Characterisation of the initial generations of recombinant inbred lines in perennial ryegrass (*Lolium perenne* L.) using molecular markers and cytogenetics.

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Ulrike C.M. Anhalt
Department of Biology
University of Leicester

September 2008

Declaration

I hereby declare that no part of this thesis has been previously submitted to this or any other university as part of the requirements for a higher degree. The content of this thesis is the result of my own work unless otherwise acknowledged in the text or by reference.

The work was conducted in the Department of Biology at the University of Leicester during the period September 2004 to September 2008.

Ulrike C.M. Anhalt, September 2008

Abstract

Characterisation of the initial generations of recombinant inbred lines in perennial ryegrass (*Lolium perenne* L.) using molecular markers and cytogenetics.

Ulrike C.M. Anhalt

In this study three early generations of a recombinant inbred line (RIL) breeding programme have been characterised by cytogenetic methods, metabolite profiling and a biomass quantitative trait loci (QTL) study.

In situ hybridization was used to analyse recombination and the structure of the parental and the F_1 genomes. A metabolite profiling study via GC/MS was conducted to characterise the metabolite activity of the parental inbred lines and the F_1 hybrid. Total metabolites were compared across a growth season in a replicated field design. A genetic map was constructed with 75 nuclear DNA markers in the F_2 generation, which consisted of 360 individuals. 63% of the markers in the F_2 population did not fit the expected Mendelian ratios and showed significant ($p<0.05$) segregation distortion. Fresh weight, dry weight, and dry matter were analysed with an alpha lattice design throughout the greenhouse and field experiments. Additionally, leaf width was recorded in the greenhouse experiment. Heritablities ranged for the biomass related traits between 78 and 95%. This study provides an insight into the recombination of the parental and the F_1 genotypes. Different levels of metabolite activities could be found among the two parental inbred lines across three harvest dates. The QTL study indicates the position of biomass QTL related traits. Major QTL with log of odds scores >3 were identified on linkage groups 2, 3, 4, 5 and 7. About 30% of overall variation could be explained.

Increased biomass yield is still one of the most important traits in any *Lolium perenne* breeding programme. The present QTL study can be used for fine mapping of biomass yield related traits in *L. perenne*. In the long term biomass yield can be eventually monitored and predicted with marker assisted selection for some of the QTL identified in the present study.

Dedicated to my parents
and to my brother Martin
with all my love, for their support, advice and love.

Acknowledgments

I would like to express sincere gratitude to my supervisors Dr. Susanne Barth and Professor Heslop-Harrison for their guidance throughout this study, motivation and for all their help during my professional development.

To Dr. Trude Schwarzacher and Dr. John Bailey for helping me with the molecular cytology work. Dr. Stephen Byrne for the helpful conversations and constructive discussions during the writing-up period.

I acknowledge Professor Piepho at the University of Hohenheim and Professor Oliver Fiehn at the University of California Davis for their corporation.

To John Hogan, The Grass Group in Oak Park, Henry O'Shea, Aurélie Guillard, Jean Liggins, Judy Horner, the administration staff in Oak Park and the clerical staff at the University of Leicester who were all always helpful and willing to assist at any occasion.

A special thanks to my close friends in Ireland Kerstin Diekmann, Bicheng Yang, Céline Tomaszewski, Finbarr Horgan and Criona Murphy for their friendship, moral support and encouragement. You made my time during the PhD so much easier.

To all my friends, especially Montse Vives, Kim and Stephan Hübner, Stefanie Latuske, Marion Mees, Gianluigi Bigio, Teo Chee How and Azhar Mohamad. Thanks for all the laughs and the cheerful times.

A sincere appreciation to Teagasc for the Walsh Fellowship and the National Development Plan for financial contributions.

Abbreviations

AFLP	amplified fragment length polymorphism
ANOVA	analysis of variance
BC	backcross
bp	base pair
BPH	best-parent heterosis
BSA	bovine serum albumin
CAPS	cleaved amplified polymorphic sequence
CIM	composite interval mapping
cM	centiMorgan
cm	centimeter
CMS	cytoplasmic male sterility
CTAB	cetyltrimethylammoniumbromid
DAPI	4',6-diamidino-2-phenylindole
dATP, dCTP, dGTP, dUTP, TTP	nucleotides
DNA	deoxyribonucleic acid
dNTP	nucleotide
EDTA	ethylenediaminetetraacetic acid
F. pratensis	*Festuca pratensis*
FISH	fluorescent *in situ* hybridization
F_n	*filial* hybrid line or genotype
g	gramme
GC	gas chromatography
GISH	genomic *in situ* hybridization
HCl	hydrochloric acid
IM	interval mapping
ISH	*in situ* hybridization
kb	kilobase
L. perenne	*Lolium perenne*
LG	linkage group
LM	liquid chromatography
LOD	logarithm of odds

M	molar
MAS	marker-assisted selection
mg	milligramme
min	minute
ml	milliliter
mM	millimolar
mM	millimolar
MPH	mid-parent heterosis
MQM	multiple QTL model
MS	mass spectrometry
ng	nanogramme
nm	nanomol
NOR	nucleolar organizer
PCA	principal coordinate analysis
PCR	polymerase chain reaction
PFa	paternal line or genotype
PM	maternal line or genotype
QTL	quantitative trait loci
RAPD	random amplified polymorphic DNA
rDNA	ribosomal deoxyribonucleic acid
RFLP	restriction fragment length polymorphism
RIL	recombinant inbred line
RNA	ribonucleic acid
Rnase	ribonuclease
rpm	rotations per minute
S	selfed generation
s	second
SCC	saline sodium citrate
SD	segregation distortion
SDS	sodium dodecyl sulfate
SIM	simple interval mapping
SSC	saline sodium citrate
SSR	simple sequence repeats

U	unit
UPGMA	unweighted pair group method with arithmetic mean
v/v	volume added to volume
w/v	weight added to volume
WSC	water-soluble carbohydrate
µg	microgramme
µl	microliter
µM	micromolar
%	percent
°C	degree Celsius
⊗	selfed

Table of content

TABLE OF FIGURES AND TABLES 3

CHAPTER 1: GENERAL INTRODUCTION 5

CHAPTER 2: *IN SITU* HYBRIDIZATION 10

2.1 INTRODUCTION 11
2.1.1 Fluorescent in situ hybridization 11
2.1.2 Genomic in situ hybridization 12
2.1.3 Chromosomes 13
2.1.4 Probe 13
2.2 MATERIAL AND METHODS 15
2.2.1 Plant material 15
2.2.2 In situ hybridization 16
2.3 RESULTS 21
2.3.1 GISH with parental genomic DNA 21
2.3.2 FISH with 18-25S-rDNA (pTa71) and 5S-rDNA (pTa794) genes 23
2.4 DISCUSSION 27
2.4.1 Genomic in situ hybridization with parental genomic DNA 27
2.4.2 Fluorescent in situ hybridization with 18-25S rDNA (pTa71) and genes 5S rDNA (pTa794): 28

CHAPTER 3: METABOLITE ANALYSIS 31

3.1 INTRODUCTION 32
3.1.1 Pathways 32
3.1.2 Metabolic profiling, target analysis, and fingerprinting 33
3.2 MATERIAL AND METHODS 35
3.2.1 Plant material 35
3.2.2 Homogenisation and extraction 35
3.2.3 Detection 36
3.2.4 Data Analysis 36
3.3 RESULTS 37
3.3.1 Comparing metabolic profile of the genotypes 37
3.3.2 Comparing metabolic profiles across the growing season 48
3.4 DISCUSSION 57

CHAPTER 4: GENETIC MAP 60

4.1. INTRODUCTION 61
4.1.1 Linkage mapping 61
4.1.2 Molecular markers 63
4.1.3 Population structure for mapping studies 65
4.1.4 Breeding and mapping in Lolium perenne 65
4.2 MATERIAL AND METHODS 66
4.2.1 Plant Material 66
4.2.2 Total DNA extraction 67
4.2.3 SSR and AFLP marker testing and analysis 68
4.2.4 Genetic map construction 71
4.2.5 Segregation distortion 72

4.3 RESULTS 72
4.3.1 AFLP and SSR testing and detection *72*
4.3.2 Genetic map construction *76*
4.3.3 Segregation distortion *80*
4.4 DISCUSSION 82
CHAPTER 5: QTL MAPPING FOR BIOMASS 92
5.1 INTRODUCTION 93
5.1.1 Yield *95*
5.1.2 Heterosis *96*
5.1.3 QTL *97*
5.1.4 Genotype x environment interaction *99*
5.1.5 Experimental design - Alpha lattice design *100*
5.1.6 Prospects of QTL cloning *100*
5.2 MATERIAL AND METHODS 101
5.2.1 Plant material *101*
5.2.2 Experimental design - Alpha lattice design *101*
5.2.3 Collection of phenotypic data *103*
5.2.4 Data analysis *104*
5.2.5 Heritability of traits *106*
5.2.6 QTL analysis *106*
5.2.7 Correlation and regression analysis *108*
5.3 RESULTS 109
5.3.1 Plant architecture *109*
5.3.2 Normal distribution of trait data *110*
5.3.3 Heritability *112*
5.3.4 Heading date *112*
5.3.5 Biomass data for each harvest and across harvests *114*
5.3.6 QTL positions for the different traits in the F_2 mapping population *118*
5.3.7 Dry weight *123*
5.3.8. Correlation and regression among traits *127*
5.4 DISCUSSION 130
5.4.1 Heterosis *130*
5.4.2 QTL studies in Lolium perenne *132*
CHAPTER 6: GENERAL CONCLUSION 137
APPENDIX 142
REFERENCES 166

Table of figures and tables

Chapter 1: General introduction

Chapter 2: *In situ* hybridization

Figures:

Figure 1: Development of the parental inbred lines 16
Figure 2: *In situ* hybridization of F_1 chromosomes with parental genomic DNA 22
Figure 3: *In situ* hybridization of maternal, F_1 and paternal chromosomes 24
Figure 4: Karyograms of maternal, F_1 and paternal chromosomes 25

Chapter 3: Metabolite analysis

Figures:

Figure 1: PCA 1 and 2 of maternal, paternal and F_1 genotype (all harvests) 38
Figure 2: Known metabolites activity during harvest times 45
Figure 3: Amino acid pathway with graphs of the measured quantity values 47
Figure 4: PCA 1 and 2 of the maternal, paternal and F_1 genotype (single harvests) 50
Figure 5: Hierarchical clustering of metabolites 52
Figure 6: Relative values of glucose over the three genotypes 55
Figure 7: Relative values of fructose over the three genotypes 56
Figure 8: Relative values of xylose over the three genotypes 56

Tables:

Table 1: Significance of t-test and fold change between the genotypes 39

Chapter 4: Genetic map

Figures:

Figure 1: Breeding scheme for Recombinant Inbred Lines 67
Figure 2: Tall fescue simple sequence repeats - marker testing 75
Figure 3: Simple sequence repeats genotyping between the F_2 genotypes 76
Figure 4: Genetic map 78
Figure 5: Frequencies of each allele in F_2 plants plotted along the linkage groups 81

Tables:

Table 1: Polymerase chain reaction programmes for SSR markers 69
Table 2: Labelled SSR markers 73
Table 3: Table shows the distribution of segregation distortion ratios 80
Table 4: Summary of maps published for pure *Lolium perenne* populations 83

Chapter 5: QTL mapping for biomass

Figures:

Figure 1: Layout of the experiments in the greenhouse and in the field 102
Figure 2: Maternal, paternal and F_1 line plant architecture and biomass production 109
Figure 3: Normal distribution charts 111
Figure 4: Distribution of heading date in the F_2 population in plants in pots 112
Figure 5: Assumed quantitative trait loci locations for heading date in the field 113
Figure 6: Heterosis data, mid-parent heterosis and best-parent heterosis 115
Figure 7: Quantitative trait loci positions on linkage groups 2, 3, 4 and 7 119
Figure 8: Greenhouse and field quantitative trait loci positions for dry weight 123

Tables:

Table 1: Summary of quantitative trait loci studies in *Lolium perenne* 94
Table 2: Broad sense heritabilities for fresh weight, dry weight, dry matter 112
Table 3: Biomass yield, dry matter and heterosis data 115
Table 4: Quantitative trait loci and their locations for all traits 121
Table 5: Quantitative trait loci data for the trait dry weight for each harvest 125
Table 6: Mean, standard deviation, minimum and maximum values for greenhouse 127
Table 7: Regression analysis of dry weight in the greenhouse 128
Table 8: Mean, standard deviation, minimum and maximum values for the field 128
Table 9: Regression analysis of dry weight in the field 129

Chapter 6: General conclusion

Appendix

Appendix Figures:

Appendix Figure 1: Weather data (temperature and rainfall) for 2005, 2006, 2007 151
Appendix Figure 2: QTL for leaf width over all harvests 155
Appendix Figure 3: Fresh weight QTL over all harvests 156
Appendix Figure 4: Dry weight QTL over all harvests 158
Appendix Figure 5: Dry matter QTL over all harvests 160
Appendix Figure 6: Dry weight QTL, LG 2 over single harvests 161
Appendix Figure 7: Dry weight QTL, LG 3 over single harvests 162
Appendix Figure 8: Dry weight QTL, LG 7 over single harvests 164

Appendix Tables:

Appendix Table 1: Metabolite quantitative values 142
Appendix Table 2: Metabolite data of an ANOVA between genotypes 146
Appendix Table 3: Biomass data for greenhouse harvest December 2005 152
Appendix Table 4: Biomass data for greenhouse harvest February 2006 152
Appendix Table 5: Biomass data for greenhouse harvest April 2006 153
Appendix Table 6: Biomass data for field harvest August 2006 153
Appendix Table 7: Biomass data for field harvest October 2006 154
Appendix Table 8: Biomass data for field harvest May 2007 154
Appendix Table 9: Biomass data for field harvest August 2007 154

Chapter 1: General introduction

World hunger and climate change require increased crop yields and the tolerance to environmental influences (like the resistance to diseases and stresses) in crop plants. The need to modify plant architecture, metabolism and the physiology of plants arises rapidly to adapt to new purposes (Vaughan *et al.* 2007). Crops were bred as food or forage plants with the additional use as energy source in the past. Nowadays, the importance of some crops as energy plants, in order to cope with diminishing fuel and energy sources, is increasing. Rapeseed and *Miscantus* are two examples for energy crops. Improvements in crop cultivation in recent years have changed plant breeding into crop engineering in order to meet the needs of humans for increased crop yield. Since early times, hybridization and selection were involved in the work of breeders. Genetic improvement by conventional breeding is often very slow (Yamada *et al.* 2005). In the present, plant breeders have technologies available that enable plants engineering (Vaughan *et al.* 2007). Approaches like quantitative trait loci (QTL) mapping, marker assisted selection (MAS) or the genetic modification of crops are some of the tools used by breeders nowadays to improve crops. Genomic and gene manipulation can complement and enhance conventional breeding programmes (Yamada *et al.* 2005).

The genera *Lolium* and *Festuca* include most of the agriculturally and economically important grasses in the temperate grassland regions (Thomas and Humphreys 1991). Approximately 90% of the farmed land area in Ireland (Connolly 2001) and 76% in the United Kingdom (Humphreys *et al.* 2005) is devoted to grassland. *Lolium perenne* (perennial ryegrass) is a desirable turf grass because of its relative fine texture, quick establishment, good density, and uniformity. *L. perenne* is used for grazing, the production of silage and hay. Numerous breeding varieties have been selected (Hubbard 1992). The major problem of *Lolium* is poor persistency and

susceptibility to various stresses and diseases (Kopecký *et al.* 2005). Species of the genus *Festuca* are known for complementary characteristics to those of *Lolium*. Good persistency and high stress tolerance are the major attributes of *Festuca* species (Thomas *et al.* 2003). Genetically, the chromatin of *Lolium* and *Festuca* can be introgressed (Thomas *et al.* 2003). The complex of species offers an enormous richness of genetic variability and potential for genetic exchange, which gives a strong background for the combination of useful quality characteristics in genotypes for grassland farming (Thomas *et al.* 2003). Breeders have exploited intergeneric *Lolium–Festuca* hybridisation and achieved numerous *Festulolium* cultivars with improved agronomic parameters compared to the parent species. But hybrids between the diploid *Lolium* and *Festuca* species are generally male sterile with low female fertility (Jauhar 1993) which makes breeding not an easy task. The advantage of *Festulolium* is the combination of the positive attributes of both genera: *Lolium* with high forage quality and *Festuca* with high persistence, tolerance to environmental influences, and disease resistance.

Lolium × *Festuca* amphidiploid hybrids are achieved by incorporation of complete genomes of the species into a single genotype. But the success in commercial varieties has been minor (Thomas *et al.* 2003) due to the fact that high level of homoeologous pairing leads to genetic instability and loss of hybridity in later generations (Yamada *et al.* 2005). An alternative approach to amphidiploid breeding is introgression. Introgression is a backcrossing procedure in which transfer of specific characters of one species into another species are achieved by means of hybridization and successive backcrossing to the parental species (Thomas *et al.* 2003). The chromosomes of *Lolium* and *Festuca* can be discriminated using genomic *in situ* hybridization (GISH). GISH is used to identify the introgressed segment of an alien

chromosome and to determine the location on the chromosome (Yamada *et al.* 2005). By combining GISH to identify plants that carry the introgressed segment with the association of molecular markers to these alien chromosome segments, it is possible to 'tag' genes responsible for the desired agronomic traits in a breeding programme (King *et al.* 1998). The aim is to introgress a limited number of alien genes into the genome of the target species by recombination and selection. The introgression-mapping approach offers researchers a tool to locate those areas on the chromosomes that are carrying the genes which determine complex traits. A strong phenotypic segregation, which is common for androgenic *Lolium/Festuca* populations, can assist breeders in selecting according to desirable agronomic traits and can be tagged in relation to closely flanking markers (Thomas *et al.* 2003). This would give marker MAS the opportunity to develop faster high quality grasses compared to conventional breeding procedures. The construction of dense genetic maps enables the transfer of genetic information from one population or even from one species to another by marker information (Thomas *et al.* 2003). MAS technologies bring precision to plant breeding and target gene combinations that control specific agronomic traits (Yamada *et al.* 2005).

Conventional breeding has already been successful in generating commercial varieties of forage grasses with traits for enhanced agricultural sustainability. Breeding objectives focus on stress resistance against drought, cold and pathogens, and on agronomic traits like nutrient use efficiency, carbohydrate content, fatty acid content, winter survival, flowering time and biomass yield (Humphreys *et al.* 2005).

The perennial ryegrass breeding programme in Oak Park focuses on varieties with improved yield and persistency to improve the productive potential of Irish grasslands. Hybrid breeding programmes are in employed. Different germplasm sources have been used in the breeding programmes: a) existing varieties developed in New

Zealand and Europe, b) gene bank collections from various climatic regions, and c) collections of germplasm in Ireland from old pastures (Connolly 2001). In order to provide control of pollination in the hybrid breeding programme genotypes were developed by Dr. V. Connolly as part of a cytoplasmic male sterility (CMS) programme. Maintainer lines in this CMS programme were originated from an inter-specific cross between meadow fescue (*Festuca pratensis*) and *L. perenne* (Connolly and Wright-Turner 1984). The initial interspecific hybrid was backcrossed for several generations to the ryegrass parent and selfed for more than ten generations. Two of these lines were used to develop the F_2 mapping population used in the present study.

The aims of this study were (1) to characterise the parental lines and initial generations for the construction of recombinant inbred lines at the molecular, molecular cytogenetic and phenotypic level, (2) to develop a genetic map of the present F_2 mapping population and (3) to find QTL for biomass related traits. In the present study, early generations of a recombinant inbred line breeding programme have been characterised using cytogenetic methods to analyse recombination and structure of the genomes (chapter 2). A metabolite profiling study was carried out in order to characterise the metabolite phenotype of the parental lines and their F_1 (chapter 3). An F_2 mapping population and a genetic map were developed (chapter 4) and agronomic traits, especially yield related traits, were studied in QTL studies (chapter 5). These findings will be of interest in a later stage for the breeding programme in Oak Park using MAS and might provide the ryegrass biotechnology and the grass breeding programme in Oak Park with a valuable mapping resource.

Chapter 2: *In situ* hybridization

2.1 Introduction

In situ hybridization (ISH) was first described by Pardue and Gall (1969) and in parallel by John *et al.* (1969) using *Drosophila* chromosomes as target and radiolabelled ribosomal DNA (rDNA) as probes. After the invention of the radioactive labelling method DNA and RNA blotting hybridization methods were developed, followed by fluorescent labels. ISH is used in plants for many purposes, including analysis of chromosomal rearrangements, to analyse cross-species chromosome homology, gene mapping or for introgression studies (Schwarzacher and Heslop-Harrison 2000).

Early chromosome studies in *Lolium* used the Giemsa C-banding technique to analyse different *Lolium* species and the *Lolium* genome (Malik and Thomas 1966, Thomas 1981). Fluorescent *in situ* hybridization (FISH) and genomic *in situ* hybridization (GISH) were applied in *Lolium* for physical mapping (Thomas *et al.* 1996, King *et al.* 2002), to analyse genome relationships between *Festuca* and *Lolium* species (Pašakinskienė *et al.* 1998; King *et al.* 2007), and to characterise *Lolium*/*Festuca* crosses (Humphreys *et al.* 1998a, Canter *et al.* 1999, Zwierzykowski *et al.* 2006).

2.1.1 Fluorescent in situ hybridization

FISH is an excellent tool for genetic analysis. FISH has direct and indirect detection techniques. Direct detection techniques make use of direct fluorophore-labelled dUTPs for labelling the probe and no immunohistochemistry is required for visualization (Schwarzacher and Heslop-Harrison 2000). In the indirect detection technique the probe is labelled with biotin or digoxigenin and detected by avidin, antibodies, avidin-streptavidin or anti-digoxigenin. The antibodies and avidins are

conjugated to the fluorophores. Fluorescent labelling has several advantages to other detection methods like higher sensitivity and stability, the method is much faster and safer, and multiple labels in multiple colours can be applied (Schwarzacher and Heslop-Harrison 2000). ISH technique provides the opportunity to use large or shorter sequences as probes using bacterial artificial chromosomes or yeast artificial chromosomes. The number of ribosomal genes (rDNA) may vary between species. Therefore, rDNA is suitable to analyse genome organisation or recombination events. FISH is also used nowadays for karyotyping (Thomas *et al.* 1996, Jiang and Gill 2006) of chromosomes by identifying chromosomes via chromosome-specific hybridization signals. An early approach for karyotyping on *Lolium* was performed by Malik and Thomas (1966) via Giemsa C-Banding.

2.1.2 Genomic in situ hybridization

GISH is a technique that can be used to identify the existence and size of segments of alien chromosomes using genomic DNA as probe. GISH was established in plants by Schwarzacher *et al.* (1989). In this study parental DNA of *Secale* and *Hordeum* as GISH probe were hybridised on F_1 chromosomes. Since then, GISH helps us to understand genomic relationships, to solve taxonomic problems, and to detect the presence of alien genomes in plant germplasm. It is possible with GISH to discriminate between genomes of many species in plants, because the chromosomes may be distinguished on the basis of divergent dispersed repetitive sequences as demonstrated in maize (Poggio *et al.* 2005). GISH was used in *Lolium* by different research groups to identify interspecific genomic rearrangements in *Festulolium* hybrids (Humphreys *et al.* 1998a, King *et al.* 1998 and 1999), hybrid genome stability over several generations in

Festulolium hybrids (Zwierzykowski *et al.* 2006) or genome relationships between *Festuca* and *Lolium* species (Pašakinskienė *et al.* 1998).

2.1.3 Chromosomes

Chromosome spreads are used for ISH. Chromosomes of different species differ in number, size, organisation, and the composition of their DNA. All these features are matter to evolutionary changes. The shape of monocentric chromosomes is determined by the location of the primary constriction. The centromere subdivides the chromosome into elements of monomorphic structure. A nucleolus organizing region (NOR), the site of 18S-25S rDNA, may help to identify chromosome either at a terminal or an interstitial position and giving rise to a secondary constriction and a distal 'satellite' (Schubert 2007). Repeated sequences could influence classification through their secondary folding structure, protein binding sites, and the condensation patterns that gives the chromosome the characteristic shape (Schwarzacher 2003). To visualise chromosome structure DNA stains like 4',6-diamidino-2-phenylindole (DAPI) are used in ISH.

2.1.4 Probe

DNA and RNA are suitable as probes for ISH. Nick translation, polymerase chain reaction (PCR) labelling or random primer labelling are some methods to label DNA probes for ISH standard protocols (Schwarzacher and Heslop-Harrison 2000). After purification of the probe from enzymes and unincorporated nucleotides, the labelled DNA is then hybridized onto metaphase and interphase cells.

The rate and the extent of hybridization can be determined by formamide, temperature and other influences like the amount of blocking DNA, the pH or hybridization time. To obtain clear, strong hybridization signals it is necessary to optimize the chromosome preparation, the probe and blocking concentration (blocking DNA: unlabelled genomic DNA used to prevent hybridisation of probe to DNA sequences in common between the probe and the target), the stringency (to limit cross-hybridization to undesired targets), the denaturation time and denaturation temperature and the counterstaining. All these steps are dependent on the species, the stage of the cell cycle, and the tissue and have to be individually optimised. Problems can occur in different ways. Low quality chromosome preparations, incorrect denaturation of preparation, poor probe labelling, use of old or contaminated solutions and materials but also incorrect microscope setup can result in no or weak hybridization signals (Schwarzacher and Heslop-Harrison 2000).

rDNA may be used as a probe and consists of motifs that are 2 to 10,000 base pairs long and are repeated many hundreds or thousands of times. 18S-25S and 5S rDNAs are represented as tandemly repeated units of genes and intergenic spacers (Campos Severi-Aguiar and Azeredo-Oliveira 2005). rDNA sites were detected in several studies in *Lolium* (Thomas *et al.* 1996, Pašakinskienė *et al.* 1998, Warpeha *et al.* 1998). They were used as landmarks to analyse chromosome construction and rearrangements and investigation of species relationships.

The aims of the present study was to use FISH and GISH to achieve a better understanding of the genome organisation and recombination in the RIL development programme to improve the background knowledge of the experimental parental, and F_1 lines, of the F_2 mapping population.

2.2 Material and Methods

2.2.1 Plant material

Lolium perenne used in the present study is diploid with seven chromosomes in the haploid set. The parental inbred ryegrass lines, which were used for the construction of an F_2 mapping population (described later in chapter four) originated from an inter-specific cross between meadow fescue (*Festuca pratensis*) and perennial ryegrass (*L. perenne*) (Connolly and Wright-Turner 1984). The inbred lines were developed by Dr. V. Connolly as part of a cytoplasmic male sterility (CMS) programme in Teagasc, Oak Park and were maintainer lines in the CMS programme (Connolly and Wright-Turner 1984). For the maternal parent (PM) of the inbred lines of the F_2 mapping population a *Festuca partensis* cultivar (P_1) was used as maternal and the ryegrass cultivar 'S24' (IGER) (P_{2A}) was used as paternal plant in the pedigree while for the paternal parent (PFa) of the inbred lines the ryegrass cultivar 'Premo' (Mommersteeg International BV) (P_{2B}) was chosen as paternal plant (Figure 1). The initial interspecific hybrid was backcrossed (BC) for several generations to the ryegrass parent and selfed (S) for nine or ten generations (Figure 1).

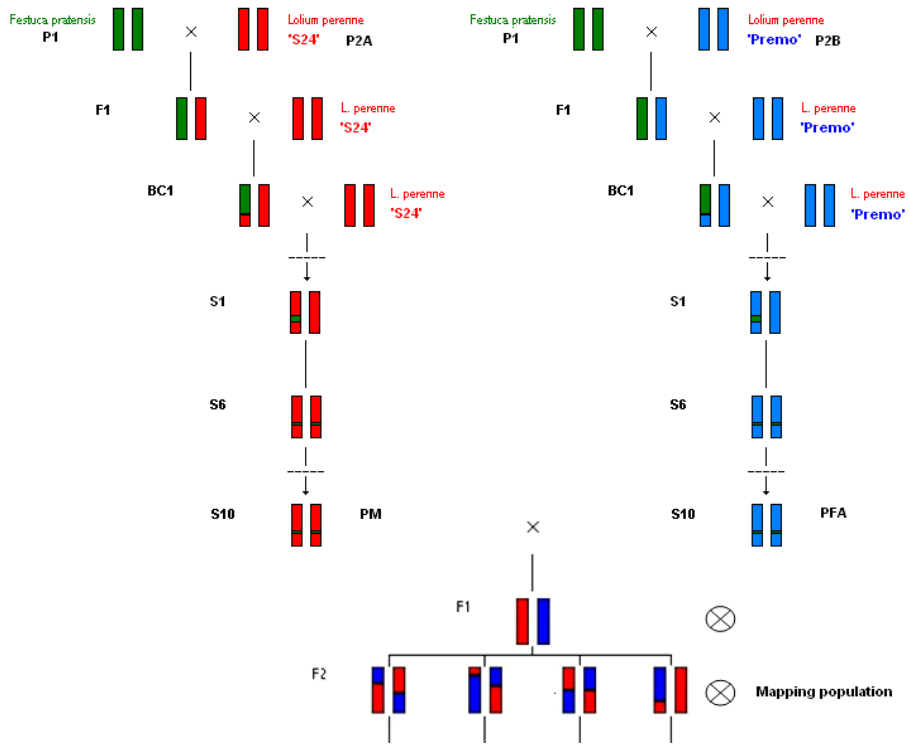


Figure 1: Development of a parental inbred lines (PM, PFA) of the F_2 mapping population. BC: backcross generation, S: selfed generation, ⊗ : selfed.

For the construction of the F_2 mapping population, the maternal parent, 'PM', was emasculated under a binocular microscope and stigmas were pollinated with pollen from the paternal plant, 'PFa'. Pollinated florets were bagged in cellophane bags and individual F_1 seed was harvested.

2.2.2 In situ hybridization

GISH and FISH were used to characterise the initial generations of the F_2 mapping population (parental lines (PM, Pfa) and F_1 genotype, figure 1). To gain insights in the genome organisation and chromosomal structures of the lines repetitive DNA segments were applied. ISH was used following the protocol of Schwarzacher and Heslop-Harrison (2000) with some modifications.

Preparation of chromosome spreads:

For the preparation of mitotic root tip spreads, root tips were treated for 24 h in ice-cold water and were fixed in 3:1 (v/v) 100% ethanol:glacial acetic acid solution. The root tips were enzymatically digested (pectinase and cellulase, Sigma-Aldrich, Taufkirchen, Germany) and meristematic cells were squashed in 75% acetic acid on a glass slide.

Preparation of the probe:

To prepare the 18S-25S rDNA genes as a probe pTa71 was cut with *Hin*dIII and for the 5S rDNA genes pTa794 was amplified by PCR with M13 primers.
(M13R: 19-mer reverse sequencing primer (-21): GGAAACAGCTATGACCATG, M13F: 17-mer sequencing primer (-20): TGACCGGCAGCAAAATG).

Digestion of pTa71 and pTa794:

DNA	(35 ng)	17	µl
Enzyme buffer	(10x)	2	µl
*Hind*III enzyme	(10 U)	1	µl

Reaction was incubated at 37ºC for 3-10 h.

PCR amplification with M13 primers:

DNA	(20 ng)	12	µl
$MgCl_2$	(50 mM)	0.5	µl
NH_4 buffer	(10x)	1.5	µl
dNTPs	(5 mM)	0.3	µl
M13 forward	(10 µM)	0.3	µl
M13 reverse	(10 µM)	0.3	µl
Taq polymerase (BioLabs, New England)	(0.5 U)	0.1	µl
Total:		15	µl

Thermal cycler programme for PCR amplification with M13 primers:

94ºC	5 mins	Initial denaturation
94ºC	30 secs	
45ºC	30 secs	35 cycles
72ºC	1 min	
72ºC	10 mins	Final extension

Both rDNA probes were labelled with the random primer labelling method.

Genomic parental DNA was sheared by autoclaving at 100ºC for 2 minutes to break DNA down to fragments of a size of around 300 bp. Genomic DNA of the parental lines, 18-25S rDNA (pTa71) and 5S rDNA (pTa794) genes were labelled by PCR random primer labelling with digoxigenin-16-dUTP (Roche, Basel, Switzerland) and biotin-11-dUTP (Roche, Basel, Switzerland).

PCR random primer labelling for the genomic DNA and rDNA probes:

9 µl DNA (35 ng) was denaturated in boiling water bath for 10 min.

Reagents were added as followed:

unlabelled dNTPs	(0.5 mM dCTP 0.5 mM dGTP 0.5 mM dATP in 100 mM Tris-HCl, pH 7.5)	3 µl
labelled dUTP mix	(Dig: 0.2 mM Dig-dUTP 0.4 mM TTP in 100 mM Tris-HCl, pH 7.5 or Bio: 0.4 mM Bio-dUTP 0.2 mM TTP in 100 mM Tris-HCl, pH 7.5)	5 µl
Hexanucleotide buffer (Roche)	(10 x)	2 µl
Klenow enzyme solution		1 µl

Reaction was incubated at 37ºC for 6-16 hours.
EDTA 2 µl stop solution (200 mM EDTA, pH 8) was added and the labelled DNA was precipitated with ethanol.

In situ hybridisation:

The slides were re-fixed, especially after a long storing period, and cleaned using 3:1 (v/v) 100% ethanol:glacial acetic acid solution for ten to 30 minutes. Slides were transferred to 100% ethanol for two times five minutes and then air-dried.

For the RNase treatment 200 µl of RNase solution (100 µg/ml) was applied to every slide and incubated for one hour at 37ºC in a humid chamber. Slides were washed afterwards with 2x saline sodium citrate (SSC) for two and then for ten minutes. For the pepsin treatment slides were incubated in 0.01M hydrochloric acid (HCl) for two minutes. 200 µl pepsin solution (5 µg/ml) was applied to each slide and incubated at 37ºC for ten minutes in a humid chamber. Afterwards slides were rinsed in distilled water and washed in 2x SSC for five minutes. For the paraformaldehyde fixation the slides were incubated for ten minutes in 0.05 g/ml paraformaldehyde solution at room temperature. Slides were washed in 2x SSC for two and ten minutes and afterwards dried in 70%, 85% and 100% ethanol each time for two minutes. The DNA probe was denatured by boiling in a water bath for 10 minutes and stored on ice. The hybridization mixture was applied to each slide and contained 4 µl unlabelled genomic parental DNA (Blocking: 1:40. 200ng genomic DNA as probe : blocking DNA (other parental DNA, not used as probe DNA)) and 2 µl labelled pTa71 or pTa794 probe in 40 µl total hybridization solution with 50% (v/v) formamide, 2x SSC, 10% (w/v) dextran sulphate, 1 µg/µl salmon sperm DNA, 0.125 mM EDTA and 0.125% SDS. Chromosomes and probes were denatured together for eight minutes at 75°C and hybridization was carried out overnight at 37°C. After the overnight incubation, slides were washed in 2x SSC for two minutes at 42ºC. Afterwards stringent washes were performed. High stringency wash baths contained 20% formamide and 0.1x SSC. Slides were incubated two times for five minutes at 42ºC and washed in 2x SSC for five minutes. Slides were transferred

into detection buffer (4x SSC and 0.2% Tween20) for 5 minutes. 200 µl of blocking solution (5% bovine serum albumin (BSA) in detection buffer) was applied to each slide and incubated at room temperature for five to 30 minutes. For the detection of signals the slides were incubated in 0.1 µg/ml Alexa Fluor 594 streptavidin (S11227, Invitrogen, Carlsbad, USA) and Anti-digoxigenin-flourescein (CAT No. 11207741910, Roche, Basel, Switzerland) in 5% (w/v) BSA in 4x SSC solution (containing 0.2% (v/v) Tween20). Afterwards slides were washed in detection buffer for two minutes and two times ten minutes at 40ºC. Chromosomes were counterstained with 4µg/ml DAPI (Fluka, Sigma-Aldrich, Buchs, Switzerland) and mounted in antifade AF1 (Citifluor, London, UK). Hybridization signals were visualized with an Zeiss Axioskop 2 *plus* epifluorescence microscope (Carl Zeiss AG, Oberkochen, Germany) with filter blocks for DAPI, FITC, and Alexa 594 and the computer programme AxioVersion 4.3 (Carl Zeiss AG, Oberkochen, Germany). The pictures were edited in Adobe Photoshop CS2 (Adobe Systems Incorporated, San Jose, California, USA.).

2.3 Results

2.3.1 GISH with parental genomic DNA

F_1 chromosomes of the F_2 mapping population were hybridized with maternal (PM) and paternal (Pfa) DNA probes to identify non-recombined blocks of either of the parental parents on the F_1 chromosomes. Figure 2 shows six different root tip squashes of the F_1 chromosomes (a) all hybridization signals and (b) with only the biotin-11-dUTP labelled genomic DNA probe.

GISH on the F_1 chromosomes could not identify any distinct regions pointing towards larger non-recombined blocks of one or the other parent. The entire chromosomes were evenly hybridized and no distinct differences could be found (Figure 2: 1 b - 6 b).

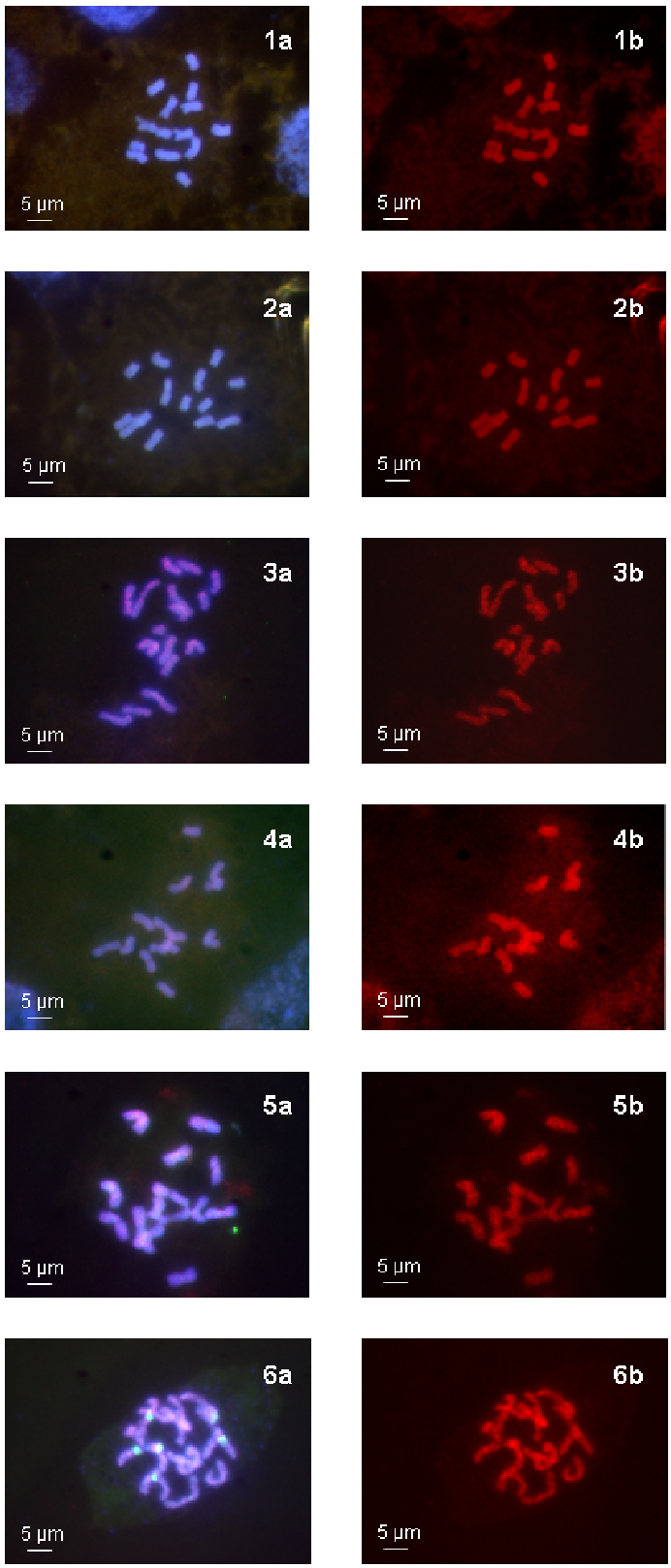


Figure 2: I*n situ* hybridization of F_1 chromosomes of the F_2 mapping population with parental genomic DNA (Pfa) of the F_2 mapping population as probe: F_1 chromosomes with 4',6-diamidino-2-phenylindole (DAPI) (blue), biotin-11-dUTP (red). 1)-4) maternal genomic DNA (PM) of the F_2 mapping population (red); 5) PM (green) and Pfa (red); 6) Pfa (red) and 18-25S rDNA (digoxigenin-16-dUTP (green)).

2.3.2 FISH with 18-25S-rDNA (pTa71) and 5S-rDNA (pTa794) genes

Two different rDNA probes (18-25S rDNA (pTa71) and 5S rDNA (pTa794)) were used to find rearrangements on the F_1 chromosomes compared to the maternal (PM) and paternal (Pfa) chromosomes and to analyse their genom organisation. The different signal pattern is shown in figures 3 and 4 for the parental and F_1 chromosomes.

The 18-25S rDNA landmarks (Figure 3: 1 a c, 2 a c d f, 3 a c d f, Figure 3) showed on the F_1 chromosomes seven bands, on the maternal chromosomes six and on the paternal chromosomes eight signals. 5S rDNA landmarks showed two hybridization signals on all three lines (Figure 3: 1 a b, 2 a b, 3 a b, Figure 4). The 5S rDNA landmarks were located on the same chromosomes but on the more distal ends together with one of the 18-25S rDNA landmarks.

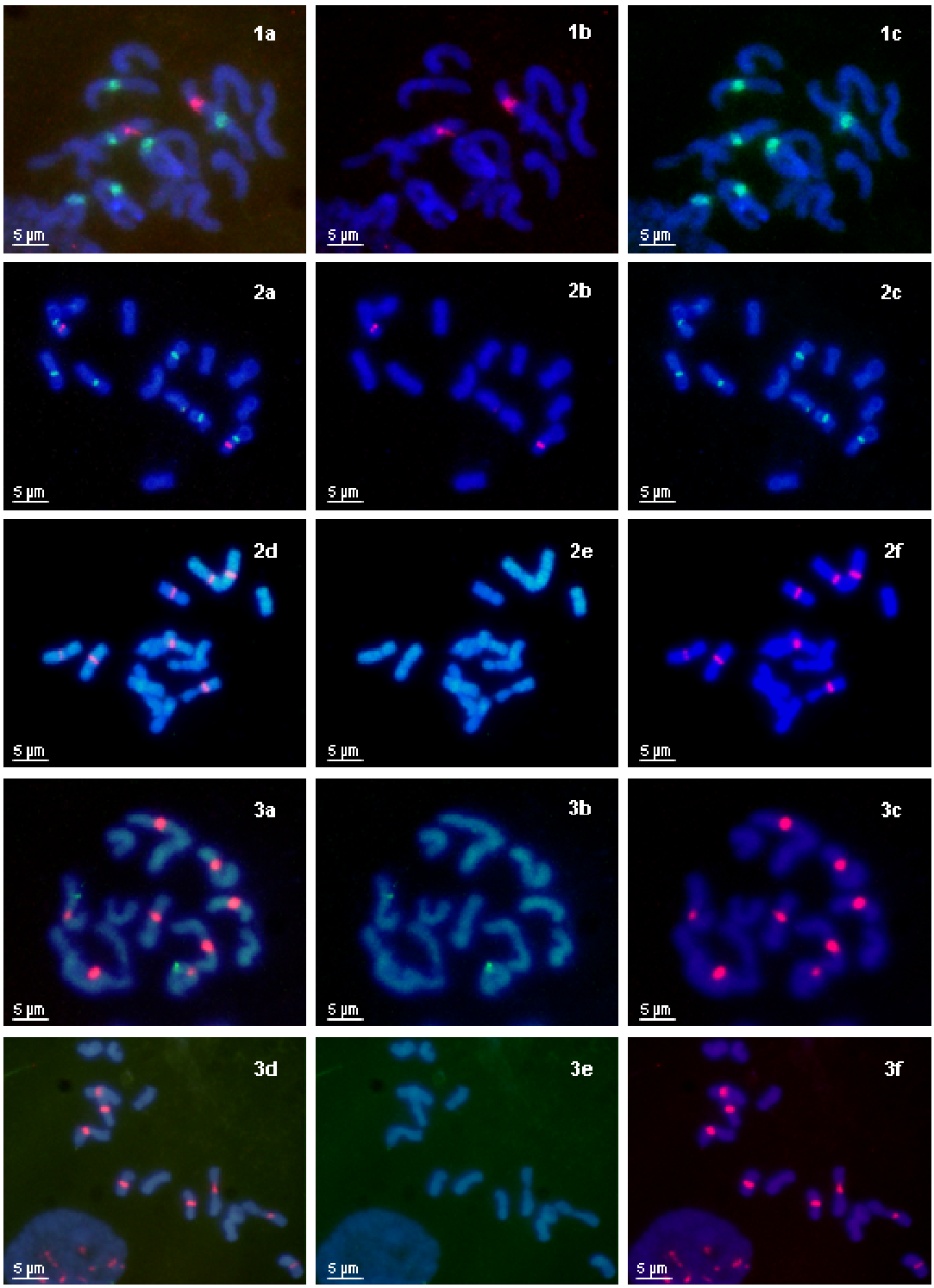


Figure 3: *In situ* hybridization of (1) maternal (PM), (2) F_1 and (3) paternal (Pfa) chromosomes labelled with 4',6-diamidino-2-phenylindole (DAPI) (blue), biotin-11-dUTP (red) and digoxigenin-16-dUTP (green). Pictures show: 1a): 5S (red) and 18-25S rDNA (green); 1b): 5S rDNA (red); 1c): 18-25S rDNA (green); 2a): 5S (red) and 18-25S rDNA (green); 2b): 5S rDNA (red); 2c): 18-25S rDNA (green); 2d): 18-25S rDNA (red) and paternal genomic DNA (green); 2e): paternal genomic DNA (green); 2f): 18-25S rDNA (red); 3a): 18-25S rDNA (red) and 5S (green); 3b): 5S (green); 3c): 18-25S rDNA (red); 3d): 18-25S rDNA (red) and parental genomic DNA (green); 3e): parental genomic DNA (green); 3f): 18-25S rDNA (red).

Karyograms (Figure 4) of the parental and F_1 chromosomes were assembled to display the results of the hybridizations clearer. Karyograms were assembled from Figures 3: 1 a, 2 a, and 3 a.

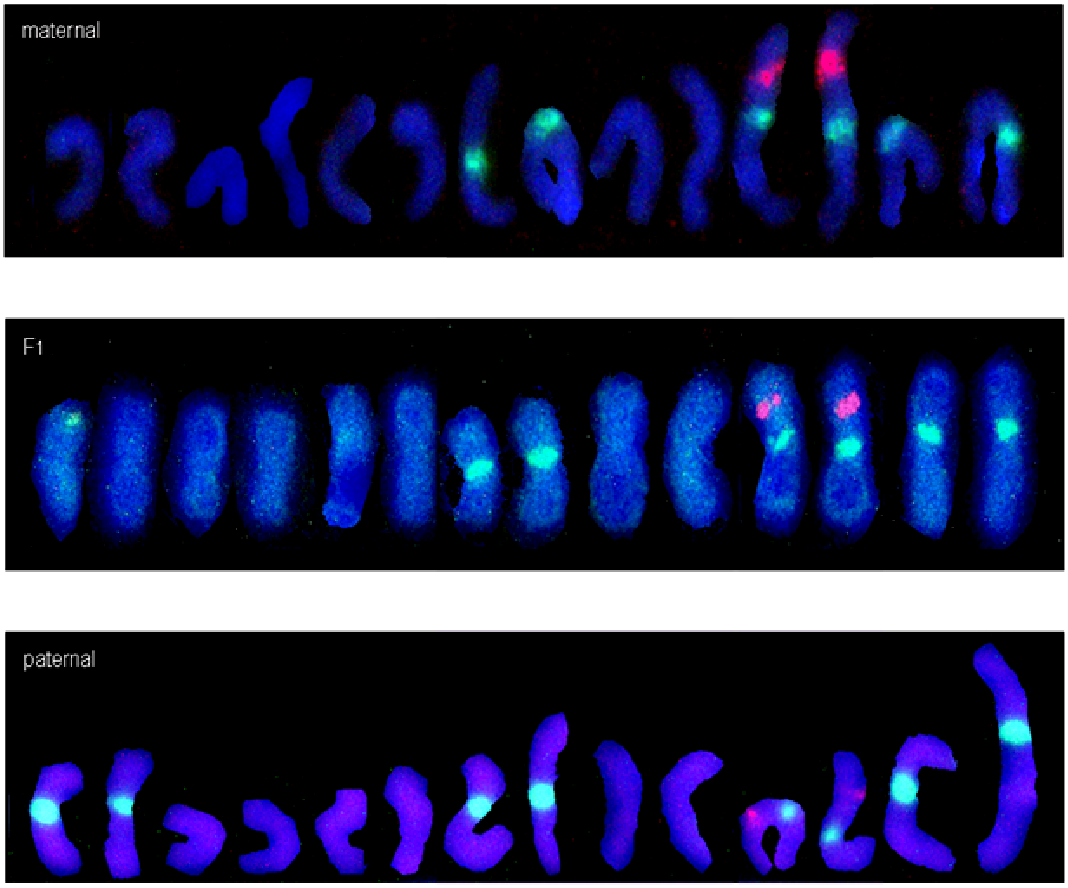


Figure 4 a: Karyograms of maternal (PM), F_1 and paternal (Pfa) chromosomes labelled with 4',6-diamidino-2-phenylindole (DAPI) (blue). Green signals are 18-25S rDNA and red signals are 5S rDNA probes. (Colours reversed from figure 3: 3 a for parental chromosomes to make comparison clearer).

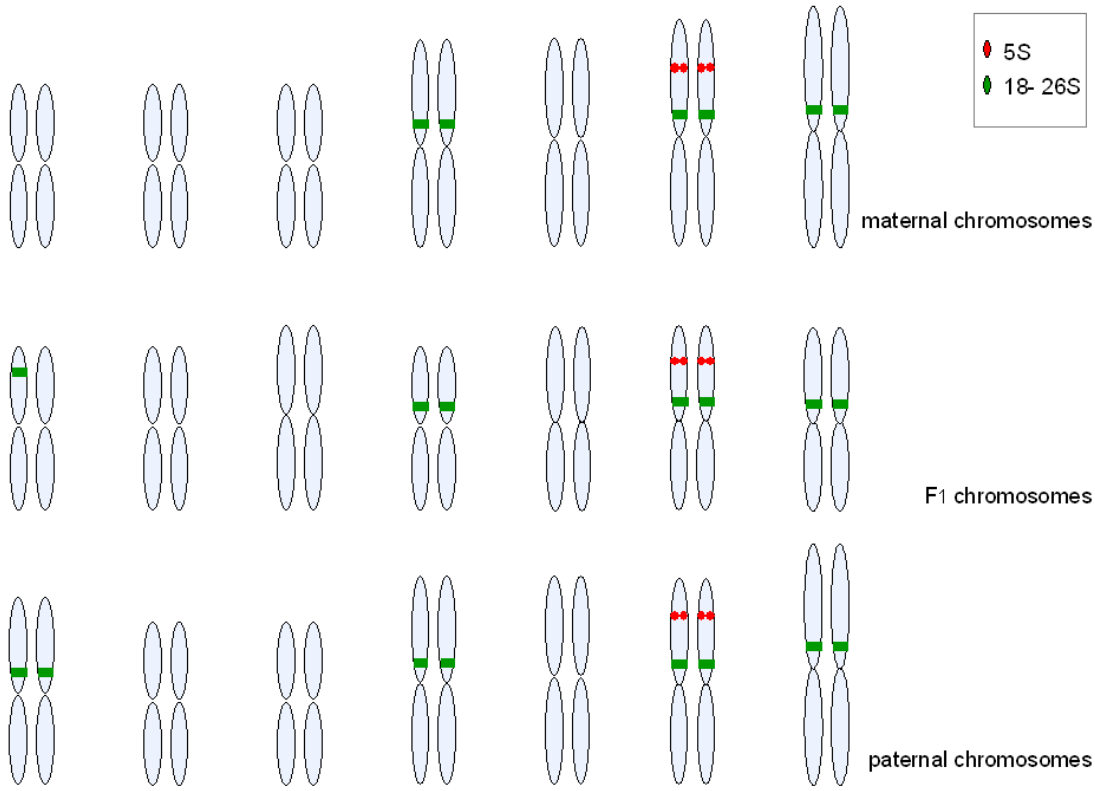


Figure 4 b: Karyograms with 5S (red) and 18-25S rDNA (green) probes of parental (PM, Pfa) and F_1 genotypes.

2.4 Discussion

2.4.1 Genomic in situ hybridization with parental genomic DNA

GISH and FISH have been shown to be useful in introgression studies, e.g. in wheat where depending on specific landrace lines small introgressed rye chromosomal segments in *Triticum aestivum* could be identified (Ribeiro-Carvalho *et al.* 1997). ISH can also detect alien chromosomes in late breeding lines, e.g. in a BC_6 line of *Pennisetum squamulatum* with *P. glaucum* introgression (Goel *et al.* 2003). Based on these approaches we deemed ISH would be suitable in our study for identifying large non-recombinant blocks of one or the other parent (PM or Pfa) in the F_1 line. The F_1 line showed recombination over the entire chromosome. No non-recombinant block via distinguished hybridization signal of one of the parents could be identified with GISH.

The use of genomic *Festuca* DNA as a probe could have been more successful to detect *Festuca* segments on the *Lolium* chromosomes. But after several generations of backcrossing to *L. perenne* during the line development, the *Festuca* proportion in the lines might have diminished and be too small to detect with FISH. Only small introgressed *Festuca* chromosome segments, not detectable with GISH, probably remained in the population presented in this study after several generations of backcrossing and selfing of the parental lines. Another hypothesis is that the *Lolium* genome has superiority over the *Festuca* genome. Extensive intergeneric recombination of *L. perenne* chromosomes might have replaced *F. pratensis* chromosomes over repeated cycles of meiosis at some stage of the life cycle (Canter *et al.* 1999). Like Canter *et al.* (1999), Zwierzykowski *et al.* (2006) found between a cross of allotetraploid hybrids of *F. pratensis* and *L. perenne* that the balance between the chromatin changed over several generations in favour of the *Lolium* chromatin. This

would suggest that also in the present study the *Festuca* proportions became too low to detect with GISH. Species specific RFLPs/AFLPs or species specific markers, which were used in the mapping procedure (later described in chapter four) could help locate *Festuca* segments on chromosomes. Rong *et al.* (2004) described, on the basis of sequence-tagged sites, functional structures and evolutionary genomic organisation of cotton (*Gossypium*). By means of RFLP sequences and SSRs the mobility of these STS loci could be shown across populations. Species specific RFLPs could provide information about the origin of the rDNA loci present in this study.

2.4.2 Fluorescent in situ hybridization with 18-25S rDNA (pTa71) and genes 5S rDNA (pTa794):

Variation in rDNA sites has been widely reported in plants. 18S-25S rDNA major loci can distinguish between subspecies as was shown in a study of different *Vigna* species where a species-specific DNA sequence for *V. unguiculata* was detected (Galasso *et al.* 1995). Interspecific polymorphism around the nucleolus organizer was shown before by Linde-Laursen (1984) in 29 barley lines by Giemsa C-banding. Twenty-seven lines showed the same signal pattern but two line showed polymorphism in the banding pattern.

The maternal (PM), paternal (Pfa) and F_1 genotypes in the present study showed two hybridization sites with 5S rDNAs but a different number of signals for the 18-25S rDNA. Thomas *et al.* (1996) found in species of *Lolium* some differences in outbreeders (*L. multiflorum*, *L. perenne*, and *L. rigidum*) compared with inbreeders (*L. tremulentum, L. remotum, L. persicum and L. rigidum*). Two 5S rDNA sites were always found but different numbers of 18-25S rDNA sites. All inbreeding *Lolium* species showed four 18-25S rDNA sites and the outbreeding genotypes of *L. multiflorum* six, *L. perenne*

seven and *L. rigidum* nine 18-25S rDNA sites. In the present study seven sites were found for the F_1 genotype as in the study of Thomas *et al.* (1996) but six in the maternal and eight in the paternal lines. *L. perenne* has variable 18-25S rDNA gene clusters (Thomas *et al.* 1996), which can lead to different numbers of 18-25S rDNA sites in *L. perenne* and could be the reason for differences in hybridization signals in the parental lines. Additionally, inbreeding might reduce the number of 18-25S rDNA sites. As shown in the study of Thomas *et al.* (1996) the inbreeding *Lolium* species had only four sites compared to the outbreeding species with more signals. This could explain why there were only six sites in the maternal inbred line in the present study. But it is probably more likely that a mutation in one of the parents led to the different number of 18-25S rDNA. DNA fragment losses like these have been described previously (Gaeta *et al.* 2007) in *Brassica* and could be an explanation for the missing signal of the 18-25S rDNA. Dubcovsky and Dvořák (1995) explained the absence of loci in *Triticum* that the loci have been reduced below the detection level of the *in situ* DNA hybridization technique or that they have been entirely eliminated during evolution. It is suggested that primate rDNA loci may move among chromosomes in a similar fashion as those in Triticeae where the major NOR loci had repeatedly changed position in the chromosome arms without rearrangements of the linkage groups (Dubcovsky and Dvořák 1995).

Genome organisation and a profound knowledge about it provide information likely to be used in plant breeding programmes (Galasso *et al.* 1995) as in the use of cytogenetic markers and high-resolution cytogenetic maps to breed superior grass cultivars (Kopecký *et al.* 2008). ISH is a powerful tool to link genetic and physical maps (King *et al.* 1998). ISH can show single recombination events (Kopecký *et al.* 2008), which can be used to visualize them in breeding programmes. This gives new

possibilities of precise working in breeding programmes and a better insight in the genome organisation.

Chapter 3: Metabolite analysis

3.1 Introduction

Plant growth and development are tightly linked to primary metabolism and are subject to natural variation. Genes apply their effect through control of metabolic processes. An attainment of a full understanding of inheritance is therefore dependent not only on knowledge of the nature of the heritable units and the process by which they act, but also on a complete knowledge of the cellular biochemistry they invoke. A central feature of this biochemical difficulty is the gathering of information on the numerous reaction pathways. The pathways can then be put together to give an overall metabolic pattern (Wagner and Mitchell 1965).

3.1.1 Pathways

The metabolism of a cell can be grouped into metabolic pathways, which comprise a sequence of chemical reactions. The compounds formed in each step along the pathway are named metabolites. Metabolites are the end products of the cellular regulatory processes, and their levels can be seen as the final reaction of biological systems to genetic or environmental changes (Fiehn 2002). Metabolic pathways can be divided into two broad types, the catabolic pathways and the anabolic pathways. Catabolic pathways have two functions. Firstly, they make the basic materials available to synthesise other molecules and secondly they provide the energy required for the activity of the cells. Anabolic pathways are dependant on catabolic pathways because they are energy-requiring and lead to the synthesis of more complex compounds from simple starting material (Karp 2008). The anabolic pathway synthesises metabolites like carbohydrates (sugars), lipids (fats, steroids, and phospholipids), proteins (like amino

acids, enzymes, hormones, and anti bodies), nucleic acids (DNA, and RNA) and phenolic compounds (like flavonoids). The number of metabolites present in the plant kingdom is estimated to exceed 200,000 (Weckwerth 2003). Different applications can be figured out for metabolomic analysis. Increasing metabolic fluxes into valuable biochemical pathways using metabolomic engineering or into the production of pharmaceuticals in plants are some examples. They are useful for the general understanding of novel metabolic pathways and to describe cellular networks *in vivo* (Fiehn 2002).

3.1.2 Metabolic profiling, target analysis, and fingerprinting

Metabolite profiling is an analytical method of parallel identification and relative quantification of a mixture of compounds or compound classes using chromatography and universal detection technologies (gas chromatography/mass spectrometry (GC/MS), liquid chromatography/mass spectrometry (LC/MS)) (Fiehn 2006) under defined conditions. Metabolomic methods imply adequate tissue sampling, homogenisation, extraction, storage, and sample preparation methods to maintain an unbiased process. Several critical parameters are accurate and make metabolite profiles a complex approach. The identification of metabolites is still a difficult task and most of the measured metabolites remain unidentified (Weckwerth 2003).

A target analysis is used when only a small number of analytes are chosen and signals for all the other components are ignored. Such targeted methods can only include a small fraction of the metabolome (Halket *et al.* 2004). GC/MS and LC/MS can also perform non-target analysis where all peaks can be characterised by their mass spectral pattern.

Fingerprinting techniques include nuclear magnetic resonance spectroscopy, Fourier transform infra-red spectroscopy, and direct infusion atmospheric pressure ionization MS. The advantage of these techniques is the high mass resolution. The development of such ultra high mass resolution instruments can overcome the current limitation on data processing and interpretation software (Halket *et al.* 2004) and provides a valuable new tool for the rapid metabolomic profiling of plants (Goodacre *et al.* 2007).

A major advantage is of metabolomics is that its mainly species-independent, therefore relatively little effort is necessary to re-optimize protocols for a new species (Schauer and Fernie 2006). Several studies has been carried out on metabolite profiling in different species in the last decade, e.g. in *Arabidopsis* (Fiehn 2000 and 2006), rice (Sato *et al.* 2004), wheat (Hamzehzarghani *et al.* 2005), alfalfa (Chen *et al.* 2003) and tomato (Schauer *et al.* 2005). Diverse metabolomic studies with different aims have been carried out in *Lolium*. Cao *et al.* (2008) looked at the metabolic effect of the symbiosis between the endophytic fungus *Neotyphodium lolii* and its host *L. perenne*. Rainey *et al.* (1987) looked at the associations between enzyme genotypes and dark respiration in *L. perenne*. Another study by Rasmussen *et al.* (2008) studied the response to nitrogen and its supply, carbohydrate content and fungal endophyte infection on different *Lolium* cultivars. Two studies looked at the fructan metabolism in *Lolium* (Chalmers *et al.* 2005, Lothier *et al.* 2007) and Amiard *et al.* (2003) published their results on a study about the putative role of fructans and sucrosyl-glactosides in relation to drought stress in *Lolium*. Changes in *Lolium* roots were investigated looking at the impact of supplying nitrogen as glycine to the roots (Thornton *et al.* 2007).

Key agronomic traits, like plant growth, or abiotic and biotic stress can be studied with the help of metabolic profiling. The specific aims of the work presented in

this chapter were (1) to compare the metabolic profiles of two inbred lines and the resulting F_1 from a cross between them and to investigate key metabolites associated with an observed heterosis effect in the F_1 and (2) to compare the metabolic profiles of these three lines at three points during the growing season.

3.2 Material and Methods

3.2.1 Plant material

F_1, maternal (PM) and paternal (Pfa) plants of the F_2 mapping population were established in 18 technical replications split between two blocks in the field in spring 2006. Leaf samples were harvested for metabolite analysis at three time points; on the 12th of June 2006, 15th of August 2006, and 19th of October 2006. Leaf samples were flash frozen immediately in liquid nitrogen. Samples were ground to a fine powder for 30 seconds at a frequency of 25 s^{-1} and stored at -80ºC. The final and intermediate harvest time points correspond to the times where the biomass harvests in the F_2 mapping population were undertaken (described in chapter five).

3.*2.2 Homogenisation and extraction*

The extraction protocol by Fiehn (2006) was followed with some modifications. 10 mg of frozen plant material was weighed, homogenized and immediately mixed with 1 ml extraction solution containing 3:3:2 (v/v) acetonitrile, isopropanol and water. Samples were vortexed for 10 seconds and shaken for 4-6 minutes at 4ºC. Samples were centrifuged for 2 minutes at 14000 rpm to remove the supernatant from the pellet.

3.2.3 Detection

Supernatant was analysed using gas chromatography coupled with mass spectroscopy (GC-MS). GC-MS was performed using a an Agilent 6890 gas chromatograph controlled by Leco ChromaTOF software version 2.32; http://www.leco.com. A Leco Pegasus IV time-of-flight mass spectrometer using Leco ChromaTOF software version 2.32 was used. Data were compared to the GC/MS Fiehn library in UCDavis (http://fiehnlab.ucdavis.edu). Data were normalized to the total metabolite content.

3.2.4 Data Analysis

Principal component analysis (PCA) was performed with GeneStat *for Windows*® 10th Edition (Payne *et al.* 2007). PCA was invented by Pearson (1901) and is used to compress data by linear combinations to derive principal vectors based on the variance inherent in the dataset (Fiehn *et al.* 2008). In this analysis a large amount of metabolite data can be summarised by means of a few parameters with a minimal loss of information. Therefore, all metabolite data for one plant can be simultaneously compared to other plants' metabolite profiles. PCA is an easy tool to compare large dataset (Schauer and Fernie 2006).

Additionally, t-tests were calculated between the genotypes to see if there are differences between the three harvests. Microsoft Office Excel 2003 (Microsoft Corporation) and GeneStat (Payne *et al.* 2007) were used for the calculation. Fold change explains in relative numbers how significant the differences between the genotypes were.

Hierarchical clustering and visualisation were performed using the software EpClust (http://bioinformaticsholstegelab.nl). The measured quantitative value levels of

the known metabolites of each genotype were calculated for each harvest relative to the median. The clustering analysis of levels above/below the median was performed.

3.3 Results

3.3.1 Comparing metabolic profile of the genotypes

As a starting point the metabolic profiles of the three genotypes used in the present study were compared. PCA using the measured quantitative values of each genotype (Appendix Table 1) showed that each harvest clustered together. PCA showed a distinct grouping of the three genotypes (Figure 1) with a variance of 20% in the PCA 1 and 14% for PCA 2. This was evidence that there was greater interaction in the metabolic profiles between the genotypes compared to the three harvest times.

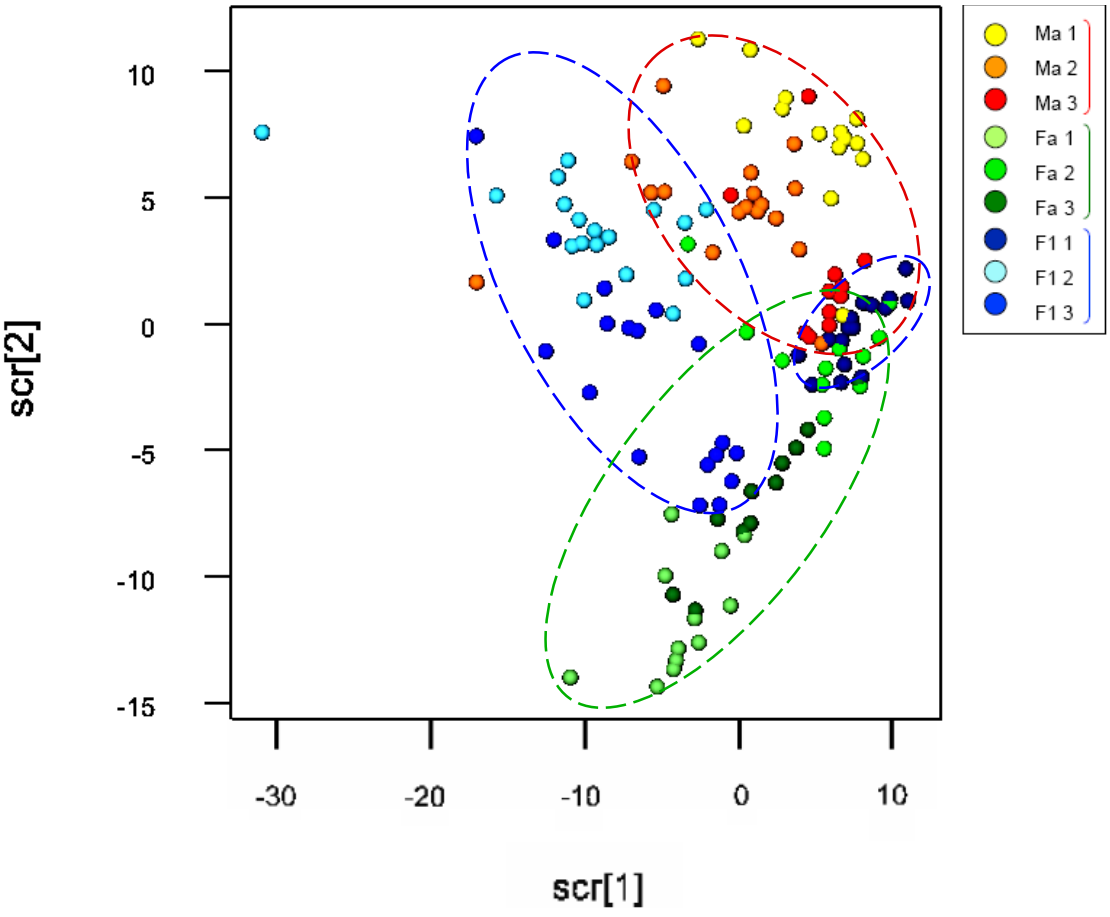


Figure 1: Principal component analysis (PCA) 1 and 2 of maternal (red), paternal (green) and F_1 (blue) genotype over all metabolites and harvest times (1 = 12.06.06, 2 = 15.08.06, 3 = 19.10.06).

Metabolites were grouped into 85 that could be identified and 167 that were unidentified. The relative quantities of the metabolites were compared between each genotype in turn using a t-test and the results are presented in tables 1 a and 1 b for known and unknown metabolites respectively. Fold change was used to give the significance in relative numbers to see how strong the significe of the results are. The metabolites are ordered from largest fold change to the smallest. There were significant differences in the metabolite profiling between the F_1 and the paternal genotypes, particularly for unknown compounds. For the known metabolites a number of compounds were identified that were significantly higher in the F_1 in comparison to both the paternal and maternal genotype (Table 1 a). These included citric acid, sorbitol,

arabitol, shikimic acid, fumarate, and aspartic acid. All these metabolites show high significant values between all genotypes and are later described in more detail.

Table 1 a: Significance of t-test (in **bold**: significant ratio: $p<0.05$) and fold change between the genotypes of known metabolites. Sorted in order of magnitude of fold change.

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
citric acid	**<0.001**	**3.0**	nicotinic acid	**<0.001**	**5.5**	tyramine	**<0.001**	**4.1**
phytol	**<0.001**	**2.9**	citric acid	**<0.001**	**2.3**	serine	**<0.001**	**3.3**
maleic acid	**<0.001**	**2.6**	galactinol	**<0.001**	**2.0**	glutamic acid	**<0.001**	**2.8**
tyramine	**<0.001**	**2.4**	fructose 2	**<0.001**	**1.9**	Glycerolphos-phate alpha	**<0.001**	**2.7**
sorbitol	**<0.001**	**2.4**	fructose 1	**<0.001**	**1.9**	threonine	**<0.001**	**2.7**
arabitol	**0.002**	**2.3**	arabitol	**0.021**	**1.9**	phenylalanine	**<0.001**	**2.6**
fumarate	**<0.001**	**2.2**	proline	0.224	1.8	phytol	**<0.001**	**2.6**
shikimic acid	**<0.001**	**2.2**	shikimic acid	**0.001**	**1.8**	inositol myo-	**<0.001**	**2.5**
glutamic acid	**<0.001**	**2.1**	sorbitol	**0.016**	**1.8**	alanine	**<0.001**	**2.4**
linolenic acid	**<0.001**	**2.1**	succinic acid	**<0.001**	**1.7**	glycine	**<0.001**	**2.4**
malate	**<0.001**	**2.0**	lactic acid	**0.013**	**1.7**	linolenic acid	**<0.001**	**2.3**
phenylalanine	**<0.001**	**2.0**	erythronic acid lactone	**0.002**	**1.7**	threonic acid	**<0.001**	**2.2**
inositol myo-	**<0.001**	**2.0**	maleic acid	**0.013**	**1.5**	tyrosine	**<0.001**	**2.2**
ribonic acid	**<0.001**	**1.9**	glycolic acid	**0.004**	**1.4**	phosphoric acid	**<0.001**	**2.2**
glucose-6-phosphate 2	**<0.001**	**1.9**	aspartic acid	**0.002**	**1.4**	stigmasterol	**<0.001**	**2.1**
aspartic acid	**<0.001**	**1.9**	fumarate	**0.003**	**1.4**	ribonic acid	**<0.001**	**2.1**
tyrosine	**<0.001**	**1.9**	xylose 1	**0.008**	**1.4**	beta alanine	**<0.001**	**2.0**
threonine	**<0.001**	**1.7**	xylose 2	**0.032**	**1.3**	glucose-6-phosphate 2	**<0.001**	**2.0**
glycerolphosphate alpha	**<0.001**	**1.7**	xylonic acid	**0.003**	**1.3**	linoleic acid	**<0.001**	**1.9**
phosphoric acid	**0.021**	**1.7**	N-acetyl-D-mannosamine	**0.016**	**1.3**	fructose-6-phosphate	**0.001**	**1.8**
erythronic acid lactone	**0.002**	**1.6**	glucose 2	0.318	1.3	GABA	**0.005**	**1.8**
linoleic acid	**<0.001**	**1.6**	caffeic acid	**0.021**	**1.2**	glyceric acid	**<0.001**	**1.8**
isofucostanol	**<0.001**	**1.6**	urea	0.445	1.2	malate	**<0.001**	**1.8**
threonic acid	**<0.001**	**1.6**	fucose 1 + rhamnose 2	0.143	1.2	lysine	**<0.001**	**1.8**
stigmasterol	**<0.001**	**1.6**	malate	0.090	1.1	maleic acid	**0.001**	**1.8**
fucose 1 + rhamnose 2	**<0.001**	**1.5**	erythritol	0.562	1.1	isofucostanol	**<0.001**	**1.7**
palmitic acid	**<0.001**	**1.5**	phytol	0.409	1.1	trehalose	0.290	1.6
serine	**<0.001**	**1.5**	hexose non-meox	0.671	1.1	fumarate	**<0.001**	**1.6**
fructose-6-phosphate	**0.030**	**1.5**	maltose 2	0.369	1.1	inulobiose 2	**0.004**	**1.5**
succinic acid	**<0.001**	**1.5**	threitol	0.452	1.1	tocopherol	**0.040**	**1.5**
benzoic acid	**0.022**	**1.4**	gluco-heptulose	0.469	1.1	glucose-1-phosphate	**<0.001**	**1.5**
glyceric acid	**<0.001**	**1.4**	putrescine	0.586	1.1	hydroxyl-amine	**0.002**	**1.5**
beta alanine	**0.002**	**1.4**	glucose 1	0.625	1.1	palmitic acid	**<0.001**	**1.5**
maltose 2	**0.003**	**1.4**	mucic acid	0.784	1.1	benzoic acid	**0.002**	**1.5**
lysine	**0.008**	**1.4**	quinic acid	0.773	1.0	oxoproline	**0.006**	**1.5**
mucic acid	0.157	1.4	xylitol	0.907	1.0	valine	**0.012**	**1.4**
galactinol	**0.038**	**1.3**	stearic acid	0.887	1.0	pelargonic acid	**0.023**	**1.4**

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
urea	0.226	1.3	inulobiose 1	0.971	1.0	levoglucosan	0.322	1.4
glucose-1-phosphate	**<0.001**	**1.3**	hexaric acid 2	0.978	1.0	leucine	0.111	1.4
GABA	**0.041**	**1.3**	palmitic acid	0.982	1.0	asparagine	0.379	1.4
hexaric acid 2	**0.035**	**1.3**	benzoic acid	0.778	1.0	hexaric acid 1	**0.045**	**1.4**
lactic acid	0.180	1.3	isofucostanol	0.461	0.9	sorbitol	0.133	1.3
gluco-heptulose	**0.033**	**1.3**	lyxitol	0.490	0.9	sucrose	0.060	1.3
alanine	0.152	1.2	glucose-6-phosphate 2	0.586	0.9	citric acid	**0.047**	**1.3**
glycolic acid	**0.045**	**1.2**	ribonic acid	0.550	0.9	aspartic acid	**0.016**	**1.3**
leucine	0.217	1.2	leucine	0.579	0.9	cyano-L-alanine	0.351	1.3
stearic acid	**0.014**	**1.2**	linolenic acid	0.340	0.9	fucose 1 + rhamnose 2	**0.027**	**1.3**
levoglucosan	0.556	1.2	sucrose	0.429	0.9	isoleucine	0.131	1.3
xylose 2	0.129	1.2	alpha ketoglutaric acid	0.516	0.9	mucic acid	0.282	1.3
sucrose	0.195	1.2	levoglucosan	0.649	0.9	hexaric acid 2	**0.048**	**1.3**
hydroxylamine	0.084	1.2	glucose-1-phosphate	**0.036**	**0.9**	glutamine	0.199	1.3
nicotinic acid	0.383	1.2	hexaric acid 1	0.270	0.9	maltose 2	**0.042**	**1.3**
tocopherol	0.385	1.2	linoleic acid	0.147	0.9	arabitol	0.424	1.2
hexaric acid 1	0.316	1.2	tyrosine	0.100	0.8	stearic acid	0.059	1.2
oxoproline	0.232	1.2	inositol myo-	**0.014**	**0.8**	shikimic acid	0.458	1.2
xylose 1	0.202	1.1	fructose-6-phosphate	0.078	0.8	hexonic acid	**0.020**	**1.2**
inulobiose 1	0.326	1.1	phosphoethanolamine	0.052	0.8	phosphoethanolamine	0.131	1.2
xylonic acid	0.157	1.1	hydroxyl-amine	0.052	0.8	gluco-heptulose	0.246	1.2
caffeic acid	0.708	1.1	isoleucine	0.106	0.8	inulobiose 1	0.290	1.1
valine	0.590	1.1	glyceric acid	**0.010**	**0.8**	urea	0.724	1.1
pelargonic acid	0.657	1.1	phosphoric acid	0.142	0.8	erythronic acid lactone	0.773	1.0
isoleucine	0.945	1.0	lysine	**0.023**	**0.8**	xylose 2	0.514	0.9
glycine	0.966	1.0	oxoproline	**0.024**	**0.8**	caffeic acid	0.601	0.9
Phosphor-ethanolamine	0.574	0.9	phenylalanine	**0.020**	**0.8**	glycolic acid	0.215	0.9
trehalose	0.762	0.9	hexonic acid	**0.001**	**0.8**	xylonic acid	0.081	0.9
inulobiose 2	0.612	0.9	tocopherol	0.107	0.8	succinic acid	**0.028**	**0.8**
hexonic acid	0.172	0.9	inulotriose 1	0.157	0.8	xylose 1	0.142	0.8
glucose 2	0.217	0.8	stigmasterol	**<0.001**	**0.8**	lactic acid	**0.019**	**0.7**
putrescine	0.053	0.8	pelargonic acid	**0.026**	**0.8**	putrescine	**0.005**	**0.7**
glucose 1	**0.002**	**0.7**	glutamic acid	**0.015**	**0.7**	alpha ketoglutaric acid	0.651	0.7
glutamine	**0.011**	**0.7**	valine	**0.021**	**0.7**	inulotriose 1	0.182	0.7
alpha ketoglutaric acid	0.474	0.6	threonic acid	**0.001**	**0.7**	galactinol	**0.027**	**0.7**
quinic acid	**<0.001**	**0.6**	GABA	0.064	0.7	glucose 1	**0.003**	**0.7**
N-acetyl-D-mannosamine	**<0.001**	**0.6**	beta alanine	**<0.001**	**0.7**	glucose 2	**0.016**	**0.6**
fructose 1	**<0.001**	**0.6**	threonine	**<0.001**	**0.7**	quinic acid	**<0.001**	**0.6**
fructose 2	**<0.001**	**0.5**	glycerolphosphate alpha	**<0.001**	**0.6**	lyxitol	**<0.001**	**0.6**
threitol	**<0.001**	**0.5**	inulobiose 2	**<0.001**	**0.6**	threitol	**<0.001**	**0.5**
lyxitol	**<0.001**	**0.5**	trehalose	0.176	0.6	hexose non-meox	**<0.001**	**0.5**
hexose non-meox	**<0.001**	**0.5**	tyramine	**<0.001**	**0.6**	N-acetyl-D-mannosamine	**<0.001**	**0.5**

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
inulotriose 1	**0.020**	**0.5**	glutamine	**<0.001**	**0.5**	xylitol	**<0.001**	**0.3**
cyano-L-alanine	**0.002**	**0.4**	alanine	**<0.001**	**0.5**	fructose 1	**<0.001**	**0.3**
xylitol	**<0.001**	**0.3**	serine	**<0.001**	**0.5**	fructose 2	**<0.001**	**0.3**
proline	**<0.001**	**0.2**	glycine	**<0.001**	**0.4**	nicotinic acid	**<0.001**	**0.2**
asparagine	**<0.001**	**0.2**	cyano-L-alanine	**<0.001**	**0.3**	erythritol	**<0.001**	**0.2**
erythritol	**<0.001**	**0.2**	asparagine	**<0.001**	**0.2**	proline	**<0.001**	**0.1**

Table 1 b: Significance of t-test (in **bold**: significant ratio: $p<0.05$) and fold change between the genotypes of unknown metabolites. Sorted in order of magnitude of fold change.

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
206528	**<0.001**	**5.7**	210894	**<0.001**	**22.2**	200450	**<0.001**	**11.3**
211894	**<0.001**	**4.1**	210901	**<0.001**	**22.0**	208662	**<0.001**	**4.6**
212679	**<0.001**	**3.7**	215344	**<0.001**	**18.8**	208845	**0.012**	**4.5**
213182	**<0.001**	**3.5**	215643	**<0.001**	**17.0**	215355	**<0.001**	**4.0**
200450	**<0.001**	**3.4**	210896	**<0.001**	**12.6**	210882	**<0.001**	**3.5**
211919	**<0.001**	**2.9**	215402	**<0.001**	**12.1**	206528	**<0.001**	**3.3**
202570	**<0.001**	**2.7**	211934	**<0.001**	**12.1**	206136	**<0.001**	**3.2**
211886	**0.033**	**2.5**	215362	**<0.001**	**12.0**	212679	**<0.001**	**3.1**
199239	**<0.001**	**2.4**	215347	**<0.001**	**11.3**	204344	**<0.001**	**3.0**
211934	**0.001**	**2.3**	212024	**<0.001**	**10.5**	199239	**<0.001**	**3.0**
200486	**<0.001**	**2.3**	211890	**<0.001**	**10.0**	200448	**0.007**	**2.7**
216472	**<0.001**	**2.3**	215978	**<0.001**	**9.8**	199317	**<0.001**	**2.5**
211890	**0.003**	**2.2**	211914	**<0.001**	**8.6**	211941	**0.012**	**2.4**
215493	**0.002**	**2.0**	203157	**<0.001**	**8.0**	199235	**<0.001**	**2.4**
208841	**<0.001**	**2.0**	216454	**<0.001**	**7.4**	208840	**0.001**	**2.4**
212208	**0.002**	**1.9**	211898	**<0.001**	**6.9**	207507	**0.012**	**2.3**
205664	**<0.001**	**1.9**	210909	**<0.001**	**6.7**	212663	**<0.001**	**2.2**
213194	**0.001**	**1.8**	215399	**<0.001**	**6.7**	200463	**<0.001**	**2.2**
200624	**0.009**	**1.7**	215529	**<0.001**	**6.3**	201051	**<0.001**	**2.1**
199235	**<0.001**	**1.7**	215504	**<0.001**	**6.1**	208841	**<0.001**	**2.0**
199338	**<0.001**	**1.7**	216424	**<0.001**	**6.0**	200491	**<0.001**	**2.0**
202834	**<0.001**	**1.7**	215563	**<0.001**	**5.8**	215062	**0.007**	**2.0**
200491	**0.002**	**1.6**	216493	**<0.001**	**5.7**	213732	**0.003**	**2.0**
208658	**0.007**	**1.6**	216098	**<0.001**	**5.3**	211962	**<0.001**	**1.9**
214405	0.212	1.6	205849	**<0.001**	**5.2**	211916	**<0.001**	**1.8**
201051	**<0.001**	**1.6**	215493	**<0.001**	**5.1**	200486	**<0.001**	**1.8**
215529	**0.017**	**1.6**	211894	**<0.001**	**4.6**	215490	**0.001**	**1.8**
213714	**0.001**	**1.6**	211980	**0.007**	**4.2**	213194	**0.022**	**1.8**
200448	**0.036**	**1.6**	212189	**<0.001**	**4.1**	200401	**0.003**	**1.8**
215355	**0.020**	**1.6**	216472	**<0.001**	**4.0**	211919	**0.013**	**1.8**
215062	**0.023**	**1.6**	210904	**<0.001**	**3.7**	200513	**<0.001**	**1.7**
211935	**0.025**	**1.6**	211886	**0.008**	**3.7**	202834	**<0.001**	**1.7**
210882	**0.007**	**1.5**	200384	**<0.001**	**3.2**	213714	**0.004**	**1.6**
204344	0.124	1.5	213155	**<0.001**	**2.9**	205664	**<0.001**	**1.6**
202178	**<0.001**	**1.5**	213182	**<0.001**	**2.8**	208701	**<0.001**	**1.6**
207750	**0.001**	**1.5**	211911	**<0.001**	**2.6**	216427	**0.003**	**1.6**
210512	0.094	1.5	202570	**0.001**	**2.6**	200531	**0.003**	**1.5**

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
212663	**0.002**	**1.5**	211935	**0.001**	**2.2**	200532	**0.001**	**1.5**
202083	**0.018**	**1.5**	213185	0.077	2.1	206318	0.171	1.5
211941	0.061	1.5	200414	**<0.001**	**2.1**	202599	0.092	1.5
211911	**0.031**	**1.5**	200624	**0.002**	**2.0**	202178	**0.011**	**1.5**
214401	**<0.001**	**1.5**	206528	**<0.001**	**1.7**	208897	**0.001**	**1.5**
211962	**<0.001**	**1.5**	200556	**0.006**	**1.7**	215466	**0.016**	**1.5**
211916	**0.006**	**1.5**	211919	**0.014**	**1.7**	212208	**0.027**	**1.5**
213310	**<0.001**	**1.5**	211917	**0.007**	**1.7**	207326	**0.006**	**1.4**
216427	**0.015**	**1.5**	210893	**0.017**	**1.6**	205673	0.083	1.4
208840	**0.037**	**1.5**	215860	0.081	1.6	199463	**0.001**	**1.4**
208662	**0.007**	**1.4**	200900	0.072	1.6	200426	**<0.001**	**1.4**
216454	0.075	1.4	215375	**0.017**	**1.5**	214405	0.311	1.4
213732	0.247	1.4	208658	**0.035**	**1.5**	214434	0.184	1.4
200532	**0.003**	**1.4**	208874	0.081	1.4	213310	**<0.001**	**1.4**
215504	0.190	1.4	202083	0.057	1.4	208651	0.208	1.3
208874	0.057	1.4	202808	**0.004**	**1.4**	214680	**0.002**	**1.3**
199223	**0.006**	**1.4**	207750	**0.017**	**1.4**	205857	0.201	1.3
200489	**<0.001**	**1.4**	199777	0.098	1.4	214401	**0.021**	**1.3**
200426	**<0.001**	**1.4**	199338	**0.002**	**1.4**	200392	**0.009**	**1.3**
208664	0.280	1.4	212735	**0.011**	**1.4**	213182	0.187	1.3
215466	**0.049**	**1.4**	210891	**0.033**	**1.4**	200567	0.139	1.2
212024	0.151	1.3	208850	0.289	1.4	210512	0.498	1.2
215492	**0.031**	**1.3**	212208	0.177	1.3	200466	0.088	1.2
211980	0.370	1.3	210912	0.093	1.3	211891	0.120	1.2
215682	**0.048**	**1.3**	200595	**0.022**	**1.3**	200489	**0.014**	**1.2**
208897	**0.001**	**1.3**	200896	0.346	1.3	199338	0.071	1.2
200513	**0.013**	**1.3**	215682	0.127	1.3	203264	0.267	1.2
211917	0.084	1.3	200511	0.077	1.2	208664	0.468	1.2
215347	0.233	1.3	210512	0.364	1.2	199177	0.410	1.2
215490	0.092	1.3	200486	**0.050**	**1.2**	215448	0.253	1.2
200392	**0.003**	**1.3**	212732	0.295	1.2	216564	0.706	1.2
200463	0.076	1.3	199328	0.062	1.2	199223	0.290	1.2
199317	**0.045**	**1.3**	214410	0.244	1.2	202838	0.519	1.1
200466	**0.042**	**1.3**	199223	0.072	1.2	210399	0.534	1.1
210891	0.085	1.2	215492	0.217	1.2	203250	0.397	1.1
212735	0.086	1.2	212679	0.467	1.2	200518	0.680	1.1
200567	0.089	1.2	214405	0.610	1.2	208658	0.404	1.1
200531	0.086	1.2	207509	0.211	1.2	211896	0.770	1.1
201832	0.108	1.2	205664	0.105	1.2	215492	0.477	1.1
199463	**0.020**	**1.2**	200874	0.288	1.2	199562	0.625	1.1
214434	0.333	1.2	214401	0.129	1.2	201832	0.568	1.1
208701	0.052	1.2	208664	0.589	1.2	207750	0.573	1.1
215344	0.317	1.2	211946	0.203	1.2	200427	0.572	1.1
202838	0.164	1.2	202573	0.530	1.1	212781	0.492	1.1
211946	0.053	1.2	211972	0.490	1.1	202083	0.678	1.1
205673	0.364	1.2	200489	**0.029**	**1.1**	202737	0.826	1.1
199328	0.077	1.2	201832	0.379	1.1	199231	0.604	1.1
210904	0.299	1.2	200521	0.514	1.1	211946	0.555	1.1
210901	0.363	1.2	199175	0.625	1.1	213143	0.804	1.1
206318	0.378	1.2	202838	0.564	1.1	202570	0.888	1.0
200401	0.239	1.2	207432	0.560	1.1	199215	0.705	1.0
216098	0.351	1.2	208770	0.408	1.1	213271	0.801	1.0

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
215402	0.410	1.2	213310	0.301	1.1	215682	0.839	1.0
211891	0.154	1.2	213271	0.618	1.1	200844	0.845	1.0
211914	0.393	1.2	200549	0.577	1.1	208770	0.857	1.0
215563	0.421	1.2	200429	0.719	1.1	199328	0.916	1.0
214680	**0.040**	**1.2**	200427	0.731	1.1	200429	0.888	1.0
210894	0.493	1.2	215555	0.936	1.0	207509	0.893	1.0
207509	0.290	1.2	200466	0.805	1.0	208874	0.893	1.0
199177	0.435	1.1	205680	0.829	1.0	215555	0.962	1.0
207326	0.263	1.1	203052	0.941	1.0	203052	0.845	1.0
200427	0.259	1.1	202178	0.927	1.0	200420	0.794	1.0
215362	0.524	1.1	200392	0.908	1.0	200540	0.869	1.0
210399	0.483	1.1	200567	0.983	1.0	210891	0.447	0.9
213155	0.506	1.1	210399	0.999	1.0	211894	0.661	0.9
203250	0.295	1.1	202834	0.993	1.0	207432	0.522	0.9
205857	0.594	1.1	203250	0.961	1.0	212735	0.430	0.9
216424	0.647	1.1	213194	0.962	1.0	200874	0.415	0.9
215399	0.591	1.1	200426	0.874	1.0	200624	0.329	0.9
208770	0.336	1.1	199562	0.936	1.0	200521	0.242	0.8
213271	0.396	1.1	208841	0.885	1.0	200549	**0.027**	**0.8**
215978	0.646	1.1	213714	0.888	1.0	214410	0.146	0.8
216493	0.700	1.1	211891	0.766	1.0	200416	0.166	0.8
199562	0.600	1.1	199177	0.880	1.0	211917	0.077	0.8
215643	0.711	1.1	212781	0.645	1.0	210912	**0.022**	**0.7**
210893	0.606	1.1	216427	0.642	0.9	211935	0.078	0.7
203157	0.723	1.1	215466	0.631	0.9	211886	0.430	0.7
212189	0.748	1.1	200532	0.459	0.9	205680	**<0.001**	**0.7**
211898	0.817	1.0	199215	0.300	0.9	199777	**0.002**	**0.7**
215448	0.675	1.0	199231	0.344	0.9	210893	**0.022**	**0.7**
206136	0.805	1.0	215448	0.379	0.9	200511	**<0.001**	**0.6**
200429	0.756	1.0	199205	0.348	0.9	211972	**0.005**	**0.6**
212781	0.778	1.0	214434	0.601	0.9	211911	**0.001**	**0.6**
200874	0.888	1.0	200416	0.384	0.9	216472	**0.047**	**0.6**
215555	0.938	1.0	214680	0.089	0.9	199175	**0.001**	**0.5**
210896	0.988	1.0	199463	0.154	0.9	199205	**<0.001**	**0.5**
210909	0.972	1.0	208897	0.208	0.9	212732	**<0.001**	**0.5**
207432	0.877	1.0	205857	0.462	0.9	202808	**<0.001**	**0.4**
203052	0.905	1.0	200844	0.134	0.9	200595	**<0.001**	**0.4**
199231	0.729	1.0	205673	0.449	0.8	215493	**<0.001**	**0.4**
214410	0.733	1.0	200420	0.271	0.8	213185	**0.007**	**0.4**
199215	0.652	1.0	211896	0.552	0.8	200414	**<0.001**	**0.4**
200521	0.635	0.9	200540	0.447	0.8	213155	**<0.001**	**0.4**
210912	0.633	0.9	200491	0.157	0.8	200896	**0.002**	**0.4**
199777	0.553	0.9	202737	0.177	0.8	208850	**<0.001**	**0.4**
208845	0.544	0.9	211916	0.137	0.8	200900	**<0.001**	**0.4**
211896	0.767	0.9	211962	**0.013**	**0.8**	215860	**<0.001**	**0.3**
200844	0.173	0.9	200531	0.082	0.8	210904	**<0.001**	**0.3**
202737	0.614	0.9	199239	**0.004**	**0.8**	211980	**<0.001**	**0.3**
200549	0.102	0.9	207326	0.070	0.8	212189	**<0.001**	**0.3**
213185	0.523	0.8	215062	0.282	0.8	215529	**<0.001**	**0.3**
208651	0.306	0.8	206318	0.365	0.8	202573	**<0.001**	**0.2**
207507	0.450	0.8	208701	**0.013**	**0.8**	212274	0.347	0.2
200414	0.149	0.8	200513	**0.022**	**0.8**	215504	**<0.001**	**0.2**

	F1 / maternal			F1 / paternal			paternal / maternal	
metabolite	t-test	fold change	metabolite	t-test	fold change	metabolite	t-test	fold change
200420	0.166	0.8	201051	**0.012**	**0.8**	211890	**<0.001**	**0.2**
216564	0.450	0.8	213732	0.171	0.7	216098	**<0.001**	**0.2**
200540	0.301	0.8	215490	**0.018**	**0.7**	215563	**<0.001**	**0.2**
200511	**0.008**	**0.8**	212274	0.217	0.7	200556	**<0.001**	**0.2**
203264	0.061	0.7	213143	0.099	0.7	211934	**<0.001**	**0.2**
213143	0.120	0.7	212663	**0.025**	**0.7**	216454	**<0.001**	**0.2**
211972	**0.024**	**0.7**	199235	**0.005**	**0.7**	216493	**<0.001**	**0.2**
202599	0.204	0.7	216564	0.262	0.7	216424	**<0.001**	**0.2**
200416	**0.013**	**0.7**	200401	**0.028**	**0.7**	200384	**<0.001**	**0.2**
205680	**<0.001**	**0.7**	208651	**0.038**	**0.6**	215399	**<0.001**	**0.2**
202808	**0.001**	**0.6**	211941	0.088	0.6	211898	**<0.001**	**0.2**
212732	**<0.001**	**0.6**	203264	**0.002**	**0.6**	210909	**<0.001**	**0.1**
200384	**0.002**	**0.6**	208840	**0.022**	**0.6**	211914	**<0.001**	**0.1**
205849	**<0.001**	**0.6**	200463	**<0.001**	**0.6**	203157	**<0.001**	**0.1**
215860	**0.001**	**0.5**	200448	0.065	0.6	212024	**<0.001**	**0.1**
199175	**<0.001**	**0.5**	204344	**<0.001**	**0.5**	215347	**<0.001**	**0.1**
200900	**0.001**	**0.5**	199317	**<0.001**	**0.5**	215978	**<0.001**	**0.1**
200595	**<0.001**	**0.5**	202599	**<0.001**	**0.5**	205849	**<0.001**	**0.1**
208850	**0.001**	**0.5**	210882	**<0.001**	**0.4**	215402	**<0.001**	**0.1**
200896	**0.001**	**0.5**	215355	**<0.001**	**0.4**	215362	**<0.001**	**0.1**
199205	**<0.001**	**0.4**	207507	**0.001**	**0.3**	210896	**<0.001**	**0.1**
200518	**<0.001**	**0.3**	206136	**<0.001**	**0.3**	215344	**<0.001**	**0.1**
200556	**<0.001**	**0.3**	208662	**<0.001**	**0.3**	215375	**<0.001**	**0.1**
202573	**<0.001**	**0.3**	200518	**<0.001**	**0.3**	215643	**<0.001**	**0.1**
212274	0.210	0.2	200450	**0.001**	**0.3**	210901	**<0.001**	**0.1**
215375	**<0.001**	**0.1**	208845	**0.004**	**0.2**	210894	**<0.001**	**0.1**

Graphs in figure 2 showing relative mean values across harvest times for six of the above mentioned metabolites.

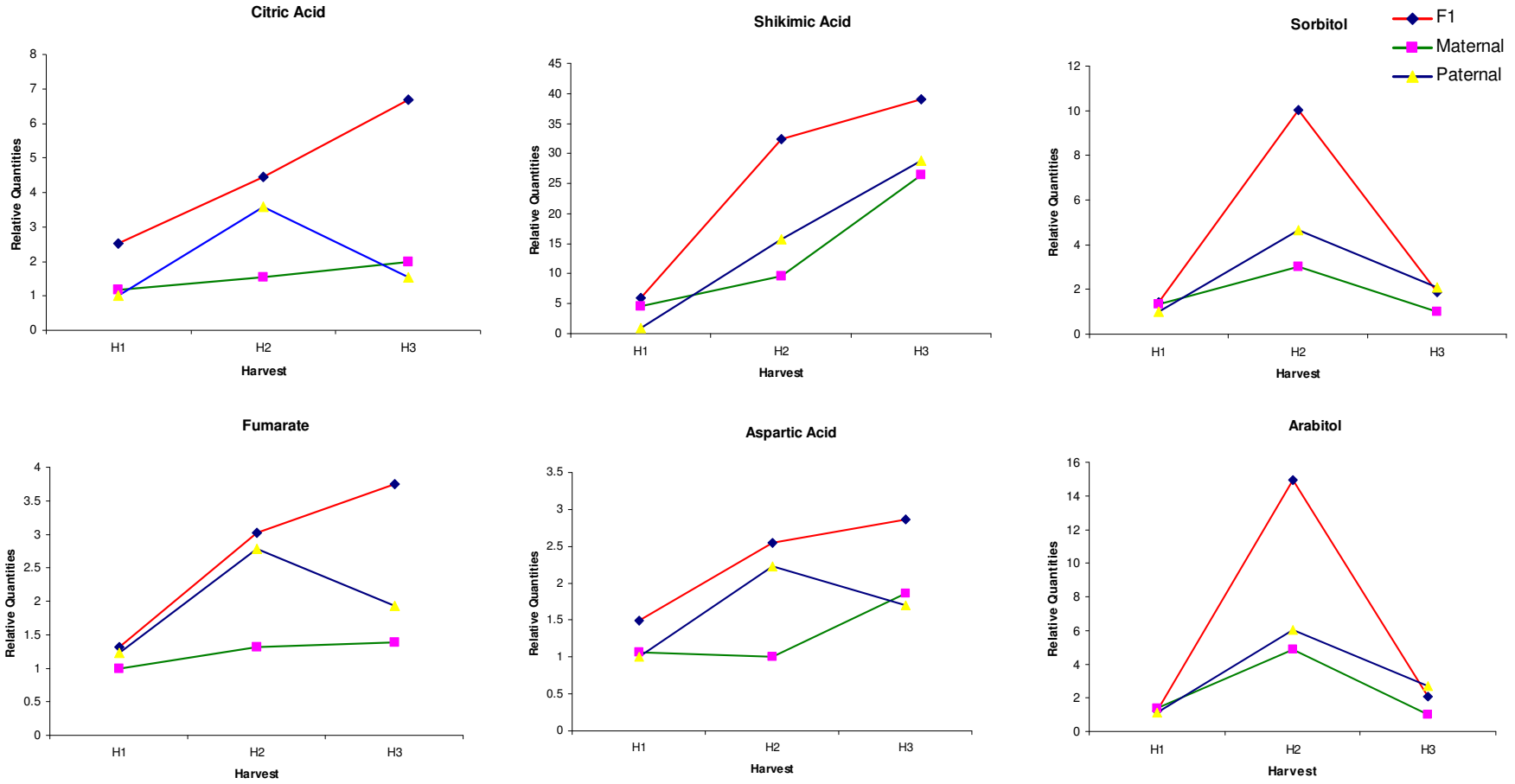


Figure 2: Known metabolites that showed some of the greatest difference between the F_1 and both the maternal and paternal genotype. Single harvests: H1) 12.06.2006, H2) 15.08.2006, and H3) 19.10.2006.

Citric acid content was seen to be higher in the F_1 at all individual harvest times, particularly at harvest three. The citric acid cycle is the central metabolic pathway for all aerobic processes and fumarate, an intermediate in the citric acid cycle, also accumulated in the F_1 (Figure 2).

The amino acid aspartic had a very similar profile to both fumarate and citric acid for all three genotypes, with the F_1 showing significantly higher accumulation. Another metabolite with a significant difference between the F_1 and parental genotypes, particularly at harvest two, was shikimic acid. It had a similar profile in the F_1 to the metabolites discussed above, with the highest levels in the third harvest (Figure 2). A large difference in the relative amount of this metabolite can be seen especially in harvest three compared to harvest one in all genotypes. It has been shown in *Vigna sinensis* plants that application of shikimic acid resulted in improved yield and yield components. In addition it was shown that it stimulated the production of Chlorophyll a, Chlorophyll b, Carotenoids and $_{14}C$ fixation during leaf growth and development (Aldesuquy *et al.* 2000).

The other two metabolites with a significant difference between the F_1 and both parental genotypes were arabitol and sorbitol. Both of these metabolites had a similar profile across the harvests with a much higher accumulation in the F_1 at harvest two. Both arabitol and sorbitol are sugar alcohols that have a role to play in adaptation to stress (Richardson *et al.* 1992). Higher sugar alcohol activity was found especially in the F_1 genotype in the second harvest. The climate was very dry and warm when the second harvest was done and it may be possible that this is the reason for the significant increase in these sugars at this harvest point.

In addition to differences between the F_1 and parental lines there were also significant differences between the parental genotypes, especially in many of the amino

acids (Table 1) and these accounted for most of the differences observed in the known metabolites. An amino acid pathway showing these metabolites is presented in figure 3.

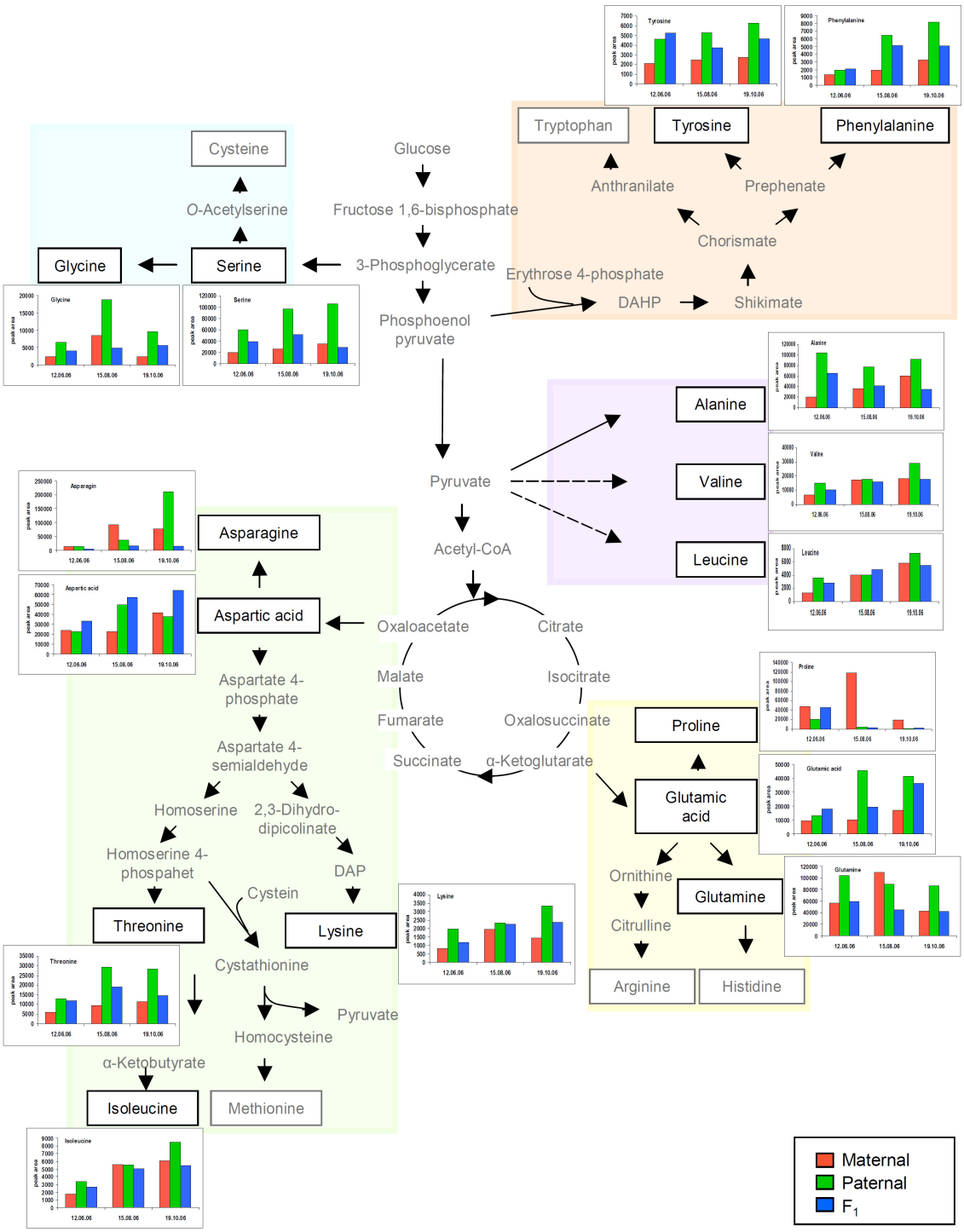


Figure 3: Amino acid pathway with individual graph of the measured quantity values over harvest times and the three genotypes
(after: http://www.uky.edu/~dhild/biochem/24/lect24.html).

In general, accumulation of amino acids was highest in the paternal genotype, e.g. glycine and serine, which was surprising considering the phenotypes of the plants. It would have been expected to find higher metabolite activity in the F_1 genotype because of its high heterosis effect (see results in chapter five).

Glycine and serine showed a higher activity in the paternal compared to the maternal and the F_1 genotypes. Glycine and serine are in the same pathway and are involved in photosynthesis. Higher activity in proline could be found in the maternal genotype especially in the second harvest. The other genotypes showed a low accumulation of proline compared to the first harvest and the maternal genotype in the second harvest. Proline has been shown to accumulate in response to osmotic stress (Verbruggen and Hermans 2008). Asparagine and glutamine displayed an inconsistent profile during the harvests. Asparagine accumulation was relatively low during the first harvest. In the second harvest the maternal genotype had the highest accumulation and in the third harvest the paternal genotype had significantly higher amounts of asparagine. Glutamine showed the same pattern as the other metabolites with the highest activity in the paternal genotype except in the second harvest where the maternal genotype had relatively higher amounts.

3.3.2 Comparing metabolic profiles across the growing season

In the previous section the metabolic profiles of the three genotypes were compared. In this section the changes in metabolic profiles of the genotypes and the harvests during the growing season are compared. Samples were harvested in mid June, mid August and mid October 2006. This represents two points (mid June and mid August) around the peak of grass growth and a point when the rate of growth was reduced (mid October).

A PCA was performed first to see if we could identify clusters according to harvest time in each of the three genotypes. It can be seen from figure 4 that in all three genotypes we see a clear clustering of samples according to harvest time. The variances explained over all harvests for maternal, paternal and F_1 genotypes were 35%, 22% and 28% in the PCA 1 and for the PCA 2: 15%, 17%, 15% respectively.

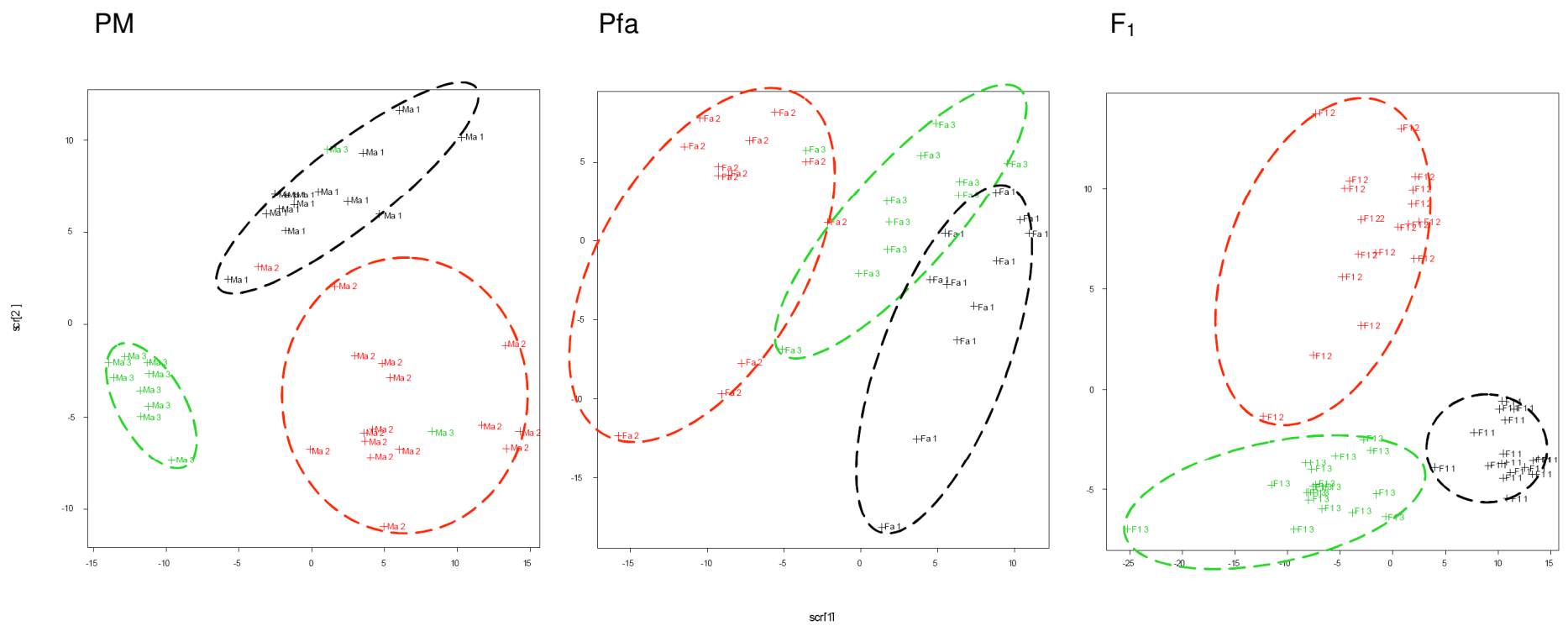


Figure 4: Principle component analysis (PCA) 1 and 2 of the maternal (PM), paternal (Pfa) and F_1 genotype. Single harvests: 1) 12.06.2006 (black), 2) 15.08.2006 (red), and 3) 19.10.2006 (green).

The next step was to identify those metabolites that clustered together with a similar profile across the harvests. A hierarchal cluster analysis was performed for each genotype separately (Figure 5 A, B and C) using only the known metabolites. Clustering was performed on averages of the replicates. An analysis of variance (ANOVA) was performed on the complete list of metabolites (known and unknown) to identify those with a significant difference across harvest times and is included in Appendix Table 2.

A)

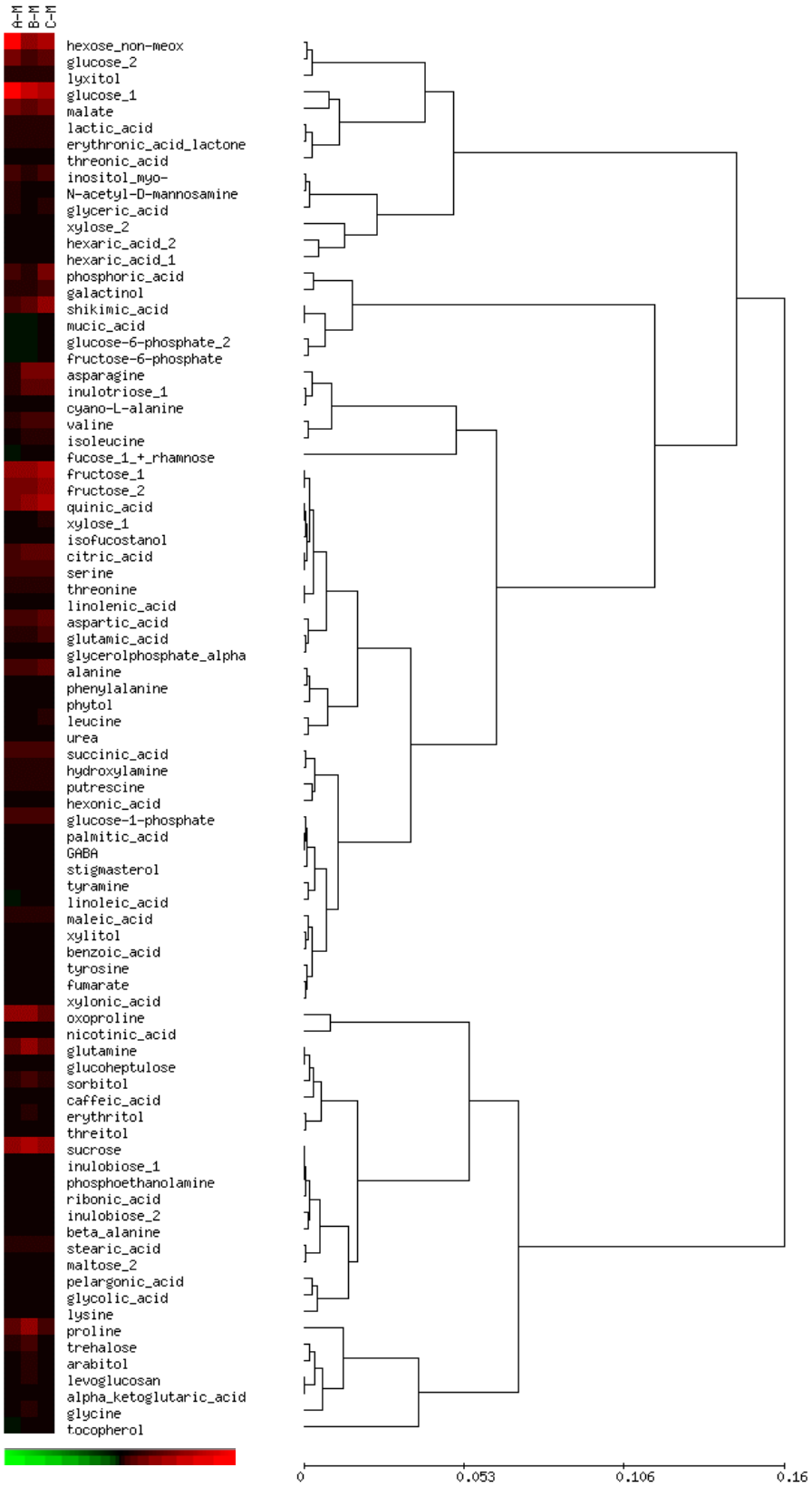


Figure 5: Hierarchical clustering with the programme EPCLUST using the metabolite levels below (green)/above (red) the median was performed. Known metabolites of the A) maternal genotype and the three harvests (A-M: 12.06.2006, B-M: 15.08.2006, C-M: 19.10.2006).

B)

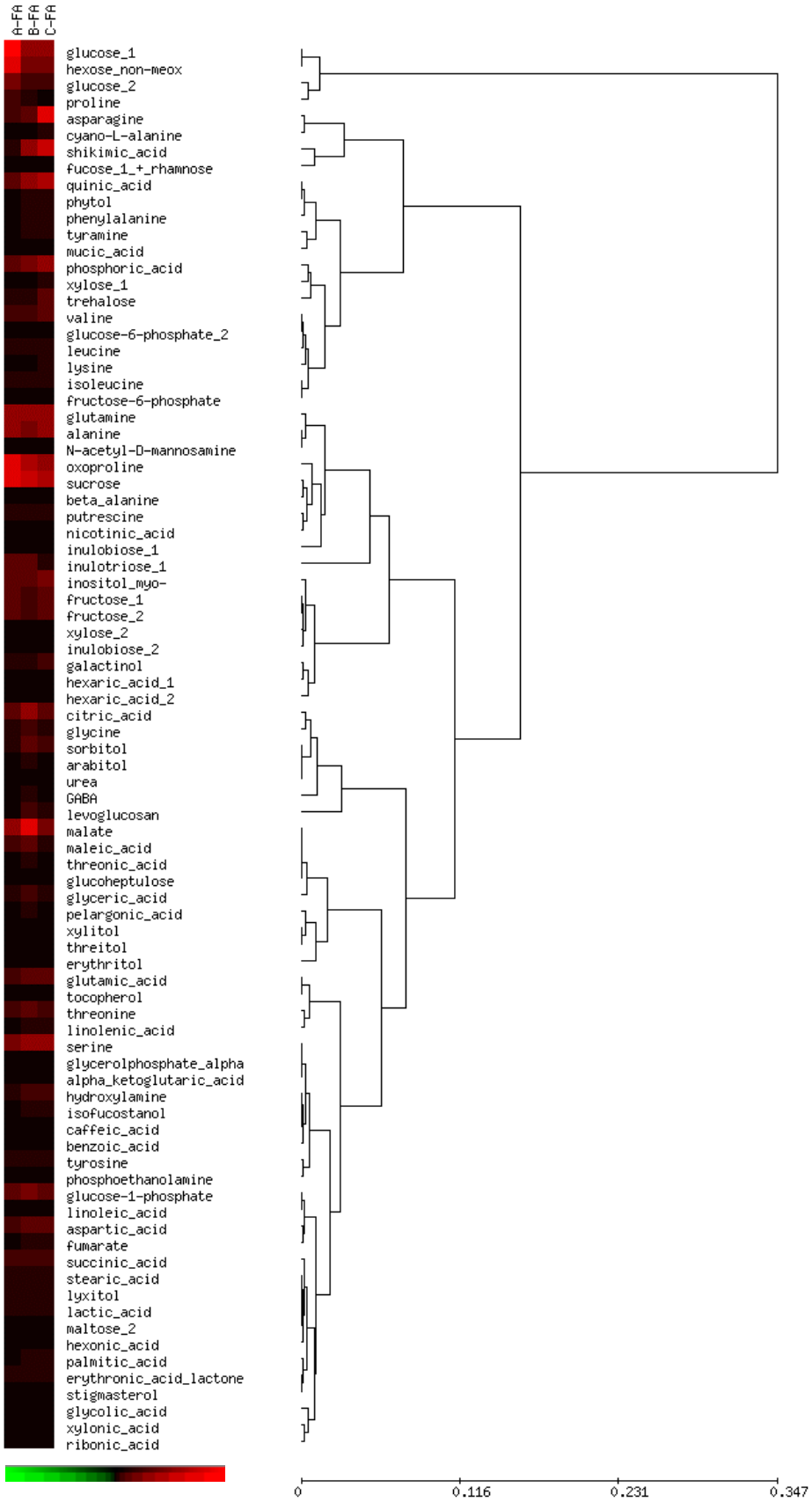


Figure 5: continued from above. B) Paternal genotype and the three harvests (A-Fa: 12.06.2006, B-Fa: 15.08.2006, C-Fa: 19.10.2006).

C)

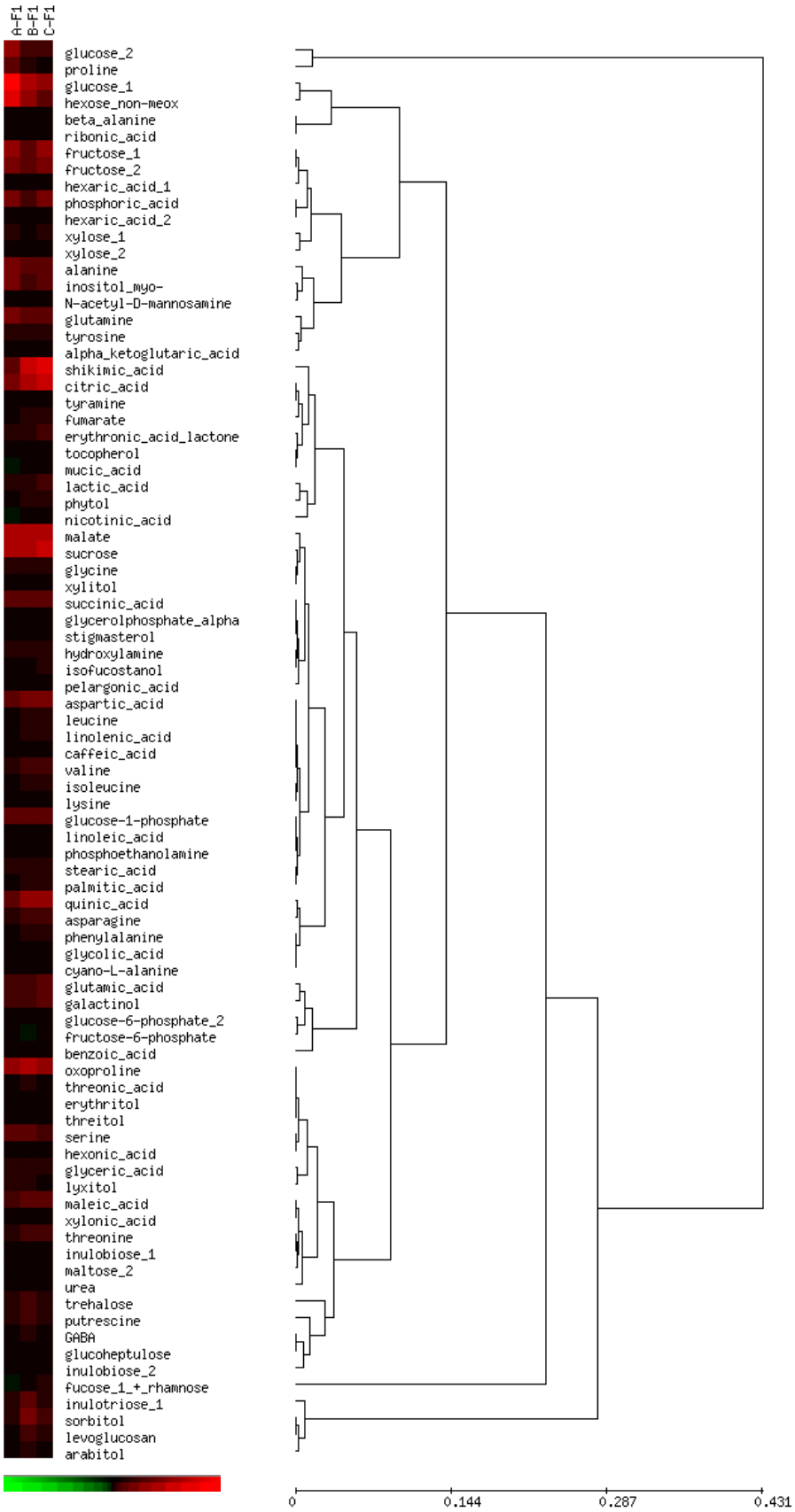


Figure 5: continued from above. C) F_1 genotype and the three harvests (A-F1: 12.06.2006, B-F1: 15.08.2006, C-Fa: 19.10.2006).

Clustering with the programme EPCLUST was performed to try to identify groups of metabolites that had similar activity profiles across the harvests in all three genotypes. It can be seen from Figure 5 A, B and C that there is no clear clustering of metabolites, which are similar across the three genotypes. This may have provided evidence for the identification of groups of metabolites varying as the growing season progresses. However, some metabolites could be identified with similar profiles across the three harvest times that are comparable in all three genotypes. A number of the metabolites showing the greatest differences between harvests and have similar profiles in the three genotypes are discussed. One such metabolite is glucose, which is a principle product of photosynthesis. All three genotypes showed the highest metabolic activity in harvest one (Figure 6), possibly indicating the highest rate of photosynthesis at this time.

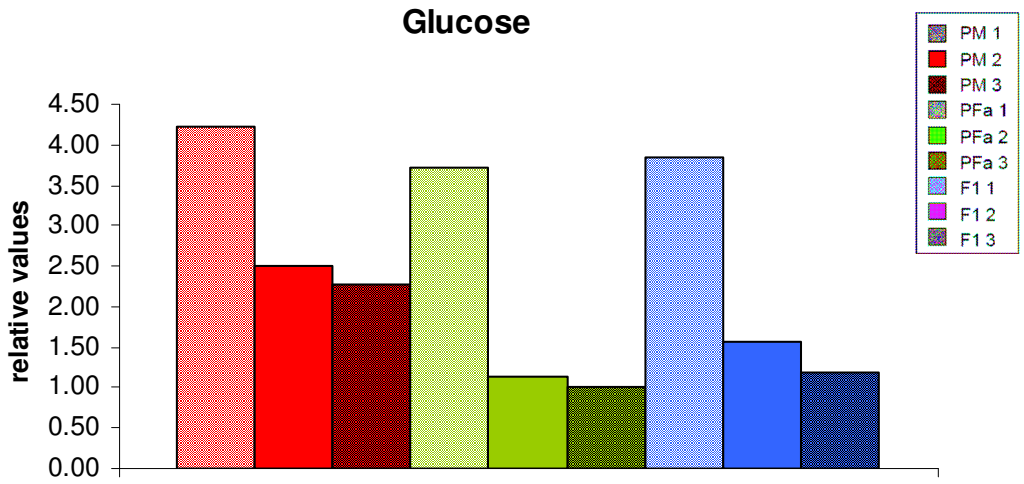


Figure 6: Relative values of glucose over the three genotypes (maternal (PM): red; paternal (PFa): green; F_1: blue) and over the three harvests (1: harvest 12.06.06, 2: harvest 15.08.06, 3: harvest 19.10.06).

Fructose is one of the most important main storage carbohydrates and therefore important for plant growth (Chalmers *et al.* 2005). Activity in the maternal genotype

was decreasing over the harvest times. Paternal and F_1 genotype showed a lower activity during the second harvest (Figure 7).

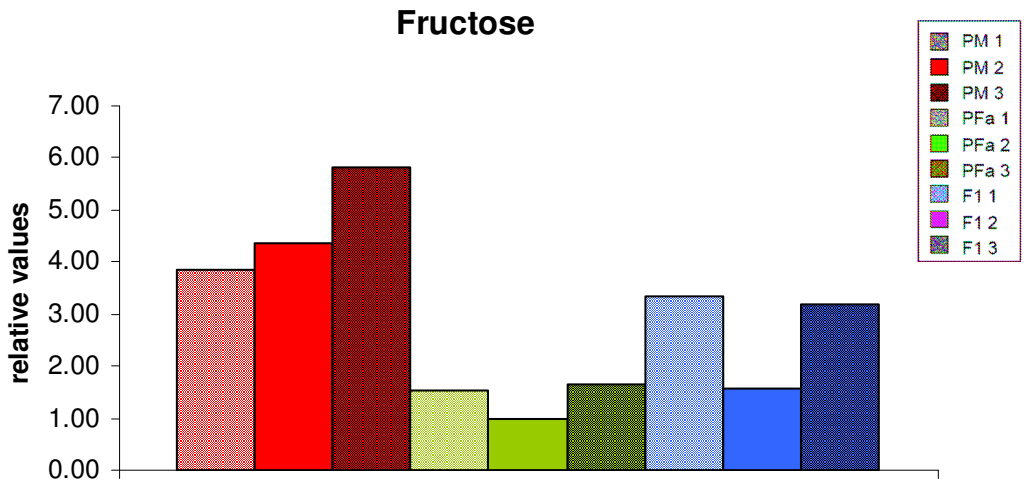


Figure 7: Relative values of fructose over the three genotypes (maternal (PM): red; paternal (PFa): green; F_1: blue) and over the three harvests (1: harvest 12.06.06, 2: harvest 15.08.06, 3: harvest 19.10.06).

Another sugar with similar profile in all genotypes was xylose. Xylose is a monosaccharide involved in biosynthetic pathways of most anionic polysaccharides. Xylose showed the highest activity in the June and in the October harvests (Figure 8).

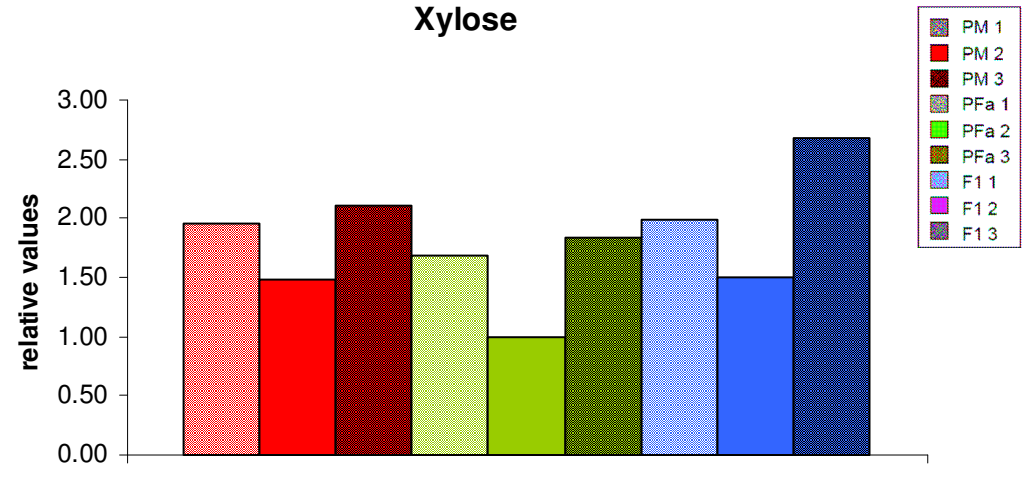


Figure 8: Relative values of xylose over the three genotypes (maternal (PM): red; paternal (PFa): green; F_1: blue) and over the three harvests (1: harvest 12.06.06, 2: harvest 15.08.06, 3: harvest 19.10.06).

Similar profiles were observed for sorbitol, arabitol and shikimic acid across the harvest times and these metabolites have been discussed above (Figure 2).

Although, a number of metabolites with similar profiles could be identified across the harvest points, no group of metabolites was identified that had a clear regulation in the three genotypes that was similar across harvests. This is possibly due to the paternal and maternal lines being inbred in comparison to the F_1, which showed a heterosis effect.

3.4 Discussion

The ultimate aim for a metabolomic study is to understand and predict the behaviour of a complex system. It should be possible to link metabolomic changes in biochemical pathways to the enzymes involved, but so far such results have not been available and current approaches of data mining and mathematical modelling are not suited to the computation of metabolic data (Fiehn 2002). In the present study, metabolites were analysed to find differences in the metabolic profile of the initial generations of the F_2 mapping population. Many metabolic traits are associated closely with yield traits (Schauer and Fernie 2006). Biomass studies are looking for the correlation between biomass and metabolic composition. A biomass study in *Arabidopsis* using RIL and introgression lines found six biomass QTL and 157 metabolite QTL (Lisec *et al.* 2008). Two of the biomass QTL found in the study corresponded with significantly more metabolite QTL than statistically expected.

PCA on all metabolites showed a separation of genotypes indicating differences in the metabolic profiles. This is not surprising because already the phenotypes showed different morphological and growth features (see results chapter 5) like high heterosis

effect in the F_1 genotype, and therefore a different plant architecture than the paternal genotypes which show high inbreeding depression. The level of heterozygosity seems to play an important role in relation to growth rate. Several studies in different species of plants and animals have shown that enzyme genotypes and the level of heterozygosity are related to growth rate (Mitton and Grant 1984). Therefore, levels of heterozygosity are also related to metabolic functions (Rainey *et al.* 1987) and an explanation for the differences between F_1 and parental genotypes. The clustering of the replicated data acts as a quality control for our data set.

Comparing the three genotypes significant differences in the accumulation of specific metabolites could be seen, especially between the F_1 and parental genotypes. Unfortunately, some of the largest differences were found for as yet unknown compounds. This is one of the present limitations of metabolic fingerprinting, in that only a proportion of the compounds are identifiable. Nevertheless, it was possible to see significant differences between a number of known compounds. These included metabolites involved in the citric acid cycle and some sugar alcohols putatively involved in abiotic stress tolerance.

In addition significant differences in the accumulation of amino acids could be seen, particularly between the paternal and maternal plants. It would be interesting to speculate that the paternal plant has a higher nitrogen use efficiency, which could explain the difference in amino acid accumulation. However, there are a number of other reasons that could account for these differences. In the case of many amino acids, the paternal plant actually accumulated these to higher concentration than the F_1, which have been unexpected considering the phenotypes of the plants.

PCA with the total metabolites for each genotype showed a clear separation of harvest times. It would be expected that plants would have variation in their metabolism

during the growing period. A study by Sathish *et al.* 2007 in *L. perenne* took a transcriptomics approach to look at seasonal specific changes in the transcriptome. In that study they succeeded in identifying genes with seasonal specific expression patterns. Metabolomic fingerprinting may be a useful tool to identify seasonal specific changes in the metabolome of *L. perenne*. In the present study the key goal was to identify differences between the metabolome of the three genotypes. However, the sampling points allowed us to take a look at changes in the metabolome at three different points, two during the high growth period and one towards the end of the growth curve. For this we only looked at known metabolites. Hierarchical clustering with each genotype independently did not identify clear clusters of metabolites with similar profiles across the three harvests in all three genotypes. The reason for the high variability of metabolites across harvests between the genotypes may be due to the phenotypes of the plants. Both the paternal and the maternal genotypes are inbred lines showing a degree of inbreeding depression in comparison to the F_1 showing strong heterosis effect. A study looking at some high yielding genotypes across the seasons may lead to the identification of season specific metabolic fingerprints.

Metabolic fingerprinting is a powerful tool to help understand complex traits of agronomic importance. The present study identified a number of metabolites showing significant differences between the genotypes. However, a large number of these metabolites remain unknown, which is a current limitation of metabolic fingerprinting. . An interesting future prospect would be to map these metabolites as QTL in the F_2 population and relate them to QTL for biomass yield identified in chapter five.

Chapter 4: Genetic map

4.1. Introduction

The advent of DNA based molecular markers has led to the creation of dense genetic linkage maps, which have greatly benefited both basic and applied research. In terms of basic research, they have allowed us to gain insight into the structure of plant genomes and understand the relationship between the genomes of different species (Jones and Pašakinskienė 2005). Applied research deals with the creation of linkage maps in populations, allowing complex traits of agricultural importance to be dissected at the genetic level (Paterson *et al*. 1988, Stuber *et al*. 1992). In grasses one of the first *Lolium* maps was developed by Hayward *et al*. (1994). The first pure *Lolium perenne* map was developed by Bert *et al.* (1999) and was based entirely on amplified fragment length polymorphism (AFLP) markers. The first map based on simple sequence repeats (SSR) markers for *L. perenne* was constructed by Jones *et al*. (2002a).

4.1.1 Linkage mapping

A linkage map is a chromosome map of a species or experimental population that shows the position of genes and/or markers relative to each other based on recombination frequencies. Genetic linkage occurs when particular genetic loci or alleles for genes are inherited jointly. If genetic loci on the same chromosome are physically connected they co-segregate during meiosis, and are thus genetically linked. Linkage group (LG) analysis is to place loci into linkage groups based on their linkage relationships. Biologically, a LG is defined as a group of genes with their loci located on the same chromosome (Liu 1998). The linkage between loci is usually calculated using the logarithm of odds (LOD) scores. The LOD score commonly used is based on

the G^2 statistic for independence in a two-way contingency table in Join Map® V3.0 software (Van Ooijen and Voorrips 2001):

$$G^2 = 2 \sum o \log(o/e)$$

o is the observed and *e* the expected number of individuals in a cell, *log* the natural logarithm, and $\sum$ the sum over all cells. LOD scores might be normally affected by segregation distortion (SD). That is not the case for the test for independence in JoinMap® V3.0. LOD values greater than three are typically used to construct linkage maps because it means that linkage at θ (practical recombination fraction) = $\hat{\theta}$ is 1000 times more likely than at $\theta = 0.5$ (Liu 1998).

Different mapping functions can be used to convert recombination fractions into centiMorgan (cM). Two mapping functions are commonly used. The Kosambi mapping function (Kosambi 1944) assumes that recombination events influence the occurrence of adjacent recombination events,

$$K(d) = \frac{1}{2} \tanh(2d)$$

with *d* the genetic map distance and *tanh* as hyperbolic tangent. The Haldane mapping function (Haldane 1919) predicts the number of crossovers from the recombination frequencies. It is a more additive measure and assumes no interference between crossover events,

$$H(d) = \frac{1}{2} (1-e^{-2d})$$

with *d* the genetic map distance and *e* as the natural logarithm.

Loci that have a recombination frequency of 50% are described as unlinked and assumed to be located far apart on the same chromosome or located on a different chromosome. The lower the frequency of recombination between two loci the closer they are situated on a chromosome. Many different programmes are available to create a

linkage map, e.g. MapManager QTX (Manly *et al*. 2001), which can be used with Kosambi or Haldane mapping function or MapMaker/EXP (Lincoln *et al*. 1992) using the Haldane mapping function. In the present study Join Map® V3.0 software was used, allowing a choice between Kosambi or Haldane mapping functions (Van Ooijen and Vorrips 2001).

4.1.2 Molecular markers

Molecular markers are essential tools for genetic mapping, genotype fingerprinting, population structure and genetic diversity studies (Saha *et al*. 2006). The development of molecular markers along with high-throughput genotyping instruments has enhanced the application of molecular markers in crop improvement with the result that molecular markers have since been widely used for cultivar identification, parental analysis, genome mapping and tagging of genetically important traits (Saha *et al*. 2005). Various marker types have been used to construct linkage maps. The most common markers are:

Random amplified polymorphic DNA (RAPD; Williams *et al*. 1990): RAPD markers derive from amplification with a single random marker (usually 8 to 12 base), which will anneal with complementary sequence at undetermined positions in the genome.

AFLP (Vos *et al*. 1995): a genomic DNA sample is digested with two enzymes followed by ligation of double stranded enzyme adaptors. Preselective amplification with markers corresponding to the adaptors and part of the restriction sites are used to amplify fragments. This step is followed by selective amplification with labelled markers and the fragments are visualised by running on polyacrylamide gels.

Restriction fragment length polymorphism (RFLP; Botstein *et al*. 1980): RFLPs take advantage of polymorphisms within restriction sites throughout the genome. Genomic DNA is digested with various restriction enzymes and hybridized with labelled probes either chosen at random or representing specific sequences. The marker is specific to a single clone/restriction enzyme combination.

Cleaved amplified polymorphic sequence (CAPS; Michaels and Amasino 1998): CAPS rely on polymorphism within restriction sites. Sequences are amplified with specific markers and digested with a restriction enzyme to identify polymorphism.

SSR (Jeffreys *et al*. 1985): SSRs or microsatellites are polymorphic DNA marker comprised of mono-nucleotides, di-nucleotides, tri-nucleotides or tetra-nucleotides that are repeated in tandem arrays and distributed throughout the genome. The best studied are CA (alternatively GT) dinucleotide repeats.

RAPD and AFLP are examples for dominant molecular markers and RFLP, CAPS, SSRs are examples of co-dominant molecular markers. Co-dominant markers have the advantage compared to dominant markers that they can discriminate between homozygotes and heterozygote genotypes. The advantage of RAPD and AFLP markers is that multiple loci can be detected with one single polymerase chain reaction (PCR). A disadvantage of AFLP as well of RAPD molecular markers is that the reproducibility is considered to be low (Collard *et al*. 2005). Another disadvantage of RFLP markers are that they are time consuming and a large amount of DNA is needed. RFLP and SSR are reliable markers and are transferable between species and populations (Jones *et al*. 1997). SSR, CAPS and RAPD are technically simple and require only a small amount of DNA. The disadvantage of SSR markers is their time consuming development but as soon as they are developed the SSR markers can be used easily and for a variety of purposes and often across different species (Saha *et al*. 2006).

4.1.3 Population structure for mapping studies

The most common mapping population structures are backcrosses or F_2 genotypes. These structures simplify mapping because genetic segregation is the result of the meiotic recombination from a single F_1 genotype. Therefore, only two alleles per locus segregate in a mapping population if inbred lines are used as parental lines (Warnke *et al*. 2004). In some cases F_1 populations based on two heterozygous parents are used to construct genetic maps. This has the advantage that parents for the mapping population can be selected solely based on the phenotype for the target trait. Another advantage of such a population is that two genetic maps for both parental lines can be constructed when marker data are scored dominantly (Faville *et al*. 2004, Cogan *et al*. 2006). The objective in this study was to construct a genetic map of a F_2 segregating population, which is the basis of a recombinant inbred line (RIL) population. RILs are constructed by crossing two inbred lines followed by repeated selfing. In this way new inbred lines can be developed with genomes which show a mosaic of the parental genomes (Broman 2005). These RILs are a powerful tool for mapping studies because (1) they enable cloning of genes due to their homozygote character and (2) they are stable lines, making it easy to select for a trait of interest between the single lines.

4.1.4 Breeding and mapping in Lolium perenne

Conventional breeding has already been successful in generating commercial varieties of forage grasses with traits for enhanced agricultural sustainability. Breeding objectives in grasses focus on stress resistance against drought, cold and pathogens, and on agronomic traits like nutrient use efficiency, carbohydrate content, fatty acid content, winter survival, flowering time and biomass yield (Humphreys *et al*. 2005). To develop successful varieties and breeding lines it is important to have freely segregating

breeding populations. The planning of breeding programmes can be greatly enhanced with knowledge of genome organization of the species and a genetic map. MAS is based upon the establishment of a linkage relationship between an easily identifiable major gene marker and a character of agronomic importance. Therefore, it may be more efficient to select in a breeding programme for the marker than for the trait itself. This method involves the development of detailed genetic maps, which can be used for determination of the location of QTL (Hayward *et al*. 1994).

The objective of the work described in this chapter was to develop a predominantly SSR based genetic linkage map of *L. perenne* and to compare it to existing maps. There are three main steps for the construction of a linkage map: (1) Development of a mapping population; (2) identification of polymorphisms within the population; (3) linkage analysis of markers. The linkage map will be utilised in subsequent chapters for the identification of QTL for traits of agricultural significance.

4.2 Material and Methods

4.2.1 Plant Material

The F_2 mapping population, which was used for the development of the genetic map, was constructed from a cross between two inbred *Lolium* lines (PM and PFA, described in chapter two) to construct RILs (Figure 1).

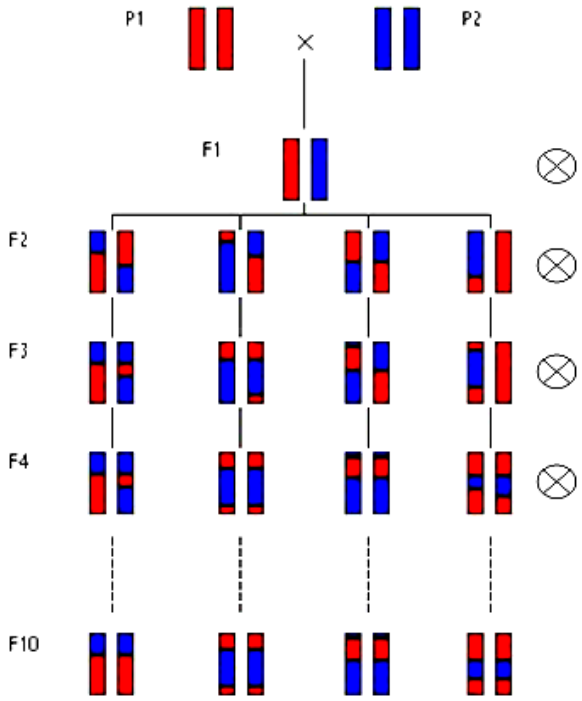


Figure 1: Breeding scheme for Recombinant Inbred Lines (RILs). P_1: maternal (PM), P_2: paternal line (Pfa). ⊗ : selfed.

The single F_1 plants were raised and self-pollinated by bagging in cellophane pollination bags to generate an independent F_2 population. One independent F_2 population was chosen and 400 F_2 individuals of this population were raised in the greenhouse. 360 genotypes were finally randomly selected for the map construction.

4.2.2 Total DNA extraction

Total genomic DNA was isolated from the parental *L. perenne* inbred lines, the F_1 and F_2 genotypes. Plant material was frozen in liquid nitrogen and DNA isolated with a Cetyltrimethylammoniumbromid (CTAB) method after Doyle and Doyle (1987):

5 g of fresh plant material was flash frozen in liquid nitrogen. The leaf material was ground with a pestle and mortar until a fine powder was obtained. The powder was collected in 50 ml Falcon tubes. 5 ml 2x CTAB buffer (2% CTAB, 200 mM Tris pH 8, 20 mM ethylenediaminetetraacetic acid (EDTA) pH 8, 1.4 M sodium chloride, 1%

polyvinylpyrrolidone, 0.28 M 2-mercaptoethanol) and 10 ml 1x CTAB buffer were added to the powder. The mixture was merged well and incubated for 1.5 hours in a water bath at 65°C and was shaken gently in between the incubation time. The powder/CTAB mix was cooled to 25ºC and 10 ml of chloroform : isoamyl alcohol (24:1) were added. The solution was incubated on a belly dancer for at least 60 minutes and centrifuged at 3750 rpm for 15 minutes. The top phase was transferred slowly into a new Falcon tube. The chloroform cleaning step was repeated with 4 ml chloroform : isoamyl alcohol (24:1). 75 µl RNase (10 mg/ml) were added to the solution and the solution was incubated for 30 minutes. Ice cold 4 ml isopropanol were slowly added to the incubation mix. The Falcon tube was gently inverted until the DNA precipitated. A sterile Pasteur pipette was used to hook up the DNA. An Eppendorf tube was filled with 1 ml ethanol-sodium acetate washing solution (76% ethanol, 0.2 M sodium acetate) and the DNA pellet hooked with the pipette and was transferred for at least 20 minutes into the ethanol solution. For a very short time the DNA was transferred for a further washing step into a second Eppendorf tube filled with 1 ml ethanol ammonium acetate (76% ethanol, 10 mM ammonium acetate). The DNA pellet on the Pasteur pipette was transferred into 1 ml 0.1x TE buffer (10 mM Tris, 1 mM EDTA pH 7.5) and was solved. The DNA concentration was determined with a spectrophotometer at a wavelength of 260 nm (Sambrook and Russell 2001).

4.2.3 SSR and AFLP marker testing and analysis

SSR markers were chosen from a number of public and non-public sources (Jensen *et al*. 2005a; Gill *et al*. 2006; Jones *et al*. 2001; Kubik *et al*. 2001; Lauvergeat *et al*. 2005; Studer *et al*. 2008; Warnke *et al*. 2004).

Tall fescue SSRs from the Robert Samuel Noble Foundation, Oklahoma, USA were optimised for amplification of *L. perenne* DNA, and cross-species amplifying and polymorphic SSRs used for further mapping work.

All SSR markers were screened for polymorphism in the inbred parental lines and F_1 line.

PCR reactions were performed in a total volume of 10 µl with 25 ng total genomic DNA as template, 2.5 µM forward and reverse primer of the marker (Applied Biosystems, Warrington, UK), 0.3 units of DNA Taq polymerase (New England Biolabs, Ipswich) and 2 mM dNTPs in a Biometra Thermocycler. PCR programmes varied for each marker sources (Table 1).

Table 1: Polymerase chain reaction programmes for simple sequence repeats markers used for the construction of the genetic map from different sources. Table displays marker source, annealing temperature, duration, cycles, and marker name.

Marker source	Temp.	Time	Cycles	Marker
Noble Foundation, *USA**[1]	95ºC	5 min	Initial denaturation	NFFa017, NFFa036, NFFa136, NFFa142, NFFa155
	95ºC	1 min		
	AT*	1 min	35 cycles	
	72ºC	1 min		
	72ºC	10 min	Final extension	
CRC, *Australia**[2]	94ºC	1 min		LpSSRH02F01, LpSSRH11G05, LpSSRK12E06, LpSSRK14F12
	AT*	30 s	10 cycles (touch	
	72ºC	1 min	down -1ºC)	
	94ºC	1 min		
	AT*	30 s	30 cycles	
	72ºC	30 s		
IGER, *UK**	96ºC	5 min	Initial denaturation	LpACT15H3, LpACT44A7, LpACTR1C5, LpHCA18A2b, LpHCA18B12
	96ºC	15 s		
	AT*	30 s	35 cycles	
	72ºC	30 s		
	72ºC	4 min	Final extension	

Marker source	Temp.	Time	Cycles	Marker
ViaLactia,				rv0029, rv0062, rv0068,
*New Zealand** [3]	95ºC	10 min	Initial denaturation	rv0134, rv0171, rv0188,
	94ºC	1 min		rv0196, rv0252, rv0262,
	AT*	30 s	10 cycles (touch	rv0307, rv0327, rv0342,
	72ºC	1 min	down -1ºC)	rv0360, rv0380, rv0562,
	94ºC	30 s		rv0674, rv0717, rv0739,
	AT*	30 s	25 cycles	rv0863, rv0983, rv1024,
	72ºC	30 s		rv1046, rv1117, rv1131,
	72ºC	10 min	Final extension	rv1133, rv1139, rv1188,
				rv1269, rv1411, rv1412
Kubik,	94ºC	1 min	Initial denaturation	DLF025, LpSSR020,
DIAS, *Denmark*	94ºC	1 min		LpSSR027, M16B, M15185,
DvP, *Belgium*	AT*	1 min	30 cycles	PR8, PR14, PRG, Rye014
DLF, *Denmark*[4]	72ºC	2 min		
	72ºC	10 min	Final extension	
Lauvergeat[5]	94ºC	3 min	Initial denaturation	B1A2, B1B6, B1C8, B3A3,
	94ºC	30 s		B3B8
	AT*	30 s	10 cycles (touch	
	72ºC	30 s	down -1ºC)	
	94ºC	30 s		
	AT*	30 s	30 cycles	
	72ºC	30 s		
	72ºC	5 min	Final extension	
Studer[6]	94ºC	5 min	Initial denaturation	G04_002, G04_030,
	94ºC	30 s		G04_043, G04_054,
	AT*	1 min	12 cycles (touch	G04_059, G04_072,
	72ºC	1 min	down -1ºC)	G04_099
	94ºC	30 s		
	AT*	1 min	30 cycles	
	72ºC	1 min		
	72 ºC	5 min	Final extension	

Marker origin via license agreement (*) or from articles: [1]Warnke *et al*. 2004; [2]Jones *et al*. 2001; [3]Gill *et al*. 2006; [4]Kubik *et al*. 2001, Jensen *et al*. 2005a; [5]Lauvergeat *et al*. 2005; [6] Studer *et al*. 2008

*AT = marker specific annealing temperature

The SSR forward primer of the markers were fluorescently 5'-labelled with 6FAM™, VIC®, NED™ or PET™ (Applied Biosystems, Warrington, UK) and GeneScan™500 LIZ® was used as internal sizing standard.

EcoACAMseCAC, EcoAGCMseCTA, EcoACAMseCTA and EcoACTMseCTA marker combinations were used to add additional loci to the F_2 map. The AFLP forward primer of the marker was either fluorescently 5'-labelled with FAM™, JOE™ or NED™ and the internal sizing standard GeneScan™500 ROX were used. The AFLP marker procedure was carried out following the Applied Biosystems protocol for AFLP™ Plant Mapping with a modification in the sample dilution of the preselective amplification product to a 1:2 dilution ($TE_{0.1}$ buffer:product).

SSR and AFLP genotyping was performed on an ABI Prism 3100 genetic analyser (Applied Biosystems, Warrington, UK) with POP-4 polymer and 36 cm capillaries. Amplification patterns were scored using GeneMapper® V3.7 software (Applied Biosystems, Warrington, UK).

4.2.4 Genetic map construction

SSR markers for the F_2 population were scored as co-dominant markers and for each allele a letter was assigned in accordance with the coding required for JoinMap® Version 3.0. AFLP markers were scored as dominantly for the presence or absence of a band. Markers were classified in two segregation types. Co-dominant markers had bands present in both parents and were expected to segregate in a 1:2:1 pattern; dominant markers were expected to segregate in a 1:3 pattern. Map construction with the SSR and AFLP marker data were performed with the software package JoinMap® Version 3.0 (Van Ooijen and Voorrips 2001). For LG calculations and determination, a LOD threshold of not lower than 4.0 was used. The calculation of the map LOD

threshold was larger than 1.0 and a jump threshold in goodness-of-fit of 5.0 were used. Kosambi's mapping function was applied to estimate genetic distances in cM. Markers on LGs were positioned with JoinMap® Version 3.0. The genetic map was drawn using MapChart Version 2.2 software (Voorrips 2002).

4.2.5 Segregation distortion

SD is defined as the deviation of genetic segregation ratios from their expected Mendelian fraction (Lyttle 1991). Expected Mendelian segregation ratios of SSR and AFLP markers were analysed using χ^2-square tests in Join Map® Version 3.0 software. If the homozygote alleles (PM, PFa) and the heterozygote allele (F_1) differed from their Mendelian segregation ratios (1:2:1; homozygote alleles: 25%; heterozygote allele: 50%), in other words: the percentage of alleles where higher than expected, they were assigned as skewed loci towards the favoured allele. The segregation of each marker was displayed in graphs for each LG. The map in cM was plotted against the percentage of allele for each loci over the whole population (Figure 6).

4.3 Results

4.3.1 AFLP and SSR testing and detection

SSR markers were tested for their polymorphism for both maternal and paternal lines (Table 2).

Table 2: Labelled SSR markers. Table displays number, polymorphism, and homozygosity and heterozygosity in the parental lines (PM: maternal, PFa: paternal).

Marker source	Total number of tested markers	Amplified markers	Amount of mono-morphic markers	Amount of poly-morphic markers	amplified markers in %	Poly-morphism in %	Number of homozygote markers		homozygote markers in %	
							PM	PFa	PM	PFa
Noble Foundation, *USA**[1]	28	20	13	7	71.4	35	17	16	85	80
CRC, *Australia**[2]	31	26	20	6	83.9	23.1	22	23	84.6	88.5
IGER, *UK**	27	12	3	9	44.4	75	6	8	50	66.7
ViaLactia, *New Zealand**[3]	109	71	34	37	65.1	52.1	62	61	87.3	85.9
Kubik *et al*. 2001	16	12	5	7	75.0	58.3	10	10	83.3	83.3
DIAS, *Denmark*[4]	12	9	6	3	75	33.3	9	9	100	100
DvP, *Belgium*[4]	3	2	1	1	66.7	50	2	1	100	50
DLF, *Denmark*[4]	3	3	2	1	100	33.3	3	3	100	100
Faville *et al*. 2004	5	-	-	-	-	-	-	-	-	-
Lauvergeat *et al*. 2005	13	12	6	6	92.3	50	12	12	100	100
Studer *et al*. 2008	20	17	10	7	85	41.2	17	17	100	100
Total	267	184	100	84	68.9	45.1	160	160	87	87

Marker origin via license agreement (*) or from articles: [1]Warnke *et al*. 2004; [2]Jones *et al*. 2001; [3]Gill *et al*. 2006, [4]Jensen *et al*. 2005a

Both, maternal and paternal inbred lines showed a homozygosity in markers of 87%, which is for an allogamous species a good range compared to an autogamous species. In general it is expected to be higher in inbred lines. Lower homozygosity levels than expected were reported by Jones *et al.* (2003) in inbred *Trifolium repens* (outbreeding species). Their findings in the inbred parents were 45% (fourth generation of inbreeding) and 40% (fifth generation of inbreeding). The expected would have been 94% and 97% homozygosity (Jones *et al.* 2003). In the current work the homozygosity for single marker groups from the different sources ranged between 85% and 100% except for markers sourced from IGER and DvP (Table 2). From DvP only three markers were tested and the numbers were therefore not representative. The IGER markers showed a high number of non amplified markers (56%), which could explain the imbalance in the homozygosity measurements. The polymorphism degree between the marker group sources ranged between 23% and 75%.

Thirty polymorphic AFLP markers were selected from the enzyme-marker combinations EcoACAMseCAC, EcoAGCMseCTA, EcoACAMseCTA and EcoACTMseCTA. Only ten of the AFLP markers were used for the construction of the F_2 inbred line derived genetic map. The other 20 markers had too high degree of SD to link them to any other marker on the LGs.

Additionally to the *Lolium* SSR markers, tall fescue SSRs from the Robert Samuel Noble Foundation, Oklahoma, USA were optimised for this study. 157 unlabelled SSR markers were tested. 28 were chosen to be labelled (Figure 2). Seven labelled markers were polymorphic and only five could be mapped on the genetic map. A low polymorphism was found in the mapping population and showed that the cross-species use of these markers did not work very successfully.

In total 267 labelled SSR markers were screened for polymorphism in the inbred parental lines and F_1 line (Figure 3, Table 2).

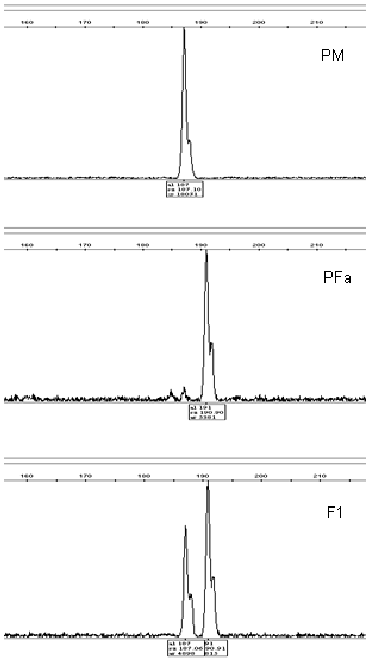


Figure 2: Tall fescue simple sequence repeats marker testing was performed on an ABI Prism 3100 genetic analyser (Applied Biosystems, Warrington, UK). Amplification patterns were scored using GeneMapper® V3.7 software (Applied Biosystems, Warrington, UK). Maternal (PM), paternal (PFa) and F_1 DNA was used as template for the marker testing.

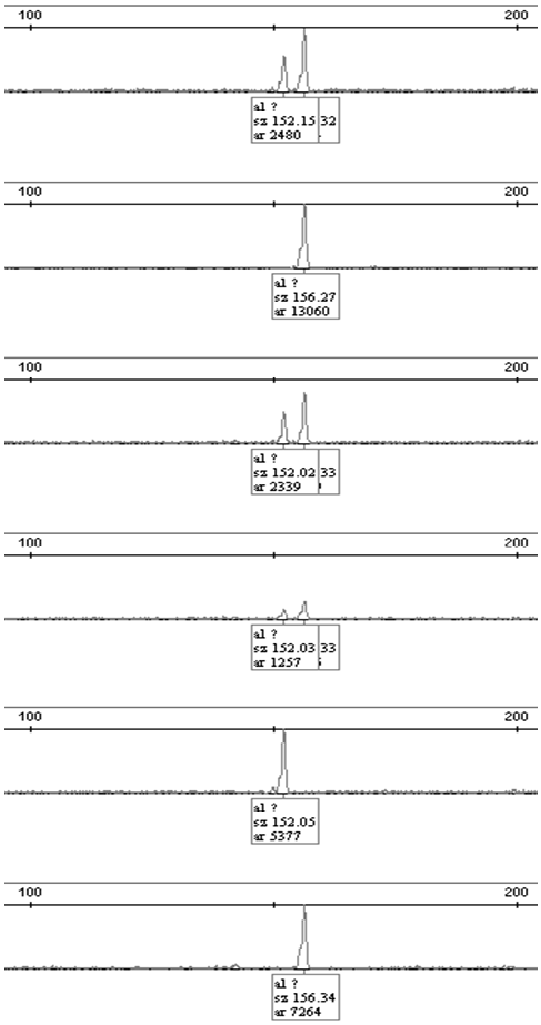


Figure 3: Simple sequence repeats genotyping between the F_2 genotypes was performed on an ABI Prism 3100 genetic analyser (Applied Biosystems, Warrington, UK). Amplification patterns were scored using GeneMapper® V3.7 software (Applied Biosystems, Warrington, UK). Basepair size of alleles in the F_2 genotypes: maternal genotype (PM) = 152 bp; paternal genotype (PFa) = 156 bp; F_1 = 152 bp, 156 bp.

4.3.2 Genetic map construction

65 out of the 84 polymorphic SSR markers and 10 out of the 30 AFLP markers could be used for the construction of the genetic linkage map. The 25 non-mapped markers (SSR and AFLP) had either a too high degree of SD or were positioned too distantly in relation to the next markers assigned to linkage groups and were presumably on the distal ends of LG 1 and 6 when the non-mapped markers were compared to the marker locations of the framework linkage map of Gill *et al.* (2006). AFLP markers had

their mapping positions in general on the distal ends of the LGs in the F_2 population and SSR markers were clustered around the centromeric regions (Figure 4 a). Seven LGs were detected for the genetic linkage maps using the software package Join Map V3.0 (Figure 4 a, b). The LGs were identified and the LG numbers assigned with the help of the framework map of Gill *et al*. (2006). The LOD scores of the genetic map for the grouping of markers ranged between LOD scores of four and six. The total map length was 592.3 cM with an average marker density of 8 cM. The biggest gap among markers was found on LG 1 with 51 cM between the two markers NFFa155 and LpSSR027 and the smallest gap was found on LG 3 with 0.6 cM between markers EacaMcac-433 and LpACT44A7 (Figure 4 a). A linkage map consisting only of SSR markers was constructed to see if there were any differences in the marker order without AFLPs. The marker loci on the LGs remained in the same order, except on LG 2 (marker G04_030 and rv0062) and LG 3 (marker LPSSRH02F01, B1A2 and LPSSRK14F12. Marker rv0863 could not be added on the map without the AFLPs) (Figure 4 b).

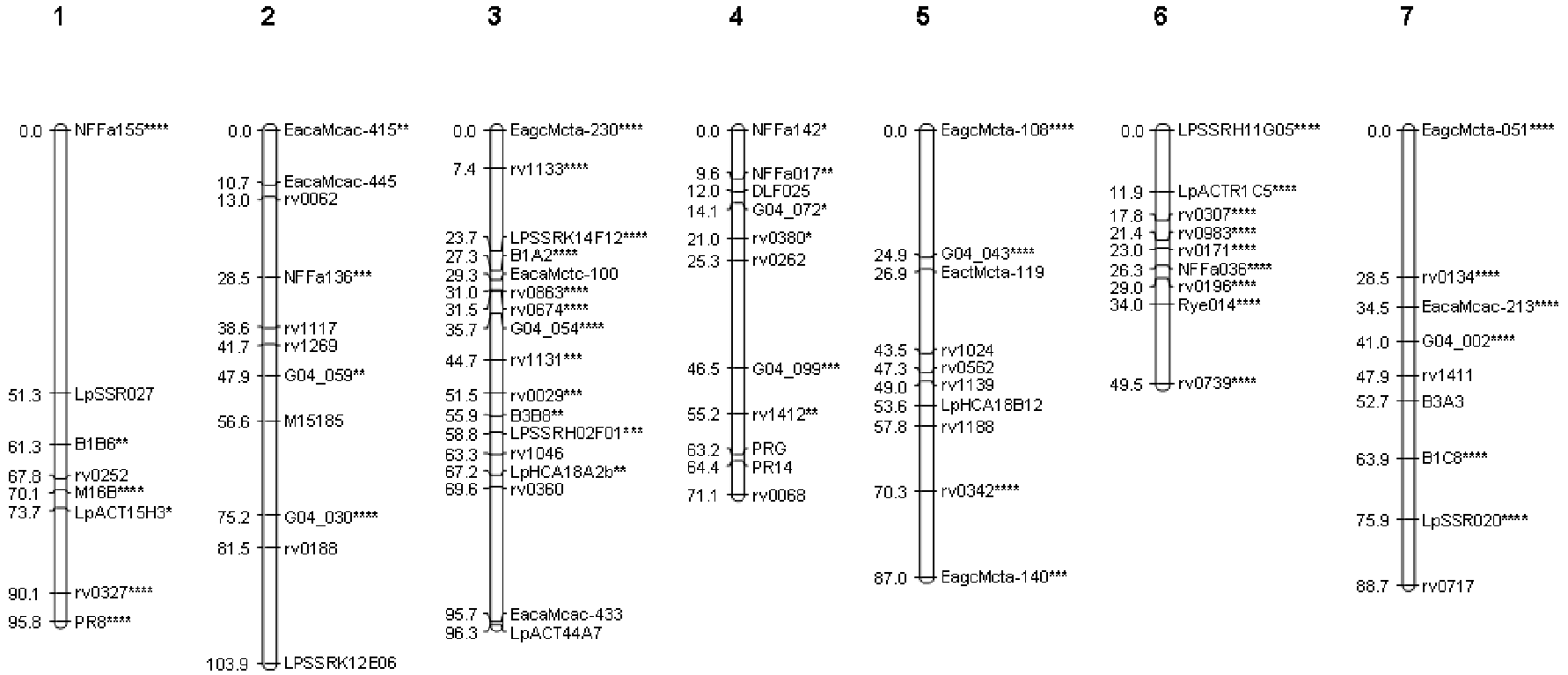


Figure 4 a: Genetic map generated in Join Map V3.0 using Kosambi's mapping function. Distances are given in cM. Asterisks indicate segregation distortion ratios (p<0.05 *, p<0.01 **, p<0.001 ***, p<0.0001 ****).

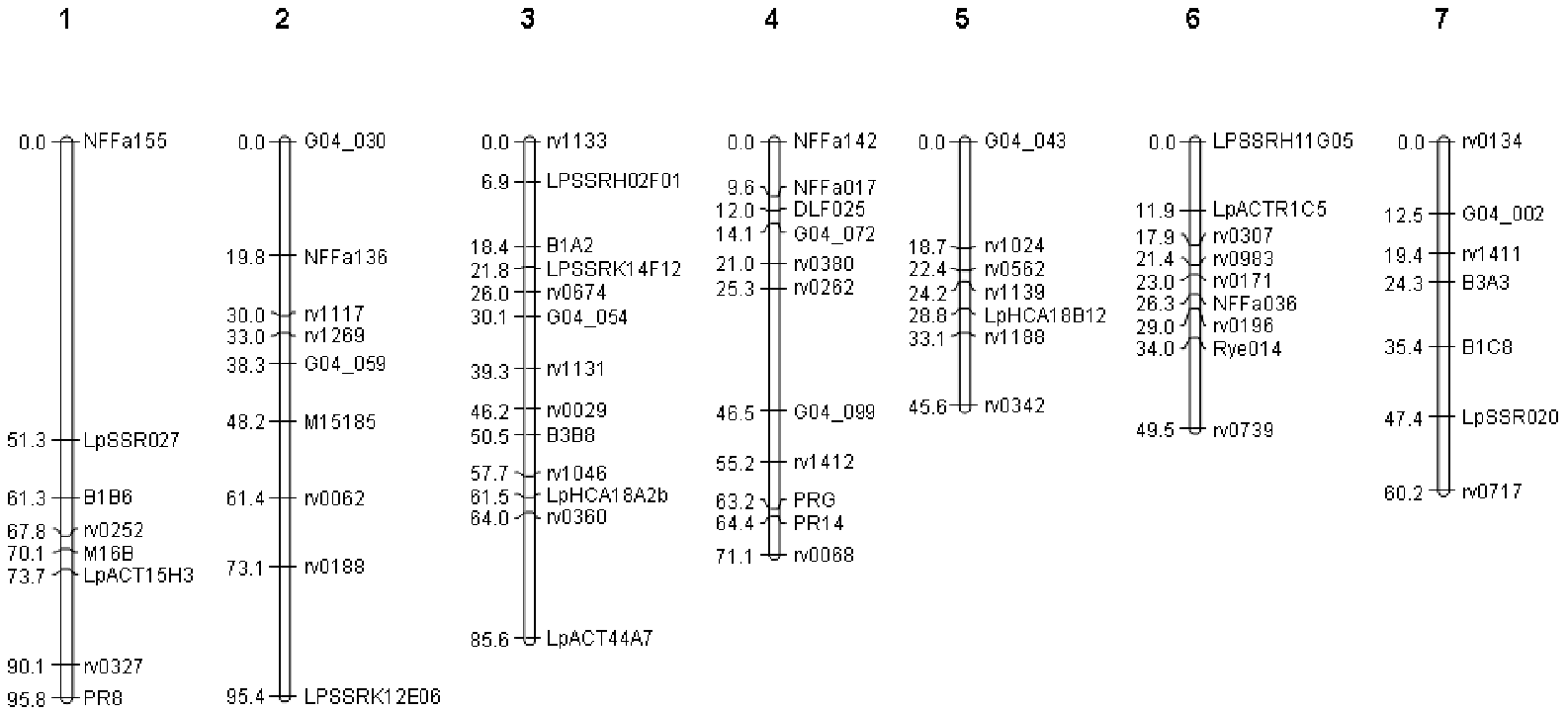


Figure 4 b: Genetic map without AFLP markers generated in Join Map V3.0 using Kosambi's mapping function. Distances are given in cM.

4.3.3 Segregation distortion

In the F_2 population, 47 (63%) out of 75 mapped loci (SSR and AFLP) showed significant (p<0.05) SD. Significance values for SD differed across the 47 distorted loci: Four loci had a p-value of <0.05, seven loci of <0.01, six loci of <0.001 and 30 loci of <0.0001. Out of these 47 loci, ten loci favoured alleles of the maternal parent (PM), 18 of the paternal parent (PFa) and one the F_1 allele. 18 loci were skewed in two allelic directions of the genotype classes at the same time. Eleven loci favoured the PM and the F_1 allele, six loci favoured PFa and F_1 and one locus favoured the PM and the PFa allele at the same time (Table 3).

Table 3: Table shows the distribution of segregation distortion (SD) ratios of distorted marker loci, which were skewed in favour of one or two genotype classes. PM (maternal), PFa (paternal)

SD ratio	Favoured genotype classes						
	PM	PFa	F_1	PM / F_1	PFa / F_1	PM / PFa	Total
p<0.05	2	2	-	-	-	-	4
p<0.01	1	4	-	2	-	-	7
p<0.001	1	2	-	3	-	-	6
p<0.0001	6	10	1	6	6	1	30
Total	10	18	1	11	6	1	47

LGs 1, 3, 5, 6 and 7 had the highest amount of SD (Figure 4 a, Table 3). Figure 5 shows alleles of each locus in F_2 plants, which are plotted along the genetic linkage map. LG 6 was completely distorted and alleles of the maternal line were favoured. LG 2 and 4 showed the lowest SD. LG 5 was distorted on both ends with a centromeric non-distorted part and LG 2 only distorted on one end of the LG (Figure 4 a and 5).

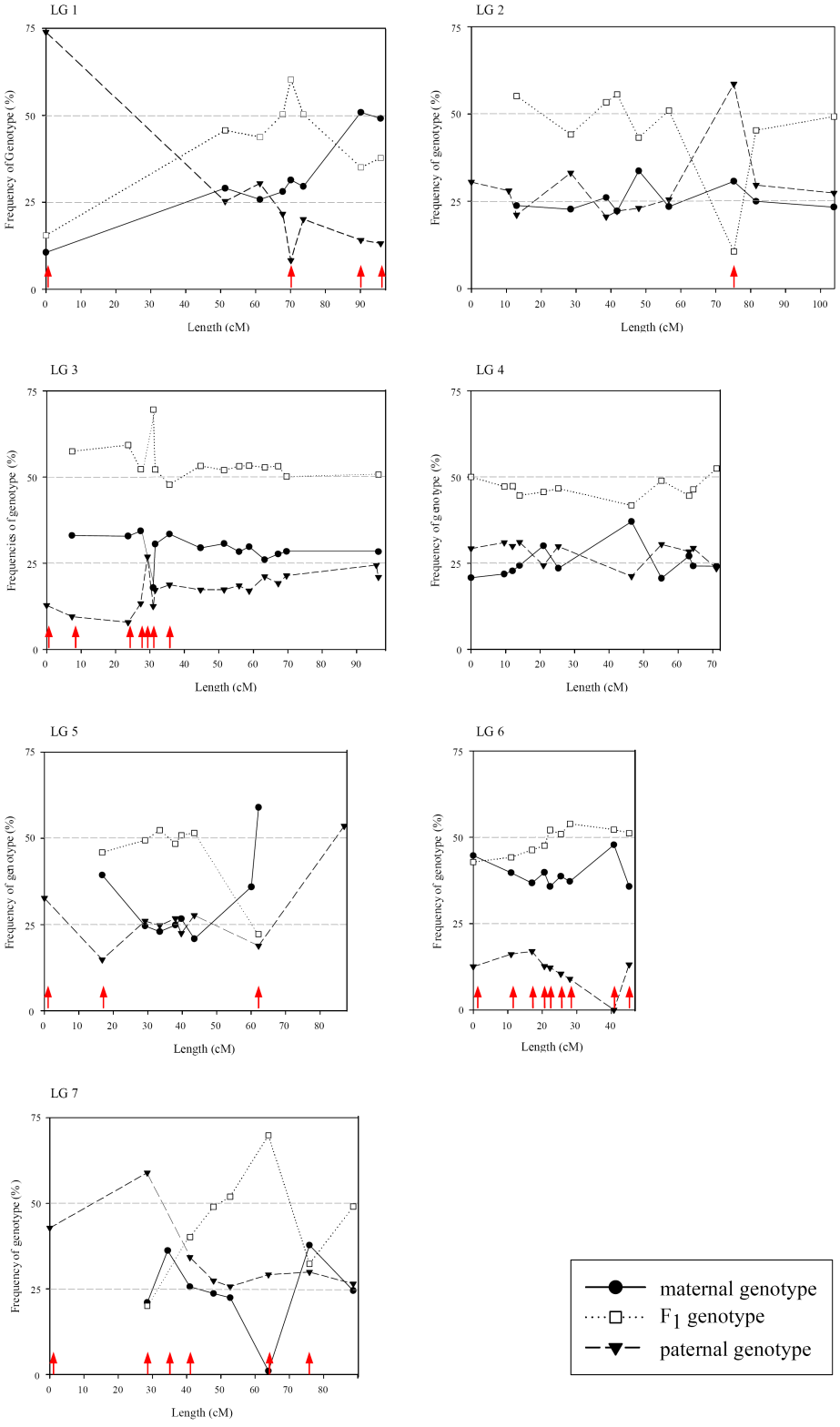


Figure 5: Frequencies of each allele in F_2 plants are plotted along the genetic linkage map; Position at 0 cM corresponds to top arm of each linkage group (LG). Theoretically without segregation distortion the homozygous genotype classes each account for 25% and the heterozygous genotype class for 50% of segregation. Arrows (↑) show, which marker was segregated distorted (*p*-value of <0.0001) on the LGs.

4.4 Discussion

The present study sets out to construct a genetic linkage map in an F_2 inbred line derived population. Genetic linkage maps are important tools for fine mapping and cloning of genes. This map is a prerequisite to carrying out QTL studies described later in the thesis. Therefore, it was important to develop a reliable map from a large segregating population that provided good genome coverage.

So far only backcross, F_2 or F_1 populations have been utilized as a basis for genetic maps in *L. perenne* (Table 4). The most used populations are the p150/112 BC_1 (ILGI) and the WSC F_2 (RASP) populations (Table 4). All together, seven mapping populations were used in the 18 published genetic maps in pure *L. perenne* breeding programmes. Two of these mapping populations are included the Oak Park breeding programme (Anhalt *et al.* 2008). One is described in the present study.

Table 4: Summary of maps published for pure *L. perenne* populations. Table displays lengths and marker numbers of LGs.

Publication	Name of the population	Total map length	LG 1		LG 2		LG 3		LG 4		LG 5		LG 6		LG 7		LG 8		Ø marker
		cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM
Anhalt *et al.* 2008	**'F_2 biomass population' *1**	**592.3**	**75**	**95.8**	**8**	**103.9**	**11**	**96.3**	**17**	**71.1**	**11**	**87**	**10**	**49.5**	**9**	**88.7**	**9**			**8**
Anhalt *et al.* 2008	'F_1 late flowering population'	225.7	60	22.7	8	68.7	8	39.9	12	39.2	8	11.5	4	21.4	10	22.3	10			3.8
Muylle *et al.* 2001	*Lp* mapping population (F_1)	550	190	65	34	104	23	111	29	100	48	58	16	69	36	43	4			3
Muylle *et al.* 2005	*Lp* mapping population (F_1)	745	164	102	48	133	42	105	30	94	32	83	22	118	17	110	33			3
Jensen *et al.* 2005 b	VrnA (F_2)	490	93	31.7	12	86.9	8	87.6	20	112.5	21	54.8	8	64.7	9	52.2	15			5
Bert *et al.* 1999	p150/112 BC_1 (ILGI)	930	471	180.1	80	177	74	156.4	95	117.6	64	106.3	33	103.2	63	89.4	62			2
Jones *et al.* 2002 b	p150/112 BC_1 (ILGI)	811	240	90	25	129	43	114	31	136	48	96	28	124	30	122	35			2
Jones *et al.* 2002 a / Cogan *et al.* 2005	p150/112 BC_1 (ILGI)	814	258	92	30	148	44	117	33	133	56	90	30	113	28	121	37			2.6
Armstead *et al.* 2002	p150/112 BC_1 (ILGI)	565	134	71	12	113	26	94	20	82	27	59	12	78	21	68	16			4
Yamada *et al.* 2004	p150/112 BC_1 (ILGI)	814	186	92	19	148	34	117	24	133	43	90	21	113	18	121	27			4.4

Publication	Name of the population	Total map length	LG 1		LG 2		LG 3		LG 4		LG 5		LG 6		LG 7		LG 8		Ø marker
		cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM	# marker	cM
Jensen *et al*. 2005 a	Consensus map*[2]	772	66	84	14	115	7	132	9	121	16	102	2	84	1	134	17			11.7
Armstead *et al*. 2002	WSC F_2 (RASP)	515	74	72	6	72	16	73	10	82	14	63	10	75	8	78	10			7
Armstead *et al*. 2004	WSC F_2 (RASP)	628	157	90	17	79	30	99	26	95	29	78	15	85	17	102	23			4
Skøt *et al*. 2005	WSC F_2 (RASP)	631	118	90	17	105	22	99	18	92	17	68	12	74	14	103	18			5
Turner *et al*. 2006	WSC F_2 (RASP)	626.9	227	86.7	27	95	40	90.5	36	111.8	37	65.2	22	83.1	26	94.6	39			3
Gill *et al*. 2006	WSC F_2 (RASP)	684.4	330	82.4	43	110.5	53	91.1	49	118	53	77.7	32	107.1	44	97.6	56			2
Faville *et al*. 2004	NA_6 (F_1)	991	175	97	28	139	19	129	30	96	24	113	14	181	29	193	25	16	6	5.7
	AU_6 (F_1)	791	140	119	22	98	20	112	18	150	15	85	17	103	24	89	21	1	3	5.7
Cogan *et al*. 2006	NA_6 (F_1)	1070	234	147.1	40	144.9	28	138.7	38	148.8	38	113.6	19	175.5	36	201.5	35			4.6
	AU_6 (F_1)	947.6	177	143.6	31	-	-	131.8	34	157.4	21	132.8	28	106.5	33	98.5	30			5.4

*[1]Population of the presented study
*[2]p150/112 BC_1 (ILGI), VrnA, SB2TC1 (Mylle *et al*. 2003), Pop8490 (Barre *et al*. 2000)

The F_2 population map had a map length of 592.3 cM. The longest reported map was based on 234 markers and had a map length of 1070 cM by Cogan *et al*. (2006). The shortest reported so far is the 'F_1 late flowering' population (Anhalt *et al.* 2008) and had 60 markers with a map length of 225.7 cM. However, the map of Bert *et al*. (1999) incorporated 471 marker loci and had a map distance of 930 cM. This would appear to indicate that good genome coverage has been achieved in the study of Cogan *et al*. (2006) and that the addition of more molecular markers would not serve to increase the map length. The published larger genetic maps have in general been created with a high amount of RFLP and AFLP markers. The advantage of RFLP and AFLP markers having multiple loci which can be detected with one single PCR, enable fast enhance genome coverage compared to maps constructed only with SSRs. Faville *et al*. (2004) and Cogan *et al*. (2006) shared the same population and map. The LGs 4, 5, 6 and 7 were the longest in the genetic map of Cogan *et al*. (2006), LGs 1, 2 and 3 were the longest in the genetic map of Bert *et al*. (1999). The LGs in the present study had an intermediate length compared to other published maps. Only LG 4 and LG 6 were much shorter compared than other maps explained at least in part by the limited number of markers used for mapping and by SD (Table 3). In general, different map length distances may occur as a result of the variation in recombination frequencies of different population structures and might also be a result of the mapping function used to construct the linkage maps (Gill *et al*. 2006). Using Haldane's mapping function instead of Kosambi's mapping function results in shorter map lengths (Jensen *et al*. 2005 b). A further explanation for map contraction of relevance here is inbreeding depression leading to an under-representation with markers in some areas of the genome (Gill *et al*. 2006) and can also be explained by the high occurrence of SD in some regions of the map in this study (Figure 4 a).

Already published SSR markers were chosen for this study because of the ease in use and their co-dominant character that allows multiallelic analysis, and because of their reproducibility. Another advantage is that they can be used in comparative mapping to compare the results to those of other *Lolium* studies. SSRs are a useful tool for alignment of trait-specific maps with reference maps (Jones *et al*. 2002 b) because of their transferability between populations and their high polymorphism rates (Warnke *et al*. 2004). AFLPs have several disadvantages because they show reduced power in genetic mapping and QTL location compared to co-dominant markers (Jones *et al*. 2002 b). The tendency for small groups of AFLP markers to cluster together has been assessed by Bert *et al*. (1999). An advantage of AFLP markers is that they tend to cluster preferably on the distal end of the chromosomes, which expanded as well the genetic map in the present study (Figure 4a). Other publications reported that AFLP markers cluster in the centromeric regions (Bert *et al*. 1999, King *et al*. 2002), which is in disagreement with the present study. SSR markers tend to map near the centromeres (Gill *et al*. 2006) and show centromeric clustering (Jones *et al*. 2002 a). Therefore it seems to be appropriate to construct a genetic map with different marker types to achieve the best genome coverage. Nevertheless, markers were clustered in some regions and gaps of 10 to 51 cM on some LGs indicating that either recombination events or mapped loci were not evenly distributed throughout the genome (Saha *et al*. 2005). In previous studies the marker density ranged between 2 cM (Bert *et al*. 1999, Jones *et al*. 2002 b) and 12 cM (Jensen *et al*. 2005 b). The marker density in the present study was 8 cM and in the range of marker density of other studies. To be sufficient for purposes of MAS for molecular breeding the genetic map distance should be 1 cM or less to the marker linked to the trait of interest (Mohan *et al*. 1997).

The maps of Gill *et al*. (2006) and Turner *et al*. (2006) shared a considerable number of common markers with the map presented here. Generally, there was good conservation of marker order between the maps. Exceptions were found for example in the marker rv0863 on LG 3, which was in the map of Gill *et al*. (2006) on the upper end of the LG and in the present study located in the centromeric region of the LG. The next marker with altered position was rv0739, which was located on the other end of LG 6 compared to the map of Gill *et al*. (2006). This could be explained for the present study by the high extent of SD in the regions of the markers on LG 3 and 6. Another explanation could be "ghost markers", which are markers with two duplicate marker loci. Equal fragment length results in an equal banding pattern, and therefore, the alleles of duplicate markers are scored in the mapping population as the alleles of one single marker. Since the locus order of LGs is determined on the basis of recombination frequencies between loci incorrect recombination values for a LG can result in an incorrect locus order for the LG (Frisch *et al*. 2004).

A large number of loci showed SD with the largest cluster occurring on LG 6 where all loci have a high level of SD ($P<0.0001$). Additionally, SD clusters were seen on LG 1 and 7 (Figure 5) and LG 5 and 3 showed regions with SD especially on the distal ends (Figure 4 a). The majority of loci on LG 2 and 4 displayed no SD (Figure 5). SD was observed in 63% of all mapped loci, which is similar to the findings of Jensen *et al*. (2005 b). The phenomenon of SD and its causes are poorly understood (Jenczewski *et al*. 1997). A single or a combination of different mechanisms may be responsible for SD in any particular case. Events leading to SD can be initiated in different developmental stages including sporogenesis, spore function, seed development and seed germination (Zamir and Tadmor 1986). SD can arise from a dysfunction of the gametes in pollen, megaspores or both (Lyttle 1991). The SD loci in

the present study were often skewed to one allele of PM, PFa or F_1 allele but as well to two alleles at the same time (Table 3). Xu *et al*. (1997), working with rice, found two directions of skewness favouring one parent or the other in all population examined. However, favouring of marker alleles towards one parent seems to be common and not unexpected in nature. In male gametes, pollen killers or pollen abortion, more frequently results in SD as compared to disturbances in female gametes (Taylor and Ingvarsson 2003). Gamete selection eliminating gametes of either sex has been previously reported (Sano 1990). SD during female meiosis can lead to genomic disorders (Jenczewski *et al*. 1997). It is reported that self-incompatibility and self-compatible loci could cause SD (Thorogood *et al*. 2002 and 2005). Regions on the genome with consistent distorted marker ratios in the homozygous genotypes can be associated with loci segregating to self-compatibility (Thorogood *et al*. 2005). There are also technical reasons for SD caused by errors during genotypic analysis or mutations within the binding site of a DNA marker. These mutations would affect only certain marker loci (Sibov *et al*. 2003) and are independent to population structure and species. Missing data and genotypic errors might occur and can lead to SD. However, a better knowledge of the influences of SD in mapping populations and breeding programmes is important. To identify and describe specific SD regions in *L. perenne*, this study reports findings of already existing genetic maps. Faville *et al*. (2004) found in their study distorted loci on the 'NA_6' map on LGs 2, 3, 4. On the 'AU_6' map distorted loci were located on LG 5, which was comparable to the findings of Bert *et al*. (1999) who reported that highly skewed markers were found on LG 5 in the ILGI population. In three studies which used a common F_2 population developed from an F_1 hybrid obtained by crossing individuals from partial inbred lines LG 5 and LG 7 were reported with the highest amount of distorted marker loci using sets of different markers (Armstead *et al*.

2004, Gill *et al*. 2006, and Turner *et al*. 2006). Armstead *et al*. (2002) found other SD affected regions in a BC_1 population using the same set of markers compared to the studies with the F_2 population (Armstead *et al*. 2004; Turner *et al*. 2006). In the BC_1 population LG 3 and LG 4 had severe SD. Markers distorted in the BC_1 population were not distorted in the F_2 population. Jones *et al*. (2002 a and b) used in two studies the ILGI F_1 population with a multiple heterozygous parent and a double haploid parent in the pedigree. One genetic map of Jones *et al*. (2002 b) was constructed with RFLPs, AFLPs, ESTs and isoenzyme markers; their second map (Jones *et al.* 2002 a) was extended with additional SSR markers but had the previously used markers in common. Both maps showed severe SD on LG 3 but the first map of Jones *et al*. (2002 b) featured an additional segregation distorted region on LG 4. All these previously reported findings share little similarities regards the extent of SD with both parental maps of the F_1 late flowering population (Anhalt *et al.* 2008) except for LG 2, which had a larger amount of SD on the 'NA_6' and 'J43' maps. On the 'J43' parental map 88% of the distorted loci were located on LGs 2 and 7 and on the 'J51' parental map 65% of the distorted loci were located on LG 1. Jensen *et al*. (2005 b) described in their publication that all their LGs showed SD but the highest amount of distorted loci were on LG 1 and 3. As well the F_2 population in the present study showed on all LGs SD with the largest number of distorted loci on LG 1, 3, 6 and 7. Concluding, all studies showed similar distorted regions in *L. perenne* genoms. No specific common hotspot regions for SD were found; although, some regions on LGs with a higher frequency of SD could be identified (LG 3, 4, 5, 7).

A breeding programme, which often takes more than ten generations to develop could benefit by the use of molecular breeding techniques but SD can influence the selection process and progress when not being recognised as a factor in the population

structure. Therefore, a deeper understanding of the causes of SD for the breeding context is required.

For a good mapping study a well studied population with a reasonable amount of genotypes is required. Among mapping populations the population size generally range from 50 to 250 genotypes. Small population sizes can lead to false negative events in screening for polymorphic markers. A false negative event is if a marker is wrongly determined to be monomorphic marker during marker testing. False negative events can happen with a high frequency if the size of the population used for screening is small (Liu 1998). Therefore, a high number of genotypes is required for a powerful mapping study. More extended populations are required for fine mapping to obtain a high resolution between marker and trait especially for recombination-poor regions such as centromeric regions. Population sizes in other *Lolium* studies varied between 95 (Bert *et al*. 1999) and 280 (Muyelle *et al*. 2001) individuals. The F_2 biomass mapping population consists of 360 individuals. This is the largest mapping population reported for *L. pere*nne so far. It is especially important to aim for a large population if the population is designed for further fine mapping and cloning of genes linked to the trait of interest. A large population study keeps the error small.

Genetic variation in a breeding programme can be greatly improved with knowledge of the genome organisation of the species concerned in the number and distribution of the genetic loci controlling traits of agronomic importance (Hayward *et al*. 1998). Genetic improvement of forage grasses by conventional breeding is hampered because of the obligate outbreeding nature of temperate grasses. MAS technology can add more exactness to plant breeding programmes. Gene combinations can be targeted with more precision, and can lead to a better management of specific agronomic traits (Yamada *et al*. 2005). Thus MAS could potentially speed up conventional breeding.

Therefore, it is necessary to identify tightly linked markers to the trait of interest. These markers should be neutral molecular markers and not environmentally regulated, that is, unaffected by the condition in which the plants are grown and useful in all stages of plant growth (Mohan *et al*. 1997). Furthermore, MAS can be used to eliminate unwanted linkage drag. Linkage drag refers to the reduction in fitness in a cultivar due to deleterious genes introduced along with the beneficial gene. When noval variation is sought from wild relatives, molecular makers offer a tool in which the amount of wild or alien DNA can be monitored. To make MAS efficient for breeding programmes the following prerequisites should receive attention: (1) the marker should be closely linked to the trait of interest, (2) the analysis of screening of a large population should be easy, cheap, and fast, and (3) the marker technique should be highly reproducible (Mohan *et al*. 1997). If all this is present, MAS will be an important and successful tool for future breeding programmes.

In this chapter the construction of a SSR and AFLP based genetic linkage map of *L. perenne* was described. The map provides good genome coverage and can be aligned to other *Lolium* maps with the selected SSR markers from different primer sources. Tall fescue cross-amplified SSR markers give the opportunity to link maps among species for comparative studies and for association mapping in ecotypes. The *Lolium* genetic map created in this work represents a good balance between marker density and population size and will serve as an important tool in the molecular breeding programme at the Oak Park Research Centre to unravel the genetic components for traits of agriculture importance.

Chapter 5: QTL mapping for biomass

5.1 Introduction

One approach to dissect the genetic basis of complex traits is QTL mapping. QTL mapping involves the creation of a genetic linkage map in a segregating population and the statistical association between molecular markers segregating in the population and segregation for the trait under investigation. The linkage between a genetic marker and a quantitative trait was already recognised by Sax (1923). With the rapid development in molecular marker technology since the 1980s, it is now possible to use molecular marker information to map major QTL (Stuber *et al.* 1992; Zeng 1994).

Traits in plants are described as being either discrete or quantitative in nature. Discrete traits are controlled by a single gene and therefore are generally more straightforward to select in a plant breeding programme. However, many traits of agricultural importance are controlled by multiple genes and behave in a quantitative manner (Snape *et al.* 2007). These traits provide challenges to plant scientists trying to improve the agricultural fitness of crops.

The improvement of biomass yield is one of the major breeding aims in forage grasses. Much progress has been made in the past in breeding for improved yield in *L. perenne* through classical breeding. Nowadays additionally molecular markers are used to follow the introgression of genes (Humphreys *et al.* 2005). It is envisaged that the use of molecular markers linked tightly to QTL associated with forage yield can be used in molecular breeding programmes to assist traditional breeding. Biomass yield-related QTL have been reported in *L. perenne* previously (Yamada *et al.* 2004; Armstead *et al.* 2008, Turner *et al.* 2008). QTL studies on biomass yield have been carried out in different species, e.g. *Arabidopsis* (Lisec *et al.* 2008) and sorghum (Lin *et al.* 1995).

Heterosis was detected in studies for different species: for example in a hybrid of two elite maize inbred lines by Stuber *et al.* (1992) or in a cross of *indica* and *japonica* rice lines by Xiao *et al.* (1995).

Several QTL studies on a range of traits have been reported for *L. perenne* (Table 1).

Table 1: Summary of quantitative trait loci (QTL) studies for different traits in *Lolium perenne* displayed by mapping populations used, linkage groups (LG) were QTL where detected and different mapping methods used.

Publication	Name of population	QTL traits	LGs	Mapping methods
Armstead *et al.* 2004	WSC F_2	Heading date	2, 4, 7	IM*[1], MQM*[1]
Yamada *et al.* 2004	p150/112	Plant height	1, 3	SIM*[1], CIM*[1]
		Tiller size	1, 3	
		Leaf length	5	
		Leaf width	3	
		Fresh weight	5	
		Plant type	4, 7	
		Spikelet per spike	1	
		Spike length	1, 3, 5	
		Heading date	4	
Cogan *et al.* 2005	p150/112	Crude protein	1, 2, 3, 4, 5	IM*[1], CIM*[1]
		In vivo dry matter	1, 3, 4, 7	
		Digestibility	1, 2, 3, 4, 5, 7	
		Neutral fibre content	1, 3, 4, 7	
		Estimated metabolism energy WSC*[2]	1, 2, 3, 5, 7	
Jensen *et al.* 2005	F_2 VrnA	Vernalization response to growth rates by days to heading	2, 4, 6, 7	IM*[1], MQM*[1]
Muylle *et al.* 2005	F_1 population for crown rust resistance	Crown rust resistance	1, 2	MQM*[1]
Turner *et al.* 2006	WSC F_2	*WSC*[*2] fractions for:*		MQM*[1]
		-Tiller base spring	1, 5	
		-Tiller base autumn	1	
		-Leaf autumn	2	
		-Leaf spring autumn	6	
Armstead *et al.* 2008	p150/112	Seed set	4[2)], 7	IM*[1],
	WSC F_2	Heading date	4, 7	MQM*[1],
		Leaf length	7	CIM*[1]
		Leaf width	7[2)]	
		Pollen viability	4[2)]	
		Anther dehiscence	4[2)]	
Byrne *et al.* 2008	Late flowering population	Heading date	2, 3, 4, 5, 7	IM*[1], MQM*[1]
		Spike length	4	
		Spikelet per spike	2, 3, 4	
King *et al.* 2008	BC_2 introgression mapping population	Heading date	4, 7	IM*[1], MQM*[1]

Publication	Name of population	QTL traits	LGs	Mapping methods
Turner *et al.* 2008	WSC F_2	Plant size (g fresh weight)	1, 4, 5	MQM*[1]
		Root amount	1, 2, 4	
		Leaf extension rate	2, 3	
		Autumn dry matter	3	
		Wiltiness	1, 2	
		Hydration (well watered)	5	
		Hydration (droughted)	6	
		Herbage survival	1, 5	
		Tiller survival	1	
		Regrowth after rewatering	1, 5	

[1)] only p150/112, [2)] only WSC F_2
*[1] IM = Interval mapping, MQM = multiple QTL models, CIM = composite interval mapping
*[2] WSC = water soluble carbohydrates

Three studies reported on biomass yield related traits in *L. perenne* so far. Yamada *et al.* (2004) reported on fresh weight and leaf width QTL, Armstead *et al.* (2008) described flag leaf width QTL, and Turner *et al.* (2008) reported on fresh weight and dry matter QTL.

5.1.1 Yield

One important quantitative trait in *L. perenne* is yield, which is together with forage quality the key objective of forage breeding programmes (Humphreys *et al.* 2005). Several factors are influencing yield. An important factor for biomass production is temperature. High temperatures generate rapid expansion and thin leaves at the expense of roots and enhance the photosynthetic fixation of carbon in leaves (Jones and Lanzenby 1988). Light also affects form and function of grass leaves and supports the plant growth, which influences leaf length and leaf width (Jones and Lanzenby 1988).

A strong association between biomass and metabolic composition was found in several studies (Meyer *et al.* 2007; Lisec *et al.* 2008). This interaction between metabolism and plant growth regulatory mechanisms may work in two ways: either a high supply of metabolites causes growth, or growth drains metabolites to a minimum tolerable level upon which growth is limited (Meyer *et al.* 2007). But many metabolites

are likely responsible for growth. Sugars like glucose and sucrose have been shown to act as metabolic signals to be involved in growth (Meyer *et al.* 2007). A very important factor in plant growth and therefore as well for biomass is nitrogen. It is a limiting nutrient in natural environments and in agriculture (Maloof 2003) and can influence biomass production. Several biotic or abiotic factors, e.g. diseases or drought stress, can limit biomass yield (Jones and Lanzenby 1988).

5.1.2 Heterosis

Heterosis was first described in plants (Darwin 1876). It is the superior performance of heterozygous F_1-hybrid plants compared to the average of their homozygous parental inbred lines (Shull 1948). Heterosis may result from either (1) true overdominance, which can be described as favourable allelic interactions at heterozygous loci that outperform either homozygous state or (2) pseudo-overdominance (Epistasis), which is a superior phenotypic expression of a trait in hybrids by interactions between non-allelic genes at two or more loci (Stuber *et al*. 1992) or (3) dominance, which is described as alleles at different loci in the two homozygous parental genomes, which are complemented in the heterozygous F_1 hybrid but only one is completely dominant (Becker 1993). Heterosis is higher in cross-pollinating crops than in self-pollinating crops and is higher when the parental plants have a highly different genetic background. Therefore, grass breeders aim for a high degree of heterozygosity in their hybrid cultivars and choose parental lines that are genetically diverse to encourage heterosis (Becker1993). There are several biomass heterosis studies in different plant species existing, e.g. in smooth bromegrass (Casler *et al.* 2005), in chickpea (Hegde *et al.* 2007), or in *Arabidopsis* (Barth *et al.* 2003).

5.1.3 QTL

QTL are locatable genetic markers that are closely linked to candidate genes affecting biological or agronomic traits. The association between the markers and the trait is used to find genetic locations of the genes controlling the trait (Winter *et al.* 2002). Accurate phenotypic and genotypic data and statistical software are required to determine QTL. Generally three steps are involved in QTL study approaches: (1) QTL detection: positioning of the QTL interval in the range of 10-30 cM on the genetic map, (2) QTL fine mapping: narrowing the QTL interval to the range of 1-2 cM on the genetic map in order to search for candidate genes, and (3) Identification of a gene or genes underlying the QTL containing the trait of interest. Marker development could be done from open reading frames (sequence of bases that could potentially encode a protein), e.g. in rice where a region on chromosome 12 contains a QTL conferring tolerance to cold-induced wilting and necrosis, which could be delimited to a contig of 55 kb (Andaya and Tai 2006).

Important is the choice of the experimental population for QTL mapping. Often backcross or F_2 population have been chosen for QTL studies (Carbonell *et al.* 1993). F_2 populations appear to be superior over backcross populations because recessive alleles in a recurrent parent cannot be detected and if dominance is present backcrosses result in biased estimates because additive and dominant effects cannot be differentiated (Carbonell *et al.* 1993). Instead of experimental populations natural populations, engaged in association mapping, can be used. The difference is in the co-segregating patterns for natural populations, which are more complex than those for experimental populations (Liu 1998). Using F_2 populations generated out of an inbred line cross for QTL studies seems to be the most powerful method because this breeding design increases the linkage between the marker and the QTL (Doerge *et al.* 1997, Liu 1998).

The QTL detection depends on the type of marker employed, the marker distribution in the genome, and cross design and magnitude of the QTL (Jansen 1993). It is important to choose the best fitting model for QTL analysis in order to avoid "ghost" QTL and to improve the power of QTL detection. Multiple QTL can be mapped more accurately by using multiple QTL models (Jansen 1993).

The provision of readily available DNA markers and the development of powerful biometrics methods have led to significant progress in QTL mapping in plants. The combined examination of genotype marker segregation and phenotypic values of individuals or lines enables the detection and location of loci affecting quantitative traits. The success of a QTL study relies on the segregating genotypes in the study. A difficulty is that phenotypic effects of individual genes associated with complex traits are often relatively small (Asins 2002). Relevant genes that determine a single trait are often located in multiple regions of the genome. Co-dominant markers have a higher power for QTL mapping than dominant markers (Liu 1998). One advantage of dominant markers, for example of AFLP markers, is the faster construction of a high density map but this provides less genetic information. To conduct a robust QTL analysis it is important to create at least a medium density map with an appropriate marker interval of 10 – 15 cM based on a population of circa 300 genotypes in an experimental design (Erickson *et al.* 2004). A large sample size besides the experimental design is necessary to identify QTL with large effects (Erickson *et al.* 2004). To improve the detection of QTL it is necessary to test in multiple environments, to decrease the genotype × environmental interactions and to detect precisely the location and effect of QTL (Deorge *et al.* 1997).

5.1.4 Genotype x environment interaction

Forage yield is a trait with a strong genotype × environment interaction and is influenced by abiotic and biotic factors (Snape *et al.* 2007). The morphogenesis of grass plants within a grazed sward also plays a key role in determining herbage yield, persistence, and recovery from grazing (Yamada *et al.* 2004). In vegetative plant growth, plant morphogenesis is described by three key variables: leaf appearance rate, leaf elongation rate, and leaf lifespan (Yamada *et al.* 2004). Structural characteristics of plants such as tiller number, leaf number, and leaf size are the result of these morphological traits, and their measurement in breeding programmes allows a dissection of the complex herbage yield trait as well as predictions of response to grazing. Different ecoclimatic regions, or pastures under different grazing regimes, may provide different selection conditions for these traits (Yamada *et al.* 2004). Environmental interactions also have a high influence on a QTL study (Asins 2002). Diseases, pests, water availability, nutrient supply, flowering time, temperature, and interaction between abiotic and biotic factors can influence the unequal phenotypic development of genotypes. Therefore, it is important to choose the most accurate experimental design for the segregating population. A sufficiently large number of genotypes and replications ensure a maximum heritability of the trait for its phenotypic component. The robustness of the genetic map and the DNA markers, and an accurate statistical approach with the fitting software are necessary to detect consistent QTL. With replicated progenies a major reduction in the number of genotypes that need to be scored can be achieved, environmental variation can be minimized, and residual variation caused by other QTL can be identified and removed from the experimental error (Asins 2002).

5.1.5 Experimental design - Alpha lattice design

The lattice design was introduced by Yates (1936). It was limited in variety number and block size. The invention of the alpha design (Patterson and Williams 1976) overcame the problem of the limited block size and created a method for constructing a class of resolvable equiblocksized designs. The alpha lattice design is an incomplete block design. An advantage of this experimental design is the presence of checks in each incomplete block and the ease in identification of block effects compared to other experimental designs.

5.1.6 Prospects of QTL cloning

Plant breeding is a dynamic area of applied sciences. Plant breeding relies on genetic variation and the selection methodology to improve plant traits and characteristics that are of interest for the breeder and the farmer. The most widely used applications of QTL analysis are MAS in breeding programmes and the location of candidate genes for a trait of interest, e.g. crown rust resistance regions in *Lolium* (Dumsday *et al.* 2003).

Map based cloning has been successful in animals and plants (Tanksley *et al.* 1995, Fridman *et al.* 2000; Johanson *et al.* 2000; Yano *et al.* 2000). A general strategy is to isolate candidate genes involved in metabolic pathways. These can be either structural or regulatory genes. This approach could connect pathways by the identification of QTL with pleiotropic gene effects. Success depends on the reliability and accuracy of the analysis itself (Asins 2002). Lastly, the direction of selection on an individual QTL may be dependent on the environment, leading to heterogeneous effects on individual QTL throughout the life span of the individual (Weinig *et al.* 2003).

The QTL cloned to date are likely to represent a biased sample of those governing the variability of target traits; only major QTL, mostly identified in wide crosses have been successfully targeted (Salvi and Tuberosa 2005). Most minor QTL will not be traceable using fine mapping (Price 2006). Lander and Botstein (1989) suggested selective genotyping as another way to fine map QTL. Some progeny contribute more linkage information than others. Only these genotypes are used for the selective genotyping analysis.

The objective of the present study was to identify QTL for biomass yield components in a segregating F_2 population of *L. perenne*. The creation of the necessary genetic linkage map for this population has been described in chapter four in this thesis. The specific traits to be mapped were (1) fresh weight, (2) dry weight (3) dry matter, and (4) leaf width.

5.2 Material and Methods

5.2.1 Plant material

The parental lines originated from two different genetic backgrounds and were chosen because of their high degree of homozygosity (see detailed description in chapter two).

5.2.2 Experimental design - Alpha lattice design

To create the layouts of the alpha lattice design Agrobase Generation 2™ (Agronomix Software, Inc., Winnipeg, Manitoba, R3N 0S4, Canada) was used. A

replicated greenhouse and a field experiment were designed. Each replication comprised 45 incomplete blocks. One block consisted of nine genotypes. In each block eight F_2 genotypes and either one of the maternal, paternal or the F_1 genotypes were included as checks. In autumn 2005 to spring 2006 a greenhouse trial in three replicates was set up in two greenhouse environments. Two out of the three replicates were planted in a greenhouse environment (Figure 1 a) with the following climatic conditions: average nightly and daily temperatures of 11ºC and 19ºC, respectively. A third replicate was planted in a second greenhouse environment (Figure 1 b) with average nightly temperatures of 6ºC and daily temperature of 19ºC.

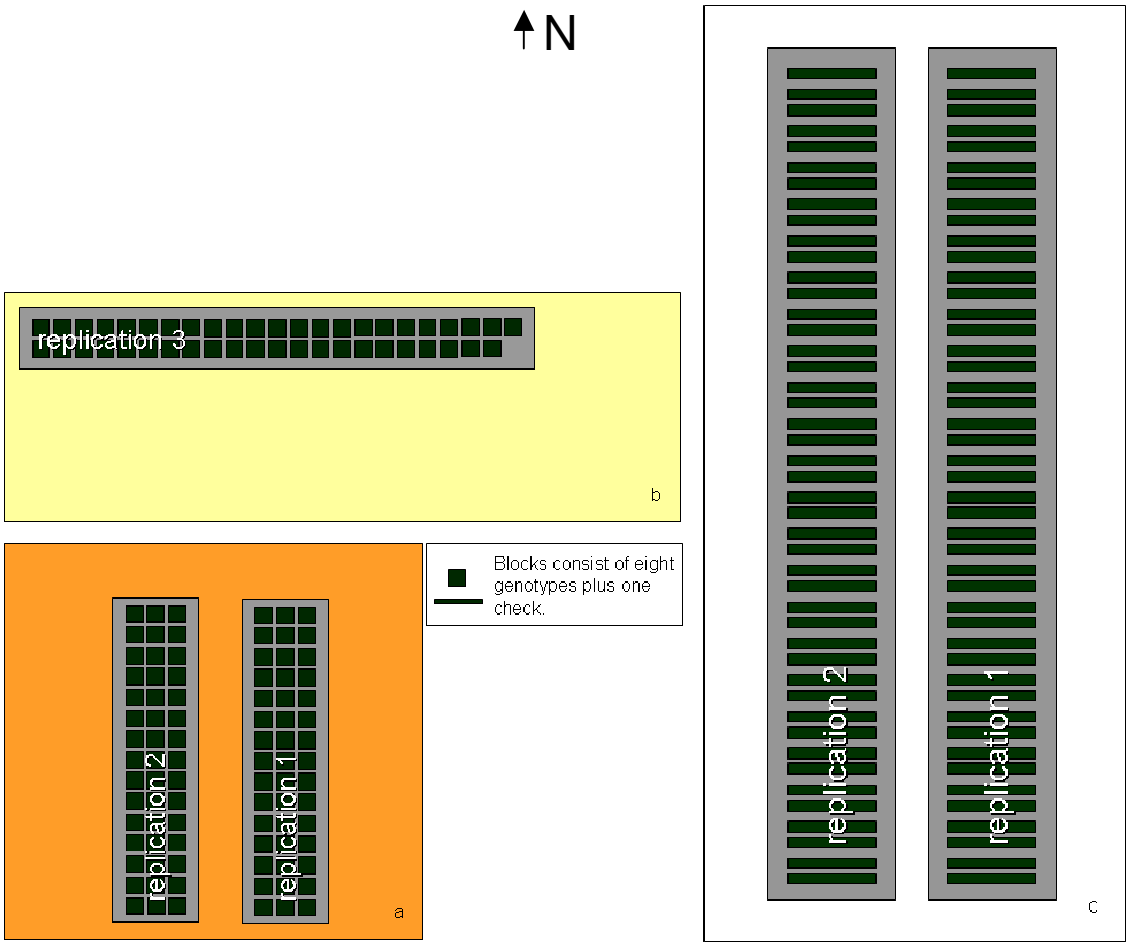


Figure 1: Layout of the experiments. a, b: Greenhouse a: first and second replication (environment 1, red); b: third replication (environment 2, yellow). c: Field: first and second replication (environment 3, white).

For the field experiment, the same experimental design as in the greenhouse was used (Figure 1 c). The field experiment consisted of two replicates. A modification was the organization of plants in mini-swards consisting of six clonal replicates of each genotype.

5.2.3 Collection of phenotypic data

Altogether, three harvests from the greenhouse experiment were carried out in December 2005, February 2006 and April 2006. For the field experiment, four harvests were taken in August 2006, October 2006, May 2007, and August 2007. Weather data were recorded with temperature (ºC) and rainfall (mm) (Appendix Figure 1). To reduce experimental errors during harvesting a measurement guide was used for the cuts. The cutting height was 8 cm for the greenhouse experiment and 5 cm for the field experiment. Since the plants in the field were smaller when the first harvest was done the cutting height in the field was lower than in the greenhouse. Plant samples were dried to constant weight in an oven at 60ºC. After each cut the plants were supplied with 80 kg N/ha. Fresh weight (g) and dry weight (g) were measured.

Dry matter (%) was calculated as follows:

dry matter = (dry weight / fresh weight)*100

Leaf width was only measured in the greenhouse (environment 1) in two replications with four random measurements on each genotype. The measurements were taken in the middle of the fully expanded leaf blades using calipers.

In addition to the measurements described above, heading date was recorded in the field experiment. However, the dataset for heading date taken in the field was incomplete because the first harvest for biomass production was taken when heading for the majority of the F_2 genotypes started. The necessity of the harvest prevented the

completion of the data set for heading date. Heading date was also recorded from potted plants under outdoor conditions every day from the 1st of April until the end of heading in 2006. Heading date was recorded because the timing of maturity has an influence on biomass yield.

5.2.4 Data analysis

The distribution of data for each trait and each harvest was calculated using SAS V9.1 using the ProcMixed model (SAS Institute Inc. 2004. SAS OnlineDoc® 9.1.3**.** Cary, NC: SAS Institute Inc.). For non-normal distributed data a log transformation was carried out and the estimates of transformed data for fresh weight, dry weight, dry matter and leaf width were used for the QTL analysis. Plants, which were existing but with a too low weight for the fine balance or under the cutting height were analysed in SAS as an absolute term value.

The terms relative mid parent heterosis (MPH) and relative best parent heterosis (BPH) describe the degree of phenotypic difference of a trait in a hybrid (F_1) compared to its parental inbred lines (P_1, P_2). MPH signifies that a trait displays a hybrid performance that is better than the average value of the two parental inbred lines. BPH explains the degree of phenotypic difference of a trait in a hybrid (F_1) compared to the homozygous parental inbred lines, which performs better (P_b) than the other parental line.

$MPH = ((F_1 - [(P_1 + P_2)/2])/[(P_1 + P_2)/2])*100$

$BPH = ((F_1 - P_b)/P_b)*100$

F_1 = phenotypic performance of a trait in a hybrid line

P_1/P_2 = phenotypic performance of a trait in the parental inbred lines

P_b = the better performer of the two parental inbred lines

An analysis of covariance was performed for the greenhouse experimental data to account for environmental influences on plants within the blocks on the three tables in the greenhouse experiment. SAS V9.1 PROC General linear models (PROC GLM) was used to calculate the influence of blocks and environment interactions.

Analysis of variance was carried out using PROC MIXED in SAS V9.1. The MIXED procedure is based on the general linear mixed models:

$$Y = X\beta + ZU + e$$

where Y is a n X 1 vector of observations, β is a p X 1 vector of fixed, unknown parameters, X is a n X p design matrix for the fixed effects, U is a q X 1 vector of unobservable random effects, Z is a n X q design matrix for random effects, and e is a n X 1 vector of residual random errors (Littell *et al.* 1998). This procedure implements random effects and permits modelling the covariance structure of data. PROC MIXED can compute efficient estimates of fixed effects and valid standard errors of the estimates. Usually there are two statements used in PROC MIXED, one is the RANDOM statement, which shows the variation between individuals and the other is the REPEAT statement, which explains the covariance between individuals (Littell *et al.* 1998). The RANDOM statement was used because of the non-covariance between individuals. Unlike GLM, where standard errors for adjusted means (Least Squares Means =LSMEANS; Means, which have been corrected for imbalances in other variables) and ESTIMATE (ESTIMATE statement computes a point estimate and the standard error for a linear combination of the parameters in the model) statements are not computed correctly and most tests of hypothesis in an ANOVA required optional specification such as TEST statements. MIXED procedure almost always makes valid

computations for tests of hypothesis and standard errors of estimates (Littell *et al.* 1998).

5.2.5 Heritability of traits

Heritability (broad sense heritability) for each trait was calculated for the greenhouse and field environment using the formula:

$$H^2 = \delta g / \delta p$$

$$= \delta g / \delta g + \delta_e$$

δg represented the genetic variance component, δp the phenotypic variance, δ_e the standard error.

5.2.6 QTL analysis

Phenotypic data of all the replicates for each trait of each experiment were calculated together for the QTL experiment to increase the power of QTL detection.

MapQTL® Version 4.0 (Van Ooijen *et al.* 2002) was used for QTL detection. MapQTL has the option to compute interval mapping (IM) and multiple QTL models (MQM) mapping.

Interval mapping (Lander and Botstein 1989):

IM is a single QTL model, which calculates at each increment (for example each cM) across the ordered markers in the genome. The result of the tests are expressed as LOD scores, which compare the evaluation of the likelihood function under the null hypothesis (no QTL) with the alternative hypothesis (QTL at the testing position) for the purpose of locating probable QTL (Doerge 2002). The ELOD value is the expected LOD score for a marker located exactly at the QTL (Lander and Botstein 1989):

$ELOD_{\text{single marker}} \approx (1 - 2\theta)^2 \, ELOD_0$

$ELOD_{\text{interval mapping}} \approx (1 - 2\theta)^2 \, ELOD_0/(1 - \psi)$,

where ψ is the recombination fraction corresponding to variable d cM, θ is the recombination fraction corresponding to ½ d cM , and $ELOD_0$ is the expected LOD score for a marker located exactly at the QTL. Interval mapping decreases the required number of progeny by a factor of $(1 - \psi)$. Interval mapping is certainly a powerful mapping method nevertheless it is a single QTL method and the one-dimensional search that does not allow interactions between multiple QTL has to be considered. MQM mapping would be more powerful for multiple QTL to state the QTL more precisely.

In an F_2 population different genetic effects can be described:

Additive effect: This can be described as:

$a = 0.5 (\mu_1 - \mu_3)$ μ_i: the estimate mean of the distribution of a quantitative trait associated with the "μ_1" (PM)-, "μ_2" (F_1)- or "μ_3" (PFa)-genotype.

A dominance effect is present if a complementary action of superior dominant alleles from both parental inbred lines at multiple loci over the corresponding unfavourable alleles exists. This can be explained as:

$d = 0.5 (2\mu_2 - \mu_1 - \mu_3)$.

If no dominance was fitted for μ_2:

$d = 0.5 (\mu_1 + \mu_3)$

Epistatic genetic effects (is a mixture of additive and dominant effects), which exists in BC populations does not exist in an F_2 population because all F_1 individuals have the same genotype.

MQM mapping (Jansen 1993 and 1994, Jansen and Stam 1994):

MQM is a model for simultaneous mapping of multiple QTL. It is based on IM but in this model selected cofactors take over the role nearby QTL. A single QTL is fitted in a

background of cofactors. Therefore, genetic effects of the separate QTL, for example the single fitted QTL plus the others as represented by cofactors, are modelled as additive fixed effects. This model assumes there are no epistatic interactions (Van Ooijen *et al.* 2002):

$$Y = A + D + I + E,$$

with A as the additive genetic, D dominant genetic, I epistatic genetic effects and E as experimental or environmental errors.

The IM function in MapQTL® Version 4.0 (Van Ooijen *et al.* 2002) was used to estimate the map position, LOD score, and phenotypic effect of potential QTL, in terms of percentage of phenotypic variance explained. Genome wide LOD significance thresholds were determined using permutation tests with 1000 iterations. This is a resampling method to obtain empirical significance threshold values (Churchill and Doerge 1994). MQM mapping is appropriate when multiple QTL are present.

5.2.7 Correlation and regression analysis

Pearson correlation and a stepwise regression analyses were carried out for the measured traits fresh weight, dry weight, dry matter and leaf width using PROC CORR and PROC REG statement in SAS V9.1. Since heading date was not a complete data set no correlation was calculated.

Pearson correlation (r) is a measure of the correlation of two variables X and Y measured on the same object or organism. It is a measure of the tendency of the variables to increase or decrease together. It is defined as the sum of the products of the standard scores of the two measures divided by the degrees of freedom.

The model of the Pearson correlation was:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{(\sum X)^2}{N})(\sum Y^2 - \frac{(\sum Y)^2}{N})}}$$

With the variable X and Y and N as the number of cases.

The model for the stepwise regression analysis was:

$y_i = ß_0 + ß\chi_i + \varepsilon_i$

y_i: value of the dependent variable
χ_i: value of the independent variable
ß: regression coefficient
ε_i: error term (followed a normal distribution centred at 0)

A stepwise regression analysis was carried out excluding stepwise non-significant parameters.

5.3 Results

5.3.1 Plant architecture

The parental lines differed in plant architecture (Figure 2). The maternal line showed an erect growth habit, which was in contrast to the paternal line that displayed a prostrate growth habit. The F_1 plant resembled more the growth habit of the maternal plant. However a striking difference to both parental lines was the visible biomass heterosis (Figure 2). The F_2 generation was segregating phenotypically in genotypes with features of the parental lines and F_1 line.

Figure 2: Maternal, paternal and F_1 line (from left to right) displaying variation in plant architecture and biomass production.

5.3.2 Normal distribution of trait data

Fresh weight, dry weight, dry matter and leaf width data were normally distributed after a log transformation (Figure 3). No significant interactions of environmental effects within the blocks were found after log transformation.

Paternal and F_1 data were in the overall data range but showed more negative residual values. Paternal and F_1 values were located generally close together in the histogram an exception being here where the F_1 dry weight data showed in the greenhouse more positive values and in the field more negative values than the data of the parental lines. The field data showed in general more negative values of the F_1 data than the greenhouse data. This may be the effect of environmental influences in the field.

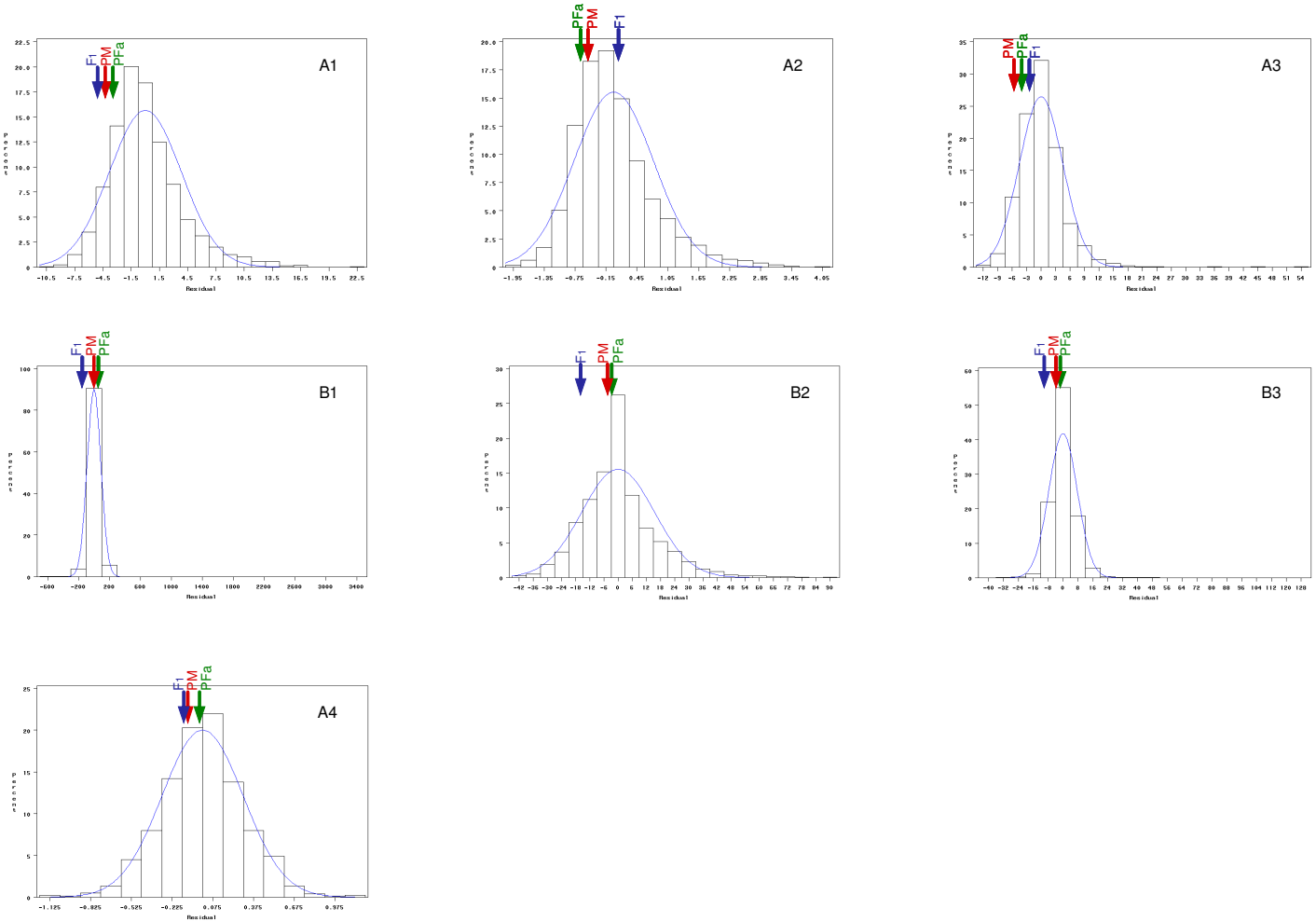


Figure 3: Normal distribution charts: Residuals of harvest data after log transformation. A) Greenhouse experiment B) Field experiment; Traits: (A1, B1) Fresh weight (g), (A2, B2) Dry weight (g), (A3, B3) Dry matter (%), (A4) leaf width (mm). Maternal line: red arrows, paternal line: green arrows, and F_1 line: blue arrows.

5.3.3 Heritability

The heritabilities for the three biomass traits of the experiment under investigation were high. For the trait fresh weight: 0.8033 for the greenhouse experiment and 0.8838 for the field experiment, for the trait dry weight: 0.9540 for the greenhouse experiment and 0.8838 for the field experiment and for the trait dry matter: 0.7780 for the greenhouse and 0.8614 for the field (Table 2) could be calculated.

Table 2: Broad sense heritabilities for fresh weight, dry weight, dry matter in the greenhouse and the field experiments.

	H^2 greenhouse experiment	H^2 field experiment
Fresh weight	0.8033	0.8838
Dry weight	0.9540	0.8838
Dry matter	0.7780	0.8614

5.3.4 Heading date

Heading date data were recorded from plants in pots (Figure 4) and from each heading plant in the field.

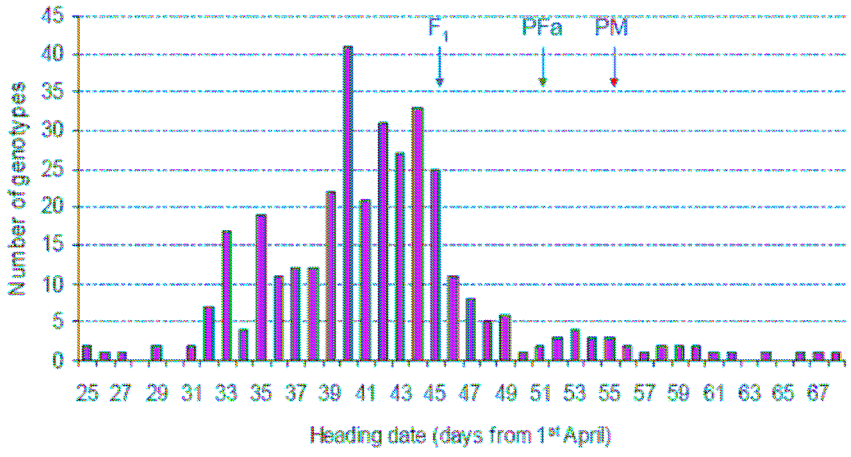


Figure 4: Distribution of heading date in the F_2 population in plants in pots outside the greenhouse. Heading date was defined as flowering time in days after the 1st of April.

Because of the incomplete field dataset for heading date QTL location analysis was less reliable. QTL for heading date were located on LG 3, 4 and 6 (Figure 5).

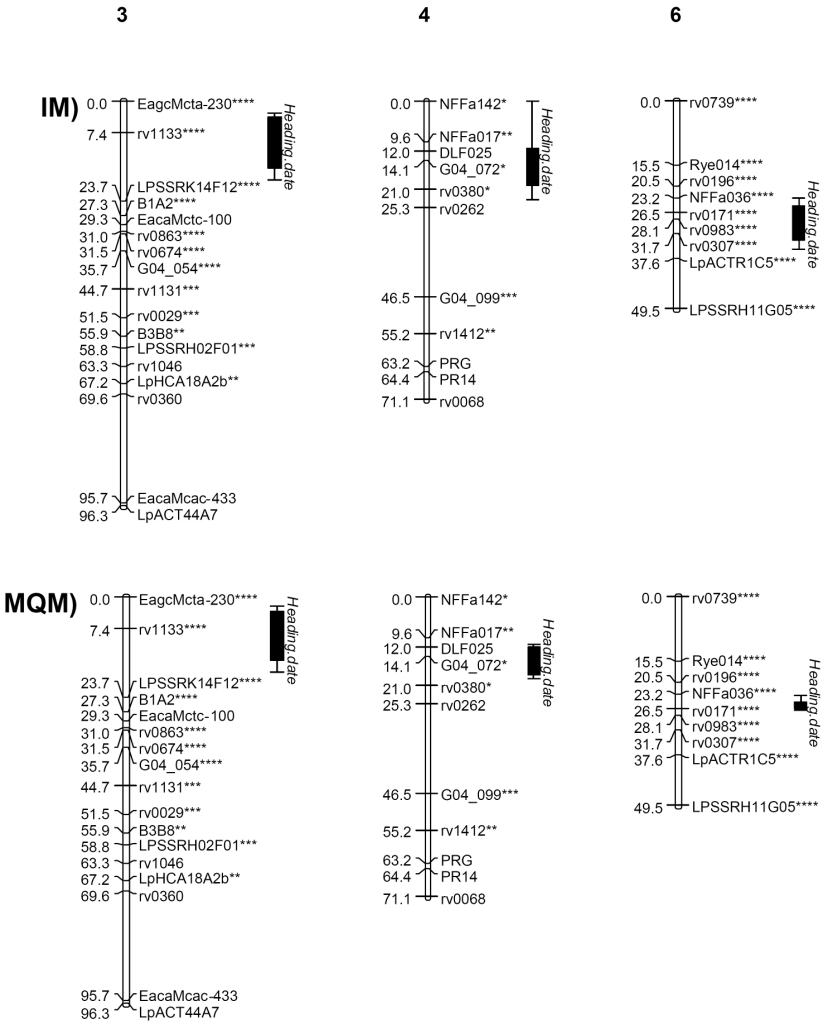


Figure 5: Assumed QTL locations for heading date in the field with an incomplete data set in the field. QTL were calculated by Interval mapping (IM) and below by multiple QTL models (MQM) mapping. QTL are shown with 1-LOD and 2-LOD support intervals.

5.3.5 Biomass data for each harvest and across harvests

MPH and BPH were calculated for the F_1 and the F_2 generations. The average values of the parental, F_1 and F_2 generations were calculated.

The highest values in the greenhouse experiment were found in the F_1 data, which ranged for fresh weight between 7 g and 12 g and for dry weight between 1 g and 3 g. The highest biomass yields for the parental genotype plants in the greenhouse were obtained during the third harvest (April 2007), which ranged for maternal fresh weight between 3 g and 9 g and for paternal fresh weight between 2 g and 8 g (Table 3 a). An increase of biomass heterosis was found for the three harvests and their replication (Each harvest and their replication data of each single harvest is separately displayed in the appendix tables 1 to 3 for the greenhouse experiment and appendix tables 4 to 7 for the field experiments). An exception was the harvest from the third harvest/third replication in the greenhouse experiment where a decrease in biomass heterosis for MPH and BPH was found. Biomass heterosis of the trait dry matter showed negative values in the greenhouse experiment except in the first harvest/third replication where positive heterosis for the F_1 was found. Heterosis values for fresh weight in the second harvest were the highest of the three greenhouse harvests (Table 3 a, Figure 6 a). They ranged from 274% for MPH F_1 and 171% for MPH F_2. BPH F_1 showed heterosis of 247%. BPH F_2 value showed the highest value in the first harvest with 154%. The highest heterosis data were obtained during the second harvest (February 2006), which is most likely the result of the winter period in between the first and the second harvest. MPH showed higher values than BPH and F_2 heterosis values were lower than F_1 heterosis values. The F_1 and F_2 genotypes had a continual decrease in fresh and dry weights.

Table 3 a: Biomass yield (in g) and dry matter (in %) data for the three greenhouse harvests averaged across the three replicates. Traits: Fresh weight (g), dry weight (g), dry matter (%). Average (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 generation and for the F_2 generation.

Greenhouse plant	December 05			February 06		
	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	2.5 ± 2.7	0.5 ± 0.4	21.4 ± 9.3	2.6 ± 2.8	0.6 ± 0.4	26.5 ± 5.5
Av PFA	2.7 ± 1.3	0.5 ± 0.2	21.3 ± 6.6	2.2 ± 1.1	0.5 ± 0.2	22.3 ± 4.8
Av F1	7.3 ± 2.4	1.3 ± 0.4	18.2 ± 2.0	9.0 ± 3.2	1.8 ± 0.6	21.2 ± 2.7
Av F2	6.8 ± 3.2	1.0 ± 0.5	15.2 ± 2.8	6.5 ± 3.8	1.2 ± 0.7	19.9 ± 4.4
MPH F1 (%)	186.2	167.8	-14.8	274.3	235.3	-13.1
BPH F1 (%)	174.8	150.1	-14.8	247.2	202.5	-19.9
MPH F2 (%)	164.1	105.1	-28.9	170.9	126.4	-18.6
BPH F2 (%)	153.6	91.6	-28.9	151.3	104.2	-24.9

Greenhouse plant	April 06		
	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	9.1 ± 5.9	2.1 ± 1.2	24.6 ± 4.2
Av PFA	8.3 ± 5.8	1.9 ± 1.3	22.5 ± 4.3
Av F1	11.9 ± 4.3	2.6 ± 0.8	22.3 ± 3.7
Av F2	10.1 ± 5.7	2.0 ± 1.1	21.2 ± 5.3
MPH F1 (%)	37.1	30.0	-5.5
BPH F1 (%)	31.3	23.7	-9.6
MPH F2 (%)	16.3	3.2	-9.9
BPH F2 (%)	11.3	-1.9	-13.8

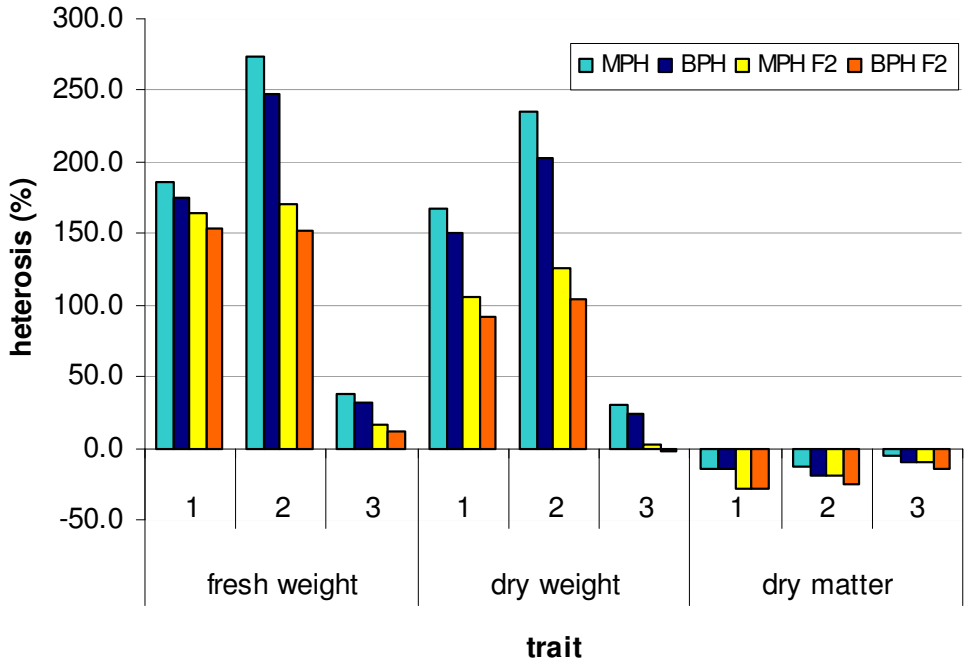


Figure 6 a: Heterosis data, mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation, for the three greenhouse harvests (1: December 05; 2: February 06; 3: April 06).

Across all harvests, biomass yields were much higher in the field than in the greenhouse experiment (Table 3 a b, Figure 6 a b). The highest values in the field experiment were found in the F_1 data, which ranged for fresh weight between 34 g and 186 g and for dry weight between 12 g and 49 g. The reason for this was that the harvests were taken from a sward of six plants per genotype and most of the plants grew better in the field with exception of the parental genotype plants. The parental inbred line genotypes performed very poorly under field conditions leading to incomplete data for the parental lines. 2006 was very dry and enabled only growth for two biomass harvests. Additionally, the plants were only planted in 2006 and an initial establishment phase was required before harvests were feasible. The most complete harvest dataset from the field experiment was gathered in late August 2007 (harvest 4) where sufficient material of the maternal and paternal genotypes could be collected. Generally, in the field differences among the replications were found. The genotypes in the second replication in the field overall performed worse than the genotypes in the first replication (Appendix Table 4-7). Data had a wider range in the field than the greenhouse, which could have been due to environmental influences like soil structure or water availability. Highest biomass heterosis yield in the field experiment were obtained in May 2007 (harvest 3) (Table 3b, Figure 6b) for MPH F_1 with 246%. The highest heterosis values of MPH F_2 with 166%, BPH F_1 with 103%, and BPH F_2 with 61% were found in the first field harvest. Plants were well established in 2007 and environmental conditions were more suitable for *L. perenne*. In the year 2006 more differences could be found between harvests. BPH had higher values than MPH.

Table 3 b: Biomass yield (in g) and dry matter (in %) data for the four field harvests. Traits: Fresh weight (g), dry weight (g), dry matter (%). Average (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2. Mid parental heterosis (MPH) and best parental heterosis (BPH) for the F_1 generation and for the F_2 generation.

Field	August 06			October 06		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	3.5	1.3	37.1	^	^	^
Av PFA	16.9 ± 16.0	5.9 ± 5.4	35.6 ± 5.0	86.5 ± 14.0	20.5 ± 1.0	24.0 ± 2.8
Av F1	34.3 ± 19.0	11.6 ± 6.4	34.5 ± 3.4	82.7 ± 44.1	18.2 ± 9.3	22.9 ± 3.9
Av F2	27.2 ± 23.1	9.4 ± 7.4	36.1 ± 5.6	63.8 ± 159.9	13.3 ± 11.0	26.3 ± 13.6
MPH F1 (%)	235.6	224.3	-5.2	^	^	^
BPH F1 (%)	102.9	98.3	-7.1	^	^	^
MPH F2 (%)	166.4	161.0	-0.7	^	^	^
BPH F2 (%)	61.0	59.6	-2.8	^	^	^

Field	May 07			August 07		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	10.9 ± 1.2	3.2 ± 0.3	29.5 ± 1.5	38.7 ± 40.1	11.5 ± 11.3	33.1 ± 11.2
Av PFA	96.7	26.5	27.4	96.0 ± 75.0	28.6 ± 23.2	29.6 ± 1.9
Av F1	186.3 ± 73.4	49.3 ± 19.5	26.5 ± 1.9	140.2 ± 43.7	39.2 ± 12.5	28.1 ± 2.8
Av F2	116.8 ± 90.2	31.0 ± 23.4	27.0 ± 2.3	114.5 ± 77.7	31.2 ± 20.2	28.4 ± 3.9
MPH F1 (%)	246.2	231.4	-6.9	108.1	95.2	-10.4
BPH F1 (%)	92.6	85.8	-10.2	46.0	36.9	-15.2
MPH F2 (%)	117.1	108.6	-5.1	69.9	55.3	-9.3
BPH F2 (%)	20.8	16.9	-8.5	19.2	8.9	-14.1

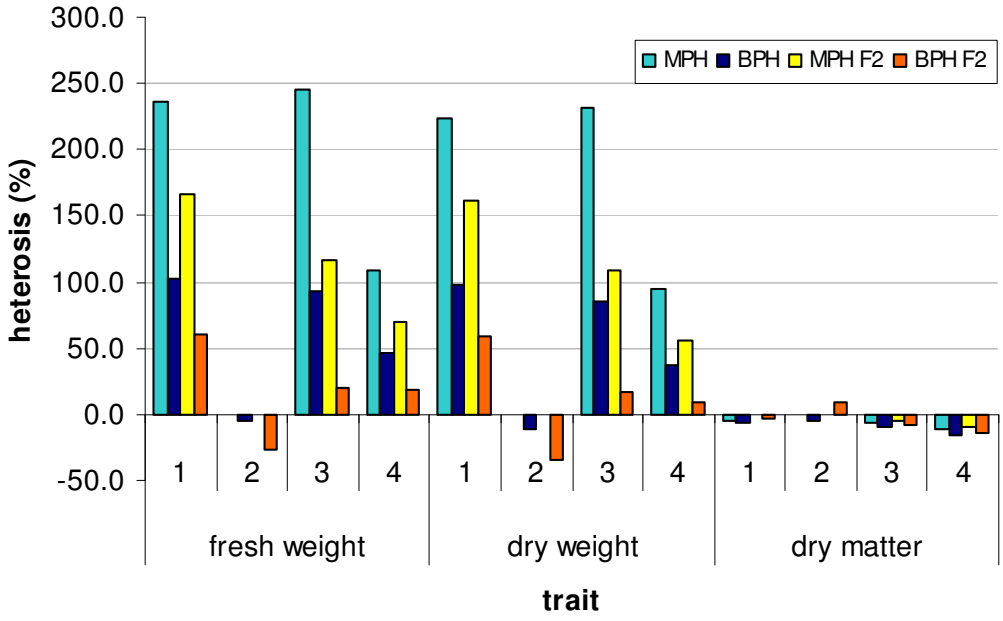


Figure 6 b: Heterosis data, mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and the F_2 generation for the four field harvests (1: August 06; 2: October 06; 3: May 07; 4: August 07).

In general, values were less high than in the greenhouse experiment. The same pattern could be seen between the different harvests except of the second harvest in October 2006, which could be due to the incomplete dataset of the parental values (Appendix data Table 5). The performance of the parental lines was low due to the very dry weather during the summer. Generally, yields were highest in the F_1 generation. The averages of the F_2 biomass heterosis values were higher than parental genotypes values, but lower than F_1 biomass heterosis values. The heterosis data of the greenhouse experiment showed compared to the field experiment higher values except for MPH of the F_1 genotype, which had in both experiments similar values.

5.3.6 QTL positions for the different traits in the F_2 mapping population

In order to detect QTL positions for the different traits (fresh weight, dry weight, dry matter and leaf width) IM and MQM mapping were applied (Figure 7 a b). QTL for each trait were calculated with the adjusted means of the ANOVA estimates with IM and MQM mapping algorithms.

For the trait leaf width QTL on LGs 2, 3 and 4 were detected. However the QTL on LG 2 did not reach the maximum LOD score threshold of 3.3 by IM. The QTL for leaf width on LG 3 consisted of four linked QTL. On LG 4 the leaf width QTL was much more distinguished with MQM mapping than with IM.

The locations for the fresh weight and the dry weight QTL consistently showed the same QTL positions with both QTL mapping methods on LGs 2, 3 and 7 (Figure 7 a b). A single QTL on LG 2 and a single QTL on LG 7 were found. In general the same QTL for dry and fresh weight could be located on LG 3. But analysis of the greenhouse fresh weight and dry weight data revealed only one QTL when using MQM mapping. An additional fresh weight QTL could be found on LG 5 in the field experiment.

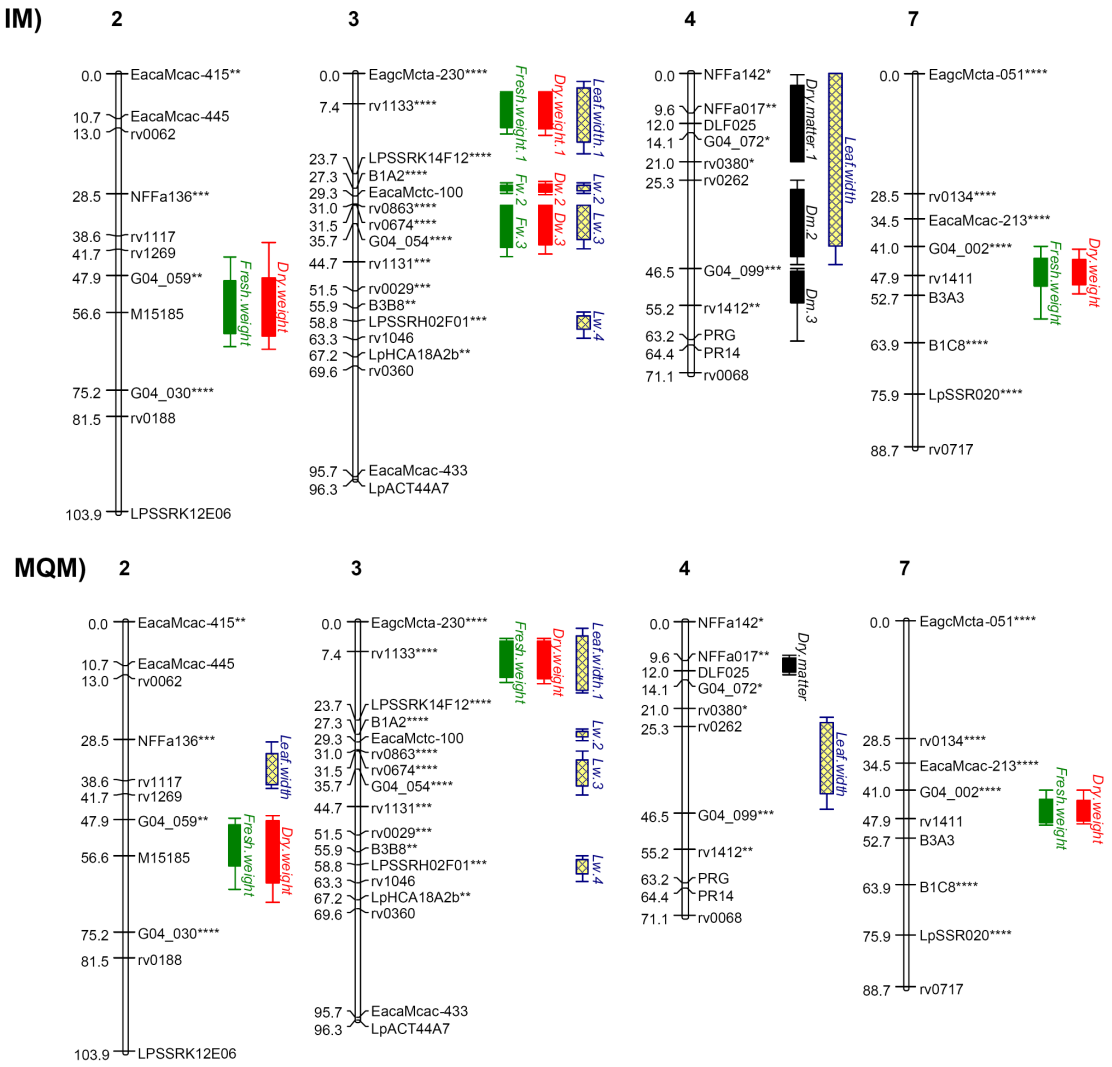


Figure 7 a: Quantitative trait loci (QTL) positions on linkage groups 2, 3, 4 and 7 of all measured traits in the greenhouse experiment. QTL were calculated by interval mapping (IM) and below by multiple QTL models (MQM) mapping. Green: fresh weight, red: dry weight, black: dry matter, yellow: leaf width. QTL are shown with 1-LOD and 2-LOD support intervals.

For the greenhouse experiment three single but neighbouring QTL for dry matter yield were identified by IM on LG 4. These QTL region were narrowed down to a single QTL by MQM. In the field experiment a dry matter QTL was detected on LG 3, but not in LG 4. Dry matter QTL had different positions in the greenhouse and field experiments.

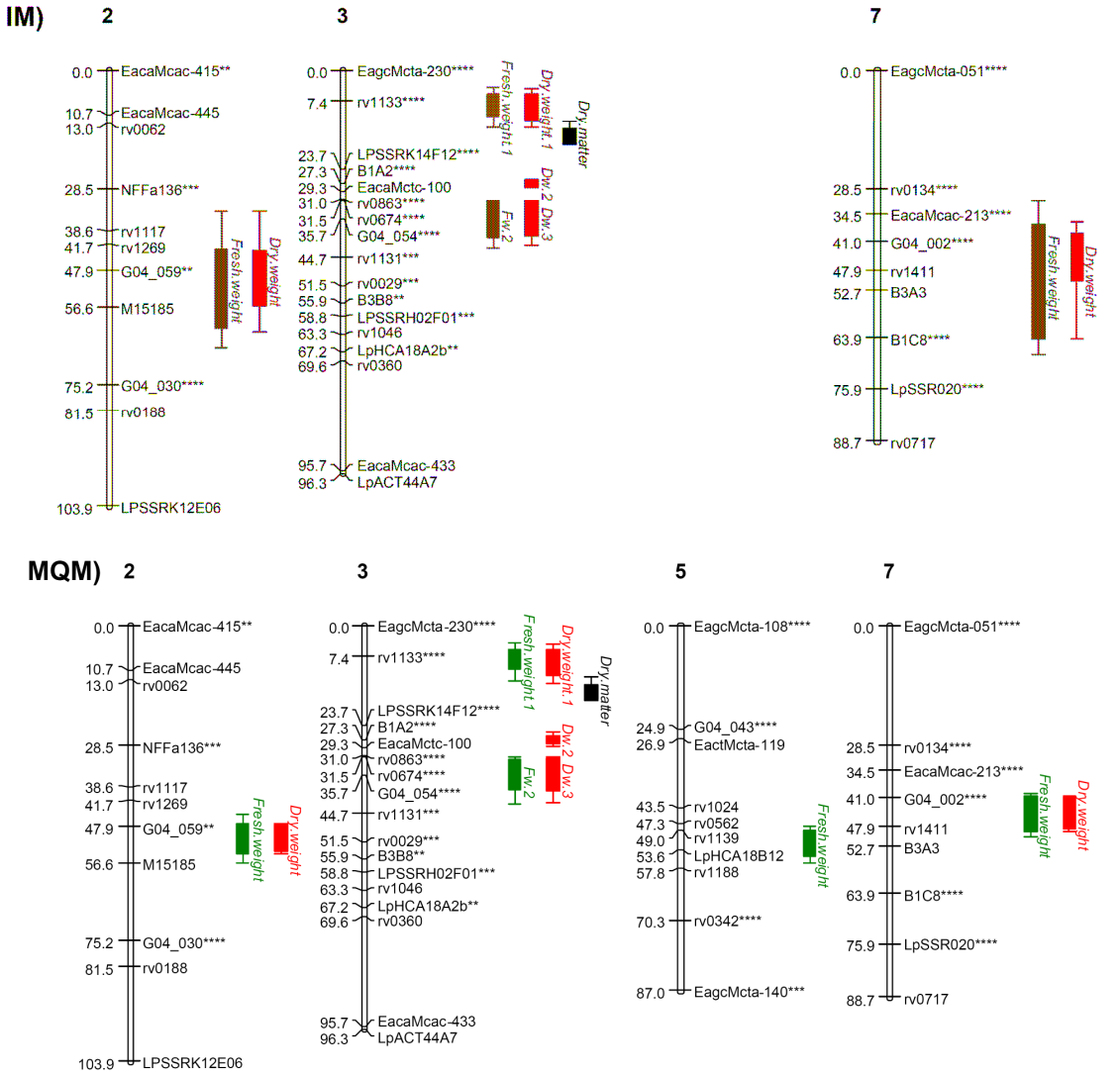


Figure 7 b: Quantitative trait loci (QTL) positions on linkage groups 2, 3, 5 and 7 of all measured traits in the field experiment. QTL were calculated with interval mapping (IM) and multiple QTL models (MQM) mapping. Green: fresh weight, red: dry weight, black: dry matter. QTL are shown with 1-LOD and 2-LOD support intervals.

QTL data were calculated and displayed by LOD scores and %explained variance of the QTL (Table 4). The location of the highest QTL LOD score was displayed by the length of the interval in cM and the marker positions on the LG. The maximum LOD scores which showed the significance level of QTL on the LG were given (Table 4).

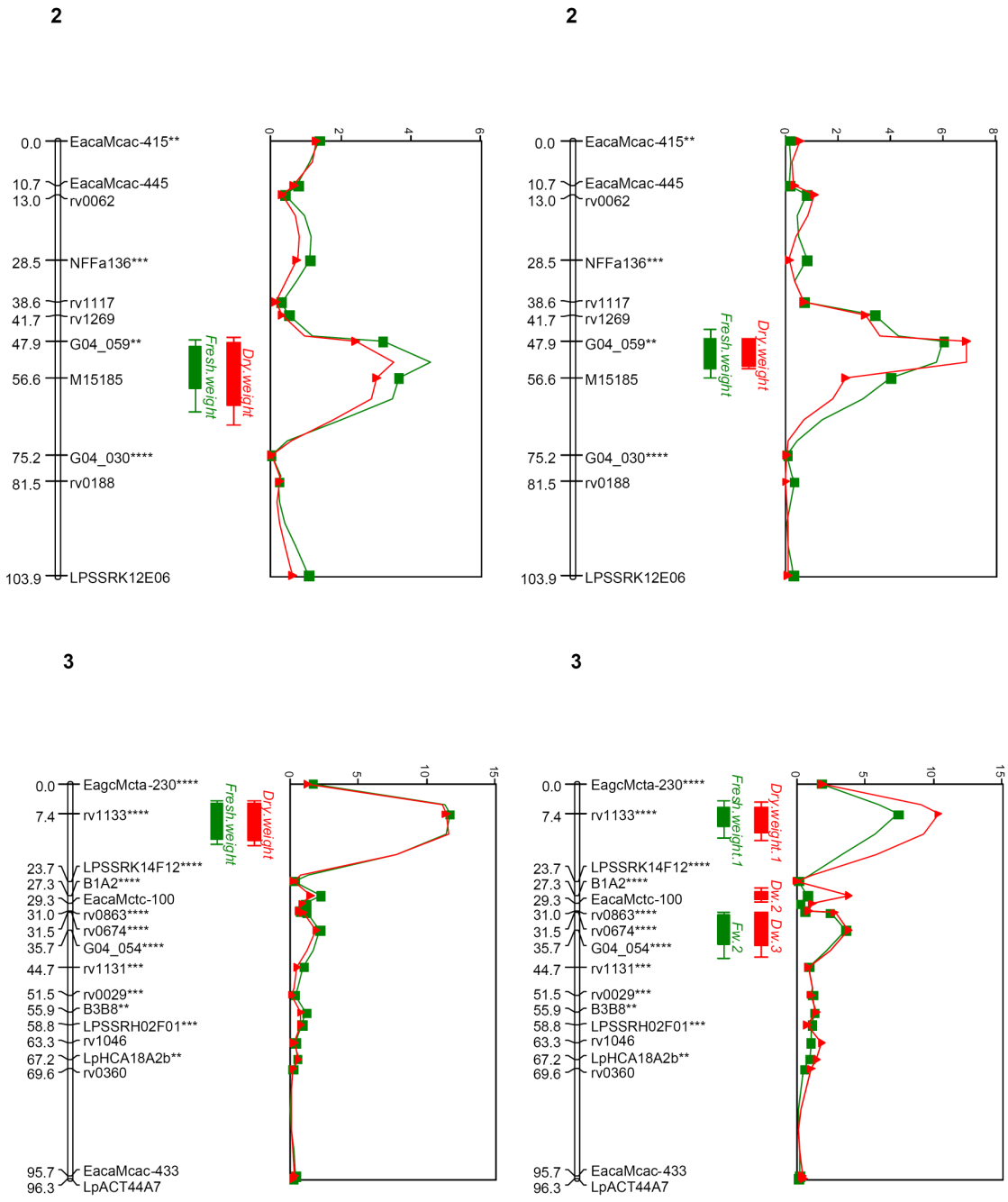
Table 4: Quantitative trait loci (QTL) and their locations for all greenhouse traits (leaf width, fresh weight, dry weight and dry matter) and for the field traits (fresh weight, dry weight and dry matter). Positions on linkage groups (LG) in cM, maximum LOD scores, percent of the phenotypic variation explained, marker intervals grouped by Interval (IM) and multiple QTL models (MQM) mapping.

		Interval mapping							MQM mapping				
		LG	Max. LOD score genome wide	LOD	%expl	cM	LOD sign. Threshold each LG	Marker interval	LOD	%expl	cM	Marker (co-factor)	Marker interval
Greenhouse	Leaf width	2	3.3	-	-	-	-	-	3.4	5.0	38.5	rv1117	NFFa136 - rv1269
		3	3.3	4.3	8.8	12.4	2.4	EagcMcta-230 - LPSSRK14F12	3.5	6.5	12.4	-	EagcMcta-230 - LPSSRK14F12
			3.3	7.1	8.7	27.3	2.4	LPSSRK14F12 - rv0863	6.1	6.5	27.3	-	LPSSRK14F12 - rv0863
			3.3	5.7	7.0	35.7	2.4	rv0674 - rv1131	5.9	6.5	35.7	-	rv0674 - rv1131
		4	3.3	6.1	12.2	30.3	2.3	rv0380 - rv1412	6.8	13.1	35.5	rv0262; G04_99	rv0380 - G04_099
	Total of variation explained				36.7					37.6			
Greenhouse	Fresh weight	2	3.4	4.5	8.6	52.9	2.3	rv1269 - G04_030	4.6	6.6	52.9	M15185	G04_059 - G04_030
		3	3.4	11.3	18.8	7.4	2.5	EagcMcta-230 - LPSSRK14F12	11.7	13.3	7.5	rv1133	EagcMcta-230 - LPSSRK14F12
			3.4	5.9	7.5	27.3	2.5	LPSSRK14F12 - rv0863					
			3.4	4.9	6.6	35.7	2.5	rv0674 - rv1131					
		7	3.4	6.4	8.7	47.9	2.5	EagcMcta-051 - LpSSR020	5.7	6.1	47.9	rv1411	G04_002 - B3A3
	Total of variation explained				50.2					26			
Field	Fresh weight	2	3.1	5.6	7.8	47.9	2.4	NFFa136 - rv0188	6.0	7.1	52.9	G04_059 / M15185	rv1117 - G04_030
		3	3.1	5.9	7.8	7.4	2.5	EagcMcta-230 - LPSSRK14F12	6.9	7.0	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
			3.1	5.4	6.8	35.7	2.5	rv0674 - rv1131	3.3	2.9	35.7		rv0674 - rv1131
		5	3.1	-	-	-	-	-	3.4	3.2	53.6	LpHCA18B12	rv1139 - rv1188
		7	3.1	3.5	4.9	47.9	2.1	rv0134 - LpSSR020	3.8	3.8	47.9	rv1411	G04_002 - B3A3
	Total of variation explained				27.3					24			

			Interval mapping						MQM mapping				
		LG	Max. LOD score genome wide	LOD	%expl	cM	LOD sign. Threshold each LG	Marker interval	LOD	%expl	cM	Marker (co-factor)	Marker interval
Greenhouse	Dry weight	2	3.4	3.9	7.4	52.9	2.4	rv1269 - G04_030	3.5	4.9	52.9	M15185	G04_059 - G04_030
		3	3.4	11.0	19.0	7.4	2.5	EagcMcta-230 - LPSSRK14F12	11.7	18.2	12.4	rv1133	EagcMcta-230 - LPSSRK14F12
			3.4	4.9	6.2	27.3	2.5	LPSSRK14F12 - rv0863					
			3.4	4.8	6.2	35.7	2.5	rv0674 - rv1131					
		7	3.4	7.9	10.9	47.9	2.2	EagcMcta-051 - LpSSR020	8.2	9.5	46.0	rv1411	G04_002 - B3A3
	Total of variation explained				49.7					32.6			
Field	Dry weight	2	3.3	6.3	9.3	46.7	2.4	NFFa136 - rv0188	6.9	7.4	47.9	G04_059 / M15185	rv1117 - G04_030
		3	3.3	8.0	12.6	7.4	2.6	EagcMcta-230 - LPSSRK14F12	10.4	10.4	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
			3.3	7.1	8.8	27.3	2.6	LPSSRK14F12 - rv0863	3.8	3.1	27.3		LPSSRK14F12 - rv0863
			3.3	6.8	8.8	31.5	2.6	rv0674 - rv1131	3.8	3.0	35.7		rv0674 - rv1131
		7	3.3	5.6	8.3	46.9	2.3	rv0134 - LpSSR020	6.3	6.8	46.0	rv1411	G04_002 - B3A3
	Total of variation explained				47.8					30.7			
Greenhouse	Dry matter	4	3.3	5.6	8.2	12.0	2.4	NFFa142 - rv0262	5.9	6.9	12.0	DLF025	NFFa142 - G04_072
			3.3	6.0	11.9	35.3	2.4	rv0262 - G04_099					
			3.3	5.5	9.0	51.5	2.4	G04_099 - rv1412					
	Total of variation explained				29.1					6.9			
Field	Dry matter	3	15.5	19.8	50.9	17.4	14.3	rv1133 - LPSSRK14F12	23.1	49.4	17.4	rv0863	rv1133 - LPSSRK14F12
	Total of variation explained				50.9					49.4			

5.3.7 Dry weight

QTL for fresh and dry weight were found on the same locations on LGs 2, 3, and 7. Dry weight showed two additional QTL in the field experiment on LG 3 (Figure 8 MQM mapping) and fresh weight were as well located on LG 5 but only in the field experiment.



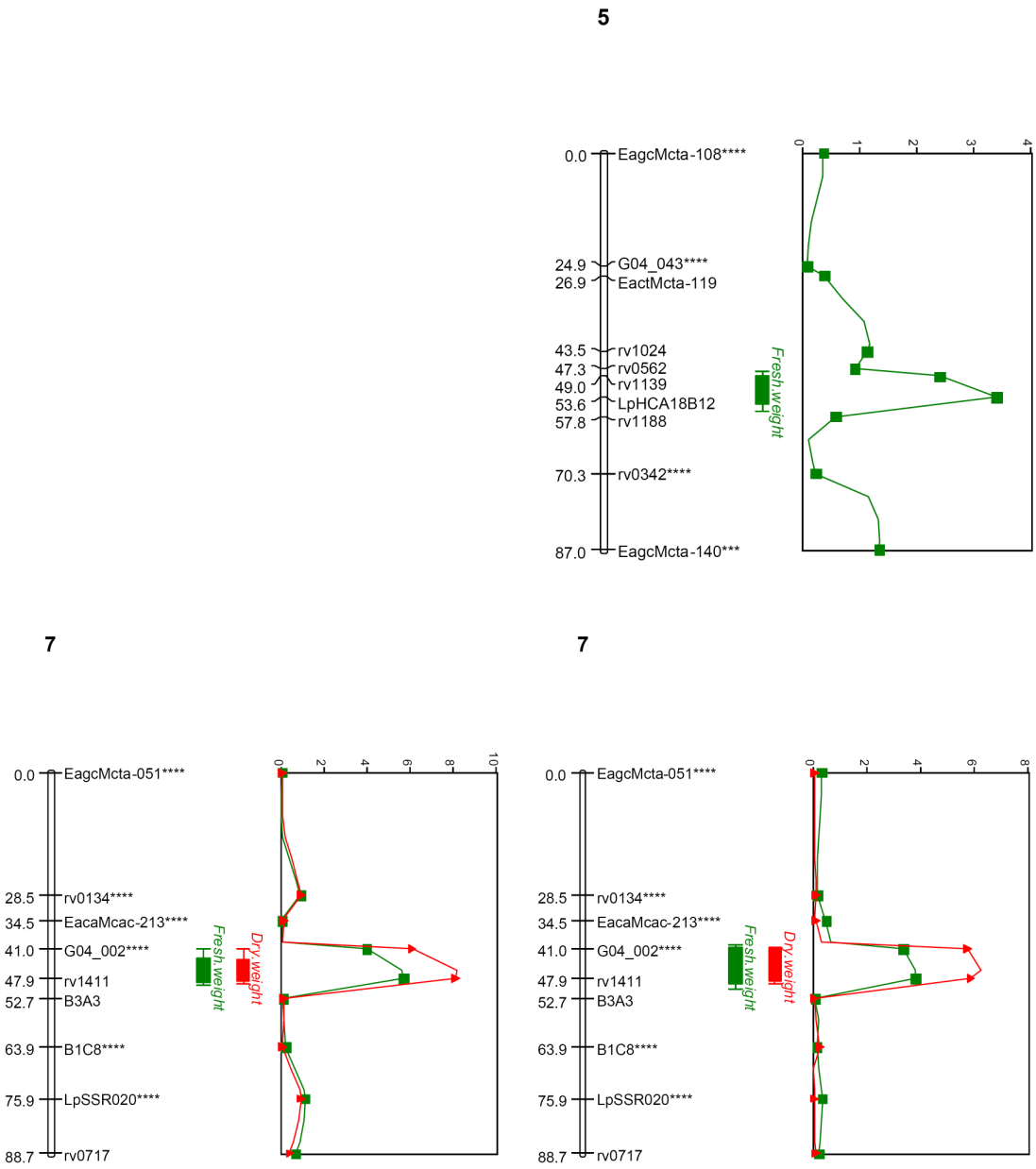


Figure 8: Greenhouse (left) and field (right) fresh weight (green) and dry weight (red) quantitative trait loci (QTL) positions (displayed with LOD scores) for linkage groups 2, 3, 5 and 7 as detected by multiple QTL models mapping.

Dry weight QTL data were reproducible over all harvests, two harvest years and environments (Table 5, Appendix Figure 6- 8).

.

Table 5: QTL data for the trait dry weight for each harvest and experiment (greenhouse and field): Positions on linkage groups (LG) in cM, maximum LOD scores, percent of the phenotypic variation explained, marker intervals grouped by Interval (IM) and multiple QTL models (MQM) mapping.

Dry weight				IM					MQM				
		LG	Max. LOD score genome wide	LOD	%expl	cM	LOD sign. threshold each LG	Marker interval	LOD	%expl	cM	Marker (co-factor)	Marker interval
Greenhouse	1. harvest	2	3.2	-	-	-	-		-	-	-	-	
	2. harvest		3.4	3.9	7.3	56.6	2.5	NFFA136 - G04_030					
	3. harvest		3.3	-	-	-	-		-	-	-	-	
	1. harvest	3	3.2	5.5	11.1	12.4	2.4	EagcMcta-230 - LPSSRK14F12	5.3	10.4	12.4	rv1133	EagcMcta-230 - LPSSRK14F12
				3.7	4.8	27.3	2.4	LPSSRK14F12 - rv0863					
				4.7	5.9	35.7	2.4	rv0674 - rv1131					
				3.9	5.0	58.8	2.4	rv0029 - rv0360					
	2. harvest		3.4	9.4	17.7	12.4	2.5	EagcMcta-230 - LPSSRK14F12	10.1	15.2	12.4	rv1133	EagcMcta-230 - LPSSRK14F12
				5.5	6.9	27.3	2.5	LPSSRK14F12 - rv0863					
				5.9	7.7	35.7	2.5	rv0674 - rv1131					
	3. harvest		3.3	7.1	13.9	7.4	2.5	EagcMcta-230 - LPSSRK14F12	7.0	8.9	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
	1. harvest	4	3.2	-	-	-	-		-	-	-	-	
	2. harvest		3.4	4.7	7.6	12.0	2.4	DLF025 - rv0380	6.9	6.7	9.6	DLF025	NFFa142 - G04_072
				4.5	11.3	30.3	2.4	rv0262 - G04_099					
	3. harvest		3.3	-	-	-	-		-	-	-	-	
	1. harvest	7	3.2	5.0	7.3	46.0	2.3	rv01234 - LpSSR020	4.5	5.5	47.8	rv1411	G04_002 - B3A3
	2. harvest		3.4	8.3	12.4	47.9	2.3	EagcMcta051 - LpSSR020	9.7	11.2	46.0	rv1411	G04_002 - B3A3
	3. harvest		3.3	3.4	4.8	47.9	2.2	EagcMcta051 - LpSSR020	3.5	4.4	46.0	rv1411	G04_002 - B3A3
Field	1. harvest	2	3.5	-	-	-	-		-	-	-	-	
	2. harvest		3.2	-	-	-	-		-	-	-	-	
	3. harvest		3.3	6.7	8.7	47.9	2.4	NFFa136 - G04_030	8.4	9.2	52.9	M15185	G04_059 - G04_030
	4. harvest		3.3	3.9	8.2	23.0	2.4	rv0062 - NFFa136	9.9	12.4	52.9	M15185	G04_059 - G04_030
				8.5	12.6	52.9	2.4	NFFa136 - rv0188					
	1. harvest	3	3.5	4.7	6.3	7.4	2.5	EagcMcta-230 - LPSSRK14F12	5.2	5.9	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
				5.5	6.9	27.3	2.5	LPSSRK14F12 - rv0863					
				5.3	7.0	31.5	2.5	rv0674 - rv1131					
				3.6	4.5	55.9	2.5	rv0029 - rv0360					
	2. harvest		3.2	-	-	-	-		3.5	4.2	7.4	rv1133	EagcMcta-230 - LPSSRK14F12

Dry weight				IM					MQM				
		LG	Max. LOD score genome wide	LOD	%expl	cM	LOD sign. threshold each LG	Marker interval	LOD	%expl	cM	Marker (co-factor)	Marker interval
	3. harvest		3.3	6.3	9.6	7.4	2.5	EagcMcta-230 - LPSSRK14F12	8.0	7.8	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
				5.0	6.4	27.3	2.5	LPSSRK14F12 - rv0863					
				5.4	7.8	35.7	2.5	rv0674 - rv1131					
	4. harvest		3.3	5.4	10.1	7.4	2.5	EagcMcta-230 - LPSSRK14F12	6.4	6.9	7.4	rv1133	EagcMcta-230 - LPSSRK14F12
				7.0	8.7	27.3	2.5	LPSSRK14F12 - rv0863	4.2	3.7	27.3		LPSSRK14F12 - rv0863
				6.4	8.0	31.5	2.5	rv0674 - rv1131	4.0	3.5	35.7		rv0674 - rv1131
	1. harvest	4	3.5	-	-	-	-		-	-	-	-	
	2. harvest		3.2	-	-	-	-		-	-	-	-	
	3. harvest		3.3	-	-	-	-		4.4	3.9	21.0	rv0380	G04_072 - rv0262
	4. harvest		3.3	-	-	-	-		-	-	-	-	
	1. harvest	5	3.5	-	-	-	-		-	-	-	-	
	2. harvest		3.2	-	-	-	-		-	-	-	-	
	3. harvest		3.3	3.6	4.7	53.6	2.4	rv0562 - rv0342	6.1	5.8	53.6	LpHCA1 8B12	rv0562 - rv1188
	4. harvest		3.3	-	-	-	-		-	-	-	-	
	1. harvest	7	3.5	4.3	5.9	47.9	2.4	rv01234 - LpSSR020	3.8	4.3	47.9	rv1411	G04_002 - B3A3
	2. harvest		3.2	4.4	6.1	47.9	3.6	rv01234 - LpSSR020	3.9	4.6	47.9	rv1411	G04_002 - B3A3
	3. harvest		3.3	6.8	13.5	39.5	2.3	EagcMcta-051	8.3	9.0	46.0	rv1411	G04_002 - B3A3
				6.9	10.6	46.0	2.3	rv01234 - LpSSR020					
	4. harvest		3.3	-	-	-	-		3.5	4.2	46.0	rv1411	G04_002 - B3A3

The graphs of the single harvests of the dry weight QTL LOD scores showed the highest LOD scores in the second harvest of the greenhouse experiment and in the third harvest of the field experiment (Table 5). These results were consistent with the heterosis data. For 2006 much weaker QTL LOD scores in the field were detected compared to 2007.

5.3.8. Correlation and regression among traits

Pearson correlation (greenhouse experiment):

Means, standard deviations, minimum and maximum values of the greenhouse data were displayed (Table 6 a). Fresh weight ranged from 1 g to 15 g, dry weight from 0.2 g to 3 g. Dry matter values ranged from 15% to 27%.

Table 6 a: Mean, standard deviation, minimum and maximum values of the greenhouse data for the traits fresh weight, dry weight, dry matter and leaf width.

Variable	Mean	Std Dev	Min	Max
Fresh weight (g)	7.7	2.6	1.0	14.8
Dry weight (g)	1.4	0.4	0.2	2.6
Dry matter (%)	18.9	1.6	15.3	27.3
Leaf width (mm)	2.6	0.5	1.3	3.7

All traits had significant correlated values (Table 6 b). Fresh weight and dry weight were significantly strong positive correlated. Dry matter and fresh weight, dry matter and dry weight, and dry matter and leaf width were negatively correlated.

Table 6 b: Pearson correlation coefficients among the traits fresh weight, dry weight, dry matter and leaf width of the greenhouse data. (Ho: Rho=0)

	Fresh weight	Dry weight	Dry matter	Leaf width
Fresh weight	1			
Dry weight	0.97695 < 0.0001	1		
Dry matter	-0.38038 < 0.0001	-0.26145 < 0.0001	1	
Leaf width	0.38908 < 0.0001	0.36486 < 0.0001	-0.30619 < 0.0001	1

Regression (greenhouse experiment):

In the stepwise regression model for the greenhouse experiment the trait leaf width was excluded as non-significant parameter.

Table 7: Parameter estimates of the regression analysis of dry weight of the greenhouse experiment data.

Variable	Parameter estimate	Standard error	t Value	Pr > |t|
Intercept	-0.64482	0.06684	-9.65	<.0001
Fresh weight	0.17930	0.00187	95.97	<.0001
Dry matter	0.03556	0.00283	12.59	<.0001
Leaf width	0.00593	0.00984	0.60	0.5473

The regression model for dry weight was stated as follows:

Y_i (dry weight) = -0.645 (intercept) + 0.179 x_{i1} (fresh weight) + 0.036 x_{i2} (dry matter)

Pearson correlation (field experiment):

Mean, standard deviations, minimum and maximum values of the field data showed a wide range between minimum and maximum values (Table 8 a). Fresh weight ranged from 0.0001 g to 591 g, dry weight from 0.0001 g to 67 g. Dry matter values ranged from 18% to 63%.

Table 8 a: Mean, standard deviation, minimum and maximum values of the field experiment data for the traits fresh weight, dry weight and dry matter.

Variable	Mean	Std Dev	Min	Max
Fresh weight (g)	79.8	53.3	0.0001	591.3
Dry weight (g)	21.0	11.8	0.0001	66.9
Dry matter (%)	29.5	4.1	17.5	62.6

Pearson correlation coefficients for fresh weight, dry weight and dry matter showed significant correlated data (Table 8 b). Fresh weight and dry weight had a strong

positive correlation, whereas dry matter and fresh weight, and dry matter and dry weight were negatively correlated.

Table 8 b: Pearson correlation coefficients among the traits fresh weight, dry weight, dry matter of the field experiment data (Ho: Rho=0)

	Fresh weight	Dry weight	Dry matter
Fresh weight	1		
Dry weight	0.86993 < 0.0001	1	
Dry matter	-0.31915 < 0.0001	-0.31978 < 0.0001	1

Regression (field experiment):

The stepwise regression model excluded for the field experiment dry matter as non-significant parameter.

Table 9: Parameter estimates of the regression analysis of dry weight of the field experiment data.

Variable	Parameter estimate	Standard error	t Value	Pr > |t|
Intercept	10.03041	2.52096	3.98	<.0001
Fresh weight	0.18770	0.00612	30.69	<.0001
Dry matter	-0.13371	0.07816	-1.71	0.0880

The regression model for dry weight was stated as followed:

Y_i (dry weight) = 10.030 (intercept) + 0.188 x_{i1} (fresh weight)

All traits were significantly correlated in the greenhouse and in the field experiment. Dry matter was always negatively correlated to other traits. Regression analysis showed that all traits were strongly dependent on each other, with the exception of leaf width in the greenhouse and dry matter in the field experiment.

5.4 Discussion

In this study 360 F_2 genotypes were used to create a genetic map and to carry out the QTL study. This mapping population is to date the largest reported *L. perenne* mapping population. QTL experiments were carried out under two different environments, with three replications for the greenhouse and two replications for the field experiments. Multiple harvests were taken for both experiments. QTL detection was carried out for the traits fresh weight, dry weight, dry matter and leaf width. All QTL for the traits fresh weight, dry weight, and dry matter were reproducible over environments, replications, and over two years, thus data derived from this study can be considered as reliable.

5.4.1 Heterosis

Values of heterosis differ between plant species. The model plant *Arabidopsis* showed in several F_1 hybrids strong heterosis effects. Biomass yield ranged between 5% and 140% for MPH and -10% and 122% for BPH (Barth *et al.* 2003). Heterosis for forage yield in seven parental clonal crosses of smooth bromegrass ranged between -5% and 28% for MPH (Casler *et al.* 2005). Biomass heterosis in chickpea hybrids ranged for MPH from -11% to 59% and for BPH from -22% to 39% (Hegde *et al.* 2007).

Heterosis for agronomic traits in *L. perenne* was described before for leaf length, leaf width, seedling dry weight, seedling tiller number, and summer regrowth (Humphreys 1992). Heterosis values in the present study for biomass yield related traits showed higher values compared to other studies with a range for fresh weight biomass of 19% to 137% (MPH) and of 31% to 247% (BPH). This reflects the inbred background of the parental lines. This is indicative of the importance of the choice of

parental genotypes for a heterosis study. All biomass yield related data were highest for fresh and dry weight in the F_1 generation, which is in accordance to the heterosis theory (Becker 1993).

Data in the present study over the different replications included frequently missing values for the parental lines since it was very difficult to keep the inbred plants alive. Therefore field data of the two replications were calculated together to obtain a data set including sufficient data for both parental lines.

During the first harvest of the greenhouse experiment positive heterosis for F_1 dry matter in the third replication was found. An explanation could be that the environmental influences on the plants in the third replication in the greenhouse experiment were different than in the other two replications. The decreasing heterosis data can be explained by the much stronger growth of the plants over time and faster development of the F_1 and F_2 plants compared to the parental plants. The weather conditions in 2007 did not affect the performance of the F_1 and F_2 plants as much as it did for the parental genotype plants in the field. Heterosis values were lower for the fourth harvest in the field experiment than the values of the previous harvest.

Negative heterosis for dry matter could be explained by increased cell sizes of the F_1 and F_2 plants compared to the parental lines. The cells could have been bigger and contained more water in F_1 and F_2 genotype, which was removed by drying. This is also in agreement with the results of the correlation analysis where dry matter content and leaf width were negatively correlated. Also in the heterosis data it seems that in the F_1 and F_2 genotypes more water was stored than in the parental plants, which would explain that dry matter is increasing when leaf width is decreasing. This could indicate again that the cells in the F_1 and F_2 were bigger and contain more water, which was removed when the material was dried in the oven. Similar findings for primary roots

were reported by Hoecker *et al.* (2006) in maize where the cortical cells were expanding in the hybrid plants. Another study in maize explained the more rapid growth rate of leaf blades in hybrids by increased cell number or enlargement of the cells (Uchimiya and Takahashi 1973).

Heterosis values under field conditions were higher for all replications when the plants had more time to grow between the harvests like in spring 2007. The same pattern can be seen in the greenhouse experiment where the heterosis values were higher in the first spring harvest compared to the harvests during winter and late spring. It seems that parental plants grew less over time and the F_1 plants generated more biomass over a longer period especially in spring when the biomass production in the F_1 was higher compared to the parental plants.

5.4.2 QTL studies in Lolium perenne

Several QTL studies have been carried out in *L. perenne* in recent years (Table 1). However, only three studies reported on traits analysed in this study. Yamada *et al.* (2004) identified fresh weight QTL on LG 5 and leaf width QTL on LG 3 in the 'ILGI' population. Turner *et al.* (2008) reported on fresh weight QTL on LG 1, 4, and 5 and autumn dry matter QTL on LG 3 using the F_2 'WSC' population. Armstead *et al.* (2008) identified QTL for flag leaf width on LG 7 in the F_2 'WSC' population. Yamada *et al.* (2004) identified QTL using simple interval mapping (SIM: Lander *et al.* 1987) and composite interval mapping (CIM: Zeng 1994) in the ILGI mapping population. The authors identified one QTL for fresh weight on LG 5 (LOD_{SIM} 4.2, expl. $variance_{SIM}$ 23%; LOD_{CIM} 5.6; expl. $variance_{CIM}$ 22.8%) and one QTL on LG 3 for leaf width (LOD_{SIM} 2.1, expl. $variance_{SIM}$ 11.3%; LOD_{CIM} 2.3; expl. $variance_{CIM}$ 10.4%). A QTL for fresh weight on LG 5 could be identified in the present study as well but only in the

field experiment and only with the MQM mapping method. The present study and the study of Yamada *et al.* (2004) share only a few common markers. It is speculative if these QTL could be identical. A QTL for leaf width was also found in this study on LG 3 in the greenhouse experiment. The leaf width QTL on LG 3 of Yamada *et al.* (2004) could be identical with the QTL for leaf width detected in the present study. Some markers in the study of Yamada *et al.* (2004) were common for both studies, e.g. the leaf width QTL region shared one common marker (LPSSRK14F12). Turner *et al.* (2008) identified one QTL for fresh weight on LG 1 (LOD_{MQM} 6.8, expl. $variance_{MQM}$ 17.6%), one on LG 4 (LOD_{MQM} 5.9, expl. $variance_{MQM}$ 13.5%) and two QTL on LG 5 (LOD_{MQM} 11.6, 7.4, expl. $variance_{MQM}$ 28.5%, 15.7%). Additionally, an autumn dry matter QTL was located on LG 3 (LOD_{MQM} 5.7, expl. $variance_{MQM}$ 11.1%) by Turner *et al*. (2008). QTL for fresh weight and for dry matter could be located as well on LG 5 in the field experiment in the present study. But no common markers were shared with the QTL found by Turner *et al*. (2008). Armstead *et al.* (2008) found one QTL for flag leaf width on LG 7 with Interval mapping and two QTL with MQM mapping (LOD_{IM} 8.9, expl. $variance_{IM}$ 20.4%; LOD_{MQM} 3.5, 2.8; expl. $variance_{MQM}$ 5.4%, 4.2%).

Dry matter QTL could be found on two different LGs in the greenhouse and in the field experiments. This is an indication that the traits are probably explained by multiple gene interactions. Additionally environmental interactions are influencing the QTL as well. Also the regression studies indicate different environmental influences in the greenhouse and in the field experiment pointed out by dry matter, which is not influencing the regression model in the field but in the greenhouse. An additional QTL for fresh weight could be detected on LG 5 in the field experiment by MQM mapping, which could be probably similar explained by either multiple gene interactions or environmental interactions.

In this study 33% over all phenotypic variance was explained for dry weight with MQM mapping in the greenhouse and 31% for the field experiments. It is difficult to decide if any unexplained variance is caused by other QTL or to environmental influences. Dry matter showed along the different environments and the two different experiments a QTL × environment interaction in the present study. Genotype x environmental interactions are very important for the expression of QTL (Zhuang *et al.* 1997) especially when stresses such as water or drought stress are components of the interaction. Breeders usually use this variation via phenotypic selection to develop varieties with general and specific acceptance (Snape *et al.* 2007). Environmental interactions influence different genes in different environments. Higher biomass yield data in the greenhouse might have been the result of conditions in the greenhouse, where higher irradiance and variable temperatures were experienced, and have been influenced by different genes in the greenhouse experiment than in the field experiment. The QTL locations on the chromosomes were still the same but some showed a drift of the shape of the QTL graphs, e.g. LG 2 for fresh or dry weight (Appendix Figures 3 and 4: A1 and B1). Also different locations of QTL in the greenhouse and the field experiment like the dry matter QTL on LG 4 in the greenhouse and on LG 3 in the field experiment (Appendix Figure 5a and 5b) guide to this conclusion. There are several examples of QTL × environmental interactions (Jansen *et al*. 1995, Snape *et al.* 2007, Stuber *et al.* 1992), showing that the expression of particular chromosome regions differ across environments. QTL of large contribution to the whole variance of the trait are generally detected through different environments and progenies (Grandillo *et al.* 1999). To be sure to have identified true QTL it is necessary to validate them under several environments with a large amount of genotypes and in several replications.

Segregation distortion might have an impact on QTL studies. Maps with high segregation distorted loci as in the present study influence the order of markers on a map or the length of the map (Hackett and Broadfoot 2003). Therefore, segregation distorted markers in a mapping population should not be ignored for further work or eliminated from further calculations. These markers distort distances of genetic markers on a map and can lead to an underestimation of the required marker numbers for fine mapping studies and therefore might reduce the reliability of a QTL position. However, since SD affected complete chromosomes or chromosomes arms, it appears to be a genuine biological phenomenon in the mapping population of the present study. Thus markers with segregation distortion cannot be removed from the genetic map for QTL analysis and segregation distortion has to be taken into account to influence the fine definition of QTL positions in the study.

It is controversial if an experiment with spaced plants is reliable for a biomass study. Forage yield measured in spaced plant are considered as a poor predictor of yield (Hayward and Vivero 1984). Although, there are studies supporting the hypothesis that spaced plant selection can be used to increase forage yield performance (Bruckner *et al.* 1991; Burton 1992). In the present study similar results were obtained from experiments using artificial mini swards in the field and single plants in the greenhouse. The QTL were not less robust from the greenhouse experiment than the QTL data from the field. Also the QTL locations were the same and showed that it was possible to use single plants to achieve a reliable QTL analysis.

The analysis present in this study gives an insight into important growth characters. The data do not explain the complete genome wide variance of the traits investigated. It would be necessary to carry out additional QTL studies in other

populations to identify all QTL affecting traits like dry weight or leaf width. However, to identify all QTL contributing to a trait can be challenging and it might be most of the time not possible since QTL can be located differently in different populations and are probably highly influenced by the environment. But the results obtained by the present study can be used for fine mapping of genetic regions containing larger QTL and cloning of the underlying genes of interest. However, they do not explain the whole genome-wide variance of the trait. Association mapping (also known as linkage disequilibrium mapping) might be an alternative to QTL mapping. But it could be difficult to map the trait biomass via association mapping since in natural populations biomass effects are smaller than in F_2 populations with inbred line parental background or in RILs. And association mapping analysis is limited by the lack of recombination in the genomic analysis (Liu 1998). Therefore, it is necessary to have a large amount of samples to map biomass related traits by an association mapping approach, which is time consuming and expensive.

The traits studied in the present study, which all lead to higher dry weight, are located on the same chromosome, which could eventually make fine mapping of the traits at the same time easier. But other studies concluded that it might be difficult to extrapolate the results of marker analysis of complex traits to populations with different backgrounds or to different environments (Reyna and Sneller 2001). The present study showed a robust dataset of heterosis data and gave clear QTL regions for biomass. The robustness of the study was as well mirrored in the reproducibility of the QTL in the replications, over environments and over the years. It is therefore very promising for MAS and for the identification of candidate genes in the future. But it will be still necessary to refine the genetic map to get closer to the desired candidate genes and to make the data more reliable.

Chapter 6: General conclusion

The aim of the *in situ* hybridization analysis described in the second chapter of this thesis was to characterise the parental and F_1 generation of a molecular breeding programme for RILs in *L. perenne*. Recombination in the F_1 generation and identification of introgressed segments of fescue into the inbred parental lines were analysed. No distinct non-recombined fragments of one of the parents and introgressed fescue segments could be identified with *in situ* hybridization. The aim of the metabolite analysis of the parental and the F_1 genotypes was to test for differences in metabolites between the parental and F_1 genotypes and the three harvests. Analysis was based on statistical methods like PCA. The goal of the fourth experimental chapter was to generate a genetic map of the F_2 population. A genetic map with 75 nuclear DNA markers with an average marker density of 8 cM and a length of 592 cM spanning seven linkage groups was developed. The quantitative trait QTL analysis used the genetic map to combine the genotypic data with phenotypic data for biomass yield trait components. Two experiments were carried out, one in the greenhouse and one in the field. Three harvests in the greenhouse experiment and four harvests in the field experiment were done. Fresh weight and dry weight were measured and dry matter was calculated. Additionally, leaf width was measured in the greenhouse experiment and in the field experiment heading date was recorded. Major QTL for dry weight and fresh weight on LGs 2, 3 and 7 were located in both experiments. Leaf width QTL were found on LGs 3 and 4. Dry matter QTL on LG 4 in the greenhouse and on LG 3 in the field experiment could be identified.

The experimental design and statistic evaluation are critical steps in any scientific project. The alpha lattice design used in the QTL analysis of the present study was laid out as an incomplete block design. It reduces the total error of a study and gives the possibility to calculate block effects through included local controls. The experiments

were carried out in replicates and samples were randomized within each replication. No significant block effects could be found within an experiment in this study. The heritability shows the reliability of the phenotypic value of a metric trait. The heritability was high in the present experiments for each of the biomass traits, which demonstrated the strength of the data for the traits. Correlation studies confirmed the relationship among biomass yield related traits. Correlations are particularly important for the selection for a trait in breeding. If a trait is correlated to another trait changes in one trait cause a corresponding change in the other trait (Acquaah 2007). Regression analysis helped to study the relationship between the variables. In the present study dry matter in the greenhouse experiment could be identified as a variable for biomass yield in the regression model for biomass yield. This result was not found in the field experiment and can be explained by multiple gene action.

The climate conditions had an important influence on the outcome of the study. The very dry weather conditions in summer 2006 after the planting of the plants influenced plant growth and adaptation in the field. An environmental influence could be seen between the two main experiments in the greenhouse and the field. There was no significant difference within the single experiments in the greenhouse or field. Between the entire greenhouse and the field experiments climatic differences resulting in different gene action for the traits fresh and dry weight on linkage group 3 and inconsistent QTL for the trait dry matter were found. Also the metabolite analysis showed differences in the activity of metabolites during the season, especially with a strong effect on the paternal genotype.

Different mating schemes can be used for mapping and QTL studies. So far only backcross, F_1 or F_2 populations were utilized for genetic maps in *L. perenne*. These breeding schemes were used successfully for genetic mapping in several *Lolium* studies

(e.g. Bert 1999, Armstead *et al.* 2004, Jensen *et al.* 2005). For a mapping study a well studied population with a reasonable amount of genotypes is required. Genotype quantities in mapping populations generally range between 50 and 250 genotypes. In the present study 360 genotypes were included in the mapping population. This gives a reputable amount of genotypes for a mapping population and is the largest mapping population for *Lolium* reported so far.

In general *in situ* hybridization is an excellent tool to connect the physical map with the genetic map and would provide additional information about the genotypes. In this study the recombination among the parental genotypes was complete. No distinct segments of one or the other parent could be found with genomic *in situ* hybridization. However, on a more fine scale level the genomes of the two parental lines must be sufficiently different leading to the expression of larger differences in the metabolites. Especially the paternal genotypes showed differences in the metabolic activity, which could explain the differential adaptation of the genotypes to the environment and therefore, different biomass yields and trait expression in overall plant architecture and development. In this study high biomass yields could also be explained by the phenomenon heterosis. Heterosis was highest in the F_1 generation, which is in agreement with the literature for hybrids of two inbred parental lines (Becker 1993).

The presented biomass QTL analysis is useful for future MAS and breeding programmes. Marker intervals flank the QTL for biomass yield but for the application in MAS the QTL areas are at present still too large. Therefore, it is necessary to identify tightly linked markers located 1 cM or probably less then 1 cM to the trait of interest (Mohan *et al*. 1997). MAS technology can add more exactness to plant breeding programmes. Gene combinations can be targeted with more precision, and can lead to a better management of specific agronomic traits (Yamada *et al*. 2005). Thus MAS could

potentially speed up the conventional breeding programme. Therefore, a fine mapping study has to be carried out to more precisely identify the QTL regions. Another option could be remapping of the biomass QTL in the RILs. RIL lines can be immortalized in comparison to the F_2 genotypes and the RILs are more homozygote than the F_2 genotypes. A long term prospect could be as well QTL cloning to identify genes, which control the trait of interest. This would be most desirable for real heterotic QTL (Lippman and Zamir 2006).

The biomass phenotyping and QTL study showed that the data were reproducible over all environments and replications. This demonstrates the reliability of the outcomes of this study and how trustable the results are. In the long term an application of the QTL study would be to trace the components of biomass to be beneficial for the conventional Oak Park breeding programme using MAS. The aim of the Oak Park breeding programme is to develop cultivars with enhanced biomass yield, quality, tolerance to environmental factors and seasonal adaptation, and diseases resistances (Connolly 2001).

Appendix

Appendix Table 1: Metabolite quantitative values. Table displays genotypes: PM: maternal; PFa: paternal; F_1 and harvest times (time 1: 12.06.06; time 2: 15.08.06; time 3: 19.10.06).

Metabolite	PM			PFa			F1		
	time1	time2	time3	time1	time2	time3	time1	time2	time3
alanine	20572.8	36287.5	60015.9	104046.4	77284.0	92276.7	64500.6	41094.0	35124.3
alpha ketoglutaric acid	505.3	2716.4	568.8	686.2	1068.6	1236.7	968.0	818.4	845.5
arabitol	1165.6	4180.3	860.7	938.3	5217.1	2309.3	1040.5	12860.8	1764.9
asparagine	12917.8	91893.1	77098.3	13244.6	37703.3	211064.2	5113.2	17113.0	15532.4
aspartic acid	23687.0	22360.6	41617.3	22369.5	49685.7	38021.0	33282.6	56888.4	63978.6
benzoic acid	917.1	1108.8	1102.4	1249.7	1743.0	1689.6	830.1	943.5	2626.8
beta alanine	1141.3	1535.1	866.2	3137.1	2393.5	1874.7	2463.1	1891.7	865.8
caffeic acid	1369.7	3231.9	1408.1	1390.6	2115.7	2097.2	1549.4	2494.7	2764.7
citric acid	36259.4	47228.2	60579.3	30503.6	108864.	47202.5	76402.6	135277.0	204059.9
cyano-L-alanine	345.3	1597.8	1552.9	368.6	833.5	3572.3	191.8	517.5	520.5
erythritol	2574.4	4425.9	1812.2	638.7	733.0	234.2	535.0	778.3	532.9
erythronic acid lactone	8111.9	7008.1	5306.1	5485.1	8177.1	6448.9	4111.3	12314.4	16813.4
fructose 1	102315.9	116869.9	155540.2	40879.6	26683.4	44320.8	89192.1	42106.7	85062.0
fructose 2	84440.9	96296.4	135434.8	32898.0	21539.1	33657.5	71779.1	34412.7	64571.4
fructose-6-phosphate	87.1	106.6	377.2	193.7	304.5	458.2	223.8	155.5	374.0
fucose 1 + rhamnose 2	1122.5	1873.7	1659.8	1189.1	3115.1	1778.8	1102.1	3259.6	2663.1
fumarate	1754.1	2317.8	2433.7	2156.3	4886.0	3390.6	2310.7	5289.0	6563.3
GABA	1405.9	2251.9	1818.7	1565.4	6542.1	1892.6	1977.0	3733.2	1447.6
galactinol	13917.5	10450.9	35024.0	12480.1	8162.7	15888.7	16323.3	16858.2	38698.4
glucoheptulose	374.5	773.9	292.8	449.3	937.6	385.3	530.8	964.3	434.2
glucose 1	355905.1	209799.1	191116.1	313847.0	96278.5	84249.4	323655.8	132983.7	100162.6
glucose 2	88252.1	35880.4	37128.6	72962.1	15459.0	13104.9	99118.0	21299.9	16010.8
glucose-1-phosphate	23182.6	35550.5	30706.5	29863.7	58644.7	49151.2	29618.3	43802.3	43118.7
glucose-6-phosphate 2	151.4	144.5	631.9	393.7	508.2	769.9	441.7	283.6	816.5
glutamic acid	9481.1	10018.9	16884.3	13172.5	45585.8	41296.9	17921.5	19359.6	36340.9
glutamine	56190.5	109350.0	43056.5	104004.5	89305.7	86400.7	59116.4	45084.9	41939.7
glyceric acid	9199.7	3901.6	6327.0	10042.9	18348.2	5455.3	10201.7	10133.3	6518.6
glycerolphosphate alpha	687.9	799.0	1296.8	1655.1	2698.1	2837.9	1244.5	1392.6	1917.0
glycine	2408.4	8453.3	2411.1	6597.3	18796.6	9543.2	4097.7	4866.3	5630.0
glycolic acid	1133.2	2189.2	1488.6	982.5	2195.2	1094.1	912.2	2513.5	2581.7
hexaric acid 1	1639.9	912.4	1769.8	2018.2	1089.9	2548.4	2182.3	795.1	1876.6
hexaric acid 2	2056.7	947.2	2518.0	2021.0	1591.4	3117.2	2560.1	1091.7	3060.7
hexonic acid	319.2	342.0	265.6	326.0	458.4	336.1	275.0	376.0	212.6
hexose non-meox	407900.6	145155.2	155670.6	211828.3	58390.6	60667.6	228810.4	101205.2	49085.3
hydroxylamine	8006.5	8526.9	8345.6	9252.0	14924.0	13485.8	7410.9	9159.9	12700.2
inositol myo-	25964.2	11002.0	20419.9	48036.2	36428.7	53796.7	54745.1	26036.1	31475.7
inulobiose 1	857.9	1310.9	824.2	1188.3	1307.3	950.4	893.2	1512.3	1043.1
inulobiose 2	1137.4	1728.0	924.0	2414.5	1423.4	2281.1	788.6	1885.8	932.3
inulotriose 1	10932.2	58766.6	54337.7	40497.8	32437.5	9606.4	10862.3	45112.1	6400.1
isofucostanol	1333.1	1937.1	2614.8	2291.7	3787.1	3715.9	2431.4	2715.6	4043.7
isoleucine	1765.6	5572.8	6034.4	3387.0	5543.4	8459.0	2681.5	5043.3	5487.3
lactic acid	12111.9	10033.9	8527.7	6466.3	9604.8	6625.4	4251.3	12155.9	21987.5
leucine	1400.7	4031.8	5809.2	3600.8	4064.2	7316.5	2786.6	4877.1	5560.7
levoglucosan	3192.0	15573.8	2981.3	1816.1	25378.2	5644.7	2495.7	21589.1	4285.3
linoleic acid	186.1	313.2	299.7	355.3	652.8	513.9	321.9	484.2	481.4
linolenic acid	1249.1	2015.8	2362.8	2297.6	5857.5	4872.6	2584.3	4241.9	4763.5
lysine	814.4	1956.0	1426.1	1972.2	2331.4	3342.1	1176.4	2253.5	2367.2
lyxitol	12103.9	4512.8	5932.7	3485.1	5558.7	3657.7	4485.2	4851.4	2729.6
malate	90865.3	59403.4	67772.6	99825.5	203170.8	76565.1	143122.3	134594.7	158619.2
maleic acid	10138.2	13566.0	13044.9	16679.6	34608.6	12468.9	22659.3	39778.0	31910.1
maltose 2	290.3	534.9	268.5	406.8	613.4	418.8	368.9	693.8	499.7
mucic acid	67.3	157.7	459.8	99.5	235.6	479.2	88.9	283.3	450.1
N-acetyl-D-mannosamine	4283.2	1600.1	3243.0	1623.3	1173.9	1329.5	2432.5	1286.7	1566.8

Metabolite	PM			PFa			F1		
	time1	time2	time3	time1	time2	time3	time1	time2	time3
nicotinic acid	1409.0	1386.2	361.7	297.7	251.7	163.4	137.9	1304.7	2391.0
oxoproline	121329.9	126224.4	56489.0	219862.8	150105.8	99835.8	109509.8	155525.4	101520.6
palmitic acid	2049.2	3323.9	2624.8	3042.2	5309.6	3864.5	2859.9	4568.6	4649.5
pelargonic acid	1360.7	3021.3	1759.8	2827.4	4072.8	2010.8	1552.9	2129.8	2989.6
phenylalanine	1325.2	1950.0	3292.0	1919.2	6417.7	8167.8	2105.6	5138.7	5069.3
phosphoethanolamine	741.6	1111.3	667.8	882.6	1083.3	1158.7	646.1	925.5	871.5
phosphoric acid	23065.4	14187.2	66194.1	46948.4	50658.9	106072.5	62211.7	21892.4	72673.5
phytol	1160.0	2131.5	3116.1	2180.5	5822.6	8265.8	2530.2	4814.0	10376.4
proline	48357.2	118302.8	19651.5	20208.9	4188.2	1592.7	45697.7	3306.4	2303.4
putrescine	11929.6	14757.5	11829.9	10976.7	10507.3	6565.1	7485.2	18070.1	4801.7
quinic acid	96205.9	123974.9	176125.3	33899.3	81053.5	114722.4	35008.7	108158.1	88237.4
ribonic acid	403.1	553.2	355.9	748.6	1278.6	770.1	1201.8	948.0	490.4
serine	20215.7	26443.8	35288.7	59763.9	96539.9	105817.5	39017.0	51348.5	29241.4
shikimic acid	25158.7	52966.7	146253.4	5540.4	87514.5	159136.6	32744.7	179619.7	216080.7
sorbitol	9611.6	21900.0	7308.8	7249.4	33630.0	15075.6	10275.8	72720.0	13781.8
stearic acid	5978.9	10795.1	5858.6	8332.3	12282.3	7867.7	6592.0	11329.2	10702.1
stigmasterol	496.8	719.1	581.9	988.0	1550.4	1258.0	781.8	906.7	1157.0
succinic acid	26588.7	28045.3	25791.9	21219.5	27891.1	18540.2	33577.5	35682.3	48985.9
sucrose	111005.2	164774.6	104285.3	208627.4	169963.1	141585.8	139611.5	158347.4	167851.2
threitol	645.2	1054.9	395.0	377.5	532.1	198.1	380.5	509.3	329.3
threonic acid	2228.1	1834.2	1236.6	2971.8	6652.4	2345.7	2528.0	3678.0	2453.3
threonine	6094.1	9365.4	11439.8	12978.5	29427.0	28291.4	11940.5	19033.2	14666.6
tocopherol	111.4	750.3	350.4	238.6	937.9	832.8	181.5	517.0	774.7
trehalose	5254.8	19590.9	2700.5	10238.8	11389.8	29834.8	4396.4	15466.6	8801.1
tyramine	532.2	968.4	831.1	951.9	3369.5	5539.7	935.1	1869.6	2683.0
tyrosine	2100.7	2468.9	2743.4	4606.4	5294.2	6272.1	5227.8	3723.9	4649.5
urea	595.1	1585.2	2076.5	467.7	2860.4	1261.7	1096.5	2529.2	1741.6
valine	6738.7	17316.4	18163.1	14807.5	17683.9	28959.3	10607.7	15935.0	17920.6
xylitol	1449.5	1688.9	1590.6	486.5	676.5	285.0	423.0	502.3	567.9
xylonic acid	1134.7	1510.7	1687.6	859.6	1744.3	1087.3	1132.8	1876.9	1618.0
xylose 1	2174.1	1650.1	2338.6	1869.8	1113.1	2044.7	2211.0	1663.2	2986.1
xylose 2	2355.5	3144.8	4340.8	2350.1	1756.2	4674.2	4221.9	2446.7	4814.8
199175	594710.8	203497.0	230086.9	314276.9	89501.7	96203.8	355952.7	145705.7	78472.4
199177	1799.9	5148.6	1070.1	2171.9	7050.2	1070.7	2045.6	6792.4	1227.2
199205	4760.8	12708.9	6433.1	3905.2	5026.5	3194.8	3389.7	5259.0	2316.8
199215	2540.0	2428.4	967.5	2565.7	2396.0	1514.9	2283.8	2639.5	1056.7
199223	247.2	344.9	726.6	389.0	497.6	542.3	397.0	564.9	745.1
199231	763.7	980.9	595.3	992.9	950.1	604.8	608.4	790.9	911.7
199235	1915.4	1899.8	1727.2	2743.3	7608.8	3141.4	1965.0	4551.6	2752.8
199239	217.8	220.2	453.2	562.6	1203.6	757.8	593.6	598.9	805.6
199317	111.0	171.9	104.2	217.9	584.7	191.7	104.6	228.9	164.1
199328	1761.6	3028.8	1863.9	1903.5	2591.4	2319.0	2100.0	3158.3	3046.6
199338	11265.2	16992.0	10512.1	8854.7	22876.9	16496.7	12716.5	31908.7	20985.3
199463	714.5	680.8	383.3	1024.6	880.7	649.0	788.2	671.0	815.2
199562	680.9	1310.9	677.0	828.8	1397.9	824.4	746.8	1266.8	986.9
199777	3968.4	1920.3	2105.7	2902.5	1301.7	939.1	4642.2	1597.5	1274.9
200384	2582.8	4194.3	2131.4	646.4	509.5	468.0	586.8	2108.2	2443.6
200392	495.5	470.5	399.8	499.9	792.2	449.1	397.8	689.2	663.5
200401	3113.2	6525.1	3580.8	3202.7	16565.9	4491.9	2876.8	9559.0	3812.5
200414	2244.0	3878.5	5900.7	940.7	2445.7	1048.9	1854.3	4015.3	3475.3
200416	836.1	1487.4	644.6	685.4	1247.1	515.9	631.2	976.2	581.0
200420	3804.3	10183.1	2543.4	4675.1	10401.6	1979.0	3326.6	9361.4	1720.5
200426	4502.5	5988.7	5111.2	5306.6	10156.6	6426.2	4947.4	9483.9	7017.2
200427	1217.9	1404.2	1148.4	1341.0	1495.6	1293.0	798.9	1799.7	1683.8
200429	1049.3	1281.5	1205.2	1191.1	1233.0	1056.1	774.2	1486.6	1370.7
200448	223.9	281.2	663.8	245.4	1966.8	736.3	305.9	914.2	485.1
200450	198.4	673.8	226.6	336.8	11794.0	1146.5	162.6	2757.5	994.2
200463	250.4	304.0	863.0	799.6	1007.3	993.5	460.3	615.9	560.2
200466	515.4	617.5	457.0	583.4	779.1	612.2	445.8	754.5	813.1
200486	300.3	449.4	363.6	506.8	942.9	620.4	430.0	1057.4	1030.6
200489	546.8	567.4	755.4	622.3	805.6	785.3	821.1	847.3	849.1
200491	784.6	466.9	191.8	1029.4	1586.3	343.5	1029.7	902.9	532.6

Metabolite	PM			PFa			F1		
	time1	time2	time3	time1	time2	time3	time1	time2	time3
200511	4381.0	4164.9	3297.2	2239.1	2538.7	2681.6	2436.7	4276.4	2511.6
200513	17367.8	41298.9	42123.2	20657.6	63321.2	91469.6	23712.9	48210.1	56352.7
200518	2206.7	8645.9	7167.7	3628.1	3186.8	14327.3	937.4	2900.4	2288.8
200521	1226.4	2044.4	893.3	891.2	1907.2	883.3	882.3	2505.8	691.8
200531	587.1	727.9	766.7	573.2	1509.7	1129.5	529.2	849.9	1148.4
200532	8469.0	10459.0	13965.1	9517.2	19627.0	20821.4	9999.3	14983.3	19922.2
200540	407.5	1275.0	640.5	565.3	1036.7	734.0	372.6	831.5	676.9
200549	211.0	343.5	345.0	195.6	290.0	248.2	198.2	281.5	287.9
200556	8518.3	6044.2	6389.5	1595.9	1016.6	1521.1	2588.7	2032.9	2383.5
200567	687.8	2496.4	1438.3	1327.7	2985.4	1654.8	1013.5	3039.8	1848.5
200595	1415.2	504.9	1069.8	466.9	331.6	350.0	752.9	379.2	409.1
200624	552.6	290.7	269.9	322.7	303.9	329.0	260.0	346.6	1267.8
200844	825.6	1120.5	871.4	917.8	1188.8	811.8	632.2	1042.3	821.3
200874	1637.5	4259.9	1255.1	1180.0	3177.4	2468.2	986.8	4395.7	2382.7
200896	1002.9	12636.6	6899.9	2620.4	2019.1	3742.3	500.1	6813.9	3014.2
200900	1611.1	2150.5	920.8	959.9	486.6	261.6	1803.3	625.2	370.9
201051	569.4	678.0	584.2	953.9	952.9	2127.1	702.0	945.2	1317.3
201832	1101.3	2028.7	622.7	1716.5	1710.4	954.6	2573.5	1611.4	914.8
202083	340.1	549.9	191.2	417.1	473.5	345.9	379.5	789.7	566.4
202178	672.9	883.7	1316.3	651.9	1771.8	1828.4	601.5	1794.8	1779.8
202570	8810.6	5556.4	13817.6	12548.8	10830.2	3791.8	13431.6	34616.6	22161.9
202573	31555.8	29926.2	48411.3	9482.0	5321.5	10793.0	12131.4	5722.7	11455.3
202599	2436.1	2576.0	405.3	2774.8	4373.0	1758.6	1585.2	1582.7	970.7
202737	879.9	1555.1	423.7	776.3	1689.5	831.4	861.1	1282.6	540.5
202808	1056.1	3266.6	4571.5	940.8	1142.4	1708.7	1426.8	1961.3	1877.9
202834	451.8	693.9	320.5	699.6	1348.4	477.4	529.7	1372.5	616.0
202838	1597.4	3811.7	1172.5	2202.4	4221.3	1564.5	1625.6	3412.7	3585.1
203052	1385.8	3636.3	1610.3	2492.3	3296.4	919.6	2296.9	3621.5	999.0
203157	88875.7	116759.6	21901.5	15535.7	8407.1	8894.5	6705.5	80223.3	169593.7
203250	506.4	771.4	337.4	504.8	951.1	441.8	340.0	684.4	845.2
203264	1965.6	2635.7	1347.4	2827.3	2173.6	2447.0	2069.4	1334.7	1259.6
204344	71925.8	17360.6	137034.2	194807.8	209655.2	205798.2	156019.9	124077.3	34139.2
205664	7022.8	6613.2	16844.4	8098.2	19280.0	18711.2	9410.1	18675.8	24736.0
205673	2173.0	2943.0	1664.2	2832.0	3346.2	4015.7	2534.2	1882.9	4153.4
205680	1176.4	1739.3	1349.2	725.9	1252.6	863.6	590.5	1107.5	1166.4
205849	3117.0	6797.9	6212.4	885.5	498.2	347.5	2247.6	3835.9	2860.1
205857	330.1	625.1	823.1	933.5	643.4	636.8	558.5	723.5	631.9
206136	197.5	422.2	277.8	893.4	826.9	1280.8	347.7	340.0	283.1
206318	279.4	463.1	226.5	355.8	468.1	767.0	319.2	460.6	436.8
206528	566.1	1871.4	430.0	941.4	5540.6	3879.0	1808.7	9022.0	6694.6
207326	849.0	1214.1	976.7	1358.5	1710.9	1379.7	919.9	1144.8	1433.8
207432	577.8	2379.1	1070.9	719.0	2115.1	1033.4	677.1	2273.2	1187.8
207507	24436.0	92928.8	27226.9	16613.5	321174.9	26749.2	12819.3	20836.8	93251.8
207509	461.4	1392.1	473.4	402.1	1216.5	852.6	492.5	1539.2	827.7
207750	3560.5	9819.1	6151.6	2126.5	10097.2	10153.3	3038.1	14873.8	12376.1
208651	395.1	1610.4	495.8	398.6	2654.2	502.6	367.0	1406.9	457.0
208658	698.4	1438.6	739.4	1056.3	1545.2	750.1	670.9	3236.2	893.0
208662	240.3	319.4	236.5	341.9	2203.2	1155.5	195.1	632.1	326.9
208664	2527.4	2759.6	1919.3	4753.1	2525.1	1307.2	8529.7	1782.0	376.3
208701	195.9	277.4	187.4	203.5	506.1	387.4	158.1	346.4	316.9
208770	821.9	1562.0	683.7	957.9	1258.5	1101.8	923.9	1644.9	999.4
208840	116.0	315.6	169.8	149.0	1119.9	248.5	102.2	417.1	379.8
208841	347.6	229.8	932.0	540.0	874.9	1375.8	721.6	506.0	1450.5
208845	383.3	423.8	693.5	285.7	1151.1	5251.9	268.4	439.0	586.6
208850	3103.0	4473.9	1758.9	1844.4	1124.3	586.6	3908.3	876.5	347.2
208874	1836.6	2542.1	1607.1	2604.7	2206.0	1137.7	4509.7	2653.4	1692.7
208897	524.4	667.0	828.3	468.9	1345.0	1141.6	494.3	1144.3	906.8
210399	572.0	1364.8	573.6	811.5	1564.7	614.5	521.2	1512.7	935.1
210512	17356.5	171049.9	20059.6	13338.4	179777.7	102328.3	2826.0	248939.3	102253.3
210882	111224.7	98098.9	337266.1	411795.6	744496.1	551618.2	426910.3	211677.4	145647.3
210891	9407.6	8938.9	6068.8	9027.3	7544.8	6028.2	5252.0	11243.0	14171.4
210893	6080.8	5671.3	4417.3	3852.4	3553.0	3334.9	2336.3	6734.0	8172.8

Metabolite	PM			PFa			F1		
	time1	time2	time3	time1	time2	time3	time1	time2	time3
210894	3940.7	5746.8	909.7	485.1	58.2	50.5	49.1	3166.4	9742.1
210896	3813.9	4965.3	825.5	448.4	150.6	229.2	98.4	3135.3	6849.9
210901	5538.3	5913.0	1139.1	506.8	133.1	93.5	76.4	4877.4	10831.2
210904	809.2	806.3	277.4	218.5	257.9	173.3	138.2	720.3	1488.1
210909	2961.9	3079.6	691.7	452.4	355.1	250.0	179.1	2192.1	4578.5
210912	5490.7	6467.1	3062.6	3960.2	3507.8	3751.2	2882.4	4129.8	7526.1
211886	379.3	357.9	890.6	231.2	551.7	234.5	206.6	1719.6	1747.3
211890	2770.8	2290.5	993.9	934.5	301.7	165.0	130.4	6041.7	7537.3
211891	310.4	438.7	207.7	431.9	511.1	261.9	420.4	442.4	325.0
211894	926.0	811.7	448.5	500.4	1155.7	368.4	354.4	3625.7	5063.9
211896	11749.4	11535.3	8232.7	11296.4	15526.1	8537.6	7780.2	10150.9	11061.8
211898	1373.1	1651.7	341.3	252.8	138.9	155.4	81.7	1161.4	2433.0
211911	1921.9	1959.4	933.2	1157.2	1008.7	700.7	638.3	2247.2	4432.2
211914	2170.3	2199.4	610.4	297.9	229.7	194.1	123.1	1898.8	4019.1
211916	628.4	624.1	973.7	1219.2	1446.0	1281.3	1074.3	1064.8	1029.6
211917	2150.9	2387.5	1434.5	1562.2	1713.9	1529.5	1199.6	2611.3	4032.6
211919	63.4	137.0	152.9	111.5	183.2	330.5	67.2	318.8	611.7
211934	1916.2	1872.9	722.7	488.1	234.8	189.1	137.3	3715.5	6862.2
211935	884.0	810.6	487.9	499.8	625.6	457.2	239.8	1377.7	1797.2
211941	1290.7	1485.7	2266.6	1886.8	5915.2	4127.7	1243.0	2990.8	3023.0
211946	2845.2	2986.5	2271.5	2393.6	3674.9	2650.7	2127.1	3141.5	4696.4
211962	250.9	244.7	410.3	404.9	510.6	715.4	425.2	353.0	514.8
211972	6391.6	6820.6	2599.5	4420.8	4376.3	1372.8	7601.3	2726.8	1882.2
211980	1651.5	2066.3	603.6	411.4	655.3	382.4	229.4	2412.6	3257.3
212024	3819.9	4232.4	1299.4	622.7	431.9	208.1	161.1	4283.4	8505.3
212189	477.6	568.4	150.1	129.2	96.5	103.3	60.5	440.2	813.7
212208	1115.4	1096.9	1134.0	1536.2	2100.0	1185.5	3568.3	1616.4	1424.7
212274	8330.9	1325.4	470.8	1411.1	406.5	590.5	842.9	642.3	247.8
212663	465.6	494.3	403.3	728.1	1904.8	349.7	760.5	1054.4	294.6
212679	1827.8	1679.9	1606.2	2427.0	7074.2	6540.9	5429.8	7603.5	5948.3
212732	1670.8	1551.0	638.0	1250.1	329.0	255.6	1443.0	584.4	346.0
212735	659.5	1027.7	591.8	678.1	884.4	565.5	589.9	1574.0	735.0
212781	639.1	1480.5	793.8	792.0	1350.2	1166.6	486.1	1515.0	1093.0
213143	406.3	266.1	106.8	545.4	223.4	70.8	364.6	125.0	125.2
213155	3424.9	4946.0	1473.6	1609.5	1272.7	1141.2	944.9	3518.5	7061.0
213182	490.6	184.7	695.6	492.0	241.4	875.0	1030.5	929.2	2439.6
213185	886.7	5000.8	4819.7	312.9	1251.2	2710.0	540.1	4719.9	3405.3
213194	115.1	115.9	291.5	133.0	113.4	643.8	144.3	181.9	510.7
213271	638.9	808.5	509.1	603.4	787.3	696.8	489.4	818.8	894.2
213310	239.0	250.8	306.9	302.6	400.4	366.8	337.6	361.7	445.5
213714	151.4	170.2	97.3	249.2	250.1	207.8	194.3	226.9	269.1
213732	338.6	693.4	415.5	960.2	1043.7	985.5	598.3	703.3	842.3
214401	540.8	797.6	489.6	802.4	874.3	739.1	592.2	1337.4	852.9
214405	1103.8	2276.2	1107.8	2062.3	2416.0	2006.6	1714.8	1893.7	3973.2
214410	1223.1	756.6	1005.9	1189.6	585.5	516.0	1389.7	531.9	944.2
214434	642.2	930.3	508.8	766.3	1017.7	1192.6	633.4	854.1	1150.9
214680	620.8	691.8	441.2	907.2	851.1	595.2	656.6	673.0	761.8
215062	1389.3	1836.3	1714.3	5137.0	3459.0	1165.8	3159.2	2782.4	1949.3
215344	7004.6	8173.7	1630.4	830.6	171.5	164.8	101.7	6751.4	14671.7
215347	5323.6	4884.4	1254.8	987.7	168.2	224.4	90.4	5973.5	9144.7
215355	2608.7	1893.6	9620.0	12569.8	21406.9	15295.4	12201.2	5493.1	2744.0
215362	2153.3	2227.0	502.7	292.9	107.5	89.0	75.4	1793.8	3891.3
215375	1533.5	2426.8	228.8	139.7	94.3	66.1	102.8	133.4	204.3
215399	1155.6	1229.6	303.2	220.1	157.0	91.9	79.2	968.5	2034.1
215402	1677.7	1612.2	377.5	220.6	80.3	78.4	52.2	1458.3	2962.5
215448	9881.8	10960.4	10855.5	10803.0	14040.7	11731.8	7528.7	10110.3	15076.0
215466	775.7	911.7	903.8	1130.8	1454.0	1236.8	973.5	1130.2	1405.2
215490	1995.4	1927.0	2015.9	4334.7	3386.7	2954.3	2512.0	2469.3	2616.5
215492	463.5	431.3	354.4	491.5	472.4	422.2	312.9	541.5	794.1
215493	1442.3	1169.4	500.5	470.3	519.4	294.5	266.3	2654.4	3412.8
215504	1452.8	992.2	355.4	237.0	232.8	202.5	148.7	702.9	3148.9
215529	638.8	742.2	240.1	191.6	139.6	107.7	82.6	807.2	1794.1

Metabolite	PM			PFa			F1		
	time1	time2	time3	time1	time2	time3	time1	time2	time3
215555	5824.6	3620.0	1359.9	8924.9	883.2	922.7	640.6	5496.7	4992.3
215563	975.0	880.2	245.2	230.4	134.4	77.7	91.7	811.7	1605.4
215643	3710.5	4412.3	793.3	425.4	99.5	80.0	53.5	3233.5	6779.7
215682	1026.6	1299.1	1122.8	1116.8	1247.6	1214.8	766.0	1934.0	1794.6
215860	862.8	935.9	1282.2	374.2	344.7	291.1	684.5	494.0	488.4
215978	746.6	1141.8	186.5	103.3	85.2	65.5	44.0	600.5	1782.8
216098	731.5	785.4	197.7	161.5	130.2	117.6	80.3	683.3	1345.2
216424	13061.7	8492.2	2832.8	2077.1	1772.4	799.5	587.1	9431.0	17342.8
216427	3245.1	3937.0	4273.9	4149.7	7537.4	6175.1	4123.4	6060.6	6257.1
216454	3036.7	2967.7	885.1	866.0	281.5	253.7	328.9	3877.4	5980.4
216472	339.5	360.8	227.7	191.2	82.8	278.2	365.5	453.6	1304.0
216493	663.5	874.4	139.8	162.8	106.3	77.0	53.9	533.4	1327.8
216564	6330.1	10816.4	5915.7	7675.1	9374.0	10838.5	4064.1	6107.8	8377.6

Appendix Table 2: Metabolite data of an ANOVA between F_1/paternal, F_1/maternal, and paternal/maternal and the harvest times (HT 1: 12.06.06, HT 2: 15.08.06, HT 3: 19.10.06). (in **bold**: $p<0.05$)

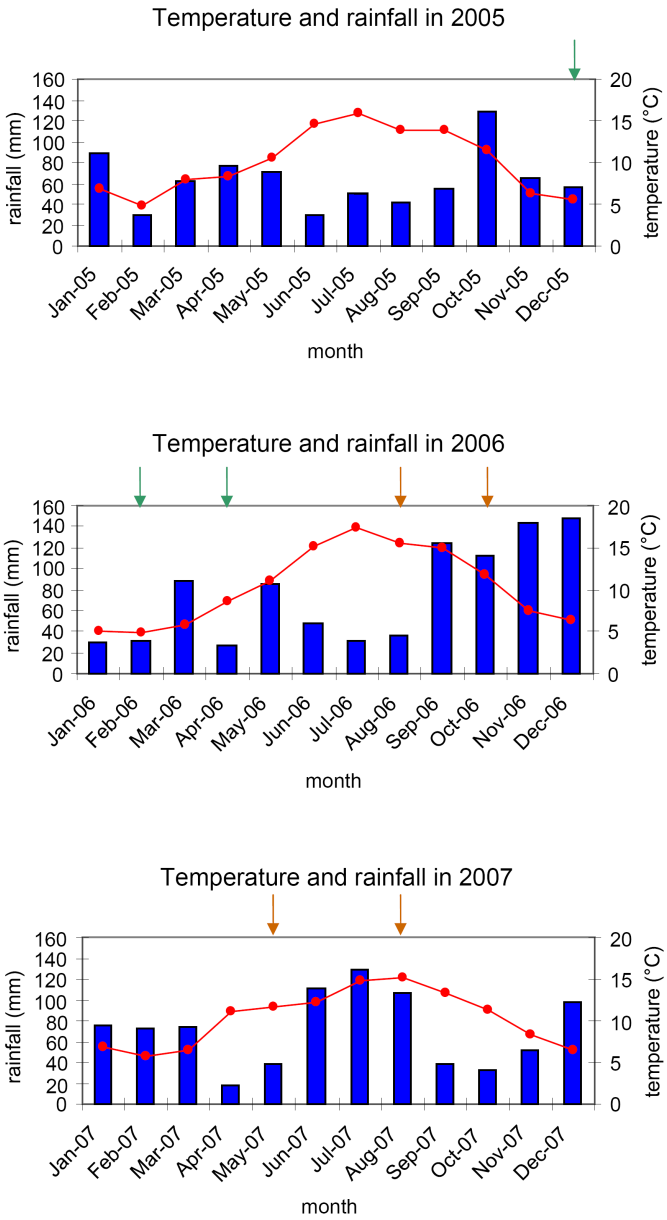
Metabolite	PM			PFa			F1		
	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3
alanine	**0.0158**	**0.0009**	**0.0334**	0.1547	0.5351	0.3683	0.0568	**0.0122**	0.4065
alpha_keto-glutaric_acid	0.3416	0.3164	0.4144	**0.0341**	0.3837	0.7905	0.2159	0.3481	0.8187
arabitol	**0.0020**	0.1635	**0.0027**	**0.0020**	0.0780	**0.0343**	**<0.0001**	**0.0300**	**<0.0001**
asparagine	**0.0001**	0.1506	0.7360	**0.0011**	**0.0003**	**0.0010**	**0.0005**	**0.0001**	0.6544
aspartic_acid	0.6371	**0.0039**	**0.0014**	**0.0001**	**0.0020**	0.0705	**0.0021**	**0.0003**	0.3534
benzoic_acid	0.2908	0.3186	0.9722	0.1827	0.1669	0.8972	0.4762	**<0.0001**	**<0.0001**
beta_alanine	**0.0396**	0.1837	**0.0043**	0.0610	**0.0005**	0.2314	**0.0079**	**<0.0001**	**<0.0001**
caffeic_acid	0.1350	0.8519	0.1966	**0.0114**	**0.0236**	0.9534	**<0.0001**	**0.0003**	0.3492
citric_acid	0.1364	**0.0194**	0.1502	**<0.0001**	0.0553	**<0.0001**	**<0.0001**	**<0.0001**	**0.0004**
cyano_L_ alanine	**0.0001**	0.1196	0.9524	**0.0014**	**<0.0001**	**0.0001**	**0.0001**	**0.0002**	0.9748
erythritol	**0.0021**	**0.0381**	**0.0004**	0.8299	0.3725	**0.0002**	**0.0022**	0.9734	**0.0005**
erythronic_acid_lactone	0.4587	**0.0185**	0.2850	**0.0206**	0.2980	0.1071	**<0.0001**	**<0.0001**	0.0880
fructose_1	0.2814	**0.0102**	0.0736	0.1861	0.7973	0.0821	**<0.0001**	0.7113	**<0.0001**
fructose_2	0.3370	**0.0077**	**0.0495**	0.2347	0.9468	0.1313	**<0.0001**	0.4299	**<0.0001**
fructose_6_ phosphate	0.4503	**0.0010**	**0.0011**	0.0539	**<0.0001**	**0.0125**	0.1448	**0.0092**	**<0.0001**
fucose_1_ rhamnose_2	**0.0008**	**0.0489**	0.3885	**<0.0001**	0.0917	**0.0006**	**<0.0001**	**<0.0001**	**0.0034**
fumarate	0.0888	**0.0147**	0.7331	**<0.0001**	**0.0264**	**0.0106**	**<0.0001**	**<0.0001**	**0.0236**
GABA	**0.0036**	0.2412	0.2391	**0.0002**	0.3259	**0.0008**	**0.0001**	0.1286	**<0.0001**
galactinol	**0.0086**	**0.0012**	**0.0001**	**0.0026**	0.0974	**0.0001**	0.8680	**<0.0001**	**<0.0001**
Gluco-heptulose	**<0.0001**	0.1225	**<0.0001**	**0.0001**	0.5080	**<0.0001**	**<0.0001**	0.1098	**<0.0001**
glucose_1	**<0.0001**	**<0.0001**	0.6065	**<0.0001**	**<0.0001**	0.4732	**<0.0001**	**<0.0001**	**0.0180**
glucose_1_ phosphate	**0.0025**	**0.0417**	0.3020	**<0.0001**	**0.0003**	0.0899	**<0.0001**	**0.0011**	0.8635
glucose_2	**<0.0001**	**<0.0001**	0.8857	**<0.0001**	**<0.0001**	0.3859	**<0.0001**	**<0.0001**	**0.0164**
glucose_6_ phosphate_2	0.8437	**0.0006**	**0.0002**	0.1804	**0.0006**	**0.0060**	**0.0253**	**0.0002**	**<0.0001**
glutamic_acid	0.8076	**0.0064**	**0.0141**	**<0.0001**	**<0.0001**	0.4981	0.5729	**<0.0001**	**<0.0001**
glutamine	**0.0256**	0.4314	**0.0174**	0.6251	0.5912	0.8732	0.2369	0.1022	0.6936
glyceric_acid	**<0.0001**	**0.0047**	**0.0014**	**<0.0001**	**0.0009**	**<0.0001**	0.9368	**<0.0001**	**<0.0001**
Glycerolphos-phate_alpha	0.4427	**0.0007**	**0.0111**	**0.0002**	**0.0041**	0.7403	0.3823	**0.0011**	**0.0089**
glycine	**0.0149**	0.9960	**0.0303**	**0.0002**	0.0566	**0.0069**	**0.0285**	**0.0003**	**0.0386**
glycolic_acid	**0.0001**	0.0602	**0.0172**	**0.0001**	0.6705	**<0.0001**	**<0.0001**	**<0.0001**	0.8029
hexaric_acid_ 1	**0.0051**	0.7707	**0.0259**	**0.0285**	0.2271	**0.0020**	**<0.0001**	0.4176	**0.0007**
hexaric_acid_ 2	**<0.0001**	0.2682	**0.0002**	0.0575	**0.0162**	**0.0028**	**<0.0001**	0.1422	**<0.0001**

Metabolite	PM			PFa			F1		
	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3
hexonic_acid	0.4248	**0.0408**	**0.0163**	**0.0315**	0.8172	**0.0493**	**0.0019**	**0.0021**	**<0.0001**
hexose_non_meox	**<0.0001**	**<0.0001**	0.7643	**0.0091**	**0.0134**	0.8474	**0.0002**	**<0.0001**	**0.0003**
hydroxylamine	0.7001	0.7936	0.8798	0.1088	0.1974	0.6203	0.2005	**0.0009**	**0.0454**
inositol_myo	**<0.0001**	**0.0303**	**0.0040**	0.0513	0.4193	**0.0011**	**<0.0001**	**0.0001**	**0.0003**
inulobiose_1	**0.0148**	0.8650	**0.0011**	0.5860	0.3236	**0.0448**	**0.0320**	0.3632	**0.0466**
inulobiose_2	0.1728	0.1220	0.1021	**0.0264**	0.7879	**0.0036**	**0.0016**	0.5864	**<0.0001**
inulotriose_1	**<0.0001**	0.1114	0.8629	0.2966	**0.0007**	**<0.0001**	**<0.0001**	0.0844	**<0.0001**
isofucostanol	**0.0037**	**0.0001**	**0.0193**	**0.0003**	**0.0001**	0.8658	0.1956	**<0.0001**	**<0.0001**
isoleucine	**0.0008**	**0.0001**	0.7212	0.0548	**0.0041**	0.1253	**0.0088**	**0.0036**	0.6886
lactic_acid	0.2994	0.1144	0.5092	0.0933	0.9209	0.0656	**<0.0001**	**0.0001**	**0.0215**
leucine	**0.0033**	**0.0001**	0.1280	0.7406	**0.0168**	0.1088	**0.0463**	**0.0160**	0.6018
levoglucosan	**0.0082**	0.8400	**0.0162**	**<0.0001**	**0.0006**	**0.0006**	**0.0002**	**0.0128**	**0.0002**
linoleic_acid	0.0990	**0.0435**	0.8700	**0.0078**	0.0912	0.1757	**0.0344**	**0.0126**	0.9684
linolenic_acid	**0.0192**	**0.0010**	0.3544	**<0.0001**	**0.0015**	0.2296	**0.0025**	**0.0001**	0.3694
lysine	**<0.0001**	**0.0034**	0.0571	0.2268	**0.0213**	0.0670	**0.0030**	**<0.0001**	0.7383
lyxitol	**<0.0001**	**0.0001**	0.0903	**0.0014**	0.7542	**0.0001**	0.5340	**0.0029**	**<0.0001**
malate	**0.0029**	**0.0263**	0.3900	**<0.0001**	**0.0485**	**<0.0001**	0.4748	0.1826	**0.0063**
maleic_acid	0.1260	0.2364	0.8328	**0.0079**	0.1803	**0.0029**	**0.0058**	0.1433	0.2838
maltose_2	**0.0007**	0.5834	**0.0010**	**0.0257**	0.8335	**0.0451**	**<0.0001**	0.0543	**0.0163**
mucic_acid	**0.0054**	**0.0001**	**0.0007**	**0.0057**	**0.0008**	**0.0200**	**0.0005**	**0.0003**	0.0683
N_acetyl_D_mannosamine	**0.0004**	0.2789	**0.0006**	**0.0316**	0.2307	0.3801	**<0.0001**	**0.0009**	**0.0195**
nicotinic_acid	0.9237	**0.0003**	**0.0002**	0.6324	0.1735	0.0564	**0.0004**	**<0.0001**	**0.0050**
oxoproline	0.8512	**0.0137**	**0.0197**	**0.0491**	**0.0001**	0.1173	**0.0208**	0.6609	**0.0048**
palmitic_acid	**0.0003**	**0.0079**	0.0574	**0.0001**	0.1051	**0.0059**	**0.0001**	**0.0001**	0.8558
pelargonic_acid	**0.0008**	0.0531	**0.0182**	0.1641	0.2909	**0.0013**	0.0708	**0.0002**	**0.0260**
phenylalanine	**0.0079**	**0.0005**	**0.0113**	**<0.0001**	**<0.0001**	**0.0367**	**<0.0001**	**<0.0001**	0.8780
phosphoethanolamine	**0.0099**	0.4852	**0.0069**	0.3582	0.1906	0.7872	0.0925	0.1021	0.7522
phosphoric_acid	0.1249	**0.0068**	**0.0002**	0.7066	**0.0003**	**0.0020**	**0.0324**	0.5825	**<0.0001**
phytol	**0.0034**	**<0.0001**	**0.0275**	**<0.0001**	**0.0003**	0.1182	**0.0001**	**<0.0001**	**<0.0001**
proline	**0.0131**	0.0786	**0.0018**	**0.0043**	**0.0016**	0.0798	**0.0006**	**0.0005**	0.4544
putrescine	0.2044	0.9588	0.2674	0.8019	**0.0296**	**0.0295**	**<0.0001**	**0.0016**	**<0.0001**
quinic_acid	**0.0230**	**<0.0001**	**0.0011**	**0.0115**	**0.0013**	0.1040	**<0.0001**	**0.0004**	0.0824
ribonic_acid	**0.0115**	0.1503	**0.0058**	**0.0185**	0.9183	**<0.0001**	0.2186	**0.0011**	**<0.0001**
serine	0.2349	**0.0198**	0.1459	**0.0023**	**<0.0001**	0.4176	**0.0201**	**0.0444**	**0.0001**
shikimic_acid	0.1312	**0.0003**	**0.0046**	**<0.0001**	**<0.0001**	**0.0121**	**<0.0001**	**<0.0001**	**0.0236**
sorbitol	**0.0004**	**0.0387**	**0.0003**	**0.0002**	**0.0417**	**0.0092**	**<0.0001**	0.1136	**<0.0001**
stearic_acid	**0.0001**	0.8464	**0.0001**	**0.0219**	0.7379	**0.0108**	**<0.0001**	**0.0001**	0.5397
stigmasterol	**0.0067**	0.0989	0.1133	**0.0001**	**0.0426**	0.0748	**0.0439**	**0.0001**	**0.0055**
succinic_acid	0.7161	0.7593	0.5936	**0.0094**	0.3016	**0.0001**	0.5520	**0.0006**	**0.0008**
sucrose	0.1048	0.7468	0.0878	0.4881	0.2048	0.4319	0.6371	0.4450	0.7112
threitol	**0.0031**	**0.0296**	**0.0001**	0.1629	0.0882	**<0.0001**	**0.0150**	0.1952	**0.0009**
threonic_acid	0.2138	**0.0018**	**0.0400**	**<0.0001**	**0.0252**	**<0.0001**	**<0.0001**	0.6600	**<0.0001**
threonine	**0.0036**	**0.0008**	0.1287	**<0.0001**	**<0.0001**	0.6836	**<0.0001**	0.1030	**0.0121**
tocopherol	**<0.0001**	**0.0057**	**0.0112**	**0.0003**	**0.0077**	0.6522	**0.0023**	**<0.0001**	**0.0199**
trehalose	**0.0031**	0.0589	**0.0021**	0.7271	0.3054	0.3275	**<0.0001**	**0.0020**	**0.0034**
tyramine	**0.0016**	**0.0101**	0.3647	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0004**
tyrosine	0.4982	0.2957	0.4307	0.4412	0.1576	0.3289	0.0907	0.4742	0.1657
urea	**0.0181**	0.0975	0.5832	**0.0008**	0.1425	0.0629	**0.0173**	0.0809	0.1186
valine	**0.0003**	**0.0001**	0.7985	0.3593	**0.0105**	**0.0498**	**0.0400**	**0.0099**	0.5169
xylitol	0.2256	0.4093	0.6260	0.5832	0.3446	0.1952	0.3768	0.1572	0.3602
xylonic_acid	**0.0083**	**<0.0001**	0.2892	**0.0002**	0.2354	**0.0001**	**<0.0001**	**<0.0001**	**0.0299**
xylose_1	0.1102	0.6935	0.1147	**0.0032**	0.7050	**0.0402**	**0.0107**	0.0567	**0.0009**
xylose_2	0.3020	**0.0014**	0.1997	0.2812	**0.0126**	**0.0019**	**0.0003**	0.3292	**<0.0001**
199175	**<0.0001**	**<0.0001**	0.5999	**0.0080**	**0.0132**	0.7294	**0.0001**	**<0.0001**	**0.0021**
199177	**0.0002**	0.0992	**0.0003**	**<0.0001**	**0.0192**	**<0.0001**	**<0.0001**	**0.0001**	**<0.0001**
199205	**<0.0001**	0.1735	**0.0007**	0.2664	0.4785	**0.0054**	**0.0011**	**0.0005**	**<0.0001**
199215	0.8034	**<0.0001**	**0.0072**	0.6007	**0.0021**	**0.0006**	0.1184	**<0.0001**	**<0.0001**

Metabolite	PM			PFa			F1		
	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3
199223	0.1403	**<0.0001**	**0.0003**	0.1475	0.1328	0.6488	**0.0417**	**0.0008**	**0.0183**
199231	0.2318	0.1829	**0.0447**	0.7457	**0.0040**	**0.0142**	0.2156	**0.0089**	0.4135
199235	0.9569	0.6265	0.6889	**<0.0001**	0.6453	**0.0001**	**<0.0001**	**0.0261**	**0.0002**
199239	0.9498	**0.0021**	**0.0029**	**<0.0001**	**0.0380**	**0.0003**	0.9226	**0.0004**	**0.0001**
199317	**0.0122**	0.6800	**0.0140**	**<0.0001**	0.5759	**0.0001**	**<0.0001**	**0.0232**	**0.0299**
199328	**0.0056**	0.8302	**0.0382**	0.1270	0.4792	0.5851	**0.0121**	**0.0459**	0.7883
199338	**0.0001**	0.5436	**0.0005**	**<0.0001**	**0.0037**	**0.0134**	**<0.0001**	**<0.0001**	**<0.0001**
199463	0.7387	**0.0003**	**0.0069**	0.3387	**0.0009**	0.1321	0.1746	0.8026	0.1874
199562	**0.0155**	0.9863	0.0560	0.1439	0.9800	0.1532	**0.0413**	**0.0373**	0.2388
199777	**<0.0001**	**0.0001**	0.5735	**0.0007**	**0.0001**	**0.0198**	**0.0001**	**<0.0001**	0.5424
200384	0.0770	0.5648	0.0935	0.6440	0.5521	0.8635	**0.0009**	**<0.0001**	0.4727
200392	0.6247	0.0612	0.1835	**0.0061**	0.4926	**0.0014**	**0.0001**	**0.0013**	0.7621
200401	**0.0001**	0.6187	**0.0166**	**<0.0001**	0.0554	**<0.0001**	**<0.0001**	**0.0027**	**<0.0001**
200414	**0.0005**	**0.0109**	0.1178	**0.0085**	0.8381	**<0.0001**	**0.0006**	**0.0040**	0.0537
200416	**0.0275**	0.2804	**0.0145**	**0.0199**	0.3662	**0.0021**	**0.0083**	0.6924	**0.0023**
200420	**<0.0001**	0.1280	**<0.0001**	**<0.0001**	**0.0109**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
200426	**0.0043**	0.1245	0.1344	**<0.0001**	0.1682	**0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
200427	0.3738	0.6914	0.1829	0.7628	0.9187	0.5672	**<0.0001**	**0.0039**	0.7080
200429	0.2112	0.2977	0.6788	0.9186	0.7227	0.5054	**0.0014**	**0.0287**	0.6928
200448	0.2253	**0.0027**	**0.0048**	**0.0079**	**0.0112**	0.0685	**0.0122**	**0.0253**	**0.0484**
200450	**<0.0001**	0.6433	**0.0002**	**<0.0001**	**0.0003**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0004**
200463	0.2279	**0.0005**	**0.0006**	0.1409	0.4076	0.9536	**0.0478**	0.1936	0.4600
200466	0.2874	0.3745	0.0972	0.2555	0.8513	0.1990	**0.0023**	**0.0069**	0.6662
200486	**0.0077**	0.1373	0.1474	**0.0001**	0.3164	**0.0128**	**<0.0001**	**<0.0001**	0.8171
200489	0.7382	**0.0208**	**0.0241**	0.0723	0.1151	0.8137	0.7237	0.7106	0.9757
200491	**0.0081**	**0.0004**	**0.0001**	**0.0279**	**0.0011**	**<0.0001**	0.5594	**0.0186**	**0.0001**
200511	0.6736	0.0834	0.1449	0.6210	0.4956	0.7798	**0.0016**	0.8669	**0.0025**
200513	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0153**	**<0.0001**	**<0.0001**	0.1072
200518	**0.0007**	0.1200	0.6601	0.6420	**0.0027**	**0.0016**	**<0.0001**	**0.0001**	0.1163
200521	**0.0004**	0.2180	**0.0004**	**0.0149**	0.9800	**0.0034**	**<0.0001**	0.1403	**<0.0001**
200531	0.4020	0.2156	0.8052	**0.0004**	**0.0080**	0.1280	**0.0424**	**<0.0001**	0.0599
200532	0.3388	**0.0063**	0.1405	**0.0059**	**0.0001**	0.7681	**0.0180**	**0.0003**	0.0535
200540	**0.0335**	0.3005	0.1574	0.2861	0.6773	0.5631	**0.0447**	0.1000	0.5555
200549	**0.0006**	**0.0022**	0.9728	**0.0103**	0.1070	0.2744	0.0507	0.0545	0.8562
200556	0.1588	0.3456	0.8467	0.3603	0.9157	0.1699	0.2520	0.7324	0.5541
200567	**<0.0001**	**0.0027**	**0.0080**	**0.0003**	0.3298	**0.0007**	**<0.0001**	**0.0010**	**<0.0001**
200595	**0.0007**	0.3044	**0.0007**	0.0956	0.1944	0.8134	**<0.0001**	**0.0002**	0.5070
200624	**0.0326**	0.0627	0.6736	0.6928	0.9070	0.6515	0.1667	**<0.0001**	**<0.0001**
200844	0.0708	0.8107	0.1202	0.1533	0.5501	0.0702	**0.0005**	0.1276	0.0975
200874	**<0.0001**	0.1560	**<0.0001**	**0.0021**	**0.0413**	0.2378	**<0.0001**	**<0.0001**	**<0.0001**
200896	**<0.0001**	**0.0001**	**0.0291**	0.7829	0.6288	**0.0147**	**<0.0001**	**<0.0001**	**0.0001**
200900	0.0996	0.1961	**0.0210**	0.0835	**0.0197**	0.0991	**0.0002**	**<0.0001**	**0.0065**
201051	0.3186	0.8633	0.4446	0.9952	**<0.0001**	**<0.0001**	**0.0341**	**<0.0001**	**0.0146**
201832	**0.0004**	0.0867	**<0.0001**	0.9891	0.1254	**0.0059**	**0.0015**	**<0.0001**	**0.0006**
202083	**0.0421**	0.0943	**0.0032**	0.6562	0.5035	0.3585	**0.0294**	**0.0106**	0.2159
202178	0.2287	**0.0045**	0.0746	**0.0057**	**0.0008**	0.8894	**<0.0001**	**<0.0001**	0.9256
202570	0.1796	0.2581	**0.0306**	0.8215	0.2494	**0.0397**	**0.0028**	0.2014	0.1291
202573	0.8259	0.0565	**0.0261**	0.4158	0.8098	**0.0066**	**0.0008**	0.8134	**0.0191**
202599	0.9024	**0.0220**	**0.0485**	0.0780	0.3014	**0.0203**	0.9967	0.2725	0.2236
202737	0.3783	**<0.0001**	0.1967	**0.0012**	0.8324	**0.0138**	0.0922	**0.0149**	**0.0015**
202808	**<0.0001**	**<0.0001**	0.0805	0.5510	0.0704	**0.0280**	0.0746	0.1208	0.6511
202834	0.0920	**0.0038**	**0.0243**	**<0.0001**	**0.0118**	**<0.0001**	**<0.0001**	0.2112	**<0.0001**
202838	**0.0052**	0.3510	**0.0023**	**0.0217**	0.2037	**0.0047**	**0.0001**	**0.0002**	0.7382
203052	**0.0005**	0.6141	**0.0063**	0.2995	**0.0144**	**0.0004**	0.0959	**0.0003**	**0.0004**
203157	0.2014	**0.0034**	**0.0001**	0.2856	0.3405	0.7342	**0.0004**	**<0.0001**	**0.0010**
203250	**0.0034**	**0.0152**	**<0.0001**	**0.0202**	0.5583	**0.0068**	**<0.0001**	**<0.0001**	0.1033
203264	0.2538	0.2212	0.0555	0.3739	0.6222	0.6290	0.1214	0.0702	0.7888
204344	0.2040	0.2670	**<0.0001**	0.8013	0.8482	0.9288	0.4162	**<0.0001**	**0.0071**
205664	0.7865	**<0.0001**	**0.0002**	**<0.0001**	**0.0004**	0.8345	**<0.0001**	**<0.0001**	**<0.0001**
205673	0.3734	0.4672	0.1467	0.6420	0.3549	0.6424	0.2926	0.1973	0.0532
205680	**0.0029**	0.4791	0.0992	**0.0011**	0.3480	**0.0218**	**<0.0001**	**<0.0001**	0.6151

Metabolite	PM			PFa			F1		
	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3
205849	**0.0005**	**0.0001**	0.6304	0.5195	0.3936	0.1809	**0.0014**	0.1538	**0.0168**
205857	0.0970	**0.0006**	0.3474	0.2959	0.2186	0.9803	0.4758	0.6807	0.6897
206136	0.0745	0.1494	0.3046	0.7156	0.1001	0.0651	0.9246	0.2901	0.4127
206318	0.1959	0.6217	0.0957	0.3904	0.2980	0.4532	0.3233	0.3857	0.8557
206528	0.3354	0.0511	0.3479	**<0.0001**	**<0.0001**	**0.0011**	**<0.0001**	**<0.0001**	**0.0020**
207326	0.0586	0.5013	0.2444	0.3885	0.9439	0.4083	0.2622	**0.0439**	0.2447
207432	**<0.0001**	**0.0169**	**0.0001**	**0.0001**	0.3001	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
207507	**0.0092**	0.8216	**0.0358**	**<0.0001**	0.3234	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
207509	**<0.0001**	0.8895	**0.0002**	**0.0003**	**0.0012**	0.0776	**<0.0001**	**0.0001**	**<0.0001**
207750	**<0.0001**	**0.0048**	**0.0227**	**<0.0001**	**<0.0001**	0.9735	**<0.0001**	**<0.0001**	**0.0408**
208651	**<0.0001**	0.5085	**<0.0001**	**<0.0001**	0.5918	**<0.0001**	**0.0001**	0.2497	**0.0001**
208658	**0.0019**	0.8399	**0.0144**	**0.0137**	**0.0227**	**0.0006**	**<0.0001**	**0.0242**	**<0.0001**
208662	0.1613	0.9531	0.0779	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0046**	**0.0001**
208664	0.8253	0.6975	0.3261	**0.0091**	**0.0007**	**0.0374**	**<0.0001**	**<0.0001**	**0.0019**
208701	**0.0154**	0.6982	**0.0156**	**<0.0001**	**0.0013**	**0.0164**	**0.0002**	**0.0002**	0.5387
208770	**0.0001**	0.2805	**<0.0001**	**0.0452**	0.4444	0.3943	**0.0003**	0.6124	**0.0003**
208840	**0.0003**	**0.0004**	**0.0135**	**<0.0001**	**0.0403**	**<0.0001**	**0.0004**	**<0.0001**	0.6863
208841	**0.0330**	**0.0007**	**<0.0001**	0.0698	**0.0004**	0.0505	0.0630	**<0.0001**	**<0.0001**
208845	0.5834	**0.0175**	**0.0207**	**0.0485**	**0.0132**	**0.0410**	0.1530	**0.0066**	0.3100
208850	0.0845	0.1776	**0.0119**	0.1408	**0.0322**	0.2123	**<0.0001**	**<0.0001**	**0.0035**
208874	0.1475	0.6978	0.1719	0.6872	0.1232	0.0713	**0.0418**	**0.0032**	**0.0201**
208897	**0.0275**	**0.0013**	0.0644	**<0.0001**	**0.0009**	0.2897	**<0.0001**	**<0.0001**	**0.0017**
210399	**0.0188**	0.9902	**0.0347**	**0.0094**	0.4072	**0.0024**	**<0.0001**	0.0684	**0.0101**
210512	**0.0002**	0.8385	**0.0014**	**<0.0001**	**0.0427**	0.1224	**<0.0001**	**<0.0001**	**<0.0001**
210882	0.7120	**0.0004**	**<0.0001**	**0.0087**	0.0932	0.1342	**0.0001**	**<0.0001**	0.1136
210891	0.7747	**0.0464**	**0.0480**	0.4948	0.1733	0.2466	**0.0035**	**<0.0001**	0.1939
210893	0.6533	0.1541	0.2891	0.8712	0.7708	0.9102	**0.0038**	**<0.0001**	0.2488
210894	0.0902	**0.0038**	**0.0001**	0.3044	0.3179	0.3402	**0.0003**	**<0.0001**	**<0.0001**
210896	0.2475	**0.0013**	**0.0003**	0.3410	0.5170	0.4151	**0.0005**	**<0.0001**	**0.0003**
210901	0.7331	**0.0005**	**0.0002**	0.2954	0.2691	0.1761	**0.0004**	**<0.0001**	**0.0004**
210904	0.9856	**0.0028**	**0.0019**	0.5229	0.4157	0.1100	**0.0006**	**<0.0001**	**0.0007**
210909	0.8657	**0.0003**	**0.0026**	0.6178	0.3145	0.1614	**0.0011**	**<0.0001**	**0.0011**
210912	0.3937	**0.0053**	**0.0051**	0.6907	0.8608	0.8120	0.2074	**<0.0001**	**0.0051**
211886	0.8424	0.4156	0.3498	**0.0036**	0.9549	**0.0107**	0.0525	**<0.0001**	0.9714
211890	0.4291	**0.0250**	0.0684	0.3410	0.2702	**0.0380**	**0.0002**	**<0.0001**	0.4012
211891	0.1571	**0.0314**	**0.0206**	0.0991	**0.0022**	**<0.0001**	0.7293	0.1500	**0.0103**
211894	0.6931	0.0793	0.2055	**0.0282**	0.5209	**0.0034**	**0.0001**	**<0.0001**	0.1713
211896	0.9746	0.5284	0.6454	0.5983	0.6955	0.3134	0.7166	0.3734	0.8775
211898	0.3173	**0.0005**	**0.0001**	0.3072	0.4081	0.6088	**0.0004**	**<0.0001**	**0.0008**
211911	0.9260	**0.0241**	**0.0117**	0.6537	0.1017	0.2538	**0.0003**	**<0.0001**	**0.0020**
211914	0.9501	**0.0046**	**0.0012**	0.7341	0.6274	0.6936	**0.0003**	**<0.0001**	**0.0002**
211916	0.9703	**0.0255**	**0.0154**	0.5646	0.8512	0.7140	0.9737	0.7816	0.8940
211917	0.5334	0.0571	**0.0162**	0.7753	0.9376	0.7545	**0.0118**	**<0.0001**	**0.0438**
211919	**0.0105**	**0.0031**	0.6716	0.1358	**0.0278**	0.1044	**<0.0001**	**<0.0001**	**0.0002**
211934	0.9231	**0.0273**	**0.0298**	0.4001	0.3427	0.3236	**0.0001**	**<0.0001**	**0.0076**
211935	0.7506	0.0710	0.1891	0.5273	0.8064	0.4161	**0.0001**	**<0.0001**	0.2290
211941	0.4404	0.2741	0.3412	0.0839	0.2619	0.5024	0.0725	**0.0001**	0.9709
211946	0.7432	0.1239	0.0921	**0.0224**	0.4835	0.1053	**0.0223**	**<0.0001**	**0.0095**
211962	0.8779	**0.0036**	**0.0012**	0.1552	**0.0003**	**0.0343**	0.1439	0.1293	**0.0065**
211972	0.6982	**0.0027**	**0.0025**	0.9708	**0.0031**	**0.0100**	**<0.0001**	**<0.0001**	**0.0352**
211980	0.4535	**0.0045**	**0.0147**	0.3817	0.7929	0.3390	**0.0294**	**0.0018**	0.4969
212024	0.6495	**0.0240**	**0.0016**	0.6558	0.3421	0.0784	**0.0002**	**<0.0001**	**0.0016**
212189	0.3138	**0.0007**	**0.0001**	0.4609	0.5913	0.7574	**0.0002**	**<0.0001**	**0.0021**
212208	0.9335	0.9440	0.8071	0.3819	0.5758	**0.0073**	**0.0152**	**0.0081**	0.3968
212274	0.3163	0.3701	0.1144	**0.0108**	0.0590	0.5128	0.5701	**0.0276**	0.0910
212663	0.7168	0.5739	0.4495	**0.0001**	**0.0015**	**<0.0001**	**0.0177**	**<0.0001**	**<0.0001**
212679	0.8767	0.7243	0.9426	**0.0324**	0.0577	0.8509	0.3689	0.8232	0.5157
212732	0.6962	**0.0007**	**0.0065**	**<0.0001**	**<0.0001**	**0.0153**	**0.0002**	**<0.0001**	**0.0004**
212735	0.0625	0.4378	0.0567	0.1022	0.3620	**0.0130**	**<0.0001**	0.0577	**<0.0001**
212781	**<0.0001**	0.2250	**0.0024**	**0.0014**	**0.0010**	0.3064	**<0.0001**	**<0.0001**	**0.0015**
213143	0.1126	**0.0003**	0.0732	**0.0035**	**0.0001**	**<0.0001**	**0.0026**	**0.0030**	0.9959

Metabolite	PM			PFa			F1		
	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3	time 1 / time 2	time 1 / time 3	time 2 / time 3
213155	0.0606	**0.0054**	**0.0002**	0.4505	0.2596	0.7352	**0.0013**	**<0.0001**	**0.0010**
213182	**0.0011**	0.1349	**<0.0001**	**0.0131**	**0.0082**	**<0.0001**	0.4087	**<0.0001**	**<0.0001**
213185	**0.0051**	**0.0007**	0.9184	**0.0372**	**0.0426**	0.2355	**0.0256**	**0.0015**	0.4814
213194	0.9722	**0.0036**	**0.0021**	0.3510	**0.0001**	**<0.0001**	0.2775	**<0.0001**	**<0.0001**
213271	0.1997	0.0959	**0.0440**	0.3286	0.5544	0.6621	**0.0213**	**0.0025**	0.6299
213310	0.7521	**0.0338**	0.1777	**0.0168**	0.2284	0.5618	0.4312	**0.0132**	**0.0341**
213714	0.5211	0.0509	**0.0080**	0.9911	0.5304	0.6365	0.3926	0.1884	0.4247
213732	0.2217	0.5376	0.3786	0.7886	0.9362	0.8453	0.6287	0.5535	0.7343
214401	**0.0263**	0.5559	**0.0187**	0.5913	0.6411	0.4160	**<0.0001**	**0.0123**	**<0.0001**
214405	0.3546	0.9910	0.4095	0.6523	0.9378	0.6109	0.8154	0.1884	0.2167
214410	**0.0018**	0.2863	0.2041	0.0870	0.0708	0.3907	**<0.0001**	**0.0075**	**0.0225**
214434	0.3930	0.6219	0.1955	0.3893	0.3079	0.6867	0.3313	0.0961	0.3081
214680	0.4215	**0.0054**	**0.0081**	0.5694	**0.0094**	**0.0217**	0.8141	0.1759	0.2393
215062	0.3633	0.6256	0.6877	0.3036	**0.0006**	0.1382	0.6929	0.1725	0.2044
215344	0.4170	**0.0012**	**0.0001**	0.2945	0.3121	0.8566	**0.0003**	**<0.0001**	**0.0006**
215347	0.6779	**0.0023**	**0.0013**	0.2303	0.2861	0.3924	**0.0003**	**<0.0001**	0.0777
215355	0.5119	**0.0003**	**<0.0001**	**0.0257**	0.3213	0.1493	**0.0001**	**<0.0001**	**0.0348**
215362	0.8705	**0.0031**	**0.0003**	0.3550	0.3312	0.4426	**0.0005**	**<0.0001**	**0.0008**
215375	0.0741	**0.0005**	**0.0001**	0.1670	**0.0297**	0.2322	0.1837	**0.0045**	**0.0411**
215399	0.7476	**0.0016**	**0.0003**	0.5506	0.2439	0.0603	**0.0003**	**<0.0001**	**0.0004**
215402	0.8343	**0.0013**	**0.0002**	0.3902	0.4060	0.9214	**0.0004**	**<0.0001**	**0.0019**
215448	0.5288	0.6547	0.9658	0.3298	0.7054	0.4677	**0.0484**	**<0.0001**	**0.0050**
215466	0.4651	0.4535	0.9690	0.4337	0.7441	0.6372	0.6275	0.1169	0.4257
215490	0.9222	0.9731	0.8880	0.4135	0.2591	0.5390	0.9174	0.8621	0.7942
215492	0.6551	0.1165	0.2453	0.8908	0.5948	0.7233	**0.0338**	**<0.0001**	**0.0475**
215493	0.2985	**0.0042**	**0.0131**	0.8243	0.3778	0.1720	**0.0001**	**<0.0001**	0.2791
215504	0.1735	**0.0023**	**0.0355**	0.9440	0.5312	0.5947	**0.0250**	**<0.0001**	**<0.0001**
215529	0.4244	**0.0104**	**0.0011**	0.5563	0.3743	0.2542	**0.0001**	**<0.0001**	**<0.0001**
215555	0.0603	**0.0009**	**0.0305**	0.2737	0.2979	0.8677	**0.0026**	**<0.0001**	0.7564
215563	0.6090	**0.0007**	**0.0019**	0.2132	0.0616	**0.0356**	**0.0002**	**<0.0001**	**0.0008**
215643	0.3640	**0.0009**	**0.0001**	0.3155	0.3106	0.4302	**0.0005**	**<0.0001**	**0.0021**
215682	0.1667	0.5463	0.3555	0.7352	0.7856	0.9096	**0.0055**	**<0.0001**	0.7205
215860	0.7108	**0.0271**	0.1154	0.8303	0.4878	0.5814	0.4968	0.4854	0.9604
215978	0.0939	**0.0030**	**0.0004**	0.5992	0.3138	0.2497	**0.0007**	**<0.0001**	**<0.0001**
216098	0.7009	**0.0007**	**0.0002**	0.6047	0.4862	0.5397	**0.0001**	**<0.0001**	**0.0008**
216424	0.1866	**0.0198**	**0.0024**	0.7926	0.2601	0.0558	**0.0007**	**<0.0001**	**0.0126**
216427	0.3769	0.2054	0.6975	**0.0437**	0.1214	0.4664	0.1592	0.1058	0.8988
216454	0.9117	**0.0039**	**0.0031**	0.1216	0.1224	0.6224	**0.0002**	**<0.0001**	0.0504
216472	0.8931	0.3808	0.3730	**0.0398**	0.2028	**0.0002**	0.4044	**<0.0001**	**<0.0001**
216493	0.1217	**0.0005**	**<0.0001**	0.3725	0.2036	0.2402	**0.0017**	**<0.0001**	**0.0004**
216564	0.3917	0.8993	0.4164	0.7936	0.5931	0.8362	0.5093	0.2402	0.4717



Appendix Figure 1: Weather data (temperature and rainfall) for 2005, 2006 and 2007. Green arrows: date of greenhouse harvests, brown arrows: date of field harvests.

Appendix Table 3: Biomass data for greenhouse harvest December 2005. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter.

Greenhouse	December 05								
	Replication 1			Replication 2			Replication 3		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	1.4 ± 0.2	0.4 ± 0.1	31.1 ± 8.7	1.1 ± 0.0	0.2 ± 0.1	19.7 ± 12.4	4.1 ± 3.8	0.6 ± 0.6	16.0 ± 1.2
Av PFA	3.2 ± 1.0	0.6 ± 0.1	18.9 ± 3.6	2.9 ± 1.6	0.6 ± 0.2	24.6± 7.9	1.8 ± 0.3	0.3 ± 0.0	17.2 ± 2.5
Av F1	7.0 ± 2.2	1.2 ± 0.4	17.5 ± 1.8	5.9 ± 1.4	1.1 ± 0.3	19.2 ± 2.3	8.9 ± 2.6	1.6 ± 0.4	18.0 ± 1.8
Av F2	6.5 ± 3.3	1.0 ± 0.5	15.2 ± 3.3	6.6 ± 2.9	1.0 ± 0.5	15.4 ± 2.8	7.3 ± 3.4	1.1 ± 0.5	14.9 ± 2.1
MPH (%)	209.7	146.4	-30.0	196.9	172.1	-13.1	203.7	232.6	8.2
BPH (%)	122.4	112.6	-43.7	106.3	85.0	-21.7	117.8	144.0	4.4
MPH F2 (%)	186.0	91.0	-39.2	228.6	139.3	-30.3	149.6	130.2	-10.3
BPH F2 (%)	105.4	64.8	-51.1	128.3	62.7	-37.3	79.1	68.9	-13.5

Appendix Table 4: Biomass data for greenhouse harvest February 2006. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter. ^) parental data incomplete.

Greenhouse	February 06								
	Replication 1			Replication 2			Replication 3		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	1.2 ± 0.1	0.3 ± 0.0	24.5 ± 0.6	^	^	^	3.0 ± 3.1	0.7 ± 0.5	27.0 ± 6.2
Av PFA	2.4 ± 0.9	0.4 ± 0.1	17.2 ± 0.6	3.1 ± 1.1	0.7 ± 0.2	24.7 ± 4.2	1.3 ± 0.4	0.3 ± 0.2	23.5 ± 5.2
Av F1	8.5 ± 3.3	1.7 ± 0.6	20.5 ± 2.6	9.0 ± 3.0	1.8 ± 0.5	20.9 ± 3.1	9.4 ± 3.5	2.0 ± 0.6	22.2 ± 2.4
Av F2	6.5 ± 4.1	1.2 ± 0.7	18.9 ± 5.1	6.3 ± 3.9	1.2 ± 0.6	19.2 ± 4.1	6.7 ± 3.4	1.4 ± 0.7	21.5 ± 3.5
MPH (%)	386.3	395.6	-1.6	^	^	^	342.8	299.4	-12.2
BPH (%)	262.8	323.3	-16.3	^	^	^	215.1	187.7	-18.0
MPH F2 (%)	273.1	248.0	-9.0	^	^	^	212.3	175.2	-15.0
BPH F2 (%)	178.3	197.2	-22.6	^	^	^	122.2	98.2	-20.6

Appendix Table 5: Biomass data for greenhouse harvest April 2006. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter. ^) parental data incomplete.

Greenhouse	April 06								
	Replication 1			Replication 2			Replication 3		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	1.3	0.4	28.8	^	^	^	9.6 ± 5.8	2.2 ± 1.1	24.3 ± 4.2
Av PFA	3.9 ± 1.5	1.0 ± 0.4	24.9 ± 0.2	3.5 ± 1.6	0.6 ± 0.3	18.3 ± 0.6	11.1 ± 5.7	2.5 ± 1.2	23.6 ± 4.5
Av F1	12.1 ± 3.5	2.6 ± 0.7	22.3 ± 2.7	10.0 ± 2.1	2.0 ± 0.3	20.1 ± 1.6	13.1 ± 5.7	2.9 ± 1.0	23.8 ± 4.9
Av F2	10.6 ± 5.9	2.1 ± 1.0	21.6 ± 5.2	8.3 ± 3.9	1.4 ± 0.6	18.2 ± 4.0	11.4 ± 6.6	2.5 ± 1.2	23.9 ± 5.0
MPH (%)	370.7	297.9	-16.9	^	^	^	26.0	22.3	-0.7
BPH (%)	211.6	173.8	-22.6	^	^	^	17.4	13.8	-2.2
MPH F2 (%)	313.0	220.4	-19.5	^	^	^	9.5	6.1	-0.3
BPH F2 (%)	173.4	120.5	-25.0	^	^	^	2.0	-1.2	-1.9

Appendix Table 6: Biomass data for field harvest August 2006. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter. ^) parental data incomplete.

Field	August 06					
	Replication 1			Replication 2		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	3.5	1.3	37.1	^	^	^
Av PFA	^	^	^	16.9 ± 16.0	5.9 ± 5.4	35.6 ± 5.0
Av F1	38.9 ± 21.3	13.4 ± 7.3	34.8 ± 2.6	30.1 ± 16.0	10.1 ± 5.3	34.1 ± 4.0
Av F2	25.1 ± 20.9	9.1 ± 7.1	37.9 ± 5.7	29.4 ± 25.0	9.7 ± 7.7	34.2 ± 4.8
MPH (%)	^	^	^	^	^	^
BPH (%)	^	^	^	^	^	^
MPH F2 (%)	^	^	^	^	^	^
BPH F2 (%)	^	^	^	^	^	^

Appendix Table 7: Biomass data for field harvest October 2006. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter. ^) parental data incomplete.

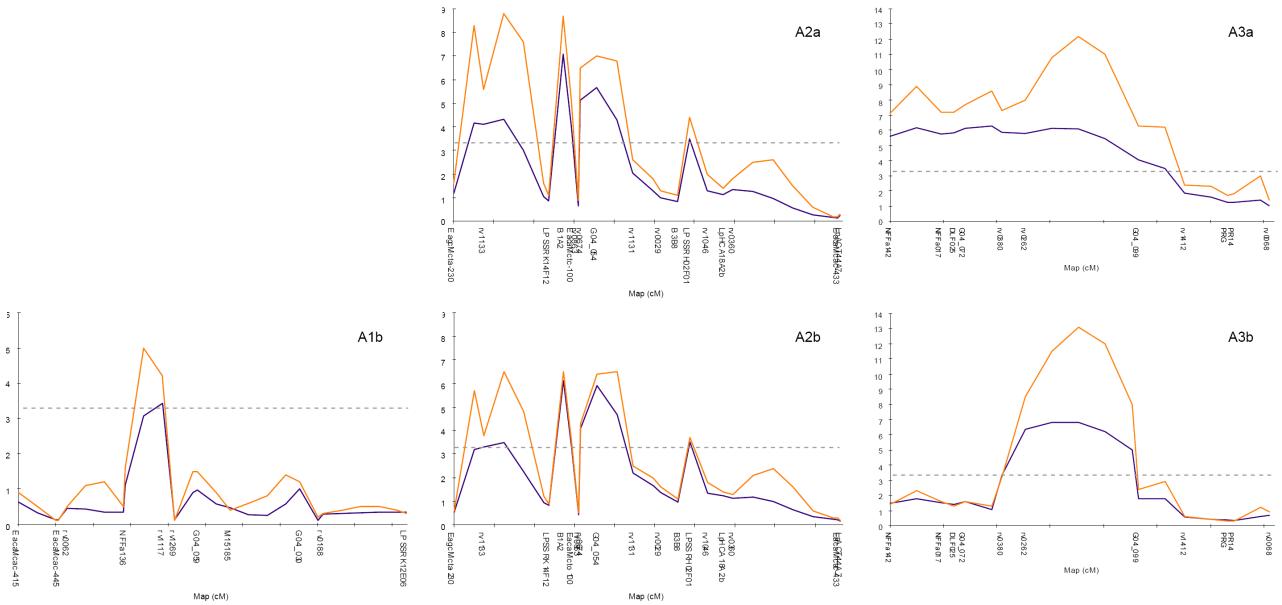
Field	October 06					
	Replication 1			Replication 2		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PM	^	^	^	^	^	^
Av PFA	^	^	^	86.5 ± 14.0	20.5 ± 1.0	24.0 ± 2.8
Av F1	99.1 ± 41.9	20.6 ± 8.0	21.2 ± 3.5	69.8 ± 42.4	16.3 ± 10.0	24.2 ± 3.7
Av F2	58.0 ± 48.4	12.7 ± 9.7	26.3 ± 16.7	69.6 ± 221.5	13.9 ± 12.2	26.2 ± 9.3
MPH (%)	^	^	^	^	^	^
BPH (%)	^	^	^	^	^	^
MPH F2 (%)	^	^	^	^	^	^
BPH F2 (%)	^	^	^	^	^	^

Appendix Table 8: Biomass data for field harvest May 2007. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%). Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter. ^) parental data incomplete.

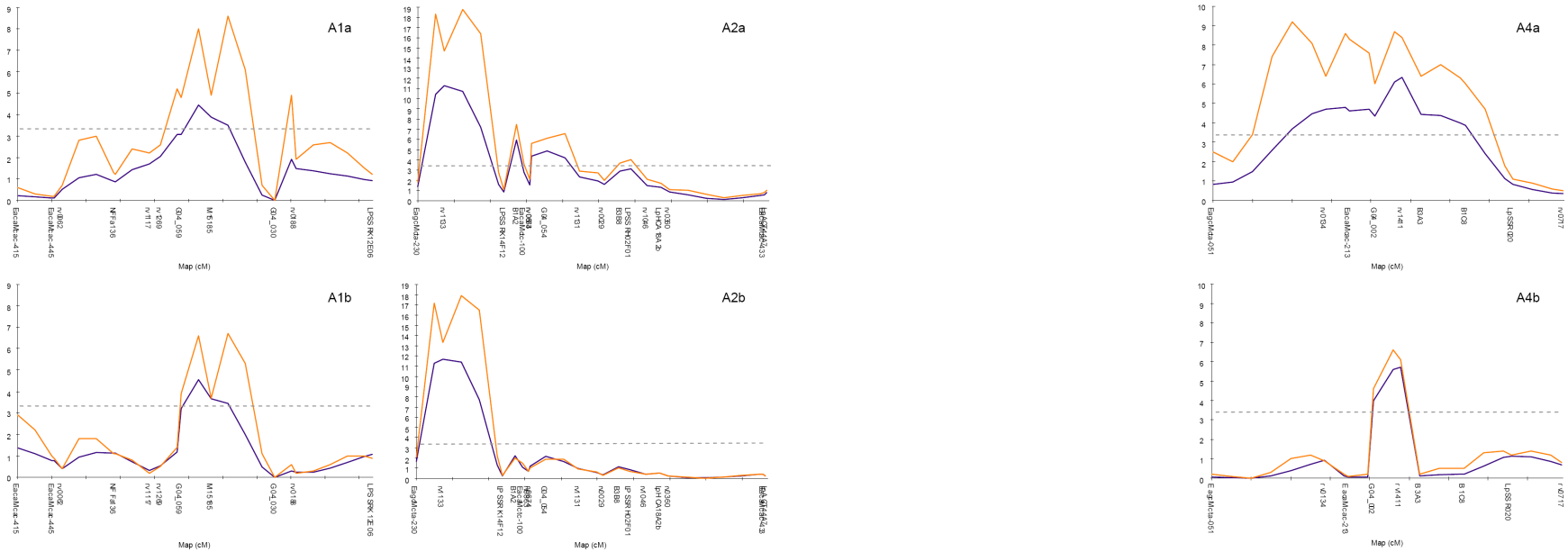
Field	May 07					
	Replication 1			Replication 2		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PFA	^	^	^	96.7	26.5	27.4
Av PM	12.4	3.7	30.1	10.4 ± 0.8	3.0 ± 0.1	29.3 ± 1.7
Av F1	192.6 ± 77.9	51.5 ± 21.4	26.7 ± 1.1	181.2 ± 71.2	47.5 ± 18.2	26.4 ± 2.4
Av F2	114.1 ± 84.7	30.7 ± 22.6	27.2 ± 1.9	119.6 ± 95.6	31.4 ± 24.2	26.8 ± 2.6
MPH (%)	^	^	^	238.3	221.5	-6.9
BPH (%)	^	^	^	87.4	79.2	-10.0
MPH F2 (%)	^	^	^	123.3	112.1	-5.4
BPH F2 (%)	^	^	^	23.7	18.2	-8.4

Appendix Table 9: Biomass data for field harvest August 2007. Replication 1 to 3. Traits: Fresh weight (g), dry weight (g), dry matter (%).Averages (Av) of parental lines (PFa: paternal and PM: maternal), F_1 and F_2 generation. Mid parent heterosis (MPH) and best parent heterosis (BPH) for the F_1 and F_2 generation for the traits fresh weight, dry weight, and dry matter.

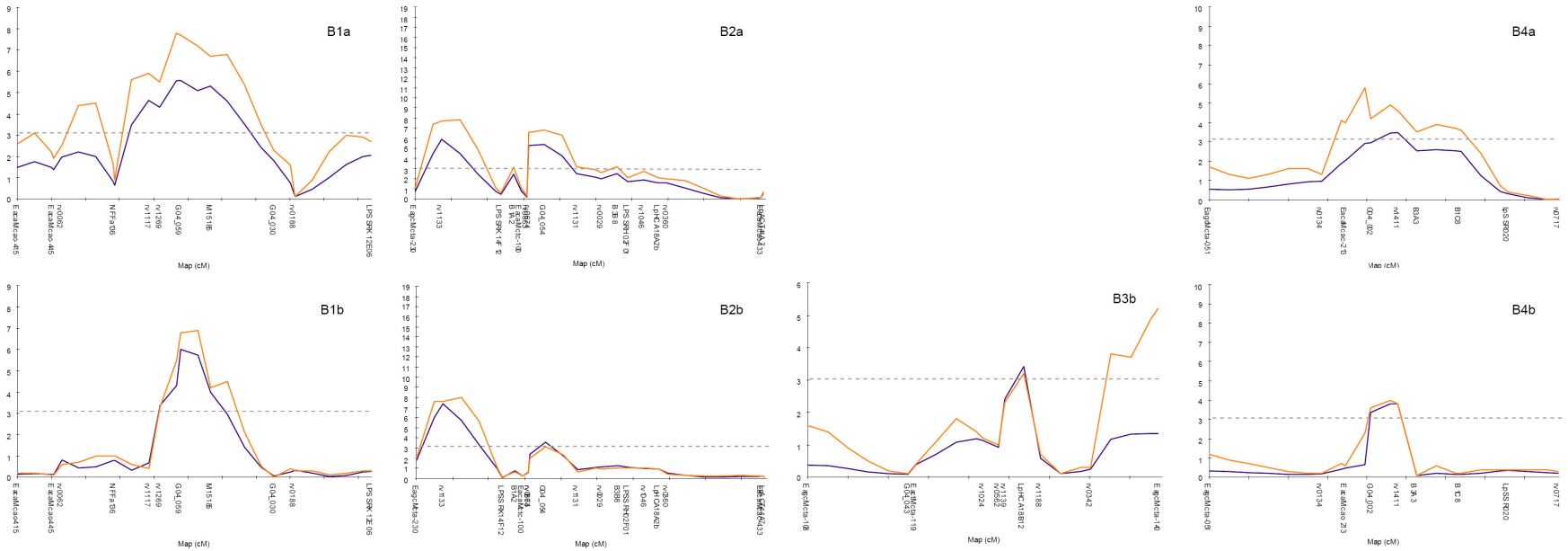
Field	August 07					
	Replication 1			Replication 2		
plant	fresh weight (g)	dry weight (g)	dry matter (%)	fresh weight (g)	dry weight (g)	dry matter (%)
Av PFA	158.9	49.9	31.4	64.6 ± 73.0	18.0 ± 20.1	28.7 ± 1.3
Av PM	40.7 ± 44.5	11.7 ± 12.6	32.4 ± 12.4	29.1	10.7	36.8
Av F1	142.2 ± 47.0	42.1 ± 13.7	29.9 ± 3.0	138.6 ± 42.1	36.9 ± 11.4	26.6 ± 1.7
Av F2	116 .2 ± 79.4	32.8 ± 21.4	29.5 ± 3.9	112.6 ± 76.0	29.5 ± 18.8	27.3 ± 3.7
MPH (%)	42.5	36.6	-6.2	196.0	157.1	-18.6
BPH (%)	-10.5	-15.7	-7.6	114.7	104.8	-27.6
MPH F2 (%)	16.5	6.5	-7.5	140.5	105.8	-16.5
BPH F2 (%)	-26.8	-34.3	-8.9	74.4	64.0	-25.7



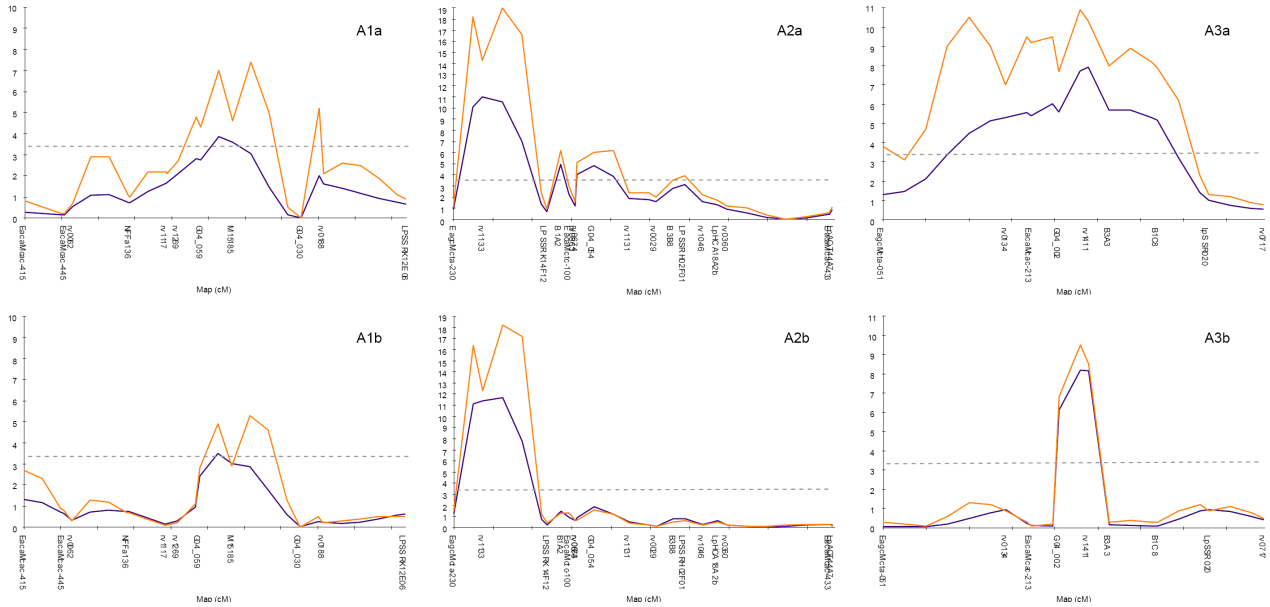
Appendix Figure 2: QTL for leaf width measured in the greenhouse. A1: LG 2; A2: LG 3; and A3: LG 4 (left to the right) and with (a) interval mapping (IM) and (b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



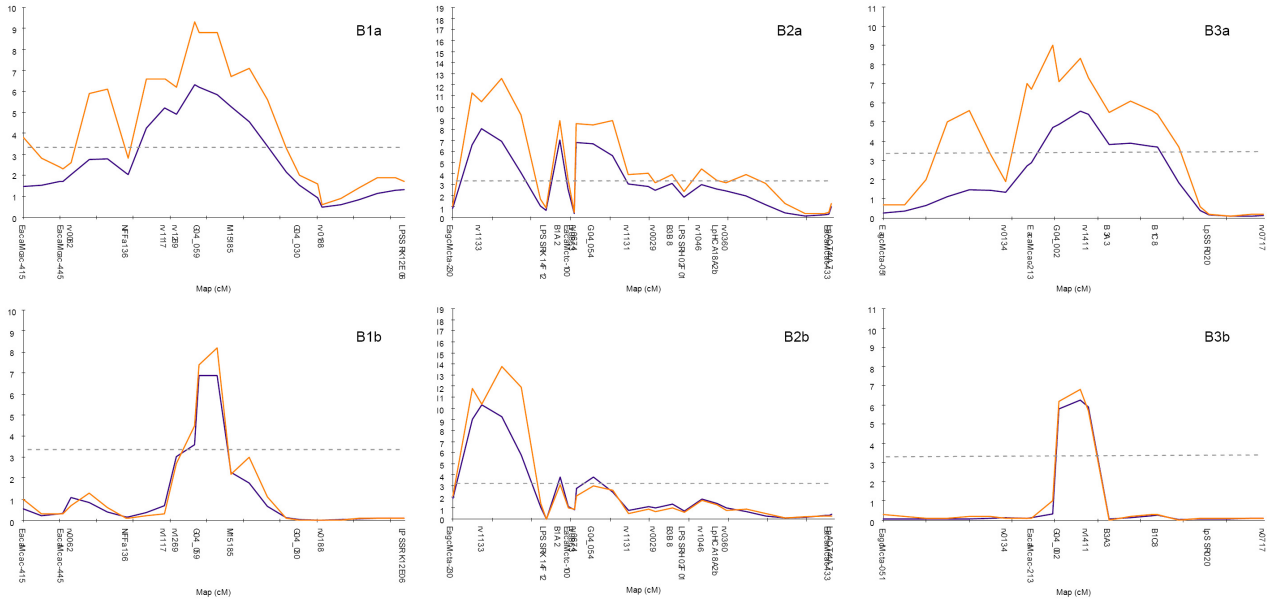
Appendix Figure 3 a: Fresh weight QTL, A1: LG 2; A2: LG3; and A4: LG7. A) Over all greenhouse harvests (a) interval mapping (IM) and (b) multiple QTL models (MQM) mapping: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



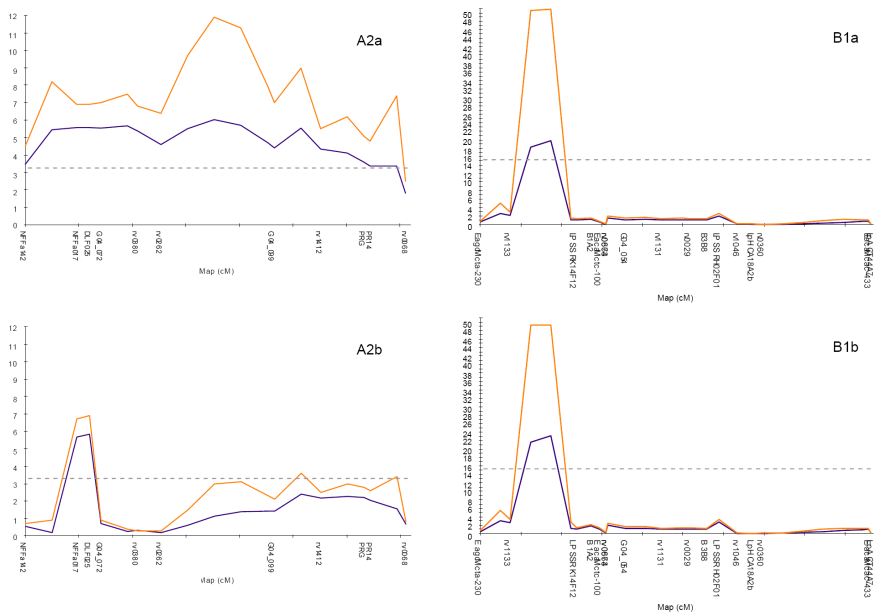
Appendix Figure 3 b: Fresh weight QTL. B) Over all field harvests a) Interval mapping (IM), A1: LG 2; A2: LG3; and A4: LG7, b) multiple QTL models (MQM) mapping, A1: LG 2; A2: LG3; A3: LG 5 and A4: LG7. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



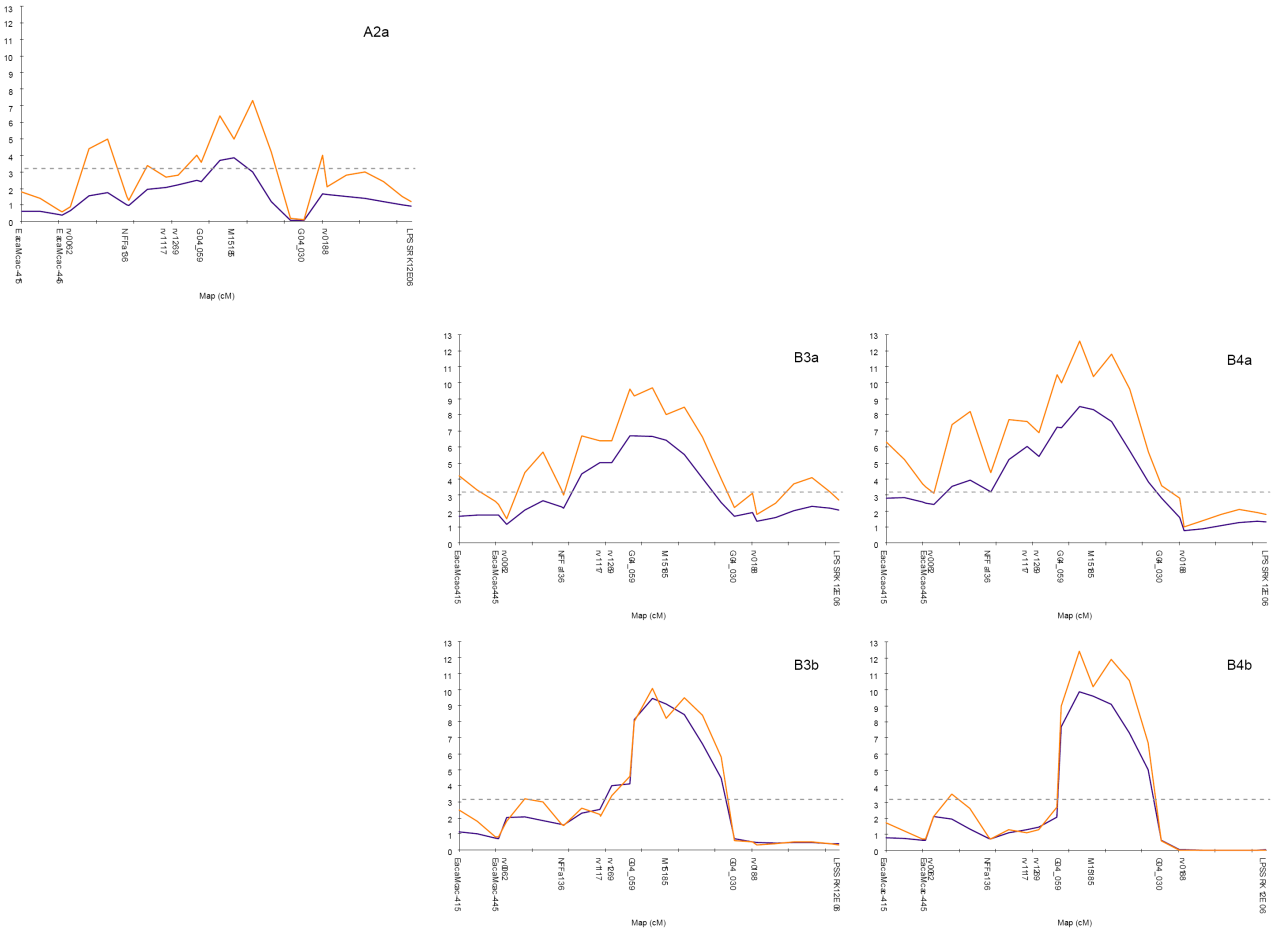
Appendix Figure 4 a: Dry weight QTL, A1: LG 2; A2: 3; A3: 7. A) Over all greenhouse harvests. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



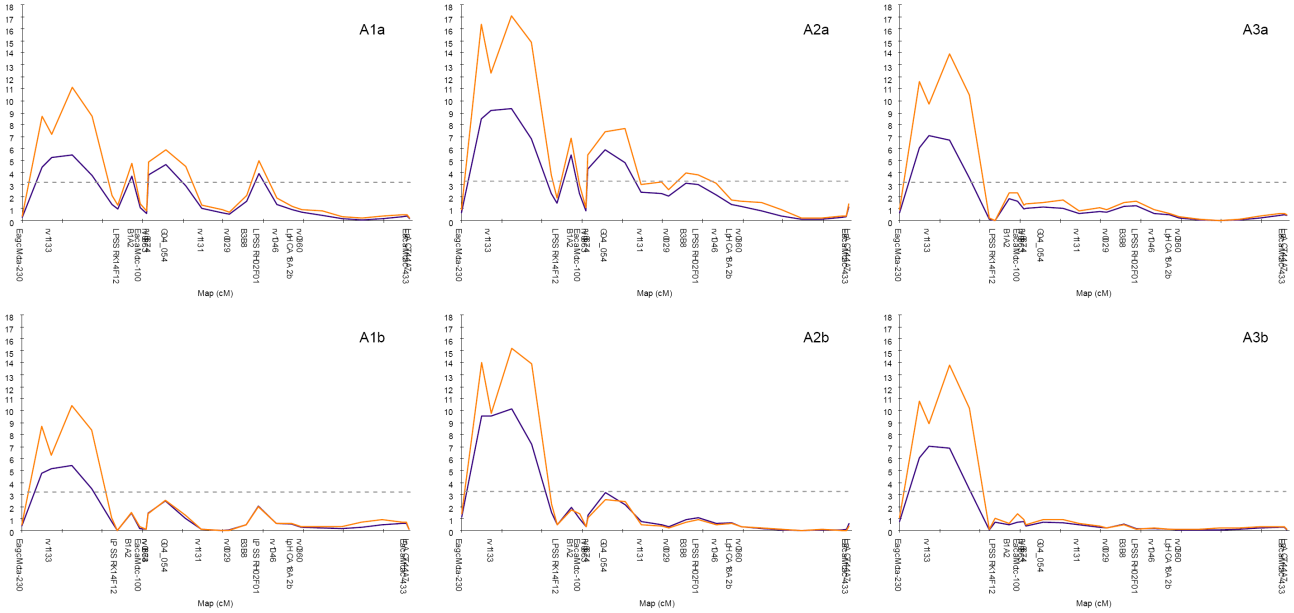
Appendix Figure 4 b: Dry weight QTL, A1: LG 2; A2: 3; A3: 7. B) Over all field harvests. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



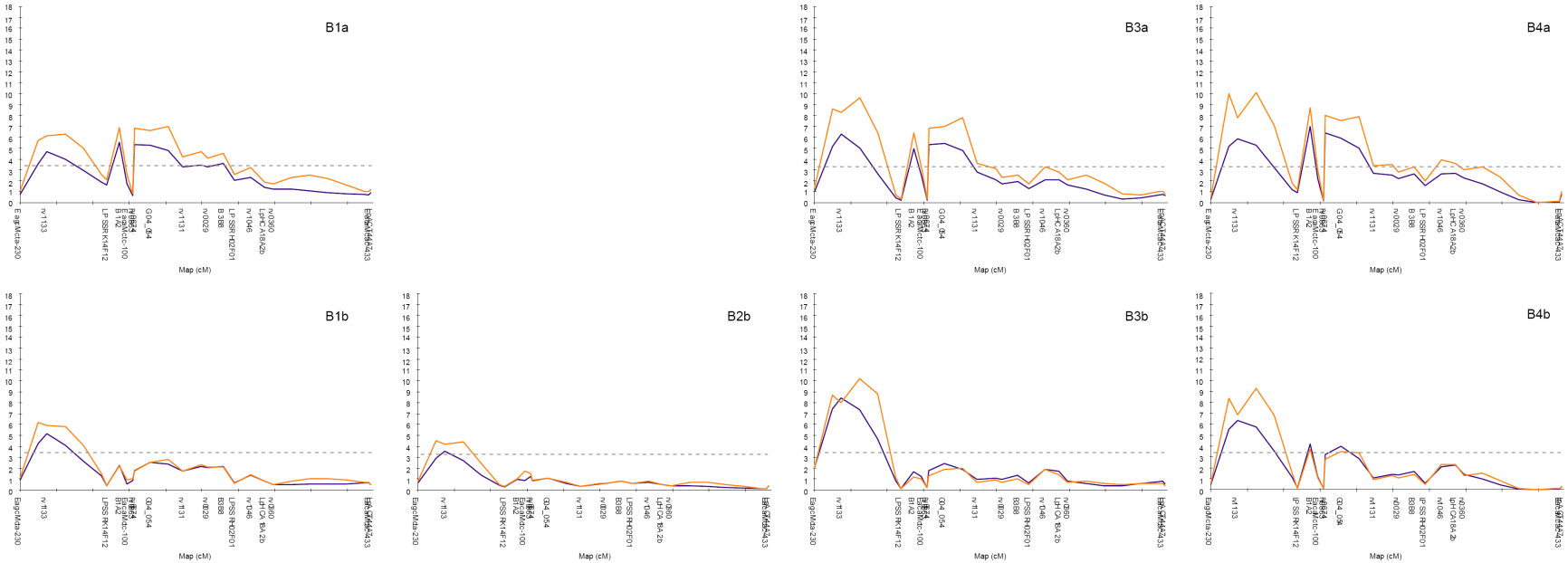
Appendix Figure 5: Dry matter QTL A) Over all greenhouse harvest. A2: LG 4. B) Over all field harvest, B1: LG 3. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



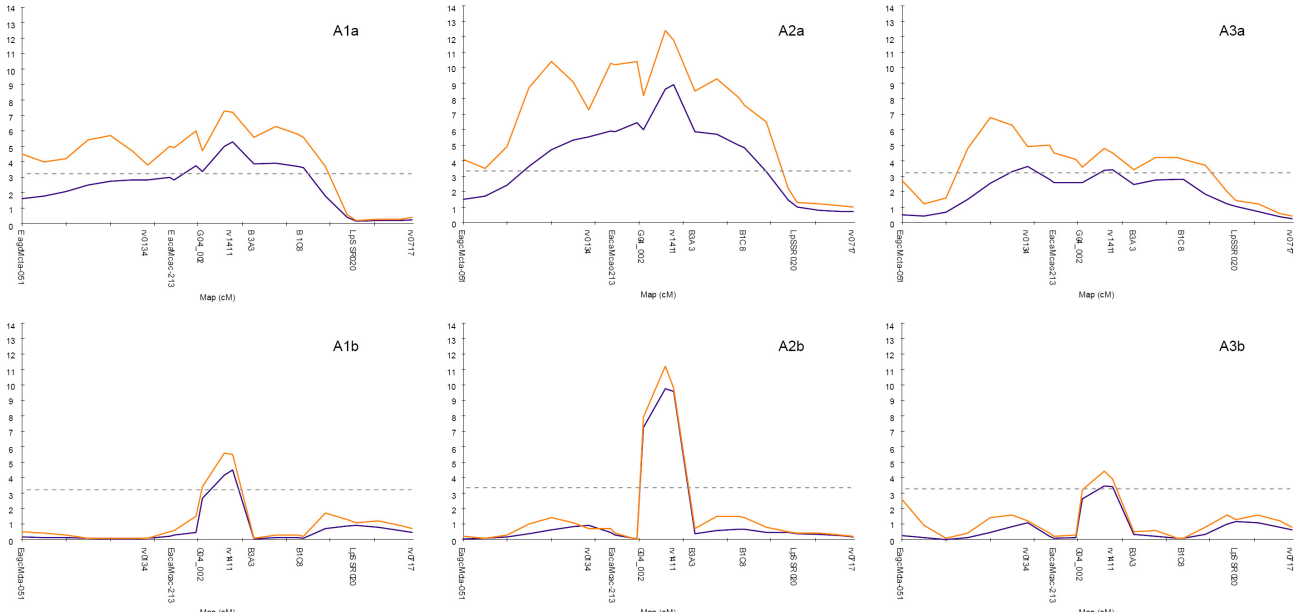
Appendix Figure 6: Dry weight QTL, LG 2. A) Single greenhouse harvests: 1) December 05 no data displayed (nd); 2) February 05; 3) April 06 (nd). B) Single field harvests: 1) August 06 (nd); 2) October 06 (nd); 3) May 07; 4) August 07. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



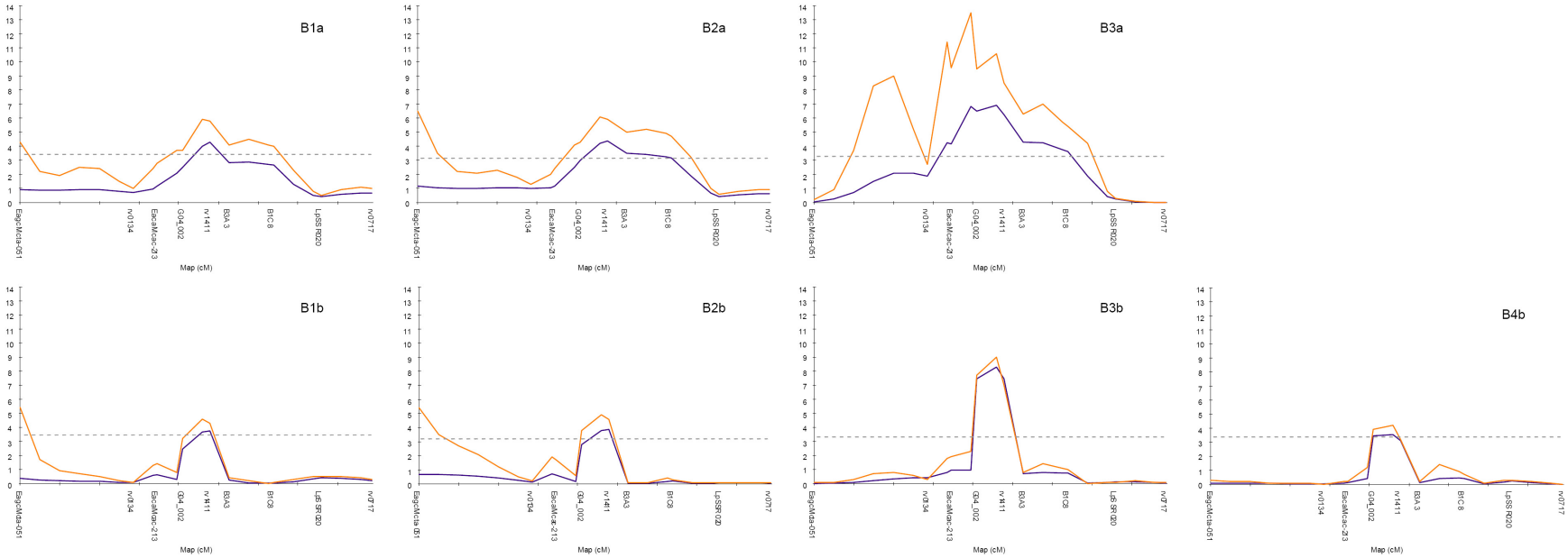
Appendix Figure 7 a: Dry weight QTL, LG 3. A) Single greenhouse harvests: 1) December 05; 2) February 05; 3) April 06. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



Appendix Figure 7 b: Dry weight QTL, LG 3. B) Single field harvests: 1) August 06; 2) October 06; 3) May 07; 4) August 07. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



Appendix Figure 8 a: Dry weight QTL, LG 7. A) Single greenhouse harvests: 1) December 05; 2) February 05; 3) April 06. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score.



Appendix Figure 8 b: Dry weight QTL, LG 7. B) Single field harvests: 1) August 06; 2) October 06; 3) May 07; 4) August 07. a) Interval mapping (IM) and b) multiple QTL models (MQM) mapping. Blue line: LOD score, orange line: % explained variance, dashed line: maximum LOD score

References

Acquaah G (2007) Principles of plant genetics and breeding. Blackwell Publishing Ltd, Malden, USA

Agronomix Software, Inc., Winnipeg, Manitoba, R3N 0S4, Canada

Aldesuquy HS, Ibrahim AHA (2000) The role of shikimic acid in regulation of growth, transpiration, pigmentation, photosynthetic activity and productivity of Vigna sinensis plants. Phyton 40(2):277-292

Amiard V, Morvan-Bertrand A, Billard J-P, Huault C, Keller F, Prud'homme M-P (2003) Fructans, but not the sucrosyl-galactosides, raffinose and loliose, are affected by drought stress in perennial ryegrass. Plant Physiol 132:2218-2229

Andaya VC, Tai TH (2006) Fine mapping of the *qCTS12* locus, a major QTL for seedling cold tolerance in rice. Theor Appl Genet 113:467–475

Anhalt UCM, Heslop-Harrison P (JS), Byrne S, Guillard A, Barth S (2008) Segregation distortion in *Lolium*: evidence for genetic effects. Theor Appl Genet 117:297–306

Armstead IP, Turner LB, Farrell M, Skøt L, Gomez P, Montoya T, Donnison IS, King IP, Humphreys MO (2004) Syntheny between a major heading-date QTL in perennial ryegrass (*Lolium perenne* L.) and the *Hd3* heading-date locus in rice. Theor Appl Genet 108:822-828

Armstead IP, Turner LB, King IP, Cairns AJ, Humphreys MO (2002) Comparison and integration of genetic maps generated from F_2 and BC_1-type mapping populations in perennial ryegrass. Plant Breed 21:501-507

Armstead IP, Turner LB, Marshall AH, Humphreys MO, King IP, Thorogood D (2008) Identifying genetic components controlling fertility in the outcrossing grass species

perennial ryegrass (*Lolium perenne*) by quantitative genetics. New Phytol 178(3):559-571

Asins MJ (2002) Present and future of quantitative trait locus analysis in plant breeding. Plant Breed 121:281-291

Barre P, Mi F, Balfourier F, Ghesquière M (2000) QTLs for morphogenetic traits and sensitivity to rusts in *Lolium perenne*. Second Molecular Breeding in Forage Crops congress, Lorne and Hamilton, Australia, p 60

Barth S, Busimi AK, Utz HF, Melchinger AE (2003) Heterosis for biomass yield and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. Heredity 91:36-42

Becker H (1993) Pflanzenzüchtung. Eugen Ulmer Verlag, Stuttgart

Bert PF, Charmet G, Sourdille P, Hayward MD, Balfourier F (1999) A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. Theor Appl Genet 99: 445-452

Botstein D, White, RL, Skolnick MH, Davis RW (1980) Construction of a linkage map in man using restriction fragment polymorphism. Am J Hum Genet 32:314-331

Broman KW (2005) The genomes of recombinant inbred lines. Genet 160:1133-1146

Bruckner PL, Raymer PL, Burton GW (1991) Recurrent phenotypic selection for forage yield in rye. Euphytica 54: 11–17

Burton GW (1992) Recurrent restricted phenotypic selection. Plant Breed Rev 9: 101–113

Byrne S, Guiney E, Barth S, Donnison I, Mur LAJ, Milbourne D (2008) Identification of coincident QTL for days to heading, spike length and spikelets per spike in *Lolium perenne* L. *in preparation*

Campos Severi-Aguiar GD, Azeredo-Oliveira MTV (2005) Localization of rDNA sites in holocentric chromosomes of three species of triatomines (Heteroptera, Triatominae). Genet Mol Res 4(4):704-709

Canter PH, Pašakinskienė I, Jones RN, Humphreys MW (1999) Chromosome substitutions and recombination in the amphiploid *Lolium perenne* x *Festuca pratensis* cv Prior (2n=4x=28). Theor Appl Genet 98:809-814

Cao M, Koulman A, Johnson LJ, Lane GA, Rasmussen S (2008) Advanced data-mining strategies for the analysis of direct-infusion ion trap mass spectrometry data from the association of perennial ryegrass with its endophytic fungus, *Neotyphodium lolii*. Plant Physiol 146:1501-1514

Carbonell EA, Asins MJ, Baselga M, Balansard E, Gerig TM (1993) Power studies in the estimation of genetic parameters of quantitative trait loci for backcross and doubled haploid populations. Theor Appl Genet 86:411-416

Casler MD, Diaby M, Stendal C (2005) Heterosis and inbreeding depression for forage yield and fiber concentration in smooth bromgrass. Crop Sci 45:44-50

Chalmers J, Lidgett A, Cummings N, Cao Y, Forster J, Spangenberg G (2005) Molecular genetics of fructan metabolism in perennial ryegrass. Plant Biotech J 3:459-474

Chen F, Duran AL, Blount JW, Sumner LW, Dixon RA (2003) Profiling phenolic metabolites in transgenic alfalfa modified in lignin biosynthesis. Phytochemistry 64:1013–1021

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Cogan NOI, Ponting RC, Vecchies AC, Drayton MC, George J, Dracatos PM, Dobrowolski MP, Sawbridge TI, Smith KF, Spangenberg GC, Foster JW (2006)

Gene-associated single nucleotide polymorphism discovery in perennial ryegrass (*Lolium perenne* L.). Mol Gen Genomics 276:101-112

Cogan NOI, Smith KF, Yamada T, Francki MG, Vecchies AC, Jones ES, Spangenberg GC, Foster JW (2005) QTL analysis and comparative genomics of herbage quality traits in perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 110:364-380

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142(1-2):169-196

Connolly V (2001) Breeding improved varieties of perennial ryegrass. End of project report 3495, Teagasc, Crops Research Centre, Carlow, Ireland, ISBN 1841701998

Connolly V, Wright-Turner R (1984) Induction of cytoplasmic male-sterility into ryegrass (*Lolium perenne*). Theor Appl Genet 68:449-453

Darwin C (editor) (1876) The effects of cross and self fertilization in the vegetable kingdom. John Murray, London.

Doerge RW, Zeng Z-B, Weir BS (1997) Statistical issues in the search for genes affecting quantitative traits in experimental populations. Stat Sci 12:(3)195-219

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet 3:43-52

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15

Dubcovsky J, Dvořák J (1995) Ribosomal RNA multigene loci: Nomads of the *Triticeae* genomes. Genetics 140:1367-1377

Dumsday JL, Smith KF, Forster JW, Jones ES (2003) SSR-based genetic linkage analysis of resistance to crown rust (*Puccinia coronata* f. sp. *lolii*) in perennial ryegrass (*Lolium perenne*). Plant Pathol 52: 628–637

Erickson DL, Fenster CB, Stenøen HK, Price D (2004) Quantitative trait locus analyses and the study of evolutionary process. Mol Ecol 13:2505-2522

Faville MJ, Vecchies AC, Schreiber M, Dryaton MC, Hughes LJ, Jones ES, Guthridge KM, Smith KF, Sawbridge T, Spangenberg GC, Bryan GT, Forster JW (2004) Functionally associated molecular genetic marker map construction in perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 110:992-993

Fiehn O (2000) Metabolite profiling for plant functional genomics. Nat Biotechnol 18:1157-1161

Fiehn O (2002) Metabolomics – the link between genotypes and phenotypes. Plant Mol Biol 48:155-171

Fiehn O (2006) Metabolite profiling in *Arabidopsis*. *Arabidopsis* protocols 2nd edition. Salinas J, Sanchez-Serrano (eds.), Methods in Molecular Biology ser., Humana Press, Totowa NJ, 439-447

Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee DY, Lu Y, Moon S, Nikolau B (2008) Quality control for plant metabolomics: reporting MSI-compliant studies. Plant J (2008) 53, 691–704

Fridman E, Pleban T, Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. PNAS 97:4718-4723

Frisch M, Quint M, Lübberstedt T, Melchinger AE (2004) Duplicate marker loci can result in incorrect locus order on linkage maps. Theor Appl Genet 109:305-316

Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403-3417

Galasso I, Schmidt T, Pignone D, Heslop-Harrison (1995) The molecular cytogenetics of *Vigna unguiculata* (L.) Walp: the physical organization and characterisation of 18*s*-5.8*s*-25*s* rRNA genes, 5*s* rRNA genes, telomere-like sequences, and a family of centromeric repetitive DNA sequences. Theor Appl Genet 91:928-935

Gill GP, Wilcox DJ, Whittaker DJ, Winz RA, Bickerstaff P, Echt CE, Kent J, Humphreys MO, Elborough KM, Gardner RC (2006) A framework linkage map of perennial ryegrass based on SSR markers. Genome 49:354-364

Goel S, Chen Z, Conner JA, Akiyama Y, Hanna WW, Ozias-Akins P (2003) Delineation by fluorescence *in situ* hybridization of a single hemizygous chromosomal region associated with aposporous embryo sac formation in *Pennisetum squamulatum* and *Cenchrus ciliaris*. Genetics 163:1069-1082

Goodacre R, Roberts L, Ellis DI, Thorogood D, Reader SM, Ougham H, King I (2007) From phenotype to genotype: whole tissue profiling for plant breeding. Metabolomics 3(4):489-501

Grandillo S, Ku HM, Tanksley SD (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor Appl Genet 99:978-987

Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 90:33-38

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 8:299-309

Halket JM, Waterman D, Przyborowska AM, Patel RKP, Paul D. Fraser PD, Bramley PM (2004) Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J Exp Bot 56(410):219-243

Hamzehzarghani H, Kushalappa AC, Dion Y, Rioux S, Comeau A, Yaylayan V Marshall WD, Mather DE (2005) Metabolic profiling and factor analysis to discriminate quantitative resistance in wheat cultivars against fusarium head blight. Physiol Mol Plant Pathol 66:119–133

Hayward MD, Mcadam NJ, Jones JG, Evans C, Evans GM, Forster JW, Ustin A, Hossain KG, Quader B, Stammers M, Will JK (1994) Genetic markers and the selection of quantitative traits in forage grasses. Euphytica 77:269-275

Hayward MD, Forster JW, Jones JG, Dolstra O, Evans C, McAdam NJ, Hossain KG, Stammers M, Will J, Humphreys MO, Evans GM (1998) Genetic analysis of *Lolium*. I. Identification of linkage groups and the establishment of a genetic map. Plant Breed 117:451-455

Hayward MD, Vivero JL (1984) Selection for yield in *Lolium perenne*. II. Performance of spaced plant selections under competitive conditions. Euphytica 33: 787–800.

Hegde VS, Yadav SS, Kumar J (2007) Heterosis and combining ability for biomass and harvest index in chickpea under a drought-prone, short-duration environment. Euphytica 157:223–230

Hoecker N, Keller B, Piepho HP, Hochholdinger F (2006) Manifestation of heterosis during early maize (Zea mays L.) root development. Theor Appl Genet 112:421-429

Hubbard CE (1992) Grasses. Third edition. Pelican Books, London, UK p.149

Humphreys MO (1992) Association of agronomic traits with isozyme loci in perennial ryegrass (*Lolium perenne* L.). Euphytica 59:141-150

Humphreys MW, Pašakinskienė I, James AR, Thomas H (1998b) Physically mapping quantitative traits for stress-resistance in the forage grasses. J Exp Bot 49(327):1611-1618

Humphreys MW, Yadav RS, Cairns AJ, Turner LB, Humphreys J, Skøt L (2005) A changing climate for grassland research. New Phytol 169(1):9-26

Humphreys MW, Zare AG, Pašakinskienė I, Thomas H, Rogers WJ, Collin HA (1998a) Interspecific genomic rearrangements in androgenic plants derived from a *Lolium multiflorum* × *Festuca arundinacea* (2*n* = 5x = 35) hybrid. Heredity 80:78-82

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205-211

Jansen RC (1994) Controlling the type I and type II errors in mapping quantitative trait loci. Genetics 138:871-881

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447-1455

Jansen RC, Van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet 91:33-37

Jauhar PP (1993) Cytogenetics of the *Festuca-Lolium* complex, relevance to breeding. Springer-Verlag, Berlin, New York

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73

Jenczewski E, Gherardi M, Bonnin I, Prosperi JM, Oliviere I, Huguet T (1997) Insight on segregation distortions in two interspecific crosses between annual species of *Medicago* (Leguminosae). Theor Appl Genet 94:682-691

Jensen LB, Andersen JR, Frei U, Xing Y, Taylor C, Holm PB, Lübberstedt T (2005b) QTL mapping of vernalisation response in perennial ryegrass (*Lolium perenne* L.) reveals co-location with an orthologue of wheat *VRN1*. Theor Appl Genet 110: 527-536

Jensen LB, Muylle H, Arsen P, Andersen CH, Holm PB, Ghesquière M, Julier B, Lübberstedt T, Nielsen K, Riek J de, Roldán-Ruiz I, Roulund N, Taylor C,Vosman B, Barre P (2005a) Development and mapping of a public references set of SSR markers in *Lolium perenne* L. Mol Ecol Notes 5(4): 951-958

Jiang J, Gill BS (2006) Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. Genome 49:1057-1068

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000) Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344-347

John H, Birnstiel M, Jones K (1969) RNA:DNA hybrids at the cytogenetical level. Nature 223:582-587

Jones ES, Dupal MP, Dumsday JL, Hughes LJ, Foster JW (2002a) An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 105:577-584

Jones ES, Dupal MP, Kölliker R, Drayton MC, Forster JW (2001) Development and characterisation of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 102: 405-415

Jones ES, Hughes LJ, Drayton MC, Abberton MT, Michaelson-Yeates TPT, Bowen C, Foster JW (2003) An SSR and AFLP molecular marker-based genetic map of white clover (*Trifolium repens* L.). Plant Sci 165:531-539

Jones MB, Lanzenby A (1988) The grass crop. The physiological basis of production. Chapman and Hall Ltd, London, New York. pp 39

Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, King IP, Kishida T, Yamada T, Balfourier F, Charmet G, Foster JW (2002b) An enhanced

molecular marker based genetic map of perennial ryegrass (*Lolium perenne*) reveals comparative relationship with other Poaceae genomes. Genome 45:282-295

Jones N, Ougham H, Thomas H (1997) Markers and mapping: we are all genetics now. New Phytol 137:165-177

Jones N, Pašakinskienė I (2005) Genome conflict in the *Gramineae*. New Phytol 165(2):391-409

Karp G (2008) Cell and molecular biology. Concepts and experiments. 5th edition, John Wiley & Sons, Inc, Danvers, USA, pp. 107

King IP, Morgan WG, Armstead IP, Harper JA, Hayward MD, Bollard A, Nash JV, Forster JW, Thomas HM (1998) Introgression mapping in the grasses. I. Introgression of *Festuca pratensis* chromosomes and chromosome segments into *Lolium perenne*. Heredity 81:462-467

King IP, Morgan WG, Harper JA, Thomas HM (1999) Introgression mapping in the grasses. II. Meiotic analysis of the *Lolium perenne*/*Festuca pratensis* triploid hybrid. Heredity 82:107-112

King J, Armstead IP, Donnison IS, Harper JA, Roberts LA, Thomas H, Ougham H, Thomas A, Huang L, King IP (2007) Introgression mapping in the grasses. Chromosome Res 15:105-113

King J, Armstead IP, Donnison IS, Thomas HM, Jones RN, Kearsey MJ, Roberts LA, Thomas A, Morgan WG, King IP (2002) Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics 161:315-324

King J, Thorogood D, Edwards KJ, Armstead IP, Roberts L, Skøt K, Hanley Z, King IP (2008) Development of a genomic microsatellite library in perennial ryegrass (*Lolium perenne*) and its use in trait mapping. Ann Bot 101:845-853

Kopecký D, Lukaszewski AJ, Doležel J (2008) Cytogenetics of Festulolium (*Festuca* × *Lolium* hybrids). Cytogenet Genome Res 120:370-383

Kopecký D, Lukaszewski AJ, Gibeault V (2005) Reduction of ploidy level by androgenesis in intergeneric Lolium-Festuca hybrids for turf grass breeding. Crop Sci 45:274-281

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172-175

Kubik C, Sawkins M, Meyer WA, Gaut BS (2001) Genetic diversity in seven perennial ryegrass (*Lolium prenne* L.) cultivars based on SSR markers. Crop Sci 41:1565-1572

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185-199

Lauvergeat V, Bonnet M, Ghesquière M (2005) Primer note. Sixty simple sequence repeats (SSR) markers for use in the *Festuca/Lolium* complex of grasses. Mol Ecol 5(2):401-405

Lin Y-R, Schertz KF, Paterson AH (1995) Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. Genetics 141:391-411

Lincoln S, Daly M, Lander ES, (1992) Constructing genetic maps with MAPMAKER/EXP 3.0, Whitehead Institute, Cambridge, Massachusetts, USA

Linde-Laursen IB (1984) Nucleolus organizer polymorphism in barley, *Hordeum vulgare* L. Hereditas 100:33-43

Lippman ZB, Zamir D (2006) Heterosis: revisiting the magic. TrendsGenet 23(2): 60-66

Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, Fiehn O, Törjék O, Selbig J, Altmann T, Willmitzer L (2008) Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL population. Plant J 53:960–972

Littell RC, Henry PR, Ammerman CB (1998) Statistical analysis of repeated measures data using SAS procedures. J Anim Sci 76:1216-1231

Liu HL (1998) Statistical genomics: linkage, mapping, and QTL analysis. CRC Press LLCBoca Raton, Florida, USA

Lothier J, Lasseur B, Le Roy K, Van Laere A, Prud'homme MP, Barre P, Van den Ende W, Morvan-Bertrand A (2007) Cloning, gene mapping, and functional analysis of a fructan 1-exohydrolase (1-FEH) from *Lolium perenne* implicated in fructan synthesis rather than in fructan mobilization. J Exp Bot 58(8):1969-1983

Lyttle TW (1991) Segregation distortion. Annu Rev Genet 25:511-557

Malik CP, Thomas PT (1966) Karyotypic studies in some *Lolium* and *Festuca* species. Caryologia 19:167-196.

Maloof JN (2003) QTL plant growth and morphology. Curr Opin Plant Biol 6:85-90

Manly KF, Cudmore Jr RH, Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome 12:930-932

Meyer RC, Steinfath M, Lisec J, Becher M, Witucka-Wall H, Törjék O, Fiehn O, Eckardt Ä, Willmitzer L, Selbig J, Altmann T (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. PNAS 104(11):4759-4764

Michaels SD, Amasino RM (1998) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J 14:381-385

Mitton JB, Grant MC (1984) Associations among protein heterozygosity, growth rate, and developmental homeostasis. Ann Rev Ecol Syst 15:479-499

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol Breed 3:87-103

Muylle H (2003) Genetic analysis of crown rust resistance in ryegrass (*Lolium* spp.) using molecular markers. PhD thesis, University of Ghent, Belgium

Muylle H, Baert J, Van Bockstaele E, Pertijs J, Roldán-Ruiz I (2005) Four QTL determine crown rust (*Puccinia coronata* f. sp. *lolii*) resistance in a perennial ryegrass (*Lolium perenne*) population. Heredity 95:348-357

Muylle H, De Loose M, Peerbolte R, Van Bockstaele E, Roldán-Ruiz I (2001) Linkage map construction in the outcrossing species *Lolium perenne*. Med Fac Landbouww Univ Gent 66(3b):401-408

Pardue ML, Gall JG (1969) Molecular hybridization of radioactive DNA to the DNA of cytological preparation. Proc Natl Acad Sci USA 64:600-604

Pašakinskienė I, Anamthawat-Jónsson, Humphreys MW, Paplauskiene V, Jones RN (1998) New molecular evidence on genome relationships and chromosome identification in fescue (*Festuca*) and ryegrass (*Lolium*). Heredity 81:659-665

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721-726

Patterson HD, Williams ER (1976) A new class of resolvable incomplete block designs. Biometrica 63(1): 83-92

Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM (2007) GenStat for Windows (10th Edition) Introduction. VSN International, Hemel Hempstead.

Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philosophical Magazine 2(6): 559–572

Poggio L, Gonzalez G, Confalonieri V, Comas C, Naranjo CA (2005) The genome organization and diversification of maize and its allied species revisited: evidences from classical and FISH-GISH cytogenetic analysis. Cytogenet Genome Res 109:259-267

Price AH (2006) Believe it or not, QTLs are accurate! Trends Plant Sci 11(5)213-216

Quarrie SA, Pekic Quarrie S, Radosevic R, Rancic D, Kaminska A, Barnes JD, Leverington M, Ceoloni C, Dodig D (2006) Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. J Exp Bot 57(11):2627–2637

Rainey DY, Mitton JB, Monson RK (1987) Associations between enzyme genotypes and dark respiration in perennial ryegrass, *Lolium perenne L.* Oecol 74:335-338

Rasmussen S, Parsons AJ, Fraser K, Xue H, Newman JA (2008) Metabolic profiles of *Lolium perenne* are differentially affected by nitrogen supply, carbohydrate content, and fungal endophyte infection. Plant Physiol 146:1440-1453

Reyna N, Sneller CH (2001) Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. Crop Sci 41:1317-1321

Ribeiro-Carvalho C, Guedes-Pinto H, Harrison G, Heslop-Harrison JS (1997) Wheat-rye chromosome translocation involving small terminal and intercalary rye chromosome segments in the Portuguese wheat landrace Barbela. Heredity 78: 539-546

Richardson MD, Chapman Jr GW, Hoveland CS, Bacon CW (1992) Sugar alcohols in endophyte-infected tall fescue under drought Crop Sci 32:1060-1061

Rong J, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, Delmonte TA, Ding X, Garza JJ, Marler BS, Park C, Pierce GJ, Rainey KM, Rastogi VK, Schulze SR, Trolinder NL, Wendel JF, Wilkins TA, Williams-Coplin TD, Wing RA, Wright RJ,

Zhao X, Zhu L, Paterson AH (2004) A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics 166:389-417

Saha MC, Cooper JD, Mian MAR, Chekhovskij K, May GD (2006) Tall fescue genomic SSR markers: development and transferability across multiple grass species. Theor Appl Genet 113:1449-1458

Saha MC, Mian R, Zwonitzer JC, Chekhovskiy K, Hopkins AA (2005) An SSR- and AFLP-based genetic linkage map of tall fescue (*Festuca arundinacea* Schreb.). Theor Appl Genet 110:323-336

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 10(6):297-304

Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edition, Cold Spring Harbor Laboratory Press

Sano Y (1990) The genic nature of gamete eliminator in rice. Genetics 125:183-191

SAS V9.1 (SAS Institute Inc. 2004**.** SAS OnlineDoc® 9.1.3**.** Cary, NC: SAS Institute Inc.).

Sathish P, Withana N, Biswas M, Bryant C, Templeton K, Al-Wahb M, Smith-Espinoza C, Roche JR, Elborough KM, Phillips JR (2007) Transcriptome analysis reveals season-specific *rbcS* gene expression profiles in diploid perennial ryegrass (*Lolium perenne* L.). Plant Biotechnol J 5(1):146 – 161

Sato S , Soga T, Nishioka T, Tomita M (2004) Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection. Plant J 40:151–163

Sax K (1923) Association of differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552-560

Schauer N, Fernie AR (2006) Plant metabolomics: towards biological function and mechanism. Trends Plant Sci 11(10):508-516

Schauer N, Zamir D, Fernie AR (2005) Metabolic profiling of leaves and fruit of wild species tomato: a survey of the *Solanum lycopersicum* complex. J Exp Bot 56:297–307

Schubert I (2007) Chromosome evolution. Curr Opin Plant Biol 10:109-115

Schubert I, Fransz PF, Fuchs J, Jong JH de (2001) Chromosome painting in plants. Meth Cell Sci 23:57-69

Schwarzacher T (2003) Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent *in situ* hybridization data in plants. J Expl Bot 54(380):11-23

Schwarzacher T, Heslop-Harrison JS (2000) Practical *in situ* hybridization, Springer, New York Berlin Heidelberg, pp 35

Schwarzacher T, Leitch AR, Bennett MD, Heslop-Harrison JS (1989) *In situ* localization of parental genomes in a wide hybrid. Ann Bot 64:315-324

Shull GH (1948) What is "heterosis"? Genetics 33:439-446

Sibov ST, De Souza jr. CL, Garcia AAF, Garcia AF, Silva AR, Mangolin CA, Benchimol LL, De Souza AP (2003) Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. Hereditas 139:96-106

Skøt L, Humphreys MO, Armstead I, Heywood S, Skot KP, Sanderson R, Thomas ID, Chorlton KH, Ruaraidh Sackville Hamilton N (2005) An association mapping approach to identity flowering time genes in natural populations of *Lolium perenne* (L.). Mol Breeding 15: 233-245

Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y, Ciavarrella M (2007) Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. Euphytica 154:401–408

Studer B, Asp T, Frei U, Hentrup S, Meally H, Guillard A, Barth S, Muylle H, Roldán-Ruiz I, Barre P, Koning-Boucoiran C, Uenk-Stunnenberg G, Dolstra O, Skøt L, Skøt KP, Turner LB, Humphreys MO, Kölliker R, Roulund N, Nielsen KK, Lübberstedt T (2008) Expressed sequence tag-derived microsatellite markers of perennial ryegrass (*Lolium perenne* L.). Molecular Breeding DOI 10.1007/s11032-007-9148-0

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823-839

Tanksley SD, Ganal MW, GB Martin (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. TIG 11(2):63-68

Taylor DR, Ingvarsson PK (2003) Common features of segregation distortion in plants and animals. Genetica 117:27-35

Thomas HM (1981) The Giemsa C-Band karyotypes of six *Lolium* species. Heredity 46(2):263-267

Thomas HM, Harper JA, Meredith MR, Morgan WG, Thomas ID, Timms E, King IP (1996) Comparison of ribosomal DNA sites in *Lolium* species by flourescence *in situ* hybridisation. Chromosome Res 4:486-490

Thomas H, Humphreys MO (1991) Progress and potential of interspecific hybrids of *Lolium* and *Festuca*. J Agric Sci Camb, 117: 1-8

Thomas HM, Morgan WG, Humphreys MW (2003) Designing grasses with a future – combining the attributes of *Lolium* and *Festuca.* Euphytica 133:19-26

Thornton B, Osborne SM, Paterson E, Cash P (2007) A proteomic and targeted metabolomic approach to investigate change in Lolium perenne roots when challenged with glycine. J Exp Bot 58(7):1581-1590

Thorogood D, Armstead IP, Turner LB, Humphreys MO, Hayward MD (2005) Identification and mode of action of self-compatibility loci in *Lolium perenne* L. Heredity 94:356-363

Thorogood D, Kaiser WJ, Jones JG, Armstead I (2002) Self-incompatibility in ryegrass 12. Genotyping and mapping the *S* and *Z* loci of *Lolium perenne* L. Heredity 88:385-390

Timmerman-Vaughan GM, Mills A, Whitfield C, Frew T, Butler R, Murray S, Lakeman M, McCallum J, Russell A, Wilson D (2005) Linkage mapping of QTL for seed yield, yield components, and developmental trait in *Pea*. Crop Sci 45:1336-1344

Turner LB, Cairns AJ, Armstead IP, Ashton J, Skøt K, Whittaker D, Humphreys MO (2006) Dissecting the regulation of fructan metabolism in perennial ryegrass (*Lolium perenne*) with quantitative trait locus mapping. New Phytol 169:45-58

Turner LB, Cairns AJ, Armstead IP, Thomas H, Humphreys MW, Humphreys MO (2008) Does fructan have a functional role in physiological traits? Investigation by quantitative trait locus mapping. New Phytol 179:765-775

Uchimiya H, Takahashi N (1973) Kinetics of heterosis in growth of the leaf blade in *Zea mays*. L Ann Bot 37:147-52

Van Ooijen JW, Boer MP Jansen RC, Maliepaard C (2002) Map QTL 4.0: software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen, The Netherlands

Van Ooijen JW, Vorrips RE (2001) Join Map® 3.1, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Vaughan DA, Balázs E, Heslop-Harrison JS (2007) From crop domestication to super-domestication. Ann Bot 100: 893–901

Verbruggen N, Hermans C (2008) Proline accumulation in plants: a review. Amino Acids DOI 10.1007/s00726-008-0061-6

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93(1):77-78

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23(21):4407-4414

Wagner RP, Mitchell HK (1965) Genetics and metabolism. Second edition. John Wiley and sons, Inc., New York, London, Sydney.

Warpeha KMF, Gilliland TJ, Capesius I (1998) An evaluation of rDNA variation in Lolium species (ryegrass). Genome 41:307-311

Warnke SE, Barker RE, Jung G, Sim S-C, Mian MAR, Saha MC, Brilman LA, Dupal MP, Forster JW (2004) Genetic linkage mapping of annual x perennial ryegrass population. Theor Appl Genet 109:294-304

Weckwerth W (2003) Metabolomics in system biology. Ann Rev Plant Biol 54:669-689

Weinig C, Dorn LA, Kane NC (2003) Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. Genetics, 165:321–329

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitary primers are useful as genetic markers. Nucleic Acids Res 18:6531-6535

Winter PC, Hickey GI, Fletcher HL (2002) Genetics, Second Edition, BIOS Scientific Publisher, Oxon, UK

Xiao J, Li J, Yuant L, Tanksley SD (1995) Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140:745-754

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F_2, backcross, double haploid and recombinant inbred population in rice (*Oryza sativa* L.). Mol Gen Genet 253:535-545

Yamada T, Foster JW, Humphreys MW, Takamizo T (2005) Genetics and molecular breeding in *Lolium*/*Festuca* grass species complex. Grassl Sci 51:89-106

Yamada T, Jones ES, Cogan NOI, Vecchies AC, Nomura T, Hisano H, Shimamoto Y, Smith KF, Hayward MD, Forster JW (2004) QTL Analysis of Morphological, Developmental, and Winter Hardiness-Associated Traits in Perennial Ryegrass. Crop Sci 44:925-935

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12:2473-2483

Yates F (1936) A new method of arranging variety trails involving a large number of varieties. J Agric Sci 26:424-455

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147(3):355-358.

Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136:1457-1468

Zhuang JY, Lin HX, Lu J, Qian HR, Hittalmani S, Huang N, Zheng KL (1997) Analysis of QTL × environment interaction for yield components and plant height in rice. Theor Appl Genet 95 : 799-808

Zwierzykowski Z, Kosmala A, Zwierzykowska E, Jones N, Jokś W, Bocianowski J (2006) Genome balance in six successive generations of the allotetraploid *Festuca pratensis* x *Lolium perenne*. Theor Appl Genet 113(3):539-547