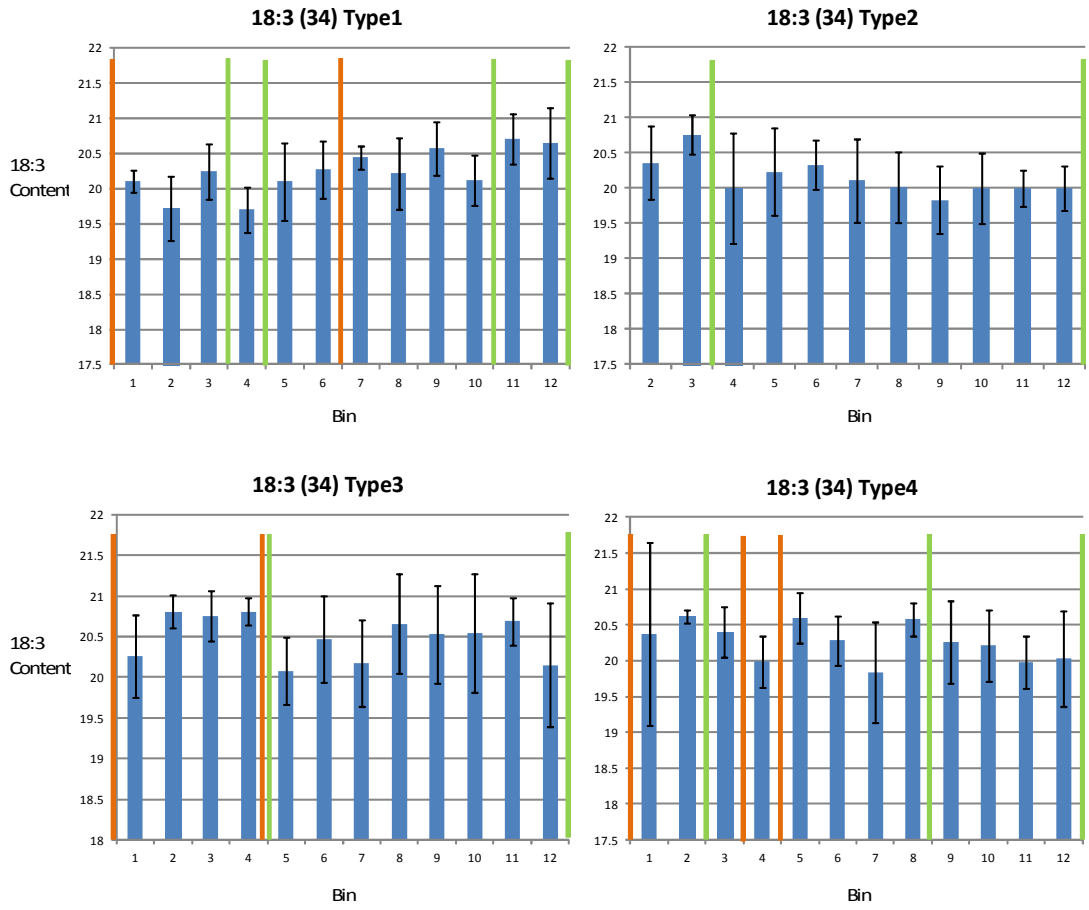


Fig.6-11 Identified shift differences between bins in population(34) for 18:3
Orange bars were analysed from bin 1 to bin 12 and green bars from bin 12 to bin 1

6.3.4 INVESTIGATION OF THE INTERACTION BETWEEN *FAD8*

After the region was narrowed down to the section between bin 6 to bin 8, a new question was posed. NIL NG28 was also grown for this QTL and although it showed a significant difference between controls statistically, the error bars of the NILs overlapped those of the controls, therefore the difference was judged insufficient to be nominated for fine mapping. It was therefore assumed that the causative locus of the QTL lay outside the introgressed segment of NG28. NIL NG32 was also grown for 18:2/18:1 and it had a much wider introgressed segment than NIL NG28 and with strong 18:2/18:1 phenotype, therefore it was chosen for fine mapping. However, the fine mapping results revealed that the causative locus of the QTL was not outside

the NG28 introgressed segment, on the contrary, NIL NG28 also contained the region between bins 6 to 8 in paternal homozygous form. This was the region identified as the strong possibility for the location of the causative locus. Therefore NIL NG28 should have had as strong a phenotype as NIL NG32.

This fact pointed to the possibility that another QTL could be present above the introgressed segment which had previously been viewed as the upper fringe of the target QTL interval and it could positively interact with the target QTL. There is a gap of 1.4 Mb between the end of chromosome 5 and the uppermost marker, nga225. Due to a lack of genetic markers within this gap, the presence of QTL at this end remains unclear. The possible candidate gene in this region is *Fatty Acid Desaturase 8 (FAD8)* which sits just 0.4Mb above the upper flanking marker. It encodes a temperature sensitive plastidic omega-3 fatty acid desaturase in the chloroplast (McConn et al., 1994). It was not included as an anticipated major gene, as the effect of this gene on seed oil profile is yet unknown and also it was not identified during the QTL analysis performed in Chapter 4 for 18:3, one of the omega-3 FAs found in *A. thaliana*.

In order to observe the effect of the upper flanking marker, the *FAD8* locus and the nga225 locus of the mapping population were genotyped. The results showed that these two loci were tightly linked together and with the upper flanking marker, having the same genotype without recombination between them.

ANOVA was performed for 18:1, 18:2, 18:3 and 18:2/18:1 to examine for significant phenotypic differences due to different genotypes at the upper flanking marker in the mapping population of all four recombination types. Due to the unequal numbers of samples in the bins, the averages for each bin were used for ANOVA.

Highly significant ($P<0.001$) phenotypic differences were found between the types of genotypes, such as paternal homozygous, heterozygous and maternal homozygous at the upper side of the flanking markers for all traits, except 18:3 ($P=0.011$) which is also significantly different, however not as significant as other traits. The effect of the lower side of the flanking markers was examined at the same time, and it did not show significant differences between genotypes.

The fine mapping performed in this study focused on the phenotype differences between bins within a recombination type where all of the plants shared the same genotype at the upper side of the flanking markers, therefore the results obtained here were independent from the effect observed above.

6.3.5 CANDIDATE GENE ANALYSIS

Following the results obtained in Section 6.3.3.3, candidate genes relating to seed lipid synthesis and FA desaturation were searched in The Arabidopsis Information Resource (TAIR) http://www.arabidopsis.org/ database. The region searched was on chromosome 5 between 5031708bp to 7063782bp which spanned bins 6, 7 and 8. The details of regions covered and the closest AGI (Arabidopsis Genome Initiative) ID at the edge of each bin are shown in the Table 6-9.

For the region of interest seven protein coding genes were identified in the TAIR database.

AT5G16230 at 5303kb contains a conserved site for the PLANT STEAROYL-ACP DESATURRASE FAMILY located in the chloroplast. AT5G16240 at 5306kb may have a similar function to AT5G16230, however the location is unknown and it is expressed widely throughout the plant.

Both AT5G15530 and AT5G16390 at 5038kb and 5360kb respectively encode BIOTIN CARBOXYL-CARRIER SUBUNIT OF THE MULTI-ENZYME PLASTIDIAL ACETYL-CoA CARBOXYLASE COMPLEX (BCCP1 and BCCP2, respectively).

The gene AT5G17400 at 5728kbp is predicted to encode an ENDOPLASMIC RETICULUM-ADENINE NUCLEOTIDE TRANSPORTER1 (ER-ANT1).

AT5G18210 at 6017kb has a domain common with the NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN found in FA synthesis.

And finally AT5G19940 at 6738kb contains a conserved domain found in a number of PLASTID LIPID ASSOCIATED PROTEINS (PAPs).

Table 6-8 Regions covered by bins 6,7and 8

	Region of chromosome 5 (bp)		
Bin 6	5031708 (AT5G15500)	to	5708080 (AT5G17330)
Bin 7	5708080 (AT5G17330)	to	6361746 (AT5G19040)
Bin 8	6361746 (AT5G19040)	to	7063782 (AT5G20840)

6.4 CONCLUSION

6.4.1 POPULATION(33) AND (34)

The NIL candidates NG32 NIL(33) and NG32 NIL(34) had different merits in their phenotypic characteristics, therefore a population was developed from each for fine mapping purposes. NIL(33) had an 18:2/18:1 ratio unexpectedly higher than those of the other NIL candidates and parental lines. However at the same time, it was recognised that this could be due to non-genetic factors such as environmental effects. NIL(34) did not have as strong a phenotype as NIL(33), however it had the expected phenotype for the NIL, based on that observed in the original RIL.

The phenotypic result of fine mapping revealed that population(33) lines, derived from NG32 NIL(33), showed weaker phenotypes than population(34) lines with differences between 0.04 to 0.06 in 18:2/18:1 values being observed. This was the opposite of what was expected based on the original NILs phenotypes. Hence the strong phenotype displayed by NG32 NIL(33) was not reproducible in a mapping population. No block effects were observed during the NIL development in NG population for any FA contents (Fig. 5-4), therefore the basis of the high phenotype observed in NIL(33) was difficult to explain. It might be due to a place change of environment during the rotation process that occurred at the right time and affected the desaturation of seed lipid.

6.4.2 FINE MAPPING

To produce the mapping populations NILs were backcrossed and selfed to allow recombination to occur within the interval between two flanking markers. Differences in recombination pattern were used to sort plants into four recombination types. The plants were then divided into 12 bins within each type according to recombination position. This binning system was to enable the detection of the position which influenced 18:2/18:1. Pinpointing the positions causing shift differences was carried out by a succession of comparisons between a bin and its adjacent group of bins, for example plants in bin 1 was compared with plants in bin 2 using a T-test with a significance level $P \leqq 0.05$. If there were no significant difference between them, plants in bin 1 and plants in bin 2 were considered to be the same and they were combined. Plants in bins 1 and 2 were compared with plants in bin 3 and if no significant difference was detected by the T-test, plants in bins 1, 2 and 3 were combined and compared with those of bin 4. This process continued until a difference was observed. Once the difference was observed for example between

the group of bins 1 to 6 and bin 7, combining bins stopped and the process was restarted by comparison of bin 7 and bin 8 and was continued until up to bin 12. Once it was reached bin 12, the comparison was started in the reverse order from bin 12 to bin 1. Since a single QTL was anticipated within the interval, the difference should occur once in the form of a shift difference.

Multiple comparisons using T-test could accumulate statistical errors which could lead to analysis type I error. The possibility of this risk can be calculated using the formula of experiment-wise error rate (EER) (Petrinovich, 1969).

When L comparisons are made and result in incorrect decisions (M),

$EER=Pr(M \geqq 1)= 1-(1-\alpha)^{L}$

Where α is the per-comparison error rate and in this study it was 0.05.

In this work six comparisons were required to identify a shift difference between bin6 and bin7, therefore the result EER=0.26, means there is a 26% possibility that one or more false positives could have occurred during the six comparisons (Iglewicz, 2005). Therefore the experiment was designed to divide the mapping populations according to the genotypes of their flanking markers ie four types as seen in Fig.6-5. If these four types showed a consistent result, then there would be an increased possibility that the result was true.

Population(33) had a number of multiple shift differences and three single shift differences. These single differences were all identified at different bins, therefore the results were not used for further analysis.

By comparison, population(34) exhibited clearer results than population(33) showing a single shift difference in all recombination types. The bins that affected the phenotype of population(34) were bins 6, 7 and 8, especially bin 7 which was identified a total of four times in this population. Therefore there was a high possibility that bin 7 could contain the target locus, even taking into account the risk of EER.

After the QTL was narrowed down to the region between bin6 to bin8, an interaction between the causative locus and *FAD8* was considered. *FAD8* is known to encode a temperature sensitive plastidic omega-3 fatty acid desaturase in the chloroplast (McConn et al., 1994). Because the QTL analysis for 18:3 and 18:2 in NG

population, as carried out in Chapter 4, did not detect any segregated QTL at the *FAD8* locus, this locus was not treated as cautiously as the major gene loci. However, the results of fine mapping suggested the possibility of an interaction between the QTL causative locus and a locus on the upper side of the QTL. *FAD8* is found in this region of the genome and it is a likely candidate as an interaction partner. Although confirming this interaction requires further detailed experiments

6.4.3 CAUSATIVE OF QTL

In order to determine whether the effect of the target QTL was dominant, recombination patterns were investigated. As seen in Fig.6-4, loci in bins Types 1 and 2 showed either heterozygous or maternal homozygous, while bins in Types 3 and 4 were either heterozygous or paternal homozygous. Therefore if the effect was dominant, Types 3 and 4 would not show the shift difference as they might already affected by the QTL. The results showed that Types 3 and 4 had a shift difference where it was coincident with the position of Type 1. Therefore the effect of the QTL was concluded to be non-dominant and it instead showed an additive effect as Types 3 and 4 showed higher 18:2/18:1 values than Types 1 and 2 as seen in Table 6-5.

To understand the characteristics of the trait 18:2/18:1, other traits closely related to 18:2/18:1 in the desaturation PUFA pathway were examined. Such traits were 18:1, 18:2 and 18:3 contents. 18:3 was included as 18:2 is a precursor of 18:3 therefore these two traits could influence each other. The positions of four of the six phase differences identified in 18:2/18:1 were coincident with those of 18:1. 18:2 did not show clear phase differences in all types. 18:3 showed four single phase differences, one these was between bin 6 and bin 7 in Type1 which was coincident with those of 18:2/18:1 and 18:1. Therefore, the effect of 18:1 on the ratio of 18:2/18:1 was stronger than on both 18:2 and 18:3. It followed the results of QTL analysis performed in the Chapter 4. The NG population had highly significant QTL for 18:1 and 18:2/18:1 at the same locus with LOD scores of 12.48 and 11.82 respectively, however no QTL was detected for neither 18:2 nor 18:3. Therefore it could be confirmed that the narrowed-down region most likely contains the same causative of QTL with the original QTL.

6.4.4 CANDIDATE GENES

Candidate genes were searched in the TAIR database between bin 6 to bin 8 where it was most likely that the target locus lay. Seven protein coding genes related to lipid biosynthesis were identified and discussed below.

AT5G16230 and AT5G16240 in bin 6 matched the protein coding region of the *PLANT STEAROYL-ACP DESATURASE FAMILY PROTEIN* which encode protein catalyzes the introduction of a double bond at the delta-9 position of stearoyl-ACP to create oleoyl-ACP, hence responsible for the desaturation between 18:0 to 18:1 (Shanklin and Somerville, 1991; Kachroo et al., 2007). It was observed in this study that 18:1 content strongly affected the target 18:2/18:1 phenotype, as both of these candidate genes are involved in determining 18:1 composition they fit well with the phenotypic results.

AT5G15530 and AT5G16390 in bin 6 matched genes which encode BIOTIN CARBOXYL-CARRIER SUBUNIT OF THE MULTI-ENZYME PLASTIDIAL ACETYL-CoA CARBOXYLASE COMPLEX1 and 2, respectively (BCCP1 and BCCP2 (Li et al., 2005; Mingjie et al., 2008). Biotin carboxylase activates CO_2 by attaching it to biotin in an ATP-dependent reaction. Biotin carboxyl carrier protein and carboxyltransferase transfer activated CO_2 from the biotin carboxylase region to acetyl-CoA to produce malonyl-CoA which is the first step of long chain FA synthesis (Ohlrogge and Browse, 1995). The FAs studied here were long chain FAs, therefore this may affect their composition.

AT5G18210 in bin 7 contains an NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN domain which is involved in oxidation and reduction in the metabolic process and is found in a number of FA synthases which catalyse the first step in the reductive modification of the beta-carbonyl centres in the growing carbon chain using NADPH to reduce 3-ketobutyryl-ACP to 3-hydroxybutyryl-ACP during the saturated FA biosynthesis in the chloroplast (Ohlrogge and Browse, 1995). This may affect the oil content but the influence on the ratio of 18:2/18:1 is unknown.

AT5G17400 in bin 7 is predicted to encode an ENDOPLASMIC RETICULUM-ADENINE NUCLEOTIDE TRANSPORTER1 (ER-ANT1) which appears to function as an ATP:ADP antiporter in *E. coli* . The fact that knock-outs of this gene contain reduced ATP requiring ER proteins suggests that it acts as an ATP transporter into the ER. Formation of malonyl-CoA from acetyl-CoA and CO_2 is ATP dependent, therefore the lack of this gene could lead to the reduction of oil content. In fact the

seeds of the mutants are smaller and seed oil content is reduced. Also the mutants have reduced apical dominance and the life cycle is prolonged (Lerocha et al., 2008). However the effect on the ratio of 18:2/18:1 is not yet known.

AT5G19940 in bin 8 contains a conserved domain found in a number of PLASTID LIPID ASSOCIATED PROTEINS (PAPs) which may participate in transportation of lipids through thylakoid membranes both inward and outward (Ting et al., 1998; Hernandez-Pinzon et al., 1999). FA biosynthesis occurs in the plastid and the lipids are transported to the ER for modification and assembly therefore transporters are fundamental for the seed lipid biosynthesis process.

From the fine mapping results, the bin most likely to contain the causative gene was bin 7, however bin 6 contained both AT5G16230 and AT5G16240 which are involved in the desaturation of 18:0 to 18:1, hence both directly relate to the trait studied in this project. More detailed fine mapping is required to determine a single candidate gene, followed by functional analysis and complimentary testing.

The positions of these candidate genes are shown in Table 6-9. Both SNP and InDel markers identified in the parental genomic sequences are sufficiently frequent to greatly facilitate further fine mapping.

Table 6-9 AGI gene code and the physical positions of candidate genes

Bin	AGI code	Region (bp)	
6	AT5G15530	5038700	5040900
6	AT5G16230	5303300	5306100
6	AT5G16240	5306800	5310000
6	AT5G16390	5360800	5363100
7	AT5G17400	5728800	5730400
7	AT5G18210	6017800	6020000
8	AT5G19940	6739600	6740900

CHAPTER 7. COMPARATIVE ANALYSIS OF QTL FOR SEED LIPID TRAITS IN *A. THALIANA* AND *B. NAPUS*

7.1 INTRODUCTION

B. napus is a recently formed amphidiploid derived from the hybridisation of *B. rapa* and *B. oleracea* known as the A and C genomes respectively (U, 1935), which can be regarded as homoeologous. Thus the sources of polymorphisms shown by molecular markers could be both from allelic variation and homoeologous polymorphism and it could be very difficult to distinguish them in a molecular marker analysis (Bancroft et al., 2011).

Moreover prior to the amphidiploidisation event, genome triplication occurred in their common progenitor ancestor, therefore *B. napus* is a polyploid which potentially contains six copies of each gene in its genome (O'Neill and Bancroft, 2000; Lysak et al., 2005).

A. thaliana, on the other hand, is a diploid species and normally contains single copies of genes in its genome, in addition it shares an ancestor with *B. napus* prior to the genome triplication and hybridization events as seen in Fig.1-3 in Chapter 1 Section 1.2.1 (Yang et al., 1999). Therefore it is an ideal model plant for the study of *Brassica* species.

A genome-wide comparative analysis between *B. napus* and the model plant *A. thaliana* was performed by Parkin using a linkage map of *B. napus* which consisted of over 1,000 linked RFLP loci that were mapped to homologous positions in the *A. thaliana* genome based on sequence similarity (Parkin et al., 2005). The study identified 21 conserved blocks with *A. thaliana* which cover almost 90% of the mapped length of *B. napus* including replications and rearrangements (Parkin et al., 2005).

To take the study further, computational analysis, next generation sequencing technology and genomic sequences of *B. rapa* (Wang et al., 2011) and *B. oleracea* (http://brassicadb.org/brad/) which were conducted by multinational consortiums were exploited for detailed comparative genomics (Trick et al., 2009; Trick et al., 2009; Bancroft et al., 2011; Harper et al., 2012).

The *B. napus* doubled haploid mapping population TNDH was developed from two *B. napus* cultivars ‘Tapidor’, a European winter oilseed rape and ‘Ningyou 7’, a

Chinese cultivar of semi-winter oilseed rape with the aim of having maximum number of polymorphisms in the population (Qiu et al., 2006). Transcriptomes of 37 Tapidor x Ningyou 7 DH (TNDH) lines were sequenced using Solexa sequencing and approximately 20 million reads were obtained for each line. Prior to this study, a set of approximately 94,000 *Brassica* unigenes were assembled from 810,000 expressed sequence tags (ESTs) sequences publically available from GenBank (Trick et al., 2009). Transcriptome sequences were aligned to this *Brassica* unigene set and scoring strings were obtained which contained SNPs derived from the two parental lines. Using these scoring strings, a linkage map with 23037 SNP markers over 9100 unigenes was constructed. All of these unigenes show consistent sequence similarity with genome sequence scaffolds derived from the two progenitor species *B. rapa* and *B. oleracea*. (Bancroft et al., 2011). This SNP linkage map was used to improve the order and orientation of genome sequence scaffolds. Chimeric scaffolds were also identified in the genomic sequences of progenitors therefore they were split into sections (Harper et al., 2012).

Finally improved genome scaffolds were assembled into “pseudomolecules” which represent the 19 chromosomes of the *B. napus* genome. Subsequently all unigenes were blasted to align the pseudomolecules to produce a hypothetically ordered unigene based map of the *B. napus* genome. Unigenes were blasted *to A. thaliana* Genome Initiative (AGI) gene models, so that each unigene was associated with an AGI gene code. The map was coloured coded to the five *A. thaliana* chromosomes or left blank (white) if unigenes did not match an AGI gene model as seen in Fig.7-1 (Harper et al., 2012*).* The process of the development of this unigene based *B. napus* map is explained as a flow chart in Fig. 7-2.

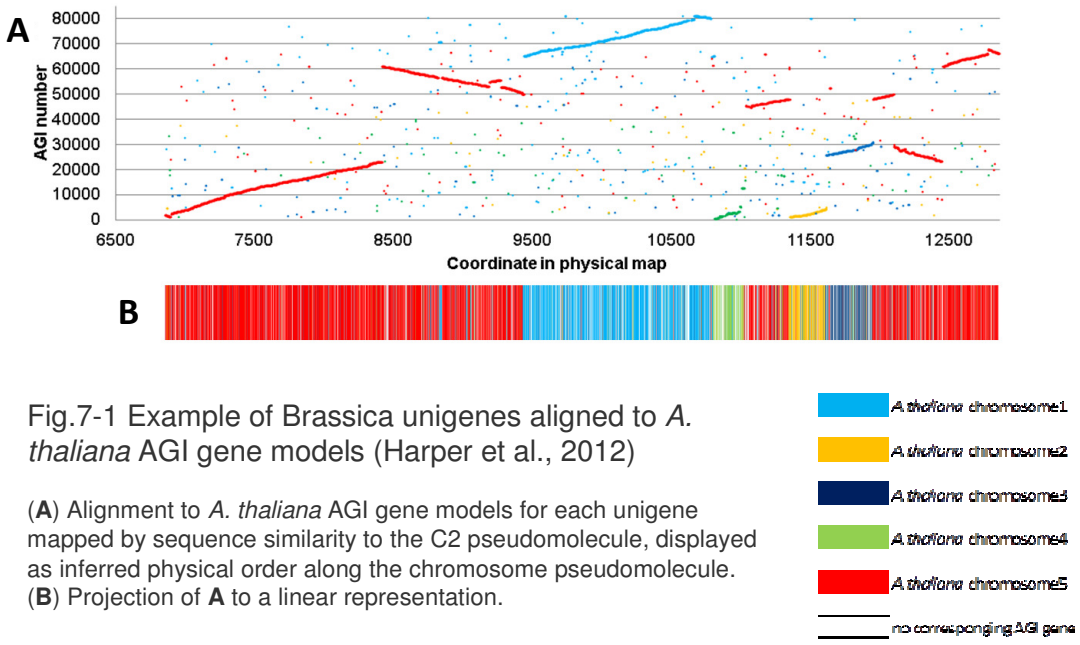


Fig.7-1 Example of Brassica unigenes aligned to *A. thaliana* AGI gene models (Harper et al., 2012)

(**A**) Alignment to *A. thaliana* AGI gene models for each unigene mapped by sequence similarity to the C2 pseudomolecule, displayed as inferred physical order along the chromosome pseudomolecule. (**B**) Projection of **A** to a linear representation.

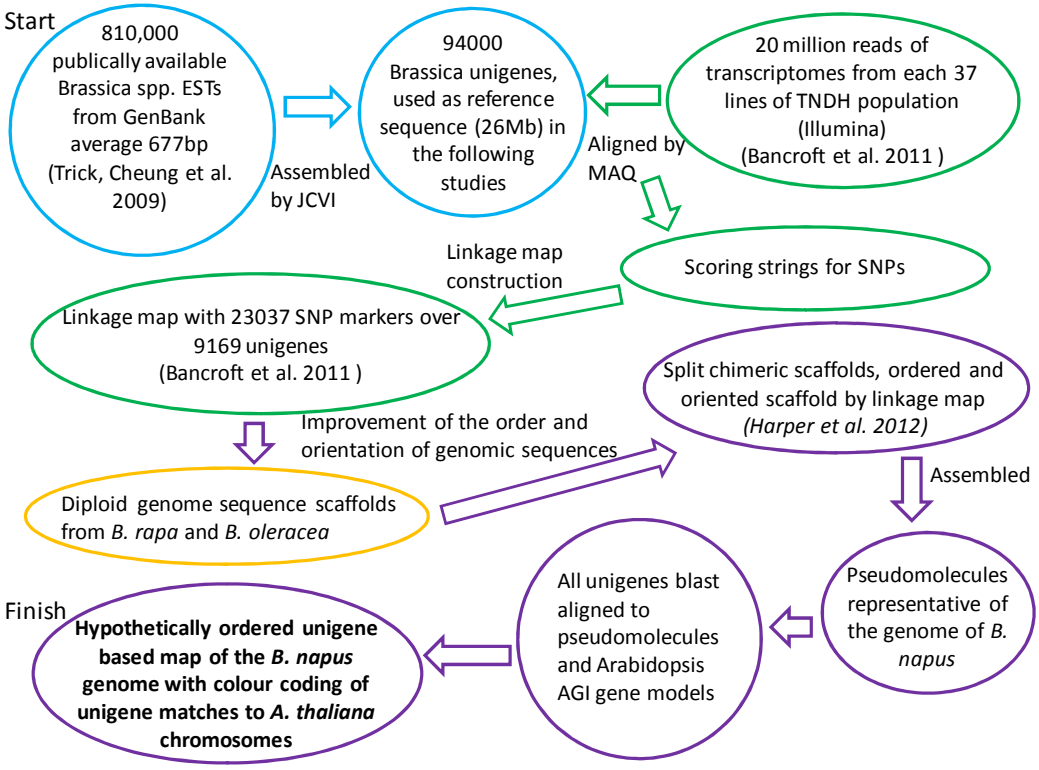


Fig.7-2 Flow chart showing the development procedures of the unigene based map of *B. napus* (pseudomolecules map) constructed from (Trick et al., 2009; Bancroft et al., 2011; Harper et al., 2012)

The potential usefulness of this informative map in comparative genomics of *Brassica* crops is enormous. It facilitates candidate gene searches of *B. napus* using an informative *A. thaliana* database. It also enables the mapping of QTL detected in *A. thaliana* onto the unigene based map of *B. napus* giving direct interpretation of the model plant result in the crop plant *B. napus.*

There is a limitation in that the linkage map used to correct the order and orientation of the genomic sequences is based on the TNDH population, and is therefore population specific. However it is still possible to compare QTL identified in different populations if they can be associated with AGI gene codes so that both can be mapped onto this pseudomolecule map.

Fig. 7-3 shows how the QTL detected in *A. thaliana* and *B. napus* were associated with the unigene based *B. napus* map in the following studies. Once QTL positions were translated onto the unigene based *B. napus* map, comparisons between them became possible.

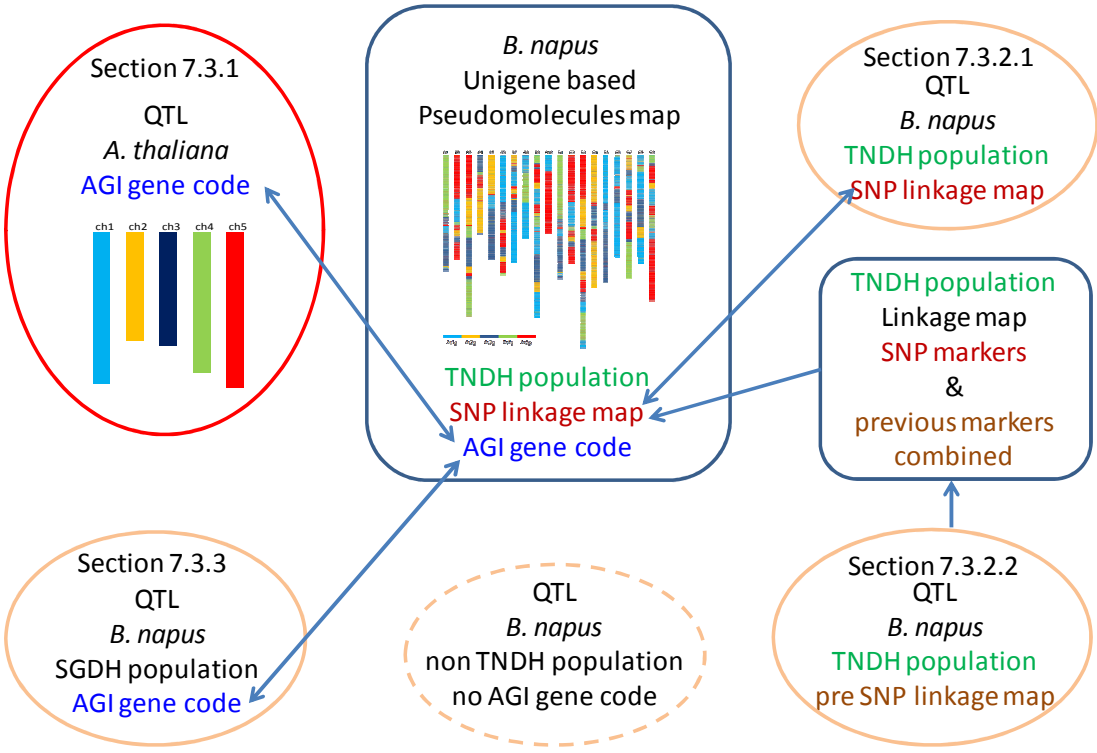


Fig.7-3 Schematic image of the association between QTL identified in *A. thaliana* (circled in red) and QTL detected in *B. napus* (circled in pale orange) with the unigene based *B. napus* map

In Chapter 5, NILs were developed for selected QTL which were identified in Chapter 4. NILs contained paternal alleles at the introgressed segments while most of the background loci were maternal in origin. The NILs were phenotyped and the differences between NILs and the controls (which had maternal alleles at the introgressed segment) were assessed by ANOVA. Of the 14 NILs developed for 13 original QTL, 12 NILs for 11 QTL showed highly significant differences (P<0.001) and a further selection was carried out to identify more distinctive differences from the controls for the use of fine mapping. Seven NILs for seven QTL were shortlisted as seen on Table 7-1 and finally the NIL developed for 18:2/18:1 in the NG population was followed for fine mapping in Chapter 6.

For this chapter, these seven QTL shortlisted for fine mapping were used for comparative mapping (Table 7-1).

Table 7-1 Details of *A thaliana* QTL used for comparative mapping

Trait	Linkage group	Position (Mb)	LOD	* %Expl	Population	Selected RIL no.
18:2	3	7.7	4.8	20.9	TJ	25
18:2/18:1	5	3.8	11.82	44.0	NG	32
18:3	1	10.9	6.32	26.6	TJ	17
18:3/18:2	1	22.57	2.97	13.6	NG	93
Oil	1	27	3.52	15.9	TJ	83
Oil	2	9.9	3.9	17.4	TJ	79
Oil	5	22.1	3.2	14.9	KB	55

* %Expl: the percentage of the QTL accounting for the phenotypic variation in the population

Firstly, the *A. thaliana* QTL were mapped onto the unigene based *B. napus* map and compared with the QTL identified in the *B. napus* TNDH population within the Bancroft group.

Also an attempt was made to compare these *A. thaliana* QTL with those *B. napus* QTL found in previous reports in the literature in order to increase the possibility of finding association between them.

7.2 MATERIAL AND METHODS

7.2.1 MAPPING *A. THALIANA* QTL ONTO THE *B. NAPUS* UNIGENE BASED MAP

QTL identified in *A. thaliana* in Chapter 5 were correlated with AGI gene codes for the region spanned in order to identify the corresponding positions in *B. napus* genome.

A QTL confidence interval is normally a 2 LOD interval of its maximum LOD score(Van Ooijen, 2004). However, according to the fine mapping carried out in Chapter 6, the bins that most influenced the phenotype were away from the locus with the highest LOD score at QTL analysis, that is from 3.8Mb to the bins spanning 5.03Mb to 7.06Mb. Therefore the region of ±3Mb from the locus with the highest LOD score was regarded as the confidence interval in this study.

Two methods of mapping were applied. The first method used all the AGI gene codes within the confidence interval of 6Mb, which consisted of approximately 1500 genes. Each gene could potentially be mapped onto six orthologous regions in the *B. napus* genome, therefore subsequent selection of appropriate positions of *A. thaliana* QTL in the *B. napus* genome became necessary. Hence, a second mapping round was introduced; where genes relating to lipid biosynthesis and FA desaturation were selected and mapped onto the *B. napus* unigene based map. Alignment of AGI gene codes to the unigene based *B. napus* map was carried out using R based software 'Mapper v.3' developed by Antony Colles. Mapper v.3 highlights corresponding positions on the pseudomolecules using AGI gene codes, unigenes and SNP markers.

Results from both alignment methods were used to identify segments more likely to represent the *A. thaliana* QTL.

7.2.2 COMPARISON WITH THE *B. NAPUS* TNDH POPULATION

Prior to this project, field trials of *B. napus* TNDH population were carried out within the Bancroft group at Barton and Cowlinge in UK 2006-2007 and FA composition was measured.

Using the phenotypic data of 75 lines from the field trial, newly developed SNP linkage maps consisting of 23000 SNP markers (Bancroft et al., 2011), and

genotype data scored from transcriptome sequences, QTL mapping was performed by Colin Morgan using single marker regression analysis within Genstat (unpublished data). QTL detected for 18:1, 18:2 and 18:3 were mapped onto the unigene based *B. napus* map using Mapper v.3, as it used the same linkage map as the unigene map, direct positioning was possible. The resulting QTL locations were compared with *A. thaliana* QTL aligned on the unigene based *B. napus* map.

A report in the literature details QTL analysis for oil content in the TNDH population using 188 lines grown in Dali and Wuhan, China, over two years (Qiu et al., 2006). The linkage map used for the QTL analysis consisted of 352 markers, which were a mixture of InDel, SNP, RFLP,SSR, and AFLP markers(Qiu et al., 2006).

To compare the QTL detected in Qiu's study with *A. thaliana* QTL, an integrated linkage map was constructed by Colin Morgan using JoinMap3.0 by combining markers from this QTL analysis with the high density SNP markers used for the development of unigene based *B. napus* map using a genotype file for a common subset of TNDH lines. This integrated linkage map facilitated the estimation of relative positions of these InDel, SNP, RFLP, SSR and AFLP markers on the unigene based *B. napus* map.

7.2.3 COMPARISON WITH PREVIOUS WORKS IDENTIFIED IN THE LITERATURE

Previous studies that conducted QTL analyses for FA composition or oil contents in *B. napus* were searched for in the literature and the QTL positions were mapped onto the unigene based *B. napus* map. Where AGI gene codes cannot be inferred for a *B. napus* QTL, the direct alignment of the QTL to the unigene based *B. napus* map is very difficult, therefore only studies which associated their identified QTL with AGI codes or could indicate close proximity to AGI codes were used.

7.3 RESULTS

In order to compare QTL found in *A. thaliana* with those of *B. napus*, linear alignment of *A. thaliana* QTL onto the *B. napus* unigene based map was performed using Mapper v.3. Subsequently comparisons were made between these and QTL detected in the TNDH population (by Morgan within the group).

The unigene based *B. napus* map shown in the following sections represents the 19 chromosomes of the *B. napus* genome, with A1 to A10 being derived from the A genome and C1 to C9 being derived from the C genome. The colour coding matches *Brassica* unigenes to their corresponding *A. thaliana* chromosome. Unigenes without a corresponding AGI gene are in white (Harper et al., 2012).

7.3.1 ALIGNMENT OF *A THALIANA* QTL WITH THE UNIGENE BASED *B. NAPUS* MAP

Two methods of alignment were used to identify the better way of representing *A. thaliana* QTL on the *B. napus* unigene based map. The first method used all of the AGI genes within the confidence interval and the second method used only those AGI genes which related to lipid synthesis and FA desaturation.

Table 7-2 List of AGI codes related to seed lipid biosynthesis and FA desaturation for seven QTL detected in *A. thaliana*

Trait		18:2	18:2/18:1	18:3	18:3/18:2	Oil	Oil	Oil
Chromosome		3	5	1	1	1	2	5
Highest LOD score position (Mb)		7.7	3.8	10.9	22.57	27	9.9	22.1
Confidence interval	Upper side (Mb)	4.7	5.03-7.06 (fine mapped)	7.9	19.57	24	6.9	19.1
	AGI code	AT3G14140		AT1G22470.	AT1G52360	AT1G64470	AT2G15830	AT5G47030
	Lower side (Mb)	10.7		13.9	25.57	30	12.9	25.1
	AGI code	AT3G28610		AT1G35190.	AT1G68180	AT1G79780	AT2G30230	AT5G62600
Number of AGI genes within the interval		1448	442	1273	1583	1532	1440	1564
AGI code related to seed lipid synthesis and FA desaturation		AT3G15850.1	AT5G15530.1	AT1G24360.1	AT1G62640.1	AT1G64400.1	AT2G16280.1	AT5G52920.1
		AT3G15870.1	AT5G16230.1	AT1G30120.1	AT1G64400.1	AT1G65290.1	AT2G19450.1	AT5G59310.1
		AT3G22960.1	AT5G16240.1	AT1G32440.1	AT1G65290.1	AT1G67730.1	AT2G22230.1	
		AT3G24650.1	AT5G16390.1		AT1G67730.1	AT1G68530.1	AT2G26640.1	
		AT3G25110.1	AT5G17400.1			AT1G76680.1	AT2G28630.1	
		AT3G25440.1	AT5G18210.1			AT1G77590.1	AT2G29980.1	
		AT3G26790.1	AT5G19940.1			AT1G79750.1	AT2G30200.1	

7.3.1.1 ALIGNMENT OF ALL AGI GENES FOUND WITHIN THE CONFIDENCE INTERVALS

Mapper v.3 enabled alignment of large numbers of genes. Approximately 1200 to 1500 genes were mapped onto the unigene based *B. napus* map. For the trait 18:2/18:1, the interval was narrowed down to 2Mb in Chapter 6, so there were only 442 genes to align as seen on Table 7-2.

The resulting maps (one example is shown in Fig.7-5 while the remainder can be found in Appendix 10) identified several positions where many AGI genes, which were indicated as short horizontal bars along the unigene map, were densely mapped forming clusters. It was assumed that the positions of these clusters represent the main parts of the target *A. thaliana* QTL. There were many other AGI genes that mapped thinly scattered across the unigene based *B. napus* map. Considering that the AGI gene models corresponded to a unigene based on the greatest sequence similarity (Harper et al., 2012), they could collectively equate to the *A. thaliana* QTL. However they could also be noise derived from non-functional segments with similar sequences resulting from factors such as transposons or unpredictable artefacts of the methods used. Additional information was required to allow clearer identification of which parts more likely represented the *A. thaliana* QTL, therefore more specific mapping was carried out in the next section.

7.3.1.2 ALIGNMENT OF AGI GENES RELATED TO LIPID SYNTHESIS AND FATTY ACID DESATURATION PATHWAY

Following the mapping carried out in the previous Section 7.3.1.1, AGI genes related to lipid biosynthesis and the FA desaturation pathway which lay within the confidence interval of the *A. thaliana* QTL were extracted from the TAIR database as seen on Table 7-2 and aligned to the unigene based *B. napus* map. This was done to identify which clusters or single matches better represented the *A. thaliana* QTL. One example of this mapping is shown in Fig.7-4. The square highlights the position of each alignment and the horizontal bar in the middle of the square indicates the exact position of the selected lipid related AGI genes. The complete results are presented in Appendix 10.

The positions which matched in both alignment methods were considered to more likely represent the *A. thaliana* QTL on the *B. napus* map.

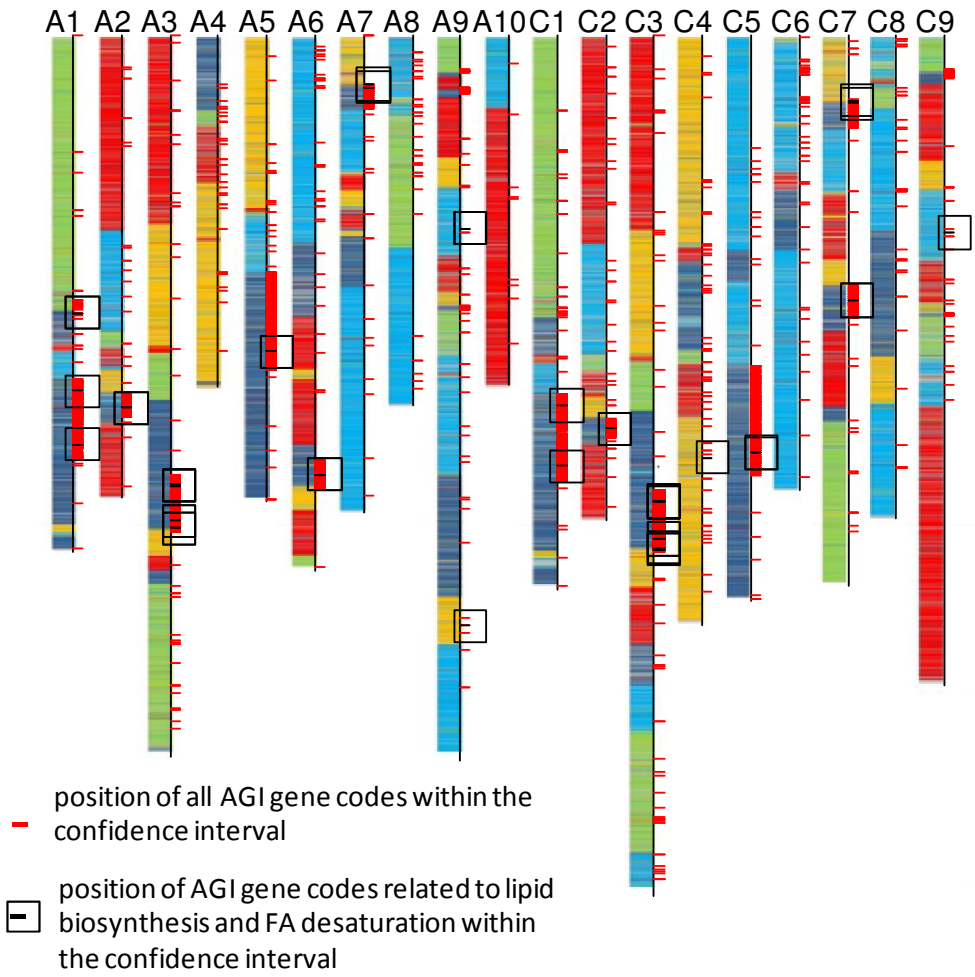


Fig.7-4 Example of an *A. thaliana* QTL (18:2) mapped onto the unigene based *B. napus* map

7.3.2 COMPARISON OF *A. THALIANA* QTL WITH QTL DETECTED IN THE *B. NAPUS* TNDH POPULATION

There are two sets of TNDH QTL mapping data available, however different linkage maps were used in each. Therefore they are explained separately in the following sections.

7.3.2.1 QTL ANALYSIS FOR 18:1, 18:2, 18:3 AS CARRIED OUT BY COLIN MORGAN (UNPUBLISHED RESULTS)

QTL detected by Morgan from 75 TNDH lines using 23000 SNP markers with phenotype data taken from Barton and Cowlinge in UK (2006 and 2007) were mapped onto the unigene based *B. napus* map as seen in Figs 7-5-a to 7-5-d. As the linkage map used in the QTL analysis was the same as the one used during the construction of the unigene based *B. napus* map, the direct alignment of QTL onto

the *B. napus* map was possible. The locations of markers with LOD scores above 2 were indicated by small horizontal bars along the unigene based *B. napus* map in Fig.7-5-a to Fig.7-5-c.

Three *A. thaliana* QTL matched *B. napus* QTL: C3 for 18:2/18:2, C3 for 18:2 and C4 for 18:3/18:2 as indicated with yellow arrows in Fig.7-5-a to Fig.7-5-c.

A QTL for 18:1 on C3 in *B. napus* was coincident with an *A. thaliana* QTL for 18:2/18:1 as seen in Fig.7-5-a. Two AGI genes which encode BIOTIN CARBOXYL CARRIER PROTEIN 1 (BCCP-1) and the PLASTID-LIPID ASSOCIATED PROTEIN (PAP) / FIBRILLING FAMILY PROTEIN were in the region common to both.

A QTL for 18:2 on C3 in *B. napus* matched with a QTL found in *A. thaliana* for 18:2 as seen in Fig.7-5-b.The two protein coding genes: *PLASTIDIAL PYRUVATE KINASE1 (PKP1)* and *FatA ACYL-ACP THIOESTERASE (AtFaTA) are* lying in this region.

Regarding 18:3/18:2, a gene encoding the MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2) in *A. thaliana* maps was within a *B. napus* QTL for 18:3 on C4 (Fig.7-5-c).

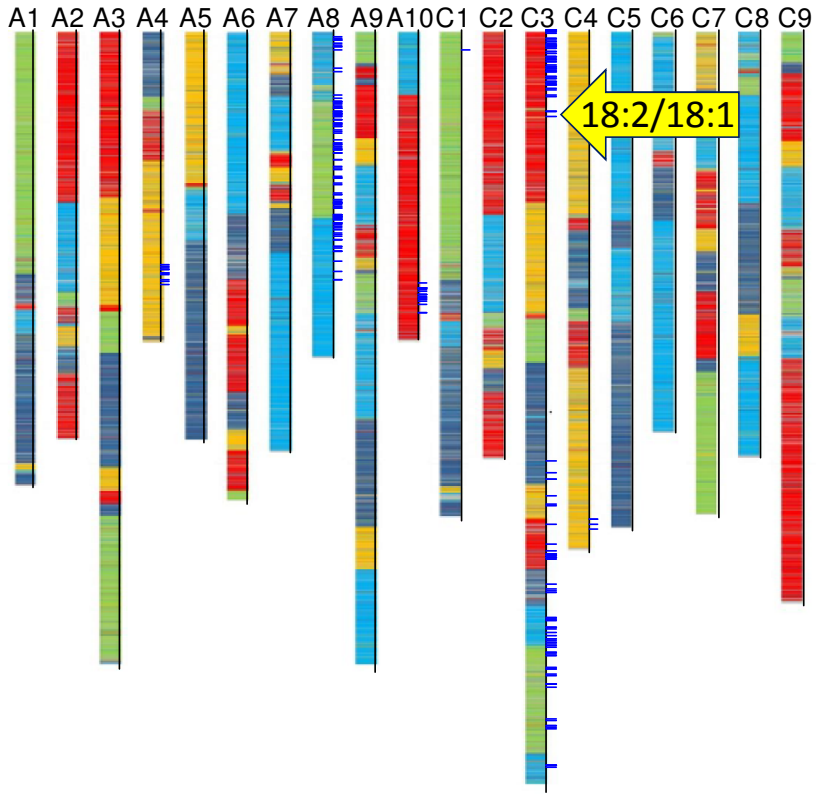


Fig.7-5-a QTL for 18:1 in the *B. napus* TNDH population shown as blue bars (markers with LOD>2.0) Coincident QTL with *A. thaliana* are indicated by a yellow arrow

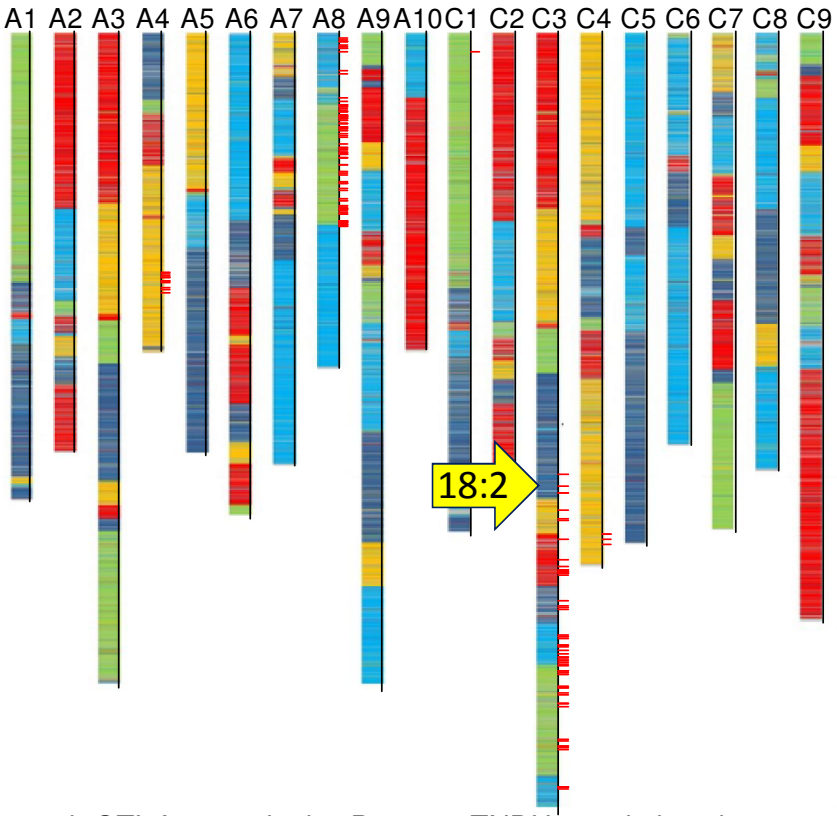


Fig.7-5-b QTL for 18:2 in the *B. napus* TNDH population shown as red bars(markers with LOD>2.0) Coincident QTL with *A. thaliana* are indicated by a yellow arrow

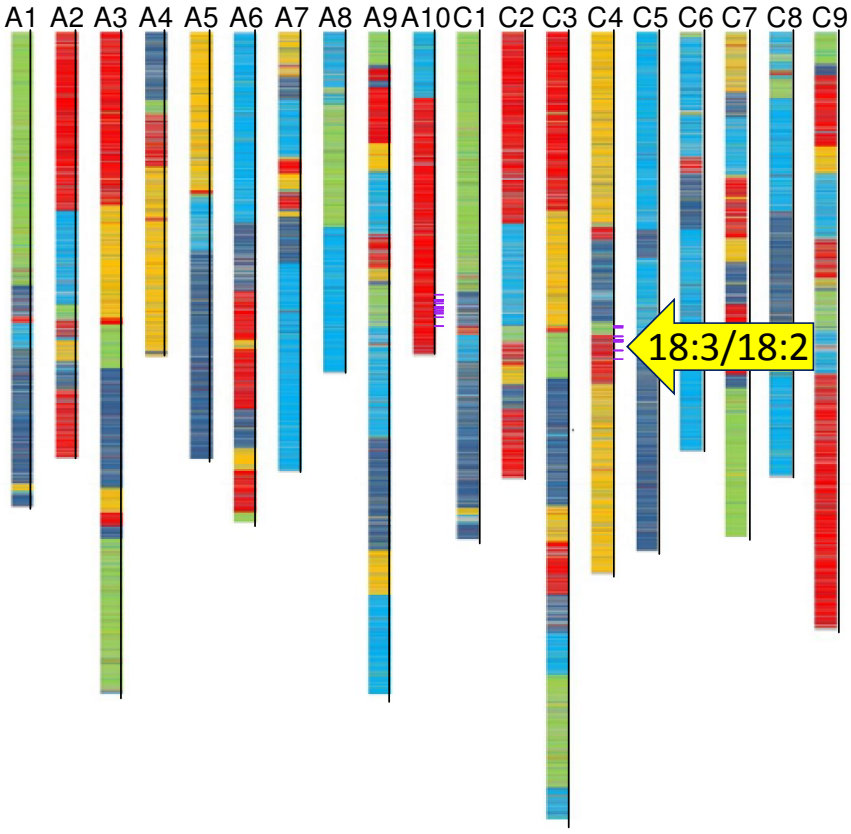


Fig.7-5-c QTL for 18:3 in the *B. napus* TNDH population shown as purple bars (markers with LOD>2.0) Coincident QTL with *A. thaliana* are indicated by a yellow arrow

7.3.2.2 QTL ANALYSIS FOR OIL CONTENT BY QIU ET AL. 2006

Nine QTL for oil content were identified in the study by Qiu et al 2006 using 188 TNDH lines grown in Dali and Wuhan, China, in 2003 and 2004. As the linkage map used for this study consisted of 352 markers which were different to those of the unigene based *B. napus* map, direct alignment was not possible, however a common subset of the TNDH population was used in both studies making integration of the two linkage maps feasible. This work was carried out by Morgan. The integrated linkage map enabled the positions of markers used in this study to be estimated. The positions of QTL showing significance ($P<0.05$) were indicated as dark blue vertical bars in Fig.7-6.

Four *B. napus* QTL matched *A. thaliana* QTL on A4, A9, A10 and C2.

A. thaliana genes related to lipid synthesis found in the coincident interval between *B napus* QTL and *A. thaliana* QTL are listed below.

On the A4 linkage group, there were two genes encoding the MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2) and BETA-HYDROXYACYL-ACYL CARRIER PROTEIN DEHYDRATA (FabA /FabZ).

On the A9 linkage group a gene encoding the ACYL-CoA:DIACYLGLYCEROL ACYLTRANSFERASE 1 (DGAT1) was found.

On the A10 linkage group, a gene encoding a DOMINANT CHLOROPLAST PYRUVATE KINASE BETA SUBUNIT (PKP-BETA1) was identified.

On the C2 linkage group, four genes were found: *PKP-BETA1*, which was also found on A10; *LIPID TRANSFER PROTEIN4, (LTP4)*; *MTACP2,* which was also found on A4 and the *A. thaliana BETA-KETOACYL REDUCTASE (AtKCR1).*

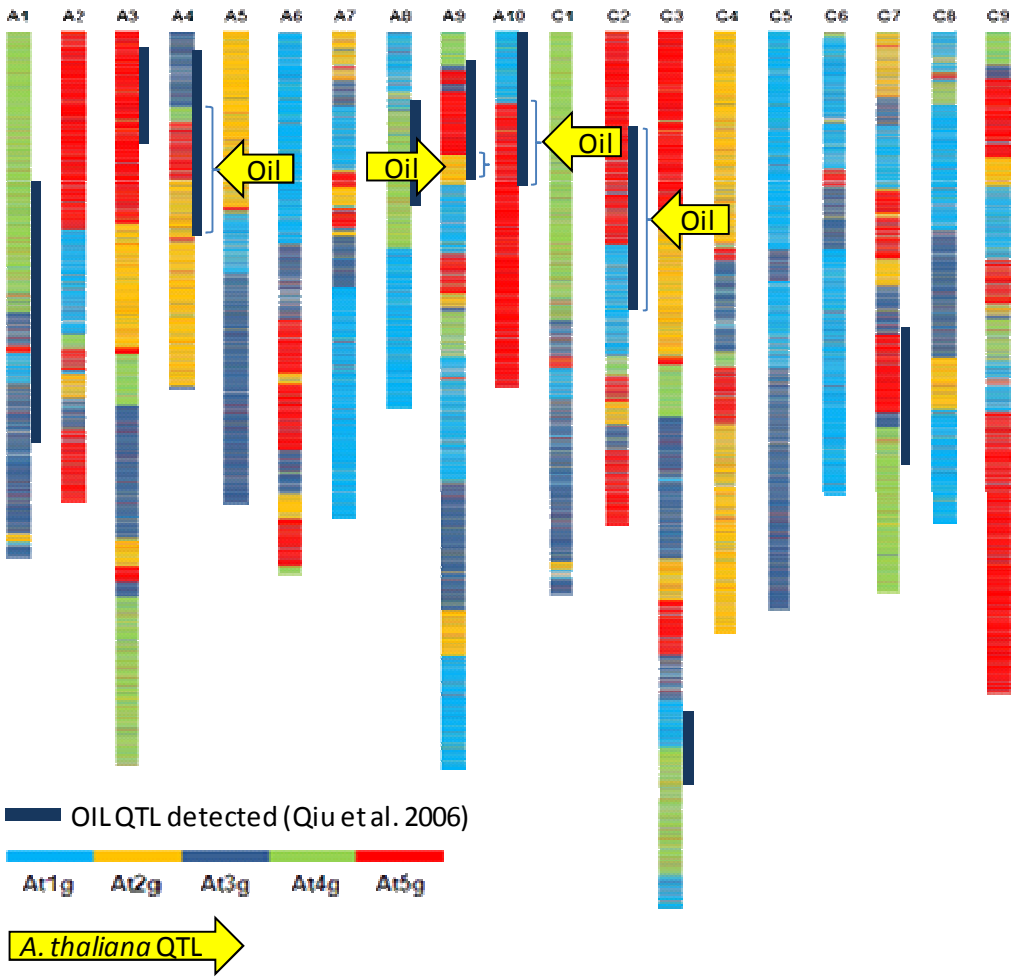


Fig.7-6 QTL for oil in the *B. napus* TNDH population (P<0.05)

Coincident QTL with *A. thaliana* QTL are indicated by a yellow arrow

7.3.3 COMPARISON OF *A. THALIANA* QTL WITH *B. NAPUS* QTL MAPPED IN DIFFERENT POPULATIONS

7.3.3.1 QTL FOR OIL CONTENT USING THE SGDH POPULATION DEVELOPED BY ZHAO ET AL. 2012

QTL analysis was carried out for oil content on a segregating DH population derived from 'Sollux' x 'Gaoyou' (SGDH population) using 282 lines and an SSR-based linkage map (Zhao et al., 2012). In total nine QTL for oil content were detected, one each on A1, C2,C3, A5, C6, A7 and A9 and two on C8. Homologous loci to 249 orthologous *A. thaliana* genes were identified throughout linkage groups for candidate gene analysis (Zhao et al., 2012).

The QTL detected in the SGDH population were aligned to the unigene based *B. napus* map (based on the TNDH population) using AGI gene codes associated with the QTL.

As *A. thaliana* genes could be mapped to multiple loci in the polyploid *B. napus* genome, only those regions and linkage groups where the original QTL were found were selected. Despite the different population, most of the QTL detected in the SGDH population mapped to similar positions on the map developed from the TNDH population, with the exceptions of the two QTL on C8.

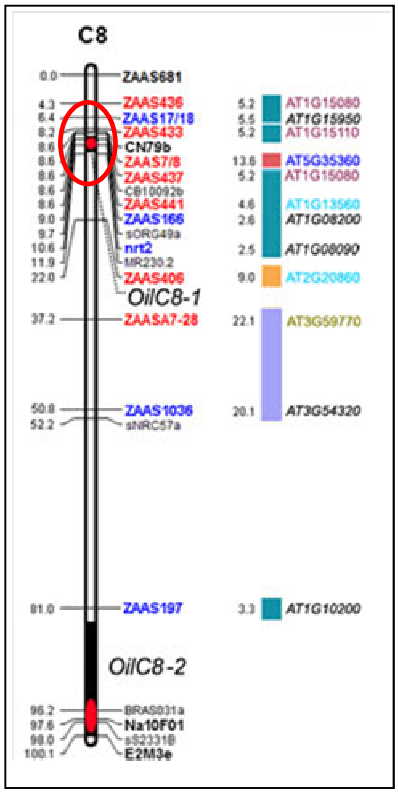


Fig.7-7 Two QTL on C8 in SGDH population as taken from (Zhao et al., 2012)

Homologous *Arabidopsis* genes of the QTL mapped on the upper arm of C8 in the SGDH population (circled in red on Fig.7-7) appeared on the lower arm of C8 in the unigene derived *B. napus* map as seen in Fig.7-8. The *Arabidopsis* gene order within the region was conserved in the unigene based B. napus map chart.

A QTL mapped on the bottom of C8 in the SGDH population had just a single AGI code, AT1G10200, corresponding to a locus near the QTL but it was not found in the TNDH population, therefore its position could not be compared in this work.

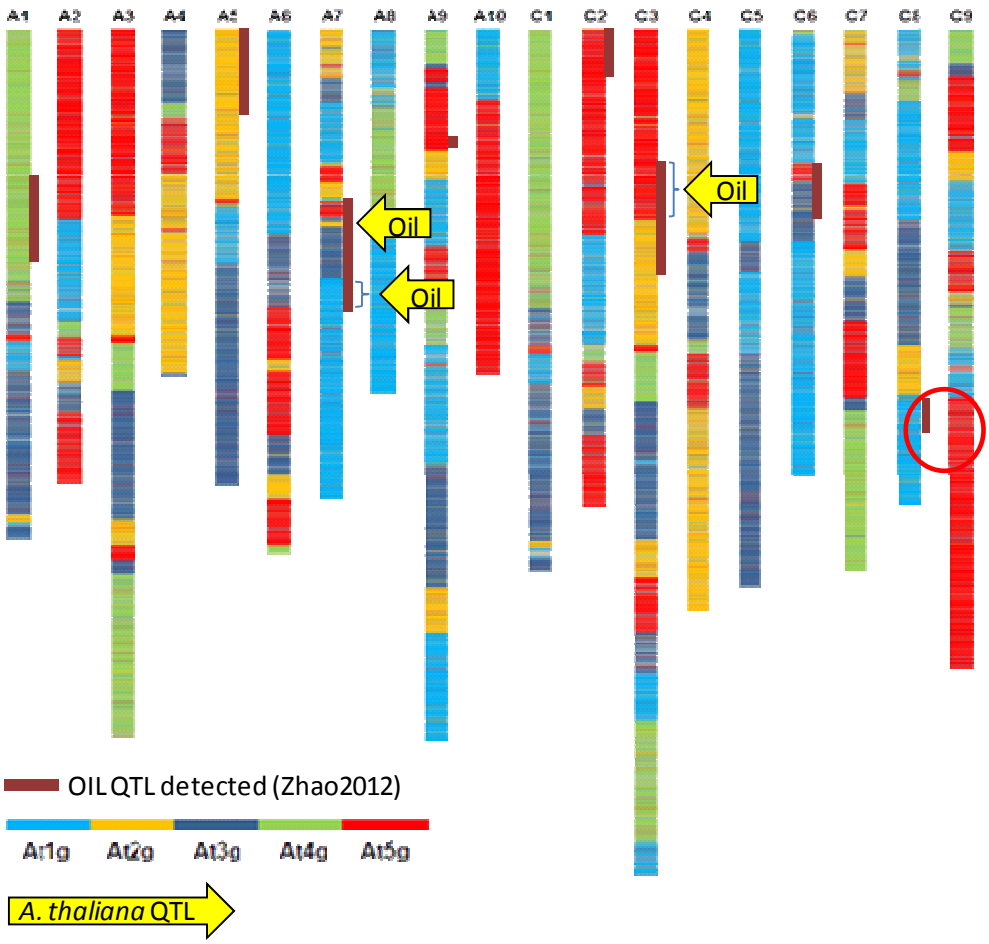


Fig.7-8 QTL detected in the SGDH population mapped onto the unigene based *B. napus* map constructed from the TNDH population. The red circle shows a position mis-match between the original SGDH map shown in Fig.7-7.

QTL detected in the SGDH population were mapped onto the unigene based *B. napus* map and they were compared with *A. thaliana* QTL for oil content aligned to the same *B. napus* map (Fig.7-8). Three QTL were identified to be coincident, two loci on A7 and a locus on C3.

Within the interval in common between *B napus* QTL and *A. thaliana* QTL on A7, three *A. thaliana* genes were found which encode DGAT1, LONG CHAIN ACYL-COA SYNTHETASE9 (LACS9) and *A thaliana* NADP-MALIC ENZYME4(ATNADP-ME4).

On C3, three *A. thaliana* genes which encode Lipid Transfer Protein 4 (LTP4), FAD3 and EMB3147 lay within the common QTL region.

All the listed AGI gene models found in the QTL regions coincident between A. thaliana and B. napus were summarized in Table 7-3.

Table 7-3 Summary of *Arabidopsis* gene model found within coincident region of *B. napus* and *A. thaliana* QTL

B. napus QTL		*A. thaliana QTL*	Lipid related genes coincident in *B. napus* and *A. thaliana*
Trait	Linkage group	Trait	
18:1	C3	18:2/18:1	*BIOTIN CARBOXYL CARRIER PROTEIN (BCCP)1*
			PLASTID-LIPID ASSOCIATED PROTEIN(PAP) / FIBRILLING FAMILY PROTEIN
18:2	C3	18:2	*PLASTIDIAL PYRUVATE KINASE1 (PKP1)*
			FatA ACYL-ACP THIOESTERASE (AtFaTA)
18:3	C4	18:3/18:2	*MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2)*
Oil (Qiu2006)	A4	Oil	*MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2)*
			BETA-HYDROXYACYL-ACYL CARRIER PROTEIN DEHYDRATA (FabA /FabZ)
	A9		*ACYL-CoA:DIACYLGLYCEROL ACYLTRANSFERASE 1 (DGAT1)*
	A10		*CHLOROPLAST PYRUVATE KINASE BETA SUBUNIT (PKP-BETA1)*
	C2		*LIPID TRANSFER PROTEIN4, (LTP4)*
			PKP-BETA1
			MTACP2
			BETA-KETOACYL REDUCTASE (AtKCR1)
Oil (Zhao2012)	A7	Oil	*DGAT1*
			LONG CHAIN ACYL-CoA SYNTHETASE (LACS9)
			NADP-MALIC ENZYME4 (ATNADP-ME4)
	C3		*LTP4*
			FAD3
			EMB3147

7.4 CONCLUSION

In order to exploit the model plant *A. thaliana* for the improvement of crop plants, in this case *B. napus*, comparison of QTL identified in the two species was carried out.

7.4.1 ALIGNMENT OF QTL

Four different types of alignment were performed. Firstly, *A. thaliana* QTL were mapped onto the unigene based *B. napus* map using AGI gene codes. Secondly, the results of *B. napus* QTL, generated with the same population and linkage map as used for the unigene map, were aligned on that map using the SNP markers as indicator. Although using the same linkage map, visual alignment of QTL on the unigene map requires special software, such as “Mapper v.3” to execute due to the large number of markers used, approximately 23,000. In both the first and second types, direct alignment to the unigene map was possible. The third type involved the use of *B. napus* TNDH QTL identified when an earlier linkage map was used. In this third alignment type, integration of the two linkage maps was carried out within the Bancroft group therefore estimation of the positions of the markers used in the earlier linkage map became possible by using the closest SNP markers as references. In the fourth type the QTL had been identified in the *B. napus* SGDH population therefore the linkage map was completely different. However, marker used for the QTL were associated with *A. thaliana* gene models, therefore by using these AGI codes, alignment of QTL detected in the different populations became feasible.

7.4.2 COMPARATIVE MAPPING

Once the QTL results were mapped on the unigene based *B. napus* map, the coincident intervals were identified and those regions were searched for candidate genes.

In the alignment type one, three *Arabidopsis* QTL for FA composition such as 18:2/18:1, 18:2 and 18:3/18:2 were found to match *Brassica* QTL, these were on: C3 for 18:1; C3 for 18:2 and C4 for 18:3, respectively in the TNDH population when the high density SNP linkage map had been used.

Lipid related genes lying within the coincident region between *B. napus and A. thaliana* were examined (summary is shown on Table 7-3). Part of the *B. napus* QTL for18:1 on C3 was coincident with an *A. thaliana* QTL for 18:2/18:1 and both spanned genes which encode the BIOTIN CARBOXYL CARRIER PROTEIN

1(BCCP1) and a PLASTID-LIPID ASSOCIATED PROTEIN (PAP) / FIBRILLING FAMILY PROTEIN. *BCCP1* is involved in the transfer of carboxyl groups to acetyl-CoA to form malonyl-CoA which is a carbon donor during growth of the FA chain and the first step of long chain FA synthesis (Ohlrogge and Browse, 1995; Feria Bourrellier et al., 2010). *PAP* is predicted to be involved in lipid transfer to and from thylakoid membranes (TAIR). These two genes are involved in FA biosynthesis not desaturation.

A *B. napus* QTL for 18:2 on C3 matched an *Arabidopsis* QTL for 18:2. The coincident region contained AGI gene models of two protein coding genes: *PLASTIDIAL PYRUVATE KINASE1 (PKP1)and FatA ACYL-ACP THIOESTERASE (AtFaTA)*. Pyruvate kinase (PK) catalyses the irreversible synthesis of pyruvate and ATP, both of which are used in multiple biochemical pathways. They are essential compounds for the sustainable production of fatty acids in the plastids of maturing *Arabidopsis* embryos (Baud et al., 2007). Acyl ACP thioesterases control carbon chain length through the termination of FA synthesis, by hydrolysis of the acyl chains from ACP. After termination, free fatty acids are activated to CoA esters, exported from the plastid, and assembled into glycerolipids in the ER (Thelen and Ohlrogge, 2002). Among the different thioesterase gene families found in plants, the *FatA*-type performs a fundamental role in the export of the C18 FA moieties that will be used for the synthesis of most plant glycerolipids. There are two copies of this gene in *A. thaliana* and a double mutant showed reduced triacylglycerol content, but no change of the other neutral lipids such as diacylglycerols. The proportions of 18:3 and 22:1 were increased in the seed oil (Moreno-Pérez et al., 2012). In *E. coli*, the strain carrying *AtFaTA* showed 68 % increase of total FA level compared to the control strains (Cao et al., 2010). Therefore this gene could be involved in both oil content and FA composition of seed lipid.

On C4 a *B. napus* QTL for 18:3 matched an *A. thaliana* QTL for18:3/18:2. The region in common contained a gene encoding MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2). The Acyl Carrier Protein (ACP) is a small (9 kilodalton) acidic protein involved in acyl transfer steps. It contains a phosphopantethein prosthetic group to which the growing acyl chain is attached as a thioester and enters saturated fatty acid biosynthesis (Ohlrogge and Browse, 1995). Therefore it may be involved in the oil content, however the effect on desaturation is not yet known.

In conclusion, At*FaTA* which lies underneath the QTL for 18:2 is more likely to affect FA desaturation traits by controlling the composition of 18C FAs and longer chain FAs.

When considering oil content a total of seven *B. napus* QTL were coincident with *A. thaliana* QTL.

Four of the *B. napus* QTL for oil content indentified by Qiu et al. in the TNDH population were coincident with *A. thaliana* QTL.

Part of a *B. napus* QTL on A4 which matched an *A. thaliana* QTL contained two genes, one encodes MITOCHONDRIAL ACYL CARRIER PROTEIN2 (MTACP2) which was found also in QTL for 18:3 on C4 and another genes contains a domain of BETA-HYDROXYACYL-ACYL CARRIER PROTEIN DEHYDRATA (FabA /FabZ). In *E. coli*, FabZ catalyzed dehydration of short chain β-hydroxyacyl-ACP and long chain saturated and unsaturated β-hydroxyacyl-ACP. *FabA* was observed most active on intermediate chain length β-hydroxyacyl-ACP and also had specificity toward both short and long chain saturated β-hydroxyacyl-ACP. Notably, *FabA* was almost inactive in the dehydration of long chain unsaturated β-hydroxyacyl-ACP. However the introduction of the double bond at any chain length of fatty acid synthesis by *FabA* or *FabZ* has not yet proved (Heath and Rock, 1996) *MTACP2* is likely to affect oil content while *FabA* might affect the PUFA composition.

A. thaliana QTL was coincident with the lower part of the *B. napus* QTL on A9 underneath an *Arabidopsis* gene encoding ACYL-COA:DIACYLGLYCEROL ACYLTRANSFERASE 1 (DGAT1) was found. DGAT1 catalyses the final step of triacylglycerol synthesis from diacylglycerol and acyl-CoA (Hobbs et al., 1999).

The lower half of the TNDH QTL on A10 for oil content overlapped an *A. thaliana* QTL, underneath a gene encoding a DOMINANT CHLOROPLAST PYRUVATE KINASE BETA SUBUNIT (PKP-BETA1) is present. PK is a glycolytic enzyme that converts phosphoenolpyruvate into pyruvate, a precursor of fatty acid synthesis. Reduced expression of this gene resulted in reduction of seed oil (Ruuska et al., 2002).

Almost the entire *B. napus* QTL on C2 for oil content was coincident with an *A. thaliana* QTL and four lipid related genes were found in that region. Those were *LIPID TRANSFER PROTEIN4, (LTP4), PKP-BETA1, MTACP2 and A. thaliana BETA-KETOACYL REDUCTASE (AtKCR1). LTP4* encodes a member of the lipid transfer protein family. The proteins can bind FAs and acyl-CoA esters and can

transfer several different phospholipids. They are localized to the cell wall and the mRNA is present in flowers and siliques (Zachowski et al., 1998; Arondel et al., 2000). *PKP-BETA1*and *MTACP2* were described above for A4 and A10, respectively. AtKCR1 most likely catalyses the first reduction of the beta- ketoacyl-CoA during very long chain FA synthesis (Beaudoin et al., 2009).

Comparison between QTL for oil content in both *A. thaliana* and those in the *B. napus* SGDH population identified three that were coincident, two on A7 and one on C3. In the coincident QTL regions on A7, in total three *A. thaliana* genes which encode DGAT1, LONG CHAIN ACYL-COA SYNTHETASE9 (LACS9) and *A. thaliana* NADP-MALIC ENZYME4(ATNADP-ME4) were found. *LACS9* encodes chloroplast long-chain acyl-CoA synthetase which catalyzes the formation of long chain fatty acyl-CoAs from free fatty acid, ATP, and CoA (Schnurr J. A. et al., 2002). The biological roles of NADP-ME for lipid biosynthesis were suggested that it could provide carbon skeletons and reducing power during the process (Smith et al., 1992; Eastmond et al., 1997).On C3 three lipids related *A. thaliana* genes were found in the coincident region of which were *LTP4*, *FAD3* and *EMB3147*. *EMB3147* is predicted to function as ACP-S-Malonyltransferase located in chloroplast according to TAIR.

In the oil content analysis, some genes were picked up repeatedly within the same study in TNDH population and also across the two studies. *MTACP2* and *PKP-BETA1* appeared twice on A4 and C2, and twice on A10 and C2, respectively in Qiu's study. This could be due to homologous regions and both copies influence that trait. *LTP4* and *DGAT1* were listed in TNDH population on C2 and C3 and SGDN populations on A9 and A7, respectively as seen on Table 7-3. These candidate genes identified on multiple occasions could be more likely to affect the phenotype than genes which appeared only once.

Due to lack of a complete genomic sequence of *B. napus,* genetic study of this species has been carried out using linkage maps where the genetic distance between markers and length of linkage group is dependent on the population used. This prevents direct comparisons being made between different populations of *B. napus*. However the unigene based *B. napus* map has associated AGI gene codes and it is possible to compare different populations if their markers are also associated with AGI gene codes to allow remapping of them onto the unigene based *B. napus* map (Trick et al., 2009; Bancroft et al., 2011; Harper et al., 2012).

There was one position where two *B napus* QTL detected in different populations overlapped on A1 for oil content. Since the AGI gene codes are known for each unigene, the candidate gene search is possible. However as it could be a highly computational and of a different nature to this project, the work will be done in different project.

In conclusion use of the unigene based *B. napus* map gives the possibility of comparison between multiple experiments carried out in different populations, at different locations and with different species through use of AGI gene codes. This means the data collected in different population can be integrated and presented in a single map.

The genetic information extracted from such meta-analysis work could increase result reliability and contribute to the progress of crop breeding programmes.

CHAPTER 8. GENERAL DISCUSSION

8.1 SUMMARY OF THESIS

This thesis reports research designed to identify novel loci involved in the quantitative control of polyunsaturated fatty acid composition and oil content in *A. thaliana* seed. In addition a comparative genomics approach to identify corresponding loci in *B. napus*, oilseed rape is also described.

QTL analyses using six *A. thaliana* RI populations derived from wild accessions were performed to detect QTL which could be involved in the regulation of lipid biosynthesis and FA desaturation pathway of seed lipid. Of these newly identified QTL, 13 QTL were selected and NILs were developed for all 13 QTL in order to test the robustness of these QTL.

Development of NILs started with selection of the most suitable RILs as starting materials from the population in which the QTL were identified. Taking account of the possible cytoplasmic effects lines used as maternal parents for the development of RI populations were also used as maternal parents in NIL development. Selected RILs should have continuous string of paternal alleles at the QTL region of interest, but less paternal alleles in the rest of the genome. The original RILs were backcrossed twice to reduce unwanted paternal alleles in their backgrounds and selfed for segregation purposes in the population at the region of interest. Plants which were paternal homozygous at the region of interest were NIL candidates for the target QTL. Plants which were maternal homozygous at the region of interest were kept as controls. NILs and controls were phenotyped and compared. Out of the 13 QTL, eleven showed significant differences in phenotype between their NILs and controls, therefore the majority of QTL successfully showed their effects on their phenotypes in this work. A further round of selection was carried out to shortlist QTL for fine mapping, as this required more distinctive differences between NILs and their controls because the material selected would be backcrossed and selfed to be used for the development of large scale mapping populations. For each QTL, NILs and controls were compared as groups and if their SDs overlapped, they were insufficiently distinct and therefore such QTL were rejected. Seven QTL were remained at the end and a single NIL was selected for each QTL for the use in fine mapping.

Of these shortlisted seven QTL, the QTL for 18:2/18:1 in the NG population was selected for fine mapping for this project and two NG32 NIL(33) and NG32 NIL(34)

derived from RIL NG32 were used to develop large scale mapping populations. The reason that two NILs were selected for fine mapping was due to NIL(33) and NIL(34) having different characteristics. NIL(33) had the most favourable genotype and an outstandingly strong phenotype. Strong phenotype is one of the most important characteristics for the NIL selection, however in this instance it was extreme considering the range of parental lines and there were a risk that this strong phenotype could be due to factors other than genetic. Therefore NIL(34) which showed a reasonably strong phenotype was also selected. These NILs were backcrossed and selfed for the development of mapping populations.

Plants in the mapping populations were sorted according to four different recombination patterns and 12 bins within each. Approximately five plants were phenotyped from each bin and their 18:2/18:1 ratios were compared. The expected phenotypic pattern in the bins was a shift difference consisting of two phases, one phase would contain the causative locus for the QTL and the other would not. To detect this phase difference, the adjacent bin was tested using a T-test. If no difference was observed between the first and second bin ($P>0.05$), they were combined and compared with those plants in the third bin. It continued until an adjacent bin showed a significant difference ($P\leqq 0.05$) from the cluster of bins. Once the difference was observed, the analysis was started again from the adjacent bin and the next one. It was repeated in both directions, from bin 1 to bin 12 and bin 12 to bin 1. Overall, population(33), which was derived from NIL(33), had similar range of phenotype with population(34) derived from NIL(34), and additionally it showed inconsistent results during the detection of shift difference between bins. Therefore the results of population(34) were taken forward.

Through this analysis, bin7 was identified as the most likely to contain the causative locus for the QTL. The adjacent bins, 6 and 8 were also included for further analysis. Finally, the region of interest was narrowed down to 2Mb and seven protein coding genes involved in seed lipid biosynthesis and FA desaturation were taken as candidate genes. These genes encode the BIOTIN CARBOXYL-CARRIER SUBUNIT OF THE MULTI-ENZYME PLASTIDIAL ACETYL-CoA CARBOXYLASE COMPLEX (BCCP1 and BCCP2), ENDOPLASMIC RETICULUM-ADENINE NUCLEOTIDE TRANSPORTER1 (ER-ANT1) and predicted to encode two PLANT STEAROYL-ACP DESATURRASE FAMILY, NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN found in FA synthesis and PLASTID LIPID ASSOCIATED PROTEINS (PAPs).

In order to translate the findings in the model plant *A. thaliana* into the crop plant *B. napus*, comparative mapping was carried out using the hypothetically ordered unigene based map of *B. napus* genome developed within the Bancroft group. Most of the unigenes were associated with an AGI gene code. Those seven QTL shortlisted for the fine mapping in Chapter 5, each of which showed distinctive differences between their NILs and controls, were used for this study. AGI gene codes were used to map *A. thaliana* QTL on the unigene based map of *B. napus*. Their positions were compared with *B. napus* QTL which could also be mapped on this unigene map.

For FA composition, the results of 75 lines of the *B napus* TNDH population grown in the UK during 2006 and 2007 were used. Three *B. napus* and *A. thaliana* QTL were found to be coincident. Among the candidate genes which lay underneath, *AtFaTA* found under a QTL for 18:2 was likely to affect FA desaturation traits by controlling the composition of 18C FAs and longer chain FAs. The remainder of the candidates were more related to total oil content.

For oil content, the results of the 188 lines of the *B. napus* TNDH population grown in China during 2003 and 2004, and the results of the 282 lines of the *B. napus* SGDH population were compared with *A. thaliana* QTL. A total of ten lipid related genes were present in the regions coincident between the two species. Notably *MTACP2* involved in acyl transfer steps and *PKP-BETA1* which provides a precursor of FA synthesis both appeared twice within the TNDH population. DGAT1 which catalyses the final step of triacylglycerol synthesis and *LTP4* which encodes a lipid transfer protein appeared in both populations. These four genes are judged likely to affect the phenotype of interest in these studies.

8.2 GENERAL DISCUSSION

The purpose of this project was to identify uncharacterised novel loci involved in the regulation of already known major effect genes in the PUFA desaturation pathway. These four well studied major effect genes encode enzymes for PUFA desaturation and they were discovered and cloned through a mutation approach (James and Dooner, 1990; Lemieux et al., 1990; Arondel et al., 1992; Miquel and Browse, 1992; Yadav et al., 1993; Lightner et al., 1994; Okuley et al., 1994; James Jr et al., 1995; Carlsson et al., 2002; Kachroo et al., 2007).

Regarding new uncharacterised QTL, to study the regulatory mechanisms of these genes, different approaches were required to shed light from different perspectives. A strategy was devised which used a combination of three approaches, wild accessions to exploit natural variation, application of QTL analysis and multiple populations derived from the bi-parental cross.

The reasoning behind the strategy is that natural variation amongst accessions could occur as a result of adaptation to change of environment and also to challenges a new environment for plants. Therefore natural variation has contributed a great deal to the genetic study of adaptation traits such as flowering time, seed and plant size, and tolerance to abiotic factors (Krannitz et al., 1991; Koornneef et al., 1998; Alonso-Blanco et al., 1999; Kobayashi and Koyama, 2002; Ungerer et al., 2002; Ghandilyan et al., 2009). Seed lipid characteristics are also regarded as adaptive traits (Linder, 2000; Sanyal and Linder, 2013). Therefore the use of different accessions other than the common lab strains was expected to facilitate the detection of new loci affecting seed lipid traits in comparison to those that could be identified from a mutation approach.

Also these adaptative traits are often polygenic and the phenotypes are expressed quantitatively (Collard et al., 2005; Keurentjes et al., 2007), therefore QTL analysis was an appropriate method for the genetic analysis of such traits.

Finally, if a single population was analysed for QTLs then it could face the same limitations as the mutation approach. The results of QTL analysis are largely population specific (Lynch and Walsh, 1998) and some degree of genetic structure unique to a population is observed in most species (Hamrick and Godt, 1996; Bohonak, 1999). Therefore QTL analysis using a single population would utilise only a small part of the natural allelic variation available within a gene pool, and could detect only a subset of the loci controlling a trait (Symonds et al., 2005). A previous report in the literature also studies the genetic control of *Arabidopsis* seed lipid synthesis using QTL analysis, however in it only a single population is used (Hobbs et al., 2004) derived from reciprocal crosses between L*er* and Cvi (Alonso-Blanco et al., 1998). The study identified new QTL at the top and bottom of Ch1 and at the bottom of Ch2. All three QTL were also found in this study in the KB and TJ populations, however direct comparison of the QTL positions is difficult as the marker positions were only described in a schematic chart according to the genetic positions (cM). The QTL on Ch2, which explained 16.7% of the variance in oil content (Hobbs et al., 2004), was also identified in this study had a LOD score 6.46 and accounted for 28.2% of the variance in oil content. Therefore it was investigated

further by developing NILs using the RIL KB55. These NILs also showed significant differences between controls. I did not chose this QTL for fine mapping, because another QTL for 18:2/18:1 identified in the NG population was more significant and also the trait fitted better with the project remit for the generation of improved biolubricant. However, since it appeared repeatedly in two studies, this QTL is worth investigating in the future. Regarding FA composition, for 18:1, 18:2 and 18:3, four QTL were detected in the Hobbs' study. QTL for 18:2 and 18:3 were mapped to the *FAD3* locus and a QTL for 18:1 was close to the *FAD2* locus. Another QTL for 18:3 mapped close to the Regulator of Fatty acid Composition loci (*RFC2*, *RFC3* and *RFC4*) which were also identified in this study however it was below the significance level. The fact that there was no polymorphism at the *FAE1* locus between the two parental lines in the Hobbs' study underlines the limitation of using a single population.

Another QTL study for the seed oil composition in *A. thaliana* was carried out by Sanyal (Sanyal and Randal, 2012) using four populations. All had L*er* as a common parent crossed with different accessions. The study also identified known major effect genes at similar regions as found in this study. Numerous QTL segregated at new loci and all the QTL for 18:1, 18:2 and 18:3 were detected at similar positions as in this study, however the positions of markers were again described in terms of genetic position, making direct comparison of the QTL positions difficult.

Comparing to these previous study, this study used more populations. Prior to the start of this project, six pairs of wild accessions were selected on the basis of geographic divergence and different seed lipid characteristics in order to maximize polymorphisms between the progenitors. The resulting six RI populations consisting of 94 lines were developed within the Bancroft group (O'Neill et al., 2003; O'Neill et al., 2008). All six populations were exploited in the search for novel loci in this project as described in Chapter 4.

QTL, which were coincident with the positions of known major genes involved in the PUFA desaturation pathway were used to test the power of detection. In total 19 large effect QTL of such type were detected across the six populations. There was no instance of a population in which all four major genes were detected, however with six populations used altogether it became possible. Therefore studying all six RI populations together proved the value for the study of the genetic architecture of seed PUFA composition.

The fact that not all the QTL coincident with these major genes appear in any one population enabled the detection of new and minor effect QTL as otherwise they could be masked by the presence of large effect QTL (Keurentjes et al., 2007). Numerous minor QTL, still with LOD scores above the significance threshold, were successfully detected in all the traits across the populations. The results of QTL analysis in this project have proven that this combination of three approaches is effective. However, if populations were to be developed specifically for the detection of new QTL for FA composition, selection of parental lines could be done differently. Genotyping the loci of known major genes could be carried out to avoid crossing accessions containing allelic variation. Therefore the population derived from parental lines with the same allele at the major gene locus could theoretically be free from QTL caused by major genes for the trait of interest.

QTL that appeared in multiple populations were assigned a higher value than those detected in a single population. Traits such as FA contents and the ratio of two contents were closely related, for example 18:1 content and the ratio of 18:2 content/18:1 content, therefore QTL identified in both traits were also deemed of greater use. At the end, 13 QTL which covered all traits in the desaturation pathway and total oil content were successfully selected for further studies. Therefore in this strategy, the combination of QTL analysis, natural variation and multiple populations worked well and successfully detected relatively minor QTL for all traits of interest.

In Chapter 5, NILs were developed for all selected QTL to evaluate their effects experimentally. Since phenotyping FA composition is labour intensive and costly, marker assisted selection was primarily applied at every generation during the development of NILs. Taking advantage of this intensive genotyping, an attempt was made to simplify the procedure of NIL development compared to the previous study (Alonso-Blanco et al., 2003). Therefore the methodology itself was also tested as to whether it could deliver the required results.

The ANOVA was performed for phenotype comparison between the NIL material and the associated controls. Eleven QTL out of the original 13 target QTL showed significant differences. Among them eight QTL showed highly significant difference ($P<0.001$), two QTL showed intermediate significances ($P<0.01$) and one showed significant difference at $P<0.05$. Only two QTL did not show significant differences ($P=0.058$ and 0.578). Therefore it was proven that QTL detected through mapping based on statistical analysis actually had an influence on the phenotype despite residual minor paternal alleles remaining in the background. At the same time this

method, a combination of intensive genotyping together with a simplified development process performed well enough to be able to deliver such results.

For fine mapping, a binning system was designed to narrow down the region of interest by identifying the bin affecting the phenotype. The region between two flanking markers was divided into 12 bins using eleven InDel markers. As it was not feasible to obtain sufficient numbers of plants with only paternal alleles in a bin, the recombination position was used to categorize plants into bins. For example those plants with a change of allele between the upper flanking marker and marker1, were sorted into bin1. As the bin position referred to where the recombination occurred, the change of phenotype should happen only once between a group of bins which contained the causative locus of a QTL and the adjacent group of bins which did not.

For fine mapping, among the several anticipated recombination patterns which could occur in the region between two flanking markers, four informative patterns were chosen and named from Type 1 to Type 4 these four different types could be used for testing for dominance effect of the QTL. For example, if the effect was dominant, all Type 3 and Type 4 bins should have the same high ratio of 18:2/18:1, otherwise their results should be similar to Types 1 and 2, respectively. This system worked well and showed that the effect of the target QTL was additive and it showed similarities between types as predicted. Analysis using these four informative recombination patterns also helped protect against the anticipated high EER due to multiple comparisons between neighbouring bins using t-test. There is about a 26% possibility that one or more false positives could have occurred during the six comparisons (Iglewicz, 2005). However by referring four results from each type, the risk could be reduced.

Heterozygous lines were used in the fine mapping populations as seen in Fig.6-5. An alternative method that could have been used was homozygous lines which would show more distinctive phenotypic differences between the bins which contained the causative locus of the QTL and those bins which didn't. In this case only two types would be sufficient instead of four. However it required additional selfing of plants from each bins and subsequent genotyping of enormous number of samples to screen homozygous lines for each bin. Therefore it would cost more and take longer to develop the population, so that it wasn't applied in this project.

Regarding the studies in the amphidiploid *B. napus*, due to the genome triplication and hybridization events (U, 1935; O'Neill and Bancroft, 2000; Lysak et al., 2005), therefore there could be six copies for each gene and positioning of genes and the number of functional genes is not straight forward. Moreover the complete genomic sequence for *B. napus* is not yet publically available. Gene copy number was not considered in this comparative study with diploid *A. thaliana*, however there is a possibility that multiple copies of genes could be the cause of the phenotypic variation as seen in the study of *rfc4* and *ife* mutants which were shown to contain additional copies of *FAD3* gene sequences within large duplicated segments of 2Mb and 1.4 Mb respectively and had increased 18:3 contents compared to the wild type (O'Neill et al., 2011).

The previous studies in *B. napus* regarding 18:1, 18:2 and 18:3 were carried out in two ways, one is the identification of loci corresponding to the known major effect genes in *A. thaliana* such as *FAD2* and *FAD3* and *FAE1* using *A. thaliana*-based candidate gene probes, and the second is QTL mapping. The positions of loci corresponding to the major genes were also described on a linkage map with genetic positions. Four orthologous loci for *FAD2* mapped to linkage groups A1, A5, C1 and C5 (Scheffler et al., 1997) and A5 (Schierholt, 2000). Later Smooker et al. (2011) mapped loci for FAD2 to C1, A5 and C5. Scheffler et al (1997) mapped five loci for *FAD3* to A3, A4, A5 and at two loci on C4, while in the work by Smooker et al. (2011) three loci were mapped, one on C3 and two on C4. The QTL mapping carried out by Burns et al (2003) detected QTL for 18:2 on C3 and C4, while Hu et al (2006) identified QTL for 18:3 on A4 and C4, and Zhao et al (2008) detected QTL for 18:3 on C4 (Burns et al., 2003; Hu et al., 2006; Zhao et al., 2008). Both Burns et al (2003) and Zhao et al (2008) detected a QTL for 18:1 on A8, which was identified as linked to *FAE1* in the study by Smooker et al (2011).

Comparative genomics between *B. napus* and *A. thaliana* were carried out exploiting the hypothetically ordered unigene based map of the *B. napus* genome associated with AGI gene codes. The key process of this analysis was the alignment of QTL detected in different species with different linkage maps or in different populations, onto this unigene based *B. napus* map using Mapper v.3. Five possible scenarios for the association of *A. thaliana* QTL with *B. napus* QTL through the use of the unigene based *B. napus* map are listed in Chapter 7 Fig.7-3. Four approaches were taken as described in Chapter 7 however, the remaining one would surely be the most common case; QTL identified in non-TNDH populations with no AGI gene

codes association. One possible way to relate such mapping with the unigene based *B. napus* map could be through the use of AGI gene codes. The sequence of genetic markers coincident with a QTL could be blasted against the *A. thaliana* genome and located on the *A. thaliana* physical map and searched for AGI gene codes closest to the markers. Then the AGI gene codes could be used to estimate the location of the QTL on the unigene map. In the SGDH population, most of the markers used in the linkage map, 280 markers, were associated with AGI gene codes. However, if only alignment with the unigene based *B. napus* map was required, it would be enough to sequence markers close to the QTL, therefore reducing both cost and time. In the future when complete *B.napus* genome sequences are publically available, alignment of QTL identified in different studies can be easily carried out by BLASTing the sequence of a region of interest against the genome sequence.

However it still proved difficult to search for candidate genes of *B napus* QTL despite all unigenes being associated with AGI gene models as the volume of data to be processed was enormous. The development of a program which allows a keyword search of the AGI gene models within a *B. napus* QTL interval would dramatically reduce the time required for the candidate gene analysis of *B. napus*.

8.3 FUTURE WORK

Further fine mapping could be performed to narrower the region using InDel markers and SNP markers detected in the genomic sequences of the parental lines, Nok-3 and Ga-0. These markers are frequent enough, therefore it may be possible to pinpoint a candidate gene if there were enough plants segregated in the region. Direct comparison of sequence data between the two parental lines could also be carried out after further narrowing down of the region, to detect obvious polymorphisms between parental lines close to the target gene.

For complete cloning, functional analysis using knock out mutants of the gene and complimentary testing using transgenic plants should be performed.

The interaction between the target QTL and another QTL that maps to the upper side of the target QTL, likely to be *FAD8,* was discussed in Chapter 6. If the result of fine mapping is to be used for breeding to manipulate 18:1 and 18:2 FA composition in oilseed rape, it is crucial to determine the basis of the interaction in the future as there is a high possibility that the target QTL could affect the ratio of 18:2/18:1

positively only in the presence of this other QTL. To determine the interaction, firstly the candidate gene of the causative of QTL should be determined, then the phenotypes of knockout mutants of the candidate gene and *fad8*, their double mutant and the wild type should be compared. If a clear interaction between the two genes is found, this could provide interesting material for study of the mechanism of regulation between the two genes.

REFERENCES AND BIBLIOGRAPHY

Al-Shehbaz IA, OKane SL (2002) Taxonomy and Phylogeny of *Arabidopsis* (*Brassicaceae*). *In Arabidopsis* Book. American Society of Plant Biologists

Alonso-Blanco C, Bentsink L, Hanhart CJ, Blankestijn-de Vries H, Koornneef M (2003) Analysis of Natural Allelic Variation at Seed Dormancy Loci of *Arabidopsis thaliana*. Genetics **164:** 711-729

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Koornneef M (1999) Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences **96:** 4710-4717

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a L*er*/Cvi recombinant inbred line population. The Plant Journal **14:** 259-271

Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science **258:** 1353-1355

Arondel V, Vergnolle C, Cantrel C, Kader J-C (2000) Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*. Plant Science **157:** 1-12

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter **9:** 208-218

Bancroft I, Morgan C, Fraser F, Higgins J, Wells R, Clissold L, Baker D, Long Y, Meng J, Wang X, Liu S, Trick M (2011) Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. Nat Biotech **29:** 762-766

Baud S, Wuillème S, Dubreucq B, De Almeida A, Vuagnat C, Lepiniec L, Miquel M, Rochat C (2007) Function of plastidial pyruvate kinases in seeds of *Arabidopsis thaliana*. The Plant Journal **52:** 405-419

Beaudoin F, Wu X, Li F, Haslam RP, Markham JE, H Z, Napier JA, Kunst L (2009) Functional Characterization of the Arabidopsis b-Ketoacyl-Coenzyme A Reductase Candidates of the *Fatty Acid Elongase1*. *In* Plant Physiology, Vol 150, pp 1174-1191

Beisson F, Koo AJK, Ruuska S, Schwender J, Pollard M, Thelen JJ, Paddock T, Salas J.J., Savage L, Milcamps A, Mhaske VB, Cho Y, Ohlrogge JB (2003) *Arabidopsis* Genes Involved in Acyl Lipid Metabolism. A 2003 Census of the Candidates, a Study of the Distribution of Expressed Sequence Tags in Organs, and a Web-Based Database. *In* Plant Physiology, Vol 132, pp 681-697

Bentsink L, Yuan K, Koornneef M, Vreugdenhil D (2003) The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana* using natural variation. TAG Theoretical and Applied Genetics **106:** 1234-1243

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *In* Q Rev Biol. , Vol 74, pp 21-45

Borevitz JO, Liang D, Plouffe D, Chang H-S, Zhu T ea (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res **13:** 10

Browse J, Somerville C (1991) Glycerolipid Synthesis: Biochemistry and Regulation. Annual Review of Plant Physiology and Plant Molecular Biology **42:** 467-506

Burns MJ, Barnes SR, Bowman JG, Clarke MHE, Werner CP, Kearsey MJ (2003) QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) Seed oil content and fatty acid composition. Heredity **90:** 39-48

Cao Y, Yang J, Xian M, Xu X, Liu W (2010) Increasing unsaturated fatty acid contents in *Escherichia coli* by coexpression of three different genes. Applied Microbiology and Biotechnology **87:** 271-280

Carlsson AS, LaBrie ST, Kinney AJ, Von Wettstein-Knowles P, Browse J (2002) A *KAS2* cDNA complements the phenotypes of the *Arabidopsis fab1* mutant that differs in a single residue bordering the substrate binding pocket. The Plant Journal **29:** 761-770

Cavell AC, Lydiate DJ, Parkinl.A .P., Dean C, and Trick, M., (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome **41:** 8

Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA, Chen H, Frazer KA, Huson DH, Schölkopf B, Nordborg M, Rätsch G, Ecker JR, Weigel D (2007) Common Sequence Polymorphisms Shaping Genetic Diversity in *Arabidopsis thaliana*. Science **317:** 338-342

Collard B, Jahufer M, Brouwer J, Pang E (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica **142:** 169-196

Department for Environment Food and Rural Affair (2006) Agricultural land use1986-2005. http://archive.defra.gov.uk/evidence/statistics/environment/land/alltables.htm

Eastmond PJ, Dennis DT, Rawsthorne S (1997) Evidence That a Malate/Inorganic Phosphate Exchange Translocator Imports Carbon across the Leucoplast Envelope for Fatty Acid Synthesis in Developing Castor Seed Endosperm. *In* Plant Physiology Vol 114, pp 851-856

Edwards KD, Lynn JR, Gyula P, Nagy F, Millar AJ (2005) Natural Allelic Variation in the Temperature-Compensation Mechanisms of the *Arabidopsis thaliana* Circadian Clock. Genetics **170:** 387-400

Fehling E, Mukherjee KD (1991) Acyl-CoA elongase from a higher plant (Lunaria annua): metabolic intermediates of very-long-chain acyl-CoA products and substrate specificity. Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism **1082:** 239-246

Feria Bourrellier AB, Valot B, Guillot A, Ambard-Bretteville F, Vidal J, Hodges M (2010) Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate–regulated by interaction of PII with the biotin carboxyl carrier subunit. Proceedings of the National Academy of Sciences **107:** 502-507

Ghandilyan A, Barboza L, Tisné S, Granier C, Reymond M, Koornneef M, Schat H, Aarts MGM (2009) Genetic analysis identifies quantitative trait loci controlling rosette mineral concentrations in *Arabidopsis thaliana* under drought. New Phytologist **184:** 180-192

Griffiths G, Stobart AK, Stymne S (1985) The acylation of sn-glycerol 3-phosphate and the metabolism of phosphatidate in microsomal preparations from the developing cotyledons of safflower (*Carthamus tinctorius L.*) seed. *In* Biochem J. , pp 379–388

Gurr MI, Harwood, J. L., Frayn, K. N (2002) Lipid Biochemistry An Introduction 5th Edition. Blackwell Science, Oxford

Hamrick JL, Godt MJW (1996) Effects of Life History Traits on Genetic Diversity in Plant Species, Vol 351. The Royal Society

Harper AL, Trick M, Higgins J, Fraser F, Clissold L, Wells R, Hattori C, Werner P, Bancroft I (2012) Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. Nat Biotech **advance online publication**

Heath RJ, Rock CO (1996) Roles of the FabA and FabZ β-Hydroxyacyl-Acyl Carrier Protein Dehydratases in *Escherichia coli* Fatty Acid Biosynthesis. Journal of Biological Chemistry **271:** 27795-27801

Hernandez-Pinzon I, Ross JHE, Barnes KA, Damant AP, Murphy DJ (1999) Composition and role of tapetal lipid bodies in the biogenesis of the pollen coat of *Brassica napus*. Planta **208:** 588-598

Hobbs DH, Flintham JE, Hills MJ (2004) Genetic Control of Storage Oil Synthesis in Seeds of *Arabidopsis*. Plant Physiol. **136:** 3341-3349

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Letters **452:** 145-149

Hoekenga OA, Vision TJ, Shaff JE, Monforte AJ, Lee GP, Howell SH, Kochian LV (2003) Identification and Characterization of Aluminum Tolerance Loci in *Arabidopsis* (Landsberg *erecta* × Columbia) by Quantitative Trait Locus Mapping. A Physiologically Simple But Genetically Complex Trait. *In* Plant Physiology Vol 132, pp 936-948

Hu X, Sullivan-Gilbert M, Gupta M, Thompson S (2006) Mapping of the loci controlling oleic and linolenic acid contents and development of fad2 and fad3 allele-specific markers in canola (*Brassica napus* L.). TAG Theoretical and Applied Genetics **113:** 497-507

Huang AHC (1992) Oil Bodies and Oleosins in Seeds. Annual Review of Plant Physiology and Plant Molecular Biology **43:** 177-200

Ichihara K (1984) sn-Glycerol-3-phosphate acyltransferase in a particulate fraction from maturing safflower seeds. *In* Arch Biochem Biophys. , Vol 1;232(2):, pp 685-698

Iglewicz B (2005) Experiment-Wise Error Rate. *In* P Armitage, Colton, T., , ed, Encyclopedia of Biostatistics, Ed 2nd Edition Vol 3. Wiley, Chichester, pp pp.1845-1846

Inaba R, Nishio T (2002) Phylogenetic analysis of *Brassiceae* based on the nucleotide sequences of the S-locus related gene, *SLR1*. TAG Theoretical and Applied Genetics **105:** 1159-1165

James DW, Dooner HK (1990) Isolation of EMS-induced mutants in *Arabidopsis* altered in seed fatty acid composition. TAG Theoretical and Applied Genetics **80:** 241-245

James Jr DW, Lim E, Keller J, Plooy I, Ralston E, Dooner HK (1995) Directed Tagging of the Arabidopsis *FATTY ACID ELONGATION1* (*FAE1*) Gene with the Maize Transposon Activator. Plant Cell **7:** 309-319

Jaworski JG, Clough, R.C., and Barnum, S.R. (1989) A cerulenin insensitive short chain 3-ketoacyl-acyl carrier protein synthase in *Spinacia oleracea* leaves. Plant Physiol**:** 41-44

Juenger TE, McKay JK, Hausmann N, Keurentjes JJB, Sen S, Stowe KA, Dawson TE, Simms EL, Richards JH (2005) Identification and characterization of QTL underlying whole-plant physiology in *Arabidopsis thaliana*: δ13C, stomatal conductance and transpiration efficiency. Plant, Cell & Environment **28:** 697-708

Kachroo A, Shanklin J, Whittle E, Lapchyk L, Hildebrand D, Kachroo P (2007) The *Arabidopsis* stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. Plant Molecular Biology **63:** 257-271

Kearsey MJ, Farquhar AGL (1998) QTL analysis in plants; where are we now? Heredity **80:** 137-142

Keurentjes JJB, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M (2007) Development of a Near-Isogenic Line Population of *Arabidopsis thaliana* and Comparison of Mapping Power With a Recombinant Inbred Line Population. Genetics **175:** 891-905

Kim KH, Kim MY, Van K, Moon J, Kim DH, Lee S (2008) Marker-Assisted Foreground and Background Selection of Near Isogenic Lines for Bacterial

Leaf Pustule Resistant Gene in Soybean. *In* J. Crop Sci. Biotech., Vol 11, p 263 ~ 268

Kimber D, McGregor DI, eds (1995) *Brassica* Oilseeds. CAB International, Oxon

Kobayashi Y, Koyama H (2002) QTL Analysis of Al Tolerance in Recombinant Inbred Lines of *Arabidopsis thaliana*. Plant and Cell Physiology **43:** 1526-1533

Koornneef M, Alonso-Blanco C, Peeters AJM, Soppe W (1998) Genetic Control of Flowering Time in *Arabidopsis*. Annual Review of Plant Physiology and Plant Molecular Biology **49:** 345-370

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally Occuriing Genetic Variation in *Arabidopsis thaliana*. Annual Review of Plant Biology **55:** 141-172

Koornneef M, Reymond M, Alonso-Blanco C (2011) Natural Variation in *Arabidopsis thaliana*. *In* Genetics and Genomics of the *Brassicaceae*. Springer pp 123-151

Krannitz PG, Aarssen LW, Dow JM (1991) The Effect of Genetically Based Differences in Seed Size on Seedling Survival in *Arabidopsis thaliana* (*Brassicaceae*). *In* American Journal of Botany, Vol 78. Botanical Society of America, pp 446-450

Lander ES, Botstein D (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. Genetics **121:** 185-199

Lemieux B, Miquel M, Somerville C, Browse J (1990) Mutants of *Arabidopsis* with alterations in seed lipid fatty acid composition. TAG Theoretical and Applied Genetics **80:** 234-240

Lerocha M, H.,, Neuhausa E, Kirchbergera S, Zimmermanna S, Melzerb M, Gerholda J, Tjadena J (2008) Identification of a Novel Adenine Nucleotide Transporter in the Endoplasmic Reticulum of *Arabidopsis*. *In* The Plant Cell pp 438-451

Li L, Li X, Ilarslan H, Wurtele E, Nikolau B (2005) Functional and Expresssion Analysis of Heteromeric Acetyl-CoA Carboxylase Subunit Genes in *Arabidopsis In* 16TH INTERNATIONAL CONFERENCE ON ARABIDOPSIS RESEARCH,

Lightner J, James DWJ, Dooner HK, Browse J (1994) Altered body morphology is caused by increased stearate levels in a mutant of *Arabidopsis*. The Plant Journal **6:** 401-412

Linder C (2000) Adaptive evolution of seed oils in plants: accounting for the biogeographic disribution of saturated and unsaturated fatty acid in seed oils. *In* The American Naturarist, Vol 156, pp 442-458

Lynch M, Walsh B (1998) Genetics and Analysis of Quantitative Traits Sinauer Associates

Lysak MA, Koch MA (2011) Phylogeny, Genome, and Karyotype Evolution of Crucifers (*Brassica*). *In* R Schmidt, Bancroft, I.,, ed, Genetics and Genomics of the *Brassicaceae*. Springer, New York, pp 1-31

Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe *Brassiceae*. Genome Res. **15:** 516-525

McConn M, Hugly S, Browse J, Somerville C (1994) A Mutation at the *fad8* Locus of *Arabidopsis* Identifies a Second Chloroplast [omega]-3 Desaturase. Plant Physiology **106:** 1609-1614

Millar AA, Kunst L (1997) Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. The Plant Journal **12:** 121-131

Mingjie C, Mooney BP, Hajduch M, Casteel J. E. , Joshi T, Zhou M, Xu D, Thelen J. J. (2008) Parallel Microarray and Proteomic Analysis of a Seed-Specific Acetyl-CoA Carboxylase Mutant Reveals Associated Changes in Metabolism. *In* 19TH INTERNATIONAL CONFERENCE ON ARABIDOPSIS RESEARCH,

Miquel M, Browse J (1992) *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis. Biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. J. Biol. Chem. **267:** 1502-1509

Mitchell-Olds T (1996) Genetic constraints on life history evolutionquantitative trait loci influencing growth and flowering in *Arabidopsis thaliana*. *In* Evolution, Vol 50, pp 140-145

Mithen R, Clarke J, Lister C, Dean C (1995) Genetics of aliphatic glucosinolates. III. Side chain structure of aliphatic glucosinolates in *Arabidopsis thaliana*. Heredity **74:** 210-215

Monforte AJ, Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. Genome **43: :** 803-813

Moreno-Pérez A, Venegas-Calerón M, Vaistij F, Salas J, Larson T, Garcés R, Graham I, Martínez-Force E (2012) Reduced expression of FatA thioesterases in *Arabidopsis* affects the oil content and fatty acid composition of the seeds. Planta **235:** 629-639

O'Neill CM, Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. alboglabra that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. The Plant Journal **23:** 233-243

O'Neill CM, Gill S, Hobbs D, Morgan C, Bancroft I (2003) Natural variation for seed oil composition in *Arabidopsis thaliana*. Phytochemistry **64:** 1077-1090

O'Neill CM, Morgan C, Hattori C, Brennan M, Rosas U, Tschoep H, Deng PX, Baker D, Wells R, Bancroft I (2012) Towards the genetic architecture of seed lipid biosynthesis and accumulation in *Arabidopsis thaliana*. Heredity **108:** 115-123

O'Neill C, Morgan C, Kirby J, Tschoep H, Deng P, Brennan M, Rosas U, Fraser F, Hall C, Gill S, Bancroft I (2008) Six new recombinant inbred populations for the study of quantitative traits in *Arabidopsis thaliana*. TAG Theoretical and Applied Genetics **116:** 623-634

O'Neill CM, Baker D, Bennett G, Clarke J, Bancroft I (2011) Two high linolenic mutants of *Arabidopsis thaliana* contain megabase-scale genome duplications encompassing the *FAD3* locus. The Plant Journal **68:** 912-918

Ohlrogge J, Browse J (1995) Lipid Biosynthesis. Plant Cell **7:** 957-970

Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J (1994) *Arabidopsis FAD2* Gene Encodes the Enzyme That Is Essential for Polyunsaturated Lipid Synthesis. Plant Cell **6:** 147-158

Oliver F, Roland M-P (2004) The effect of fatty acid composition on biodiesel oxidative stability. EUROPEAN JOURNAL OF LIPID SCIENCE AND TECHNOLOGY **106:** 837-843

Paran I, Zamir D (2003) Quantitative traits in plants: beyond the QTL. Trends in Genetics **19:** 303-306

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics**:** genetics.105.042093

Petrinovich LFH, Curtis D. (1969) Error rates for multiple comparison methods: Some evidence concerning the frequency of erroneous conclusions. *In* Psychological Bulletin, Vol 71, pp 43-54

Qiu D, Morgan C, Shi J, Long Y, Liu J, Li R, Zhuang X, Wang Y, Tan X, Dietrich E, Weihmann T, Everett C, Vanstraelen S, Beckett P, Fraser F, Trick M, Barnes S, Wilmer J, Schmidt R, Li J, Li D, Meng J, Bancroft I (2006) A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. TAG Theoretical and Applied Genetics **114:** 67-80

Quesada V, García-Martínez S, Piqueras P, Ponce MR, Micol JL (2002) Genetic Architecture of NaCl Tolerance in *Arabidopsis*. *In* Plant Physiology, Vol 130, pp 951-963

Rana D, van den Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP, Bancroft I (2004) Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. The Plant Journal **40:** 725-733

Rashotte AM, Jenks MA, Nguyen TD, Feldmann KA (1997) Epicuticular wax variation in ecotypes of *Arabidopsis thaliana*. *In* Phytochemistry, Vol 45, pp 251-255

Robbelen G, Downey RK, Ashri A (1989) Oil crops of the world. McGraw-Hill, Inc, USA

Ruuska SA, Girke T, Benning C, Ohlrogge JB (2002) Contrapuntal Networks of Gene Expression during *Arabidopsis* Seed Filling. Plant Cell **14:** 1191-1206

Sanyal A, Linder C (2013) Plasticity and constraints on fatty acid composition in the phospholipids and triacylglycerols of *Arabidopsis* accessions grown at different temperatures. BMC Plant Biology **13:** 63

Sanyal A, Randal L, C. (2012) Quantitative trait loci involved in regulating seed oil composition in *Arabidopsis thaliana* and their evolutionary implications. Theoretical and Applied Genetics **124:** 723-738

Scarisbrick DH, Daniels RW, eds (1986) Oilseed rape. Collins, London

Scheffler JA, Sharpe AG, Schmidt H, Sperling P, Parkin IAP, Luhs W, Lydiate DJ, Heinz E (1997) Desaturase multigene families of *Brassica napus* arose through genome duplication. Theoretical and Applied Genetics **94:** 583-591

Schierholt A, Becker, H.C., Ecke, W.G. (2000) Mapping a high oleic acid mutation in winter oilseed rape (*Brassica napus L.*). *In* Theor Appl Genet 101, Vol 101, pp 897-901

Schmidt R, Acarkan A, Boivin K (2001) Comparative structural genomics in the *Brassicaceae* family. Plant Physiology and Biochemistry **39:** 253-262

Schnurr J. A., Shockey J. M., De Boer G, A. BJ (2002) Fatty Acid Export from the Chloroplast. Molecular Characterization of a Major Plastidial Acyl-Coenzyme A Synthetase from *Arabidopsis*. *In* Plant Physiology, Vol 129, pp 1700-1709

Shanklin J, Somerville C (1991) Stearoyl-acyl-carrier-protein desaturase from higher plants is structurally unrelated to the animal and fungal homologs. Proceedings of the National Academy of Sciences of the United States of America **88:** 2510-2514

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotech **26:** 1135-1145

Slate J (2005) Quantitative trait locus mapping in natural populations: progress,caveatsandfuturedirections. Mol.Ecol. **14:** 363-379.

Smith RG, David AG, David TD, David HT (1992) Malate- and Pyruvate-Dependent Fatty Acid Synthesis in Leucoplasts from Developing Castor Endosperm. *In* Plant Physiology, Vol 98, pp 1233-1238

Stratton DA (1998) Reaction norm functions and QTL-environment interactions for flowering time in *Arabidopsis thaliana*. Heredity **81:** 144-155

Stryer L (1988) Biochemistry Third Edition. W.H. Freeman and Company, New York

Swarup K, Alonso-Blanco C, Lynn JR, Michaels SD, Amasino RM, Koornneef M, Millar AJ (1999) Natural allelic variation identifies new genes in the *Arabidopsis circadian* system. The Plant Journal **20:** 67-77

Symonds VV, Godoy AV, Alconada T, Botto JF, Juenger TE, Casal JJ, Lloyd AM (2005) Mapping Quantitative Trait Loci in Multiple Populations of *Arabidopsis thaliana* Identifies Natural Allelic Variation for Trichome Density. Genetics **169:** 1649-1658

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796-815

Thelen JJ, Ohlrogge JB (2002) Metabolic Engineering of Fatty Acid Biosynthesis in Plants. Metabolic Engineering **4:** 12-21

Ting JTL, Wu SSH, Ratnayake C, Huang AHC (1998) Constituents of the tapetosomes and elaioplasts in *Brassica campestris* tapetum and their degradation and retention during microsporogenesis. Plant Journal **16:** 541-551

Tonsor SJ, Alonso-Blanco C, Koornneef M (2005) Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. Plant, Cell & Environment **28:** 2-20

Trick M, Cheung F, Drou N, Fraser F, Lobenhofer E, Hurban P, Magusin A, Town C, Bancroft I (2009) A newly-developed community microarray resource for transcriptome profiling in *Brassica* species enables the confirmation of *Brassica*-specific expressed sequences. BMC Plant Biology **9:** 50

Trick M, Long Y, Meng J, Bancroft I (2009) Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. Plant Biotechnology Journal **7:** 334-346

Trontin C, Tisné S, Bach L, Loudet O (2011) What does *Arabidopsis* natural variation teach us (and does not teach us) about adaptation in plants? Current Opinion in Plant Biology **14:** 225-231

Tuinstra MR, Ejeta G, Goldsbrough PB (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. Theoretical and Applied Genetics **95:** 1005-1011

Tuinstra MR, Ejeta G, Goldsbrough PB (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. TAG Theoretical and Applied Genetics **95:** 1005-1011

U N (1935) Genomic analysis of *Brassica* with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilization. . Japanese Journal of Botany **7:** 63

Ungerer MC, Halldorsdottir SS, Modliszewski JL, Mackay TFC, Purugganan MD (2002) Quantitative Trait Loci for Inflorescence Development in *Arabidopsis thaliana*. Genetics **160:** 1133-1151

Van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity **83:** 613-624

Van Ooijen JW (2004) MapQTL v.5 Software for the mapping of quantitative trait loci in experimental populations., Wageningen, Netherlands

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun J-H, Bancroft I, Cheng F, Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires JC, Paterson AH, Chalhoub B, Wang B, Hayward A, Sharpe AG, Park B-S, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, Peng C, Geng C, Koh C, Lin C, Edwards D, Mu D, Shen D, Soumpourou E, Li F, Fraser F, Conant G, Lassalle G, King GJ, Bonnema G, Tang H, Wang H, Belcram H, Zhou H, Hirakawa H, Abe H, Guo H, Wang H, Jin H, Parkin IAP, Batley J, Kim J-S, Just J, Li J, Xu J, Deng J, Kim JA, Li J, Yu J, Meng J, Wang J, Min J, Poulain J, Hatakeyama K, Wu K, Wang L, Fang L, Trick M, Links MG, Zhao M, Jin M, Ramchiary N, Drou N, Berkman PJ, Cai Q, Huang Q, Li R, Tabata S, Cheng S, Zhang S, Zhang S, Huang S, Sato S, Sun S, Kwon S-J, Choi S-R, Lee T-H, Fan W, Zhao X, Tan X, Xu X, Wang Y, Qiu Y, Yin Y, Li Y, Du Y, Liao Y, Lim Y, Narusaka Y, Wang Y, Wang Z, Li Z, Wang Z, Xiong Z, Zhang Z (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet **43:** 1035-1039

Weiss EA, ed (1983) Oilseed Crops. Longman, London and New York

Yadav NS, Wierzbicki A, Aegerter M, Caster CS, Perez-Grau L, Kinney AJ, Hitz WD, Booth Jr JR, Schweiger B, Stecca KL, Allen SM, Blackwell M, Reiter RS, Carlson TJ, Russell SH, Feldmann KA, Pierce J, Browse J

(1993) Cloning of Higher Plant [omega]-3 Fatty Acid Desaturases. Plant Physiol. **103:** 467-476

Yang T, Kim J, Kwon S, Lim K, Choi B, Kim J, Jin M, Park J, Lim M, Kim H, Lim Y, Kang J, Hong J, Kim C, Bhak J, Bancroft I, Park B (2006) Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. Plant Cell **18:** 1339 - 1347

Yang YW, Lai KN, Tai PY, Li WH (1999) Rates of Nucleotide Substitution in Angiosperm Mitochondrial DNA Sequences and Dates of Divergence Between *Brassica* and Other Angiosperm Lineages. Journal of Molecular Evolution **48:** 597-604

Young ND (1996) QTL Mapping and Qwantitative Disease Resistance in Plants. Annual Review of Phytopathology **34:** 479

Young ND, Zamir D, Ganal MW, Tanksley SD (1988) Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the tm-2a gene in tomato. Genetics **120:** 579-585

Zachowski A, Guerbette F, Grosbois M, Jolliot-Croquin A, Kader J-C (1998) Characterisation of acyl binding by a plant lipid-transfer protein. European Journal of Biochemistry **257:** 443-448

Zhao J, Dimov Z, Becker H, Ecke W, Möllers C (2008) Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. Molecular Breeding **21:** 115-125

Zhao J, Huang J, Chen F, Xu F, Ni X, Xu H, Wang Y, Jiang C, Wang H, Xu A, Huang R, Li D, Meng J (2012) Molecular mapping of Arabidopsis thaliana lipid-related orthologous genes in ;*Brassica napus*. TAG Theoretical and Applied Genetics **124:** 407-421

APPENDICES CONTENTS

Appendix 1 Linkage mapping 4
Calculation option settings 4
Appendix 2 QTL mapping population phenotype data analysis and MAP charts 7
ANOVA results of six populations 7
Proportion of heritable variation of six population 8
CA population histogram 9
CA population normality test 10
CA population QTL map charts 18:1 11
CA population QTL map charts 18:2 12
CA population QTL map charts 18:2 MQM mapping 13
CA population QTL map charts 18:2/18:1 14
CA population QTL map charts 18:3 15
CA population QTL map charts 18:3 MQM mapping 16
CA population QTL map charts 18:3/18:2 17
CA population QTL map charts 18:3/18:2 18
CA population QTL map charts oil content 19
SG population histogram 20
SG population normality test 21
SG population QTL map chart 18:1 22
SG population QTL map chart 18:1 MQM mapping 23
SG population QTL map chart 18:2 24
SG population QTL map chart 18:2 MQM mapping 25
SG population QTL map chart 18:2/18:1 26
SG population QTL map chart 18:2/18:1 MQM mapping 27
SG population QTL map chart 18:3 28
SG population QTL map chart 18:3/18:2 29
SG population QTL map chart 18:3/18:2 MQM mapping 30
SG population QTL map chart oil content 31
WC population histogram 32
WC population normality test 33
WC population QTL map chart 18:1 34
WC population QTL map chart 18:2 35
WC population QTL map chart 18:2/18:1 36
WC population QTL map chart 18:3 37
WC population QTL map chart 18:3/18:2 38

WC population QTL map chart oil content 39
KB population histogram 40
KB population normality test 41
KB population QTL map chart 18:1(log10) 42
KB population QTL map chart 18:1(log10) MQM mapping 43
KB population QTL map chart 18:2 44
KB population QTL map chart 18:2 MQM mapping 45
KB population QTL map chart 18:2/18:1 46
KB population QTL map chart 18:2/18:1 MQM mapping 47
KB population QTL map chart 18:3 48
KB population QTL map chart 18:3/18:2 49
KB population QTL map chart 18:3/18:2 MQM mapping 50
KB population QTL map chart oil content 51
NG population histogram 52
NG population normality test 53
NG population QTL map chart 18:1 54
NG population QTL map chart 18:2 55
NG population QTL map chart 18:2/18:1 56
NG population QTL map chart 18:3 57
NG population QTL map chart 18:3/18:2 58
NG population QTL map chart oil content 59
TJ population histogram 60
TJ population normality test 61
TJ population QTL map chart 18:1 62
TJ population QTL map chart 18:2 63
TJ population QTL map chart 18:2/18:3 64
TJ population QTL map chart 18:3 65
TJ population QTL map chart 18:3/18:2 66
TJ population QTL map chart il content 67
Appendix 3 Genotype files for RIL selection for NIL development 68
TJ25 18:2 68
NG28 and NG32 18:2/18:1 69
TJ79 18:2/18:1 70
SG93 18:3 71
TJ17 18:3 72
NG32 18:3/18:2 73
NG93 18:3/18:2 74
SG10 18:3/18:2 75
KB55 oil 76
KB91 oil 77

TJ79 oil Ch2 and Ch5 78
TJ83 oil 79
Appendix 4 Histogram of RIL population (phenotype data were provided by Carmel O'Neill measured in 2008) 80
Appendix 5 Genotype file of NIL candidates and MH plants and phenotype data 83
TJ25 18:2 83
NG28 18:2/18:1 85
NG32 18:2/18:1 87
TJ79 18:2/18:1 89
TJ17 18:3 91
NG32 18:3/18:2 93
NG93 18:3/18:2 95
SG10 18:3/18:2 97
KB55 oil 99
TJ79 oil Ch2 101
TJ79 oil Ch5 103
TJ83 oil 105
Appendix 6 Mapping population binned 107
Appendix 7 Scoring of genotype by GeneMapper v 3.7 109
Appendix 8 Fatty acid methyl Ester Preparation method 112
First methods 112
Second method 115
GC settings 116
Appendix 9 Samples for parental line genomic sequence 117
Appendix 10 *A. thaliana* QTL mapped on unigene base *B. napus* map 118
18:2 118
18:2/18:1 119
18:3 120
18:3/18:2 121
Oil content (all QTL) 122
Oil content QTL on Ch1 123
Oil content QTL on Ch2 124
Oil content QTL on Ch5 125

APPENDIX 1 LINKAGE MAPPING

CALCULATION OPTION SETTINGS

Procedure

1. Convert heterozygote into unclassified.

2. Calculate with default setting.

3. If the default setting didn't form a single group, change the calculate option step by step.

① Change LOD Grouping threshold "Lower" from 2.0 to 0

② Change mapping parameters "LOD larger than" from 1.00 to 0

③ Change mapping parameters "REC smaller than" from 0.4 to 0.5

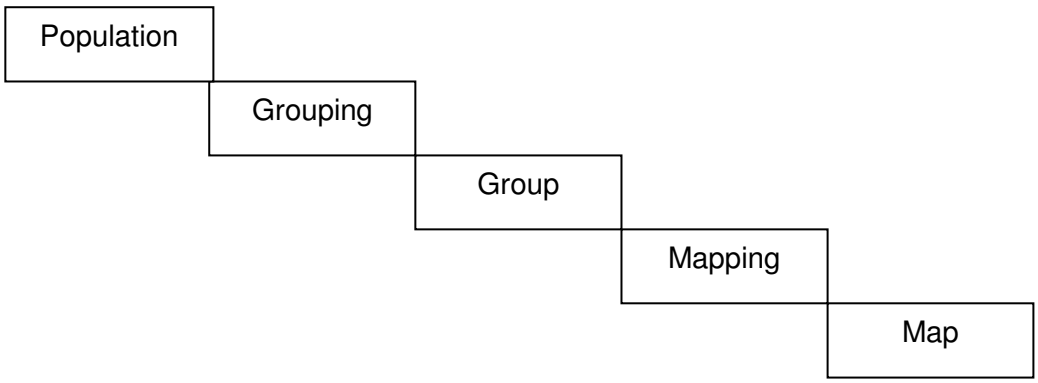
④ Fix the order according to the physical map

⑤ Remove troublesome marker (detail in page X)

⑥ If it still did not form a single group, calculate in segments and combine by fixing order as a group.

Check point

To see whether the map is statistically approved



Population

Tab "Locus genot. freq"

It shows how the segregation of allele is distorted from normal 50:50 ratio for a hybrid of RIL.

Group

Tab “Weak linkages”

It shows pairs of markers which are far away each other as the recombination rate is close to 0.5.

Tab “Maximum linkages”

It shows pairs of marker which has smaller recombination rate. That means the pair is located closer.

Tab “suspicious linkages”

The weakest recombinant rate is 0.5, however there were pairs of markers which show more than 0.5, therefore there is a suspicion that there could be error in the data.

Mapping

Tab Session log

This can tell which calculate option was used, such as ”lod threshold” and “rec threshold”.

Also by looking at mean of each round, the value can tell how the map has been improved or not when changing the calculation option and when excluding troublesome markers.

Map

Tab “Locus genot. freq”

The same information can be seen in the population section. It shows how allele segregation is distorted from normal 50:50 ratio for a hybrid of RIL.

Tab “Mean chisquare contribs”

It shows how well the marker fits with other markers.

The lower the value, the better the quality of map. If there are markers with outstanding large figure, they are the candidate markers to be removed.

Tab ”Genotype probabilities”

It shows any individual which has statistically impossible combination over a very short genotypic distance. The frequently appearing markers could be suspicious whether it was correctly genotyped or not. If requiring attention, there is a significant mark on the right side.

Tab”Locus average”

It shows which locus is better mapped. The markers with lower figures are better mapped.

Removing troublesome markers

To remove troublesome markers, check these three tabs. ⑴Mean chisquare contribs, ⑵Locus genot. freq and ⑶Genotype probabilities in Map section.

If one marker tends to cause a problem it mght not be the one in Genotype probabilities but in the chi square contribs or in Locus genot. freq. See all 3 check points and test each one, and then check the value of mean in tab session in Mapping section whether the quality of map has been improved by removing the marker

APPENDIX 2 QTL MAPPING POPULATION PHENOTYPE DATA ANALYSIS AND MAP CHARTS

ANOVA RESULTS OF SIX POPULATIONS

population	CA		SG		WC	
	residual m.s.	line x block v.r.	residual m.s.	line x block v.r.	residual m.s.	line x block v.r.
df	8	1 per line 1 per block	42	1 per line 2 per block	8	1 per line 1 per block
18:1	1.6240	0.44 ns	0.2121	0.84 ns	1.3910	0.01 ns
18:2	0.7432	0.21 ns	1.3200	0.98 ns	2.5380	0 ns
18:2/18:1	0.0992	0.2 ns	0.0111	2.59 ns	0.0612	0.04 ns
18:3	0.0662	0.56 ns	1.1100	1.09 ns	1.0810	0.06 ns
18:3/18:2	0.0030	0.02 ns	0.0032	2.7 ns	0.0008	0.19 ns
oil	9.1740	1.51 ns	0.5721	1.69 ns	2.4450	1.25 ns

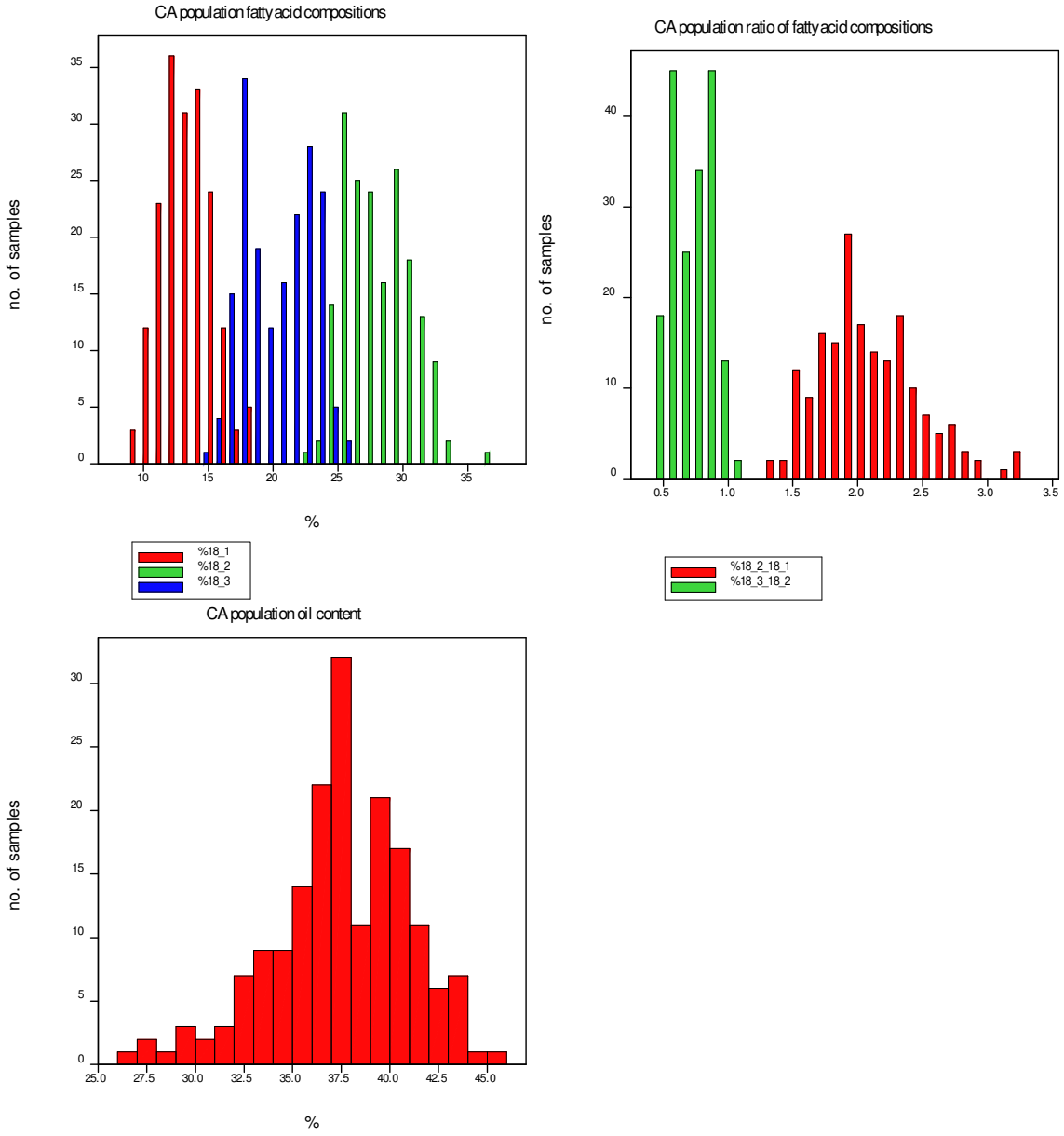
population	KB		NG		TJ	
	residual m.s.	line x block v.r.	residual m.s.	line x block v.r	residual m.s.	line x block v.r.
df	45	1 per line 2 per block	47	1 per line 2 per block	42	1 per line 2 per block
18:1	0.4908	3.34 *	0.5029	0.32 ns	0.2354	1.05 ns
18:2	0.8527	1.75 ns	0.5188	0.6 ns	0.7121	0.42 ns
18:2/18:1	0.0140	0.63 ns	0.0139	0.07 ns	0.0065	1.51 ns
18:3	0.2769	7.63 **	0.9367	3.29 *	0.2013	2.68 ns
18:3/18:2	0.0004	17.46 ***	0.0009	3.54 *	0.0007	0.13 ns
oil	1.3930	1.17 ns	6.1850	3.51 *	1.0080	1.76 ns

PROPORTION OF HERITABLE VARIATION OF SIX POPULATION

Population	Trait	Variety ms*	Residual ms*	Block	Variance from genotype	Proportion of heritable variation
CA	18:1	5.8790	1.6050	2	2.1370	57.1
	18:2	11.1080	1.2640	2	4.9220	79.6
	18:2/18:1	0.2364	0.0670	2	0.0847	55.8
	18:3	12.9821	0.7384	2	6.1219	89.2
	18:3/18:2	0.0420	0.0036	2	0.0192	84.1
	Oil	14.7960	9.7160	2	2.5400	20.7
SG	18:1	8.5950	1.1870	3	2.4693	67.5
	18:2	7.7010	4.4940	3	1.0690	19.2
	18:2/18:1	0.0925	0.0486	3	0.0146	23.1
	18:3	2.6370	2.0800	3	0.1857	8.2
	18:3/18:2	0.0107	0.0017	3	0.0030	64.4
	Oil	6.4460	1.3850	3	1.6870	54.9
WC	18:1	5.1170	2.0520	2	1.5325	42.8
	18:2	2.2670	2.2510	2	0.0080	0.4
	18:2/18:1	0.1263	0.0846	2	0.0208	19.8
	18:3	3.4190	1.9460	2	0.7365	27.5
	18:3/18:2	0.0073	0.0023	2	0.0025	51.6
	Oil	12.3560	6.0040	2	3.1760	34.6
KB	18:1	11.2656	0.7076	3	3.5193	83.3
	18:2	14.1600	2.9880	3	3.7240	55.5
	18:2/18:1	0.1928	0.0343	3	0.0529	60.7
	18:3	4.5490	1.9360	3	0.8710	31.0
	18:3/18:2	0.0216	0.0032	3	0.0061	65.4
	Oil	6.7620	1.2090	3	1.8510	60.5
NG	18:1	5.9040	1.1910	3	1.5710	56.9
	18:2	3.0250	1.0920	3	0.6443	37.1
	18:2/18:1	0.1203	0.0232	3	0.0324	58.2
	18:3	3.9920	2.6230	3	0.4563	14.8
	18:3/18:2	0.0131	0.0040	3	0.0030	43.4
	Oil	6.3320	5.5400	3	0.2640	4.5
TJ	18:1	3.5660	1.0920	3	0.8247	43.0
	18:2	2.7830	1.5260	3	0.4190	21.5
	18:2/18:1	0.0356	0.0196	3	0.0053	21.4
	18:3	3.4250	1.2150	3	0.7367	37.7
	18:3/18:2	0.0098	0.0042	3	0.0019	31.3
	Oil	5.5520	1.9170	3	1.2117	38.7

* ms=mean square

CA POPULATION HISTOGRAM



CA POPULATION NORMALITY TEST

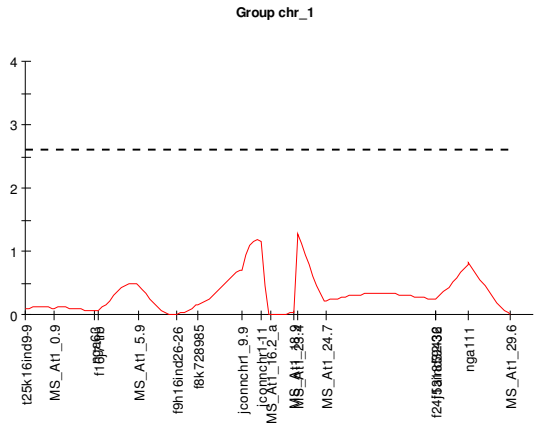
CA original data
Shapiro-Wilk test for Normality

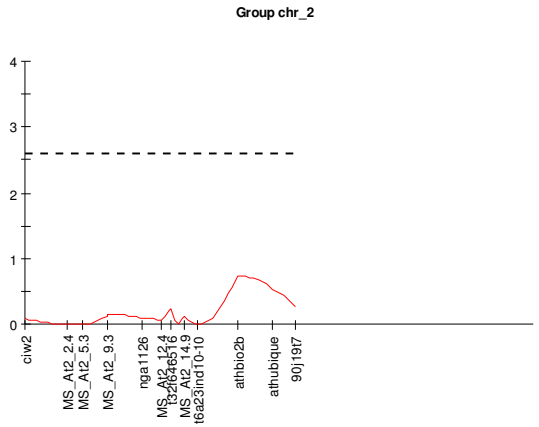
Data variate:	18:1
Test statistic W:	0.9844
Probability:	0.33
Data variate:	18:02
Test statistic W:	0.9678
Probability:	0.02
Data variate:	18:2/18:1
Test statistic W:	0.9773
Probability:	0.101
Data variate:	18:03
Test statistic W:	0.932
Probability:	<0.001
Data variate:	18:3/18:2
Test statistic W:	0.9295
Probability:	<0.001
Data variate:	Oil
Test statistic W:	0.9895
Probability:	0.664

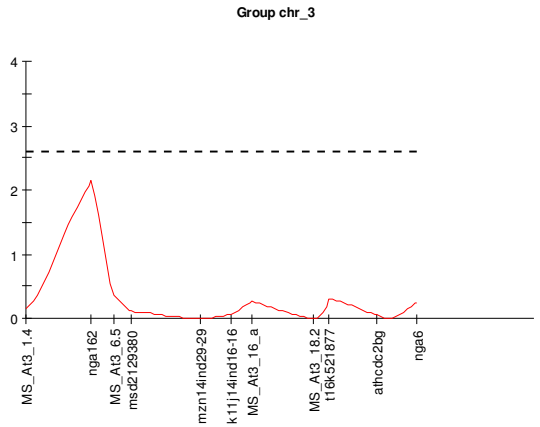
CA transformed data
Shapiro-Wilk test for Normality

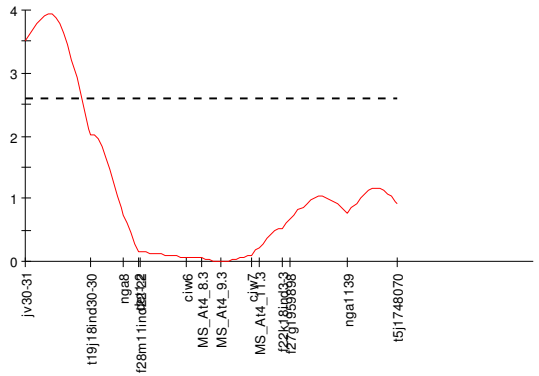
Data variate:	square root:18:2
Test statistic W:	0.9706
Probability:	0.032
Data variate:	square root:18:3
Test statistic W:	0.9319
Probability:	<0.001
Data variate:	square root:18:3/18:2
Test statistic W:	0.9281
Probability:	<0.001
Data variate:	log10:18:2
Test statistic W:	0.9727
Probability:	0.046
Data variate:	log10:18:3
Test statistic W:	0.931
Probability:	<0.001
Data variate:	log10:18:3/18:2
Test statistic W:	0.925
Probability:	<0.001

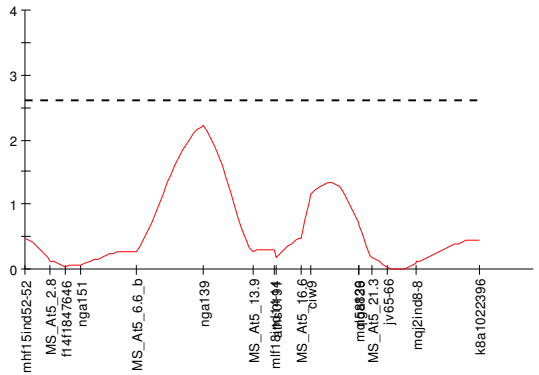
CA POPULATION QTL MAP CHARTS 18:1



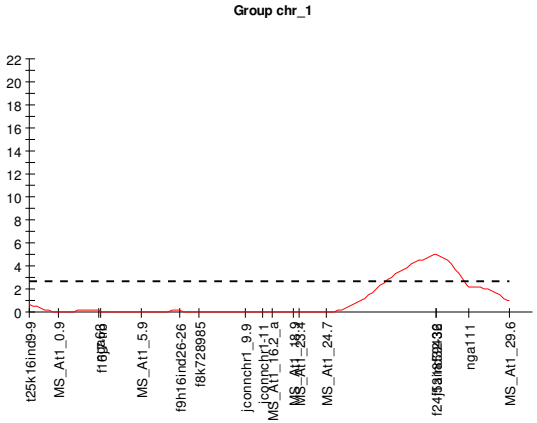


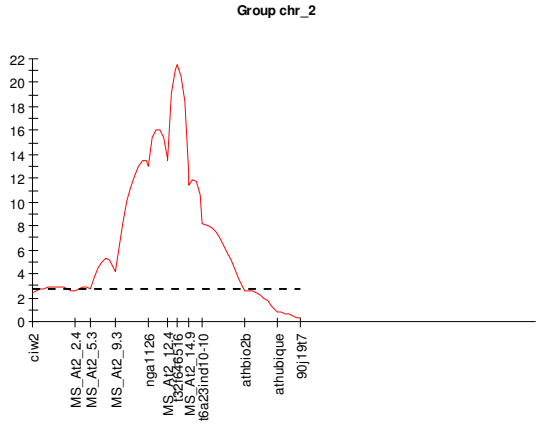


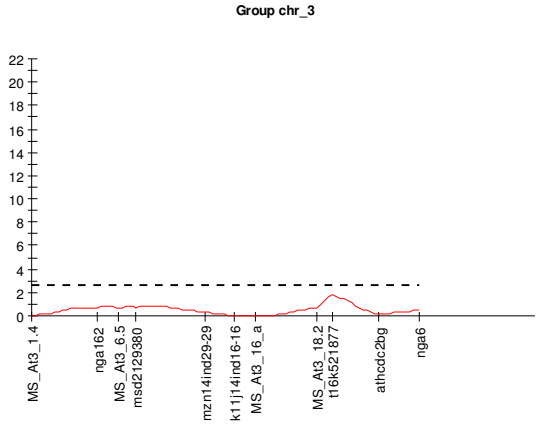


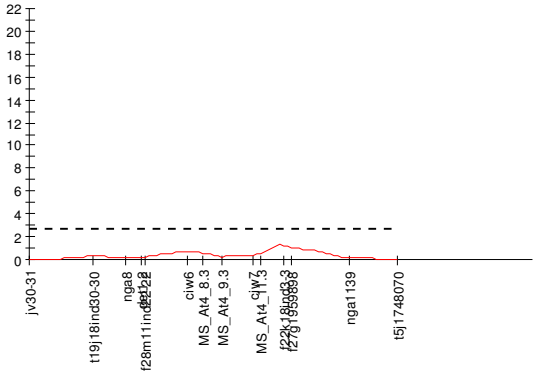


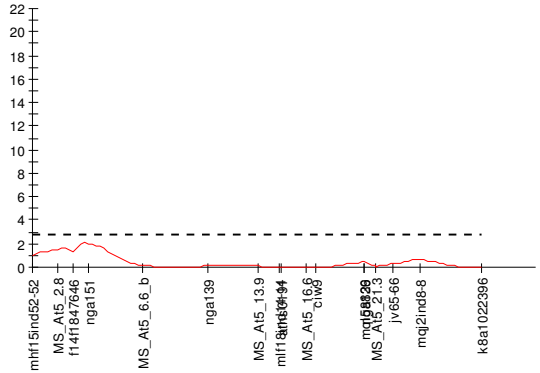
CA POPULATION QTL MAP CHARTS 18:2



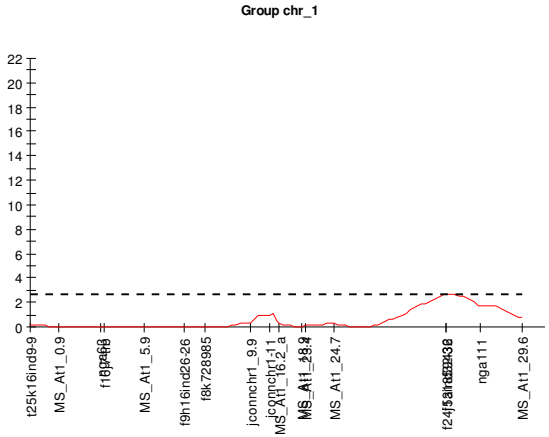


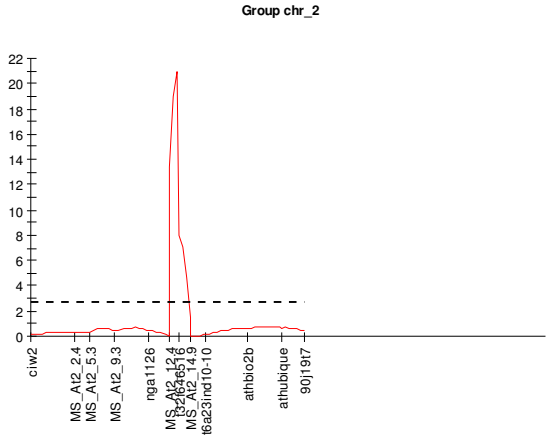


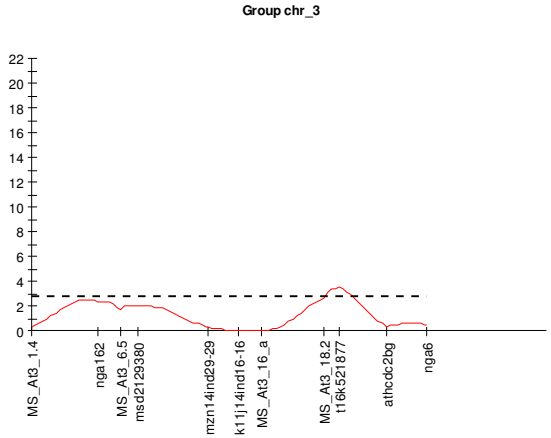


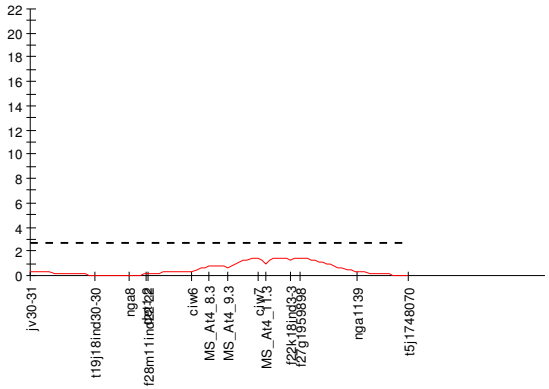


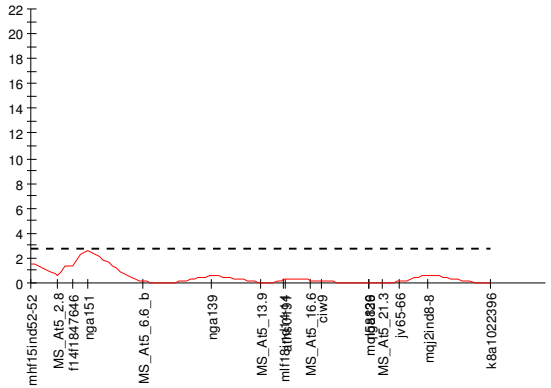
CA POPULATION QTL MAP CHARTS 18:2 MQM MAPPING







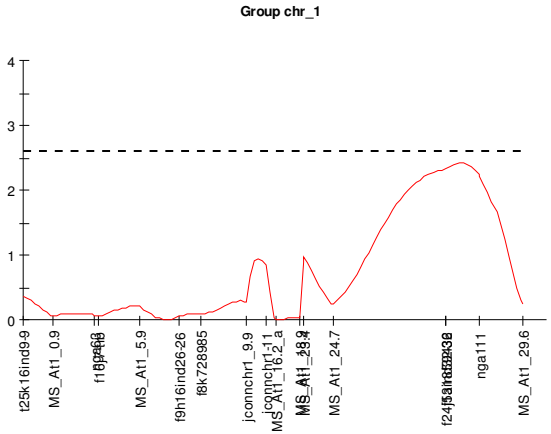


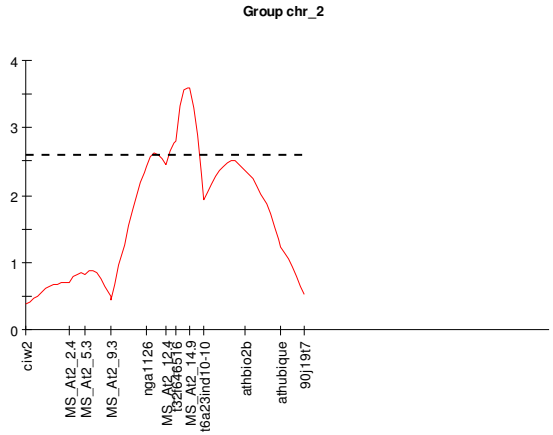


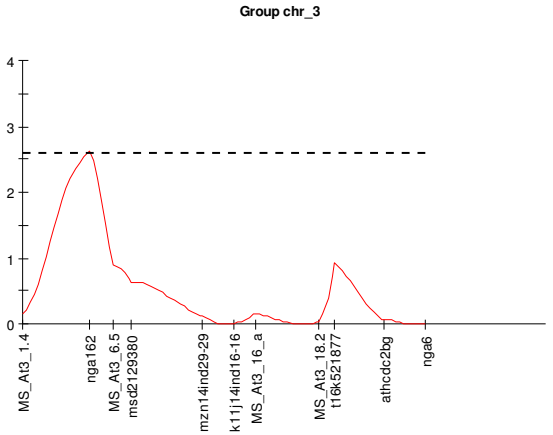
Cofactors (LG2) *FAD3* loci

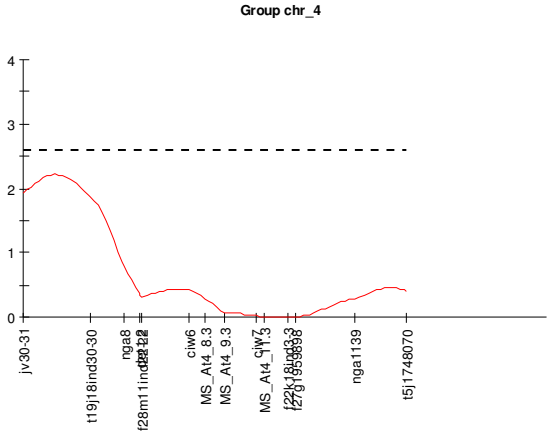
MS_At2_12.4 and
t32f646516

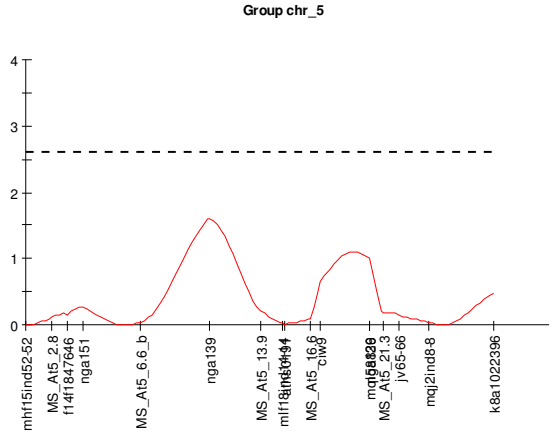
CA POPULATION QTL MAP CHARTS 18:2/18:1



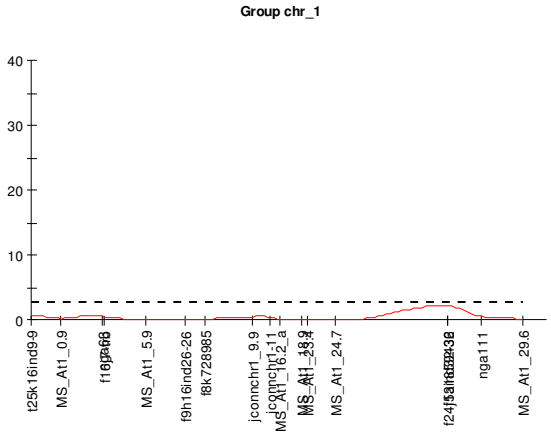


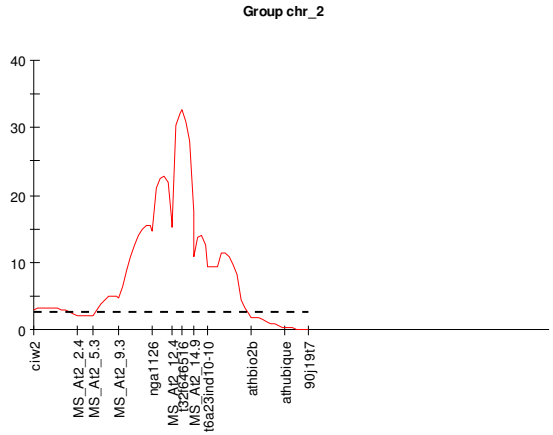


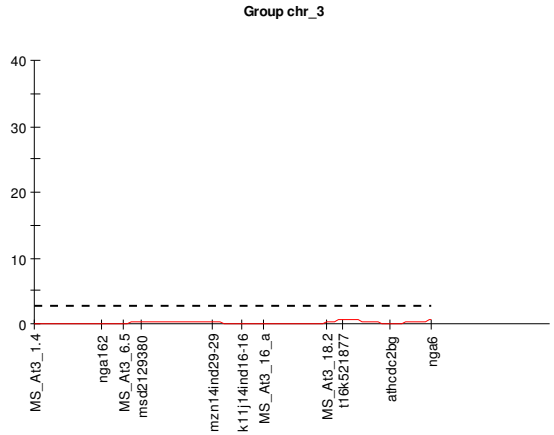


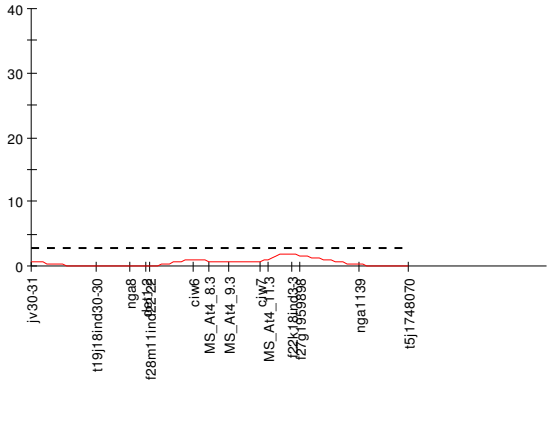


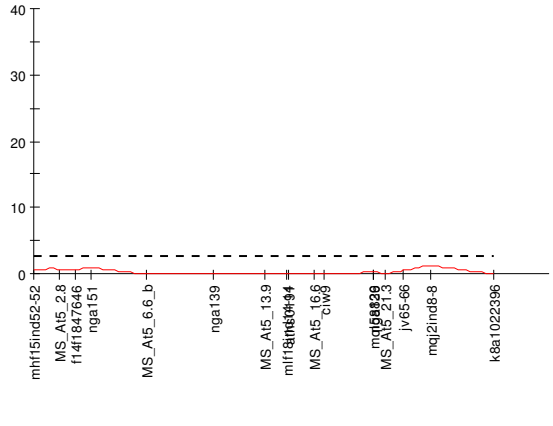
CA POPULATION QTL MAP CHARTS 18:3



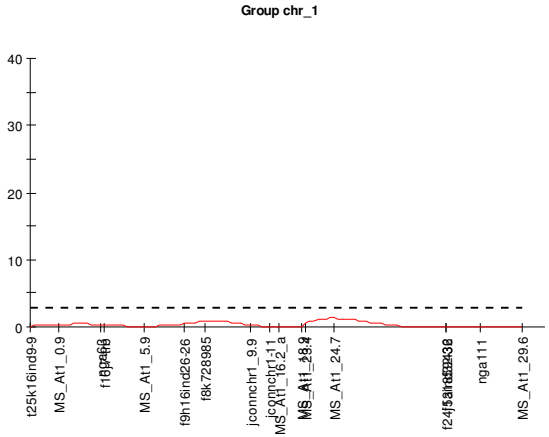


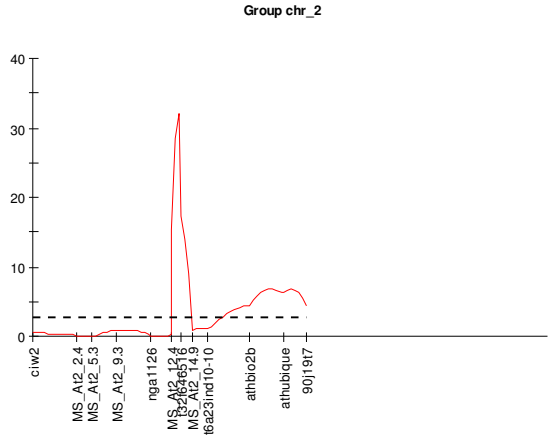


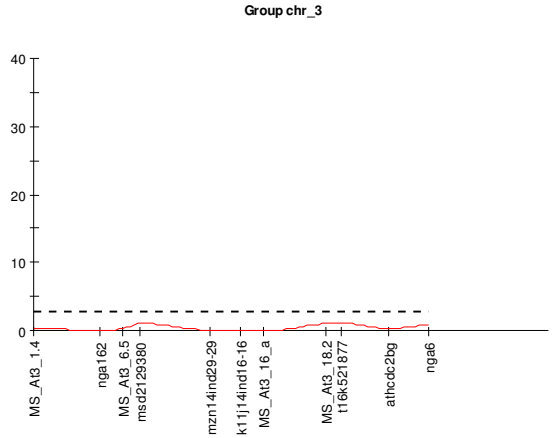


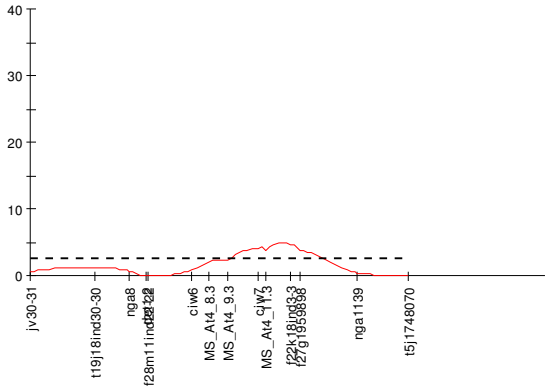


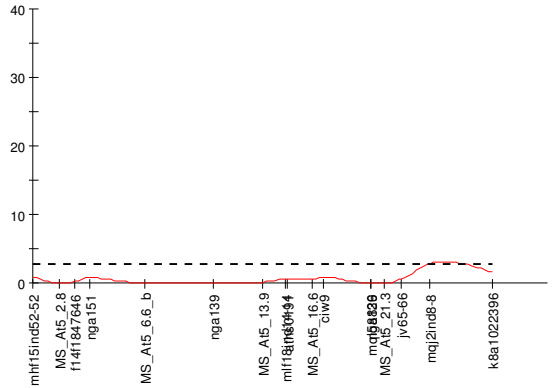
CA POPULATION QTL MAP CHARTS 18:3 MQM MAPPING







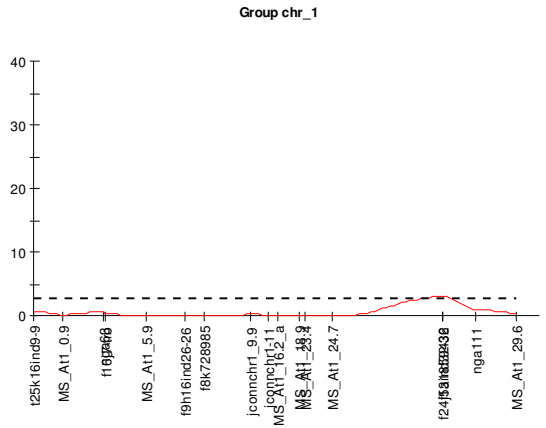


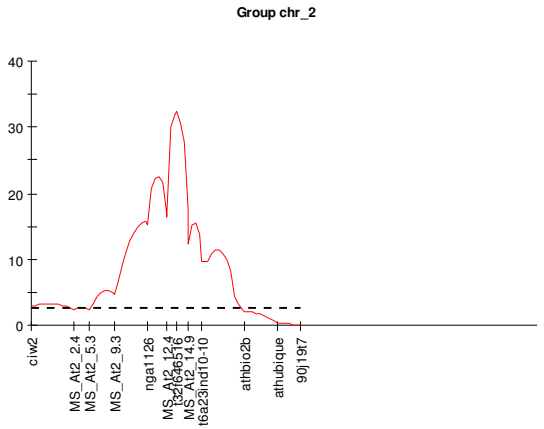


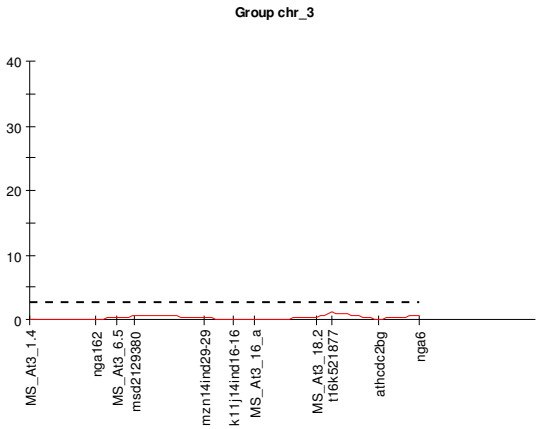
Cofactors (LG2) *FAD3* loci

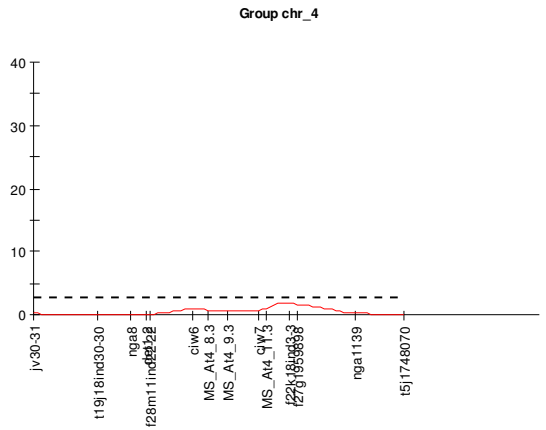
MS_At2_12.4 and t32f646516

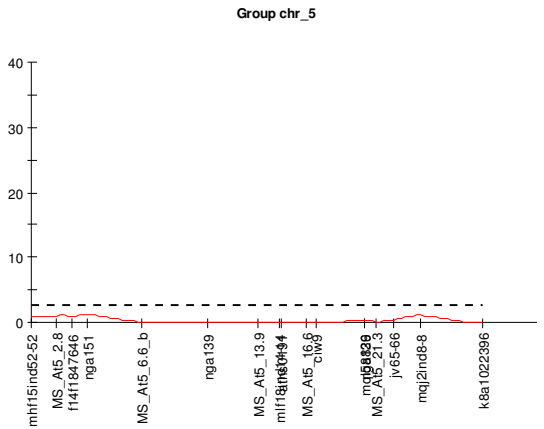
CA POPULATION QTL MAP CHARTS 18:3/18:2



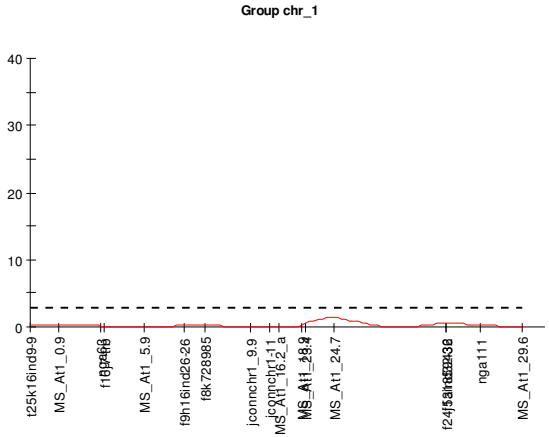


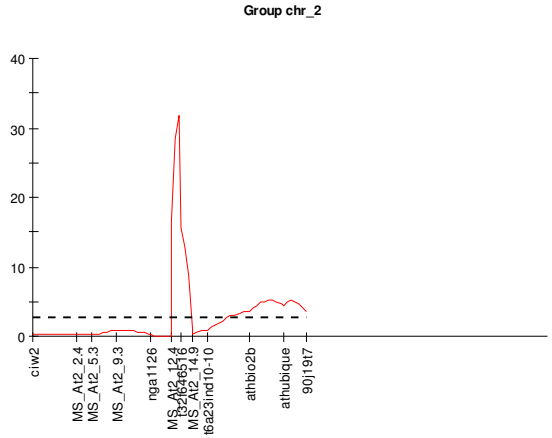


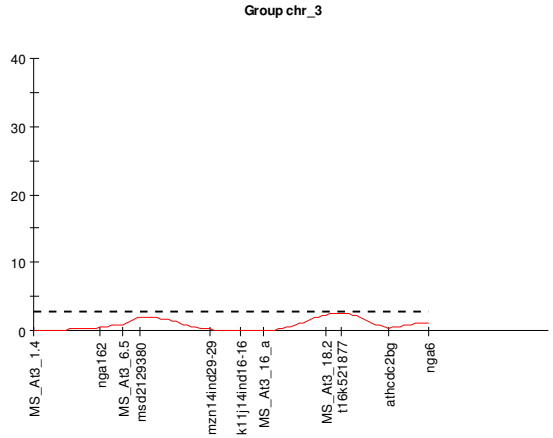


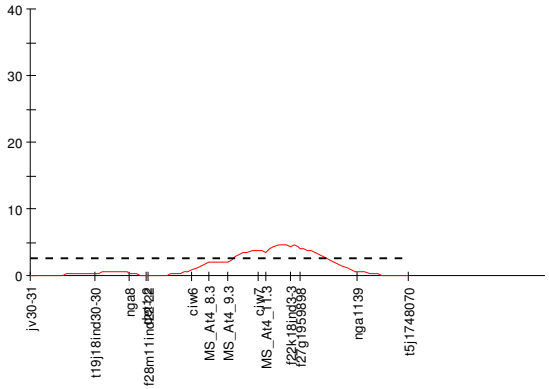


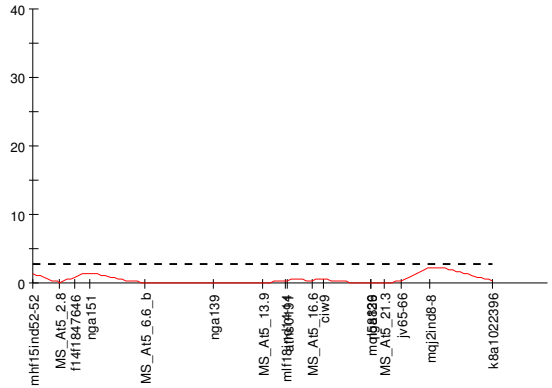
CA POPULATION QTL MAP CHARTS 18:3/18:2







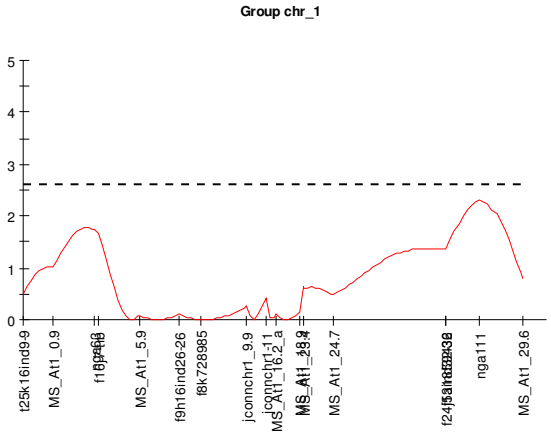


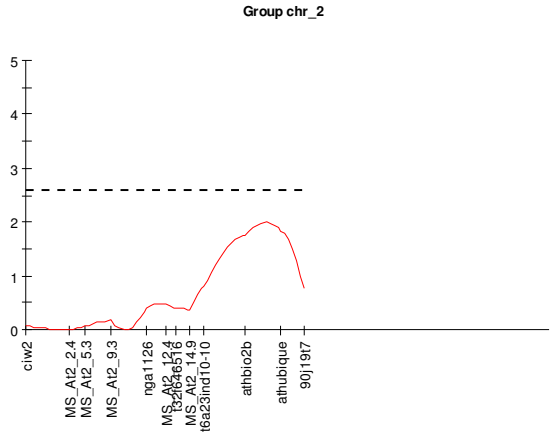


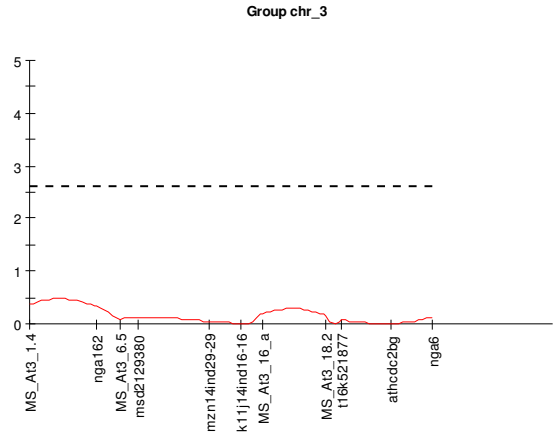
Cofactors (LG2) *FAD3* loci

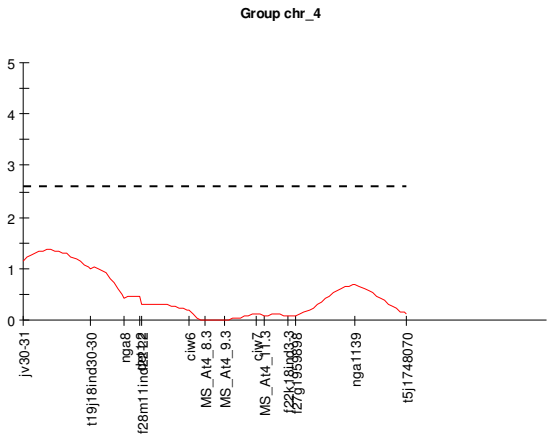
MS_At2_12.4 and t32f646516

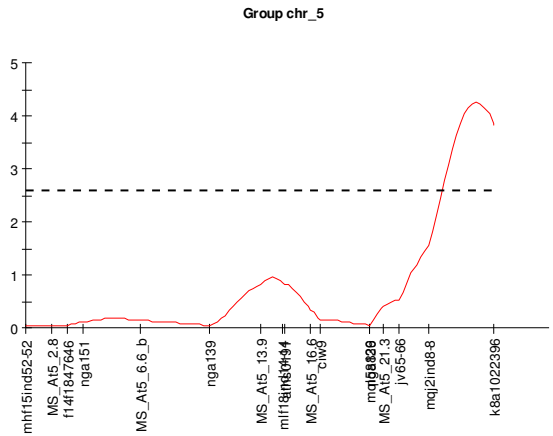
CA POPULATION QTL MAP CHARTS OIL CONTENT



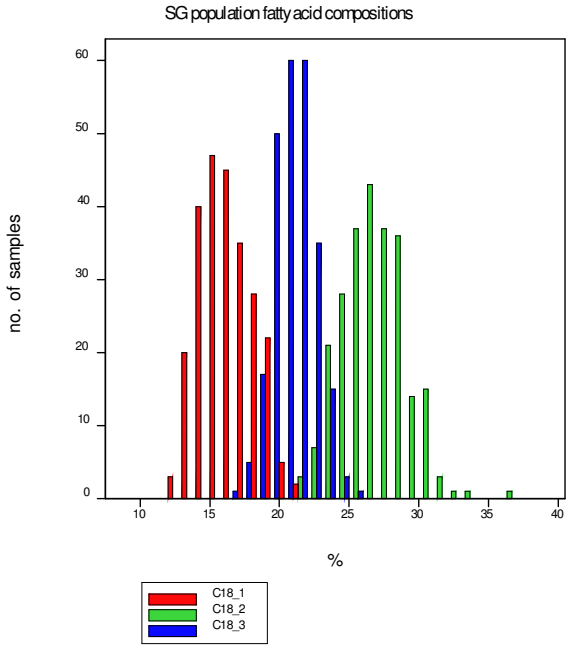


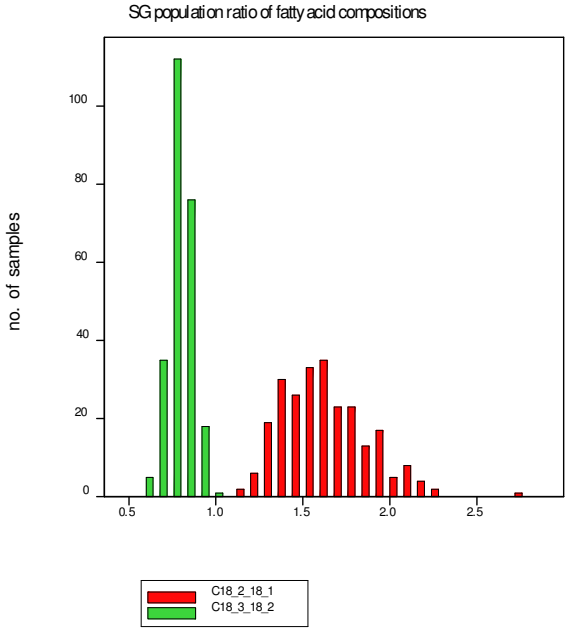


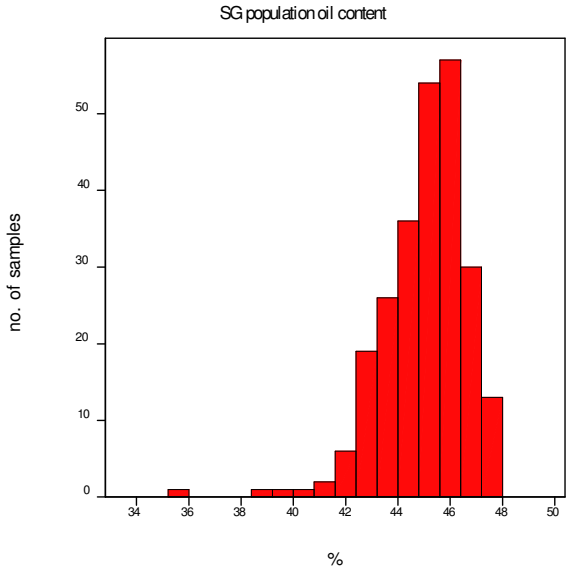




SG POPULATION HISTOGRAM







SG POPULATION NORMALITY TEST

SG original data
Shapiro-Wilk test for Normality

Data variate:	18:01
Test statistic W:	0.9464
Probability:	<0.001

Data variate:	18:02
Test statistic W:	0.9839
Probability:	0.302

Data variate:	18:2/18:1
Test statistic W:	0.984
Probability:	0.309

Data variate:	18:03
Test statistic W:	0.9884
Probability:	0.584

Data variate:	18:3/18:2
Test statistic W:	0.9772
Probability:	0.1

Data variate:	Oil
Test statistic W:	0.9177
Probability:	<0.001

SG transformed data
Shapiro-Wilk test for Normality

Data variate:	square root:18:1
Test statistic W:	0.9521
Probability:	0.002

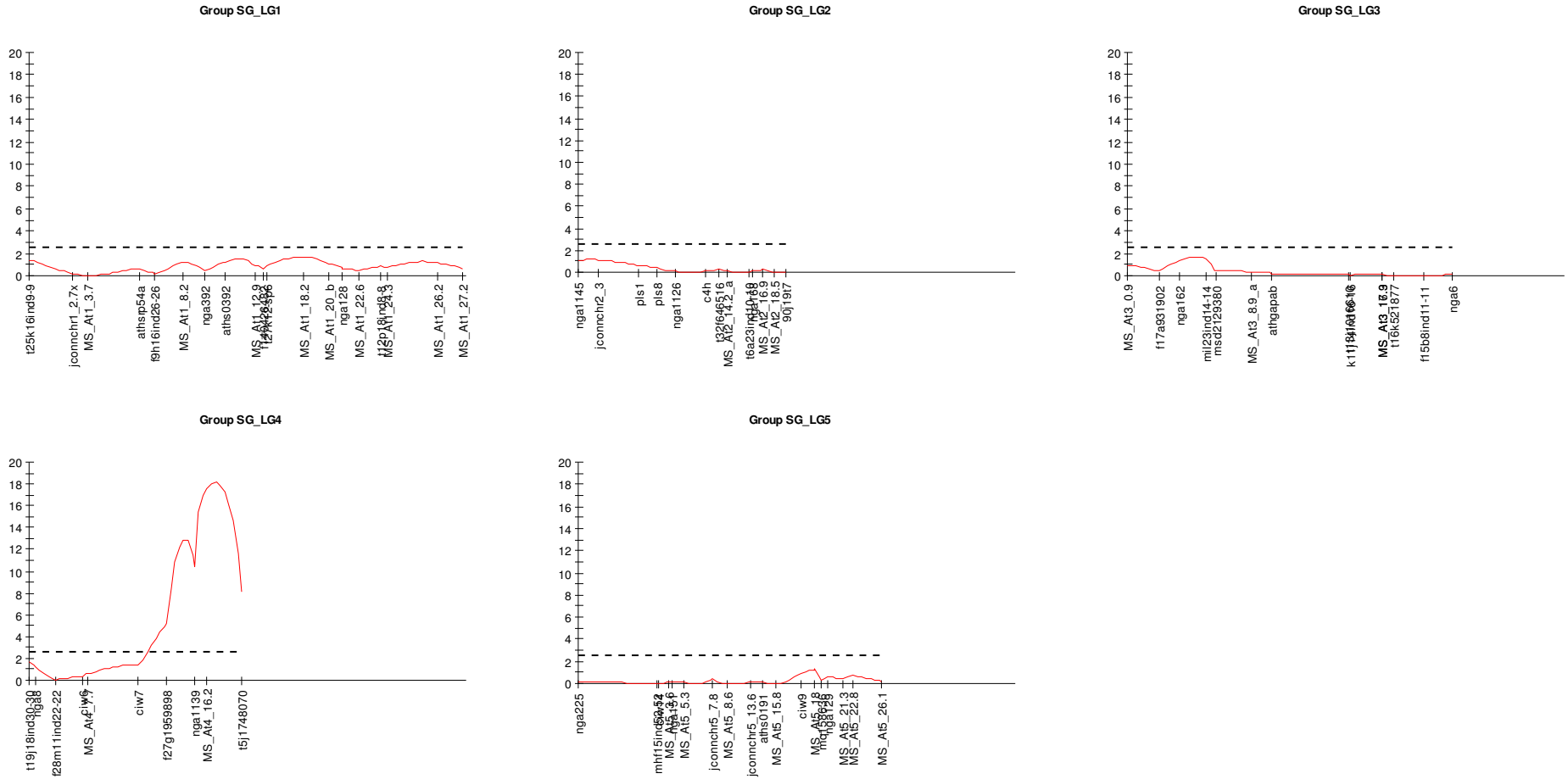
Data variate:	square root:Oil
Test statistic W:	0.9092
Probability:	<0.001

Data variate:	log10:18:1
Test statistic W:	0.9568
Probability:	0.003

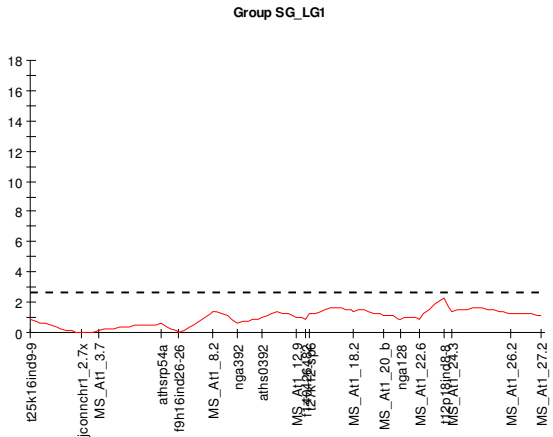
Data variate:	log10:Oil
Test statistic W:	0.9
Probability:	<0.001

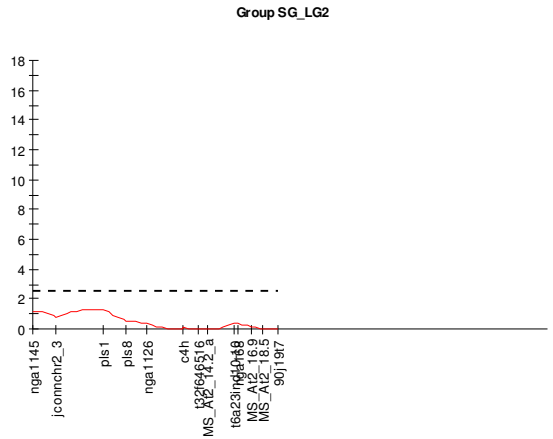
Data variate:	Oil
Test statistic W:	0.9802
Probability:	0.176

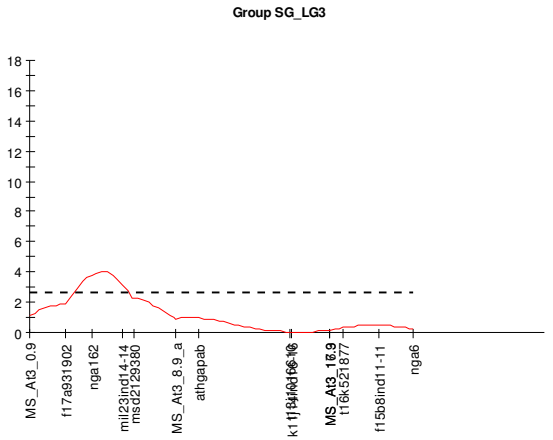
SG POPULATION QTL MAP CHART 18:1

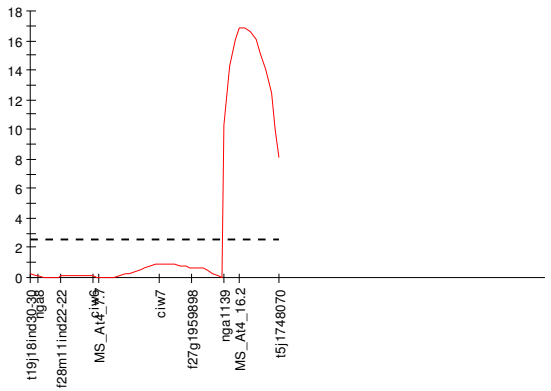


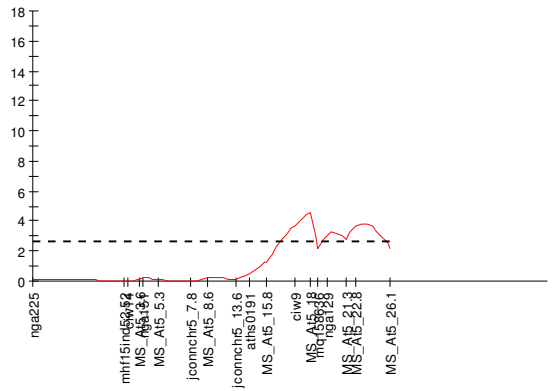
SG POPULATION QTL MAP CHART 18:1 MQM MAPPING







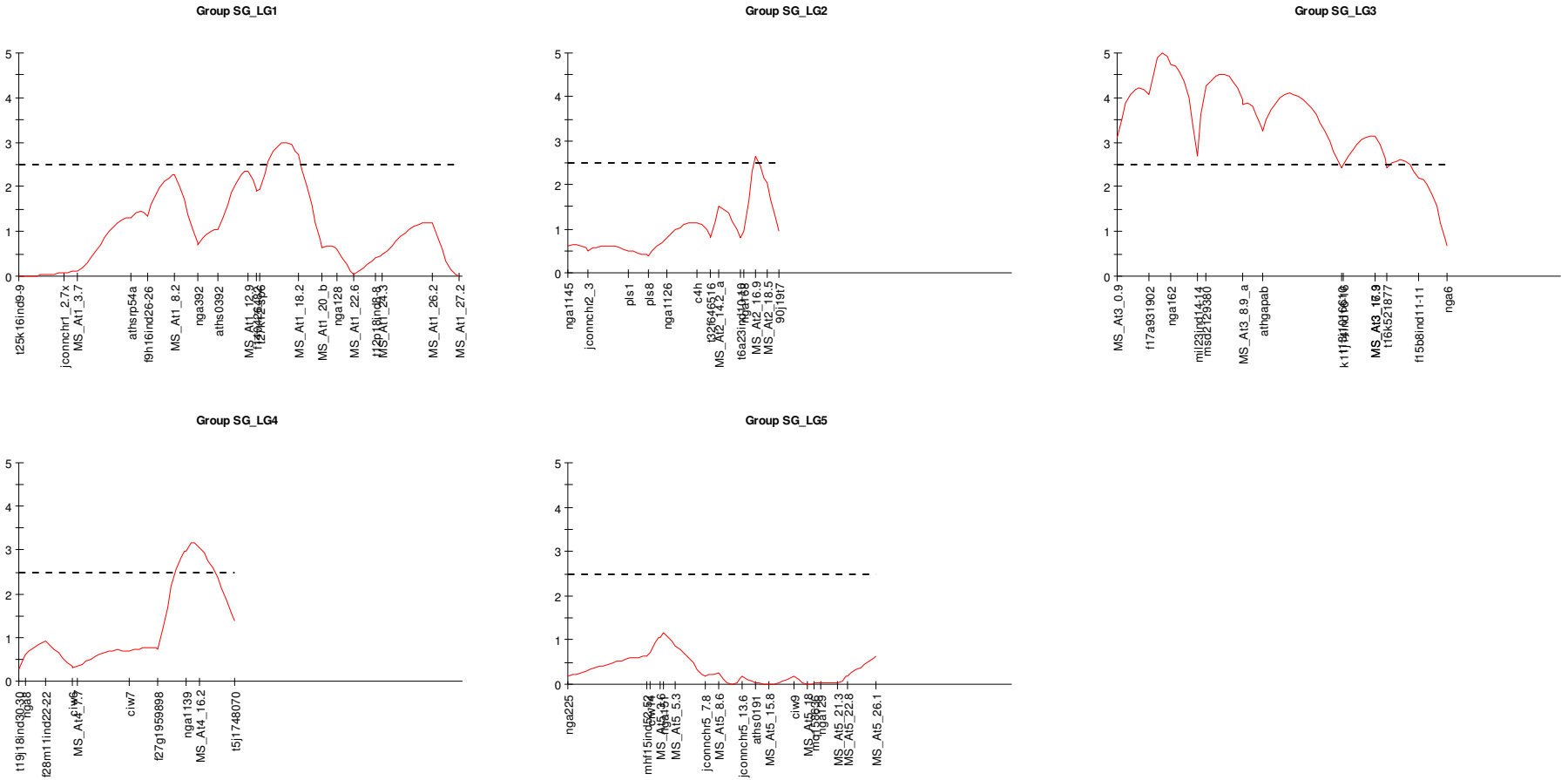




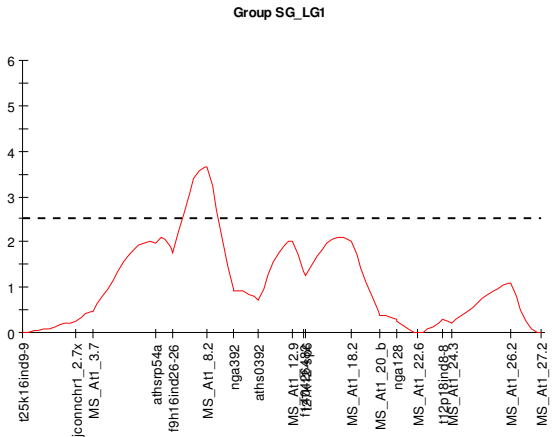
Cofactor *FAE1* locus

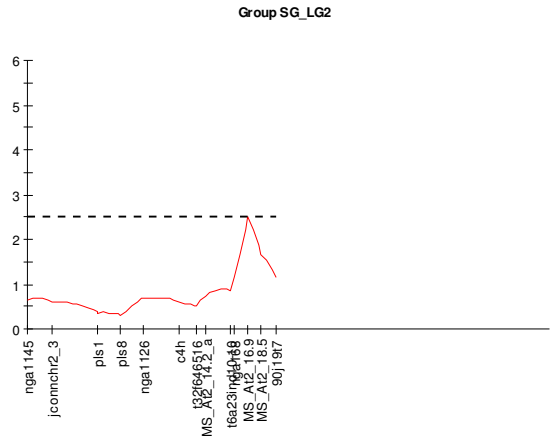
MS_At4_16.2 on LG4

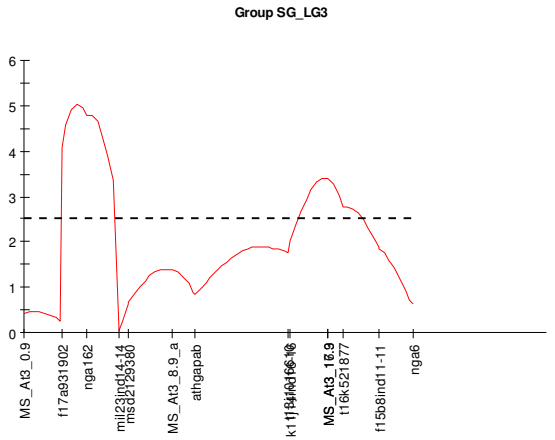
SG POPULATION QTL MAP CHART 18:2

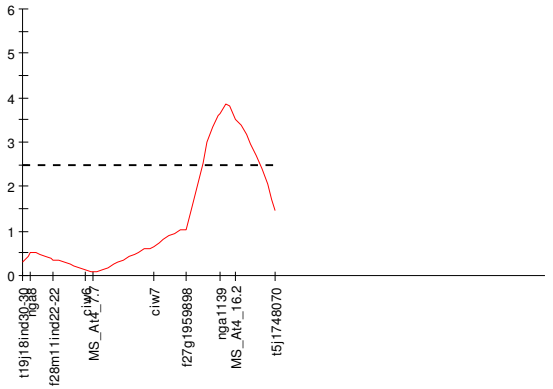


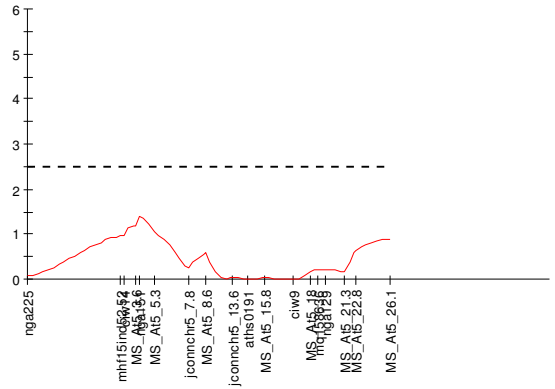
SG POPULATION QTL MAP CHART 18:2 MQM MAPPING







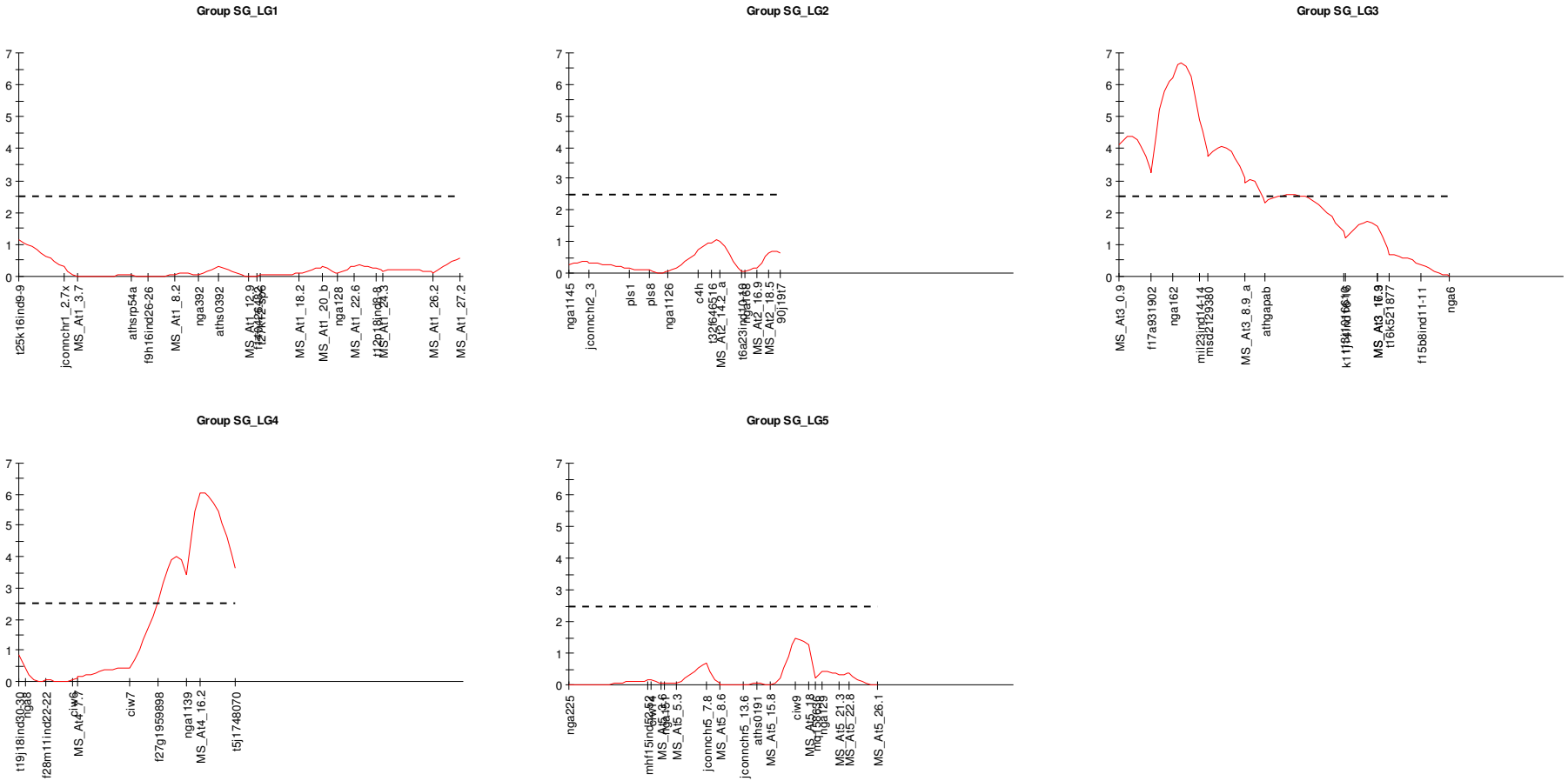




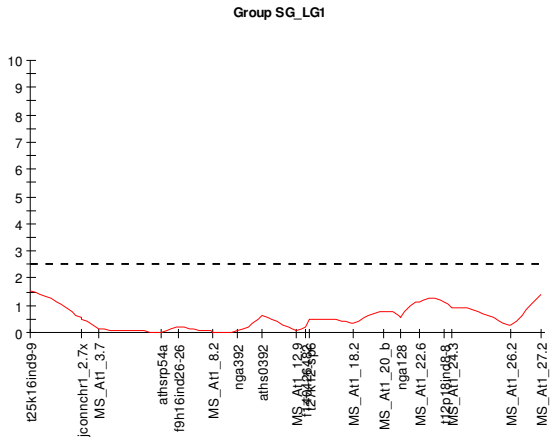
Cofactor *FAD2* locus

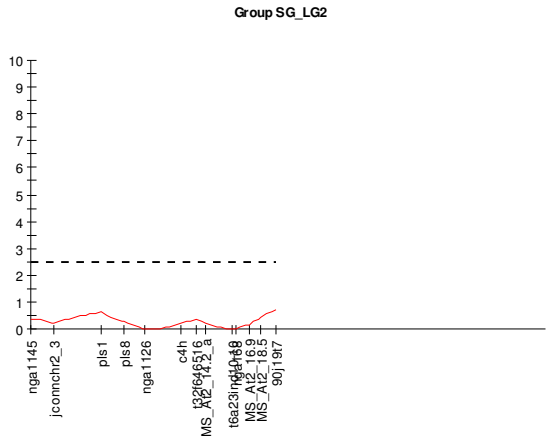
nga162 on LG3

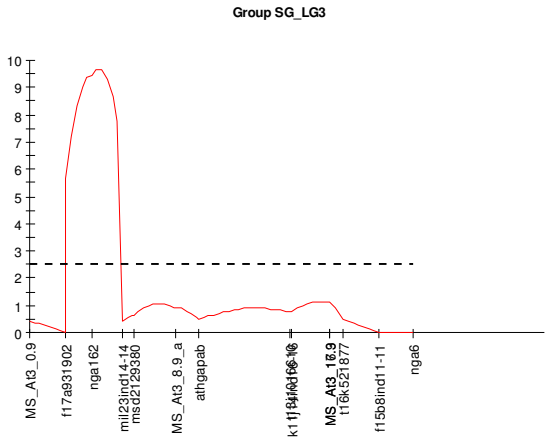
SG POPULATION QTL MAP CHART 18:2/18:1

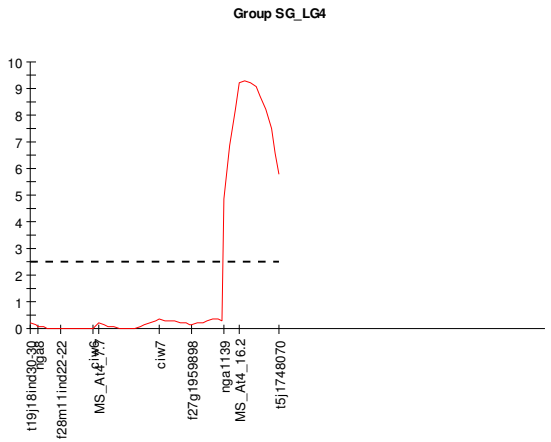


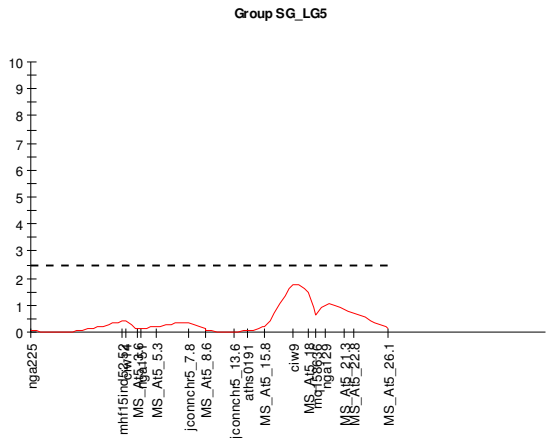
SG POPULATION QTL MAP CHART 18:2/18:1 MQM MAPPING









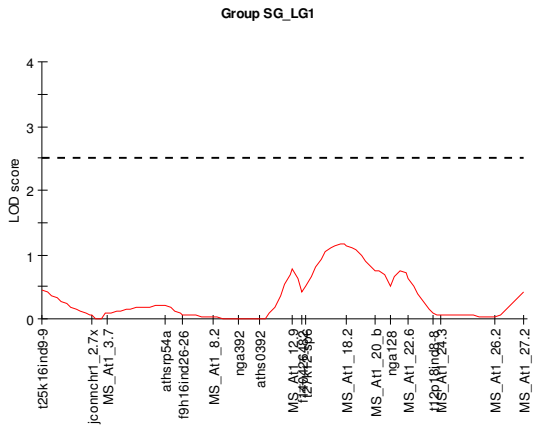


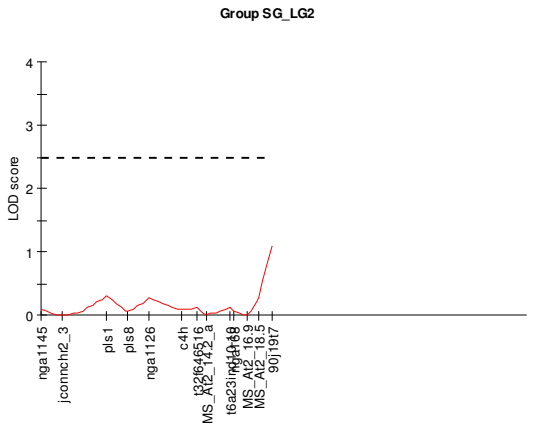
Cofactor *FAD2* locus

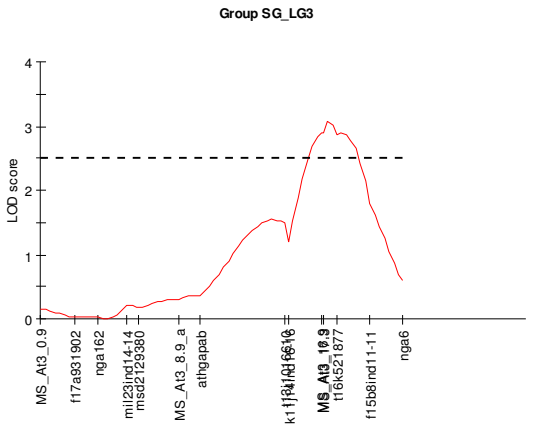
nga162 on LG3

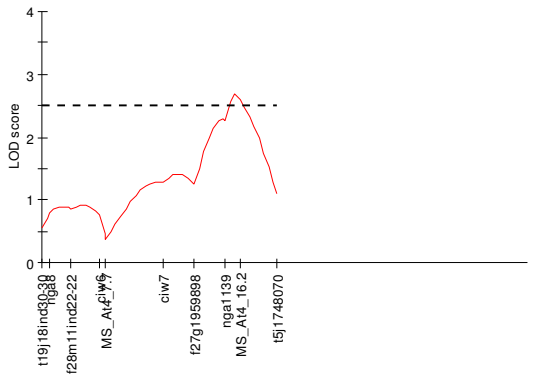
MS_At4_16.2 on LG4

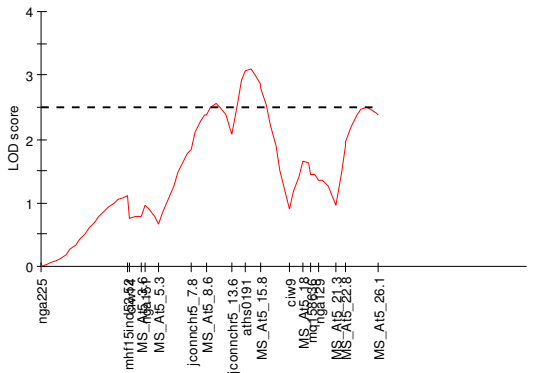
SG POPULATION QTL MAP CHART 18:3



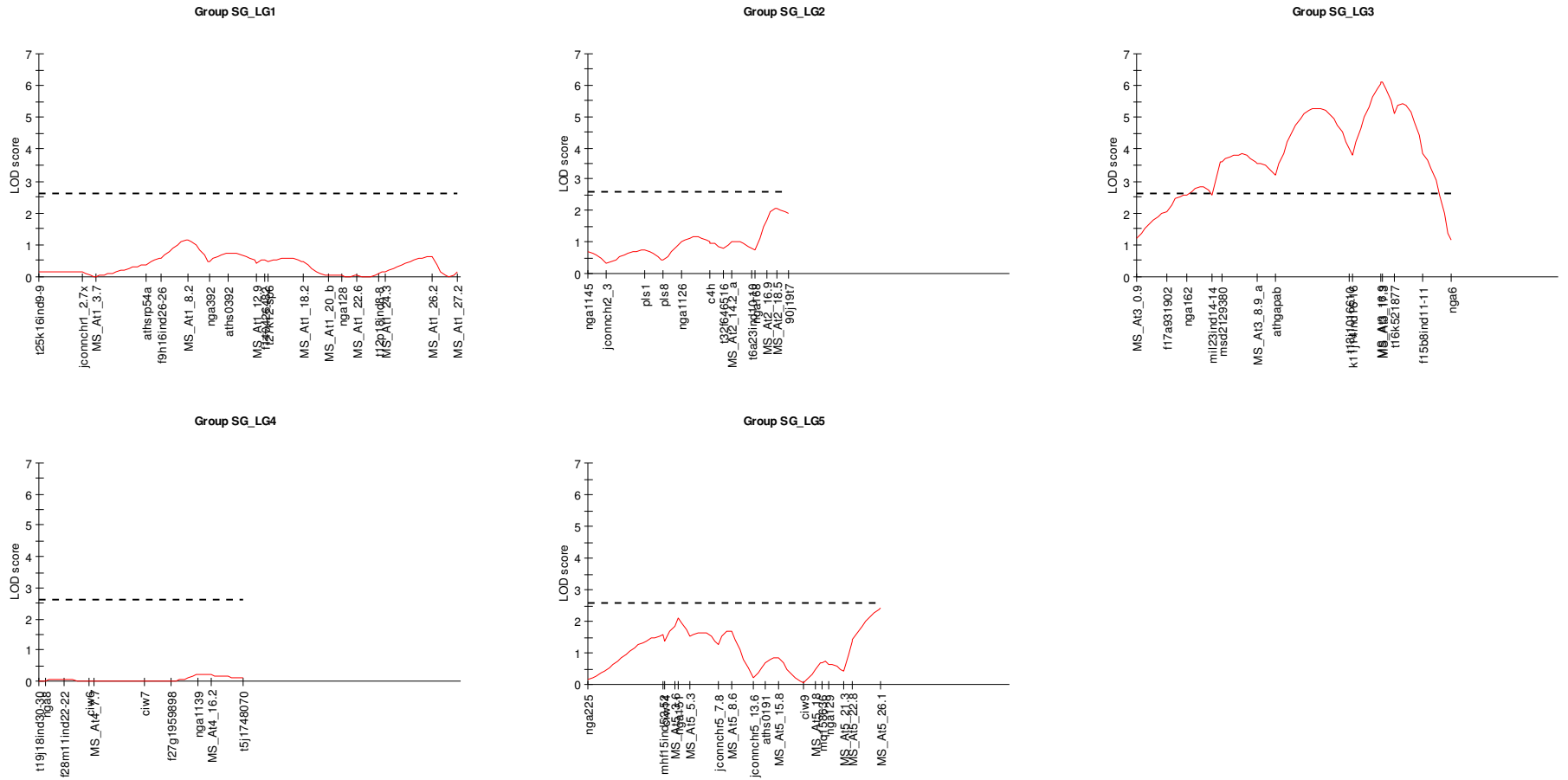




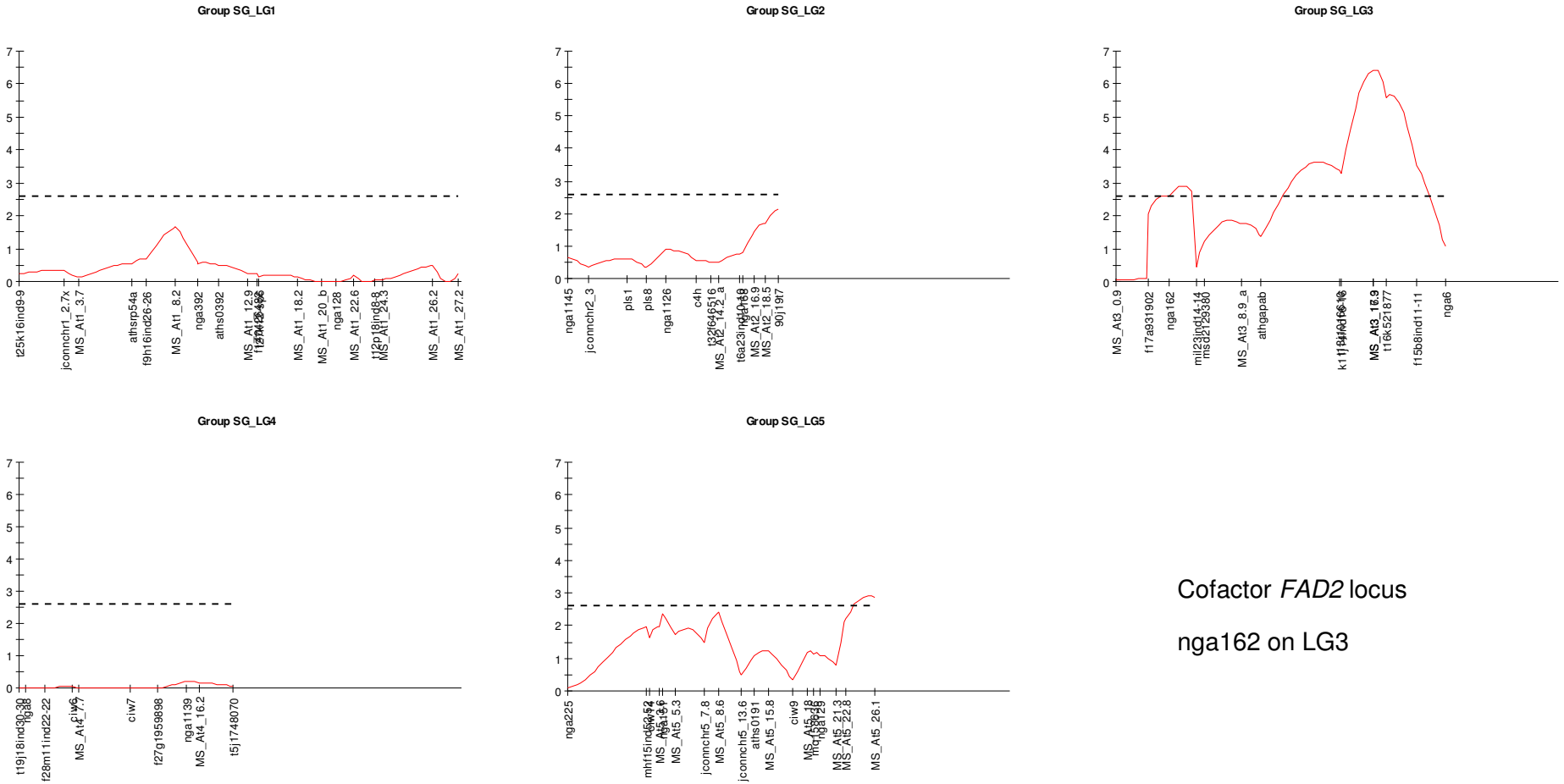




SG POPULATION QTL MAP CHART 18:3/18:2



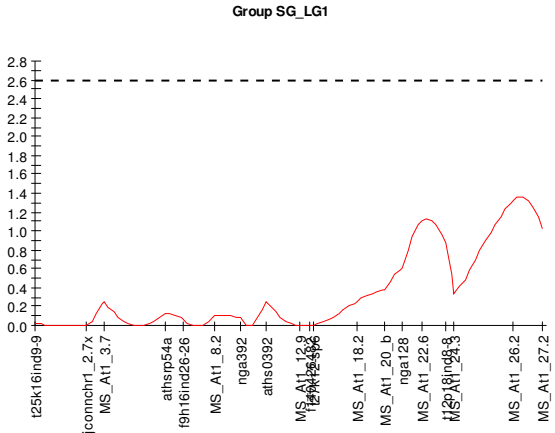
SG POPULATION QTL MAP CHART 18:3/18:2 MQM MAPPING

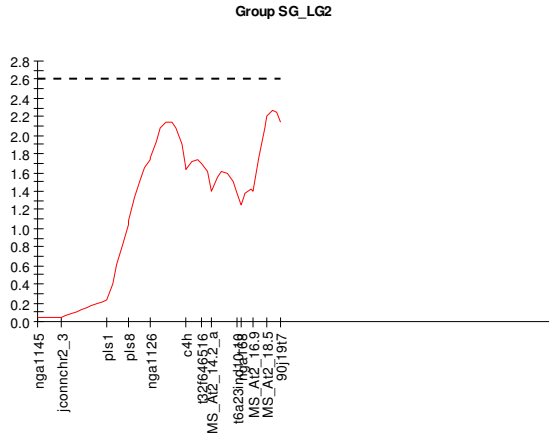


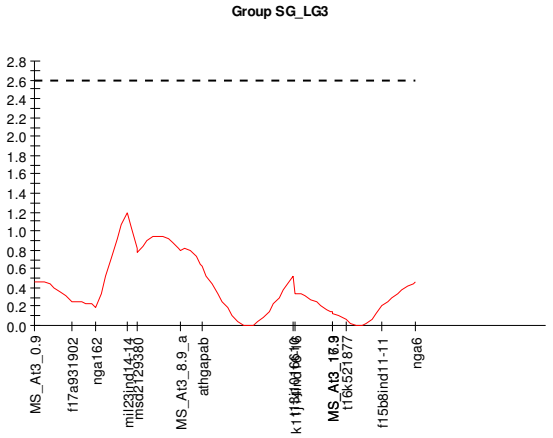
Cofactor *FAD2* locus

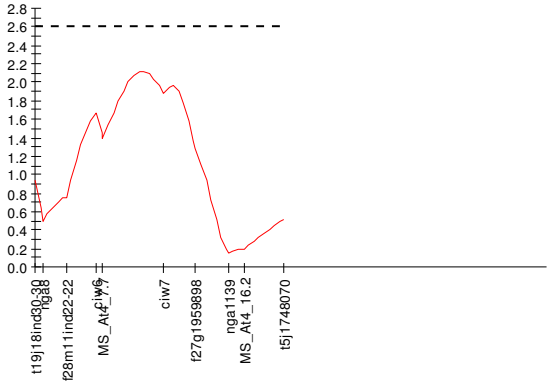
nga162 on LG3

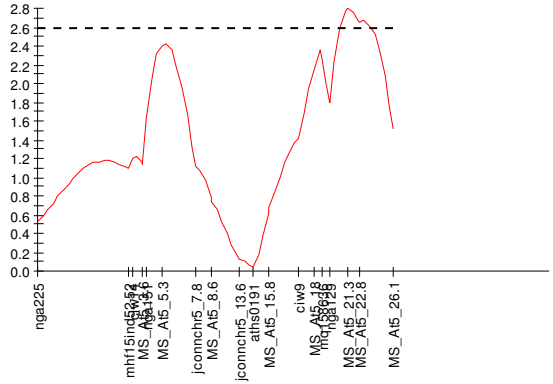
SG POPULATION QTL MAP CHART OIL CONTENT



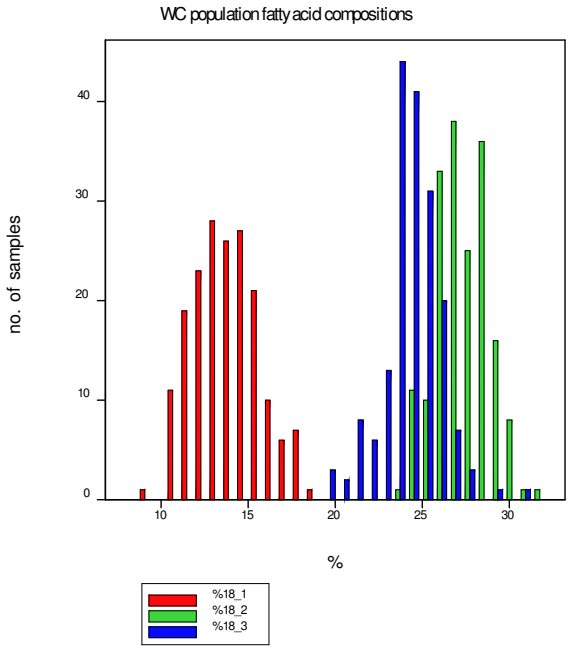


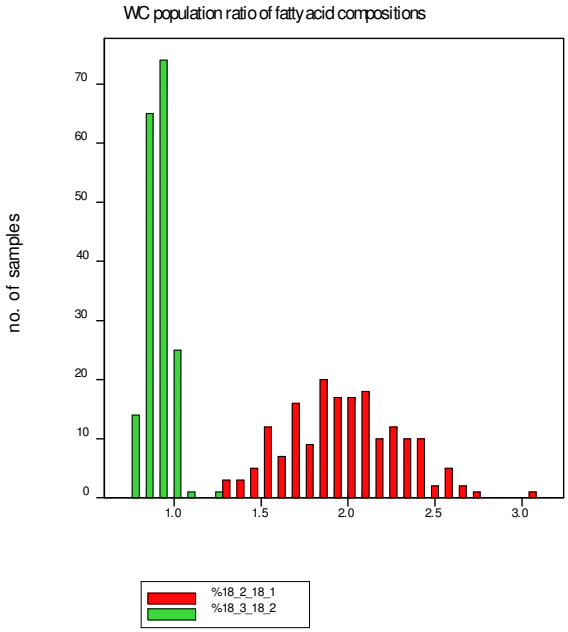


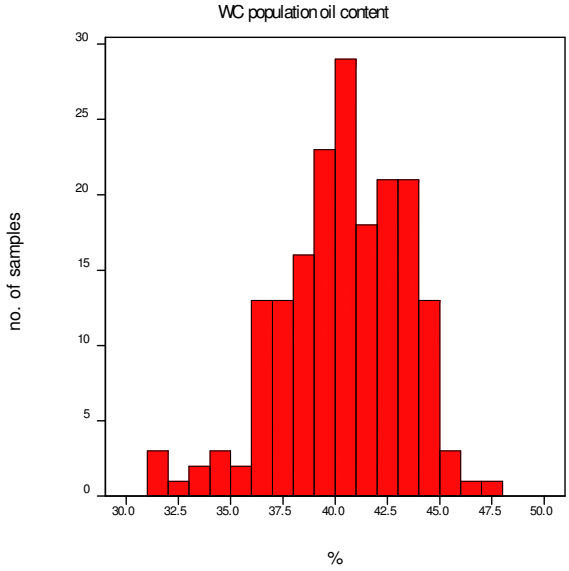




WC POPULATION HISTOGRAM







WC POPULATION NORMALITY TEST

WC original data
Shapiro-Wilk test for Normality

Data variate:	18:01
Test statistic W:	0.9925
Probability:	0.879

Data variate:	18:02
Test statistic W:	0.9903
Probability:	0.724

Data variate:	18:2/18:1
Test statistic W:	0.9916
Probability:	0.821

Data variate:	18:03
Test statistic W:	0.9706
Probability:	0.032

Data variate:	18:3/18:2
Test statistic W:	0.9866
Probability:	0.455

Data variate:	Oil
Test statistic W:	0.98
Probability:	0.161

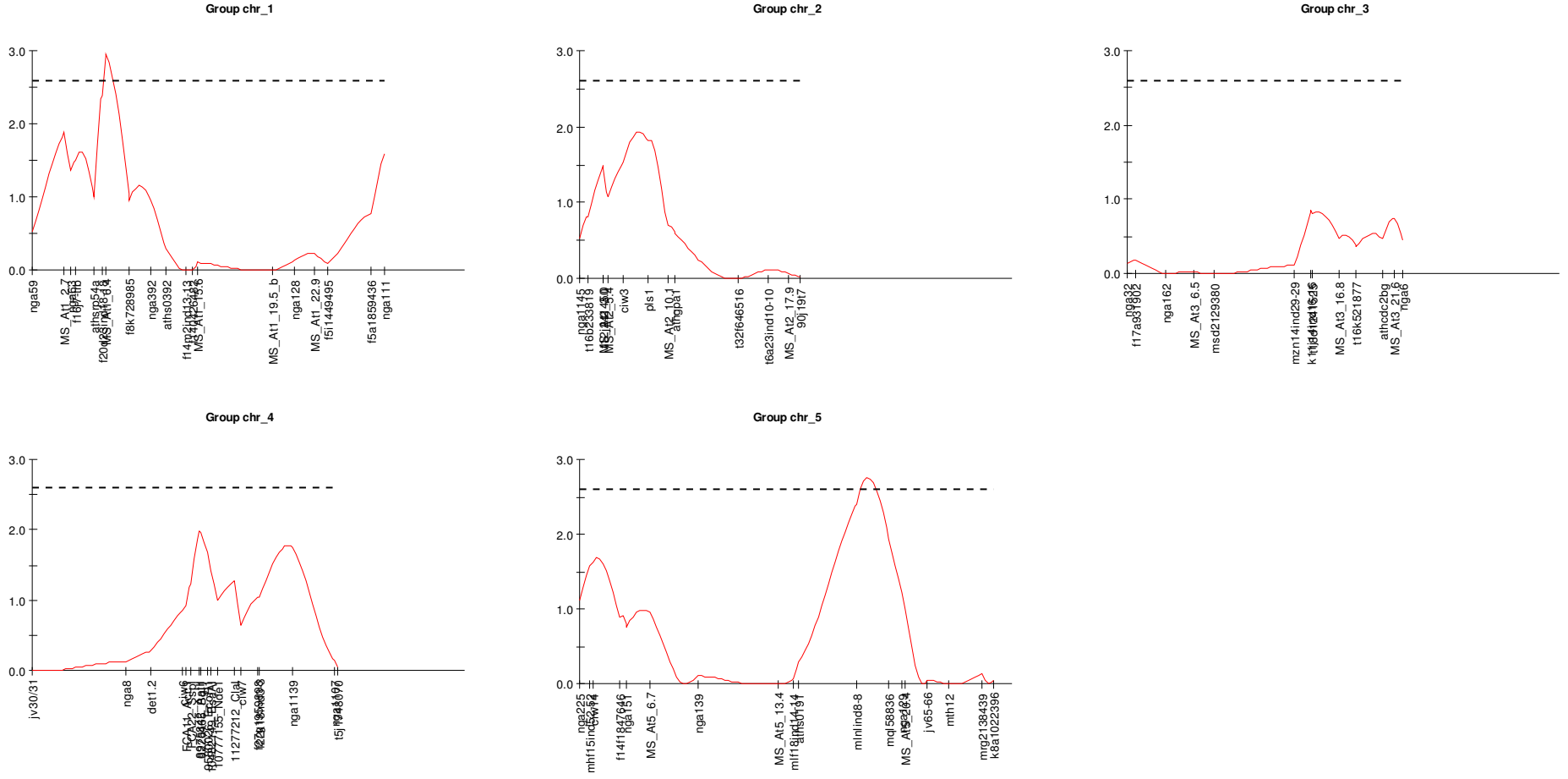
WC transformed data
Shapiro-Wilk test for Normality

Data variate:	square root:18:3
Test statistic W:	0.9726
Probability:	0.045

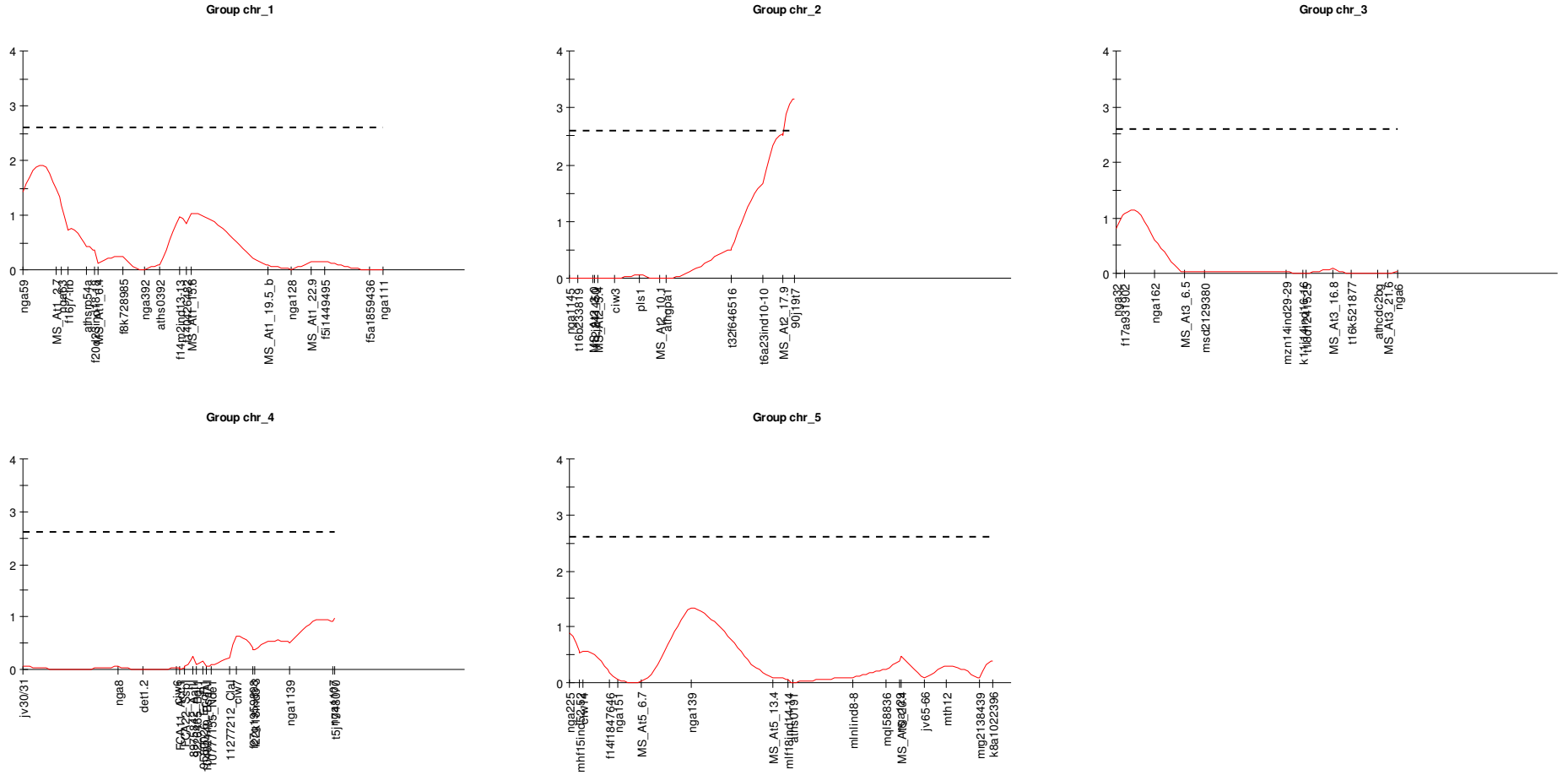
Data variate:	log10:18:3
Test statistic W:	0.973
Probability:	0.049

Data variate:	18:03
Test statistic W:	0.9801
Probability:	0.167

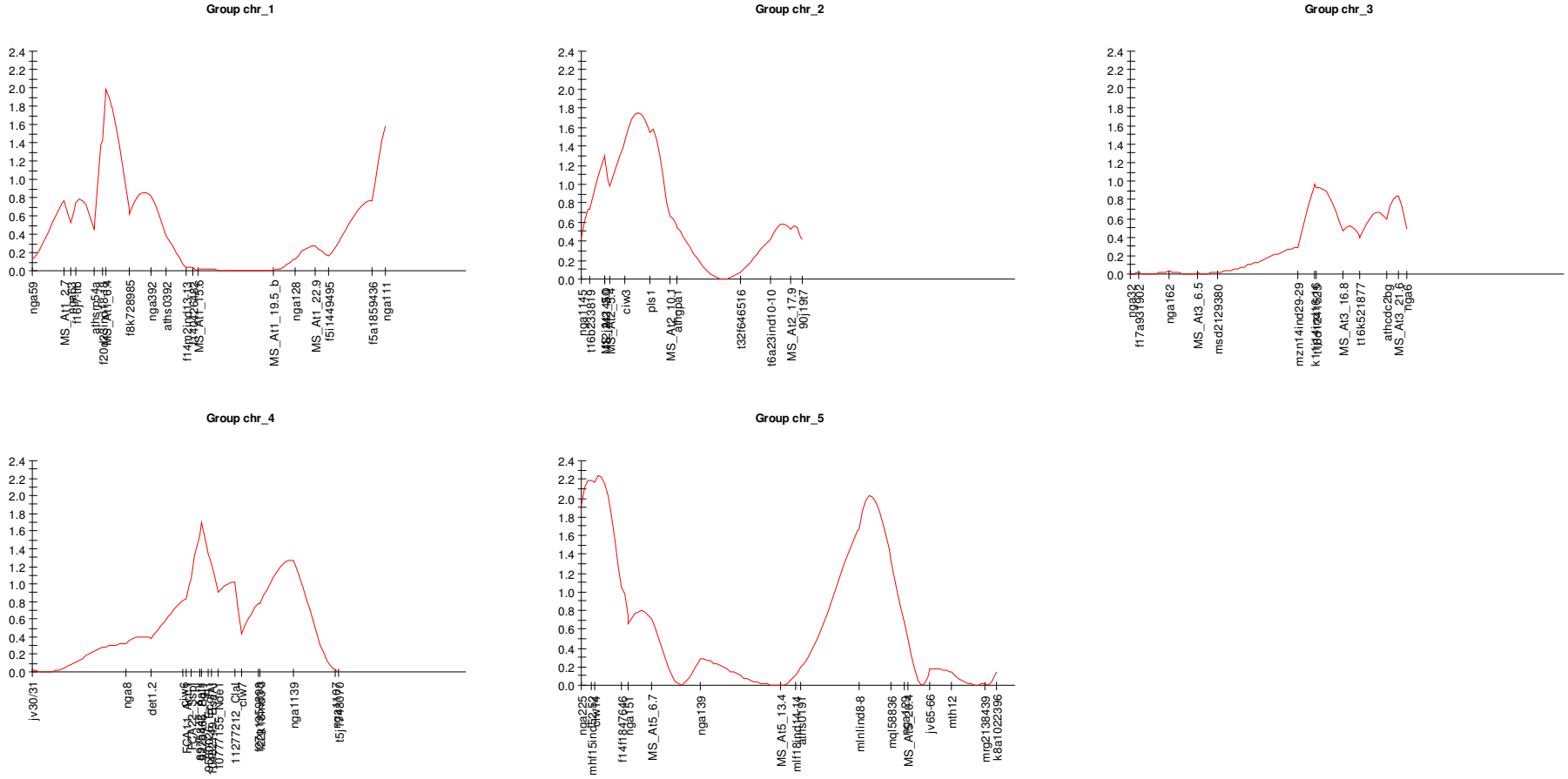
WC POPULATION QTL MAP CHART 18:1



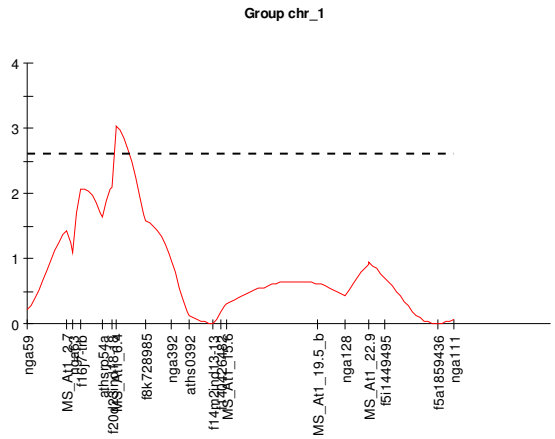
WC POPULATION QTL MAP CHART 18:2

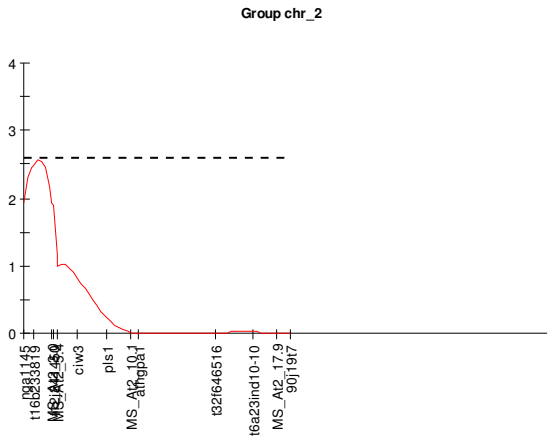


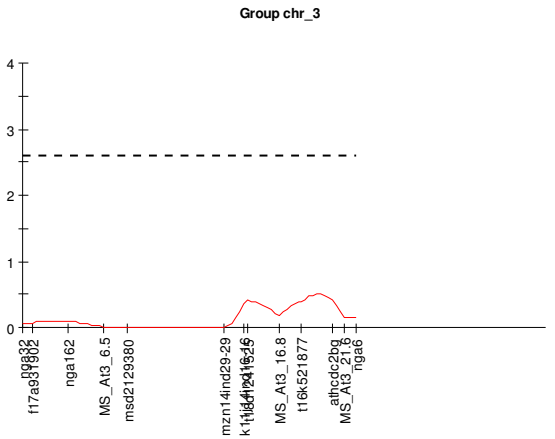
WC POPULATION QTL MAP CHART 18:2/18:1

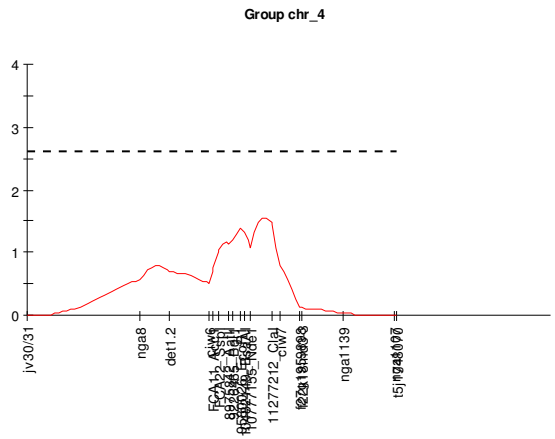


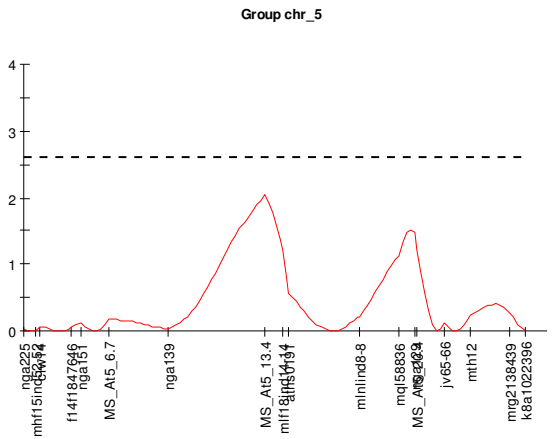
WC POPULATION QTL MAP CHART 18:3



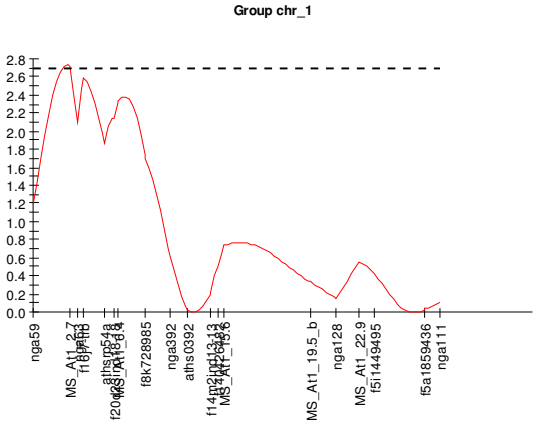


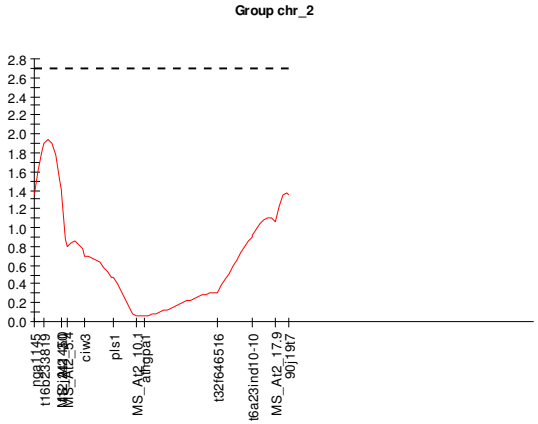


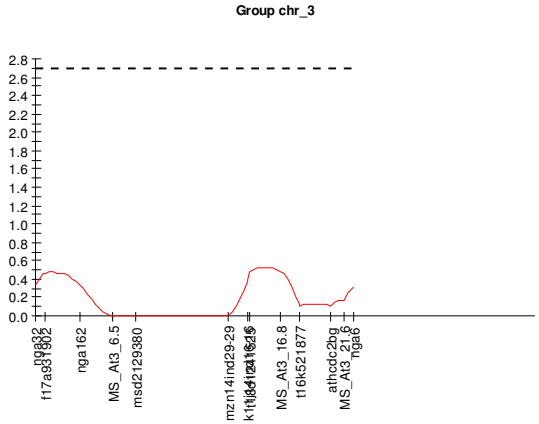


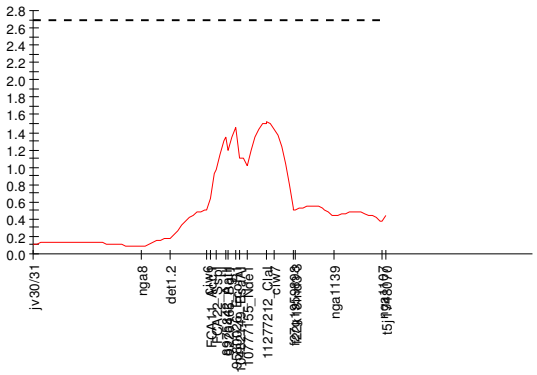


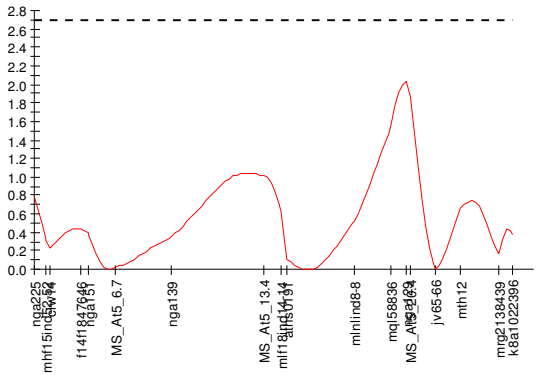
WC POPULATION QTL MAP CHART 18:3/18:2



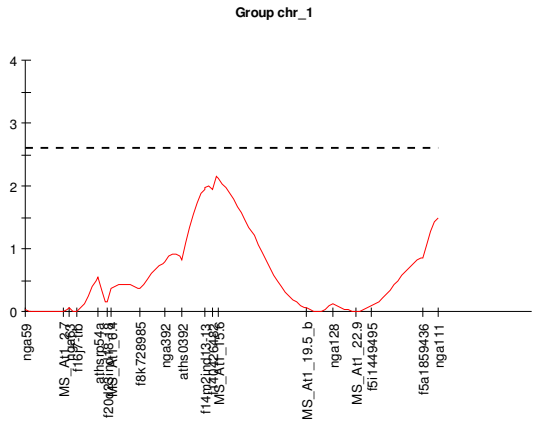


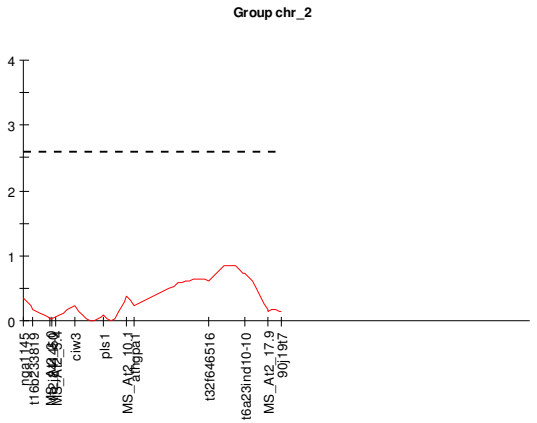


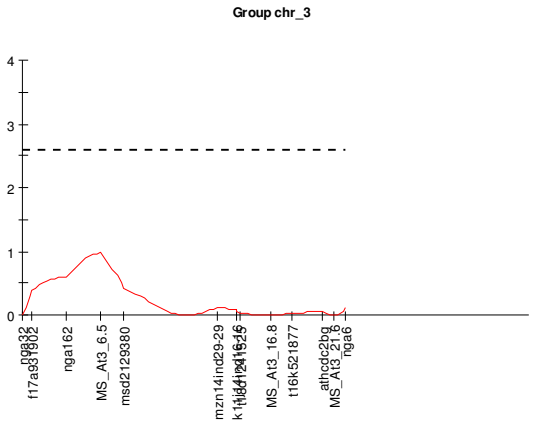


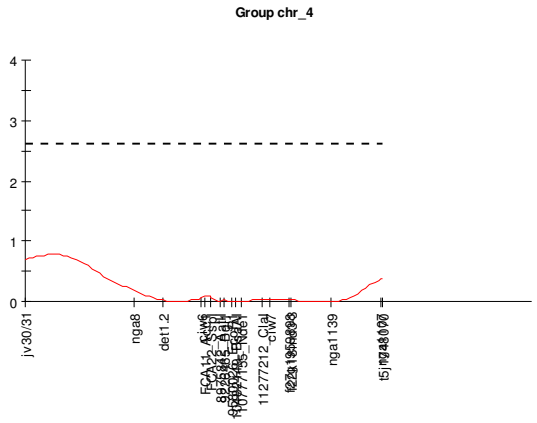


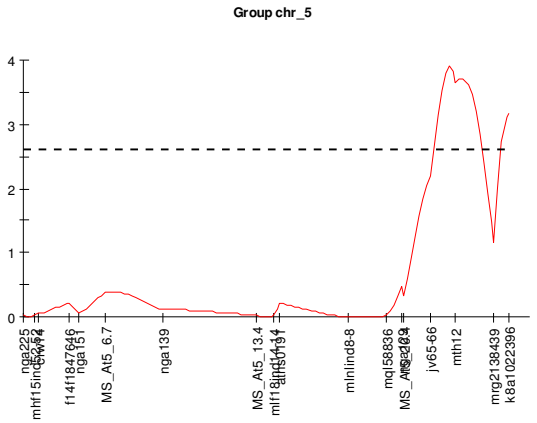
WC POPULATION QTL MAP CHART OIL CONTENT



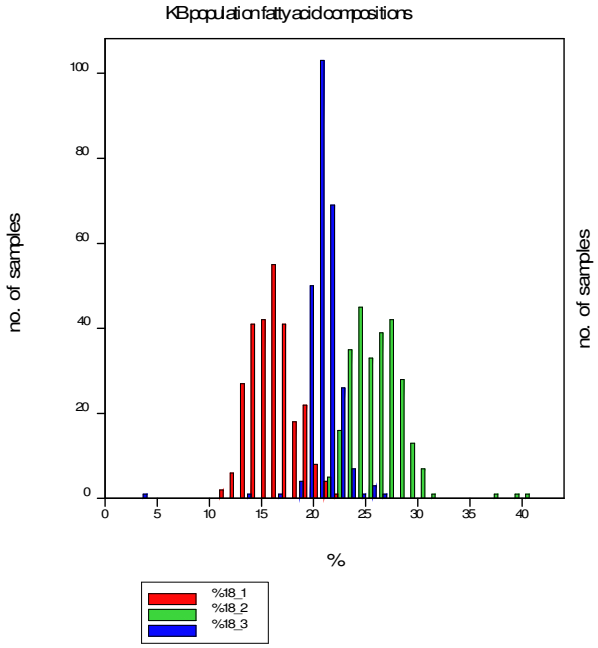


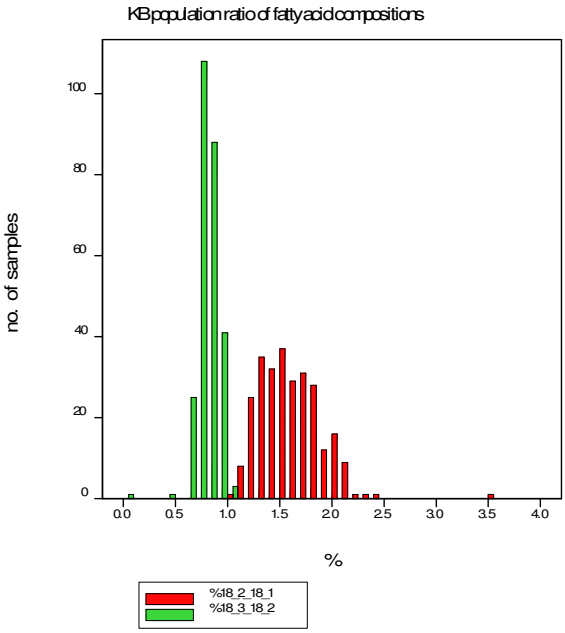


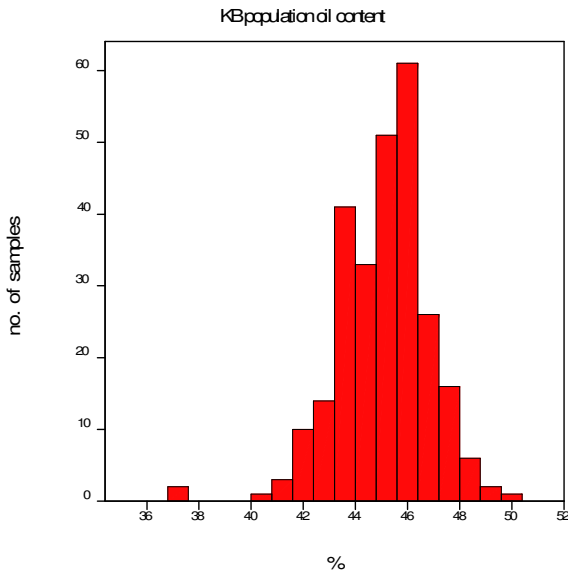




KB POPULATION HISTOGRAM







KB POPULATION NORMALITY TEST

KB original data
Shapiro-Wilk test for Normality

Data variate:	18:01
Test statistic W:	0.9669
Probability:	0.02

Data variate:	18:02
Test statistic W:	0.9832
Probability:	0.289

Data variate:	18:2/18:1
Test statistic W:	0.9783
Probability:	0.132

Data variate:	18:03
Test statistic W:	0.8319
Probability:	<0.001

Data variate:	18:3/18:2
Test statistic W:	0.987
Probability:	0.509

Data variate:	Oil
Test statistic W:	0.988
Probability:	0.574

KB transformed data
Shapiro-Wilk test for Normality

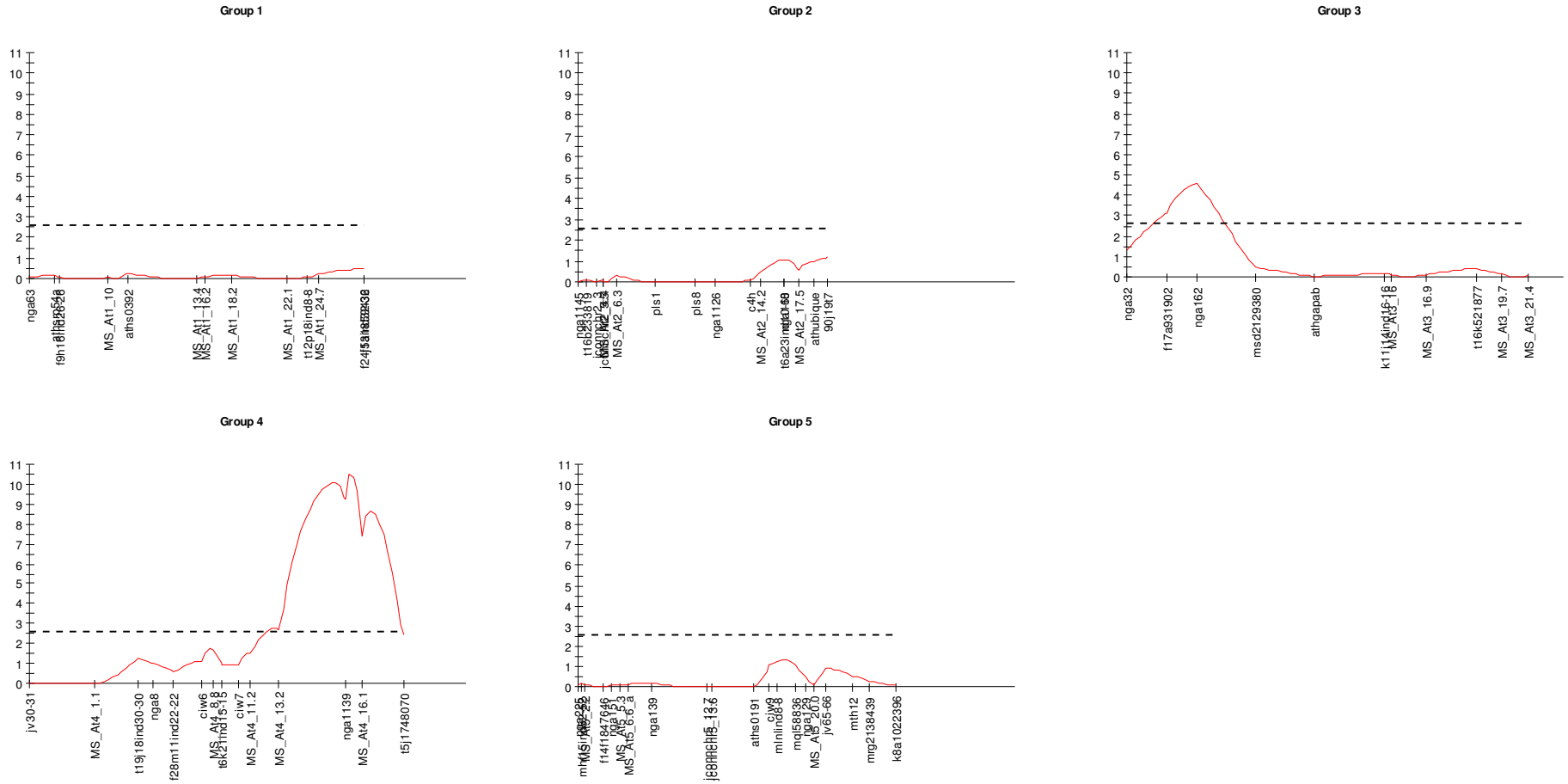
Data variate:	square root:18:1
Test statistic W:	0.973
Probability:	0.055

Data variate:	square root:18:3
Test statistic W:	0.8004
Probability:	<0.001

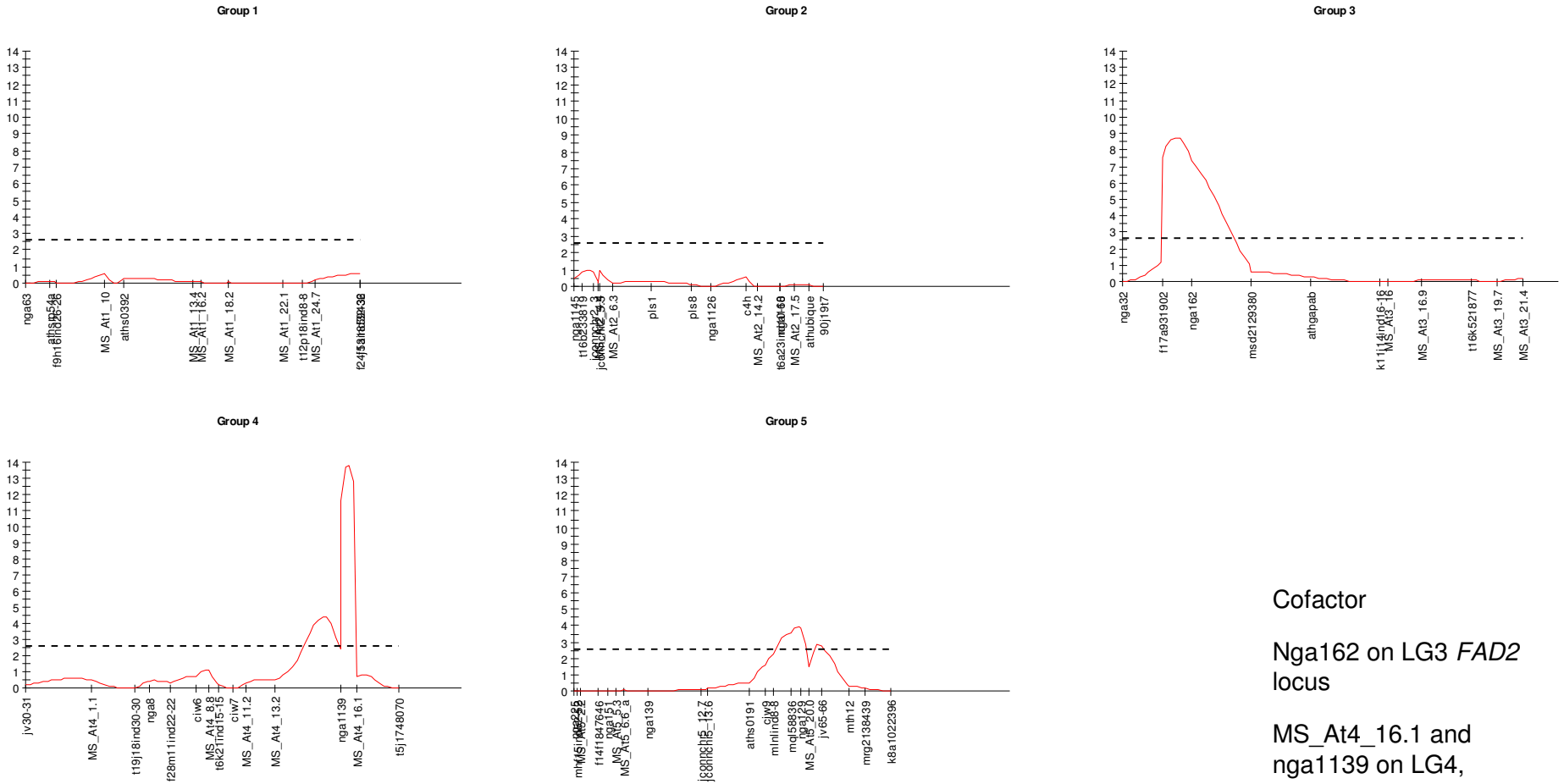
Data variate:	log10:18:1
Test statistic W:	0.977
Probability:	0.107

Data variate:	log10:18:3
Test statistic W:	0.765
Probability:	<0.001

KB POPULATION QTL MAP CHART 18:1(LOG10)



KB POPULATION QTL MAP CHART 18:1(LOG10) MQM MAPPING

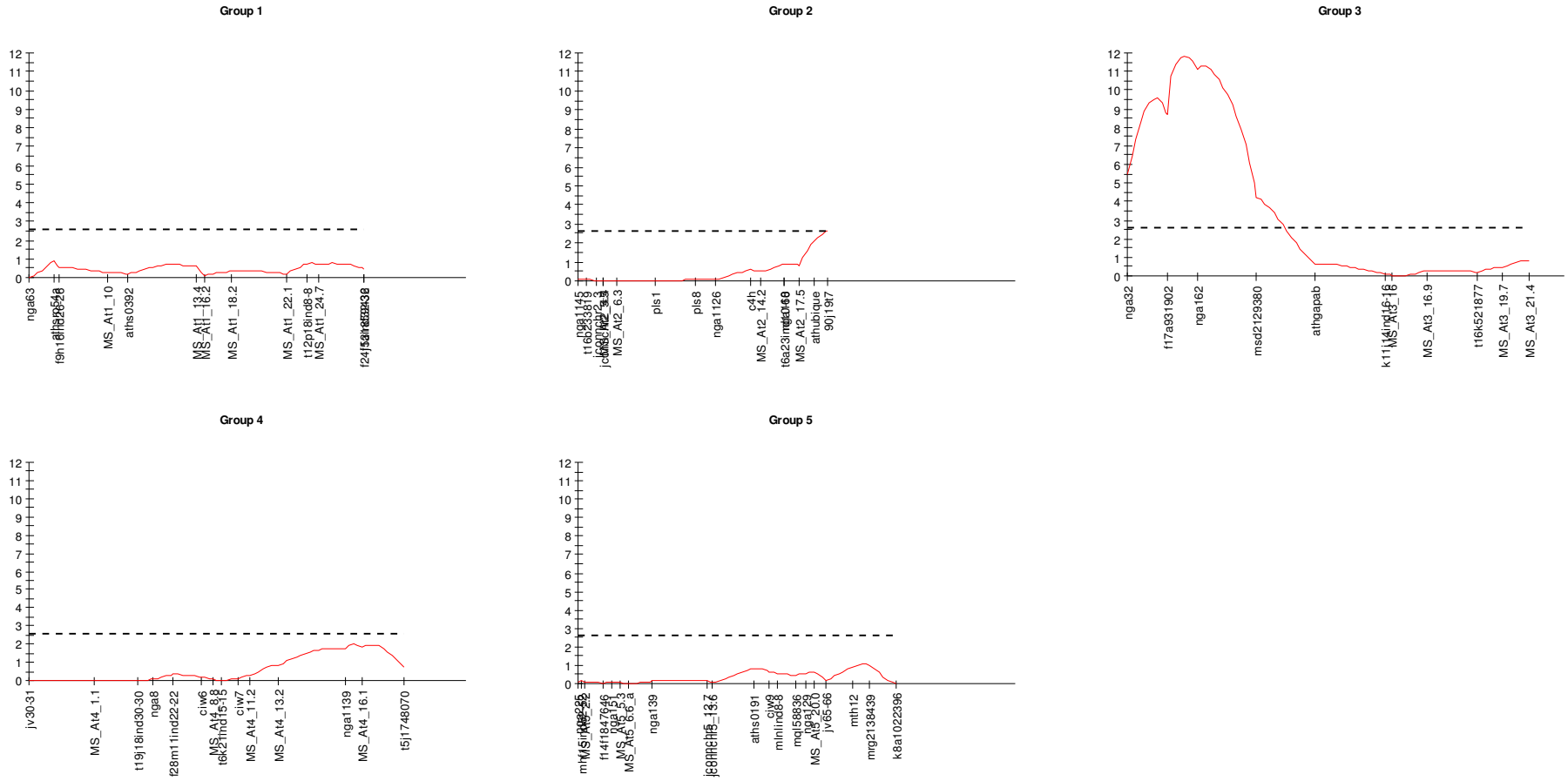


Cofactor

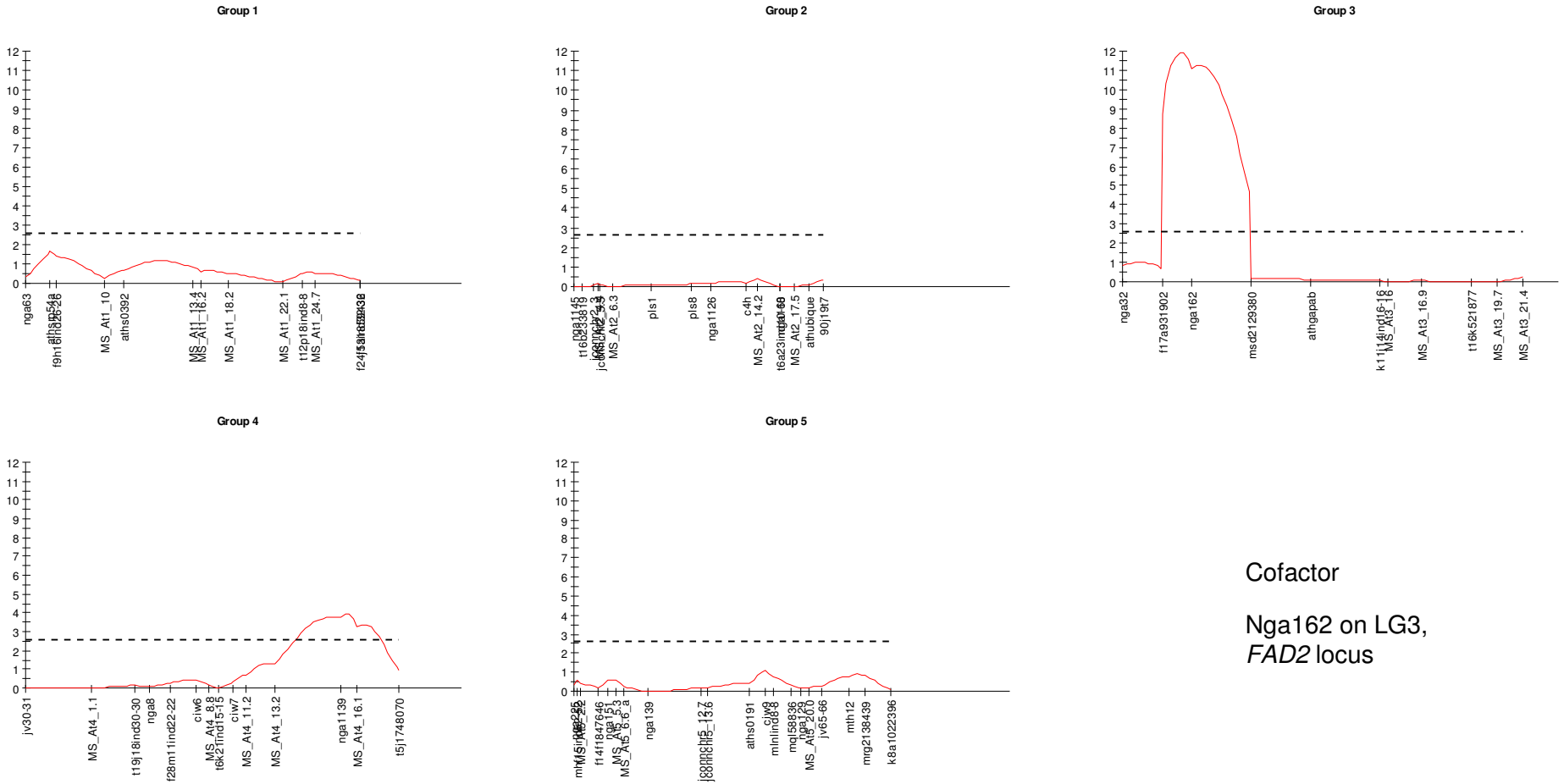
Nga162 on LG3 *FAD2* locus

MS_At4_16.1 and nga1139 on LG4, *FAE1* loci

KB POPULATION QTL MAP CHART 18:2



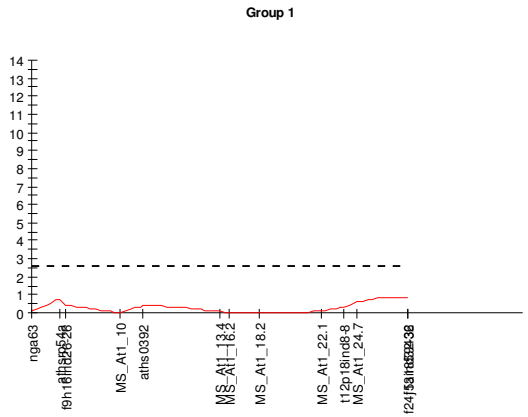
KB POPULATION QTL MAP CHART 18:2 MQM MAPPING

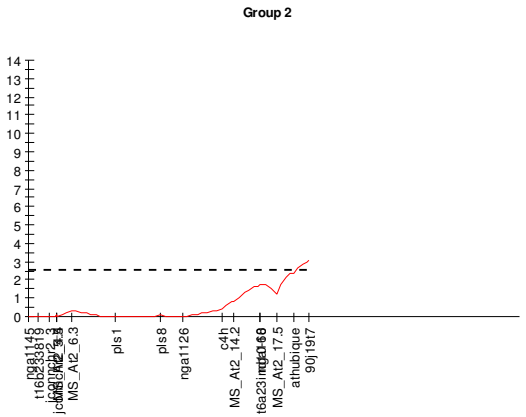


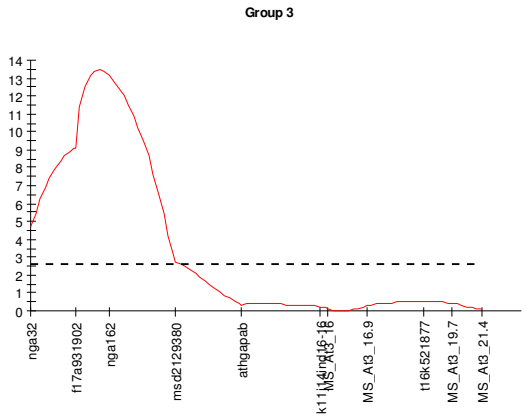
Cofactor

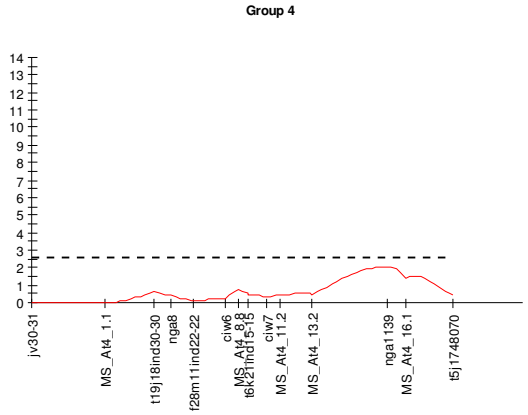
Nga162 on LG3,
FAD2 locus

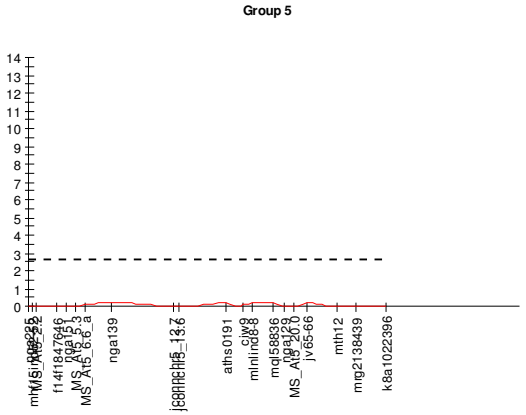
KB POPULATION QTL MAP CHART 18:2/18:1



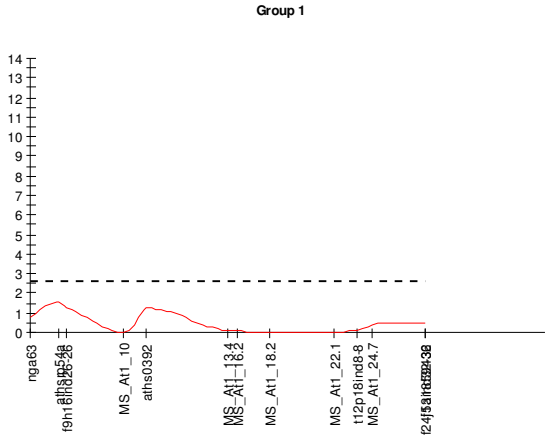


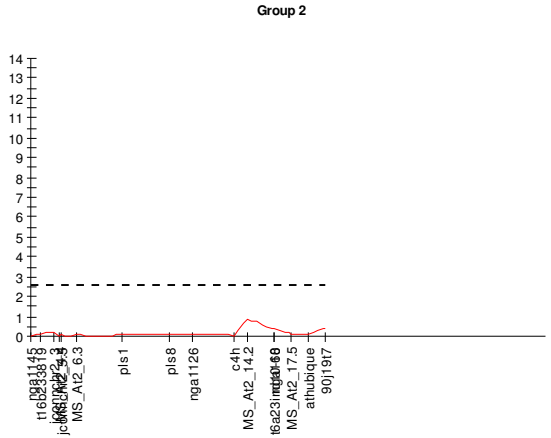


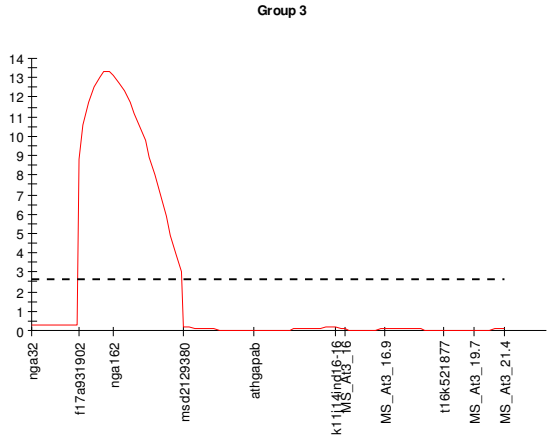


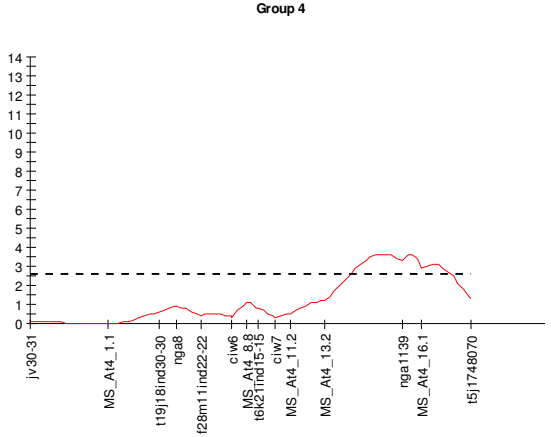


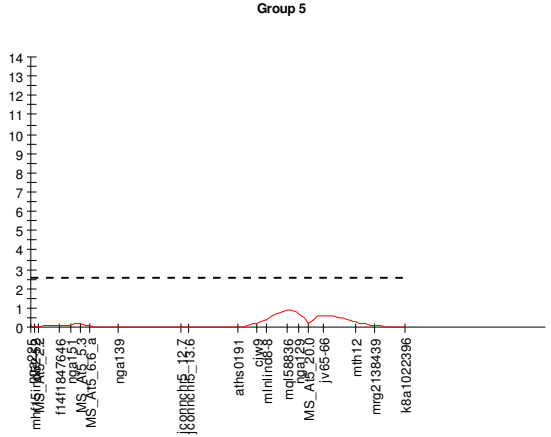
KB POPULATION QTL MAP CHART 18:2/18:1 MQM MAPPING







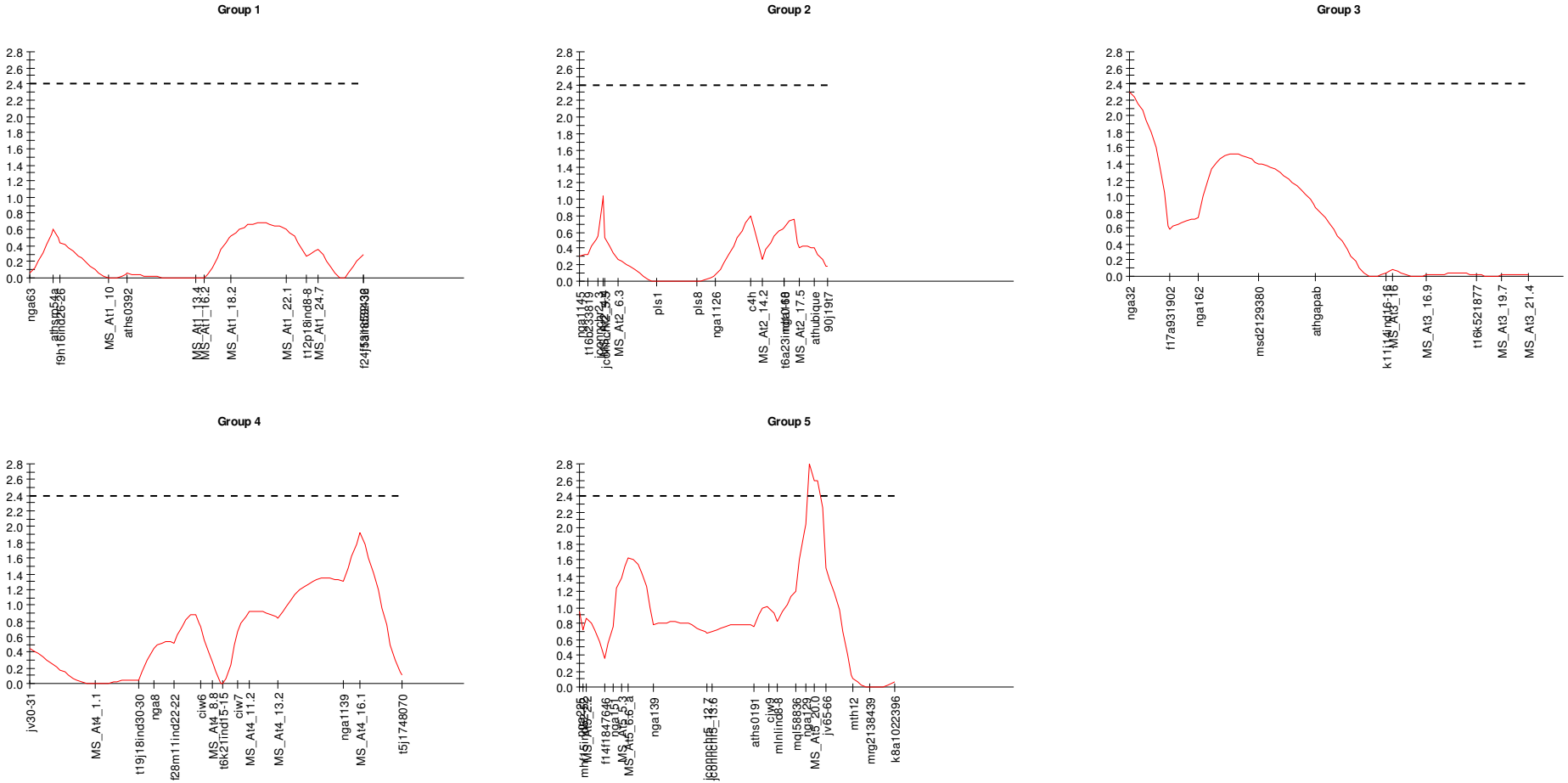




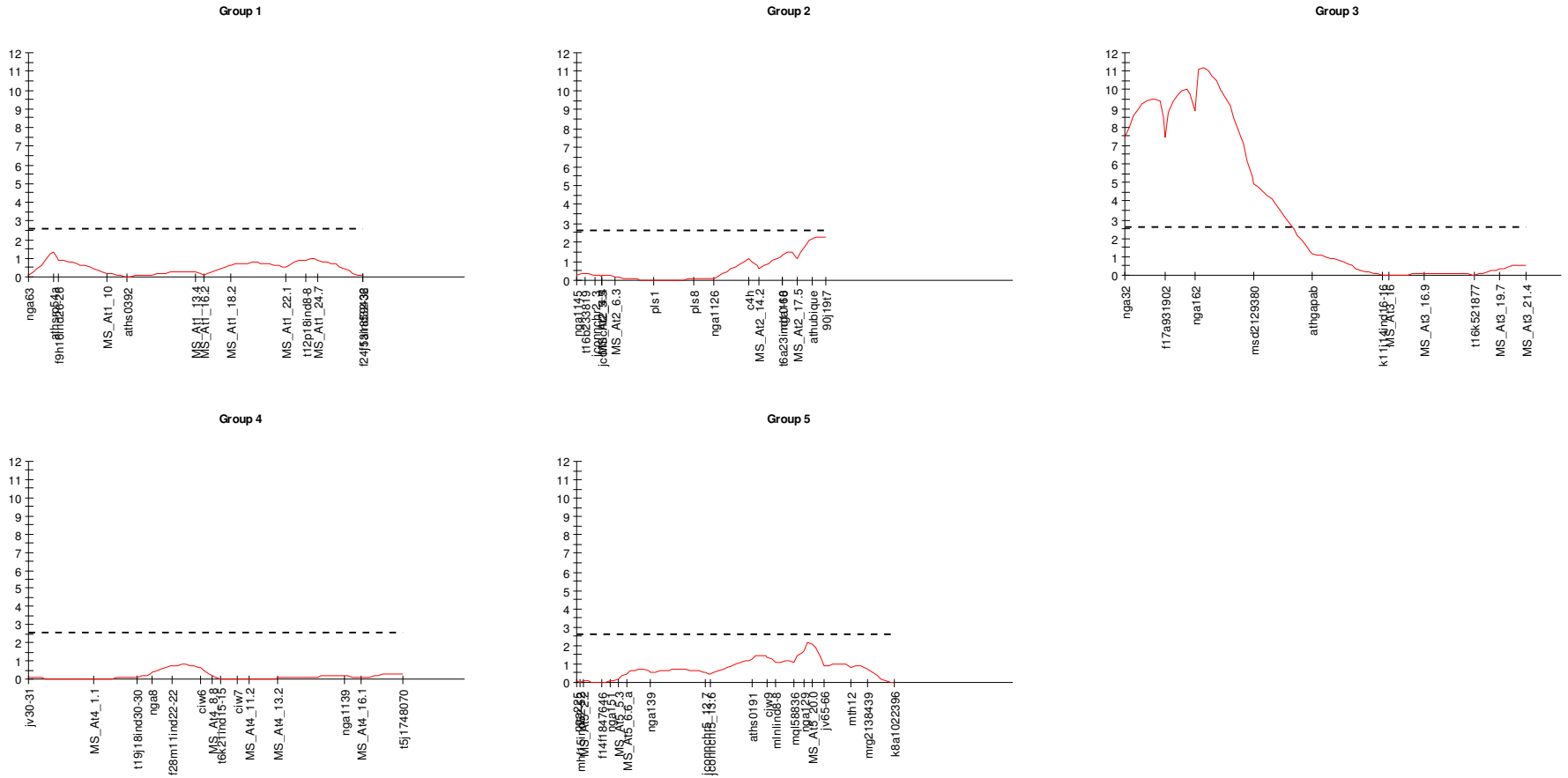
Cofactor

Nga162 on LG3,
FAD2 locus

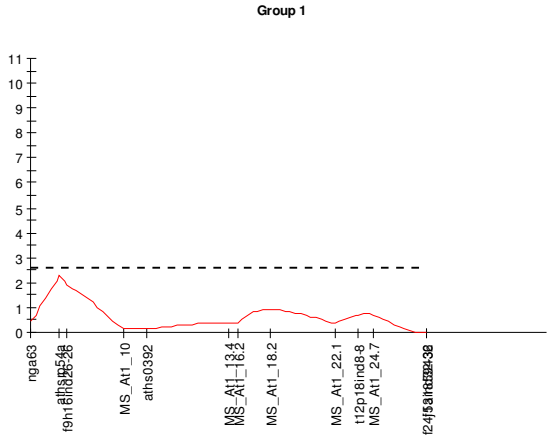
KB POPULATION QTL MAP CHART 18:3

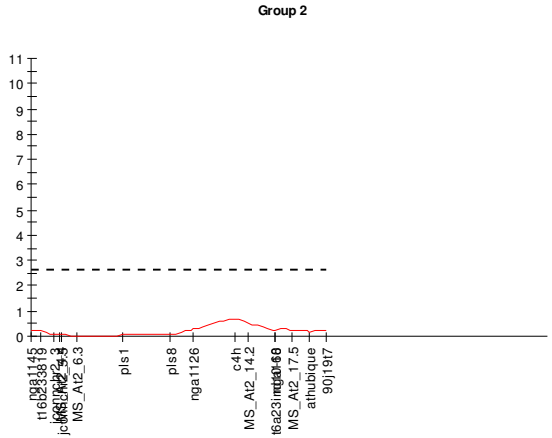


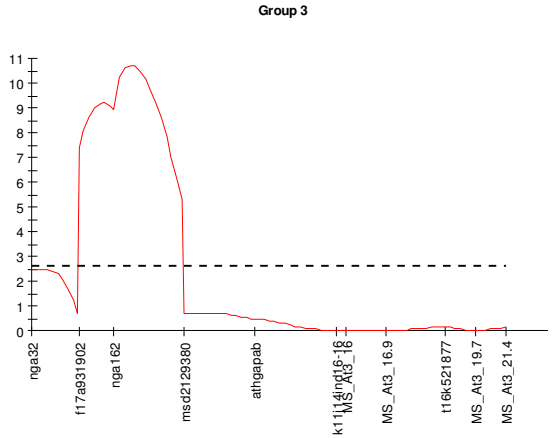
KB POPULATION QTL MAP CHART 18:3/18:2

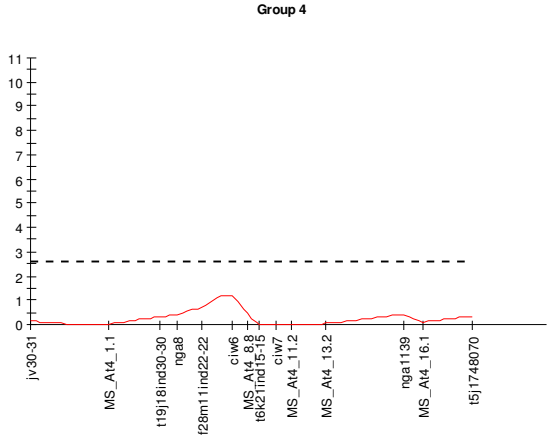


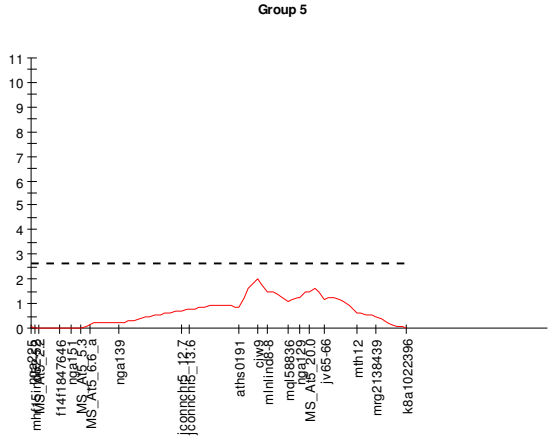
KB POPULATION QTL MAP CHART 18:3/18:2 MQM MAPPING







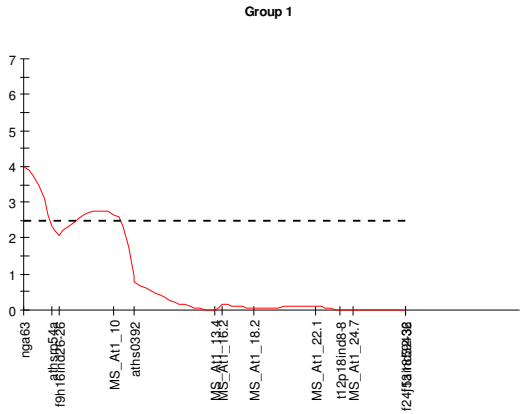


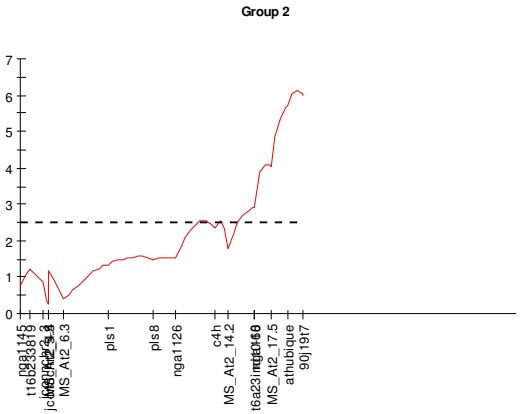


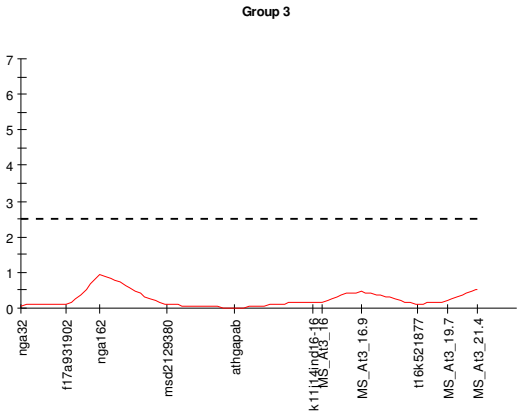
Cofactor

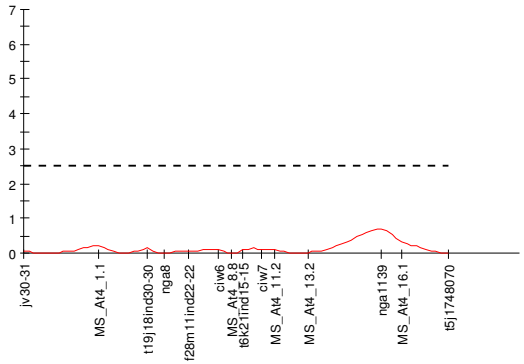
Nga162 on LG3,
FAD2 locus

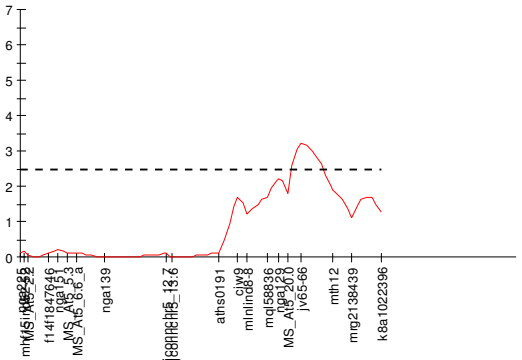
KB POPULATION QTL MAP CHART OIL CONTENT



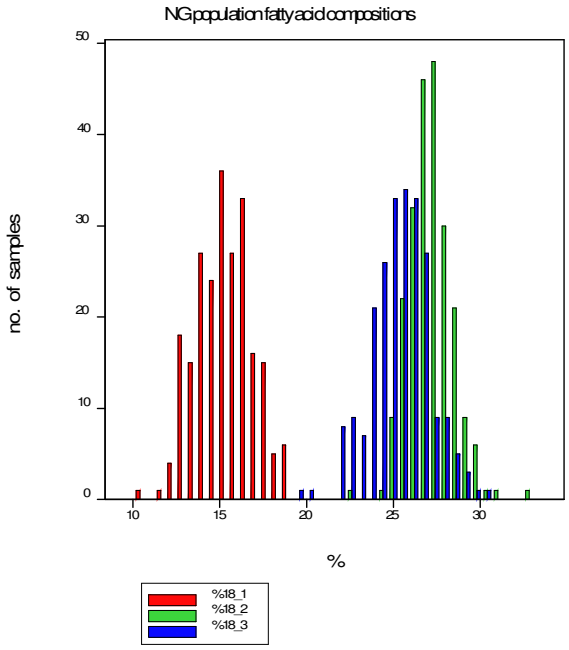


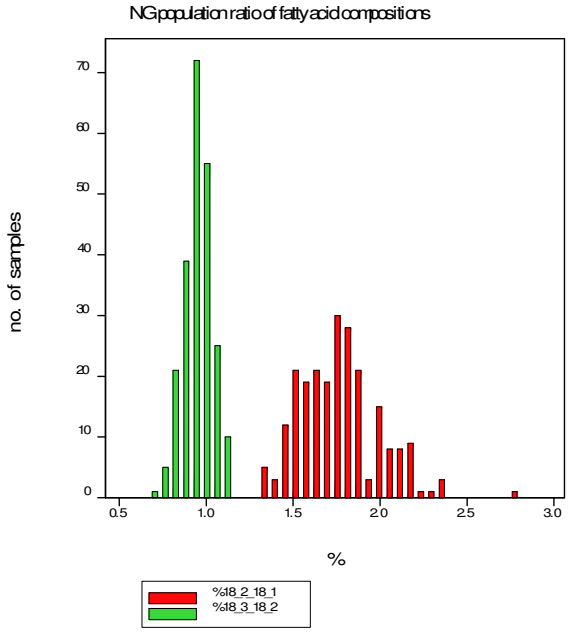


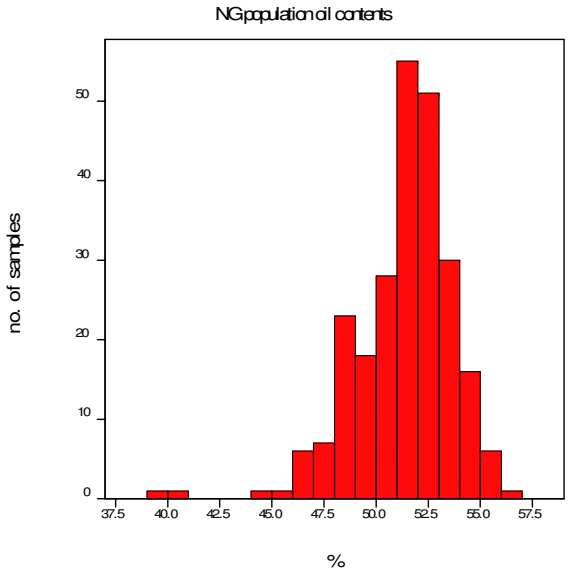




NG POPULATION HISTOGRAM







NG POPULATION NORMALITY TEST

NG original data
Shapiro-Wilk test for Normality

Data variate:	18:01
Test statistic W:	0.9891
Probability:	0.639

Data variate:	18:02
Test statistic W:	0.9819
Probability:	0.219

Data variate:	18:2/18:1
Test statistic W:	0.9779
Probability:	0.113

Data variate:	18:03
Test statistic W:	0.992
Probability:	0.851

Data variate:	18:3/18:2
Test statistic W:	0.9887
Probability:	0.604

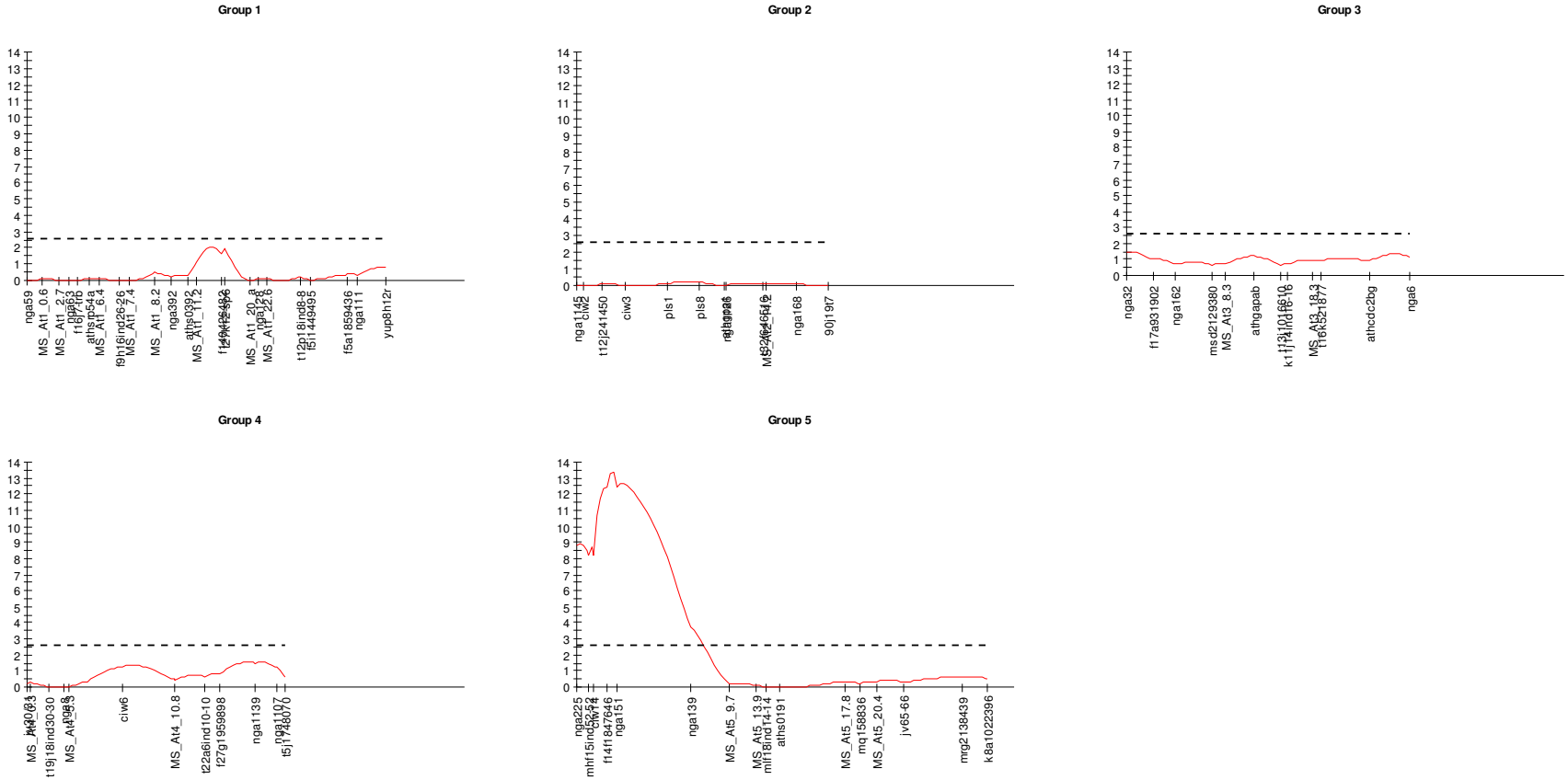
Data variate:	Oil
Test statistic W:	0.9571
Probability:	0.004

NG transformed data
Shapiro-Wilk test for Normality

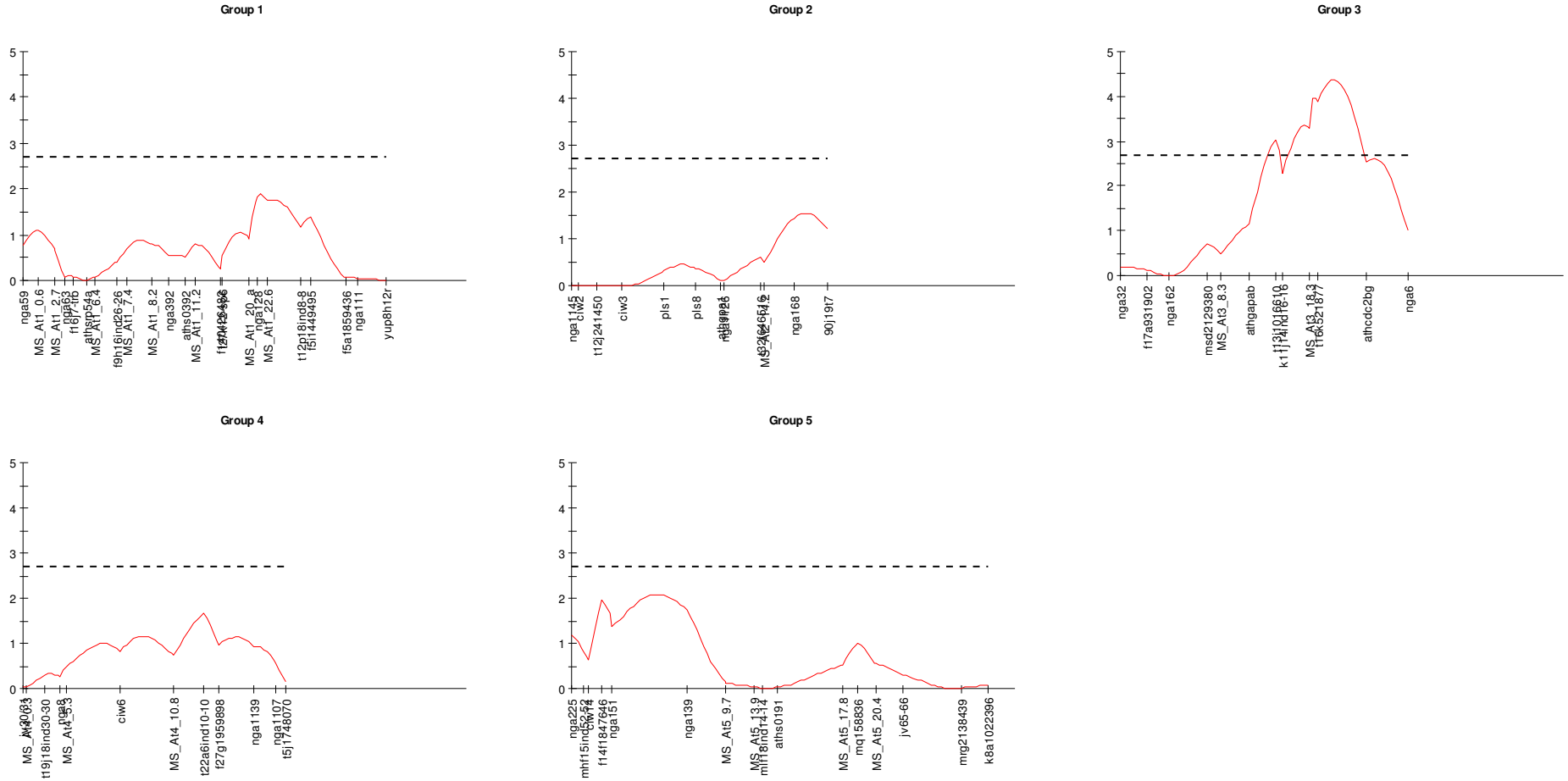
Data variate:	square root:Oil
Test statistic W:	0.9545
Probability:	0.002

Data variate:	log10:Oil
Test statistic W:	0.9518
Probability:	0.002

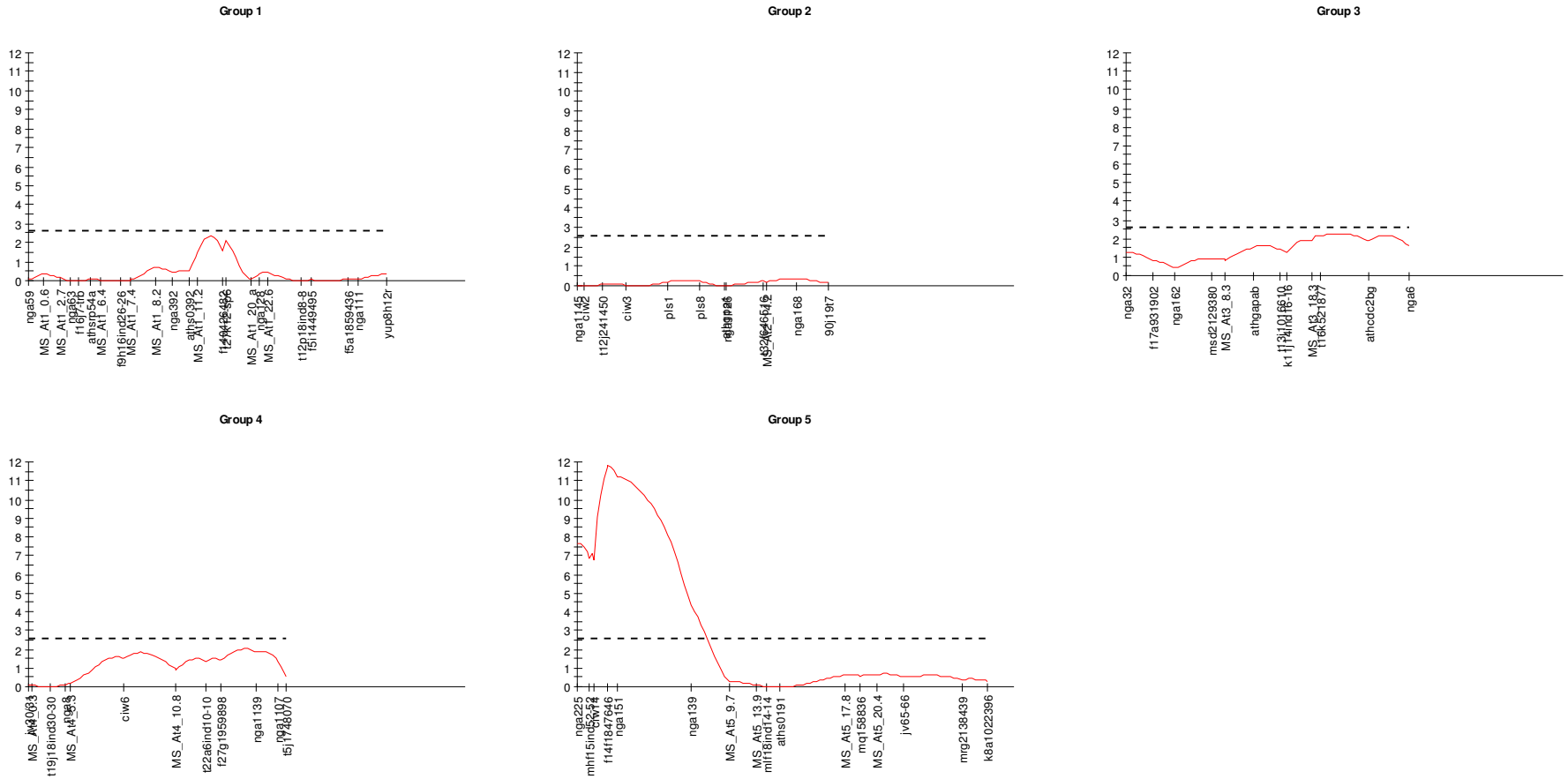
NG POPULATION QTL MAP CHART 18:1



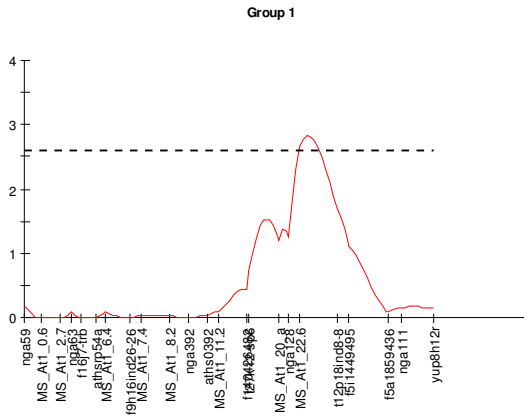
NG POPULATION QTL MAP CHART 18:2

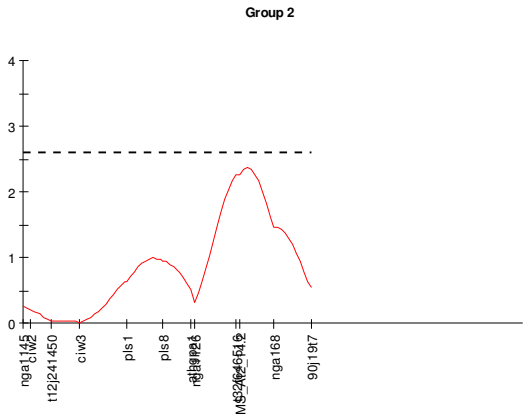


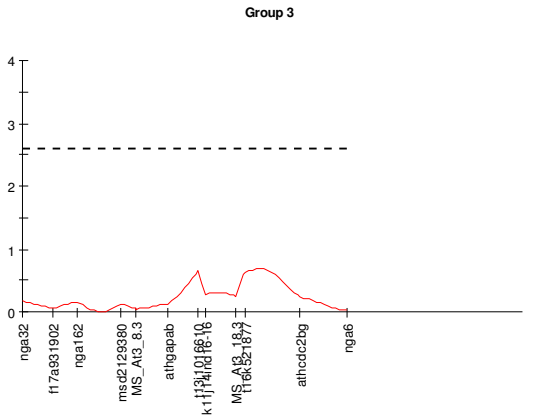
NG POPULATION QTL MAP CHART 18:2/18:1

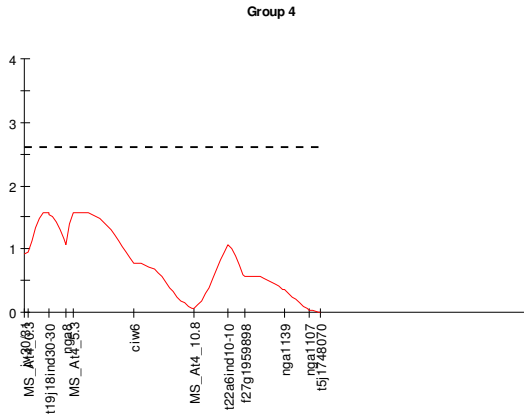


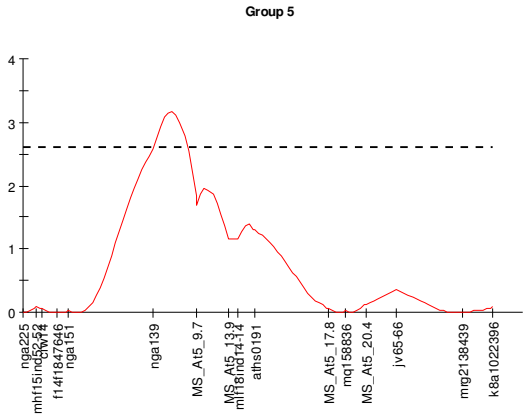
NG POPULATION QTL MAP CHART 18:3



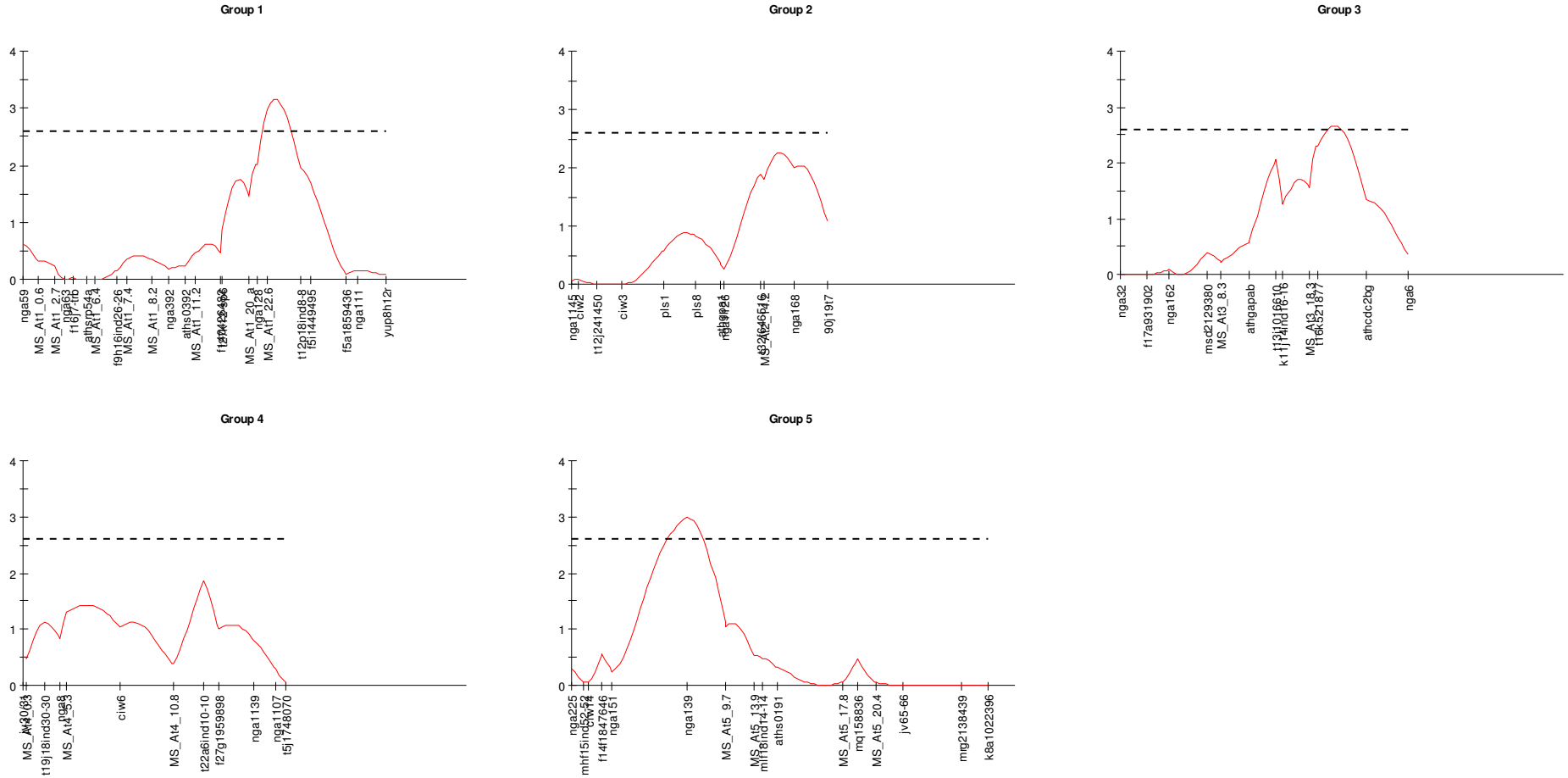




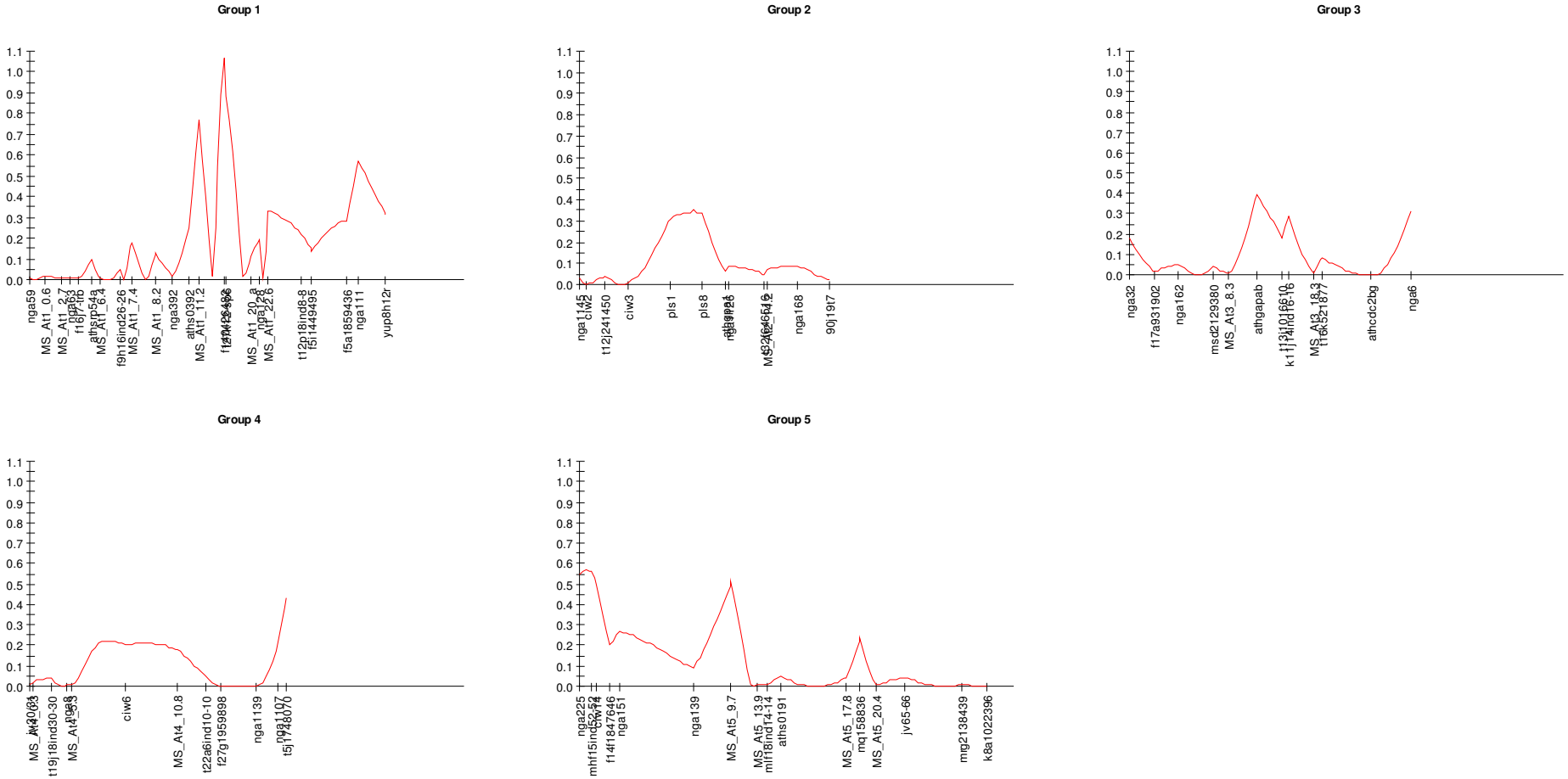




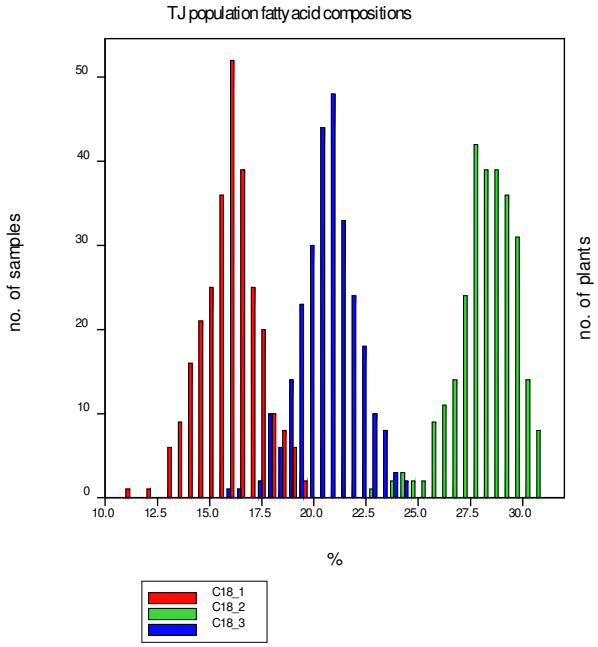
NG POPULATION QTL MAP CHART 18:3/18:2

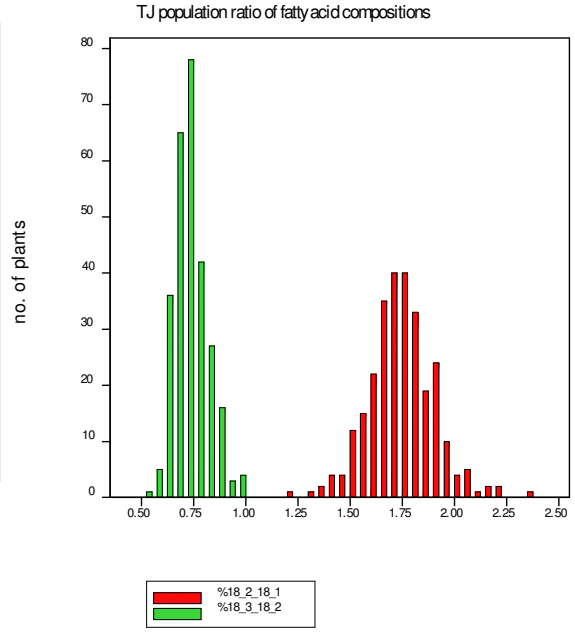


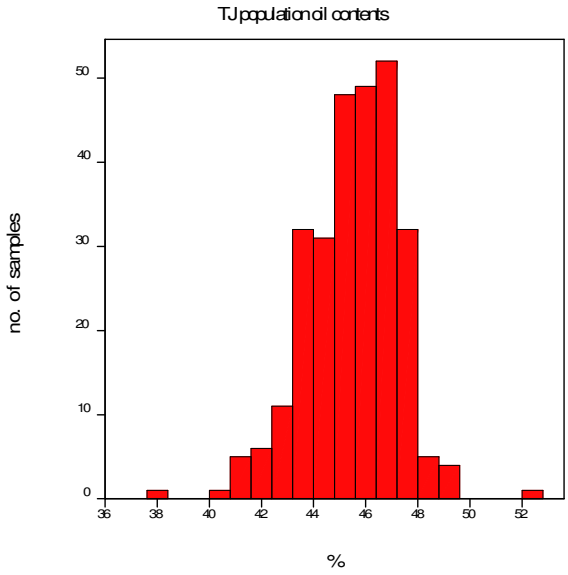
NG POPULATION QTL MAP CHART OIL CONTENT



TJ POPULATION HISTOGRAM







TJ POPULATION NORMALITY TEST

TJ original data
Shapiro-Wilk test for Normality

Data variate: C18:1
Test statistic W: 0.996
Probability: 0.994

Data variate: C18:2
Test statistic W: 0.9924
Probability: 0.875

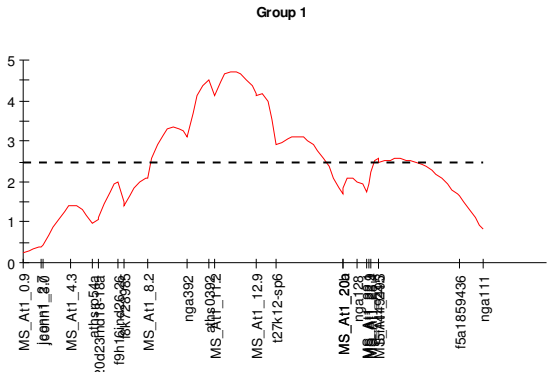
Data variate: 18:2/18:1
Test statistic W: 0.9846
Probability: 0.338

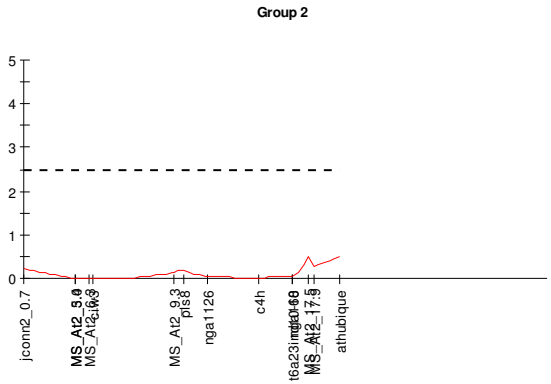
Data variate: C18:3
Test statistic W: 0.9885
Probability: 0.59

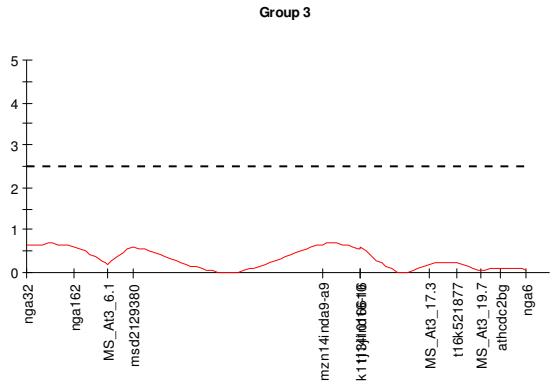
Data variate: 18:3/18:2
Test statistic W: 0.9853
Probability: 0.374

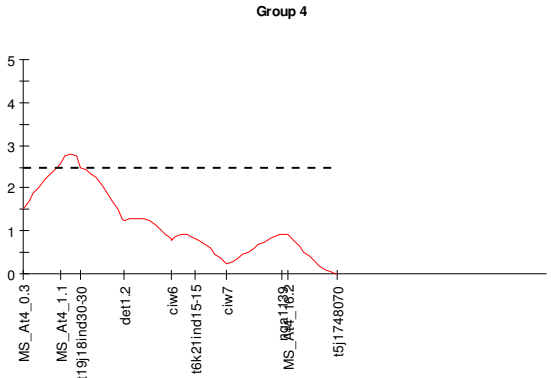
Data variate: Oil
Test statistic W: 0.9806
Probability: 0.178

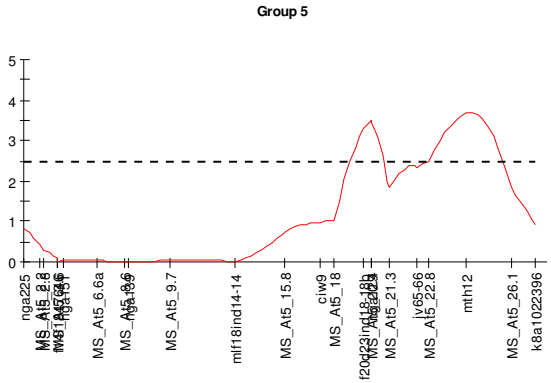
TJ POPULATION QTL MAP CHART 18:1



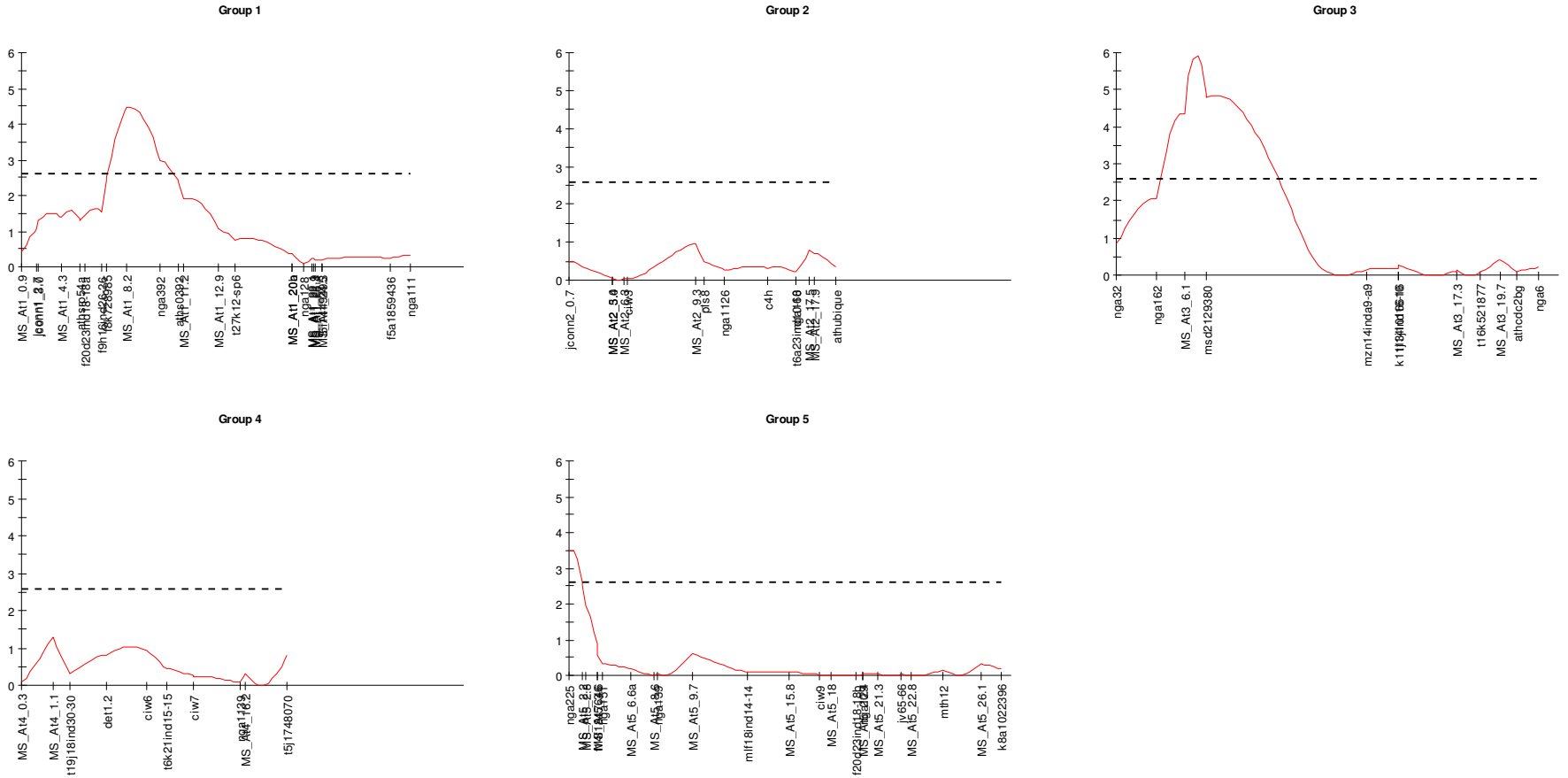




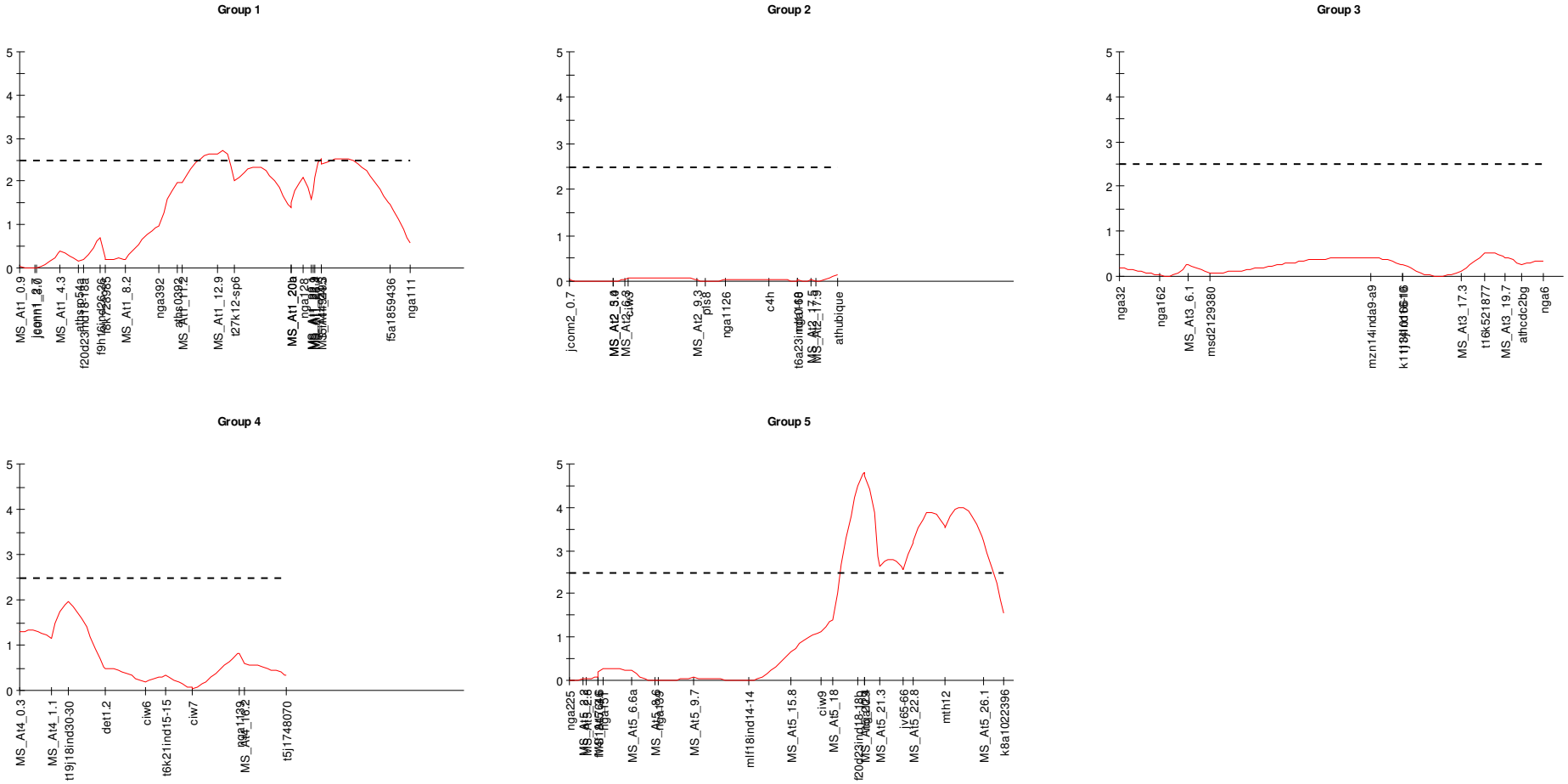




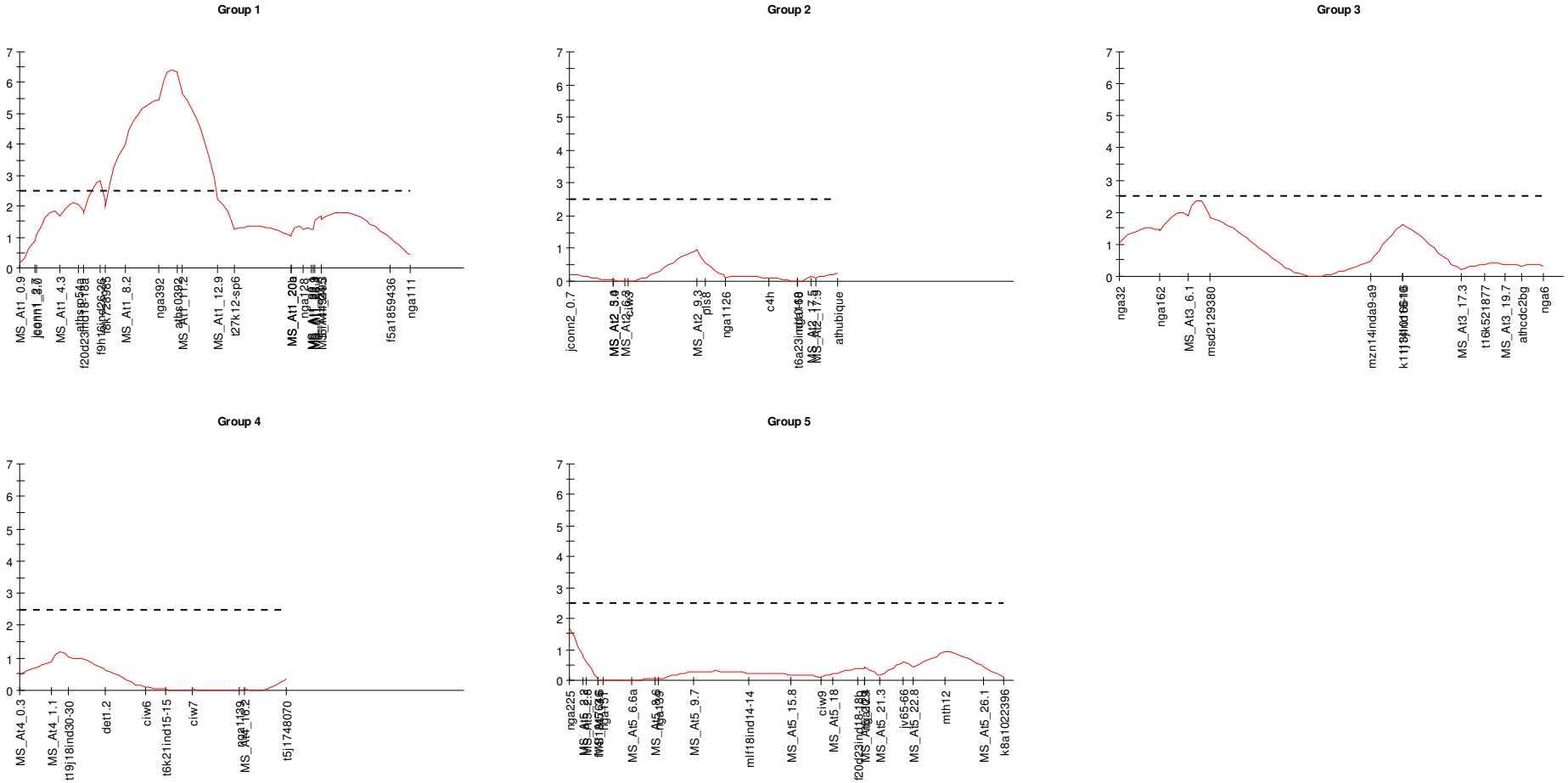
TJ POPULATION QTL MAP CHART 18:2



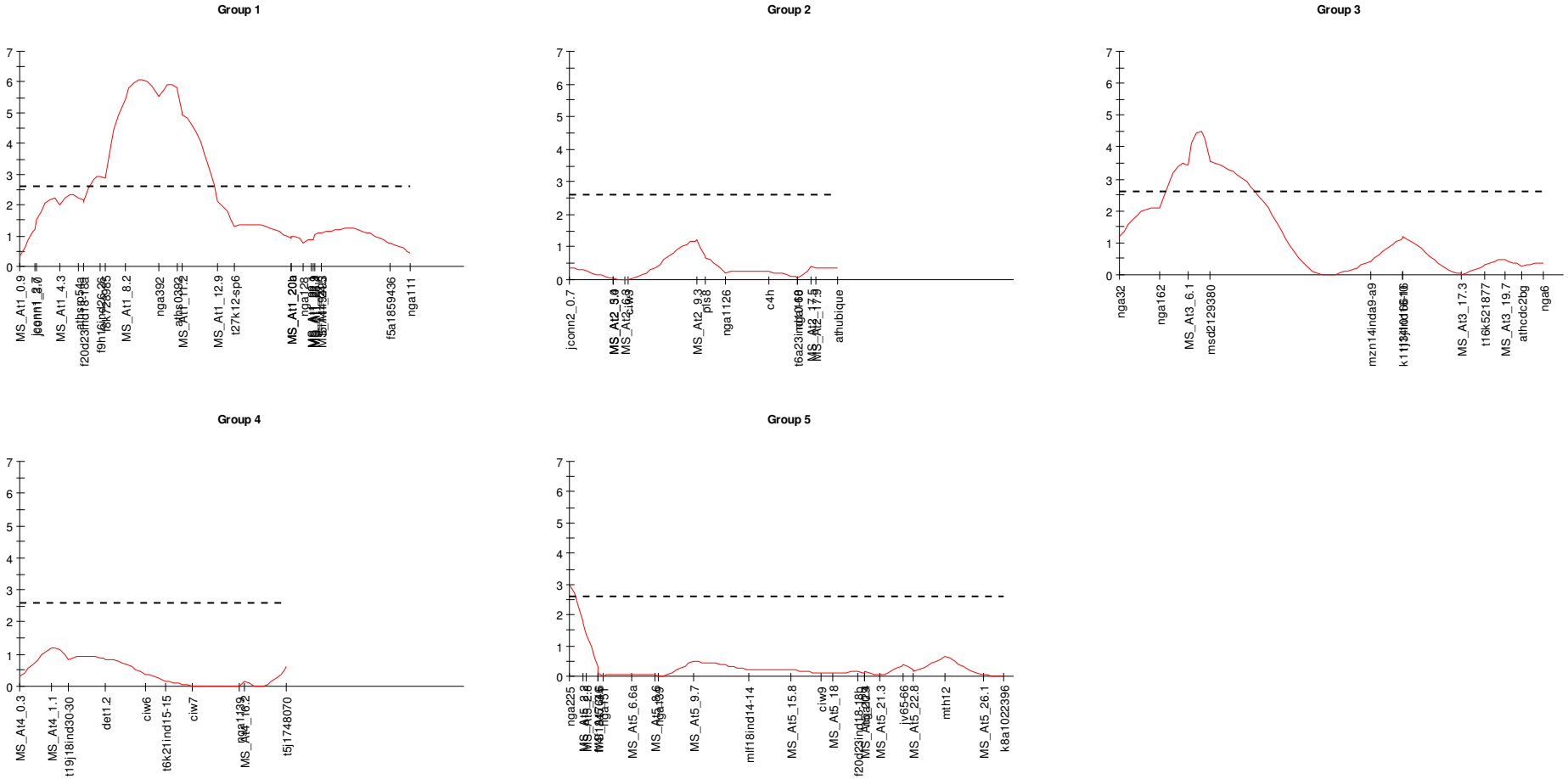
TJ POPULATION QTL MAP CHART 18:2/18:3



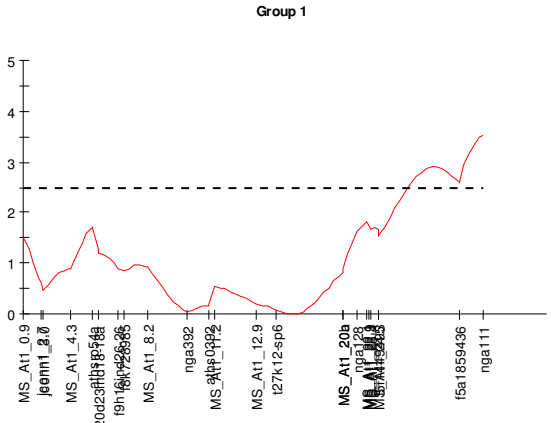
TJ POPULATION QTL MAP CHART 18:3

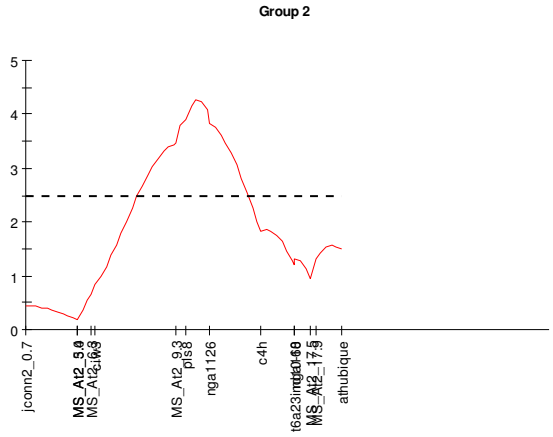


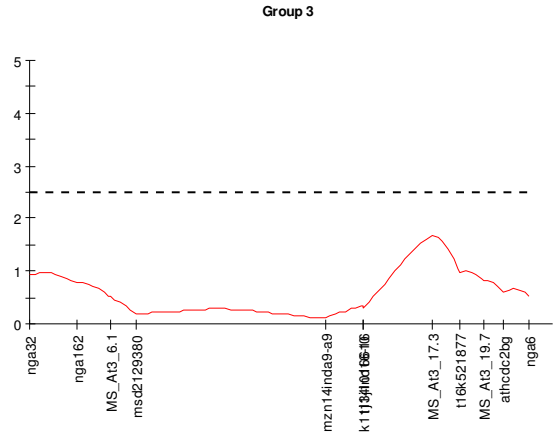
TJ POPULATION QTL MAP CHART 18:3/18:2

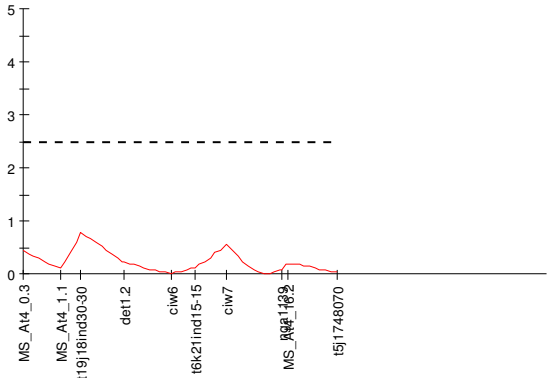


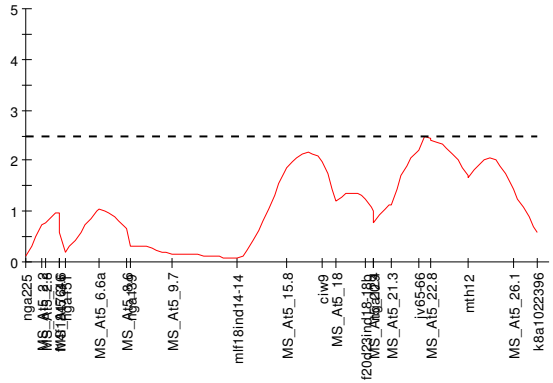
TJ POPULATION QTL MAP CHART IL CONTENT











APPENDIX 3 GENOTYPE FILES FOR RIL SELECTION FOR NIL DEVELOPMENT

TJ25 18:2

;Ts-5 x Mz-0 recombinant F8 population genotype file;
;A= Ts-5,B = Mz-0,U = unclassified,
B is increasing allele
18:2

rep1	24.5	27.92	27.82	25.21	27.97
rep2	26.7	28.01	28.89	28.11	*
rep3	27.43	28.90	28.86	27.11	28.06
ave.	**26.21**	**28.28**	**28.52**	**26.81**	**28.02**
no. of A	70	58	52	39	34

markers	LOD	LG	69	25	76	93	74
MS_At1_0.9	0.4	1	A	A	A	A	A
jconn1_2.7	1.1	1	A	U	A	U	A
jconn1_3.0	1.3	1	A	A	A	U	A
MS_At1_4.3	1.4	1	A	A	A	A	A
athsrp54a	1.3	1	A	A	A	A	A
f20d23ind18-18a	1.4	1	A	A	A	A	A
f9h16ind26-26	1.5	1	A	A	A	U	A
f8k728985	2.5	1	A	A	A	A	U
MS_At1_8.2	4.5	1	A	A	A	A	A
nga392	3.0	1	A	A	A	A	A
aths0392	2.4	1	A	A	A	A	B
MS_At1_11.2	1.9	1	A	A	A	A	B
MS_At1_12.9	1.1	1	A	B	A	A	B
t27k12-sp6	0.8	1	A	B	A	A	B
MS_At1_20b	0.4	1	A	B	B	A	B
MS_At1_20a	0.4	1	A	B	B	A	B
nga128	0.1	1	A	B	B	A	B
MS_At1_22.1	0.2	1	A	A	B	A	B
MS_At1_22.9	0.2	1	A	A	B	A	B
MS_At1_23.4	0.2	1	A	A	B	A	B
f5i1449495	0.2	1	A	A	B	A	B
MS_At1_24.3	0.2	1	A	A	B	A	B
f5a1859436	0.2	1	A	A	B	A	B
nga111	0.3	1	A	A	B	A	B
jconn2_0.7	0.5	2	B	B	B	B	B
MS_At2_3.0	0.1	2	A	B	B	B	B
MS_At2_5.4	0.1	2	A	B	B	B	B
MS_At2_6.3	0.0	2	A	B	B	B	B
jconn2_3.0		2	A	B	B	B	B
jconn2_5.2		2	A	B	B	B	B
ciw3	0.0	2	A	B	B	B	B
MS_At2_9.3	1.0	2	A	B	A	B	B
pls8	0.5	2	A	B	A	B	B
nga1126	0.3	2	A	B	A	B	B
c4h	0.3	2	A	A	A	B	B
t6a23ind10-10	0.2	2	A	A	B	B	B
nga168	0.2	2	A	A	B	B	U
MS_At2_17.5	0.8	2	A	A	B	B	A
MS_At2_17.9	0.7	2	A	A	B	B	A
athubique	0.3	2	A	A	B	B	A
nga32	0.8	3	A	B	A	B	B
nga162	2.1	3	A	B	A	B	U
MS_At3_6.1	4.3	3	A	B	B	B	B
msd2129380	4.8	3	B	B	B	B	B
mzn14inda9-a9	0.1	3	A	A	B	B	B
t13j1016610	0.2	3	A	A	B	B	A
k11j14ind16-16	0.3	3	A	A	B	B	A
MS_At3_17.3	0.1	3	A	A	B	B	A
t16k521877	0.1	3	A	A	A	A	A
MS_At3_19.7	0.4	3	A	A	A	A	A
athcdc2bg	0.1	3	A	A	A	A	A
nga6	0.2	3	A	A	A	A	A
MS_At4_0.3	0.1	4	B	A	A	A	A
MS_At4_1.1	1.3	4	B	A	A	A	A
t19j18ind30-30	0.3	4	B	A	A	A	B
nga8		4	B	A	A	A	U
det1.2	0.8	4	B	A	B	B	B
f28m1-ind22-22		4	B	A	B	B	B
ciw6	0.9	4	A	A	B	B	B
t6k21ind15-15	0.5	4	A	A	B	B	A
ciw7	0.2	4	A	A	B	B	A
nga1139	0.1	4	A	B	B	B	B
MS_At4_16.2	0.3	4	A	B	B	B	B
t5j1748070	0.8	4	B	B	B	B	B
nga225	3.5	5	A	A	A	A	A
MS_At5_2.2	2.5	5	A	A	A	A	A
MS_At5_2.8	1.9	5	A	A	A	A	A
MS_At5_3.6	0.9	5	A	A	A	B	A
f14f1847646	0.6	5	A	A	A	B	A
nga151	0.3	5	B	A	A	B	A
MS_At5_6.6a	0.2	5	B	A	A	B	A
MS_At5_6.6b	0.2	5	B	A	A	B	A
MS_At5_8.6	0.0	5	B	B	A	H	A
nga139	0.0	5	B	B	A	B	A
MS_At5_9.7	0.6	5	B	B	A	B	A
mlf18ind14-14	0.1	5	B	B	A	A	B
MS_At5_15.8	0.1	5	A	B	A	A	B
ciw9	0.0	5	A	B	A	A	B
MS_At5_18	0.0	5	A	A	A	A	B
f20d23ind18-18b	0.0	5	A	A	A	A	B
nga129	0.0	5	A	A	A	A	B
MS_At5_20.4	0.0	5	A	A	A	A	B
MS_At5_21.3	0.0	5	A	A	A	B	B
jv65-66	0.0	5	A	A	A	B	B
MS_At5_22.8	0.0	5	A	A	A	B	B
mth12	0.1	5	A	A	A	B	B
MS_At5_26.1	0.3	5	B	A	A	B	B
k8a1022396	0.2	5	B	B	A	B	B

;Nok-3 x Ga-0 RI F8 population, 94 individuals, genotype file,
;A = Nok-3 B = Ga-0 U = unclassified & heterozygotes
B is increasing allele
18:2/18:1

rep1	2.02	2.04	2.19	2.39	1.63	*	2.13	2.18	2.18	2.39	2.16	*	1.99	1.80	2.77	2.20	2.02	
rep2	1.91	1.90	2.08	*	1.78	1.81	1.66	1.85	2.04	*	*	2.11	1.70	1.61	1.82	1.84	*	
rep3	2.00	2.06	2.14	2.34	1.85	1.87	2.21	1.97	2.08	2.21	2.11	2.03	1.84	1.82	1.90	1.55	1.80	
ave.	**1.97**	**2.00**	**2.13**	**2.37**	**1.75**	**1.84**	**1.98**	**1.99**	**2.09**	**2.30**	**2.14**	**2.07**	**1.84**	**1.74**	**2.10**	**1.84**	**1.91**	
no. of A	55	54	54	51	50	50	49	48	48	47	47	46	46	46	46	44	43	

markers	LOD	LG	9	36	59	32	6	72	29	17	41	13	28	3	12	16	78	4	21
nga59	0.1	1	A	A	A	B	A	A	B	A	A	A	B	U	A	A	A	B	A
MS_At1_0.6	0.4	1	B	B	A	B	A	A	B	A	A	A	B	B	A	A	A	B	A
MS_At1_2.7	0.2	1	B	B	A	B	A	B	B	A	A	A	B	B	A	A	A	A	A
nga63	0.0	1	B	A	A	B	A	A	B	A	A	A	B	U	A	A	A	A	A
f16j7-trb	0.0	1	U	A	A	B	A	B	B	U	A	A	B	A	A	A	A	A	A
athsrp54a	0.1	1	U	A	A	A	A	B	B	A	A	A	B	A	U	A	A	A	A
MS_At1_6.4	0.0	1	B	A	A	A	A	B	A	A	A	A	U	A	A	A	A	A	A
f9h16ind26-26	0.0	1	U	A	A	A	A	A	B	U	A	A	B	A	A	A	B	A	A
MS_At1_7.4	0.1	1	A	A	A	A	A	A	B	A	A	B	B	A	A	A	B	A	A
MS_At1_8.2	0.7	1	A	A	A	A	A	A	B	B	A	A	B	A	A	A	B	A	A
nga392	0.4	1	A	A	A	A	A	A	B	B	A	A	B	U	U	A	B	A	A
aths0392	0.5	1	A	A	A	A	A	A	B	B	A	A	B	A	A	A	B	U	A
MS_At1_11.2	1.5	1	A	B	A	A	A	A	B	B	A	A	B	A	A	A	B	A	A
f140426482	1.6	1	A	A	A	A	A	A	A	A	A	B	A	A	A	A	B	A	A
t27k12-sp6	2.1	1	A	A	A	A	A	A	A	A	A	B	A	A	A	A	B	U	A
~~MS_At1_16.4~~		1	A	A	A	A	A	A	A	A	A	U	A	A	A	A	B	A	A
MS_At1_20_a	0.1	1	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A
nga128	0.4	1	A	B	A	A	A	A	A	A	A	A	A	B	A	A	A	U	A
MS_At1_22.6	0.5	1	A	A	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A
t12p18ind8-8	0.0	1	U	A	A	A	A	A	A	A	A	B	A	A	U	A	A	A	A
f5i1449495	0.1	1	A	B	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A
f5a1859436	0.1	1	A	B	A	A	A	A	A	A	A	B	A	B	A	B	B	U	A
nga111	0.1	1	A	B	A	A	U	A	A	A	A	B	A	B	A	B	B	A	A
yup8h12r	0.3	1	A	B	A	A	A	A	A	A	A	B	A	B	A	B	B	A	A
nga1145	0.0	2	A	A	B	B	B	A	A	U	B	B	A	A	A	B	A	B	A
ciw2	0.0	2	A	A	U	B	B	A	A	B	B	B	A	A	A	B	A	B	A
t12j241450	0.1	2	A	A	A	B	B	A	A	U	B	B	A	A	A	A	A	B	A
ciw3	0.0	2	A	A	A	B	B	A	A	B	B	A	A	A	A	A	A	B	A
pls1	0.2	2	A	A	A	B	B	A	A	U	A	A	A	A	A	A	A	B	B
pls8	0.3	2	A	A	A	A	A	A	B	U	B	A	B	A	A	A	A	B	B
athgpa1	0.0	2	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	B	B
nga1126	0.0	2	A	A	A	A	A	A	A	B	A	A	U	A	A	A	A	B	B
t32f646516	0.2	2	A	A	B	A	A	A	A	A	B	A	A	A	A	A	A	B	B
MS_At2_14.2	0.2	2	A	A	B	A	A	A	A	A	B	A	A	A	A	A	A	B	B
nga168	0.4	2	A	A	U	A	A	A	B	A	B	A	U	B	A	A	A	B	B
90j19t7	0.1	2	B	B	U	A	A	B	B	A	B	A	U	B	A	A	A	B	B
nga32	1.3	3	A	A	A	A	A	B	A	A	A	B	A	A	A	A	A	A	B
f17a931902	0.8	3	A	A	A	A	A	B	B	A	A	B	B	A	A	B	A	A	B
nga162	0.4	3	A	B	A	A	A	B	B	A	A	B	B	A	A	B	B	A	B
msd2129380	0.9	3	A	A	A	U	U	B	A	A	A	A	A	A	B	B	B	B	B
MS_At3_8.3	0.8	3	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B
athgapab	1.5	3	A	B	U	A	B	A	A	A	A	A	A	A	B	B	B	A	A
t13j1016610	1.4	3	A	B	A	A	U	A	A	A	A	A	A	A	B	B	B	A	B
k11j14ind16-16	1.3	3	A	A	A	A	B	A	A	A	A	A	A	A	B	B	B	A	B
MS_At3_18.3	1.9	3	U	A	A	A	B	A	A	A	A	A	A	A	B	B	A	B	B
t16k521877	2.1	3	A	B	A	A	B	A	A	A	A	A	A	A	B	B	A	B	B
athcdc2bg	1.9	3	A	A	A	A	B	A	A	A	A	U	A	A	B	B	A	A	B
nga6	1.6	3	A	A	A	A	B	A	A	A	B	A	A	A	B	B	A	A	B
jv30/31	0.1	4	B	A	U	A	B	B	B	B	B	B	U	B	B	B	U	B	A
MS_At4_0.3	0.1	4	B	A	U	A	B	B	B	B	B	B	B	B	B	B	A	B	A
t19j18ind30-30	0.0	4	U	A	B	A	B	B	B	U	B	B	B	B	B	B	B	B	A
nga8	0.1	4	A	A	A	A	A	B	B	B	A	B	B	B	B	B	A	B	A
MS_At4_5.3	0.2	4	B	A	A	A	A	B	B	A	A	B	B	B	B	B	A	B	A
ciw6	1.5	4	A	A	A	B	A	B	A	B	A	A	A	B	B	B	A	B	A
MS_At4_10.8	0.9	4	A	A	U	A	A	B	A	B	A	A	A	B	A	B	A	A	A
t22a6ind10-10	1.4	4	A	B	A	A	A	B	A	B	B	A	A	A	A	B	A	A	A
f27g1959898	1.4	4	A	A	U	A	A	B	A	B	B	A	A	B	A	A	A	A	A
~~t27e11ind12-12~~		4	A	B	A	A	A	B	A	B	B	A	A	A	A	A	B	A	A
nga1139	1.9	4	A	A	A	A	A	B	A	U	A	A	A	B	A	A	B	A	B
nga1107	1.3	4	A	A	A	B	A	A	A	A	A	A	A	B	A	A	B	A	B
t5j1748070	0.6	4	A	A	A	B	A	A	B	A	A	A	B	B	A	A	B	A	B
nga225	7.6	5	B	B	U	B	B	B	A	B	B	B	A	B	B	B	B	B	B
mhf15ind52-52	6.9	5	B	B	B	B	B	B	A	B	B	B	A	B	B	B	B	B	B
ciw14	6.7	5	B	B	B	B	B	B	A	B	B	B	A	B	B	B	B	B	B
f14f1847646	11.8	5	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
nga151	11.2	5	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
nga139	4.4	5	A	B	U	B	B	B	B	A	B	B	B	B	B	A	B	A	B
MS_At5_9.7	0.3	5	A	A	U	B	A	A	A	A	B	B	A	A	B	A	A	A	B
MS_At5_13.9	0.1	5	A	U	B	B	A	U	A	A	B	A	A	A	B	A	B	A	B
mlf18ind14-14	0.0	5	A	A	B	B	A	A	A	A	B	A	A	A	U	A	A	A	B
aths0191	0.0	5	A	B	B	A	A	A	A	A	B	U	A	A	U	A	A	A	B
MS_At5_17.8	0.6	5	A	A	A	A	A	A	A	A	B	A	A	A	B	A	A	A	B
mq158836	0.6	5	A	A	A	U	U	A	A	A	B	A	A	A	B	A	A	A	B
MS_At5_20.4	0.6	5	A	A	A	A	B	A	A	A	B	A	A	A	A	A	A	A	B
jv65-66	0.5	5	U	A	A	A	B	A	B	A	A	A	B	A	U	A	A	A	A
mrg2138439	0.4	5	B	A	A	B	B	A	B	A	A	A	B	A	A	B	A	A	A
k8a1022396	0.3	5	B	A	U	B	A	A	A	A	A	A	A	A	A	B	A	A	A

TJ79 18:2/18:1

;Ts-5 x Mz-0 recombinant F8 population genotype file;
;A = Ts-5,B = Mz-0,U = unclassified,
A is increasing allele
18:2/18:1

rep1	1.85	1.55	1.8	1.53	1.67	1.85	1.22	1.77
rep2	1.96	1.68	1.93	1.72	1.81	1.81	1.67	1.96
rep3	1.80	1.72	1.55	1.76	1.88	1.94	1.74	1.86
ave.	**1.87**	**1.65**	**1.74**	**1.67**	**1.78**	**1.87**	**1.54**	**1.86**
no. of A	65	60	52	51	48	44	33	32

markers	LOD	LG	68	79	35	83	24	15	80	46
MS_At1_0.9	0.0	1	A	A	A	A	B	A	B	B
jconn1_2.7	0.0	1	A	A	A	A	B	A	B	B
jconn1_3.0	0.0	1	A	A	A	A	B	A	B	B
MS_At1_4.3	0.4	1	A	A	A	A	B	A	B	B
athsrp54a	0.2	1	A	A	A	A	B	A	B	B
f20d23ind18-18a	0.2	1	A	A	A	A	U	A	B	B
f9h16ind26-26	0.7	1	A	U	A	U	B	A	B	B
f8k728985	0.2	1	A	A	A	A	B	A	B	B
MS_At1_8.2	0.2	1	A	A	A	A	B	A	B	B
nga392	1.0	1	A	A	A	U	B	A	B	B
aths0392	2.0	1	B	A	A	A	B	A	U	B
MS_At1_11.2	2.0	1	B	A	A	A	B	A	B	B
MS_At1_12.9	2.6	1	B	A	A	B	B	A	B	B
t27k12-sp6	2.0	1	B	A	A	B	B	A	B	B
MS_At1_20b	1.4	1	A	A	A	B	A	A	B	A
MS_At1_20a	1.5	1	A	A	A	B	A	A	B	A
nga128	2.1	1	A	A	A	B	A	A	B	A
MS_At1_22.1	1.6	1	A	A	A	B	A	U	B	A
MS_At1_22.9	1.8	1	A	A	A	B	A	A	B	A
MS_At1_23.4	2.1	1	A	A	A	B	A	A	B	A
f5i1449495	2.5	1	A	A	A	B	A	U	B	B
MS_At1_24.3	2.4	1	A	A	A	B	A	A	B	B
f5a1859436	1.5	1	A	A	A	B	B	B	U	A
nga111	0.6	1	A	A	A	B	B	B	A	A
jconn2_0.7	0.0	2	A	A	A	A	A	A	A	A
MS_At2_3.0	0.0	2	A	A	A	A	A	A	A	B
MS_At2_5.4	0.0	2	A	A	A	A	A	A	A	B
MS_At2_6.3	0.1	2	A	A	A	U	A	U	A	B
~~jconn2_3.0~~		2	A	A	A	A	A	A	A	B
~~jconn2_5.2~~		2	A	A	A	A	A	A	A	B
ciw3	0.1	2	A	U	A	A	A	A	A	B
MS_At2_9.3	0.1	2	A	B	A	A	A	B	A	B
pls8	0.0	2	A	B	A	A	A	B	U	B
nga1126	0.0	2	A	B	A	A	A	B	A	B
c4h	0.1	2	A	B	A	A	A	B	A	B
t6a23ind10-10	0.0	2	A	A	A	B	A	B	A	B
nga168	0.0	2	A	A	A	B	A	B	A	B
MS_At2_17.5	0.1	2	A	B	A	B	A	U	A	B
MS_At2_17.9	0.0	2	A	B	A	B	A	B	A	U
athubique	0.2	2	A	B	A	B	A	B	A	B
nga32	0.2	3	A	U	A	A	B	B	B	A
nga162	0.0	3	A	B	A	B	A	B	A	A
MS_At3_6.1	0.3	3	A	A	A	B	A	B	A	A
msd2129380	0.1	3	A	A	A	A	A	B	A	A
mzn14inda9-a9	0.4	3	B	A	U	B	A	B	U	A
t13j1016610	0.3	3	B	A	B	B	A	B	A	U
k11j14ind16-16	0.3	3	B	A	U	B	A	B	A	A
MS_At3_17.3	0.1	3	A	A	B	A	A	A	A	A
t16k521877	0.5	3	A	A	B	A	A	A	A	A
MS_At3_19.7	0.4	3	A	A	B	A	A	A	A	B
athcdc2bg	0.3	3	A	A	B	A	U	A	A	B
nga6	0.3	3	A	A	B	A	A	A	A	B
MS_At4_0.3	1.3	4	B	B	A	A	A	U	B	A
MS_At4_1.1	1.2	4	B	B	B	A	A	U	B	A
t19j18ind30-30	2.0	4	B	B	B	A	A	A	B	A
~~nga8~~		4	B	B	B	A	A	A	B	A
det1.2	0.5	4	B	B	B	A	B	A	A	A
~~f28m11ind22-22~~		4	B	B	B	A	B	A	A	A
ciw6	0.2	4	A	A	B	A	B	A	A	A
t6k21ind15-15	0.3	4	A	A	B	A	B	A	B	A
ciw7	0.1	4	A	A	B	A	B	A	B	B
nga1139	0.8	4	U	A	B	B	B	A	B	B
MS_At4_16.2	0.6	4	A	A	B	B	B	U	B	B
t5j1748070	0.3	4	A	A	A	B	A	A	B	B
nga225	0.0	5	A	A	B	A	B	B	B	A
MS_At5_2.2	0.0	5	A	A	B	A	B	B	B	B
MS_At5_2.8	0.0	5	A	A	B	A	B	B	B	B
MS_At5_3.6	0.1	5	A	A	B	A	B	B	B	B
f14f1847646	0.2	5	A	A	B	A	B	B	B	B
nga151	0.3	5	A	U	B	A	B	B	B	B
MS_At5_6.6a	0.2	5	A	A	B	A	B	B	B	B
MS_At5_6.6b	0.2	5	A	A	B	A	B	B	B	B
MS_At5_8.6	0.0	5	A	B	B	B	A	U	B	B
nga139	0.0	5	U	B	B	B	A	A	B	B
MS_At5_9.7	0.1	5	A	A	B	B	A	U	B	B
mlf18ind14-14	0.0	5	B	A	A	B	U	A	B	B
MS_At5_15.8	0.7	5	B	A	B	A	A	U	B	A
ciw9	1.1	5	B	B	B	A	A	B	B	B
MS_At5_18	1.4	5	B	B	B	B	A	U	B	B
f20d23ind18-18b	4.5	5	B	B	B	B	U	B	B	B
nga129	4.8	5	B	B	B	B	B	B	B	B
MS_At5_20.4	4.7	5	B	B	B	B	B	U	B	B
MS_At5_21.3	2.6	5	B	B	B	B	B	U	A	B
jv65-66	2.6	5	A	U	A	A	B	B	A	A
MS_At5_22.8	3.2	5	A	B	A	A	B	U	A	A
mth12	3.5	5	A	A	A	A	A	A	A	A
MS_At5_26.1	3.2	5	A	A	A	A	A	U	B	A
k8a1022396	1.6	5	A	A	A	A	A	A	B	A

;Sorbo x Gy-0 recombinant F8 population genotype file;
;A = Sorbo,B = Gy-0,U = unclassified,
B is increasing allele
18:3

rep1	21.90	20.95	21.27	21.38	21.13	22.33	22.35	21.45	20.62	22.40	21.63
rep2	21.32	22.03	*	22.30	22.66	23.50	23.03	23.11	21.00	22.67	*
rep3	19.54	20.17	20.11	19.31	20.37	20.36	20.70	19.90	19.55	19.81	20.55
ave.	**20.92**	**21.05**	**20.69**	**21.00**	**21.39**	**22.07**	**22.03**	**20.66**	**20.26**	**21.63**	**21.09**
no. of A	53	53	49	48	47	45	45	42	42	42	41

markers	LOD	LG	76	78	52	34	93	8	10	33	43	75	61
t25k16ind9-9	0.5	1	B	B	A	A	U	B	B	B	B	A	A
jconnchr1_2.7x	0.1	1	B	B	A	A	B	A	A	B	B	A	A
MS_At1_3.7	0.1	1	B	B	A	A	B	A	A	B	B	A	A
athsrp54a	0.2	1	A	A	A	A	U	A	A	B	B	A	A
f9h16ind26-26	0.1	1	A	A	A	A	B	A	A	B	B	A	A
MS_At1_8.2	0.0	1	A	A	B	B	B	A	A	B	B	A	A
nga392	0.0	1	A	A	A	B	B	U	A	B	B	U	B
aths0392	0.0	1	A	A	A	B	B	A	A	B	B	A	B
MS_At1_12.9	0.8	1	A	A	B	A	B	A	A	B	B	A	B
f140426482	0.4	1	A	A	A	A	B	A	A	B	A	A	B
t27k12-sp6	0.5	1	A	A	A	A	B	A	A	B	A	A	B
MS_At1_18.2	1.1	1	A	A	A	A	A	A	A	A	A	A	B
MS_At1_20_b	0.8	1	A	A	A	A	A	A	A	A	A	B	B
nga128	0.5	1	A	A	B	A	A	A	A	A	A	B	A
MS_At1_22.6	0.6	1	A	A	A	A	A	A	A	A	A	B	A
t12p18ind8-8	0.1	1	A	A	B	A	U	A	A	A	B	U	A
MS_At1_24.3	0.1	1	A	A	A	A	A	A	A	A	B	B	A
MS_At1_26.2	0.0	1	A	A	A	B	A	A	A	A	B	B	A
MS_At1_27.2	0.4	1	A	A	A	B	A	A	A	A	B	B	A
nga1145	0.1	2	A	A	A	A	B	B	U	A	A	A	A
jconnchr2_3	0.0	2	A	A	A	A	B	B	B	A	A	A	A
pls1	0.3	2	A	A	A	A	A	A	B	A	A	A	A
pls8	0.1	2	A	A	A	A	A	A	A	A	A	A	A
nga1126	0.3	2	A	A	A	A	U	A	A	A	A	A	A
c4h	0.1	2	A	A	A	A	A	A	A	A	A	A	A
t32f646516	0.1	2	A	A	A	A	A	A	A	A	A	A	A
MS_At2_14.2_a	0.0	2	A	A	A	A	A	A	A	A	A	A	A
t6a23ind10-10	0.1	2	A	U	A	A	U	A	A	A	A	A	B
nga168	0.1	2	A	A	A	A	U	A	A	A	A	A	B
MS_At2_16.9	0.0	2	A	A	A	A	A	A	A	A	A	A	B
MS_At2_18.5	0.3	2	A	A	A	A	A	A	A	A	A	A	B
90j19t7	1.1	2	A	A	A	A	A	A	A	A	A	A	B
MS_At3_0.9	0.2	3	A	A	B	A	B	B	A	A	B	A	A
f17a931902	0.0	3	A	A	A	A	A	U	A	A	B	A	A
nga162	0.0	3	A	A	A	A	A	A	B	A	B	A	A
mil23ind14-14	0.2	3	A	A	A	A	A	B	B	A	B	A	A
msd2129380	0.2	3	A	A	A	A	A	B	U	A	B	A	A
MS_At3_8.9_a	0.3	3	A	A	B	A	A	U	A	B	B	A	B
athgapab	0.4	3	A	A	A	A	U	A	A	B	B	U	B
t13j1016610	1.5	3	A	A	A	A	A	A	A	A	A	A	A
k11j14ind16-16	1.2	3	A	A	A	A	U	A	A	A	A	A	A
MS_At3_16.9	2.9	3	A	A	B	A	A	U	B	A	A	A	A
MS_At3_17.3	2.9	3	A	A	B	A	A	B	B	A	A	A	A
t16k521877	2.9	3	A	A	A	A	A	B	B	A	A	A	A
f15b8ind11-11	1.8	3	A	A	A	A	A	B	B	A	A	B	A
nga6	0.6	3	A	A	A	A	A	B	B	A	A	A	A
t19j18ind30-30	0.6	4	B	B	A	U	B	A	A	A	B	A	B
nga8	0.8	4	B	B	A	A	A	A	A	A	B	A	B
f28m11ind22-22	0.9	4	B	B	A	A	A	B	B	A	B	A	B
ciw6	0.8	4	B	B	A	A	A	B	B	A	B	B	B
MS_At4_7.7	0.4	4	B	B	A	A	A	B	B	A	B	B	B
ciw7	1.3	4	B	B	A	A	A	B	B	B	B	B	A
f27g1959898	1.2	4	B	B	A	A	A	B	U	B	B	B	A
nga1139	2.3	4	A	A	A	B	A	A	A	B	B	A	A
MS_At4_16.2	2.6	4	A	A	A	B	A	A	A	B	B	A	A
t5j1748070	1.1	4	U	A	A	B	A	A	B	B	B	A	U
nga225	0.0	5	U	B	B	B	A	A	B	A	A	B	A
mhf15ind52-52	1.1	5	B	B	B	B	A	B	B	B	A	B	A
ciw14	0.8	5	B	B	B	B	A	B	B	B	A	B	A
MS_At5_3.6	0.8	5	B	B	B	B	A	B	B	B	A	B	B
nga151	1.0	5	B	B	B	B	A	B	B	B	A	B	A
MS_At5_5.3	0.7	5	B	B	B	B	A	B	B	B	A	B	A
jconnchr5_7.8	1.8	5	B	B	A	B	A	B	B	B	A	B	B
MS_At5_8.6	2.4	5	B	B	B	B	A	B	B	B	A	B	B
jconnchr5_13.6	2.1	5	B	B	B	B	B	B	B	B	B	B	B
aths0191	3.1	5	B	B	B	B	B	B	B	B	B	B	B
MS_At5_15.8	2.8	5	A	A	B	B	B	B	B	B	A	B	B
ciw9	0.9	5	A	A	B	B	B	A	A	B	A	B	B
MS_At5_18	1.7	5	A	A	B	B	B	A	A	B	A	B	B
mq158636	1.4	5	A	A	B	U	U	A	A	A	A	B	B
nga129	1.3	5	A	A	B	B	A	A	A	A	A	B	B
MS_At5_21.3	1.0	5	A	A	B	B	A	A	A	A	A	B	B
MS_At5_22.8	1.9	5	A	A	B	B	A	A	A	B	A	B	B
MS_At5_26.1	2.4	5	A	A	B	A	A	A	A	B	A	B	B

;Ts-5 x Mz-0 recombinant F8 population genotype file;
;A = Ts-5,B = Mz-0,U = unclassified,
A is increasing allele
18:3

rep1	22.67	20.71	20.49	20.9	21.02	21.21	22.36	21.21
rep2	21.14	21.27	18.77	20.24	20.67	20.79	21.58	19.24
rep3	20.34	19.47	18.80	18.61	19.72	20.77	16.48	19.33
ave.	**21.38**	**20.48**	**19.35**	**19.92**	**20.47**	**20.92**	**20.14**	**19.93**
no. of A	65	64	54	53	51	48	48	45

markers	LOD	LG	68	53	17	45	18	24	57	1
MS_At1_0.9	0.1	1	A	A	A	B	B	B	B	B
jconn1_2.7	0.9	1	A	A	A	B	U	B	B	U
jconn1_3.0	1.1	1	A	A	B	B	B	B	B	U
MS_At1_4.3	1.7	1	A	A	A	B	B	B	B	B
athsrp54a	2.1	1	A	A	A	B	A	B	B	B
f20d23ind18-18a	1.8	1	A	A	B	B	A	U	B	B
f9h16ind26-26	2.8	1	A	A	B	B	A	B	B	B
f8k728985	2.0	1	A	A	B	B	B	B	B	B
MS_At1_8.2	4.0	1	A	B	B	B	B	B	B	B
nga392	5.4	1	A	B	B	B	B	B	B	B
aths0392	6.3	1	B	B	B	B	B	B	B	B
MS_At1_11.2	5.7	1	B	B	B	B	B	B	B	B
MS_At1_12.9	2.2	1	B	B	B	B	B	B	B	B
t27k12-sp6	1.3	1	B	B	B	B	B	B	A	B
MS_At1_20b	1.0	1	A	B	B	A	A	A	A	B
MS_At1_20a	1.1	1	A	B	B	A	A	A	A	B
nga128	1.3	1	A	A	B	A	A	A	A	B
MS_At1_22.1	1.3	1	A	A	B	A	A	A	A	B
MS_At1_22.9	1.3	1	A	A	B	A	A	A	A	B
MS_At1_23.4	1.5	1	A	A	B	A	A	A	A	B
f5i1449495	1.7	1	A	A	B	A	A	A	A	B
MS_At1_24.3	1.6	1	A	A	B	A	A	A	A	B
f5a1859436	1.0	1	A	A	B	A	A	B	A	A
nga111	0.4	1	A	A	B	A	A	B	A	A
jconn2_0.7	0.2	2	A	A	A	B	A	A	B	B
MS_At2_3.0	0.0	2	A	A	A	A	A	A	B	B
MS_At2_5.4	0.0	2	A	A	A	A	A	A	B	B
MS_At2_6.3	0.0	2	A	A	A	A	A	A	B	B
jconn2_3.0		2	A	A	A	A	A	A	B	B
jconn2_5.2		2	A	A	A	A	A	A	B	B
ciw3	0.0	2	A	A	A	U	A	A	B	U
MS_At2_9.3	1.0	2	A	A	B	A	A	A	A	A
pls8	0.6	2	A	A	B	A	A	A	A	A
nga1126	0.1	2	A	A	B	A	A	A	A	A
c4h	0.1	2	A	A	B	A	A	A	A	A
t6a23ind10-10	0.0	2	A	A	B	A	A	A	A	A
nga168	0.0	2	A	A	B	A	A	A	A	A
MS_At2_17.5	0.2	2	A	A	B	A	A	A	A	A
MS_At2_17.9	0.1	2	A	A	B	A	A	A	A	A
athubique	0.2	2	A	A	B	A	A	A	A	A
nga32	1.1	3	A	A	A	B	B	B	B	A
nga162	1.4	3	A	B	B	B	B	A	B	B
MS_At3_6.1	1.9	3	A	B	B	B	B	A	B	B
msd2129380	1.9	3	A	B	B	B	B	A	B	B
mzn14inda9-a9	0.5	3	B	B	U	A	A	A	B	B
t13j1016610	1.6	3	B	A	U	U	A	A	B	B
k11j14ind16-16	1.6	3	B	A	A	A	A	A	B	B
MS_At3_17.3	0.2	3	A	A	A	A	A	A	B	A
t16k521877	0.4	3	A	A	A	A	A	A	B	A
MS_At3_19.7	0.4	3	A	A	A	A	A	A	B	A
athcdc2bg	0.3	3	A	A	A	B	A	U	A	A
nga6	0.3	3	A	A	A	A	A	A	A	A
MS_At4_0.3	0.5	4	B	U	A	B	U	A	A	A
MS_At4_1.1	0.9	4	B	B	A	B	B	A	A	A
t19j18ind30-30	1.0	4	B	B	A	B	U	A	A	A
nga8		4	B	B	A	B	U	A	A	A
det1.2	0.6	4	B	B	A	B	B	B	U	A
f28m11ind22-22		4	B	B	A	B	B	B	A	A
ciw6	0.1	4	A	B	A	B	A	B	A	B
t6k21ind15-15	0.0	4	A	B	A	B	A	B	A	A
ciw7	0.0	4	A	B	A	B	A	B	B	A
nga1139	0.0	4	U	B	A	A	A	B	U	A
MS_At4_16.2	0.0	4	A	B	A	A	A	B	B	B
t5j1748070	0.4	4	A	B	A	A	A	A	B	A
nga225	1.7	5	A	A	A	A	B	B	A	A
MS_At5_2.2	0.8	5	A	A	A	A	B	B	A	A
MS_At5_2.8	0.6	5	A	A	A	A	B	B	A	A
MS_At5_3.6	0.1	5	A	A	A	A	B	B	A	A
f14f1847646	0.0	5	A	A	A	A	B	B	A	A
nga151	0.0	5	A	A	A	U	B	B	A	A
MS_At5_6.6a	0.0	5	A	A	A	A	B	B	A	A
MS_At5_6.6b	0.0	5	A	A	A	A	B	B	A	A
MS_At5_8.6	0.1	5	A	A	A	A	B	A	A	A
nga139	0.1	5	U	A	A	A	B	A	A	A
MS_At5_9.7	0.3	5	A	A	A	A	B	A	A	A
mlf18ind14-14	0.2	5	B	A	A	A	B	U	A	A
MS_At5_15.8	0.2	5	B	A	A	A	B	A	A	A
ciw9	0.1	5	B	A	A	B	B	A	U	A
MS_At5_18	0.2	5	B	A	A	B	B	A	A	A
f20d23ind18-18b	0.4	5	B	A	A	A	A	U	A	A
nga129	0.4	5	B	A	A	A	A	B	U	A
MS_At5_20.4	0.4	5	B	A	A	A	A	B	A	A
MS_At5_21.3	0.2	5	B	A	A	A	A	B	A	B
jv65-66	0.6	5	A	A	A	U	A	B	A	B
MS_At5_22.8	0.4	5	A	A	A	A	A	B	A	B
mth12	0.9	5	A	A	A	A	A	A	B	B
MS_At5_26.1	0.4	5	A	A	A	A	A	A	B	B
k8a10222396	0.1	5	A	A	A	A	B	A	B	B

;Nok-3 x Ga-0 RI F8 population, 94 individuals, genotype file,
;A = Nok-3 B = Ga-0 U = unclassified & heterozygotes
A is increasing allele
18:3/18:2

rep1	0.96	0.90	0.97	*	0.96	1.13	0.94	0.97	0.96	0.96	0.86
rep2	0.87	*	0.93	0.92	0.82	1.03	0.89	*	*	0.86	0.77
rep3	1.02	0.81	0.94	0.93	0.88	1.01	0.92	0.92	0.93	0.98	0.81
ave.	**0.95**	**0.85**	**0.95**	**0.93**	**0.89**	**1.05**	**0.92**	**0.94**	**0.94**	**0.93**	**0.81**
no. of A	54	51	50	50	49	49	48	48	47	47	47

markers	LOD	LG	36	32	6	72	29	47	41	88	28	37	91
nga59	0.6	1	A	B	A	A	B	A	A	A	B	A	B
MS_At1_0.6	0.3	1	B	B	A	A	B	A	A	A	B	A	B
MS_At1_2.7	0.2	1	B	B	A	B	B	A	A	A	B	A	U
nga63	0.0	1	A	B	A	A	B	A	A	A	B	A	B
f16j7-trb	0.0	1	A	B	A	B	B	A	A	A	B	A	B
athsrp54a	0.0	1	A	A	A	B	B	A	A	A	B	A	B
MS_At1_6.4	0.0	1	A	A	A	B	A	A	A	A	U	A	B
f9h16ind26-26	0.1	1	A	A	A	A	B	A	A	A	B	A	B
MS_At1_7.4	0.4	1	A	A	A	A	B	A	A	A	B	A	B
MS_At1_8.2	0.3	1	A	A	A	A	B	A	A	A	B	A	A
nga392	0.2	1	A	A	A	A	B	A	A	A	B	A	U
aths0392	0.2	1	A	A	A	A	B	A	A	A	B	A	A
MS_At1_11.2	0.5	1	B	A	A	A	B	A	A	A	B	A	A
f140426482	0.5	1	A	A	A	A	A	A	A	A	A	A	A
t27k12-sp6	0.9	1	A	A	A	A	A	A	A	A	A	A	A
MS_At1_16.4		1	A	A	A	A	A	A	A	A	A	A	A
MS_At1_20_a	1.5	1	A	A	A	A	A	A	A	A	A	A	A
nga128	2.0	1	B	A	A	A	A	A	A	A	A	A	A
MS_At1_22.6	3.0	1	A	A	A	A	A	A	A	A	A	A	A
t12p18ind8-8	2.0	1	A	A	A	A	A	A	A	U	A	A	B
f5i1449495	1.7	1	B	A	A	A	A	A	A	U	A	A	B
f5a1859436	0.1	1	B	A	A	A	A	A	A	B	A	A	A
nga111	0.2	1	B	A	U	A	A	A	A	B	A	A	A
yup8h12r	0.1	1	B	A	A	A	A	B	A	A	A	A	A
nga1145	0.1	2	A	B	B	A	A	B	B	A	A	B	B
ciw2	0.1	2	A	B	B	A	A	B	B	U	A	B	B
t12j241450	0.0	2	A	B	B	A	A	B	B	A	A	B	B
ciw3	0.0	2	A	B	B	A	A	B	B	A	A	B	B
pls1	0.6	2	A	B	B	A	A	A	A	A	A	A	B
pls8	0.8	2	A	A	A	A	B	A	B	A	B	B	A
athgpa1	0.4	2	A	A	A	A	A	A	A	A	A	B	A
nga1126	0.3	2	A	A	A	A	A	A	A	A	U	B	A
t32f646516	1.9	2	A	A	A	A	A	A	B	A	A	B	A
MS_At2_14.2	1.8	2	A	A	A	A	A	A	B	A	A	B	A
nga168	2.0	2	A	A	A	A	B	A	B	A	U	B	A
90j19t7	1.1	2	B	A	A	B	B	A	B	A	U	A	B
nga32	0.0	3	A	A	A	B	A	A	A	B	A	B	A
f17a9931902	0.0	3	A	A	A	B	B	A	A	B	B	B	A
nga162	0.1	3	B	A	A	B	B	A	A	B	B	B	B
msd2129380	0.4	3	A	U	U	B	A	A	A	B	A	A	A
MS_At3_8.3	0.2	3	A	A	A	A	A	A	A	B	A	A	A
athgapab	0.6	3	B	A	B	A	A	A	A	B	A	A	A
t13j1016610	2.1	3	B	A	U	A	A	B	A	A	A	A	A
k11j14ind16-16	1.3	3	A	A	B	A	A	B	A	A	A	A	A
MS_At3_18.3	1.6	3	A	A	B	A	A	B	A	A	A	A	A
t16k521877	2.3	3	B	A	B	A	A	B	A	A	A	A	A
athcdc2bg	1.3	3	A	A	B	A	A	B	A	A	A	A	B
nga6	0.4	3	A	A	B	A	A	A	B	A	A	B	B
jv30/31	0.5	4	A	A	B	B	B	B	B	A	U	B	A
MS_At4_0.3	0.5	4	A	A	B	B	B	B	B	A	B	B	A
t19j18ind30-30	1.1	4	A	A	B	B	B	B	B	A	B	B	A
nga8	0.8	4	A	A	A	B	B	B	A	A	B	B	A
MS_At4_5.3	1.3	4	A	A	A	B	B	B	A	A	B	B	A
ciw6	1.0	4	A	B	A	B	A	A	A	A	A	B	A
MS_At4_10.8	0.4	4	A	A	A	B	A	A	A	A	A	B	A
t22a6ind10-10	1.9	4	B	A	A	B	A	A	B	B	A	B	A
f27g1959898	1.0	4	A	A	A	B	A	A	B	B	A	A	A
t27e11ind12-12		4	B	A	A	B	A	A	B	B	A	A	A
nga1139	0.8	4	A	A	A	B	A	A	A	B	A	A	A
nga1107	0.3	4	A	B	A	A	A	A	A	B	A	B	A
t5j1748070	0.1	4	A	B	A	A	B	A	A	B	B	B	A
nga225	0.3	5	B	B	B	B	A	B	B	A	A	A	B
mhf15ind52-52	0.1	5	B	B	B	B	A	B	B	A	A	A	A
ciw14	0.1	5	B	B	B	B	A	B	B	A	A	A	A
f14f1847646	0.6	5	B	B	B	B	B	A	B	U	B	A	A
nga151	0.2	5	B	B	B	B	B	B	B	A	B	A	A
nga139	3.0	5	B	B	B	B	B	B	B	B	B	B	B
MS_At5_9.7	1.0	5	A	B	A	A	A	B	B	B	A	B	B
MS_At5_13.9	0.5	5	U	B	A	U	A	B	B	B	A	B	B
mlf18ind14-14	0.5	5	A	B	A	A	A	B	B	B	A	B	B
aths0191	0.3	5	B	A	A	A	A	B	B	B	A	B	B
MS_At5_17.8	0.1	5	A	A	A	A	A	B	B	B	A	A	A
mq158836	0.5	5	A	U	U	A	A	B	B	B	A	A	A
MS_At5_20.4	0.1	5	A	A	B	A	A	B	B	B	A	A	A
jv65-66	0.0	5	A	A	B	A	B	B	A	B	B	A	B
mrg2138439	0.0	5	A	B	B	A	B	A	A	B	B	B	B
k8a1022396	0.0	5	A	B	A	A	A	A	A	B	A	A	B

;Nok-3 x Ga-0 RI F8 population, 94 individuals, genotype file,
;A = Nok-3 B = Ga-0 U = unclassified & heterozygotes
A is increasing allele
18:3/18:2

rep1	*	0.94	*	*	0.99
rep2	0.86	0.92	0.94	0.86	*
rep3	0.88	0.91	0.99	0.86	0.94
average	**0.87**	**0.92**	**0.97**	**0.86**	**0.97**
no. of A	56	40	37	29	29

markers	LOD	LG	93	24	81	51	84
nga59	0.6	1	A	A	A	B	B
MS_At1_0.6	0.3	1	A	A	A	B	B
MS_At1_2.7	0.2	1	A	A	A	B	B
nga63	0.0	1	A	A	A	B	B
f16j7-trb	0.0	1	A	A	A	B	B
athsrp54a	0.0	1	A	A	B	B	B
MS_At1_6.4	0.0	1	A	A	A	B	B
f9h16ind26-26	0.1	1	A	A	A	B	B
MS_At1_7.4	0.4	1	U	A	A	B	B
MS_At1_8.2	0.3	1	A	A	A	B	A
nga392	0.2	1	A	A	A	B	A
aths0392	0.2	1	A	A	A	B	A
MS_At1_11.2	0.5	1	A	A	A	B	A
f140426482	0.5	1	A	A	A	A	A
t27k12-sp6	0.9	1	A	A	A	A	A
MS_At1_16.4		1	A	A	A	A	A
MS_At1_20_a	1.5	1	B	B	B	A	B
nga128	2.0	1	B	B	B	B	B
MS_At1_22.6	3.0	1	B	B	B	B	B
t12p18ind8-8	2.0	1	B	U	A	B	B
f5i1449495	1.7	1	B	B	A	B	B
f5a1859436	0.1	1	B	A	B	A	B
nga111	0.2	1	B	A	B	A	A
yup8h12r	0.1	1	B	A	B	A	B
nga1145	0.1	2	B	B	A	A	B
ciw2	0.1	2	B	U	A	A	B
t12j241450	0.0	2	A	B	A	A	B
ciw3	0.0	2	A	B	A	A	B
pls1	0.6	2	A	A	A	A	U
pls8	0.8	2	A	A	U	B	A
athgpa1	0.4	2	A	A	A	B	A
nga1126	0.3	2	A	A	A	B	A
t32f646516	1.9	2	A	A	A	B	A
MS_At2_14.2	1.8	2	A	A	A	B	A
nga168	2.0	2	A	A	A	B	A
90j19t7	1.1	2	A	A	A	B	A
nga32	0.0	3	A	B	B	A	B
f17a9931902	0.0	3	A	B	B	A	B
nga162	0.1	3	A	B	B	B	B
msd2129380	0.4	3	A	B	B	B	B
MS_At3_8.3	0.2	3	A	B	B	B	B
athgapab	0.6	3	A	B	B	B	B
t13j1016610	2.1	3	A	B	U	B	B
k11j14ind16-16	1.3	3	A	B	B	B	B
MS_At3_18.3	1.6	3	A	B	B	B	B
t16k521877	2.3	3	A	U	B	B	B
athcdc2bg	1.3	3	B	B	U	A	B
nga6	0.4	3	B	B	B	A	B
jv30/31	0.5	4	B	A	B	B	B
MS_At4_0.3	0.5	4	B	A	B	B	B
t19j18ind30-30	1.1	4	B	B	B	B	B
nga8	0.8	4	B	A	A	B	B
MS_At4_5.3	1.3	4	U	A	A	B	B
ciw6	1.0	4	A	A	A	B	B
MS_At4_10.8	0.4	4	A	A	A	B	A
t22a6ind10-10	1.9	4	A	B	B	A	B
f27g1959898	1.0	4	A	B	B	A	B
t27e11ind12-12		4	A	B	B	A	B
nga1139	0.8	4	A	B	B	A	B
nga1107	0.3	4	A	B	B	B	B
t5j1748070	0.1	4	A	B	B	B	B
nga225	0.3	5	A	B	A	B	A
mhf15ind52-52	0.1	5	A	B	A	B	A
ciw14	0.1	5	A	U	B	B	A
f14f1847646	0.6	5	A	A	A	B	A
nga151	0.2	5	A	A	A	A	A
nga139	3.0	5	A	A	A	A	A
MS_At5_9.7	1.0	5	A	A	B	A	A
MS_At5_13.9	0.5	5	U	A	B	A	A
mlf18ind14-14	0.5	5	A	A	B	A	A
aths0191	0.3	5	A	A	U	B	A
MS_At5_17.8	0.1	5	A	B	B	A	B
mq158836	0.5	5	A	B	B	A	B
MS_At5_20.4	0.1	5	A	B	B	A	B
jv65-66	0.0	5	A	B	B	A	A
mrg2138439	0.0	5	B	B	B	B	A
k8a1022396	0.0	5	B	B	B	B	A

;Sorbo x Gy-0 recombinant F8 population genotype file;
;A = Sorbo,B = Gy-0,U = unclassified,
B is increasing allele
18:3/18:2

rep1	0.75	0.79	*	0.80	0.80	0.74	0.81	0.78	0.74	0.75	0.73
rep2	0.71	*	0.79	0.84	0.78	0.71	0.76	0.81	0.85	0.74	0.73
rep3	0.75	0.77	0.78	0.77	0.80	0.73	0.80	0.82	0.81	0.69	0.75
ave	**0.74**	**0.78**	**0.78**	**0.80**	**0.79**	**0.73**	**0.79**	**0.80**	**0.80**	**0.73**	**0.74**
no. of A	55	49	46	45	45	44	44	39	38	36	36

markers	LOD	LG	91	52	66	8	10	30	42	26	90	27	79
t25k16ind9-9	0.1	1	U	A	A	B	B	B	A	A	U	B	A
jconnchr1_2.7x	0.1	1	B	A	A	A	A	A	B	A	B	B	A
MS_At1_3.7	0.0	1	B	A	A	A	A	A	B	A	B	B	A
athsrp54a	0.4	1	A	A	A	A	A	A	B	A	A	B	B
f9h16ind26-26	0.6	1	A	A	A	A	A	A	B	A	A	B	B
MS_At1_8.2	1.2	1	A	B	A	A	A	A	A	A	A	B	B
nga392	0.5	1	A	A	A	U	A	A	A	A	A	B	B
aths0392	0.7	1	A	A	A	A	A	A	A	A	A	B	B
MS_At1_12.9	0.4	1	A	B	A	A	A	A	A	A	A	B	B
f140426482	0.5	1	A	A	A	A	A	A	U	A	A	B	B
t27k12-sp6	0.5	1	A	A	A	A	A	A	A	A	A	B	U
MS_At1_18.2	0.5	1	A	A	A	A	A	A	B	A	A	B	B
MS_At1_20_b	0.0	1	A	A	A	A	A	A	A	A	A	A	B
nga128	0.1	1	A	B	U	A	A	B	A	A	U	U	B
MS_At1_22.6	0.1	1	A	A	A	A	A	B	A	A	A	A	B
t12p18ind8-8	0.1	1	A	B	A	A	A	U	A	A	A	A	B
MS_At1_24.3	0.2	1	A	A	A	A	A	B	A	A	A	A	B
MS_At1_26.2	0.6	1	A	A	A	A	A	A	A	A	A	A	A
MS_At1_27.2	0.2	1	A	A	A	A	A	A	A	A	A	A	A
nga1145	0.7	2	A	A	B	B	U	A	B	A	A	A	A
jconnchr2_3	0.3	2	A	A	B	B	B	A	B	A	A	B	A
pls1	0.8	2	A	A	B	A	B	A	A	A	B	B	A
pls8	0.4	2	A	A	B	A	A	A	A	A	B	B	A
nga1126	1.0	2	A	A	B	A	A	A	A	A	B	A	A
c4h	1.0	2	A	A	A	A	A	B	A	A	B	A	A
t32f646516	0.8	2	A	A	A	A	A	B	A	A	B	A	U
MS_At2_14.2_a	1.0	2	A	A	A	A	A	B	A	A	B	A	B
t6a23ind10-10	0.8	2	U	A	B	A	A	U	A	U	U	A	B
nga168	0.7	2	U	A	B	A	A	U	A	U	U	A	B
MS_At2_16.9	1.7	2	A	A	B	A	A	A	A	A	B	A	B
MS_At2_18.5	2.1	2	A	A	B	A	A	A	A	B	B	A	A
90j19t7	1.9	2	B	A	B	A	A	A	A	B	B	A	A
MS_At3_0.9	1.2	3	A	B	A	B	A	B	B	B	A	B	B
f17a931902	2.0	3	U	A	A	U	A	B	B	B	A	B	A
nga162	2.6	3	A	A	U	A	B	B	A	A	U	U	A
mil23ind14-14	2.6	3	A	A	A	B	B	B	B	U	A	A	A
msd2129380	3.6	3	A	A	A	B	U	B	B	A	A	A	A
MS_At3_8.9_a	3.5	3	A	B	B	U	A	B	B	B	B	A	A
athgapab	3.2	3	A	A	B	A	A	U	B	B	B	B	A
t13j1016610	4.0	3	B	A	B	A	A	B	B	B	B	B	A
k11j14ind16-16	3.8	3	B	A	B	A	A	B	B	B	B	B	B
MS_At3_16.9	6.1	3	B	B	B	U	B	B	B	B	B	B	B
MS_At3_17.3	6.1	3	B	B	B	B	B	B	B	B	B	B	B
t16k521877	5.1	3	B	A	B	B	B	U	B	B	B	B	B
f15b8ind11-11	3.9	3	B	A	B	B	B	A	U	B	B	B	B
nga6	1.1	3	B	A	B	B	B	A	U	B	B	B	B
t19j18ind30-30	0.0	4	A	A	B	A	A	B	B	A	A	B	B
nga8	0.0	4	A	A	U	A	A	B	B	A	A	B	B
f28m11ind22-22	0.1	4	A	A	A	B	B	B	B	A	A	B	B
ciw6	0.0	4	A	A	A	B	B	U	A	U	A	B	B
MS_At4_7.7	0.0	4	A	A	A	B	B	A	U	A	A	B	B
ciw7	0.0	4	A	A	A	B	B	A	A	A	A	B	B
f27g1959898	0.0	4	A	A	B	B	U	A	A	A	A	B	B
nga1139	0.2	4	A	A	B	A	A	B	B	A	A	B	B
MS_At4_16.2	0.2	4	A	A	B	A	A	U	A	A	A	B	B
t5j1748070	0.1	4	B	A	B	A	B	U	B	B	A	B	B
nga225	0.2	5	A	B	A	A	B	B	A	A	A	B	A
mhf15ind52-52	1.6	5	A	B	A	B	B	A	A	B	A	A	A
ciw14	1.4	5	A	B	A	B	B	A	U	B	A	A	A
MS_At5_3.6	1.9	5	A	B	A	B	B	A	A	B	A	A	A
nga151	2.1	5	A	B	U	B	B	A	A	B	A	A	A
MS_At5_5.3	1.5	5	A	B	A	B	B	A	A	B	A	A	A
jconnchr5_7.8	1.3	5	B	A	A	B	B	A	A	B	B	A	B
MS_At5_8.6	1.7	5	B	B	A	B	B	A	A	B	B	A	B
jconnchr5_13.6	0.2	5	B	B	A	B	B	A	A	B	A	A	A
aths0191	0.7	5	A	B	A	B	B	A	U	B	B	A	A
MS_At5_15.8	0.8	5	A	B	A	B	B	A	A	B	B	A	A
ciw9	0.0	5	A	B	A	A	A	A	A	B	B	A	A
MS_At5_18	0.5	5	A	B	A	A	A	A	A	B	B	A	A
mq158636	0.7	5	B	B	A	A	A	A	A	B	B	A	A
nga129	0.6	5	A	B	A	A	A	A	A	B	B	A	A
MS_At5_21.3	0.4	5	A	B	A	A	A	A	A	B	B	A	A
MS_At5_22.8	1.4	5	A	B	A	A	A	A	A	B	B	A	A
MS_At5_26.1	2.4	5	A	B	A	A	A	A	B	B	B	A	A

;Kondara x Br-0 recombinant F8 population genotype file;
;A = Kondara,B = Br-0, U = unclassified,
B is increasing
Oil content

rep1	43	43	42	44	43	38	42	42	46	44	*
rep2	43	43	42	44	43	38	42	42	46	44	*
rep3	45	45	46	46	41	46	42	45	48	47	44
ave	**43.93**	**43.66**	**43.99**	**44.49**	**42.38**	**40.24**	**42.18**	**43.17**	**46.40**	**44.88**	**43.90**
no. of A	45	42	39	39	38	37	34	34	33	32	32

markers	LOD	LG	55	49	28	78	51	58	38	94	77	15	39
nga63	4.0	1	A	A	B	A	A	A	B	A	B	U	A
athsrp54a	2.4	1	A	A	B	A	A	A	B	A	B	B	A
f9h16ind26-26	2.1	1	A	A	B	A	A	A	B	U	B	B	A
MS_At1_10	2.7	1	A	A	B	B	A	A	B	A	B	B	A
aths0392	0.8	1	A	A	B	B	A	B	B	U	B	B	A
MS_At1_13.4	0.0	1	A	A	B	A	B	B	B	A	A	B	A
MS_At1_16.2	0.2	1	A	A	B	B	B	B	B	A	A	B	A
MS_At1_18.2	0.0	1	A	A	B	A	B	B	B	A	A	B	A
MS_At1_22.1	0.1	1	A	A	B	A	B	B	B	A	A	A	A
t12p18ind8-8	0.0	1	A	A	B	A	A	B	B	A	B	U	A
MS_At1_24.7	0.0	1	A	A	B	A	A	B	B	A	B	U	A
f5a1859436	0.0	1	A	A	U	A	A	B	B	A	B	A	A
f24j13ind32-32	0.0	1	A	A	U	A	A	B	B	A	B	A	A
nga1145	0.8	2	A	A	A	A	A	A	A	B	B	B	B
t16b233819	1.2	2	A	A	A	A	A	A	A	B	B	B	B
jconnchr2_3	0.9	2	A	A	A	A	A	A	A	B	A	U	B
MS_At2_4.4	0.3	2	A	A	A	A	A	A	A	B	A	B	B
jconnchr2_5.5	1.2	2	A	A	A	A	U	A	A	B	A	B	B
MS_At2_6.3	0.4	2	A	A	A	A	B	A	A	B	A	B	B
pls1	1.4	2	A	A	A	B	A	A	A	B	A	U	A
pls8	1.5	2	A	A	A	B	A	A	A	A	A	B	A
nga1126	1.5	2	A	A	A	B	A	A	A	A	A	B	A
c4h	2.4	2	A	A	A	B	A	A	A	A	A	A	A
MS_At2_14.2	1.8	2	A	A	A	B	A	B	A	A	A	B	A
t6a23ind10-10	2.9	2	A	A	A	A	A	B	A	A	A	B	A
nga168	2.9	2	A	A	A	A	A	B	A	A	A	B	A
MS_At2_17.5	4.1	2	A	A	A	A	A	A	A	A	A	A	A
athubique	5.7	2	A	A	A	A	A	A	A	A	A	B	A
90j19t7	6.0	2	A	A	A	A	A	A	A	A	A	A	A
nga32	0.1	3	A	A	B	A	A	B	A	A	B	A	A
f17a931902	0.1	3	A	U	B	A	U	U	A	B	B	A	A
nga162	1.0	3	A	A	A	A	A	B	A	U	B	A	A
msd2129380	0.1	3	A	A	A	A	U	B	A	B	B	U	B
athgapab	0.0	3	A	A	A	A	B	B	A	B	B	U	B
k11j14ind16-16	0.2	3	A	A	B	A	B	B	A	B	B	A	A
MS_At3_16	0.2	3	B	A	B	A	B	B	A	B	B	A	A
MS_At3_16.9	0.5	3	B	A	B	A	B	A	A	B	A	A	A
t16k521877	0.1	3	B	B	B	A	B	A	A	A	A	A	A
MS_At3_19.7	0.2	3	B	B	B	A	A	U	A	A	A	A	B
MS_At3_21.4	0.5	3	A	B	B	A	B	U	A	A	A	A	B
jv30-31	0.0	4	B	B	A	A	B	A	B	B	B	A	A
MS_At4_1.1	0.2	4	B	B	B	A	B	A	B	A	B	B	B
t19j18ind30-30	0.2	4	U	B	B	A	B	U	B	A	A	A	B
nga8	0.0	4	B	B	A	A	B	A	B	A	A	A	B
f28m11ind22-22	0.0	4	U	B	A	A	B	A	B	A	A	A	B
ciw6	0.1	4	A	B	A	B	B	A	A	A	U	A	B
MS_At4_8.8	0.0	4	B	B	A	B	B	A	A	A	A	A	B
t6k21ind15-15	0.1	4	B	B	A	B	B	A	A	A	A	B	B
ciw7	0.1	4	B	B	A	B	B	A	B	A	A	U	B
MS_At4_11.2	0.1	4	B	B	A	B	B	A	B	A	A	B	B
MS_At4_13.2	0.0	4	U	B	A	B	B	A	B	A	A	B	B
nga1139	0.7	4	B	A	A	A	U	A	A	B	A	B	B
MS_At4_16.1	0.3	4	B	A	A	A	B	A	A	B	A	B	B
t5j1748070	0.0	4	B	A	A	B	A	A	A	B	A	B	B
nga225	0.1	5	A	A	A	B	A	B	B	B	B	A	B
mhf15ind52-52	0.1	5	A	A	A	B	A	B	B	B	B	A	B
MS_At5_2.2	0.1	5	A	A	A	B	A	B	B	B	B	A	B
f14f1847646	0.1	5	A	B	A	B	A	A	B	B	B	A	B
nga151	0.2	5	A	B	A	B	A	A	B	B	B	A	B
MS_At5_5.3	0.1	5	A	B	A	B	A	A	B	B	B	A	B
MS_At5_6.6_a	0.1	5	A	B	A	B	A	A	B	B	B	A	B
nga139	0.0	5	A	B	B	B	A	A	B	B	B	B	B
jconnchr5_12.7	0.1	5	B	B	B	B	A	B	B	B	B	B	B
jconnchr5_13.6	0.0	5	B	B	B	B	A	B	B	B	B	B	B
aths0191	0.1	5	B	B	B	B	B	B	B	B	B	A	B
ciw9	1.7	5	B	B	B	B	A	B	B	B	B	A	B
mlnlind8-8	1.2	5	B	B	B	B	B	B	B	B	B	A	B
mql58836	1.7	5	B	B	B	B	B	B	B	B	B	A	B
nga129	2.2	5	B	B	A	B	B	B	B	B	B	B	B
MS_At5_20.0	1.8	5	B	B	U	B	B	B	B	B	B	U	B
jv65-66	3.2	5	B	B	B	B	B	B	B	B	B	B	B
mth12	1.9	5	B	B	B	B	B	B	B	B	B	B	B
mrg2138439	1.1	5	B	B	B	B	B	B	B	B	B	B	B
k8a1022396	1.3	5	B	B	B	B	B	B	A	B	B	B	A

;Kondara x Br-0 recombinant F8 population genotype file;
;A = Kondara,B = Br-0, U = unclassified,
B is increasing
Oil content

rep1	45	43	44	44	46	44	45	45	46
rep2	45	43	44	44	46	44	45	45	46
rep3	47	47	46	45	46	44	46	46	46
ave.	**45.42**	**44.55**	**45.00**	**44.34**	**45.79**	**43.97**	**45.15**	**45.23**	**45.91**
no. of A	39	33	32	31	29	29	28	28	27

markers	LOD	LG	26	91	86	6	17	92	11	79	67
nga63	4.0	1	B	A	B	A	A	B	B	B	U
athsrp54a	2.4	1	B	A	B	A	A	A	U	B	B
f9h16ind26-26	2.1	1	B	A	B	A	U	A	B	B	U
MS_At1_10	2.7	1	B	A	B	A	B	B	B	B	U
aths0392	0.8	1	B	A	B	A	U	A	B	B	U
MS_At1_13.4	0.0	1	A	A	B	A	B	B	B	B	B
MS_At1_16.2	0.2	1	A	A	B	A	B	B	B	B	B
MS_At1_18.2	0.0	1	A	A	B	B	B	B	B	B	B
MS_At1_22.1	0.1	1	A	B	B	B	A	B	B	B	B
t12p18ind8-8	0.0	1	A	B	B	B	A	B	B	B	B
MS_At1_24.7	0.0	1	A	B	B	B	A	A	B	B	B
f5a1859436	0.0	1	A	A	A	A	A	A	B	B	A
f24j13ind32-32	0.0	1	A	A	A	A	A	A	B	B	A
nga1145	0.8	2	B	A	B	B	A	A	U	A	A
t16b233819	1.2	2	A	A	B	B	A	A	B	A	A
jconnchr2_3	0.9	2	A	A	B	B	A	A	B	U	A
MS_At2_4.4	0.3	2	A	A	B	B	A	A	B	A	A
jconnchr2_5.5	1.2	2	A	A	U	B	A	A	B	U	A
MS_At2_6.3	0.4	2	U	A	B	B	A	A	B	U	A
pls1	1.4	2	U	B	B	A	A	B	B	U	A
pls8	1.5	2	B	B	B	B	B	B	B	B	A
nga1126	1.5	2	B	B	B	B	B	B	A	B	B
c4h	2.4	2	B	B	B	B	B	B	A	B	B
MS_At2_14.2	1.8	2	B	B	B	B	B	A	A	B	B
t6a23ind10-10	2.9	2	B	B	B	B	B	B	A	B	B
nga168	2.9	2	B	B	B	B	B	B	A	B	B
MS_At2_17.5	4.1	2	B	B	B	B	B	A	A	B	B
athubique	5.7	2	B	B	B	B	B	B	B	B	B
90j19t7	6.0	2	B	B	B	B	B	B	B	B	B
nga32	0.1	3	A	B	B	A	B	U	B	B	B
f17a931902	0.1	3	A	B	B	A	B	U	B	B	B
nga162	1.0	3	B	B	B	A	U	B	B	B	U
msd2129380	0.1	3	B	A	B	A	B	B	B	U	B
athgapab	0.0	3	B	A	A	B	B	A	B	A	A
k11j14ind16-16	0.2	3	B	A	U	B	B	A	B	A	B
MS_At3_16	0.2	3	B	A	A	B	B	B	B	A	B
MS_At3_16.9	0.5	3	B	A	B	B	B	B	B	A	B
t16k521877	0.1	3	A	A	B	B	B	A	B	A	B
MS_At3_19.7	0.2	3	A	B	A	B	B	B	B	A	B
MS_At3_21.4	0.5	3	A	B	A	A	B	B	B	A	B
jv30-31	0.0	4	A	B	B	A	B	B	U	B	B
MS_At4_1.1	0.2	4	A	B	U	A	B	B	B	B	B
t19j18ind30-30	0.2	4	A	B	B	A	B	B	A	B	B
nga8	0.0	4	B	B	B	A	B	B	A	B	B
f28m11ind22-22	0.0	4	B	B	A	A	B	B	A	B	B
ciw6	0.1	4	B	B	A	A	B	B	A	B	B
MS_At4_8.8	0.0	4	B	B	A	A	B	B	A	B	B
t6k21ind15-15	0.1	4	B	B	U	A	B	B	U	B	A
ciw7	0.1	4	B	B	A	A	B	B	A	A	A
MS_At4_11.2	0.1	4	B	B	A	A	B	B	B	B	B
MS_At4_13.2	0.0	4	B	A	A	A	B	B	A	A	B
nga1139	0.7	4	A	B	A	B	A	B	U	A	A
MS_At4_16.1	0.3	4	A	B	A	B	A	B	A	A	B
t5j1748070	0.0	4	A	B	B	B	A	B	B	A	A
nga225	0.1	5	A	B	A	B	B	B	A	B	B
mhf15ind52-52	0.1	5	A	B	A	B	B	B	A	B	B
MS_At5_2.2	0.1	5	A	B	A	B	B	B	A	B	B
f14f1847646	0.1	5	A	B	A	B	B	B	U	B	U
nga151	0.2	5	A	B	A	B	B	B	U	B	B
MS_At5_5.3	0.1	5	A	B	A	B	B	A	A	B	B
MS_At5_6.6_a	0.1	5	A	B	A	B	B	A	A	A	B
nga139	0.0	5	A	B	A	B	B	B	A	B	A
jconnchr5_12.7	0.1	5	A	B	A	B	A	B	A	A	A
jconnchr5_13.6	0.0	5	A	B	A	U	A	B	A	A	A
aths0191	0.1	5	A	A	A	B	A	A	A	A	A
ciw9	1.7	5	B	A	A	B	A	A	U	A	A
mlnlind8-8	1.2	5	A	A	A	B	A	A	B	A	A
mql58836	1.7	5	A	A	A	A	A	A	A	A	A
nga129	2.2	5	A	A	A	A	A	A	U	A	A
MS_At5_20.0	1.8	5	A	A	A	A	A	A	A	A	B
jv65-66	3.2	5	A	A	A	A	A	A	A	A	A
mth12	1.9	5	B	A	A	A	A	A	U	A	A
mrg2138439	1.1	5	B	A	U	B	A	A	A	A	A
k8a1022396	1.3	5	B	A	B	B	A	A	A	A	A

TJ79 OIL CH2 AND CH5

;Ts-5 x Mz-0 recombinant F8 population genotype file;
;A = Ts-5,B = Mz-0,U = unclassified,
B is increasing allele
Oil content ch2 ch5

rep1	47	44	*	45	45	45	47	*	44
rep2	48	46	44	47	44	46	46	47	45
rep3	47	45	44	47	44	44	46	45	42
ave.	**47.20**	**45.12**	**43.99**	**46.51**	**44.25**	**44.86**	**46.65**	**45.57**	**43.82**
no. of A	60	58	55	51	50	47	47	46	40

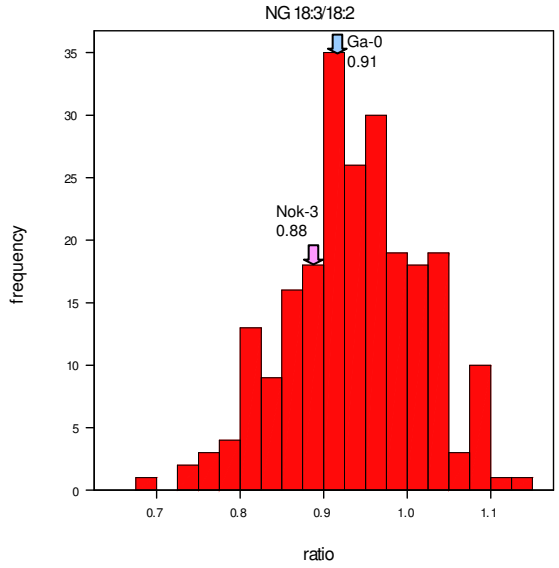
markers	LOD	LG	79	25	84	64	48	43	65	12	66
MS_At1_0.9	1.5	1	A	A	A	B	B	B	A	A	A
jconn1_2.7	0.6	1	A	U	A	B	B	A	A	A	A
jconn1_3.0	0.5	1	A	A	A	B	B	A	A	A	A
MS_At1_4.3	0.9	1	A	A	A	B	B	A	A	A	A
athsrp54a	1.7	1	A	A	A	B	B	A	A	A	B
f20d23ind18-18a	1.2	1	A	A	A	B	B	A	A	A	B
f9h16ind26-26	0.9	1	U	A	B	U	A	A	A	A	B
f8k728985	0.8	1	A	A	B	B	A	A	A	A	B
MS_At1_8.2	0.9	1	A	A	B	B	A	A	A	A	B
nga392	0.0	1	A	A	A	A	A	A	A	A	B
aths0392	0.2	1	A	A	A	A	A	A	A	A	B
MS_At1_11.2	0.6	1	A	A	A	A	A	A	A	A	B
MS_At1_12.9	0.2	1	A	B	A	A	A	A	A	A	B
t27k12-sp6	0.1	1	A	B	A	A	A	A	B	A	B
MS_At1_20b	0.8	1	A	B	A	A	B	A	B	A	B
MS_At1_20a	0.9	1	A	B	A	A	B	A	B	A	B
nga128	1.6	1	A	B	A	A	B	A	B	A	B
MS_At1_22.1	1.8	1	A	A	A	A	B	A	B	A	B
MS_At1_22.9	1.7	1	A	A	A	A	B	A	B	A	B
MS_At1_23.4	1.7	1	A	A	A	A	B	A	B	A	B
f5i1449495	1.7	1	A	A	A	A	A	A	B	A	B
MS_At1_24.3	1.6	1	A	A	A	A	A	A	B	A	B
f5a1859436	2.6	1	A	A	A	A	A	A	A	A	A
nga111	3.5	1	A	A	A	A	A	A	A	A	A
jconn2_0.7	0.4	2	A	B	B	A	B	A	B	A	A
MS_At2_3.0	0.2	2	A	B	B	A	B	A	A	A	A
MS_At2_5.4	0.2	2	A	B	B	A	B	A	A	A	A
MS_At2_6.3	0.7	2	A	B	U	A	B	U	U	A	A
jconn2_3.0		2	A	B	B	A	B	A	A	A	A
jconn2_5.2		2	A	B	B	A	B	A	A	A	A
ciw3	0.8	2	U	B	B	A	B	B	A	A	A
MS_At2_9.3	3.5	2	B	B	B	B	B	B	B	B	B
pls8	3.9	2	B	B	B	B	B	B	B	B	B
nga1126	3.8	2	B	B	B	B	B	B	B	B	B
c4h	1.8	2	B	A	B	B	B	B	B	B	B
t6a23ind10-10	1.2	2	A	A	B	B	B	B	B	A	B
nga168	1.3	2	A	A	B	B	B	B	B	A	B
MS_At2_17.5	1.0	2	B	A	B	B	B	B	B	A	B
MS_At2_17.9	1.3	2	B	A	B	U	A	U	B	A	U
athubique	1.5	2	B	A	A	B	A	B	B	A	B
nga32	0.9	3	U	B	A	B	A	A	B	A	A
nga162	0.8	3	B	B	B	A	A	A	B	U	A
MS_At3_6.1	0.5	3	A	B	U	A	A	A	B	A	B
msd2129380	0.2	3	A	B	B	A	A	A	B	A	A
mzn14inda9-a9	0.1	3	A	A	A	A	A	U	A	A	A
t13j1016610	0.3	3	A	A	A	A	A	B	A	A	U
k11j14ind16-16	0.3	3	A	A	A	A	A	B	A	A	A
MS_At3_17.3	1.7	3	A	A	A	A	U	B	A	B	A
t16k521877	1.0	3	A	A	A	A	A	B	B	A	A
MS_At3_19.7	0.8	3	A	A	A	A	A	B	B	A	A
athcdc2bg	0.6	3	A	A	A	A	A	B	B	A	A
nga6	0.5	3	A	A	A	A	A	B	B	A	A
MS_At4_0.3	0.4	4	B	A	B	A	A	B	U	B	B
MS_At4_1.1	0.1	4	B	A	B	A	A	B	U	B	B
t19j18ind30-30	0.8	4	B	A	B	A	A	B	U	B	B
nga8		4	B	A	B	A	A	B	U	B	B
det1.2	0.2	4	B	A	B	A	A	B	A	B	B
f28m11ind22-22		4	B	A	B	A	A	B	A	B	B
ciw6	0.0	4	A	A	B	A	A	B	A	B	A
t6k21ind15-15	0.1	4	A	A	B	A	A	B	A	B	A
ciw7	0.6	4	A	A	B	A	A	B	A	B	B
nga1139	0.1	4	A	B	B	A	A	B	A	B	U
MS_At4_16.2	0.2	4	A	B	B	A	A	B	A	U	B
t5j1748070	0.0	4	A	B	B	A	B	B	B	B	B
nga225	0.1	5	A	A	A	B	B	A	B	B	B
MS_At5_2.2	0.8	5	A	A	A	B	B	A	B	B	A
MS_At5_2.8	0.8	5	A	A	A	B	B	A	B	B	A
MS_At5_3.6	1.0	5	A	A	A	B	B	B	B	B	A
f14f1847646	0.6	5	A	A	A	B	B	B	B	B	A
nga151	0.2	5	U	A	A	B	B	B	B	B	A
MS_At5_6.6a	1.0	5	A	A	A	A	B	B	B	B	A
MS_At5_6.6b	1.0	5	A	A	A	A	B	B	B	B	A
MS_At5_8.6	0.7	5	B	B	A	A	A	B	A	B	A
nga139	0.3	5	B	B	A	A	A	B	A	B	A
MS_At5_9.7	0.1	5	A	B	A	A	A	A	A	B	A
mlf18ind14-14	0.1	5	A	B	A	B	A	A	A	B	A
MS_At5_15.8	1.9	5	A	B	A	B	A	A	A	B	A
ciw9	2.0	5	B	B	A	B	A	A	A	B	A
MS_At5_18	1.2	5	B	A	A	B	A	A	A	B	B
f20d23ind18-18b	1.2	5	B	A	A	B	A	A	A	B	B
nga129	1.0	5	B	A	A	B	A	A	A	B	B
MS_At5_20.4	0.8	5	B	A	A	B	A	A	A	U	B
MS_At5_21.3	1.1	5	B	A	A	B	A	A	A	B	B
jv65-66	2.2	5	U	A	A	B	A	A	A	B	B
MS_At5_22.8	2.4	5	B	A	A	B	A	A	A	B	B
mth12	1.7	5	A	A	A	B	A	A	A	B	B
MS_At5_26.1	1.4	5	A	A	A	B	B	U	A	B	A
k8a1022396	0.6	5	A	B	A	A	B	U	A	B	A

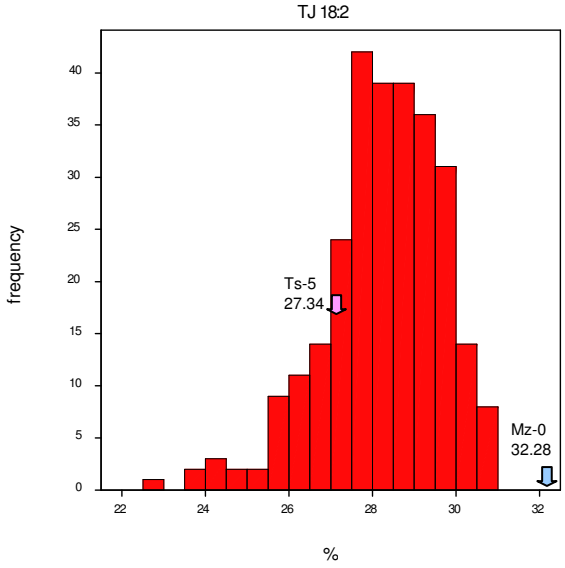
;Ts-5 x Mz-0 recombinant F8 population genotype file;
;A = Ts-5,B = Mz-0,U = unclassified,
B is increasing allele
Oil content

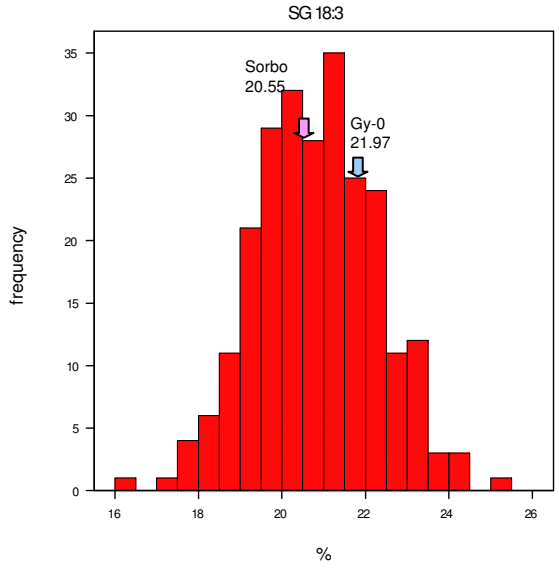
rep1	*	44	53	47	38	45	44	46	47	44	45
rep2	43	47	46	45	46	44	46	44	45	45	45
rep3	46	46	46	45	45	46	43	44	47	45	45
ave.	**44.63**	**45.52**	**47.90**	**45.79**	**43.06**	**45.05**	**44.34**	**44.86**	**46.17**	**44.49**	**45.17**
no. of A	56	53	52	51	50	48	43	42	42	42	40

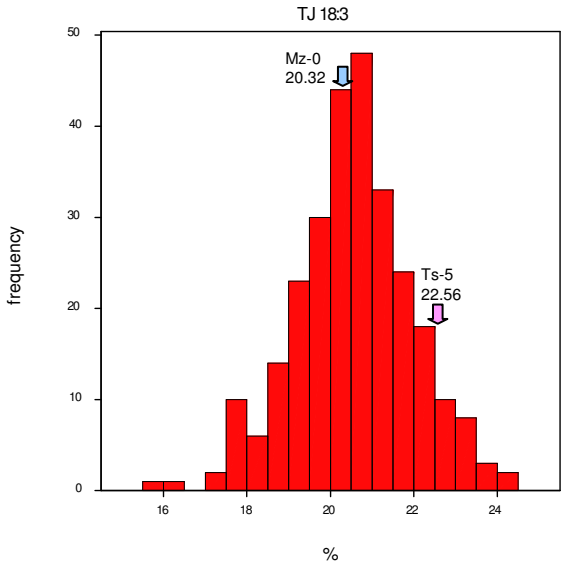
markers	LOD	LG	32	22	76	83	70	24	10	9	41	89	28
MS_At1_0.9	1.5	1	B	B	A	A	B	B	B	A	B	B	U
jconn1_2.7	0.6	1	U	B	A	A	B	B	U	A	B	B	B
jconn1_3.0	0.5	1	B	B	A	A	B	B	B	A	B	B	B
MS_At1_4.3	0.9	1	B	B	A	A	B	B	B	A	B	B	A
athsrp54a	1.7	1	B	B	A	A	B	B	B	A	B	B	A
f20d23ind18-18a	1.2	1	B	B	A	A	B	U	B	A	B	B	A
f9h16ind26-26	0.9	1	B	B	A	U	B	B	B	A	B	B	A
f8k728985	0.8	1	B	B	A	A	B	B	U	B	B	B	A
MS_At1_8.2	0.9	1	B	B	A	A	B	B	B	B	B	B	A
nga392	0.0	1	A	B	A	U	A	B	B	B	A	A	A
aths0392	0.2	1	A	A	A	A	A	B	B	B	A	A	A
MS_At1_11.2	0.6	1	A	A	A	A	B	B	B	B	A	A	A
MS_At1_12.9	0.2	1	A	A	A	B	B	B	A	B	A	B	A
t27k12-sp6	0.1	1	U	U	A	B	B	B	A	U	A	B	A
MS_At1_20b	0.8	1	A	B	B	B	A	A	A	A	A	B	B
MS_At1_20a	0.9	1	A	B	B	B	A	A	A	A	A	B	B
nga128	1.6	1	A	B	B	B	A	A	A	A	A	B	B
MS_At1_22.1	1.8	1	A	B	B	B	A	A	A	A	A	B	B
MS_At1_22.9	1.7	1	A	B	B	B	A	A	A	U	A	B	B
MS_At1_23.4	1.7	1	A	B	B	B	A	A	A	A	A	B	B
f5i1449495	1.7	1	B	B	B	B	U	A	A	U	A	B	B
MS_At1_24.3	1.6	1	B	B	B	B	U	A	A	A	A	B	B
f5a1859436	2.6	1	B	A	B	B	B	B	B	A	B	B	B
nga111	3.5	1	B	B	B	B	B	B	B	B	B	B	B
jconn2_0.7	0.4	2	A	A	B	A	A	A	B	B	A	A	A
MS_At2_3.0	0.2	2	A	A	B	A	A	A	B	B	A	A	A
MS_At2_5.4	0.2	2	A	A	B	A	A	A	B	U	A	A	A
MS_At2_6.3	0.7	2	A	A	B	U	A	A	B	A	A	U	A
~~jconn2_3.9~~		2	A	A	B	A	A	A	B	B	A	A	A
~~jconn2_5.2~~		2	A	A	B	A	A	A	B	B	A	A	A
ciw3	0.8	2	A	A	B	A	A	A	B	A	A	A	A
MS_At2_9.3	3.5	2	A	A	A	A	A	A	A	A	A	A	A
pls8	3.9	2	A	A	A	A	A	A	A	A	A	A	A
nga1126	3.8	2	B	A	A	A	A	A	A	A	A	A	B
c4h	1.8	2	B	A	A	A	A	A	A	A	A	A	B
t6a23ind10-10	1.2	2	B	A	B	B	A	A	B	A	A	A	B
nga168	1.3	2	B	A	B	B	A	A	U	A	A	A	B
MS_At2_17.5	1.0	2	B	A	B	B	A	A	B	A	A	A	B
MS_At2_17.9	1.3	2	U	A	B	B	A	A	B	A	A	A	U
athubique	1.5	2	B	A	B	B	A	A	B	A	A	A	A
nga32	0.9	3	B	A	A	A	A	B	B	A	B	B	B
nga162	0.8	3	A	A	A	B	A	A	U	B	B	B	B
MS_At3_6.1	0.5	3	A	U	B	B	A	A	A	B	B	B	B
msd2129380	0.2	3	A	B	B	A	A	A	A	B	B	B	B
mzn14inda9-a9	0.1	3	A	B	B	B	A	A	A	U	B	A	B
t13j1016610	0.3	3	U	U	B	B	A	A	A	U	B	U	B
k11j14ind16-16	0.3	3	A	B	B	B	A	A	A	U	B	A	B
MS_At3_17.3	1.7	3	U	B	B	A	A	A	A	B	A	A	B
t16k521877	1.0	3	A	A	A	A	A	A	A	B	A	A	B
MS_At3_19.7	0.8	3	A	A	A	A	A	A	A	B	A	A	B
athcdc2bg	0.6	3	A	A	A	A	A	U	A	B	A	A	B
nga6	0.5	3	A	A	A	A	A	A	A	B	A	A	B
MS_At4_0.3	0.4	4	A	B	A	A	A	A	A	A	B	B	A
MS_At4_1.1	0.1	4	A	A	A	A	A	A	A	A	B	B	A
t19j18ind30-30	0.8	4	A	A	A	A	B	A	B	A	B	B	A
~~nga8~~		4	A	B	A	A	A	A	U	A	B	B	A
det1.2	0.2	4	B	B	B	A	A	B	U	A	A	B	A
~~f28m11ind22-22~~		4	B	B	B	A	A	B	A	A	A	B	A
ciw6	0.0	4	A	B	B	A	A	B	A	B	A	A	A
t6k21ind15-15	0.1	4	A	B	B	A	A	B	A	B	A	A	A
ciw7	0.6	4	A	A	B	A	A	B	A	B	A	B	A
nga1139	0.1	4	A	A	B	B	A	B	U	B	A	B	A
MS_At4_16.2	0.2	4	A	A	B	B	A	B	A	B	A	B	A
t5j1748070	0.0	4	A	A	B	B	A	A	A	B	A	A	A
nga225	0.1	5	B	B	A	A	B	B	B	B	B	B	B
MS_At5_2.2	0.8	5	B	B	A	A	B	B	B	B	B	B	B
MS_At5_2.8	0.8	5	B	B	A	A	B	B	B	B	B	B	B
MS_At5_3.6	1.0	5	B	A	A	A	B	B	U	B	B	B	B
f14f1847646	0.6	5	B	A	A	A	B	B	A	B	B	B	B
nga151	0.2	5	B	A	A	A	B	B	U	B	B	B	B
MS_At5_6.6a	1.0	5	A	A	A	A	B	B	B	B	B	B	B
MS_At5_6.6b	1.0	5	A	A	A	A	B	B	B	B	B	B	B
MS_At5_8.6	0.7	5	A	A	A	B	B	A	B	U	B	B	B
nga139	0.3	5	A	A	A	B	B	A	B	U	B	B	B
MS_At5_9.7	0.1	5	A	A	A	B	B	A	B	A	B	A	B
mlf18ind14-14	0.1	5	A	A	A	B	A	U	B	B	B	A	B
MS_At5_15.8	1.9	5	A	A	A	A	A	A	B	B	B	A	B
ciw9	2.0	5	A	A	A	A	A	A	U	A	B	A	B
MS_At5_18	1.2	5	A	A	A	B	B	A	A	A	B	A	B
f20d23ind18-18b	1.2	5	A	A	A	B	B	U	A	A	B	A	B
nga129	1.0	5	A	A	A	B	B	B	A	A	B	A	A
MS_At5_20.4	0.8	5	A	A	A	B	B	B	A	A	B	A	A
MS_At5_21.3	1.1	5	A	A	A	B	B	B	A	A	B	A	A
jv65-66	2.2	5	A	A	A	A	B	B	A	A	B	A	A
MS_At5_22.8	2.4	5	A	A	A	A	B	B	A	A	B	A	U
mth12	1.7	5	A	A	A	A	B	A	A	U	B	A	A
MS_At5_26.1	1.4	5	A	A	A	A	B	A	A	A	B	A	A
k8a1022396	0.6	5	A	A	A	A	B	A	A	B	B	A	A

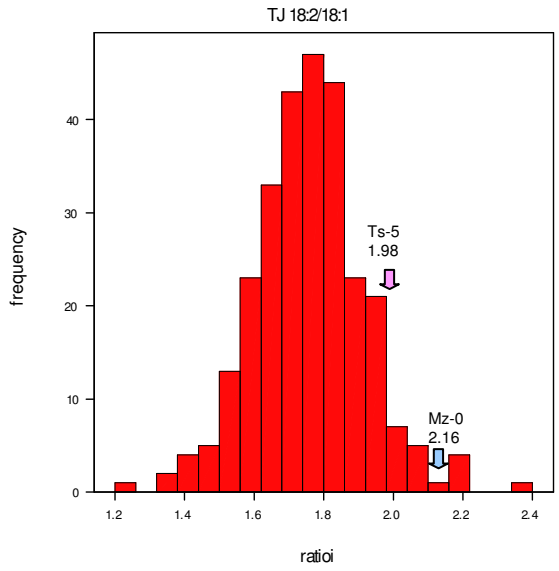
APPENDIX 4 HISTOGRAM OF RIL POPULATION (PHENOTYPE DATA WERE PROVIDED BY CARMEL O'NEILL MEASURED IN 2008)

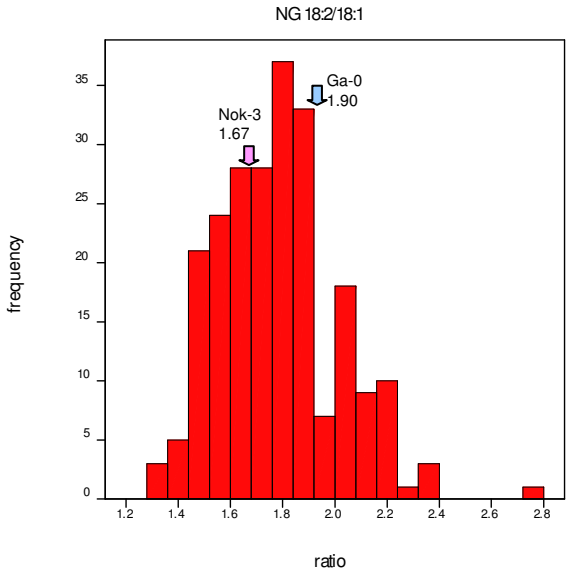


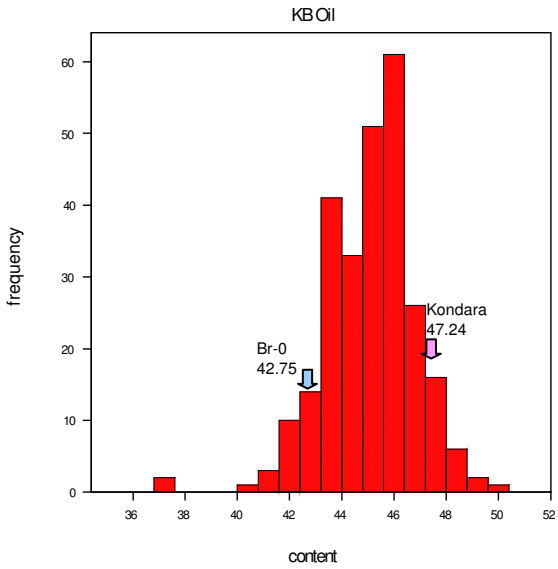


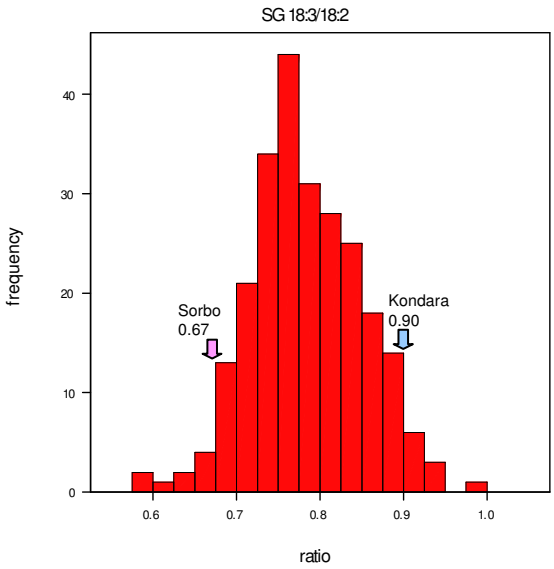


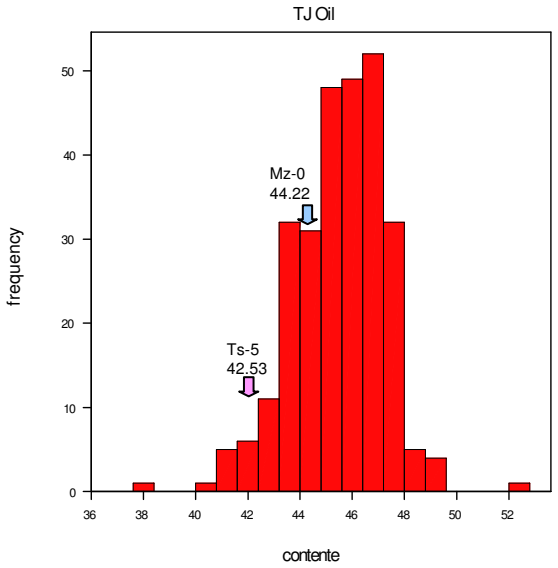







KB Oil
frequency
content
Br-0
42.75
Kondara
47.24


SG 18:3/18:2
frequency
ratio
Sorbo
0.67
Kondara
0.90


TJ Oil
frequency
contente
Mz-0
44.22
Ts-5
42.53

APPENDIX 5 GENOTYPE FILE OF NIL CANDIDATES AND MH PLANTS AND PHENOTYPE DATA

TJ25 18:2

TJ25 18:2

B is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ25	7	9	10	21	35	39	2	15	19	26	44	45	Ts-5	Mz-0
MS_At1_0.9	0.43	1	0.9	0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconn1_2.7	1.05	1		3															
jconn1_3.0	1.28	1		4	A														
MS_At1_4.3	1.38	1	4.3	9	A														
athsrp54a	1.31	1	5.2	13	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f20d23ind18-18a	1.44	1	5.5	14	A														
f9h16ind26-26	1.54	1	7.0	18	A														
f8k728985	2.53	1	7.6	19	A														
MS_At1_8.2	4.47	1	8.2	24	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga392	2.98	1	9.8	31	A														
aths0392	2.37	1	10.9	36	A														
MS_At1_11.2	1.89	1	11.2	37	A														
MS_At1_12.9	1.06	1	12.9	45	B	B	H	A	H	A	H	H	H	H	A	H	B	A	B
t27k12-sp6	0.76	1	15.5	48	B														
MS_At1_20b	0.36	1	20.0	61	B														
MS_At1_20a	0.37	1	20.0	61	B														
nga128	0.09	1	20.2	64	B	B	H	A	H	H	H	B	B	H	A	H	B	A	B
MS_At1_22.1	0.21	1	22.1	66	A														
MS_At1_22.9	0.21	1	22.9	66	A														
MS_At1_23.4	0.19	1	23.4	66	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f5i1449495	0.2	1	24.0	68	A														
MS_At1_24.3	0.2	1	24.3	68	A														
f5a1859436	0.23	1	26.2	83	A														
nga111	0.33	1	27.0	88	A														
jconn2_0.7	0.5	2		0	B														
MS_At2_3.0	0.06	2	3.0	10	B	A	A	H	B	H	H	A	A	B	H	A	H	A	B
MS_At2_5.4	0.06	2	5.4	10	B														
MS_At2_6.3	0.03	2	6.3	12	B														
~~jconn2_3.0~~		2			B														
~~jconn2_5.2~~		2			B														
ciw3	0.03	2	6.4	13	B	A	A	H	B	_	H	A	A	B	H	A	H	A	B
MS_At2_9.3	0.96	2	9.3	29	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	0.49	2	9.9	31	B														
nga1126	0.26	2	11.6	35	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
c4h	0.32	2	12.9	45	A														
t6a23ind10-10	0.22	2	16.1	51	A														
nga168	0.21	2	16.2	51	A														
MS_At2_17.5	0.78	2	17.5	54	A														
MS_At2_17.9	0.72	2	17.9	55	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athubique	0.34	2	19.3	60	A														
nga32	0.84	3	0.3	0	B	B	B	B	B	B	H	A	A	H	H	H	H	A	B
nga162	2.06	3	4.6	9	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
MS_At3_6.1	4.34	3	6.1	15	B														
msd2129380	4.8	3	7.7	20	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
mzn14inda9-a9	0.14	3	10.7	56	A														
t13j1016610	0.2	3	11.5	64	A														
k11j14ind16-16	0.27	3	13.4	64	A														
MS_At3_17.3	0.12	3	17.3	77	A														
t16k521877	0.08	3	18.9	82	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At3_19.7	0.41	3	19.7	87	A														
athcdc2bg	0.1	3	20.5	91	A														
nga6	0.22	3	23.0	95	A														
MS_At4_0.3	0.1	4	0.3	0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_1.1	1.29	4	1.1	7	A														
t19j18ind30-30	0.32	4	2.3	11	A														
~~nga8~~		4			A														
det1.2	0.8	4	5.3	19	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
~~f28m11ind22-22~~		4			A														
ciw6	0.93	4	7.0	28	A														
t6k21ind15-15	0.46	4	9.9	33	A														
ciw7	0.24	4	10.6	39	A														
nga1139	0.07	4	15.4	49	B														
MS_At4_16.2	0.29	4	16.2	51	B	B	H	H	B	H	H	H	_	H	A	H	A	A	B
t5j1748070	0.8	4	17.7	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga225	3.5	5	1.4	0	A														
MS_At5_2.2	2.53	5	2.2	3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_2.8	1.94	5	2.8	4	A														
MS_At5_3.6	0.87	5	3.6	6	A														
f14f1847646	0.56	5	3.8	6	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga151	0.33	5	4.7	7	A														
MS_At5_6.6a	0.18	5	6.6	14	A														
MS_At5_6.6b	0.18	5	6.6	14	A														
MS_At5_8.6	0	5	8.6	19	B	H	H	B	A	H	H	H	H	H	A	A	A	A	B
nga139	0.03	5	8.6	20	B														
MS_At5_9.7	0.6	5	9.7	28	B	H	H	H	A	H	H	B	H	H	A	A	A	A	B
mlf18ind14-14	0.08	5	14.6	40	B														
MS_At5_15.8	0.09	5	15.8	50	B														
ciw9	0.02	5	17.2	57	B	A	H	H	B	H	B	H	H	H	H	A	A	A	B
MS_At5_18	0	5	18.0	59	A														
~~f20d23ind18-18b~~	0.01	5		65	A														
nga129	0.03	5	19.7	67	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_20.4	0.03	5	20.4	67	A														
MS_At5_21.3	0.03	5	21.3	70	A														
jv65-66	0.03	5	22.1	75	A														
MS_At5_22.8	0	5	22.8	78	A														
mth12	0.13	5	24.1	85	A														
MS_At5_26.1	0.31	5	26.1	93	A														
k8a1022396	0.17	5	26.9	98	B	H	A	H	B	B	B	H	H	B	H	H	A	A	B

Fig. A5-1 Genotype file of NIL candidates and MH plants of TJ25 for 18:2

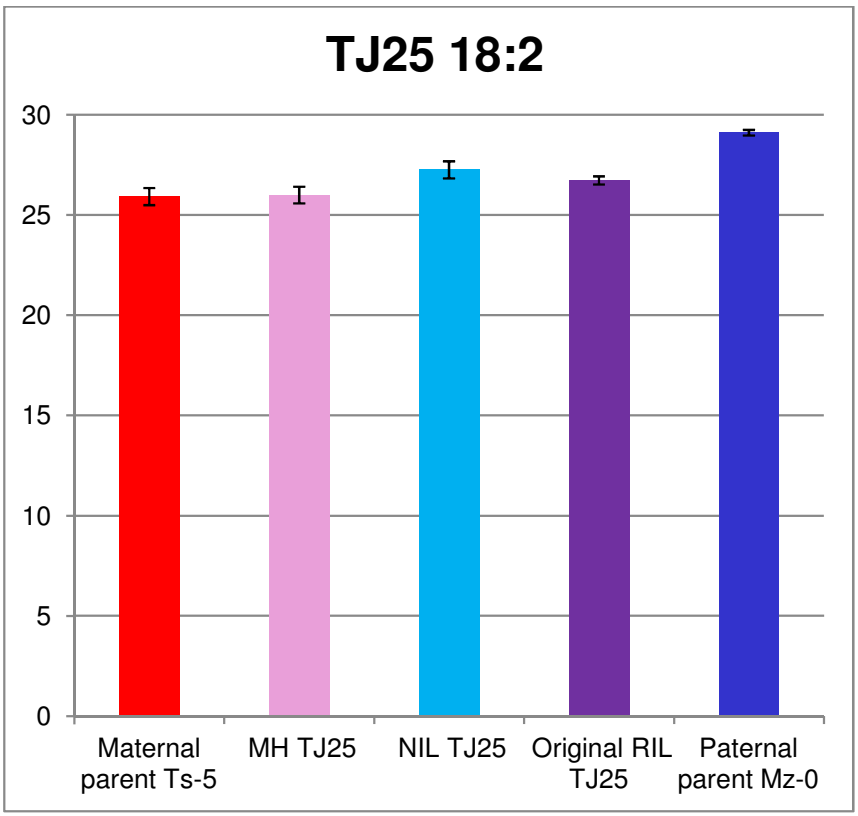


Fig. A5-2 Average of 18:2 contents of NIL candidates and MH plants of TJ25

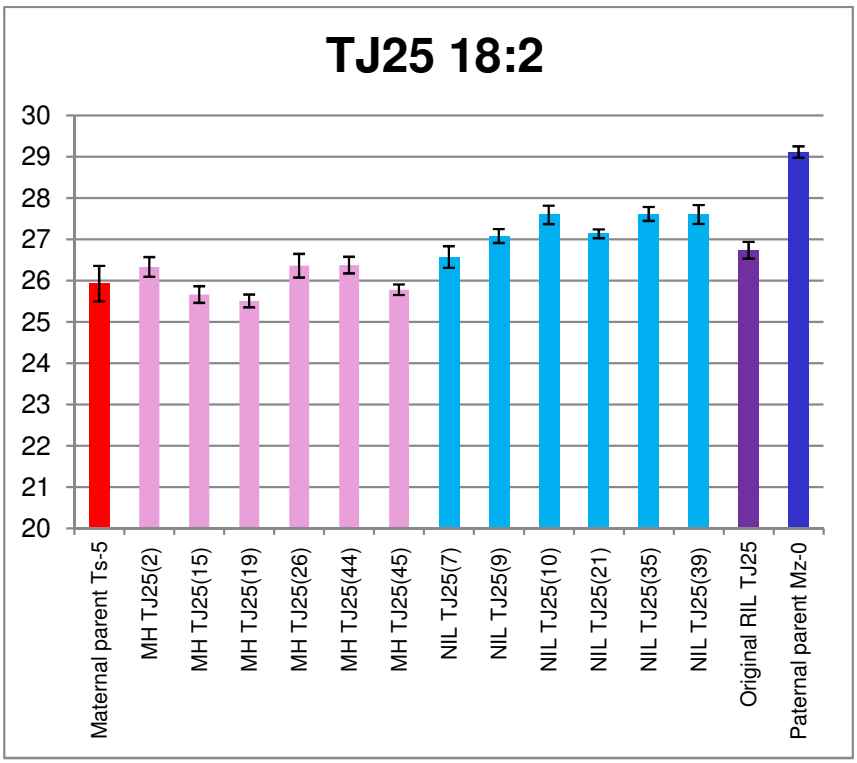


Fig. A5-3 Average of18;2 conetns of individual NIL candidates and MH plants of TJ25

NG28 18:2/18:1

B is increasing allele					RIL	NIL candidates					MH					Parents	
Markers	**LOD**	**LG**	**Mb**	**cM**	**NG28**	**5**	**16**	**19**	**27**	**32**	**4**	**9**	**14**	**30**	**44**	**Nok-3**	**Ga-0**
nga59	0.07	1	0.0	0	B												
MS_At1_0.6	0.39	1	0.6	4	B	-	A	A	A	A	A	A	A	A	A	A	B
MS_At1_2.7	0.15	1	2.7	10	B												
nga63	0	1	3.2	12	B	A	A	A	A	A	A	A	A	A	A	A	B
f16j7-trb	0	1	3.8	15	B												
athsrp54a	0.12	1	5.2	18	B												
MS_At1_6.4	0.04	1	6.4	21	U	A	A	A	A	A	A	A	A	A	A	A	B
f9h16ind26-26	0.01	1	7.0	27	B												
MS_At1_7.4	0.05	1	7.4	30	B												
MS_At1_8.2	0.71	1	8.2	38	B	-	A	A	A	A	A	A	A	A	A	A	B
nga392	0.42	1	9.8	43	B												
aths0392	0.52	1	10.9	48	B												
MS_At1_11.2	1.48	1	11.2	50	B	A	A	A	A	A	-	A	A	A	A	A	B
f140426482	1.57	1	13.2	58	A												
t27k12-sp6	2.07	1	15.5	59	A	A	A	A	A	A	A	A	A	A	A	A	B
~~MS_At1_16.4~~		1	~~16.4~~		A												
MS_At1_20_a	0.05	1	20.0	66	A												
nga128	0.36	1	20.2	69	A												
MS_At1_22.6	0.46	1	22.6	71	A	A	A	A	A	A	A	A	A	A	A	A	B
t12p18ind8-8	0	1	23.5	81	A												
f5i1449495	0.05	1	24.0	84	A												
f5a1859436	0.09	1	26.2	95	A												
nga111	0.08	1	27.0	98	A												
yup8h12r	0.34	1	29.1	106	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1145	0	2	0.7	0	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw2	0	2	1.2	2	A												
t12j241450	0.06	2	3.6	8	A												
ciw3	0.04	2	6.4	15	A												
pls1	0.21	2	8.2	27	A												
pls8	0.25	2	9.9	36	B	H	B	B	H	A	H	H	H	H	B	A	B
athgpa1	0	2	11.1	44	A												
nga1126	0.03	2	11.6	45	U	A	A	A	A	A	-	-	-	-	-	-	-
t32f646516	0.24	2	13.8	55	A												
MS_At2_14.2_b	0.21	2	14.2	56	A												
nga168	0.35	2	16.2	65	U	H	H	H	H	H	B	H	A	H	B	A	B
90j19t7	0.14	2	19.6	75	U												
nga32	1.26	3	0.3	0	A	A	A	A	A	A	A	A	A	A	A	-	B
f17a931902	0.8	3	2.1	8	B	B	H	H	H	B	A	H	A	B	B	A	B
nga162	0.41	3	4.6	14	B												
msd2129380	0.9	3	7.7	25	A												
MS_At3_8.3_b	0.84	3	8.3	30	A												
athgapab	1.51	3	9.8	38	A	A	A	A	A	A	A	A	A	A	A	A	B
t13j1016610	1.41	3	11.5	46	A												
k11j14ind16-16	1.27	3	13.4	48	A												
MS_At3_18.3	1.88	3	18.3	55	A												
t16k521877	2.11	3	18.9	58	A												
athcdc2bg	1.86	3	20.5	72	A	A	A	A	A	A	A	A	A	A	A	A	B
nga6	1.57	3	23.0	84	A												
jv30/31	0.1	4	0.0	0	U	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_0.3	0.11	4	0.3	1	B	A	A	A	A	A	-	A	A	A	A	A	B
t19j18ind30-30	0	4	2.3	6	B												
nga8	0.08	4	4.6	11	B												
MS_At4_5.3	0.21	4	5.3	13	B	-	-	-	-	-	-	-	-	-	-	-	-
ciw6	1.51	4	7.0	29	A												
MS_At4_10.8	0.9	4	10.8	44	A	A	A	A	A	A	A	A	A	A	A	A	B
t22a6ind10-10	1.35	4	12.6	53	A												
f27g1959898	1.39	4	12.8	57	A												
~~t27e11ind12-12~~		4			A												
nga1139	1.87	4	15.4	68	A												
nga1107	1.33	4	17.0	74	A												
t5j1748070	0.55	4	17.7	77	B	A	A	A	A	A	A	A	A	A	A	A	B
nga225	7.64	5	1.4	0	A												
mhf15ind52-52	6.88	5	1.7	4	A												
ciw14	6.73	5	2.1	5	A	A	A	A	A	A	A	A	A	A	A	A	B
f14f1847646	11.82	5	3.8	9	B	B	B	B	B	B	A	A	A	A	A	A	B
nga151	11.17	5	4.7	12	B												
nga139	4.35	5	8.6	34	B	B	B	B	B	B	A	A	A	A	A	A	B
MS_At5_9.7	0.28	5	9.7	45	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_13.9	0.08	5	13.9	53	A												
mlf18ind14-14	0.01	5	14.6	56	A												
aths0191	0	5	15.1	60	A												
MS_At5_17.8	0.63	5	17.8	80	A	A	A	A	A	A	A	A	A	A	A	A	B
mq158836	0.56	5	19.3	84	A												
MS_At5_20.4	0.63	5	20.4	89	A												
jv65-66	0.49	5	22.1	97	B	H	B	B	B	H	A	A	H	B	H	A	B
mrg2138439	0.39	5	25.2	114	B	H	H	A	H	H	H	H	H	B	H	A	B
k8a1022396	0.31	5	26.9	122	A												

Fig. A5-4 Genotype file of NIL candidates and MH plants of NG28 for 18:2/18:1

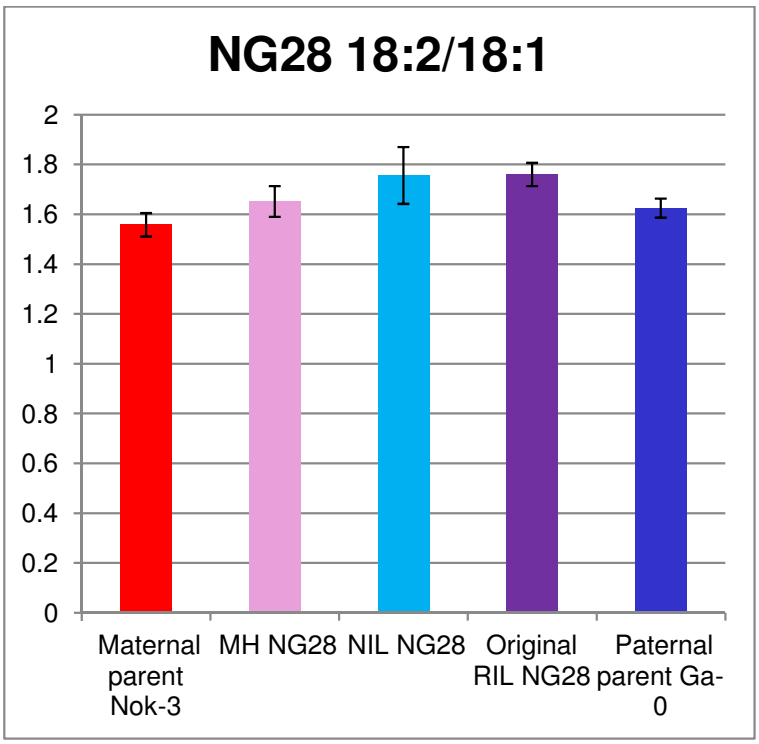


Fig. A5-5 Average of 18;2/18:1 ratio of NIL candidates and MH plants of NG28

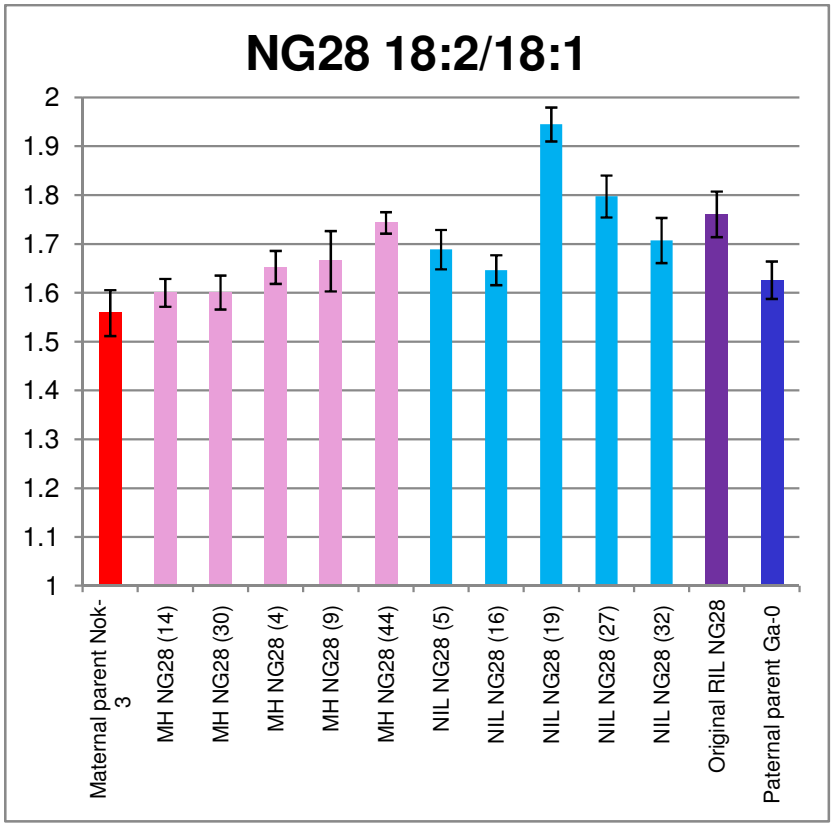


Fig. A5-6 Average of 18;2/18:1 ratio of individual NIL candidates and MH plants of NG28

NG32 18:2/18:1

A is increasing allele					RIL	NIL candidates						MH					Parents	
Markers	LOD	LG	Mb	cM	NG32	23	27	31	33	34	40	1	25	32	39	42	Nok-3	Ga-0
nga59	0.07	1	0.0	0	B	B	B	A	A	B	B	B	B	A	B	B	A	B
MS_At1_0.6	0.39	1	0.6	4	B													
MS_At1_2.7	0.15	1	2.7	10	B	H	A	A	A	H	B	H	H	-	A	B	A	B
nga63	0	1	3.2	12	B													
f16j7-trb	0	1	3.8	15	B	H	A	A	A	H	B	H	H	A	A	B	A	B
athsrp54a	0.12	1	5.2	18	A													
MS_At1_6.4	0.04	1	6.4	21	A													
f9h16ind26-26	0.01	1	7.0	27	A													
MS_At1_7.4	0.05	1	7.4	30	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_8.2	0.71	1	8.2	38	A													
nga392	0.42	1	9.8	43	A													
aths0392	0.52	1	10.9	48	A													
MS_At1_11.2	1.48	1	11.2	50	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f140426482	1.57	1	13.2	58	A													
t27k12-sp6	2.07	1	15.5	59	A													
~~MS_At1_16.4~~		1	~~16.4~~		A													
MS_At1_20_a	0.05	1	20.0	66	A													
nga128	0.36	1	20.2	69	A													
MS_At1_22.6	0.46	1	22.6	71	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t12p18ind8-8	0	1	23.5	81	A													
f5i1449495	0.05	1	24.0	84	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f5a1859436	0.09	1	26.2	95	A													
nga111	0.08	1	27.0	98	A													
yup8h12r	0.34	1	29.1	106	A	A	H	A	H	B	B	H	H	A	H	B	A	B
nga1145	0	2	0.7	0	B	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw2	0	2	1.2	2	B													
t12j241450	0.06	2	3.6	8	B	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw3	0.04	2	6.4	15	B													
pls1	0.21	2	8.2	27	B	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	0.25	2	9.9	36	A													
athgpa1	0	2	11.1	44	A													
nga1126	0.03	2	11.6	45	A													
t32f646516	0.24	2	13.8	55	A													
MS_At2_14.2_b	0.21	2	14.2	56	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga168	0.35	2	16.2	65	A													
90j19t7	0.14	2	19.6	75	A													
nga32	1.26	3	0.3	0	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f17a931902	0.8	3	2.1	8	A													
nga162	0.41	3	4.6	14	A													
msd2129380	0.9	3	7.7	25	U	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At3_8.3_b	0.84	3	8.3	30	A													
athgapab	1.51	3	9.8	38	A													
t13j1016610	1.41	3	11.5	46	A	A	A	A	A	A	A	A	A	A	A	A	A	B
k11j14ind16-16	1.27	3	13.4	48	A													
MS_At3_18.3	1.88	3	18.3	55	A													
t16k521877	2.11	3	18.9	58	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athcdc2bg	1.86	3	20.5	72	A													
nga6	1.57	3	23.0	84	A													
jv30/31	0.1	4	0.0	0	A													
MS_At4_0.3	0.11	4	0.3	1	A	A	A	A	A	A	A	-	A	A	A	A	A	B
t19j18ind30-30	0	4	2.3	6	A													
nga8	0.08	4	4.6	11	A													
MS_At4_5.3	0.21	4	5.3	13	A													
ciw6	1.51	4	7.0	29	B	A	A	A	A	A	A	-	A	A	A	A	A	B
MS_At4_10.8	0.9	4	10.8	44	A													
t22a6ind10-10	1.35	4	12.6	53	A													
f27g1959898	1.39	4	12.8	57	A	A	A	A	A	A	A	A	A	A	A	A	A	B
~~t27e11ind12-12~~		4			A													
nga1139	1.87	4	15.4	68	A													
nga1107	1.33	4	17.0	74	B	A	A	H	H	A	A	H	H	A	A	A	A	B
t5j1748070	0.55	4	17.7	77	B													
nga225	7.64	5	1.4	0	B													
mhf15ind52-52	6.88	5	1.7	4	B	B	B	H	B	B	B	A	A	A	A	A	A	B
ciw14	6.73	5	2.1	5	B													
f14f1847646	11.82	5	3.8	9	B	B	B	H	B	B	B	A	A	A	A	A	A	B
nga151	11.17	5	4.7	12	B	B	B	B	B	B	B	A	A	A	A	A	A	B
nga139	4.35	5	8.6	34	B													
MS_At5_9.7	0.28	5	9.7	45	B	B	B	B	B	B	B	A	A	A	A	A	A	B
MS_At5_13.9	0.08	5	13.9	53	B	B	B	B	B	B	B	A	A	-	A	A	A	B
mlf18ind14-14	0.01	5	14.6	56	B	B	B	B	B	B	B	H	A	H	A	A	A	B
aths0191	0	5	15.1	60	A													
MS_At5_17.8	0.63	5	17.8	80	A													
mg158836	0.56	5	19.3	84	U	A	A	A	A	A	A	A	A	A	-	A	A	B
MS_At5_20.4	0.63	5	20.4	89	A													
jv65-66	0.49	5	22.1	97	A													
mrg2138439	0.39	5	25.2	114	B													
k8a1022396	0.31	5	26.9	122	B	H	H	B	A	B	H	B	A	H	H	H	A	B

Fig. A5-7 Genotype file of NIL candidates and MH plants of NG32 for 18:2/18:1

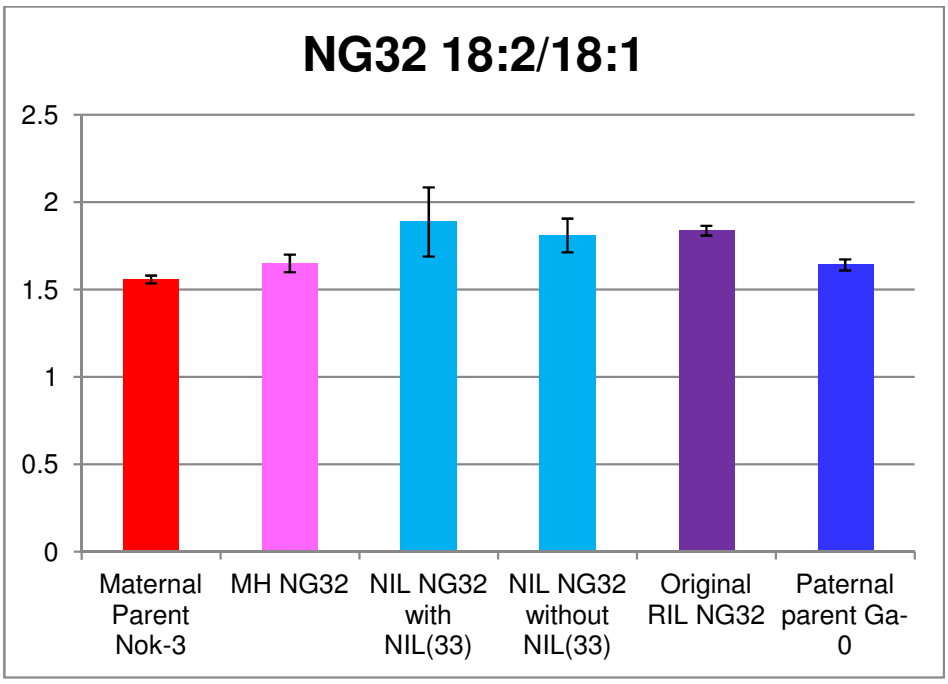


Fig. A5-8 Average of 18;2/18:1 ratio of NIL candidates and MH plants of NG32

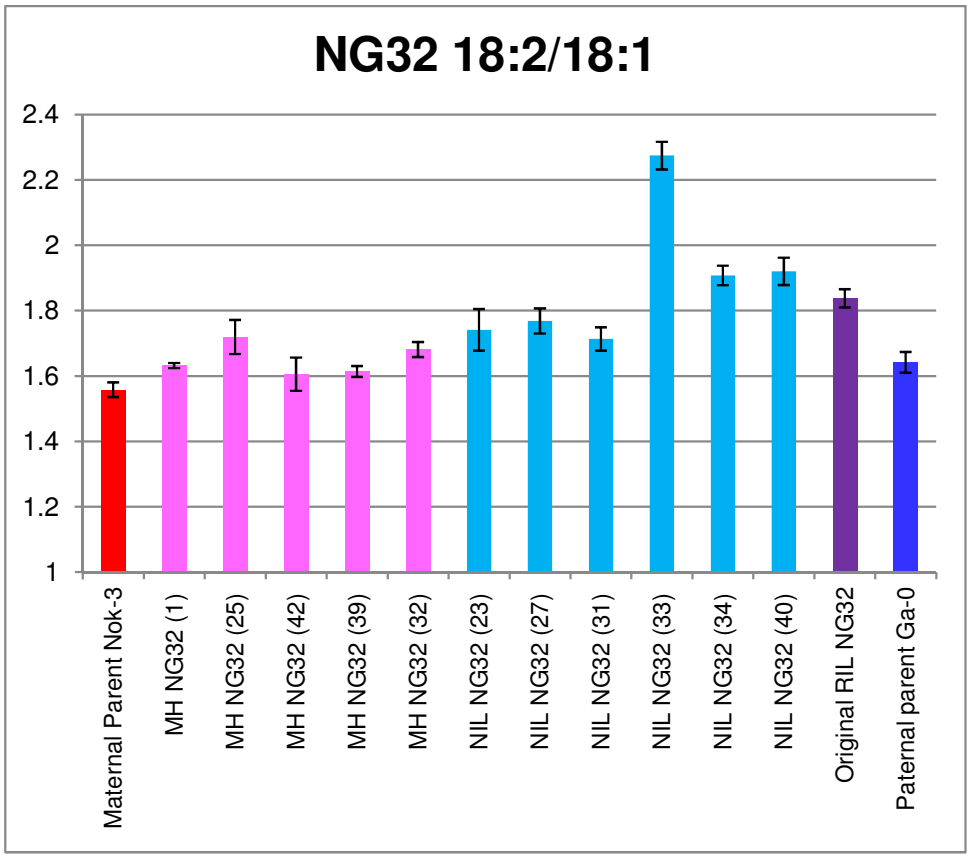


Fig. A5-9 Average of 18;2/18:1 ratio of individual NIL candidates and MH plants of NG32

TJ79 18:2/18:1

A is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ79	14	18	21	29	42	45	5	17	25	30	33	38	Ts-5	Mz-0
MS_At1_0.9	0.02	1	0.9	0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconn1_2.7	0	1		3	A														
jconn1_3.0	0.01	1		4	A														
MS_At1_4.3	0.37	1	4.3	9	A														
athsrp54a	0.16	1	5.2	13	A														
f20d23ind18-18a	0.2	1	5.5	14	A														
f9h16ind26-26	0.71	1	7.0	18	U	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f8k728985	0.19	1	7.6	19	A														
MS_At1_8.2	0.21	1	8.2	24	A														
nga392	0.96	1	9.8	31	A														
aths0392	1.99	1	10.9	36	A														
MS_At1_11.2	1.99	1	11.2	37	A														
MS_At1_12.9	2.62	1	12.9	45	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t27k12-sp6	2	1	15.5	48	A														
MS_At1_20_b	1.39	1	20.0	61	A														
MS_At1_20_a	1.5	1	20.0	61	A														
nga128	2.09	1	20.2	64	A														
MS_At1_22.1	1.6	1	22.1	66	A														
MS_At1_22.9	1.8	1	22.9	66	A														
MS_At1_23.4	2.11	1	23.4	66	A														
f5i1449495	2.51	1	24.0	68	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_24.3	2.41	1	24.3	68	A														
f5a1859436	1.49	1	26.2	83	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga111	0.57	1	27.0	88	A														
jconn2_0.7	0.02	2		0	A														
MS_At2_3.0	0	2	3.0	10	A														
MS_At2_5.4	0	2	5.4	10	A	A	A	A	A	A	A	A	A	A	A	-	A	A	B
MS_At2_6.3	0.05	2	6.3	12	A														
~~jconn2_3.0~~		2			A														
~~jconn2_5.2~~		2			A														
ciw3	0.08	2	6.4	13	U														
MS_At2_9.3	0.05	2	9.3	29	B	A	A	-	A	A	A	-	A	A	A	A	A	A	B
pls8	0	2	9.9	31	B														
nga1126	0.03	2	11.6	35	B														
c4h	0.05	2	12.9	45	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t6a23ind10-10	0	2	16.1	51	A														
nga168	0.02	2	16.2	51	A														
MS_At2_17.5	0.05	2	17.5	54	B														
MS_At2_17.9	0	2	17.9	55	B	A	A	A	A	A	A	A	A	A	A	A	A	A	-
athubique	0.16	2	19.3	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga32	0.19	3	0.3	0	U														
nga162	0.03	3	4.6	9	B	A	H	H	A	H	A	A	H	A	H	B	A	A	B
MS_At3_6.1	0.27	3	6.1	15	A														
msd2129380	0.06	3	7.7	20	A														
mzn14inda9-a9	0.41	3	10.7	56	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t13j1016610	0.25	3	11.5	64	A														
k11j14ind16-16	0.27	3	13.4	64	A														
MS_At3_17.3	0.11	3	17.3	77	A														
t16k521877	0.52	3	18.9	82	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At3_19.7	0.42	3	19.7	87	A														
athcdc2bg	0.25	3	20.5	91	A														
nga6	0.32	3	23.0	95	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_0.3	1.28	4	0.3	0	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_1.1	1.16	4	1.1	7	B														
t19j18ind30-30	1.95	4	2.3	11	B	A	A	A	A	A	A	A	A	A	A	A	A	A	-
~~nga8~~		4			B														
det1.2	0.5	4	5.3	19	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
~~f28m11ind22-22~~		4			B														
ciw6	0.17	4	7.0	28	A														
t6k21ind15-15	0.32	4	9.9	33	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw7	0.05	4	10.6	39	A														
nga1139	0.81	4	15.4	49	A														
MS_At4_16.2	0.58	4	16.2	51	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t5j1748070	0.34	4	17.7	60	A														
nga225	0	5	1.4	0	A														
MS_At5_2.2	0.02	5	2.2	3	A														
MS_At5_2.8	0.03	5	2.8	4	A														
MS_At5_3.6	0.06	5	3.6	6	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f14f1847646	0.18	5	3.8	6	A														
nga151	0.28	5	4.7	7	U														
MS_At5_6.6a	0.22	5	6.6	14	A														
MS_At5_6.6b	0.22	5	6.6	14	A														
MS_At5_8.6	0	5	8.6	19	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga139	0.01	5	8.6	20	B														
MS_At5_9.7	0.06	5	9.7	28	A														
mlf18ind14-14	0	5	14.6	40	A														
MS_At5_15.8	0.66	5	15.8	50	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw9	1.13	5	17.2	57	B	B	B	B	H	B	B	A	A	A	A	A	A	A	B
MS_At5_18	1.4	5	18.0	59	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
~~f20d23ind18-18b~~	4.5	5		65	B														
nga129	4.81	5	19.7	67	B														
MS_At5_20.4	4.72	5	20.4	67	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
MS_At5_21.3	2.62	5	21.3	70	B	-	-	-	-	-	-	-	-	-	-	-	-	-	-
jv65-66	2.56	5	22.1	75	B														
MS_At5_22.8	3.23	5	22.8	78	B	H	B	B	B	H	B	A	A	A	A	A	A	A	B
mth12	3.54	5	24.1	85	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_26.1	3.22	5	26.1	93	A														
k8a1022396	1.56	5	26.9	98	A														

Fig. A5-10 Genotype file of NIL candidates and MH plants of TJ79 for 18:2/18:1

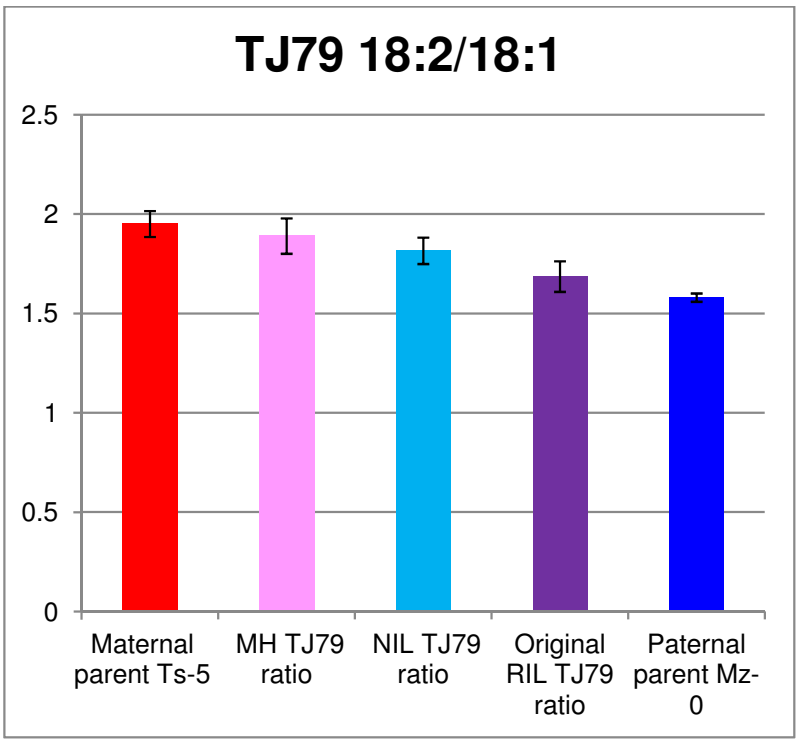


Fig. A5-11 Average of 18:2/18:1 ratio of NIL candidates and MH plants of TJ79

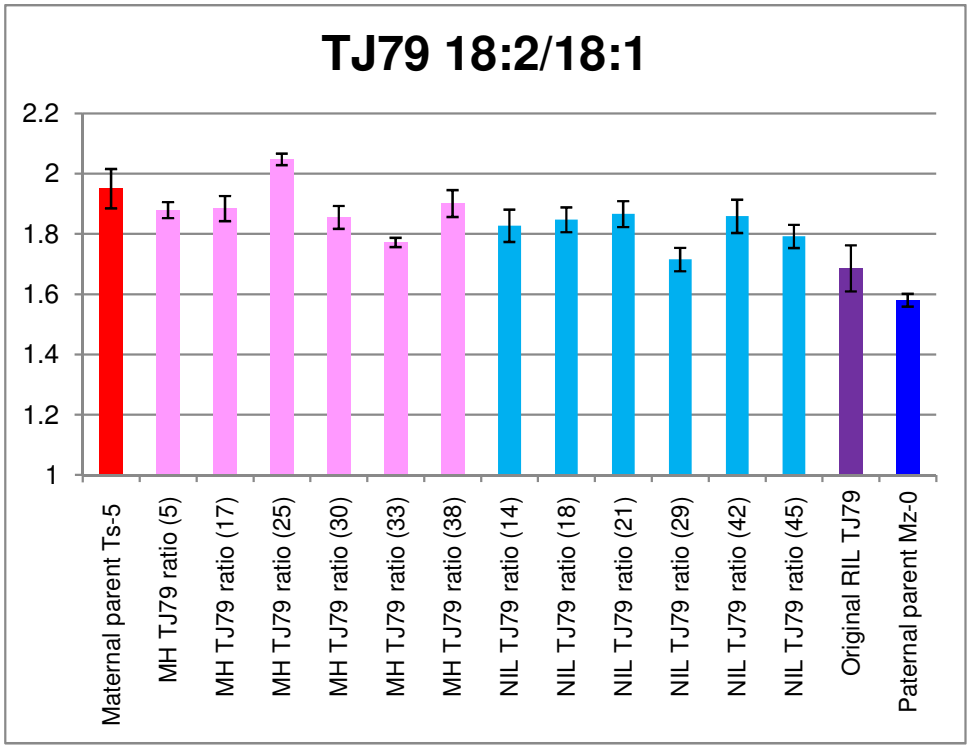


Fig. A5-12 Average of 18:2/18:1 ratio of individual NIL candidates and MH plants of TJ79

TJ17 18:3

A is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ17	1	7	12	15	24	32	11	16	18	19	26	28	Ts-5	Mz-0
MS_At1_0.9	0.14	1	0.9	0	A	A	A	A	A	A	-	A	A	-	A	A	A	A	B
jconn1_2.7	0.87	1		3	A														
jconn1_3.0	1.1	1		4	B														
MS_At1_4.3	1.69	1	4.3	9	A	A	A	A	A	A	-	A	A	-	A	A	A	A	B
athsrp54a	2.05	1	5.2	13	A														
f20d23ind18-18a	1.81	1	5.5	14	B	H	H	H	H	H	H	A	H	A	A	A	A	A	H
f9h16ind26-26	2.81	1	7.0	18	B														
f8k728985	2.02	1	7.6	19	B	B	B	B	B	B	H	A	H	A	A	A	A	A	B
MS_At1_8.2 (ril25)	4.01	1	8.2	24	B	B	B	B	B	B	B	A	A	A	A	A	A		
nga392	5.42	1	9.8	31	B														
aths0392	6.32	1	10.9	36	B														
MS_At1_11.2 (ril106)	5.66	1	11.2	37	B	B	B	B	B	B	B	A	A	A	A	A	A		
MS_At1_12.9	2.21	1	12.9	45	B	H	B	B	-	B	-	A	A	-	-	A	A	A	B
t27k12-sp6	1.26	1	15.5	48	B														
MS_At1_20b	1.02	1	20.0	61	B														
MS_At1_20a	1.1	1	20.0	61	B	H	B	B	B	A	-	H	H	-	H	A	H	A	B
nga128	1.27	1	20.2	64	B														
MS_At1_22.1	1.25	1	22.1	66	B														
MS_At1_22.9	1.25	1	22.9	66	B														
MS_At1_23.4	1.53	1	23.4	66	B														
f5i1449495	1.67	1	24.0	68	B														
MS_At1_24.3	1.6	1	24.3	68	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f5a1859436	1	1	26.2	83	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga111	0.41	1	27.0	88	B														
jconn2_0.7	0.21	2		0	A														
MS_At2_3.0	0.03	2	3.0	10	A														
MS_At2_5.4	0.03	2	5.4	10	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-
MS_At2_6.3	0	2	6.3	12	A														
~~jconn2_3.0~~		2			A														
~~jconn2_5.2~~		2			A														
ciw3	0.01	2	6.4	13	A														
MS_At2_9.3	0.95	2	9.3	29	B	A	A	A	-	H	A	H	H	A	H	H	H	H	B
pls8	0.57	2	9.9	31	B														
nga1126	0.12	2	11.6	35	B	H	A	H	H	H	A	B	H	A	B	H	A	A	B
c4h	0.12	2	12.9	45	B														
t6a23ind10-10	0.01	2	16.1	51	B														
nga168	0	2	16.2	51	B	H	A	A	H	H	A	B	H	A	B	A	A	A	B
MS_At2_17.5	0.16	2	17.5	54	B														
MS_At2_17.9	0.11	2	17.9	55	B														
athubique	0.23	2	19.3	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga32	1.07	3	0.3	0	A														
nga162	1.43	3	4.6	9	B	H	B	H	A	B	A	A	H	H	A	B	A	A	B
MS_At3_6.1	1.9	3	6.1	15	B														
msd2129380	1.85	3	7.7	20	B	H	B	H	H	B	H	A	H	A	A	B	H	A	B
mzn14inda9-a9	0.48	3	10.7	56	U	A	A	A	A	A	A	A	A	A	-	A	A	A	B
t13j1016610	1.62	3	11.5	64	U	-	-	-	-	-	-	-	-	-	-	-	-	-	-
k11j14ind16-16	1.61	3	13.4	64	A														
MS_At3_17.3	0.21	3	17.3	77	A														
t16k521877	0.38	3	18.9	82	A														
MS_At3_19.7	0.38	3	19.7	87	A	A	A	A	-	A	A	A	-	-	-	-	-	A	B
athcdc2bg	0.33	3	20.5	91	A														
nga6	0.32	3	23.0	95	A														
MS_At4_0.3	0.49	4	0.3	0	A														
MS_At4_1.1	0.88	4	1.1	7	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t19j18ind30-30	1.02	4	2.3	11	A														
~~nga8~~		4			A														
det1.2	0.64	4	5.3	19	A														
~~f28m11ind22-22~~		4			A														
ciw6	0.09	4	7.0	28	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-
t6k21ind15-15	0.03	4	9.9	33	A														
ciw7	0.03	4	10.6	39	A														
nga1139	0	4	15.4	49	A														
MS_At4_16.2	0.04	4	16.2	51	A	A+477	A	A	A	A	A	A	A	A	A	A	A	A	B
t5j1748070	0.36	4	17.7	60	A														
nga225	1.68	5	1.4	0	A														
MS_At5_2.2	0.77	5	2.2	3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_2.8	0.59	5	2.8	4	A														
MS_At5_3.6	0.05	5	3.6	6	A														
f14f1847646	0	5	3.8	6	A														
nga151	0.02	5	4.7	7	A														
MS_At5_6.6a	0.01	5	6.6	14	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_6.6b	0.01	5	6.6	14	A														
MS_At5_8.6	0.08	5	8.6	19	A														
nga139	0.05	5	8.6	20	A														
MS_At5_9.7	0.26	5	9.7	28	A														
mlf18ind14-14	0.23	5	14.6	40	A														
MS_At5_15.8	0.18	5	15.8	50	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw9	0.11	5	17.2	57	A														
MS_At5_18	0.19	5	18.0	59	A														
~~f20d23ind18-18b~~	0.39	5		65	A														
nga129	0.39	5	19.7	67	A														
MS_At5_20.4	0.41	5	20.4	67	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_21.3	0.16	5	21.3	70	A														
jv65-66	0.6	5	22.1	75	A														
MS_At5_22.8	0.44	5	22.8	78	A														
mth12	0.94	5	24.1	85	A														
MS_At5_26.1	0.44	5	26.1	93	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
k8a1022396	0.12	5	26.9	98	A														

Fig. A5-13 Genotype file of NIL candidates and MH plants of TJ17 for 18:3

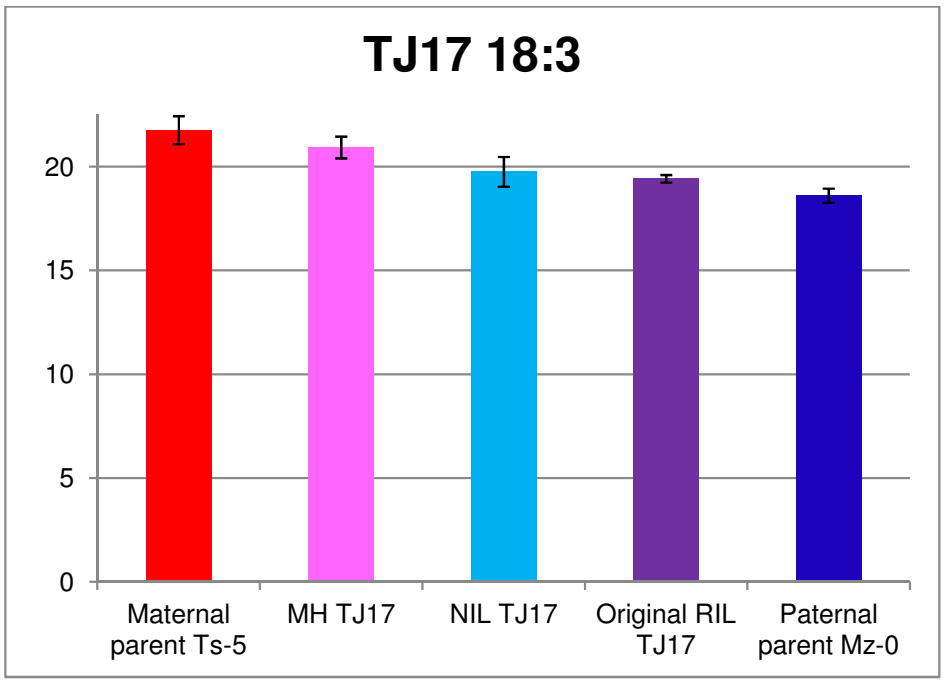


Fig. A5-14 Average of 18:3 contends of NIL candidates and MH plants of TJ17

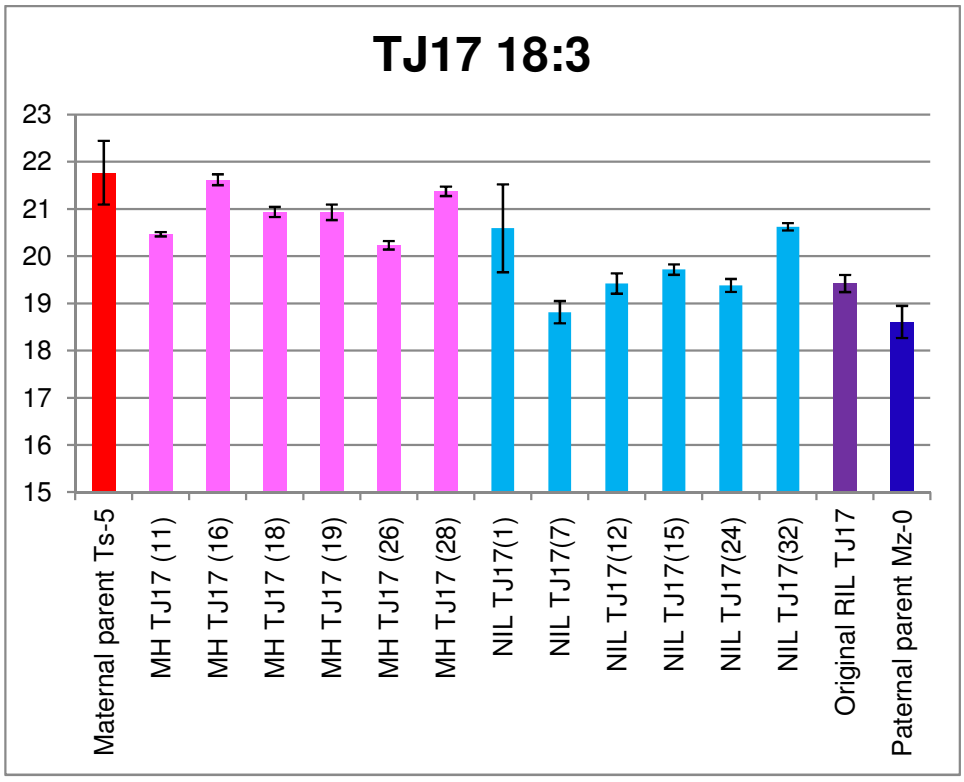


Fig. A5-15 Average of 18:3 contents of individual NIL candidates and MH plants of TJ17

NG32 18:3/18:2

A is increasing allele					RIL	NIL candidates							MH					Parents	
Markers	**LOD**	**LG**	**Mb**	**cM**	**NG32**	**23**	**27**	**31**	**33**	**34**	**40**	**1**	**25**	**32**	**39**	**42**	**Nok-3**	**Ga-0**
nga59	0.6	1	0.0	0	B	B	B	A	A	B	B	B	B	A	B	B	A	B
MS_At1_0.6	0.33	1	0.6	4	B													
MS_At1_2.7	0.22	1	2.7	10	B	H	A	A	A	H	B	H	H	-	A	B	A	B
nga63	0	1	3.2	12	B													
f16j7-trb	0.02	1	3.8	15	B	H	A	A	A	H	B	H	H	A	A	B	A	B
athsrp54a	0	1	5.2	18	A													
MS_At1_6.4	0	1	6.4	21	A													
f9h16ind26-26	0.14	1	7.0	27	A													
MS_At1_7.4	0.36	1	7.4	30	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_8.2	0.34	1	8.2	38	A													
nga392	0.18	1	9.8	43	A													
aths0392	0.23	1	10.9	48	A													
MS_At1_11.2	0.46	1	11.2	50	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f140426482	0.47	1	13.2	58	A													
t27k12-sp6	0.87	1	15.5	59	A													
~~MS_At1_16.4~~		1	~~16.4~~		A													
MS_At1_20_a	1.46	1	20.0	66	A													
nga128	2.02	1	20.2	69	A													
MS_At1_22.6	2.97	1	22.6	71	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t12p18ind8-8	1.96	1	23.5	81	A													
f5i1449495	1.68	1	24.0	84	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f5a1859436	0.1	1	26.2	95	A													
nga111	0.16	1	27.0	98	A													
yup8h12r	0.08	1	29.1	106	A	A	H	A	H	B	B	H	H	A	H	B	A	B
nga1145	0.07	2	0.7	0	B	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw2	0.08	2	1.2	2	B													
t12j241450	0	2	3.6	8	B	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw3	0	2	6.4	15	B													
pls1	0.58	2	8.2	27	B	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	0.83	2	9.9	36	A													
athgpa1	0.35	2	11.1	44	A													
nga1126	0.25	2	11.6	45	A													
t32f646516	1.9	2	13.8	55	A													
MS_At2_14.2_b	1.81	2	14.2	56	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga168	2	2	16.2	65	A													
90j19t7	1.1	2	19.6	75	A													
nga32	0.01	3	0.3	0	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f17a931902	0	3	2.1	8	A													
nga162	0.08	3	4.6	14	A													
msd2129380	0.4	3	7.7	25	U	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At3_8.3_b	0.21	3	8.3	30	A													
athgapab	0.59	3	9.8	38	A													
t13j1016610	2.07	3	11.5	46	A	A	A	A	A	A	A	A	A	A	A	A	A	B
k11j14ind16-16	1.25	3	13.4	48	A													
MS_At3_18.3	1.56	3	18.3	55	A													
t16k521877	2.3	3	18.9	58	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athcdc2bg	1.34	3	20.5	72	A													
nga6	0.37	3	23.0	84	A													
jv30/31	0.49	4	0.0	0	A													
MS_At4_0.3	0.47	4	0.3	1	A	A	A	A	A	A	A	-	A	A	A	A	A	B
t19j18ind30-30	1.14	4	2.3	6	A													
nga8	0.83	4	4.6	11	A													
MS_At4_5.3	1.29	4	5.3	13	A													
ciw6	1.04	4	7.0	29	B	A	A	A	A	A	A	-	A	A	A	A	A	B
MS_At4_10.8	0.38	4	10.8	44	A													
t22a6ind10-10	1.86	4	12.6	53	A													
f27g1959898	1.01	4	12.8	57	A	A	A	A	A	A	A	A	A	A	A	A	A	B
~~t27e11ind12-12~~		4			A													
nga1139	0.81	4	15.4	68	A													
nga1107	0.29	4	17.0	74	B	A	A	H	H	A	A	H	H	A	A	A	A	B
t5j1748070	0.06	4	17.7	77	B													
nga225	0.31	5	1.4	0	B													
mhf15ind52-52	0.06	5	1.7	4	B	B	B	H	B	B	B	A	A	A	A	A	A	B
ciw14	0.05	5	2.1	5	B													
f14f1847646	0.55	5	3.8	9	B	B	B	H	B	B	B	A	A	A	A	A	A	B
nga151	0.24	5	4.7	12	B	B	B	B	B	B	B	A	A	A	A	A	A	B
nga139	3	5	8.6	34	B													
MS_At5_9.7	1.03	5	9.7	45	B	B	B	B	B	B	B	A	A	A	A	A	A	B
MS_At5_13.9	0.54	5	13.9	53	B	B	B	B	B	B	B	A	A	-	A	A	A	B
mlf18ind14-14	0.48	5	14.6	56	B	B	B	B	B	B	B	H	A	H	A	A	A	B
aths0191	0.34	5	15.1	60	A													
MS_At5_17.8	0.07	5	17.8	80	A													
mg158836	0.46	5	19.3	84	U	A	A	A	A	A	A	A	A	A	-	A	A	B
MS_At5_20.4	0.05	5	20.4	89	A													
jv65-66	0	5	22.1	97	A													
mrg2138439	0	5	25.2	114	B													
k8a1022396	0	5	26.9	122	B	H	H	B	A	B	H	B	A	H	H	H	A	B

Fig. A5-16 Genotype file of NIL candidates and MH plants of NG32 for 18:3/18:2

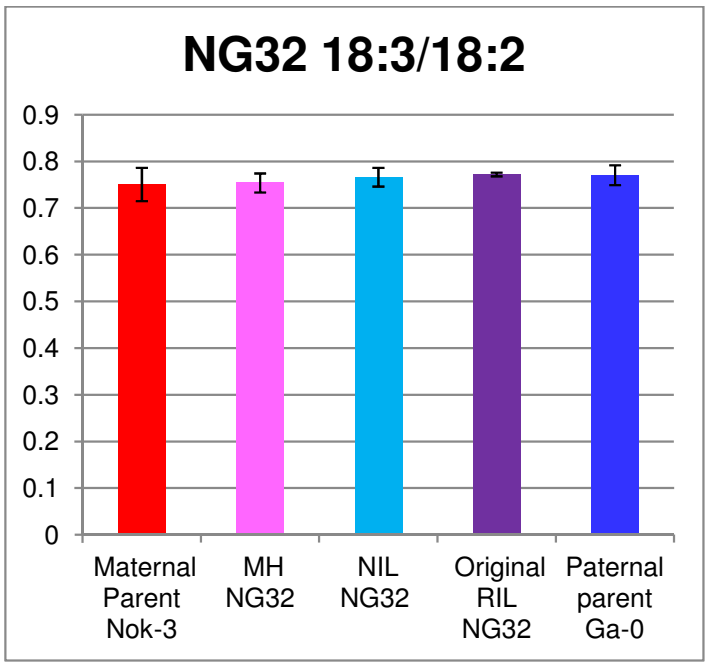


Fig. A5-17 Average of 18:3/18:2 ratio of NIL candidates and MH plants of NG32

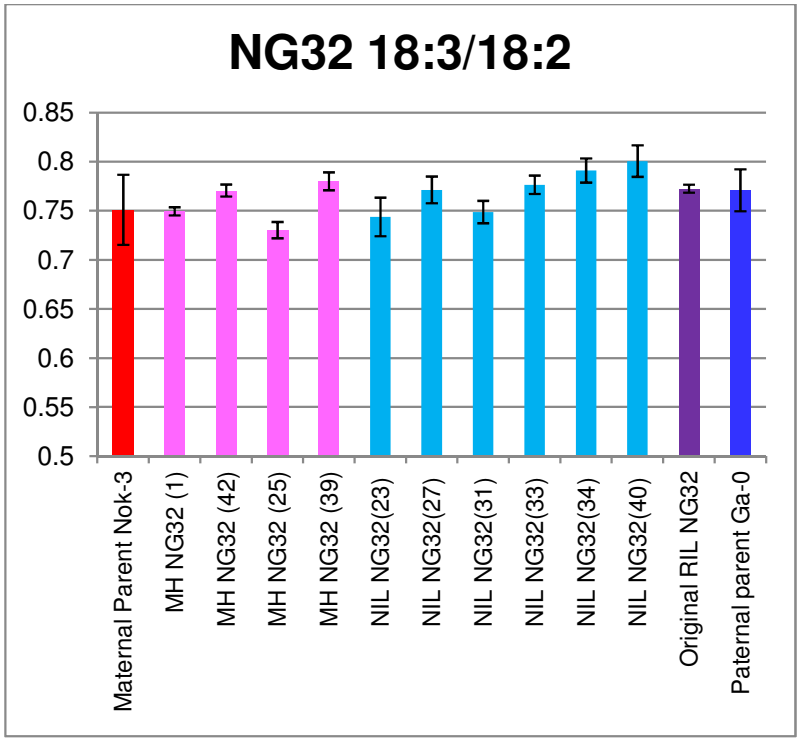


Fig. A5-18 Average of 18:3/18:2 ratio of individual NIL candidates and MH plants of NG32

NG93 18:3/18:2

A is increasing allele					RIL	NIL candidates							MH					Parents	
Markers	**LOD**	**LG**	**Mb**	**cM**	**NG93**	**2**	**5**	**14**	**16**	**23**	**26**	**38**	**1**	**4**	**9**	**15**	**29**	**Nok-3**	**Ga-0**
nga59	0.6	1	0.0	0	A	A	A	A	A	A	A	A	A	A	A	_	A	A	B
MS_At1_0.6	0.33	1	0.6	4	A														
MS_At1_2.7	0.22	1	2.7	10	A														
nga63	0	1	3.2	12	A	_	A	A	A	A	A	A	A	A	A	_	A	A	B
f16j7-trb	0.02	1	3.8	15	A														
athsrp54a	0	1	5.2	18	A														
MS_At1_6.4	0	1	6.4	21	A														
f9h16ind26-26	0.14	1	7.0	27	A														
MS_At1_7.4	0.36	1	7.4	30	U	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_8.2	0.34	1	8.2	38	A														
nga392	0.18	1	9.8	43	A														
aths0392	0.23	1	10.9	48	A														
MS_At1_11.2	0.46	1	11.2	50	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f140426482	0.47	1	13.2	58	A														
t27k12-sp6	0.87	1	15.5	59	A														
~~MS_At1_16.4~~		1	~~16.4~~		A														
MS_At1_20_a	1.46	1	20.0	66	B	B	B	B	B	B	B	B	A	A	A	A	A	A	B
nga128_20.2	2.02	1	20.2	69	B														
MS_At1_22.6	2.97	1	22.6	71	B														
t12p18ind8-8	1.96	1	23.5	81	B														
f5i1449495	1.68	1	24.0	84	B	B	B	B	B	B	B	B	A	A	A	A	A	A	B
f5a1859436	0.1	1	26.2	95	B														
nga111	0.16	1	27.0	98	B	A	A	A	A	A	A	A	A	A	A	_	A	A	B
yup8h12r	0.08	1	29.1	106	B	A	A	A	A	A	A	A	A	A	A	_	A	A	B
nga1145	0.07	2	0.7	0	B	A	A	A	A	A	A	A	A	A	A	_	A	A	B
ciw2	0.08	2	1.2	2	B														
t12j241450	0	2	3.6	8	A														
ciw3	0	2	6.4	15	A														
pls1	0.58	2	8.2	27	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	0.83	2	9.9	36	A														
athgpa1	0.35	2	11.1	44	A														
nga1126	0.25	2	11.6	45	A														
t32f646516	1.9	2	13.8	55	A	_	A	A	A	A	A	A	A	A	A	_	A	A	B
MS_At2_14.2_b	1.81	2	14.2	56	A														
nga168	2	2	16.2	65	A														
90j19t7	1.1	2	19.6	75	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga32	0.01	3	0.3	0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f17a931902	0	3	2.1	8	A														
nga162	0.08	3	4.6	14	A	A	A	A	A	A	A	A	A	A	A	_	A	A	B
msd2129380	0.4	3	7.7	25	A														
MS_At3_8.3_b	0.21	3	8.3	30	A														
athgapab	0.59	3	9.8	38	A														
t13j1016610	2.07	3	11.5	46	A	A	A	A	A	A	A	A	A	A	A	_	A	A	B
k11j14ind16-16	1.25	3	13.4	48	A														
MS_At3_18.3	1.56	3	18.3	55	A														
t16k521877	2.3	3	18.9	58	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athcdc2bg	1.34	3	20.5	72	B														
nga6	0.37	3	23.0	84	B	B	H	H	B	A	A	A	H	A	B	_	B	A	B
jv30/31	0.49	4	0.0	0	B	A	A	A	A	A	A	A	A	A	A	_	A	A	B
MS_At4_0.3	0.47	4	0.3	1	B														
t19j18ind30-30	1.14	4	2.3	6	B														
nga8	0.83	4	4.6	11	B														
MS_At4_5.3	1.29	4	5.3	13	U	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw6	1.04	4	7.0	29	A														
MS_At4_10.8	0.38	4	10.8	44	A														
t22a6ind10-10	1.86	4	12.6	53	A														
f27g1959898	1.01	4	12.8	57	A														
~~t27e11ind12-12~~		4			A														
nga1139	0.81	4	15.4	68	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1107	0.29	4	17.0	74	A														
t5j1748070	0.06	4	17.7	77	A	A	A	A	A	A	A	A	A	A	A	_	A	A	B
nga225	0.31	5	1.4	0	A														
mhf15ind52-52	0.06	5	1.7	4	A														
ciw14	0.05	5	2.1	5	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f14f1847646	0.55	5	3.8	9	A														
nga151	0.24	5	4.7	12	A														
nga139	3	5	8.6	34	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_9.7	1.03	5	9.7	45	A														
MS_At5_13.9	0.54	5	13.9	53	U	A	A	A	A	A	A	A	A	A	A	_	A	A	B
mlf18ind14-14	0.48	5	14.6	56	A														
aths0191	0.34	5	15.1	60	A														
MS_At5_17.8	0.07	5	17.8	80	A														
mq158836	0.46	5	19.3	84	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_20.4	0.05	5	20.4	89	A														
jv65-66	0	5	22.1	97	A														
mrg2138439	0	5	25.2	114	B	A	B	A	H	H	H	H	B	B	H	H	A	A	B
k8a1022396	0	5	26.9	122	B														

Fig. A5-19 Genotype file of NIL candidates and MH plants of NG93

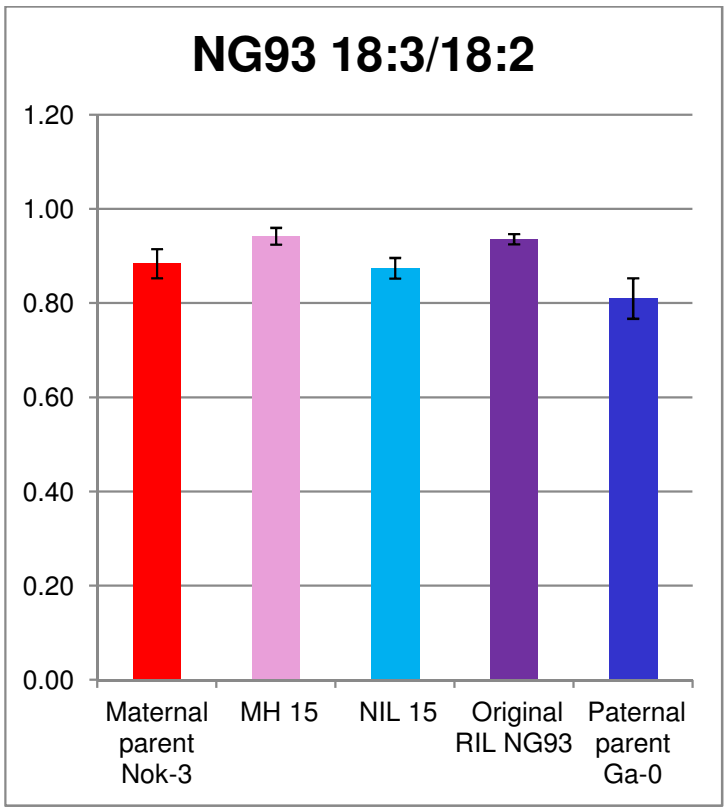


Fig. A5-20 Average of 18:3/18:2 ratio of NIL candidates and MH plants of NG93

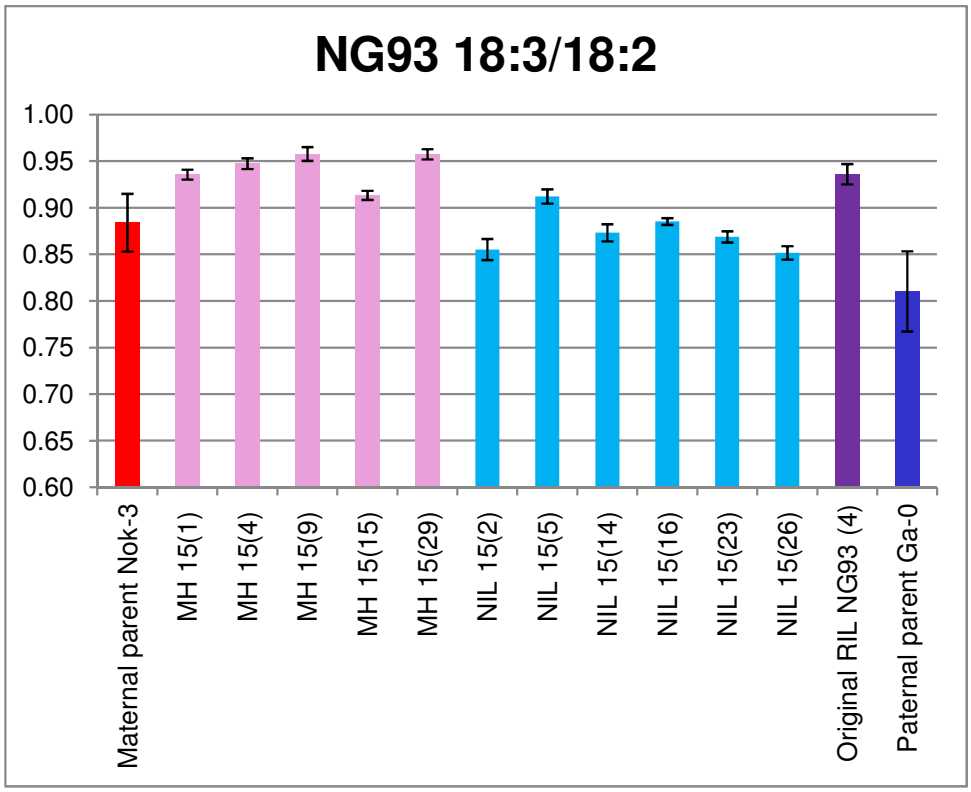


Fig. A5-21 Average of 18:3/18:2 ratio of individual NIL candidates and MH plants of NG93

SG10 18:3/18:2

B is increasing					RIL	NIL candidates									MH							Parents	
Markers	LOD	LG	Mb	cM	SG10	4	7	10	25	26	28	30	36	39	11	24	31	32	34	40	45	Sorbo	Gy-0
t25k16ind9-9	0.14	1	0.0	0.0	B	A	H	H	B	B	B	H	H	H	H	H	B	A	H	A	H	A	B
jconnchr1_2.7x	0.14	1		10.1	A																		
MS_At1_3.7	0.02	1	3.7	13.3	A																		
athsrp54a	0.37	1	5.2	25.4	A																		
f9h16ind26-26	0.57	1	7.0	28.9	A																		
MS_At1_8.2	1.16	1	8.2	35.3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga392	0.48	1	9.8	40.3	A																		
aths0392	0.72	1	10.9	45.0	A																		
MS_At1_12.9	0.44	1	12.9	51.6	A																		
f140426482	0.51	1	13.2	53.7	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t27k12-sp6	0.47	1	15.5	54.4	A																		
MS_At1_18.2	0.45	1	18.2	62.9	A																		
MS_At1_20_b	0.03	1	20.0	68.5	A																		
nga128	0.06	1	20.2	71.8	A																		
MS_At1_22.6	0.07	1	22.6	75.6	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t12p18ind8-8	0.12	1	23.5	80.4	A																		
MS_At1_24.3	0.18	1	24.3	82.0	A																		
MS_At1_26.2	0.63	1	26.2	93.4	A																		
MS_At1_27.2	0.15	1	27.2	99.2	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1145	0.67	2	0.7	0.0	U	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconnchr2_3	0.31	2		4.6	B																		
pls1	0.75	2	8.2	13.6	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	0.4	2	9.9	18.0	A																		
nga1126	0.99	2	11.6	22.1	A																		
c4h	0.98	2	12.9	29.1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t32f646516	0.77	2	13.8	32.2	A																		
MS_At2_14.2_a	1.01	2	14.2	4.1	A																		
t6a23ind10-10	0.78	2	16.1	38.9	A																		
nga168	0.74	2	16.2	39.7	A																		
MS_At2_16.9	1.68	2	16.9	42.3	A																		
MS_At2_18.5	2.05	2	18.5	44.8	A	A	A	A	A	A	_	A	A	A	A	A	A	A	A	A	A	A	B
90j19t7	1.92	2	19.6	47.5	A																		
MS_At3_0.9	1.19	3	0.9	0.0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f17a931902	2.03	3	2.1	7.2	A																		
nga162	2.56	3	4.6	12.1	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
mil23ind14-14	2.56	3	7.5	18.0	B																		
msd2129380	3.61	3	7.7	20.1	U	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At3_8.9_a	3.53	3	8.9	28.5	A																		
athgapab	3.17	3	9.8	32.8	A																		
t13j1016610	4.03	3	11.5	50.5	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
k11j14ind16-16	3.79	3	13.4	51.0	A																		
MS_At3_16.9	6.1	3	16.9	58.1	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B
MS_At3_17.3	6.11	3	17.3	58.3	B																		
t16k521877_18.9Mb	5.1	3	18.9	60.9	B																		
f15b8ind11-11	3.86	3	21.4	68.0	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B
nga6	1.14	3	23.0	74.4	B	B	B	H	B	B	B	B	H	B	A	A	A	A+	A+	A+	A	A	B
t19j18ind30-30	0	4	2.3	0.0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga8	0.02	4	4.6	1.4	A																		
f28m11ind22-22	0.06	4	6.3	6.0	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw6	0	4	7.0	12.1	B																		
MS_At4_7.7	0.02	4	7.0	13.3	B																		
ciw7	0	4	10.6	25.1	B																		
f27g1959898	0	4	12.8	31.4	U	_	_	_	_	_	_	_	_	_	A	A	A	A	A	A	_	_	_
nga1139	0.21	4	15.4	37.8	A																		
MS_At4_16.2	0.19	4	16.2	40.8	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t5j1748070	0.09	4	17.7	48.5	B																		
nga225	0.17	5	1.4	0.0	B																		
mhf15ind52-52	1.58	5	1.7	17.9	B	H	H	A	B	H	B	H	H	H	B	B	B	A	H	H	H	A	B
ciw14	1.38	5	2.1	18.4	B																		
MS_At5_3.6	1.85	5	3.6	20.6	B																		
nga151	2.12	5	4.7	21.4	B	_	A	A	A	A	A	A	_	_	A	A	A	A	A	A	_	_	_
MS_At5_5.3	1.53	5	5.3	24.2	B																		
jconnchr5_7.8	1.27	5	7.8	30.8	B																		
MS_At5_8.6	1.7	5	8.6	34.0	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconnchr5_13.6	0.23	5		39.4	B																		
aths0191	0.69	5	15.1	42.2	B																		
MS_At5_15.8	0.84	5	15.8	45.4	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw9	0.04	5	17.2	51.0	A																		
MS_At5_18	0.47	5	18.0	54.0	A																		
mg158636	0.71	5	19.3	55.5	A																		
nga129	0.64	5	19.7	57.1	A	_	A	A	A	A	A	A	_	_	A	_	A	A	A	A	_	_	_
MS_At5_21.3	0.42	5	21.3	60.7	A																		
MS_At5_22.8	1.4	5	22.8	62.9	A																		
MS_At5_26.1	2.42	5	26.1	69.5	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B

Fig. A5-22 Genotype file of NIL candidates and MH plants of SG10 for 18:3/18:2

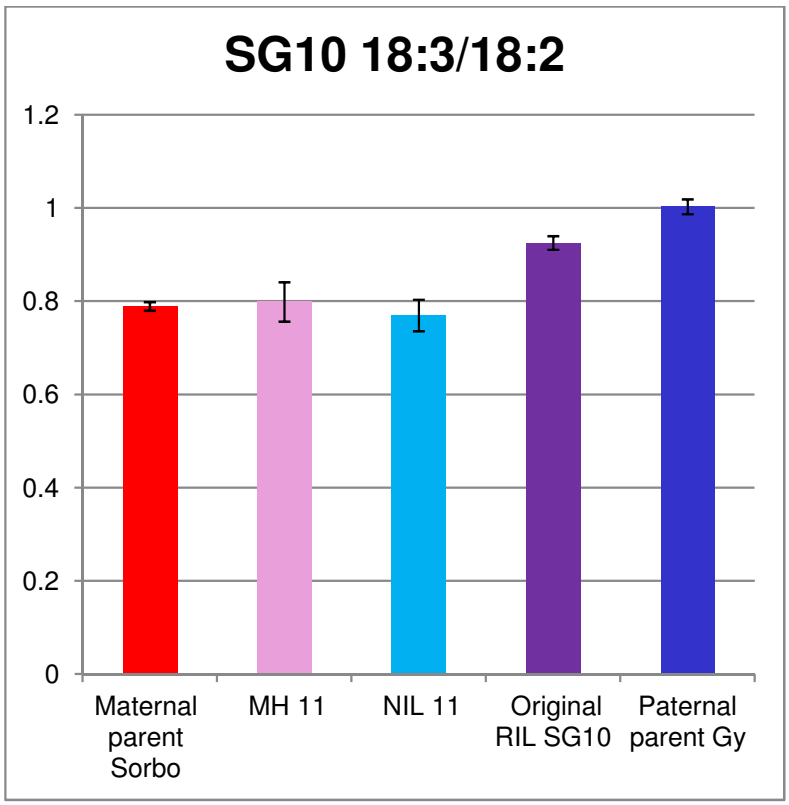


Fig. A5-23 Average of 18:3/18:2 ratio of NIL candidates and MH plants of SG10

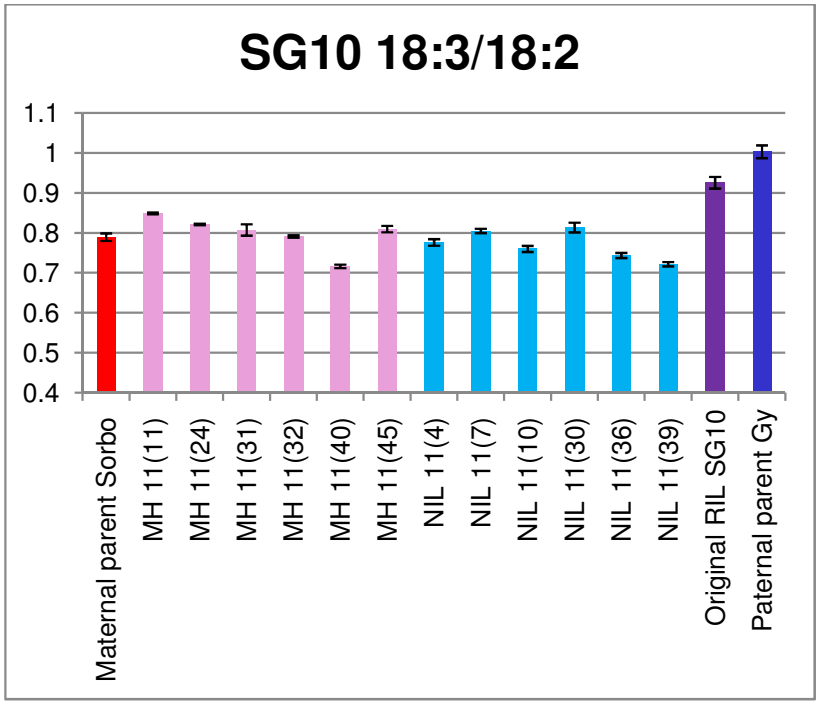


Fig. A5-24 Average of 18:3/18:2 ratio of individual NIL candidates and MH plants of SG10

KB55 oil

B is increasing					RIL	NIL candidates									MH								Parents	
Markers	LOD	LG	Mb	cM	KB55	2	4	17	22	26	33	38	41	42	20	23	24	25	28	29	32	39	Kondara	Br-0
nga63	3.99	1	3.2	0.0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athsrp54a	2.35	1	5.2	5.6	A																			
f9h16ind26-26	2.1	1	7.0	7.0	A																			
MS_At1_10	2.65	1	10.0	18.2	A																			
aths0392	0.77	1	10.9	22.4	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_13.4	0	1	13.4	38.3	A																			
MS_At1_16.2	0.15	1	16.2	40.1	A																			
MS_At1_18.2	0.04	1	18.2	46.3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_22.1	0.11	1	22.1	58.9	A																			
t12p18ind8-8	0	1	3.6	63.6	A																			
MS_At1_24.7	0	1	24.7	66.4	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f5a1859436	0	1	26.2	76.8	A																			
f24j13ind32-32	0	1	26.5	76.8	A																			
nga1145	0.75	2	0.7	0.0	A																			
t16b233819	1.23	2	1.4	1.8	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconnchr2_3	0.85	2	3.0	4.4	A																			
MS_At2_4.4	0.28	2	4.4	5.6	A																			
jconnchr2_5.5	1.18	2	5.5	5.8	A																			
MS_At2_6.3	0.4	2	6.3	8.7	A																			
pls1	1.35	2	8.2	17.7	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
pls8	1.49	2	9.9	27.0	A																			
nga1126	1.54	2	11.6	31.4	A																			
c4h	2.35	2	12.9	39.3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At2_14.2	1.81	2	14.2	41.9	A																			
t6a23ind10-10	2.9	2	16.1	47.2	A																			
nga168	2.9	2	16.2	47.3	A																			
MS_At2_17.5	4.05	2	17.5	50.4	A																			
athubique	5.74	2	19.3	53.9	A																			
90j19i7	6	2	19.6	57.1	A																			
nga32	0.07	3	0.3	0.0	A																			
f17a931902	0.09	3	2.1	9.1	A																			
nga162	0.96	3	4.6	16.0	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
msd2129380	0.08	3	7.7	29.6	A																			
athgapab	0.01	3	9.8	43.0	A																			
k11j14ind16-16	0.18	3	13.4	59.0	A																			
MS_At3_16	0.15	3	16.0	60.7	B	H	H	H	H	B	H	H	H	H	B	H	H	B	H	B	H	B	A	B
MS_At3_16.9	0.45	3	16.9	68.5	B																			
t16k521877	0.1	3	18.9	80.0	B	H	A	A	H	B	A	A	B	B	B	A	A	B	H	B	A	B	A	B
MS_At3_19.7	0.19	3	19.7	85.9	B																			
MS_At3_21.4	0.51	3	21.0	91.9	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jv30-31	0.04	4	0.0	0.0	B	A	H	H	B	H	_	H	H	H	H	A	H	B	H	H	B	_	A	
MS_At4_1.1	0.22	4	4.0	15.1	B	H	H	H	B	H	B	H	H	H	H	A	H	B	H	H	B	B	A	B
t19j18ind30-30	0.16	4	12.5	25.0	U																			
nga8	0.02	4	4.6	28.4	B																			
f28m11ind22-22	0.04	4	6.3	33.1	U																			
ciw6	0.09	4	7.0	39.3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_8.8	0	4	8.8	42.0	B																			
t6k21ind15-15	0.11	4	9.9	44.2	B																			
ciw7	0.13	4	10.6	47.8	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_11.2	0.11	4	11.0	50.6	B																			
MS_At4_13.2	0.01	4	2.0	57.2	U	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1139	0.72	4	15.4	72.2	B																			
MS_At4_16.1	0.32	4	16.1	6.2	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t5j1748070	0	4	17.7	85.7	B																			
nga225	0.08	5	1.4	0.0	A																			
mhf15ind52-52	0.14	5	1.7	0.7	A																			
MS_At5_2.2	0.06	5	2.2	1.6	A																			
f14f1847646	0.11	5	3.8	5.7	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga151	0.2	5	4.7	7.6	A																			
MS_At5_5.3	0.13	5	5.3	9.5	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_6.6_a	0.09	5	6.6	11.3	A																			
nga139	0	5	8.6	17.0	A																			
jconnchr5_12.7	0.08	5	12.7	29.3	B																			
jconnchr5_13.6	0	5	13.6	30.6	B																			
aths0191	0.13	5	15.1	40.2	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
ciw9	1.71	5	17.2	43.8	B																			
mlnlind8-8	1.22	5	17.8	45.8	B																			
mql58836	1.7	5	19.3	49.6	B																			
nga129	2.21	5	19.7	52.0	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	B
MS_At5_20.0	1.81	5	20.0	53.9	B																			
jv65-66	3.2	5	22.1	56.7	B																			
mth12_24.1Mb	1.92	5	24.1	62.9	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	B
mrg2138439	1.13	5	25.2	66.9	B																			
k8a1022396	1.28	5	26.9	72.8	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	B

Fig. A5-25 Genotype file of NIL candidates and MH plants of KB55 for oil conetns

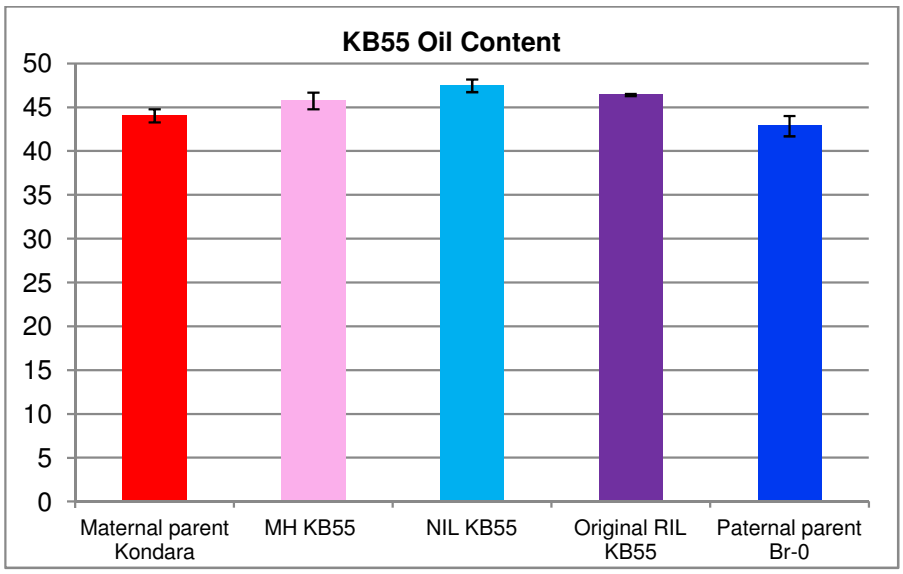


Fig. A5-26 Average of oil contents of NIL candidates and MH plants of KB55

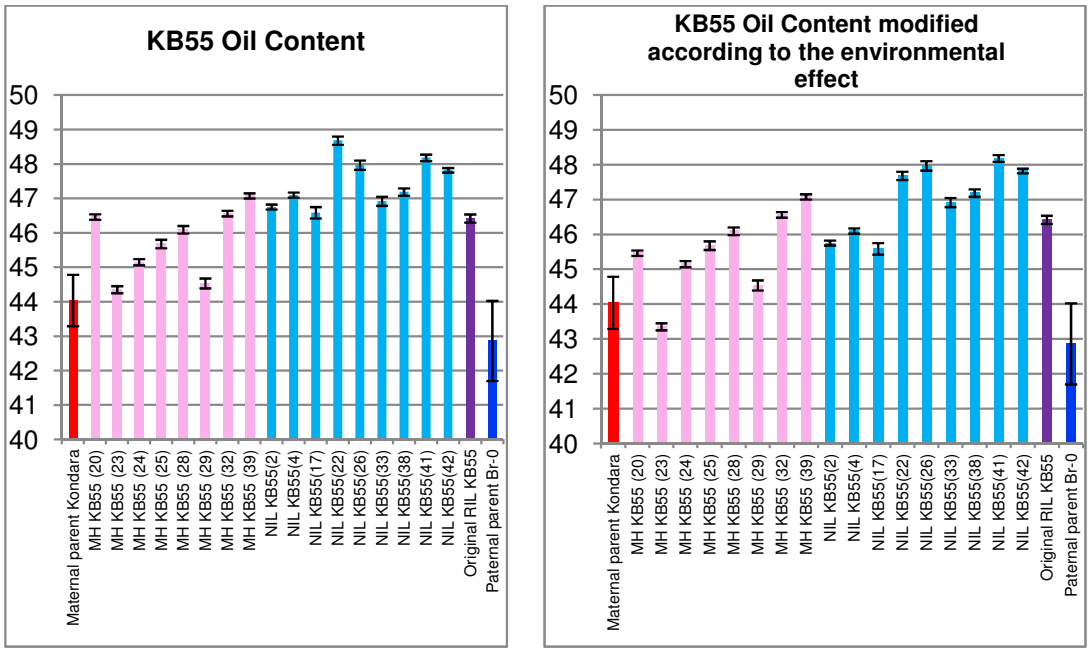


Fig. A5-27 Average of oil contents of individual NIL candidates and MH plants of KB55, original(left) and modified(right). 1% was deducted from MH(20), MH (23), NIL(2), NIL(4), NIL(17) and NIL(22)

TJ79 OIL CH2

TJ79 oil ch2

B is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ79	3	18	23	24	25	29	4	6	10	20	21	37	Ts-5	Mz-0
MS_At1_0.9	1.5	1	0.9	0	A														
jconn1_2.7	0.63	1		3	A														
jconn1_3.0	0.48	1		4	A														
MS_At1_4.3	0.91	1	4.3	9	A	-	A	A	A	A	A	A	A	A	A	A	A	A	B
athsrp54a	1.7	1	5.2	13	A														
f20d23ind18-18a	1.19	1	5.5	14	A														
f9h16ind26-26	0.9	1	7.0	18	U														
f8k728985	0.84	1	7.6	19	A														
MS_At1_8.2	0.93	1	8.2	24	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga392	0.03	1	9.8	31	A														
aths0392	0.18	1	10.9	36	A														
MS_At1_11.2	0.55	1	11.2	37	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_12.9	0.21	1	12.9	45	A														
t27k12-sp6	0.09	1	15.5	48	A														
MS_At1_20b	0.81	1	20.0	61	A														
MS_At1_20a	0.91	1	20.0	61	A														
nga128	1.63	1	20.2	64	A														
MS_At1_22.1	1.83	1	22.1	66	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At1_22.9	1.74	1	22.9	66	A														
MS_At1_23.4	1.65	1	23.4	66	A														
f5i1449495	1.66	1	24.0	68	A														
MS_At1_24.3	1.55	1	24.3	68	A														
f5a1859436	2.61	1	26.2	83	A														
nga111_27Mb	3.52	1	27.0	88	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
jconn2_0.7	0.43	2		0	A														
MS_At2_3.0	0.19	2	3.0	10	A	-	A	A	A	A	A	-	-	A	-	A	A	A	B
MS_At2_5.4	0.2	2	5.4	10	A														
MS_At2_6.3	0.67	2	6.3	12	A														
jconn2_3.0		2			A														
jconn2_5.2		2			A														
ciw3	0.84	2	6.4	13	U	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At2_9.3 (ril91)	3.45	2	9.3	29	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
pls8	3.9	2	9.9	31	B														
nga1126	3.84	2	11.6	35	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
c4h	1.84	2	12.9	45	B	-	-	-	-	-	-	A	A	A	A	A	A	A	B
t6a23ind10-10	1.19	2	16.1	51	A														
nga168	1.33	2	16.2	51	A														
MS_At2_17.5	0.96	2	17.5	54	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At2_17.9	1.31	2	17.9	55	B														
athubique	1.49	2	19.3	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga32	0.93	3	0.3	0	U														
nga162	0.77	3	4.6	9	B	H	H	H	H	H	H	A	A	A	A	B	H	B	B
MS_At3_6.1	0.51	3	6.1	15	A														
msd2129380	0.19	3	7.7	20	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
mzn14inda9-a9	0.11	3	10.7	56	A														
t13j1016610	0.33	3	11.5	64	A														
k11j14ind16-16	0.31	3	13.4	64	A														
MS_At3_17.3	1.69	3	17.3	77	A	-	A	A	A	A	A	A	A	A	A	-	A	A	B
t16k521877	0.97	3	18.9	82	A														
MS_At3_19.7	0.82	3	19.7	87	A														
athcdc2bg	0.59	3	20.5	91	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga6	0.53	3	23.0	95	A														
MS_At4_0.3	0.43	4	0.3	0	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_1.1	0.1	4	1.1	7	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t19j18ind30-30	0.77	4	2.3	11	B														
nga8		4			B														
det1.2	0.21	4	5.3	19	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f28m11ind22-22		4			B														
ciw6	0.01	4	7.0	28	A														
t6k21ind15-15	0.12	4	9.9	33	A														
ciw7	0.56	4	10.6	39	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1139	0.09	4	15.4	49	A														
MS_At4_16.2	0.2	4	16.2	51	A														
t5j1748070	0.02	4	17.7	60	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga225	0.13	5	1.4	0	A														
MS_At5_2.2	0.75	5	2.2	3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_2.8	0.79	5	2.8	4	A														
MS_At5_3.6	0.98	5	3.6	6	A														
f14f1847646	0.57	5	3.8	6	A														
nga151	0.19	5	4.7	7	U	A	A	A	A	A	A	A	A	A	A	A	A	A	-
MS_At5_6.6a	1.04	5	6.6	14	A														
MS_At5_6.6b	1.04	5	6.6	14	A														
MS_At5_8.6	0.66	5	8.6	19	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga139	0.32	5	8.6	20	B														
MS_At5_9.7	0.14	5	9.7	28	A														
mlf18ind14-14	0.06	5	14.6	40	A														
MS_At5_15.8	1.86	5	15.8	50	A														
ciw9	1.99	5	17.2	57	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_18	1.19	5	18.0	59	B														
f20d23ind18-18b	1.24	5		65	B														
nga129	1	5	19.7	67	B														
MS_At5_20.4	0.77	5	20.4	67	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_21.3	1.13	5	21.3	70	B														
jv65-66__22.1Mb	2.22	5	22.1	75	B														
MS_At5_22.8	2.4	5	22.8	78	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
mth12	1.68	5	24.1	85	A														
MS_At5_26.1	1.42	5	26.1	93	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
k8a1022396	0.57	5	26.9	98	A														

Fig. A5-28 Genotype file of NIL candidates and MH plants of TJ79 Ch2 for oil contents

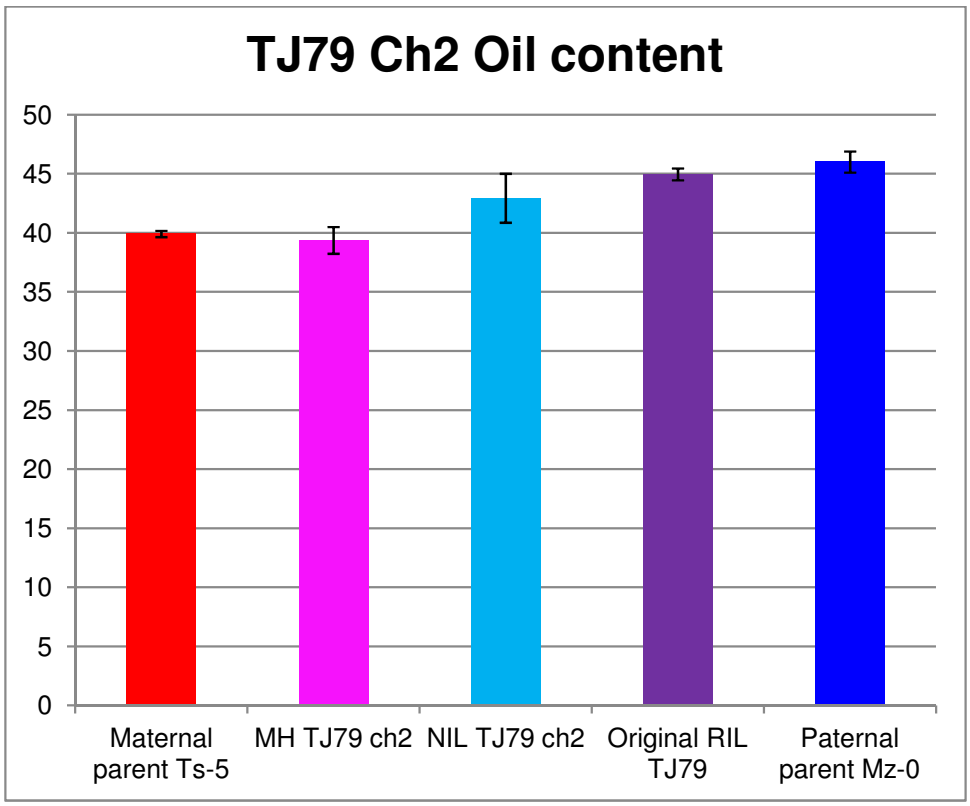


Fig. A5-29 Average of oil contents of NIL candidates and MH plants of TJ79 Ch2

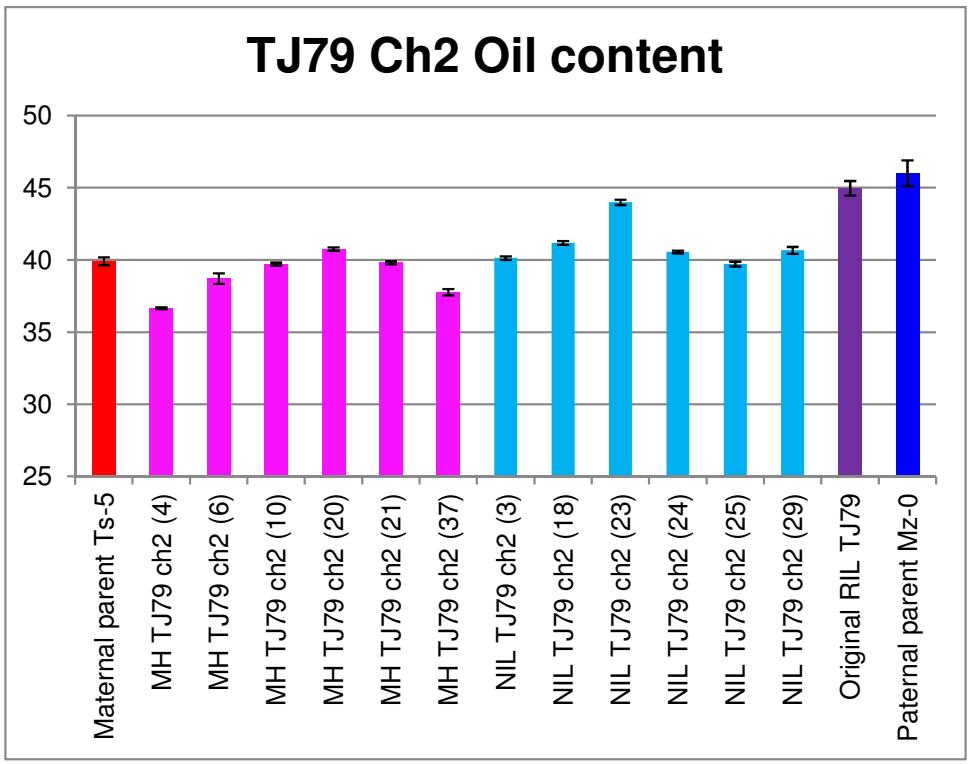


Fig. A5-30 Average of oil contents of individual NIL candidates and MH plants of TJ79 Ch2

TJ79 oil ch5

B is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ79	20	23	24	30	35	40	3	8	18	26	28	31	Ts-5	Mz-0
MS_At1_0.9	1.5	1	0.9	0	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
jconn1_2.7	0.63	1		3	A														
jconn1_3.0	0.48	1		4	A														
MS_At1_4.3	0.91	1	4.3	9	A														
athsrp54a	1.7	1	5.2	13	A														
f20d23ind18-18a	1.19	1	5.5	14	A														
f9h16ind26-26	0.9	1	7.0	18	U														
f8k728985	0.84	1	7.6	19	A														
MS_At1_8.2	0.93	1	8.2	24	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
nga392	0.03	1	9.8	31	A														
aths0392	0.18	1	10.9	36	A														
MS_At1_11.2	0.55	1	11.2	37	A														
MS_At1_12.9	0.21	1	12.9	45	A														
t27k12-sp6	0.09	1	15.5	48	A														
MS_At1_20b	0.81	1	20.0	61	A	A	A	A	A	A	A	A	A	A	A	A	A	-	-
MS_At1_20a	0.91	1	20.0	61	A														
nga128	1.63	1	20.2	64	A														
MS_At1_22.1	1.83	1	22.1	66	A														
MS_At1_22.9	1.74	1	22.9	66	A														
MS_At1_23.4	1.65	1	23.4	66	A														
f5i1449495	1.66	1	24.0	68	A														
MS_At1_24.3	1.55	1	24.3	68	A														
f5a1859436	2.61	1	26.2	83	A														
nga111_27Mb	3.52	1	27.0	88	A	A	A	A	A	A	A	A	A	A	A	A	A	-	-
jconn2_0.7	0.43	2		0	A														
MS_At2_3.0	0.19	2	3.0	10	A														
MS_At2_5.4	0.2	2	5.4	10	A														
MS_At2_6.3	0.67	2	6.3	12	A	A	A	A	A	A	A	-	-	-	-	-	-	A	B
jconn2_3.0		2			A														
jconn2_5.2		2			A														
ciw3	0.84	2	6.4	13	U														
MS_At2_9.3 (ril91)	3.45	2	9.3	29	B														
pls8	3.9	2	9.9	31	B	A	A	A	A	A	A	A	A	A	A	A	A	-	-
nga1126	3.84	2	11.6	35	B														
c4h	1.84	2	12.9	45	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t6a23ind10-10	1.19	2	16.1	51	A														
nga168	1.33	2	16.2	51	A														
MS_At2_17.5	0.96	2	17.5	54	B														
MS_At2_17.9	1.31	2	17.9	55	B	A	A	A	A	A	A	A	A	A	A	A	A	-	-
athubique	1.49	2	19.3	60	B	A	A	A	A	A	A	A	A	A	A	A	A	B	B
nga32	0.93	3	0.3	0	U	A	A	A	A	A	A	A	A	A	A	A	A	-	-
nga162	0.77	3	4.6	9	B	A	A	A	A	A	A	A	A	A	A	A	A	-	B
MS_At3_6.1	0.51	3	6.1	15	A														
msd2129380	0.19	3	7.7	20	A														
mzn14inda9-a9	0.11	3	10.7	56	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
t13j1016610	0.33	3	11.5	64	A														
k11j14ind16-16	0.31	3	13.4	64	A														
MS_At3_17.3	1.69	3	17.3	77	A														
t16k521877	0.97	3	18.9	82	A														
MS_At3_19.7	0.82	3	19.7	87	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athcdc2bg	0.59	3	20.5	91	A														
nga6	0.53	3	23.0	95	A														
MS_At4_0.3	0.43	4	0.3	0	B	H	A	A	A	A	H	H	H	B	B	H	H	B	B
MS_At4_1.1	0.1	4	1.1	7	B														
t19j18ind30-30	0.77	4	2.3	11	B	A	A	A	A	A	A	A	A	A	-	A	A	-	-
nga8		4			B														
det1.2	0.21	4	5.3	19	B	H	A	A	A	A	H	H	H	H	B	H	H	B	B
f28m11ind22-22		4			B														
ciw6	0.01	4	7.0	28	A														
t6k21ind15-15	0.12	4	9.9	33	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
ciw7	0.56	4	10.6	39	A														
nga1139	0.09	4	15.4	49	A														
MS_At4_16.2	0.2	4	16.2	51	A	A	A	A	A	A	A	-	A	A	A	A	A	B	B
t5j1748070	0.02	4	17.7	60	A														
nga225	0.13	5	1.4	0	A														
MS_At5_2.2	0.75	5	2.2	3	A														
MS_At5_2.8	0.79	5	2.8	4	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_3.6	0.98	5	3.6	6	A														
f14f1847646	0.57	5	3.8	6	A														
nga151	0.19	5	4.7	7	U														
MS_At5_6.6a	1.04	5	6.6	14	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
MS_At5_6.6b	1.04	5	6.6	14	A														
MS_At5_8.6	0.66	5	8.6	19	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga139	0.32	5	8.6	20	B														
MS_At5_9.7	0.14	5	9.7	28	A														
mlf18ind14-14	0.06	5	14.6	40	A	A	A	A	A	A	A	-	A	A	A	A	A	B	B
MS_At5_15.8	1.86	5	15.8	50	A														
ciw9	1.99	5	17.2	57	B	-	-	-	A	-	-	A	A	A	A	A	A	-	-
MS_At5_18	1.19	5	18.0	59	B														
f20d23ind18-18b	1.24	5		65	B														
nga129	1	5	19.7	67	B														
MS_At5_20.4	0.77	5	20.4	67	B	B	B	B	H	B	B	H	A	A	A	A	A	B	B
MS_At5_21.3 (ril52)	1.13	5	21.3	70	B	B	B	B	-	B	B	A	A	A	A	A	A	A	B
jv65-66__22.1Mb	2.22	5	22.1	75	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
MS_At5_22.8 (ril53)	2.4	5	22.8	78	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
mth12	1.68	5	24.1	85	A														
MS_At5_26.1	1.42	5	26.1	93	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B
k8a1022396	0.57	5	26.9	98	A														

Fig. A5-31 Genotype file of NIL candidates and MH plants of TJ79 Ch5 for oil contents

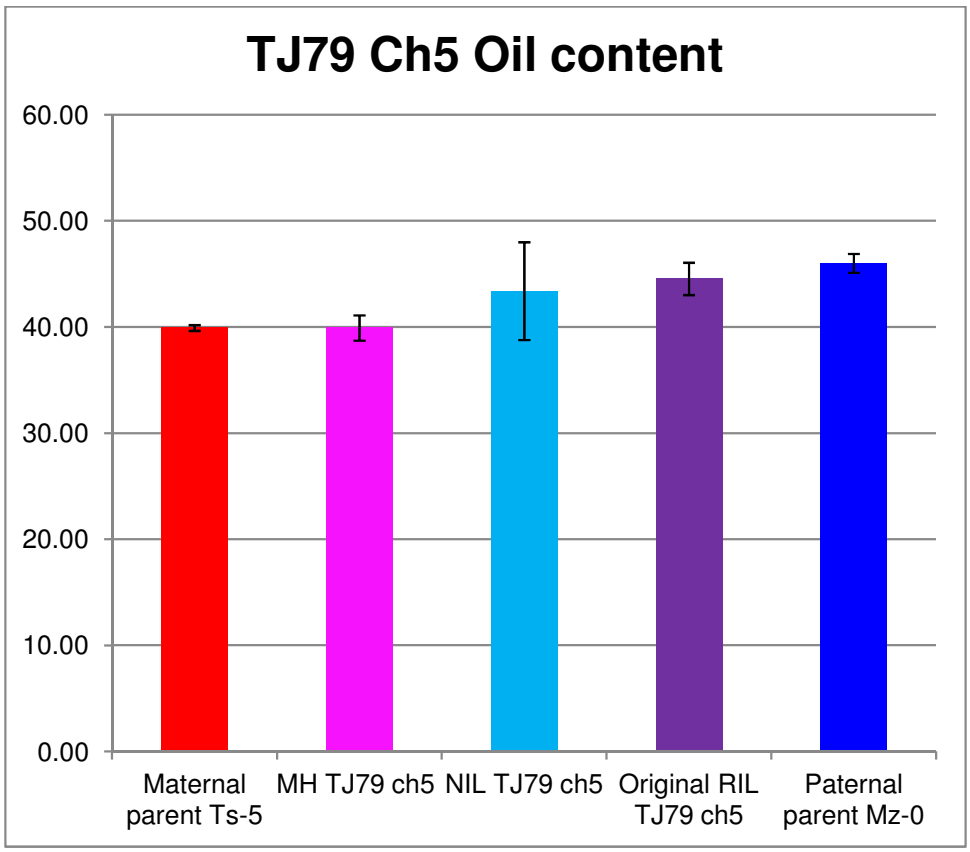


Fig. A5-32 Average of oil contents of NIL candidates and MH plants of TJ79 Ch5

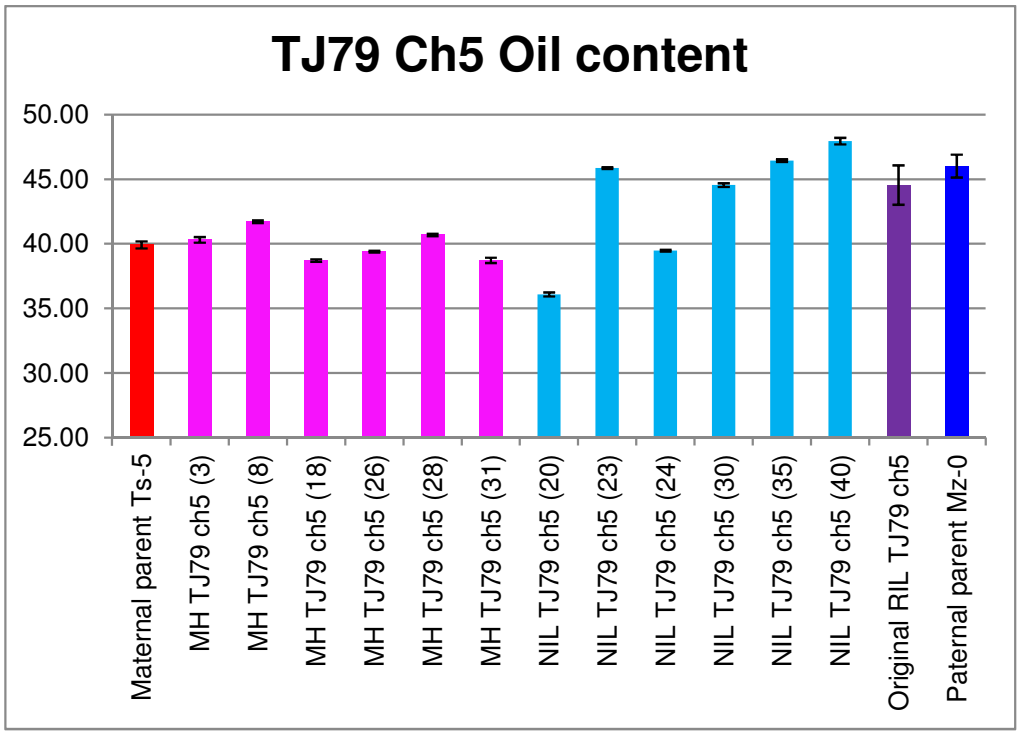


Fig. A5-33 Average of oil contents of individual NIL candidates and MH plants of TJ79 Ch5

TJ83 oil

B is increasing allele					RIL	NIL candidates						MH						Parents	
Markers	LOD	LG	Mb	cM	TJ83	10	22	31	33	40	43	11	12	13	15	21	26	Ts-5	Mz-0
MS_At1_0.9	1.5	1	0.9	0	A														
jconn1_2.7	0.63	1		3	A														
jconn1_3.0	0.48	1		4	A														
MS_At1_4.3	0.91	1	4.3	9	A	A	A	A	A	A	A	-	A	A	A	A	-	A	-
athsrp54a	1.7	1	5.2	13	A														
f20d23ind18-18a	1.19	1	5.5	14	A														
f9h16ind26-26	0.9	1	7.0	18	U	H	H	B	B	H	A	B	H	H	H	B	A	A	-
f8k728985	0.84	1	7.6	19	A														
MS_At1_8.2	0.93	1	8.2	24	A														
nga392	0.03	1	9.8	31	U	A	A	A	A	A	A	A	A	A	A	A	A	A	B
aths0392	0.18	1	10.9	36	A														
MS_At1_11.2	0.55	1	11.2	37	A														
MS_At1_12.9	0.21	1	12.9	45	B	H	H	H	B	H	A	H	H	A	H	A	-	A	-
t27k12-sp6	0.09	1	15.5	48	B														
MS_At1_20b	0.81	1	20.0	61	B														
MS_At1_20a	0.91	1	20.0	61	B	B	B	B	B	B	B	A	A	H	A	A	-	A	-
nga128	1.63	1	20.2	64	B														
MS_At1_22.1	1.83	1	22.1	66	B														
MS_At1_22.9	1.74	1	22.9	66	B														
MS_At1_23.4	1.65	1	23.4	66	B														
f5i1449495	1.66	1	24.0	68	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
MS_At1_24.3 (ril30)	1.55	1	24.3	68	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
f5a1859436	2.61	1	26.2	83	B														
nga111_27Mb	3.52	1	27.0	88	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B
jconn2_0.7	0.43	2		0	A														
MS_At2_3.0	0.19	2	3.0	10	A	A	A	A	A	A	A	A	A	A	A	A	-	A	-
MS_At2_5.4	0.2	2	5.4	10	A														
MS_At2_6.3	0.67	2	6.3	12	U	A	A	A	A	A	A	A	A	A	A	A	-	A	-
jconn2_3.0		2			A														
jconn2_5.2		2			A														
ciw3	0.84	2	6.4	13	A														
MS_At2_9.3	3.45	2	9.3	29	A														
pls8	3.9	2	9.9	31	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga1126	3.84	2	11.6	35	A														
c4h	1.84	2	12.9	45	A														
t6a23ind10-10	1.19	2	16.1	51	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga168	1.33	2	16.2	51	B														
MS_At2_17.5	0.96	2	17.5	54	B														
MS_At2_17.9	1.31	2	17.9	55	B														
athubique	1.49	2	19.3	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga32	0.93	3	0.3	0	A														
nga162	0.77	3	4.6	9	B	H	H	A	H	H	B	A	B	H	A	A	H	A	B
MS_At3_6.1	0.51	3	6.1	15	B	B	B	A	B	B	-	-	-	-	-	-	-	-	-
msd2129380	0.19	3	7.7	20	A														
mzn14inda9-a9	0.11	3	10.7	56	B	H	H	A	A	H	H	A	B	A	H	B	A	A	-
t13j1016610	0.33	3	11.5	64	B														
k11j14ind16-16	0.31	3	13.4	64	B	H	H	A	A	H	H	A	B	A	H	B	A	A	-
MS_At3_17.3	1.69	3	17.3	77	A														
t16k521877	0.97	3	18.9	82	A														
MS_At3_19.7	0.82	3	19.7	87	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
athcdc2bg	0.59	3	20.5	91	A														
nga6	0.53	3	23.0	95	A														
MS_At4_0.3	0.43	4	0.3	0	A														
MS_At4_1.1	0.1	4	1.1	7	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
t19j18ind30-30	0.77	4	2.3	11	A														
nga8		4			A														
det1.2	0.21	4	5.3	19	A														
f28m11ind22-22		4			A														
ciw6	0.01	4	7.0	28	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-
t6k21ind15-15	0.12	4	9.9	33	A														
ciw7	0.56	4	10.6	39	A														
nga1139	0.09	4	15.4	49	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At4_16.2	0.2	4	16.2	51	B														
t5j1748070	0.02	4	17.7	60	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
nga225	0.13	5	1.4	0	A														
MS_At5_2.2	0.75	5	2.2	3	A														
MS_At5_2.8	0.79	5	2.8	4	A														
MS_At5_3.6	0.98	5	3.6	6	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f14f1847646	0.57	5	3.8	6	A														
nga151	0.19	5	4.7	7	A														
MS_At5_6.6a	1.04	5	6.6	14	A														
MS_At5_6.6b	1.04	5	6.6	14	A														
MS_At5_8.6	0.66	5	8.6	19	B	H	A	A	A	A	H	B	H	H	A	H	H	A	B
nga139	0.32	5	8.6	20	B														
MS_At5_9.7	0.14	5	9.7	28	B														
mlf18ind14-14	0.06	5	14.6	40	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
MS_At5_15.8	1.86	5	15.8	50	A														
ciw9	1.99	5	17.2	57	A														
MS_At5_18	1.19	5	18.0	59	B	A	A	A	A	A	A	A	A	A	A	A	A	A	B
f20d23ind18-18b	1.24	5		65	B														
nga129	1	5	19.7	67	B														
MS_At5_20.4	0.77	5	20.4	67	B														
MS_At5_21.3	1.13	5	21.3	70	B	A	H	B	A	A	H	H	B	B	H	H	H	A	B
jv65-66	2.22	5	22.1	75	A														
MS_At5_22.8 (ril53)	2.4	5	22.8	78	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B
mth12	1.68	5	24.1	85	A														
MS_At5_26.1	1.42	5	26.1	93	A														
k8a1022396	0.57	5	26.9	98	A														

Fig. A5-34 Genotype file of NIL candidates and MH plants of TJ79 ch5

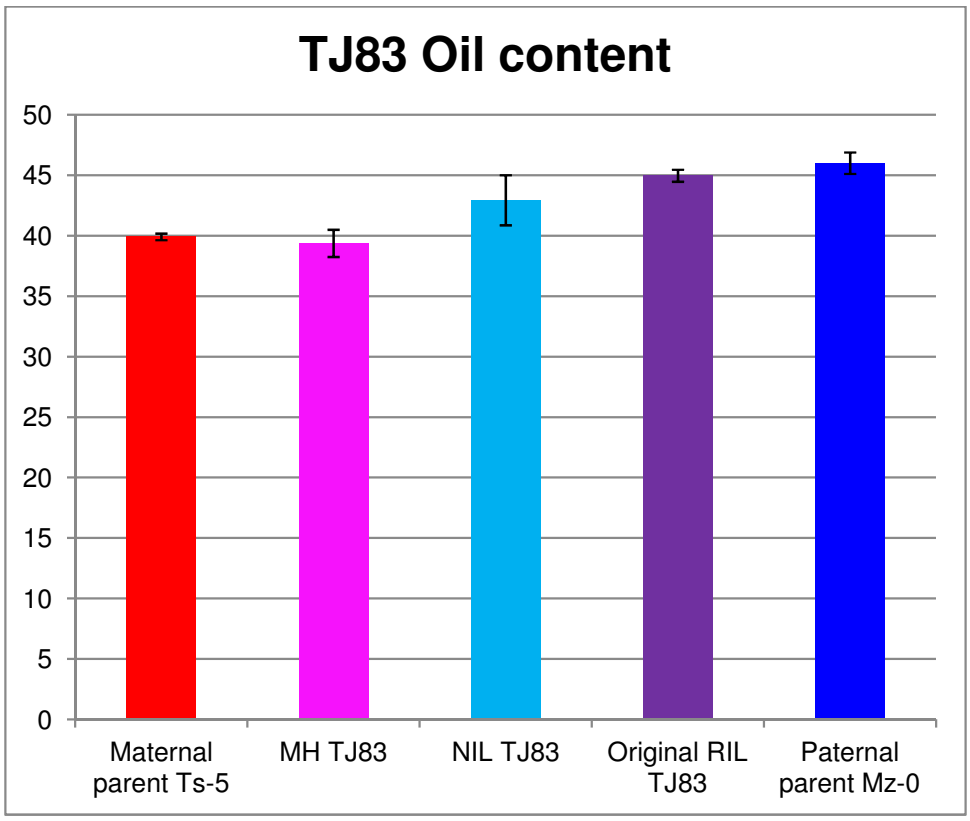


Fig. A5-35 Average of oil contents of NIL candidates and MH plants of TJ83

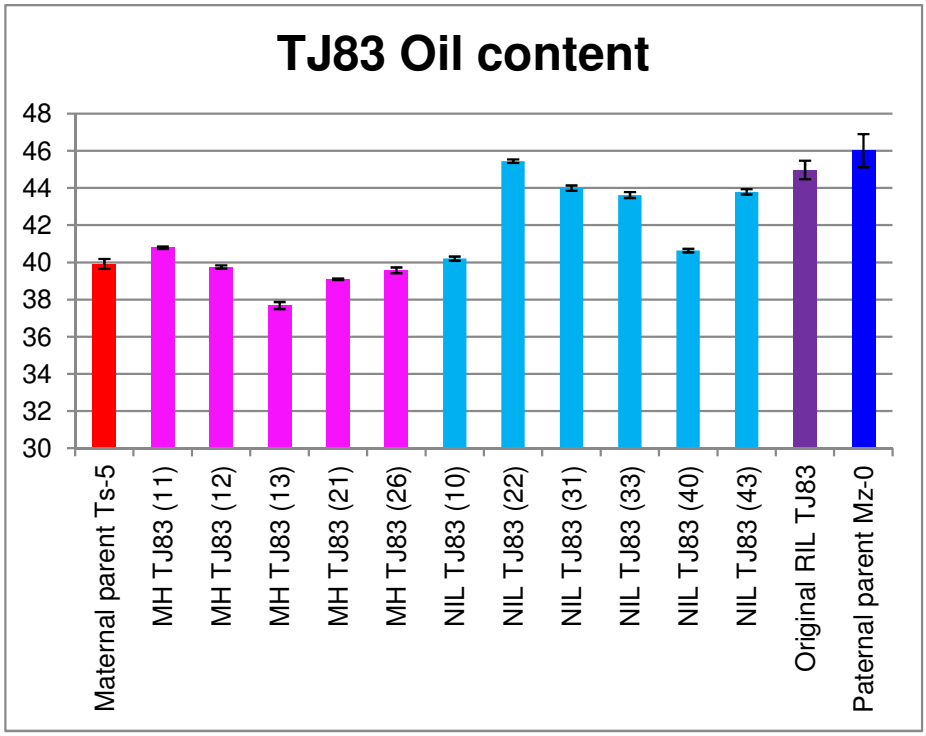


Fig. A5-36 Average of oil contents of individual NIL candidates and MH plants of TJ83

APPENDIX 6 MAPPING POPULATION BINNED

Population (33)

type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin	type		s. no	bin
1	7	41	1	1	15	375	2	1	9	54	3	1	2	2	4	1	21	171	5	1	13	359	6	1	1	1	7	1	4	25	8	1	10	333	9	1	8	42	10	1	1	207	11	1	3	239	12
1	7	318	1	1	20	162	2	1	12	112	3	1	2	218	4	1	22	478	5	1	19	433	6	1	3	17	7	1	5	511	8	1	15	125	9	1	18	155	10	1	5	30	11	1	6	33	12
1	9	324	1	1	18	814	2	1	12	343	3	1	4	259	4	1	23	718	5	1	16	799	6	1	24	201	7	1	13	113	8	1	17	404	9	1	21	461	10	1	8	320	11	1	23	482	12
1	10	67	1	1	23	880	2	1	20	443	3	1	6	311	4	1	9	551	4,5	1	20	860	6	1	4	745	7	1	14	372	8	1	24	489	9	1	22	185	10	1	11	73	11	1	11	575	12
1	5	536	1					1	1	730	3	1	14	114	4	1	18	629	5or6	1	16	131	6	1	9	773	7	1	16	398	8	1	17	624	9					1	18	421	11	1	3	736	12
1	6	768	1					1	3	534	3	1	17	136	4									1	10	775	7	1	14	592	8	1	24	722	9					1	19	160	11	1	14	795	12
								1	16	597	3or4	1	2	733	4									1	10	558	7	1	13	790	8	1	5	767	9					1	19	645	11				
												1	11	781	4									1	21	681	7or8	1	15	796	8	1	21	869	9					1	22	871	11				
												1	2	530	4													1	17	813	8																
												1	4	535	4													1	19	819	8																
												1	13	588	4													1	24	886	8									1	11	342	11or12				
												1	1	892	4or5													1	23	192	8or9																
												1	6	546	4or5													1	15	595	8or9																
												1	12	587	4or5																														
												1	20	646	4or5																														
2	47	429	1	2	30	71	2	2	25	4	3	2	42	393	4	2	44	410	5	2	27	27	6	2	30	295	7	2	27	287	8	2	28	291	9	2	25	267	10	2	26	279	11	2	48	265	12
2	42	602	1	2	40	191	2	2	29	292	3					2	45	235	5	2	31	296	6	2	37	149	7	2	28	44	8	2	29	50	9	2	32	313	10	2	33	119	11	2	27	475	12
2	30	674	1	2	43	217	2	2	34	326	3					2	37	567	5	2	39	368	6	2	38	367	7	2	32	103	8	2	34	128	9	2	33	316	10	2	35	129	11	2	29	500	12
2	38	710	1	2	46	420	2	2	41	214	3					2	44	616	5	2	45	619	6	2	40	370	7	2	35	349	8	2	39	161	9	2	36	350	10	2	43	613	11	2	35	549	12
2	46	839	1	2	40	729	2	2	48	438	3													2	31	514	7	2	36	133	8	2	44	230	9	2	37	363	10	2	26	650	11	2	36	565	12
								2	46	625	3									2	31	75	6or7	2	27	651	7	2	43	403	8	2	25	643	9	2	38	151	10	2	28	656	11	2	34	701	12
2	46	251	1or2					2	32	699	3									2	32	518	6or7	2	45	816	7	2	30	502	8	2	37	709	9	2	41	386	10	2	36	704	11	2	35	703	12
								2	40	586	3																	2	38	568	8	2	44	815	9	2	42	216	10	2	48	851	11				
								2	25	451	3													2	45	419	7?	2	41	791	8					2	47	253	10								
																								2	33	526	7?	2	47	848	8					2	26	455	10	2	26	26	11or12				
																												2	34	544	8or9					2	28	499	10	2	41	598	11or12				
																												2	43	806	8or9					2	39	570	10	2	31	698	11or12				
																																				2	48	642	10								
																																				2	29	672	10								
																																				2	39	711	10								
																																				2	42	800	10								
3	61	480	1	3	49	327	2	3	50	331	3	3	54	37	4	3	52	361	5	3	50	21	6	3	52	28	7	3	60	198	8	3	55	78	9	3	54	377	10	3	56	385	11	3	49	9	12
3	69	304	1	3	56	169	2	3	51	23	3	3	65	243	4	3	58	434	5	3	53	29	6	3	53	374	7	3	68	984	8	3	58	178	9	3	64	234	10	3	57	126	11	3	59	445	12
3	55	630	1	3	59	195	2	3	55	380	3	3	65	523	4	3	62	222	5	3	68	539	6	3	60	446	7	3	63	755	8	3	63	497	9	3	66	246	10	3	57	177	11	3	66	525	12
3	66	761	1	3	72	327	2	3	67	272	3	3	67	531	4	3	63	224	5	3	72	573	6	3	61	221	7					3	57	666	9	3	69	542	10	3	62	486	11	3	71	564	12
				3	54	609	2	3	68	294	3	3	52	596	4	3	51	831	5	3	51	584	6	3	67	762	7					3	68	774	9	3	61	692	10	3	70	555	11	3	70	802	12
				3	59	678	2	3	56	867	3	3	52	836	4	3	54	856	5	3	69	776	6	3	49	824	7					3	71	820	9	3	53	850	10	3	71	310	11	3	61	928	12
				3	60	691	2	3	67	973	3					3	55	866	5	3	50	827	6	3	57	884	7									3	58	888	10	3	62	742	11	3	62	929	12
				3	60	914	2	3	69	985	3	3	65	758	4,5,6,	3	70	991	5	3	72	1024	6	3	59	895	7									3	64	943	10					3	66	961	12
				3	63	938	2	3	72	822	3,4,5,	3	51	360	4or5	3	71	994	5					3	65	950	7													3	64	519	11or12				
				3	70	309	2	3	56	644	3or4	3	49	574	0r5or	3	50	582	5or6					3	53	604	7																				
								3	64	756	3or4													3	58	675	7																				
4	73	382	1	4	83	84	2	4	84	92	3	4	79	66	4	4	85	315	5	4	77	414	6	4	75	406	7	4	73	3	8	4	74	18	9	4	75	34	10	4	76	412	11	4	74	388	12
4	91	496	1	4	93	508	2	4	86	353	3	4	88	371	4	4	93	180	5	4	87	132	6	4	84	298	7	4	79	428	8	4	76	35	9	4	80	74	10	4	78	418	11	4	94	510	12
4	84	618	1	4	94	1018	2	4	92	172	3	4	85	640	4	4	77	550	5	4	78	552	6	4	85	104	7	4	81	269	8	4	78	55	9	4	86	115	10	4	81	79	11	4	80	576	12
4	88	676	1					4	94	229	3	4	74	521	4,5,6					4	79	560	6	4	89	679	7	4	90	468	8	4	80	459	9	4	88	145	10	4	82	278	11	4	90	721	12
4	75	833	1					4	96	821	3									4	95	818	6	4	80	857	7	4	92	503	8	4	87	365	9	4	90	154	10	4	83	286	11	4	91	723	12
4	82	876	1					4	84	908	3									4	79	844	6	4	83	889	7	4	86	648	8	4	89	467	9	4	91	156	10	4	89	147	11	4	93	748	12
								4	82	81	3									4	90	995	6					4	87	974	8	4	77	838	9	4	76	537	10	4	75	524	11	4	76	834	12
																				4	96	1052	6									4	81	858	9	4	87	652	10	4	92	725	11	4	85	909	12
																												4	95	1021	8?					4	94	764	10	4	73	825	11	4	89	989	12
																												4	77	51	8or9					4	91	1002	10	4	74	830	11				
																																				4	92	1011	10	4	78	841	11	4	73	513	11or12
																																								4	86	957	11	4	81	594	11or12
																																							4	88	981	11	4	82	605	11or12
																																							4	93	1017	11				

Population(34)

type			bin	type			bin	type			bin	type			bin	type			bin	type			bin	type			bin	type			bin	type			bin	type			bin
1	5	15	1	1	13	420	2	1	1	4	3	1	5	320	4	1	7	360	5	1	3	312	6	1	1	569	7	1	4	14	8	1	3	13	9	1	1	309	10	1	10	76	11	1	6	334	12
1	13	132	1	1	24	297	2	1	7	66	3	1	10	368	4	1	18	221	5	1	11	369	6	1	11	804	7	1	4	319	8	1	11	97	9	1	2	5	10	1	19	231	11	1	12	131	12
1	16	191	1	1	6	596	2	1	8	68	3	1	21	270	4	1	20	269	5	1	20	507	6	1	19	840	7	1	6	36	8	1	22	530	9	1	2	310	10	1	2	771	11	1	14	140	12
1	16	648	1	1	12	626	2	1	14	427	3	1	23	544	4	1	22	279	5	1	18	669	6					1	9	366	8	1	23	281	9	1	12	397	10	1	22	719	11	1	24	568	12
1	17	661	1	1	24	756	4,5,6	1	18	495	3	1	4	776	4	1	14	823	5									1	15	162	8	1	3	773	9	1	2	572	10					1	1	770	12
1	9	71	1					1	6	782	3	1	9	608	4	1	15	644	5									1	15	434	8	1	10	610	9	1	3	575	10	1	8	365	1or12				
1	23	730	1					1	9	789	3	1	15	824	4													1	17	486	8	1	13	634	9	1	4	578	10	1	16	470	1or12				
								1	10	791	3	1	21	712	4													1	19	505	8	1	18	839	9	1	5	592	10	1	17	218	1or12				
								1	13	813	3	1	23	886	4													1	21	509	8	1	20	848	9	1	5	781	10								
								1	14	639	3	1	16	836	4or5													1	21	861	8					1	11	617	10								
								1	19	672	3	1	24	900	4or5																			1	12	808	10								
																																				1	20	701	10								
2	30	390	1	2	34	411	2	2	26	300	3	2	27	19	4	2	27	325	5	2	26	12	6	2	25	286	7	2	35	129	8	2	29	52	9	2	37	152	10	2	29	379	11	2	25	2	12
				2	35	418	2	2	33	106	3	2	38	153	4	2	48	265	5	2	28	333	6	2	31	395	7	2	39	168	8	2	31	84	9	2	41	194	10	2	32	399	11	2	28	37	12
				2	41	438	2	2	34	125	3	2	45	459	4	2	38	892	5	2	44	237	6	2	38	429	7	2	42	211	8	2	36	419	9	2	46	468	10	2	37	428	11	2	36	141	12
				2	43	442	2	2	44	445	3	2	34	852	4					2	25	496	6	2	43	232	7	2	42	439	8	2	40	436	9	2	27	745	10	2	39	489	11	2	45	241	12
				2	28	524	2	2	47	471	3	2	48	955	4or5					2	26	735	6	2	46	254	7	2	25	733	8	2	26	504	9	2	34	566	10	2	47	262	11	2	33	556	12
				2	30	539	2	2	48	482	3									2	31	543	6	2	32	809	7	2	31	803	8	2	27	520	9	2	35	579	10	2	29	538	11	2	43	912	12
				2	30	802	2	2	28	764	3									2	32	548	6	2	37	890	7	2	37	619	8	2	36	581	9	2	35	866	10	2	29	768	11	2	46	951	12
				2	45	702	2	2	44	924	3									2	33	816	6					2	40	898	8	2	38	629	9	2	39	896	10	2	36	877	11				
								2	30	74	3or4									2	39	658	6					2	41	675	8	2	42	910	9					2	43	696	11				
																				2	42	677	6					2	41	902	8	2	32	105	8or9					2	45	938	11				
																				2	40	673	6or7					2	44	698	8	2	48	721	9or10					2	47	954	11				
																												2	46	716	8																
3	56	346	1	3	49	272	2	3	53	46	3	3	50	28	4	3	51	287	5	3	54	314	6	3	68	226	7	3	66	207	8	3	57	348	9	3	49	26	10	3	52	305	11	3	55	332	12
3	65	402	1	3	51	43	2	3	58	88	3	3	55	60	4	3	58	350	5	3	59	109	6	3	50	533	7	3	49	528	8	3	60	356	9	3	54	50	10	3	57	75	11	3	56	72	12
3	68	458	1	3	66	446	2	3	64	178	3	3	61	367	4	3	61	137	5	3	63	378	6	3	51	555	7	3	49	784	8	3	62	376	9	3	62	155	10	3	60	133	11	3	53	822	12
				3	50	790	2	3	69	473	3	3	65	201	4	3	69	713	5	3	72	263	6	3	57	862	7	3	52	586	8	3	72	483	9	3	51	805	10	3	63	170	11	3	61	939	12
				3	55	595	2	3	71	253	3	3	67	452	4	3	69	1016	5	3	54	828	6	3	59	618	7	3	60	918	8	3	61	647	9	3	57	606	10	3	71	480	11	3	67	979	12
				3	65	680	2	3	62	649	3	3	69	229	4	3	56	597	4or5	3	59	906	6	3	68	695	7	3	66	685	8	3	68	1000	9	3	60	638	10	3	52	811	11				
				3	71	1071	2	3	63	653	3	3	70	240	4	3	72	1084	4or5	3	62	947	6					3	67	686	8					3	63	952	10	3	55	830	11				
				3	56	835	2					3	54	593	4					3	64	676	6													3	70	1069	10	3	50	274	1or12				
												3	58	614	4					3	72	783	6													3	67	208	0,11,1	3	64	953	1or12				
												3	58	884	4					3	64	382	6,7,8													3	52	45	10or11								
												3	65	963	4					3	53	589	6or7																								
												3	66	967	4																														
												3	70	737	4																														
												3	71	746	4																														
4	79	688	1	4	73	336	2	4	91	615	3	4	75	359	4	4	76	59	5	4	83	490	6	4	83	177	7	4	75	32	8	4	74	31	9	4	81	469	10	4	82	176	11	4	78	403	12
4	93	842	1	4	86	209	2	4	92	288	3	4	76	388	4	4	80	136	5	4	91	259	6	4	87	570	7	4	77	398	8	4	79	130	9	4	85	523	10	4	93	631	11	4	86	527	12
				4	94	1170	2	4	91	815	3	4	79	435	4	4	82	487	5	4	85	984	6	4	90	255	7	4	78	113	8	4	84	494	9	4	88	585	10	4	94	316	11	4	75	652	12
4	77	70	2or3	4	89	238	3or2	4	95	1240	3	4	73	645	4	4	85	184	5	4	96	865	6	4	82	966	7	4	81	144	8	4	88	225	9	4	89	590	10	4	74	874	11	4	76	663	12
				4	88	1047	2					4	78	937	4	4	84	982	5	4	80	453	6or7	4	83	977	7	4	84	180	8	4	78	678	9	4	92	621	10	4	75	897	11	4	80	943	12
												4	77	671	4,5,6	4	87	785	5					4	94	855	7	4	87	219	8	4	86	774	9	4	94	633	10	4	80	690	11	4	83	710	12
																4	89	800	5									4	93	311	8	4	89	1052	9	4	73	873	10	4	81	962	11	4	85	758	12
																4	90	814	5									4	77	931	8					4	74	646	10	4	84	738	11	4	86	1018	12
																4	91	1138	5									4	88	794	8					4	76	923	10	4	74	342	1or1	4	87	1027	12
																4	95	859	5																	4	79	941	10	4	90	591	1or1	4	92	1156	12
																																				4	81	694	10	4	96	1249	1or12				
																																				4	82	709	10								
																																				4	90	1121	10								
																																				4	92	826	10								
																																				4	93	1166	10								

APPENDIX 7 SCORING OF GENOTYPE BY GENEMAPPER V 3.7

Manual scoring method

1, Unzip the fsa files sent from genotyping and save in U drive.

2, Open Gene Mapper v.3.7 and click file tab and select “Add samples to Project”. Select required folder and click the fsa file to Add To List” and click “Add”.

3, To sort the samples according to the order of submission, click “Edit” tub and select “Sort” by Sample ID.

4, Select analysis method “Microsatelite Default” and size standard is “GS500LIZ”.

5, Click the green arrow icon on the tool bar to initiate analysis and name the project.

6, Select sample to examine and click the icon “Display Plot” on the toolbar. The electropherograms of markers appear on the monitor.

7, Change the colour of label ppearing on the monitor when necessary.

8. Adjust the number of panes on the tool bar and zoom the range of size of interest and peak size within the pane.

9, After checking the parental lines, select samples according to the order to print and adjust as necessary to score.

10, Print the panes into PDF file for record.

Automated scoring method

The method of Automated scoring involves hierarchical naming of file at each processes and setting the marker sizes. Once it has set up, it can be used for other samples if it used the same markers.

As the process is complexed. it require detailed information to fully function. Therefore the method is explained in Appendix

1, Go through the process in manual scoring method and record the size range of parental lines of all the markers to be analysed.

2, In the Tools menu select panel manager and a new window appears. On the left side of new window, there is panel manager pane and “Panel Manager” icon is on the top of the list. Click it and the list of panels already exist appears in the window. In the file menu, select “New Kit” and name the kit. For example the kit was named “Chie First screen” and “Chie Bin”.

3, The way to organize the panels and markers determines the way of output. If a new panel was named the same with Chie Bin and I add all 11 markers in the same panel, the results will be shown in order of sample 1 with all 11 markers results, then sample 2 with all 11 markers result and so on.

If I name the panel individual marker name, then the marker in the pane is one marker of the same name with panel. Then the outcome becomes sample 1 with one marker result, then sample 2 with the same marker results.

Click on the newly created kit in the left side of window and click in the file menu on the tool bar to select “New Panel”. Name the panel with the name of the marker.

4, Click on the kit in the panel manager pane and in the “Bins” menu on the tool bar, select “New Bin Set”. Different bin set of the same marker can be made for the different population. For example, when the product size of marker 1 is 102 and 106 in KB population and 160 and 164 in SG population, marker 1 can have two bin sets. The same population can share the same name of bin set in the different markers. In my case, I used only one population for fine mapping, therefore I had only one bin set.

5, Click the marker and the new window for bin panel will open. Right click in the bin panel and “Add Bin to ‘ selected marker’ ” window will pop up. Name the bin for the name of parental allele such as A or B and decide the location according to the size of PCR product recorded manually. Offset should not overlap with other markers with the same inflorescence colour. Add another bin by right click in the same bin panel and name it with another parental allele and type in size range.

6, Click OK and the panel manager window closes.

7, Back to the window with a spreadsheet, click “Analysis” menu on the menu bar and select “Analysis Method Editor”. Select the “Allele” tab and select the bin set created in the process 4.

8, In the main spreadsheet, there is a column for panel and select the panel for each sample by right click in the cell and bring the drop down menu. To ill down, select a few top cells one by one and then right click the title of the column ”Panel” which select all the column, then Control+D will select all the same panel.

9, Click the green arrow analysis icon and analysis starts.

10, Click the “Genotype” tab to see the spreadsheet of genotyping results.

11, In the “Tools” menu, select “Table Setting Editor” and set the table to match the needs.

12, Export the table from file menu and save as a tab delimited text file which can be opened in Excel.

13, The results should be reviewed to eliminate miss-call of alleles and judge the allele of the sample with low peak manually.

APPENDIX 8 FATTY ACID METHYL ESTER PREPARATION METHOD

The procedure evolved twice along the progress of the project. The first and second procedures described here were applied during NIL development. The original protocol was written to use GC MS machine at JIC with split needle, therefore samples were required to be concentrated to up to the optimum concentration of the machine(Seed lipid methanolysis Reference No: CG-IB-019 by O'Neill 2010). However in order to achieve accurate quantification of FAME, GC FID is more suitable than GC MS, therefore the protocol was adjusted to the use of CG FID at Rothamsted Institute which require only a tenth of concentration of samples used for GC MS. Therefore the concentration process was removed. The third method was developed to deal with a large number of samples for fine mapping using smaller disposable vials and written in section 3.7.1.

FIRST METHODS

This method was used for the phenotype of NG28 population.

Equipment

-15ml Pyrex screw cap tubes

-Hot block

-Pipettor

-Nitrogen shower (Univap)

-Waterbath

-Bench-top centrifuge

-Dispensers x 2

-Vortex

-Pasteur pipettes 2 x the number of sample + extra

-Set of vials for GC analysis

2 ml vials (Chromacol Cat# 2-SVW) with

200 microliter glass inserts (Chromacol Cat# 02-MTVWG)

Blue Screw Caps (Chromacol Cat# 9-SC(B)-ST1)

Plastic foot/adapter (Chromacol Cat# MTS-1)

-GC vials and crimp caps to store back up sample

-Crimper

Materials

-3N Methanolic HCl (Supelco 33050-U 400ml, store @ 4℃)

-iso-Hexane high purity reagent (Fischer Scientific H/0408/17 2.5l store @ room temperature)

-Methanol Analytic reagent (Fischer Scientific M/4000/PP17 2.5l store @ room temperature)

-0.9% w.v.NaCl in dispenser

Method

Preparation

1. Wash Pyrex tube with chloroform and dry for at least 30 minutes.

2. Wash three glass beakers with Chloroform and air dry.

3. Assemble vials for GC analysis. Label two sets of Pyrex tubes and vials for GC analysis and back up.

4. Measure 50mg of seeds into labeled chloroform cleaned Pyrex tube.

FAME extraction

1. Place hot block in the fume hood cabinet and preheat to 80℃.

2. Dilute methanolic HCl (3N) to 1N.

3. Add 1ml of 1N methanolic HCl into each Pylex tube with seeds by Pasteur pipette and screw the lid tightly to avoid evaporation.

4. Place pyrex tubes in the hot block and check the lids after 5 minutes, as they can loosen due to the expansion of the lid by the heat.

5. Heat the sample for two hours in total.

6. Take out the samples and leave them to cool completely for approximately 15minutes.

7. Open the lid and add 1.5ml of 0.9% of NaCl by pump dispenser.

8. Add 1ml of iso-hexane by pump dispenser and close the lid.

9. Vortex for 20 seconds.

10. Spin in bench top centrifuge at 3000rpm for 5 minutes at 21℃.

11. While waiting for the centrifuge to finish, pour water into the water bath attached to the nitrogen shower and turn on to heat the water to 40℃

12. Move tubes from centrifuge to the fume cabinet.

13. Decant off the clear and colourless supernatant solution into a new chloroform cleaned labeled Pyrex tubes. If it was contaminated by the lower red aqueous layer, decant again into fresh tube.

14. Place tubes in the heated water bath. Occasionally Pylex tubes were slightly larger than the hole to support the tube in the bath. In such case transfer the contents into tube which fits the size of the hole.

15. Lower the needles to position in the middle of each tube correctly then slowly release nitrogen from the tank adjusting the pressure carefully.

16. Blow off for 5 to 10 minutes to evaporate iso-hexane.

17. Using Pasteur pipette add 150µl iso- hexane per tube to the FAME prep and transfer a drop into a labeled vial for analysis and the rest into another vial as a back up.

18. Add 9 drops of iso-hexane to dilute the sample into a tenth.

19. Carefully close the lids tight and store in -20℃ freezer.

SECOND METHOD

This method was used to phenotype the remainder of the NILs. The sample preparation up to step 12 was the same as the first method and the sample was diluted rather than concentrated. The calculation of dilution rate is explained after the second method of FAME extraction.

FAME preparation

(Preparation and 1-12 was the same with first method)

13. Aliquot 50µl of the clear and colourless supernatant solution into a sample vial and the remainder into a backup vial.

14. Add 50µl of iso-hexane into a sample vial to dilute by half.

15. Close the lids tightly and store in -20℃ Freezer.

Calculation of dilution rate for the second method

1. If the concentration of original FAME sample of 1ml of hexane (process 8 in first method) was μmol/l and 700µl of hexane was recovered in process13 in first method, there was 700/1000 μ / 1000 = 7/10000 μmol of FAME was captured.

2. After N_2 shower, 150µl of hexane was added to the FAME sample at process 17 in the first method. Therefore the concentration of the final output from first method is

7/10000 μmol x (1000x1000)/150=700/150 μmol/l = 4.66 μmol/l

3. Required concentration is a tenth of 4.66 μmol/l, therefore 0.466 μmol/l

4. To achieve this concentration from the original sample in process 10 in the first method which was μmol/l, dilution to half was done.

GC SETTINGS

The detail of GC settings below were provided by Frederic Beaudoin who performed GC analysis for this project.

Inlet and detector temperature was set to 250℃ and 1 µml of each sample was analysed using splitless injection and a constant flow rate of 2 ml/min. The starting oven temperature was 50℃ and increased rapidly to 190℃ at 40℃/min followed by a slower increase to 240℃ at 1.5℃/min. This temperature cycle was repeated for every sample. FAMEs were detected by a flame ionization detector (FID).

The Chromatograms were analysed by the Agilent ChemStation software. The retention time and identity of each FAME peak was calibrated using the FAME Mix Rapeseed oil standard (Supelco, cat# 07756) supplemented with 2% w/w methyl 11,14-eicosadienoate(C20:2 n-6; Fluka Cat# 17272). 1 mM Methyl heptadecanoate (C17:0; Sigma cat# H4515) was added to some samples as an internal standard.

APPENDIX 9 SAMPLES FOR PARENTAL LINE GENOMIC SEQUENCE

These are the detail information of samples used for genomic sequencing of parental lines of RIL. Selected lines and their concentration and amount of DNA were shown below.

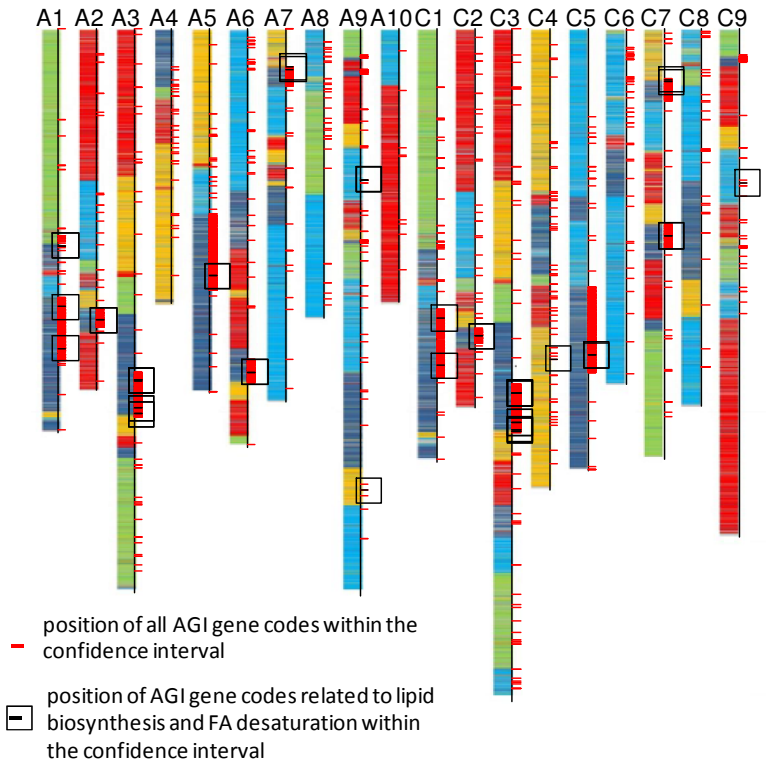
Kondara(1)	13/01/2011	17.3	µg/mL	1ug
Gy-0(1)	13/01/2011	17.7	µg/mL	1ug
Br-0(1)	13/01/2011	18.8	µg/mL	1.1ug
Ga-0(1)	13/01/2011	22.4	µg/mL	1.3ug
Mz-0(1)	13/01/2011	15.3	µg/mL	918ng
Ts-5(4)	27/01/2011	20.8	µg/mL	1019.2ng
Sorbo (1)	16/02/2011	16.1	µg/mL	2400ng
Nok-3 (3)	16/02/2011	24.8	µg/mL	3700ng

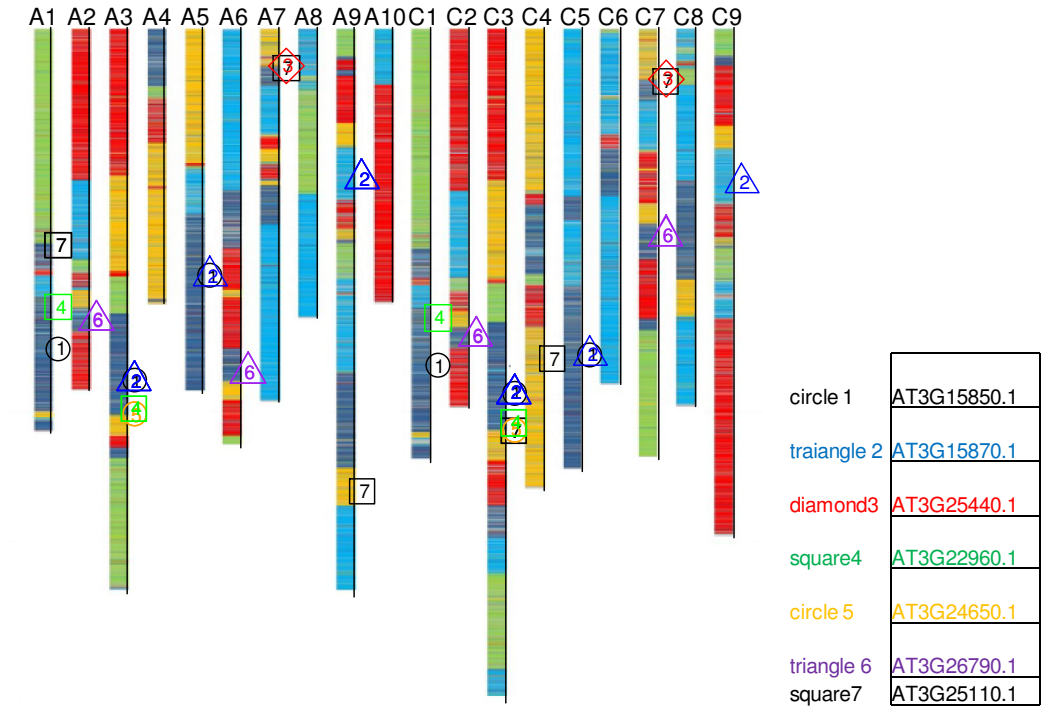
IB Kondara = 384bp @ 33.2nM
IB Gy-0 = 424bp @ 26.3nM
IB Br-0 = 430bp @ 7.6nM
IB Ga-0 = 392bp @ 27.2nM
IB Mz-0 = 407bp @ 4.4nM
IB Sorbo = 476bp @ 66.8nM
IB Nok-3 = 441bp @ 17.5nM
IB Ts-5 = 404bp @ 5.3nM

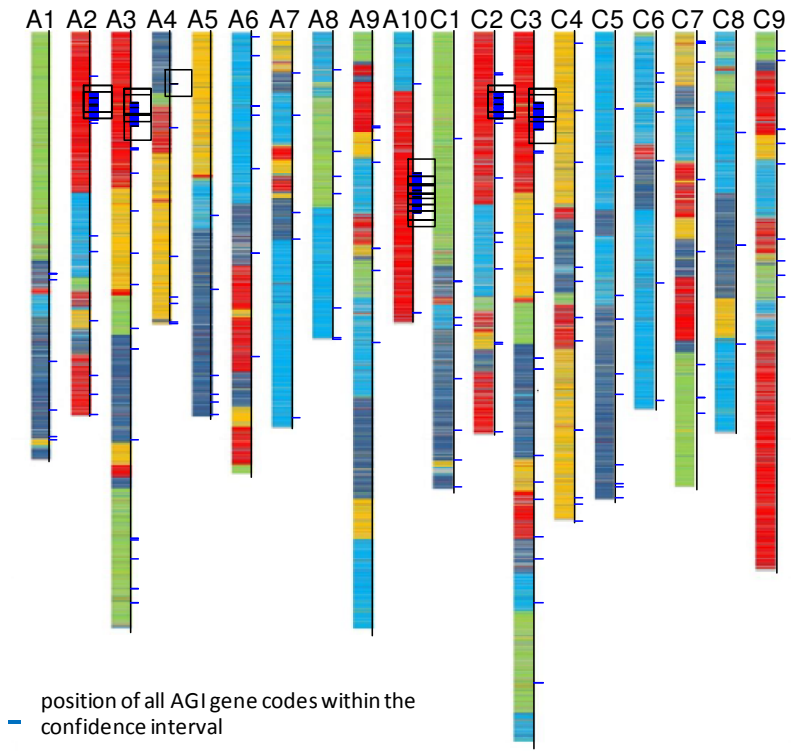
The size of the libraries and concentration were descrived below. The insert size can be calculated by deducting 122bp from the library size.

APPENDIX 10 *A. THALIANA* QTL MAPPED ON UNIGENE BASE *B. NAPUS* MAP

18:2

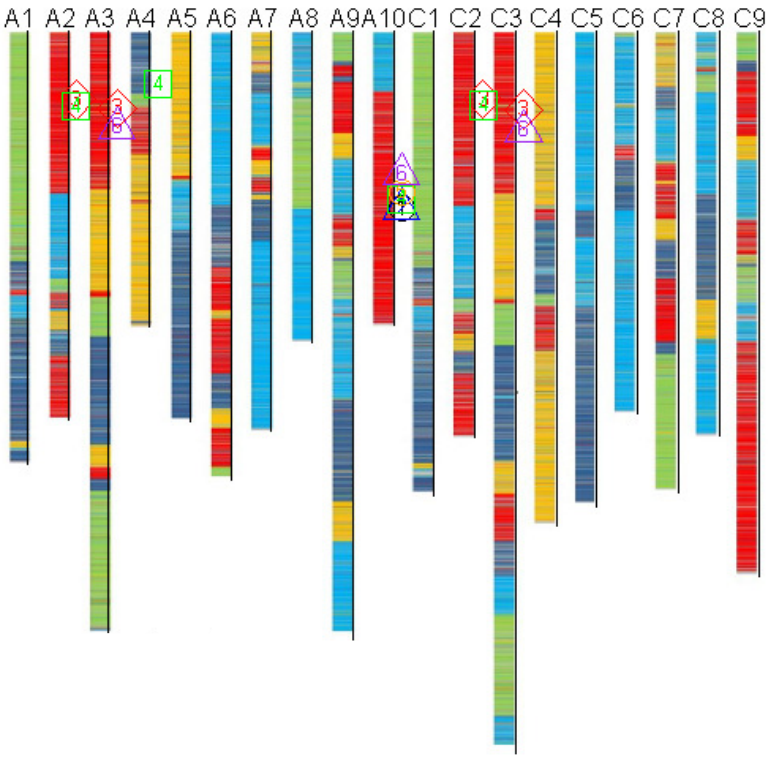


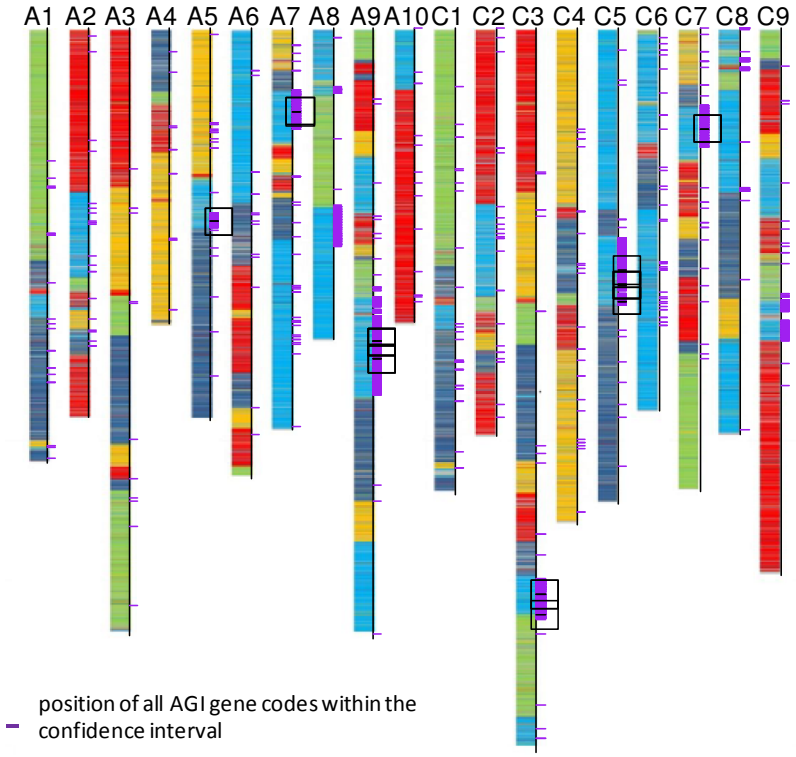




position of all AGI gene codes within the confidence interval

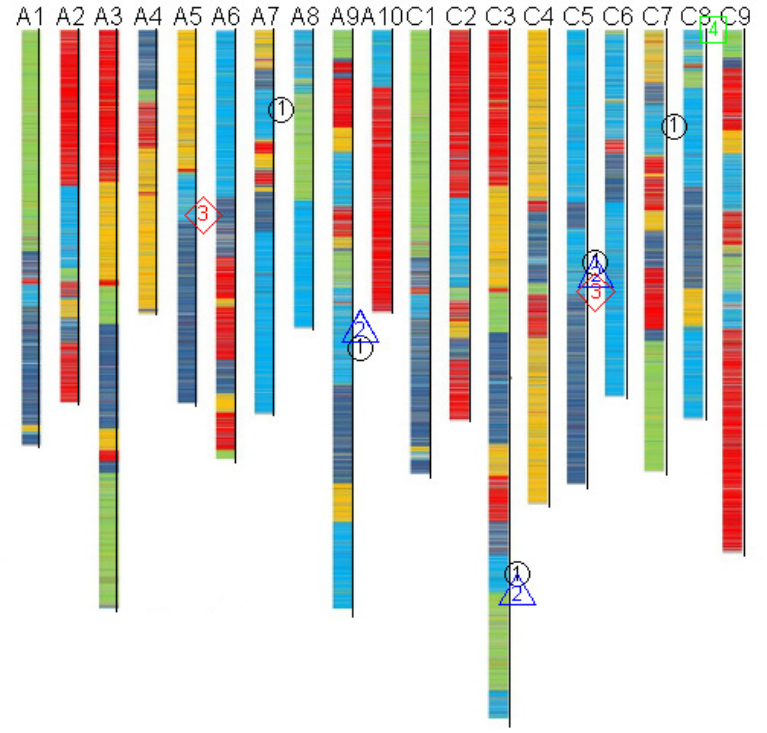
position of AGI gene codes related to lipid biosynthesis and FA desaturation within the confidence interval



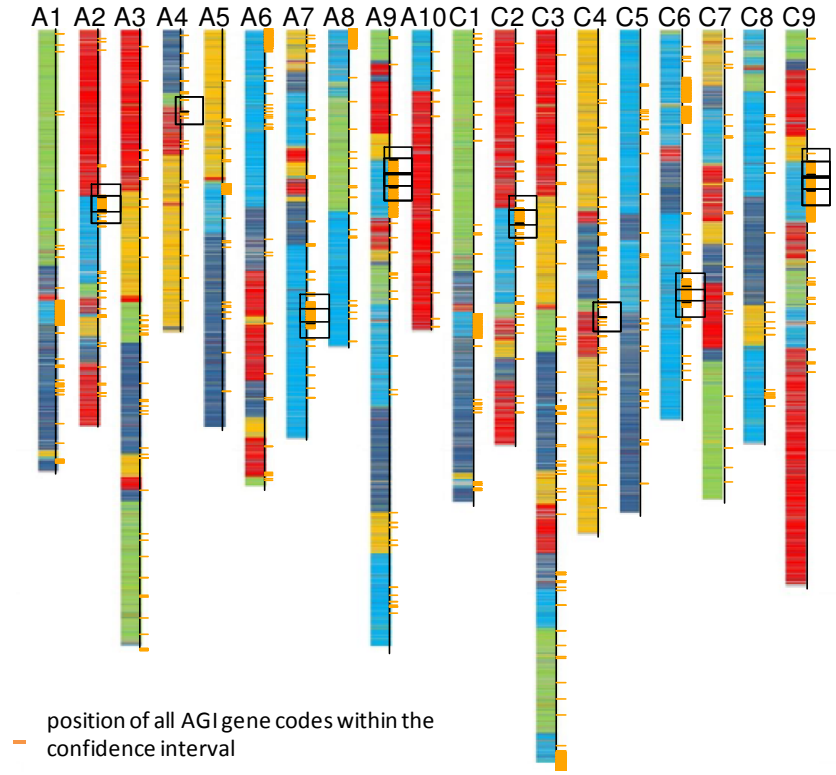


position of all AGI gene codes within the confidence interval

position of AGI gene codes related to lipid biosynthesis and FA desaturation within the confidence interval

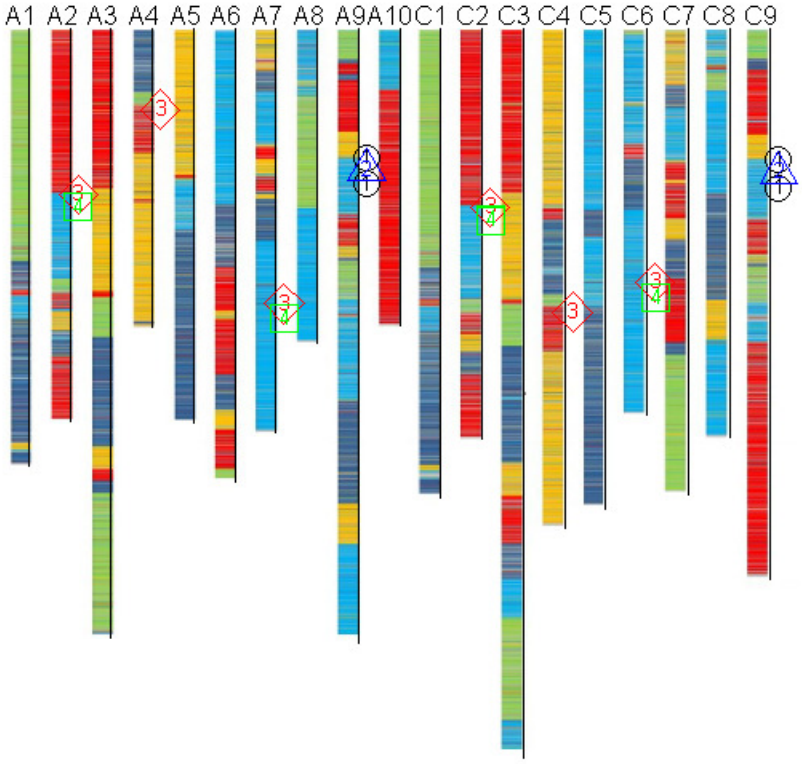


AT1G24360.1
AT1G30120.1
AT1G32440.1



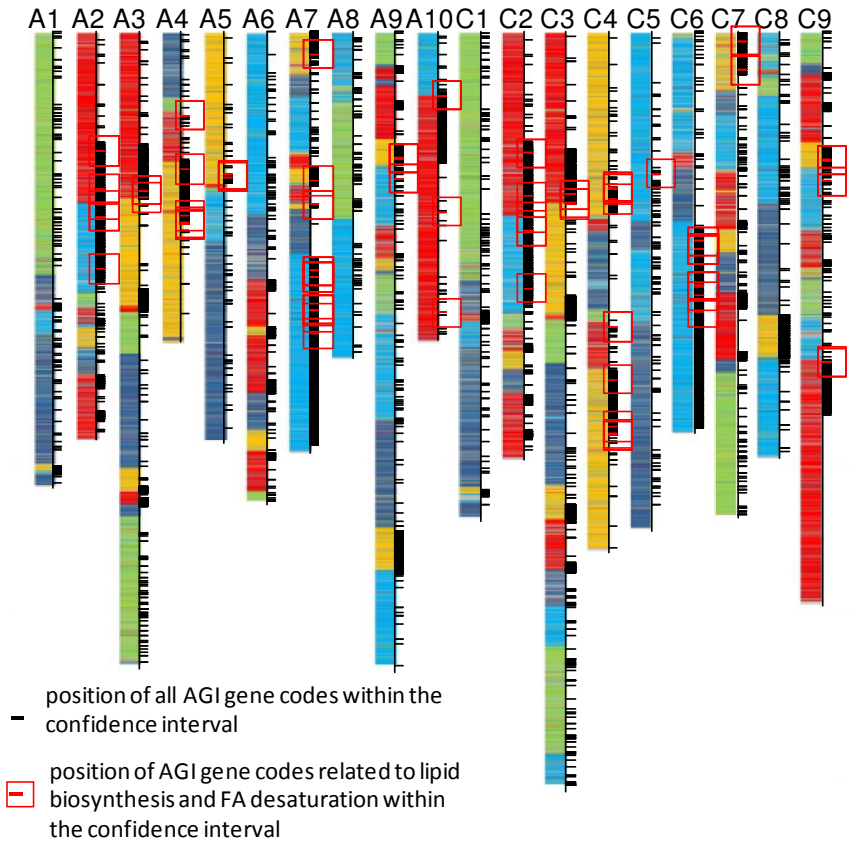
position of all AGI gene codes within the confidence interval

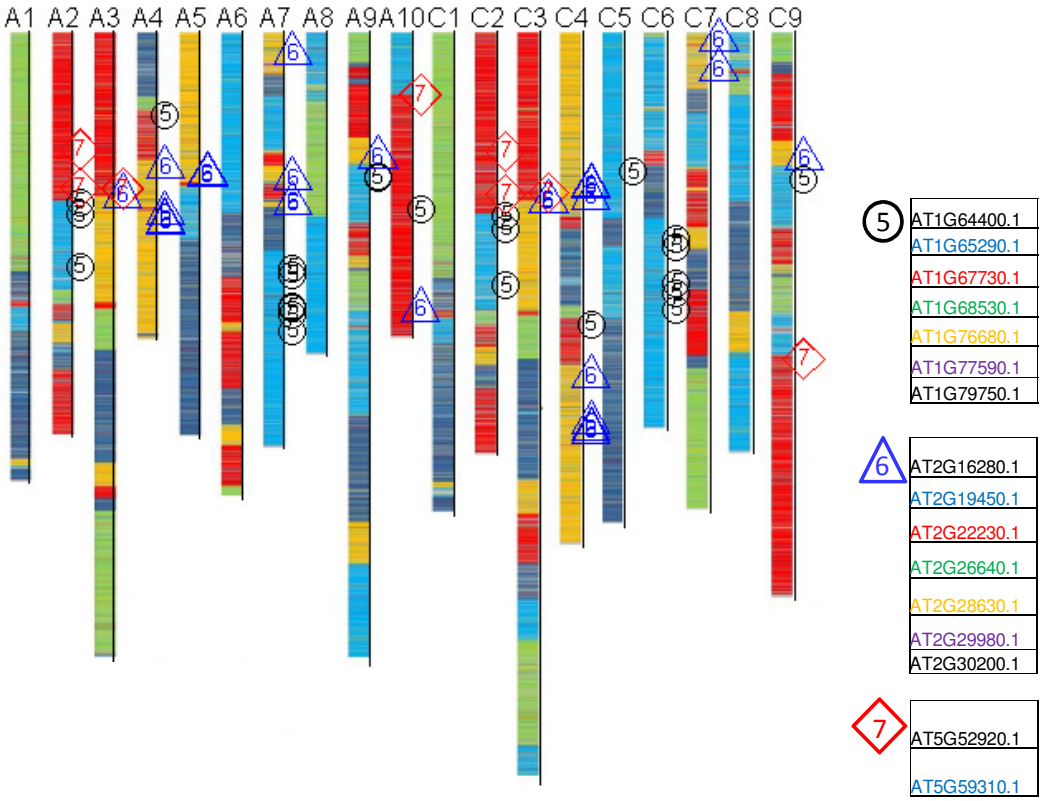
position of AGI gene codes related to lipid biosynthesis and FA desaturation within the confidence interval



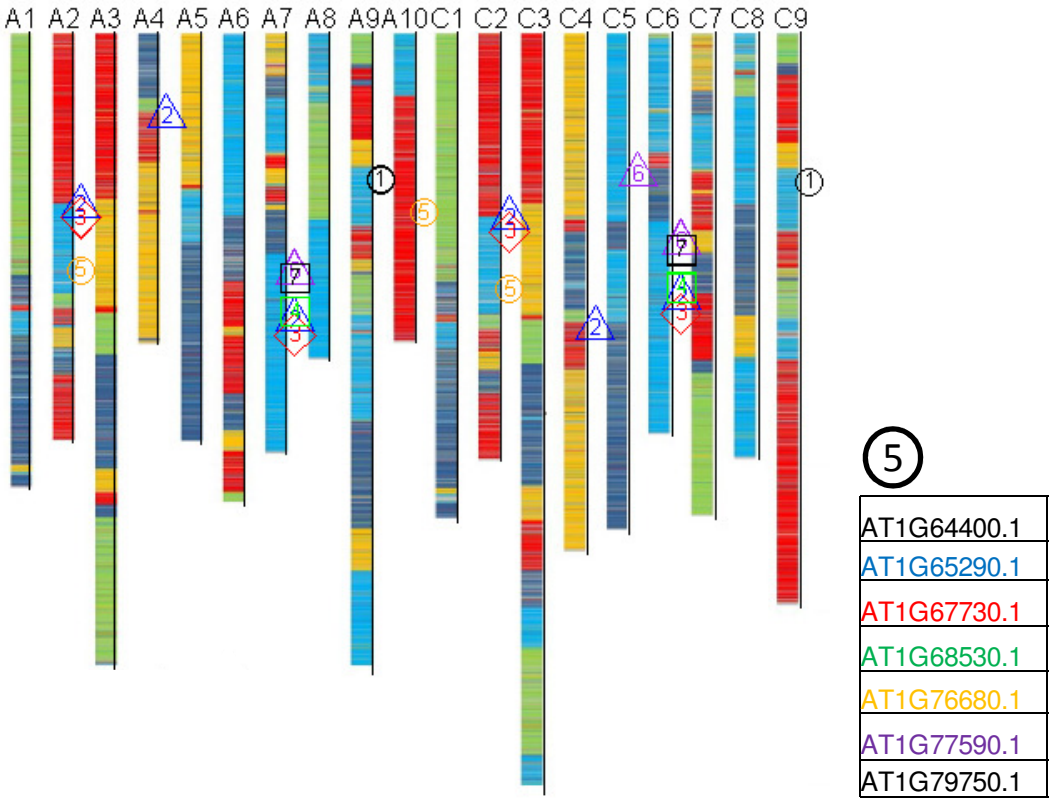
AT1G62640.1
AT1G64400.1
AT1G65290.1
AT1G67730.1

OIL CONTENT (ALL QTL)

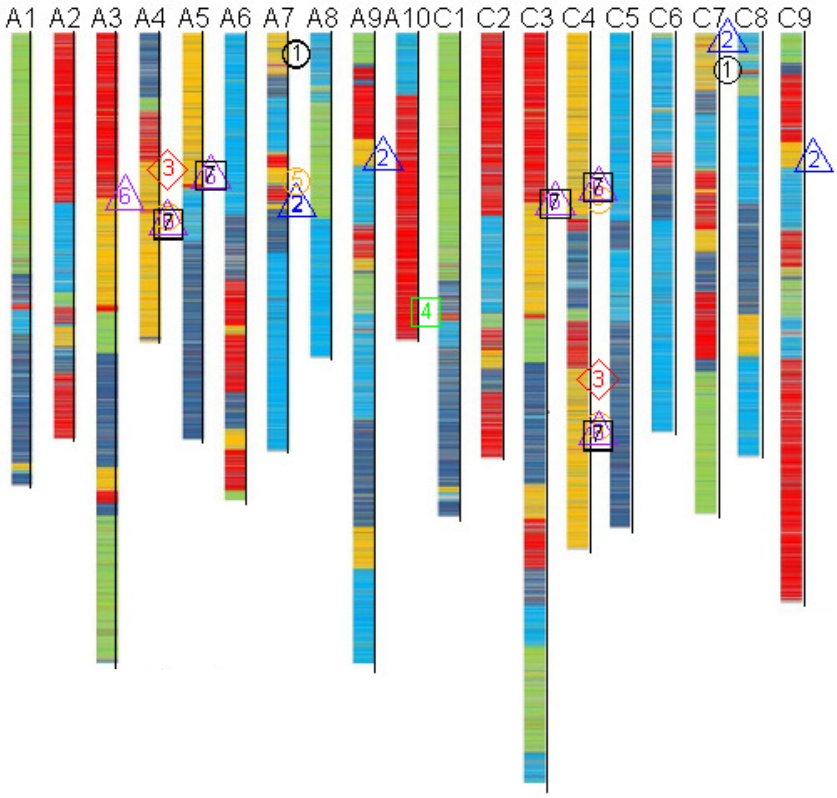




OIL CONTENT QTL ON CH1



OIL CONTENT QTL ON CH2



OIL CONTENT QTL ON CH5