Rubens Marschalek

Marker Assisted Selection for the development of intervarietal substitution lines in rapeseed (*Brassica napus* L.) and the estimation of QTL effects for glucosinolate content

Marker Assisted Selection for the development of intervarietal substitution lines in rapeseed (*Brassica napus* L.) and the estimation of QTL effects for glucosinolate content

Doctoral Dissertation submitted for the degree of
Doctor of Agricultural Sciences
of the Faculty of Agricultural Sciences
Georg-August-Universität zu Göttingen
Germany

by
Rubens Marschalek
born in Blumenau
Federal State of Santa Catarina
Brazil

Göttingen, July 2003

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar

1. Aufl. - Göttingen : Cuvillier, 2003
Zugl.: Göttingen, Univ., Diss., 2003
ISBN 3-89873-849-3

D7

Referent: Prof. Dr. H. C. Becker
Korreferent: Prof. Dr. Reiner Finkeldey
Tag der mündlichen Prüfung: 17 Juli 2003

D7

1st examiner: Prof. Dr. Heiko. C. Becker
2nt examiner: Prof. Dr. Reiner Finkeldey
Date of oral examination: July 17th 2003

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2003
Gedruckt auf säurefreiem Papier

ISBN 3-89873-849-3

Diese Arbeit ist dir gewidmet, meine liebe ***Julia***
(This Dissertation is dedicated to you, my dear Julia)

MORGENROT

Text und Musik: Herbert Grönemeyer, 1993
The use of Morgenrot´s Lyrics in this dissertation was kindly granted by
Mr. Herbert Grönemeyer (Grönland Records LTD – London)

ich bau dir ein bett aus rosen
die wände aus glanzpapier
das zimmer hat einen goldenen boden
und der regenbogen endet genau hier

alle fenster gehen nach süden
mit blick auf's glitzernde meer
ich glätte täglich die wogen
tauche versunkenen träumen hinterher
tausche sehnsucht gegen perlmutt
werde lachen und weinen um dich
hab immer 'n trumpf im ärmel, der sticht

ich werde dir die liebe versprechen
wenn dir das wasser bis zum halse steht
werde in zerrütteten zeiten
dir ein netz ausbreiten
stell mich mit in den sturm, bis der wind dreht

ich lüge dir das blaue vom himmel
rede dir jede tragik schön
verjag den kummer ein für alle mal
trauer kommt vor's tribunal
paß auf, daß die zeiten für dich gut stehn

alle wünsche gehen direkt in erfüllung
auch schon früher, wenn du willst
stehe tag und nacht zur verfügung
bin verschwiegen und halte still

das alltagsgrau kipp ich in den ausguß
zweifel ersticke ich im keim
für dich soll es sterntaler regnen
und du kannst eitler als der sonnenschein sein
laß die luft knistern für dich
bau dir traumschlösser ins morgenrot
böse geister werden aufgemischt

ich werde dir die liebe versprechen...

die zulassung kriegen die guten launen
die schlechten werden sofort entehrt
herzschmerz verfüttert an die friedenstauben
probleme unter den fliegenden teppich gekehrt

ich werde dir die liebe versprechen...

Für meine lieben Kinder,

Edgar und Sarah

Wir sollten viel öfter von ganzem Herzen etwas tun,
das kein Ziel verfolgt, keine Eile hat,
und sich nicht lohnen muß.

Jochen Mariss

I am really grateful for the kindly support of my dear country:

Santa Catarina State Agricultural Research and
Rural Extension Agency (SC – Brazil)

National Council for Scientific and Technological Development
(Ministry for Science and Technology of Brazil)

&

also to BMBF (Federal Ministry of Education and Research - Germany)
through the GABI programme (Grant no. 0312286C)

for supporting this research work.
Special thanks to the Institute of Agronomy and Plant Breeding

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AFLP	Amplified Fragment Length Polymorphism
BC	Backcross
cM	Centimorgan
DH Line	Double haploid line
GSL	Glucosinolates
HPLC	High Performance Liquid Chromatography
LG	Linkage Group
NIRS	Near infrared spectroscopy
MAS	Marker assisted selection
mg	Marker group (see Appendix)
PE	Phenotypic Effect
pers. com.	Personal communication
PV	Phenotypic Value
QTL	Quantitative Trait Loci
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RL	Restriction-Ligation
SSR	Simple Sequence Repeat

In the beginning God created
the heavens and the earth.
Genesis 1:1 (Bible)

1. INTRODUCTION

Our earth is about 4,600 million years old and about 3,500-4,000 million years ago the first organic molecules have got the ability to reproduce and so the first unicellular organisms have arisen (Welter-Schultes & Krätzner, 1999). Nevertheless, flowering plants only arose later, about 80 or 90 million years ago (Goth, 2002), during the cretaceous period. Looking through this period it is also possible to find the first fossils of many insect groups, modern mammal and bird groups. The Cretaceous was thus the time in which life, as it now exists on Earth, came together (Welter-Schultes & Krätzner, 1999; Museum of Paleontology – Uni. of California, 2003)

According to Allard (1999) there is an agreement that humanlike creatures had evolved in Africa by about 3 to 4 million years before the present (b.p.), although new reports (Ziegler, 2002; Gibbons, 2002) show that it may have happened earlier (6 to 7 million years b.p.). The genus *Homo* seems to be appeared around 2 million years ago (Leakey and Walker, 1997) being the ancestor of all modern people probably an early *Homo erectus* in Africa who lived at least 1.8 million years ago (O'Neil, 2003). Nearly all specialists on human origin agree that “anatomically” modern humans originated relatively recently, perhaps about 200,000 – 270,000 years b.p. (Allard, 1999; O'Neil, 2003; Bräuer, 2003), and that the modern human traits are not older as 100.000 years (Bräuer, 2003). Concerning the activities of humans however, the agricultural economies developed only about 15,000 to 10,000 years ago (Allard, 1999).

Humans started with quite simple agricultural activities, and at the same time they started to interfere definitively in the way some species evolved. That means, humans activities started to change dramatically the fate of other species, or even changed the evolution of such species. Plants were affected directly by humans since that their necessities imply in the search for some traits in the plants. In this way humans have, for a long time maybe unconsciously or empirically, used selection to obtain from the nature what they were looking for. History proves that they have succeeded and nowadays we are using different and quite extraordinary methods to still following persistently our aims.

Some of the important plant traits show a clear discrete variation. Such traits are called qualitative traits since different classes can be easily distinguished. Other characters show a continuous variation and are called quantitative traits.

It remained to Johannsen, Nilsson-Ehle, and East to provide, early in the twentieth century, convincing evidence that alleles of Mendel's particulate "factors" or "elements", now called genes, were responsible not only for discretely inherited characters but also for continuously varying characters. It is important to emphasize that both, continuous and sharply discontinuous variation are observed in many characters and this establishes that the distinction between qualitative and quantitative characters is not clear-cut (Allard, 1999). Generally it could be said that qualitative traits are controlled by one or few genes, while quantitative characters are typically controlled by many genes, usually termed quantitative trait loci (QTL), and they are usually much more affected by the environment than the qualitative ones.

To breed new and better plant genotypes we should improve our knowledge about the inheritance of desired traits. Regarding this aspect two points are interesting to consider nowadays: identification and localization of genes responsible for the characters and estimation of the genotypic effects of the alleles found at these loci. These are questions to be solved in all important agricultural species.

Rapeseed is one of the more important oil crops of the world and today, after a decline of over 2 million ha in the last two years, world rapeseed area has regained 1.3 million ha in 2002/03 (Fapri, 2002), giving a total of about 26 million ha, which is representing in 2002 10.1% of the harvested area from oilseed crops (FAO, 2003). It is estimated that the grown area will increase annually by about 0.3% (Fapri, 2002). The world production of rapeseed in 2002/03 is estimated to be around 31,604,000 t (Oil World, 2003). The most important oil seed crop nowadays is soybean but in the temperate zones rapeseed is more important. A better understanding of the genetics and inheritance of characters in rapeseed is therefore an important and desired aim of plant breeding science.

After extraction of the oil, which is the most valuable seed component, the resulting meal is an important protein source for animal feed. Since some components present in the meal are detrimental to animal nutrition, like glucosinolates, it still remains a breeding aim to develop varieties with lower glucosinolate content. Glucosinolates are nitrogen- and sulphur-containing natural plant products which have different biological effects, ranging from antimicrobial and cancer preventing function to inflammatory and goitrogenic activities or antithyroid activity (Wittstock and Halkier, 2002). The goitrogenic activities appear since the glucosinolates of rapeseed meal increase iodine requirement in animals (Schöne, 1999).

Regarding to the plant itself, the benefits of glucosinolates in the defences against insects and pathogens should not be neglected (Wittstock and Halkier, 2002).

The negative effects of the antinutrients in rapeseed meal can be reduced or eliminated by plant breeding, proper processing or a combination of breeding and processing (Jensen, 1999). With respect to glucosinolates, the wide spread growing of double low rapeseed (<25 µmol glucosinolates/g seed) has greatly reduced the negative effect of glucosinolates on animal performance and health. However, even the double low rapeseed varieties are only used in restricted amounts to monogastric animals (Sørensen, 1988, cited after Jensen, 1999). Therefore, in spite of having been a trait submitted to breeding efforts since long time, it is still a breeding aim to have varieties with low glucosinolate content. Older forms of rapeseed have a glucosinolate content above 80 µmol/g in the seed. Presently cultivated low-glucosinolate forms of oilseed rape have less that 25 µmol/g of seed. The first low-glucosinolate cultivar was the Canadian spring cultivar Tower released in 1974, which contained alleles for low glucosinolate content derived from the Polish fodder rape cultivar Bronowski (Campos de Quiroz and Mithen, 1996).

Six QTL for glucosinolate content have been identified until now in rapeseed, three of them have been reported to have an important effect while the other 3 showed only smaller effects (Uzunova *et al.*, 1995; Weißleder, 1996; Fischer and Ecke, 1997; Gül, 2002). Despite of the use of low glucosinolate genotypes in modern breeding programmes it seems to be common that genotypes appear, through segregation, which carry higher levels of glucosinolates than each of the parents (pers. com. H.C.Becker). Also Rücker and Röbbelen (1994) reported that even in crosses between genotypes with less than 20 µmol/g seed, sufficient genetic variation is present enabling selection. All this indicates that a more detailed study about the inheritance of glucosinolates in rapeseed would be very useful to better understand and control this character. In this way, the glucosinolate inheritance could be used as a model for better understanding of other traits.

To reach a better level of knowledge on the rapeseed genetics and more specifically about the inheritance and effects of the glucosinolate genes, as a model for other traits, some aspects should be considered, i.e. identification and localization of genes and estimation of the phenotypic effects of such genes using intervarietal substitution lines. An intervarietal substitution line carries a single segment of a donor genotype, in an otherwise genetic background of one recurrent genotype. A complementary set of intervarietal substitution lines represents the whole donor genome divided into a limited number of distinct segments, each carried by a different intervarietal substitution line in a common genetic background. Trying to improve information about the localization and effects of genes related

to the glucosinolates in rapeseed the specific aims of the present work consists respectively of two points:

1) **The development of substitution lines as a tool to study with more detail the QTL:** using the backcross method and molecular markers to develop a complementary set of intervarietal substitution lines. These lines will be used in mapping and studying different traits.

2) **Investigate glucosinolate inheritance:** estimation of the effect of the alleles of 6 QTL (according to the literature, responsible for the seed glucosinolate content in *Brassica napus* L.) based on field data.

1.1 The subject species: rapeseed (*Brassica napus* L.)

Rapeseed is classified as:
Order: Capparales
Family: Brassicaceae – Mustard family
Genus: Brassica L. – mustard
Species: *Brassica napus* L. – rapeseed
(USDA-Natural Resources Conservation Service, 2002)

Brassica napus L. is an allopolyploid with 19 pairs of chromosomes (n=19), derived from the A and C genomes of *B. rapa* and *B. oleracea*, respectively. It is an allopolyploid, which means an organism originated from a combination of two or more sets of chromosomes derived from different parental species. Brassica crops consists of six species, of which *B. nigra* (n=8; B genome), *B. oleraceae* (n=9; C genome) and *B. rapa* (n=10; A genome) are diploid monogenomic species. The other three, *B. carinata* (n=17), *B. juncea* (n=18) and *B. napus* (n=19; AACC genome) are species which evolved in nature through hybridization between any two of the diploid species. All three genomes are partially homologous; the genetic information in all three genomes is similar, only its organization and distribution on the chromosomes is different (Prakash, *et al.*, 1999 in Gómez-Campo, 1999).

Brassica napus has a genome with approximately 1,129 to 1,235 Mbp. (Prakash, *et al.*, 1999 in Gómez-Campo, 1999). More recent studies show *Brassica napus* genome having 1,127 Mbp (1C), corresponding to 1.15 pg (1C) (Bennett and Leitch, 1995; Bennett and Leitch, 2001). C-value measures the amount of DNA in the haploid genome of an organism (MayHospi.com, 2000).

Brassica napus is not known to occur truly wild in nature though it often occurs as an escape. The first reference to rapeseed (*B. napus*) was by Dodoens (1578, cited after Gómez-Campo 1999). As a crop it appeared around the year 1600. Cultivation of rapeseed

started with oleiferous varieties of *B. rapa*, but *Brassica napus* has progressively taken the supremacy in this role. This has occurred to a point that *B. napus*, only 400 years old as a cultivated species, has now climbed to the second or third place in economic importance among edible crops in several countries in temperate zones such as Canada and some Central European countries (Gómez-Campo, 1999), including Germany.

Two factors were mainly responsible for the expansion of rapeseed. The first is the increasing commercial yields which made it into a crop more attractive to cultivation. The second factor is the improvement of oil and meal qualities through breeding programmes, leading to the emergence of the double-zero (canola) varieties, which are almost free from erucic acid and have low glucosinolate content (Goffman, 2000; Gómez-Campo, 1999). The old varieties show a poor quality, with oil high in erucic acid (> 20-50%) and meal high in glucosinolates (>70 μmol/g). High levels of erucic acid are nutritionally undesirable, while meals containing high levels of glucosinolates are goitrogenic, causing palatability and nutritional problems in non-ruminant animals (Goffman, 2000).

The changes from high to zero erucic acid content of the oil and from high to low content of glucosinolates in the meal have opened almost unlimited avenues into the food and feed markets. Today rapeseed oil quality is in the top class compared to other major oilseeds (Becker *et al.*, 1999) and so, because of its qualities and multiple uses, rapeseed has been a crop intensively studied and breed.

1.2 Use and Development of Intervarietal Substitution Lines

1.2.1 Review

Traditionally, the genetic analysis of quantitative traits has been restricted to the statistical approaches of biometry (Mather and Jinks, 1982), which deal with the average effects of loci and provide little information about the number and location of QTL or the relative effects of allelic variation at specific quantitative trait loci. The absence of such kind of information do not trouble the breeding programmes too much, at least in the case of qualitative traits, but if the information is available, especially when more than one trait is considered, it can be interesting since then genes responsible for different traits are sometimes correlated because they are located on the same chromosome. In this case, monitoring the alleles with marker assisted selection (MAS) in breeding programmes could be better understood if the location of such genes is known.

Recently, the development of molecular markers has allowed the construction of dense genetic maps for many crops. The development of such genetic maps has led to the

development of new approaches to QTL mapping (Lander and Botstein, 1989). Most of these approaches are based on interval mapping, using the information of two linked markers to test for the presence of a QTL in the intervening interval and to estimate the QTL effects. A number of QTL mapping methods are currently available (simple interval mapping, SIM; composite interval mapping, CIM, etc) relying on statistical techniques to find relationships between molecular markers and QTL in segregating populations. However, all of these methods can suffer from lack of precision of QTL mapping owing to a number of factors. These factors include: heritability of the trait, the total number of QTL governing the trait, the distribution of these QTL over the genome, their interactions with each other, the random variation because of the environment and other unlinked genetic factors, the type and size of the population studied, the genome size, and the number and distribution of marker loci.

Therefore, interval mapping only gives a rough estimate of QTL position. Depending on the size of the segregating population, the total variance of the character analysed and the QTL effect, confidence intervals have been estimated to be in the range of several ten cM (Darvasi *et al.* 1993). These limitations make interval mapping unsuitable for a genome wide analysis of allelic variation or a fine mapping of QTL and the identification of the genes that underlie the observed QTL effects. Computer simulations, used to test the precision and reliability of QTL mapping, have shown that segregating populations often underestimate the number of QTL (Hyde *et al.*, 1995), result in large confidence intervals associated with QTL position, and overestimate the size of the QTL effects, particularly at small population sizes and low heritability (Thomas *et al.*, 2000; Burns *et al.*, 2003). Even in CIM, which is often preferable to then SIM, it is unsure whether the effects detected are real or due to over parameterisation (Thomas *et al.*, 2000). It is also difficult to distinguish between are two closely linked QTL, one affecting for example heading date and the other yield, or just one QTL with pleiotropic effect.

The reliable scoring of agronomic traits generally requires plot trials involving genetically uniform individuals within each plot. QTL analysis of such traits benefits from the use of double haploid or recombinant inbred lines because such populations allow the replication of an individual line, resulting in greater precision of QTL mapping. However, these specialised set of lines require a substantial investment in time in their construction, while studies have shown that large population sizes are needed to detect QTL which exhibit only small effects, the scale of which are typically outside the size of the normal experimental field design when plot trials are used (Burns *et al.*, 2003).

Many limitations of interval mapping can be overcome by using intervarietal substitution lines for the analysis of quantitative traits. In its ideal state, a substitution line will carry a single segment of a donor genotype, in an otherwise pure genetic background of the

recurrent genotype. However, studies using lines with more than one introgressed region are still in use (Rae *et al.*, 1999). A complementary set of intervarietal substitution lines ("substitution library", according to Burns *et al.*, 2003) represents the whole donor genome divided into a limited number of distinct segments, each carried by a different line in a common genetic background. The donor and the recurrent genome are two different varieties contrasting for the traits of interest. In experimental field trials, the performance of a substitution line can be compared to the recurrent parent and any significant differences found between the two lines can be attributed to QTL within the defined introgressed region (Burns *et al.*, 2003). An example is the research carried out by Ramsay *et al.* (1996) in *Brassica oleracea* where the aim was the precision mapping of quantitative trait loci using a set of substitution lines. Substitution lines are obtained using the backcross method, and in this sense, molecular markers have been proven to be very useful in improving backcross breeding schemes, particularly, markers allow breeders to estimate the genomic composition of individuals, and selection on markers can speed up the recipient genome recovery on noncarrier chromosomes (background selection) (Servin and Hospital, 2002).

The development of substitution lines in successive backcross generations consists of a number of criteria, including the inheritance of a large proportion of the recurrent parent genome and the inheritance of relatively large and unintercepted donor segments. In studies done by Cermakova *et al.* (1999) the number of introgressed segments has been steadily reduced from 5-10 present in BC_2 families (backcross families), through 1-6 present in BC_3 families, to a single segment in subsequent generations. Lines heterozygous for a single segment have been recovered after three or four rounds of backcrossing with marker-assisted selection followed by one round of self-pollination to "fix" the desired genotypes as homozygotes (Cermakova *et al.* 2001). However, plants homozygous for the desired segment may not be homozygous for all the rest of the genome if some regions are not properly covered by markers.

Segregating populations of even several hundred lines are likely to give misleading results when used for QTL analysis (Beavis *et al.*, 1994 and Hyne *et al.* 1995 cited after Howell *et al.* 1996). In contrast, substitution lines make it possible to assay the whole genome with increased confidence using a small number of highly replicated lines (Law *et al.*, 1987, cited after Howell *et al.* 1996). Homozygous substitution lines can be multiplied indefinitely to improve precision in QTL mapping experiments, eliminate the genetic "noise" of segregating and unlinked loci (Burns *et al.*, 2003).

Interactions between donor alleles are limited to those between genes on the same homozygous substituted segment, simplifying calculations of the significance and magnitude of the mean effects of each segment. Furthermore, since the entire donor genome can be

represented by relatively few lines (around 100), substitution lines offer the opportunity for large-scale replication, increasing the power of detection for QTL and allowing for a precise analysis of “QTL x environment” interactions. In addition, individual lines can be analysed independently from the whole set. If substitution lines want to be used to fine mapping, more lines will be necessary, in this way, substitution lines are also suitable for a high-resolution mapping of QTL because donor segments can be subdivided by recombination with the recurrent parent in additional backcross generations. The precision of this mapping is limited only by the availability of markers in the region analysed (Paterson *et al.*, 1990) and in the case of fine mapping additional lines have to be developed for the region of interest.

In tomato a set of 50 substitution lines, called introgression lines by the authors, covering the whole donor genome was produced by marker assisted selection in a backcross approach from a cross between a wild tomato, *Lycopersicon pennellii*, and *L. esculentum* with the wild species as donor parent (Eshed and Zamir 1994, 1995). With this set a total of 104 QTL could be mapped for six traits in a comparatively small field trial using plots with only 6 homozygous plants (for each of the 49 introgression lines) and 12 plants/plot for the F_1 hybrid between the 49 lines and the line A8 (Eshed and Zamir, 1995). The number of QTL that could be mapped with this approach was approximately twice the number of what had been mapped by interval mapping in earlier studies. One of the regions containing a QTL for fruit mass was analysed in greater detail, using new lines with smaller segments that were produced by an additional backcross step. The results of this analysis indicated the mapped QTL to be comprised of 3 linked genes (Eshed and Zamir, 1995). In an additional study, crosses between selected substitution lines were used to analyse epistatic interactions between mapped QTL (Eshed and Zamir, 1996).

Sets of intervarietal substitution lines have also been developed in *Brassica oleracea* (Ramsay *et al.*, 1996), rapeseed (Howell *et al.*, 1996; Cermakova *et al.*, 1999), and rice (Kubo *et al.*, 1999). The results from the work in rapeseed indicate that four backcross generations will be required to develop a full set of substitution lines covering all of the rapeseed genome.

Nevertheless, the development of “intervarietal substitution lines” has some disadvantages, being the long time needed to develop such lines compared to the development of DH lines. With DH lines it is in a short time possible not only to get the lines, but to get enough seeds from them to realize field experiments. With the substitution lines all this needs still more time, and the MAS with many plants has to be carried out in each generation, which means around 3 or 4 rounds of selection using markers, whereas when using DH lines only these lines need to be analysed by markers.

It is however necessary to distinguish the concept of substitution lines described above from the traditional concept of substitution lines used in cytogenetic, in which substitution lines mean genotypes that carry usually one pair of chromosomes from a different parent (often from a different related species) in a common background of chromosomes. This term is still in use nowadays (Odenbach, 1997; Clua *et al.*, 2002). A similar approach to the cytogenetic concept is the one used in animals studies in which the term “chromosome substitution strains” appear. These strains are also used in QTL mapping and in this case each of these strains has a single chromosome from the donor strain (mouse) substituting for the corresponding chromosome in the host strain (Nadeau *et al.*, 2000).

For the term “substitution lines”, as used in the present study, some synonyms are found in the literature, like “Recombinant Chromosome Substitution Lines” (RCSLs) which are defined as an overlapping set of nearly isogenic lines in a common genetic background (Matus *et al.*, 2003 in press; Matus *et al.*, 1998; Thomas *et al.*, 2000). Even a variation of this term could be found as “recombinant substitution lines” (RSLs), which was used by Rousset *et al.* (2001) in wheat; “substitution lines” as used by Yano and Yamamoto (1997) in rice; or “single segment substitution lines (SSSLs)” (Zhang *et al.*, 2002). Also the term “chromosomal segment substitution lines (CSSLs)” appear in the literature, sometimes referring to lines obtained from the same species (Yano *et al.*, 2002) and sometimes obtained from different species (Doi *et al.*, 2002). A similar term used is “chromosome segment substitution lines” (CSSLs) in rice (Kubo *et al.*, 1999; Miura *et al.*, 2002).

Similar to this nomenclature is the term used by Shah *et al.* (1999): “Recombinant Inbred Chromosome Lines”. These RICLs are, according to the authors, also called “Recombinant Substitution Lines”, and whereas the “Chromosome Substitution Lines” allow the effect of a whole chromosome to be studied, the RICLs allow chromosomal effects to be partitioned into chromosomal-segment effects.

In the same way, these concepts do not differ from the term “Near Isogenic Lines” (NIL) (Paterson, 1996; Han *et al.*, 1999; Howell *et al.*, 1996) and “Near Isogenic Introgression Line” (NIIL) (Shen *et al.*, 1999). The difference that supposes to exist between the term “substitution lines” and NILs is that the substitution lines should cover all the donor genome, whereas the NILs do not necessary cover it. Indeed, even Howell *et al.* (1996) call the substitution lines created by them as “material carrying small defined homozygous substitutions (i.e., near isogenic lines)”. Another method that originate NIL in a similar way, that means, using backcrosses, is the “Advanced Backcross QTL analysis”, which is a method for the simultaneous discovery and transfer of valuable QTL from un-adapted

germplasm into elite breeding lines utilizing BC generations, followed by selection of elite NIL for variety production (Tanksley and Nelson, 1996).

All these lines are developed by backcrossing, e.g., crossing a donor genotype carrying a specific trait of interest, to a “recipient parent” with generally desirable attributes (Paterson, 1996; Han *et al.*, 1999). By recurrently selecting for the trait of interest, and repeatedly crossing to the recipient, donor chromatin is progressively eliminated except for a small amount which is closely linked to the trait under selection. By comparing the backcross-derived stock to the original recurrent parent, one can determine the likely position of the target gene simply by identifying DNA markers which reveal the donor allele in the backcross-derived stock (Paterson, 1996).

This short review shows the many possibilities of names given in principle to the same thing and should advise about the importance of reach a consensus in the scientific nomenclature as soon as possible, not only in this case, but also for sure in many others.

1.2.2 The main tool: AFLP markers (review)

One of the aims of this work was to get a number of intervarietal substitution lines which cover all the donor genome despite that each line will carry only a little segment of the donor DNA. To carry out the genotyping and selection processes the marker technology chosen was the AFLP since it proves to be an appropriate method for this purpose (Savelkoul *et al.*, 1999). The AFLP markers were applied in the selection using the backcross method.

Amplified fragment-length polymorphism (AFLP) analysis is currently the most powerful and efficient technique for the generation of large numbers of anonymous DNA markers in plant and animal genomes (Vos and Kuiper, 1997; Myburg *et al.*, 2001). The efficiency and relatively low cost of AFLP analysis have enabled *de novo* genetic map construction in many species, saturation of existing linkage maps, and high resolution mapping of genomic regions of interest. In addition, this marker system has provided a fast, low-cost approach for studying genetic diversity and obtaining molecular phylogenies (Myburg *et al.*, 2001).

Amplified fragment-length polymorphism (AFLP) or its fluorescent version (sometimes called “fAFLP”) is a polymerase chain reaction (PCR)-based fingerprinting technology. In its most basic form, AFLP involves the restriction of genomic DNA. This digestion is carried out with two restriction enzymes, one with an average (sometimes called “rare”) cutting frequency (like *Eco*RI), and a second one with a higher cutting frequency (like *Mse*I). The next step is the ligation of complementary double-stranded adapters to the ends

of the restricted fragment (restriction sites) generating a template suitable for PCR amplification. These double-stranded oligonucleotide adapters are designed in such a way that the initial restriction site is not restored after ligation, which allows simultaneous restriction and ligation, while religated fragments are cleaved again (Savelkoul *et al.*, 1999).

Depending on genome characteristics, the restriction-ligation procedure may generate thousands of adapted fragments. For sufficient fragment resolution after denaturing polyacrylamide gel electrophoresis, the number of differing fragments must be reduced, otherwise the pattern or final picture would be too complex to be successfully analysed. This is accomplished by PCR amplification of a subset of the original fragments in which adapter-specific primers that have at their 3´ends an extension of one to three nucleotides are used (Savelkoul *et al.*, 1999; Berres, 2002).

Two rounds of highly stringent (because PCR amplifications are performed with high stringency, primers differing by only a single nucleotide base amplify a different subset of adapted fragments) PCR amplification need to be done: the first one is called preselective (preamplification), with a single-nucleotide addition (which amplifies 1 out of 4 ligated fragments); the second one is called selective (amplification) and is performed with possibly multiple nucleotide additions with primers complementary to the adapter sequences (usually three selective nucleotides in both primers are used and amplify 1 out of 4096 ligated fragments). The PCR primer which spans the average-frequency restriction site is labelled with a fluorescent dye. Only fragments containing a priming site complementary to the labelled primer will be visualized (Savelkoul *et al.*, 1999; Berres, 2002).

These amplified fragments are visualized on denaturing polyacrylamide gels either through auto radiographic or fluorescence methodologies and a highly informative pattern of 40 to 200 bands is obtained. The availability of many different restriction enzymes and corresponding primer combinations provides a great deal of flexibility, enabling the direct manipulation of AFLP fragment generation for defined applications (e.g. polymorphism screening, QTL analysis, genetic mapping) (Savelkoul *et al.*, 1999; Berres, 2002). The number of selective nucleotides required for optimum fragment distribution is highly dependent on the complexity of the target DNA which varies greatly among classes of organisms. (Berres, 2002). By using combinations of primers with different selective nucleotides, a series of AFLP amplifications will sample loci from a large fraction of the genome. With the ability to control the number of selectively amplified fragments, an optimal number of fragments may be generated thereby avoiding complications associated with DNA smears or unacceptable levels of fragment co migration (Berres, 2002).

The AFLP technique differs importantly from other random fingerprinting techniques and also from RFLP or microsatellites, by its robustness, resolution, reproducibility and time

efficiency (Vos and Kuiper, 1997; Savelkoul *et al.*, 1999; Berres, 2002). Probably the single greatest advantage of the AFLP technology is its sensitivity to polymorphism detection at the total-genome level. With all of these assets, AFLP markers are fast becoming a molecular standard for investigations ranging from systematics to population genetics (Berres, 2002). AFLP markers are usually considered and analysed as being dominant markers (Ferreira and Grattapaglia, 1998). A review in which the AFLP analysis is compared to other methods with respect to reproducibility and robustness, discriminatory power, and operational aspects was published by Savelkoul *et al.* (1999). Since relatively small amounts of DNA are digested and detection of AFLP fragments does not depend on hybridization, partial digestion and faint patterns, which are sources of irreproducibility with RFLP genotyping, can easily be avoided. Furthermore, the possibility of using stringent PCR annealing temperatures renders the AFLP analysis method more reproducible and robust than RAPD analysis. With respect to the discriminatory power, according to many studies reviewed by Savelkoul *et al.* (1999), AFLP analysis has been found to be more informative than RAPD analysis, RFLP analysis and SSR analysis (Savelkoul *et al.*, 1999).

A unique feature of AFLP analysis is that it can be adapted to the DNA of any organism via the use of selective nucleotides. However, a basic limitation of AFLP analysis (and other genomic typing procedures) is that the organism to be typed must be isolated, since DNA from other sources disturbs the AFLP pattern. The operational aspects of AFLP can be described as advantageous. Since AFLP analysis is a PCR-based assay, only a small amount of DNA is required, however this DNA must be pure and double stranded, but its exact concentration seems to be less critical than it is for RAPD analysis. Oligonucleotide adapters and primers can be custom synthesized and besides, commercial kits for AFLP are also available. The subsequent digestions, ligations, amplifications and electrophoresis are relatively time-consuming, but with fluorescent primers and analysis on an automatic sequencer, AFLP analysis can be performed within 24 h (Savelkoul *et al.*, 1999).

1.3 Glucosinolates

1.3.1 General comments

Glucosinolates are nitrogen- and sulfur-containing natural plant components found mainly in several plant families of the order Capparales (Malvaceae, Bombacaceae, Sterculiaceae, Cucurbitaceae, Brassicaceae, Ericaceae, Ebenaceae, Primulaceae) (Wittstock and Halkier, 2002), of which the most significant from an agricultural and culinary

viewpoint are the Brassicaceae (Wallsgrove and Bennett, 1995) and the Cucurbitaceae. Glucosinolates are secondary metabolites (Thangstad *et al.*, 2001).

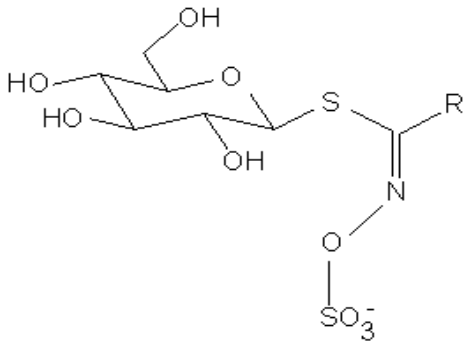


Fig. 1 Core structure of glucosinolates (R = Alkenyl, Aromatic or Indolyl)

To date, more than 120 different glucosinolates have been detected in hundreds of plant species in the order Capparales and in the Genus *Drypetes* (Euphorbiales) (Wittstock and Halkier, 2002). All glucosinolates have a common core structure (Fig. 1) comprising the R-C-N structure derived from an amino acid, a glucose moiety attached via a thioester link, and the sulphate group linked to the nitrogen (Wittstock and Halkier, 2002; Wallsgrove and Bennett, 1995).

Glucosinolates and the thioglucosidases (also known as myrosinases) are the main components of the glucosinolate-myrosinase system, which provides plants with an effective defence against generalist herbivores and probably also against pathogens (Bones and Rossiter, 1996; Raybould and Moyes, 2001; Kessler and Bakdwin, 2002; Zhao and Meng, 2003). Glucosinolates are probably also used by the plant as sulphur source, since there are already some indications that double-zero oilseed rape is more sensitive to sulphur deficiency than single-zero plants (Bones and Rossiter, 1996) A model has been proposed by Lüthy and Matile (1984, cited after Thangstad *et al.*, 2001) for the co-localization of glucosinolates and myrosinase, known as the “mustard oil bomb”. According to this model, myrosinase is separated from its substrate, the glusosinolates, by the vacuolar membrane. Whereas in “the mustard oil bomb” model the myrosin cells were not considered, the myrosinase was later shown to be localized in vacuoles (myrosin grains) of the myrosin cells (studies from 1990 to 1998 cited by Thangstad *et al.*, 2001). Kelly *et al.* (1998) showed glucosinolates to be localized in all cells of cotyledons except myrosin cells.

Glucosinolates occur throughout the tissues of all plant organs, whereas myrosinases are localized in scattered myrosin cells, which seem to be glucosinolate free

(Kelly *et al.,* 1998, cited after Wittstock and Halkier, 2002). Nevertheless, Kelly *et al.* (1998) have reported that some myrosinase could be observed in vacuoles of non-myrosin cells of cotyledons. Thus, the important thing is that the two components of the system are separated until autolysis or tissue damage brings them into contact. The precise localization of glucosinolates is not known (and has been hampered by their water solubility, mobility, hydrolysis by myrosinase, and lack of specific markers) but they have been reported to be stored in vacuoles (Thangstad *et al.*, 2001; NTNU Plant Genetics Lab, 2003).

Compartmentalization of the glucosinolate-myrosinase system has been shown recently in *Arabidopsis* by identification of sulfur-rich cells (S-cells) between the phloem and the endodermis of the flower stalk, which presumably contain high concentrations of glucosinolates, and by localizing myrosinase in the neighbouring cells (Wittstock and Halkier, 2002). Nevertheless, glucosinolates are found in all parts of the plant and up to fifteen different glucosinolates have been found in the same plant. Generally, levels in the seed are high (up to ten per cent of the dry weight), whereas the levels in the leaf, stem and root are approximately five to ten times lower (Cloissais-Besnard and Larher, 1991). Concentrations differ according to tissue type, physiological age, plant health and nutrition (NTNU Plant Genetics Lab, 2003).

Considering the compartmentalization of glucosinolates, they are non-toxic but, upon tissue damage (e.g. by cutting or chewing) they come in contact with myrosinases and are hydrolysed into unstable aglucones, which rearrange into a range of biologically active and sometimes toxic compounds, typically isothiocyanates, nitriles and other. The hydrolysis products are responsible for the characteristic flavour of brassicaceous vegetables. They have different biological effects, ranging from antimicrobial and cancer preventing to inflammatory and goitrogenic activities (Wittstock and Halkier, 2002). Goitrogens are foods which suppress thyroid function. Presence of glucosinolates in the diets leads to hyperthyroidism in animals. It also reduces the level of circulating thyroid hormones and alters the ratio between thyroxine (T4) and triiotathyronine (T3) in blood (Zeb, 1998).

The first toxic effects of isothiocyanates and other hydrolytic products from glucosinolates that were identified were goitre and a general inhibition of iodine uptake by the thyroid. Numerous studies have indicated that the hydrolytic products of at least three glucosinolates, 4-methyl-sulfinylbutyl (glucoraphanin), 2-phenylethyl (gluconasturtiin) and 3-indolylmethyl (glucobrassicin) have anticarcinogenic activity. Indole-3-carbinol, a metabolite of glucobrassicin, has shown inhibitory effects in studies of human breast and ovarian cancers. S-methyl cysteine sulfoxide, another sulfur-containing phytochemical found in Brassica, and its metabolite methyl methane thiosulfinate were shown to inhibit chemically-induced genotoxicity in mice. Thus, the cancer chemopreventive effects of Brassica

vegetables that have been shown in human and animal studies may be due to the presence of both types of sulfur-containing phytochemicals (i.e. certain glucosinolates and S-methyl cysteine sulfoxide) (Stoewsand, 1995).

The dual roles of glucosinolates and their degradation products as deterrents against generalist herbivores and as attractants to insects that are specialized feeders on glucosinolate-containing plants have also been reported (Lambrix *et al.*, 2001; Wittstock and Halkier, 2002). Nevertheless, some insects can use plant secondary metabolites in defence by using the same methodology as plants, i.e., compartmentalization. The specialist brassica feeders, *Brevicorne brassicae* (cabbage aphid) and *Lipaphi erysimi* (turnipo aphid) can sequester glucosinolates from their host plants, yet avoid the generation of toxic degradation products by compartmentalizing myrosinase into crystalline microbodies. In that way, maybe insects death or damage by predators cause disruption of compartmentalized myrosinase, which results in the release of isothiocyanate that acts as a synergist for the alarm pheromone E-beta-farnesene (Bridges *et al.*, 2002). All this agrees with evolutionary theory stating that specialist insects may overcome host plant chemical defenses, whereas generalists will be sensitive to these same defenses, as also observed by Kliebenstein *et al.* (2002) studying two different insects in *Arabidopsis thaliana* concerning glucosinolates.

One of the major objectives of oil-seed rape breeding programmes has been to reduce the level of glucosinolates in seeds. Following oil extraction from rapeseed, the quality of the resultant meal for livestock is reduced by the presence of these sulphur-containing glycosides. The major component, 2-hydroxy-3-butenyl glucosinolate (progoitrin) degrades to goitrogenic products while other glucosinolates such as 3-butenyl and 4-pentenyl produce isothiocyanates which reduce meal palatability (Campos de Quiroz and Mithen, 1996).

Depending on the nature of the glucosinolates (GSL), temperature and pH, various compounds are formed (Huisman and Tolman, 1992, Etienne and Dourmad, 1994 cited after McGee, 1998). These compounds are toxic, in that they can adversely effect feed intake, feed conversion efficiency and growth performance of animals. Evidence indicates that diet palatability can be adversely affected by the inclusion of rapeseed meal and that the response, while related to the level of GLS, is variable depending on species, age and growth status. Ruminants seem less sensitive than non-ruminants in accepting feeds containing rapeseed meal (Hill, 1991, cited after McGee, 1998; Derycke *et al.*, 1999) and pigs seem to be more sensitive than poultry to levels of GLS in rapeseed (McGee, 1998). Younger animals, piglets, calves and chicks seem to be more severely affected than older ones (McGee, 1998).

1.3.2 The Glucosinolate genes (QTL) and the effects of it's alleles

The sequences of the first genes involved in the biosynthesis of glucosinolates were published in 2000. To date, ten genes from *Arabidopsis* have been identified that control reactions of the three major phases of glucosinolates biosynthesis, and the corresponding enzymes have been characterized (Wittstock and Halkier, 2002). In addition, the identification and characterization of an epithiospecifier protein from Arabidopsis has added important information to our knowledge of glucosinolate catabolism by myrosinases and the accompanying proteins, and of the role of glucosinolates in plant-herbivore interactions (reviewed in Wittstock and Halkier, 2002).

In rapeseed, six QTL for glucosinolate content have been mapped in previous studies (Uzunova *et al.*, 1995; Weißleder, K., 1996; Fischer and Ecke, 1997; Gül, M.K., 2002) in a segregating doubled haploid population derived from a cross between the old cultivar 'Mansholt' and the canola quality winter rapeseed variety 'Samourai'. (see Results, Tab. 7). Three of these QTL have been reported to have major effects while the remaining 3 showed only minor effects. A similar research was done by Toroser *et al.* (1995) using 99 DH lines obtained by the cross between the cultivars Stellar (low glucosinolate) and Major (high glucosinolate). In this study two major loci, with the largest influence on total seed aliphatic-glucosinolates, were mapped onto linkage group (LG) 20 and LG 1. This two important QTL identified in this case were the same ones identified by Uzunova *et al.*(1995), that means, equivalent to LG 16 and LG 18, and were even the same identified by Campos de Quiroz and Mithen (1996) in two other different crosses. In other Brassica species, like *Brassica juncea,* the total glucosinolate content was found to be under control of seven genes (Sodhi e*t al.*, 2002)

Another study reported by Cermakova *et al.* (1999) was done by crossing the double low quality winter rapeseed variety 'Tapidor' with the winter variety 'Bienvenu'. A marker assisted backcrossing were carried out using the "TapDH1" (double haploid derivate from 'Tapidor') as the recurrent parent and RFLP as marker technology. In this case, also three loci influencing the amount of seed glucosinolates have been mapped by QTL analysis to linkage groups N9, N12 and N19, controlling together 90% of the variation of glucosinolates levels. Other studies using different crosses in *Brassica napus* also confirm that at least three important QTL control the glucosinolate content (Howell *et al.*, 2003; Zhao and Meng, 2003).

Few genes controlling a great amount of the phenotypic variation of glucosinolate content were also observed in the model plant Arabidopsis, where a single locus appeared to control a majority (nearly 75%) of the observed quantitative variation (Kessler and Baldwin, 2002).

Magrath *et al.* (1993) have described six unlinked loci which determine the aliphatic glucosinolate profile of *B. napus.* Many other studies have been carried out to investigate the biochemical genetics of glucosinolates in rapeseed (Hall *et al.*, 2001; Giamoustaris and Mithen, 1996; Magrath *et al.,* 1994).

Another point to stress out is that usually in breeding programmes the glucosinolate content is considered to be sporophytic determined. However Clossais-Besnard and Larher (1991) show that the absolute quantity of glucosinolates in the shoots is not sufficient to explain the final content in the seeds. In addition, they report that the presence in seeds of specific compounds such as but-3-enyl (gluconapin) and pent-4-enylglucosinolates (glucobrassicanapin) suggest an *in situ* synthesis, and this could occur either in pod shells or in the seed, although Magrath and Mithen (1993) cited after Parkin *et al*. (1994) have shown that there is no *de novo* aliphatic glucosinolate biosynthesis in seeds, and that all aliphatic glucosinolates in seeds are derived from the maternal pod tissue. De March *et al.* (1989, cited after Clossais-Besnard and Larher, 1991) reported that an increase of glucosinolates in the seeds was temporarily associated with a decrease in pod shells, however, the relationship was not quantitative, and transport from other parts or synthesis in the seed seemed possible. Because of the existence of seed-specific glucosinolates it is suggested that vegetative parts mainly provide precursors and that the final steps for glucosinolate synthesis occur in the seed (Clossais-Besnard and Larher, 1991). If this trait really behaves as described, such pollen effect should maybe possible to assume, but it was not what some other studies have shown (pers. com. Schulz).

In the present study a BC_2 population derived from the same cross with Samourai as recurrent parent was genetically characterized using markers. Based on the marker information of BC_2 plants, BC_3 families segregating for only one of the six QTL were selected and grown in field trials and seed glucosinolate content was evaluated.

You can not acquire experience by making experiments.
You can not create experience. You must undergo it.
Albert Camus (1913 - 1960), French existentialist philosopher

2. MATERIAL AND METHODS

2.1 Material

2.1.1 Plant Material

2.1.1.1 Plant Material for the development of intervarietal substitution lines

A segregating doubled haploid population of 151 lines derived from a cross between doubled haploid lines of the winter rapeseed varieties 'Mansholt's Hamburger Raps' (DH 5.1 and DH 5.2) and 'Samourai' (DH 11.4) was the starting point for this research work (Uzunova *et al.*, 1995). Mansholt is an old cultivar with a high content of both erucic acid and glucosinolates. Samourai is a new French variety of canola quality. The doubled haploid lines were derived from microspores of F_1 plants (Fig. 2). All of the lines have been characterised for 250 mapped RFLP and RAPD markers and a subset of 96 lines was also characterised for 213 AFLP markers (19 primer combinations) (pers. com. Ecke).

Based on the marker information from the DH Lines, 10 DH lines could be selected so that they represent together the whole genome of the donor parent (Mansholt). Those lines were crossed with Samourai (recurrent parent) to generate plants equivalent to BC_1. Therefore 10 BC_1 genotypes could be originated. Each of these ten BC_1 genotypes was crossed with Samourai in December 1999 in the greenhouse (Göttingen) to produce BC_2 seeds. The BC_2 seeds were harvested in February 2000.

In July 2000 30 BC_2 seeds from each of the selected BC_1 plants were planted in the greenhouse and were grown under illumination of Phillips-400 W-Lamps (day/night periods of 16/8 hours) and 18°C. The ten families have got the following numbers: 1036, 1053, 1089, 1097, 2034, 2078, 2080, 2098, 2123, and 2127. 15-21 days after sowing, three leaf samples were taken from each plant. Each sample consists of 0.1 g leaf material, and was immediately put in a 1.5 ml tube (Micro Test Tubes Safe-Lock 1.5 ml) which had been autoclaved and labelled with the plant number.

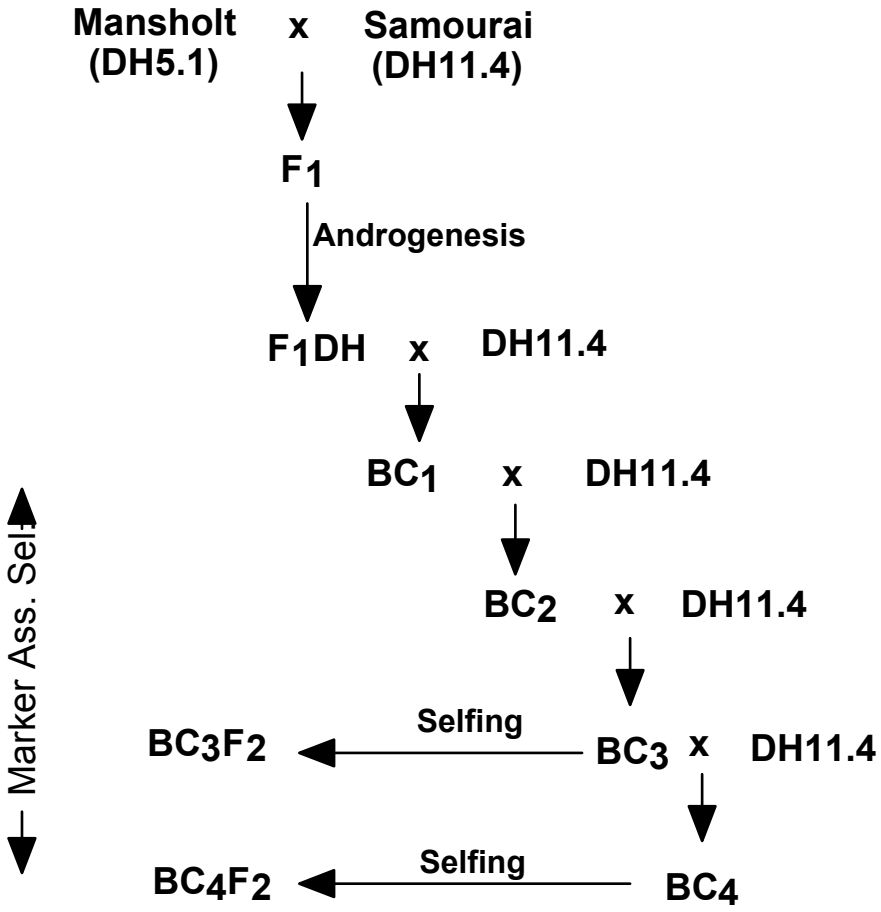


Fig. 2 Scheme for the production of intervarietal substitution lines.

After storaged in the tubes, the leaf material was frozen in liquid nitrogen (-196°C) and then kept at -30°C. In the same way, leaf material was taken from Mansholt's plants (DH 5.1 and DH 5.2) and Samourai's ones (DH 11.4). Approximately 4 weeks after sowing, the plants were put into a container (4°C) for vernalization during approximately 9 weeks.

The backcrosses to generate the BC_3 population were performed in November 2000. It was necessary to backcross all the BC_2 plants to Samourai because at this time the AFLP method was not optimized at the institute and the MAS could not be finished before flowering. From January-October 2001 the BC_2 plants were analysed by molecular markers, mainly AFLP. The marker analysis was done in 2001 and with the marker information 20 out of 300 BC_2 genotypes (plants) could be selected that cover the whole donor genome.

The BC_3 plants were grown under greenhouse conditions under the same conditions as the BC_2 plants. DNA from 300 BC_3 was isolated (Dec 2001), corresponding to 15 plants per each BC_2 original selected genotype (20 genotypes).

The backcrosses to generate the BC_4 population were performed in April 2002. 300 plants were crossed using Samourai (DH 11.4) as recurrent parent. As it was expected to get the first genotypes carrying only one segment from the donor genome in the BC_3 (Cermakova *et al.*, 1999), all the plants were also selfed, so that also BC_3F_2 seeds could be generated. In this way, genotypes homozygous for one donor segment could have a chance to appear. Such genotypes, if they exist, will constitute the aimed substitution lines. BC_3F_2 genotypes offer the opportunity to find genotypes homozygous for one donor segment, which would represent substitution lines their selves.

After getting the marker information, selection was performed based on two criteria: (a) the selected genotypes should have the smallest possible number of donor fragments; (b) this segments should be as long as possible. For the final graphical representation of the selected genotypes in each generation, BC_2 and BC_3 respectively, the GGT software was used (Van Berloo, 1999), which was developed to enable representation of molecular marker data by simple chromosome drawings in several ways.

2.1.1.2 Plant Material for the field trials

The field trials were based on BC_3 seeds originated from the backcrosses above described. The 300 BC_2 plants backcrossed to Samourai have originated 300 BC_3 families, which were tested on two field experiments. The 300 BC_2 plants have been analysed by molecular markers using 114 map positions (See Results, Tab. 1). For details see also chapter 2.2.2.

2.1.2 Chemicals, Enzymes and Oligonucleotides

A list of companies and respective products used is found in Appendix 9.1.

2.1.3 Laboratory Equipment, Material and Software

A list of companies and respective products used is found in Appendix 9.2.

2.2 Methodology

2.2.1 Genotyping and Selection using Molecular Markers

2.2.1.1 DNA Isolation and measurement of the DNA concentration

DNA was isolated from 15-21 days old seedlings from which leaf samples were taken. Isolation was done using the Nucleon PhytoPure Kit for small samples (Amersham, 1997), i.e., 0.1 g/sample, according to the Nucleon Extraction & Purification Protocols (Appendix 9.3.1).

After adding TE buffer, the DNA was kept at 4°C for approximately one week to achieve full resuspension, than the DNA concentration was measured by a Fluorescent DNA Quantification method by using a fluorometer with the fluorochrome dye Hoechst 33258 (bisbenzimide). This is a sensitive and simple method for quantifying DNA. The Hoechst's dye binds to the minor groove of DNA with a preference for AT sequences. Upon binding to DNA, the efficiency and the maximum wavelength of the fluorescence shifts. The fluorescence changes was measured using an excitation wavelength of 360 nm (optical filter excitation, EX 360/40 – 340-380 nm) and an emission wavelength of 460 nm (optical filter emission, EM 460/10 – 455-485 nm). The dye preferential binds to DNA in presence of high salt and neutral pH and allows the DNA to be quantified in the presence of RNA, proteins (< 100 μg/ml), nucleotides, and diluted buffer reagents (details see in Appendix 9.3.2). All the samples were then diluted to a standard concentration of 50 ng/μl; samples with lower concentration were diluted to 25 ng/μl.

The DNA amount that could be extracted varied from 810 ng to 29 μg, showing 7 μg as the mean value in the BC_2 population. The DNA quality was verified by gel electrophoresis, performed with a 1% agarose gel prepared with TAE buffer (40 mM Tris-acetat, 1 mM EDTA, pH 8.0). Each sample used for the gel was prepared taking 10μl from the DNA and adding 10 μl loading buffer (30%). The loading buffer stock solution (100%) was composed by 0.25% bromophenol blue, 0.25% xylene cyanol II, 40% saccharose, and water.

The electrophoresis was conducted under 30-100 V in 1x TAE-buffer. After this step, the gel was put into an ethidium bromide solution (1.0 mg/l H_2O) for 20 minutes to be stained followed for at least 10 minutes by incubation in water to remove excess of ethidium bromide. DNA was visualised under UV light (λ=254nm) and a photo was taken from the gel with a

Polaroid MP-4 Land Camera (Polaroid film 667) or with a video camera. If the sample showed a clearly defined band it was considered representing good DNA quality.

The good quality of the DNA samples was verified by gel electrophoresis (Fig. 3 as an example) and could even be confirmed by looking at the significantly low number of observed failure rate (less than 2%) during the restriction-ligation step of the AFLP procedure. In Fig. 3 the gel performed to verify the DNA quality shows clear bands without smear indicating that the DNA is still present and has good quality.

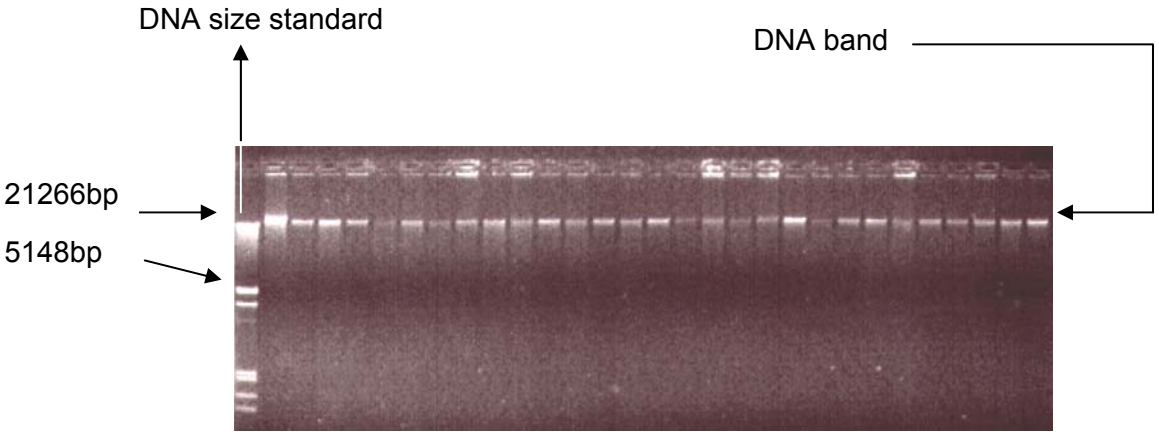


Fig. 3 Gel electrophoresis with total DNA from BC_2 plants of family 1097. DNA size standard can be seen in the first column (left). On top of the picture the slots can be seen (Gel: 1% agarose; 100V, Running time: 2 hours)

2.2.1.2 The AFLP Analysis

Restriction, ligation, and preamplification reactions were performed following the principles contained on the original protocol of Vos *et al.* (1995), with some modifications.

a) Restriction:

Genomic DNA was digested with the enzymes *Eco*RI and *Mse*I by incubation at 37°C in a thermocycler for 1h and 30 minutes. The reaction samples were prepared as follows:

Reaction:

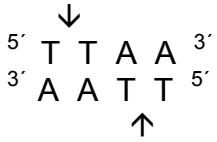
Genomic DNA	250 ng
*Eco*RI	4 U
*Mse*I	4 U
Restriction-ligation buffer	10%
add H_2O to:	30 µl

Restriction-Ligation buffer consists of TrisHAc (10 mM), MgAc (10 mM), KAc (50 mM), DTT [Dithiothreitol ($C_4H_{10}O_2S_s$)] (5 mM), pH 7,5. The restriction-ligation buffer was stored at –20°C.

The *Eco*RI enzyme has the function to cut the DNA rarely, and the *Mse*I cuts it frequently, as shown below:

*Eco*RI

*Mse*I



b) Ligation:

The ligation reaction was performed as a second step immediately after the restriction reaction by adding 10 µl ligation 'master mix' to the restriction reaction sample (30 µl). The 10 µl ligation 'master mix' consisted of:

*Eco*RI Adapter	5 pmol
*Mse*I Adapter	50 pmol
T4 DNA Ligase	1 U
ATP	0,25 mM
Restriction-Ligation buffer	10%
Add H_2O to:	10 µl

The ligation was done on the thermocycler using the programme below and stored by 4°C:

1) 03h 10´	37.0°C	4) 00h 04´	26.0°C
2) 00h 03´	33.5°C	5) 00h 15´	22.0°C
3) 00h 03´	30.0°C	6) Final temperature hold at 4°C	

This programme was designed to increase the activity of the T4 DNA ligase at least at the latest 15 minutes of the reaction (since the best temperature for the enzyme is around 20°C), while at the first 3 hours and 10´ the aim was the maintenance of the activity of the *Eco*RI and *Mse*I, so that if during the ligation two or more fragments get ligated to each other (restriction enzyme sites will be reconstituted) instead of with the corresponding adapters, the new ligated fragment could again be cut by these two enzymes and so avoiding the occurrence of "false" AFLP fragments.

After finished the restriction-ligation reaction, this product was stored at 4°C. The preamplification was done using as template 5μl restriction-ligation product (RL-DNA) which had been diluted 1:5 with TE buffer.

What concerns the adapters, they came as single strands. To combine them to double stranded DNA each 2 pairs of single strands, that means, EA1 and EA2, and MA1 and MA2 respectively, were put together as follows: to build double stranded DNA the solutions (under 1x TE) buffer were put on 0,2ml tubes/8 tubes strips (ThermoStrip™) on a thermocycler using a programme that heats the solutions until 56°C. When the solutions reached this temperature the thermocycler was switched off and the tubes were left in it for 2 hours, so that the temperature decreased slowly and the single strands had time enough to bind and form double stranded adapters. The final solution was then diluted to the final concentration of 5 pmol/μl for *Eco*RI adapter and 25 pmol/μl for the *Mse*I Adapter using 1x TE buffer and then stored at 4°C.

The adapter's structure can be seen below:

*Eco*RI Adapter:

5´ CTC GTA GAC TGC GTA CC 3´
3´CTG ACG CAT GGT TAA 5´

*Mse*I Adapter:

5´ GAC GAT GAG TCC TGA G 3´
3´ TA CTC AGG ACT CAT 5´

c) Preamplification (first PCR reaction):

The preamplification was performed using a 'master mix' using following components:

Aliquot of RL-DNA (diluted 1:5)	5 μl
*Eco*RI Primer E01	0.67-1.0 pmol
*Mse*I Primer M02	0.58-0.87 pmol
dNTP´s, each	0.2 mM
Taq-DNA-Polymerase	1-1.5 U
PCR buffer	10%
Add H_2O to:	20 μl

10x PCR buffer consists of Tris-HCl (100 mM); $MgCl_2$ (15 mM); KCl (500 mM).

The restricted-ligated DNA samples were placed in 0.2 ml tubes/8 tubes strips (ThermoStrip™) and the reaction runs on a thermocycler. For details about the primers see appendix 9.4.1.

The PCR was performed using the following profile: for the preamplification a first step of 30 s at 94°C was used to denaturate the DNA; after that, 20 cycles of denaturation for

30 s at 94°C, annealing for 30 s at 56°C, and extension for 1 min at 72°C. A final step of 5 min of 72°C was performed at the end of the cycling programme. The PCR product was diluted 1:10 with TE buffer and stored at 4°C.

d) Selective Amplification (second PCR reaction):

The selective amplification was done using a 'master mix' for each *Eco*RI/*Mse*I primer combination to be used. The 'master mix' (consisting of all the components listed below with the exception of the preamplification aliquot) was prepared in an 1.5 ml Safe-Lock tube in the way that it was enough for the samples needed to be amplified more 20% samples extra. The 'master mix' for one sample contains:

Aliquot of Preamplification (diluted 1:10)	5 µl
*Eco*RI Primer E+A+NN	1 pmol
*Mse*I Primer M+C+NN	6 pmol
dNTP´s, each	0.2 mM
Taq-DNA-Polymerase	0.5 U
PCR buffer	10%
Add H_2O to:	20 µl

[*10x PCR buffer:* 100 mM Tris-HCl; 15 mM $MgCl_2$; 500 mM KCl]

All the primers used came lyophilised and were diluted with water to get a stock solution of 100 pmol/µl. Working solutions for all the primers were prepared with a 10 pmol/µl concentration. For details about the primers see Appendix 9.4.2.

18 primer combinations were used to analyse the BC_2 population of 300 plants:

E32M47***	E33M49	E38M59	E32M48
E32M49	E33M59***	E38M62	E32M50
E32M59	E33M61***	E38M61	E35M60
E32M61***	E33M62***	E40M60	
E32M62	E33M47	E33M48	

Primer combinations were selected based on AFLP markers previously found by Keygene which used 20 primer combinations in the radioactive AFLP method (pers. com. W. Ecke).

Some of the combinations (***) always produced weak bands. For these specific combinations and for those that did not produce good signals in the first amplification and electrophoresis, an alternative set of conditions were applied for a second PCR:

Aliquot of Preamplification (diluted 1:10)	5 – 7 µl
*Eco*RI Primer E+A+NN	1.1 pmol
*Mse*I Primer M+C+NN	6.6 pmol
dNTP´s	0.2 – 0.24 mM
Taq-DNA-Polymerase	1.0 U
PCR buffer	10%
Add H_2O to:	20 µl

The amplification used for the BC_2 population was done using the following programme: a first step of 30 s at 94°C was used to denature the DNA; after that, 13 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 65°C in the first cycle and decreasing 0.8°C in each cycle (reaching at the end 56°C), and extension for 1 min at 72°C. Followed by 23 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 56°C, and extension for 1 min at 72°C. A last step of 5 min of 72°C was performed at the end of the programme.

In the BC_3 population a slightly different programme was applied trying to favour the amplification of longer fragments (Myburg *et al.*, 2001). It consist of the following profile: a first step of 30 s at 94°C was used to denaturate the DNA; 13 cycles of 30 s at 94°C, 30 s at 65°C (reduced 0.7°C per cycle) and 1 min at 72°C, followed by 25 cycles 30 s at 94°C, 30 s at 56°C and 1 min (extended 1 s per cycle) at 72°C. A final step of 5 min of 72°C was performed at the end.

The final PCR reaction products (20 µl amplification) were mixed with 10 µl loading buffer (Appendix 9.4.2), denatured for 4 minutes at 94°C in a thermocycler and then transferred to ice before loading. Samples could be stored for several months by 4°C in the dark.

e) Electrophoresis on the DNA Analyser Gene Readir 4200 (Li-Cor):

Electrophoresis and detection were performed on a one-dye, model 4200 Li-Cor® DNA Analyser. AFLP fragments were resolved on 6% polyacrylamide gels (25 cm x 0.2 mm).

The following solution was used to prepare the gels (see also Appendix 9.4.3):

- 1.386 M urea (NF-urea Rotiphorese®)
- 1x Long Run 10x TBE buffer
- 12% Long Ranger™ (50% Gel solution)

- 0.7% Ammoniumpersulfat (10% w/v)
- 0.01% Temed

Composition of the Long Run 10x TBE buffer: TrisHCl (1,340 mM), Boric acid (450 mM), EDTA (25 mM). Buffer's pH should be around 9.2.

Each gel electrophoresis on the Li-Cor® 4200 starts with a "Pre run" (Pre-electrophoresis) of 8 minutes. The pre-run flushes impurities from the gel and creates an even temperature across the gel. During the pre-run the microscope is focused. In the one-dye systems only the 800 nm channel is used. The Auto Gain values are used to adjust the fluorescence detector in order to optimize the average background intensity and noise of the electrophoresis image (Li-Cor, 1999a; Li-Cor, 1999b). The values shown below were the ones used in this work:

Focus: 800 Auto gain: 800 Average: 5.0 Noise: 2.0

After the pre-run, a volume of at least 2.0 μl from each sample was loaded using a 20 μl pipette (Gilson). A size standard (1 μl) was always loaded on each gel. The data collection settings were like below:

Signal filter:	3	Images:	800
Bin size:	17	Motor speed:	3
Pixel size:	16		

The electrophoresis was performed for around 5 hours and was carried out with the following parameters:

Voltage:	1000 V	Current:	37 mA
Power:	40 W	Temperature:	45°C

When a gel was used for a second electrophoresis the following parameters were changed as shown:

Voltage: 1100 V Current: 42 mA Power: 45 W

f) Scoring AFLP Gel Images:

First of all some elucidations have to be done concerning the dominant and codominant analysis of AFLP patterns. The expectation, if the DH line is homozygous for a dominant allele (Fig. 4) from the donor parent at a locus (AA), is that the BC_1 genotype is certainly heterozygous for this locus (Aa), and a segregation ratio of 1:1 is expected to

appear in the BC_2 offspring, i.e., 50% BC_2 genotypes should have a band (Aa) and in 50% the band should be absent (aa). On the contrary, assuming that one DH line is already homozygous for the recessive allele from a marker locus (Fig. 5), represented by “aa”, and the dominant allele comes from the recurrent parent, then it is expected that the BC_1 looks like “Aa” but although a 1:1 (Aa : AA) segregation appear in the BC_2, it is a different situation what concerns to the analysis. In this case such segregation represented a great disadvantage since both kind of genotypes show the band and so they are undistinguishable at the first glance, making the analysis of such cases more difficult.

However, according to that what can be expected from a PCR reaction, a genotype “Aa” has normally to show, at the end of PCR, the half amount of amplified fragment compared to genotype “AA”, since in the first one only one allele is present and at the second one two alleles are present and available for the amplification. Based on this, if it is possible to distinguish between the band intensities, then it should be possible to do a codominant analysis of AFLP fragments.

Digital AFLP gel images from the Li-Cor sequencer were scored using the software AFLP-Quantar™ *Pro* 1.0 (Keygene, 2000). The appropriate way to prepare the 16 bit TIFF images from the Li-Cor sequencer for use in AFLP Quantar *Pro* is to import the images in Adobe®Photoshop 5.5. The image is then initially inverted, that means, transformed in an image with dark bands on a white/grey background, and than flipped vertically, producing a mirror image, with the short fragments at the bottom of the image. This enables that the image can be seen from top to the bottom with the lanes displayed from the left to the right. The primer front (unincorporated labelled primers) on bottom of the image was removed with the Adobe®Photoshop.

This image was then opened by the AFLP-Quantar™Pro programme and the lane definition and the other steps required to the analysis were performed as described in the “Users Guide – Part I” (Keygene, 2000). The “band finding parameters” on the sequencer (Li-Cor®) used to screen the BC_2 and BC_3 can be found at Appendix 9.4.4.

Many markers were found to be very weak and difficult to score. Scoring in these cases was only possible by using either AFLP-Quantar™Pro or Adobe®Photoshop to modify the settings of the picture regarding contrast curve and other adjustments.

Semi-automated scoring was performed by manually looking for the AFLP markers previously found by Keygene using 20 primer combinations in the radioactive AFLP method (pers. com. W. Ecke). From this 20, only 13 were chosen at the beginning to be used to screen the BC_2 plants looking for 131 mapped loci (175 Markers), corresponding to 84 markers which the dominant allele coming from the recurrent parent and 91 which the

dominant allele coming from the donor parent. The number of mapped loci (map positions) does not coincide with the number of markers because many markers cosegregate.

Since it was not possible to get markers for all of the 131 loci (because some could not be reproduced) the genome coverage was not sufficient, meaning that only 114 map positions (140 markers) could be found. Therefore it was decided to apply 5 more primer combinations. With the finally 18 primer combinations 154 markers could be scored, which represent 127 map positions. The additional 5 primer combinations were however only used with the 20 BC_2 selected genotypes that have produced the BC_3 population. On the other hand, in the BC_3 population all 18 primer combinations were applied.

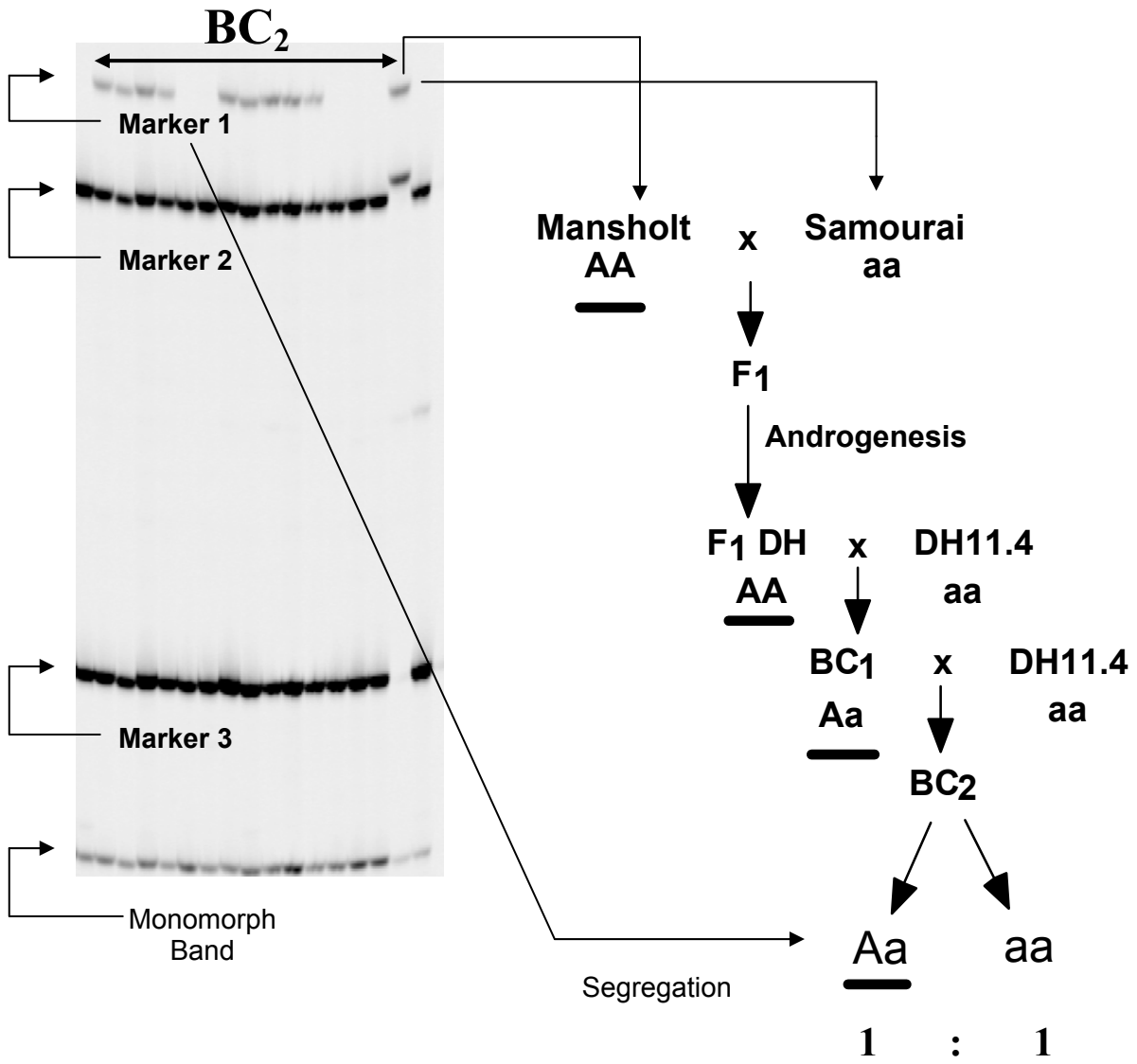


Fig. 4 AFLP marker analyses when the dominant allele comes from the donor parent. As an example, marker 1 is segregating in BC_2.

2.2.1.3 The SSR Analysis

The aim of using SSR markers was to get markers that would help to cover the glucosinolate QTL regions from the linkage groups 2 and 9 (based on map information from: Uzunova *et al.*, 1995; Rudolph, 2001; pers. com. W. Ecke). The amplification was done using a thermocycler performing a programme with an initial step of denaturation for 60 s at 94°C. The initial annealing step was done maintaining for 30 s a temperature of 65°C. Subsequently the temperature was reduced by 1°C every two cycles until a level of 55°C was reached. This annealing temperature was then maintained for 19 cycles. The extension step was always performed for 45 s at 72°C. Two primer pairs (CB10278 and MR13) were applied to the BC_2 and BC_3 populations.

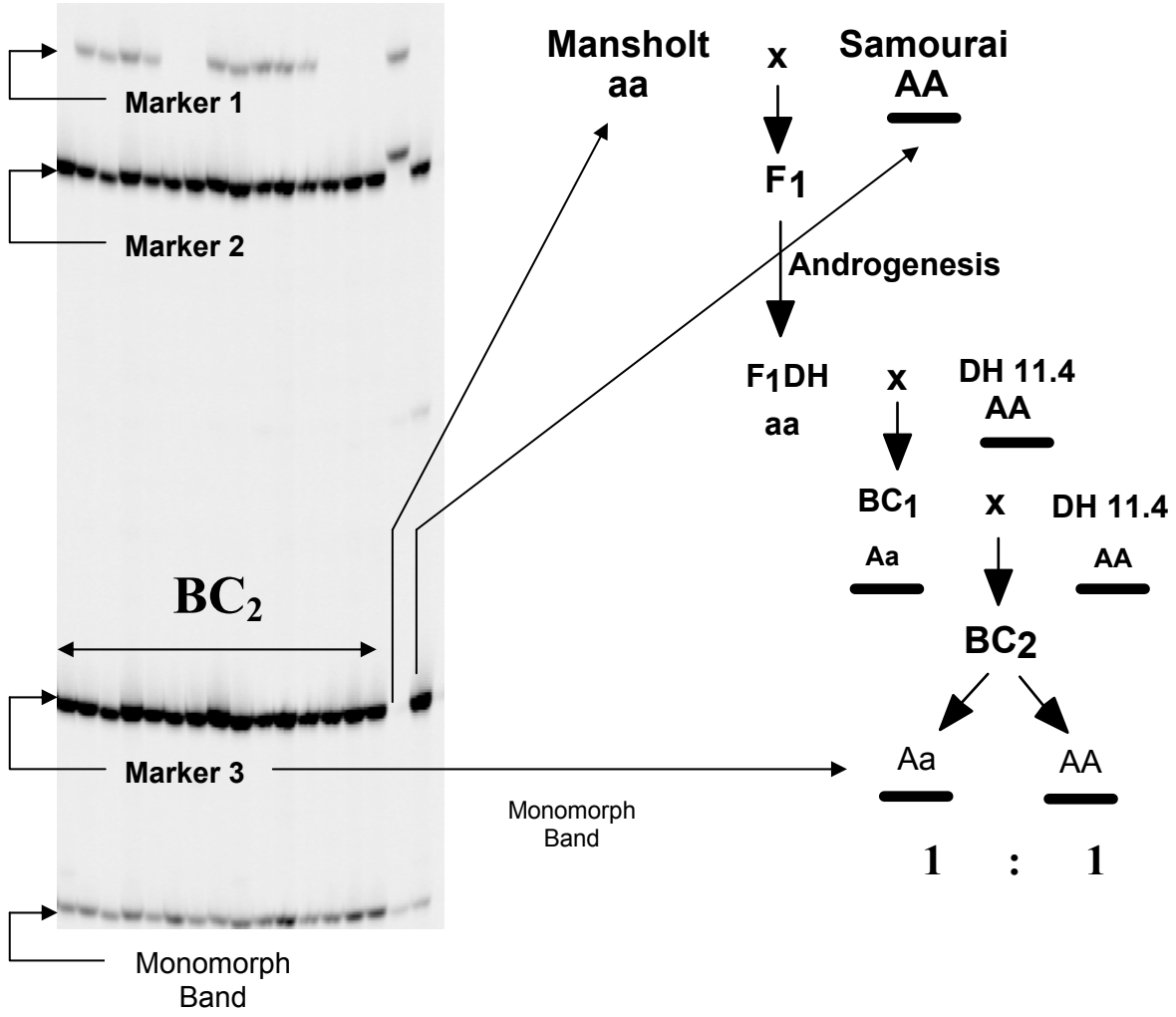


Fig. 5 AFLP marker analyses when the dominant allele comes from the recurrent parent. As an example, marker 3 segregates in BC_2, although visually it is not possible to distinguish the two segregating genotypes (Aa and AA), since both show a band.

The PCR volume in the SSR reactions was 12.5 µl consisting of:

Genomic DNA	25 ng
Forward primer	0.7 µM
Reverse primer	0.7 µM
dNTP´s, each	0.2 mM
Taq-DNA-Polymerase	0.5 U
PCR buffer	10%
$MgCl_2$	1.5 mM
DMSO (Dimethyl Sulfoxide, C_2H_6SO)	2%
Add H_2O to:	12.5 µl

10x PCR buffer consists of Tris-HCl (100 mM); $MgCl_2$ (15 mM); KCl (500 mM).

The fragment analysis was done by an agarose gel electrophoresis using 4.34% Low Melting MetaPhor™ Agarose gels (30 cm x 18.8 cm), which means that 10 g of the agarose was dissolved into 230 ml 1x TBE buffer by heating in a microwave oven. After the gel solution had cooled to around 65°C and was completely clear it was poured into plate with the combs. After the gel was solidified it was put at 4°C for at least 30 minutes. By remelting the same agarose gel could be utilized up to 4 times.

The 12.5µl PCR product was mixed with 5 µl loading buffer (0.25% bromphenol blue; 0.25% xylene cyanole FF, 40% saccharose, H_2O). The total amount of 17.5 µl was then loaded into the gel which was inside an ECONO-Submarine Gel Unit and the electrophoresis was done using 80 V for 13-15 hours in 1x TBE buffer.

After electrophoresis the gel was incubated for 20 minutes in an ethidium bromide solution (1.0 mg/l H_2O), followed by a treatment by water for 20 minutes. The DNA fragments were then visualized through UV-Transluminator (UV-Rays at λ= 254 nm) and photographed using a video-camera. The pictures were also saved as files using the software 'Herolab E.A.S.Y. store'. The analysis of the two SSR markers was done by visually scoring the bands (present or absent) in the digital image.

2.2.1.4 The RAPD Analysis

The aim of using one RAPD marker (primer OPQ9) was to get a better coverage of linkage group 6. The part of linkage group 6, which contains the possible minor QTL for glucosinolate, was not properly covered by the AFLP markers in the map.

The PCR reaction was performed using a thermocycler performing the following programme: a first step of 30 s at 94°C was used to denaturate the DNA; 46 cycles of 30 s at 94°C, 1 min at 35°C and 2 min at 72°C, followed by a final step of 5 min of 72°C.

The PCR volume was 25 µl consisting of:

Genomic DNA	25 ng
Primer (Appendix 9.5)	0.16 mM
dNTP´s	0.4 mM
Taq-DNA-Polymerase	1.25 U
PCR buffer	10%
$MgCl_2$	3 mM
Add H_2O to:	20 µl

10x PCR buffer consists of Tris-HCl (100 mM); $MgCl_2$ (15 mM); KCl (500 mM).

The fragment analysis was done by an agarose gel-electrophoresis (1.8% PeqGoldUniversal agarose) in which 6.3 g of the agarose was dissolved in 350 ml 1x TAE buffer and heated in a microwave oven. When the gel solution had cooled to around 45°C it was poured into the plate (30 cm x 18.8 cm) with the combs.

The 20 µl PCR product was mixed with 5 µl loading buffer (0.25% bromphenol blue; 0.25% xylene cyanole FF, 40% saccharose, H_2O). The total amount of 25 µl was then loaded onto the gel and the electrophoresis performed like described for the microsatellites using 40 V for about 14 hours.

After the electrophoresis the gel was treated in the same way as explained for the microsatellites and the band capture and scoring done using the procedures also described for the microsatellites.

2.2.2 Field Experiments with the BC_3 plants

In the present study the BC_2 population (300 plants) was genetically characterized using 154 AFLP markers, 2 SSR markers and 1 RAPD marker (130 loci = 1,325 cM). Near almost 300 BC_3 families were grown in two field trials. Not all 300 families could be evaluated due to limitations in the available amount of seeds. Based on the marker information on BC_2, BC_3 families segregating for only one or two of the six QTL were selected. After harvest of the BC_3 plants, seed glucosinolate content was evaluated using Near-Infrared-Reflection-Spectroscopy (NIRS).

Two field experiments were planted in Göttingen in August 2001 to examine the glucosinolate content in seeds in open pollinated plants. In the first experiment (Reinshof/Göttingen) a total of 333 plots were planted. Each plot consisted of two rows of 25 plants of one BC_3 family each (the rows were 2.5 m long, spaced 33 cm; the within-row spacing was 10 cm), arising from one BC_2 mother plant which has been backcrossed to DH 11.4 (Samourai). The progeny from a total of 279 BC_2 plants, that means, 279 BC_3 families were submitted to field conditions without replication. Ten plots having the ten DH lines that were the parents of the backcross populations were also evaluated as well as the respective BC_1 genotypes. Between every 10 double rows, one double row of DH 11.4 (Samourai) was planted. The donor parent, the DH 5.1 (Mansholt) was also included in the experiment. This first experiment was planted on August 27th, 2001 and harvested on July 16th, 2002. The agronomic practices were as usual.

The second experiment (Elliehausen/Göttingen) was build up in a similar way, but using only 228 BC_3 families due to limitations in the available amount of seeds of some families. This experiment was planted on August 28th, 2001 and harvested on July 20th, 2002. All the seeds used in the field experiments have arisen from backcrosses done in the greenhouse. Seeds from the DH lines and corresponding BC_1 genotypes used in the field trials have been produced before.

According to the marker information (AFLP) from the BC_2 plants, BC_3 families were selected. For the selection plants were chosen which contain only one or a maximum of two Mansholt's alleles of the six QTL for glucosinolate known from previous studies. Therefore, plots in which it was known that the BC_2 mother plant carries more than two Mansholt's alleles were not harvested as individual plants, but only as a bulked sample from all plants in this plot. Seeds from each plant were collected from the main branch (distal) of the plant since according to Clossais-Besnard and Larher (1991) seeds of different sampling positions contain quite similar total glucosinolate concentrations; however, the individual behaviour of each glucosinolate seems to vary.

2.2.3 Determination of seed glucosinolate content by NIRS and HPLC

The glucosinolate content was determined in intact-seed samples of 300 mg by near infrared reflectance spectroscopy (NIRS) using a Foss Tecator 6500 instrument and the software ISI version 1.04. Some of those samples were later also analysed by HPLC (High Performance Liquid Chromatography). The spectra were standardized using the goeneu.std equation (Institut für Pflanzenbau und Pflanzenzüchtung, Universität Göttingen) and the calibration used was the Raps2001.eqa (Peter Tillmann, LUFA, Kassel).

Measurement of glucosinolate content by HPLC was carried out as described by Buchner (1988, cited after Herrmann, 1992) using a sulfatase for desulfatation. However, it should be pointed out that this method can not detect rare GSL with negative bycharged groups others than the sulfat group (Herrmann, 1992). Details can be seen in appendix 9.6.

2.2.4 Estimation of glucosinolate content and QTL effect

The BC_3 plants were tested in the field and if the parental BC_2 plant carries one allele from the three major QTL, it is expected that a 1 : 1 segregation in the BC_3 plants will occur. That means that it should be possible to separate one group (50% of plants) showing a high level of glucosinolate content, while the other group should show low glucosinolate content. In those cases, the donor allele effect of the QTL was estimated calculating the difference between the means of the groups with high and low glucosinolate content. Since the lower groups correspond to Samourai, this one was used for calculating the difference mentioned above.

With the minor QTL such great differences between groups are not expected since the effect of such QTL are weak and more affected by the environment, which then contribute to transform a clear genetic 1 : 1 segregation in a diffuse cloud of points (Fig. 6). In those cases in which no clear segregations could be observed, a model was used to estimate the QTL effect. In this case, all the genotypes measured resulted in a cloud of points with mean called “z” (Fig. 6). It can be expected that the mean of all BC_3 genotypes in one family is composed by those which do not carry the Mansholt’s allele and genotypes which carry this allele. As an example Fig. 6 shows the graph from the offspring (BC_3) of the BC_2 genotype 1053.2.07 (in Reinshof) which carries one Mansholt’s allele at the QTL on LG 9, which thus will segregate. Although this is a major QTL the distribution of values do not permit a clear separation into two groups.

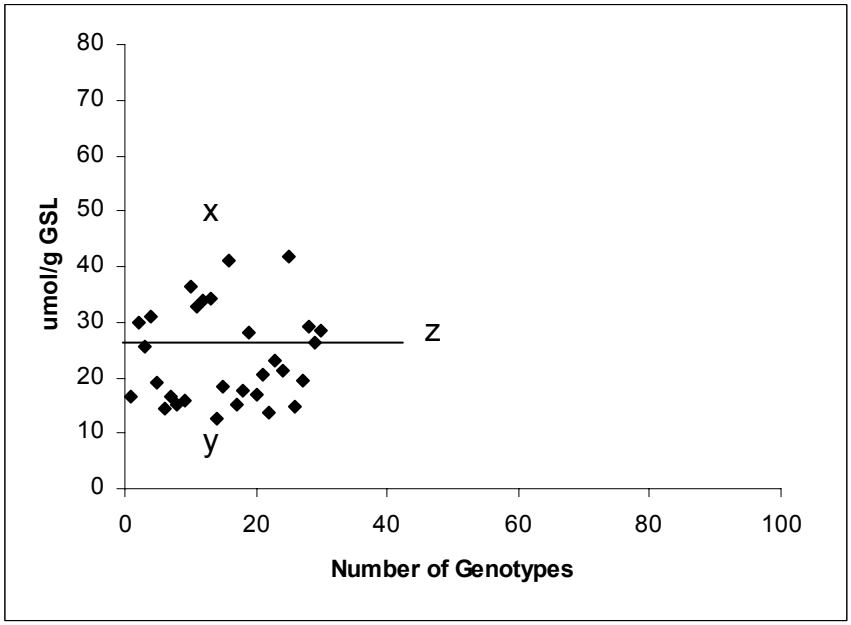


Fig. 6 Distribution of glucosinolate values from one BC_3 family which segregates for one QTL for glucosinolate content.

With “x” und “y” being the mean glucosinolate content of the heterozygous and homozygous genotypes, respectively (model 1):

$$\bar{z} = \frac{x + y}{2}$$

Since “y” is the glucosinolate content of Samourai (DH 11.4), than:

$$x = (2 * \bar{z}) - y$$

In that way, “x” represents the estimated phenotypic value (PV) from the genotypes that carry one Mansholt’s allele at QTL on LG 9. The genotypes represented by “y” correspond to Samourai since in these genotypes two Samourai’s alleles are present at the QTL on LG 6, whereas other parts of the genome still have components of Mansholt. “y” only can be used as Samourai if the results show that Samourai is not significantly different from the genotypes carrying two Samourai’s alleles at the QTL on LG 6 but with other parts of the genome from Mansholt, what should be verified in the data set from the experiments. Finally, QTL effect, called “Phenotypic Effect” (PE) was calculated for each QTL as the difference between phenotypic classes, i.e. between the heterozygous class (x), and the homozygous class (y), which corresponds to Samourai.

Following such model, a similar model (2) can be applied if two donor alleles of two different QTL are present in the BC_2 plant. In this case usually all genotypes measured result in a cloud of points which mean is "z", but now four groups contribute to such mean. Taking as example one BC_2 genotype carrying Mansholt's alleles for the QTL on LG 2 and 3, such a model should be (model 2):

$$\bar{z}_{(3x2)} = \frac{1}{4}D_{(3x2)} + \frac{1}{4}C_{(3)} + \frac{1}{4}B_{(2)} + \frac{1}{4}A_{(Samourai)}$$

$\bar{z}_{(3x2)}$ = mean value from the BC_3 genotypes coming from one BC_2 mother plant which carries two donor alleles for QTL on LG 3 and 2.

$D_{(3x2)}$ = phenotypic value from BC_3 genotypes carrying donor alleles for QTL 3 and 2, therefore the QTL on LG 3 and 2 will be in heterozygous state.

$C_{(3)}$ = phenotypic value from BC_3 genotypes carrying the donor allele for QTL on LG 3. This "C" value could be calculated using the model 1. The same procedure is valid for the genotype B. The genotype "A" corresponds to Samourai, like in the model 1.

Since "z" is the value that is derived from the field data (plot mean), and considering that the other 3 genotypes (C, B and A) can also be estimated from other families, it is possible to estimate the phenotypic value (PV) of the genotype carrying two alleles from the two QTL (D).

Two kinds of variances were calculated. The first one is the "variance within families" and refers to the mean of the variances calculated separately for each of the BC_3 families segregating for the same QTL. It was also calculated for the parents, DH lines, BC_1 or for the segregating groups inside one BC_3 family. It gives information about the dispersion around the mean of the different plants in the different BC_3 families.

The second variance (between families) calculated refers to the variance of the means of the different BC_3 families or genotypes (Parents, DH and BC_1) which have the same status which refers to the presence of donor alleles at the six QTL. It gives information about the reliability of the estimated mean.

Nevertheless, sometimes the variances between families could not be calculated since only one family was harvested. In the same way sometimes it was impossible to estimate the variance within families, since some families, especially the ones segregating for more than one QTL, were harvested as a block, therefore making impossible an estimation of the variance within the family.

The epistatic effects (Tab. 8) were estimated using Samourai (mean from Reinshof and Elliehausen) as the reference value, since the QTL effects were calculated in the same way, i.e., taken Samourai as reference basic value too. That means that for testing the presence of epistasis between two donor alleles, for example for QTL on LG 2 and LG 3, the PV from this genotype (2x3) was compared to the value obtained summing up the value of Samourai to the individual effects of the donor alleles on LG 2 and LG 3, as estimated in families segregating for only one of the two QTL. If the resulting value surpasses the PV than it is an indication that negative epistatic effects are present, that means that the interaction between these nonallelic alleles in the two involved genes is causing a reduction of the phenotypic value when compared to the expected value assuming only additive effects. If the calculated value is smaller than the PV than positive epistatic effects are present. If the two values are approximately the same there is no evidence for epistatic effects.

2.2.5 Field Experiments with DH Lines (1998/99 and 1999/00)

During 1998/1999 and 1999/2000 (both in two locations and with two replications) experiments were carried out by Gül (2002). Since the DH lines used in such experiments are coming from the same cross (Mansholt x Samourai) as the BC_3 families evaluated in 2001/2002, it will be possible to consider such data in the analysis and discussion. The difference is that in the two experiments of 2001/02 heterozygous material (BC_3 plants) was evaluated, and in Gül's case, homozygote one, which will permit some interesting comparisons. Since Gül has worked with nitrogen levels, only the data results from nitrogen level N1 (240 kg N/ha) will be used at the present work.

Each plot consisted of two rows (the rows were 2,5 m long, spaced 30 cm; the within-row spacing was 5 cm) and 142 DH lines were tested, but only 37 are suitable for the proposed aims, that means, only 37 lines have marker data without doubts in all 6 QTL regions, and all of them have zero to four Mansholt's alleles present at the six mentioned QTL.

2.2.6 QTL mapping

142 DH lines were evaluated by Gül (2002) during 1998/1999 and 1999/2000 (both in two locations and with two replications) and the data from nitrogen level N1 (240 kg N/ha) were used for QTL mapping. The marker information were the same used by Gül (2002), i.e., a map derived from the original from Uzunova *et al.* (1995) having 185 RFLP markers, which originated a map with 20 linkage groups covering 1,739 cM of the rapeseed genome. QTL mapping was carried out using the programme *QTLMapper Version 1.1* (Wang *et al.* 1999a)

which implements a mixed model composite interval mapping approach (Wang *et al.* 1999b). Following this model, main additive effects, additive x additive epistatic effects were estimated.

The inclusion and exclusion significance threshold used for the selection of main effect and interaction markers by stepwise regression was P = 0.005. Positions of QTL or pairs of loci with epistatic interactions were identified by a LR-test (likelihood ratio test) testing for significant additive and additive x environment interaction effects or additive x additive epistatic interactions and epistasis x environment interaction effects, respectively. For declaring a QTL, a LR value of 7.8 was set as significance threshold, equivalent to a probability level of P = 0.005. The distance between two adjacent testing points in mapping was 2 cM. The positions of local maximum LR values were selected and presented as putative additive QTL and/or epistatic locus pairs. In a final step, estimates of main additive or epistatic effects and environment interaction effects as well as estimates of the significance of a QTL or epistatic locus pair (probability threshold P = 0.005) were derived by a Jackknife test. Main additive effects are expressed as difference in glucosinolate content due to the substitution of a Samourai's allele by a Mansholt allele. Epistatic effects (AA_{ij}) with a negative sign indicate that recombinant allele combinations at the two loci involved in epistasis increase phenotypic values, while a positive sign indicates that parental allele combinations increase phenotypic trait expression.

Nothing shocks me. I'm a scientist.

Harrison Ford (1942 -), as Indiana Jones

3. RESULTS

3.1 Experiences with marker technology

3.1.1 AFLP markers and scoring of AFLP images

The detection of AFLP fragments using infrared technology offers several advantages over conventional detection methods using autoradiography. First, especially because the use of radioactivity is eliminated, and gel images are available for analysis immediately after gel electrophoresis, as TIF files. The electrophoresis image represented by Fig.7 shows the elementary components of such TIF files: (a) One size standard lane can be found, used to facilitate the localization of the markers; (b) The parents (Mansholt and Samourai) have always been loaded; (c) The primer front can be seen; (d) Markers which the dominant allele comes from the donor parent can be found, like 148M, 208M, 214M. In those cases a band can be seen in Mansholt but it is absent in Samourai; (e) One codominant scored marker ("recurrent marker") can be seen, as an example; (f) One monomorph band, as an example (there are many more in the Fig. 7).

More PCR reactions and gel electrophoresis than originally planned were carried out both in BC_2 and BC_3 populations. The original work with AFLP markers, used to build the AFLP map, was made using the standard AFLP method with radioactive labelling. The original marker analysis was not carried out in Göttingen but was done by Keygene. Therefore differences among the two kinds of methods could be expected and exactly this was observed since some bands that were quite good to see in the autoradiography from Keygene disappeared or were too weak to allow a good scoring in the non radioactive method. The inverse was also common.

Therefore, the question was to find and reproduce the markers previously defined by Keygene in this population. When some markers could not be found in one electrophoresis file, in spite of the fact that the image looked pretty good, the PCR was repeated, sometimes under a little different condition, trying to get these missing bands as far as possible.

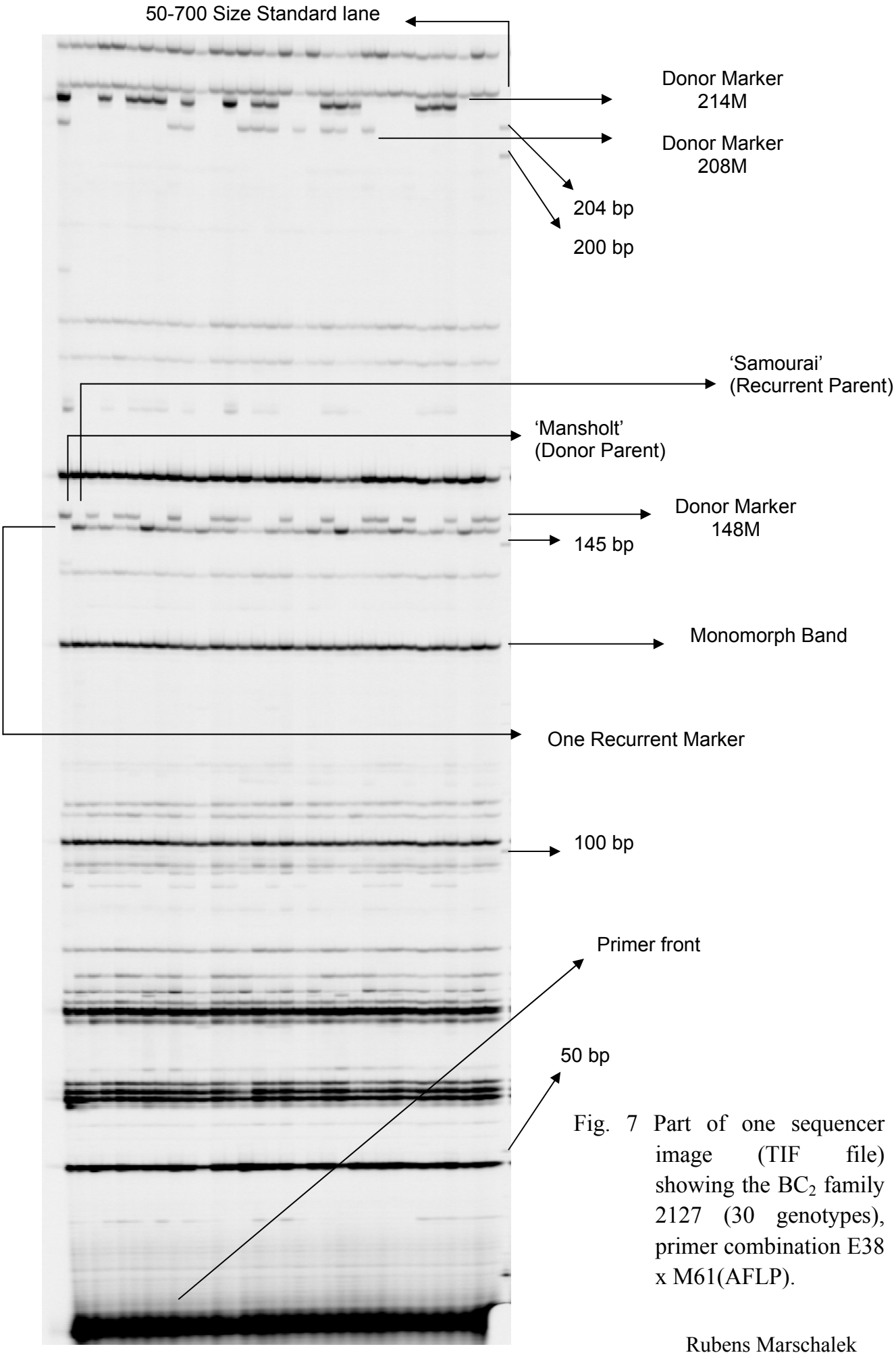


Fig. 7 Part of one sequencer image (TIF file) showing the BC_2 family 2127 (30 genotypes), primer combination E38 x M61(AFLP).

The markers that could not be reproduced were located by chance in the gel, that means, sometimes quite short fragments were missing, sometimes larger fragments were missing. Even comparing AFLP banding patterns from the same method, like using the 'Li-Cor' sequencers, some differences can be found among different analyses, reaching for example 1.15% non-reproducible data points in the work from Myburg *et al.* (2001).

Consequently, in total it was expected to get 213 markers with the 18 primer combinations, but only 161 markers were obtained, because 52 markers were not reproducible or reproducible only in 60% or less of the families, giving an overage reproducibility of 75.6%.

Scoring of AFLP images

The scoring of AFLP markers with AFLP Quantar *Pro* itself took at least 3-4 hours per gel image. The long time required for each image in this work, compared to other references (Myburg *et al.*, 2001) is probably due to the fact that it was necessary to find exactly the previously defined markers, and even try to score very weak markers.

It was always important to check if the scoring results from the AFLP-Quantar™ *Pro* software were really error free what, unfortunately, was often not the case. For markers with the dominant allele coming from the donor parent ("donor markers"), it was not rare that the software classified a genotype showing a clear band as a homozygote "aa", i.e. as a genotype without band. Corrections "by eye" were necessary in each image and in each individual marker and genotype, since the software results could not be relied on.

With regard to the markers with the dominant allele coming from the recurrent parent ("recurrent markers"), it was a still more complicate situation. AFLP-Quantar™ *Pro* is a software based on the theoretical premise that AFLP technology allows the quantitative measurement of the degree of amplification of a fragment.

In many cases however the software could find the desirable distribution showing a histogram with two curves which would mean that two different groups of genotypes could be identified, one with low intensities, and another one with higher intensities if compared to the first one. An example is shown in Fig. 8, but such perfect situations only occurred in 26% of the markers scored in a codominant way in the BC_2 generation.

Quite often the distribution of recurrent markers was so that many genotypes were classified as "U" (unknown), that means, they represent genotypes that carry intensities which are located in the area between two border lines where none of the distributions are predominant. "U" means that the genotype can not be classified as heterozygote or homozygote in the case of a "recurrent marker" and it was very difficult to handle the markers

which show such genotypes in great number. Sometimes more than 50% of the genotypes of one marker were fit with an “U”. Examples from such situations can be found below, as presented on Fig. 9 and 10, in which two different situations can be seen, both presenting many genotypes classified as “unknown”. The situation showed in Fig. 10 is worse, since too many genotypes are scored as “U”.

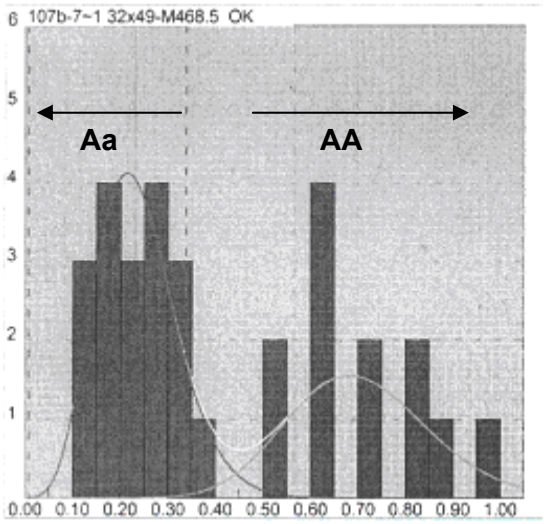


Fig. 8 On the left side a group of genotypes have a low band intensity (Aa), and could be clearly separated from the other group of genotypes (on the right side), that showed a higher band intensity (AA). (marker at the position 468.5 bp from the picture 107b – BC_2 – primer combination E32xM49).

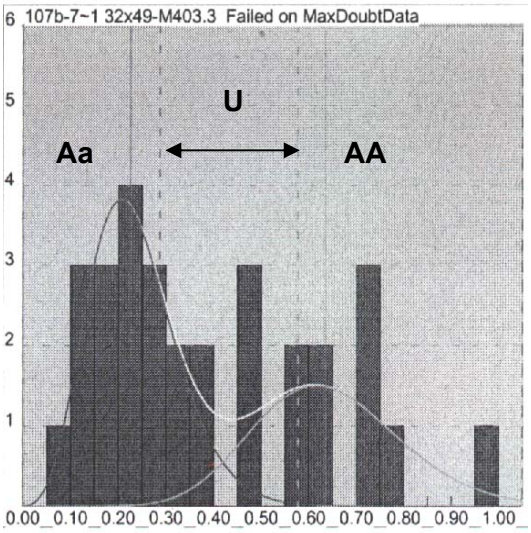


Fig. 9 On the left side a not very good defined group of genotypes that shows a low band intensity, and going on to higher intensities many other genotypes appear without forming a clear second group. That’s what the software calls (middle intensity region) as “U” genotypes (unknown) (marker at the position 403.3 bp from the picture 107b – BC_2 – primer combination E32xM49)

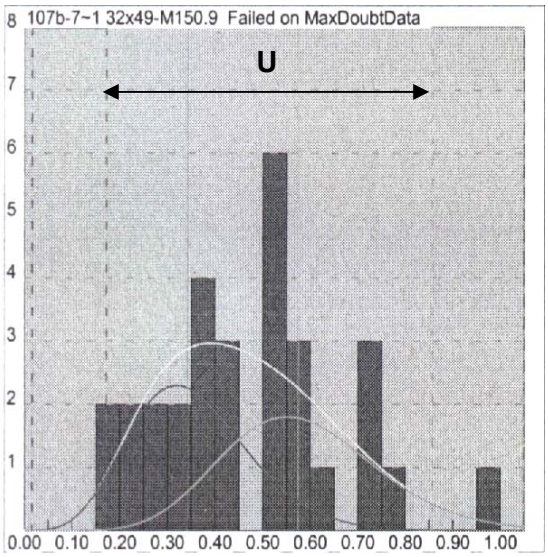


Fig. 10 Showing a frequent "normal distribution" from a group of BC_2 genotypes that should give a 1:1 segregation between the genotypes AA : Aa. In this case the software is not able to separate AA genotypes from Aa ones showing the message "Failed" (marker at 150.9 bp from the picture 107b – BC_2 – primer combination E32xM49)

3.1.2 SSR markers

Fig. 11 shows the microsatellite marker MR13A in some genotypes, as an example. The aim of using SSR markers was to get two markers that would help to cover properly the glucosinolate QTL regions from the linkage group 2 and 9, respectively mapped as markers CB10278 and MR13A. The marker CB10278 constitutes a band of 243 bp for the donor parent (Mansholt) and at 228 bp for the recurrent parent (Samourai). In the other case (marker MR13A), the primer generates a band at 155 bp for the donor parent and is absent in the recurrent parent (Rudolph, 2001). These primers were applied to the BC_2 and BC_3 populations.

3.1.3 RAPD markers

Fig. 12 shows the RAPD marker OQ1590 for some BC_3 genotypes as an example. The aim of using one RAPD marker was to get one marker (OQ9.1590, band present for the donor parent at 1590bp and absent for the recurrent parent) at the end of LG 6. This part of LG 6, which contains one possible minor QTL for glucosinolate content, was not properly covered by the AFLP markers in the map. This primer was applied to the BC_2 and BC_3 populations.

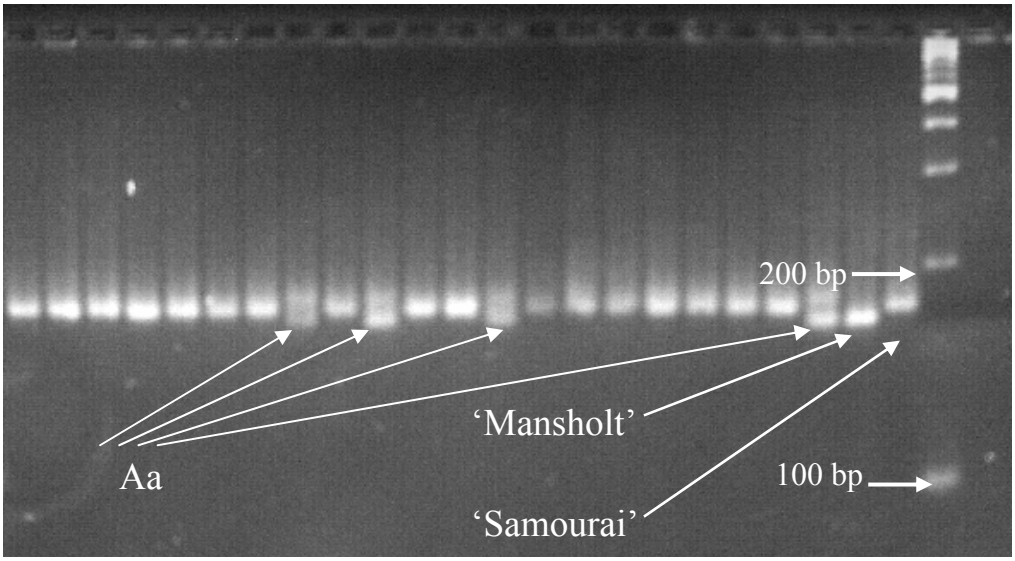


Fig. 11 From the right to the left side the DNA Size-Standard can be seen, and after that the recurrent parent and donor parent respectively, the last one showing the expected band at the position 155 bp (marker MR13A). Getting closer to the left side of the gel, 4 more plants showing this band can be found (Aa), while the other ones do not show it (aa).

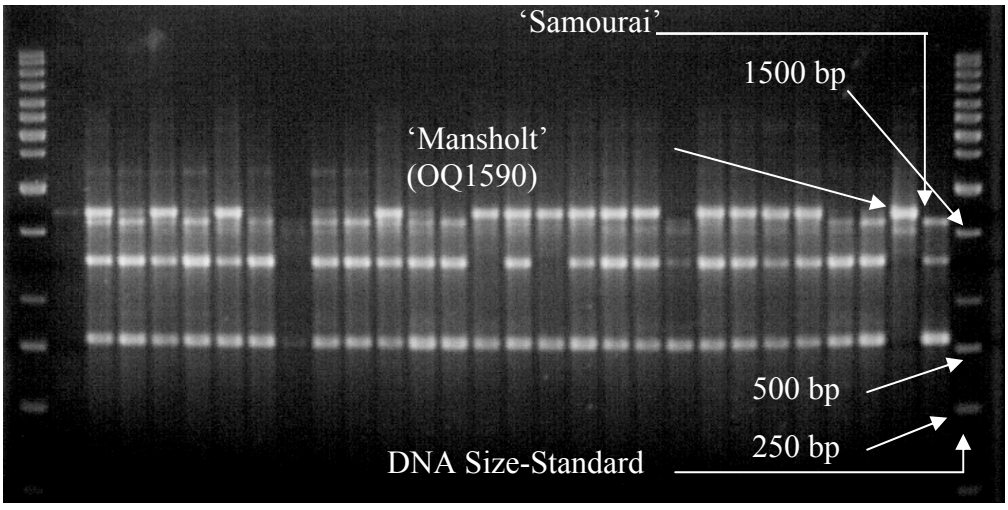


Fig. 12 From the right to the left side the DNA Size-Standard can be seen, and after that the recurrent parent and then the donor parent, the last one showing the expected band (marker OQ1590).

3.2 MAS in BC generations (developing substitution lines)

3.2.1 Genotypes selected in BC_2

In this generation, 300 BC_2 plants originated from the 10 selected BC_1 genotypes were analysed with molecular markers and with the marker data 20 out of 300 BC_2 plants could be selected that cover the whole donor genome. However, the genome was not completely covered by the 140 Markers (114 map positions) generated by the initial 13 primer combinations (Tab. 1), and so 5 more primer combinations were applied to the 20 BC_2 selected plants so that they could be better genotyped. So, it sums up to 18 primer combinations. This resulted in 161 AFLP markers (Tab. 1), representing 127 map positions. The 130 map positions correspond to (164 markers) reported in Tab. 1 include more two SSR markers and one RAPD marker.

Additional analyses were done using 2 SSR and one RAPD marker to get a better coverage of the regions on LG 2, 9 and 6, containing QTL for glucosinolate content. The SSR and RAPD were used for the BC_2 and BC_3 populations. At the end a total of 164 markers (Tab. 1), which represent 130 map positions, was reached.

The basic genetic map used in this study includes RFLP, RAPD and SSR markers and covers about 1,800 cM of the rapeseed genome. The 130 map positions used for marker assisted selection in BC_2 and BC_3 cover 1,325 cM of this map, so that in the development of substitution lines 72.86% of the mapped genome could be covered. In the representations of the linkage groups in appendix 9.7 using the GGT Software (Van Berloo, 1999) these regions are presented in black and dark grey color.

The 20 selected BC_2 genotypes had 10, 11, 13, 8, 9, 12, 10, 11, 8, 12, 18, 9, 9, 13, 8, 8, 13, 13, 11 and 7 donor segments respectively, which results in a mean of 11 segments with a variance of 6.87. General, for all 300 BC_2 plants from this population the mean number of donor fragments present were 11.22 (variance = 7.4), with 19 the highest and 5 the lowest number (Fig. 13).

3.2.2 Genotypes selected in BC_3

Each of the 20 selected BC_2 plants produced 15 plants as offspring, and from these 300 BC_3 plants 52 were selected in the marker assisted selection. Contrary to the BC_2 population, in the BC_3 population all the 18 primer combinations as well as the two SSR and one RAPD markers were applied to screen all 300 BC_3 plants giving genetic information on a total of 120 map positions (Tab. 1).

Tab. 1 Number of donor and recurrent markers and loci in BC_2 and BC_3 generations

		BC_2			BC_3		
Kind of markers	Units (No.)	Donor	Recur.	Total	Donor	Recur.	Total
a) 13 AFLP p.c.	markers	79	61	140	–	–	–
b) 13 AFLP p.c.	map positions	74	40	114	–	–	–
c) 18 AFLP p.c.+2 SSR+1 RAPD	markers	103	61	164	105	27	132
d) 18 AFLP p.c.+2 SSR+1 RAPD	map positions	95	35	130	96	24	120

p.c. = primer combinations

c and d = only in the 20 selected BC_2 plants but in the whole BC_3 population

Donor = dominant allele comes from the donor parent (AFLP markers scored dominantly)

Recur.= dominant allele comes from the recurrent parent (AFLP markers scored codominantly)

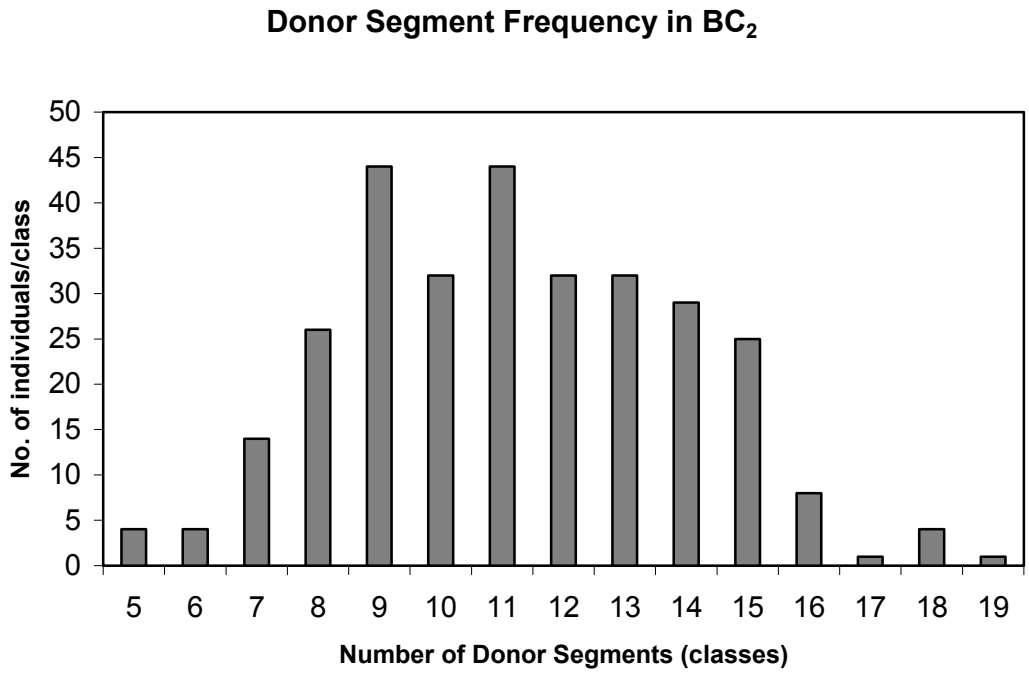


Fig. 13 Distribution of donor segment number in the 300 BC_2 plants (number of individuals in each donor segments class)

The reduction from 130 loci (164 markers) used in the BC_2 generation to 120 loci (132 markers) in the BC_3 is due to the elimination of some doubtful, codominantly scored markers and the addition of the marker e3560.4-344M on linkage group 16. The change in marker number did not change the genome coverage, which was still around 1,325 cM (Tab. 2). Still in Tab. 2 the percentages of recurrent and donor covering are present considering

the haploid genome, that means that for example if the donor genome is 3.8% this means that in 3.8% from all loci there are a donor allele present, because the genotypes are heterozygous, i.e., the percentage of the donor genome is this number divided by two, since one chromatide is coming from the recurrent parent. So, the donor genome present in the 52 selected BC_3 plants is 1.9% (3.8 / 2), which means that in average 3.8% of the loci of each plant have one donor allele present.

General, for all 300 BC_3 plants from this population (Fig. 14) the mean number of donor fragments present were 7.1 (variance = 5.5), with 13 the highest and 2 the smallest number. The 52 selected genotypes are shown graphically in appendix 9.8 using the GGT Software (Van Berloo, 1999). The 52 selected genotypes show a mean of 4.9 donor segments in the genome, with 2.44 variance, giving as maximal value 8 and 2 as minimum (Fig. 15).

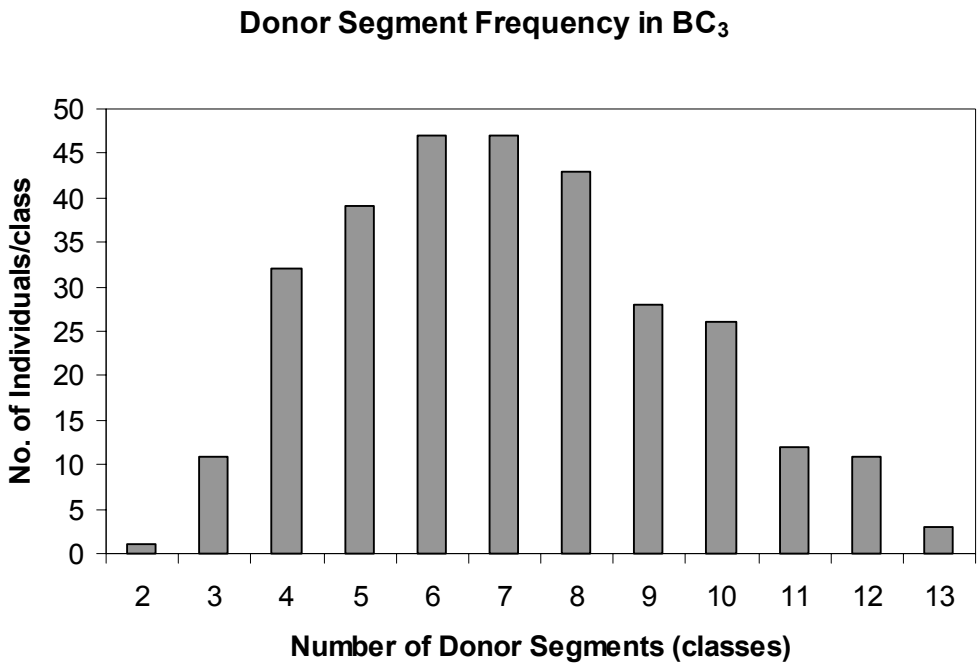


Fig. 14 Distribution of donor segment number in the 300 BC_3 plants (number of individuals in each donor segments class)

Tab. 2 Percentage of the genome coverage estimated from the 1,800 cM of the complete map and the 1,325 cM effectively covered in this study. Presented are the percentage of homozygous recurrent (Recur.) and heterozygous donor segments (Donor) in the different BC generations.

Genome lenght	Gene-	n° of sel.	% of genome coverage [1]		
cM	ration	plants	Donor	Recur.	Unk. [2]
1800 [3]	BC_1	10	42.62	55.55	1.82
1800 [3]	BC_2	20	7.7	78.9	13.5
1800 [3]	BC_3	52	3.15	84	12.95
1325 [4]	BC_3	52	3.8	92.55	3.75

[1] Average % of donor, recurrent or unkown genome covered per plant in each generation (with respect to the genome covered by the selection markers)
[2] Unk.=unkown
[3] % of genome coverage calculated taken the total known lenght of the map as 100%
[4] % of genome coverage calculated taken the covered map as 100%

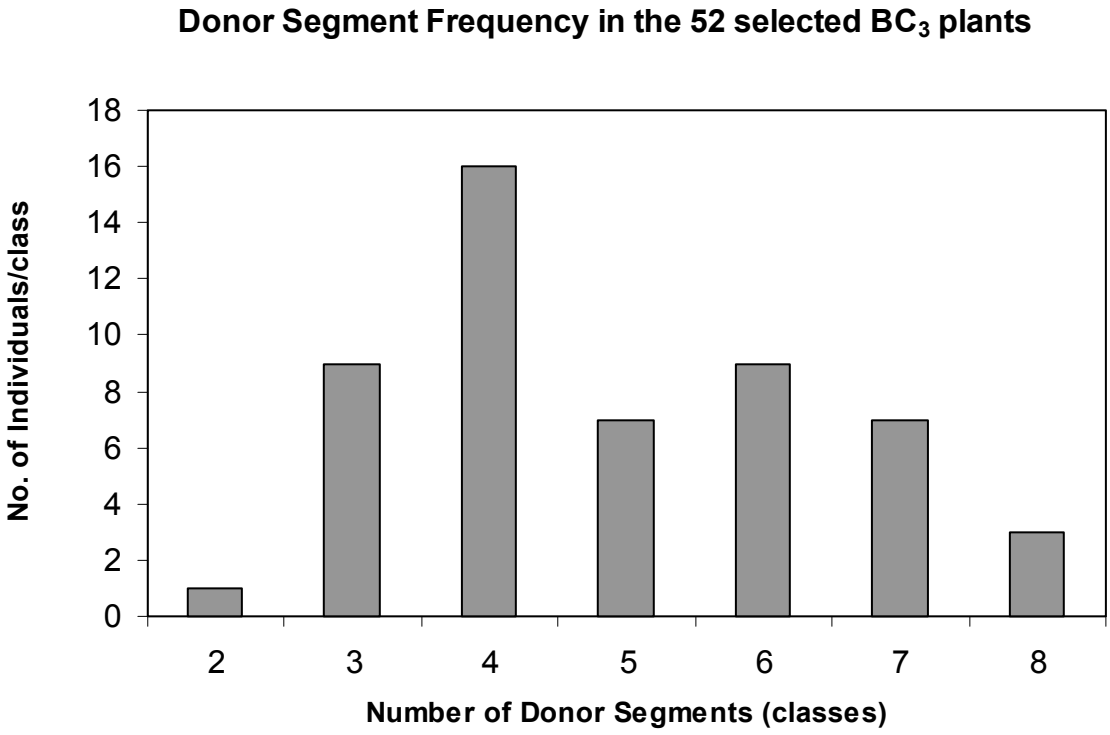


Fig. 15 Number of individuals in each donor segment class in the 52 BC_3 selected plants.

3.3 Glucosinolate content

3.3.1 Glucosinolate content in Reinshof and Elliehausen

Tab. 3 and 4 present the results (means and variances) for glucosinolate content from the two field experiments carried out in Reinshof and Elliehausen (Göttingen). Included in the experiments were the BC_3 families, the ten DH lines that were the parents of the backcross populations as well as the respective BC_1 genotypes. Between every 10 double rows, one double row of DH 11.4 (Samourai) was planted. These tables show, for each location, the kind of genotypes, the number of BC_3 families segregating for the indicated QTL or, the number of plots of the specified genotype. The phenotypic mean column in both tables show the phenotypic value got in each plot, as a mean of the individual plants harvested in a plot, or as the value of the bulk harvest of all the plants in the plot. Since some plots (families) has been harvested as a mixture sample of plants the "number of families or plots" used to calculate the phenotypic mean (4^{th} column) is usually superior to the number of families used to estimate the variance within families.

In each genotype, the variance between the mentioned families or plots is shown, and even the mean of the variances within this families or plots. Genotypes in which only one segregating family was studied do not present the variance between families, since there is only one. In other cases, the variance within families are not presented, since those families were harvested as a bulk and the glucosinolate content was not analysed in individual plants in the family but as a mixture sample from the whole plants on this family (plot), like has been explained in material and methods. The phenotypic value (PV), calculated according to model 1 or model 2, is given in Tab. 3 and Tab. 4 as well as the respective phenotypic effect (PE) over the basic genotype, which is Samourai. The PE, when one QTL is taken isolated, corresponds to the QTL effect. PE is the result of taking the "phenotypic value" and subtracting the value of the recurrent parent. This will thus represent really how big the effect is assuming that Samourai is the lowest reference value.

Tab. 3 Phenotypic Values and Phenotypic effects of genotypes and BC_3 families at Reinshof 2001/2002.

Genotype	Donor alleles at QTL on LG	No. of Families or plots [1]	Family Phenot. mean [2] µmol/g	Variance between families [3] µmol/g	No. of Families or plots [4]	Variance within families [5] µmol/g	PV [6] µmol/g	PE [7] µmol/g
BC_3 family	0	25	16.02	3.31	22	8.02	16.02 6c	-0.85
BC_3 family	6	4	16.07	0.41	4	8.99	15.26	-1.61
BC_3 family	3	12	18.98	7.44	9	10.83	21.09	4.22
BC_3 family	2	9	16.58	11.77	9	12.27	16.28	-0.59
	16 H	4	42.91	19.68	4	38.09	42.91 6b	26.04
	16 L	4	17.02	10.00	4	8.79	17.02 6b	0.14
BC_3 family	16	7	31.46	15.10	5	195.09	46.05	29.18
	18 H	4	29.72	12.63	4	17.97	29.72 6b	12.85
	18 L	4	15.52	8.06	4	4.98	15.52 6b	-1.35
BC_3 family	18	5	23.22	9.30				
BC_3 family	9	4	26.29	10.37	2	99.43	35.72	18.84
BC_3 family	2x3	2	19.18	0.91	2	9.85	22.49	5.62
BC_3 family	2x16	4	35.37	31.47	3	178.08	65.41	48.54
BC_3 family	9x16	3	35.44	18.73	1	151.55	46.24	29.37
BC_3 family	16x18	2	35.86	25.74	1	103.36	54.11	37.23
BC_3 family	9x18	1	33.94				53.59	36.72
BC_3 family	2x18	1	22.93		1	53.33	29.01	12.14
BC_3 family	3x18	1	22.36	0.60	1	70.80	21.93	5.06
Samourai	0	27	16.87	4.96	16	8.18	16.87 6c	0.00
Mansholt	all six	5	63.96	7.70	5	50.12	63.96 6c	47.09
DH2078	3	1	21.49		1	17.24	21.49 6c	4.62
BC_1-2078	3	1	21.12		1	10.70	21.12 6c	4.24
DH-1097	2 ?	1	13.28		1	3.07	13.28 6c	-3.59
BC_1-1027	2 ?	1	13.26		1	2.14	13.26 6c	-3.61
DH-2034	6x16	1	49.53		1	20.06	49.53 6c	32.66
DH-2127	6x16	1	50.09		1	27.24	50.09 6c	33.21
BC_1-2127	6x16	1	47.57		1		47.57 6c	30.70

[1] Number of plots used to estimate the phenotypic mean of one type of genotype; in the case of BC_3 families it indicates the number of BC_3 families segregating for the indicated QTL.

[2] Family Phenotypic Mean = the phenotypic mean of the genotypes belonging to this group

[3] This Variance is the variance of the "means" from all BC_3 families (plots) belonging to this group

[4] Number of families used to estimate the mean variance within families

[5] Average of the variances from the BC_3 families (plots)

[5]Variance within families could not be measured when the plot was harvested as a mixture of plants

[6] PV = Phenotypic Value estimated by Model 1 or 2 (in case of BC_3 fam. in which no clear segregation could be found)

[6b] If a clear segregation was observed in the BC_3 fam., the PV was the group value itself

[6c] If the genotype is a DH line, parent or BC_1 the PV is the plot mean (phenotypic mean) already got on the field

[7] PE = Phenotypic Effect = (PV - Samourai) = QTL effect

? = it is not sure that Mansholt´s Allele at LG 2 is present

H = group of genotypes that show high glucosinolate content (comming from segretating BC_3 families)

L = group of genotypes that show low glucosinolate content (comming from segretating BC_3 families)

Tab. 4 Phenotypic Values and Phenotypic effects of genotypes and BC_3 families at Elliehausen 2001/2002.

Genotype	Donor alleles at QTL on LG	No. of Families or plots [1]	Phenot. mean[2] µmol/g	Variance between families[3] µmol/g	No. of Families or plots [4]	Variance within families[5] µmol/g	PV [6] µmol/g	PE [7] µmol/g
BC_3 family	0	18	15.44	2.57	15	5.60	15.44 6c	-1.31
BC_3 family	6	3	17.38	1.55	3	8.85	18.01	1.26
BC_3 family	3	5	17.52	7.57	2	8.40	18.29	1.53
BC_3 family	2	7	14.75	4.28	7	9.40	12.75	-4.00
	16 H	4	37.50	1.48	4	40.46	37.50 6b	20.75
	16 L	4	14.97	1.40	4	10.00	14.97 6b	-1.78
BC_3 family	16	4	25.05	2.42	4	150.63	35.71	18.96
	18 H	1	25.18				25.18 6b	8.43
	18 L	1	15.40				15.40 6b	-1.35
BC_3 family	18	3	22.34	5.76	3	28.34	27.93	11.18
BC_3 family	9	1	25.92		1	86.10	35.10	18.35
BC_3 family	2x3	1	15.99		1	2.65	16.18	-0.57
BC_3 family	2x16	3	30.93	69.16	2	132.42	56.72	39.97
BC_3 family	9x16	2	34.60	10.36	1	120.00	49.01	32.26
BC_3 family	16x18	2	32.40	9.19	1	176.81	47.39	30.64
BC_3 family	2x18	1	21.03		1		26.71	9.96
BC_3 family	3x18	2	22.50	1.52	1	48.77	27.02	10.27
Samourai	0	23	16.75	4.47	16	6.75	16.75 6c	0.00
Mansholt	all six	4	57.79	30.14	4	87.76	57.79 6c	41.04
DH2078	3	1	16.71		1	5.54	16.71 6c	-0.04
BC_1-2078	3	1	17.72				17.72 6c	0.97
DH-1097	2 ?	1	11.81				11.81 6c	-4.94
BC_1-1097	2 ?	1	14.20		1	5.80	14.20 6c	-2.55
DH-2034	6x16	1	42.45		1	33.07	42.45 6c	25.70
DH-2127	6x16	1	47.47		1	46.40	47.47 6c	30.72
BC_1-2127	6x16	1	47.46				47.46 6c	30.71

[1] Number of plots used to estimate the phenotypic mean of one type of genotype; in the case of BC_3 families it indicates the number of BC_3 families segregating for the indicated QTL.

[2] Family Phenotypic Mean = the phenotypic mean of the genotypes belonging to this group

[3] This Variance is the variance of the "means" from all BC_3 families (plots) belonging to this group

[4] Number of families used to estimate the mean variance within families

[5] Average of the variances from the BC_3 families (plots)

[5]Variance within families could not be measured when the plot was harvested as a mixture of plants

[6] PV = Phenotypic Value estimated by Model 1 or 2 (in case of BC_3 fam. in which no clear segregation could be found)

[6b] If a clear segregation was observed in the BC_3 fam., the PV was the group value itself

[6c] If the genotype is a DH line, parent or BC_1 the PV is the plot mean (phenotypic mean) already got on the field

[7] PE = Phenotypic Effect = (PV - Samourai) = QTL effect

? = it is not sure that Mansholt´s Allele at LG 2 is present

H = group of genotypes that show high glucosinolate content (comming from segretating BC_3 families)

L = group of genotypes that show low glucosinolate content (comming from segretating BC_3 families)

In both experiments (Tab. 3 and Tab. 4) the difference between the parents (Mansholt and Samourai) is similar, varying from 47.09 µmol/g seed in Reinshof to 41.04 µmol/g seed in Elliehausen. This difference may be attributed to the variance within families, especially in this case, due to the higher variance (within families) in Mansholt, due probably to environment effects.

The correlation between Reinshof's (Tab. 3) and Elliehausen's (Tab. 4) GSL phenotypic effects (PE) is 0.96, which even allow to compare the two experiments. Also the values for Samourai in Reinshof and Elliehausen (16.87 ± 2.22 and 16.75 ± 2.11) are nearly identical (non significant using t-test; P=0.84), as well as the means of the Mansholt plots in the two experiments (63.96 ± 2.77 and 57.79 ± 5.48; P=0.11 at t-test).

In other hand, in Tab. 5 the general results (means) of both experiments are presented, in which the QTL effects of the six QTL for glucosinolate content were estimated (PE Column in Tab. 5, in bold letter). The phenotypic value of Samourai was found to be not different from the genotypes (no. 4 on Tab. 5) which contain one Mansholt's allele at LG 6, although a difference was found between Samourai and those genotypes equivalent to Samourai (genotypes no. 3 on Tab. 5), i.e. genotypes which do not carry any Mansholt allele in the six QTL. Even insignificant was the difference between Samourai and the genotypes which carry one Mansholt allele at LG 2. Concerning the QTL on LG 3, a significant difference could be identified by t-test when comparing the phenotypic effect of the genotypes carrying this allele from Mansholt with Samourai. Other expected statistic differences could be found comparing Samourai and Mansholt, which is quite obviously, and in the same way the differences should be highly significant if calculated between Samourai and the genotypes carrying Mansholt's alleles for the other 3 major QTL.

Tab. 5 Field Experiment Results (Means) from 2001/2002 in two locations (Reinshof and Elliehausen) under 170 kg N/ha fertilization

No	Genotyp	Donor alleles at QTL on LG No.	No. of Fam.[1] Rein.	No. of Fam.[2] Elli.	µmol/g PV [3]		PE [4]	Var. between fam. [5]	Var. within fam. [6]
3	BC_3 family		25	18	15.73	**	-1.08	2.94	6.81
4	BC_3 family	6	4	3	16.64	ns	**-0.18**	0.98	8.92
5	BC_3 family	3	12	5	19.69	*	**2.88**	7.51	9.62
6	BC_3 family	2	9	7	14.51	ns	**-2.30**	8.03	
		16H	4	4	40.20		**23.39**	10.58	39.28
		16L	4	4	15.99		-0.82	5.70	9.40
	BC_3 family	16	7	4	40.88		24.04	8.76	172.86
		18H	4	1	27.45		10.64		10.94
		18L	4	1	15.46		-1.35		3.71
	BC_3 family	18	4	3	28.83		**12.02**	7.53	28.34
	BC_3 family	9	4	1	35.41		**18.60**		92.77
	BC_3 family	2x3	2	1	19.34		2.52		6.25
	BC_3 family	2x16	4	4	61.06		44.27	50.31	155.25
	BC_3 family	9x16	2	2	47.64		30.83	14.54	135.78
	BC_3 family	16x18	2	2	50.77		33.94	17.46	140.09
	BC_3 family	9x18	1	0	53.59		18.36		
	BC_3 family	2x18	1	1	27.86		11.05		53.33
	BC_3 family	3x18	1	2	24.48		7.67	1.06	70.80
1	Samourai		27	23	16.81		**0.00**	4.71	7.46
2	Mansholt	all	5	4	60.88	**	44.07	18.92	68.94
	DH2078	3	1	1	19.10		2.29		11.39
	BC1-2078	3	1	1	19.42		2.61		10.70
	DH1097	2 ?	1	1	12.55		-4.26		3.07
	BC1-1027	2 ?	1	1	13.73		-3.08		3.97
	DH2034	6x16	1	1	46.13		29.18		26.60
	DH2127	6x16	1	1	48.78		31.97	20.08	36.80
	BC1-2127	6x16	1	1	47.52		30.71		

[1,2] No. of families used in each location for estimation of the mean (not for variance)

[3] PV = Phenotypic Value (= 2 x Fam. Phen. Mean - Samourai) [see chapter 2.2.4]

[4] PE = Phenotypic Effect (PV - Samourai)

[5] Variance between families: Variance of the means of BC_3 families or parents

[6] Variance within families: Mean of the variances between plants belowing to respective BC_3 fam. or plot

? = it is not sure that Mansholt´s Allele at LG 2 is present

H = High level of glucosinolate in this segregating group

L = Low level of glucosinolate in this segregating group

t-test between some genotpypes (PV):

1 and 2	Significant (P=0.01)	1 and 5	Significant (P=0.02)
1 and 3	Significant (P=0.01)	1 and 6	Non-significant (P=0.18)
1 and 4	Non-significant (P=0.68)	3 and 4	Non-significant (P=0.10)

From the 3 major QTL for glucosinolate content, segregation into two clearly separated phenotypic classes could be observed for the QTL on LG 16 (Fig. 16) and the one on LG 18 (Fig. 17). This segregation could be observed in such families in Reinshof and Elliehausen. Therefore, for these two QTL, the phenotypic value (PV) of the group of plants carrying the Mansholt's allele have not been calculated using the model 1, since this group of plants could be identified and separated from the other one which is homozygous for the recurrent parent at this QTL (legend 6b on Tab. 3 and Tab. 4). In the same way, model 1 has also not been applied to the parents (Samourai and Mansholt), to the DH lines, since they do not segregate, and BC_1 genotypes, which also do not segregate because all BC_1 plants in one plot have arisen from a cross between a DH line and Samourai, and therefore are all identically heterozygous (legend 6c on Tab. 3 and Tab. 4).

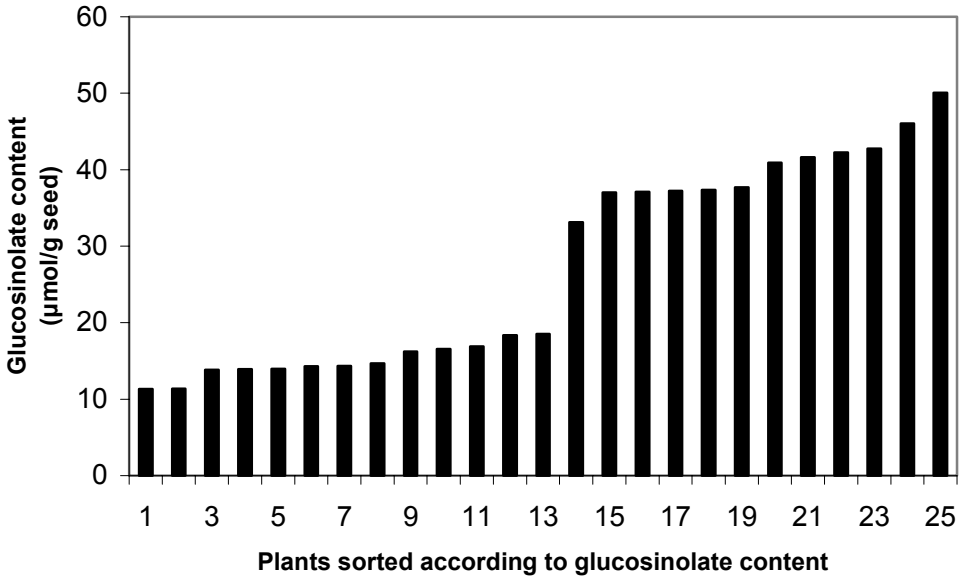


Fig. 16 Distribution of glucosinolate content in BC_3 family 2034-2-18 (in Reinshof) segregating for QTL on linkage group 16.

The QTL present on LG 9 usually did not show a clear segregation, and therefore the formula (model 1) in chapter 2.2.4 was used. With the minor QTL a discrete phenotypic segregation was not expected to appear since the effects of these QTL are small and therefore also "model 1" was applied. Despite of the fact that the genetic segregation is still 1:1, in these cases the segregation is hidden by the cloud of points (not discrete classes). However, in such a cloud of points (see Fig. 6) it is expected to find a higher variance, if one QTL is segregating, than if no QTL is segregating. This is the case for example when the

variance within families (Tab. 5) from those families segregating for the QTL on LG 9 is compared to the variance within Samourai plots, i.e., 92.77 compared to 7.46, respectively.

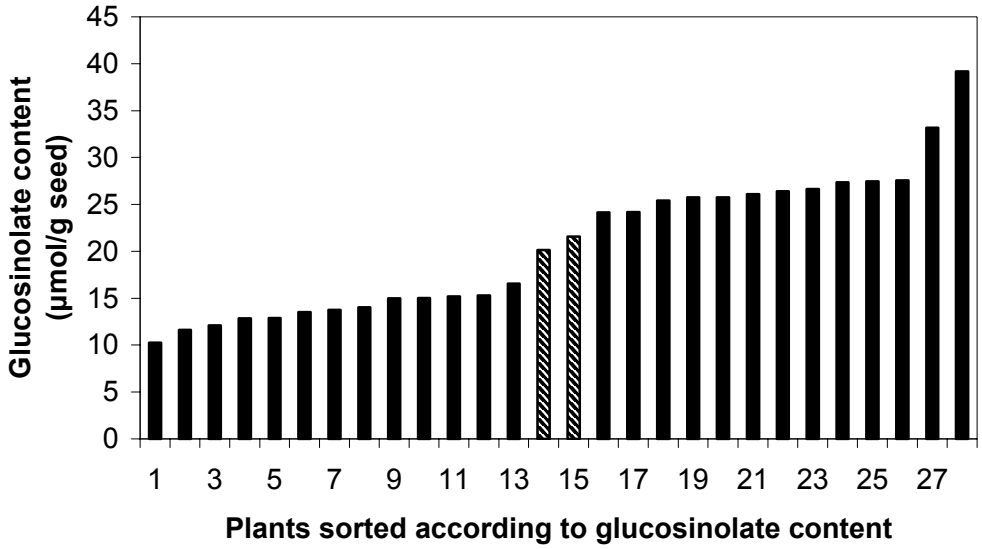


Fig. 17 Almost discrete distribution of glucosinolate content in BC_3 family 1089-2-22 (in Reinshof) segregating for QTL on linkage group 18 (the two striped genotypes in the middle are intermediate ones between the two groups).

3.3.2 GSL data from Reinshof and Ellienhausen compared with literature data

Tab. 6 shows the phenotypic values of both parents in the present study compared to the values got by Gül (2002), which allow to conclude that the values are quite similar, especially when the standard deviation of the present study is considered.

Reviewing previous studies carried out on the same population (Mansholt x Samourai), the QTL effects estimated through composite interval mapping by different authors are summarized in Tab. 7, which even shows the QTL effects estimated and showed in Tab. 5.

Tab. 6 Field Experiment Results from 2001/2002 and 1999/2000

Genotype	Mansholt's alleles present at QTL on linkage group no.	Mean Reinshof/ Elliehausen 2001/02 µmol/g GSL	Mean 1999/2000 (Gül, 2002) µmol/g GSL
Samourai (DH11.4)		16.81 ± 2.17	16.51*
Mansholt (DH5.2)	2, 3, 6, 9, 16, 18	60.88 ± 4.35	65.15*

*Standard deviation not mentioned by Gül (2002).

Tab. 7 QTL effects for glucosinolate content [µmol/g seed] in different studies in the cross Mansholt x Samourai and phenotypic values of the parents

Linkage Group (LG)	QTL effects estimation by interval mapping [1] Uzunova *et al.* [3]	Weiß-leder [4]	Fisher and Ecke [5]	Gül [6]	Mean	Effects in BC_3 fam. [2] 2002	N° of families analysed
Major QTLs							
LG9	4.30	4.64	5.38	6.83	5.28	18.60	4
LG16	7.80	8.29	8.31	10.42	8.70	23.39	4
LG18	3.50	5.90	6.45	6.10	5.48	12.02	4
Minor QTLs							
LG2	3.40				3.40	-2.30	9
LG3	detected[7]	2.46	3.15		2.81	2.88	12
LG6				-2.83	-2.83	-0.18	4
Mansholt [8]	62.30	52.40	51.87	65.15	57.93	60.88	
Samourai [8]	11.20	13.70	13.45	16.51	13.72	16.81	

[1] Additive effect due to the substitution of a "Samourai" allele by a "Mansholt" allele.
[2] Additive + dominance effects in BC_3 families.
[3] 1995, [4] 1996, [5] 1997, [6] 2002
[7] Not significant
[8] Phenotypic values

So, in Tab. 7 it becomes clear that the effect of QTL on LG 16 is the biggest one among the 3 major QTL for glucosinolate, and this in both cases, considering the mean across the four studies presented and in the present study. Comparing the values obtained by the mean of the four previous studies with the one obtained through the BC_3 families, it is clear that a great difference exists between them. Nevertheless, the first ones indicate the additive effect of one Mansholt's allele, and the values got from the BC_3 families involve both the additive effect of such allele and the dominance effect.

3.3.3 Epistasis

Tab. 8 shows the estimated phenotypic value of some genotypes carrying two different QTL in the present study, compared with the estimated values got through the "*additive model*". This model implies that Samourai, as it always has been used in the present study, is once more used as the basic value. Summing up to Samourai's phenotypic value (16.81µmol/g seed) the individual effects of the other QTL (Tab. 5 and Tab. 7), the result is a new estimated phenotypic value. For example, to estimate the glucosinolate phenotypic value (PV) of the genotype 2x16 (which contains one Mansholt's allele on LG 2 and on LG 16), one should sum up 16.81 µmol/g to -2.30 µmol/g and to 23.39 µmol/g, which results in 37.90 µmol/g. This value should be compared to the PV estimated through the BC_3 families. In this case, the data on Tab. 8 demonstrates that some genotypes (marked in Tab. 8 as "E") show significant statistical deviation from the additive model, indicating possible epistatic effects.

Tab. 8 Detection of epistasis by comparisons between phenotypic values of BC_3 genotypes carrying Mansholt's alleles at two QTL and the sum of the values of genotypes with Mansholt's alleles at either one of these loci alone.

QTL present in BC_3 genotype	No. of Genot. Reinsh.	No. of Genot. Ellienh.	Estimated PV on BC_3 fam. µmol/g	Estimation of PV using aditive model µmol/g		Conclusion
2x3	2	1	19.34	17.39		No Epistasis
2x16	4	4	61.06	37.90	E	Epistasis
9x16	2	2	47.64	58.80	E	Epistasis
16x18	2	2	50.77	52.22		No Epistasis
9x18	1	0	53.59	47.43		No Epistasis
2x18	2	1	27.86	26.53		No Epistasis
3x18	1	2	24.48	31.71	E	Epistasis
6x16 [2]	1	1	46.13	40.02	E	Epistasis
6x16 [3]	1	1	48.78	40.02	E	Epistasis

[1] Estimation upon the Additiv model (for QTL on LG 6 the value 16.81 was considered, as a basic value; the effects of the other QTL were taken from Tab. 5 (PE) and have been summed to 16.81
[2] DH2034
[3] DH2127
E = Some possible epistatic effects (Significant in Chi-Square Test, P= 0.005)

In the same way, Tab. 9 shows 37 DH lines evaluated by Gül (2002) which have good marker information for all 6 QTL regions, i.e., for these lines the flanking markers allow

to be sure about which allele is present in all six QTL regions, what is not the case in all 142 DH lines analysed by Gül (2002). Taken Samourai as the basic value (16.51 µmol/g seed) and adding the respective QTL effects estimated in this study (Tab. 5 and Tab. 7), a second column was calculated which shows the estimated phenotypic value (PV). Again like in Tab. 8, some estimated values are far away from the observed phenotypic values (Tab. 9) so that those genotypes were market with an “E”, indicating that possible epistatic effects could be present. A chi-square test proves that the additive model used to estimate the phenotypic value is not statistically different from the values got by Gül (2002) if the genotypes suspected to be influenced by epistasis were taken out of the chi-square test.

3.3.4 Using composite interval mapping on Gül´s data

Tab. 10 show the results of composite interval mapping estimations carried out with Gül’s (2002) original N1 (nitrogen level 1) data (142 DH Lines). Four (Tab. 10a) out of six expected QTL for glucosinolate content could be identified (6, 9, 16 and 18), among them, the 3 major known QTL mapped on LG 9, LG 16 and LG 18. Together, the four QTL found explain 89.7% of the phenotypic variance, and when only the 3 major QTL are taken 86.1% of the variance is explained. Also three epistatic interactions (Tab. 10b) show significance (Prob $_{ij}$ = 0.00), i.e. between LG 7 and LG 14 (F-L); between LG 9 and LG 16 (G-M); between LG 13 and LG 16 (H-N). These three regions explain 9.5% of the total phenotypic variance, and together with the three major QTL estimated explain 99.2% of this variance. It is interesting to observe that, some regions itself, seem not to influence the glucosinolate content, like for example the marker intervals F and L, respectively LG 7 and 14, but both together have a significant interaction (Prob. $_{ij}$), what points to epistasis. Even when one LG like 16 is involved, for example in the case of the interaction between intervals G and M, it is important to note that in this case the interval which refers to LG 16 (M) is not the same interval in which the main isolated effect from the major QTL on LG 16 (C) is localized. So, one has to be aware that some regions not identified before by any study or even by Gül (2002) seem to play a role, at least at the epistatic level, in the glucosinolate content. Once more, even in Gül’s data, some epistatic effects could be found corroborating in some way with the data presented in Tab. 8 and 9, which also demonstrate the presence of epistatic effects between different QTL.

Tab. 9 Glucosinolate content of different DH lines (cross Mansholt x Samourai) evaluated by Gül (2002). Estimation from PV was done using the Allele Effects from the present study (Tab. 7).

Genotyp	Donor alleles at QTL on LG No.	No. of DH Lines 1999/00 (Gül)	Mean (Gül) 1999/00 DH Lines GSL µmol/g	Estimation of PV using aditive model[1]	
Samourai	0		16.51		
Mansholt	2,3,6,9,16,18		65.15	70.92	
	6	2	18.32	16.33	
	3x6	1	15.18	19.21	
	3	1	24.12	19.39	
	18	1	42.05	28.53	
	6x18	1	34.96	28.35	
	9	1	38.51	35.11	
	6x9x16	3	62.76	58.32	
	2x16x18	1	51.62	49.62	
	6x16x18	1	56.26	51.74	
	9x18	1	53.81	47.12	
	2x18	3	36.78	28.53	
	3x9x18	2	55.94	50.00	
	3x9	4	42.91	37.98	
	2x9	1	35.51	32.81	
	2x3x9	1	33.39	37.98	
	6x9x16x18	1	65.42	70.34	
	2x3x9x16x18	1	62.50	71.09	
	2x9x16x18	1	71.46	68.22	
	3x9x16x18	1	67.67	73.39	
DH2034	6x16	1	55.22	39.73	E
DH2127	6x16	1	59.57	39.73	E
	2x3x9x18	2	57.75	47.70	E
	2x6x9x16x18	2	65.92	49.45	E
	2x16x18	1	64.08	49.62	E
	2x3x18	1	40.15	29.10	E
	2x3x16	1	58.83	40.48	E

Aditive Model[1]: model based on Effects calculated at Table 5
Aditive Model[1]: PV of Samourai + QTL effect from QTL present, as example,
Mansholt is the sum of 16.51+(-0.18)+2.88+(-2.30)+23.39+12.02+18.60 = 70.92
E = deviations from additive model due to probably epistatic effects
Chi-Square Test with all data: ** (P=0.0001), i.e., obtained data differ from expected (model[1]) data.
Chi-Square Test without "E" data: non significant (P=0.44), i.e, obtained data did not differ from expected (additive model) data.

Tab. 10 Composite interval mapping using data from Gül (2002) at nitrogen level N1: additive effects of marked QTL have been estimated (10a) and epistatic interactions (10b)

Tab. 10a

LG_i	$Interval_i$	Site cM	LOD	A	Prob
6	A	2	5.66	-2.83	0.00
9	B	2	20.03	6.83	0.00
16	C	2	33.63	10.42	0.00
18	D	2	17.56	6.10	0.00

Tab. 10b

LG_i	$Interval_i$	$Site_i$ cM	LG_j	$Interval_j$	$Site_j$ cM	LOD	A_i	$Prob_i$	A_j	$Prob_j$	AA_{ij}	$Prob_{ij}$
5	E	0	6	A	2	5.80	-0.40	0.45	-2.77	0.00	-0.15	0.78
7	F	4	14	K	16	5.33	-0.43	0.47	0.58	0.34	3.15	0.00
9	G	0	16	L	14	19.58	6.23	0.00	-0.33	0.59	-2.52	0.00
13	H	34	16	C	2	29.69	0.15	0.81	10.36	0.00	0.51	0.39
15	I	0	18	D	0	19.25	1.24	0.03	5.76	0.00	-2.23	0.00

Marker intervals:

A	MG23-OD3.1055	E	MG18-MG19	I	R150.E1-R1362.E1
B	R1100.E1-R825.H1	F	R1202.H1-R318b.E1	K	W4D10.E1-OAI16.14
C	W3F7.H1-R1175.H1	G	R1100.E1-R825.H1	L	R1516.E1-R1360.E1
D	W7A8.H1-W4E12.H1	H	cRT21.E1-OAJ10.66		

LG = linkage group

Interval = marker interval

Site = position of QTL in cM from lefthand marker of the indicated interval

LOD = LOD Score

A or A i = additive effect of QTL at site "i"

$Prob_i$ = statistical significance of the additive effect at site "i"

A_j = additive effect of QTL at site "j"

$Prob_j$ = statistical significance from the additive effect of site "j"

A_{ij} = epistatic effects (interaction) of loci at site "i x j"

$Prob_{ij}$ = statistical significance of the epistatic effect of the interaction between loci at site "i x j"

3.3.5 Using HPLC to analyse the two trials of 2002 (Reinshof and Elliehausen)

Tab. 11 presents results of HPLC analyses which have been carried out in part of the material harvested in 2002 in the two field trails. The table shows different comparisons between different materials and the aim was to find out whether the two groups compared are different in the amount of glucosinolate types produced or in the percentage of production of each one. The first comparison, between Samourai plots and BC_3 families containing neither Mansholt allele in the six QTL, show that, contrary to the results obtained by NIRS (Tab. 5), this two groups show not statistical differences what concerns to the total

glucosinolate content. Nevertheless some significant differences appear when the glucosinolate types are analysed one by one, be it concerning the amount (µmol/g) or the percentage of each type. In this case, 6 components show significant changes between these two groups.

Tab. 11 Comparisons of HPLC analyses of glucosinolates between different materials from the two trials of 2002 (Reinshof and Elliehausen) tested upon statistical differences using the t-test.

Kind of Comparison		n°	PRO	SIN	GNL	ALY	GNA	4OH	GBN	ERU	GBC	NAS	4ME	NEO	HPLC[2] Total umol/g
Samourai **vs** 0 QTL[1]			*	ns	ns	*	ns	ns	ns	ns	ns	ns	**	**	ns
Samourai	Mean (µmol/g)	41	9.22	0.28	0.23	0.16	4.33	3.62	0.63	0.03	0.45	0.47	0.07	0.01	19.53
O QTL[1]	Mean (µmol/g)	40	8.35	0.38	0.18	0.31	4.50	3.42	0.46	0.02	0.44	0.42	0.12	0.00	18.61
			*	*	ns	*	**	ns	ns	ns	ns	ns	**	*	
Samourai	Mean (%)	41	47.38	1.41	1.21	0.83	22.28	18.37	3.06	0.15	2.31	2.43	0.35	0.08	100
O QTL[1]	Mean (%)	40	44.97	2.04	0.97	1.60	24.18	18.47	2.26	0.10	2.41	2.26	0.63	0.00	100
QTL LG 16 **vs** QTL on LG 18 (only Reinshof)			**	ns	ns		**	ns	*		ns	*	ns	ns	**
QTL 16	Mean (µmol/g)	4	28.05	0.73	1.28	0.00	13.28	2.58	2.45	0.00	0.33	0.95	0.13	0.05	49.85
QTL 18	Mean (µmol/g)	4	16.68	1.85	0.68	0.00	10.03	2.13	1.08	0.00	0.30	0.68	0.10	0.08	33.60
			ns	ns	ns		*	ns	ns		ns	ns	ns	ns	
QTL 16	Mean (%)	4	56.27	1.46	2.48	0.00	26.76	5.14	4.89	0.00	0.65	1.90	0.25	0.09	100
QTL 18	Mean (%)	4	49.69	5.36	1.96	0.00	29.87	6.52	3.12	0.00	0.90	2.00	0.30	0.22	100
Mansholt **vs** Samourai			**	ns	**	**	**	ns	**		ns	*	ns	ns	**
Mansholt	Mean (µmol/g)	9	40.01	0.86	2.17	0.00	15.58	2.94	4.18	0.02	0.57	0.73	0.09	0.03	67.20
Samourai	Mean (µmol/g)	40	9.22	0.28	0.23	0.16	4.33	3.62	0.63	0.03	0.45	0.47	0.07	0.01	19.53
			**	ns	*	**	ns	**	**	**	**	**	**	ns	
Mansholt	Mean (%)	9	59.49	1.23	3.12	0.00	23.45	4.37	6.16	0.03	0.86	1.08	0.13	0.05	100
Samourai	Mean (%)	40	47.38	1.41	1.21	0.83	22.28	18.37	3.06	0.15	2.31	2.43	0.35	0.08	100

[1] Group of plants denominated as n° 3 in the Tab. 5 (carry neither allele of Mansholt at the six QTL)
[2] umol GSL/g seed

T-Test:

ns	Non-significant	PRO	Progoitrin	GNA	Gluconapin	GBC	Glucobrassicin
*	Significant at 0.05	SIN	Sinitrin	4OH	4-Hydroxyglucobrassicin	NAS	Gluconasturtiin
**	Significant at 0.01	GNL	Gluconapoleiferin	GBN	Glucobrassicanapin	4ME	4-Methoxyglucobrassicin
N° = number of genotypes		ALY	Glucoalyssin	ERU	Glucoerucin	NEO	Neoglucobrassicin

Concerning the differences between the Mansholt's alleles at the QTL on LG 16 and LG 18 it can be seen (Tab. 11) that there are no large different glucosinolate profile between these two QTL, since generally the QTL on LG 16 is producing larger amounts of progoitrin, gluconapin, glucobrassicanapin and gluconasturtiin, but only gluconapin show a different rate of production.

When Samourai and Mansholt are compared only sinitrin and neoglucobrassicin do not alter both the amount and the percentage rate. Other comparisons, although they could be very interesting, could not be performed since with exception of the groups classified as H (high, i.e., heterozygote) in the LG 16 and LG 18, in the other BC_3 families the genotypes carrying the Mansholt's allele could not be separated from those one which do not carry them, and therefore could not be separately analysed and compared.

The important thing is not to stop questioning.

Albert Einstein (1879-1955)

4. DISCUSSION

4.1 Codominant scoring of AFLP markers

As commented in the results, with regard to the codominant scoring of AFLP markers, i.e. with the dominant allele coming from the recurrent parent, it has proven to be difficult to perform the codominant analysis, despite of the use of an appropriate software. AFLP-Quantar™ *Pro* is a software based on the theoretical premise that AFLP technology allows the quantitative measurement of the degree of amplification of a fragment, called "band intensity" (Jansen *et al.*, 2001) or "optical density" (Piepho and Koch, 2000). The major problem of utilizing such a software is that it is not possible to really verify the accuracy of the results, since the only way to do so would be to confirm such results by using other truly codominant markers (SSR, RFLP) or by selfing the genotypes that are supposed to be heterozygous (or even the ones that are supposed to be homozygous) to confirm the genotyping done by the software by segregation analysis in the progeny. It was tried always to look upon the flanking markers (if there were some) to verify the probability that the codominant scores were coherent or not, but even this is not a reliable way to do it, because it depends on how far reliable "donor markers" are from one "recurrent marker", which on the other hand influences the recombination rate. Nevertheless, sometimes even such kind of procedure was impossible since in some regions of the genome only recurrent markers were available. All the points stressed out demonstrate that the results coming from the software will contain errors which can only be detected in the next level of this project. Nowadays there are very few reports in the literature showing that a codominant analysis of AFLP markers has been successful (Castiglioni *et al.*, 1999)

Some other authors have tried, through different ways, to use AFLP markers (Piepho and Koch, 2000) and even other typically dominant markers, like RAPD, as codominant ones (Vandemark and Miklas, 2002). The first study has used the optical density of the band on the gel and by this way used statistical approaches to find out the heterozygous genotypes. The second one has used a better way to quantify directly the PCR product by using real-time detection of fluorescent-labelled DNA fragments which contains a passive reference dye.

4.2 Development of substitution lines

Several studies (Hospital *et al.* 1992 and Visscher *et al.* 1996, cited after Servin and Hospital, 2002) have shown that few markers, i.e. 2-4 markers/Morgan, are necessary to control genetic background in marker assisted backcrossing. According to Visscher (1996), for a randomly mating backcrossing population, it was shown that nearly all the variance can be explained by placing three or more markers per chromosome. Yet, more than 2-4 markers/Morgan are generally available, which is also the case in this study, since there are marker information on 130 loci available in BC_2 and 120 available in BC_3 (Tab. 1). When the BC_3 generation is considered, a mean number of 6 marker loci (=map position) per chromosome is available in the present study (Tab. 1 and 2), which results in 6.6 marker loci per Morgan (120 loci/18 Morgan).

On the other hand, not all linkage groups shown in Appendix 9.7 and 9.8, at least not all regions of those linkage groups, are covered by such a good number of marker loci, but when the linkage groups are simply classified in three groups according to the coverage by markers (good covered by markers, sufficiently covered, and poorly covered), then only two linkage groups (7, 11) have a unsatisfying situation because for those groups only one marker locus (= map position) is available. From the other 17 linkage groups, 5 could be classified as “sufficiently covered” (5, 6, 16, 17 and 19), so that in the other 12 a good situation was found with respect to the coverage by markers (linkage group 20 was not considered in one of these 3 groups because it is too small). Hence, the situation that could be found in this study, i.e., 120 map positions (1,325 cM) that should cover a total genome of about 1,800 cM, is good, concerning the possibilities of selection against remaining donor segments.

It is evident that not only the number of markers present in each linkage group is important, but the distribution and localization of them across the linkage group. Hence, what is really important is not whether markers will be fixed or not through the backcrosses, but how much of the genome outside the markers will be fixed for recurrent type by the time the markers are fixed. Based on this considerations Servin and Hospital (2002) have proposed to define the “optimal marker positions” as the positions that maximize the genome-wide proportion of loci that are fixed for homozygous recipient type once the markers are fixed for homozygous recipient type (i.e., selection on markers has been successful). This was evaluated by these authors as the expected probability that any locus on the genome is of recipient type, given that all markers are of recipient type. According to these authors the optimal position of two markers on a chromosome of 100 cM are about 20 cM from the telomeres (from 18.6 cM in BC_1 to 22.9 cM in BC_3).

In the present study the absolute majority of the linkage groups have a length about 100 cM, and with the exception of LG 7, 11, 16 and 19, all linkage groups (Appendix 9.7 and 9.8) have more than two marker loci, although not always localized at the optimal positions mentioned above. Using more markers on a chromosome leads to better control of the return to the recipient genome because the regions controlled by the markers overlap. Thus better control of the genomic background can be achieved either by using more markers, that can be sub optimally placed, or by using fewer markers, optimally placed. Besides, for a given number of markers, the impact of suboptimal positioning of markers is less important when the backcross generation is more advanced (Servin and Hospital, 2002).

The proportion of recipient genome that can be recovered in marker assisted selection with optimally placed markers compared to the average values obtained when no selection on markers is performed is high. For example, a noncarrier chromosome presenting three optimal placed markers that are of recipient type in BC_3 will have 99.2% of recipient genome. Without selection on markers, the same return of the recipient genome would be obtained only in the BC_6 (Servin and Hospital, 2002). Based on this theoretical point of view, at the present study a good performance of the 120 to 130 map positions is expected, and a good control of the recipient genome is evident.

The aim of a backcross selection programme, which is essentially not different from the one applied in this study to form the substitution lines, is that in any locus the gene (in the present case, the segment) introgressed from the donor line returns to a homozygous recipient type. Even without background selection on markers, this is just a matter of time (i.e., of the number of backcross generations). The aim of selection on markers is to go faster toward fixation than without selection on markers, like explained even in the example above (Servin and Hospital, 2002). What concerns this aspect, i.e., the amount of donor genome still present, a normal backcross programme without selection should result in the presence of 6.25% donor genome at the BC_3 generation. That means that in the BC_3 generation 12.5% of the loci have still one donor allele present, since in the backcross genotypes the donor segments present are in heterozygous state, i.e., still 12.5% of the genome in a BC_3 population is expected to be covered by donor segments. Indeed, this value is much higher than the values can be found in Tab. 2 (3.8%). On the other hand, a significant amount of ‘unknown’ type of genome segments (3.75% in Tab. 2) still remains present in the 52 selected BC_3 plants in the present study. Assuming that from this 3.75%, theoretically 12.5% (12.5% from 3.75 = 0.47) could be donor genome, than an amount of 4.27% (3.8+0.47) of donor genome is possible to be covering these 52 plants. This value is lower than 12.5%, which is the amount that corresponds to an entire BC_3 population, while 4.27% is related only to the 52 selected BC_3 plants, and not to the entire BC_3 population used in this study.

If the whole BC_3 population would be analysed concerning the amount of donor and recurrent genome, it is possible therefore that a higher amount of genome covered by donor segments would be found. If without MAS an amount of 12.5% of the loci being still heterozygous (6.25% of donor genome present) is theoretically expected, with MAS this value should be lower. Since it is exactly what could be observed (Tab. 2) the conclusion is that MAS is effective in reducing the donor genome in the backcross generations. However, it is important to remind that the aim of the present project is not only to obtain lines (through backcrossing), which will contain some target genes, but the aim is to get a set of lines, in which all the donor segments together should represent the whole donor genome. This is a more complex situation than simple transfer of one gene and recovering the recurrent genome, since during MAS not a simple selection against the donor genome can be carried out, but it should be done keeping in mind that the selected genotypes, together, should represent the whole donor genome. This aim to get the entire donor genome to be covered during the MAS provoke the unavoidable selection of many genotypes containing larger donor segments, and those genotypes therefore may still inflate the amount of donor genome present despite of MAS, i.e., the donor genome present could be even smaller if the aim of getting substitution lines were not realized.

The present results are even in fact not exactly comparable to the values obtained by Ribaut *et al.* (2002), since these authors have worked with special conditions at computer simulations, but they have got results in BC_3 ranging from 1.5% to 15.3% of donor genome. The first step, in Ribaut *et al.* (2002) study, was to identify the genotypes that are heterozygous at the target loci, reducing the screened populations size 'N' to the 'N*sl* '. The second step was to identify within the N*sl* individuals those presenting the most suitable genomic composition at the nontarget loci, i.e., individuals carrying lesser donor genome and still carrying the target gene were preferred. The simulations done by these authors were performed for the introgression of one target gene. Obtaining 1.5% donor genome had required screening a population of 200 individuals (N) at a single locus, followed by a screening of 109 markers (11 per chromosome) at nontarget loci at each BC on the 100 N*sl* genotypes. When the MAS on nontarget genes is not done in all BC generations, then 15.3% of donor genome is expected at the end (for one target gene), i.e. more then the usually 6.25% of donor genome expected in a BC_3 generation if no MAS is done at all (no selection). The result of 1.5% of donor genome (3% of the genome covered by donor segments, i.e. 3% of the loci having one donor allele in heterozygous state) was not reached in the present study, but it should even not be achieved since in the simulations (Ribaut *et al.,* 2002) size populations of 200 (N) individuals and N*sl* populations of 100 individuals were used in each BC generation. Since in the present study not only one target gene is still under focus, but many target regions are tried to be maintained throughout the BC generations so that the

entire donor genome is covered, a quite different situation is expected, so that the 3.8% donor genome coverage got in the selected BC_3 plants is not only acceptable, compared to 3% (above), but could be considered a good result. Such result is especially acceptable if it is compared to the expected 12.5% which should cover the genome if no selection is done through the backcross generations.

Concerning the average number of donor segments present in plants along the backcross generations, compared to a previous study done by Cermakova *et al.* (1999), the number of introgressed segments was larger than expected. Cermakova *et al.* (1999) have found 5-10 donor segments present in BC_2 families, through 1-6 present in BC_3 families, to a single segment in subsequent generations. In the study carried out using the Mansholt x Samourai population, the mean number of donor fragments present were 11.22 at the BC_2 generation (Fig. 13) and 7.1 donor fragments were still present in the BC_3 (Fig. 14). This difference may be due to the use of a relatively short (450 cM, Sharpe and Lydiate, 2003) RFLP map comprising only 77 polymorphic marker loci by Cermakova *et al.* (1999), which can lead to mistakes by not recognizing donor segments still present. In other hand, when the pedigree of the two parents, which contrast for glucosinolate content, are analysed, they are found to be related. This explains partially the short map obtained and used. Since in the present work a more detailed map was used, with 130 loci at the BC_2 and 120 at the BC_3 generation, corresponding to 1,325 cM of the rapeseed genome, it can be expected that more segments from the donor genome can be identified, and therefore the values presented here are higher.

When the present data are compared to that one obtained by Howell *et al.* (1996), in which 158 RFLP loci result in a map of 1,204.5 cM, (equivalent to 66% of the rapeseed genome, according to these authors) it is surprising to verify that these authors could find genotypes carrying only one donor segment in BC_3 generation, which was not the case in the present study. This happened probably because already in the BC_1 a set of selected genotypes was selected in that two small regions of the genome were missing. Later on, in the BC_2 generation, the selection was made disregarding linkage groups 9 and 12 and part of linkage group 5, which were not represented by donor segments in the set of selected BC_2 plants. From the 17 BC_2 selected genotypes that were advanced to BC_3, only 8 were used to originate substitution lines, and more 2 others originally not selected were used. From these 10 BC_3 families, 19 out of 250 BC_3 plants were selected to originate substitution lines, however they were selected disregarding almost 9 entire linkage groups, which are then definitively not represented by donor segments. This led the authors to the situation and conclusion that these BC_3 individual genotypes constituted heterozygous substitution lines that together spanned 33% of the mapped genome. By such a reduction of coverage, it is clear that lines carrying only one segment of the donor genome will be reached earlier than

in the present study, in which all segments defined by the 120-130 loci (1,325 cM) used in the selection have been retained throughout the BC_1 to BC_3 generations, representing 100% of the mapped genome.

Another point is that from the 288 BC_2 plants in Howell *et al.* (1996), 17 have five or less donor segments present, although in the present study only four such plants could be found in the BC_2 generation (Fig. 13). A reason for having so few of such plants could be the codominant analysis of AFLP markers, which could led to errors in the score procedure, i.e., scoring one marker locus as being heterozygous instead of scoring it correctly, that means, as a homozygous. Considering that only 26% of the codominantly scored markers in the BC_2 could be scored without major ambiguities, indicates that the results of many markers are not free from errors. If the codominant analysis mostly makes mistakes in this direction this can be speculated as the reason for the few number of plants with five or less segments in the present study. In this way many plants classified as having 6 or 7 segments of donor genome could in fact have only five or even less than five segments present.

In the study done by Ramsay *et al.* (1996) in *Brassica oleracea*, it was possible to reach genotypes containing only one segment of donor genome in the BC_2S_1 population, but in this case, although the entire donor genome could be represented through the selected BC_1 individuals, only 82.6% of the donor genome was represented by the 77 selected BC_2 genotypes, which were carrying at that time only one or two donor segments in their genome. Later on, the selfs of a subset of these families, together with other BC_2S_1 and BC_2S_2 families, enable the development of recombinant backcross lines with single homozygous introgressed tracts. These lines cover evidently not more than 82.6%, and even do not cover the entire linkage group 6. Since the genetic map of *B. oleracea* has only a length of 747 cM (smaller than *B. napus*), and taking in consideration the selective choice of some parts of the genome to be represented, it is clear that lines with only one donor segment will be easier obtained in earlier backcross generations.

Another recent study (Rae *et al.*, 1999) could get substitution lines after selfing two or three times the BC_2 plants. In this case, a set of 79 substitution lines was produced in which each line contains from 1 to 4 introgressed segments, together representing a maximum of 91% of the donor genome. Indeed, substitution lines carrying more segments of donor genome are also an approach of the original concept, in which each line should carry only one segment.

In this way the number of donor segments obtained in the present study through the backcross generations, although different from the literature, are coherent, since in the present study the whole mapped genome has been covered. The studies cited above could reach genotypes carrying only one donor segment at BC_3 generation but the selected groups

of genotypes did not cover the whole genome, which was the aim maintained in this study, and therefore a different situation was encountered in BC_3, i.e. that no plants with only one donor segment could be identified in this generation.

Coming back to the substitution lines and summarizing, up to present stage (BC_3 selected plants, from which BC_4 seeds are still available), the best way to reach a complete set of substitution lines in the Mansholt x Samourai population is to analyse with markers the BC_4 generation, which has already been produced by backcrossing the BC_3 plants to Samourai. Once done before flowering, the selected plants carrying only one or two segments of donor genome should be used to produce a set of DH lines via microspore culture. The use of DH lines is recommended by Ramsay *et al.* (1996) because they have observed that some segregation is still present in BC_2S_1 lines homozygous for the donor segment, so that lines coming from the microspore will guarantee their homozygosity and avoid such kind of problems.

The BC_4 generation would be analysed with the AFLP markers only scoring the markers where the dominant allele is coming from the donor parent, since the codominant scoring of the other markers have the difficulties explained before. Since a sufficiently good codominant analysis demands at least around 15 BC_4 individuals per selected BC_3 plant, a big population of BC_4 plants (780 plants) should be analysed, what exceed the capacities of the proposed project. In this way, the use of DH lines made directly from the selected BC_4 plants should be the best way to obtain the set of substitution lines. Besides, if it would be decided to originate the substitution lines by the original proposed way, using selfing progenies like explained in Fig. 2, it will lead to obligatory codominant analysis of all the AFLP markers, since they all will segregate in the 1:2:1 ratio. Since this will be avoided, the best way is to create DH lines from the selected BC_4 plants.

In this way, it is probably better to take the 26 selected BC_3 plants that have four or less donor segments, analyse the BC_4 offspring of them and from those BC_4 ones select the candidate genotypes from which DH lines should be won. At that time, a good estimation can be carried out (based on the BC_4 marker analysis) to verify which donor regions from genome are not covered. Then, additional DH lines can be developed from the BC_5 plants coming from the 26 BC_3 selected plants that still have 5 to 8 donor segments aiming to cover such regions. In BC_5 or more precisely, in the DH lines coming from the BC_5 plants, one can probably get genotypes with only one donor segment.

Many of the donor segments are twice or more represented in the 52 BC_3 plants, so that the development of some DH lines per selected BC_4 and BC_5 plants should be enough to represent all segments in the lines, covering the whole genome. The great advantage of the use of DH lines is that all the AFLP markers can be scored without problems since that

the AFLP markers will be only used as dominant ones. If at the end, the set could not be completed because some region(s) are still not covered, it will be time consuming but it is relatively easy to go back and rescue more lines from new growing BC_4 plants, since there are enough BC_4 seeds from the selected 52 BC_3 plants remaining.

4.3 The effect estimation of QTL for glucosinolate content

In the present study a BC_2 population derived from the cross with Samourai as recurrent parent was genetically characterized using 163 markers. Based on this marker information BC_3 families segregating for one and two of the six know QTL for total glucosinolate content were selected. The BC_3 plants were grown in field trials and after harvest seed glucosinolate content was evaluated using NIRS and HPLC. It is important to stress out that the two methods used for measure the glucosinolate content showed good correlation, that means, NIRS and HPLC has given a correlation of 0.94.

Considering the data shown in Tab. 3, Tab. 4, Tab. 5 and Tab. 6:

1. The data show very good relationship, what concerns the values from the parents Samourai and Mansholt, which are very similar across the 2 different data sets (Reinshof + Elliehausen, and Gül), principally when the standard deviation (Tab. 6) is considered. In this case, the values got in 2001/2002 reach the ones got in 1999/2000 indicating that the deviation is probably not significant. Moreover, Samourai and Mansholt are not significantly different across Reinshof and Elliehausen (t-test).
2. The heritability of glucosinolate is high (0.96 by Gül 2002, 0.97 by Weißleder 1996 and between 0.87 to 0.96 by pers. com. Fischer and Ecke), so, glucosinolate content usually tends to show small interaction between genotypes and environments, although it has sometimes been reported (Gül, 2002).
3. Supposing that only the 6 known QTL reported are responsible for the GSL content.

Accepting as true the points cited above it should be possible to consider that data from at least Reinshof (2001/02) and Elliehausen (2001/02) are comparable. With respect to Gül's data, it is also in part comparable to the results obtained in 2001/02, but one has to be aware by any comparisons between BC genotypes and DH lines that the last ones have probably a much larger amount of donor genome whereas the BC_3 plants should have an

average amount of 6.25%. Even when they still carry the same QTL for glucosinolate content, the DH lines carry much more donor segments than the BC_3 plants, offering the possibility that in those segments still other minor genes which have been not mapped can be found, and therefore could act modifying the final GSL content.

Starting the specific comments on the QTL effects, QTL on LG 16 shows the largest effect and therefore, the best segregation pattern. It was always possible to find either in 4 BC_3 families analysed in Reinshof and Elliehausen (2001/02) the clear 1 : 1 expected segregation (Fig. 16), i.e., since the BC_2 mother plants of such families were carrying the Mansholt's allele on this QTL, the two groups could be found, one with the allele from Mansholt present at this QTL and the other group without this allele.

The mean from the plants carrying Mansholt's allele at the QTL on LG 16 is 40.2 µmol GSL/g (Tab. 5). Subtracting the value from Samourai (16.81 µmol GSL/g), an effect of 23.39 µmol GSL/g seed is achieved (Tab. 5). It has to be pointed out that such genotypes are of course heterozygous for the QTL 16, and therefore they are called "16H" (16 high GSL). The mean additive effect calculated by interval mapping was 8.7 µmol GSL/g seed for one allele of Mansholt. When this value is multiplied by two, the GSL content (17.4 µmol GSL/g) is still less than the result obtained in this study, whereas, indeed, 23.39 µmol GSL/g is the effect of only one allele of QTL 16. This result could indicate overdominance. Still observing the data on Tab. 5, the mean value of two DH lines carrying the QTL on LG 16 is also interesting. Those DH lines contain 47.45 µmol GSL/g seed (mean of 46.13 and 48.78), having two alleles of QTL 16 (i.e., homozygous state). Specifically the DH-2127 contains 48.78 ± 4.48 µmol GSL/g seed whereas the respective BC_1-2127, carrying only one allele of QTL on LG 16, contains 47.52 ± 3.25 µmol GSL/g seed. As conclusion, this result supports the existence of complete dominance in this locus, and not overdominance. In this way, one may also ask why then the value from the "16H" (16 high) genotypes (40.20 µmol GSL/g) is smaller then the BC_1-2127 (47.52 µmol GSL/g) containing the same allele for this QTL and one more, at the QTL on LG 6, but this last one has near to no effect (Tab. 5), as it will detailed later on. One reason could be that a BC_1 plant, and even his respective DH line, indeed contains much more amount of donor genome than BC_3 plants, so, this difference may be due to unknown additional genes in Mansholt which alleles increase the GSL content (Tab. 5) since QTL 6 has near to no effect, as said before. Another explanation for the higher results of the BC_1 and DH lines is that epistatic effects could exist between the alleles at the QTL on LG 16 and LG 6, which already has been observed also in Gül's data (Tab. 9, DH2034 and DH2127, marked as "E"). Such epistatic effects will be discussed on the following pages. On the other hand, the simple standard deviation (Var. between families) at DH2127 (Tab. 5) and the "16H" group (Tab. 5) could be the reason for this difference, in this

case excluding the hypothesis that other alleles from Mansholt are influencing the glucosinolate content.

In that way, comparing the estimation of this QTL effect (LG 16) through the BC_3 and the estimated value (mean among studies) derived by interval mapping (Tab. 7), overdominance or epistasis could be the hypothesis. On the other hand, looking after the comparisons of DH lines and BC_1 genotypes, dominance should be postulated as the best explanation.

Looking to the next QTL, the effect from the Mansholt allele on the QTL at LG 18 is 12.02 µmol GSL/g seed. Genotypes carrying one Mansholt's allele for this QTL have an estimated glucosinolate content of 28.83 µmol GSL/g seed (Tab. 5), i.e., one QTL effect from 12.02. This is almost the same effect estimated by composite interval mapping which was found to have an additive effect from 5.48 µmol GSL/g seed for one allele, since 2 alleles of 5.48 µmol GSL/g result in a phenotypic value for the homozygous class from 10.96 µmol GSL/g, which is similar to 12.02 µmol GSL/g. This will imply in the presence of complete dominance intra allelic interaction at this locus. In other hand some doubts remains present when the phenotype of BC_3 families (28.83 µmol GSL/g) is compared to one DH Line (42.05 µmol GSL/g) from Gül (2002) which is homozygote for this QTL (Tab. 9). Summing the phenotypic value of the heterozygotes (28.83 µmol GSL/g) and effect of one more allele (12.02 µmol GSL/g) a result of 40.85 µmol GSL/g is obtained, which should to be the phenotypic value of the homozygotes. This value seems to be apparently similar to 42.05 µmol GSL/g (Tab. 9), which on other hand is the value of the homozygote genotype (DH line) measured by Gül (2002) and presented in Tab. 9. In this way of looking across the data, partial dominance could be postulated. Another comparison that endorse such conclusion, and even point to complete dominance at QTL on LG 18 as well, is when a BC_3 genotype which carries Mansholt's allele at the QTL on LG 9 and 18 (Tab. 5) have approximately the same glucosinolate content (53.59 µmol GSL/g seed) as the respective DH line (53.81 µmol GSL/g seed) carrying such alleles (Tab. 9). So, the doubt between complete and partial dominance at the QTL on LG 18 will remain.

The effect from the Mansholt's allele at the QTL on LG 9 is 18.60 µmol GSL/g seed. The dominant intra allelic interaction is confirmed when the phenotype (35.41 µmol GSL/g) of the genotypes carrying Mansholt's allele at the QTL on LG 9 (Tab. 5) is compared to one DH line evaluated by Gül (2002) with the same QTL composition, which has 38.51 µmol GSL/g seed (Tab. 9). But of course, comparisons between BC_3 families and DH lines have to be taken with care for the reasons already mentioned. The allele from Mansholt at QTL on LG 9 do not provoke a clear 1 : 1 segregation within the BC_3 families, as it could be observed for QTL 16 and QTL 18, but instead, a clearly much higher variance (within families) could be

observed in the families segregating for the Mansholt's allele at QTL on LG 9, when compared to the variance present in Samourai (Tab. 3 and 4).

The intra allelic interaction from the three major QTL seems to vary from overdominance to partial dominance, but could not be finally clarified since some interesting DH lines have not been evaluated together with the BC_3 families. Nevertheless also Weißleder (1996) has pointed out to partial dominance in this character, but has taken his conclusions based on F_1 data, and therefore did not offer estimations on the different QTL, what could be in part possible at the present study. Also Rücker and Röbbelen (1994) report dominant gene action for glucosinolate content in crosses between genotypes with high and low GSL content. Furthermore, much better information can arise when substitution lines are available, especially if some of them carry Mansholt's segments of two QTL for GSL. Also F_1 between substitution lines carrying Mansholt's alleles for different QTL can be tested on the field. Such study could produce better and more reliable results.

Moving to the minor QTL for glucosinolates, the QTL on LG 6 was detected only in the study done by Gül (2002) and in that case the allele from Samourai, the recurrent parent, was responsible for increasing the GSL content. This QTL shows in Gül's work one small additive effect. In this work, Samourai was taken as a standard, since it has the lowest glucosinolate content (Tab. 5). In that way, all three experiments in Tab. 5 and Tab. 6 show results around 16.5 µmol GSL/g seed for Samourai. Taken Samourai as reference, a phenotypic effect (QTL effect) of -0.18 shown for QTL on LG 6 could be estimated (Tab. 5), but in other hand the phenotypic value of the 7 genotypes carrying Mansholt's allele at the QTL on LG 6 are not statistically different from Samourai (t-test). That means that at the present study the estimated effect of QTL on LG 6 can be considered as not significant.

From Gül's data can be postulated that Mansholt's allele at QTL on LG 6 increases the glucosinolate content. That can be seen in Tab. 9, where two genotypes containing Mansholt's allele for QTL 6 show a higher (18.32 µmol GSL/g) GSL content than Samourai (16.51 µmol GSL/g). Nevertheless, probably also the 15.73 µmol GSL/g and the 18.32 µmol GSL/g are part of the variance promoted by the environmental effects, and therefore do not mean a large difference. Such questions would only be clarified when the intervarietal substitution lines are available and a wide range of field experiments can be done. The trustful conclusion in the present study is that QTL 6 has only a small effect or none (Tab. 5).

The effect from Mansholt's allele at the QTL on LG 2 is negative (-2.30 µmol GSL/g seed), which is contradictory to the results of Uzunova *et al.* (1995) where a positive effect could be found. The negative effect found in the present BC_3 data can be verified also through two other genotypes, which also have negative effects compared to Samourai. The problem is that in those genotypes (Tab. 5), respectively DH 1097 (12.55 µmol GSL/g) and

BC_1 1097 (13.73 µmol GSL/g), one can not be sure about the presence of these Mansholt's alleles, because of doubtful marker information at LG 2. In that way, a doubt is still present but since the effect is also negative, the probability that this allele is present is high, since the results of such DH and BC_1 are coherent to the results of the BC_3.

On the other hand, the effect from Mansholt's allele at QTL on LG 3 is positive, but small (2.88 µmol GSL/g seed), and agrees very well with the effects estimated in previous studies (Tab. 7). As the difference between the phenotypic value of Samourai and the phenotypic value of the genotypes heterozygous for the QTL on LG 3 is significant (t-test at Tab. 5), one can conclude that the respective effect estimated for this QTL is also significant. Almost the same small effect can be observed in two other genotypes (Tab. 5), DH 2078 and its BC_1, which show similar results (19.10 and 19.42 µmol GSL/g seed respectively) and confirm so the effect estimated through the BC_3 families. Dominance (complete dominance) is probably the best explanation in this case observing the results of the later two mentioned genotypes (Tab. 5), but even here one can not be totally sure since the standard deviation between families means in these genotypes are 2.7 µmol (variance = 7.51, Tab. 5).

Indeed, the real effect of such minor QTL is difficult to determine, even with BC_3 genotypes, because they still contain considerable amount of donor genome besides the normal segments which contain the allele at the QTL of interest. In this sense, another interesting comparison is useful. One set of 43 genotypes (25 in Reinshof and 18 in Elliehausen) in the present study show a phenotypic value of 15.73 µmol GSL/g (Tab. 5), which is statistically different from Samourai (t-test). The lower glucosinolate content of such BC_3 families, which have the QTL on LG 6 in homozygote state (two Samourai's alleles) can be a weak indication that there could be other minor genes in Mansholt that, if separated from the major genes for high glucosinolate content, can promote a reduction of glucosinolate content. However, if the variance of plants within the plots of Samourai is compared to the variance within plots in those 43 BC_3 families, a low variance can be found in both, indeed lower in those BC_3 than in Samourai (Tab. 5). If hypothetically other genes that affect the GSL content are really still present in Mansholt, the variance of such BC_3 families should be higher than Samourai. This, in his turn, suggests that the difference between 16.81 and 15.73 µmol GSL/g is maybe due to environmental effects or by chance and possible no more major or minor genes are segregating in those BC_3 families, which obviously still contain different parts of the donor genome. This is confirmed also by the HPLC data shown in Tab. 11 where the total glucosinolate content from a subset of 41 Samourai plots (out of 50 plots in Tab. 5, i.e., 27 + 23) in Reinshof and Elliehausen was found to be not different (t-test) from a subset of 40 (out of 43) plots of BC_3 families which are also homozygote for the recurrent allele at QTL on LG 6. Obviously the 43 BC_3 families do not represent the entire possibilities of different genetic composition of the lines with

reference to donor segments, and therefore can not exclude the possibility that in some regions of Mansholt's genome some minor genes can be found that reduce or increase glucosinolate content.

Summarizing, if the effect of the three major QTL were taken as additive ones (like it could be calculated in Reinshof+Elliehausen) the expected model following the effects estimated in the BC_3 families should be good correlated (Tab. 9) with the observed phenotypic values got by Gül (2002). Nevertheless, this is not always true since many genotypes classified as "E" show possible epistatic effects (Tab. 9), and therefore are deviations of such "*additive model*" (see chapter 3.3.3). Some epistasis were also supposed to exists, according to BC_3 families (Tab. 9), between Mansholt's alleles 9 and 16 (negative epistatic effect), between 9 and 18 (small positive epistatic effect), 3 and 18 (negative epistatic effect), and 2 and 16 (strong positive epistatic effect).

Tab. 8 shows that a clear negative epistasis is observed between the two Mansholt's alleles from QTL on LG 16 and 9, like it was described by Weißleder (1996) and confirmed by the results of QTL-mapping by Gül's data (Tab. 10). The other combinations of Mansholt's alleles for the major QTL are giving almost no epistatic effects or small ones (Tab. 8). Nevertheless, strong epistatic effects exist between Mansholt's alleles from the QTL on LG 2 and 16 and between 3 and 18. So, sometimes interaction effects reduce (or increase) phenotypic effects of Mansholt's alleles when these alleles are present at more than one of the QTL. Such interaction effects may also have led to an under or overerestimation of additive effects by interval mapping since many of the doubled haploid lines of the mapping population have Mansholt's alleles at more than one of the QTL and interaction effects were not included in the model used for QTL mapping in previous studies.

One must be aware that such epistatic effects should be confirmed in further studies using DH lines or intervarietal substitution lines which carry the different alleles one by one and in pairs. But in any case, these epistatic effects explain why genotypes containing the QTL on LG 9 and 16 contain less glucosinolate than genotypes carrying Mansholt's alleles at LG 16 and 18, since if no epistasis were present, the contrary should be expected according to the individual effect of such QTL.

Still concerning epistasis, the original data from Gül (2002) was reviewed and new estimations were done (Tab. 10). Three significant interactions (epistasis) could be detected, and one of those was between the QTL on LG 9 and one region of LG 16, although with respect to LG 16 it must be pointed out that it is not the region containing the QTL mapped for glucosinolate content on this linkage group. Another interaction occurs between two linkage groups (7 and 14) in which no QTL could be located before. One more interaction also could be verified between a locus on LG 15 and the QTL on LG 18.

Another last interesting thing to point out from the results of Reinshof and Elliehausen is that usually only two major alleles from Mansholt, respectively at two out of the three major QTL, are sufficient for ensure a glucosinolate content precisely around 51 µmol/g seed. This can be seen if the genotypes 9x16, 16x18 and 9x18 are compared (Tab. 5).

The present results confirm also the position of the three major QTL, which on his turn also confirm the results obtained in 5 different crosses (7 studies) where 9 different QTL (three major and six minor QTL) influencing this trait could be detected (Uzunova *et al.*, 1995; Toroser *et al.*, 1995; Weißleder, K., 1996; Campos de Quiroz and Mithen, 1996; Fischer and Ecke, 1997; Gül, M.K., 2002; Howell *et al.*, 2003). Through map alignment six out of seven studies (three different crosses) could identify the same three major QTL, while Toroser *et al.* (1995) and Campos de Quiroz and Mithen (1996) have identified two out of these three major QTL, i.e., those corresponding to the LG 16 and LG 18 presented in this study. Finally, since only two locations without replications were used as the source of the BC_3 families data, the results can sometimes only show some tendencies but much better and secure answers can arise in the future, with the use of the desired intervarietal substitution lines.

Therefore, as a final conclusion, for the total glucosinolate content also in the present study three major QTL are involved. At least two of these QTL (LG 16 and LG 18) seem to appear even across different crosses mentioned in the literature. In this study even the position of such three major QTL could be confirmed, since at the expected position mapped in previous studies (through composite interval mapping) the undoubted effects could be found at the BC_3 families selected with molecular markers. With respect to the three minor QTL, some effect could be found for each of them, although that only the QTL on LG 3 was significant. Dominance is the more probably intra allelic interaction present concerning the glucosinolate content but the results show also that some other factors are guiding this trait, like epistasis, so that the complete elucidation of the genetic control of glucosinolates in rapeseed still needs more time and investigation.

5. SUMMARY

Rapeseed (*Brassica napus* L.) is one of the most important oil crops of the world and in 2002 represents 10.1% (26 million ha) of the harvested area from oilseed crops, producing 31,604,000 t seeds. A better understanding of the genetics and inheritance of characters in rapeseed is therefore an important objective of plant breeding research. After extraction of the oil, which is the most valuable seed component, the resulting meal is an important protein source for animal feed. Since some components present in the meal are detrimental to animal nutrition, like glucosinolates, it still remains a breeding aim to have varieties with lower glucosinolate content.

Six QTL for glucosinolate content have been identified until now in rapeseed in the cross Mansholt x Samourai. Three of them have been reported to have an important effect (major QTL) while the other three showed only smaller effects (minor QTL). Despite the use of low glucosinolate genotypes in modern breeding programmes it seems to be common that genotypes appear, through segregation, which carry higher levels of glucosinolates than each of the parents.

Trying to improve information which respects to the glucosinolate content in rapeseed the specific aims of the present work consists of two points. The first objective was the production of basic material for the development of intervarietal substitution lines. The second objective was to investigate the inheritance of glucosinolate content, i.e., the estimation of the effect of the alleles of six QTL based on field trials in which BC_3 families segregating for one or two of the QTL were tested.

With reference to the first aim mentioned, using marker assisted backcrossing a set of intervarietal substitution lines should be developed. Sets of such intervarietal substitution lines represent a genetically well characterised material that is ideally suited for a genome wide analysis of the effects of allelic variation on the phenotypic and molecular levels. Within the common genetic background of the recurrent parent, each line will contain a small, well-defined segment of the genome of the donor parent. Across all lines of a series, these segments will, partially overlapping, represent the entire donor genome. The work began with a double haploid (DH) population developed from a cross between doubled haploid lines of the two winter rapeseed varieties ‘Mansholt’s Hamburger Raps’ and ‘Samourai’. In the development of the substitution lines ‘Mansholt’, an old Dutch variety, was used as donor parent and ‘Samourai’, a current French variety, as recurrent parent. Ten DH lines carrying donor segments covering the whole donor genome were selected through marker assisted selection. These ten DH lines were crossed with Samourai resulting in ten BC_1 genotypes which were again backcrossed with Samourai producing a BC_2 population with 300 plants.

These plants were analysed with 140 AFLP markers (114 loci) and 20 out of 300 could be selected which donor segments cover the entire donor genome. Because a better coverage of the genome was desired the 20 selected plants were analysed with more AFLP markers as well as with two SSR and one RAPD markers, giving a total of 164 markers (130 loci) which cover 1,325 cM of the rapeseed genome.

The 20 selected BC_2 genotypes were also backcrossed to Samourai and a population of 300 BC_3 plants was obtained. These plants were analysed by 132 markers (120 loci) and 52 plants were selected which donor segments cover the mapped Mansholt's genome. However among these 300 plants none could be found having only one donor segment. The number of donor segments in the BC_3 plants ranged from 2 to 13. This was, according to literature information, not expected. Previous studies could find already in BC_3 populations genotypes carrying only one donor segment. But looking into these studies more carefully it became clear that this was only possible through the fact that not the whole mapped donor genome was covered. In the present study, the whole mapped genome was covered.

Concerning to the second objective of this study, the backcross populations used for the development of the substitution lines were also used for the estimation of QTL effects of the six QTL for glucosinolate content. The offspring of the BC_2 population analysed by markers were evaluated in two field trials in Göttingen. 279 BC_3 families, the 10 DH lines initially selected for the development of the intervarietal substitution lines and their derived BC_1 genotypes as well as the parental lines (Samourai and Mansholt), were grown in 2001/2002 in field trials at two locations in plots with 50 plants. Seed glucosinolate content was determined in intact-seed samples by near infrared reflectance spectroscopy (NIRS). Based on the marker information on BC_2, BC_3 families segregating for only one or two of the six QTL were selected and seed glucosinolate content was determined from individual plants. From such families the plants were individually harvested and analysed by NIRS, to look after the segregation within each family. QTL effects in BC_3 families segregating for only one of the QTL were calculated as the difference between phenotypic classes.

For the two major QTL on linkage groups 16 and 18 clear 1 : 1 segregations into two phenotypic classes were observed, which was not the case for the major QTL on linkage group 9 and for the three minor QTL, and therefore another strategy was necessary to estimate the QTL effects in these cases. In families were phenotypic classes could not be clearly distinguished the QTL effect was calculated by comparing the phenotypic mean (z) of the BC_3 family with the phenotypic mean (y) of the recurrent parent Samourai under the assumption that "z" can be calculated as "z =(x + y)/2" with "x" being the phenotypic mean of the heterozygous class and "y" of the homozygous class, which corresponds to Samourai. In

this way, in BC_3 families segregating for only one QTL the phenotypic mean of the heterozygous class can be calculated as “x = (2z) – y”, and the QTL effect as “x-y”.

In this study the positions of the 3 major QTL could be confirmed, since in the expected regions in which they had been mapped in previous studies (by composite interval mapping), clear effects on glucosinolate content in the BC_3 families were observed. In the BC_3 families the QTL effect estimated is the sum of additive effects and dominance effects, and even epistatic effects are not excluded. A comparison with the additive effects estimated by composite interval mapping indicates the presence of strong dominance effects. For the QTL on linkage group 18 additive and dominance effects are nearly equal. For the other two major QTL the dominance effects are larger than the additive effects. With respect to the 3 minor QTL, some effects could be found for each of them, although only the effect of the QTL on LG 3 was statistically significant. The results show also that some other factors are guiding this trait, like epistasis, so that the complete elucidation of the genetic control of glucosinolates in rapeseed still needs more time and investigation. Epistatic effects could also be found in BC_3 families segregating for 2 QTL.

6. ZUSAMMENFASSUNG

Markergestützte Selektion für die Entwicklung von intervarietalen Substitutionslinien bei Raps (*Brassica napus* L.) und die Schätzung von QTL-Effekten für Glucosinolatgehalt.

Raps (*Brassica napus* L.) ist eine der wichtigsten Ölpflanzen der Welt. In 2002 lag der Anteil von Raps an der weltweiten Anbaufläche für Ölpflanzen bei 10,13%, und die Produktion betrug 36.371.000 t. Ein besseres Wissen über die Genetik und Vererbung von Merkmalen bei Raps ist deshalb von großer Bedeutung. Nachdem das Öl extrahiert ist, dient das verbleibende Schrot als hochwertige Proteinquelle für die Tierernährung. Da aber einige Komponenten im Mehl schädliche Wirkung haben, wie die Glucosinolate, bleibt es immer noch ein Ziel der Züchtung Sorten zu entwickeln die einen geringeren Anteil an Glucosinolaten haben.

Sechs QTL für Glucosinolatgehalt wurden bisher identifiziert in einer Kreuzung zwischen ‚Mansholt' und ‚Samourai'. Drei davon zeigen größere Effekte (major QTL) und drei nur kleine Effekte (minor QTL). In modernen Züchtungsprogrammen kommt es immer wieder vor, dass aus Kreuzungen zwischen Eltern die beide einen niedrigen Glucosinolatgehalt haben, Nachkommen entstehen, die einen höheren Gehalt als die Eltern haben. Deshalb ist ein besseres Verstehen der Vererbung dieses Merkmals erwünscht.

In diesem Sinne wurde im Rahmen dieser Studie versucht, mehr über dieses Merkmal zu erfahren, in dem man zwei Ziele verfolgte. Das erste Ziel war die Erstellung eines Basismaterials für die Herstellung von Substitutionslinien („Invervarietal Substitution Lines"). Das zweite Ziel war die Schätzung der Effekte der sechs genannten QTL durch Feldversuche mit spaltenden BC_3 Familien.

Für das erste Ziel sollten über markergestützte Rückkreuzungen mit Hilfe von AFLP-Markern eine Serie sog. „intervarietal substitution lines" hergestellt werden. Diese Substitutionslinien enthalten vor dem gemeinsamen Hintergrund des rekurrenten Elters jeweils ein einzelnes genau definiertes Segment des Donorgenoms. In einer Serie decken diese Segmente, z. T. überlappend, das gesamte Genom des Donorelters ab. Serien solcher Substitutionslinien repräsentieren daher ein genetisch sehr gut charakterisiertes Material, das ideal geeignet ist für eine genomweite Charakterisierung alleler Variation sowohl auf phänotypischer als auch auf molekularer Ebene. Die Arbeit begann mit einer Population von doppelt haploid Linien (DH Linien) hergestellt aus einer Kreuzung zwischen zwei Winterrapssorten, ´Mansholt's Hamburger Raps´ und ´Samourai´. Zehn dieser DH Linien mit Donorsegmente die das gesamte Genom abdecken, wurden durch markergestützte Selektion ausgewählt. Diese zehn Linien wurden mit Samourai rückgekreuzt um zehn BC_1 Genotypen zu erzeugen, die

dann ein zweites Mal mit Samourai rückgekreuzt wurden, so dass eine BC_2 Population mit 300 Pflanzen entstand. Diese wurde dann mit 140 AFLP Markern (114 loci) untersucht und 20 Pflanzen konnten selektiert werden deren Donorsegmente wiederum das gesamte Donorgenom abdeckten. Da eine bessere Abdeckung des Genoms erwünscht war, wurden in den 20 selektierten Pflanzen noch einige AFLP Primerkombinationen eingesetzt, sowie zwei Mikrosatelliten-Markern und ein RAPD-Marker, wodurch sich eine Gesamtzahl von 164 Markers (130 loci) ergab, die etwa 1325 cM des Rapsgenoms abdecken.

Die 20 selektierten BC_2 Genotypen wurden abermals mit Samourai rückgekreuzt um eine Population von 300 BC_3 Pflanzen zu erzeugen. Diese wurde dann mit 132 Markern (120 loci) analysiert. Insgesamt 52 BC_3 Pflanzen wurden selektiert deren Donorsegmente des Mansholt Genom wiederum abdecken. Allerdings wurde unter den 300 BC_3 Genotypen keiner gefunden der nur noch ein Donorsegment enthält, sondern die verschiedene BC_3 Pflanzen enthielten zwischen 2 und 13 Donorsegmente. Das Ergebnis war nach vorliegenden Studien unerwartet. In diesen Studien waren in der BC_3 schon Genotypen aufgetreten, die nur ein Donorsegment enthielten. Aber eine genauere Betrachtung dieser Untersuchungen zeigte dass bei der Selektion nicht das gesamte kartierte Genom berücksichtig worden war, also mit den hergestellten Substitution Linien wurde nicht das gesamte kartierte Donorgenom abgedeckt. In der vorliegenden Studie wurde dagegen immer das ganze kartierte Genom betrachtet sofern die Markerabdeckung es ermöglichte.

Was es dem zweiten Ziel angeht, die Rückkreuzungspopulationen die herstellt wurden für die Entwicklung des Substitutionslinien, wurden auch für die Schätzung der Effekte der sechs QTL für Glucosinolatgehalt benutzt. Die Nachkommenschaften der mit Markern analysierten BC_2 Pflanzen wurden in zwei Feldversuchen in Göttingen getestet. 279 BC_3 Familien, die 10 selektierten DH Linien und die entsprechenden BC_1 Genotypen sowie die Eltern der Kreuzung wurden im Jahr 2001/2002 in Parzellen mit je 50 Pflanzen angebaut. Der Glucosinolatgehalt in den Körnern wurde in ganzen Korn durch NIRS (near infrared reflectance spectroscopy) gemessen. Von den angebauten Familien wurden, nach Abschluss der Markerselektion, nur die Familien ausgewählt, die für einen oder zwei QTL segregierten. Von diesen Familien wurden dann einzeln Pflanzen geerntet und mit NIRS gemessen um innerhalb der Familien die Spaltung erkennen zu können. Der QTL Effekt bei den BC_3 Familien, die für nur einen QTL segregierten, wurde aus der Differenz zwischen den phänotypischen Klassen, die sich durch die Segregation ergaben, bestimmt.

Für die zwei major QTL auf den Kopplungsgruppen 16 und 18 wurde eine klare 1:1 Spaltung erkennbar, was allerdings nicht bei der Kopplungsgruppe 9, oder bei den minor QTL der Fall war. In Familien wo die phänotypischen Klassen nicht klar zu erkennen waren wurde der QTL Effekt geschätzt mittels eines Vergleichs zwischen dem phänotypischen Mittelwert (z) der BC_3 Familie und dem phänotypischen Mittelwert (y) des rekurrenter Elters (Samourai)

unter der Annahme das “z” durch folgende Formel errechnet werden kann: “z = (x + y)/2”. Hier steht “x” für den der phänotypischen Wert der heterozygoten Klasse und “y” für den homozygoten Klasse, die Samourai entspricht. Der phänotypische Wert der heterozygoten Klasse der spaltenden BC_3 Familien (die für nur einen QTL spalten) kann dann nach Umformung als “x = (2z) – y” errechnet werden, und der QTL Effekt als die Differenz zu Samourai, also als “x-y”.

Die Effekte der 2 major QTL (auf Kopplungsgruppen 16 und 18), aber auch der dritte QTL (Kopplungsgruppen 9), sind so deutlich, dass dessen genaue Lokalisierung (aus früheren Studien durch ‚composite interval mapping’) bestätigt werden konnte. In den BC_3 Familien ist der QTL Effekt die Summe von Additiv- und Dominanzeffekt, wobei epistatische Effekte nicht ausgeschlossen sind. Wenn man diese QTL Effekte mit den durch Intervall Kartierung geschätzten Werten vergleicht, erkennt man dass deutliche Dominanzeffekte vorhanden sind. Für den QTL auf Kopplungsgruppe 18 sind Additiv- und Dominanzeffekte etwa gleich groß. Bei den anderen zwei QTL sind die Dominanzeffekte größer als die Additiveffekte. Für die 3 minor QTL konnten Effekte gefunden werden aber nur der Effekt auf Kopplungsgruppe 3 war statistisch signifikant. Epistatische Effekte konnten ebenfalls nachgewiesen werden, und zwar in den BC_3 Familien die für 2 QTL spalten.

7. RESUMO

Seleção assistida por marcadores para o desenvolvimento de linhas de substituição invervarietais em colza (*Brassica napus* L.) e estimativa do efeito dos QTL para teor de glucosinolatos.

A colza (*Brassica napus* L.) economicamente, é uma das plantas oleaginosas mais importantes do mundo, onde, em 2002 ela representou 10,1% da área colhida de oleaginosas, com uma produção de 36.371.000 toneladas de grãos. Considerando isto, um conhecimento mais detalhado da genética e hereditariedade dos caracteres nesta espécie é portanto um importante objetivo de pesquisa no melhoramento vegetal. Após a extração do óleo, que é o componente mais importante nos grãos, a farelo resultante é uma boa fonte de proteína para ração animal. O farelo, no entanto, ainda contém alguns componentes nocivos à nutrição animal, como os glucosinolatos, tornando a geração de variedades com um menor teor de glucosinolatos um importante e constante alvo no melhoramento desta espécie.

Até o momento foram identificados seis QTL (Quantitative Trait Loci) para teor de glucosinolatos no cruzamento 'Mansholt' x 'Samourai'. Destes, três apresentam grande efeito (QTL maiores) e três apresentam um efeito menor (QTL menores). Em modernos programas de melhoramento é comum que, do cruzamento de parentais com baixo teor de glucosinolatos surjam, devido à segregação, descendentes com teor superior aos dos parentais. Assim, um melhor conhecimento a respeito da herança deste caracter é desejável.

Neste sentido, tentou-se através deste estudo obter maiores informações sobre a herança dos glucosinolatos, buscando especificamente atingir dois objetivos. O primeiro objetivo visou obter material básico para o desenvolvimento de "linhas de substituição intervarietais". Estas linhas seriam basicamente constituídas pelo genoma do parental recorrente e segmentos pequenos e bem definidos do parental doador. O segundo objetivo visou estimar o efeito dos seis QTL mencionados acima, através de experimentos de campo com famílias BC_3, que segregariam para um ou dois dos QTL citados.

Considerando o primeiro objetivo, visava-se obter, utilizando retrocruzamentos assistidos por marcadores AFLP, um conjunto de plantas denominado "série de linhas de substituição". Numa série (biblioteca) os segmentos do parental doador estariam parcialmente sobrepostos, de modo que cobririam todo o genoma deste. Estas séries representam um conjunto de "linhas de substituição" que por sua vez constituem-se em um material genético bem definido, ideal para a caracterização de alelos presentes no genoma em nível fenotípico bem como molecular. Iniciou-se com uma população duplo haplóide

(linhas DH) derivada de um cruzamento entre duas variedades de colza de inverno, 'Mansholt's Hamburger Raps' e 'Samourai'. A primeira, usada como parental doador, é uma antiga cultivar holandesa, e 'Samourai', usada como parental recorrente, é uma moderna cultivar francesa. Dez linhas DH, cobrindo todo o genoma do parental doador, foram escolhidas através da seleção assistida por marcadores (SAM). Estas dez linhas foram retrocruzadas com Samourai originando plantas BC_1, que por sua vez foram novamente retrocruzadas com Samourai constituindo assim uma população de 300 plantas BC_2. Estas foram então analisadas com 140 marcadores AFLP (114 loci), sendo que 20 plantas BC_2 foram então selecionadas abrangendo a totalidade do genoma do parental doador. Considerando-se que era desejável um detalhamento da informação de marcadores, decidiu-se aplicar nas 20 plantas BC_2 selecionadas mais algumas combinações de primers AFLP, bem como dois microsatélites e um RAPD. Como decorrência, obteve-se 164 marcadores (130 loci) que permitiram a cobertura de 1.325 cM do genoma da colza.

As 20 plantas BC_2 selecionadas foram também retrocruzadas com Samourai originando uma população de 300 plantas BC_3. Estas plantas foram analisadas com 132 marcadores moleculares (120 loci) e destas, 52 foram selecionadas, de modo que os segmentos do parental doador cobrissem a totalidade do genoma (Mansholt) mapeado. No entanto, entre as 52 plantas não foi possível encontrar alguma que possuísse somente um segmento do parental doador. O conjunto das 300 plantas apresentou de 2 a 13 segmentos do parental doador por planta. De acordo com a literatura, já na geração BC_3 poderiam ser obtidos genótipos que portassem somente um segmento do parental doador. Uma análise pormenorizada revelou, no entanto, que isso somente seria possível se se ignorasse parte do genoma. Isto significa que com as linhas de substituição desenvolvidas nesses estudos, não foi possível abranger a totalidade do genoma do parental doador. Entrentanto, no presente estudo a totalidade do genoma mapeado foi considerada.

No que se refere ao segundo objetivo, isto é, estimar o efeito dos seis QTL para teor de glucosinolatos, o material produzido através dos retrocruzamentos visando a obtenção de linhas de substituição foi concomitantemente utilizado para estimar o efeito dos seis QTL. Assim, os descendentes das plantas BC_2, analisadas com marcadores, foram avaliados em dois experimentos de campo em Göttingen. As 279 famílias BC_3, as 10 linhas DH parentais selecionadas inicialmente para o desenvolvimento de linhas de substituição, e seus correspondentes genótipos BC_1, bem como os parentais Mansholt e Samourai, foram avaliados no ano 2001/2002 em parcelas de 50 plantas. O teor de glucosinolatos foi avaliado nos grãos intactos, utilizando-se NIRS (técnicas estereoscópicas de infravermelho próximo). Do total de famílias testadas no campo, após a seleção assistida por marcadores, somente foram selecionadas aquelas que segregam para um ou dois QTL. Destas famílias fez-se então a colheita de plantas individuais, as quais foram avaliadas por NIRS para

verificar a existência ou não de segregação nas diversas famílias. O efeito do QTL nas famílias BC_3 que só segregavam para um dos seis QTL, foi calculado a partir da diferença entre as classes fenotípicas que decorreram da segregação.

Para os dois QTL maiores, nos grupos de ligação 16 e 18, uma clara segregação 1 : 1 em duas classes fenotípicas foi observada, o que no entanto não ocorreu com o outro QTL maior localizado no grupo nove e tampouco nos três QTL menores. O cálculo do efeito destes QTL, por sua vez, necessitou de um procedimento especial. Nas famílias onde uma clara segregação não pode ser observada, o efeito do QTL foi estimado através da comparação do valor fenotípico médio da família BC_3 em questão (z), e o valor fenotípico médio do parental recorrente (y) (Samourai), considerando que "z" poderia ser calculado pela seguinte equação: "z=(x + y)/2". Nesta equação, o valor de "x" é o valor fenotípico médio da classe heterozigota para o QTL em questão, enquanto "y" representa o valor fenotípico médio da classe homozigota, que corresponde ao parental Samourai. Assim, estimou-se o fenótipo (valor fenotípico médio) da classe heterozigota através da igualdade "x = (2z) – y", sendo finalmente o efeito do QTL a diferença desta classe para Samourai, ou seja, "x-y".

No presente estudo, a localização (posição) dos três QTL maiores pode ser confirmada, uma vez que nas regiões em que foram mapeados (através de "composite interval mapping"), claros efeitos, quanto ao teor de glucosinolatos, puderam ser encontrados nas famílias BC_3. Nestas, o efeito do QTL é a soma dos componentes aditivos e de dominância, sendo que efeitos epistáticos não estão excluídos. Quando comparam-se os efeitos dos QTL estimados pelo "interval mapping" com os obtidos neste estudo através das famílias BC_3, conclui-se que efeitos de dominância estão significativamente presentes. Para o QTL do grupo 18 os efeitos aditivos e de dominância praticamente se igualam. Nos outros dois QTL maiores, os efeitos de dominância são maiores do que os aditivos. No que se refere aos QTL menores, efeitos também puderam ser estimados, embora apenas o QTL no grupo de ligação três tenha apresentado um efeito estatísticamente significativo. Efeitos epistáticos foram verificados nas famílias BC_3 que apresentavam segregação em dois QTL.

8. LITERATURE

Allard, R. W. (1999). Principles of Plant Breeding 2nd ed. John Wiley & Sons Inc., New York.

Amersham International plc. (1997). Nucleon Extraction & Purification Protocols. 20p.

Beavis, W.D., Smith, O.S., Grant, D., Fincher, R. (1994). Identification of quantitative trait loci using a small sample of topcrossed and F4 progeny from maize. Crop. Sci. 34:882-896

Becker, H.C.; Löptien, H.; Röbbelen, G. (1999). Breeding: an overview. In: Gómez-Campo, C. Biology of Brassica Coenospecies. Elsevier Science B.V., Amsterdam. p. 413-460.

Bennett, M.D.; Leitch, I.J. (1995). Nuclear DNA amounts in angiosperms. Annals of Botany 76: 113-176

Bennett, M.D.; Leitch, I.J. (2001). Plant DNA C-values Database (homepage). Royal Botanic Gardens, Kew, release 1.0, Sep. 2001. http://www.rbgkew.org.uk/cval/homepage.html

Berres, M.E. (2002). http://ravel.zoology.wisc.edu/sgaap/AFLP_html/fAFLP_Introduction.htm.

Bones, A.M.; Rossiter, J.T. (1996). The myrosinase-glucosinolate system, its organization and biochemistry. Physiologia Plantarum 97:194-208.

Bräuer, G. (2003). Der Ursprung lag in Afrika. Spektrum der Wissenschaft, 03/03.

Bridges, M.; Jones, A.M.; Bones, A.M.; Hodgson, C.; Cole, R.; Bartlet, E.; Wallsgrove, R.; Karapapa, V.K.; Watts, N.; Rossiter, J.T. (2002). Spatial organization of the glucosinolate-myrosinase system in Brassica specialist aphids is similar to that of the host plant. Proc. R. Soc. Lond. B., 269:187-191.

Buchner, R. (1988). Analyse und Biologie der Glucosinolate in Raps (*Brassica napus* L.). Dissertation, Georg August Universität, Göttingen.

Burns, M.J.; Barnes, S.R.; Bowman, J.G.; Clarke, M.H.E.; Werner, C.P.; Kearsey, M.J. (2003). QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) Seed oil content and fatty acid composition. Heredity, 90:39-48.

Campos de Quiroz, H.; Mithen, R. (1996). Molecular markers for low-glucosinolate alleles in oilseed rape (*Brassica napus* L.). Molecular Breeding, 2:277-281.

Castiglioni, P.; Ajmone-Marsan, P.; van Wijk, R. (1999). AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor. Appl. Genet., 99:425-431.

Cermakova, L.; Sharpe, A.; Trick, M.; Bechyne, M.; Lydiate, D. (1999). Genetic analysis of quantitative traits in *Brassica napus* using substitution lines. 10th International Rapeseed Congress, Camberra, Australia, Article at: 1999. http://www.regional.org.au/au/gcirc/4/542.htm

Clossais-Besnard, N. and Larher, F. (1991). Physiological Role of Glucosinolates in *Brassica napus*. Concentration and Distribution Pattern of Glucosinolates among Plant Organs during a Complete Life Cycle. Journal of the science and food and agric. 56:25-38

Clua, A.A.; Castro, A.M.; Giménez, D.O.; Tacaliti, M.S.; Worland, A.J. (2002). Chromosomal effects in the endogenous contents of non-structural carbohydrates and proteins measured in wheat substitution lines. Plant Breeding 121: 141-145

Darvasi, A.; Weinreb, A.; Minke, V.; Weller, J.I.; Soller, M. (1993). Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134:943-951

De March, G.; McGregor, D.I. (1989). Glucosinolate content of maturing pods and seeds of high and low glucosinolate summer rape. Can. J. Plant Sci. 69: 929-932

Derycke, G.; Mabon, N.; Mandiki, S.N.M.; Bister, J.L.; Wathelet, J.-P.; Marlier, M.; Paquay, R. (1999). Chemical changes and influences of rapeseed antinutritional factors on lamb physiology and performance. Animal Feed Science and Technology, 81:81-91.

Dodoens, R. (1578). A nieuve Herbal... Antverpiae, transl. H. lyte. London

Doi, K.; Sobrizal; Ikeda, K.; Sanches, P.L.; Kurakazu, T.; Nagai, Y.; Yoshimura, A. (2002). Development and evaluation of interspecific chromosome segment substitution lines in rice. International Rice Congress, Beijing, 16-20.09.2002. Abstract: p.73.

Ecke, W.; Uzunova, M.; Weißleder, K. (1995). Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid synthesis and seed oil content. Theor. Appl. Genet. 91:972-977

Eshed, Y., Zamir, D. (1994). A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. Euphytica 79:175-179

Eshed, Y.; Zamir, D .(1996). Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143: 1807-1817

Eshed, Y.; Zamir, D. (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141:1147-1162

Eshed, Y.; Zamir, D. (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141:1147-1162

Etienne M. and Dourand J.Y. (1994). Effects of zearalenone or Glucosinolates in the diet on reproduction in sows, a review. Livestock Production Science 40: 99-113.

Fao, (2003). FAO statistical data .http://apps.fao.org/page/collections?subset=agriculture

Fapri (2002). World Oilseeds and Products: FAPRI 2002 Agricultural Outlook. 336p PDF file at www.fapri.org

Ferreira, M. E.; Grattapaglia, D. (1998). Introdução ao uso de Marcadores Moleculares em Análise Genética. 3 ed. Brasilia: EMBRAPA-CENARGEN, 220p.

Fischer, S., Ecke, W. (1997). Personal communication on “Nutzung von Rückreuzungen einer spaltenden DH-Population zur Schätzung der Dominanzeffekte kartierter Gene und QTL bei Raps (*Brassica napus* L.). Diplomarbeit, Georg August Universität, Göttingen”

Giamoustaris, A.; Mithen, R. (1996). Genetics of aliphatic glucosinolates. IV. Side-chain modification in *Brassica oleracea*. Theor. Appl. Genet., 93:1006-1010.

Gibbons, A. (2002). First Member of Human Family Uncovered. Science, 297:171-172

Goffman, F.D. (2000). Chemical Analysis, Genetic Variability and Inheritance of Tocopherol Contents in Seeds of Oilseed Rape (*Brassica napus* L.). PhD Dissertation. Göttingen: Cuvillier, 2000. 29p.

Gómez-Campo, C. (1999). Biology of Brassica Coenospecies. Elsevier Science B.V., Amsterdam. Chap. 13, p. 413-460.

Goth, K. (2002). Neue Erkenntnisse über den Ursprung der Blütenpflanzen. Naturwissenschaftliche Rundschau, 55(10): 552-3.

Gül, M.K. (2002). QTL-Kartierung und Analyse von QTL x Stickstoff Interaktionen beim Winterraps (*Brassica napus* L.). Dissertation, Georg August Universität, Göttingen.

Hall, C.; McCallum, D.; Prescott, A.; Mithen, R. (2001). Biochemical genetics of glucosinolate modification in Arabidopsis and Brassica. Theor. Appl. Genet., 102:369-374.

Han, F.; Ullrich, S.E.; Clancy, J.A.; Romagosa, I. (1999). Inheritance and fine mapping of a major barley seed dormancy QTL. Plant Science, 143:113-118. Abstract at: http://www.elsevier.nl/locate/plantsci

Herrmann, M. (1992). Die Einlagerung von Indolglucosinolaten in den reifenden Rapssamen (*Brassica napus*). Dissertation, Georg August Universität, Göttingen.

Hill R. (1991). Rapeseed meal in the diets of ruminants. Nutrition Abstracts and Reviews (Series B) 61(3): 139-155.

Hospital, F.; Chevalet, C.; Mulsant, P. (1992). Using Markers in Gene Introgression Breeding Programs. Genetics 132: 1199-1210.

Howell, P.M.; Marshall, D.F.; Lydiate, D.J. (1996). Towards developing intervarietal substitution lines in *Brassica napus* using marker-assisted selection. Genome 39:348-358

Howell, P.M.; Sharpe, A.G.; Lydiate, D.J. (2003). Homoeologus loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). Genome, 46(3): 454-460.

Huisman, J.; Tolman G.H. (1992). Antinutritional factors in the plant proteins of diets for non-ruminants. 'Recent Advances in Animal Nutrition 1992' Chp.1: 3-31.

Hyne, V.; Kearsey, M.J.; Pike, D.J.; Snape, J.W. (1995). QTL analysis: unreliability and bias in estimation procedures. Mol. Breed. 1:273-282

Jansen, R.; Geerlings, H.; Jan Van Oeveren, A.; Van Schaik, R.C. (2001). A comment on codominant scoring of AFLP Markers. Genetics, 158(2):925

Jensen, S.K. (1999). Improvement of the nutritive value of rapeseed by selecting varieties with very low glucosinolate content. 10th International Rapeseed Congress, Camberra, Australia, Article at: 1999. http://www.regional.org.au/au/gcirc/1/240.htm#TopOfPage

Kelly, P.J.; Bones, A.; Rossiter, J.T. (1998). Sub-cellular immunolocalization of the glucosinolate sinigrin in seedlings of *Brassica juncea*. Planta, 206:370-377.

Kessler, A.; Baldwin, I.T. (2002). Plant Responses to Insect Herbivory: the Emerging Molecular Analysis. Annu. Rev. Plant Biol., 53: 299-328.

Keygene (Products B.V.) (2000). AFLP-Quantar™ *Pro* 1.0 Part I, Users guide. Keygene products B.V., Wageningen, The Netherlands.

Kliebenstein, D.; Pedersen, D.; Barker, B.; Mitchell-Olds, T. (2002). Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidosis thaliana*. (Abstract) Genetics, 161:325-332.

Kubo, T.; Nakamura, K.; Yoshimura, A. (1999). Development of a series of Indica chromosome segment substitution lines in Japonica background of rice. Rice Genetics Newsletter, 16: 104. http://www.gramene.org/newsletters/rice_genetics/rgn16/v16p104.html

Lambrix, V.; Reichelt, M.; Mitchell-Olds, T.; Kliebenstein, J.; Gershenzon, J. (2001). The Arabidopsis Epithiospecifier Protein Promotes the Hydrolysis of Glucosinolates to Nitriles and Influences *Thichoplusia in* Herbivory. The Plant Cell, 13: 2793-2807.

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

Law, C.N.; Snape, J.W.; Worland, A.J. (1987). Aneuploidy in wheat and its use in genetic analysis. *In* Lupton, F.G.H. Wheat breeding and its scientific basis. Chapman and Hall, London.

Li-Cor (1999a). Remote connection guide: Global Edition IR^2 System. Li-Cor Publication n. 9910-157. 31p.

Li-Cor (1999b). DNA Sequencing Manual – Global Edition IR^2 System. LI-COR, Inc. Remote connection guide: Global Edition IR^2 System. (1999). Li-Cor Publication n. 9910-150.

Lüthy, B.; Matile, P.H. (1984). The mustard oil bomb: rectified analysis of the subcellular organization of the myrosinase system. Biochem. Physiol. Pflanz. 179: 5-12.

Magrath, R.; Mithen, R. (1993). Maternal effects on the expression of individual aliphatic glucosinolates in seeds and seedlings of *Brassica napus*. Plant Breeding, 111:249-252

Magrath, R.; Herron, C.; Giamoustaris, A.; Mithen, R. (1993). The inheritance of Aliphatic Glucosinolates in *Brassica napus*. Plant Breeding, 111:55-72

Magrath, R.; Bano, F.; Morgner, M.; Parkin, I.; Scarpe, A.; Lister, C.; Dean, C.; Turner, J.; Lodiate, D.; Mithen, R. (1994). Genetics of aliphatic glucosinolates. I. Side chain elongation in Brassica napus and Arabidopsis thaliana. Heredity, 72:290-299

Mather, K.; Jinks, J.L. (1982). Biometrical genetics. Chapman and Hall, London.

Matus, I.; Castro, A.; Corey, A.; Flichkina, T.; Hayes, P.; Marquez-Cedillo, L.; Riera-Lizarazu, O.; Rossi, C.; Sato, K.; Vales, I.; VonZitzevitz, J. (1998). Exploiting *Hordeum vulgare spp spontaneum* Genetic Resources: Diversity Analysis, Recombinant Chromosome Substitution Lines and Linkage Maps. Barley Newsletter (16th American Barley Researchers Workshop) Abstract on http://grain.jouy.inra.fr/ggpages/BarleyNewsletter/42/post18.html

Matus, I.; Corey, A.; Filichkin, T.; Hayes, P. M.; Kling, J.; Powell, W.; Riera-Lizarazu, O.; Sato, K.; Vales, M.I.;Waugh, R. (2003). Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *sulgare* background. Genome. In press

MayHospi.com-On Line Dictionary (2000). Hospipac, collaborative health integrated network. http://www.mayhospi.com/dictionary/mdresults.asp?text1=C

McGee, M. (1998). Rapeseed meal as a feed ingredient. R & H Hall Technical Bulletins http://www.rhhall.ie/technical.htm

Meave, L.; Walzer, A. (1997). Frühe Hominiden. Spektrum der Wissenschaft, 8/1997.

Miura, K.; Lin, S.Y.; Yano, M.; Nagamine, T. (2002). Mapping quantitative trait loci controlling seed longevity in rice (*Oryza sativa* L.). Theor. Appl. Genet. 104:981-986

Museum of Paleontology (2003). University of California, Berkeley. http://www.ucmp.berkeley.edu/mesozoic/cretaceous/cretaceous.html

Myburg, A.A.; Remington, D.L. (2000). Protocol for High-Throughput AFLP Analysis using LI-COR *IR*^{2} Automated Sequencers. NCState Forest Biotech AFLP Protocol-Sep2000. http://www.up.ac.za/academic/fabi/eucgenomics/euc_mapping/AFLP_protocol.pdf

Myburg, A.A.; Remington, D.L.; O´Malley, D.M.; Sederoff, R.R.; Whetten, R.W. (2001). High-Throughput AFLP Analysis Using Infrared Dye-Labeled Primers and an Automated DNA Sequencer. BioTechniques, 30:348-357.

Nadeau, J.H.; Singer, J.B.; Matin, A.; Lander, E.S. (2000). Analysing complex genetic traits with chromosome substitution strains. Nature Genetics, 24(3):221-225.

NTNU Plant Genetics Lab, (2003). http://boneslab.chembio.ntnu.no/Paal/glucosin.htm

O'Neil, D. (2003). Early Modern Homo sapiens. http://anthro.palomar.edu/homo2/modern_humans.htm

Odenbach, W. (1997). Biologische Grundlagen der Pflanzenzüchtung. Parey Buchverlag, Berlin. 381p.

Oil World (2003). ISTA Mielke GmbH – Hamburg. http://www.oilworld.de

Parkin, I.; Magrath, R.; Keith, D.; Sharpe, A.; Mithen, R.; Lydiate, D. (1994). Genetics of aliphatic glucosinolates. II. Hydroxylation of alkenyl glucosinolates in *Brassica napus*. Heredety, 72:594-598.

Parkin, I.A.P., Sharpe, A.G., Keith, D.J., Lydiate, D.J. (1995). Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome 38:1122-1131

Paterson, A.H. (1996). Genome Mapping in Plants. R.G. Landes Company and Academic Press, Inc, San Diego. 330p.

Paterson, A.H., DeVerna, J.W., Lanini, B., Tanksley, S.D. (1990). Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124:735-742

Piepho, H.; Koch, G. (2000). Codominant analysis of banding data from a dominant marker system by normal mixtures. Genetics, 155(3):1459-1468

Prakash, S.; Takahata, Y.; Kirti, P.B.; Chopra, V.L. (1999). Cytogenetics. In:Gómez-Campo, C. Biology of Brassica Coenospecies. Elsevier Science B.V., Amsterdam. Chap. 3, p. 59-106.

Rae, A.M.; Howell, E.C.; Kearsey, M.J. (1999). More QTL for flowering time revealed by substitution lines in *Brassica oleracea*. Heredity, 83: 586-596.

Ramsay, L.D., Jennings, D.E., Bohuon, E.J.R., Arthur, A.E., Lydiate, D.J., Kearsey, M.J., Marshall, D.F. (1996). The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. Genome 39:558-567

Raybould, A.F.; Moyes, C.L. (2001). The ecological genetics of aliphatic glucosinolates. Heredity 87:383-391

Ribaut, J.-M.; Jiang, C.; Hoisington, D. (2002). Simulation Experiments on Efficiencies of Gene Introgression by Backcrossing. Crop Science 42:557-565

Rousset, M.; Brabant, P.; Kota, R.S.; Dubcovsky, J.; Dvorak, J. (2001). Use of recombinant substitution lines for gene mapping and QTL analysis of bread making quality in wheat. Euphytica 119: 81-87.

Rücker, B.; Röbbelen, G. (1994). Inheritance of total and individual glucosinolate contents in seeds of winter oilseed rape (*Brassica napus* L.). Plant Breeding, 113: 206-216.

Rudolph, B. (2001). Entwicklung, Charakterisierung und genetische Kartierung von Mikrosatelliten-Markern beim Raps (*Brassica napus* L.). Dissertation, Georg August Universität, Göttingen.

Savelkoul, P.H.M.; Aarts, H.J.M.; de Haas, J.; Dijkshoorn, L.; Duim, B.; Otsen, M.; Rademaker, J.L.W.; Schouls, L.; Lenstra, J.A. (1999). Amplified-Fragment Length Polymorphism Analysis: the state of an art. Journal of Clinical Microbiology, 37(10): 3083-3091.

Schöne, F. (1999). Jodunterversorgung, Jodbedarf und Jodüberversorgung des Nutztieres – Untersuchungen mit wachsenden Schweinen. Berl. Münch. Tierärztl. Wschr. 112: 64-70 (Abstract).

Schluz, J. (2003). Personal communication on “Einfluss des Pollen-Genotyps auf die wertbestimmenden Inhaltstoffe des Rapssamens. Diplomarbeit, Georg August Universität, Göttingen”.

Seiffert, B. (2000). Molekularebiologische Untersuchungen zur Stickstoffeffizienz bei *Brassica napus* L. Dissertation, Georg August Universität, Göttingen.

Servin, B.; Hospital, F. (2002). Optimal Positioning of Markers to Control Genetic Background in Marker-Assisted Backcrossing. The Journal of Heredity, 93: 214-217

Shah, M.M.; Gill, K.S.; Baenziger, P.S.; Yen, Y.; Kaeppler, S.M.; Ariyarathne, H.M. (1999). Molecular mapping of loci for agronomic traits of Chromosome 3A of Gread Wheat. Crop Science, 39:1728-1732.

Sharpe, A.G. ; Lydiate, D.J. (2003). Mapping the mosaic of ancestral genotypes in a cultivar of oilseed rape (Brassica napus) selected via pedigree breeding. Genome, 46(3): 461-468.

Shen, Z.L.; Courtois, B.; Lafitte, R. (1999). Development of near isogenic introgression line (NIIL) sets for QTL associated with drought tolerance in rice. In: Workshop on Molecular Approaches for the genetic improvement of Cereals for Stable Production in Water-Limited Environments. Mexico: CIMMYT. http://www.cimmyt.mx/ABC/map /research_tools_results/wsmolecular/workshopmolecular/WsdroughtDevelopmentnear.htm

Sodhi, Y.S.; Mukhopadhyay, A.; Arumugam, N.; Verma, J.K.; Gupta, V.; Pental, D.; Pradhan, A.K. (2002). Genetic analysis of total glucosinolate in crosses involving a high glucosinolate Indian variety and a low glucosinolate line of *Brassica juncea*. Plant Breeding 121:508-511

Sørensen, H. (1988). Analysis of glucosinolates and acceptable concentrations of glucosinolates in oilseed rape and products thereof used as feed to different animals. GCIRC Bulletin 4: 17-19.

Stoewsand, G.S. (1995). Bioactive Organosulfur Phytochemicals in Brassica oleracea Vegetables – A Review. Food and Chemical Toxicology, 33(6):537-543, Abstract.

Tanksley, S.D.; Nelson, J.C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet., 92:191-203

Teutonico, R.A.; Osborn, T.C. (1994). Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. Theor. Appl. Genet., 89:885-894

Thangstad, O.P.; Bones, A.M.; Hotan, S.; Moen, L.; Rossiter, J.T. (2001). Microautoradiographic localisation of a glucosinolate precursor to specific cells in Brassica napus L. embryos indicates a separate transport pathway into myrosin cells. Planta, 213:207-213 http://link.springer.de/link/service/journals/00425/contents/00/00491/

Thies, W. (1988). Isolation of Sinigrin and Glucotropaeolin from cruciferous seeds. Fat Sci. Tech. 90(8):311-314 (cited by Herrmann, 1992)

Thomas, W.T.B.; Newton, A.C.; Wilson, A.; Booth, A.; Macaulay, M.; Keith, R. (2000). Development of Recombinant Chromosome Substitution Lines – a barley resource. Scottish Crop Research Institute. Annual Report 1999/2000: 99-100.

Toroser, D.; Thormann, C.E.; Osborn, T.C.; Mithen, R. (1995). RFLP mapping of quantitative trait loci controlling seed aliphatic-glucosinolate content in oilseed rape (*Brassica napus* L.). Theor. Appl. Genet. 91: 802-808.

USDA (2002). - Natural Resources Conservation Service, http://plants.usda.gov/cgi_bin/plant_profile.cgi?symbol=BRNA

Uzunova, M.; Ecke, W.; Weißleder, K.; Röbbelen, G. (1995). Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor. Appl. Genet. 90:194 – 204

Van Berloo, R. (1999). GGT: software for the display of graphical genotypes. Journal of Heredity 90: 328-329. (http://www.dpw.wau.nl/pv/pub/ggt/)

Vandemark, G.J.; Miklas, P.N. (2002). A fluorescent PCR assay for the codominant interpretation of a dominant SCAR marker linked to the virus resistance allele *bc-1*^{2} in common bean. Molecular breeding, 10:193-201

Visscher, P.M. (1996). Proportion of the variation in genetic composition in backcrossing programs explained by genetics markers. Journal of Heredity, 87: 136-138.

Visscher, P.M.; Haley, C.S.; Thompson, R. (1996). Marker-Assisted Introgression in Backcross Breeding Programs. Genetics, 144: 1923-1932.

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Vandelee, T.; Hornes, M.; Frijters, A.; Pot, J. *et al.* (1995). AFLP – a new technique for DNA-fingerprinting. Nucleic Acids Research, 23: 4407-4414.

Vos, P.; Kuiper, M. (1997) AFLP analysis. In: Caetano-Anolles, G; Gresshoff, P.M. DNA Markers: protocols, applications and overviews. Wiley-Liss, Inc., p.115-131.

Wallsgrove, R.; Bennet, R.N..(1995). The biosynthesis of glucosinolates in Brassicas. In: Wallsgrove, R.M. Amino acids and their derivates in higher plants. Cambridge: Cambridge University Press,. p. 243-259

Wang, D. L.; Zhu, J.; Li, Z.K.;Patterson, A.H. (1999a). Mapping QTL with effects and QTL x environment interactions by mixed linear model approaches. Theor. Appl. Genet. 99: 1255-1264.

Wang, D. L.; Zhu, J.; Li, Z.K.;Patterson, A.H. (1999b). User Manual for QTLMapper Version 1.01b. A Computer Software for Mapping Quantitative Trait Loci (QTL) with Main Effects, Epistatic Effects and QTL x Environment Interactions. Department of Agronomy, Zhjiang University, Hangzhou 310029, China.

Weißleder, K. (1996). Genetische Kartierung von Loci für züchterisch bedeutsame Merkmale beim Winterraps (*Brassica napus* L.). Dissertation, Georg August Universität, Göttigen.

Weißleder, K. (1993). Spaltungs- und Kopplungsanalysen an einer Winterrapskreuzung zur Entwicklung phänotypischer Marker für die Genomkartierung. Diplom-Arbeit, FB Agrar-wissenschaften, Universität Göttingen, 133 S

Weißleder, K.; Uzunova, M.; Ecke, W. (1996). Genetische Kartierung von Loci für phänotypische Merkmale bei Winterraps (*Brassica napus* L.). Vortr. Pflanzenzüchtung 32:88-90

Welter-Schultes, F.W.; Krätzner, R. (1999). Lebendiger Planet – 4600 Millionen Jahre auf einen Blick. Göttingen: Planet Poster Edition. 1. Auflage. 34p.

Wittstock, U.; Halkier, B.A. (2002). Glucosinolate research in the *Arabidopsis* era. Trends in Plant Science, 7(6): 263-270.

Yano, M.; Lin, H.X.; Takeuchi, Y.; Nonoue, Y.; Ando, T.; Shomura, A.; Kono, I.; Takahashi, Y.; Liang, Z.W.; Ueda, T.; Sasaki, T. (2002). Marker-assisted dissection and pyramiding of complex traits in rice. International Rice Congress, Beijing, 16-20.09.2002. Abst.p.59

Yano, M.; Yamamoto, T. (1997). Construction of Substitution Lines with Overlapping Chromosomal Segments for Genetic Studies in Rice. Rice Genomic Newsletter 6(1):4-5 http://rgp.dna.affrc.go.jp/rgp/ricegenomenewslet/nl11/04.html

Zeb, A. (1998). Possibilities and Limitations of Feeding Rapeseed Meal to Broiler Chicks. Dissertation. Göttingen, http://webdoc.gwdg.de/diss/1998/zeb/inhalt.htm

Zhang, G.Q.; Ding, X.H.; Zeng, R.Z.; Zhang, Z.M.; Li, W.T.; Shah, B.; Ramalingam, J.; Liu, G.M.; Talukdar, F.H.; He, F.H.; Lu, Y.G. (2002). Progress of the rice molecular breeding program in SCAU. International Rice Congress, Beijing, 16-20.09.2002. Abstract: p.126

Zhao, J.; Meng, J. (2003). Detection of loci controlling seed glucosinolate content and their association with *Sclerotinia* resistance in *Brassica napus*. Plant Breeding 122:19-23

Ziegler, R. (2002). Ältester Hominide aus dem Tschad. Naturwissenschaftliche Rundschau, 55:613-614.

9. APPENDIX

9.1 Chemicals, Enzymes and Oligonucleotides

Adapters	Mwg Biotech AG
APS (Ammoniumpersulfat)	Serva
ATP	Sigma-Aldrich Chemie GmbH
Bromphenol blue	Sigma Chemical CO
Chloroform	Carl Roth GmbH
Deionized formamid	Merck KgaA
Dextran blue	Carl Roth GmbH Co
DMSO (Dimethyl Sulfoxide, C_2H_6SO)	Sigma-Aldrich Chemie GmbH
dNTP´s	PeqLab Biotechnologie GmbH
DTT Dithiothreitol	Sigma-Aldrich Chemie GmbH
Dye Hoechst 33258 (bisbenzimide)	Sigma-Aldrich Chemie GmbH
*Eco*RI	MBI Fermentas
EDTA	Carl Roth GmbH+Co
Ethidium Bromide	Sigma Chemical CO
Isopropanol	Carl Roth GmbH
Long Ranger™ 50% Gel solution BMA	BioWhittaker Molecular Applications Inc.
Low Melting MetaPhor™ Agarose	Biozym
Methanol LiChrosolv®	Merck
$MgCl_2$	PeqLab Biotechnologie GmbH
*Mse*I	BioLabs Inc.
Nucleon PhytoPure (RPN 8510)	Amersham Life Science
PeqGold Universal agarose	PeqLab Biotechnologie GmbH
Primers (AFLP and SSR)	Mwg Biotech AG
Primer MR13ULG (SSR primer)	Pharmacia Biotech
RAPD Primer	Qiagen - Operon
Rhodamin B	Merck
Size Standard IRDye 800 50-700 bp	Li-Cor Inc.
Sulfatase	Sigma-Aldrich Chemie GmbH
T4 DNA Ligase	Promega GmbH
Taq-DNA-Polymerase	PeqLab Biotechnologie GmbH
Temed	Sigma-Aldrich Chemie GmbH
Tris (hydroxymethil-) aminomethan	ICN Biomedicals INC.
Urea (NF-urea Rotiphorese®)	Carl Roth GmbH
Xylene cyanole FF	Sigma Chemical CO

9.2 Laboratory Equipment, Material and Software

Adobe®Photoshop 5.5	Adobe Systems Incorporated
AFLP-Quantar™ *Pro* 1.0	KeyGene Products B.V.
Biofuge 13R	Heraeus Sepatech
Biometra T-Gradient Thermoblock	Biometra Biomedizinische Analytik GmbH
CA-membrane Filter 0.45 µm	Sartorius AG
Cuvette (10x10x48 mm)	Cuvettes Sarstedt Aktien. & Co
DNA Analyser Gene Readir 4200	Li-Cor® (USA)
ECONO-Submarine Gel Unit SGE-300	C.B.S. Scientific CO – DelMar, CA, USA
Fluorometer (Versa Fluor™)	System 170-2402 (Bio-Rad Lab., Inc.)
FOSS NIRS instrument	Foss NIRS Systems/Tecator (6500)
Greenhouse Lamps	Phillips-400W, Sont Agro, 48000 lm
Gynkothek-HPLC-Machine	Gynkothek (Dionex Softron GmbH)
Kim Wipes	Kimberly-Clark Corporation
Micro Test Tubes Safe-Lock 1.5 ml	Eppendorf-Netheler-Hinz-GmbH
Pipette	Gilson S.A.
Polaroid MP-4 Land Camera	Polaroid
Polaroidfilm 667	Polaroid
Software ISI version 1.04	Infrasoft International, LLC.
Syringe (50ml)	B-D® - Becton Dickinson
ThermoStrip™ Standard Strip Tubes	PeqLab Biotechnologie GmbH
UV Transluminator	Bachofer Laboratoriunsgeräte
Videocamera	Mitsubishi, model P78E, Intas

9.3 DNA isolation

9.3.1 The method of DNA extraction

The DNA extraction was performed adding mercaptoethanol to the reagent n° 1 of the Nucleon PhytoPure kit to a concentration of 10 mM. It was also performed an RNA digestion at the initial step of the extraction by adding RNAse to a concentration of 20 µg/ml in the reagent n.1 and then it was incubated at 37°C for 30 minutes. 600 µl of the reagent n.1 was added to the 0.1g leaf material which has been previously grinded in liquid nitrogen until a flowing powder was reached. After mixing, 200 µl from reagent n. 2 was added. The whole material was then mixed again inverting the tubes several times. Followed an incubation at 65°C in a shaking water bath for 10 minutes. The next step was to place the samples on ice for 20 minutes

The samples were then removed from the ice and 500 µl cold (-20°C) chloroform (Carl Roth GmbH) was added, followed by 100 µl of Nucleon Phytopure DNA extraction resin

suspension. Samples were placed in a shaker (inverting the tubes) for 10 minutes. A centrifugation at 1,300g in the Biofuge 13R for 10 minutes was the next step; without disturbing the resin suspension layer, the upper DNA containing phase (above the brown resin layer) was transferred into a fresh tube using a pipette (Gilson). The upper phase was centrifuged again using 22,619 g (15,000 rpm Hettich-Zentrifugen GmbH) so that the original green and cloudy upper phase was cleared. In the bottom of the tube a little pellet of impurities was formed. The upper phase was again transferred to a new tube (2 ml) and an equal volume of cold isopropanol (Carl Roth GmbH) was added; the tubes were gently inverted to allow the precipitation of the DNA. The precipitation was performed overnight at 4°C.

The next day a centrifugation was done under 22,619 g (15,000 rpm) for 5 minutes to pellet the DNA. The upper isopropanol phase was discarded after centrifugation and new 70% ethanol was introduced to wash the DNA pellet. The tubes were submitted again to a new centrifugation at 22,619 g for 5 minutes. The supernatant was discarded, and the DNA pellet air-dry for 30 minutes. DNA was resuspended in TE Buffer (10 mM TrisHCl pH 8.0, 1 mM EDTA) and, according to Seiffert (2000) kept during one hour on a 65°C water bath to ensure a good resuspention. The TE buffer was added in a different volume to each sample depending on the size of the pellet, ranging from 40 μl for very small pellets until 100 μl for large pellets.

9.3.2 The measurement of the DNA concentration

For the DNA quantification all reagents were used following BIO-RAD´s Catalogue Number 170-2480 aiming at DNA concentrations between 200 ng/μl to 10 μg/μl. The materials used in the DNA quantification are described below.

The dye solution (0.1 μg/ml Hoechst 33258):

- 1 mg/ml Hoechst 33258 2.5 μl
- 10x TEN assay buffer 2.5 ml
- Sterile water 22.5 ml

The 10x TEN buffer consists of Tris (100 mM), NaCl (2 M), EDTA (10 mM), pH 7,4.

From this basic solution (the total volume needed was adjusted according to the number of samples to be measured) 2 ml were taken for each sample to be measured. The solution was always made fresh before use and placed in a bootle covered with aluminium foil (to be kept in the dark). Before starting the measurements a calibration is necessary with a 7 point standard curve, made using two basic DNA standard solutions, one of them with 100 μg/ml Calf Thymus DNA and the other one with 10 μg/ml.

DNA Standard solutions:

a) 100 μg/ml DNA Standard

- 1 mg/ml Calf Thymus DNA 100 μl
- 10x TEN assay buffer 100 μl
- Sterile water 800 μl

b) 10 μg/ml DNA Standard

- 1 mg/ml Calf Thymus DNA 10 μl
- 10x TEN assay buffer 100 μl
- Sterile water 890 μl

The solutions above were mixed in 1.5 ml tubes and stored at 4°C.

The standard curve was made measuring 7 samples which contained DNA amounts as follow:

1. 1000 ng
2. 500 ng
3. 200 ng
4. 100 ng
5. 50 ng
6. 20 ng
7. Empty (without DNA) (to set the fluorometer to „zero“).

The fluorometer was set to “zero” using the blank sample, which only contains the 2 ml of dye solution mentioned above. Immediately after setting it to zero, the 1000 ng sample cuvette (10x10x48 mm) was placed into the instrument and after waiting for 10 seconds the „range“ was set to 1000, given the maximal reference value. Then the other 5 samples were measured and the regression curve was calculated which was then used to adjust the values obtained for the plant samples.

9.4 AFLP: some more details

9.4.1 Preamplification

Detailed information on the preamplifications components:

*Eco*RI Primer E01:
5´ CTG CGT ACC ATT TCA 3´

*Mse*I Primer M02:
5´ GAT GAG TCC TGA GTA AC 3´

9.4.2 Amplification

*Eco*RI Primers used:

E32:
5´ CTG CGT ACC AAT TCA AC 3´

E33:
5´ CTG CGT ACC AAT TCA AG 3´

E38:
5´ CTG CGT ACC AAT TCA AT 3´

E35:
5´ CTG CGT ACC AAT TCA CA 3´

E40:
5´ CTG CGT ACC AAT TCA GC 3´

*Mse*I Primers used:

M47:
5´ GAT GAG TCC TGA GTA ACA A 3´

M49:
5´ GAT GAG TCC TGA GTA ACA G 3´

M59:
5´ GAT GAG TCC TGA GTA ACT A 3´

M50:
5´ GAT GAG TCC TGA GTA ACA T 3´

M61:
5´ GAT GAG TCC TGA GTA ACT G 3´

M62:
5´ GAT GAG TCC TGA GTA ACT T 3´

M48:
5´ GAT GAG TCC TGA GTA ACA C 3´

M60:
5´ GAT GAG TCC TGA GTA ACT C 3´

M51:
5´ GAT GAG TCC TGA GTA ACC A 3´

The loading buffer added to the amplified samples was prepared as follow:

- 1,9 ml deionized formamid (Merck KGaA)
- 40 µl EDTA (0,5M) (Carl Roth GmbH+Co)
- 20 µl NaOH (1M)
- 0.0025 g Rhodamin B (Merck KGaA)

9.4.3 Electrophoresis

The work on the Li-Cor IR^2 Gene Readir 4200 normally began with fixing the glass plates, always using the same sides of the two glass plates of a set as the gel side. The gel side from each glass plate was cleaned two times with distilled water and a white absorbent

(kitchen roll) paper, and two times at least with double-distilled ethanol using the special Kim wipes (Kimberly-Clark).

The two spacers (0.2 mm) were placed on either side, towards the margins of the large glass plate. After that, the small glass plate was taken and placed over the large one with the two spacers in position. The glass plates were fixed to the left and right rails.

The gels were prepared adding the components one by one (see chapter 2.2.1.2) to a 50 ml glass using a magnetic stirrer for mixing it. Polymerisation of the gel started immediately after the addition of APS and Temed, so it was always necessary to work fast after this point. A 50 ml syringe was used to draw the gel solution. Putting a CA - membrane filter (yellow filter 0.45 μm) on the syringe, the gel solution was then applied between the glass plates, so that it runs in between the two plates towards the bottom. A quick localization and removing of air bubbles with the bubble hook was always important to get good quality gels. After this step, the comb (64 tooth, was placed upside down in between the two glass plates to form a great slot. The casting plate was placed and the knobs from the rails were tightened. The gels take about two hours to polymerise.

After polymerisation, the 2 glass plates were cleaned again from the outside, first with water and paper and then with ethanol and paper so that the laser and microscope from the sequencer have clear view of the gel. The glass plates were put into the Li-Cor.

A solution of 1 x TBE buffer was prepared from a 10x concentrated one. With the 1x TBE buffer solution both, the upper and the lower buffer tanks, were filled. The comb was carefully removed and washed free (in H_2O) of gel particles. Using Kim wipes, the rests of the gel from the glass plate were removed and using a syringe with 1x TBE buffer from the upper tank the large slot made by the comb was also cleaned removing gel particles. An important detail to get good loads was apply, with a 20 μl pipette, around 20 μl loading buffer (0,1 g dextran blue and 2 ml deionized formamid) in the slot so that the limit (bottom) of the slot/gel could be easily identified. This step also was very useful to identify rests of gel inside the slot, and so it could be carefully removed with the bubble hook without piercing the gel line with the hook.

After this procedure, the slot was cleaned once more using the syringe and 1x TBE buffer. The loading buffer (Dextranblue + Formamide) was applied once more again to confirm the quality of the slot, and if it was good, the comb was introduced with the comb side down until it touches the gel line, and then start the pre run.

9.4.4 Scoring Gel Images

Following band finding parameters were used to screen the BC_2 generation:

Lane Profiles:

Half profile with	3
Profile smooth length	1
Background line length	500

Band finding:

	Top of image	*Bottom of image*
Min. band above background	500	500
Min. dist. between bands	4	8
Max. valley ratio	0.80	0.80
Max. Background percentage	40.00	40.00
Min. horiz. symmetry ratio	0.70	0.70
Max. horiz. symmetry ratio	1.30	1.30
Min. top valley	150	150
Max. shape match error	400	400
Band max search window	1	2

Score parameters:

	Top of image	*Bottom of image*
Min. distance between markers	7	3
Marker cursor vert. window	2	3

Percentage lane width for quantification: 0.80

Minimum band ratio for scoring a “?” (unknown): 0.20 (If it is decreased the Quantar Pro will find weaker bands not as a minus mark (-), but as a question mark (?)). Increases number of “?” scores.

Minimum band ratio for scoring a “+” (band present): 0.35 (If it is decreased the Quantar Pro will find weaker bands not as a minus (-) / question (?) mark, but as a positive). Increases number of “+” scores.

Market name mode: fragment size. That means, all markers are shown as a number indicating is size in base pairs, considering a fragment size standard as reference.

Fit parameters:

Population type: BC (Backcross)	Minimum likelihood: 0.20
Normalization: lane intensity	Maximum doubt data: 0.25
Transformation: square foot	Minimum class chance: 0.05
Fit confidence: 0.98	Minimum mu difference: 0.20

Following Band finding parameters were modified for the BC_3 generation according to Myburg and Remington (2000) and Myburg *et al.* (2001).

Lane Profiles:

Half profile with	2
Profile smooth length	1
Background line length	50

Band finding:

	Top of image	*Bottom of image*
Min. band above background	400	600
Min. dist. between bands	10	2
Max. valley ratio	0.9	0.9
Max. background percentage	50.00	50.00
Min. horiz. symmetry ratio	0.70	0.70
Max. horiz. symmetry ratio	0.5	0.5
Min. top valley	40	40
Max. shape match error	400	400
Band max search window	2	1

Score parameters:

	Top of image	*Bottom of image*
Min. distance between markers	10	2
Marker cursor vert. window	2	1

Percentage lane width for quantification: 0.50

Minimum band ratio for scoring a “?” (Unknown): 0.18

Minimum band ratio for scoring a “+” (band present): 0.22

Market name mode: fragment size.

Fit parameters:

Population type: BC (Backcross)	Minimum likelihood: 0.20
Normalization: lane intensity	Maximum doubt data: 0.25
Transformation: square foot	Minimum class chance: 0.05
Fit confidence: 0.98	Minimum mu difference: 0.20

9.5 The RAPD Analysis

The sequence of the primer is:
OQ09
$^{5'}$ GGCTAACCGA $^{3'}$

9.6 HPLC Procedures

Procedures:

1) For each sample a quantity of approximately 200mg seeds was taken and milled for 5 s with IKA mill with reduced volume of the milling room.
2) Ca. 200 mg milled seeds was put into a 70/12 PP (Polypropylene) tube, noticing the exact weight.
3) The tubes were heated in a water bath at 75°C for 1 minute.
4) 2 ml 70% Methanol was added in each tube.
5) 200 µl internal standard solution (6 mmol glucotropaeolin / l H_20)
6) Mixed on the vortex mixer
7) Extraction for 10 min, but mixed again after the first 5 min
8) Centrifugation for 4 min at about 2400 g
9) Supernatant transferred to new 70/12 PP tubes
10) To the remaining rest in the bottom of the old tubes 2 ml 10% Methanol is added and the steps 6 to 9 are repeated
11) Combine and mix the supernatants
12) Pippetting of the extract on top of a 20 mg sephadex DEAE-A-25 column (shortened Pasteur pipette) in the formiate form
13) Wash the columns twice with water
14) Add to each column 100 µl purified sulfatase (purified to 3.33* mg/ml) solution (1* : 2.5 water): desulfatation over night at 40°C.
15) Elution of the desulfoglucosinolates by adding 3 times 500 µl HPLC Water**/ column. Capture the effluent fraction in a 3 ml PS (Polystyrol) tube (**obtained using the ELGASTAT® UHQ).
16) From each sample (tube) 1.0 ml solution were taken with a syringe and the solution was filtered using a 45µm filter adapted to the bottom of the syringe.
17) 30-70 µl of the filtered solution was used to injection into the HPLC machine.

The analyses were done at room temperature using the GYNKOTHEK-HPLC-Machine, composed by the automatic sampler GINA 160, the UV-detector SP6, the pump

M480G and the SHIMADZU-Integrator C-R6A. The detection was made using a wavelength of 229 nm.

Separation was performed on a 250 mm and 3 mm column (Lichrosphy 100RP 18EC-5), at a flow rate of 0.9 ml/min, using a 5 x 4 mm pre-column (Lichrosphy), and a linear gradient like described below:

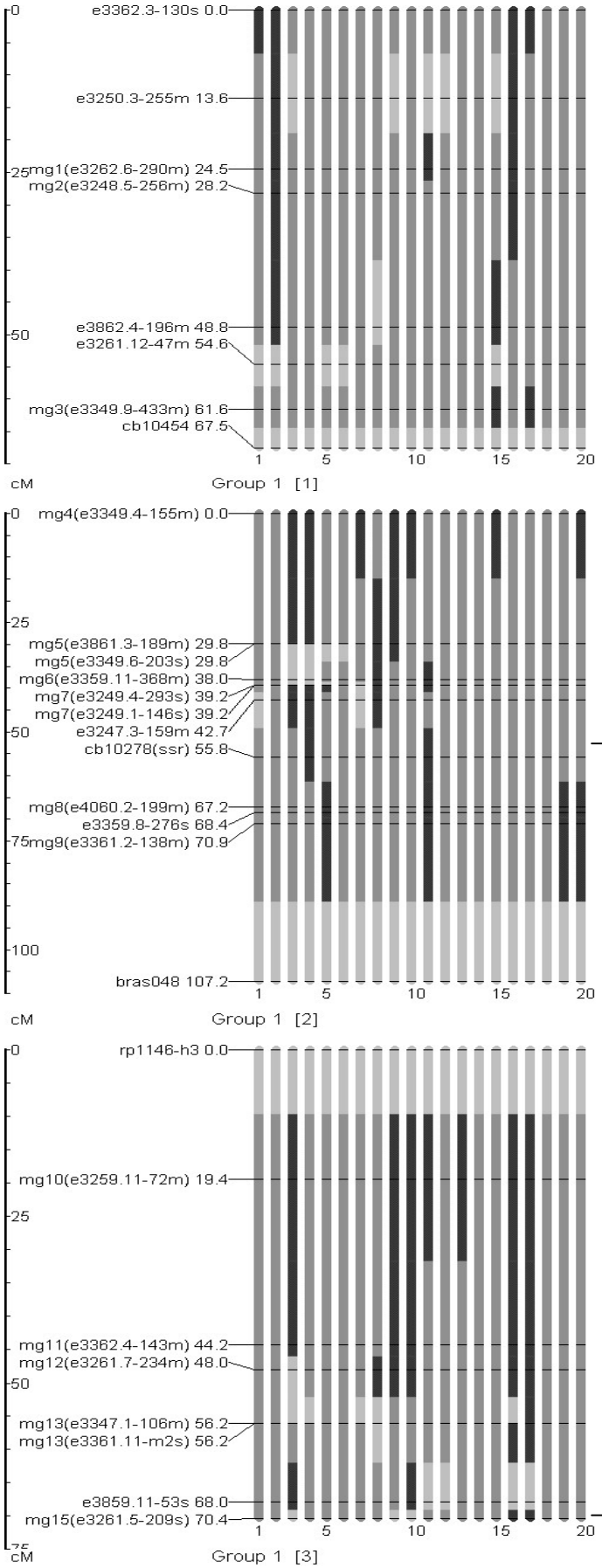
0-18.5 min:	1% AN	→	19% AN
18.5-20.5 min:	19% AN	→	1% AN
20.5-25.0 min		→	1% AN

AN = Acetonitrile in water

The column temperature was maintained at 35°C using the thermostat Techlab K1.

Glucotropaeolin (GTL) (isolated according to Thies, 1988, cited after Herrmann, 1992) in form of TMA-salt (Glucotropaeolin tetramethylamonium salt) were used as internal standard. The variety Linetta was used as standard for the method. The glucosinolate contend was calculated according to Buchner (1988) cited after Herrmann (1992).

9.7 Genotypes of the 20 selected BC_2 plants



QTL 2

QTL 3

Marker description:

Example: *E3262.3-130s*

AFLP Marker originated by the primer combination E33 vs M62. This marker has the number 3 and its length is 130 bp. The “s” at the end means that the dominant allele comes from the recurrent parent (Samourai), i.e., this marker was codominantly scored. On the other hand when a “m” appears, it means that the dominant allele comes from the donor parent (Mansholt), i.e., this marker was scored as a dominant one.

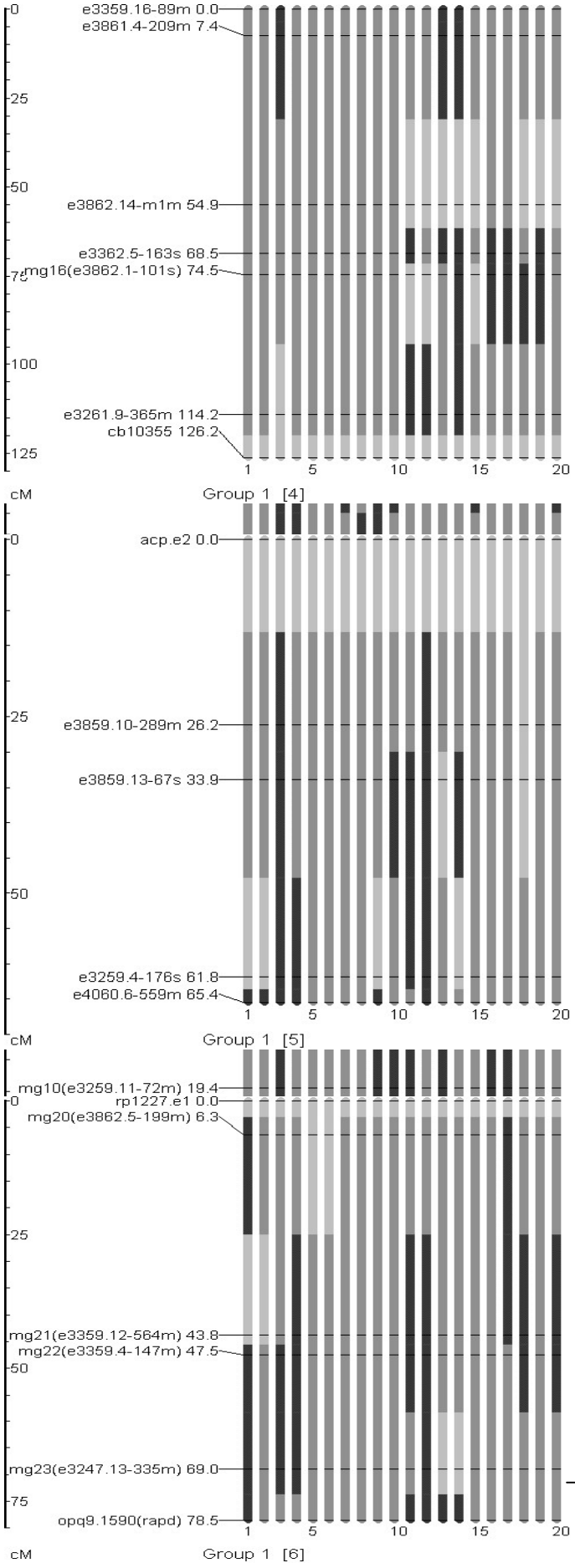
Colors:

Black: Donor Genome
Dark Grey: Recurrent Genome
Medium Grey: unknown

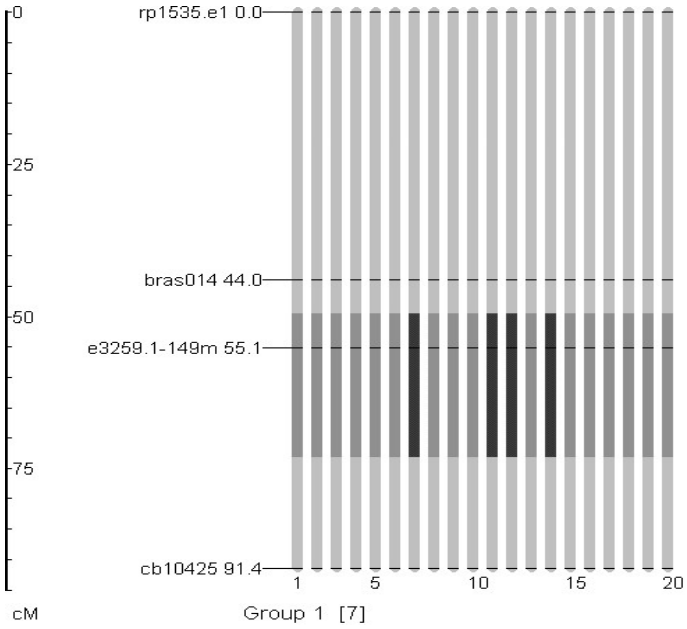
Abreviations:

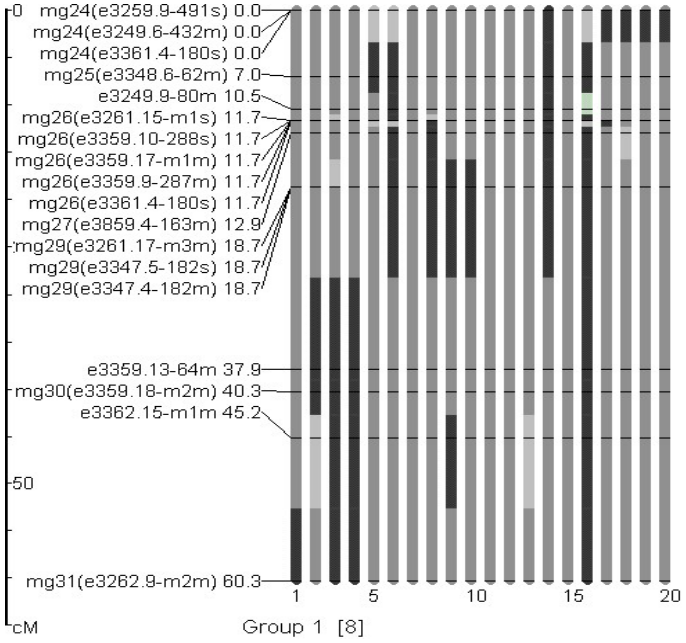
mg = marker group (that means that more markers are cosegregating)

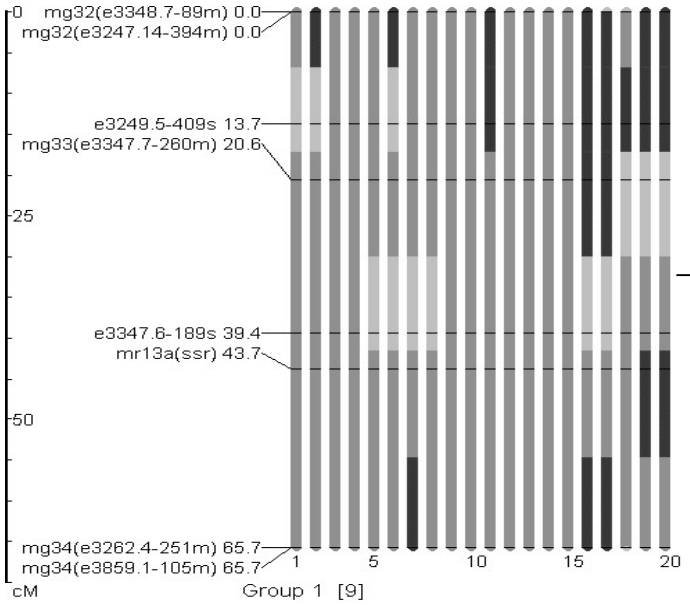
cb/bras/rp/acp/opq/mr/wg/opb/crt/db = markers not used for genotyping, only available here to show the mapped length of each linkage group (exceptions are the two SSR and one RAPD marker mentioned in Material and Methods).


0
e3359.16-89m 0.0
e3861.4-209m 7.4
25
50
e3862.14-m1m 54.9
e3362.5-163s 68.5
mg16(e3862.1-101s) 74.5
75
100
e3261.9-365m 114.2
cb10355 126.2
125
1 5 10 15 20
cM
Group 1 [4]
0
acp.e2 0.0
25
e3859.10-289m 26.2
e3859.13-87s 33.9
50
e3259.4-176s 61.8
e4060.6-559m 65.4
1 5 10 15 20
cM
Group 1 [5]
mg10(e3259.11-72m) 19.4
0
rp1227.e1 0.0
mg20(e3862.5-199m) 6.3
25
mg21(e3359.12-564m) 43.8
mg22(e3359.4-147m) 47.5
50
mg23(e3247.13-335m) 69.0
75
opq9.1590(rapd) 78.5
1 5 10 15 20
cM
Group 1 [6]

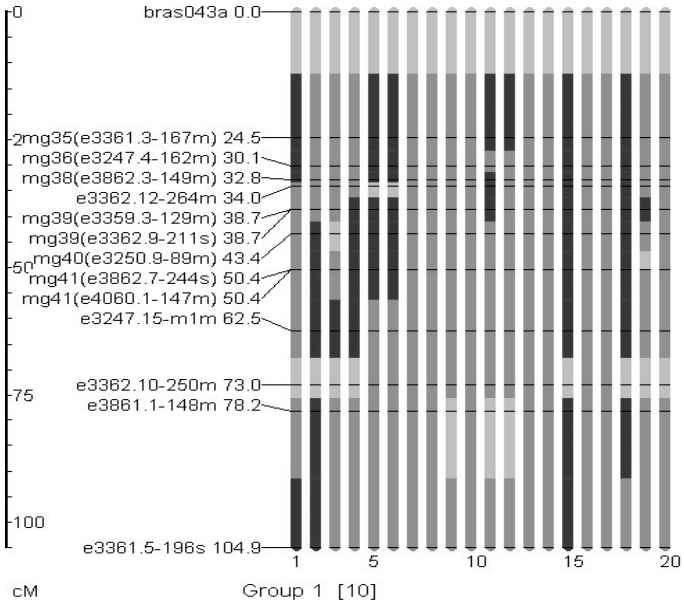
QTL 6

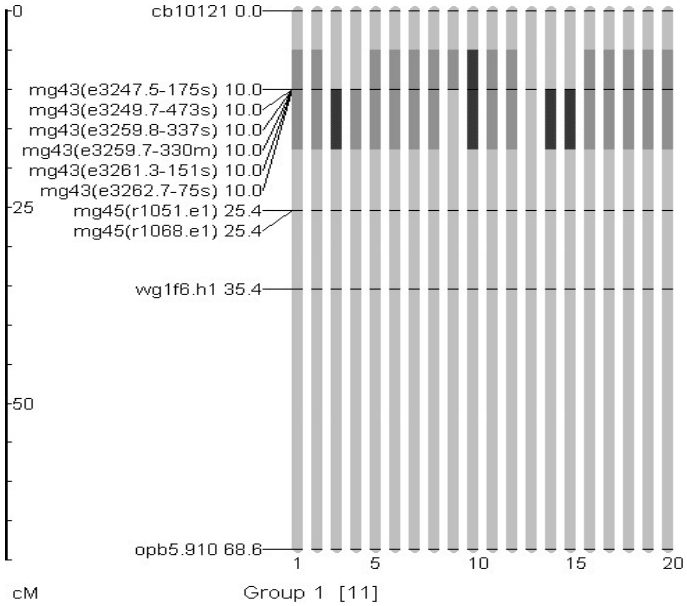

0
rp1535.e1 0.0
25
bras014 44.0
50
e3259.1-149m 55.1
75
cb10425 91.4
1
5
10
15
20
cM
Group 1 [7]

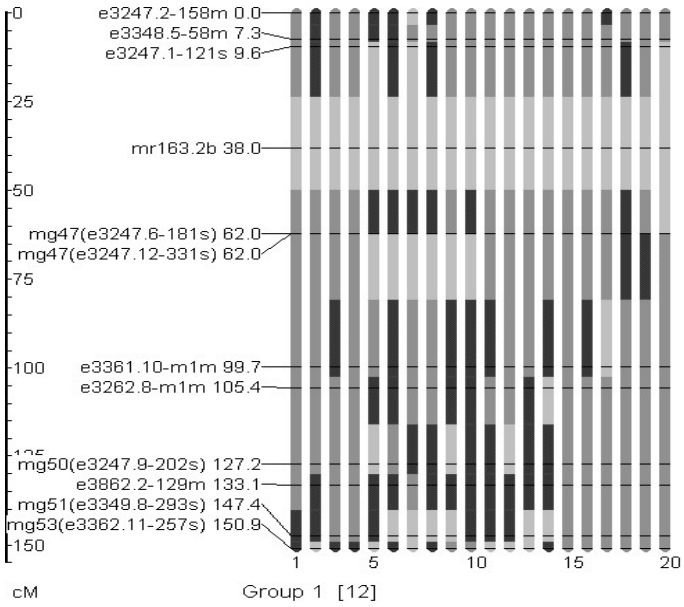

0 mg24(e3259.9-491s) 0.0
mg24(e3249.6-432m) 0.0
mg24(e3361.4-180s) 0.0
mg25(e3348.6-62m) 7.0
e3249.9-80m 10.5
mg26(e3261.15-m1s) 11.7
mg26(e3359.10-288s) 11.7
mg26(e3359.17-m1m) 11.7
mg26(e3359.9-287m) 11.7
mg26(e3361.4-180s) 11.7
mg27(e3859.4-163m) 12.9
mg29(e3261.17-m3m) 18.7
mg29(e3347.5-182s) 18.7
mg29(e3347.4-182m) 18.7
e3359.13-64m 37.9
mg30(e3359.18-m2m) 40.3
e3362.15-m1m 45.2
50
mg31(e3262.9-m2m) 60.3
1
5
10
15
20
cM
Group 1 [8]

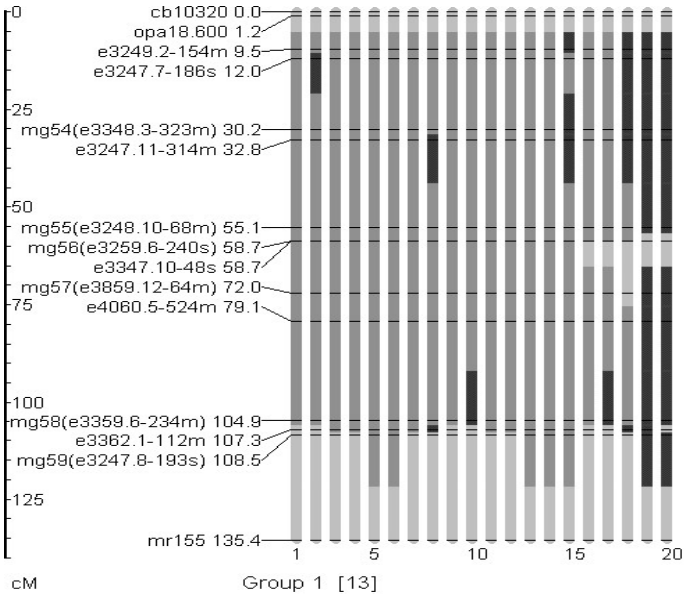

0 mg32(e3348.7-89m) 0.0
mg32(e3247.14-394m) 0.0
e3249.5-409s 13.7
mg33(e3347.7-260m) 20.6
25
e3347.6-189s 39.4
mr13a(ssr) 43.7
50
mg34(e3262.4-251m) 65.7
mg34(e3859.1-105m) 65.7
1
5
10
15
20
cM
Group 1 [9]

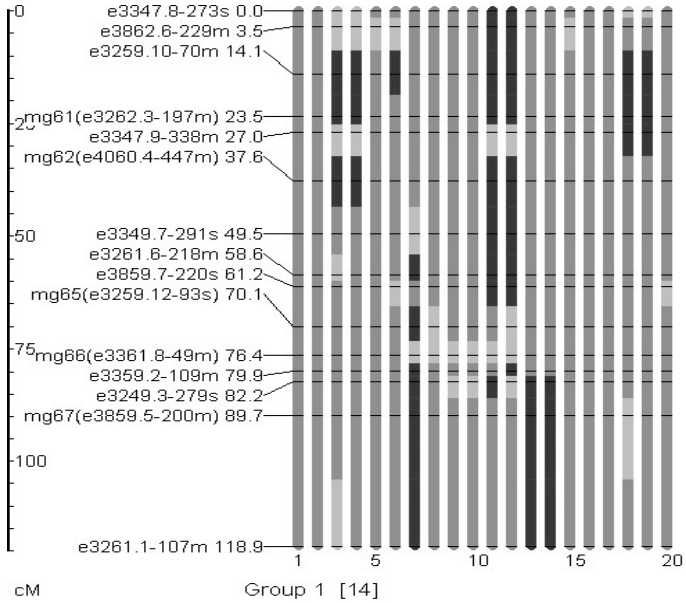
QTL 9

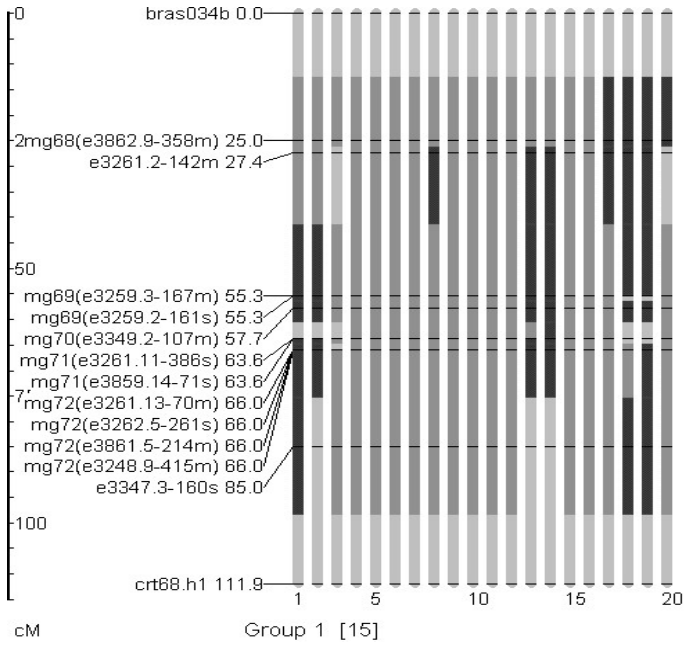

0
bras043a 0.0
mg35(e3361.3-167m) 24.5
mg36(e3247.4-162m) 30.1
mg38(e3862.3-149m) 32.8
e3362.12-264m 34.0
mg39(e3359.3-129m) 38.7
mg39(e3362.9-211s) 38.7
mg40(e3250.9-89m) 43.4
mg41(e3862.7-244s) 50.4
mg41(e4060.1-147m) 50.4
e3247.15-m1m 62.5
e3362.10-250m 73.0
e3861.1-148m 78.2
e3361.5-196s 104.9
cM
Group 1 [10]

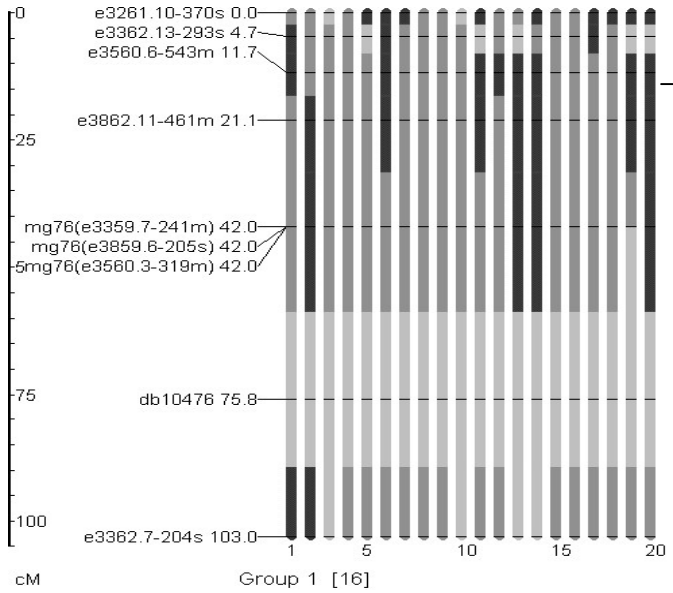

0
cb10121 0.0
mg43(e3247.5-175s) 10.0
mg43(e3249.7-473s) 10.0
mg43(e3259.8-337s) 10.0
mg43(e3259.7-330m) 10.0
mg43(e3261.3-151s) 10.0
mg43(e3262.7-75s) 10.0
mg45(r1051.e1) 25.4
mg45(r1068.e1) 25.4
wg1f6.h1 35.4
opb5.910 68.6
cM
Group 1 [11]


0
e3247.2-158m 0.0
e3348.5-58m 7.3
e3247.1-121s 9.6
mr163.2b 38.0
mg47(e3247.6-181s) 62.0
mg47(e3247.12-331s) 62.0
e3361.10-m1m 99.7
e3262.8-m1m 105.4
mg50(e3247.9-202s) 127.2
e3862.2-129m 133.1
mg51(e3349.8-293s) 147.4
mg53(e3362.11-257s) 150.9
cM
Group 1 [12]

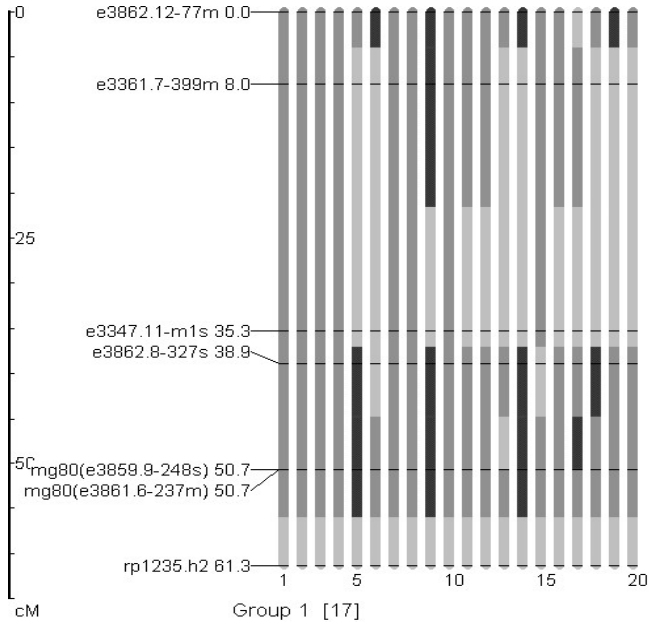

0
cb10320 0.0
opa18.600 1.2
e3249.2-154m 9.5
e3247.7-186s 12.0
25
mg54(e3348.3-323m) 30.2
e3247.11-314m 32.8
50
mg55(e3248.10-68m) 55.1
mg56(e3259.6-240s) 58.7
e3347.10-48s 58.7
mg57(e3859.12-64m) 72.0
75
e4060.5-524m 79.1
100
mg58(e3359.6-234m) 104.9
e3362.1-112m 107.3
mg59(e3247.8-193s) 108.5
125
mr155 135.4
1
5
10
15
20
cM
Group 1 [13]

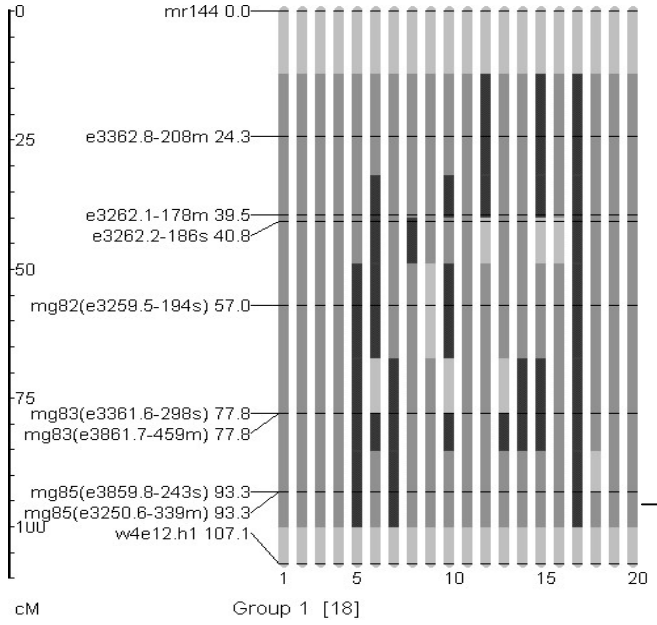

0
e3347.8-273s 0.0
e3862.6-229m 3.5
e3259.10-70m 14.1
mg61(e3262.3-197m) 23.5
e3347.9-338m 27.0
mg62(e4060.4-447m) 37.6
50
e3349.7-291s 49.5
e3261.6-218m 58.6
e3859.7-220s 61.2
mg65(e3259.12-93s) 70.1
75
mg66(e3361.8-49m) 76.4
e3359.2-109m 79.9
e3249.3-279s 82.2
mg67(e3859.5-200m) 89.7
100
e3261.1-107m 118.9
1
5
10
15
20
cM
Group 1 [14]


0
bras034b 0.0
mg68(e3862.9-358m) 25.0
e3261.2-142m 27.4
50
mg69(e3259.3-167m) 55.3
mg69(e3259.2-161s) 55.3
mg70(e3349.2-107m) 57.7
mg71(e3261.11-386s) 63.6
mg71(e3859.14-71s) 63.6
mg72(e3261.13-70m) 66.0
mg72(e3262.5-261s) 66.0
mg72(e3861.5-214m) 66.0
mg72(e3248.9-415m) 66.0
e3347.3-160s 85.0
100
crt68.h1 111.9
1
5
10
15
20
cM
Group 1 [15]

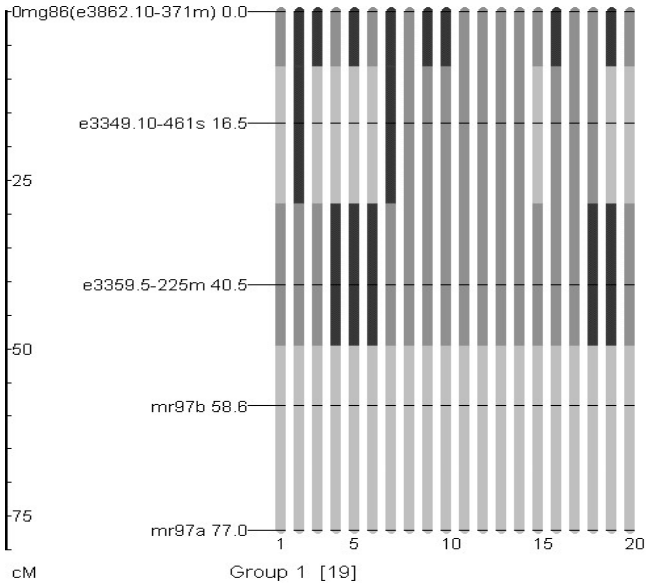

0
e3261.10-370s 0.0
e3362.13-293s 4.7
e3560.6-543m 11.7
e3862.11-461m 21.1
25
mg76(e3359.7-241m) 42.0
mg76(e3859.6-205s) 42.0
mg76(e3560.3-319m) 42.0
75
db10476 75.8
100
e3362.7-204s 103.0
1 5 10 15 20
cM
Group 1 [16]

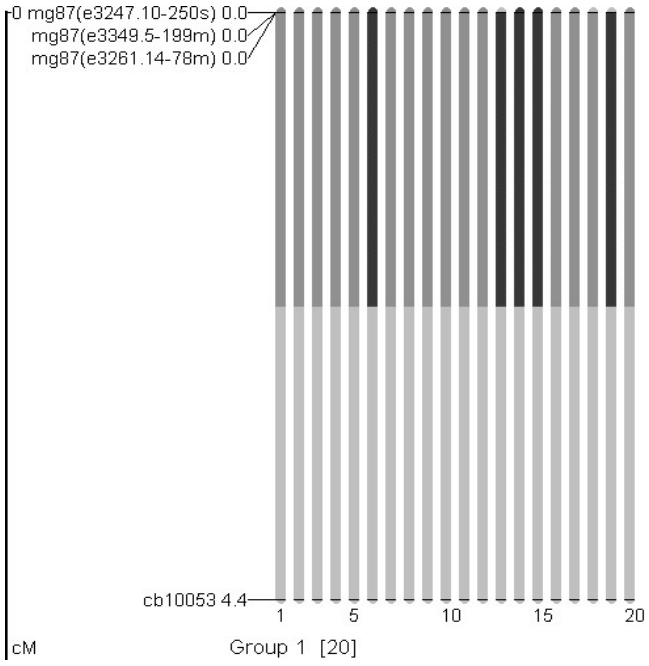
QTL 16


0
e3862.12-77m 0.0
e3361.7-399m 8.0
25
e3347.11-m1s 35.3
e3862.8-327s 38.9
mg80(e3859.9-248s) 50.7
mg80(e3861.6-237m) 50.7
rp1235.h2 61.3
1 5 10 15 20
cM
Group 1 [17]

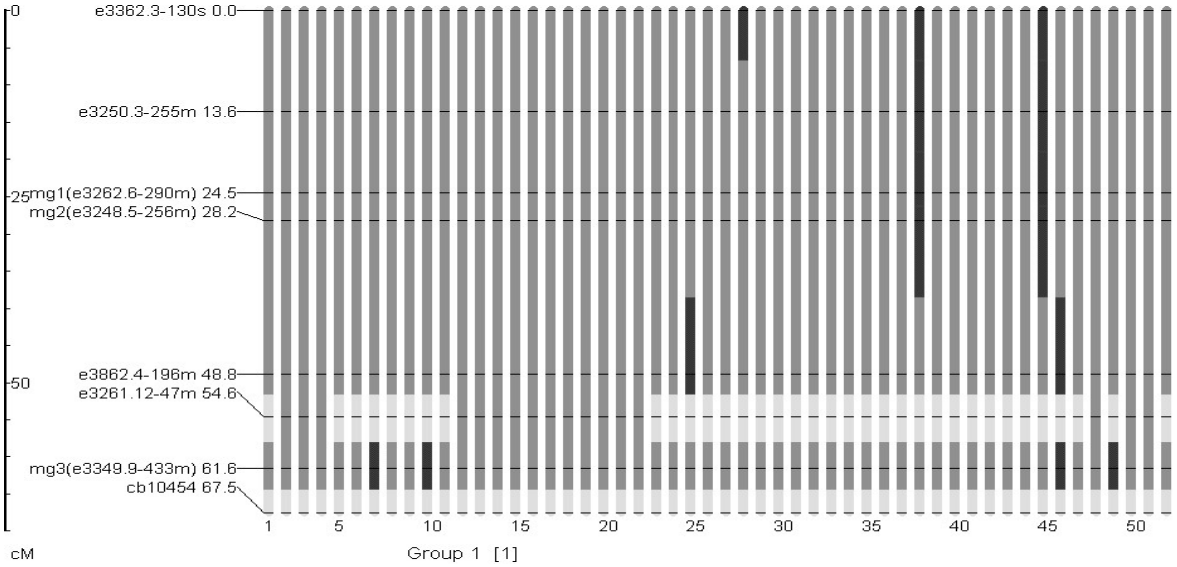

0
mr144 0.0
25
e3362.8-208m 24.3
e3262.1-178m 39.5
e3262.2-186s 40.8
50
mg82(e3259.5-194s) 57.0
75
mg83(e3361.6-298s) 77.8
mg83(e3861.7-459m) 77.8
mg85(e3859.8-243s) 93.3
mg85(e3250.6-339m) 93.3
100
w4e12.h1 107.1
1 5 10 15 20
cM
Group 1 [18]

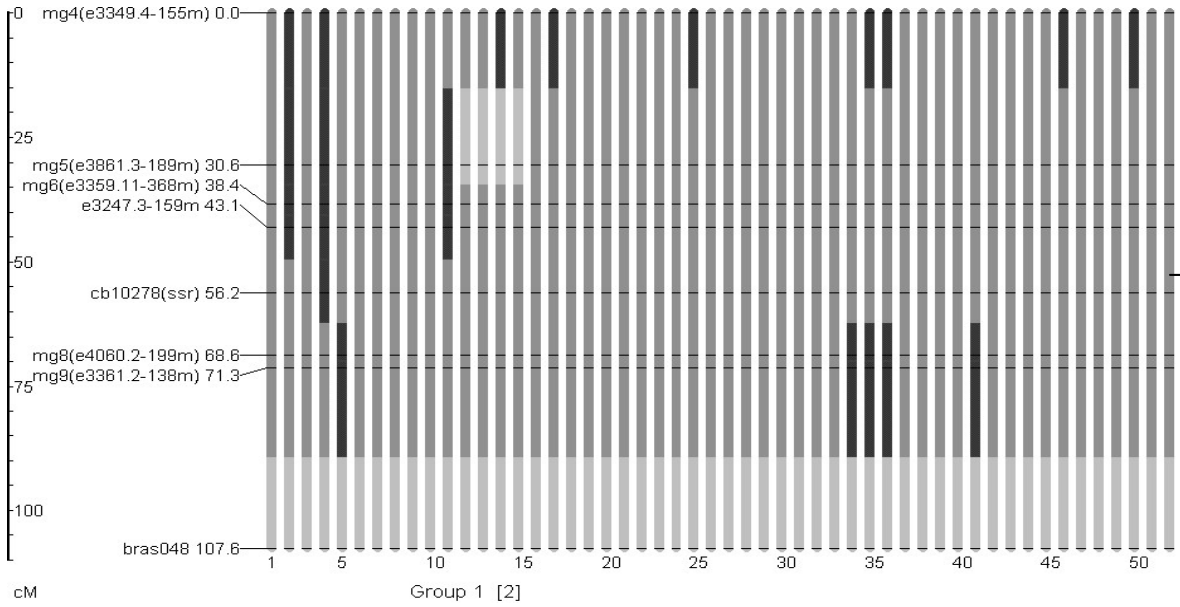
QTL 18


0mg86(e3862.10-371m) 0.0
e3349.10-461s 16.5
25
e3359.5-225m 40.5
50
mr97b 58.6
75
mr97a 77.0
1
5
10
15
20
cM
Group 1 [19]

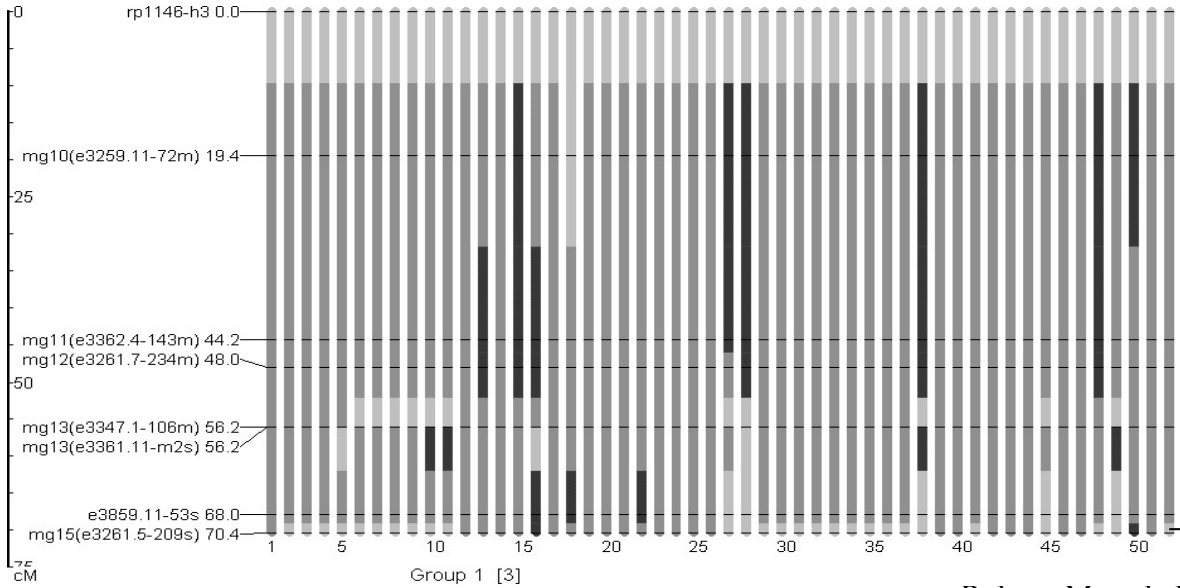

0 mg87(e3247.10-250s) 0.0
mg87(e3349.5-199m) 0.0
mg87(e3261.14-78m) 0.0
cb10053 4.4
1
5
10
15
20
cM
Group 1 [20]

9.8 Genotypes of the 52 selected BC_3 plants



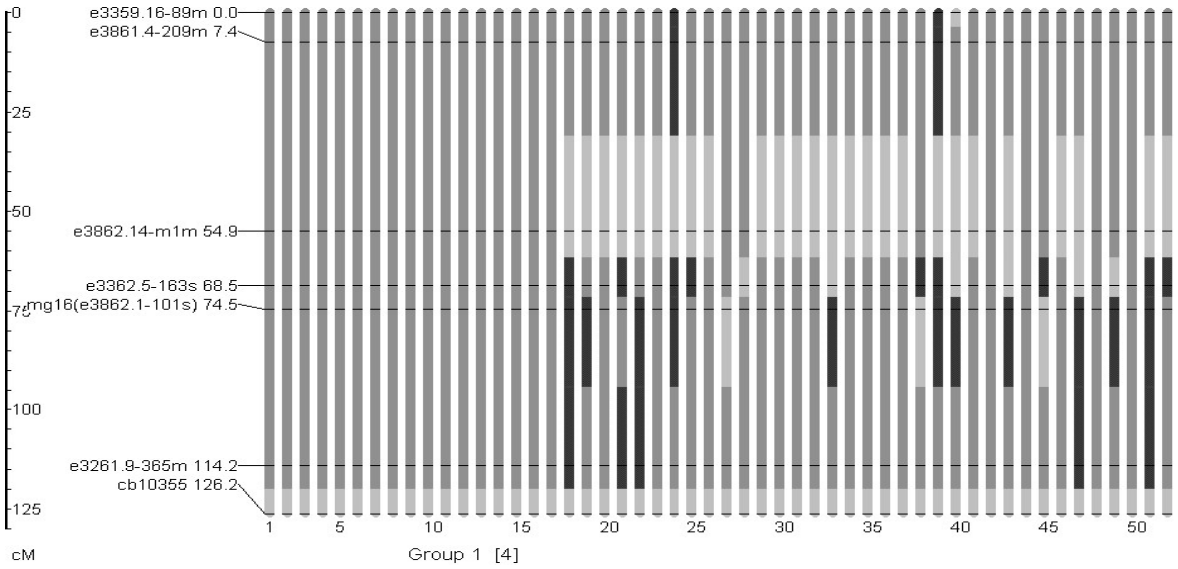


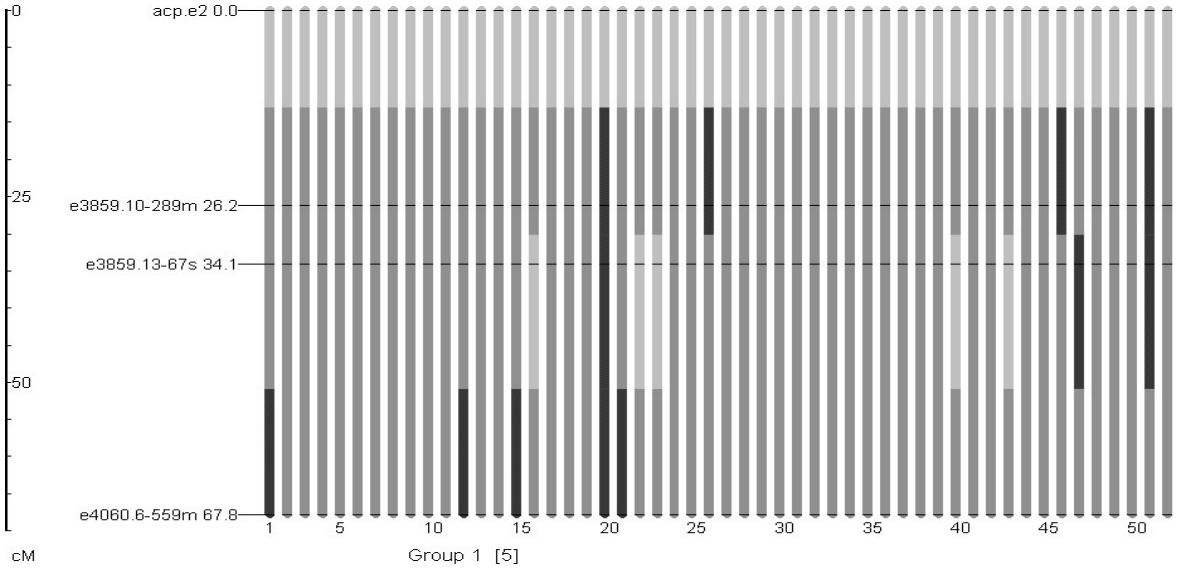
QTL 2

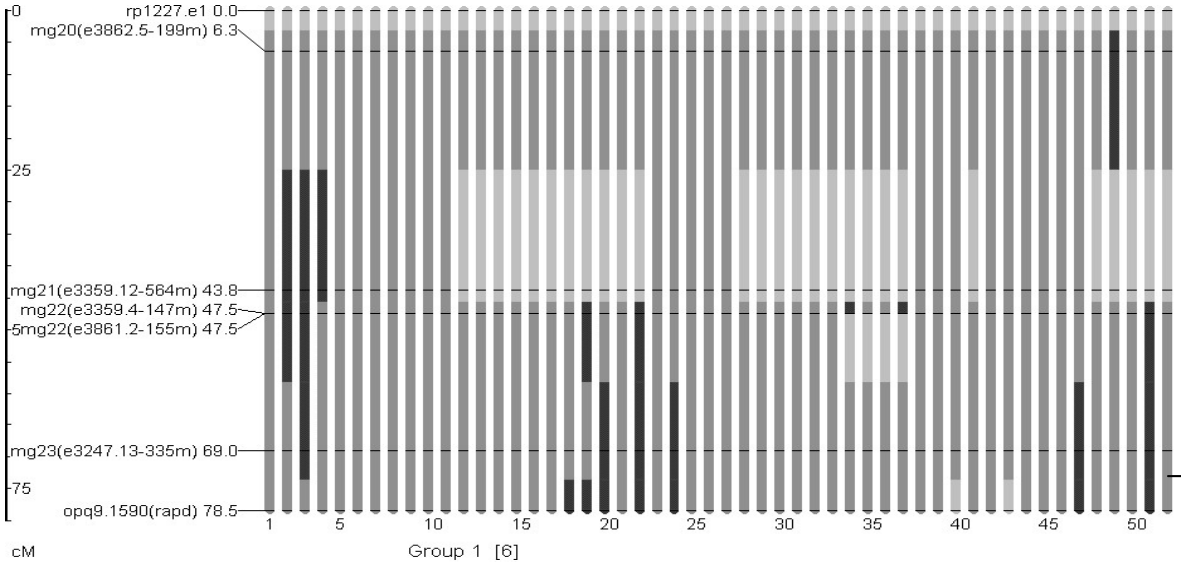


QTL 3

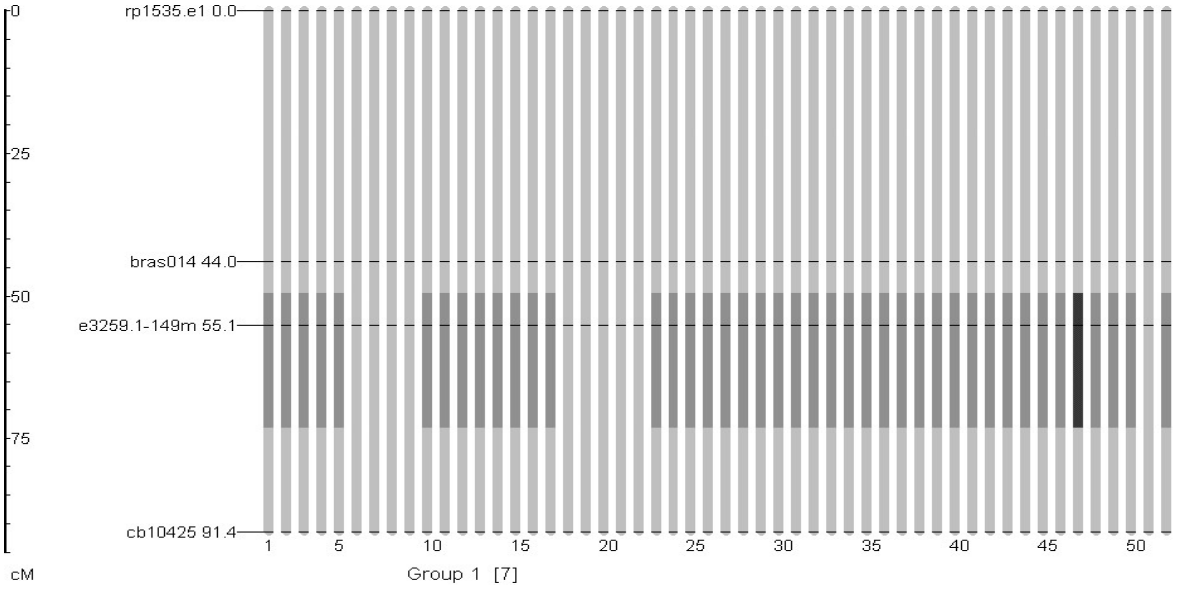
Rubens Marschalek

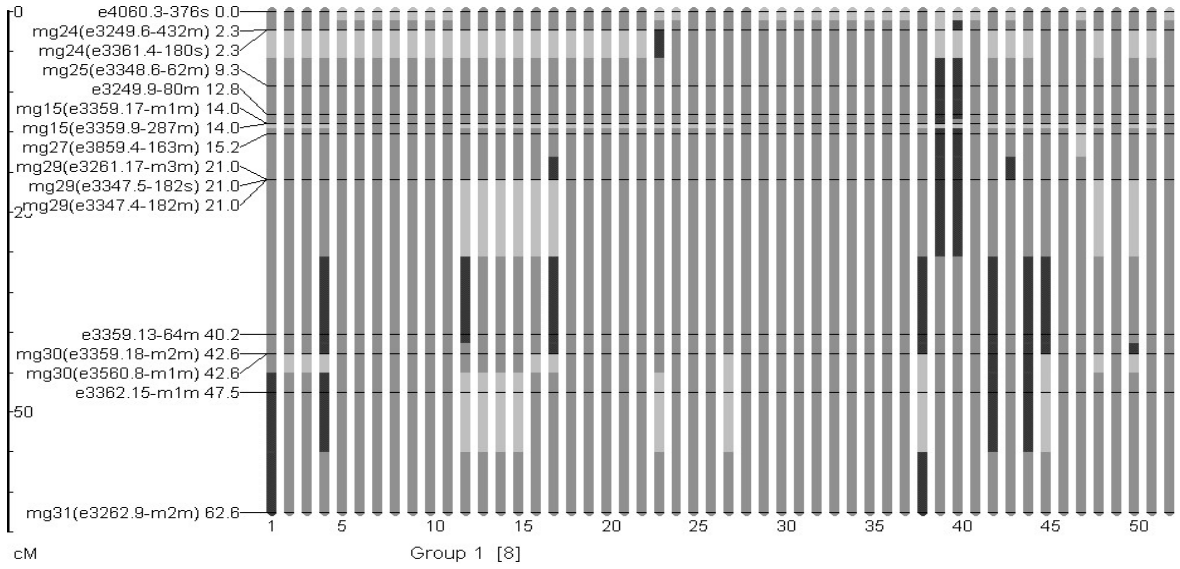

0
e3359.16-89m 0.0
e3861.4-209m 7.4
25
50
e3862.14-m1m 54.9
e3362.5-163s 68.5
mg16(e3862.1-101s) 74.5
75
100
e3261.9-365m 114.2
cb10355 126.2
125
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [4]

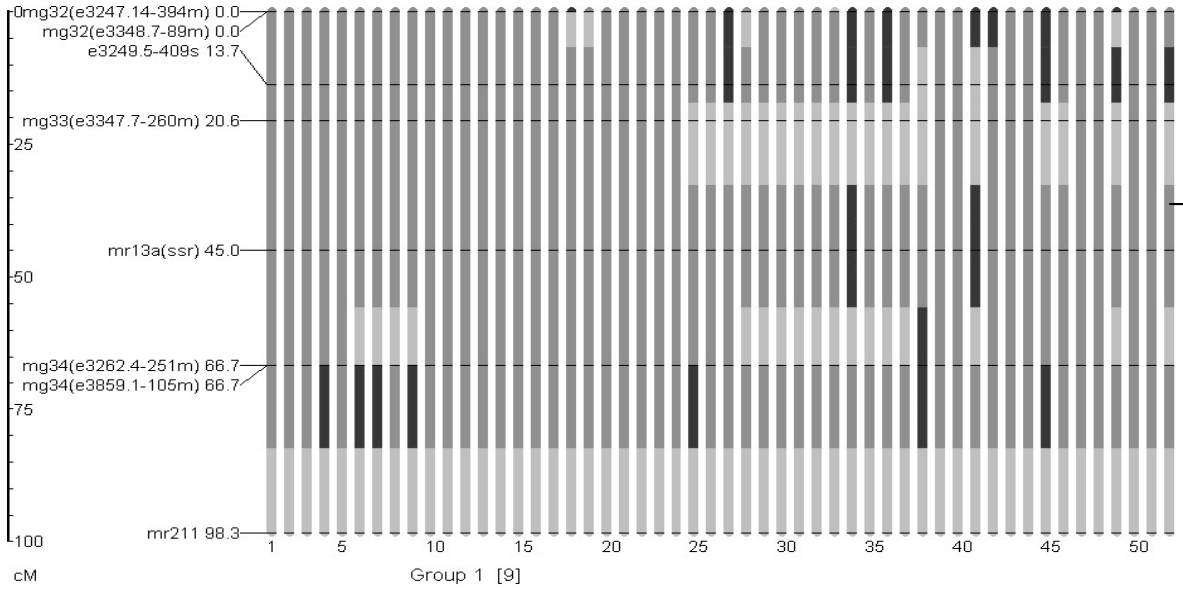

0
acp.e2 0.0
25
e3859.10-289m 26.2
e3859.13-67s 34.1
50
e4060.6-559m 67.8
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [5]


0
rp1227.e1 0.0
mg20(e3862.5-199m) 6.3
25
mg21(e3359.12-564m) 43.8
mg22(e3359.4-147m) 47.5
mg22(e3861.2-155m) 47.5
mg23(e3247.13-335m) 69.0
75
opq9.1590(rapd) 78.5
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [6]

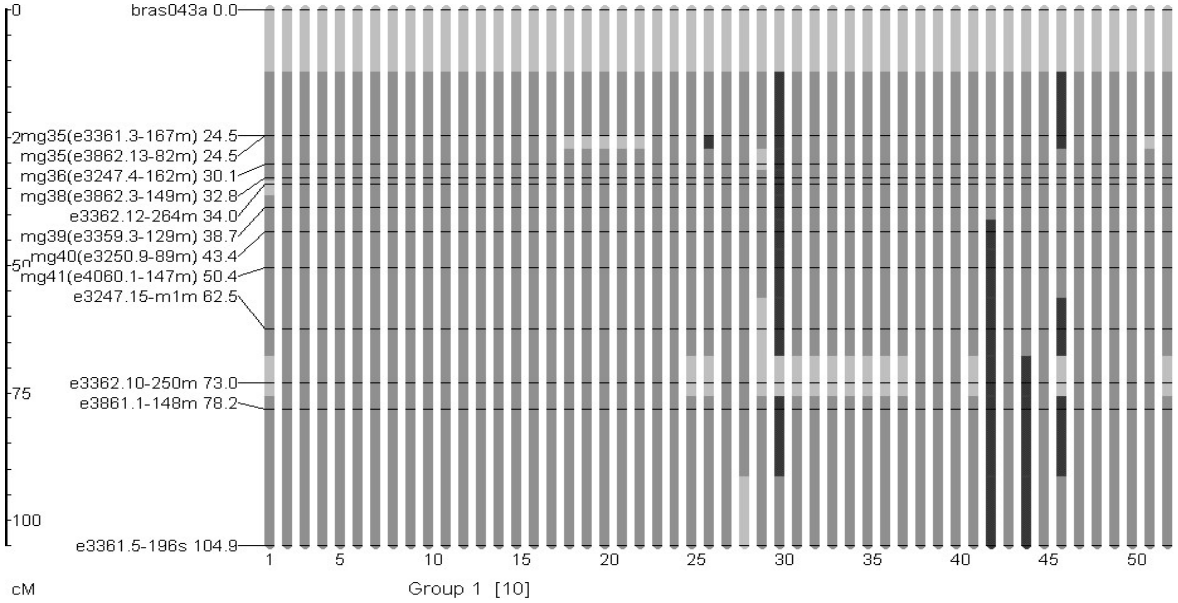
QTL 6

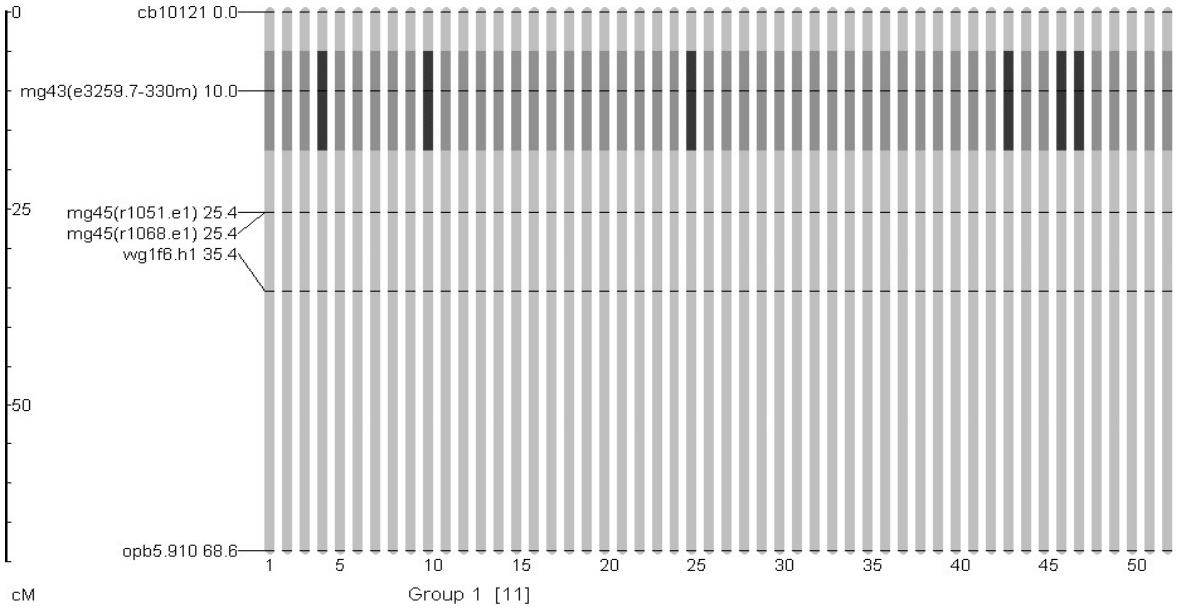

0
rp1535.e1 0.0
25
bras014 44.0
50
e3259.1-149m 55.1
75
cb10425 91.4
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [7]

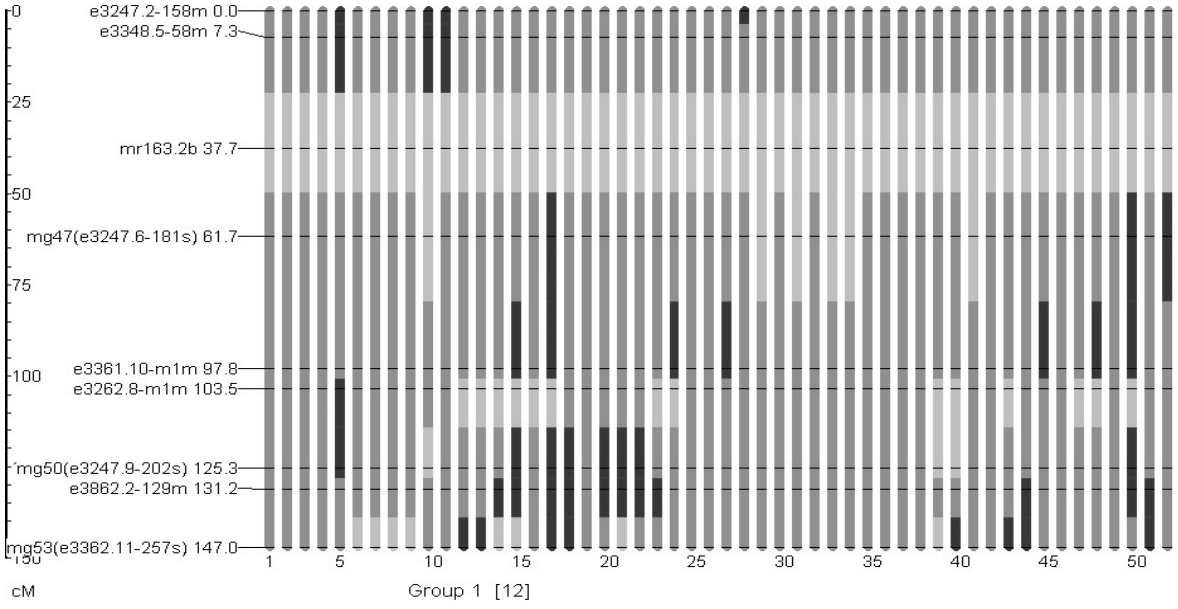

0
e4060.3-376s 0.0
mg24(e3249.6-432m) 2.3
mg24(e3361.4-180s) 2.3
mg25(e3348.6-62m) 9.3
e3249.9-80m 12.8
mg15(e3359.17-m1m) 14.0
mg15(e3359.9-287m) 14.0
mg27(e3859.4-163m) 15.2
mg29(e3261.17-m3m) 21.0
mg29(e3347.5-182s) 21.0
mg29(e3347.4-182m) 21.0
e3359.13-64m 40.2
mg30(e3359.18-m2m) 42.6
mg30(e3560.8-m1m) 42.6
e3362.15-m1m 47.5
50
mg31(e3262.9-m2m) 62.6
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [8]

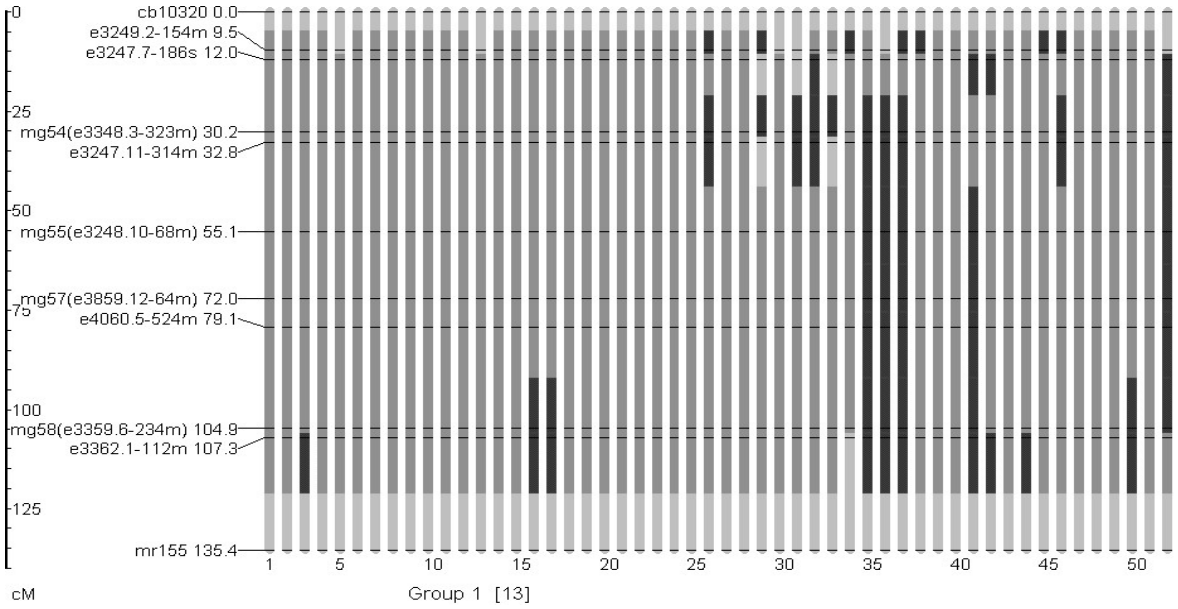

0mg32(e3247.14-394m) 0.0
mg32(e3348.7-89m) 0.0
e3249.5-409s 13.7
mg33(e3347.7-260m) 20.6
25
mr13a(ssr) 45.0
50
mg34(e3262.4-251m) 66.7
mg34(e3859.1-105m) 66.7
75
mr211 98.3
100
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [9]

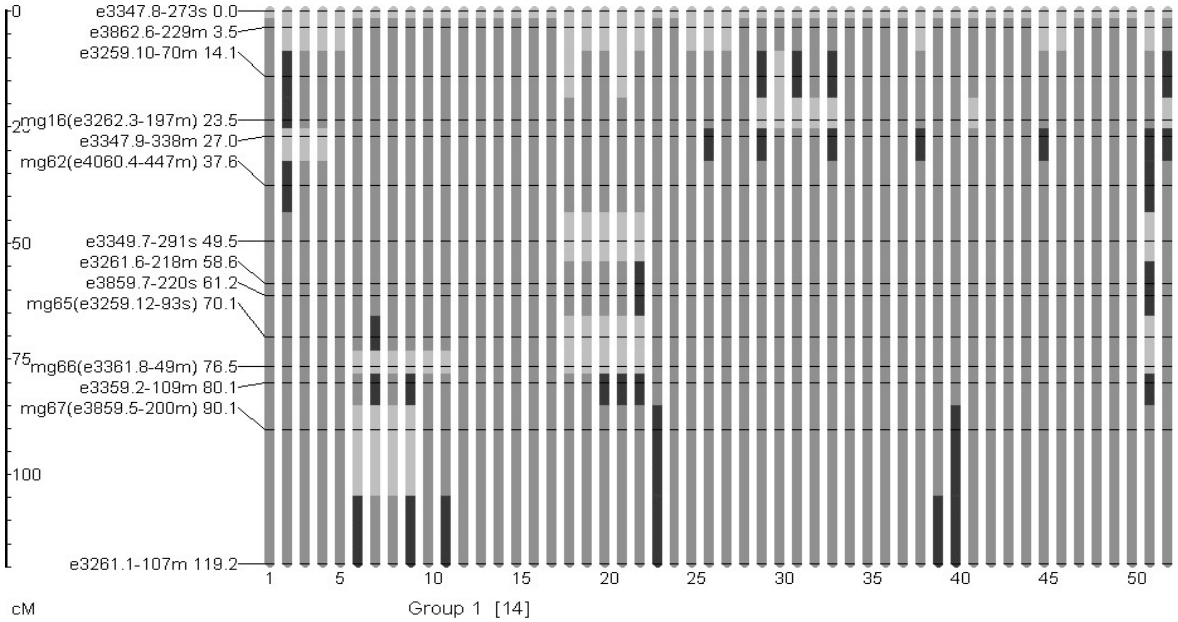
→ QTL 9

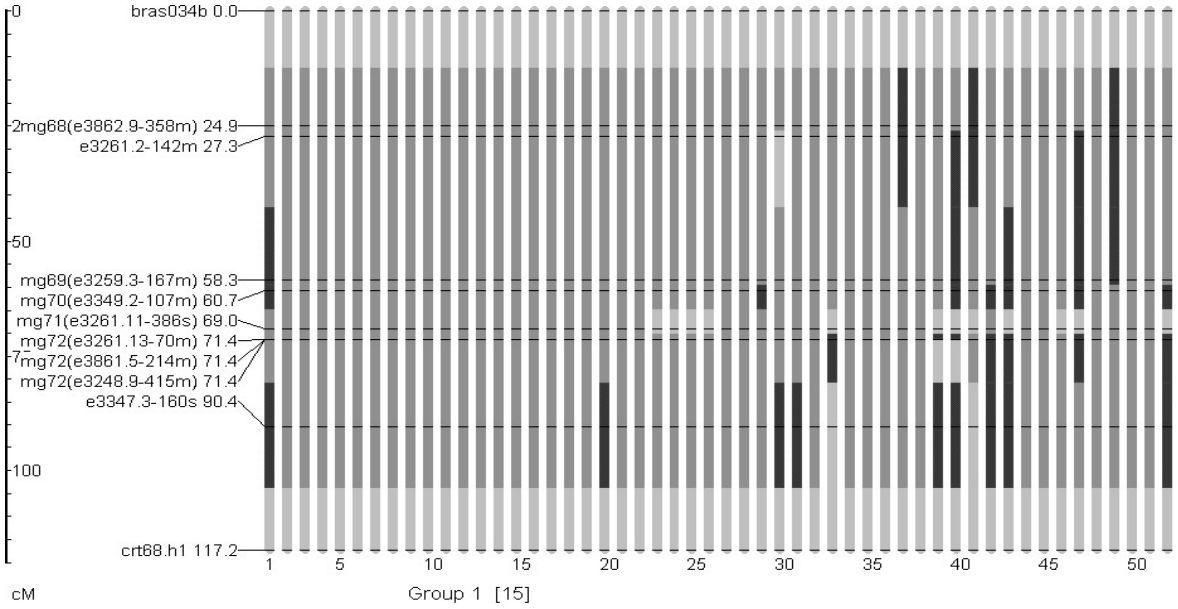

0
bras043a 0.0
mg35(e3361.3-167m) 24.5
mg35(e3862.13-82m) 24.5
mg36(e3247.4-162m) 30.1
mg38(e3862.3-149m) 32.8
e3362.12-264m 34.0
mg39(e3359.3-129m) 38.7
mg40(e3250.9-89m) 43.4
mg41(e4060.1-147m) 50.4
e3247.15-m1m 62.5
e3362.10-250m 73.0
e3861.1-148m 78.2
e3361.5-196s 104.9
cM
Group 1 [10]

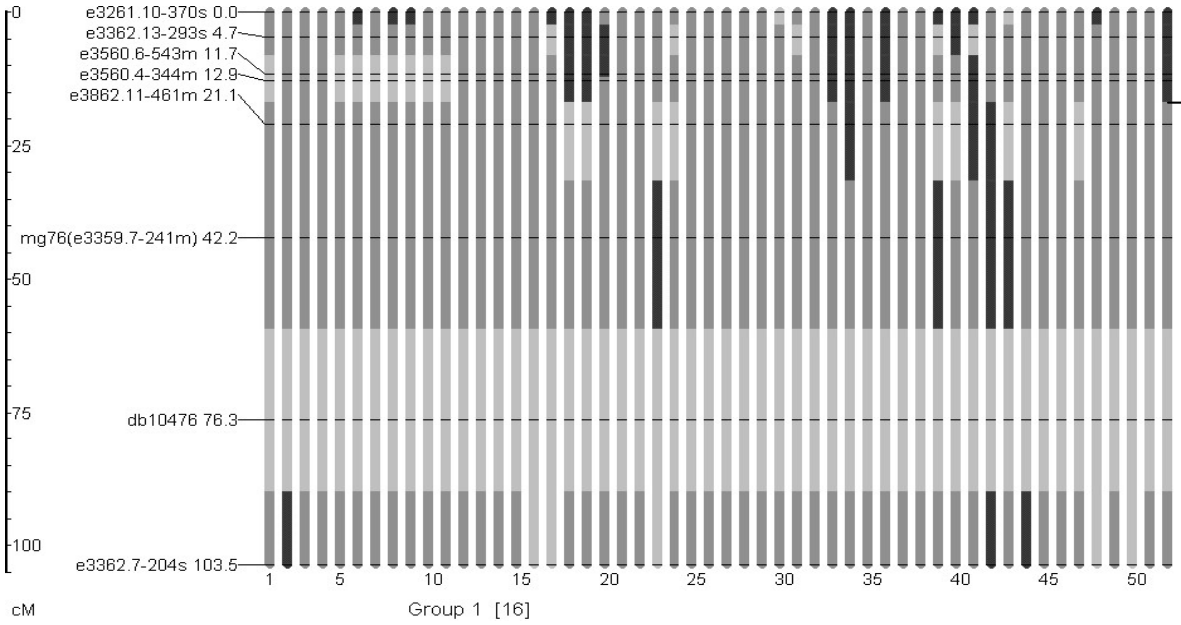

0
cb10121 0.0
mg43(e3259.7-330m) 10.0
mg45(r1051.e1) 25.4
mg45(r1068.e1) 25.4
wg1f6.h1 35.4
opb5.910 68.6
cM
Group 1 [11]


0
e3247.2-158m 0.0
e3348.5-58m 7.3
mr163.2b 37.7
mg47(e3247.6-181s) 61.7
e3361.10-m1m 97.8
e3262.8-m1m 103.5
mg50(e3247.9-202s) 125.3
e3862.2-129m 131.2
mg53(e3362.11-257s) 147.0
cM
Group 1 [12]

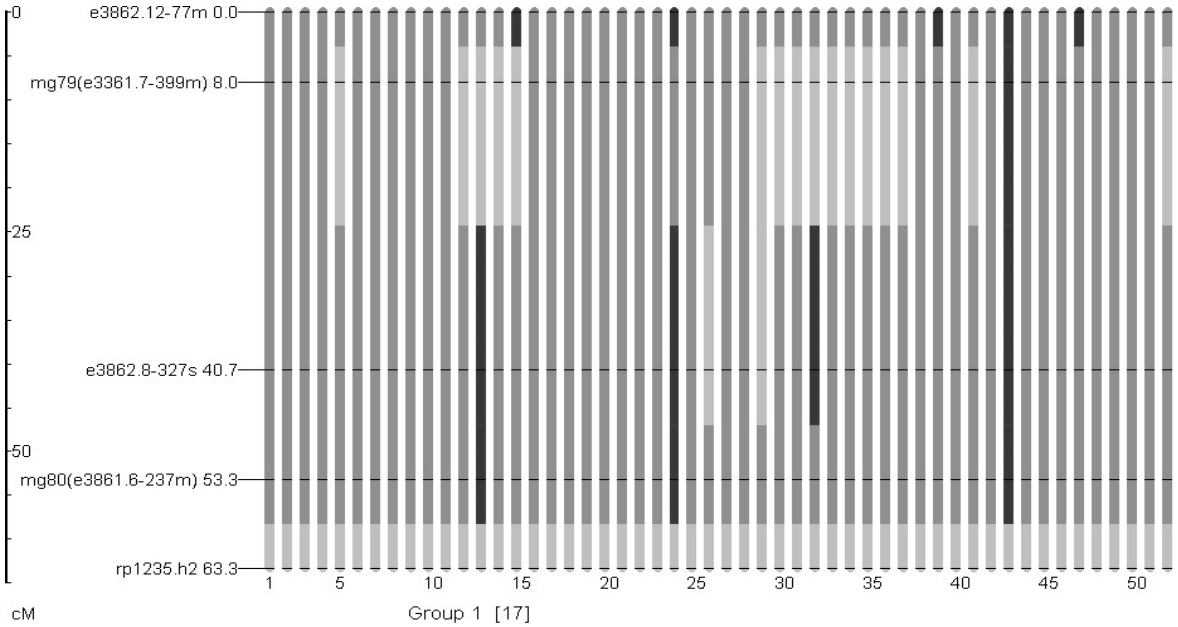

0
cb10320 0.0
e3249.2-154m 9.5
e3247.7-186s 12.0
25
mg54(e3348.3-323m) 30.2
e3247.11-314m 32.8
50
mg55(e3248.10-68m) 55.1
mg57(e3859.12-64m) 72.0
75
e4060.5-524m 79.1
100
mg58(e3359.6-234m) 104.9
e3362.1-112m 107.3
125
mr155 135.4
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [13]

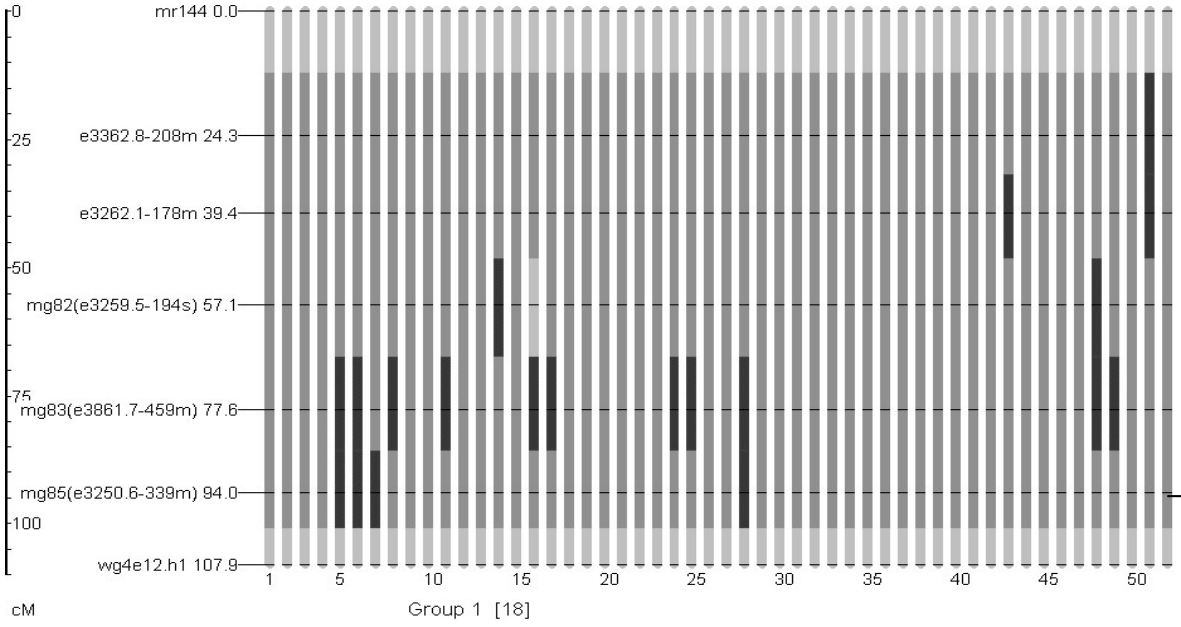

0
e3347.8-273s 0.0
e3862.6-229m 3.5
e3259.10-70m 14.1
mg16(e3262.3-197m) 23.5
25
e3347.9-338m 27.0
mg62(e4060.4-447m) 37.6
50
e3349.7-291s 49.5
e3261.6-218m 58.6
e3859.7-220s 61.2
mg65(e3259.12-93s) 70.1
75
mg66(e3361.8-49m) 76.5
e3359.2-109m 80.1
mg67(e3859.5-200m) 90.1
100
e3261.1-107m 119.2
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [14]


0
bras034b 0.0
2
mg68(e3862.9-358m) 24.9
e3261.2-142m 27.3
50
mg69(e3259.3-167m) 58.3
mg70(e3349.2-107m) 60.7
mg71(e3261.11-386s) 69.0
mg72(e3261.13-70m) 71.4
75
mg72(e3861.5-214m) 71.4
mg72(e3248.9-415m) 71.4
e3347.3-160s 90.4
100
crt68.h1 117.2
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [15]

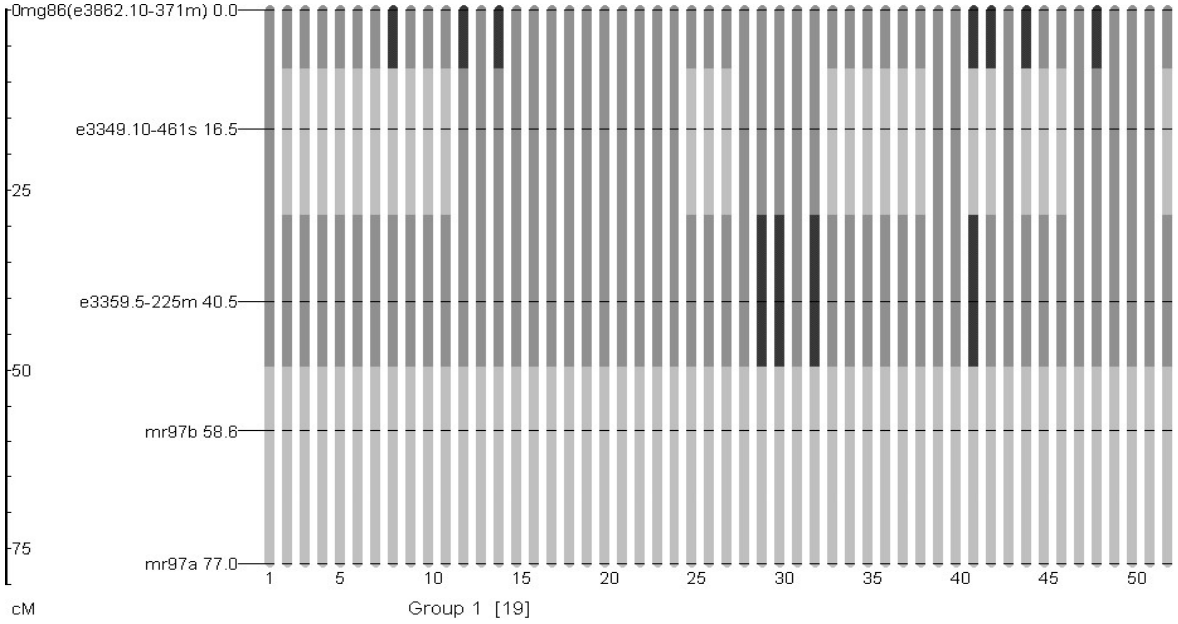

0
e3261.10-370s 0.0
e3362.13-293s 4.7
e3560.6-543m 11.7
e3560.4-344m 12.9
e3862.11-461m 21.1
25
mg76(e3359.7-241m) 42.2
50
75
db10476 76.3
100
e3362.7-204s 103.5
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [16]

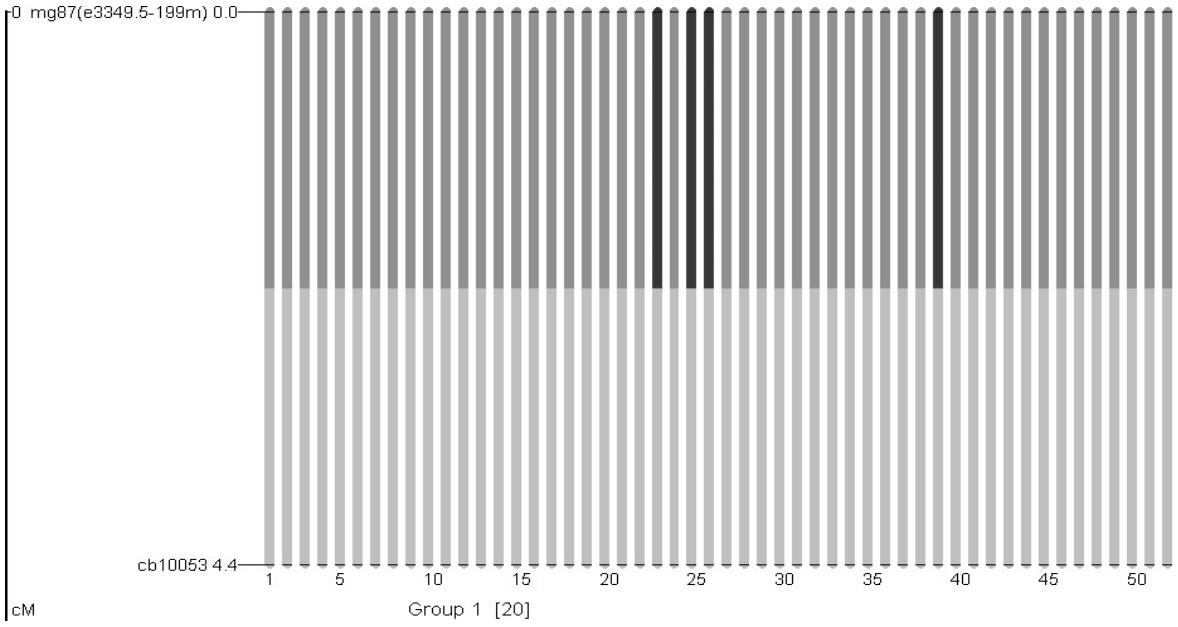
QTL 16


0
e3862.12-77m 0.0
mg79(e3361.7-399m) 8.0
25
e3862.8-327s 40.7
50
mg80(e3861.6-237m) 53.3
rp1235.h2 63.3
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [17]


0
mr144 0.0
25
e3362.8-208m 24.3
e3262.1-178m 39.4
50
mg82(e3259.5-194s) 57.1
75
mg83(e3861.7-459m) 77.6
mg85(e3250.6-339m) 94.0
100
wg4e12.h1 107.9
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [18]

QTL 18


0mg86(e3862.10-371m) 0.0
e3349.10-461s 16.5
25
e3359.5-225m 40.5
50
mr97b 58.6
75
mr97a 77.0
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [19]


0 mg87(e3349.5-199m) 0.0
cb10053 4.4
1 5 10 15 20 25 30 35 40 45 50
cM
Group 1 [20]

Plans fail for lack of counsel,
but with many advisers they succeed.
Proverbs 15:22 (Bible, New International Version)

Die Pläne werden zunichte, wo man nicht miteinander berät;
wo aber viele Ratgeber sind, gelingen sie.
Sprüche 15:22 (Bibel, Deutsche Bibelgesellschaft Stuttgart)

ACKNOWLEDGMENTS

I will be grateful for all my life to my wife Julia and my children Edgar and Sarah for his support, love, patience, encouragement and endless comprehension during these hard, but wonderful years. Without you I could never have accomplished it. I love you indeed! I hope, from now on, to have more time for you and for life.

Thanks God for providing me and my family with healthiness and so many blessings and for giving me enough persistence during these 4 years.

I want to express my sincerely thanks to my supervisors, Professor Dr. Heiko C. Becker and Dr. Wolfgang Ecke.

> Thank you Heiko for your encouragement and kindness. Thanks for your attention, support, and for keeping in mind one of my aims: the thesis. Thanks for providing suggestions and ideas during these years and for your valuable comments, corrections and help in this research and present dissertation.
>
> Wolfgang, I am grateful for your supervision. Thanks for trusting in a student without previous laboratory experience. Thank you also for the valuable suggestions on this dissertation and also for providing it with much better English.
>
> “Ich habe viel von Euch gelernt. Herzlichen Dank!“

Thanks also to Professor Dr. Reiner Finkeldey (Institute of Forest Genetics and Forest Tree Breeding) for giving me the possibility to attend his lectures and especially for his valuable suggestions and ideas regarding to this dissertation.

I am grateful to DAAD for the support during the language course in Göttingen (1999).

My special thanks to CNPq (National Council for Scientific and Technological Development – Ministry for Science and Technology of Brazil) for the scholarship provided during this PhD course. My sincerely thanks especially to Ms. Eli Ribeiro da Silva and Ms. Alcina Taitson Queiroz for the endless courtesy and kindness during this period.

Thanks to EPAGRI (Santa Catarina Agricultural Research and Rural Extension Agency – SC – Brazil) for the kind support during these years and for providing me the opportunity to

attend this course. Special thanks to: Tania B., Voltaire, Laércio, Dr. Piana, Dr. Murito, Carlos P. V., Euclides S., and to Dr. Satoru Yokoyama (in memorian).

Thanks also to Professor Dr. Heiko C. Becker and Professor Dr. Wolfgang Link for their very interesting lectures, and to the scientific staff at our institute for the valuable comments and discussions in the seminars. My acknowledgments also to Professor Dr. G. Röbbelen.

I would like to express my gratitude to Karla Galvão Bezerra dos Santos, Anja Kramer and Mahmoud Zeid for so many suggestions and advices. Thank you!!!

Thanks to: Bettina O., Andreas H., Gerald, Bernd, Andreas G., Volker, Nicole, Mary, Britta, Blazenka, Ingrid F., Sylvia, Christian, Ulrike, Adefris, Kemal, Memen, Andi, Agung, Seraphin, Brigitte, Hom, Sasongko (Joe), Alexandra, John, Thomas, Rosi, Antje, Bernd H., Nigussie, Joseph and Annie, Heise, S. Grube, Tala, Tesfaye, Jianyi, Lixi, Beriso, Seif, Christiane, Stefan. Thanks to the staff of the Göttingen Institute of Agronomy and Plant Breeding.

Thanks to Uwe Ammermann for his help on the HPLC analyses. My acknowledgments to Uwe, Dr. Johannes Hippe and to my brother-in-law, Carlos Eduardo P. Nascimento, for giving me so much support concerning the PC matters. Thank you!

Thanks to Thomas zum Felde for helping me with the NIRS.

I am grateful to Dr. Muhammet-Kemal Gül for the use of some field trials data.

I and my family are especially grateful to Henrich Kele, Margareta Keleova and their children Petra and Paula for the wonderful friendship during these years. You mean so much to us!

We, as a family, are deeply grateful to Andreas and Christel, Sabine and Walter, Hyung-Joo and Sun-Yong. Thank you very much for your friendship.

We were happy with the many good and unforgettable opportunities in which we could relax and simple keep laughing with Stephanie, Germano, Luciana and Sophia; Maria Helena and Ulli; Gisele, Henrique, Mariana and Yasmin; Karla, Martin and Isabelle; Paulina, Bernardo, Mayra and Marvin; Valderes, Gianna and Jacqueline.

Rosalba Badillo and Ayumu Endo, those memorable days during the German course will still remind me about this amazing time spent with you.

Thanks also to: Eveline and Herbert, Ellen and Helmut, Méri and Mark, Neri and Monika.

Thanks to my grandparents: Herta and Gustav August Isleb (in memorian), Erna (in memorian) and Erich Marschalek (in memorian). I wish you could be here!

Thanks to my sisters: Ruth and Irene; my sisters-in-law: Betina and Daniela.

Thanks to my parents, Heinz and Edeltraud Doracy, for giving me the chance to study so much. Thanks for your lovely support.

Rubens Marschalek

CURRICULUM VITAE

Rubens Marschalek

Birth date	December 13th 1965 in Blumenau, Santa Catarina State (SC), Brazil
Elementary School	1972-1979 in Blumenau - SC - Brazil
Secondary School	1980-1983 - Agricultural Technical School (Escola Técnica do Vale do Itajaí) ETEVI-FURB - Blumenau - SC - Brazil.
B. Sc. Degree	1984-1988 - Agronomy Engineering at the "Universidade Federal do Paraná" - UFPR - Curitiba - PR - Brazil
Professional Activities	June 1989 to August 1990 - Producer Cooperative Cooperjuriti, Massaranduba - SC - Brazil. Irrigated rice seed production and technical assistance.
Rural Extension Service	September 1990 to February 1992. ACARESC (called EPAGRI since 1991) Cities: Ibirama and Ascurra - SC - Brazil.
M. Sc. Degree	February 1992 to March 1995 at the "Universidade de São Paulo (USP), Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Departamento de Genética" - Piracicaba - SP - Brazil. Area: Plant Breeding and Genetics.
Rural Extension Service	October 1993 to March 1996. Land Management Project - EPAGRI/ IBRD World Bank - Massaranduba - SC - Brazil; Water and Soil Conservation.
Research	April 1996 to July 1999. Cassava Breeding. Itajaí Experiment Station - EPAGRI. Itajaí - SC - Brazil.
Faculty member	1997-1998. Teacher of Forest Tree Breeding at the "Universidade Regional de Blumenau – FURB" (Forest Engineering Department) Blumenau - SC – Brazil
Dr. Sc. Degree	October 1999 to September 2003 at the Institute of Agronomy and Plant Breeding - Georg August Universität - Göttingen - Germany

Itajaí Experiment Station – EPAGRI
www.epagri.rct-sc.br
Rodovia Antônio Heil km 6, Cx. Postal 277; Itajaí - SC - Brazil
CEP 88301-970 Telefone: 0055 (0) 47-341-5244 FAX: (0) 47-341-5255
rubensm@epagri.rct-sc.br rmarschalek@hotmail.com