Aus dem Institut für Pflanzenbau und Pflanzenzüchtung I
der Justus-Liebig-Universität Giessen,
Lehrstuhl für Pflanzenzüchtung

Genome analysis and molecular breeding of *Brassica* oilseed crops

---

Habilitationsschrift

zur Erlangung des akademischen Grades eines Dr. habil.

und der *venia legendi*, verliehen durch den Fachbereich 09 Agrarwissenschaften,
Ökotrophologie und Umweltmanagement der Justus-Liebig-Universität Giessen,

für das Lehrgebiet

„Molekulare Pflanzenzüchtung und Genomanalyse“

Vorgelegt von

Dr. rer. nat. Rodney John SNOWDON

Akademischer Rat am Institut für Pflanzenbau und Pflanzenzüchtung I

Giessen, 2009

If we knew what it was we were doing, it would not be called research, would it? Albert Einstein

Table of Contents

Foreword

Genome analysis and molecular breeding of *Brassica* oilseed crops

Brassica oilseeds have become a significant agricultural product during the past 30 years and are now the world's third-leading source of both vegetable oil (after soybean and oil palm) and oil meal (after soybean and cotton). In Europe and Asia oilseed rape (predominantly from *Brassica napus* L.) is today the most important source of vegetable oil for human nutrition and as a renewable fuel. The dramatic increase in demand for *Brassica* oils places increasing pressure on plant breeders to continuously improve seed and oil yields, the overall agronomical performance and the quality of the oil and of the meal that remains after oil extraction.

Molecular genetic techniques today play an important role in breeding strategies for improvement of all major agricultural crops. On the one hand, marker-assisted breeding can potentially reduce the time and cost of trait introgression. On the other hand, genome analysis tools provide a means to dissect and better understand complex traits, and to isolate, characterise and ultimately utilise the responsible genes.

The work described here focuses on the development and application of genome analysis tools for molecular breeding and trait dissection in *Brassica* oilseed crops. The introduction gives an overview of the current status and future trends in production, breeding, genetics and genomics. Chapters 2 to 6 present one invited review article and nine original scientific articles that provide diverse examples for genome analysis and molecular breeding in oilseed rape, based on molecular genetic and molecular cytogenetic techniques. In Chapter 7 this work is summarised in the context of worldwide molecular breeding and genomics of *Brassica* oilseeds, including an outlook on future work incorporating current developments in high-throughput genomics and next-generation sequencing. Chapter 10 lists nine additional book chapters or monographs, a further 24 original scientific articles in reviewed journals, and five additional review articles. These publications cover a broad range of topics related to molecular genetics and genome analysis in *Brassica* oilseeds and other species.

Publications included in this work

1. Snowdon R.J., 2007: Cytogenetics and genome analysis in *Brassica* crops. Chromosome Res. 15: 85-95
2. Snowdon R.J., T. Friedrich, W. Friedt & W. Köhler, 2002: Identifying the chromosomes of the A and C genome diploid *Brassica* species *B. rapa* and *B. oleracea* in their amphidiploid *B. napus*. Theor. Appl. Genet. 104: 533-538
3. Wang Y.P., X.X. Zhao, K. Sonntag, P. Wehling & R.J. Snowdon, 2005: GISH analysis of BC1 and BC2 2 progenies derived from somatic hybrids between *Brassica napus* and *Sinapis alba*. Chromosome Res. 13: 819-826
4. Hasan M., F. Seyis, A.G. Badani, J. Pons-Kuhnemann, W. Lühs, W. Friedt & R.J. Snowdon, 2006: Surveying genetic diversity in the *Brassica napus* gene pool using SSR markers. Genet. Res. Crop Evol. 53: 793–802
5. Hasan M., W. Friedt., N.M. Freitag, K. Link, J. Pons-Kühnemann & R.J. Snowdon, 2008: Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). Theor. Appl. Genet. 116:1035-1049
6. Rygulla W., R.J. Snowdon, C. Eynck, A. von Tiedemann, W. Lühs & W. Friedt, 2007: Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridisation. Phytopathology 97: 1391-1396
7. Snowdon R.J., H. Winter, A. Diestel & M.D. Sacristan, 2000: Development and characterisation of *Brassica napus-Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*. Theor. Appl. Genet. 101: 1008-1014
8. Rygulla W., R.J. Snowdon, W. Friedt, I. Happstadius, W. Cheung & D. Chen, 2008: Identification of quantitative trait loci for resistance against *Verticillium longisporum* in *Brassica napus* L. Phytopathology 98: 215-221
9. Badani A.G., R.J. Snowdon, R. Baetzel, F.D. Lipsa, B. Wittkop, R. Horn, A. De Haro, R. Font, W. Lühs & W. Friedt, 2006: Co-localisation of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). Genome 49: 1499-1509
10. Obermeier C., B. Hosseini , W. Friedt & R.J. Snowdon, 2009: Global gene expression profiling via LongSAGE in a non-model plant species: A case study in seeds of *Brassica napus*. BMC Genomics, in press

1 Introduction: *Brassica* oilseed crops and genome analysis

1.1 Oilseed rape (*Brassica napus* L.) and related crop species

The species *Brassica napus* L. originated through spontaneous interspecific hybridisation between turnip rape (*Brassica rapa* L., syn. *campestris*; genome AA, 2n = 20) and cabbage (*Brassica oleracea* L.; genome CC, 2n = 18), resulting in an amphidiploid genome comprising the full chromosome complements of its two progenitors. Because no wild *B. napus* forms are known, it is assumed that the species arose relatively recently, when the parental species began being cultivated in geographical proximity due to anthropogenic influences. The occurrence of spontaneous chromosome doubling in crosses among closely-related *Brassica* diploid species is well documented; the related amphidiploids Indian or brown mustard (*Brassica juncea*; genome AABB, 2n = 36) and Abyssinian or Ethiopian mustard (*Brassica carinata*; genome BBCC, 2n = 34) arose in the same manner after crosses of black mustard (*Brassica nigra*, genome BB, 2n = 16) with *B. rapa* and *B. oleracea*, respectively. *Brassica* vegetables and oilseeds were among the earliest plants to be systematically cropped by mankind. There are indications that a vegetable crucifer was widely cultivated as early as 10,000 years ago. In India records have been identified which suggest that oilseed *Brassica*s (probably *B. rapa*) were being used as early as 4000 B.C., and 2000 years ago their use had spread into China and Japan. Swedes (*B. napus* ssp. *napobrassica*) were known in Europe at the time of the Romans, and utilization (probably of *B. rapa*) for oil purposes in northern Europe is thought to have begun around the 13th century. By the 16th century, rapeseed was the major source of lamp oil in Europe, although it was not until the 18th century that significant cultivation areas of the crop were recorded (Kroll 1994; Kimber and McGregor 1995). For winter oilseed rape only three distinct local landraces are known. These evolved in different European climate zones and hence display variation in vegetative growth and winter hardiness. The first released cultivar 'Lembkes', selected in Germany from a Mecklenburg landrace in the early 20th century, was extensively exploited in French, Swedish, German and Polish breeding programs. Spring-sown oilseed rape was first grown in Canada in the mid 20th century. Large-scale worldwide production of oilseed rape did not begin until the mid 1970s, however, when the value of rapeseed oil and seed meal was significantly improved as a result of intensive breeding efforts. Today, oilseed rape is the most heavily produced oilseed crop in Europe and only soybean has a greater importance worldwide. Oilseed rape production is dominated by North America (particularly Canada), Western Europe and China, however *Brassica* oilseed crops also play a major role in Eastern Europe, the Indian subcontinent and Australia.

1.2 Major breeding achievements in oilseed rape

Oilseed rape has become a major international crop only over the course of the past three decades. This rapid advance to one of the major arable crops is a result of spectacular breeding success. The oil from rapeseed and most other *Brassica*s naturally contains a high quantity of erucic acid (C22:1, cis 13-docosenoic acid), which has a bitter taste and in high doses has been implicated in cardiac health problems. This serious limitation of rapeseed oil as a foodstuff was overcome only by the development of “0” and “00” oilseed rape varieties in the 1970s (Stefansson 1983; Downey and Röbbelen 1989; Downey 1990). The first major breakthrough came with the initial 0-quality cultivars with erucic acid levels of less than 1% (Stefansson and Hougen 1964). Earlier rapeseed cultivars contained up to 50% erucic acid in the seed oil. 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. The value of the crop was still suppressed by the presence of high quantities of glucosinolates in the seed, however, which made rapeseed meal unsuitable as a livestock feed. In monogastric animals the digestion of glucosinolates results in the release of toxic by-products that can cause liver and kidney damage along with lymph dysfunction. 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 material. The result was the release in 1974 of the first 00-quality spring oilseed rape variety, ‘Tower’, with zero erucic acid and low glucosinolate content. Thus began the advance of oilseed rape (also known as canola) to one of the most important oil crops in temperate regions.

1.3 Genetic mapping and genome analysis in *Brassica* oilseeds

Molecular markers have been widely used to map agronomically important genes in oilseed rape and in many cases play an important role in breeding and selection procedures. The complete sequencing of *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000) and the development of comparative genetic and physical maps between *B. napus* and Arabidopsis (e.g. Parkin et al. 2005) has enormous potential for application in gene identification and breeding in oilseed rape. The genome of *B. rapa* is expected to be completely sequenced within the near future (for progress see http://www.brassica.info), and current technological developments in the field of ultra-fast DNA sequencing are beginning to revolutionise the fields of polymorphism discovery, genome analysis and molecular breeding. The number of expressed sequence tag (EST) sequences available for *Brassica* species has skyrocketed in the past few years as sequencing costs have diminished, enabling DNA sequence mining to become extremely useful for the identification and development of single nucleotide polymorphism (SNP) markers in oilseed rape. In the near future it can be expected that high-density *B. napus* SNP arrays will play an important role in development of dense genetic maps for oilseed rape. Next-generation sequencing technologies are also set to rapidly accelerate SNP discovery, so that ultra-high density SNP maps will probably become available in the relatively near future. High-throughout SNP screening methods will also be a valuable resource for whole-genome allele-trait association studies, which can potentially play a major role in the identification of genes contributing to complex traits.

The first genetic map for *B. napus* was developed by Landry et al. (1991) using restriction fragment length polymorphism (RFLP) markers. In the subsequent two decades a large number of *B. napus* genetic maps were generated worldwide using different crosses, and considerable efforts have been invested in the localisation of genes and QTL controlling agronomically relevant traits. The most extensive *B. napus* genetic map published to date was an ultradense map of 13,551 sequence-related amplified polymorphism (SRAP) markers that were assembled into an ultra-dense bin map by Sun et al. (2007). A marker density of 8.45 SRAPs per cM was achieved, which according to the authors could correspond to more than one marker per 100 kb. Although these markers are anonymous and therefore not directly useful for annotations to a physical map, this achievement demonstrates the great potential of high-throughput marker techniques for genome mapping. The availability of ultra-dense maps based on sequence-annotated SNP markers derived from next-generation sequencing or EST-SNP arrays will — in the very near future — hugely expand the opportunities for rapid discovery of candidate genes for both simple and complex traits.

A detailed summary of *B. napus* crosses, mapping populations, marker systems, map details and the quantitative traits that were studied until 2006 is given by Snowdon et al. (2006). In the following I describe some of the most recent genetic mapping studies, including QTL analysis of traits that had previously been less intensively studied, along with some novel approaches for identification of genetic markers and candidate genes closely linked to important traits in oilseed rape. In some cases marker-assisted genome scans were implemented to introgress novel genetic diversity into oilseed rape breeding lines.

1.4 Genetic analysis of seed yield and heterosis

As the importance of hybrid cultivars has increased over the past decade, there has been growing interest in identifying the mechanisms and potential genomic loci responsible for the manifestation of heterosis in oilseed rape. A number of different strategies have been used to analyse the quantitative genetics of heterosis for yield and related traits with the assistance of genetic maps. For example, Radoev et al. (2008) mapped QTL contributing to additive, dominant and overdominant heterosis effects in a mapping population of 250 doubled haploid lines that were tested in two-year, multi-location field trials along with a corresponding set of test hybrids from each of the DH lines with a common male-sterlile tester parent. Heterosis levels of up to 30% (for grain yield) were used to map QTL involved in heterosis for yield and related seed traits. A large number of epistatic interactions were found to interact with dominance and overdominance effects to control expression of heterosis. Further studies from our own work revealed that important QTL clusters contributing to yield heterosis are also involved in heterosis for plant development during seedling development (Basunanda et al. in review). This raises the possibility to identify key regulatory loci that may function as global regulators of heterosis in different tissues and at different stages of plant development. Dissection of yield and yield component traits are another important aspect that has been analysed extensively in oilseed rape by QTL analysis. For example, Chen et al. (2007) recently reported on the detection of numerous QTL for yield and yield-related traits in DH and immortalised F2 populations, and found some QTL that contributed significantly to numerous yield-related traits and could be interesting targets for yield improvement.

Technologies for highly-dense genetic mapping today make it possible to more accurately integrate and compare map and QTL data from different populations. This will be of great interest to see whether important QTL related to yield co-localise in different materials, and whether such QTL may interact with epistatic loci involved in yield heterosis. Meta-analyses with multiple mapping populations and large, common marker sets will hopefully enable such determinations in the not-too-distant future. Use of genome-wide marker screens can also be useful for the introduction of novel genetic diversity for the exploitation of heterosis in hybrid breeding. For example Li et al. (2006) described a marker-assisted approach to develop new types of *B. napus* with introgressions of A genome chromosomes from *B. rapa* and C genome chromosome segments from *B. carinata*. When crossed with conventional *B. napus* these new types demonstrated elevated levels of so-called “inter-subgenomic heterosis” for seed yield and related traits. In our own work, dense whole-genome marker scanning was used to identify DH lines in which the genes for zero erucic acid along with QTL for low glucosinolate content were introgressed from a 00-quality variety into a novel genetic background of semi-synthetic, double-high quality rapeseed (Basunanda et al. 2007). Test hybrids generated using these genetically diverse introgression lines as pollinators showed high mid-parent heterosis for seed yield (Basunanda et al. 2007, Gehringer et al. 2007).

1.5 New male sterility genes

Considerable advances have been made in recent years in mapping and marker development for genes controlling genic and cytoplasmic male sterility systems in oilseed rape. For example, Yi et al. (2006), Lei et al. (2007) and Huang et al. (2007) described the fine-mapping of three recessive genic male-sterility genes using amplified fragment length polymorphisms (AFLP®: Keygene, Wageningen, Holland) and amplified consensus genetic markers (AGGM) using large segregating populations. Tightly-linked markers were anchored to previous *B. napus* genetic maps. In each case flanking marker sequences covering a region of well under 1 cM in *B. napus* were used to delineate syntenic chromosome regions in Arabidopsis that may contain the orthologs to the respective sterility genes. Hong et al. (2008), Xie et al. (2008) and Xiao (2008) described the development of sequence-based markers with tight linkage to an epistatic genic male sterility suppressor gene, while He et al. (2008) generated sequence-characterised markers liked to a cytoplasmic male sterility fertility restoration gene. In each of these cases AFLP markers and bulked-segregant analysis played an important role in whole-genome marker saturation to identify sequences with very lose linkage to the responsible genes. Sequence annotations to Arabidopsis and an often well-conserved synteny can assist greatly to identify potential candidates in corresponding chromosome regions, and bulked-segregant analyses have proved a valuable method to fine-map and clone genes for simple-segregating genes involved in male sterility and fertility restoration. In a different approach based on differential gene expression, Wu et al. (2007) used suppressive subtractive cDNA techniques and cDNA microarray hybridisation to try to identify candidate genes for a dominant genic male sterility in *B. napus*. A number of genes involved in male gametogenesis pathways were among the differentially expressed genes between fertile and sterile near-isogenic lines.

1.6 Genetic improvement of seed quality

Much interest has developed recently in breeding of yellow-seeded oilseed rape and canola with improved seed meal quality. The yellow-seed trait in *B. napus* is generally associated with a reduced seed coat thickness; this leads to a reduced contribution of the seed coat to the seed meal after oil extraction and a consequent lowering of anti-nutritive crude fibre and phenolic compounds. At the same time the proportional contribution of the embryo to the total seed is increased, raising the sum content of oil plus protein. Unfortunately, seed colour itself is difficult to use as a morphological marker for improved meal quality, because the accumulation of seed coat tannins is highly sensitive to temperature, light intensity and other environmental factors. Therefore there is a considerable effort to identify major genes contributing to reduced seed coat in different yellow-seeded materials and to develop screening tools and markers for effective breeding of high-performing light-seeded varieties. In one example from our own work, Badani et al. (2006) localised a major QTL with a large contribution on seed colour and acid detergent fibre content in two different yellow-seeded winter rapeseed sources. The gene was flanked by markers originating from *B. napus* chromosome N18, although later work showed that the chromosome region containing the major QTL may be a non-reciprocal translocation from chromosome N09. In Chinese oilseed rape with a completely different genetic background to the above study, Fu et al. (2007) also found a major dominant QTL that appeared to be localised on chromosome N09. Xiao et al. (2007) also developed closely linked markers to a major gene for yellow seed colour in Chinese oilseed rape. It will be of great interest to determine whether the same major genes are influencing seed colour-related traits in these genetically diverse materials. Furthermore, the availability of markers and candidate genes will enable us to identify new allelic diversity for seed meal quality and ultimately to clone the responsible genes.

Synteny to Arabidopsis and increasing quantities of aligned genomic sequence data from *Brassica* species are particularly valuable for identification of candidate genes and potentially gene-linked markers for important seed quality traits in oilseed rape. To demonstrate the power of synteny-based marker development, we developed potentially gene-linked markers for four seed glucosinolate candidate genes from *A. thaliana*. Via structure-based allele-trait association studies in genetically diverse *B. napus* genotypes, we demonstrated that othologous copies of all four candidate genes were associated with significant effects on total seed glucosinolate content (Hasan et al. 2008). 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.

Identification and utilization of important genes contributing to oil content is one of the major aims of seed quality breeding in oilseed rape. A recent publication compared oil content QTL in different mapping populations and revealed that some major gene loci appear to influence this complex trait in different genetic backgrounds. Delourme et al. (2006) localised oil content QTL in two large, genetically divergent mapping populations and compared their locations to previously mapped QTL from earlier published studies. In some cases the QTL were found to be consistently revealed across different genetic backgrounds. In particular, a QTL on N03 was revealed in all the studies and other QTL on N01, N08 and N13 were found in three out of five different studies. Other QTL were located in homoeologous genome regions, while some were specific to a particular genetic background and potentially carry novel alleles. These results show the potential for combination of favourable alleles at different QTL to increase seed oil content. Furthermore, examples were given of how Arabidopsis genomic data could be used to derive markers and identify candidate genes for oilseed rape QTL. The study also demonstrated the added value of consolidated information from different segregating populations, in order to identify meta-QTL involved in a specific trait in different genetic backgrounds.

1.7 Genetic mapping of resistance loci

Mapping and marker development for resistance genes to biotic stress factors represent ongoing goals for oilseed rape breeding. In recent years, continued progress has been made in the map-based cloning of genes contributing to resistance against blackleg disease, the major disease of oilseed rape worldwide (see Rimmer et al. 2006). Mayerhofer et al. (2005) described the fine mapping of loci involved in seedling resistance to blackleg in two different canola cultivars. Both loci localised to the same position on *B. napus* chromosome N07, and a collinear chromosome region was able to be identified in Arabidopsis. A complex pattern of tandem duplications was identified in the *B. napus* genome region containing these loci. Apparently, duplication and sequence divergence during the polyploidisation events that led to *Brassica* species may also have played a major role in the evolution of resistance to major pathogens. Candidate genes for blackleg resistance were also identified in Arabidopsis by Staal et al. (2006); fine-mapping was performed in recombinant inbred lines to identify two genes that were associated with resistance and contained typical resistance gene sequence motifs. The contribution to resistance was confirmed by reverse genetics. For a number of other diseases of oilseed rape breeding efforts have been hindered by a lack of resistance sources. In some cases this can be overcome by introduction of resistance genes from exotic *B. napus* materials, resynthesised rapeseed or other interspecific crosses, however the available of useful selection markers is a prerequisite for effective combination of quantitative resistances in elite germplasm. For example, we identified QTL-linked markers associated with resistance against *Verticillium longisporum* introduced from *B. oleracea* (Rygulla et al. 2007), while Werner et al. (2008) described QTL involved in resistance against clubroot disease (*Plasmodiophora brassica*e). Resistance to *Sclerotinia sclerotiorum* is a major breeding aim in most of the major oilseed rape growing areas of the world, however little resistance has yet been identified against this disease in *B. napus*. QTL analysis of a partial resistance in Chinese rapeseed lines was characterised by Zhao et al. (2003), and gene expression profiles produced by resistant and susceptible genotypes in response to *S. sclerotiorum* infection was analysed by Zhao et al. (2007) using microarray analysis. Early response genes to pathogen inoculation were integrated into the QTL map, leading to the identification of a number of candidate genes for the defence reaction. Among the genes that co-localised with interesting resistance QTL, some plant cell wall-related proteins and WRKY transcription factors were identified as potential contributors to defence against sclerotinia rot.

Abiotic stress resistance is also gaining increasing attention in oilseed rape breeding, although the regulatory mechanisms involved in whole-plant reactions to drought conditions, nutrient deprivation or cold stress can be extremely complex. Considerable efforts have been made to investigate genes involved in vernalisition requirement and flowering time in *B. napus* and related species (reviewed in detail by Snowdon et al. 2006), and in recent years it has become clear that some of the major genes controlling flowering traits may play an important global role in gene expression regulation in general. The role of FLOWERING LOCUS C (FLC) homologs and associated genes in the regulation of flowering time and related traits, and their involvement in relevant QTL for these traits, has recently been confirmed in *B. napus* and its diploid progenitors (Pires et al. 2004, Kim et al. 2007, Lou et al. 2007, Okazaki et al. 2007, Razi et al. 2008). Such global regulatory genes are particularly interesting with regard to overall yield improvement through targeted manipulation of development and heterosis.

1.8 New developments in marker applications and molecular breeding

The astonishing current developments in next-generation sequencing technologies (see Dalton 2006) offer unprecedented opportunities for new genomics-based breeding and selection strategies. In particular, as soon as the sequencing of the reference A genome of *B. rapa* is completed it will be possible to re-sequence large portions of the *B. oleracea* C genome and consequently the A and C genomes of *B. napus*. Even without a reference sequence, the next-generation sequencing technologies enable large-scale comparative sequencing of BAC libraries from elite breeding lines for a relatively low (and continuously decreasing) cost, so that whole-genome selection in oilseed rape and *other* major crops is likely to become a reality in the foreseeable future. This new sequence-based genomics era is likely to completely change the way that genetic mapping, genome analysis and marker-assisted selection are performed in crop plants.

Considerable interest is arising in the potential use of whole-genome selection techniques based on ultra-high thoughput, genome-wide SNP marker screening. Whole-genome selection is today already being utilized in commercial animal breeding programs, and as genotyping costs continue to fall they could also become an integral part of breeding strategies for *Brassica* oilseeds and other major crops in the foreseeable future. Next-generation sequencing techniques will doubtless be a pivotal technology in the application of whole-genome association genetics methods for the identification and utilisation of genes involved in important complex traits. Already, the use of genome-wide transcriptome analysis has enabled the identification of potential global gene expression regulators that could be useful for significant yield increases. Ultradeep expression profiling via next-generation sequencing, and comparisons to whole-genome sequence variation, have the potential to take this kind of gene discovery to a completely new level.

Although oilseed rape genome research has long profited from the close relationship to Arabidopsis, one of the most intensively studied plant species, in the near future it is likely that a vast array of genomic tools will also be available for crop brassicas. One major limitation to genome-assisted breeding is the current lack of associations between genomic data and detailed, reliable phenotypic data for factors contributing to major traits. Intensive phenotyping, including high-thoughput physiological and metabolite profiling, may be the most important key to understanding important complex traits like oil content, seed development, biotic and abiotic stress tolerance and the manifestation of yield characters in oilseed rape. Novel and high-thoughput phenotpying technologies should therefore be an important priority in coming years, in order to facilitate the identification of genomic and transcriptomic variation associated with economically important characters.

1.9 Cited literature

Badani AG, Snowdon RJ, Baetzel R, Lipsa FD, Wittkop B, Horn R, De Haro A, Font R, Lühs W, Friedt W (2006) Co-localisation of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). Genome 49: 1499-1509

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

Basunanda P, Radoev M, Ecke W, Friedt W, Becker H, Snowdon RJ (2009) Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). Theor. Appl. Genet., in review

Chen W, Zhang Y, Liu XP, Chen BY, Tu JX, Fu TD (2007) Detection of QTL for six yield-related traits in oilseed rape (*Brassica napus*) using DH and immortalized F_2 populations. Theor Appl Genet 115: 849–858

Dalton CR (2006) Sequencers step up to the speed challenge. Nature 443: 258–259

Delourme R, Falentin C, Huteau V,Clouet V, Horvais R, Gandon B, Specel S. Hanneton L, Dheu JE, Deschamps M, Margale E, Vincourt P, Renard M (2006) Genetic control of oil content in oilseed rape (Brassica napus L.). Theor Appl Genet 113: 1331–1345

Downey RK (1990) *Brassica* oilseed breeding: achievements and opportunities. Plant Breed Abstr 60: 1165–1170

Downey RK, Röbbelen G (1989) *Brassica* species. In: Röbbelen G, Downey R, Ashri A (eds) Oil Crops of the World. McGraw-Hill, New York, pp 339-382

Fu FY, Liu LZ, Chai YR, Chen L, Yang T, Jin MY, Ma AF, Yan XY, Zhang, ZS, Li JN (2007) Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments, Genome 50: 840-845

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 marginal conditions. Breeding Sci 57: 315-320

Hasan M, Friedt W, Freitag NM, Link K, Pons-Kühnemann J, 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

He J, Ke L, Hong D, Xie Y, Wang G, Liu P, Yang G (2008) Fine mapping of a recessive genic male sterility gene (Bnms3) in rapeseed (*Brassica napus*) with AFLP- and Arabidopsis-derived PCR markers. Theor Appl Genet 117: 11-18

Hong DF, Liu J, Yang GS, He QB (2008) Development and characterization of SCAR markers associated with a dominant genic male sterility in rapeseed. Plant Breeding 127: 69-73

Huang Z, Chen YF, Yi B, Xiao L, Ma CZ, Tu JX, Fu TD (2007) Fine mapping of the recessive genic male sterility gene (Bnms3) in *Brassica napus* L. Theor Appl Genet 115: 113-118

Kim SY, Park BS, Kwon SJ, Kim J, Lim MH, Park YD, Kim DY, Suh SC, Jin YM, Ahn JH, Lee YH (2007) Plant Cell Rep 26: 327–336

Kimber D, McGregor DI (1995) *Brassica* Oilseeds: Production & Utilisation. CAB, Wallingford, UK

Kroll H (1994) Ein archäologischer Rapsfund des 16. Jahrhunderts, entdeckt in Heide in Holstein, Norddeutschland. J Agron Crop Sci 173: 17-21

Landry BS, Hubert N, Etoh T, Harada JJ, Lincoln SE (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genome 34: 543-552

Lei SL, Yao XQ, Bin Y, Chen W, Ma CZ, Tu JX, Fu TD (2007) Towards map-based cloning: fine mapping of a recessive genic male-sterile gene (BnMs2) in *Brassica napus* L. and syntenic region identification based on the *Arabidopsis thaliana* genome sequences. Theor Appl Genet 115: 643-651

Li M, Chen X, Meng J (2006) Intersubgenomic heterosis in rapeseed production with a partial new-typed *Brassica napus* containing subgenome A^r from *B. rapa* and C^c from *Brassica carinata*. Crop Sci 46: 234-242

Lou P, Zhao J, Kim JS, Shen S, Del Carpio DP, Song X, Jin M, Vreugdenhil D, Wang X, Koornneef M, Bonnema G (2007) Quantitative trait loci for flowering time and morphological traits in multiple populations of *Brassica rapa*. J Exp Bot 58: 4005–4016

Mayerhofer R, Wilde K, Mayerhofer M, Lydiate D, Bansal V, Good A, Parkin I (2005) Complexities of chromosome landing in a highly duplicated genome: Towards map based cloning of a gene controlling blackleg resistance in *Brassica napus*. Genetics 171: 1977-1988

Okazaki K, Sakamoto K, Kikuchi R, Saito A, Togashi E, Kuginuki Y, Matsumoto S, Hirai M (2007) Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea*. Theor Appl Genet 114: 595–608

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

Pires JC, Zhao J, Schranz ME, Leon EJ, Quijada PA, Lukens LN, Osborn TC (2004) Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids (Brassicaceae). Biol J Linnean Soc 82: 675-688

Radoev M, Becker HC, Ecke W (2008) Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. Genetics 179: 1547–1558

Razi H, Howell EC, Newbury HJ, Kearsey MJ (2008) Does sequence polymorphism of FLC paralogues underlie flowering time QTL in *Brassica oleracea*? Theor Appl Genet 116:179–192

Stefansson BR (1983) In: Kramer JKG, Sauer FD, Pigden WJ (eds) High and low erucic acid rapeseed oils. Academic, New York, pp 143-160

Stefansson BR FW Hougen (1964) Selection of rape plants (*Brassica napus*) with seed oil practically free of erucic acid. Can J Plant Sci 44: 359–364

The Arabidopsis Genome Initiative (2000) Analysis of the genome of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815

Werner S, Diederichsen E, Frauen M, Schondelmaier J, Jung C (2008) Genetic mapping of clubroot resistance genes in oilseed rape. Theor Appl Genet 116: 363-372

Wu J, Shen J, Mao X, Liu K, Wei L, Liu P, Yan G (2007) Isolation and analysis of differentially expressed genes in dominant genic male sterility (DGMS) *Brassica napus* L. using subtractive PCR and cDNA microarray. Plant Science 172: 204-211

Xiao L (2008) Molecular markers linked to Bn;rf: a recessive epistatic inhibitor gene of recessive genic male sterility in *Brassica napus* L.. Euphytica 164: 377-384

Xiao S, Xu J, Li Y, Zhang L, Shi S, Wu J, Liu K (2007) Generation and mapping of SCAR and CAPS markers linked to the seed coat color gene in *Brassica napus* using a genome-walking technique. Genome 50: 611-618

Xie YZ, Hong DF, Xu ZH, Liu PW, Yang GS (2008) Identification of AFLP markers linked to the epistatic suppressor gene of a recessive genic male sterility in rapeseed and conversion to SCAR markers. Plant Breeding 127: 145-149

Yi B, Chen YN, Lei SL, Tu JX, Fu TD (2006) Fine mapping of the recessive genic male-sterile gene (*Bnms1*) in *Brassica napus* L. Theor Appl Genet 113: 643–650

Zhao JW, Meng JL (2003) Genetic analysis of loci associated with partial resistance to *Sclerotinia sclerotiorum* in rapeseed (*Brassica napus* L.). Theor Appl Genet 106: 759-764

Zhao J, Wang J, An L, Doerge RW, Chen ZJ, Grau CR, Meng J, Osborn TC (2007) Analysis of gene expression profiles in response to *Sclerotinia sclerotiorum* in *Brassica napus*. Planta 227: 13-24

2 Molecular cytogenetics of *Brassica* and related species

This chapter begins with a review article describing recent advances in the application of cytogenetic and genomic techniques in *Brassica* oilseed crops, with particular emphasis on the structure and evolution of *Brassica* genomes and the consequences for genome analysis. The second article in the chapter represents the first compete molecular karyotpe of *B. napus* based on fluorescence in situ hybridisation (FISH). The use of FISH techniques to identify physical chromosomes and assign them to *Brassica* genetic maps is today an integral part of the *Brassica* rapa genome sequencing project. The third article presents one example for the use of genomic *in situ* hybridisation (GISH) to identify and characterise alien genome introgressions in *Brassica* interspecific hybrids. Since the development of the methodology to distinguish different genome components in *Brassica* amphidiploids and interspecific hybrids (Snowdon et al. 1997), GISH has become the method of choice for initial analysis of chromosome additions and introgressions carrying genes of interest from a related donor species. Such techniques play a particularly important role in the transfer of new resistance genes to *B. napus* from related crucifer species.

2.1 Cytogenetics and genome analysis in *Brassica* crops

Snowdon R.J., 2007. Chromosome Res. 15: 85-95

Chromosome Research (2007) 15:85–95
DOI: 10.1007/s10577-006-1105-y

Cytogenetics and genome analysis in *Brassica* crops

Rod 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; Tel: +49-641-9937423; Fax: +49-641-9937429; E-mail: Rod.Snowdon@agrar.uni-giessen.de

Key words: *Brassica*, comparative genomics, cytogenetics, genome analysis

Abstract

The genus *Brassica* contains a wide range of diploid and amphipolyploid species including some of the most important vegetable, condiment and oilseed crops worldwide. As members of the Brassicaceae family the brassicas are the closest crop relatives to the model plant *Arabidopsis thaliana*, 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. In this review examples are shown from recent studies that demonstrate the potential for intergenome navigation from model to crop plant and for comparisons among genetic and cytogenetic maps between the model and crop species and among different crop brassicas. The use of interspecific and intergeneric hybridization for introgression of novel traits into *Brassica* genomes from the secondary and tertiary crucifer genepools is described. In this context the use of the *Brassica* triangle of three diploid species and their corresponding amphiploids as an excellent model system for studying the mechanisms and control of homeologous recombination and polyploidization is discussed from a crop breeding perspective.

Introduction

The chromosomal relationships among the A, B and C genomes of the diploid species *Brassica rapa* (genome AA, 2n = 20; turnip rape, turnip, Chinese cabbage), *B. nigra* (genome BB, 2n = 16; black mustard) and *B. oleracea* (genome CC, 2n = 18; cabbage, cauliflower, broccoli, kale, kohlrabi, brussel sprouts) and their natural spontaneous amphidiploids *B. carinata* (genome AABB, 2n = 34; Abyssinian or Ethiopian mustard), *B. napus* (genome AACC, 2n = 38; oilseed rape, swede) and *B. juncea* (genome BBCC, 2n = 36; Indian or brown mustard) were elucidated through interspecific crosses and meiotic analyses by the Asian cytogeneticists Morinaga and U in the early 20th century (Morinaga 1933, 1934, U N 1935). Because the *Brassica* amphidiploid species can be generated synthetically with the help of embryo rescue techniques, this complex of three diploid species and their corresponding polyploids (Figure 1) is today one of the most useful model systems for investigations of polyploidy in crop plants (e.g. Song *et al.* 1995, Lukens *et al.* 2006). Colchicine treatment can also be used to artificially synthesize autotetraploid brassicas, which can potentially be used to compare the corresponding effects of gene dosage, autoploidy, alloploidy and amphiploidy on gene regulation and expression.

Brassica–Arabidopsis comparative genomics: model-based breeding tools

Comparative genome analyses between *Brassica* and the model crucifer *Arabidopsis thaliana* have

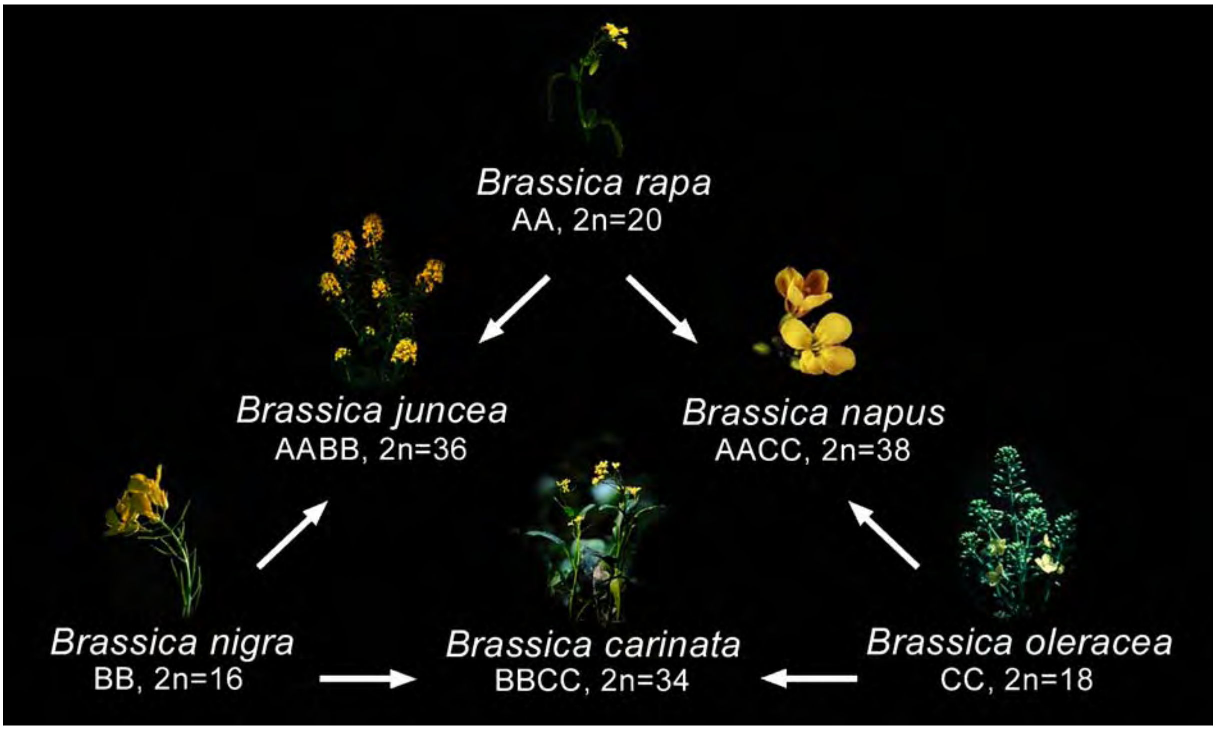


Figure 1. The *Brassica* triangle of species, as described by U N (1935), representing the A, B and C genomes and their respective amphidiploids that arose from spontaneous chromosome doubling via meiotic nondisjunction after interspecific hybridizations in regions of overlapping geographical distribution of the respective diploid progenitors.

revealed that *Brassica* chromosomes show a complex rearrangement in comparison to the *A. thaliana* genome (Lagercrantz & Lydiate 1996), presumably the result of numerous rounds of polyploidization. As a consequence the synteny and microsynteny relationships between the more or less triplicated diploid *Brassica* genomes (Schmidt *et al.* 2001, Lan *et al.* 2000, Lysak *et al.* 2005, Parkin *et al.* 2005, Ziolkowski *et al.* 2006) and the paleopolyploid genome of *Arabidopsis* are extremely complicated. Detailed comparative genomic information is therefore needed for directed utilization of genome data from *Arabidopsis* in *Brassica* molecular breeding.

The most intensively studied *Brassica* amphidiploid is the most economically important crop, oilseed rape (*B. napus* ssp. *napus*). The A and C genome chromosomes in *B. napus* have remained essentially unaltered after their amphipolyploidization (Olsson & Hagberg 1955, Attia *et al.* 1987, Sharpe *et al.* 1995, Snowdon *et al.* 2002). Hence it is possible to compare genetic maps, chromosomes and gene positions among the three species and to use the smaller diploid genomes to obtain genomic data that can be extrapolated to the amphidiploid. A large number of independent genetic maps among different *B. napus* crosses have been generated (reviewed by Snowdon *et al.* 2006) with a particular emphasis on investigation of quantitative traits, development of markers for marker-assisted breeding, and potential map-based cloning of genes involved in agronomically important traits. The ability to navigate between the *Arabidopsis* sequence and the major *Brassica* genomes is improving constantly as the genomic sequencing of *B. rapa* (Bancroft 2006, Yang *et al.* 2006, see also http://www.brassica.info) progresses and the annotation of the new sequence data to the *Arabidopsis* genome is updated. On the other hand, comparative genomics from *Arabidopsis* sequence information is a valuable asset for *in-silico* chromosome walking within the *B.* rapa sequencing project (Bancroft 2006, Lim *et al.* 2006). For example, evenly spaced seed BAC (bacterial artificial chromosomes) for the sequencing of *B. rapa* chromosome 1 were selected using mapped expressed sequence tags (EST). The physical positions and ordering of these BAC on chromosome 1 were confirmed by fluorescence *in-situ* hybridization (FISH) to mitotic and meiotic chromosomes. Because the majority of the BAC show collinearity to the corresponding homoeologous chromosome region in *Arabidopsis* it is possible to validate the positions of the *B. rapa* sequencing clones by *in-silico* comparative physical mapping (Yang *et al.* 2006).

A detailed comparative genetic map between the chromosomes of *A. thaliana* and *B. napus* was generated by Parkin *et al.* (2005) by localization of orthologous sequences from mapped *B. napus* RFLP markers to the corresponding chromosome positions in *A. thaliana*. A total of 21 chromosomal blocks were identified in the genome of *A. thaliana* that could be duplicated and rearranged to reconstruct the basic chromosome structure of *B. napus*. Each block contained several closely linked homologous loci with more or less conserved order in both the *A. thaliana* and *B. napus* genomes. For each *B. napus* chromosome numerous blocks of conserved synteny were found between *B. napus* and *A. thaliana*, each presumably representing chromosomal segments that have been maintained since the divergence of *Arabidopsis* and *Brassica* from a common ancestor. According to Parkin *et al.* (2005) a minimum of 74 gross rearrangements (38 in the A genome and 36 in the C genome) appear to have occurred since the divergence of the brassicas from the *Arabidopsis* lineage some 14–24 million years ago (Koch *et al.* 2000). In similar studies Lukens *et al.* (2003) and Lan *et al.* (2000) compared the positions of sequenced loci with a known position on a *B. oleracea* genetic map to the physical positions of their putative orthologues in *A. thaliana*. By distinguishing orthologous from paralogous loci and establishing criteria to identify significant regions of collinearity between the genomes a total of 34 significant *A. thaliana* regions were found by Lukens *et al.* (2003) to be collinear with 28% of the *B. oleracea* genetic map. The large number of macrosynteny breakpoints coupled with the identification of extensive duplications in the *B. oleracea* genome also confirmed the extremely high degree of chromosomal rearrangement since divergence of the *Brassica* diploids from *A. thaliana*. Nevertheless it is potentially possible to utilize rough macrosyntenic data to navigate between particular regions of interest in the *Arabidopsis* genome and *Brassica* genetic maps. On the other hand evidence from microsyntenic studies of gene order in duplicated chromosome regions suggests that interspersed gene loss is prevalent in homoeologous *Brassica* chromosome regions (O'Neill & Bancroft 2000, Rana *et al.* 2004, Town *et al.* 2006).

The conservation of gross macrosynteny can prove extremely useful in the identification and characterization of *Brassica* genes involved in important quantitative traits based on comparative analysis with chromosome regions in *Arabidopsis* corresponding to QTL in the crop species. For example, considerable knowledge has been gained on the genetic control of vernalization-responsive flowering time in *B. rapa* and *B. napus* through QTL analysis and comparative mapping of genes associated with this trait in homoeologous *Brassica* and *Arabidopsis* genome regions. Osborn *et al.* (1997) found that a chromosome region containing a homoeologous QTL for flowering time in *B. rapa* and *B. napus*, respectively, was collinear with a region of *A. thaliana* chromosome 5 containing the flowering time genes *FLC*, *FY* and *CO*. A second QTL region showed fractured collinearity with several regions of the *Arabidopsis* genome, including the top of chromosome 4 where another flowering-time gene, *FRI*, is located. Detailed analysis enabled identification of *FLC* and *FRI* as the major candidate genes for regulation of flowering time in the *Brassica* species.

On a sequence level high similarity is found between the exons of putative orthologous genes in *Arabidopsis* and *Brassica* (Schmidt 2002), meaning that knowledge from *Arabidopsis* is highly relevant for gene isolation and characterization in *Brassica* crops. With an ever-growing resource of *Brassica* sequence data it is today becoming increasingly possible – despite the complex rearrangements among *Brassica* genomes – to annotate, align and compare chromosomal and genomic data between the crop brassicas and the model species and to use this new information for genomic studies in the comparatively large crop genomes. Navigation between *Brassica* and *Arabidopsis* physical maps using published genome annotation and synteny data uncovers an enormous wealth of tools for fine-mapping, synteny-based gene cloning and marker development for marker-assisted selection. For example, we have successfully used online SSR search engines to scan *Arabidopsis* chromosome regions flanking candidate genes of interest or major QTL positions. Many of the *Brassica* SSR primers we identified in this manner were found to amplify polymorphic markers at one or more homologous loci in oilseed rape. The linkage of these markers to the trait of interest could often be confirmed by re-mapping to QTL regions or by allele-trait association analysis in genetically diverse material (unpublished results). If such markers are in linkage disequilibrium with the gene of interest, this strategy can be extremely useful for indirect mapping of candidate genes on *Brassica* chromosomes. Because SSR markers are codominant, this approach also has the potential to enable map localization of duplicated copies of a given candidate gene, for example to allow comparisons with major QTL positions.

Integration of karyotypes with genetic and physical maps

Development of classical cytogenetic resources for *Brassica* crop species is difficult due to the small size of the chromosomes and the lack of distinct karyological features that can be readily identified in metaphase preparations. This restricts the exact cytological characterization of *Brassica* addition, substitution and particularly introgression lines. However, *Brassica* interspecific addition and introgression lines can be generated with relative ease via interspecific hybridization and recurrent backcrossing using embryo rescue techniques in early generations. Characterization of addition lines with molecular markers has to a certain extent enabled the successful characterization of specific additions or introgressions for genome analysis or associations with particular traits of interest from the donor genome. Quiros *et al.* (1987) generated a set of monosomic addition plants and disomic addition lines by crossing and backcrossing *B. napus* to *B. rapa* (syn. *campestris*) and confirmed the genome-specificity of the additions using genome-specific molecular markers. The marker analyses revealed extensive gene duplications on individual addition chromosomes (McGrath *et al.* 1990) and the presence of deletions in some chromosomes (Hu & Quiros 1991), a phenomenon that may be a contributing factor to the extensive chromosomal rearrangement that has occurred during the evolution of the Brassicaceae. In a similar manner Chen *et al.* (1992, 1997a) generated addition lines containing monosomic addition chromosomes from the C-genome donor *Brassica alboglabra* in a *B. rapa* background by backcrossing a resynthesized *B. napus* containing the donor genome to its parental *B. rapa* line. One addition line was found to contain, on the same addition chromosome, independent genes for erucic acid biosynthesis, white flower colour and an isoenzyme of leucine aminopeptidase. Again, intergenomic recombination was frequently observed in the monosomic addition line and resulted in the introgression of one or two loci from the alien chromosome into the *B. rapa* genome. Another addition line was found to contain a gene controlling seed colour in a region where a homoeologous recombination or chromosomal deletion was detectable (Chen *et al.* 1997b). Disomic *B. napus* + *B.* nigra addition lines were described by Chèvre *et al.* (1991) and Struss *et al.* (1991), respectively. A number of other studies have reported individual addition or substitution lines between *B. napus* and related genera that have been generated in an effort to introgress specific traits into oilseed rape (e.g. Sjödin & Glimelius 1989, Skarzhinskaya *et al.* 1998, Snowdon *et al.* 2000, Voss *et al.* 2000, Peterka *et al.* 2004). In comparison to cereals, however, where chromosome additions, translocations and introgressions are well characterized and can be closely integrated with genetic maps, such detailed cytogenetic information is not yet available for *Brassica* materials and broad, well-defined cytological stocks are not available.

The use of FISH techniques offers new potential not only for more reliable chromosome identification in *Brassica*, but also in terms of potential information regarding the integration of genetic and physical maps, for ordering molecular markers and measuring physical genome distances, and for structural and functional chromosome analyses. Methods for the accurate *in-situ* localization of repetitive DNA sequences at chromosomal sub-arm level, particularly repetitive DNA sequences, have enabled a considerably more accurate identification of chromosomes and the elucidation of karyotypes for diploid and amphidiploid *Brassica* species (e.g. Figure 2, see also Armstrong *et al.* 1998, Fukui *et al.* 1998, Snowdon *et al.* 2002, Maluszynska & Hasterok 2005). FISH hybridization of BAC clones to *B. oleracea* (Howell *et al.* 2002) and *B. rapa* chromosomes (Jackson *et al.* 2000) represents a first step towards integration of physical and genetic maps with the karyograms of the diploid species and their amphidiploid hybrid *B. napus*. The first integration of a complete *Brassica* genetic map with the corresponding mitotic chromosome karyotype was achieved by Howell *et al.* (2002), who assigned all nine linkage groups of a *B. oleracea* genetic map to the nine mitotic metaphase chromosomes using FISH. The probes were

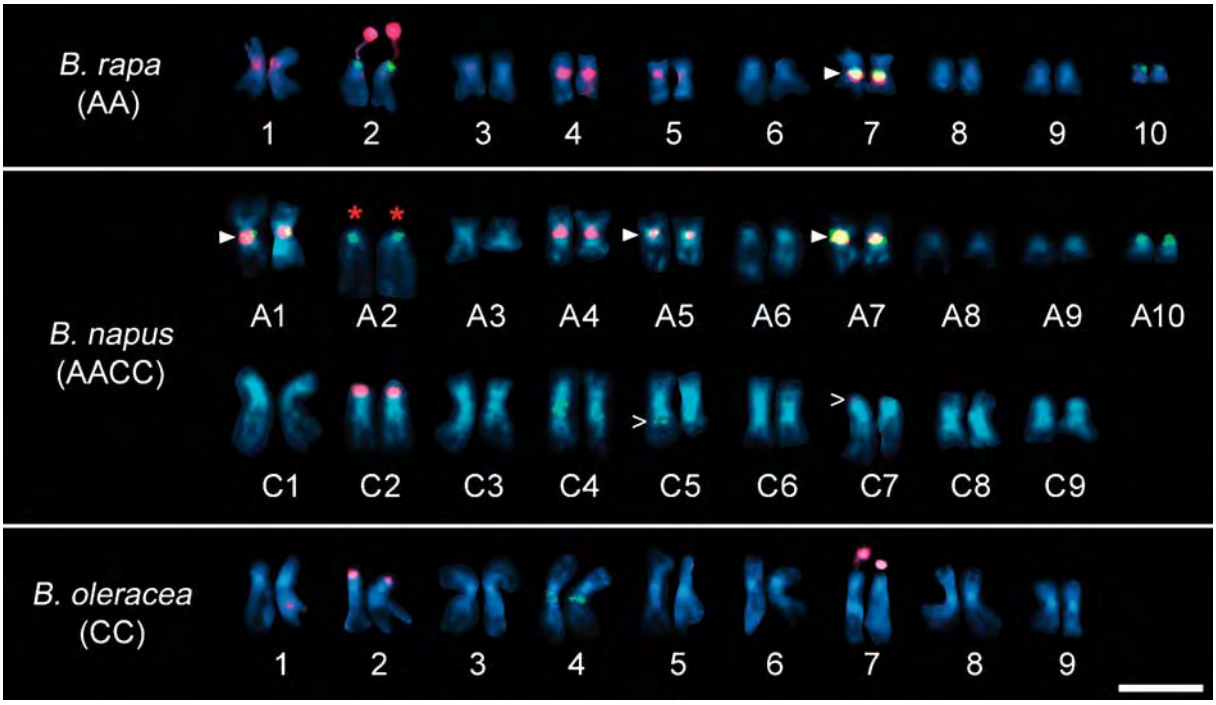


Figure 2. Karyotypes based on fluorescence *in-situ* hybridization patterns with 5S (green) and 25S (red) rDNA probes and DAPI staining (blue), for *Brassica rapa*, *B. oleracea* and their amphidiploid *B. napus*. Closed arrowheads indicate co-localization of 5S and 25S loci, whereas open arrowheads show a small 5S locus and a small 25S locus on *B. napus* chromosomes C5 and C7, respectively. The red asterisks represent the position of a large 25S rDNA locus located on the satellite of *B. napus* chromosome A2, which in this spread was lost during chromosome preparation. The *B. napus* karyotype is divided into two sets of chromosomes with differing chromatin condensation patterns resembling, respectively, those of *B. rapa* (**A**) and *B. oleracea* (**C**). Each *B. napus* chromosome is aligned and numbered in accordance with its putative homologue in the *B. rapa* or *B. oleracea* genome. Image from Snowdon *et al.* (2002).

mainly *B. oleracea* BAC clones that could be assigned to linkage map positions through development of locus-specific PCR assays. A total of 22 probes representing 19 loci were used to integrate the cytogenetic and genetic linkage maps and compare the orientation of the chromosomes with their respective linkage groups. Such cytogenetic maps represent an important new resource for locating genomic sequences with unknown genetic map positions and to analyse the relationships between genetic and cytogenetic maps. A similar strategy based on FISH localization of genetically and physically mapped BAC clones is currently being followed for integration of the genetic, physical and cytological maps of *B. rapa* within the *B. rapa* whole-genome sequencing project (Yang *et al.* 2006). Hybridization of locus-specific probes to meiotic pachytene chromosomes can considerably increase the resolution of the cytological mapping. Ziolkowski & Sadowski (2002) demonstrated the application of this technique in *B.* oleracea using ribosomal DNA probes and BAC clones from *A. thaliana*. Because of the considerably higher resolution than is possible in *Brassica* mitotic chromosomes, cytological mapping to pachytene bivalents offers the opportunity to potentially compare physical and genetic distances among selected markers. Together with synteny data and comparative genome annotations this technique should play a major role in increasing the feasibility and success of map-based gene cloning in *Brassica* crops.

Repetitive sequences and genome evolution

Centromeric regions of *B. rapa and B. oleracea* chromosomes contain two divergent 176 bp centromeric repeat sequences (Hallden *et al.* 1987, Iwabuchi *et al.* 1991, Sibson *et al.* 1991, Xia *et al.* 1993) with a degree of specificity for different chromosomes within the diploid genomes (Harrison & Heslop-Harrison 1995). Lim *et al.* (2005) identified two classes of these 176 bp repeat sequences in *B. rapa*. Similar 176 bp repeats with more than 90% sequence similarity are also predominant in distant crucifer relatives including *Diplotaxis* (Harbinder & Lakshmikumaran 1990), although less than 80% sequence identity is found in the corresponding repeats in *Sinapis* and *Raphanus* (Capesius 1983, Grellet *et al.* 1986, Sibson *et al.* 1991, Xia *et al.* 1993). In *B. nigra* two different tandem repeat sequences specific to the B genome were identified by Gupta *et al.* (1992), one of which is unrelated to either A or C genome sequences. *In-situ* hybridization to *B. juncea* metaphase chromosomes showed that this sequence is located at all B genome centromeres and enables B genome chromosomes to be distinguished from A and C genome chromosomes by GISH (Schelfhout *et al.* 2004).

Detailed analysis of sequence diversity in repetitive sequences can offer interesting insight into the evolutionary relationships among and between *Brassica* species and their crucifer relatives, and may give useful information regarding the potential for intergeneric gene transfer to brassicas from more distant relatives via homoeologous recombination (see below). For example, Lenoir *et al.* (1997) and Tatout *et al.* (1999) studied the molecular phylogeny of short interspersed element (SINE) Sl retrotransposons in numerous different *Brassica* diploids and related crucifer species. Even among closely related species significant sequence divergence was found that appears to have arisen from sequence-specific surges in S1 amplification (Lenoir *et al.* 1997). SINE amplification near or in genes can cause post-transcriptional regulation or even give rise to novel gene domains (Bennetzen 2000). Furthermore, the amplification of SINES can result in unequal homologous recombination between cis-Sl elements (Lenoir *et al.* 1997), potentially leading to chromosomal rearangements that can be a driving factor in speciation. The genomic organization and diversity of retrotransposons in *Brassica* diploids and allotetraploids was investigated in detail by Alix *et al.* (2004, 2005). Sequence analysis revealed distinct Ty1/copia and LINE-like elements, whereby the latter were present at only very low copy numbers in the genomes investigated. A third clade could be subdivided into Ty3/*gypsy*, Athila and virus-like branches. Phylogenies based on the sequence comparisons showed no correlation with the known genome relationships among the species of the *Brassica* triangle, indicating that members of the element families were present in a common ancestor (Alix *et al.* 2004). On the other hand some sub-families appeared to be amplified in individual species. Fluorescent *in-situ* hybridization of representative reverse transcriptase domains from the different retroelements to *B. oleracea* (Alix *et al.* 2005) showed characteristic chromosomal distributions for each group, suggesting that the different retrotransposons have preferential amplification sites and possibly different insertion/excision control mechanisms.

In order to study evolutionary relationships among different members of the Brassicaceae, Lysak *et al.* (2005) performed chromosome painting with an almost 9 Mbp long BAC contig from *A. thaliana* chromosome 4 to trace homoeologous chromosome regions in 21 different species of the family Brassicaceae, including species representing *Brassica* crops. Homoeologues were identified in all three *Brassica* amphidiploids in six copies corresponding to the *Arabidopsis* segment, whereby rearrangements caused by inversions or translocations could be observed in the homoeologous copies within the *Brassica* genomes. Phylogenetic studies based on comparative sequencing of conserved genes indicated that species containing three or six copy pairs descended from a common hexaploid ancestor with basic genomes similar to that of *Arabidopsis*. The presumed hexaploidization event was shown to have occurred after the *Arabidopsis–Brassiceae* split, between 7.9 and 14.6 Mya.

Interspecific hybridization and homoeologous chromosome pairing

One strategy to broaden the genetic basis of oilseed rape breeding material is the production of resynthesized rapeseed by crossing the original ancestors, *B. oleracea* and *B. rapa*. This has the potential not only to increase genetic variability with a view to hybrid breeding, but also to broaden the genetic base with respect to pest and disease resistances in the narrow gene pool of modern oilseed rape. Interspecific and intergeneric incompatibility barriers can be successfully overcome in crosses between *Brassica* crop species and their relatives by embryo rescue techniques or protoplast fusion. In some cases resynthesized rape forms subjected to backcrossing with elite breeding material have resulted in successful release of cultivars carrying novel resistance genes from the diploid donor species. Synthetic *Brassica* polyploids can also offer important insights into the genetic and epigenetic changes that occur during polyploidization. For example, Song *et al.* (1995) reported extensive and rapid genome change in the form of loss or gain of restriction fragments in early generations of polyploids generated from interspecific crosses among the A, B and C genome diploid species. The alterations also revealed divergence among different genotypes with the same parental origin, demonstrating the ability of *Brassica* polyploids to generate novel genetic diversity in only a few generations. By analysing the methylation status of a large number of isogenic resynthesized rapeseed lines and their two common parental genotypes, Lukens *et al.* (2006) demonstrated that this high degree of divergence is caused by extensive alteration in DNA methylation patterns after polyploidization, whereas few deletions or insertions could be detected. In other words, polyploidy in *B. napus* is accompanied by considerable regulation of epigenetic changes rather than by major genetic changes.

The observed variability in the degree of homoeologous pairing in different *B. napus* haploids demonstrates the potential for extensive homoeologous recombination after amphipolyploidization among *Brassica* diploids. By analysing different *B. napus* doubled-haploid (DH) mapping populations with co-dominant RFLP markers, Udall *et al.* (2005) identified chromosomal rearrangements that could be classified into *de-novo* homoeologous non-reciprocal translocations (HNRT), pre-existing HNRT and homoeologous reciprocal translocations (HRT). A total of 99 *de-novo* HNRT were identified as duplications of particular chromosomal regions in a small number of lines accompanied by a loss of the corresponding homoeologous region. These *de-novo* HNRT were more prevalent in a population that had a resynthesized *B. napus* as a parent, indicating a higher rate of homoeologous recombination in early generations of new polyploids. Nine pre-existing HNRT were identified by fragment duplication or fragment loss in DH lines from three populations involving natural *B. napus* parents, indicating a segregation of HNRT that already existed in one of the parents. This study suggests that chromosomal rearrangements caused by homoeologous recombination are apparently widespread in *B. napus*, and this phenomenon can be speculated to be also prevalent in the other *Brassica* amphidiploids *B. juncea* and *B. carinata*. One of the most interesting aspects of this from a crop-breeding perspective is that such non-reciprocal translocations can cause changes in allele dosage, which in some cases have been shown to have a demonstrable effect on the additive expression of important agronomic traits such as pathogen resistance (Zhao *et al.* 2005).

Other *Brassica* species and even less closely related genera are also important as potential sources of disease resistance for oilseed rape breeding. A prime example for this is the use of interspecific and intergeneric hybrids as a source for new resistance against blackleg (*Leptosphaeria maculans*), the most serious disease of oilseed rape worldwide. The B-genome

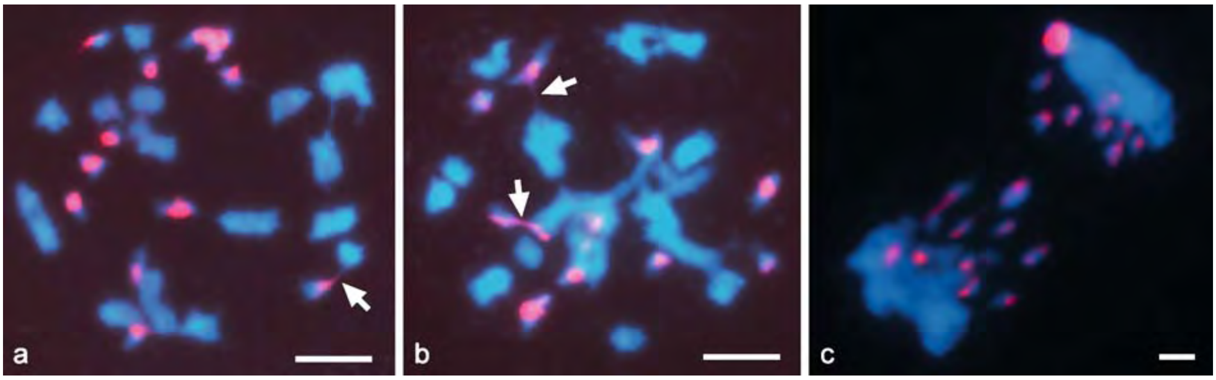


Figure 3. Detection of interspecific homoeologous recombination events (arrows) in sexual progeny of asymmetric *Brassica napus* × *C. abyssinica* hybrids (from Wang *et al.* 2005, used with permission from NRC Research Press, Canada) by GISH analysis in meiotic metaphase I. *Crambe* chromosomes are labelled red with Cy3, whereas non-labelled *B. napus* chromosomes are stained blue with DAPI. (**a, b**) In metaphase I Cy3-labelled chromatin strands (arrows) can be observed between *C. abyssinica* univalents and *B. napus* bivalents. (**c**) At early anaphase I a late disjunction of *C. abyssinica* chromosomes is observed; however, introgressions on *B. napus* chromosomes can still potentially segregate into the daughter cells. Scale bars = 10 µm.

species *B. nigra* and *B. juncea* exhibit strong resistance against blackleg and have been extensively used in an attempt to introgress resistance genes into oilseed rape (Roy 1978, Sacristán & Gerdemann 1986, Sjödin & Glimelius 1989, Chèvre *et al.* 1996, Struss *et al.* 1996). Similarly, related genera including *Sinapis* are also potential donors for transfer of blackleg resistance into *B. napus* via homoeologous recombination (Snowdon *et al.* 2000). Even distantly related crucifer genera can potentially be used for introgression of genes of interest into *Brassica* crop species via sexual or somatic hybridization, for example *Raphanus* (Voss *et al.* 2000, Benabdelmouna *et al.* 2003, Peterka *et al.* 2004), *Lesquerella* (Skarzhinskaya *et al.* 1998) or *Eruca* (Fahleson *et al.* 1997). The primary (intraspecific), secondary (intrageneric) and tertiary (intergeneric) gene pools for *Brassica* vegetable and oilseed crops contain an enormous diversity of species, and hundreds of examples have shown the potential for interspecific or intergeneric hybridizations with *Brassica* crop plants (see Warwick *et al.* 2000 for detailed information). As an example, Wang *et al.* (2005) demonstrated the successful transfer of new allelic variants of the *FAE1* gene controlling erucic acid biosynthesis from the distantly related, triploid oilseed plant *Crambe abyssinica* ($n = 3x = 45$) into *B.* napus by somatic hybridization. Via genomic *in-situ* hybridization (GISH) to meiotic chromosomes it was possible to detect the occurrence of homoeologous pairing between *C. abyssinica* and *B. napus* chromosomes (Figure 3), demonstrating that homoeologous recombination can potentially occur even between chromosomes from distantly related genera. On the other hand, homoeologous introgression of genes of interest even from closely related species of the same genera is not always successful, underlining the fact that pairing does not depend simply on genome homoeology.

The prerequisite for interspecific or intergeneric gene transfer is a mechanism for control of homoeologous chromosome pairing. In wheat the pairing regulation locus *Ph1* (Riley & Chapman 1958) suppresses homoeologous pairing. The successful utilization of alien introgressions in wheat breeding is particularly due to suppression of the effect of this locus. In *B. napus* the chromosomes of the A and C genomes each pair in a bivalent, disomic manner, suggesting a similar suppression of homoeologous pairing in favour of homologous pairing. On the other hand, haploid *B. napus* plants are known to show considerable genetic variation in the degree of chromosome pairing, with common formation of chiasmata among rod-shaped and ring-shaped bivalents along with multivalents (Attia & Röbbelen 1986). By investigating segregation for pairing behaviour at meiotic metaphase I in haploids produced from F1 hybrids between high-pairing and low-pairing lines, Jenczewski *et al.* (2003) found that suppression of homoeologous chromosome pairing in *B. napus* is, as in wheat, largely controlled by a single major gene, coined *PrBn* (*Pairing regulator in Brassica napus*). According to Jenczewski *et al.* (2003), however, the mode of action of *PrBn* appears to be quite different from that of the wheat *Ph1* gene. In particular, *Ph1* suppresses homoeologous pairing at both the haploid and diploid stage, but even *B. napus* haploids with a high level of homoeologous pairing show regular disomic inheritance and normal bivalent chro-

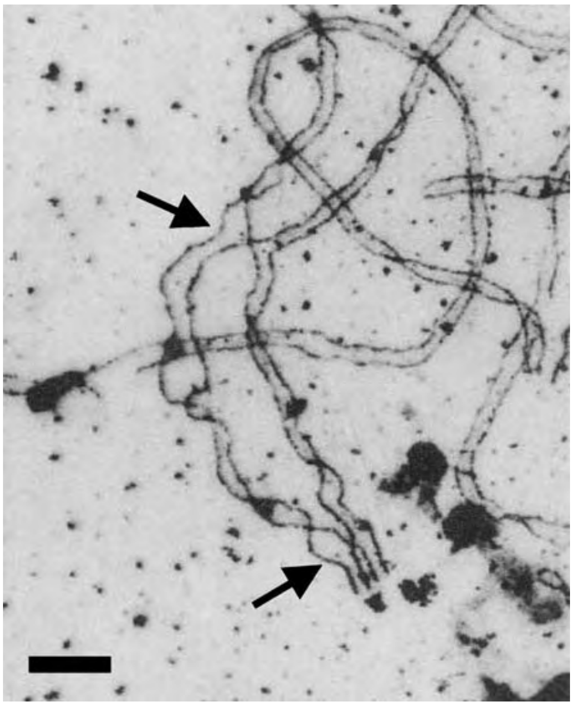


Figure 4. Cytological evidence for a reciprocal translocation between *Brassica napus* linkage groups N7 and N16 (from Fig. 3 in Osborn *et al.* 2003 , kindly provided by John Parker and used with permission from the Genetics Society of America). The image shows meiotic chromosome behaviour in part of a synaptonemal complex spread from an F1 hybrid showing quadrivalent formation between two chromosome pairs. Two synapsis exchange points are indicated by arrows. Comparative genetic mapping of N16 and analysis of marker segregation in two different crosses revealed a homoeologous reciprocal translocation with N7. Scale bar = 2 μm.

mosomes in their diploid form. Furthermore, almost no natural polymorphism has been discovered for *Ph1* in wheat, whereas allelic variability in *PrBn* is clearly detected among high-pairing and low-pairing *B. napus* phenotypes. In other words although *Ph1* is vital for chromosome fertility and stability in wheat (Sánchez-Morán *et al.* 2001), in *B. napus* the suppression of homoeologous pairing by *PrBn* is not essential. Map-based cloning on *PrBn* has the potential to offer considerable new insight into the genetic mechanisms involved in homoeologous pairing in polyploid plant species.

Homoeologous chromosomal translocations can be readily detected in *B. napus* using co-dominant marker systems for which homoeologous loci are assigned to different linkage groups in the genetic map. In this manner Osborn *et al.* (2003) detected a homoeologous reciprocal translocation between *B. napus* linkage groups N7 and N16 using segregating populations of doubled haploid lines. Pairs of homoeologous RFLP loci from the two chromosomes had identical alleles in the parental lines in regions expected to be homoeologous, and an interstitial reciprocal translocation was confirmed by cytological analysis of synaptomenal complexes in F1 hybrids (Figure 4). Although this translocation included approximately one-third of the physical length of the N7 and N16 chromosomes, few recombination events within the region were recovered in the progenies of the hybrids. Higher seed yields in offspring exhibiting the parental configurations of the rearrangement in segregating progenies point to a possible selective advantage of allopolyploidy through the fixation of intergenomic heterozygosity.

Outlook

The growing collection of physical genome resources and genomic sequence data for *Brassica* crop species, combined with the expansive genomic and transcriptomic data from the model crucifer *Arabidopsis*, today enable increasingly detailed annotation and navigation between the *Arabidopsis* and *Brassica* genomes. With the expected completion of the first complete genome sequence of a *Brassica* crop plant in the coming years, the opportunity will be greater than ever to utilize the close relationship between the model plant and its nearest crop relatives, the brassicas, for applied genomics and breeding biotechnology. Of particular interest will be the possibility this allows us to gain considerably more insight into the genetic functionality underlying intricate biochemical pathways and metabolic expression patterns involved in important *Brassica* vegetable and seed compounds – based on primary knowledge from *Arabidopsis* – and perhaps to dissect extremely complex traits such as yield or heterosis into genetic components that are manageable and usable from a practical breeding perspective. Finally, identification of the genetic factors controlling homoeologous recombination in *Brassica* chromosomes may in future enable more flexibility in the targeted introgression of novel germplasm into the gene pools of the crop brassicas, opening the way for a broadening in the genetic diversity that is available for breeding towards a sustainable production in future.

References

Alix K, Heslop-Harrison JS (2004) The diversity of retroelements in diploid and allotetraploid *Brassica* species. *Plant Mol Biol* **54**: 895–909.

Alix K, Ryder C, Moore J, King GJ, Heslop-Harrison JS (2005) The genomic organization of retrotransposons in *Brassica oleracea*. *Plant Mol Biol* **59**: 839–851.

Armstrong SJ, Fransz P, Marshall DF, Jones GH (1998) Physical mapping of DNA repetitive sequences to mitotic and meiotic chromosomes of *Brassica oleracea* var. *alboglabra* by fluorescence in situ hybridisation. *Heredity* **81**: 666–673.

Attia T, Röbbelen G (1986) Cytogenetic relationship within cultivated *Brassica* analyzed in amphihaploids from the three diploid ancestors. *Can J Genet Cytol* **28**: 323–329.

Attia T, Busso C, Röbbelen G (1987) Digenomic triploids for an assessment of chromosome relationships in the cultivated diploid *Brassica* species. *Genome* **29**: 326–330.

Bancroft I (2006) The multinational *Brassica* genome project. *Acta Hort* **706**: 65–67.

Benabdelmouna A, Guéritaine G, Abirached-Darmency M, Darmency H (2003) Genome discrimination in progeny of interspecific hybrids between *Brassica napus* and *Raphanus raphanistrum*. *Genome* **46**: 469–472.

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. *Plant Mol Biol* **42**: 251–269.

Capesius I (1983) Sequence of the cryptic satellite DNA from the plant *Sinapis alba*. *Biochim Biophys Acta* **739**: 276–280.

Chen BY, Simonsen V, Lannér-Herrera C, Heneen WK (1992) A *Brassica campestris-alboglabra* addition line and its use for gene mapping, intergenomic gene transfer and generation of trisomics. *Theor Appl Genet* **84**: 592–599.

Chen BY, Cheng BF, Jörgensen RB, Heneen WK (1997a) Production and cytogenetics of *Brassica campestris-alboglabra* chromosome addition lines. *Theor Appl Genet* **94**: 633–640.

Chen BY, Jørgensen RB, Cheng BF, Heneen WK (1997b) Identification and chromosomal assignment of RAPD markers linked with a gene for seed colour in a *Brassica campestris-alboglabra* addition line. *Hereditas* **126**: 133–138.

Chèvre AM, This P, Eber F *et al.* (1991) Characterization of disomic addition lines *Brassica napus–Brassica nigra* by isozyme, fatty acid, and RFLP markers. *Theor Appl Genet* **81**: 43–49.

Chèvre AM, Eber F, This P *et al.* (1996) Characterization of *Brassica nigra* chromosomes and of blackleg resistance in *B. napus – B. nigra* addition lines. *Plant Breed* **115**: 113–118.

Fahleson J, Lagercrantz U, Mouras A, Glimelius K (1997) Characterization of somatic hybrids between *Brassica napus* and *Eruca sativa* using species-specific repetitive sequences and genomic in situ hybridization. *Plant Sci* **123**: 133–142.

Fukui K, Nakayama S, Ohmido N, Yoshiaki H, Yambe M (1998) Quantitative karyotyping of three diploid *Brassica* species by imaging methods and localization of 45SrDNA loci on the identified chromosomes. *Theor Appl Genet* **96**: 325–330.

Grellet F, Delcasso D, Panabires F, Delseny M (1986) Organization and evolution of a higher plant alphoid-like satellite DNA sequence. *J Mol Biol* **187**: 495–507.

Gupta V, Lakshmisita G, Shaila MS, Jagannathan V, Lakshmikumaran MS (1992) Characterization of species-specific repeated DNA sequences from *B. nigra*. *Theor Appl Genet* **84**: 397–402.

Hallden C, Bryngelsson T, Sail T, Gustafsson M (1987) Distribution and evolution of a tandemly repeated DNA sequence in the family Brassicaceae. *J Mol Evol* **25**: 318–323.

Harbinder S, Lakshmikumaran M (1990) A repetitive sequence from *Diplotaxis erucoides* is highly homologous to that of *Brassica campestris* and *B. oleracea*. *Plant Mol Biol* **15**: 155–156.

Harrison GE, Heslop-Harrison JS (1995) Centromeric repetitive DNA sequences in the genus *Brassica*. *Theor Appl Genet* **90**: 157–165.

Howell EC, Barker GC, Jones GH *et al.* (2002) Integration of the cytogenetic and genetic linkage maps of *Brassica oleracea*. *Genetics* **161**: 1225–1234.

Hu J, Quiros CF (1991) Molecular and cytological evidence of deletions in alien chromosomes for two monosomic addition lines of *Brassica campestris-oleracea*. *Theor Appl Genet* **81**: 221–226.

Iwabuchi M, Itoh K, Shimamoto K (1991) Molecular and cytological characterization of repetitive DNA sequences in *Brassica*. *Theor Appl Genet* **81**: 349–355.

Jackson SA, Cheng ZK, Wang ML, Goodman HM, Jiang JM (2000) Comparative fluorescence *in situ* hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. *Genetics* **156**: 833–838.

Jenczewski E, Eber F, Grimaud A *et al.* (2003) *PrBn*, a major gene controlling homeologous pairing in oilseed rape (*Brassica napus*) haploids. *Genetics* **164**: 645–653.

Koch MA, B. 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.

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in *Brassica*. *Genetics* **144**: 1903–1910.

Lan TH, Del Monte TA, Reischmann KP *et al.* (2000) An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. *Genome Res* **10**: 776–788.

Lenoir A, Cournoyer B, Warwick S, Picard G, Deragon J-M (1997) Evolution of SINE Sl retroposons in Crucifeme plant species. *Mol Biol Evol* **14**: 934–941.

Lim KB, De Jong H, Yang TJ, Park JY, Jin YM, Park BS (2005) Characterisation of rDNAs and tandem repeats in heterochromatin of *Brassica rapa*. *Mol Cells* **19**: 436–444.

Lim YP, Plaha P, Choi SR *et al.* (2006) Toward unraveling the structure of *Brassica rapa* genome. *Physiol Plant* **126**: 585–591.

Lukens L, Zou F, Lydiate D, Parkin I, Osborn T (2003) Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics* **164**: 359–372.

Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L, Osborn T (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. *Plant Physiol* **140**: 336–348.

Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassiceae. *Genome Res* **15**: 516–525.

Maluszynska J, Hasterok R (2005) Identification of individual chromosomes and parental genomes in *Brassica juncea* using GISH and FISH. *Cytogenet Genome Res* **109**: 310–314.

McGrath JM, Quiros CF, Harada JJ, Landry BS (1990) Identification of *Brassica oleracea* monosomic alien chromosome addition lines with molecular markers reveals extensive gene duplication. *Mol Gen Genom* **223**: 198–204.

Morinaga T (1933) Interspecific hybridisation in *Brassica*: 5. The cytology of F1 hybrid of *B. carinata* and *B. alboglabra*. *Jpn J Bot* **6**: 467–475.

Morinaga T (1934) Interspecific hybridisation in *Brassica*: 6. The cytology of *B. juncea* and *B. nigra*. *Cytologia* **6**: 62–67.

Olsson G, Hagberg A (1955) Investigations on haploid rape. *Hereditas* **41**: 227–237.

O'Neill C, Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J* **23**: 233–243.

Osborn TC, Kole C, Parkin IA *et al.* (1997) Comparison of flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. *Genetics* **146**: 1123–1129.

Osborn TC, Pires JC, Birchler JA *et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* **19**: 141–147.

Parkin IA, Gulden SM, Sharpe AG *et al.* (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **171**: 765–781.

Peterka H, Budahn H, Schrader O, Ahne R, Schütze W (2004) Transfer of resistance against the beet cyst nematode from radish (*Raphanus sativus*) to rape (*Brassica napus*) by monosomic chromosome addition. *Theor Appl Genet* **109**: 30–41.

Quiros CF, Ochoa O, Kianian SF, Douches D (1987) Analysis of the *Brassica oleracea* genome by the generation of *B. campestris-oleracea* chromosome addition lines: characterization by isozymes and rDNA genes. *Theor Appl Genet* **74**: 758–766.

Rana D, Van Den Boogaart T, O'Neill CM *et al.* (2004)

Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. *Plant J* **40**: 725–733.

Riley R, Chapman V (1958) Genetic control of the diploid behavior of hexaploid wheat. *Nature* **182**: 713–715.

Roy NN (1978) A study on disease variation in the populations of an interspecific cross of *Brassica juncea* L. × *B. napus* L. *Euphytica* **27**: 145–149.

Sacristán MD, Gerdemann M (1986) Different behavior of *Brassica juncea* and *B. carinata* as sources of *Phoma lingam* resistance in experiments of interspecific transfer to *B. napus*. *Z Pflanzenzüchtg* **97**: 304–314.

Sánchez-Morán E, Benavente E, Orellana J (2001) Analysis of homoeologous-pairing (ph) mutants in allopolyploid wheat. *Chromosoma* **110**: 371–377.

Schelfhout CJ, Snowdon RJ, Cowling WA, Wroth JM (2004) A PCR based B-genome specific marker in *Brassica* species. *Theor Appl Genet* **109**: 917–921.

Schmidt R (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol Biol* **48**: 21–37.

Schmidt R, Acarkan A, Boivin K (2001) Comparative structural genomics in the Brassicaceae family. *Plant Physiol Biochem* **39**: 253–262.

Sharpe AG, Parkin IAP, Keith DJ, Lydiate DJ (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). *Genome* **38**: 1112–1121.

Sibson DR, Hughes SG, Bryant JA, Fitchett PN (1991) Sequence organization of simple, highly repetitive DNA elements in *Brassica* species. *J Exp Biol* **42**: 243–249.

Sjödin C, Glimelius K (1989) Transfer of resistance against *Phoma lingam* to *Brassica napus* by asymmetric somatic hybridization combined with toxin selection. *Theor Appl Genet* **78**: 513–520.

Skarzhinskaya M, Landgren M, Glimelius K (1998) Production of intertribal somatic hybrids between *Brassica napus* L. and *Lesquerella fendleri* (Gray) Wats. *Theor Appl Genet* **93**: 1242–1250.

Snowdon RJ, Winter H, Diestel A, Sacristan MD (2000) Development and characterisation of *Brassica napus–Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*. *Theor Appl Genet* **101**: 1008–1014.

Snowdon RJ, Friedrich T, Friedt W, Köhler W (2002) Identifying the chromosomes of the A and C genome diploid *Brassica* species *B. rapa* and *B. oleracea* in their amphidiploid *B. napus*. *Theor Appl Genet* **104**: 533–538.

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

Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. *Proc Natl Acad Sci USA* **92**: 7719–7723.

Struss D, Bellin U, Röbbelen G (1991) Development of B-genome chromosome addition lines of *B. napus* using different interspecific *Brassica* hybrids. *Plant Breed* **106**: 209–214.

Struss D, Quiros CF, Plieske J, Röbbelen G (1996) Construction of *Basilica* B genome synteny groups based on chromosomes extracted from three different sources by phenotypic, isozyme and molecular markers. *Theor Appl Genet* **93**: 1026–1032.

Tatout C, Warwick S, Lenoir A, Deragon J-M (1999) SINE insertions as clade markers for wild crucifer species. *Mol Biol Evol* **16**: 1614–1621.

Town CD, Cheung F, Maiti R *et al.* (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. *Plant Cell* **18**: 1348–1359.

U N (1935) Genomic analysis of *Brassica* with special reference to the experimental formation of *B. napu*s and peculiar mode of fertilization. *Jpn J Bot* **7**: 389–452.

Udall JA, Quijada PA, Osborn TC (2005) Detection of chromosomal rearrangements derived from homoeologous recombination in four mapping populations of *Brassica napus* L. *Genetics* **169**: 967–979.

Voss A, Snowdon RJ, Lühs W, Friedt W (2000) Intergeneric transfer of nematode resistance from *Raphanus sativus* into the *Brassica napus* genome. *Acta Hort* **539**: 129–134.

Wang YP, Zhao XX, Sonntag K, Wehling P, Snowdon RJ (2005) GISH analysis of BC1 and BC2 progenies derived from somatic hybrids between *Brassica napus* and *Sinapis alba*. *Chromosome Res* **13**: 819–826.

Warwick SI, Francis A, La Fleche J (2000) Guide to wild germplasm of *Brassica* and allied crops (tribe Brassiceae, Brassicaceae), 2nd edn. Agriculture and Agri-Food Canada. http://www.brassica.info/resources/crucifer_genetics/guidewild.htm

Xia X, Selvaraj G, Bertrand H (1993) Structure and evolution of a highly repetitive DNA sequence from *Brassica napus*. *Plant Mol Biol* **21**: 213–224.

Yang TJ, Kim JS, Lim KB *et al.* (2006) An advanced strategy for *Brassica* genome sequencing using comparative genomics with *Arabidopsis*. *Acta Hort* **706**: 73–76.

Zhao J, Udall J, Quijada P, Grau C, Meng J, Osborn T (2005) Quantitative trait loci for resistance to *Sclerotinia sclerotiorum* and its association with a homeologous non-reciprocal transposition in *Brassica napus* L. *Theor Appl Genet* **7**: 1–8.

Ziolkowski PA, Sadowski J (2002) FISH-mapping of rDNAs and *Arabidopsis* BACs on pachytene complements of selected *Brassicas*. *Genome* **45**: 189–197.

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.

2.2 Identifying the chromosomes of the A and C genome diploid *Brassica* species *B. rapa* and *B. oleracea* in their amphidiploid *B. napus*

Snowdon R.J., T. Friedrich, W. Friedt & W. Köhler, 2002. Theor. Appl. Genet. 104: 533-538

Theor Appl Genet (2002) 104:533–538

R.J. Snowdon · T. Friedrich · W. Friedt · W. Köhler

Identifying the chromosomes of the A- and C-genome diploid *Brassica* species *B. rapa* (syn. *campestris*) and *B. oleracea* in their amphidiploid *B. napus*

Received: 13 July 2001 / Accepted: 23 August 2001

Abstract Oilseed rape (*Brassica napus* L.) is an amphidiploid species that originated from a spontaneous hybridisation of *Brassica rapa* L. (syn. *campestris*) and *Brassica oleracea* L., and contains the complete diploid chromosome sets of both parental genomes. The metaphase chromosomes of the highly homoeologous A genome of *B. rapa* and the C genome of *B. oleracea* cannot be reliably distinguished in *B. napus* because of their morphological similarity. Fluorescence in situ hybridisation (FISH) with 5S and 25S ribosomal DNA probes to prometaphase chromosomes, in combination with DAPI staining, allows more dependable identification of *Brassica* chromosomes. By comparing rDNA hybridisation and DAPI staining patterns from *B. rapa* and *B. oleracea* prometaphase chromosomes with those from *B. napus*, we were able to identify the putative homologues of *B. napus* chromosomes in the diploid chromosome sets of *B. rapa* and *B. oleracea*, respectively. In some cases, differences were observed between the rDNA hybridisation patterns of chromosomes in the diploid species and their putative homologue in *B. napus*, indicating locus losses or alterations in rDNA copy number. The ability to reliably identify A and C genome chromosomes in *B. napus* is discussed with respect to evolutionary and breeding aspects.

Keywords *Brassica* · Chromosome · Karyotype · rDNA · Fluorescence in situ hybridisation

Introduction

The genetics of oilseed rape (*Brassica napus* L. 2n=38, genome AACC), now one of the most important oilseed crops worldwide, are complex because it is an amphidiploid composed of two related genomes which themselves are thought to represent ancestral hexaploids (Lagercrantz and Lydiate 1996; Lagercrantz et al. 1996; see also Truco et al. 1996; Quiros 1999). *B. napus* originated from a spontaneous hybridisation of *Brassica rapa* L. (syn. *campestris*; AA, 2n=20) and *Brassica oleracea* L. (CC, 2n=18) and contains the entire diploid chromosome sets of both parental genomes. The chromosomes of the highly homoeologous A genome of *B. rapa* and the C genome of *B. oleracea*, which diverged relatively recently from a common ancestor (Lagercrantz and Lydiate 1996), are difficult to reliably distinguish in *B. napus* because of their small size and morphological similarity (Röbbelen 1960; Olin-Fatih and Heneen 1992). Fluorescence in situ hybridisation (FISH) with ribosomal DNA probes, however, considerably improves the identification of *Brassica* chromosomes, particularly in prometaphase. FISH with 25S ribosomal DNA probes has been used to assist chromosome identification in the three diploid *Brassica* species *B. rapa*, *Brassica nigra* (BB, 2n=16) and *B. oleracea* (Fukui et al. 1998), and to speculate on the evolution of rDNA loci in modern-day *Brassica* amphidiploids (Maluszynska and Heslop-Harrison 1993; Snowdon et al. 1997a; Hasterok and Maluszynska 2000). Moreover, multicolour FISH with 25S and 5S rDNA probes has been applied to reliably identify a number of chromosomes in *B. oleracea* (Armstrong et al. 1998) and *B. napus* (Kamisugi et al. 1998; Schrader et al. 2000; Snowdon et al. 2000). Despite this, the association of individual *B. napus* chromosomes with their homologues in *B. rapa* and *B. oleracea* remains problematic to-date.

Parkin et al. (1995) and Sharpe et al. (1995) demonstrated that the molecular linkage groups of *B. rapa* and *B. oleracea* could be identified in *B. napus* with only relatively minor reorganisation. Most linkage groups were more or less intact, indicating that recombination between A- and C-genome chromosomes has not been extensive in the evolution of modern *B. napus* (Parkin and Lydiate 1997). In a similar study, Axelsson et al. (2000) by com-

Communicated by C. Möllers

R.J. Snowdon (✉) · T. Friedrich · W. Friedt · W. Köhler
Institut für Pflanzenbau und Pflanzenzüchtung,
Justus-Liebig-Universität Giessen
Heinrich-Buff-Ring 26–32, D-35392 Giessen, Germany
e-mail. Rod.Snowdon@agrar.uni-giessen.de
Tel: +49 (641)-9937-438, Fax: 49-(641)-9937-429

parative mapping of re-synthesised and natural *Brassica juncea* demonstrated that very little recombination has occurred between A- and B-genome chromosomes during the evolution of the AABB-genome *Brassica* amphidiploid. This could also be effectively demonstrated by the fact that genomic in situ hybridisation (GISH) allowed the respective A- and B-genome chromosome sets to be clearly distinguished in *B. juncea* (Snowdon et al. 1997b). The same was true for the B and C genomes of *Brassica carinata*, but the A and C genomes of *B. napus*, on the other hand, were too highly homoeologous to label specifically by GISH (Snowdon et al. 1997b). Chromatin condensation patterns have been shown to differ in *B. rapa* and *B. oleracea* prometaphase chromosomes (Cheng et al. 1995; Chen et al. 1997), and using image-analysis equipment Kamisugi et al. (1998) were able to divide *B. napus* chromosomes into those with *rapa*-type and those with *oleracea*-type chromatin condensation. A direct association of *B. rapa* and *B. oleracea* chromosomes with their homologues in *B. napus* was however not made.

Besides providing information on chromosome and genome evolution, the ability to reliably identify A- and C-genome chromosomes is important for the genomic location of transgene inserts in genetically transformed oilseed rape. The well-documented ability of *B. napus* to outcross with *B. rapa* poses a risk for transgene spread from transgenic oilseed rape to wild *B. rapa* populations (see Mikkelsen et al. 1996a). On the other hand, it has long been known that sexual crosses between *B. napus* and *B. oleracea* are extremely difficult to generate with *B. napus* as the pollen donor (Roemer 1935); hence the probability of transgene spread from *B. napus* to the *B. oleracea* genome is by comparison very low. On this basis it has been suggested that transgenes introgressed into the C genome of *B. napus* would be considerably less likely to be introgressed into wild *Brassica* populations than those present on A-genome chromosomes, because *B. napus*×*B. oleracea* hybrids are extremely rare and transgene-carrying C-genome chromosomes will be eliminated from wild *B. napus*×*B. rapa* backcross offspring due to the absence of homologues (Mikkelsen 1996b; Metz et al. 1997; but see also Tomiuk et al. 2000). In combination with the methods described here, localisation of introgressed transgenes by FISH (see Moscone et al. 1996; Ten Hoopen et al. 1999; Snowdon et al. 2001) will assist in the selection of *B. napus* plants with transgene inserts on C-genome chromosomes.

In this study we compared 5S/25S rDNA hybridisation and DAPI staining patterns of *B. rapa* and *B. oleracea* chromosomes with those from *B. napus*, in order to match the A- and C-genome chromosomes of the amphidiploid *B. napus* with their putative homologues in the diploid species.

Materials and methods

Plant material

Seeds from *B. napus* L. subsp. *oleifera* (winter oilseed rape cv “Lirajet”), *B. rapa* L. subsp. *trilocularis* (yellow sarson, accession number YSPb-24) and *B. oleracea* L. var. *capitata* (white cabbage, cv ”Braunschweiger”) were germinated on moist filter paper until the primary roots were 2–3 cm long. Whole seedlings were treated for 2 h at room temperature and 2 h at 4°C in 2 mM of 8-hydroxyquinoline to accumulate metaphases, then fixed in ethanol-acetic acid fixative (3:1). Root tips were excised and chromosome spreads were made from protoplast suspensions using a method described previously (Snowdon et al. 2000, adapted from Schwarzacher et al. 1994).

Fluorescence in situ hybridisation (FISH)

The 25S rDNA subclone pAt27 from *Arabidopsis thaliana* and a 5S rDNA subunit from *Beta vulgaris* were directly labelled by nick translation with the fluorochromes Cy3 and fluorescein-isothiocyanate (FITC), respectively. FISH methods followed Snowdon et al. (2000). After hybridisation, slides were washed at 42°C for 5 min in 2×SSC and 10 min in 0.2×SSC. Chromosomes were counter-stained with DAPI and fluorescence was visualised using a Leica DM-R microscope. At least five to ten well-spread prometaphases were observed for each species and images were obtained using a Cohu 4912 uncooled CCD camera and Leica QFISH software. Individual chromosomes were dissected from the FISH images and karyotypes were constructed using the Adobe Photoshop.

Results

Based on 5S and 25S rDNA hybridisation patterns with DAPI counterstaining, the prometaphase chromosomes of *B. rapa*, *B. oleracea* and *B. napus*, respectively, could be identified and paired. Figure 1 shows the chromosome sets of the three species, aligned according to their FISH and DAPI staining patterns, chromosome size and morphology. The chromosomes of *B. rapa* and *B. oleracea* were numbered according to Fukui et al. (1998), whose chromosome assignment for *B. oleracea* matches that of Armstrong et al. (1998) with the exception of chromosomes 2 and 4 which are exchanged. In accordance with Armstrong (1998) the chromosome with the larger 25S rDNA locus was designated chromosome 7. Chromosome-arm length ratios were not calculated here, because the intense DAPI staining at *Brassica* centromeres only permitted the approximate localisation of the centromere.

In *B. rapa*, four 25S rDNA loci were observed near the centromeres of the metacentric chromosomes 1, 4, 5 and 7. Chromosome 1 could be easily differentiated from chromosomes 4 and 5 by size, whereas the large 25S locus on chromosome 4 was readily distinguishable from the small locus on chromosome 5. The NOR-bearing chromosome 2 contains the fifth largest 25S rDNA locus, which extended over the NOR and chromosome satellite. The large 25S locus located interstitially on chromosome 7 co-localised with a large 5S rDNA locus, giving a strong double-hybridisation signal that allowed this chromosome to be readily identified. Two further 5S loci were localised on the short arms of chromosomes 2 and 10, the largest and the smallest acrocentric *B. rapa* chromosomes, respectively.

As described previously by Armstrong et al. (1998), the *B. oleracea* 5S rRNA genes were located in two

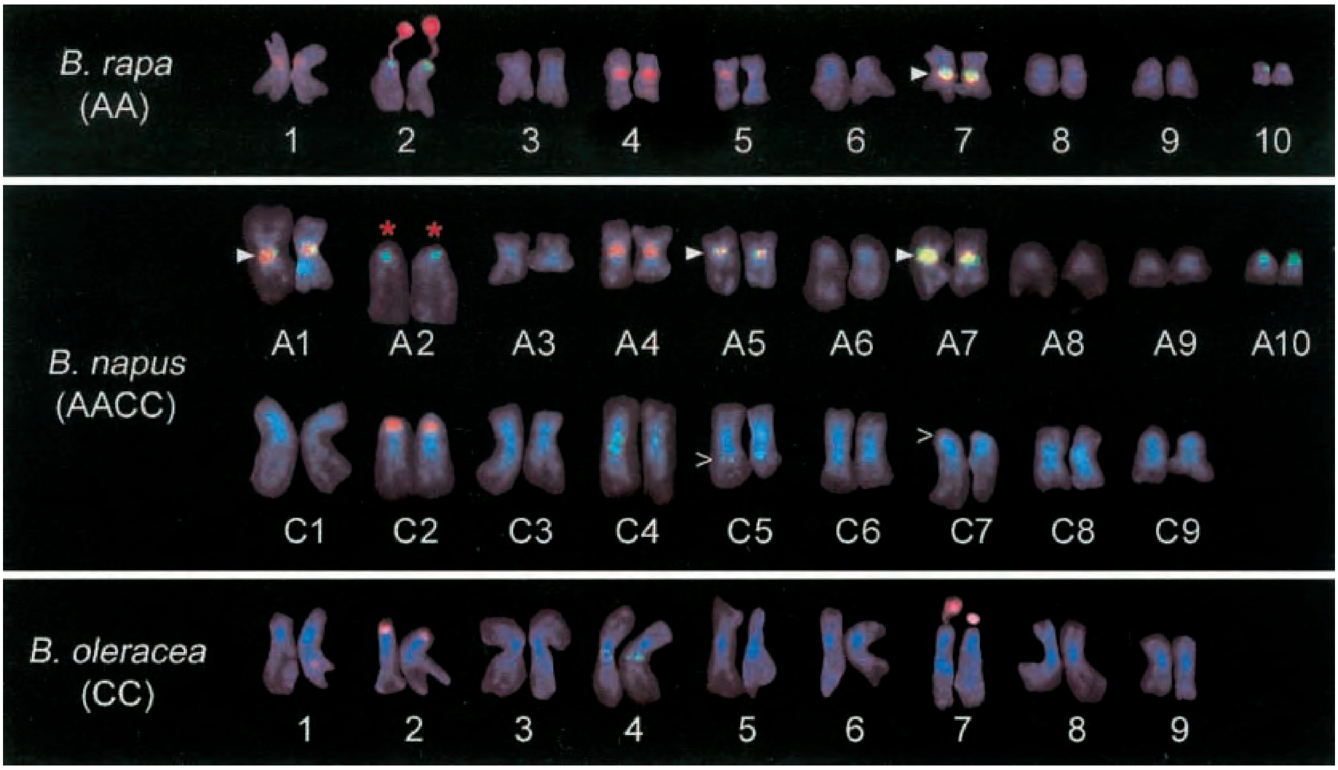


Fig. 1 Karyotypes based on fluorescence in situ hybridisation patterns with 5S (*green*) and 25S (*red*) rDNA probes and DAPI staining (*blue*), for *B. rapa* L., *B. oleracea* L. and their amphidiploid *B. napus* L. Closed arrowheads indicate co-localisation of 5S and 25S loci, whereas *open arrowheads* show a weak 5S locus and a weak 25S locus on *B. napus* chromosomes C5 and C7, respectively. The *red asterisks* represent the position of a large 25S rDNA locus located on the satellite of *B. napus* chromosome A2, which in this spread was lost during chromosome preparation. The *B. napus* karyotype is divided into two sets of chromosomes with differing chromatin condensation patterns resembling, respectively, those of *B. rapa* (A) and *B. oleracea* (C). Each *B. napus* chromosome is aligned and numbered in accordance with its putative homologue in the *B. rapa* or *B. oleracea* genome

closely adjacent loci on the long arm of a single large submetacentric chromosome, designated here and by Fukui (1998) as chromosome 4. The two acrocentric, satellite-carrying *B. oleracea* chromosomes (2 and 7) both had 25S loci at the terminal ends of their short arms and extending over the satellite.

The chromosomal positions of rDNA loci in *B. napus* have been described previously (see Kamisugi et al. 1998; Snowdon et al. 2000) and the *B. napus* karyotype in Fig. 1 derives from the chromosome spread shown by Snowdon et al. (2000). In the present study, however, the large 25S locus covering the chromosome satellite of *B. rapa* chromosome 2 led us to investigate the number of 25S loci in *B. napus* in more detail, because it appeared that this locus was absent from the corresponding chromosome in *B. napus*. After observation of a large number of metaphases it was revealed that chromosome A2 in fact also possesses a chromosome satellite with a large 25S locus adjacent to the 5S locus, corresponding to *B. rapa* chromosome 2. In prometaphase, however, this satellite can be highly distended (Maluszynska and Heslop-Harrison 1993) or alternatively completely lost during chromosome preparation. This was the case for *B. napus* chromosome A2 in the karyotype shown, hence the location of this 25S locus has been indicated by asterisks in Fig. 1.

When the chromosome sets from the diploid species *B. rapa* (A genome) and *B. oleracea* (C genome) were compared to those from *B. napus*, the AACC amphidiploid, chromosomes could be identified in *B. napus* which closely resembled all rDNA-carrying chromosomes from the two diploid species. Chromosomes not containing rDNA loci could also be matched to their putative homologues due to similarities in size, centromere position and DAPI staining pattern. In prometaphase, different chromatin condensation patterns allow *B. rapa* chromosomes to be distinguished from *B. oleracea* chromosomes (Cheng et al. 1995; Chen et al. 1997). Hence, the *B. napus* karyotype could be clearly divided into ten A-genome (*B. rapa*) and nine C-genome (*B. oleracea*) chromosomes, as described previously by Kamisugi et al. (1998). Moreover, based on rDNA hybridisation and DAPI staining patterns, the chromosomes of *B. napus* could be aligned with their putative homologues in the respective diploid species (Fig. 1), and were designated A1–A10 and C1–C9 to reflect their putative origin from either *B. rapa* or *B. oleracea*, respectively.

Minor differences were observed between the rDNA hybridisation patterns of some of the presumed homologues, with respect to signal strengths (indicating differences in copy number) or absence of loci. Details of these discrepancies are listed in Table 1.

Table 1 Differences observed in 5S and 25S rDNA occurrence and signal strength between chromosomes of *B. napus* (genome AACC) and their putative homologues in *B. rapa* (AA) and *B. oleracea* (CC)

rDNA locus	*B. napus* (genome AACC)		Diploid genome homologue	
	Chromosome	FISH signal strength	*B. rapa* (AA)	*B. oleracea* (CC)
5S	A1	Weak	Absent	
5S	A5	Low	Absent	
5S	A10	Strong	Weak	
5S	C5	Low		Absent
25S	A1	Strong	Weak	
25S	C7	Weak		Strong

Discussion

Metaphase chromosomes from *Brassica* species are extremely difficult to reliably identify in standard cytological preparations because they are very small and normally highly condensed. They can be more-readily identified in prometaphase, but because heterochromatin is located almost solely in the pericentromeric regions of *Brassica* chromosomes few useful cytogenetic landmarks are available using classical banding techniques (Olin-Fatih and Heneen 1992). Because of this, the ability to recognise chromosomes of the highly homoeologous diploid genomes of *B. rapa* and *B. oleracea* in their amphidiploid *B. napus* has remained elusive to-date.

Various comparative mapping studies (e.g. Parkin et al. 1995; Sharpe et al. 1995; Parkin and Lydiate 1997) have demonstrated that no more than minor recombination has occurred between the A- and C-genome linkage groups since the spontaneous origin of *B. napus*. This implies that the diploid karyotypes should be largely intact in the amphidiploid, as has been found by GISH to be the case in *B. juncea* and *B. carinata*, the amphidiploids that contain the *Brassica* B genome (Snowdon et al. 1997b). The high level of homoeology between the A- and C-genomes prevents their differentiation by GISH, however (Snowdon et al. 1997b), and might be expected to promote intergenomic recombination. Here we provide molecular cytogenetic evidence that the physical karyotypes of *B. rapa* and *B. oleracea* are largely intact in *B. napus*.

In this study we were able to identify and distinguish the A- and C-genome chromosomes in *B. napus* by multicolor FISH with rDNA probes in combination with DAPI staining. The resulting karyotype for *B. napus* closely resembles that produced by Kamisugi et al. (1998), whereby the inclusion of 5S rDNA loci aided considerably in comparing *B. napus* chromosomes to those of *B. rapa* and *B. oleracea*. Schrader et al. (2000) note that such a comparison should preferably be performed in meiotic prophase or mitotic prometaphase, where a higher resolution of FISH signals is possible than in *Brassica* mitotic metaphase. This was confirmed in the present study, where mitotic prometaphase chromosomes gave clear, reproducible hybridisation signals that could be used effectively for chromosome identification and interspecies comparisons. All rDNA-carrying chromosomes in the diploid species had a characteristic 5S and/or 25S rDNA hybridisation pattern closely resembling a putative homologue with a similar centromere position and chromosome size in the amphidiploid. The remaining chromosomes could be identified by the DAPI staining pattern, chromosome size and morphology, and could be allocated to either the A or C genome due to differences in DAPI staining patterns caused by differential chromatin condensation (Cheng et al. 1995; Chen et al. 1997). Most chromosomes not containing rDNA could also be matched to a putative homologue in *B. napus* with a similar DAPI staining pattern and chromosome morphology.

In three cases *B. napus* possessed a 5S rDNA locus that was not present on the putative homologues in the diploid species, and strong variation in signal strength was apparent for a further 5S locus and two 25S loci (see Table 1). Losses or additions of rDNA loci and alterations in copy number are quite common in polyploid plant genomes (e.g. Dubcovsky and Dvorak 1995; Osuji et al. 1998; Taketa et al. 1999), and have been reported previously in *Brassica* amphidiploids (Maluszynska and Heslop-Harrison 1993; Snowdon et al. 1997a). In *Brassica*, differences are often observed between the number of 25S rDNA loci in different subspecies or varieties of the same species. For example, a third, interstitial 25 S locus is present in *B. oleracea* var. *alboglabra* (Armstrong et al. 1998) and various studies report the presence of either six (Malusynska and Heslop-Harrison 1993; Schrader et al. 2000) or seven (e.g. Kamisugi et al. 1998) loci in *B. napus*. The latter may arise from differences in hybridisation conditions in different laboratories, which can affect the resolution of small loci in particular. In any case, the *B. rapa* and *B. oleracea* varieties investigated in the present study do not represent the original parental genotypes of modern oilseed rape, hence minor discrepancies in rDNA locus and copy number are not unexpected. On the other hand, it was shown that on the whole the *B. rapa* and *B. oleracea* chromosome sets are largely conserved in *B. napus*, indicating that the overall karyotype structures of the respective A and C genomes have not changed excessively during the course of *B. napus* evolution.

Comparative analysis of the *Brassica* A, B and C genomes has unveiled a remarkably high degree of intergenomic similarity and intragenome duplication, suggesting that the three genomes each possess full but rearranged copies of what is now thought to be a hexaploid ancestral genome (Lagercrantz and Lydiate 1996). Through chromosome fusions and rearrangements, the genome of the *Brassica* predecessor has been reorganised into 8, 9 and 10 chromosomes in *B. nigra, B. oleracea* and *B. rapa*, respectively. The results presented here, and by Parkin et al. (1995), suggest that, despite their high homoeology, the evolutionary reorganisation of the respective diploid genomes has been sufficiently extensive and divergent to result in physically distinct chromosome sets that have kept intergenomic recombination to a minimum, not only in *B. carinata* (Snowdon et al. 1997b) and *B. juncea* (Snowdon et al. 1997b; Axelsson et al. 2000) but also between the A and C genomes in *B. napus*. The distinct A- and C-genome chromosomes *in B. napus* mean that the narrow genetic variability present in *B. napus* breeding material can be expanded by interspecific crosses with *B. rapa* and *B. oleracea*, which can be valuable sources for resistance (e.g. Crouch et al. 1994; Scholze and Hammer 1998) or to increase heterotic potential (Seyis et al. 2001).

Although the presence of rDNA loci as physical markers provided convincing evidence for matching chromosome homologues between species, the identification of homologues in the absence of hybridisation signals still needs to be viewed with some caution. For example, *B. rapa* chromosomes 8 and 9 were distinguished by the presence of a DAPI-stained region at the distal end of the long arm of chromosome 8. Chromosomes A8 and A9 in *B. napus*, on the other hand, show extremely similar DAPI staining and could be distinguished only by the larger long arm of A8. Therefore, further markers will be necessary before A8 and A9 can be confirmed as the homologues of *B. rapa* chromosomes 8 and 9, respectively. Similarly, the absence of markers on six *B. oleracea* chromosomes means that the identification of their putative homologues in the *B. napus* C genome can presently only be based on similarities in chromosome morphology. In order to generate physical markers to positively identify all *B. oleracea* chromosomes, Armstrong et al. (2001) are developing FISH probes derived from anchored BAC clones with which the physical karyotype of *B. oleracea* can be aligned with its molecular-marker linkage groups. Such markers should also enable confirmation of *B. oleracea* chromosome homologues in *B. napus*.

As further cytogenetic markers become available, the accuracy of chromosome identification will increase and it should also become possible to physically identify regions of interest on specific chromosomes, delivering much more detailed information about A- and C-genome chromosome structure. This will enable the accurate designation of transgene introgressions to A- or C-genome chromosomes or specific chromosomal regions, and will also bring benefits for studies of *Brassica* genome evolution and for physical mapping.

Acknowledgements The authors acknowledge the helpful comments on this manuscript from two anonymous reviewers. The work presented here was funded by a grant from the Deutsche Forschungsgemeinschaft.

References

Armstrong SJ, Fransz P, Marshall DF, Jones GH (1998) Physical mapping of DNA repetitive sequences to mitotic and meiotic chromosomes of *Brassica oleracea* var. *alboglabra* by fluorescence in situ hybridisation. Heredity 81:666–673

Armstrong SJ, Howell EC, Jones GH, Kearsey MJ, Ryder CD, King GJ (2001) Integrating the genetic and physical chromosome maps of *Brassica oleracea*. Proceedings, Plant and Animal Genome IX, January 13–17 2001. http://www.intl-pag.org/pag/9/abstracts/W16_05.html

Axelsson T, Bowman CM, Sharpe AG, Lydiate DJ, Lagercrantz U (2000) Amphidiploid *Brassica juncea* contains conserved progenitor genomes. Genome 43:679–688

Chen BY, Cheng BF, Jorgensen RB, Heneen WK (1997) Production and cytogenetics of *Brassica campestris-alboglabra* chromosome addition lines. Theor Appl Genet 94:633–640

Cheng BF, Heneen WK, Chen BY (1995) Mitotic karyotypes of *Brassica campestris* and *Brassica alboglabra* and identification of the *B. alboglabra* chromosome in an addition line. Genome 38:313–319

Crouch JH, Lewis BG, Mithen RF (1994) The effect of A genome substitution on the resistance of *Brassica napus* to infection by *Leptosphaeria maculans*. Plant Breed 112:265–278

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

Fukui K, Nakayama S, Ohmido N, Yoshiaki H, Yamabe M (1998) Quantitative karyotyping of three diploid *Brassica* species by imaging methods and localization of 45 s rDNA loci on the identified chromosomes. Theor Appl Genet 96:325–330

Hasterok R, Maluszynska J (2000) Nucleolar dominance does not occur in root tip cells of allotetraploid *Brassica* species. Genome 43:574–579

Kamisugi Y, Nakayama S, O'Neil CM, Mathias RJ, Trick M, Fukui K (1998) Visualization of the *Brassica* self-incompatibility S-locus on identified oilseed rape chromosomes. Plant Mol Biol 38:1081–1087

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in *Brassica*. Genetics 4:1903–1910

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine-scale genome collinearity and congruence of genes controlling flowering time. Plant J 9:13–20

Maluszynska J, Heslop-Harrison JS (1993) Physical mapping of rDNA loci in *Brassica* species. Genome 36:774–781

Metz PLJ, Jacobsen E, Nap JP, Pereira A, Stiekema WJ (1997) The impact on biosafety of the phosphinothricin-tolerance transgene in inter-specific *B. rapa*×*B. napus* hybrids and their successive backcrosses. Theor Appl Genet 95:442–450

Mikkelsen TR, Andersen B, Jorgensen RB (1996a) The risk of crop transgene spread. Nature 380:31

Mikkelsen TR, Jensen J, Jorgensen RB (1996b) Inheritance of oilseed rape (*Brassica napus*) RAPD markers in a backcross progeny with *Brassica campestris*. Theor Appl Genet 92:492–497

Moscone EA, Matzke MA, Matzke AJM (1996) The use of combined FISH/GISH in conjunction with DAPI counterstaining to identify chromosomes containing transgene inserts in amphidiploid tobacco. Chromosoma 105:231–236

Olin-Fatih M, Heneen WK (1992) C-banded karyotypes of *Brassica campestris*, *B.oleracea* and *B. napus*. Genome 35:583–589

Osuji JO, Crouch J, Harrison GE, Heslop-Harrison JS (1998) Molecular cytogenetics of *Musa* species, cultivars and hybrids: location of 18S-5.8S-25S and 5S rDNA and telomere-like sequences. Ann Bot 82:243–248

Parkin IAP, Lydiate DJ (1997) Conserved patterns of chromosome pairing and recombination in *Brassica napus* crosses. Genome 40:496–504

Parkin IAP, Sharpe AG, Keith DJ, Lydiate DJ (1995) Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome 38:1122–1131

Quiros CF (1999) Genome structure and mapping. In: Gómez-Campo C (ed) Biology of *Brassica* coenospecies. Elsevier Science, Amsterdam, pp 217–245

Röbbelen G (1960) Beiträge zur Analyse des *Brassica*-genomes. Chromosoma 11:205–228

Roemer W (1935) Fruchtbarkeits- und Vererbungsstudien bei *Brassica*-Artkreuzungen. Z Pflanzenzüchtung 20:377–416

Scholze P, Hammer K (1998) Evaluation of resistance to *Plasmodiophora brassicae*, *Alternaria* and *Phoma* in *Brassicaceae*. Acta Hort 459:363–369

Schrader O, Budahn H, Ahne R (2000) Detection of 5S and 25S rRNA genes in *Sinapis alba*, *Raphanus sativus* and *Brassica napus* by double fluorescence in situ hybridization. Theor Appl Genet 100:665–669

Schwarzacher T, Leitch AR, Heslop-Harrison JS (1994) DNA:DNA in situ hybridization – methods for light microscopy. In: Harris N, Oparka KJ (eds) Plant cell biology: a practical approach. Oxford University Press, Oxford, pp 127–155

Seyis F, Friedt W, Lühs W (2001) Resynthese-Raps als genetische Ressource für die Qualitäts- und Ertragszüchtung. In: Hammer K, Gladis T (eds) Nutzung genetischer Ressourcen – ökologischer Wert der Biodiversität. Zentralstelle für Agrardokumentation und –information (ZADI), Bonn, IGR-Schriftenreihe Band 16, pp 91–112

Sharpe AG, Parkin IA, Keith DJ, Lydiate DJ (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38:1112–1121

Snowdon RJ, Köhler W, Köhler A (1997a) Chromosomal localization and characterization of rDNA loci in the *Brassica* A and C genomes. Genome 40:582–587

Snowdon RJ, Köhler W, Friedt W, Köhler A (1997b) Genomic in situ hybridization in *Brassica* amphidiploids and interspecific hybrids. Theor Appl Genet 95:1320–1324

Snowdon RJ, Friedt W, Köhler A, Köhler W (2000) Molecular cytogenetic localisation and characterisation of 5S and 25S rDNA loci for chromosome identification in oilseed rape (*Brassica napus* L.). Ann Bot 86:201–204

Snowdon RJ, Böttinger P, Pickardt T, Köhler W, Friedt W (2001) Physical localisation of transgenes on *Vicia faba* chromosomes. Chromosome Res 9:607–610

Taketa S, Harrison GE, Heslop-Harrison JS (1999) Comparative physical mapping of the 5S and 18S-25S rDNA in nine wild *Hordeum* species and cytotypes. Theor Appl Genet 98:1–9

Ten Hoopen R, Montijn BM, Veuskens JTM, Oud OJL, Nanninga N (1999) The spatial localisation of T-DNA insertions in petunia interphase nuclei: consequences for chromosome organisation and transgene insertion sites. Chromosome Res 7:611–623

Tomiuk J, Hauser TP, Bagger-Jorgensen R (2000) A- or C-chromosomes, does it matter for the transfer of transgenes from *Brassica napus*. Theor Appl Genet 100:750–754

Truco MJ, Hu J, Sadowski J, Quiros CF (1996) Inter- and intra-genomic homology of the *Brassica* genomes: implications for their origin and evolution. Theor Appl Genet 8:1225–1233

2.3 Behaviour of *Sinapis alba* chromosomes in a *Brassica napus* background revealed by genomic *in situ* hybridization

Wang Y.P., X.X. Zhao, K. Sonntag, P. Wehling & R.J. Snowdon, 2005. Chromosome Res. 13: 819-826

Chromosome Research (2005) 13:819–826
DOI: 10.1007/s10577-005-1017-2

Behaviour of *Sinapis alba* chromosomes in a *Brassica napus* background revealed by genomic *in-situ* hybridization

Y. P. Wang[1], X. X. Zhao[1], K. Sonntag[2], P. Wehling[2] & R. J. Snowdon[3]*
[1]*College of Bioscience and Biotechnology, Yangzhou University, 225009 Yangzhou, China;* [2]*Federal Centre for Breeding Research on Cultivated Plants, Institute of Agricultural Crops, 18190 Groß Lüsewitz, Germany;* [3]*Institute for Plant Breeding and Crop Science I, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany; Tel: +49-641-9937423; Fax: +49-641-993742; E-mail: Rod.Snowdon@agrar.uni-giessen.de*
* *Correspondence*

Received 9 August 2005. Received in revised form and accepted for publication by Pat Heslop-Harrison 1 November 2005

Key words: *Brassica napus* L., genomic *in-situ* hybridization (GISH), hybrid progeny, *Sinapis alba* L.

Abstract

Genomic *in-situ* hybridization (GISH) was applied to study the behaviour of addition chromosomes in first and second backcross (BC) progenies of hybrids between *Brassica napus* sp. napus L. (AACC, 2n = 38) and *Sinapis alba* L. (SS, 2n = 24) produced by electrofusion. With GISH using genomic DNA of *S. alba* as a probe it was possible to clearly distinguish both of the parental genomes and effectively monitor the fate of *S. alba* chromosomes in the BC_1 and BC_2 progenies. GISH analysis confirmed the sesquidiploid genome composition (AACCS) of the BC_1 progenies, which contained 38 chromosomes from *B. napus* and 12 chromosomes from *S. alba*. Genome painting in the pollen mother cells (PMCs) of the BC_1 plants revealed intergenomic association between *B. napus* and *S. alba* chromosomes, whereby a maximum of 4 trivalents between AC and S chromosomes were identified at metaphase I. In the BC_2 progenies, aneuploids with different numbers of additional chromosomes from *S. alba*, ranging from 1 to 7, were confirmed. Three putative monosomic alien addition lines were characterized, and the results are discussed with respect to the potential for intergenomic chromosome recombination.

Introduction

Sinapis alba L. (genome SS, 2n = 24) is phylogenetically close to *Brassica* species and possesses desirable agronomic characteristics such as yellow seed colour, reduced pod shattering and resistance to various diseases including black spot (*Alternaria brassicae* Berk., Hansen & Earle 1997), beet cyst nematode (*Heterodera schachtii* Schm.) and clubroot (*Plasmodiophora brassicae* Wor., Lelivelt *et al.* 1993), as well as tolerance to flea beetles (Bodnaryk & Lamb 1991, Brown *et al.* 2004). It is also tolerant to high temperatures and drought stress (Brown *et al.* 1997), therefore *S. alba* has considerable promise as an alternative cruciferous oilseed crop in areas with short, dry growing seasons such as those found in the semi-arid regions of Western Canada and Australia (Rakow *et al.* 2000). On the other hand, the seed and oil yield are relatively low in comparison to oilseed rape/canola (*Brassica napus* ssp. *napus*, AACC, 2n = 38), therefore the interspecific transfer of these positive traits to *B. napus* is desirable to expand the genetic variability for these traits in the more important crop species. Successful sexual hybridization between *B. napus* and *S. alba* has been reported through embryo rescue or ovary culture (Ripley &

Arnison 1990, Mathias 1991, Lelivelt *et al.* 1993, Chevre *et al.* 1994, Brown *et al.* 1997) and by protoplast fusion (Primard *et al.* 1988, Lelivelt *et al.* 1993, Wang *et al.* 2005b). In our previous work, somatic hybridization was performed between *B. napus* and *S. alba* via electrofusion, and seven somatic hybrids were produced (Wang *et al.* 2005b) with the aim of enriching sources of disease resistance of *B. napus* and introduction of the yellow seed trait from *S. alba*. In order to create a full set of monosomic alien addition lines of *S. alba* in a *Brassica napus* genomic background, and to eliminate undesirable traits, backcrosses with *B. napus* were subsequently carried out.

Genomic *in-situ* hybridization (GISH) enables not only the distinction of the parental chromosomes in a large number of interspecific and intergeneric hybrids, but also the detection of genomic constitution and chromosome rearrangements (Schwarzacher *et al.* 1989, Anamthawat-Jonsson *et al.* 1990, Jacobsen *et al.* 1995, Kamstra *et al.* 1999, Takahashi *et al.* 1997, Stevenson *et al.* 1998, Karlov *et al.* 1999, Ji *et al.* 2004). GISH has also been applied successfully for identification of *Brassica* intergeneric hybrids (Fahleson *et al.* 1997, Skarzhinskaya *et al.* 1998, Snowdon *et al.* 2000, Benabdelmouna *et al.* 2003, Wang *et al.* 2004); however, it is difficult to detect intergenomic rearrangements in *Brassica* chromosomes. Also, due to a concentration of heterochromatin around the centromeric regions, and the extensive intergenomic homoeology among the Brassicaceae, genomic probes often do not hybridize uniformly across the entire length of the chromosomes (Snowdon *et al.* 1997). On the other hand, we have shown that GISH on meiotic preparations can allow the detection of intergenomic recombination between *B. napus* and *Crambe abyssinica* (Wang *et al.* 2004).

The transfer of genes and chromosomes from alien species and genera has contributed a great deal to the improvement of numerous crops in the past (Lim *et al.* 2000), and interspecific hybrids can also be used to generate novel *B. napus* genotypes with genetically diverse polyploidy genome components (Li *et al.* 2004). GISH has proved invaluable in monitoring the fate of alien chromatin through subsequent generations of wide hybrids (Raina & Rani 2001). In order to establish the number of *B. napus* and *S. alba* chromosomes, and to determine whether homoeologous recombination has occurred in the BC_1 and BC_2 plants, mitotic and meiotic GISH was performed in plants of the BC_1 and BC_2 progenies from hybrids between *B. napus* and *S. alba*.

Materials and methods

Plant materials

Somatic hybrids obtained by electrofusion of *Brassica napus* L. cv. 'Maplus' (AACC, 2n = 38) protoplasts with *Sinapis alba* L. protoplasts have been described previously (Wang *et al.* 2005b). Sexual progeny of the hybrid was obtained by backcrossing twice with *B. napus* cv. 'Maplus'. Root tips from seeds harvested from BC_1 and BC_2 plants were used for mitotic GISH analysis, whereas flower buds of the BC_1 and BC_2 plants were used as experimental material for meiotic GISH. The crossing scheme for the development of hybrid progeny is shown in Figure 1. All plants were grown in the greenhouse.

Chromosome preparation

For study of mitotic metaphase complements, the root tips of young seedlings from BC_1 and BC_2 seeds were collected, and the whole seedlings were pretreated in 2 mM 8-hydroxyquinoline for 2 h at 25°C followed by 2 h at 10°C. Material was fixed in Farmer's solution (acetic acid: ethanol = 1:3) and stored at −20°C until use. Flower buds from BC_1 and BC_2 plants were fixed directly in Farmer's solution. Both root tips and anthers were incubated in an enzyme mixture containing 2% (v/w) cellulase and 20% (v/v) pectinase in 4 mmol/L citrate buffer (pH 4.8) for about 1.5–2 h at 37°C and subjected to a 45 min treatment in 75 mmol/L KCl. Subsequently

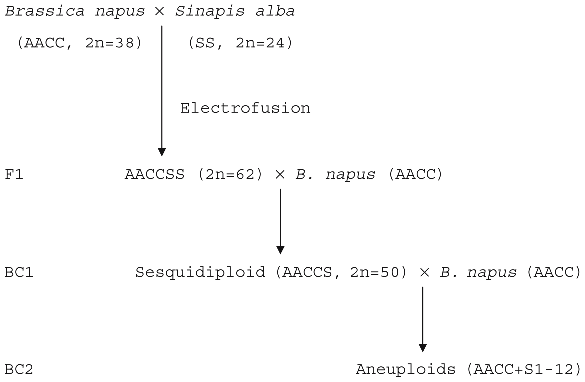


Figure 1. Crossing scheme for the development of hybrid progeny.

each tip or anther was transferred to chilled slides directly using a pipette, and 60% acetic acid was added to clear the cytoplasm followed by washing with ice-cold Farmer's solution to spread the cells on the slides. Slides were air-dried before further use.

Probe preparation and in-situ *hybridization*

Total genomic DNA was extracted from young leaves of *B. napus* and *S. alba* plants using the DNeasy Plant Maxi Kit (Qiagen, Germany). Genomic DNA of *S. alba* was labelled with fluorescein-12-dUTP using a nick-translation kit (Catalogue No. 976776, Roche, Germany) according to the manufacturer's instructions. To prevent non-specific intergenomic cross-hybridization, a 30-fold excess of sheared genomic DNA from *B. napus* was added to the hybridization solution. The DNA was sheared by autoclaving (5 min, 1 bar), yielding fragments of around 300–500 bp in size. Labelled probe and chromosomes were denatured simultaneously on cleaned microscope slides at 80°C for 4 min and hybridized overnight at 37°C. After hybridization the slides were washed at 42°C for 5 min each in 2 × SSC and 0.4 × SSC, respectively. Chromosomes were counterstained with propidium iodide (PI) and fluorescence was visualized using an Olympus BX51 microscope. At least 5 cells were observed for each preparation. Photographs were taken using a computer-assisted cooled charge-coupled device (CCD) camera and images were merged with Image-Pro Plus Version 5.0 software.

Results

GISH analysis of BC_1 plants

Using genomic DNA of *S. alba* as a probe, the *B. napus* and *S. alba* genomes were clearly distinguished in the hybrid progeny. As expected, the GISH studies showed that the chromosome constitution in the seeds harvested from F_1 hybrids via backcrosses with *B. napus* (BC_1) was 2n = 50 (AACCS, sesquidiploid), i.e., 38 chromosomes of *B. napus* origin were strongly painted in red, while 12 chromosomes of *S. alba* origin fluoresced in yellow (Figure 2a). We examined the possibility of meiotic pairing between *B. napus* (AC) and *S. alba* (S) chromosomes in the BC_1 plants. At diakinesis I stage of the pollen mother cells (PMCs), *B. napus* chromosomes formed 19 normal bivalents, while the 12 *S. alba* chromosomes formed univalents and no intragenomic pairing was observed. However, chromatin association between AC and S chromosomes was detected by GISH to meiotic preparations. At metaphase I, AC chromosomes from *B. napus* aligned on the equatorial plate and paired preferentially, whereby between one and four trivalents between AC and S chromosomes were also formed (Table 1). The example shown in Figure 2b shows eight S chromosomes present as univalents and distributed around the exterior of the PMC, while four S chromosomes form trivalents with AC genome chromosomes and were typically found to be oriented more centrally. At anaphase I the 12 chromosomes from *S. alba* were randomly distributed to both poles. Among 65 PMCs scored, 4.6%, 7.7%, 6.2%, 18.5%, 35.4% and 27.7%, respectively, were counted at the ratios of 1:11, 2:10, 3:9, 4:8, 5:7 and 6:6 (Table 1). The example given in Figure 2c shows separation in a ratio of 5:7, with two chromosomes from *S. alba* exhibiting association with *B. napus* chromosomes (arrow). Figure 2d shows a PMC at telophase I with seven chromosomes from *S. alba*, one of which is combined with an AACC chromosome (arrow).

GISH analysis of BC_2 plants

Seeds harvested from BC_1 plants were also examined by GISH on mitotic root tip preparations. Among 36 seeds tested, all had more than 38 chromosomes, ranging from 39 to 45, and were confirmed as aneuploids. Thirty-eight chromosomes from *B. napus* and 1, 3, 4, 5, 6 and 7 additional chromosomes from *S. alba*, respectively, were observed (Table 2, Figure 2e, 2g). Of the progeny investigated, three were

Table 1. Chromosome behaviour of BC_1 plants at diakinesis I or metaphase I and anaphase I analyzed by GISH.

	No. of trivalents per PMC at diakinesis I or metaphase I				Separation ratio of *S. alba* chromosomes at anaphase I					
Meiosis	1	2	3	4	1:11	2:10	3:9	4:8	5:7	6:6
No. of PMCs scored	12	18	24	11	3	5	4	12	23	18
Frequency (%)	18.5	27.7	36.9	16.9	4.6	7.7	6.2	18.5	35.4	27.7

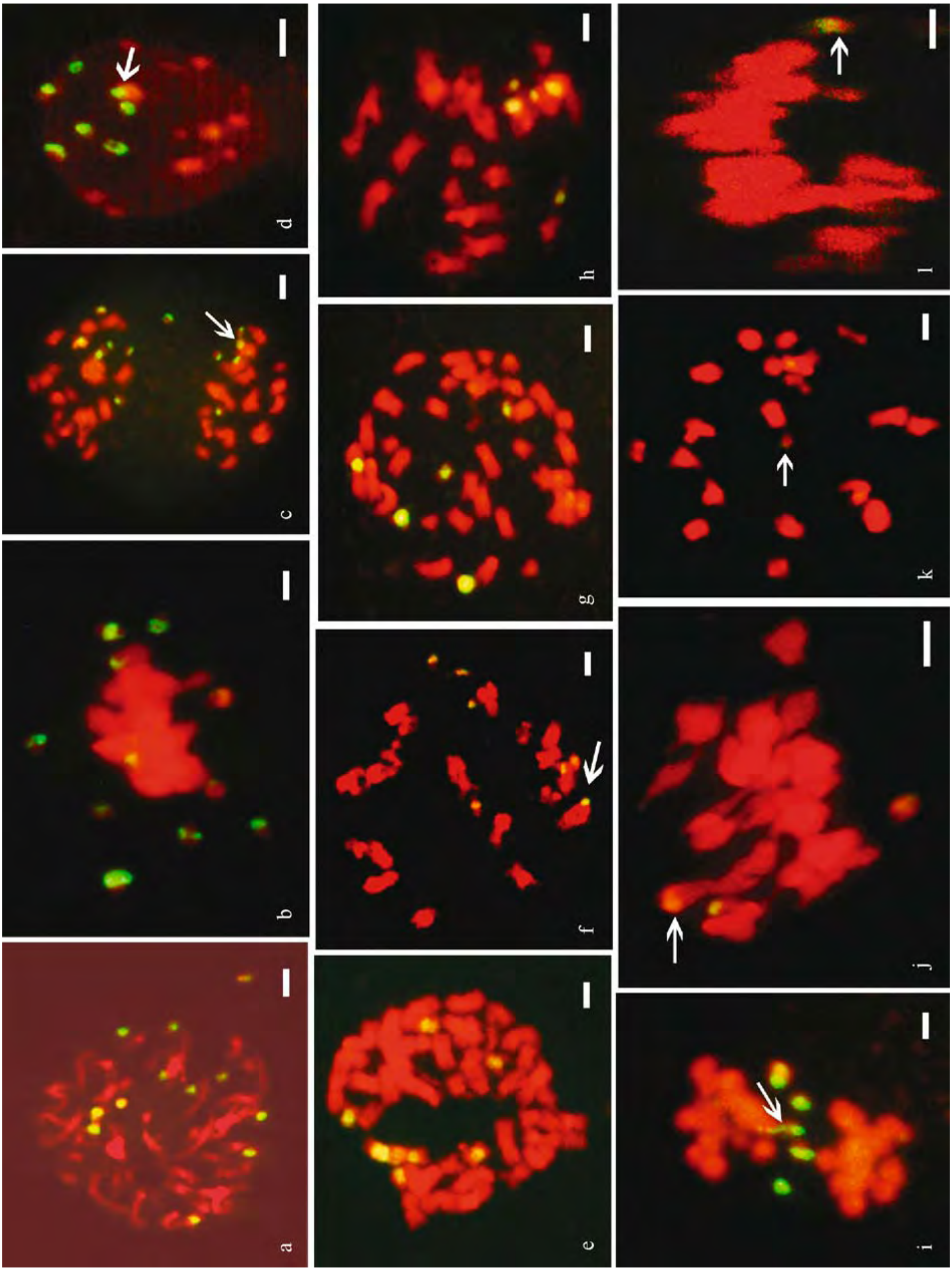

a
b
c
d
e
f
g
h
i
j
k
l

Table 2. Chromosome constitution of somatic hybrids and the BC_1 and BC_2 progenies analysed by GISH.

Genotype	Chromosome number (2n)	Genomic constitution	Genome constitution Chromosomes from *B. napus*	*S. alba*
B. napus	38	AACC	38	
S. alba	24	SS		24
Somatic hybrids (F_1)[a]	62	AACCSS	38	24
BC1 (F1 × *B. napus*)	50	AACCS	38	12
BC2 (BC × *B. napus*)		AACCS(1–12)		
BC2-1			38	1 (3)[b]
BC2-2			38	3 (8)
BC2-3			38	4 (9)
BC2-4			38	5 (7)
BC2-5			38	6 (4)
BC2-6			38	7 (5)

[a] For details see Wang *et al.* (2006).

[b] Figures in parentheses = frequency.

confirmed to contain 39 chromosomes comprising 38 from *B. napus* and one monosomic addition from *S. alba*. Selected BC_2 plants were used for further meiotic GISH analysis. Figure 2f shows a PMC with seven additional chromosomes from *S. alba*, two of which were associated with *B. napus* chromosomes at diakinesis I (arrows). Figures 2h and 2i indicate chromosome behaviour of PMCs with five additional chromosomes of *S. alba*, three of which are possibly associated with AC chromosomes (Figure 2h). Five laggards of *S. alba* were clearly displayed at anaphase I, one of them with chromatin bridge between *S. alba* and *B. napus* (Figure 2i, arrow). Figure 2j shows a PMC with three additional *S. alba* chromosomes, one of which appears to be undergoing chromatin recombination with *B. napus* chromatin (arrow). Potential *B. napus* + *S. alba* monosomic alien addition lines with one *S. alba* chromosome were identified in the BC_2 populations. At diakinesis I, 19 bivalents derived from *B. napus* and 1 univalent derived from *S. alba* were clearly detected (Figure 2k). At metaphase I, one univalent of alien *S. alba* chromosome was observed (Figure 2l).

Morphological characterization of the monosomic alien addition lines

Considerable morphological variation was observed among the BC_1 and BC_2 progenies. Three putative monosomic alien addition lines (MAALs) detected by GISH grew vigorously, were taller than *B. napus* and were differentiated morphologically from each other. Plant BC2-1-1 had very dark green leaves without wax, while plant BC2-1-2 had a thick stem with numerous trichomes and plant BC2-1-3 had a compact shape resembling *S. alba*. The MAALs had a higher pollen fertility, ranging from 82% to 90%. The average seed set of the MAALs by self-

Figure 2. Genomic *in situ* hybridization (GISH) to mitotic and meiotic chromosomes of BC_1 and BC_2 progeny from somatic hybrids between *Brassica napus* ssp. *napus* (AACC, 2n = 38) and *Sinapis alba* L. (SS, 2n = 24). Chromatin of *S. alba* is labelled yellow with FITC, while *B. napus* chromatin is counterstained red with PI. (a–d) BC_1 plants: (**a**) mitotic chromosomes of sesquidiploid BC_1 plants with 38 *B. napus* chromosomes and 12 *S. alba* chromosomes; (**b**) metaphase I in the sesquidiploid hybrid. In this example the *B. napus* chromosomes are oriented on the equatorial plate and eight *S. alba* chromosomes separate to form univalents, while four *S. alba* chromosomes are involved in putative trivalents with *B. napus* chromosomes; (**c**) example of anaphase I showing a 5:7 separation of *S. alba* chromosomes and association of AC and S chromosomes (arrow); (**d**) example of telophase I showing seven chromosomes from *S. alba*, one of which shows association with a *B. napus* chromosome (arrow). (**e–l**) BC_2 plants: (**e, f**) plants with seven additional chromosomes from *S. alba*; (**e**) mitotic cell with 38 *B. napus* chromosomes and seven *S. alba* chromosomes (yellow); (**f**) example of diakinesis showing a PMC with 19II from *B. napus*, five *S. alba* univalents and two putative AC-S trivalents (arrow); (**g–i**) plants with five additional chromosomes from *S. alba*; (**g**) mitotic cell with 38 *B. napus* chromosomes and five *S. alba* chromosomes; (**h**) example of diakinesis in a PMC with 19II *B. napus*, two *S. alba* univalents and three putative AC-S trivalents (arrows); (**i**) example of anaphase I showing five *S. alba* laggards, one of which exhibits a chromatin bridge with a *B. napus* chromosome (arrow); (**j**) metaphase I in a PMC with three additional *S. alba* chromosomes, one of which shows an intergenomic recombination with a *B. napus* chromosome (arrow); (**k, l**) monosomic addition line carrying a single *S. alba* chromosome at diakinesis I (**k**) and metaphase I (**l**) in the form of a univalent. Scale bar represents 2μm.

pollination was 2–3 seeds per pod, in backcrosses with *B. napus* 5–7 seeds per pod were produced. Siliques of all MAALs exhibited a long beak characteristic of *S. alba*. Most of the seeds harvested from the MAALs were larger than that of *B. napus*, and the plant BC2-1-3 produced brown–yellow seed.

Discussion

Genome differentiation by GISH

The results of this study demonstrate the utility of GISH for genome discrimination in hybrid nuclei for an analysis of intergenomic relationships. Furthermore, the transmission and recombination of *S. alba* and *B. napus* chromosomes through meiotic divisions of intergeneric hybrid progeny was also able to be determined by GISH. At meiosis, *B. napus* chromosomes were shown to pair preferentially and some *S. alba* chromosomes formed trivalents with *B. napus* chromosomes, meaning that recombinant chromosomes are present in the nuclei of backcross progeny (Figure 2j). This demonstrates the existence of partial chromosome homology between the genomes of *B. napus* and *S. alba*, as was found previously in somatic hybrids between *B. juncea* and *S. alba* (Gaikwad *et al.* 1996), and therefore the potential for intergenomic recombination. Therefore, the introgression of alien genes from *S. alba* to *B. napus* can be achieved through meiotic cross-over in backcross progenies. This important prerequisite for the integration of agronomically relevant traits from related crucifers into *Brassica* crops is not always fulfilled: In intergeneric hybrids between *B. napus* and the closely related crucifer oil radish (*Raphanus sativus*), for example, radish addition chromosomes were maintained more or less unaltered in the background of the *Brassica* genomes (Peterka *et al.* 2004) and desired resistance traits could not be introgressed from *R. sativus* to the *B. napus* genome despite numerous rounds of backcrossing (Voss *et al.* 2000, Peterka *et al.* 2004).

Chromosome associations in hybrid progeny

The identification of intergeneric or interspecific recombination by GISH in mitotic preparations is questionable in small genomes with a relatively low proportion of medium and highly repetitive DNA families, such as *Brassica* species (Snowdon *et al.* 1997, 2000, Wang *et al.* 2004). In this case, detectable GISH signals are mainly restricted to pericentromeric heterochromatin blocks and sometimes to nucleolus organizers (NORs) where repetitive DNA sequences are clustered. In contrast, our results using GISH to meiotic preparations were able to effectively reveal intergenomic recombination and homoeology. Homoeologous associations between *B. napus* and *S. alba* chromosomes were identified, and in some cases recombinant chromosomes could be clearly detected. The hybrid progenies had a higher seed set when backcrossed with *B. napus*. This result differs from the study of Lelivelt *et al.* (1993), who reported infertility in backcross offspring from somatic hybrids of *B. napus* and *S. alba*. On the other hand, however, GISH can fail to identify very short recombinant segments in interspecific *Brassica* hybrids and the exact size and the position of introgressions can be extremely difficult or impossible to estimate. In such cases, analysis with chromosome-specific markers along the chromosome may assist in more exact chromosomal localization and characterization of the introgressions (Ali *et al.* 2001, Peterka *et al.* 2004).

Utility of the MAALs

S. alba and *B. napus* are closely related species, hence it is possible to transfer individual chromosomes from *S. alba* to *B. napus* to produce monosomic alien addition lines. Obviously, this opens the prospect for establishing a complete set of monosomic *S. alba* alien addition lines within the genomic background of cultivated rapeseed. Our study shows that *S. alba* chromosomes were decreased in subsequent backcrosses with *B. napus*, whereby after only two backcrosses three individuals were detected with only a single *S. alba* chromosome. The identity of individual addition chromosomes can be established in some cases by hybridization with labelled 45S and 5S rDNA probes or other chromosome-specific markers (Schrader *et al.* 2000, Ali *et al.* 2005), or with the help of chromosome-specific PCR markers. This should enable the rapid generation of a complete set of *B. napus* + *S. alba* MAALs as a tool to localize genes of interest controlling relevant agronomic traits in *S. alba*, and to transfer these in a targeted manner into the genome of oilseed rape (Ali *et al.* 2001). MAALs of *B. napus* with alien chromosomes from *S. arvensis* exhibiting resistance to *Leptosphaeria maculans* were successfully obtained and the resistance was successfully introgressed into *B. napus* by backcrossing and selfing (Snowdon *et al.* 2000). Five different multiple *B. napus* + *R. sativus* (oil radish) addition chromosome lines (a–i) were selected by Peterka *et al.* (2004) and used to identify the oil radish chromosome containing genes for beet cyst nematode resistance. More recently, we produced two MAALs of *B. napus* + *C. abyssinica*, and a fertile, stable doubled haploid disomic addition line was obtained via microspore culture (Wang *et al.* 2006). The evidence presented herein of chromosome recombination and association between *B. napus* and *S. alba* demonstrates in principle the feasibility of gene transfer from *S. alba* to *B. napus*. In previous work (Wang *et al.* 2005b), *fae1* gene introgression from *S. alba* to *B. napus* was confirmed by GISH and cleaved amplified polymorphic sequence (CAPS) analysis of the *fae1* gene in F_1 plants of the hybrid and their progenies (F_2 and BC_1). Our eventual goal is to develop a set of *S. alba* introgression lines or MAALs in a *B. napus* background. Many studies show that loci controlling complex traits are numerous, widespread, and intensively interact. The application of MAALs offers more prospects to dissect these loci and ultimately transform them into simple Mendelian factors via backcrosses, greatly facilitating map-based cloning of the genes in the wild relatives (Fridman *et al.* 2004). A full set of *B. napus* + *S. alba* MAALs containing enhanced traits from the donor species would provide us with a potentially powerful tool to identify and transfer genes of interest from *S. alba* to *B. napus* and ultimately to generate stable rapeseed lines containing disease resistance and other relevant traits, such as yellow seed colour and pod shattering resistance, from the donor genome on small chromosome introgressions with minimal genetic drag.

MAALs are also useful for detecting structural variations in homoeologous chromosomes of related species. By creation of a full set of MAALs from *S. alba* to cultivated rapeseed, genes of interest from *S. alba* could be fine-mapped and tagged with closely linked markers. For meaningful utilization of these valuable materials in further introgression breeding, accurate identification of recombinant chromosomes will be necessary using molecular marker analyses and assessment by fluorescence *in situ* hybridization (FISH) with genomic and chromosome-specific DNA as probes (Wang *et al.* 2005a). Furthermore, the production of substitution lines for recombinant segments is required. Similar studies have been conducted in *Festuca–Lolium* hybrids where some of the agronomic traits have been assigned to specific chromosome segments (King *et al.* 1998). Ultimately, this approach enables exploitation of the *S. alba* gene pool for use in rapeseed breeding.

Acknowledgements

This research was supported by grants from the National Natural Science Foundation of China (No. 30571175) and the Natural Science Foundation of Jiangsu Province (BK2005048).

References

Ali SNH, Ramanna MS, Jacobsen E, Visser RGF (2001) Establishment of a complete series of a monosomic tomato chromosome addition lines in the cultivated potato using RFLP and GISH analyses. *Theor Appl Genet* **103**: 687–695.

Ali HBM, Lysak MA, Schubert I (2005) Chromosomal localization of rDNA in the Brassicaceae. *Genome* **48**: 341–346.

Anamthawat-Jonsson K, Schwarzacher T, Leitch AR, Bennett MD, Heslop-Harrison JS (1990) Discrimination between closely related Triticeae species using genomic DNA as a probe. *Theor Appl Genet* **79**: 721–728.

Benabdelmouna A, Guéritaine G, Abirached-Darmency M, Darmency H (2003) Genome discrimination in progeny of interspecific hybrids between *Brassica napus* and *Raphanus raphanistrum*. *Genome* **46**: 469–472.

Bodnaryk RP, Lamb RJ (1991) Mechanisms of resistance to the flea beetle, *Phyllotreta cruciferae* (Goeze), in yellow mustard seedlings, *Sinapis alba* L. *Can J Plant Sci* **71**: 13–20.

Brown J, Brown AP, Davis JB, Erickson D (1997) Intergeneric hybridization between *Sinapis alba* and *Brassica napus*. *Euphytica* **93**: 163–168.

Brown J, McCaffrey JP, Brown DA, Harmon DA, Harmon BL, Davis JB (2004) Yield reduction in *Brassica napus, B. rapa, B. juncea,* and *Sinapis alba* caused by flea beetle (Phyllotreta cruciferae (Goeze) (Coleoptera: Chrysomelidae)) infestation in northern Idaho. *J Econ Entomol* **97**: 1642–1647.

Chevre AM, Eber F, Margale E *et al.* (1994) Comparison of somatic and sexual *Brassica napus–Sinapis alba* hybrids and their progeny by cytogenetic studies and molecular characterization. *Genome* **37**: 367–374.

Fahleson J, Lagercrantz U, Mouras A, Glimelius K (1997) Characterization of somatic hybrids between *Brassica napus* and *Eruca sativa* using species-specific repetitive sequences and genomic *in situ* hybridization. *Plant Sci* **123**: 133–142.

Fridman E, Carrari F, Liu YS, Fernie AR, Zamir D (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* **305**: 1786–1789.

Gaikward K, Kirti PB, Sharma A, Prakash S, Chopra VL (1996)

Cytogenetical and molecular investigations on somatic hybrids of *Sinapis alba* and *Brassica juncea* and their backcross progeny. *Plant Breed* **115**: 480–483.

Hansen LN, Earle ED (1997) Somatic hybrids between *Brassica oleracea* L. and *Sinapis alba* L. with resistance to *Alternaria brassicae* (Berk.) Sacc. *Theor Appl Genet* **94**: 1078–1085.

Jacobsen E, De Jong JH, Kamstra SA, Van den Berg PM, Ramanna MS (1995) Genomic *in situ* hybridization (GISH) and RFLP analysis for the identification of alien chromosomes in the backcross progeny of potato (+) tomato fusion hybrids. *Heredity* **74**: 250–257.

Ji Y, Pertuze R, Chetelat RT (2004) Genome differentiation by GISH in interspeci¢c and intergeneric hybrids of tomato and related nightshades. *Chromosom Res* **12**: 107–116.

Kamstra SA, Ramanna MS, De Jeu MJ, Kuipers GJ, Jacobsen E (1999) Homoeologous chromosome pairing in the distant hybrid *Alstroemeria aurea* × A. inodora and the genome composition of its backcross derivatives determined by fluorescent *in situ* hybridization with species-specific probes. *Heredity* **82**: 69–78.

Karlov GI, Khrustaleva LI, Lim KB, Van Tuyl JM (1999) Homoeologous recombination in 2n-gamete producing interspecific hybrids of *Lilium* (Liliaceae) studied by genomic *in situ* hybridization (GISH). *Genome* **42**: 681–686.

King IP, Morgan WG, Armstead IP *et al.* (1998) Introgression mapping in the grasses. I. Introgression of *Festuca pratensis* chromosomes and chromosome segments into *Lolium perenne*. *Heredity* **81**: 462–467.

Lelivelt CLC, Leunissen EHM, Frederiks HJ, Helsper JPFG, Krens FA (1993) Transfer of resistance to the beet cyst nematode (*Heterodera schachtii* Schm.) from *Sinapis alba* L. (white mustard) to the *Brassica napus* L. gene pool by sexual and somatic hybridization. *Theor Appl Genet* **85**: 688–696.

Li M, Qian W, Meng J, Li Z (2004) Construction of novel *Brassica napus* genotypes through chromosomal substitution and elimination using interploid species hybridization. *Chromosom Res* **12**: 417–426.

Lim KB, Chung JD, Kronenburg BCE, Ramanna MS, Jong JH, Tuyl JM (2000) Introgression of *Lilium rubellum* Baker chromosomes into *L. longiflorum* Thunb.: a genome painting study of the F_1 hybrid, BC_1 and BC_2 progenies. *Chromosom Res* **8**: 119–125.

Mathias R (1991) Improved embryo rescue technique for intergeneric hybridization between *Sinapis* species and *Brassica napus*. *Crucif Newsl* **14/15**: 90–92.

Peterka H, Budahn H, Schrader O, Ahne R, Schütze W (2004) Transfer of resistance against the beet cyst nematode from radish (*Raphanus sativus*) to rape (*Brassica napus*) by monosomic chromosome addition. *Theor Appl Genet* **109**: 30–41.

Primard C, Vedel F, Mathieu C, Pelletier G, Chevre AM (1988) Interspecific somatic hybridization between *Brassica napus* and *Brassica hirta* (*Sinapis alba* L.). *Theor Appl Genet* **75**: 546–552.

Raina SN, Rani V (2001) GISH technology in plant genome research. *Methods Cell Sci* **23**: 83–104.

Rakow G, Potts D, Raney P, Katepa-Mupondwa F (2000) Designing oilseed crops for the Canadian dry prairie. In: 3rd Int Crop Sci Congr. Book of Abstracts, p 221, session 4C, poster no. 34. Hamburg, Germany.

Ripley VL, Arnison PG (1990) Hybridization of *Sinapis alba* L. and *Brassica napus* L. via embryo rescue. *Plant Breed* **104**: 26–33.

Schrader O, Budahn H, Ahne R (2000) Detection of 5S and 25S rRNA genes in *Sinapis alba, Raphanus sativus* and *Brassica napus* by double fluorescence *in situ* hybridization. *Theor Appl Genet* **100**: 665–669.

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

Skarzhinskaya M, Landgren M, Glimelius K (1998) Production of intertribal somatic hybrids between *Brassica napus* L. and *Lesquerella fendleri* (Gray) Wats. *Theor Appl Genet* **93**: 1242–1250.

Snowdon RJ, Köhler W, Friedt W, Köhler A (1997) Genomic *in situ* hybridization in *Brassica* amphidiploids and interspecific hybrids. *Theor Appl Genet* **95**: 1320–1324.

Snowdon RJ, Winter H, Diestel A, Sacristán MD (2000) Development and characterization of *Brassica napus–Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*. *Theor Appl Genet* **101**: 1008–1014.

Stevenson M, Armstrong SJ, Ford-Lloyd BV, Jones GH (1998) Comparative analysis of crossover exchanges and chiasmata in *Allium cepa* × *fistulosum* after genomic *in situ* hybridization (GISH). *Chromosom Res* **6**: 567–574.

Takahashi C, Leitch IJ, Ryan A, Bennett MD, Brandham PE (1997) The use of genomic *in situ* hybridization (GISH) to show transmission of recombinant chromosomes by partially fertile bigeneric hybrid, *Gasteria lutzii* × *Aloe arstata* (Aloaceae), to its progeny. *Chromosoma* **105**: 342–348.

Voss A, Snowdon RJ, Lühs W, Friedt W (2000) Intergeneric transfer of nematode resistance from *Raphanus sativus* into the *Brassica napus* genome. *Acta Hortic* **539**: 129–134.

Wang YP, Snowdon RJ, Rudloff E, Wehling P, Friedt W, Sonntag K (2004) Cytogenetic characterization and fae1 gene variation in progenies from asymmetric somatic hybrids between *Brassica napus* and *Crambe abyssinica*. *Genome* **47**: 724–731.

Wang J, Xiang F, Xia G (2005a) *Agropyron elongatum* chromatin localization on the wheat chromosomes in an introgression line. *Planta* **221**: 277–286.

Wang YP, Sonntag K, Rudloff E, Chen J (2005b) Intergeneric somatic hybridization between Brassica napus and *Sinapis alba*. *J Integrative Plant Biol* **47**: 84–91.

Wang YP, Sonntag K, Rudloff E, Wehling P, Snowdon RJ (2006) GISH analysis of disomic *Brassica napus–Crambe abyssinica* chromosome addition lines produced by microspore culture from monosomic addition lines. *Plant Cell Rep* (in press). DOI: 10.1007/s00299-005-0031-3.

3 Analysis and exploitation of genetic diversity in *Brassica napus*

The following two articles describe the analysis and utilisation of genetic diversity in winter-type and spring-type *Brassica napus* including oilseed, fodder and vegetable rape varieties of diverse geographical origin. In the first article, genome-wide simple-sequence repeat (SSR) markers were used to estimate genetic diversity in *B. napus* genebank collections. The second article describes quantitative trait analysis of an important seed quality trait with a complex mode of inheritance using population structure-based allele-trait association analysis, including an innovative approach to marker development using *Brassica*-Arabidopsis sequence synteny. This paper was the first description of association mapping techniques for analysis of a quantitative trait in *B. napus*, and is also the first to use synteny information to develop potentially gene-linked markers in oilseed rape.

3.1 Surveying genetic diversity in the *Brassica napus* gene pool using SSR markers

Hasan M., F. Seyis, A.G. Badani, J. Pons-Kuhnemann, W. Lühs, W. Friedt & R.J. Snowdon, 2006. Genet. Res. Crop Evol. 53: 793–802

Genetic Resources and Crop Evolution (2005)
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, 2*n* = 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′-AGGGTTTTCCCAGTCACGAC-GTT-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 arithmetic 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

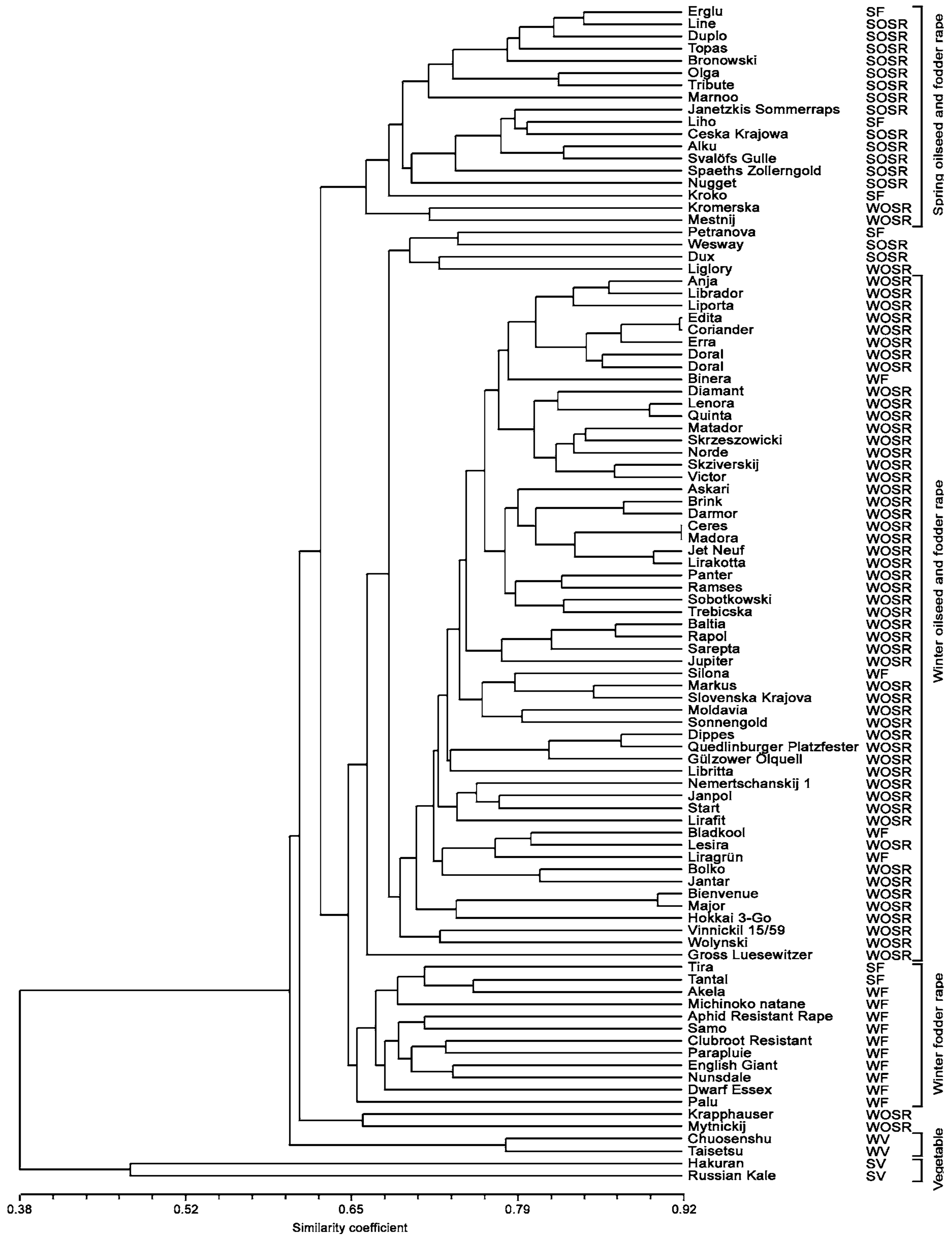


Figure 1. Phenogram showing DICE genetic similarity for a diverse set of 96 oilseed rape accessions including winter-type (W) and spring-type (S) fodder (F), vegetable (V) and oilseed rape (OSR) *Brassica napus* L. accessions from European genebanks (see Table 1), revealed by UPGMA clustering based on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

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

Hasan M., W. Friedt, N.M. Freitag, K. Link, J. Pons-Kühnemann & R.J. Snowdon, 2008. Theor. Appl. Genet. 116: 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

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

---

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

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 $P \leq 0.05$)

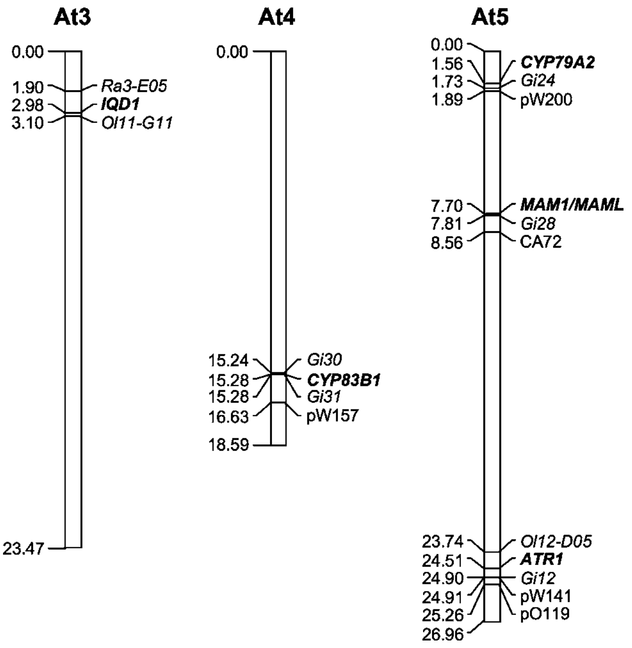


Fig. 1 Chromosomal positions in *Arabidopsis thaliana* (numbers in Mbp) of orthologs for potentially gene-linked *Brassica* SSR markers (*italics*) in comparison to potential candidate genes for glucosinolate biosynthesis (***bold italics***) and RFLP markers (*non-italic, non-bold*) located at major 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

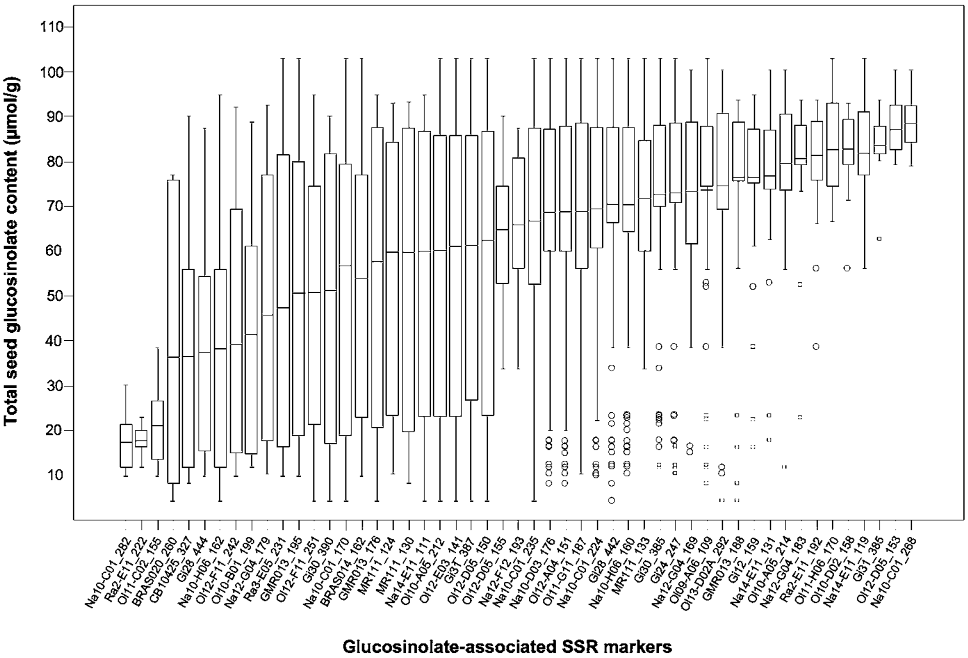


Fig. 2 *Boxplots* showing distributions of total seed glucosinolate content within 94 genetically diverse *B. napus* gene bank accessions for 51 SSR marker alleles with significant association ($P \leq 0.05$) to glucosinolate content. Boxes cover the inter-quartile range around the mean (*horizontal 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

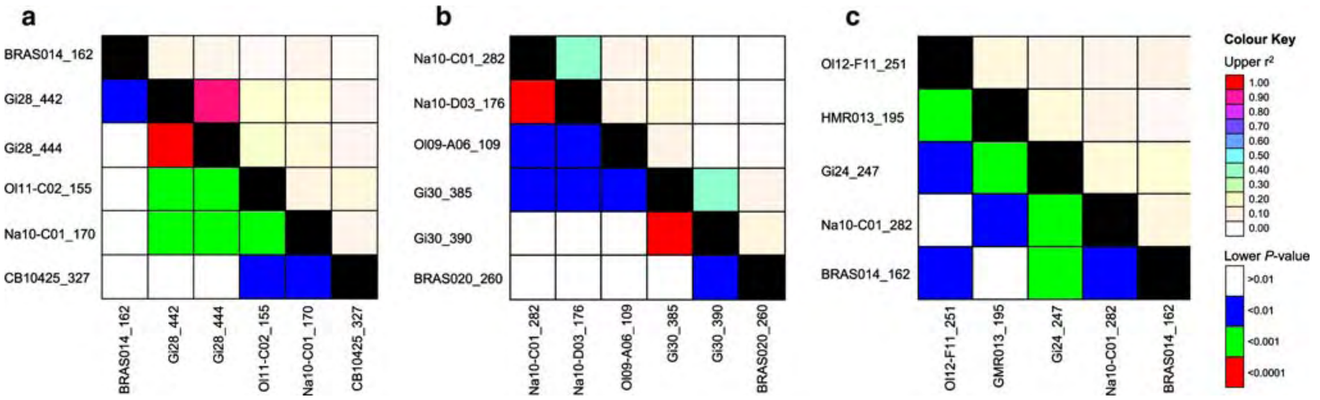


Fig. 3 Linkage disequilibrium (*LD*) around the gene-linked SSR markers **a** Gi28, **b** Gi30, and **c** Gi24 in the 94 *B. napus* gene bank accessions. Cells above the diagonal show the squared allele frequency correlation r^2, while the cells below the diagonal represent the significance level of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

4 Interspecific hybridisation for oilseed rape resistance breeding

The two articles that comprise this chapter give important practical examples for the generation of interspecies hybrids via tissue culture techniques. This helps to broaden the *B. napus* gene pool for resistance to major diseases of oilseed rape. The first paper describes resynthesised *B. napus* genotypes, produced by embryo rescue-assisted interspecific hybridisation of *B. rapa* and *B. oleracea*, that exhibit novel resistance against the soil-borne fungal pathogen *Verticillium longisporum*. Verticillium wilt caused by this pathogen is today emerging as a significant new threat to oilseed rape production in Europe, because fungicide treatment cannot control the disease and current varieties possess only limited resistance. The second article describes oilseed rape lines containing novel resistance genes against blackleg disease caused by *Leptosphaeria maculans*, the most important disease of rapeseed on a worldwide scale. In this case the blackleg resistance was introduced from the closely related mustard species *Sinapis arvensis* by recurrent backcrossing of interspecific hybrids that were generated by protoplast fusion with *B. napus*. The value of the new germplasm for resistance breeding is discussed.

4.1 Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridisation

Rygulla W., <u>R.J. Snowdon</u>, C. Eynck, A. von Tiedemann, W. Lühs & W. Friedt, 2007a: Phytopathology 97: 1391-1396

Broadening the Genetic Basis of *Verticillium longisporum* Resistance in *Brassica napus* by Interspecific Hybridization

W. Rygulla, R. J. Snowdon, C. Eynck, B. Koopmann, A. von Tiedemann, W. Lühs, and W. Friedt

First, second, sixth, and seventh authors: Department of Plant Breeding, Research Center for BioSystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, D-35392 Giessen, Germany; and third, fourth, and fifth authors: Department of Crop Sciences, Division of Plant Pathology and Crop Protection, Georg-August-University, Grisebachstrasse 6, D-37077 Göttingen, Germany.

Accepted for publication 16 July 2007.

ABSTRACT

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

Verticillium wilt caused by the vascular fungal pathogen *Verticillium longisporum* is one of the most important pathogens of oilseed rape (*Brassica napus* sp. *oleifera*) in northern Europe. Because production of this major oilseed crop is expanding rapidly and no approved fungicides are available for *V. longisporum*, long-term control of the disease can only be achieved with cultivars carrying effective quantitative resistance. However, very little resistance to *V. longisporum* is available within the gene pool of oilseed rape, meaning that interspecific gene transfer from related species is the only possibility for broadening levels of resistance in current varieties. The amphidiploid species *B. napus* can be resynthesized by crossing the two progenitor species *Brassica oleracea* and *Brassica rapa*, hence resistant accessions of these two diploid species can be used as resistance donors. In this study a total of 43 potential *B. rapa* and *B. oleracea* resistance donors were tested with regard to their reaction to a mixture of two aggressive *V. longisporum* isolates, and resistances from diverse lines were combined by embryo rescue-assisted interspecific hybridization in resynthesized rapeseed lines. Progenies from crosses of the two *B. rapa* gene bank accessions 13444 and 56515 to the *B. oleracea* gene bank accessions BRA1008, CGN14044, 8207, BRA1398, and 7518 showed a broad spectrum of resistance in pathogenicity tests. Of 45 tested resynthesized lines, 41 lines exhibited a significantly higher level of resistance than the moderately *V. longisporum*-tolerant oilseed rape cultivar Express. These lines represent a promising basis for the combination of different resistance resources in new varieties.

Verticillium wilt caused by the fungal pathogen *Verticillium longisporum* (13) is one of the most important pathogens of oilseed rape (*Brassica napus* sp. *oleifera*) in northern Europe. The disease is especially prominent in northern Germany and Sweden, but it has also been reported to occur in Poland, France, southern Russia, and Ukraine (12). The fungus infects oilseed rape by colonizing the xylem vessels. Due to the production of fungal material and host reaction products the vascular system becomes blocked, meaning that heavily infected plants die due to water stress. At the same time microsclerotia are produced in the plant, causing symptomatic blackening of leaf veins or stems. With the decomposition of dead plant material the microsclerotia are released to the soil, where the fungus can survive for many years in this dormant form (19).

Oilseed rape is the most important oilseed crop in Europe and the production area is continuing to rise dramatically due to the high demand for oilseed rape oil, both as a healthy vegetable oil, and a renewable resource for the oleochemical industry. Because no approved fungicides are available for *V. longisporum*, long-term control of the disease can only be sustained with cultivars carrying effective quantitative resistance. Up to 70% yield loss due to disease-induced premature ripening has been observed in individual plants (9), and yield losses in the field can also be severe. Currently available commercial varieties are generally susceptible or exhibit only slight tolerance to *V. longisporum*.

Intensive screening of available *B. napus* germplasm failed to identify resistance resources within the primary gene pool (11). On the other hand, because *B. napus* (genome AACC, $2n = 38$) is an amphidiploid species that is derived from interspecific hybridizations between turnip rape (*Brassica rapa* L., syn. *campestris*; AA, $2n = 20$) and cabbage (*Brassica oleracea* L.; CC, $2n = 18$), the diverse gene pools of these two progenitor species can also be used as potential gene donors to broaden the resistance spectrum of the amphidiploid species. Nothing is known about the inheritance of *Verticillium* resistance in the diploid progenitor species. Due to the self-incompatibility system, self-fertilization is prevented and homozygous lines cannot be produced for inheritance studies.

With the help of embryo rescue techniques it is possible to generate novel, resynthesized (RS) oilseed rape lines comprising the intact genomes of the respective *B. rapa* and *B. oleracea* donor lines. Such RS *B. napus* represents an important source of novel genetic variation for specific traits in oilseed rape. In previous investigations this strategy has been adopted for introduction of resistance genes into *B. napus* against *Turnip yellows virus* (TuYV) (8), *Turnip mosaic virus* (TMV) (22), and also important fungal pathogens including *Phoma lingam* (3) and *Plasmodiophora brassicae* (5). A similar strategy has great promise for *V. longisporum*-resistance breeding, since numerous *B. oleracea* gene bank accessions are known with a high level of resistance to this pathogen, and their transfer to *B. napus* was found to improve the resistance of some of the resulting RS oilseed rape lines (11).

Widespread use of specific resistances from single donor sources induces a strong selection pressure on the pathogen that can result in the resistance being rapidly overcome. This has been shown previously for some of the above mentioned resistances

Corresponding author: R. J. Snowdon
E-mail address: Rod.Snowdon@agrar.uni-giessen.de

doi:10.1094/PHYTO-97-11-1391

derived only from *B. rapa* or *B. oleracea* sources (10,20). It is therefore desirable to combine different resistance resources with potentially different mechanisms to generate long-term resistance, and to avoid the formation of new pathotypes with enhanced aggressiveness. The aim of the present study was to identify new resistance sources against *V. longisporum* among *B. rapa* gene bank material, and to combine these with known *B. oleracea* resistance sources in novel RS *B. napus* lines via sexual interspecific hybridization with the help of embryo rescue. The material that was developed is discussed regarding the potential for resistance breeding in oilseed rape and further investigations in Verticillium research.

MATERIALS AND METHODS

Plant materials. Twenty-nine *B. oleracea* and 12 *B. rapa* gene bank accessions (Table 1) obtained from eight different gene banks and a commercial seed supplier, were screened for their resistance response to *V. longisporum* in order to identify resistant genotypes for combination of potentially different resistances in RS *B. napus*. Parallel to the resistance screening of the diploid progenitors, a set of potentially resistant RS *B. napus* genotypes were generated by interspecific hybridizations. The parents of these crosses were different plants belonging to two *B. rapa* accessions that were chosen based on prior observations which indicated that these accessions could contain potential resistance to *V. longisporum*, and five *B. oleracea* gene bank accessions that were reported by Happstadius et al. (11) to have a high level of resistance to *V. longisporum*. The origins of the diploid accessions used for the interspecific crosses along with the pedigrees of the resulting RS lines are shown in Table 2.

The *B. rapa* genotypes were emasculated and hand-pollinated 2 days later with pollen from the respective *B. oleracea* parent, except in one case where the *B. oleracea* accession was used as the maternal parent and pollinated by *B. rapa*. To enhance the efficiency of the interspecific crossing, the embryo rescue technique was used to regenerate haploid plants in vitro. Immature pods were harvested 2 weeks after pollination and surface sterilized with NaOCl (3%) for 10 min and washed with sterile water. Ovules were dissected and cultivated on Murashige and Skoog (MS) media with sucrose (10 g/liter) (15) in a growth chamber at 20°C, 75% humidity and a day length of approximately 16 h. Ovules that did not germinate after 6 weeks were transferred to new media, while germinated ovules that had already developed into a small plant were transferred to MS media containing saccharose (30 g/liter). Regenerated, vigorous plants were transferred to soil, and cuttings from the haploid hybrids were treated with colchicine (0.05% colchicine, 0.15% dimethyl sulfoxide) for 12 h in the dark in order to obtain amphidiploid *B. napus* plants. All resulting RS plants were self-pollinated for seed multiplication by bagging with super-micro perforated Crispac bags (Baumann Saatzuchtbedarf, Waldenburg, Germany) before the onset of flowering. A total of 45 RS plants derived from 8 different *B. rapa* × *B. oleracea* cross combinations produced sufficient seeds (>50) for the resistance screening.

Resistance screening. Resistance tests were performed using the *V. longisporum* isolates VL 43 and VL 40, which originate from *B. napus* grown in northern Germany (23). Long-term

TABLE 1. Accession numbers and sources of *Brassica rapa* and *B. oleracea* gene bank accessions that were phenotyped for identification of potential resistance to *Verticillium longisporum*

Species	Accession names	Source[a]
B. rapa	13444, 56429, 56515	BAZ
	G454 (Chinese cabbage ‘Granat’), G684 (‘Runde weiße rotköpfige Herbstrüben’)	Cultivars
	YS Pb-24 (Yellow Sarson cultivar)	HAU
	Q3F, ‘Reward’, SWSP, 00 *B. rapa*	NPZ
	03-510001 (‘Salut’), 03-510002 (‘Credit’)	SW
B. oleracea	2398, 2808, 3775, 4250, 4622, 4707, 5085, 5394, 6206, 6704, 7518, 7799, 8207, 10590, 10618	HRI
	BRA504, BRA544, BRA723, BRA809, BRA1008, BRA1355, BRA1398, BRA1427	IPK
	CGN07069, CGN07089, CGN14044	CGN
	1227	Li et al. 2001 (14)
	1428, 2070	NIV

[a] BAZ: Gene bank of the Federal Centre for Breeding Research on Cultivated Plants, Braunschweig, Germany; CGN: Centre for Genetic Resources, Wageningen, Netherlands; HAU: Haryana Agricultural University, Hisar, India; HRI: Horticulture Research International Genetic Resources Unit, Warwick, UK; IPK: Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany; NIV: Gene bank of the NI Vavilov Institute of Plant Industry (VIR), St. Petersburg, Russia; NPZ: Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth, Germany; SW: Svalöf Weibull AB, Svalöf, Sweden.

TABLE 2. Origin of resynthesized oilseed rape lines and progenitor species with resistance to *Verticillium longisporum*

Parental species accession designations	Botanical names
13444	*B. rapa* sp. *pekinensis*
56515	*B. rapa* sp. *chinensis*
8207	*B. oleracea* convar. *acephala*
BRA1398	*B. oleracea* sp. *oleracea* convar. *botrytis* var. *botrytis*
BRA1008	*B. oleracea* convar. *acephala* var. *sabellica*
7518	*B. oleracea* sp. *oleracea* convar. *botrytis* var.*alboglabra*
CGN14044	*B. oleracea* sp. *oleracea* convar. *botrytis* var. *italica*
Resynthesized oilseed rape lines	♀ × ♂
K160.1.1, K160.1.2, K160.6.1, K160.6.2, S160.7.1, K160.7.2, S160.5.1	13444 × 8207
K242.2.1, K242.2.2, S175.5.1, S175.5.2, S237.20.1, S237.23.1	13444 × BRA1398
K147.34.1, K239.42.2, K241.20.2	13444 × BRA1008
K162.2.2	13444 × 7518
K195.1.2, K208.1.2, K208.4.1, K208.4.2, K228.5.1, K228.9.2, K233.2.1, K233.2.2, K233.4.1, K233.4.2, S189.3.1, S219.3.2, S228.8.1, S228.9.1, S228.10.2, S233.1.2, S228.13.1, S228.13.2, S228.21.2, S233.1.1, S233.5.1	56515 × 8207
K199.8.2, K199.16.1, K199.16.2	56515 × BRA1398
K138.1.1	BRA1008 × 56515
K142.2.2, S194.6.1	56515 × CGN14044

storage was performed as conidial suspensions in a concentration of 1 to 3 × 10^6 conidia/ml in Czapek-Dox medium (4,7) supplemented with 25% glycerol. Inoculum was produced by adding 500 µl of spore stock solutions to 250 ml of potato dextrose broth. The cultures were subsequently incubated for 7 days at 23°C on a rotary shaker (100 RPM). The resulting suspension was filtered through sterile gauze. Spore concentration was determined with a hemacytometer and diluted to give 1 × 10^6 spores ml^{-1}.

The winter oilseed rape cultivars, ‘Express’ (moderately tolerant) and ‘Falcon’ (highly susceptible), were used as controls in all experiments. Seeds of the control and test plants were surface-sterilized by immersion in 70% ethanol for 15 min. Subsequently, the seeds were washed in sterilized tap water before being sown in double-autoclaved silica sand. Seedlings appeared after 3 to 7 days, and 7 days later the roots were carefully washed out of the sand. Inoculations were performed by cutting 2 cm off the root apex and submerging the capped roots for 30 min in a mixed conidial suspension (10^6 ml^{-1}) of both *V. longisporum* isolates VL 43 and VL 40. Roots of control plants were also cut and submerged in tap water for the same length of time. A total of 24 inoculated seedlings and 24 control plants of each accession were transferred into a mixture of sand, peat, and compost (1:1:2) in pots containing two plants each. Plants were grown at 23°C in a greenhouse with a lighting regime of 14/10 h (light/dark). Plants were not fertilized during the investigations. Each plant was scored weekly for disease symptoms over a 4-week time period using an assessment key (Table 3) slightly modified from Zeise (24).

TABLE 3. Assessment key for scoring disease severity of *Verticillium longisporum* in *Brassica* species using the root dip inoculation method

Score	Symptom development
1	No symptoms
2	Symptoms (yellowing, black veins) on the youngest leaves
3	Symptoms on the next older leaves
4	≤50% of leaves show symptoms
5	>50% of leaves show symptoms
6	≤50% of leaves are dead
7	Only shoot meristem still alive
8	Plant is dead

Statistical analysis of data. For each accession tested, the area under the disease progress curve (AUDPC) was calculated from the disease severity values according to the following formula from Campbell and Madden (1): AUDPC = $\Sigma(y_i + y_{i+1}/2) \times (t_{i+1} - t_i)$, where y_i is the disease severity value for observation number i, t_i is the number of days after inoculation at the time point of observation number i, and n is the number of observations. To compensate for a fluctuating infection level between the trials a correction of the AUDPC value for each accession was made based on the reaction of the internal controls. The corrected value ($AUDPC_{corr}$) was calculated as follows:

$$AUDPC_{corr} = \frac{AUDPC}{(AUDPC_{Express} + AUDPC_{Falcon})/2}$$

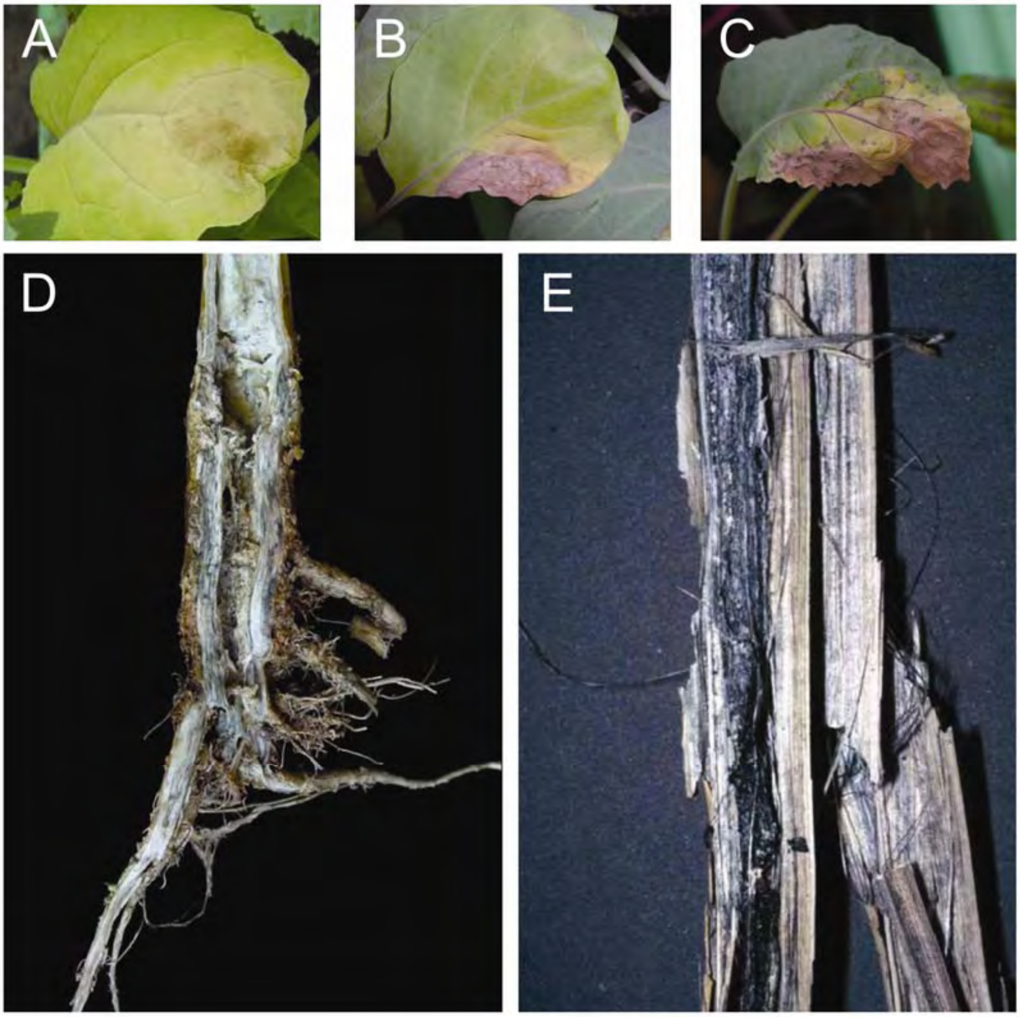


Fig. 1. Typical *Verticillium longisporum* disease symptoms on **A, B, and C,** young leaves, **D,** roots, and **E,** stem of *Brassica napus* plants.

Phenotype data of RS lines was statistically analyzed with the software package SAS 9.1 (SAS Institute, Cary, NC). A general linear model procedure was used to identify RS lines with a significantly lower AUDPC$_{corr}$ value than cultivar Express, which in comparison to other currently available European winter oilseed rape cultivars shows a relatively high tolerance to *V. longisporum* (11).

RESULTS

Resistance screening. A broad range of resistance was observed among the tested materials. Examples of disease symptoms in susceptible oilseed rape plants are shown in Figure 1. Resistance responses of the tested *B. rapa* and *B. oleracea* genotypes, along with the oilseed rape controls, 'Express' and 'Falcon', measured by AUDPC$_{corr}$ value, are shown in Figure 2. The tested plants of the *B. oleracea* accessions showed predominantly good resistance. With the exception of five accessions, the resistance phenotype was verified in all of the *B. oleracea* lines that were reported as *V. longisporum*-resistant by Happstadius et al. (11). In addition, two new resistant *B. oleracea* genotypes were identified, namely accession 1428 (white cabbage 'Kashirka 202') and accession 2070 (white cabbage 'Ladozhskaya'). The tested plants of both of these accessions showed consistently good resistance, with mean AUDPC$_{corr}$ values comparable to the most resistant cabbage accessions identified by Happstadius et al. (11).

In comparison to the relatively high level of resistance among the *B. oleracea* accessions, the *B. rapa* accessions exhibited a predominantly susceptible or only moderately resistant phenotype. The highest level of susceptibility was observed in the Indian Yellow Sarson cv. YS Pb-24. On the other hand, the accessions 13444 and G454 'Granat' showed moderate resistance phenotypes. However, of these two potentially promising *B. rapa* resistance donors, only the accession 13444 gave rise to RS plants that produced sufficient seed for resistance screening. A total of 18 RS lines with sufficient seed were produced from *B. rapa* accession 13444, crossed with various resistant *B. oleracea* accessions. The remaining 27 RS lines that produced sufficient seed for pathogenicity testing all derived from the moderately susceptible *B. rapa* accession 56515, suggesting that the other *B. rapa* accessions that were used may have a negative effect on fertility or fecundity of the resulting RS genotypes.

The 45 RS lines that were tested in the pathogenicity screening showed a large variation in their response to *V. longisporum* infection (Fig. 3). This reflects the heterozygous nature of the *B. rapa* and *B. oleracea* gene bank accessions, which were in most cases self-incompatible. Statistical data analysis of AUDPC$_{corr}$ values of all RS lines in relation to the control cultivar Express, confirmed that 37 RS lines (82%) showed significantly lower symptom scores ($P \leq 0.0194$) than the moderately tolerant control. Unexpectedly, the RS lines derived from the moderately susceptible *B. rapa* accession 56515 showed significantly higher resistance ($P = 0.001$) than RS lines derived from crosses with the more resistant *B. rapa* accession 13444, irrespective of the *B. oleracea* resistance donor. This implies that the resistance factors from the respective *B. oleracea* donors are more effective in combination with resistance from the *B. rapa* accession 56515. The seven most resistant RS lines were derived from crosses between plants of *B. rapa* accession 56515 and *B. oleracea* accession 8207. Interestingly, none of the RS lines derived from

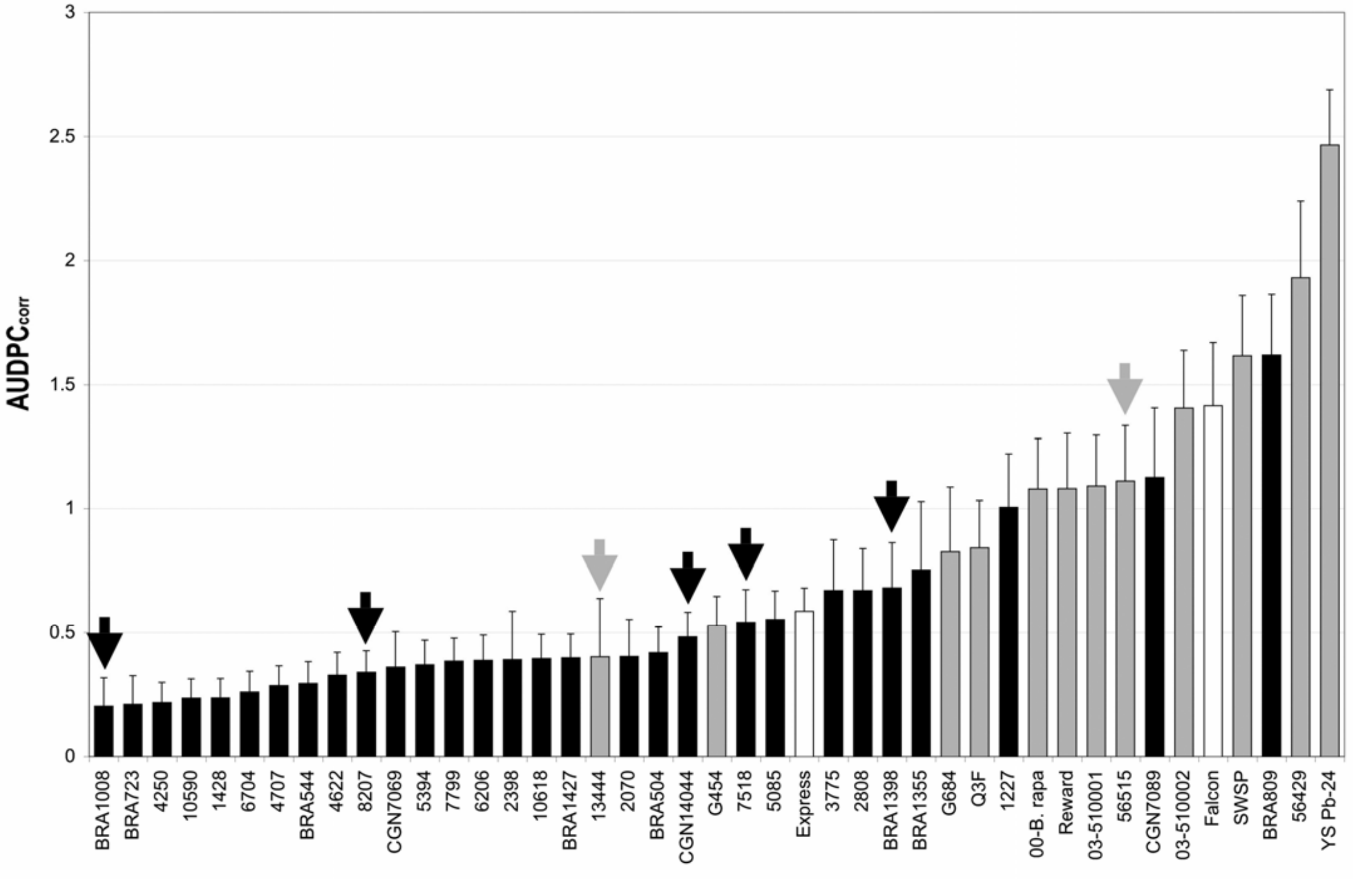


Fig. 2. *Verticillium longisporum* resistance responses measured by area under the disease progress curve (AUDPC$_{corr}$) for *Brassica rapa* (gray) and *B. oleracea* (black) accessions compared to the reference oilseed rape cultivars Express and Falcon (white). Columns and bars represent mean values and standard errors, respectively, from 20 infected plants of each accession. Arrows show accessions from which resistant resynthesized *B. napus* lines were successfully regenerated via embryo rescue-assisted interspecific hybridization.

the cross *B. rapa* 13444 × *B. oleracea* 8207 showed a high level of resistance, suggesting a possible antagonism between the respective resistance mechanisms of these two donors. On the other hand, the high resistance of the RS lines from *B. rapa* 56515 × *B. oleracea* 8207 in comparison to the lines from *B. rapa* 13444 × *B. oleracea* 8207 appears to confirm that we have succeeded in effectively combining two different *B. oleracea* and *B. rapa* resistance sources in RS oilseed rape genotypes with a high level of quantitative resistance.

DISCUSSION

Broadening the narrow genetic basis of *B. napus* through interspecific or intergeneric hybridization is an important strategy for resistance breeding. The *Brassica B*-genome species *B. juncea*, *B. carinata*, and *B. nigra* have been used in many studies to transfer resistance against *Leptosphaeria maculans* to oilseed rape (2,6,17,18,21), however the development of new oilseed rape cultivars with B-genome resistance has proved difficult in many cases. On the other hand the use of *B. rapa* and *B. oleracea* genotypes as resistance donors can circumvent this problem because these are the ancestral parents of *B. napus* and their chromosomes have remained more or less intact in the amphidiploid genome (16). Crouch et al. (3) and Diederichsen and Sacristan (5) successfully transferred resistance against *L. maculans* and *Plasmodiophora brassicae*, respectively, from *B. rapa* to *B. napus*. These two studies led to the development of new winter oilseed rape cultivars carrying resistance to blackleg and clubroot disease, respectively. However, in both cases the resistance was overcome in a relatively short time by virulent pathogen races. On the other hand, the identification of genetically diverse resistance sources in different *B. rapa* and *B. oleracea* accessions opens the possibility for direct combination of different resistance genes. Combination or pyramiding of resistance sources can potentially provide a more durable and sustainable resistance.

In the present study two previously unknown *B. oleracea* sources of *Verticillium* resistance (accessions 1428 and 2070) were identified and the resistance phenotypes of other cabbage accessions were verified. Moreover, a resistance screening of *B. rapa* accessions led to the discovery of two lines (accessions 13444 and G545) with moderate resistance to *V. longisporum*. The combination of different *B. oleracea* and *B. rapa* resistance in RS *B. napus* resulted in a high level of resistance that in some cases surpassed the expectation based on the resistance levels observed in the respective parental lines. The best-performing interspecific hybrids showed considerably higher resistance than the variety Express, one of the most *V. longisporum*-tolerant cultivars available in Germany. Particularly high resistance was achieved in RS lines derived from combination of *B. rapa* accession 56515 and *B. oleracea* 8207, although other *B. rapa* parents showed higher levels of resistance

Standard errors of the means were clearly higher in the resistance experiment of the *B. rapa* and *B. oleracea* lines, which can be explained by the heterogeneous condition of the material. Diploid *Brassica* species possess a sporophytic homomorphic self-incompatibility system that prevents self-fertilization and promotes outbreeding. For this reason it can be assumed that many of the donor plants that were used for the interspecific hybridizations may have been heterozygous at the relevant resistance gene loci. This suggestion is supported by the large variation of resistance response among the RS lines derived from the same resistant accessions. By crossing these heterozygous plants, many different combinations of resistance sources are possible in the progenies. The most resistant RS phenotypes can be expected to represent an accumulation of the most effective genes in a single homozygous *B. napus* genotype. The sporophytic self-incompatibility common in the diploid *Brassica* species is generally suppressed in the amphidiploid *B. napus*. However, we observed a great variation among the RS plants in their self-fertility and fecundity. The resistant lines described here, how-

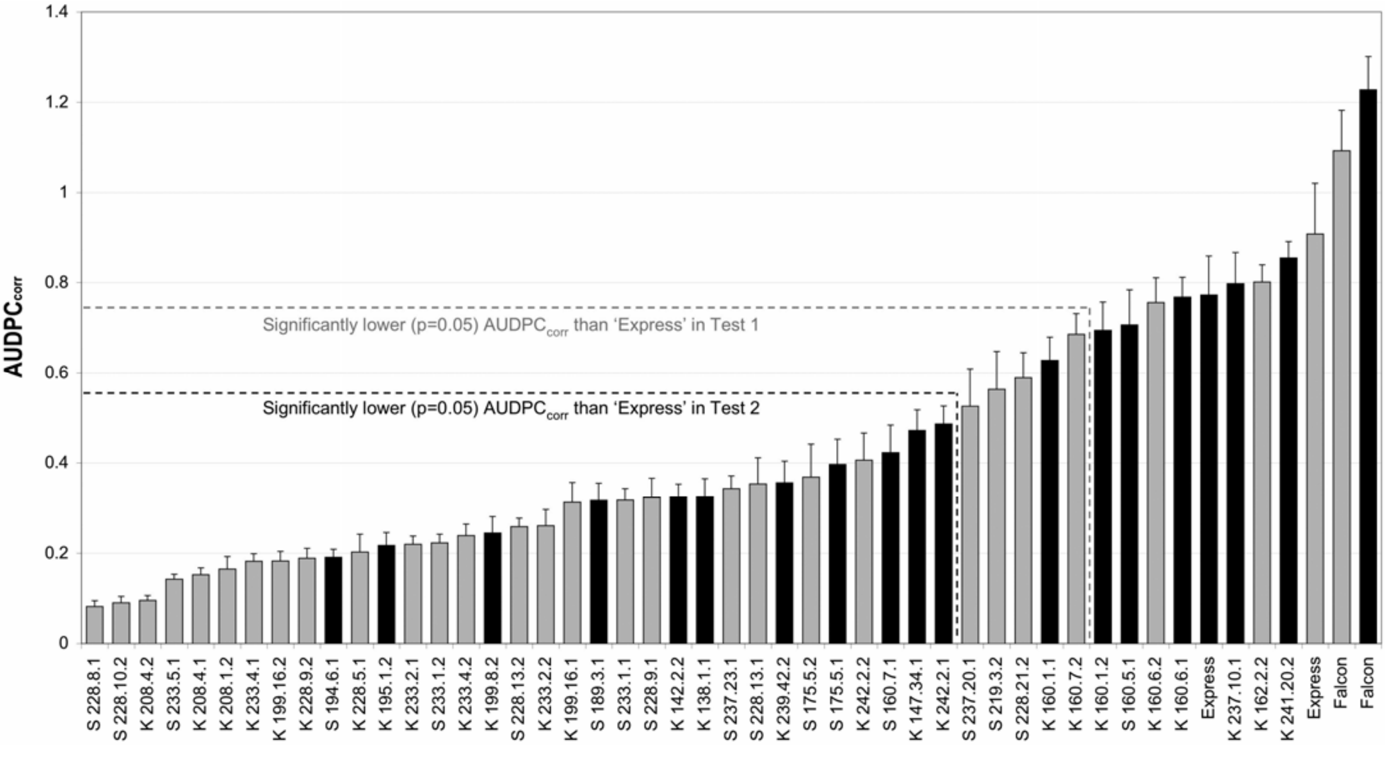


Fig. 3. *Verticillium longisporum* resistance responses of 45 resynthesized oilseed rape lines, along with the reference oilseed rape cultivars Express and Falcon, measured by normalized areas under the disease progress curve (AUDPC$_{corr}$). The light and dark bars represent results from two independent resistance experiments (test 1 and 2) along with the respective controls from each of the two tests. Columns and bars represent mean values and standard errors, respectively, from 20 infected plants of each accession. The dashed horizontal lines show the significance levels (P = 0.05) in each of the two tests for significantly lower AUDPC$_{corr}$ than Express in the respective resistance tests.

ever, showed a good self-fertility and high seed-set, hence they are useful candidates for backcross breeding towards introgression of the resistances into elite oilseed rape lines.

Simultaneous introgression of numerous different resistance loci can be difficult, particularly in the case of quantitative resistance. To ensure that resistant genotypes have as many contributing resistance loci as possible, the development of molecular markers for pyramiding of different A- and C-genome resistances (with potentially complementary resistance mechanisms) is desirable. For this purpose the different RS *B. napus* lines identified in the current study are presently being crossed to susceptible lines for the generation of segregating mapping populations. By testing such material both in greenhouse tests and under field conditions this should allow us to (i) study and compare the genetic basis of the different resistances, (ii) map quantitative trait loci for identification of selection markers closely linked to major resistance loci, and (iii) to select promising resistant oilseed rape lines for the development of new cultivars with resistance to *V. longisporum*.

ACKNOWLEDGMENTS

This work was funded by the Association for Promotion of Private German Plant Breeding (GFP) and the Agency for Renewable Resources (FNR) of the Federal Ministry for Nutrition, Agriculture, and Consumer Protection (BMELV). We thank R. Kürschner, S. Sahner, and J. Schaper for excellent technical assistance.

LITERATURE CITED

1. Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley, New York.
2. Chèvre, A. M., Eber, F., This, P., Barret, P., Tanguy, X., Brun, H., Delseny, M., and Renard, M. 1996. Characterization of *Brassica nigra* chromosomes and of blackleg resistance in *B. napus-B. nigra* addition lines. Plant Breed. 115:113-118.
3. Crouch, J. H., Lewis, B. G., and Mithen, R. F. 1994. The effect of A genome substitution on the resistance of *Brassica napus* to infection by *Leptosphaeria maculans*. Plant Breed. 112:265-278.
4. Czapek, F. 1902. Untersuchungen über die Stickstoffgwinnung und Eiweißbildung der Pflanzen. Breitr. Chem. Physiol. Pathol. 2:540-560.
5. Diederichsen, E., and Sacristan, M. D. 1996. Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to *Plasmodiophora brassicae.* Wor. Plant Breed. 115:5-10.
6. Dixelius, C. 1999. Inheritance of the resistance to *Leptosphaeria maculans* of *Brassica nigra* and *B. juncea* in near-isogenic lines of *B. napus*. Plant Breed. 118:151-156.
7. Dox, A. W. 1910. The intracellular enzymes of Penicillium and Aspergillus with special references to those of *P. camenberti*. U.S. Dep. Agric. Bur. Anim. Ind. Bull. 120:1-170.
8. Dreyer, F., Graichen, K., and Jung, C. 2001. A major quantitative trait locus for resistance to Turnip yellows virus (TuYV, syn. beet western yellows virus, BWYV) in rapeseed. Plant Breed. 120:457-462.
9. Dunker, S., Keunecke, H., and von Tiedemann, A. 2006. *Verticillium longisporum* in winter oilseed rape—Impact on plant development and yield. Integrated Control Oilseed Crops 29:365-374.
10. Fähling, M., Graf, H., and Siemens, J. 2003. Pathotype separation of *Plasmodiophora brassicae* by the host plant. J. Phytopathol. 151:425-430.
11. Happstadius, I., Ljungberg, A., Kristiansson, B., and Dixelius, C. 2003. Identification of *Brassica oleracea* germplasm with improved resistance to Verticillium wilt. Plant Breed. 122:30-34.
12. Heale, J. B., and Karapapa, V. K. 1999. The Verticillium threat to Canada's major oilseed crop: Canola. Can. J. Plant Pathol. 21:1-7.
13. Karapapa, V. K., Bainbridge, B. W., and Heale, J. B. 1997. Morphological and molecular characterization of *Verticillium longisporum* comb. Nov, pathogenic to oilseed rape. Mycol. Res. 101:1281-1294.
14. Li, L., Paolillo, D. J., Parthasarathy, M. V., DiMuzio, E. M., and Garvin, D. F. 2001. A novel gene mutation that confers abnormal patterns of β-carotene accumulation in cauliflower (*Brassica oleracea* var. *botrytis*). Plant J. 26:59-67.
15. Murashige, T., and Skoog, F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant 15:473-497.
16. Parkin, I. A. P., Sharpe, A. G., Keith, D. J., and Lydiate, D. J. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome 38:1122-1131.
17. Roy, N. N. 1978. A study on disease variation in the populations of an interspecific cross of *Brassica juncea* L. × *B. napus* L. Euphytica 27:145-149.
18. Sacristán, M. D., and Gerdemann, M. 1986. Different behavior of *Brassica juncea* and *B. carinata* as sources of *Phoma lingam* resistance in experiments of interspecific transfer to *B. napus*. Z. Pflanzenzuecht. 97:304-314.
19. Schnathorst, W. 1981. Life cycle and epidemiology of *Verticillium*. Pages 81-111 in: Fungal Wilt Diseases of Plants. M. E. Mace, A. A. Bell, and C. H. Beckmann, eds. Academic Press, London, UK.
20. Sprague, S. J., Balesdent, M. H., Brun, H., Hayden, H. L., Marcroft, S. J., Pinochet, X., Rouxel, T., and Howlett, B. J. 2006. Major gene resistance in *Brassica napus* (oilseed rape) is overcome by changes in virulence of populations of *Leptosphaeria maculans* in France and Australia. Eur. J. Plant Pathol. 114:33-40.
21. Struss, D., Quiros, C. F., Plieske, J., and Röbbelen, G. 1996. Construction of *Brassica* B genome synteny groups based on chromosomes extracted from three different sources by phenotypic, isozyme and molecular markers. Theor. Appl. Genet. 93:1026-1032.
22. Walsh, J. A., Sharpe, A. G., Jenner, C. E., and Lydiate, D. J. 1999. Characterization of resistance to turnip mosaic virus in oilseed rape (*Brassica napus*) and genetic mapping of TuRB01. Theor. Appl. Genet. 99:1149-1154.
23. Zeise, K., and von Tiedemann, A. 2001. Morphological and Physiological Differentiation among vegetative compatibility groups of *Verticillium dahliae* in relation to *V. longisporum*. J. Phytopathol. 149:469-475.
24. Zeise, K. 1992. Gewächshaustest zur Resistenzprüfung von Winterraps (*Brassica napus* L. var. *oleifera* Metzger) gegen den Erreger der Rapswelke *Verticillium dahliae* Kleb. Nachrichtenbl Deutscher Pflanzenschutzd 44:125-128.

4.2 Development and characterisation of *Brassica napus-Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*

Snowdon R.J., H. Winter, A. Diestel & M.D. Sacristan, 2000a: Theor. Appl. Genet. 101: 1008-1014

Theor Appl Genet (2000) 101:1008–1014

R.J. Snowdon · H. Winter · A. Diestel
M.D. Sacristán

Development and characterisation of *Brassica napus-Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*

Received: 20 September 1999 / Accepted: 25 March 2000

Abstract Blackleg caused by *Leptosphaeria maculans* is one of the most important diseases affecting oilseed rape worldwide. *Sinapis arvensis* is valuable for the transfer of blackleg resistance to oilseed rape (*Brassica napus*) because this species contains high resistance against various aggressive isolates of the blackleg fungus. These include at least one Australian isolate which has been found to overcome resistance originating from species with the *Brassica* B genome, until now the major source for interspecific transfer of blackleg resistance. Backcross offspring from intergeneric crosses between *Brassica napus* and *S. arvensis* were subjected to phytopathological studies and molecular cytogenetic analysis with genomic *in situ* hybridisation (GISH). The BC_3S progenies included fertile plants exhibiting high seedling (cotyledon) and adult plant resistance associated with the presence of an acrocentric addition chromosome from *S. arvensis*. In addition, some individuals with adult plant resistance but cotyledon susceptibility were observed to have a normal *B. napus* karyotype with no visible GISH signals, indicating possible resistant introgression lines. Phytopathological analysis of selfing progenies from 3 different highly resistant BC_3 plants showed that seedling and adult plant resistance are probably conferred by different loci.

Key words *Brassica napus* · *Sinapis arvensis* · *Leptosphaeria maculans* · Intergeneric hybrids · Blackleg resistance · Genomic *in situ* hybridisation

Communicated by H.C. Becker

R.J. Snowdon and H. Winter contributed equally to this article

R.J. Snowdon
Institut für Pflanzenbau und Pflanzenzüchtung II,
Justus-Liebig-Universität Giessen, IFZ, Heinrich-Buff-Ring 26,
35392 Giessen, Germany

H. Winter · A. Diestel · M.D. Sacristán (✉)
Institut für Biologie – Angewandte Genetik, Freie Universität Berlin,
Albrecht-Thaer-Weg 6, 14195 Berlin, Germany
Fax: +49 (030) 838 54345
e-mail: sacris@zedat.fu-berlin.de

Introduction

Leptosphaeria maculans (Desm.) Ces. et De Not. [anamorph *Phoma lingam* (Tode ex Fr.) Desm.], the causal agent of blackleg, induces severe damage worldwide to basal parts (stem canker) of oilseed rape (*Brassica napus* L., genome AACC, 2n=38) and other susceptible plants of the family Brassicaceae. Moreover, it can also provoke lesions and necroses on leaves, pods and seeds. The genetic basis of blackleg resistance in *B. napus* in Europe is narrow and originates for the most part from the French cultivar Jet Neuf, which possesses a partial, polygenically controlled adult plant resistance not expressed at the seedling stage (Cargeeg and Thurling 1980). In contrast, all *Brassica* species containing the B genome exhibit an absolute and stable resistance to most of the aggressive pathogen isolates studied to date. B-genome resistance is mono- or oligogenically controlled (see Rimmer and van den Berg 1992; Dixelius 1999) and efficient from the seedling stage onwards. Thus, B-genome donors like *B. nigra* (L.) Koch (BB, 2n=16) and *B. juncea* (L.) Czern (BBCC, 2n=36) are often used as a genetic pool for the development of resistant oilseed rape (Roy 1978; Sacristán and Gerdemann 1986; Sjödin and Glimelius 1989; Chèvre et al. 1996a; Struss et al. 1996; Plieske et al. 1998; Dixelius 1999).

Some aggressive isolates of the pathogen have recently been shown to overcome the resistance of *B. juncea* (Purwantara et al. 1998; Winter et al. 1999). *L. maculans* exhibits a broad variation in virulence, giving it the potential to adapt quickly to a given resistance (Kuswinanti et al. 1999). The generation of a durable resistance therefore necessitates the application of a broad spectrum of resistance sources in oilseed rape breeding. For this reason, interspecific and intergeneric transfer of blackleg resistance from wild crucifers like *Sinapis arvensis* L. (wild mustard, genome SarSar, 2n=18) is becoming increasingly important. *S. arvensis* possesses resistance in all developmental stages to various *L. maculans* isolates (Plümper 1995), including one from Australia that overcomes the resistance of *B. juncea* (see Winter et al. 1999).

Table 1 Scales for evaluation of blackleg symptoms on cotyledons and adult plants

Score	Symptoms	Class[a]
Cotyledon		
1	Punctiform necrosis at inoculation site	R_C
2	Minor chlorotic lesion (< 2 mm diameter)	R_C
3	Moderate chlorotic lesion (2–5 mm)	S_C[b]
4	Major chlorotic lesion (> 5 mm), usually with pycnidia formation at later stages	S_C
Adult plant		
1	Limited necrosis around inoculation site	R_A
2	Minor damage to the basal parts of the stem (≤ 25% girdling)	R_A
3	Moderate damage on the basal parts of the stem (>25% and ≤50% girdling)	S_A
4	Major damage on the basal parts of the stem (>50% and <100% girdling)	S_A
5	Severe damage on the basal parts of the stem (100% girdling)	S_A
6	Plant collapse	S_A

[a] Classes: R = resistant, S = susceptible, C = cotyledon, A = adult plant.
[b] Classified as R_C when lesion develops only at late stage

Sexual hybridisation between *S. arvensis* and *B. napus* has been reported by various authors (Kerlan et al. 1993; Plümper 1995; Bing et al. 1996; Chèvre et al. 1996b; Lefol et al. 1996). Although such hybrids between *Brassica* species and their close relatives are generally easy to produce, until now they have been used in only a few cases for the breeding of new rapeseed varieties with traits of agronomical importance (see Friedt and Lühs 1998). The generation of new lines containing the character of interest in the desired genetic background is accelerated considerably when the transfer of donor chromatin can be effectively monitored (Jiang and Gill 1996). Classical cytogenetic investigations are difficult in *Brassica*, however, due to a lack of cytological markers for a reliable identification of chromosomes. Molecular cytogenetic techniques like genomic *in situ* hybridisation (GISH), on the other hand, provide an effective alternative for identifying chromosome additions and introgressions in hybrid offspring without the need for time-consuming and expensive molecular genetic investigations. The suitablility of GISH for the characterisation of intergeneric *Brassica* hybrids has been demonstrated (Fahleson et al. 1997; Sharzhinskaya et al. 1998; Snowdon et al. 1998). Despite this, genetic investigations of hybrids carrying introduced blackleg resistance have until now been mainly limited to molecular marker or classical cytogenetic analyses in interspecific crosses between *Brassica* species (Chèvre et al. 1996a, 1997; Plieske et al. 1998; Dixelius 1999).

In the study reported here, selfing progenies (BC_3S) of resistant BC_3 plants from hybrids between *B. napus* and *S. arvensis* were characterised with respect to their resistance behaviour, at different developmental stages, to the aggressive *L. maculans* isolate W4. Molecular cytogenetic analysis using GISH was applied to investigate the genomic composition of resistant and susceptible individuals. Preliminary results have been reported by Snowdon et al. (1999) and Winter et al. (1999).

Materials and methods

Plant material

Offspring from intergeneric crosses between *B. napus* cv. Madora (winter oilseed rape) and *S. arvensis* (origin: Biologische Bundesanstalt Braunschweig, gene bank accession No. 22529) were backcrossed three times with winter oilseed rape cv. Ceres. Following blackleg resistance tests (see below) of backcross offspring, BC_3 plants exhibiting both cotyledon and adult plant resistance were selfed to generate BC_3S progenies. Plants were grown in a greenhouse at a minimum temperature of 18°C, with additional illumination during the months October-April. The material was vernalised for 6–8 weeks at 4°C.

Crosses, ovule and ovary cultures (embryo rescue), along with classical mitotic chromosome analyses of the original hybrids and plants of the backcross generations, have been described by Plümper (1995), who used methods modified from Sacristán and Gerdemann (1986). In some cases, 3- to 6-mm-long styles from young flower buds were used instead of root tips for chromosome counts, as suggested by Wu et al. (1997).

In addition to the resistance tests on different BC_3 and BC_3S progenies, susceptible control genotypes (*B. napus* cvs Madora, Ceres and Lesira) and plants of the resistance donor *S. arvensis* were also included as positive and negative infection controls, respectively. Lesira was integrated in the tests because it is one of the most susceptible oilseed rape cultivars.

Resistance tests

Plants were inoculated with a pycnidiospore suspension of the aggressive German *L. maculans* isolate W4 (Tox^+, Hassan et al. 1991) obtained from cultures grown on V8-agar medium (Sacristán 1982). Cotyledon and adult plant resistance tests with double inoculation were carried out in the greenhouse as follows. Cotyledons of 5- to 9-day-old plants were punctured centrally and inoculated with 5 µl of a standard spore suspension (10^7 spores/ml). Plants were incubated for 3 days in a transparent tent under high-humidity conditions. Maximum lesion sizes were scored at 2-day intervals between the 14th and 24th day after inoculation using a scale (see Table 1) modified from Sacristán (1982). Plants with scores 1, 2 and only late occurrence of 3 were classified as cotyledon resistant (R_C); the remaining cases were classified as cotyledon-susceptible (S_C).

At the age of 4–5 weeks the same plants were inoculated a second time. A cellulose pad, soaked with 100 µl of the standard spore suspension, was applied with a strip of parafilm to a wound at least 1 cm long and 1–2 mm deep that had been inflicted with a lancet across the border between the hypocotyl and epicotyl. The pads were removed after 10 days. External blackleg symptoms on adult plants (see Table 1) were evaluated 6–7 weeks after stem in-

Table 2 Chromosome counts and blackleg resistance response frequencies in hybrid plants and negative (-) and positive (+) infection controls. *Sinapis arvensis* was the resistance donor, while Ceres, Madora and Lesira are blackleg susceptible *B. napus* cultivars

Genotype	Chromosome number (2n)		Plants tested	Resistance[a] response frequencies							
				R_C	S_C	R_A	S_A	$R_{C/A}$	S_C/R_A	R_C/S_A	$S_{C/A}$
BC_3 9.10	38, 39	BC_3S offspring	19	6	13	11	8	4	7	2	6
BC_3 13.6	37, 39, 41		19	4	15	12	7	4	8	0	7
BC_3 16.3	40–43[b]		25	6	19	17	8	5	12	1	7
BC_3 31.2	36, 38		19	0	19	0	19	0	0	0	19
S. arvensis (-)	18	Control genotypes	25	25	0	25	0	25	0	0	0
Ceres (+)	38		21	0	21	3	18	0	3	0	18
Madora (+)	38		17	0	17	2	15	0	2	0	15
Lesira (+)	38		19	0	19	0	19	0	0	0	19

[a] Resistance classes are described in Table 1; C/A = combined cotyledon and adult plant resistance or susceptibility, respectively

[b] Exact determination of BC_3 chromosome number was not possible

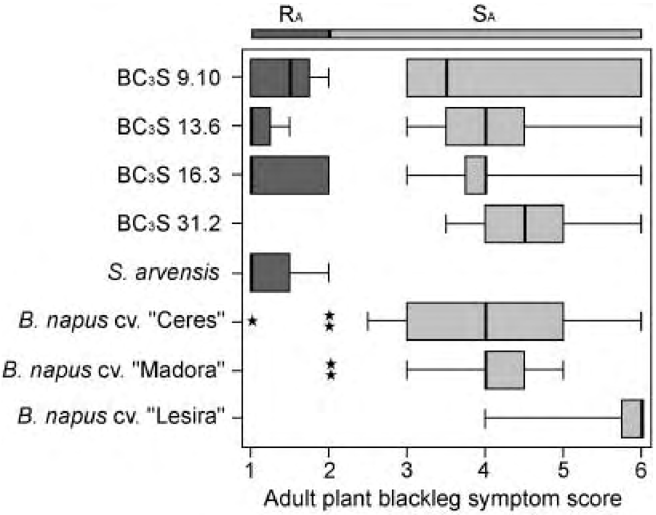


Fig. 1 Boxplot (SPSS) showing adult plant blackleg symptom scores of resistant (R_A) and susceptible (S_A) plants in 4 BC_3S progenies from the interspecific cross *B. napus* × *S. arvensis*, along with control genotypes. *Boxes* cover the interquartile range, with median scores shown by the *thick vertical lines*. *Whiskers* cover the remaining variation, with the exception of 5 resistant *B. napus* individuals that are represented by *stars*

oculation. Selected plants with no or only small external lesions were chosen as candidates for resistance. These plants were vernalised and evaluated once more during seed ripening to give a post-vernalisation score. Plants with a maximum external score of 2 or less were classed as adult plant resistant (R_A); others were grouped as adult plant susceptible (S_A). Distributions of adult plant blackleg symptom scores in different BC_3S progenies and control genotypes were compared with a boxplot generated by the software SPSS.

For the majority of the plants, internal lesions were scored simultaneously, by cross-sectioning of the basal stem parts, using a scale based on that of Hammond and Lewis (1987). Because the internal lesion scores usually reflected the external symptoms, this data is omitted here.

Genomic *in situ* hybridisation

Genome composition in resistant and susceptible backcross offspring was analysed by GISH following methods described previously (Snowdon et al. 1997). Chromosome spreads were produced from protoplast suspensions from root tips of young plants or cuttings, then fixed with formaldehyde and denatured using standard techniques. For the GISH probe, genomic DNA from *S. arvensis* was directly labelled with the fluorochrome Cy3 by nick translation and resuspended in hybridisation solution with a 50-fold excess of unlabelled, sheared *B. napus* competitor DNA. The probe was pre-annealed for 20 min at 37°C prior to hybridisation. Following overnight *in situ* hybridisation (see Schwarzacher et al. 1994) and a stringent wash for 10 min in 0.2 ´ SSC at 42°C, chromosomes were counter-stained in DAPI and examined using a Leica DM 3 fluorescence microscope. At least ten metaphases were examined for each hybridisation.

Results

Out of 222 tested BC_3 plants, with chromosome numbers from 2n=38–43, 15 exhibited both cotyledon and adult plant resistance; 98 were classed as susceptible on both levels. Results of phytopathological tests on BC_3S progenies derived from 3 double resistant BC_3 plants (9.10, 13.6 and 16.3) and 1 double susceptible BC_3 individual (31.2), respectively, along with positive and negative control genotypes, are outlined in Table 2 and Figs. 1 and 2. Tested offspring from the BC_3S individual 31.2 were all double susceptible, and plant collapse occurred in some instances, although the average adult plant susceptibility was not as high as for the very susceptible control *B. napus* cv. Lesira (Figs. 1, 2a, 2e). BC_3S progenies derived from the BC_3 plants 9.10, 13.6 and 16.3 exhibited cotyledon and/or adult plant resistance (Figs. 2a, 2d, 2e). The majority of BC_3S offspring from these 3 individuals were adult plant resistant but cotyledon susceptible. Only a small number of plants showed the reciprocal resistance behaviour.

Two out of seventeen tested plants from the original cross *B. napus* parent Madora and 3 out of 21 plants from the backcross parent *B. napus* cv. Ceres were scored as adult plant resistant (see Discussion), while most of the others were only moderately adult plant susceptible, similar to the majority of adult plant susceptible BC_3S progeny (Figs. 1, 2c, 2e left). All *S. arvensis* individuals showed a clear cotyledon (Fig. 2a) and adult plant resistance phenotype (Figs. 1, 2b).

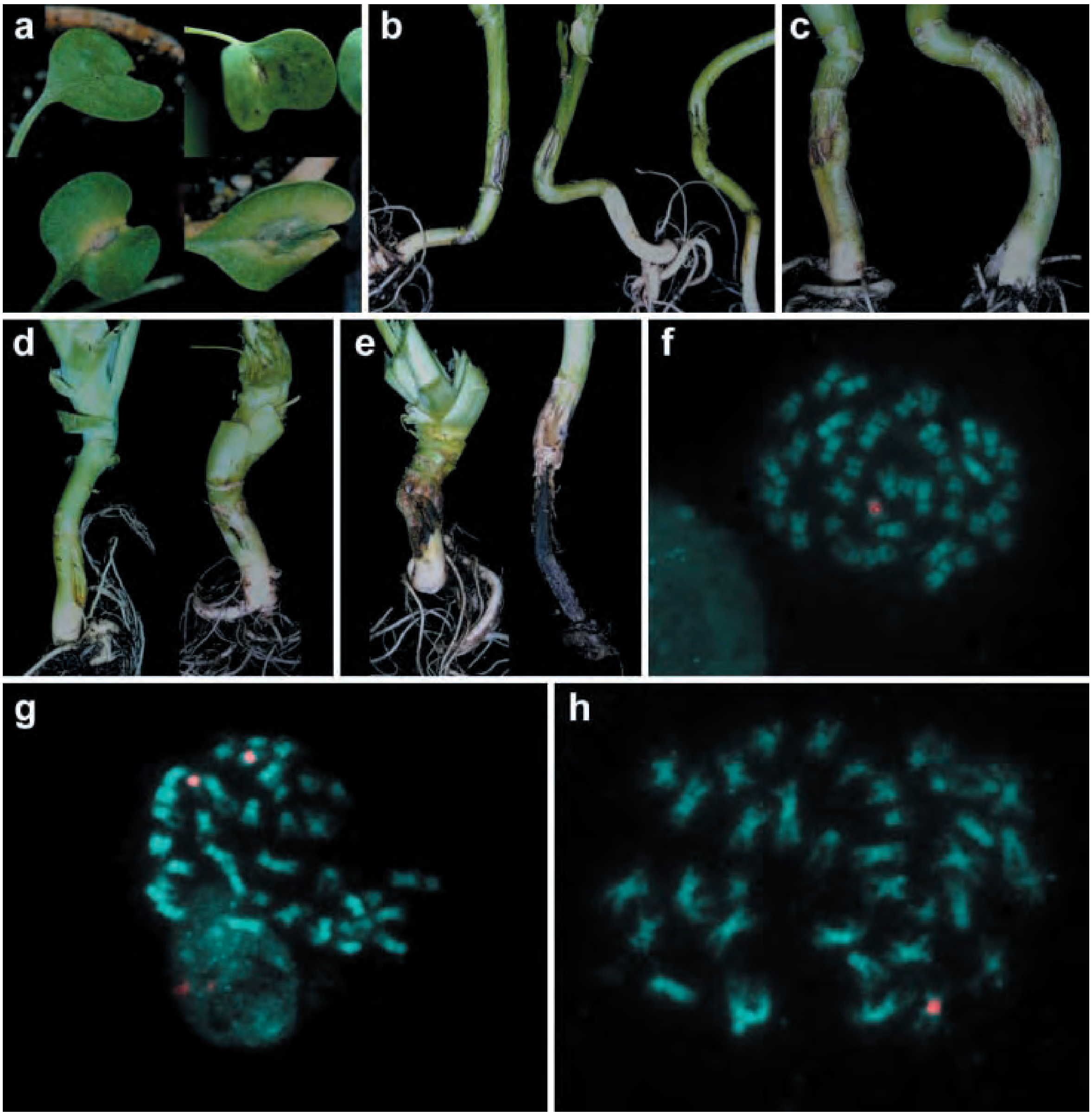


Fig. 2a–h Phytopathological and molecular cytogenetic characterisation of hybrid plant material. **a** Cotyledon resistance (*above*) and susceptibility (*below*) phenotypes in BC_3S progeny. The *above-left* phenotype is typical for *S. arvensis*, while the *below-right* symptoms are typical for the *B. napus* cultivars Ceres, Madora and Lesira. **b** Adult plant resistance of *S. arvensis*. **c** Moderate adult plant susceptibility of *B. napus* cv. Ceres. **d** Adult plant resistance phenotypes in BC_3S progeny. **e** Adult plant susceptibility phenotypes in BC_3S progeny. The phenotype on the *left* is representative for the majority of adult plant susceptible BC_3S offspring, while the stem on the *right* shows plant collapse, typical for *B. napus* cv. Lesira. **f-h** Genomic *in situ* hybridisation results in three BC_3S offspring plants. *S. arvensis* chromosomes are labelled *red* with Cy3, while *B. napus* chromosomes show no hybridisation signals and are stained *blue* with DAPI. **f, g** Individuals showing both cotyledon and adult plant resistance and corresponding respectively to the *left* and *right* plants in **d**, with **f** a monosomic, acrocentric addition, and **g** one acrocentric and a second, metacentric *S. arvensis* addition chromosome. **h** An adult plant sensitive individual with a monosomic, metacentric addition chromosome, corresponding to the *right-hand* plant in **e**

Results from GISH with selected BC_3S offspring are given in Table 3, while Fig. 2 (f–h) shows examples of GISH images from resistant and susceptible BC_3S plants with monosomic or double addition chromosomes. Two individuals showed adult plant resistance but cotyledon susceptibility and a stable chromosome number of 2n=38 with no visible GISH signals, indicating the possible presence of small introgressions carrying the resistance gene(s). A further plant with the same resistance behaviour and no GISH signals had variable chromosome numbers (2n=36–38). The BC_3 parents studied also showed somatic variations in mitotic chromosome number (Table 2), however chromosome counts in the BC_3S offspring were generally more uniform than in the BC_3. All remaining adult plant resistant individuals tested contained an acrocentric addition chromosome, and in 3 plants that were also cotyledon resistant this addition was monosomic (Fig. 2f). One adult plant resistant individual and 2

Table 3 Karyotpyes of BC_3S plants investigated by GISH detection of *S. arvensis* chromatin, grouped according to their respective resistance response (see Table 1)

		Chromosome number (2n)					
GISH karyotype:		36–37 No GISH signal	38 No GISH signal	38 Monosomic metacentric substitution chromosome	39 Monosomic metacentric addition chromosome	39 Monosomic acrocentric addition chromosome	40 Metacentric plus acrocentric addition
Resistance response	R_A	1[a]	3[a]			1	1
	$R_{C/A}$					3	2
	$S_{C/A}$		5	1	1		

[a] Variable chromosome number within one plant

plants also exhibiting cotyledon resistance had both an acrocentric and a metacentric addition (Fig. 2g). Plants that possessed only a metacentric addition (Fig. 2h), or in one case a metacentric substitution chromosome, were fully susceptible. All other susceptible plants tested had 38 chromosomes and no visible GISH signals.

Discussion

This study describes the characterisation of blackleg resistant backcross offspring from a cross between *B. napus* and *S. arvensis*. The use of the GISH method provides what is to the best our knowledge the first clear evidence of monosomic and double addition chromosomes in crosses between species of the genera *Brassica* and *Sinapis*. Moreover, the results presented here indicate the possible occurrence of intergenomic recombination, resulting in plants with a normal *B. napus* karyotype (2n=38; no *S. arvensis* chromatin detected by GISH) exhibiting resistance introgressed from the donor genome.

In intergeneric hybrids between *S. arvensis* and the three diploid *Brassica* species containing the A, B and C genomes, respectively, Mizushima (1950) reported three allosyndetic bivalents in ASar and CSar hybrids and seven in BSar hybrids. This is in agreement with data published by Kerlan et al. (1993) and Chèvre et al. (1996b) who found few paired chromosomes in ACSar hybrids. Investigations using chemotaxonomic markers (Tsukamoto et al. 1993) and on the DNA level (Song et al. 1988; Warwick and Black 1991; Kapila et al. 1996) also showed the relatively close relationship of *S. arvensis* to *Brassica* species containing the B genome.

S. arvensis is particularly interesting for oilseed rape breeding because of its high level of blackleg resistance to various *L. maculans* isolates. Furthermore, its resistance response differs largely from that of *Brassica* species with the B genome with respect to the timing of phytoalexin induction after inoculation with the pathogen (Storck and Sacristán 1995), making it a potentially useful complementation to existing resistance sources. This study confirms the suitability of *S. arvensis* as a donor plant for blackleg resistance transfer into oilseed rape and reinforces the value of GISH for the investigation of intergeneric hybrids between *B. napus* and its close relatives.

Blackleg resistance tests like those used in this study, involving a double inoculation with the same isolate, are the method of choice for accurate comparative resistance evaluation at both the seedling (cotyledon) and adult plant stages. Because double inoculation induces a very severe phytopathological response, it widely prevents escapes and levels differences in the reaction within one genotype. It is possible with this test, however, to overlook specific resistances; for example in the lamina of the cotyledons or leaves, or in the petioles (Hammond and Lewis 1987; Pang and Halloran 1996). Nevertheless, in comparison with tests of adult plants inoculated only on cotyledons or basal parts of the stem, enhanced effects were observed in tests with double inoculation (data not shown), especially in susceptible or only moderately resistant genotypes. This indicates that systemic acquired resistance, as described by Mahuku et al. (1996) for coinfection of weakly and highly virulent *L. maculans* isolates, does not occur in this system.

The oilseed rape cultivars Madora and Ceres, used as the respective cross and backcross parents in the present study, showed a certain degree of adult plant resistance in comparison to the other positive control, *B. napus* cv. Lesira (Fig. 1). The resistance behaviour of these genotypes, which both derive from the cultivar Jet Neuf, demonstrates that adult plant tests with double inoculation are preferable to clearly differentiate a significant resistance response. Moreover, the results presented here confirm previous reports on oilseed rape, which showed that the cotyledon response is not a suitable indicator for adult plant resistance behaviour (Cargeeg and Thurling 1979; Sacristán 1982; Pang and Halloran 1996).

The segregation pattern in selfing progenies derived from 3 different, highly resistant BC_3 individuals suggests that adult plant resistance is inherited more readily than cotyledon resistance. Probably, more gene loci conferring adult plant resistance are present in the *S. arvensis* genome than genes for cotyledon resistance. Alternatively, it is also possible that the genes for adult plant and cotyledon resistance in the hybrid material might be carried on addition chromosomes with variable transmission rates. It is, however, impossible to draw firm conclusions regarding the number of resistance genes or their inheritance because aneuploids do not show Mendelian segregation.

The addition chromosomes of the different BC_3S plants could not be reliably compared due to the small size of the chromosomes and the lack of cytogenetic markers in these species. Nevertheless, chromosome morphology provides strong evidence that some or all of the genes for adult plant resistance are found on a single, acrocentric *S. arvensis* chromosome. With the exception of plants containing 38 chromosomes or less, all BC_3S individuals with adult plant resistance were found to possess a similar acrocentric *S. arvensis* addition chromosome. A second, metacentric addition chromosome was sometimes also present, however plants with only a metacentric addition showed no resistance. Of course it cannot be ruled out that the resistance might in each case be carried on one or more translocations, independent of the addition chromosome(s), that cannot be detected by GISH.

The BC_3 plants used in this study, along with 1 BC_3S individual, showed indications of mitotic instability leading to somatic variations in chromosome number. This phenomenon, probably caused by different parental cell cycles influencing the respective *B. napus* and *S. arvensis* chromosomes, can result in incongruous chromosome count results, particularly in earlier backcross offspring with more alien chromosomes where more mitotic disturbances might be expected. This could make a cytogenetic selection of resistant individuals with a minimal number of donor chromosomes difficult. The BC_1 plant from which all the BC_3 plants in the present study originated were originally scored, using classical cytogenetic techniques, as having chromosome numbers from 2n=37–39 (Plümper 1995). The higher chromosome counts in some later backcross offspring could be explained by the mixoploid nature of the BC_1 plant.

While GISH has been shown, both in this study and previously (e.g. Fahleson et al. 1997; Sharzhinskaya et al. 1998; Snowdon et al. 1998), to be very effective for the detection of addition chromosomes in intergeneric hybrids between *Brassica* species and close relatives, its utility for localising small translocations in such hybrids is perhaps questionable. Chromosome arms in *Brassica* and species of related genera contain unusually low copy numbers of dispersed repeat sequences (Heslop-Harrison and Schwarzacher 1996). Because such dispersed repeats generally form the basis of chromosome ‘‘painting’’ and GISH signals, GISH in *Brassica* is therefore normally characterised by strong signals at centromeric heterochromatin and only very weak hybridisation on chromosome arms (cf. Fig. 2f-h). A translocation in a backcross individual from a *B. napus-Raphanus sativus* hybrid has been successfully detected by GISH (Snowdon et al. 1999), however translocations in non-heterochromatic regions may be beyond the resolution limits of this technique in *Brassica*.

Because of the two phenomena described above, it cannot be stated with absolute certainty that the individuals observed in this study with adult plant resistance, but no visible GISH signals, indeed contain introgressed blackleg resistance genes. Alternatively, the resistance might be carried on one or more addition chromosomes that were not seen due to mitotic instability. This appears unlikely, however, because for the plants in question a large number of metaphases and interphases were scored with no variation in the GISH results. Conclusive evidence and characterisation of the chromosome introgressions, where present, will be obtained by detailed molecular genetic analysis of the plant material.

Acknowledgements The authors thank Bernhard Plümper for producing and analysing the original hybrids and early backcross generations, and Hannelore Lehmann and Yvonne Bernhardt for technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (DFG Ko701–15/1 and Sa 488/8–1) and the Professor Werner Schulze-Stiftung im Stifterverband für die Deutsche Wissenschaft.

References

Bing DJ, Downey RK, Rakow GFW (1996) Hybridizations among *Brassica napus*, *B. rapa* and *B. juncea* and their two weedy relatives *B. nigra* and *Sinapis arvensis* under open pollination conditions in the field. Plant Breed 115: 470–473

Cargeeg LA, Thurling N (1979) Seedling and adult plant resistance to blackleg [*Leptosphaeria maculans* (Desm.) Ces. et de Not.] in spring rape (*Brassica napus* L.). Aust J Agric Res 30: 37–46

Cargeeg LA, Thurling N (1980) Contribution of host-pathogen interactions to the expression of the blackleg disease of spring rape (*Brassica napus* L.) caused by *Leptosphaeria maculans* (Desm.) Ces. et De Not. Euphytica 29: 465–476

Chèvre AM, Eber F, Baranger A, Kerlan MC, Barret P, Festoc G, Vallée P, Renard M (1996a) Characterization of *Brassica nigra* chromosomes and of blackleg resistance in *B. napus-B. nigra* addition lines. Plant Breed 115: 113–118

Chèvre AM, Eber F, This P, Barret P, Tanguy X, Brun H, Delseny M, Renard M (1996b) Interspecific gene flow as a component of risk assessment for transgenic *Brassicas*. (Proc Int Symp *Brassicas*/9th Crucifer Genet Workshop). Acta Hortic 407: 169–179

Chèvre AM, Barret P, Eber F, Dupuy P, Brun H, Tanguy X, Renard M (1997) Selection of stable *Brassica napus-B. juncea* recombinant lines resistant to blackleg (*Leptosphaeria maculans*). 1: Identification of molecular markers, chromosomal and genomic origin of the introgression. Theor Appl Genet 95: 1104–1111

Dixelius C (1999) Inheritance of the resistance to *Leptosphaeria maculans* of *Brassica nigra* and *B. juncea* in near-isogenic lines of *B. napus*. Plant Breed 118: 151–156

Fahleson J, Lagercrantz U, Mouras A, Glimelius K (1997) Characterization of somatic hybrids between *Brassica napus* and *Eruca sativa* using species-specific repetitive sequences and genomic in situ hybridization. Plant Sci 123: 133–142

Friedt W, Lühs WW (1998) Recent developments and perspectives of industrial rapeseed breeding. Fat-Lipid 100: 219–226

Hammond KE, Lewis BG (1987) Variation in stem infections caused by aggressive and non-aggressive isolates of *Leptosphaeria maculans* on *Brassica napus* var. oleifera. Plant Pathol 36: 53–65

Hassan AK, Schulz C, Sacristán MD, Wöstemeyer J (1991) Biochemical and molecular tools for the differentiation of aggressive and non-aggressive isolates of the oilseed rape pathogen, *Phoma lingam*. J Phytopathol 131: 120–136

Heslop-Harrison JS, Schwarzacher T (1996) Genomic Southern and *in situ* hybridization for plant genome analysis. In: Jauhar PP (ed) Methods of genome analysis in plants. CRC Press, Boca Raton, Fla., pp 163–179

Jiang J, Gill BS (1996) Current status and potential of fluorescence *in situ* hybridization in plant genome mapping. In: Patterson AH (ed) Genome mapping in plants. Academic Press, San Diego, pp 127–135

Kapila R, Negi MS, This P, Delseny M, Srivastava PS, Lakshmikumaran M (1996) A new family of dispersed repeats from *Brassica nigra*: characterization and localization. Theor Appl Genet 93:1123–1129

Kerlan MC, Chèvre AM, Eber F (1993) Interspecific hybrids between transgenic rapeseed (*Brassica napus*) and related species: cytogenetical characterization and detection of the transgene. Genome 36: 1099–1106

Kuswinanti T, Koopmann B, Hoppe HH (1999) Virulence pattern of aggressive isolates of *Leptosphaeria maculans* on an extended set of *Brassica* differentials. J Plant Dis Protect 106: 12–20

Lefol E, Danielou V, Darmency H (1996) Predicting hybridization between transgenic oilseed rape and wild mustard. Field Crops Res 45: 153–161

Mahuku GS, Hall R, Goodwin (1996) Co-infection and induction of systemic acquired resistance by weakly and highly virulent isolates of *Leptosphaeria maculans* in oilseed rape. Physiol Mol Plant Pathol 49: 61–72

Mizushima U (1950) Karyogenetic studies of species and genus hybrids in the tribe Brassiceae of Cruciferae. Tohoku J Agric Res 1: 1–14

Pang ECK, Halloran GM (1996) The genetics of blackleg [*Leptosphaeria maculans* (Desm.) Ces. et De Not.] resistance in rapeseed (*Brassica napus* L.). II. Seedling and adult-plant resistance as quantitative traits. Theor Appl Genet 93: 941–949

Plieske J, Struss D, Röbbelen G (1998) Inheritance of resistance derived from the B-genome of *Brassica* against *Phoma lingam* in rapeseed and the development of molecular markers. Theor Appl Genet 97: 929–936

Plümper B (1995) Somatische und sexuelle Hybridisierung für den Transfer von Krankheitsresistenzen auf *Brassica napus* L. PhD thesis, Freie Universität Berlin

Purwantara A, Salisbury PA, Burton WA, Howlett BJ (1998) Reaction of *Brassica juncea* (Indian mustard) lines to Australian isolates of *Leptosphaeria maculans* under glasshouse and field conditions. Eur J Plant Pathol 104: 895–902

Rimmer SR, van den Berg CGJ (1992) Resistance of oilseed *Brassica* spp. to blackleg caused by *Leptosphaeria maculans*. Can J Plant Pathol 14: 56–66

Roy NN (1978) A study on disease variation in the populations of an interspecific cross of *Brassica juncea* L. × *B. napus* L. Euphytica 27: 145–149

Sacristán MD (1982) Resistance responses to *Phoma lingam* of plants regenerated from selected cell and embryogenic cultures of haploid *Brassica napus*. Theor Appl Genet 61: 193–200

Sacristán MD, Gerdemann M (1986) Different behavior of *Brassica juncea* and *B. carinata* as sources of *Phoma lingam* resistance in experiments of interspecific transfer to *B. napus*. Plant Breed 97: 304–314

Schwarzacher T, Leitch AR, Heslop-Harrison JS (1994) DNA: DNA *in situ* hybridization – methods for light microscopy. In: Harris N, Oparka KJ (eds) Plant cell biology: a practical approach. Oxford University Press, Oxford, pp 127–155

Sharzhinskaya M, Fahleson J, Glimelius K, Mouras A (1998) Genome organization of *Brassica napus*, *Lesquerella fendleri* and analysis of their somatic hybrids using genomic *in situ* hybridization. Genome 41: 691–701

Sjödin C, Glimelius K (1989) Transfer of resistance against *Phoma lingam* to *Brassica napus* by asymmetric somatic hybridization combined with toxin selection. Theor Appl Genet 78: 513–520

Snowdon RJ, Köhler W, Friedt W, Köhler A (1997) Genomic in situ hybridization in *Brassica* amphidiploids and interspecific hybrids. Theor Appl Genet 95: 1320–1324

Snowdon RJ, Köhler W, Friedt W, Lühs W, Köhler A (1998) Detecting R-genome chromatin in nematode-resistant *Brassica napus* × *Raphanus sativus* hybrids. In: Lelley T (ed) Current topics in plant cytogenetics related to plant improvement. WUV-Universitätsverlag, Vienna, pp 220–226

Snowdon RJ, Köhler A, Köhler W, Friedt W (1999) FISHing for new rapeseed lines: the application of molecular cytogenetic techniques to *Brassica* breeding. In: Groupe Consultatif International de Recherche Sur Le Colza (GCIRC), Canberra, Australia

Song KM, Osborn TC, Williams PH (1988) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor Appl Genet 75: 784–794

Storck M, Sacristán MD (1995) The role of phytoalexins in the seedling resistance to *Leptosphaeria maculans* in some crucifers. Z Naturforsch 50c: 15–20

Struss D, Quiros CF, Plieske J, Röbbelen G (1996) Construction of *Brassica* B genome synteny groups based on chromosomes extracted from three different sources by phenotypic, isozyme and molecular markers. Theor Appl Genet 93: 1026–1032

Tsukamoto C, Furuya M, Chikayasu K, Okubo K, Hinata K (1993) Chemotaxonomic markers in *Brassica* seeds at the species and subspecies levels. Biosci Biotechnal Biochem 57: 653–654

Warwick SI, Black, LD (1991) Molecular systematics of *Brassica* and allied genera (subtribe Brassicinae, Brassiceae) – chloroplast genome and cytodeme congruence. Theor Appl Genet 82: 81–92

Winter H, Gaertig S, Diestel A, Sacristán MD (1999) Blackleg resistance of different origin transferred into *Brassica napus*. In: Groupe Consultatif International de Recherche Sur Le Colza (GCIRC), Canberra, Australia

Wu JG, Li Z, Liu Y, Liu HL, Fu TD (1997) Cytogenetics and morphology of the pentaploid hybrid between *Brassica napus* and *Orychophragmus violaceus* and its progeny. Plant Breed 116: 251–257

5 Genetic mapping and quantitative trait analysis

The following two articles present examples for genetic mapping and the analysis of quantitative trait loci (QTL) in oilseed rape. Both papers present new *B. napus* genetic maps including numerous previously unmapped simple sequence repeat (SSR) markers. The first article describes the first successful localisation of QTL for resistance against Verticillium wilt caused by *Verticillium longosiporum*. Due to the difficulties of accurate resistance screening in this complex host-pathogen system, the availability of effective markers is vital for breeding efforts. This paper therefore represents an important step towards marker-assisted introgression and combination of new resistances. This will help combat the growing threat of *V. longisporum* to sustainable, large-scale oilseed rape cultivation. The second paper was the first publication to describe co-localising QTL for seed colour and seed coat fibre compounds in oilseed rape. Yellow seed colour is associated with a considerable reduction in antinutritive phenolic compounds in the seed meal, however breeding and selection are difficult due to complex genotype-environment interactions. This paper represents a major step in the development of molecular markers for breeding and towards the map-based gene cloning of key genes involved in seed meal quality traits.

5.1 Identification of quantitative trait loci for resistance against *Verticillium longisporum* in oilseed rape (*Brassica napus* L.)

Rygulla W., R.J. Snowdon, W. Friedt, I. Happstadius, W. Cheung & D. Chen, 2008. Phytopathology 98: 215-221

Identification of Quantitative Trait Loci for Resistance Against *Verticillium longisporum* in Oilseed Rape (*Brassica napus*)

W. Rygulla, R. J. Snowdon, W. Friedt, I. Happstadius, W. Y. Cheung, and D. Chen

First, second, and third authors: Department of Plant Breeding, Research Centre for BioSystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, D-35392 Giessen, Germany; fourth author: Svalöf Weibull AB, S-268 81 Svalöv, Sweden; and fifth and sixth authors: DNA Landmarks Inc., Saint-Jean-sur-Richelieu 84, Quebec, Canada.

Accepted for publication 19 September 2007.

ABSTRACT

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

Verticillium longisporum is one of the major pathogens of oilseed rape (*Brassica napus*; genome AACC, 2n = 38) in Europe. Current European cultivars possess only a low level of resistance against *V. longisporum*, meaning that heavy infection can cause major yield losses. The aim of this study was to identify quantitative trait loci (QTL) for resistance against *V. longisporum* as a starting point for marker-assisted breeding of resistant cultivars. Resistance QTL were localized in a segregating oilseed rape population of 163 doubled haploid (DH) lines derived by microspore culture from the F_1 of a cross between two *B. napus* breeding lines, one of which exhibited *V. longisporum* resistance derived by pedigree selection from a resynthesized *B. napus* genotype. A genetic map was constructed comprising 165 restriction fragment length polymorphism, 94 amplified fragment length polymorphism and 45 simple sequence repeats (SSR) markers covering a total of 1,739 cM on 19 linkage groups. Seedlings of the DH lines and parents were inoculated with *V. longisporum* isolates in four greenhouse experiments performed in Sweden during autumn 1999. In three of the experiments the DH lines were inoculated with a mixture of five isolates, while in the fourth experiment only one of the isolates was used. The intention was to simulate four different environments with variable disease pressure, while still maintaining uniform conditions in each environment to enable reliable disease scoring. The disease index (DI) was calculated by scoring symptoms on a total of 21 inoculated plants per line in comparison to 21 noninoculated plants per line. Using the composite interval mapping procedure a total of four different chromosome regions could be identified that showed significant QTL for resistance in more than one environment. Two major QTL regions were identified on the C-genome linkage groups N14 and N15, respectively; each of these QTL consistently exhibited significant effects on resistance in multiple environments. The presence of flanking markers for the respective QTL was associated with a significant reduction in DI in the inoculated DH lines.

Verticillium wilt caused by *Verticillium longisporum* (ex. *Verticillium dahliae* var. *longisporum* Stark) is one of the major diseases of oilseed rape (2n = 38, genome AACC) in Europe. Current European cultivars possess only a low level of resistance against *V. longisporum*, meaning that heavy infection can cause major yield losses (9,34). Because long-term crop rotation is not always possible and no efficient chemical control is available, the disease is becoming a serious problem in some areas of Germany, France, Sweden, and other major European rapeseed-producing countries (14). The fungus infects the plant via the roots by penetrating the root epidermis in the root hair zone, and establishes itself in the xylem (10,35). Once it has reached the xylem vessels it forms conidiospores, which spread inside the xylem throughout the vegetative parts of the plant. Symptoms usually appear late in the growing season on mature plants, when diseased plants wilt and often die. The fungus produces microsclerotia in plant debris that can survive in the soil for long periods of time.

Brassica napus is a young species that originated through spontaneous interspecific hybridizations between *B. rapa* L. (syn. *campestris*; genome AA, 2n = 20) and *B. oleracea* L. (CC, 2n = 18) (26). After intensive breeding for oil and seed quality traits, the spring and winter oilseed forms of *B. napus* now represent one of the most important sources of vegetable oil worldwide. Due to an extremely strong bottleneck selection for zero seed erucic acid and low seed glucosinolate content, current oilseed rape cultivars have a comparably narrow genetic basis (12); in particular, the availability of novel germplasm for resistance breeding is often extremely limited. On the other hand, sources of *Verticillium* resistance that can potentially be used in rapeseed breeding have been identified in *B. oleracea* and *B. rapa* (6,11,24,25). For example, Happstadius et al. (11) selected the most resistant *B. oleracea* accessions identified in a greenhouse screening and demonstrated that their resistance was effective against *V. longisporum* when introduced into resynthesized (RS) *B. napus* breeding lines by interspecific hybridization. We have produced further RS rapeseed breeding lines with high levels of resistance from cabbage accessions combined with oilseed turnip rape (24,25), while Debode et al. (6) identified resistance in cauliflower cultivars and were able to show that resistant cultivars can prevent the pathogen from spreading systemically inside the plant. The genetic basis of these resistance sources is not known, however, and molecular markers linked to resistance loci and therefore exploitable for marker-assisted selection have not previously been described for this disease. The aim of this study was to identify quantitative trait loci (QTL) for resistance against *V. longisporum* in order to gain information on the genetics of new, resistant *B. napus* breeding lines. The study represents a first step towards marker-assisted breeding of resistant oilseed rape cultivars and, in the long term, towards the map-based cloning of the responsible resistance genes.

MATERIALS AND METHODS

Resistance donor and mapping population. The original resistance donor for the mapping population used in the QTL

Corresponding author: R. J. Snowdon
E-mail address: Rod.Snowdon@agrar.uni-giessen.de

doi:10.1094/PHYTO-98-2-0215

analysis was a tetraploid white cabbage (*B. oleracea* convar. *capitata*; genome CCCC, $2n = 4\times = 36$), which was crossed with a tetraploid, low erucic acid winter turnip rapeseed (*B. rapa* ssp. *olerifera*; genome AAAA, $2n = 4\times = 40$) to generate a synthetic *B. napus*. This breeding line was produced in Sweden in 1975 to 1976 with the aim of increasing the genetic variability in *B. napus* germplasm, and the resistance was subsequently transferred to oilseed rape breeding lines by pedigree selection. The parents of the mapping population were homozygous doubled haploid (DH) breeding lines that were identified by greenhouse screenings of two DH populations for resistance to *V. longisporum*. The parental line 307-406-1 was found to be susceptible to Verticillium wilt, while the parental line 307-230-2 was considered to be resistant/tolerant to the disease. Three F_1 donor plants were used for microspore cultivation as described in Zhou et al. (36). A total of 3,453 in vitro embryos were induced, 478 shoots were regenerated, and 398 plants were transplanted to soil. At the three- to four-leaf stage the ploidy number of each plant was analyzed by flow cytometry using standard methods (33). In 40% of the plantlets, the genome had doubled spontaneously, 6% were polyploids, and 54% were still haploid and consequently treated with colchicine to induce diploidization. After vernalization, 329 plants were transferred to the greenhouse and selfed by bagging. One hundred and sixty-three randomly chosen DH lines were used for the QTL mapping study.

Resistance screening. Seedlings of the DH lines and parents were inoculated with *V. longisporum* isolates in four separate greenhouse experiments during autumn 1999 at Svalöf Weibull AB, Sweden. The intention was to simulate four different environments with variable disease pressure, while still maintaining uniform conditions in each environment to enable reliable disease scoring. Since the resistance was expected to be quantitative in nature, this strategy was used to enable more reliable identification of QTL that were stable over different environments.

In three of the experiments a mixture of the Swedish *V. longisporum* isolates Vd1, Vd4, Vd11, Vd12, and Vd13 in equal densities was used for the inoculation. The characteristics of these five isolates, which show a similar virulence spectrum and moderate to high aggressiveness, have been described in detail previously (27,28). In the fourth experiment, the inoculation was performed with the single-spore isolate Vd13, in order to determine if any identified QTL might be explained by a race-specific resistance. Vd13 shows consistent, moderate disease symptoms, disease severity, and DI with different inoculation techniques in susceptible *B. napus* cultivars (28). The same overall inoculum density (2×10^6 spores/ml + mycelium fragments) was used for the inoculations with the isolate mixture and for the single-spore isolate.

Each experiment included a total of 21 inoculated and 21 noninoculated plants for each of the DH lines, along with 10 repetitions of 21 inoculated and 21 noninoculated plants of each of the parental lines. Furthermore, 21 inoculated and 21 noninoculated plants of the susceptible *B. napus* 'Libraska' were included in each experiment as an infection control, making a total of 3,654 inoculated and 3,654 noninoculated plants in each of the four experiments. The plants were grown in individual pots with the genotypes arranged in a completely random design. The inoculations were carried out using a root-dip inoculation method that was shown in a comparison of different *Verticillium* inoculation techniques to give the most consistent results (15): Bottom-broken peat pots containing 14-day-old seedlings were dipped into 40 ml of inoculum. The pots were left in the inoculum suspension for 30 min and then replaced at the original position in a planting tray. The remaining inoculum was recovered with rinsing water and poured into the container with the inoculated seedlings. The seedlings were then grown at 20°C in a greenhouse with a light regime of 16 h day/8 h night; the relative humidity was not controlled in order to allow environmental variation between the four experiments in the interests of the QTL analysis. Disease scoring was carried out between 8 to 10 weeks after inoculation, when 80 to 90% of the inoculated plants from *B. napus* Libraska showed wilting symptoms.

Disease severity scores (0 to 5) for each inoculated plant followed a scale (28) based on the percentage of leaves showing wilting and chlorosis, along with the degree of stunted growth (height in mm) in comparison to the mean of the noninoculated plants of the same genotype: 0 = no symptoms, 1 = wilting and/or stunting <25%, 2 = wilting and/or stunting 25 to 50%, 3 = wilting and/or stunting 51 to 90%, 4 = wilting and stunting ≥91%, and 5 = dead plant. From the number of plants in each class, DI was calculated as follows for each of the DH lines and parents, as described by Happstadius et al. (11):

$$\mathrm{DI} = \frac{(0\times N_0)+(1\times N_1)+(2\times N_2)+(3\times N_3)+(4\times N_4)+(5\times N_5)}{(N_0+N_1+N_2+N_3+N_4+N_5)} - \left(\mathrm{DI}_{\mathrm{n,\,control}} - \mathrm{DI}_{\mathrm{total,\,control}}\right)$$

where N_n is the number of plants in the respective class, $\mathrm{DI}_{\mathrm{n,\,control}}$ represents the mean symptom score of the noninoculated control plants in the *n*th trial, and $\mathrm{DI}_{\mathrm{total,\,control}}$ is the mean symptom score of the noninoculated controls over all trials.

Statistical analyses of the data were performed using the software package SAS 9.1. Analysis of variance (ANOVA) was performed using the general linear model procedure in order to estimate differences within the population and among the four different experiments (environments). ANOVA was also used to analyze genotype-environment interactions of the two parental lines, for which 10 DI values (each calculated from 21 inoculated and 21 noninoculated plants) per line per experiment were available. Variation in DI within and among the four experiments was further quantified by calculation of least significant differences among the means of the parents and the DH lines, along with Pearson's correlation coefficients for the individual genotypes in the four different environments.

Restriction fragment length polymorphism (RFLP) probes and Southern analysis. Plant genomic DNA was extracted for RFLP analysis as described by Cheung et al. (4). Three restriction enzymes (*Eco*RI, *Eco*RV, and *Hin*dIII) were applied to digest the genomic DNA. The DNA digests were separated in 0.8% agarose gels and transferred to nylon membranes (H-bond N+). Six sets of DNA membranes for each enzyme digest were prepared for Southern analysis. DNA probes coded DLM2 to DLM372 correspond to anonymous clones from a seedling cDNA library of *B. napus*. Procedures for Southern analysis were described by Cheung et al. (4). Duplicated loci detected by a single clone were designated by the same name followed by a different lowercase letter, as described by Landry et al. (17).

Amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) analysis. Marker saturation of the genetic map was achieved using AFLP markers, while previously mapped loci from publicly-available SSR primer combinations were also included in the map to enable identification of linkage groups. For the polymerase chain reaction (PCR) marker analyses new genomic DNA samples were isolated from dry frozen leaf material according to Doyle and Doyle (8). AFLP amplification products were generated using the *Eco*RI-*Mse*I-based AFLP kit of Invitrogen (AFLP Core Reagent Kit, Invitrogen Life Technologies Inc., Carlsbad, CA) according to the manufacturer's instructions. *Eco*RI primers were fluorescently labeled with IRD700 or IRD800. DNA from DH lines and their parents were amplified with a set of 10 AFLP primer combinations known either from the literature (18) or from previous mapping studies (1,2) to reveal a high polymorphism rate in oilseed rape.

A total of 127 SSR primer combinations for which genetic map positions of one or more loci in *B. napus* are known (2,19,23), were screened for polymorphism among the mapping parents. Of these primer combinations 65 (51.2%) amplified at least one polymorphic locus and were screened in the DH lines. For all SSR analyses the M13-tailing procedure described by Berg and Olaisen (3) was used. In this method the fluorescently labeled 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 cycle of amplification the PCR fragments are subsequently amplified by the labelled universal primer. All AFLP and SSR amplification products were separated using a LI-COR 4200 DNA Analyzer (LI-COR Biosciences, Lincoln, NE) and scored visually.

Segregation analysis and map construction. A genetic map was constructed using the program JoinMap 3.0 (31). The fit of allelic segregation to the expected 1:1 segregation ratio was tested for each marker by a chi-square test, and map construction was performed in three steps based on the suggestions of Cloutier et al. (5) for *B. napus* mapping. (i) Selection of loci whose segregation fit the expected ratio ($P < 0.01$) to identify linkage groups with high confidence; these markers were anchored as "fixed orders". (ii) Addition of markers with segregation other than the expected 1:1 ratio ($P \geq 0.01$) using a minimum logarithm of odds (LOD) threshold of 2.0. These markers were only added when they did not influence the grouping of the fixed-order markers and did not form noticeable clusters with other skewed markers. (iii) Identification of linkage groups based on SSR markers from consensus maps (1,2,23). Map distances measured in cM between markers were derived from the Kosambi function (16).

Composite interval mapping of resistance QTL. QTL conferring resistance to *V. longisporum* in the DH lines were localized by composite interval mapping with cofactors, using the computer software PLABQTL (30) Version 1.2. All markers proposed by the default option SELECT were used as cofactors. Phenotype data from each of the four experiments were first analyzed separately, then the data from the three experiments with the isolate mixture were averaged to compare the locations of QTL detected over different environments under the same infection conditions. Initially, all putative QTL for resistance were identified using an LOD threshold of 3.0. From these putative QTL, the ones exhibiting a significant likelihood ratio that a QTL is present versus absent were identified based on the multiple regression of the final simultaneous fit function of PLABQTL. After a further regression analysis based on the phenotypic variance, only QTL that showed a significant LOD score in combination with a significant phenotypic effect ($\theta = 0.02$) were selected, and we considered these to be trait-relevant QTL only when the confidence intervals overlapped in two or more of the four experiments.

RESULTS AND DISCUSSION

Phenotypic evaluation of the mapping population. Figure 1 shows the distribution of DI values for the DH lines in the four experiments. In all cases, a continuous distribution was observed, with the majority of DH lines having an intermediate DI between the two parents; this indicates a quantitative resistance for both the isolate mixture and the single-spore isolate. For both the single-spore isolate and the isolate mixture, a number of DH lines showed a lower DI than the resistant parent, or a higher susceptibility than the susceptible parent, suggesting a transgressive segregation with some resistance alleles also being contributed by the susceptible parent. In the experiment with the single-spore isolate, the DI values in the DH population showed a more or less normal distribution, with the majority of the DH lines showing an intermediate resistance phenotype and few lines being highly susceptible or highly resistant. On the other hand, in the three experiments with the isolate mixture, both the distribution and the mean DI values of the two parents were strongly skewed towards greater susceptibility. This indicates that the isolate mixture raises the disease pressure and may overcome part of the resistance against the single-spore isolate.

Considerable variation was sometimes observed among the symptom scores of individual plants within the same DH line,

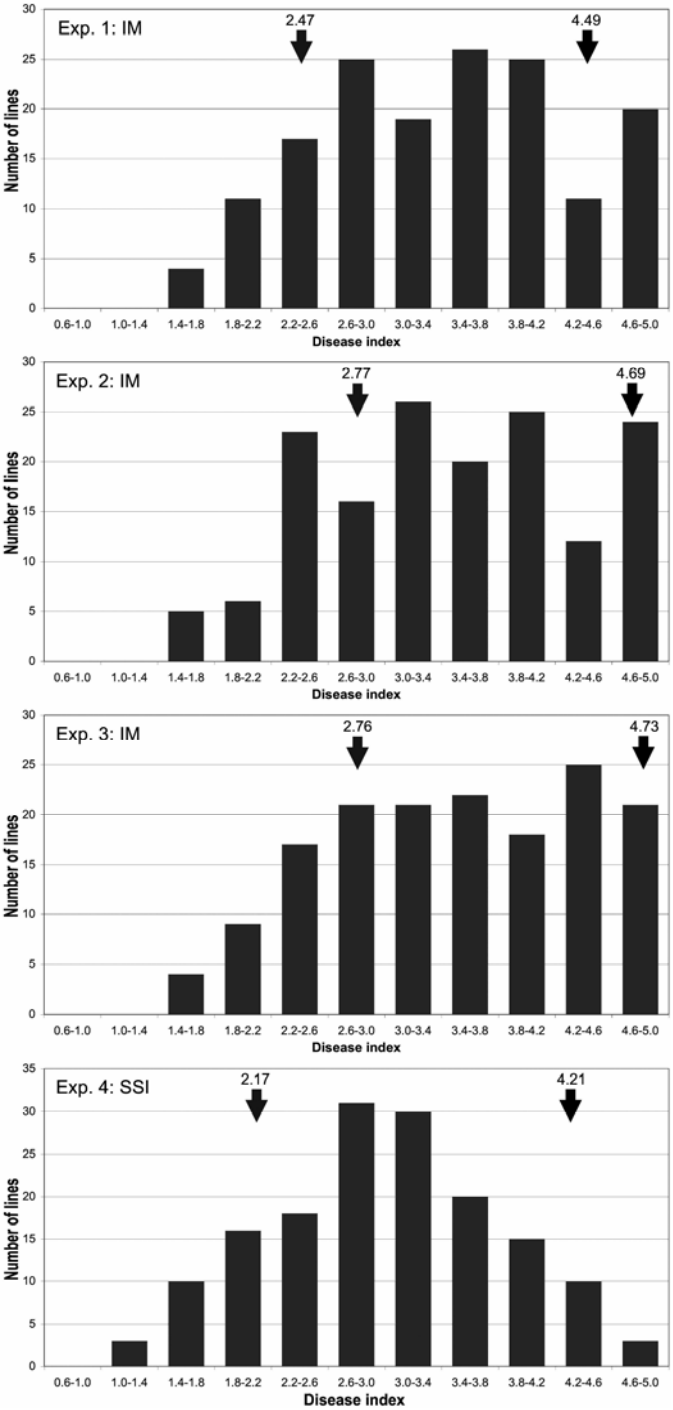


Fig. 1. Distributions of *Verticillium longisporum* disease index (DI) values in a population of 163 doubled-haploid *Brassica napus* lines derived from a single F_1 individual from the population 307-230-2 (resistant)/307-406-1 (susceptible). The disease index was calculated from symptom scores on 21 plants per line after root-dip inoculation, compared to 21 noninoculated plants per line, in four greenhouse experiments in autumn 1999. In experiments 1 to 3, an isolate mixture was used, while experiment 4 was performed by inoculation with a single spore isolate. The left and right arrows show the mean DI values of the resistant and susceptible parents, respectively; parental means were calculated from 10 repetitions of 21 inoculated and 21 noninoculated plants in each of the four experiments. Exp.: experiment; IM: isolate mixture; SSI: single-spore isolate.

possibly indicating that variable amounts of inoculum may be taken up during the disease screening. Despite this variation at a single-plant level, statistical analysis of the DI data showed that a reliable and reproducible quantitative estimate of the disease score was still obtained by calculating the DI from 21 inoculated plants, compared to 21 noninoculated plants in each experiment. Since the mean DI values of the parents and the DH lines showed high reproducibility over the three different experiments with the isolate mixture, we decided to also include the DI values from the inoculation with the single-spore isolate in the QTL analysis, even though only one experimental repetition was available in this case. In the ANOVA, the DI values of the parents and the DH lines were significantly lower ($P < 0.01$) after inoculation with the single-spore isolate than after the inoculations with the isolate mixture; this difference can be clearly seen in the distributions shown in Figure 1 and was confirmed by the results of the least significant difference (LSD) test (Table 1). Despite this difference in overall symptom level, however, the mean DI values of the individual genotypes infected with the single-spore isolate were still significantly correlated ($P < 0.01$) to the mean DI values after the inoculations with the isolate mixture (Table 2). This is presumably because a large number of DH lines were resistant against both the single-spore isolate and the isolate mixture. The ANOVA also revealed that there was no significant genotype × environment interaction ($P = 0.96$) over the four experiments.

Linkage map. A total of 96 polymorphic SSR markers, 247 polymorphic AFLP markers, and 346 polymorphic RFLP markers were scored in the DH population 307-230-2/307-406-1. Due to the allopolyploid nature of the *B. napus* genome a large proportion of duplicated loci are expected, which can distort the marker segregation and lead to mapping errors (5); these can be recognized because the marker segregation ratios deviate significantly from the expected 1:1 ratio in a DH population. In order to obtain the best possible map, we therefore performed a stringent selection for markers with correct or only moderately skewed segregation, and to facilitate the composite interval mapping of QTL we further removed mapped markers with less than 1 cM intervals. Overall these measures resulted in the exclusion of 53.1% of the available SSR markers, 61.1% of the AFLP markers, and 52.3% of the RFLP markers from the mapping. The resulting genetic map comprised 165 RFLP, 94 AFLP, and 45 SSR markers on 19 linkage groups (LG) covering 1,739 cM, with an average marker distance of 5.7 cM (Fig. 2). A total of 68 markers (22%) with moderately skewed distribution ($P < 0.01$) were included in the map, but only when they did not form noticeable clusters. Based on comparisons to other maps, all linkage groups could be associated to defined *B. napus* chromosomes based on the accepted N1-N19 nomenclature for *B. napus* (22). The size of the groups varied between 46.3 and 158.1 cM, while the number of molecular markers per group ranged from 3 to 35. A maximum of seven SSR markers were identified for each linkage group.

QTL for *V. longisporum* resistance. Four chromosome regions were identified that contained significant QTL for resistance to *V. longisporum* (defined as significant LOD scores after multiple regression combined with significant phenotypic effects at a threshold of $\theta = 0.02$) in at least two of the four experiments (Fig. 2). In all cases the resistance alleles at the detected QTL were derived from the resistant parent 307-230-2. The relative degree of differentiation in disease response within this mapping population was comparatively low, however together the four significant QTL explained a total of 45.7% (sum of partial R^2) of the phenotypic variance in the *V. longisporum* DI. To our knowledge this is the first report of QTL for resistance to *V. longisporum* in *B. napus*, hence these results represent an important first step in the investigation of the genetic basis of currently available resistance sources.

On N15, a QTL region with overlapping confidence intervals was detected in two of the three experiments with the isolate mixture, and also with the mean data for the three experiments with the isolate mixture (Table 3). A QTL whose confidence interval was adjacent to this region was detected in the other experiment with the isolate mixture. No significant QTL was detected at this position in the experiment with the single spore isolate, however the marker BRMS030 flanking the QTL had significant effects (*t* test, $P < 0.05$) on the mean DI for the single-spore isolate. The power of QTL detection may have been lower in the experiment with the single-spore isolate, since the majority of the DH lines showed an intermediate resistance response; this could explain why no significant QTL was detected in this case.

On N14 another QTL region was found with overlapping QTL confidence intervals for one of the three experiments with the isolate mixture, the mean data over all environments with the isolate mixture, and for the experiment with the single-spore isolate. Although no significant QTL were detected at this position in the other two experiments with the isolate mixture, the flanking markers (DLM138b and CB10079) still both had significant effects on the mean DI in these replications (*t* test, $P < 0.01$), suggesting that this region also contains one or more genes with a significant effect on the resistance reaction. The QTL for the experiment with the single-spore isolate was at a slightly different position to the QTL from the experiments with the isolate mixture, which might indicate a race-specific resistance gene that was no longer effective against the isolate mixture.

In the mean data from the three experiments with the isolate mixture, the two QTL on N14 and N15 accounted for 9.9 and 12.5% of the phenotypic variation (partial R^2), respectively. Because these two QTL had consistently significant effects on the

TABLE 1. Comparison of mean least significant differences in *Verticillium longisporum* disease index values in a population of 163 doubled haploid *Brassica napus* lines, and their parents 307-230-1 (susceptible) and 307-230-2 (resistant), after root-dip inoculation in four experiments performed in the greenhouse in autumn 1999

	Mean disease index[v]			
Experiment	DH[w] lines	307-230-1	307-406-2	LSD[x] ($\alpha = 0.05$)
Exp. 1: IM[y]	3.45^a	2.47^c	4.49^e	0.63
Exp. 2: IM[y]	3.50^a	2.77^c	4.69^e	0.64
Exp. 3: IM[y]	3.53^a	2.76^c	4.73^e	0.63
Exp. 4: SSI[z]	3.00^b	2.17^d	4.21^f	0.66
LSD[x] ($\alpha = 0.05$)	0.30	0.37	0.25	

[v] The disease index for each doubled haploid line in each experiment was calculated from the symptom scores of 21 inoculated and 21 noninoculated plants per genotype, while the mean disease index of the parental lines was calculated from 10 repetitions of 21 inoculated and 21 noninoculated plants in each experiment. Different mean values that show nonsignificant differences are assigned the same superscript letter.

[w] Doubled-haploid.

[x] Least significant difference.

[y] Isolate mixture.

[z] Single spore isolate.

TABLE 2. Pearson's correlation coefficient for comparisons among *Verticillium longisporum* disease index values of 163 doubled haploid *Brassica napus* lines and their population parents tested by root-dip inoculation in three greenhouse experiments performed in autumn 1999 with an isolate mixture and in one experiment with a single-spore isolate[x]

Experiment	Exp. 2: IM[y]	Exp. 3: IM[y]	Exp. 1-3: Mean	Exp. 4: SSI[z]
Exp. 1: IM[y]	0.64	0.68	0.87	0.64
Exp. 2: IM[y]		0.76	0.90	0.57
Exp. 3: IM[y]			0.91	0.57
Exp. 1-3: Mean				0.66

[x] All correlations were significant with $P < 0.01$.

[y] Isolate mixture.

[z] Single spore isolate.

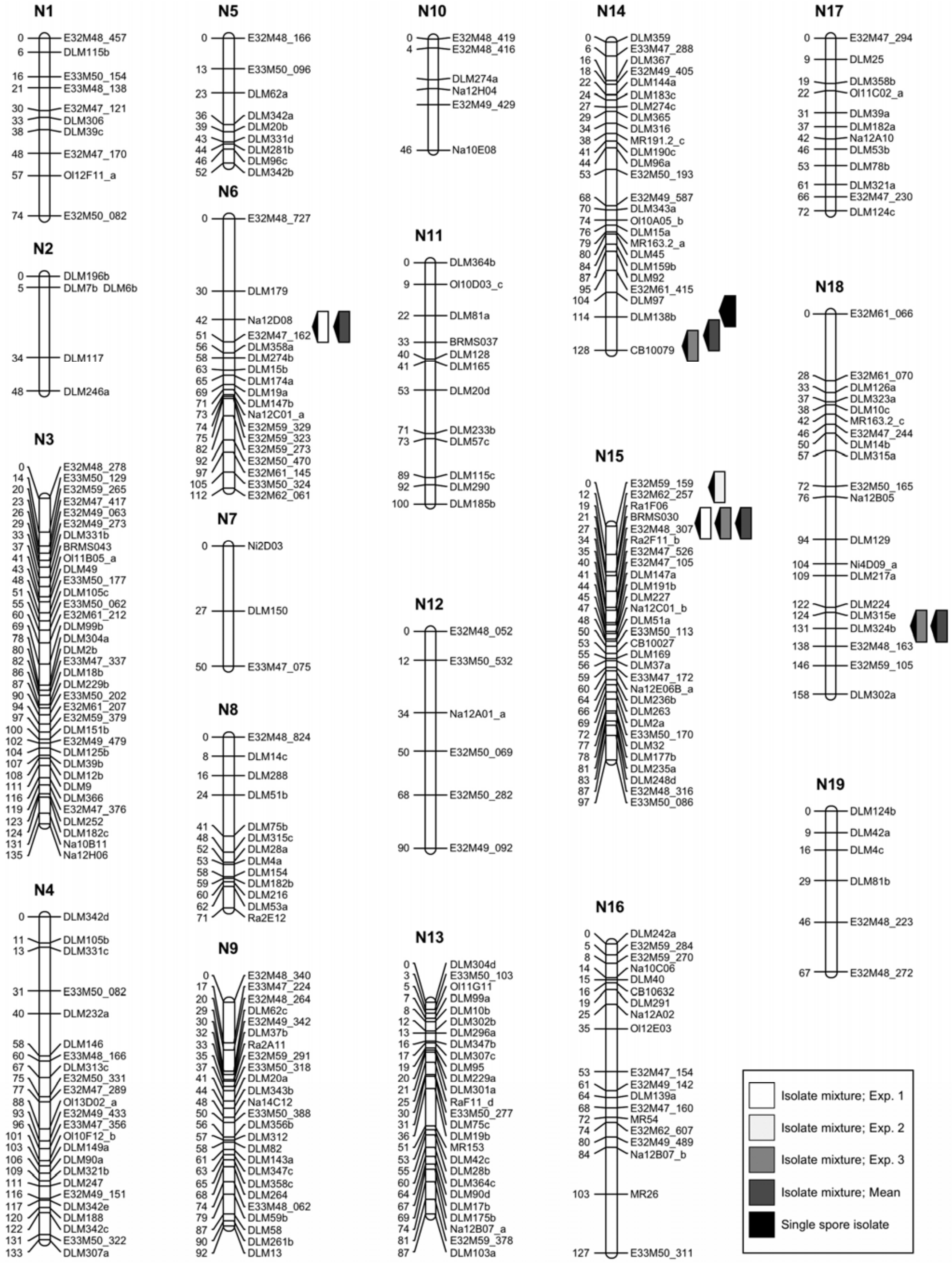


Fig. 2. Genetic linkage map (1,739 cM) for an oilseed rape mapping population (163 doubled haploid lines) derived from the population 307-230-2/307-406-1. The shaded boxes show quantitative trait loci for *Verticillium longisporum* disease index measured in the greenhouse in autumn 1999 after three root-dip inoculation experiments with a *V. longisporum* isolate mixture and one experiment with the single-spore isolate VD13. The map contains 165 restriction fragment length polymorphism, 94 amplified fragment length polymorphism and 45 simple sequence repeat markers. Genetic distances are given in centimorgans.

resistance phenotype in multiple experiments with controlled but not identical inoculation and environmental conditions, we considered these to be environmentally stable QTL with potential for utilization in breeding. The *B. napus* chromosomes N14 and N15 are both derived from the *Brassica* C-genome (22), which is consistent with the origin of the resistance from a *B. oleracea* donor.

On the other hand, linkage groups N6 and N18 each contained one QTL region that showed significant phenotypic effects on resistance in a single experiment with the isolate mixture. Although the genetic effects of these QTL are apparently lower than for the QTL on N14 and N15, they nevertheless represent chromosomal regions of interest for future studies of *V. longisporum* resistance in oilseed rape. The QTL on N6, which accounted for 12.4% of the phenotypic variation (partial R^2) in the resistance reaction in the three experiments with the isolate mixture, is of particular interest because chromosome N6 in *B. napus* is derived from the A genome of *B. rapa* (21) rather than the C-genome of the original *B. oleracea* resistance donor. This could be due to a homoeologous or nonhomoeologous translocation between A and C genome chromosomes, which are known to occur commonly in newly synthesised *B. napus* (29), or alternatively it could indicate a further resistance locus contributed by the *B. rapa* parent of the original RS *B. napus* resistance donor.

Expanding the genetic basis for resistance breeding in oilseed rape. The addition of new sources of resistance against *V. longisporum* to the germplasm of *B. napus* is vital for the sustainable production of oilseed rape in Europe, since there is only limited resistant germplasm currently available for breeding and Verticillium wilt is becoming increasingly important in some rapeseed production areas. Introgression of germplasm from *B. rapa* and *B. oleracea* via RS rapeseed is a breeding strategy that has been successfully used in the past to generate new genetic variation for resistance against important diseases of oilseed rape. For example, protoplast fusion has been successfully used to transfer resistance to clubroot disease (*Plasmodiophora brassicae* Wor.) from *B. rapa* to *B. napus* (7). Through advanced backcrossing a race-specific resistance was subsequently transferred from RS rapeseed to elite winter oilseed rape germplasm, and the winter oilseed rape cultivars Mendel and Tosca derived from this material were released in the early 2000s to specifically combat this disease in affected areas of Britain and Germany. Synthetic lines of *B. napus* carrying resistance to blackleg derived from *B. rapa* were generated via embryo culture (20). The resistance was subsequently introgressed into spring canola, resulting in the release of ‘Surpass’ in the late 1990s.

The introgression of novel resistance genes normally requires extensive backcrossing accompanied by effective resistance screening. However, selection for *V. longisporum* resistance in the field can be difficult due to a high environmental influence on the symptom development, and symptom scoring is further complicated by the frequent co-occurrence of blackleg disease caused by *Leptosphaeria maculans* (Desm) Ces. & de Not. [anamorph *Phoma lingam* (Tode ex Fr.) Desm.] on plants showing *V. longisporum* symptoms (I. Happstadius, *personal observation*, 2007). As shown in Table 4, the combination of *V. longisporum* resistance alleles from different QTL in the present study was associated with increased resistance in the DH lines. This underlines the potential for a marker-assisted pyramiding strategy to assist the effective transfer of polygenic *V. longisporum* resistance from RS rapeseed donors into elite breeding lines.

TABLE 4. Numbers and mean *Verticillium longisporum* disease index values of doubled haploid *Brassica napus* lines containing zero, one, two, three, or four resistance alleles at the four quantitative trait loci (QTL) on *B. napus* chromosomes N6, N14, N15, and N18, compared with the disease index values of the susceptible 307-230-1 and the resistant 307-230-2[v]

Genotypes	Number of resistance alleles present	Number of DH[w] lines[x]	Mean DI[y] value
(307-230-2 / 307-230-1)	0	1	4.62
	1	58	3.69
	2	47	3.39
	3	15	2.74
	4	4	2.36
307-230-1	0	–	4.64[z]
307-230-2	4	–	2.66[z]

[v] The disease index values were calculated from the symptom scores of 21 inoculated and 21 noninoculated plants per genotype. Data is from three root-dip inoculation experiments with an isolate mixture. Thirty-eight double haploid lines contained alleles from both parents at the QTL-flanking markers and were not counted.

[w] Doubled haploid.

[x] Lines were counted as containing a resistance allele when the alleles for both markers flanking the QTL were derived from 307-230-2. The resistance allele was counted as absent when the alleles for both markers flanking a QTL were derived from 307-406-1.

[y] Disease index.

[z] Mean DI values for the parental lines were calculated from 10 repetitions of 21 inoculated and 21 noninoculated plants in each experiment.

TABLE 3. Details of significant quantitative trait loci for resistance against *Verticillium longisporum* estimated from mean disease index data in a population of 163 doubled-haploid lines from the *Brassica napus* population 307-230-2/307-406-1

Experiment[t]	Linkage group	Peak position (cM)[u]	LOD[v] at QTL[w] position	Confidence interval (cM)[u]	Flanking markers[x]	Partial R^2 (%)	Additive effects
Exp. 1: IM[y]	N6	50	6.32	46–56	Na12D08 (41.8) – E32M47_162 (51.2)	9.5	0.29
	N15	22	4.98	18–24	BRMS030 (21.4) – E32M48_307 (26.6)	6.6	0.24
Exp. 2: IM[y]	N15	2	3.08	0–8	E32M59-159 (0) – E32M62_257 (11.8)	7.6	0.30
Exp. 3: IM[y]	N14	128	6.64	120–128	DLM138b (114.3) – CB10079 (128.5)	8.6	0.24
	N15	22	6.68	18–24	BRMS030 (21.4) – E32M48_307 (26.6)	9.7	0.27
	N18	130	3.30	126–136	DLM315e (124.3) – DLM324b (131.5)	7.3	0.23
Exp. 1-3: Mean	N6	50	4.87	46–56	Na12D08 (41.8) – E32M47_162 (51.2)	12.4	0.27
	N14	124	8.46	118–128	DLM138b (114.3) – CB10079 (128.5)	9.9	0.24
	N15	22	10.06	20–24	BRMS030 (21.4) – E32M48_307 (26.6)	12.5	0.27
	N18	130	4.20	124–134	DLM315e (124.3) – DLM324b (131.5)	10.9	0.24
Exp. 4: SSI[z]	N14	112	6.50	104–122	DLM97 (102.5) – DLM138b (112.9)	7.1	0.25

[t] The population was evaluated in four experiments representing different environments. Disease index of each genotype in each experiment was calculated by comparing disease symptoms of 21 plants inoculated by a root-dip procedure with 21 noninoculated plants. Only resistance loci with significant logarithm of odds scores after regression analysis in combination with significant phenotypic effects in more than one experiment are shown.

[u] Centimorgan.

[v] Logarithm of odds.

[w] Quantitative trait loci.

[x] In all cases, the marker alleles associated with increased resistance were derived from the resistant parent 307-230-2.

[y] Isolate mixture.

[z] Single spore isolate.

ACKNOWLEDGMENTS

This work was partially supported by the Society for Promotion of Private German Plant Breeding (GFP) and the Agency for Renewable Resources (FNR) of the Federal Ministry for Nutrition, Agriculture and Consumer Protection (BMELV). We thank D. Baetzel for excellent technical assistance and four anonymous reviewers for constructive comments on an earlier version of this manuscript.

LITERATURE CITED

1. Badani, A. G., Snowdon, R. J., Baetzel, R., Lipsa, F. D., Wittkop, B., Horn, R., De Haro, A., Font, R., Lühs, W., and Friedt, W. 2006. Co-localization of a partially dominant gene for yellow seed color with a major QTL influencing acid detergent fiber (ADF) content in different crosses of oilseed rape (*Brassica napus*). Genome 49:1499-1509.
2. Basunanda, P., Spiller, T. H., Hasan, M., Gehringer, A., Schondelmaier, J., Lühs, W., Friedt, W., and Snowdon, R. J. 2007. Marker-assisted increase of genetic diversity in a double-low seed quality winter oilseed rape genetic background. Plant Breed. 126:581-587.
3. Berg, E. S., and Olaisen, B. 1994. Hybrid PCR sequencing – sequencing of PCR products using a universal primer. Biotechniques 17:896.
4. Cheung, W. Y., Champagne, G., Hubert, N., and Landry, B. S. 1997. Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. Theor. Appl. Genet. 94:569-582.
5. Cloutier, S., Cappadocia, M., and Landry, B. S. 1997. Analysis of RFLP mapping inaccuracy in *Brassica napus* L. Theor. Appl. Genet. 95:83-91.
6. Debode, J., Declercq, B., and Höfte, M. 2005. Identification of cauliflower cultivars that differ in susceptibility to *Verticillium longisporum* using different inoculation methods. J. Phytopathol. 153:257-263.
7. Diederichsen, E., and Sacristan, M. D. 1996. Disease response of resynthesized *Brassica napus* L. lines carrying different contributions of resistance to *Plasmodiophora brassicae* Wor. Plant Breed. 115:5-10.
8. Doyle, J. J., and Doyle, J. L. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.
9. Dunker, S., Keunecke, H., and von Tiedemann, A. 2006. *Verticillium longisporum* in winter oilseed rape – Impact on plant development and yield. Integrated Control in Oilseed Crops 29:365-374.
10. Eynck, C., Koopmann, B., Grunewaldt-Stoecker, G., Karlovsky, P., and von Tiedemann, A. 2007. Differential interactions of *Verticillium longisporum* and *V. dahliae* with *Brassica napus* detected with molecular and histological techniques. Eur. J. Plant Pathol. 118:259-274.
11. Happstadius, I., Ljungberg, A., Kristiansson, B., and Dixelius, C. 2003. Identification of *Brassica oleracea* germplasm with improved resistance to Verticillium wilt. Plant Breed. 122:30-34.
12. Hasan, M., Seyis, F., Badani, A. G., Pons-Kuhnemann, J., Lühs, W., Friedt, W., and Snowdon, R. J. 2006. Surveying genetic diversity in the *Brassica napus* gene pool using SSR markers. Genet. Res. Crop Evol. 53:793-802.
13. Johansson, A., Staal, J., and Dixelius, C. 2006. Early responses in the *Arabidopsis-Verticillium longisporum* pathosystem are dependent on NDR1, JA- and ET-associated signals via cytosolic NPR1 and RFO1. Mol. Plant-Microbe Interact. 19:958-969.
14. Karapapa, V. K., Bainbridge, B. W., and Heale, J. B. 1997. Morphological and molecular characterization of *Verticillium longisporum* comb. nov., pathogenic to oilseed rape. Mycol. Res. 101:1281-1294.
15. Koike, S. T., Subbarao, K. V., Davis, R. M., Gordon, T. R., and Hubbard, J. C. 1994. Verticillium wilt of cauliflower in California. Plant Dis. 78:1116-1121.
16. Kosambi, D. D. 1944. The estimation of map distances from recombination values. Ann. Eugen. 12:172-175.
17. Landry, B. S., Hubert, N., Etoh, T., Harrada, J. J., and Lincoln, S. E. 1991. A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genome 34:543-552.
18. Lombard, V., Baril, C. P., Dubreuil, P., Blouet, F., and Zhang, D. 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: Consequences for varietal registration. Crop Sci. 40:1417-1425.
19. 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.
20. Mithen, R. F., and Magrath, R. 1992. Glucosinolates and resistance to *Leptosphaeria maculans* in wild and cultivated *Brassica* species. Plant Breed. 108:60-68.
21. Parkin, I. A. P., Sharpe, A. G., Keith, D. J., and Lydiate, D. J. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome 38:1122-1131.
22. Parkin, I. A. P., Sigrun, M. G., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., and Lydiate, D. J. 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana.* Genetics 171:765-781.
23. Piquemal, J., Cinquin, E., Couton, F., Rondeau, F., Seignoret, E., Doucet, I., Perret, D., Villeger, M. J., Vincourt, P., and Blanchard, P. 2005. Construction of an oilseed rape (*Brassica napu*s L.) genetic map with SSR markers. Theor. Appl. Genet. 111:1514-1523.
24. Rygulla, W., Seyis, F., Eynck, C., Friedt, W., Lühs, W., Snowdon, R., and von Tiedemann, A. 2007. Generation of zero erucic acid resynthesised rapeseed lines with resistance to *Verticillium longisporum*. Plant Breed. 126:596-602.
25. Rygulla, W., Snowdon, R. J., Eynck, C., von Tiedemann, A., Lühs, W., and Friedt, W. 2007. Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridization. Phytopathology 97:1391-1396.
26. Snowdon, R. J., Lühs, W., and Friedt, W. 2006. Oilseed rape. Genome Mapping and Molecular Breeding, Vol. 2. Pages 55-114 in: Oilseeds. C. Kole, ed. Springer Verlag, Heidelberg, New York, Berlin, Toyko.
27. Steventon, L. A., Fahleson, J., Hu, Q., and Dixelius, C. 2002. Identification of the causal agent of Verticillium wilt of winter oilseed rape in Sweden, *V. longisporum*. Mycol. Res. 106:570-578.
28. Steventon, L. A., Happstadius, I., Okori, P., and Dixelius, C. 2002. Development of a rapid technique for the evaluation of the response of *Brassica napus* to Verticillium wilt. Plant Dis. 86:854-858.
29. Udall, J. A., Quijada, P. A., and Osborn, T. C. 2005. Detection of chromosomal rearrangements derived from homoeologous recombination in four mapping populations of *Brassica napus* L. Genetics 169:967-979.
30. Utz, H. F., and Melchinger, A. E. 1996. PLABQTL: A program for composite interval mapping of QTL. J. Quant. Trait Loci 2:1.
31. Van Ooijen, J. W., and Voorrips, R. E. 2001. JoinMap 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands.
32. Veronese, P., Narasimhan, M. L., Stevenson, R. L., Zhu, J. K., Weller, S. C., Subbarao, K. V., and Bressan, R. A. 2003. Identification of a locus controlling Verticillium disease symptom response in *Arabidopsis thaliana*. Plant J. 35:574-587.
33. Weber, S., Ünker, F., and Friedt, W. 2005. Production of *Brassica napus* (L.) doubled haploids by microspore in vitro colchicine treatment and determination of ploidy by flow cytometry. Plant Breed. 124:511-513.
34. Zeise, K., and von Tiedemann, A. 2002. Host specialization among vegetative compatibility groups of *Verticillium dahliae* in relation to *Verticillium longisporum*. J. Phytopathol. 150:112-119.
35. Zhou, L., Hu, Q., Johansson, A., and Dixelius, C. 2006 *Verticillium longisporum* and *V. dahliae*: Infection and disease in *Brassica napus*. Plant Pathol. 55:137-144.
36. Zhou, W. J., Tang, G. X., and Hagberg, P. 2002. Efficient production of doubled haploid plants by immediate colchicine treatment of isolated microspores in winter *Brassica napus*. Plant Growth Regul. 37:185-192.

5.2 Co-localisation of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*)

Badani A.G., R.J. Snowdon, R. Baetzel, F.D. Lipsa, B. Wittkop, R. Horn, A. De Haro, R. Font, W. Lühs & W. Friedt, 2006a. Genome 49: 1499-1509

Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*)

Ana Gloria Badani, Rod J. Snowdon, Benjamin Wittkop, Florin D. Lipsa, Roland Baetzel, Renate Horn, Antonio De Haro, Rafael Font, Wilfred Lühs, and Wolfgang Friedt

Abstract: Quantitative trait loci (QTLs) contributing to yellow seed colour and acid detergent fibre (ADF) were localized and compared in 3 mapping populations developed from 2 crosses (designated 'YE1' and 'YE2') between 2 distinct sources of true-breeding yellow-seeded oilseed rape (*Brassica napus*) and 2 different black-seeded genotypes. A clear correlation was observed between seed colour and ADF content in both crosses. In all 3 populations, a major QTL, with a large effect on both seed colour and ADF in multiple environments, was detected at the same position on chromosome N18. In YE1, a second minor QTL, with a small effect on seed colour but not on ADF content, was localized on chromosome N1. In YE2, no QTL was observed on N1; however, 2 minor seed-colour loci were localized to N15 and N5. A second major QTL for ADF was localized in YE1 on N13; in YE2, no other QTLs for ADF were detected. Combined QTL and segregation data for seed colour and ADF content in the different populations suggest that a partially dominant *B. napus* gene for seed colour on N18 contributes to a reduction in fibre content in different yellow-seeded *B. napus* genotypes. The other QTLs that were identified appear to represent different genes in the 2 yellow-seeded rapeseed sources, which, in each case, affect only fibre content or seed colour, respectively. Potential candidate genes and implications for marker-assisted breeding of oilseed rape with reduced seed dietary fibre content are discussed.

Key words: Brassica napus, oilseed rape, seed colour, acid detergent fibre, ADF, QTL mapping.

Résumé : Les QTL déterminant la couleur jaune des graines et la teneur en fibres (ADF ou résidu de détergent acide) ont été localisés et comparés chez trois populations de cartographie dérivées de deux croisements (appelés 'YE1' et 'YE2') entre deux sources pures différentes de colza (*Brassica napus*) à graines jaunes et deux génotypes à graines noires. Une corrélation nette a été observée entre la couleur des graines et la teneur en fibres dans les deux croisements. Chez les trois populations, un QTL majeur ayant un grand effet tant sur la couleur des graines que sur la teneur en fibre dans plusieurs environnements a été détecté au même endroit sur le chromosome N18. Au sein du croisement YE1, un QTL mineur ayant un effet limité sur la couleur des graines mais pas sur la teneur en fibre a été détecté sur le chromosome N1. Au sein du croisement YE2, aucun QTL n'a été détecté sur le chromosome N1, mais deux QTL mineurs pour la couleur des graines ont été observés sur N15 et N5. Un second QTL majeur pour la teneur en fibre a été localisé sur N13 dans le croisement YE1, alors qu'aucun autre QTL pour la teneur en fibre n'a été décelé chez YE2. Ensemble, les données QTL et la ségrégation pour la couleur des graines et de la teneur en fibre chez les différentes populations suggèrent l'existence chez le *B. napus* d'un gène partiellement dominant déterminant la couleur des graines et situé sur N18. Ce même gène contribuerait à une réduction de la teneur en fibre chez divers génotypes à graines jaunes du *B. napus*. Les autres QTL qui ont été identifiés semblent correspondre à différents gènes qui, chez les deux sources distinctes de couleur jaune des graines, affectent soit uniquement la teneur en fibre ou la couleur des graines, respectivement. Des gènes candidats potentiels et les implications pour la sélection assistée de colza oléagineux ayant une teneur réduite en fibres alimentaires sont discutés.

Received 15 April 2005. Accepted 9 July 2006. Published on the NRC Research Press Web site at http://genome.nrc.ca on 20 February 2007.

Corresponding Editor: J.P. Gustafson.

A.G. Badani, R.J. Snowdon,[1] B. Wittkop, F.D. Lipsa, R. Baetzel, W. Lühs, and W. Friedt. Department of Plant Breeding, Research Centre for BioSystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, D-35392 Giessen, Germany.
R. Horn. Institute of Biological Sciences, Division Plant Genetics, University of Rostock, Albert-Einstein-Str. 3, D-18057 Rostock, Germany.
A. De Haro and R. Font. Department of Agronomy and Plant Breeding, Institute of Sustainable Agriculture (CSIC), Alameda del Obispo s/n, E-14080 Cordoba, Spain.

[1] Corresponding author (e-mail: rod.snowdon@agrar.uni-giessen.de).

doi:10.1139/G06-091

Mots clés : Brassica napus, colza oléagineux, couleur des graines, résidu de détergent acide, ADF, cartographie QTL.

[Traduit par la Rédaction]

Introduction

The combination of genetic and environmental factors confers a continuous range of seed colour in *Brassica* species, varying from black to yellow. Brown and yellow seeds are of particular interest to breeders of oilseed rape (*Brassica napus* ssp. *napus*, 2*n* = 38, genome AACC) because of their association with a thinner seed coat, which results in reduced dietary fibre content. This considerably improves the feed quality of rapeseed meal after oil extraction (Shirzagedan and Röbellen 1985, Slominski et al. 1994, 1999). It is thought that light seed colour and low fibre content coincide because the biochemical pathways leading to lignin and pigment synthesis have common precursors, such as p-cumarate (Theander et al. 1977, Whetten et al. 1998). The reduction in testa thickness in yellow-seeded oilseed rape is also associated with increased seed oil and (or) protein content per dry weight (Xiao and Liu 1982; Piotrowska et al. 2003), because of the proportional increase in the contribution of the embryo to seed volume. Acid detergent fibre (ADF) is the fraction of ground seed that remains intact after digestion with acid detergent compounds. Because it comprises mainly cellulose and lignin, this fraction represents one of the major nondigestible components of seed meal.

Selection for the yellow-seeded phenotype is difficult because of strong environmental effects, particularly temperature (cf. Van Deynze et al. 1993) and, to date, effective genetic markers linked to the genes controlling the trait are not available for winter oilseed rape. The mutated gene loci that affect the synthesis and (or) accumulation of dark pigments (proanthocyanidins, or condensed tannins) in the seed coat of *Arabidopsis thaliana*, collectively named transparent testa (*tt*) loci (Shirley et al. 1995; Haughn and Chaudhury 2005), represent interesting candidate genes for the analogous character of oilseed rape. The seed-specific genes *TT1* (Sagasser et al. 2002) and *TT16* (Nesi et al. 2002) are involved in the development of the seed endothelium, in which proanthocyanidins accumulate. Transcription factors TT2 and TT8 also regulate gene expression in the seed-coat epidermis, where they are involved in the control of flavonoid biosynthetic genes, such as BANYULS (Nesi et al. 2000, 2001; Baudry et al. 2004).

Several other loci (*TT3*, *TT4*, *TT5*, *TT6*, and *TTG*) encode enzymes involved in the accumulation of purple anthocyanins in leaves and stems (see Winkel-Shirley 2001); *TTG* is also involved in trichome and root-hair development (Koornneef 1994). Numerous *A. thaliana* flavonoid biosynthetic enzymes have been identified and correlated with specific *TT* loci, such as chalcone synthase (CHS) with *TT4* (Feinbaum and Ausubel 1988), chalcone isomerase (CHI/CFI) with *TT5* (Shirley et al. 1992), dihydroflavonol 4-reductase (DFR) with *TT3* (Shirley et al. 1992), flavanone 3-hydroxylase (F3H) with *TT6* (Wisman et al. 1998), and flavanone 3′-hydroxylase (F3′H) with *TT7* (Schoenbohm et al. 2000). Although these *TT* genes are present in only a single copy in *Arabidopsis*, multiple copies of their orthologues have been detected in *B. napus* with bacterial artificial chromosome (BAC) library screening, using physical functional markers within 50 kb of the gene sequences (Lotz et al. 2003).

Different mechanisms have been proposed for the inheritance of yellow seed colour in crop brassicas. According to Shirzagedan (1986), Henderson and Pauls (1992), and Van Deynze et al. (1995), 3 genes are involved in the genetic control of testa colour in *B. napus*, and only lines that are homozygous for recessive alleles at all 3 loci will breed true for yellow seed colour. Heneen and Brismar (2001) suggested a combination of maternal and embryonal control of seed colour in the diploid species *B. alboglabra* (2*n* = 18), with loci on 3 different C-genome chromosomes; Lionneton et al. (2004) identified 2 Mendelian trait loci responsible for the inheritance of yellow seeds in *B. juncea* (2*n* = 36, AABB). In contrast, Somers et al. (2001) localized a single major gene, explaining 72% of the variance in seed colour, in a cross between a yellow-seeded and a black-seeded *B. napus*; they suggested that this gene interacts in a dominant manner with 2 additional epistatic loci, each of which explains around 10% of the variance in seed colour. This finding was supported by Liu et al. (2005*a*), whose analyses of segregation ratios from crosses between black-seeded and resynthesized yellow-seeded rape showed that a partially dominant gene for yellow seed colour exhibits epistatic effects on 2 dominant genes for black seed colour. However, neither the responsible genes nor their putative effect are known to date. Also not known is whether the same gene loci are involved in the expression of yellow seed colour in oilseed rape derived from different yellow-seeded material. In particular, no information has been gathered regarding associations between the genetic control of *B. napus* seed colour and the dietary fibre content of the seed meal.

The aim of this study was to localize and compare the quantitative trait loci (QTLs) controlling seed colour and ADF in mapping populations from 2 distinct sources of yellow-seeded rapeseed crossed with different black-seeded winter oilseed rape parents, respectively. The results are discussed with respect to the putative effects of the responsible gene loci, potential candidate genes, and the implications for the breeding of yellow-seeded rapeseed with reduced dietary fibre content.

Materials and methods

Plant material

Genetic mapping and segregation analyses were performed in 3 segregating populations derived from 2 different yellow-seeded rapeseed sources. The first source was a *B. napus* winter-type inbred line (‘25629-3’) with yellow seed colour introgressed from a Chinese *B. rapa* ssp. *oleifera* accession. The other yellow-seeded line (‘1012/98’) originates from a resynthesized *B. napus* line created at the University of Göttingen, Germany, by Gland (1982).

The first population (YE1-DH) comprised 118 doubled haploid (DH) lines produced by microspore cultivation from the F_1 of a cross between the yellow-seeded line '25629-3' with double-low (00, canola) quality and the black-seeded DH line 'K26-96,' possessing high erucic acid and high glucosinolate content (++). A second segregating population (YE2-F_2), comprised of 1210 F_2 individuals, was derived from a cross between an inbred line ('Express 617') of the black-seeded German winter oilseed rape cultivar 'Express' and the true-breeding, yellow-seeded line '1012/98,' both with 00-seed quality. The same cross was also used to generate a population of 458 DH lines (YE2-DH).

The YE1-DH population was grown in 2001 in the greenhouse and over the 2001/2002 growing season in a field trial at 1 location with 2 replicates. The YE2-F_2 population was evaluated in the field in the same year; YE2-DH was grown at 2 locations in 2002/2003 and 2003/2004, respectively. Seeds were harvested from self-pollinated plants for seed colour and fibre analysis.

Measurement of seed colour, ADF, and seed quality parameters

Seed colour was assessed with the digital-optical measurement of seeds from F_3 plants or DH lines. The quantitative seed colour measurements were averaged over replicates within each population and environment. To get an initial impression of the correlation between seed-coat fibre and seed colour in YE1-DH and YE2-F_2, a visible near-infrared spectroscopy (NIRS) calibration for ADF content was developed, based on a calibration for *B. juncea*, *B. carinata*, and *B. napus* (Font et al. 2003). Samples used for the new calibration were selected from the full set of DH lines from YE1 and F_3 seeds from the YE2 cross, on the basis of their spectral features, after principal component analysis. On the basis of the Mahalanobis distance (H) of each spectrum from the mean spectrum of the population, 98 samples were selected as most variable. These samples were analysed for ADF content, using the method described by Goering and Van Soest (1970). New NIRS calibrations were then developed for ADF by regressing reference values against spectra, using modified partial least squares regression (Font et al. 2005). Finally, NIRS estimations of ADF were obtained for 80 YE1-DH lines and F_3 seeds from 170 YE2-F_2 individuals where sufficient seeds were available for spectroscopic measurement. The NIRS spectra were also used with standard calibrations to estimate oil and protein content in the investigated lines. Statistical analysis of the phenotype data was performed using the SPSS 6.0 software package (SPSS Inc., 1998).

Genetic mapping and construction of consensus maps

Genetic maps were generated using 105 randomly chosen DH lines from YE1-DH, 179 F_2 individuals from YE2-F_2, and 166 DH lines from YE2-DH. Maps were generated using 28 amplified fragment length polymorphism (AFLP) primer combinations, along with a set of simple sequence repeat (SSR) markers from public and commercial sources. AFLP analyses were performed using an *Eco*RI-*Mse*I-based AFLP kit (Invitrogen, Carlsbad, Calif.), in accordance with the manufacturer's instructions. *Eco*RI primers were fluorescently labelled with IRD700 or IRD800, and labelled AFLP bands were detected with electrophoresis, using LI-COR 4200 DNA analysis systems. For SSR analyses, the M13-tailing procedure described by Berg and Olaisen (1994) was used. 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 amplifications, the PCR fragments are subsequently amplified by the labelled universal primer.

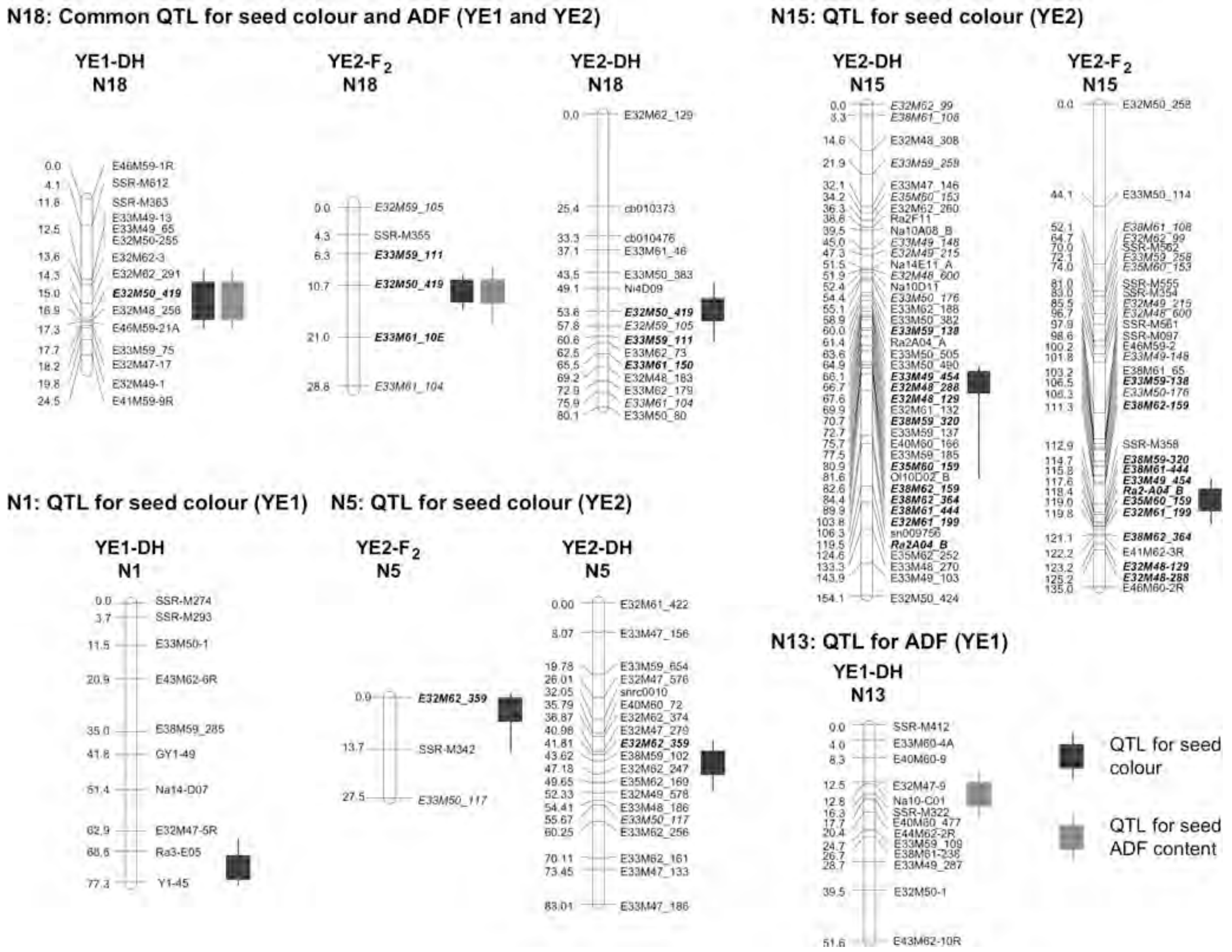
During marker scoring, the parents of the 'Mansholts Hamburger Raps' × 'Samourai' map, from the University of Göttingen (Ecke et al. 1995; Uzunova et al. 1995), were included on all gels to assist in the subsequent map alignment and the identification of linkage groups. Allelic data and raw-marker images from the mapping of the Göttingen population with AFLP markers were kindly provided by Dr. Wolfgang Ecke, University of Göttingen. Genetic linkage maps were constructed with JoinMap 3.0 linkage software (Kyazma, Netherlands), using a logarithm of odds (LOD) linkage threshold of 7.0 and a maximum likelihood distance of 40 cM. Kosambi's mapping function was used to compute the distances, in centiMorgans, from the recombination fractions. Linkage group designations were confirmed by comparing the map positions of publicly available SSR markers (Lowe et al. 2004; Piquemal et al. 2005). QTL analysis was performed with the PLABQTL, v. 1.2, software package (H.F. Utz, University of Hohenheim, Germany; http://www.uni-hohenheim.de/~ipspwww/soft.html), using composite interval mapping. Only QTLs showing significant effects ($\theta = 0.02$) with the multiple regression of the *final simultaneous fit* calculation were considered.

Results

A total of 144 AFLP and 49 SSR markers, covering 923.3 cM, were localized in the genetic map for YE1-DH, whereas 212 AFLP and 51 SSR markers were mapped over 1186.3 cM in YE2-F_2. The YE2-DH map contained 262 AFLP and 85 SSR markers, over a total map length of 1720.8 cM. The linkage groups were designated using the standard N1–N18 nomenclature for *B. napus*. All 19 *B. napus* chromosomes could be aligned and identified in the 3 maps, either directly using mapped SSR markers or markers common to the 3 maps, or indirectly on the basis of reference markers from 3 other maps (Piquemal et al. 2005; P. Basunanda 2006, personal communication; and W. Ecke 2004, personal communication). In the YE2-F_2 and YE2-DH maps, chromosome N18 grouped into 2 separate linkage groups that could not be joined, even after reducing the linkage threshold and increasing the maximum likelihood distance. Because both groups contained a number of markers localized on N18 in other maps, we designated these 2 groups N18a and N18b, respectively.

Figure 1 and Table 1 show details of all significant QTLs for seed colour and ADF content that were detected in different environments in the respective populations. A major QTL for yellow seed colour colocalized with a major locus for ADF content at the same position on chromosome N18 in all 3 populations investigated, and the same AFLP marker

Fig. 1. Comparison of quantitative trait loci (QTLs) for seed colour and content of acid detergent fibre (ADF) in 2 distinct crosses of different yellow- and black-seeded oilseed rape genotypes. Seed-colour QTLs were localized in a doubled-haploid (DH) population derived from the ‘25929’ × ‘K26–96’ cross (YE1-DH), and in an F_2 and a DH population derived from the ‘1012/98’ × ‘Express 617’ cross (YE2-F_2 and YE2-DH, respectively). QTL analysis for ADF content was performed only in YE1-DH and YE2-F_2. Markers shared by the corresponding linkage groups from the different mapping populations are shown in italics; markers near QTL regions that are shared by the different populations are shown in bold.



was localized within the QTL in all 3 cases. The positive alleles for yellow seed colour and low ADF content at this locus were always derived from the respective yellow-seeded parent. In YE2-F_2 and YE2-DH, a second QTL for seed colour was localized on chromosome N15, and a third QTL, with a small effect on seed colour, was found on chromosome N5. In YE1-DH, no QTLs were observed on chromosomes N15 or N5; however, a second significant seed colour locus was localized to chromosome N1. In YE1-DH, a second major QTL for ADF was localized on chromosome N13, whereas in YE2-F_2, no other QTLs for ADF could be detected. In YE1-DH, the QTL on chromosome N18 was the major contributing factor to seed colour, with a partial R^2 of 53.8%; however, its contribution to ADF content (partial R^2 = 24.0%) was matched by the second QTL for ADF on chromosome N13 (partial R^2 = 20.7). In YE2-DH, the QTL on chromosome N18 also had a very strong effect on seed colour (partial R^2 = 52.7%), and in YE2-F_2, this locus had a very strong effect on both seed colour (partial R^2 = 51.6%) and ADF content (partial R^2 = 47.5%). Other minor QTLs were detected that showed significant but low LOD values; however, these QTLs were not consistently detected in different environments, nor were they detected in the F_2 and DH populations of the YE2 cross. After multiple regression analysis, these minor QTLs were also found to have insignificant effects on seed colour and ADF content; hence, they were excluded from the further interpretation of the results.

Histograms illustrating segregation for seed colour in the 3 populations (Fig. 2*a*) show 2 peaks, 1 for genotypes with black and and 1 for genotypes with nonblack (brown or yellow) seeds. This corresponds to results in crosses between other yellow- and black-seeded *B. napus* genotypes (Somers et al. 2001; Liu et al. 2005*a*). Because no clear discrimination between yellow- and brown-seeded phenotype classes was possible, we followed the example of those authors in grouping yellow and brown seeds together, and compared the segregation ratios of black and nonblack phenotypes. A segmentation procedure, based on maximization of the corresponding *t* value, was used to identify the respective cutoff points separating black and nonblack phenotype classes in the different distributions. The same procedure was used to divide the respective segregations for ADF content (Fig. 2*b*) in YE1-DH and YE2-F_2 into groups representing the 2 peaks (“high ADF” and “reduced ADF,” respectively) observed for each of these distributions. Chi-square analyses were used to test the fit of the seed colour and ADF data from each population with putative segregation patterns expected, based on the respective QTLs in the 2 crosses.

Table 2 shows putative genetic explanations for the observed segregations between nonblack- and black-seeded genotypes in the respective populations, based on comparable findings by Liu et al. (2005*a*), who obtained similar results in 2 crosses of a Chinese yellow-seeded line and 2 different black-seeded parents. The segregation pattern for seed colour in YE1-DH was found to fit most closely to a 3:1 ratio. In this cross, 1 major QTL and 1 minor QTL were found to have a significant effect on seed colour. Hence, the observed segregation ratio can be explained by inheritance from a major, partially dominant gene for yellow seed colour (corresponding to the Y gene hypothesized by Liu et al. 2005*a*), with a second epistatic gene causing black seed colour only when the first locus is recessive. In YE2-DH, the segregation for seed colour fit closely to a 5:3 ratio. Based on the detection in YE2-DH of 1 major QTL with a large effect on seed colour, along with 2 minor QTLs, this can be explained if the partially dominant yellow-seed Y gene is influenced, in this case, by 2 epistatic genes (which we call B2 and B3) that cause black seed colour when 1 or both are dominantly expressed and Y is homozygous recessive. This suggested model also fits the 49:15 segregation ratio for seed colour observed in YE2-F_2 (see Table 2). Because different minor QTLs were detected in the 2 crosses (YE1 and YE2), we concluded that B1, B2, and B3 are different genes, or at least different homoeologous copies of the same gene or genes.

As expected from the colocalization of the major QTLs, ADF content was highly correlated with seed colour (YE1-DH, r = 0.88; YE2-F_2, r = 0.74). In both crosses, 2 clearly distinguishable classes were confirmed by the segmentation procedure mentioned above and, as was the case for seed colour, the segregation of “high ADF” vs. “reduced ADF” in the different populations was tested against the putative segregation patterns suggested by the respective QTL results. The segregation for reduced ADF content in YE1-DH was found to correspond closely (χ^2 = 0.129, p = 0.79, n = 80) to a 3:1 ratio. This is consistent with a diallelic inheritance, assuming that the 2 major QTLs for ADF content represent 2 major genes with similar effects on the trait. In YE2-F_2, the observed segregations suggests a more complex inheritance than is implied by the single detected QTL. This could indicate that loci not detected by the QTL analysis are involved in seed fibre accumulation in this cross. Nevertheless, the same significant single major locus with a strong phenotypic effect on ADF content could be localized on chromosome N18 in both crosses.

Discussion

Previous investigations of seed colour using QTL analyses in Canadian spring-type *B. napus* (Somers et al. 2001), and segregation studies of a yellow-seeded DH rapeseed line of Chinese origin (Liu et al. 2005*a*), indicate, in each case, that a major dominant gene was responsible for the majority of the genetic variation in seed colour observed in these different yellow-seeded *B. napus* materials. After combining QTL information with segregation data for 2 independent crosses derived from 2 other distinct sources of yellow seed colour, we detected a major QTL that appears to contribute significantly to both seed colour and fibre content in both of these new sets of material, and which presumably corresponds to the major gene described in the previous studies. We tested the markers closely linked to the common major QTL for seed colour in a third, independent cross, and the verification of the markers confirmed that the same major locus has a major effect on seed colour in all 3 yellow-seeded sources. In one of the yellow-seeded genotypes we used for our study, however, not 1 but 2 major QTLs were identified, each of which had a more or less equivalent effect on the content of dietary fibre (ADF) in the seeds. Furthermore, different minor loci with less significant effects on seed colour were also found in the different materials. This suggests

Fig. 2. (*a*) Histograms showing segregation for seed colour in seeds from YE1-DH, YE2-F$_2$, and YE2-DH, 3 segregating populations from 2 distinct crosses of different yellow- and black-seeded oilseed rape genotypes. (*b*) Segregation for seed ADF content in the YE1-DH and YE2-F$_2$ populations. The cut-off points used to count observed frequencies in the respective phenotype classes (numbers on either side of dashed lines) were determined using a segmentation procedure based on maximization of the corresponding *t* value. DM, dry mass.

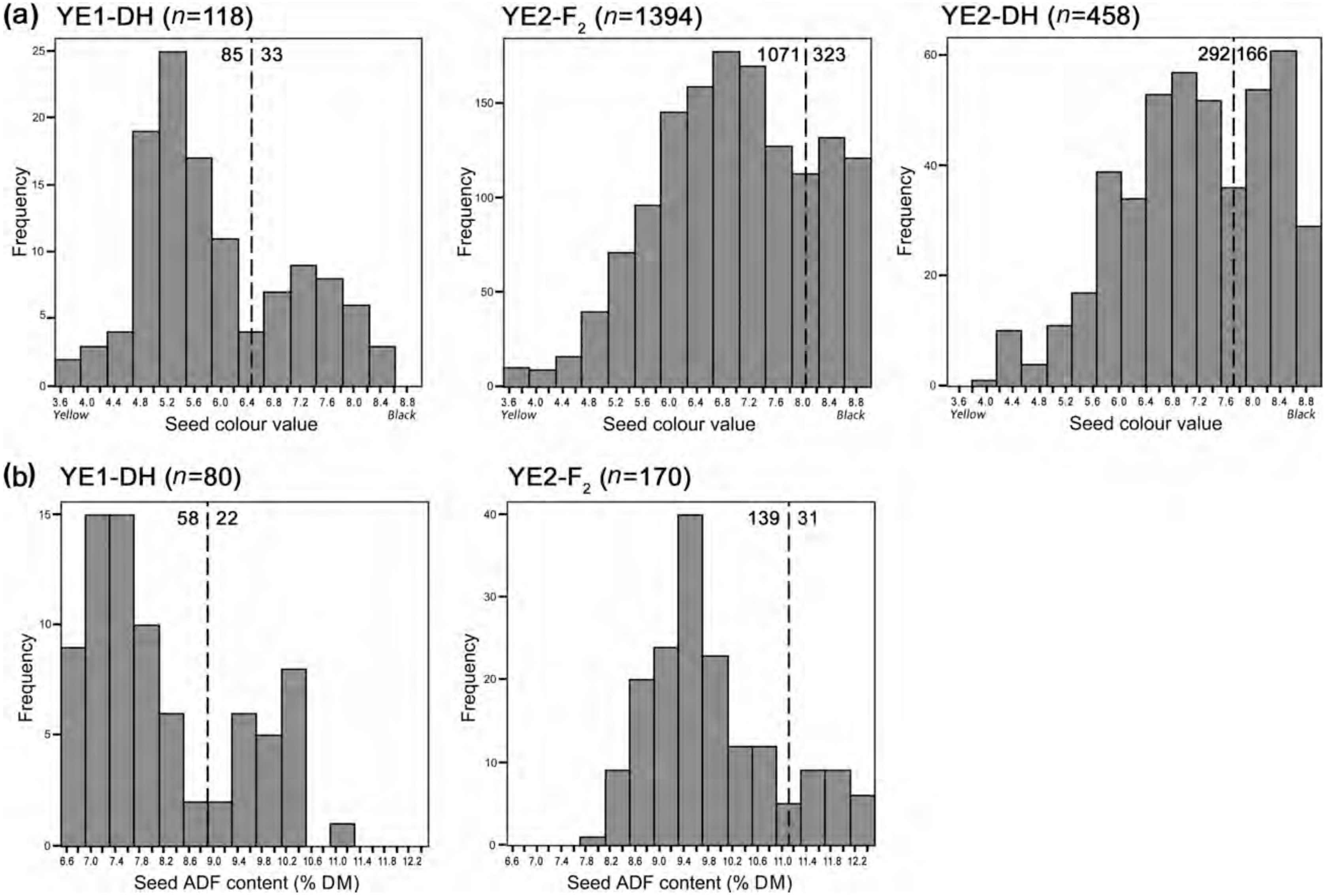


Table 1. Quantitative trait loci (QTLs) for seed colour and acid detergent fibre (ADF) content in F_2 and (or) doubled-haploid (DH) populations of the *Brassica napus* YE1 (between the yellow-seeded parent '25929' and the black-seeded parent 'K26-96') and YE2 (between the yellow-seeded parent '1012/98' and the black-seeded parent 'Express 617') crosses. For ADF, only a subset of the YE1-DH and YE2-F_2 populations with sufficient seed for near-infrared spectroscopy (NIRS) analysis were measured.

Trait	Population	Chromosome	LOD	R^2 (%)	Partial R^2 (%)	Origin of positive allele
Seed colour	YE1-DH (*n* = 105)	N18	5.53	21.4	53.8	'25629–3'
		N1	5.03	19.8	4.8	'25629–3'
	YE2-F_2 (*n* = 179)	N18	3.84	9.4	51.6	'1012/98'
		N15	5.98	14.3	13.8	'1012/98'
		N5	4.06	10.0	7.1	'1012/98'
	YE2-DH (*n* = 166)	N18	12.10	30.3	52.7	'1012/98'
		N15	10.40	26.6	19.4	'1012/98'
		N5	6.70	18.1	5.9	'1012/98'
ADF content	YE1-DH (*n* = 80)	N18	11.90	42.5	24.0	'25629-3'
		N13	9.72	36.4	20.7	'25629-3'
	YE2-F_2 (*n* = 170)	N18	8.06	19.6	47.5	'1012/98'

Note: LOD, logarithm of odds.

that different yellow-seeded *B. napus* genotypes have a different genetic background, at least at some of the loci responsible for seed colour and ADF content.

Although a large number of genes give rise to *transparent testa* mutations in *Arabidopsis*, it appears that, in *B. napus*, the same major gene locus plays a central role in different yellow-seeded sources. The presence of 1 or more dominant alleles at this locus always results in a nonblack seed phenotype; furthermore, this locus also contributes to a reduction in dietary fibre. This accounts, at least in part, for the strong correlation between seed colour and ADF content that we observed. Brown and yellow seeds always exhibit a reduction in seed testa thickness (Neubert et al. 2003); hence, it appears that this major gene in some way affects the development of the seed coat rather than simply reducing the expression of condensed tannins. This finding suggests that, in some cases, true-breeding brown- or yellow-seeded rapeseed with reduced dietary fibre can be bred by selection with respect to this single dominant gene, provided no segregation occurs in the accompanying epistatic loci. This corresponds to the results presented by Liu et al. (2005*b*), who observed segregation only at the partially dominant yellow-seed locus in crosses of the same yellow-seeded line investigated by Liu et al. (2005*a*) and 2 different black-seeded parents. In one of our crosses, however, another major gene contributes significantly to the reduction of ADF content. For the development of true-breeding yellow-seeded varieties with reduced fibre content, it would be beneficial to identify the genes underlying the responsible QTLs. In the meantime, the markers identified in this study closely linked to seed colour and ADF content should be valuable tools in breeding for reduced dietary fibre in brown- or yellow-seeded oilseed rape.

The seed-specific *TT1* gene (At1g34790) encodes a *WIP* zinc-finger-containing protein domain involved in the development of the seed endothelium, the tissue layer in which tannin pigments accumulate (Sagasser et al. 2002). The *Arabidopsis* B-Sister MADS domain protein encoded by *TT16* is required for normal endothelium development, as well as for seed pigmentation in the endothelium body (Nesi et al. 2002). Because of their association with the endothelium, these 2 seed-specific regulatory factors represent potential candidates for the putative major gene controlling ADF fibre accumulation. The QTLs and segregation data for ADF suggest that inheritance of this trait is different in different *B. napus* crosses, but that the same partially dominant gene plays a major role in the reduction of ADF content in different yellow-seeded materials. Genetic mapping of the respective homoeologous copies of *TT1* and *TT16* in *B. napus* should help establish whether they are associated with the major QTLs for seed colour and ADF.

In terms of seed colour, the major consequence of a reduction in testa thickness is a reduction in spatial accumulation capacity for proanthocyanidins (condensed tannins) in the endothelium cells. Assuming that one of the above genes reduces the thickness of the endothelium cell layer, a fluctuation in proanthocyanidin synthesis would likely result in a noticeable effect on seed colour. The epistatic loci detected in this study, and by Somers et al. (2001) and Liu et al. (2005*a*), might, therefore, be independent genes involved in the biosynthesis of condensed tannins. Such genes probably act additively during the accumulation of pigments in the testa, with the quantity (and hence the degree of pigmentation) being regulated by the testa (endothelium) thickness under control of the ADF major gene(s). A strong genotype–environment influence on the expression of pigment-synthesis genes might contribute to the difficulty in differentiating yellow and brown colour classes. This effect is not as obvious in black-seeded genotypes, probably because a saturation of condensed tannins in the thicker testa masks smaller variations in pigment quantity.

The most promising candidates to fit the genetic model presented here are the *A. thaliana TT* genes that code for key enzymes in the flavonoid biosynthesis pathway (for details, see http://www.genome.jp/kegg/pathway/ath/ath00941.html). Of particular interest are genes involved in the latter stages of proanthocyanadine biosynthesis. As described by Baudry et al. (2004), the expression of the seed-coat-specific

Table 2. Comparison of observed phenotype ratios for black and nonblack seed colour in segregating populations from the YE1 cross, with 1 major quantitative trait locus (QTL) and 1 minor (QTL) for seed colour, and from the YE2 cross, with 2 major QTLs and 1 minor QTL.

(a) YE1: One major, one minor QTL for seed colour				
Parents	‘25629–3’		‘K26–96’	
Seed colour	Yellow		Black	
Genotype	Y1Y1 b1b1		y1y1 B1B1	
F_1		Y1y1 B1b1: yellow-brown		
DH genotypes	Y1Y1 b1b1	Y1Y1 B1B1	y1y1 B1B1	y1y1 b1b1
Seed colour	Yellow	Yellow/brown	Black	Yellow/brown
Ratio	1	1	1	1
Expected (exp) segregation: 88 non-black: 30 black (3:1 segregation, n = 118)				
Observed (obs) segregation: 85 non-black: 33 black				
Test result (exp. vs. obs.): $\chi^2 = 0.402$, $p = 0.526$				

(b) YE2: Two major, one minor QTLs for seed colour								
Parents	‘1012/98’		‘Express 617’					
Seed colour	Yellow		Black					
Genotype	Y1Y1 b2b2 b3b3		y1y1 B2B2 B3B3					
F_1		Y1y1Y2y2Bb: brown						
YE2-DH genotypes	Y1Y1 B2B2 B3B3	Y1Y1 B2B2 b3b3	Y1Y1 b2b2 B3B3	Y1Y1 b2b2 b3b3	y1y1 B2B2 B3B3	y1y1B2B2b3b3	y1y1b2b2B3B3	y1y1b2b2 b3b3
Seed colour	Yellow/brown	Yellow/brown	Yellow/brown	Yellow	Black	Black	Black	Yellow/brown
Ratio	1	1	1	1	1	1	1	1
Expected (exp.) segregation: 286 non-black: 172 black (5:3 segregation, n = 458)								
Observed (obs.) segregation: 292 non-black: 166 black								
Test result (exp. vs. obs.): $\chi^2 = 0.589$, $p = 0.443$								
YE2-F2 genotypes	Y1–B2–B3–	Y1–B2–b3b3	Y1–b2b2 B3–	Y1–b2b2 b3b3	y1y1B2–B3–	y1y1B2–b3b3	y1y1b2b2 B3–	y1y1b2b2 b3b3
Seed colour	Yellow-brown	Yellow-brown	Yellow-brown	Yellow	Black	Black	Black	Yellow-brown
Ratio	27	9	9	3	9	3	3	1
Expected (exp.) segregation: 1067 non-black: 327 black (49:15 segregation, n = 1394)								
Observed (obs.) segregation: 1071 non-black: 323 black								
Test result (exp. vs. obs.): $\chi^2 = 0.064$, $p = 0.800$								

Note: Putative genotype classes and the resulting expected phenotypic segregations in each case are based on QTLs observed in the respective crosses. In the suggested model, Y1 is a putative partially dominant gene for yellow seed colour common to both crosses, whereas B1, B2, and B3 are the different putative epistatic genes that give black seed colour when Y1 is not dominantly homozygous. DH, doubled haploid.

BANYULS gene (GenBank accession No. At1g61720), which plays a central role in the accumulation of proanthodyanidins in the seed coat, is regulated by *TTG1* (acc. No. At5g24520), *TT2* (acc. No. At5g35550), and *TT8* (acc. No. At4g09820), which makes all of these genes potential candidates to have a significant effect on seed colour. However, the flavanone 3-hydroxylase enzymes, encoded by *TT6* (acc. No. At3g51240), and flavonoid 3′-monooxygenase, encoded by *TT7* (acc. No. At5g07990), act in tandem on 2 parallel branches at a central position within the flavonoid pathway. Together, these 2 enzymes are responsible for the synthesis of dihydroquercetin. This, in turn, is a dihydroflavanol, the class of substances believed to be converted by dihydroflavonol 4-reductase (*TT3*) into flavan-3,4,-diols (leucoanthocyanidins). Because the latter are the final precursors of both anthocyanins and proanthocyanidins (condensed tannins), the enzymes involved in their synthesis are strong candidates to have an effect on seed pigment biosynthesis. Nonexpression of flavanone 3-hydroxylase, flavonoid 3′-monooxygenase, or dihydroflavonol 4-reductase would severely impair condensed tannin production; however, mutants with reduced expression of one or more of these genes under seed-specific control could result in pigmentation reduction in a quantitative manner. Furthermore, the activity of flavonoid 3′-monooxygenase is dependent on cytochrome P450 (Schoenbohm et al. 2000) and, hence, affected by UV irradiation, which presumably accounts, at least in part, for the environmental variation in the accumulation of testa pigments in *B. napus*. Seed-specific regulation of the expression of one or more of these genes by a seed- or tissue-specific transcription factor could potentially regulate testa pigmentation without affecting anthocyanin pigmentation in the stems and leaves. In *B. carinata* dihydroflavonol 4-reductase expression was found to be significantly lower in yellow-seeded genotypes than in brown-seeded genotypes (Marles et al. 2003). Other genes directly involved in flavonoid biosynthesis have been shown to have at least 4 (chalcone isomerase, *TT5*) or, in the case of chalcone synthase (*TT4*), more than 10, homoeologous gene copies in *B. napus* (Lotz et al. 2003). In *A. thaliana*, only a single copy of these genes are present, and nonfunctional mutants display the *transparent testa* phenotype. In the ancestral polyploid *B. napus*, however, functional mutations at single loci of genes with multiple copies are likely compensated for by the continued expression at homoeologous loci. Interestingly, Van Deynze et al. (1995) located a restriction fragment length polymorphism marker from cDNA derived from the chalcone synthase gene at a position near a *B. napus* QTL, which contributed 11.5% of the observed variation in seed colour in their segregating DH mapping population.

Our hypothesis, that *transparent testa* genes are involved in the analogous yellow-seed trait in oilseed rape, is contradicted by the fact that all of the *TT* genes in *A. thaliana* are inherited in a recessive manner. Conversely, in the *B. napus* crosses investigated here, and in Somers et al. (2001), Liu et al. (2005*a*), and Liu et al. (2005*b*), inheritance can be explained, at least in part, by a partially dominant gene for yellow-seed colour. In contrast, studies of other yellow-seeded genotypes have indicated a triple-recessive inheritance of yellow-seededness (Shirzagedan 1986; Henderson and Pauls 1992), and QTL analysis revealed a considerably lower effect of these 3 loci on seed colour variation (Van Deynze et al. 1995). This apparent contradiction can be explained by a multiallelic control of the major gene locus, with (at least) 2 different *B. napus* alleles for yellow seed colour. It might be that 1 allelic form acts in a recessive manner toward the black-seeded allele (as it does in *A. thaliana*) and that a different allele acts in a dominant manner over the black-seeded allele (as described here). However, it is possible that the different situations result from independent mutations in different genes that each influence fibre content and endothelium thickness.

The reduction in testa thickness in light-seed genotypes increases the relative contribution of the cotyledon to the overall dry weight of the seed and, hence, the relative amount of oil and (or) protein. Therefore, yellow-seeded genotypes should normally show an increase in seed oil and (or) protein content. In our study, seed colour was correlated with protein content in YE1-DH (r = –0.378) but not in YE2-F_2 (r = –0.0037). Conversely, a correlation was observed between seed colour and oil content in YE2-F_2 (r = 0.340) but not in YE1-DH (r = –0.004). When protein + oil content was compared with ADF content, using partial correlation without consideration of the seed colour, a significant negative correlation was observed in both crosses (r = –0.55 in YE1, r = –0.56 in YE2). It should be noted, however, that the parental lines used for the crosses investigated here were selected on the basis of segregation for seed colour, independent of their oil or protein content. After introducing alleles for reduced fibre content into elite winter rapeseed breeding material, it is not surprising that oil and (or) protein levels are higher. Marker-assisted breeding, based on the results of this study, should improve the progress of true-breeding light-seeded oilseed rape varieties with improved seed meal nutritional quality.

Acknowledgements

This work was partially funded by the German Federal Ministry of Education and Research (BMBF grant No. 0312282), as part of the consortium GABI-GARS (“Genome Analysis in Rapeseed”). The authors thank Daniela Bernhardt for excellent technical assistance, and Jutta Ahlemeyer and Joern Pons-Kuehnemann for help with the statistical analyses. We also thank Deutsche Saatveredelung GmbH, KWS Saat AG, Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, and SW Seed GmbH for assistance with the mapping populations, and Saaten-Union Resistenzlabor GmbH and KWS Saat AG for providing commercial SSR primers. We also thank Panjisakti Basunanda and Wolfgang Ecke for providing unpublished marker and map data.

References

Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. 2004. TT2, TT8 and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant J. **39**: 366–380. doi:10.1111/j.1365-313X.2004.02138.x. PMID:15255866.

Berg, E.S., and Olaisen, B. 1994. Hybrid PCR sequencing: sequencing of PCR products using a universal primer. Biotechniques, **17**: 896–901. PMID:7840970.

Ecke, W., Uzunova, M., and Weissleder, K. 1995. Mapping the genome of rapeseed (*Brassica napus* L.) 2: Localization of genes controlling erucic acid synthesis and seed oil content. Theor. Appl. Genet. **91**: 972–977.

Feinbaum, R.L., and Ausubel, F.M. 1988. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. Mol. Cell. Biol. **8**: 1985–1992. PMID:3386631.

Font, R., Del Rio, M., Fernandez, J.M., and De Haro, A. 2003. Acid detergent fibre analysis in oilseed *Brassicas* by near-infrared spectroscopy. J. Agric. Food Chem. **51**: 2917–2922. doi:10.1021/jf021011+. PMID:12720371.

Font, R., Badani, A.G., Wittkop, B., del Río-Celestino, M., Lühs, W., Friedt, W., and de Haro-Bailón, A. 2005. Toward a global calibration for acid detergent fiber in rapeseed by visible and near-infrared spectroscopy. Plant Breed. **124**: 410–412. doi:10.1111/j.1439-0523.2005.01138.x.

Gland, A. 1982. Gehalt und Muster der Glucosinolate in den Samen von resynthesierten Rapsformen. Z. Pflanzenzüchtg. **88**: 242–254.

Goering, H.K., and Van Soest, P.J. 1970. Forage fiber analysis; apparatus, reagents, procedures, and some applications. USDA-ARS Agric. Handbook No. 379, US Govt. Print Office, Washington DC.

Haughn, G., and Chaudhury, A. 2005. Genetic analysis of seed coat development in Arabidopsis. Trends Plant Sci. **10**: 472–477. doi:10.1016/j.tplants.2005.08.005. PMID:16153880.

Henderson, C.A.P., and Pauls, K.P. 1992. The use of haploidy to develop plants that express several recessive traits using light-seeded canola (*Brassica napus*) as an example. Theor. Appl. Genet. **83**: 476–479.

Heneen, W.K., and Brismar, K. 2001. Maternal and embryonal control of seed colour by different *Brassica alboglabra* chromosomes. Plant Breed. **120**: 325–329. doi:10.1046/j.1439-0523.2001.00614.x.

Koornneef, M. 1994. Arabidopsis genetics. *In* Arabidopsis. *Edited by* E.M. Meyerowitz and C.R. Somerville. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. pp. 89–120.

Lionneton, E., Aubert, G., Ochatt, S., and Merah, O. 2004. Genetic analysis of agronomic and quality traits in mustard (*Brassica juncea*). Theor. Appl. Genet. **109**: 792–799. doi:10.1007/s00122-004-1682-0. PMID:15340689.

Liu, X.P., Tu, J.X., Chen, B.Y., and Fu, T.D. 2005*a*. Identification and inheritance of a partially dominant gene for yellow seed colour in *Brassica napus.* Plant Breed. **124**: 9–12. doi:10.1111/j.1439-0523.2004.01051.x.

Liu, Z.W., Fu, T.D., Tu, J.X., and Chen, B.Y. 2005*b*. Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). Theor. Appl. Genet. **110**: 303–310. PMID:15565377.

Lotz, T., Snowdon, R., Horn, R., Dewal, G., Weisshaar, B., Friedt, W., et al. 2003. Molecular analysis of *Arabidopsis thaliana tt*-genes in *Brassica napus. In* Proceedings of the 11th International Rapeseed Congress. 6–10 July 2003. Vol. 1. *Edited by* H. Sørensen. Dept. of Chemistry, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark. pp. 109–111.

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. doi:10.1007/s00122-003-1522-7. PMID:15067397.

Marles, M.A., Gruber, M.Y., Scoles, G.J., and Muir, A.D. 2003. Pigmentation in the developing seed coat and seedling leaves of *Brassica carinata* is controlled at the dihydroflavonol reductase locus. Phytochemistry, **62**: 663–672. doi:10.1016/S0031-9422(02)00488-0. PMID:12620317.

Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M., and Lepiniec, L. 2000. The TT8 gene encodes a basic helix–loop–helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. Plant Cell, **12**: 1863–1878. doi:10.1105/tpc.12.10.1863. PMID:11041882.

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L. 2001. The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. Plant Cell, **13**: 2099–2114. doi:10.1105/tpc.13.9.2099. PMID:11549766.

Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M., and Lepiniec, L. 2002. The *TRANSPARENT TESTA16* locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. Plant Cell, **14**: 2463–2479. doi:10.1105/tpc.004127. PMID:12368498.

Neubert, A.B., Lühs, W., Baetzel, R., and Friedt, W. 2003. Influence of seed colour on seed vigour in *Brassica napus. In* Proceedings of the 11th International Rapeseed Congress. 6–10 July 2003. Vol. 1. *Edited by* H. Sørensen. Dept. of Chemistry, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark. pp. 253–256.

Piotrowska, A., Krymanski, J., Bartkowiak-Broda, I., and Krotka, K. 2003. Characteristic of yellow-seeded lines of winter oilseed rape. *In* Proceedings of the 11th International Rapeseed Congress 6–10 July 2003. Vol. 1. *Edited by* H. Sørensen. Dept. of Chemistry, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark. pp. 247–249.

Piquemal, J., Cinquin, E., Couton, F., Rondeau, C., Seignoret, E., Doucet, I., et al. 2005. Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. Theor. Appl. Genet. **111**: 1514–1523. doi:10.1007/s00122-005-0080-6. PMID:16187118.

Sagasser, M., Lu, G.H., Hahlbrock, K., and Weisshaar, B. 2002. *A. thaliana* TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. Genes Dev. **16**: 138–149. doi:10.1101/gad.212702. PMID:11782451.

Schoenbohm, C., Martens, S., Eder, C., Forkmann, G., and Weisshaar, B. 2000. Identification of the *Arabidopsis thaliana* flavonoid 3′-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol. Chem. **381**: 749–753. doi:10.1515/BC.2000.095. PMID:11030432.

Shirley, B.W., Hanley, S., and Goodman, H.M. 1992. Effect of ionizing radiation on a plant genome: analysis of two *Arabidopsis transparent testa* mutations. Plant Cell, **4**: 333–347. doi:10.1105/tpc.4.3.333. PMID:1354004.

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M., and Goodman, H.M. 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J. **8**: 659–671. doi:10.1046/j.1365-313X.1995.08050659.x. PMID:8528278.

Shirzagedan, M. 1986. Inheritance of seed color in *Brassica napus.* Z. Pflanzenzüchtung, **96**: 140–146.

Shirzagedan, M., and Röbellen, G. 1985. Influence of seed colour and hull proportions on quality properties of seeds in *Brassica napus* L. Fette Seifen Anstrichm. **87**: 235–237.

Slominski, B.A., Campbell, L.D., and Guenter, W. 1994. Carbohydrates and dietary fibre components of yellow and brown-seeded canola. J. Agric. Food Chem. **42**: 704–707. doi:10.1021/jf00039a020.

Slominski, B.A., Simbaya, J., Campbell, L.D., Rakow, G., and Guenter, W. 1999. Nutritive value for broilers of meals derived from newly developed varieties of yellow-seeded canola. Anim.

Feed Sci. Technol. **78**: 249–262. doi:10.1016/S0377-8401(99)00003-6.

Somers, D.J., Rakow, G., Prabhu, V.K., and Friesen, K.R.D. 2001. Identification of a major gene and RAPD markers for yellow seed coat colour in *Brassica napus.* Genome, **44**: 1077–1082. doi:10.1139/gen-44-6-1077. PMID:11768211.

Theander, O., Aman, P., Miksche, G.E., and Yasuda, S. 1977. Carbohydrates, polyphenols, and lignin in seed hulls of different colors from turnip rapeseed. J. Agric. Food Chem. **25**: 270–273. doi:10.1021/jf60210a042.

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

Van Deynze, A.E., Beversdorf, W.D., and Pauls, K.P. 1993. Temperature effects on seed color in black- and yellow-seeded rapeseed. Can. J. Plant Sci. **73**: 383–387.

Van Deynze, A.E., Landry, B.S., and Pauls, K.P. 1995. The identification of restriction fragment length polymorphisms linked to seed colour genes in *Brassica napus.* Genome, **38**: 534–542.

Whetten, R.W., Mackay, J.J., and Sederoff, R. 1998. Recent advances in understanding lignin biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. **49**: 585–609. doi:10.1146/annurev.arplant.49.1.585. PMID:15012247.

Wisman, E., Hartmann, U., Sagasser, M., Baumann, E., Palme, K., Hahlbrock, K., et al. 1998. Knock-out mutants from *En-1* mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proc. Natl. Acad. Sci. U.S.A. **95**: 12432–12437. doi:10.1073/pnas.95.21.12432. PMID:9770503.

Winkel-Shirley, B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology. Plant Physiol. **126**: 485–493. doi:10.1104/pp.126.2.485. PMID:11402179.

Xiao, D.R., and Liu, H.L. 1982. Correlation analysis of seed colour and seed oil in *Brassica napus* L. Acta Agronomica Sinica, **8**: 24–27.

6 Global transcriptome analysis

This article presents a comprehensive analysis of gene expression during seed development in *B. napus* using serial analysis of gene expression (SAGE). This technique uses digital tag profiling of expressed sequences to quantify global gene expression, giving data similar to but potentially more powerful and accurate than microarray analysis. We developed methods and data mining tools to enable the use of EST-tagging for gene expression profiling in *B. napus*, despite the absence of a completed genome sequence. This is of particular interest in the context of novel, next-generation sequencing (NGS) technologies that enable cheap, ultrafast, massively-parallel sequencing of short nucleotide molecules. As NGS technologies move into a new generation, in the foreseeable future it is likely that digital tag expression profiling or ultradeep EST sequencing will supersede microarrays as the method of choice for global transcriptional profiling.

6.1 Global gene expression profiling via LongSAGE in a non-model plant species: A case study in seeds of *Brassica napus*

Obermeier C., B. Hosseini, W. Friedt & R.J. Snowdon, 2009. BMC Genomics, in press

Gene expression profiling via LongSAGE in a non-model plant species: A case study in seeds of *Brassica napus*

Christian Obermeier*, Bashir Hosseini, Wolfgang Friedt, Rod Snowdon

Justus Liebig University Giessen, Department of Plant Breeding,
Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding author

Submitted January 2009, accepted June 2009

Abstract

Background

Serial analysis of gene expression (LongSAGE) was applied for gene expression profiling in seeds of oilseed rape (*Brassica napus* ssp. *napus)*. The usefulness of this technique for detailed expression profiling in a non-model organism was demonstrated for the highly complex, neither fully sequenced nor annotated genome of *B. napus* by applying a tag-to-gene matching strategy based on *Brassica* ESTs and the annotated proteome of the closely related model crucifer *A. thaliana*.

Results

Transcripts from 3,094 genes were detected at two time-points of seed development, 23 days and 35 days after pollination (DAP). Differential expression showed a shift from gene expression involved in diverse developmental processes including cell proliferation and seed coat formation at 23 DAP to more focussed metabolic processes including storage protein accumulation and lipid deposition at 35 DAP. The most abundant transcripts at 23 DAP were coding for diverse protease inhibitor proteins and proteases, including cysteine proteases involved in seed coat formation and a number of lipid transfer proteins involved in embryo pattern formation. At 35 DAP, transcripts encoding napin, cruciferin and oleosin storage proteins were most abundant. Over both time-points, 18.6 % of the detected genes were matched by *Brassica* ESTs identified by LongSAGE tags in antisense orientation. This suggests a strong involvement of antisense transcript expression in regulatory processes during *B. napus* seed development.

Conclusions

This study underlines the potential of transcript tagging approaches for gene expression profiling in *Brassica* crop species via EST matching to annotated *A. thaliana* genes. Limits of tag detection for low-abundance transcripts can today be overcome by ultra-high throughput sequencing approaches, so that tag-based gene expression profiling may soon become the method of choice for global expression profiling in non-model species.

Background

Seed developmental processes of plants are complex and follow coordinated gene expression programs. Some genes are exclusively expressed during seed development [1]. Understanding the genetics of developmental processes and regulatory networks involved in embryogenesis and seed development is crucial for improvement of seed quality in crop plants. Gene expression during seed development has been studied intensively in the model plant *Arabidopsis thaliana* using mutagenesis [2, 3] and microarray analyses (e.g., [4, 3]). These studies revealed major gene expression changes during seed filling and desiccation, along with distinct expression patterns related to carbohydrate metabolism, lipid biosynthesis and storage protein accumulation. However, the regulatory mechanisms that ensure the proper execution of seed development in *A. thaliana* and other plants remain largely unknown. Furthermore, generalisation of findings regarding gene regulation from model to crop plants is difficult because most major crops have considerably more complex genomes.

Oilseed rape (*Brassica napus* ssp. *napus*), the closest major crop relative of *A. thaliana*, is the world's second most important oilseed crop. The high-value nutritional oil is also an excellent substrate for biodiesel production, whereas the protein-rich seed meal remaining after oil extraction is a valuable livestock feed. Oligonucleotide microarrays constructed for *A. thaliana* have been used in the past for expression profiling in *B. napus*, but have not provided optimal signal intensity and reproducibility [4, 5, 6, 7]. Recently the total number of ESTs from *Brassica* species deposited in public databases has risen dramatically to more than 800,000 entries with about 280,000 from seed developmental stages. 67,000 ESTs from seed developmental stages have been used to develop a *B. napus* cDNA microarray for analysis of seed gene expression patterns [7].

An alternative for accurate, quantitative global expression profiling is serial analysis of gene expression (SAGE; [8]), which in contrast to microarray hybridization allows the detection of new transcripts. SAGE is an expression profiling technique that simultaneously measures the levels of thousands of genes expressed in a given tissue. The method is based on the excision of short tags from poly A+ RNAs and end-to-end ligation of ditags to form high molecular weight concatemers. This allows cost-effective high-throughput cloning and sequencing of concatemers. Matching of tags to genomic sequences is a critical step in SAGE data analysis, and normally this requires the availability of high quality genome annotation data [9]. SAGE was first used for quantification of gene expression in human using 13-15 bp tags [8]. Modifications of the original SAGE protocols producing 21 bp tags (LongSAGE; [9]) and 26 bp tags (SuperSAGE; [10]) have been developed to enable more efficient and unambiguous tag-to-gene assignment in higher organisms with more complex transcriptomes. SAGE is commonly used in animal genomics, but has been increasingly used in plant species and tissues [11]. The intention of the present study was to adapt the LongSAGE technique for analysis of global gene expression in *B. napus* and other similarly complex polyploid plant genomes where the complete genome sequence and annotation are not yet available. A data processing pipeline was adapted by matching *B. napus* tags via *Brassica* ESTs to annotated *A. thaliana* gene loci, including detection of tags matching in sense and antisense orientation.

Results

B. napus seed LongSAGE libraries

Two LongSAGE libraries were produced from *B. napus* seeds harvested at 23 and 35 DAP, respectively. A total of 3,136 clones from the 23 DAP library yielded 34,984 ditags (22 tags/clone), while 3,168 sequence traces from the 35 DAP library yielded 15,919 ditags (10 tags/clone) (Table 1). The relatively low average number of tags per clone resulted from a high number of cloned empty vector plasmids (28.2 % and 60.3 % of all high-quality reads for the 23 DAP and 35 DAP libraries, respectively). The vector plasmid pZErO (Invitrogen) was designed to prevent cloning of empty vector by using a LacZa-ccdB fusion cDNA insert that should destroy bacteria containing only the vector. However, as described by other authors (e.g. [12]) we found that pZErO could efficiently re-ligate without containing inserts. Thus, PCR screening of plasmid libraries for clones containing an empty vector might considerably reduce total Sanger sequencing costs and increase total tag sample sizes in LongSAGE transcription profiling experiments.

Table 1 summarises the numbers of accepted and rejected tags in the two libraries. In total, 73,697 tags comprising 37,367 unique tags were extracted from both libraries. Valid tags from 23 DAP and 35 DAP were deposited in the Gene Expression Omnibus (GEO) public domain under accession number GSE14313. The total number of accepted tags from the 23 DAP library was 48,514 comprising 24,654 unique tags. The total number of accepted tags from the 35 DAP library was 25,183 tags comprising 12,713 unique tags. 67.0 % and 64.9 % of the accepted total tags were represented by five or less copies in the 23 DAP and 35 DAP libraries, respectively.

Matching *B. napus* tags to the *A. thaliana* proteome via *Brassica* ESTs

Table 2 summarises the results from the tag-to-gene maching analysis of the two libraries. Initially, all ESTs from a redundant database consisting of 834,732 *Brassica* EST entries (from the A and C genome brassicas *B. napus*, *B. oleracea*, *B. rapa*) were oriented by blastx alignment against the *A. thaliana* TAIR7 proteome (cut-off value 1e-6). Virtual tags were then extracted from all successfully oriented ESTs and assigned via these ESTs to the annotated *A. thaliana* proteins. A total of 3,563,603 virtual 21 bp tags were extracted from all positions of the ESTs. From the 32,395 different observed tags, 15,955 (49.3 %) were matched under the applied conditions to *Brassica* A or C genome ESTs (Table 2). Matching success was reduced about 30 % for singleton tags compared to tags with counts ≥ 2. For tags with counts ≥ 2 the matching success amounted to about 80 % (Table 2). This discrepancy might be due to the reduced numbers of ESTs in the database derived from transcripts expressed at low levels, compared to transcripts expressed at high levels within *B. napus* seeds. Broadly different estimates exist on the proportion of errors within tag sequences caused by experimental artefacts, ranging from 1.67 % to 15.6 % of LongSAGE tags [13, 14] with at least one erroneous base-pair introduced by sequencing. However, the singleton tag category might contain a higher proportion of tags with sequencing artefacts, also contributing to the different ratio of successfully matched tags in the singleton and ≥ 2 count categories. 12,728 tags from library 23 DAP were matched to 7,208 of 20,462 *Brassica* EST-matched *A. thaliana* proteins. 7,232 tags from library 35 DAP were matched to 4,875 *A. thaliana* proteins (Table 2).

Comparison of LongSAGE and Real-time RT-PCR analysis

A good correlation was observed between expression levels estimated by LongSAGE and Real-time RT-PCR for most of the selected genes (Table 3), although some minor differences in estimated fold-change were observed (e.g. for CCR1). This could arise from the more selective specificity of Real-time RT-PCR for transcripts derived from closely-related gene loci, since specific RT-PCR primers were designed from contigs of *B. napus* ESTs putatively derived from different *B. napus* gene loci. Furthermore, Real-time RT-PCR may be more sensitive in detecting low abundant transcripts than LongSAGE under the experimental conditions and sequencing depth applied in the present study.

Transcript diversity

Average tag frequencies for the two analysed seed libraries differed significantly (Table 1). Whereas tags with medium frequency (2-20) were more abundant in the 23 DAP library, the 35 DAP library was considerably more enriched for high copy-number tags with a frequency ≥100. Nine unique tags with ≥100 counts each represented 3.2 % of the total tag number at 35 DAP, whereas the 8 tags with ≥100 counts each at 23 DAP comprised only 2.1 % of the total tags. Accordingly, the relative diversity of tags was greater at 23 DAP than at 35 DAP. We detected 48 % (7,208 compared to 4,875) more expressed genes at 23 DAP than at 35 DAP (Table 2), which is consistent with the number of different tags observed in each of the two libraries.

To validate the LongSAGE method and estimate the diversity of transcripts, LongSAGE tags derived from selected tags with 2-32 counts in the 35 DAP library were used as primers for RT-PCR in combination with anchored oligo-dT primers. In 17 out of 17 cases a PCR product was amplified and sequenced that aligned in blastn analysis with *Brassica* or *A. thaliana* ESTs or cDNAs. A further validation was performed with 9 different LongSAGE tags that had 3 to 325 counts in the 35 DAP library and matched members of the *A. thaliana* 2S storage gene family or ESTs/cDNAs from the *Brassica* napin gene family. In all cases the PCR amplification products, ranging from 75 to 457 bp in length, were related to *Brassica* napin genes or other protease inhibitor/seed storage/lipid transfer protein (LTP) family protein genes (56 to 100 % identity). A high diversity of 3' termini of clones were obtained using the same napin tag pimer, with up to 14 % divergence, within 200-400 bp of the 3' untranslated region (UTR), from up to 8 clones per amplicon. Interestingly, diverse 3'-UTR sequences were also obtained using different napin tag primers, with up to 30 % divergence, within 200 bp of the 3'-UTR, from 16 clones. These observations suggest that a high number of distinct protease inhibitor/seed storage/lipid transfer protein (LTP) family protein genes or transcript variants are expressed in *B. napus* seeds. Also these data support the increasing recognition of the high complexity and variability of 3' UTR regions in genes from higher organisms [13, 15, 16].

Based on the cDNA library production using polyA+ trapped mRNA, LongSAGE tags are expected to match at the anchoring enzyme site closest to the 3' end of full-length ESTs expressed in sense orientation (canonical position). However, only 95.3 % of the matched tags in the sense EST dataset were matched in the last canonical position. There are numerous possible explanations for the high frequency of tags matching a non-canonical position. In particular, they could arise from (a) cloning artefacts caused by partial digestion with the anchoring enzyme *Nla*III during SAGE library preparation, (b) alternative transcript variants with varying length derived from one gene locus, or (c) EST cloning artefacts creating chimeric EST molecules derived from two or more different genes. Also, due to partial EST sequences present in the dataset, a non-canonical position might be falsely annotated to the canonical position. Except in the third case, non-canonical tags are still valid tags for accurate annotation. We attempted to reduce annotation artefacts (c) by removing putative chimeric ESTs from analysis (see section 'Genes matched in sense and antisense orientation' for details). This increased the ratio of tags matched to the canonical position of oriented *Brassica* ESTs from 95.3 % to 98.3 %. However, even after filtering out putative experimental artefacts using this approach, only 77 % of all 243,528 ESTs matched by tags in the canonical position are matched exlusively at this position. Instead, 23 % of these individual ESTs are additionally matched by tags at other non-canonical positions (up to the 9th before last position), suggesting that the usage of alternative transcripts (e.g. by alternative polyadenylation) is a common mechanism. In other words, a high diversity exits in the 3' UTR of many expressed transcripts, as has been described before in humans [13]. No information exists about transcription start points for polyA+ containing transcripts expressed in antisense orientation. Therefore different proportions of matched antisense tags might be expected at canonical and non-canonical positions. In contrast to the tags matched in sense orientation, only 65.5 % of tags were matched to canonical positions for the tags matched in antisense orientation.

Comparison of the most abundantly expressed tags and assigned genes

Tag-to-EST matches based on 15,955 tags were assigned to a total of 8,255 *A. thaliana* gene loci at both time points (Table 2). About 85 % of the tags (13,641) were assigned exclusively to a single *A. thaliana* gene locus. The other 15 % of the tags matched 2 to 45 loci, with some tags matching more than 10 genes exhibiting a polyA-like homopolymer composition. Due to the positioning of the anchoring enzyme restriction site close to the polyA tail in some transcripts, these contained stretches with up to 16 adenosine bases derived from phylogenetically unrelated genes. Other tags matching more than 2 genes were clearly derived from groups of phylogenetically related genes. For example, the tag CATGAACAGTTTCATCAACGA matched to six different members of the histone gene family (AT2G37470, AT3G53650, AT5G22880, AT5G59910, AT1G07790 and AT5G02570).

On the other hand, some *A. thaliana* loci were matched by up to 96 different *B. napus* tags via the *Brassica* ESTs. Moreover, some loci were not only matched by different tags, but these tags also matched at different positions within one particular matched EST molecule. These multiple matches underline the complex paleopolyploid structure of the *B. napus* genome in comparison to *A. thaliana* [17]. For example, the 4 *A. thaliana* gene loci of the 2S seed storage protein family were matched by 47 different tags, including 9 of the 30 most abundant tags, at 35 DAP. Due to this complexity, a three-step procedure was applied for calculating average values of the relative abundance of gene expression, based on the tag-to-gene matching results for each single *A. thaliana* gene locus. In a first step, the measured counts were evenly distributed to the matched gene loci if a tag matched more than one gene locus. In a second step, tag counts were added together if different tags matched the same gene locus. In a final step the summed tag counts for each gene locus were normalized to a total tag count of 1,000,000 for both libraries and the relative abundances were calculated.

Average cumulative counts and frequencies for the 20 most abundantly expressed genes at 23 and 35 DAP are shown in Table 4. The most abundant transcripts at 23 DAP, with 1.59 % of all matched tag counts, correspond to the *A. thaliana* senescence-associated cysteine-type protease SAG12 (AT5G45890), which was also among the 20 most highly-expressed genes at 35 DAP. Two protease inhibitor/seed storage/lipid transfer family proteins (AT5G38195, AT1G48750) were also among the 20 most highly expressed genes at both time-points. Other highly expressed transcripts at 23 DAP are a number of catalytic enzymes and other genes that are involved in diverse biological processes. At 35 DAP, transcripts related to the four closely-related *A. thaliana* 2S seed storage protein genes showed the highest counts, with a cumulative abundance of about 10 %. Other highly expressed transcripts at 35 DAP are related to other storage proteins (cruciferin, oleosin), protease inhibitor or lipid transfer proteins and genes involved in fatty acid biosynthesis. For many of the highly expressed genes several different tags were found (e.g. 44 for AT5G45890 at 23 DAP) matching to a large number of different ESTs (e.g. 5,452 for AT5G45890 at 23 DAP) that were aligned to a particular *A. thaliana* gene at <1e-6.

Differentially expressed genes

620 tags showed differential expression at $p \leq 0.05$ between 23 and 35 DAP. These represent only 1.7 % of the accepted unique tags, but 21.7 % of all normalized tag counts. 498 of 620 (80.3 %) differentially expressed unique tags were successfully matched to 490 *A. thaliana* genes (5.9 % of all 8,255 matched genes, Table 2), and these differentially expressed genes were matched by 18.2 % and 38.1 % of all matched tags in 23 DAP and 35 DAP libraries, respectively (see additional file 1: 490_differentially_expressed_genes.xls). This indicates that, in the 35 DAP library but not in the 23 DAP library, many of the differentially expressed transcripts are highly abundant. Of these 490 genes, 194 genes were upregulated at 23 DAP and 296 genes were upregulated at 35 DAP. The 20 most abundant differentially expressed genes at 23 DAP include 6 genes encoding proteases, 5 genes coding for protease inhibitor/ seed storage/lipid transfer protein (LTP) family proteins and 2 genes coding for defensin-like family proteins. SAG12 is the most abundant differentially expressed gene at 23 DAP with 1.591 % and is down-regulated about 3-fold at 35 DAP with 0.566 % total abundance. The 20 most abundant differentially expressed genes at 35 DAP cover a narrower range of genes coding for 2S seed storage proteins, oleosins, cruciferins and protease inhibitor proteins. The 5 most abundant differentially expressed genes are the 2S seed storage proteins 1 to 4 and cruciferin 3. The 2S seed storage protein 1 is the most abundant differentially expressed gene with 2.657 % abundance and is up-regulated 374-fold compared to time-point 23 DAP. From all 490 genes showing a significantly different expression level at the two timepoints, 23 were exclusively expressed at 23 DAP and 54 at 35 DAP (see additional file 1: 490_differentially_expressed_genes.xls). A number of cases were observed where different matched *A. thaliana* genes with similar functional annotations were expressed at the two different time-points, e.g. different *A. thaliana* genes belonging to the large cysteine protease and lipid transfer protein families were exclusively expressed at 23 DAP and 35 DAP, respectively. In contrast to cruciferin A (CRA1) and cruciferin 3 (CRU3), cruciferin 2 (CRU2) was exclusively expressed at 35 DAP.

Expression data produced by microarray hybridization for 79 developmental stages of *A. thaliana* were downloaded from EBI ArrayExpress (accession number E-AFMX-9, [18]) and expression profiles were compared with expression profiles obtained by LongSAGE for *B. napus* seeds using hierarchical cluster analysis and Spearman's rank correlations according to Lu et al. [19]. Both *B. napus* SAGE data sets showed weak correlations with expression value estimates from *A. thaliana* microarray data sets with the highest correlation for the seed developmental stages ATGE_77 (0.522 for 23 DAP) and ATGE_78 (0.440 for 35 DAP). Sample ATGE_77 was derived from *A. thaliana* seed stage 4, early to late heart stage embryos and sample ATGE_78 from seed stage 5, late heart to mid torpedo embryos. Embryos dissected from sample 23 DAP were classified to be at the mid torpedo stage and from sample 35 DAP to be at the early curled cotyledon stage. Genes that were differentially expressed between two different *A. thaliana* seed stages (ATGE_77 and ATGE_78) were identified at different stringencies and compared with the 490 genes identified to be differentially expressed between 23 DAP and 35 DAP in *B. napus* as described above. From the 490 genes identified by LongSAGE 76 (15 %) were not represented on the microarray chip. From the remaining 414 genes a maximum of 29 % (122) overlapping genes were detected that were mainly highly abundant genes (e.g. 2S seed storage proteins). The percentage of overlapping genes identified here between the two different crucifer species during morphological similar seed developmental stages appears to be high when taking into account that the correlation of microarray hybridization and LongSAGE data analysis has been found to be low (below 0.5) even when using the same RNA preparations and LongSAGE has been characterized to be more efficient at identifying differentially expressed tags than microarray technology [19].

Gene ontology enrichment analysis

A total of 7,208 of the 8,255 matched genes were expressed at 23 DAP, while 4,875 genes were expressed at 35 DAP (Table 2). 3,828 of 8,255 genes (46.4 %) were expressed at both time points. Figure 1 gives a comparison of GO categories that are statistically over-represented for the genes expressed at 23 DAP compared to all other *A. thaliana* genes. From the 490 genes matched by statistically significant differentially expressed tags between 23 and 35 DAP, 194 were up-regulated at 23 DAP and 296 were up-regulated at 35 DAP. Statistically enriched GO terms found for the differentially expressed genes either up-regulated at 23 or up-regulated at 35 DAP were calculated using GOEAST [20] and details of the results are provided in additional file 2: Enriched_GO_terms_490_differentially_expressed_genes.xls. Within the category 'Biological Process', the GO terms 'developmental process', 'localization' and 'metabolic process' are enriched at both time-points. Generally, GO terms that are statistically enriched at the highly general information level 2 are similar for the two time-points,. Comparison of statistically enriched GO terms for 490 differentially expressed genes at 23 and 35 DAP reveals their involvement in seed maturation, regulation of meristem organization and photomorphogenesis, seed coat development, water and fluid transport, cell-cell signaling, cell wall modification and glycerolipid, neutral lipid, acylglycerol, triglyceride, and carbohydrate metabolic processes at 23 DAP, and in protein processing, protein targeting, photosynthesis, fatty acid biosynthesis, lipid localization, storage and lipid metabolic processes at 35 DAP.

Genes matched in sense and antisense orientation

In the 23 DAP library, 6 of the 30 most abundant tags matched in sense as well as in antisense orientation to the *Brassica* EST dataset, oriented based on the *A. thaliana* proteome. The same was true for 10 of the 30 most abundant tags at 35 DAP. In most of these cases 500 to 1,500 ESTs were matched by a single tag in one direction, whereas only 1 to 13 ESTs were matched by the same tag in the opposite direction. A detailed analysis of these matches revealed that in all cases the smaller fraction of matched ESTs were chimeric sequences composed of two regions, derived from different gene loci, that aligned to the *A. thaliana* proteome or to *B. napus* full-length cDNAs in different orientations. These apparently chimeric ESTs may represent cloning artefacts derived from tail-to-tail or head-to-head ligations during cDNA preparation for EST library production. In some cases a putative chimeric composition was also found for individual ESTs, with two different regions derived from different genes being aligned in sense and antisense orientation to the *A. thaliana* proteome or to oriented *B. napus* full-length cDNAs. To remove these putative chimeric ESTs from our analysis, tag-to-EST-to-*A. thaliana* gene locus matches were only accepted if at least 4 tag-matched *Brassica* ESTs matching a particular *A. thaliana* locus were found. In addition, tag-to-EST-to-gene matches of particular tags matching in sense as well as in antisense orientation were only included in the analysis if the frequency of ESTs matched in one direction was at least 1 % of the frequency of ESTs matched in the other direction. After removal of putative chimeric ESTs, the number of accepted tags that were successfully matched via *Brassica* ESTs to *A. thaliana* genes was reduced stongly from 8,255 to 5,120 (Table 2), due to the exclusion of many low-copy ESTs from the analysis. The occurence of chimeric ESTs in databases has been documented previously. For example, Hillier *et al*. [21] found chimeric ESTs in a dataset of 280,000 human ESTs at a frequency of up to 1.04 %. The strategy we applied resulted in removal of all 72 chimeric ESTs that were identified by manual analysis of the 47,446 ESTs matched by the 30 most abundant tags in the two LongSAGE libraries.

Based on the tag-to-gene matching strategy described above it was found that 24.3 % of all matched tags matched *A. thaliana* genes via oriented *Brassica* ESTs in antisense orientation (corresponding to 1762 of 5,120, 34.4 % of all matched genes, Table 2). In the 23 and 35 DAP libraries, only 68.3 % and 74.6 % of the respective genes were matched exclusively by sense tags (Table 2). On the other hand, 6.6 % and 6.8 % of all genes were matched exclusively by antisense tags at 23 DAP and 35 DAP, respectively, while 25.1 % (23 DAP) and 18.6 % (35 DAP) of the genes were matched by both sense and antisense tags. Genes expressed with high abundance in sense orientation, and particularly members of complex gene families, often also showed coexpression of antisense transcripts at a lower frequency (Figure 2, 275 differentially expressed genes matched in sense and antisense orientation). A detailed analysis of tags matching to the four 2S seed storage proteins of *A. thaliana* (AT4G27140, AT4G27150, AT4G27160, AT4G27170) was performed by aligning them to *Brassica* spp. napin cDNAs and genomic napin sequences. This revealed that some tags matched in antisense orientation to highly conserved regions of napin transcripts, derived from multiple *B. napus* loci, whereas other tags matched in sense orientation to highly diverse regions of napin transcripts derived from a limited number of *B. napus* loci.

Detection of napin-related sense and antisense transcript expression by Real-time RT-PCR

At 35 DAP, a high number of tags matched in sense orientation to *Brassica* ESTs that were similar to *A. thaliana* proteins containing the AAI_SS domain (NCBI Conserved Domain Database cd00261). This protein domain is found in the alpha-amylase inhibitors (AAIs) and seed storage (SS) protein subfamily including plant lipid transfer proteins, seed storage proteins and trypsin-alpha amylase inhibitors. At 35 DAP the most abundantly expressed tags were matched via *Brassica* ESTs in sense orientation to the four *A. thaliana* 2S seed storage proteins 1-4 (napins) AT4G27140, AT4G27250, AT4G27160 and AT4G27170 (37,182 to 41,806 tags per milliion). The four *A. thaliana* 2S seed storage proteins share 75-91 % amino acid (aa) sequence identity with each other. The four 2S seed storage protein genes and 7 other protease inhibitor/seed storage/lipid transfer protein (LTP) family protein genes were also among the genes with the highest tag counts matched in antisense orientation (see additional file 3: All_matched_tags.xls).

For comparison of SAGE results with quantitative Real-time RT-PCR detection, two sets of primers and their antisense reverse-complemented sequences were derived from SAGE primer amplified napin 3’-termini (see above) and from assembled contigs of 8,977 *B. napus* ESTs aligning with an e-value cut-off of 1e-6 to *A. thaliana* 2S seed storage protein 1 (AT4G27140). One primer set (napin) was designed specifically for the major cluster of *B. napus* assembled EST contigs most similar to the *A. thaliana* 2S seed storage proteins 1-4 (59 % to 70% on the aa level). Another set of primers (napin-related) was designed specifically for the minor cluster of *B. napus* assembled EST clusters grouping in-between the four *A. thaliana* 2S seed storage proteins and the related protease inhibitor/seed storage/lipid transfer (LTP) family protein AT5G54740 (about 47 % to 57 % identity on the aa level). AT5G54740 represents the next closely related protease inhibitor/seed storage/lipid transfer protein (LTP) family protein gene in *A. thaliana,* which like other more distantly related proteins of this large family (about 113 members in *A. thaliana*) exhibits an AAI_SS domain and shares about 53-57 % aa identity with the *A. thaliana* 2S seed storage proteins 1-4. Fig 3 shows the expression profile from days 17 to 70 after pollination from seeds for sense and antisense napin and napin-related transcript expression. The coexpression of sense and antisense transcripts was confirmed for a minor subgroup of napin-related transcripts by strand-specific Real-time RT-PCR, but not for the major napin transcript group as indicated by LongSAGE.

Discussion

B. napus and *A. thaliana* are members of the Brassicaceae family and exhibit around 87 % conservation of their protein-coding sequences [22]. Based on this close phylogenetic relationship, the tag-to-gene matching strategy applied in this study for LongSAGE expression profiling in the non-model plant *B. napus* uses the information available for the well-annotated proteome of the model species *A. thaliana*. Our results show that expression profiling can be achieved for *B. napus* seeds using publically available EST resources and the annotated proteome of *A. thaliana* for tag-to-gene assignment.

Major factors influencing the accuracy of the first step of the SAGE analysis are the sequencing error rate, the tag length and the transcriptome complexity. The accuracy of SAGE depends on the ability to unambiguously match the tags to the genes of origin. Theoretical calculations show that >99.8 % of 21 bp LongSAGE tags are expected to occur only once in large genomes such as the human genome, which is about twice the size of the *B. napus* genome, with the remaining tags matching duplicated genes or repeated sequences [9]. In a polyploid plant genome the proportion of duplicated genes is generally considerably higher, and our strategy of matching tags via ESTs to a related model plant proteome will not *per se* allow differentiation between closely related duplicated genes. However, matching to multiple members of gene families still allows biological inferences from the LongSAGE data, since the predicted function in each case is identical. In our case only around 90 % of LongSAGE tags were assigned to a unique *A. thaliana* gene. This underlines the highly duplicated nature of the *B. napus* and *A. thaliana* genomes [17] and reflects the well-documented ancient hexaploidization events in major angiosperm phyla [23]. Despite this we found no example where a *B. napus* tag matched multiple *A. thaliana* genes with apparantly unrelated functions. This suggests that the matching strategy we applied is specific enough with 21 bp *B. napus* tags to draw biological conclusions on the expression levels of functionally related gene families in *B. napus* seeds. Raising the blastx e-value stringency to 1e-100 for matching of *Brassica* ESTs to the *A. thaliana* proteome was found to reduce the average number of genes per gene family to which the observed tags were assigned. However, this also resulted in a higher number of tags that could no longer be assigned to *A. thaliana* genes and thus could not be annotated (data not shown). For LongSAGE analysis from *B. napus* we found an e-value of >1e-6 useful; this is also the cut-off value used by NCBI in the UniGene database for EST-to-protein matching to create clusters of transcript sequences that are expected to come from the same transcription locus.

The development of oilseeds can be broadly divided into three phases, (1) the pattern formation/cell proliferation phase, (2) the maturation phase, with lipid and storage protein accumulation, and (3) the dessication phase. At 23 DAP *B. napus* seed has reached the end of phase 1 and is beginning the transition to phase 2, while at 35 DAP the seed is in the middle of the maturation phase. The most abundantly expressed gene detected at 23 DAP, SAG12, is a cysteine protease expressed during leaf senescence in *A. thaliana* [24]. In *B. napus* SAG12 is encoded by two orthologous copies, BnSAG12-1 and BnSAG12-2, involved in leaf senescence and seed development, respectively. BnSAG12-2 and BnCysP1, a cysteine protease with around 91 % identity to BnSAG12-2, are expressed in the early phase of *B. napus* seed development within the inner integument of the testa and are predicted to function in the disposal of proteins in the inner integument cells commited to programmed cell death [25]. Other genes encoding proteases and protease inhibitor genes are also strongly represented within the 30 most abundantly expressed genes at 23 DAP and 35 DAP, e.g. a δ-vascular processing enzyme (δ-VPE) at 23 DAP. VPEs are cysteine-type proteases which were originally discovered to be responsible for maturation of seed storage proteins. In *A. thaliana* δ-VPE was found to be involved in developmental cell death during embryogenesis [26]. In general, proteases are considered to be key regulators of plant development involved in post-translational modification or activation of proteins and enzymes of diverse biological processes. Proteases play a role both in general protein turnover and in highly specific regulation of plant development. Dong *et al*. [27] also described an elevated expression of genes coding for subtilisin-like protease and vacuolar processing enzymes (VPE), consistent with seed endosperm development, along with early expression of lipid transfer proteins that are required for embryo pattern formation [28, 29]. The high transcription levels of genes encoding seed storage proteins are also consistent with the results of Dong *et al*. [27] based on cDNA cloning and Northern blot expression profiling. On the other hand, our LongSAGE analysis revealed that transcripts encoded by the peroxidase gene AT4G21960 (Protein ID CAA66957) had a high abundance at 23 DAP and a subsequent drop in expression at 35 DAP, whereas Northern blot analysis appeared to reveal the opposite expression pattern. Such differences might be due to different specificities and resolutions of the two techniques, e.g. cross-hybridization of Northern blot probes with other closely related peroxidase transcripts.

Antisense transcripts and sense-antisense transcript pairs (natural antisense transcripts, NATs) were found to be prelvalent at both developmental stages investigated. Antisense transcripts were observed earlier in *A. thaliana* and rice cDNA/EST databases and confirmed using tag-based expression profiling approaches like SAGE or Massively Parallel Signature Sequencing [30], with matching to the sequenced genomes. In some of these studies it was noted that the sense and antisense transcripts of an overlapping NAT pair (*cis*-NATs) tend to be expressed in different tissues or different conditions. Furthermore, in cases where the sense and antisense transcripts of a NAT pair were expressed in the same library, one type of transcript was usually more abundant than the other. One speculation is that *cis*-antisense transcripts could be involved in down-regulating the expression levels of their target mRNA to achieve a low protein concentration by interfering with the transcription of their sense transcript. Due to the matching strategy we applied, the sense-antisense pair matches we found to the *A. thaliana* proteome might either indicate natural sense-antisense pair expression from one *B. napus* gene, or a concurrent expression of sense and antisense transcripts from paralogous *B. napus* genes. Either way, the results clearly indicate that a significant number of antisense transcripts are expressed during *B. napus* seed development. The coexpression of sense and antisense transcripts was confirmed for a subgroup of napin-related transcripts coding for proteins from the large group of protease inhibitor/seed storage/lipid transfer protein (LTP) family by strand-specific Real-time RT-PCR. Although Real-time RT-PCR confirms the detection of antisense transcripts that are related to protease inhibitor/seed storage/lipid transfer protein (LTP) family proteins by LongSAGE, it also indicates that the use of the proteome of *A. thaliana* for annotation of tag-matched *B. napus* transcripts in case of diverse gene families limits the resolution and the detailed analysis of the genome and transcriptome structure and availability of more genetic resources for *B. napus* will help to increase the resolution in the future. The detection of transcripts for a diverse number of genes in antisense orientation by LongSAGE at both time-points suggests strong involvement of antisense transcripts in regulatory processes during *B. napus* seed development.

The comparison of LongSAGE and Real-time RT-PCR data suggests that the sequencing depth applied in this experiment limits the LongSAGE detection to transcripts with medium to high copy numbers in *B. napus* seeds. On the other hand, recent developments in next-generation sequencing (NGS) technologies allow ultra-deep transcriptome profiling using transcript tag-based techniques (e.g. DeepSAGE; [31]) that can realistically achieve transcriptome saturation and accurate quantification of low-copy transcripts. The data analysis techniques developed in the present study represent a valuable platform for future NGS-based transcriptome tagging studies in *Brassica* species. As suggested by Shendure [32], such approaches may ultimately replace microarrays as the method of choice for quantitative global transcriptome profiling as a basis for genetical genomics or systems genetics approaches. In the present study we have demonstrated that tag-based transcriptome profiling can also be effectively applied in a non-model, complex polyploid plant species.

Conclusions

This study underlines the potential of transcript tagging approaches for gene expression profiling in *Brassica* crop species via EST matching to annotated *A. thaliana* genes. The data processing pipeline adapted by matching *B. napus* tags via *Brassica* ESTs to annotated *A. thaliana* gene loci enabled differential expression profiling of during seed development in the complex *B. napus* genome, and furthermore detected an unexpectedly high proportion of EST tags matching in antisense orientation. This suggests a strong involvement of antisense transcript expression in regulatory processes during *B. napu*s seed development. Limits of tag detection for low-abundance transcripts can today be overcome by ultra-high throughput sequencing approaches, so that tag-based gene expression profiling may soon become the method of choice for global expression profiling in non-model species.

Methods

Plant materials and RNA isolation

Plants of the homozygous male-sterile winter oilseed rape maternal line ‘MSL-Express’ were cultivated under controlled growth chamber conditions (16 hour, 20°C day and 8 hour, 16°C night, 60% relative humidity) and manually fertilised with pollen from plants of the isogenic male-fertile line ‘Express 617’. Pods were harvested weekly from the main racemes of pollinated plants at time points 14 days up to 70 days after pollination (DAP). The pods were shock-frozen in liquid nitrogen and stored at -80°C, and seeds were collected by splitting the frozen pods. Total RNA was extracted from seeds using RNA Reagent (Invitrogen) and purified from DNA residues by incubation with RQ1 RNAse-free DNAse (Promega), according to the manufacturer’s recommendations.

LongSAGE library production

Total RNA extracted from two seed developmental time points, 23 DAP and 35 DAP, was processed using the I-SAGELong Kit (Invitrogen) with several modifications to allow more efficient library production and cloning. Briefly, 50 µg of polyA+ RNA was captured on the surface of magnetic oligo(dT) beads and double stranded cDNA was synthesized using SuperScript II reverse transcriptase. Double stranded cDNA was digested for 2.5 hours with *Nla*III, and adapters containing *Mme*I recognition sites were ligated overnight in ligation buffer with 15% polyethylene glycol (PEG) to the 5’ end of the cDNAs. Adaptor-linked tags were released from the beads by restriction with *Mme*I overnight, then precipitated and washed five times with 70% ethanol before being ligated overnight in pairs to form adaptor-linked ditags. These were amplified by PCR and 200 PCR reactions were pooled. Subsequently, 130 bp fragments were purified from 20 lanes on 16 % polyacrylamide gels. Purified adapter-linked ditags were redigested overnight with *Nla*III to remove adapters and 34 bp ditags were purified from a 16 % polyacrylamide gel. The ditags were ligated overnight according to Kenzelmann and Mühlemann [33] to form concatemers, then partially digested with 2 U of *Nla*III for 1 min at 37°C according to Gowda *et al*. [15] to create linearized clonable concatemers of suitable sizes. The partially-digested ligation mixture was separated on a 6% polyacrylamide gel. DNA fragments between 0.8 and 2.5 kb were isolated and 1.6 µg concatemers were ligated to 25 ng of linearized pZErO-1 vector digested with *Sph*I. For each of the 23 and 35 DAP libraries more than 3,000 plasmid clones were sequenced in one direction by SeqWright (Houston, TX, USA).

SAGE data processing

Ditags were extracted from sequence traces using the perl script sage-prhed.pl [34]. Due to the 1% error rate associated with single pass Sanger sequencing [14], Phred sequence trace quality scores were used to remove all tags of low sequence quality (containing bases with a Phred score <20). 21 bp monotags were extracted from both ends of all correctly-sized ditags. Contaminating linker and vector sequences and low-complexity polyA tags were identified and removed by screening a local tag database using formatdb and blast (valid tags in Table 1; <0.06 % removed). A total of 819,455 *Brassica* ESTs were downloaded on Sept, 15th 2007, from NCBI dbEST (http://www.ncbi.nlm.nih.gov/dbEST/) using the keyword search (Brassica napus[Organism] OR Brassica oleracea[Organism] OR Brassica rapa[Organism] AND EST[Keyword]). About 30 % of all *Brassica* ESTs downloaded from the NCBI dbEST database were produced from seed developmental stages. The orientation of all *Brassica* ESTs was determined by blastx analysis against the TAIR7 *A. thaliana* proteome accepting only matches with an e-value of less then 1e-6. Any *Brassica* ESTs that did not match the *A. thaliana* proteome using these criteria were removed from the dataset. ESTs that did match the NCBI VecScreen vector database using blastn with pre-set parameters were removed from the analysis. The remaining 834,732 ESTs from NCBI and PBI were annotated based on the alignment with the *A. thaliana* proteome and orientated in sense using ReadSeq [35]. Identitag perl scripts were modified so that virtual tags from all positions were extracted from the *Brassica* ESTs, and the modified Identitag perl scripts were used to create a MySQL database of virtual tags [36]. To differentiate between tags that matched *B. napus* seed transcripts in sense and in antisense orientation, the oriented sense EST dataset was used in combination with an oriented reverse-complemented EST dataset (antisense EST dataset). For identification of differentially expressed tags the online-tool DiscoverySpace (http://www.bcgsc.ca/discoveryspace; [37]) was used and significance calculated according to Audic and Claverie [38]. The online-tools WEGO (http://wego.genomics.org.cn; [39]) and GOEAST (http://omicslab.genetics.ac.cn/GOEAST; [20]) were used for gene ontology (GO) enrichment analysis. Hierarchical cluster analysis with average linkage (UPGMA) and Euclidean distance was performed for GC-RMA processed, normalized data using the software program HCE 3.0 [40]. Spearman’s rank correlations were used to compare Affymetrix chip and LongSAGE [19]. Differentially expressed genes were identified using the software program SAM [41].

Cloning of 3’ends of transcripts using SAGE tag primers

For validation, the last 17 base pairs of observed LongSAGE tags (excluding the CATG restriction site) were used as primers for RT-PCR amplification together with anchored oligo-dT primers, as described by Xiao *et al*. [42], and 2.5 µg total RNA from time point 35 DAP. cDNA synthesis was performed using ThermoX Reverse Transcriptase (Invitrogen) according to the manufacturer’s instructions. The PCR used 0.5 U *Taq* Polymerase (Eppendorf) and 250 nM of each primer, with 2 min at 95°C followed by 30 cycles of 1 min 95°C, 1 min 50°C, 2 min 72°C and final extension at 72°C for 10 min.

Strand-specific Real-time RT-PCR

Total RNA was extracted as described above and additionally purified using RNAeasy columns (Qiagen). DNase digestion was performed using RQ1 RNAse-free DNAse (Promega) according to manufacter's recommendations. cDNA was synthesized after denaturation for 10 min at 70°C using 1 µg of total RNA, 100 nM of gene and strand-specific reverse primer and Omniscript Reverse Transcriptase (Qiagen) according to the manufacturer’s recommendations in a total volume of 10 µl. After the cDNA synthesis reaction was finished the enzyme was inactivated for 15 min at 95°C. Real-time PCR was performed using a 7500 Fast Real-time PCR system (Applied Biosystems) with a total volume of 10 µl using 1 µl of cDNA (1:10 diluted), 5 µl of Power SYBR Green PCR master mix (Applied Biosystems) and 100 nM each of the forward and reverse primers. Amplification was for 2 min at 50°C, 10 min at 95°C followed by 40 cycles of 15 sec at 95°C and 1 min at 60°C. Controls included PCR amplification reactions using cDNA preparations without adding the reverse transcriptase enzyme to check for genomic DNA contamination, and PCR amplification reactions without adding reverse primers during cDNA synthesis to check for a loss in strand-specificity by RNA self priming leading to primer-independent cDNA synthesis [43]. Primers were designed by using *Brassica* ESTs that were matched using blastx at an e-value cut-off of 1e-6 with selected *A. thaliana* proteins. All matched *Brassica* ESTs were assembled using the software CAP3 [44] and edited manually. Alignments of coding DNA were contructed from aligned amino acid sequences using the software RevTrans [45]. Contigs from these assemblies were used for primer design using the software PrimerExpress (Applied Biosystems). Primer pairs used for sense transcript detection were Napin_for (5’ACGAGCTCCACCAGGAAGAG-3’) and Napin_rev (5’-ACGGCTTTGGATGCTCCTT-3’), which amplify position 320 to 383 relative to the coding region of gene napB [GenBank:X14492], and Napin-related_for (5’-AACAAGCAGCCAAGTCAGTTAGG-3’) and Napin-related_rev (5’-GATGCGGGTGGACTGGAAT-3’) which amplify position 323 to 384. For antisense transcript detection the reverse complemented sequences of the primers were used. Expression fold-changes were calculated relative to the endogeneous reference gene cyclophilin (GenBank accession M55018) according to the $2^{-\Delta\Delta Ct}$ method using three replicates [46].

Authors’ contributions

RS conceived the project and WF provided the plant material. CO and BH generated the SAGE libraries. CO adapted and applied bioinformatic tools for analysis of LongSAGE tags. CO designed the primers for Realtime RT-PCR and BH performed the experiments. CO and RS developed the experimental strategy, and were responsible for the preparation of the manuscript. RS and WF participated in the coordination of the study. All authors read and approved the final manuscript.

Acknowledgements

This work was partially supported by the German Research Foundation (DFG) grant SN14/7-2. The plant material was generated by Roland Baetzel and Rosa Allerdings during the BMBF-funded project GABI-GARS: Genome Analysis in Rapeseed. We thank Raju Datla, Dustin Cram and Jacek Nowak from the Plant Biotechnology Institute, NRC, Saskatoon, Canada, and Isobel Parkin, Wayne Clarke and Steve Robinson from Agriculture and Agri-Food Canada, Saskatoon, Canada, for helpful discussions on tag-to-gene matching strategies and bioinformatic support. Stefanie Wellner provided assistance with the cloning and Real-time PCR analysis.

References

1. Ruuska SA, Girke T, Benning C, Ohlrogge JB: **Contrapuntal networks of gene expression during Arabidopsis seed filling.** *Plant Cell* 2002, **14**:1191-1206.
2. Wobus U,. Weber H: **Seed maturation: genetic programmes and control signals.** *Curr Opin Plant Biol* 1999, **2**:33-38.
3. McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, Tossberg J, Nickle T, Levin JZ, Law M, Meinke D, Patton D: **Insertional Mutagenesis of Genes Required for Seed Development in *Arabidopsis thaliana*.** *Genetics* 2001, **159**:1751-1763.
4. Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J: **Microarray Analysis of Developing Arabidopsis Seeds.** *Plant Physiol*. 2000, **124**:1570-1581.
5. Li F, Wu X, Tsang E, Cutler AJ: **Transcriptional profiling of imbibed *Brassica napus* seed.** *Genomics* 2005, **86**:718-730.
6. Hudson ME, Brugginkd T, Changa SH, Yuc W, Hana B, Wanga X, van der Toornd P, Zhua T: **Analysis of Gene Expression during *Brassica* Seed Germination Using a Cross-Species Microarray Platform.** *Crop Sci* 2007, **47**:96-112.
7. Xiang D, Datla R, Li F, Cutler A, Malik MR, Krochko JE, Sharma N, Fobert P, Georges F, Selvaraj G, Tsang E, Klassen D, Koh C, Deneault JS, Nantel A, Nowak J, Keller W, Bekkaoui, F: **Development of a *Brassica* seed cDNA microarray.** *Genome* 2008, **51**:236-42.
8. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: **Serial analysis of gene expression.** *Science* 1995, **270**:484-487.
9. Saha S, Sparks AB, Rago C, Akmaev V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE: **Using the transcriptome to annotate the genome.** *Nat Biotechnol*. 2002, **20**:508-512.
10. Matsumura H, Reich S, Ito A, Saitoh H, Kamoun S, Winter P, Kahl G, Reuter M, Kruger DH, Terauchi R: **Gene expression analysis of plant host-pathogen interactions by SuperSAGE.** *Proc Natl Acad Sci U S A* 2003, **100**:15718-15723.
11. Ibrahim AFM, Hedley PE, Cardle L, Kruger W, Marshall DF, Muehlbauer GJ, Waugh R: **A comparative analysis of transcript abundance using SAGE and Affymetrix assays.** *Functional and Integrative Genomics* 2005, **5**:163-174.
12. Angelastro JM, Ryu EJ, Törõcsik B, Fiske BK, Greene LA: **Blue-white selection step enhances the yield of SAGE concatemers.** *BioTechniques* 2002, **32**:484-486.
13. Chen J, Sun M, Lee S, Zhou G, Rowley JD, Wang SM: **Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags.** *Proc Natl Acad Sci USA* 2002, **99**: 12257-12262.
14. Akmaev VR, Wang C J: **Correction of sequence-based artifacts in serial analysis of gene expression.** *Bioinformatics* 2004, **20**:1254-1263.
15. Gowda M, Venu RC, Li HM, Jantasuriyarat C, Chen SB, Bellizzi M, Pampanwar V, Kim HR, Dean RA, Stahlberg E, Wing R, Soderlund C, Wang GL: ***Magnaporthe grisea***

Infection Triggers RNA Variation and Antisense Transcript Expression in Rice. *Plant Physiol.* 2007 144: 524-533.

16. Strausberg RL, Levy S: **Promoting transcriptome diversity.** *Genome Res.* 2007, **17**: 965-968.
17. Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ: **Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*.** *Genetics* 2005, **171**:765-781.
18. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann J: **A gene expression map of Arabidopsis development**. *Nature Genetics* 2005, **37**:501-506.
19. Lu J, Lal A, Merriman B, Nelson S, Riggins G: **A comparison of gene expression profiles produced by SAGE, long SAGE, and oligonucleotide chips**. *Genomics* 2004, *84*:631-636.
20. Zheng Q, Wang XJ: **GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis.** *Nucleic Acids Research* 2008, **36**: Web Server issue W358–W363.
21. Hillier LD, Lennon G, Becker M, Bonaldo MF, Chiapelli B, Chissoe S, Dietrich N, DuBuque T, Favello A, Gish W, Hawkins M, Hultman M, Kucaba T, Lacy M, Le M, Le N, Mardis E, Moore B, Morris M, Parsons J, Prange C, Rifkin L, Rohlfing T, Schellenberg K, Bento Soares M, Tan F, Thierry-Meg J, Trevaskis E, Underwood K, Wohldman P, Waterston R, Wilson R, Marra M: **Generation and analysis of 280,000 human expressed sequence tags.** *Genome Res*. 1996, **6**:807-828.
22. Cavell AC, Lydiate DJ, Parkin AIP, Dean C, Trick M: **Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome.** *Genome* 1998, **41**:62-69.
23. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M., Adam-Blondon AF, Weissenbach J, Quétier F, Wincker P, the French-Italian Public Consortium for Grapevine Genome Characterization: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla.** *Nature 2007,* **449**:463-467.
24. Lohman KN, Gan S, John MC, Amasino RM: **Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*.** *Physiol. Plant.* 1994, **92**:322-328.
25. Wan L, Xia Q, Qiu X, Selvaraj G: **Early stages of seed development in *Brassica napus*: a seed coat-specific cysteine proteinase associated with programmed cell death of the inner integument.** *The Plant Journa*l 2002, **30**:1-10.
26. Hara-Nishimura I, Hatsugai N, Nakaune S, Kuroyanagi M, Nishimura M: **Vacuolar processing enzyme: an executor of plant cell death.** *Current Opinion in Plant Biology* 2005, **8**:404-408.
27. Dong J, Keller WA, Yan W, Georges F: **Gene expression at early stages of *Brassica napus* seed development as revealed by transcript profiling of seed-abundant cDNAs**. *Planta* 2004, **218**:483-491.

28. Vroemen CW, Langeveld S, Mayer U, Ripper G, Jürgens G, Kammen AV, de Vries SC**: Pattern formation in the Arabidopsis embryo revealed by position-specific lipid transfer protein gene expression.** *Plant Cell* 1996, **8**:783-791.
29. Kader JC: **Lipid-transfer proteins in plants.** *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:627-54.
30. Wang XJ, Gaasterland T, Chua N-H: **Genome-wide prediction and identification of cis-natural antisense transcripts in *Arabidopsis thaliana*.** *Genome Biology* 2005, **6**:R30.
31. Nielsen KL, Høgh AL, Emmersen J: **DeepSAGE-digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples.** *Nucleic Acids Research* 2006, **34**:e133.
32. Shendure J: **The beginning of the end for microarrays?** *Nature Methods* 2008, **5**:585-587.
33. Kenzelmann M, Mühlemann K: **Substantially enhanced cloning efficiency of SAGE (Serial Analysis of Gene Expression) by adding a heating step to the original protocol.** *Nucleic Acids Research* 1999, 27:917-918.
34. Emmersen J, Heidenblut AM, Høgh AL, Hahn SA, Welinder KG, Nielsen KL: **Discarding duplicate ditags in LongSAGE analysis may introduce significant error.** *BMC Bioinformatics* 2007, **8**:92.
35. Gilbert D: **ReadSeq: read & reformat biosequences** 2001, [http://iubio.bio.indiana.edu/].
36. Keime C, Damiola F, Mouchiroud D, Duret L, Gandrillon O: **Identitag, a relational database for SAGE tag identification and interspecies comparison of SAGE libraries.** *BMC Bioinformatics* 2004, **5**:143.
37. Robertson N, Oveisi-Fordorei M, Zuyderduyn SD, Varhol RJ, Fjell C, Marra M, Jones S, Siddiqui, A: **DiscoverySpace: an interactive data analysis application.** *Genome Biol* 2007, **8**:R6.
38. Audic S, Claverie JM: **The significance of digital gene expression profiles.** *Genome Res* 1997, **7**:986-995
39. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J: **WEGO: a web tool for plotting GO annotations.** *Nucleic Acids Research* 2006, **34**: Web Server issue W293–W297.
40. Seo J, Bakay M, Chen Y-W, Hilmer S, Shneiderman B, Hoffman EP: **Interactively optimizing signal-to-noise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays**. *Bioinformatics* 2004, **20**:2534-2544.
41. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response**. *PNAS* 2001, **98**: 5116-5121.
42. Xiao N-Z, Ba L, Holm PB, Wang X-Z, Bowra S: **Quantitative transcript analysis in plants: improved first strand cDNA synthesis.** *Acta Biochimica et Biophysica Sinica* 2005, **37**:429-434.
43. Haddad F, Qin AX, Giger JM, Guo H, Baldwin KM: **Potential pitfalls in the accuracy of analysis of natural sense-antisense RNA pairs by reverse transcription-PCR.** *BMC Biotechnology* 2007, **7**:21.
44. Huang,X, Madan A: **CAP3: A DNA sequence assembly program**. *Genome Res* 1999, **9**:868-877.

45. Wernersson R, Pedersen AG: **RevTrans - Constructing alignments of coding DNA from aligned amino acid sequences**. *Nucl Acids R*es 2003, **31**: 3537-3539.
46. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.** *Methods* 2001, **25**:402-408.

Figure 1: Statistically enriched Gene Ontology terms level 2 at $p \leq 0.05$ in *B. napus* seeds for LongSAGE tag-matched *A. thaliana* gene loci at 23 DAP, compared with level 2 GO terms for all other *A. thaliana* gene loci. The plot was produced using the online tool WEGO (Ye *et al*., 2006).

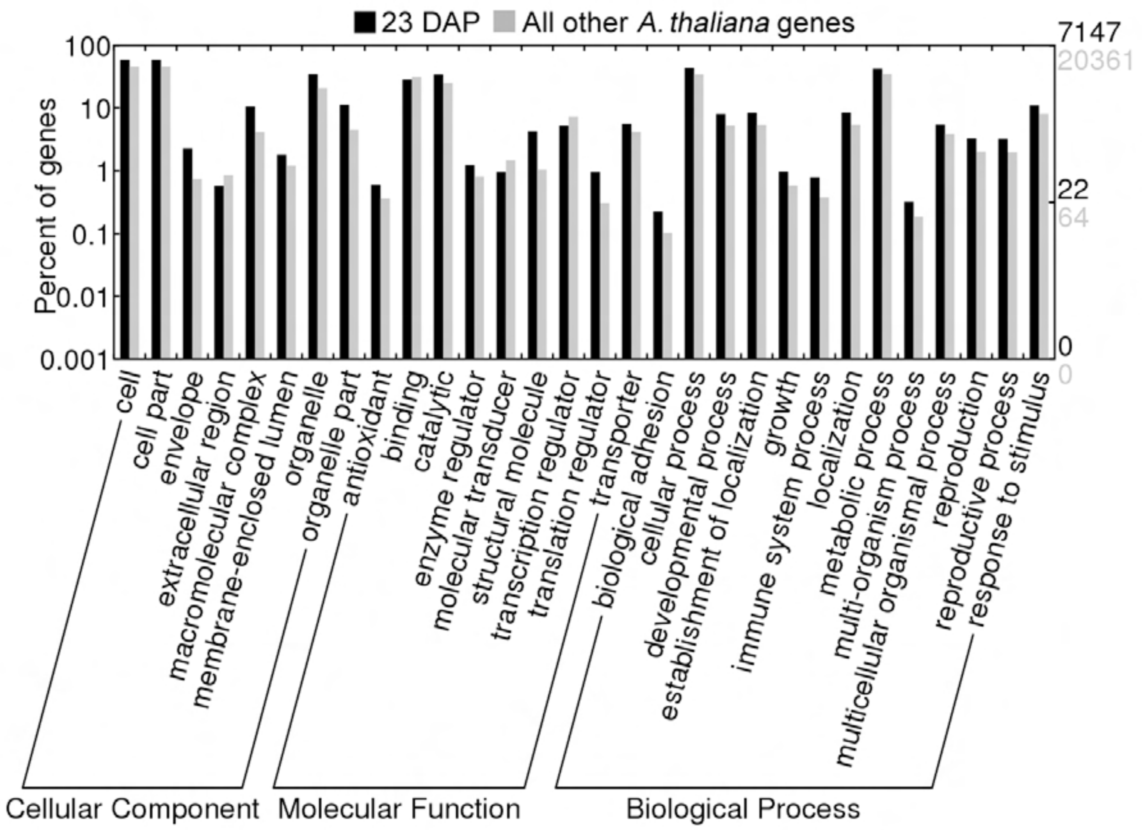


Figure 2: Heat map of normalized LongSAGE counts for 375 differentially expressed *A. thaliana* genes matched by LongSAGE tags in sense and/or antisense orientation at 23 DAP and 35 DAP. Expression is represented on a logarithmic scale. The descriptions of the 10 most abundantly expressed genes in the respective libraries are highlighted in red for 23 DAP, in blue for 35 DAP or in black for genes that were among the 10 most abundant at both time-points. If no expression was detected counts were set to 0.1 (black squares).

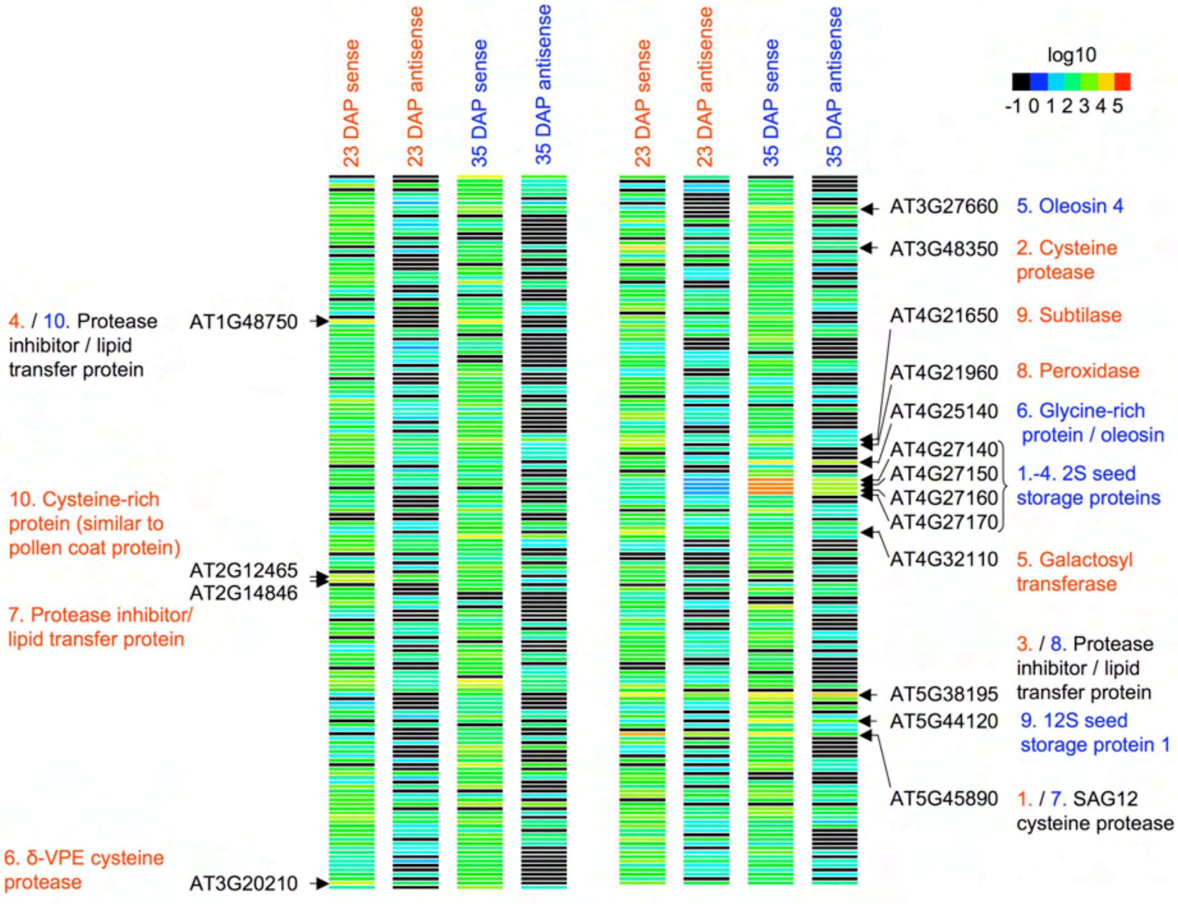


Figure 3: Real-time RT-PCR expression profile during *B. napus* seed development using primers specific for the coding (A - sense transcripts) and non-coding strands (B - antisense transcripts) of a group of napin and napin-related transcripts. Square symbols represent napin transcripts most similar to 2S seed storage proteins AT4G27140, AT4G27150, AT4G27160, AT4G27170), while triangular symbols represent napin-related transcripts most similar to protease inhibitor/seed storage/lipid transfer protein AT5G54740.

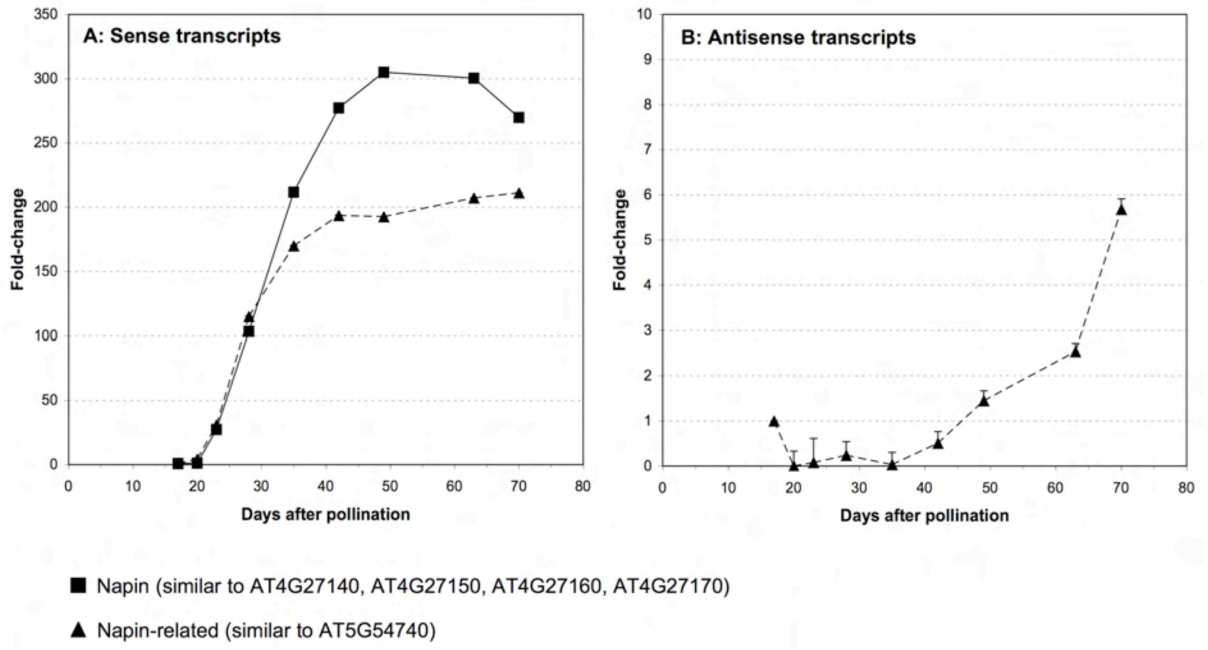


Table 1: Summary of LongSAGE tags sequenced from *B. napus* seeds at 23 and 35 days after pollination (DAP).

	23 DAP library		**35 DAP library**		**Combined data**	
Category	**Unique sequences**	**All sequences**	**Unique sequences**	**All sequences**	**Unique sequences**	**All sequences**
Total extracted ditags	-	34,984	-	15,919	-	50,903
Rejected ditags (Phred score ≤ 20[a])	-	10,685	-	3,306	-	13,991
Accepted ditags (Phred score > 20[a])	23,996	24,299	11,846	12,613	35,842	36,912
Total extracted tags	24,655	48,534	12,714	25,198	37,369	73,732
Accepted tags (excluding polyA tags)	24,654	48,514	12,713	25,183	37,367	73,697
Copy number of accepted tags						
1	18,183 (73.8. %)	18,183 (37.5 %)	9,741 (76.6 %)	9,741 (38.7 %)		
2-5	5,329 (21.6 %)	14,320 (29.5 %)	2,510 (19.7 %)	6,596 (26.2 %)		
6-20	1,016 (4.1 %)	9,450 (19.5 %)	406 (3.2 %)	3,668 (14.6 %)		
21-99	118 (0.5 %)	4,437 (9.1 %)	46 (0.4 %)	1,997 (7.9 %)		
>100	8 (0.03 %)	2,124 (4.4 %)	9 (0.07 %)	3,181 (12.6 %)		

[a] Using Sage-phred.pl, including duplicated ditags

Table 2: Summary of LongSAGE tag-to-gene mapping data for two libraries from.*B napus* seeds at 23 and 35 days after pollination (DAP).

Analysis including all singleton tags	23 DAP library	35 DAP library	combined data
Number of Brassica ESTs matched to the *A. thaliana* proteome		834,732	
Number of *A. thaliana* genes matched by Brassica ESTs		20,462	
Number of tags accepted for processing	24,654	12,713	32,395
Number of Brassica ESTs matched by tags	256,420	214,981	301,180
Total number of accepted tags matched to Brassica ESTs	12,728 of 24,654 (51.6 %)	7,232 of 12,713 (56.9 %)	5,955 of 32,395 (49.3 %)
Number of accepted tags not matched to any Brassica EST	11,926 of 24,654 (48.4 %)	5,481 of 12,713 (43.1 %)	6,400 of 32,395 (50.6 %)
Number of accepted singleton tags matched to Brassica ESTs	7,615 of 18,183 (41.9 %)	4,819 of 9,741 (49.5%)	11,421 of 26,498 (43.1 %)
Number of accepted tags with ≥2 counts matched to Brassica ESTs	5,513 of 6,471 (85.2 %)	2,413 of 2,972 (81.2 %)	6,010 of 7,727 (77.8 %)
Number of *A. thaliana* genes matched via Brassica ESTs and tags	7,208	4,875	8,255
Number of *A. thaliana* genes matched and oriented via Brassica ESTs and tags (putative chimeric ESTs removed)	4,478	3,005	5,120
Number of genes matched in both directions	1,125 (25.1 %)	560 (18.6 %)	1,453 (28.4 %)
Number of genes matched in sense orientation exclusively	3,057 (68.3 %)	2,242 (74.6 %)	3,358 (69.5 %)
Number of genes matched in antisense orientation exclusively	296 (6.6 %)	203 (6.8 %)	309 (6.1 %)

Table 3: Comparison of fold-changes for selected genes between 23 and 35 DAP in *B. napus* seeds measured by LongSAGE and Real-time RT-PCR. Primers for Real-time RT-PCR were designed using assembled *B. napus* ESTs matched to *A. thaliana* proteins using blastx with *e*-values greater than 1e-6.

***A. thaliana* gene locus**	**Gene**	**Cumulated counts,35:23 DAP LongSAGE (non-normalized)**	**Fold-change 35:23 DAP LongSAGE (normalized)**	**Fold-change 35:23 DAP Real-time RT-PCR**	***B. napus* EST contigs assembled by CAP3**
AT2G40890	Coumarate 3-hydroxylase (C3H)	2:2	1.9	2.0 4.1	Contig 1 from 5 ESTs Contig 2 from 2 ESTs
AT5G48930	Hydroxycinnamoyl-Coenzyme A shikimate/quinate hydroxycinnamoyl-transferase (HCT)	5:3	3.2	2.3	1 EST
AT1G15950	Cinnamoyl-CoA reductase 1 (CCR1)	2.2:1	4.3	0.9	Contig from 4 ESTs
AT2G23910	Cinnamoyl-CoA reductase-related (CCR1-related)	0:0	-	1.2 1.7	Contig 1 from 22 ESTs Contig 2 from 13 ESTs
AT3G19450	Cinnamoyl alcohol dehydrogenase (CAD1)	8:9.5	1.6	2.1	Contig 3 from 4 ESTs
AT4G27140 AT4G27150 AT4G27160 AT4G27170	2S seed storage proteins (napin)	669.1:3.4 560.1:1.9 575.3:1.9 636.5:7.9	374.1 564.4 579.6 155.0	1224	Contig from SAGE primer-amplified cloned 3’-terminal napin gene fragments

Table 4: Summary of the 20 most abundant genes expressed in *B. napus* seeds at 23 and 35 days after pollination (DAP) based on LongSAGE data analysis.

Timepoint	Rank	*A. thaliana* locus	No. of matched *Brassica* ESTs at 1e-6	No. of different tags matched to *Brassica* ESTs	Mean tag count per locus	Frequency of tag counts (%, normalized)	Description of matched *A. thaliana* gene locus
23 DAP	1	AT5G45890	4,452	44	772	1.590	SAG12 (SENESCENCE-ASSOCIATED GENE 12); cysteine-type peptidase
	2	AT5G50260	109	15	320	0.660	cysteine proteinase, putative
	3	AT3G48350	14	7	299	0.615	cysteine proteinase, putative
	4	AT5G38195	1,640	36	263	0.541	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	5	AT1G48750	8	4	209	0.430	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	6	AT3G20210	630	13	156	0.323	DELTA-VPE (delta vacuolar processing enzyme); cysteine-type endopeptidasee
	7	AT2G12465	160	11	134	0.275	LCR50 (Low-molecular-weight cysteine-rich 50)
	8	AT4G32110	852	20	127	0.263	transferase, transferring glycosyl groups
	9	AT3G04120	814	30	124	0.256	GAPC (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C SUBUNIT)
	10	AT5G37474	35	10	120	0.247	Encodes a defensin-like (DEFL) family protein
	11	AT2G14846	191	13	118	0.244	protease inhibitor/seed storage/lipid transfer protein (LTP) family proteinperoxidise
	12	AT4G21960	1,312	16	118	0.243	PRXR1 (peroxidase 42); peroxidase
	13	AT5G60390	2,771	30	113	0.233	elongation factor 1-alpha / EF-1-alpha
	14	AT4G32105	79	13	109	0.225	galactosyltransferase
	15	AT1G58055	23	5	100	0.206	Encodes a defensin-like (DEFL) family protein
	16	AT1G04645	502	11	97	0.200	self-incompatibility protein-related
	17	AT1G05850	426	11	96	0.198	POM1; chitinase
	18	AT1G08830	410	14	95	0.196	CSD1 (copper/zinc superoxide dismutase 1)
	19	AT4G12960	351	6	93	0.191	gamma interferon responsive lysosomal thiol reductase family protein / GILT
	20	AT4G21650	53	8	92	0.189	subtilase family protein
35 DAP	1	AT4G27140	16,413	96	669	2.657	2S seed storage protein 1 / 2S albumin storage protein / NWMU1-2S albumin 1
	2	AT4G27170	6,691	67	637	2.528	2S seed storage protein 4 / 2S albumin storage protein / NWMU2-2S albumin 4
	3	AT4G27160	2,402	34	575	2.284	2S seed storage protein 3 / 2S albumin storage protein / NWMU2-2S albumin 3
	4	AT4G27150	249	24	560	2.224	2S seed storage protein 2 / 2S albumin storage protein / NWMU2-2S albumin 2
	5	AT4G28520	126	13	195	0.776	CRU3 (CRUCIFERIN 3)
	6	AT3G27660	1,226	15	183	0.728	OLEO4 (OLEOSIN4)
	7	AT5G38195	1,640	36	161	0.640	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	8	AT4G25140	1,119	15	146	0.580	OLEO1 (OLEOSIN1)
	9	AT5G44120	1,112	17	143	0.567	CRA1 (CRUCIFERINA); nutrient reservoir
	10	AT5G45890	4,452	44	143	0.566	SAG12 (SENESCENCE-ASSOCIATED GENE 12); cysteine-type peptidase
	11	AT1G03880	1,791	12	138	0.549	CRU2 (CRUCIFERIN 2); nutrient reservoir
	12	AT1G48750	8	4	136	0.538	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	13	AT1G79870	5	3	123	0.487	oxidoreductase family protein
	14	AT1G12920	1	2	113	0.447	ERF1-2 (EUKARYOTIC RELEASE FACTOR 1-2); translation release factor
	15	AT5G35530	401	6	102	0.406	40S ribosomal protein S3 (RPS3C)
	16	AT3G12580	31	8	102	0.405	HSP70 (heat shock protein 70); ATP binding
	17	AT3G01570	170	6	99	0.393	glycine-rich protein / oleosin
	18	AT1G69830	1	2	97	0.385	AMY3/ATAMY3 (ALPHA-AMYLASE-LIKE 3); alpha-amylase
	19	AT3G05020	401	14	86	0.340	ACP1 (ACYL CARRIER PROTEIN 1)
	20	AT5G09440	31	3	82	0.326	phosphate-responsive protein, putative

Additional data files

Additional file 1

File name: 490_differentially_expressed_genes.xls

File format: Microsoft Office Excel 2003

Title of data: 490 differentially expressed genes between 23 and 35 DAP

Description of data: Description and normalized counts for 490 loci matched by LongSAGE that were differentially expressed in *B. napus* seeds at 23 or 35 days after pollination (DAP)

Additional file 2

File name: Enriched_GO_terms_490_differentially_expressed_genes.xls

File format: Microsoft Office Excel 2003

Title of data: Enriched GO terms for 490 differentially expressed genes between 23 and 35 DAP

Description of data: Statistically enriched Gene Ontology (GO) terms exclusively found for one of the time-points 23 DAP and 35 DAP of *B. napus* seed development ($p<0.05$) for 490 differentially expressed genes calculated using GOEAST (Zheng and Wang, 2008).

Additional file 3

File name: All_matched_genes.xls

File format: Microsoft Office Excel 2003

Title of data: All LongSAGE tags extracted from *B. napus* seed at 23 and 35 days after pollination (DAP) and matched via Brassica ESTs to *A. thaliana* genes.

Description of data: Description and normalized counts for all *A. thaliana* loci matched by LongSAGE that were expressed in *B. napus* seeds at 23 or 35 days after pollination (DAP) and matched in sense and/or antisense orientation. Genes are sorted by locus name.

7 Closing discussion

Over the past few decades, molecular genetics and genome analysis techniques have played an ever-increasing role in the breeding of important agricultural and horticultural crops. The utilisation of genome mapping techniques was initially restrained by technical bottlenecks that limited the density of genetic maps and consequently made it difficult to develop useful molecular selection markers for all but the simplest monogenic traits. These constraints have been overcome in recent years through the development of high-throughput SNP marker assays and high-density genetic and physical mapping resources for many of the major crop species. Furthermore, the availability of the first complete genome sequences from model plant species enabled unprecedented insights into the genetic architecture of plants and allowed detailed descriptions of gene functions. In countless cases, the relationship of important model plants like Arabidopsis and rice to many of the major arable crop species has made it possible to transfer knowledge on genes involved in related traits from the model species to crops. In the past decade this has led to discovery of a large number of genes involved in important traits in crop genomes, and a vastly improved understanding of the genetics, inheritance and variation for a large number of traits in all important crops.

On the other hand, the majority of important agronomical traits are characterised by a complex, quantitative inheritance. The identification and practical utilisation of genes and markers involved in highly complex traits remains elusive, and there are still only few examples where marker-assisted selection is applied in breeding programs to improve selection for complex traits. In most cases these are based on individual major-gene loci contributing to the respective trait. In this respect, molecular breeding is only just beginning to contribute significantly to the overall improvement of crop plants. However, the newest developments in high-throughput genomics and whole-genome selection procedures could soon also lead to major advances in breeding for complex traits.

The works presented in this compilation describe diverse examples for the development and application of genome analysis tools for molecular breeding and trait dissection in oilseed rape. This crop, the source of a highly nutritional vegetable oil and at the same time a valuable renewable resource for the biofuel industry, is currently the most important oilseed in Europe, Canada and Asia, and the second most important oilseed crop worldwide. Due to this high economic importance, the breeding of oilseed rape/canola is accompanied in all the major producing countries by significant genomic research activities to develop resources for molecular breeding. All major international seed companies involved in oilseed rape or canola breeding actively implement biotechnological and genomic resources in their breeding programs. In Germany, numerous public-private research consortia, both on a national level and in international collaborations including partners in France, Canada, China and the UK, are involved in *B. napus* trait improvement using genomics techniques. The strong investment in oilseed rape genome research by breeding companies underlines the high priority of *B. napus* genomic studies for commercial breeding, and therewith the practical relevance of the work presented here for molecular breeding of oilseed rape.

Of particular importance for practical breeding is the generation of novel genetic variation for traits where little germplasm is available from current breeding materials. Due to the time, cost and poor economic return involved, practical plant breeders are often reluctant to invest in the development of primary germplasm for new or emerging traits (so-called “pre-breeding”, for example through interspecific hybridisation and marker-assisted backcrossing). The same is true for screening of unadapted germplasm in genebank collections or wild relatives of crop species. By taking on these tasks and generating or identifying new primary germplasm for breeding, the research community plays an important role in sustaining the progress of breeding success, since useful genetic variation is the basis of any crop breeding program.

The identification, development and genetic characterisation of novel germplasm for *B. napus* breeding is therefore a key part of the work presented in this volume, and in the additional publications listed in chapter 10. In cooperation with scientists and breeders in Germany, Australia and China, a broad range of novel *B. napus* germplasm from interspecific hybrids was generated. These materials contain agronomically interesting variation for numerous resistance and quality traits, including nematode resistance (Snowdon et al. 1998b, Voss et al. 1999, 2000), blackleg resistance (Snowdon et al. 2000a), fatty acid composition (Wang et al. 2004, 2006), pod shattering and drought tolerance (Schelfhout et al. 2004, 2006), Verticillium resistance (Rygulla et al. 2007a, 2007b), seed glucosinolate content (Hasan et al. 2008), seed colour, seed fibre and seed oil content (Li et al. 2009). In numerous cases this material is now being used in commercial breeding programs to broaden the genetic basis of breeding lines for important resistance and seed quality traits.

Molecular cytogenetic techniques are not only useful for chromosomal analysis of interspecific hybrids or alien introgression lines, but also provide a means for unambiguous assignment and ordering of genomic sequences to specific chromosomes. The techniques we developed for chromosome identification and fluorescence *in situ* hybridisation in the *Brassica* A and C genomes (Snowdon et al. 1997a, 1997b, 1998, 2000a) today represent an important platform within the multinational *B. rapa* sequencing project. Before progressing with sequencing, all BAC clones from the sequencing contigs allocated to each each chromosome are hybridised in pairs to *B. rapa* pachytene chromosomes to confirm their chromosomal positions and orders. False assignment or ordering of markers are key problems in genetic mapping in the complex polyploid genome of *B. napus*, therefore our FISH methods are making an important contribution to the efficiency of the *B. rapa* sequencing project.

Particularly when useful traits must be introgressed from less adapted materials into high-yielding winter oilseed rape varieties, the availability of diagnostic molecular genetic selection markers is a key to rapid and effective transfer of responsible genes without excessive linkage drag for non-desired traits. Through generation of genetic linkage maps and tightly-linked molecular markers, we have provided novel selection tools to complement phenotypic screening for seed colour and fibre content (Badani et al. 2006a, 2006b), seed glucosinolate content (Basunanda et al. 2007, Hasan et al. 2008), Verticillium resistance (Rygulla et al. 2008) and turnip yellows mosaic virus resistance (Jürgens et al. 2009). In our current work we are using these genetic maps as a basis for map-based cloning of genes of interest. In particular, we have identified a 250 kb BAC sequence contig covering a major QTL contributing to seed colour and fibre content (unpublished results). Interesting seed-expressed candidate genes that are potentially involved in seed coat cell wall biosynthesis are currently being investigated by reverse genetics in an attempt to identify and clone the responsible gene or genes underlying this QTL.

Besides the direct application of germplasm and markers for practical breeding, genetic mapping and genome analysis also provide a basis for more fundamental research into complex traits. A good example is heterosis, a complex phenomenon whose control is genetically only poorly understood but which has important implications for improvement of seed yield. Today, F_1 hybrid varieties with improved yield stability and adaptability have the largest market share in most major oilseed rape and canola growing areas worldwide. Despite this market dominance, the yield gain of *B. napus* hybrids compared to homozygous open-pollinated varieties is relatively low compared to classical hybrid crops like sunflower or maize. In official plot trials in Germany, winter oilseed rape hybrids currently achieve a mean yield advantage of only around 6% compared to open-pollinated varieties (cumulative data from state variety trials), while in practical winter oilseed rape production the 10-year mean yield advantage of hybrids is currently only around 11% (data from Kleffmann Group/Norddeutsche Pflanzenzucht H.G. Lembke KG). Further improvement of yield performance through more efficient exploitation of heterosis is therefore an important breeding aim. The genetic basis of heterosis in *B. napus* is not well understood, however, and few studies have systematically compared the expression and control of heterosis in different genetic backgrounds. Our work on comparative mapping of heterosis-related QTL for seedling and yield traits (Basunanda et al. 2009) was the first study to investigate genetic relationships between the expression of heterosis in different developmental stages in *B. napus*. Interestingly, we revealed a number of QTL clusters involved in different traits, and at different developmental stages, in two genetically divergent test-hybrid mapping populations. A co-localisation of *per se* QTL with heterosis-related QTL indicated the presence of regulatory loci that may also contribute to fixed heterosis in the highly duplicated *B. napus* genome. Given the key role of epistatic interactions in the expression of heterosis in oilseed rape, these QTL hotspots might harbour genes involved in regulation of heterosis (including fixed heterosis) for different traits throughout the plant lifecycle, including a significant overall influence on heterosis for seed yield. Based on these observations, the discovery of *trans*-acting regulatory loci corresponding to these hotspots by expression QTL (eQTL) analysis or gene network co-expression studies seems a promising approach towards identification of potential candidate genes for global regulation of heterosis and yield. Since we could demonstrate that heterosis is manifested at a very early stage in *B. napus* seedlings, and that in some cases the same regulatory loci may be involved in expression of heterosis for seedling and yield traits, this opens the possibility to use seedlings as a promising controlled system for eQTL analysis of yield-relevant gene expression in winter oilseed rape. To further target such genes we have established an international systems genetics consortium to identify global regulatory genes in *B. napus*, collaborating with other relevant research groups from Germany, England and Canada, within the framework of the ERANET Plant Genomics program (see http://www.erapg.org/everyone/16790/18613/19533/19534).

We also published the first genetic and QTL map for the alternative spring oilseed crop *Camelina sativa* (Gehringer et al. 2006). This camelina mapping population and genetic map data are meanwhile being used in different international projects aimed at establishing camelina as a new oilseed crop for sustainable biofuel production in low-input cropping systems. Compared to other oilseed plants, camelina is particularly competitive in semi-arid regions and in low-fertility or saline soils. Furthermore, it exhibits a high adaptability to adverse environmental conditions and possesses a high level of resistance against insect pests and plant pathogens. This means that it can be produced in low-input cropping systems with reduced N-fertilization and minimal fungicide or pesticide applications. Moreover, the comparatively short vegetation period of only around 120 days makes it particularly suitable as an alternative annual crop for renewable resource production within tight crop rotations. Among the international research activities that have been initiated, using our *C. sativa* genetic map and mapping population as a basis, is a planned collaborative project between US and Canadian scientists aimed at developing high-throughput genomic tools for camelina breeding.

One area where international *B. napus* genome research has been somewhat slow in comparison to other major crop species, for example rice, barley, or maize, is in the availability of public marker and EST collections and the integration of genetic and physical maps via standardised nomenclature. To combat this deficit, the so-called “Multinational *Brassica* Genome Project” (MBGP) was established by international *Brassica* researchers to coordinate genomics activities and pool resources to achieve common goals in molecular breeding of *Brassica* crops. The primary aim of this initiative is the provision of freely available genetic resources for *Brassica* genome analysis, including mapping populations, integrated genetic maps, DNA marker sequences, genomic libraries, genomic sequences and gene expression data. Some of the genetic mapping data and one of the mapping populations described in this volume have already been deposited as public resources with the MBGP, via the internet portal http://www.brassica.info, and recently we have begun two new research projects that aim to contribute significant quantities of new genomic and transcriptomic data along with new genotyping tools and germplasm to this platform. Included in the resources we are developing are a large set of new SNP markers developed using next-generation sequencing techniques, along with a new high-density SNP map that will be sequence-annotated to the *B. rapa* genome sequence and to an ultradeep-expression map of *B. napus* seed development. Furthermore we are generating a substantial collection of over 450 genetically diverse *B. napus* inbred lines for association genetic studies. Included in this genotype diversity set are gene-bank core collections of winter and spring oilseed rape, fodder rape and swede varieties, a substantial set of genetically diverse, modern 00-quality winter oilseed rape varieties, and a collection of exotic lines containing rare alleles from Eastern European and Asian oilseed and vegetable accessions. The diversity set will be genotyped by colleagues at the Max Planck Institute for Breeding Research in Cologne with a large set of genome-wide microsatellite markers, and with a new, public, high-density SNP chip that is being developed by partners at the NRC-Plant Biotechnology Institute in Saskatoon, Canada. Together, this material and data will represent the most comprehensive public resource for *B. napus* association genetics that has been developed to date. As phenotypic, metabolomic and transcriptomic data for the diversity population is accumulated over the coming years and decades, it is anticipated that this association genetics platform will play a key role in the elucidation of important complex traits in *B. napus*, for example oil content, seed yield and quality traits along with resistance to biotic and abiotic stress factors.

In another approach, we are developing new techniques for ultradeep transcriptome analysis during *B. napus* seedling and seed development, based on multiplexed next-generation sequencing of short EST tag sequences. The publication of Obermeier et al. (2009) was the first to describe serial analysis of gene expression (SAGE) for global transcriptome analysis in *B. napus*. In that work, however, we used conventional Sanger technology to sequence concatenated EST-tags, using a standard LongSAGE procedure. Due to the restrictive cost of Sanger sequencing, this strongly limited the depth of transcriptome coverage we were able to achieve. By implementing next-generation sequencing platforms that are predestined for sequencing of short DNA fragments, like the

EST sequence tags generated via SAGE, sequencing costs are no longer a limiting factor. For around the same amount of money with which we obtained a total of only around 100,000 17bp LongSAGE tags in the previous study, we can today sequence more than 500 times as many SuperSAGE tags (24bp in length) using Illumina/Solexa sequencing. In contrast to conventional microarray analyses, this offers the opportunity to accurately quantify all expressed genes in a given transcriptome, including very low-copy transcripts and even unknown genes. We are currently using this method for ultradeep expression studies of seed development and for eQTL analysis during *B. napus* seedling development. The global transcriptome data we are generating will also be made publicly available via the MBGP in the form of a relational database incorporating SNP-maps, QTL and eQTL data, SNP marker sequences annotated to the *B. rapa* and Arabidopsis genome sequences, and the aforementioned ultradeep seed expression map.

With the breathtaking current developments in high-throughput genomics and sequencing technologies, it can be expected that the coming decade will see enormous advances in genome analysis and molecular breeding in oilseed rape. The public consortium that is currently competing the genome sequencing of *B. rapa* plans to achieve their target before 2011, however a private initiative led by the Beijing Genomics Institute expects to publish more or less complete sequences of both *B. rapa* and *B. oleracea* before the end of 2009, using data from high-depth Illumina/Solexa sequencing to fill existing gaps in the public sequence. The completion of the *Brassica* A and C genome sequences will immediately facilitate re-sequencing in *B. napus*, and it is not inconceivable that third-generation "single-molecule" sequencing technologies will in just a few years make "genotyping by sequencing" a reality. Even before this vision is achieved, however, the near future will without doubt see the development and implementation of ultra-high throughout assays for whole-genome and whole-transcriptome polymorphism analysis in major crop species. This is likely to change the focus of genetic research from discovery and analysis of individual genes underlying traits of interest, to a broader investigation of complete expression networks underlying complex traits and metabolic expression patterns. The increasing availability and accessibility of *Brassica* genomics tools and data, along with the close relationship between *B. napus* and the model plant Arabidopsis, mean that oilseed rape breeding stands to benefit enormously from the newest technological platforms for DNA sequencing and functional genomics.

8 Summary

The articles presented in this compilation, and the additional works listed in Chapter 10, describe numerous examples for the development and application of genome analysis tools for molecular breeding and trait dissection in *Brassica* oilseed crops. The main focus is on molecular breeding of oilseed rape (*Brassica napus*). This crop, the source of a highly nutritional vegetable oil and at the same time a valuable renewable resource for the biofuel industry, is currently the most important oilseed crop in Europe and Asia, and the second most important oilseed worldwide. Patterns of genetic and phenotypic diversity were studied in winter-type and spring-type *B. napus* gene bank collections, including oilseed, fodder and vegetable rape varieties of diverse geographical origin. Based on these data a new research initiative has begun to establish, genotype and phenotype a large set of genetically diverse *B. napus* germplasm with over 450 accessions for association mapping studies.

The narrow gene pool available for oilseed rape breeding can be broadened by the generation of interspecies hybrids via tissue culture techniques. Such strategies are vital for introduction and maintenance of effective resistance against important pathogens, since a sustainable production of major crops like oilseed rape is only possible when a broad range of resistant cultivars are available. We used interspecific hybridisation to successfully transfer new resistances against *Verticillium longisporum* and *Leptosphaeria maculans*, two of the major diseases of oilseed rape in Europe, into *B. napus* breeding lines. Information obtained by molecular genetic and cytogenetic characterisation of the introgression lines and their progeny are now contributing to the utilisation of these new resistance sources in commercial breeding programs. Similar strategies were used to develop and characterise interspecific hybrids for transfer of new seed quality alleles into oilseed rape from related crucifer species.

Genetic mapping and quantitative trait analysis are important tools for marker development and potential gene discovery in crop plants. The publications included here include the first published QTL analyses for *V. longisporum* resistance, seed colour and seed fibre traits in oilseed rape, along with comparative QTL analysis for heterosis in seedling and seed yield traits. We also generated the first genetic map for the alternative oilseed crop *Camelina sativa*, and used this map to analyse QTL for important agronomic and seed quality traits. These genetic maps and QTL data represent the first step towards map-based identification and cloning of major genes underlying important quantitative traits. Work is underway to develop a high-density SNP map and link important QTL regions from this map via the *B. rapa* genome sequence to a global gene expression map for seed development. This approach will incorporate next-generation sequencing technologies for gene expression tag profiling, using the methods and database tools we developed to annotate genes with the help of *B. napus* ESTs and *B. rapa* genomic sequence data.

Molecular cytogenetic techniques are not only useful for chromosomal analysis of interspecific hybrids or alien introgression lines, but also provide a means for unambiguous assignment and ordering of genomic sequences to specific chromosomes. The techniques we developed for chromosome identification and fluorescence *in situ* hybridisation (FISH) in the *Brassica* A and C genomes are making an important contribution to the efficiency of the multinational *B. rapa* sequencing project. Before progressing with sequencing, all BAC clones from the sequencing contigs allocated to each chromosome are hybridised in pairs to *B. rapa* pachytene chromosomes by FISH to confirm their chromosomal positions and orders.

As a member of the crucifer family, oilseed rape is closely related to the model plant *Arabidopsis thaliana*. Genome analysis in the complex, polyploid *B. napus* genome can benefit greatly from this close relationship, since there is a high sequence similarity of around 80% in gene-coding regions between the two species and a certain degree of conserved synteny exists between the genomes. Based on this relationship, an innovative approach was developed for marker identification, using *Brassica*-Arabidopsis sequence synteny to find molecular markers with close physical linkage to putative seed glucosinolate biosynthesis genes. Use of these markers in population structure-based allele-trait association analyses confirmed the association of the candidate genes with the trait of interest, providing a basis for sequence-based selection techniques to identify and use new low-glucosinolate alleles from gene bank resources.

With the astounding current developments in high-throughput genomics and sequencing technologies, it can be expected that the coming decade will see enormous advances in genome analysis in all major crop plants. In our own work we are shifting increasingly to the use of next-generation sequencing for high-throughput genome and transcriptome analyses. These technologies will have even greater implications for *B. napus* genome analysis and breeding when the *Brassica* A and C genome sequences are completed in the next few years. Re-sequencing of *B. napus* genotypes will without doubt play a key role in polymorphism discovery, whole-genome selection and molecular breeding of oilseed rape in the foreseeable future.

9 Zusammenfassung

Die hier zusammengestellten Veröffentlichungen sowie die in Kapitel 10 aufgeführten Artikel beschreiben zahlreiche Beispiele für die Genomanalyse und molekulare Züchtung von *Brassica*-Ölsaaten. Schwerpunkt dabei stellt die molekulare Züchtung von Ölraps (*Brassica napus*) dar. Der Raps liefert ein ernährungsphysiologisch höchst wertvolles Pflanzenöl, das aber auch gleichzeitig in Form von Biodiesel eine wichtige Ressource für die Treibstoffindustrie ist. Aufgrund dieser vielseitigen Verwendung ist Raps derzeit die wichtigste Ölsaatpflanze in Europa und Asien, weltweit die Zweitwichtigste nach Soja. In den vorliegenden Arbeiten wurden u.a. genetische und phänotypische Variationsmuster in Winter- und Sommerrapsgenotypen diverser geographischer Herkunft aus Genbanksammlungen ermittelt, darunter Ölraps-, Futterraps- und Gemüserapstypen. Auf diesen Daten basierend wurde eine neue Forschungsinitiative gegründet, bei der ein großes *B. napus*-Diversitätsset mit über 450 Akzessionen für assoziationsgenetische Studien zur Verfügung gestellt und umfangreich charakterisiert werden soll.

Der enge Genpool für die Ölrapszüchtung wurde desweiteren durch die Erstellung neuer interspezifischer Hybriden mittels Gewebekurturtechniken erweitert. Solche Strategien sind zur Einführung und Erhaltung effektiver Resistenzen gegenüber wichtigen Pathogenen unabdingbar. Auf Dauer kann die immer intensiver werdende Rapsproduktion nur durch die Verfügbarkeit einer breiten Auswahl resistenter Sorten gesichert werden. In den vorliegenden Arbeiten wurde die interspezifische Hybridisierung erfolgreich genutzt, um z.B. neue Resistenzen gegen *Verticillium longisporum* und *Leptosphaeria maculans*, zwei der wichtigsten Krankheiten des Rapses in Europa, in *B. napus*-Zuchtlinien zu übertragen. Informationen aus den hier dargestellten Arbeiten zur genetischen und zytogenetischen Charakterisierung von Introgressionslinien und deren Nachkommenschaften werden wesentlich zur effektiven Nutzung dieser neuen Resistenzquellen in der praktischen Rapszüchtung führen. Ähnliche Strategien wurden verwendet, um neue Allele für bestimmte Saatgutqualitätseigenschaften in den Raps zu übertragen.

Die genetische Kartierung und Analyse von *quantitative trait loci* (QTL) sind wichtige Werkzeuge zur Markerentwicklung für agronomisch wichtige Merkmale sowie für die Aufdeckung der beteiligten Gene. Zu den hier aufgeführten Publikationen gehören u.a. die ersten veröffentlichten QTL-Analysen für Verticillium-Resistenz, Samenfarbe sowie Samenfasergehalt in *B. napus* sowie vergleichende QTL-Analysen von Heterosis für Keimlingsentwicklung sowie Kornertrag und verwandte Merkmale. Darüber hinaus wurde die erste genetische Karte der alternativen Ölsaatpflanze *Camelina sativa* veröffentlicht und für die QTL-Analyse wichtiger agronomischer Eigenschaften und Saatgutqualitätsmerkmale verwendet. Diese genetischen Karten und QTL-Studien stellen einen ersten Schritt zur Identifizierung und Klonierung von zentralen Genen für die Ausprägung wichtiger, quantitativer Merkmale dar. In weiterführenden Arbeiten wird derzeit eine hochdichte Karte des Rapsgenoms auf Basis von *Single-Nucleotide Polymorphisms* (SNPs) generiert, anhand derer bedeutende QTL-Regionen mit der *B. rapa*-Genomsequenz sowie mit einer neuen, globalen Genexpressionskarte für die Samenentwicklung geknüpft werden können. Bei diesen Arbeiten werden die neuesten *Next-Generation* Sequenzierungstechnologien für Genexpression-*Tag-Profiling* unter Einsatz unserer bereits entwickelten Gen-Annotationsmethoden verwendet.

Molekularzytogenetische Methoden sind nicht nur zur Analyse der Chromosomen von interspezifischen Hybriden oder Introgressionslinien nützlich, sondern auch zur Bestimmung der chromosomalen Position und Anordnung von genomischen Sequenzen. Die von uns beschriebenen Methoden zur Fluoreszenz-*in situ*-Hybridisierung (FISH) in den A- und C-Genomen von *B. rapa*, *B. oleracea* bzw. *B. napus* spielen heute eine entscheidende Rolle beim multinationalen *B. rapa*-Sequenzierungsprojekt. Vor der Sequenzierung werden alle BAC-Klone aus den BAC-Contigs der einzelnen Chromosomen paarweise mittles FISH auf *B. rapa*-Pachytenchromosomen hybridisiert. Somit können ihre chromosomale Zuordnung sowie ihre Reihenfolge in Bezug zur Sequenzassemblierung bestätigt werden.

Als Mitglied der Familie Brassicaceae ist *B. napus* sehr eng mit dem Modellkreuzblütler *Arabidopsis thaliana* verwandt. Genomanalyse im komplexen, polyploiden Raps-Genom kann von dieser Beziehung stark profitieren, da es eine hohe Sequenzähnlichkeit von ca. 80% zwischen den beiden Arten gibt. Basierend auf dieser Verwandtschaft haben wir einen innovativen Ansatz zur Markeridentifizierung entwickelt, bei der die *Brassica*-Arabidopsis Sequenzidentität genutzt wird, um PCR-Marker mit enger Kopplung zu Kandidatengenen für Samenqualitätsmerkmale zu identifizieren. Durch den Einsatz dieser Marker für Assoziationsanalysen konnten wir die Assoziation dieser Kandidatengene mit dem relevanten Merkmal bestätigen und somit nützliche Marker für die markergestützte Selektion zur Verfügung stellen. Mit der ständig steigenden Menge und Verfügbarkeit von Genomsequenzen wird die Rolle solcher Techniken in der Genomanalyse und molekularen Züchtung von wichtigen Kulturpflanzen eine immer bedeutendere Rolle spielen.

Die derzeitigen Entwicklungen im Bereich der Hochdurchsatzgenomik und der DNA-Sequenzierung werden im kommenden Jahrzehnt zu enormen Sprüngen bei der Genomanalyse aller bedeutenden Kulturpflanzen führen. In unseren Arbeiten setzen wir vermehrt die *Next-Generation*-Sequenzierung für die Markerentwicklung und die Transkriptomanalyse ein. Diese Technologien werden eine sogar noch wichtigere Rolle spielen, wenn die *Brassica* A- und C-Genome in den nächsten Jahren vollständig sequenziert sind. In absehbarer Zukunft wird die Resequenzierung von *B. napus*-Genotypen ohne Zweifel eine Schlüsselrolle bei der Aufdeckung von Polymorphismen, der genomweiten Selektion sowie der molekularen Züchtung von Ölraps spielen.

10 Further publications

The following list of additional publications include nine monographs or book chapters covering molecular and cytogenetic analyses of crop plants, with special emphasis on breeding, genetic mapping and genome analysis in oilseed *Brassica* species. A further 29 original articles cover various aspects of molecular breeding in oilseed rape and other crops, including applied examples from Indian mustard (*Brassica juncea*), camelina (*Camelina sativa*), white mustard (*Sinapis alba*), faba bean (*Vicia faba*) and asparagus (*Asparagus officinalis*). Two publications describe molecular genetic variation in a pathogenic fungus (*Alternaria solani*) and a vertebrate species (rainbow trout, *Onchoryhnchus mykiss*), respectively. These "exotic" examples underline the global applicability of molecular genetic analysis techniques to organisms from all eukaryotic kingdoms.

8.1 Monographs and book chapters (chronological)

Friedt W. & R. Snowdon, 2009: Oilseed rape. In: Vollmann J. & R. Istvan (eds) Handbook of Plant Breeding, Vol. 4: Oil crops breeding. Springer Verlag, in press

Ahlemeyer J., R.J. Snowdon, F. Ordon & W. Friedt, 2006: Agro-diversity - Genetic diversity in crops and cropping systems. In: Benckiser G. & S. Schnell (eds) Biodiversity in agricultural production systems. CRC Press, Boca Raton, Fl, USA, pp. 21-40

Snowdon R.J., W. Lühs & W. Friedt, 2006a: *Brassica* oilseeds. In: Singh R. (ed) Genetic Resources, Chromosome Engineering, and Crop Improvement. Volume 3: Oilseed Crops. CRC Press, Boca Raton, Fl., USA, pp. 195-230

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

Friedt W., W. Lühs & R. Snowdon, 2003: Raps – eine moderne Kulturpflanze. In: Justus Liebig (1803-1873): Seine Zeit und unsere Zeit. Chemie – Landwirtschaft – Ernährung. Ausstellungskatalog, Universität Giessen, pp. 244-249

Voss A., W. Lühs, R.J. Snowdon & W. Friedt, 1999: Development and molecular characterization of nematode-resistant rapeseed (*Brassica napus* L.). In: Scarascia-Mugnozza G.T., E. Porceddu & M.A. Pagnotta (eds) Genetics and Breeding for Crop Quality and Resistance. Kluwer Academic Publishers, Netherlands, pp. 195-202

Snowdon R.J. & A. Langsdorf, 1998a: An introduction to DNA fingerprinting using RFLP and RAPD techniques. In: Tietz D. (ed) Nucleic Acid Electrophoresis. Springer, Berlin/ Heidelberg, pp. 99-128

Snowdon R.J., W. Köhler, W. Friedt, W. Lühs & A. Köhler, 1998b: Detecting R-genome chromatin in nematode-resistant *Brassica napus* x *Raphanus sativus* hybrids. In: Lelley, T (ed): Current topics in plant cytogenetics related to plant improvement. WUV-Universitätsverlag, Wien, pp 220-226

Snowdon, R.J., 1997: Fluorescence *in situ* hybridisation techniques for *Brassica* species: Methodological development and practical applications. Dissertation, Justus-Liebig-Universität Gießen, 1997

8.2 Original scientific publications in peer-reviewed journals (chronological)

Basunanda P., M. Radoev, W. Ecke, W. Friedt, H. Becker & R.J. Snowdon, 2009: Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). Theor. Appl. Genet., in press

Juergens M., I. Krämer, M. Zahn, C. Paetsch, R.J. Snowdon, W. Friedt, F. Rabenstein, J. Schondelmaier, & F. Ordon, 2009: Genetic analyses and development of molecular markers for resistance to Turnip Yellows Virus (TuYV) in rapeseed (*Brassica napus* L.). Theor. Appl. Genet., in press

Li A.M., C.X. Wei, J.J. Jiang, Y.T. Zhang, R.J. Snowdon & Y. P. Wang, 2009: Phenotypic variation in progenies from somatic hybrids between *Brassica napus* and *Sinapis alba*. Euphytica, in press

Wittkop B., R.J. Snowdon & W. Friedt, 2009: Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. Euphytica, in press

Basunanda P., T.H. Spiller, M. Hasan, A. Gehringer, J. Schondelmaier, W. Lühs, W. Friedt & R.J. Snowdon, 2007: Marker-assisted increase of genetic diversity in a double-low seed quality winter oilseed rape genetic background. Plant Breeding 126: 581-587

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

Badani A.G., B. Wittkop, W. Lühs, R. Baetzel, R. Horn, A. De Haro, R. Font, W. Friedt & R.J. Snowdon, 2006b: Seed colour in *Brassica napus*: QTL mapping, candidate genes and associations with quality traits. Acta Hort. 706: 203-209

Gehringer A., W. Friedt, W. Lühs & R.J. Snowdon, 2006: Genetic mapping of agronomic traits in false flax (*Camelina sativa* subsp. *sativa*). Genome 49: 1555-1563

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

Schelfhout C.J, R.J. Snowdon, W.A. Cowling & J.M. Wroth, 2006: Tracing B-genome introgressions in *Brassica napus* x *B. juncea* interspecific hybrids. Genome 49: 1490-1497

Wang Y.P., K. Sonntag, E. Rudloff, P. Wehling & R.J. Snowdon, 2006: GISH analysis of disomic *Brassica napus-Crambe abyssinica* chromosome addition lines produced by microspore culture from monosomic addition lines. Plant Cell Rep. 25: 35-40

Schelfhout C.J., R.J. Snowdon, W.A. Cowling & J.M. Wroth, 2004: A PCR based B-genome specific marker in *Brassica* species. Theor. Appl. Genet. 109: 917-921

Wang Y.P., R.J. Snowdon, E. Rudloff, P. Wehling, W. Friedt & K. Sonntag, 2004: Cytogenetic characterization and *fae1* gene variation in progenies from asymmetric somatic hybrids between *Brassica napus* and *Crambe abyssinica*. Genome 47: 724-731

Snowdon R.J. & W. Friedt, 2004: Molecular markers in *Brassica* oilseed breeding: Current status and future possibilities. Plant Breeding 123: 1-8

Peres S., R.J. Snowdon & J. Pons-Kühnemann, 2003: AFLP analysis of genetic variation in Cuban and international populations of *Alternaria solani*. Eur. J. Plant Path. 110: 399-409

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

Snowdon R.J., P. Böttinger, T. Pickardt, W. Köhler & W Friedt, 2001: Physical localisation of transgenes on *Vicia faba* chromosomes. Chromosome Res. 9: 607-610

Snowdon R.J., W. Friedt, A. Köhler & W. Köhler, 2000b: Molecular cytogenetic localisation and characterisation of 5S and 25S rDNA loci for chromosome identification in oilseed rape (*Brassica napus* L.). Ann. Bot. 86: 201-204

Voss A., R.J. Snowdon, W. Lühs & W. Friedt, 2000: Intergeneric transfer of nematode resistance from *Raphanus sativus* into the *Brassica napus* genome. Acta Hort. 539: 129-134

Snowdon R.J., W. Köhler, W. Friedt & A. Köhler, 1998: Fluorescence *in situ* hybridization techniques for *Brassica*: Methodological development and practical applications. Acta Hort. 459: 127-134

Jaag T., A. Langsdorf, R.J. Snowdon, W. Köhler & H.D. Hartmann, 1998: Random amplified polymorphic DNA variation among German and Dutch asparagus cultivars. Plant Breeding 117: 290-292

Snowdon R.J., W. Köhler & A. Köhler, 1997a: Chromosomal localization and characterization of rDNA loci in the *Brassica* A and C genomes. Genome 40: 582-587

Snowdon R.J., W. Köhler, W. Friedt & A. Köhler, 1997b: Genomic *in situ* hybridization in *Brassica* amphidiploids and interspecific hybrids. Theor. Appl. Genet. 95: 1320-1324

Snowdon R.J. & K.D. Adam, 1992: Allozyme variation in rainbow trout from central North Island, New Zealand. Trans. Am. Fish. Soc. 121: 810-814

8.3 Review articles (chronological)

Snowdon R.J. & Friedt W., 2008: Die Erfolgsgeschichte Raps geht weiter. GenomExpress 2.08: 25-27

Friedt W., R. Snowdon, F. Ordon & J. Ahlemeyer, 2006: Assessment of genetic diversity in crop plants and its exploitation in breeding. Progress in Botany 69:151-178

Snowdon R.J., K. Link, A.G. Badani & W. Friedt, 2005: Recent advances in molecular breeding of oilseed rape (*Brassica napus* L.). Progress in Botany 66: 144-163

O'Neill R., R.J. Snowdon & W. Köhler, 2003: Population genetics aspects of biodiversity. Progress in Botany 64: 115-137

Snowdon R.J., B. Kusterer & R. Horn, 2002: Structural genome analysis using molecular cytogenetic techniques. Progress in Botany 63: 55-79

11 Acknowledgements

The work described here was carried out at Justus Liebig University, Giessen, in the Department of Biometry and Population Genetics (1998 until 2001) and the Department of Plant Breeding (2001 until 2009). I am indebted to all of my former and current colleagues, collaboration partners and co-authors without whom the publications included herein would not have eventuated. I particularly thank Prof. Dr. Wolfgang Köhler for his support at the beginning of my postdoc work, and Prof. Dr. Wolfgang Friedt for giving me the opportunity to lead an enthusiastic and successful research group and for providing the excellent infrastructure with which this work was possible. Very special thanks are due to the highly motivated postdocs, PhD students, MSc students and technical staff who have contributed to the success of this research and to a highly enjoyable work environment. In particular, the following colleagues and cooperation partners made significant contributions to the publications included or listed in this work: Ana Gloria Badani, Panjisakti Basunanda, Anke Gehringer, Maen Hasan, Christian Obermeier, Rob O'Neill, Simon Peres, Jörn Pons-Kühnemann, Wilfried Lühs, Mladen Radoev, Wiebke Rygulla, Chris Schelfhout, Fatih Seyis, Tobias Spiller, Youping Wang, Hendrik Winter and Benjamin Wittkop. A large number of others who contributed valuable technical assistance are acknowledged in the publications herein.

The work described in the aforementioned publications was made possible by generous funding from the German Research Foundation (DFG), the Society for the Promotion of Private German Plant Breeders (GFP e.V.), the GABI plant genomics network of the Federal Ministry of Education and Research (BMBF), the German Agency for Renewable Resources (FNR) and the German Academic Exchange Service (DAAD). Collaborating institutions and companies included the University of Western Australia, the Free University of Berlin, Georg August University, Göttingen, Yangzhou University, China, the Federal Centre for Breeding Research on Cultivated Plants, Groß Lüsewitz, Deutsche Saatveredelung AG, Lippstadt, KWS Saat AG, Einbeck, Norddeutsche Pflanzenzucht H.-G. Lembke KG, Hohenlieth, Saaten-Union Resistenzlabor GmbH, Leopoldshöhe, and SW Seed GmbH, Hadmersleben.