Justus-Liebig-University Giessen
Research Centre for BioSystems, Land Use and Nutrition,
Department of Plant Breeding
Head: Prof. Dr. Dr. h.c. Wolfgang Friedt

Genetic diversity in *Brassica napus* and association studies with seed glucosinolate content

Dissertation
Submitted for the degree of Doctor of Agricultural Science
Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management
Justus-Liebig-University Giessen

Submitted by
Maen K. A. Hasan
from
Amman, Jordan

Giessen, 2008

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree of Doctor of Agricultural Science by Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management Justus-Liebig-University Giessen

Date of defence: July 02, 2008

Chairman of the examination committee: Prof. Dr. Ernst-August Nuppenau

Supervisor:	Prof. Dr. Dr. h.c. Wolfgang Friedt
Co-supervisor:	Prof. Dr. Wolfgang Köhler
Examiner:	Prof. Dr. Sylvia Schnell
Examiner:	Prof. Dr. Andreas Vilcinskas

To My Parents with Love

TABLE OF CONTENTS

I INTRODUCTION

1.1 Rapeseed

Rapeseed (*Brassica napus* L., genome AACC, 2n = 38) is a relatively young species that originated in a limited geographic region through spontaneous hybridisations between turnip rape (*B. rapa* L., AA, 2n = 20) and cabbage (*B. oleracea* L., CC, 2n = 18) genotypes (Kimber and McGregor 1995), resulting in an amphidiploid genome comprising the full chromosome complements of its two progenitors. Because no wild *B. napus* forms are known, it is believed that the species arose relatively recently, in the Mediterranean region of south-western Europe (Cruz et al. 2007, Friedt et al. 2007). The species is divided into two subspecies, namely *B. napus* ssp. *napobrassica* (swedes) and *B. napus* ssp. *napus*, which includes winter and spring oilseed, fodder and vegetable forms. The latter include the distinct leaf rapes (*B. napus* ssp. *napus* var *pabularia*), which used to be common as a winter-annual vegetable (Snowdon et al. 2006). Rapeseed cultivars are classified as winter or spring types according to their vernalisation requirement in order to induce flowering.

Winter cultivars are usually higher yielding than spring cultivars, but they can only be grown profitably in areas where they regularly survive the winter (Butruille et al. 1999). Oilseed rape is cultivated predominantly as winter or semi-winter forms in Europe and Asia, respectively, whereas spring-sown canola types are more suited to the climatic conditions in Canada, northern Europe and Australia (Friedt et al. 2007).

The majority of oilseed rape cultivars are pure lines derived from breeding schemes designed for self-fertilizing crops, i.e. pedigree selection or modifications thereof (Snowdon et al. 2006). Although for many years the emphasis in oilseed rape breeding was strongly focussed on open-pollinating varieties, up to 30% heterosis for seed yield has been reported for *B. napus* (e.g. Schuster 1969, Grant and Beversdorf 1985, Lefort-Buson et al. 1987, Brandle and McVetty 1989, Gehringer et al. 2007), and for both winter rapeseed and spring canola hybrid

varieties have rapidly gained importance over the past decade as effective systems for controlled pollination were developed. In Germany the first restored winter rapeseed hybrids were released in 1995 (Snowdon et al. 2006).

Today oilseed rape (*B. napus* ssp. *napus*) is the most important source of vegetable oil in Europe and the second most important oilseed crop in the world after soybean (data from FAOstat: http://faostat.fao.org/). The seeds of modern varieties typically contain 40 to 45% oil, which provides a raw material for many other products ranging from rapeseed methyl ester (biodiesel) to industrial lubricants and hydraulic oils, tensides for detergent and soap production and biodegradable plastics (Friedt et al. 2007). After oil extraction the residual meal, which contains 38-44% of high quality protein, is used in livestock feed mixtures. However the nutritional value of rapeseed meal is compromised by the presence of glucosinolates, a group of secondary compounds typical for crucifer plant species. Leaf glucosinolates play an important role in interactions with insect pests and pathogens. On the other hand, high intakes of seed meal glucosinolates and their degradation products in livestock feeds can cause problems of palatability and are associated with goitrogenic, liver and kidney abnormalities (Walker and Booth 2001). This particularly limits the use of the rich-protein meal from seeds of oilseed rape as a feed supplement for monogastric livestock. Hence, there is a strong interest in seed-specific regulation and optimisation of glucosinolate levels and composition, in order to improve the nutritional value of rapeseed meal without compensating the disease and pest resistance properties in the crop (Wittstock and Halkier 2002). In contrast to soybean meal, rapeseed meal is not widely used for human consumption (Snowdon et al. 2006).

1.2 Genetic diversity in crop plants

About 12,000 years ago, a group of humans living in the historic Fertile Crescent made the first shift from hunter-gathering to cultivating plants for sustained survival (Salamini et al. 2002), giving rise to the domesticated breeds that today form the foundation of the world's food supply (McCouch 2004). These food crops were first domesticated from wild species (Tanksley and McCouch 1997). The repeated cultivation and maintenance of these selected plants under site-specific conditions led to landraces that are highly adapted to their respective growing conditions and production methods (Friedt et al. 2007). Landrace varieties are the

earliest form of cultivars and hence represent the first step in the domestication process (McCouch 2004). However, many crop landraces were lost as farmers throughout the world shifted to growing high-yielding varieties (Zamir 2001).

The general trend of agriculture during the past half-century has been the release and cultivation of improved cultivars of most major and minor crop species (Rao and Hodgkin 2002). These cultivars generally carry only a fraction of the variation present in the gene pool of the respective species (Fernie et al. 2006). Unlike the highly heterogeneous landraces, which were selected for subsistence agricultural environments where uniformity was not a major selection criterion, modern cultivars tend to be highly uniform (Fernie et al. 2006). Low levels of genetic diversity in cultivars grown in a particular region increase the potential vulnerability to pests and abiotic stresses, which can cause major losses in the production of most or all cultivars of a crop (Graner et al. 1994, Jordan et al. 1998).

The challenges that face modern plant breeders are to develop higher yielding, nutritious and environmentally friendly varieties that improve our quality of life without harnessing additional natural habitats to agricultural production (Zamir 2001). Without a broad base of heterogeneous plant material, it is impossible for plant breeders to produce cultivars that meet the changing needs regarding adaptation to growing conditions, resistance to biotic and abiotic stresses, product yield or specific quality requirements (Friedt et al. 2007). Therefore, the most efficient way to farther improve the performance of crop varieties is to access to large diverse pool of genetic diversity.

1.2.1 Genetic diversity in *Brassica napus*

Like wheat (*Triticum aestivum*), oilseed rape originated as a result of interspecific hybridisation followed by polyploidisation. However, in comparison to *T. aestivum* and most other major crop species, *B. napus* is a comparatively young species that probably originated only a few centuries ago. It is thought that traders travelling between Asia, Europe and Africa transported *B. rapa* from eastern Europe and Asia to the Mediterranean region, and probably also brought *B. oleracea* eastwards, enabling for the first time the interspecific hybridisation that led to the origin of *B. napus*. The limited geographic range and intensive breeding of rapeseed has led to a comparatively narrow genetic basis in current breeding material. In

particular, the gene pool of elite oilseed rape breeding material has been eroded by an emphasis on specific oil and seed quality traits, with particularly strong bottleneck selection for zero seed erucic acid (C22:1) content and low seed glucosinolate content (so-called 00, double-low or canola seed quality). As a consequence, genetic variability in modern oilseed rape cultivars is severely restricted with regard to many characters of value for breeding purposes.

Owing to their generally unsuitable seed characters, in particular high contents of seed erucic acid, glucosinolates, and other anti-nutritive substances, fodder and vegetable rape forms have been generally overlooked for breeding of oilseed cultivars in recent decades. On the other hand, genetically diverse material is a potentially valuable source of improved pathogen and pest resistance, and introduction of previously unused germplasm into breeding lines also has the potential to considerably improve heterotic potential of hybrid varieties. Because of linkage drag for seed yield and quality traits associated with non-oilseed *B. napus* types, identification of genetically diverse germplasm amongst the respective gene pools of winter and spring oilseed forms is of particular interest.

Traditionally, morphological, phenological and agronomical traits have been employed as criteria for the introgression of new variation into oilseed rape breeding lines. In recent years, molecular genetic techniques to detect DNA polymorphisms have been increasingly used to characterise and identify novel germplasm for use in crop breeding (O'Neill et al. 2003). A number of previous studies have dealt with genetic diversity in *B. napus*, however most have investigated a limited range of genotypes. For example, Thormann et al. (1994) used restriction fragment length polymorphism (RFLP) and randomly amplified polymorphic DNA (RAPD) markers to determine genetic distances in and between cruciferous species. Halldén et al. (1994) compared *B. napus* breeding lines using RFLP and RAPD markers, while Diers and Osborn (1994) compared RFLP patterns in 61 winter and spring rapeseed genotypes and concluded that the two forms constitute two genetically different groups, and Lombard et al. (2000) also clearly discriminated between 83 spring and winter rapeseed cultivars using amplified fragment length polymorphism (AFLP) markers. Simple sequence repeat (SSR; microsatellite) markers were used by Plieske and Struss (2001) to differentiate 29 winter and 3 spring rapeseed varieties and breeding lines in a cluster analysis. In a recent study, Shengwu et al. (2003) found considerable genetic diversity between European and Chinese oilseed rape

using RAPD markers. Using SSR makers, Zhou et al. (2006) also clearly distinguished between 11 Chinese and 12 Swedish spring rapeseed genotypes.

Resynthesis of *B. napus* genotypes through interspecific hybridization between the diploid parents, assisted by embryo rescue, has repeatedly been shown to be useful for broadening the genetic basis of rapeseed. For example, Becker et al. (1995) compared the genetic diversity in rapeseed cultivars with resynthesised lines using allozyme and RFLP markers, and concluded that resynthesised rapeseed genotypes are a suitable resource for broadening the genetic base of the species. In another study, Seyis et al. (2003b) described genetic diversity in a large set of resynthesised rapeseed lines and spring rape varieties and found that the resynthesised *B. napus* showed an extremely high genetic divergence from the modern varieties. Hybrids produced from crosses between genetically diverse resynthesised rapeseed and adapted oilseed types showed a high yield potential (Seyis et al. 2006). The relationship between genetic distance and heterosis in oilseed rape was investigated by Diers et al. (1996) using RFLP markers, while a similar study was performed by Riaz et al. (2001) using sequence-related amplified polymorphic (SRAP) markers. The latter study demonstrated that crosses between genotypes from genetically divergent clusters tended to show higher levels of heterosis for seed yield and other traits.

1.3 Glucosinolates

Glucosinolates are secondary plant metabolites synthesized by species in the family Brassicaceae, which includes all of the *Brassica* crop species, related mustard crops and the model plant *Arabidopsis thaliana*. More than 120 different glucosinolate compounds have been identified in sixteen families of dicotyledonous angiosperms, including a large number of edible species (Fahey et al. 2001). The various glucosinolate compounds are designated aliphatic, aromatic and indole glucosinolates depending on whether they originate from aliphatic amino acids (methionine, alanine, valine, leucine, isoleucine), aromatic amino acids (tyrosine, phenylalanine) or tryptophan, respectively. Biosynthesis of glucosinolates proceeds in three phases: (i) side chain elongation of amino acids by incorporation of methylene groups, (ii) formation of the glucone moiety to produce primary glucosinolates, and (iii) secondary modifications of the side chain to generate the known spectrum of glucosinolate compounds (Grubb et al. 2004, Levy et al. 2005).

Together with the thioglucosidase enzymes (also known as myrosinases), glucosinolates form the glucosinolate-myrosinase system (Wittstock and Halkier 2002), which is generally believed to be part of the plant's defence system against insects and possibly also against pathogens (Rask et al. 2000). In intact tissues the thioglucosidase enzymes are stored separately from glucosinolates (Bones and Rossiter 1996), however when plant tissue is damaged the glucosinolates are hydrolysed by the enzymes to release a range of toxic defence compounds from substrate cells (Mithen et al. 2000), These toxins include nitriles, thiocyanates, isothiocyanates, oxozaladines and epithioalkanes (Kliebenstein et al. 2001, Wittstock and Halkier 2002). Besides genetic variation within and among different species, the pattern of hydrolysis products depends on numerous factors, and reaction conditions such as pH, availability of ferrous ions and presence of myrosinase-interacting proteins determine the final composition of the product mix (Mithen et al. 2000, Wittstock and Halkier 2002).

High levels of glucosinolates present in rapeseed meal have been found to reduce feed intake and growth rate, induce iodine deficiency, goitrogenicity, impair fertility, and furthermore can lead to liver, kidney and thyroid hypertrophy (Burel et al. 2000, Kermanshahi and Abbasi Pour 2006, McNeill et al. 2004, Mawson et al. 1994a, 1994b, Schöne et al. 1997). In spite of the above negative effects, certain degradation products, e.g. isothiocyanate, exhibit strong anticarcinogenic properties (Keck and Finley 2004). Negative effects of glucosinolates on animals are relative to their concentration in the diet (Maroufyan and Kermanshahi 2006), and deleterious effects also depend on the type and age of the animal. Ruminants are less sensitive to high glucosinolate intakes than non-ruminants, for example, and adult ruminants are more tolerant compared to young animals (Mawso et al. 1994, Derycke et al., 1999) because their hepatic pathways and rumen activity can detoxify glucosinolate breakdown products more efficiently (Mandiki et al. 2002). Tripathi and Mishra (2007) quote tolerance levels of total glucosinolate content in ruminants, pigs, rabbits, poultry and fish at 1.5-4.22, 0.78, 7.0, 5.4 and 3.6 $\mu mol.g^{-1}$ diet, respectively.

Various processing techniques can be applied to remove or reduce glucosinolate content in rapeseed meal in order to minimize glucosinolate-associated deleterious effects on animal health and production (recently reviewed by Tripathi and Mishra 2007). However, most of these methodologies include hydrolysis or decomposition of glucosinolate via heat treatment and the high energy costs that is needed mean that it is not economical to generate low-glucoinolate rapeseed meal from cultivars with high glucosinolate content. Production of

oilseed rape / canola meal is therefore limited to 00 varieties with low concentrations of total seed glucosinolates. In 1969 the Polish spring rape variety ‘Bronowski’ was identified as a low-glucosinolate form, and this cultivar provided the basis for an international backcrossing program to introduce this polygenic trait into high-yielding erucic acid-free breeding lines. A result was the release in 1974 of the first 00-quality spring rapeseed variety, ‘Tower’. Today the overwhelming majority of modern spring and winter oilseed rape varieties have 00-quality (“canola”). However, residual segments of the 'Bronowski' genotype in modern cultivars are believed to cause reductions in yield, winter hardiness, and oil content (Sharpe and Lydiate 2003), therefore finding new allelic sources for low-glucosinolate content will be beneficial.

1.3.1 Genetics of glucosinolate biosynthesis

Glucosinolate biosynthesis is a complex process that is influenced by interactions among a large number of genes and also by the environment, so that the glucosinolate content of any given tissue is quantitatively inherited. The low seed glucosinolate trait derived from the *B. napus* cultivar ‘Bronowski’ has been investigated in a number of studies via quantitative trait locus (QTL) analysis of total seed glucosinolate content in different oilseed rape crosses. For example, in a cross between the cultivar ‘Major’ (high seed glucosinolate content) and a doubled-haploid line derived from the low seed glucosinolate cultivar ‘Stellar’, Toroser et al. (1995) identified two major loci (*GSL-1* and *GSL-2*) with a large influence on total seed aliphatic-glucosinolates, and three further loci with minor effects (*GSL-3*, *GSL-4* and *GSL-5)*. In another study Uzunova et al. (1995) identified four QTL for seed glucosinolate content (designated *gsl-1*, *gsl-2*, *gsl-3* and *gsl-4*) in a cross between the old German winter rapeseed landrace ‘Mansholt's Hamburger Raps’ and the French winter rapeseed cv. ‘Samourai’. Four QTL for seed glucosinolate content were also localised by Howell et al. (2003) in a population derived from the cross ‘Victor’ × ‘Tapidor’. These QTL (*GLN1*, *GLN2*, *GLN3* and *GLN4*) mapped to *B. napus* chromosomes N9, N12, N19 and N17, respectively, and the first three of these loci were found to co-localise with seed glucosinolate QTL in another cross, between ‘Bienvenu’ and ‘Tapidor’. Furthermore, according to Howell et al. (2003), *GLN1*, *GLN2* and *GLN4* correspond to the QTL *GSL-1, GSL-2* and *GSL-4* from the study of Toroser et al. (1995), while *GLN1*, *GLN3*, and *GLN2*, respectively, correspond to *gsl-1*, *gsl-2* and *gsl-3* from the study of Uzunova et al. (1995). Other studies using different mapping parents also localised major QTL for seed glucosinolate content on chromosomes N9, N12 and N19

(Sharpe and Lydiate 2003, Zhao and Meng 2003, Basunanda et al. 2007). The QTL found on these three chromosomes in the different studies show a high degree of homology amongst the respective genomic regions, suggesting that the reduction in seed glucosinolate accumulation seen in cultivars derived from 'Bronowski' are controlled to a large extent by multiple homeologous copies of the same genes (Howell et al. 2003).

By screening QTL-associated markers in different cultivars, Howell et al. (2003) also demonstrated that low-glucosinolate genotypes can often carry high-glucosinolate alleles at one or more of the major quantitative trait loci (QTL) controlling seed glucosinolate accumulation. With effective molecular markers for marker-assisted selection these genotypes could be used to introduce new genetic variation for low seed glucosinolate content into breeding programs. The major genes that contribute to these major QTL for seed glucosinolate content in *B. napus* are not yet known, however the biosynthesis pathways for different glucosinolate compounds are well-characterised in the model crucifer *A. thaliana* and comparative mapping studies with close relatives can give insights into potential candidate genes for major loci in *Brassica* species. For example, the *GSL-ELONG* locus in *B. oleracea*, which is involved in methionine side-chain elongation during the biosynthesis of aliphatic glucosinolates, is thought to contain duplicated orthologs of the *A. thaliana* gene *METHYLTHIOALKYLMALATE SYNTHASE* (*MAM1/MAM-L*; Kroymann et al. 2001, Li and Quiros 2002). Identification of key genes involved in other major *B. napus* seed glucosinolate QTL could assist in identification of additional low-glucosinolate alleles for use in development of new low-glucosinolate gene pools for hybrid breeding.

1.4 Marker-trait association studies

Many agronomically important traits such as seed yield, product quality and some forms of disease resistance are controlled by interactions among many genes and are known as quantitative traits (Collard et al. 2005), Currently the most common method for mapping plant genes involves generation of a segregating population derived from a bi-parental cross (e.g. F_2-F_3 families, backcross progeny, recombinant inbred lines or doubled haploids), and estimation of recombination frequencies between molecular marker loci and genes of interest in the segregating offspring (Hansen et al. 2001, Peleman and van der Voort 2003). However, in a cross between two pure-breeding parental lines only two alleles segregate per locus, and

only phenotypes that segregate within the population can be analysed (Peleman and van der Voort 2003). Markers for QTL detected by classical genetic mapping in individual crosses are therefore not necessarily transferable to other material, and the utility of QTL-linked markers for marker-assisted selection is limited by the relative effects of individual loci on the trait of interest (Snowdon and Friedt 2004).

An alternative method for identification of molecular markers linked to quantitative traits of interest is marker-trait association analysis, also known as linkage disequilibrium (LD) mapping or association mapping, which is a population-based survey technique used to identify trait-marker relationships based on LD (Flint-Garcia et al. 2003). Association mapping incorporates all allelic and phenotypic variation in a given set of materials, and also considers all recombinations since the mutation at a polymorphic locus occurred; this can considerably increase the precision of the estimated QTL localisation (Meuwissen and Goddard 2000). Association genetics approaches rely on the decay of LD, initially present in a population, at a rate determined by the genetic distance between loci and the number of generations since it arose (Mackay and Powell 2007).

Detection of marker-trait associations based on linkage disequilibrium in genetically diverse materials can identify alleles with strong linkage to genes showing significant effects on the trait. In plant breeding populations the technique has seldom been used for marker development (Breseghello and Sorrells 2006), although association approaches can be particularly suitable for identification of useful allelic variation in genetically diverse genotype collections (Flint-Garcia et al. 2003). Distinguishing as many alleles as possible at loci of interest and determining phenotypic values for these alleles should greatly improve the predictive power of selection markers and enable marker-assisted combination of positive alleles for different loci (Peleman and van der Voort 2003). Association can be an effective approach for closing the gap between QTL analysis and marker-assisted selection (Breseghello and Sorrells 2006). A major advantage over classical QTL mapping studies is that no mapping populations must be created, which means that association tests can be performed relatively quickly and inexpensively when suitable markers, along with phenotype data from suitable sets of materials, are available (Francia et al. 2005).

One of the methodologies proposed for association mapping is a whole-genome scanning with a sufficient number of markers to detect regions associated with the phenotype of interest. An

alternative is to use a small number of genes already suspected of being involved in the trait of interest, and to investigate allelic variation within the gene sequences in a genetically diverse population (Rafalski 2002). Only polymorphisms with extremely tight linkage to a locus with phenotypic effects are likely to be significantly associated with the trait in a randomly mating population, providing much finer resolution than traditionally QTL localisation in single-cross genetic maps (Remington et al. 2001).

1.5 *Brassica*-Arabidopsis comparative genome analysis

The *Arabidopsis thaliana* genome-sequencing project was partly prompted by the prospect of transferring information on genome structure to closely related crop species within the Brassicaceae family (Ziolkowski et al. 2006). The *Brassica* species are the closest crop relatives to the model crucifer, and hence are major beneficiaries from the vast array of Arabidopsis molecular genetic and genomic tools and the increasingly good annotation to major *Brassica* crop genomes. Oilseed rape *(B. napus* ssp. *napus*) is the most intensively studied *Brassica* amphidiploid and the most economically important crop among the Brassicaceae family. Using extensive comparative genetic mapping, Parkin et al. (2005), identified 21 syntenic blocks that are shared by *Brassica napus* and *A. thaliana* genomes, corresponding to 90% of the mapped length of the *B. napus* genome. These conserved blocks represent co-linear regions that have been maintained since the divergence of the brassicas from the Arabidopsis lineage some 14-24 million years ago (Koch et al. 2000).

With increasingly detailed *Brassica*-Arabidopsis comparative genomics data it is becoming possible to navigate between and among the chromosomes of *A. thaliana* and *B. napus*. In some cases this can enable the map positions of *B. napus* QTL for traits of agronomic importance to be compared with the positions of potential candidate genes in the model genome. *Brassica* sequences with homology to the corresponding *A. thaliana* regions can then potentially be used for database-oriented identification of new markers for fine mapping, association studies or marker-assisted selection towards trait improvement. Moreover, it is also potentially possible to identify relevant candidate genes for important traits in oilseed rape, based on their positions in syntenic maps compared to important QTL.

1.6 Simple Sequence Repeat (SSR) markers

Simple sequence repeats (SSR) markers, also known as microsatellites, are short, tandemly-repeated sequence motifs, consisting of 1-6 nucleotide repeats, that are found in all prokaryotic and eukaryotic genomes (Zane et al. 2002). They can be found anywhere in the genome, in both coding and non-coding regions (Tóth et al. 2000, Varshney et al. 2005), and are usually characterized by a high degree of length polymorphism (Zane et al. 2002). Microsatellite loci have high mutation rates, ranging from 10^{-6} to 10^{-2} (Bachtrog 1999, Schlötterer 2000), whereby two alternative mutation mechanisms are proposed to explain their polymorphism: The first possibility is slippage of the DNA polymerase, while the second is unequal crossing-over caused by mispairing during DNA replication (Sia et al. 1997, Tóth et al. 2000, Li et al. 2002). Each of these mechanisms can potentially alter the microsatellite length by insertion or deletion of one or more repeat motifs (Wierdl et al. 1997, Kruglyak et al. 1998). The DNA sequences flanking motif repeats are generally conserved among individuals of the same species, allowing the selection of polymerase chain reaction (PCR) primers that will amplify the intervening microsatellite in all genotypes (Rongwen 1995). The resulting PCR products can be separated according to size by gel electrophoresis in a high-resolution agarose gel; however, because SSR polymorphisms can also result from the addition or deletion of only a single copy of the repeat motif, in the case of dimeric or trimeric repeat motives it is often advisable to separate amplification products by polyacrylamide or capillary gel electrophoresis (O'Neill et al. 2003).

In comparison with other molecular marker techniques SSR markers are useful for a variety of applications in plant genetics and breeding because they are numerous, highly polymorphic and informative, codominant, technically simple and reproducible. Furthermore they are relatively inexpensive when primer information is available, and SSR markers are thought to be frequent in gene-rich genome regions, increasing their potential relevance for allele-trait association studies in well-characterised genome regions containing quantitative trait loci.

The number of publicly available *Brassica* microsatellite primers is constantly increasing (see http://www.brassica.info/ssr/SSRinfo.htm), however in comparison with other important crop species relatively few SSR markers are freely available (Snowdon and Friedt 2004). Moreover, the traditional method for the development of SSR marker is time-consuming, labour-intensive and costly, involving the construction of a small-insert genomic DNA

library, the subsequent hybridization with tandemly repeated probes and the sequencing of hybridized positive clones (Thiel et al. 2003). With the increase in the availability of DNA sequence information, an automated process to identify and design PCR primers for amplification of SSR loci would be a useful tool in plant breeding programs (Robinson et al. 2004). Several computational tools are currently available for the identification of SSRs within sequence data as well as for the design of PCR primers suitable for the amplification of specific loci (Robinson et al. 2004). The identification of SSRs from genomic sequence data provides a potentially rich source of valuable molecular markers distributed across the genome. Large scale genomic sequencing also provides the opportunity to evaluate the abundance and relative distribution of SSR markers (Burgess et al. 2006) and to develop new SSR primers for use in plant breeding programs (Robinson et al. 2004). With the currently ongoing genomic sequencing of *B. rapa* (see http://www.brassica.info/) a large quantity of *Brassica* genomic sequence data is becoming available that can be screened in this way for new SSR markers. Since *Brassica* sequences have a high homology to Arabidopsis it is often possible to align genomic sequences, and any SSR markers they may contain, to the *A. thaliana* genome. This could provide a useful resource for development of new, potentially gene-linked SSR markers (Hasan et al. 2008)

1.7 Objectives

Knowledge of genetic diversity of a species has an important impact on the improvement of crop productivity as well as the conservation of genetic resources. In recent years more attention has been given to the genetic analysis of diverse genotype sets, which are particularly attractive for association analysis of qualitative traits such as disease resistance or special quality characteristics. Such genotype sets encompass a wide allelic and phenotypic diversity and association studies can potentially identify useful alleles for use in breeding.

The objectives of the present study were:

1. To investigate genetic diversity in the *Brassica napus* gene pool using SSR markers;

2. To investigate the potential use of *Brassica*-Arabidopsis comparative genomics data for marker and gene identification in oilseed rape based on sequence orthology to *A. thaliana*;

3. To identify potentially gene-linked markers for important seed glucosinolate loci via structure-based allele-trait association studies in genetically diverse *B. napus* genotypes;

4. To test the viability of association studies to identify key genes for quantitative traits in oilseed rape.

II ARTICLE ONE

Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers

Hasan M., Seyis F., Badani A.G., Pons-Kühnemann J., Friedt W., Lühs W. and Snowdon R.J.

Published in *Genetic Resources and Crop Evolution* (2006)
Volume 53, pages 793–802

Genetic Resources and Crop Evolution (2006) 53: 793–802
DOI 10.1007/s10722-004-5541-2

Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers

M. Hasan, F. Seyis, A.G. Badani, J. Pons-Kühnemann, W. Friedt, W. Lühs and R.J. Snowdon*
*Institute of Crop Science and Plant Breeding I, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany; *Author for correspondence (e-mail: rod.snowdon@agrar.uni-giessen.de; phone: +49-641-99-37421; fax: +49-641-99-37429)*

Received 28 June 2004; accepted in revised form 22 October 2004

Key words: Brassica napus, Genetic diversity, Microsatellites, Rapeseed, SSR markers

Abstract

Genetic diversity throughout the rapeseed (*Brassica napus* ssp. *napus*) primary gene pool was examined by obtaining detailed molecular genetic information at simple sequence repeat (SSR) loci for a broad range of winter and spring oilseed, fodder and leaf rape gene bank accessions. The plant material investigated was selected from a preliminary *B. napus* core collection developed from European gene bank material, and was intended to cover as broadly as possible the diversity present in the species, excluding swedes (*B. napus* ssp. *napobrassica* (L.) Hanelt). A set of 96 genotypes was characterised using publicly available mapped SSR markers spread over the *B. napus* genome. Allelic information from 30 SSR primer combinations amplifying 220 alleles at 51 polymorphic loci provided unique genetic fingerprints for all genotypes. UPGMA clustering enabled identification of four general groups with increasing genetic diversity as follows (1) spring oilseed and fodder; (2) winter oilseed; (3) winter fodder; (4) vegetable genotypes. The most extreme allelic variation was observed in a spring kale from the United Kingdom and a Japanese spring vegetable genotype, and two winter rape accessions from Korea and Japan, respectively. Unexpectedly the next most distinct genotypes were two old winter oilseed varieties from Germany and Ukraine, respectively. A number of other accessions were also found to be genetically distinct from the other material of the same type. The molecular genetic information gained enables the identification of untapped genetic variability for rapeseed breeding and is potentially interesting with respect to increasing heterosis in oilseed rape hybrids.

Introduction

Brassica napus L. is a relatively young species that originated in a limited geographic region through spontaneous hybridisations between turnip rape (*B. rapa* L. s.str.; AA, $2n = 20$) and cabbage (*B. oleracea* L. p.p.; CC, 2n = 18) genotypes (Kimber and McGregor 1995). Today oilseed rape (*B. napus* ssp. *napus*) is the most important source of vegetable oil in Europe and the second most important oilseed crop in the world after soybean. However, its limited geographic range and intensive breeding has led to a comparatively narrow genetic basis in current breeding material. The gene pool of elite oilseed rape breeding material has been further eroded by an emphasis on specific oil and seed

quality traits. As a consequence, genetic variability in this important crop is restricted with regard to many characters of value for breeding purposes.

Rapeseed cultivars are classified as winter or spring types according to their vernalisation requirement in order to induce flowering. Besides spring and winter oilseed types, rapeseed is often also grown as a fodder crop or as green manure. Swede cultivars are also relatively common, particularly in Great Britain and Scandinavia, and a small number of kale vegetable forms are also known. Owing to their generally unsuitable seed characters, in particular high contents of seed erucic acid, glucosinolates, and other anti-nutritive substances, fodder and vegetable rape forms have been generally overlooked for breeding of oilseed cultivars in recent decades. On the other hand, genetically diverse material is a potentially valuable source for improved pathogen and pest resistance, and introduction of untapped germplasm into breeding lines also has the potential to improve heterotic potential. Because of linkage drag for seed yield and quality traits associated with non-oilseed rape morphotypes, identification of genetically diverse germplasm amongst the respective gene pools of winter and spring oilseed forms is of particular interest.

Traditionally, morphological, phenological and agronomical traits have been employed as criteria for the introgression of new variation into oilseed rape breeding lines. In recent years, molecular genetic techniques using DNA polymorphism have been increasingly used to characterise and identify novel germplasm for use in crop breeding (see O'Neill et al. 2003 for a review). A number of previous studies have dealt with genetic diversity in *B. napus*, however most have investigated a limited range of genotypes. Thormann et al. (1994) used restriction fragment length polymorphism (RFLP) and RAPD markers to determine genetic distances in and between cruciferous species, Halldén et al. (1994) compared *B. napus* breeding lines with RFLPs and RAPDs, while Diers and Osborn (1994) compared RFLP patterns in 61 winter and spring rapeseed genotypes and concluded that the two forms constitute two genetically different groups. Becker et al. (1995) compared the genetic diversity in rapeseed cultivars with resynthesised lines using allozyme and RFLP markers. Seyis et al. (2003) described genetic diversity in a large set of resynthesised rapeseed lines and spring rape varieties. Diers et al. (1996) investigated the relationship between genetic distance and heterosis in oilseed rape using RFLP markers, and Riaz et al. (2001) performed a similar study with sequence-related amplified polymorphic (SRAP) markers. Simple sequence repeat (SSR; microsatellite) markers were used by Plieske and Struss (2001) to differentiate 29 winter and 3 spring rapeseed varieties and breeding lines in a cluster analysis. In comparison with other molecular marker techniques SSR markers are numerous, highly polymorphic and informative, codominant, technically simple and reproducible, and they are relatively inexpensive when primer information is available. Furthermore, SSR markers often occur in gene-rich genome regions, increasing their potential relevance for allele–trait association studies in well-characterised genome regions containing quantitative trait loci.

In this study SSR genotype information was compiled for loci spread throughout the *B. napus* genome in a core set of gene bank accessions representing a broad range of the gene pool, including spring and winter fodder, oilseed and vegetable types from diverse origins. Ninety-six morphologically diverse genotypes were chosen from a *B. napus* core set described by a European project on *Brassica* genetic resources (RESGEN; see Lühs et al. 2003a). Swede accessions were not included in the analysis. The genotypes were genetically characterised using 30 informative mapped SSR primer combinations from the public domain. The genetic diversity of the material is discussed with respect to its potential to expand the narrow gene pool for oilseed rape breeding.

Materials and methods

Plant material

In a previous study (see Lühs et al. 2003a) around 1500 oilseed, fodder and vegetable rape accessions, representing the *B. napus* material present in European gene banks, were evaluated in field trials to investigate their variation for descriptive characters like emergence date, vernalisation requirement and winter hardiness, flowering time and duration, fatty acid composition and seed glucosinolate content. Based on the data collected and available pedigree knowledge, a preliminary core

collection of around 180 genotypes was established for more detailed agronomic investigations to identify germplasm of interest for oilseed rape breeding. The core collection is intended to represent as broadly as possible the genetic diversity present in the species as a whole (Poulsen et al. 2004).

In order to quantify the genetic variation present in the preliminary core collection 96 accessions were selected that covered as broadly as possible the observed morphological variation, with the exception of swede (*B. napus* ssp. *napobrassica* (L.) Hanelt) accessions for which no material was available. The 96 genotypes were grown in the field to confirm homogeneity of the accessions. Total cellular DNA was extracted from young true leaves of a representative plant from each accession using a standard CTAB extraction protocol (Doyle and Doyle 1990). The accessions investigated and their origins are listed in Table 1.

SSR analyses

Sixty *Brassica* SSR primer combinations selected from the collection available in the public domain (Lowe et al. 2002, 2004; see www.brassica.info/ssr/SSRinfo.htm) were tested for their suitability. After pre-screening 30 primer pairs were chosen that gave clear, reproducible and polymorphic amplification products at one or more loci in *B. napus*.

PCR reactions were carried out in a volume of 15 μL containing 20 ng of DNA template, 0.75 pmol of each primer, 0.2 mM dNTP mix, 1 mM $MgCl_2$, 1 × PCR reaction buffer (with 15 mM $MgCl_2$, Qiagen) and 0.25 unit of *Taq* DNA polymerase (Qiagen). Amplifications were performed using a standard amplification cycle in a GeneAmp PCR System 9700 thermal cycler, and SSR polymorphisms were separated and visualised using a LI-COR GeneReadir 4200 (MWG Biotech, Ebersberg). To reduce primer labelling costs PCR products were labelled with the M13-tailing technique described by Berg and Olaisen (1994). In this method the fluorescently labelled universal M13 primer5′-AGGGTTTTCCCAGTCACGACGTT-3′ is added to the PCR reaction, and the forward primer of each SSR is appended with the sequence 5′-TTTCCCAGTCACGACGTT-3′. After the first round of amplification the PCR fragments are subsequently amplified by the labelled universal primer. For further savings on polymerase costs and workload, duplex or triplex PCR assays were developed in cases where primers with the same annealing temperature gave amplification products of sufficiently different size. PCR primers along with the annealing temperatures used are shown in Table 2.

Data analysis

Up to five polymorphic loci were amplified for each SSR primer combination, reflecting the complex polymorphic nature of the *B. napus* genome, and as many as 11 alleles per locus were amplified. This resulted at times in complex banding patterns that made the assignment of alleles to loci difficult and in some cases impossible. Furthermore, homologous *Brassica* SSR loci can display identical alleles that cannot be resolved quantitatively. This prohibited the use of allele-scoring techniques designed for simple diploid organisms that assign SSR amplification products to individual loci and calculate genetic distance matrices based on allele frequencies. Instead, for each primer combination, the absence/presence of each polymorphic amplification product in each plant was determined and data was recorded in a 0/1 matrix, as described for example by Alamerew et al. (2004) for analysis of polyploid wheat. Alleles were scored digitally using the software package RFLP-SCAN (Version 2.01, Scanalytics Inc., Fairfax, VA, USA) through matching with a size-calling standard, however all genotypes were confirmed manually. The genotype data was converted to a similarity matrix using the Dice similarity index Dice 1945), described as follows by Nei and Li (1979):

$$S = 2N_{ab}(2N_{ab} + N + N),$$

where N_{ab} is the number of bands shared by genotypes a and b in each pairwise comparison, and N_a and N_b are the numbers of bands present in the respective genotypes. Dice genetic similarities were calculated using the software WINDIST (I.W. Yap, Univ. of Washington, USA), and the SAHN module of NTSYSpc (Version 2.01, Exeter Software, Setauket, NY, USA) was used to generate a dendrogram based on the UPGMA algo-

Table 1. The 96 *Brassica napus* ssp. *napus* accessions for the study were selected from a preliminary core collection (Lühs et al. 2003a) representing the genetic variability present in *Brassica napus* L. collections in European genebanks.

Accession name	Type	Genebank[a]	Accession number	Country of origin
Erglu	SF	BS	16237	Germany
Kroko	SF	BS	32545	Germany
Liho	SF	IPK	CR704	Germany
Petranova	SF	BS	16223	Germany
Tantal	SF	CGN	CGN06893	France
Tira	SF	BS	30285	Germany
Alku	SOSR	NGB	2336	Sweden
Bronowski	SOSR	IPK	CR270	Poland
Ceska Krajova	SOSR	IPK	CR 280	Czech Republic
Duplo	SOSR	IPK	CR303	Germany
Dux	SOSR	BS	53721	Germany
Janetzkis Sommerraps	SOSR	BS	31208	Poland
Line	SOSR	NGB	1794	Sweden
Marnoo	SOSR	BS	34719	Australia
Nugget	SOSR	IPK	CR807	Canada
Olga	SOSR	NGB	8678	Sweden
Spaeths Zollerngold	SOSR	BS	16225	Germany
Svalöf's Gulle	SOSR	NGB	585	Sweden
Topas	SOSR	NGB	2695	Sweden
Tribute	SOSR	IPK	2005	Canada
Wesway	SOSR	IPK	CR1055	Australia
Hakuran	SV	CGN	CGN06897	Japan
Russian Kale	SV	HRI	6234	Great Britain
Akela	WF	BS	7310	Germany
Aphid Resistant Rape	WF	IPK	CR 167/92	New Zealand
Binera	WF	IPK	CR 182	Germany
Bladkool	WF	IPK	CR 192	Netherlands
Clubroot Resistant	WF	IPK	CR283/98	New Zealand
Dwarf Essex	WF	IPK	CR304/95	Great Britain
English Giant	WF	HRI	3258	Great Britain
Liragrün	WF	IPK	CR 718	Germany
Michinaku natane	WF	IPK	CR 774	Japan
Nunsdale	WF	HRI	5133	Great Britain
Palu	WF	IPK	CR 834	Italy
Parapluie	WF	IPK	CR837/75	France
Samo	WF	NGB	2767	Sweden
Silona	WF	NGB	2769	Sweden
Anja	WOSR	IPK	CR 164	Germany
Askari	WOSR	IPK	CR 172/92	Germany
Baltia	WOSR	CGN	6868	Poland
Bienvenue	WOSR	IPK	CR181/86a	France
Bolko	WOSR	IPK	CR853/2000	Poland
Brink	WOSR	IPK	CR 267/82a	Sweden
Ceres	WOSR	IPK	CR279/88a	Germany
Coriander	WOSR	IPK	CR 286	Germany
Darmor	WOSR	HRI	9139	France
Diamant	WOSR	BS	7299	Germany
Dippes	WOSR	IPK	CR298/92	Germany
Doral	WOSR	BS	29619	Germany
Doral	WOSR	IPK	CR 301/96	Germany
Edita	WOSR	IPK	CR 307	Germany
Erra	WOSR	IPK	CR 318	Germany
Groß-Lüsewitzer	WOSR	BS	34435	Germany
Gülzower Ölquell	WOSR	IPK	CR 1191/81	Germany
Hokkai 3-Go	WOSR	IPK	CR 646/92	Japan

Table 1. Continued.

Janpol	WOSR	IPK	CR659/97	Poland
Jantar	WOSR	IPK	CR743/90	Poland
Jet Neuf	WOSR	CGN	CGN07227	France
Jupiter	WOSR	NGB	2703	Sweden
Krapphauser	WOSR	IPK	CR 2187	Germany
Kromerska	WOSR	CGN	CGN06869	Czech Republic
Lenora	WOSR	BS	7295	Germany
Lesira	WOSR	IPK	CR689/86a	Germany
Librador	WOSR	BS	28959	Germany
Libritta	WOSR	BS	61797	Germany
Liglory	WOSR	IPK	CR561/96	Germany
Liporta	WOSR	BS	65330	Germany
Lirafit	WOSR	IPK	CR715/89	Germany
Lirakotta	WOSR	BS	61794	Germany
Madora	WOSR	IPK	CR 740	Germany
Major	WOSR	BS	28659	France
Markus	WOSR	BS	28661	France
Matador	WOSR	NGB	594	Sweden
Mestnij	WOSR	CGN	CGN06871	Soviet Union
Moldavia	WOSR	IPK	CR299/73	Moldova
Mytnickij	WOSR	CGN	CGN06880	Ukraine
Nemertschanskij 1	WOSR	IPK	CR787/92	Soviet Union
Norde	WOSR	NGB	593	Sweden
Panter	WOSR	NGB	592	Sweden
Quedlinburger Platzfester	WOSR	IPK	CR 862	Germany
Quinta	WOSR	BS	7302	Germany
Ramses	WOSR	IPK	CR 872	France
Rapol	WOSR	BS	7306	Germany
Sarepta	WOSR	IPK	CR 904	France
Skrzeszowicki	WOSR	IPK	CR 925	Poland
Skziverskij	WOSR	CGN	CGN06885	Soviet Union
Slovenska Krajova	WOSR	IPK	CR 948	Czech Republic
Sobotkowski	WOSR	IPK	CR 950	Poland
Sonnengold	WOSR	IPK	CR 954	Germany
Start	WOSR	IPK	CR972/96	Poland
Trebicska	WOSR	IPK	CR1026	Czech Republic
Victor	WOSR	IPK	CR 1038/98	Sweden
Vinnickij 15/59	WOSR	CGN	CGN06881	Soviet Union
Wolynski	WOSR	IPK	CR 1060/95	Soviet Union
Chuosenshu	WV	IPK	CR 2198	Korea
Taisetsu	WV	IPK	CR 1008/92	Japan

The material included a broad range of winter-type (W) and spring-type (S) fodder (F), vegetable (V) and oilseed rape (OSR) forms from diverse origins.

[a]BS, Braunschweig, Germany; IPK, Gatersleben, Germany; HRI, Wellesbourne, United Kingdom; CGN, Wageningen, Netherlands; NGB, Nordic Gene Bank, Alnarp, Sweden.

rithm (unweighted pair group method with arithmetric average). Because UPGMA clustering sometimes results in discrepancies depending on the choice of similarity index and can be biased by rare alleles, genetic relationships among the genotypes were also represented using a principal coordinate (PCO; Backhaus et al. 1989) analysis, using the software SPSS (version 10.0.7, SPSS Inc., Chicago, USA), in order to confirm the positions of outlier genotypes.

Results

The 30 SSR primer combinations used amplified a total of 51 scoreable polymorphic loci with 220 alleles in the 96 genotypes. The polymorphic loci gave unique genetic fingerprints for all 96 accessions. Figure 1 shows the UPGMA dendrogram representing genetic similarity among the accessions. As expected two major clusters were formed representing winter and spring accessions. The

Table 2. Allelic diversity at SSR loci amplified by primer used for the genetic diversity analysis.

SSR-Primer	Annealing temperature (°C)	Polymorphic loci detected	No. of alleles detected	Linkage group/s (Lowe et al. 2002, 2004)
Na10-C01	55	5	25	N13, N14
Na10-D03	55	1	2	N3
Na10-E02	55	3	11	N5, N6, N8, N13
Na10-F06	55	1	3	N3, N14
Na12-D10	55	1	4	N15
Na12-E06A	55	1	8	N9, N15
Na12-F12	55	1	3	N13
Na14-D07	55	1	2	N1
Na14-E11	55	4	14	N14
Na14-G02	55	1	8	N3
Ol09-A06	55	2	6	N12
Ol10-A05	58	2	7	N2
Ol10-B01	55	1	7	N17
Ol11-B03	55	2	8	N10, N19
Ol11-B05	55	2	7	N3
Ol11-G11	55	2	9	N13
Ol11-H02	55	1	6	N4, N14
Ol11-H05	50	2	10	N15, N19
Ol12-A04	55	1	4	N19
Ol12-D05	55	1	7	N18
Ol12-E03	55	1	4	N7
Ol12-F11	55	2	6	N1, N10
Ol13-D02A	55	1	7	N3, N14
Ra1-F06	60	3	9	N6
Ra2-A04	50	1	5	N5, N15
Ra2-A05	55	1	4	N7
Ra2-E11	55	1	11	N13
Ra2-E12	55	1	5	N8
Ra2-F11	50	3	10	N12, N13, N15, N19
Ra3-E05	55	2	8	N11
	Total	51	220	

Primer sequences were obtained from the collection available in the public domain (http://ukcrop.net/perl/ace/search/BrassicaDB), which were developed by Lowe et al. (2002, 2004). The 18 bp M13-tail sequence 5′-TTTCCCAGTCACGACGTT-3′ was added to 5′ end of each forward primer for detection of PCR products with a labelled universal M13 primer (see Methods). Because of this the annealing temperatures differ in some cases from those recommended by Lowe et al. (2002, 2004). S, spring; W, winter; F, fodder; V, vegetable; OSR, oilseed rape.

winter cluster differentiated clearly into two genetically distinct sub-clusters. The first sub-cluster contained mainly winter oilseed rape, however, four winter fodder varieties ('Binera', 'Silona', 'Bladkool' and 'Liragrün') also grouped with the oilseed accessions. The second winter sub-cluster was formed from more genetically diverse fodder rape accessions, whereby the spring fodder varieties 'Tantal' and 'Tira' also grouped in this sub-cluster. Unexpectedly, a number of oilseed genotypes did not group with the rest of the spring and summer forms: The German winter oilseed variety 'Liglory', the old spring oilseed cultivar 'Wesway', from Australia, and the German spring oilseed rape 'Dux' grouped separately between the winter and spring clusters, together with the German spring fodder rape 'Petranova'. The winter oilseed cultivars 'Mestnij' from Mongolia and 'Kromerska' from the Czech Republic were more closely related to the spring material than the other winter accessions. The old winter oilseed varieties 'Mytnickij' from Ukraine and 'Krapphauser' from Germany were highly genetically distinct from all other oilseed genotypes.

The vegetable varieties 'Hakuran' from Japan and 'Russian Kale' from Great Britain were the most genetically diverse genotypes, with a genetic similarity coefficient of only around 0.38, and the Asian accessions 'Chuosensu' from Korea and 'Taisetsu' from Japan were also extremely distinct

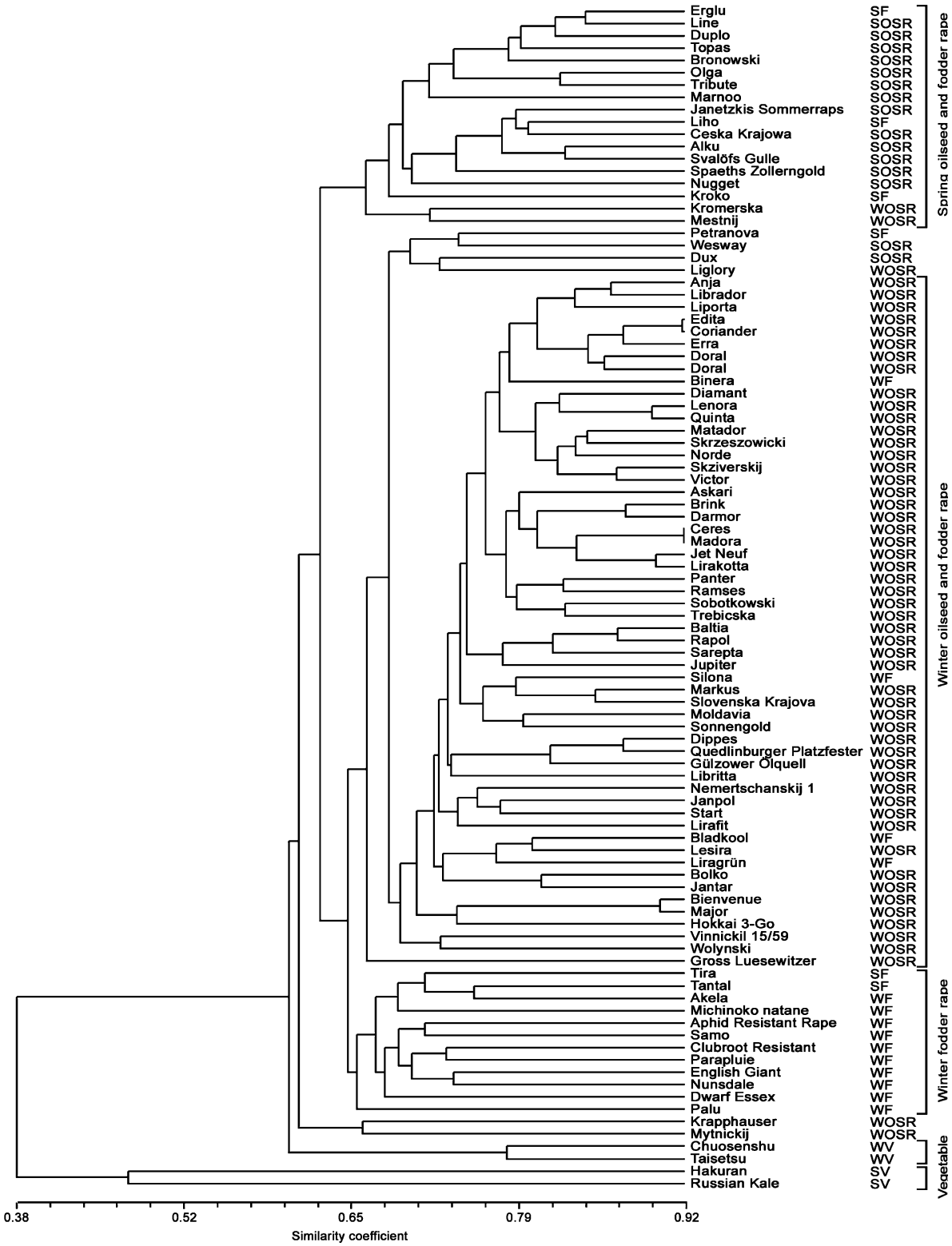

Figure 1. Phenogram showing DICE genetic similarity for a diverse set of 96 oilseed rape accessions including winter-type (W) and spring-type (S) fodder (F), vegetable (V) and oilseed rape (OSR) *Brassica napus* L. accessions from European genebanks (see Table 1), revealed by UPGMA clustering based on genetic fingerprints from 51 polymorphic SSR loci.

from all oilseed and fodder genotypes. Genetically distinct within the major clusters were the old German winter oilseed rape cultivar 'Groß Lüsewitzer' and the Italian winter fodder type 'Palu', which each showed a similarity index of around 0.65 to their respective cluster.

At the other extreme, the narrow clustering of specific genotype pairs with very similar SSR fingerprints (e.g. 'Lenora' and 'Quinta', 'Bienvenue' and 'Major', 'Brink' and 'Darmor', 'Coriander' and 'Edita') reflects the close pedigree relationships among these varieties. The German winter oilseed varieties 'Ceres' and 'Madora' showed identical SSR fingerprints for all but 5 of the 60 loci, suggesting a very close pedigree relationship. The set of investigated varieties contained one repetition of two different accessions of the variety 'Doral'. These two accessions clustered closely together however, were not identical, as might be expected after independent regeneration in different gene banks.

The PCO analysis reflected the UPGMA clustering (results not shown), confirming the positions of the unexpected outliers described above. The winter outliers 'Liglory', 'Mestnij' and 'Kromerska', and the spring outliers 'Dux', 'Petranova' and 'Wesway' were all located at intermediate positions between the respective spring and winter groupings. The spring outliers 'Tira' and 'Tantal' were confirmed to cluster with the winter fodder rape and the extreme divergence of the winter oilseed varieties 'Mestnij' and particularly 'Krapphauser' was also confirmed.

Overall allelic diversity within the respective groups of accessions was estimated based on the number of alleles and of unique alleles per group, respectively (Table 3). As expected the four vegetable accessions exhibited an extremely high allelic diversity, with more than 20% of the alleles in this group being absent from the other groups of accessions. The grouping of spring fodder and oilseed rape showed the lowest number of unique alleles.

Discussion

The results of this study demonstrate the suitability of SSR data for analysis of genetic diversity in *B. napus* genotypes. Unique genetic fingerprints were obtained for all genotypes, and the genotype data enabled differentiation between winter rapeseed, spring rapeseed and vegetable accessions. Furthermore, a group of winter fodder accessions was found to be genetically distinct from the winter oilseed rape. Based on the genetic data it appeared that the accessions 'Tantal' and 'Tira' are in fact winter forms (*B. napus* forma *biennis*), and this was confirmed by winter and summer field trials studying the vernalisation requirement and seasonality of the accessions (Lühs et al. 2003a). Presumably these varieties are listed as spring varieties because their use exclusively for fodder purposes means they do not flower before harvesting.

It is possible that unexpected genetic fingerprints can arise from mislabelling of samples or unintended substitution of seed among the original gene bank accessions. During the original field evaluation of the full set of European *B. napus* gene bank material a number of duplicated accessions from different gene banks were found to have markedly different morphological characteristics (Lühs et al. 2003a), obviously a result of incorrect labelling. The accession of the winter oilseed variety 'Liglory' we tested in this study was found to be a yellow-seeded type, although the original 'Liglory' has brown-black seeds. This appears to explain the unexpected genetic clustering of this genotype. Such information can be conveyed to the gene bank in question to enable them to update their collection with seed from the original variety where possible.

Although in many cases well-known rapeseed varieties can be identified based on existing phenological knowledge, when two modern varieties exhibit relatively similar characteristics an accidental substitution of seed cannot be ruled out despite the best controls of the responsible gene banks. We tested two accessions of the variety 'Doral' from different gene banks and confirmed that these were indeed closely related. The fact that their genetic fingerprints were not genetically identical reflects the fact that *B. napus* is a facultative outcrossing species, and regeneration in gene banks can be expected to introduce minor variation in accessions that generally does not however, diminish the overall value of the collections.

A number of both spring and winter oilseed genotypes could be identified which showed a high genetic divergence from the remaining elite spring or winter oilseed rape breeding material, respec-

tively. During the respective winter and spring field evaluations (Lühs et al. 2003a) the accessions in question all showed the expected vegetative or non-vegetative morphotype. One barrier preventing the direct use of diverse germplasm to improve the heterotic potential of oilseed rape is the generally poor adaptability or performance of exotic genotypes to the predominant climatic conditions in the major rapeseed-growing regions. Furthermore, serious linkage drag for seed yield and quality characters reduce the suitability of exotic vegetable and fodder rape germplasm for integration in oilseed rape breeding programs. Genetically divergent genotypes that belong to the appropriate spring or winter gene pool can potentially be much more readily utilised for oilseed breeding. The diverse spring and winter oilseed genotypes identified in this study may therefore represent a useful resource for improving heterotic potential in spring and winter oilseed rape, respectively. Butruille et al. (1999) described significant yield increases in spring oilseed rape hybrids through introgression of winter germplasm. However, this also requires backcrossing to re-establish the desired seasonality. In this regard the highly genetically distinct winter oilseed varieties 'Mytnickij', 'Kromerska', 'Mestnij' and 'Krapphauser' represent a potentially valuable resource for winter oilseed rape breeding, whereas 'Dux' (Germany) and 'Wesway' (Australia) could be of interest for diversifying the spring oilseed rape gene pool.

The exotic vegetable accessions we tested were found to be extremely genetically diverse in comparison with the other material. 'Hakuran' is a heading vegetable derived from a resynthesised rapeseed developed by interspecific crosses between Chinese cabbage (*B. rapa* ssp. *pekinensis* (Lour.) Hanelt in J. Schultze-Motel) and white cabbage (*B. oleracea* ssp. *oleracea* convar. *capitata* (L.) Alef) (Nishi 1980). The extreme genetic diversity of resynthesised rapeseed genotypes in comparison to conventional oilseed rape has been described previously by Becker et al. (1995) and Seyis et al. (2003). 'Russian Kale' belongs to leaf rape representing a distinct form (*B. napus* L. var. *pabularia* (DC.) Rchb. in Mössler, Handb. Gewächskde) of ssp. *napus* with a very leafy shoot (Diederichsen 2001). Formerly leaf rape (also commonly known as Siberian kale or Hanover salad; German *Schnittkohl*; French *chou à faucher*, Chinese *xi yang you cai*) was used more often as a winter-annual vegetable. From phenological studies it is known that swede accessions (*B. napus* ssp. *napobrassica* (L.) Hanelt), which were not included in the present study, are also extremely diverse. As mentioned above, however, seed yield and quality aspects seriously hinder the potential use of such exotic material for improvement of heterosis in oilseed rape hybrids. On the other hand, the high divergence among the vegetable and fodder rape genotypes was reflected in field evaluations that identified a considerable degree of variation for other relevant traits like disease and insect resistance (Lühs et al. 2003b). In this respect it would also be interesting to analyse the genetic variability of the *B. napus* ssp. *napobrassica* (Swede) accessions from the core collection with the same SSR markers, in order to compare their genetic diversity in comparison to *B. napus* ssp. *napus* and to evaluate the degree of genetic diversity found within this group.

Table 3. Number of alleles and unique alleles, respectively, detected within accessions representing the respective spring fodder (SF) and spring oilseed rape (SOSR), winter fodder (WF), winter oilseed rape (WOSR), and spring and winter vegetable rape (SV, WV) gene pools.

Classification	Number of accessions	Number of alleles	Unique alleles
SF/SOSR	21	126	7
WF	18	135	9
WOSR	53	154	14
WV/SV	4	119	27
Total	96	220	57

Assignment of alleles to the mapped loci described by Lowe et al. (2002, 2004) was not always possible because of the high allelic diversity and variation in allele sizes in comparison to the published data. For this reason it was also not feasible to calculate polymorphic information content (PIC) values for the individual loci.

For a number of primer combinations more polymorphic loci were detected than were mapped by Lowe et al. however, this is not unexpected when comparing a large set of genotypes with selected cross parents. On the other hand, for a few SSRs not all mapped loci could be reliably identified in the present study. This could be due to differences in PCR conditions leading to more stringent amplification, or possibly an altered specificity of the M13-

tailed primers used in the present study. We are using the primer set used here for genetic mapping studies in different oilseed rape populations and hope thus to increase the amount of available information regarding the polymorphic loci and their map positions. Detailed allelic information on mapped SSR loci, in combination with candidate gene association studies based on linkage disequilibrium in trait-relevant genome regions, has the potential to identify genes that contribute to such agronomically important traits in exotic *B. napus* germplasm, and ultimately to assist in the development of molecular markers for marker-assisted transfer of these traits into elite breeding material.

Acknowledgements

The authors thank contributors to the BBSRC database for their publicly available SSR primer information. Daniela Bernhardt and Liane Renno provided valuable technical assistance.

References

Alamerew S., Chebotar S., Huang X., Röder M. and Börner A. 2004. Genetic diversity in Ethiopean hexaploid and tetraploid wheat germplasm assessed by microsatellite markers. Genet. Resour. Crop. Evol. 51: 559–567.

Backhaus K., Erichson B., Plinke W. and Weiber R. 1989. Multivariate Analysemethoden. Springer-Verlag, Heidelberg.

Becker H.C., Engqvist G.M.and Karlssom B. 1995. Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. Theor. Appl. Genet. 91: 62–67.

Berg E.S. and Olaisen B. 1994. Hybrid PCR sequencing – sequencing of PCR products using a universal primer. Biotechniques 17: 896–901.

Butruille D.V., Guries R.P. and Osborn T.C. 1999. Increasing yield of spring oilseed rape hybrids through introgression of winter germplasm. Crop Sci. 39: 1491–1496.

Dice L.R. 1945. Measures of the amount of ecologic association between species. Ecology 26: 297–302.

Diederichsen A. 2001. *Brassica*. In: Hanelt P. and Institute of plant genetics and crop plant research (eds.), Mansfeld's Encyclopedia of Agricultural and Horticultural Crops 3. Springer, Heidelberg, pp. 1435–1465.

Diers B.W. and Osborn T.C. 1994. Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphisms. Theor. Appl. Genet. 88: 662–668.

Diers B.W., McVetty P.B.E. and Osborn T.C. 1996. Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.). Crop Sci. 36: 79–83.

Doyle J.J. and Doyle J.L. 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Halldén C., Nilsson N.O., Rading I.M. and Säll T. 1994. Evaluation of RFLP and RAPD markers in comparison of *Brassica napus* breeding lines. Theor. Appl. Genet. 88: 123–128.

Kimber D.S. and McGregor D.I. 1995. The species and their origin, cultivation and world production. In: Kimber D. and McGregor D.I. (eds.), Brassica Oilseeds: Production and Utilization. CABI Publishing, Wallingford, UK, pp. 1–9.

Lowe A.J., Jones A.E., Raybould A.F., Trick M., Moule C.L. and Edwards K.J. 2002. Transferability and genome specificity of a new set of microsatellite primers among *Brassica* species of the U triangle. Mol. Ecol. Notes 2: 7–11.

Lowe A.J., Moule C.L, Trick M. and Edwards K.J. 2004. Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theor. Appl. Genet. 108: 1103–1112.

Lühs W., Seyis F., Frauen M., Busch H., Frese L., Willner E., Friedt W., Gustafsson M. and Poulsen G. 2003a. Development and evaluation of a *Brassica napus* core collection. In: Knüpffer H. and Ochsmann J. (eds.), Rudolf Mansfeld and Plant Genetic Resources. Proceedings of a symposium dedicated to the 100th birthday of Rudolf Mansfeld, Gatersleben, Germany, 8–9 October 2001. Schriften zu Genetischen Ressourcen 19. ZADI/IBV, Bonn, pp. 284–289. (http://www.genres.de/infos/rei-bd22.htm).

Lühs W., Seyis F., Frauen M., Ulber B., Stemann G., Busch H., Friedt W., Gustafsson M. and Poulsen G. 2003b. Evaluation of a *Brassica napus* core collection and prospects of its exploitation through rapeseed breeding. 11th International Rapeseed Congress 6–10 July 2003, Vol. 2. Copenhagen, Denmark, pp. 371–374.

Nei M. and Li W. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl Acad. Sci. USA 76: 5269–5273.

Nishi S. 1980. Differentiation of *Brassica* crops in Asia and the breeding of 'Hakuran', a newly synthesized leafy vegetable. In: Tsunoda S., Hinata K. and Gómez-Campo C. (eds.), *Brassica* Crops and Wild Allies, Biology and Breeding. Japan Scientific Societies Press, Tokyo, pp. 133–150.

O'Neill R., Snowdon R.J. and Köhler W. 2003. Population genetics aspects of biodiversity. Prog. Bot. 64: 115–137.

Plieske J. and Struss D. 2001. Microsatellite markers for genome analysis in *Brassica*. I. Development and abundance in *Brassica* species. Theor. Appl. Genet. 102: 689–694.

Poulsen G., Busch H., Frauen M., Frese L., Friedt W., Gustafsson M., Ottosson F., Seyis F., Stemann G., Ulber B., Willner E. and Lühs W. 2005. The European *Brassica napus* core collection – Characterisation, evaluation and establishment. Cruciferae Newslett. (in press).

Riaz A., Li G., Quresh Z., Swati M.S. and Quiros C.F. 2001. Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breed. 120: 411–415.

Seyis F., Snowdon R.J., Lühs W. and Friedt W. 2003. Molecular characterization of novel resynthesized rapeseed (*Brassica napus*) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars. Plant Breed. 122: 473–478.

Thormann C.E., Ferreira M.E., Camargo L.E.A., Tivang J.G. and Osborn T.C. 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88: 973–980.

III ARTICLE TWO

Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*)

Hasan M., Friedt W., Pons-Kühnemann J., Freitag N.M., Link K. and Snowdon R.J.

Published in *Theoretical and Applied Genetics* (2008)
Volume 116, pages 1035–1049

Theor Appl Genet (2008) 116:1035–1049
DOI 10.1007/s00122-008-0733-3

ORIGINAL PAPER

Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*)

M. Hasan · W. Friedt · J. Pons-Kühnemann · N. M. Freitag · K. Link · R. J. Snowdon

Received: 16 June 2007 / Accepted: 9 February 2008 / Published online: 6 March 2008

Abstract Breeding of oilseed rape (*Brassica napus* ssp. *napus*) has evoked a strong bottleneck selection towards double-low (00) seed quality with zero erucic acid and low seed glucosinolate content. The resulting reduction of genetic variability in elite 00-quality oilseed rape is particularly relevant with regard to the development of genetically diverse heterotic pools for hybrid breeding. In contrast, *B. napus* genotypes containing high levels of erucic acid and seed glucosinolates (++ quality) represent a comparatively genetically divergent source of germplasm. Seed glucosinolate content is a complex quantitative trait, however, meaning that the introgression of novel germplasm from this gene pool requires recurrent backcrossing to avoid linkage drag for high glucosinolate content. Molecular markers for key low-glucosinolate alleles could potentially improve the selection process. The aim of this study was to identify potentially gene-linked markers for important seed glucosinolate loci via structure-based allele-trait association studies in genetically diverse *B. napus* genotypes. The analyses included a set of new simple-sequence repeat (SSR) markers whose orthologs in *Arabidopsis thaliana* are physically closely linked to promising candidate genes for glucosinolate biosynthesis. We found evidence that four genes involved in the biosynthesis of indole, aliphatic and aromatic glucosinolates might be associated with known quantitative trait loci for total seed glucosinolate content in *B. napus*. Markers linked to homoeologous loci of these genes in the paleopolyploid *B. napus* genome were found to be associated with a significant effect on the seed glucosinolate content. This example shows the potential of *Arabidopsis-Brassica* comparative genome analysis for synteny-based identification of gene-linked SSR markers that can potentially be used in marker-assisted selection for an important trait in oilseed rape.

Keywords *Brassica napus* · Seed glucosinolates · Candidate genes · SSR markers · Allele-trait association · Synteny

Communicated by M. Sillanpää.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0733-3) contains supplementary material, which is available to authorized users.

M. Hasan · W. Friedt · N. M. Freitag · K. Link · R. J. Snowdon (✉)
Department of Plant Breeding,
Research Centre for BioSystems,
Land Use and Nutrition, Justus Liebig University,
Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany
e-mail: rod.snowdon@agrar.uni-giessen.de

J. Pons-Kühnemann
Department of Biometry and Population Genetics,
Research Centre for BioSystems, Land Use and Nutrition,
Justus Liebig University, Heinrich-Buff-Ring 26-32,
35392 Giessen, Germany

Introduction

Oilseed rape (*Brassica napus* ssp. *napus*; genome AACC, $2n = 38$) is the most important source of vegetable oil in Europe and the second most important oilseed crop in the world after soybean. *Brassica napus* is a relatively young species that originated in a limited geographic region through spontaneous hybridisations between turnip rape (*B. rapa*; AA, $2n = 20$) and cabbage (*B. oleracea*; CC, $2n = 18$) genotypes (Kimber and McGregor 1995). The gene pool of elite oilseed rape breeding material has been depleted by breeding for specific oil and seed quality traits, with particularly strong bottleneck selection for zero seed erucic acid (C22:1) and low seed glucosinolate content (so-called

double-low, 00 or canola quality). The first erucic acid-free variety, derived from a spontaneous mutant of the German spring rapeseed cultivar "Liho", was released in Canada in the early 1970s. In 1969 the Polish spring rape variety "Bronowski" was identified as a low-glucosinolate form, and this cultivar provided the basis for an international backcrossing program to introduce this polygenic trait into high-yielding erucic acid-free breeding lines. The result was the release in 1974 of the first 00-quality spring rapeseed variety, "Tower". Today the overwhelming majority of modern spring and winter oilseed rape varieties have 00-quality. However, residual segments of the "Bronowski" genotype in modern cultivars are believed to cause reductions in yield, winter hardiness, and oil content (Sharpe and Lydiate 2003). Furthermore, the restricted genetic variability in modern 00-quality oilseed rape (Hasan et al. 2006) is particularly relevant with regard to the development of genetically diverse heterotic pools of adapted genotypes for hybrid breeding. For this purpose *B. napus* genotypes containing high levels of erucic acid and seed glucosinolates (so-called ++ seed quality) represent a comparatively genetically divergent source of germplasm (Röbbelen 1975; Thompson 1983; Schuster 1987).

Glucosinolates are secondary plant metabolites synthesized by species in the family Brassicaceae, which includes a large number of economically important *Brassica* crops and the model plant *Arabidopsis thaliana*. The various glucosinolate compounds are designated aliphatic, aromatic and indole glucosinolates depending on whether they originate from aliphatic amino acids (methionine, alanine, valine, leucine, isoleucine), aromatic amino acids (tyrosine, phenylalanine) or tryptophan, respectively. Together with the myrosinase enzymes (also known as thioglucosidases) glucosinolates form the glucosinolate-myrosinase system (Wittstock and Halkier 2002), which is generally believed to be part of the plant's defence against insects and possibly also against pathogens (Rask et al. 2000). When plant tissue is damaged the glucosinolates are hydrolysed by the myrosinases to release a range of defence compounds from substrate cells (Mithen et al. 2000).

After oil extraction from the seeds of oilseed rape the residual meal, which contains 38–44% of high quality protein, is used in livestock feed mixtures. However, high intakes of glucosinolates and their degradation products in rapeseed-based meals can cause problems of palatability and are associated with goitrogenic, liver and kidney abnormalities (Walker and Booth 2001). This particularly limits the use of the rich-protein meal as a feed supplement for monogastric livestock. Seed-specific optimisation of the glucosinolate content and composition would help to improve the nutritional value of rapeseed meal without compensating the disease and pest resistance properties in the crop (Wittstock and Halkier 2002). Genetic control of glucosinolate accumulation is polygenic, and the biosynthesis pathways for different glucosinolate compounds are well characterised in *A. thaliana*. Furthermore, Howell et al. (2003) demonstrated through comparative mapping that high-glucosinolate rapeseed genotypes often carry low-glucosinolate alleles at one or more of the major quantitative trait loci (QTL) controlling seed glucosinolate accumulation. With effective molecular markers for marker-assisted selection these genotypes could be used to introduce new genetic variation for low seed glucosinolate content into breeding programs. A number of studies have described detection of QTL for total seed glucosinolate content in different oilseed rape crosses (Uzunova et al. 1995; Howell et al. 2003; Sharpe and Lydiate 2003; Zhao and Meng 2003; Basunanda et al. 2007). Four QTL on *B. napus* chromosomes N9, N12, N17, and N19 were detected independently in different studies, indicating that these QTL represent major loci that influence seed glucosinolate content in different materials. The QTL on N9, N12 and N19 were found by Howell et al. (2003) to be homoeologous loci.

Markers for QTL detected by classical genetic mapping in individual crosses are not necessarily transferable to other material, and the utility of QTL-linked markers for marker-assisted selection is limited by the relative effects of individual loci on the trait of interest (Snowdon and Friedt 2004). On the other hand, detection of marker-trait associations based on linkage disequilibrium in genetically diverse materials can identify alleles with direct linkage to genes showing significant effects on the trait. In plant breeding populations the technique has seldom been used for marker development (Breseghello and Sorrells 2006), although association approaches can be particularly suitable for identification of useful allelic variation in genetically diverse genotype collections (Flint-Garcia et al. 2003). To date association studies in plants have mainly been performed in species for which extensive sequence data is available. For example, genome-wide analysis was used by Aranzana et al. (2006) to confirm trait associations of flowering time and disease resistance genes in *A. thaliana*, and sequence diversity in trait-relevant candidate genes has also been used to uncover allele-trait associations in *Arabidopsis* (Hagenblad and Nordborg 2002; Balasubramanian et al. 2006; Ehrenreich et al. 2007), rice (Bao et al. 2006; Iwata et al. 2007) and maize (Thornsberry et al. 2001; Wilson et al. 2004; Yu et al. 2006). On the other hand, genome-wide and candidate gene association studies have also been successful in crops with less well-characterised genomes, for example potato (Gebhardt et al. 2004). Oesterberg et al. (2002) identified associations with flowering time in sequence variants of the *COL1* gene in *Brassica nigra*, but to date this remains the only report of an association study in a brassica crop.

In recent years considerable progress in the accumulation and distribution of *Brassica* genome data has been

Table 1 Results of Bayesian clustering within two sets of genetically diverse *Brassica napus* genotypes

Cluster allocation by STRUCTURE 2.1	*Brassica napus* gene bank acccessions	Mean total seed glucosinolate content (μmol/g)	Type
Gene bank accessions, group 1	Spaeths Zollerngold	62.5	SOSR
	Marnoo	21.3	SOSR
	Bronowski	11.9	SOSR
	Ceska Krajova	60.1	SOSR
	Duplo	14.7	SOSR
	Nugget	33.8	SOSR
	Wesway	51.9	SOSR
	Tribute	15.1	SOSR
	Svalöf's Gulle	61.2	SOSR
	Olga	66.5	SOSR
	Alku	52.3	SOSR
	Topas	16.3	SOSR
	Line	9.8	SOSR
	Erglu	30.2	SF
	Janetzkis Sommerraps	74.4	SOSR
	Dux	22.8	SOSR
	Liho	52.8	SF
	Kroko	69.2	SF
	Petranova	103.1	SF
	Mestnij	87.1	WOSR
Gene bank accessions, group 2	Mytnickij	82.6	WOSR
	Aphid Resistant Rape	77.0	WF
	Hokkai 3-go	79.2	WOSR
	Taisetsu	71.3	WV
	Krapphauser	83.7	WOSR
	Chuosenshu	56.1	WV
	Akela	91.6	WF
	English Giant	88.9	WF
	Nunsdale	86.5	WF
	Clubroot Resistant	88.9	WF
	Dwarf Essex	71.2	WF
	Michinaku natane	92.6	WF
	Palu	85.9	WF
	Parapluie	70.4	WF
	Samo	66.1	WF
	Hakuran	87.4	SV
	Tira	100.5	SF
	Kromerska	79.0	WOSR
	Lirafit	22.3	WOSR
	Vinnickij 15/59	93.6	WOSR
Gene bank accessions, group 3	Lenora	94.8	WOSR
	Diamant	84.7	WOSR
	Quinta	61.7	WOSR
	Rapol	93.1	WOSR
	Major	87.0	WOSR
	Markus	81.5	WOSR
	Librador	10.3	WOSR
	Doral	84.2	WOSR
	Groß-Lüsewitzer	89.5	WOSR
	Libritta	11.7	WOSR
	Liporta	20.0	WOSR
	Lirakotta	74.8	WOSR
	Baltia	73.4	WOSR
	Skziverskij	93.8	WOSR
	Jet Neuf	93.2	WOSR
	Jupiter	81.5	WOSR
	Panter	86.7	WOSR
	Norde	76.2	WOSR
	Matador	76.0	WOSR
	Darmor	16.2	WOSR
	Askari	69.5	WOSR
	Quedlinburger Platzfester	87.8	WOSR
	Ramses	88.4	WOSR
	Sarepta	75.7	WOSR
	Skrzeszowicki	80.1	WOSR
	Slovenska Krajova	90.8	WOSR
	Sonnengold	85.2	WOSR
	Trebicska	83.2	WOSR
	Victor	86.9	WOSR
	Gülzower Ölquell	91.1	WOSR
	Bienvenue	56.0	WOSR
	Brink	77.1	WOSR
	Ceres	12.2	WOSR
	Coriander	88.8	WOSR
	Doral	87.6	WOSR
	Edita	85.7	WOSR
	Erra	77.4	WOSR
	Liglory	17.7	WOSR
	Janpol	90.2	WOSR
	Lesira	38.5	WOSR
	Madora	23.1	WOSR
	Jantar	4.2	WOSR
	Bolko	8.1	WOSR
	Silona	92.1	WF
	Bladkool	93.1	WF
	Dippes	81.8	WOSR
	Anja	77.0	WOSR
	Binera	17.6	WF
	Liragrün	75.8	WF
	Moldavia	90.8	WOSR

Table 1 continued

Cluster allocation by STRUCTURE 2.1	*Brassica napus* gene bank acccessions	Mean total seed glucosinolate content (μmol/g)	Type
	Nemertschanskij 1	84.1	WOSR
	Start	23.3	WOSR
	Sobotkowski	83.3	WOSR
	Wolynski	85.8	WOSR
Winter rapeseed genotypes, group 1	Campari	13.6	WF
	Caramba	11.6	WF
	Licapo	6.7	WF
	Resyn. H048	70.7	RS
	Resyn. H226	32.9	RS
	Sollux	81.8	WOSR
	1012–98	16.5	RS
	Erox	44.8	WOSR
	Resyn. Gö S4	81.8	RS
Winter rapeseed genotypes, group 2	Amor	27.6	WOSR
	Ascona	9.8	WOSR
	Askari	80.3	WOSR
	Bienvenue	7.5	WOSR
	Cobra	21.7	WOSR
	DH1from Apex x Mohican	10.2	WOSR
	DH7 from Apex x Mohican	12.2	WOSR
	Duell	8.9	WOSR
	Jessica	8.1	WOSR
	Lion	8.7	WOSR
	Lipid	20.8	WOSR
	Lirabon	8.1	WOSR
	Lisek	10.0	WOSR
	Magnum	9.5	WOSR
	Phil	44.9	WOSR
	Samourai	26.1	WOSR
	Orlando	9.7	WOSR
	Quinta	12.7	WOSR
	Flip	8.7	WOSR
	Maplus	39.0	WOSR
Winter rapeseed genotypes, group 3	Aviso	11.0	WOSR
	Bristol	10.8	WOSR
	Columbus	16.8	WOSR
	Contact	11.8	WOSR
	Express	18.0	WOSR
	Falcon	9.6	WOSR
	Idol	61.4	WOSR
	Jet Neuf	8.6	WOSR
	Leopard	21.2	WOSR
	Lirajet	51.5	WOSR
	Lisabeth	12.4	WOSR
	Pollen	11.8	WOSR
	Prince	25.7	WOSR
	Wotan	9.1	WOSR
	Zenith	9.0	WOSR
	Capitol	11.0	WOSR
	Mansholt's Hamburger Raps	20.0	WOSR

The first genotype set comprises 94 *B. napus* gene bank accessions including old oilseed, fodder and vegetable rape varieties from different geographic origins. For accession numbers and origins of the gene bank accessions see Hasan et al. (2006). The second genotype set contains 46 winter-type varieties and breeding lines

SOSR spring oilseed rape, *SF* spring-type fodder rape, *WF* winter-type fodder rape, *SV* spring-type vegetable rape, *WOSR* winter oilseed rape, *WV* winter-type vegetable rape, *RS* resynthesised rapeseed

made by participants in the Multinational Brassica Genome Project (see http://www.brassica.info/). With the increasing amount of *Brassica-Arabidopsis* comparative genomics data it is becoming possible to navigate between and among the chromosomes of *A. thaliana* and *B. napus*. In some cases this can enable the map positions of *B. napus* QTL for traits of agronomic importance to be compared with the positions of potential candidate genes in the model genome. *Brassica* sequences with homology to the corresponding *A. thaliana* regions can then potentially be used for database-oriented identification of new markers for fine mapping, association studies or marker-assisted selection towards trait improvement. Moreover, it is also potentially possible to identify relevant candidate genes for important traits in oilseed rape, based on their positions in syntenic maps compared to important QTL.

According to Peleman and van der Voort (2003), distinguishing as many alleles as possible at loci of interest and determining phenotypic values for these alleles should greatly improve the predictive power of selection markers and enable marker-assisted combination of positive alleles for different loci. Because *B. napus* is a facultative outcrosser, a high degree of heterozygosity would be expected in natural populations. However, cultivars and gene bank collections of this amphipolyploid species are maintained as pure-breeding lines by self-pollination, so that genetically diverse genotype collections are effectively homozygous inbred lines and therefore ideal for allele-trait association studies. In this study we performed structure-based association studies for seed glucosinolate content in two divergent sets of *B. napus* genotypes. For the association studies a set

of new simple-sequence repeat (SSR) markers was developed whose closest orthologs in *A. thaliana* are physically closely linked to promising candidate genes for seed glucosinolate biosynthesis. In order to incorporate information on the population structure into the association analysis, the potentially gene-linked markers were supplemented with a large set of SSR markers distributed throughout the genome. Furthermore, we also tested trait associations of previously mapped SSR markers for which homologous loci were localised near major QTL for seed glucosinolate content. This research tests the utility of association studies based on gene-linked and QTL-linked markers to detect seed glucosinolate content in *B. napus*. At the same time we describe a technique for synteny-based identification of gene-linked SSR markers for marker development in oilseed rape.

Materials and methods

Plant materials

Two different sets of genetically diverse *B. napus* genotypes were used for the allele-trait association studies (Table 1). The primary genotype set comprised 94 genetically diverse *B. napus* gene bank accessions from a *B. napus* "core collection" which spans the genetic diversity present in European gene bank collections of winter and spring oilseed, fodder and vegetable rape varieties. The core collection was selected based on phenotypic descriptors that were assessed during a European project on genetic diversity in *Brassica* crop species (Lühs et al. 2003; Poulsen et al. 2004), in combination with available pedigree information. The genetic diversity within the core collection has been described previously (Hasan et al. 2006). A second set of genotypes was used to further investigate markers that showed significant associations with glucosinolate content in the gene bank accessions. The second set of material comprised 46 winter-type, predominantly oilseed rape genotypes that were chosen based on pedigree knowledge to cover as broadly as possible the genetic and phenotypic variation present in current western European cultivars. Thirty-two of the 46 genotypes were cultivars or breeding lines with low seed glucosinolate content.

The gene bank accessions were grown in field trials in Rauischholzhausen, Germany, in 2003 and 2004, while the second set of genotypes were grown in Einbeck, Germany, from 2003 to 2005. Seeds were harvested from five to six self-pollinated plants per genotype and mean total seed glucosinolate content was estimated by near infrared reflectance spectroscopy (NIRS). Approximately 2 g seeds per sample were measured by monochromator analysis in a spinning cell at all wavelengths between 1,100 and 1,800 nm. For the molecular marker analyses genomic DNA samples were extracted from young leaves of five pooled plants per genotype using a standard CTAB extraction protocol (Doyle and Doyle 1990).

Potentially gene-linked SSR markers identified by comparative genome analysis

Twelve new *Brassica* SSR primer combinations were identified in sequences with homology to *A. thaliana* chromosome regions containing relevant candidate genes for glucosinolate content. First, interesting *Arabidopsis* chromosome regions with putative associations to glucosinolate QTL in *B. napus* were identified by *in silico* localisation of the closest *A. thaliana* orthologs for RFLP marker sequences from three major homoeologous *B. napus* glucosinolate QTL. Sequences for the RFLP probes CA72, pO119, pW141, pW200, and pW157, which were reported by Howell et al. (2003) to label loci belonging to homoeologous QTL on *B. napus* chromosomes N9, N12 and N19, were obtained from the EMBL database of the European Bioinformatics Institute (http://www.ebi.ac.uk/embl/). Four *A. thaliana* chromosome regions containing orthologous sequences to one or more of the abovementioned markers were identified based on the BLASTn annotations reported by Parkin et al. (2005). By searching the biological process "glucosinolate biosynthesis" in the gene ontology database of the Arabidopsis Information Resource (TAIR: http://www.arabidopsis.org/) the genes *cytochrome P450 monooxygenase 83B1* (*CYP83B1*: At4g31500), *cytochrome P450 79A2* (*CYP79A2*: At5g05260), *methylthioalkylmalate synthase* (*MAM1/MAML*: tandem duplication At5g23010/At5g23020) and *altered tryptophan regulation* (*ATR1*: At5g60890) were identified as the physically closest potential candidates to the QTL-marker orthologs in the four relevant chromosome regions on *A. thaliana* chromosomes 4 and 5, respectively.

The "SSR Search" tool of the Brassica ASTRA database from the Plant Genetics and Genomics platform of Primary Industries Research Victoria, Australia (http://hornbill.cspp.latrobe.edu.au/cgi-binpub/brassica/index.pl) was used to search *A. thaliana* genome regions up to 500 kbp upstream and downstream of the four selected candidate genes for potentially gene-linked SSR sequences. A total of thirty-two putative *Brassica* SSR primer combinations were identified in the different candidate gene regions and all primers were tested for suitability in *B. napus*. Twelve of the primer pairs gave clear, reproducible and polymorphic amplification products at one or more loci in *B. napus* and were used to screen for allelic polymorphisms in the 94 gene bank accessions. Sequences for these new SSR primers are available in Supplementary Table 1. All of the four putative candidate genes were represented by these potentially gene-linked SSR markers.

Four publicly-available *Brassica* SSRs (BRAS014, CB10425, Ol10-D03 and Ol11-C02) were also included in the association analysis in the primary genotype set. These four primers amplify SSR markers that are known to be linked to the seed glucosinolate QTL on *B. napus* N17 (Basunanda et al. 2007; F. Lipsa and R. Snowdon, unpublished results), for which no tightly-linked RFLP markers with clear synteny to *Arabidopsis* regions containing putative candidate genes were available.

Genome-wide SSR markers

Population structure among the 94 gene bank accessions was analysed using allelic data from 46 publicly available *Brassica* SSR primer combinations that amplify loci dispersed throughout the entire *B. napus* genome. Thirty of these primer combinations were also used previously to screen the genetic diversity in these genotypes (Hasan et al. 2006). For population structure analysis in the 46 winter oilseed rape genotypes, allelic data from a total of 104 SSR primer combinations that amplified 559 marker alleles were kindly provided by the breeding companies KWS Saat AG, SW Seed GmbH and Saaten-Union Resistenzlabor GmbH. This data was generated as part of the project GABI-BRIDGE: *Brassica napus* allelic diversity in candidate genes.

SSR analyses

PCR reactions were performed in a GeneAmp PCR System 9700 thermal cycler in a volume of 15 µL containing 20 ng of DNA template, 0.75 pmol of each primer, 0.2 mM dNTP mix, 1×PCR reaction buffer containing 15 mM $MgCl_2$, a further 1 mM $MgCl_2$ and 0.25 units of *Taq* DNA polymerase (Qiagen, Hilden, Germany). To reduce primer-labelling costs, PCR products were labelled with the M13-tailing technique described by Berg and Olaisen (1994). In this method the fluorescently labelled universal M13 primer 5′-AGGGTTTTCCCAGTCACGACGTT-3′ is added to the PCR reaction, and the forward primer of each SSR is appended with the sequence 5′-TTTCCCAGTCACGACGTT-3′. After the first round of amplification the PCR fragments are subsequently amplified by the labelled universal primer. A touch-down PCR cycle was modified from the procedure described by Xu et al. (2005) as follows: An initial denaturation was performed at 95°C for 2 min, followed by five cycles of denaturation for 45 s at 95°C, annealing for 5 min beginning at 68°C and decreasing by 2°C in each subsequent cycle, and extension for 1 min at 72°C. Then five cycles were performed with 45 s denaturation at 95°C, 1 min annealing beginning at 58°C and decreasing 2°C in each subsequent cycle, and 1 min of extension at 72°C. The PCR was then completed with an additional 27 cycles of 45 s denaturation at 94°C, 2 min of annealing at 47°C, and 30 s of extension at 72°C, with a final extension at 72°C for 10 min. The SSR polymorphisms were separated and visualised using a LI-COR GeneReadir 4200 (MWG Biotech, Ebersberg, Germany). Allele sizes including the 23 bp labelled M13 tail primer were scored with the software RFLP-SCAN (Version 2.01, Scanalytics Inc., Fairfax, VA, USA) based on a labelled length standard.

Analysis of population structure

A potential problem for every population-based association study is the presence of undetected population structure that can mimic the signal of association and lead to false positives or to missed real effects (Marchini et al. 2004). We analysed the population structure with the model-based Bayesian clustering approach in the software STRUCTURE 2.1 (Pritchard et al. 2000) using allelic data from genome-wide SSR markers. Many *Brassica* SSR primer combinations amplify different marker alleles at multiple loci in the paleopolyploid *B. napus* genome, and homoplasic alleles may be amplified at different loci. This means it can be difficult or impossible to assign the different marker alleles to individual loci in genotypes with high allelic diversity. Hence all SSR alleles were scored dominantly as present or absent in each genotype, and no information on marker linkage could be included in the population structure model. Therefore the model of no admixture was applied for the analysis of population structure, as stipulated by the user instructions for STRUCTURE 2.1. The basis of the Bayesian clustering method is the allocation of individual genotypes to groups in such a way that Hardy–Weinberg equilibrium and linkage equilibrium are valid within clusters, whereas these forms of equilibrium are absent between clusters. For each of the two genotype sets the optimum number of clusters (K) was selected after ten independent runs of a burn-in of 100,000 iterations, followed by 100,000 iterations using a model allowing for no admixture and correlated allele frequencies. We tested for $K = 1$–10 in the gene bank accessions and $K = 1$–5 in the set of winter rapeseed genotypes. A summary of the average logarithm of the probability of data likelihoods ($LnP(D)$) for both sets of genotypes is given in Table 2.

Structure-based association analysis

Due to the high allelic diversity, the clear population structure and an expectation of low familial relatedness due to the way the genotype collections were selected, we performed structured association tests rather than using a mixed-model approach (Yu et al. 2006) to control for false positives (type I errors) caused by the population structure. Associations between the marker data and the total seed

Table 2 Summary of the average logarithm of the probability of data likelihoods (*LnP*(*D*)) for two distinct sets of genetically diverse *Brassica napus* genotypes

Genotype set	*K*	Average *Ln P*(*D*)	SD
94 gene bank accessions	1	−11,084.47	0.57
	2	−10,372.74	1.35
	3	−10,100.25	18.35
	4	−10,146.21	140.23
	5	−10,076.47	115.89
	6	−10,043.32	116.17
	7	−10,028.03	217.97
	8	−10,386.28	547.07
	9	−10,835.44	2083.44
	10	−10,331.89	232.89
46 winter oilseed rape genotypes	1	−10,316.79	5.82
	2	−9,872.74	8.79
	3	−9,719.00	8.21
	4	−9,660.20	19.76
	5	−10,090.76	467.30

Likelihoods were averaged over ten independent runs of a burn-in of 100,000 iterations, followed by 100,000 iterations using a model allowing for no admixture and correlated allele frequencies. The set of 94 gene bank accessions were tested for $K = 1$–10 subpopulations, while the 46 winter oilseed rape genotypes were tested for $K = 1$–5 subpopulations

glucosinolate content were tested using the logistic regression approach of Pritchard et al. (2000), as modified by Thornsberry et al. (2001) in order to deal with quantitative traits. This procedure is implemented in the software package TASSEL 2 (http://www.maizegenetics.net/). The response variable was the presence or absence of the SSR polymorphism, while the quantitative trait (total seed glucosinolate content) and the population structure (Q-matrix) were used as independent variables. In the null hypothesis, candidate polymorphisms are independent of the seed glucosinolate content (only the Q-matrix is included in the model), whereas in the alternative hypothesis the candidate polymorphisms are associated with the seed glucosinolate content (the quantitative trait and the Q-matrix are both included in the model). The test statistic Λ derives from the ratio between these two likelihoods and indicates the degree of association between individual polymorphisms and the quantitative trait. The null distribution of random markers was simulated by 1,000 permutations of the quantitative trait data over all genotypes. The *P* value for individual polymorphisms was calculated as the proportion of observed Λ greater then the maximal permuted Λ. This approach enables evaluation of associations involving quantitative traits while controlling for population structure. Only markers with an allele frequency of 5% or greater were included in the association analysis. In order to account for type I error bias the *P* values were adjusted for multiple tests using a procedure proposed by Whitt and Buckler (2003) based on the permuted *P* values of random markers. The rescaled *P* value accounts for the proportion of random markers with a permuted *P* value less than or equal to 0.05. According to Thornsberry et al. (2001) the true test statistic probably lies somewhere between the rescaled *P* value and *P*(Λ), since some of the random markers are probably truly associated with the trait. Therefore *P*(Λ) provides an overview of markers with potential association to trait, while the rescaled *P* value is a conservative test to reduce the likelihood of false-positive associations.

Map positions of markers with significant associations to seed glucosinolate content

Where possible the map positions of markers with significant associations to seed glucosinolate content were identified in existing *B. napus* genetic maps. For SSR primers where the allele sizes were not given in published maps, the positions of all known loci were recorded. Annotations of public *Brassica* SSR markers to the *A. thaliana* genome were obtained from the public microsatellite database at http://brassica.bbsrc.ac.uk/cgi-bin/ace/searches/browser/BrassicaDB. Glucosinolate-associated SSR markers from the set of new, synteny-based markers were screened for polymorphisms among the parents of three different doubled-haploid (DH) mapping populations and integrated into the maps of these populations where possible. The genetic mapping procedure followed Basunanda et al. (2007). Markers that deviated significantly ($P > 0.01$) from the expected 1:1 segregation in the DH populations were presumed to represent two or more homoeologous loci with identical allele sizes and hence could not be mapped.

Linkage disequilibrium

In order to gain information about the putative map positions of the gene-linked SSR markers in cases where these markers could not be directly mapped in available mapping populations, we used TASSEL to analyse linkage disequilibrium (LD) based on the parameter r^2 (the squared allele frequency correlation). The significance of the LD between marker pairs was determined by Fisher's exact test. Due to the pre-selection for the association analysis only markers with a minimum allele frequency of 0.05 were included in the LD analysis, as recommended by Thornsberry et al. (2001). In a first step the LD was calculated among all markers that were significantly associated with seed glucosinolate content, in order to identify previously mapped markers with high LD to new, unmapped markers. Subsequently, the LD was recalculated within groups of markers with significant LD. Levels of LD were expected to be somewhat underestimated by the available SSR allele data,

because in a paleopolyploid like *B. napus* it is known that identical alleles can be amplified by multiple loci. Therefore no presumption was made that two markers amplified by the same primer combination must necessarily belong to the same locus, even when these showed high LD.

Results

As expected from previous investigations (Hasan et al. 2006) a significant population structure was observed among the 94 gene bank accessions. As seen in Table 2, the highest average likelihoods for the population structure in this set of material were observed with K values between 3 and 7, whereby the most stable prediction (standard deviation = 18.35) was obtained with three groups ($K = 3$). These groups comprised (1) twenty genotypes of predominantly spring-type oilseed rape, (2) twenty genotypes of mainly fodder or vegetable rape, and (3) fifty-four predominantly winter oilseed rape genotypes, respectively. The most stable and high average likelihoods for population structure amongst the 46 winter rapeseed genotypes were obtained with $K = 2$ or 3. Seventeen oilseed genotypes were strongly assigned to the same group in both cases, while the remaining 29 genotypes were divided with $K = 3$ into another group of 20 oilseed types and a group of nine exotic genotypes, including fodder rape varieties and resynthesised (RS) rapeseed lines. Such material is known to represent a divergent *B. napus* gene pool in comparison to oilseed *B. napus* genotypes (Seyis et al. 2003), and since most of these exotic genotypes also exhibited high glucosinolate content this grouping was expected to be particularly relevant for the association analysis with glucosinolate content. We therefore used the respective Q-matrix outputs of the three-subpopulation runs ($K = 3$) for the structure-based association analyses in both sets of genotypes. A broad range in total seed glucosinolate content was observed among the 94 gene bank accessions, whereas the winter rapeseed set included 32 genotypes with low seed glucosinolate content (<25 µmol/mg dry weight). Details of the groupings of the accessions along with the mean total seed glucosinolate data used for the association analyses are given in Table 1.

Using the complete set of 62 polymorphic SSR primer combinations, a total of 348 polymorphic SSR marker alleles were amplified in the 94 gene bank accessions. Of these, a total of 51 marker alleles from 29 SSR primer combinations were found to exhibit a significant association ($P \leq 0.05$) to total seed glucosinolate content in the 94 *B. napus* gene bank accessions. Ten of the markers also exhibited significant association using the rescaled P values, indicating that these associations are not likely to be caused by type I errors. All markers with significant associations to seed glucosinolate content are described in detail in Table 3, including information (where available) on map positions and annotations to the *A. thaliana* genome. Positions of glucosinolate-associated markers with known physical linkage to relevant candidate genes in *A. thaliana* are shown in Fig. 1. The phenotypic distributions of the genotypes with the 51 marker alleles showing significant associations to seed glucosinolate content in the gene bank accessions are illustrated by box-plots in Fig. 2. Allelic data for all SSR markers with significant associations to total seed glucosinolate content are available in Supplementary Table 2.

In order to get an idea of the abundance of these glucosinolate-associated markers in European winter rapeseed, and particularly in material with low seed glucosinolate content, we re-screened all of the significantly associated SSRs in the set of 46 winter rapeseed genotypes. Interestingly, many of the significantly associated marker alleles were only found at very low frequencies (<5%) in the winter rapeseed set, and only three markers with frequencies of greater than 5% also showed significant association to seed glucosinolate content among these 46 genotypes. All three of these markers were associated with low glucosinolate content. In two cases (Na12-G04 and Ol10-D02), different marker alleles amplified by the same SSR primer combinations showed significant associations in the two different sets of materials. The marker Gi31_387 was the only marker allele that was found to be significantly associated to seed glucosinolate content in both sets of materials. As seen in Table 3, the sequence of Gi31 is located in *A. thaliana* only 736 bp downstream of the gene *CYP83B1*. In the gene bank accessions, the two marker alleles amplified by this primer combination, Gi31_385 and Gi31_387, are associated with significantly increased and decreased total seed glucosinolate content, respectively. The allele-trait association of both alleles together with the very short physical distance to the candidate gene strongly support the potential involvement of homoeologous *CYP83B1* copies in biosynthesis of seed glucosinolates in *B. napus*. Two of four marker alleles from the SSR sequence Gi30, which is located somewhat further away from *CYP83B1*, were also significantly associated with total seed glucosinolate content in the gene bank accessions.

For the three other candidate genes we were also able to identify putatively linked SSR markers with significant associations to seed total glucosinolate content (Table 3). The SSR Gi24, located 166 kbp from *CYP79A2* in *A. thaliana*, amplified a single band whose presence in the gene bank accessions was associated with an increased glucosinolate content. The SSR Gi12, although located 382 kbp away from *ATR1* in *A. thaliana*, amplified a single band that was associated with a mean decrease in total glucosinolate content. Two of three bands amplified by the SSR

Table 3 Details of SSR marker alleles showing significant associations (*P* values) to seed glucosinolate (GSL) content in a set of 94 genetically diverse *Brassica napus* gene bank accessions

Markers with potential physical linkage to GSL candidate genes[a]

SSR primer	Number of amplified bands	*A. thaliana* chromosome and position in Mbp	Linked GSL candidate gene	Distance from gene in *A. thaliana* (bp)	GSL-associated marker allele(s)	Mean seed GSL content (μmol/g) in 94 gene bank accessions	Allele frequency	*P* value
Gi30	4	At4: 15237362	CYP83B1	37,954	Gi30_385	72.79	72.83	0.004
					Gi30_390	51.02	20.65	0.024
Gi31	3	At4: 15276052	CYP83B1	736	Gi31_385	83.70	9.57	0.013
					Gi31_387	61.25	75.53	0.000*
Gi24	1	At5: 1727906	CYP79A2	166,140	Gi24_247	73.05	68.28	0.010
Gi12	1	At5: 24895600	ATR1	382,023	Gi12_159	76.49	40.45	0.010
Gi28	3	At5: 7808676	*MAM1/MAML*	101,782	Gi28_442	70.36	80.00	0.000*
					Gi28_444	37.66	18.82	0.000*

SSR markers randomly distributed in the *B. napus* genome[b]

SSR primer	Number of amplified bands	Known map positions of homoeologous loci	Known annotations to *Arabidopsis thaliana*	GSL-associated marker allele(s)	Mean seed GSL content (μmol/g) in 94 gene bank accessions	Allele frequency	*P* value
BRAS014	5	N6/N17		BRAS014_162	0.540	35.00	0.044
BRAS020	13	N9		BRAS020_260	0.365	5.32	0.019
CB10425	5	N17		CB10425_327	0.366	17.07	0.016
GMR013	6	N9/N19		GMR013_176	0.577	59.57	0.000*
				GMR013_188	0.763	47.87	0.025
				GMR013_195	0.507	42.55	0.000*
MR111	4	N13/N14/N19		MR111_124	0.597	26.60	0.021
				MR111_130	0.598	12.77	0.047
				MR111_133	0.717	19.15	0.001*
Na10-C01	25	N4/N13/N14/N17	At5g31905	Na10-C01_170	0.519	55.32	0.000*
				Na10-C01_224	0.694	66.74	0.015
				Na10-C01_235	0.663	92.55	0.037
				Na10-C01_268	0.886	8.51	0.006
				Na10-C01_282	0.172	6.38	0.003
Na10-D03	2	N13		Na10-D03_176	0.688	90.43	0.019
Na10-H06	4	N11		Na10-H06_160	0.705	81.72	0.002
		N11		Na10-H06_162	0.382	15.00	0.002
Na12-F12	3	N13		Na12-F12_193	0.657	8.51	0.041
Na12-G04	10	N9/N14/N19		Na12-G04_169	0.740	32.61	0.007
				Na12-G04_179	0.455	15.38	0.002
				Na12-G04_183	0.807	25.00	0.002
Na14-E11	14	N4/N14/N15		Na14-E11_111	0.601	65.96	0.030
				Na14-E11_119	0.830	19.15	0.003
				Na14-E11_131	0.768	28.72	0.049
Ol09-A06	6	N12	At2g41710	Ol09-A06_109	0.736	56.38	0.033
Ol10-A05	7	N2/N14		Ol10-A05_212	0.601	70.21	0.037
				Ol10-A05_214	0.798	25.53	0.037
Ol10-B01	7	N17		Ol10-B01_199	0.424	14.89	0.035
Ol10-D02	11	N9/N15	At1g32240	Ol10-D02_158	0.830	14.13	0.040
Ol11-C02	2	N17		Ol11-C02_155	0.210	11.70	0.000*
Ol11-G11	9	N3/N13	At3g10040	Ol11-G11_187	0.694	61.70	0.047

Table 3 continued

SSR markers randomly distributed in the *B. napus* genome[b]							
SSR primer	Number of amplified bands	Known map positions of homoeologous loci	Known annotations to *Arabidopsis thaliana*	GSL-associated marker allele(s)	Mean seed GSL content (μmol/g) in 94 gene bank accessions	Allele frequency	*P* value
Ol11-H06	8	N19		Ol11-H06_170	0.827	5.32	0.008
Ol12-A04	4	N19	At5g18500	Ol12-A04_151	0.690	86.17	0.041
Ol12-D05	7	N18	At3g58730	Ol12-D05_150	0.622	78.72	0.002
				Ol12-D05_153	0.871	10.64	0.015
				Ol12-D05_155	0.638	9.57	0.009
Ol12-E03	4	N7/N16	At1g69260	Ol12-E03_141	0.612	76.60	0.014
Ol12-F11	6	N1/N11	At4g24210	Ol12-F11_242	0.393	14.89	0.008
		N1/N11		Ol12-F11_251	0.508	35.10	0.004
Ol13-D02A	7	N14		Ol13-D02A_292	0.745	39.36	0.011
Ra2-E11	11	N13		Ra2-E11_192	0.812	32.98	0.001*
		N13		Ra2-E11_222	0.178	5.32	0.006
Ra3-E05	8	N1/N11/N16	At3g06270	Ra3-E05_231	0.472	28.72	0.001*

[a] New SSR markers developed based on potential physical linkage to promising candidate genes for glucosinolate biosynthesis in *Arabidopsis thaliana*

[b] SSR markers dispersed throughout the *B. napus* genome. Primer sequences of the potentially gene-linked SSRs, along with allelic data for all markers associated with seed glucosinolate content, are available in Supplementary Tables 1 and 2, respectively

* Markers that still showed significant association after rescaling (rescaled $P \leq 0.05$)

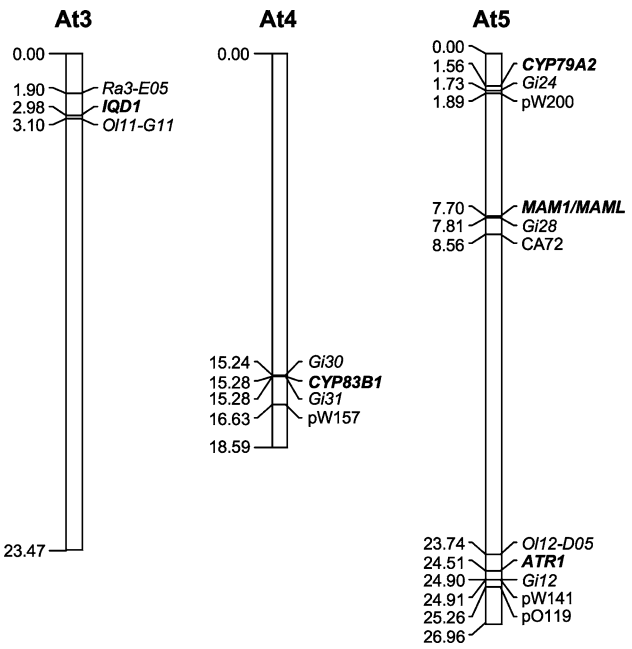

Fig. 1 Chromosomal positions in *Arabidopsis thaliana* (numbers in Mbp) of orthologs for potentially gene-linked *Brassica* SSR markers (*italics*) in comparison to potential candidate genes for glucosinolate biosynthesis (*bold italics*) and RFLP markers (*non-italic*, *non-bold*) located at major seed glucosinolate QTL (Uzunova et al. 1995; Howell et al. 2003). Markers with the prefix *Gi* are new SSR sequences identified by synteny studies in candidate gene regions. Primer sequences of the new SSR markers are available in Supplementary Table 1

Gi28, which is derived from a sequence near the duplicated *MAM1/MAML* gene locus in *A. thaliana*, were associated with increased and reduced total seed glucosinolate content, respectively, in the gene bank accessions. Both of the latter markers, along with Gi31_387, also showed significant associations with rescaled *P*-values, meaning that a type I error is unlikely.

Where sequence and annotation information were available, the glucosinolate-associated SSRs from the genome-wide marker set were also compared with the *Arabidopsis* genome to establish further potential physical linkages to candidate genes. For example, Ol11-G11 amplifies a marker allele with significant association to glucosinolate content, although this SSR maps to two loci on *B. napus* N03 and N13 (Basunanda et al. 2007; Rygulla et al. 2008) where no major QTL for total seed glucosinolate content are known. As shown in Fig. 1, the sequence of Ol11-G11 is annotated in *A. thaliana* to the sequence At3g10040 (data from BrassicaDB), which is located on *A. thaliana* chromosome 3 only 117 kbp downstream from the gene *IQ-DOMAIN 1* (*IQD1*: At3g09710). A further glucosinolate-associated marker, Ra3-E05, also annotates nearby on *A. thaliana* chromosome 3. This chromosome regions shows no obvious homology to *Brassica* regions involved in seed glucosinolate QTL, however *IQD1* is nevertheless a further interesting candidate gene for this trait because it is known to modulate expression

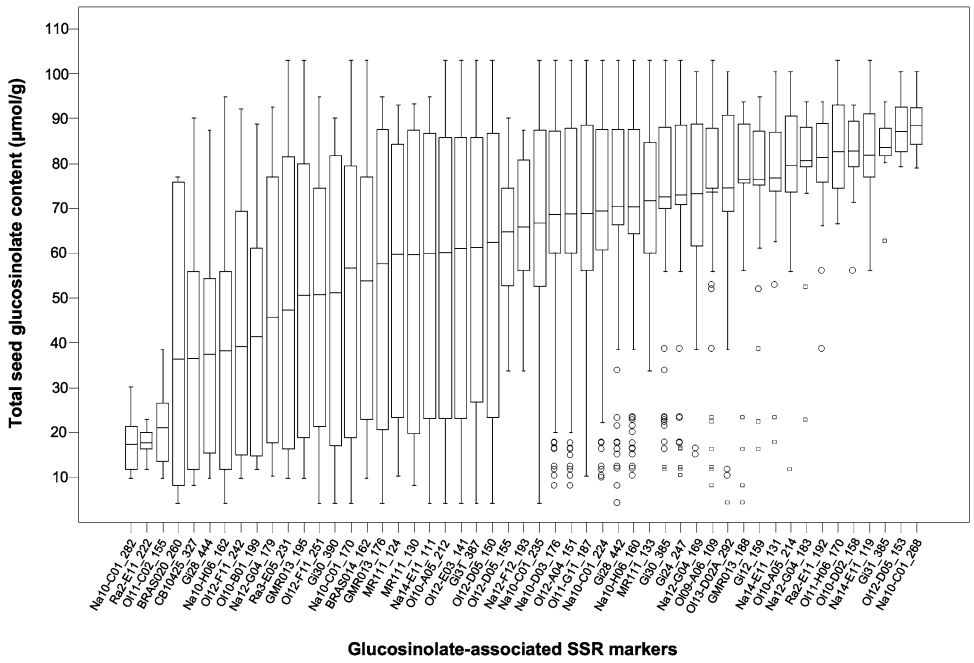

Fig. 2 *Boxplots* showing distributions of total seed glucosinolate content within 94 genetically diverse *B. napus* gene bank accessions for 51 SSR marker alleles with significant association ($P \leq 0.05$) to glucosinolate content. Boxes cover the inter-quartile range around the mean (*horizontal lines*), while the *vertical whiskers* cover the remaining variation with the exception of outliers (*open circles*) and extreme values (*open squares*)

Table 4 Details of SSR marker alleles associated with seed glucosinolate (GSL) content in 46 winter rapeseed varieties and breeding lines

GSL-associated marker allele(s)	Number of amplified bands	GSL-associated marker allele	Mean seed GSL content (µmol/g) in 46 winter rapeseed lines	Allele frequency	*P* value
Gi31	2	Gi31_387	18.59	0.696	0.013*
Na12-G04	3	Na12-G04_165	15.08	0.413	0.036*
Ol10-D02	9	Ol10-D02_155	13.64	0.370	0.03*

Only SSR primers already found to be significantly associated with the trait in the 96 gene bank accessions were screened, in order to assess how much allelic variation for low GSL content from the broader gene pool is present in modern varieties. Detailed allelic data is available in Supplementary Table 2. In all cases the *P* values still showed significant association after rescaling (rescaled $P \leq 0.05$)

of numerous glucosinolate pathway genes. Gain-of-function and loss-of-function *IQD1* alleles in *A. thaliana* are correlated with increased and decreased total glucosinolate accumulation, respectively (Levy et al. 2005).

Knowledge about homoeologous *B. napus* map positions of the SSR markers that showed significant associations to seed total glucosinolate content in the gene bank accessions allowed these to be sorted into groups of markers that might putatively be linked to the same responsible gene or genes, either in the same chromosome region or in common homoeologous chromosome regions. For gene-linked SSR markers with unknown map positions, significant LD to a marker with a known map position was used to infer the putative map positions of the new marker. Results of LD analysis for three groups of markers showing significant LD within their respective group are shown in Fig. 3. The relatively low LD within the groups was reflected by generally low correlation coefficients among the putative neighbouring alleles; however, weak LD patterns could nevertheless be observed. For example one group of markers (Fig. 3a) showed linkage disequilibrium flanking two presumably allelic SSR markers (Gi28_442 and Gi28_444, respectively) derived from sequences near the duplicated *MAM1/MAML* locus in *A. thaliana*. Since all of these markers are known to have loci on *B. napus* chromosome N17 (see Table 3), and in the case of BRAS014_162 and CB10425_327 are known to map near the seed glucosinolate QTL on this chromosome (Basunanda et al. 2007), this indicates that orthologs of *MAM1* and/or *MAML* might be among the responsible genes influencing variation in total seed glucosinolate at this locus. On the other hand the groups of markers exhibiting significant LD did not always derive from the same linkage group. This is not unexpected, since homoeologous SSR loci in *B. napus* often exhibit homoplasic alleles. The consequence is that it is apparently inconclusive to use this method to infer map positions, since it is very difficult to relate specific SSR alleles from a diversity study to unique chromosome locations.

A description of LD might nevertheless be useful to detect potential copies of candidate genes contributing to different QTL. For example, four markers showing LD on either side of the putative gene-linked markers Gi30_385

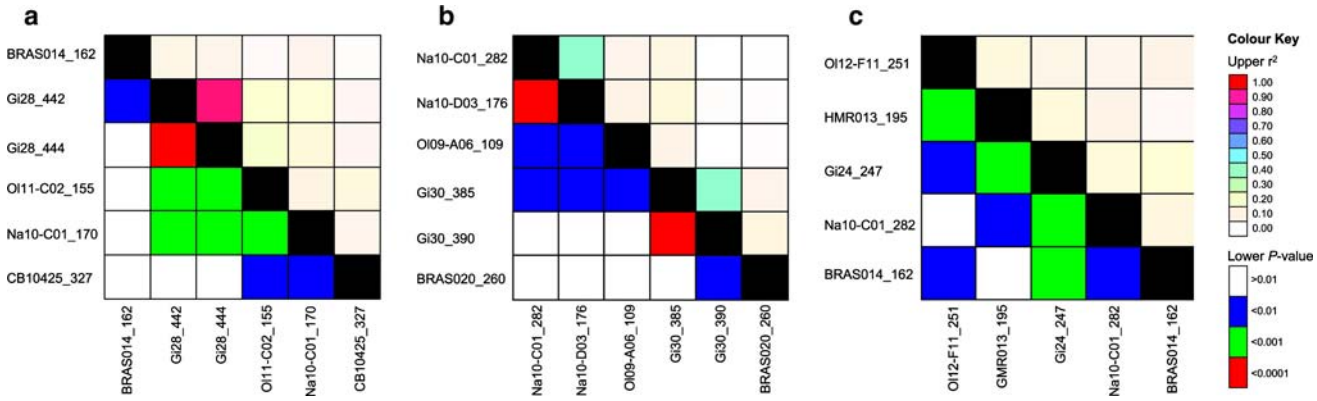

Fig. 3 Linkage disequilibrium (*LD*) around the gene-linked SSR markers **a** Gi28, **b** Gi30, and **c** Gi24 in the 94 *B. napus* gene bank accessions. Cells above the diagonal show the squared allele frequency correlation r^2, while the cells below the diagonal represent the significance level of the LD determined by Fisher's exact test

and Gi30_390 (Fig. 3b) map to different chromosomes: Na10_C01-282 on N14, Na10-D03_176 on N13, Ol9-A06_109 on N12 and BRAS020_260 on N9. The Gi30 sequence is separated in *A. thaliana* by a physical distance of only 30 kb from the glucosinolate biosynthesis gene *CYP83B1*, and the neighbouring RFLP marker pW157 has been mapped in *B. napus* to loci on N1, N11, N9 and N19. Neighbouring markers in the latter two chromosomes have homology to N12. A similar example is shown in Fig. 3b for the trait-associated markers that show LD around the marker Gi24_247, whose sequence in *A. thaliana* is located near the gene *CYP79A2*. These markers have known homoeologous loci on a number of different *B. napus* chromosomes including N1, N4, N6, N9, N11, N13, N14, N17 and N19. Again this suggests that a considerable number of homoeologous chromosome regions might contain copies of CYP79A2, along with different copies of the linked marker loci. The supposition of homologous gene loci associated to glucosinolate content was supported by the fact that many of the new, potentially gene-linked markers deviated significantly from the expected 1:1 segregation in different DH populations (data not shown), which in turn also prevented us from genetically mapping these markers.

Discussion

The aim of this work was to investigate the potential use of *Brassica-Arabidopsis* comparative genomics data for marker and gene identification in oilseed rape based on sequence orthology to *A. thaliana*. Using marker sequences from important *B. napus* QTL for seed total glucosinolate content, along with comparative mapping data, we were able to navigate to potential orthologous genome regions in *A. thaliana*. This enabled us to identify four promising candidate genes with putative physical linkage to homoeologous *Brassica* genome regions involved in seed glucosinolate biosynthesis. Through *in silico* screening of neighbouring *Brassica* genomic and EST sequences a number of new *Brassica* SSR sequences with putative close physical linkage to these four genes were identified, and hom(oe)ologous markers for many of these SSR primers showed significant associations with seed glucosinolate content when screened in *B. napus* gene bank accessions. A number of the trait-associated markers showed skewed segregation in DH mapping populations, indicating the presence of two or more homologous copies of the markers and their putatively linked genes.

In addition, whole-genome association analyses were performed with SSR markers dispersed throughout the *B. napus* genome. This approach also led to the identification of numerous markers with significant associations to glucosinolate content. In some cases the markers were mapped in available *B. napus* genetic maps, and numerous markers showed no apparent relationship to known QTL regions for seed glucosinolate content. This indicates that we may have identified novel allelic variation for this important trait, which should be of considerable interest for breeding purposes. The orthologous sequences in *A. thaliana* for two of the genome-wide SSR markers are closely physically linked to a further promising candidate gene for glucosinolate biosynthesis, *IQD1*. The successful identification of new markers associated to an important seed quality trait underlines the great promise of *in silico* mapping data for gene discovery in oilseed rape based on intergenomic comparisons to *A. thaliana*. Marker sequences, QTL and association data from the crop plant can potentially be used to discover or confirm potential candidate genes in the model species. Furthermore, with the growing resource of *Brassica* genomic sequence data and its alignment to the *A. thaliana* genome, it is now also possible to identify new molecular markers in linkage disequilibrium to genes of interest in *Arabidopsis*. Of particular interest for practical plant breeding is the possibility to identify gene-linked SSR

markers: SSRs are robust, highly polymorphic markers that are relatively cheap and easy to use, and hence predestined for use in marker-assisted selection. On the other hand, recent developments in high-throughput sequencing technologies may soon enable large-scale re-sequencing of candidate gene orthologs. At present it is still difficult to develop locus-specific assays for single-nucleotide polymorphisms (SNPs) in polyploid species like *B. napus*, which can contain many orthologous and paralogous gene copies. However in the near future it is likely that high-throughput SNP discovery will become an important tool for gene discovery and association genetics in oilseed rape. High-density SNP maps will considerably improve our knowledge of LD in *B. napus* and enable much more accurate use of *Arabidopsis-Brassica* comparative genomics data. At present little is known about the extent of LD in *B. napus*.

Seed glucosinolate content in *B. napus* is governed by complex biochemical interactions that make it difficult to predict the actions of individual genes. Because specific pathway branches control the synthesis of different aliphatic, aromatic and indole glucosinolates, dissection of QTL for total glucosinolate content into the individual components is desirable to gain more information about which QTL may involve global pathway genes and which QTL might be more specific for individual compounds. The five candidate genes we identified in this study are well characterised in *A. thaliana*, and the use of selection markers with putative linkage to these genes might enable selection for specific glucosinolate pathway chains. For example, the cytochrome P450 monooxygenase enzyme *CYP83B1* (Hoecker et al. 2004) catalyses the N-hydroxylation of tryptophan-derived indole-3-acetaldoxime, an intermediate in the biosynthesis of indole glucosinolates (Bak et al. 2001; Hansen et al. 2001). The enzyme encoded by *CYP79A2* catalyses the conversion of L-phenylalanine to phenylacetaldoxime, a precursor of the aromatic benzylglucosinolates in *A. thaliana* (Wittstock and Halkier 2000), whereas *ATR1* encodes a transcription factor which activates the expression of tryptophan synthesis genes as well as the tryptophan-metabolizing genes *CYP79B2*, *CYP79B3*, and *CYP83B1*; *ATR1* therefore plays a central regulatory role in the production of indole-3-acetic acid and indole glucosinolates (Celenza et al. 2005). On the other hand, the two tandemly duplicated loci *MAM1* and *MAML* encode genes that catalyze the condensation reactions of the first two cycles in methionine side-chain elongation in *A. thaliana*, therefore they play a vital role in methionine chain elongation and the biosynthesis of aliphatic glucosinolates (Kroymann et al. 2001; Textor et al. 2004). At present it is not known if different homoeologous methylthioalkamate synthase loci in *B. napus* also carry the gene duplication seen in *A. thaliana*.

All but a few of the the glucosinolate-associated markers we identified in the 94 *B. napus* gene bank accessions were found at only very low frequencies in the set of winter rapeseed genotypes. This appears to indicate that the glucosinolate-associated alleles we identified represent novel allelic diversity for this trait that is not present in current European 00-quality oilseed rape cultivars. On the other hand, these results underline the finding of Howell et al. (2003) that most low-glucosinolate cultivars still contain alleles at some loci that in fact are associated with increased total glucosinolate content. The markers we identified will potentially help to further reduce glucosinolate content in existing elite 00-quality oilseed rape, and to introduce new genetic diversity into the comparatively narrow gene pool of 00-quality rapeseed.

Overall the results of this study give strong indications that genetically linked homologous copies of a small number of key biosynthetic and regulatory genes play a major role in the accumulation of aliphatic, aromatic and indole glucosinolates in *B. napus* seeds. By identifying gene-linked SSR markers with significant associations to total seed glucosinolate content in genetically diverse oilseed rape germplasm, we hope to provide a simple molecular tool for marker-assisted combination of positive alleles in new, low-glucosinolate genotypes. This has considerable interest for breeding because the markers should enhance the identification of high-glucosinolate accessions carrying desirable alleles that until now have been largely ignored in breeding of 00-quality oilseed rape. Inter-crossing of different high-glucosinolate genotypes that contain complementary marker alleles associated with reduced total glucosinolate content at different gene loci should result in transgressive segregation with the possibility for marker-assisted pyramiding of positive alleles at all major loci. Ultimately this could open the way for the development of new, genetically diverse heterotic pools for hybrid breeding.

Acknowledgments The authors thank Fatih Seyis, Wilfried Lühs, and KWS Saat AG for providing seed material for this study. We also thank Liane Renno and Petra Degen for valuable technical assistance. Some of the SSR primers were kindly provided by Wolfgang Ecke, University of Göttingen, and Jörg Schondelmaier, SaatenUnion Resistenzlabor GmbH, Leopoldshöhe. Genome-wide SSR data for the 46 winter rapeseed genotypes genotypes was made available by KWS Saat, SaatenUnion Resistenzlabor and SW Seed GmbH as part of the BMBF-funded project GABI-BRIDGE. We thank two anonymous reviewers and the subject editor for constructive comments that led to considerable improvements in the manuscript.

References

Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M (2006) Genome-wide association mapping in *Arabidopsis* identifies previously

known flowering time and pathogen resistance genes. PLOS Genet 1:531–539

Bak S, Tax FE, Feldmann KA, Galbraith DW, Feyereisen R (2001) CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in Arabidopsis. Plant Cell 13:101–111

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J, Weigel D (2006) The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*. Nat Genet 38:711–715

Bao JS, Corke H, Sun M (2006) Nucleotide diversity in *starch synthase IIa* and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.). Theor Appl Genet 113:1171–1183

Basunanda P, Spiller TH, Hasan M, Gehringer A, Schondelmaier J, Lühs W, Friedt W, Snowdon RJ (2007) Marker-assisted increase of genetic diversity in a double-low seed quality winter oilseed rape genetic background. Plant Breed 126:581–587

Berg ES, Olaisen B (1994) Hybrid PCR sequencing—sequencing of PCR products using a universal primer. Biotechniques 17:896–901

Breseghello F, Sorrells ME (2006) Association analysis as a strategy for improvement of quantitative traits in plants. Crop Sci 46:1323–1330

Celenza JL, Quiel JA, Smolen GA, Merrikh H, Silvestro AR, Normanly J, Bender J (2005) The *Arabidopsis* ATR1 Myb transcription factor controls indolic glucosinolate homeostasis. Plant Physiol 137:253–262

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Ehrenreich IM, Stafford PA, Purugganan MD (2007) The genetic architecture of shoot branching in *Arabidopsis thaliana*: A comparative assessment of candidate gene associations vs. quantitative trait locus mapping. Genetics 176:1223–1236

Flint-Garcia SA, Thornsberry JM, Buckler ES IV (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schüler K (2004) Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol Breed 13:93–102

Hagenblad J, Nordborg M (2002) Sequence variation and haplotype structure surrounding the flowering time locus *Fr1* in *Arabidopsis thaliana*. Genetics 161:289–298

Hansen CH, Du L, Naur P, Olsen CE, Axelsen KB, Hick AJ, Pickett JA, Halkier BA (2001) CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis*. J Biol Chem 276:24790–24796

Hasan M, Seyis F, Badani AG, Pons-Kuhnemann J, Lühs W, Friedt W, Snowdon RJ (2006) Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. Genet Resour Crop Evol 53:793–802

Hoecker U, Toledo-Ortiz G, Bender J, Quail PH (2004) The photomorphogenesis-related mutant red1 is defective in CYP83B1, a red light-induced gene encoding a cytochrome P450 required for normal auxin homeostasis. Planta 219:195–200

Howell PM, Sharpe AG, Lydiate DJ (2003) Homoeologous loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). Genome 46:454–460

Iwata H, Uga Y, Yoshioka Y, Ebana K, Hayashi T (2007) Bayesian association mapping of multiple quantitative trait loci and its application to the analysis of genetic variation among *Oryza sativa* L. germplasms. Theor Appl Genet 114:1437–1449

Kimber DS, McGregor DI (1995) The species and their origin, cultivation and world production. In: Kimber D, McGregor DI (eds) Brassica oilseeds: production and utilization. CABI Publishing, Wallingford, pp 1–9

Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T (2001) A gene controlling variation in *Arabidopsis thaliana* glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol 127:1077–1088

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S (2005) *Arabidopsis* IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defence. Plant J 43:79–96

Lühs W, Seyis F, Frauen M, Busch H, Frese L, Willner E, Friedt W, Gustafsson M, Poulsen G (2003) Development and evaluation of a *Brassica napus* core collection. In: Knüpffer H, Ochsmann J (eds) Rudolf Mansfeld and Plant Genetic Resources. Proceedings of a symposium dedicated to the 100th birthday of Rudolf Mansfeld, Gatersleben, Germany, 8–9 October 2001. Schriften zu Genetischen Ressourcen 19. ZADI/IBV, Bonn, pp 284–289 (http://www.genres.de/infos/rei-bd22.htm)

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004) The effects of human population structure on large genetic association studies. Nat Genet 36:512–517

Mithen RF, Dekker M, Verkerk R, Rabot S, Johnson IT (2000) The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. J Sci Food Agric 80:967–984

Oesterberg MK, Shavorskaya O, Lascoux M, Lagercrantz U (2002) Naturally occurring indel variation in the *Brassica nigra COL1* gene is associated with variation in flowering time. Genetics 161:299–306

Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171:765–781

Peleman JD, van der Voort JR (2003) Breeding by design. Trends Plant Sci 8:330–334

Poulsen G, Busch H, Frauen M, Frese L, Friedt W, Gustafsson M, Ottosson F, Seyis F, Stemann G, Ulber B, Willner E, Lühs W (2004) The European *Brassica napus* core collection—characterisation, evaluation and establishment. Cruciferae Newsl 25:115–116

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rask L, Andréasson E, Ekbom B, Eriksson S, Pontoppidan B, Meijer J (2000) Myrosinase: gene family evolution and herbivory defense in Brassicaceae. Plant Mol Biol 42:93–113

Röbbelen G (1975) Totale Sortenumstellung bei Körnerraps. Bericht über die Arbeitstagung der Arbeitsgemeinschaft der Saatzuchtleiter, Gumpenstein. Arbeitsgemeinschaft der Saatzuchtleiter, Austria, pp 119–146

Rygulla W, Snowdon RJ, Friedt W, Happstadius I, Cheung WY, Chen D (2008) Identification of quantitative trait loci for resistance against Verticillium longisporum in oilseed rape (*Brassica napus*). Phytopathology 98:215–221

Schuster W (1987) Die Entwicklung des Anbaues und der Züchtung von Ölpflanzen in Mitteleuropa. Fat Sci Technol 89:15–27

Seyis F, Snowdon RJ, Lühs W, Friedt W (2003) Molecular characterisation of novel resynthesised rapeseed (*Brassica napus* L.) lines and analysis of their genetic diversity in comparison to spring rapeseed cultivars. Plant Breed 122:473–478

Sharpe AG, Lydiate DJ (2003) Mapping the mosaic of ancestral genotypes in a cultivar of oilseed rape (*Brassica napus*) selected via pedigree breeding. Genome 46:461–468

Snowdon RJ, Friedt W (2004) Molecular markers in Brassica oilseed breeding: current status and future possibilities. Plant Breed 123:1–8

Textor S, Bartram S, Kroymann J, Falk KL, Hick A, Pickett J.A, Gershenzon J (2004) Biosynthesis of methionine-derived

glucosinolates in *Arabidopsis thaliana*: recombinant expression and characterization of methylthioalkylmalate synthase, the condensing enzyme of the chain elongation cycle. Planta 218:1026–1035

Thompson KF (1983) Breeding winter oilseed rapeseed *Brassica napus*. Adv Appl Biol 7:1–104

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES IV (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

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

Walker KC, Booth EJ (2001) Agricultural aspects of rape and other Brassica products. Eur J Lipid Sci Technol 103:441–446

Whitt SR, Buckler ES IV (2003) Using natural allelic diversity to evaluate gene function. In: Grotewald E (ed) Plant functional genomics: methods and protocols. Humana Press, New York, pp 123–140

Wilson LM, Whitt SR, Ibáñez AM, Rocheford TR, Goodman MM, Buckler ES (2004) Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16:2719–2733

Wittstock U, Halkier BA (2000) Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. catalyzes the conversion of L-phenylalanine to phenylacetaldoxime in the biosynthesis of benzylglucosinolate. J Biol Chem 275:14659–14666

Wittstock U, Halkier BA (2002) Glucosinolate research in the *Arabidopsis* era. Trends Plant Sci 7:263–270

Xu XY, Bai GH, Carver BF, Shaner GE (2005) Mapping of QTLs prolonging the latent period of *Puccinia triticina* infection in wheat. Theor Appl Genet 110:244–251

Yu J, Pressior G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

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

IV DISCUSSION

4.1 Genetic diversity in *Brassica napus* genebank collections

In conservation programs for plant genetic resources, the availability of characterization data and information on available genetic diversity can help germplasm users to identify accessions of interest and also provide plant breeders with initial data regarding materials for use in crop improvement programs (Cruz et al. 2007). Information on the genetic diversity in *B. napus* germplasm collections can give breeders and geneticists important information on the allelic diversity present in genebank materials and may help to identify genetically diverse pools for use in cross combinations to improve important agronomic traits or to better exploit heterosis (Diers and Osborn, 1994).

Predicting heterosis is of utmost importance in hybrid breeding. Although genetic distances can be easily determined by molecular markers, to predict heterosis by the relatedness of the parents is still a difficult challenge (Qian et al. 2007). Diers et al. (1996) and Riaz et al. (2001), respectively, reported a significant correlation between genetic distance and heterosis for seed yield in *Brassica* inbred lines. The diverse spring and winter oilseed genotypes identified in the study described in this thesis may therefore represent a useful resource for improving heterotic potential in spring and winter oilseed rape, respectively. In this regard the highly genetically distinct winter oilseed varieties ‘Mytnickij’, ‘Kromerska’, ‘Mestnij’ and ‘Krapphauser’ represent a potentially valuable resource for winter oilseed rape breeding, whereas ‘Dux’ (Germany) and ‘Wesway’ (Australia) could be of interest for diversifying the spring oilseed rape gene pool.

Although winter oilseed rape germplasm is unsuitable for spring-sown cropping regimes due to its vernalisation requirement, the winter rapeseed gene pool nevertheless represents an untapped source of genetic diversity for spring canola breeding, in order to broaden its genetic base and potentially boost seed yields of spring hybrids (Quijada et al. 2006). In particular,

winter rapeseed may contain superior alleles at some loci that could improve specific traits when introgressed into the adapted germplasm (Udall et al. 2006). Significant yield increases in spring oilseed rape hybrids through introgression of winter germplasm have been reported previously by Butruille et al. (1999), Quijada et al. (2004) and Udall et al. (2006).

On the other hand, genetically diverse vegetable and fodder rape genotypes represent a potentially valuable source for improved pathogen and pest resistance (Lühs et al. 2003a, 2003b), and the same is true for the introduction of resistance genes from the *B. rapa* and *B. oleracea* gene pools via resynthesized *B. napus* (e.g. Rygulla et al. 2007a, 2007b, 2008). Furthermore, introduction of untapped germplasm into breeding lines also has the potential to improve heterotic potential (Seyis et al. 2006, Basunanda et al. 2007, Gehringer et al. 2007). However, despite the availability of genetically diverse *B. napus* germplasm from such sources the construction of distinct genetic pools, as used for example in maize hybrid breeding, has not been achieved for oilseed rape to date. Because of linkage drag for seed yield and quality traits associated with non-oilseed rape morphotypes, identification of exotic germplasm amongst the respective gene pools of winter and spring oilseed forms is of particular interest in this respect. In this study a considerable genetic variation was identified in *B. napus* vegetable and fodder rape genotypes compared to the gene pools of conventional spring and winter oilseed material (Hasan et al. 2006). Similar extreme genetic variation in resynthesized rapeseed compared to conventional rapeseed cultivars was reported previously by Becker et al. (1995) and Seyis et al. (2003b). In the latter study, the genetic differences were correlated to heterotic yield potential in experimental hybrids (Seyis et al. 2006). Similarly, Udall et al. (2004) also reported that introgression of alleles from resynthesized *B. napus* into a spring canola hybrid combination can improve the seed yield. Such exotic material must be viewed from a long-term perspective with regard to use in oilseed rape breeding, however genetic diversity analyses using molecular markers have the potential to identify novel genetic variation that might assist in future improvement of heterotic potential in *B. napus.*

4.2 Molecular characterization of the *B. napus* core collection

Chapter 2 of this thesis (Hasan et al. 2006) describes the molecular genetic characterization of a *B. napus* core collection and the use of the data to describe the molecular genetic diversity.

The intention of core collections is to represent the genetic diversity of a species as broadly as possible within a limited number of gene bank accessions that contain a maximum degree of allelic diversity. Such genotype sets are particularly attractive for association analysis of qualitative traits, such as disease resistance or special quality characteristics (Breseghello and Sorrells 2006). To investigate the phenotypic variation present in *B. napus* genebank collections and generate the above mentioned *B. napus* core collection, Lühs et al. (2003a) evaluated a set of around 1500 oilseed, fodder and vegetable rape accessions, representing the *B. napus* material present in European gene banks, in field trials. Variation in the materials was described for a number of descriptive characters like morphotype, emergence date, vernalisation requirement and winter hardiness, flowering time, flowering duration and seed quality traits. Based on the data collected and available pedigree knowledge, a core collection of 150 genotypes was established for more detailed agronomic investigations to identify germplasm of interest for oilseed rape breeding. Descriptor data on the *B. napus* core collection (along with other *Brassica* species investigated in the same project), are available from http://documents.plant.wur.nl/cgn/pgr/brasedb/brasresgen.htm.

The *B. napus* core collection studied here is intended to represent as broadly as possible the genetic diversity present in the species as a whole (Poulsen et al. 2004). When used together with the molecular marker data collected in the study described in this thesis (Hasan et al. 2006), phenotypic data from field trials with the core collection will represent a useful resource for further association studies on specific traits using relevant candidate genes. Detailed allelic information on mapped SSR loci, in combination with candidate gene association studies based on linkage disequilibrium in trait-relevant genome regions, has the potential to identify genes that contribute to agronomically important traits in exotic *B. napus* germplasm. Ultimately this will assist in the development of molecular markers for marker-assisted transfer of these traits from exotic genotypes into elite breeding material. The genome-wide SSR markers used in the present study for population structure inference could also be used to speed up the recovery of the recurrent parent genome in backcrossing programs (Breseghello and Sorrells 2006).

4.3 Association mapping in *Brassica napus*

Plant genetics has an important and challenging goal of identifying the genetic variants that underlie complex traits (Agrama et al. 2007), and in this context association mapping has emerged as a powerful method for generating genetic information from population samples (Marcano et al. 2007). The basic objective of association mapping studies is to detect correlations between genotypes and phenotypes in a sample of genetically diverse individuals (e.g. a genotype core collection) on the basis of linkage disequilibrium (LD; Zondervan and Cardon 2004). The study outlined in Chapter 3 of this thesis (Hasan et al. 2008) represents the first application of association mapping approach to investigate a quantitative trait in oilseed rape and demonstrates that the *B. napus* core collection of Lühs et al. (2003a, 2003b) can be effectively used to identify useful allelic variation for an important seed quality trait in oilseed rape.

Association genetics approaches are vital for the exploitation of the phenotypic diversity present in genotype core collections, because they allow us to sort through this diversity in order to find the alleles and polymorphisms that are beneficial for crop improvement (Buckler and Thornsberry 2002). The ability to interpret patterns of molecular genetic diversity and to relate them to phenotypic variation makes it possible to correlate the genetic diversity with phenotypic variation and allow the identification of actual genes underlying quantitative traits (Garris et al. 2003). Allele-trait association mapping approaches (LD mapping) detect and locate quantitative trait loci (QTL) based on the strength of the correlation between genetic markers and trait phenotypes. The key advantages of association tests in well-described, genetically diverse populations include their speed, because no dedicated, segregating mapping population need be created for a specific trait, as is the case in classical QTL mapping. Furthermore, if high-density mapped marker data is available then the resolution of association mapping approaches can be much higher than that of traditional QTL mapping (Buckler and Thornsberry 2002).

The primary obstacle to successful association mapping is the presence of population structure, particularly the presence of subgroups with an unequal distribution of alleles, which can mimic the signal of association and lead to false positives or missed real effects (Pritchard et al. 2000b, Marchini et al. 2004, Simko et al. 2004, Iwata et al. 2007). This is an important aspect to consider in association mapping in *B. napus*, since the species shows a distinct

population structure divided into the different cultivated forms (Hasan et al. 2006). Thus, a crucial first step in LD mapping is to define a set of subpopulations. To control for this type of structure, Pritchard et al. (2000a) developed a statistical approach that assigns membership to various subpopulations by determining the amount of genotypic correlation based on unlinked random markers. Association mapping without consideration of population structure would have a high rate of Type I error (false positive) because of spurious associations between non-linked loci (Agrama et al. 2007). Pritchard et al. (2000b) developed a test for association in a case control study that is valid in the presence of population structure. To account for population structure in the present study (Hasan et al. 2008) a modification of the Pritchard approach was used that was developed by Thornsberry et al. (2001) in order to deal with quantitative traits. Thornsberry et al. (2001) previously used this modified method to study a candidate gene for the control of flowering time in structured maize populations.

Association mapping requires a large amount of genotype (marker) data to detect regions associated with the phenotype of interest. However this limitation can be overcome by using a small number of genes, or defined genome regions, that are already suspected of being involved in the trait of interest. In this case a subset of markers is used to report on neighboring linked markers within the genome, and to investigate allelic variation within the gene sequences in a genetically diverse population (Rafalski 2002, Smith and O'Brien 2005). In the present study (Hasan et al. 2008), a number of new *Brassica* SSR sequences with putative close physical linkage to four candidate genes for seed glucosinolate content were identified through *in silico* screening of neighbouring *Brassica* genomic and EST sequences. Hom(oe)ologous markers for many of these SSR primers showed significant associations with seed glucosinolate content when screened in *B. napus* gene bank accessions. In addition, whole-genome association analyses were performed with SSR markers dispersed throughout the *B. napus* genome. This approach also led to the identification of numerous markers with significant associations to glucosinolate content. In some cases the markers were mapped in available *B. napus* genetic maps, and numerous markers showed no apparent relationship to known QTL regions for seed glucosinolate content. This indicates the presence of novel allelic variation for this important trait, which should be of considerable interest for breeding purposes. The orthologous sequences in *A. thaliana* for two of the genome-wide SSR markers were found to be closely physically linked to a promising candidate gene for glucosinolate biosynthesis, *IQD1*, which had not been previously considered as a candidate for seed glucosinolate reduction in oilseed rape. As described by Frisch and Melchinger (2005) such

markers with significant association to the trait of interest can be useful for introgression of this novel allelic variation from genebank accessions into elite germplasm through marker-assisted backcrossing.

The genome-wide SSR marker data and population structure information described in this work (Hasan et al. 2006, 2008) provide a useful starting point for structure-based association analyses of other phenotypic traits in this *B. napus* core collection. Furthermore, a new procedure was developed and successfully applied for the identification of potentially gene-linked SSR markers based on *Brassica*-Arabidopsis comparative genome analysis (Hasan et al. 2008). In future this method could be applied for marker development, allele-trait association studies and marker-assisted selection for many important quantitative traits in *B. napus* and other *Brassica* crop species.

4.4 Reduction of oilseed rape seed glucosinolate content

Successful commercial utilization of the meal by-product of *Brassica* oilseed crops requires the cultivation of cultivars with low glucosinolate content in the seeds (Márquez-Lema et al. 2006), and in oilseed rape this has been achieved through stringent selection in breeding programs (Mithen et al. 2000) All low-glucosinolate cultivars available today are derived from the introgression of genes from the low seed glucosinolate cv. ‘Bronowski’, initially into Canadian spring rape cultivars and then European winter rape cultivars (Mithen et al. 2000). Hence ‘Bronowski’ is the primary donor conferring low glucosinolate alleles in all modern *B. napus* varieties (Kondra and Stefansson 1970, Ma and Li 2007). However, residual segments of the 'Bronowski' genotype in modern cultivars are believed to cause reductions in yield, winter hardiness and oil content (Sharpe and Lydiate 2003), which suggests that some undesirable alleles for these traits are linked in repulsion phase to some canola quality alleles (Quijada et al. 2006). It is known that high-glucosinolate rapeseed genotypes often carry low-glucosinolate alleles at one or more of the major QTL controlling seed glucosinolate accumulation (Howell et al. 2003 Basunanda et al. 2007). With effective molecular markers for marker-assisted selection these genotypes could be used to introduce new genetic variation for low seed glucosinolate content into breeding programs. However, the introgression of novel genetic diversity from ++ rapeseed normally requires extensive backcrossing to eliminate all unwanted alleles contributing to high seed erucic acid and glucosinolate

contents. With each recurrent backcross the genomic contribution of the genetically diverse ++ parent is reduced, and with it the proportion of novel alleles that can potentially contribute to heterosis in hybrids derived from such lines. *B. napus* genotypes containing high levels of erucic acid and seed glucosinolates (so-called ++ seed quality) represent a comparatively genetically divergent source of germplasm for the generation of new genetic pools for hybrid breeding (Röbbelen 1975, Thompson 1983, Schuster 1987). The transfer of these latent low glucosinolate alleles (together with the flanking chromosomal regions) to modern low glucosinolate varieties might reduce the effect of the hypothesized `Bronowski` bottleneck (Howell et al. 2003.) In the present study (Hasan et al. 2008), a total of 51 marker alleles from 29 SSR primer combinations were found to exhibit a significant association ($P \leq 0.05$) to total seed glucosinolate content in a core collection of 94 *B. napus* gene bank accessions. Ten of the markers also exhibited significant association using rescaled *P*-values, indicating that these associations are not likely to be caused by type I errors. The trait associations found in candidate gene-linked markers indicate that genetically linked homologous copies of a small number of biosynthetic and regulatory genes play a major role in the accumulation of aliphatic, aromatic and indole glucosinolates in *B. napus* seeds. These gene-linked SSR markers, with significant associations to total seed glucosinolate content in genetically diverse oilseed rape germplasm, provide a simple molecular tool for marker-assisted combination of positive alleles in new, low-glucosinolate genotypes. This has considerable interest for breeding because the markers should enhance the identification of high-glucosinolate accessions carrying desirable alleles that until now have been largely ignored in breeding of 00-quality oilseed rape. Inter-crossing of different high-glucosinolate genotypes that contain complementary marker alleles associated with reduced total glucosinolate content at different gene loci should result in transgressive segregation with the possibility for marker-assisted pyramiding of positive alleles at all major loci. Ultimately this could open the way for the development of new, genetically-diverse heterotic pools for hybrid breeding.

V SUMMARY

Today oilseed rape (*B. napus* ssp. *napus*) is the most important source of vegetable oil in Europe and the second most important oilseed crop in the world after soybean. However, breeding of oilseed rape has evoked a strong bottleneck selection towards double-low (00, canola) seed quality with zero erucic acid and low seed glucosinolate content. Genetic diversity throughout the rapeseed primary gene pool was examined by obtaining detailed molecular genetic information at simple sequence repeat (SSR) loci for a broad range of winter and spring oilseed, fodder and leaf rape gene bank accessions. A set of 96 genotypes was characterised using publicly available mapped SSR markers spread over the *B. napus* genome. Allelic information from 30 SSR primer combinations amplifying 220 alleles at 51 polymorphic loci provided unique genetic fingerprints for all genotypes. UPGMA clustering enabled the identification of four general groups with increasing genetic diversity as follows (1) spring oilseed and fodder; (2) winter oilseed; (3) winter fodder; (4) vegetable genotypes. The diverse spring and winter oilseed genotypes identified in the study may represent a useful resource for improving heterotic potential in spring and winter oilseed rape, respectively. In this regard the highly genetically distinct winter oilseed varieties ‘Mytnickij’, ‘Kromerska’, ‘Mestnij’ and ‘Krapphauser’ represent a potentially valuable resource for winter oilseed rape breeding, whereas ‘Dux’ (Germany) and ‘Wesway’ (Australia) could be of interest for diversifying the spring oilseed rape gene pool. Moreover a considerable genetic variation was identified in *B. napus* vegetable and fodder rape genotypes compared to the gene pools of conventional spring and winter oilseed material. The molecular genetic information gained enables the identification of untapped genetic variability for rapeseed breeding and is potentially interesting with respect to increasing heterosis in oilseed rape hybrids. The results of this study demonstrate the suitability of SSR data for analysis of genetic diversity in *B. napus* genotypes.

Using this set of genetically diverse genotypes, structure-based allele-trait association studies were conducted to identify potentially gene-linked markers for important seed glucosinolate loci. The analyses included a set of new SSR markers whose orthologs in *Arabidopsis thaliana* are physically closely linked to promising candidate genes for glucosinolate

biosynthesis. Using 62 polymorphic SSR primer combinations, a total of 348 polymorphic SSR marker alleles were amplified in the 94 gene bank accessions. Associations between the marker data and the total seed glucosinolate content were tested using the logistic regression approach. A total of 51 marker alleles from 29 SSR primer combinations were found to exhibit a significant association ($P \leq 0.05$) to total seed glucosinolate content. The results indicate that four genes involved in the biosynthesis of indole, aliphatic and aromatic glucosinolates might be associated with known QTL for total seed glucosinolate content in *B. napus*. Markers linked to homoeologous loci of these genes in the paleopolyploid *B. napus* genome were found to be associated with a significant effect on the seed glucosinolate content. Inter-crossing of different high-glucosinolate genotypes that contain complementary marker alleles associated with reduced total glucosinolate content at different gene loci should result in transgressive segregation with the possibility for marker-assisted pyramiding of positive alleles at all major loci. This could open the way for the development of new, genetically-diverse heterotic pools for hybrid breeding.

Genome-wide SSR marker data and population structure information were generated that provide a useful starting point for structure-based association analyses of other phenotypic traits in this *B. napus* core collection. Furthermore, a new procedure was developed and successfully applied for the identification of potentially gene-linked SSR markers based on *Brassica*-Arabidopsis comparative genome analysis. In the future, this method could be applied for marker development, allele-trait association studies and marker-assisted selection for numerous important quantitative traits in *B. napus* and other *Brassica* crop species.

VI ZUSAMMENFASSUNG

Raps (*B. napus* ssp. *napus*) stellt die wichtigste Ölsaat in Europa und nach der Sojabohne die zweitwichtigste Ölsaat weltweit dar. Die Züchtung von Qualitätsraps, d.h. 00-Sorten weitgehend ohne Erucasäure (C22:1) und mit einem geringen Glucosinolat-Gehalt, haben im aktuellen Zuchtmaterial zu einem „Flaschenhals-Effekt“, d.h. zu einer gewissen Einschränkung der Diversität, geführt. Aus diesem Grunde sollte in der vorliegenden Arbeit die genetische Diversität innerhalb des primären Raps-Genpools anhand eines breiten Sortimentes von Winter- und Sommerraps, Futterraps und Genbankakzessionen mithilfe von Mikrosatelliten (*simple-sequence-repeats*, SSR) untersucht werden, um detaillierte genetische Informationen zu den betreffenden genetischen Loci zu erhalten. Mittels öffentlich verfügbarer SSR-Marker, die das gesamte Genom von *B. napus* abdecken, wurde ein Set von 96 Rapsgenotypen charakterisiert. Insgesamt 30 SSR Primerkombinationen mit 220 Allelen, von denen 51 polymorph waren, lieferten allelische Informationen, die die Darstellung eines individuellen genetischen Fingerabdrucks für jeden einzelnen Genotyp ermöglichten.

Mit diesen Daten konnten anhand einer Cluster-Analyse (UPGMA-Cluster) vier Hauptgruppen mit einer erhöhten genetischen Diversität zueinander identifiziert werden: (1) Sommerraps und Futterraps; (2) Winterraps; (3) Winterfutterraps und (4) Gemüseraps. Die identifizierten Winter- und Sommerraps-Genotypen repräsentieren eine nützliche Ressource für eine verbesserte Nutzung der Heterosis (Bastardwüchsigkeit) in dem Zuchtmaterial. In diesem Zusammenhang stellen die genetisch stark diversen Winterrapssorten ‘Mytnickij’, ‘Kromerska’, ‘Mestnij’ und ‘Krapphauser’ eine potentielle, wertvolle Quelle für die Winterrapszüchtung dar. Dagegen dürften die Sommerrapssorten ‘Dux’ (Deutschland) und ‘Wesway’ (Australien) sehr nützlich sein, um die genetische Diversität innerhalb des Genpools als Voraussetzung für eine effektive Sommerrapszüchtung zu steigern. Weiterhin wurde im Vergleich zu den Genpools konventioneller Sommer- und Winterrapsformen in den Gemüseraps- und Futterrapsgenotypen eine beachtliche genetische Variation identifiziert. Die molekulargenetischen Daten ermöglichten die Identifizierung bisher nicht genutzter genetischer Variabilität für die Rapszüchtung, was in Bezug auf die Züchtung von Hybridsorten mit einem gesteigerten Heterosiseffekt von großer Bedeutung ist. Die

Ergebnisse dieser Arbeit demonstrieren, dass SSR-Marker nützliche Werkzeuge für die Rapszüchtung sind, insbesondere um die genetische Diversität in *B. napus*-Genotypen aufzudecken.

Des weiteren wurden die 96 Genotypen genutzt, um für prominente Glucosinolat-Loci genomstrukturbasierte Allel-Merkmal Assoziationsstudien durchzuführen, die zur Identifizierung von gekoppelten Markern dienten. Diese Analysen wurden mit einem neuen Set aus SSR-Markern durchgeführt, bei denen die orthologen Sequenzen in *Arabidopsis thaliana* eng mit aussichtsreichen Kandidatengenen für die Glucosinolat-Biosynthese gekoppelt sind. Mittels 62 polymorpher SSR-Primerkombinationen konnten in 94 Genotypen insgesamt 348 Allele amplifiziert werden. Die Assoziationen zwischen Alleldaten und dem Glucosinolatgehalt im Samen wurden mit dem *logistic regression* Ansatz ermittelt. Insgesamt zeigten sich 29 Primerkombinationen mit dem Gehalt an Glucosinolaten (GSL) als signifikant assoziiert ($p \leq 0.05$). Die Ergebnisse deuten darauf hin, dass 4 Gene, die in die Biosynthese von indolischen, aliphatischen und aromatischen Glucosinolaten involviert sind, mit bekannten QTLs für den Glucosinolat-Gehalt in *B. napus* assoziiert sind. Marker, die mit homoeologen Loci dieser Gene in dem paleopolyploiden Organismus *B. napus* gekoppelt sind, waren signifikant mit dem Glucosinolat-Gehalt im Samen assoziiert. Mithin sollten Kreuzungen zwischen Genotypen mit hohem bzw. niedrigem Glucosinolat-Gehalt, die gesichert assoziierte komplementäre Markerallele besitzen, in einer transgressiven Segregation resultieren und damit eine Marker-basierte Pyramidisierung von positiven Allelen an relevanten GSL-Loci ermöglichen. Dies eröffnet einen neuen Weg für die Entwicklung von genetisch diversen, heterotischen Pools für die Hybridzüchtung.

Insgesamt wurden im Rahmen der vorliegenden Arbeit genomweite SSR-Markerdaten und Informationen über Populationsstrukturen ermittelt, die einen nützlichen Ansatz für weitere strukturbasierte Assoziationsanalysen mit phänotypischen Merkmalen innerhalb des hier genutzten *B. napus* Coresets darstellen. Weiterhin wurde eine neuartige Herangehensweise entwickelt und erfolgreich angewendet, um durch eine vergleichende Genomanalyse zwischen *Brassica* und Arabidopsis potentielle, gekoppelte Marker zu identifizieren. In Zukunft kann diese Methode in *B. napus* und anderen *Brassica*-Spezies für die Markerentwicklung, für Assoziationsstudien und die markerbasierte Selektion auf zahlreiche Merkmale genutzt werden.

VII REFERENCES

Agrama HA, Eizenga GC, Yan W (2007) Association mapping of yield and its components in rice cultivars. Mol. Breed. 19: 341-356

Bachtrog D, Weiss S, Zangerl B, Brem G, Schlötterer C (1999) Distribution of dinucleotide microsatellites in the *Drosophila melanogaster genome.* Mol. Biol. and Evol. 16: 602-610

Basunanda P, Spiller TH, Hasan M, Gehringer A, Schondelmaier J, Lühs W, Friedt W, Snowdon RJ (2007) Marker-assisted increase of genetic diversity in a double-low seed quality winter oilseed rape genetic background. Plant Breed. 126: 581-587

Becker HC, Engqvist GM, and B, Karlsson B (1995) Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. Theor. Appl. Genet. 91: 62-67

Bones AM, Rossiter JT (1996) The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol. Plant. 97: 194-208

Brandle JE, McVetty PBE (1989) Heterosis and combining ability in hybrids derived from oilseed rape cultivars and inbred lines. Crop Sci. 29:1191-1195

Breseghello F, Sorrells ME (2006) Association analysis as a strategy for improvement of quantitative traits in plants. Crop Sci. 46: 1323-1330

Buckler ES, Thornsberry JM (2002) Plant molecular diversity and applications to genomics. Curr. Opin. Plant Biol. 5: 107-111

Burel C, Boujard T, Escaffre AM, Kaushik SJ, Boeuf G, Mol KA, Van der Geyten S, Kuhn ER (2000) Dietary low glucosinolate rapeseed meal affect thyroid status and nutrient utilization in rainbow trout (*Oncorhynchus mykiss*). Brit. J. Nutr. 83: 653-664

Burgess B, Mountford H, Hopkins CJ, Love C, Ling A, Spangenberg GC, Edwards D, and Batley J (2006) Identification and characterization of simple sequence repeat (SSR) markers derived *in silico* from *Brassica oleracea* genome shotgun sequences. Mol. Ecol. Notes. 6: 1191-1194

Butruille DV, Guries RP, Osborn TC (1999) Increasing yield of spring oilseed rape hybrids through introgression of winter germplasm. Crop Sci. 39: 1491-1496

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

Cruz VMV, Luhman R, Marek LF, Rife CL, Shoemaker RC, Brummer EC, Gardner CAC (2007) Characterization of flowering time and SSR marker analysis of spring and winter type *Brassica napus* L. germplasm. Euphytica. 153: 43-57

Derycke G, Mabon N, Mandiki SNM, Bister JL, Wathelet J-P, Marlier M, Paquay R (1999) Chemical changes and influence of rapeseed antinutritional factor on lamb physiology and performance. Part 1. Animal performance and thyroid physiology. Anim. Feed Sci. Technol. 81: 81-91

Diers BW, McVetty PBE, Osborn TC (1996) Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.). Crop Sci. 36: 79-83

Diers BW, Osborn TC (1994) Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphisms. Theor. Appl. Genet. 88: 662-668

Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56:5-51

Fernie AR, Tadmor Y, Zamir D (2006) Natural genetic variation for improving crop quality. Curr. Opin. Plant Biol. 9: 196-202

Flint-Garcia SA, Thornsberry JM. Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54: 357-374

Francia E, Tacconi G, Crosatti C, Barabaschi D, Bulgarelli D, Dall'Aglio E, Valè G (2005) Marker assisted selection in crop plants. Plant Cell Tissue Organ Cult. 82: 317-342

Friedt W, Snowdon R, Ordon F, Ahlemeyer J (2007) Plant Breeding: assessment of Genetic Diversity in Crop Plants and its Exploitation in Breeding. Progress in Botany. 168: 152-177 or 178)

Frisch M, Melchinger AE (2005) Selection theory for marker-assisted backcrossing. Genetics. 170: 909-917

Garris A, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa5* locus of rice (*Oryza sativa* L.). Genetics. 165: 759-769

Gehringer A, Spiller T, Basunanda P, Snowdon R, Friedt W (2007) New oilseed rape (*Brassica napus*) hybrids with high levels of heterosis for seed yield under nutrient-poor conditions. Breeding Sci. 57: 315-320

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. Crop Sci. 34: 1199-1205

Grant I, Beversdorf WE (1985) Heterosis and combining ability estimates in spring-planted oilseed rape (*Brassica napus* L.). Can. J. Gen. Cytol. 27: 472-478.

Grubb CD, Zipp BJ, Ludwig-Müller J, Masuno MN, Molinski TF, Abel S (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J. 40: 893-908

Halldén C, Nilsson N-O, Rading IM, Säll T (1994) Evaluation of RFLP and RAPD markers in comparison of *Brassica napus* breeding lines. Theor. Appl. Genet. 88: 123-128

Hansen M, Kraft T, Ganestam S, Säll T, Nilsson NO (2001) Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. Genet. Res. 77: 61-66

Hasan M, Friedt W, Pons-Kühnemann J, Freitag NM, Link K, Snowdon RJ (2008) Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). Theor. Appl. Genet. 116:1035–1049

Hasan M, Seyis F, Badani AG, Pons-Kuhnemann J, Lühs W, F riedt W, Snowdon RJ (2006) Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. Genet. Res. Crop. Evol. 53: 793-802

Howell PM, Sharpe AG, Lydiate DJ (2003) Homoeologous loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). Genome. 46: 454-60

Iwata H, Uga Y, Yoshioka Y, Ebana K, Hayashi T (2007) Bayesian association mapping of multiple quantitative trait loci and its application to the analysis of genetic variation among *Oryza sativa* L. germplasms. Theor. Appl. Genet. 114: 1437-1449

Jordan DR, Tao YZ, Godwin ID, Henzell RG, Cooper M, McIntyre CL (1998). Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. Euphytica 102: 1-7

Keck AS, Finley JW (2004) Cruciferous vegetables: cancer protective mechanisms of glucosinolate hydrolysis products and selenium. Integr. Cancer Ther. 3: 5-12

Kermanshahi H, Abbasi Pour AR (2006) Replacement value of soybean meal with rapeseed meal supplemented with or without a dietary NSP-degrading enzyme on performance, carcass traits and thyroid hormones of broiler chickens. Int. J. Poult. Sci. 5: 925-930

Kimber DS, McGregor DI (1995) The species and their origin, cultivation and world production. In: Kimber D. and McGregor D.I. (eds.), Brassica Oilseeds: Production and Utilization. CABI Publishing, Wallingford, UK. pp. 1-9

Kliebenstein DJ, Gershenzon J, Mitchell-Olds T (2001) Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. Genetics 159: 359-370

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (*Brassicaceae*). Mol. Biol. Evol. 17: 1483-1498

Kondra ZP, Stefansson BR (1970) Inheritance of the major glucosinolates of rapeseed (*Brassica napus*) meal. Can. J. Plant Sci. 50: 643-647

Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T (2001) A gene controlling variation in *Arabidopsis* glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol. 127:1077-1088

Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc. Natl. Acad. Sci. USA. 95: 10774-10778

Lefort-Buson M, Guillot-Lemoine B, Dattée Y (1987) Heterosis and genetic distance in rapeseed (*Brassica napus* L.): crosses between European and Asiatic selfed lines. Genome. 29: 413-418

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11: 2453-65

Li G, Quiros CF (2002) Genetic Analysis, Expression and Molecular Characterization of *BoGSL-ELONG*, a Major Gene Involved in the Aliphatic Glucosinolate Pathway of Brassica Species. Genetics. 162: 1937-1943

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S (2005) Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defence. Plant Journal 43: 79-96

Lombard V, Baril CP, Dubreuil P, Blouet F, Zhang D (2000) Genetic relationships and fingerprinting of rapeseed cultivars using AFLP: consequences for varietal registration. Crop Sci. 40: 1417-1425

Lühs W, Seyis F, Frauen M, Busch H, Frese L, Willner E, Friedt W, Gustafsson M, Poulsen G (2003a) Development and evaluation of a *Brassica napus* core collection. In: Knüpffer H, Ochsmann J (eds) Rudolf Mansfeld and Plant Genetic Resources. Proceedings of a symposium dedicated to the 100th birthday of Rudolf Mansfeld, Gatersleben, Germany, 8-9 October 2001. Schriften zu Genetischen Ressourcen. ZADI/IBV, Bonn, 19: 284-289 (http://www.genres.de/infos/rei-bd22.htm)

Lühs W, Seyis F, Frauen M, Ulber B, Stemann G, Busch H, Friedt W, Gustafsson M, Poulsen G (2003b) Evaluation of a *Brassica napus* core collection and prospects of its exploitation through rapeseed breeding. In: Proc 11th Int Rapeseed Congress, Copenhagen. 2: 371-374

Ma N, Li Z-y (2007) Development of Novel *Brassica napus* lines with canola quality and higher levels of oleic and linoleic acids derived from intergeneric hybrids between *B. napus* and *Orychophragmus violaceus*. Euphytica 157: 231-238

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops Trends in Plant Science. 12: 57-63

Mandiki SNM, Derycke G, Bister JL, Mabon N, Wathelet JP, Marlier M, Paquay R (2002) Chemical changes and influence of rapeseed antinutritional factor on gestating and lactating ewes. Part 1. Animal performances and plasma hormones and glucose. Anim. Feed Sci. Technol. 98: 25-35

Marcano M, Pugh T, Cros E, Morales S, Páez EAP, Courtois B, Glaszmann JC, Engels JMM, Phillips W, Astorga C, Risterucci AM, Fouet O, González V, Rosenberg K, Vallat I, Dagert M, Lanaud C (2007) Adding value to cocoa (*Theobroma cacao* L.) germplasm information with domestication history and admixture mapping. Theor. Appl. Genet. 114: 877-884

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004) The effects of human population structure on large genetic association studies. Nat. Genet. 36: 512-517

Maroufyan E, Kermanshahi H (2006) Effect of different levels of rapeseed meal supplemented with calcium iodate on performance, some carcass traits, and thyroid hormones of broiler chickens. Inter. J. Poul. Science. 5: 1073-1078

Márquez-Lema A, Fernández-Martínez JM, Pérez-Vich B, Velasco L (2006) Transgressive segregation for reduced glucosinolate content in *Brassica carinata* A. Braun. Plant Breeding. 125: 400-402

Mawson R, Heaney RK, Zdunczyk Z, Kozlowska H (1994a) Rapeseed meal-Glucosinolates and their antinutritional effects. Part 4. Goitrogenicity and internal organs abnormalities in animals. Die Nahrung. 38: 178-191

Mawson R, Heaney RK, Zdunczyk Z, Kozlowska H (1994b) Rapeseed meal-glucosinolates and their antinutritional effects. Part 5. Animal reproduction. Die Nahrung. 38: 588-598

McCouch S (2004) Diversifying selection in plant breeding. PLoS Biol. 2:e347

McNeill L, Bernard K, MacLeod MG (2004) Food intake, growth rate, food conversion and food choice in broilers fed on diets high in rapeseed meal and pea meal with observations of the resulting poultry meat. Brit. Poultry Sci. 45: 519-523

Meuwissen THE, Goddard ME (2000) Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci, Genetics 155: 421-430

Mithen RF, Dekker M, Verkerk R, Rabot S, Johnson IT (2000) The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. Journal of the Science of Food and Agriculture. 80: 967-984

O'Neill R, Snowdon RJ, Köhler W (2003) Population genetics aspects of biodiversity. Prog. Botany. 64: 115-137

Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics. 171: 765-781

Peleman JD, van der Voort JR (2003) Breeding by design. Trends Plant Sci. 8: 330-334

Plieske J, Struss D (2001) Microsatellite markers for genome analysis in Brassica. I. Development in *Brassica napus* and abundance in Brassicaceae species. Theor. Appl. Genet. 102: 689-694

Poulsen G, Busch H, Frauen M, Frese L, Friedt W, Gustafsson M, Ottosson F, Seyis F. Stemann G. Ulber B. Willner E. Lühs W (2004) The European *Brassica napus* core collection – Characterisation, evaluation and establishment. Cruciferae Newsl. 25: 115-116

Pritchard JK, Stephens M, Donnelly P (2000a) Inference of population structure using multilocus genotype data. Genetics. 155: 945-959

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b) Association mapping in structured populations. Am. J. Hum. Genet. 67: 170-181

Qian W, Sass O, Meng J, Li M, Frauen M, Jung C (2007) Heterotic patterns in rapeseed (*Brassica napus* L.): I. Crosses between spring and Chinese semi-winter lines. Theor. Appl. Genet. 115: 27-34

Quijada PA, Udall JA, Lambert B, Osborn TC (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 1. Identification of genomic regions from winter germplasm. Theor Appl Genet. 113: 549-561.

Quijada PA, Udall JA, Polewicz H, Vogelzang RD, Osborn TC (2004) Phenotypic effects of introgressing French winter germplasm into hybrid spring canola. Crop Sci. 44: 1982-1989

Rafalski JA (2002) Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci. 162: 329-333

Rao VR, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell Tissue Organ Cult. 68: 1-19

Rask L, Andréasson E, Ekbom B, Eriksson S, Pontoppidan B, Meijer J (2000) Myrosinase: gene family evolution and herbivore defense in Brassicaceae. Plant Mol. Biol. 42: 93-113

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. USA. 98: 11479-11484

Riaz A, Li G, Quresh Z, Swati MS, Quiros CF (2001) Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breed. 120: 411-415

Robinson AJ, Love CG, Batley J, Barker G, Edwards D (2004) Simple sequence repeat marker loci discovery using SSR primer. Bioinformatics. 20: 1475-1476

Röbbelen G (1975) Totale Sortenumstellung bei Körnerraps. Bericht über die Arbeitstagung der Arbeitsgemeinschaft der Saatzuchtleiter, Gumpenstein. Arbeitsgemeinschaft der Saatzuchtleiter, Austria. pp. 119-146

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90: 43-48

Rygulla W, Friedt W, Seyis F, Lühs W, Eynck C, von Tiedemann A, Snowdon RJ (2007a) Combination of resistance to *Verticillium longisporum* from zero erucic acid Brassica oleracea and oilseed Brassica rapa genotypes in resynthesized rapeseed (*Brassica napus*) lines. Plant Breed. 126: 596-602

Rygulla W, Snowdon RJ, Eynck C, Koopmann B, von Tiedemann A, Lühs W, Friedt W (2007b) Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridization. Phytopathology. 97: 1391-1396

Rygulla W, Snowdon RJ, Friedt W, Happstadius I, Cheung WY, Chen D (2008) Identification of quantitative trait loci for resistance against *Verticillium longisporum* in oilseed rape (*Brassica napus*). Phytopathology. 98: 215-221

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nat. Rev. Genet. 3: 429-441

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma. 109: 365-371

Schöne F, Groppel B, Hennig A, Jahreis G (1997). Rapeseed meals, methimazole, thiocyanate and iodine affect growth and thyroid. Investigations into glucosinolate tolerance in the pig. J. Sci. Food Agric. 74: 69-80

Schuster W (1969) Vergleich von zwei Zuchtverfahren in der Erhaltungszüchtung bei Winterraps. Zeitschrift für Pflanzenzüchtung 62:47-62

Schuster W (1987) Die Entwicklung des Anbaues und der Züchtung von Ölpflanzen in Mitteleuropa. Fat Sci. Technol. 89: 15-27

Seyis F, Friedt W, Lühs W (2003a) Resynthesised *Brassica napus* as a genetic resource in rapeseed improvement for quality and agronomic performance. In:Knüpffer H,Ochsmann J (eds)RudolfMansfeld and PlantGeneticResources. Schriften zu Genetischen Ressources, ZADI/IBV, Bonn. 19: 334-338

Seyis F, Friedt W, Lühs W (2006) Yield of *Brassica napus* L. hybrids developed using resynthesised rapeseed material. Field Crops Res. 96: 176-180.

Seyis F, Snowdon RJ, Lühs W, Friedt W (2003b) Molecular characterisation of novel resynthesised rapeseed (*Brassica napus* L.) lines and analysis of their genetic diversity in comparison to spring rapeseed cultivars. Plant Breed. 122: 473-478

Sharpe AG, Lydiate DJ (2003) Mapping the mosaic of ancestral genotypes in a cultivar of oilseed rape. *Brassica napus*. selected via pedigree breeding. Genome 46: 461-468.

Shengwu H, Ovesná J, Kučera L. Kučera V. Vyvadilová M (2003) Evaluation of genetic diversity of *Brassica napus* germplasm from China and Europe assessed by RAPD markers. Plant Soil Environ. 49: 106-113

Sia EA, Jinks-Robertson S, Petes TD (1997) Genetic control of microsatellite stability. Mutat. Res. 383: 61-70

Simko L, Costanzo S, Haynes KG, Christ BJ, Jones RW (2004) Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor. Appl. Genet. 108: 217-224

Smith MW, O'Brien SJ (2005) Mapping by admixture linkage disequilibrium: advances, limitations, and guidelines. Nat. Rev. Genet. 6: 623-632.

Snowdon RJ, Friedt W (2004) Molecular markers in Brassica oilseed breeding: current status and future possibilities. Plant Breeding 123:1-8

Snowdon, RJ, W Lühs, W Friedt (2006) Oilseed rape. In: Kole C.(ed) Genome Mapping and Molecular Breeding, Vol. 2: Oilseeds. Springer Verlag, Heidelberg, New York, Berlin, Toyko, pp 55-114

Tanksley SD, McCouch SR, (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science. 277: 1063-1066

Thiel T, Michalek V, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 106: 411-422

Thompson KF (1983) Breeding winter oilseed rapeseed *Brassica napus*. Adv. Appl. Biol. 7: 1-104

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994) Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88:973-980

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. 28: 286-289

Toroser D, Thormann CE, Osborn TC, 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

Tóth G, Gáspári Z, Jurka J (2000) Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. Genome Res. 10: 967-981

Tripathi MK, Mishra AS (2007) Glucosinolates in animal nutrition: A review. Anim. Feed Sci. Technol. 132: 1-27

Udall JA, Quijada PA, Lambert B, Osborn TC (2006) Identification of alleles from unadapted germplasm affecting seed yield and other quantitative traits in hybrid spring oilseed *Brassica napus* L. Theor. Appl. Genet. 113: 597-609

Udall JA, Quijada PA, Polewicz H, Vogelzang R, Osborn TC (2004) Phenotypic eVects of introgressing Chinese winter and resynthesized Brassica napus L. germplasm into hybrid spring canola. Crop Sci 44: 1990-1996

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

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotech. 23: 48-55

Walker KC, Booth EJ (2001) Agricultural aspects of rape and other Brassica products. Eur J Lipid Sci Technol 103:441–446

Wierdl M, Dominska M, Petes TD (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics. 146: 769-779

Wittstock U, Halkier BA (2002) Glucosinolate research in the Arabidopsis era. Trends Plant Sci. 7: 263-270

Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nat. Rev. Genet. 2: 983-989

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol. Ecol. 11: 1-16

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

Zhou WJ, Zhang GQ, Tuvesson S, Dayteg C, Gertsson B (2006) Genetic survey of Chinese and Swedish oilseed rape (*Brassica napus* L.) by simple sequence repeats (SSRs). Genet. Resour. Crop Evol. 53: 443-447

Ziolkowski PA, Kaczmarek M, Babula D, Sadowski J (2006) Genome evolution in Arabidopsis/*Brassica*: conservation and divergence of ancient rearranged segments and their breakpoints. Plant J. 47: 63-74

Zondervan KT, Cardon LR (2004) The complex interplay among factors that influence allelic association. Nat. Rev. Genet. 5: 89-100

LIST OF ABBREVIATIONS

A.	*Arabidopsis*
AA	*Brassica* A-Genome
00	Double-low or canola seed quality
++	High levels of erucic acid and seed glucosinolates
AFLP	Amplified Fragment Length Polymorphism
B.	*Brassica*
CC	*Brassica* C-Genome
(C22:1)	erucic acid
cv.	cultivar
DNA	Desoxyribonucleic Acid
EST	Expressed Sequence Tags
e.g.	For example
F	Filial generation
g	Gram
GSL	Glucosinolate
LD	Linkage Disequilibrium
µ	Micro
MAM1	METHYLTHIOALKYLMALATE SYNTHASE 1
n	Number of chromosomes in the haploid genome
%	Percent
P	Probability
PCR	Polymerase chain reaction
QTL	Quantitative Trait Locus
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
ssp.	Subspecies
SSR	Simple Sequence Repeat (microsatellite)
UPGMA	Unweighted Pair Group Method using Arithmetic Averages

EIDESSTATTLICHE ERKLÄRUNG

„Ich erkläre: Die vorgelegte Dissertation habe ich selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

Giessen, 13 May 2008

ACKNOWLEDGMENTS

I would like to express my deepest gratitude and appreciation to my supervisor Prof. Wolfgang Friedt, for his guidance, encouragement, and continuous supervision throughout the course of research work and preparation of the thesis. Herzlichen Dank.

Special thanks are expressed to Prof. Wolfgang Köhler for consenting to be the co-advisor for this work.

My thanks are also extended to the chairman of the examination committee Prof. Ernst-August Nuppenau, and the members of the examination committee, Prof. Sylvia Schnell and Prof. Dr. Andreas Vilcinskas, for their valuable comments and reviewing the thesis.

I would like to take this opportunity to express my sincere gratitude to Dr. Rod Snowdon for his valuable suggestions, and advices. Rod you were my source of inspiration for the great ideas. Thank you.

I would like to thank Dr. Jörn Pons-Kühnemann, for his help and advice on statistical analysis.

I would like also to thank Stavros Tzigos, Daniela Bernhardt, Liane Renno and Petra Degen for valuable technical assistance.

I am particularly grateful to all my colleagues at the Department of Plant Breeding. It was a pleasure for me to work in such a warm and friendly atmosphere.

I can not offer enough thanks to my parents, sister, brothers, and my friends for their generous help, encouragement and support.

Last but not least, thanks also to those who contributed in any way but are not mentioned here.

Above all, I am sincerely thankful to Almighty God, the Creator and the Giver of life and knowledge, Who sustained me throughout the duration of this study. May the fruit of this work be used to His glory. Ameen